Recent developments in machine learning force fields: Foundation models, symmetries, constraints, and long-range interactions

Marcel F. Langer
Laboratory of Computational Science and Modeling
École Polytechnique Fédérale de Lausanne
1015 Lausanne, Switzerland
marcel.langer@epfl.ch

Molecular dynamics simulations are ubiquitous in physics, chemistry, and biology. While they have been traditionally driven by force fields, machine learning interatomic potentials (MLIPs) trained on first-principles reference data have emerged as a data-driven and more accurate, albeit less computationally efficient, alternative. Over the last few years, the field has developed rapidly. This talk will provide an overview of some recent developments and discusses their impact in practice.

Rather than training models for specific systems, so-called universal potentials, or foundation models, have been proposed recently.[1,2] They are trained on large databases with diverse chemistry and are expected to predict the dynamics of unseen systems to good accuracy. Operating at scale, these models have put into question many design principles for MLIPs that were previously considered essential. For instance, some models discard rotational invariance in favour of simpler and scalable architectures, while others even disregard the notion of energy-conserving forces. In this talk, I will discuss our recent work[3,4] on probing the impact of these design choices, finding that while rotational symmetry can be learned relatively easily from data, conservative forces are essential for physically meaningful and stable simulations.

While force fields have included long-range electrostatics and dispersion corrections for a long time, MLIPs have only recently begun to consider such effects, in part due to the lack of efficient and differentiable implementations[5]. Different mechanisms for including long-range interactions, ranging from simple physical terms added to the total energy to purely data-driven transformers, have been proposed. I present an overview of relevant benchmarks and discuss our recently long-range message passing model LOREM[6].

References

1. Batatia, I.; *et al.* A Foundation Model for Atomistic Materials Chemistry. arXiv March 1, 2024. https://doi.org/10.48550/arXiv.2401.00096.
2. Mazitov, A.; Bigi, F.; Kellner, M.; Pegolo, P.; Tisi, D.; Fraux, G.; Pozdnyakov, S.; Loche, P.; Ceriotti, M. PET-MAD, a Universal Interatomic Potential for Advanced Materials Modeling. arXiv March 18, 2025. https://doi.org/10.48550/arXiv.2503.14118.
3. Langer, M. F.; Pozdnyakov, S. N.; Ceriotti, M. Probing the Effects of Broken Symmetries in Machine Learning. Mach. Learn.: Sci. Technol. 2024, 5 (4), 04LT01. https://doi.org/10.1088/2632-2153/ad86a0.
4. Bigi, F.; Langer, M. F.; Ceriotti, M. The Dark Side of the Forces: Assessing Non-Conservative Force Models for Atomistic Machine Learning. In Forty-second international conference on machine learning; 2025.
5. Loche, P.; Huguenin-Dumittan, K. K.; Honarmand, M.; Xu, Q.; Rumiantsev, E.; How, W. B.; Langer, M. F.; Ceriotti, M. Fast and Flexible Long-Range Models for Atomistic Machine Learning. The Journal of Chemical Physics 2025, 162 (14), 142501. https://doi.org/10.1063/5.0251713.
6. Rumiantsev, E.; Langer, M.F.; Sodjargal, T.-E.; Ceriotti, M.; Loche, P., Learning Long-Range Representations with Equivariant Messages