Equivariant Transformers for Neural Network based Molecular Potentials

https://www.compscience.org/

Philipp Thölke[1,2,3] and Gianni de Fabritiis[1,4]

philipp.thoelke@posteo.de g.defabritiis@gmail.com

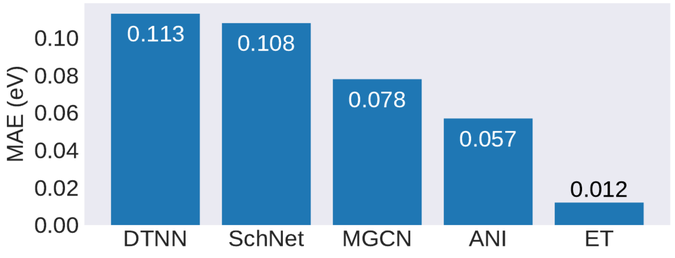
Pompeu Fabra University[1]
Osnabrück University[2]
Université de Montréal[3]
Institució Catalana de Recerca i Estudis Avançats (ICREA)[4]

Overview

- attention-centered Neural Network Potential (NNP)
- molecular dynamics simulation or quantum mechanical property prediction
- layered architecture allows efficient encoding of many-body interactions
- improved force predictions through the use of rotationally equivariant features
- fully differentiable, CUDA-optimized, conservative force field
- evaluation on benchmarks MD17 and ANI-1

ANI-1

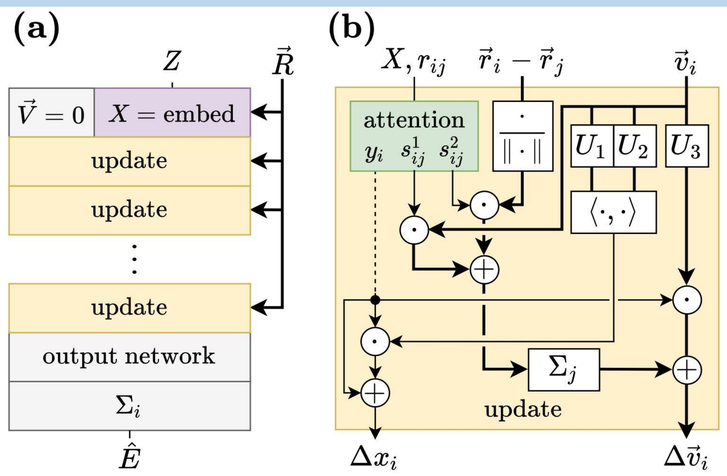
- > 22 million configurations of molecules with up to 8 heavy atoms (train: 80%, val: 5%)
- normal mode sampling of > 50,000 molecules
- DFT level of accuracy (only energies)

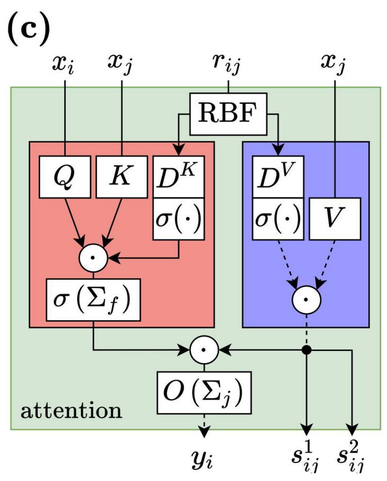


MD17

- MD trajectories of 8 molecules
- benchmark to test data efficiency (train: 950, val: 50)
- DFT labels (energies and forces)
- table shows MAE in kcal/mol (energy) and kcal/mol/Å (forces)

Molecule		SchNet	PhysNet	DimeNet	PaiNN	NequIP	ET
Aspirin	*energy*	0.37	0.230	0.204	0.167	-	**0.123**
	forces	1.35	0.605	0.499	0.338	0.348	**0.253**
Benzene	*energy*	0.08	-	0.078	-	-	**0.058**
	forces	0.31	-	**0.187**	-	**0.187**	0.196
Ethanol	*energy*	0.08	0.059	0.064	0.064	-	**0.052**
	forces	0.39	0.160	0.230	0.224	0.208	**0.109**
Malondialdehyde	*energy*	0.13	0.094	0.104	0.091	-	**0.077**
	forces	0.66	0.319	0.383	0.319	0.337	**0.169**
Naphthalene	*energy*	0.16	0.142	0.122	0.116	-	**0.085**
	forces	0.58	0.310	0.215	0.077	0.097	**0.061**
Salicylic Acid	*energy*	0.20	0.126	0.134	0.116	-	**0.093**
	forces	0.85	0.337	0.374	0.195	0.238	**0.129**
Toluene	*energy*	0.12	0.100	0.102	0.095	-	**0.074**
	forces	0.57	0.191	0.216	0.094	0.101	**0.067**
Uracil	*energy*	0.14	0.108	0.115	0.106	-	**0.095**
	forces	0.56	0.218	0.301	0.139	0.173	**0.095**



