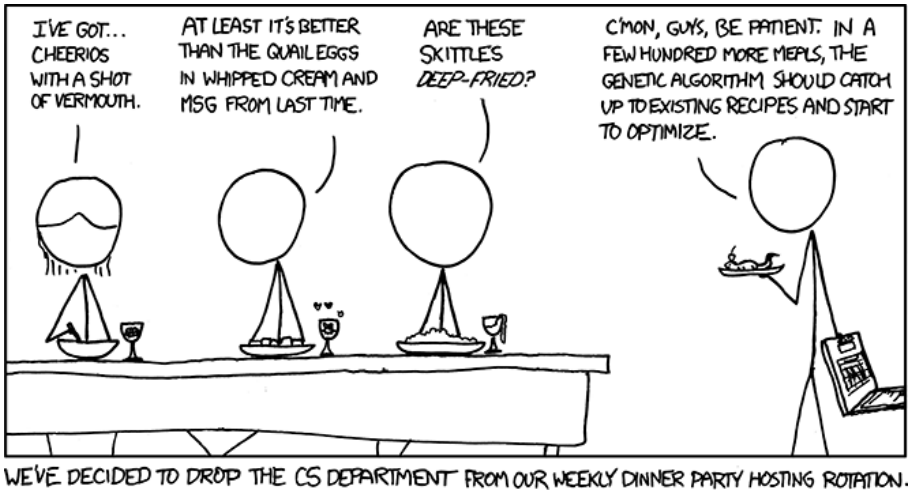


Machine learned interatomic potentials

Sai Gautam Gopalakrishnan, Debsundar Dey, and Tejus Rohatgi

Materials Engineering, Indian Institute of Science

saigautamg@iisc.ac.in; https://sai-mat-group.github.io

AI/ML for Materials Science Workshop
Jan 8, 2025

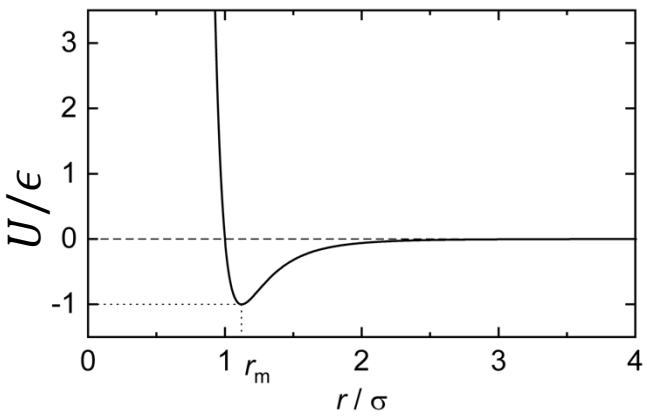
Why interatomic potentials?

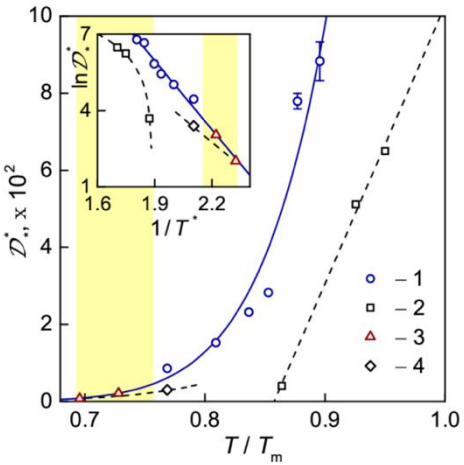
Interatomic potentials: simulate ‘large’ length-scale or ‘long’ time-scale phenomena

- Classical force-fields
- Length: ~nm, Time: ~ns (with molecular dynamics)
- Interfaces, diffusivities, rapid phase transitions (→phase diagrams)
- Underlying structure can change (vs. lattice models)
- Computational cost-accuracy trade-off

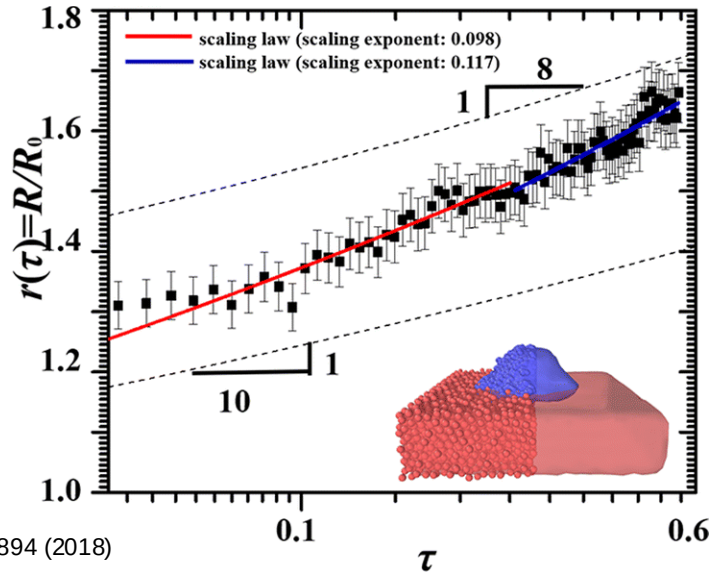
Interatomic potentials model the potential energy surface of a given material

Lennard-Jones: $$U(r) = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right]$$





Tipeev et al., J. Phys. Chem. C 122, 28884-28894 (2018)

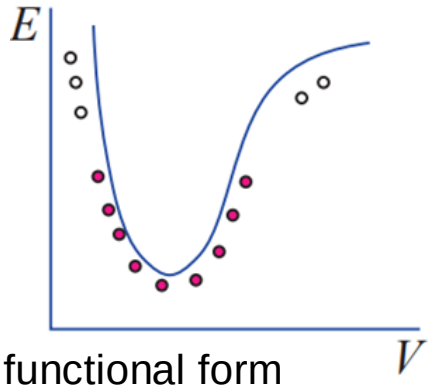


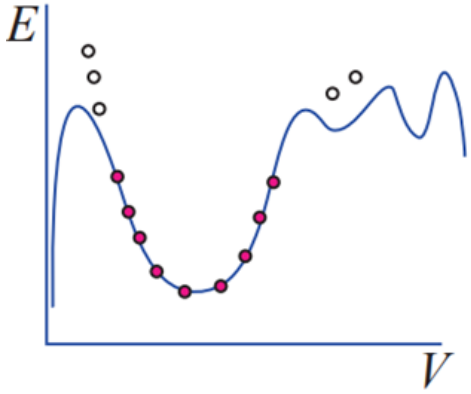
Miao and Yuan, Phys. Chem. Chem. Phys. 25, 7487-7495 (2023)

Why machine learned interatomic potentials (MLIPs)?

Classical force-fields have difficulties in modelling ‘complex’ potential energy surfaces

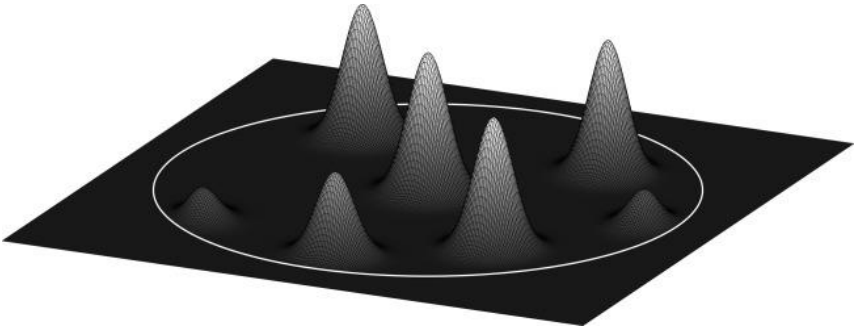
- Diversity of species and bonding environments
- Limited accuracy vs. DFT





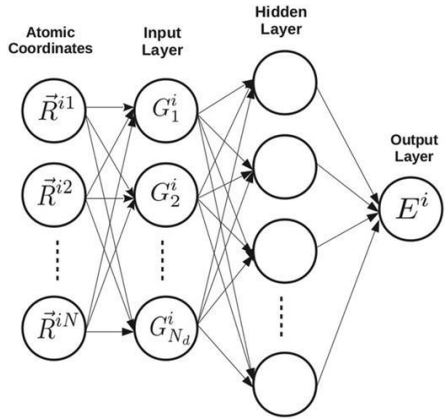
MLIPs: Flexible functional form

- Can handle diversity of species and bonding environments
- Introduce permutation, rotation invariance
- Improved accuracy vs. DFT compared to classical force-fields



Bartók and Csányi, Int. J. Quantum Chem. 116, 1049 (2016)

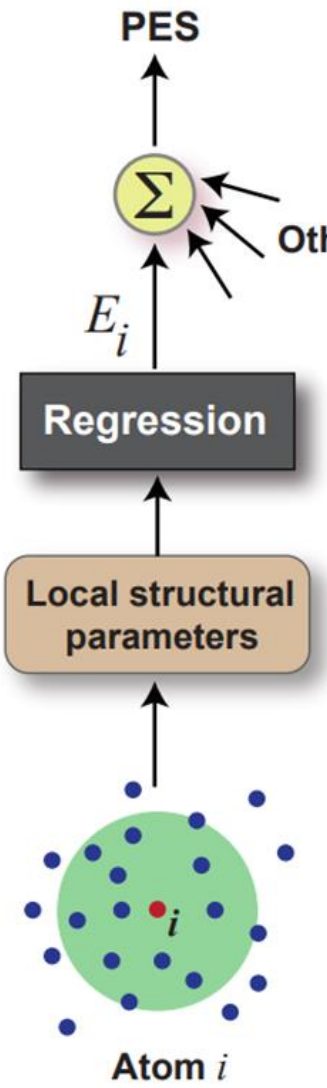
Mishin, Acta Mater. 214, 116980 (2014)



Kocer et al., J. Chem. Phys. 150, 154102 (2019)

Fingerprint a local environment around a reference atom
+ machine-learning model
= (classic) MLIP

How do classical MLIPs work?

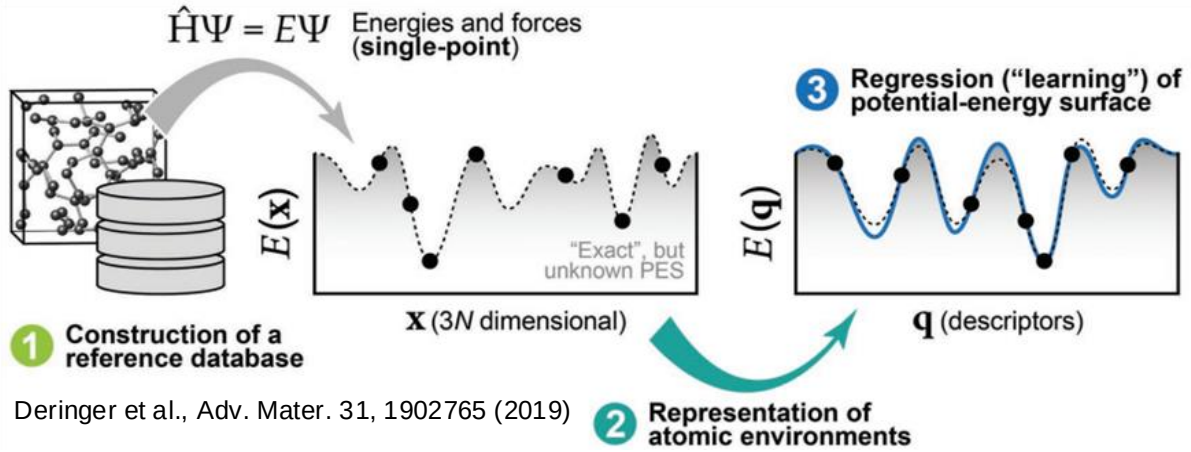


PES: Potential energy surface = Sum of atomic energies

Weights, biases, and hyperparameters

Fingerprint the local environment

Define a neighborhood of interest for each atom



Deringer et al., Adv. Mater. 31, 1902765 (2019)

Typically MLIPs are trained on total energies, atomic forces, and lattice stresses of several different structures in a chemical space

Popular MLIPs:
- Artificial neural network potential (ANNP)
- Gaussian approximation potential (GAP)
- Moment tensor potential (MTP)
- Spectral neighbor analysis potential (SNAP)

Mishin, Acta Mater. 214, 116980 (2014)

Breakdown of classical MLIPs

MLIP	Moment tensor potential (MTP)	Spectral neighbor analysis potential (SNAP) Or Quadratic SNAP (qSNAP)	Gaussian approximation potential (GAP)	Artificial neural network potential (ANNP)
Idea	Many-body interactions represented via moment-tensors	Local atomic density projected on a 4D hypersphere	Local atomic density modelled via smooth overlap of atomic positions	Local environment as input layer in feed-forward neural network
Descriptor	Moment tensors	Bispectrum components	Weighted sum of Gaussians	Radial and angular distribution functions (or symmetry functions)
Training algorithm	BFGS	Linear	Gaussian process	L-BFGS
Basis functions	Chebychev	Hyperspherical harmonics	Gaussians, spherical harmonics	Chebychev

