Received: 1 July 2021 | Revised: 21 September 2021 | Accepted: 27 September 2021

DOI: 10.1002/msd2.12021

REVIEW ARTICLE

IJMSD International Journal of Mechanical System Dynamics Open Access WILEY

Machine-learning-based interatomic potentials for advanced manufacturing

Wei Yu[1] | Chaoyue Ji[2] | Xuhao Wan[1] | Zhaofu Zhang[3] | John Robertson[1,3] | Sheng Liu[2,4] | Yuzheng Guo[1,2]

[1]School of Electrical Engineering and Automation, Wuhan University, Wuhan, Hubei, China

[2]Institute of Technological Sciences, Wuhan University, Wuhan, Hubei, China

[3]Engineering Department, Cambridge University, Cambridge, UK

[4]School of Power and Mechanical Engineering, Wuhan University, Wuhan, Hubei, China

Correspondence

Prof. Sheng Liu, Institute of Technological Sciences, Wuhan University, Wuhan 430061, Hubei, China.
Email: shengliu@whu.edu.cn

Prof. Yuzheng Guo, School of Electrical Engineering and Automation, Wuhan University, Wuhan 430061, Hubei, China.
Email: yguo@whu.edu.cn

Funding information

National Natural Science Foundation of China, Grant/Award Numbers: 51727901, U1501241, 62174122; National Key R&D Program of China, Grant/Award Number: 2017YFB1103904; Wuhan University Junior Faculty Research, Grant/Award Number: 2042019KF0003; Hubei Provincial Natural Science Foundation of China, Grant/Award Number: 2020CFA032

Abstract

This paper summarizes the progress of machine-learning-based interatomic potentials and their applications in advanced manufacturing. Interatomic potential is essential for classical molecular dynamics. The advancements made in machine learning (ML) have enabled the development of fast interatomic potential with ab initio accuracy. The accelerated atomic simulation can greatly transform the design principle of manufacturing technology. The most widely used supervised and unsupervised ML methods are summarized and compared. Then, the emerging interatomic models based on ML are discussed: Gaussian approximation potential, spectral neighbor analysis potential, deep potential molecular dynamics, SCHNET, hierarchically interacting particle neural network, and fast learning of atomistic rare events.

KEYWORDS

advanced manufacturing, interatomic potential, machine learning, molecular dynamics

1 | INTRODUCTION

Advanced manufacturing using new materials and emerging technologies encompasses all the aspects of industrial production. In many fields, such as microelectronics and nanoelectronics, the cost of developing and producing the tools for manufacture has become a major limitation for the whole system. Therefore, it has become increasingly important to include simulation early in the design process. Atomic simulation techniques such as molecular dynamics (MD) and density functional theory (DFT) make it possible to explore materials with minimal investment in the conventional and expensive trial-and-error experiments. However, the current length and time scale of atomic simulation are still markedly limited by the computational resource and algorithm. It is a trade-off between speed and accuracy, limiting its application to macroscopic lengths and time frames.

The accuracy and performance prediction capabilities of classic MD simulations depend on the empirical force field. The development of a reliable potential function for a specific system is critical for generating results with scientific robustness. For silicon and other fully researched materials, it is a common practice to perform large-scale simulations by

selecting experienced/classical potential. However, many alloys and amorphous materials have complex structures and interactions, which makes the construction of the classical potential challenging.

The electronic structure calculations utilizing DFT can provide useful insights into properties, while the simulation of large-scale system is computationally expensive. Typical examples are calculations of properties of phase-change alloys, thermal transport phenomena of disordered systems, and ductility and fracture by complex atomic-scale mechanisms.

Machine learning (ML) technology to construct interatomic potentials with (quasi) ab initio accuracy and computational efficiency is (almost) comparable to classical potentials. ML promotes the application of MD simulation in terms of two aspects. It can accelerate the development of traditional empirical potential functions as well as directly represent MD systems by using ML potentials. For traditional potential functions, genetic algorithms and other ML methods have considerably shortened the life cycle of parameter optimization. Machine-learned spectral neighbor analysis potential (SNAP) models can well establish the high-performing embedded atom method (EAM) and modified EAM potentials for face-centered cubic (fcc) Cu and Ni.[1] ReaxFF parameterization is an efficient artificial neural network (ANN)-based method to simulate the reactive dynamics of the zinc-oxide model.[2] An Intelligent ReaxFF (I-ReaxFF) model in terms of matrix (or tensor) operations based on the TensorFlow coding platform can optimize ReaxFF parameters automatically.[3] More importantly, ML potentially establishes a direct relationship between the atomic structure and the system energy. It uses electronic structure information to model the MD system and does not make any physical approximations to the functional form.

The machine-learning-based interatomic potential has already been developed and is widely used in various fields of materials research. This review is divided into the following sections. The most widely used ML methods are summarized in Section 2. A few emerging methods to generate interatomic potentials are introduced in Section 3. A few applications in machine-learning-based interatomic potentials for advanced manufacturing are compared in Section 4.

2 | ML METHODS

Recently, artificial intelligence (AI) has been attracting worldwide attention, benefiting from the development of computer science and data processing technology. As the most significant part of AI, rapid developments have been achieved in the ML field in the last 40 years.[4,5] ML shows outstanding performance in many fields such as pattern recognition,[6] image segmentation,[7] medical diagnosis,[8] catalytic reactions, and autonomous driving cars[9] by reorganizing existing knowledge structures and mining implicit relationships. For example, the possible synthetic route of chemicals based on the results of deep neural networks has a similar feasibility to those formulated by human experts.[10] ML can even collect valuable data from failed experiments.[11]

A typical workflow to apply the ML commonly begins with unsolved problems, then acquisition of sufficient data samples, extraction of appropriate features, and finally, establishment of the ML model.[12] The ML model can be adopted in the prediction problem and even mechanism analysis. Several ML algorithms have already been proposed and developed. Considering the research target and the sample size, parts of the ML algorithms are more suitable than others in different research topics. Several popular ML algorithms are briefly introduced herein, and their suitability is discussed in the following paragraphs.

Linear regression is a basic strong algorithm commonly used to illustrate the relations and solve the problem for predicting target values in the corresponding research.[13] Linear regression algorithms usually consist of least absolute shrinkage and selection operator regression (LASSO),[14] elastic net regression (ENR),[15] and so forth. The basis of the linear regression algorithm is the ordinary least squares (OLS) method as shown in Figure 1A. The aim of the OLS method is to minimize the sum of squares of the differences in the ML-predicted and experimental values (L) by adjusting $\hat{\beta}$ as shown in the following equations:

$$\hat{y} = \hat{\beta} X, \tag{1}$$

$$L = \sum_i |y_i - \hat{y}_i|^2. \tag{2}$$

To avoid the challenge of overfitting for OLS owing to the features of multicollinearity and weak relation, regularization linear algorithms have been exploited. LASSO and ridge regularization are the most widespread regularized linear models that are established by $\mathcal{L}_{L1}$ and $\mathcal{L}_{L2}$ regularization methods, respectively. In addition, the ENR algorithm is combined with the above two regularization algorithms and has the advantages of both linear algorithms. The linear model is commonly used to establish models at first because it is useful to gain beneficial insights into the feature relations. Generally, linear models can reveal the linear feature interactions, but may perform poorly when the features have multiple collinearities.[13]

Kernel regression algorithms, as shown in Figure 1B, can investigate the nonlinear correlation of a data set by combining the kernel scheme with the nonlinear kernel.[17] Support vector regression (SVR),[18] Gaussian process regression (GPR),[19] and kernel ridge regression (KRR)[20] are the most widespread kernel algorithms. Kernel regression can map the data points to a higher-dimensional feature space at first by calculating $k^*(x \cdot x)$ (k represents the kernel function).[21] Then, the mapping relations can be applied to replace the $\hat{\beta}$ in Equation (1). KRR is a typical ridge regularization linear algorithm combined with the kernel method. SVR is similar to KRR, but its loss function is completely different. The kernel regression models are excellent in terms of the feature of multicollinearity, but it is difficult to choose a suitable kernel, and it is generally not universal.[22]

Decision tree regression is also a common used supervised algorithm in the ML field.[23] Decision tree regression usually consists of k-neighbor regression (KNR),[24] random forest regression (RFR),[25] extra trees regression (ETR),[26] and so forth. As shown in Figure 1C, a decision tree can construct a continuous decision boundary by deciding numerous "if-else" problems as the branches of the tree, until hitting a terminal leaf node. Then, those decisions divide the entire feature space into different regions and obtain the decision threshold by the loss function.[27] Finally, the model can be used to predict the target values from the average value

