Nuclear Materials and Energy 22 (2020) 100724

Contents lists available at ScienceDirect

Nuclear Materials and Energy

journal homepage: www.elsevier.com/locate/nme

On the classification and quantification of crystal defects after energetic bombardment by machine learned molecular dynamics simulations

F.J. Domínguez-Gutiérrez[*,a], J. Byggmästar[b], K. Nordlund[b,c], F. Djurabekova[b,c], U. von Toussaint[a]

[a] Max-Planck Institute for Plasma Physics, Boltzmannstrasse 2, Garching 85748, Germany
[b] Department of Physics, University of Helsinki, Helsinki, PO Box 43, FIN 00014, Finland
[c] Helsinki Institute of Physics, Helsinki, Finland

ARTICLE INFO

Keywords:
Tungsten
MD simulations
Descriptor vectors
Machine learning
Material damage analysis
Gaussian approximation potentials

ABSTRACT

The analysis of the damage on plasma facing materials (PFM), due to their direct interaction with the plasma environment, is needed to build the next generation of nuclear fusion reactors. After systematic analyses of numerous materials over the last decades, tungsten has become the most promising candidate for a nuclear fusion reactor. In this work, we perform molecular dynamics (MD) simulations using a machine learned interatomic potential, based on the Gaussian Approximation Potential framework, to model better neutron bombardment mechanisms in pristine W lattices. The MD potential is trained to reproduce realistic short-range dynamics, the liquid phase, and the material recrystallization, which are important for collision cascades. The formation of point defects is quantified and classified by a descriptor vector (DV) based method, which is independent of the sample temperature and its constituents, requiring only modest computational resources. The locations of vacancies are calculated by the k-d-tree algorithm. The analysis of the damage in the W samples is compared to results obtained by Finnis–Sinclair and Tersoff–Ziegler–Biersack–Littmark potentials, at a sample temperature of 300 K and a primary knock-on atom (PKA) energy range of 0.5–10 keV, where a good agreement with the reported number of Frenkel pair is observed. Our results provide information about the advantages and limits of the machine learned MD simulations with respect to the standard ones. The formation of dumbbell and crowdion defects as a function of PKA energy were identified and distinguished by our DV method.

1. Introduction

In order to design the next generation of fusion reactors, the analysis of different types of crystal defects in plasma facing material (PFM) is necessary to better understand the effects of plasma irradiation on several physical and chemical properties of the materials. The materials of the first wall of a fusion machine is exposed to a hostile environment due to the plasma interaction, high temperatures, and energetic neutron irradiation, to mention a few [1,2]. Tungsten has emerged as the primary PFM due to its physical and chemical properties like low erosion rates, small tritium retention, and high melting point [3]. When an atom of the sample materials receives a higher kinetic energy than the threshold displacement energy [4,5] it can produce permanent point or extended defects [6,7]. For example, it has been observed that at high initial energies of the primary knock-on atom (PKA), defect clusters can be formed directly in crystalline materials [3,8], whereas simple point defects like self-interstitial-atoms (SIA) are commonly formed at low impact energies. Specifically, in body-center-cubic (bcc) W lattice samples, SIA's are commonly observed in an atomic arrangement known as dumbbells and crowdions [9,10].

Molecular dynamics (MD) simulations are frequently used to model collision cascades during neutron bombardment in fission or fusion reactors [11]. The better the interatomic potentials on which the MD simulation is based are, the better the obtained results can predict the induced damage in crystalline materials. Recently, machine learning potentials have been proposed to improve the accuracy on the modelling of point defects formation in damaged PFMs. W. Szlachta et al. [12] recently developed the interatomic potential based on the Gaussian Approximation Potential (GAP) framework [13,14] to investigate tungsten in the bcc crystal phase and its defects. However, this GAP potential cannot be utilized in the study of material damage by neutron bombardment due to the lack of information of the repulsive region to treat short distance interactions realistically. In the current work, we take into use a very recently developed machine learning interatomic potential for tungsten based on GAP [15], that includes relevant physical properties for collision cascades simulations in the training data

* Corresponding author.
E-mail address: javier.dominguez@ipp.mpg.de (F.J. Domínguez-Gutiérrez).

https://doi.org/10.1016/j.nme.2019.100724
Received 4 November 2019; Accepted 16 December 2019
Available online 20 December 2019

set. Usually, the resulting damaged samples obtained from MD simulations are analyzed by Wigner-Seitz cell or Voronoi diagram methods to quantify the number of Frenkel pairs (interstitials and vacancies) formed after the cascade [16–20]. Nevertheless, formation of complex defects and thermal motion have not been well studied or modeled by these methods [8]. Therefore, we analyze the damaged material by the descriptor vector (DV) based method, which is developed by us [21] and is capable to assign a probability of being a point defect to each atom in the sample.

Our paper is organized as follows: in Section 2 we discuss the theory to develop the machine learned (ML) potential [15] for W, and the DV based method to identify, classify and quantify standard and uncommon crystal defects [21]. In Section 3, we present the analysis of point defects and more complex defects formation in W samples at the PKA range of $0.5 - 10$ keV. In order to provide an insight of the limitations and advantages of our new ML interatomic potential, we compare our results to those obtained by commonly used Finnis-Sinclair (Embedded Atom Method (EAM)-like [22] and Tersoff-ZBL [23] potentials. Finally, in Section 4, we provide concluding remarks.

2. Theory

2.1. Machine learned interatomic potential

Machine Learned (ML) interatomic potentials are not restricted to an analytical form and can be systematically improved towards the accuracy of the training data set. In order to model collision cascades, the ML potential must be able to treat realistic short-range dynamics defined by its repulsive part. In addition, the structure of the liquid phase and re-crystallization process (including elastic energies) should be well described, to accurately emulate atomic mixing together with defect creation and annihilation during the collision cascade. In this work, we use a ML potential recently developed [15] within the Gaussian Approximation Potential (GAP) framework [13,14]. Here, the total energy of a system of N atoms is expressed as

$$E_{\text{tot}} = \sum_{i<j}^{N} V_{\text{pair}}(r_{ij}) + \sum_{i}^{N} E^{i}_{\text{GAP}}, \tag{1}$$

where V_{pair} is a purely repulsive screened Coulomb potential, and E_{GAP} is the machine learning contribution. E_{GAP} is constructed using a two-body and the many-body Smooth Overlap of Atomic Positions (SOAP) descriptor [13], and given by

