Merits of unconstrained atomistic machine learning

Sergey Pozdnyakov[1], Michele Ceriotti[1]
[1]Laboratory of Computational Science and Modelling, Institute of Materials, Ecole Polytechnique Fédérale de Lausanne, Lausanne 1015, Switzerland

sergey.pozdnyakov@epfl.ch

In the last decade, methods incorporating rotational symmetry into their functional form have dominated the field of atomistic machine learning. Two main considerations have driven this design choice. Firstly, it was widely believed that rotational symmetry, when intrinsically built into a machine learning architecture, plays the role of necessary inductive bias, thus being crucial for a model's performance and generalizability. Secondly, rigorous rotational invariance might be necessary for atomistic simulations such as molecular dynamics to avoid subtle artifacts. In this talk, I will challenge the first belief by presenting an unconstrained model, Point Edge Transformer (PET), which is not rotationally invariant and instead relies on rotational augmentations during fitting. As Figure and Table 1 show, PET not only achieves state-of-the-art performance on multiple benchmark datasets of molecules and solids but also improves faster with the increase of the training data compared to other methods. To address the second consideration, we introduce a general symmetrization method that a-posteriori enforces rigorous rotational equivariance for any backbone architecture, which might even be superfluous given how accurate PET's learned rotational equivariance is. In the final part, I will present a few general considerations of why unconstrained architectures are likely more efficient, especially in the data-rich regime, which might explain PET's excellent and often superior performance and more favorable scaling laws. These include: (i) in contrast to the equivariant world, **each** shallow layer of an unconstrained architecture is provably and trivially a universal approximator; (ii) there is no need in the expensive SO(3) algebra, which allows to achieve an unlimited angular resolution inexpensively and favors the computational efficiency overall.

References

[1] Pozdnyakov, Sergey, and Michele Ceriotti. "Smooth, exact rotational symmetrization for deep learning on point clouds." Advances in Neural Information Processing Systems 36 (2024).

Figures

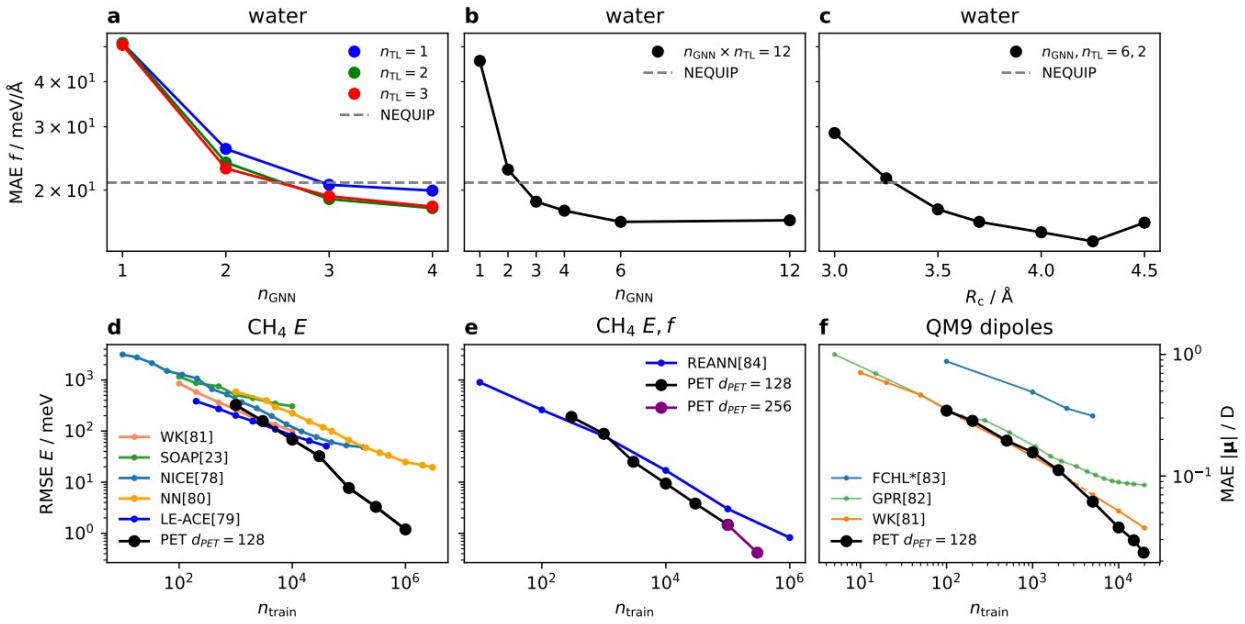

Figure 1. (a-c) Accuracy of PET potentials on liquid water, compared with NEQUIP[35]. (d-f) Learning curves for different molecular data sets, comparing symmetrized PET models with several previous works[23, 78–84], including the current state of the art. (d) Random CH4 dataset, training only on energies; (e) Random CH4 dataset, training on energies and forces; (f) Vectorial dipole moments in the QM9 dataset[82]. References in the legend and caption match the ones of Ref[1].

dataset	COLL		MnO		HME21		HEA	
metric	MAE f	MAE E	RMSE f	RMSE E/at.	MAE $\|f\|$	MAE E/at.	MAE f	MAE E/at.
SOTA	26.4[17]	47[86]	125[88]	1.11[88]	138[89]	15.7[89]	190[90]	10[90]
model	GemNet	DimeNet++	mHDNNP		MACE		HEA25-4-NN	
PET (y_0)	23.1	12.0	22.7	0.312	140.5 ± 2.0	17.8 ± 0.1	60.2	1.87
PET (y_S)	23.1	11.9	22.7	0.304	141.6 ± 1.9	17.8 ± 0.1	60.1	1.87
PET (ens.)					128.5	16.8		

Table 1. Comparison of the accuracy of PET and current state-of-the-art models for the COLL, MnO, HM21 and HEA data sets. Energy errors are given in meV/atom, force errors in meV/Å. y_0 and y_S indicate unsymmetrized and symmetrized models, respectively. References match the ones of Ref[1].