Multiscale machine learning: from quantum chemistry to dislocation dynamics.

Boris Kozinsky[1,2]
[1] Harvard University School of Engineering and Applied Sciences, Cambridge, MA USA
[2]Bosch Research, Watertown, MA USA

bkoz@g.harvard.edu

Discovery and understanding of next-generation materials requires a challenging combination of the high accuracy of first-principles calculations with the ability to reach large size and time scales. We pursue a multi-tier method development strategy in which machine learning (ML) algorithms are combined with exact physical symmetries and constraints to significantly accelerate computations of electronic structure and atomistic dynamics. First, density functional theory (DFT) is the cornerstone of modern computational materials science, but its current approximations fall short of the required accuracy and efficiency for predictive calculations of defect properties, band gaps, stability and electrochemical potentials of materials for energy storage and conversion. To advance the capability of DFT we introduce non-local charge density descriptors that satisfy exact scaling constraints and learn exchange functionals called CIDER [1]. These models are orders of magnitude faster in self-consistent calculations for solids than hybrid functionals but similar in accuracy. On a different level, we accelerate molecular dynamics (MD) simulations by using machine learning to capture the potential energy surfaces obtained from quantum calculations. We developed NequIP [2] and Allegro [3], as well as extensions the first deep equivariant neural network interatomic potential models, whose Euclidean symmetry-preserving layer architecture achieves state-of-the-art data efficiency and accuracy for simulating dynamics of molecules and materials. In parallel, we implement autonomous active learning of interactions in reactive systems, with the FLARE algorithm that constructs accurate and uncertainty-aware Bayesian force fields on-the-fly from a molecular dynamics simulation, using Gaussian process regression [4]. These MD simulations are used to explore long-time dynamics of phase transformations and heterogeneous reactions.

References

[1] K. Bystrom, B. Kozinsky, arXiv:2303.00682 (2023)
[2] A. Musaelian, S. Batzner et al, Nature Comm. 14, 579 (2023)
[3] J. Vandermause et al, Nature Comm. 13 (1), 5183 (2022)
[4]

Figures

Figure 1. Large-scale heterogeneous catalytic reaction dynamics simulated using Bayesian machine learning force fields with FLARE, capable of reaching 1 trillion atoms at quantum accuracy.