Machine Learning Based On-The-Fly Kinetic Monte Carlo Simulations of Sluggish Diffusion in Ni-Fe Concentrated Alloys

Wenjiang Huang and Xian-Ming Bai*

Department of Materials Science and Engineering

Virginia Polytechnic Institute and State University

Blacksburg, Virginia, USA

Abstract

Defect diffusion in concentrated alloys plays a key role on governing their unique mechanical and physical properties. In such alloys, defect diffusion depends on its complex local atomic environment and varies from site to site due to the chemical disorder. On-the-fly determination of the defect migration barrier at every site using the standard nudged elastic band (NEB) method is computationally expensive and often impractical. In this work, we couple machine learning and kinetic Monte Carlo (KMC) to study vacancy-mediated sluggish diffusion in concentrated Ni-Fe model alloys. Based on about 32,000 pre-calculated NEB barriers, an artificial neural network (ANN) based machine learning model is developed to accurately predict the vacancy migration barriers for arbitrary local atomic environments, including both random solution configurations and alloys with short-range orders. The ANN model is then coupled with KMC (ANN-KMC) to determine the vacancy migration barriers on-the-fly, enabling an efficient study of the vacancy diffusion in the full composition range at a wide range of temperatures. In addition, a composition and temperature dependent jump attempt frequency model is developed. Upon calibration, the ANN-KMC modeling can predict nearly identical vacancy diffusivities as those obtained from independent molecular dynamics (MD) and temperature accelerated dynamics (TAD) simulations at their accessible temperatures. The sluggish diffusion mechanisms in this specific alloy system at both high and low temperatures are discussed based on the ANN-KMC results.

*Corresponding author: xmbai@vt.edu

1. Introduction

Compositionally disordered alloys including high-entropy alloys, also known as multi-component solid solution concentrated alloys (CSAs), have been attracting significant interest in recent years because of their unique properties, such as excellent mechanical properties at both low and high temperatures [1, 2], good corrosion resistance [3], and enhanced radiation tolerance [4, 5]. Of particular interest is the concept of "sluggish diffusion" in CSAs because defect or solute diffusion kinetics is directly related to these unique material properties. However, the existence or the extent of sluggish diffusion in concentrated alloys is still under debate and contradictory results have been reported in literature. For instance, some experimental studies showed evidence of "sluggish diffusion" [6-8], whereas some others raised questions about this phenomenon [9-11]. On the modeling side, many atomistic simulations [12-15] have been conducted with the aim of understanding the mechanisms of atomic diffusion in such chemically disordered structures. Similar as in experimental studies, both existence and non-existence of sluggish diffusion have been reported, depending on the chosen simulation systems. In many studies, it is assumed that the alloying elements are randomly distributed within CSAs and the alloys form a single-phase (such as face-centered-cubic (fcc) or body-centered-cubic (bcc) structures) solid solution. However, numerous experimental observations have demonstrated that CSAs may contain some extent of short-range orders (SROs) or other ordered phases, and these ordered structures cannot be overlooked [16-18].

Computer modeling is a powerful tool for understanding and predicting the defect/solute diffusion in materials. However, there are some challenges for modeling CSAs. For example, some theoretical models such as multi-frequency models [19] can be used for predicting the atomic diffusion. However, these analytical models are typically applicable to dilute alloys [20, 21]. Given the uncountable number of nonequivalent atomic configurations in CSAs, these models may not be easily generalized to concentrated alloys. An alternative tool for studying atomic-level defect diffusion is the molecular dynamics (MD) method [13, 14, 22], in which optimal diffusion paths are automatically chosen based on the local potential energy landscape, provided suitable and accurate interatomic potentials are available. If a simulated trajectory is sufficiently long to ensure the statistics of defect or solute jumps, MD can accurately simulate their diffusion events. However, MD is limited by its inherently short timescale, typically up to a few tens of nanoseconds. As a result, many slow but important diffusion processes (e.g., rare events) that happen at low temperatures of interest may not be captured by MD. The kinetic Monte-Carlo (KMC) method, which implicitly treats atomic vibrations and only considers the kinetics of defects or solutes, can extend the atomic-level simulation to the experimentally accessible timescale [23, 24]. In the conventional rigid-lattice KMC, atoms are restricted to the predefined positions such as fcc lattice sites. The reliability of KMC modeling depends heavily on the accurate description of migration barriers of defect jumps, E_m, which need to be provided to the model as *a priori* knowledge. However, in CSAs the defect migration barrier changes from site to site due to the varying local atomic environment. Therefore, it is nearly impossible to determine all the migration barriers beforehand and provide the complete event table to KMC. On the other hand, it is possible to use the off-lattice KMC [25] such as the kinetic activation-

relaxation technique (k-ART) [26, 27] to determine the migration barriers on-the-fly in CSAs, as demonstrated by Osetsky et al. [28]. However, the on-the-fly determination of migration barriers using the nudged elastic band (NEB) method [29-31] or the activation relaxation technique nouveau (ARTn) [32] entails a high computational cost.

Recently, machine learning techniques such as the artificial neural networks (ANNs) have attracted researchers' significant interests for studying various materials science problems [33-37]. Machine learning is a powerful tool for analyzing high-dimensional and complex problems to identify the underlying correlations with some material features or descriptors, if sufficient and accurate training data are available. For the defect or solute diffusion in CSAs that is the focus of this work, it is envisioned that one may develop an advanced regression algorithm or machine learning model for predicting the migration barriers in CSAs based on the local atomic configurations (LACs) around a defect or solute, because its migration barrier strongly depends on its neighboring chemical species as well as its jumping pathway. After such a model is established, one may replace the computationally expensive on-the-fly calculation of migration barriers in KMC with such a more computationally efficient model.

In this work, we aim to develop an ANN model to predict vacancy migration barriers for arbitrary atomic configurations in CSAs and couple it with the conventional KMC to study the vacancy diffusion at a wide range of alloy compositions and temperatures. The concentrated fcc $Ni_{1-x}Fe_x$ ($x = 0 - 1$) alloy system is used as a model system because it has been shown to have strongest sluggish diffusion at the percolation threshold of $x_{Fe} \approx 0.2$ [28]. Two types of atomic configurations are considered: fully random distribution of alloying elements (a chemically disordered structure) as in previous MD studies [13, 28], and the Metropolis Monte-Carlo (MMC) optimized structures that encompass local SROs as characterized by the Warren-Cowley SRO parameter [38]. A finite number of NEB calculations are conducted to obtain necessary training data for the ANN-based machine learning model. The trained ANN model is then coupled with KMC to predict the vacancy migration barrier on-the-fly and calculate the effective vacancy diffusivity in this alloy system. A composition- and temperature-dependent jump attempt frequency model, rather than a fixed frequency that is commonly used in many KMC studies [28, 39], is also developed. The calibrated ANN-KMC model can predict similar vacancy diffusivities as independent results by MD at high temperatures and by temperature accelerated dynamics (TAD) [40] at moderate temperatures, demonstrating that the approach can achieve very high computational efficiency without compromising its accuracy. The approach thus enables us to study defect diffusion in CSAs at low temperatures that may not be accessible by other atomistic simulation methods.

2. Methods

A number of complementary simulation methods are used in this work: NEB, ANN based machine learning, KMC, MMC, MD, and TAD. For the NEB, MMC, MD, and TAD, the Ni-Fe interatomic potential developed by Bonny et al. [41] is used. These methods are briefly described as follows.

2.1. Alloy structures and NEB database of vacancy migration barriers

As mentioned earlier, two groups of atomic configurations for the $Ni_{1-x}Fe_x$ (x = 0 – 1) fcc alloys are considered: random distribution of alloying elements and MMC-optimized structures that contain SROs. In each group, eight compositions are studied: x_{Fe} = 10%, 20%, 25%, 35%, 50%, 65%, 80%, 90%. To create the first group (denoted as the "Random" group), Ni and Fe atoms are randomly mixed in a $10a_0 \times 10a_0 \times 10a_0$ fcc lattice containing 4,000 atoms at each composition. To create the second group (denoted as "MMC" group), the hybrid MC+MD method implemented in the LAMMPS package [42] is applied to optimize each generated random structure. During the optimization, a randomly selected Ni is swapped with a randomly selected Fe atom, and the acceptance probability of this swap depends on the change of the system potential energy. Following every swap, a short (0.01 ps or 5 timesteps) MD simulation is performed in an NVT (constant number of atoms, volume, and temperature) ensemble at 300 K to relax the system. This procedure is repeated for a total of 400,000 MMC steps. The resulting alloy structures in the MMC group have much lower potential energies than the counterparts in the Random group and an example is shown in Fig. 1(a). The MMC-optimized structures exhibit some extent of short-range orders. To quantify the degree of ordering, the Warren-Cowley SRO parameter [38] is calculated up to four nearest neighboring (nn) shells for each composition, as shown in Fig. 1(b). Clearly, the Random structures have no discernible ordering because the SRO values of all shells are nearly zero, while the MMC structures have distinctive ordering, as indicated by the large positive (clustering of Ni-Ni or Fe-Fe pairs) or negative (clustering of Ni-Fe pairs) SRO values in the second and third shells.

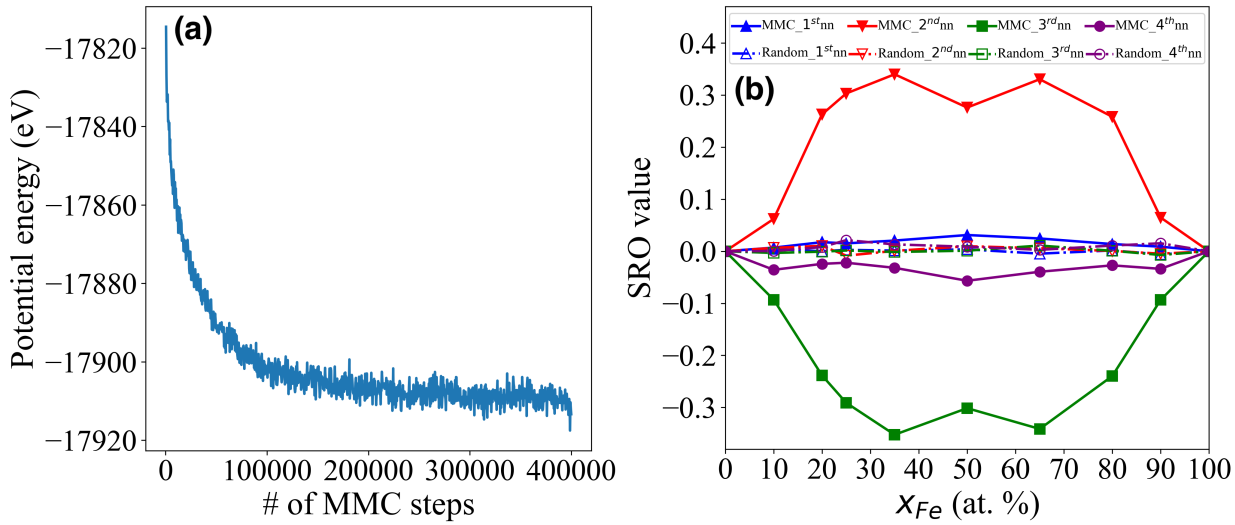

Fig. 1 (a) Potential energy change during a hybrid MC + MD simulation (denoted as MMC optimization) in an $Ni_{80}Fe_{20}$ alloy that has an initially random atomic configuration. (b) Warren-Cowley SRO values up to the 4th nn for both Random and MMC structures at different compositions.

