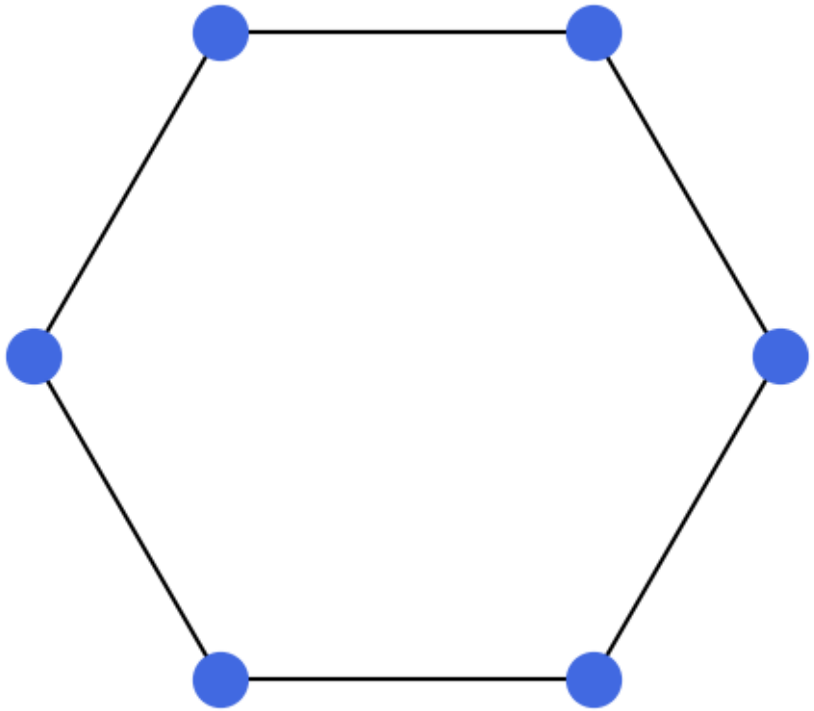
Performance of an Attention-based Model on Atomic Systems

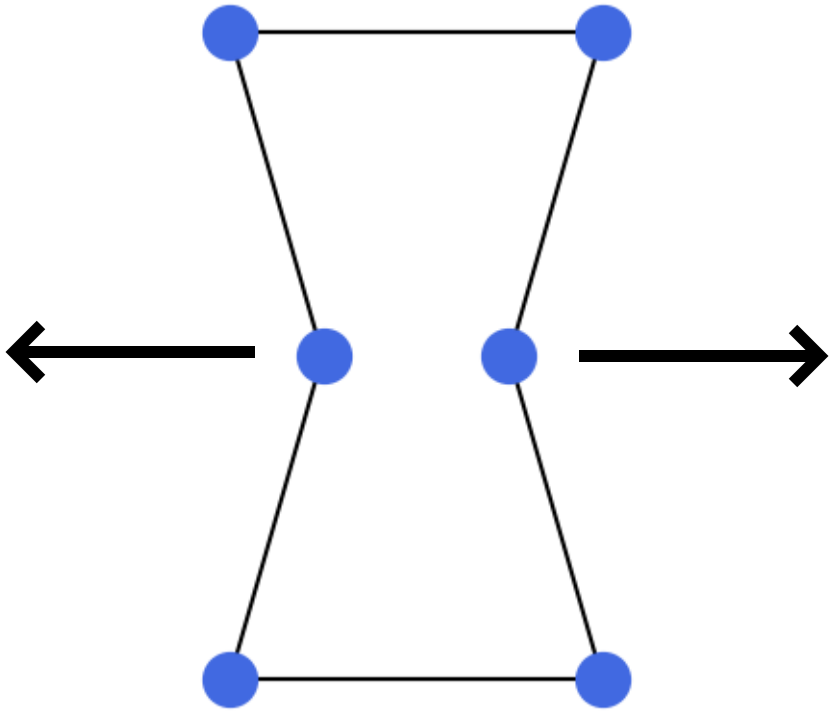
Praharsh Suryadevara

How do atoms configure themselves?



C_1

??



