Deep neural network potential for simulating hydrogen blistering in tungsten

Xiao-Yang Wang [1], Yi-Nan Wang,[2,3] Ke Xu,[4,5] Fu-Zhi Dai,[6,3] Hai-Feng Liu,[1] Guang-Hong Lu,[4,5] and Han Wang [1,7,*]

[1]*Laboratory of Computational Physics, Institute of Applied Physics and Computational Mathematics,*
Fenghao East Road 2, Beijing 100094, People's Republic of China

[2]*School of Mathematical Sciences, Peking University, No. 5 Yiheyuan Road Haidian District, Beijing 100871, People's Republic of China*

[3]*AI for Science Institute, No. 150 Chengfu Road, Haidian District, Beijing 100084, People's Republic of China*

[4]*School of Physics, Beihang University, No. 37 Xueyuan Road, Haidian District, Beijing 100191, People's Republic of China*

[5]*Beijing Key Laboratory of Advanced Nuclear Materials and Physics, Beijing 100191, People's Republic of China*

[6]*DP Technology, Beijing 100080, People's Republic of China*

[7]*HEDPS, CAPT, College of Engineering, Peking University, Beijing 100871, People's Republic of China*

Tungsten is a promising candidate for the plasma-facing material in fusion energy facilities, however, the low-energy, high-flux hydrogen plasma causes severe blistering in tungsten, which gives rise to safety concerns. By far, the formation mechanism of intragranular hydrogen blisters is still unclear. Large-scale atomistic simulations are crucial for improving the understanding, however, the available empirical interatomic potentials are mostly defective in predicting the formation of hydrogen self-clusters in tungsten, thus may lead to wrong blister formation mechanisms. In this work, we develop a machine-learning potential, DP-WH, for the tungsten-hydrogen binary system based on the Deep Potential method. We demonstrate that the DP-WH potential is able to describe, as accurately as *ab initio* calculations, the basic properties of bcc tungsten, the solute hydrogen properties in tungsten, adsorption and migration of hydrogen on tungsten free surfaces, interactions between hydrogen atoms and vacancy, dislocations, the interaction between neighboring interstitial hydrogen atoms, and the formation energy of H self-clusters. By using DP-WH, we perform nanosecond-long molecular dynamics simulations and report the formation of planar self-cluster of tetrahedral-interstitial-site hydrogen atoms normal to {001} tungsten planes at a hydrogen concentration of ≈10 at.%. This form of the H self-cluster is highly possible to be the early nucleates of the crack-shaped H blisters observed in recent experiments. The DP-WH is thus proven as a good candidate potential for the atomistic simulations to unveil the formation mechanisms of the intragranular hydrogen blisters in tungsten under the relevant working conditions.

I. INTRODUCTION

Tungsten (W) is a promising candidate for plasma-facing materials (PFMs) in the International thermonuclear experimental reactor (ITER) [1] project, due to its desirable properties such as high melting point, high thermal conductivity, and low sputtering yield. In ITER operations, great challenges are imposed on PFMs of fusion devices by the harsh service environment, including the high thermal load, and high flux [10^{22} to 10^{24} ions/(m² s)] plasma of hydrogen (H) and its isotopes (deuterium and tritium) at low energy (from tens to hundreds of eV) [2]. Interaction with H plasma causes hydrogen retention and severe blistering in W [3]. H blistering damages the material integrity and causes dramatic degradation of the mechanical properties (e.g., H-induced hardening and embrittlement [4]) and increases safety concerns. Moreover, W dust from the bursting of the blisters may contaminate the core plasma and disturb the steady discharges [5]. Thus investigating the behavior of hydrogen and its isotopes in tungsten has become one of the most important subjects in the research of PFMs. Understanding how the H blisters form is of central importance.

H blisters in W can be categorized into two types: the H intergranular blisters and the H intragranular blisters. It is well understood that the intergranular blisters, which are often observed in the form of large cavities between two grains, arise from the H accumulation at the grain boundaries [6,7]. Whereas the formation mechanisms of the intragranular H blisters (the H blisters in the grains) are far more complex and not fully understood. Up to now, three explanations for the formation of intragranular H blisters in W can be summarized from existing studies: vacancy-trapping, dislocation-trapping, and self-trapping.

Vacancy-trapping refers to the formation of intragranular blisters via the binding between H atoms and vacancy-type defects. Under working conditions with high doses of irradiation, energetic particles, such as neutrons and high-energy ions, produce a large number of vacancies in W. Vacancies are strong trapping sites for H atoms [8]. The binding between H and vacancies in W was extensively studied both experimentally [9,10] and theoretically [11–13]. And it is well-established that H-vacancy composite clusters are the nucleates of the H blister, which could develop into a

*wang_han@iapcm.ac.cn

gas-filling bubble via other mechanisms such as loop punching [8]. However, the formation of intragranular H blisters is also observed in W exposed to low-energy high-flux deuterium plasma that mimics the plasma-wall interaction conditions expected in the ITER divertor [14]. Under such conditions, the plasma energy is well below the displacement threshold of W, thus the production of vacancies by irradiation is unlikely. Moreover, the range of temperatures (300 to 1000 K) in recent experiments [2,3,14,15] only leads to a negligible equilibrium concentration of vacancies in W [14], which excludes the H trapping in the equilibrium vacancies in W. Thus the vacancy-trapping formation mechanism plays a less important role in the formation of the H blisters in W under exposure to low-energy plasma.

The dislocation-trapping refers to the formation of intragranular H blisters via the binding between H atoms and dislocation-related microstructures. It was supported by the experimental studies reporting the dislocation decorated by H bubbles [16] and the significantly enhanced H retention in W due to plastic deformation [14]. Theoretic studies also reported strong binding interaction between H and dislocations in W [17] and dislocation may also facilitate H blister nucleation via the "jog-punching" mechanism [18]. Additionally, based on the facts that W under H plasma exposure has a change of microstructure with a substantial increase of dislocation density [19,20], it was proposed by Guo *et al.* [3] and Chen *et al.* [21] that the intragranular H blisters could possibly nucleate on the binary junctions with a $\langle 100 \rangle$ burgers vector, which is formed due to the burgers vector conservation law of dislocation reaction: $\frac{1}{2}\langle 111 \rangle + \frac{1}{2}\langle \bar{1}\bar{1}1 \rangle = \langle 001 \rangle$ [22]. However, a feasible explanation is still missing for the origin of the dislocations in W under exposure to H plasma. It was suggested that the dislocations could be generated due to the thermo-load caused by the temperature gradient [15]. But it was observed in later experiments that the majority of dislocation is related to the growth of the H blister itself [2]. Thus whether the H blister nucleates on dislocations or the dislocations are generated from the H blisters is still undetermined, which put the basis of the dislocation trapping mechanisms in question.

Self-trapping mechanism refers to the spontaneous agglomeration of H atoms in neighboring interstitial sites of W, which may also serve as the formation mechanisms of the intragranular blister. Unlike the strong self-trapping of helium (He) [4] in W, the interactions between a pair of interstitial H atoms in bcc W are mostly repulsive. Thus the self-trapping of H in W was once ruled out [14]. However, a recent ab initio study by Hou [23] *et al.* reveals that the planar H self-clusters, with the H atoms filling tetrahedral interstitial sites (TIS) on a single plane, are energetically more favorable than the solution of H in TISs of W lattices. Contrary to this result, it was found by Smirnov [24] *et al.* that high concentrations of H atoms can spontaneously form planar self-clusters with H atoms located at octahedral interstitial sites (OIS) in W. There are early experiments suggesting that low-temperature H retention saturation is dictated by self-trapping [25]. However, the direct experimental proof of the self-trapping mechanisms is still lacking.

It should be noted that the atomistic details of the intragranular H blister nucleation processes are difficult or even impossible to observe in experiments [8]. Alternatively, the W-H interaction at atomistic scales can be modeled from quantum mechanical principles like the density functional theory (DFT), which has been proven to achieve high accuracy and reliability [26]. However, the computationally affordable sample size of DFT is usually limited to no more than 10^3 atoms due to the typical $\mathcal{O}(N^3)$ computational complexity (N is the number of atoms). Moreover, the timescale required to simulate the H diffusion, trapping, and agglomeration processes in W is usually up to tens of nanoseconds [24], which is prohibitively expensive for *ab initio* molecular dynamics (AIMD). Large-scale classic MD simulations, which models atomic interactions by empirical potentials, can effectively access the microscopic origin and the underlying atomistic mechanisms involving up to millions of atoms, thus they were frequently used to study the behaviors of H isotopes in W [21,24,27–31].

The reliability of the MD simulation depends on the accuracy of the interatomic potential, which is used to describe the potential energy surface (PES) in the simulation. The quality of the interatomic potential is assessed by the consistency between the material properties predicted by the potential and those calculated by DFT or obtained experimentally. To explore the intragranular H blister nucleation mechanisms, a good interatomic potential should be accurate in a wide range of properties concerning the W-H interaction, such as the basic W properties, H solution and migration energies in W, the interaction between H and W free surfaces, the interaction between H and W lattice defects (vacancies and dislocation), formation of the H_2 molecules, interaction among solute H atoms in W lattices, and the formation of H self-clusters. Up to now, the interatomic potential of the W-H system is rare. There are only five empirical potentials for the W-H binary system reported in the literature: the analytical bond-order potentials (BOP) proposed by Juslin *et al.* [32] (BOP-Juslin), and Li *et al.* [33] (BOP-Li), the two embedded-atom model (EAM) potentials proposed by Bonny *et al.* [34] (EAM-Bonny-1 and EAM-Bonny-2) and one by Wang *et al.* [35] (EAM-Wang). These potentials are based on simple, analytical forms with only a few adjustable parameters, which may not be flexible enough to fit a broad range of fitting targets concerning the complexity of the W-H interaction. As a result, all the empirical potentials are more or less defective. For example, none of these potentials have reasonable descriptions of the interaction between H atoms and W surfaces. All of the empirical potentials overestimate the migration energy of H at compressive strain. And all of the empirical potentials have inaccurate descriptions of the H self-clusters in W. Thus, with these empirical potentials, the reliability of the MD simulations results concerning the formation of intragranular H blister, and the implications from these MD simulations, are put in question.

The recent development of machine learning (ML) potential [36–46] methods have brought new possibilities for modeling atomic interaction with high accuracy. The Deep Potential [44,45] (DP), as one of the popular ML potential methods, has been proven successful in numerous fields of metallic materials, such as Al-Mg binary [47] and Al-Mg-Cu ternary alloys [48], Ag-Au nanoalloys [49], Ti [50], V [51], W [26], and Fe-He [52]. Benefiting from the excellent ability of neural networks of fitting high-dimensional,

multivariant functions, DP has shown a great capability of describing multi-component materials' properties [47–49,53–61]. In comparison with empirical potentials, DP has a significant advantage in representability, i.e., the ability to reproduce the properties in the training database. The enhanced representability enables DP to take advantage of relatively large training datasets, which further enhances the generalization ability of the potential [26] based on the abundant information of the local atomistic environments contained in the training database. It is thus promising to construct an ML potential for the W-H binary system using the DP method, to facilitate the atomistic simulation, and unveil the mystery of the formation mechanism of the intragranular H blister in W.

In this work, a new ML potential for W-H binary system is developed with the DP method. This potential, namely DP-WH, shows significant advantages over empirical potentials by accurately describing a wide range of properties including the basic properties of W, solute H properties in W, H_2 molecule formation energy, the interaction between H and W surfaces, vacancy, dislocations, and formation of H self-clusters, while guarantees the basic properties which are already satisfied by the empirical potentials. In the MD simulation using this ML potential, we find a high concentration of H atoms in W tend to spontaneously form the H planar self-cluster with H atoms located between two adjacent (001) W planes [23]. This form of self-cluster facilitates the development of cracks along the cleavage plane, and is highly likely to be the nucleates of the crack-shaped H blisters observed in recent experiments [2,3,21].

This manuscript is organized as follows. In Sec. II, we present the technical details of the construction of the training database of DP-WH and the training of the ML potential. Then in Sec. III, we conduct a comprehensive benchmark of the potential, while making comparisons with the existing empirical potentials. In Sec. IV, using DP-WH, we run MD simulations to study the behaviors of high concentrations of H atoms in W. Finally, we draw conclusive remarks in Sec. V.

II. METHOD

A. Construction of the training dataset

We use the concurrent learning strategy implemented by the DP-GEN [62] software to generate the most compact and adequate dataset to guarantee a uniform accuracy of DP in the relevant configuration space. DP-GEN runs iteratively with the *exploration*, *labeling*, and *training* steps. The first DP-GEN interaction begins with *training* an ensemble of DP models with the *initial dataset*. Then, using one of the models, *exploration* of the configuration space is implemented by running MD simulations. The prediction error on the explored configurations is estimated by the deviation of the predicted forces by the ensemble of DP models, and only a small subset of the configurations with large errors are selected for *labeling*, i.e., for the DFT calculations of energy, forces, and virial tensors. The labeled data are then appended to the dataset, which is used in the *training* step of the next iteration.

a. Initial dataset. The initial dataset is composed of six subsets.

(1) Equilibrated bcc W with dilute hydrogen solutions. In a $2 \times 2 \times 2$ supercell of bcc W, 0-4 H atoms occupy random TISs. All configurations are optimized to zero pressure. All data generated during the optimizations are collected and appended to the dataset.

(2) Locally perturbed structures. Based on the optimized configurations in subset 1, strains and perturbations are performed. The applied strain ranges from -2% to $+2\%$. The mean perturbation distance of atoms is set to 0.03 Å, and the mean perturbation fraction of cell vectors is set to 1.0%. The labels of the strained and perturbed structures are obtained by single-point DFT calculation.

(3) Solution of 1 H atom in W lattice under strain. $3 \times 3 \times 3$ bcc supercells, with an H atom occupying the TIS and the OIS, are hydrostatically strained. The applied strain ranges from -4% to $+4\%$. The structures under strain are optimized, and all the data generated during the optimizations are collected.

(4) Hydrogen-vacancy complex clusters. In a $3 \times 3 \times 3$ bcc W supercell with a vacancy, 0–8 hydrogen atoms are bonded to the vacancy, with the seventh and eighth H forming a H_2 molecule. All nine structures are equilibrated to zero pressure. All the data generated during the optimizations are collected.

(5) 2 H atoms in neighboring TISs. In $3 \times 3 \times 3$ bcc W supercells, one H atom is placed in a TIS, and the other is placed at nine different neighboring TIS positions relative to the first H. All structures are equilibrated to zero pressure. The data generated during the optimizations are collected.

(6) H-H dimer with the bond-length ranging from 0.5 to 1.0 Å, with a step of 0.01 Å. The energy and forces of the dimer configurations are calculated by single-point DFT.

b. Exploration. The LAMMPS package [63] compiled with the DEEPMD-KIT [64] support is employed to perform Deep Potential molecular dynamics (DPMD) [44] simulations for the exploration of the configuration space. The exploration uses five types of structures as the initial configurations for the MD simulations, which are conducted with various settings and ensembles.

(1) The strained and perturbed $2 \times 2 \times 2$ supercell bcc bulk W with 0–4 H atoms randomly filling its TISs. NPT ensemble is employed with the pressure ranging from -0.5 to 2 GPa, and the temperature ranging from 50 to 1800 K.

(2) The strained and perturbed $2 \times 2 \times 2$ supercell bcc bulk W with 8 to 16 H atoms randomly filling the TISs. NPT ensemble is employed with the pressure ranging from -0.5 to 2 GPa, and the temperature ranging from 300 to 900 K.