After both Random and MMC structures are created, the NEB method [30, 31] implemented in LAMMPS is used to calculate the vacancy migration barriers (E_m) for different atomic environments. At each Ni-Fe composition, a vacancy is created at every lattice site, one at a time. At each vacancy site, the vacancy migration path is randomly chosen between the vacancy and one of its twelve 1st nn atoms and the corresponding migration barrier (E_m) is calculated. This results in 4,000 barriers at each composition and 32,000 barriers in each structural group (because there are eight compositions in each group). In all NEB calculations, the number of NEB images (including the initial and final states) is 13 and the spring constant is set to 1.0 eV/Å^2.

2.2. LAC definition and ANN model

In an $Ni_{1-x}Fe_x$ fcc alloy, a vacancy can diffuse to one of its 12 nearest neighbor sites, as schematically illustrated in Fig. 2(a). The vacancy migration barrier of each pathway depends on not only the atom type at the neighboring site along that pathway, but also the local atomic configuration (LAC) around that neighboring site. Therefore, a well-defined LAC is a premise for an ANN model that can accurately predict the vacancy migration barrier (E_m). Figure 2(b) schematically shows two possible vacancy migration pathways between the vacancy "V" and atom "$A_{i=1,2}$". The LAC of an atom "A_i" includes all atoms as well as the vacancy site within a cutoff radius (large dash circle) from A_i. In the relaxed $Ni_{1-x}Fe_x$ alloys, atoms can have small displacements from the perfect fcc lattice sites due to the mixing effect. Although the NEB calculations include all local chemistry and lattice distortion, the migration events in the conventional KMC are still described in a rigid lattice framework. To keep this advantageous simplicity of the KMC model, the small off-site displacement of an atom in the equilibrium state is neglected in our ANN model if it stays within the original Wigner-Seitz cell [43] (see Fig. 2(b)). Therefore, all sites in the same nearest neighboring shell have the same distance from the central atom/site. The description of a LAC is represented by a string of integers for each neighboring shell in consideration, e.g., 1 for Ni, 2 for Fe and 0 for vacancy. Four neighboring shells are taken into consideration for constructing LAC vectors. A closer shell is endowed with a higher weight. For example, the first two neighboring shells around a moving atom "A_i" can be defined as,

$$1nn\ LAC = W_{1nn}[0, 1, 2, 1, 2, 1, 1, 1, 2, 1, 1, 1],$$

$$2nn\ LAC = W_{2nn}[1, 2, 2, 2, 1, 1],$$

where W_{1nn}, W_{2nn} are the weights for 1st and 2nd neighboring shells, respectively, and $W_{1nn} > W_{2nn}$.

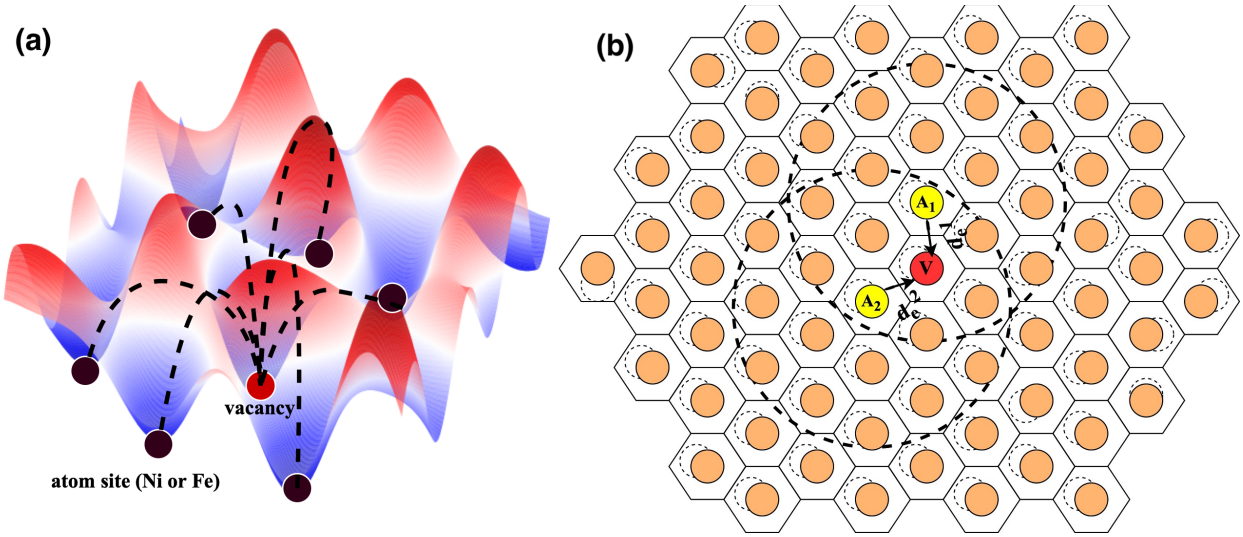

Fig. 2 (a) Schematic illustration of vacancy migration pathways in the potential energy landscape. (b) 2D schematic illustration of two possible vacancy migration pathways (from "V" to "$A_{i=1,2}$") in a rigid lattice approximation of a distorted crystal structure. Hexagons represent Wigner-Seitz cells, within which small dash circles denote real equilibrium positions of atoms and filled circles are rigid lattice sites. Here "A_i" represents the moving atom (Ni or Fe), "V" is the vacancy site and "d_e^i" represents the atom movement direction of the pathway. The large dash circle around each A_i shows the cutoff radius within which all neighboring sites are included in constructing its LAC.

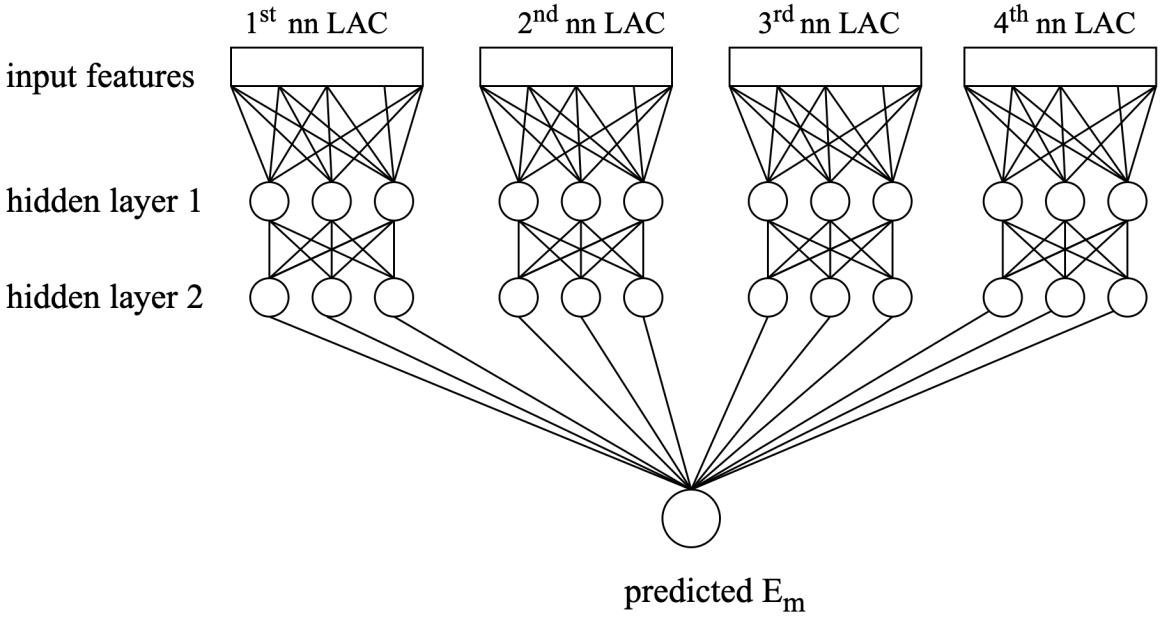

Fig. 3 Schematic of the ANN architecture. There are three neurons (solid circles) in the hidden layers 1 and 2 for each input shell vector to integrate all information to predict E_m values.

The ANN architecture is shown in Fig. 3. The training logic is: (i) The LAC input features from the 1st nn to 4th nn shells are separately connected to the first hidden layer; (ii) The first hidden layer is implemented to integrate the information for each shell; (iii) The second hidden layer desegregate the information passed by the first hidden layer to give the predicted E_m. The number of neurons and layers are determined empirically by trial-and-error. The training algorithm uses the Levenberg-Marquardt method [44], a hybrid technique that uses both Gauss-Newton update and gradient descent to converge to an optimal solution, and its superior performance for similar applications has been demonstrated [45, 46].

2.3. ANN-KMC modeling

The trained ANN model is coupled with the rigid-lattice KMC using the KMCLib package [47] to simulate the vacancy diffusion in $Ni_{1-x}Fe_x$ alloys for both MMC and Random groups. In the lattice KMC model, atom diffusion is solely driven by the thermally activated migration jumps that a single vacancy takes. The rate of a vacancy jump event from an initial position *i* to one of its neighboring sites *j* is expressed by the standard transition state theory,

$$\Gamma_{ij} = \nu_o \exp\left(-\frac{E_{m,ij}}{k_B T}\right), \quad (1)$$

where ν_o is the jump attempt frequency, k_B is the Boltzmann's constant, T is the absolute temperature, and $E_{m,ij}$ is the vacancy migration barrier from site *i* to *j*. At each KMC step, the ANN model predicts $E_{m,ij}$ for all 12 possible jumping pathways on-the-fly with a negligible computational cost. A random number r_1 in the range of (0, 1] will pick a path *s* among 12 possible pathways following,

$$\sum_{j=1}^{s} \Gamma_{ij} > r_1 \cdot \sum_{j=1}^{12} \Gamma_{ij}, \quad (2)$$

and the system evolves into a new state by the path *s*. The time elapsed is calculated by another random r_2 in the range of (0, 1] according to the residence time algorithm [24],

$$\Delta t = -\frac{\ln(r_2)}{\sum_{j=1}^{12} \Gamma_{ij}}. \quad (3)$$

In this work, 10,000 – 30,000 KMC steps are conducted for each simulation. The diffusion coefficient is calculated using the Einstein relation [48],

$$D = \frac{\Delta r_s^2(t)}{6t} \quad (4)$$

where $\Delta r_s^2(t)$ is the sum of the atomic square displacements (ASDs) of all atoms (for total diffusivity) or individual type of atoms (for partial diffusivity) in 3D, and *t* is the simulation time.