’Easy’ to train
‘Fast’ to run (CPUs)
Reasonably accurate on system trained

Don’t generalize well, ‘short-sighted’
Poor learning rate (need ‘large’ data)
Poor scaling with number of elements

Breakdown of classical MLIPs

MLIP	Moment tensor potential (MTP)	Spectral neighbor analysis potential (SNAP) Or Quadratic SNAP (qSNAP)	Gaussian approximation potential (GAP)	Artificial neural network potential (ANNP)
Idea	Many-body interactions represented via moment-tensors	Local atomic density projected on a 4D hypersphere	Local atomic density modelled via smooth overlap of atomic positions	Local environment as input layer in feed-forward neural network
Descriptor	Moment tensors	Bispectrum components	Weighted sum of Gaussians	Radial and angular distribution functions (or symmetry functions)
Training algorithm	BFGS	Linear	Gaussian process	L-BFGS
Basis functions	Chebychev	Hyperspherical harmonics	Gaussians, spherical harmonics	Chebychev

Let’s take a deeper look at MTP

Moment tensor potential

$$E^{\mathrm{mtp}}(\mathrm{cfg}) = \sum_{i=1}^{n} V(\mathfrak{n}_i)$$

n_i - atomic environment (within a cut-off radius) comprising of a reference atom, its neighbours, and their relative positions

V: function invariant to permutations, rotations, and reflections

- Smooth with respect to exchange of atoms from neighborhood

$$V(\mathfrak{n}_i) = \sum_{\alpha} \xi_\alpha B_\alpha(\mathfrak{n}_i)$$

Weights to be fit

Basis functions: written up to a maximum 'level' of 'contracted' moment tensors

Moment tensor:

$$M_{\mu,\nu}(\mathfrak{n}_i) = \sum_j f_\mu(|r_{ij}|, z_i, z_j) \underbrace{\mathbf{r}_{ij} \otimes \ldots \otimes \mathbf{r}_{ij}}_{\nu \text{ times}}$$

$$\mathrm{lev} M_{\mu,\nu} = 2 + 4\mu + \nu$$

$$\mathrm{lev}(\mathrm{M}_{1,2}{:}M_{0,2}) = (2+4+2) + (2+0+2) = 12$$

Radial component

Angular component

Expanded via radial basis functions: pair-wise

$$f_\mu(|r_{ij}|, z_i, z_j) = \sum_{\beta=1}^{N_Q} c^{(\beta)}_{\mu,z_i,z_j} Q^{(\beta)}(|r_{ij}|)$$

Weights to be fit

Chebychev polynomials ×smooth cut-off function

Expanded via tensors: many-body

$\nu = 0$ → Scalar

$\nu = 1$ → Vector; $\mathbf{r}_{ij} = (x_{ij}, y_{ij}, z_{ij})$

$\nu = 2$ → Tensor; $\mathbf{r}_{ij} \otimes \mathbf{r}_{ij} = \begin{pmatrix} x_{ij}^2 & x_{ij}y_{ij} & x_{ij}z_{ij} \\ y_{ij}x_{ij} & y_{ij}^2 & y_{ij}z_{ij} \\ z_{ij}x_{ij} & z_{ij}y_{ij} & z_{ij}^2 \end{pmatrix}$

Novikov et al. Mach. Learn.: Sci. Technol. 2, 025002 (2021)

MTP: fitting

$$\sum_{k=1}^{K}\left[w_e\left(E^{\mathrm{mtp}}(\mathrm{cfg}_k;\theta)-E^{\mathrm{qm}}(\mathrm{cfg}_k)\right)^2+w_f\sum_{i=1}^{N_k}\left|\mathbf{f}_i^{\mathrm{mtp}}(\mathrm{cfg}_k;\theta)-\mathbf{f}_i^{\mathrm{qm}}(\mathrm{cfg}_k)\right|^2+w_s\left|\sigma^{\mathrm{mtp}}(\mathrm{cfg}_k;\theta)-\sigma^{\mathrm{qm}}(\mathrm{cfg}_k)\right|^2\right]\rightarrow\min_{\theta},$$

Set of *k* configurations in the training set

θ: parameters to be fit (ξ, c)

qm: DFT or other quantum mechanical tools

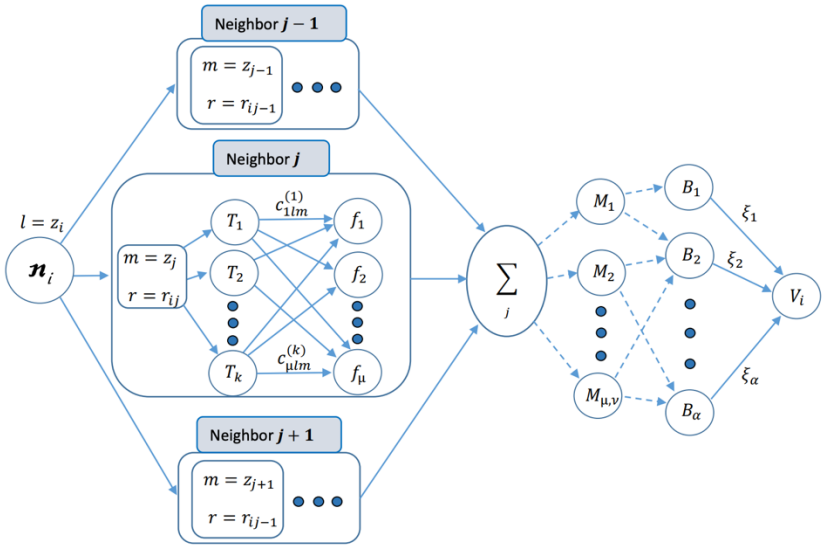
Energies, forces, and stresses considered within loss function

Hyperparameters

$$\mathrm{RMSE}(E)^2=\frac{1}{K}\sum_{k=1}^{K}\left(\frac{E^{\mathrm{mtp}}(\mathrm{cfg}_k;\theta)}{N^{(k)}}-\frac{E^{\mathrm{qm}}(\mathrm{cfg}_k)}{N^{(k)}}\right)^2,$$

$$\mathrm{RMSE}(\mathbf{f})^2=\frac{1}{K}\sum_{k=1}^{K}\frac{1}{3\,N^{(k)}}\sum_{i=1}^{N_k}\left|\mathbf{f}_i^{\mathrm{mtp}}(\mathrm{cfg}_k;\theta)-\mathbf{f}_i^{\mathrm{qm}}(\mathrm{cfg}_k)\right|^2$$

$$\mathrm{RMSE}(\sigma)^2=\frac{1}{K}\sum_{k=1}^{K}\frac{1}{9}\left|\sigma^{\mathrm{mtp}}(\mathrm{cfg}_k;\theta)-\sigma^{\mathrm{qm}}(\mathrm{cfg}_k)\right|^2.$$



Novikov et al. Mach. Learn.: Sci. Technol. 2, 025002 (2021)
https://www.skoltech.ru/app/data/uploads/2019/09/THESIS_FINAL.pdf

MTP: fitting

$$\sum_{k=1}^{K}\left[w_e\left(E^{\mathrm{mtp}}(\mathrm{cfg}_k;\theta)-E^{\mathrm{qm}}(\mathrm{cfg}_k)\right)^2+w_f\sum_{i=1}^{N_k}\left|\mathbf{f}_i^{\mathrm{mtp}}(\mathrm{cfg}_k;\theta)-\mathbf{f}_i^{\mathrm{qm}}(\mathrm{cfg}_k)\right|^2+w_s\left|\sigma^{\mathrm{mtp}}(\mathrm{cfg}_k;\theta)-\sigma^{\mathrm{qm}}(\mathrm{cfg}_k)\right|^2\right]\rightarrow\min_{\theta},$$

Set of *k* configurations in the training set

θ: parameters to be fit (ξ, c)

qm: DFT or other quantum mechanical tools

Energies, forces, and stresses considered within loss function

Hyperparameters

Once MTP is fit, can be used for both static and dynamic runs

- Using 'LAMMPS' for example

Also has ability to perform active learning during predictions

- Using an 'extrapolation grade'
- Structures outside a confidence interval can be calculated with density functional theory and the potential retrained

Novikov et al. Mach. Learn.: Sci. Technol. 2, 025002 (2021)
https://www.skoltech.ru/app/data/uploads/2019/09/THESIS_FINAL.pdf

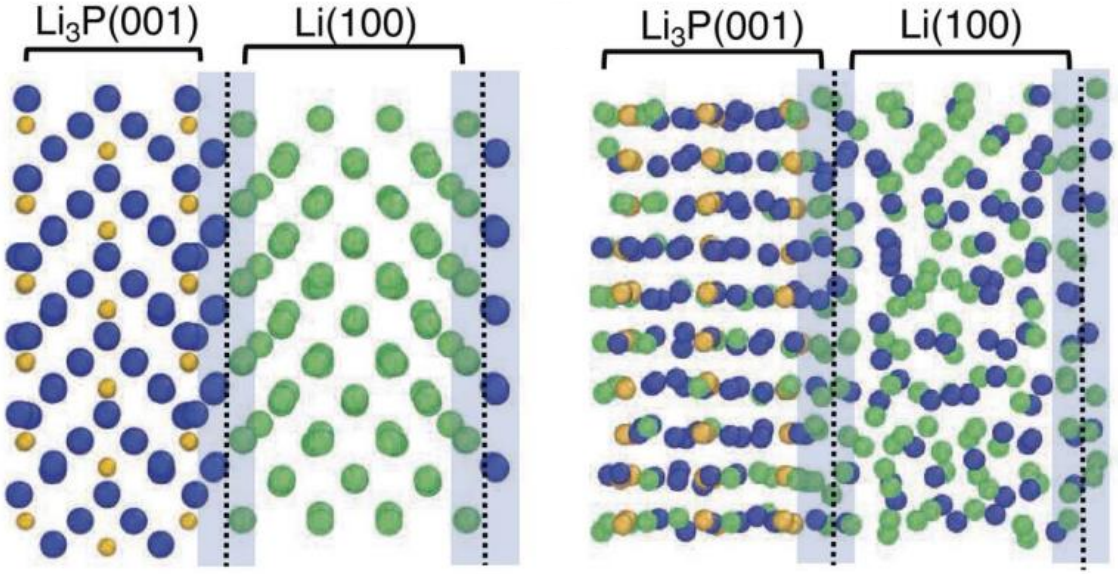
MTP in action

Predicting Li migration energies for cathode coating materials

Composition	MTP E_a (eV)	Experimental E_a (eV)
$Li_3Sc_2(PO_4)_3$	0.62 ± 0.04	0.65
$Li_2B_6O_9F_2$	0.79 ± 0.10	0.92
LiCl	1.11 ± 0.13	0.83

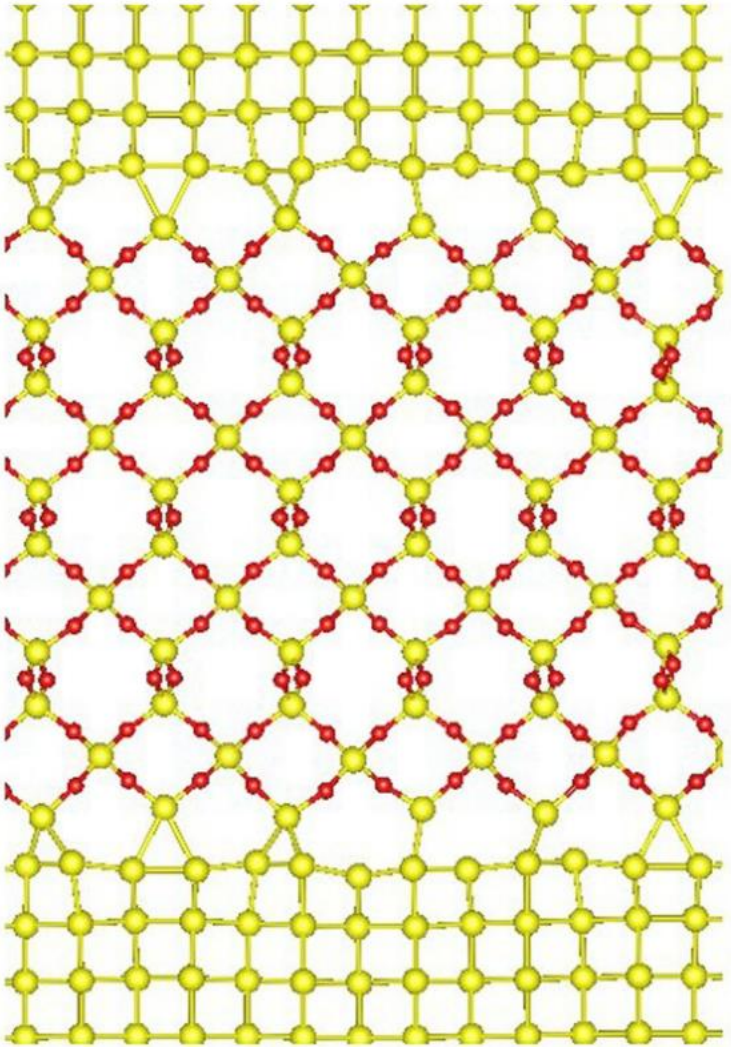
Wang et al., Chem. Mater. 32, 3741–52 (2020)