(3) The equilibrated $2 \times 2 \times 2$ supercell bcc bulk W with one vacancy, filled with 0–6 hydrogen atoms. NVT ensemble is employed, with the temperature from 50 to 1800 K.

(4) Free surfaces with miller indexes (111), (110), and (100), and these surfaces with the adsorption of one H atom. NVT ensemble is employed. Temperature ranges from 50 to 1800 K.

(5) The strained and perturbed $2 \times 2 \times 2$ supercell bcc bulk W. NPT ensemble is employed with the pressure ranging from -0.5 to 2 GPa, and the temperature from 1800 to 5500 K.

During the MD exploration, the deviations of predicted forces by the four DP models, trained with identical hyperparameters but different random seeds, are used to estimate the error in the force predictions. If the maximal deviation of atomic forces is between 0.15 eV/Å (the lower bound) and 0.35 eV/Å (the upper bound) in an explored frame, the frame is considered as a candidate configuration and may be sent for *labeling*. In each iteration at most 100 candidate configurations are labeled. If the estimated errors of 99.9% of the explored configurations are below the lower bound, the *labeling* step is skipped.

c. Labeling. The labels of the candidate configurations, i.e., the energies, forces, and virial tensors, are computed by DFT with exchange-correlation modeled by the generalized gradient approximation (GGA) proposed by Perdew, Burke, and Ernzerhof (PBE) [65]. The DFT calculations are conducted using the VASP package [66,67]. The Brillouin zone is sampled by the Monkhorst-Pack method with a grid spacing of 0.08 Å^{-1}. The projector-augmented-wave (PAW) method is used and the energy cutoff of the plane-wave basis set is set to 1000 eV. For tungsten, the pseudopotential (W_SV) with the semicore states in the electronic structure of a W atom is used. The $6s$, $5d$, and $5p$ electrons are considered valence electrons. For hydrogen, the pseudopotential (H_H) with the outermost core radius 0.423 Å is used. The convergence criterion for the self-consistent field iteration is set to 10^{-6} eV. The same DFT parameters are also used for labeling the initial dataset.

d. Training. In each iteration, four models are trained simultaneously using the same dataset and hyperparameters, with the only difference being the random seeds employed to initialize the model parameters. We use the two-body embedding descriptor developed in the DP model in the *training* steps of the iterations. The principles of the DP models are described in Refs. [44,45], and the brief introduction of the hyperparameters is presented in Appendix of this manuscript. The Adam stochastic gradient descent method [68] with default hyperparameter settings provided by the TENSORFLOW package [69] is used to train the DP models. During the training processes, the weights and biases of the embedding nets ($\{W_k^e, b_k^e\}_{k=0}^m$) and those of the fitting net ($\{W_k^f, b_k^f\}_{k=0}^l$) are trained to minimize the loss function:

$$\mathcal{L} = \frac{1}{|\mathcal{B}|} \sum_{k \in \mathcal{B}} \left(p_\epsilon \frac{1}{N} |\hat{E}^k - E^k|^2 + p_f \frac{1}{3N} \sum_{i\alpha} |\hat{F}_{i\alpha}^k - F_{i\alpha}^k|^2 \right. \\ \left. + p_\xi \frac{1}{9N} \sum_{i\alpha} |\hat{\Xi}_{\alpha\beta}^k - \Xi_{\alpha\beta}^k|^2 \right), \tag{1}$$

which measures the accuracy of the model according to the differences between the DFT calculated energy $\hat{E}^k$, forces $\hat{F}_{i\alpha}^k$ and the virial tensor $\hat{\Xi}_{\alpha\beta}^k$, and those predicted by the DP model. In Eq. (1), $\mathcal{B}$ denotes a mini-batch of datasets, and $|\mathcal{B}|$ is the batch size. The superscript k denotes the index of the training data in the minibatch. Each training datum contains a configuration (including the coordinates of atoms, the box basis vectors, and the element types) and its corresponding labels. Prefactors (p_ϵ, p_f, p_ξ) are a set of hyperparameters determining the relative importance of the energy, forces, and virial tensor during the training. The prefactors are gradually adjusted according to the learning rate $r_l(t)$, which

exponentially decays with the training step t:

$$r_l(t) = r_l^0 k_d^{\frac{t}{t_d}}, \tag{2}$$

where r_l^0 is the learning rate at the beginning of the training, t_d denotes the typical timescale of the learning rate decaying, and k_d denotes the decay rate. The prefactors vary with the learning rate in the following way:

$$p_\alpha(t) = p_\alpha^{\text{limit}} \left[1 - \frac{r_l(t)}{r_l^0} \right] + p_\alpha^{\text{start}} \left[\frac{r_l(t)}{r_l^0} \right], \quad \alpha = \epsilon, \ f, \text{ or } \xi, \tag{3}$$

where $p_\alpha(t)$ is either of the three prefactors (p_ϵ, p_f, p_ξ) at training step t. p_α^{start} and p_α^{limit} are the prefactors at the beginning and at an infinitely small learning rate, respectively. In practice, a relatively larger force prefactor at the beginning and relatively balanced prefactors at the end of the training can make the best use of the training datasets and achieve relatively good accuracy [43]. Thus, in the *training* steps of the DP-GEN iterations, we set $p_\epsilon^{\text{start}} = 0.02$, $p_\epsilon^{\text{limit}} = 1.00$, $p_f^{\text{start}} = 1000.00$, $p_f^{\text{limit}} = 1.00$, $p_\xi^{\text{start}} = 0.00$, $p_\xi^{\text{limit}} = 0.00$, $t_d = 4000$, $k_d = 0.95$, and $r_0^l = 1 \times 10^{-3}$. The number of training steps is 800 000.

The DP-GEN iterations converge when estimated errors of more than 99.9% of the explored configurations are below the lower bound. Using the initial dataset and the dataset collected with DP-GEN, the productive models using the hybrid descriptor [26] mentioned in Eq. (A14) are trained. To ensure high training quality, the productive models are firstly trained by 2.40×10^7 steps with $p_e^{\text{start}} = 0.02$, $p_e^{\text{limit}} = 1.00$, $p_f^{\text{start}} = 1000.00$, $p_f^{\text{limit}} = 1.00$, $p_\xi^{\text{start}} = 0.02$, $p_\xi^{\text{limit}} = 1.00$, $t_d = 1.20 \times 10^5$, and $r_0^l = 1.00 \times 10^{-3}$. Then, with the model parameters initialized from the result of productive training, the models are additionally trained by 8.0×10^6 steps with $p_e^{\text{start}} = 10.00$, $p_e^{\text{limit}} = 10.00$, $p_f^{\text{start}} = 1.00$, $p_f^{\text{limit}} = 1.00$, $p_\xi^{\text{start}} = 0.02$, $p_\xi^{\text{limit}} = 1.00$, $t_d = 4.00 \times 10^4$, and $r_0^l = 1.00 \times 10^{-4}$ to obtain a higher training accuracy on the energy.

e. Refinement. The productive models exhibit satisfactory accuracy on the training dataset generated by DP-GEN. However, the accuracy of the productive models is not necessarily guaranteed for certain crucial material properties that may not have been adequately sampled during the exploration of DP-GEN. In order to enhance the overall quality and broaden the range of properties accurately captured by the productive models, two additional rounds of refinement training procedures are performed.

During the first refinement training, the refining dataset includes (1) the planar H self-cluster, (2) the SIA structures in pure W, (3) the generalized stacking fault (GFS) structures in pure W, and (4) the equilibrated W surface structures with the adsorption of H atoms. The refining dataset is appended to the training dataset before the refinement training. The subset of hydrogen self-cluster is generated via additional DP-GEN iterations. The initial configurations used during the exploration MD of the self-clusters are established in the following approach: in $1 \times 1 \times 4$ and $1 \times 4 \times 1$ supercells, with cell vectors aligned along the [110], [1$\bar{1}$0], and [001] directions, H planar self-clusters composed of 16 W atoms

TABLE I. Hyperparameters used during the productive training, additional productive training, the first refinement training, and the second refinement training. The superscript (2) and (3) denote the hyperparameter used in two-body embedding descriptors $\mathcal{D}^{(2)}$ and three-body embedding descriptors $\mathcal{D}^{(3)}$, respectively.

Hyper-parameter	Productive training	Additional productive training	First Refinement	Second Refinement
$r_c^{(2)}$ (Å)	6.0	6.0	6.0	6.0
$r_{cs}^{(2)}$ (Å)	0.5	0.5	0.5	0.5
$N_m^{(2)}$	100 W,100 H	100 W,100 H	100 W,100 H	100 W,100 H
layers of $\mathcal{G}^{(2)}$	20, 40, 80	20, 40, 80	20, 40, 80	20, 40, 80
M_2	80	80	80	80
$M^<$	16	16	16	16
$r_c^{(3)}$ (Å)	4.0	4.0	4.0	4.0
$r_{cs}^{(3)}$ (Å)	0.5	0.5	0.5	0.5
$N_m^{(3)}$	25 W,20 H	25 W,20 H	25 W,20 H	25 W,20 H
layers of $\mathcal{G}^{(3)}$	4, 8, 16	4, 8, 16	4, 8, 16	4, 8, 16
M_3	16	16	16	16
M_F	240	240	240	240
r_l^0	0.001	0.0001	0.0001	0.001
k_d	0.95	0.95	0.95	0.95
t_d	1.2×10^5	4×10^4	4×10^4	2×10^4
$p_\epsilon^{\text{start}}$	0.02	10.0	10.0	10.0
$p_\epsilon^{\text{limit}}$	1.0	10.0	10.0	10.0
p_f^{start}	1000.0	1.0	1.0	1.0
p_f^{limit}	1.0	1.0	1.0	1.0
p_ξ^{start}	0.02	0.02	0.02	0.02
p_ξ^{limit}	1.0	1.0	1.0	1.0
Training Steps	2.4×10^7	8×10^6	8×10^6	4×10^6

and 4 H atoms are established (W16H4). Based on these configurations, two more (W16H6) and four more (W16H8) H atoms are randomly introduced into the TISs in bulk W. Using these initial configurations, exploration DPMD runs up to 10 ps. The NPT ensemble is applied in the explorations at -0.5 to 2 GPa, with the temperature ranging from 300 to 600 K. The subset of the SIAs includes optimized structures of three types of SIA (namely, $\langle 111 \rangle$, $\langle 110 \rangle$, and $\langle 100 \rangle$ dumbbells) in $3 \times 3 \times 3$ bcc supercells. The subset of GFS structures is obtained directly from the GFS calculations using DFT, along $\langle 111 \rangle$ directions on $(1\bar{1}0)$ and $(11\bar{2})$ planes. The subset of the H surface adsorption structures is obtained by placing H atoms on the energetically most stable binding sites on the (100), (110), and (111) free surfaces, and optimizing the structures using DFT.

The refined models are trained for 8×10^6 steps with model parameters initialized by the productive model. During the first refinement training, the starting learning rate is set to 1×10^{-4}, and the starting prefactors of energy, force, and virial tensors, $p_\epsilon^{\text{start}}$, p_f^{start}, and p_ξ^{start} are set to 10.0, 1.0, and 0.02, respectively. t_d is set to 4×10^4. The other hyperparameters are the same as those used to train the productive model.

When conducting the benchmark of the model following the first refinement training, we observed occasional instability with the presence of a very high local concentration of H in the simulation system, which necessitates the second round of refinement training. This instability may be attributed to the lack of knowledge regarding local structures with such a high H concentration. Thus we performed additional DP-GEN

iterations, exploring a $2 \times 2 \times 2$ bcc W supercell containing 15 clustered H atoms (W16H15). The exploration DPMD simulations run up to 1 ps. The NPT ensemble is applied in the explorations at 0 to 2 GPa, with temperatures ranging from 300 to 600 K. The subset is appended to the training dataset. Additionally, we included SIA configurations of $\langle 11\xi \rangle$ dumbbell [70] in the refining dataset. In the second refinement training, $r_0^l = 1.00 \times 10^{-3}$, t_d is set to 2×10^4, and the models are trained for 4×10^6 steps, with other hyperparameters kept the same as those used in the first refinement training. Following the completion of the second refinement training, we obtained the final potential model, DP-WH.

B. List of hyperparameters

In Table I, we list the hyperparameters used in the productive and refinement training processes. The introductions of these hyperparameters are presented in Appendix. All hyperparameters related to the model construction remain the same during all these training processes. However, we adjust the hyperparameters related to model training, such as the learning rate at the beginning of training r_l^0, the typical timescale of the learning rate decay t_d, and the prefactors of energy, forces, and virial tensors (p_ϵ, p_f, p_ξ), in various training procedures.

C. Summary of the training dataset and the training accuracy

In Table II, we summarize the composition of the training dataset and the corresponding root mean square error (RMSE) on energy (RMSE$_E$) and forces (RMSE$_F$) of the

TABLE II. Summary of the training dataset and training RMSE of DP-WH.

Category	Subsets	$RMSE_E$ (eV/atom)	$RMSE_F$ (eV/Å)	N_{frame}
a. Initial dataset	1. Equilibrated W with dilute H	9.21×10^{-4}	2.66×10^{-2}	101
	2. Perturbed W with dilute H	9.51×10^{-4}	2.19×10^{-2}	619
	3. Strained W with 1 H	3.25×10^{-4}	1.75×10^{-2}	235
	4. H-vacancy clusters	1.28×10^{-3}	4.26×10^{-2}	148
	5. 2H in neighboring TISs	4.36×10^{-4}	1.84×10^{-2}	142
	6. H-H dimer	1.07×10^{-2}	9.18×10^{-2}	28
b. DP-GEN dataset	1. W with dilute H	2.71×10^{-4}	1.13×10^{-1}	3442
	2. W with high concentration H	4.87×10^{-3}	1.95×10^{-1}	557
	3. H atoms in vacancy	5.48×10^{-3}	1.61×10^{-1}	1615
	4. H atoms on surface	3.89×10^{-4}	1.25×10^{-1}	3222
	5. Pure W up to 5500 K	1.23×10^{-2}	3.38×10^{-1}	980
c. First refinement dataset	1. Self-clusters	4.30×10^{-3}	1.46×10^{-1}	656
	2. SIA($\langle 111 \rangle$, $\langle 110 \rangle$, $\langle 100 \rangle$ dumbbell)	4.41×10^{-4}	5.90×10^{-1}	54
	3. GSF	1.07×10^{-4}	4.48×10^{-1}	954
	4. Surfaces with H adsorption	2.30×10^{-3}	5.40×10^{-2}	57
d. Second refinement dataset	1. High local concentration	3.51×10^{-3}	1.82×10^{-1}	291
	2. SIA($\langle 11\xi \rangle$ dumbbell)	4.31×10^{-4}	7.15×10^{-2}	9

DP-WH on each subset. The entire training dataset consists of 13 110 frames of DFT data. The overall RMSE on energy and forces are 4.69×10^{-3} eV/atom and 1.27×10^{-1} eV/Å, respectively.