To obtain reliable statistics, six independent simulations are conducted at each condition and the average diffusivities are reported in this work.

Regarding the jump attempt frequency (v_o), it was typically set to a fixed value close to the phonon frequency ($10^{12} \sim 10^{13}$ s^{-1}) in many previous KMC simulations [49, 50]. However, it could be affected by both temperature and composition. For example, in some studies it was found to be two orders of magnitude higher or lower than the phonon frequency [51, 52]. In this work, a temperature- and composition-dependent v_o is developed because a constant v_o leads to unsatisfactory results. The details are discussed in the Results section.

2.4. MD simulations

To compare with ANN-KMC results at high temperatures, MD simulations of vacancy diffusion in the $Ni_{1-x}Fe_x$ alloys are conducted at 900 – 1300 K using LAMMPS [42]. The structures in both Random and MMC groups are studied. The effects of atomic configurations and other material properties on the vacancy diffusivities will be presented in detail elsewhere. Here only the calculated diffusivities are reported, which are used to validate ANN-KMC results. The MD simulations are conducted in an NPT ensemble (constant number of atoms, pressure, and temperature). The Nose-Hoover thermostat and barostat [53] are used to control the system temperature and pressure (at zero bar), respectively. The time step is set to 2 fs. The simulations at lower temperatures, such as 900 K and 1000 K, are conducted up to 100 ns each for ten independent simulations to improve the statistics at each condition; Whereas the simulations at higher temperatures (> 1000 K), which have faster diffusivities, are completed over 40 ns each for six independent simulations. Same as in KMC, the total and partial diffusion coefficients are calculated from the atomic square displacements (Eq. 4).

2.5. TAD simulations

The vacancy diffusivities calculated by MD are limited to high temperatures. In order to obtain diffusivities at a moderate temperature (800 K) for further validating the KMC results, the TAD method [40] implemented in LAMMPS is used. TAD is one type of accelerated molecular dynamics (AMD) methods [54, 55], and has the ability to reach much longer timescale than MD but still maintains the full atomic fidelity as MD [56, 57]. TAD conducts basin-constrained MD at a high temperature to accelerate the search of many candidate transition events from the current state (basin). Unlike in the rigid-lattice KMC, TAD does not require *a priori* knowledge of these candidate transition paths or mechanisms. Once a candidate transition event is detected, the NEB method is used to calculate the transition barrier on-the-fly. The system is then brought back to the current basin and new searches continue, until a stopping criterion is met. For a candidate transition event j, the time at the low temperature of interest ($t_{Low,j}$) is extrapolated from its high-temperature MD time ($t_{High,j}$) based on the transition state theory,

$$t_{Low,j} = t_{High,j} exp\left[E_{m,j}\left(\frac{1}{k_B T_{Low}} - \frac{1}{k_B T_{High}}\right)\right], \quad (5)$$

where $E_{m,j}$ is the transition barrier obtained by NEB (e.g., the vacancy migration barrier in this work) along the path j, T_{Low} is the low temperature of interest, and T_{High} is the high temperature for accelerating the transition. Among the detected candidate events after the stopping criterion is met, the event with the shortest time at T_{Low} is accepted. Therefore, the candidate events that should not happen at T_{Low} are filtered out. Then the system advances to a new basin and the process repeats. In this work, T_{High} and T_{Low} are set to 1600 K and 800 K, respectively. The simulation system contains 864 atoms of different compositions in a $6a_0 \times 6a_0 \times 6a_0$ fcc lattice. The time step for the high-temperature MD is 2 fs. A transition is defined when the displacement of any atom is greater than 0.6 Å and the transition is checked every 50 steps. The desired confidence level for stopping criterion is set to δ=0.05. The minimum pre-exponential factor is set to $\nu_{min} = 2\times10^{12}$ s^{-1} (or $t_{max} = 0.5$ ps). The simulation time is over 60 ns at T_{High}, which results in μs-scale time at T_{Low}. Same as in KMC and MD, the atomic square displacements (Eq. 4) are used to calculate the diffusion coefficients. In addition, the transition barriers of all accepted events during the vacancy diffusion are recorded by TAD, which will be presented in the Results section.

3. Results

3.1 ANN model for vacancy migration barriers

To train the ANN model of vacancy migration barriers (E_m), the two groups (Random and MMC) of databases mentioned above are not identically utilized. In this work, the barriers in the MMC group, which contain some SROs, are used to train the ANN model. Specifically, the data in the MMC group are split to serve the functions of training, validation, and testing with a ratio of 70%:15%:15%. Once an ANN model has been adequately trained, e.g., obtained a satisfactory performance on the testing set, the data in the Random group, equivalent to unseen data to the trained ANN model, are tested to verify the transferability of the ANN model. The trained ANN model is accepted when it also reaches satisfactory performances on the data in the Random group. Figure 4(a) shows a good correlation between ANN predicted and NEB calculated E_m on the MMC group dataset. It is adequately accurate with the average prediction error within 0.040 eV, and R^2 of ~95%. Figure 4(b) presents an average error of 0.043 eV and R^2 of 93% on the Random group dataset. Some outliers do exist, possibly related to the rigid lattice approximation. However, the accuracy of the KMC model is not affected by these extreme and rare cases because they make little difference to the statistics.

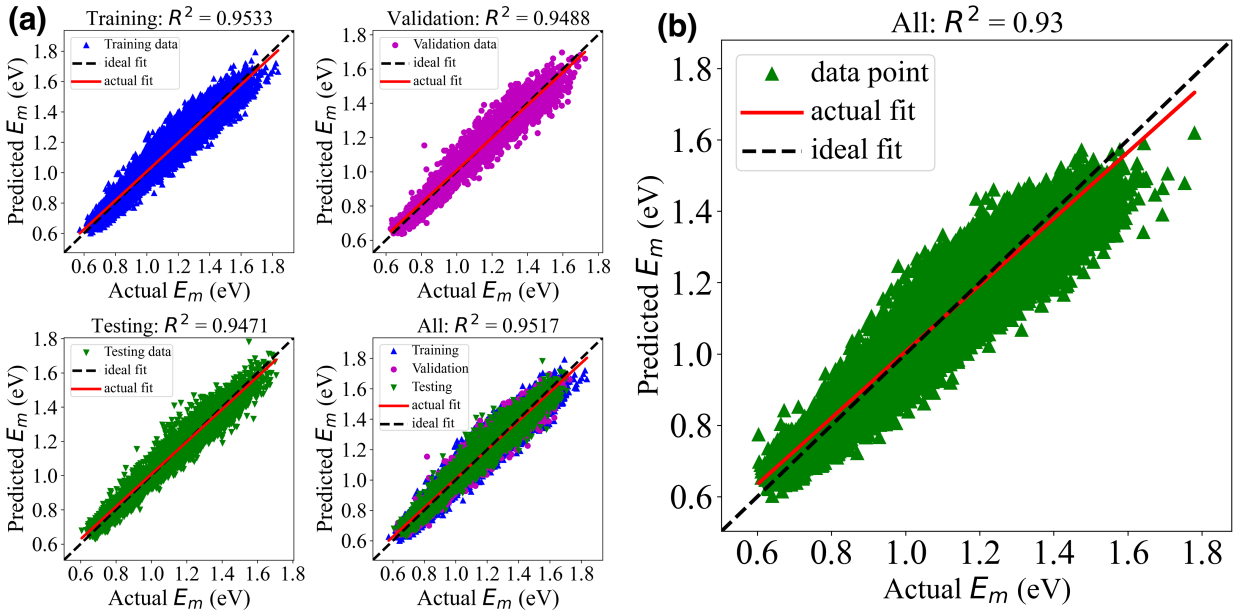

Fig. 4 (a) The trained ANN model of vacancy migration barriers that only uses the MMC group's data. Four plots indicate regression accuracies on training, validation, testing, and overall dataset. (b) The performance of the trained ANN model in (a) on the unseen data from the Random group.

3.2 ANN-KMC results of the MMC group

The ANN model developed in the previous section is coupled with KMC to calculate the vacancy diffusivities in the MMC-group structures that contain SROs. The first test is using a fixed jump attempt frequency, v_o = 10^{13} s^{-1}, as in many previous studies [28, 39]. The calculated diffusion coefficients are compared with the MD results as a function of alloy composition at 1300 K, as shown in Fig. 5(a). Clearly, there is a large discrepancy between ANN-KMC and MD results when the Fe concentration $x_{Fe} < 80\%$. The most distinctive difference is the pure Ni, which is independent of the ANN model because it only has a single vacancy migration barrier. Therefore, the fixed v_o value is likely the reason that induces such discrepancies. To solve this problem, in this work v_o is treated as both composition- and temperature-dependent. According to Kong and Lewis [58], the temperature-dependent attempt frequency has the form,

$$v_o = \frac{k_B T}{h} \exp\left(\frac{-\Delta F_{vib}}{k_B T}\right), \tag{6}$$

where h is the Planck's constant, ΔF_{vib} is the vibrational free energy difference between the transition state and the initial state. For simplicity, the exponential term in Eq. (6) is replaced by a polynomial function of composition,

$$v_o = \frac{k_B T}{h} f(x_{Fe}), \tag{7}$$

where $f(x_{Fe})$ will be calibrated with the MD diffusivity data, as discussed below.