C_2

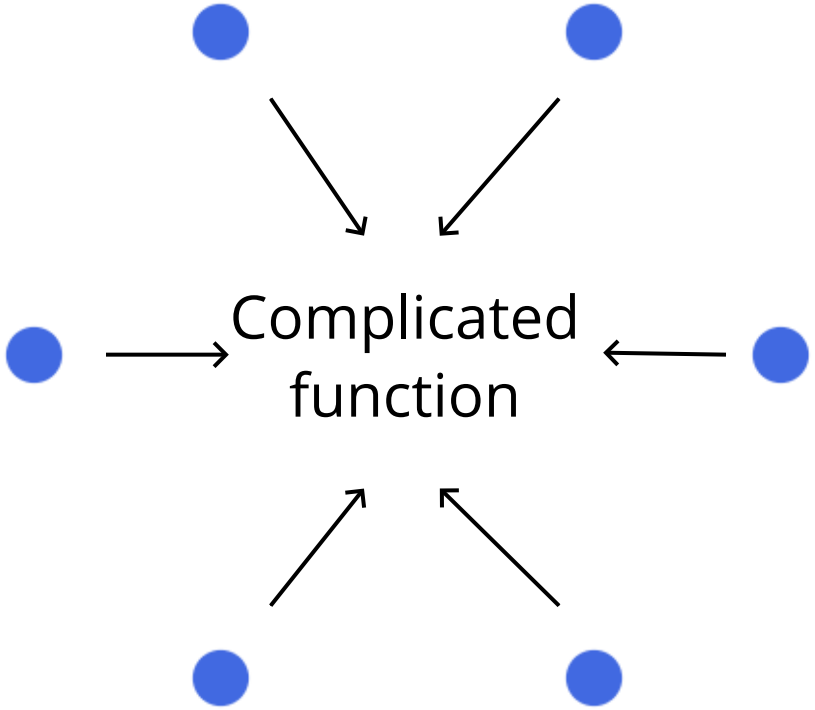
$$E(C_1) < E(C_2)$$

$$\vec{F}(C_1) = 0$$

This is consequential and hard

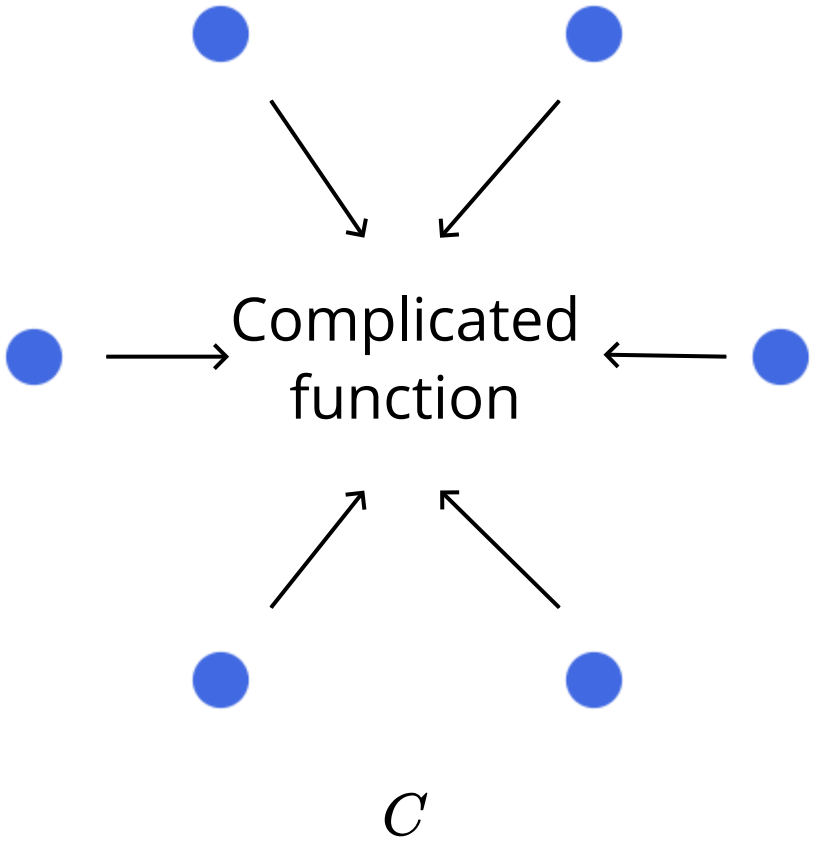
Drug design

New photovoltaic materials



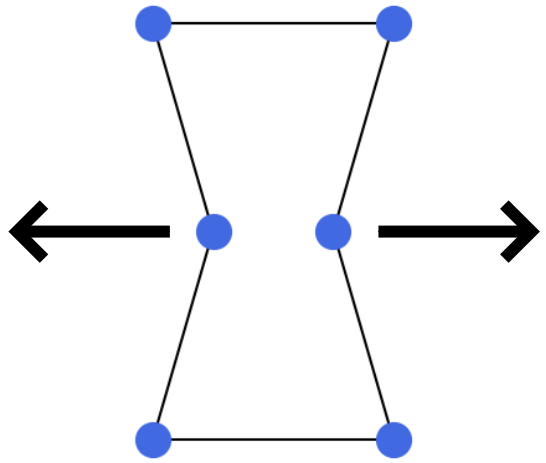
Evaluating $E(C_1)$ is $\mathcal{O}(d^{n_e})$ exactly and $\mathcal{O}(n_e^3)$ approximately

