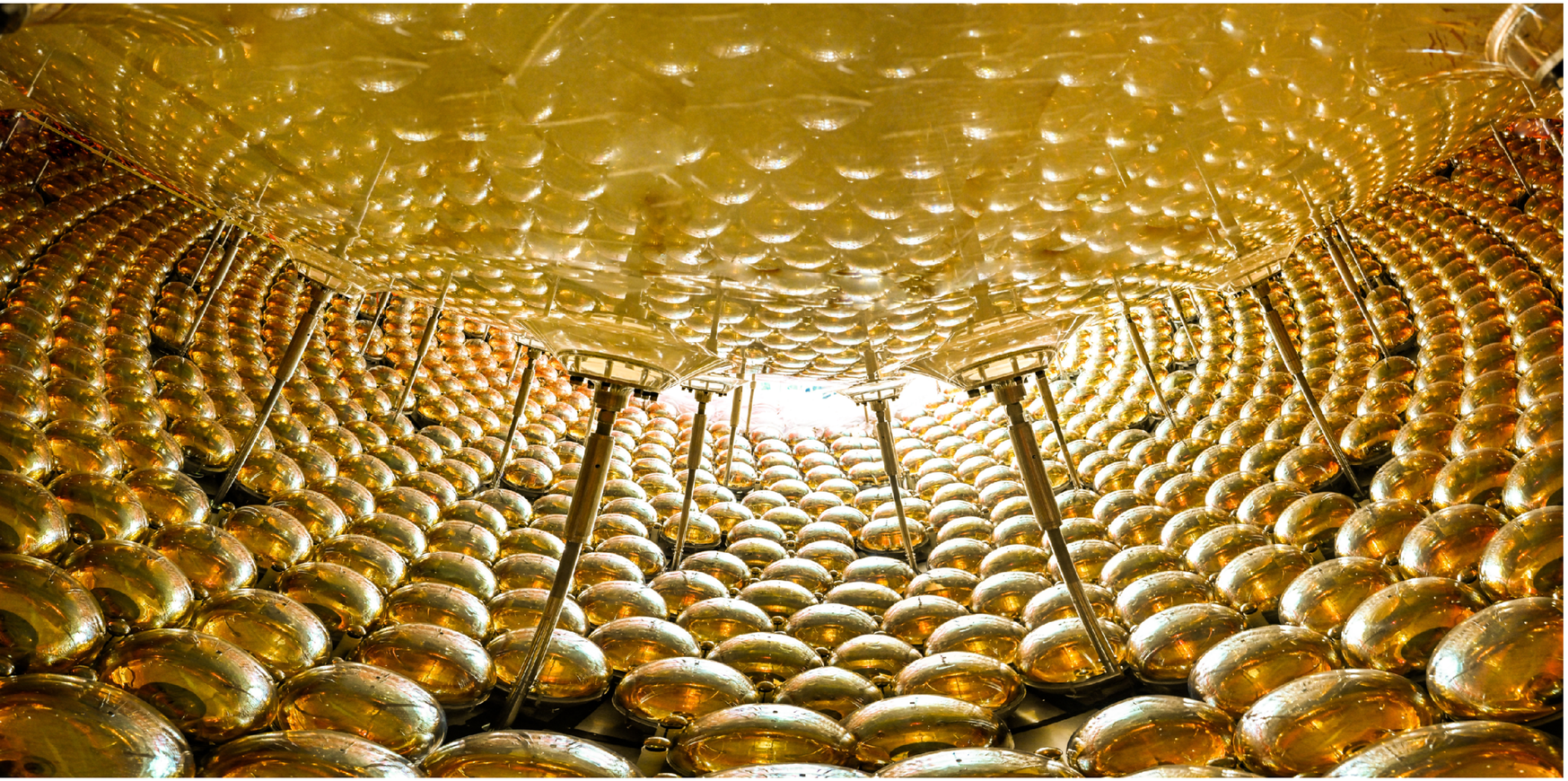
Atmospheric Neutrino PID in JUNO

First-year PhD progress report

Zekun Yang | ***Supervisors: Costas Andreopoulos · Xianguo Lu (Warwick)***



Outline

- Background
- Methodology
- 3-class PID
- Secondary triggers
- Summary

28 Oct 2026

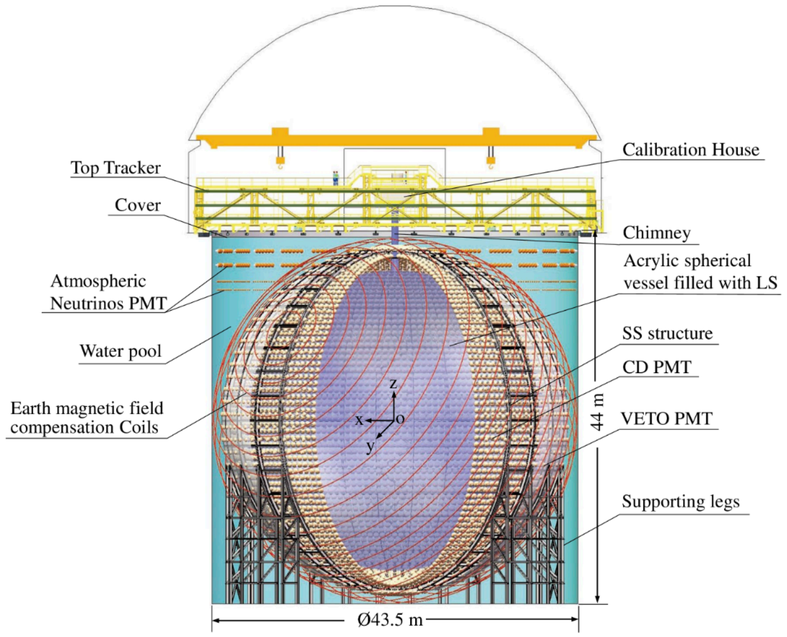
Background

Atmospheric Neutrinos at JUNO

