PHYSICAL REVIEW B **102**, 144107 (2020)

Modeling of metal nanoparticles: Development of neural-network interatomic potential inspired by features of the modified embedded-atom method

Feifeng Wu,[1] Hang Min,[1] Yanwei Wen,[1] Rong Chen,[2] Yunkun Zhao,[3] Mike Ford,[4,*] and Bin Shan[1,†]
[1]*State Key Laboratory of Material Processing and Die and Mould Technology and School of Materials Science and Engineering, Huazhong University of Science and Technology, Wuhan, Hubei 430074, People's Republic of China*
[2]*State Key Laboratory of Digital Manufacturing Equipment and Technology, School of Mechanical Science and Engineering, Huazhong University of Science and Technology, Wuhan 430074, Hubei, People's Republic of China*
[3]*State Key Laboratory of Advanced Technologies for Comprehensive Utilization of Platinum Metal, Kunming Institute of Precious Metals, Kunming 650106, Yunnan, People's Republic of China*
[4]*School of Mathematical and Physical Sciences, University of Technology Sydney, Ultimo, Sydney, New South Wales 2007, Australia*

(Received 25 February 2020; revised 5 August 2020; accepted 29 September 2020; published 15 October 2020)

Interatomic potential plays a key role in ensuring the accuracy and reliability of molecular-dynamics simulation results. While most empirical potentials are benchmarked against a set of carefully chosen bulk material properties, recent advances in machine learning have seen the emergence of neural-network-based mathematical potentials capable of describing highly complex potential energy surfaces for a variety of systems. We report here the development of a neural-network interatomic potential (NNIP) with modified embedded-atom method background density as fingerprint functions, which could accurately model the energetics of metallic nanoparticles and clusters (Cu as a representative example) widely used in catalysis. To appropriately account for the diverse chemical environments encountered in nanoparticles/nanoclusters, an extensive set of atomic configurations (totaling 18 084) were calculated using density-functional-theory (DFT) at the Perdew-Burke-Ernzerhof level. In addition to standard bulk properties such as cohesive energies and elastic constants, the sampled configurations also include a substantial number of differently oriented crystal facets and differently sized nanoparticles and nanoclusters, greatly expanding the value range of NNIP features that was otherwise quite limited. The complex energy potential surface of Cu can be faithfully reproduced, with an average error of 0.011 eV/at for energy states within 3 eV of the ground state. As an illustration, the developed NNIP is used to simulate the molecular dynamics of copper nanoparticles, and good agreement is achieved between DFT and the NNIP.

DOI: 10.1103/PhysRevB.102.144107

I. INTRODUCTION

Computational materials design through advanced first-principles or molecular-dynamics (MD) simulations has contributed greatly to materials discovery, innovation, manufacturing, and commercialization in recent years. We are now entering an era in which the paradigm of material development is shifting from trial-and-error experimentation to computer-aided rational design [1–5]. It has become common practice to investigate the physical and chemical behavior of materials via computation, since some properties are difficult or costly to obtain directly from experiments [6,7]. This paradigm shift also makes possible rapid screening of virtual materials candidates with desired physical and chemical properties before they are synthesized, opening up new avenues of research [8–11]. The first-principles method [12] has been one of the core tools for describing interactions and dynamic processes of materials on an atomic scale. However, such a method relies on solving the electronic wave function to determine the force exerted on each atom to evolve the system, the computational cost of which is prohibitive for large systems. MD [13,14] has been one of the most effective approaches for larger systems, and it is widely used in studying atomic processes such as defect formation, crack propagation, and nanoparticle segregation [15–18]. The accuracy and reliability of MD hinges critically on the interatomic potential's capability to faithfully describe potential energy surfaces and to conform to the actual laws of the physical model. Parameters involved in potential functionals are in general fitted to experimental data or higher-level density functional theory (DFT) calculations.

Among the different types of interatomic potentials developed over the past few decades [19–24], the embedded-atom method (EAM) proposed by Daw *et al.* stands out as a successful model for describing metallic systems [24]. Researchers have since developed different flavors of analytical EAM potentials, such as the Johnson EAM the Cai-Ye EAM [23,25,26]. One of the deficiencies of the EAM is the lack of angular dependence in background density, making it challenging to correctly predict the ground state and surface properties of non-close-packed systems [25]. The modified embedded-atom method (MEAM) [25,26] was proposed as a direct extension to the EAM by incorporating angular terms to the background density distribution, and it has been

---

[*] mike.ford@uts.edu.au

[†] bshan@mail.hust.edu.cn

successfully applied to modeling metal surfaces and impurities [26–28]. Recent progress toward a more accurate MEAM includes the formulation of the second nearest-neighbor MEAM (2NN-MEAM) [29,30]. Our group also developed the lattice inversion MEAM (LI-MEAM) [31], which uses the Chen-Mobius inversion method to take into account the contribution of farther neighbors. Nonetheless, the analytical MEAM employs preset embedding and pairwise functional forms based on physical intuitions that are parametrized mostly against bulk structure properties [23,31]. This makes it less transferrable in predicting the energies of small nanoparticles, whose local atomic environment is considerably different from that of the bulk counterpart [32–35]. By using mathematical potentials based on machine learning (ML) [34,36–39], extended datasets can be used to construct complex potential energy surfaces with reasonable accuracy and transferability. It is desirable to combine the flexibility of mathematical potential with MEAM-derived input features that could reflect the physics based upon which the total energy is assigned [40].

We report here the development of a neural-network interatomic potential (NNIP) that uses MEAM background density as fingerprint features, which preserves vector invariance under spatial translation and rotation and has proven to be capable of modeling diverse local atomic environments. Copper, being a widely used material in catalysis and other fields, has been chosen as an illustrative example to demonstrate the flexibility and transferability of the developed NNIP. Our DFT training set contains a total of 18 084 structures with 471 300 unique atomic environments that can be roughly categorized into bulk, surface, nanoparticle, and cluster configurations. The overall energy deviation across this diverse dataset is less than 0.02 eV/at, well within the DFT accuracy limit. In particular, the NNIP shows good agreement with DFT data for nanoparticle and cluster configurations, which are generally not well treated in the original MEAM formalism. Our work extends the applicability of MEAM descriptors to nanoparticles and clusters with improved accuracy and reliability.

II. METHOD

A. Fingerprint construction from MEAM background density

In the original MEAM formalism, the total energy of the system E equals the sum of individual atomic energies E_i, further divided into an embedding part and a pairwise interaction, which is given by

$$E = \sum_i E_i, \tag{1}$$

$$E_i = F_i(\rho_i) + \frac{1}{2}\sum_{j\neq i} V(R_{ij}). \tag{2}$$

The embedding function F_i is the energy to embed atom i into the background density at site i. R_{ij} denotes the distance vector between atoms j and i. ρ_i and $V(r_{ij})$ denote the superposition of the background density for all neighboring atoms j at site i and a pair interaction separated by a distance R_{ij}.

In the original MEAM formalism, Baskes designed the expressions for the partial background densities $\rho_i^{(0)}$(radial term) and $\rho_i^{(1)}$, $\rho_i^{(2)}$, $\rho_i^{(3)}$(angular terms) to reflect the local atomic environments [26]. They are given by the following expression:

$$\rho_i^{(0)} = \sum_j \rho_j^{\alpha(0)}(R_{ij}), \tag{3}$$

$$\left(\rho_i^{(1)}\right)^2 = \sum_\alpha \left[\sum_j x_{ij}^\alpha \rho_j^{a(1)}(R_{ij})\right]^2, \tag{4}$$

$$\left(\rho_i^{(2)}\right)^2 = \sum_{\alpha,\beta} \left[\sum_j x_{ij}^\alpha x_{ij}^\beta \rho_j^{a(2)}(R_{ij})\right]^2 - \frac{1}{3}\left[\sum_j \rho_j^{a(2)}(R_{ij})\right], \tag{5}$$

$$\left(\rho_i^{(3)}\right)^2 = \sum_{\alpha,\beta,\gamma} \left[\sum_j x_{ij}^\alpha x_{ij}^\beta x_{ij}^\gamma \rho_j^{a(3)}(R_{ij})\right]^2, \tag{6}$$

where $x_{ij}^\alpha = R_{ij}^\alpha / R_{ij}$, and ${R_{ij}}^\alpha$ denotes the component of R_{ij} on x, y, or z. $\rho_j^{\alpha(0)}(R_{ij})$ represents the atomic electron density of the atom j at the distance R_{ij} from the site i. The different partial background density components are constructed to relate to different local atomic environments.

