

GNN Track Matching Update

Daniel Xing

DUNE-ND/2x2 ML

July 22nd, 2026

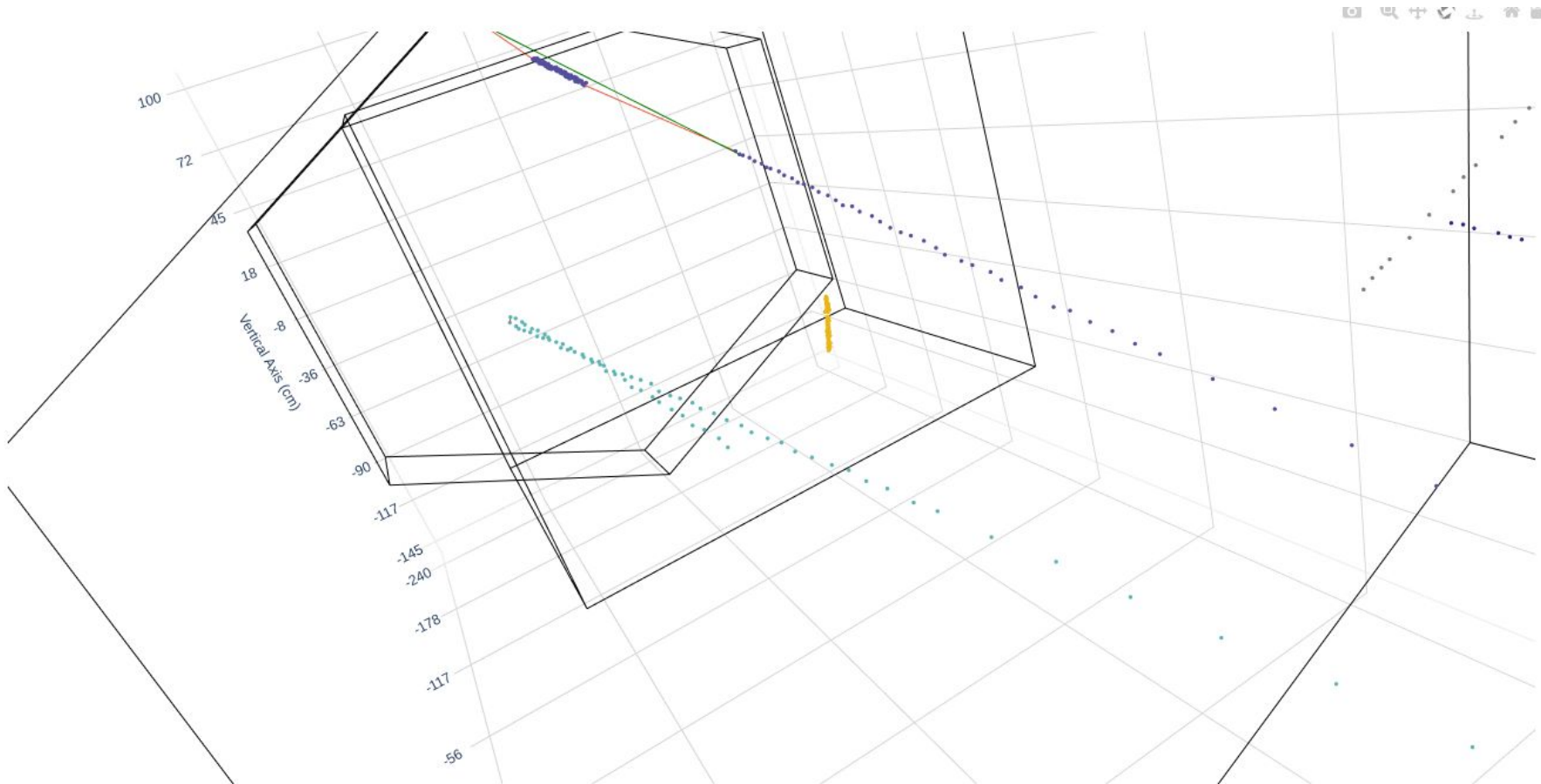
Introduction

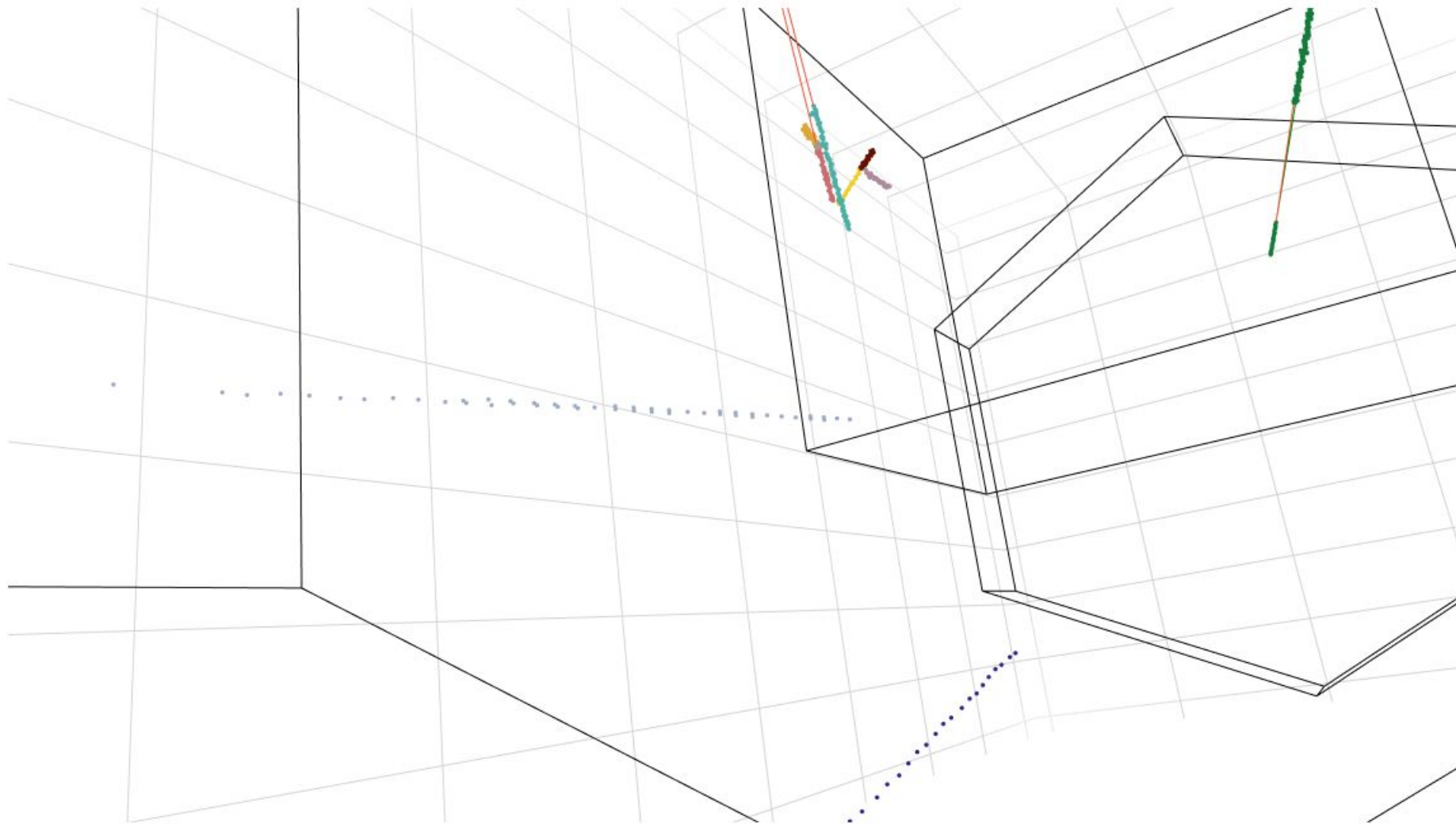
- Have been training GNNs to match tracks between Mx2 and 2x2
- Reasonable performance, but some edge cases like true matches within MINERvA or true cross-detector matches being labelled incorrectly
- Purely geometric so far (no energy/timing info)

True Label	Not Connected	Connected
	Not Connected	Connected
Not Connected	253507 (99.6%)	929 (7.5%)
Connected	1060 (0.4%)	11480 (92.5%)

