Enhancing High-Fidelity Neural Network Potentials through Low-Fidelity Sampling

Gang Seob Jung[1†]

[1]Computational Sciences and Engineering Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

†Email: jungg@ornl.gov

Abstract: The efficacy of neural network potentials (NNPs) critically depends on the quality of the configurational datasets used for training. Prior research using empirical potentials has shown that well-selected liquid-solid transitional configurations of a metallic system can be translated to other metallic systems. This study demonstrates that such validated configurations can be relabeled using density functional theory (DFT) calculations, thereby enhancing the development of high-fidelity NNPs. Training strategies and sampling approaches are efficiently assessed using empirical potentials and subsequently relabeled via DFT in a highly parallelized fashion for high-fidelity NNP training. Our results reveal that relying solely on energy and force for NNP training is inadequate to prevent overfitting, highlighting the necessity of incorporating stress terms into the loss functions. To optimize training involving force and stress terms, we propose employing transfer learning to fine-tune the weights, ensuring the potential surface is smooth for these quantities composed of energy derivatives. This approach markedly improves the accuracy of elastic constants derived from simulations in both empirical potential-based NNP and relabeled DFT-based NNP. Overall, this study offers significant insights into leveraging empirical potentials to expedite the development of reliable and robust NNPs at the DFT level.

Keywords: Neural Network Potential, Empirical Potential, Configurational Sampling, Transfer Learning, Active Learning, Multicanonical ensemble

1. Introduction

Machine Learning (ML) has significantly propelled advancements across numerous scientific and engineering disciplines.[1-6] Utilizing ML in computational models for materials screening can expedite the development and design of new materials, such as alloys, pharmaceuticals, and polymers or proteins.[7-9] While numerous hurdles remain in transitioning from computational discovery to the actual synthesis and application of these materials, the initial discovery phase remains a crucial step.[10,11] Recent advances in physics-based computational atomistic modeling and simulations with techniques provide a novel avenue for exploring hypothetical materials to narrow the design space for materials with desired target properties.[12,13]

Recent advancements in machine learning-based interatomic potentials (MLIPs) have shown that they can predict potential energy surfaces (PESs) for atomic configurations with accuracy on par with ab initio electronic structure methods yet operate several orders of magnitude faster. [14-17] Neural network potentials (NNPs), a subset of MLIPs, employ neural networks to model interatomic interaction energies.[18] These potentials have rapidly emerged due to their flexibility, accuracy, and efficiency.

Reliable and robust NNPs can advance the materials discovery and screening more efficiently. However, NNPs, like other MLIPs, usually perform poorly outside their training domain and typically fail to predict unseen configurations without appropriate data. The active learning[19] can alleviate the problem by improving the accuracy and exploring new structural data by combining NNPs with enhanced sampling methods.[20] However, continuously acquiring more data and re-training still incurs significant computational costs. Therefore, it is highly desirable to efficiently sample configurations cover wide range of chemical and physical states such as phases, temperature, pressure, and strain.

The Multiorder-Multithermal (MOMT) ensemble molecular dynamics, an advanced variant of the multicanonical ensemble method,[21-23] is capable of sampling potential transition states between solid and liquid phases while accurately estimating the density of states.[24] However, such enhanced sampling typically necessitates extended simulation durations, which are impractical for density functional theory (DFT) calculations due to their high computational demands. A previous study has shown that configurations sampled using the MOMT ensemble can effectively encompass a variety of temperature states and different crystal structures through empirical potentials. Furthermore, the study indicates that carefully chosen configurations can be applicable to different types of metals.[25]

This study investigates the use of empirical potentials (low-fidelity) to develop neural network potentials (NNPs) that achieve the accuracy of density functional theory (DFT) calculations (high-fidelity). While the concept appears straightforward and feasible, it has not yet been successfully demonstrated in previous research. Notably, there are existing methods that employ transfer learning to elevate from DFT calculations to more complex computational levels.[26] However, these approaches often fall short when it comes to incorporating more intricate configurations, underlining that DFT calculations alone may not provide comprehensive coverage with enhanced sampling. Consequently, there is a pressing need to develop a methodology that efficiently derives high-quality configurations for high-fidelity models using sampling at the empirical potential level.

2

Figure 1 illustrates the comprehensive workflow used to enhance NNPs by combining low-fidelity data sampling with relabeled high-fidelity data. The process starts with collecting a substantial amount of low-fidelity data using empirical potentials, in this study, specifically employing the EAM potential for the nickel FCC system.[27,28] By leveraging techniques such as the MOMT ensemble molecular dynamics, this dataset encompasses a wide spectrum of states essential for capturing the diverse behaviors of materials. This extensive data is then refined through active learning and data distillation processes to effectively train NNPs. These NNPs are thoroughly evaluated for their predictions of energy, force, and stress, alongside essential physical properties like elastic constants derived from simulations. Following this, selected data is relabeled through high-fidelity DFT calculations to obtain DFT-based NNP. In our study, we identified limitations with both DFT and EAM-based NNPs in predicting shear modulus from the initial dataset, leading us to design specific new sampling strategies aimed at enhancing data crucial for calculating this modulus and verifying its effectiveness using EAM. This framework allows for adaptation based on different applications and sampling strategies through empirical potentials first, facilitating continuous improvements in NNP performance within this iterative cycle.

2. Method
2.1 Data and Sampling
2.1.1 Multiorder-Multithermal Ensemble Molecular Dynamics
The dataset was created in a prior study[25] using MOMT ensemble molecular dynamics (MD) with the Wang-Landau algorithm. [29] The EAM potential[27] for nickel was employed in these simulations. MOMT MD simulations were conducted with 32 and 108 nickel atoms, utilizing order parameters defined by reciprocal vectors.[24] This technique effectively samples both the liquid and solid phases of silicon and MgO.[23,30] A notable advantage of this method is its ability to enhance random-walk sampling of energy and defined order parameters. For a comprehensive explanation of the methods, please refer to the previous study.[24] The number of total data is 20,000 for each 32 and 108 atoms system, sampled through MOMT MD simulations. We only utilized the last 2,000 data from 32 atoms system for the initial training and fine-tuning. The other 18,000 data from the 32 atoms system and 20,000 data from the 108 atoms systems are utilized as a test set to evaluate the generalization of EAM-based NNPs.

2.1.2 Molecular Dynamics Simulation for Uniaxial and Shear Deformation
The molecular dynamics simulations were conducted using the LAMMPS package.[31] To obtain configurations related to uniaxial and shear deformations for calculating elastic constants, we specifically designed this sampling approach. We found that the 32-atom system provided sufficient coverage of the configurations, so we retained it. Initially, the structure was relaxed through energy minimization. We then made only the *lx* (uniaxial component) and *xy* (shear component) dimensions deformable in the cell. The system was further relaxed at 300K and sampled configurations every 100 steps at 300K over 200 ps with a 1 fs timestep, resulting in a total of 2,000 configurations. This sampling was designed to improve the elastic constants and different from MOMT the section 2.1.1, which covers both solid and liquid phases.

2.1.3 Data Distillation through Active Learning
Uncertainty quantification (UQ) of the NNPs is a crucial aspect of active learning, as it helps identify valuable data that are likely to be informative and worth labeling with new calculations. In this study, we use an ensemble-based approach,[32] employing the same NNP

structure but different training and validation sets for each model.[25] The energy predictions from each model are used to estimate the standard deviation, serving as a measure of UQ. This process was only applied to the 2,000 configurations sampled from the new sampling design described in 2.1.2. Based on the UQ from five models trained through the MOMT data set (Section 2.1.1), we trained the NNP with five models. For each model, the atomic energy predictions were obtained, and the standard deviation of these predictions was used as the UQ measure. Configurations with atomic UQ values greater than $\mu_{UQ} + 3\sigma_{UQ}$ were selected for further training. About 500 configurations were selected based on the defined criterion.

2.1.4 Density Functional Theory

The selected configurations were relabeled through DFT calculations using Quantum Espresso (QE).[33,34] The pseudopotentials library was employed,[35] utilizing the projector augmented-wave (PAW) method[36] with the Perdew-Burke-Ernzerhof (PBE) functional.[37] An energy cutoff of 50 Ry was set for the plane waves. The Brillouin zone was sampled using 3x3x3 Monkhorst-Pack grids for the 32-atom system. The convergence threshold for self-consistency was set to 1e-8. Gaussian smearing was applied with a Gaussian spreading of 0.022 Ry. We utilized ASE library[38] to collect the energy, forces, and stress term from QE outputs, and the smearing contribution in the total energy from the output was subtracted.