$$\begin{aligned} E^{i}_{\text{GAP}} &= \delta^{2}_{2\text{b}} \sum_{j}^{M_{2\text{b}}} \alpha_{j,2\text{b}} K_{2\text{b}}(\vec{q}_{i,2\text{b}}, \vec{q}_{j,2\text{b}}) \\ &+ \delta^{2}_{\text{mb}} \sum_{j}^{M_{\text{mb}}} \alpha_{j,\text{mb}} K_{\text{mb}}(\vec{q}_{i,\text{mb}}, \vec{q}_{j,\text{mb}}), \end{aligned} \tag{2}$$

where $\delta^{2}_{2,\text{mb}}$ is the standard deviation of the Gaussian process that sets the energy ranges of the training data, which contains the energy information and is chosen by systematic convergence tests [15]; $K_{2,\text{mb}}$ is the kernel function representing the similarity between the atomic environment of the ith and jth atoms; α is a coefficient obtained from the fitting process; and $\vec{q}$ is the normalized descriptor vector of the local atomic environment of the ith atom (See Section 2.3). In the computation of the ML potential the descriptors for two bodies, 2b, is utilized to take into account most of the interatomic bond energies, while the atomic environment due to the many-body, mb, contributions are treated by the SOAP descriptor.

The GAP method has been applied by Szlachta et al. [12] to develop the ML interatomic potential for tungsten to reproduce the properties of screw dislocations and vacancies. However, this potential lacks of information for the structures relevant modeling of collision cascades (See Appendix A) such as self-interstitial atom formation, the liquid phase, and realistic repulsive interactions. The new ML potential developed in Ref. [15] included these types of structures, and is therefore suitable for performing MD simulations of material damage due to neutron bombardment, for example. The elastic response of bcc W is also included in the training data of the new ML potential, which is an important property to treat the recrystallization of the highly affected target region during a collision cascade. More details about the development of this new potential can be found in Ref. [15].

2.2. MD simulations

In order to explore the advantages and limitations of our new ML interatomic potential, MD simulations are performed to emulate a neutron bombardment process by using both the ML potential and well known analytical interatomic potentials [24]. Then, a comparison between the obtained results, under the same numerical and physical conditions, is carried out. For all potentials, we first define a simulation box as a pristine W lattice sample based on a body-centered-cube (bcc) unit cell with a lattice constant of $a = 3.16$ Å [25]. Then, the sample is subjected to a process of energy optimization and thermalization to 300 K using the Langevin thermostat, with the time constant of 100 fs. [26]. Most of the experiments of tungsten damaging are done at room temperature, which is used in our simulation to perform MD simulations as close as possible to them [27,28].

Every MD simulation starts by assigning a chosen PKA energy in a range of 0.5–10 keV to a W atom, which is located at the center of the numerical cell. We use ten velocity directions for each PKA: ⟨001⟩, ⟨011⟩, ⟨111⟩, and 7 cases for $\langle r_1 r_2 r_3 \rangle$, where r_i are random numbers uniformly distributed in an interval of [0,1]. We utilize the Velocity Verlet integration algorithm to model the collision dynamics, which is performed for 10 ps, followed by an additional relaxation run for 5 ps. In Tab. 1, we present the dimensions of the numerical boxes as the number of unit cells with a side length of a; the number of W atoms in each numerical box; and the time step used in the simulations as a function of the PKA. The MD simulations were done in a traditional desktop computer by using the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) [29] with the Quantum mechanics and Interatomic Potential package (QUIP) [30] that is used as an interface to implement machine learned interatomic potentials based on GAP [13]. We also perform MD simulation by using the reactive interatomic potential for the ternary system W-C-H by Juslin et al. [23] referred as J-T-ZBL in our work, which is based on an analytical bond-order scheme. This potential has been used to study neutron damage in polycrystalline tungsten [31], trapping and dissociation processes of H in tungsten vacancies [32]; and cumulative bombardment of low energetic H atom of W samples for several crystal orientations [33]. In addition, we use a second standard interatomic potential for MD simulations based on the embedded-atom method-like Finnis-Sinclair model with modification by Ackland et al. [22], and the repulsive potential fit reported in Ref. [34]. This MD potential is denoted as AT-EAM-FS in this work and has been applied to study Frenkel pair formation as a function of the PKA in pristine tungsten [35] and self-sputtering of tungsten in a wide impact energy range [36].

Table 1
Size of the numerical boxes based on a bcc unit cell as a function of the impact energy (PKA velocity), which is used in the MD simulations. The box size is reported as the number of unit cells with side length of $a = 3.16$ Å, that is the lattice constant of W at 300 K.

PKA	Num. atoms	Box size [a]	Δt (ps)
0.5	35 152	(25, 25, 25)	10^{-3}
1	35 152	(25, 25, 25)	10^{-3}
2	35 152	(25, 25, 25)	10^{-3}
5	124 722	(38, 38, 40)	10^{-4}
10	235 008	(47, 47, 50)	10^{-4}

2.3. Descriptor vectors based method

The quantification and classification of point defects in a damaged sample starts by computing the descriptor vector (DV) of all the atoms in the material sample. The DV of the ith atom of the sample, $\vec{\xi}^{\,i}$ (defined below), is invariant to rotation, reflection, translation, and permutation of atoms of the same species, but sensitive to small changes in the local atomic environment [14]. It can be considered as a fingerprint of the particular atomic environment of an ith atom, which is expressed by a sum of truncated Gaussian density functions as [14],

$$\rho^i(\vec{r}) = \sum_{j}^{\text{neigh.}} \exp\left(-\frac{|\vec{r} - \vec{r}^{\,ij}|^2}{2\sigma^2_{\text{atom}}}\right) f_{\text{cut}}(|\vec{r}^{\,ij}|), \tag{3}$$

where $\vec{r}^{\,ij}$ is the difference vector between the atom positions i and j. The term σ^2_{atom} defines the broadening of the atomic position, which is set according to the lattice constant of the sample. Finally, $f_{\text{cut}}(|\vec{r}^{\,ij}|)$ is a smooth cutoff function, that limits the considered neighborhood of an atom. The function $\rho^i(\vec{r})$ can also be defined in terms of expansion coefficients, c_{nlm}, that corresponds to the ith-atom in the lattice as [12],

$$\rho^i(\vec{r}) = \sum_{nlm}^{NLM} c^{(i)}_{nlm} g_n(r) Y_{lm}(\hat{r}), \tag{4}$$

where $c^{(i)}_{nlm} = \langle g_n Y_{lm} | \rho^i \rangle$, $\hat{r}$ is a unit vector in the $\vec{r}$ direction, $g_n(r)$ is a set of orthonormal radial basis functions $\langle g_n(r) | g_m(r) \rangle = \delta_{nm}$, and $Y_{lm}(\hat{r})$ are the spherical harmonics with the atom positions projected onto a unit sphere. Thus, Eq. (4) is averaged over all possible rotations to be invariant against rotations, by the product of the $c_{n'lm}$ with its complex conjugate coefficient c^*_{nlm}, summed over all m. Then the DV of the ith atom, $\vec{\xi}^{\,i}$, is defined as [12]

$$\vec{\xi}^{\,i} = \left\{ \sum_{m} (c^i_{nlm})^* c^i_{n'lm} \right\}_{n,n',l}, \tag{5}$$

where each component of the vector corresponds to one of the index triplets $\{n, n', l\}$.

