Developing Lorentz Equivariant Graph Neural Networks for Top Quark Tagging

Jason Wong (jtwong71@berkeley.edu) Savannah Thais Daniel Murnane
Sofia Graziano

May 7, 2021

1 Background

Top quark tagging is the process by which data from particle collisions is labeled by the type of particle a jet represents. For example, there is a public dataset which contains over a million training entries of jets that are either top quarks, light quarks, or gluons. Being able to distinguish between these types of particles in a collision experiment is important to understand and assess physical models of particle theory.

Neural networks have recently been implemented to solve this top quark tagging problem. Graph Neural Networks (GNNs) have also been introduced as an extension to traditional neural networks by taking advantage of potential relationships between parts of the data; GNNs have been successfully implemented for particle track reconstruction with great accuracy and efficieny [1, 3].

As seen in other research, incorporating symmetries into the neural network as a prior constraint greatly improves the efficiency of the algorithm. For example, Convolution Neural Networks (CNNs), which are known to establish local translational invariance, have been extended to incorporate rotational symmetry [4]. In addition, Lorentz-invariant neural networks have been implemented and prove to have fewer parameters to tune due to the higher constraints of the problem, and hence results in a much more robust model [2].

GNNs nowadays have permutation invariance to ensure that the order in which nodes are enumerated does not matter. However, there is not much literature on enforcing other symmetries into GNNs, especially in the context of particle physics. By incorporating symmetries into GNN models, we reduce the number of trainable parameters by a significant amount. Moreover, with symmetries already established into a model, less training is required to reach similar results to a model that has to be trained to learn particular symmetries.

2 Proposed Project

We propose an extension of the Lorentz-equivariant neural network by creating a Lorentz-equivariant graph neural network for top quark tagging as used in the former network model. Thus, we would use the same dataset as the Lorentz paper to try and achieve greater accuracy than previous networks using fewer trainable parameters. So far, networks that have not enforced symmetry into the network design perform with higher accuracy for this dataset, but require between 10 to 100 times more trainable parameters, and hence take much longer to train. As GNNs have been shown to perform with high accuracy compared to traditional networks, a Lorentz-equivariant GNN would hopeful have greater accuracy, as well as having fewer parameters to train. After constructing this equivariant GNN, we can move onto similar problems that could benefit from having Lorentz-invariance inherently built into the network structure. This project would be conducted remotely with Dr. Savannah Thais and Dr. Daniel Murnane and in collaboration with Sofia Graziano to implement this Lorentz equivariant GNN. Dr. Thais and Dr. Murnane both have experience with GNNs, and have connections with other researchers working on equivariant networks, so we can seek additional help if needed. This research would fall under the innovative algorithms goal of the IRIS-HEP program.

3 Timeline and Deliverables

Month 1: Defining the problem and understanding how Lorentz-invariance is enforced in models

Month 2: Developing a GNN with Lorentz-invariance for top quark tagging

Month 3: Tweaking and applying Lorentz-equivariant GNN to similar data.

4 Student Background

I am a math and physics double major, and so the importance of symmetry and invariance is deeply rooted in my research interests. I have conducted theoretical astrophysics research with a professor from George Mason University in developing an x-ray pulsar model using Mathematica. In addition, I have developed custom machine learning algorithms, such as an auto-tuner for a Model Predictive Controller that I developed a couple years ago to enhance the performance of autonomous driving for a holonomic robot. Lastly, I am currently using TensorFlow to improve speedsolving methods for the rubik's cube.

References

[1] Farrell S., Anderson D., and Cerati G. 2017, ”Particle Track Reconstruction with Deep Learning,” *Deep Learning for Physical Sciences*, viewed 18 April 2021, from https://dl4physicalsciences.github.io/files/nips_dlps_2017_28.pdf

[2] Bogatskiy A., Anderson B., Offermann J.T., Roussi M., Miller D.W., and Kondor R. 2020, ”Lorentz Group Equivariant Neural Network for Particle Physics,” *37th International Conference on Machine Learning*, PMLR 108, viewed 18 April 2021, from https://arxiv.org/pdf/2006.04780.pdf

[3] Verma Y. and Jena S. 2020, ”Particle Track Reconstruction using Geometric Deep Learning,” viewed 18 April 2021, from https://arxiv.org/pdf/2012.08515.pdf

[4] Marcos D., Volpi M., and Tuia D. 2016, ”Learning rotation invariant convolutional filters for texture classification”, viewed 18 April 2021, from https://arxiv.org/pdf/1604.06720.pdf