LAMMPS implementation of rapid artificial neural network derived interatomic potentials

Christopher Barrett,

Mashroor Nitol,

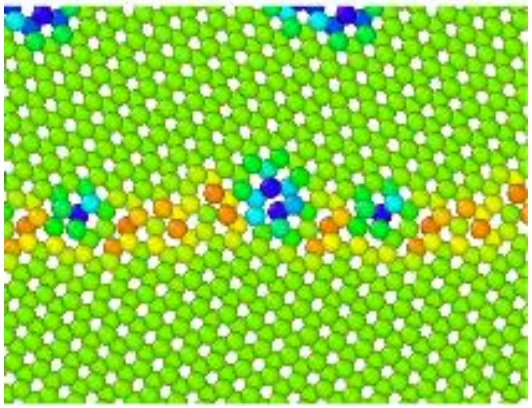
Doyl Dickel

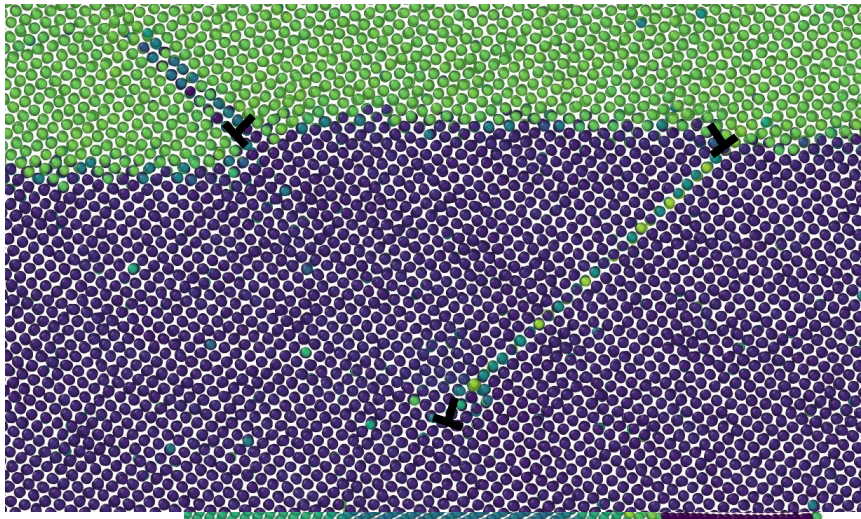
Uses of atomistic potentials in metals

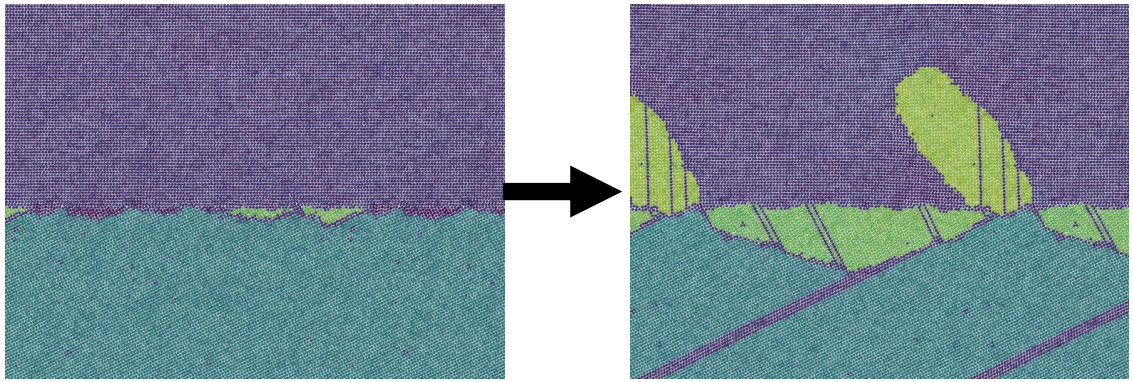
Over the past 30 years, atomistic simulations have been demonstrated to be a powerful tool for developing insight between the electronic scale and mesoscale.

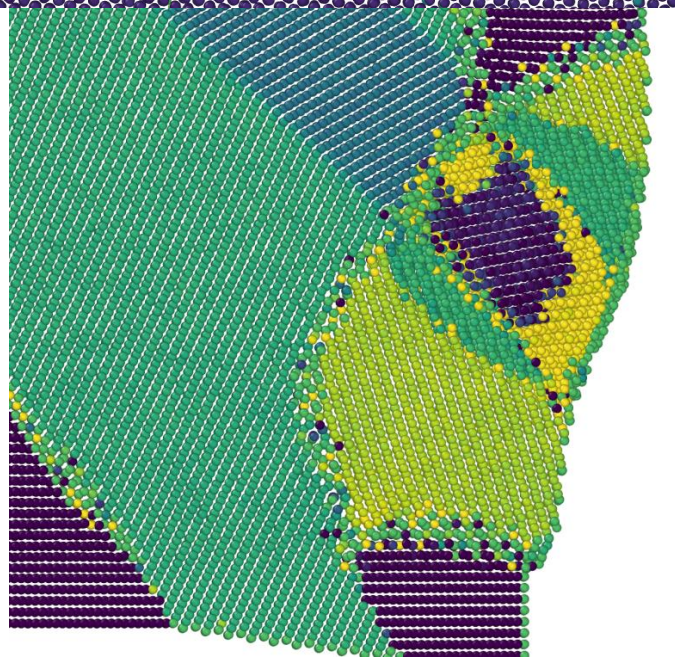
They work well for characterizing:

- Grain boundary structures
- Twinning and martensitic transformations
- Dislocation reactions
- Segregation energies
- Etc.