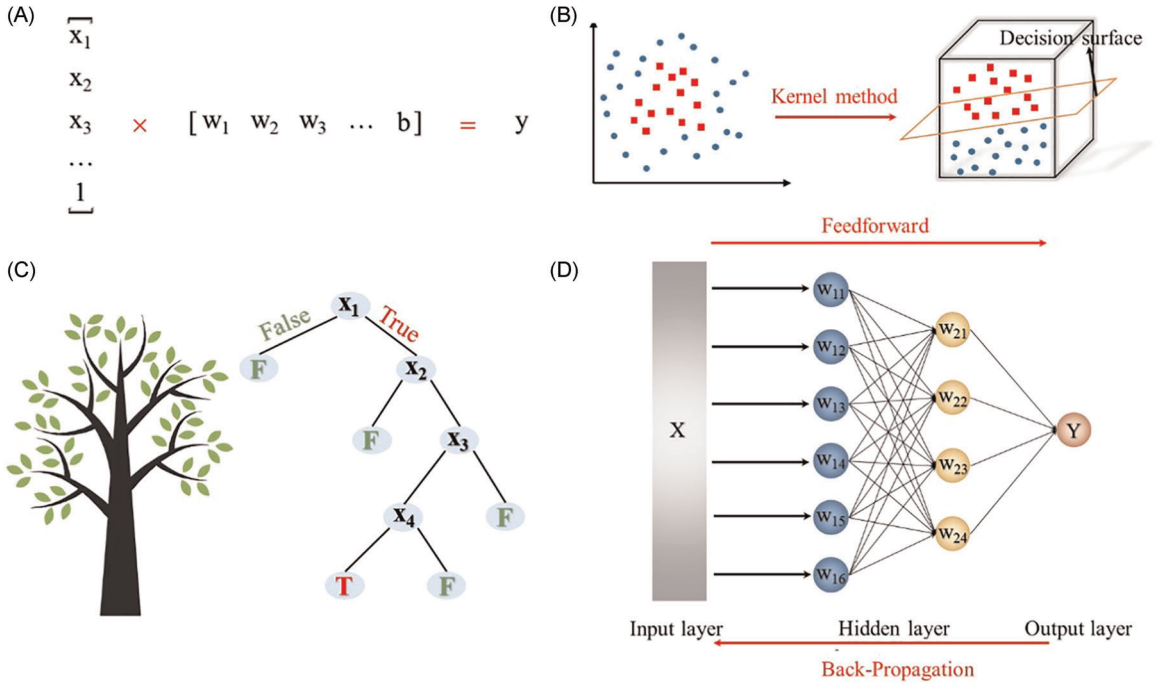

FIGURE 1 Sketch map of (A) the linear regression algorithms, (B) kernel regression algorithms, (C) decision tree algorithms, and (D) neural networks algorithms in the machine learning framework, respectively. Reproduced under terms of the CC-BY license.[16] Copyright 2021, The Authors, published by AIP

located in the selected area of the feature space. Decision tree algorithms are favorable when both the linear and kernel ML models show poor performance for the studied practical problems. When the linear model and the kernel model have low accuracy for the studied system, the decision tree algorithm is a good choice. The parameters tuning of the decision tree model is easy, and multicollinearity does not affect the final results much. However, the decision tree model is a non-parametric model, so that it is difficult to analyze the model, and it cannot be easily adapted to test data.[28] Besides, the ensemble method, which combines multiple basic ML models to establish a single optimal model, is also popular in the engineering field. Gradient boosting regression (GBR) is the most well-known ensemble method.[29,30] Decision tree algorithms such as random forest and extra trees can also improve the ML model by selecting the average prediction values from a single classifier.

The neural networks (NNs) algorithm is a rising star in the field of ML. It is powerful as it is inspired by the behavior of the brain.[31] Feedforward neural network (FNN) is one of the strongest NN models, which is commonly used.[32] As Figure 1D shows, the training data are transferred to the nodes (perceptron), which are linked to each other by edges with different defined weight parameters and thresholds. Each individual node is its own linear regression model, including input data, weights, threshold, and an output. After the input layer is determined, the weights are also determined. The weights can help evaluate the significance to the output compared to other inputs. The nodes gather together and can be divided into different separate hidden layers, in which weighted features are summed, and then the sum is transmitted into a nonlinear activation function (such as Relu and Sigmoid) to optimize the model.[33] If the output of the individual node is larger than the threshold value, the node will be activated and the data will be sent to the next layer. Otherwise, no data are passed along. The last layer called the output layer can output target values depending on the practical problem. NNs models are usually trained through a cyclic process called back-propagation, which is a repeatable optimization process of the weights of hidden layers. NNs algorithms have the characteristics of high applicability and excellent performance, but the weight parameters are also difficult to understand, and there are few training samples, which are not suitable for neural network models.[31]

Table 1 briefly presents the category of those aforementioned algorithms. The appropriate data set size of each model is also listed. For few-shot learning (10–200), linear algorithms and kernel regression models are the most suitable, while for medium size (e.g., 100–500), the decision trees models are the best. Considering that the performance of NNs models improves with increasing data set size, NNs models are especially appropriate for those large data sets (>500).

3 | EMERGING METHODS TO GENERATE INTERATOMIC POTENTIAL

ML interatomic potentials have shown comparable accuracy to ab initio MD methods, while the length and time scales of the computation have been markedly improved by orders of magnitude in MD simulations.[34] In this section, we will briefly introduce some commonly used methods to generate ML interatomic potentials, including NN- and kernel-based regression models.[35]

TABLE 1 Four common machine learning algorithm categories and their suitable data set size

Category	Algorithms	Sample size
Linear models	LASSO, ENR	Small
Kernel regression	SVR, GPR, KRR	Small
Decision trees	KNR, RFR, ETR	Medium
Neural networks	FNN	Large

Note: The GBR is an ensemble algorithm that uses many other machine learning algorithms.

3.1 | Gaussian approximation potential (GAP)

GAP[36] is a kernel-based method that can generate interatomic potential automatically from the data of the atomic configurations and the corresponding energies and forces, which are calculated using DFT-based methods.[37]

In this method, a bispectrum space is built to represent the atomic neighborhood by extending the 4D sphere using the specified coefficients. The bispectrum is given by[36]

$$B_{j_1 j_2 j} = \sum_{m_1, m_1 = -j_1}^{j_1} \sum_{m_2', m_2 = -j_2}^{j_2} \sum_{m', m = -j}^{j} \left(c^{j}_{m'm}\right)^* C^{jm}_{j_1 m_1 j_2 m_2} \times C^{jm'}_{j_1 m_1' j_2 m_2'} c^{j_1}_{m_1' m_1} c^{j_2}_{m_2' m_2}, \tag{3}$$

where C denotes the ordinary Clebsch–Gordan coefficients.[38] The local atomic environment with $j, j_1, j_2 \le J_{\max}$ can be described by the truncated version of the bispectrum. Then, nonparametric Gaussian process (GP) regression[39,40] is used to predict the atomic energy by[36]

$$E(b) = \sum_n \alpha_n e^{-(1/2)\sum_l [(b_l - b_{n,l})/\theta_l]^2} \equiv \sum_n \alpha_n G(b, b_n), \tag{4}$$

where n denotes the reference configurations, l the bispectrum components, and θ_l are the hyperparameters. The local atomic energy function is generated by a kernel function, which measures the similarity between the local environment and directly determines the success of the generated GAP. The configurations from two adjacent atoms or one atom during the two sequential MD steps are very similar, leading to strong correlation of the data set. The sparsification procedure is adopted to solve the problem by randomly sampling the configurations from the data set.

A widely used kernel function to train the GAP is the smooth overlap of atomic positions (SOAP) kernel,[41] which is computationally efficient. The computation cost can be greatly reduced by about one order of magnitude, with the prediction accuracy unchanged. The SOAP-GAP[42] model can also be applied to predict the molecular properties.

3.2 | Spectral neighbor analysis potential

The local environments are represented by the bispectrum components in SNAP,[43] which is the same as GAP. However, unlike GAP, the potential energy and bispectrum components are assumed to be linearly dependent in SNAP. This assumption makes it possible to use the linear functions of the SNAP coefficients to describe the total energy, atom forces, and stress tensor. The total energy can be written as[43]

$$E_{\mathrm{SNAP}}(r^N) = N\beta_0 + \beta \cdot \sum_{i=1}^{N} B^i, \tag{5}$$

where β_0 is a constant, indicating the element-specific contribution for each atom i; β and B^i are both the k-dimensional vectors of SNAP coefficients and bispectrum components, respectively.

The specific values of the hyperparameters chosen in SNAP directly determine the performance of the SNAP potential. As shown in Figure 2, the quantum mechanics (QM) data are fed into SNAP and processed to training data in the native large-scale atomic/molecular massively parallel simulator (LAMMPS) input format. LAMMPS then uses the input to calculate the energy and forces of the QM training data with first-principle accuracy, while SNAP predicts the energy and forces of the training data with the hyperparameters given by the DAKOTA[44] optimization toolkit. The errors between energy predicted by SNAP and calculated by LAMMPS are then aggregated into the loss function to optimize the hyperparameters of SNAP by DAKOTA.

3.3 | Deep potential molecular dynamics (DPMD)

DPMD[45] is an NN-based interatomic potential method that models the relationship between a sum of atomic energy $E = \sum_i E_i$ and the environment of all atoms by deep neural networks (DNN).[46] The energy E_i of atom i is determined by the local environment within a cutoff radius R_c around the atom. This can meet the requirement to preserve translation, rotation, and permutation invariance, as shown in Figure 3.