To determine $f(x_{Fe})$, the v_o values are obtained by fitting to the MD data at 1300 K for all compositions, as shown in Fig. 5(b). At this temperature, the coefficient $\frac{k_B T}{h} = 2.71\times10^{13}$ s^{-1}, and a 4th order polynomial function is used to fit $f(x_{Fe})$,

$$f(x_{Fe} = x) = 1.764 - 8.92 \times 10^{-3}x + 3.96 \times 10^{-5}x^2 - 4.16 \times 10^{-6}x^3 + 3.2 \times 10^{-8}x^4. \quad (8)$$

To determine v_o for other temperatures T, Eqs. (7-8) are directly used and no other fitting is needed. Essentially this scales the v_o values at 1300 K with a factor of T/1300, i.e., $v_{o,T} = \frac{T}{1300} v_{o,1300K}$ for the corresponding compositions. It should be noted that the MD data at x_{Fe} = 25% (i.e., $Ni_{75}Fe_{25}$) is not included in the fitting but $f(x_{Fe})$ still works well for this composition. Also noted that although $f(x_{Fe})$ is only fitted to the MD data of the MMC group at 1300 K, it works well at other temperatures for both MMC and Random groups, as shown below.

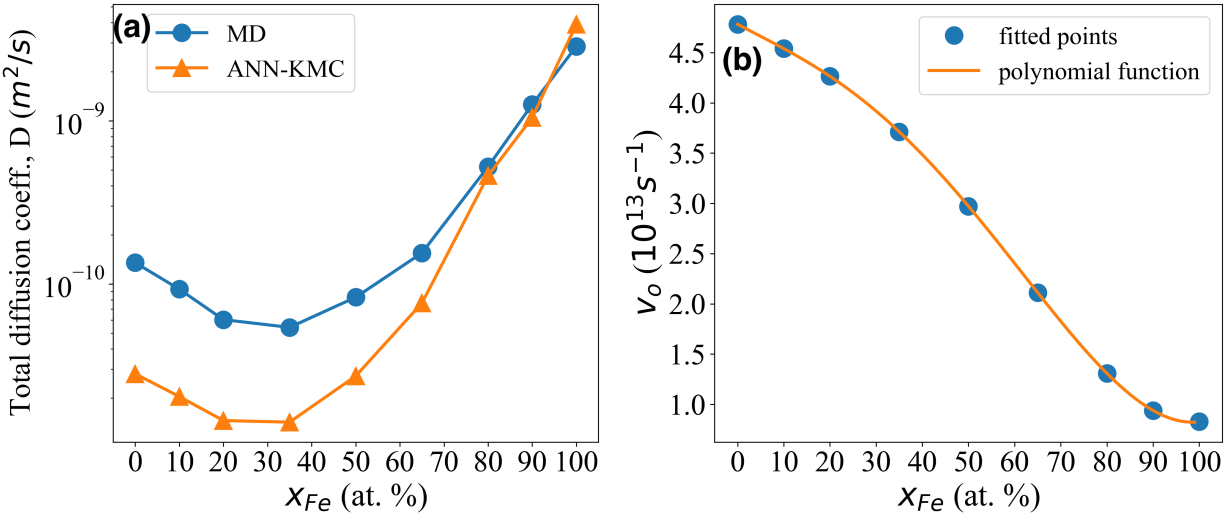

Fig. 5 (a) Total diffusion coefficients obtained from ANN-KMC at 1300 K using a constant jump attempt frequency, v_o = 10^{13} s^{-1}, which show large discrepancies with the MD results at x_{Fe} < 80%. (b) The fitted v_o values (filled circles) against the MD data at 1300 K. The solid line shows the fitted 4th order polynomial function (Eqs. (7 – 8)).

Figure 5(b) shows that v_o decreases monotonically with increasing Fe concentration. It also shows that v_o in pure Ni is about 5 times higher than in pure Fe, even though they both have an fcc structure in this work, consistent with many previous studies that it is material dependent [58]. After the composition-dependent v_o is calibrated at 1300 K, it is used in ANN-KMC to predict the vacancy diffusivities at lower temperatures. Figure 6(a) shows a complete comparison of vacancy diffusivities obtained via ANN-KMC and MD simulations for temperatures from 900 K to 1300 K. The error bar of each MD data point represents the standard deviation from 6 – 10 independent simulations. Each diffusivity value from ANN-KMC is averaged from 6 independent simulations, and its error bar is ignored due to very small variance. Clearly, the

ANN-KMC predicts nearly identical vacancy diffusivities as the MD results in a wide range of compositions and temperatures, even though v_o is merely fitted to the MD data at 1300 K. There are some discrepancies under some conditions, primarily at low temperatures (e.g., $Ni_{65}Fe_{35}$ at 900 K). This is likely due to the insufficient statistics in MD simulations because the short timescale in MD poses an inherent challenge for studying the slow diffusion kinetics at low temperatures. At all temperatures, the vacancy diffusivity first decreases then increases with the increasing Fe concentration, consistent with Osetsky et al.'s work [28]. The two end points (i.e., pure Ni and pure Fe) at each temperature indicate that Fe is a faster diffuser than Ni. The minimum diffusivity occurs around $x_{Fe} = 25{\sim}35\%$ for the structures of the MMC group in this work, slightly different from ~20% (which is the Fe percolation threshold) in Osetsky et al.'s work [28].

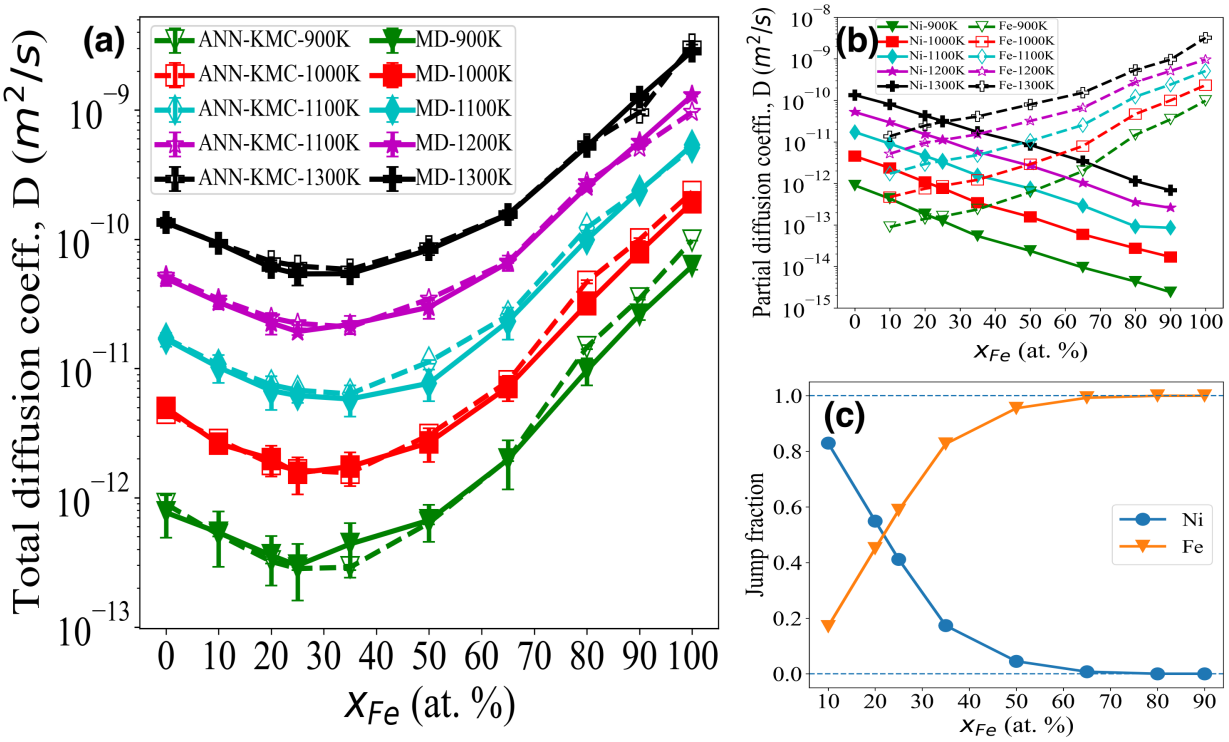

Fig. 6 (a) Total vacancy diffusivities as a function of composition in the MMC-optimized $Ni_{1-x}Fe_x$ model alloys calculated from ANN-KMC and MD simulations at 900 – 1300 K. Note only the MD data at 1300 K are used to calibrate the v_o in ANN-KMC. Solid lines with filled symbols are MD data, and dashed lines with unfilled symbols are ANN-KMC data. Each error bar represents the standard deviation of 6 – 10 MD results. The error bars in ANN-KMC results are ignored due to the negligible variances. (b) Partial diffusivities as a function of alloy composition from ANN-KMC results at different temperatures. (c) Jump fractions of Ni and Fe as a function of composition at 1200 K.

Figure 6(b) shows the partial diffusion coefficients of Ni and Fe from the ANN-KMC modeling. At all temperatures, Ni diffusivity (D_{Ni}^{MMC}) decreases while Fe diffusivity (D_{Fe}^{MMC})

increases with increasing Fe concentration, because Fe is a faster diffuser. The intersection between D_{Ni}^{MMC} and D_{Fe}^{MMC} is located at x_{Fe} = 20 ~ 25% and slightly shifts rightward (to higher x_{Fe}) with the increasing temperature. This crossover Fe concentration is slightly lower than that for the minimal total diffusivity in Fig. 6(a), which is in the range of $x_{Fe} = 25{\sim}35\%$. At x_{Fe} = 20% (in the $Ni_{80}Fe_{20}$ alloy), the ratio of D_{Fe}^{MMC} / D_{Ni}^{MMC} is 0.79, 0.72, 0.65, 0.58, 0.55 for 900 ~ 1300 K, respectively. This indicates that the contribution from the faster diffuser Fe increases as the temperature decreases. Figure 6(c) shows the fractions of vacancy jump types (i.e., a vacancy jump via Ni or Fe atom) as a function of Fe concentration at 1200 K. It can be seen that the dominant vacancy jump type changes from Ni to Fe in the range of x_{Fe} = 20 ~ 25% (about x_{Fe}= 22%), similar to the crossover concentration for the partial diffusivities in Fig. 6(b). After the crossover, the jump via Fe becomes more and more dominant. When x_{Fe} > 65%, the Fe jumps dominate completely and there are almost no Ni jumps.

