


Machine learning for metallurgy II. A neural-network potential for magnesiumMarkus Stricker ^{*}, Binglun Yin, Eleanor Mak , and W. A. Curtin*Laboratory for Multiscale Mechanics Modeling, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland* (Received 24 June 2020; revised 18 August 2020; accepted 9 September 2020; published 1 October 2020)

Interatomic potentials are essential for studying fundamental mechanisms of deformation and failure in metals and alloys because the relevant defects (dislocations, cracks, etc.) are far above the scales accessible to first-principles studies. Existing potentials for non-fcc metals and nearly all alloys are, however, not sufficiently quantitative for many crucial phenomena. Here machine learning in the Behler-Parrinello neural-network framework is used to create a broadly applicable potential for pure hcp magnesium (Mg). Lightweight Mg and its alloys are technologically important while presenting a diverse range of slip systems and crystal surfaces relevant to both plasticity and fracture that present a significant challenge for any potential. The machine learning potential is trained on first-principles density-functional theory (DFT) computable metallurgically relevant properties and is then shown to well predict metallurgically crucial dislocation and crack structures and competing phenomena. Extensive comparisons to an existing very good modified embedded atom method potential are made. These results demonstrate that a single machine learning potential can represent the wide scope of phenomena required for metallurgical studies. The DFT database is openly available for use in any other machine learning method. The method is naturally extendable to alloys, which are necessary for engineering applications but where ductility and fracture are controlled by complex atomic-scale mechanisms that are not well predicted by existing potentials.

DOI: [10.1103/PhysRevMaterials.4.103602](https://doi.org/10.1103/PhysRevMaterials.4.103602)**I. INTRODUCTION**

Metal alloys are widely useful for many structural applications due to the presence and behavior of the underlying fundamental defects in the crystalline lattice. That is, the atomistic structures of dislocations, interfaces, crack tips, grain boundaries, precipitates, and vacancies, and their evolution and interactions determine the plastic flow behavior, creep, fatigue, fracture toughness, radiation resistance, etc., that ultimately control the macroscopic material performance. Understanding those structures, and their dependence on chemistry, is crucial for optimizing the use of existing alloys and for designing new higher-performance alloys. This necessitates atomic-scale simulations at the scales of the defects, which are often far too large for the use of first-principles methods such as density-functional theory (DFT). The development of semiempirical interatomic potentials that accurately capture the structures, energies, and motion of the various defects is thus essential.

The development of interatomic potentials has a long history, with successes and limitations, that is briefly reviewed in a companion paper [1]. Existing potentials for metals become increasingly inaccurate with increasing complexity in both crystal structure and chemistry. Relevant to the present paper is the case of elemental hexagonally close packed (hcp) metals, where many potentials struggle to capture the full range of operative slip systems (basal, prismatic, and pyramidal I and II) and with limited quantitatively accurate

extensions to alloys [2–6]. The general issues with traditional potentials are leading to the application of machine learning methods to fit the potential energy surface (PES) of a metal without imposing a highly restricted functional form [7–12]. The construction of a machine learned potential consists of (i) choosing a suitable class of geometric representations (the descriptors) to describe local atomic environments [13–17], (ii) developing a database of energies and forces of atomic structures using first-principles methods (the training dataset), and (iii) applying a regression algorithm (e.g., neural network, kernel ridge regression) to optimize the parameters in the machine learning framework to best-match the training data. Since the number of descriptors and/or parameters is unlimited, the machine learning approaches provide a “parameter-rich” space that can capture the training dataset well. However, since it is pure regression, machine learning potentials are not suitable for extrapolation to structures that differ notably from those in the training dataset. It must be recognized, however, that traditional interatomic potentials having fixed functional forms with limited parameters (i) are also intrinsically limited in their ability to accurately fit many properties, (ii) depend on the target properties and regression algorithms, and (iii) involve user-imposed decisions regarding which properties are most important because all desired properties cannot be achieved with sufficient accuracy.

For machine learning, a careful balance must be struck in selecting an appropriate limited set of descriptors, a limited set of fitting parameters, and a sufficiently large and diverse training set, so as to achieve broad accurate performance without the overfitting that exacerbates extrapolation errors. To circumvent the tedious task of curating an

^{*}markus.stricker@epfl.ch

exhaustive training dataset while maintaining transferability, several semiautomated protocols have been developed to sample the phase space of a material. *Active learning* methods have used random perturbations of bulk crystalline structures to sample the phase space of Al, Mg, and an Al-Mg alloy space [18]. *Self-guided learning* has explored the phase space by using randomized unit cells paired with a selection of the most diverse structures and applied to C, Si, and Ti [19]. *On-the-fly learning* methods have combined DFT calculations with machine learning for the calculation of melting points for Al, Si, Ge, Sn, and MgO [20]. A new hybrid approach combines an analytical bond order potential (BOP) form with a neural network that adjusts the BOP parametrization depending on the specific environment [10]. This approach can improve extrapolations because the analytical form is smooth.

A key, and limiting, aspect of nearly all of the machine learning (ML) potentials generated to date is that the training data, and fitness of the potential, are mainly demonstrated on basic properties of the bulk crystalline material [21]. Very few, if any, defects are considered. Structures in the training set and the predicted properties are associated with the equilibrium geometry, elastic response of the bulk, vibrational properties, vacancies, surfaces, etc., and, in some cases, liquid-state information. These features are necessary but far from sufficient for performing metallurgically useful studies of the behavior of defects in metals. For instance, the stable stacking fault energy (SSFE) for both fcc Ni and Cu is very poorly predicted by all the machine learning methods analyzed in Ref. [21] because the nature of the training set did not include configurations near this structure. The aforementioned self-guided semiautomatic approaches using random perturbations are useful to sample very-near bulk configurations but will not produce, e.g., a stacking fault or a vacancy. Yet, the SSFE is essential for modeling dislocations, and so none of the machine learning potentials in Ref. [21] would be suitable for quantitative plasticity modeling, although that was not the purpose of Ref. [21]. There are a few efforts extending beyond basic properties. Kobayashi *et al.* [9] developed a neural-network potential (NNP) for the Al-Mg-Si system including many intermetallics, solute-solute interactions, and interfaces and showed good predictions for edge and screw dislocation structures, solute or dislocation interactions, and *in situ* precipitates; further development was shown by Imbalzano *et al.* [22]. The Gaussian approximation potential (GAP) potentials for Fe [23] and W [24] 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 [25]. The hybrid physically informed neural network (PINN) approach [10] showed application to a face-centered-cubic (fcc) edge dislocation in Al. A NNP developed for Ti presented transformation pathways from body-centered-cubic (bcc) to fcc crystal structure [12]. These recent works highlight the promise of machine learning methods, but still remain fairly limited in metallurgical scope.

In this paper we study magnesium (Mg) as both an excellent complex testbed for machine learning and because Mg is a very desirable structural material [26–28] due to a combination of low density, high specific strength, and availability [29]. Complexity arises because the crystal structure is hcp so that the many different slip systems must be activated to

enable general plasticity according to the von Mises criterion [30], and various possible crack geometries may be susceptible to cleavage [31]. These issues are intimately tied to the low ductility and low toughness of pure Mg, which severely limits its practical application. Recent work shows that alloying Mg with dilute (<1%) additions of rare-earth elements (Y, Nd, Ce, Gd), Ca, and Mn improves ductility at room temperature substantially. This has been attributed to solute-accelerated cross-slip and double cross-slip of pyramidal ($\mathbf{c} + \mathbf{a}$) dislocations [32–34], demonstrating the intimate connection between detailed atomic-scale dislocation processes in the alloy and the macroscopic behavior. But direct simulations require scales far above those accessible by DFT. Simulation of this mechanism and other mechanisms associated with flow and fracture remain necessary but thus require potentials suitable for the simulations of hundreds of thousands of atoms. Furthermore, the potentials must have an accuracy sufficient to capture very subtle energy differences between different dislocations that ultimately govern the mechanical behavior of Mg. Early embedded atom method (EAM) potentials for Mg failed dramatically in their descriptions of the pyramidal dislocations [35] but modified embedded atom method (modified embedded atom method (MEAM)) potentials for pure Mg exist [4,5] and the most recent versions [5,6] are overall quite good. However, even the most recent MEAM for the important Mg-Y [36] is not sufficiently accurate for all necessary quantities, making it unsuitable for studying critical mechanisms in Mg alloys.

As the first essential step in overcoming the current limitations to atomistic studies of Mg and its alloys, here we present a family of Behler-Parrinello neural network (BPNN) [13] (or short NNP hereafter) machine learning potentials for pure Mg using the implementation of the library n2p2 [37]. The potentials are fitted with extensive first-principles DFT calculations of metallurgically relevant properties. However, we intentionally restrict the training database to (nearly) the same set previously used to fit MEAM potentials and then make side-by-side comparisons of MEAMs and NNP. We control every step of the fitting process with state-of-the-art methods from the choice of the representation (symmetry functions), the choice of the neural-network topology, and the curation of the training data. Results demonstrate that the family of NNP, and detailed analysis of one selected NNP denoted as NNP63, is broadly superior to the best MEAM potentials when evaluated across a wide range of metallurgical properties. The crucial pyramidal II ($\mathbf{c} + \mathbf{a}$) dislocation structure remains imperfect, and restrictions on the use of the current potential are discussed. However, the NNP in general can be continually improved with an expanded training database, augmented to avoid some of the extreme extrapolation problems, and, most importantly, easily extended to Mg alloys.

The remainder of this paper is organized as follows. In Sec. II the BPNN potential framework is summarized and the fitting procedure is described. Additional details about the DFT implementation used to build the training database, and the range of structures considered is presented. Section III compares the predictions of both NNP63 and MEAM potentials against the DFT for a range of properties directly derivable from DFT and relevant for mechanical properties. In more detail, Sec. III A presents the results from the fitting

the neural network; Sec. III B presents material properties of which most underlying structures were part of the training set; Secs. III C and III D each present results on training set data (stacking fault energy curves and decohesion, respectively) and the related applications to dislocation and fracture which demonstrate transferability of the potential. Predictions using NNP63 are then made for dislocation geometries, basal dislocation Peierls stress, and fracture (crack-tip phenomena) for a range of orientations, with comparisons to available DFT, experiments, and/or theoretical predictions. Section IV provides further discussion of our results and summarizes the work including future possibilities for Mg-alloy potentials.

II. METHODS

A. Neural-network architecture and implementation

In this section we briefly outline the structure of a NNP in the formulation of Behler-Parrinello [13]. All details and implementation of the neural-network method employed here are well presented in the recent literature and the reader is referred to, e.g., Refs. [13,37–40].

$$E_n = f_1^3 \left\{ b_\ell^3 + \sum_{\ell=1}^{M_{\text{layer},2}} w_{k\ell}^{23} f_k^2 \left[b_k^2 + \sum_{k=1}^{M_{\text{layer},1}} w_{jk}^{12} f_j^1 \left(b_j + \sum_{i=1}^{M_{\text{sym}}} w_{ij}^{01} G_i \right) \right] \right\}, \quad (2)$$

where $f(\cdot)$ are the so-called activation functions, $M_{\text{layer},i}$ denotes the number of nodes in the i th hidden layer, and M_{sym} is the number of symmetry functions used to represent local atomic environments. The quantities $\{w_{ni}\}$ and $\{b_i\}$ are the so-called weights and biases, which are determined by fitting the total energy modeled according to Eq. (1) to a training dataset of structures, total energies, and individual atomic force components. The i th component $F_{n,i}$ of the force on atom n is computed as $F_{n,i} = -\partial E_{\text{tot}}/\partial x_{n,i}$.

The structural representations of the atomic environment are dictated by the symmetry functions. The BPNN framework defines radial and angular symmetry functions of the forms

$$G_i^{\text{rad}} = \sum_{i \neq j} e^{-\eta(r_{ij}-r_s)} f_c(r_{ij}), \quad (3)$$

$$G_i^{\text{ang}} = 2^{1-\zeta} \sum_{j,k \neq i} (1 + \lambda \cos \theta_{ijk})^\zeta e^{-\eta(r_{ij}^2 + r_{ik}^2 + r_{jk}^2)} \times f_c(r_{ij}) f_c(r_{ik}) f_c(r_{jk}), \quad (4)$$

where $r_{ij} = |\mathbf{r}_j - \mathbf{r}_i|$ is the distance between two atoms i and j ; θ_{ijk} the angle between three atoms i, j, k ; f_c is a smooth cutoff function; and $\eta, r_s, \lambda,$ and ζ are predefined hyperparameters. Here we employ 27 radial and 5 angular symmetry functions; this is a heuristic design choice based on goals of avoiding overfitting and having acceptable computational costs (cf. Appendix A for the specific hyperparameters). The selection of the specific functions is made as follows. An initially very large number M of possible symmetry functions is considered. The training dataset of atomic structures is then specified. Each atom in each structure has a local atomic environment; across the entire training set there are

Machine learning potentials developed to date first assume that the total potential energy of a system of N atoms can be represented as a sum of the energies of each atom $n = 1, \dots, N$, with the energy of each atom depending on the local environment around the atom,

$$E_{\text{tot}} = \sum_{n=1}^N E_n. \quad (1)$$

This assumption is also underlying any other empirical potential and enables efficient application of the methods to large system sizes, as opposed to specific potentials for problems with a fixed size [41]. The neural-network potential formulation of Behler and Parrinello [13] consists of the choice of the atomistic structural representation called the symmetry functions, the number of hidden layers in the neural network, and the number of hidden nodes per layer in the network. For a dense two-layer neural network as used here with the symmetry functions denoted as G_i , the local atomic energy E_n is described as

a total of N atoms and thus N local environments (some of which may be identical). By evaluating the M candidate symmetry functions centered on each atom in the training dataset a so-called *feature matrix* is created that consists of the M symmetry functions (columns) by N atomic environments (rows). An unsupervised selection algorithm based on a CUR matrix decomposition [22] then determines the 32 most *valuable* symmetry functions ($\ll M$ columns) of the feature matrix, i.e., those with the highest information content with respect to all the N environments in the training structures. This selection constitutes the input layer of the NNP. The CUR method could also be used to provide an error measure for the selection, from which the number of symmetry functions would then be an outcome. Note that the selection of symmetry functions is intimately tied to the training dataset; if the training dataset is increased or filtered, then the optimal set of symmetry functions might be different. In practice, the use of a large number of bulklike atomic environments in metals leads to fairly small differences in the selected set of symmetry functions when the training dataset is extended by including additional structures. The final chosen hyperparameters for the 32 symmetry functions are presented in Appendix Tables III and IV.

