Equivariant Graph Neural Networks for Charged Particle Tracking

Ameya Thete[1], Daniel Murnane[2], Savannah Thais[3]

[1] BITS, Pilani - KK Birla Goa Campus, [2] Lawrence Berkeley National Laboratory, [3] Columbia University

Motivation

Track reconstruction is essential to many physics reconstruction tasks!

- The HL-LHC demonstrates the need for new tracking algorithms with reduced latency and improved performance in high-pileup environments [1].
- Traditional Kalman filter-based algorithms scale poorly with detector occupancy.
- Enforcing expected equivariance is a proposed technique to decrease model size while maintaining performance.
- Our work focuses on rotation-equivariant edge-classifying GNNs
- We explore the rotation groups, specifically SO(3) and SO(2) symmetry groups.

Dataset

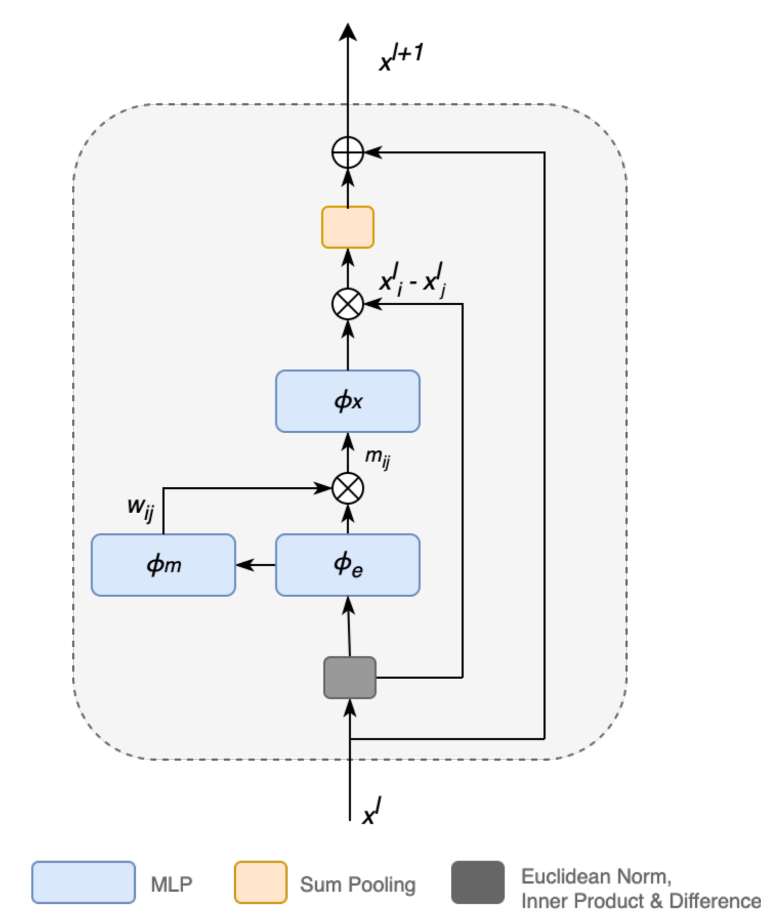
- **TrackML [2]:** ACTS simulated $p-p$ collision events with ~200 pileup interactions.
- Each event contains 3D hit position (x, y, z) and truth information about the particle that generated them.
- We only include hits generated in the pixel layers in the innermost region of the tracker.
- 3 filters applied to the dataset:
  - A $p_T^{\min}$ filter to reject hits generated by particles with $p_T < p_T^{\min}$
  - Noise filter to reject noise hits
  - Same-layer filter to ensure only one hit per particle per layer
- Graph constructed by mapping hits to nodes and edges to possible track segments
  - Node features are 3D hit positions
  - Edge e_{ij} constructed if it satisfies constraints on the geometric quantities

$$z_0 = z_i - r_i \frac{z_j - z_i}{r_j - r_i} \text{ and } \phi_{\text{slope}} = \frac{\phi_j - \phi_i}{r_j - r_i}$$

Model Architecture

- To enforce equivariance, the exchanged messaged must be constructed using equivariant information.
- Our architecture, EuclidNet, consists of repeating equivariant blocks (EB).
- Two variants are developed: SO(2) and SO(3), for each of the symmetry group being studied.

