Integrating Atomistic Simulations and Machine Learning to Predict Unstable Stacking Fault Energies of Refractory Non-dilute Random Alloys

RICHARD BRINLEE,[1] AMIN POOZESH,[1] ANVESH NATHANI,[1] ANSHU RAJ,[1] XIANG-GUO LI,[2] IMAN GHAMARIAN,[1] and SHUOZHI XU [1,3]

1.—School of Aerospace and Mechanical Engineering, University of Oklahoma, Norman, OK 73019-1052, USA. 2.—School of Materials, Shenzhen Campus of Sun Yat-sen University, No. 66, Gongchang Road, Guangming District, Shenzhen 518107, Guangdong, People's Republic of China. 3.—e-mail: shuozhixu@ou.edu

This study utilizes three machine learning (ML) models—XGBoost, graph neural networks (GNN), and graph attention networks (GATs)—to predict unstable stacking fault energies (USFEs) in over 1000 refractory non-dilute random alloys, including mono, binary, ternary, quaternary, and quinary systems. Data for training are generated through atomistic simulations. In ML models, each chemical composition is encoded as a five-dimensional numerical vector to facilitate the representation of features, which are then normalized to balance the input scales during model training. Although explicit inter-composition connections are not included, the models learn compositional trends through node features and internal attention mechanisms. Because the lattice parameter is a byproduct in calculating the USFE in atomistic simulations, we also build three models for the lattice parameters. All models achieve reasonable accuracy in predicting lattice parameters but show variability in USFE predictions. Among the three models, GAT achieves the highest accuracy with average errors of 0.68% for USFE and 0.09% for the lattice parameter, outperforming GNN (1.24%, 0.11%) and XGBoost (0.46%, 0.54%). This study highlights the potential of graph-based ML models in predicting material properties from simulation-derived datasets, offering insights that could enhance traditional methods in materials science.

INTRODUCTION

Alloy design has traditionally relied on a straightforward approach: choosing one primary metal as the base and adding small amounts—usually < 10 atomic percent—of other elements to fine-tune properties like strength, ductility, and corrosion resistance.[1] These dilute alloys are typically positioned at the edges of multi-component phase diagrams, where solute atoms are sparsely distributed and have limited interactions with each other or the matrix.[2] Therefore, the performance of such alloys is largely determined by the dominant element, with the solutes playing only a secondary role. While this method has proven effective in many industrial applications, it overlooks the potential for more complex and beneficial interactions among multiple principal elements when used in higher concentrations.[3]

In contrast, non-dilute random alloys (NDRAs) embrace a more compositionally balanced philosophy.[4,5] These alloys consist of two or more principal elements in near-equiatomic ratios, arranged randomly on simple crystal lattices such as face-centered cubic and body-centered cubic (BCC).[6] We remark that the concept of NDRA encompasses high entropy alloy, which has five or more principal elements, medium entropy alloy, which has three or

four principal elements, and multi-principal element alloy or complex concentrated alloy, which has three or more principal elements. The absence of a single dominant matrix leads to a highly disordered atomic structure, where local chemical environments vary significantly.[7] This complexity gives rise to unique mechanical and thermodynamic behaviors, including exceptional high-temperature strength, thermal stability, and resistance to radiation—properties that cannot be predicted by simply averaging those of the individual elements.[8,9] In recent years, NDRAs that are primarily composed of refractory elements like Mo, Nb, Ta, W, and V, known as refractory NDRAs (RNDRAs), have attracted significant research attention.[10–12] Because of their excellent mechanical and thermal properties, RNDRAs are emerging as strong candidates for demanding applications in aerospace, nuclear energy, and other extreme environments, where conventional refractory materials often fall short.[13–15]

In this paper, we focus on the unstable stacking fault energies (USFEs), which play a critical role in determining the deformation mechanisms,[16,17] of RNDRAs. Recent work by Ma et al.[18] has shown that in NDRAs, SFEs can vary significantly across different regions because of local chemical fluctuations—variations that strongly impact dislocation behavior and mechanical response. To account for this complexity without relying on computationally expensive methods, LaRosa et al.[19] introduced a bond energy model that predicts both average and distributed SFEs based on bulk phase energy differences. Similarly, Jagatramka et al.[20] developed a set of analytical tools to statistically describe potential energy fluctuations in NDRAs, offering valuable insights into cohesive and fault energy landscapes that help clarify the role of local environments in defect energetics.

Accurately assessing USFEs would offer valuable insight into the underlying atomic-scale behavior of RNDRAs and provide a foundation for tailoring their mechanical performance through composition design.[21,22] One of the most commonly used approaches to calculate USFEs of metallic alloys is the atomistic simulation method. However, high-fidelity atomistic simulations, particularly those enabled by interatomic potentials with quantum-level accuracy, can be computationally expensive. Additionally, before calculating the USFE for a specific crystal, the lattice parameters must first be determined to construct appropriate structures, further increasing the overall computational cost.[7]

As a result, in recent years, machine learning (ML) models have been increasingly employed as surrogates to predict material properties. Various ML algorithms, including random forest (RF), Gaussian process regression (GPR), and neural networks (NN), have demonstrated considerable predictive capabilities for properties across diverse alloy systems.[23–26] For example, RF models have

been effective in handling large datasets and capturing non-linear dependencies among compositional and structural features.[23] GPR and NN methods have also achieved notable success, especially when trained on smaller but high-quality datasets, as they can incorporate uncertainty quantification and non-linear mapping capabilities.[27,28]

For lattice parameter prediction, models such as RF, GPR, and NN have been employed using compositional and elemental descriptors—such as atomic radii, electronegativity, and valence electron concentration—as inputs.[23–25] For USFE prediction, studies have applied support vector machines (SVM), deep NN (DNN), gradient boosting machines (GBM), and convolutional NN (CNN).[29,30] These approaches have demonstrated promising performance in capturing the structural trends in alloys, especially when trained on well-curated datasets. However, they often struggle to generalize across different alloy families because of limited compositional diversity and insufficient atomic-scale descriptors.[31] Furthermore, these models tend to underperform when exposed to non-linear phenomena influenced by electronic structure, magnetic interactions, or phonon contributions.[32,33]