The employed neural network consists of an input layer with 32 symmetry functions, two hidden layers with 20 nodes each, and the final layer with one neuron representing the energetic contribution an atomic environment [Eq. (2)]. The hyperbolic tangent is used as an activation function. This chosen topology results in 1101 fitting parameters, which is the combined number of weights and biases of the network.

After selecting the symmetry functions and fixing the topology of the neural network, the determination of the

weights $\{w_{ni}\}$ and biases $\{b_i\}$ is done with supervised learning to minimize an error function Γ equal to the sum of the squares of the differences between the NNP and DFT energies and forces. Specifically, for N_{struct} structures in the training set [38], the error measure is

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

where E and F represent the energy of, and forces on individual atoms in, a structure, respectively. β (in \AA^2) is a fixed parameter during training that allows for relative weighting of forces versus energies; here we use $\beta = 10 \text{\AA}^2$. The quality of the optimization is then measured by separately evaluating of the root-mean-square errors (RMSE) of the energy and forces,

$$\text{RMSE}(E) = \sqrt{\frac{1}{N_{\text{struct}}} \sum_{i=1}^{N_{\text{struct}}} (E_{\text{NNP}}^i - E_{\text{DFT}}^i)^2}, \quad (6)$$

$$\text{RMSE}(F) = \sqrt{\frac{1}{N_{\text{struct}}} \sum_{i=1}^{N_{\text{struct}}} \sum_{j=1}^{3N_{\text{atom}}^i} \frac{1}{3N_{\text{atom}}^i} (F_{j,\text{NNP}}^i - F_{j,\text{DFT}}^i)^2}. \quad (7)$$

The fitting (or *training*) of NNPs is done with a Kalman filter [42] as implemented in n2p2 [37] for 400 epochs (iterations). The number of epochs is based on initial tests. For the chosen symmetry functions, neural-network topology, and the dataset the gradient of the error function (Eq. (5)) with respect to epochs becomes sufficiently small at 400 epochs while overfitting is avoided. Changes in the initial values of $\{w_{ni}\}$ and $\{b_i\}$ and/or of the subset of structures used in the fitting lead to different final NNPs with different results for material properties. Comparisons among these NNPs enables assessment of the broader capability of the NNPs and for determining when the NNPs are being used in regions of inaccurate extrapolation. If the different NNPs deviate by some meV/atom on any given structure, then the differences are well below the typical errors between the NNPs and the DFT (a few meV/atom). The different NNPs are also used to compute material properties that involve energy differences between a defect or deformed and a reference state. If the errors in the absolute energies of both reference and defect states differ in sign, then the material property can deviate more than anticipated from the general RMSE of the NNP. Assessing these latter issues is essentially validation steps outside the generic machine learning formulation itself.

The choice of the training dataset with the function for optimization [Eq. (5)] defines the mathematical optimization problem. This might seem trivial but is important. Choosing many similar structures implicitly steers the optimization toward favoring a low error on those structures. Also, the loss function here [Eq. (5)] contains an implicit weighting of structures with high energies (typically structures with large numbers of atoms). Any step in the optimization that

reduces the error of a large structure also reduces the total error by a factor scaling with the number of atoms. Most of the energy of a solid-state structure is dominated by the cohesive energy [cf. Figs. 1(a) and 1(b)] and reducing that error for large structures reduces the overall root-mean-squared error (RMSE). However, defect formation energies are calculated as the difference between the defect structure and the reference structure. The defect structure may involve many (10s or 100s) atoms that are away from the defect and near the bulk reference configuration. This issue points to the value of adding atomic forces through nonequilibrium structures into the overall optimization as a means capturing local behavior. Many of these points cannot be addressed here and exist because the standard methods such as the BPNN as implemented in n2p2 [37] were mainly developed for applications to molecules rather than solids such as bulk metals. Contrary to a database of bulk metals, a database of molecules is often very homogeneous with respect to the number of atoms per structure. As the application of ML methods to solids expands, future optimization methods and codes may address these issues.

B. Training dataset and DFT details

A suitable training dataset should span the atomic environments encountered in a wide range of crystalline defects so that the potential is best suited to interpolate between those environments. Here we use a training dataset for the NNP that is nearly the same as used for prior MEAM-type potentials. Typical input data for fitting a MEAM potential comprises values for lattice constants, elastic constants, stacking faults, generalized stacking fault energy curves, decohesion curves, and surface energies for various crystallographic planes, often with a combination of first-principles calculations and experimental values [5]. Here, for the NNP, we use the same underlying data but fit only energies and forces, from which material properties are then derived. Thus, instead of using a *big data* approach, we incorporate metallurgical knowledge by choosing relevant structures for mechanical properties. A big data approach might include many types of random structures which would not be used in a traditional approach, e.g., liquid structures which are snapshots from a trajectory at a specific temperature. But, as noted earlier, the inclusion of large structures may drive the NNP toward capturing those structures, and those structures may not be sufficient for accurate representation of crystalline defects. The addition of further structures to the training dataset is always possible.

The training dataset developed here consists of atomic structures whose energies are used for calculating the energy-volume curve, elastic constants (C_{ij}), cohesive energy, generalized stacking fault energy (GSFE) curves (basal, pyramidal I, and pyramidal II), stable stacking fault energies (basal, pyramidal I, and pyramidal II) that can involve relaxations missing in the standard GSFE, decohesion curves (basal, pyramidal I, and II), and fully relaxed surface energies. In addition, beyond the above data typically used for the MEAM potential, we find it necessary to include rod and cuboidal structures containing corners and edges of several high symmetry planes to obtain physical behavior at atomically sharp crack tips. The complete

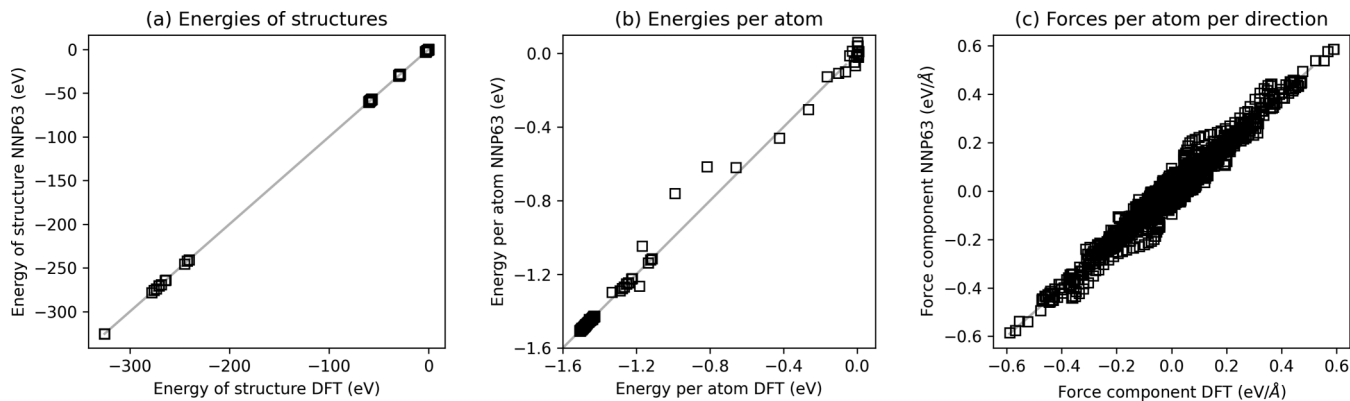


FIG. 1. NNP63 vs. DFT calculated energies of structures (a), energies per atom (b), and forces (c) after training.

DFT dataset consists of 443 structures with a total of $N = 12\,268$ atom environments. As discussed earlier, the number of atomic environments in the dataset defines the N rows of the feature matrix. Similar or identical atomic environments, e.g., from a bulk material, do not need a different set of optimal symmetry functions; they can be described with the same symmetry functions without loss of information. But given the current function [Eq. (5)], the number of individual atoms (and therefore the environments) does change the optimization by implicit weighting. The entire dataset is accessible on MaterialsCloud [43].

All first-principles calculations are performed using DFT as implemented in the Vienna Ab Initio Simulation Package (VASP) [44,45] within the generalized gradient approximation (GGA) and using the Perdew-Burke-Ernzerhof (PBE) exchange correlation functional [46]. Core electrons are replaced by the projector augmented wave (PAW) pseudopotentials [47] with Mg (3s) as the valence state. The valence-electron eigenstates are expanded using a spin-polarized plane-wave basis set with a cutoff energy 400 eV. In reciprocal space, a Γ -centered Monkhorst-Pack [48] k -mesh is used with line density consistent across all geometries. The interval between the neighboring k -points along each reciprocal lattice vector \mathbf{b}_j is $0.02\pi \text{ \AA}^{-1}$ (in VASP, $\mathbf{a}_i \cdot \mathbf{b}_j = \delta_{ij}$) and the k -mesh for the Mg 2-atoms primitive hcp cell is $36 \times 36 \times 19$. A second-order Methfessel-Paxton method [49] with a smearing parameter of 0.2 eV is used. The employed DFT parameters yield lattice constants ($a = 3.198 \text{ \AA}$, $c/a = 1.627$) and elastic constants (see Table I) in very good agreement with experiments [50].

It is important to use well-converged and consistent DFT calculations. Any computational inconsistencies across the DFT are translated directly into the resulting NNP, i.e., the machine learning potential is trying to learn inconsistencies yet treating them as real. As usual, approximations in the DFT carry over into the NNP; the NNP can only be as accurate as the DFT training set.

III. RESULTS

In the following we mainly compare one fitted NNP63 to the DFT dataset and to various material properties directly derivable from DFT energies. However, results for all 129 NNP are also shown in a number of cases. The sin-

gle choice of NNP63 for the majority of studies presented here is based on several assessments. First, the RMSEs of energies and forces [Eqs. (6) and (7)] are on the order of meV/atom and meV/Å, respectively. Second, the material properties (cf. Table I) match reasonably against DFT values. Additional tests, not shown here, with rod and cuboid shapes were also performed and no unexpected or anomalous behavior was found. In comparisons below, we start with basic material properties and then extend to properties outside of the training set. We then examine stacking fault energies and predicted dislocation structures, followed by decohesion energies and predicted crack-tip behavior relevant to fracture.

All the following simulations are performed using LAMMPS [51] with n2p2 [37], and Ovito [52] is used for the visualization of atomic structures.

A. NNP63 versus DFT training set

Figure 1 presents the energies, energies per atom, and forces for all of the structures in the DFT dataset as computed using our final single choice for the Mg NNP63 versus the DFT training and testing inputs. The fit is good, especially given the range of energies [Fig. 1(a)] for the different structures: RMSEs are 6.68 meV/atom and 12.8 meV/Å for energy and forces, respectively. Large deviations from the ideal energy in the center of Fig. 1(b) belong to structures with two atoms only and are very high in total energy, i.e., a large difference does not matter here, since the energies are on the order of 1 eV above bulk equilibrium (cf. discussion along Fig. 4). Similar results are obtained for all other NNP generated. These errors are in expected range with machine learned potentials for the prediction of atomic energies and forces for metals based on diverse training sets [9,23–25]. The total energies of each structure are shown, rather than energy per atom, because it is the total energies that enter the loss function [Eq. (5)]. Forces are presented as per atom per direction, again because that is how they enter the loss function [Eq. (5)]. As mentioned in Sec. II, structures with large energies bias the optimization toward minimizing the absolute errors of these structure, not the relative errors. Our training set includes structures containing from 2 to 216 atoms, with 2-atom structures related to bulk properties and the 216-atom structures related to structures with corners, i.e., rod shapes

TABLE I. Pure Mg material properties from experiments, *ab initio* calculations (evaluated on the same dataset used for training the neural-network potential), NNP63, the empirical MEAM potential (here denoted MEAM1 [53]), and another MEAM fit (here denoted as MEAM2 [6]). Stacking fault values are for relaxed ($\sigma_{3j} = 0$) as described in Ref. [54]. All values other than DFT have been calculated with the same protocol. The additional values in the lower part of the table are not part of the training set.

| Mg properties | Expt. | DFT | NNP63 | MEAM1 [5] | MEAM2 [6] |
|--|--------------------|------------|-------|-----------|-----------|
| a (Å) | 3.209 (300 K) [55] | 3.190 | 3.186 | 3.187 | 3.186 |
| c/a | 1.624 (300 K) [55] | 1.627 | 1.613 | 1.623 | 1.622 |
| E_c (eV/atom) | -1.51 [56] | -1.51 | -1.51 | -1.51 | -1.51 |
| C_{11} (GPa) | 63.5 [57] | 61 | 72 | 66 | 66 |
| C_{12} (GPa) | 25.9 [57] | 28 | 31 | 24 | 25 |
| C_{13} (GPa) | 21.7 [57] | 22 | 27 | 20 | 22 |
| C_{33} (GPa) | 66.5 [57] | 64 | 68 | 71 | 70 |
| C_{44} (GPa) | 18.4 [57] | 18 | 19 | 18 | 18 |
| Stacking fault energies (mJ/m ²) | | | | | |
| γ_{sf} Basal II | 33 [58] | 34 | 31 | 23 | 20 |
| γ_{sf} Pyr. I SF2 | — | 161 | 157 | 169 | 154 |
| γ_{sf} Pyr. II SF | — | 165 | 155 | 200 | 169 |
| Surface energy (mJ/m ²) | | | | | |
| Basal (0001) | — | 549 | 583 | 568 | 562 |
| Prism (10 $\bar{1}$ 0) | — | 624 | 644 | 583 | 569 |
| Pyramidal I (10 $\bar{1}$ 1) | — | 640 | 663 | 616 | 622 |
| Pyramidal II (11 $\bar{2}$ 1) | — | 732 | 746 | 652 | 639 |
| Additional values | | | | | |
| Tension twin | | | | | |
| Interface E (mJ/m ²) | — | 128.1 [59] | 118.8 | 148.0 | 132.3 |
| Vacancy energies (eV) | | | | | |
| Formation | 0.79 ± 0.03 [60] | 0.796 | 0.735 | 0.908 | 0.667 |
| Migration barrier basal | — | 0.397 | 0.436 | 0.616 | 0.601 |
| Migration barrier pyr | — | 0.416 | 0.470 | 0.641 | 0.616 |
| Melting point T_m (K) | 923 | — | ~900 | — | — |

(surfaces in two directions, periodic in one direction) and cuboids (surfaces in all three directions). The implicit weighting of structures with a higher number of atoms is, however, partially compensated by the larger number of bulk-related structures as compared to surfaces and corners in the present dataset.