The coordinates R_i of atoms i inside R_c are fed into a fully connected network to generate the representation D_{ij} of the local environment of atom i. The $1/R_{ij}$ factor is introduced to the local environment D_{ij}, meaning that the weight is reduced linearly with an increase in distance, as the input of a feed-forward DNN model. D_{ij} is fed into the DNN from the input layer, and then flows to multiple hidden layers through a linear transformation. Following a nonlinear activation function,[47] DNN outputs the atomic energy of E_i for each atom. The potential energy E is obtained by the linear sum of all E_i in the final step. The loss function of DPMD is defined as[45]

$$L(p_\varepsilon, p_f, p_\xi) = p_\varepsilon \Delta\varepsilon^2 + \frac{p_f}{3N}\sum_i |\Delta F_i|^2 + \frac{p_\xi}{9}\|\Delta\xi\|^2, \tag{6}$$

where $p_\varepsilon, p_f, p_\xi$ are adjustable coefficients, Δ denotes the error between the DPMD prediction and the training data, N is the atomic

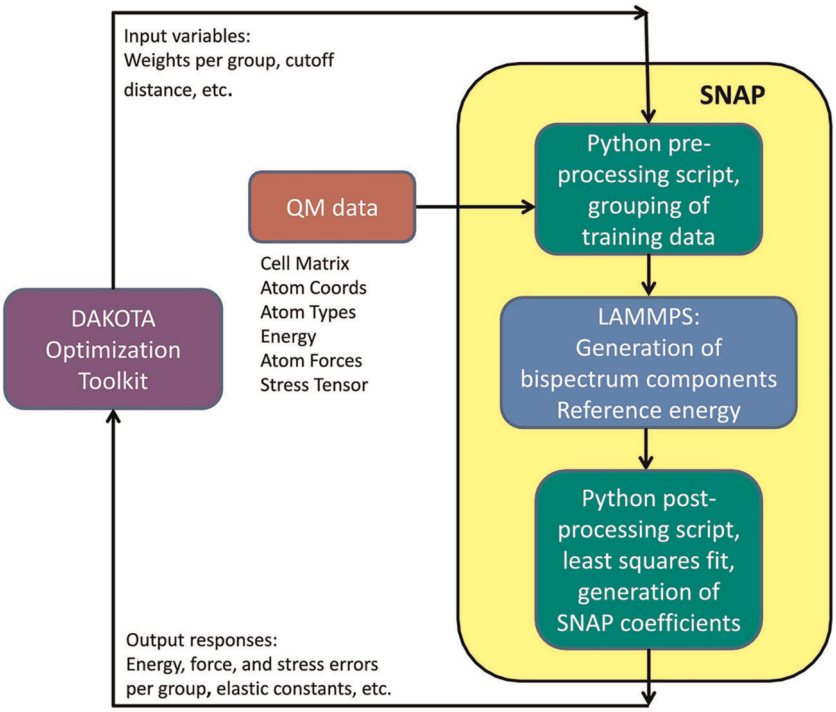

FIGURE 2 Generation of the SNAP potential in LAMMPS. Reproduced with permission.[43] Copyright 2015, Elsevier. SNAP, Spectral neighbor analysis potential

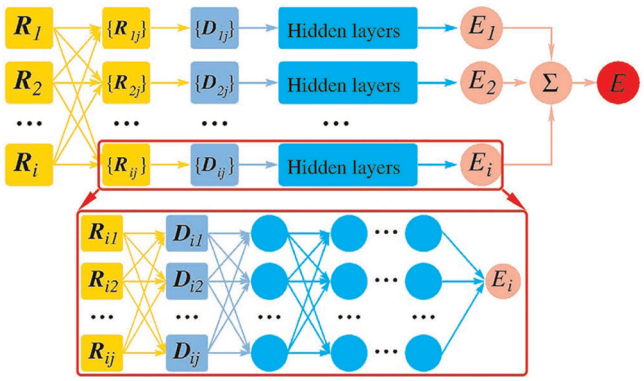

FIGURE 3 Network architecture of the DPMD model. Reproduced with permission.[45] Copyright 2018, American Physical Society. DPMD, Deep potential molecular dynamics

numbers, F_i denotes the force on the atom i, and ξ denotes the virial tensor.

The data-generated first-principles calculations will be fed into the DNN model, and the Adam[48] optimization method is applied to tune the weights w_{kl} and bias b_l of DNN. After training on adequate data sets of atomic environments and corresponding potential energies, the DNN model can propose large-scale MD simulations with an accuracy comparable with that of the AI-MD.

3.4 | SchNet

The deep learning architecture SchNet can model complex atomic interactions to predict the potential energies of the atomistic systems.[49] SchNet is designed to learn the representations for the local atomic environment that satisfy the invariant under all physical symmetries.

As shown in Figure 4, the network structure of SchNet is similar to deep tensor neural networks (DTNNs),[50] constructed by embedding layers, interaction layers, and atom-wise layers. Every atom in the system is expressed by a set of the charges $Z = (Z_1, \ldots, Z_n)$ and positions $R = (r_i, \ldots, r_i)$ of atoms. Through the embedding layers, the sets will be embedded into the feature tuples $X^l = (X_1^l, \ldots, X_n^l)$. The atom-wise layer is a fully connected layer. In interaction layers, SchNet applies continuous-filter convolutions with filter-generating networks to model the interaction term as[49]

$$x_i^{l+1} = (X^l * W^l)_i = \sum_{j=0}^{n_{\text{atoms}}} x_j^l \circ W^l(r_j - r_i), \tag{7}$$

where "$\circ$" denotes the element-wise multiplication.[49] This operation can reduce the dimension of the feature set to improve the computational efficiency.[51]

The Adam optimizer is applied to train SchNet for the target energy E and force F by minimizing a combined loss function[49]

$$\begin{aligned} &L((\hat{E}, \hat{F}_1, \ldots, \hat{F}_N), (E, F_1, \ldots, F_n)) \\ &= \rho \| E - \hat{E} \|^2 + \frac{1}{n_{\text{atoms}}} \sum_{i=0}^{n_{\text{atoms}}} \left\| F_i - \left(-\frac{\partial \hat{E}}{\partial R_i} \right) \right\|^2, \end{aligned} \tag{8}$$

where ρ is a tunable coefficient, which trades off between energy and force. The weights and bias of the SchNet model are trained using the mini-batch stochastic gradient descent method.

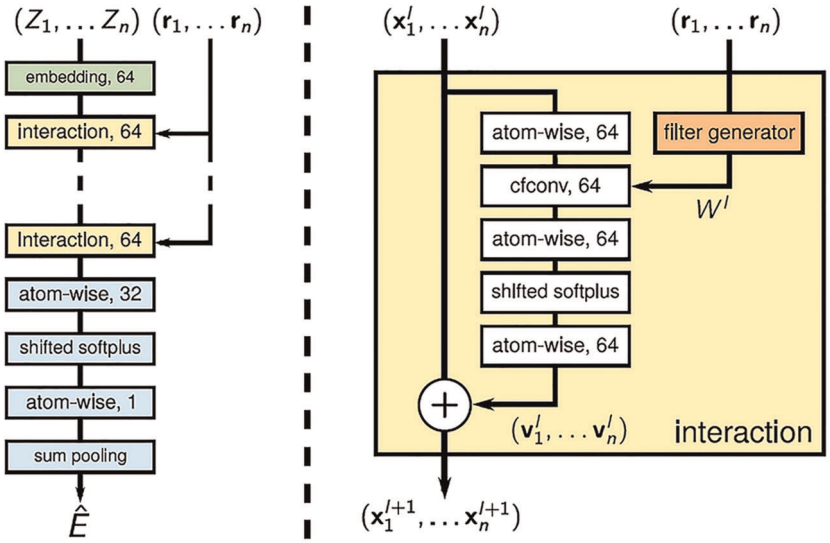

FIGURE 4 Architecture of the SchNet (left) and the interaction blocks (right) in SchNet. Reproduced with permission.[49] Copyright 2018, AIP Publishing

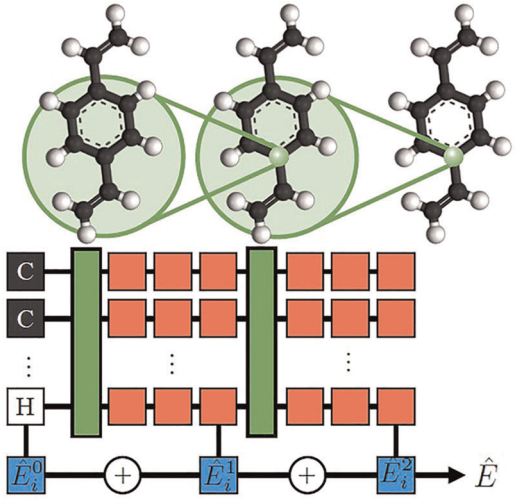

FIGURE 5 Representation (upper) and architecture (lower) of HIP-NN. Reproduced with permission.[52] Copyright 2018, AIP Publishing

3.5 | Hierarchically interacting particle neural network (HIP-NN)

HIP-NN[52] is an NN-based method to model the potential energy from the ab initio quantum mechanical calculation. The total energy E is a sum of local contributions of each atom, $E = \sum_i E_i$. Moreover, the potential energy is also decomposed as a sum over n hierarchical terms generated from a corresponding NN,[52]

$$\hat{E}_i = \sum_{n=0}^{N_{\text{interaction}}} \hat{E}_i^n. \tag{9}$$

Figure 5 shows the process of HIP-NN predicting molecular energy and other properties. The molecular configuration consisted of the number of atoms Z_i and the relative atomic coordinates r_i. Also, only the atoms inside the cutoff radius (in the green circle) are considered, $r_{ij} < R_c$.

This molecular configuration is then input to the HIP-NN, and flows from left to right to generate consecutive sets of atomic features. The red boxes denote on-site layers, and consist of fully connected layers. The activation function of the on-site layers is the softplus function[53]

$$f(x) = \log(1 + e^x). \tag{10}$$

However, at the last layer, ResNet[54] transformation is applied to obtain the final atomic features[52]

$$z_{i,a}^{l+1} = \sum_b (\tilde{W}_{ab}^l \bar{z}_b^{l+1} + \tilde{M}_{ab}^l z_{i,b}^l) + \tilde{B}_a^l, \tag{11}$$

where $\tilde{W}_{ab}^l$, $\tilde{M}_{ab}^l$, and $\tilde{B}_a^l$ are all learned parameters.

The green boxes denote interaction layers, in which the spatial sensitivities are selected to collect information from the atoms in the cutoff radius[52]

$$\varphi_c(r) = \begin{cases} \left[\cos\left(\frac{\pi}{2}\frac{r}{R_c}\right)\right]^2 & r \le R_c \\ 0 & r > R_c \end{cases}. \tag{12}$$

In blue boxes, linear regression is adopted to model the relationship between the atomic features and the local hierarchical energies. The total energy from the final output of HIP-NN is a sum of contributions of all the local hierarchical energies.