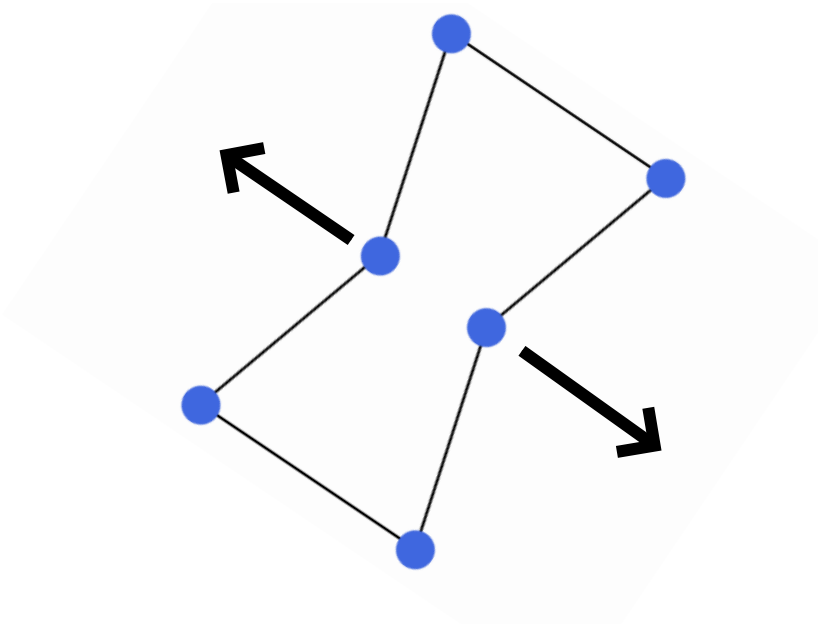
Task



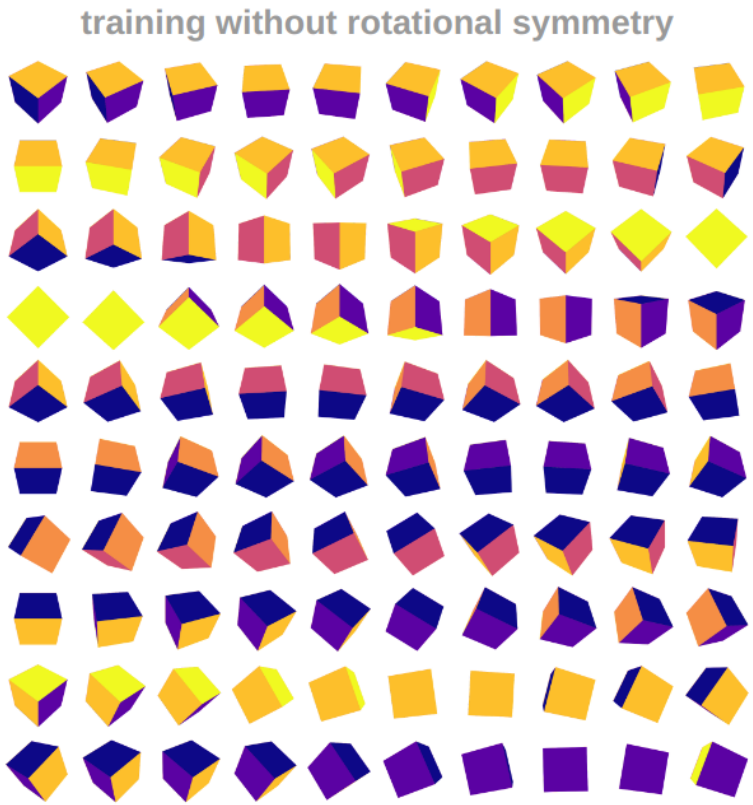
Given positions of atoms C predict $\vec{F}(C)$ and $E(C)$