2.2 Neural Network Potential

2.2.1 Neural Network Structure

We employed the ANI-type neural network potential (NNP) using the TorchANI library.[39] The neural network structure used in this study is detailed in Table 1. We applied the Gaussian error linear unit (GELU) activation function, as it has been shown to improve the stress-strain behavior of graphene. The atomic environment vectors (AEV), also known as symmetry functions,[40] capture the atomic environment features for the neural network input. We adopted the AEV parameters from the ANI-2x model,[41] with the exception of using a longer radius cutoff of 6.9 Å and additional Gaussian centers.

For the initial training, 80% of the data was used to train the model, while 20% was reserved for validation, with a mini-batch size of 64. The data was shuffled during loading.

The loss function is defined as follows:

$$Loss = \frac{\alpha}{N_{data}}\sum\frac{(E_{NNP}-E_{ref})^2}{\sqrt{N_{atom}}} + \frac{\beta}{N_{data}}\sum\frac{(\vec{F}_{NNP}-\vec{F}_{ref})^2}{N_{atom}} + \frac{\gamma}{N_{data}}\sum(V\overrightarrow{\sigma_{NNP}} - V\overrightarrow{\sigma_{ref}})^2$$

(3)

,where E, F, V, and σ are energy, forces, volume, and stress, respectively. α , β, and γ control the contributions of the defined loss from energy, force, and stress.

2.2.2 Initial Training

The maximum number of epochs was set to 400, and we selected the best parameters based on the root mean square error (RMSE) of energy in the validation set during training. For the initial training, the hyperparameters α, β, and γ were set to 1.0, 0.1, and 0.025, respectively. When training the model without the stress term, γ was set to zero.

The Adam optimizer, with weight decay for the weights[42,43] (decays of 1e-5 and 1e-6 for the two hidden layers, respectively, and default values for others), was used alongside stochastic gradient descent (SGD) [46] for the biases, with a learning rate of 1e-4. The weights were

initialized using Kaiming initialization[47] with a normal distribution, and the initial values of the biases were set to zero.

A learning rate scheduler ("ReduceLROnPlateau" in PyTorch) was applied for both Adam and SGD optimizers, with a factor of 0.5, patience of 100, and a threshold of 0. However, the learning rate did not change significantly during the 400 epochs.

2.2.3 Transfer Learning for Fine-Tuning

The trained model can be further refined using a smaller learning rate and transfer learning, which can improve the accuracy of energy, force, and stress terms. Although it is possible to design training with a much larger number of epochs and control the contribution weights using an optimization scheduler, we propose transfer learning for the fine-tuning process. After the initial training described in Section 2.2.2, the model was further trained with the input and output layer weights and biases fixed.

In this phase, α, β, and γ were adjusted to 0.2, 0.4, and 0.4, respectively. When the stress term was excluded from the loss function, γ was set to zero and β to 0.8. The goal of these adjustments was to increase the contribution from the derivatives of energy (forces and stress) to fine-tune the model while fixing feature extraction from the symmetry function and the output layer to avoid overfitting to the derivative terms. The scheduler plays a more crucial role in this stage than in the initial training. The learning rates for the inner layers were set to 1e-4 for weights and 1e-5 for biases. A learning rate scheduler with a factor of 0.8, patience of 50, and a threshold of 0 was used for 1,000 epochs and select the best model based on the RMSE of the stress term.

We found that this approach maintained generalization, even when using a combined training and validation set from the initial training phase as the training and validation set during this transfer learning. This method is particularly effective with smaller datasets because it leverages the initial model's generalization capabilities while refining its accuracy with additional focused training. By keeping the feature extraction layers fixed, the model avoids overfitting to the data, ensuring robust performance across different configurations. The validation of the training process was conducted using the entire dataset, including configurations of 32 atoms and 108 atoms obtained from MOMT ensemble MD with the EAM potential.

2.3 Elastic Constants

To evaluate the NNPs, we performed structural optimization and computed the elastic constants using LAMMPS. For both the EAM and NNP, we prepared a 32-atom system and performed energy minimization. To avoid starting from a pre-relaxed structure, we expanded the FCC structure by 2-3%, resulting in more than 100 steps for structural optimization. Based on the obtained lattice parameters, the elastic constants were then measured assuming orthotropic symmetry, calculating C_{11}, C_{22}, C_{33}, C_{12}, C_{13}, C_{23}, C_{44}, C_{55}, and C_{66}.[44]

For the DFT calculations, we prepared an 8-atom system with increased K-sampling to 6x6x6. To ensure consistency in calculating elastic constants, we used a previously developed interface[20] between Python and LAMMPS to employ Quantum Espresso (QE) through the ASE library, allowing us to compare the elastic constants calculated with LAMMPS using the same process.[45] The deformation strain was set to ±1.0%.

We report only C_{11}(which is similar to C_{22} and C_{33}), C_{12} (similar to C_{13} and C_{23}), and C_{44} (similar to C_{55} and C_{66}), as other components were close to zero. The similarity of some elastic constant values is due to the inherent symmetry in the crystal structure, leading to equivalent elastic responses along different directions.

3. Results and Discussions

We selected the last 2,000 data points from a full dataset of 20,000 configurations, which were sampled using empirical potentials and the Multiorder-Multithermal (MOMT) ensemble molecular dynamics. The MOMT ensemble MD facilitates enhanced random-walk sampling across an expanded order-enthalpy space, increasing in diversity with continued sampling. Notably, the previous research initiated data distillation through active learning using the initial 2,000 data points.[25] In contrast, our findings indicate that the last 2,000 data points sampled offer superior performance than the initial 2,000, highlighting the effectiveness of late-stage data in capturing the various configurations of the system.

Therefore, we trained the selected 2,000 configurational data points to establish our initial model, integrating the stress term in the loss function to improve the accuracy and robustness of our predictions (refer to Methods section for details). Utilizing an ensemble approach, we developed five distinct models using varied combinations of training and validation sets. We then selected the best-performing model based on its force accuracy, a crucial indicator for mitigating overfitting, as highlighted in previous studies.[46] The initial results, depicted in Figure 2a, demonstrated that the predictions for energy, force, and stress were closely aligned with the reference data, affirming that the NNPs adeptly captured the essential physics of the metallic systems. However, despite these promising indications, the subsequent calculation of elastic constants from the trained NNPs did not yield satisfactory accuracy.

To address this, we implemented a transfer learning strategy aimed at fine-tuning the model to emphasize force and stress terms more effectively, while the weights of the input and output layers were fixed, focusing the training efforts on the hidden layers (refer to Methods for detailed setup). In this phase, the contributions of force and stress errors in the loss function were increased to 0.4 each from 0.1, and the contribution from energy was reduced to 0.2 from 0.9. We initialized the weights using those from the previously trained model, and the starting learning rate was set to be ten times smaller than that used in the initial training. During this process, we continuously adjusted the learning rate based on a predefined schedule, selecting the best model based on its performance on stress-related terms. This approach was designed to refine the estimation of the potential surface, enhancing its accuracy and smoothness by strategically adjusting learning rates and incorporating derivative information, such as force and stress, during the fine-tuning process.

Figure 2b demonstrates marked improvements in model performance with reduced MAE errors of all energy, force, and stress. During the fine-tuning phase, we maintained the hypothesis that generalization would remain robust, even when utilizing the full dataset, because the input and output layers were fixed. However, this assumption requires further verification. Here, the benefits of using a low-fidelity model, specifically the EAM potential, become evident. It's important to note that our 2,000 data points represent only 10% of the total sampled data. We can further validate this hypothesis using the remaining 90% of the data. If the outcomes are consistent, it would suggest that the designed training/finetuning

6

processes are sufficiently robust with just 2,000 data points, based on the EAM results. This implies that using re-labeled data from DFT calculations on these points could be considered reliable.

Figure 3 illustrates the performance of the trained NNP using re-labeled data from DFT calculations. We observed a similar improvement in accuracy through the fine-tuning process. While the MAEs are slightly higher compared to those from EAM, this outcome is expected; EAM values are derived from analytic potentials within a specified cutoff, whereas DFT calculations account for electronic density effects over distances extending beyond these cutoffs. Nonetheless, the close alignment in the results confirms that the radius cutoff approximation remains effective in capturing the behavior of metallic systems.