The mean absolute error (MAE) and the root-mean-square error (RMSE)[55] are considered simultaneously in a loss function,[52]

$$\mathcal{L} = \frac{1}{\sigma_E}(\text{MAE} + \text{RMSE}) + \mathcal{L}_{L_2} + \mathcal{L}_R, \tag{13}$$

where σ_E denotes the standard deviation of the total energy. The two regularization terms $\mathcal{L}_{L_2}$ and $\mathcal{L}_R$ are applied to avoid overfitting and promote the hierarchical R of energy contributions, respectively. HIP-NN also uses the Adam algorithm to tune the learned parameters.

HIP-NN is then performed on the QM9[56] data set to predict the potential energy of the organic molecules. The MAEs of prediction are well under 1 kcal/mol, which are less than the SchNet and DTNN, achieving state-of-the-art accuracy.[52]

3.6 | Fast learning of atomistic rare events (FLARE)

FLARE[57] is an "on the fly" (also called active learning[58]) ML interatomic potential method based on the GP algorithm. The GP interatomic potential is trained while the DFT-based MD simulation is running. The main feature of FLARE is that the prediction is not, like most ML methods, uncertain energy or forces, but the distributions.

FLARE uses the interatomic distances between atoms $d_{i,i_1,\ldots,i_{n-1}}$ to preserve the atomic information. As shown in Figure 6A, the local energy E_i only considers the contributions of atoms within a cutoff distance $R_{\text{cut}}^{(n)}$ from atom i. Figure 6B shows that a DFT-based MD simulation is called at the beginning of the algorithm with an arbitrary structure. Then, a subset of atoms with the highest uncertainty is chosen as input to train the GP model. After that, the forces on all atoms can be predicted by the GP model as an MD step. Whether to use the predictions by the GP model or to perform a DFT calculation as the next MD step will be decided by the epistemic uncertainties. The epistemic uncertainties measure the dissimilarity between the environment of the atoms in the current step and the local environments in the training set. The predictions of the GP model will be accepted as the next MD step if the epistemic uncertainty of any atom does not exceed the current noise uncertainty. Otherwise, DFT will be called instead and the environment of the atoms with the largest epistemic uncertainties will be added to the training set.

The number of DFT calls will decrease with the above process. After training on a sufficient amount of data, the DFT calculation is no longer needed. The trained GP model will be mapped to highly efficient tabulated force fields.

Most of the machine-learning-based interatomic potential approaches consist of three main ingredients. The first includes training data, which provide the structure information (atom coordinates, atom types, atom numbers, etc.) and the reference potential energies. The reference potential is generated from first principles calculations, using quantum electronic structure software packages (VAPS, Quantum ESPRESSO, CP2K, LAMMPS, etc.). The second is a representation of the atomic structure, which can maintain the translation, rotation, and permutation invariance. The third is a regression algorithm, which is used to determine the relationship between the representation and the potential energy.

The regression model will output the potential energy, once the input structures in the training data are fed into the regression model. The loss function will be designed based on the difference between the output potential energy and the training data. Then, the interatomic potentials will be obtained by optimizing the hyperparameters of the regression model using optimizers, such as gradient descent (GD), Nesterov accelerated gradient (NAG), AdaGrad, Adam, and so forth.

4 | APPLICATIONS

The applications of ML-generated interatomic potentials in selected problems in materials science have been reported: alloys, phase change materials (PCMs), and thermal performance calculations.

4.1 | Single element

A single element is the simplest form for most elements. Although it could be predicted quite accurately by DFT, some complex processes such as phase transition and surface/interface properties are still beyond the capacity of quantum methods and require the interatomic

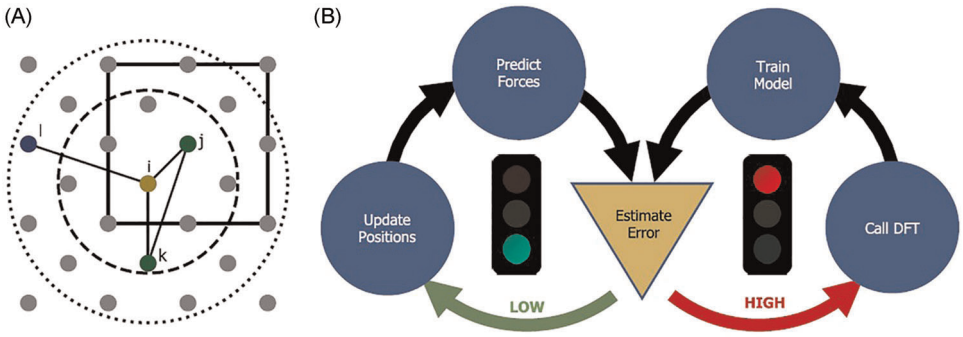

FIGURE 6 Overview of the on-the-fly learning algorithm.[57] (A) Representation of the local environment of each atom. (B) Automatic training process of FLARE during MD. Reproduced under terms of the CC-BY license.[57] Copyright 2020, The Authors, published by Springer

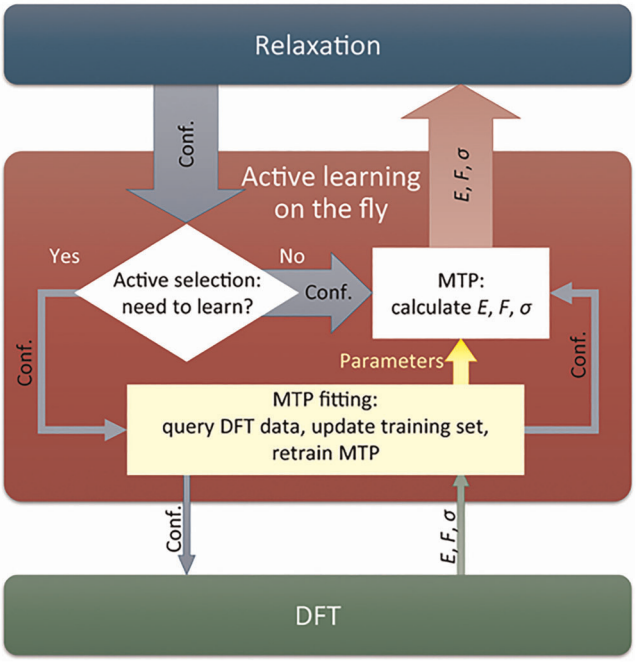

FIGURE 7 Scheme of ML interatomic potentials actively learning on-the-fly. An active learning approach determines the degree of extrapolation for each sampling configuration. If the extrapolation grade is sufficiently high, then the configuration is learned; otherwise, the energy, forces, and stresses are calculated using the Moment Tensor Potentials (MTPs) model. Reproduced with permission.[63] Copyright 2019, American Physical Society

potential for classical MD. For materials such as Al, Mo, Ti, and U, phonon properties are evaluated using the temperature-dependent effective potential method.[59–62] Automatic creation of an interatomic interaction model from scratch is performed to predict the crystal structure of carbon, the high-pressure phase of sodium, and the allotrope of boron[63] (see Figure 7). The generation processing of the interatomic potential of the bulk and the multilayer structure of hexagonal boron nitride use the GAP ML algorithm.[64] An ML data set construction method in which the data set is regularly retrained is used on elemental aluminum (ANI-Al).[65] The physically informed neural network (PINN) method is used to obtain the potential function for Al and accurately predicts some physical properties.[66] The grain boundary energies of fcc element metals (e.g., Al, Cu, Ag, Au, Pd, and Pt) are predicted by ML potential.[67] The semi-empirical force field potential for titanium is trained by an ANN.[68] A set of ML interatomic potentials for bcc structure transition metals Mo, Nb, Ta, V, and W using the GAP framework is used to study melting and liquid properties.[69] Several ML methods suitable for high-dimensional systems used in the construction of PESs are also summarized.[70]

4.2 | Alloy

Alloying is one of the most commonly used methods to tune the properties of materials. The complex local environment of alloys hinders the traditional fitting generation of interatomic potential. For example, the fabrication of a high-entropy alloy, usually with at least 5 different kinds of metals, requires a much more complicated interatomic potential. The 5 alloys require at least 10 combinations of pair potentials, which is far beyond the capacity of the fitting. Therefore, the application of ML for alloys was one of the first in the field. Several groups have proposed different methods for various high-entropy alloy systems.

It has been reported that linearized pairwise and angular-dependent MLIPs are robustness for 31 elemental metals, and angular-dependent descriptors are important for transition metals.[71] The ML framework is applied to predict the segregation energy of more than 250 metal-based binary alloys.[72]

Deep potential generators (DP-GENs) can produce potential energy surface (PES) models in Mg, Al, and Mg–Al alloys.[73] The potential generated by the NN-ML approach of the Pd–Si system can describe both liquid and crystal structures.[74] The Gaussian approximate potential framework is used to train the potential of the W_xMo_{1-x} random alloys.[75]

A multi-fidelity (MF) ML framework leveraging GP calculated the bulk modulus of an aluminum–niobium–titanium (Al–Nb–Ti) ternary alloy composition space.[76] The moment tensor potentials (MTPs) are used to predict the high-temperature elastic properties of GUM Ti-based alloys.[77] Some bulk, interface, and defect characteristics are calculated in the range from low temperatures to those close to the melting point of Ti-based alloys.[78]

The ML framework is also applicable for high-entropy alloy systems. The short-range sequencing of an equiatomic Co–Cr–Fe–Ni high-entropy alloy is calculated,[79] as shown in Figure 8. For multi-principal element alloys (MPEAs) and high-entropy alloys (HEAs), the calculation results of local lattice distortion affect the elastic properties.[82] The phase stability of a bcc Al–Nb–TiV refractory high-entropy alloys is also studied.[83]

4.3 | Phase change materials

PCMs are extensively used in thermal management and dynamics, especially microelectronic storage devices such as optical disc and non-volatile memory cells. Although some transition-metal oxides such as VO_2 are widely used for thermal/optical applications, most PCMs are chalcogenide with three to four elements and amorphous structures, which is far beyond the calculation capacity of DFT. Unlike alloying, where the metallic bonding can be treated as a free-electron approximation and the crystal structure is simple and clear, the covalent bonding in chalcogenides is orientation dependent and can form a complex network. Therefore, a more precise interatomic potential is necessary.