Recent advancements in graph-based deep learning offer promising directions. Specifically, graph attention networks (GATs) provide a framework for encoding relational information between constituent elements in a material system, allowing more nuanced understanding of inter-element interactions.[26] Unlike traditional ML models that rely on handcrafted descriptors or tabular data, GATs learn from structured graphs representing compositional and structural similarity, making them ideal for modeling non-dilute alloys.[34,35] In the literature, GATs have shown promising results in predicting a variety of material properties. For example, Park et al.[36] demonstrated that GATs outperform other graph NN (GNN) variants in modeling the formation energy and band gap of crystalline materials. Similarly, Fung et al.[37] used GATs to effectively model atomic interactions, capturing subtle nuances in structural and electronic environments. The present research leverages a GAT model to improve predictive performance for both lattice parameters and USFE, offering a flexible and generalizable architecture. Performance of the GAT model is compared with those of XGBoost and GNN to highlight GAT's capability of capturing complex compositional dependencies inherent to RNDRAs.

MATERIALS AND METHODS

Atomistic Simulations

We consider pure metals and RNDRAs consisting of one or more of these five refractory elements: Mo, Nb, Ta, V, and W. In total, there are 5 pure metals, 990 binaries, 10 ternaries, 5 quaternaries, and 1 quinary. All alloys are summarized in Table 1. A recently developed ML-based moment tensor

Integrating Atomistic Simulations and Machine Learning to Predict Unstable Stacking Fault Energies of Refractory Non-dilute Random Alloys

8129

Table 1. All RNDRAs studied in this paper. For binaries, x varies from 0.01 to 0.99. All ternaries, quaternaries, and quinary are equimolar

990 binaries	Mo_xNb_{1-x}, Mo_xTa_{1-x}, Mo_xV_{1-x}, Mo_xW_{1-x}, Nb_xTa_{1-x}, Nb_xV_{1-x}, Nb_xW_{1-x}, Ta_xV_{1-x}, Ta_xW_{1-x}, V_xW_{1-x}
10 ternaries	MoNbTa, MoNbV, MoNbW, MoTaV, MoTaW, MoVW, NbTaV, NbTaW, NbVW, TaVW
5 quaternaries	MoNbTaV, MoNbTaW, MoNbVW, MoTaVW, NbTaVW
1 quinary	MoNbTaVW

potential (MTP) is used to describe the interactions among elements.[38] LAMMPS is employed for all atomistic simulations.[39]

To calculate the lattice parameter, we utilize simulation cells that are oriented such that the [100], [010], and [001] crystallographic directions are aligned with the x-, y-, and z-axes, respectively. Atoms are randomly distributed on a BCC lattice based on their compositional ratio. Periodic boundary conditions (PBCs) are applied in all three directions. We use the volume-energy method. Beginning with an initial trial lattice parameter of 3.3 Å, each structure was incrementally stretched or shrunk, and the total bulk energy was computed at each step. The actual lattice parameter a_0 was identified as the value corresponding to the lowest energy.[40] A total of 1011 LAMMPS simulations are conducted to obtain all lattice parameter data, comprising the sum of values listed in Table 1, plus those for the five pure metals.

Next, we build atomistic structures for the USFE calculations based on a_0. We build simulation cells such that the x, y, and z axes are respectively aligned with $[11\bar{2}]$, $[111]$, and $[1\bar{1}0]$. PBCs are applied along the x and y axes but not the z axis.[41] Along the z direction, we divide the simulation cell into an upper half and a lower half. To obtain USFE in the studied NDRAs, we calculate generalized SFE (GSFE) curves by applying incremental shear displacements to the upper half of each simulation cell relative to the lower half.[42] The displacement is performed within the xy plane along the y axis until it reaches $\sqrt{3}a_0/2$. At each increment, the system's total energy was minimized while (1) fixing the top two and bottom two layers of atoms and (2) allowing all remaining atoms to move along the z axis only. In RNDRAs, the stacking fault energy is highly sensitive to local atomic configurations, resulting in variability across different slip planes.[43,44] To ensure robustness and statistical relevance, GSFE calculations are repeated on 20 distinct, randomly chosen xy planes for each alloy composition. The USFE was then defined as the peak value from each GSFE curve.[45] As a result, a total of 20,125 LAMMPS calculations are performed to obtain all USFE data. For each alloy, only the mean USFE value, among the 20 data points, will be used in subsequent analyses. To support the reliability of the selected MTP, we refer to the work by Mubassira et al.,[42] where a good agreement between MTP and DFT-derived USFE values for representative alloys was presented.

The accuracy of calculated properties in NDRAs can be quite sensitive to the cell size. Xu et al.[46] found that lattice parameters vary depending on the atomic arrangement used in simulations; thus, if a single configuration is used, it must be sufficiently large, as in the current work. In another study, Xu et al.[47] showed that USFE is also affected by the cell size—particularly the cross-sectional area of the cell. It was found that while the mean USFE is largely size-independent, the standard deviation becomes negligible when the area is larger than about 10 nm². Since the current work only concerns the mean USFE value, the size dependence is of no concern.

Development of Machine Learning Models

Based on atomistic simulation data in the previous section, we develop ML models to predict both the lattice parameter and USFE from the compositional data of 1006 RNDRAs and 5 pure metals. The dataset consists of the combination name, lattice parameter, and mean USFE. To make these data usable in ML pipelines, we apply a feature extraction step that converts chemical formulas into fixed-length numerical vectors based on a predefined elemental basis [Mo, Nb, Ta, V, W]. For example, $Mo_{0.1}Nb_{0.9}$ becomes [10, 90, 0, 0, 0], with atomic percentages represented as integers.[23,24] Feature normalization is performed using z-score standardization: $xnorm = (x - \mu)/\sigma$, where μ and σ are computed from the training set only to prevent data leakage. For compositional features, we additionally apply a compositional data transformation using the centered log ratio to account for the constant-sum constraint. This dual normalization strategy ensures that the model treats all elemental contributions equitably while properly addressing the compositional nature of the input space (Table 2).