Based on these results, we aimed to evaluate the impact of including stress terms in the loss functions. To do this, we replicated the training and fine-tuning processes, this time omitting the stress term from the loss function. In compensation, we increased the weight of the force contribution to 0.9 from 0.1 and decreased the energy contribution from 1.0 to 0.1, selecting the model that performed best in terms of force error. The results, displayed in Figure 4 after fine-tuning, reveal a significant increase in the MAE of stress terms—approximately 2 to 3 times larger—while force errors are substantially reduced. This outcome underscores that excluding stress terms can lead to overfitting of the force terms, which is likely a primary cause of instability in ML-based potentials during MD simulations, as was clearly demonstrated in a previous study.[47]

We conducted a detailed comparison of the lattice parameters, elastic constants, and mechanical properties for models trained using both EAM and DFT, as presented in Tables 2 and 3. Both models displayed similar trends; however, the accuracy of the shear modulus (G) was notably poorer compared to other elastic constants. Additionally, an evaluation of the NNP_{EAM} model using the remaining 90% of the data, as shown in Figure 5a, revealed clear issues with some data points of the stress term near zero. This observation motivated an additional design of our sampling strategy, which is depicted in Figure 5b.

Since our MOMT sampling was limited to cubic cells at high temperatures—the reference temperature being 2000K, though the MOMT ensemble is capable of sampling across a wider temperature range—we opted to fix the cell at the equilibrium lattice constant and conducted conventional MD sampling at a lower temperature (Methods). During this sampling, we permitted deformations only along the l_x and xz axes. Following this, we collected 2,000 data points from this new additional sampling using EAM. From these, we then selected approximately 500 configurations through data distillation, employing a high-uncertainty quantification strategy that utilized the atomic energy predictions from five previously trained models.[25] Additionally, only these 500 configurations were re-labeled using DFT calculations.

Figure 5b illustrates the performance of the initially fine-tuned models using the additional data. While the energy and force predictions appear satisfactory, the stress predictions for both EAM and DFT data are inadequate. This discrepancy further underscores that the accuracy of energy and force alone is not enough to uncover the limitations of the models. Consequently, we incorporated these 500 new configurations into the dataset and retrained/fine-tuned the models using both EAM and DFT data. The enhanced performance is depicted in Figure 6. Here, Data1 represents the initial dataset, while Data2 consists of the additional data. The previously observed anomalies in the stress term with Data2 have been completely resolved, and the overall accuracy for Data1 has also shown improvement.

7

The elastic constants for the models have also shown significant improvement (Tables 4 and 5), validating our hypothesis regarding the models' limitations and demonstrating the effectiveness of the newly designed sampling approach (Also see Tables S1-S4 to compare). It is important to note that this sampling strategy was not our initial attempt; we tested other sampling and ultimately selected the most effective one based on the EAM potentials. Once it met our objectives, we then re-labeled these selected configurations through DFT calculations.

Finally, we confirmed the generalization of the trained NNP based on EAM by evaluating it with the other 90% of the MOMT-sampled configurations, as shown in Figure 7a. Compared to the earlier discrepancies observed in the stress term in Figure 4a, the mismatches near the lower stress values have been resolved. One significant advantage of using empirical potentials is their computational speed. We conducted tests using a MOMT-sampled configuration with 108 atoms for the nickel FCC structure under very similar conditions. The performance of the model was assessed using data that was never included in the training phase. As depicted in Figure 7b, the performance was commendable, confirming that the models are well-generalized.

The results effectively showcase how sampled and selected configurations using empirical potentials can be directly utilized for DFT calculations, leading to significant improvements in NNPs derived from both EAM and DFT data. The methodology applied here, particularly with the FCC nickel crystal, would be effective for examining the transition temperatures between liquid and crystal phases and assessing elastic properties. However, this model may not adequately capture phenomena such as dislocation dynamics, plastic deformation, or mechanical failure, as configurations pertinent to these processes might be excluded from the training set. Typically, these properties demand larger system sizes and extended timescales.

Nonetheless, employing EAM allows for the development of efficient and straightforward methods to sample such configurations for NNP training, as demonstrated in this study. Additionally, like other empirical potentials, EAM calculates atomic energy directly, providing a basis for comparing trained atomic energies. Instead of relying solely on uncertainty quantification, the difference in atomic energy could serve as a more effective indicator for selecting new configurations. While this study focuses on a single metal type, the approach has the potential to expedite the development of NNPs for more complex systems, including binary and ternary alloys as well as metal oxide-type materials.

4. Conclusion

In summary, this study has demonstrated the robustness and versatility of neural network potentials (NNPs) trained on empirical potential data and further refined through density functional theory (DFT) calculations. Our findings underscore the critical importance of including stress terms in the loss function to avoid overfitting and enhance model stability during molecular dynamics simulations. Through strategic data distillation and fine-tuning processes, we significantly improved the predictive accuracy of NNPs, particularly in capturing the mechanical properties of metallic systems, as evidenced by the elastic constants. The successful application of these models to a large percentage of additional MOMT-sampled data further validated the generalization capability of our approach. By optimizing sampling strategies and leveraging both low- and high-fidelity data, we have developed NNPs that not only meet computational efficiency demands but also maintain high accuracy, making them valuable tools for advancing materials science research and development.

5. Acknowledgements

Supplementary Information

ORCID

Gang Seob Jung: 0000-0002-8047-6505

CODE/DATA AVAILABILITY

Sample inputs, data, and source codes will be available
https://github.com/gsjung0419/Low2High

AUTHOR CONTRIBUTION STATEMENTS

G.S.J. conceived the idea, developed codes, and performed simulations, training, and evaluations, wrote, and edited the draft.

6. REFERENCES

1   Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589, doi:10.1038/s41586-021-03819-2 (2021).

2   Schmidt, J., Marques, M. R. G., Botti, S. & Marques, M. A. L. Recent advances and applications of machine learning in solid-state materials science. *npj Computational Materials* **5**, 83, doi:10.1038/s41524-019-0221-0 (2019).

3   Butler, K. T., Davies, D. W., Cartwright, H., Isayev, O. & Walsh, A. Machine learning for molecular and materials science. *Nature* **559**, 547-555, doi:10.1038/s41586-018-0337-2 (2018).

4   Pfau, D., Spencer, J. S., Matthews, A. G. D. G. & Foulkes, W. M. C. Ab initio solution of the many-electron Schr\"odinger equation with deep neural networks. *Physical Review Research* **2**, 033429, doi:10.1103/PhysRevResearch.2.033429 (2020).

5   Vamathevan, J. *et al.* Applications of machine learning in drug discovery and development. *Nature Reviews Drug Discovery* **18**, 463-477, doi:10.1038/s41573-019-0024-5 (2019).

6   Wang, H. *et al.* Scientific discovery in the age of artificial intelligence. *Nature* **620**, 47-60, doi:10.1038/s41586-023-06221-2 (2023).

7   Hu, M. *et al.* Recent applications of machine learning in alloy design: A review. *Materials Science and Engineering: R: Reports* **155**, 100746, doi:https://doi.org/10.1016/j.mser.2023.100746 (2023).

8   Meyers, J., Fabian, B. & Brown, N. De novo molecular design and generative models. *Drug Discovery Today* **26**, 2707-2715, doi:https://doi.org/10.1016/j.drudis.2021.05.019 (2021).

9   Yan, C. & Li, G. The Rise of Machine Learning in Polymer Discovery. *Advanced Intelligent Systems* **5**, 2200243, doi:https://doi.org/10.1002/aisy.202200243 (2023).

10   Dara, S., Dhamercherla, S., Jadav, S. S., Babu, C. H. M. & Ahsan, M. J. Machine Learning in Drug Discovery: A Review. *Artificial Intelligence Review* **55**, 1947-1999, doi:10.1007/s10462-021-10058-4 (2022).

11   Tabor, D. P. *et al.* Accelerating the discovery of materials for clean energy in the era of smart automation. *Nature Reviews Materials* **3**, 5-20, doi:10.1038/s41578-018-0005-z (2018).