“Ghost Tracks”

Parallel reco tracks originating from the same hits/true track



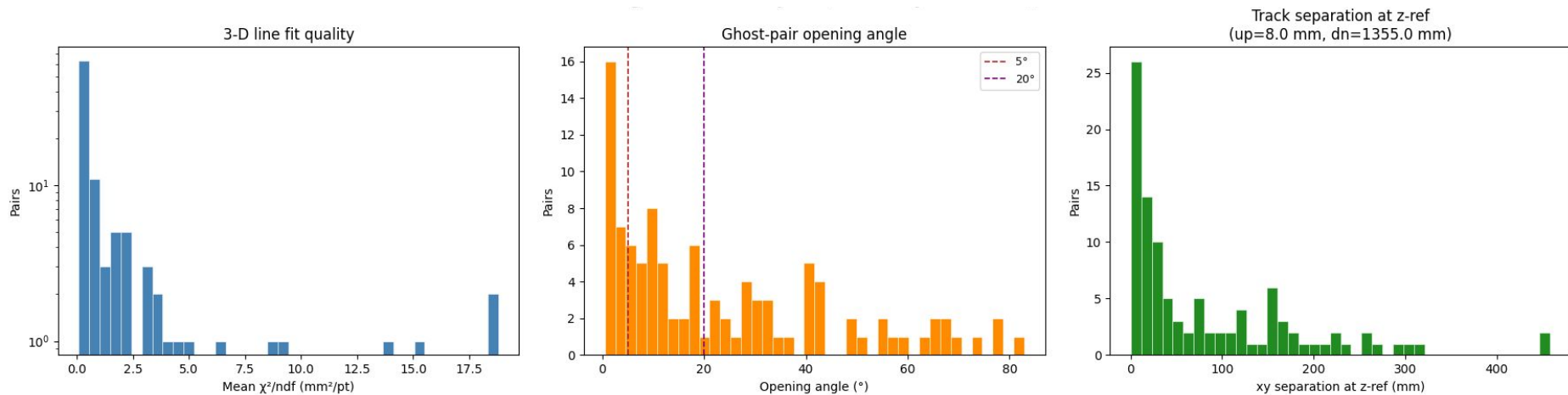


Further Investigation

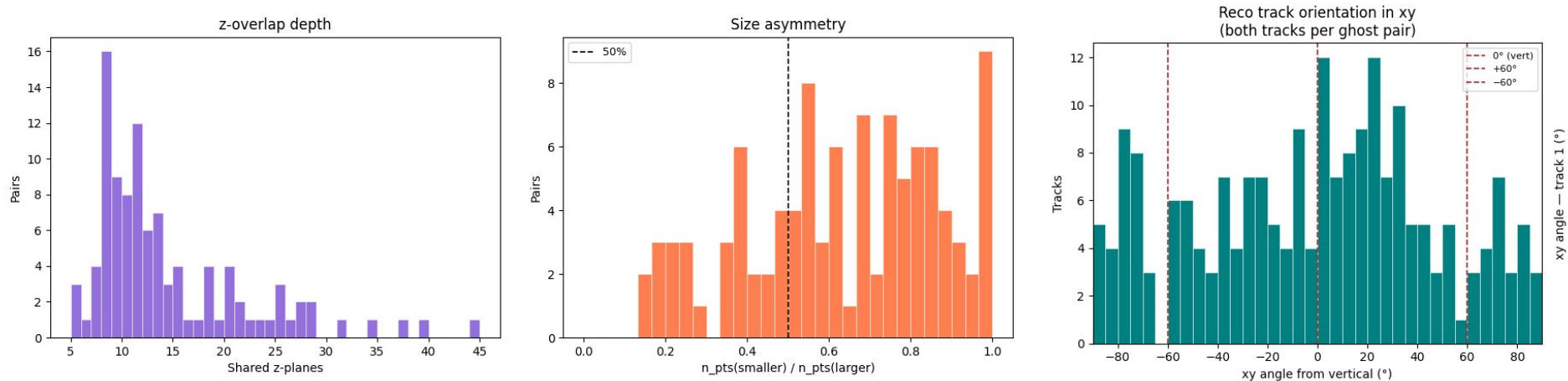
- In the validation set of 40 files (6757 events) where the GNN made 6803 connections, there were about 100 of these ghost tracks ($\sim 1.5\%$)
- Found by selecting reco tracks matched to the same true particle had overlapping z points in MINERvA (minimum 5)
- Looking at hits / energy depositions, these ghost tracks often have activity near the beginning of the track, which might confuse the MINERvA reco to think there are 2 tracks