As a baseline, we train an XGBoost model, which is one of the best implementations of the GBM framework. It is known that XGBoost is effective for tabular data and capable of modeling non-linear dependencies.[23] Here, in addition to using only compositional vectors as inputs, the XGBoost model employs matminer's ElementProperty featurizer with the "Magpie" preset, generating 132 features per composition, including elemental statistics (mean, standard deviation, maximum, minimum) for properties like atomic radius, electronegativity, and valence electrons. we employ 100 estimators with a maximum depth of 6, learning rate $\eta = 0.10$,

Table 2. Numerical encoding of example alloy compositions used as *node features* in the GNN and GAT models. Each vector lists the atomic-percent of the five elements in the fixed order [Mo, Nb, Ta, V, W]

Combination name	Feature vector
$Mo_{0.1}Nb_{0.9}$	[10, 90, 0, 0, 0]
$Mo_{0.3}Ta_{0.7}$	[30, 0, 70, 0, 0]
$Nb_{0.5}V_{0.5}$	[0, 50, 0, 50, 0]
$Ta_{0.25}W_{0.75}$	[0, 0, 25, 0, 75]
$V_{0.4}W_{0.6}$	[0, 0, 0, 40, 60]
MoNbV	[33, 33, 0, 33, 0]
MoNbVW	[25, 25, 0, 25, 25]
MoNbTaVW	[20, 20, 20, 20, 20]

subsample ratio of 0.80, and column sampling ratio of 0.80. The model uses squared error loss optimized with RMSE as the evaluation metric. Early stopping with 10 epochs patience prevents overfitting.

To capture relational similarities among alloy compositions, we construct a graph where each node represents a unique composition, while edges are defined using Euclidean distance in composition space. The resulting graph structure allows for the application of graph-based learning models capable of modeling inter-compositional relationships. It follows that we train a GNN model using three GCNConv layers with hidden dimensions of 128 and ReLU activation. The GNN employs fixed mean aggregation weights, contrasting with GAT's adaptive attention mechanism. Despite using the same optimizer settings (AdamW, learning rate: 1×10^{-4}, weight decay: 5×10^{-4}) and dynamic loss weighting scheme, the GNN's inability to distinguish neighbor importance results in 15–20% higher prediction errors for quaternary and quinary alloys compared to GAT. Although GNNs can operate on graph-structured data, they aggregate neighbor information using fixed weights, which can limit their ability to model asymmetric and context-dependent interactions between elements.[26,35]

On the other hand, GAT introduces learnable attention weights, enabling the model to assign different levels of importance to neighboring nodes during message passing.[36,37] This capability is particularly useful for modeling RNDRAs, where small variations in local chemical environments can significantly affect material properties.

Figure 1 illustrates the overall GAT model workflow used in this study, including feature extraction from alloy compositions, graph construction based on compositional similarity, and the training pipeline. The GAT architecture employs a multi-scale attention mechanism: the first two GATConv layers use four attention heads each (concatenated for a 512-dimensional representation), while the final layer uses a single head to produce focused predictions. This multi-head design enables the model to

simultaneously learn different types of compositional relationships—experimental visualization confirms that different heads specialize in capturing size-based, electronic, and structural similarities among elements. A hidden dimension of 128 and dropout ($p = 0.5$) are used to improve generalization. A multi-layer perceptron is appended after the attention layers to serve as a regression head. The model is optimized using the AdamW optimizer with a learning rate of 1×10^{-4} and weight decay of 5×10^{-4}. The model is trained for 300 epochs with early stopping based on validation loss (patience = 20). A dynamic loss weighting scheme is used to jointly optimize USFE and lattice parameter predictions; this weighting typically converges to a stable 0.85:0.15 (USFE:LatPara) ratio, reflecting the greater complexity of predicting USFE. During training, attention weights begin to stabilize after approximately 150 epochs. Post hoc analysis reveals that different attention heads specialize in distinct compositional patterns: head 1 focuses on nearest-neighbor interactions, heads 2 and 3 capture medium-range dependencies, and head 4 identifies long-range similarities across alloy compositions.

During training for GNN and GAT, we employ a dynamic loss weighting scheme to adaptively balance errors between the USFE and lattice parameter predictions. The total loss was defined as:

$$\mathcal{L}_{total} = \frac{w_{USFE} \cdot \mathcal{L}_{USFE} + w_{Lat} \cdot \mathcal{L}_{Lat}}{w_{USFE} + w_{Lat}}$$
$$\text{where} \quad w_i = \frac{1}{\mathcal{L}_i + \varepsilon} \qquad (1)$$

where $\mathcal{L}_{USFE}$ and $\mathcal{L}_{Lat}$ are the individual mean squared errors, and ε is a small constant to prevent division by zero. This formulation directs more attention toward harder-to-learn targets. Additionally, L2 regularization is applied to further mitigate overfitting.

To ensure reliable model training and evaluation, the dataset of 1011 compositions was randomly split into training and testing subsets with an 80:20 ratio. We confirm that the random sampling was performed without bias toward any specific range of composition, lattice parameter, or USFE. As a result, the training set samples the alloy composition space broadly and includes a wide range of property values rather than being linearly or uniformly distributed.

