A COMPARATIVE STUDY OF MOLECULAR DYNAMICS APPROACHES FOR SIMULATING IONIC CONDUCTIVITY IN SOLID LITHIUM ELECTROLYTES

Dounia Shaaban Kabakibo
Département de physique, Université de Montréal
Mila – Quebec AI Institute
Montréal, Canada
dounia.shaaban.kabakibo@umontreal.ca

Félix Therrien
Mila – Quebec AI Institute
Montréal, Canada

Yoshua Bengio
Département d'informatique et de recherche opérationnelle, Université de Montréal
Mila – Quebec AI Institute
Montréal, Canada

Michel Côté
Département de physique, Université de Montréal
Mila – Quebec AI Institute
Montréal, Canada

Hongyu Guo & Homin Shin
National Research Council Canada
Ottawa, Canada

Alex Hernandez-Garcia
Département d'informatique et de recherche opérationnelle, Université de Montréal
Mila – Quebec AI Institute
Institut Courtois, Université de Montréal
Montréal, Canada

ABSTRACT

Accurate prediction of ionic conductivity is critical for the design of high-performance solid-state electrolytes in next-generation batteries. We benchmark molecular dynamics (MD) approaches for computing ionic conductivity in 21 lithium solid electrolytes for which experimental ionic conductivity has been previously reported in the literature. In particular, we compare simulations driven by density functional theory (DFT) and by universal machine-learning interatomic potentials (uMLIPs), namely a MACE foundation model. We find comparable performance between DFT and MACE, despite MACE on one GPU more than 350 times faster than DFT on a 64-CPU node. The framework developed here is designed to enable systematic comparisons with additional uMLIPs and fine-tuned models in future work.

1 INTRODUCTION

Solid-state electrolytes are a key component of next-generation energy storage technologies, offering the promise of improved safety and higher energy density compared to liquid electrolytes (Kim et al., 2015). A central property governing their performance is ionic conductivity, which quantifies how efficiently mobile ions migrate through a solid under an applied electric field (Mehrer, 2007). While experimental measurements remain the reference standard, computational approaches play an important role in screening candidate materials and providing physics-based insights into ion transport. Molecular dynamics (MD) simulations offer a direct route to computing ionic conductivity from atomic trajectories, but the reliability of the results depends critically on the quality of the force model and the accessible system sizes and simulation times.

Ab initio molecular dynamics based on density functional theory (DFT) provide a principled description of interatomic forces, yet its high computational cost severely limits the length and scale of simulations. Recent advances in universal machine-learning interatomic potentials (uMLIPs) have opened new possibilities for large-scale MD simulations. Models such as the MACE foundation model (Batatia et al., 2023) are trained on diverse DFT data and can, in principle, be applied across a wide range of chemistries, on systems involving thousands of atoms at the scale of nanoseconds and in a fraction of the time needed for DFT simulations. However, it remains unclear how accurately MD simulations driven by uMLIPs reproduce emergent transport properties such as ionic conductivity. This motivates a systematic benchmarking against ab initio molecular dynamics.

1.1 CONTRIBUTIONS

Our approach introduces **three key contributions**:

1. **A consistent dataset:** We consider a relatively large set of materials for which ionic conductivities have been experimentally measured, enabling a direct and systematic comparison between simulations and experiments.
2. **A unified procedure:** We propose a unified procedure for ionic conductivity estimation that standardizes parameter choices and uncertainty quantification, providing a statistically meaningful basis for data-driven comparison across methods.
3. **A comparison of multiple simulation approaches:** With this dataset and our proposed procedure, we assess the performance of two simulation methods—ab initio molecular dynamics as well as a general uMLIP—and how they compare to experimental measurements using an identical simulation protocol.

Previous work with connections to ours is reviewed in Appendix A.1.