In this work, we refer to the DV as the normalized vector $\vec{q}^{\,i} = \vec{\xi}^{\,i}/|\vec{\xi}^{\,i}|$ for the local environment of the ith atom. Depending on the choice of the expansion orders in Eq. (4) for the spherical harmonics and the radial basis functions, the number of components of $\vec{q}^{\,i}$ varies. In order to compute the DVs, we used the SOAP descriptor tool in QUIP with a cutoff distance of 3.1 Å, which allows us to describe the local atomic environment and to identify lattice distortions and defects at the first nearest neighbors. This parameter is chosen according to the W lattice constant at 0 K, 3.16 Å. The values for the spherical harmonics are $N = 4$, and $L = 4$ (with $-L \leq m \leq L$) which results in a vector with $k = 51$ $(0 \cdots 50)$ components.

2.4. Identification of point defects

The difference of two local environments of the ith atom and jth atom can be computed by calculating the distance, d, between two DVs, $d = d(\vec{q}^{\,i}, \vec{q}^{\,j})$. However, we keep in mind that some vector components may be more fluctuating than others and an appropriate measure to compare the DVs is done as follows: We define a small simulation box with hundreds of W atoms to apply a Langevin thermostat, which generates a thermalized tungsten bcc lattice without defects to a desire sample temperature, $T = 300$ K, in our case. Then, we compute the DVs of all the W atoms to calculate a mean reference DV, $\vec{v}(T) = \frac{1}{N}\sum_{i=1}^{N} \vec{q}^{\,i}(T)$, for defect-free environments; as well as the associated covariance matrix, $\Sigma(T)$, which highly depends on the temperature of the sample. Therefore, the distance difference between a thermalized atomic environment and a damaged one is computed by the Mahalanobis measure as [37]

$$d^M(T)(\vec{q}^{\,i}, \vec{v}(T)|\Sigma) = \sqrt{(\vec{q}^{\,i} - \vec{v}(T))^{\mathrm{T}} \Sigma^{-1}(T)(\vec{q}^{\,i} - \vec{v}(T))}, \tag{6}$$

where $(\vec{q}^{\,i} - \vec{v}(T))^{\mathrm{T}}$ is the transpose vector. This provides us information about the presence of an unexpectedly large distortion of the local environments [21]. In order to detect common types of defects, a similar approach has been chosen at $T = 0$ K. A small simulation box containing the defect of interest (e.g. an interstitial, an atom next to a single vacancy) is prepared and the DVs of all the atoms are then computed, subsequently acting as a fingerprint for this specific type of defect (see Section 3).

The definition of the distance difference, d^M(T), between two local atomic environments provides a probabilistic interpretation of the obtained results. Thus, the probability, $P(\vec{q}^{\,i}|\vec{v}(T))$, of an ith atom being in a locally undistorted lattice can be computed using

$$P(\vec{q}^{\,i}|\vec{v}(T)) = P_0 \exp\left[-\frac{1}{2} d^M(T)^2\right], \tag{7}$$

where P_0 is the normalization factor. Therefore, all the atoms in a damaged material sample have an assigned probability of being in a lattice position and atoms with the lowest probability will be labeled as point defects in the sample, following a type defect classification [21]. Here, atoms with lower probability define the distorted region around the permanent defects, which provides a good visualization of the damaged in the material

3. Results

In order to test the advantages and limitations of the new ML interatomic potential [15] to traditional ones like J-T-ZBL and AT-EAM-FS potentials; we performed MD simulations at a primary knock-on atom (PKA) energy of 1 keV, in the ⟨001⟩ velocity direction, with a sample temperature of 300 K. Then, the damaged sample was analyzed by applying our DV based method to identify the formation of standard point defects (e.g. interstitial, vacancies) and unforeseen point defects [21]. The comparison of our results with those obtained by using AT-EAM-FS [22] and J-T-ZBL [23] potentials, under the same numerical and physical conditions, serves as a test for our ML potential.

In Fig. 1, we present the distance difference between the reference DV vector of a W atom in an interstitial site to all the W atoms in the damaged sample after collision cascade. These results are obtained by Eq. (6), considering a reference DV of an interstitial site at $T = 0$ K, $\vec{v}_{\text{I}}(T = 0)$. We observe that the shape of the histograms in Fig. 1(a) and

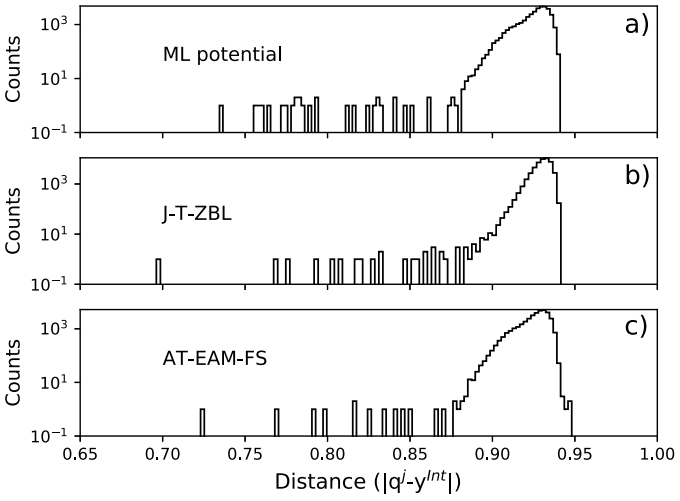

Fig. 1. Histogram of the distance difference between the interstitial DV and the W atoms in the damaged sample after relaxation process. MD simulation were performed by the ML potential in a), J-T-ZBL in our previous work [21], and the AT-EAM-FS potential in c).

