

Self-Consistent Fine-Tuning of Machine Learning Interatomic Potentials by Targeted Exploration of Configuration Space

J. Steffen¹

¹*Friedrich-Alexander-University Erlangen-Nürnberg, Germany*

Universal machine learning interatomic potentials (uMLIPs) have shown an astonishing ability of describing almost arbitrary chemical systems reasonably. They are, however, not accurate enough for the calculation of quantities like reaction rate constants that require a very precise representation of the potential energy surface. Fortunately, uMLIPs can be fine-tuned for a system of interest, by training them on system-specific ab initio reference data. The generation of such reference data is nontrivial since it requires a broad sampling of the relevant configuration space, often out of reach for direct ab initio molecular dynamics (AIMD) [1]. Self-consistent fine-tuning is a solution for this: long MD samplings are done with the uMLIP, a relevant subset of sampled structures is recalculated with the ab initio method, and this is iterated with the successively fine-tuned MLIP to cure the slightly incorrect configurational sampling of the uMLIP [2].

In this contribution, I show how self-consistent fine-tuning of uMLIPs can be combined with umbrella samplings along complex reaction coordinates with the Caracal program package [3]. This enables the black-box setup of highly accurate targeted MLIPs for different chemical reaction mechanisms in arbitrary environments. I present the application of this method on a benchmark study of small proton exchange reactions in the gas phase [4], and then continue with real-world applications - based on actual ab initio data - for larger gas phase systems and hydrogen diffusions on different metal surfaces. In all examples, ring polymer molecular dynamics (RPMD) is used to consider nuclear quantum effects. Finally, an outlook is given on how to generalize the method to arbitrary chemical systems and reaction mechanisms.

[1] J. Steffen, A. Alibakshi, J. Chem. Phys. 161, 184116 (2024).

[2] J. Steffen, J. Phys. Chem. C, 129, 13513 (2025).

[3] J. Steffen, J. Chem. Theory Comput., 19, 5334 (2023).

[4] J. Steffen, submitted (DOI: 10.26434/chemrxiv.15002012/v1).