Due to its invariant nature with respect to spatial translation and rotation, MEAM-based background densities serve as good candidates for symmetry functions to NNIP. We adopted here a slightly modified version of partial background density components G_s, G_p, G_d, and G_f as features to describe the local chemical environment:

$$G_s^i = \left[\sum_j e^{-\eta {R_{ij}}^2} \cdot f_c(R_{ij})\right]^2, \tag{7}$$

$$G_p^i = \sum_\alpha \left[\sum_j \frac{R_{ij}^\alpha}{R_{ij}} e^{-\eta {R_{ij}}^2} \cdot f_c(R_{ij})\right]^2, \tag{8}$$

$$G_d^i = \sum_{\alpha,\beta} \left[\sum_j \frac{R_{ij}^\alpha R_{ij}^\beta}{R_{ij}^2} e^{-\eta {R_{ij}}^2} \cdot f_c(R_{ij})\right]^2, \tag{9}$$

$$G_f^i = \sum_{\alpha,\beta,\gamma} \left[\sum_j \frac{R_{ij}^\alpha R_{ij}^\beta R_{ij}^\gamma}{R_{ij}^3} e^{-\eta {R_{ij}}^2} \cdot f_c(R_{ij})\right]^2, \tag{10}$$

$$f_c(R_{ij}) = \begin{cases} \frac{1}{2}\left[\cos\left(\frac{\pi R_{ij}}{R_c}\right) + 1\right], & R_{ij} < R_c, \\ 0 & \text{otherwise}, \end{cases} \tag{11}$$

where R_{ij} denotes the vector of atom i pointing to atom j, and α, β, and γ denote the vector of R_{ij} along the axis of x, y, and z, respectively. R_c and η denote the cutoff radius and decay rate in units of Å^{-1}. $f_c(R_{ij})$ is a cutoff function that smoothly truncates the potential beyond R_c. Note that the angular shielding function S_{ij} in the original MEAM has been removed as it was primarily introduced in the MEAM formalism to simplify the deduction of the analytical pair function. Based on Eqs. (3)–(6), a smooth cutoff term $f_c(R_{ij})$ [Eq. (11)] is added to ensure the continuity of the function and the efficiency of the transformation; a decay term $e^{-\eta {R_{ij}}^2}$ is used to differentiate contributions of neighboring atoms to local environments. The computational complexity associated with the computation of the MEAM fingerprint function for the n-atom system is $O(n^2)$, compared to the $O(n^3)$ complexity in calculating the angular term symmetry

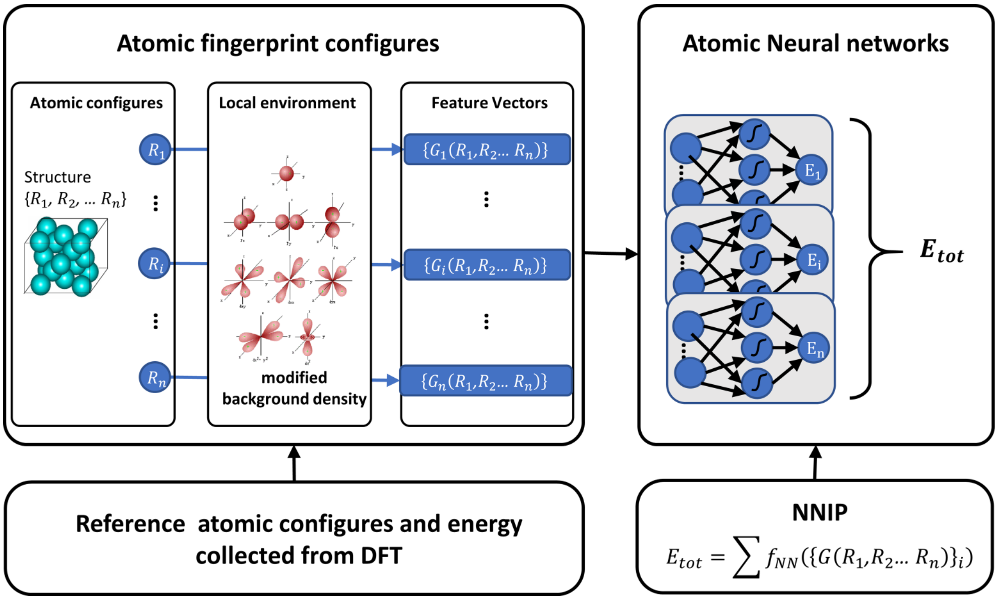

FIG. 1. The schematic of the construction of the NNIP. Using DFT to obtain an accurate dataset, the symmetry functions convert the atomic coordinates into modified background density as the inputs of NN, and the NN directly constructs the mapping relationship between the structure and the energy as the main framework.

function for high-dimensional neural network potentials (HDNNPs) [34].

B. Neural-network interatomic potential formalism

Regarding the neural-network structure, a four-layer NN is constructed that enables us to take into account various types of structural and energetic information with different local chemical environments. ML frameworks Tensorflow [41] is used to train the NNIP. The feedforward F_{NN} [42] is used to represent each atomic energy of the system, which can be written as follows:

$$E_i = F_{\mathrm{NN}}\left(\left\{G_s^i, G_p^i, G_d^i, G_f^i\right\}\right), \tag{12}$$

where F_{NN} denotes the neural network with each atom corresponding to an atomic NN (Fig. 1), with input features being G_s^i, G_p^i, G_d^i, and G_f^i.The output is the atomic energy E_i. All the neurons are connected to the network with appropriate weighting factors. A hyperbolic tangent (tanh) function is selected as the activation function, and optimization of the NN weights leads to a minimization of the target loss function.

For the model fitting, a loss function $L(y_i, f(x_i))$ is introduced to evaluate the difference between the predicted total energy $f(x_i)$ and actual energy y_i. Among the many types of loss functions, Huber loss [Eqs. (13)] is chosen for its continuous gradient and good robustness toward off-group points. It is given by

$$L(y_i, f(x_i)) = \begin{cases} \frac{1}{2}[y_i - f(x_i)]^2, & |y_i - f(x_i)| < \delta, \\ \delta|y_i - f(x_i)| - \frac{1}{2}\delta^2 & \text{otherwise,} \end{cases} \tag{13}$$

where y_i denotes the system energy calculated by DFT, and $f(x_i)$ denotes the system energy predicted by the neural network. N is the number of samples, and λ and δ are weighting factors. To prevent the model from overfitting and to improve the generalization ability of the model, a regularization term is added to the loss function. The final regularized loss function is given by

$$\text{loss} = \min \frac{1}{N}\sum_{i=1}^{N} L(y_i, f(x_i)) + \lambda J(f), \tag{14}$$

where $J(f)$ is an additional penalty term to prevent overfitting, including L1 (Lasso regression) [43] and L2 (ridge regression) [44]. λ is a weighting factor that balances between underfitting (large λ) and overfitting (small λ). With appropriate choice of λ, one can train a model to have both good accuracy and generalization capability. Since a large number of deformation configurations has been included in the training set, implicitly taking the force into account [45,46], the force is not explicitly included due to efficiency considerations. The batch gradient descent (BGD) optimization algorithm is used to search for the minimum value of loss to get optimized NNIP parameters.

