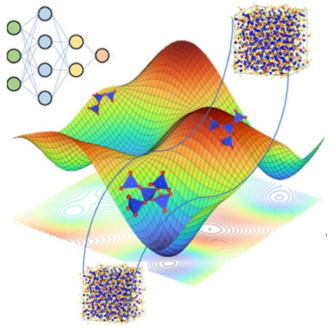
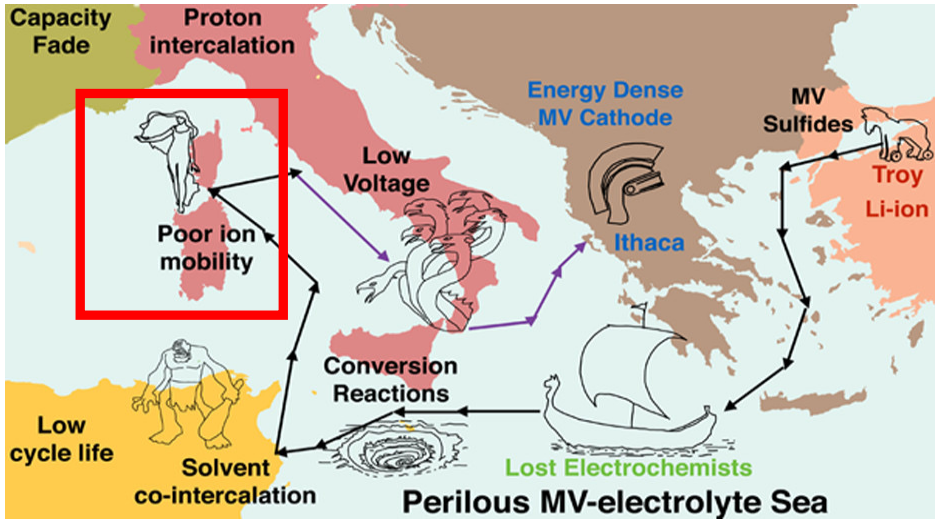
Evaluation of foundational machine learned interatomic potentials for migration barrier predictions

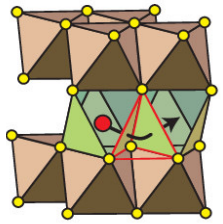
[1]Achinthya Krishna Bheemaguli, [2]Penghao Xiao, [1]Sai Gautam Gopalakrishnan
[1]Materials Engineering, Indian Institute of Science
[2]Physics and Atmospheric Sciences, Dalhousie University

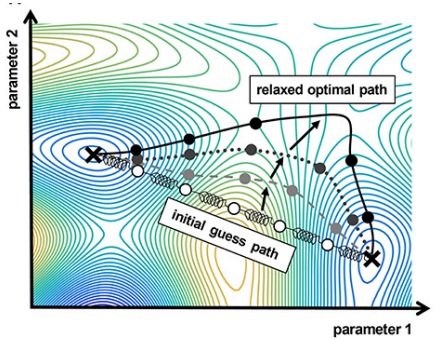
achinthyab@iisc.ac.in

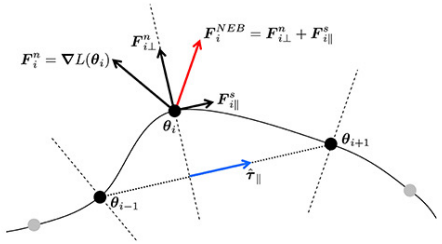


Migration barriers: A central challenge for multivalent cathodes









- Multivalent chemistries (e.g. Mg, Ca-ion etc) offer high volumetric energy density, less toxic and earth earth-abundant
- Ionic Diffusivity (D) within the bulk is a major factor in the rate performance of intercalation electrodes

$$D = D_0 \exp\left(\frac{-Em}{k_B T}\right)$$

D_0: Diffusivity pre-factor
k_B: Boltzmann constant
T: Temperature
E_m: Migration barrier

- Computational technique to predict E_m: ab initio molecular dynamics, nudged elastic band (NEB) with density functional theory (DFT)
- DFT-NEB is computationally expensive and often exhibits convergence difficulties

van der Ven et al. **Ann. Rev. Mater. Res.** 48, 27-55 (2018)
Canepa et al. **Chemical reviews.** 117(5):4287–4341 (2017)