Fig. 1(b) are similar, however the results for the ML potentials present a clear distance gap between the W atoms in a lattice position and the ones in an interstitial site and the distorted region, at a distance difference of 0.88. This makes the identification of the W atoms in the vicinity of the interstitial atoms simple, and works for a good test for this new MD potential. The W atoms can recover to their lattice positions (material re-crystallization) after collision cascade during the relaxation process due to the elastic energy. The histogram reported in Fig. 1(b) and in our previous work [21] shows a narrow shape and the W atoms that are in the vicinity of the interstitial atoms are defined as atoms in a distorted region. Finally, our results are compared to those computed by the Wigner-Seitz cell analysis [16], which is implemented in OVITO [19]. Although, this analysis is limited by the definition of spatial region around the W atoms, we have a good agreement by finding the two Frenkel pairs (single vacancy and a single self-interstitial atom) formed at the location in the damaged W lattice. These W atoms, called SIA, have the lowest probability to be at a lattice position [21]. However, our method is capable to identify the W atoms that are in the vicinity of the SIAs. This visualization can be done *via* OVITO and choosing different distance thresholds manually.

Since the formation of interstitials is well modeled by the new ML interatomic potential, it is interesting to investigate the formation of different point defects as a function of the simulation time by considering ten different velocity directions (i.e. 10 MD simulations). In Fig. 2 we present the quantification and classification of material defects formation during collision dynamics (0–10 ps) by using the new ML potential, J-T-ZBL, and AT-EAM-FS in the MD simulations. The defects remain in the material during the relaxation process (10–15 ps). The total number of SIA and atoms in their distorted region (Fig. 2a) after the collision cascade presented by the ML potential is similar to those performed with the standard potentials, under the same numerical parameters. Nevertheless, W atoms tend to adapt to their interstitial site gradually during the collision cascade simulation. While the J-T-ZBL and AT-EAM-FS simulations show more W atoms as interstitial in the time interval of 1–3 ps. In the same Fig. 2(a), we add a fitting curve to the number of SIA and atoms in the distorted region as: $f(t) = f_0 \exp(-\alpha t) + \beta$, with $f_0 = 31$, $\alpha = 0.65$ ps^{-1}, and $\beta = 17.0$. The second classification is defined as a W atom next to a single vacancy, $\overrightarrow{v_V}(T = 0)$, in a bcc unit cell, at 0 K [21]. In Fig. 2(b) we notice that the number W atoms that belong to this classification is similar for the ML potential and AT-EAM-FS results. A third classification is a type-A defect, $\overrightarrow{v_A}(T = 0)$, which is defined as a W atom in the vicinity of a split vacancy or di-vacancy [21]. In Fig. 2(d) the formation of this point defect is observed at the beginning of the MD simulations by the three interatomic potentials, however only the ML potential does not preserve this defect after collision cascade. It is known that this defect is energetically unstable according to DFT calculations [38]. Therefore, it is important to notice that the formation of a type-A defect after the annealing process is not observed in MD simulations that uses the machine learned potential [15]. Providing an advantage over the standard MD potentials.

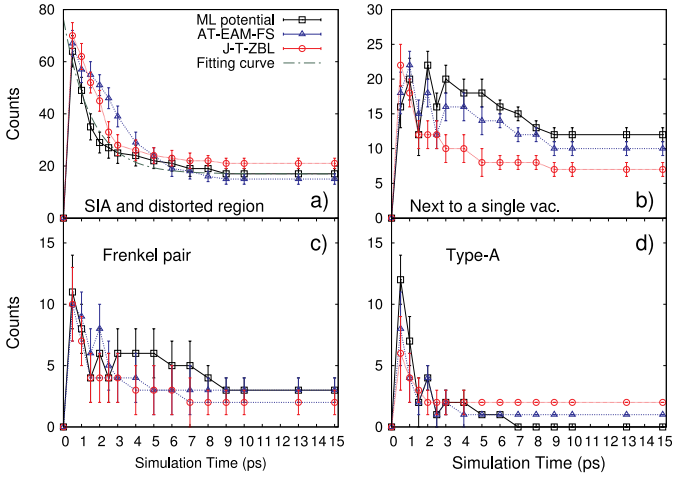

Fig. 2. (Color online) Quantification and classification of crystal defects formation as a function of the time for 10 MD simulations. We follow the formation of point defects during and after collision dynamics obtained by using the new ML potential, J-T-ZBL, and AT-EAM-FS. We add a fitting curve to the SIA and distorted region counting as $f(t) = f_0 \exp(-\alpha t) + \beta$ in a).

In order to identify the location of vacancies and to obtain a visualization of their spatial volume, we first define a sampling grid by 200 points of lateral dimension and a spatial step of 0.5 Å that fits the numerical box of the damage sample. Then the nearest neighbor distance between the spatial position of the atoms and the grid points can be calculated by a k-d-tree algorithm [21,39] with the KDTREE code ver. 2 [40]. Then, squared distances larger than the lattice constant, 3.16 Å for W, are used to identify the spatial volume around a vacancy in the damaged material [21]. Sampling grid points with the largest distances are associated to the location of a single vacancy. These results are presented in Fig. 2(c) as Frenkel pairs. We notice that at the beginning of the MD simulations the number of vacancies and W atoms next to single vacancy have the same trend (Fig. 2b–c). Then, the ML potential and AT-EAM-FS potential reach an agreement for the number of vacancies formed, however the J-T-ZBL results present a lower number of vacancies in the sample. The detailed analysis of the material damage due to neutron bombardment by the DV based method and k-d-tree algorithm show that standard interatomic potentials have some unexpected errors during the modeling of collision cascades. We provide the visualization of the point defects formation during collision dynamics in the supplementary material. In conclusion, the ML potential shows its first advantage at a PKA of 1 keV, regardless of the good agreement of the number of point defects with the standard potentials.

