Machine-Learned Interatomic Potential of Pure Crystalline Yttrium

Andrew Trepagnier[1]

Michael W. Hall School of Mechanical Engineering, Mississippi State University
Center for Advanced Vehicular Systems, Mississippi State University

Abstract

Pure crystalline Yttrium (Y) has been a material of great interest for solid-solution hardening of Magnesium (Mg) alloys due to its enhancement of Mg's high strength-to-weight ratio. Molecular dynamics simulations have been a cornerstone of material science research for systems like Mg-Y for characterizing the behavior of various configurations on the atomic level. However, one challenge of modern molecular dynamics is the inflexibility of conventional interatomic potential formulations such as Modified Embedded Atom Methods (MEAM) and Embedded Atom Methods (EAM) due to their limited number of characterizing parameters. The use of neural network architectures for interatomic potential fitting has recently become a computation strategy of great interest in material science due to its superior abilities to train over large, memory-intensive datasets in reasonable time and to tailor the number of characterizing parameters to the needs of the dataset. In this study, an interatomic potential of pure Y was trained using a machine-learned rapid artificial neural network (RANN). The performance of the RANN model was characterized by testing the machine-learned interatomic potential (MLIP) with the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS). The model was validated based on two metrics of the potential – energy-volume relation, and elastic constants. For the energy-volume relation, the model achieved accuracies on the order of meV/atom compared to Density Functional Theory (DFT) energies. However, the predicted elastic constants showed significant deviations of up to 50% difference from both DFT and MEAM-based calculations, suggesting the need for further refinement in training datasets and metaparameters.

Background

Density Functional Theory

Electronic structure calculations such as DFT have set superior benchmarks in approximating material property behaviors for small ensembles of atoms (lattice structures fewer than a hundred atoms). However, DFT simulation performance significantly decreases as larger structures are modeled. Larger structures consisting of hundreds of atoms become too computationally expensive even for supercomputers due to the larger volumes of data and calculation complexity [1].

Many-Body Potential Formalisms

To address this, many-body interatomic potential calculations like Modified Embedded Atom Methods (MEAM) and Embedded Atom Methods (EAM) have traditionally scaled electronic structure data to molecular dynamics and statics simulations from hundreds to millions of atoms. Despite their semi-empirical premise, these formalisms cannot capture the breadth or complexity of other interatomic energy potentials. As a result, MEAM and EAM produce inflexible inputs for molecular dynamics simulations due to a limiting 13 free parameters for structural characterization [2]

Prior: MEAM Potential Expression

The MEAM formalism is the best current method of characterizing the atomic environment as it makes use of the angular screening between neighboring atoms [3]. The expression takes the form:

$$E = \sum_i \left[F_i(\rho_i) + \frac{1}{2} \sum_{j \neq i} \phi_{ij}(r_{ij}) \right]$$

Where the embedding function, $F_i(\rho_i)$ is parametrized empirically.

Numerical Methods

Rapid Artificial Neural Network

Feed-forward multi-layer perceptron (MLP) neural networks, as illustrated in Figure 1 below, have demonstrated impressive benchmarks of computing speed and accuracy for potential fitting [3].

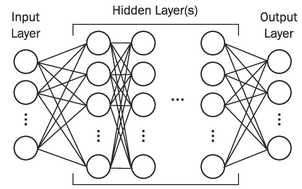
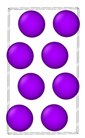


Figure 1: Illustration of Neural Network Architecture [2]

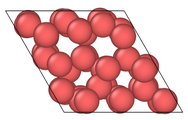
The model uses a Rapid Artificial Neural Network (RANN) with embedded structural fingerprints based on the MEAM formalism in the input layer.

Pair-Interaction Structural Fingerprint

Simple pair interactions are considered and summed over all the neighbors of a given atom as A_r^α described as:



$$A_r^\alpha = \sum^{\beta} \left(\frac{r^{\alpha\beta}}{r_e} \right)^n e^{-\alpha_n \frac{r^{\alpha\beta}}{r_e}} f_c \left(\frac{r_c - r^{\alpha\beta}}{\Delta r} \right)$$



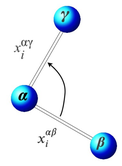
and the **cutoff radius**, r_c, is a MEAM-based piecewise:

$$r_c(x) = \begin{cases} 1, & x > 1 \\ (1-(1-x)^4)^2, & 0 \leq x \leq 1 \\ 0, & x < 0 \end{cases}$$

Where β labels the neighboring atoms within a cutoff radius of α. The cutoff radius varies from 0 to 1 to account for negligible effects from atoms at far displacements from atom α.