2 METHODS

2.1 IONIC CONDUCTIVITY FROM MOLECULAR DYNAMICS SIMULATIONS

Room temperature ionic conductivity σ is a key property that characterizes ion transport in solid state materials. One approach to estimate σ is to use MD simulations, which track the time evolution of atoms in a material based on Newton's equations of motion. These simulations allow the study of ionic diffusion at finite temperatures by modeling atomic interactions through different levels of theory: DFT, empirical potentials, or MLIPs. After performing an MD simulation, the self diffusion coefficient D for mobile ions (in this case, lithium ions) is obtained using the Einstein relation

$$D = \lim_{t\to\infty} \frac{\text{MSD}(t)}{2d\,t}, \tag{1}$$

where $d = 3$ is the system's dimensionality and $\text{MSD}(t)$ is the per ion mean square displacement at time t. In practice, D is calculated from the slope of the $\text{MSD}(t)$ plot over a finite time window. Once D is determined, the ionic conductivity σ is obtained via the Nernst-Einstein relation

$$\sigma(T) = \frac{Nq^2 D(T)}{V k_B T}, \tag{2}$$

where N is the total number of lithium ions, q is the charge of the diffusing ion, V is the cell volume, k_B is the Boltzmann constant and T is the temperature. Because room-temperature ionic conductivity requires long MD timescales, which are computationally prohibitive, especially for ab initio MD (AIMD), simulations are instead performed at higher temperatures. The diffusion coefficients estimated at these temperatures are used to fit the Arrhenius equation:

$$D(T) = D_0 \exp\left(-\frac{E_a}{k_B T}\right), \tag{3}$$

where E_a is the activation energy. This relation is then extrapolated to estimate the conductivity at room temperature. Despite the widespread use of the Arrhenius fitting, strict Arrhenius behavior is not universal and deviations have been observed (Qi et al., 2021).

2.2 MATERIALS AND PARAMETERS CHOICE

A set of 21 solid lithium electrolytes covering a wide range of ionic conductivities were selected from the OBELiX database (Therrien et al., 2025) which is a collection of synthesized solid electrolytes and their experimentally measured room-temperature conductivities compiled from the literature. Further details are provided in Appendix A.3.

We considered two classes of force calculators: fully *ab initio* (DFT) and uMLIPs. In the present work, we focus on a single uMLIP, namely the MACE model (Batatia et al., 2023), using the `medium-mpa-0` checkpoint.

In our procedure, MD simulations were performed on supercells with minimum dimensions of $10\text{Å} \times 10\text{Å} \times 10\text{Å}$ [1]. From these simulations, for each material, we obtained the diffusion coefficient D at 5 different temperatures that we then used to fit equation 3 and obtain the extrapolated room temperature diffusion coefficient $D(T = 300\text{K})$. Simulations were carried out at temperatures ranging from 800 to 1200 K in 100 K increments in the NVT ensemble using a Nosé–Hoover thermostat and a time step of 2 fs. Each MD trajectory was 100 ps long, with the first 5 ps discarded prior to diffusion analysis. MD simulations using MACE models were carried out with the Atomic Simulation Environment (ASE) (Larsen et al., 2017). Parameters used for AIMD simulations are detailed in Appendix A.4.