Modelling solid electrolyte||anode interfaces



Wang et al., J. Mater. Chem. A 10, 19732-19742 (2022)

Simulating Si||SiO_2 interface



Zongo et al., npj Comput. Mater. 10, 218 (2024)

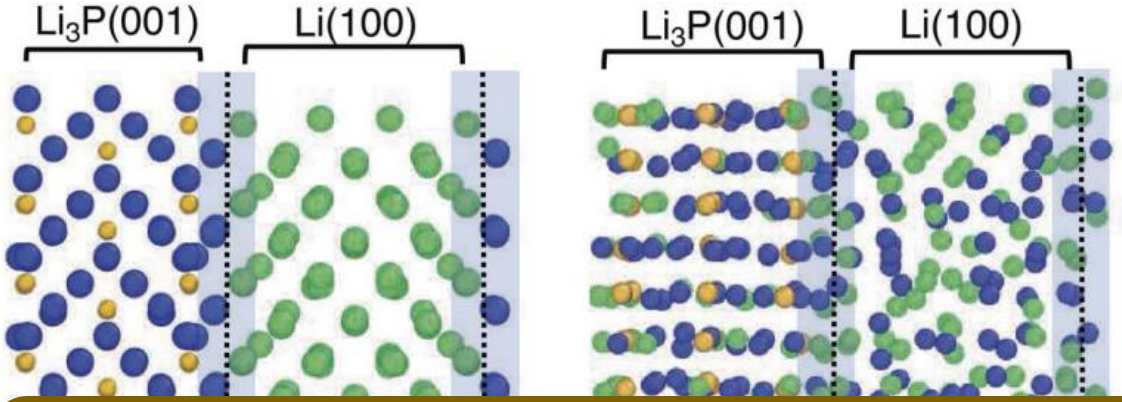
MTP in action

Predicting Li migration energies for cathode coating materials

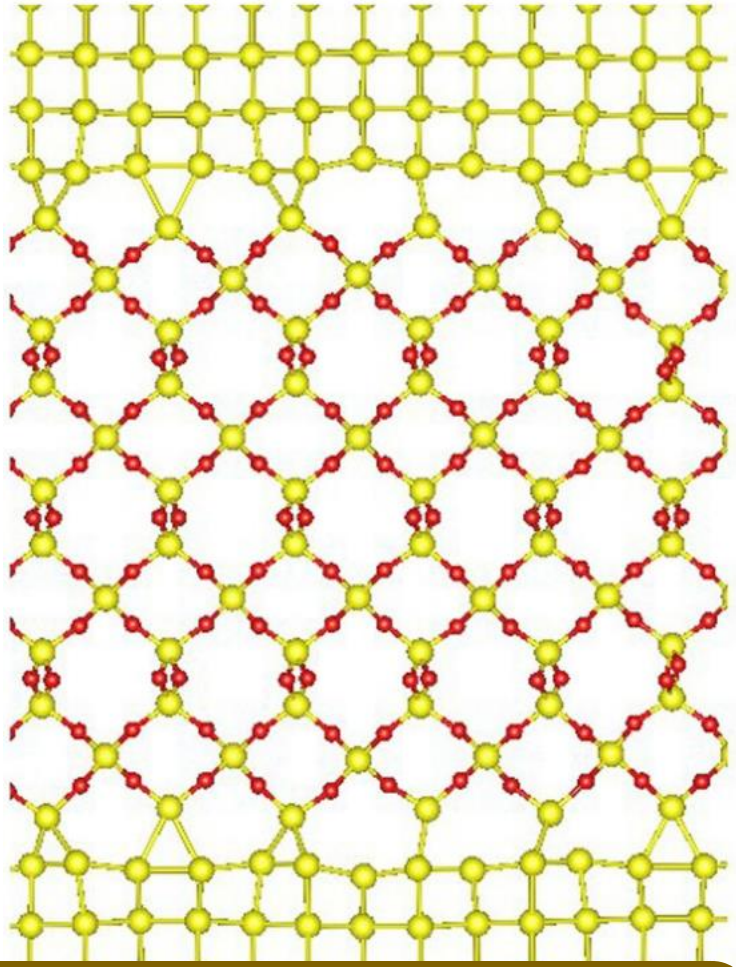
Composition	MTP E_a (eV)	Experimental E_a (eV)
$Li_3Sc_2(PO_4)_3$	0.62 ± 0.04	0.65
$Li_2B_6O_9F_2$	0.79 ± 0.10	0.92
LiCl	1.11 ± 0.13	0.83

Wang et al., Chem. Mater. 32, 3741–52 (2020)

Modelling solid electrolyte||anode interfaces

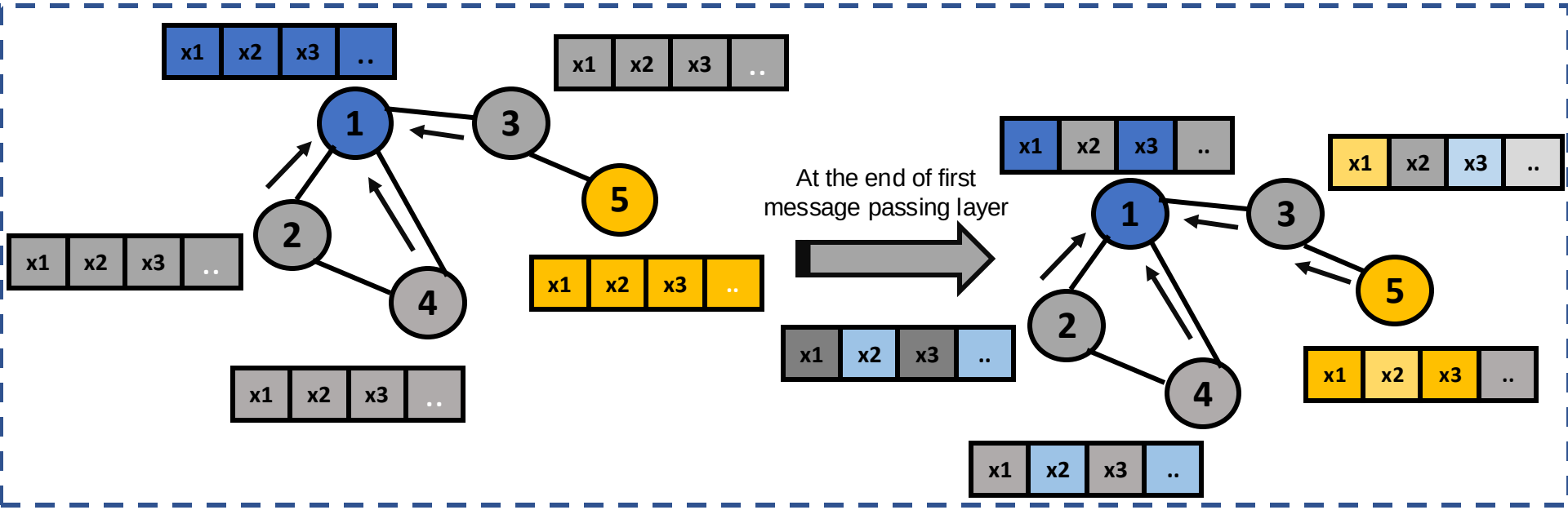


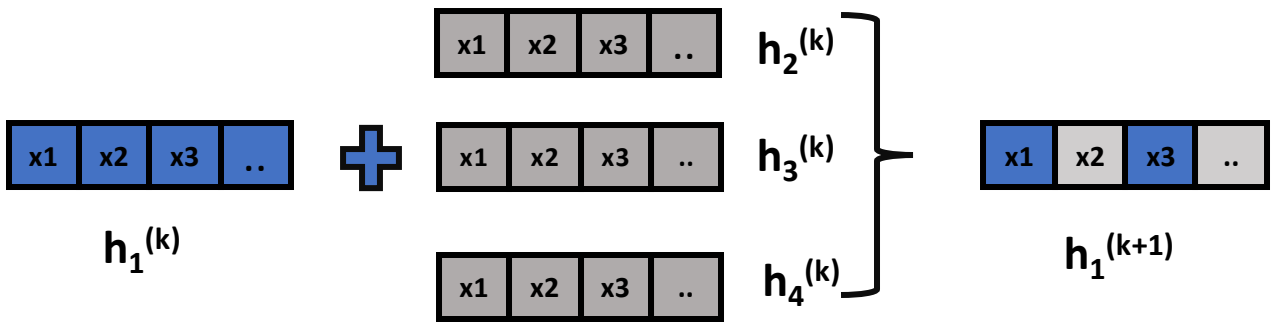
Simulating Si||SiO_2 interface

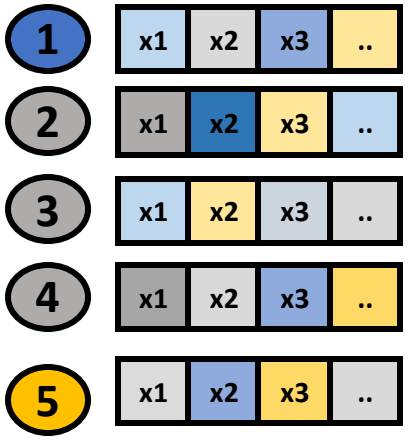


How to mitigate some of the cons of ‘classical’ MLIP? → Use graphs

Recap: graphs and messages

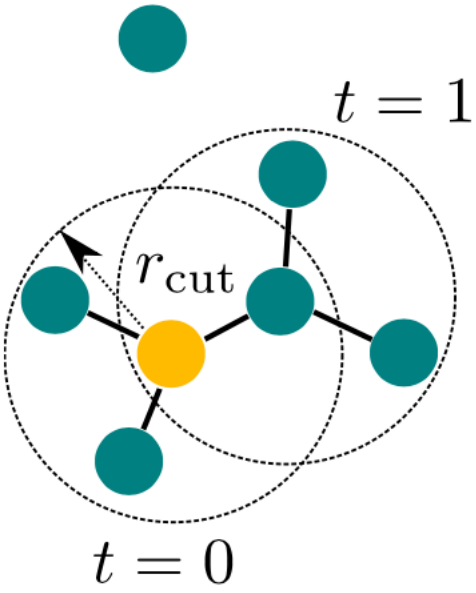




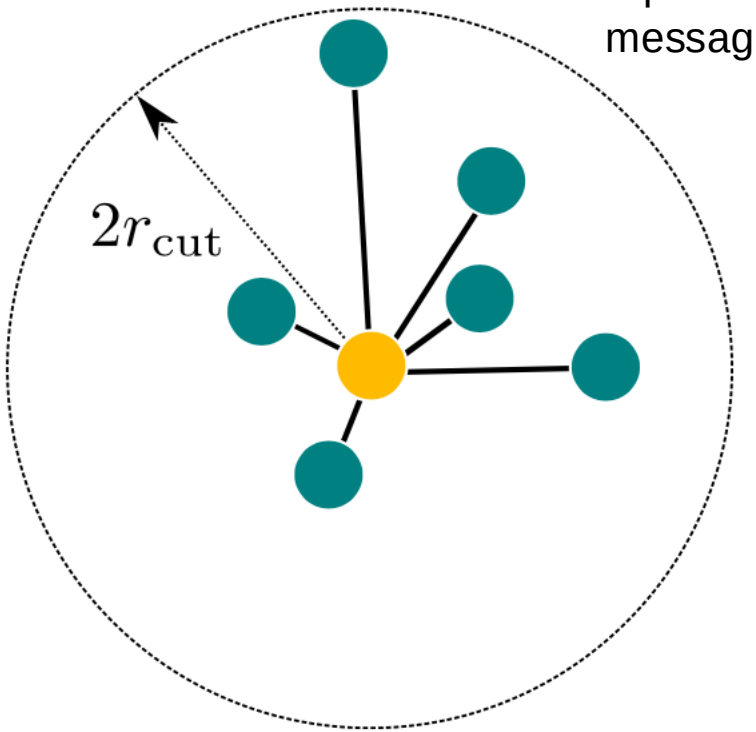


Message passing is quite useful

With message passing
(t: iteration)



Atom-centered representation (without message passing)



