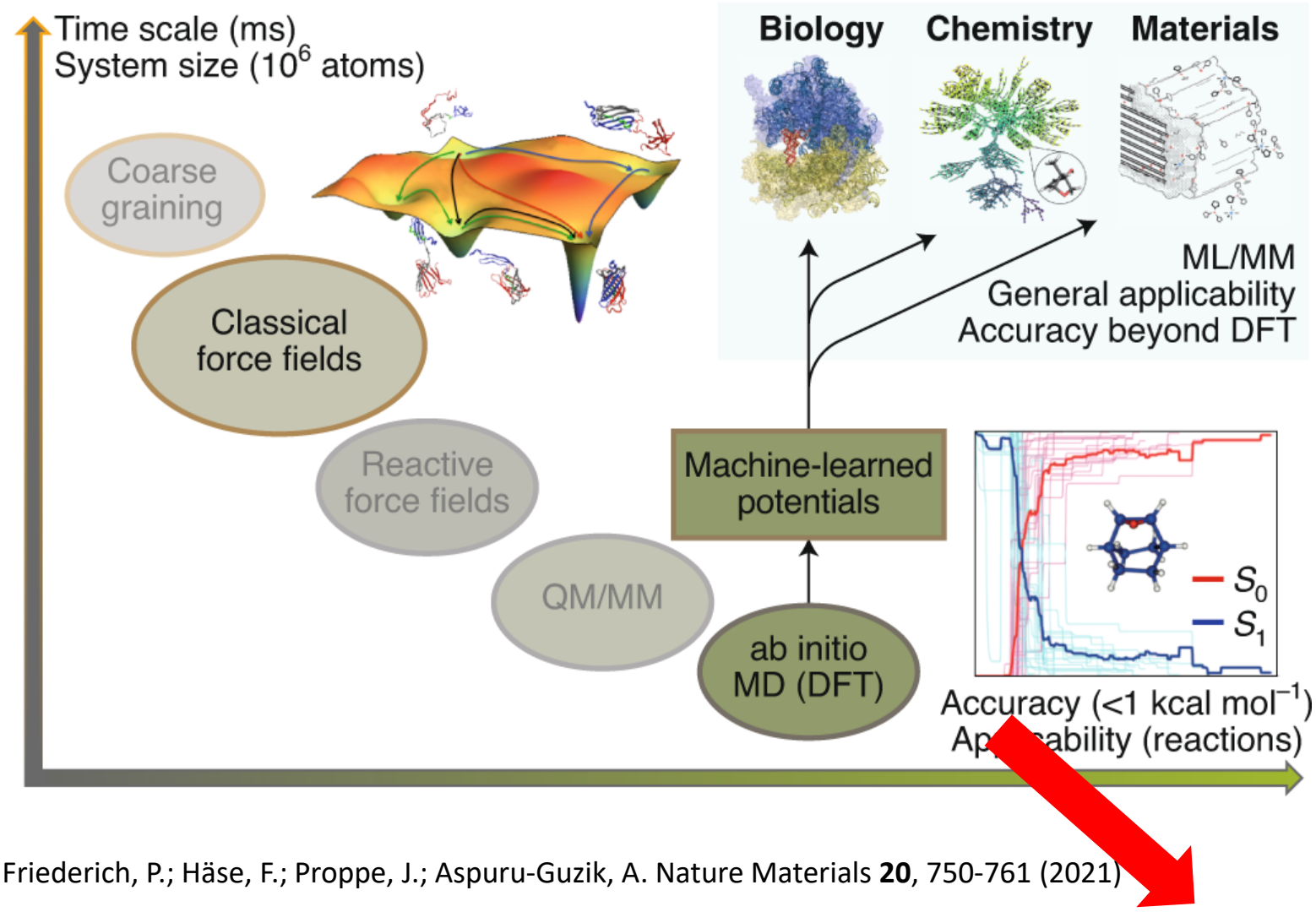
Feasibility Study of *GNN/ACE* for Chemistry

