High-Throughput Machine Learning - Kinetic Monte Carlo Framework for Diffusion Studies in Equiatomic and Non-equiatomic FeNiCrCoCu High-Entropy Alloys

Wenjiang Huang , Diana Farkas , Xian-Ming Bai

PII: S2589-1529(23)00293-4
DOI: https://doi.org/10.1016/j.mtla.2023.101966
Reference: MTLA 101966

To appear in: *Materialia*

Received date: 23 August 2023
Accepted date: 16 November 2023

Please cite this article as: Wenjiang Huang , Diana Farkas , Xian-Ming Bai , High-Throughput Machine Learning - Kinetic Monte Carlo Framework for Diffusion Studies in Equiatomic and Non-equiatomic FeNiCrCoCu High-Entropy Alloys, *Materialia* (2023), doi: https://doi.org/10.1016/j.mtla.2023.101966

High-Throughput Machine Learning - Kinetic Monte Carlo Framework for Diffusion Studies in Equiatomic and Non-equiatomic FeNiCrCoCu High-Entropy Alloys

Wenjiang Huang, Diana Farkas, Xian-Ming Bai*

Department of Materials Science and Engineering, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, USA

*Corresponding author.

xmbai@vt.edu

Abstract

Sluggish diffusion is postulated as an underlying mechanism for many unique properties in high-entropy alloys (HEAs). However, its existence remains a subject of debate. Due to the challenges of exploring the vast composition space, to date most experimental and computational diffusion studies have been limited to equiatomic HEA compositions. To develop a high-throughput approach to study sluggish diffusion in a wide range of non-equiatomic compositions, this work presents an innovative artificial neural network (ANN) based machine learning model that can predict the vacancy migration barriers for arbitrary local atomic configurations in a model FeNiCrCoCu HEA system. Remarkably, the model utilizes the training data exclusively from the equiatomic HEA while it can accurately predict barriers in non-equiatomic HEAs as well as in the quaternary, ternary, and binary sub-systems. The ANN model is implemented as an on-the-fly barrier calculator for kinetic Monte Carlo (KMC) simulations, achieving diffusivities nearly identical to the independent molecular dynamics (MD) simulations but with far higher efficiency. The high-throughput ANN-KMC method is then used to study the diffusion behavior in 1,500 non-equiatomic HEA compositions. It is found that although the sluggish diffusion is not evident in the equiatomic HEA, it does exist in many non-equiatomic compositions. The compositions, complex potential energy landscapes (PEL), and percolation effect of the fastest diffuser (Cu) in these sluggish compositions are analyzed, which could provide valuable insights for the experimental HEA designs.

Keywords

High-entropy alloys; Sluggish diffusion; Machine learning informed kinetic Monte Carlo; Non-equiatomic compositions; High-throughput modeling.

1. Introduction

High-entropy alloys (HEAs), which typically contain five or more principal elements with intrinsic chemical disorder [1, 2], have garnered immerse attention in recent years, ushering in what can be aptly described as a "High-Entropy Materials" age [3]. The heightened interest stems from their exceptional mechanical and structural properties, such as strength retention at high temperatures [4], exceptional fracture toughness at cryogenic temperatures [5, 6], good corrosion resistance [7, 8], and outstanding radiation resistance [9]. In addition, some dual-phase HEAs are among the few alloys known to overcome the strength-ductility tradeoff [10]. It has been suggested that some of these excellent properties may be tied to the possible sluggish diffusion kinetics of HEAs [11]. Specifically, the random arrangement of principal elements in HEAs leads to substantial fluctuations in the potential energy landscape (PEL), creating enormous atomic traps and barriers that could result in a reduced diffusivity. As such, the concept of "sluggish diffusion" has been commonly quoted to explain many of the outstanding properties in HEAs [11-13]. However, numerous experimental and computational studies have shown that sluggish diffusion in HEAs or in concentrated alloys is a rather complex phenomenon, yielding conflicting results [14-19]. To date, the existence of sluggish diffusion remains a subject of ongoing debate. Therefore, understanding the mechanism of sluggish diffusion — if present, and designing HEAs with enhanced performance continue to be a challenging endeavor, given the complexity of HEAs and the conflicting results presented in various studies.

Understanding the controlling mechanism of sluggish diffusion is fundamentally rooted in deciphering how atomic transport is influenced by the complex PEL or local atomic configuration (LAC) in HEAs. While the experimental techniques such as diffusion couples or tracer techniques can accurately measure the diffusion coefficients [15, 20], they provide little insights about the underlying diffusion mechanisms. As an alternative, atomic-level simulation methods such as molecular dynamics (MD) are often employed to study the diffusion mechanisms in concentrated alloys including HEAs [16, 18, 19]. Previous simulation studies in binary or ternary concentrated alloys provided valuable insights about the sluggish diffusion mechanisms, which could inform our approach as we delve deeper into the intricacies of HEAs. For example, Osetsky et al. have shown that the compositional-dependent diffusivity in a Ni-Fe concentrated alloy reaches its minimum at the percolation threshold, about 20 at.% of Fe, instead of the equiatomic composition. The coincidence of retarded diffusion with the percolation threshold suggests that the local trapping effects could be caused by the faster-diffusing species, Fe, before it reaches the percolation threshold [16, 21]. However, recent work on concentrated CoCrNi ternary alloys revealed that the percolation threshold does not always result in sluggishness, and the role of PEL may be more critical [22]. In our recent MD studies of diffusion in an equiatomic FeNiCrCoCu HEA, no sluggish effect is observed in the vacancy-mediated bulk diffusion [18], although the diffusion in a $\Sigma5(210)$ grain boundary is sluggish [19]. These findings suggest a multitude of factors can affect defect/species diffusion in HEAs, many of which remain to be explored. For example, sluggish diffusion could exist at some non-equiatomic compositions in HEAs. Given their enormously large compositional space, it is very costly and unfeasible to use experimental techniques to search for the sluggish compositions.

Another challenge for experiments is the lack of a fair reference system for determining the sluggish effect, while in modeling a hypothetical average atom (AA) material [23] can be used, which predicts the same average bulk properties as in the HEA but without HEA's compositional complexity [18]. Therefore, to solve these challenges it is imperative to design a high-throughput modeling framework to study the diffusion behavior in HEAs of both equiatomic and non-equiatomic compositions.

Although MD is a powerful tool for investigating atomic-level diffusion mechanisms, its applicability is limited by its inherently short timescale (typically at a few tens of nanoseconds). For vacancy-mediated diffusion studies, the applicable temperature range is typically limited to high temperatures [18] due to its inherent timescale limitation. However, multiple diffusion mechanisms can happen simultaneously at high temperatures while some of them may not be relevant at low temperatures of interest. Even though the timescale of MD can be extended to microseconds through modern supercomputers, the problem may still persist for studying the diffusion in concentrated alloys. For instance, in a 4000-atom Ni-Fe alloy at a high temperature of 1200 K, a microsecond-scale MD simulation (a very long simulation time for MD) captured only about 4500 vacancy jumps [24]. Consequently, the complex atomic environment in concentrated alloys may not be visited sufficiently. In addition, MD simulations do not perform well at low temperatures of interest so that some rare but important diffusion events might not be captured by MD. In this regard, the kinetic Monte Carlo (KMC) method has been employed in attempts to extend the diffusion simulation timescale as well as the temperature range for HEAs, as reported in [25, 26]. However, conventional KMC encounters difficulties of describing the rough PEL in HEAs because the vacancy migration barrier varies significantly with the LAC. Several approaches have been developed to approximate the rough PEL, such as random sampling method and species-average migration barrier method [25-27]. However, these approaches over-simplify the effects of the varying atomic environment surrounding the diffusing objects, which may lead to inaccurate conclusions. A reliable KMC simulation requires an accurate description of the candidate migration barriers as a prior knowledge at each defect jump event. This, in turn, demands an on-the-fly calculation of migration barriers using the nudge elastic band (NEB) [28] or the activation relaxation technique nouveau (ARTn) [29], which leads to an immense computational cost and is often impractical.

Alternatively, machine learning (ML) can serve as an efficient tool to map the dependency of PEL or migration barrier on the LAC surrounded by a defect and/or the migrating solute. Once a well-trained ML model is established, the computationally expensive on-the-fly calculation of migration barriers in KMC can be replaced, enabling long-time and accurate simulations of defect/species diffusion in concentrated alloys including HEAs. Previous studies have demonstrated that the artificial neural network (ANN), which is a widely used ML method, can effectively serve as an on-the-fly barrier calculator for KMC modeling in concentrated alloys with up to four components [22, 30-32], including our recent ANN-KMC modeling of vacancy-mediated diffusion in concentrated Ni-Fe alloys for different compositions, short-range orders (SROs), and temperatures [33]. The accuracy of the ML model is a prerequisite for developing such a reliable ML-KMC framework. However, previous studies have shown that the prediction

performance tends to decline with the rapidly increasing complexity of LACs as the number of alloying species increases [22, 30]. For instance, Xu et al. reported that their ANN models achieved mean absolute errors (MAEs) of 15, 35, and 88 meV for their binary, ternary, and quaternary alloys, respectively [30]. Therefore, particular attention should be given to constructing an accurate ML model for HEAs when more components are present. In addition, one limitation of current ML models is their limited predictability, because they might only be applied to the systems covered in the training dataset. Specifically, a ML model trained on the equiatomic composition may not be capable for non-equiatomic compositions or their sub-systems. As a result, a different ML model should be trained for each distinct composition, leading to significant inefficiency when it comes to deal with the vast composition space of HEAs.

In this work, we present the first comprehensive effort to combine ANN-based ML method and KMC for studying the possible sluggish diffusion in five-component HEAs, using FeNiCrCoCu as a model HEA system and with the focus on its non-equiatomic compositions. This HEA system is described by an EAM (Embedded Atom Method) potential developed by one of present authors (Farkas) [34]. There are a few reasons for choosing this HEA system. First, it predicts the stable face-centered-cubic (FCC) phase over a wide range of compositions so that the multi-phase complexity in some HEA systems [35] is excluded. Second, it has a low heat of mixing (-0.0002 kJ/mol) so that the ordering effect is small. Third, its pure components have a wide spread of diffusivities (e.g., Cu diffuses much faster than Ni and Fe in this model system) [18] so that the role of different species diffusivities on sluggish diffusion can be explored. To develop a robust ML model for predicting the vacancy migration barriers, special considerations have been given to the construction of LACs, which serve as training features (descriptors) for our ANN model. Remarkably, our ANN model is trained using a modest dataset of about 17,000 migration barrier samples only from the equiatomic FeNiCrCoCu composition, yet it is capable of accurately predicting non-equiatomic compositions as well as its respective sub-systems, i.e., quaternary, ternary and binary alloys. Based on the success of our previous ANN-KMC modeling work for the Ni-Fe binary system [33], a species- and temperature-dependent jump attempt frequency model has also been developed for the current HEA system, which is derived only from pure components. In addition, an average-atom KMC (AA-KMC) model has been developed to approximate the AA materials without the need of developing a distinct AA interatomic potential [23] for each HEA composition, serving as references for determining the existence of sluggish diffusion. The vacancy-mediated diffusivity obtained from our ANN-KMC and AA-KMC models show excellent agreement with the independent MD simulations, at least within MD's accessible temperature ranges. The ANN-KMC and AA-KMC methods are then used to study the $(FeNiCrCo)_{1-x}Cu_x$ alloys with x = 0 – 30%, with the aim of determining if the percolation of the fastest diffuser (Cu) leads to sluggish diffusion. More importantly, the two high-throughput methods are utilized to explore the extensive compositional space of HEAs (1,500 distinctive non-equiatomic HEA compositions) to study the controversial sluggish diffusion effect.

2. Methods

2.1 Generating datasets of vacancy migration barriers

To collect the training dataset for the ANN model, the NEB method implemented in the LAMMPS [36] was used to calculate vacancy migration barriers in an equiatomic FeNiCrCoCu HEA. A $10a_0 \times 10a_0 \times 10a_0$ FCC crystal structure with randomly populated Fe, Ni, Cr, Co, and Cu elements at the equimolar ratio (20 at.% each) was constructed. As mentioned earlier, the EAM potential developed by Farkas et al. was used to describe the interatomic interactions [34]. In each NEB calculation, a vacancy was created by removing an atom and then one of its twelve nearest neighboring atoms was randomly selected to be the moving atom. Once this vacancy-moving atom path was established, the path remained fixed while the atoms (including the moving atom) in the entire system were randomly assembled to constitute the target composition. This random assembly process was repeated about 17,000 times to generate different LACs around the vacancy migration path. Both the initial and final states of a migration event were relaxed by the conjugate gradient algorithm with the force stopping tolerance of 10^{-2} eV/Å and energy convergence criterion of 0.0 eV. The maximum number of iterations was 2,000. In all NEB calculations, the number of NEB images (excluding the initial and final states) was 13 and the spring constant was set to 1.0 eV/Å^2. In this manner, a dataset of ~17,000 samples was collected from the equiatomic FeNiCrCoCu composition, serving as the basis of training and testing data for the ANN model. To assess the generalizability of the trained ANN model, multiple supplementary datasets were also produced using the same NEB approach. Specifically, NEB calculations were carried out on an assortment of randomly chosen non-equiatomic FeNiCrCoCu alloys, yielding approximately 2,500 migration barrier samples. In addition, 2,000 migration barrier samples were collected for each equiatomic sub-system of the FeNiCrCoCu alloy, including five quaternary (i.e., 10,000 barriers in total), ten ternary, and ten binary alloys. It should be emphasized that these supplementary datasets were not included in the ANN model training. Instead, they were used as independent data to verify the predictability of the ANN model.