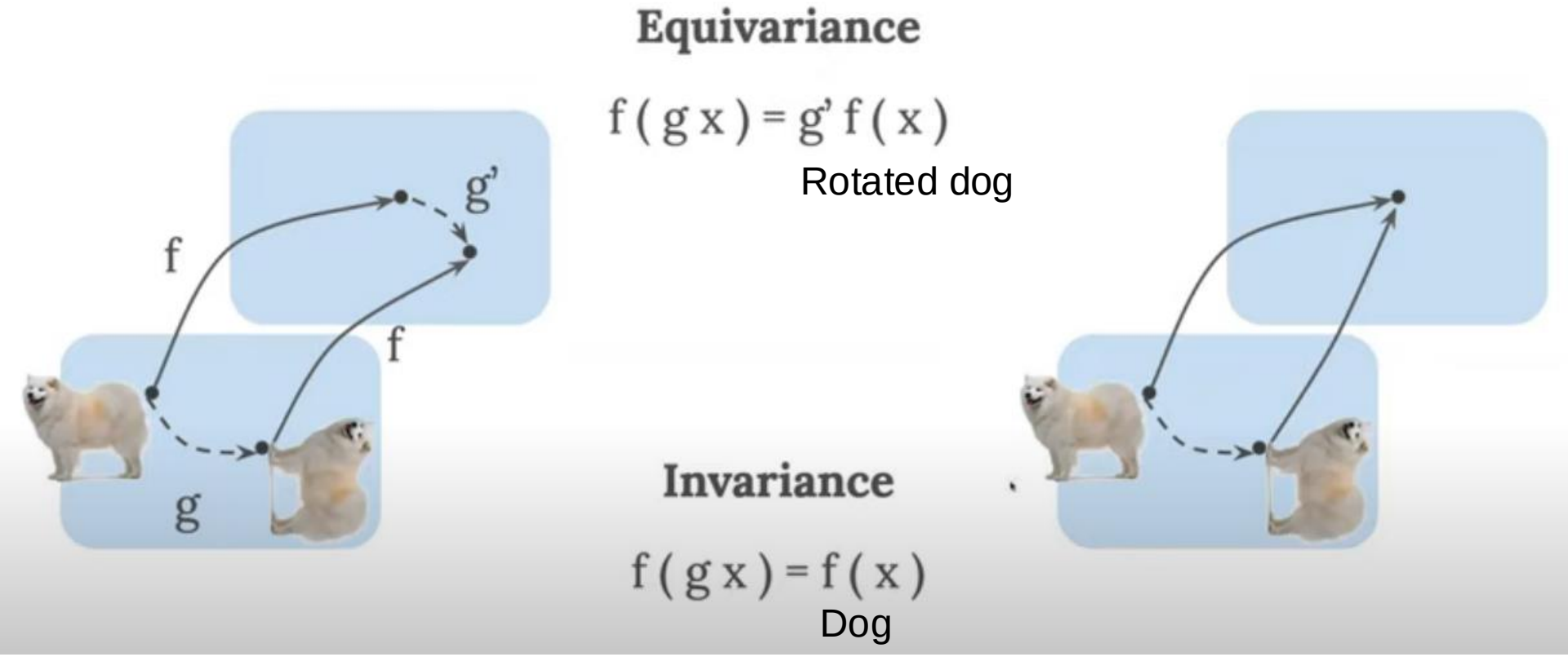
Message passing helps learn long-range interactions

- Effective interaction from $t \times r_{cut}$
- Computationally efficient
- Eliminates unnecessary neighbors

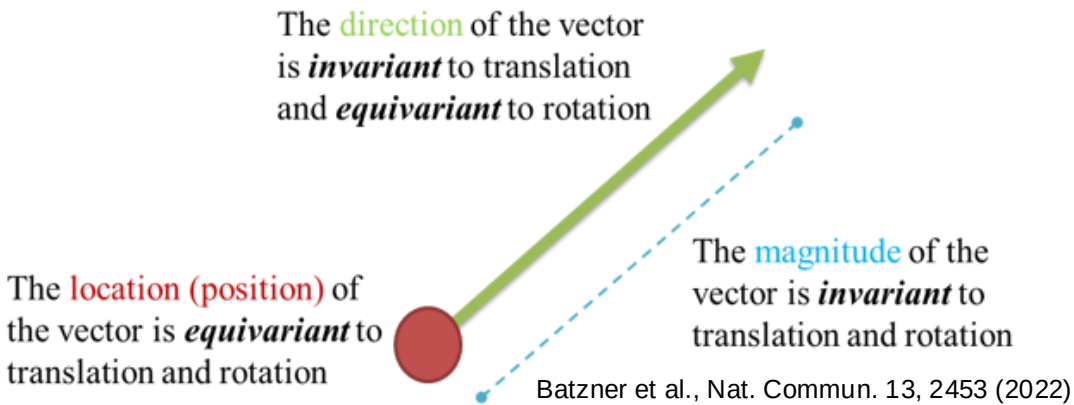
MLIPs incorporating message passing should have higher learning rates and describe longer range interactions better

Figure: Batatia et al.,, arXiv, 2205.06643v2 (2022)

Invariance vs. equivariance



https://datascience.stackexchange.com/questions/16060/what-is-the-difference-between-equivariant-to-translation-and-invariant-to-tr



Batzner et al., Nat. Commun. 13, 2453 (2022)

In materials parlance:

- Scalars (energies) are invariant
- Vectors (forces) and tensors (stresses) are equivariant
- Several useful material properties are equivariant

E(3): Translation, Rotation, Reflection

Introducing equivariance: radial basis and spherical harmonics

State of a node at any iteration (t) in a graph: $\sigma_i^{(t)} = \left(r_i, \theta_i, h_i^{(t)}\right)$

Position Elemental Learnable features

$h_i^{(t)}$ updated through messages ($m_i^{(t)}$) via an update function ($U_t(\sigma_i, m_i^{(t)})$)

Obtained via 'pooling' of neighboring $\sigma_i^{(t)}$ and $\sigma_j^{(t)}$ $\quad \sigma_i^{(t+1)} = \left(r_i, \theta_i, U_t\left(\sigma_i^{(t)}, m_i^{(t)}\right)\right)$

Bake equivariance within $m_i^{(t)}$ by expanding with a basis that is equivariant

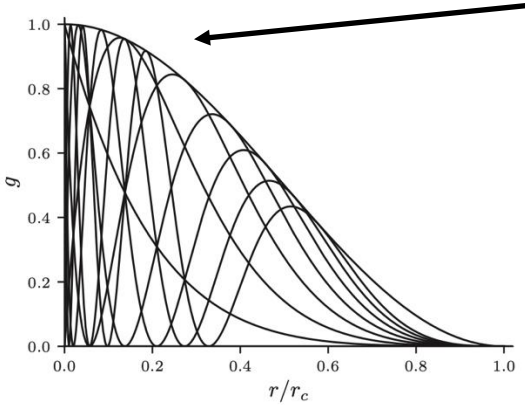
$$m_{i,L}^{(t)}\left(Q \cdot (r_1, r_2, \cdots, r_N)\right) = D^L(Q) m_{i,L}^{(t)}(r_1, r_2, \cdots r_N)$$

Symmetry level of equivariance; Arbitrary rotation; Wigner D-matrix

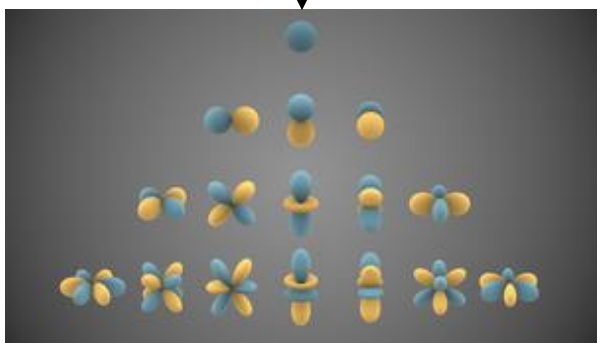
Possible basis: one-particle functions, $\phi_{nlm}(r_{ji}) = R_{nl}(r_{ji}) Y_l^m(\vec{r}_{ji})$

Radial basis: Bessel/Chebychev polynomials with smooth cut-off

Translational equivariance



arXiv 2205.06643v2 (2022); Phs. Rev. B 99, 014104 (2019)



https://en.wikipedia.org/wiki/Spherical_harmonics

Legendre polynomials angular momenta

Rotational and reflectional equivariance

Neural equivariant interatomic potential (NequIP): equivariance + message passing

Based on using deep, graph neural networks to construct interatomic potentials

Every atom has a feature vector of different orders (scalars, vectors, and tensors)

$$E_{pot} = \sum_{i \in N_{atoms}} E_{i,atomic}$$

$$\vec{F}_i = -\nabla_i E_{pot}$$

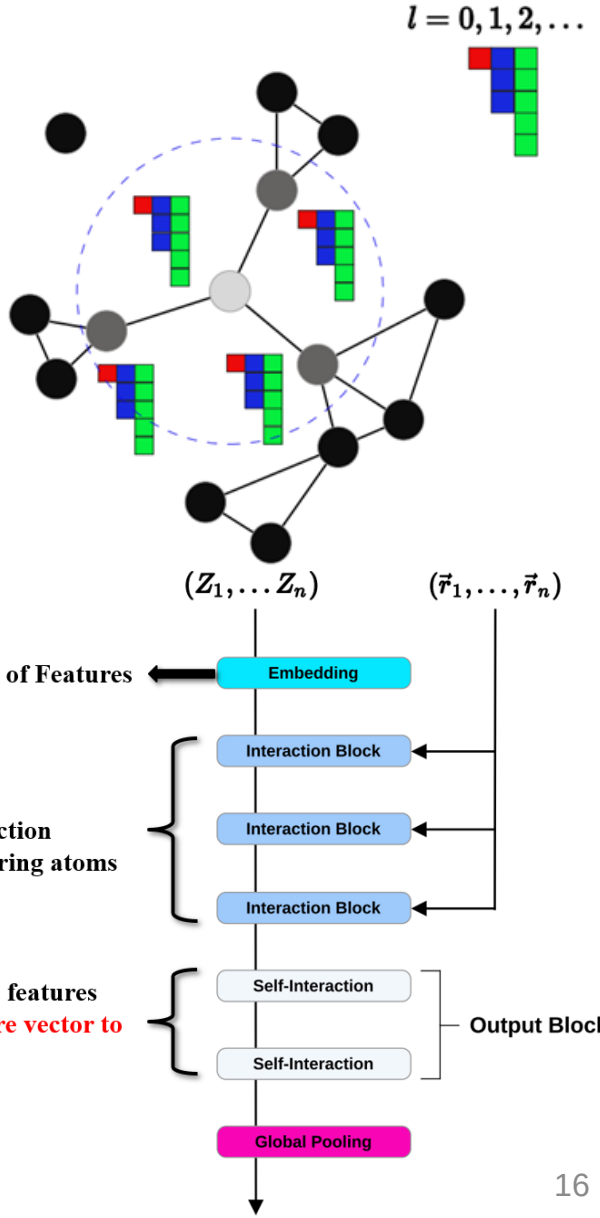
Convolution filters: $S_m^{(l)}(\vec{r}_{ij}) = R(r_{ij}) Y_m^{(l)}(\hat{r}_{ij})$

$$R(r_{ij}) = W_n \sigma(...\sigma(W_2 \sigma(W_1 B(r_{ij}))))$$

$$B(r_{ij}) = \frac{2}{r_c} \frac{\sin(\frac{b\pi}{r_c} r_{ij})}{r_{ij}} f_{env}(r_{ij}, r_c)$$

