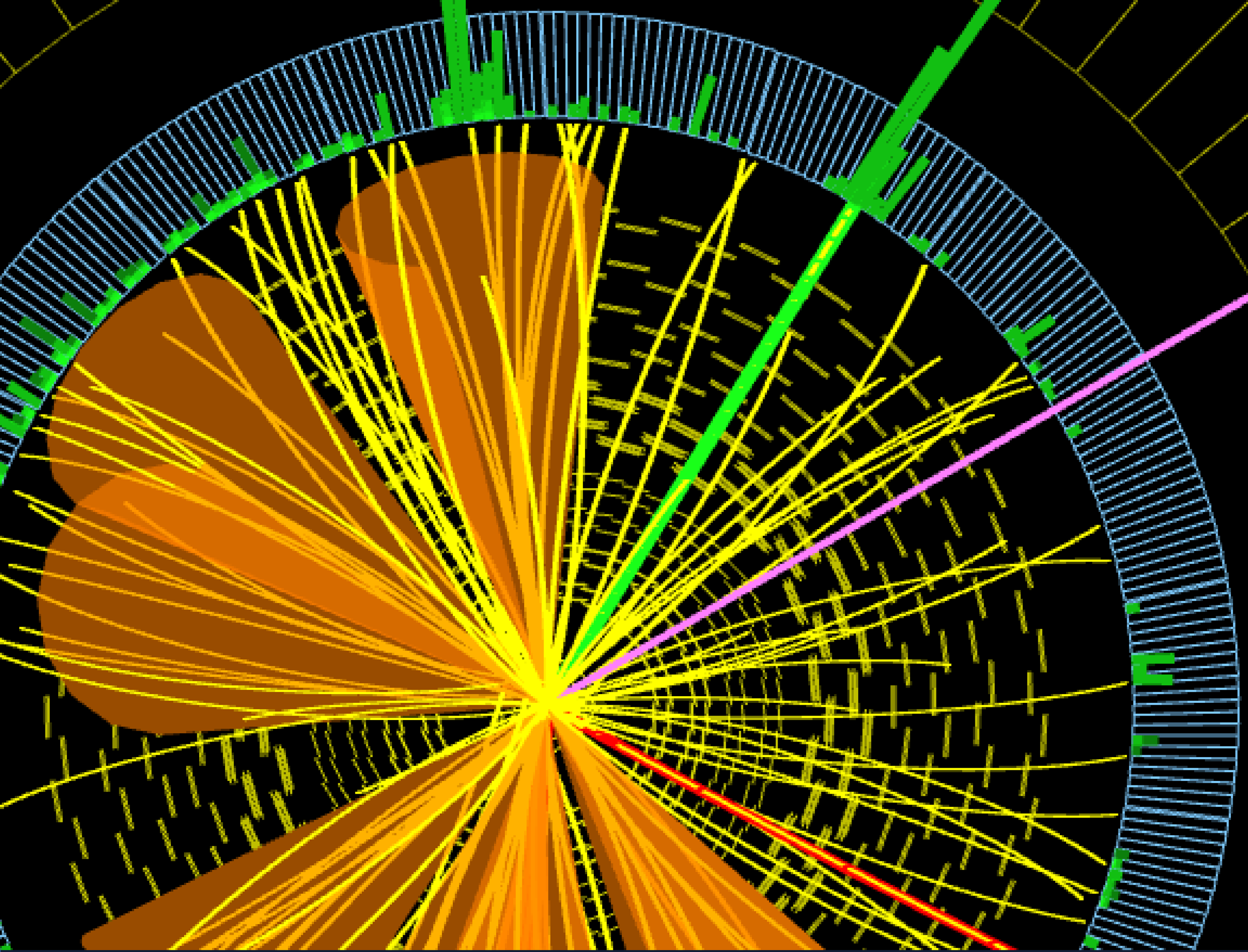