III. VALIDATION OF POTENTIAL

In this section, a comprehensive benchmark is conducted to evaluate the accuracy of the DP-WH on critical properties of the W-H interaction. Comparisons are made among the material properties calculated by DP-WH and empirical potentials. We first examine the basic W properties and then assess the performances of these potentials on solute H properties and these properties under the effect of the strain. Then, we benchmark the behavior of H on low-index W free surfaces. Next, we examine the interaction between the H atoms and the monovacancy, and the binding energy of a H atom with dislocations. Finally, we benchmark the interaction of H atoms at neighboring interstitial sites and the binding energy of H in the H self-cluster composed of interstitial H atoms. Calculations of the properties are implemented by the LAMMPS code [63]. Visualizations are conducted by the OVITO software [71].

A. Basic W properties

The benchmark results of the basic W properties calculated by DP-WH and other empirical potentials are listed in Table III. For fare comparisons with the empirical potentials, we also listed the previous DFT and experimental results, which were used as the fitting target of the empirical potentials. The lattice constant and the cohesive energy of bcc W determined by experiments is 3.165 Å and −8.9 eV. The present DFT predicts close lattice constant and cohesive energy to experimental values. For the sake of self-consistency of the training dataset, DP-WH is fit to the DFT values. Thus DP-WH slightly disagrees with the experiments on the lattice constant and cohesive energy of bcc W. In comparison, BOP potentials by bJuslin *et al.* [32] and Li *et al.* [33], which

are directly fit to experimental values, well reproduce the experimental lattice constant and cohesive energy. The EAM potentials by Bonny *et al.* [34] and Wang *et al.* [35] use the EAM-2 of Marinica-13 [72] as their W-W potential, which correctly reproduces the experimental cohesive energy, but not the lattice constant. All listed potentials reproduce the elastic constants well. The formation energy of point defects, i.e., monovacancy and self-interstitial atoms (SIAs) with $\langle 100 \rangle$, $\langle 110 \rangle$, and $\langle 111 \rangle$ dumbbell configurations, are well-predicted by all potentials except BOP-Juslin [32]. According to a recent work [70], the most stable SIA structure in W is $\langle 11\xi \rangle$ dumbbell, where ξ is roughly 0.5. DP-WH can well reproduce the $\langle 11\xi \rangle$ dumbbell as the most stable SIA, while the empirical potentials predict $\langle 11\xi \rangle$ dumbbell to be unstable. None of the empirical potentials are fitted to the surface energies, thus the correct predictions of formation energy of low-index free surfaces of W are not expected for the empirical potentials. DP-WH not only predicts the correct surface energies but also reproduces the reconstruction of the (100) free surface of W [73]. The accurately predicted surface properties serve as the basis for the correct description of the H adsorption sites and migration behavior of H atoms on W surfaces, which are presented in later benchmarks. The equation of states (EOS) curves of bcc W predicted by the DP-WH and the empirical potentials are shown in Fig. 1. As the DP-WH and empirical potentials predict different ground-state lattice constants, the minimum on the EOS curve predicted by the potentials corresponds to different atomistic volumes. The changes of energy with volume predicted by BOP-Juslin and BOP-Li show similar tendencies with DFT and DP-WH. However, the EOS curves predicted by EAM potentials increase faster than DFT with smaller and larger atomistic volumes. Thus we argue that the EOS predicted by the EAM potentials is less accurate than the other potentials.

The generalized stacking fault energy (GSFE or γ line) is the variation of energy on displacing one part of the crystal against the other on a specific plane. We calculate the γ-line for bcc W along the [111] direction on the $(1\bar{1}0)$ and $(11\bar{2})$ plane using DFT, DP-WH, and empirical potentials, and the

TABLE III. Basic properties of bcc W predicted by various potentials, including the lattice constant a_0, cohesive energy of bcc W $E_{\text{coh}}^{\text{bcc}}$, independent components of elastic constant C_{ij}, formation energy of SIA E_f^{SIA}, formation energy of monovacancy E_f^{vac}, and low-index free surface E_f^{surf}.

Potential	BOP-Juslin [32]	BOP-Li [33]	EAM [34,35]	Reference	DFT	DP-WH
$a_{0,\text{bcc}}$ (Å)	3.165	3.165	3.140	3.185 [74], 3.172 [75]	3.185	3.186
$E_{\text{coh}}^{\text{bcc}}$ (eV)	−8.890	−8.906	−8.900	−8.900 [76], −8.412 [74]	−8.412	−8.413
C_{11} (GPa)	542 [33]	515	544	517 [75], 523 [72]	–	516
C_{12} (GPa)	191 [33]	203	208	198 [75], 203 [72]	–	201
C_{44} (GPa)	162 [33]	162	160	142 [75], 160 [72]	–	146
$E_f^{\text{SIA}}\langle 100\rangle$ (eV)	8.93 [33]	12.01	12.86	11.49 [77], 12.87 [72]	11.84	12.03
$E_f^{\text{SIA}}\langle 110\rangle$ (eV)	8.77 [33]	9.53	10.82	9.84 [77], 10.83 [72]	10.14	10.35
$E_f^{\text{SIA}}\langle 111\rangle$ (eV)	9.62 [33]	9.33	10.52	9.55 [77], 10.53 [72]	9.86	10.16
$E_f^{\text{SIA}}\langle 11\xi\rangle$ (eV)	unstable	unstable	unstable	10.25 [78]	9.84	10.13
E_f^{vac} (eV)	1.68	3.52	3.49	3.11 [79], 3.56 [33]	3.56	3.30
$E_f^{\text{surf}}(100)$ (J/m^2)	1.446 [33]	3.157 [80]	2.721 [80]	4.635 [81], 4.021 [75]	3.933	3.900
reconstruction	N	N	N	–	Y	Y
$E_f^{\text{surf}}(110)$ (J/m^2)	0.931 [33]	2.319 [80]	2.306 [80]	4.005 [81], 3.268 [75]	3.348	3.302
$E_f^{\text{surf}}(111)$ (J/m^2)	1.720	3.222 [80]	2.963 [80]	4.452 [81], 3.556 [75]	3.525	3.563

comparison is made in Fig. 2. DP-WH agrees well with DFT on nearly the entire γ-lines. The EAM potentials predict reasonable γ-lines, but the values are lower than the DFT results. The γ-line predicted by the BOP-Juslin potential is notably lower than DFT. Such a low GSFE even leads to unphysical dissociation of edge dislocation cores [82]. BOP-Li predicts a slightly lower γ-line on $(1\bar{1}0)$ plane and a higher γ-line on $(11\bar{2})$ plane. Among the empirical potentials, BOP-Li has the best prediction on the γ-lines.

B. H solution in W and formation of hydrogen molecule

The properties of solute H in bcc W lattices predicted by the empirical potentials and DP-WH are listed in Table IV. The solution energy of H at TIS and OIS in bcc W is calcu-
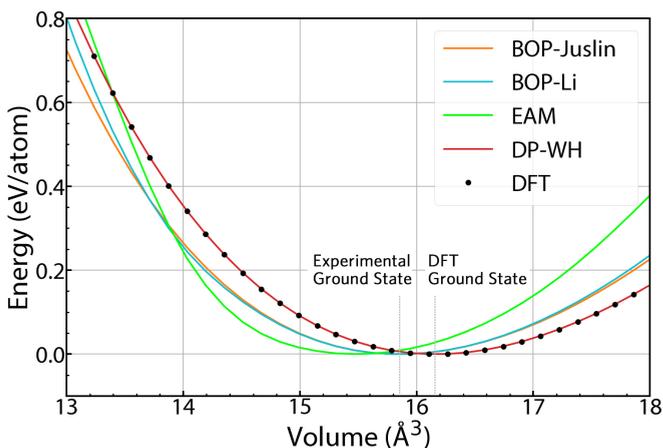

FIG. 1. Equation of states of bcc W predicted by various potentials. On the ground-state volume, the BOP potentials agree with experiments, and DP-WH agrees with the present DFT. The change of EOS curve predicted by the BOP and DP-WH have similar tendencies. EAM potentials are less accurate in predicting the EOS curve than other potentials.

lated by

$$E_s^{\text{tet}} = E_{\text{tot}}^{\text{tet}}(\text{WH}) - N_W E(\text{W}) - \frac{1}{2}E(\text{H}_2) \quad (4)$$

and

$$E_s^{\text{oct}} = E_{\text{tot}}^{\text{oct}}(\text{WH}) - N_W E(\text{W}) - \frac{1}{2}E(\text{H}_2), \quad (5)$$

where $E_{\text{tot}}^{\text{tet}}(\text{WH})$ and $E_{\text{tot}}^{\text{oct}}(\text{WH})$ are the total energies of configurations with a H atom in a TIS and an OIS in the W supercells, N_W is the number of W atoms, $E(\text{W})$ is the energy contribution of a W atom in the ground-state bcc W lattice, and $\frac{1}{2}E(\text{H}_2)$ stands for the half of the total energy of an isolated hydrogen molecule. The solution energies of H in TIS and OIS were the fitting targets of the empirical potentials except for the EAM-Bonny. BOP-Li is fitted to the solution energy calculated by DFT. BOP-Juslin and EAM-Wang are fitted to the solution energy with the zero-point-energy (ZPE) correction. In previous literature, ZPE correction is evaluated by summing up the ground-state vibration energy of the H atom in W [23]. In this work, DP-WH is fitted to TIS H and OIS H solution energy calculated by DFT without the ZPE correction. Also the ZPE correction is separately evaluated by

$$\Delta E_{\text{ZPE}}^{\text{tet}} = E_{\text{ZPE}}^{\text{tet}}(\text{WH}) - E_{\text{ZPE}}(\text{W}), \quad (6)$$

where $E_{\text{ZPE}}^{\text{tet}}(\text{WH})$ is the ZPE of the bcc W supercell with one H in a TIS, and $E_{\text{ZPE}}(\text{W})$ is the ZPE of the W supercell without the H atom. $E_{\text{ZPE}}^{\text{tet}}(\text{WH})$ and $E_{\text{ZPE}}(\text{W})$ are calculated by summing up the energy of all vibration modes of all W and H atoms, which is implemented by the PHONOLAMMPS [83] code. The $E_{\text{ZPE}}^{\text{tet}}(\text{WH})$ predicted by DP-WH in Table IV shows a small difference with the previous DFT result. The difference between solution energies of OIS H and TIS H, ΔE_s, is slightly underestimated by all previous empirical potentials but is well reproduced by DP-WH. The migration energy along the minimum energy paths (MEP) for hydrogen diffusion in W is calculated using the nudged elastic band (NEB) method [84]. Except for BOP-Juslin, all potentials well-reproduce the migration energy for H from TIS to another nearby TIS, which is the most feasible pathway for H diffusion in W. The formation of a H_2 molecule, $E_f^{H_2}$, can be
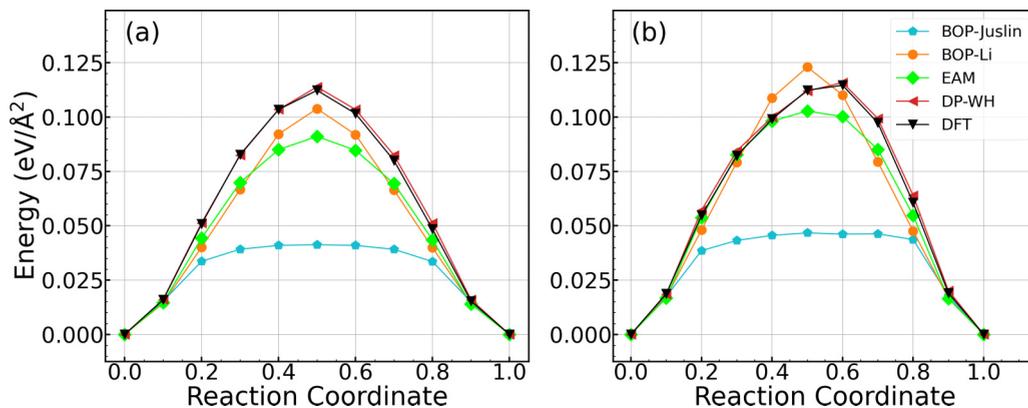
FIG. 2. GSFE of bcc W predicted by various potentials and DFT. (a) γ-line along the [111] direction on the $(1\bar{1}0)$ plane. (b) γ-line along the [111] direction on the $(11\bar{2})$ plane.

reasonably reproduced by BOP-Juslin, BOP-Li, EAM-Wang, and DP-WH. The formation energy with respect to $E_f^{H_2}$, as a function of the H-H dimer bond length predicted by all the potentials is shown in Fig. 3. BOP-Juslin and BOP-Li use the same H-H potential proposed by Brenner *et al.* [88], which accurately describes the H-H dimer energy. EAM-Wang qualitatively reproduces the H-H dimer energy, but the energy increases too fast with a longer hydrogen bond length. With more focus on the H-defect interactions in W and the H-He interactions, the two EAM-Bonny potentials disregard the formation of a hydrogen molecule [34]. Thus the accurate description of H-H dimer energy by EAM-Bonny is not expected. Ignoring $E_f^{H_2}$ also leads to the unreasonable solution energy of the H in W TISs and OISs calculated by Eq. (4). DP-WH well-reproduces the H-H dimer energy since it is fitted to the DFT result.

C. Solute H properties under strain

Due to the complexity of the fusion service condition, lattice strains are inevitably induced by many factors such as the presence of defects [24], thermal stresses caused by the temperature gradient [15], and the formation of blisters itself [2]. It is proved by DFT that strain is a key factor determining the solubility and diffusivity of H in W [89]. It is thus important for a potential to accurately describe the effect of lattice strain on H solution energy and migration energy.

The solution energy of the H atom in a TIS and OIS of a hydrostatically strained bcc W predicted by various potentials is shown in Fig. 4. For the solution energy of a TIS H shown in Fig. 4(a), a nearly constant shift is observed between the DFT/DP-WH calculated results and the previous DFT results calculated by Zhou [89] *et al.* The shift is attributed to the ZPE correction. And DP-WH agrees well with the present DFT since the solution energy under strain is included in the training dataset. The E_s^{tet} calculated by the empirical potentials with respect to lattice strain is obtained from Ref. [35] For the two BOP potentials, E_s^{tet} decreases faster with strain than the DFT results. EAM-Wang agrees well with the ZPE corrected DFT result at positive strain but overestimates E_s^{tet} at negative strains. The solution energies calculated by EAM-Bonny potentials are shifted to 0.0 at zero pressure. The change of E_s^{tet} with respect to the strain predicted by EAM-Bonny-1 shows a similar tendency to that predicted by EAM-Wang. But EAM-Bonny-2 predicts that the E_s^{tet} increase with either the tensile or compressive strain, which disagrees with DFT and all other potentials.