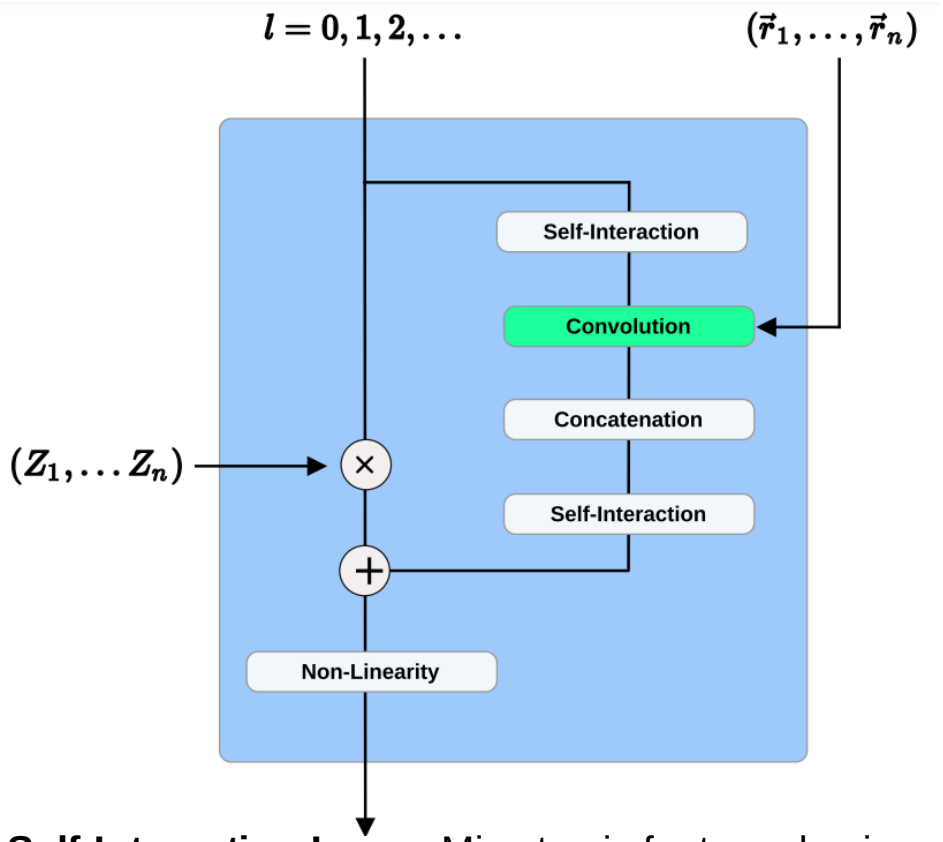
Spherical harmonics

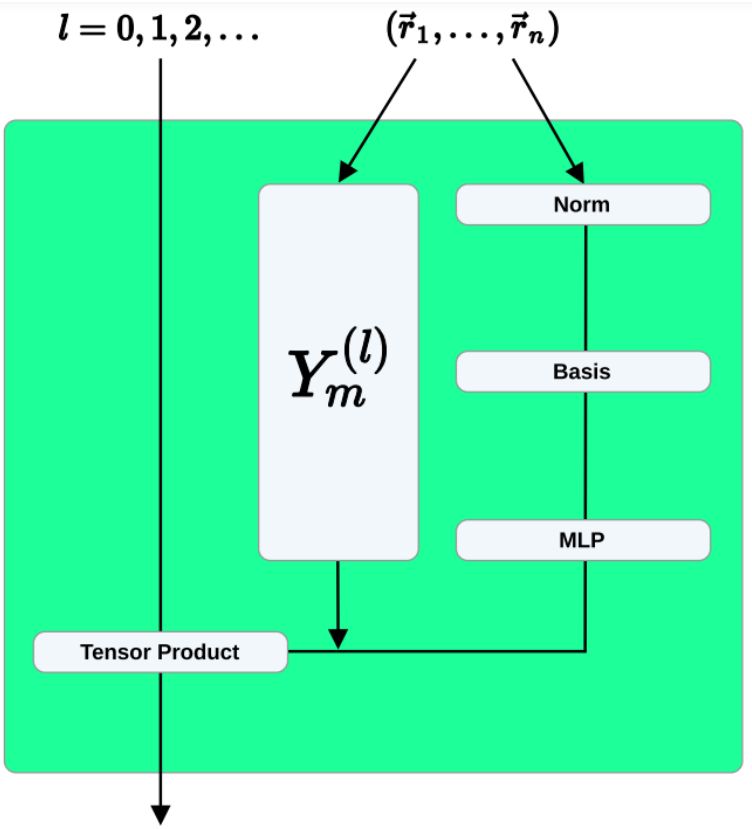
$$\mathcal{L} = \lambda_E ||\hat{E} - E||^2 + \lambda_F \frac{1}{3N} \sum_{i=1}^{N} \sum_{\alpha=1}^{3} \left|\left| -\frac{\partial \hat{E}}{\partial r_{i,\alpha}} - F_{i,\alpha} \right|\right|^2$$



Batzner et al., Nat. Commun. 13, 2453 (2022)

NequIP: code blocks





Self-Interaction Layer: Mix atomic features having same order and mirror parity, reduces dimensionality

Convolution Layer: Rotational equivariance

Concatenation: Recombines feature vectors to form new feature vectors

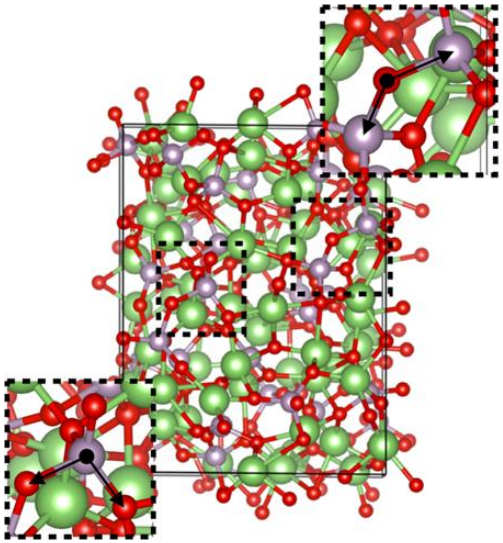
$$R(r_{ij}) = W_n \sigma(\ldots \sigma(W_2 \sigma(W_1 B(r_{ij}))))$$

$$B(r_{ij}) = \frac{2}{r_c} \frac{\sin(\frac{b\pi}{r_c} r_{ij})}{r_{ij}} f_{env}(r_{ij}, r_c)$$

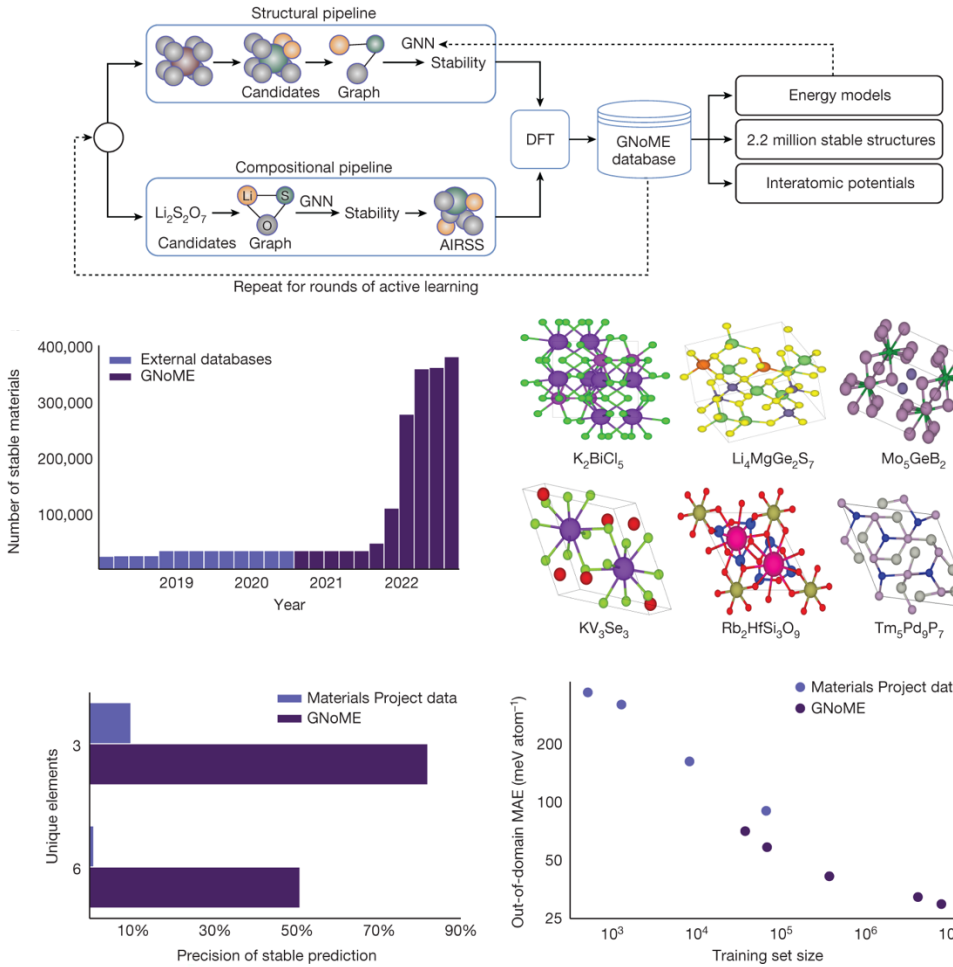
Batzner et al., Nat. Commun. 13, 2453 (2022)

NequIP in action

Simulations of glassy $Li_4P_2O_7$



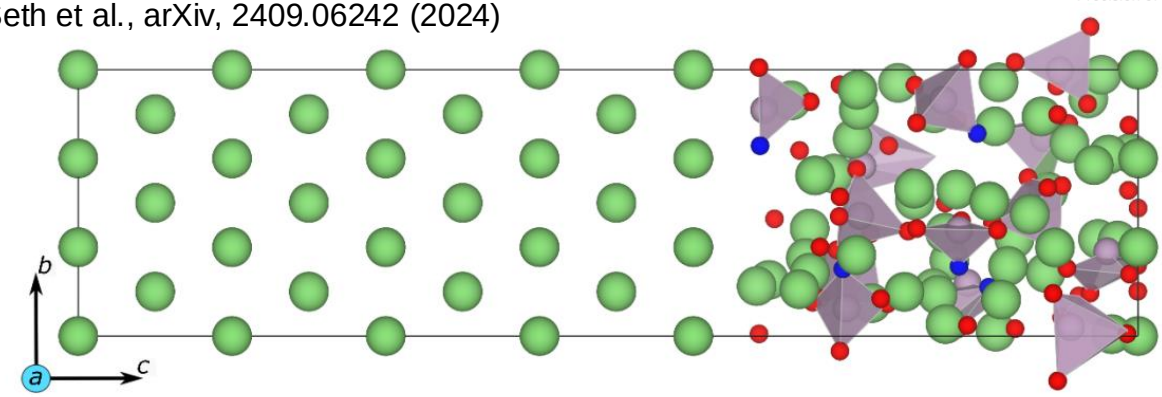
Batzner et al., Nat. Commun. 13, 2453 (2022)



GNoME materials discovery pipeline

Merchant et al., Nature 624, 80-85 (2023)

Seth et al., arXiv, 2409.06242 (2024)



Simulations of amorphous LiPON||Li metal interface

Step back and questions?

- Classical MLIPs: atom-centered, different descriptors to fingerprint local environment, linear/non-linear in energy (and forces/stresses)
  - MTP is linear
  - Generally invariant

- Equivariance and message passing seem important
  - Graph networks for message passing
  - Radial basis and spherical harmonics for equivariance
  - Generally encode 2-body information (bonds)

- Questions?

Towards generalization: Atomic cluster expansion (ACE)

- Are classical and message passing frameworks fundamentally different in construction?

- Are there systematic ways to generate better MLIPs?
  - Or is it just feature engineering or hyperparameter optimization?

- Can MLIPs be foundational?
  - One MLIP for the entire periodic table and the combinations of elements?
  - Replace classical computations with swifter, accurate models?

Need to visit the ACE formalism

ACE: basics

Energy of a collection of atoms can be written (arbitrarily as),

$$E = V_0 + \sum_i V^{(1)}(r_i) + \frac{1}{2}\sum_{ij} V^{(2)}(r_i, r_j) + \frac{1}{3!}\sum_{ijk} V^{(3)}(r_i, r_j, r_k) + \frac{1}{4!}\sum_{ijkl} V^{(4)}(r_i, r_j, r_k, r_l) + \cdots$$

The total energy can be decomposed into atomic contributions,

$$E_i = V^{(1)}(r_i) + \frac{1}{2}\sum_{j} V^{(2)}(r_i, r_j) + \frac{1}{3!}\sum_{jk} V^{(3)}(r_i, r_j, r_k) + \frac{1}{4!}\sum_{jkl} V^{(4)}(r_i, r_j, r_k, r_l) + \cdots$$

Extremely difficult to compute, $K+1$ term costs $\approx N^K$ computational time (N-average # of neighbors)

Need mathematical tricks to make computations tractable → introduce alternate basis functions

Energies depend on bonds (relative positions of atoms) more than absolute positions

- Introduce particle basis → orthogonal and complete by construction
- Group of bonds → cluster → expand energies based on different clusters

$$\Phi_{\alpha\nu} = \phi_{\nu_1}(r_{j_1 i})\phi_{\nu_2}(r_{j_2 i})\cdots\phi_{\nu_K}(r_{j_K i})$$

α is a cluster consisting of K bonds
Each ϕ depends on a distinct bond length
ϕ → expanded with radial basis and spherical harmonics

$$E_i(\sigma) = J_0 + \sum_{\alpha\nu} J_{\alpha\nu}\Phi_{\alpha\nu}(\sigma)$$

σ → configuration of atoms
ν → collection of bonds within cluster α

We haven't solved the computational problem yet, merely transferred it to a ϕ basis!