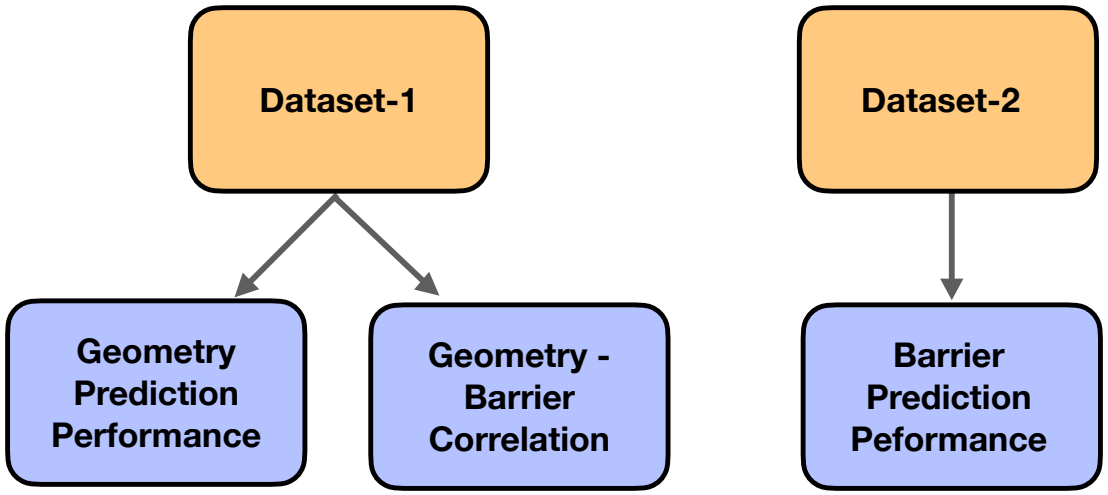
Questions we ask

How accurate are the MLIP-NEB predicted E_m? **(Barrier prediction performance)**

Can MLIP-NEB relaxed image geometries be a better initial guess for subsequent DFT-NEB? **(Geometry prediction performance)**

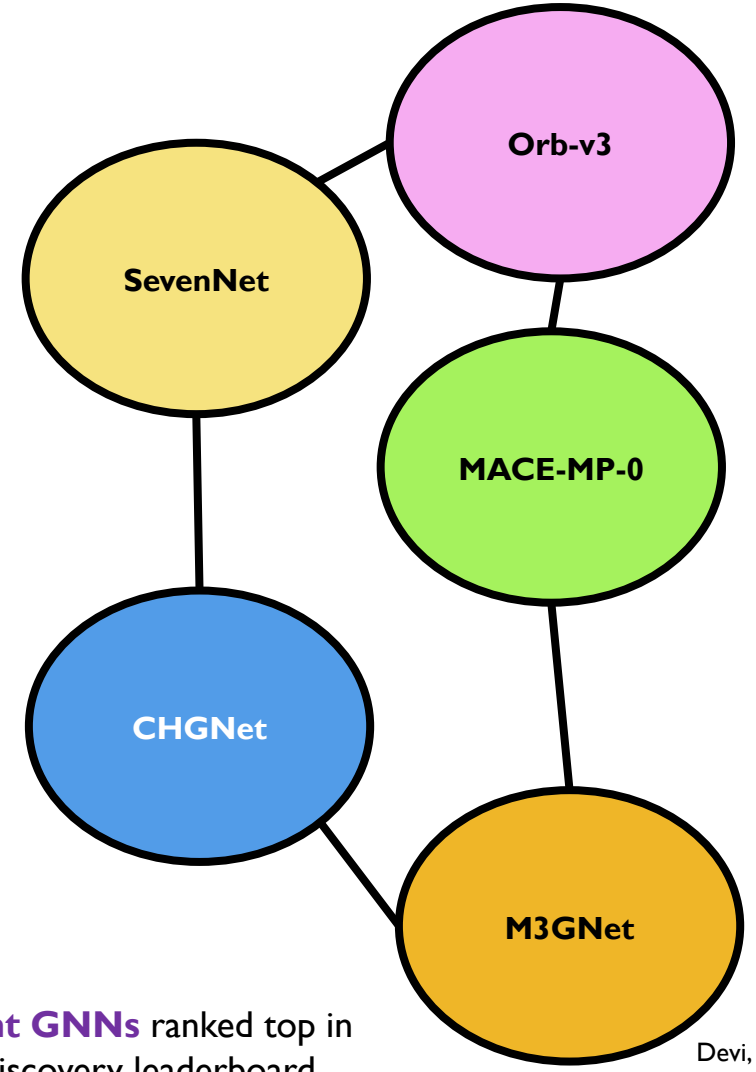
Are geometry and E_m predictions correlated? **(Geometry-barrier correlation)**