In terms of computational efficiency, the GAT model requires approximately 3.2 s per epoch on an NVIDIA A100 GPU with our dataset of 1011 compositions. The total training time of 45 min (including hyperparameter optimization) represents a $1000\times$ speedup compared to the 20,125 LAMMPS simulations (approximately 750 CPU-hours) required to generate the USFE data, demonstrating the practical value of ML surrogates for high-throughput alloy screening.
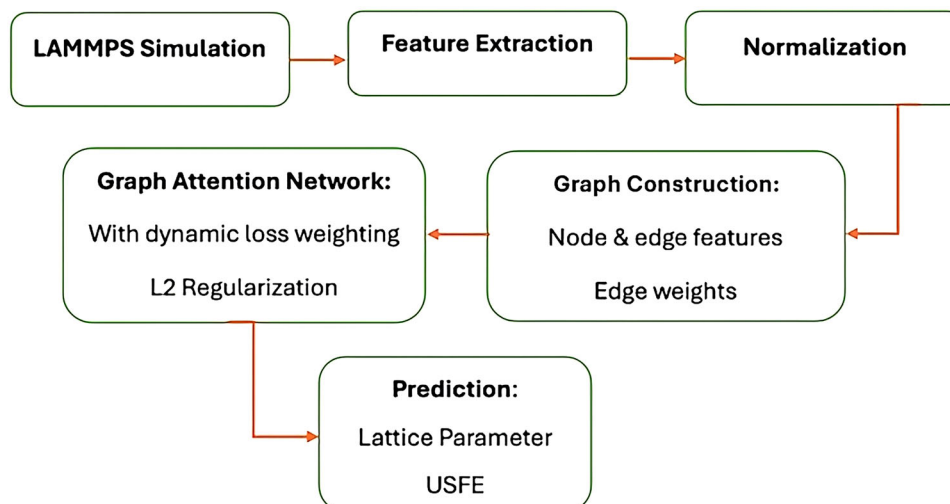
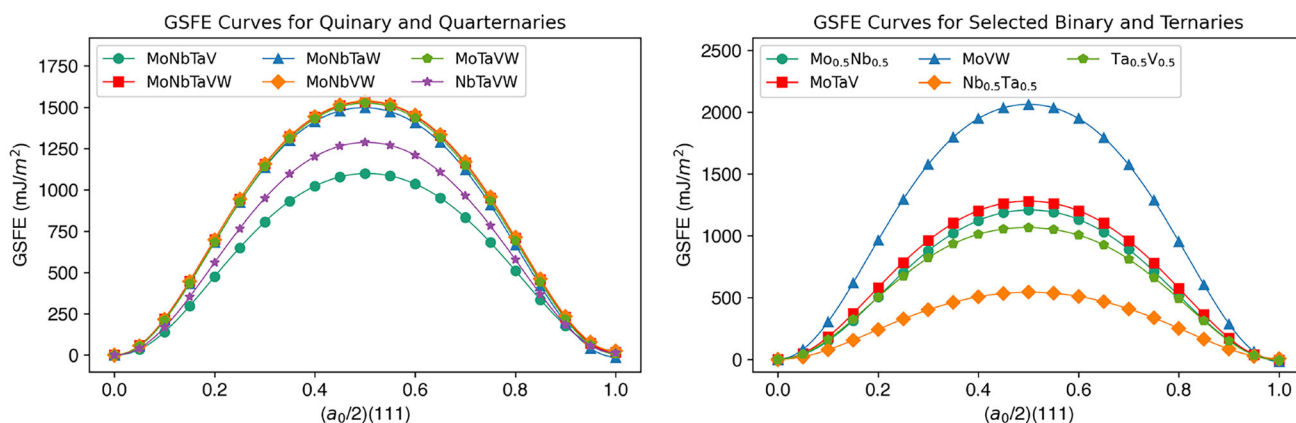Fig. 1. Schematic representation of the GAT workflow.

Fig. 2. GSFE curves for all quinary and quaternary systems, along with selected binary and ternary alloys.

RESULTS AND DISCUSSION

Atomistic Simulation Results

First, we present GSFE curves in multiple representative alloys in Fig. 2. It is found that while the GSFE curves across different alloys all have a common unimodal shape, the peak values, i.e., USFEs, vary greatly with the composition. Notably, alloys containing Mo and W tend to exhibit higher USFE values, suggesting enhanced resistance to dislocation glide, whereas those rich in Nb and V typically show lower USFE peaks. In addition to peak values, the breadth of the GSFE curves also varies; some alloys display broader curves, indicating a more gradual increase and decrease in fault energy along the slip plane. This may reflect a wider spread in the atomic-scale resistance to shear displacement. These variations in both the magnitude and shape of the GSFE curves have direct implications for mechanical behavior: alloys with higher and broader USFE profiles are generally expected to exhibit reduced dislocation mobility,

potentially leading to higher yield strength but lower ductility.