JUNO

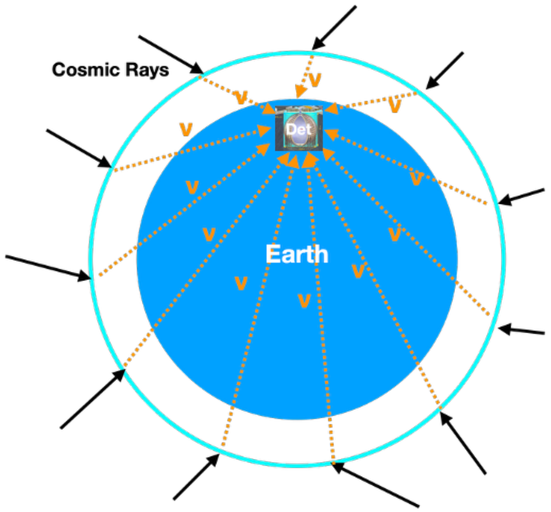
A 20 kt liquid-scintillator detector.

Designed to determine neutrino mass ordering (NMO) with MeV reactor antineutrinos.



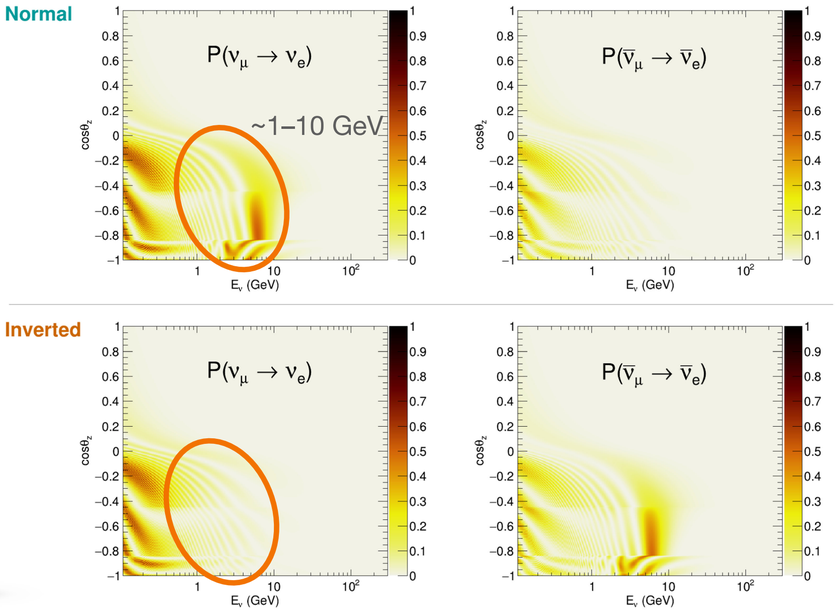
Atmospheric Neutrinos

JUNO also detects GeV atmospheric neutrinos, whose oscillations are modified by matter effects as they travel through the Earth.



