Accelerating the Discovery of Multi-Component Solid Electrolytes using High-Performance Machine Learning Interatomic Potentials

Yanhao Deng [1] **Yan Li** [1] **Zeyu Deng** [1]

[1]*Department of Materials Science and Engineering, National University of Singapore, Singapore, Singapore*. Correspondence to: Zeyu Deng msedz@nus.edu.sg.

1. Introduction

MLIPs have opened up new directions in computational materials chemistry. They offer a compelling combination of accuracy, flexibility, and computational efficiency [1, 2]. ML-driven simulations are enabling new insights into the structure and behavior of amorphous materials, bridging fundamental understanding and real device applications [3, 4]. As their reliability and generalizability continue improving, ML potentials rapidly evolve into mainstream simulation tools for materials modeling [5, 6]. A foundation model has been proposed using state-of-the-art MACE architecture[7, 8], which can achieve qualitative and at times quantitative accuracy on a huge dataset including diverse molecules and atoms.

This study investigates the application of MLIPs based on MACE models to multicomponent solid-state electrolytes (SSEs), a critical class of materials for next-generation batteries. We apply this method to study all-solid-state batteries, where liquid electrolytes are replaced by solid fast ion conductors, offering a promising pathway for safer commercial lithium- and sodium-based batteries. Ion transport in mixed polyanion solid electrolytes has been previously investigated using both computational and laboratory methods. We are using computational ways to study here.

We first benchmark MLIP performance against established SSE families, $Na_{1+x}Zr_2Si_xP_{3-x}O_{12}$ and $Li_{4x}Ge_xP_{1-x}O_{4-4x}S_{4x}$, confirming their general applicability while identifying key considerations for robust potential development in chemically diverse systems.

Then the predictive power of this approach is demonstrated by constructing a high-performance MLIP for the new halide system $Li_3In_xY_{1-x}Br_{6y}Cl_{6-6y}$. This potential facilitated an efficient exploration of its two-dimensional compositional space, leading to the identification of the $Li_3In_xY_{1-x}Br_{6y}Cl_{6-6y}$ stoichiometry with the most favorable predicted. Our results underscore the capability of MLIPs to accelerate the discovery cycle of complex functional materials and provide a robust computational framework for designing advanced solid electrolytes.

2. Methodology

2.1 Dataset generation

The distribution of training dataset is a critical factor governing the performance of MLIP models. We designed a workflow depicted in Fig. 1 to build a comprehensive training dataset for a multi-component system.

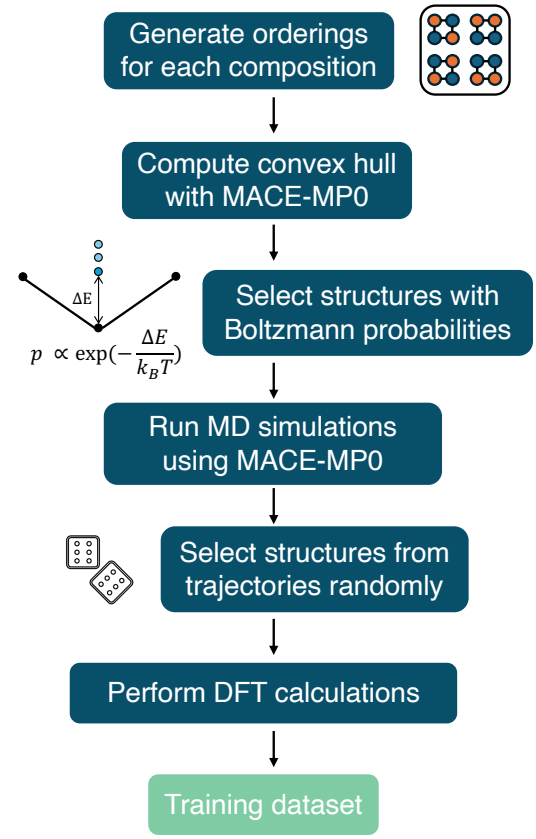

Fig. 1: Workflow to build training dataset.

First, we generate different orderings for each composition, and their energies are estimated using the foundation model (MACE-MP0) to construct a convex hull. These structures are then selected based on their energy above the hull ΔE, with a probability proportaional to the Boltzmann distribution, $p \propto \exp -\frac{\Delta E}{k_B T}$. T is set to 800K in this work. This probability balances low-energy structures and thermodynamically accessible higher-energy structures realistically. The structure with minimum energy is always included to ensure baseline coverage. Subsequently, for all these chosen structures, molecular dynamics simulations are performed using the same foundation model (MACE-MP0),in the NVT ensemble at 800K for 15ps. This step aims to explore the possible configuration space effectively. Then we randomly select about 2500 structures from these MD trajectories after 1 ps, and perform DFT calculations to build an accurate training dataset efficiently.

