ML4Jets 2021

Equivariant Energy Flow Networks for jet tagging

Ayodele Ore
The University of Melbourne
ayodeleo@student.unimelb.edu.au

July 6, 2021

Based on work with M. J. Dolan
Phys. Rev. D 103 074022 [2012.00964]

Jet representations

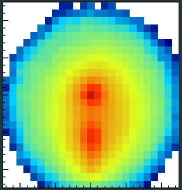
List

$\tau_1 = 1.7$
$\tau_2 = 1.2$
$\tau_3 = 1.0$
$\vdots$
$\tau_N = 2.1$

↓

DNN / RNN

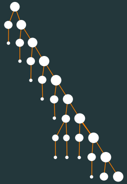
Image



↓

CNN

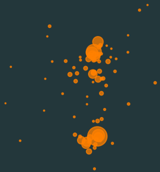
Tree



↓

RecNN

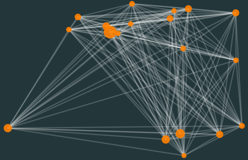
Point cloud



↓

DeepSets

Graph



↓

GNN

A permutation-invariant function can be composed into two functions:

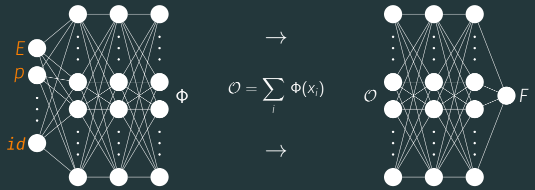
$$g(\{x_i\}) : R^{|x|} \times R^{\text{in}} \to R^{\text{out}}$$

$$\Phi : R^{\text{in}} \to R^{\ell}$$

$$F : R^{\ell} \to R^{\text{out}}$$

Join with a sum: $g(\{x_i\}) = F\left(\sum_i \Phi(x_i)\right)$

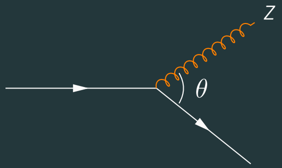
Applying to jets gives PFN [1810.05165]:



IRC safety and Energy Flow Networks

1810.05165

IRC safety is invariance to splittings with small opening angle or emission energy.



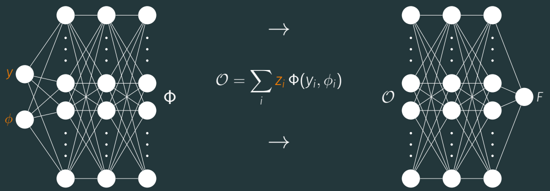
$$dP \sim \frac{dz}{z}\frac{d\theta}{\theta}$$

$z \to 0$ "IR"

$\theta \to 0$ "C"

IRC safety of EFN output can be enforced $\left[\Phi = \Phi(\vec{p}) \text{ and } \Phi \propto z\right]$

$$EFN(\{p_i\}) = F\left(\sum_i z_i\, \Phi(y_i, \phi_i)\right)$$



Quantify importance of IRC-safe info + Get new IRC-safe observables

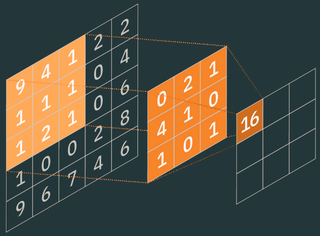
Equivariance in neural networks

A function $f : \boldsymbol{R}^n \to \boldsymbol{R}^n$ is equivariant to a transformation π iff

$$f(\pi(x)) = \pi(f(x))$$

In symmetric tasks, reduces parameters via weight sharing:

- Translation equivariance in CNNs
- Lorentz group equivariance in LGNs [2006.04780]



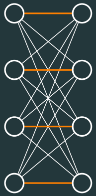
Reduces burden on network to learn the symmetry.

Permutation-equivariant network layers

1703.06114

For a neural network layer $f_\Theta(x) = \sigma\left(\Theta \cdot x\right)$ permutation equivariance requires

$$\Theta = \lambda I + \gamma(11^T) \quad \sim \quad \begin{pmatrix} \lambda+\gamma & \gamma & \gamma & \gamma \\ \gamma & \lambda+\gamma & \gamma & \gamma \\ \gamma & \gamma & \lambda+\gamma & \gamma \\ \gamma & \gamma & \gamma & \lambda+\gamma \end{pmatrix} \quad \sim$$