Minsik Cho
(Mentor: Dr. Emmanuel Lujan)

Computational Chemistry

Cost-accuracy tradeoff limits simulation of *large* systems in *long* time scale



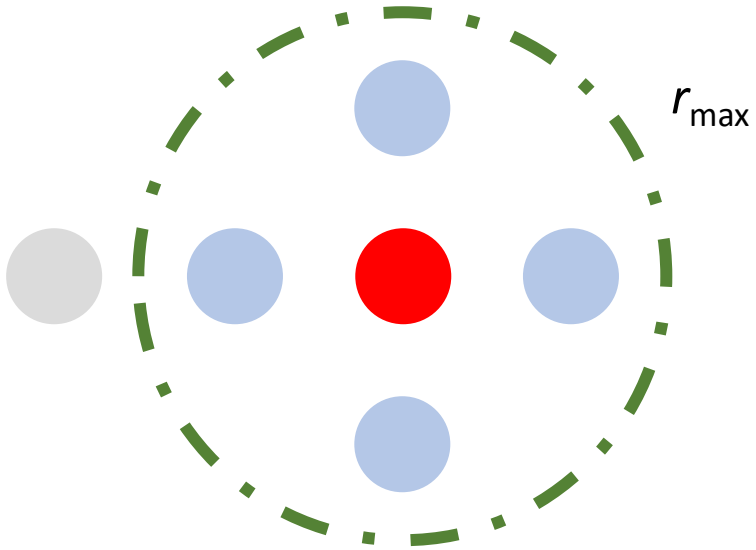
For AIMD,
interatomic potentials
are the quantities that matter

MD requires updates
for each time step

Friederich, P.; Häse, F.; Proppe, J.; Aspuru-Guzik, A. Nature Materials **20**, 750-761 (2021)

Atomic Cluster Expansion

Expand the energies (and forces) as a many-body expansion of atoms within some cutoff distance



$$E_i = \varepsilon_{(i)}^{(1)} + \frac{1}{2}\sum_j \varepsilon_{(i,j)}^{(2)} + \frac{1}{6}\sum_{j,k} \varepsilon_{(i,j,k)}^{(3)} + \cdots$$

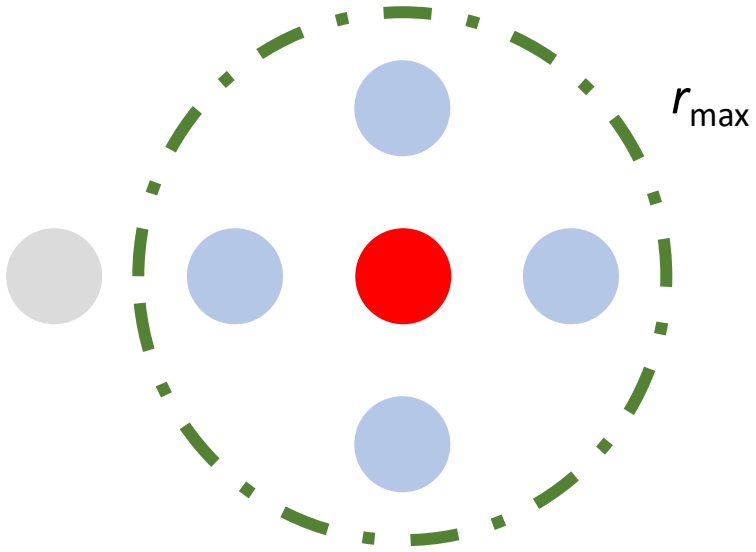
Rotational and Translational Invariance
Permutation Invariance
…

through tensor product states of hydrogen-like wavefunctions

Drautz, R. PRB **99**, 14104 (2019)

Atomic Cluster Expansion

Expand the energies (and forces) as a many-body expansion of atoms within some cutoff distance



$$E_i = \varepsilon_{(i)}^{(1)} + \frac{1}{2}\sum_j \varepsilon_{(i,j)}^{(2)} + \frac{1}{6}\sum_{j,k} \varepsilon_{(i,j,k)}^{(3)} + \cdots$$