12   Alberi, K. *et al.* The 2019 materials by design roadmap. *Journal of Physics D: Applied Physics* **52**, 013001, doi:10.1088/1361-6463/aad926 (2018).

13   Chen, A., Zhang, X. & Zhou, Z. Machine learning: Accelerating materials development for energy storage and conversion. *InfoMat* **2**, 553-576, doi:https://doi.org/10.1002/inf2.12094 (2020).

14   Behler, J. & Csányi, G. Machine learning potentials for extended systems: a perspective. *The European Physical Journal B* **94**, 142, doi:10.1140/epjb/s10051-021-00156-1 (2021).

15   Kocer, E., Ko, T. W. & Behler, J. Neural Network Potentials: A Concise Overview of Methods. *Annual review of physical chemistry* (2022).

16   Pinheiro, M., Ge, F., Ferré, N., Dral, P. O. & Barbatti, M. Choosing the right molecular machine learning potential. *Chemical Science* **12**, 14396-14413, doi:10.1039/D1SC03564A (2021).

17   Unke, O. T. *et al.* Machine Learning Force Fields. *Chemical Reviews* **121**, 10142-10186, doi:10.1021/acs.chemrev.0c01111 (2021).

18   Lorenz, S., Groß, A. & Scheffler, M. Representing high-dimensional potential-energy surfaces for reactions at surfaces by neural networks. *Chemical Physics Letters* **395**, 210-215 (2004).

19   Settles, B. Active learning literature survey. (2009).

20   Jung, G. S., Choi, J. Y. & Lee, S. M. Active learning of neural network potentials for rare events. *Digital Discovery*, doi:10.1039/D3DD00216K (2024).

21   Nakajima, N., Nakamura, H. & Kidera, A. Multicanonical Ensemble Generated by Molecular Dynamics Simulation for Enhanced Conformational Sampling of Peptides. *The Journal of Physical Chemistry B* **101**, 817-824, doi:10.1021/jp962142e (1997).

22   Okumura, H. & Okamoto, Y. Multibaric–multithermal ensemble molecular dynamics simulations. *Journal of Computational Chemistry* **27**, 379-395, doi:https://doi.org/10.1002/jcc.20351 (2006).

23   Yoshimoto, Y. Extended multicanonical method combined with thermodynamically optimized potential: Application to the liquid-crystal transition of silicon. *The Journal of Chemical Physics* **125**, 184103, doi:10.1063/1.2363987 (2006).

24   Jung, G. S., Yoshimoto, Y., Oh, K. J. & Tsuneyuki, S. Extended Ensemble Molecular Dynamics for Thermodynamics of Phases. *arXiv preprint arXiv:2308.08098* (2023).

25   Jung, G. S., Lee, S. & Choi, J. Y. Data Distillation for Neural Network Potentials toward Foundational Dataset. *arXiv preprint arXiv:2311.05407* (2023).

26   Smith, J. S. *et al.* Approaching coupled cluster accuracy with a general-purpose neural network potential through transfer learning. *Nature Communications* **10**, 2903, doi:10.1038/s41467-019-10827-4 (2019).

27   Zhang, Y., Ashcraft, R., Mendelev, M., Wang, C. & Kelton, K. Experimental and molecular dynamics simulation study of structure of liquid and amorphous Ni62Nb38 alloy. *The Journal of chemical physics* **145**, 204505 (2016).

28   Mendelev, M. I., Kramer, M. J., Becker, C. A. & Asta, M. Analysis of semi-empirical interatomic potentials appropriate for simulation of crystalline and liquid Al and Cu. *Philosophical Magazine* **88**, 1723-1750, doi:10.1080/14786430802206482 (2008).

29   Wang, F. & Landau, D. P. Efficient, Multiple-Range Random Walk Algorithm to Calculate the Density of States. *Physical Review Letters* **86**, 2050-2053, doi:10.1103/PhysRevLett.86.2050 (2001).

30   Yoshimoto, Y. Melting of MgO studied using a multicanonical ensemble method combined with a first-principles calculation. *Journal of the Physical Society of Japan* **79**, 034602 (2010).

31   Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *Journal of Computational Physics* **117**, 1-19, doi:https://doi.org/10.1006/jcph.1995.1039 (1995).

32   Ying, Z., Jun, G. & Xuezhi, Y. in *2005 International Conference on Neural Networks and Brain.* 438-442.

33   Giannozzi, P. *et al.* QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *Journal of Physics: Condensed Matter* **21**, 395502, doi:10.1088/0953-8984/21/39/395502 (2009).

34   Giannozzi, P. *et al.* Advanced capabilities for materials modelling with Quantum ESPRESSO. *Journal of Physics: Condensed Matter* **29**, 465901, doi:10.1088/1361-648x/aa8f79 (2017).

35   Dal Corso, A. Pseudopotentials periodic table: From H to Pu. *Computational Materials Science* **95**, 337-350, doi:https://doi.org/10.1016/j.commatsci.2014.07.043 (2014).

36   Blöchl, P. E. Projector augmented-wave method. *Physical Review B* **50**, 17953-17979, doi:10.1103/PhysRevB.50.17953 (1994).

37   Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Physical Review Letters* **77**, 3865-3868, doi:10.1103/PhysRevLett.77.3865 (1996).

38   Hjorth Larsen, A. *et al.* The atomic simulation environment—a Python library for working with atoms. *Journal of Physics: Condensed Matter* **29**, 273002, doi:10.1088/1361-648X/aa680e (2017).

39   Gao, X., Ramezanghorbani, F., Isayev, O., Smith, J. S. & Roitberg, A. E. TorchANI: A Free and Open Source PyTorch-Based Deep Learning Implementation of the ANI Neural Network Potentials. *Journal of Chemical Information and Modeling* **60**, 3408-3415, doi:10.1021/acs.jcim.0c00451 (2020).

40   Behler, J. & Parrinello, M. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. *Physical Review Letters* **98**, 146401, doi:10.1103/PhysRevLett.98.146401 (2007).

41   Devereux, C. *et al.* Extending the Applicability of the ANI Deep Learning Molecular Potential to Sulfur and Halogens. *Journal of Chemical Theory and Computation* **16**, 4192-4202, doi:10.1021/acs.jctc.0c00121 (2020).

42   Kingma, D. P. & Ba, J. Adam: A Method for Stochastic Optimization. *arXiv [cs.LG]* (2017).

43   Loshchilov, I. & Hutter, F. Decoupled Weight Decay Regularization. *arXiv [cs.LG]* (2019).

44   Jung, G. S. & Buehler, M. J. Multiscale Mechanics of Triply Periodic Minimal Surfaces of Three-Dimensional Graphene Foams. *Nano Letters* **18**, 4845-4853, doi:10.1021/acs.nanolett.8b01431 (2018).

45   Jung, G. S., Irle, S. & Sumpter, B. G. Dynamic aspects of graphene deformation and fracture from approximate density functional theory. *Carbon* **190**, 183-193, doi:https://doi.org/10.1016/j.carbon.2022.01.002 (2022).

46   Jung, G. S., Myung, H. & Irle, S. Artificial neural network potentials for mechanics and fracture dynamics of two-dimensional crystals**. *Machine Learning: Science and Technology* **4**, 035001, doi:10.1088/2632-2153/accd45 (2023).

47   Fu, X. *et al.* Forces are not Enough: Benchmark and Critical Evaluation for Machine Learning Force Fields with Molecular Simulations. *arXiv preprint arXiv:2210.07237* (2022).

48   Hendrycks, D. & Gimpel, K. Gaussian Error Linear Units (GELUs). *arXiv [cs.LG]* (2020).

Table 1. The neural network structures for Ni in this study used the Gaussian error linear unit (GELU) activation function[48] to introduce non-linearity between the AEV-1st, 1st-2nd, and 2nd-3rd layers. The radius cutoff for the radial part was set to 6.9 Å.

