Machine Learning for Accelerated Discovery of Superionic Solids

Artem Maevskiy[1], Alexandra Carvalho[1, 2], Emil Sataev[3], Volha Turchyna[3], Keian Noori[1, 2], Aleksandr Rodin[4, 5], A. H. Castro Neto[1, 2, 5, 6], Andrey Ustyuzhanin[1, 7]

[1]National University of Singapore, Institute for Functional Intelligent Materials, NUS S9 Building, 4 Science Drive 2, 117544 Singapore
[2]National University of Singapore, Centre for Advanced 2D Materials, 6 Science Drive 2, 117546 Singapore
[3]HSE University, Faculty of Computer Science, Pokrovsky Boulevard 11, 109028 Moscow, Russian Federation
[4]Yale-NUS College, 16 College Avenue West, 138527 Singapore
[5]National University of Singapore, Department of Materials Science Engineering, 9 Engineering Drive 1, 117575 Singapore
[6]National University of Singapore, Department of Physics, 2 Science Drive 3, 117551 Singapore
[7]Constructor University, Bremen, Campus Ring 1, 28759, Germany

maevskiy@nus.edu.sg

The discovery of novel solid electrolytes is crucial for the development of next-generation energy storage devices, as they offer enhanced charge capacity, stability, and safety [1]. However, the traditional approach of relying on density functional theory (DFT) calculations to predict the ionic conductivity is computationally intensive and has limited scalability. To overcome this bottleneck, data-driven approaches can be employed to accelerate the discovery process. One promising approach to accelerate solid electrolyte discovery is to replace DFT calculations with efficient machine-learned (ML) interatomic potentials. Recent advances have led to the development of universal potential models based on graph neural networks (GNNs) which can infer energies and forces across a broad range of chemical systems [2]. However, when applied to molecular dynamics (MD) or Nudge Elastic Band (NEB) simulations, commonly used to estimate ionic mobility and migration barriers, these models exhibit increased generalization errors as the studied systems are driven further away from the training domain. This can make the predictions less reliable and limit the overall screening efficiency of such methods.

In this work, we explore two alternative data-driven approaches for screening promising superionic candidates. The first approach addresses the computational bottleneck associated with calculating ion mobility from MD simulations, which requires numerous consecutive evaluations of the force field to estimate diffusion coefficients. We aim to mitigate this by training an ML model that can predict diffusion coefficients from a limited number of force field evaluations. To achieve this, we leverage a E3-equivariant graph neural network (GNN) with node features being simply the atomic accelerations under a given force field for a given configuration. For the basic proof-of-principle demonstration, we constructed a dataset of diffusion coefficients at $T = 1000\,K$ for 179 structures from MD simulations driven by a universal ML potential. To make a prediction, we evaluate the force field at a random atomic displacement from the equilibrium configuration and feed the corresponding coordinates and accelerations into the GNN. Importantly, we don't provide the atomic species as input to prevent the model from relying on chemical information. The final prediction is then calculated as average over several random evaluations. The learning curves for this model, shown in Fig. 1, along with predictions compared with the ground truth diffusion in Fig. 2, demonstrate that even this basic setup yields predictive power, with coefficient of determination value above 0.5.

Our second method, described in [3], leverages a universal potential in a setting that minimizes its generalization errors by confining the atomic systems as close as possible to the training domain via the *frozen framework* approximation. Within this approximation, the potential energy surface (PES) is inferred from the interatomic potential model by varying the position of a single mobile ion, while keeping the framework ions fixed at their equilibrium positions. The obtained PES is then used to calculate physically motivated descriptors designed to correlate with ionic mobility based on the labeled data [4, 5]. Finally, these descriptors are used to screen potential superionic materials. The overall logical pipeline for this approach is shown in Fig. 3. Ab initio MD validation results for our top predicted candidates are shown in Fig. 4. Our top prediction, LiB_3H_8, demonstrates an impressive ionic conductivity estimate of $1.6 \pm 0.4\,S/cm$ at $T = 500\,K$.

References

[1] Qing Zhao et al. "Designing solid-state electrolytes for safe, energy-dense batteries". In: *Nature Reviews Materials* 5.3 (Mar. 2020), pp. 229–252. ISSN: 2058-8437. DOI: 10.1038/s41578-019-0165-5.

[2] Janosh Riebesell et al. "Matbench Discovery – A framework to evaluate machine learning crystal stability predictions". 2023. DOI: 10.48550/arXiv.2308.14920. arXiv: 2308.14920.

[3] Artem Maevskiy et al. *Predicting ionic conductivity in solids from the machine-learned potential energy landscape*. 2024. arXiv: 2411.06804 [cond-mat.mtrl-sci]. URL: https://arxiv.org/abs/2411.06804.

[4] Leonid Kahle, Aris Marcolongo, and Nicola Marzari. "High-throughput computational screening for solid-state Li-ion conductors". In: *Energy Environ. Sci.* 13 (3 2020), pp. 928–948. DOI: 10.1039/C9EE02457C.

[5] Forrest AL Laskowski, Daniel B McHaffie, and Kimberly A See. "Identification of potential solid-state Li-ion conductors with semi-supervised learning". In: *Energy & Environmental Science* 16.3 (2023), pp. 1264–1276.