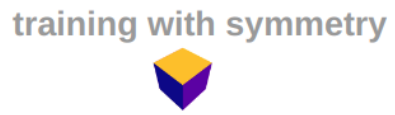
Equivariance: Rotational symmetry





$$\vec{F}(R(C)) = R(\vec{F}(C))$$





In $3d$ 500x the cost

https://e3nn.org/

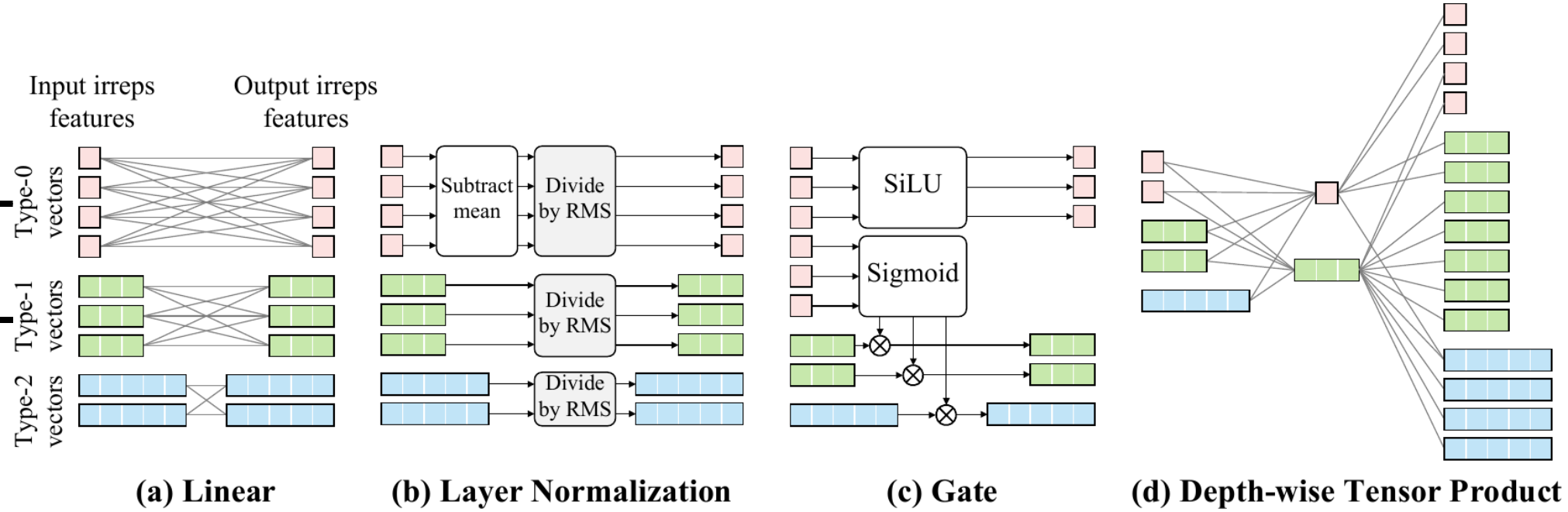
Equivariance: Rotational symmetry

Every Layer preserves rotational information

Decompose atom-atom interactions into
Type-L vectors

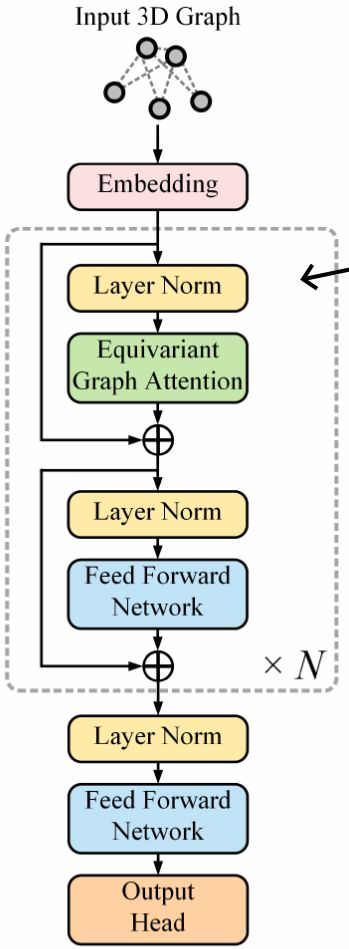
Example: energy

Example: Force

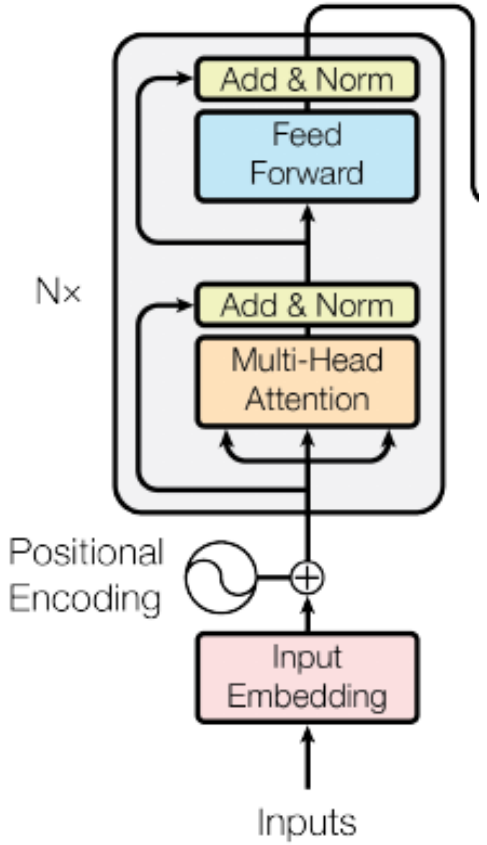


Equiformer: Equivariant Graph Attention Transformer for 3D Atomistic Graphs, Yi-Lun Liao et al