NN Model	1st	2nd	3rd	Output (Energy)
Ni	224	192	160	1

Table 2. Elastic constants and the lattice parameter of FCC obtained from the initial NNP_{EAM} model.

	l_a (Å)	C_{11} (GPa)	C_{12} (GPa)	C_{44} (GPa)	K (GPa)	G (GPa)	v
EAM	3.516	214.4	158.9	124.9	177.4	27.7	0.426
NNP_{EAM}	3.511	219.2	202.6	133.3	208.1	8.3	0.48
Error (%)	-0.14	2.24	27.50	6.73	17.31	-70.04	12.68

Table 3. Elastic constants and the lattice parameter of FCC obtained from the initial NNP_{DFT} model.

	l_a (Å)	C_{11} (GPa)	C_{12} (GPa)	C_{44} (GPa)	K (GPa)	G (GPa)	v
DFT	3.511	284.1	188.3	131.2	220.4	48.0	0.399
NNP_{DFT}	3.522	250.8	193.7	129.4	212.8	28.5	0.436
Error (%)	0.31	-11.72	2.87	-1.37	-3.45	-40.63	9.27

Table 4. Elastic constants and the lattice parameter of FCC obtained from the final NNP_{EAM} model.

	l_a (Å)	C_{11} (GPa)	C_{12} (GPa)	C_{44} (GPa)	K (GPa)	G (GPa)	v
EAM	3.516	214.4	158.9	124.9	177.4	27.7	0.426
NNP_{EAM}	3.515	192.9	154.3	122.1	167.2	19.3	0.444
Error (%)	-0.03	-10.03	-2.89	-2.24	-5.75	-30.32	4.23

Table 5. Elastic constants and the lattice parameter of FCC obtained from the final NNP_{DFT} model.

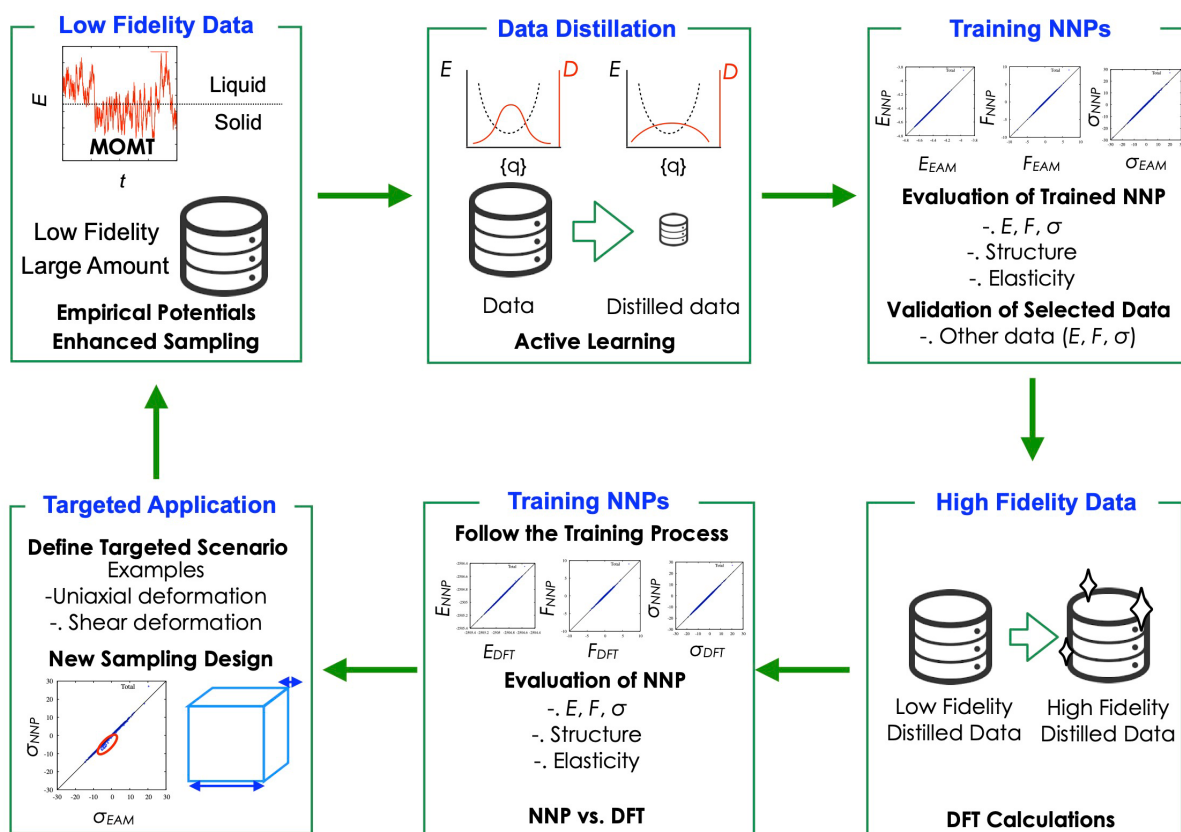|	l_a (Å)	C_{11} (GPa)	C_{12} (GPa)	C_{44} (GPa)	K (GPa)	G (GPa)	v
DFT	3.511	284.1	188.3	131.2	220.4	48.0	0.399
NNP_{DFT}	3.512	280.4	197.6	130.6	225.2	41.4	0.413
Error (%)	0.03	-1.30	4.89	-0.46	2.18	-13.75	3.51

Figure 1. Workflow for enhancing neural network potentials (NNPs) through data processing and training. The process begins with low-fidelity, large-quantity data from empirical potentials, which undergoes data distillation through active learning to refine and enhance its utility. This distilled data is then used for training NNPs, followed by an evaluation against high-fidelity DFT calculations. The workflow is based on the cycle of using targeted application scenarios for new sampling design, feeding into the training process, and validation to ensure the accuracy and robustness of the NNPs. The integration of low and high-fidelity data ensures a comprehensive approach to reliable and efficient development of high-fidelity NNPs for complex metallic systems.
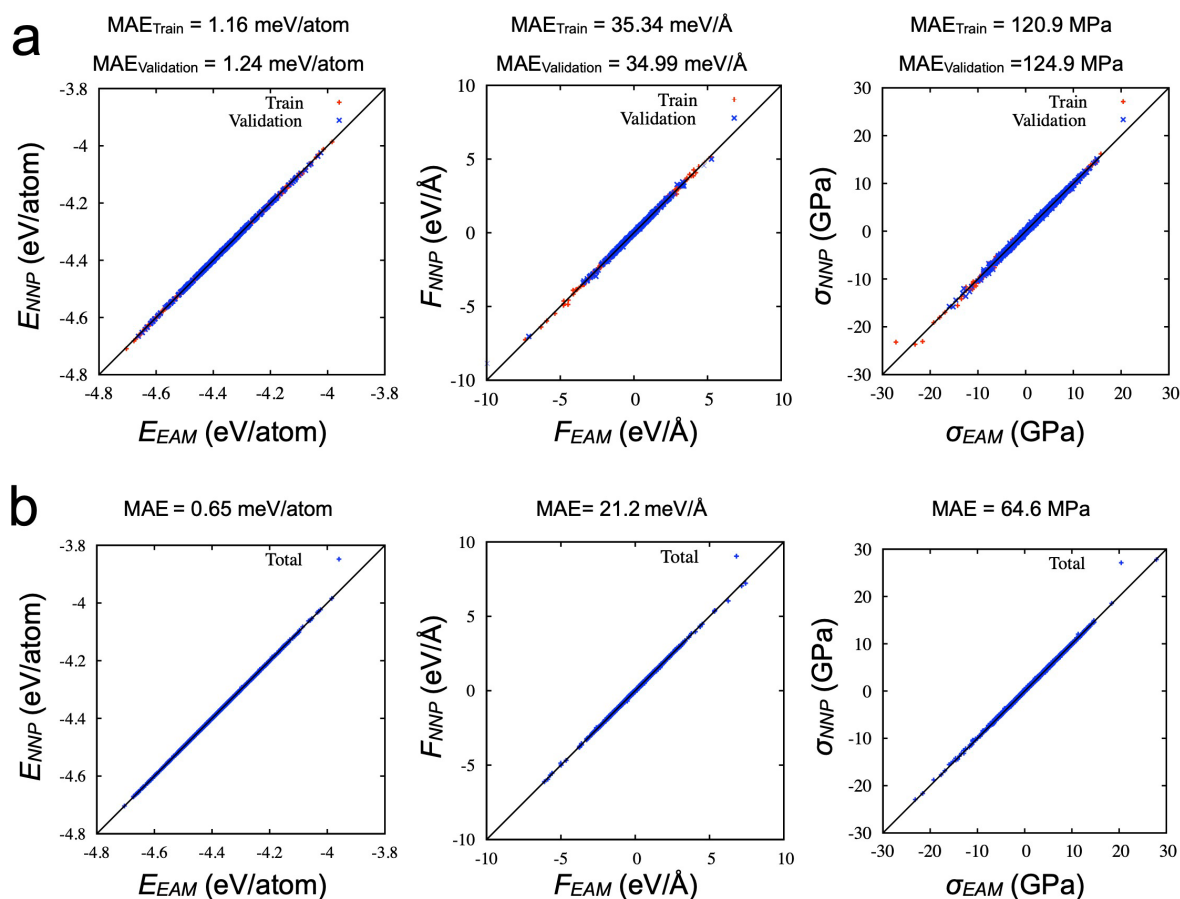
14