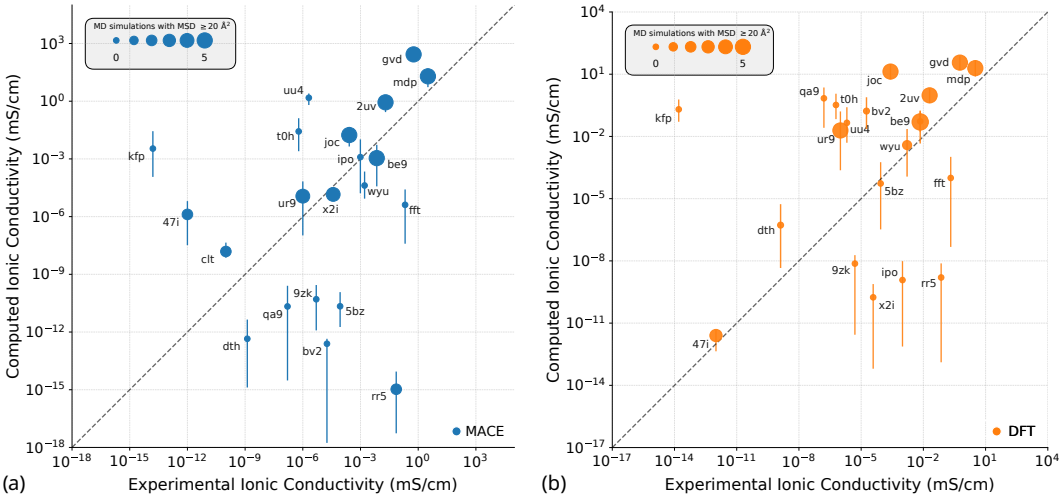

Figure 1: Parity plot comparing ionic conductivity computed using (a) MACE and (b) DFT foundation model with experimental values reported by Therrien et al. (2025). The size of the circles reflect the number of simulations (at 5 different temperatures) that reached at least $\text{MSD} = 20\ \text{Å}^2$.

Uncertainty estimation for diffusion coefficients was carried out using the `Kinisi-2.0.1` package (McCluskey et al., 2024a;b). In this framework, ionic diffusion is statistically analyzed by modeling atomic displacements over time and propagating their uncertainties to obtain confidence intervals on the diffusion coefficient. The parameters used in the documentation examples were used throughout and no ablation study was conducted to assess sensitivity to these settings.

For each material, as a measure of reliability, we also report how many of the five simulations (one for each temperature) had at least some ionic movement ($\text{MSD} = 20\ \text{Å}^2$). This threshold corresponds to a characteristic displacement of ~ 4.5 Å, a length scale comparable to typical interatomic distances in inorganic crystals, indicating motion beyond local vibrational fluctuations.

[1] Exceptions were made for AIMD simulations. See Appendix A.4.

Table 1: Aggregated computational cost over all materials per temperature. Times are reported in `dd:hh:mm`. Only maximum available GPU memory usage is reported.

Method	**# of devices**		**Total time**			**Max. mem. (GB)**	
	GPU	CPU	Median	Min	Max	GPU	CPU
MACE	1	1	47m	20m	2h 27m	40	1.2
DFT	0	64	9d 21h 09m	2d 3h 0m	42d 09h27m	0	32

3 RESULTS

In this section, we present results for the analysis discussed above. Figure 1 shows the parity plot of the estimated ionic conductivity using both DFT[2] and the MACE foundation model versus the experimental values included in Therrien et al. (2025). The materials are labeled with their corresponding names in OBELiX. The chemical formulas are given in Appendix A.3.

The figures also show how many of the MD simulations at 5 temperatures had MSD above threshold for each material. We observe that most materials passing all checks have estimated conductivities close to the experimental values. Conversely, the estimations from simulations that fail to meet the displacement threshold tend to diverge from the experimental values. Interestingly, some materials which are known to exhibit high diffusion experimentally, namely `fft` (β-Li_3N) and `rr5` ($LiTi_2(PO_4)_3$), show little diffusion in our simulations. We anticipate that a diffusion mechanism mediated by a small concentration of defects—either not significant enough to be reported in lab experiments or not directly measurable—is necessary for diffusion in these cases. Future work will extend this analysis. Additionally, it should be noted that the experimental value $\sigma = 10^{-12}$ mS/cm we use for `47i` ($Li_4P_2O_7$) may be lower than the actual value, as the literature only reports an upper bound ($\sigma_{\text{exp}} < 10^{-7}$ mS/cm). Figures providing additional insight on the previous results can be found in Appendix A.5.

In Table 1, we report details about the computational resources used by both methods. For the DFT simulations, the number of CPUs used for each material was optimized based on system size (see Appendix A.4). For the MACE simulations, a single GPU was used for all materials using a `1x NVIDIA A100-SXM4-40GB` paired with one CPU core. Each reported value corresponds to the computational cost of a single MD simulation (one material at one temperature). Since MD simulations at different temperatures are independent, they can be run concurrently provided sufficient computational resources are available. The reported per-simulation runtimes therefore approximate the expected wall-clock time of one MD simulation under parallel execution and optimized resources.