C. DFT training dataset for a Cu element

Similar to any other NN applications, the completeness of the dataset affects directly the reliability of the NNIP potential. Types of structures included in the training set determine the applicable range of the NNIP potential, as only configurations that have similar counterparts in the training set can be predicted with high confidence. As one of the major motivations for the development of NNIP potential is to enable the modeling of metallic clusters and nanoparticles, our datasets purposely include the potential energy surfaces of a large number of regular and deformed Cu structures. Self-consistent DFT calculations are performed with the Perdew-Burke-Ernzerhof exchange-correlation functional as implemented in the Vienna Ab Initio Simulation (VASP) package [47]. The Kohn-Sham electronic wave function is

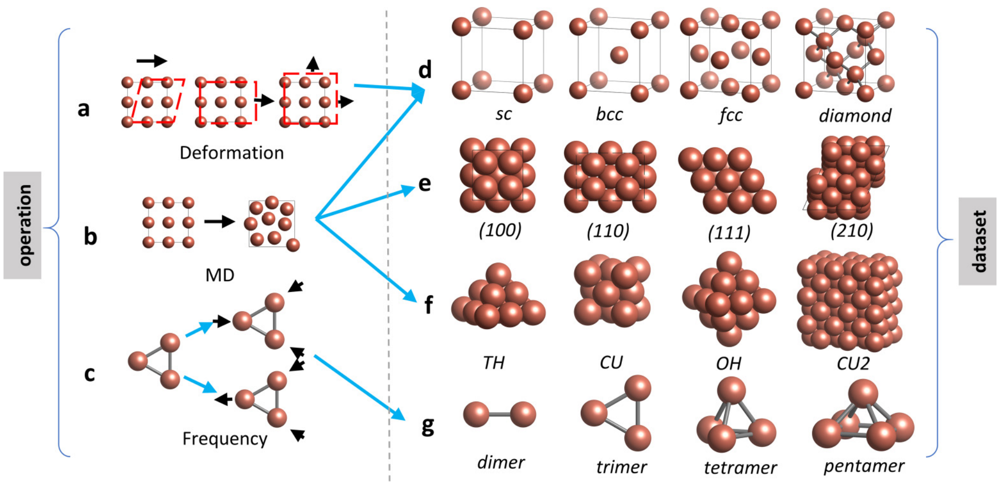

FIG. 2. A total of 15 427 structures are included in the dataset. They can be roughly divided into four categories: bulk, slab, nanoparticle, and cluster. The initial atomic configurations are established, and the dataset is extended by using deformation, MD, and interpolated frequency vibration modes.

expended by plane waves with an energy cutoff of 400 eV, and the total energy converges to less than 10^{-3} meV. Cu is used as an illustrative example to demonstrate the process of dataset construction and the applicability of NNIP as a whole (Fig. 2).

A total of 18 084 geometries included in the training set can be roughly divided into four categories: bulk, surface, nanoparticles, and clusters. For bulk structures, in addition to the ground state of fcc structure, we have included other symmetrical structures with different packing forms, such as bcc, sc, hcp, and diamond. A 9*9*9 *k*-point mesh was used to sample the first Brillouin zone of the bulk copper. A series of configurations are sampled along the equation of state for Cu under different stress/strain conditions, including the isotropic expansion, pure shear, and uniaxial expansion. Vacancy formation energy is also incorporated into the dataset as it dictates the behavior of materials structural evolution in phase transformation, crystal growth, and atomic diffusion. In addition, MD at relatively high temperatures (below melting point) is conducted to generate deformed bulk configurations, supplementing the aforementioned bulk properties data with more distorted structures.

Surface properties are also of great importance as they account for chemical processes such as catalysis and corrosion. To model the surface properties, we have used a slab geometry with six layers. $p(1 \times 1)$, $p(2 \times 1)$, and $p(2 \times 2)$ primitive surface cells for the simulation of crystal planes (100), (110), and (111), respectively. (210) and (211) planes are also considered as they possess distinctive coordinately unsaturated sites that would be very similar to edge sites in nanoparticles. The surface energy calculations are done with the bottom three layers fixed at their crystallographic positions, while atoms in the upper three layers are allowed to fully relax. A 9*9*1 *k*-point mesh is used with a vacuum thickness of 10 Å to prevent slab image interactions along the z axis. This part of the datasets is also supplemented with structures from MD runs equilibrated at appropriate volumes and temperatures.

On an even smaller scale, atoms in nanoparticles show even richer local chemical environments. Nanoparticles have distinct terrace, edge, and vertex sites. One of the essential properties of a nanoparticle that determines its stability with respect to other phases is the size-dependent chemical potential, which can be defined as

$$E_{\text{che}} = \frac{1}{n}(E_{\text{tot}} - nE_{\text{bulk}}), \tag{15}$$

where E_{tot} is total energy of the structure, E_{bulk} is the cohesive energy, and n denotes the number of atoms in the structure. In general, the chemical potential would become more positive going toward smaller nanoparticles, reflecting the higher chemical potential of smaller nanoparticles.

The final part of the dataset is small-sized nanoclusters consisting of only a few atoms. The defined MEAM input features for nanoclusters would deviate most from bulk values. By including these geometries, we significantly extended the range of feature values, especially in the low background density part and the higher-order component (p, d, and f) input vectors, as they are mostly small numbers due to symmetry restrictions in bulk cases. To ensure that the nanoclusters are locally stable geometries, we chose three dominant vibrational modes in each cluster, and we calculated the energy curve with respect to distortions along these vibrational modes. This way, the NNIP would predict correctly the local optimum even in geometries that deviate significantly from bulk environments.

Through the dataset construction strategy, we constructed a database of initial configurations and expanded the dataset by means of deformation, MD, frequency mode, etc. Detailed information is shown in Fig. 2 and Table I.

III. RESULTS AND DISCUSSION

A. Feature space and correlation

G_s, G_p, G_d, and G_f as defined in Eqs. (7)–(10) describe the partial background density components of different symmetries from neighboring atoms in the system. This forms a

TABLE I. The label and number corresponding to different structures in the dataset.

Label	Detail	Number
bulk	deformation information of bulk structure	3284
bulk$_{\mathrm{md}}$	disordered bulk structure	5000
NP	typical shapes of large size cluster	1200
NP$_{\mathrm{md}}$	disordered structure of large size cluster	5000
nano	small size cluster	2400
slab	slab structure contains the information of surface	1200

unique set of features for each atom in different local chemical environments. Overall, the characteristic quantization mark indicates the type of structure we are modeling [Fig. 3(a)]. For example, bulk configurations have a larger number of nearest neighbors, so G_s in bulk structure is typically larger than that in the surface, nanoparticles (NPs), and nanoclusters. G_p and G_f are relatively small as the nearest neighbors around a bulk atom tend to form a symmetrical spatial distribution, and the magnitude of G_d is similar to that of G_s because of the peaks at 180° [Fig. 4(c)]. In many cases, such as ideal fcc/bcc/sc structures, all other components except G_s and G_d are zero due to cubic symmetry. Under distorted conditions (shear and uniaxial distortion, and configurations obtained by MD), there is a small increment of G_p and G_f components, but overall they remain relatively small. One exception to this rule is the bulk diamond structure, where there is a large G_f component due to the existence of tetrahedral interstitial atoms. For slab configurations, the atoms in the surface layer (slab) are situated in an asymmetric local environment, where the G_p and G_f contribution from the vacuum side is completely missing, leading to a large G_p and G_f value. For atoms located on the surface of the NPs, they form fewer bonds with neighboring atoms than bulk, so their characteristics are similar to those of the surface atoms. The final scenario is nanoclusters, where we can see that G_s decreases considerably due to the smaller number of atoms around. From the above discussion, it can be seen that features in the NNIP establish a unique mapping between energy and local atomic environment, which helps to accurately capture the nonlinear scaling of atomic energies with respect to local coordination numbers [46].

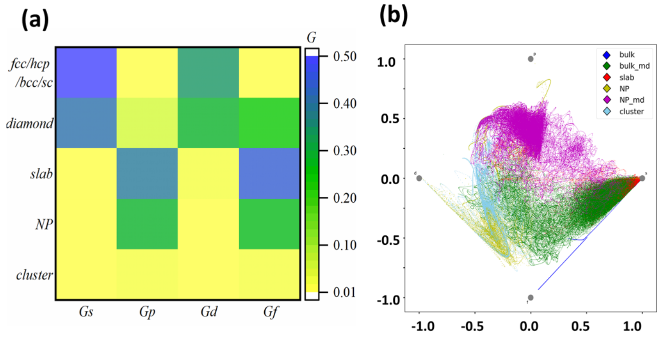

FIG. 3. (a) The distribution of features for different structures. (b) The distribution of each sample in the dataset is in G-space and shows the diversity and uniformity of the dataset release.