Then $(\Theta x)_i = \lambda x_i + \gamma \sum x$

In general, just need $f_i(x) = \sigma\left[\lambda x_i + \gamma \mathcal{Q}(x)\right]$ with $\mathcal{Q} \in \{\text{sum}, \text{mean}, \text{max}, \cdots\}$

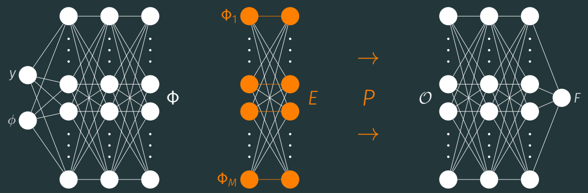
Equivariant Deep Sets

Take advantage of:

1. Equivariant layers can be stacked
2. Can construct permutation invariant operation

Then can recast decomposition:

$$F\left(\sum_i \Phi(x_i)\right) \quad \rightarrow \quad F \circ P \circ E\left(\{\Phi(x_i)\}\right)$$



Enforcing IRC safety

The new network is

$$\text{EV-EFN}(\{p_i\}) = F \circ P \circ E\Big(\{\Phi(y_i, \phi_i)\}\Big) \equiv F(\mathcal{O})$$

The steps to IRC safety are:

1. Choose same P as EFN
2. Parameterise E with jet observable
3. Choose IRC safe observable

$$\mathcal{O} = P \circ E(\Phi)$$

1 → $P(E) = \sum_i z_i E_i$

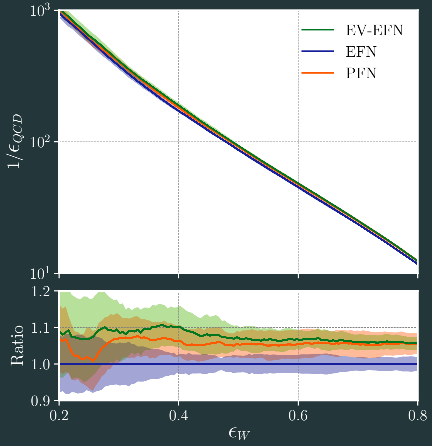
2 → $E_i(\Phi) = \sigma\left(\lambda\Phi_i + \gamma\mathcal{Q}(\Phi)\right)$

3 → $\mathcal{Q}(z, \Phi) = \sum_i z_i \Phi_i$

In full, observables are

$$\mathcal{O} = \sum_i z_i \sigma \left(\lambda\Phi_i + \gamma \sum_j z_j \Phi_j \right)$$

Results



Model	AUC	$1/\varepsilon_{QCD}$ at $\varepsilon_W = 0.5$
EFN	0.9339 ± 0.0007	87.4 ± 5.7
EV-EFN	0.9367 ± 0.0009	93.4 ± 2.5
PFN	0.9366 ± 0.0012	93.2 ± 4.6
EV-PFN	0.9333 ± 0.0026	80.1 ± 3.8

Recovering an EFN from an EV-EFN

What limits on the equivariant layer recover an EFN?

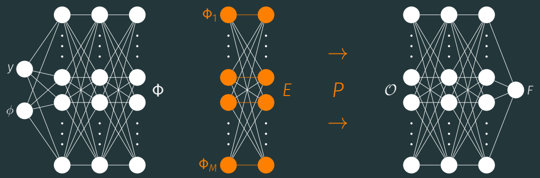
There are three such limits:

1. Taking $\sigma \equiv id \rightarrow$ Redefinition of Φ
2. Taking $\gamma \equiv 0 \rightarrow$ Additional layers in Φ
3. Taking $\lambda \equiv 0 \rightarrow$ Additional layers in F

Model	AUC	$1/\varepsilon_{QCD}$ at $\varepsilon_W = 0.5$
$\text{EV-EFN}\big\rvert_{\sigma\equiv id}$	0.9320 ± 0.0009	83.5 ± 5.2
$\text{EV-EFN}\big\rvert_{\gamma\equiv 0}$	0.9327 ± 0.0013	85.0 ± 1.9
$\text{EV-EFN}\big\rvert_{\lambda\equiv 0}$	0.9337 ± 0.0008	87.6 ± 2.4
EFN	0.9339 ± 0.0007	87.4 ± 5.7
EV-EFN	0.9367 ± 0.0009	93.4 ± 2.5

Summary

- Found IRC-safe implementation of permutation-equivariant layers in EFN architecture.
- Performance of new model on W boson tagging matches that of IRC-unsafe PFN.
- Verified that network performance degrades when restricting equivariance.



Mass sculpting