4 DISCUSSION AND CONCLUSION

While the number of materials considered in this study remain limited to draw robust statistical conclusions, several trends emerge. First, molecular dynamics simulations using MACE require only a fraction of the computational time needed for DFT-based simulations, provided that GPU resources are available. Specifically, MACE on one GPU is on average 378 times faster than DFT on one 64-CPU node. Although DFT simulations can in principle also be accelerated on GPUs, the expected speedup is not as substantial as for uMLIPs. Moreover, many researchers in the materials science community still rely primarily on CPU-based DFT calculations, which further motivates our DFT simulations being CPU–based.

Despite this large difference in computational cost, MACE and DFT exhibit comparable predictive performance. Quantitatively, we evaluate performance by computing the mean absolute error (MAE) between the logarithm of the predicted and the experimentally measured ionic conductivities. Over the 21 materials studied with MACE, we obtain an MAE of 4.11. For the 20 materials for which DFT results are available, the MAE is 4.10. Restricting the analysis to the materials passing all five MSD reliability checks — five for MACE and six for DFT — the MAE decreases to 1.55

[2]DFT simulations for 1 material could not be completed within a reasonable amount of time.

for MACE and 2.35 for DFT. It is important to emphasize that molecular dynamics results become severely unreliable when little to no ionic motion is observed, as the extracted transport properties in such cases mainly reflect thermal fluctuations rather than genuine diffusion. Although the statistics over the five most reliable materials remain limited, these results suggest that MACE achieves performance comparable to DFT.

To put these results in perspective, Therrien et al. (2025) found that the mean absolute variation over experimental ionic conductivity measurements of the same material was about 0.4 orders of magnitude and that a random forest model trained on the OBELiX dataset could achieve an MAE of about 1.6 on their test set.

These findings do not diminish the value of molecular dynamics simulations, which remain essential for understanding diffusion mechanisms and the underlying physics of ionic transport. However, they suggest that, in a high-throughput screening context, MD simulations should be performed selectively: primarily using uMLIPs such as MACE, and only occasionally with DFT when higher-fidelity calculations are strictly necessary. Beyond screening, these results are also particularly relevant in the context of machine-learning–driven materials discovery. In practice, the amount of high-quality experimental data on solid-state electrolytes remains limited, and models trained to predict ionic conductivity directly might exhibit restricted transferability to unseen chemistries. In this setting, multi-fidelity active learning frameworks (Hernandez-Garcia et al., 2023) that combine large numbers of low-cost uMLIP-based MD simulations with a smaller number of higher-fidelity DFT calculations, offer a promising path forward. Future work directions are discussed in Appendix A.2.

ACKNOWLEDGMENTS

The authors thank Gabor Csanyi and Mickael Dolle for valuable discussions. The authors acknowledge support from the National Research Council Canada (NRC) through a collaborative R&D grant (AI4D-core-132), Calcul Québec and the Digital Research Alliance of Canada. D.S.K. acknowledges support by the NSERC Postgraduate Scholarships – Doctoral program and the FRQ Nature and technologies Sector – Doctoral Research Scholarship Program. M.C. is a member of the Regroupement québécois sur les matériaux de pointe (RQMP).

REFERENCES

Luis Barroso-Luque, Muhammed Shuaibi, Xiang Fu, Brandon M Wood, Misko Dzamba, Meng Gao, Ammar Rizvi, C Lawrence Zitnick, and Zachary W Ulissi. Open materials 2024 (omat24) inorganic materials dataset and models. *arXiv preprint arXiv:2410.12771*, 2024.