3-Body Structural Fingerprint

The 3-Body fingerprint form is similar to the partial electron densities used in the MEAM formalism, which accounts for the bond angle between atom α, atom β, and atom γ. It is described as:



$$A_b = \sum^{\beta,\gamma} \left(\frac{x^{\alpha\beta} \cdot x^{\alpha\gamma}}{r_e^2} \right)^m e^{-\alpha_k \frac{r^{\alpha\beta} + r^{\alpha\gamma}}{r_e}} f_c\left(r^{\alpha\beta}\right) f_c\left(r^{\alpha\gamma}\right)$$

Where the double sum dot product,

$$\theta_{\alpha\beta\gamma} = \left(\frac{x^{\alpha\beta} \cdot x^{\alpha\gamma}}{r_e^2} \right)$$

yields the **bond angle** between radii $r^{\alpha\beta}$ and $r^{\alpha\gamma}$.

Nonlinear Activation Function

Succeeding the implementation of the structural fingerprints is a vector of features describing the entire atomic configuration. In the instance of pure Y, an HCP structure, the input vector size for each atom has 29 features. The input vector of a single atom could be visualized as such:

$${}^0\mathbf{A}^\alpha = \begin{bmatrix} {}^0A_0{}^\alpha \\ {}^0A_1{}^\alpha \\ {}^0A_2{}^\alpha \\ \vdots \\ {}^0A_{29}{}^\alpha \end{bmatrix}$$

Within the hidden layer of the network, each node applies a weight matrix and bias vector that was previously determined in training using a Levenberg-Marquardt algorithm [3]. The weighted sums take the form,

$$\mathbf{Z}_1 = {}^0\mathbf{A}_1\mathbf{W}_0 + {}^0\mathbf{A}_1\mathbf{B}_0$$

A nonlinear activation function, $g(x)$ is applied to each weighted sum and passed to the output layer. The nonlinear activation function is given by:

$$g(x) = g(\mathbf{Z}_1) = \frac{x}{10} + \frac{9}{10} log(e^x + 1)$$

The output layer returns a predicted energy for each atom in the configuration. A regularization term of $\lambda = 1 \times 10^{-4}$ controlled the weight decay in the following loss function:

$$L_{\text{MSE}} = \frac{1}{m} \sum \left(\hat{Y} - Y \right)^2 + \frac{\lambda}{2m} \sum \|W\|_F^2$$

Validation

Primary validation consisted of overlaying 10% of the reserved DFT energies to test for overfitting. An RMSE analysis between the training data and validation data were 1.1412 meV/atom and 1.6302 meV/atom, respectively.

The role of the neural network is to predict the energy of each atom given its respective local environment; therefore, the accuracy of the fitted potential was evaluated on two common metrics, the **elastic constants** and **energy-volume relation**.

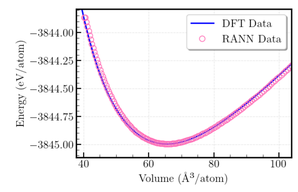
Table 1: Elastic Constants for Y [GPa]

Property	RANN(this work)	Exp.	2NN MEAM	DFT
C_{11}	55.16	83.40	77.74	86.27
C_{12}	17.32	29.10	30.00	25.73
C_{13}	15.85	19.00	28.21	19.41
C_{33}	49.10	80.10	75.32	69.51
C_{44}	15.98	26.90	21.71	24.10

Five independent elastic constants of pure Y in its HCP crystal structure.

Figure 2: Energy-Volume Relation

As seen in Figure 2, the RANN fitted potential replicated equilibrium point of the DFT energy-volume data with a minimum target/atom and maximum error/atom value of 0.4367 and 0.3688, respectively.



Conclusion

In conclusion, the RANN model effectively reproduced energy-volume relations with DFT-level accuracy but failed to improve elastic constant predictions compared to MEAM-based calculations. The performance of machine-learned interatomic potentials on complex mechanical properties highlights opportunities to develop more flexible and accurate models for atomic-scale simulations.

Future Work

Future studies should explore the effect of randomized seeds to assess the robustness of the RANN model, investigate FCC and BCC Yttrium structures to determine structural dependency, validate the model using generalized stacking fault energies for a more comprehensive mechanical characterization, and attempt a wider range of metaparameters to optimize network performance and enhance predictive accuracy.

References

[1] Nakata, A., Baker, J. S., Mujahed, S. Y., Poulton, J. T., Arapan, S., Lin, J., ... & Bowler, D. R. (2020). Large scale and linear scaling DFT with the CONQUEST code. The Journal of chemical physics, 152(16).

[2] Dickel, D., Nitol, M., Barrett, C. D. (2021). LAMMPS implementation of rapid artificial neural network derived interatomic potentials. Computational Materials Science, 196, 110481.

[3] Zhou, H., Dickel, D. Barrett, C.D. Improving stability and transferability of machine learned interatomic potentials using physically informed bounding potentials. Journal of Materials Research 38, 5106–5113 (2023).

[4] Nitol, Mashroor Shafat Dickel, Doyl Barrett, Christopher. (2021). Machine learning models for predictive materials science from fundamental physics: An application to titanium and zirconium

[5] Nichols, M., Barrett, C. D., Dickel, D. E., Nitol, M. S., Fensin, S. J. (2024). Predicting Ti-Al Binary Phase Diagram with an Artificial Neural Network Potential. arXiv preprint arXiv:2411.07960.