3.1. Classification and quantification of crystal defects as a function of the PKA energy

We calculate the number of crystal defects at different PKAs as a test of our ML potential and DV based method. For this, we perform MD simulations for an impact energy range of 0.5–10 keV at 10 different velocity directions to count the remaining crystal defects at the last frame of the simulation. Collision cascade is performed for 10 ps and the MD simulation run for 5 ps to model the relaxation process. However, at 10 keV of PKA simulations are performed for 20 ps due to the longer lifetime of the collision cascade at this highest energy. In Fig. 3, we report (a) the average of the number of Self-Interstitial-Atom (SIA) and atoms in its distorted region; (b) W atoms next to a single vacancy; (c) Frenkel pair formation and (d) type-A (W atom in the vicinity of a split-vacancy or di-vacancy [21]) defects, all as a function of the impact energy. A fitting curve to the energy dependence of point defects formation in different metals has been proposed by Bacon et al. and Stoller et al. [41,42] to be Counts $= \alpha E_{\mathrm{PKA}}^{\beta}$; where E_{PKA} is the PKA energy, and α and β are fitting parameters. Recently, Nordlund et al. have used this fitting law in an analysis of realisitc atomic displacement simulations with physically realistic material damage [43]. We apply here the damped least-square method to fit this functional form to our results for the number of atoms in a distorted region, obtaining the fitting parameters as $\alpha = 18.49$ and $\beta = 0.553$, with a correlation factor of 0.99. Besides that, a Frenkel pair is a typical defect, where the formation of an SIA is related to the creation of a vacancy, thus we can compare to the fitting curve reported by Setyawan et al. [25]. This fitting law is expressed as a function of a reduced cascade energy as: $0.49(E_{\mathrm{PKA}}/E_{\mathrm{d}})^{0.74}$ at a sample temperature of 300 K with $E_{\mathrm{d}} = 128$ eV. Although the authors performed MD simulations by using corrected semi-empirical potentials by Finnis and Sinclair [22], this fitting law is in good agreement with our results. In the Table 2, we report the

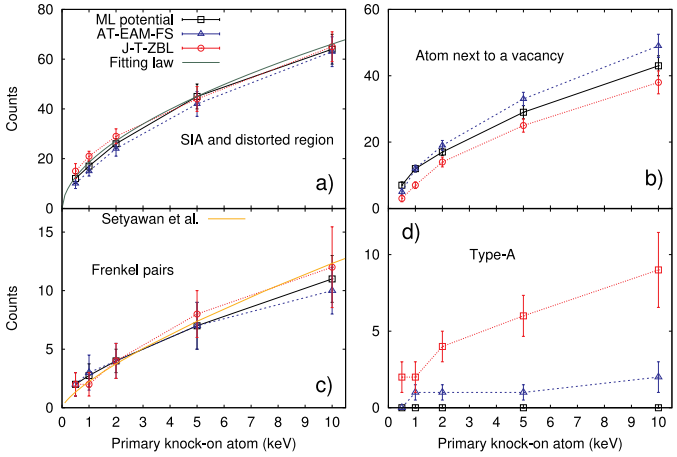

Fig. 3. (Color online) Number of crystal defects as a function of the PKA energy, E_{PKA}. We include a fitting curve to the average number of SIA and atoms in its distorted region as: Counts $= \alpha E_{PKA}^{\beta}$ [41] with $\alpha = 18.49$ and $\beta = 0.553$ with a correlation factor of 0.99. Results at 1 keV were obtained in our previous work [21]. The number of vacancies are in good agreement with the reported results by Setyawan et al. [25] .

Table 2
Average number of point defects and vacancies as a function of the PKA, which are identified by our DV based method. Interstitials are counted at the total number of SIA + atoms in its distorted region. SIA are identified as W atoms with the highest probability to be in an interstitial site, reported into parentheses. Total number of defects is calculated as: Interstitials + Next to vac. + type-A .

	PKA (keV)				
Defect	0.5	1	2	5	10
			ML potential		
Interstitial	12 ± 2	17 ± 2	26 ± 3	45 ± 5	64 ± 6
	(2 ± 1)	(3 ± 1)	(4 ± 1)	(7 ± 2)	(11 ± 2)
Next to vac.	7 ± 1	12 ± 1	17 ± 1	29 ± 2	43 ± 3
type-A	0	0	0	0	0
Total	19 ± 2	30 ± 2	43 ± 3	74 ± 5	107 ± 7
Vacancy	2 ± 1	3 ± 1	4 ± 1	7 ± 2	11 ± 2
			J-T-ZBL		
Defect	**0.5**	**1**	**2**	**5**	**10**
Interstitial	15 ± 2	21 ± 2	29 ± 3	44 ± 5	65 ± 6
	(2 ± 1)	(2 ± 1)	(4 ± 1)	(8 ± 2)	(12 ± 3)
Next to vac.	3 ± 1	7 ± 1	14 ± 1	25 ± 2	38 ± 3
type-A	2 ± 1	2 ± 1	4 ± 1	6 ± 1	9 ± 2
Total	20 ± 2	30 ± 2	47 ± 3	75 ± 5	112 ± 7
Vacancy	2 ± 1	2 ± 1	4 ± 1	8 ± 2	12 ± 3
			AT-EAM-FS		
Defect	**0.5**	**1**	**2**	**5**	**10**
Interstitial	10 ± 2	15 ± 2	24 ± 3	42 ± 5	63 ± 6
	(2 ± 1)	(3 ± 1)	(4 ± 1)	(7 ± 2)	(10 ± 2)
Next to vac.	5 ± 1	10 ± 1	17 ± 2	33 ± 2	49 ± 4
type-A	0	1 ± 1	1 ± 1	1 ± 1	2 ± 1
Total	15 ± 2	28 ± 2	48 ± 4	76 ± 5	114 ± 7
Vacancy	2 ± 1	3 ± 1	4 ± 1	7 ± 2	10 ± 2

average of the number of crystal defects as a function of the PKAe energy. Interstitials are counted as SIA and atoms in its local neighborhood, and only those atoms with the maximum probability are counted as interstitial sites (Frenkel pair) reported into parentheses. In good agreement with the number of vacancies formation, quantified and identified by the k-d-tree method. The total number of defects is defined as: Total = Interstitials + Next to vac. + type-A.

There is a couple of common and complex materials defects that are formed by several atoms in their interstitial sites. A dumbbell defect, where two atoms share a lattice site, is the most likely material defect to be found in a bcc unit cell based material [9,44], this type of defect oriented on $\langle 11\xi \rangle$ with $\xi \approx 0.5$ is the most stable one according to DFT calculations [44]. It is well modelled by our new ML potential and found in our MD simulations with an orientation of $\langle 11\xi \rangle$ with $0.55 \leq \xi \leq 1$ due to the thermal motion. In Fig. 4(a), we show the structure of a dumbbell defect found in our MD simulations after collision cascade at 2 keV of PKA; where W atoms represented by black sphere correspond to the dumbbell atomic geometry and blue spheres are included to have a better visualization of this type of defect. The average distance between the W atoms 1 and 2 is 2.18 Å. This atomic arrangement is used to count the number of dumbbell defects found in the W sample at different velocity direction and PKA. Another common defect where four atoms share three lattice sites is called a Crowdion [45], which is stable at the $\langle 111 \rangle$ direction and found it in our MD simulations at this orientation. Fig. 4-(b) shows a snapshot of this particular material defect at the end of the MD simulation for a PKA of 1 keV. W atoms illustrated as golden spheres representing the geometry of a crowdion defect, while W atoms depicted as light-blue spheres are considered as atoms in their lattice position. The average inter-nuclear distance between the W atoms that define a crowdion is 2.3 Å. The geometries of the crystal defects are reported in the Supplementary material. In Fig 4 (c) and (d), we report the number of dumbbells and crowdions as a function of the PKA obtained by the ML potential [15]. A comparison to results given by MD simulations with the J-T-ZBL potential shows the absence of a crowdion defects formation at low impact energies, where the machine learned MD simulations predicts the formation of this type of defects in the whole PKA range. The higher the PKA value is, the bigger the number of crowdion defects is. Besides that, the J-T-ZBL potentials are able to model dumbbell defects, and its quantification agrees with the results obtained by using the ML potential. A second comparison to the results obtained by AT-EAM-FS for the identification of these types of defects is presented in the same figure. The same number of crowdion defects, in average, is formed after collision cascade by the ML potential and AT-EAM-FS potentials. Also, the formation of dumbbells defects is observed in the MD simulations by these two potentials, but the total number of defects is different.