Ilyes Batatia, Philipp Benner, Yuan Chiang, Alin M. Elena, Dávid P. Kovács, Janosh Riebesell, Xavier R. Advincula, Mark Asta, William J. Baldwin, 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, Edvin Fako, Andrea C. Ferrari, Annalena Genreith-Schriever, Janine George, Rhys E. A. Goodall, Clare P. Grey, 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, 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, Eric Sivonxay, Tamás K. Stenczel, Viktor Svahn, Christopher Sutton, 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. 2023.

Chi Chen, Dan Thien Nguyen, Shannon J Lee, Nathan A Baker, Ajay S Karakoti, Linda Lauw, Craig Owen, Karl T Mueller, Brian A Bilodeau, Vijayakumar Murugesan, et al. Accelerating computational materials discovery with machine learning and cloud high-performance computing: from large-scale screening to experimental validation. *Journal of the American Chemical Society*, 146 (29):20009–20018, 2024.

Artem D Dembitskiy, Innokentiy S Humonen, Roman A Eremin, Dmitry A Aksyonov, Stanislav S Fedotov, and Semen A Budennyy. Benchmarking machine learning models for predicting lithium ion migration. *npj Computational Materials*, 11(1):131, 2025.

Bowen Deng, Peichen Zhong, KyuJung Jun, Janosh Riebesell, Kevin Han, Christopher J. Bartel, and Gerbrand Ceder. Chgnet as a pretrained universal neural network potential for charge-informed atomistic modelling. *Nature Machine Intelligence*, pp. 1–11, 2023. doi: 10.1038/s42256-023-00716-3.

Natascia L Fragapane and Volker L Deringer. Li-ps electrolyte materials as a benchmark for machine-learned interatomic potentials. *arXiv preprint arXiv:2511.16569*, 2025.

Alex Hernandez-Garcia, Nikita Saxena, Moksh Jain, Cheng-Hao Liu, and Yoshua Bengio. Multi-fidelity active learning with GFlownets. *arXiv preprint arXiv:2306.11715*, 2023.

Anubhav Jain, Geoffroy Hautier, Charles J Moore, Shyue Ping Ong, Christopher C Fischer, Tim Mueller, Kristin A Persson, and Gerbrand Ceder. A high-throughput infrastructure for density functional theory calculations. *Computational Materials Science*, 50(8):2295–2310, 2011.

Leonid Kahle, Aris Marcolongo, and Nicola Marzari. High-throughput computational screening for solid-state li-ion conductors. *Energy & Environmental Science*, 13(3):928–948, 2020.

Joo Gon Kim, Byungrak Son, Santanu Mukherjee, Nicholas Schuppert, Alex Bates, Osung Kwon, Moon Jong Choi, Hyun Yeol Chung, and Sam Park. A review of lithium and non-lithium based solid state batteries. *Journal of Power Sources*, 282:299–322, 2015.

Georg Kresse and Jürgen Furthmüller. Computational mater. *Sci*, 6(1):15, 1996a.

Georg Kresse and Jürgen Furthmüller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Physical review B*, 54(16):11169, 1996b.

Georg Kresse and Jürgen Hafner. Ab initio molecular dynamics for liquid metals. *Physical review B*, 47(1):558, 1993.

Georg Kresse and Daniel Joubert. From ultrasoft pseudopotentials to the projector augmented-wave method. *Physical review b*, 59(3):1758, 1999.

Ask Hjorth Larsen, Jens Jørgen Mortensen, Jakob Blomqvist, Ivano E Castelli, Rune Christensen, Marcin Dułak, Jesper Friis, Michael N Groves, Bjørk Hammer, Cory Hargus, et al. The atomic simulation environment—a python library for working with atoms. *Journal of Physics: Condensed Matter*, 29(27):273002, 2017.

Artem Maevskiy, Alexandra Carvalho, Emil Sataev, Volha Turchyna, Keian Noori, Aleksandr Rodin, AH Neto, and Andrey Ustyuzhanin. Predicting ionic conductivity in solids from the machine-learned potential energy landscape. *arXiv preprint arXiv:2411.06804*, 2024.