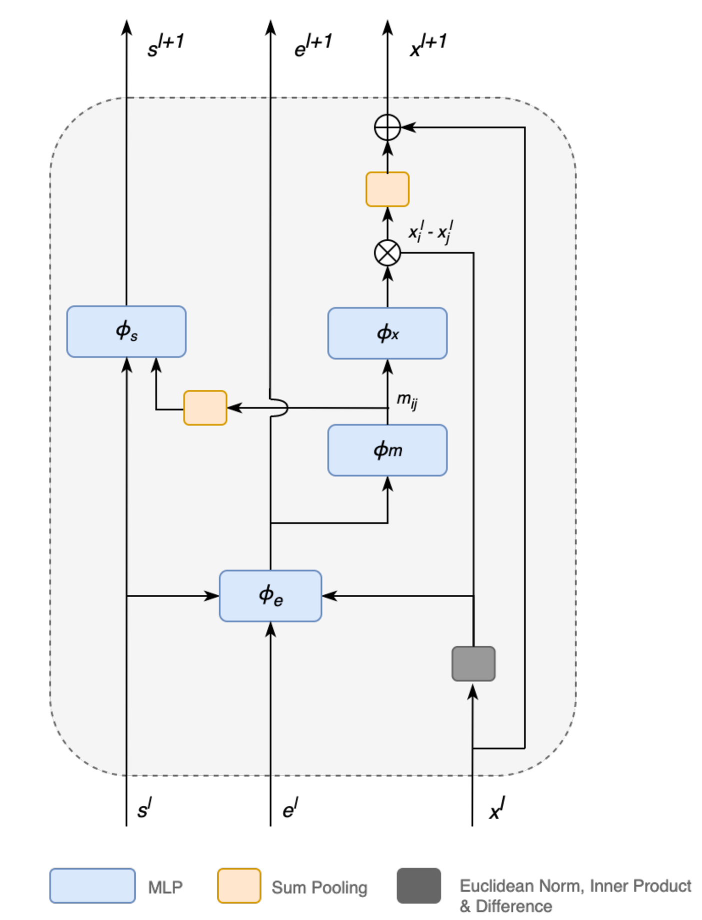
SO(3) EuclidNet EB



$$m_{ij}^l = \phi_e\left(\psi(\langle x_i^l, x_j^l \rangle),\ \psi(||x_i^l - x_j^l||^2)\right)$$

$$x_i^{l+1} = x_i^l + c \sum_{j \in [N]} \phi_x(m_{ij}^l) \cdot (x_i^l - x_j^l)$$

SO(2) EuclidNet EB



$$m_{ij}^l = \phi_e\left(\psi(\langle x_i^l, x_j^l \rangle),\ \psi(||x_i^l - x_j^l||^2), s_i^l, s_j^l, e_{ij}^l\right)$$

$$x_i^{l+1} = x_i^l + c \sum_{j \in [N]} \phi_x(m_{ij}^l) \cdot (x_i^l - x_j^l)$$

$$s_i^{l+1} = s_i^l + \phi_s(s_i^l, \sum_{j \in [N]} m_{ij}^l)$$

$$e_{ij}^{l+1} = \phi_e(m_{ij}^l)$$

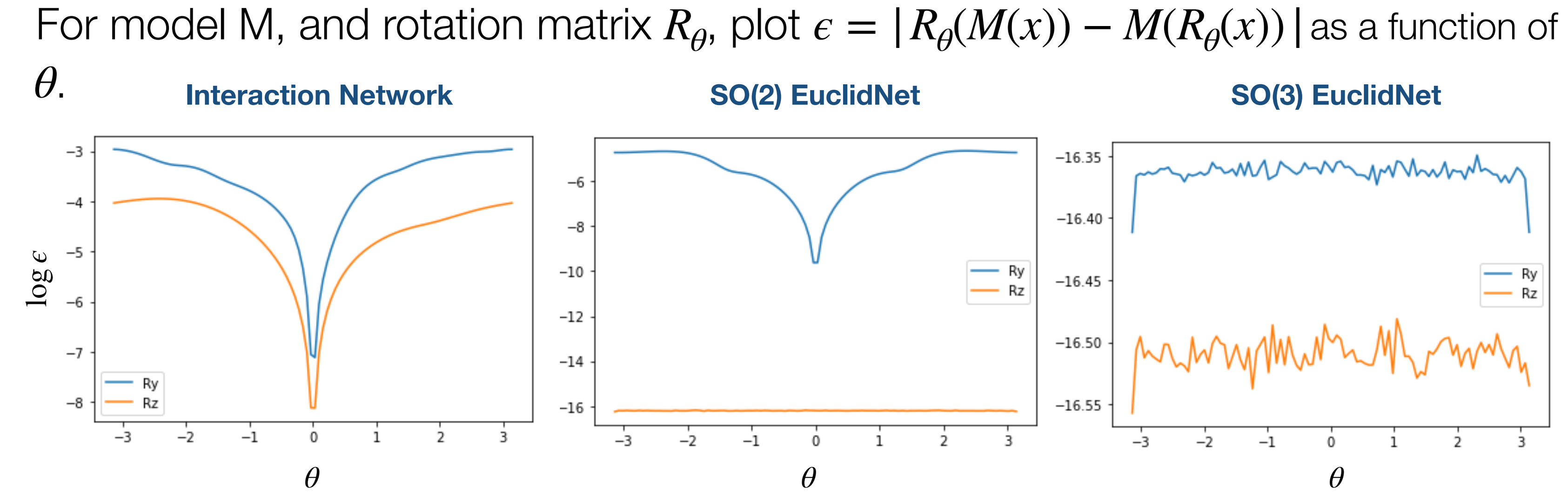
- Depending on which quantities are pooled and passed to the decoding MLP, the output can be chosen to be equivariant or non-equivariant