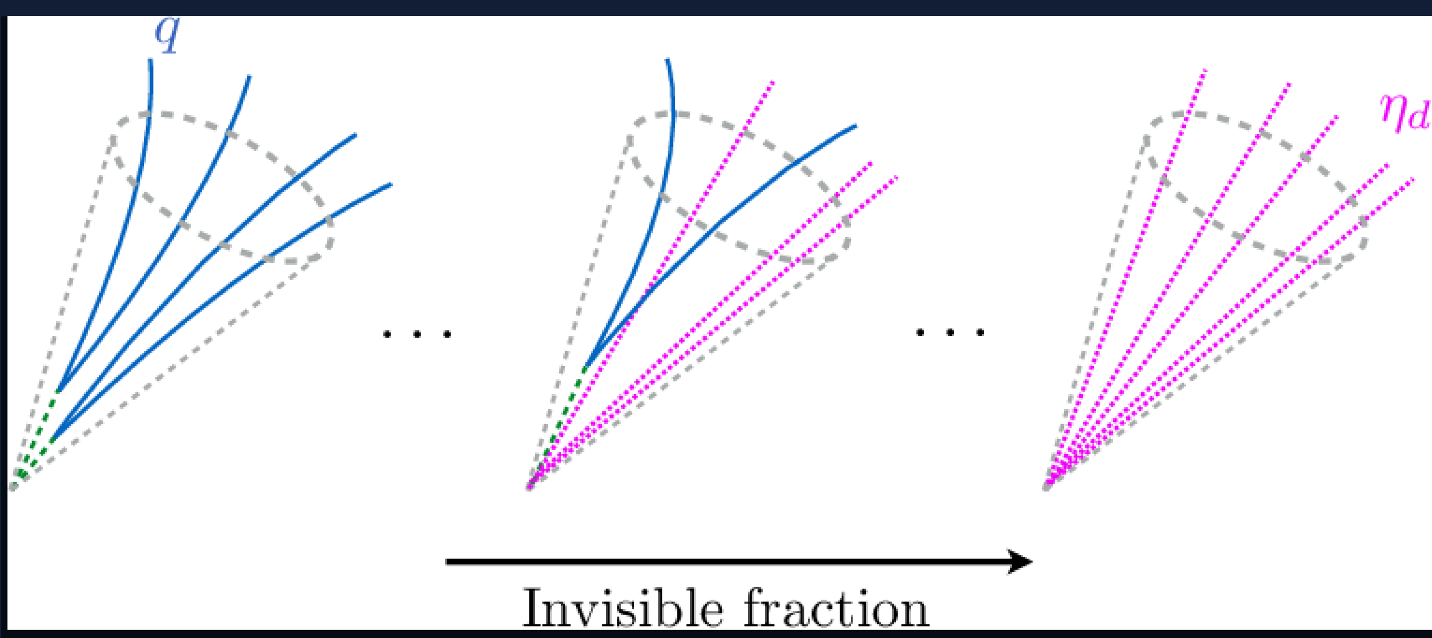