For the solution energy of an OIS H shown in Fig. 4 (b), DP-WH also agrees with the present DFT result on the solution energy and the effect of the strain. The BOP-Juslin potential has the best agreement with DFT among the empirical potentials. E_s^{oct} predicted by EAM-Wang potential is overestimated at compressive strains but is underestimated

TABLE IV. Solute H properties in bcc W and H-H dimer properties. Solute H properties include the solution energy of H atom at TIS E_s^{tet}, OIS E_s^{oct} and their energy differences ΔE_s, the ZPE correction of a H atom at the TIS ΔE_{ZPE}^{tet}, H migration energy from TIS to its first-nearest-neighbor TIS, $E_{mig}^{t \to t}$. H-H dimer properties include the formation energy $E_f^{H_2}$ and the equilibrium bond length $L_c^{H_2}$ of an isolated H_2 molecule.

| Potential	BOP-Juslin	BOP-Li	EAM-Bonny-1	EAM-Bonny-2	EAM-Wang [35]	Reference	DFT	DP-WH
E_s^{tet} (eV)	1.04 [35]	0.86 [35]	3.65 [35]	−19.86 [35]	1.05 [35]	0.85 [85], 0.87 [86]	0.88	0.95
ΔE_{ZPE}^{tet} (eV)	−	−	−	−	−	0.22 [86]	−	0.24
E_s^{oct} (eV)	1.40 [35]	1.18 [35]	3.99 [35]	−19.48 [35]	1.40 [35]	1.26 [85], 1.33 [87]	1.27	1.35
ΔE_s (eV)	0.36	0.32	0.34	0.34	0.35	0.41 [85], 0.45 [87]	0.39	0.40
$E_{mig}^{t \to t}$ (eV)	0.34 [35]	0.23 [35]	0.21 [35]	0.21 [35]	0.22 [35]	0.20 [85], 0.21 [87]	0.20	0.20
$E_f^{H_2}$ (eV)	−4.75	−4.75	−	−	−4.72	−	−4.53	−4.53
$L_b^{H_2}$ (eV)	0.75	0.75	−	−	0.74	−	0.75	0.75
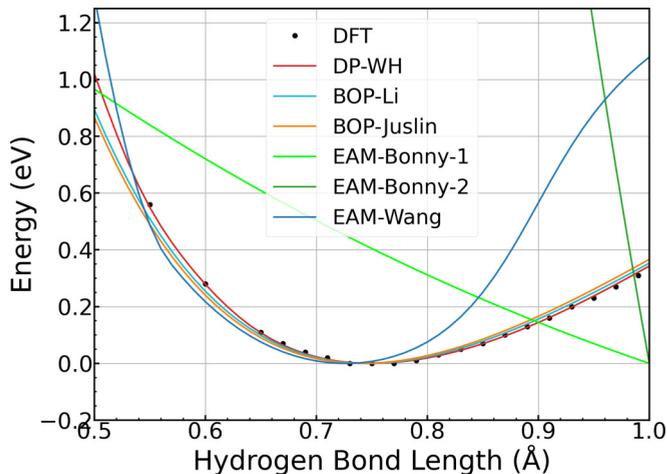
FIG. 3. Relative energy of a H-H dimer as a function of the bond lengths. BOP-Juslin, BOP-Li, and DP-WH reproduce the DFT values. Though EAM-Wang is accurate in describing the equilibrium H_2 molecule bond length, the H-H dimer energy predicted by EAM-Wang deviates from DFT results with longer bond length. EAM-Bonny potentials ignore the H-H dimer energy and $E_f^{H_2}$, thus their description of the H-H dimer energy is unreasonable.

at large tensile strains. E_s^{oct} predicted by BOP-Li decreases significantly faster with strain than DFT results. Also, the E_s^{oct} predicted by the EAM-Bonny potentials is shifted to 0.0 at zero stress. The change of E_s^{oct} with respect to the strain predicted by EAM-Bonny-1 is close to that predicted by EAM-Wang. EAM-Bonny-2 predicts higher E_s^{oct} at compressive or large (>2%) tensile strain than at zero stress, which disagrees with DFT and all other potentials.

The migration energy of a solute H atom in a W lattice and the effect of the hydrostatic lattice strain is shown in Fig. 5. These migration paths are not explicitly included in the training database of DP-WH. In Fig. 5(a), we show the energies along two MEPs for hydrogen migration in W. One is from a TIS to a first-nearest-neighbor (1NN) TIS ("t-t"), and another is from a TIS, via a nearby OIS to a second-nearest-neighbor (2NN) TIS ("t-o-t"). For both paths, the migration energy and the energy along the MEP predicted by DP-WH agree well with the DFT results by Liu *et al.* [85]. The migration barrier for the "t-t" path is nearly 0.2 eV. It is only half of the barrier for the "t-o-t" path (nearly 0.4 eV). Thus, in MD simulation using DP-WH, the H migrates in W mainly via the "t-t" path. The migration barrier along the "t-o-t" shown in Fig. 5(a) is very close to the ΔE_s in Table IV, which indicates a H in an OIS is a metastable state with a very shallow energy basin.

The changes in the migration energies along the "t-t" path with the hydrostatic strain are shown in Fig. 5(b). The migration barrier predicted by BOP-Juslin is 0.34 eV at zero stress, much higher than the migration barrier predicted by other potentials and DFT. Thus we omitted BOP-Juslin in Fig. 5(b). Benefiting from the good generalization ability, DP-WH shows a good agreement with the DFT result obtained from Ref. [35]. The migration energies predicted by all empirical potentials increase faster with the compressive strain than the DFT results. It is thus expected that all the

empirical potentials might underestimate H diffusivity under compressive strain.

D. H adsorption and migration on low-index W free surfaces

Interactions between H atoms and W surfaces are fundamental to understanding many aspects of H behaviors in W. The critical processes to the intragranular H blister, such as energetic H impinging on W [90], H diffusion and interaction beneath W surfaces [91], formation of internal H cracks [3], and development of H blisters [2], involve the interaction between H with W free surfaces. (100) is the experimental cleavage plane of W, and (110) is the index with the lowest surface energy. H behaviors on these low miller-index surfaces, and to a less extent, (111), have received essential attention [73]. It is thus important for a potential to accurately describe the interaction between a H atom with these low-index free surfaces. We benchmark the adsorption sites, binding energy, and the migration barrier of a H atom on these low-index surfaces using various potentials. The results are shown in Fig. 6.

The binding energy of a H atom on W free surface is calculated by

$$E_b^{surf} = E_{tot}^{surf}(WH) - E_{tot}^{surf}(W) - \tfrac{1}{2}E(H_2), \quad (7)$$

where $E_{tot}^{surf}(WH)$ is the total energy of the binding configurations, $E_{tot}^{surf}(W)$ is the total energy of the surface configuration without H and $E(H_2)$ is the total energy of an isolated hydrogen molecule. A negative value of E_b^{surf} suggests an energetically favorable binding state. As shown in Fig. 6(a), the most stable adsorption site on the reconstructed (100) surface by DFT is the "short-bridge (sb)" site, with a binding energy of -0.90 eV. It is in agreement with the -0.91 eV predicted by previous DFT calculations [73]. The barrier of H migration from a "sb" site to a nearby "sb" site is 0.44 eV [73]. DP-WH is the only potential that predicts the reconstruction of the (100) free surface and correctly reproduces the H binding energy, H adsorption sites, and a 0.53 eV H migration barrier on the (100) free surface. On the (110) free surface, as shown in Fig. 6(b), the DP-WH predicts the lowest binding energy -0.76 eV at the "threefold (3f)" site, in agreement with the present DFT calculated -0.73 eV, and -0.75 eV from a previous DFT result [73]. The migration barrier from a "3f" site on the (110) surface to a nearby "3f" site is 0.08 eV as predicted by DP-WH, which is also in good agreement with previous DFT results [73]. On the (111) free surface, DFT results in Ref. [73] give the most stable configuration *at* the "bond-centered (bc)" site between the first W layer and the second W layer with a -0.62 eV binding energy. DP-WH and the present DFT predict that the most stable H binding configuration is *above* the "bc" site. As shown in Fig. 6(c), DP-WH predicts the binding energy to be -0.81 eV, which underestimates the present DFT result -0.64 eV. The migration energy predicted by DP-WH is 0.29 eV, which is higher than the 0.27 eV from the previous DFT calculation [73].

Remind that all W-H empirical potentials have poor descriptions of the free surface energy, and thus the accuracy of their surface-related properties is not expected. Nevertheless, for comparisons with DP-WH, we benchmark their H binding energies on surfaces and H adsorption sites. The results are
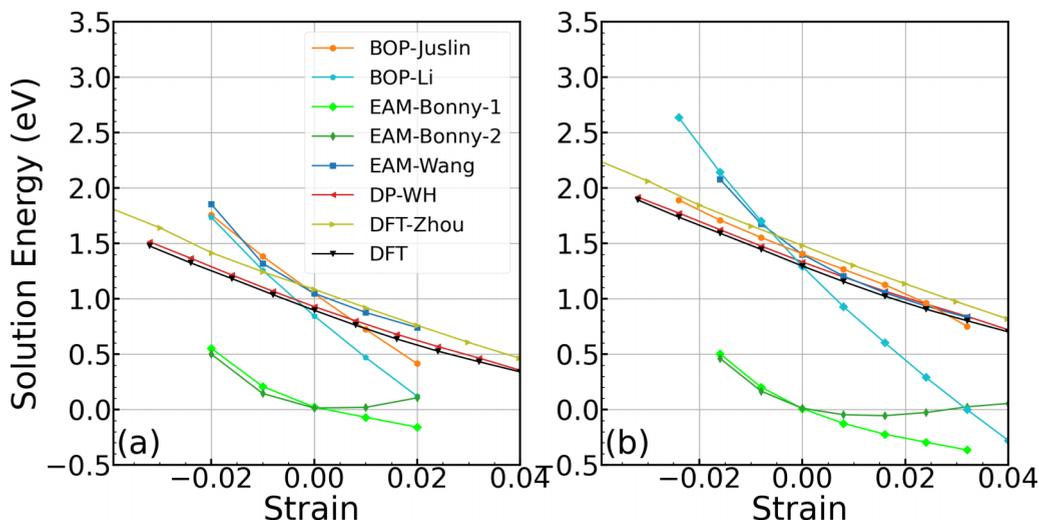
FIG. 4. Solution energy of interstitial H atoms in hydrostatically strained W lattices. (a) TIS H. The results of the empirical potentials are obtained from Ref. [35]. (b) OIS H. The DFT results with ZPE corrections are obtained from Ref. [89].

presented in Figs. 6(d)–6(f). We exclude the binding configuration predicted by EAM-Bonny-2 because it predicts that the binding of H on the W surface is energetically less favorable than the H solution in TISs of W bulk, which is contrary to the present DFT result. Moreover, as neither EAM-Bonny potentials give correct energy $E(H_2)$, we exclude their H binding energies on surfaces in Fig. 6, and only show the adsorption sites predicted by EAM-Bonny-1. As shown in Fig. 6(d), none of the empirical potentials predict the reconstructed (100) free surface. EAM-Bonny-1 and BOP-Li predict the "two-fold (2f)" adsorption site, and BOP-Juslin and EAM-Wang predict the "four-fold (4f)" site. On (110) surfaces shown in Fig. 6(e), however, EAM-Wang, EAM-Bonny-1, and BOP-Li correctly predict the "three-fold" (3f) binding site. BOP-Juslin incorrectly predicts H binding at the "sb" site. On (111) free surfaces shown in Fig. 6(f), none of the empirical potentials predict the correct H adsorption sites. Only the BOP-Li predicts somehow reasonable H binding energy on the (110) and (111) free surfaces. Except that, none of the empirical

potentials give reasonable predictions of H-binding energy on the surfaces. Thus we omit the calculations of migration energy of H atom on free surfaces using empirical potentials, due to their poor descriptions of the H-surface interactions.

E. Interaction between H atoms and monovacancy

It is revealed experimentally that vacancy-type defects, which are produced massively under irradiation or at high temperatures, significantly enhance hydrogen retention [9,92]. Due to the importance of vacancy-hydrogen interactions in W, there are abundant results concerning the interaction between H and vacancy-type defects [8,11,13,17], and advanced models describing the H retention rate with the presence of vacancy-type defects already exist [27]. Although it was believed that the exposure to low-energy high-flux H plasma does not generate a large number of vacancies in W [2,3,14,15,20,21], accurate description of the interaction between a vacancy and H atoms is still an important
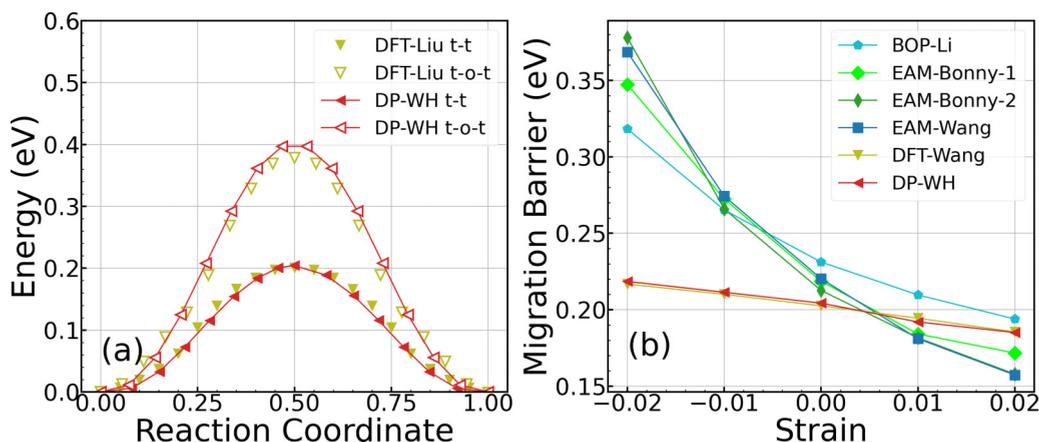

FIG. 5. Migration properties of a solute H atom in bcc W. (a) Energy on the H migration MEP along the "t-t" path and "t-o-t" path. Comparisons are made between the result predicted by DP-WH and DFT calculation by Liu *et al.* [85] (b) The "t-t" path migration energy as a function of hydrostatic lattice strain predicted by various potentials. The DFT results by Wang [35] *et al.* is plotted for comparison.
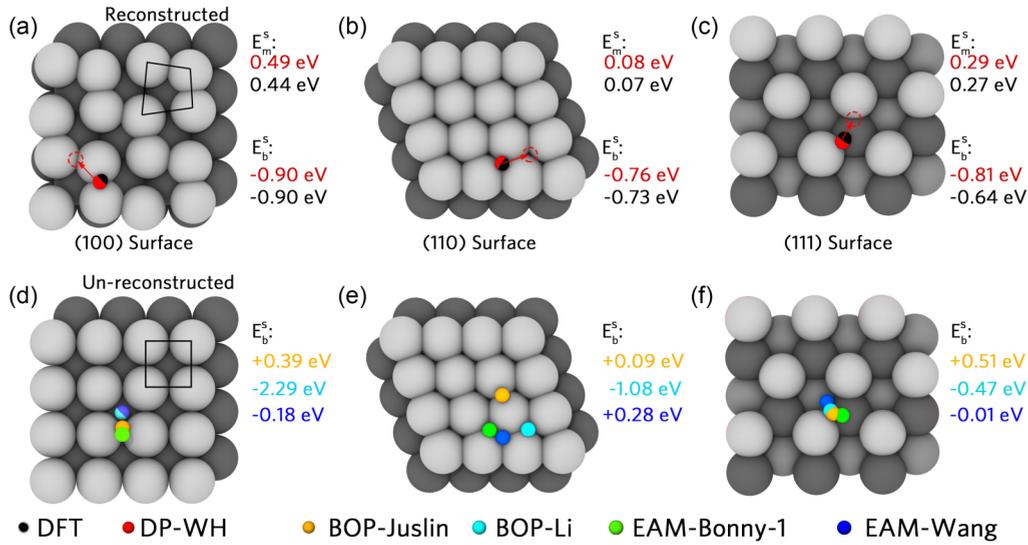
FIG. 6. H adsorption and migration behaviors on low-index surfaces of W predicted by various potentials. E_b^s is the binding energy of the H atoms to the surfaces, and E_m^s is the H migration energy on W surfaces. [(a)–(c)] DP-WH and DFT predicted adsorption sites, binding energy, and migration energy. The DFT binding energy is calculated in this work, and the migration energy is obtained from Ref. [73]. [(d)–(f)] Adsorption sites and binding energy predicted by empirical potentials. Migration energies predicted by the empirical potentials are not presented, as their accuracies are not expected.