Figures

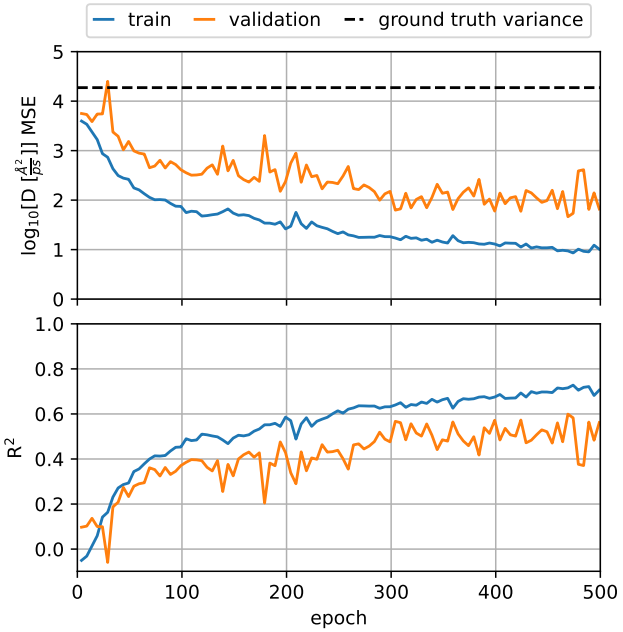

Figure 1: Learning curves for the baseline model for Li diffusivity prediction from the force field

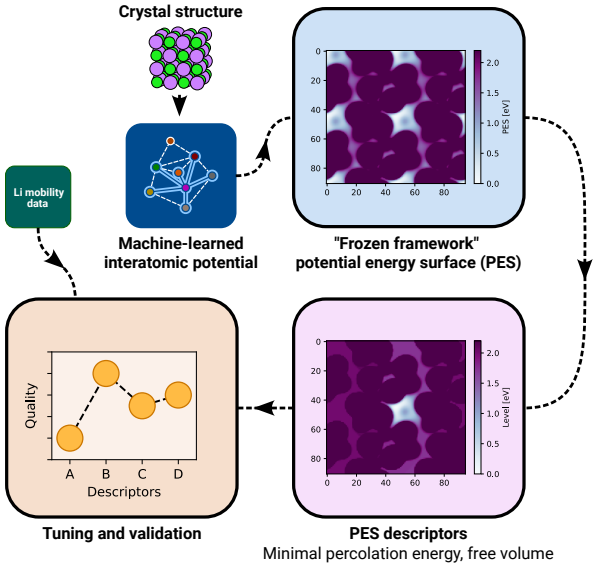

Figure 3: Analysis pipeline. See [3] for more details.

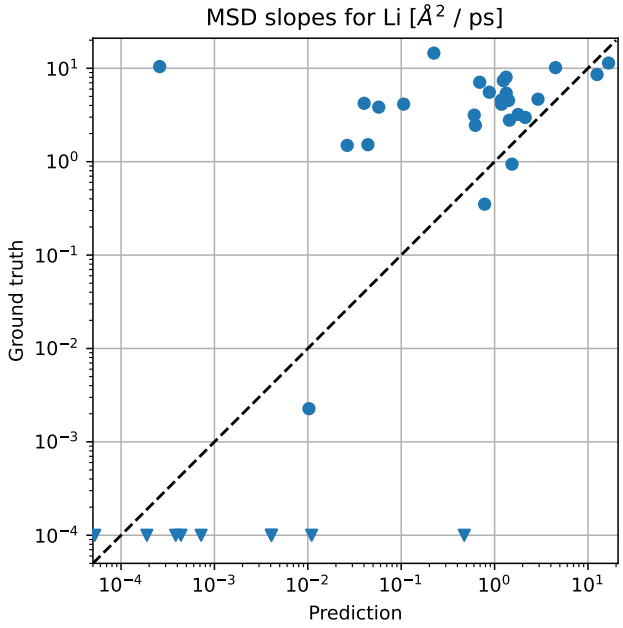

Figure 2: Baseline model predictions compared to the ground truth labels for predicting Li diffusivity from the force field (epoch 500, validation $R^2 = 0.53$). Ground truth MSD slopes are clipped from below at 10^{-4} Å^2/ps (shown with triangles). The dashed black line corresponds to $X = Y$.

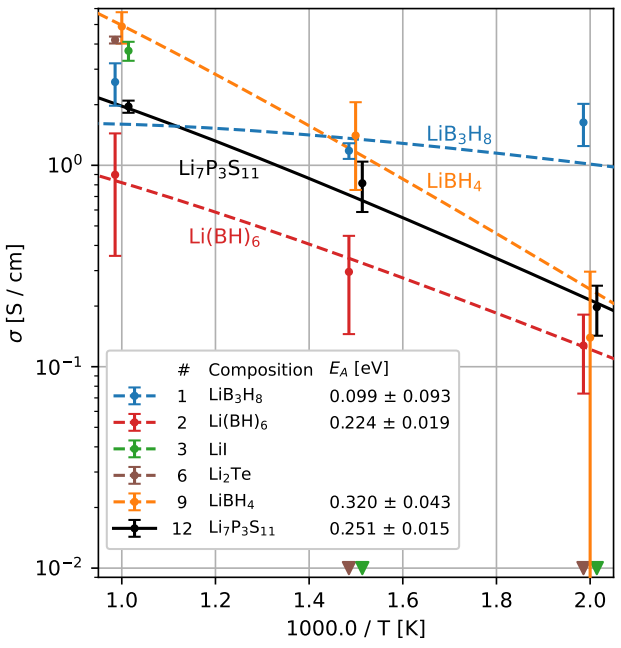

Figure 4: Results of AIMD validation for the top predicted candidates [3]