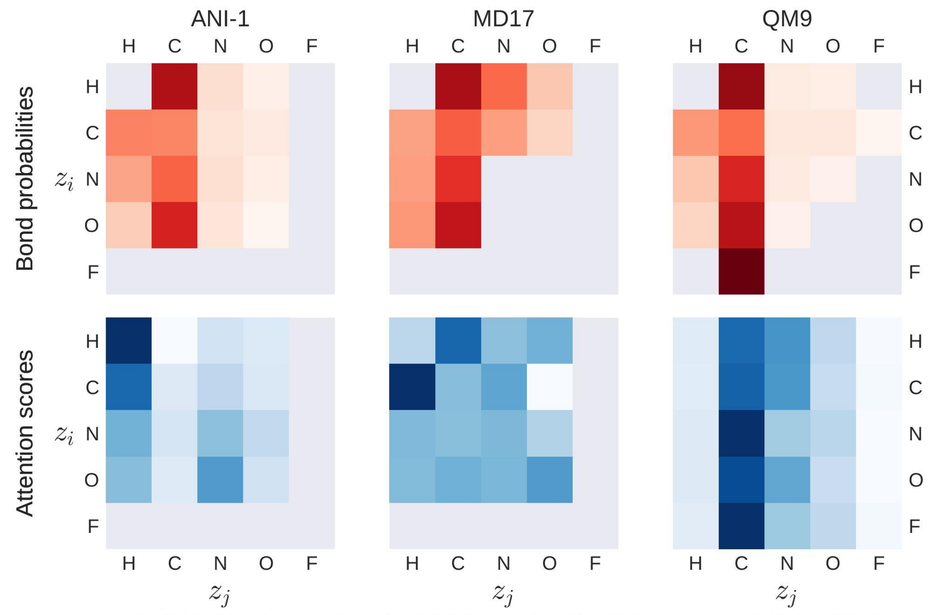


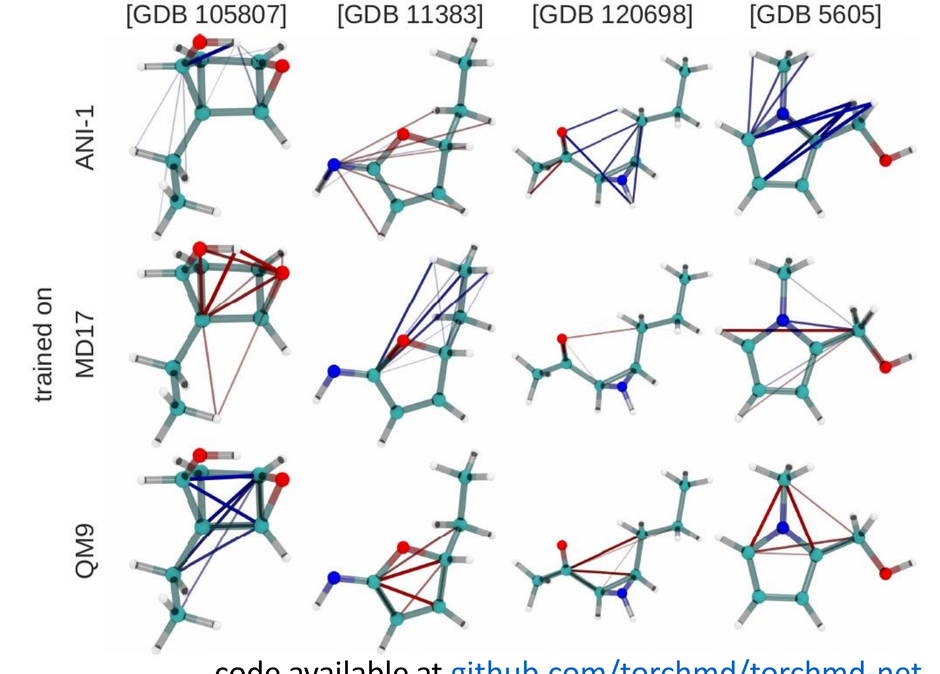
Architecture

a) equivariant Transformer (ET) architecture overview
b) update layer with attention mechanism and exchange between scalar and equivariant features
c) modified attention mechanism using edge-wise distance features

Attention Weight Analysis

- attention weights highlight differences in QM9, MD17 and ANI-1
- core structural atoms are most important in static data (QM9)
- hydrogen is attended most when trained on dynamical data (MD17, ANI-1)
- ET learned chemical representations of the underlying data





code available at github.com/torchmd/torchmd-net