4. Concluding remarks

In this paper, we performed molecular dynamics simulations to emulate neutron bombardment on Tungsten samples in an impact energy range of 0.5–10 keV, and a temperature of 300 K. For this, we use a new machine learning (ML) interatomic potential based on the Gaussian Approximation Potential framework. This new ML potential is accurately trained to the liquid phase, which is important to model the highly affected collision target; the short-range interatomic dynamics by including an accurate repulsive potential; and some samples to better model the re-crystallization of the molten region. The damage in the W material sample is analyzed by the classification and identification of point defects with our descriptor vector (DV) based method, which is based on the calculation of the rotation and translation invariant DV that describes the unique atomic neighborhood of each W atom in the material sample. Common point defects like self-interstitial-atoms and W atom next to a vacancy, and vacancy formation are quantified and classified as a function of the PKA energy. We found that the formation of W atoms as SIA and those in their distorted local environment follow a law of $18.49 E_{PKA}^{0.553}$ with E_{PKA} is the PKA energy. Point defects as crowdion shapes and W atoms next to a single vacancy are formed in the whole impact energy range. Our results have, in average, a good agreement with reported results by standard potentials. However, some energetically unstable point defects are corrected in the training data set for the ML potential to improve the accuracy of the MD simulations. Finally, these two methods are quite general and can be applied to develop efficient machine learning interatomic potentials for bcc metals and the damaged material samples are analyzed by the DV

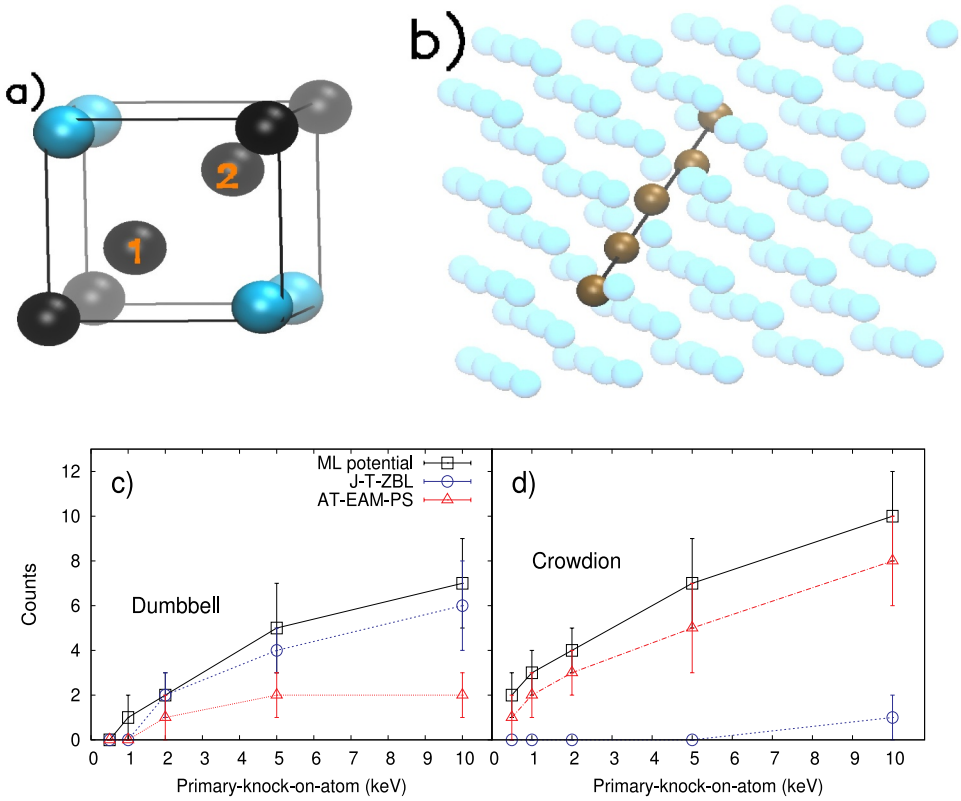

Fig. 4. (Color online) A dumbbell defect is shown in a) and a crowdion line defect is presented in b), identified at the final snapshot frame of the MD simulation with the ML potential at 2 and 1 keV of PKA, respectively. W atoms depicted as black (dumbbell) and golden (crowdion) spheres represent the atomic arrangement of the defects and atoms in a lattice position are illustrated as blue (light-blue) spheres. These defects are identified by the reference DV for an interstitial site, $\vec{v}_1(T=0)$ with a ⟨111⟩ orientation. The quantification of these defects is presented in c) and d). We compare results to those obtained by J-T-ZBL and AT-EAM-PS potentials. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

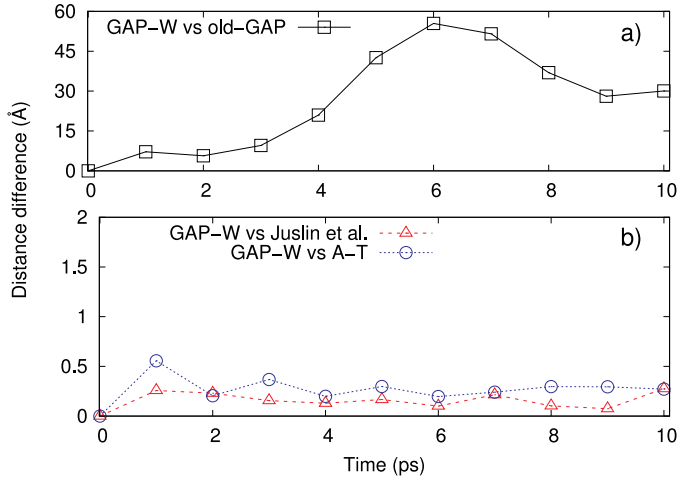

Fig. 5. Projectile trajectory comparison, as a function of the time, between the results obtained by new ML potential (GAP-W) and the one with the original GAP training data (old-GAP) in a), which shows the need of repulsion information in the training data set to model collision cascades. We also compare the GAP results to those by Juslin et al [23] potential and Ackland–Thetford (A–T) potentials [22] in b).

based method, which is a future work for our research group.