Our atomistic simulation results across over 1000 RNDRAs reveal the USFE's dependence on chemical composition. To illustrate this, we present USFEs in all 990 binaries and 5 pure metals in Fig. 3. In six of the ten sets of binaries, the USFE varies nearly linearly with the composition; as a result, the binaries' USFEs are between those of two consistent pure metals. In the four remaining sets of binaries, however, some binaries have higher or lower USFE than both constituent pure metals. Two notable examples are Nb_xV_{1-x} and Ta_xV_{1-x}, whereas alloys with intermediate compositional ratios such as $Nb_{0.3}V_{0.7}$ and $Ta_{0.32}V_{0.68}$ have the lowest USFEs among all binaries in their individual binary sets. To further understand this behavior, we evaluated the lattice distortion in these systems and found values of 0.0433 for $Nb_{0.5}V_{0.5}$ and 0.0438 for $Ta_{0.5}V_{0.5}$—among the highest observed across all equimolar binaries. This elevated distortion indicates significant local atomic strain, which likely contributes to the unusually low USFEs.
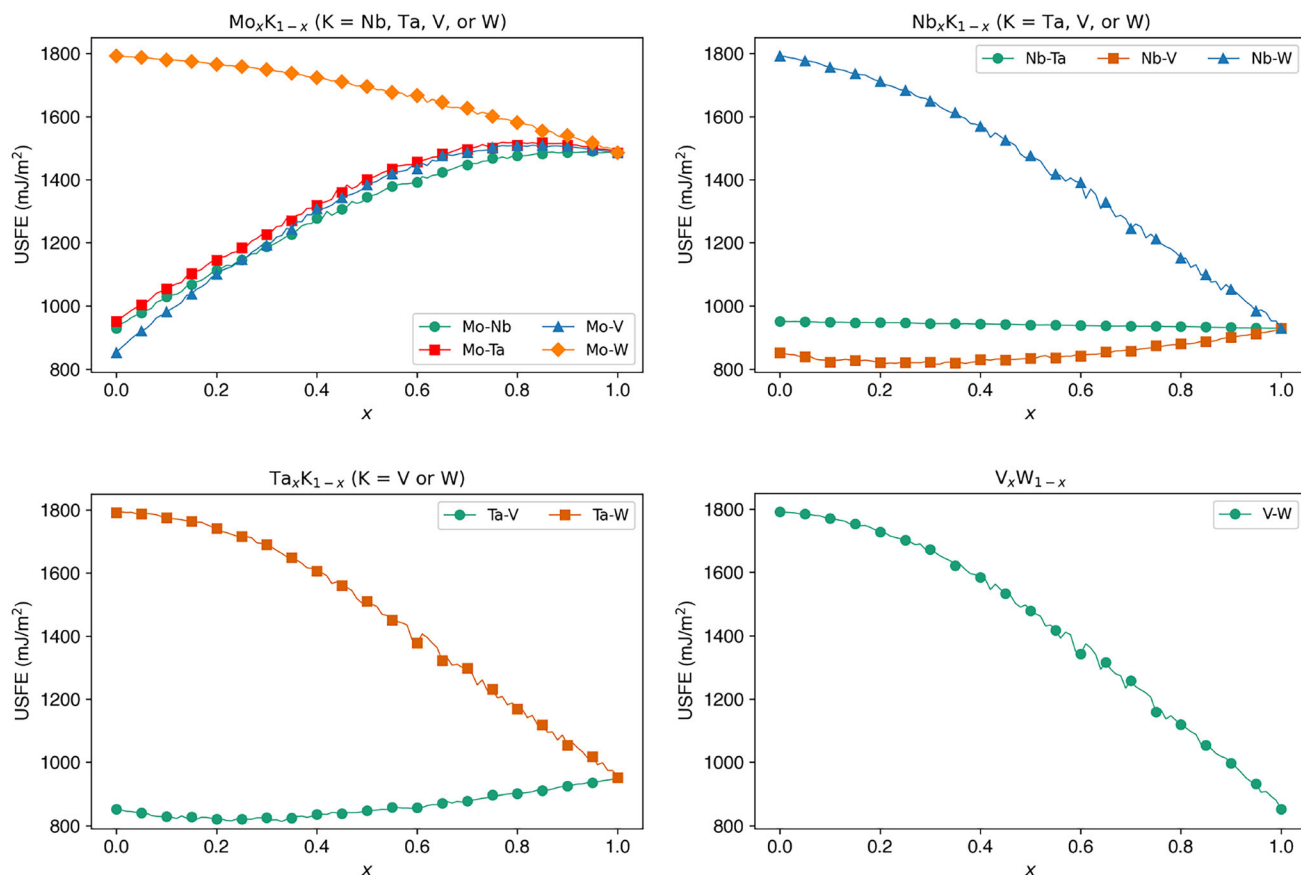
Fig. 3. Variation of USFE as a function of x in ten binary alloy systems.

Machine Learning Models

The ML models' performances in predicting materials' properties are mainly assessed using the calculated average and maximum percent differences between predicted and actual values.

First, we analyze the XGBoost model. Despite employing matminer's Magpie featurizer to generate 132 compositional descriptors—while the other two graph-based ML models do not—the XGBoost model achieves only moderate accuracy. The model shows an average percent difference of 0.46% for USFE and 0.54% for lattice parameter, with maximum errors reaching 2.41% and 1.32%, respectively. The overall MSE combining the USFE and lattice parameter of 31.52 reflects the model's limitations in capturing complex compositional dependencies without relational information. Notably, these errors represent normalized values; when considering the full range of USFE values (about 800–1200 mJ/m^2), the effective error rate increases substantially. For the lattice parameter, the slight upward shift arises naturally from its squared-error objective and regularized tree ensemble. Results are shown in Fig. 4.

Second, we discuss the GNN model. The GNN achieves an average percent difference of 1.24% for USFE and 0.11% for lattice parameter, with

maximum errors of 6.29% and 0.60%, respectively. While the lattice parameter prediction shows excellent accuracy ($R^2 = 0.997$), the USFE prediction exhibits poorer performance ($R^2 = 0.947$). This performance gap suggests that fixed aggregation weights fail to capture the varying importance of different compositional neighbors—a limitation addressed by GAT's attention mechanism. Results of the GNN model are shown in Fig. 5.