Unsupervised Machine Learning in the Search for Dark and Semi-visible Jets

Prof. Deepak Kar, Roy Gusinow



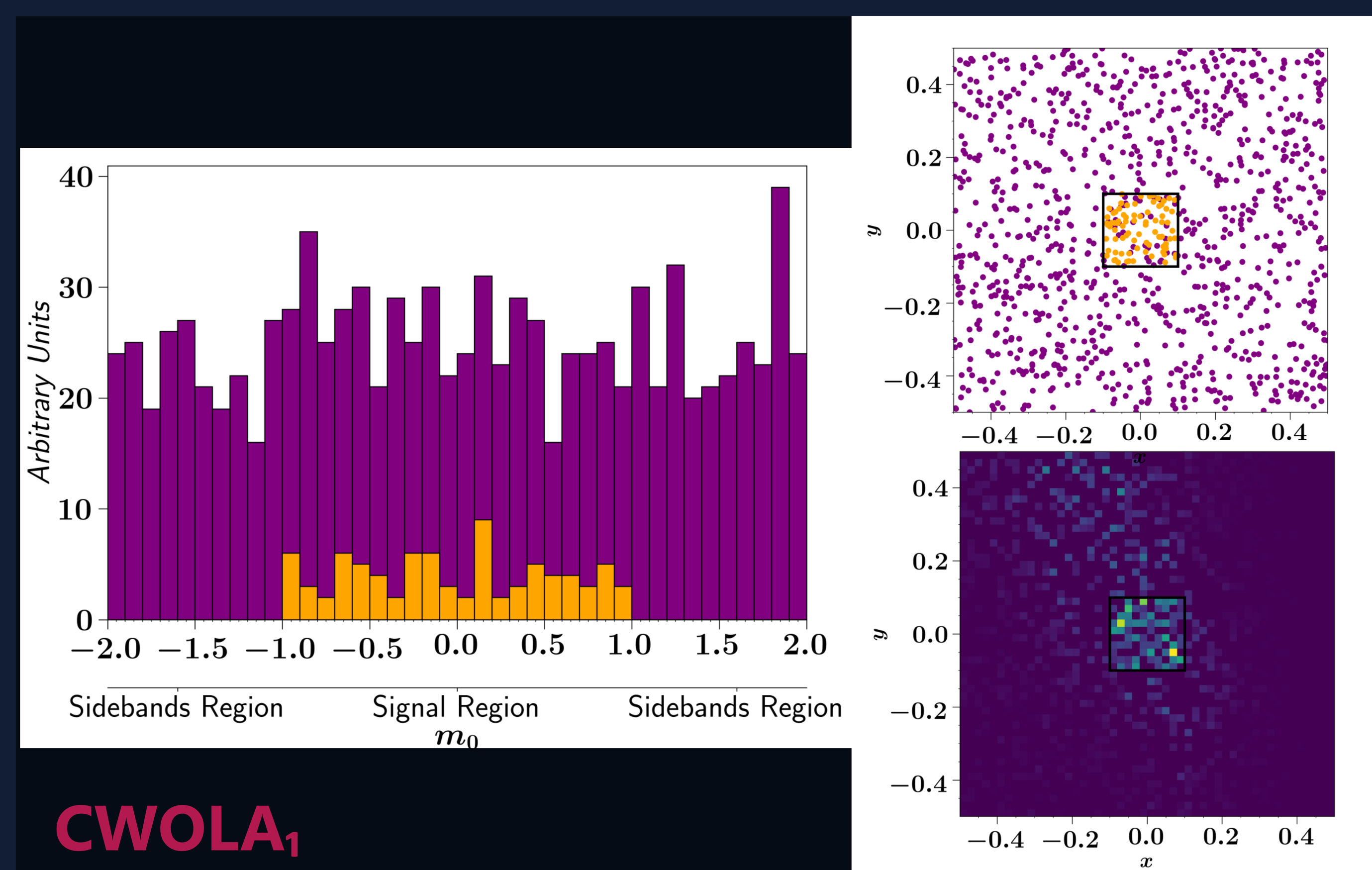
WHAT ARE DARK AND SEMI-VISIBLE JETS?

Much of DM research has focused on DM candidate particles which are heavy and have little interaction with baryonic matter. However, many theories have proposed DM candidates that do indeed interact with observable matter, particularly resulting in the formation of jets. In certain models, only a portion of dark hadrons produced in a collision will decay back to SM quarks, while the rest will pass through the detector undetected. ***Semi-visible jets (SVJ) occur when dark hadrons only partially decay to SM hadrons, while for dark jets (DJR), the dark hadrons decay fully.***

WHY UNSUPERVISED LEARNING OR ANOMALY DETECTION?