B. Basic material properties

The accuracy of an interatomic potential is assessed mainly by its ability to reproduce material properties that are either known from experiments or are computed or derived from DFT energies. For machine learning potentials, only DFT is

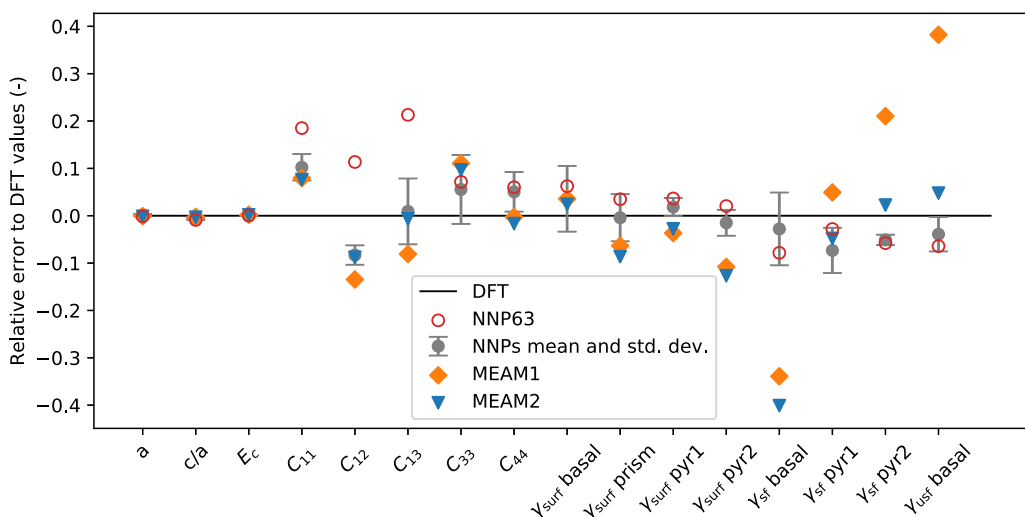


FIG. 2. Relative errors for various material properties with respect to DFT of NNP63 (red open circles), the mean and standard deviation of the entire family of NNPs (gray), and two MEAMs (orange diamond and blue triangle).

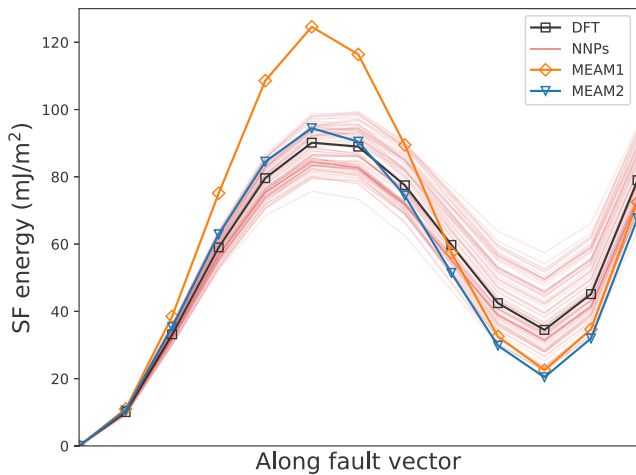


FIG. 3. Generalized stacking fault energy curves for the basal plane from DFT (black), multiple NNPs (red), and two MEAM potentials (orange and blue).

the reference; deviations of DFT-derived properties must be resolved independently of any DFT-derived machine learned potential. For elemental metals, the main material properties considered by most traditional potentials are the lattice constants, elastic constants, cohesive energy, vicinal surface energies, and vacancy formation energy. Occasionally, the stable and unstable stacking fault energies for important slip modes are also evaluated or fitted.

Furthermore, the standard error assessment of a NNP after training (e.g., Fig. 1) is not necessarily meaningful. Standard material properties are defined in terms of energy differences between states. So, errors in absolute energy having a different sign may lead to a poor property prediction. Or, when the errors in absolute energy are not small compared to the energy difference, a poor property prediction is obtained. The latter is especially relevant for elastic constants, which require energy differences between structures that differ only by very small

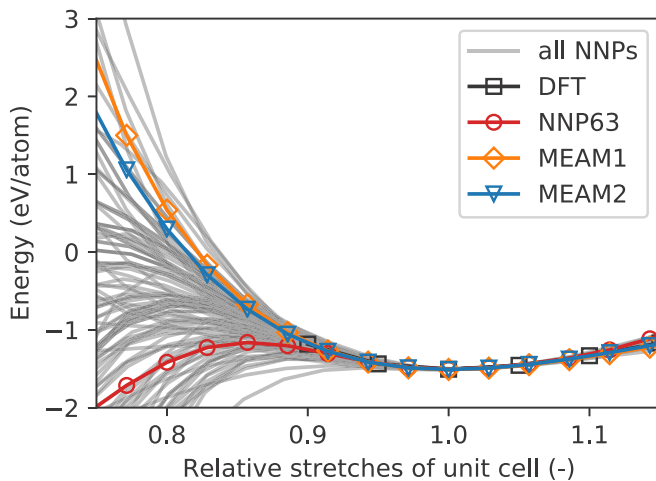


FIG. 4. Equation of state for all NNPs, the DFT reference (black squares), the selected NNP63 (red circles), and the two MEAMs (orange diamond and blue triangle).

strains—the use of larger strains introduces nonlinearities that violate linear elasticity. Hence, potentials must be carefully validated on various material properties. Standard potentials are often directly fit to the material properties rather than fitting energies from which properties are then derived, as done for the NNP.

Table I presents a range of basic material properties of Mg from experiments, as computed from DFT, our selected NNP63 potential, and two existing MEAM potentials developed previously by our group [5,6]. Overall, the properties of NNP63 agree well with the DFT reference values, which in turn agree well with experiments. One notable deviation is a slightly smaller c/a ratio; deviations in c/a from the ideal value determine the Burgers vector associated with twinning dislocations and so this deviation in c/a is not as negligible as it might otherwise appear. Another notable deviation is the larger value of C_{11} . Since elastic constants determine elastic energies, this deviation can affect dislocation structures and other defect properties. On the other hand, the stable stacking fault energies for the pyramidal I and pyramidal II planes are much closer to DFT than either of the MEAM potentials, and these quantities are crucial for establishing the relative stability of pyramidal I and II dislocations that are believed to hold the key to ductility in Mg and its alloys. The agreement for properties in the upper half of the table is expected because the associated structures are contained in the training dataset.

Table I also shows some additional properties that are not derived directly from structures and energies in the training set. Examination of such structures provides some indication of transferability of a potential to unknown atomic environments. Here the tension twin interface, which is a very low-energy planar structure, is well predicted by NNP63 while the MEAM potentials shows a larger error. Similarly, the vacancy formation and migration energies—involving coordinations that are not directly contained in the training set structures—are also well predicted by NNP63. In contrast, the MEAM values are much worse, with the migration barriers being very poor and not sufficient for any realistic studies of diffusion-related phenomena.

Finally, while our primary interest is in solid-state Mg and the crystalline defects that control mechanical properties, we have also examined the melting point of NNP63 to further demonstrate the scope of possible applications. We estimate the melting point by performing a periodic dual-phase crystal-liquid system with 8064 atoms and a nominally flat crystal-liquid interface with the normal coinciding with the c axis, i.e., $\mathbf{n} = [0001]$. A constant average interface position over time is achieved at an estimated melting point of $T_m \approx 900$ K. No liquid or random configurations of atoms were part of the training set and therefore these simulations additionally demonstrate transferability of the potential even beyond the solid state.

The above results are shown for the selected NNP63. Executing the optimization algorithm using different training sets (90% of the total set of structures and energies) leads to different final weights and biases and hence different final NNPs. Limited studies indicate that variations in the initial choices of the weights and biases, but using a fixed training set, result in small variations in the final RMSE. The variations among NNPs thus arise from the differences in the

choice of the 90% of structures used for training. Using our overall set of structures and energies, we have developed 129 different NNPs for Mg that should be nominally identical. Figure 2 shows the mean and standard deviation of a range of material properties, including those in Table I, computed from these 129 NNPs, along with the properties obtained from the MEAM potentials; here we show the relative deviation versus DFT-computed properties rather than the properties themselves. The average of NNPs predictions are, overall, in better agreement with DFT than the MEAMs across most material properties. The MEAMs does provide better properties in a few cases, generally those that were directly fit. The standard deviation in properties across all 129 NNPs is typically $\pm 10\%$ or less, and the NNPs deviate from DFT by no more than $\pm 12\%$ across all material properties. We note clearly that this standard deviation is not a true statistical measure—the different NNPs are all trained on very similar training sets and so results are not representative of statistically independent random variables or samplings. A more relevant standard deviation quantity can be estimated using the methodology in Ref. [61]. Here the standard deviation is merely used to indicate the range of property values starting from different specific training sets. Overall, most of the 129 NNPs perform better than the MEAMs, and the MEAMs can show substantial deviations from DFT, particularly for some surface and stacking fault energies.

The MEAM predictions for the stable stacking fault energy of the basal plane γ_{sf} show particularly large relative errors of $\approx 40\%$. Such an error has significant consequences for plasticity in Mg and its alloys because basal slip is the dominant slip system in Mg. The low MEAM stacking fault energy leads to a wider dissociation of the basal $\langle \mathbf{a} \rangle$ dislocation. This may affect the strengthening of basal slip on alloying, which is important for reducing the plastic anisotropy (ratio of pyramidal strength to basal strength) that can be important for ductility. In addition, prismatic $\langle \mathbf{a} \rangle$ slip, having the same Burgers vector as the basal $\langle \mathbf{a} \rangle$ slip, can be important in the plasticity of textured Mg sheet materials. But it is understood to occur only via cross-slip of screw dislocations from the basal plane to the prism plane. This cross-slip requires thermally activated recombination of the basal partial dislocations, and the energy barrier scales inversely with the stacking fault energy. We make these remarks to demonstrate that a deep understanding of metallurgical mechanisms is critical to the assessment of any interatomic potential and that a broadly applicable potential must be sufficiently accurate for many properties.

As another example showing the generally good agreement of the 129 NNPs versus reference training data is presented in Fig. 3. This figure shows the generalized stacking fault energy curve for basal slip as predicted by DFT, the two MEAMs potentials, and multiple NNPs. All the NNPs perform much better than the original MEAM1 potential over the entire range of slip but especially for the unstable stacking fault energy γ_{usf} that is important for fracture (see Sec. IIID). Many of the NNPs provide a better basal stable stacking fault energy than either of the MEAMs, as noted above. Figure 3 further shows that it is not only the extremal values presented in Fig. 2 that are in good agreement but rather the entire curves. The entire GSFE curve enters into the widely used Peierls-Nabarro

model for the dislocation structure (distribution of Burgers vector along the slip plane) and so is valuable to capture well.

As with traditional potentials, the choice of one particular NNP over any of the other NNPs is based on heuristics associated with preferred applications of the potential. It would be possible to determine a single *best* overall NNP by minimizing the relative error with respect to DFT across all basic material properties computed, but this also implies a relative weighting of the importance of structures according to their energies. So there is no unbiased way to select a preferred NNP.

It is inherent in the formulation that NNPs can fail in extrapolation, with the typical example being failure under high compression. Figure 4 shows the equation of state, i.e., the energy per atom versus volume at fixed c/a ratio, as predicted by all the NNPs and with the reference or training DFT values also shown. Indeed, below the minimum relative volume 0.9 in the DFT training set, the NNPs show wide variability, and with some cases showing unstable collapse of the structure below ~ 0.85 . Such behavior does not occur for the two MEAMs potentials because these potentials impose an ad hoc strongly repulsive pair potential precisely to achieve reasonable behavior in compression. In fact, in the original EAM formalism, the pair potential was purely repulsive [62,63]. Three aspects about this behavior merit discussion. First, in many application cases outside of shock loading, this unphysical domain will never be encountered—in the region where NNPs remain accurate, the energies per atom and forces are too high for configurations to enter into the unstable domain. For instance, we have executed an extended molecular dynamics (MD) simulation at a temperature slightly below melting and examined the global and local energy fluctuations as predicted by the selected NNP63 and by the MEAM1 potential. We find no statistical differences at all. So no atom pairs are ever encountering the close separations at which NNP63 becomes unstable. Second, it is trivial to add a strongly repulsive pair potential to the NNPs formalism that operates only in this rarely sample region of phase space. For instance, a Morse potential [64] of the form $V(r) = D_e [1 - \exp(-(r - r_0)/a_0)]^2 - 1$ with $r_0 = 0.9 a$, a small a_0 , and large D_e will provide a strong repulsive contribution that will only be invoked if rare atom pairs attempt to move to within $0.9 a$ (again, rare under nearly all situations). Third, the lowest value of the DFT results presented in Fig. 4 is at 0.9 relative stretch of the unit cell. The reason for this is that lower values of relative stretches do not converge in DFT. Hence adding DFT data points below 0.9 is not possible. One way to add data points below 0.9 would be by manually extrapolating the repulsive part, which is manual intervention based on the assumption of repulsion without knowing the actual values. A test training using such a strategy yielded the expected result: The repulsive part is well predicted. Thus, this common failure of the NNPs in high compression is easily rectified but also of very limited importance in realistic applications.