Finally, we present results from the GAT model. For the lattice parameter, the average percent difference was 0.09%, reflecting excellent agreement between predicted and experimental values. For the USFE, the average percent difference was 0.68%, indicating a high degree of accuracy in predictions. However, it is important to note that while these results suggest good overall performance, there were instances where larger discrepancies emerged. The maximum percent differences observed are 0.24% for the lattice parameter and 4.82% for the USFE. These findings highlight the variability in predictions, particularly for USFE values. The relatively small maximum percent difference for the lattice parameter suggests that the model is highly consistent in capturing minor variations within the materials dataset. In contrast, the slightly higher maximum percent difference for USFE implies that the model may struggle with
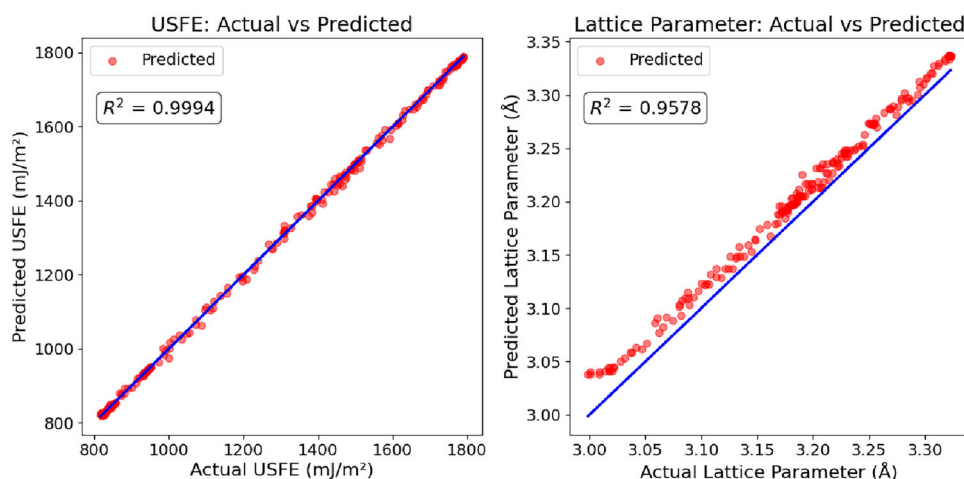
Fig. 4. A comparison between actual and predicted USFEs and lattice parameters based on the XGBoost model.
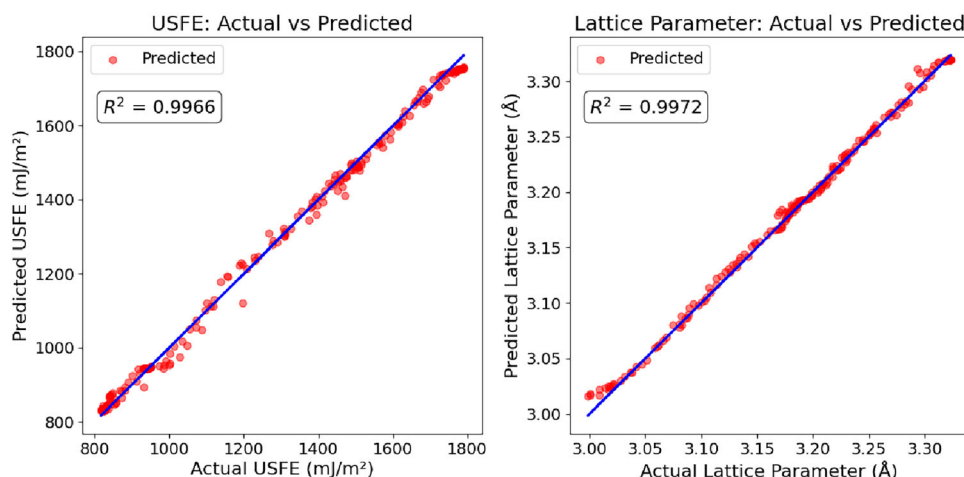

Fig. 5. A comparison between actual and predicted USFEs and lattice parameters based on the GNN model.

more predictions that have a higher variation. Results are shown in Fig. 6.

Overall, the findings from this study suggest that ML models hold significant potential for predicting materials' properties when trained on small amounts of high-quality data obtained through traditional methods. In particular, the GAT model, due to its ability to assign adaptive importance weights to neighboring nodes through attention mechanisms, outperforms both the XGBoost and GNN models.

Table 3 summarizes the comparative performance of all three models. The GAT's superior performance can be attributed to its attention mechanism, which dynamically adjusts the importance of neighboring compositions based on their relevance to the target node. This is particularly beneficial for RNDRAs where small compositional variations can lead to significant property changes. The attention weights analysis reveals that the model learns to assign higher importance to compositionally similar alloys, effectively creating a soft

clustering in the latent space. Although the GAT generalizes well across most compositions, it can underperform in sparsely sampled quaternary and quinary regions. To address this, we can run targeted atomistic simulations to obtain USFE and lattice parameters for non-equimolar quaternaries and quinaries. Adding physics-informed descriptors and applying transfer learning or active learning loops will further sharpen predictions on these rare compositions.

The computational complexity of our models varies significantly with dataset size N. XGBoost scales as $O(N \log N)$ due to recursive tree construction. The GNN complexity reduces to $O(NF^2)$ in our one-node graph setting (each graph carries only a self-edge); GAT adds an extra factor H for the number of attention heads, yielding $O(NkHF^2)$. While GAT requires roughly $2\times$ more computation than GNN, the accuracy gains—particularly for quaternary and quinary systems—justify the cost. Note that all ML models offer $> 1000\times$ speedup relative to high-fidelity atomistic simulations.
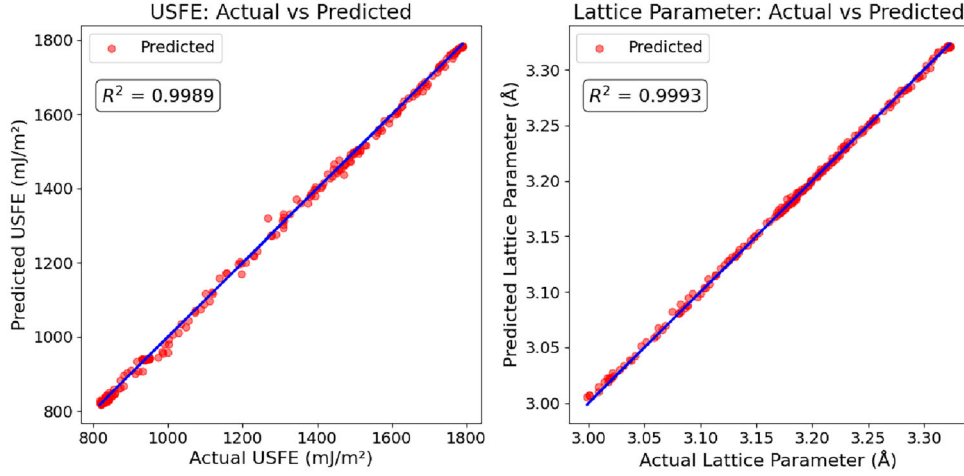
Fig. 6. A comparison between actual and predicted USFEs and lattice parameters based on the GAT model.