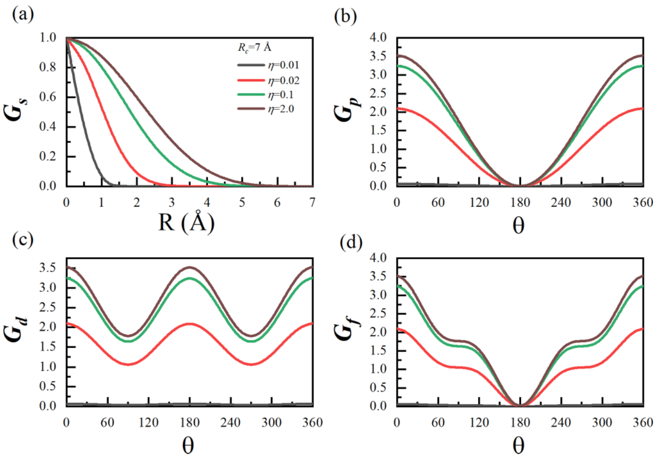

FIG. 4. (a) The change of G_s with distance R; (b)–(d) the change of G_p, G_d, and G_f with angle θ.

One requirement of a good dataset is to have diversity and uniformity of data points across a majority of the feature space. In our high-dimensional feature space, each data point reflects the distribution of particular geometries transformed into local chemical environment features. Here, we use the RadViz chart to plot the feature vector collection [Fig. 3(b)] that summarizes all the datasets in one concise graph. Along the four corners of the unit circle, each point represents an attribute (G_s, G_p, G_d, and G_f). The position of the data point on the plane is determined by an equilibrium position assuming simultaneous spring connection to each corner. Samples of different classes are represented by different colors. We can see that bulk properties only expand a very limited feature space, mostly centered around s (since G_p and G_f are small) and selected lines between s and f (diamond structures). Generating distorted bulk configurations with MD expands the feature space by occupying an enlarged portion of the space with G_s, G_d, and G_f, yet G_p is mostly negligible (indicated by the green dots). The reason for this is that neighboring atoms tend to distribute uniformly on both sides of an atom, canceling out the G_p component. The upper RadViz space is mostly filled up by slab and nanoparticle configurations, where the existence of a large portion of surface atoms leads to a substantial G_p component due to asymmetricity. Most of the nanocluster configurations take up the feature space on the left side, indicating a lack of symmetry and a smaller magnitude of G_s. Overall, by sampling over the four categories of structures, we have obtained a G-space that is fairly well taken up as compared to the limited G-space spanned by bulk configurations only. The dataset (Table I) contains the necessary structural information, and the corresponding energy information is obtained through standard DFT calculations.

B. The effect of parameters on accuracy

The transformation from Cartesian coordinates to features has been defined above to quantify the local atomic environment [Eqs. (7)–(10)], and the selection of the parameters R_c and η relates to details of the NNIP features. R_c determines the range of the interatomic interaction, and η dictates the rate of background density decay that can effectively distinguish

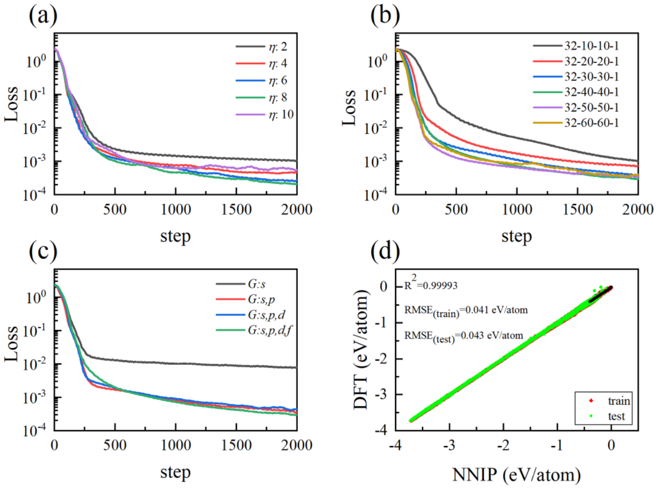

FIG. 5. (a) The effect of η parameter on the loss function. (b) Comparison of fitting errors for several neural-network (NN) architectures. (c) G_s, G_p, G_d, and G_f contribution for the final result. (d) Comparison of the error between the NNIP potential and the DFT; the red dot represents the training set, and the green square represents the test set.

TABLE II. Symmetry function parameter values list in NNIP.

The parameter of symmetry functions G_s, G_p, G_d, and G_f								
R_c(Å)	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0
η(Å^{-1})	0.01	0.02	0.05	0.1	0.2	0.5	1.0	1.5

contributions from different neighbors. Each set of parameters with different η reflects a different aspect of the local atomic structure, and when combined they provide a comprehensive picture of the local chemical environment. Each of G_s, G_p, G_d, and G_f contains partial information about the local environment, similar to the Fourier series expansion. Different parameters of G_s, G_p, G_d, and G_f are combined into a unique feature vector to represent the local atomic environment. G_s explicitly contains the information about the radial R_{ij} [Fig. 4(a)], and η determines the width of the nonzero region with different values representing different neighbors. G_p, G_d, and G_f not only contain the radial term R_{ij}, but they also indirectly give the local atomic environment information containing angular terms in R_{ij}^{α}/R_{ij}. To understand the variation of features (G_p, G_d, and G_f) with θ, the three-atom system is discussed as an illustrative example. η plays the same role as in G_s to distinguish the atomic environment information of different neighbors. Since the increase in the number of series in the angular direction R_{ij}^{α}/R_{ij} leads to the differences in contributions at different angles in Figs. 4(b)–4(d), the angular information on a part of the local atomic environment can be independently provided, and it can also be seen that G_p, G_d, and G_f are complementary to each other.

A few R_c and η combinations are selected, and the root-mean-square error (RMSE) between the predicted NNIP energies and DFT calculations is evaluated. The range of $[0, r_c]$ is divided into N regions corresponding to $\{\eta_i\}$, and the results are shown in Fig. 5(a). As can be seen from Fig. 5(a), when the number of η is larger than 8, the accuracy does not improve much. Thus, we chose $R_c = 7$ Å, and the eight η parameters are listed in Table II. G_s, G_p, G_d, and G_f partial electron densities are added progressively to test the contribution of the individual component to the final result [Fig. 5(c)]. It can be seen that G_s alone could reduce the average energy error to the order of 0.1 eV/at. Utilizing G_s only is conceptually equivalent to the EAM. The addition of G_p and G_d further reduces the error to 0.01 eV/at, primarily improving the description of non-close-packed systems and nonbulk systems. The inclusion of G_f improves redundancy, making it more flexible to describe lattice structures such as diamond.

For the optimization of NNIP parameters, we tested varying numbers of neurons in the two middle hidden layers. A total of five different neural-network structures are evaluated and compared [Fig. 5(b)]. It can be seen that when the number of neurons in the middle layers is larger than 30, it has a negligible influence on either accuracy or convergence. Taking a reasonable balance between flexibility and computational cost, we used a 32-40-40-1 NN. After determining the neural-network structure, the training details are straightforward: the weight of Adam is optimized using the minibatch method algorithm. The required potential function model is obtained at the end of the optimization.

Figure 5(d) shows the correlation between DFT and NNIP data for all the geometries in our training dataset. The point distribution is on the 45° line, with a standard deviation of 0.041 and 0.043 eV for the training set and the test set, respectively. The small and comparable MSD value for the two sets indicates that the fitting is not overfitted or biased. The overall correlation coefficient is 0.999 93 across the energy range from −3.8 to 0 eV.

