Tau Identification & Classification with GNN – How to Improve

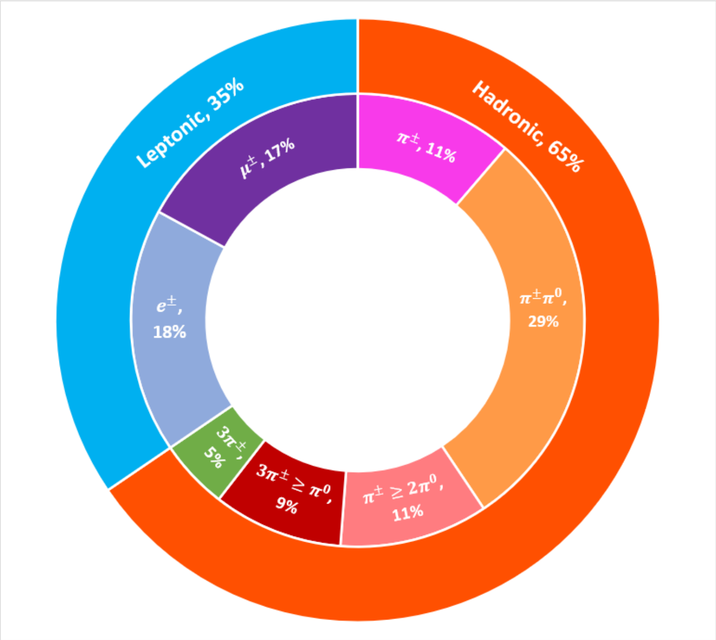
Robert McNulty

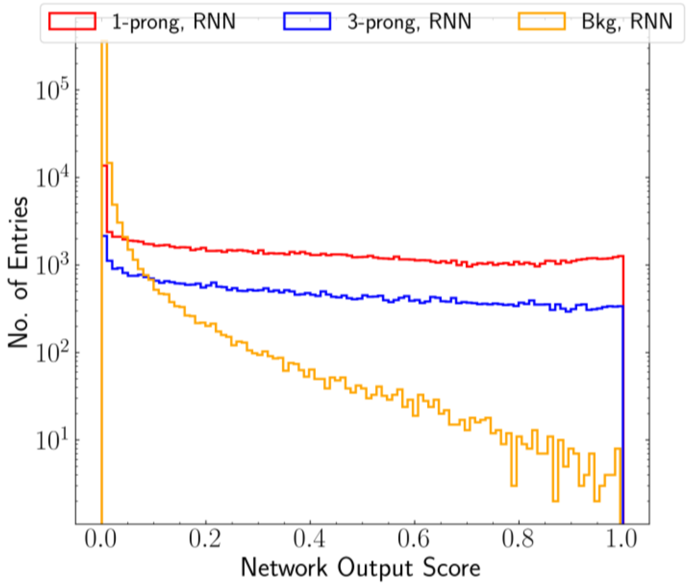
Supervisors:
Dr. Nikolaos Rompotis
Prof. Monica D'Onofrio
With some contributions from: *Dr. Joseph Carmignani*

τ-Leptons and The Motivations for τ_{had} ID and Decay Mode Classification

- ID important for several areas of research, such as:
  - $H \rightarrow \tau\tau$ production [2]
  - Di-Higgs searches with $b\bar{b}\tau^+\tau^-$

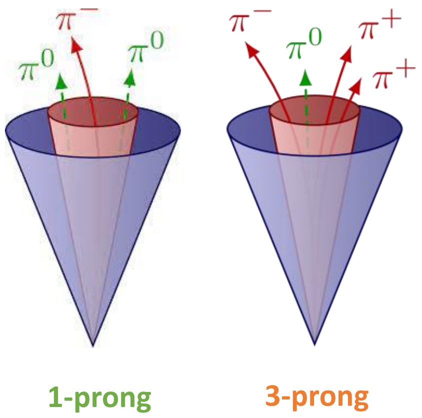
- Goal is to obtain greater background rejection and signal efficiency for ID and Decay Mode Classification





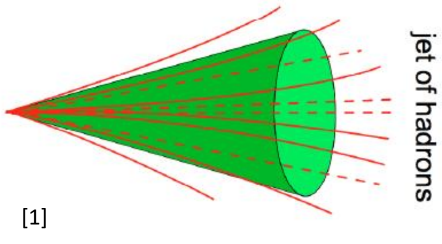
τ_{had} Decay Cones:

- Highly collimated – narrow cone
- Small cross-section
- Low multiplicity



Dijet Production & Cone

- Main background source of fake τ_{had} are jets from QCD
- Shower shape can mimic/drown-out the shower shape of π's from τ_{had}
- Fragment into multiple hadrons (high multiplicity)
- High production cross-section for dijets
- Wider cone area