2.2 Training the ANN model

In our previous work, it has been demonstrated that an ANN-based ML model is able to accurately predict vacancy migration barriers in a concentrated binary Ni-Fe alloy of any composition using LAC information as a descriptor [33]. In this work, a more robust LAC representation is introduced: first a raw LAC is constructed then it is transformed to an encoded LAC. The raw LAC collects all element types up to the i^{th} nearest neighboring (NN) shell surrounding both the vacancy site and the migrating atom. Given the one-to-one correspondence between the input training features and neurons in the ANN model, it becomes necessary to maintain the consistency of atom sequence in the raw LACs. In this regard, the "Adaptive Path Alignment (APA)" method is designed, as depicted in Fig. 1. Initially, a "reference path" is created to represent the migration path for the pre-determined vacancy-moving-atom pair described in Section 2.1. For other migration paths, their local coordinate systems are rotated to

align with the chosen "reference path". Note that the choice of the "reference path" does not impact the final ANN performance and the primary objective is to ensure that the input LAC sequence maintains a consistent relative ordering. A similar approach called the "Coordinate System Transformation (CST)" method, has been previously documented in [32].

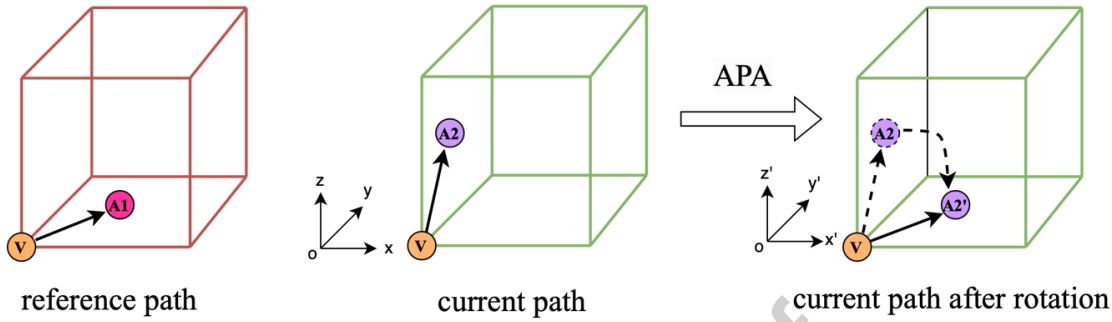


Fig. 1. Schematic illustration of the "Adaptive Path Alignment" (APA) method. The reference path (V-A1) corresponds to the pre-determined vacancy-moving-atom pair in Section 2.1. For the current migration path (V–A2), the local coordinate is rotated to a new position (V-A2'), which aligns with the reference path.

Following the APA method, with a given migration event, the raw LAC around the vacancy-moving-atom pair up to i^{th} NN shell can be established. As an example, the first two NN shells, as well as the total raw LAC can be represented below,

$LAC(raw, 1^{st}nn) = W_{1nn}[3\ 2\ 4\ 2\ 5\ 3\ 2\ 5\ 3\ 4\ 3\ 3\ 2\ 5\ 1\ 5\ 2\ 5\ 5]$,

$LAC(raw, 2^{nd}nn) = W_{2nn}[1\ 4\ 5\ 5\ 4\ 4\ 2\ 2\ 2\ 1\ 1\ 4]$,

$LAC(raw, total) = [LAC(raw, 1^{st}nn), LAC(raw, 2^{nd}nn), \ldots, LAC(raw, i^{th}nn)]$,

where the numbers in each list represent the element types, with Fe, Ni, Cr, Co, and Cu denoted by 1, 2, 3, 4, and 5, respectively. W_{1nn} and W_{2nn} signify the initial training weight coefficients for the first and second NN shells, respectively. A closer NN shell is assigned with a larger initial weight, which is then trainable during the ANN model optimization. The 1st NN shell LAC combines the 1st NN atoms around the vacancy and the moving atom. Since four of them overlap, there are a total of 19 atoms, including the moving atom (which is the 1st atom in the list) but not the vacancy site. The 2nd NN shell LAC consists of 12 atoms, some of which overlap with atoms in the 1st NN shell LAC. This overlap occurs because an atom could simultaneously serve as the first NN atom for the vacancy and as the second NN atom for the moving atom, and vice versa. Therefore, even though those atoms are already included in the 1st NN LAC, they still appear in the 2nd NN LAC. The final total raw LAC is the combination of all considered NN shells.

While the raw LAC representation works well for simpler systems, such as in a binary A-B alloy system where "1" represents A and "2" represents B [33, 37, 38], it can cause a significant imbalance in describing HEAs. In particular, using numerical values like "5" for Cu and "1" for Fe unintentionally emphasizes the role of Cu while downplays the importance of Fe,

leading to a biased training during the gradient descent optimization process. To address this issue and ensure that all species are treated equally, we employed the "one-hot" representation with a dimension of five to encode the raw LACs. As a result, the aforementioned raw LAC for the first NN shell ([3 2 4 2 5 3 ...]) is reformulated in this new encoding scheme,

$$LAC(1^{st}nn) = W_{1nn}\begin{bmatrix}000000\\010100\\100001\cdots\\001000\\000010\end{bmatrix},$$

where each column represents one neighboring atom and the position of "1" indicates its type. The rest of NN shells follow the same encoding method. To choose the ANN architecture, a straightforward three-layer structure was constructed, consisting of an input layer, a hidden layer with ten neurons, and an output layer. The encoded LACs serve as the input features for the ANN model. To train and optimize the parameters in the ANN model, the Levenberg-Marquardt method [39] was used, which is a hybrid method that combines the Gauss-Newton update and gradient descent approaches to search for an optimal solution.

To determine the optimal number of NN shells, the number of NN shells in LACs was progressively increased, from only the first NN shell up to a maximum of nine NN shells. It showed that the ANN performance improved with the number of NN shells, as expected. However, a satisfactory performance can be obtained by including only the first three NN shells. Details will be provided in the **Results** section. After the ANN model was developed based on the equiatomic composition, multiple supplementary (but independent) datasets as mentioned in Section 2.1 were used to evaluate the generalizability of the ANN model for the non-equiatomic HEAs and equiatomic sub-systems that were not covered in the ANN model training and testing.

2.3 Coupling ANN and KMC models

Leveraging the well-trained ANN model, the migration barrier for any given LAC can be accurately predicted in an efficient manner. This allows us to treat the ANN model as an on-the-fly barrier calculator for the KMC model (ANN-KMC), which is implemented in the KMCLib package [40] by the present author (Huang). The size of the ANN-KMC system is also $10a_0 \times 10a_0 \times 10a_0$. For each vacancy jump event *i*, the ANN model collects its LAC information and predicts the migration barriers across all twelve possible migration paths, which are then used to calculate corresponding jump rates,

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

where $E_{m,ij}$ is the ANN-predicted migration barrier for the jump event *i* along a candidate migration path *j*; ν_o is the jump attempt frequency, k_B is the Boltzmann constant, and *T* is the absolute temperature. The KMC then selects a migration path based on the jump rates and the elapsed time is calculated based on the residence time algorithm [41]. In this work, 300,000

ANN-KMC steps were simulated for each simulation condition to ensure sufficient jump statistics. Using the resulting atomic trajectories, the vacancy diffusion coefficient can be calculated using the sum of atomic square displacements (ASD) of all atoms in the system [33],

$$D = \frac{ASD}{6t}, \quad (2)$$

where t is the simulation time and the factor of 6 is for the 3D diffusion. For each condition, six independent simulations were conducted and the average diffusivity is reported in this work. It should be noted that the diffusivity calculated here is the vacancy mobility rather than the self-diffusivity. For the latter, the vacancy formation energy distribution should be included, which can be predicted by Zhang et al.'s model [42]. Our ANN model does include the formation energy difference along each migration path because the forward and backward migration barriers are different, giving a direction to the migration path. However, the inclusion of the complete vacancy formation energy distribution in the ANN-KMC modeling is beyond the scope of this work and can be a future research topic.

2.4 MD simulations

To validate the reliability of the ANN-KMC results, independent MD simulations of single-vacancy diffusion in a $10a_0 \times 10a_0 \times 10a_0$ FCC system were conducted as benchmarks for considered systems within the accessible temperature range of MD, from 1600 to 2000 K, using LAMMPS [36]. It should be noted that the EAM potential used in this work predicts the equilibrium melting temperature (solid-liquid coexistence temperature) for the equiatomic HEA is about 2070 K [18]. This is somewhat higher than the melting temperatures of many HEAs in experiments, which are typically below 1800 K [11, 43, 44]. We confirmed that all the simulation systems remained crystalline. The MD simulations were performed in a NPT ensemble (constant number of atoms, pressure, and temperature), with the Nose-Hoover-style thermostat and barostat [45, 46] controlling the system temperature and pressure (at zero bar), respectively. A time step of 2 fs was employed in all MD simulations. Four independent simulations were conducted for each condition, each lasting up to 20 ns of simulation time. Same as in ANN-KMC approach, the diffusion coefficients were calculated using the Eq. (2).

3. Results

3.1 ANN model performance

As mentioned in the **Methods** section, the dataset used for training and testing the ANN model was obtained exclusively from the equiatomic FeNiCrCoCu HEA. This dataset was then randomly divided into training (including validation) and testing sets, with a ratio of 9 : 1. The performance of the ANN model was evaluated based on the MAE on the testing set. Figure 2 shows the relationship between MAE and the number of NN shells utilized for constructing the LACs. As expected, the MAE decreases as more NN shells are included. For instance, the ANN model considering only the first NN shell, which includes 19 atoms, yields a MAE of 47 meV,

while the ANN model including nine NN shells, encompassing 331 atoms, results in a MAE of 25 meV. It is worth noting that during the training of the ANN model incorporating more than six NN shells, overfitting can easily occur due to the increasing degree of freedom of the LACs. To remedy this issue, the number of neurons in the hidden layer was reduced from ten to five. Although including more NN shells does improve the accuracy to some extent, it also raises the computational cost. In addition, it may be challenging to generalize the approach for simulation systems of small sizes, such as those used for density functional theory (DFT) calculations. In fact, it can be seen in Fig. 2 that the MAE decreases slowly when three or more NN shells are included. To balance the accuracy and efficiency, the ANN model that incorporates the first three NN shells was found to deliver a satisfactory performance with a MAE of 32 meV and was chosen for this study. Figures 3(a – c) show the good performance of this ANN model on training set, testing set and the entire dataset that were obtained from the equiatomic FeNiCrCoCu HEA.

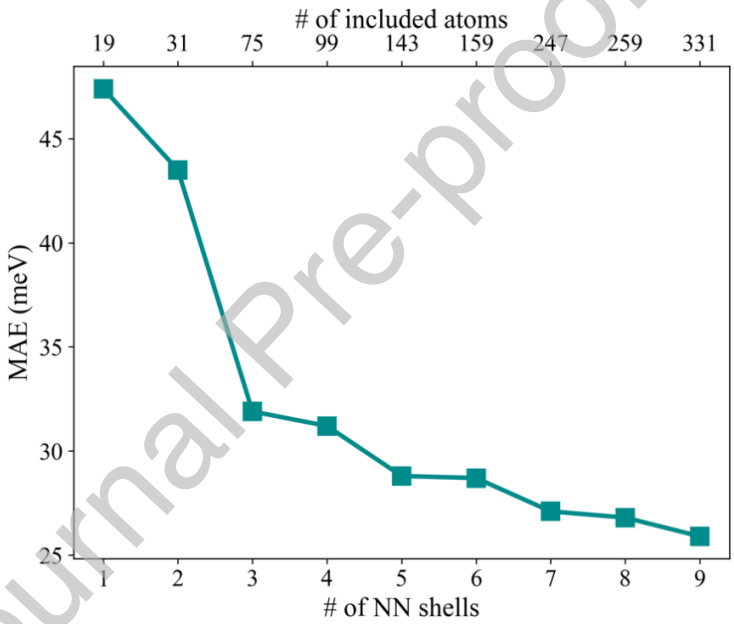


Fig. 2. Mean absolute error (MAE) of the ANN model on the testing set in relation to the number of NN shells included in LACs.

Upon finalizing the ANN model, its generalizability was evaluated from two perspectives: (1) its ability to generalize from equiatomic to non-equiatomic HEA compositions, and (2) extension of its predictability to its sub-systems. For the first assessment, the ANN model was applied on a dataset that includes ~2,500 migration barrier samples from randomly selected non-equiatomic FeNiCrCoCu HEAs (as mentioned in Section 2.1). Figure 3(d) presents the performance of the trained ANN on this supplemental dataset. It can be seen that the ANN model achieves a satisfactory accuracy with a MAE of 42 meV, despite this extra dataset being entirely independent to the dataset used to train the ANN model. This underscores the predictive

capability of the ANN model to extend beyond the equiatomic composition, demonstrating its potential for studying non-equiatomic HEA systems.