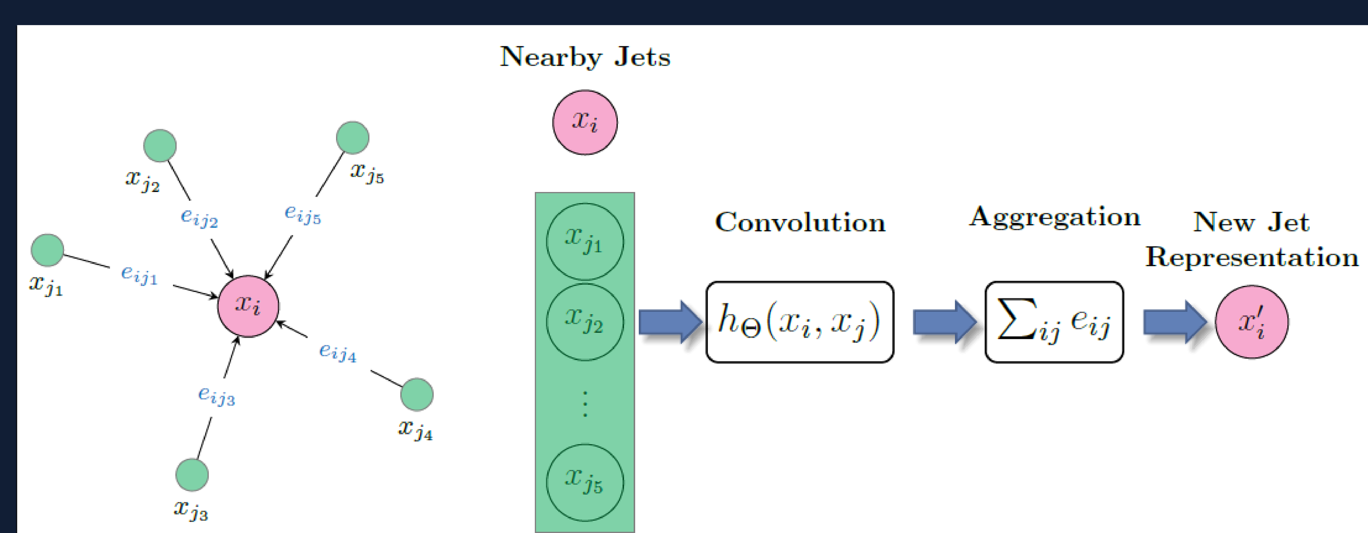
SVJ and DJR have final-state signatures visible to collider detectors, and therefore it is difficult to differentiate from jets originating from Standard Model processes. Since the final states involve unusual topologies, searches using traditional methods prove challenging to find evidence of resonant signal for dark and semi-visible jets. One proposed idea is to greatly utilise the ***jet substructure observables*** in order to assist such searches.

Anomaly detection seeks to tag few events within a dataset which are anomalous to the primary trend of the entire sample set. We expect DJR and SVJ slightly varied in structure compared to jet generated by SM particles, anomaly detection can be used effectively.



CWOLA[1]

Classification without Labels is a recent successful AD algorithm being introduced in the field. By comparing events with high concentration of signal in the Signal Region, to low concentration in the Sideband Region, we can train a neural network to approach the optimal classifier between signal and background, by training with signal and background rich sample sets.