The applications and potential of use of ML technology to determine with the functional characteristics of PCMs are summarized in this study.[84] The phase transition behavior between incommensurate host-guest structures (KIII) and close-packed fcc of alkali metal potassium is reported to occur in a complex and diffusive

manner.[85] The electronic structures of the medium gap state of the prototypical phase-change alloy, $Ge_2Sb_2Te_5$, were identified and characterized by the ML potential,[86] as shown in Figure 9. The phase transition and plastic properties for iron were researched by a classical potential trained by the GAP framework.[87]

4.4 | Amorphous material

An ML model constructed using the GAP methodology is obtained for amorphous carbon.[88] Deformation of the representative volume element of amorphous carbon is achieved by melt quenching using bond order and ML of the interaction potential between atoms.[89]

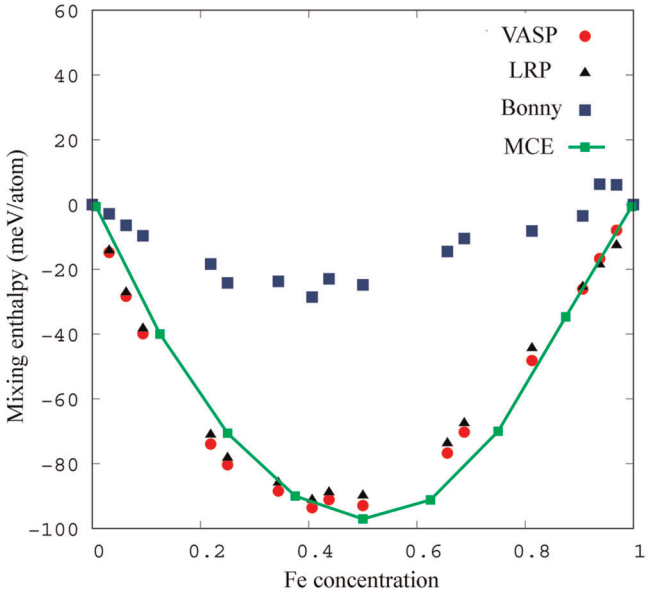

FIGURE 8 Composition dependence of the mixing enthalpy of FeNi. The comparison results are obtained from DFT, LRP, and an EAM[80] potential and magnetic cluster expansion (MCE).[81] Reproduced with permission.[79] Copyright 2019, Elsevier

The structure and elastic properties of low-speed quenched amorphous carbon are investigated.[90]

Different cooling rates influence short- and intermediate-range orders of the a-Si network, which is similar to the experimental annealing process.[91] An ANN model for silicon-carbon systems and ceramic matrix composites (CMCs) based on silicon carbide is proposed.[92] MTPs combined with the charge-equilibration (MTP+QEq) cannot improve MTP potentials alone for alpha-quartz SiO_2-based structures.[93] The method combining ab initio MD simulations and Bayesian optimization agrees well with the ab initio simulations by considering the example of glassy silica.[94]

A statistical learning model was trained using the least absolute shrinkage and selection operator with a gradient boost machine (GBM-LASSO), which only needs to be trained using a data set consisting of only binary and ternary glass samples,[95] as shown in Figure 10. The local structural order in $Al_{90}Tb_{10}$ metallic glass is studied using an interatomic potential trained by a deep ANN.[96] The effectiveness of the potential for ZrB_2 ultra-high-temperature ceramic (UHTC) materials is extended from room temperature to the ultra-high-temperature region.[97]

4.5 | Heat transfer

Heat transfer is one of the most important processes for advanced manufacturing. However, it is difficult to accurately measure the heat transfer experimentally due to the limitation of detection. MD is widely used to predict the thermal response of various structures with different boundary conditions. The time and length scales associated with heat transfer phenomena are usually far beyond the scope of ab initio calculations.[98] ML of interatomic potentials can be used as a solution for this challenge.[99] Meanwhile, the heat transfer process is determined by the phonon vibration of the lattice, which requires an accurate description of the interatomic force. Using ML-based interatomic potentials to analyse the heat transfer

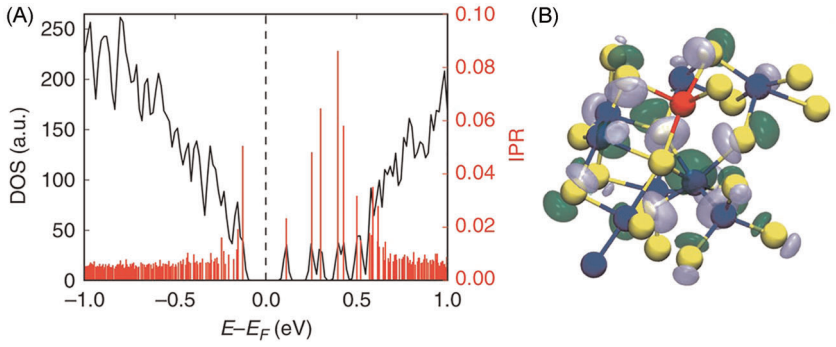

FIGURE 9 A 900-atom amorphous $Ge_2Sb_2Te_5$ model is generated by classical molecular dynamics simulations with a Gaussian approximation potential, and then the electronic structure is calculated. (A) Total electronic density of states (DOS) and Kohn–Sham orbitals. (B) Molecular orbital and atomic structure associated with the medium gap electronic state. Reproduced under terms of the CC-BY license.[86] Copyright 2019, The Authors, published by Springer Nature

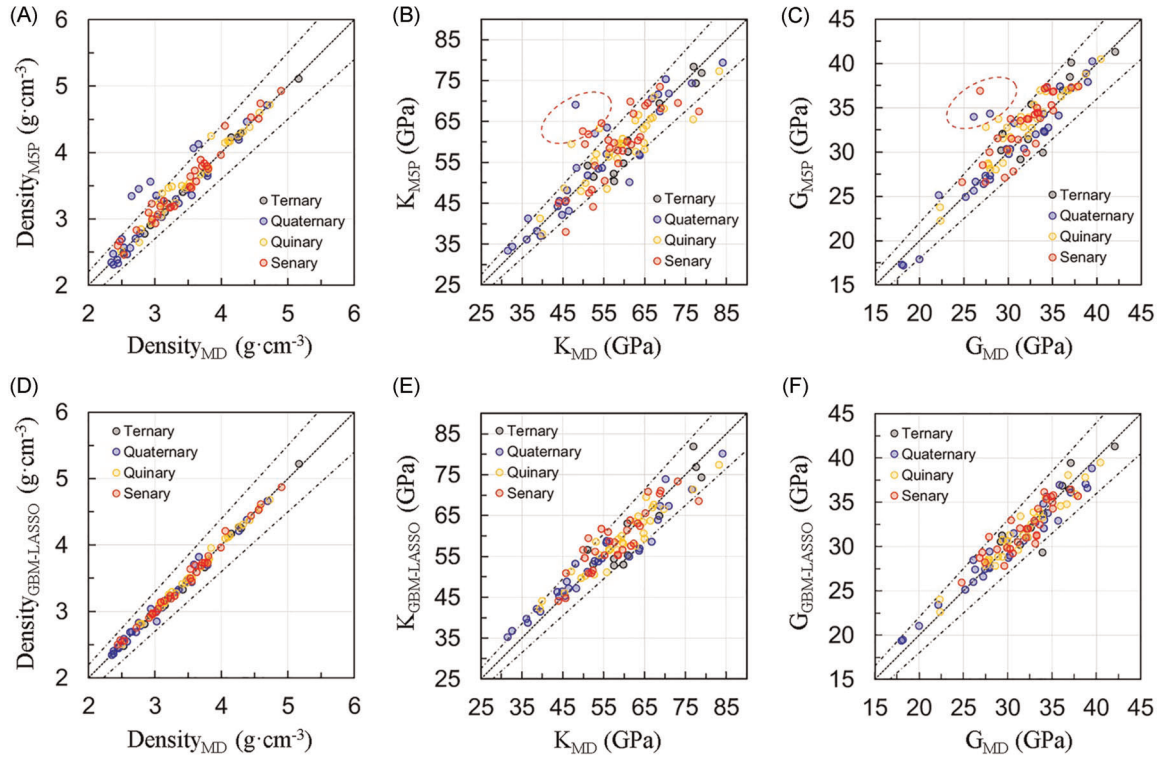

FIGURE 10 Predictions of density bulk and shear modulus of SiO_2-based glasses by machine learning. (A–C) Results from a statistical learning model training by implementing the least absolute shrinkage and selection operator with GBM-LASSO. (D–F) Results from a traditional decision tree-based model (M5P). 101 randomly chosen ternary, quaternary, and senary systems are used to evaluate the prediction capabilities of the machine-learning model. Reproduced under terms of the CC-BY license.[95] Copyright 2020, The Authors, published by Springer Nature

performance has been widely used, including crystalline and amorphous silicon,[100] GeTe nanowires,[101] CaF_2,[102] $MoS_{2(1-x)}Se_{2x}$ alloys,[103] and beta-Ga_2O_3.[104] MTPs are used to calculate vacancy diffusivity in Al, Mo, and Si.[105]