Declaration of Competing Interest

None.

Acknowledgments

F.J.D.G gratefully acknowledges funding from A. von Humboldt Foundation and C. F. von Siemens Foundation for research fellowship. Simulations were performed using the Linux cluster at the Max-Planck Institute for plasma physics. KN, FD and JB acknowledge that their part of this work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training programme 2014–2018 under grant agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

Appendix A. Test of original GAP potential

A machine learned interatomic potential for tungsten based on the Gaussian Approximation Potential formalism was developed by W. Szlachta et al. [12]. However, it lacks of information about the repulsive potential, so that the projectile is expected to travel freely when a primary knock-on-atom is assigned to it, in a MD simulation. The new ML potential [15] includes a realistic short-range repulsion to correctly simulate collision cascades. In order to test our new ML potential, we perform a MD simulation at 1 keV of PKA with a sample temperature of 300 K. The original GAP [12], Juslin et al. (J-T-ZBL) [23], Ackland–Thetford (AT-EAM-FS) [22], and our ML potentials are used to compare the projectile trajectory as a function of the simulation time.

In Fig. 5(a), we present the comparison between the projectile trajectory calculated by the original GAP and the new ML potentials. The distance difference between two projectile trajectories is calculated as $\sqrt{\sum_i (\xi_i(t) - \Xi_i(t))^2}$ where $\vec{\xi}(t)$ and $\vec{\Xi}(t)$ are the projectile trajectory obtained by different MD potentials, with i= x, y, z. We observe a remarkable difference, where in the original GAP the projectile travels freely in the material sample during the whole simulation. This result is caused by the original GAP potential not having a high-energy repulsive part. In Fig. 5(b), a similar comparison is done to the results obtained by using the J-T-ZBL and AT-EAM-FS potentials (which do have the high-energy repulsive part), as a function of the time. The distance difference is smaller than 1 Å for the complete MD simulation and the final position of the projectile is the same for the three cases. This result shows that the high energy collisional interactions are well treated by our new ML potential.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.nme.2019.100724 .

References

[1] S.J. Zinkle, N.M. Ghoniem, Prospects for accelerated development of high performance structural materials, Journal of Nuclear Materials 417 (1) (2011) 2–8. Proceedings of ICFRM-14

[2] K. Ehrlich, E. Bloom, T. Kondo, International strategy for fusion materials development, Journal of Nuclear Materials 283-287 (2000) 79–88, https://doi.org/10.1016/S0022-3115(00)00102-1. 9th Int. Conf. on Fusion Reactor Materials

[3] J. Marian, C.S. Becquart, C. Domain, S.L. Dudarev, M.R. Gilbert, R.J. Kurtz, D.R. Mason, K. Nordlund, A.E. Sand, L.L. Snead, T. Suzudo, B.D. Wirth, Recent advances in modeling and simulation of the exposure and response of tungsten to fusion energy conditions, Nucl. Fusion 57 (9) (2017) 92008.

[4] P. Lucasson, The production of Frenkel defects in metals, in: M.T. Robinson, F.N. Young Jr. (Eds.), Fundamental Aspects of Radiation Damage in Metals, ORNL, Springfield, 1975, pp. 42–65.

[5] K. Nordlund, J. Wallenius, L. Malerba, Molecular dynamics simulations of threshold energies in fe, Nucl. Instr. Methods Phys. Res. B 246 (2) (2005) 322–332.

[6] G. Federici, C. Skinner, J. Brooks, J. Coad, C. Grisolia, A. Haasz, A. Hassanein, V. Philipps, C. Pitcher, J. Roth, W. Wampler, D. Whyte, Plasma-material interactions in current tokamaks and their implications for next step fusion reactors, Nucl. Fusion 41 (12) (2001) 1967–2137.

[7] B.D. Wirth, X. Hu, A. Kohnert, D. Xu, Modeling defect cluster evolution in irradiated structural materials: focus on comparing to high-resolution experimental characterization studies, J. Mater. Res. 30 (9) (2015) 1440.

[8] A.E. Sand, S.L. Dudarev, K. Nordlund, High-energy collision cascades in tungsten: dislocation loops structure and clustering scaling laws, EPL (Europhys. Lett.) 103 (4) (2013) 46003.

[9] D. Nguyen-Manh, A.P. Horsfield, S.L. Dudarev, Self-interstitial atom defects in bcc transition metals: group-specific trends, Phys. Rev. B 73 (2006) 20101.

[10] S.L. Dudarev, P.-W. Ma, Elastic fields, dipole tensors, and interaction between self-interstitial atom defects in bcc transition metals, Phys. Rev. Mater. 2 (2018) 33602, https://doi.org/10.1103/PhysRevMaterials.2.033602.

[11] K. Nordlund, S.J. Zinkle, A.E. Sand, F. Granberg, R.S. Averback, R. Stoller, T. Suzudo, L. Malerba, F. Banhart, W.J. Weber, F. Willaime, S. Dudarev, D. Simeone, Primary radiation damage: a review of current understanding and models, J. Nucl. Mater. 512 (2018) 450–479.

[12] W.J. Szlachta, A.P. Bartók, G. Csányi, Accuracy and transferability of gaussian approximation potential models for tungsten, Phys. Rev. B 90 (2014) 104108.

[13] A.P. Bartók, M.C. Payne, R. Kondor, G. Csányi, Gaussian approximation potentials: the accuracy of quantum mechanics, without the electrons, Phys. Rev. Lett. 104 (2010) 136403.

[14] A.P. Bartók, R. Kondor, G. Csányi, On representing chemical environments, Phys. Rev. B 87 (2013) 184115.

[15] J. Byggmästar, A. Hamedani, K. Nordlund, F. Djurabekova, Machine-learning interatomic potential for radiation damage and defects in tungsten, Phys. Rev. B 100 (14) (2019) 144105, https://doi.org/10.1103/PhysRevB.100.144105.