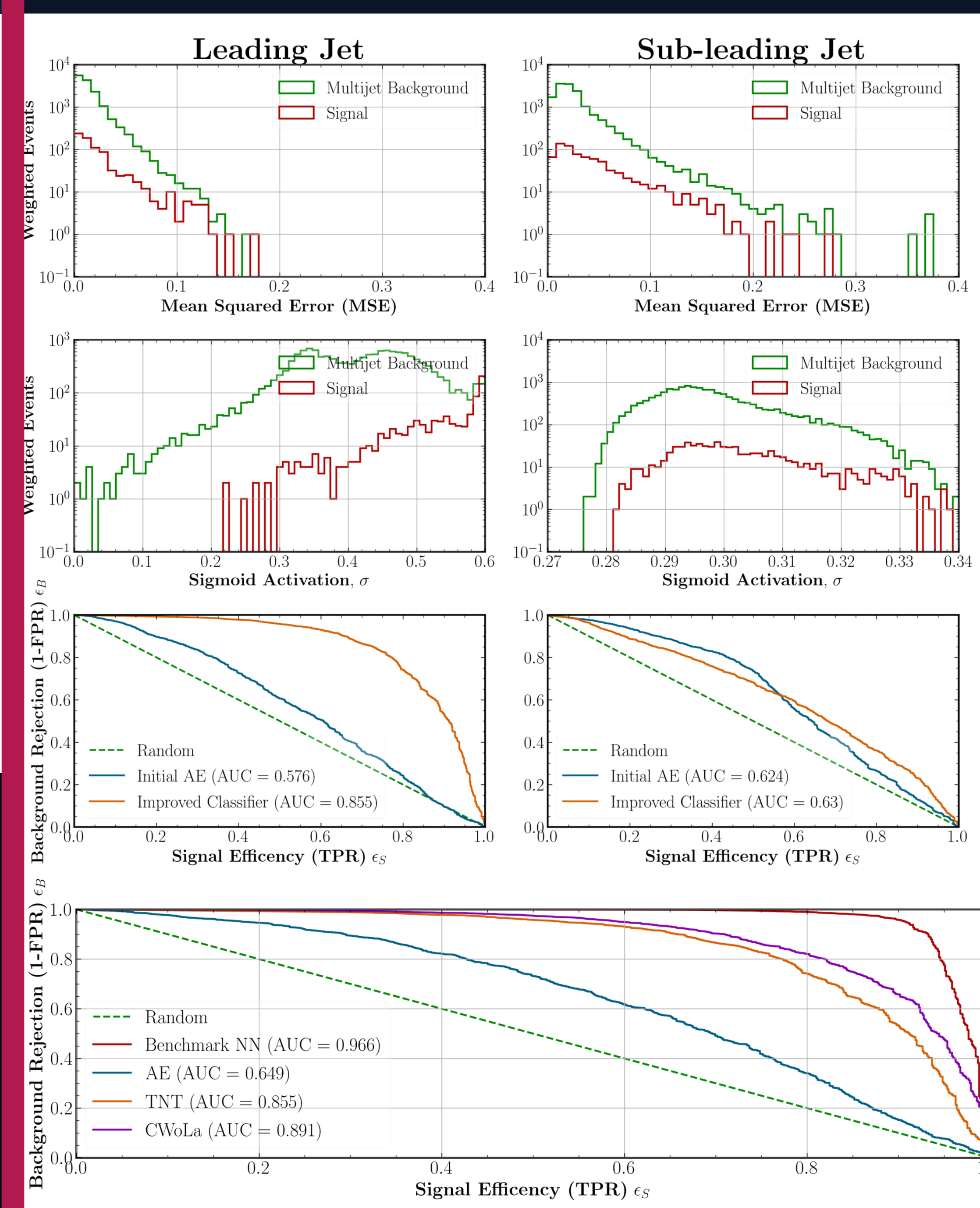


Dynamic Graph Convolutional Neural Network[2,3]