4.6 | Defect

In most cases, the mechanical and electronic properties are limited or have defects. For example, electronic properties such as conductivity can be tuned by defects or doping by orders of magnitude. Atomic simulation serves as a powerful tool to track the motion of dopants or defects and as an essential tool to assess thermodynamic parameters such as diffusion barrier and formation energy. For the study of the basic mechanisms of deformation and failure of metals and alloys, the interatomic potential is crucial[106] because related defects (dislocations, cracks, etc.) are far beyond the scope of first-principles research.[107]

SNAP is used to calculate the generalized stacking fault energies (GSFEs) and Peierls stresses of screw and edge dislocations in four bcc refractory metals.[108] A universal strategy for determining crystal defects based on statistical outlier detection is proposed.[106] The potential from the Behler–Parrinello NN framework is used to predict metallurgical critical dislocations and cracks for pure hcp magnesium.[107]

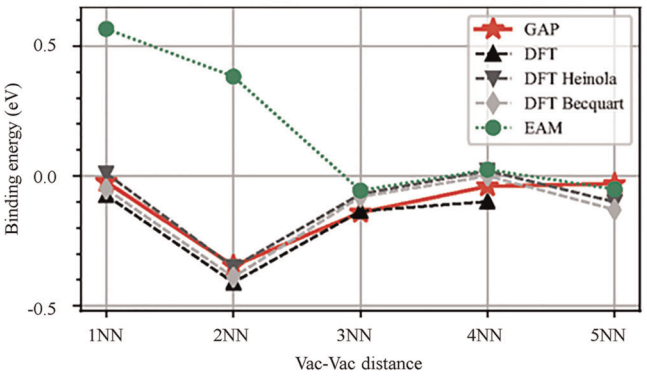

FIGURE 11 An ML interatomic potential for tungsten using the GAP framework is used to calculate the binding energies of a divacancy in bcc W at different nearest-neighbor separations, which compares with DFT from References [111,112], our own DFT, and EAM potentials. Reproduced with permission[109,110] Copyright 2020, American Physical Society. DFT, density functional theory; EAM, embedded atom method; GAP, Gaussian approximation potential

4.7 | Damage

Damage is the key process governing a material's mechanical properties, such as crack and creep, and requires a large space and time scale for simulation. The ML interatomic potential of

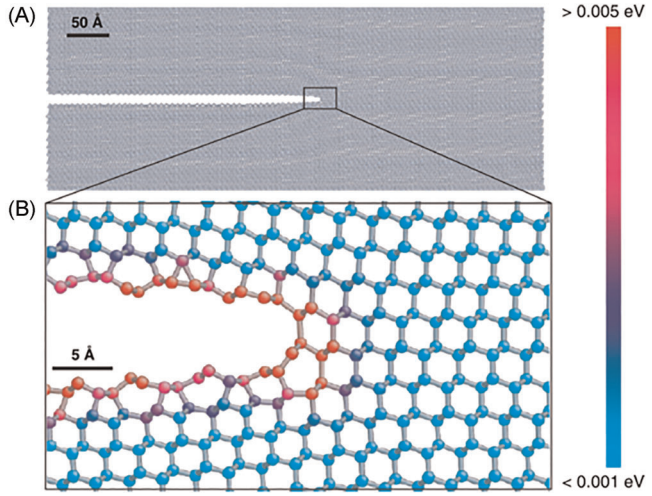

FIGURE 12 (A) Snapshot from a quasistatic simulation of fracture in the Si (111) [$11\bar{2}0$] cleavage system calculated by GAP. The calculated strain energy release rate is $G = 5.13\ J/m^2$, similar to the results of DFT-based hybrid simulations. (B) The close-up of the crack tip, which shows a crack-tip reconstructions. The atoms with different colors represnets the error between strain energy predicted by GAP and calculated by DFT. Reproduced under terms of the CC-BY license.[119] Copyright 2020, The Authors, published by American Physical Society

tungsten is calculated by the GAP framework, which enables simulation of radiation damage in tungsten,[109,110] as shown in Figure 11. Similarly, the evolution of silicon primary damage and collision cascades are also analyzed.[113] Neutron bombardment is studied in a primary knock-on atom (PKA) of crystalline molybdenum material samples by classical MD using a potential based on the GAP framework.[114] A tungsten-beryllium potential is developed by SNAP and used to simulate high-energy Be atom implantation onto the (001) surface of solid tungsten.[115]

Extensive atomic simulations were performed on the deposition and growth of amorphous carbon (a-C) thin films using a GAP model to describe the interaction between atoms.[116] Neural-network potentials (NNPs) are applied to simulation early-stage nucleation and growth of the Al-Cu system.[117] A charge-transfer ionic potential (CTIP) model for the Cu/Hf/O alloy system can accurately predict the surface properties of both oxides and metals.[118]

The GAP of silicon can be calculated for dynamic brittle fracture,[119] as shown in Figure 12. ML potential applies to the formation energy of the interface phase.[120]

5 | CONCLUSIONS

The use of ML in the generation of interatomic potential can considerably increase the efficiency of MD simulation and drastically transform advanced manufacturing by incorporating atomic simulation early into the design process. We have discussed a few promising new approaches based on descriptors and their potential applications in various manufacturing processes. Most of the methods discussed here are well presented with ready packages for testing in well-developed MD packages such as LAMMPS. This is important for their further application in experimental comparisons.

There are several promising directions for future development. The faster interatomic potential models will increase the length and time scale of MD simulation, which is vital for simulation in advanced manufacturing. Automatic parameterization can also benefit the simulation of complex systems such as HEAs with 10 different elements and new chemical compositions.

ACKNOWLEDGMENTS

This study was supported by the Wuhan University Junior Faculty Research (2042019KF0003), the National Natural Science Foundation of China (51727901, U1501241, and 62174122), the National Key R&D Program of China (2017YFB1103904), and the Hubei Provincial Natural Science Foundation of China (2020CFA032). The numerical calculations in this article have been performed on the supercomputing system in the Supercomputing Center of Wuhan University.

CONFLICT OF INTEREST

The authors declare that there are no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding authors upon reasonable request.

ORCID

Yuzheng Guo https://orcid.org/0000-0001-9224-3816

REFERENCES

1. Fogel DB. *Evolving artificial intelligence*. PhD thesis, University of California; 1993.
2. Jordan MI. *Artificial intelligence—the revolution hasn't happened yet*. Medium; 2018.
3. Minsky M. Steps toward artificial intelligence. *Proc IRE*. 1961;49(1):8-30.
4. Jordan MI, Mitchell TM. Machine learning: trends, perspectives, and prospects. *Science*. 2015;349(6245):255-260.
5. Michie D. "Memo" functions and machine learning. *Nature*. 1968;218(5136):19-22.
6. Sun Y, Wang X, Tang X. Protective effects of ginsenoside Rg1 on astrocytes and cerebral ischemic-reperfusion mice. *Biol Pharm Bull*. 2014;37:1891-1898.
7. Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. *IEEE Trans Pattern Anal Mach Intell*. 2017, 39:640-651.
8. Deo RC. Machine learning in medicine. *Circulation*. 2015;132(20):1920-1930.
9. Tsugawa S. Vision-based vehicles in Japan: machine vision systems and driving control systems. *IEEE Trans Ind Electron*. 1994;41(4):398-405.