Define neighborhood density

$$\rho(i,\mu) = \sum_{j \neq i} \delta_{\mu,\mu_j} \delta(r - r_{ji})$$

Project neighborhood density into hydrogen-like orbitals

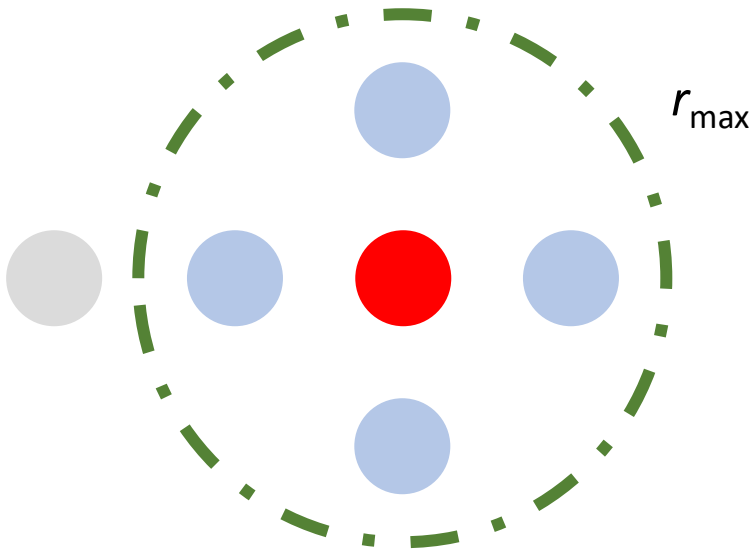
$$\phi_{\mu_i,\mu_j,n,l,m} = R_{n,l}^{\mu_i,\mu_j}(r_{ji}) Y_{l,m}(r_{ji})$$

$$A_{i,(\mu_i,\mu_j,n,l,m)} = < \rho(i,\mu) | \phi_{\mu_i,\mu_j,n,l,m} > = \sum_j \phi_{\mu_i,\mu_j,n,l,m}(r_{ji})$$

Drautz, R. PRB **99**, 14104 (2019)

Atomic Cluster Expansion

Expand the energies (and forces) as a many-body expansion of atoms within some cutoff distance



$$E_i = \varepsilon_{(i)}^{(1)} + \frac{1}{2}\sum_j \varepsilon_{(i,j)}^{(2)} + \frac{1}{6}\sum_{j,k} \varepsilon_{(i,j,k)}^{(3)} + \cdots$$

Take tensor products for n-th order cluster terms

$$A_{i,v} = \prod_{t=1}^{\nu} A_{i,(\mu_t,n_t,l_t,m_t)}$$

over all possible n-body combinations

Set up energy estimator based on the basis

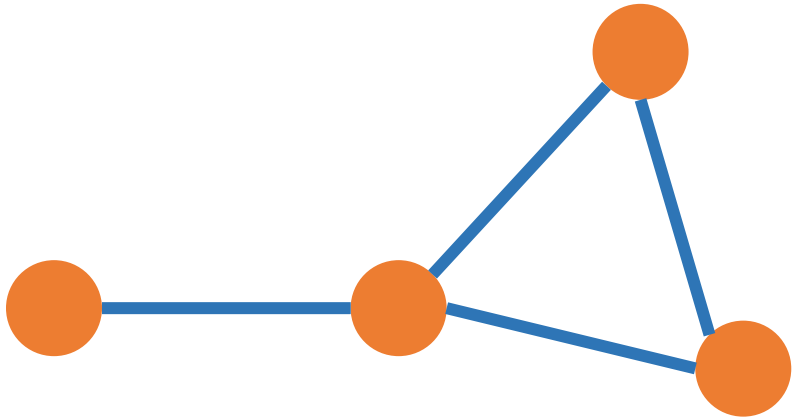
$$\psi_i^{(p)} = \sum_v c_v^{(p)} B_{i,v} = \sum_v c_v^{(p)} \sum_{v'} c_{v,v'} A_{i,v}$$