benchmark for a high-quality interatomic potential. The basis of the H-vacancy interaction is the binding between H atoms and a monovacancy. The sequential binding energy of a vacancy with $n = 1$–8 H atoms is plotted in Fig. 7(a). The sequential binding energy is determined by

$$E_b^{seq}(H_n V_1) = E_{tot}(H_n V_1) - E_{tot}(H_{n-1}V_1) - \left(E_s^{tet} + \tfrac{1}{2}E(H_2)\right),$$
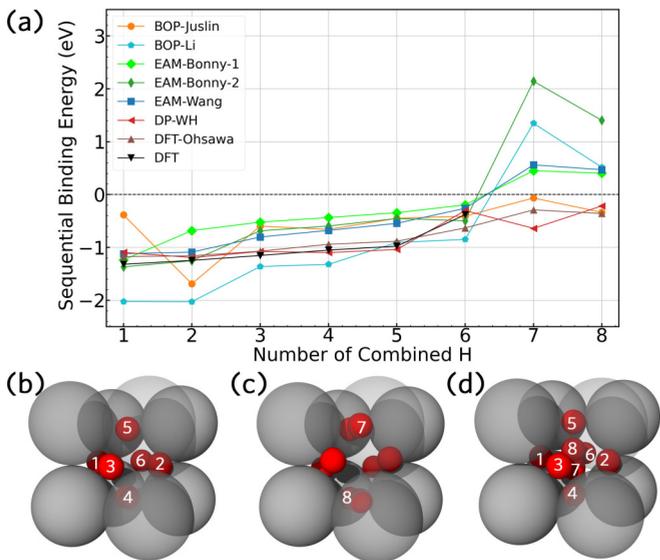(8)

FIG. 7. (a) Sequential binding energy of a series of H atoms to a monovacancy in W predicted by various potentials. (b) The binding sequence of 1–6 H atoms in a vacancy in W. (c) The binding sequence of the seventh and eighth H atoms in a vacancy in W. (d) The binding configuration of a metastable $H_8 V_1$ complex cluster, with the seventh and eighth H forming a H_2 molecule.

where the $E_{tot}(H_n V_1)$ is the total energy of the binding configuration of a monovacancy with n H atoms, $E_{tot}(H_{n-1}V_1)$ is the total energy of the binding configuration with one less H atom, E_s^{tet} is the solution energy of an H atom in the tetrahedral lattice site in bulk W, and $E(H_2)$ is the energy of an isolated hydrogen molecule. The last term in Eq. (8), $E_s^{tet} + \tfrac{1}{2}E(H_2)$, is the energy contribution of an H atom in a TIS of bcc W. $E_b^{seq}(H_n V_1)$ measures the binding energy of a $H_{n-1}V_1$ with a TIS H. The numbers in Figs. 7(b) and 7(c) show the binding sequence. In a vacancy in bcc W, H first takes the six positions near the "face centers" (OIS positions if the vacancy is absent). The seventh hydrogen and the eighth hydrogen slightly distort the binding configuration. As can be observed from Fig. 7(a), most empirical potentials overestimate the sequential binding energy of a vacancy with less than seven hydrogen atoms [35], except that BOP-Li underestimates the sequential binding energy. BOP-Juslin underestimates the binding energy of the second hydrogen to $H_1 V_1$ because the cutoff radius between W and H is too short, which makes BOP-Juslin not suitable for describing H interaction with vacancy [35]. The binding configurations of a vacancy with 1–6 H atoms are explored by DP-GEN, thus DP-WH shows better agreement with the present and previous DFT results [13,87] in predicting the binding energy with less than 6 H atoms than the empirical potentials, though the binding between the first H and the monovacancy predicted by DP-WH is slightly weaker than DFT result. According to the binding energy shown in Fig. 7(a), most empirical potentials predict that a vacancy in W accommodates up to 6 H atoms. DP-WH predicts that a vacancy accommodates up to 9 H atoms, with the sequential binding energy of the nineth and tenth H being -0.35 and $+0.43$ eV. The previous DFT result shows that the maximum number of H atoms accommodated by a monovacancy is 12 [13], which is underestimated by all potentials including DP-WH. DP-WH predicts that the minimum number of H atoms
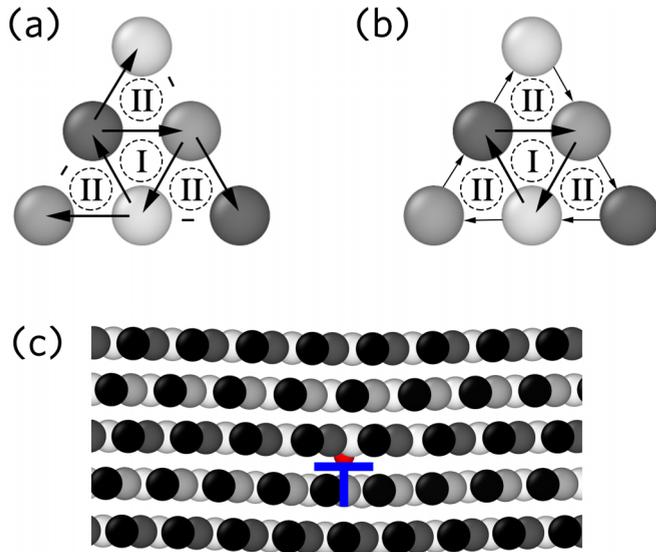
FIG. 8. Binding configurations of dislocations with H atoms. (a) The "degenerate core" predicted by BOP potentials. (b) The "nondegenerate core" is predicted by EAM potentials, DP-WH, and DFT. "Screw Dislocation I" indicates the H binding site within the dislocation core, and "Screw Dislocation II" indicates the H binding site adjacent to the dislocation core. (c) The binding configuration of the edge dislocation core with H.

required to stabilize a H_2 molecule is 8. It is in agreement with previous DFT calculations [11]. The binding configuration is the metastable H_8V_1 cluster composed of a H_6V_1 cluster and an H_2 molecule at the center of the vacancy, as shown in Fig. 7(d). Configurations involving many H atoms, such as the binding configurations of a monovacancy with more than eight atoms, and the binding configurations of multiple vacancies binding with H atoms, are not included in the training database. High accuracies on such structures are not expected for DP-WH. It should be noted that the interaction between H atoms and vacancy-type defects in W was extensively studied [8,27]. Moreover, the H-vacancy interaction plays a less critical role under working conditions where W is exposed to low-energy, high-fluence H plasma. Thus reproducing all aspects of interactions between H and vacancy-type defects are beyond the scope of this work.

F. Interaction between an H atom with dislocations

Dislocations, the major carriers of plastic deformation in bcc W, are massively produced beneath surfaces in W under exposure to H plasma [15,20]. Dislocations also contribute substantially to the retention of hydrogen [14]. Theoretically, dislocations are believed to be the important trapping sites for hydrogen atoms and are suggested to facilitate H blistering [18]. It is thus critical for a W-H interatomic potential to accurately predict the binding configuration and binding energy of an H atom to dislocations in atomistic simulations. Empirical potentials predict different core structures of screw dislocations. BOP-Juslin and BOP-Li potentials predict the "degenerate core" as shown in Fig. 8(a). EAM potentials, DP-WH, and DFT predict the "nondegenerate core" as shown in Fig. 8(b). The schematics of H binding sites *within* the

screw dislocation core ("Screw Dislocation I") and *adjacent* to the screw dislocation core ("Screw Dislocation II") are shown in Figs. 8(a) and 8(b), and the binding configuration of the edge dislocation core with an H atom is shown in Fig. 8(c).

The binding energies of an H atom to dislocations predicted by various potentials are shown in Table V and are determined by

$$E_b^d = E_{tot}^d(WH) - E_{tot}^d(W) - \left(E_s^{tet} + \tfrac{1}{2}E(H_2)\right), \quad (9)$$

where $E^d(WH)$ is the total energy of the binding configuration of the dislocation with an H atom, $E_{tot}^d(W)$ is the total energy of the dislocation without binding the H atom, and $E_s^{tet} + \tfrac{1}{2}E(H_2)$ is the energy contribution of an H atom at a TIS in bulk W. According to the DFT results in Ref. [17], the binding energy between H and the screw dislocation in the screw dislocation I site is -0.55 eV, which is slightly lower than the -0.54 eV in the screw dislocation II site. Among the empirical potentials, only EAM-Wang predicts the correct order of the binding energies between an H atom with screw dislocation at both sites. The BOP-Juslin and BOP-Li underestimate the binding energy at the screw dislocation II site. The EAM-Bonny potentials predict nearly zero binding energy at the screw dislocation I site, which disagrees with DFT. There is no information on the binding between H and dislocations in the training datasets of DP-WH. Nevertheless, with a good generalization ability, DP-WH predicts that the binding energies at the two sites in screw dislocations are -0.48 and -0.47 eV. The DP-WH slightly overestimates the DFT results but gives the correct order of the binding energies. As predicted by all the potentials, the binding of the edge dislocation with a H atom is always stronger than that of the screw dislocation. Up to now, there is no accessible DFT data on the binding between edge dislocations with H. However, the experimental thermal desorption spectrum (TDS) results reveal that a desorption peak between 0.8–0.9 eV is attributed to the desorption of H isotope from dislocations [14]. The predictions by DP-WH, EAM-Wang, and EAM-Bonny-2 are within the range. And the binding energy between the edge dislocation with an H atom is underestimated by the BOP potentials.

G. Interaction between TIS H atoms and the formation of H self-clusters

With the exposure to low-energy high-flux H plasma, high concentrations of H are introduced in W [21]. The interaction between H atoms at neighboring interstitial sites in bcc W has strong implications for the behaviors of H at high concentrations. The binding energy of a pair of interstitial H has been extensively studied [23,85,93]. The binding energy between H at neighboring TISs in the W lattices predicted by various potentials is shown in Table VI and is determined by

$$E_b^{2H} = E_{tot}(W2H) - N_W E(W) - 2\left(E_s^{tet} + \tfrac{1}{2}E(H_2)\right), \quad (10)$$

where $E_{tot}(W2H)$ is the total energy of the binding configuration of two neighboring H atoms in W.

The binding configuration is shown in Fig. 9(a). According to the present DFT calculation and the reference DFT results [23,85,93], a TIS H is strongly repulsive to another TIS H at the "A" site and the "B" site. Weakly attractive interactions are between a TIS H atom and another at the "D", "E", "F," and "H" sites. At the rest of the neighboring TISs, H atoms are

TABLE V. Binding energy of an H atom with screw/edge dislocations corresponding to the binding sites shown in Fig. 8 (unit: eV).

Potential	BOP-Juslin [32]	BOP-Li [33]	EAM-Bonny-1 [34]	EAM-Bonny-2 [34]	EAM-Wang [35]	Reference (DFT)	DP-WH
Screw dislocation I	−0.50	−0.41 [17]	0.0 [17]	0.0 [17]	−0.71	−0.55 [16]	−0.48
Screw dislocation II	−1.08	−1.03 [17]	−0.42 [17]	−0.66 [17]	−0.66	−0.54 [16]	−0.47
Edge dislocation	−1.61	−1.64 [17]	−0.63 [17]	−0.89 [17]	−0.84	−	−0.86

weakly repulsive. The BOP-Juslin predicts very strong repulsive interactions of an H atom with another at the "A" site but predicts attractive interactions at nearly all other neighboring TISs. The EAM-Bonny-1 and BOP-Li predict weak attractive interactions only at the "H" site and the EAM-Bonny-2 predicts repulsive H interactions at all neighboring TISs. The EAM-Wang potential is the only empirical potential fitted to
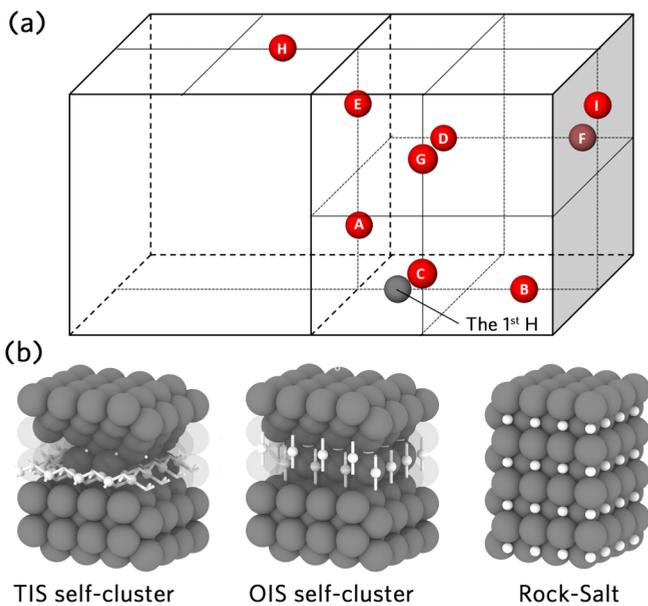

FIG. 9. The binding configurations corresponding to the binding energies shown in Table VI. (a) The configurations of interaction between a pair of H atoms at neighboring TIS in W. (b) The structures of the H self-clusters in W and the rock-salt structure.

the neighboring H binding energies and agrees well with DFT. DP-WH is also able to reproduce the interaction between neighboring TIS H atoms, but DP-WH slightly underestimates the binding energy at the "E" site.

It is proposed recently that high concentrations of H atoms can form a planar self-cluster by filling the TISs along a (001) plane [23]. H atoms in this TIS planar self-cluster are energetically more favorable than the random solution in TISs of bcc W. Additionally, H at OISs can also be stabilized by forming planar self-clusters [8]. It was also predicted by an MD simulation using EAM-Wang that the self-cluster comprised of multiple layers of OIS H along W (001) planes can spontaneously form [24], with W atoms in the self-clusters transforming into fcc structure via the reverse bain-path. This structure is identical to the rock-salt structure [8]. It was later revealed by DFT that the rock-salt structure is energetically more stable than both of the TIS/OIS planar H self-clusters [8].

The binding energies of H atoms in the self-clusters and the rock-salt structure predicted by various potentials are shown in Table VI, and are determined by

$$E_b^{\text{SC}} = \left[E_{\text{tot}}^{\text{SC}}(\text{WH}) - N_{\text{H}}\left(E_s^{\text{tet}} + \tfrac{1}{2}E(\text{H}_2)\right) - N_{\text{W}}E(\text{W}) \right]/N_{\text{H}}, \tag{11}$$

where $E_{\text{tot}}^{\text{SC}}(\text{WH})$ is the total energy of the self-cluster configuration, N_{H} is the number of H atoms in the self-cluster. SC stands for TIS self-cluster, OIS self-cluster, or the rock-salt W-H structure. DP-WH can accurately predict the H binding energies in the TIS/OIS self-clusters and shows a good agreement with the present DFT. The BOP-Juslin underestimates the binding energy of H in TIS self-clusters. OIS H self-cluster is unstable as described by BOP-Juslin. The BOP-Li

TABLE VI. Upper panel: Binding energy between a pair of H atoms in TISs of W. Lower panel: the binding energies of H in the TIS/OIS planar self-clusters and the rock-salt structures (unit: eV).