Figure 2. Comparative evaluation of NNPs trained by EAM potentials before and after transfer learning (fine-tuning). (a) presents the initial training results of NNPs, with scatter plots comparing the predicted energy per atom, forces, and stress against empirical potential estimates. Each plot includes mean absolute error (MAE) metrics for training and validation sets. See Table S1 for lattice parameters and elastic constants. (b) illustrates the outcomes after fine-tuning through transfer learning, where increased contributions of forces and stress in the loss function with a learning rate schedule, significantly enhance model performance, as shown by reduced MAEs and improved alignment in the total data set. These graphs underscore the effectiveness of transfer learning in refining the predictive capabilities of NNPs. See Table 2 for lattice parameters and elastic constants.
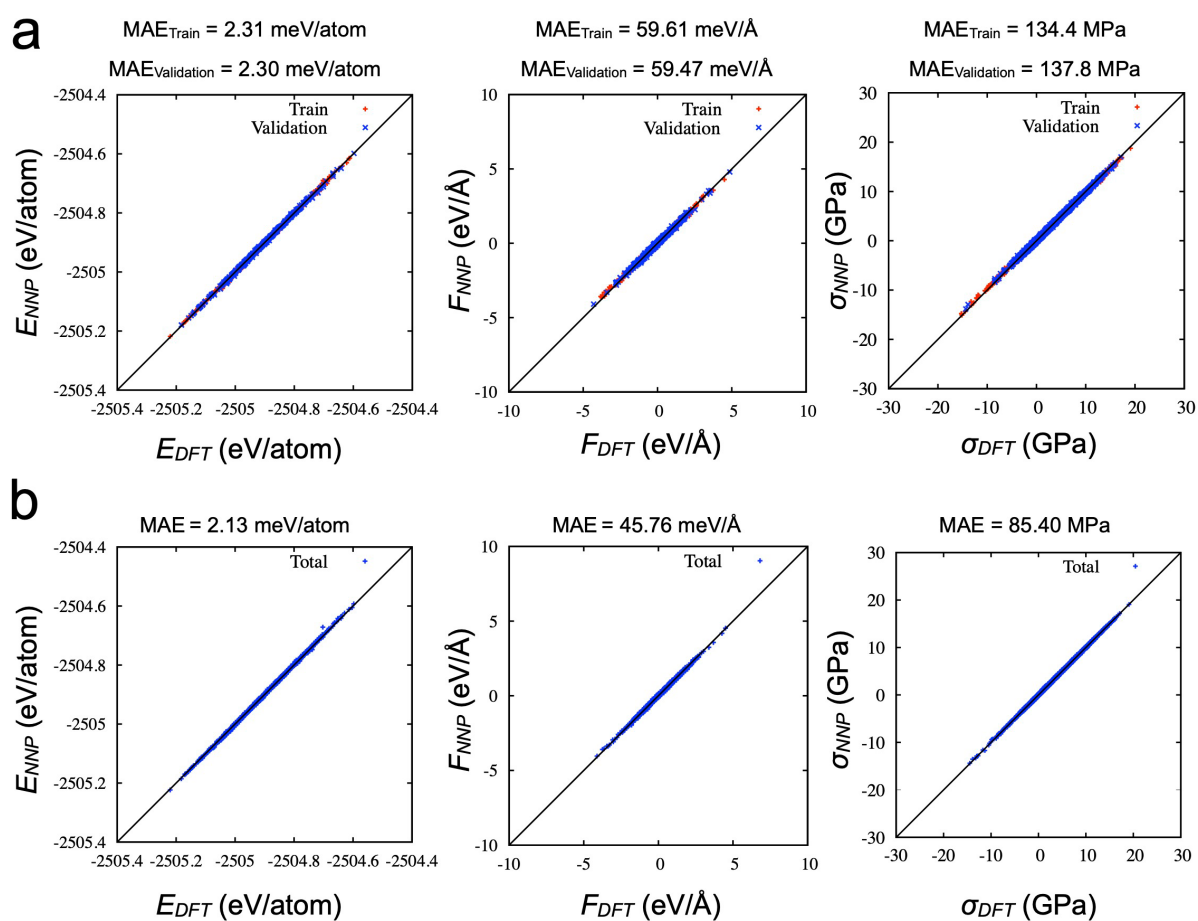
15

Figure 3. Validation of NNPs using DFT data before and after transfer learning. (a) displays the initial comparisons between DFT-calculated and NNP-predicted values for energy per atom, forces, and stress. See Table S2 for lattice parameters and elastic constants. (b) shows the results after the transfer learning (fine-tuning), applying the same strategy of increased contributions of forces and stress in the loss function, along with a learning rate schedule. This approach demonstrates consistent performance improvements, indicated by reduced MAEs and better alignment of the predicted values with high-fidelity DFT data, affirming the efficacy of transfer learning. See Table 3 for lattice parameters and elastic constants.
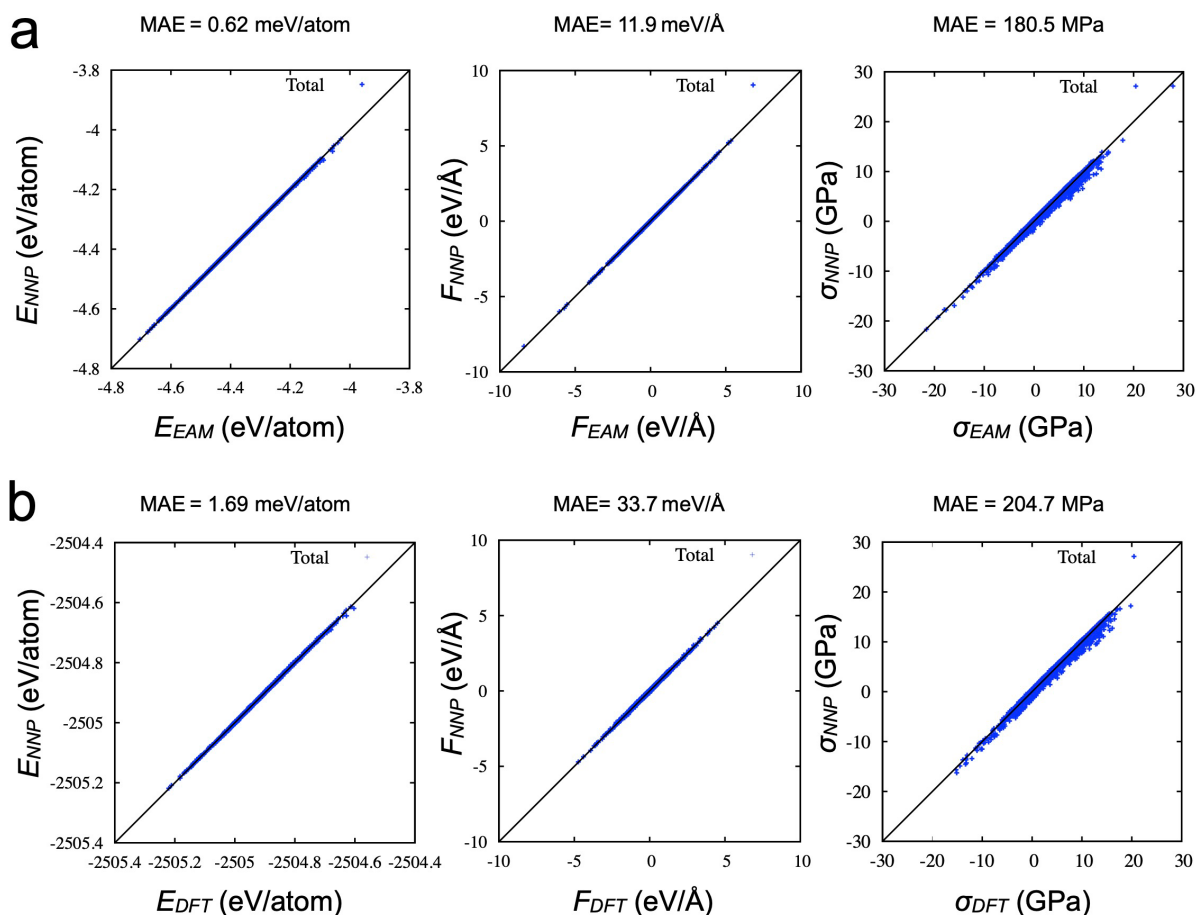
16