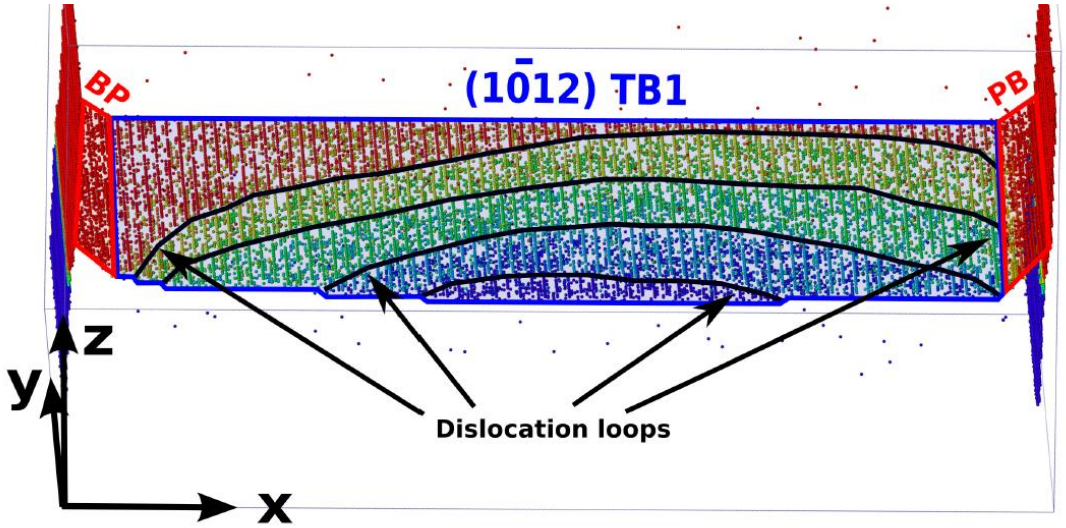


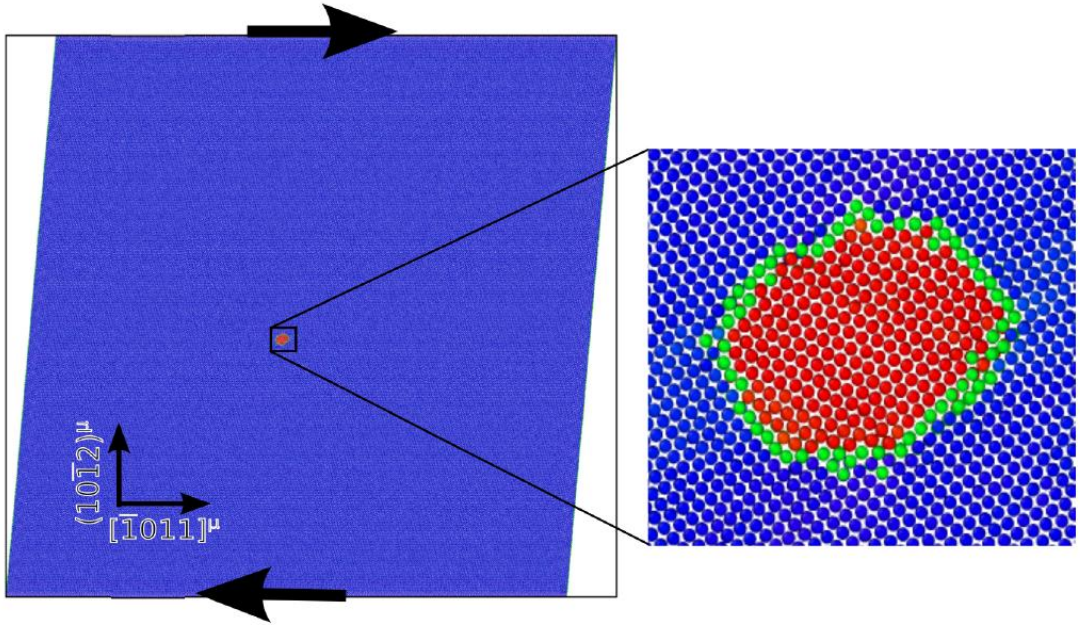


Typical limitations of atomistic potentials

Disadvantages:

- Timescale is extremely limited – cannot characterize diffusion processes
- Spatial scale is also very limited – periodic and free surface simulations generally contain major deviations from experimental results because of scale
- Difficult to fit to multiple different kinds of properties
- Development of alloy potentials is very time consuming





Types of atomistic potentials

Simple and efficient potentials:
- Lennard-Jones
- Embedded Atom Method

A little slower, but more accurate:
- MEAM
- MEAM-spline

Potentials designed for covalent bonding:
- ReaxFF
- Tersoff

Machine Learning:
- SNAP
- HDNNP (N2P2)
- QUIP (GAP)

Many other potential types also exist

Why do we want another one?