For the second evaluation, separate tests were conducted on 5 quaternary, 10 ternary, and 10 binary equiatomic sub-systems, each system containing 2,000 migration barrier samples (as mentioned in Section 2.1). Again, these migration barriers obtained from the sub-systems are invisible to the trained ANN model. Figure 4 illustrates the ANN prediction performance for the five quaternary alloys, FeNiCrCo, FeNiCoCu, FeCrCoCu, FeNiCrCu, NiCrCoCu, with respective MAEs of 34, 47, 39, 38, and 30 meV, and Pearson correlation coefficient (R) of 0.99 in all cases (note R = 1 means the perfect prediction). The overall MAE for the quaternary datasets is 37 meV, indicating an excellent level of prediction accuracy. Figure 5 presents the prediction performance for all ten ternary alloys, i.e., CrCoCu, FeCrCo, FeNiCo, FeNiCr, NiCrCo, NiCrCu, NiCoCu, FeNiCu, FeCrCu, and FeCoCu, with respective MAEs of 31, 39, 54, 33, 25, 25, 28, 57, 49, and 73 meV, and R > 0.95 for most cases. The overall MAE for the ternary datasets is 41 meV, which is still a commendable level of accuracy. Figure 6 displays the prediction performance for ten binary systems: NiCr, NiCu, FeCr, CoCu, FeNi, CrCo, CrCu, FeCo, FeCu, and NiCo, with respective MAEs of 20, 32, 50, 40, 43, 25, 29, 127, 127, and 40 meV, and R > 0.95 for all cases except for NiCo. The MAEs for FeCo and FeCu are a little bit large. For FeCu, the large MAE could be caused by the difference in lattice constant: the lattice constant for the equiatomic FeNiCrCoCu alloy used for training the ANN model is 3.55 Å, and most sub-systems have the values close to it, whereas that for FeCu is 3.59 Å. The reason for the large MAE in the FeCo is unclear. Nevertheless, the slightly large errors are still considered acceptable. The overall MAE for the binary datasets is 54 meV, which still represents a satisfactory level of accuracy, particularly when considering the substantial compositional disparity between binary and quinary systems. Based on the good performance on these sub-systems, it is reasonable to state that the trained ANN model can be effectively extended to such sub-systems.

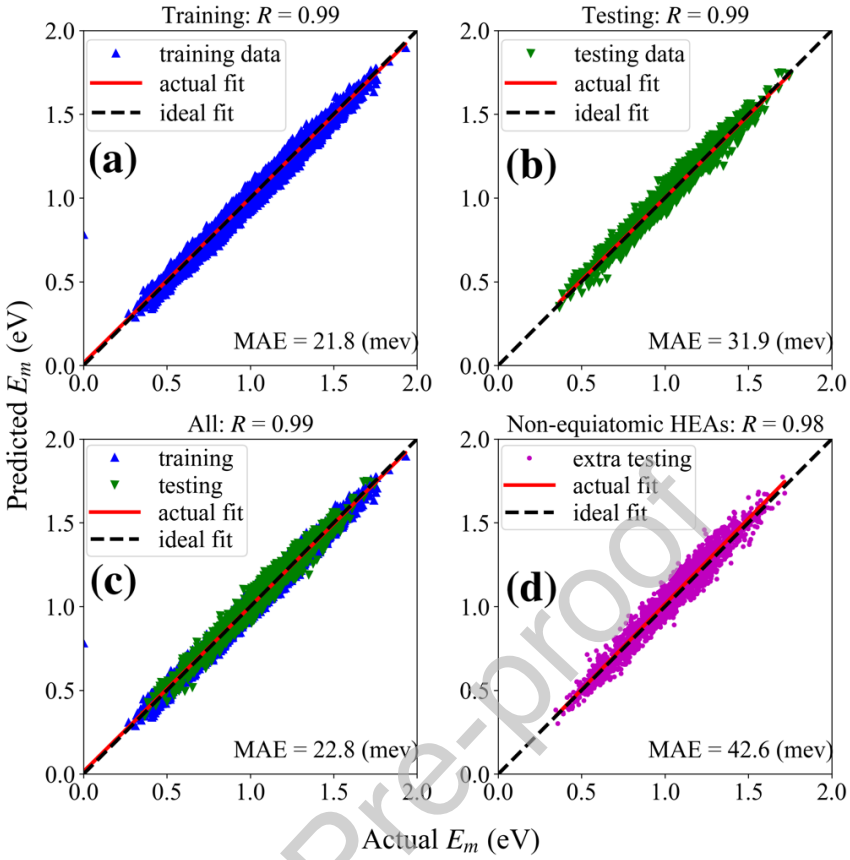


Fig. 3. Prediction performance of the ANN model incorporated with the first three NN shells on (a) training set, (b) testing set, (c) the entire dataset from the equiatomic FeNiCrCoCu HEA, and (d) independent supplemental dataset from non-equiatomic FeNiCrCoCu HEAs.

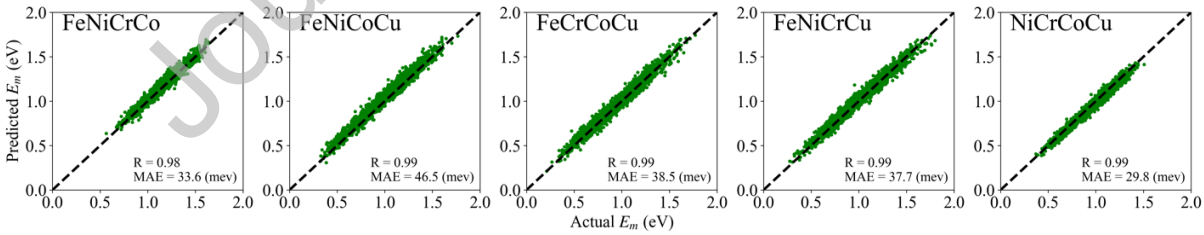


Fig. 4. Prediction performance of the ANN model on five independent quaternary systems.

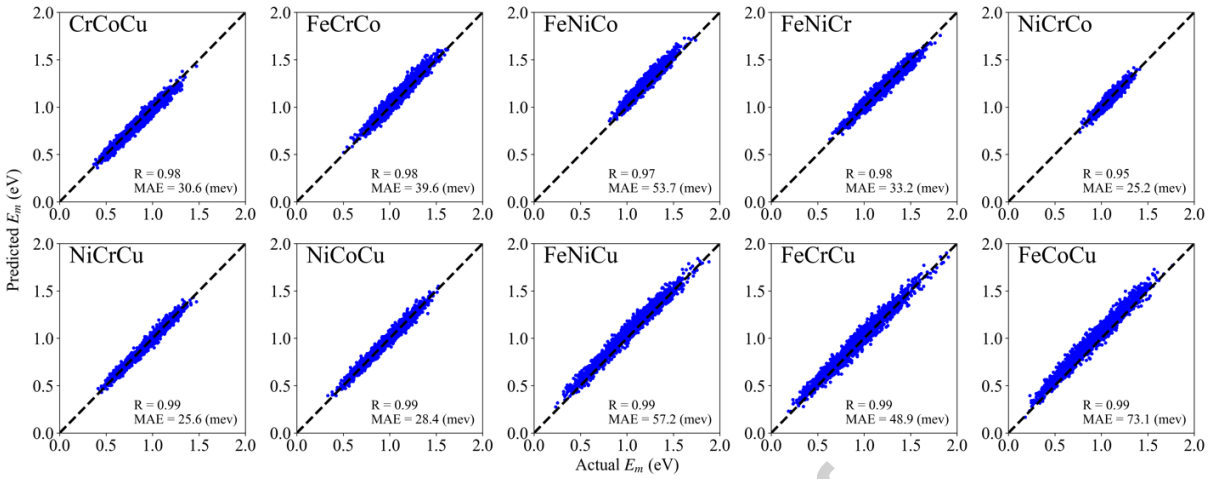


Fig. 5. Prediction performance of the ANN model on ten independent ternary systems.

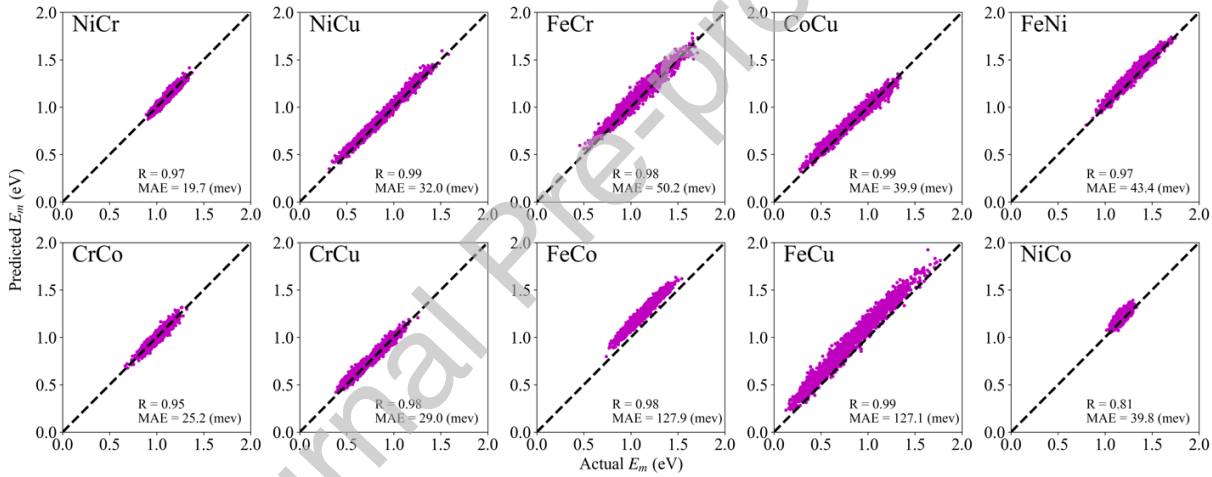


Fig. 6. Prediction performance of the ANN model on ten independent binary systems.

3.2 ANN-KMC and AA-KMC Results

After the well-trained ANN model is developed, it is coupled with KMC (ANN-KMC) to investigate vacancy-mediated diffusion in various HEAs. While many conventional KMC simulations commonly employ a constant jump attempt frequency (v_o = 10^{12} ~ 10^{13} s^{-1}) [47, 48], our recent study has demonstrated that v_o can be affected significantly by the temperature and composition in concentrated alloys [33]. To determine v_o for our ANN-KMC simulations of HEAs, the conventional KMC simulations were performed first with a fixed v_o = 10^{13} s^{-1} for each of the five pure components of the FeNiCrCoCu HEA, which are entirely independent of the ANN model since the migration barrier is constant for each pure component.

Figure 7(a) displays the diffusion coefficients of five pure components obtained from the conventional KMC using the fixed v_o in the Arrhenius plots. The results from independent MD simulations are also shown as benchmark data. Evidently, there are some discrepancies between KMC and MD results for some pure components, such as Fe and Cu, suggesting that the fixed v_o value is likely the reason for these discrepancies. By fitting the results of KMC to MD for these pure components, the v_o of each pure component can be approximated as a linear function with temperature, as shown in Fig. 7(b). The fitted functions of v_o for the five pure components are provided in Eqs. (3) – (7). The jump attempt frequency for an alloy, v_o^{alloy}, is then treated as the weighted sum of its components based on the alloy composition, as expressed in Eq. (8).

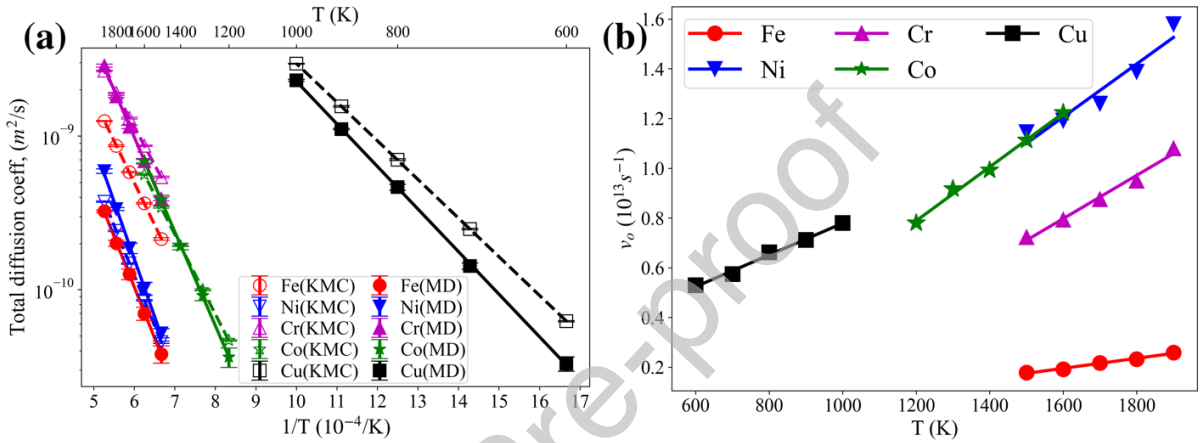


Fig. 7. (a) Diffusion coefficients for five pure components obtained from the conventional KMC simulations (open symbols and dashed lines) at different temperatures in the Arrhenius plots, using a fixed v_o value of 10^{13} s^{-1}. The independent MD results (filled symbols and solid lines) are shown for comparison. (b) The v_o derived from fitting the conventional KMC to MD results for each pure component at their respective temperature ranges.

$$v_o^{Fe}(T) = 2.01 \times 10^{-4}T - 0.126, \quad (3)$$

$$v_o^{Ni}(T) = 1.07 \times 10^{-3}T - 0.499, \quad (4)$$

$$v_o^{Cr}(T) = 8.71 \times 10^{-4}T - 0.596, \quad (5)$$

$$v_o^{Co}(T) = 1.08 \times 10^{-3}T - 0.511, \quad (6)$$

$$v_o^{Cu}(T) = 6.37 \times 10^{-4}T + 0.141, \quad (7)$$

$$v_o^{alloy}(T) = \sum_{X \in alloy} C_X v_o^X(T). \quad (8)$$

In Eq. (8), C_X is the concentration of the component X in the alloy. The unit is 10^{13} s^{-1} in all of the Eqs. (3) – (8). There are two possible strategies to implement the composition- and temperature-dependent v_o model for alloys in the ANN-KMC calculations. The first approach is to specify the corresponding v_o^X values for different moving atoms according to their species. For

example, v_o^{Fe} is applied if Fe is the moving atom, and so forth. The second approach employs the v_o^{alloy} as described in Eq. (8) for the entire alloy system. Both strategies yielded comparable results in terms of the total diffusion coefficients, which is the primary focus in this work. Consequently, the second approach (Eq. (8)) is adapted due to its simplicity in the KMC modeling.