Figure 4. Performance evaluation of NNPs excluding the stress terms in the loss function during the initial training and transfer learning. (a) presents results from NNPs trained and fine-tuned without stress terms, using EAM potential for energy per atom, forces, and stress. (b) displays the corresponding results using DFT data, following the same training process. Both panels demonstrate the impact of excluding stress terms, with notably higher MAEs, particularly in stress predictions, illustrating that including stress terms in the loss function is crucial to prevent overfitting and ensure the accuracy and generalizability of NNPs. See Tables S3 and S4 for lattice parameters and elastic constants from the models.
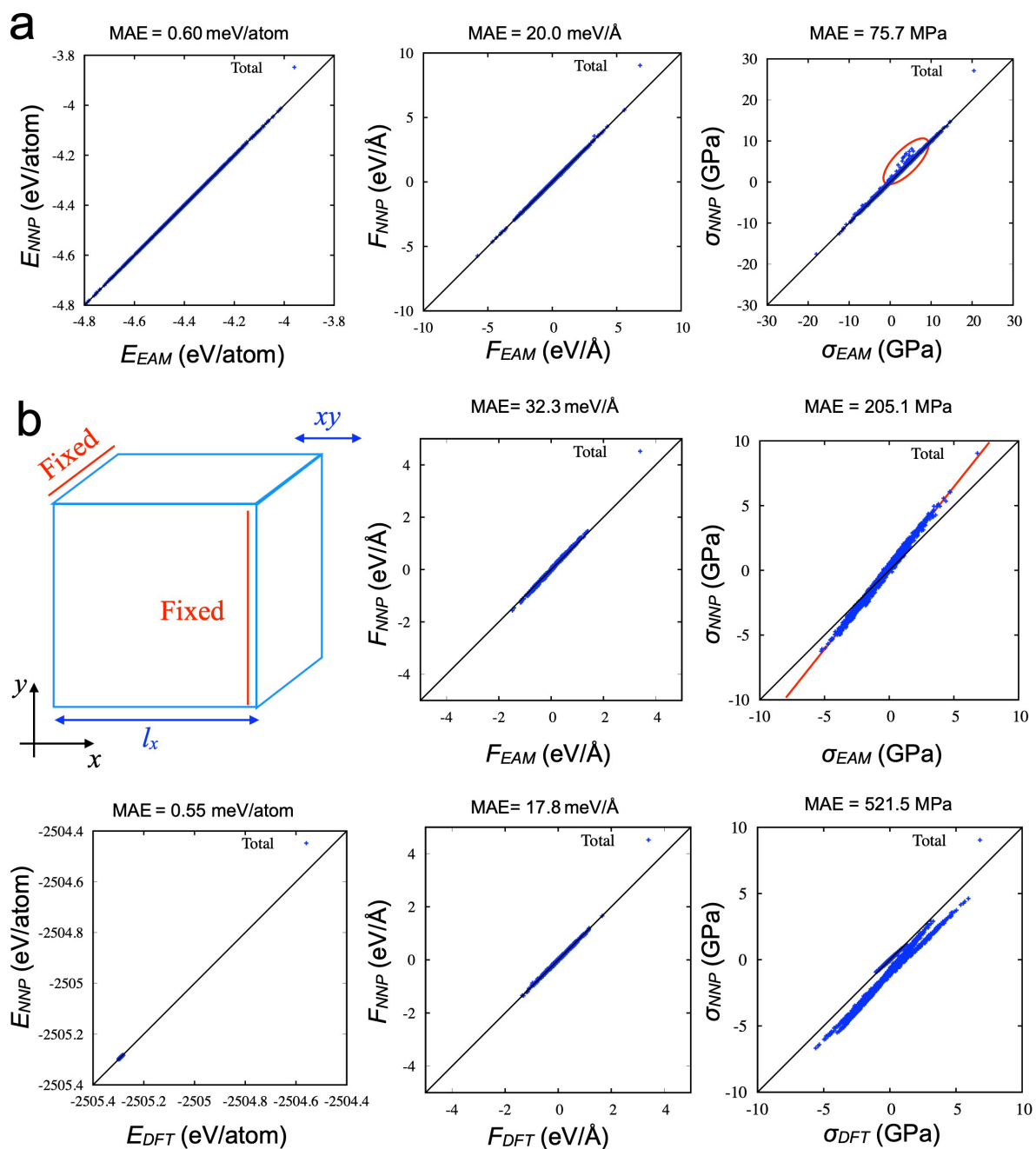
Figure 5. (a) Performance evaluation utilizes all available data (20,000 points) from the EAM potential, plotting the NNPs' predicted energy per atom, forces, and stress against EAM-calculated values. Notably, stress predictions show deviations, especially at lower values, indicating potential issues with model accuracy in these regions as confirmed by suboptimal elasticity calculations in Tables 2. (b) A new sampling design focused on shear and uniaxial deformation from energy-minimized structures, aimed at enhancing model training in these specific scenarios. The results reveal persistent challenges in accurately representing stress terms in both EAM and DFT frameworks, highlighting the need for further model optimization to address discrepancies in stress prediction across newly sampled configurations.
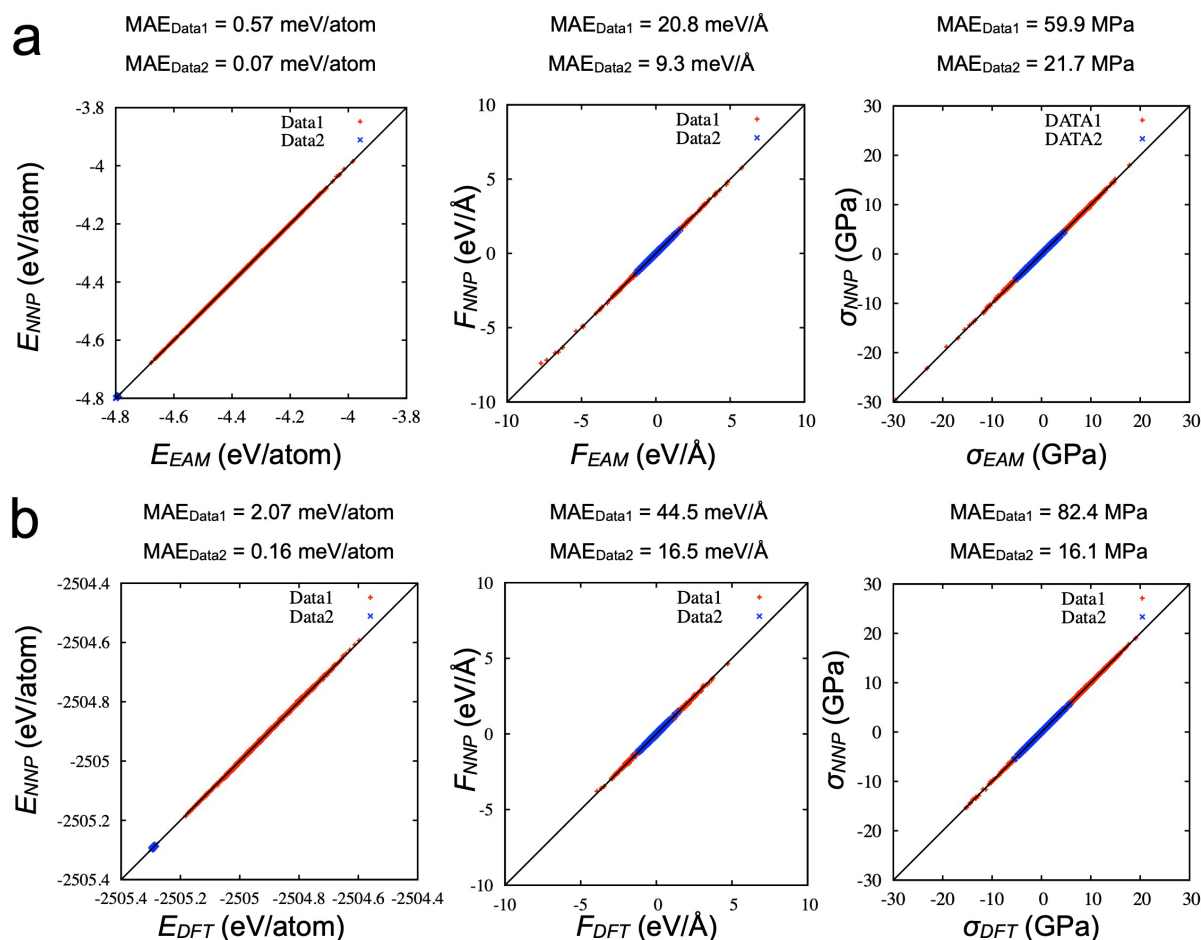
Figure 6. (a) compares NNPs' predictions against EAM data, showing two distinct datasets: Data1, the initial model dataset consisting of ~2000 points, and Data2, an additional ~500 points from newly designed sampling scenarios. (b) demonstrates a similar analysis using DFT data, where the inclusion of the new dataset (Data2) results in significant improvements in model performance, particularly evident in the reduced MAEs for Data2 from 521MPa (Figure 5) to 16.1MPa. This enhanced performance correlates with improved elasticity properties, detailed in Tables 4 and 5, illustrating the efficacy of integrating specifically designed samples through initial training and subsequent processes.
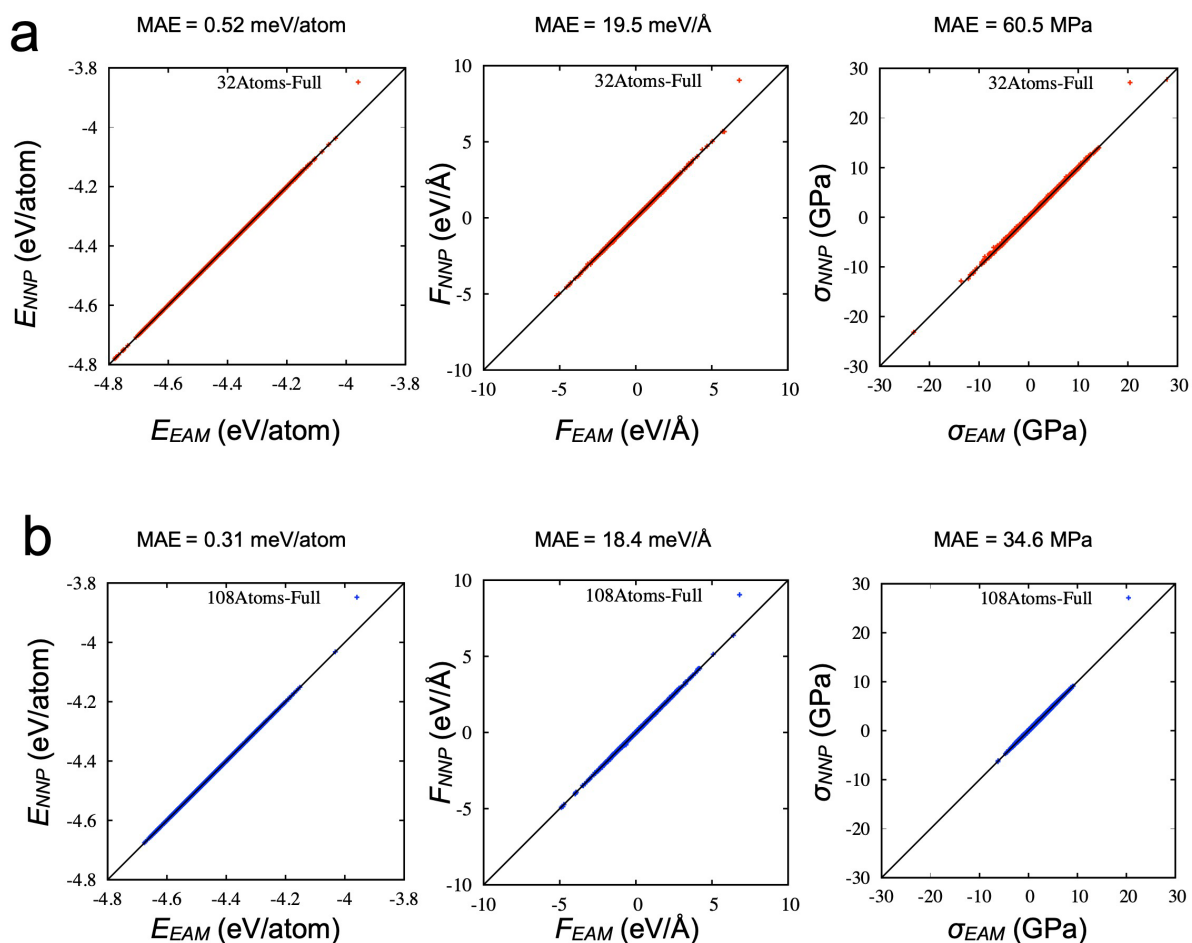
Figure 7. (a) illustrates the performance of the final NNP model trained using EAM potentials on the full dataset (20,000 points) comprising full 32 atoms. The model demonstrates strong predictive accuracy for energy, force, and stress, as indicated by low MAEs across all metrics. Also, a limitation observed in <mark>Figure 5a</mark> was resolved (b) extends the evaluation to a larger system of 108 atoms (never included in the training), showcasing that the NNP maintains high accuracy in predicting physical properties. Although these data were not derived from DFT, the consistency in performance across different system sizes suggests that similar transferability can be expected when applying the model to DFT-based systems.