Sensitivity to mass ordering

Matter effects produce different oscillation probabilities for normal and inverted mass ordering, providing information complementary to reactor neutrinos.



Background

Challenge

Particle identification (PID) distinguishes atmospheric-neutrino interaction types. However, it is challenging in JUNO:

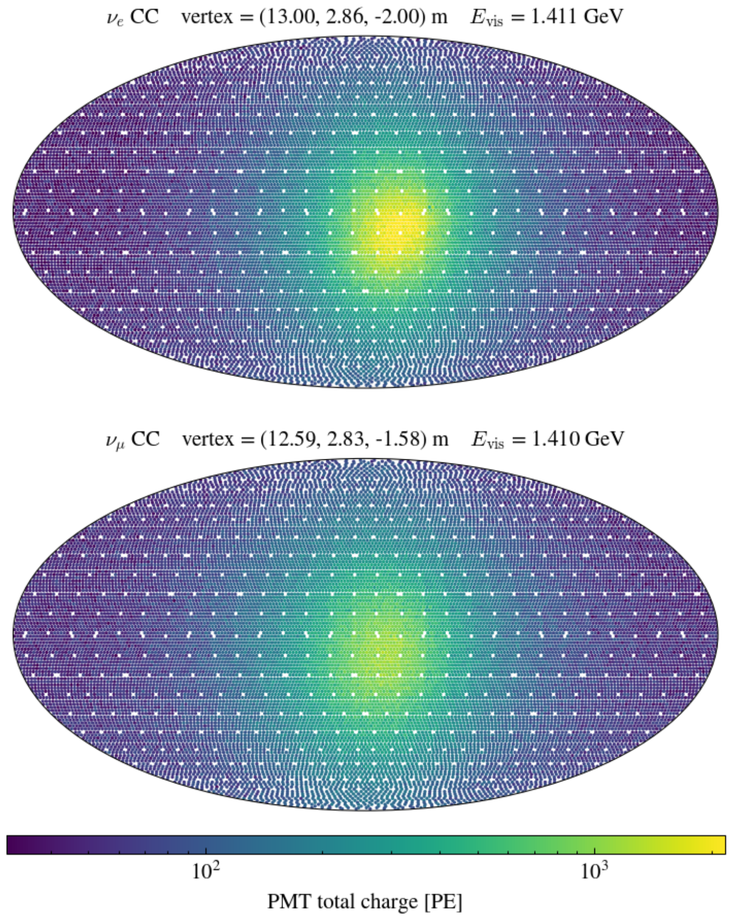
1. Electron–muon separation

Nearly isotropic scintillation light providing no direct track information.

2. Neutrino–antineutrino separation

JUNO cannot directly measure the lepton charge sign.

PID must exploit both spatial and temporal differences in the detected light.

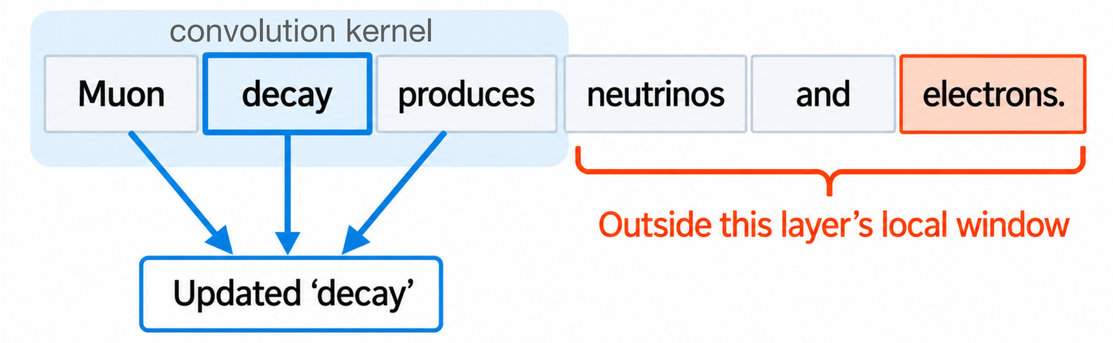


Total charge patterns (similar vertex and visible energy)
Top: ν_e-CC, Bottom: ν_μ-CC