The material properties for the chosen potential NNP63 in Fig. 2 demonstrate that using the mean and standard deviation of a family of NNPs can be misleading. These quantities suggests that one can find a single potential for which all properties fall within the mean and standard deviation, which is not the case in general. Selecting a single NNP thus involves some

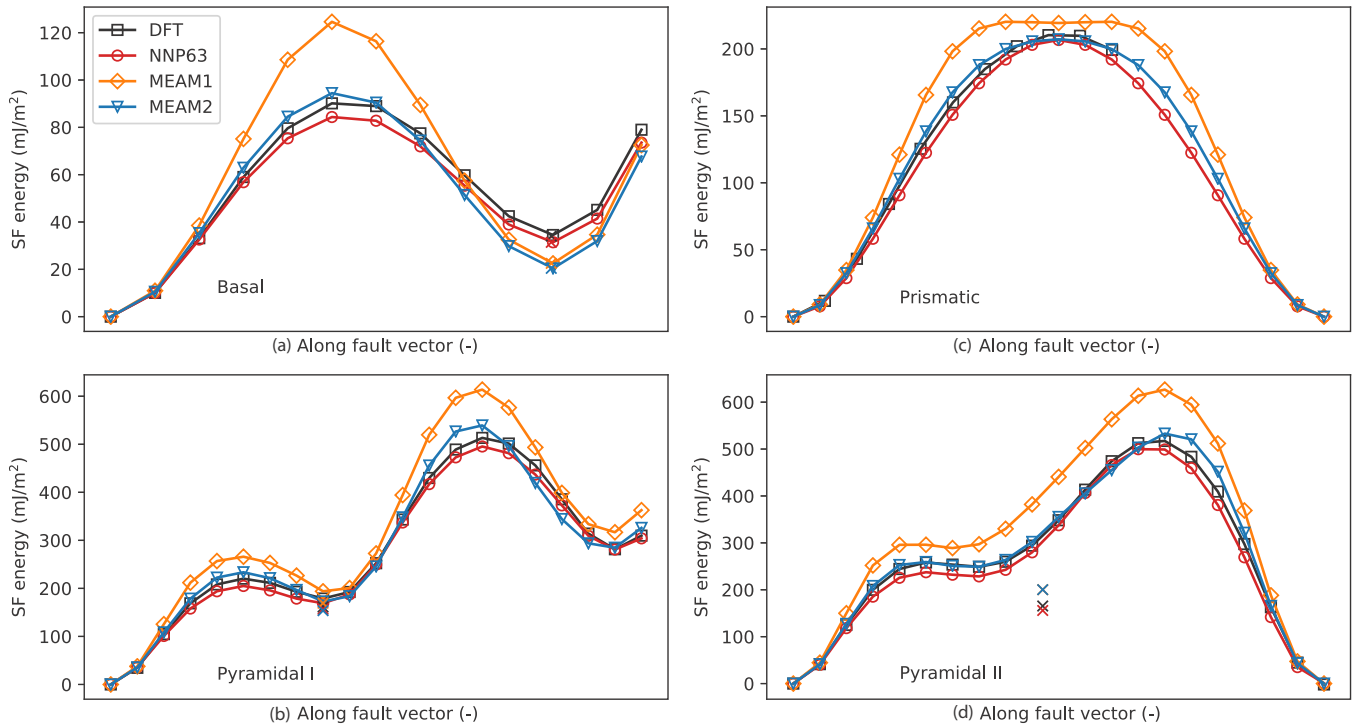


FIG. 5. Generalized stacking fault energy curves for basal, prism, pyramidal I, and pyramidal II planes; comparison among DFT, NNP63, and the two MEAM potentials. The prismatic DFT values are not part of the training set. Fully relaxed stable stacking faults are indicated with an “x” in the respective color.

external decision by the developer. For NNP63, the lattice parameters (a , c/a) and cohesive energy (E_c) all show a very low relative error but the elastic constants and stacking fault energies have large errors well outside the standard deviation across the entire family of NNPs. It should also be noted, however, that the relative errors of the MEAM potentials also show some significant errors across all the properties studied and require an external decision by the developer. An overview of the flexibility of classical interatomic potentials in fitting material parameters can be found in Ref. [65] where the authors optimize traditional fixed functional forms and assess the trade-offs between accurate modeling of different material properties.

C. Plasticity: Stacking faults and dislocations

Plastic deformation in crystalline metals is mediated by dislocations. Many aspects of dislocation structures depend on the stable and unstable stacking fault energies and on the overall generalized stacking fault energies for slip over entire crystallographic surfaces associated with dislocation slip planes. Thus, the first step in accurate modeling of actual dislocations is accurate modeling of the GSFs and stable stacking faults. We note that the GSFs are computed by relative in-plane sliding of blocks of material, with relaxation only of the stress normal to the fault plane. This procedure does not allow atomistic relaxations on either side of the fault, but such relaxations are crucial for pyramidal slip in hcp materials [32]. Figure 5 shows the GSFs for the basal, prism, pyramidal I, and pyramidal II planes. The NNP63 predictions

are in generally very good agreement with the DFT, whereas the MEAM1 potential is rather poor in all cases, while the recent MEAM2 is nearly comparable to the NNP but still with some notable deviations. As noted earlier for basal slip, the errors in unstable stacking fault energies for the MEAM1 potential have notable consequences for fracture behavior, as discussed in the next subsection. The stable stacking fault energies are also shown in Fig. 5 (indicated by small x’s); the pyramidal II case in particular shows a difference of a factor 1.5–2 for the out-of-plane relaxations vs. rigid shift value. The differences versus DFT were discussed earlier.

The GSFs for prismatic slip was not in the training set and so is a true prediction. The absence of any local minimum at the symmetric midpoint has key consequences for the stability and structure of the prismatic dislocations. The spurious shallow local minimum predicted for the MEAM1 potential leads to the prediction of a stable $\langle a \rangle$ Burgers vector screw dislocation on the prism plane, while DFT shows such a dislocation to be unstable relative to the $\langle a \rangle$ screw dislocation on the basal plane. In contrast, the NNP63 potential is in agreement with DFT, predicting that the $\langle a \rangle$ prism screw dislocation is not stable. The prismatic $\langle a \rangle$ edge dislocation is stable for all potentials. We simulate the critical resolved shear stress, or Peierls stress, to initiate glide for the prismatic edge dislocation as $\sigma_P \approx 4$ MPa, in line with expectations from the literature [66].

We now examine some of the stable dislocation core structures predicted by NNP63. Dislocations control essentially all of the plastic flow phenomena in metals and so are responsible for the remarkable combination of strength and toughness of

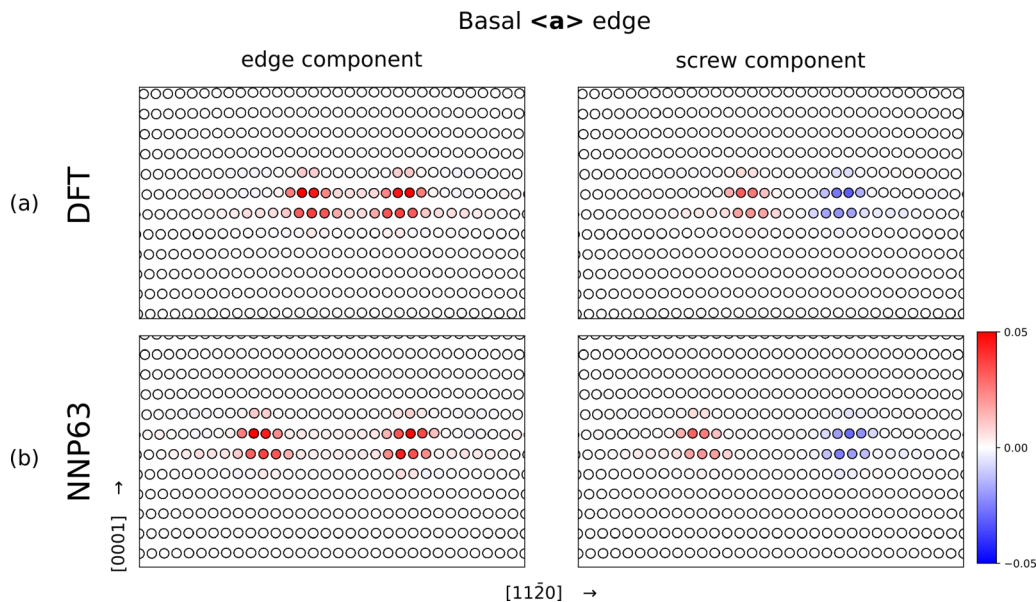


FIG. 6. Dissociated core structure of the basal (a) edge dislocation with superimposed Nye tensor components of the edge (left column) and screw (right column) components from DFT (a) and NNP63 (b). DFT results in (a) are from Ref. [72].

metal alloys. The mechanical behavior at the macroscopic scale is then directly determined by the detailed atomistic structure of the dislocations in the metal.

All dislocation geometries are constructed using the same protocol. For a specified Burgers vector \mathbf{b} and line direction ζ , a cuboidal simulation cell is constructed with x along the glide direction, y along the line direction ζ , and $z = \mathbf{b} \times \zeta$ normal to the glide plane. The length along y is the minimum periodic length for the given line direction and Burgers vector, and periodic boundary conditions are applied in the line direction. The other two dimensions are $L_x = L_z = 300$ nm. Two partial dislocations are introduced at a separation distance of ~ 10 Å [54] and all atoms are displaced according to the anisotropic elastic Volterra fields of the two partial dislocations, as implemented in Ref. [67]. The positions of all atoms within a distance $\sim 2r_c$ (where r_c is the cutoff of the potential) from the outer boundaries are held fixed at the anisotropic elastic Volterra solution according to a perfect dislocation. The remaining interior atoms are then relaxed to the minimum-energy configuration via the conjugate gradient method. Only the pyramidal I screw dislocation is obtained differently, as follows. A relaxed dissociated pyramidal II screw dislocation core (obtained as described above) is heated up to 100 K and MD simulation is executed with the outer atoms held fixed (the elastic field of the pyramidal I and II screws being identical). Snapshots from the MD trajectory are saved periodically. Each of those snapshots is then relaxed to $T = 0$ K minimum-energy configuration and the structure is analyzed. This procedure allows for a meaningful comparison of energy differences between pyramidal I and II dislocation cores generated with exactly the same boundary conditions. Since the pyramidal II core has a lower energy, it is the most common outcome, but pyramidal I cores do occur.

For all dislocations, the resulting structures are analyzed using the Nye tensor [68] as implemented in Ref. [67], which

compares the dislocation structure to a reference structure and reveals the flux of Burgers vector passing through any closed loop normal to the dislocation line. The edge and screw components of the Nye tensor are computed at each atomic position. We note that analyses based on the differential displacement [69], common neighbor analysis [70], or the dislocation extraction algorithm [71] would also be possible and yielded similar results in limited tests.

Figure 6 and Figure 7 show the dislocation core structures with superimposed Nye tensor components of the basal (a) edge and screw dislocations, respectively, as predicted by DFT and NNP63. The Nye tensor components for DFT are recomputed from Ref. [72]. For ease of comparison, we have aligned the right partials in all figures.

For the basal (a) edge dislocation in Fig. 6 the Nye tensor components for NNP63 agree extremely well with DFT for both edge and screw components. A summation of the edge component yields the exact Burgers vector, while a summation over the screw component results in zero, as required. Compared with DFT, NNP63 predicts a larger partial dislocation separation consistent with a combination of a slightly lower stable stacking fault energy and slightly higher elastic constants. The Nye tensor components for the basal (a) screw dislocation are shown in Fig. 7. As for the edge dislocation, the NNP63 core agrees very well with DFT. And as required, the edge components sum to zero and the screw components sum to the exact Burgers vector. A small difference is noticeable in the distribution. While there is no indication of partial dislocation separation for both cores, the Nye tensor components from NNP63 are slightly less compact than the reference DFT.

Predictions of the core structure of the MEAM1 (not shown here) for the basal (a) dislocations compare less favorably to the DFT reference. While the partial dislocation structures are predicted well, both the edge and screw dislocation cores

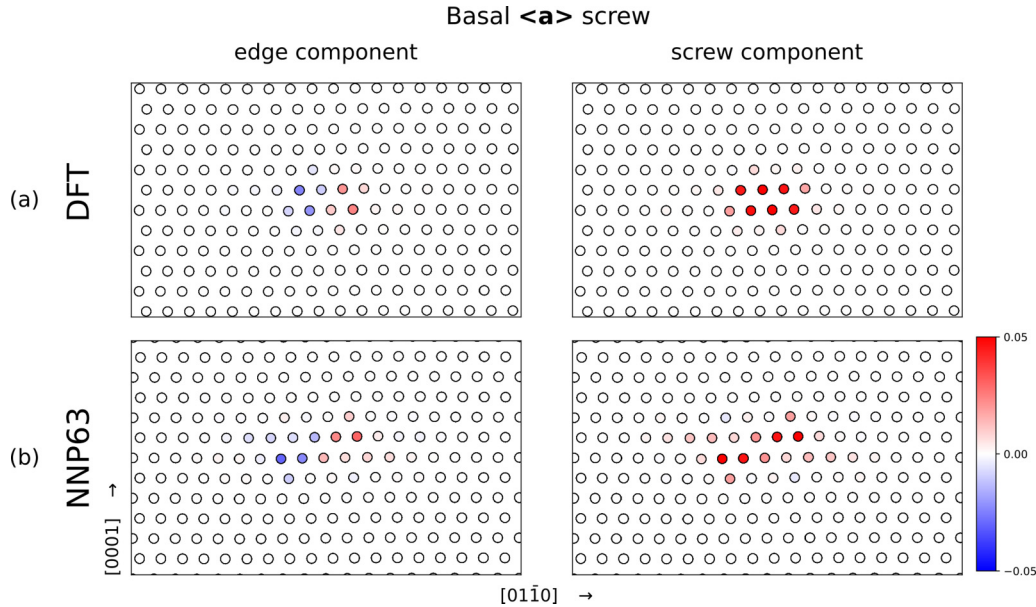


FIG. 7. Dissociated core structure of the basal (a) screw dislocation with superimposed Nye tensor components of the edge (left column) and screw (right column) components from DFT (a) and NNP63 (b). DFT results in (a) are from Ref. [72].

exhibit a much larger separation distance, see, e.g., Ref. [5] for comparison, due to the large error in basal stable stacking fault energy.