Table 3. Comparison of ML models for predicting USFEs and lattice parameters

Metric	XGBoost	GNN	GAT
Avg. % error (USFE)	4.30	2.70	0.68
Max % error (USFE)	12.10	6.40	4.82
R^2 (USFE)	0.9994	0.9977	0.9989
Avg. % error (latt. para.)	0.48	0.32	0.09
Max % error (latt. para.)	1.90	0.80	0.24
R^2 (latt. para.)	0.9578	0.9963	0.9993

Stratified error analysis reveals systematic trends for the GAT. Pure elements exhibit $< 0.05\%$ error for both USFEs and lattice parameters. Errors increase with compositional complexity: binaries (0.4%, 0.07%), ternaries (0.8%, 0.10%), quaternaries (1.1%, 0.13%), and quinary systems (1.5%, 0.17%) for USFE and lattice, respectively. Compositions near binary extremes (e.g., $Mo_{0.99}Nb_{0.01}$) show up to $2\times$ higher errors, reflecting reduced model confidence in extrapolation regions. Additionally, alloys containing V consistently yield $\sim 20\%$ higher USFE errors, likely due to V's small atomic radius, which is not captured by composition alone. The lattice distortion values of $Nb_{0.5}V_{0.5}$ (0.0433) and $Ta_{0.5}V_{0.5}$ (0.0438)—among the highest in our dataset—suggest that V introduces considerable local atomic strain, attributed to the model's reduced accuracy in predicting USFE for V-containing systems.

CONCLUSIONS

The study demonstrates the potential of GATs in predicting material properties, such as lattice parameter and USFE, of RNDRAs. By leveraging data from LAMMPS simulations across mono, binary, ternary, quaternary, and quinary systems,

the research successfully preprocesses material combinations into numerical features, enabling the model to capture relationships among different materials. Compared with the XGBoost and GNN models, the GAT model achieves a higher accuracy in predicting both lattice parameter and USFE. Between the two material parameters, all ML models are effective for simple structural properties like the lattice parameter, yet they may struggle with more complex properties like USFE. Despite these limitations, the study highlights the promising potential of ML in materials science, offering a robust framework for predicting material features from limited datasets. In this context, our model serves as a practical tool to estimate USFEs across complex alloy systems, helping to identify promising candidates for further study. While USFEs do not fully describe dislocation-mediated plasticity, they still offer valuable insight into dislocation behavior, especially in the early stages of alloy screening.[48] By combining physical understanding with data-driven methods, this approach provides an efficient path for exploring large compositional spaces without the need for costly simulations.

Future research will refine the model architecture, enrich the input space with local structural and electronic descriptors, and test transferability to broader refractory systems, e.g., Cr-containing corrosion-resistant alloys and Hf-containing ductile alloys. Despite the GAT's strong performance, three limitations remain:

1. *Feature set* Reliance on composition alone omits physics-relevant terms such as atomic-size mismatch and electronegativity spread.
2. *Graph scope* Each composition is treated as a one-node graph without inter-composition edges. This simplifies batching but prevents message passing across samples. Future work should explore persistent global graphs.

Integrating Atomistic Simulations and Machine Learning to Predict Unstable Stacking Fault
Energies of Refractory Non-dilute Random Alloys

8135

3. *External validation* The model is untested for RNDRAs that contain non-BCC elements such as Hf, Zr, and Re. Multi-fidelity training that blends atomic-scale simulation and experimental data, together with physics-informed constraints, should enhance accuracy and generalizability.

Addressing these points—particularly the larger errors observed for a few extreme end-member compositions—will be essential for deploying graph-based surrogates in high-throughput alloy design.

ACKNOWLEDGEMENTS

DATA AVAILABILITY

The datasets generated and/or analyzed during the current study are available in this GitHub repository: https://github.com/shuozhixu/JOM_2025