10. Segler MH, Preuss M, Waller MP. Planning chemical syntheses with deep neural networks and symbolic AI. *Nature*. 2018; 555(7698):604-610.
11. Raccuglia P, Elbert KC, Adler PDF, et al. Machine-learning-assisted materials discovery using failed experiments. *Nature*. 2016;533(7601): 73-76.
12. Feurer M, Klein A, Eggensperger K, Springenberg JT, Blum M, Hutter F. Auto-sklearn: efficient and robust automated machine learning. In: *Automated Machine Learning*. Neural information processing systems; 2019:113-134.
13. Cook RD. Detection of influential observation in linear regression. *Technometrics*. 1977;19(1):15-18.
14. Zou H. The adaptive lasso and its oracle properties. *J Am Stat Assoc*. 2006;101(476):1418-1429.
15. Zou H, Hastie T. Regularization and variable selection via the elastic net. *J R Stat Soc: Ser B (Stat Methodol)*. 2005;67(2): 301-320.
16. Wan X, Zhang Z, Yu W, Guo Y. A density-functional-theory-based and machine-learning-accelerated hybrid method for intricate system catalysis. *Mater Rep Energy*. 2021;1(3):100046.
17. John R. Boundary modification for kernel regression. *Commun Stat Theory Methods*. 1984;13(7):893-900.
18. Smola AJ, Schölkopf B. A tutorial on support vector regression. *Stat Comput*. 2004;14(3):199-222.
19. Quinonero-Candela J, Rasmussen CE. A unifying view of sparse approximate Gaussian process regression. *J Mach Learn Res*. 2005; 6:1939-1959.
20. Zhang Y, Duchi J, Wainwright M. Divide and conquer kernel ridge regression: a distributed algorithm with minimax optimal rates. *J Mach Learn Res*. 2015;16(1):3299-3340.
21. Mack Y-P, Silverman BW. Weak and strong uniform consistency of kernel regression estimates. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*. 1982;61(3):405-415.
22. Derville S, Torres LG, Iovan C, Garrigue C. Finding the right fit: comparative cetacean distribution models using multiple data sources and statistical approaches. *Divers Distrib*. 2018;24(11):1657-1673.
23. Quinlan JR. Induction of decision trees. *Mach Learn* 1986;1(1): 81-106.
24. Devroye L, Gyorfi L, Krzyzak A, Lugosi G. On the strong universal consistency of nearest neighbor regression function estimates. *Ann Stat*. 1994;22:1371-1385.
25. Liaw A, Wiener M. Classification and regression by randomForest. *R News*. 2002;2(3):18-22.
26. Geurts P, Ernst D, Wehenkel L. Extremely randomized trees. *Mach Learn*. 2006;63(1):3-42.
27. Safavian SR, Landgrebe D. A survey of decision tree classifier methodology. *IEEE Trans Syst Man Cybern*. 1991;21(3):660-674.
28. Vayssières MP, Plant RE, Allen-Diaz BH. Classification trees: an alternative non-parametric approach for predicting species distributions. *J Veg Sci*. 2000;11(5):679-694.
29. Friedman JH. Greedy function approximation: a gradient boosting machine. *Ann Stat*. 2001;29:1189-1232.
30. Friedman JH. Stochastic gradient boosting. *Comput Stat Data Anal*. 2002;38(4):367-378.
31. Crick F. The recent excitement about neural networks. *Nature*. 1989;337(6203):129-132.
32. Sanger TD. Optimal unsupervised learning in a single-layer linear feedforward neural network. *Neural Netw*. 1989;2(6): 459-473.
33. Leshno M, Lin VY, Pinkus A, Schocken S. Multilayer feedforward networks with a nonpolynomial activation function can approximate any function. *Neural Netw*. 1993;6(6):861-867.
34. Mueller T, Hernandez A, Wang C. Machine learning for interatomic potential models. *J Chem Phys*. 2020;152(5):050902.
35. Deringer VL, Caro MA, Csanyi G. Machine learning interatomic potentials as emerging tools for materials science. *Adv Mater*. 2019;31(46):1902765.
36. Bartok AP, Payne MC, Kondor R, Csanyi G. Gaussian approximation potentials: the accuracy of quantum mechanics, without the electrons. *Phys Rev Lett*. 2010;104(13):136403.
37. Mattsson TR, Lane JMD, Cochrane KR, et al. First-principles and classical molecular dynamics simulation of shocked polymers. *Phys Rev B Condens Matter Mater Phys* 2010;81(5):054103.
38. Varshalovich DA, Moskalev AN, Khersonskii VKM. *Quantum Theory of Angular Momentum*. World Scientific; 1988.
39. Quinonero-Candela J, Rasmussen CE, Williams CK. Approximation methods for Gaussian process regression. In: *Large-Scale Kernel Machines*. MIT Press; 2007:203-223.
40. MacKay DJ, Mac Kay DJ. *Information Theory, Inference and Learning Algorithms*. Cambridge University Press; 2003.
41. Bartók AP, Kondor R, Csányi G. On representing chemical environments. *Phys Rev B Condens Matter Mater Phys*. 2013;87(18): 184115.
42. Yang L, Dacek S, Ceder G. Proposed definition of crystal substructure and substructural similarity. *Phys Rev B Condens Matter Mater Phys*. 2014;90(5):054102.
43. Thompson AP, Swiler LP, Trott CR, Foiles SM, Tucker GJ. Spectral neighbor analysis method for automated generation of quantum-accurate interatomic potentials. *J Comput Phys*. 2015;285:316-330.
44. Eldred MS, Giunta AA, van Bloemen Waanders BG, et al. DAKOTA, a multilevel parallel object-oriented framework for design optimization, parameter estimation, uncertainty quantification, and sensitivity analysis version 3.0 developers Manual (title change from electronic posting). 2002.
45. Zhang L, Han J, Wang H, Car REW. Deep potential molecular dynamics: a scalable model with the accuracy of quantum mechanics. *Phys Rev Lett*. 2018;120(14):143001.
46. Kim KG. Book review: deep learning. *Healthc Inform Res*. 2016; 22(4):351.
47. Bengio Y, Lamblin P, Popovici D, Larochelle H. Greedy layer-wise training of deep networks. *Adv Neural Inf Process Syst*. 2007;19:153.
48. Kingma DP, Ba J. Adam: a method for stochastic optimization. *arXiv*. preprint arXiv:1412.6980. 2014.
49. Schütt KT, Sauceda HE, Kindermans P-J, Tkatchenko A, Müller K-R. SchNet—a deep learning architecture for molecules and materials. *J Chem Phys*. 2018;148(24):241722.
50. Schutt KT, Arbabzadah F, Chmiela S, Muller KR, Tkatchenko A. Quantum-chemical insights from deep tensor neural networks. *Nat Commun*. 2017;8:13890.
51. Chollet F. Xception: deep learning with depthwise separable convolutions. 2017:1251-1258.
52. Lubbers N, Smith JS, Barros K. Hierarchical modeling of molecular energies using a deep neural network. *J Chem Phys*. 2018;148(24): 241715.
53. Dugas C, Bengio Y, Bélisle F, Nadeau C, Garcia R. Incorporating second-order functional knowledge for better option pricing. *Adv Neural Inf Process Syst*. 2001:472-478.
54. Targ S, Almeida D, Lyman K. Resnet in resnet: generalizing residual architectures. *arXiv*. preprint arXiv:1603.08029. 2016.
55. Willmott CJ, Matsuura K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Clim Res*. 2005;30(1):79-82.
56. Ramakrishnan R, Dral PO, Rupp M, von Lilienfeld OA. Quantum chemistry structures and properties of 134 kilo molecules. *Sci Data*. 2014;1:140022.
57. Vandermause J, Torrisi SB, Batzner S, et al. On-the-fly active learning of interpretable Bayesian force fields for atomistic rare events. *npj Comput Mater*. 2020;6(1):20.