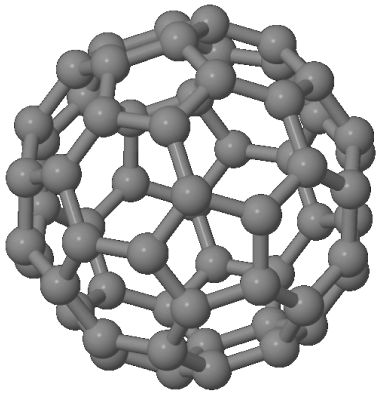
$$E_i = \mathcal{F}(\psi_i^{(0)}, \cdots, \psi_i^{(p)}, \cdots)$$

Drautz, R. PRB **99**, 14104 (2019)

Graph Neural Networks

Input data are graphs…





represented as adjacency matrices

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

each **node**, **edge**, and **graph** can have attributes

Molecular geometry as input graphs,
ACE local descriptors as node attributes
Energies and **forces** as graph attributes (output)

Sanchez-Lengeling, B. Reif, E., Pearce, A., Wiltschko, A. Distill (2021)

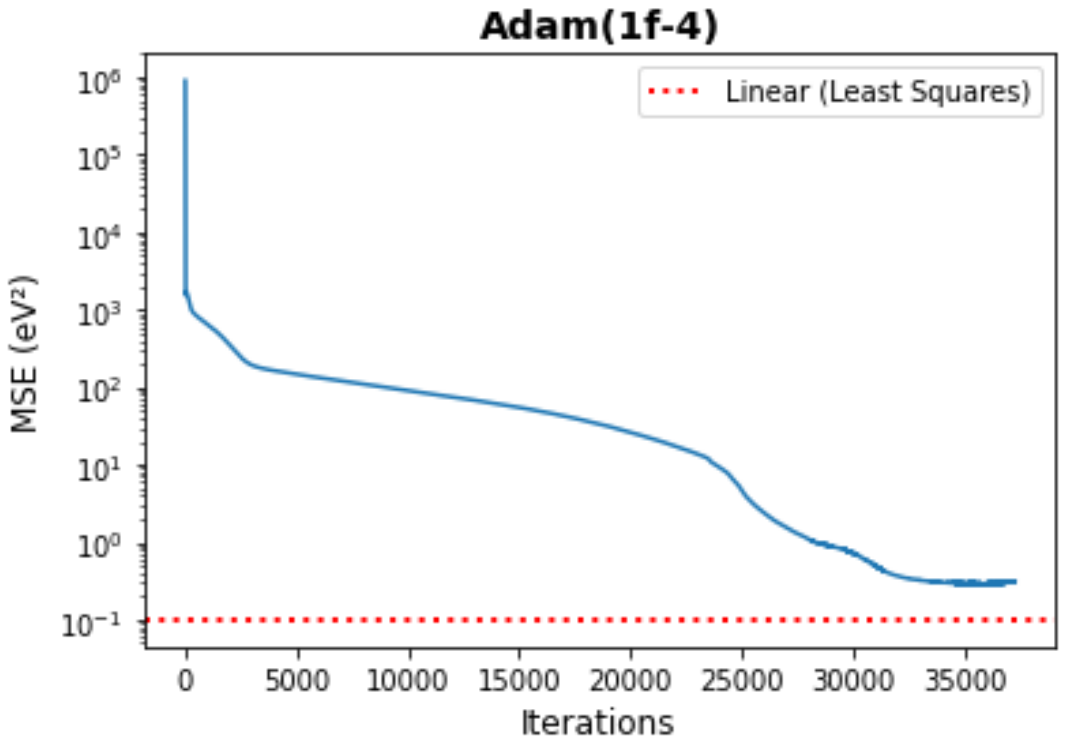
Workflow

Training / Test dataset : 1) molecular geometry 2) DFT energies	`PotentialLearning.jl (xyz.jl)`
Expressed in ACE basis set, with local descriptors	`PotentialLearning.jl (ace.jl)`
Build normalized adjacency matrix	
Define simple GNN model	`GraphNeuralNetworks.jl`
Train the neural net	`Flux.jl`
Compare with linear (least squares) solution	