Binding Position	BOP-Juslin [32]	BOP-Li [33]	EAM-Bonny-1 [34]	EAM-Bonny-2 [34]	EAM-Wang [35]	Reference	DFT	DP-WH
A	4.213	0.639	0.470	0.340	0.512	0.47 [93], 0.45 [85]	0.442	0.383
B	−0.375	0.168	0.158	0.095	0.102	0.11 [93], 0.09 [85]	0.093	0.068
C	−0.461	0.138	0.063	0.042	0.026	0.03 [93], 0.01 [85]	0.019	0.003
D	−0.243	0.195	0.018	0.097	−0.022	−0.01 [93], −0.02 [85]	−0.014	−0.016
E	−0.196	0.160	0.020	0.187	−0.005	0.00 [93], −0.01 [85]	−0.011	−0.033
F	−0.174	0.120	0.020	0.132	0.003	0.00 [23], 0.01 [85]	−0.01	−0.008
G	−0.079	0.197	0.038	0.156	0.027	0.03 [93], 0.03 [85]	0.017	0.009
H	−0.054	−0.122	−0.008	0.026	−0.039	0.05 [23], 0.11 [93]	−0.013	−0.008
I	0.024	0.305	0.077	0.179	0.082	0.06 [85]	0.047	0.032
Self-cluster								
TIS planar	−1.030	0.114	0.068	0.358	−0.011	−0.38 [23] (with ZPE)	−0.245	−0.297
OIS planar	unstable	−0.190	0.003	0.442	−0.398	−	−0.009	−0.076
rock-salt	−1.404	−0.015	−0.174	−0.172	−0.693	−0.55 [8] (with ZPE)	−0.388	−0.449

TABLE VII. Arrhenius fits of diffusion data of hydrogen. Data of the empirical potentials are taken from Ref. [35]. As did by Ahlgren *et al.* [96], the pre-exponential factor from Ref. [96] is multiplied by $\sqrt{2}$ to convert from deuterium to hydrogen.

Potential	BOP-Juslin	BOP-Li	EAM-Bonny-1	EAM-Bonny-2	EAM-Wang	Experiment [95] (Corrected [96])	DP-WH
Pre-exponential factor (m^2/s)	3.9×10^{-8}	4.7×10^{-8}	1.1×10^{-7}	5.5×10^{-8}	5.3×10^{-8}	4.1×10^{-7} (1.6×10^{-7})	1.4×10^{-7}
Activation Energy (eV)	0.37	0.19	0.26	0.20	0.23	0.39 (0.25)	0.24

predicts that the binding energy of H in TIS self-clusters is positive, which is contrary to DFT results. The EAM-Bonny potentials predict that none of the planar H self-clusters are energetically more favorable than the solute H. Although EAM-Wang predicts negative binding energy for both TIS and OIS planar self-clusters, it significantly overestimates the binding energy of H in the TIS planar self-cluster and underestimates the binding energy of H in the OIS planar self-cluster. The binding energy of H in the rock-salt structure is around -0.55 eV per H atom by DFT with ZPE correction [8]. According to the present DFT calculation, the binding energy is around -0.39 eV without the ZPE correction. The stability of the rock-salt structure is significantly overestimated by BOP-Juslin, slightly overestimated by EAM-Wang, and underestimated by BOP-Li, EAM-Bonny-1, and EAM-Bonny-2. The rock-salt structure is not explicitly included in the training database of DP-WH, but the binding energy of H in the rock-salt structure predicted by DP-WH shows a satisfying agreement with DFT, which can be attributed to the generalization ability of DP-WH.

H. Diffusivity of dilute H at finite temperature

The diffusivity of H in W and its dependence on temperature is of central importance to many key issues in PFM studies, such as hydrogen retention [94]. It is also an important benchmark to validate the finite-temperature property of the potential. To simulate the thermal diffusion of H in W, we conduct a DPMD simulation in a $12 \times 12 \times 12$ bcc W supercell, containing 3456 W atoms and 1 H atom. In an *NPT* ensemble, the diffusion simulations run up to 500 ps with a time step of 0.1 fs. The pressure is kept zero during the simulation. The mean-square-displacement (MSD) H is measured at $T = 900$, 1200, 1500, and 1800 K to calculate H diffusivity. The diffusivity and temperature are then fitted to the Arrhenius law: $D = D_0 \exp(\frac{-E_A}{k_B T})$, to obtain the pre-exponential factor D_0 and the activation energy E_A. k_B is the Boltzman constant.

The D_0 and E_A predicted by DP-WH, empirical potentials, and experimental data are presented in Table VII. All potentials predict that the activation energy for H diffusion in W is close to the energy barrier calculated by NEB as shown in Sec. II C. According to the experimental result [95], the diffusivity of H is $D = 4.1 \times 10^{-7} \exp(\frac{-0.39 \text{ eV}}{k_B T})$ m^2/s. It seems that all potentials except BOP-Juslin underestimate the activation energy and the pre-exponential factors of H diffusion in W. However, it was then suggested by Ahlgren [96] *et al.* that the two low-temperature points in Ref. [95] are underestimated due to the trapping effect [97], and the true experimental activation energy should be 0.25 eV. In comparison with the corrected experimental data, all potentials predict reasonable activation energy. DP-WH has the most accurate description of the pre-exponential factors.

EAM-Bonny-1 also has good accuracy on the pre-exponential factors, which is underestimated by other empirical potentials.

IV. MD SIMULATIONS ON THE FORMATION OF H SELF-CLUSTERS

As DP-WH is designed to unveil the possible formation mechanism of intragranular H blister in W, the potential is used in an MD simulation based on a bcc W supercell with a high concentration of H.

We equilibrate a $5 \times 5 \times 20$ bcc W supercell by a 1.0 ns MD simulation at 600 K and zero pressure in the NPT ensemble, then randomly introduce H atoms into the TISs at a concentration of 10.7 at.%. A similar concentration (10 at.%) is used in Ref. [24], and was reported in experiments [98]. Then an NPT simulation of 2.5 ns is conducted at zero pressure with the temperature linearly decreasing from 600 to 400 K during the simulation. The decreasing temperature enables fast diffusion of the randomly distributed H atoms at the beginning and stabilizes the resultant structures at the end of the simulation. Moreover, the range of temperature in this simulation covers the temperatures used in recent experimental studies [2,3,14,15,20,21]. The mass of H atoms in all simulations is set to 3.0 a.m.u. to allow a timestep of up to 1 fs. The heavier mass only affects the dynamics of the H atoms, but not their equilibrium distribution in W.

The MD snapshots using DP-WH are shown in Fig. 10(a). The hydrogen atoms within a radius of 2.8 Å is considered as a cluster. The largest H cluster is colored red. At the beginning of the simulation, since H atoms in W have a migration barrier of only 0.2 eV, their fast diffusion is observed. Beginning from 0.50 ns, a large H cluster gradually forms into a planar region with a thickness of a few atom layers along the (001) plane. The planar self-cluster gradually grows by absorbing surrounding H atoms and fills the entire (001) plane. Further evolution of H atoms leads to the reduction of the width of the planar agglomeration region. As can be observed from the upper panels of Fig. 10(a), from 1.5 ns to 2.0 ns, the H atoms gradually fill all the "four-fold" TIS observed along the [001] view. It is accompanied by the H atoms in the planar agglomeration region gradually occupying the TIS on a single plane between two (001) W planes. The TIS planar self-cluster remains stable until the end of the simulation. In the enlarged figures in the rightmost panels of Fig. 10(a), we show the atomistic details of the TIS planar self-cluster at the end of the MD simulation. It can be observed that almost all H atoms in the self-cluster are located in the "fourfold" sites from the viewpoint along the [001] direction. These sites are TISs located between two adjacent (001) atomistic planes of W. From the perspective viewpoint, most of the H atoms in the self-clusters are located on the single plane normal to the [001] direction. Like previous DFT results [23], the TIS planar
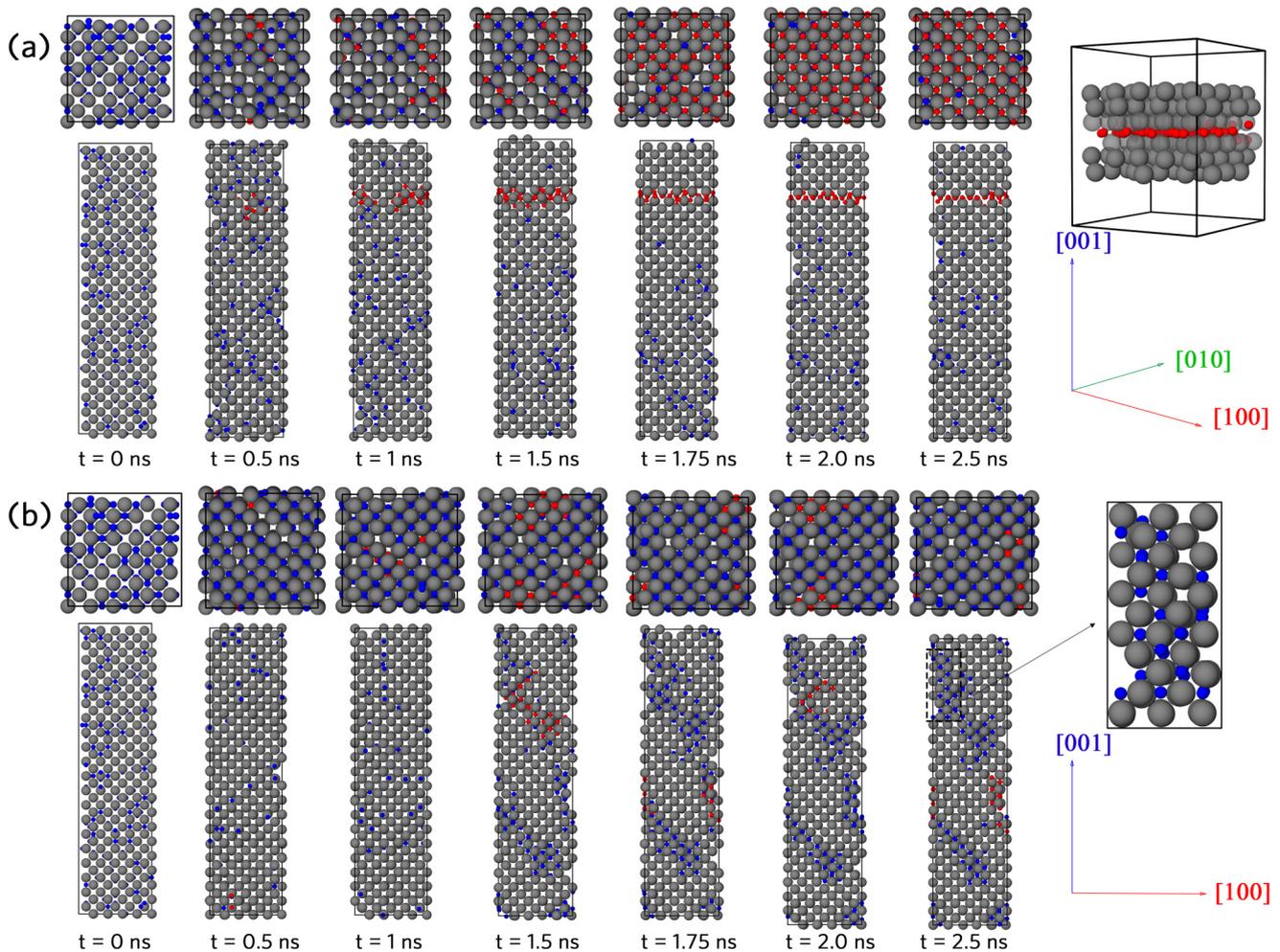
FIG. 10. MD simulations based on a W supercell with 10.7% at. H. Red atoms are H in the largest cluster. Blue atoms are the rest of the H atoms. (a) In the MD simulation using DP-WH, the largest H self-cluster at the end of the simulation is the TIS planar cluster between two (001) W planes. (b) In the MD simulation using EAM-Wang, the majority of clustered H atoms are located in OIS, forming either the OIS planar self-cluster or the small nucleate of the rock-salt structure. The dashed rectangle indicates the rock-salt structure formed in the MD simulation.

self-cluster of H widens the interplanar distances between the two adjacent W (001) planes. The formation of other forms of the self-clusters, such as the planar OIS self-cluster, or the rock-salt structures, are not observed in the MD simulation shown in Fig. 10(a). Remind that the rock-salt structure has a higher binding energy with the solute H than the TIS planar self-cluster, and DP-WH can accurately predict the H binding energy in the TIS/OIS planar self-cluster and the rock-salt structure. Thus the absence of the rock-salt structure in the MD simulation implies that the rock-salt structure is kinetically less feasible, which may be attributed to the strong lattice distortions of W to transform from bcc to fcc.

The spontaneous formation of the TIS planar self-cluster is not observed in the MD simulation using EAM-Wang [24] [see Fig. 10(b)]. The simulation protocols of the EAM-Wang is identical to those of the DP-WH. The behavior of H atoms in this MD simulation is featured by the formation of a H OIS self-cluster along the (110) plane and a small nucleate of the rock-salt structure. According to the comparisons in Table VI, we attribute the observed H clusters to the underestimated H binding energies of the OIS self-clusters and the rock-salt

structure and the overestimated binding energy of H in TIS self-clusters by the EAM-Wang potential.

To rule out the possibility that the formation of planar H self-cluster is due to the small-size effect, we run additional DPMD simulations in an $8 \times 8 \times 32$ supercell. The MD snapshots of this simulation are shown in Fig. 11. H atoms begin to cluster at 1.0 ns [Fig. 11(b)]. Once formed, the cluster accumulates nearby H atoms and grows into a planar shape with a thickness less than one lattice parameter of bcc tungsten, as shown in the left panel of Fig. 11(c). In this MD simulation, the H planar self-cluster grows along the (010) plane, instead of the (001) plane, whose cross-section is the smallest. The planar cluster remains stable and gradually grows till the end of the simulation. In the rightmost panel of Fig. 11(e) where we removed all atoms except the largest H cluster, a finite-sized planar-shaped H self-cluster is observed. It thus proves that the formation of the planar-shaped H self-cluster is not due to the small-size effect caused by the small cross-section. However, although all H atoms in the self-clusters are located in TISs, the finite-sized planar-shaped cluster along the (010) plane has not transformed into a single-plane structure.
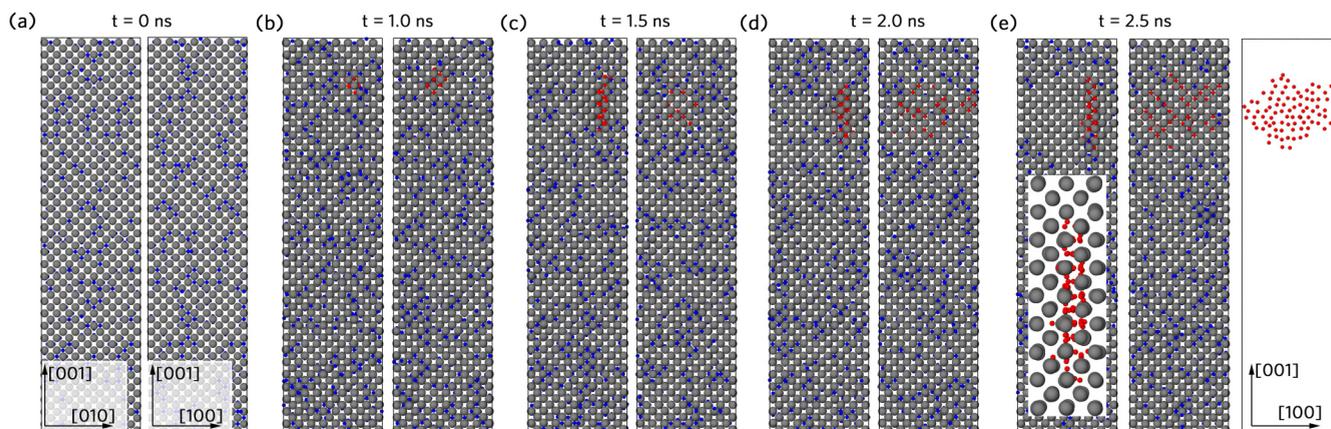
FIG. 11. DPMD simulation of H self-cluster formation in a $8 \times 8 \times 32$ BCC W supercell. Red atoms are H in the largest cluster. Blue atoms are the rest of H atoms. (a) MD snapshots at $t = 0$, (b) 1.0, (c) 1.5, (d) 2.0, and (e) 2.5 ns. The rightmost panel shows the [010] view of the finite-sized planar H cluster.