Methodology

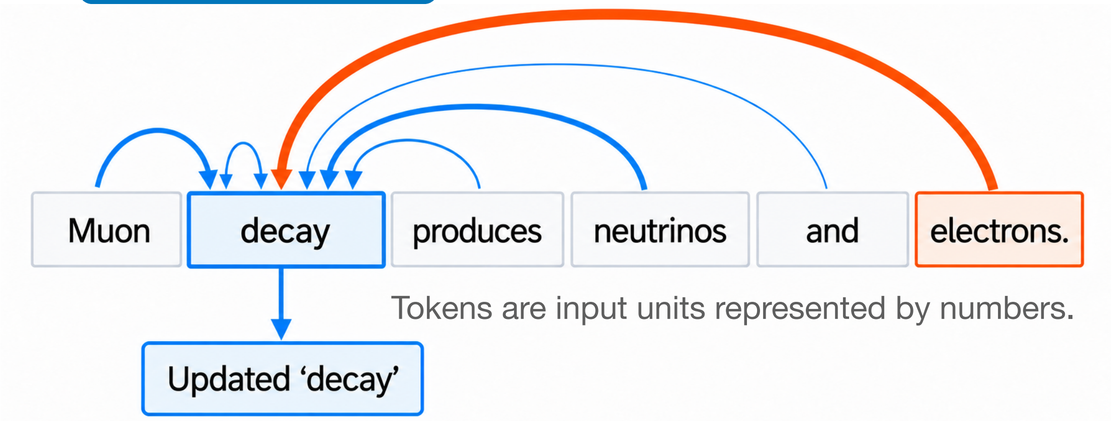
QUESTION: What charged particle is produced after muon decay?

Convolution



- One local layer cannot directly link “electrons" to “decay".
- Distant information needs more layers to propagate.