Andrew R McCluskey, Samuel W Coles, and Benjamin J Morgan. Accurate estimation of diffusion coefficients and their uncertainties from computer simulation. *Journal of Chemical Theory and Computation*, 21(1):79–87, 2024a.

Andrew R McCluskey, Alexander G Squires, Josh Dunn, Samuel W Coles, and Benjamin J Morgan. kinisi: Bayesian analysis of mass transport from molecular dynamics simulations. *Journal of Open Source Software*, 9(94):5984, 2024b.

Helmut Mehrer. *Diffusion in solids: fundamentals, methods, materials, diffusion-controlled processes*. Springer, 2007.

Juno Nam, Sulin Liu, Gavin Winter, KyuJung Jun, Soojung Yang, and Rafael Gómez-Bombarelli. Flow matching for accelerated simulation of atomic transport in crystalline materials. *Nature Machine Intelligence*, pp. 1–11, 2025.

Shyue Ping Ong, William Davidson Richards, Anubhav Jain, Geoffroy Hautier, Michael Kocher, Shreyas Cholia, Dan Gunter, Vincent L Chevrier, Kristin A Persson, and Gerbrand Ceder. Python materials genomics (pymatgen): A robust, open-source python library for materials analysis. *Computational Materials Science*, 68:314–319, 2013.

Ji Qi, Swastika Banerjee, Yunxing Zuo, Chi Chen, Zhuoying Zhu, ML Holekevi Chandrappa, Xiangguo Li, and Shyue Ping Ong. Bridging the gap between simulated and experimental ionic conductivities in lithium superionic conductors. *Materials Today Physics*, 21:100463, 2021.

Austin D Sendek, Ekin D Cubuk, Evan R Antoniuk, Gowoon Cheon, Yi Cui, and Evan J Reed. Machine learning-assisted discovery of solid li-ion conducting materials. *Chemistry of Materials*, 31(2):342–352, 2018.

Félix Therrien, Jamal Abou Haibeh, Divya Sharma, Rhiannon Hendley, Leah Wairimu Mungai, Sun Sun, Alain Tchagang, Jiang Su, Samuel Huberman, Yoshua Bengio, et al. Obelix: A curated dataset of crystal structures and experimentally measured ionic conductivities for lithium solid-state electrolytes. *arXiv preprint arXiv:2502.14234*, 2025.

A Appendix

A.1 Related work

High-throughput and machine-learning-based approaches are now central to the efforts for the discovery of solid ionic conductors. Early screening frameworks relied on workflows based on physics and molecular dynamics to estimate lithium diffusion across large materials databases (Kahle et al., 2020). Machine learning was subsequently introduced as a data-driven pre-screening tool to prioritize promising Li-ion conductors and reduce the cost of first-principles simulations (Sendek et al., 2018), and later combined with cloud high-performance computing to enable screening of tens of millions of candidates with experimental validation of a few selected solid electrolytes (Chen et al., 2024). More recently, machine-learned interatomic potentials have been used to model ion diffusion directly through MD in representative electrolytes (Fragapane & Deringer, 2025), while other studies have leveraged uMLIPs or ML models to infer transport trends (Maevskiy et al., 2024) or predict migration barriers (Dembitskiy et al., 2025), without running MD. Generative modeling has also been used to accelerate MD simulations in Lithium solid electrolytes (Nam et al., 2025).

A.2 FUTURE WORK

While the present results focus on simulations using a MACE uMLIP, the full framework—intended to be explored in future work—is designed to enable consistent comparisons across additional approaches, including other uMLIPs as well as material-specific models fine-tuned on DFT data. Such study would provide a systematic way to quantify the performance of uMLIPs and their fine-tuning strategies for long-time-scale diffusion simulations involving large numbers of atoms, and how they compare to fully ab initio simulations. The present study is limited to fully occupied crystal structures. In many solid electrolytes, however, ionic diffusion is known to be mediated by defects. Extending this work to partially occupied structures, such as those available in the OBELiX database, would therefore be of significant interest. In addition, one could consider relaxing the experimental structures prior to running MD simulations to assess the sensitivity of the computed conductivities to structural relaxation. Additional factors not examined here include the impact of longer simulation times and the number of temperature points used in the Arrhenius fits. Finally, introducing criteria to detect possible structural melting would be an important consideration, especially for simulations conducted at high temperatures.