We have also examined the Peierls stress of the basal $\langle \mathbf{a} \rangle$ edge dislocation. The basal $\langle \mathbf{a} \rangle$ dislocation is the prevalent system in pure Mg, with a very low Peierls stress that controls the measured yield strength in all but idealized crystal orientations. Using standard procedures, NNP63 exhibits a $T = 0$ K Peierls stress of $\sigma_p \approx 0.05$ MPa for the basal edge dislocation. This is smaller than experiments (~ 0.5 MPa); other potentials show results higher than experiments, but all values remain quite small so that alloying to increase the basal strengthening is usually necessary for an engineering material.

Dislocations with $\langle \mathbf{c} + \mathbf{a} \rangle$ Burgers vector on the pyramidal I and II planes are important in Mg because they provide the accessible plastic slip in the $\langle \mathbf{c} \rangle$ axis direction. Figure 8 shows the Nye tensor analysis for the pyramidal II edge core from DFT and NNP63. The pyramidal II edge core is relevant for investigating solute strengthening effects [73,74]. Dissociation is clearly along a pyramidal II plane. Comparing the Nye components of the two partial cores yields a very good match between DFT and NNP63. The edge components match especially well while the screw components are smaller and agree slightly less well. The partial dislocation separation is slightly larger for the NNP63 prediction, again in line with the slightly smaller stacking fault energy.

Of perhaps most importance for the $\langle \mathbf{c} + \mathbf{a} \rangle$ Burgers vector dislocations is the energy difference $\Delta E_{I-II} = \Delta E_I - \Delta E_{II}$ between pyramidal I and II screw dislocations, where a positive value indicates that pyramidal II is more stable (lower energy). This energy difference has been identified as the key material parameter for ductility in Mg due to its role in controlling a cross-slip process that enables significant $\langle \mathbf{c} + \mathbf{a} \rangle$ slip [34]. However, this energy difference is very small and so is not well established in DFT itself. Using a quadrupolar DFT cell [75] reported an energy difference of $\Delta E_{I-II} = 7 \pm$

20 meV/nm (pyramidal II more stable); the small cell size and periodic boundary conditions can have a nonnegligible effect on the total energy, however. We have performed DFT using a very large (up to 2000 atoms) cylindrical geometry with outer boundaries fixed at the anisotropic elastic displacement field for a Volterra $\langle \mathbf{c} + \mathbf{a} \rangle$ screw dislocation using the DFT lattice and elastic constants. We find that the sign of the energy difference varies with increasing cell radius, with $\Delta E_{I-II} = -30$ meV/nm at the largest size (pyramidal I favorable, in conflict with experiments). So even DFT is uncertain here, and the NNP is trained only on DFT structures that do not include actual dislocations. Additional complexity related to the pyramidal slip systems arise during the stable stacking fault calculations during which complex shuffling processes are observed which lead to a lower energy, see Ref. [54], with the actual stable fault positions and energies marked with \times 's in Fig. 5 in the pyramidal II GSFE. Actual dislocations are even more complex in geometry and might include further shuffling processes during relaxation that are not present in the training dataset by design and hence are probably not fully captured by the NNPs or any other potential.

With the above background, Figs. 9(a) and 9(b) show the screw dislocation core structures for the pyramidal I dislocation with the edge and screw Nye tensor components superimposed as obtained from our DFT calculation (noting that the structure is quite robust with respect to cell size) and as predicted by NNP63. Both cores are split into two partial dislocations and the partial cores show a very good agreement: large screw components and small edge components. A summation of the screw components yields the expected Burgers vector, and the summation over the edge components is zero. The screw components show a dissociation along the pyramidal I plane with a short section along the pyramidal II plane. The relative positions of the partial dislocations differ slightly. The segment along the pyramidal II plane is slightly longer for the DFT geometry, which results in the

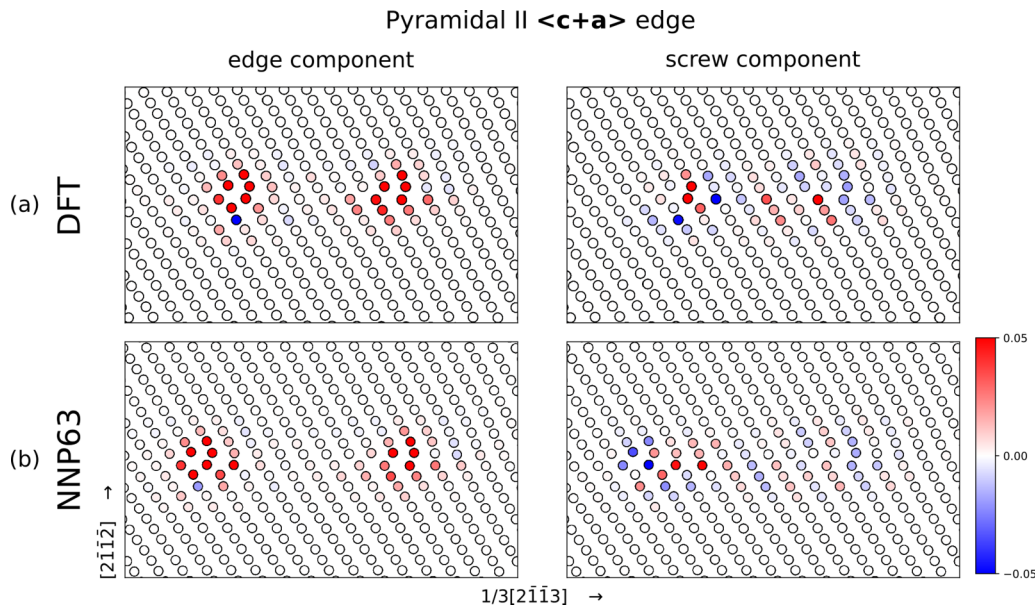


FIG. 8. Dissociated core structures of the pyramidal II $\langle \mathbf{c} + \mathbf{a} \rangle$ edge dislocations with superimposed Nye tensor components of the edge (left column) and screw (right column) components; (a) DFT from Ref. [73]; (b) NNP63.

right partial dislocation being two pyramidal I planes below the left partial as opposed to just one for the NNP63 geometry. Overall, the level of agreement for this complex defect is good.

Figures 9(c) and 9(d) show the structure and Nye tensor components for the pyramidal II screw dislocation from DFT and NNP63, respectively. The DFT results are reanalyzed from Ref. [76]. The partial separation distance is similar but there are notable differences in structure. Most importantly, the DFT core structure shows a dissociation in the pyramidal II plane while the NNP63 core shows what appears to be some dissociation on the pyramidal I plane. That is, the left and right partials in Fig. 9(d) are not similar, with the left partial of the NNP63 geometry extending further along a pyramidal I plane. This quasimixed pyramidal II-pyramidal I structure obtained for the NNP63 pyramidal II core presumably reflects the delicate energetic competition between pyramidal II and pyramidal I cores, which NNP63 is not able to fully resolve. The pyramidal II core as predicted by the MEAM1 (not shown), on the other hand, shows dissociation only on the pyramidal II plane (cf. Ref. [6]) and so its structure is in better agreement with DFT. However, the pyramidal II core as predicted by the MEAM2 potential bears some similarity to that predicted by NNP63, i.e., the left partial extends slightly along the pyramidal I plane, and the right partial is in less-good agreement with DFT. Thus, the possible error in structure for NNP63 may not be unique to the type of potential.

The key pyramidal I-pyramidal II energy difference ΔE_{I-II} varies with the potential. The MEAM1 potential predicts an energy difference of $\Delta E_{I-II} = 27$ meV/nm, while the MEAM2 potential gives an energy difference of $\Delta E_{I-II} = 54$ meV/nm (pyramidal II stable for both, consistent with experiments). Another version of the MEAM2 was developed to make pyramidal I energetically favorable with $\Delta E_{I-II} = -28$ meV/nm with negligible changes to other properties [6]. Both the uncertainty in the DFT studies and the fact that the

MEAM can be tweaked to change the absolute stability of pyramidal I vs. pyramidal II is an indication that the energy balance is very subtle. The NNP63 potential based solely on DFT inputs predicts a pyramidal I-pyramidal II energy difference of 170 meV/nm so that pyramidal II is much more stable than obtained by other estimates. This is in spite of the observation that the relaxed pyramidal II structure appears to be a mix of a pyramidal II-like partial and a pyramidal I-like partial. Limited tests with other NNPs show similar dislocation cores and energy differences.

The procedure to obtain the pyramidal I dislocation screw core by the annealing process described earlier yields many screw dislocation cores in pyramidal I and II planes as well as the mixed type in the case of NNP63. Their structures are generally very similar with undetectable differences but exhibit differences in energy. Those energy differences are presented in Fig. 10 for NNP63 and both MEAMs. The reference energy for NNP63 is the energy corresponding to the mixed-type pyramidal II structure presented in Fig. 9(d). The references for the two MEAMs are the pyramidal II dislocation cores. Nearly all of the cores created with NNP63 are found to be either identical to, or mostly less than ~ 10 meV/nm above, the lowest-energy structure. The energies cluster into just a few distinct values. One structure had an energy ≈ 170 meV/nm that is close to the pyramidal I screw energy, and one structure had an energy ≈ 125 meV/nm. Overall, however, and unlike assumptions that ML potentials may have many local spurious minima, NNP63 is quite robust for this defect. In contrast, the energy differences among the MEAMs are much more widely distributed, with a range of distinct energies spanning up to ≈ 70 meV/nm for the MEAM1 and ≈ 50 meV/nm for the MEAM2 but with some structures with even higher energies extending up to the maximum found for NNP63. Thus, it is actually the MEAM potentials that have many metastable structures for this complex defect.

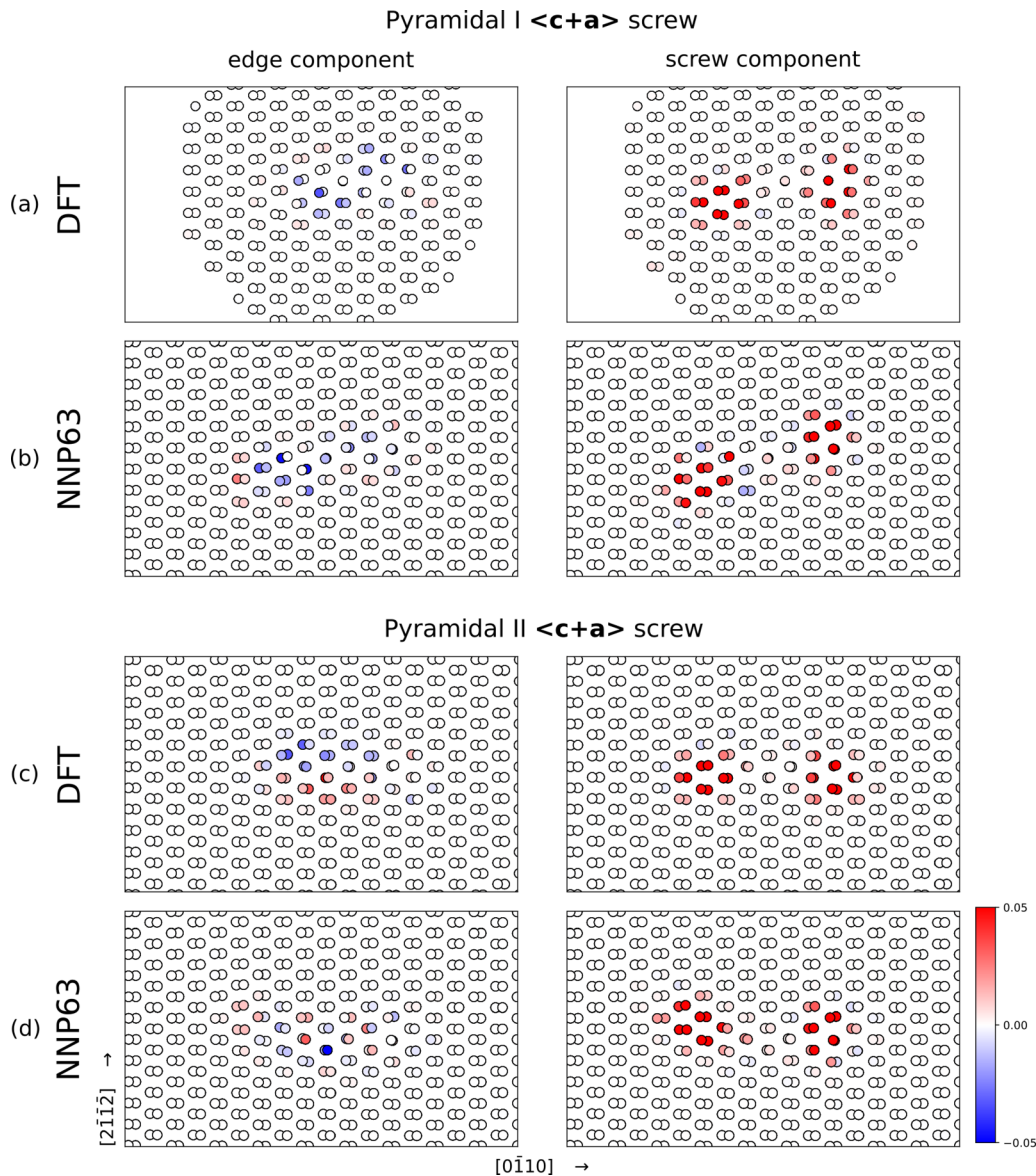


FIG. 9. Dissociated core structure of the pyramidal I and II $\langle c+a \rangle$ screw dislocation with superimposed Nye tensor components of the edge (left column) and screw (right column) components; (a) pyramidal I core from DFT; (b) NNP63 pyramidal I core; (c) pyramidal II core from DFT; and (d) the NNP63 prediction. The missing atoms in subfigure (a) are related to the procedure in DFT; the reader is referred to the text for an explanation. Results for the pyramidal II in (b) from DFT are replotted from Ref. [73]. DFT geometries for pyramidal I in panel (a) are our own results.