The excellent agreement between ANN-KMC and MD results at 900 ~ 1300 K suggests that ANN-KMC may be used to predict the diffusivities at low temperatures. However, as mentioned earlier the events that happen at high temperatures in MD may not happen at the lower temperatures of interest. Therefore, to further validate the ANN-KMC model, the TAD method is used to calculate the total vacancy diffusivities in the $Ni_{1-x}Fe_x$ alloy at a moderate temperature, T = 800 K. The comparison between TAD and ANN-KMC results at different compositions at 800 K is shown in Fig. 7(a). Note that the TAD results are completely independent (blind) to the ANN-KMC model because the system size in TAD is different from that in ANN-KMC and the simulated structures are also different. Nevertheless, the agreement between ANN-KMC and TAD results is reasonably good. Even though there are slight discrepancies, the error bars (standard deviations) in TAD results overlap with ANN-KMC results at many compositions. Both TAD and ANN-KMC give similar minimal diffusivities at $x_{Fe} = 20 \sim 25\%$, which is slightly different from the minimum diffusivities occurring at $x_{Fe} = 25 \sim 35\%$ at higher temperatures (900 – 1300 K) shown in Fig. 6(a). Since ANN-KMC results have good agreement with MD and TAD results at both high and moderate temperatures, the ANN-KMC is used to predict the total vacancy diffusivities at a low temperature (T = 500 K) that is inaccessible to MD, as shown in Fig. 7(a). Again, the minimal diffusivity happens at $x_{Fe} = 20 \sim 25\%$, similar as at T = 800 K. It should be noted that the sluggish diffusion effect at this composition is more pronounced than at high temperatures (Fig. 6(a)) because the minimum vacancy diffusivity at 500 K is about one and five orders of magnitude than the diffusivities in pure Ni and Fe, respectively.

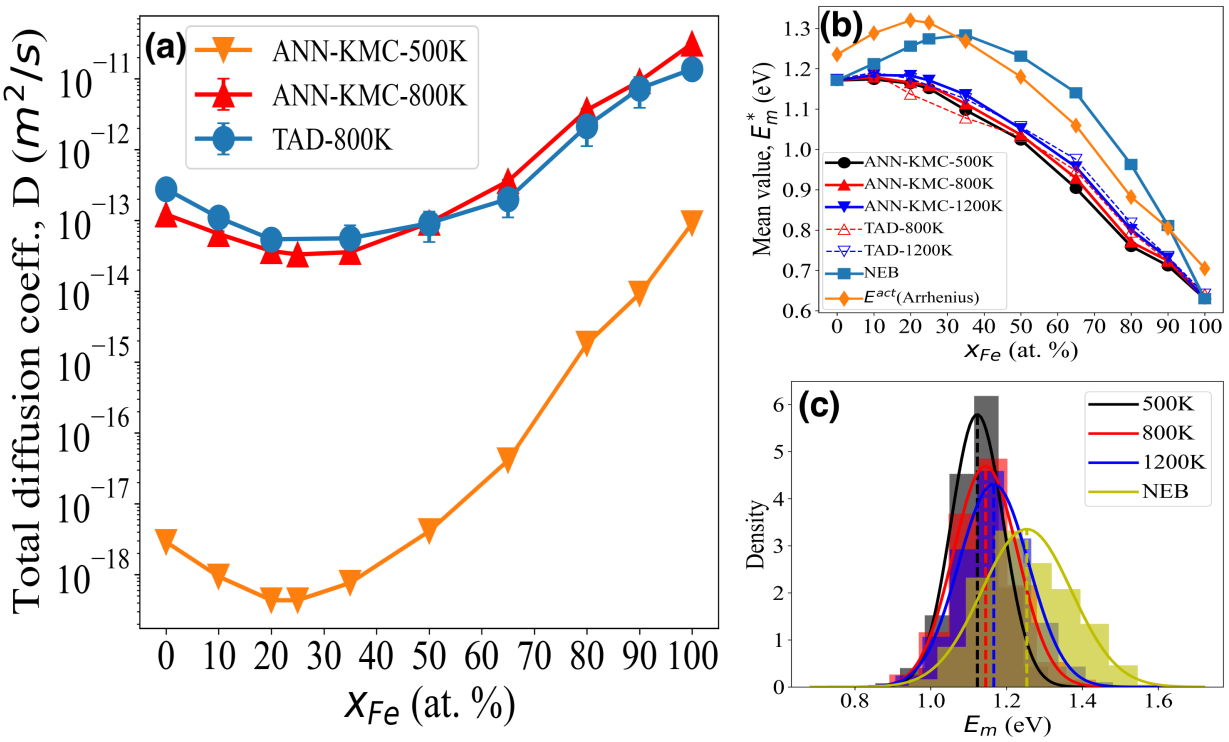

Fig.7 (a) Comparison of the total diffusivities in the MMC group alloys obtained by TAD and ANN-KMC at 800K as a function of composition. The ANN-KMC results at a low temperature (500 K) that is unreachable to MD is also shown. (b) Two methods for estimating the effective vacancy migration energies at different compositions: Arrhenius fitting of ANN-KMC diffusivities at 500, 800 – 1300 K; The average of the migration barriers from executed (accepted) jumps during ANN-KMC and TAD simulations. For comparison, the average of the static NEB barriers at 0 K is also shown at each composition. (c) The migration barrier distributions from static NEB calculations at 0 K, and the ANN-KMC simulations at 500K, 800K, and 1200K for the $Ni_{80}Fe_{20}$ alloy.

Here, two methods are used to present the effective migration energies. The first is the activation energy extracted from an Arrhenius-type treatment of ANN-KMC diffusivities at all temperatures, E^{act}, as seen in previous studies [28, 59], and the second is to record each executed migration barrier in ANN-KMC and TAD, and the averaged value is displayed as E_m^*, shown in Fig. 7(b). In addition, the average of all static NEB barriers at 0 K is shown as function of composition, even though many of them may have never been visited during the vacancy diffusion. Figure 7(b) shows that E^{act} increases initially and reaches a maximum value at the percolation threshold $x_{Fe} = 20\%$, which was claimed in Ref. [28] to be one of the reasons for the sluggish diffusion in Ni-Fe alloys. The E^{act} trend is similar as the average of the NEB barriers, although for the later the maximum barrier occurs at a higher Fe concentration ($x_{Fe} \sim 35\%$). However, the E_m^* values based on the average of accepted barriers either from ANN-KMC or TAD simulations do not show a clear maximum. Instead, they exhibit a monotonically

decreasing trend. Closer examination also shows that both KMC and TAD result in higher E_m^* values at higher temperatures at each composition. Figure 7(c) presents the distribution of executed migration barriers in the $Ni_{80}Fe_{20}$ alloy (which shows significant sluggish diffusion) at 500 K, 800 K, and 1200 K from KMC simulations, as well as the distribution of the static NEB barriers. Each spectrum is fitted to a Gaussian function in which the vertical dash line indicates its mean value. It shows apparent decline in the average migration energy as well as the distribution variance as the temperature decreases. Therefore, at low temperatures vacancy jumps are primarily along the low-barrier paths, and many of them are through the fast diffuser, Fe. At $x_{Fe} = 20\%$, however, there are only limited number of percolated paths that enable the long-range Fe diffusion. Therefore, vacancy diffusion may be trapped by Fe and sluggish diffusion occurs.

3.3 ANN-KMC results of the Random group

To exclude the statistical errors potentially induced by different atomic configurations, first the random configurations in ANN-KMC simulations are created as same as those used in the MD (note the small off-lattice distortions presented in the MD structures are ignored in ANN-KMC). In the next section, our further tests demonstrate that the choice of random configurations has a negligible effect on the diffusivities predicted by ANN-KMC. Figure 8(a) shows the total vacancy diffusivities of the Random group at different compositions predicted by the ANN-KMC from 900 K – 1300 K. The MD results are also shown for comparison. It should be noted that the ANN-KMC results are independent to the MD data because the jump attempt frequencies are the same as those in the Section 4.2 (for the MMC group) and no new fitting has been conducted. Nevertheless, ANN-MKC predicts vacancy diffusivities very close to the independent MD results, demonstrating the robustness and transferability of the ANN-KMC model. The minimum diffusivity happens at x_{Fe} = 20 ~ 25% at these temperatures, which is slightly lower than the MMC group (Fig. 6(a)). Similar to the partial diffusivities in the MMC group (Fig. 6(b)), in the Random group D_{Ni}^{random} and D_{Fe}^{random} also intersect at x_{Fe} = 20 ~ 25% and the cross-over composition slightly shifts rightward (to higher Fe concentrations) as the temperature increases, as shown in Fig. 8(b). The cross-over compositions are also close to the minimum diffusivities in Fig. 8(a). Figure 8(c) shows the jump fractions of Ni and Fe at different compositions at 1200 K. Again the Fe fraction increases while Ni fraction decreases with the increasing Fe concentration. The cross-over composition is at x_{Fe} = 20 ~ 25%, which is the same as that for the minimum total diffusivities shown in Fig. 8(a) and the cross-over composition for D_{Ni}^{random} and D_{Fe}^{random} shown in Fig. 8(b).

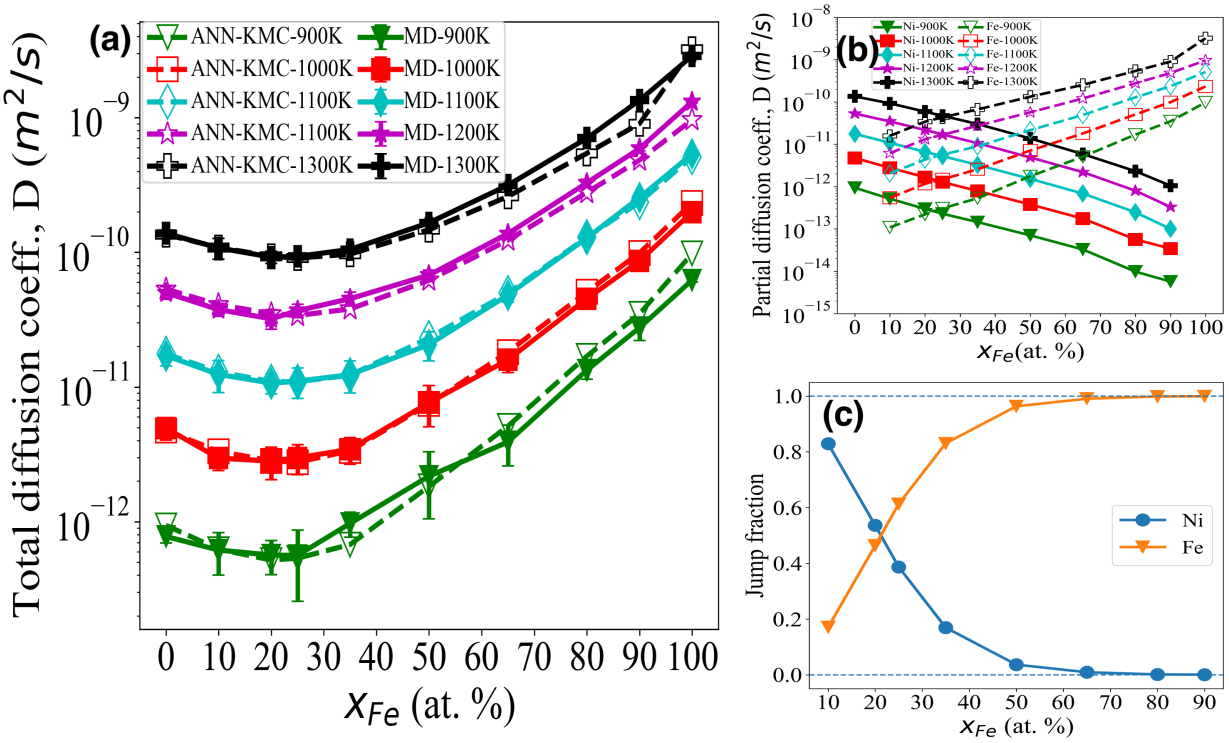

Fig. 8 (a) Total vacancy diffusivities at different Fe concentrations for the Random group alloys calculated from ANN-KMC. The independent MD results are also shown for comparison. Solid lines are MD results, and dash lines are ANN-KMC results. Each error bar represents the standard deviation of 6 – 10 MD simulations. The error bars in ANN-KMC results are ignored due to the negligible variances. (b) Partial diffusivities of Ni and Fe as a function of Fe concentration from the ANN-KMC simulations at different temperatures. (c) Jump fractions of Ni and Fe at different Fe concentrations from the ANN-KMC simulations at 1200 K.