A.3 MATERIALS SELECTION

As mentioned before, the materials were selected from the OBELiX database (Therrien et al., 2025). As a first filter, only materials with fully occupied sites were considered to avoid complexities related to partial occupancies. For this subset, the range of ionic conductivities was divided on a logarithmic scale into 20 equal bins, from $\sigma = 10^{-7}$ mS/cm^2 to the highest measured value in the dataset, 25 mS/cm^2. Two bins were empty, and one material was randomly selected from each of the remaining 18 bins. To include poorly conducting materials, the lower conductivity range, $\sigma = 10^{-15}$–10^{-7} mS/cm^2, was further divided into four equal bins, from which one material per bin was selected, resulting in a total of 22 materials. MD simulations using the MACE uMLIP failed for $LiClC_3H_7NO$, and therefore this material was excluded from further analysis. This procedure resulted in a total of 21 materials. Table 2 gives the chemical formulas of the selected materials. Cross-referencing our material selection against the MPtraj (Deng et al., 2023) and sAlexandria (Barroso-Luque et al., 2024) datasets on which the Mace `medium-mpa-0` model was trained returned matches for 11 Obelix IDs within the 5% lattice parameter and 5° angle difference thresholds: `qa9`, `bv2`, `be9`, `5bz`, `gvd`, `fft`, `9zk`, `47i`, `clt` , and `joc`.

A.4 AIMD PARAMETERS

AIMD simulations were performed using the Vienna *Ab initio* Simulation Package (VASP) (Kresse & Furthmüller, 1996a;b; Kresse & Hafner, 1993; Kresse & Joubert, 1999), employing a minimal Γ-centered $1 \times 1 \times 1$ k-point grid to reduce computational cost and an energy cutoff of 520 eV. All calculations were non-spin-polarized and used the Perdew–Burke–Ernzerhof (PBE) exchange–correlation functional. Projector Augmented Wave (PAW) pseudopotentials were selected following the Materials Project recommended mapping, as implemented in the `MPRelaxSet` VASP input set from `pymatgen` (Ong et al., 2013; Jain et al., 2011). These choices were made to maximize consistency with the DFT data on which the Mace `medium-mpa-0` model was trained, which was computed at the PBE+U level with PAW pseudopotentials. An exception that we note is that no Hubbard U correction was applied here, which in practice only affects $LiVO_3$ among the materials studied. Another notable difference is that portions of both training datasets include spin-polarized calculations, whereas all AIMD runs reported here are non-spin-polarized. Finally, the training data was generated using dense k-point meshes, while the present calculations use only the Γ point, which is well-justified for the large supercells employed but may introduce small systematic offsets in absolute energies and forces relative to the training references. Regarding the supercell size, exceptions to minimum dimensions of $10\text{Å} \times 10\text{Å} \times 10\text{Å}$ were made for AIMD simulations of materials Li_5GaO_4, $LiTi_2(PO_4)_3$, and $LiZr_2(PO_4)_3$, for which only primitive cells were used to keep the number of atoms computationally feasible, and for $Li_6CuB_4O_{10}$, where one lattice dimension was not expanded beyond 10 Å. In all cases, the remaining dimensions were at least 8 Å. With 64 CPUs available per node, the number of nodes was determined as $n_{\text{nodes}} = \max\left(1, n_{\text{atoms}}//64\right)$, which in practice always resulted in a single node.

Table 2: Selected materials and their chemical formulas.