Equiformer = Transformer + Equivariance



Extra norm in the beginning

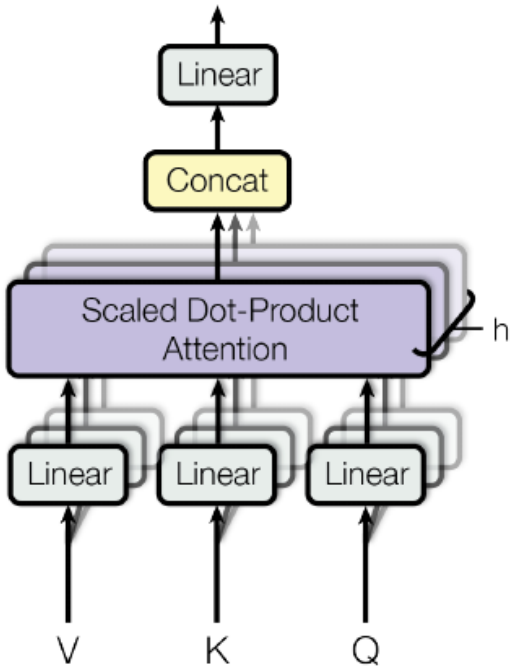


(a) Equiformer Architecture

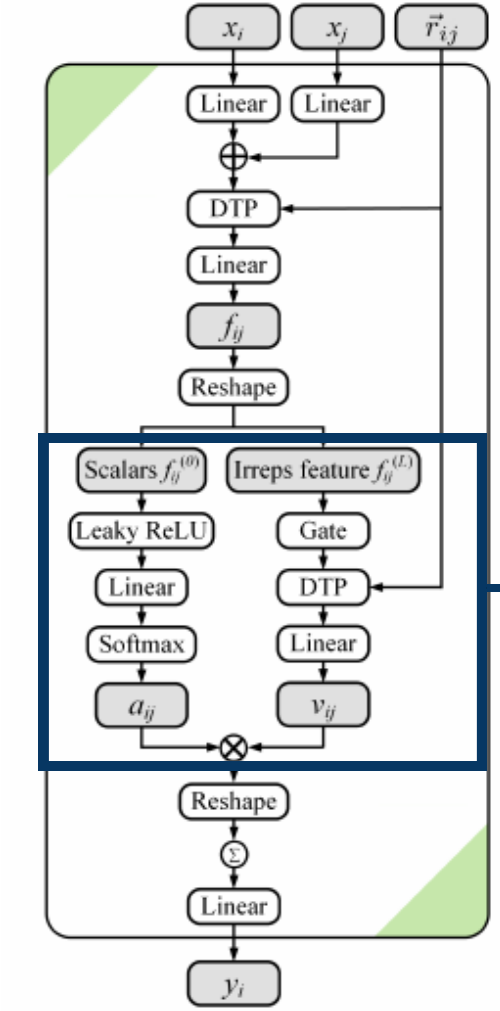
Equiformer: Equivariant Graph Attention Transformer for 3D Atomistic Graphs, Yi-Lun Liao et al