Using the v_o^{alloy} from Eq. (8), the vacancy diffusivities in the equiatomic FeNiCrCoCu HEA, and the five quaternary equiatomic sub-systems were calculated from the ANN-KMC simulations from 1600 K to 2000 K, with an interval of 100 K. To validate the ANN-KMC model, diffusivities were also calculated by the independent MD simulations in these alloys. Figure 8 shows a complete comparison of the results between ANN-KMC and MD using Arrhenius treatments in the temperature range from 1600 K to 2000 K. The error bar for each diffusivity data indicates the standard deviation derived from six independent simulations in ANN-KMC or four independent simulations in MD. The activation energies for diffusion are extracted for both ANN-KMC and MD results, as indicated in each sub-figure. Overall, the ANN-KMC results agree very well with the independent MD results, in terms of both diffusivities and activation energies, across various alloy systems and temperatures. Slight discrepancies can be seen in two quaternary sub-systems: FeCrCoCu and NiCrCoCu. However, given the fact that the ANN model is trained using the dataset exclusively from the equiatomic five-component FeNiCrCoCu HEA, and the attempt jump frequency (v_o^{alloy}) in the KMC is a weighted sum of pure components (Eq. (8)), those small discrepancies are deemed as reasonable and acceptable. Experimentally, HEAs typically do not contain Cu so there are no experimental data to directly compare with our model HEA results. For quaternary HEAs, some experimental self-diffusivity data are available for the equiatomic FeNiCrCo HEA [49]. Since the EAM potential predicts a different melting temperature from the realistic HEA, here only the comparison of the activation energy is made. The self-diffusion activation energy of the FeNiCrCo HEA is about 280 kJ/mol in experiment, which is comparable to 295 kJ/mol obtained in this modeling work. Note to compare the self-diffusion activation energy, the mean vacancy formation energy of 1.62 eV or 156 kJ/mol is added to the vacancy diffusion activation energy shown in Fig. 8, because the vacancy in our simulations is pre-created. For individual components, their activation energies are also comparable between experiments and this modeling work, although the order among the components has some discrepancies. Similar observation was also found in Seoane et al.'s work [18]. Overall, the model HEA system used in this work captures many essential features of realistic HEAs such as single FCC phase, complex PEL, and variation in component diffusivities. Therefore, it is expected that the scientific conclusions obtained from this model HEA can be generalized to many realistic HEA systems.

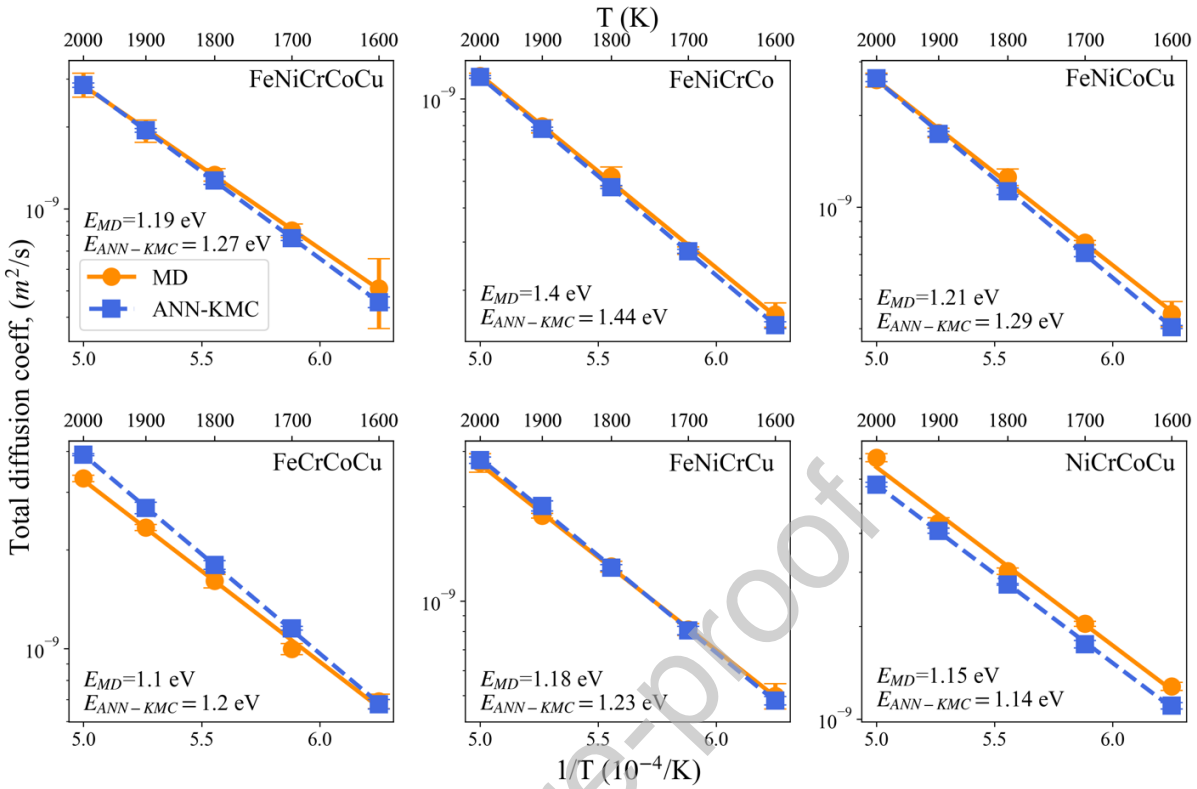


Fig. 8. Arrhenius plots of vacancy diffusion coefficients for the equiatomic FeNiCrCoCu HEA and its five equiatomic quaternary alloys obtained from ANN-KMC and MD simulations in the temperature range from 1600 to 2000 K.

One challenge in determining if diffusion is "sluggish" in HEAs or concentrated alloys arises from the absence of a universal criterion. For example, Daw and Chandross proposed four different criteria and each of them can be used to define sluggish diffusion [50]. Experimentally, the HEA diffusivities were typically compared with those of pure components or simpler alloys to determine if the diffusion is sluggish [51, 52]. However, such comparisons might not be very fair because a HEA is a very different material from its pure components or simpler alloys. As mentioned earlier, in this work a “sluggish diffusion” is determined if the diffusivity in an alloy is slower than the counterpart predicted by the average-atom (AA) model [23], similar to our previous work [18, 19]. This criterion is similar as comparing with the rule of mixture of the diffusivities of pure components [18], which is one of Daw et al.’s criteria [50]. The AA model is an EAM potential for a hypothetical single element. It predicts the average bulk properties nearly identical as the HEA while it does not have HEA’s compositional complexity. For example, Seoane et al. [18] reported that an AA material predicts similar lattice constant, elastic constants, vacancy formation and migration energies as the average values in an equiatomic FeNiCrCoCu HEA. Therefore, using the AA material as the reference can isolate the effect of HEA’s inherent chemical disorder and elucidate if such an effect leads to a sluggish diffusion. It should be noted

that such a comparison is only achievable through computer modeling because a hypothetical AA material does not exist experimentally. However, each HEA composition requires a distinct AA EAM potential. Therefore, it is impractical to develop numerous AA models for the vast non-equiatomic HEA compositional space. To tackle this challenge, another high-throughput method, termed as AA-KMC, is developed in this work to approximate the diffusion properties in AA materials. Based on the definition of the AA material [23], this approach sets the vacancy migration barrier in the AA-KMC modeling the same as the mean migration barrier of its corresponding HEA to substitute developing an AA EAM potential. By utilizing our well-trained ANN model (Section 3.1), the static migration barrier distribution and its mean in a HEA of any composition can be easily obtained, without the need of conducting computationally expensive NEB calculations. Subsequently, conventional KMC modeling can be conducted to calculate the diffusivities for AA materials, using the same jump attempt frequency v_o^{alloy} as described in Eq. (8).

To validate the AA-KMC approach, Figure 9 presents the diffusivities and activation energies obtained from the AA-KMC calculations and the independent MD counterparts using the AA EAM potential. The corresponding ANN-KMC and MD diffusivity data for the equiatomic FeNiCrCoCu HEA are also shown for comparison, which are the same as those shown in the first sub-figure of Fig. 8. It is evident that the AA-KMC method yields diffusivities nearly identical to the independent MD results based on the corresponding AA EAM potential, demonstrating the feasibility of the AA-KMC method. Consistent with our previous results that are solely based on MD simulations [18], the comparison between the HEA and AA material (regardless via MD or KMC) confirms the absence of vacancy-mediated sluggish diffusion in the equiatomic FeNiCrCoCu HEA, which is also consistent with experimental observations [53].

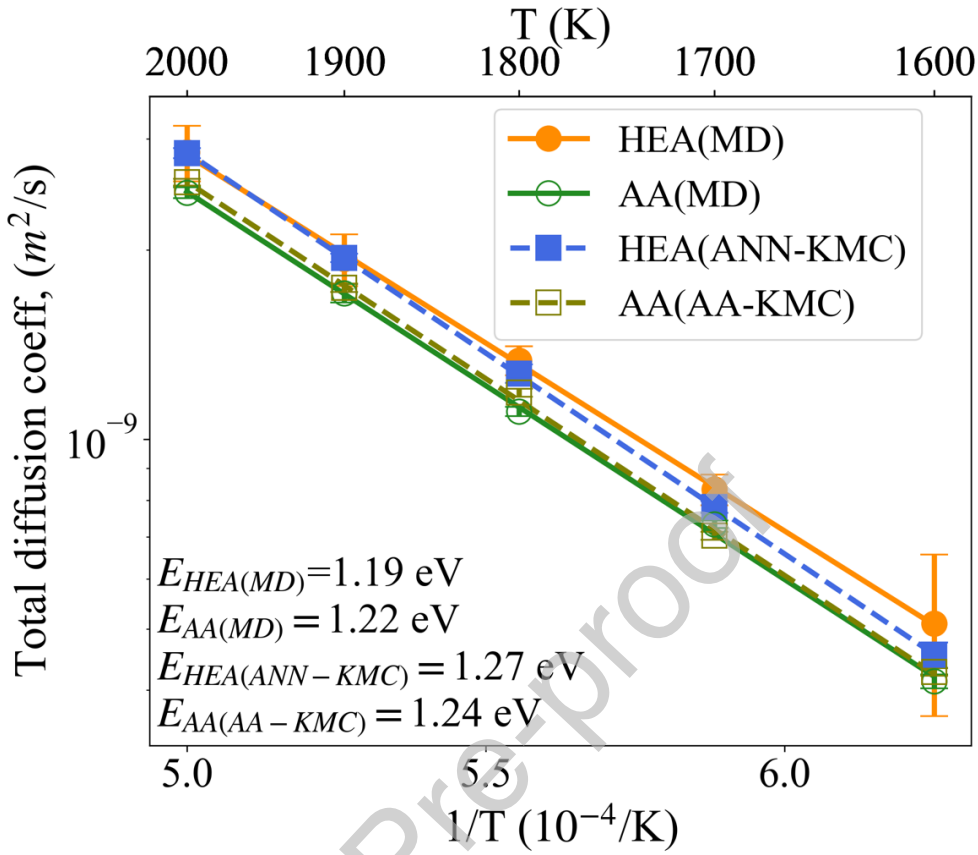


Fig. 9. Arrhenius plots of the vacancy-mediated diffusivities for the equiatomic FeNiCrCoCu HEA and its corresponding AA material using MD, ANN-KMC, and AA-KMC simulations at the temperatures ranging from 1600 K to 2000 K.

3.3 Percolation effects of fastest diffuser – Cu

Osetsky et al. [16, 21] argued that the faster diffuser in the binary Ni-Fe alloys, Fe, can cause a local trapping effect when its concentration is below the percolation threshold in an FCC crystal, 20 at.%. As a result, the total diffusivity reaches a minimum at this threshold. To investigate if such a percolation effect also exists in HEAs, a series of non-equiatomic FeNiCrCoCu alloys were prepared in which the concentration of the fastest-diffusing species (i.e., Cu) was varied from x = 0 – 30 at.% with an interval of 5 at.% and the other four components were in equiatomic ratios, denoted as $(FeNiCrCo)_{100-x}Cu_x$. An additional x = 1 at.% of Cu composition is also studied. Figure 10(a) shows the total diffusion coefficients of $(FeNiCrCo)_{100-x}Cu_x$ and its corresponding AA materials, as determined by MD, ANN-KMC, and AA-KMC simulations at 1800 K. Here the AA EAM potentials for those $(FeNiCrCo)_{100-x}Cu_x$ HEAs were developed based on the method described in [23]. Again, there is good agreement between ANN-KMC and MD results for HEAs over the entire Cu concentrations, as well as between AA-KMC and MD results for AA materials. Notably, both HEA and AA materials

exhibit a monotonic and nearly linear increase in diffusivity with the increasing Cu concentration, indicating that the percolation threshold (x = 20 at.%) does not guarantee the slowest diffusivity in the $(FeNiCrCo)_{100-x}Cu_x$. This is consistent with the observation found in the ternary CoCrNi alloy [22]. However, when compared with AA materials, HEAs show marginally slower diffusivities at 1800 K when the Cu concentration is ≤ 15 at.%, indicating the presence of sluggish diffusion according to the criterion discussed in the previous section. Figure 10(b) presents the total diffusivity for the same systems obtained from ANN-KMC and its corresponding AA-KMC at 900 K. Since MD cannot provide reliable diffusivity calculations at this low temperature, no MD results were obtained for HEAs or AA materials. At 900 K, the slower diffusivities in HEAs compared to AA materials become more evident with the decreasing Cu concentration, suggesting that the local trapping effects by Cu are more pronounced at lower temperatures. The results also demonstrate the necessity of using low-temperature simulations to magnify the presence of sluggish diffusion in this HEA system, consistent with an experimental study in which the sluggish diffusion in an FeNiCrCoMn HEA grain boundary is more evident at low temperatures [54]. These observations suggest a competitive interplay between the local trapping effect and the enhanced total diffusivity due to the incorporation of the fastest diffuser – Cu. However, unlike in the binary Ni-Fe system [16, 21], the competition does not lead to a diffusivity minimum, probably because Cu diffusivity is a few orders of magnitude faster than other components (Fig. 7(a)) so that the diffusion enhancement overpowers the local trapping effect.