[16] E. Wigner, F. Seitz, On the constitution of metallic sodium, Phys. Rev. 43 (1933) 804.

[17] Y.-N. Liu, T. Ahlgren, L. Bukonte, K. Nordlund, X. Shu, Y. Yu, X.-C. Li, G.-H. Lu, Mechanism of vacancy formation induced by hydrogen in tungsten, AIP Adv. 3 (12) (2013) 122111.

[18] A. Okabe, B. Boots, K. Sugihara, S.N. Chiu, Spatial Tessellations: Concepts and Applications of Voronoi Diagrams, John Wiley and Sons, Inc., New York, NY, 2000.

[19] A. Stukowski, Visualization and analysis of atomistic simulation data with OVITO–the open visualization tool, Modell. Simul. Mater. Sci.Eng. 18 (1) (2009) 15012.

[20] J. Fikar, R. Schublin, D.R. Mason, D. Nguyen-Manh, Nano-sized prismatic vacancy dislocation loops and vacancy clusters in tungsten, Nucl. Mater. Energy 16 (2018) 60–65.

[21] F. Domínguez-Gutiérrez, U. von Toussaint, On the detection and classification of material defects in crystalline solids after energetic particle impact simulations, J. Nucl. Mater. 528 (2019) 151833, https://doi.org/10.1016/j.jnucmat.2019.151833.

[22] G.J. Ackland, R. Thetford, An improved n-body semi-empirical model for body-centred cubic transition metals, Philos. Mag. A 56 (1) (1987) 15–30.

[23] N. Juslin, P. Erhart, P. Träskelin, J. Nord, K.O.E. Henriksson, K. Nordlund, E. Salonen, K. Albe, Analytical interatomic potential for modeling nonequilibrium processes in the w-c-h system, J. Appl. Phys. 98 (12) (2005) 123520.

[24] G. Bonny, D. Terentyev, A. Bakaev, P. Grigorev, D.V. Neck, Many-body central force potentials for tungsten, Modell. Simul. Mater. Sci.Eng. 22 (5) (2014) 53001.

[25] W. Setyawan, G. Nandipati, K.J. Roche, H.L. Heinisch, B.D. Wirth, R.J. Kurtz, Displacement cascades and defects annealing in tungsten, part i: defect database from molecular dynamics simulations, J. Nucl. Mater. 462 (2015) 329–337.

[26] F. Domínguez-Gutiérrez, P. Krstić, Sputtering of lithiated and oxidated carbon surfaces by low-energy deuterium irradiation, J. Nucl. Mater. 492 (2017) 56–61.

[27] G.M. Wright, M. Mayer, K. Ertl, G. de Saint-Aubin, J. Rapp, Hydrogenic retention in irradiated tungsten exposed to high-flux plasma, Nucl. Fusion 50 (2010) 075006.

[28] A. Herrmann, H. Greuner, N. Jaksic, M. Balder, A. Kallenbach, et al., Solid tungsten divertor-III for ASDEX Upgrade and contributions to ITER, Nucl. Fusion 55 (2015) 063015.

[29] S. Plimpton, Fast parallel algorithms for short-range molecular dynamics, J. Comput. Phys. 117 (1) (1995) 1–19.

[30] 2018, (http://libatoms.github.io/QUIP/).

[31] U. von Toussaint, S. Gori, A. Manhard, T. Höschen, C. Hschen, Molecular dynamics study of grain boundary diffusion of hydrogen in tungsten, Phys. Scr. 2011 (T145) (2011) 14036.

[32] B. Fu, M. Qiu, J. Cui, M. Li, Q. Hou, The trapping and dissociation process of hydrogen in tungsten vacancy: a molecular dynamics study, J. Nucl. Mater. 508 (2018) 278–285.

[33] B. Fu, M. Qiu, L. Zhai, A. Yang, Q. Hou, Molecular dynamics studies of low-energy atomic hydrogen cumulative bombardment on tungsten surface, Nucl. Instrum. Methods Phys. Res.Sect. B (2018).

[34] Y. Zhong, K. Nordlund, M. Ghaly, R.S. Averback, Defect production in tungsten: a comparison between field-ion microscopy and molecular-dynamics simulations, Phys. Rev. B 58 (5) (1998) 2361–2364, https://doi.org/10.1103/PhysRevB.58.2361.

[35] A. Sand, J. Dequeker, C. Becquart, C. Domain, K. Nordlund, Non-equilibrium properties of interatomic potentials in cascade simulations in tungsten, J. Nucl. Mater. 470 (2016) 119–127.

[36] E. Salonen, K. Nordlund, J. Keinonen, C.H. Wu, Enhanced erosion of tungsten by atom clusters, J. Nucl. Mater. 305 (1) (2002) 60–65.

[37] P. Mahalanobis, On tests and measures of group divergence i. theoretical formulae, J. Proc. Asiat. Soc. Bengal 26 (1930) 541.

[38] K. Heinola, F. Djurabekova, T. Ahlgren, On the stability and mobility of di-vacancies in tungsten, Nucl. Fusion 58 (2) (2017) 26004.

[39] J.L. Bentley, Multidimensional binary search trees used for associative searching,

Commun. ACM 18 (9) (1975) 509–517.
[40] M.B. Kennel, Kdtree 2: Fortran 95 and c++ software to efficiently search for near neighbors in a multi-dimensional euclidean space, 2004, arXiv:physics/0408067.
[41] D. Bacon, A. Calder, F. Gao, V. Kapinos, S. Wooding, Computer simulation of defect production by displacement cascades in metals, Nucl. Instrum. Methods Phys. Res.Sect. B 102 (1) (1995) 37–46.
[42] R.E. Stoller, The role of cascade energy and temperature in primary defect formation in iron, J. Nucl. Mater. 276 (1) (2000) 22–32.
[43] K. Nordlund, S.J. Zinkle, A.E. Sand, F. Granberg, R.S. Averback, R. Stoller, T. Suzudo, L. Malerba, F. Banhart, W.J. Weber, F. Willaime, S. Dudarev, D. Simeone, Improving atomic displacement and replacement calculations with physically realistic damage models, Nat. Commun. 9 (2018) 1084.
[44] P.-W. Ma, S.L. Dudarev, Symmetry-broken self-interstitial defects in chromium, molybdenum, and tungsten, Phys. Rev. Mater. 3 (2019) 43606.
[45] P.M. Derlet, D. Nguyen-Manh, S.L. Dudarev, Multiscale modeling of crowdion and vacancy defects in body-centered-cubic transition metals, Phys. Rev. B 76 (2007) 054107.