TAD simulations are also conducted to study the vacancy diffusivities in the Random alloys. Figure 9(a) shows the good agreement between ANN-KMC and independent TAD results at all compositions at 800 K, where the minimum total diffusivity is observed at x_{Fe} = 20%. The ANN-KMC is also used to calculate the total diffusivities of Random group alloys at 500 K. In comparison to the results at 500 K in the MMC group alloys shown in Fig. 7(a), the sluggish diffusion is less pronounced, indicating that the atomic configuration or SRO can be another important factor for affecting the sluggish diffusion. The ANN-KMC results in the Random alloys at 500 K are also compared with the work of on-the-fly KMC coupled with k-ART by Osetsky et al. [28], in which a constant v_o of 10^{13} s^{-1} was used for all compositions. Although the overall trend is similar between the two studies, the diffusivities in our work are about 5 times faster than in their work. The discrepancies could be due to different v_o values as well as the different methodologies for calculating vacancy migration energies in the two studies. Similar to the analysis for the MMC group (Fig. 7(b)), the Arrhenius-type treatment of activation energy (E^{act}) and the average migration energies (E_m^*) from the accepted events in ANN-KMC and TAD

simulations at different compositions are shown in Fig. 9(b). Again, while the E^{act} curve has a convex shape with a maximal value at about x_{Fe} = 20%, the E_m^* curves for both ANN-KMC and TAD simulations decrease monotonically with x_{Fe}. Both E^{act} and E_m^* are different from the average of the static NEB barriers at 0 K. Figure 9(c) shows the distribution of the accepted migration energies in the $Ni_{80}Fe_{20}$ alloy obtained by ANN-KMC at 500 K, 800 K, and 1200 K, along with the distribution of the static NEB barriers. It can be clearly seen that many high-energy NEB barriers have never been visited by the ANN-KMC. In addition, the average migration energy of executed events (E_m^*) decreases as temperature decreases. For example, E_m^* = 1.16 eV at 1200 K while E_m^* = 1.07 eV at 500 K. Again, the results indicate that vacancy diffusion is mainly through low-barrier paths at low temperatures. If the number of such low-barrier paths is limited (e.g., at x_{Fe} = 20%), sluggish diffusion occurs.

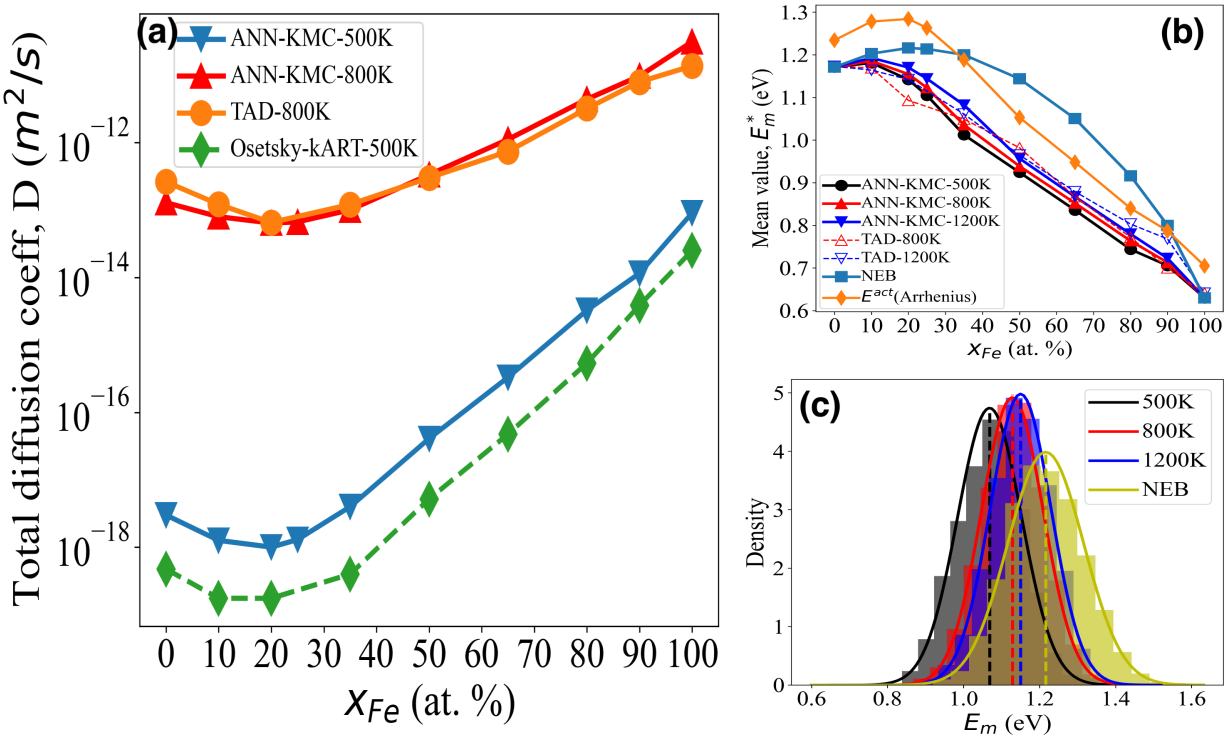

Fig. 9 (a) Comparison of the total diffusivities in the Random alloys obtained by TAD and ANN-KMC at 800 K at different compositions. The ANN-KMC results at 500 K are compared with a previous on-the-fly k-ART + KMC study [28]. (b) Two methods for estimating the effective vacancy migration energies at different compositions: Arrhenius-like treatment of ANN-KMC diffusivities at 500, 800 – 1300 K; The average of migration barriers from executed jumps during ANN-KMC and TAD simulations. For comparison, the average of the static NEB barriers at 0 K is also shown at each composition. (c) Migration barrier spectra of the static NEB barriers at 0 K and from the ANN-KMC simulations at 500K, 800K, and 1200K in the $Ni_{80}Fe_{20}$ random alloy.

4. Discussion

A comparison of the results from two structural groups reveals some commonalities and discrepancies. First, both MMC and Random structures have partial diffusivities and jump fractions crossing at x_{Fe} = 20~25%, which is close to the percolation threshold of Fe at x_{Fe} = 20%. However, while these cross-over values are identical to the minimal diffusivities in Random structures, MMC structures have the minimal diffusivities at x_{Fe} = 25~35%. The main factor that causes this discrepancy could be the presence of SROs in MMC structures, which can be a non-eligible factor in diffusion processes. For example, Zhao et al. have shown that the existence of ordered phases in concentrated alloys resulted in a significant delay in defect evolution [60]. As shown in Fig. 1(b), the MMC-optimized $Ni_{65}Fe_{35}$ (x_{Fe} = 35%) alloy has the largest extent of SROs. The SROs may hinder the vacancy diffusion, changing the minimum diffusivity from x_{Fe} = 20~25% in Random structures to 25~35% in the MMC structures. Second, both MMC and Random structures have a convex-shape curve of the effective migration energy (E^{act}) from the Arrhenius-type treatment, and the maximum value occurs at x_{Fe} = 20% in both cases, which coincides with the cross-over compositions for the partial diffusivities and jump fractions discussed above. The MMC and Random structures, therefore, may have a similar site percolation threshold around x_{Fe} = 20%, even though the MMC structures contain significant SROs. Third, the average migration barrier (E_m^*) for both MMC and Random structures, obtained by averaging the accepted migration barriers in ANN-KMC and TAD simulations, decreases monotonically with the increasing x_{Fe} or decreasing temperature. In addition, the E_m^* values are lower than the NEB and E^{act} counterparts for the same composition. Similar trends were also observed in Ref. [28] in which E_m^* from k-ART was 0.96 eV for $Ni_{50}Fe_{50}$ at 500 K while E^{act} and NEB values were 1.08 eV and 1.20 eV, respectively. The discrepancy between the average of the accepted barriers and the average of all NEB barriers indicates that only a fraction of NEB barriers are actually visited during vacancy diffusion, and most of the accepted barriers are at the smaller barrier side. Many of these smaller barriers are related to the jump of faster diffuser, Fe. When there are sufficient diffusion paths for Fe such as at high x_{Fe}, the overall diffusivity is high. However, if the Fe diffusion paths are limited such as near the Fe percolation threshold (x_{Fe} = 20%), the vacancy diffusion may be trapped by these local and disconnected small-barrier events, which leads to sluggish diffusion. As discussed earlier, however, this work also suggests that the percolation threshold may not be the sole factor for determining sluggish diffusion. Other factors including the presence of SROs can also influence it.

This work demonstrates that the combination of lattice KMC model and ANN-based on-the-fly determination of vacancy migration energy can achieve accurate results with high computational efficiency. However, the accuracy of the on-the-fly KMC is highly dependent on the quality of the ANN model. Here, the ANN-KMC approach is tested from three perspectives in order to evaluate its transferability, reliability, and practicality. First, the ANN model should be capable of predicting the migration energies for arbitrarily random configurations in a robust manner. Second, the accuracy of the ANN-KMC model should not be affected by the size of simulating system. Third, the effort required to obtain a database for training an ANN model should be feasible from a computational standpoint. To address the first concern, a set of random

structures with the same compositions and number of atoms (4,000 atoms) as those in Section 4.3, but with different random arrangements of elements, are created. Regarding the second concern, a set of larger Ni-Fe random alloy system consisting of 108,000 atoms are created. Figure 10 shows the total diffusivities calculated by ANN-KMC for the two new sets of alloys at 1000 K, along with the results for the original random alloys studied in Section 4.3. Clearly, both new sets of alloys have nearly the same diffusivities as those reported in Section 4.3 over the entire composition range, while the slight discrepancies are likely caused by the statistical errors. The results indicate that the ANN-KMC model is able to predict robust results for Ni-Fe alloys of different atomic configurations and sizes.