Datasets and ground-truth



- **Dataset-1**: 60 DFT-NEB E_m of GGA level accuracy (Includes intermediate relaxed images)
- **Dataset-2** : 574 DFT-NEB E_m of GGA level accuracy compiled from literature
- Both the datasets span over wide variety of battery-relevant chemistries and structures

MLIPs



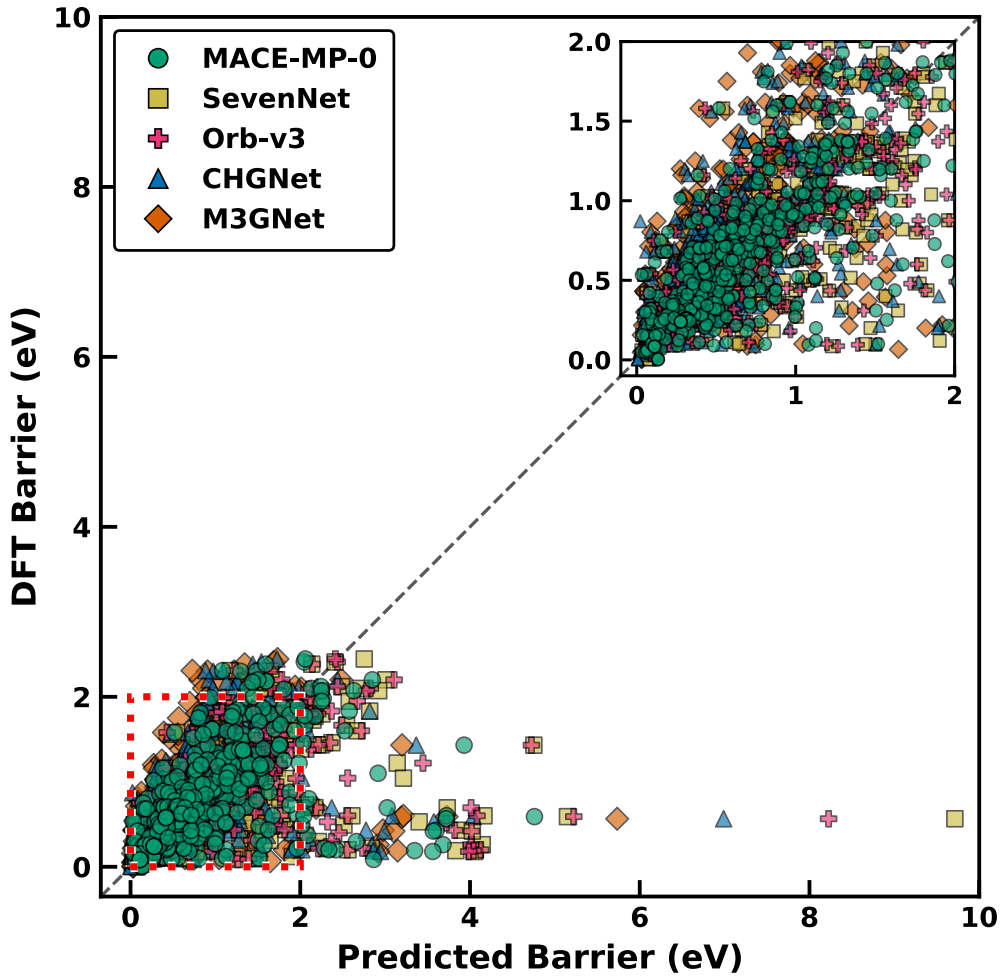
- **Equivariant GNNs** ranked top in matbench discovery leaderboard