[1]

[1] Joern Mahlstedt and the ATLAS collaboration 2014 *J. Phys.: Conf. Ser.* **513** 012021, **DOI** 10.1088/1742-6596/513/1/012021

Current Workflow used by ATLAS [3] & Input Variables Used

1) Track Reconstruction - Recurrent Neural Network (RNN):
- Takes jet track information as input
- Identifies the type of jet into four categories, including τ-tracks

2) τ Identification – Recurrent Neural Network (RNN):
- Takes track ID and uses information on shower shape and other properties to discriminate τ_{had} candidates against dijets

3) Decay Mode Classifier (DMC) – DeepSet Neural Network (DSNN):
- Used to determine the type of decay, via the number of π^0 associated with the decay

Track inputs

$p_T^{\tau_{\text{had-vis}}}$
p_T^{track}
$\Delta\eta^{\text{track}}$
$\Delta\phi^{\text{track}}$
N_{IBL}
N_{pixel}
N_{SCT}
$z_0^{\text{TJVA}}\sin\theta$
d_0^{TJVA}
$S(z_0^{\text{TJVA}}\sin\theta)$
$S(d_0^{\text{TJVA}})$

Cluster inputs

$p_T^{\tau_{\text{had-vis}}}$
E_T^{cluster}
$\Delta\eta^{\text{cluster}}$
$\Delta\phi^{\text{cluster}}$
$\langle\lambda_{\text{centre}}\rangle$
$\langle\lambda^2\rangle$
$\langle r^2\rangle$

High-level inputs

$p_T^{\tau_{\text{had-vis}}}$
$f_{\text{cent}}^{\Delta R<0.1}$
$f_{\text{lead-track}}^{-1}$
$\Delta R_{\text{max}}(\text{track}, \tau_{\text{had}})$
S_T^{flight}
$f_{\text{track}}^{\text{iso}}$
$f_{\text{track}}^{\text{EM}}$
$p_T^{\text{EM+track}}/p_T$
$m^{\text{EM+track}}$
m^{track}

* $p_T(\tau_{\text{had}})$
$p_T(\text{object})$
$\Delta\phi(\text{object}, \tau_{\text{had}})$
$\Delta\eta(\text{object}, \tau_{\text{had}})$
$\Delta\phi(\text{object}, \text{trackECal})$
$\Delta\eta(\text{object}, \text{trackECal})$

π^0 Candidates only:

$\langle\eta^1\rangle$
$\log(\langle r^2\rangle)$
$\Delta\theta$
$\log(\lambda_{\text{centre}})$
$\langle\lambda^2\rangle$
$\log(\langle\rho^2\rangle)$
f_{core}
$f_{\text{core}}^{\text{EM1}}$
$N_{\text{pos,EM1}}$
$N_{\text{pos,EM2}}$
E_{EM1}
E_{EM2}
$\langle\eta^1_{\text{EM1}}\rangle$ w.r.t. cluster
$\langle\eta^1_{\text{EM2}}\rangle$ w.r.t. cluster
$\log(\langle\eta^2_{\text{EM1}}\rangle)$ w.r.t. cluster
$\log(\langle\eta^2_{\text{EM2}}\rangle)$ w.r.t. cluster

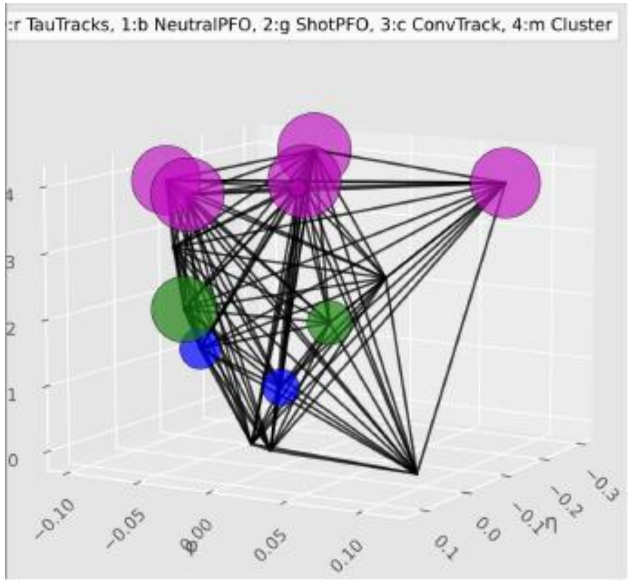
* Object refers to a candidate for either a τ_{had} Track, π^0, photon shot or conversion track