Overall, there is currently no resolution of the pyramidal I-pyramidal II screw competition based on any computational method (DFT, NNP, MEAM). This problem thus requires considerably more study at the first-principles level. The apparent mixed structure of the pyramidal II screw core as predicted by NNP63 makes it of questionable use for studies involving this dislocation. Limited studies with other NNPs in our family of 129 potentials also do not exhibit a clear pyramidal II core dissociation with the described procedure. This remains an issue for further study and resolution.

In summary, even though dislocation geometries were not part of the training dataset, the NNP63 generally reproduces well the various dislocation geometries and Burgers vector distributions found in DFT studies. The Peierls stresses for basal and prism $\langle a \rangle$ are also in reasonable agreement with

experiments. Only the pyramidal II screw dislocation shows a different behavior that may be related to the inability of NNP63 to capture the (uncertain, even in DFT) energy differences between the pyramidal I and II partial dislocations. While we are able to compare to available DFT studies, such studies are very computationally intensive, and are not at all feasible for properties beyond the basic structure of a straight periodic dislocation line with high-symmetry character (e.g., edge or screw). The success here of the NNP opens the study of many other plasticity phenomena at metallurgically relevant length and timescales.

D. Fracture: Decohesion and intrinsic ductility

Fracture is a complex process involving very high multi-axial stresses around the crack tip with atomistic regions

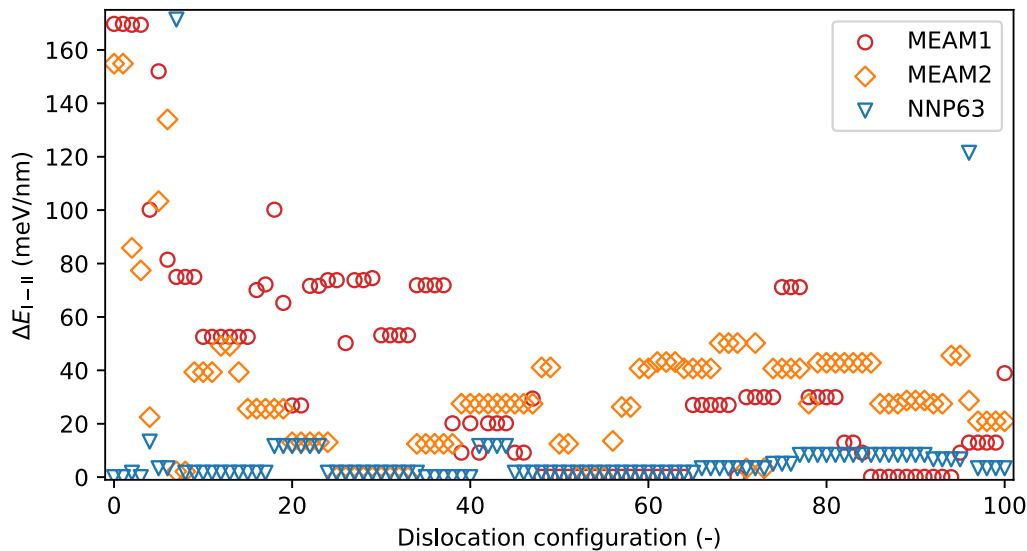


FIG. 10. Energy differences to pyramidal II cores for various relaxed screw dislocation geometries in on pyramidal planes from the two MEAMs and NNP63.

spanning from fully decohered to moderately deformed perfect crystals. Capturing proper fracture behavior, and even simply avoiding totally unphysical behavior at the crack tip, represents a high challenge for any interatomic potential. Various physically motivated potentials (EAM [63], MEAM [77], Stillinger-Weber [78], etc.) and the machine learning GAP Fe potential [23] all give unphysical crack-tip behavior in spite of making good predictions for many other material properties and defects. Fracture is thus a very demanding test for any interatomic potential.

Moreover, the precise behavior at a crack tip determines whether a material is intrinsically brittle or intrinsically ductile. An intrinsically brittle material will cleave and create new surfaces. An intrinsically ductile material will emit dislocations, blunting the crack and preventing the brittle cleavage failure. It is computationally prohibitive to study cracks accurately with DFT because large sizes are needed, the geometry is nonperiodic, and many load levels are needed to find the cleavage or emission load. Limited studies to date have thus relied on multiscale methods [79–83]. This makes the framework of anisotropic linear elastic fracture mechanics (LEFM) valuable for predicting both cleavage and dislocation emission. For Mg in particular, the cleavage and emission phenomena are predicted to occur at nearly the same load levels for a number of crack orientations, making it a particularly sensitive case. Modeling fracture behavior in Mg and Mg alloys thus again requires highly accurate interatomic potentials.

Fracture in mode I (tensile) loading is controlled by the stress intensity factor K_I , which is the strength of the square-root singularity in the crack-tip stress field. The thermodynamic critical stress intensity factor for cleavage in mode I for an atomically sharp crack is given by [84]:

$$K_{Ic} = \sqrt{\frac{2\gamma_s}{\Lambda_{22}}}, \quad (8)$$

where γ_s is the surface energy and Λ is an elasticity parameter [85]. A critical stress intensity for dislocation emission in

mode I cannot be determined exactly but has been accurately estimated as [86,87]

$$K_{Ic} = \sqrt{G_{Ic}^{\text{first}}(\theta, \varphi_{\text{first}}) o(\theta, \varphi) / |F_{12}(\theta)|}, \quad \text{with} \quad (9)$$

$$G_{Ic}^{\text{first}} = \begin{cases} 0.145 \gamma_s^e + 0.5 \gamma_{\text{usf}}^e, & \gamma_s^e > 3.45 \gamma_{\text{usf}}^e \\ \gamma_{\text{usf}}^e, & \text{otherwise,} \end{cases} \quad (10)$$

where γ_s^e and γ_{usf}^e are the surface energy and unstable stacking fault energy of the emission plane, respectively; o is an anisotropic elastic coefficient; θ and φ are the inclinations of the slip plane and Burgers vectors; F_{ij} is the resolved applied K field to an effective shear along the dislocation slip plane; and indices 1 and 2 refer to the crack growth direction and normal to the crack plane, respectively [88]. The LEFM analysis thus shows that the critical material parameters for assessing cleavage and emission are the surface energies, the unstable stacking fault energies, and the elastic constants (which enter both phenomena, so differences mainly depend on the anisotropy, not the absolute values). With this background, we note that we have already examined the predictions of NNP63 for the elastic constants, surface energies, and unstable stacking fault energies. NNP63 provides much better accuracy for the surface energies as compared to both MEAM1 and MEAM2 and is better than MEAM1 for the unstable stacking fault energies. Since the intrinsic ductility depends on both quantities, we can expect NNP63 to be more realistic for assessing the fracture behavior of Mg.

Before examining atomistic crack tips, we first study the atomistic decohesion energy, i.e., the energy versus separation distance for rigid separation of two blocks of material across specified surface. This decohesion curve is the “cohesive zone” that determines the crack-tip shape during loading [89]. Figure 11 shows the decohesion energy versus normal separation distance as computed via DFT, NNP63, and the two MEAMs for the basal, prismatic, pyramidal I, and pyramidal II planes. The DFT data were contained in the training set but here, for clarity, not all data points are shown. The

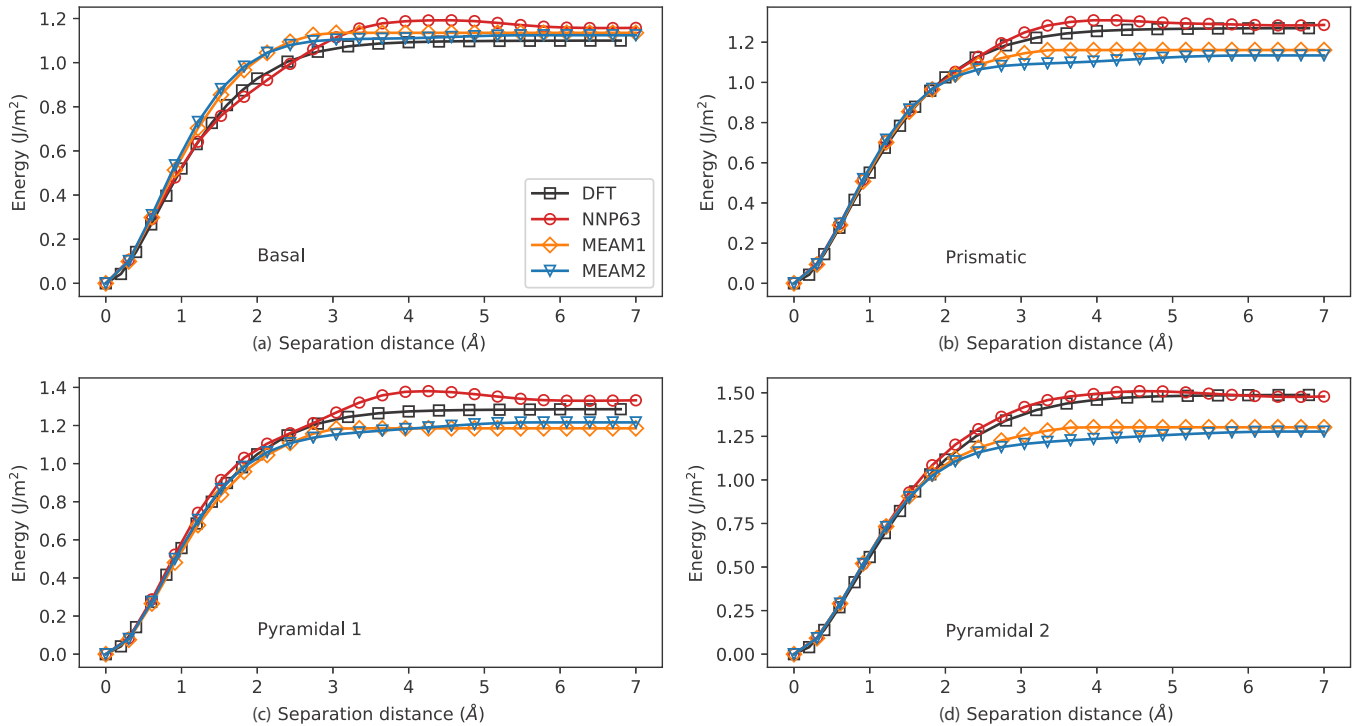


FIG. 11. Decoherence curves for basal, prism, pyramidal I, and pyramidal II planes; comparison among DFT, NNP63, and the two MEAM potentials. All DFT data points are part of the training set.

energy at maximum separation is twice the unrelaxed surface energy and shows deviations for the MEAM potentials that are comparable to the deviations found for the fully relaxed surfaces (Table I). In contrast, results for NNP63 generally agree better with DFT except for basal decohesion. The maximum cohesive stress corresponds to the maximum slope in the decohesion curve and is well captured by all potentials except for the MEAM potentials under basal separation. NNP63 shows a small maximum in the energy prior to full separation. This is unphysical; beyond this maximum, the surfaces are being pushed apart rather than pulled together. However, the effect is small and does not have broader significance.

We now examine the phenomena occurring at the tip of a sharp crack via simulation. In this study, we assess the whether the NNP gives physical behavior and, if so, whether the intrinsic ductility of Mg for relevant fracture orientations is consistent with theoretical predictions. We use the semi-infinite crack “ K -test” for high accuracy [88] with an atomistic domain size of $L_x = 500$ nm, $L_y = 500$ nm, and $L_z = 10$ nm with x the crack growth direction, y the normal to the crack plane, and z the crack line direction with periodicity imposed in this direction. The theoretical critical values of K_{Ic} and K_{IIc} are computed, and all atoms are displaced according to the anisotropic K -field solution at a value of K_I just below the lower of the two critical values. Atoms within $\sim 2r_c$ of the outer x - y boundaries are held fixed and the energy is then minimized by relaxing all interior atoms. Increments in K_I are then imposed until some simulated critical value is reached at which the system undergoes either a cleavage or emission event. As a first result, no anomalous or unphysical behavior is observed for any crack geometry when using NNP63; this is already a very positive result.

Table II shows the predicted and observed results obtained using NNP63. In all cases, the observed or simulated behavior coincides with the predicted behavior using the DFT material properties, cf. Table V. The simulated and predicted critical values differ slightly, but this is typical due to the well-known existence of lattice trapping for cleavage fracture and the nonexact (although generally quite accurate) theory for emission. The only notable quantitative deviation between simulation and theory is for the pyramidal II orientation case where the simulated emission occurs at $0.82 K_{IIc}^{\text{th}}$, i.e., much lower than the theoretical value. NNP63 thus provides very good predictions for all fracture orientations studied.

Table II also shows the events obtained from MEAM1 simulations as reported previously [90]. While simulations using MEAM1 agree with theoretical predictions using the material properties of MEAM1 [90] (not shown here), the operative MEAM1 event differs from the NNP63 event for some orientations while the NNP63 event agrees with the event predicted using DFT material properties. Specifically, the MEAM1 predicts three orientations (prismatic I, prismatic II, and pyramidal II) to be brittle (cleavage), whereas NNP63 predicts them all to be ductile (emission).