Figures 11 and 12 illustrate several key diffusion factors for studying the underlying diffusion mechanisms in $(FeNiCrCo)_{100-x}Cu_x$. Figure 11(a) showcases the migration barriers of AA materials at different Cu concentrations using NEB calculations based on the EAM potentials, alongside the mean migration barriers of HEAs predicted by the ANN model. To obtain the latter, a vacancy was sequentially created at each lattice site in a HEA and a random migration path was chosen from the twelve nearest neighboring atoms of the vacancy. The migration barrier was then predicted by the ANN model based on the LAC around the migration path. Repeating this process, 4000 migration barriers in each HEA were obtained and the mean value is reported in Fig. 11(a). The figure shows that the average migration barrier of each non-equiatomic HEA closely resembles the migration barrier of the corresponding AA material, further validating the assumption of the migration barrier in the AA-KMC method. Also, a noticeable decrease in the migration barrier for both HEA and AA is observed as the Cu concentration increases, consistent with the inverse trend of the diffusivity in Fig. 10. Figure 11(b) demonstrates the absolute jump fraction by each species at 1800 K. When Cu concentration is 1 at.%, its jump fraction is already large (~15%) but does not dominate. When Cu concentration is 5 at.%, its jump fraction becomes the largest among other components, nearly 50%. As the Cu concentration increases further, it increasingly dominates the vacancy jumps, such as over 80% in $(FeNiCrCo)_{70}Cu_{30}$. Figure 11(c) shows the normalized jump ratios of the HEA components with respect to their compositions at different Cu concentrations at 1800 K. Clearly, the fastest diffusing species (Cu) "over-jumps" significantly with respect to its composition, and the over-jump ratio increases with the decreasing Cu concentration. Again, the

result demonstrates that the vacancy prefers to exchange with the fastest-diffusing species. If many of such preferential jumps are back-and-forth ones, they can lead to a local trapping effect.

The correlation factor for diffusion ($0 \leq f_c \leq 1$) is a useful parameter to characterize if the defect jumps lead to a local trapping effect (small f_c) or a long-range diffusion (large f_c). Here the correlation factor for each species in a HEA is calculated using the formula [22],

$$f_c^X = \frac{ASD_X}{N_X \Delta^2}, \quad (9)$$

where ASD_X and N_X are the ASD and the number of jumps of the species X in a HEA, respectively, and Δ is the vacancy jump distance (the 1st nearest neighbor distance). The same equation is also used to calculate the correlation factors in the entire HEAs (f_c^{HEA}) and AA materials (f_c^{AA}) during the ANN-KMC and AA-KMC simulations. Figure 12 shows the calculated correlation factors as a function of Cu concentration at 1800 K and 900 K, respectively. At both temperatures, the correlation factor for the single-component AA material is $f_c^{AA} \approx 0.78$ at all Cu concentrations, which is in excellent agreement with the theoretical value of 0.78146 for a 3D vacancy diffusion in a pure FCC crystal [55]. On the other hand, f_c^{HEA} is significantly lower than f_c^{AA}, suggesting that the vacancy jumps in HEAs can result in a trapping effect due to their rough PEL. For the five components in the HEAs, their correlation factors reversely correspond to their pure-component diffusivities (Fig. 7(a)), namely, a component with a lower diffusivity yields a higher correlation factor. Furthermore, the addition of Cu leads to a slight increase in the correlation factor for the other four components. At all Cu concentrations, the Cu correlation factor is very low, implying its strongest ability to interchange with the vacancy and many of them are back and forth jumps, thus potentially creating local traps. In terms of temperature effect, the Cu correlation factor at 900 K (Fig. 12(b)) is much lower than that at 1800 K (Fig. 12(a)). This result indicates that the local trapping effect by Cu is stronger at low temperatures, consistent with the more pronounced sluggish diffusion at 900 K than at 1800 K (Fig. 10). As previously discussed, due to the significantly higher diffusivity of Cu, the effect of diffusion enhancement likely predominates over localized trapping effects. The overall correlation factor of the HEA has its lowest value at the equiatomic composition (i.e., 20 at.% Cu), which could be attributed to the maximum compositional complexity at this composition. However, it is observed that such compositional complexity does not directly impact the total diffusivity in the equiatomic HEA in comparison with its AA counterpart as they yield similar diffusivities at this composition.

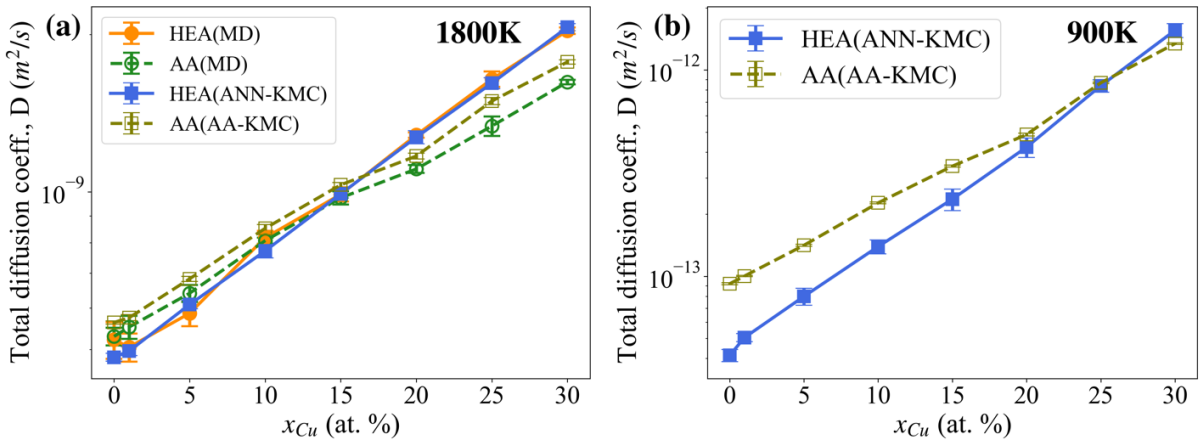


Fig. 10. Total diffusivities of $(FeNiCrCo)_{100-x}Cu_x$ and AA materials at different Cu concentrations, (a) determined by MD, ANN-KMC, and AA-KMC simulations at 1800 K, and (b) determined by ANN-KMC and AA-KMC simulations at 900 K.

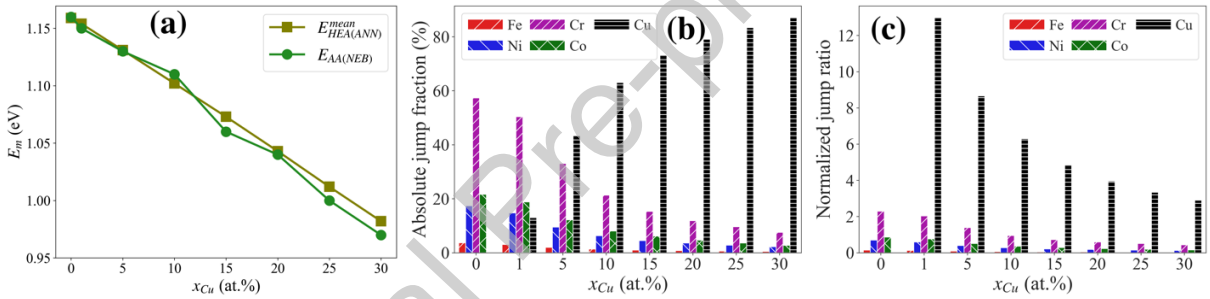


Fig. 11. Diffusion governing factors in the $(FeNiCrCo)_{100-x}Cu_x$ HEAs and AA materials: (a) migration barriers of AA materials (by NEB) and mean migration barriers of HEAs (by ANN), (b) absolute jump fractions of individual components at 1800 K, and (c) normalized jump ratios of individual components with respect to their compositions at 1800 K.

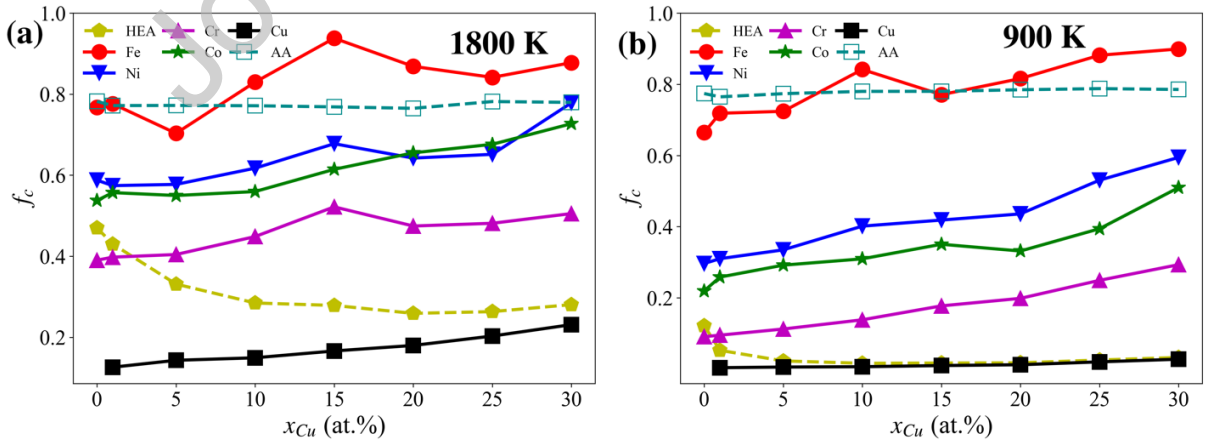


Fig. 12. Total and elemental correlation factors at (a) 1800 K and (b) 900 K.

3.4 High-throughput investigation of the vast compositional space of HEAs

The results presented in the previous sections have demonstrated the reliability of the ANN-KMC and AA-KMC methods, and revealed the existence of sluggish diffusion in non-equiatomic HEAs as Cu concentration varies, especially at a low temperature of 900 K. Therefore, the two methods are used to further explore the more diverse compositional space of HEAs and search for those exhibiting greater degrees of sluggish diffusion at 900 K. To achieve this, the concentrations of all five species were randomly varied with a 5 at.% interval to create 1,500 unique non-equiatomic HEAs (e.g., $Fe_{25}Ni_{15}Cr_{50}Co_5Cu_5$), each of these HEAs has the same system size as the equiatomic HEA. Taking advantage of the well-trained ANN model, the static migration barrier distribution for each HEA composition is obtained using the same way as that for obtaining the mean barriers in Fig. 11(a). From the distribution, its mean barrier ($\mu_{E_m}^{static}$) and standard deviation ($\sigma_{E_m}^{static}$) can be obtained, with the former ($\mu_{E_m}^{static}$) employed for AA-KMC simulations. In addition, during an ANN-KMC simulation, the executed vacancy hops and their corresponding migration barriers were recorded, resulting in a dynamic barrier distribution with its respective mean ($\mu_{E_m}^{dynamic}$) and standard deviation ($\sigma_{E_m}^{dynamic}$).