NEB parameters

- Image Dependent Pair Potential (IDPP) generated images connected via k = 5 eV/Å^2 spring
- Tangent defined as per Elastic Band (EB) implementation
- Number of images identical to ground-truth
- All calculations run using atomic simulation environment (ASE)

Devi, Reshma et al. **Scientific Data**. 12.1 (2025): 1922

Barrier prediction performance

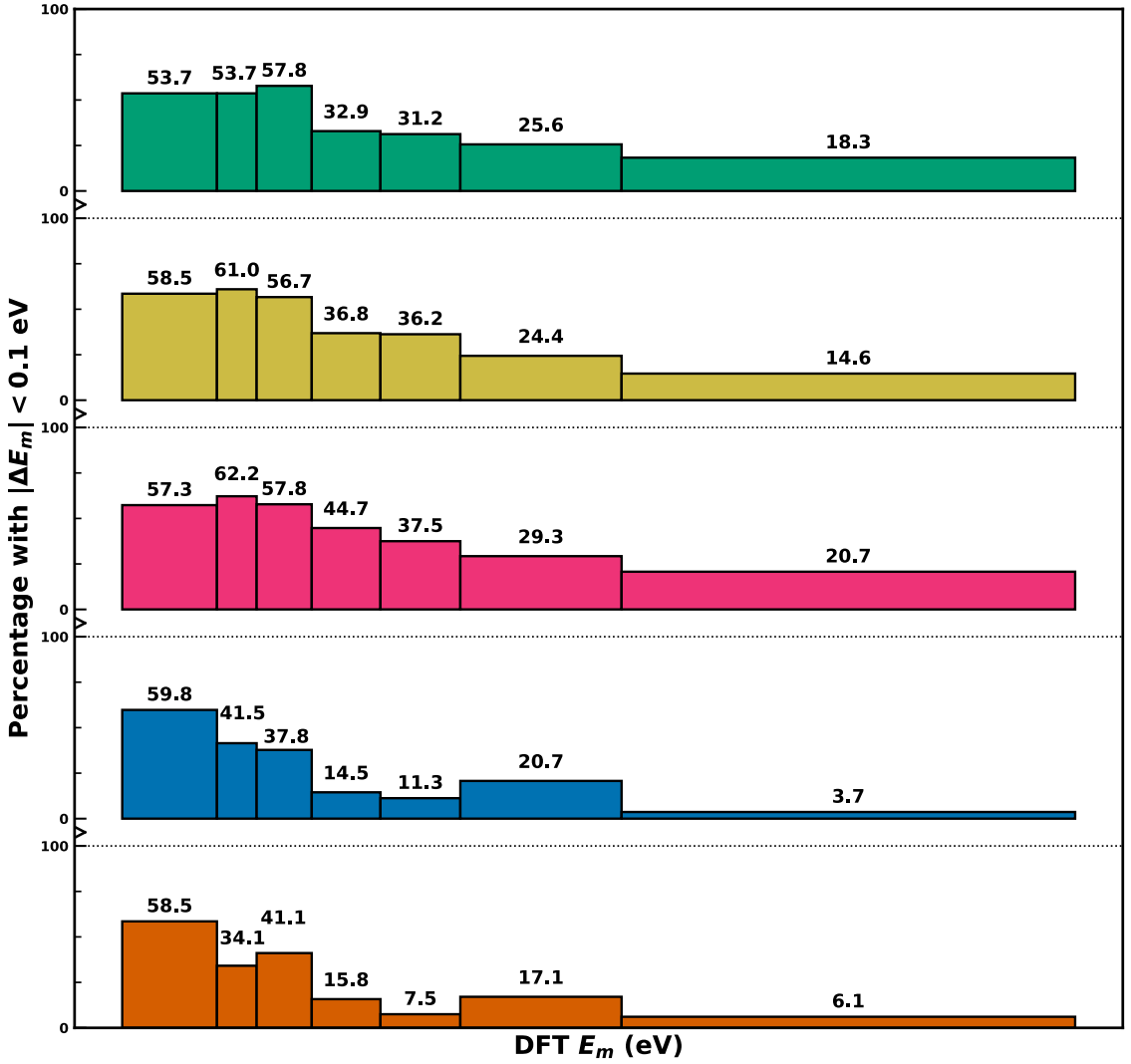


- MACE-MP-0 (M3GNet) shows the lowest (highest) mean absolute error (MAE) of 0.31 eV (0.349 eV) while considering the entire dataset