2.2 Model training

Fine-tuning has the advantage of requiring only a small dataset to achieve stable results, but it heavily relies on the quality and suitability of the underlying foundation model. In some cases, significant additional data may be needed to correct model biases or limitations. Fine-tuning on a multi-head model can prevent catastrophic forgetting that occurs sometimes during naive fine-tuning.

In this work, we perform a multi-head fin-tuning on the materials project model MACE-MP0 with medium size, rather than naive fine-tuning, to avoid catastrophic forgetting. The model is trained for a maximum of 500 epochs, with early stopping applied if the validation loss does not improve for 20 consecutive epochs. The learning rate is 0.01.

References

[1] Choongseok Chang, Volker L. Deringer, Kalpana S. Katti, Veronique Van Speybroeck, and Christopher M. Wolverton. Simulations in the era of exascale computing. *Nature Reviews. Materials*, 8(5):309–313, 2023.

[2] Volker L. Deringer, Miguel A. Caro, and Gábor Csányi. Machine Learning Interatomic Potentials as Emerging Tools for Materials Science. *Advanced Materials*, 31(46):1902765, 2019. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/adma.201902765.

[3] Linus C. Erhard, Jochen Rohrer, Karsten Albe, and Volker L. Deringer. Modelling atomic and nanoscale structure in the silicon–oxygen system through active machine learning. *Nature Communications*, 15(1):1927, March 2024. Publisher: Nature Publishing Group.

[4] Yuxing Zhou, Wei Zhang, En Ma, and Volker L. Deringer. Device-scale atomistic modelling of phase-change memory materials. *Nature Electronics*, 6(10):746–754, October 2023. Publisher: Nature Publishing Group.

[5] Zakariya El-Machachi, Damyan Frantzov, A. Nijamudheen, Tigany Zarrouk, Miguel A. Caro, and Volker L. Deringer. Accelerated First-Principles Exploration of Structure and Reactivity in Graphene Oxide. *Angewandte Chemie International Edition*, 63(52):e202410088, 2024. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/anie.202410088.

[6] Chiheb Ben Mahmoud, John L. A. Gardner, and Volker L. Deringer. Data as the next challenge in atomistic machine learning. *Nature Computational Science*, 4(6):384–387, June 2024. Publisher: Nature Publishing Group.

[7] Ilyes Batatia, Philipp Benner, Yuan Chiang, Alin M. Elena, Dávid P. Kovács, Janosh Riebesell, Xavier R. Advincula, Mark Asta, Matthew Avaylon, William J. Baldwin, Fabian Berger, Noam Bernstein, Arghya Bhowmik, Samuel M. Blau, Vlad Cărare, James P. Darby, Sandip De, Flaviano Della Pia, Volker L. Deringer, Rokas Elijošius, Zakariya El-Machachi, Fabio Falcioni, Edvin Fako, Andrea C. Ferrari, Annalena Genreith-Schriever, Janine George, Rhys E. A. Goodall, Clare P. Grey, Petr Grigorev, Shuang Han, Will Handley, Hendrik H. Heenen, Kersti Hermansson, Christian Holm, Jad Jaafar, Stephan Hofmann, Konstantin S. Jakob, Hyunwook Jung, Venkat Kapil, Aaron D. Kaplan, Nima Karimitari, James R. Kermode, Namu Kroupa, Jolla Kullgren, Matthew C. Kuner, Domantas Kuryla, Guoda Liepuoniute, Johannes T. Margraf, Ioan-Bogdan Magdău, Angelos Michaelides, J. Harry Moore, Aakash A. Naik, Samuel P. Niblett, Sam Walton Norwood, Niamh O'Neill, Christoph Ortner, Kristin A. Persson, Karsten Reuter, Andrew S. Rosen, Lars L. Schaaf, Christoph Schran, Benjamin X. Shi, Eric Sivonxay, Tamás K. Stenczel, Viktor Svahn, Christopher Sutton, Thomas D. Swinburne, Jules Tilly, Cas van der Oord, Eszter Varga-Umbrich, Tejs Vegge, Martin Vondrák, Yangshuai Wang, William C. Witt, Fabian Zills, and Gábor Csányi. A foundation model for atomistic materials chemistry, March 2024. arXiv:2401.00096 [cond-mat, physics:physics].

[8] Ilyes Batatia, Dávid Péter Kovács, Gregor N. C. Simm, Christoph Ortner, and Gábor Csányi. MACE: Higher Order Equivariant Message Passing Neural Networks for Fast and Accurate Force Fields, January 2023. arXiv:2206.07697 [cond-mat, physics:physics, stat].