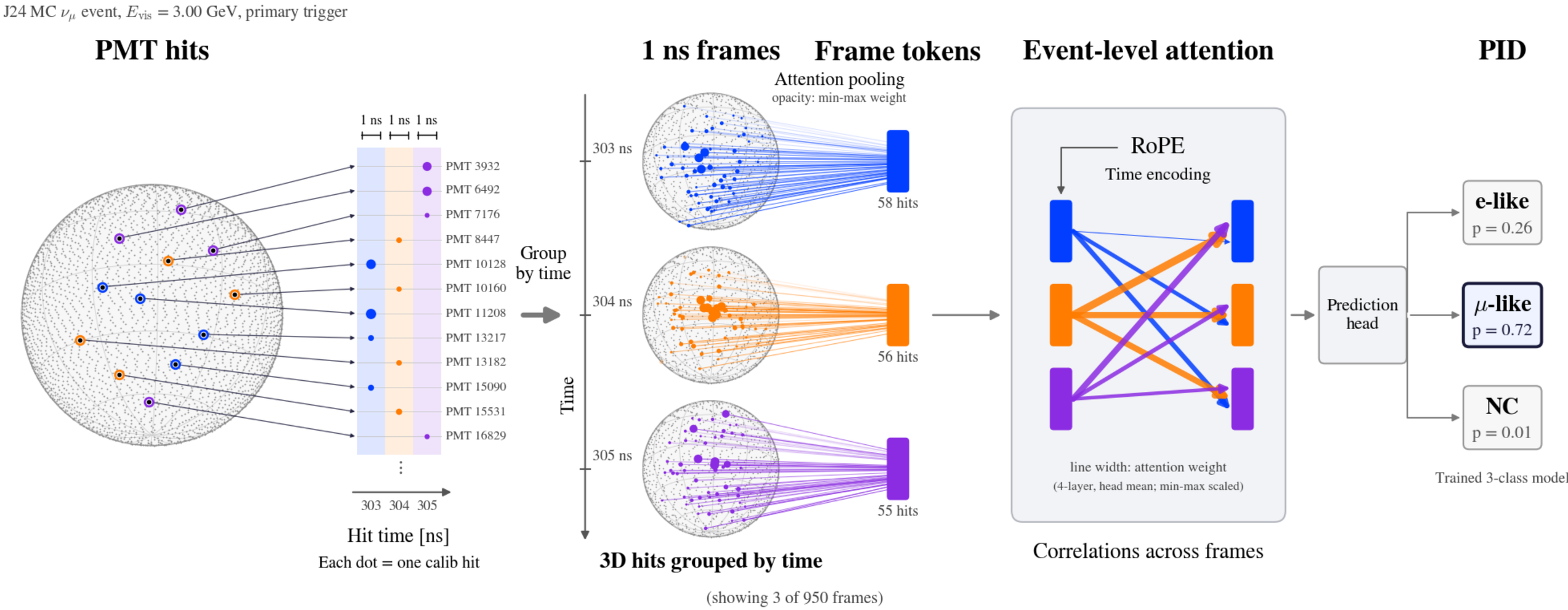
Attention



- Attention directly connects distant tokens.
- Connection weights adapt to the input content.
- Transformer is an architecture built around Attention.

Transformers Around Us

Model Architecture

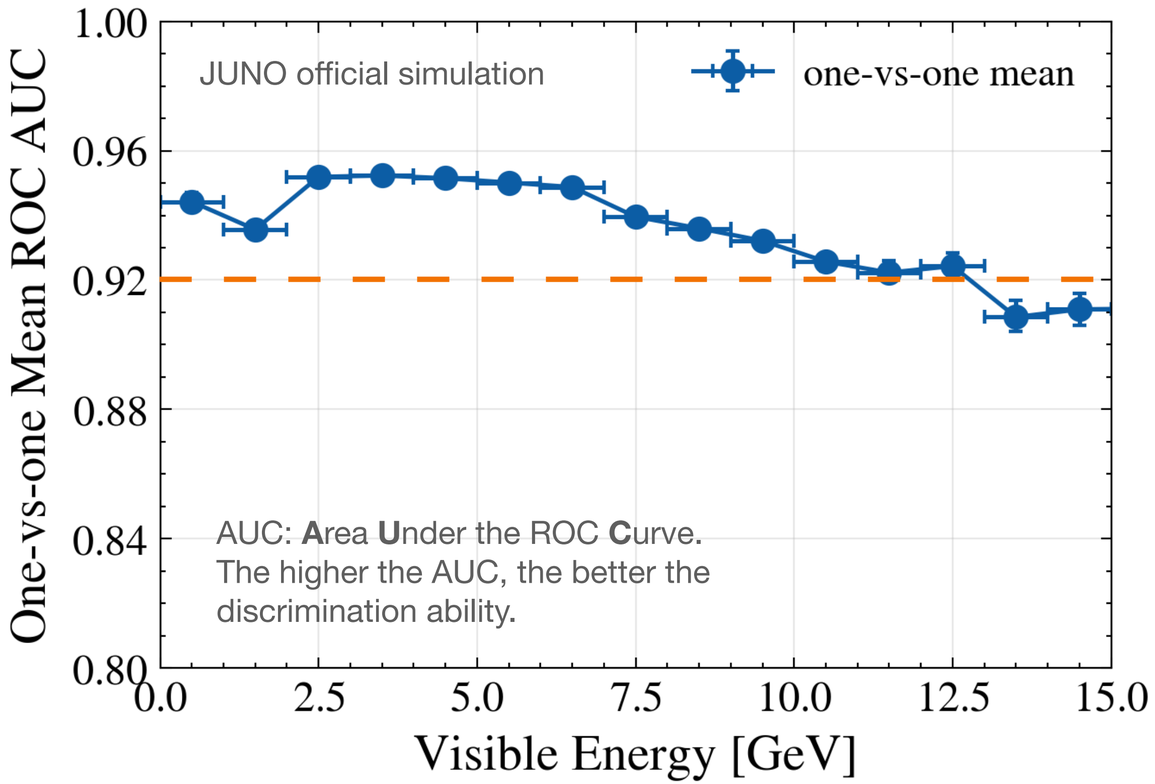


Grouping hits into 1 ns frame greatly reduces inputs to attention, cutting computation and memory use while preserving the event's temporal structure.