Attention Is All You Need

Equivariant graph attention



Attention Is All You Need

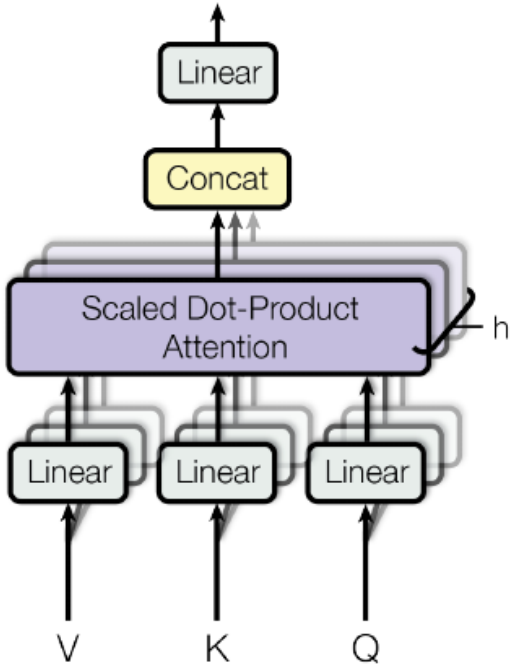


(b) Equivariant Graph Attention

Nonlinear Message passing

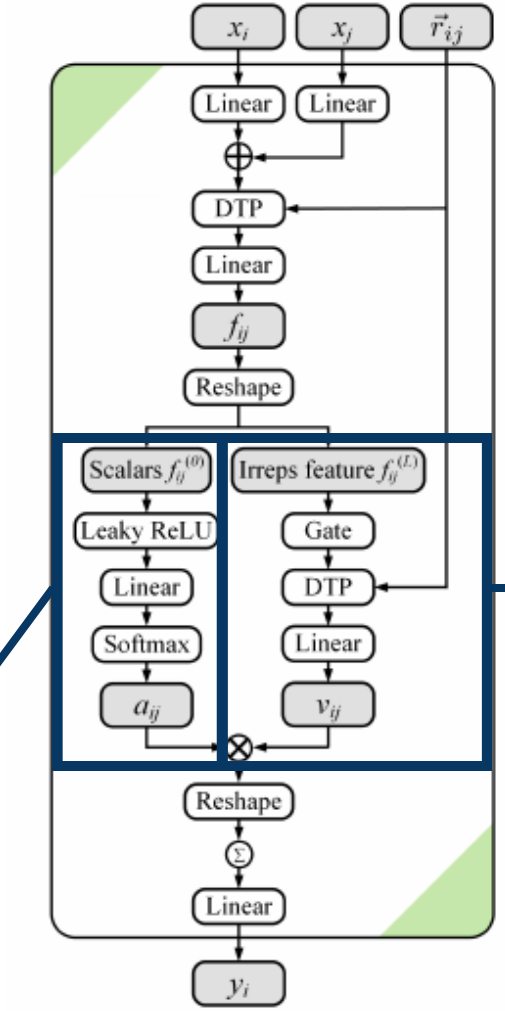
Equiformer: Equivariant Graph Attention Transformer for 3D Atomistic Graphs, Yi-Lun Liao et al