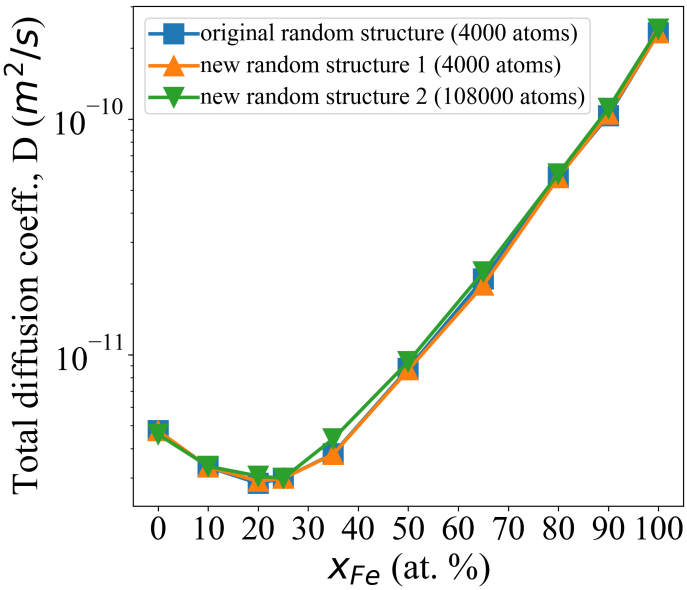

Fig. 10. Total vacancy diffusivities for two new sets of random structures of different system sizes (4,000 and 108,000) calculated by ANN-KMC at 1000 K. The results are compared with the diffusivities for the original Random structures reported in Section 4.3 at all compositions.

To address the third concern, a fraction of the NEB barrier database rather than the entire database (i.e., 32,000 data points in the MMC group) is chosen to complete the training, validation, and testing for obtaining an ANN model. Figure 11(a) illustrates the relationship between the ANN performance (using R^2 as the performance metrics) and the percentage of the database used. A two-fold effect is observed: On the one hand, the performance of the ANN model improves with the increasing percentage of database usage, as expected; On the other hand, the improvement is only significant at the early stage (for instance, from 10% to 20%); After that, the improvement is unremarkable. Next, the two ANN models that are trained with 10% and 20% of the database, respectively, are coupled with KMC to calculate total diffusivities in a random Ni-Fe system. The results are compared with the original ANN-KMC model that is based on the whole database in Section 4.3, as shown in Fig. 11(b). It is evident that the diffusivities obtained from these two new models only differ slightly from those obtained from

the original full model, and the discrepancies are more evident for the ANN model using the 10% of database. Despite of this, both models yield similar trends in diffusivity, with an overall deviation less than 20% from the full model (note the large deviation mainly appears in the 10% case), which can be considered acceptable in practice. Such a trade-off could be useful when the computational cost for obtaining the NEB barriers is high, for example, in density functional theory (DFT) calculations.

A final remark is that to achieve an accurate prediction of E_m across all compositions, the ANN model should be trained on the NEB data that cover different compositions. For example, it is found that training an ANN model with only $Ni_{50}Fe_{50}$ data is not able to predict E_m accurately across all compositions. Thus, a good strategy for building the database is to use a subset of data from each composition and include many different compositions.

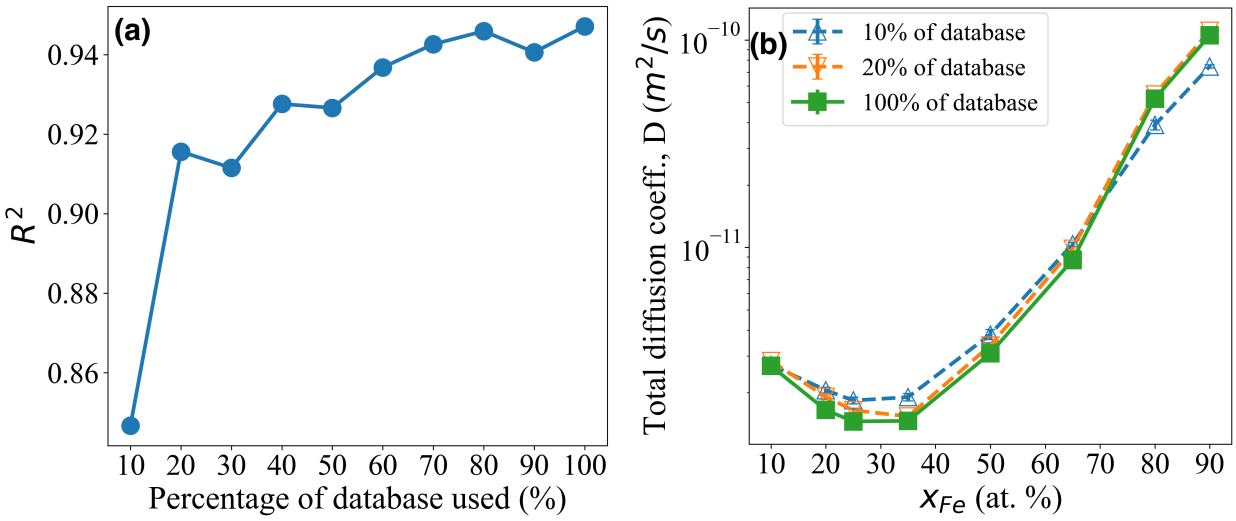

Fig. 11. (a) The R^2 performance of the ANN model using different percentages of the NEB database for obtaining an ANN model. (b) Total diffusivities predicted by the ANN-KMC models using 10%, 20%, and 100% of NEB database, respectively.

5. Conclusions

In this work, the ANN based machine learning method is coupled with KMC to study the sluggish vacancy diffusion in an fcc $Ni_{1-x}Fe_x$ (x = 0 – 1) concentrated model alloy system, in which the atomic-environment-dependent vacancy migration energy is determined by the ANN model on-the-fly. Two different groups of atomic configurations are studied: MMC-optimized structures that contain SROs and Random structures with randomly distributed Ni and Fe atoms. Using only the NEB migration barriers from the MMC structures for training, the ANN model can predict the vacancy migration barriers in the Random structures with satisfactory accuracy. A composition- and temperature-dependent jump attempt frequency model is also developed in this work, which is calibrated only using the MD data of the MMC structures at 1300 K. Using

this new jump attempt frequency model, the ANN-KMC model can predict the vacancy diffusivities in excellent agreement with independent MD and TAD results at both high and moderate temperatures, for both MMC-optimized and Random structures, and across all compositions, indicating the model has good robustness and transferability. This novel and computationally efficient ANN-KMC approach enables the simulation of vacancy diffusion at low temperatures and long timescale that are unreachable by other methods such as MD. This study also evaluates the practicability of the ANN-KMC approach by using a subset of database. The results indicate that using 20% of training data can result in a satisfactory accuracy. Therefore, the ANN-KMC approach may be applied to other studies in which the collection of the NEB barrier database is computationally expensive, such as in DFT calculations. Although this work uses a binary fcc concentrated alloy system as a model system, it is expected that the ANN-KMC approach can be generalized to other concentrated alloys with more components such as high-entropy alloys as well as other crystal structures.

For the $Ni_{1-x}Fe_x$ (x = 0 – 1) concentrated model alloy system studied in this work, a "sluggish diffusion" phenomenon is observed, with the lowest diffusivity being x_{Fe} = 20~25% for Random structures while x_{Fe} = 25~35% for MMC-optimized structures that contain SROs. Although these results support the argument that the percolation threshold (x_{Fe} = 20%) is an important factor for affecting sluggish diffusion, this work also shows that other factors such as SRO can influence the sluggish diffusion behavior. At low temperatures such as 500 K, the sluggish diffusion becomes more pronounced, particularly in the MMC-optimized structures. The average of the accepted migration barriers from ANN-KMC modeling is smaller than the average of all NEB barriers. In addition, it decreases with the decreasing temperature. The results indicate that smaller-barrier events, which are mainly related to the faster diffuser Fe, play significant roles on affecting the total diffusivity. When the Fe migration paths are limited and disconnected such as at $x_{Fe} < 25\%$, the smaller-barrier events may trap the vacancy locally to inhibit its long-range diffusion and thus induce sluggish diffusion. When the Fe migration paths are abundant at high Fe concentrations, these smaller-barrier events can dominate the long-range diffusion and lead to fast diffusivities.

Acknowledgements

This material is based upon work supported by the U.S. National Science Foundation under Grant No. 1847780. The authors also acknowledge the Advanced Research Computing (ARC) at Virginia Tech for providing the high performance computing resources to this work.

References

[1] B. Gludovatz *et al*., A fracture-resistant high-entropy alloy for cryogenic applications. *Science* **345**, 1153-1158 (2014).

[2] Y. F. Ye, Q. Wang, J. Lu, C. T. Liu, Y. Yang, High-entropy alloy: challenges and prospects. *Materials Today* **19**, 349-362 (2016).

[3] Y. Qiu, S. Thomas, M. A. Gibson, H. L. Fraser, N. Birbilis, Corrosion of high entropy alloys. *npj Materials degradation* **1**, 1-18 (2017).

[4] Y. W. Zhang *et al*., Atomic-level heterogeneity and defect dynamics in concentrated solid-solution alloys. *Current Opinion in Solid State & Materials Science* **21**, 221-237 (2017).

[5] C. Lu *et al*., Enhancing radiation tolerance by controlling defect mobility and migration pathways in multicomponent single-phase alloys. *Nature Communications* **7**, 1-8 (2016).

[6] J. Dabrowa *et al*., Interdiffusion in the FCC-structured Al-Co-Cr-Fe-Ni high entropy alloys: Experimental studies and numerical simulations. *Journal of Alloys and Compounds* **674**, 455-462 (2016).

[7] K. Y. Tsai, M. H. Tsai, J. W. Yeh, Sluggish diffusion in Co-Cr-Fe-Mn-Ni high-entropy alloys. *Acta Materialia* **61**, 4887-4897 (2013).