Drautz, Phys. Rev. B 99, 014104 (2019)

Solving computational complexity: product basis

Slightly modify definition of particle basis, $\phi_{kv} = R_{kcl}(r_{ij})Y_l^m(\overrightarrow{r_{ij}})T_{kc}(\theta_i, \theta_j)$

k: uncoupled indices (element identity and position)
c, l, m: coupled indices (element agnostic, capture of bonds, ensures equivariance)

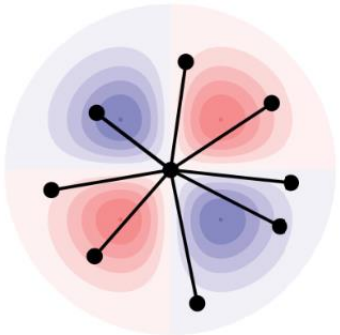
Introduce an atomic basis for each atom, i.e., a density projection

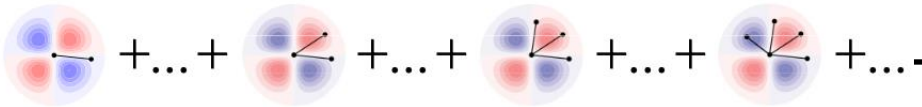
$$A_{i,kv} = \sum_{j \epsilon \mathrm{N}(i)} \phi_{kv}$$

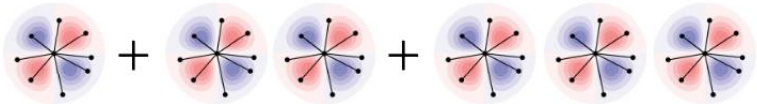
Counts all bonds within a cut-off

Higher body interactions can now be captured via products of A, i.e., a product basis

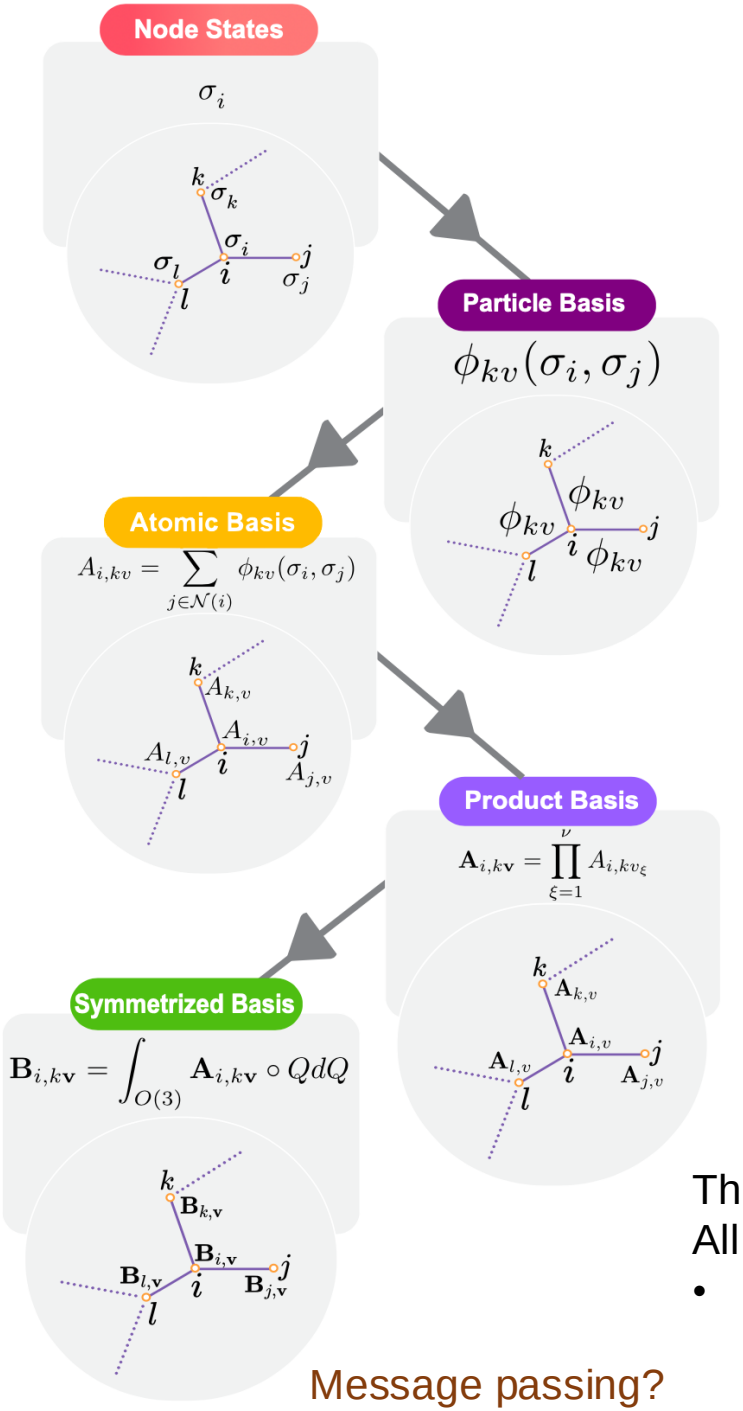
$$A_{i,k\boldsymbol{v}} = \prod_{\zeta=1}^{\nu} A_{i,kv_\zeta} \qquad \boldsymbol{v} = (v_1, v_2, \cdots, v_\nu)$$



$E =$  $\longrightarrow N^K$ scaling

$E =$  $\longrightarrow$ Linear scaling!

Hold on, we aren't done yet!



Change of basis: symmetrized basis

Final change-of-basis: add equivariance to product basis

- Result is the symmetrized basis

$$B_{i,k\boldsymbol{v},\alpha} = \int_{O(3)} (\boldsymbol{D}(Q)^{-1} e_\alpha) \boldsymbol{A}_{i,k\boldsymbol{v}} \left(\{ Q \cdot (\sigma_i, \sigma_j) \}_{j \in \mathcal{N}(i)} \right) dQ$$

In practice, $B_{i,k\eta,LM} = \sum_{\boldsymbol{v}} \mathcal{C}_{\eta,\boldsymbol{v}}^{LM} \boldsymbol{A}_{i,k\boldsymbol{v}}$

Determines which set of A-products are equivariant

Atomic energy can written as a linear expansion of symmetrized basis

- (Linear) ACE

$$E_i(\sigma) = \sum_{\eta,k} \omega_{k\eta L} B_{i,k\eta,LM}$$

Learnable weights

This linear summation of energy → surprisingly similar to MTP!
All classical MLIPs can be derived as special cases of ACE

- Choice of basis, equivariance, and body-order

Message passing + ACE: MACE

Instead of the total energy, define linear summation of symmetrized basis as the message at iteration t

$$m_{i,kLM}^{(t)} = \sum_{\nu} \sum_{\eta_\nu} W_{z_i kL,\eta_\nu}^{(t)} \boldsymbol{B}_{i,\eta_\nu kLM}^{(t)}$$

Update learnable features:

$$h_{i,kLM}^{(t+1)} = U_t^{kL}(\sigma_i^{(t)}, \boldsymbol{m}_i^{(t)}) = \sum_{\tilde{k}} W_{kL,\tilde{k}}^{(t)} m_{i,\tilde{k}LM} + \sum_{\tilde{k}} W_{z_i kL,\tilde{k}}^{(t)} h_{i,\tilde{k}LM}^{(t)}$$

Read-out energies:

$$E_i = E_i^{(0)} + E_i^{(1)} + \ldots + E_i^{(T)}, \qquad \text{where}$$

$$E_i^{(t)} = \mathcal{R}_t\left(\boldsymbol{h}_i^{(t)}\right) = \begin{cases} \sum_{\tilde{k}} W_{\text{readout},\tilde{k}}^{(t)} h_{i,\tilde{k}00}^{(t)} & \text{if } t < T \\ \text{MLP}_{\text{readout}}^{(t)}\left(\left\{h_{i,k00}^{(t)}\right\}_k\right) & \text{if } t = T \end{cases}$$

Multi-ACE (MACE): Message passing on ACE symmetrized basis with linear update functions and non-linear readout functions

Any message passing MLIP can be considered a subset of MACE!

- Choice of basis+update functions, body-order and equivariance

Batatia et al., arXiv, 2205.06643v2 (2022)

Comparison of message passing networks

	SchNet	NequIP	Linear ACE
Message function M_t	$R_k^{(t)}\left(\|\|r_j - r_i\|\|\right) h_{j,k}^{(t)}$	$R_{kl_1l_2L}^{(t)}(r_{ji}) Y_{l_1}^{m_1}(\hat{\boldsymbol{r}}_{\boldsymbol{ji}}) h_{j,kl_2m_2}^{(t)}$	$R_n(r_{ji}) Y_l^m(\hat{\boldsymbol{r}}_{ji}) \delta_{z_i\theta_i}\delta_{z_j\theta_j}$
Symmetric pooling $\bigoplus_{j\in\mathcal{N}(i)}$	$\sum_{j\in\mathcal{N}(i)}$	$\sum_{l_1m_1l_2m_2} \mathcal{C}_{l_1m_1l_2m_2}^{LM} \sum_{j\in\mathcal{N}(i)}$	$\sum_\eta w_\eta \ \sum_{\boldsymbol{v}} \mathcal{C}_{\eta,\boldsymbol{v}}^{00} \prod_{\xi=1}^{\nu} \sum_{j\in\mathcal{N}(i)}$
Update function U_t	$\boldsymbol{h}_i^{(t)} + \tanh\left(W^{(t)}\boldsymbol{m}_i^{(t)} + \boldsymbol{b}^{(t)}\right)$	$\boldsymbol{h}_i^{(t)} + \tanh\left(\|W^{(t)}\boldsymbol{m}_i^{(t)}\|^2\right) W^{(t)}\boldsymbol{m}_i^{(t)}$	-

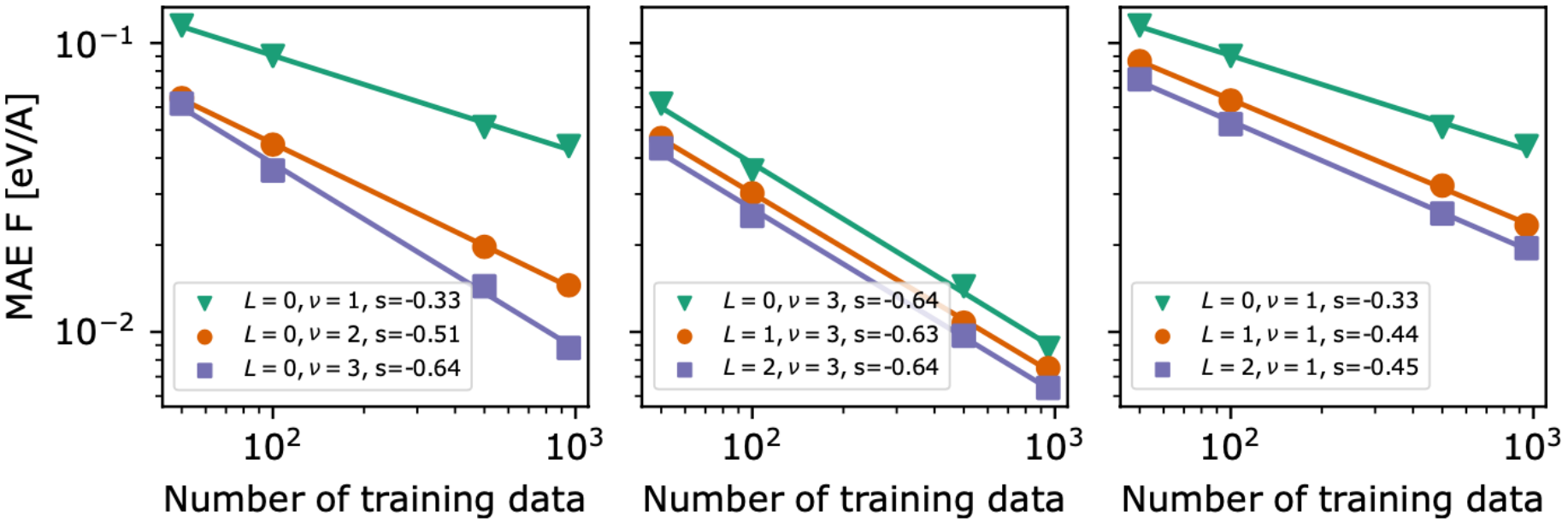
	$l_{\max}$	Update $L_{\max}$	Local correlation order (ν)	Number of layers (T)	Total correlation order	$T_{kc}^{(t)}\left(\boldsymbol{h}_j^{(t)}, \theta_i, \theta_j\right)$	Coupling (v)
SOAP [2]	≥ 3	0	2	1	≥ 3	$\delta_{z_i\theta_i}\delta_{z_j\theta_j}$	nlm
Linear ACE [11]	≥ 1	0	≥ 1	1	≥ 3	$\delta_{z_i\theta_i}\delta_{z_j\theta_j}$	nlm
SchNet [15]	0	0	1	T ≥ 2	T	$h_{j,kl=0}^{(t)}$ (Scalars)	$\emptyset$
DimeNet [17]	0	0	2	T ≥ 2	2T	$h_{j,l=0}^{(t)}$ (Scalars)	$\emptyset$
Cormorant [46]	≥ 1	≥ 1	1	T ≥ 2	T	$h_{j,klm}^{(t)}$ (Spherical Vec.)	lm
NequIP [21]	≥ 1	≥ 1	1	T ≥ 2	T	$h_{j,klm}^{(t)}$ (Spherical Vec.)	$l_1m_1l_2m_2$
GemNet [25]	≥ 1	≥ 1	3	T ≥ 2	T	$h_{j,klm}^{(t)}$ (Spherical Vec.)	$l_1m_1l_2m_2$
NewtonNet [24]	1	1	1	T ≥ 2	T	Cartesian Vectors	-
EGNN [22]	1	1	1	T ≥ 2	T	Cartesian Vectors	-
PaINN [23]	1	1	1	T ≥ 2	T	Cartesian Vectors	-
TorchMD-Net [26]	1	1	1	T ≥ 2	T	Cartesian Vectors	-