Figures 13(a) and (b) show the diffusivities for those 1,500 HEA compositions at 900 K calculated from ANN-KMC simulations, in relation to $\mu_{E_m}^{static}$ and $\sigma_{E_m}^{static}$, as well as $\mu_{E_m}^{dynamic}$ and $\sigma_{E_m}^{dynamic}$, respectively. Here two types of sluggish diffusion are categorized: "strong" sluggish diffusion if the HEA diffusivity is more than one order of magnitude slower than its AA counterpart (i.e., $D_{HEA} < \frac{1}{10} D_{AA}$), and "mild" sluggish diffusion if the HEA diffusivity is slower than at least half of its AA counterpart (i.e., $\frac{1}{10} D_{AA} < D_{HEA} < \frac{1}{2} D_{AA}$). Note here if the HEA diffusivity is only slightly lower than its AA counterpart (i.e., $\frac{1}{2} D_{AA} < D_{HEA} < D_{AA}$), it is not considered as sluggish because the AA-KMC method could exhibit some uncertainties in calculating the AA diffusivities. This treatment increases the comparison margin and reduces the uncertainties in identifying the sluggish compositions. It should be noted that the above definitions of strong and mild sluggish diffusion are somewhat arbitrary. However, the main conclusions of this study remain unchanged if these thresholds are modified slightly, e.g., defining strong sluggish as $D_{HEA} < \frac{1}{5} D_{AA}$. While many HEAs show mild sluggish diffusion, six HEAs display strong sluggish diffusion: $Fe_{60}Ni_5Cr_{20}Co_{10}Cu_5$, $Fe_{60}Ni_{20}Cr_5Co_{10}Cu_5$, $Fe_{60}Ni_5Cr_5Co_{10}Cu_{20}$, $Fe_{60}Ni_{25}Cr_5Co_5Cu_5$, $Fe_{55}Ni_5Cr_{10}Co_5Cu_{25}$, and $Fe_{45}Ni_{20}Cr_{10}Co_{20}Cu_5$, as marked by red circles in Fig. 13. In Fig. 7(a), the diffusivity order of pure components is: Fe < Ni < Co ≈ Cr < Cu. Therefore, these strong sluggish HEAs generally contain a high concentration of slowest-diffusing species, Fe. In terms of the sum of the two slowest species (Fe + Ni), it exceeds 60% in all cases. Figure 13(a) also reveals that these sluggish HEAs (both strong and mild) possess a relatively higher $\mu_{E_m}^{static}$. However, the $\sigma_{E_m}^{static}$ does not exhibit a clear influence on diffusivity, as the sluggish HEAs have both large and moderate $\sigma_{E_m}^{static}$. This observation suggests that the variations in static migration barriers is not critical in shaping the diffusion

behavior, in agreement with the previous study [25]. Conversely, Figure 13(b) shows a different trend, where a sluggish HEA generally has a larger $\sigma_{E_m}^{dynamic}$ while its $\mu_{E_m}^{dynamic}$ spans from low to moderate. In particular for the 4 out of 6 strong sluggish HEAs, their $\sigma_{E_m}^{dynamic}$ are extremely large. The higher $\sigma_{E_m}^{dynamic}$ indicates a greater variation in the executed migration barriers in a HEA. This variation could result in regions with both low and high migration barriers that act as traps for diffusing atoms. In this scenario, either vacancy hops via high-energy barrier sites could occur, or the likelihood of localized back-and-forth hops via low-energy barriers may increase, ultimately leading to a reduction in the overall diffusivity.

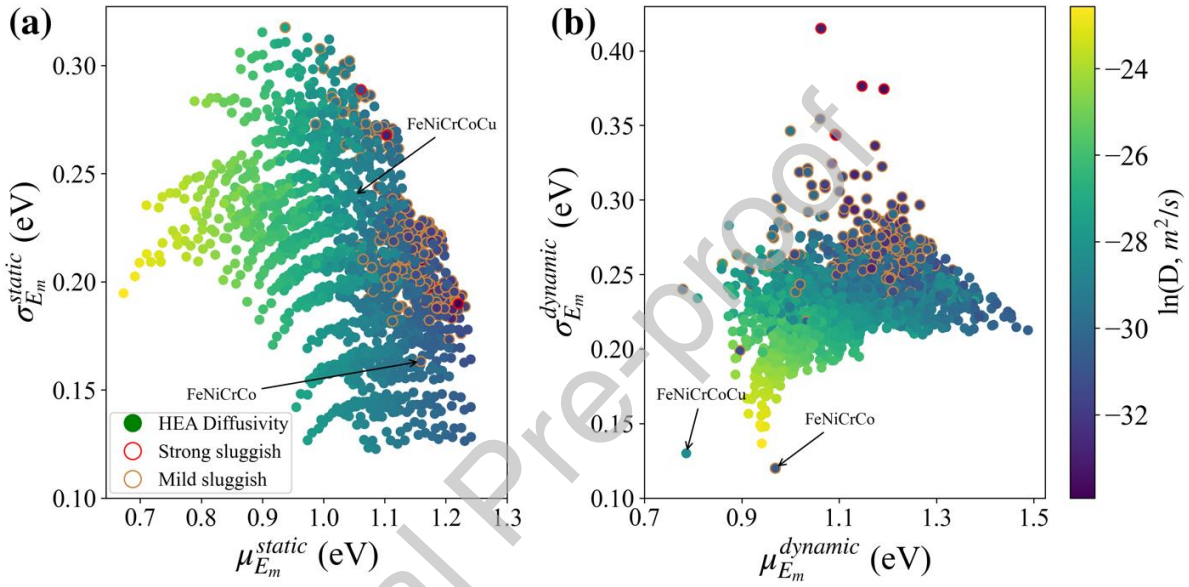


Fig. 13. Diffusivities of 1,500 unique non-equiatomic HEA compositions calculated by ANN-KMC at 900 K in relation to (a) $\mu_{E_m}^{static}$ and $\sigma_{E_m}^{static}$, and (b) $\mu_{E_m}^{dynamic}$ and $\sigma_{E_m}^{dynamic}$. The color of each filled circle indicates the magnitude of diffusivity. HEAs exhibiting sluggish diffusion with respect to their AA counterparts are classified as strong ($D_{HEA} < \frac{1}{10} D_{AA}$, red circle line) or mild ($\frac{1}{10} D_{AA} < D_{HEA} < \frac{1}{2} D_{AA}$, orange circle line) degree, respectively.

In prolonged diffusion scenarios, short-range orders (SROs) may develop and impact the diffusion behavior [31]. Evidence from both our recent work and other studies have demonstrated that SROs can introduce extra trapping effects and further slower the diffusion [33, 56]. However, in this study, the utilized EAM potential, characterized by its low heat of mixing, ensures that the tendency of chemical ordering remains minimal even after long-time KMC simulations. To provide a quantitative assessment of the possible SROs development, the Warren-Cowley parameter [57] for each pair of elements is computed, as defined by

$$\alpha_k^{ij} = 1 - \frac{\langle N_k^{ij} \rangle}{c_j N_k}, \quad (10)$$

where i and j represent i^{th}-type and j^{th}-type of elements (e.g., Fe, Cr), k represents the k^{th} nearest-neighbor shell, $\langle N_k^{ij} \rangle$ represents the average number of j^{th}-type atoms surrounding i^{th}-type atoms at its k^{th}-nearest-neighbor shell, N_k is the total number of atoms in the k^{th} nearest-neighbor shell, and c_j is the concentration of j^{th}-type atoms in the entire system. α_k^{ij} should approach zero in a random solid-solution alloy, while a negative value means favorable *i-j* pair formation and vice versa. As an illustrative example, the SRO values at the 1st nearest-neighbor shell for the $Fe_{55}Ni_5Cr_{10}Co_5Cu_{25}$ before and after 300,000 KMC steps are shown in Figs. 14(a) and (b), respectively, which exhibits strong sluggish diffusion in Fig. 13. Clearly, the SRO values for most of elemental pairs in the initial structure (Fig. 14(a)) are close to zero, indicating the initial atomic configuration is random. The SRO value for the Co-Co pair deviates from zero slightly, possibly due to insufficient statistics from its low concentration (5 at.%). After the KMC simulation (Fig. 14(b)), the SRO values only change slightly for these pairs, confirming that an extended KMC simulation does not introduce significant SROs. This finding suggests that the trapping effect discussed here is not significantly impacted by SROs.

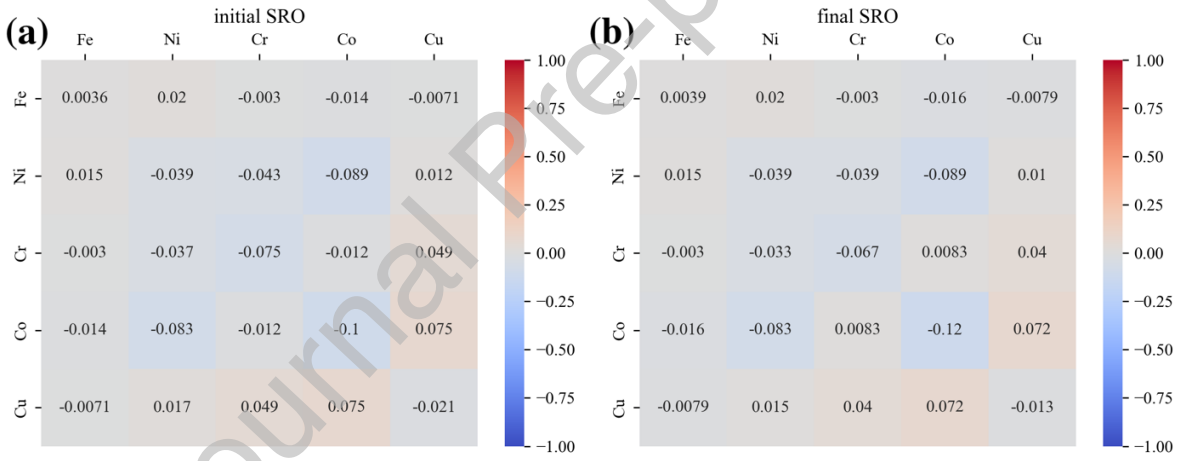


Fig. 14. Warren-Cowley SRO parameters of all element pairs in the $Fe_{55}Ni_5Cr_{10}Co_5Cu_{25}$ alloy for the (a) initial structure, and (b) final structure after 300,000 KMC steps at 900 K. Note this alloy shows a strong sluggish diffusion in Fig. 13.

4. Discussion

Molecular dynamics (MD) offers a wealth of advantages for studying diffusion properties in alloys, granting atomic-scale insights into the complex diffusion mechanisms in HEAs. However, it is computationally expensive for exploring the extensive compositional space in HEAs. Moreover, MD simulations are limited to high temperatures and short timescales, which pose challenges for exploring the complex potential energy landscape (PEL) in HEAs. Theoretical models, such as random barrier/trap (RB/RT) model or continuous-time random

walk model [58, 59], are unable to accurately describe the complex PEL in HEAs. On-the-fly off-lattice KMC method [16] is able to explore the complex PEL in concentrated alloys, but it is computationally expensive because the migration barriers need to be dynamically calculated at every step. In this regard, the coupling of machine learning (ML) with KMC as used in this work presents a promising solution for studying the complex diffusion problems in HEAs, in which the well-trained ANN-based ML model is able to capture the non-linear dependency of PEL on local atomic configurations (LACs). As such, there has been growing interest in employing ML for similar purposes in concentrated alloys [60-63]. While some ML-based studies have achieved promising accuracies in predicting the vacancy migration barriers for a specific HEA composition, generalizing these ML models to other HEA compositions or their subsystems remains challenging. This limitation leads to a dilemma where one must either collect extensive datasets encompassing various compositions or build separate ML models for each composition. The present work shows that it is possible to use a small dataset (e.g., equiatomic composition) as the training data and use the trained ANN model to predict other compositions (e.g., non-equiatomic compositions and sub-systems). The performance of a ML model is highly dependent on the quality of dataset and the choice of descriptors. In this study, the ANN model performance was systematically assessed with respect to the number of nearest neighboring (NN) shells used for constructing LACs, as shown in Fig. 2. While the ANN model presented here is built using LACs surrounding the vacancy-migration atom pair up to the third NN shells, it should be noted that our ANN model can still achieve an acceptable MAE of 47 meV and 42 meV when considering only the first NN atoms or up to the second NN atoms, respectively. Therefore, this approach could be extended to other smaller-size systems such as those for density functional theory (DFT) calculations, which typically have less than two hundred atoms. However, more work is needed in the future to determine if the advantages (e.g., small dataset and fewer NN shells for training) of our ANN model can be applied to other HEA systems. In particular, this HEA model system predicts the stable FCC phase for a wide range of random solid-solution compositions. In some aluminum (Al) containing HEAs, multiple phases can coexist [35]. Applying the ANN model developed in this work for such complex systems would be challenging. Nevertheless, it is expected that the LAC-ANN model developed in this work can be extended to some other HEA systems with DFT-calculated migration barriers. If a good DFT-ANN-KMC model can be developed, the development of EAM potentials is no longer necessary and the migration barriers may be more accurate than those obtained from interatomic potentials. In turn, the calculated diffusivities could be more comparable with experiments.

To test how the calculated diffusivity is sensitive to the accuracy of the ANN model, three "artificial" ANN models are constructed by introducing Gaussian noises with a mean of zero and varying standard deviations (0.1, 0.2, and 0.3 eV). This treatment widens the deviation in ANN predictions while keeping the mean migration barrier nearly unchanged, as indicated by the red linear regression lines in Figure 15(a-d). Figure 15(b-d) contrasts the performance of the three artificial ANN models with the actual model (Fig. 15(a)), as the MAE value increases from the original 31.9 meV to 85.7, 163.3, and 244.3 meV, respectively. These modified ANNs are coupled with KMC to calculate the diffusivities of the equiatomic FeNiCrCoCu at different temperatures, and the results are compared with the original model, as shown in Fig. 15(e-h).

Notably, as seen in Fig. 15(f), even with a relatively large MAE of 85.7 meV (artificial-1), the ANN-KMC model still agrees reasonably well with the independent MD results. It should be noted that this artificial MAE is much larger than all of those in the actual quinary and quaternary systems (Figs. 3-4), which are used for diffusivity calculations. Although it is close to those in a few ternary and binary compositions (Figs. 5-6), they are not used for any diffusivity calculations in this work. Therefore, the small MAEs in the ANN models for the quinary and quaternary systems should not significantly affect the accuracy of predicted diffusivities. However, if amplified errors are introduced (artificial-2 and artificial-3), the predicted diffusivities notably deviate from the MD data even though the mean migration barrier is similar. This sensitivity analysis demonstrates the risk of using the mean migration barrier in the conventional KMC simulations for HEAs while ignoring its variance, which is a common practice in literature but could yield inaccurate results.