Ghost Track Distributions

Fit lines to the ghost tracks, calculated opening angle and separation distance at the most downstream plane of the detector



Ghost Track Distributions

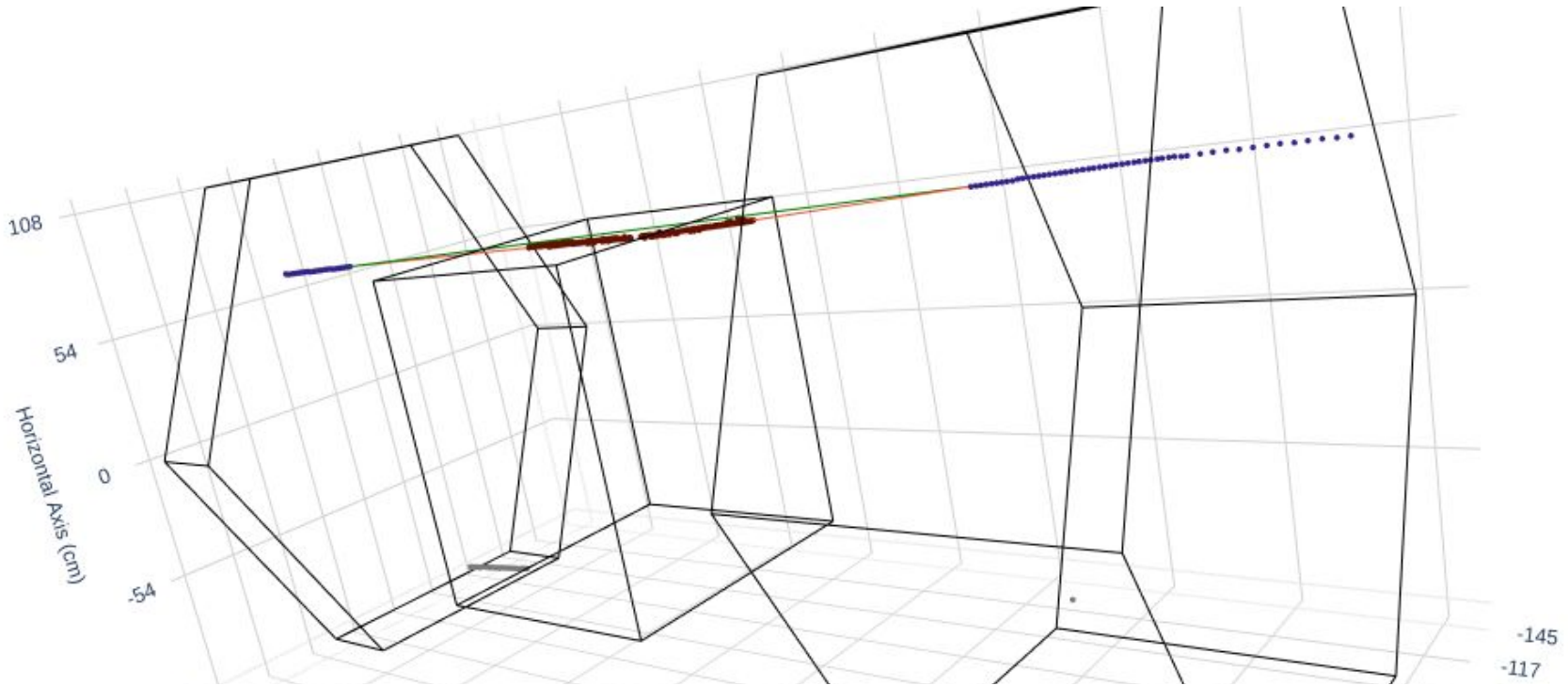


Also calculated the number of planes the two ghost tracks overlapped for, the size asymmetry, and angle of each track projected in the xy plane

Maybe some deficit at ± 60 deg, indicating it has something to do with the strip orientation?