$\nu = 1$: Pairs (or 2 body-order) included

Total correlation order ~ $\nu \times T$

MACE: higher body-order is important

Dataset: rMD17



Higher body-order changes the order (slope) of learning rate

- Faster learning with fewer datapoints by including body-order up to 4

Equivariance leads to shift in learning rate, not changing the slope

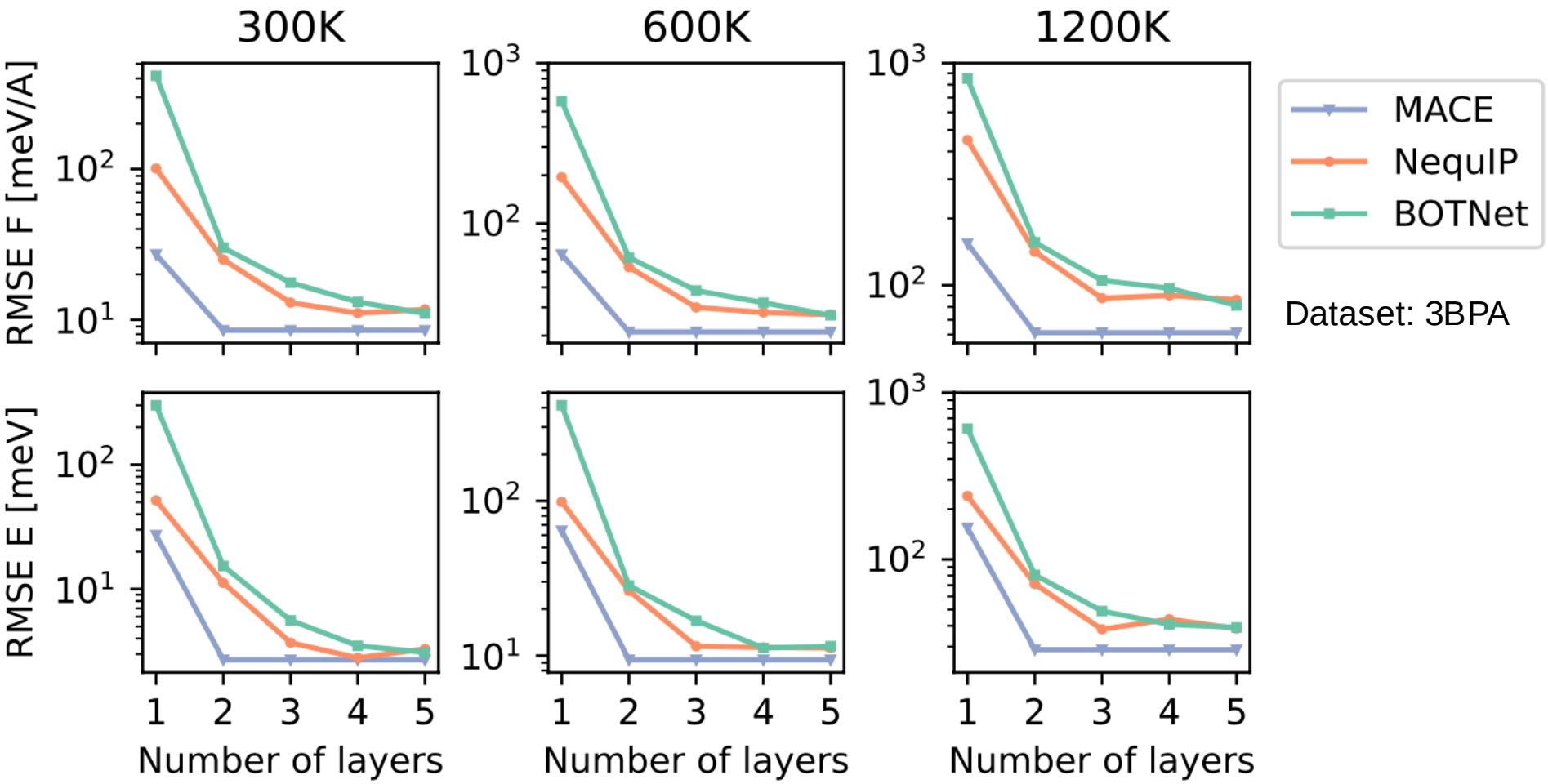
- At high body-order

Equivariance can lead to shift in slope of learning rate at small body order

- Effect of equivariance saturates at the vector-level

Batatia et al., NeurIPS (2022)

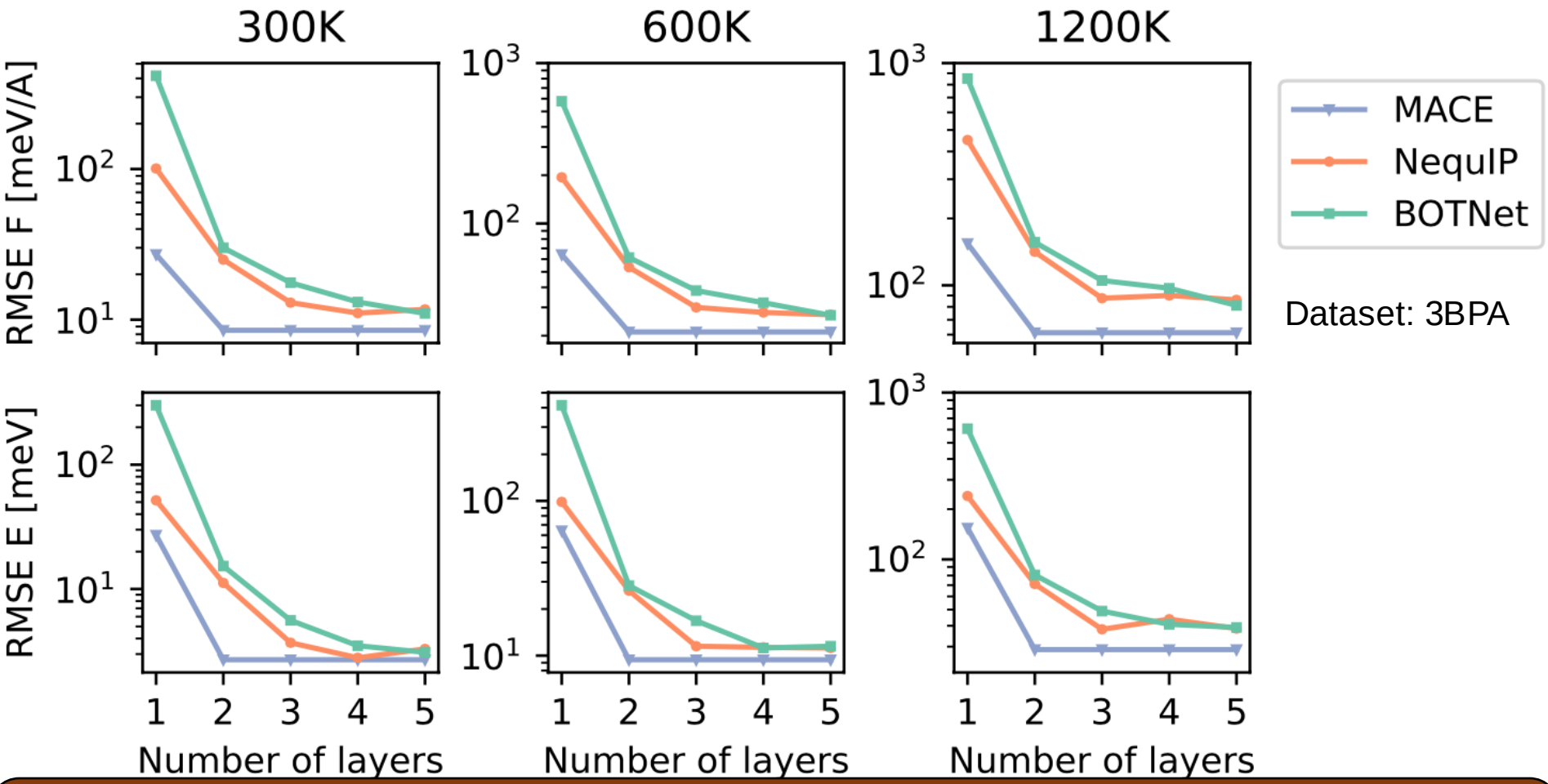
MACE: quicker learning with fewer layers



Fewer (message passing) layers: faster energy/force evaluations, faster molecular dynamics

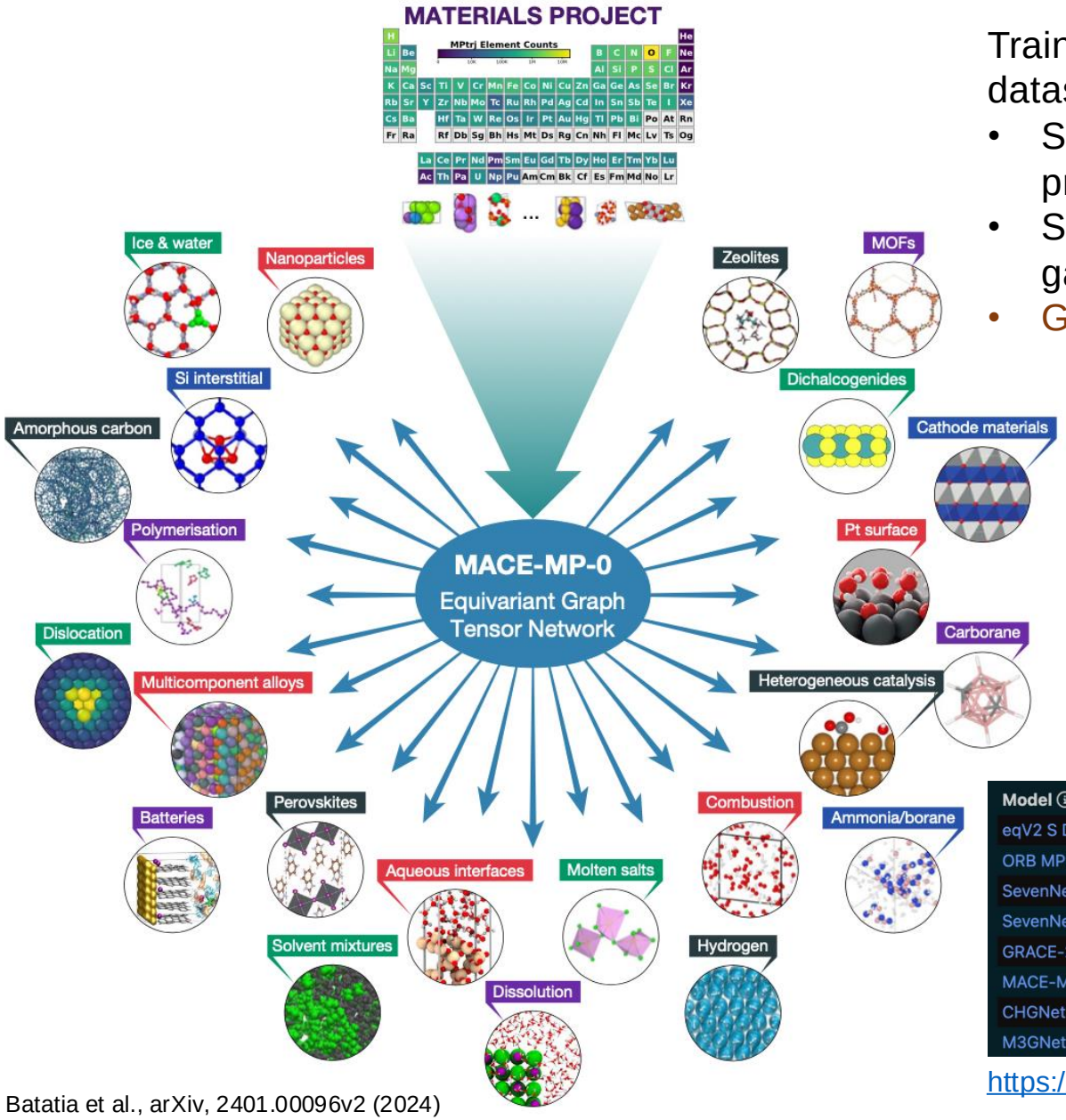
Batatia et al., NeurIPS (2022)

MACE: quicker learning with fewer layers



Given MACE provides a theoretical basis for combining message passing, equivariance and high body-order within MLIPs: can foundational models be constructed?