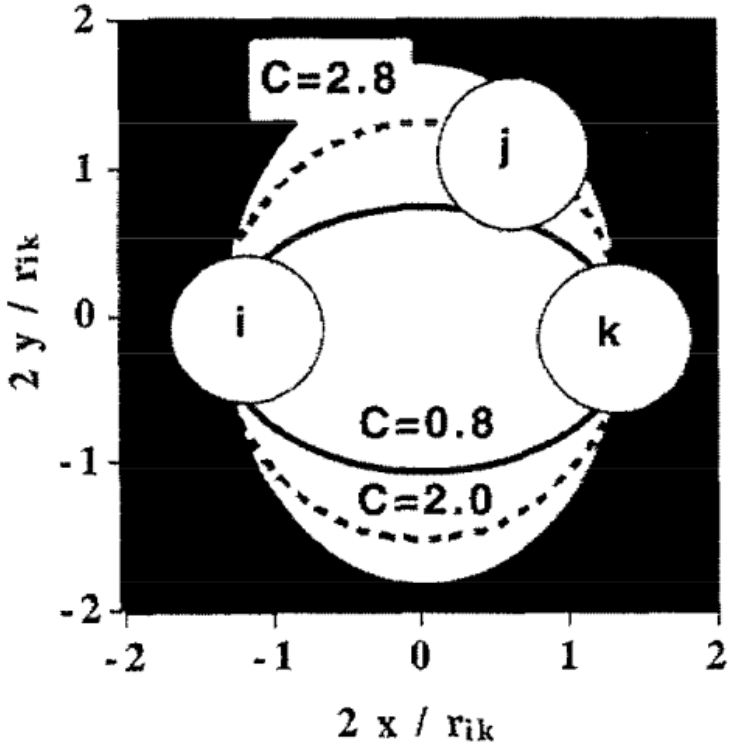
- Existing classical potentials have limited accuracy and are time-consuming to fit.
- Machine learning potentials can have much higher accuracy and the fitting can be streamlined.
- Therefore we have developed rapid artificial neural network (RANN) potentials

Why do we need another kind of Machine Learning potential?

- GAP is very slow
- SNAP is much faster, but has limited complexity due to the lack of hidden layers
- N2P2 was developed concurrently to our approach and is similar
    - RANN potentials are usually faster than N2P2
    - RANN is more user-friendly than N2P2
- No other existing ML method has screening

Screening

- Pair effects between a central atom and its neighbors are reduced or eliminated by other neighbors which are in between them.
- Our screening implementation is borrowed from the MEAM formulation.



$$cos\left(\theta^{\alpha\beta\gamma}\right) = \frac{\mathbf{r}^{\alpha\beta} \cdot \mathbf{r}^{\alpha\gamma}}{r^{\alpha\beta} r^{\alpha\gamma}}$$

$$X^{\gamma\beta} = \left(\frac{r^{\gamma\beta}}{r^{\alpha\beta}}\right)^2$$

$$X^{\alpha\gamma} = \left(\frac{r^{\alpha\gamma}}{r^{\alpha\beta}}\right)^2$$

$$C = \frac{2\left(X^{\alpha\gamma} + X^{\gamma\beta}\right) - \left(X^{\alpha\gamma} - X^{\gamma\beta}\right)^2 - 1}{1 - \left(X^{\alpha\gamma} - X^{\gamma\beta}\right)^2}$$

$$f_c(x) = \begin{cases} 1 & x \geq 1 \\ \left(1 - (1-x)^4\right)^2 & 0 < x < 1 \\ 0 & x \leq 0 \end{cases}$$

$$S^{\alpha\beta\gamma} = f_c\left(\frac{C - C_{min}}{C_{max} - C_{min}}\right)$$

$$S^{\alpha\beta} = \prod_{\gamma} S^{\alpha\beta\gamma}$$

Baskes, Materials Chemistry and Physics, 50(2), 152-158, (1997).

Structural Fingerprints

- Describe local environment in efficient, comprehensive and physically meaningful way
- Reduce dependence on exact number of neighbors
- Enforce ordering invariance and rotation invariance

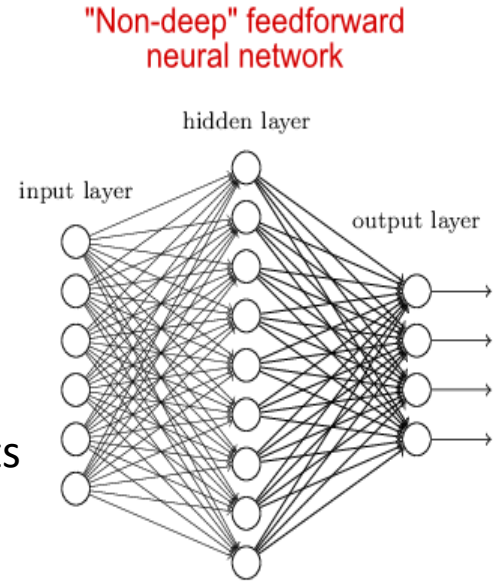
$$A_n = \sum_i \left(\frac{r_i}{r_e}\right)^n \exp\left(-\alpha \frac{r_i}{r_e}\right) f(r_i)$$

$$A_{m,k} = \sum_{i,j} g_m(\cos\theta_{ij}) \exp\left(-\alpha_k \frac{r_i + r_j}{r_e}\right) f(r_i) f(r_j)$$

- We use two types of structural fingerprints, which primarily incorporate the effects of neighbor radial distribution and bond angle respectively
- Exponentially decaying intensity reflects underlying physics
- Physical motivation implies simpler network (fewer necessary inputs and neurons) and greater transferability
- These fingerprints are very similar to intermediate quantities computed by the MEAM potential