[8] M. H. Tsai *et al*., Thermal Stability and Performance of NbSiTaTiZr High-Entropy Alloy Barrier for Copper Metallization. *Journal of the Electrochemical Society* **158**, H1161-H1165 (2011).

[9] A. Paul, Comments on "Sluggish diffusion in Co-Cr-Fe-Mn-Ni high-entropy alloys" by KY Tsai, MH Tsai and JW Yeh, Acta Materialia 61 (2013) 4887-4897. *Scripta Materialia* **135**, 153-157 (2017).

[10] W. Kucza *et al*., Studies of "sluggish diffusion" effect in Co-Cr-Fe-Mn-Ni, Co-Cr-Fe-Ni and Co-Fe-Mn-Ni high entropy alloys; determination of tracer diffusivities by combinatorial approach. *Journal of Alloys and Compounds* **731**, 920-928 (2018).

[11] D. B. Miracle, O. N. Senkov, A critical review of high entropy alloys and related concepts. *Acta Materialia* **122**, 448-511 (2017).

[12] M. S. Daw, M. Chandross, Sluggish diffusion in random equimolar FCC alloys. *Physical Review Materials* **5**, 043603 (2021).

[13] A. Seoane, D. Farkas, X.-M. Bai, Influence of compositional complexity on species diffusion behavior in high-entropy solid-solution alloys. *Journal of Materials Research* **37**, 1403-1415 (2022).

[14] Y. N. Osetsky, L. K. Béland, R. E. Stoller, Specific features of defect and mass transport in concentrated fcc alloys. *Acta Materialia* **115**, 364-371 (2016).

[15] S. Zhao, Y. Osetsky, Y. Zhang, Preferential diffusion in concentrated solid solution alloys: NiFe, NiCo and NiCoCr. *Acta Materialia* **128**, 391-399 (2017).

[16] F. X. Zhang *et al*., Local Structure and Short-Range Order in a NiCoCr Solid Solution Alloy. *Physical Review Letters* **118**, 205501 (2017).

[17] J. Y. He *et al*., A precipitation-hardened high-entropy alloy with outstanding tensile properties. *Acta Materialia* **102**, 187-196 (2016).

[18] T. Yang *et al*., Multicomponent intermetallic nanoparticles and superb mechanical behaviors of complex alloys. *Science* **362**, 933-937 (2018).

[19] A. Le Claire, Solute diffusion in dilute alloys. *Journal of Nuclear Materials* **69**, 70-96 (1978).

[20] S. Choudhury *et al*., Ab-initio based modeling of diffusion in dilute bcc Fe–Ni and Fe–Cr alloys and implications for radiation induced segregation. *Journal of Nuclear Materials* **411**, 1-14 (2011).
[21] V. Ramunni, M. Pascuet, Mobility of U solutes in fcc Al: a theoretical model. *Procedia Materials Science* **8**, 451-460 (2015).
[22] X.-M. Bai, Y. Zhang, M. R. Tonks, Strain effects on oxygen transport in tetragonal zirconium dioxide. *Physical Chemistry Chemical Physics* **15**, 19438-19449 (2013).
[23] C. C. Fu, J. Dalla Torre, F. Willaime, J. L. Bocquet, A. Barbu, Multiscale modelling of defect kinetics in irradiated iron. *Nature Materials* **4**, 68-74 (2005).
[24] A. F. Voter, "Introduction to the kinetic Monte Carlo method" in Radiation effects in solids. (Springer, 2007), pp. 1-23.
[25] M. Trochet, N. Mousseau, L. K. Béland, G. Henkelman, Off-lattice kinetic Monte Carlo methods. *Handbook of Materials Modeling: Methods: Theory and Modeling*, 715-743 (2020).
[26] F. El-Mellouhi, N. Mousseau, L. J. Lewis, Kinetic activation-relaxation technique: An off-lattice self-learning kinetic Monte Carlo algorithm. *Physical Review B* **78**, 153202 (2008).
[27] L. K. Béland, P. Brommer, F. El-Mellouhi, J.-F. Joly, N. Mousseau, Kinetic activation-relaxation technique. *Physical Review E* **84**, 046704 (2011).
[28] Y. N. Osetsky, L. K. Beland, A. V. Barashev, Y. W. Zhang, On the existence and origin of sluggish diffusion in chemically disordered concentrated alloys. *Current Opinion in Solid State & Materials Science* **22**, 65-74 (2018).
[29] H. Jónsson, G. Mills, K. W. Jacobsen, "Nudged elastic band method for finding minimum energy paths of transitions" in Classical and quantum dynamics in condensed phase simulations. (World Scientific, 1998), pp. 385-404.
[30] G. Henkelman, H. Jónsson, Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points. *The Journal of chemical physics* **113**, 9978-9985 (2000).
[31] G. Henkelman, B. P. Uberuaga, H. Jonsson, A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *Journal of Chemical Physics* **113**, 9901-9904 (2000).
[32] N. Mousseau *et al*., The activation-relaxation technique: Art nouveau and kinetic art. *Journal of Atomic and Molecular Physics* **2012**, 14 (2012).
[33] K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev, A. Walsh, Machine learning for molecular and materials science. *Nature* **559**, 547-555 (2018).
[34] T. Mueller, A. G. Kusne, R. Ramprasad, Machine learning in materials science: Recent progress and emerging applications. *Reviews in computational chemistry* **29**, 186-273 (2016).
[35] Z. Fan, B. Xing, P. Cao, Predicting path-dependent diffusion barrier spectra in vast compositional space of multi-principal element alloys via convolutional neural networks. *Acta Materialia* **237**, 118159 (2022).
[36] B. Xu *et al*., Exploring the influence of percolation on vacancy-mediated diffusion in CoCrNi multi-principal element alloys. *Materials & Design*, 111238 (2022).
[37] W. Huang, P. Martin, H. L. Zhuang, Machine-learning phase prediction of high-entropy alloys. *Acta Materialia* **169**, 225-236 (2019).

[38] J. M. Cowley, An Approximate Theory of Order in Alloys. *Physical Review* **77**, 669-675 (1950).
[39] Z. Shen *et al*., Kinetic Monte Carlo simulation framework for chemical short-range order formation kinetics in a multi-principal-element alloy. *Computational Materials Science* **198**, 110670 (2021).
[40] M. R. Sorensen, A. F. Voter, Temperature-accelerated dynamics for simulation of infrequent events. *The Journal of chemical physics* **112**, 9599-9606 (2000).
[41] G. Bonny, D. Terentyev, R. C. Pasianot, S. Ponce, A. Bakaev, Interatomic potential to study plasticity in stainless steels: the FeNiCr model alloy. *Modelling and Simulation in Materials Science and Engineering* **19**, 085008 (2011).
[42] S. Plimpton, Fast parallel algorithms for short-range molecular dynamics. *Journal of computational physics* **117**, 1-19 (1995).
[43] E. Wigner, F. Seitz, On the constitution of metallic sodium. *Physical Review* **43**, 804 (1933).
[44] S. M. Shakhno, O. P. Gnatyshyn, Algorithm for the solution of a nonlinear least squares problem under secondary conditions. *Zeitschrift Fur Angewandte Mathematik Und Mechanik* **81**, S1023-S1024 (2001).
[45] N. Castin, J. R. Fernandez, R. C. Pasianot, Predicting vacancy migration energies in lattice-free environments using artificial neural networks. *Computational Materials Science* **84**, 217-225 (2014).
[46] N. Castin, L. Malerba, Prediction of point-defect migration energy barriers in alloys using artificial intelligence for atomistic kinetic Monte Carlo applications. *Nuclear Instruments & Methods in Physics Research Section B-Beam Interactions with Materials and Atoms* **267**, 3148-3151 (2009).
[47] M. Leetmaa, N. V. Skorodumova, KMCLib: A general framework for lattice kinetic Monte Carlo (KMC) simulations. *Computer Physics Communications* **185**, 2340-2349 (2014).
[48] H. Mehrer, *Diffusion in solids: fundamentals, methods, materials, diffusion-controlled processes* (Springer Science & Business Media, 2007), vol. 155.
[49] V. A. Borodin, P. V. Vladimirov, A. Moslang, Lattice kinetic Monte-Carlo modelling of helium-vacancy cluster formation in bcc iron. *Journal of Nuclear Materials* **367**, 286-291 (2007).
[50] C. S. Deo *et al*., Helium bubble nucleation in bcc iron studied by kinetic Monte Carlo simulations. *Journal of Nuclear Materials* **361**, 141-148 (2007).
[51] D. Terentyev, N. Juslin, K. Nordlund, N. Sandberg, Fast three dimensional migration of He clusters in bcc Fe and Fe-Cr alloys. *Journal of Applied Physics* **105**, 103509 (2009).
[52] M. I. Pascuet, N. Castin, C. S. Becquart, L. Malerba, Stability and mobility of Cu-vacancy clusters in Fe-Cu alloys: A computational study based on the use of artificial neural networks for energy barrier calculations. *Journal of Nuclear Materials* **412**, 106-115 (2011).
[53] W. G. Hoover, Canonical dynamics: Equilibrium phase-space distributions. *Physical Review A* **31**, 1695 (1985).
[54] A. F. Voter, F. Montalenti, T. C. Germann, Extending the time scale in atomistic simulation of materials. *Annual review of materials research* **32**, 321-346 (2002).
[55] B. P. Uberuaga, A. F. Voter, "Accelerated molecular dynamics methods" in Radiation Effects in Solids. (Springer, 2007), pp. 25-43.

[56] X.-M. Bai, A. F. Voter, R. G. Hoagland, M. Nastasi, B. P. Uberuaga, Efficient annealing of radiation damage near grain boundaries via interstitial emission. *Science* **327**, 1631-1634 (2010).
[57] X.-M. Bai, A. El-Azab, J. Yu, T. R. Allen, Migration mechanisms of oxygen interstitial clusters in UO2. *Journal of Physics: Condensed Matter* **25**, 015003 (2012).
[58] L. Kong, L. J. Lewis, Transition state theory of the preexponential factors for self-diffusion on Cu, Ag, and Ni surfaces. *Physical Review B* **74**, 073412 (2006).
[59] K. Ferasat *et al*., Accelerated kinetic Monte Carlo: A case study; vacancy and dumbbell interstitial diffusion traps in concentrated solid solution alloys. *Journal of Chemical Physics* **153**, 074109 (2020).
[60] S. J. Zhao, Y. Osetsky, Y. W. Zhang, Diffusion of point defects in ordered and disordered Ni-Fe alloys. *Journal of Alloys and Compounds* **805**, 1175-1183 (2019).