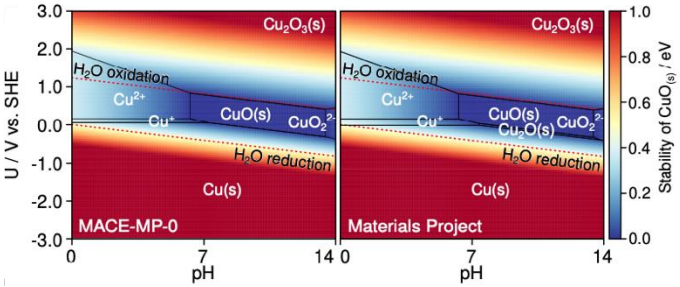
Foundational models: MACE-MP-0



Batatia et al., arXiv, 2401.00096v2 (2024)

Trained on Materials Project trajectory dataset (~1.5M structures)

- Stable performance on 30 different property predictions/application areas
- Stable dynamics in solids, liquids, and gases
- GPU; limited system size



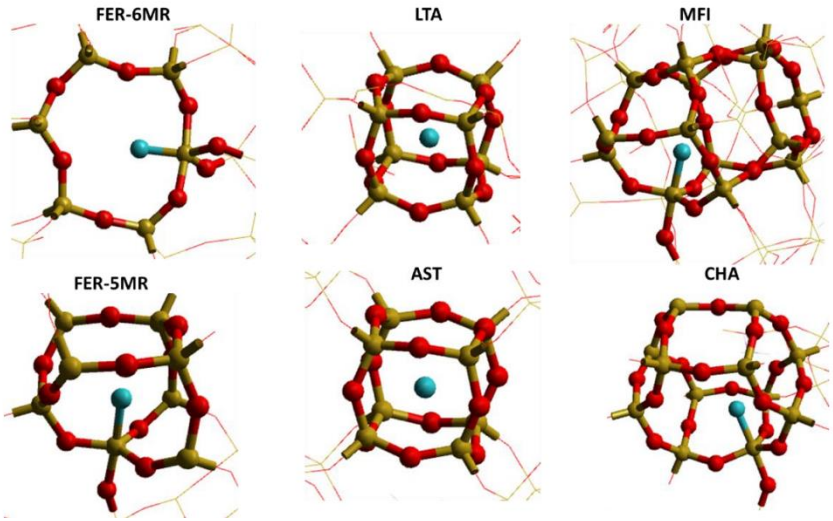
Many more to come…

Model	F1 ↑	DAF ↑	Prec ↑	Acc ↑	MAE ↓	R^2 ↑	κ_{SRME} ↓	Training Set
eqV2 S DeNS	0.815	5.042	0.771	0.941	0.036	0.788	1.665	146k (1.58M) (MPtrj)
ORB MPtrj	0.765	4.702	0.719	0.922	0.045	0.756	1.725	146k (1.58M) (MPtrj)
SevenNet-l3i5	0.76	4.629	0.708	0.92	0.044	0.776	0.55	146k (1.58M) (MPtrj)
SevenNet-0	0.724	4.252	0.65	0.904	0.048	0.75	0.767	146k (1.58M) (MPtrj)
GRACE-2L (r6)	0.691	4.163	0.636	0.896	0.052	0.741	0.525	146k (1.58M) (MPtrj)
MACE-MP-0	0.669	3.777	0.577	0.878	0.057	0.697	0.647	146k (1.58M) (MPtrj)
CHGNet	0.613	3.361	0.514	0.851	0.063	0.689	1.717	146k (1.58M) (MPtrj)
M3GNet	0.569	2.882	0.441	0.813	0.075	0.585	1.412	62.8k (188k) (MPF)

https://matbench-discovery.materialsproject.org/

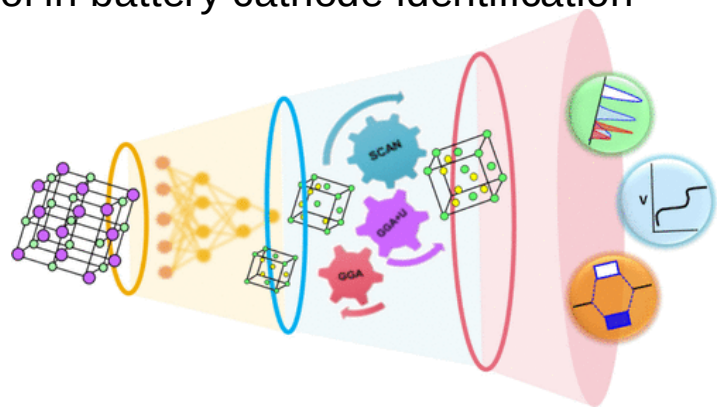
MACE in action

Modelling zeolites (using MACE-ML-IP model)



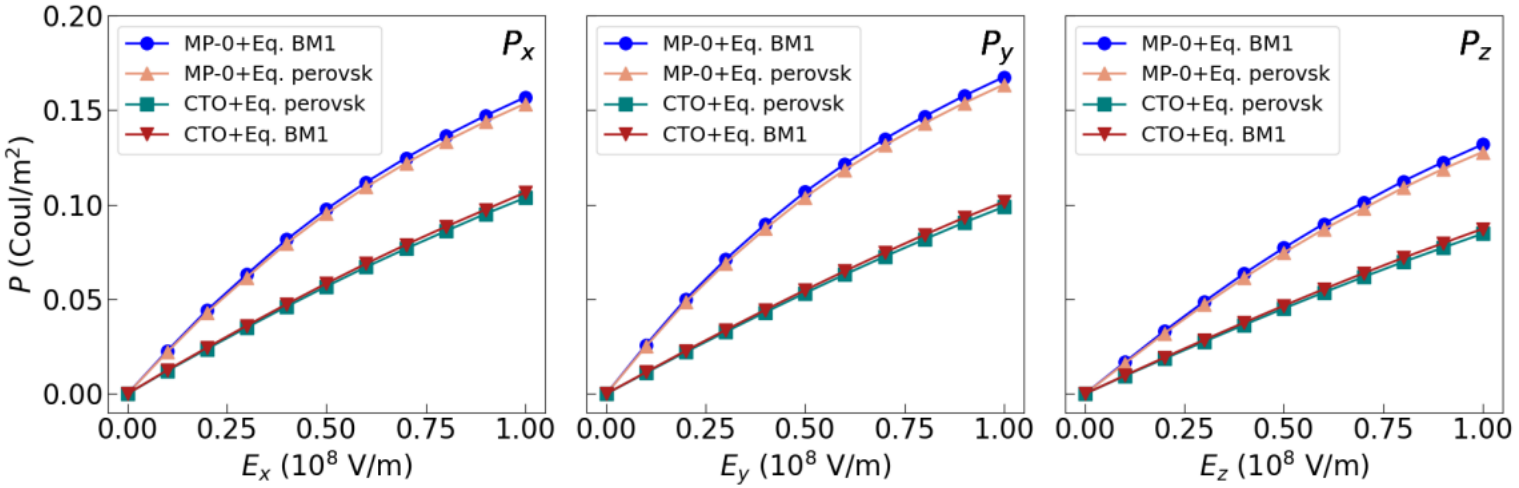
Nasir et al., arXiv, 2411.00436 (2024)

Using MACE-MP-0 as a pre-screening tool in battery cathode identification



Singh et al., ACS Appl. Electron. Mater. 6, 7065-7074 (2024)

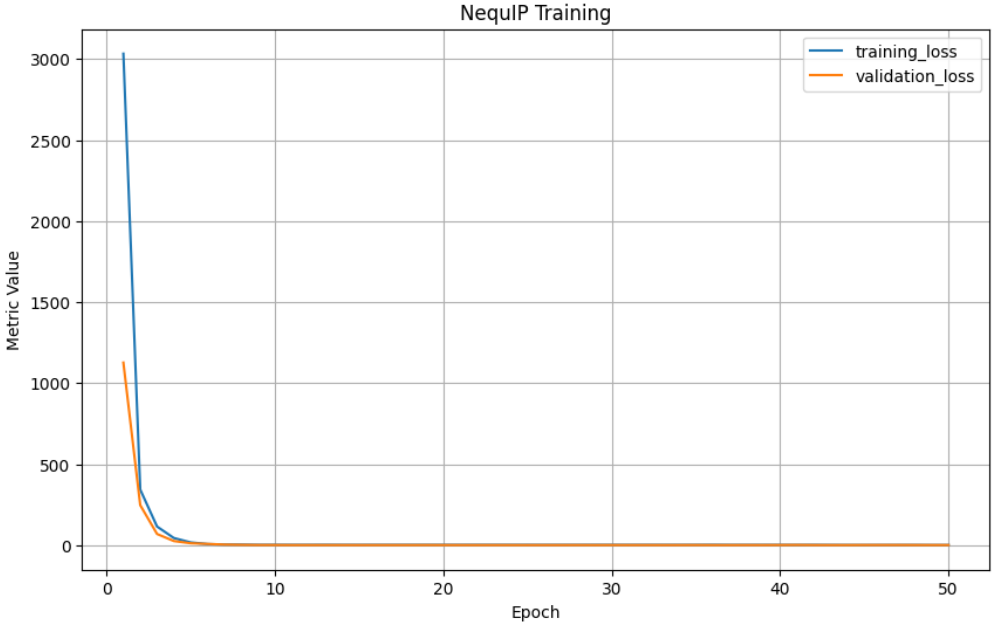
Polarization of $CaTiO_3$ with applied electric field (MACE-MP-0 and custom models)

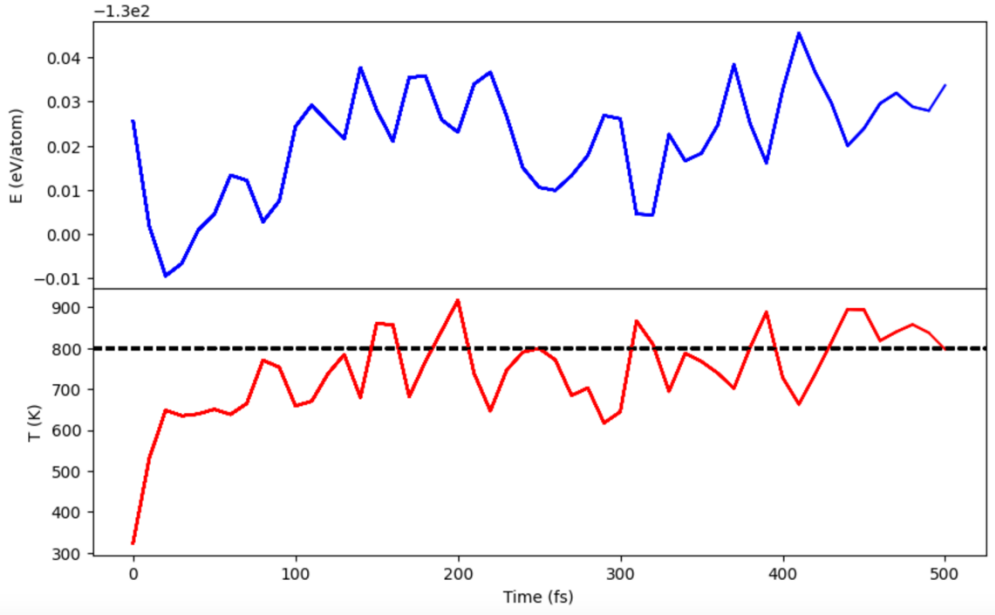


Kutana et al., arXiv, 2412.03541 (2024)

Hands—on session?

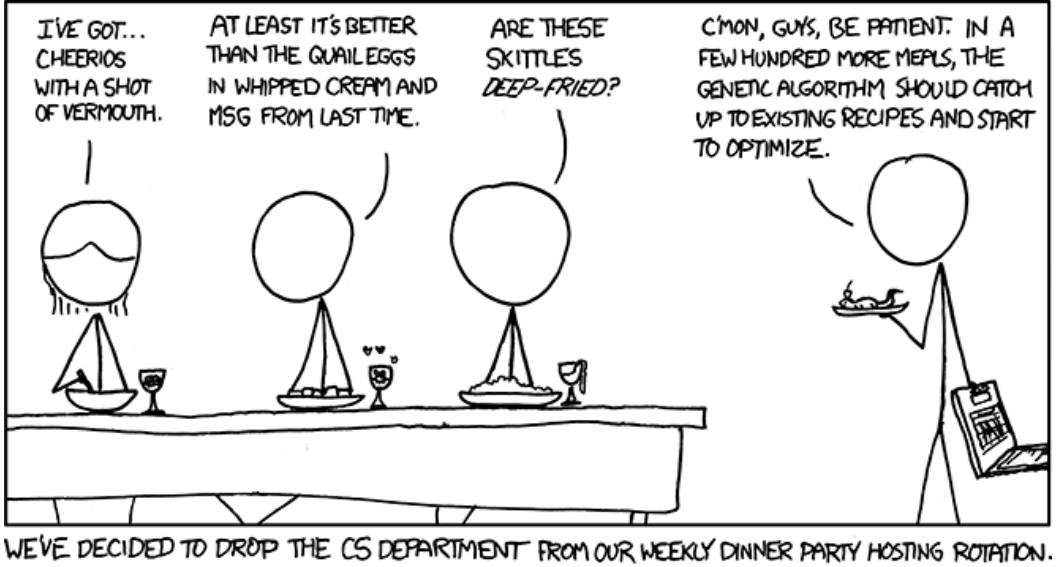
Build NequIP and MACE, run MD





Hopefully get some molecules to explode!

Summary



- Interatomic potentials are important for simulating dynamics in systems
  - MLIPs provide better accuracies than classical force fields at similar computational costs
- Classical MLIPs: atom-centered
  - Accuracy improvement largely due to feature engineering and including non-linearity
- Message passing, equivariance, and high body-order: important for accuracy improvement
  - MACE: offers a platform and systematic ways to create foundational models
  - NequIP: highly efficient message passing, equivariant models for large system sizes

saigautamg@iisc.ac.in; https://sai-mat-group.github.io; https://github.com/sai-mat-group