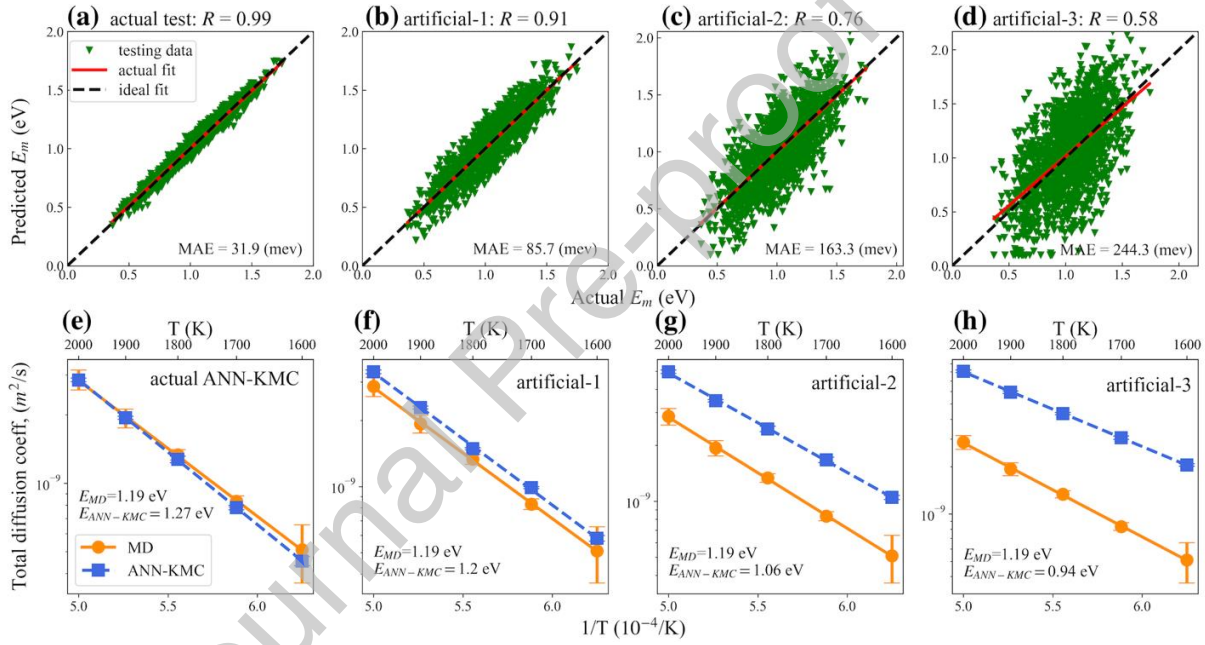


Fig. 15. (a) - (d) Performance comparison of the actual ANN model with three "artificial" ANN models for the equiatomic FeNiCrCoCu, constructed with Gaussian noises with a mean value of zero and varying standard deviations (0.1, 0.2, and 0.3 eV). (e) - (h) ANN-KMC calculated diffusivities using the actual and three artificial ANN models at different temperatures. The same set of MD data is shown in each sub-figure as a reference.

In this work, the determination of sluggish diffusion is based on the comparison between HEAs and their corresponding AA materials. The comparison aims to isolate the effect of compositional complexity in HEAs and understand how it influences the diffusion behavior, similar as in our previous studies [18, 19]. Building upon prior studies [18, 22] that suggest neither equiatomic compositions nor percolation thresholds guarantee slowest or sluggish diffusion in HEAs, our results also confirm that the equiatomic FeNiCrCoCu HEA exhibits no

sign of sluggishness when compared to its AA counterpart, as demonstrated in Fig. 9. While the percolation threshold of the fastest diffuser does not yield the sluggish diffusion in the $(FeNiCrCo)_{100-x}Cu_x$ HEAs, they can be deemed as sluggish when Cu concentration is below the percolation threshold (< 20 at.% Cu). The sluggish effect becomes more evident at low temperatures (e.g., 900 K). Although the sluggish diffusion is observed as compared to AA counterparts, the overall diffusivity in the $(FeNiCrCo)_{100-x}Cu_x$ HEAs increases with the increasing Cu concentration (Fig. 10(b)), which seems contradictory. This could be due to the two competing effects by Cu: trapping and enhanced diffusion. When Cu is not percolated, the back-and-forth exchange (through low migration barriers) between Cu and vacancy can cause a local trapping effect, leading to a sluggish diffusion (in comparison with the AA counterparts). On the other hand, since Cu diffusion is so fast, it can overpower the trapping effect and enhance the total diffusivity (in comparison with 0 at.% Cu composition) as its concentration increases.

This work shows that even though the diffusion in the equiatomic FeNiCrCoCu HEA is not sluggish, the diffusion in non-equiatomic compositions could be sluggish. Therefore, the high-throughput modeling approach developed in this work is useful for exploring the large composition space and study how the complex PEL influences the diffusion kinetics. For example, Figure 13(a) shows that these sluggish HEAs tend to have high average static migration barriers ($\mu_{E_m}^{static}$), while their variances ($\sigma_{E_m}^{static}$) can be either high or moderate; Figure 13(b) shows that these sluggish HEAs tend to have large variances ($\sigma_{E_m}^{dynamic}$) in the accepted migration barriers, while their means ($\mu_{E_m}^{dynamic}$) span from low to moderate. The discrepancies between the static and dynamic (accepted) barrier distributions indicate that sluggish diffusion has a complex relationship with HEA composition and thus the PEL. A seemingly rough static PEL (i.e., having large $\sigma_{E_m}^{static}$), which is based on all atomic sites in the system regardless they are visited or not, does not necessarily lead to a sluggish diffusion. This is because a vacancy could find some alternative paths to "bypass" the trapping or high-barrier sites so that those sites may never be visited during the diffusion. On the other hand, if the system consists of effective local trapping sites that a vacancy cannot escape easily, such as some low-barrier sites enclosed by high-barrier sites, the sluggish diffusion effect could be strong. In such case, the dynamic PEL should have large $\sigma_{E_m}^{dynamic}$.

In Section 3.4, six compositions exhibiting strong sluggish diffusion (at least 10 times slower than the AA counterpart) have been identified. The general trend is that they contain high concentrations of slowest diffusing species (Fe and Ni) and low concentrations of fastest diffusing species (Cu). To check if the mild sluggish diffusion ($\frac{1}{10}D_{AA} < D_{HEA} < \frac{1}{2}D_{AA}$) also follows the same trend, Figure 16 shows a 3-D plot of the diffusion behavior with respect to Cu, Ni, and Fe concentrations. Clearly, the majority of the sluggish compositions contain less than 10 at.%Cu but more than 40 at.%Fe. It should be noted that this composition analysis is consistent with Fig. 13. When the system contains more slow diffusing species such as Fe and Ni, the mean static barrier ($\mu_{E_m}^{static}$) will be high, consistent with Fig. 13(a). On the other hand, if the system contains a small amount of fast diffusing species such as Cu, the fast diffuser may form non-

percolated local traps and thus increase the variances of the accepted barriers ($\sigma_{E_m}^{dynamic}$), consistent with Fig. 13(b). The results from this work could provide some useful insight for experimental alloy design. To design a HEA with sluggish diffusion, our suggestion is that the alloy should contain high-concentration slow diffusing species and low-concentration fast diffusing species. However, the fast diffuser should not be as fast as Cu to avoid the predominance of its diffusion enhancement over its local trapping effect. If the fast diffuser can form some non-percolated local precipitates due to its low solubility, the local trapping effect and thus the sluggish diffusion could be even stronger. Of course, these suggested alloy design guidelines need to be experimentally validated by the research community in the future.

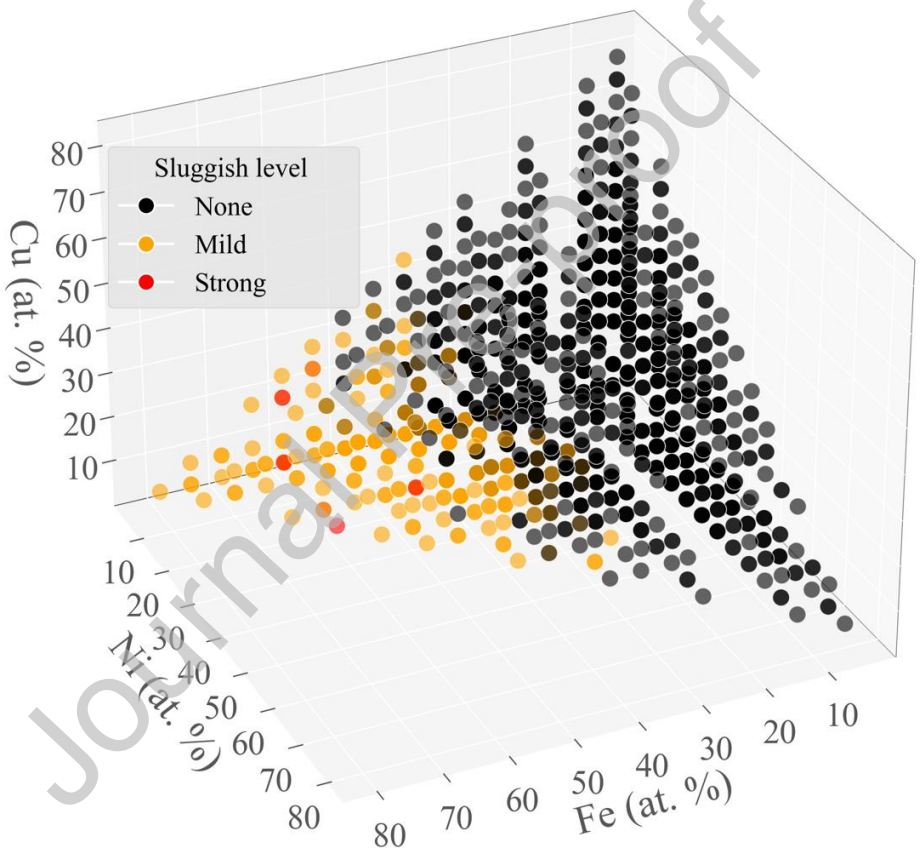


Fig. 16. Relation between the diffusion behavior in 1500 non-equiatomic HEA compositions with faster diffuser Cu, and slower diffuser (Ni, Fe) concentrations. Here strong sluggish ($D_{HEA} < \frac{1}{10} D_{AA}$), mild sluggish ($\frac{1}{10} D_{AA} < D_{HEA} < \frac{1}{2} D_{AA}$), and non-sluggish ($D_{HEA} > \frac{1}{2} D_{AA}$) compositions are represented by red, orange, and black filled circles, respectively.

5. Conclusions

- *A high-throughput machine learning informed KMC modeling framework has been developed to study the possible sluggish vacancy diffusion in a wide range of FeNiCrCoCu HEA compositions*

In this work, an innovative artificial neural network (ANN) based machine learning model has been developed to predict local-atomic-configuration (LAC) dependent vacancy migration barriers in a quinary FeNiCrCoCu HEA system. Despite the ANN model being trained solely on the data from the equiatomic HEA, it can predict barriers in many non-equiatomic compositions and its subsystems with a reasonable accuracy. By integrating this ANN model as an on-the-fly barrier calculator for KMC modeling, termed as ANN-KMC in this work, it can calculate vacancy-mediated diffusivities in alignment with independent MD simulations across a range of HEA compositions, but with a much higher efficiency. In addition, a composition- and temperature-dependent jump attempt frequency model has also been developed for KMC modeling of the HEA system. To evaluate the possible existence of sluggish diffusion, the HEA diffusivities are compared to its average atom (AA) counterparts, which describes a hypothetical element having nearly identical bulk properties as the HEA but without HEA's compositional complexity. To avoid the development of a specific AA EAM potential for every HEA composition, a high-throughput AA-KMC model has also been developed, which can predict the vacancy diffusivities comparable with those from independent MD simulations using AA EAM potentials. Using the high-throughput ANN-KMC and AA-KMC methods, the diffusivities in many non-equiatomic HEA compositions have been calculated at a wide range of temperatures and long timescales to search for the possible sluggish diffusion. It is expected that the high-throughput approaches can be applied to DFT-calculated barriers, enabling the prediction of more accurate diffusivities without the need of developing interatomic potentials for different HEAs.

- *The fastest diffuser (Cu) can lead to both sluggish diffusion due to the local trapping effect and diffusion enhancement due to its fast diffusivity when its concentration is below the percolation threshold*

To study the percolation effects of the fastest-diffuser (Cu) on the overall diffusion behavior in the FeNiCrCoCu HEA system, its concentration was systematically varied from 0 to 30 at.%. It is found that the equiatomic composition or the percolation threshold (20% Cu) in this HEA system does not result in either a minimum diffusivity or sluggish diffusion, which is in stark contrast to the binary Ni-Fe alloy system [16, 21]. Instead, the total diffusivity in the HEA system increases nearly linearly with the increasing Cu concentration. On the other hand, the HEA diffusion is slower than its AA counterpart when the Cu concentration is below its percolation threshold. The seemingly contradictory results suggest a complex interplay between the local trapping effects and enhanced global diffusion due to the incorporating of the ultrafast diffusing element (Cu). The former results in a sluggish diffusion with respect to the AA material when Cu is not percolated and can exchange with vacancy back and forth, while the latter overpowers the local trapping effect and leads to the enhancement of overall diffusivity in comparison to the 0% Cu composition.