- MAE excluding the common outliers (> 1eV error) shows the same trend with MACE-MP-0 (M3GNet) exhibiting the lowest (highest) 0.239 eV (0.290 eV)
- Excluding outliers specific to each model:

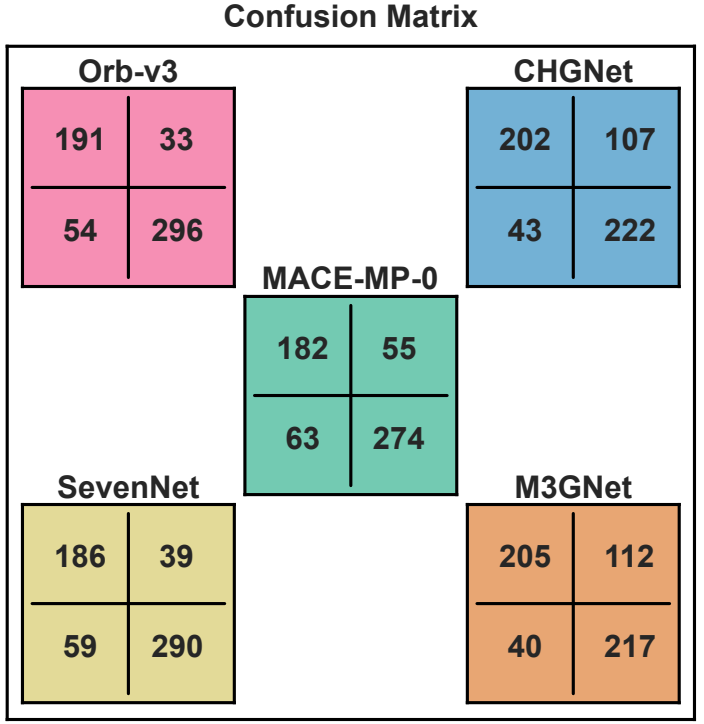
Orb-v3: 0.198 eV
MACE-MP-0: 0.202 eV
SevenNet: 0.203 eV
CHGNet: 0.248 eV
M3GNet: 0.257 eV

- DFT-NEB usually carry a **~60 meV error** in their predictions
- 60 meV change of E_m corresponds to an order of magnitude change in D at 298K
- We define **0.1 eV** as **'acceptable error'**

‘Sweet spot’ in E_m range? Classify good/bad conductors?