58. Settles B. *Active learning literature survey*. Computer Sciences Technical Report 1648, University of Wisconsin–Madison; 2009.
59. Li XG, Hu CZ, Chen C, Deng Z, Luo J, Ong SP. Quantum-accurate spectral neighbor analysis potential models for Ni-Mo binary alloys and fcc metals. *Phys Rev B Condens Matter Mater Phys*. 2018;98(9):094104.
60. Daksha CM, Yeon J, Chowdhury SC, Gillespie JW. Automated ReaxFF parametrization using machine learning. *Comput Mater Sci*. 2021;187(9):110107.
61. Guo F, Wen YS, Feng SQ, et al. Intelligent-ReaxFF: evaluating the reactive force field parameters with machine learning. *Comput Mater Sci*. 2020;172(11):109393.
62. Ladygin VV, Korotaev PY, Yanilkin AV, Shapeev AV. Lattice dynamics simulation using machine learning interatomic potentials. *Comput Mater Sci*. 2020;172(8):109333.
63. Podryabinkin EV, Tikhonov EV, Shapeev AV, Oganov AR. Accelerating crystal structure prediction by machine-learning interatomic potentials with active learning. *Phys Rev B Condens Matter Mater Phys*. 2019;99(6):064114.
64. Thiemann FL, Rowe P, Muller EA, Michaelides A. Machine learning potential for hexagonal boron nitride applied to thermally and mechanically induced rippling. *J Phys Chem C*. 2020;124(40):22278-22290.
65. Smith JS, Nebgen B, Mathew N, et al. Automated discovery of a robust interatomic potential for aluminum. *Nat Commun*. 2021;12(1):1257.
66. Pun GPP, Yamakov V, Hickman J, Glaessgen EH, Mishin Y. Development of a general-purpose machine-learning interatomic potential for aluminum by the physically informed neural network method. *Phys Rev Mater*. 2020;4(11):113807.
67. Nishiyama T, Seko A, Tanaka I. Application of machine learning potentials to predict grain boundary properties in fcc elemental metals. *Phys Rev Mater*. 2020;4(12):123607.
68. Dickel D, Francis DK, Barrett CD. Neural network aided development of a semi-empirical interatomic potential for titanium. *Comput Mater Sci*. 2020;171(8):109157.
69. Byggmastar J, Nordlund K, Djurabekova F. Gaussian approximation potentials for body-centered-cubic transition metals. *Phys Rev Mater*. 2020;4(9):093802.
70. Behler J. Perspective: machine learning potentials for atomistic simulations. *J Chem Phys*. 2016;145(17):170901.
71. Takahashi A, Seko A, Tanaka I. Linearized machine-learning interatomic potentials for non-magnetic elemental metals: limitation of pairwise descriptors and trend of predictive power. *J Chem Phys*. 2018;148(23):234106.
72. Wagih M, Larsen PM, Schuh CA. Learning grain boundary segregation energy spectra in polycrystals. *Nat Commun*. 2020;11(1):6376.
73. Zhang LF, Lin DY, Wang H, Car R, E WN. Active learning of uniformly accurate interatomic potentials for materials simulation. *Phys Rev Mater*. 2019;3(2):023804.
74. Wen T, Wang CZ, Kramer MJ, et al. Development of a deep machine learning interatomic potential for metalloid-containing Pd-Si compounds. *Phys Rev B Condens Matter Mater Phys*. 2019;100(17):174101.
75. Nikoulis G, Byggmastar J, Kioseoglou J, Nordlund K, Djurabekova F. Machine-learning interatomic potential for W-Mo alloys. *J Phys Condens Matter*. 2021;33(31):315403.
76. Tran A, Tranchida J, Wildey T, Thompson AP. Multi-fidelity machine-learning with uncertainty quantification and Bayesian optimization for materials design: application to ternary random alloys. *J Chem Phys*. 2020;153(7):074705.
77. Shapeev AV, Podryabinkin EV, Gubaev K, Tasnadi F, Abrikosov IA. Elinvar effect inβ-Ti simulated by on-the-fly trained moment tensor potential. *New J Phys*. 2020;22(11):113005.
78. Lopanitsyna N, Ben Mahmoud C, Ceriotti M. Finite-temperature materials modeling from the quantum nuclei to the hot electron regime. *Phys Rev Mater*. 2021;5(4):043802.
79. Meshkov EA, Novoselov II, Shapeev AV, Yanilkin AV. Sublattice formation in CoCrFeNi high-entropy alloy. *Intermetallics*. 2019;112(7):106542.
80. Bonny G, Pasianot R, Malerba L. Fe–Ni many-body potential for metallurgical applications. *Modell Simul Mater Sci Eng*. 2009;17(2):025010.
81. Lavrentiev MY, Wróbel J, Nguyen-Manh D, Dudarev S. Magnetic and thermodynamic properties of face-centered cubic Fe–Ni alloys. *Phys Chem Chem Phys* 2014;16(30):16049-16059.
82. Jafary-Zadeh M, Khoo KH, Laskowski R, Branicio PS, Shapeev AV. Applying a machine learning interatomic potential to unravel the effects of local lattice distortion on the elastic properties of multi-principal element alloys. *J Alloy Compd*. 2019;803:1054-1062.
83. Kormann F, Kostiuchenko T, Shapeev A, Neugebauer J. B2 ordering in body-centered-cubic AlNbTiV refractory high-entropy alloys. *Phys Rev Mater*. 2021;5(5):053803.
84. Sosso GC, Bernasconi M. Harnessing machine learning potentials to understand the functional properties of phase-change materials. *MRS Bull*. 2019;44(9):705-709.
85. Zhao L, Zong HX, Ding XD, Sun J, Ackland GJ. Commensurate-incommensurate phase transition of dense potassium simulated by machine-learned interatomic potential. *Phys Rev B Condens Matter Mater Phys*. 2019;100(22):220101.
86. Konstantinou K, Mocanu FC, Lee TN, Elliott SR. Revealing the intrinsic nature of the mid-gap defects in amorphous $Ge_2Sb_2Te_5$. *Nat Commun*. 2019;10:3065.
87. Maillet JB, Denoual C, Csanyi G. Machine-learning based potential for iron: plasticity and phase transition. In: Chau R, Germann TC, Lane JMD, Brown EN, Eggert JH, Knudson MD, eds. *Shock Compression of Condensed Matter—2017. Amer Inst Physics*; 2018. AIP Conference Proceedings.
88. Rowe P, Deringer VL, Gasparotto P, Csanyi G, Michaelides A. An accurate and transferable machine learning potential for carbon. *J Chem Phys*. 2020;153(3):034702.
89. Jana R, von Lautz J, Khosrownejad SM, Andrews WB, Moseler M, Pastewka L. Constitutive relations for plasticity of amorphous carbon. *J Phys Mater*. 2020;3(3):035005.
90. Jana R, Savio D, Deringer VL, Pastewka L. Structural and elastic properties of amorphous carbon from simulated quenching at low rates. *Modell Simul Mater Sci Eng*. 2019;27(8):085009.
91. Li WW, Ando Y. Dependence of a cooling rate on structural and vibrational properties of amorphous silicon: a neural network potential-based molecular dynamics study. *J Chem Phys*. 2019;151(11):114101.
92. Kubo A, Umeno Y. Machine-learning-based atomistic model analysis on high-temperature compressive creep properties of amorphous silicon carbide. *Materials*. 2021;14(7):1597.
93. Novikov IS, Shapeev AV. Improving accuracy of interatomic potentials: more physics or more data? a case study of silica. *Mater Today Commun*. 2019;18:74-80.
94. Liu H, Fu ZP, Li YP, Sabri NFA, Bauchy M. Parameterization of empirical forcefields for glassy silica using machine learning. *MRS Commun*. 2019;9(2):593-599.
95. Hu YJ, Zhao G, Zhang M, et al. Predicting densities and elastic moduli of SiO_2-based glasses by machine learning. *Npj Comput Mater*. 2020;6(1):25.
96. Tang L, Yang ZJ, Wen TQ, Ho KM, Kramer MJ, Wang CZ. Short- and medium-range orders in Al90Tb10 glass and their relation to the structures of competing crystalline phases. *Acta Mater*. 2021;204:116513.

97. Zhang YH, Lunghi A, Sanvito S. Pushing the limits of atomistic simulations towards ultra-high temperature: a machine-learning force field for ZrB2. *Acta Mater*. 2020;186:467-474.
98. Sosso GC, Deringer VL, Elliott SR, Csanyi G. Understanding the thermal properties of amorphous solids using machine-learning-based interatomic potentials. *Mol Simul* 2018;44(11):866-880.
99. Korotaev P, Novoselov I, Yanilkin A, Shapeev A. Accessing thermal conductivity of complex compounds by machine learning interatomic potentials. *Phys Rev B Condens Matter Mater Phys*. 2019;100(14):144308.
100. Qian X, Peng S, Li X, Wei Y, Yang R. Thermal conductivity modeling using machine learning potentials: application to crystalline and amorphous silicon. *Mater Today Phys*. 2019;10(7):100140.
101. Bosoni E, Campi D, Donadio D, Sosso GC, Behler J, Bernasconi M. Atomistic simulations of thermal conductivity in GeTe nanowires. *J Phys D Appl Phys*. 2020;53(5):054001.
102. Faraji S, Allaei SMV, Amsler M. Thermal conductivity of CaF_2 at high pressure. *Phys Rev B Condens Matter Mater Phys*. 2021;103(13):134301.
103. Gu XK, Zhao CY. Thermal conductivity of single-layer $MoS_{2(1-x)}Se_{2x}$ alloys from molecular dynamics simulations with a machine-learning-based interatomic potential. *Comput Mater Sci*. 2019;165:74-81.
104. Li R, Liu Z, Rohskopf A, et al. A deep neural network interatomic potential for studying thermal conductivity of β-Ga_2O_3. *Appl Phys Lett*. 2020;117(15):152102.
105. Novoselov, II, Yanilkin AV, Shapeev AV, Podryabinkin EV. Moment tensor potentials as a promising tool to study diffusion processes. *Comput Mater Sci*. 2019;164:46-56.
106. Goryaeva AM, Lapointe C, Dai CD, Deres J, Maillet JB, Marinica MC. Reinforcing materials modelling by encoding the structures of defects in crystalline solids into distortion scores. *Nat Commun*. 2020;11(1):4691.
107. Stricker M, Yin BL, Mak E, Curtin WA. Machine learning for metallurgy II. A neural-network potential for magnesium. *Phys Rev Mater*. 2020;4(10):103602.
108. Wang XW, Xu SZ, Jian WR, Li XG, Su YQ, Beyerlein IJ. Generalized stacking fault energies and Peierls stresses in refractory body-centered cubic metals from machine learning-based interatomic potentials. *Comput Mater Sci*. 2021;192(16):110364.
109. Byggmastar J, Hamedani A, Nordlund K, Djurabekova F. Machine-learning interatomic potential for radiation damage and defects in tungsten. *Phys Rev B Condens Matter Mater Phys*. 2019;100(14):144105.
110. Dominguez-Gutierrez FJ, Byggmastar J, Nordlund K, Djurabekova F, von Toussaint U. On the classification and quantification of crystal defects after energetic bombardment by machine learned molecular dynamics simulations. *Nucl Mater Energy*. 2020;22(8):100724.
111. Heinola K, Djurabekova F, Ahlgren T. On the stability and mobility of di-vacancies in tungsten. *Nucl Fusion*. 2017;58(2):026004.
112. Becquart C, Domain C. Ab initio calculations about intrinsic point defects and He in W. *Nucl Instrum Methods Phys Res Sect B Beam Interact Mater Atoms*. 2007;255(1):23-26.
113. Hamedani A, Byggmästar J, Djurabekova F, et al. Insights into the primary radiation damage of silicon by a machine learning interatomic potential. *Mater Res Lett*. 2020;8(10):364-372.
114. Dominguez-Gutierrez FJ, Byggmastar J, Nordlund K, Djurabekova F, von Toussaint U. Computational study of crystal defect formation in Mo by a machine learning molecular dynamics potential. *Modell Simul Mater Sci Eng*. 2021;29(5):055001.
115. Wood MA, Cusentino MA, Wirth BD, Thompson AP. Data-driven material models for atomistic simulation. *Phys Rev B Condens Matter Mater Phys*. 2019;99(18):184305.
116. Caro MA, Csanyi G, Laurila T, Deringer VL. Machine learning driven simulated deposition of carbon films: from low-density to diamondlike amorphous carbon. *Phys Rev B Condens Matter Mater Phys*. 2020;102(17):174201.
117. Marchand D, Jain A, Glensk A, Curtin WA. Machine learning for metallurgy I. A neural-network potential for Al-Cu. *Phys Rev Mater*. 2020;4(10):103601.
118. Sasikumar K, Chan H, Narayanan B, Sankaranarayanan S. Machine learning applied to a variable charge atomistic model for Cu/Hf binary alloy oxide heterostructures. *Chem Mat*. 2019;31(9):3089-3102.
119. Bartok AP, Kermode J, Bernstein N, Csanyi G. Machine learning a general-purpose interatomic potential for silicon. *Phys Rev X*. 2018;8(4):041048.
120. Korolev V, Mitrofanov A, Kucherinenko Y, Nevolin Y, Krotov V, Protsenko P. Accelerated modeling of interfacial phases in the Ni-Bi system with machine learning interatomic potential. *Scr Mater*. 2020;186:14-18.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Yu W, Ji C, Wan X, et al. Machine-learning-based interatomic potentials for advanced manufacturing. *Int J Mech Syst Dyn*. 2021;1:159-172. doi:10.1002/msd2.12021