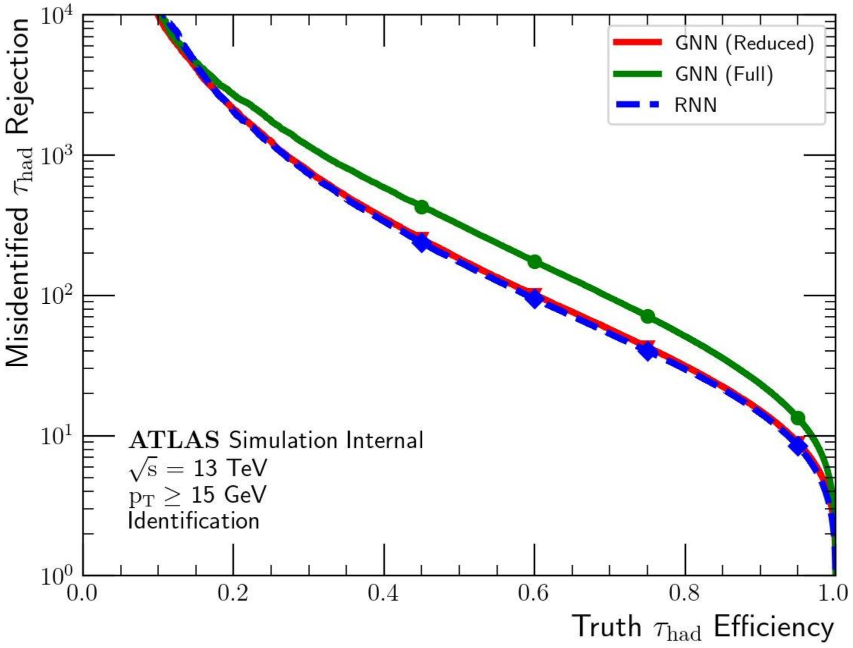
[3] ATL-PHYS-PUB-2022-044

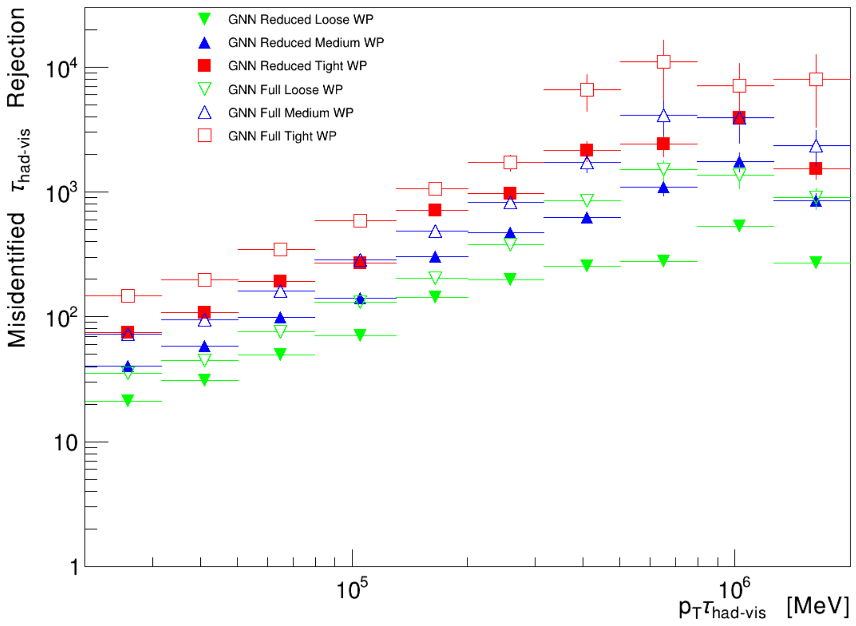
Unified Approach: Graph Neural Network (GNN)

(With Dr. Joe Carmignani, University of Liverpool)

Example graph of 3-prong event







Objects are represented as nodes:
Red for TauTracks (Layer 0)
Blue for NeutralPFOs (Layer 1)
Green for ShotPFOs (Layer 2)
Cyan for ConvTracks (Layer 3)
Magenta for TauClusters (Layer 4)

- Reduced GNN performed slightly better than RNN
- Increasing number of different types of objects increases performance in misidentified τ_{had} rejection

- Reduced GNN data shows lower p_T rejection across the entire p_T distribution

Next Steps

1. Use edge attributes
2. Training for 1-prong data
3. Expand to include classifications for decay mode classes
4. Additional preprocessing on the data
5. Additional checks on size of graphs for impact on training

Thank You for Listening