Experiments

- The Interaction Network (IN) from [3] is used as our benchmark.
- We systematically vary the different hyperparameters (HP) using the Weights & Biases API, and select the best HP set.
- To demonstrate the effect of model size, we study models with two different hidden channel dimensions: 8, and 16.
- To compare model performance, we track the model's ROC AUC (area under the curve), efficiency, and purity. The efficiency and purity are defined as:
  - Efficiency = TP/(TP + FP)
  - Purity = TP/(TP + FN)

We run an equivariance test to measure the robustness of the baked-in symmetries.

For model M, and rotation matrix R_θ, plot $\epsilon = |R_\theta(M(x)) - M(R_\theta(x))|$ as a function of θ.



Results

SO(3) EuclidNet

	n_hidden = 8				n_hidden = 16			
Model	**# Parameter**	**AUC**	**Efficiency**	**Purity**	**# Parameters**	**AUC**	**Efficiency**	**Purity**
EN	**780**	0.9439 ± 0.002	0.8684 ± 0.011	**0.9551 ± 0.028**	**2580**	0.9547 ± 0.004	0.8398 ± 0.024	**0.9453 ± 0.041**
IN	1432	**0.9849 ± 0.006**	**0.9314 ± 0.021**	0.7319 ± 0.052	4392	**0.9932 ± 0.004**	**0.9575 ± 0.019**	0.8168 ± 0.073

We see that the **SO(3) symmetry is too restrictive** — the model is unable to compete with the baseline

SO(2) EuclidNet

	n_hidden = 8				n_hidden = 16			
Model	**# Parameter**	**AUC**	**Efficiency**	**Purity**	**# Parameters**	**AUC**	**Efficiency**	**Purity**
EN	**957**	**0.9913 ± 0.004**	**0.9459 ± 0.022**	**0.7955 ± 0.040**	**2580**	0.9932 ± 0.003	0.9530 ± 0.014	**0.8194 ± 0.033**
IN	1432	0.9849 ± 0.006	0.9314 ± 0.021	0.7319 ± 0.052	4392	0.9932 ± 0.004	**0.9575 ± 0.019**	0.8168 ± 0.073

SO(2) EN **marginally better** than the IN at small scales! It also **uses fewer parameters**!

Conclusions

In this study, we applied a Euclidean rotation-equivariant GNN to the particle tracking problem. An SO(3)-equivariant model performed poorly and was unable to compete with the baseline IN. The SO(2)-equivariant model offers a marginal improvement (AUC = 0.9913) over the current SOTA benchmark (AUC = 0.9849) at small model scales (< 1000 parameters), but the results are within one standard deviation of each other. More work is needed to concretely establish the reasons for this result. Possible future directions of work include studying the problem's equivariance and identifying non-equivariant facets, if any, as well as investigating the quality of the learnt symmetry. However, if the dataset inherently contains non-equivariant features, any symmetric model will always underperform. Models which partially relax the constraints imposed by symmetry groups might be able to learn the non-equivariant aspects of the tracking dataset.

Link to code

Contact
Daniel Murnane,
dtmurnane@lbl.gov

[1] https://arxiv.org/abs/1810.06111
[2] https://zenodo.org/record/4730167
[3] https://arxiv.org/abs/2103.16701

This work was supported by IRIS-HEP through the U.S. National Science Foundation (NSF) under Cooperative Agreement OAC-1836650 and by Department of Energy grant DE-SC0007968. This research used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility located at Lawrence Berkeley National Laboratory, operated under Contract No. DE-AC02-05CH11231 using 2022 NERSC award ERCAP-0021226.