Instead, H atoms are located in three planes: two adjacent [010] W planes, and the plane between them. During the simulation, H atom jump among these planes but rarely jump out of the planar cluster. The interplanar distance between the two [010] W planes is also slightly widened due to the presence of H self-cluster, as shown in the enlarged figure of Fig. 11(e).

According to the theory of H embrittlement [99], H atoms between crystallographic planes play an important role in the hydrogen-enhanced decohesion [100]. The H TIS planar self-cluster is a typical form of planar segregation, which is assumed to reduce the interatomic bond strength, weakening the cohesion between the planes [100], and facilitating the development of cracks along the (001) cleavage plane. In recent experiments where W is exposed to the low-energy H plasma [2,3,21], the crack-shaped H blisters normal to the (001) planes are observed. This coincidence makes us believe that the H behaviors in the present MD simulation using DP-WH are strongly related to the crack-shaped H blisters observed in the experiments and implies that the nucleation of the intragranular H blister may begin with the formation of the TIS H self-cluster, instead of the energetically more favorable rock-salt structures.

V. CONCLUSIVE REMARKS

In this work, we develop an ML potential for W-H binary system, named DP-WH, based on the DP method. The potential is designed for conducting high-reliability atomistic simulations to unveil the formation mechanisms of intragranular H blisters in W under exposure to low-energy high-flux H plasma. Benefiting from the excellent fitting ability of the deep neural networks, DP-WH is trained to a large database constructed by DP-GEN, with the training error of energy being less than 5 meV/atom. According to our benchmark results, DP-WH shows good agreement with DFT in a wide range of properties including basic W properties, solute H properties in W, the interaction between H and W free surfaces, vacancies and dislocations, the interaction between

H atoms at neighboring TISs, and the formation energy of interstitial H self-clusters. The correct prediction of all the benchmarked properties with a single potential, which is by far challenging for the empirical potentials, is achieved in the benchmark results of DP-WH.

Up to now, the DP-WH is the only W-H binary potential that reproduces the H binding energy in the H TIS planar self-cluster, OIS planar self-cluster, and the rock-salt structure with the DFT accuracy. Using the DP-WH, the MD simulation results reveal that high concentrations of H atoms tend to form the TIS planar self-cluster normal to the {001} planes of W, instead of the OIS planar self-clusters, or the rock-salt structure reported by previous MD studies [24,35]. The TIS self-cluster is a typical form of planar H segregation, which is believed to be responsible for the H-induced decohesion and facilitate the development of cracks along the (001) cleavage plane [100]. The TIS self-cluster is thus highly likely to be the early nucleate of the crack-shaped H blister observed in recent experiments [2,3,21]. Therefore, we believe that the DP-WH is a good candidate potential for atomistic simulation that reveals the formation mechanisms of H blister in W under ITER working conditions.

In the future, we will use the DP-WH potential to conduct atomistic simulations at longer timescale and larger length scales and to further investigate the atomistic details of the development of the H self-cluster into H blisters.

Although the accuracy of DP-WH is comparable to the DFT on the benchmarked properties, it is unlikely that the current version of DP-WH can be used to simulate the processes like the collision cascades induced by high-energy protons [101], because the training dataset of DP-WH includes little information on very short-range W-H and W-W interactions. To improve the accuracy in short-range interaction, coupling with the short-range repulsion models [102] is suggested. The extension of the DP-WH potential to the simulation of primary irradiation damage is beyond the scope of the present work and will be investigated in future works.

The DP-WH model can be obtained in Ref. [103].

APPENDIX: THE DEEP POTENTIAL MODEL

The deep potential (DP) model [45,64] assumes the total energy of a system, represented by E, is the summation of the energy contributions of each atom in the system, denoted as E_i, i.e.,

$$E = \sum_{i=1}^{N} E_i, \tag{A1}$$

$$f_c(r) = \begin{cases} 1 & r < r_{cs} \\ u^3(-6u^2 + 15u - 10) + 1 & r_{cs} \leqslant r < r_c, \\ 0 & r_c \leqslant r \end{cases} \quad u = \frac{r - r_{cs}}{r_c - r_{cs}}. \tag{A3}$$

According to this definition, the switching function f_c smoothly decreases from 1 to 0 within the range $r_{cs} \leqslant r \leqslant r_c$.

The environment matrix $\mathcal{R}_i$ consists of N_m rows, where N_m represents the maximum number of neighbors any atom in the system can have. If the actual number of neighbors of atom i is less than N_m, the remaining entries in the environment matrix are filled with zeros.

In the DP models, the local environment matrix is first mapped onto a descriptor, denoted as $\mathcal{D}$. This descriptor is designed to retain the symmetries related to rotation, translation, and permutation of atoms. Subsequently, the descriptor $\mathcal{D}$ is further mapped onto the corresponding energy contribution E_i using a fitting net referred to as $\mathcal{F}$.

$$E_i = \mathcal{F}(\mathcal{D}(\mathcal{R}_i)). \tag{A4}$$

a. Descriptors. Two types of descriptors: the two-body embedding descriptor $\mathcal{D}_i^{(2)}$ [45] and the three-body embedding descriptor $\mathcal{D}_i^{(3)}$ [26] are used in this study.

In the construction of $\mathcal{D}_i^{(2)}$, the *generalized environment matrix* $\tilde{\mathcal{R}}_i$ is established. This matrix shares the same number of rows as the environment matrix, but includes an additional column consisting of $s(r_{ij})$. The jth row of $\tilde{\mathcal{R}}_i$ is defined as

$$\{\tilde{\mathcal{R}}_i\}_{j,\cdot} = s(r_{ij}) \times \left(1, \frac{x_{ij}}{r_{ij}}, \frac{y_{ij}}{r_{ij}}, \frac{z_{ij}}{r_{ij}}\right). \tag{A5}$$

The *two-body embedding matrix* $\mathcal{G}_i^{(2)}$ involves the embedding of two-atom distances. The embedding matrix $\mathcal{G}_i^{(2)}$ has N_m lines and M_2 columns. Its jth line is

$$\left(\mathcal{G}_i^{(2)}\right)_{j,\cdot} = \left(G_1^{(2)}(s(r_{ij}), Z_j), \cdots G_{M_2}^{(2)}(s(r_{ij}, Z_j))\right), \tag{A6}$$

where Z_j is the chemical species of a neighbor atom j. $G^{(2)}$, the *two-body embedding net*, is a fully connected deep neural

where N represents the number of atoms in the system, the energy contribution E_i is dependent on the local environment of each atom. The local environment matrix $\mathcal{R}_i$ for a central atom i consists of the relative positions of all its neighbors j within a specified cutoff radius r_c. The jth row of the environment matrix corresponds to the relative position of the central atom with respect to its jth neighbor, i.e.,

$$\{\mathcal{R}_i\}_{j,\cdot} = s(r_{ij}) \times \left(\frac{x_{ij}}{r_{ij}}, \frac{y_{ij}}{r_{ij}}, \frac{z_{ij}}{r_{ij}}\right), \tag{A2}$$

where $\boldsymbol{r}_i$ is the position of atom i, $\boldsymbol{r}_{ij} = \boldsymbol{r}_j - \boldsymbol{r}_i$ is the relative position between i and j, and (x_{ij}, y_{ij}, z_{ij}) denotes the three Cartesian coordinates of the vector $\boldsymbol{r}_{ij}$, and $r_{ij} = |\boldsymbol{r}_{ij}|$ stands for the distance between the neighbor j and the central atom i. In Eq. (A2), the term $s(r_{ij})$ is defined as $s(r_{ij}) = f_c(r_{ij})/r_{ij}$, where f_c is a switching function that smoothly varies from 1 at the smooth cutoff distance r_{cs} to 0 at the cutoff distance r_c. The construction of the switching function is

network, that maps the scaler $s(r_{ij})$ onto M_2 outputs. The embedding net $G^{(2)}$ has $m + 1$ layers, and can be written as

$$G^{(2)}(x) = \mathcal{L}_m^e \circ \mathcal{L}_{m-1}^e \circ \cdots \circ \mathcal{L}_1^e \circ \mathcal{L}_0^e(x), \tag{A7}$$

where $\circ$ denotes the function composition. The first hidden layer $\mathcal{L}_0^e$ takes a scalar as input and outputs a vector of size s_0. It is defined by

$$\mathcal{L}_0^e(x) = \tanh\left(x \cdot W_0^e + b_0^e\right), \tag{A8}$$

where $W_0^e \in \mathbb{R}^{s_0}$ are the weights, represented by a vector of size s_0, $b_0^e \in \mathbb{R}^{s_0}$ denote the biases and the activation function tanh applies to the vector $x \cdot W_0^e + b_0^e$ in a componentwise way. Other hidden layers are expressed as

$$\mathcal{L}_k^e(x) = (x, x) + \tanh\left(x \cdot W_k^e + b_k^e\right), \quad 1 < k \leqslant m, \tag{A9}$$

where (x, x) denotes the concatenation of two input vectors x. The weights are denoted as $W_k^e \in \mathbb{R}^{s_{k-1} \times s_k}$, and the biases are represented by $b_k^e \in \mathbb{R}^{s_k}$. We set the output size of the k-th hidden layer to be twice the input size, i.e., $s_k = 2s_{k-1}$. The output size of the final layer, denoted as s_m, is equal to M_2, which is the same as the number of columns in the embedding matrix $\mathcal{G}_i^{(2)}$. The parameters in the embedding net, $\{W_k^e, b_k^e\}_{k=0}^m$ will be trained together with those in the fitting net.

The two-body embedding descriptor $\mathcal{D}_i^{(2)}$ is constructed as

$$\mathcal{D}_i^{(2)} = \frac{1}{N_m^2} \left(\mathcal{G}_i^{(2),<}\right)^T \tilde{\mathcal{R}}_i (\tilde{\mathcal{R}}_i)^T \mathcal{G}_i^{(2)}, \tag{A10}$$

where the superscript T denotes the matrix transpose. The superscript $<$ on $\mathcal{G}_i^{(2),<}$ means that $\mathcal{G}_i^{(2),<}$ is a sub-matrix of $\mathcal{G}_i^{(2)}$, which takes the first $M^<$ columns of $\mathcal{G}_i^{(2)}$. The output of descriptor $\mathcal{D}_i^{(2)}$ is a matrix of shape $M^< \times M_2$, and is reshaped into a vector before it is passed to the fitting net.

The three-body embedding descriptor $\mathcal{D}_i^{(3)}$ is differentiated from the two-body embedding descriptor $\mathcal{D}^{(2)}$ by the embedding matrix used. The notation $(\theta_i)_{jk}$ is a shorthand representation for the elements of the product of the environment matrix $\mathcal{R}_i$ with its transpose.

$$(\theta_i)_{jk} \equiv \{\mathcal{R}_i(\mathcal{R}_i)^T\}_{jk} = s(r_{ij})s(r_{ik})\frac{\mathbf{r}_{ij} \cdot \mathbf{r}_{ik}}{r_{ij}r_{ik}}. \qquad (A11)$$

For any atom i, the three-body embedding tensor $\mathcal{G}_i^{(3)}$, is defined as

$$\left(\mathcal{G}_i^{(3)}\right)_{jk,\cdot} = \left(G_1^{(3)}((\theta_i)_{jk}, Z_j, Z_k), \cdots, G_{M_3}^{(3)}((\theta_i)_{jk}, Z_j, Z_k)\right). \qquad (A12)$$

The three-body embedding tensor $\mathcal{G}i^{(3)}$ is constructed by mapping the scalar $(\theta_i)_{jk}$ to a vector of dimension M_3 using the three-body embedding net $G^{(3)}$. The three-body embedding net has the same architecture as the two-body embedding net defined by Eqs. (A7)–(A9). The number of layers in the three-body embedding net is denoted by $m_t + 1$, and its trainable parameters are represented by $W_k^t, b_k^t{}_{k=0}^{m_t}$. In the embedding tensor $\mathcal{G}_i^{(3)}$, the first two indices j and k range from 1 to N_m.

The three-body embedding descriptor $\mathcal{D}_i^{(3)}$ is constructed as

$$\mathcal{D}_i^{(3)} = \frac{1}{N_m^2}\theta_i : \mathcal{G}_i^{(3)} = \frac{1}{N_m^2}\sum_{jk=1}^{N_m}(\theta_i)_{jk}\left(\mathcal{G}_i^{(3)}\right)_{jk}, \qquad (A13)$$

where : denotes the double contraction operation. The output of the descriptor is a vector of dimension M_3.

In practice, we usually use the hybridization of the two-body embedding descriptor $\mathcal{D}_i^{(2)}$ and the three-body

embedding descriptor $\mathcal{D}_i^{(3)}$, i.e.,

$$\mathcal{D}_i = \left(\mathcal{D}_i^{(2)}, \mathcal{D}_i^{(3)}\right), \qquad (A14)$$

where both $\mathcal{D}^{(2)}$ and $\mathcal{D}^{(3)}$ are treated as vectors, and the two vectors are concatenated to form a new vector. Thus the hybrid descriptor $\mathcal{D}_i$ has a total number of $M^< \times M_2 + M_3$ outputs. The parameters needed to be optimized in the training are $\{W_k^e, b_k^e\}_{k=0}^m, \{W_k^t, b_k^t\}_{k=0}^{m_t}$.

b. Fitting net. The fitting net is responsible for mapping the descriptor $\mathcal{D}_i$ to the energy contribution E_i of each atom i. It is a fully connected deep neural network with skip connections, consisting of l hidden layers.

$$\mathcal{F}(x) = \mathcal{L}_l^f \circ \cdots \circ \mathcal{L}_1^f \circ \mathcal{L}_0^f(x). \qquad (A15)$$

The layers of the fitting net are defined as

$$\mathcal{L}_0^f(x) = \tanh\left(x \cdot W_0^f + b_0^f\right), \qquad (A16)$$

$$\mathcal{L}_k^f(x) = x + \tanh\left(x \cdot W_k^f + b_k^f\right), \quad 1 \leqslant k < l, \qquad (A17)$$

$$\mathcal{L}_l^f(x) = x \cdot W_l^f + b_l^f. \qquad (A18)$$

In the first hidden layer $\mathcal{L}_0^f$, the weights W_0^f have a size of $(M^< \times M_2 + M_3) \times M_F$, and the bias b_0^f has a size of M_F. For layers $1 \leqslant k < l$, the input and output are vectors of the same length M_F, which allows for a skip connection as described in Eq. (A17). The weights and biases for these layers have a size of $M_F \times M_F$ and M_F, respectively. The output layer $\mathcal{L}_l^f$ is a linear transformation that maps a vector of length M_F to a scalar. The weights W_l^f form a vector of size M_F, and the bias is a scalar. All the parameters $\{W_k^f, b_k^f\}_{k=0}^l$ are optimized together with the parameters in the descriptors during training.