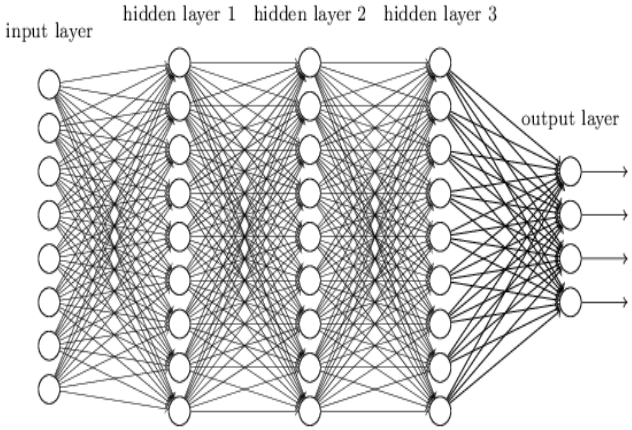
Machine Learning in LAMMPS

Machine learning works by having inputs fed through a series of “neurons” to produce a final output signal



Structural Fingerprints

Atomic Energies



Each neuron receives signals from neurons in the previous layer which are weighted and summed, then biases the result and feeds it through an activation function

$$Z_1 = \sum_i W_{n;i} A_{n;i} + B_n$$

$A_{0;i}$ = Inputs (Structural fingerprints)

$$A_{n;i} = g_i(Z_{n;i})$$

$$A_{n,i} = g_i\left(\sum_j W_{n-1;i,j} A_{n-1;j} + B_{n-1;i}\right)$$

Each layer may have a different number of neurons and the number of layers can be varied depending upon the problem requirements

The calibration process

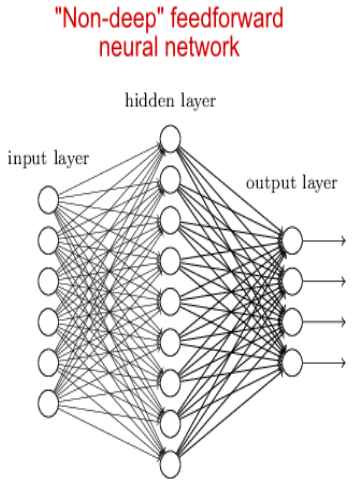
The individual weights and biases of each neuron are the fitting parameters for the network

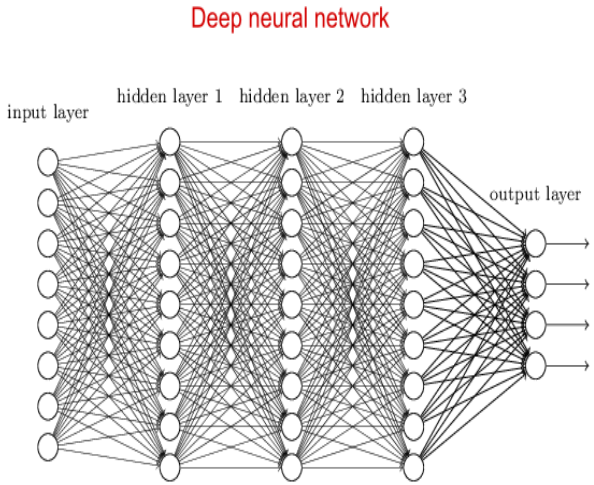
$$\frac{\partial Z_{n,i}}{\partial W_{n-2;j,k}} = \sum_j W_{n-1;i,j} \frac{\partial A_{n-1;j}}{\partial W_{n-2;j,k}}$$

$$A_{n;i} = g_i(Z_{n;i}) \qquad dA_{n;i} = dg_i(Z_{n;i})dZ_{n;i}$$

Because of the matrix-based structure of each neuron, the gradients with respect to all of the weights and biases are easily produced by backpropagation.

The large number of parameters indicates that overfitting is often a problem. Overfitting is avoided by checking the network performance on data which it was not trained with.





Derivation of Atomic Forces

The force on atom x in the η direction is defined as the derivative of the system energy with respect to the η coordinate of atom x's position

$$F_{x;\eta} = \frac{\partial E}{\partial x_\eta}$$

Using backpropagation, this may be derived analytical through the network

$$\frac{\partial Z_{n,i}}{\partial x_\eta} = \sum_j W_{n-1;i,j} \frac{\partial A_{n-1;j}}{\partial x_\eta} \qquad dA_{n;i} = dg_i(Z_{n;i}) dZ_{n;i}$$

Finally, the derivatives of the structural fingerprints must be obtained

$$dA_{0;n} = \sum_i \left(\frac{r_i}{r_e}\right)^n \exp\left(-\alpha \frac{r_i}{r_e}\right) f(r_i) \left(\frac{n}{r_i} - \frac{\alpha}{r_e} + \frac{df(r_i)}{f(r_i)}\right) dr_i$$

$$dA_{0;m,k} = \sum_{i,j} g_m(\cos\theta_{ij}) \exp\left(-\alpha_k \frac{r_i + r_j}{r_e}\right) f(r_i) f(r_j) \left(\frac{dg_m(\cos\theta_{ij})}{g_m(\cos\theta_{ij})} - \frac{\alpha}{r_e} + \frac{df(r_i)}{f(r_i)}\right) dr_i$$

$$dr_i = \frac{x_\eta}{r_i} dx_\eta$$

Making alloy potentials

- Each element has multiple sets of fingerprints depending on which neighbors are included

$$A_{n,X} = \sum_{i \in X} \left(\frac{r_i}{r_e}\right)^n \exp\left(-\alpha \frac{r_i}{r_e}\right) f(r_i)$$

$$A_{m,k,X,Y} = \sum_{i \in X, j \in Y} g_m(\cos\theta_{ij}) \exp\left(-\alpha_k \frac{r_i + r_j}{r_e}\right) f(r_i) f(r_j)$$