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

1. A. Brezini, R. Bouamrane, F. Hamdache, and C. Depolier, *Phys. Status Solidi* 188, 697 (1995).
2. Y. Ye, Q. Wang, J. Lu, C. Liu, and Y. Yang, *Mater. Today* 19, 349 (2016).
3. R.A. Romero, S. Xu, W.-R. Jian, I.J. Beyerlein, and C.V. Ramana, *Int. J. Plast.* 149, 103157 https://doi.org/10.1016/j.ijplas.2021.103157 (2022).
4. J.-W. Yeh, S.-K. Chen, S.-J. Lin, J.-Y. Gan, T.-S. Chin, T.-T. Shun, C.-H. Tsau, and S.-Y. Chang, *Adv. Eng. Mater.* 6, 299 (2004).
5. B. Cantor, I.T.H. Chang, P. Knight, and A.J.B. Vincent, *Mater. Sci. Eng. A* 375–377, 213 https://doi.org/10.1016/j.msea.2003.10.257 (2004).
6. S. Nag, and W.A. Curtin, *Acta Mater.* 200, 659 https://doi.org/10.1016/j.actamat.2020.08.011 (2020).
7. X. Wang, S. Xu, W.-R. Jian, X.-G. Li, Y. Su, and I.J. Beyerlein, *Comput. Mater. Sci.* 192, 110364 (2021).
8. M. Laurent-Brocq, L. Perrière, R. Pirès, F. Prima, P. Vermaut, and Y. Champion, *Mater. Sci. Eng. A* 696, 228 (2017).
9. Z. Li, S. Zhao, R.O. Ritchie, and M.A. Meyers, *Prog. Mater. Sci.* 102, 296 (2019).
10. O.N. Senkov, D.B. Miracle, K.J. Chaput, and J.-P. Couzinié, *J. Mater. Res.* 33, 3092 (2018).
11. J.-P. Couzinié, and G. Dirras, *Mater. Charact.* 147, 533 (2019).
12. O.N. Senkov, S. Gorsse, and D.B. Miracle, *Acta Mater.* 175, 394 (2019).
13. E. Mak, B. Yin, and W. Curtin, *J. Mech. Phys. Solids* 152, 104389 (2021).
14. Z. Wang, and S. Zhang, *Coatings* 13, 1916 https://doi.org/10.3390/coatings13111916 (2023).
15. A. Kumar, A. Singh, and A. Suhane, *J. Mater. Res. Technol.* 17, 2431 https://doi.org/10.1016/j.jmrt.2022.01.141 (2022).
16. X. An, S. Wu, Z. Wang, and Z. Zhang, *Prog. Mater. Sci.* 101, 1 (2019).
17. A. Raj, S. Mubassira, and S. Xu, *High Entropy Alloys Mater.* 3, 203 https://doi.org/10.1007/s44210-025-00060-8 (2025).
18. E. Ma, *Scr. Mater.* 181, 127 https://doi.org/10.1016/j.scriptamat.2020.02.021 (2020).
19. C.R. LaRosa, and M. Ghazisaeidi, *Acta Mater.* 238, 118165 https://doi.org/10.1016/j.actamat.2022.118165 (2022).
20. R. Jagatramka, C. Wang, and M. Daly, *Comput. Mater. Sci.* 214, 111763 https://doi.org/10.1016/j.commatsci.2022.111763 (2022).
21. Q.-J. Li, H. Sheng, and E. Ma, *Nat. Commun.* 10, 3563 (2019).
22. A. Jarlöv, W. Ji, Z. Zhu, Y. Tian, R. Babicheva, R. An, H.L. Seet, M.L.S. Nai, and K. Zhou, *J. Alloy. Compd.* 905, 164137 (2022).
23. L. Ward, R. Liu, A. Krishna, V.I. Hegde, A. Agrawal, A. Choudhary, and C. Wolverton, *Npj Comput. Mater.* 2, 16028 https://doi.org/10.1038/npjcompumats.2016.28 (2016).
24. D. Jha, L. Ward, A. Paul, W.-K. Liao, A. Choudhary, C. Wolverton, and A. Agrawal, *Sci. Rep.* 8, 17593 https://doi.org/10.1038/s41598-018-35934-y (2018).
25. C. Ruiz, A. Raj, and S. Xu, *APL Mach. Learn.* 2(2), 026107 (2024).
26. T. Xie, and J.C. Grossman, *Phys. Rev. Lett.* 120, 145301 https://doi.org/10.1103/PhysRevLett.120.145301 (2018).
27. X. Feng, Y. Zhang, Y. Zhao, Y. Zhu, and C. Chen, *J. Appl. Phys.* 126, 154303 https://doi.org/10.1063/1.5118831 (2019).
28. Y. Zhao, J. Liu, H. Dong, and X. Li, *Comput. Mater. Sci.* 199, 110722 https://doi.org/10.1016/j.commatsci.2021.110722 (2021).
29. X. Dai, C. Wei, Y. Zhou, and J. Zhang, *Mater. Des.* 191, 108640 https://doi.org/10.1016/j.matdes.2020.108640 (2020).
30. P. Zhang, Y. Lu, and Y. Wang, *Npj Comput. Mater.* 8, 107 https://doi.org/10.1038/s41524-022-00791-w (2022).
31. H. Singh, and A. Mishra, *Comput. Mater. Sci.* 226, 112001 https://doi.org/10.1016/j.commatsci.2023.112001 (2023).
32. A. Gupta, N. Singh, and S. Sharma, *Mater. Today: Proc.* 26, 2634 https://doi.org/10.1016/j.matpr.2020.02.558 (2020).
33. S. Lu, T. Li, L. Wang, and W. Yang, *Comput. Mater. Sci.* 210, 111444 https://doi.org/10.1016/j.commatsci.2022.111444 (2022).
34. Z. Liu, C. Wang, Y. Xu, V. Hegde, and A. Agrawal, *Comput. Mater. Sci.* 196, 110504 https://doi.org/10.1016/j.commatsci.2021.110504 (2021).
35. A. Jain, S.P. Ong, and C. Chen, *NPJ Comput. Mater.* 7, 94 https://doi.org/10.1038/s41524-021-00565-y (2021).
36. C. Park, H. Kim, and J. Lee, *Mater. Today Commun.* 32, 103999 https://doi.org/10.1016/j.mtcomm.2022.103999 (2022).
37. V. Fung, D. Jha, J. Yoo, J. Tran, K. Yao, C. Wolverton, and E. Kim, *Chem. Mater.* 35, 1234 https://doi.org/10.1021/acs.chemmater.2c03837 (2023).
38. T. Wang, et al., *NPJ Comput. Mater.* 10, 143 https://doi.org/10.1038/s41524-024-01330-6 (2024).
39. A. Thompson, et al., *Comput. Phys. Comm.* 271, 108171 https://doi.org/10.1016/j.cpc.2021.108171 (2022).
40. Y. Zhang, Y.J. Zhou, J.P. Lin, G.L. Chen, and P.K. Liaw, *Adv. Eng. Mater.* 10, 534 (2008).
41. Y. Su, S. Xu, and I.J. Beyerlein, *J. Appl. Phys.* 126, 105112 (2019).
42. S. Mubassira, M. Fani, A. Raj, C. Hirt, R.S. Brinlee, A. Poozesh, W.-R. Jian, S.Z. Chavoshi, C. Lee, and S. Xu, *Comput. Mater. Sci.* 248, 113587 https://doi.org/10.1016/j.commatsci.2024.113587 (2025).
43. J. Ding, Q. Yu, M. Asta, and R.O. Ritchie, *Proc. Natl. Acad. Sci.* 115, 8919 (2018).
44. S. Zhao, Y. Osetsky, G.M. Stocks, and Y. Zhang, *NPJ Comput. Mater.* 5, 13 (2019).
45. S. Xu, E. Hwang, W.-R. Jian, Y. Su, and I.J. Beyerlein, *Intermetallics* 124, 106844 https://doi.org/10.1016/j.intermet.2020.106844 (2020).
46. S. Xu, S.Z. Chavoshi, and Y. Su, *Comput. Mater. Sci.* 202, 110942 (2022).

47. S. Xu, W.-R. Jian, and I.J. Beyerlein, *APL Mater.* 10, 111107 (2022).
48. Y.-J. Hu, A. Sundar, S. Ogata, and L. Qi, *Acta Mater.* 210, 116800 https://doi.org/10.1016/j.actamat.2021.116800 (2021).