Figure 12 shows the geometries and structures after these events, highlighting the fundamentally different physical behavior. These differences are important because they indicate that Mg is not as brittle as suggested by studies based on the MEAM1 potential. The difference in phenomenon is directly traceable to the differences in underlying material properties. The MEAM1 potential provides poorer predictions of the surface and/or unstable stacking fault energies as compared to the NNP63 potential, and this leads to different behavior in simulations. Furthermore, while both potentials

TABLE II. Stress intensity factors K_I for cleavage and emission for various crack orientations as computed and simulated using NNP63 and its material properties. The observed events using the MEAM1 potential are also shown [90]. Critical K values in bold indicate the predicted event. Cases where the MEAM1 event differs from the NNP63-observed and DFT-predicted event are indicated in italics.

| Crack plane | Predicted critical K and event | | | Observed NNP63 | | NNP63 vs. Pred $K^{\text{NNP}}/K^{\text{pred}}$ | MEAM1 Event [90] |
|--------------|----------------------------------|----------------------|----------|----------------|----------|--|---------------------|
| | K_{Ic}^{th} | K_{Ie}^{th} | Event | K | Event | | |
| Basal I | 0.265 | 0.284 | Cleavage | 0.280 | Cleavage | 1.06 | Cleavage |
| Basal II | 0.265 | 0.291 | Cleavage | 0.263 | Cleavage | 0.99 | Cleavage |
| Prismatic I | 0.297 | 0.253 | Emission | 0.236 | Emission | 0.93 | <i>Cleavage</i> |
| Prismatic II | 0.307 | 0.282 | Emission | 0.283 | Emission | 1.00 | <i>Cleavage</i> |
| Pyramidal I | 0.284 | 0.237 | Emission | 0.250 | Emission | 1.05 | Emission |
| Pyramidal II | 0.302 | 0.286 | Emission | 0.220 | Emission | 0.82 | <i>Cleavage</i> |

predict the basal orientation to be brittle, NNP63 predicts a much closer competition, with $K_{Ie}^{\text{th}}/K_{Ic}^{\text{th}} \approx 0.93$ for NNP63 as compared to ≈ 0.73 for MEAM1 [90]. This suggests the possibility that dilute alloying could have a more important effect in changing the brittle basal behavior than obtained by recent studies on a model Mg-Y alloy using an alloy MEAM potential [90].

IV. DISCUSSION

We have presented a neural-network interatomic potential for magnesium using the Behler-Parrinello symmetry functions for the description of the atomic environments within the n2p2 [37] framework. Magnesium has been studied due to (i) its important technological value, (ii) the complexities associated with dislocations and slip in the hcp crystal structure, (iii) the close competition between intrinsic brittleness and intrinsic ductility, and (iv) the need to understand the atomistic mechanisms that enhance the performance of dilute Mg alloys relative to pure Mg, which will require interatomic potentials for complex alloy systems. Compared to existing very good MEAM potentials, NNP63 generally performs better for many material properties and crystalline defects that are critical to mechanical performance. These improvements were accomplished with a rather small DFT training dataset, extended

only beyond a relatively standard set of structures to enable accurate fracture modeling. The addition of further selected data to improve on specific defect properties (as was done here for fracture) is easily accomplished within the parameter-rich NNP framework. Because the NNP formalism is regression, extrapolations could be highly inaccurate, and so the present potential was carefully assessed in many scenarios relevant for plasticity and fracture.

The current potential remains less than ideal for the pyramidal II screw dislocation structure. The important pyramidal I-II screw energy difference will likely require improvements as well. However, accurate DFT reference data are not yet available and may remain challenging due to the very small energy difference. The Mg potential NNP63 was also not tested for grain boundary structures or interstitial atoms and so is not yet a fully general potential for Mg. Nonetheless, the delicate tension twin boundary energy (and structure) is well predicted although not in the training set. Both the vacancy formation and migration energies are also in good agreement with DFT while, along with the dislocation and crack structures, not being in the training dataset. The melting point of NNP63 is also in good agreement with experiments. These results are encouraging not only for Mg but also for the broader prospects for machine learning potentials.

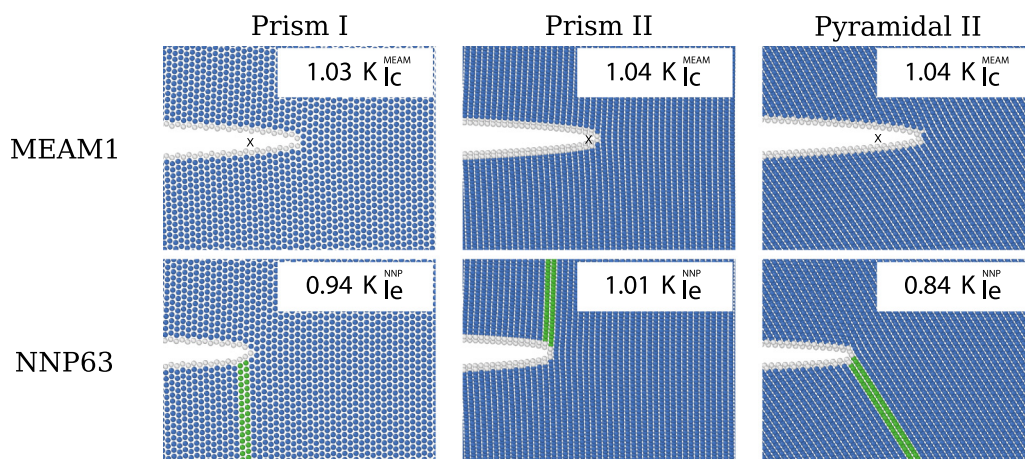


FIG. 12. Cross-sectional view of crack tips post fracture for prismatic I, prismatic II, and pyramidal II crack planes, as observed in simulations for the MEAM1 and NNP63. The corresponding K values at each event are shown relative to K_{Ic} for the MEAM and K_{Ie} for NNP63 in the respective crack plane. Visualization used the common neighbor analysis [70] (blue = hcp, green = fcc, gray = other) and the black “x” denotes the initial crack position for cleavage events.

TABLE III. Hyperparameters for the radial symmetry functions.

| Type | η | r_2 | r_c |
|------|------------------------|---------------------|---------------------|
| 2 | 2.630×10^{-2} | 0.0 | 8.0 |
| 2 | 1.560×10^{-2} | 0.0 | 8.0 |
| 2 | 2.630×10^{-2} | 0.0 | 8.0 |
| 2 | 6.900×10^{-3} | 0.0 | 1.200×10^1 |
| 2 | 4.420×10^{-2} | 0.0 | 8.0 |
| 2 | 1.560×10^{-2} | 0.0 | 8.0 |
| 2 | 7.430×10^{-2} | 0.0 | 8.0 |
| 2 | 7.430×10^{-2} | 0.0 | 8.0 |
| 2 | 4.420×10^{-2} | 0.0 | 8.0 |
| 2 | 1.560×10^{-2} | 0.0 | 8.0 |
| 2 | 1.117×10^{-1} | 1.301×10^1 | 1.600×10^1 |
| 2 | 1.821×10^{-1} | 5.657 | 8.0 |
| 2 | 4.420×10^{-2} | 0.0 | 8.0 |
| 2 | 1.170×10^{-2} | 0.0 | 1.200×10^1 |
| 2 | 1.081×10^{-1} | 1.322×10^1 | 2.0×10^1 |
| 2 | 7.150×10^{-2} | 1.626×10^1 | 2.0×10^1 |
| 2 | 9.110×10^{-2} | 8.0 | 1.600×10^1 |
| 2 | 4.550×10^{-2} | 1.131×10^1 | 1.600×10^1 |
| 2 | 1.166×10^{-1} | 7.071 | 2.0×10^1 |
| 2 | 5.830×10^{-2} | 1.0×10^1 | 2.0×10^1 |
| 2 | 2.910×10^{-2} | 1.414×10^1 | 2.0×10^1 |
| 2 | 3.900×10^{-3} | 0.0 | 1.600×10^1 |
| 2 | 2.500×10^{-3} | 0.0 | 2.0×10^1 |
| 2 | 5.560×10^{-2} | 0.0 | 1.200×10^1 |
| 2 | 8.700×10^{-3} | 0.0 | 2.0×10^1 |
| 2 | 1.350×10^{-2} | 0.0 | 1.600×10^1 |
| 2 | 1.960×10^{-2} | 0.0 | 1.200×10^1 |

The present well-converged and consistent DFT training set developed here is openly available [43]. This enables its application to other machine learning methods and/or other optimization functions that might better handle the structurally inhomogeneous datasets typical for metallurgical applications. In particular, it would be valuable to develop a loss function that is formulated in terms of energy differences between structures rather than absolute energies of individual structures. In any case, the quality of a potential for metallurgical applications must not only be measured by the RMSE. Substantial tests outside of the training dataset and/or to challenging but realistic structures (e.g., dislocations, crack tips) must be included. The performance comes at a cost. Compared to the MEAM formalism as implemented in LAMMPS, the n2p2 implementation [37] for BPNNs is $10\times$ more expensive computationally. Adding more elements to create interatomic potentials for alloys increases the cost significantly [1]. However, if the aim of atomistic studies is to be quantitatively correct for specific alloy compositions, then it is likely that only machine learning methods will provide an accuracy approaching that of DFT while at a tiny fraction of the computational cost, even if this cost significantly exceeds that of traditional potentials.

Finally, although the potential presented here is not a general-purpose Mg potential, this work is among the first to demonstrate the broad application of a machine learned potential to a cross section of metallurgically relevant properties, structures, and behaviors. NNP63 clearly improves on

TABLE IV. Hyperparameters for the angular symmetry functions.

| Type | η | λ | ζ | r_c |
|------|------------------------|-----------|---------|---------------------|
| 3 | 6.900×10^{-3} | 1 | 1.0 | 1.200×10^1 |
| 3 | 6.900×10^{-3} | 1 | 4.0 | 1.200×10^1 |
| 3 | 6.900×10^{-3} | -1 | 1.0 | 1.200×10^1 |
| 3 | 1.170×10^{-2} | 1 | 1.0 | 1.200×10^1 |
| 3 | 1.560×10^{-2} | 1 | 1.0 | 8.0 |

existing traditional potentials that have been highly optimized and that are quantitatively among the best potentials among all those developed to date for pure metals. This work sets the baseline for extensions to Mg alloys, which are very challenging for interatomic potentials having fixed functional forms and limited flexibility and yet are essential for the technological application of lightweight Mg alloys in energy-efficient structural components. The present potential can be used to study other defect-related problems in Mg which includes, e.g., the prismatic-basal cross-slip mechanism. This is beyond the scope of the current paper and will be presented in future work.

ACKNOWLEDGMENTS

M.S., B.Y., and W.A.C. acknowledge support by the NCCR MARVEL funded by the Swiss National Science Foundation. We also acknowledge support for high-performance computing provided by the Scientific IT and Application Support (SCITAS) at EPFL. E.M. acknowledges support from the Swiss National Science Foundation through Project No. 200021_18198/1. M.S. thanks G. Imbalzano for providing the CURSEL algorithm, M. Ghazisaeidi for providing the pyramidal II dislocation core geometry from DFT, and A. Jain for providing the vacancy migration barriers.

APPENDIX A: SYMMETRY FUNCTION HYPERPARAMETERS

Symmetry function choice optimized with CUR decomposition as described in Sec. II.

APPENDIX B: FRACTURE BEHAVIOR PREDICTION FOR DFT

The discussion of the predictions and observations in Table II include a reference to the event prediction based on

TABLE V. Stress intensity factors K_{Ic} for cleavage (c) and emission (e) with material properties from the DFT for the event prediction in Table II. DFT values are taken as presented in Table I; surface energy for the prismatic II plane is taken from Ref. [91].

| Plane | K_{Ic} | K_{Ie} | Event prediction |
|--------------|--------------|--------------|------------------|
| Basal I | 0.242 | 0.255 | Cleavage |
| Basal II | 0.242 | 0.264 | Cleavage |
| Prismatic I | 0.255 | 0.229 | Emission |
| Prismatic II | 0.273 | 0.253 | Emission |
| Pyramidal I | 0.259 | 0.214 | Emission |
| Pyramidal II | 0.277 | 0.241 | Emission |

DFT values for unstable stacking fault energies and surface energies. The values behind these predictions are given here in Table V.

The event prediction is based on the minimum between the stress intensity factor for cleavage and emission (K_{Ic}, K_{IIc}) and is highlighted in the table in bold numbers.