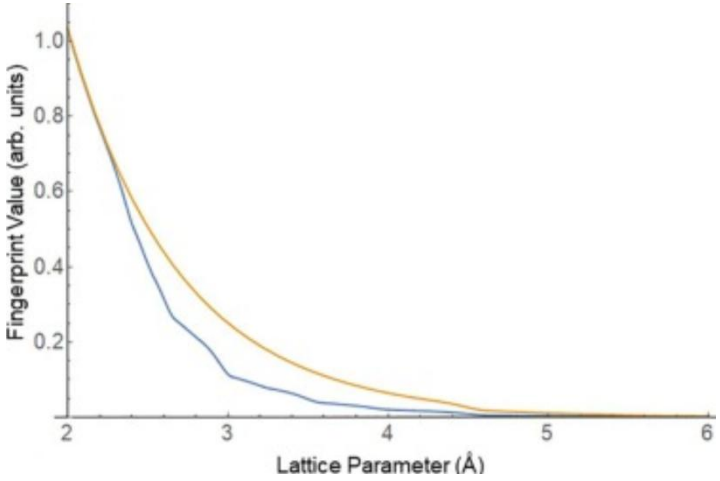
- List of inputs expands from 32 to 104 for two different networks
- In general, complexity increases like $\mathcal{O}(N^2)$ (N = number of element types)

Choice of fingerprints and architecture

Effect of varying 3-body terms and hidden layer size on calculation error and speed for a Mg potential:

k_{tot}	m_{tot}	#SF	r_c	C_{min}	#NN	Training RMSE	Validation RMSE	Calculation speed
*3	4	13	6	0.49	20	0.90	2.25	21.77
3	4	13	8	0.49	20	0.67	1.64	15.74
3	4	13	10	0.49	20	0.69	1.71	10.71
2	4	13	12	0.25	38	0.89	1.71	5.56
2	4	13	12	0.49	20	1.09	1.05	8.26
2	4	13	12	0.70	18	0.93	1.08	9.34
3	4	17	12	0.25	38	0.80	1.03	5.07
3	4	17	12	0.49	20	0.71	1.38	8.04
3	4	17	12	0.70	18	0.80	1.06	8.75
5	3	20	12	0.25	38	1.57	2.95	4.91
5	3	20	12	0.49	20	2.80	3.08	7.88
5	3	20	12	0.70	18	2.88	3.10	8.32
MEAM [33]								61.963
N2P2 [27]								5.12

Effect of screening on a single fingerprint as the lattice parameter is scaled:



The effect of screening on computation speed can be positive or negative, but generally speeds up potentials with normal cutoff radii.

Functional Form of the Neural Network Potential for Ti

n = 1 through 4, m = 0 through 6, k = 1 through 4

The total size of the input layer is 32
$= n_{max} + (m_{max} + 1)k_{max}$

$$A_{0;n} = \sum_i \left(\frac{r_i}{r_e}\right)^n \exp\left(-\alpha \frac{r_i}{r_e}\right) f(r_i)$$

$$A_{0;m,k} = \sum_{i,j} g_m(\cos\theta_{ij}) \exp\left(-\alpha_k \frac{r_i + r_j}{r_e}\right) f(r_i) f(r_j)$$

$$r_e = 2.91^5$$
$$r_c = 6.75$$
$$dr = 0.5$$
$$\alpha = 4.65$$
$$\alpha_k = [1\ 2\ 4\ 8]$$

Architecture:
Input layer: 32 neurons
Hidden Layer 1: 20 neurons
Output Layer: 1 neuron

Activation Functions:
$$g_i(x) = 0.1x + 0.9\ln(e^x + 1)$$
$$g_2(x) = x$$

$$A_{n,i} = g_i(\sum_j W_{n-1;i,j} A_{n-1;j} + B_{n-1;i})$$

681 individual weights and biases are fit

How to use RANN in LAMMPS

Input script:

Syntax like eam/alloy and many other pairs:

pair_style rann

pair_coeff * * potential_file.nn Element1 …

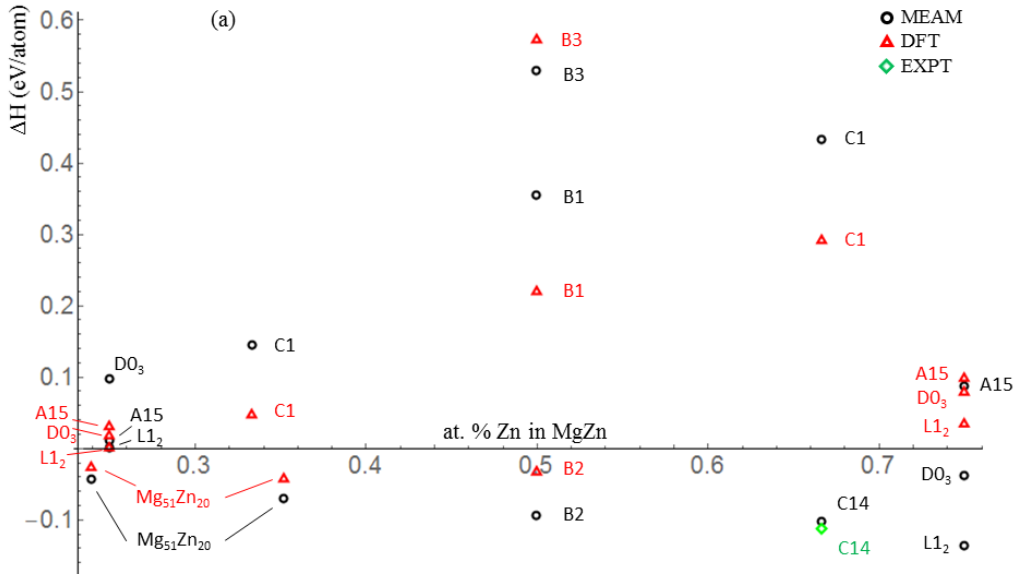
Potential file syntax:

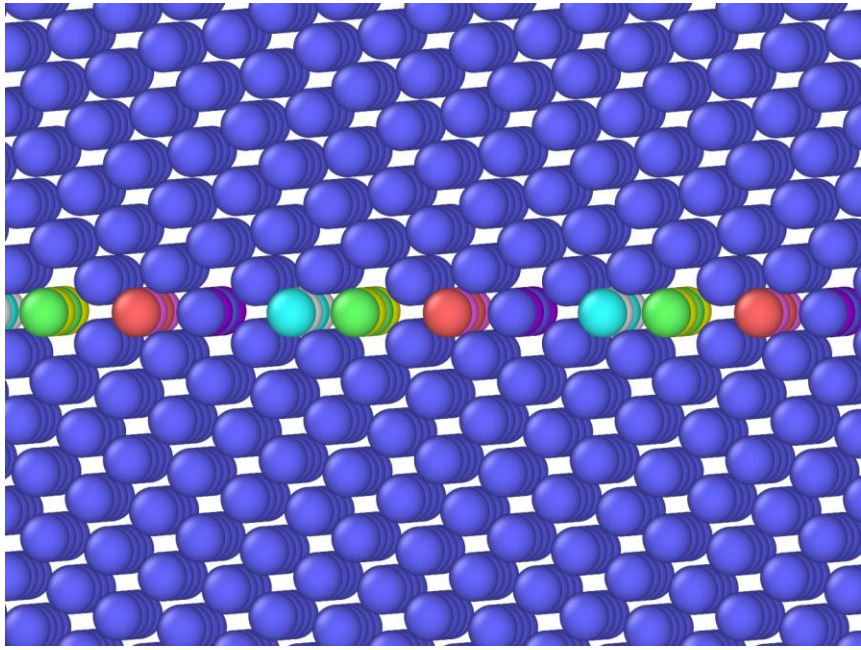
The potential file is designed to be human readable. It is divided into sections by keywords:

- atomtypes
- mass
- fingerprintsperelement
- fingerprints
- fingerprintconstants
- screening (optional)
- networklayers
- layersize
- weight
- bias
- activationfunctions
- calibrationparameters (ignored)

```
atomtypes:
Mg
mass:Mg:
24.305
fingerprints:Mg_Mg:
 radialpower_0
fingerprints:Mg_Mg_Mg:
 bondpower_0
fingerprintconstants:Mg_Mg:radialpower_0:re:
3.1936
fingerprintconstants:Mg_Mg:radialpower_0:rc:
6
fingerprintconstants:Mg_Mg:radialpower_0:dr:
2.8064
fingerprintconstants:Mg_Mg:radialpower_0:n:
3
fingerprintconstants:Mg_Mg:radialpower_0:o:
-1
fingerprintconstants:Mg_Mg:radialpower_0:alpha:
5.52        5.52        5.52        5.52        5.52
fingerprintconstants:Mg_Mg_Mg:bondpower_0:re:
3.1936
fingerprintconstants:Mg_Mg_Mg:bondpower_0:rc:
6
fingerprintconstants:Mg_Mg_Mg:bondpower_0:dr:
2.8064
fingerprintconstants:Mg_Mg_Mg:bondpower_0:m:
8
fingerprintconstants:Mg_Mg_Mg:bondpower_0:alphak:
1  2  6  9
fingerprintconstants:Mg_Mg_Mg:bondpower_0:k:
4
networklayers:Mg:
3
layersize:Mg:0:
37
layersize:Mg:1:
20
layersize:Mg:2:
1
weight:Mg:0:
  -2.854926896534886    1.551245978744016   -5.299486715905226    1.30303239306658
-3.927350992087088    0.806210218127504   -1.428595777065406   -0.168066964538803
-0.728351260519369   -0.157450640068210    5.491434639296052    1.080334461269439
-2.130535923458539   -1.156477049555650   -2.051010848588853   -0.782611130585220
-4.606523573915344   -2.387644217672982   -5.576794096713175   -3.318269168403370
-6.292539058097229   18.703592187192875   -2.381590380244950    2.799241528343801
-2.407391034332221    4.405517388343755   -9.491423045727968    8.229571895198635
2.527344035625121   -0.067244382960372    0.170484581466223    1.282379910053745
0.065724915799647    1.113377435138116   -0.573535002947138    1.560506886857219
```