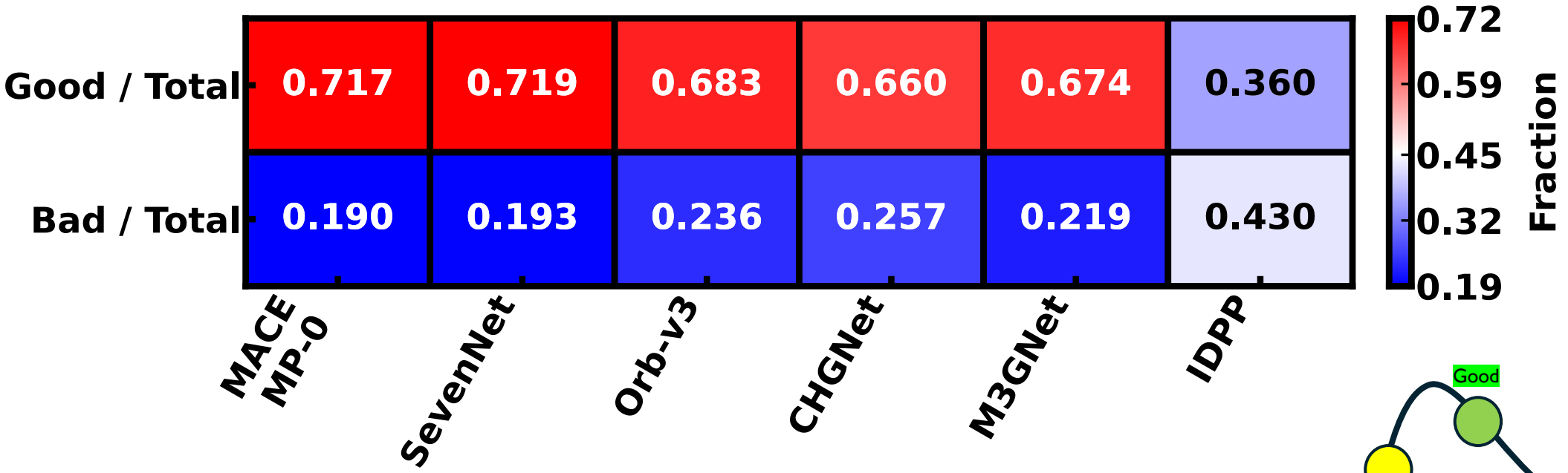


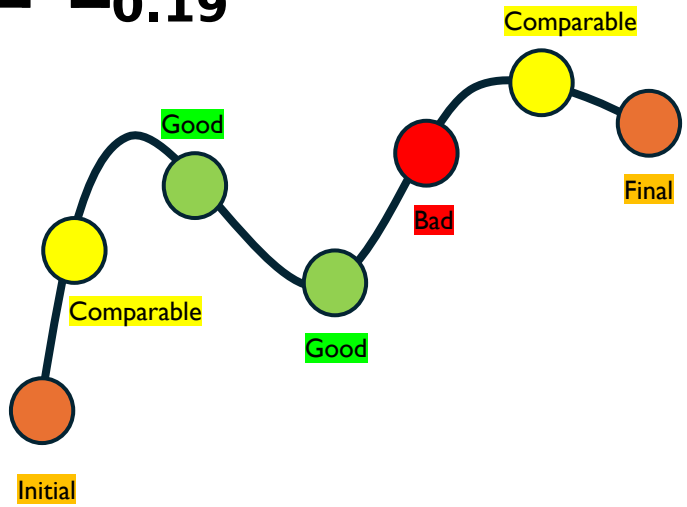
- All models struggle in predictions for systems with high E_m
- Orb-v3 exhibits the slowest degradation
- ‘Simpler’ models show peak performance in a narrow range of low E_m values



- Define < 500 meV as ‘good’ ionic conductor
- Prediction considered as TP (TN) if both the DFT-computed and model-predicted E_m are less than (greater than or equal to) 500 meV
- Orb-v3 reliably classifies 85% of the systems