C. Application of NNIP to copper

1. *Properties of Cu bulk*

We first discuss the reliability of the developed NNIP potential in describing the properties of bulk Cu. Binding energy curves obtained by the NNIP potential, DFT, and the original MEAM potential are compared in Fig. 6. These structures represent a range of different coordination environments. For the ground state of fcc Cu, both MEAM and NNIP agree well with DFT data. The average error between the NNIP and DFT is 0.006 eV. The discrepancy at lattice constants larger than 5.4 Å for MEAM could be attributed to its shorter cutoff distance The advantage of NNIP primarily lies in the description of bulk copper other than fcc structure. As can be seen from Fig. 6, the lattice constant of MEAM deviates substantially from DFT for non-close-packed structures such as bcc and sc, with an underestimation as high as 3.4%, while NNIP very well reproduces both the lattice constant and the cohesive energy of non-close-packed Cu structures. The mean absolute error (MAE) between the NNIP and DFT is 0.044 eV/at and the root-mean-square error (RMSE) is 0.054 eV/at. For comparison, MAE between MEAM and DFT is 0.297 eV/at and RMSE is 0.288 eV/at. In Table III, we quantitatively compare the lattice constants and equilibrium structure energy differences of NNIP and MEAM. Overall, MEAM gives comparable accuracy to NNIP for fcc ground states, but

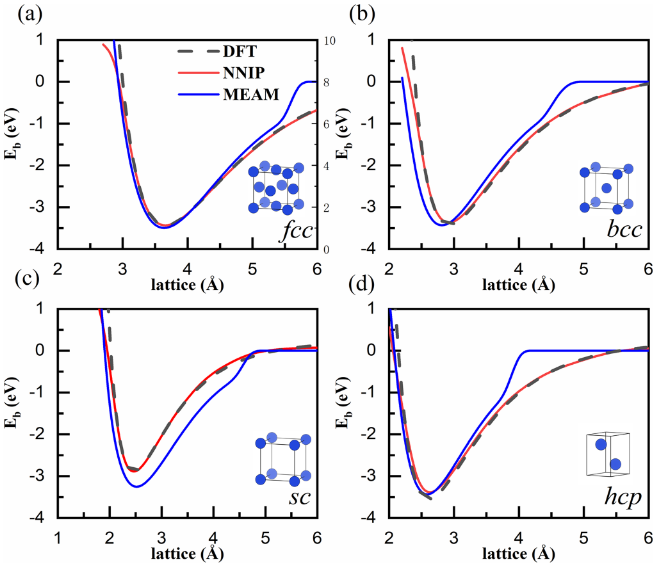

FIG. 6. The binding energy curves obtained by NNIP, DFT, and MEAM. The black dotted line represents the DFT calculation, the red line represents the NNIP potential calculation, and the blue line represents the MEAM potential calculation. (a) fcc binding energy curves; (b) bcc binding energy curves; (c) sc binding energy curves; (d) hcp binding energy curves.

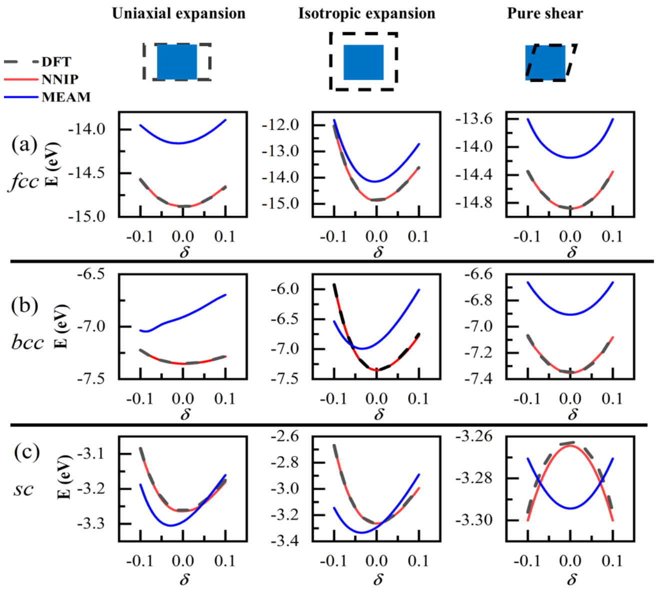

FIG. 7. The energy curves of DFT, NNIP, and MEAM when fcc, bcc, and sc are different strains (uniaxial expansion, isotropic expansion, and pure shear and). (a) The energy change curves of fcc during deformation. (b) The energy change curves of bcc during deformation. (c) The energy change curves of sc during deformation.

significantly worse results for crystals with non-close-packed lattice systems.

In addition to the binding energy curve, bulk structures under different types of deformation and distortion (uniaxial tension, isotropic expansion, and pure shear) are studied. The range of variation is set to be ±10% (Fig. 7). From Fig. 7, it can be seen that the energy curves obtained by the NNIP potential agree well with DFT data, with a RMSE of around 0.01 eV. The lowest point obtained by the MEAM potential in Figs. 7(b), 7(e), 7(g), and 7(h) has an apparent offset from the DFT results, mostly due to the error stemming from non-close-packed systems, and the curves of Figs. 7(d), 7(e), and 7(g) have completely different curvatures with DFT. In addition, both the DFT calculation and NNIP show that the sc structure of Cu is a transition state structure [Fig. 7(i)], while the MEAM potential predicts that it is a local minimum. In general, the MEAM potential performs well for fcc structures that are in the process of deformation, but it does not apply to other crystal symmetries. In addition, we also calculated the elastic constants, which requires precision, and the results of the comparison of NNIP and MEAM with DFT are listed in Table IV. The average deviation of NNIP from DFT is 4.6 GPa while that of MEAM is around 16.3 GPa. It is also worth noting that NNIP correctly captures the instability of sc structure with respect to shear distortion, while MEAM incorrectly labels it as a local minimum. Overall, NNIP performs better than MEAM in predicting elastic constants.

TABLE III. List of results from NNIP, DFT, and MEAM calculations of lattice constants and structural energy differences between different crystals of Cu.

	Lattice (Å)			$\Delta E_{\text{structure}}$ (eV)		
	fcc	bcc	sc	$\Delta E_{\text{fcc}\to\text{bcc}}$	$\Delta E_{\text{fcc}\to\text{sc}}$	$\Delta E_{\text{fcc}\to\text{hcp}}$
DFT	3.637	2.889	2.413	0.046	0.458	0.008
NNIP	3.647	2.894	2.402	0.040	0.456	0.003
MEAM	3.620	2.790	2.330	0.044	0.205	0.013
δ (NNIP)	0.28%	0.17%	0.46%	13.04%	0. 44%	62.50%
δ (MEAM)	0. 47%	3.43%	3.44%	10.00%	55.04%	333.33%

2. *Properties of Cu surfaces*

Surface energy is one of the most basic properties of material. It is defined as the energy required to cleave a surface of a bulk material as follows:

$$\gamma = \frac{1}{2A}(E_{\text{slab}} - NE_{\text{bulk}}), \tag{16}$$

where E_{slab} denotes the energy of the symmetric slab model, E_{bulk} denotes the energy of the atom in the bulk, and A denotes surface area of slab.

We first investigate the surface energy of fcc Cu. As can be seen from Fig. 8(a), the order of stability of the fcc surface predicted by the NNIP follows the order (111) > (100) > (110) > (211) > (210), consistent with DFT calculations. This is also in line with the physical intuition that a close-packed (111) facet is most stable and has the lowest surface energy. The RMSE of the difference between the surface energies obtained by the NNIP and DFT is around 0.003 eV/Å^2. For comparison, the MEAM potential correctly predicts the stability of most surfaces, but it deviates significantly for a non-close-packed (100) facet. Overall, both MEAM and NNIP predict with reasonable accuracy the surface energies of fcc Cu. To further test the robustness of MEAM and NNIP for surface energies of Cu with other crystallographic symmetry, we chose non-close-packed bcc Cu as another case [Fig. 8(c)].

TABLE IV. List of results from NNIP, DFT, and MEAM calculations of elastic constants between different crystals of Cu.

	fcc (GPa)			bcc (GPa)			sc (GPa)		
	C11	C12	C44	C11	C12	C44	C11	C12	C44
DFT	178.45	115.97	91.66	114.13	142.81	81.26	284.48	6.48	saddle point
NNIP	180.56	110.61	81.23	129.27	127.52	98.58	276.49	12.40	saddle point
MEAM	169.44	115.04	63.99	33.68	129.49	87.28	259.93	−20.14	local minima

The order predicted by the NNIP potential is (110) > (210) > (100) > (211) > (111), which only deviates slightly from the DFT on the (100) facet. In contrast, the MEAM potential grossly overestimated the surface energy of bcc Cu, where the surface energy error reaches as high as 40%, without any correlation to the stability order as predicted by DFT. The relative error in surface energy for MEAM benchmarked against DFT is 13% compared to that of 4% for NNIP.