- [1] D. Marchand, A. Jain, A. Glensk, and W. A. Curtin, Machine learning for metallurgy I. A neural-network potential for Al-Cu, *Phys. Rev. Materials* **4**, 103601 (2020).
- [2] X.-Y. Liu, J. B. Adams, F. Ercolessi, and J. A. Moriarty, EAM potential for magnesium from quantum mechanical forces, *Modell. Simul. Mater. Sci. Eng.* **4**, 293 (1996).
- [3] D. Y. Sun, M. I. Mendeleev, C. A. Becker, K. Kudin, T. Haxhimali, M. Asta, J. J. Hoyt, A. Karma, and D. J. Srolovitz, Crystal-melt interfacial free energies in hcp metals: A molecular dynamics study of mg, *Phys. Rev. B* **73**, 024116 (2006).
- [4] Y.-M. Kim, N. J. Kim, and B.-J. Lee, Atomistic modeling of pure mg and mg-al systems, *Calphad* **33**, 650 (2009).
- [5] Z. Wu, M. F. Francis, and W. A. Curtin, Magnesium interatomic potential for simulating plasticity and fracture phenomena, *Modell. Simul. Mater. Sci. Eng.* **23**, 015004 (2015).
- [6] R. Ahmad, Z. Wu, and W. Curtin, Analysis of double cross-slip of pyramidal i ($c + a$) screw dislocations and implications for ductility in mg alloys, *Acta Mater.* **183**, 228 (2020).
- [7] A. P. Bartók, R. Kondor, and G. Csányi, On representing chemical environments, *Phys. Rev. B* **87**, 184115 (2013).
- [8] J. Behler, Representing potential energy surfaces by high-dimensional neural network potentials, *J. Phys.: Condens. Matter* **26**, 183001 (2014).
- [9] R. Kobayashi, D. Gíofré, T. Junge, M. Ceriotti, and W. A. Curtin, Neural network potential for al-mg-si alloys, *Phys. Rev. Mater.* **1**, 053604 (2017).
- [10] G. Pun, R. Batra, R. Ramprasad, and Y. Mishin, Physically-informed artificial neural networks for atomistic modeling of materials, *Nat. Comm.* **10**, 2339 (2019).
- [11] H. Wang, L. Zhang, J. Han, and W. E. Deepmd-kit: A deep learning package for many-body potential energy representation and molecular dynamics, *Comput. Phys. Commun.* **228**, 178 (2018).
- [12] D. Dickel, D. Francis, and C. Barrett, Neural network aided development of a semi-empirical interatomic potential for titanium, *Comput. Mater. Sci.* **171**, 109157 (2020).
- [13] J. Behler and M. Parrinello, Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces, *Phys. Rev. Lett.* **98**, 146401 (2007).
- [14] D. Rogers and M. Hahn, Extended-connectivity fingerprints, *J. Chem. Inf. Model.* **50**, 742 (2010).
- [15] G. Montavon, M. Rupp, V. Gobre, A. Vazquez-Mayagoitia, K. Hansen, A. Tkatchenko, K.-R. Müller, and O. A. von Lilienfeld, Machine learning of molecular electronic properties in chemical compound space, *New J. Phys.* **15**, 095003 (2013).
- [16] A. P. Bartók and G. Csányi, Gaussian approximation potentials: A brief tutorial introduction, *Int. J. Quantum Chem.* **115**, 1051 (2015).
- [17] F. A. Faber, L. Hutchison, B. Huang, J. Gilmer, S. S. Schoenholz, G. E. Dahl, O. Vinyals, S. Kearnes, P. F. Riley, and O. A. von Lilienfeld, Prediction errors of molecular machine learning models lower than hybrid dft error, *J. Chem. Theory Comput.* **13**, 5255 (2017).
- [18] L. Zhang, D.-Y. Lin, H. Wang, R. Car, and W. E, Active learning of uniformly accurate interatomic potentials for materials simulation, *Phys. Rev. Mater.* **3**, 023804 (2019).
- [19] N. Bernstein, G. Csányi, and V. L. Deringer, De novo exploration and self-guided learning of potential-energy surfaces, *npj Comput. Mater.* **5**, 99 (2019).
- [20] 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).
- [21] Y. Zuo, C. Chen, X. Li, Z. Deng, Y. Chen, J. Behler, G. Csányi, A. V. Shapeev, A. P. Thompson, M. A. Wood, and S. P. Ong, Performance and cost assessment of machine learning interatomic potentials, *J. Phys. Chem. A* **124**, 731 (2020).
- [22] G. Imbalzano, A. Anelli, D. Gíofré, 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).
- [23] 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**, 013808 (2018).
- [24] 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).
- [25] 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).
- [26] B. Mordike and T. Ebert, Magnesium: Properties - applications - potential, *Mater. Sci. Eng. A* **302**, 37 (2001).
- [27] H. Friedrich and S. Schumann, Research for a “new age of magnesium” in the automotive industry, *J. Mater. Process. Technol.* **117**, 276 (2001).
- [28] T. M. Pollock, Weight loss with magnesium alloys, *Science* **328**, 986 (2010).
- [29] K. K. Turekian and K. H. Wedepohl, Distribution of the elements in some major units of the earth’s crust, *Geol. Soc. Am. Bull.* **72**, 175 (1961).
- [30] D. Hull and D. J. Bacon, *Introduction to Dislocations*, 4th ed. (Elsevier, Amsterdam, 2011).
- [31] Z. Wu and W. Curtin, Brittle and ductile crack-tip behavior in magnesium, *Acta Mater.* **88**, 1 (2015).
- [32] B. Yin, Z. Wu, and W. Curtin, First-principles calculations of stacking fault energies in mg-y, mg-al and mg-zn alloys and implications for $(c+a)$ activity, *Acta Mater.* **136**, 249 (2017).
- [33] Z. Wu, R. Ahmad, B. Yin, S. Sandlöbes, and W. A. Curtin, Mechanistic origin and prediction of enhanced ductility in magnesium alloys, *Science* **359**, 447 (2018).
- [34] R. Ahmad, B. Yin, Z. Wu, and W. Curtin, Designing high ductility in magnesium alloys, *Acta Mater.* **172**, 161 (2019).
- [35] T. Nogaret, W. Curtin, J. Yasi, L. Hector, and D. Trinkle, Atomistic study of edge and screw $(c + a)$ dislocations in magnesium, *Acta Mater.* **58**, 4332 (2010).

- [36] R. Ahmad, S. Groh, M. Ghazisaeidi, and W. A. Curtin, Modified embedded-atom method interatomic potential for Mg–Y alloys, *Modell. Simul. Mater. Sci. Eng.* **26**, 065010 (2018).
- [37] A. Singraber, J. Behler, and C. Dellago, Library-based lammmps implementation of high-dimensional neural network potentials, *J. Chem. Theory Comput.* **15**, 1827 (2019).
- [38] J. Behler, Constructing high-dimensional neural network potentials: A tutorial review, *Int. J. Quant. Chem.* **115**, 1032 (2015).
- [39] A. Singraber, Designing and training neural network potentials for molecular dynamics simulations: Application to water and copper sulfide, Ph.D. thesis, Universität Wien, 2018.
- [40] 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).
- [41] S. Lorenz, A. Groß, and M. Scheffler, Representing high-dimensional potential-energy surfaces for reactions at surfaces by neural networks, *Chem. Phys. Lett.* **395**, 210 (2004).
- [42] J. B. Witkoskie and D. J. Doren, Neural network models of potential energy surfaces: Prototypical examples, *J. Chem. Theory Comput.* **1**, 14 (2005).
- [43] B. Yin, M. Stricker, and W. Curtin, Pure magnesium DFT calculations for interatomic potential fitting, materials cloud archive, *MaterialsCloud* (2020), doi: 10.24435/materialscloud:2020.0046/v1.
- [44] G. Kresse and J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set, *Phys. Rev. B* **54**, 11169 (1996).
- [45] G. Kresse and D. Joubert, From ultrasoft pseudopotentials to the projector augmented-wave method, *Phys. Rev. B* **59**, 1758 (1999).
- [46] J. P. Perdew, K. Burke, and M. Ernzerhof, Generalized Gradient Approximation Made Simple, *Phys. Rev. Lett.* **77**, 3865 (1996).
- [47] P. E. Blöchl, Projector augmented-wave method, *Phys. Rev. B* **50**, 17953 (1994).
- [48] H. J. Monkhorst and J. D. Pack, Special points for brillouin-zone integrations, *Phys. Rev. B* **13**, 5188 (1976).
- [49] M. Methfessel and A. T. Paxton, High-precision sampling for brillouin-zone integration in metals, *Phys. Rev. B* **40**, 3616 (1989).
- [50] C. Barrett and T. Massalski, *Structure of Metals*, *International Series on Materials Science and Technology*, 3rd ed. (Pergamon, London, 1980).
- [51] S. Plimpton, Fast parallel algorithms for short-range molecular dynamics, *J. Comput. Phys.* **117**, 1 (1995).
- [52] A. Stukowski, Visualization and analysis of atomistic simulation data with OVITO—The open visualization tool, *Modell. Simul. Mater. Sci. Eng.* **18**, 015012 (2009).
- [53] Z. Wu and W. A. Curtin, The origins of high hardening and low ductility in magnesium, *Nature* **526**, 62 (2015).
- [54] B. Yin, Z. Wu, and W. Curtin, Comprehensive first-principles study of stable stacking faults in hcp metals, *Acta Mater.* **123**, 223 (2017).
- [55] G. Walker and M. Marezio, Lattice parameters and zone overlap in solid solutions of lead in magnesium, *Acta Metall.* **7**, 769 (1959).
- [56] E. Wachowicz and A. Kiejna, Bulk and surface properties of hexagonal-close-packed be and mg, *J. Phys.: Condens. Matter* **13**, 10767 (2001).
- [57] G. Simmons and H. Wang, *Single Crystal Elastic Constants and Calculated Aggregate Properties: A Handbook*, 2nd ed. (MIT Press, Cambridge, MA, 1971).
- [58] J. Han, X. Su, Z.-H. Jin, and Y. Zhu, Basal-plane stacking-fault energies of mg: A first-principles study of li- and al-alloying effects, *Scr. Mater.* **64**, 693 (2011).
- [59] Z. Pei, X. Zhang, T. Hickel, M. Friák, S. Sandlöbes, B. Dutta, and J. Neugebauer, Atomic structures of twin boundaries in hexagonal close-packed metallic crystals with particular focus on mg, *npj Comput. Mater.* **3**, 6 (2017).
- [60] P. Tzanetakis, J. Hillairet, and G. Revel, The formation energy of vacancies in aluminium and magnesium, *Phys. Status Solidi B* **75**, 433 (1976).
- [61] 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).
- [62] M. S. Daw and M. I. Baskes, Semiempirical, Quantum Mechanical Calculation of Hydrogen Embrittlement in Metals, *Phys. Rev. Lett.* **50**, 1285 (1983).
- [63] M. S. Daw and M. I. Baskes, Embedded-atom method: Derivation and application to impurities, surfaces, and other defects in metals, *Phys. Rev. B* **29**, 6443 (1984).
- [64] P. M. Morse, Diatomic molecules according to the wave mechanics. ii. Vibrational levels, *Phys. Rev.* **34**, 57 (1929).
- [65] E. J. Ragasa, C. J. O'Brien, R. G. Hennig, S. M. Foiles, and S. R. Phillpot, Multi-objective optimization of interatomic potentials with application to MgO, *Modell. Simul. Mater. Sci. Eng.* **27**, 074007 (2019).
- [66] I. Shin and E. A. Carter, Simulations of dislocation mobility in magnesium from first principles, *Int. J. Plast.* **60**, 58 (2014).
- [67] AtomMan, <https://github.com/usnistgov/atomman> (2020).
- [68] C. Hartley and Y. Mishin, Characterization and visualization of the lattice misfit associated with dislocation cores, *Acta Mater.* **53**, 1313 (2005).
- [69] V. Vitek, R. C. Perrin, and D. K. Bowen, The core structure of $\frac{1}{2}(111)$ screw dislocations in b.c.c. crystals, *Philos. Mag.* **21**, 1049 (1970).
- [70] D. Faken and H. Jónsson, Systematic analysis of local atomic structure combined with 3d computer graphics, *Comput. Mater. Sci.* **2**, 279 (1994).
- [71] A. Stukowski and K. Albe, Extracting dislocations and non-dislocation crystal defects from atomistic simulation data, *Modell. Simul. Mater. Sci. Eng.* **18**, 085001 (2010).
- [72] J. A. Yasi, T. Nogaret, D. R. Trinkle, Y. Qi, L. G. Hector, and W. A. Curtin, Basal and prism dislocation cores in magnesium: Comparison of first-principles and embedded-atom-potential methods predictions, *Modell. Simul. Mater. Sci. Eng.* **17**, 055012 (2009).
- [73] M. Ghazisaeidi, L. Hector, and W. Curtin, First-principles core structures of $\langle c+a \rangle$ edge and screw dislocations in mg, *Scr. Mater.* **75**, 42 (2014).
- [74] R. Ahmad, Z. Wu, S. Groh, and W. Curtin, Pyramidal ii to basal transformation of $\langle c+a \rangle$ edge dislocations in mg-y alloys, *Scr. Mater.* **155**, 114 (2018).
- [75] M. Itakura, H. Kaburaki, M. Yamaguchi, and T. Tsuru, Novel Cross-Slip Mechanism of Pyramidal Screw Dislocations in Magnesium, *Phys. Rev. Lett.* **116**, 225501 (2016).
- [76] M. Ghazisaeidi, L. Hector, and W. Curtin, Solute strengthening of twinning dislocations in mg alloys, *Acta Mater.* **80**, 278 (2014).

- [77] M. I. Baskes, Modified embedded-atom potentials for cubic materials and impurities, *Phys. Rev. B* **46**, 2727 (1992).
- [78] F. H. Stillinger and T. A. Weber, Computer simulation of local order in condensed phases of silicon, *Phys. Rev. B* **31**, 5262 (1985).
- [79] G. Lu, E. B. Tadmor, and E. Kaxiras, From electrons to finite elements: A concurrent multiscale approach for metals, *Phys. Rev. B* **73**, 024108 (2006).
- [80] D. H. Warner, W. A. Curtin, and S. Qu, Rate dependence of crack-tip processes predicts twinning trends in f.c.c. metals, *Nat. Mater.* **6**, 876 (2007).
- [81] D. Warner and W. Curtin, Origins and implications of temperature-dependent activation energy barriers for dislocation nucleation in face-centered cubic metals, *Acta Mater.* **57**, 4267 (2009).
- [82] X. Zhang, G. Lu, and W. A. Curtin, Multiscale quantum/atomistic coupling using constrained density functional theory, *Phys. Rev. B* **87**, 054113 (2013).
- [83] V. Yamakov, D. Warner, R. Zamora, E. Saether, W. Curtin, and E. Glaesgen, Investigation of crack tip dislocation emission in aluminum using multiscale molecular dynamics simulation and continuum modeling, *J. Mech. Phys. Solids* **65**, 35 (2014).
- [84] A. A. Griffith, Vi. the phenomena of rupture and flow in solids, *Philos. Trans. R. Soc. Lond. A* **221**, 163 (1921).
- [85] T. T. C. Ting, *Anisotropic Elasticity: Theory and Applications* (Oxford University Press, Oxford, 1996).
- [86] J. R. Rice, Dislocation nucleation from a crack tip: An analysis based on the peierls concept, *J. Mech. Phys. Solids* **40**, 239 (1992).
- [87] P. Andric and W. Curtin, New theory for mode i crack-tip dislocation emission, *J. Mech. Phys. Solids* **106**, 315 (2017).
- [88] P. Andric and W. Curtin, Atomistic modeling of fracture, *Modell. Simul. Mater. Sci. Eng.* **27**, 013001 (2018).
- [89] J. Song, W. Curtin, T. Bhandakkar, and H. Gao, Dislocation shielding and crack tip decohesion at the atomic scale, *Acta Mater.* **58**, 5933 (2010).
- [90] E. Mak and W. A. Curtin, Intrinsic fracture behavior of mg-y alloys, *Modell. Simul. Mater. Sci. Eng.* **28**, 035012 (2020).
- [91] J.-J. Tang, X.-B. Yang, L. OuYang, M. Zhu, and Y.-J. Zhao, A systematic first-principles study of surface energies, surface relaxation and friedel oscillation of magnesium surfaces, *J. Phys. D* **47**, 115305 (2014).