Erratum: Equivariant graph attention



Attention Is All You Need

MLP Attention



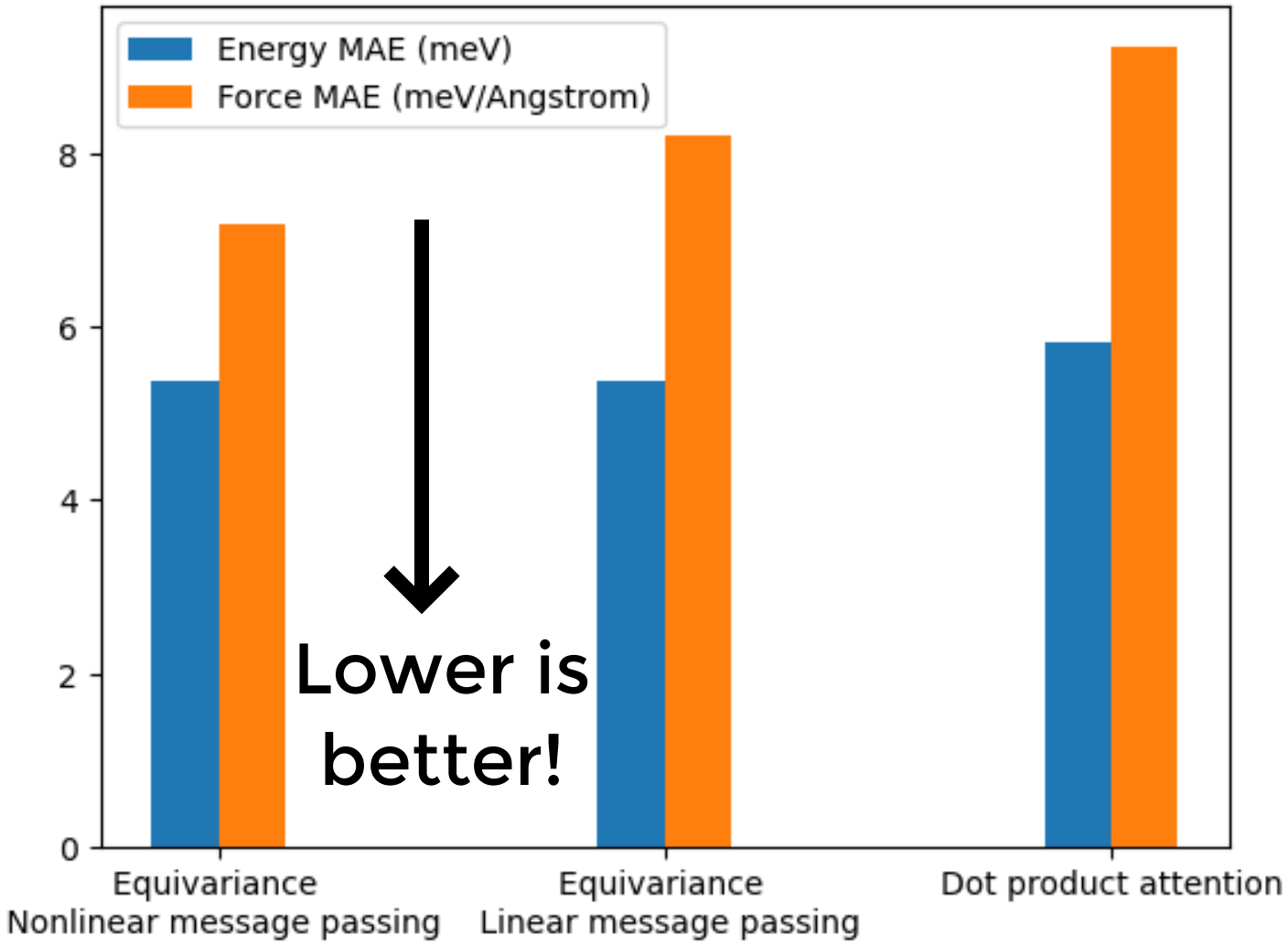
(b) Equivariant Graph Attention

Equiformer: Equivariant Graph Attention Transformer for 3D Atomistic Graphs, Yi-Lun Liao et al

Nonlinear Message passing

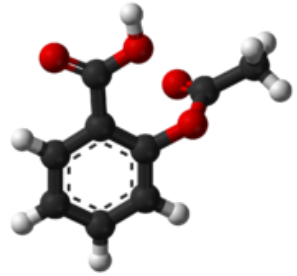
Results: Aspirin MD17

MAE on Test Set



Best!