We further compare the relationship between G and surface energy in Figs. 8(a) and 8(c). This is well reflected in the value of G_s where the (111) facet for fcc is higher than any other facets, and the order of G_s is consistent with the trend of surface energy change. We can also see that the surface energy is no longer directly correlated to the G_s background density on non-close-packed structures (bcc). We can infer that the surface energy is not only related to G_s, but also to G_p, G_d, and G_f. It is found from regression analysis that the surface energy and G are multivariate linearly correlated [Figs. 8(b) and 8(d)], where R^2 for fcc is 0.999 92 and for bcc is 0.999 96. From the multiple regression analysis of the two surface structures, the weights of G_p and G_f are considerably different. Because G_f has a narrower range of angle contribution than G_p and there are many small-angle atom pairs in the close-packed structures, the weight of G_f is smaller than G_p, but the opposite situation exists in non-close-packed structures.

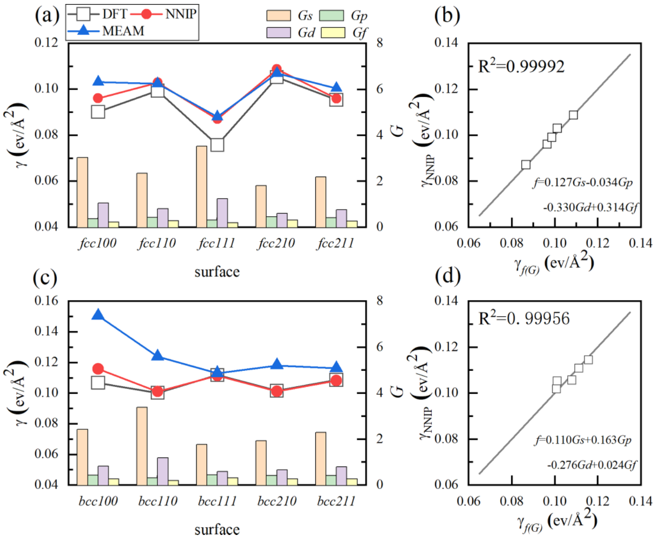

FIG. 8. (a) The surface energy of fcc surfaces obtained by NNIP, DFT, and MEAM. (b) The correlation between G and fcc surface energy. (c) The surface energy of bcc surfaces obtained by NNIP, DFT, and MEAM. (d) The correlation between G and bcc surface energy.

3. Chemical potentials of Cu nanoparticles

This section evaluates the accuracy of the NNIP potential simulation for nanoparticle configurations, which has a large ratio of surface atoms that are critically important for catalysis and other surface-related processes. In addition to terrace atoms that are similar to surface atoms, NPs have a large number of edge and vertex atoms that are even lower coordinated. We compare the chemical potential energy of polyhedral nanoparticles of different sizes as obtained from DFT and NNIP, which is defined as Eqs. (15).

It can be seen from Fig. 9 that the smaller the size of the nanoparticles is, the higher is the chemical potential, indicating the instability of nanoparticles going to smaller sizes. Comparing the chemical potential curves from MEAM and NNIP, we can see that both potentials reproduce faithfully the DFT chemical potentials for nanoparticles larger than 50 atoms. Reducing the size to under 50, the electronic effect of the asymmetric structure is underestimated in MEAM, which leads to the underestimation of the chemical potential as compared to the DFT data. This is in line with the phenomenon observed in Ref. [48], from which we can explain the variation of the size and melting point of the nanoparticles. Moreover, RMSE obtained by NNIP and DFT is 0.020 eV/at while the origin MEAM is an order of magnitude higher at 0.214 eV/at, so it can be considered that the NNIP is superior in terms of accuracy in simulating the configuration with small nanoparticle sizes.

4. Subnanometer Cu clusters

Since the cluster consists of only a small number of atoms, the total energy is strongly dependent on the relative position of atoms, where even small changes in structure can cause

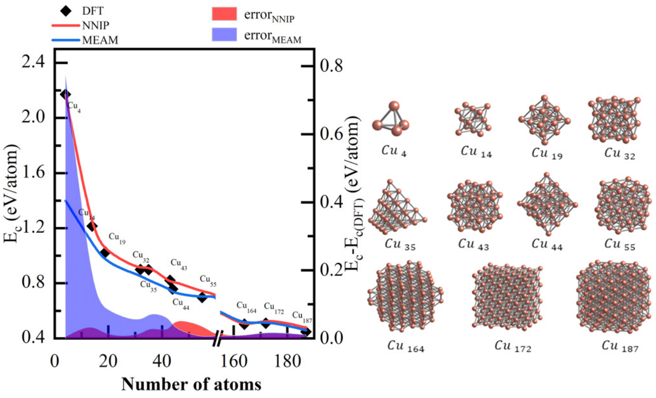

FIG. 9. Left: chemical potential curves of different nanoparticles changing with size and the error of different nanoparticles changing with size. Right: structural graph of Cu_n ($4 \leqslant n \leqslant 187$).

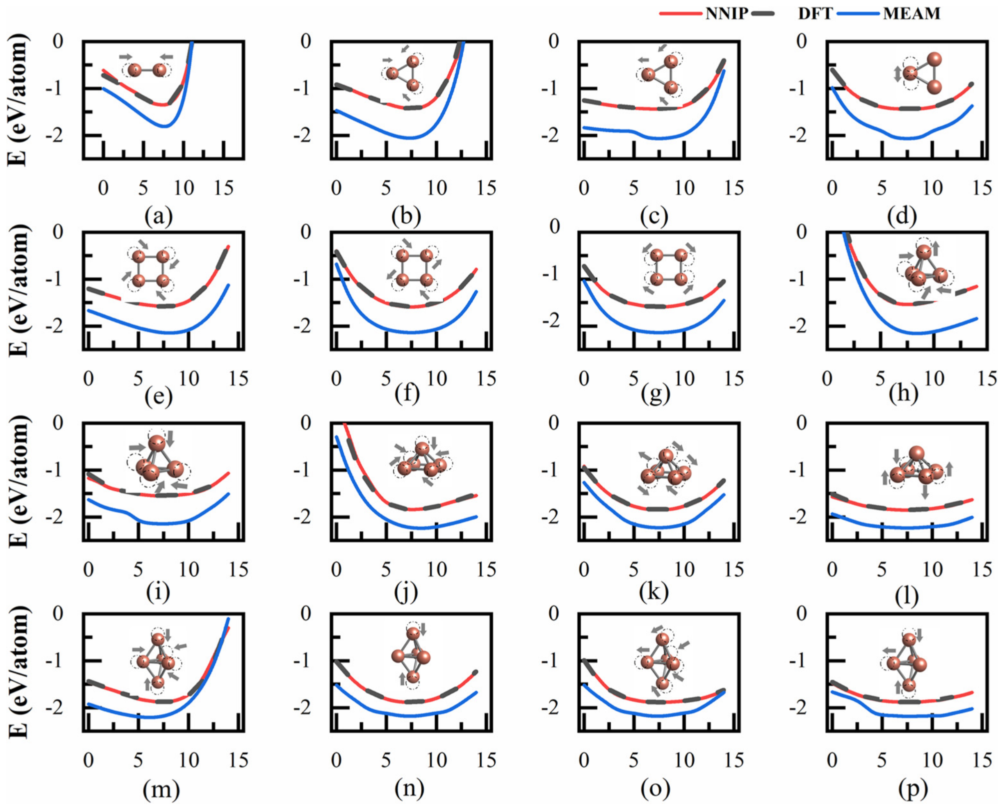

FIG. 10. The energy curves when small clusters (up to Cu$_5$) are distorted along certain vibration modes, with gray arrows indicating the mode of vibration.

significant energy changes. At the same time, the small cluster will exhibit different properties from the large-sized nanoparticles [49,50], so we use Cu clusters as the ultimate test of the robustness of the NNIP. Besides predicting correctly the energy of these small clusters at a fixed geometry, another critical test is the geometric stability of the cluster. To validate that the NNIP has a similar potential energy surface profile around the optimized local configurations, we distort each Cu cluster along typical vibration modes and trace their energy change to evaluate the reliability of the NNIP.