Case Example: Single Element Zn





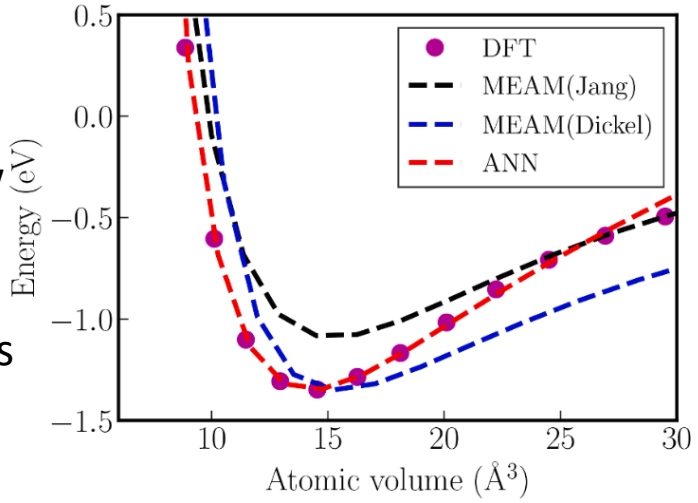
Case Example: Single Element Zn

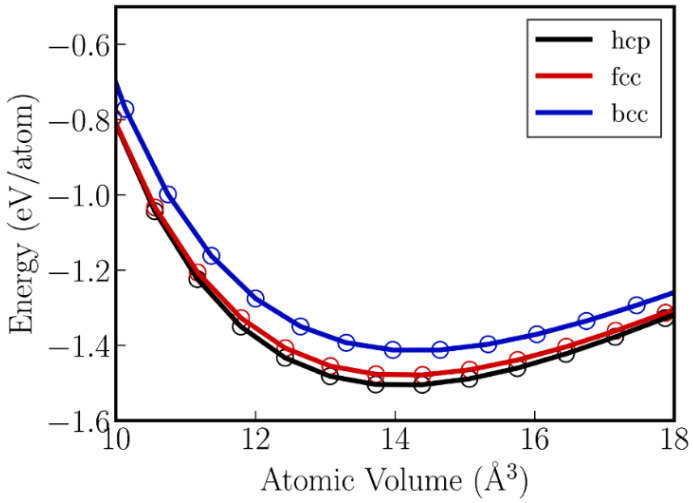
- Two new MEAM potentials for Zn were published in 2018 independently by two different groups
- Zn is a critical alloying element in many Iron, Aluminum and Magnesium materials
- Because of its very high c/a ratio, Zn is very difficult to accurately simulate
- BJ Lee et al's potential captures the energy difference from fcc to hcp well, but has very bad lattice parameters and elastic constants
- Dickel et al's potential captures the lattice parameters and elastic constants well, but incorrectly predicts fcc as more stable than hcp

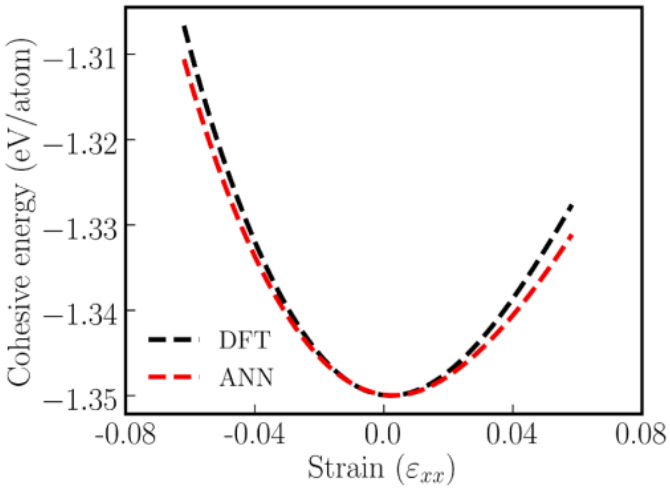
New RANN potential for zinc

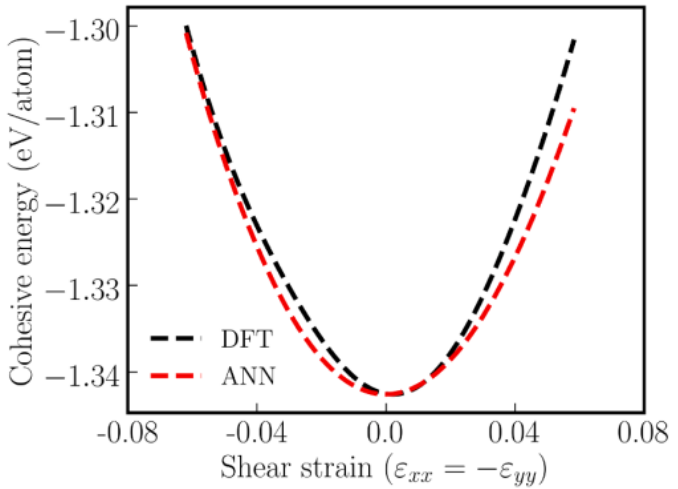
In a first for Zn, our new potential has:

- Accurate c/a ratio
- Accurate GSFE curves
- Accurate EV curves
- Reasonable Elastic constants
- Reasonable melting temperature
- Accurate thermal expansion coefficients including the expansion anisotropy