Parent Matching

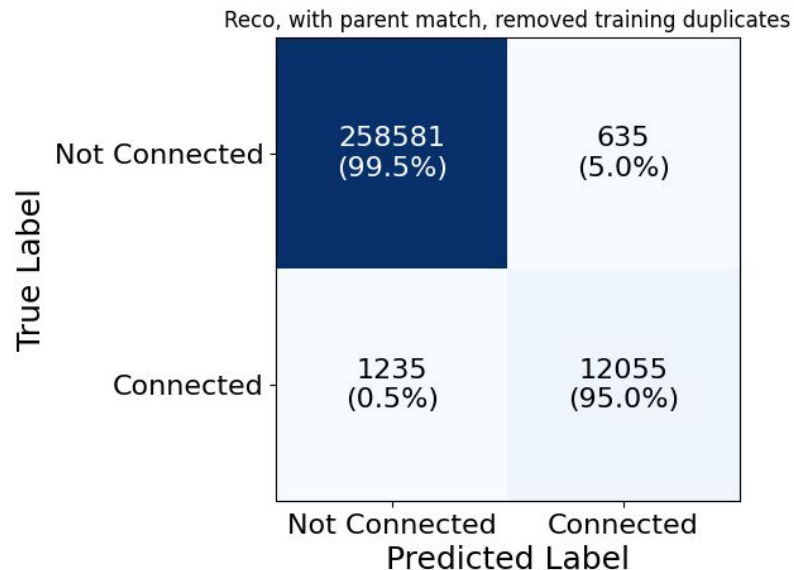
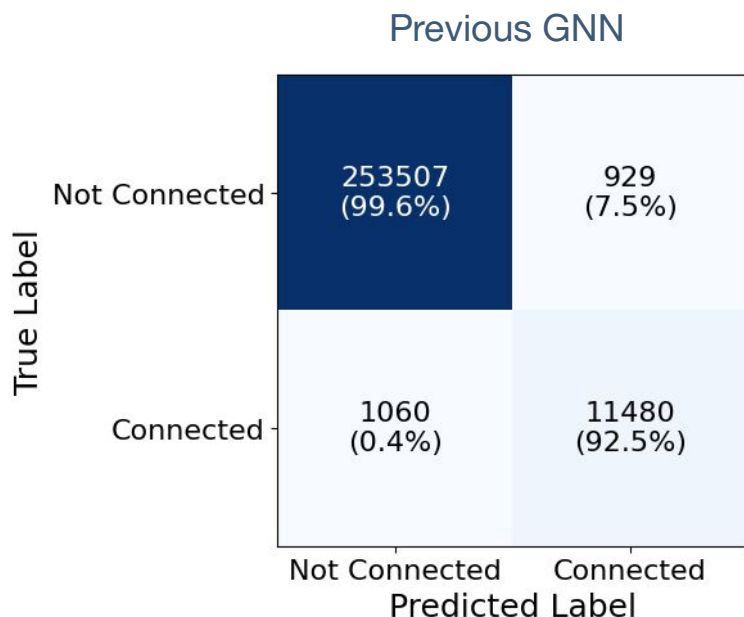
The maroon 2x2 track and purple MINERvA tracks were not matched in truth - MINERvA tracks were labelled as an electron



However, the parent particle of the MINERvA tracks was the same as the 2x2 track

Newest Training

- Removed any ghost tracks from the training sample (501 tracks in ~35k events) but kept in validation
- Implemented parent matching - two tracks are matched if either matches with the other's parent particle (1005 events)
- While purity is better, efficiency is worse and so purity*efficiency and F1 score are about the same (P^*E : 0.859 old vs 0.861 new)



Next Steps: Adding Energy and Timing

- Plan to use reco_particle depositions from 2x2 SPINE files, and clus_id_energy from MINERvA .dst files
 - Different materials, spacing, thresholds, etc. Can I just plug energies in, or should I rescale by detector somehow?
- Timing: MINERvA dst files have many timing variables, but couldn't find any in 2x2 files other than flashes/light information. Ideas?