In Fig. 10, a total of 16 typical vibration modes for Cu$_n$ $(2 \leqslant n \leqslant 5)$ clusters are listed. Their corresponding energy curves are obtained by DFT, the MEAM potential, and NNIP, respectively. The energy trends obtained by DFT and NNIP are very consistent and form smooth curves, with an average error of 0.008 eV/atom, well within the DFT accuracy limit. The MEAM potential has an overall overestimation of the binding energy value, as most of the curves are located below the DFT. We also observe in Figs. 10(c), 10(d), 10(i), and 10(p) that there is a certain degree of curvature change in MEAM that is not consistent with DFT. The NNIP shows smaller errors compared to MEAM for smaller nanoparticles, providing a unified formalism capable of simulating metallic systems from bulk to small clusters.

5. *Comparison to HDNNPs*

To evaluate the efficiency and accuracy of the MEAM inspired NNIP, we compare it against Behler's HDNNP [34]. For consistency, we used the same dataset and the same neural network structure, with two hidden 40-node layers to train the HDNNP and NNIP model, respectively. In the HDNNP formalism, a total of 51 atom-center symmetry functions (the radial terms *G1*, *G2*, and *G3* and the angular terms *G4* and *G5*) were required to describe the copper-copper interactions in both bulk and surface systems. In NNIP, the background density of the MEAM potential as symmetry functions helps transformthe atomic coordinates into feature vectors that represent local chemical environments, allowing the use of fewer descriptors (only 32 descriptors) while retaining accuracy. The use of MEAM-inspired feature vectors not only makes optimization easier due to the smaller number of

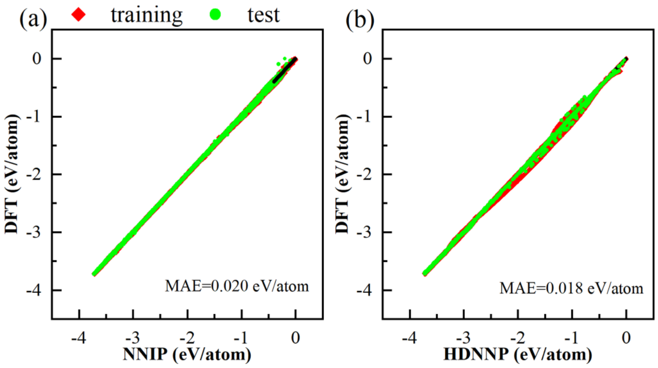

FIG. 11. Comparison of the error between the NNIP (a) and HDNNP (b) potential and DFT; the red dot represents the training set, and the green square represents the test set.

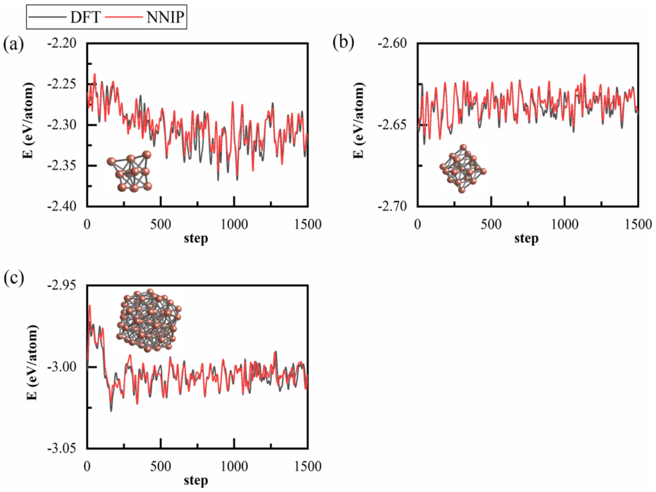

FIG. 12. MD at *NVT* ensemble for simulating the heating process of nanoparticles Cu_{10} (a), Cu_{19} (b), and Cu_{55} (c) at 500 K.

hyperparameters (Rc and η), but it also reduces the algorithmic complexity of $O(n^3)$ involved in the calculation of symmetry function $G5$ in HDNNP (which requires three atoms per group to compute the angle term) to $O(n^2)$. Our preliminary testing showed that the construction of symmetry functions derived from the MEAM potential is faster than HDNNP for reasonably large systems [51].

The accuracy of HDNNP and NNIP was further compared to illustrate the reliability of the MEAM-derived features. Figure 11 show the correlation between these two mathematical potentials with respect to DFT datasets, with an average error of 0.018 eV/at for HDNNP and 0.011 eV/at for NNIP. There are no obvious outliers in either case, demonstrating the capability of symmetry functions (G_s, G_p, G_d, and G_f) in accurately fitting the complex PES and achieving comparable accuracy to HDNNP with fewer features.

6. Application example: MD of small Cu clusters

To assess the accuracy of the NNIP potential on disordered structures such as nanoparticles, we conducted three sets of MD simulations with an *NVT* ensemble at 500 K for 10-, 19-, and 55-atom nanoparticles, respectively (Fig. 12). All these intermediate configurations are not in the training dataset. As can be seen from Fig. 12, the potential energies along trajectories for NNIP and DFT are very close, with a mean error of 0.011 eV/at. This suggests that the NNIP potential has good generalization capability and can perform robustly in treating nanoparticle structures.

IV. SUMMARY

We constructed NNIP using MEAM background density as symmetry functions and extended its application to nanoparticles and cluster. The method can be applied to different structures such as bulk, slab, and nanoparticles, with reasonable accuracy and reliability. NNIP inherits the insights of MEAM feature construction and combines the high flexibility of the NN, making it a highly transferrable potential with a solid theoretical foundation. NNIP can be applied to the study of metal catalysts in molecular-dynamics simulations.

ACKNOWLEDGMENTS

The work is supported by the National Natural Science Foundation of China (Grants No. 51871103 and No. 51835005), the Major Science and Technology Programs of Yunnan (202002AB080001-1), program for HUST Academic Frontier Youth Team (2018QYTD03), and the Fundamental Research Funds for the Central Universities (HUST:2020JYCXJJ054). Thanks to SCTS/CGCL HPCC of HUST for providing computing resources and technical support.

---

[1] P. Raccuglia, K. C. Elbert, P. D. F. Adler, C. Falk, M. B. Wenny, A. Mollo, M. Zeller, S. A. Friedler, J. Schrier, and A. J. Norquist, Machine-learning-assisted materials discovery using failed experiments, Nature (London) **533**, 73 (2016).

[2] A. Agrawal and A. Choudhary, Perspective: Materials informatics and big data: Realization of the "fourth paradigm" of science in materials science, APL Mater. **4**, 053208 (2016).

[3] G. Hautier, C. C. Fischer, A. Jain, T. Mueller, and G. Ceder, Finding nature's missing ternary oxide compounds using machine learning and density functional theory, Chem. Mater. **22**, 3762 (2010).

[4] C. C. Fischer, K. J. Tibbetts, D. Morgan, and G. Ceder, Predicting crystal structure by merging data mining with quantum mechanics, Nat. Mater. **5**, 641 (2006).

[5] S. Curtarolo, D. Morgan, K. Persson, J. Rodgers, and G. Ceder, Predicting Crystal Structures with Data Mining of Quantum Calculations, Phys. Rev. Lett. **91**, 135503 (2003).

[6] F. F. Abraham, R. Walkup, H. Gao, M. Duchaineau, T. D. De La Rubia, and M. Seager, Simulating materials failure by using up to one billion atoms and the world's fastest computer: Work-hardening, Proc. Natl. Acad. Sci. USA **99**, 5783 (2002).

[7] A. E. Garcia and K. Y. Sanbonmatsu, α-Helical stabilization by side chain shielding of backbone hydrogen bonds, Proc. Natl. Acad. Sci. USA **99**, 2782 (2002).

[8] X.-Q. Chen, Boosting the discovery of 3D topological materials: mixing chemistry with physics via a two-step computational screening strategy, Natl. Sci. Rev. **5**, 316 (2018).