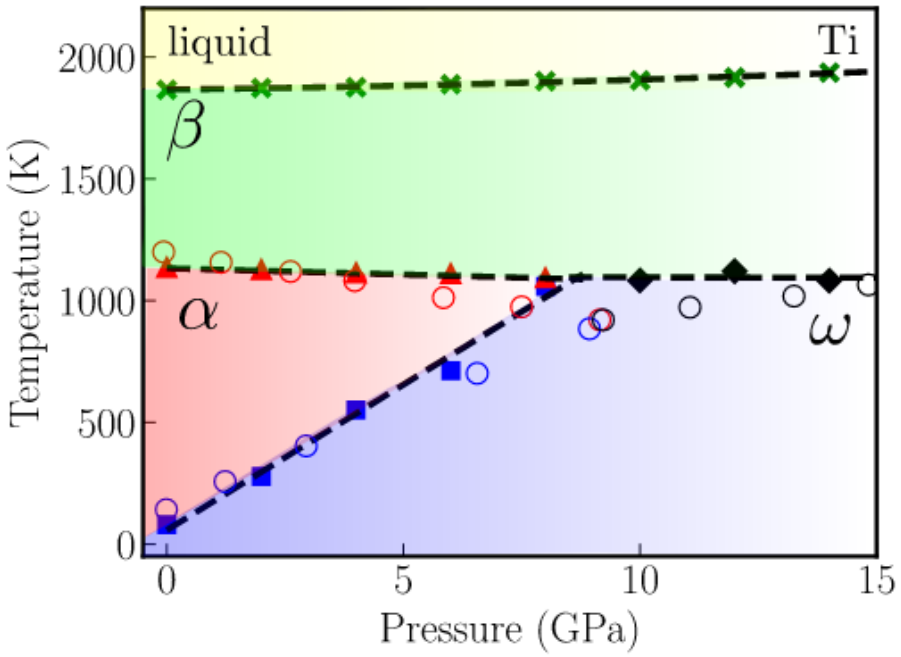


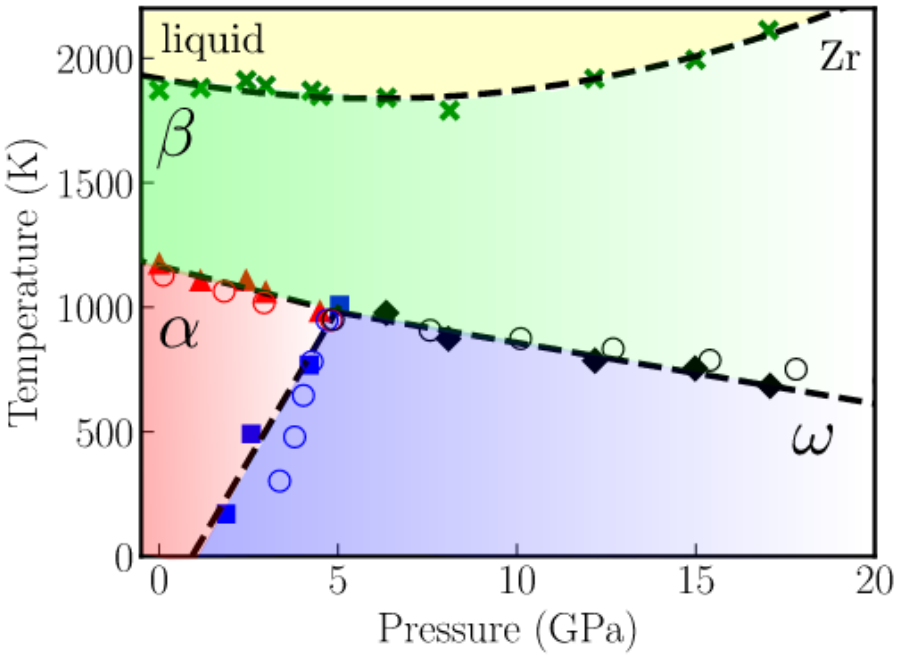




Nitol, Dickel, and Barrett. *Comp. Mat. Sci.* 188 (2021): 110207.

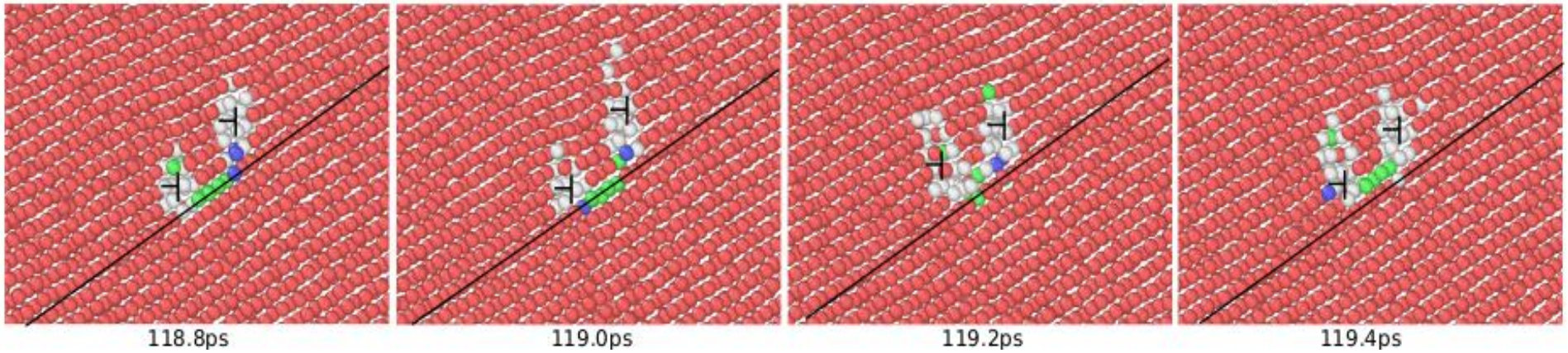
Applications for Ti and Zr





New potentials for Ti and Zr correctly capture not only melting but also phase transitions between three solid phases as a function of temperature and pressure!

Applications for Mg



- 2nd order pyramidal <c+a> slip in Mg is an important plastic mechanism
- Recent understanding has been that the edge dislocation of this mode spontaneously dissociate, becoming sessile.
- Our new potential shows the dislocations can still migrate at low stress even after dissociating.
- A recent N2P2 potential by Stricker et al. confirms the same results
- Other MEAM and EAM potentials for Mg cannot produce this mechanism.