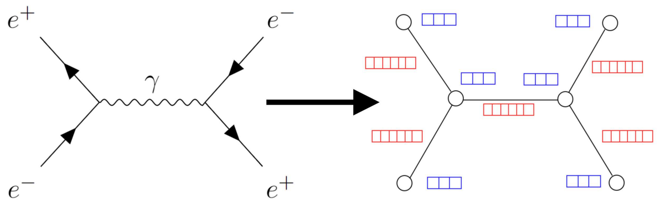
Learning Feynman Diagrams using Graph Neural Networks

Harrison Mitchell
hgwm2@cantab.ac.uk
Alexander Norcliffe
alin2@cam.ac.uk
Pietro Liò
pl219@cam.ac.uk

UNIVERSITY OF CAMBRIDGE
NEURAL INFORMATION PROCESSING SYSTEMS

Introduction:

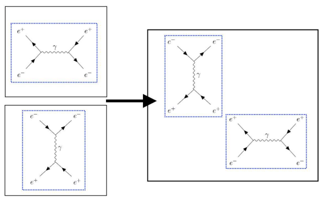
This work finds a new application of geometric deep learning, beyond experimental analysis [1], on Feynman diagrams to make accurate and fast matrix element predictions with the potential to be used in analysis of quantum field theory. This research makes use of a natural graph representation of Feynman diagrams and uses the graph attention layer to makes matrix element predictions to 1 significant figure accuracy above 90% of the time. Peak performance was achieved in making predictions to 3 significant figure accuracy over 10% of the time with less than 200 epochs of training, serving as a proof of concept on which future works can build upon for better performance. Finally, a procedure is suggested, to use the network to make advancements in quantum field theory by constructing Feynman diagrams with effective particles that represent non-perturbative calculations. The hope is that upon fully learning the Feynman rules, the network can be used to reverse engineer a graph that generate the correct matrix elements when passed through the GNN and can be interpreted as physical processes.



Schematic of the natural graph representation of Feynman diagrams. Each element of the edge embedding encodes the properties of the particle.

$$\text{edge vector} = [m, S, I_L^3, Y_L, I_R^3, Y_R, r, g, b, \bar{r}, \bar{g}, \bar{b}]$$

Each node encodes whether the particle is an initial state, virtual, or final state particle.



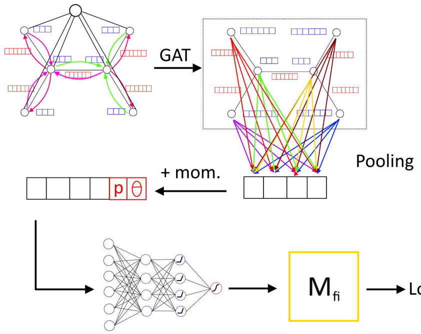
Mini-batching process for the graphs. No additional edges constructed so that the attention mechanism treats the graphs as disconnected.

Work summary:

- A program to convert Feynman diagrams into graph valued data
- Demonstrated the use of GNNs in QED.
- Initiated testing on QCD for momentum dependence decoding
- Tested the GAT layer for use beyond experimental analysis in physics.

Architecture:

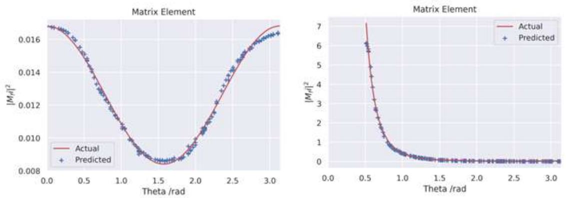
The network architecture consists of a Graph Attention Layer (GAT) [2], whose node embeddings are pooled to make a graph representation. Since QED Feynman rules are independent of momenta, the momentum is concatenated after producing the graph representation. Finally, the representation vector is passed to a fully connect network (FCN) to predict the matrix element and compute the loss.



Outline of the architecture of the network. The GAT and hidden linear layers create new node embeddings. The embeddings are pooled to create a graph representation. Finally momentum is concatenated before being passed to an FCN to get a matrix element prediction. Back propagation is conducted with the loss.

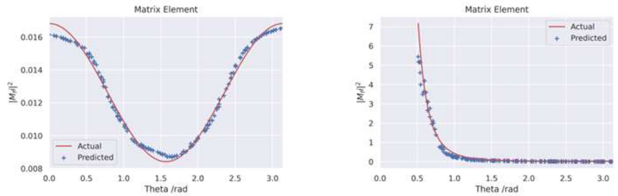
$e^-e^+ \to \mu^-\mu^+$ and $e^-e^+ \to e^-e^+$ results:

This is the first experiment on electron-positron to muon-antimuon s-channel diagrams; and the two Bhabha electron-positron to electron-positron s and t-channel diagrams, batched into one graph with ultra-relativistic matrix element. This exhibits the models performance at distinguishing between graphs. The predictions are good with a small discrepancy for the muon matrix element at large angles. This serves as a baseline for performance.



The predictions for the $e^-e^+ \to \mu^-\mu^+$ (left) and $e^-e^+ \to e^-e^+$ (right) processes. An initial benchmark to test the expected performance of the network when the dataset is increased.

Extending dataset to tree-level QED with electrons and muons:

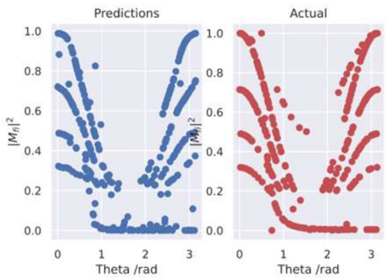


The same predictions for the $e^-e^+ \to \mu^-\mu^+$ and $e^-e^+ \to e^-e^+$ processes again but this time after having been trained on a larger dataset involving other processes like $\mu^-\mu^+ \to \mu^-\mu^+$.

Main experiment on training on a larger dataset. This dataset contains all tree-level Feynman diagrams involving muons and electrons. Performance indicates the potential to fit a wide range of matrix elements. Requires some improvements in implementation to create better fits. To 1 decimal place, accuracy of 96.75% was reached with an overall test L[1] test loss of 0.015 on a normalized target in the range [0,1].

Preliminary testing on QCD:

Initial testing of QCD to probe the FCN's ability to distinguish momentum dependence from the graph representation. Performance drops but can clearly see the generic shape emerge. Improvements in testing are needed but this experiment illustrates the ability to create momentum dependent Feynman rules.



Testing the performance of the GNN on a $u\bar{u} \to t\bar{t}$ s-channel via gluon propagator. The mass of the top quark means that the matrix elements are no longer momentum independent and so different momentum concatenations produce different targets.

References

[1]M. Feickert and B. Nachman, "A living review of machine learning for particle physics", 10.48550/ARXIV.2102.02770 (2021).

[2]P. Velickovic, G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio, "Graph attention networks", 10.48550/ARXIV.1710.10903 (2017).