- *This work demonstrates that although sluggish diffusion may not exist at the equiatomic composition, it could exist at non-equiatomic compositions. In addition, a rough dynamic migration barrier distribution is more important than a static one for sluggish diffusion*

To search for the sluggish compositions, the diffusivities of 1500 non-equiatomic HEAs and their AA counterparts were calculated by the high-throughput ANN-KMC and AA-KMC, respectively. It is found that many non-equiatomic compositions exhibit sluggish diffusion, and six of them show very strong sluggish diffusion (more than 10 times slower than the AA counterpart). To understand the underlying mechanisms, both static and dynamic migration barrier distributions of the 1500 HEA compositions have been analyzed. The former ones are based on every atomic sites regardless of being visited or not, while the latter ones are only from those accepted migration events during diffusion. A key observation is that the variance in the static migration barrier distribution has a limited influence on diffusion, suggesting that a seemingly rough PEL is not necessary to lead a sluggish diffusion. This is because vacancies may find alternative paths to bypass high-barrier or trapping sites, leaving those sites unvisited during the diffusion. Conversely, it is found that the accepted/executed migration barriers have distinctly large variances in those sluggish compositions. This observation suggests that the effective trapping sites are those with localized low-energy barriers surrounded by high-barrier sites. In such cases, vacancies may perform many back-and-forth hops before they can escape.

- *The results from this work could provide valuable insights for experimental HEA design*

The compositions in these sluggish HEAs have been analyzed in detail (Fig. 16). It is found that they typically contain high-concentrations of slowest diffusing species (Fe or Ni in this work) and low-concentrations of fastest diffusing species (Cu in this work). In such cases, the local trapping effect is expected to be strong. These findings may provide useful insights for HEA design. It is envisioned that the high-throughput modeling framework developed in this work could be used to down-select sluggish compositions first, which can be validated experimentally in the future.

CRediT authorship contribution statement

W. Huang: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing – original draft, Visualization.

D. Farkas: Conceptualization, Methodology, Formal analysis, Writing – review & editing.

X.M. Bai: Conceptualization, Methodology, Writing – review & editing, Supervision, Funding acquisition.

Data availability

Data will be made available on request.

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.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] J. W. Yeh *et al.*, Nanostructured high - entropy alloys with multiple principal elements: novel alloy design concepts and outcomes. *Advanced engineering materials* **6**, 299-303 (2004).

[2] Y. Jien-Wei, Recent progress in high entropy alloys. *Ann. Chim. Sci. Mat* **31**, 633-648 (2006).

[3] P. Brown, H. Zhuang, Quantum machine-learning phase prediction of high-entropy alloys. *Materials Today* **63**, 18-31 (2023).

[4] S. Praveen, H. S. Kim, High - entropy alloys: potential candidates for high - temperature applications‐an overview. *Advanced engineering materials* **20**, 1700645 (2018).

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

[6] D. Liu *et al.*, Exceptional fracture toughness of CrCoNi-based medium-and high-entropy alloys at 20 kelvin. *Science* **378**, 978-983 (2022).

[7] H. Cheng *et al.*, Corrosion-resistant high-entropy alloy coatings: a review. *Journal of The Electrochemical Society* **168**, 111502 (2021).

[8] Y. Shi, B. Yang, P. K. Liaw, Corrosion-resistant high-entropy alloys: A review. *Metals* **7**, 43 (2017).

[9] C. Lu *et al.*, Enhancing radiation tolerance by controlling defect mobility and migration pathways in multicomponent single-phase alloys. *Nature communications* **7**, 13564 (2016).

[10] Z. Li, K. G. Pradeep, Y. Deng, D. Raabe, C. C. Tasan, Metastable high-entropy dual-phase alloys overcome the strength–ductility trade-off. *Nature* **534**, 227-230 (2016).

[11] 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).

[12] C. Lu *et al.*, Radiation-induced segregation on defect clusters in single-phase concentrated solid-solution alloys. *Acta Materialia* **127**, 98-107 (2017).

[13] H.-W. Chang *et al.*, Influence of substrate bias, deposition temperature and post-deposition annealing on the structure and properties of multi-principal-component (AlCrMoSiTi) N coatings. *Surface and Coatings Technology* **202**, 3360-3366 (2008).

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

[15] 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).

[16] 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).

[17] A. Mehta, Y. Sohn, Investigation of sluggish diffusion in FCC Al0. 25CoCrFeNi high-entropy alloy. *Materials Research Letters* **9**, 239-246 (2021).

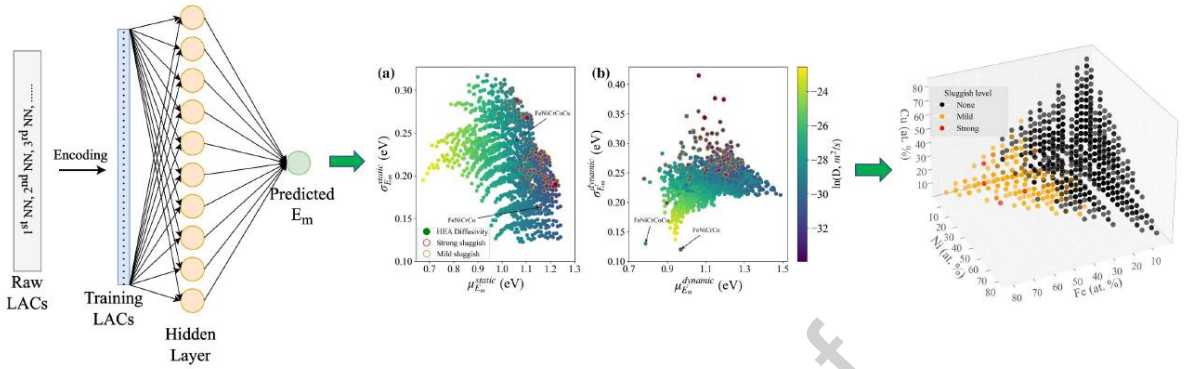
[18] A. Seoane, D. Farkas, X.-M. Bai, Influence of compositional complexity on species diffusion behavior in high-entropy solid-solution alloys. *Journal of Materials Research* **37**, 1403-1415 (2022).

[19] A. Seoane, D. Farkas, X.-M. Bai, Molecular dynamics studies of sluggish grain boundary diffusion in equiatomic FeNiCrCoCu high-entropy alloy. *Journal of Materials Science*, 1-17 (2023).
[20] M. Vaidya, S. Trubel, B. Murty, G. Wilde, S. V. Divinski, Ni tracer diffusion in CoCrFeNi and CoCrFeMnNi high entropy alloys. *Journal of Alloys and Compounds* **688**, 994-1001 (2016).
[21] Y. Osetsky, A. V. Barashev, Y. Zhang, Sluggish, chemical bias and percolation phenomena in atomic transport by vacancy and interstitial diffusion in NiFe alloys. *Current Opinion in Solid State and Materials Science* **25**, 100961 (2021).
[22] B. Xu *et al.*, Exploring the influence of percolation on vacancy-mediated diffusion in CoCrNi multi-principal element alloys. *Materials & Design* **223**, 111238 (2022).
[23] C. Varvenne, A. Luque, W. G. Nöhring, W. A. Curtin, Average-atom interatomic potential for random alloys. *Physical Review B* **93**, 104201 (2016).
[24] 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).
[25] S. L. Thomas, S. Patala, Vacancy diffusion in multi-principal element alloys: The role of chemical disorder in the ordered lattice. *Acta Materialia* **196**, 144-153 (2020).
[26] J. Kottke *et al.*, Experimental and theoretical study of tracer diffusion in a series of (CoCrFeMn) 100− xNix alloys. *Acta Materialia* **194**, 236-248 (2020).
[27] K. Ferasat *et al.*, Accelerated kinetic Monte Carlo: A case study; vacancy and dumbbell interstitial diffusion traps in concentrated solid solution alloys. *The Journal of Chemical Physics* **153**, 074109 (2020).
[28] 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).
[29] N. Mousseau *et al.*, The activation-relaxation technique: Art nouveau and kinetic art. *Journal of Atomic and Molecular Physics* **2012** (2012).
[30] B. Xu *et al.*, Mechanism of sluggish diffusion under rough energy landscape. *Cell Reports Physical Science* (2023).
[31] B. Xu *et al.*, Influence of short-range order on diffusion in multiprincipal element alloys from long-time atomistic simulations. *Physical Review Materials* **7**, 033605 (2023).
[32] B. Xu *et al.*, Revealing the crucial role of rough energy landscape on self-diffusion in high-entropy alloys based on machine learning and kinetic Monte Carlo. *Acta Materialia* **234**, 118051 (2022).
[33] W. Huang, X.-M. Bai, Machine learning based on-the-fly kinetic Monte Carlo simulations of sluggish diffusion in Ni-Fe concentrated alloys. *Journal of Alloys and Compounds* **937**, 168457 (2023).
[34] D. Farkas, A. Caro, Model interatomic potentials and lattice strain in a high-entropy alloy. *Journal of Materials Research* **33**, 3218-3225 (2018).
[35] M. Asadikiya *et al.*, A review of the design of high-entropy aluminum alloys: a pathway for novel Al alloys. *Journal of Materials Science* **56**, 12093-12110 (2021).
[36] S. Plimpton, Fast parallel algorithms for short-range molecular dynamics. *Journal of computational physics* **117**, 1-19 (1995).
[37] N. Castin, L. Malerba, Calculation of proper energy barriers for atomistic kinetic Monte Carlo simulations on rigid lattice with chemical and strain field long-range effects using artificial neural networks. *The Journal of Chemical Physics* **132** (2010).

[38] 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).
[39] 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).
[40] M. Leetmaa, N. V. Skorodumova, KMCLib: A general framework for lattice kinetic Monte Carlo (KMC) simulations. *Computer Physics Communications* **185**, 2340-2349 (2014).
[41] A. F. Voter, "Introduction to the kinetic Monte Carlo method" in Radiation effects in solids. (Springer, 2007), pp. 1-23.
[42] Y. Zhang, A. Manzoor, C. Jiang, D. Aidhy, D. Schwen, A statistical approach for atomistic calculations of vacancy formation energy and chemical potentials in concentrated solid-solution alloys. *Computational Materials Science* **190**, 110308 (2021).
[43] J. Cieslak *et al.*, Multi-phase nature of sintered vs. arc-melted CrxAlFeCoNi high entropy alloys-experimental and theoretical study. *Journal of Alloys and Compounds* **801**, 511-519 (2019).
[44] A. Munitz, M. Kaufman, R. Abbaschian, Liquid phase separation in transition element high entropy alloys. *Intermetallics* **86**, 59-72 (2017).
[45] W. G. Hoover, Canonical dynamics: Equilibrium phase-space distributions. *Physical review A* **31**, 1695 (1985).
[46] W. Shinoda, M. Shiga, M. Mikami, Rapid estimation of elastic constants by molecular dynamics simulation under constant stress. *Physical Review B* **69**, 134103 (2004).
[47] V. Borodin, P. Vladimirov, A. Möslang, Lattice kinetic Monte-Carlo modelling of helium–vacancy cluster formation in bcc iron. *Journal of nuclear materials* **367**, 286-291 (2007).
[48] 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).
[49] M. Vaidya, K. Pradeep, B. Murty, G. Wilde, S. Divinski, Bulk tracer diffusion in CoCrFeNi and CoCrFeMnNi high entropy alloys. *Acta Materialia* **146**, 211-224 (2018).
[50] M. S. Daw, M. Chandross, Sluggish diffusion in random equimolar FCC alloys. *Physical Review Materials* **5**, 043603 (2021).
[51] S. V. Divinski, A. V. Pokoev, N. Esakkiraja, A. Paul, A mystery of" sluggish diffusion" in high-entropy alloys: the truth or a myth? *Diffusion foundations* **17**, 69-104 (2018).
[52] J. Dąbrowa, M. Danielewski, State-of-the-art diffusion studies in the high entropy alloys. *Metals* **10**, 347 (2020).
[53] J. Dąbrowa *et al.*, Demystifying the sluggish diffusion effect in high entropy alloys. *Journal of Alloys and Compounds* **783**, 193-207 (2019).
[54] M. Vaidya, K. Pradeep, B. Murty, G. Wilde, S. Divinski, Radioactive isotopes reveal a non sluggish kinetics of grain boundary diffusion in high entropy alloys. *Scientific reports* **7**, 12293 (2017).
[55] K. Compaan, Y. Haven, Correlation factors for diffusion in solids. *Transactions of the Faraday Society* **52**, 786-801 (1956).
[56] B. Xing, X. Wang, W. J. Bowman, P. Cao, Short-range order localizing diffusion in multi-principal element alloys. *Scripta Materialia* **210**, 114450 (2022).

[57] J. M. Cowley, An Approximate Theory of Order in Alloys. *Physical Review* **77**, 669-675 (1950).
[58] J. W. Haus, K. W. Kehr, Diffusion in regular and disordered lattices. *Physics Reports* **150**, 263-406 (1987).
[59] I. Webman, Effective-medium approximation for diffusion on a random lattice. *Physical review letters* **47**, 1496 (1981).
[60] N. Castin, L. Messina, C. Domain, R. C. Pasianot, P. Olsson, Improved atomistic Monte Carlo models based on ab-initio-trained neural networks: Application to FeCu and FeCr alloys. *Physical Review B* **95**, 214117 (2017).
[61] A. Manzoor *et al.*, Machine learning based methodology to predict point defect energies in multi-principal element alloys. *Frontiers in Materials* **8**, 673574 (2021).
[62] L. Qiao, Y. Liu, J. Zhu, A focused review on machine learning aided high-throughput methods in high entropy alloy. *Journal of Alloys and Compounds* **877**, 160295 (2021).
[63] 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).

Graphical abstract



High-throughput machine learning + KMC

Sluggish diffusion mechanisms in non-equiatomic HEAs

HEA design guideline for non-equiatomic compositions