---

[1] T. Hirai, F. Escourbiac, S. Carpentier-Chouchana, A. Durocher, A. Fedosov, L. Ferrand, T. Jokinen, V. Komarov, M. Merola, and R. Mitteau, Phys. Scr. **2014**, 014006 (2014).

[2] W. Chen, X. Wang, Y.L. Chiu, T. Morgan, W. Guo, K. Li, Y. Yuan, B. Xu, and W. Liu, Acta Mater. **193**, 19 (2020).

[3] W. Guo, L. Ge, Y. Yuan, L. Cheng, S. Wang, X. Zhang, and G.H. Lu, Nucl. Fusion **59**, 026005 (2019).

[4] W. Chen, X. Xiao, B. Pang, S. Si, Y. Jia, B. Xu, T. Morgan, W. Liu, and Y. Chiu, J. Nucl. Mater. **522**, 11 (2019).

[5] T. Wang, X. Wang, X. Ma, L. Cheng, Y. Yuan, W. Guo, K. Xu, M. Liu, Z. Xie, L. Guo, and G.H. Lu, J. Mater. Sci. Technol. **139**, 245 (2023).

[6] Y. Jia, W. Liu, B. Xu, G.N. Luo, S. Qu, T. Morgan, and G. De Temmerman, J. Nucl. Mater. **477**, 165 (2016).

[7] J. Roth and K. Schmid, Phys. Scr. **2011**, 014031 (2011).

[8] J. Hou, X. S. Kong, C. S. Liu, and J. Song, Acta Mater. **201**, 23 (2020).

[9] S. Ryabtsev, Y. Gasparyan, M. Zibrov, A. Shubina, and A. Pisarev, Nucl. Instrum. Methods Phys. Res., Sect. B **382**, 101 (2016).

[10] P. Wang, Q. Cao, J. Hou, X.S. Kong, L. Chen, and Z. M. Xie, J. Nucl. Mater. **561**, 153576 (2022).

[11] Y.L. Liu, Y. Zhang, H. B. Zhou, G.H. Lu, F. Liu, and G. N. Luo, Phys. Rev. B **79**, 172103 (2009).

[12] N. Fernandez, Y. Ferro, and D. Kato, Acta Mater. **94**, 307 (2015).

[13] K. Ohsawa, J. Goto, M. Yamakami, M. Yamaguchi, and M. Yagi, Phys. Rev. B **82**, 184117 (2010).

[14] D. Terentyev, G. De Temmerman, T. Morgan, Y. Zayachuk, K. Lambrinou, B. Minov, A. Dubinko, K. Bystrov, and G. Van Oost, J. Appl. Phys. **117**, 083302 (2015).

[15] D. Terentyev, A. Dubinko, A. Bakaeva, G. De Temmerman, and B. Unterberg, Phys. Scr. **2017**, 014064 (2017).

[16] D. Terentyev, V. Dubinko, A. Bakaev, Y. Zayachuk, W. Van Renterghem, and P. Grigorev, Nucl. Fusion **54**, 042004 (2014).

[17] P. Grigorev, D. Terentyev, G. Bonny, E. E. Zhurkin, G. Van Oost, and J. M. Noterdaeme, J. Nucl. Mater. **465**, 364 (2015).

[18] P. Grigorev, D. Terentyev, V. Dubinko, G. Bonny, G. Van Oost, J. M. Noterdaeme, and E. E. Zhurkin, Nucl. Instrum. Methods Phys. Res., Sect. B **352**, 96 (2015).

[19] D. Terentyev, A. Dubinko, A. Bakaeva, and G. De Temmerman, Fusion Eng. Des. **124**, 405 (2017).

[20] A. Dubinko, D. Terentyev, A. Bakaeva, M. Hernández-Mayoral, G. De Temmerman, L. Buzi, J. M. Noterdaeme, and B. Unterberg, Appl. Surf. Sci. **393**, 330 (2017).

[21] W. Chen, X. Wang, K. Li, Y. Wang, T. Morgan, B. Xu, Y. Chiu, and W. Liu, Scr. Mater. **187**, 243 (2020).

[22] P. M. Anderson, J. P. Hirth, and J. Lothe, *Theory of Dislocations* (Cambridge University Press, 2017).

[23] J. Hou, X. S. Kong, J. Sun, Y. W. You, X. Wu, C. S. Liu, and J. Song, Nucl. Fusion **58**, 096021 (2018).

[24] R. Smirnov and S. Krasheninnikov, Nucl. Fusion **58**, 126016 (2018).

[25] O. Ogorodnikova, J. Roth, and M. Mayer, J. Nucl. Mater. **373**, 254 (2008).

[26] X. Wang, Y. Wang, L. Zhang, F. Dai, and H. Wang, Nucl. Fusion **62**, 126013 (2022).

[27] T. Ahlgren, K. Heinola, K. Vörtler, and J. Keinonen, J. Nucl. Mater. **427**, 152 (2012).

[28] U von Toussaint, S. Gori, A. Manhard, T. Höschen, and C. Höschen, Phys. Scr. **2011**, 014036 (2011).

[29] N. Juslin and B. Wirth, J. Nucl. Mater. **438**, S1221 (2013).

[30] Y. N. Liu, T. Wu, Y. Yu, X. C. Li, X. Shu, and G.H. Lu, J. Nucl. Mater. **455**, 676 (2014).

[31] X. Yu, F. Gou, and X. Tian, J. Nucl. Mater. **441**, 324 (2013).

[32] N. Juslin, P. Erhart, P. Träskelin, J. Nord, K. O. Henriksson, K. Nordlund, E. Salonen, and K. Albe, J. Appl. Phys. **98**, 123520 (2005).

[33] X. C. Li, X. Shu, Y. N. Liu, F. Gao, and G.H. Lu, J. Nucl. Mater. **408**, 12 (2011).

[34] G. Bonny, P. Grigorev, and D. Terentyev, J. Phys.: Condens. Matter **26**, 485001 (2014).

[35] L. F. Wang, X. Shu, G. H. Lu, and F. Gao, J. Phys.: Condens. Matter **29**, 435401 (2017).

[36] A. P. Thompson, L. P. Swiler, C. R. Trott, S. M. Foiles, and G. J. Tucker, J. Comput. Phys. **285**, 316 (2015).

[37] A. V. Shapeev, Multiscale Model. Simul. **14**, 1153 (2016).

[38] J. Behler and M. Parrinello, Phys. Rev. Lett. **98**, 146401 (2007).

[39] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, Phys. Rev. Lett. **104**, 136403 (2010).

[40] S. Chmiela, A. Tkatchenko, H. E. Sauceda, I. Poltavsky, K. T. Schütt, and K. R. Müller, Sci. Adv. **3**, e1603015 (2017).

[41] K. Schütt, P. J. Kindermans, H. E. S. Felix, S. Chmiela, A. Tkatchenko, and K. R. Müller, in *Advances in Neural Information Processing Systems* (Curran Associates, Inc., New York, 2017), pp. 992–1002.

[42] J. S. Smith, O. Isayev, and A. E. Roitberg, Chem. Sci. **8**, 3192 (2017).

[43] J. Han, L. Zhang, R. Car, and W. E, Commun. Comput. Phys. **23**, 629 (2018).

[44] L. Zhang, J. Han, H. Wang, R. Car, and W. E, Phys. Rev. Lett. **120**, 143001 (2018).

[45] L. Zhang, J. Han, H. Wang, W. A. Saidi, R. Car, and W. E, in *Advances of the Neural Information Processing Systems (NIPS)* (2018).

[46] M. A. Wood and A. P. Thompson, arXiv:1702.07042 [physics.comp-ph].

[47] L. Zhang, D. Y. Lin, H. Wang, R. Car, and W. E, Phys. Rev. Mater. **3**, 023804 (2019).

[48] W. Jiang, Y. Zhang, L. Zhang, and H. Wang, Chin. Phys. B **30**, 050706 (2021).

[49] Y. Wang, L. Zhang, B. Xu, X. Wang, and H. Wang, Model. Simul. Mater. Sci. Eng. **30**, 025003 (2022).

[50] T. Wen, R. Wang, L. Zhu, L. Zhang, H. Wang, D. J. Srolovitz, and Z. Wu, npj Comput. Mater. **7**, 206 (2021).

[51] R. Wang, X. Ma, L. Zhang, H. Wang, D. J. Srolovitz, T. Wen, and Z. Wu, Phys. Rev. Mater. **6**, 113603 (2022).

[52] K. C. Pitike and W. Setyawan, J. Nucl. Mater. **574**, 154183 (2023).

[53] L. Zhang, H. Wang, R. Car, and W. E, Phys. Rev. Lett. **126**, 236001 (2021).

[54] D. Marchand, A. Jain, A. Glensk, and W. A. Curtin, Phys. Rev. Mater. **4**, 103601 (2020).

[55] A. Bahramian, Surf. Interface Anal. **45**, 1727 (2013).

[56] T. Urbańczyk and J. Koperski, Spectrochim. Acta, Part A **189**, 502 (2018).

[57] L. Hu, R. Su, B. Huang, and F. Liu, arXiv:1901.01638.

[58] W. Li, Y. Ando, and S. Watanabe, J. Phys. Soc. Jpn. **86**, 104004 (2017).

[59] R. Kobayashi, D. Giofré, T. Junge, M. Ceriotti, and W. A. Curtin, Phys. Rev. Mater. **1**, 053604 (2017).

[60] N. Artrith and A. Urban, Comput. Mater. Sci. **114**, 135 (2016).

[61] N. Artrith and J. Behler, Phys. Rev. B **85**, 045439 (2012).

[62] Y. Zhang, H. Wang, W. Chen, J. Zeng, L. Zhang, H. Wang, and W. E, Comput. Phys. Commun. **253**, 107206 (2020).

[63] S. Plimpton, J. Comput. Phys. **117**, 1 (1995).

[64] H. Wang, L. Zhang, J. Han, and W. E, Comput. Phys. Commun. **228**, 178 (2018).

[65] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[66] G. Kresse and J. Furthmüller, Phys. Rev. B **54**, 11169 (1996).

[67] G. Kresse and J. Furthmüller, Comput. Mater. Sci. **6**, 15 (1996).

[68] D. P. Kingma and J. Ba, arXiv:1412.6980 [cs.LG].

[69] M. Abadi, in *Proceedings of the 21st ACM SIGPLAN International Conference on Functional Programming* (ACM, New York, 2016), pp. 1–1.

[70] P. W. Ma and S. L. Dudarev, Phys. Rev. Mater. **3**, 043606 (2019).

[71] A. Stukowski, Modell. Simul. Mater. Sci. Eng. **18**, 015012 (2010).

[72] M. C. Marinica, L. Ventelon, M. R. Gilbert, L. Proville, S. L. Dudarev, J. Marian, G. Bencteux, and F. Willaime, J. Phys.: Condens. Matter **25**, 395502 (2013).

[73] Z. J. Bergstrom, C. Li, G. D. Samolyuk, B. P. Uberuaga, and B. D. Wirth, J. Phys.: Condens. Matter **31**, 255002 (2019).

[74] J. Byggmästar, A. Hamedani, K. Nordlund, and F. Djurabekova, Phys. Rev. B **100**, 144105 (2019).

[75] W. J. Szlachta, A. P. Bartók, and G. Csányi, Phys. Rev. B **90**, 104108 (2014).

[76] H. Landolt and Börnstein, *Landolt-Börnstein: Numerical Data and Functional Relationships in Science and Technology: Refractory Metal Systems* (Springer, 1991).

[77] D. Nguyen-Manh, A. P. Horsfield, and S. L. Dudarev, Phys. Rev. B **73**, 020101(R) (2006).

[78] P. W. Ma and S. L. Dudarev, Phys. Rev. Mater. **3**, 013605 (2019).

[79] C. Becquart and C. Domain, Nucl. Instrum. Methods Phys. Res., Sect. B **255**, 23 (2007).

[80] J. Hao, S. Jin, G. H. Lu, and H. Xu, Comput. Mater. Sci. **184**, 109893 (2020).

[81] L. Vitos, A. Ruban, H. L. Skriver, and J. Kollár, Surf. Sci. **411**, 186 (1998).

[82] Y. Wang, Q. Li, C. Li, G. Shu, B. Xu, and W. Liu, J. Nucl. Mater. **496**, 362 (2017).

[83] phonolammps, code and data can be download at https://github.com/abelcarreras/phonolammps (2019).

[84] G. Henkelman and H. Jónsson, J. Chem. Phys. **113**, 9978 (2000).

[85] Y. L. Liu, Y. Zhang, G. N. Luo, and G. H. Lu, J. Nucl. Mater. **390-391**, 1032 (2009).

[86] D. F. Johnson and E. A. Carter, J. Mater. Res. **25**, 315 (2010).

[87] K. Heinola and T. Ahlgren, J. Appl. Phys. **107**, 113531 (2010).

[88] D. W. Brenner, Phys. Rev. B **42**, 9458 (1990).

[89] H. B. Zhou, S. Jin, Y. Zhang, G. H. Lu, and F. Liu, Phys. Rev. Lett. **109**, 135502 (2012).

[90] B. Fu, M. Qiu, L. Zhai, A. Yang, and Q. Hou, Nucl. Instrum. Methods Phys. Res., Sect. B **450**, 220 (2019).

[91] L. Yang and B. Wirth, J. Appl. Phys. **125**, 165105 (2019).

[92] M. Zibrov, S. Ryabtsev, Y. Gasparyan, and A. Pisarev, J. Nucl. Mater. **477**, 292 (2016).

[93] C. Becquart and C. Domain, J. Nucl. Mater. **386-388**, 109 (2009).

[94] T. Tanabe, Phys. Scr. **2014**, 014044 (2014).

[95] R. Frauenfelder, J. Vac. Sci. Technol. **6**, 388 (1969).

[96] T. Ahlgren and L. Bukonte, J. Nucl. Mater. **479**, 195 (2016).

[97] K. Heinola, T. Ahlgren, K. Nordlund, and J. Keinonen, Phys. Rev. B **82**, 094102 (2010).

[98] L. Gao, W. Jacob, U. von Toussaint, A. Manhard, M. Balden, K. Schmid, and T. Schwarz-Selinger, Nucl. Fusion **57**, 016026 (2017).

[99] V. Gavriljuk, V. Shyvaniuk, and S. Teus, in *Hydrogen in Engineering Metallic Materials: From Atomic-Level Interactions to Mechanical Properties* (Springer, 2022), pp. 201–274.

[100] A. A. Guzmán, J. Jeon, A. Hartmaier, and R. Janisch, Materials **13**, 5785 (2020).

[101] I. Ipatova, R. W. Harrison, P. Wady, S. Shubeita, D. Terentyev, S. E. Donnelly, and E. Jimenez-Melero, J. Nucl. Mater. **501**, 329 (2018).

[102] H. Wang, X. Guo, L. Zhang, H. Wang, and J. Xue, Appl. Phys. Lett. **114**, 244101 (2019).

[103] https://www.aissquare.com/models.