Force MAE matches exactly, energy matches upto ≈ 0.1 meV

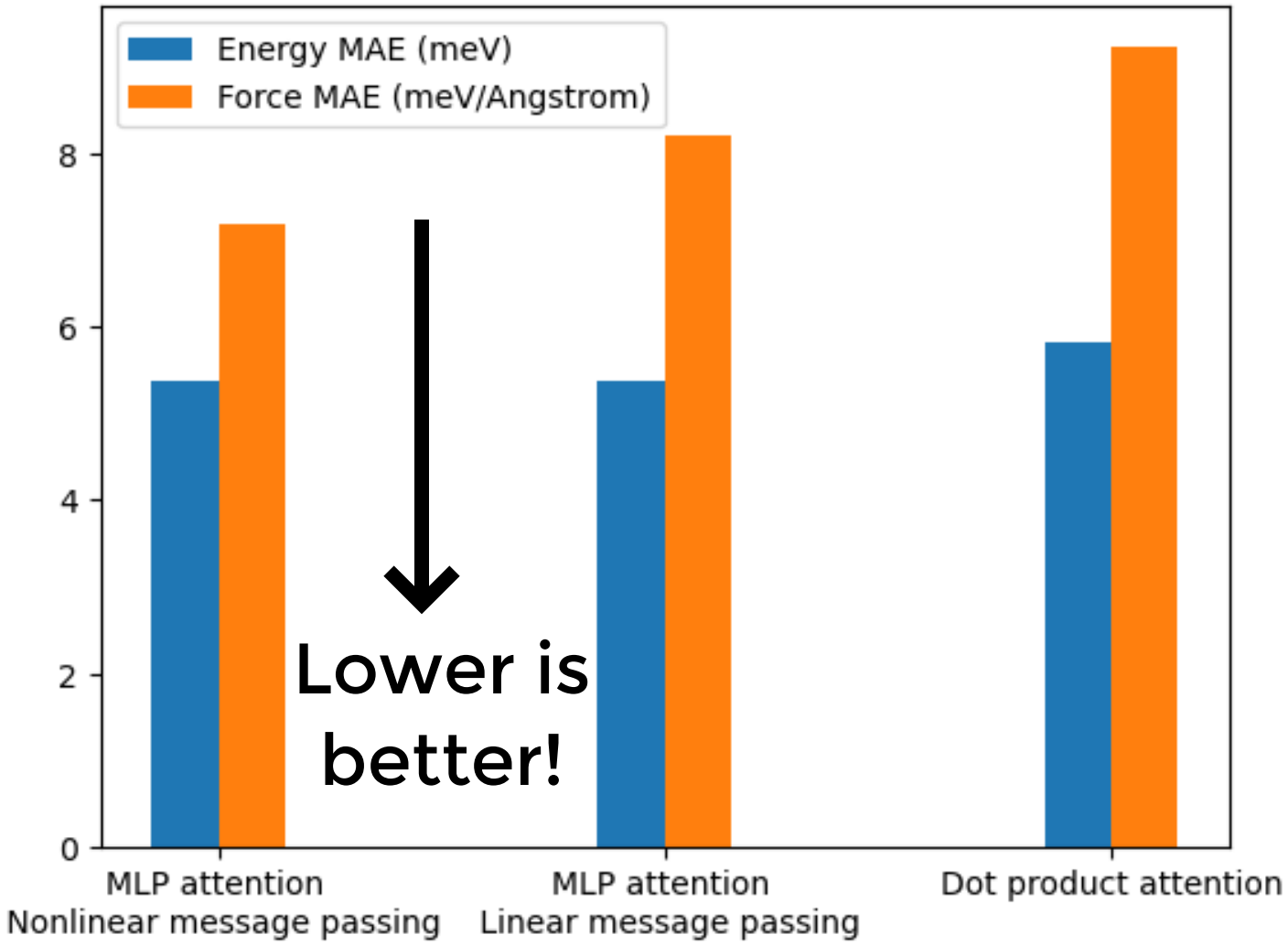


1500 epochs

1. Attention models with equivariance gives SOTA on atomic force and energy predictions
2. Ablation studies show equivariance and non-linear message passing improve performance!

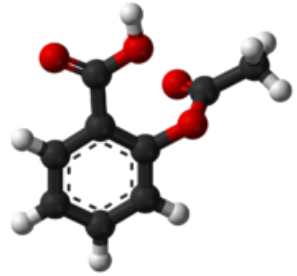
Erratum: Results: Aspirin MD17

MAE on Test Set



Best!

Force MAE matches exactly, energy matches upto ≈ 0.1 meV



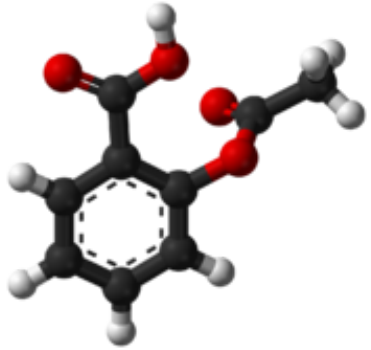
1500 epochs

1. Attention models with equivariance gives SOTA on atomic force and energy predictions
2. Ablation studies show MLP attention and non-linear message passing improve performance!

Backup

Results: Aspirin

Model	Energy MAE	Force MAE	Energy MAE (original)	Force MAE (original)	Parameters
Non-linear message passing + MLP	**5.4**	**7.2**	5.3	7.2	3.5 million
Linear message passing + MLP	**5.4**	8.2	-	-	2.9 million
Dot product attention	5.8	9.2	-	-	3.3 million



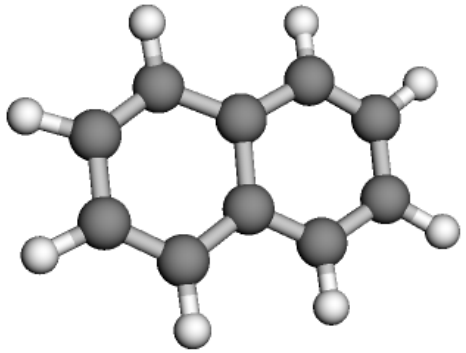
1500 epochs: ~1.5 days per run

Ablation studies show MLP and non-linear message passing make a difference!

Results: Other

MD17

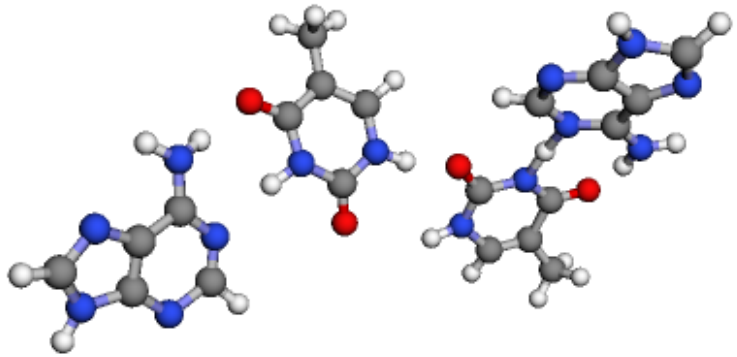
1. Training for ~950 epochs done for full model on Ethanol, Malonaldehyde, Naphthalene, Salicyclic_acid
2. Hit GPU hour limits



Naphthalene

MD22

1. Training attempted on DNA base pairs and Ac-Ala3-NHMe
2. Hit Memory limits



DNA base pair (AT-AT)

Acknowledgements

Nitish Joshi

NYU HPC