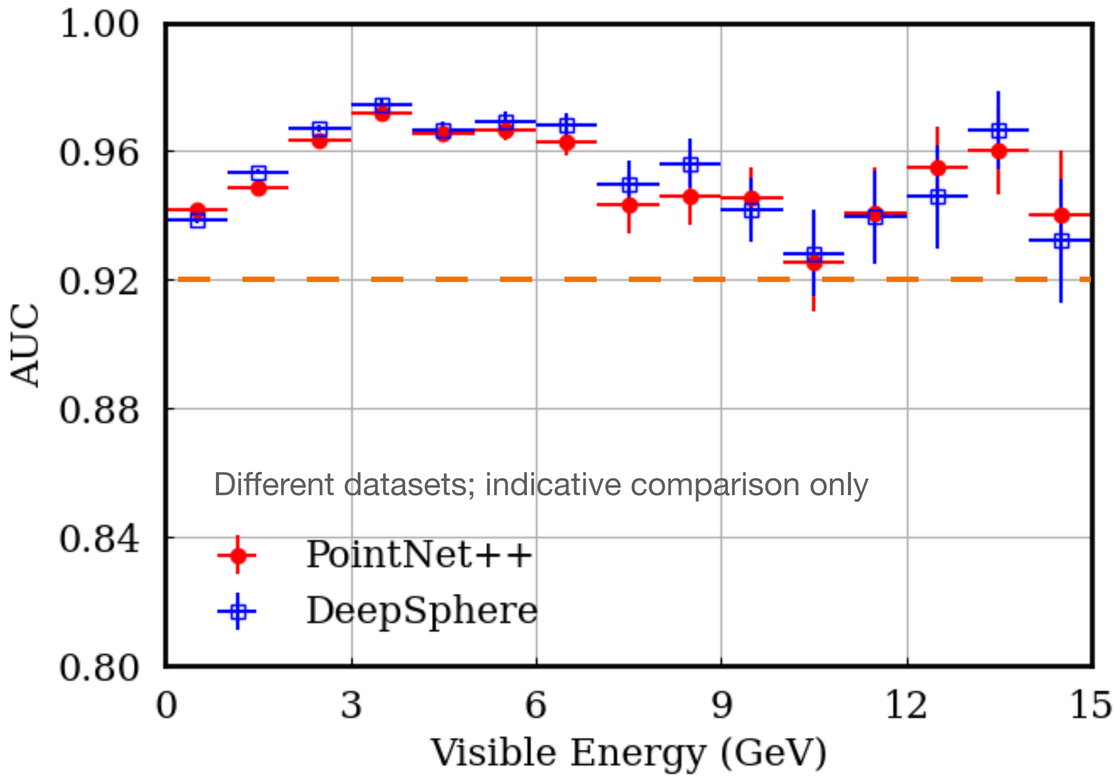
3-class PID

Mean OVO AUC >0.92 in most bins

This work



Other ML methods in JUNO

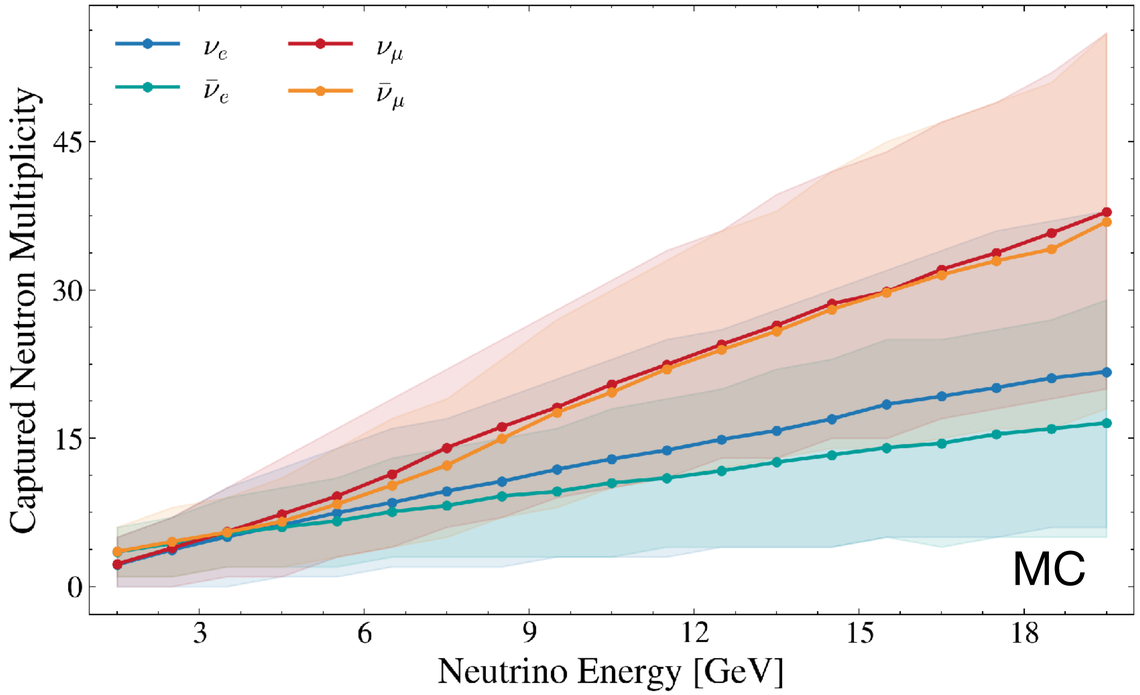


My model achieves broadly comparable AUC values with other ML models in the ~1–10 GeV range.

Secondary triggers