[9] J. M. Cole, K. S. Low, H. Ozoe, P. Stathi, C. Kitamura, H. Kurata, P. Rudolf, and T. Kawase, Data mining with molecular design rules identifies new class of dyes for dye-sensitised solar cells, Phys. Chem. Chem. Phys. **16**, 26684 (2014).

[10] J. Kibsgaard, C. Tsai, K. Chan, J. D. Benck, J. K. Nørskov, F. Abild-Pedersen, and T. F. Jaramillo, Designing an improved transition metal phosphide catalyst for hydrogen evolution using experimental and theoretical trends, Energy Environ. Sci. **8**, 3022 (2015).

[11] V. Sharma, C. Wang, R. G. Lorenzini, R. Ma, Q. Zhu, D. W. Sinkovits, G. Pilania, A. R. Oganov, S. Kumar, G. A. Sotzing, S. A. Boggs, and R. Ramprasad, Rational design of all organic polymer dielectrics, Nat. Commun. **5**, 4845 (2014).

[12] W. Koch and M. C. Holthausen, *A Chemist's Guide to Density Functional Theory* (Wiley, Weinheim, 2015).

[13] J. M. Haile, *Molecular Dynamics Simulation, Elementary Methods* (John Wiley & Sons, Inc. New York, 1992); Adv. Mater. **5**, 223 (1993).

[14] W. G. Hoover, *Molecular Dynamics* (Springer-Verlag, Berlin, 1986).

[15] A. Janotti, Van de Walle, and G. Chris, Native point defects in ZnO, Phys. Rev. B **76**, 165202 (2007).

[16] V. Padmanabhan, A. L. Frischknecht, and M. E. Mackay, Effect of chain stiffness on nanoparticle segregation in polymer/nanoparticle blends near a substrate, Macromol. Theor. Simul. **21**, 98 (2012).

[17] J. Nord, K. Nordlund, J. Keinonen, and K. Albe, Molecular dynamics study of defect formation in GaN cascades, Nucl. Instrum. Methods Phys. Res. Sect. B **202**, 93 (2003).

[18] J.-H. Shim, H.-J. Lee, and B. D. Wirth, Molecular dynamics simulation of primary irradiation defect formation in Fe-10%Cr Alloy, J. Nucl. Mater. **351**, 56 (2006).

[19] A. K. Rappé, C. J. Casewit, K. S. Colwell, W. A. Goddard III, and W. M. Skiff, UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations, J. Am. Chem. Soc. **114**, 10024 (1992).

[20] J. Tersoff, New Empirical Model for the Structural Properties of Silicon, Phys. Rev. Lett. **56**, 632 (1986).

[21] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, A second generation force field for the simulation of proteins, nucleic acids, and organic molecules, J. Am. Chem. Soc. **117**, 5179 (1995).

[22] J. Tersoff, New empirical approach for the structure and energy of covalent systems, Phys. Rev. B **37**, 6991 (1988).

[23] J. Cai and Y. Y. Ye, Simple analytical embedded-atom-potential model including a long-range force for fcc metals and their alloys, Phys. Rev. B **54**, 8398 (1996).

[24] M. S. Daw and M. I. Baskes, Semiempirical, Quantum Mechanical Calculation of Hydrogen Embrittlement in Metals, Phys. Rev. Lett. **50**, 1285 (1983).

[25] M. I. Baskes, J. S. Nelson, and A. F. Wright, Semiempirical modified embedded-atom potentials for silicon and germanium, Phys. Rev. B **40**, 6085 (1989).

[26] M. I. Baskes, Modified embedded-atom potentials for cubic materials and impurities, Phys. Rev. B **46**, 2727 (1992).

[27] Y.-M. Kim, B.-J. Lee, and M. I. Baskes, Modified embedded-atom method interatomic potentials for Ti and Zr, Phys. Rev. B **74**, 014101 (2006).

[28] J.-M. Zhang, F. Ma, and K.-W. Xu, Calculation of the surface energy of FCC metals with modified embedded-atom method, Appl. Surf. Sci. **229**, 34 (2004).

[29] B.-J. Lee and M. I. Baskes, Second nearest-neighbor modified embedded-atom-method potential, Phys. Rev. B **62**, 8564 (2000).

[30] B.-J. Lee, M. I. Baskes, H. Kim, and Y. K. Cho, Second nearest-neighbor modified embedded atom method potentials for bcc transition metals, Phys. Rev. B **64**, 184102 (2001).

[31] X. Duan, B. Zhou, R. Chen, H. Zhou, Y. Wen, and B. Shan, Development of lattice inversion modified embedded atom method and its applications, Curr. Appl. Phys. **14**, 1794 (2014).

[32] B. Lee and K. Cho, Extended embedded-atom method for platinum nanoparticles, Surf. Sci. **600**, 1982 (2006).

[33] B. Shan, L. Wang, S. Yang, J. Hyun, N. Kapur, Y. Zhao, J. B. Nicholas, and K. Cho, First-principles-based embedded atom method for PdAu nanoparticles, Phys. Rev. B **80**, 035404 (2009).

[34] J. Behler, Atom-centered symmetry functions for constructing high-dimensional neural network potentials, J. Chem. Phys. **134**, 074106 (2011).

[35] A. Takahashi, A. Seko, and I. Tanaka, Conceptual and practical bases for the high accuracy of machine learning interatomic potential, Phys. Rev. Materials **1**, 063801 (2017).

[36] M. Rupp, A. Tkatchenko, K.-R. Müller, and O. A. von Lilienfeld, Fast and Accurate Modeling of Molecular Atomization Energies with Machine Learning, Phys. Rev. Lett. **108**, 058301 (2012).

[37] L. Zhang, J. Han, H. Wang, R. Car, and Weinan E., Deep Potential Molecular Dynamics: A Scalable Model with the Accuracy of Quantum Mechanics, Phys. Rev. Lett. **120**, 143001 (2018).

[38] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, Without the Electrons, Phys. Rev. Lett. **104**, 136403 (2010).

[39] A. P. Bartók and G. Csányi, Gaussian approximation potentials: a brief tutorial introduction, Int. J. Quantum Chem. **115**, 1051 (2015).

[40] J. Behler, Perspective: Machine learning potentials for atomistic simulations, J. Chem. Phys. **145**, 170901 (2016).

[41] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, and M. Devin, Tensorflow: large-scale machine learning on heterogeneous distributed systems, arXiv:1603.04467.

[42] G. Cybenko, Approximation by superpositions of a sigmoidal function, Math. Control Signals Syst. **5**, 455 (1992).

[43] F. Santosa and W. W. Symes, Linear inversion of band-limited reflection seismograms, SIAM J. Sci. Stat. Comput. **7**, 1307 (1986).

[44] A. Y. Ng, Feature selection, L 1 vs. L 2 regularization, and rotational invariance, in *Proceedings of the Twenty-First International Conference on Machine Learning* (Association for Computing Machinery, New York, 2004), p. 78.

[45] O. T. Unke and M. Meuwly, A reactive, scalable, and transferable model for molecular energies from a neural network approach based on local information, J. Chem. Phys. **148**, 241708 (2018).

[46] N. Artrith and A. Urban, An implementation of artificial neural-network potentials for atomistic materials simulations: performance for TiO2, Comput. Mater. Sci. **114**, 135 (2016).

[47] J. Hafner, Ab-Initio simulations of materials using VASP: density-functional theory and beyond, J. Comput. Chem. **29**, 2044 (2008).

[48] Ph. Buffat and J.-P. Borel, Size effect on the melting temperature of gold particles, Phys. Rev. A **13**, 2287 (1976).

[49] O. Varnavski, G. Ramakrishna, J. Kim, D. Lee, and T. Goodson, Optically excited acoustic vibrations in quantum-sized monolayer-protected gold clusters, ACS Nano **4**, 3406 (2010).

[50] K. Judai, S. Abbet, A. S. Wörz, U. Heiz, and C. R. Henry, Low-temperature cluster catalysis, J. Am. Chem. Soc. **126**, 2732 (2004).

[51] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevB.102.144107 for more details.