OBELiX ID	Reduced chemical formula
qa9	Li_3BN_2
t0h	Li_2ZrO_3
ur9	$LiZr_2(PO_4)_3$
uu4	$LiVO_3$
9zk	Li_5GaO_4
bv2	$LiYO_2$
x2i	$LiBiO_2$
5bz	Li_3BN_2
joc	Li_3PS_4
ipo	$Li_7La_3Hf_2O_{12}$
wyu	$Li_7La_3Zr_2O_{12}$
be9	$LiGaBr_4$
2uv	Li_4SnSe_4
rr5	$LiTi_2(PO_4)_3$
goi	$LiClC_3H_7NO$
fft	β-Li_3N
gvd	α-Li_3N
mdp	$Li_7P_3S_{11}$
dth	Li_4GeO_4
clt	$Li_6CuB_4O_{10}$
47i	$Li_4P_2O_7$
kfp	$Li_2BaP_2O_7$

A.5 Additional analysis

In Figure 2, we present a parity plot comparing DFT and MACE predictions. In principle, DFT can be regarded as a reference for MACE given that the latter is trained on DFT calculations. However, it should be noted that MACE is trained on DFT data generated using computational settings that differ from those employed in this work.

Figure 3 shows the prediction errors of MACE and DFT with respect to experimental measurements.

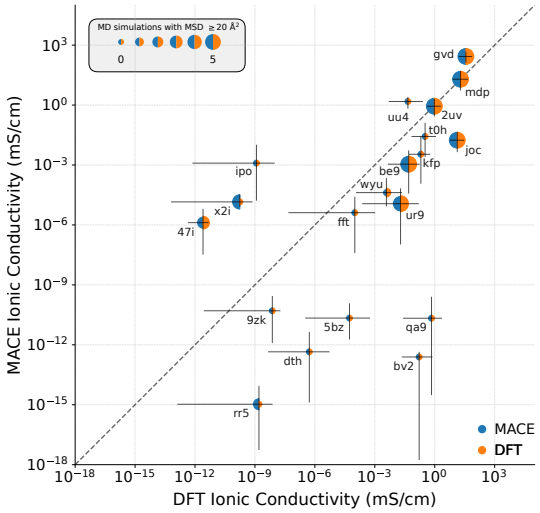

Figure 2: Parity plot comparing ionic conductivity computed using DFT and MACE foundation model. The size of the half circles reflect the number of simulations (at 5 different temperatures) that reached at least $\mathrm{MSD} = 20$ Å^2.

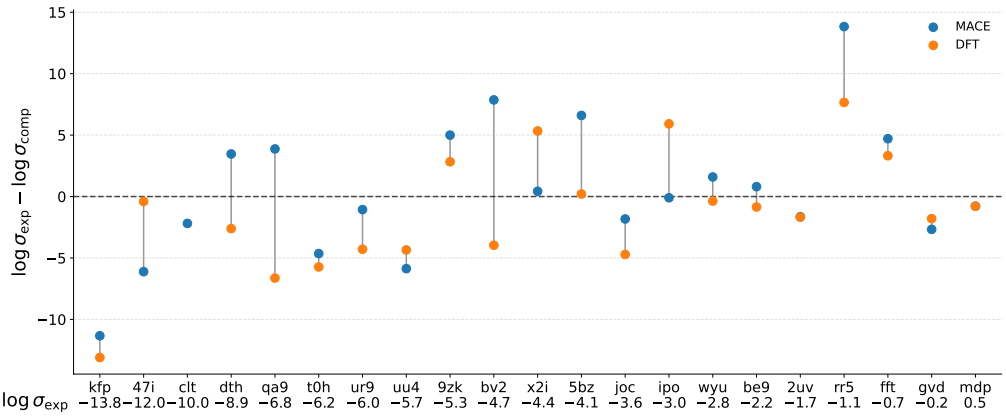

Figure 3: Logarithmic error between experimental and computed ionic conductivities for each material. The horizontal dashed line indicates perfect agreement. The x-axis lists materials equally spaced for clarity, with experimental $\log_{10} \sigma_{\text{exp}}\,[\text{mS/cm}]$ values shown below the labels. Materials are ordered by increasing experimental conductivity.