Training

a-HfO2: 96 atoms (either Hf or O) in periodic cell

Least squares fit of 200 input data, 100 test data: MSE = 0.10 eV (67 sec)



Initial training with small fixed learning rate
(200 input data, 100 test data)
~ hrs time scale

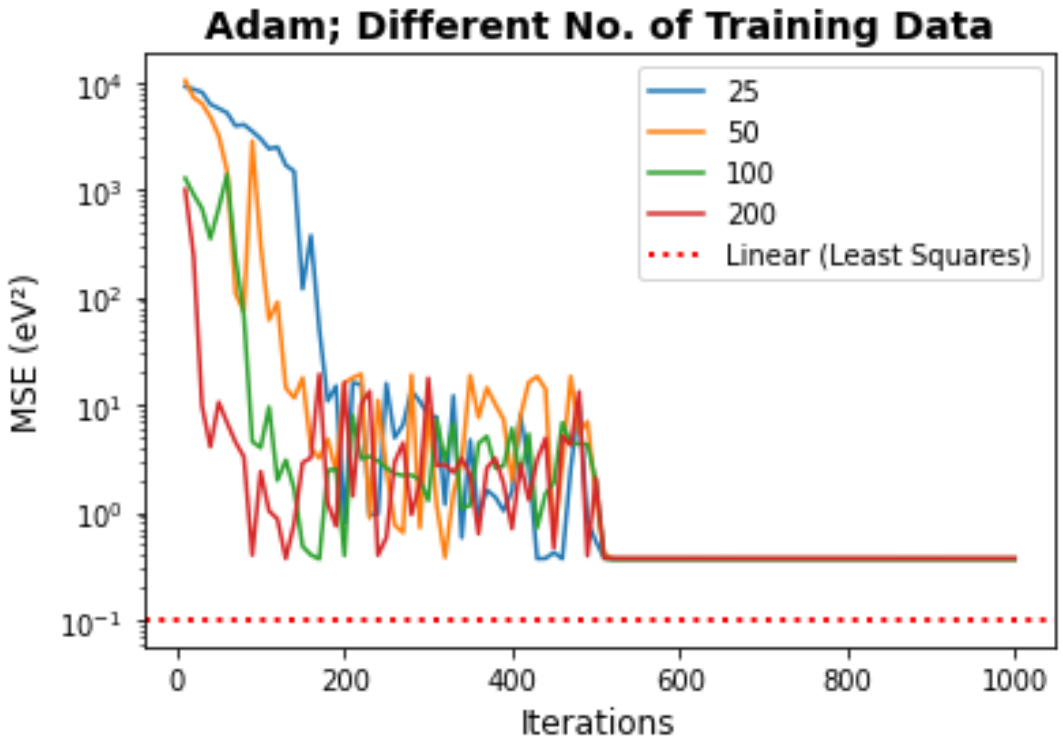
To improve convergence, use

- Normalized adjacency matrix
- Xavier initialization
- Larger learning rate, change momentum

Training

a-HfO2: 96 atoms (either Hf or O) in periodic cell

Least squares fit of 200 input data, 100 test data: MSE = 0.10 eV (67 sec)