Supplementary Information

Enhancing High-Fidelity Neural Network Potentials through Low-Fidelity Sampling

Gang Seob Jung[1†]

[1]Computational Sciences and Engineering Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

†Email: jungg@ornl.gov

Table S1. Elastic constants and the lattice parameter of FCC obtained from the initial NNP$_{EAM}$ model without fine-tuning.

	l_a (Å)	C_{11} (GPa)	C_{12} (GPa)	C_{44} (GPa)	K (GPa)	G (GPa)	v
EAM	3.516	214.4	158.9	124.9	177.4	27.7	0.426
NNP$_{EAM}$	3.534	184.7	129.4	126.1	147.8	27.7	0.412
Error (%)	0.51	-13.85	18.57	0.96	-16.69	0.0	-3.29

Table S2. Elastic constants and the lattice parameter of FCC obtained from the initial NNP$_{DFT}$ model without fine-tuning.

	l_a (Å)	C_{11} (GPa)	C_{12} (GPa)	C_{44} (GPa)	K (GPa)	G (GPa)	v
DFT	3.511	284.1	188.3	131.2	220.4	48.0	0.399
NNP$_{DFT}$	3.536	230.3	181.4	109.9	197.7	24.4	0.441
Error (%)	0.71	-18.94	-3.66	-16.31	-10.30	-49.19	10.53

Table S3. Elastic constants and the lattice parameter of FCC obtained from the initial NNP$_{EAM}$ model without stress contribution in the loss function after fine-tuning.

	l_a (Å)	C_{11} (GPa)	C_{12} (GPa)	C_{44} (GPa)	K (GPa)	G (GPa)	v
EAM	3.516	214.4	158.9	124.9	177.4	27.7	0.426
NNP$_{EAM}$	3.527	204.3	189.6	124.4	194.5	7.34	0.481
Error (%)	0.31	-4.71	19.32	-0.4	9.64	-73.5	12.91

Table S4. Elastic constants and the lattice parameter of FCC obtained from the initial NNP$_{DFT}$ model without stress contribution in the loss function after fine-tuning.

	l_a (Å)	C_{11} (GPa)	C_{12} (GPa)	C_{44} (GPa)	K (GPa)	G (GPa)	v
DFT	3.511	284.1	188.3	131.2	220.4	48.0	0.399
NNP$_{DFT}$	3.526	283.7	219.4	118.5	240.9	32.1	0.436
Error (%)	0.43	-0.14	16.52	-9.68	9.3	-33.13	9.27