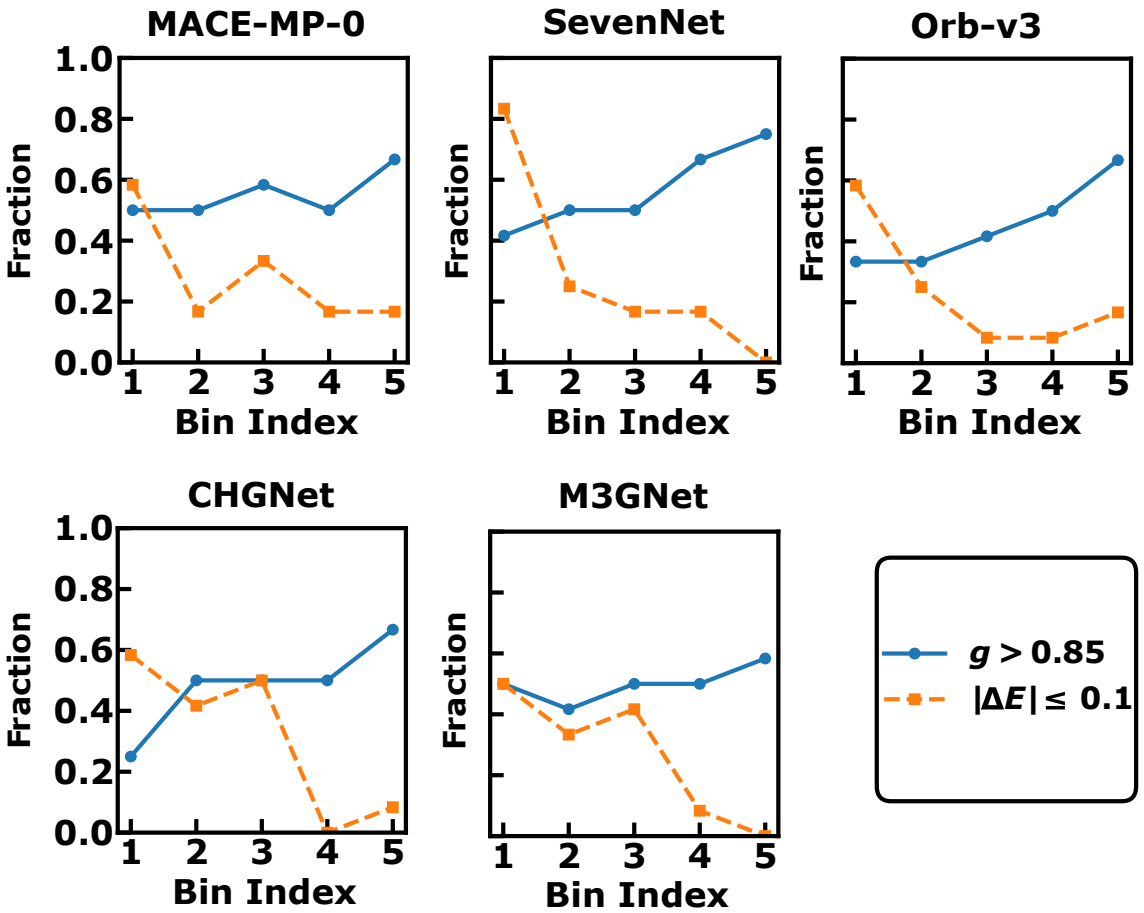
Geometry prediction peformance





- MACE-MP-0 and SevenNet generate better geometries most frequently
- Any MLIP MEP as an initial guess can lead to a good geometry in at least 66% cases
- IDPP and LI show similar performance

Geometry-barrier correlation



- Highest fraction of acceptable E_m predictions at low E_m range
- Highest fraction of 'good' path prediction at high E_m range
- No evident correlation
- The potential energy surfaces are likely 'flat' with variations in local geometries
- Good E_m prediction doesn't necessarily imply good geometry prediction, and vice-versa

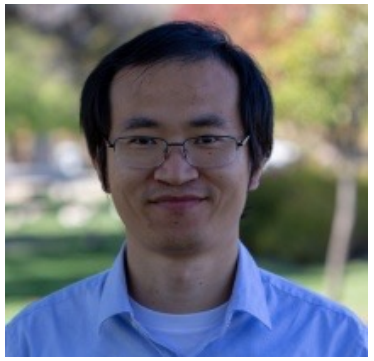
Conclusions

- Foundational MLIPs in their absolute prediction of E_m are not reliable
  - MLIPs perform better if the system can move ions easily
  - Orb-v3 can reliably classify materials, and hence could be used for high-throughput screening
- MLIP-NEB relaxed images as an initial guess can potentially accelerate the subsequent DFT-NEB
- No evident correlation between E_m and geometry prediction accuracy

ML with ~~instead of~~ DFT

Acknowledgments

Dr. Sai Gautam Gopalakrishnan



Dr. Penghao Xiao

SERC (IISc)

DALHOUSIE UNIVERSITY

Jureca- Julich

Digital Research Alliance of Canada | Alliance de recherche numérique du Canada

Compute Canada

Bheemaguli, Achinthya Krishna, Penghao Xiao, and Gopalakrishnan Sai Gautam. "Evaluation of Foundational Machine Learned Interatomic Potentials for Migration Barrier Predictions." arXiv preprint arXiv:2512.03642 (2025).

achinthyab@iisc.ac.in