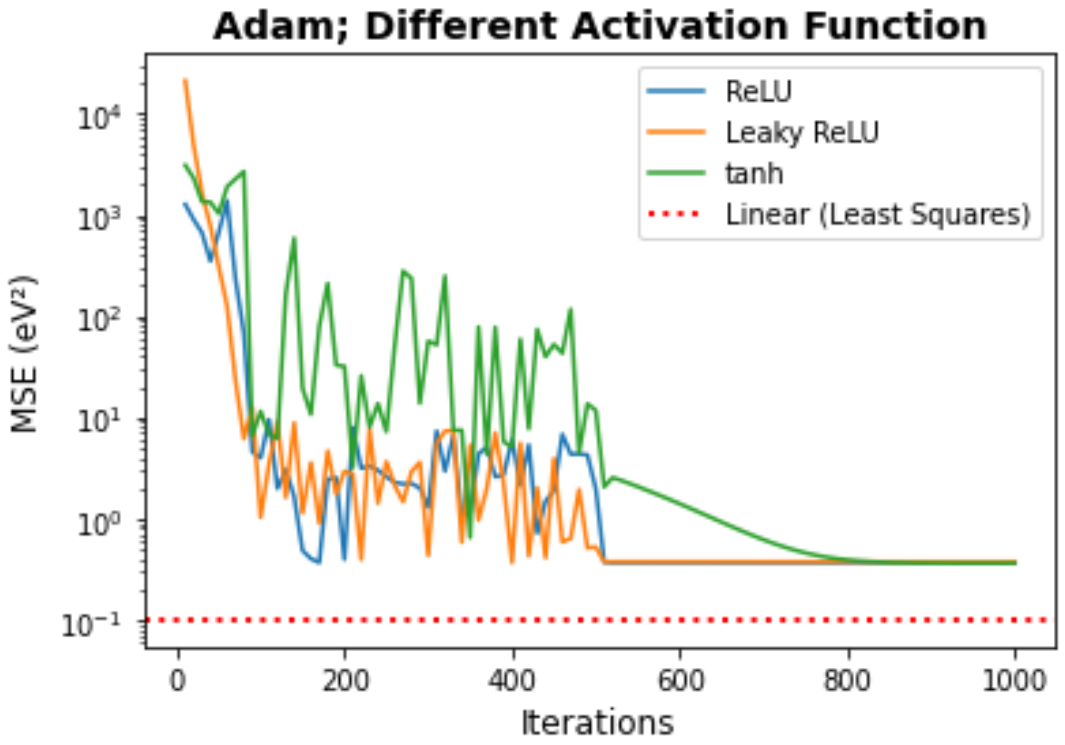
Improved training (~ minutes time scale)

GPU code tested, but difference insignificant (compared to 20-threaded CPU runs)

Training

a-HfO2: 96 atoms (either Hf or O) in periodic cell

Least squares fit of 200 input data, 100 test data: MSE = 0.10 eV (67 sec)



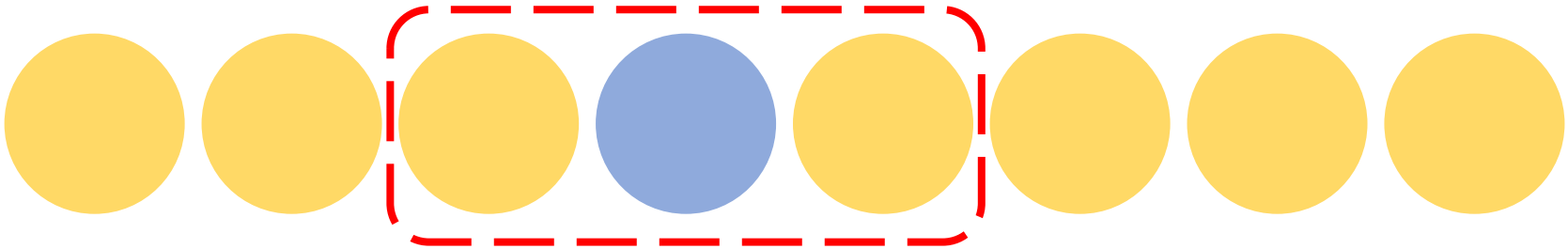
Modifications to the activation function do not cause significant changes

Next Steps…

Verify with other dataset
(i.e. FitSNAP)

Connect edge embedding in GNN & cluster expansion terms with quantum embedding methods

- Edge embedding over multiple edges
- Handle incomplete attributes



Experiment larger GNN models with full dataset