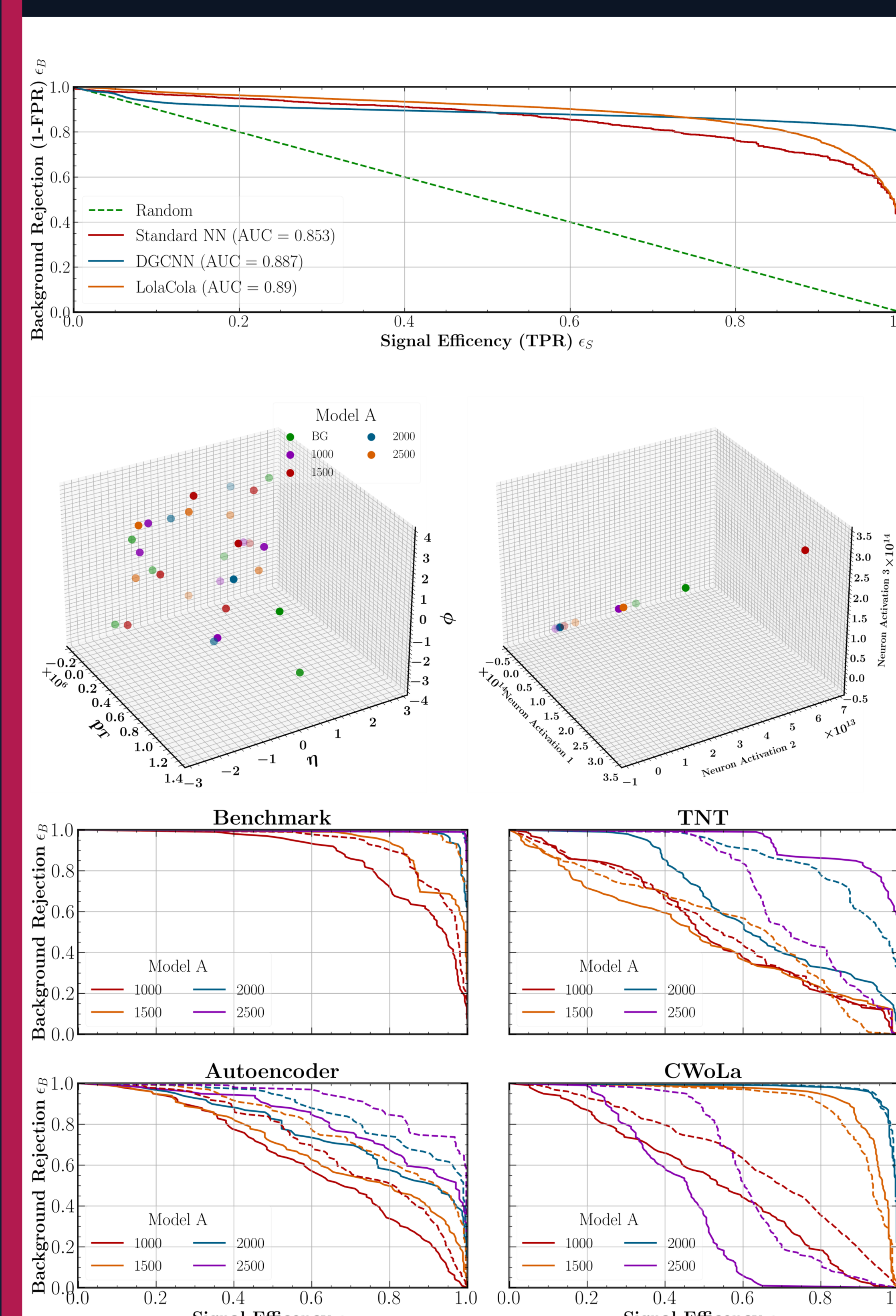
Instead of a jet image, we may consider constructing a ***single event as an undirected graph, where each node is a jet with corresponding jet substructure*** and kinematic variables as the attributes of the node.

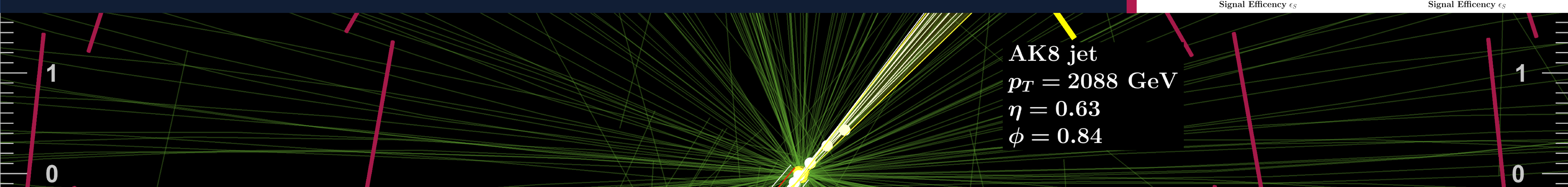
The considered jet and its neighbours are used to find the edges, through a multi-layer perceptron, ***h***. Θ is the set of trainable weights that are set to be optimized to produces the most relevant edge features. These edges are then aggregated in order to produce a new feature representation of the considered jet.

SEMI-VISIBLE JET RESULTS



DARK JET RESULTS





[1] Metodiev, Eric M., Benjamin Nachman, and Jesse Thaler. "Classification without labels: Learning from mixed samples in high energy physics." Journal of High Energy Physics 2017.10 (2017): 1-18.
[2] Wang, Yue, et al. "Dynamic graph cnn for learning on point clouds." Acm Transactions On Graphics (tog) 38.5 (2019): 1-12.
[3] Qu, Huilin, and Loukas Gouskos. "Jet tagging via particle clouds." Physical Review D 101.5 (2020): 056019.