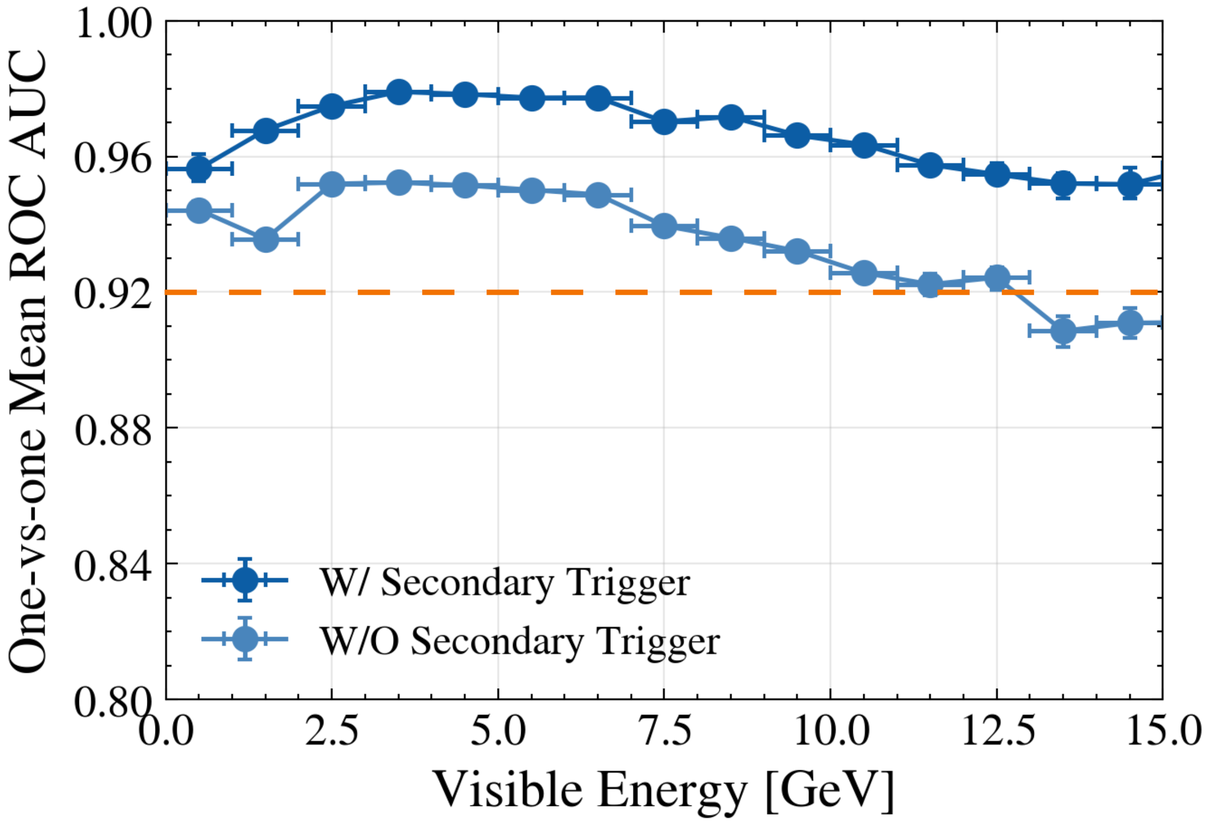
Extended the model to combine primary and secondary triggers:

- Delayed signals, including neutron captures, provide additional information to distinguish neutrino flavours.
- A pre-selection is implemented to select signals by neutron captures with ~88% efficiency.

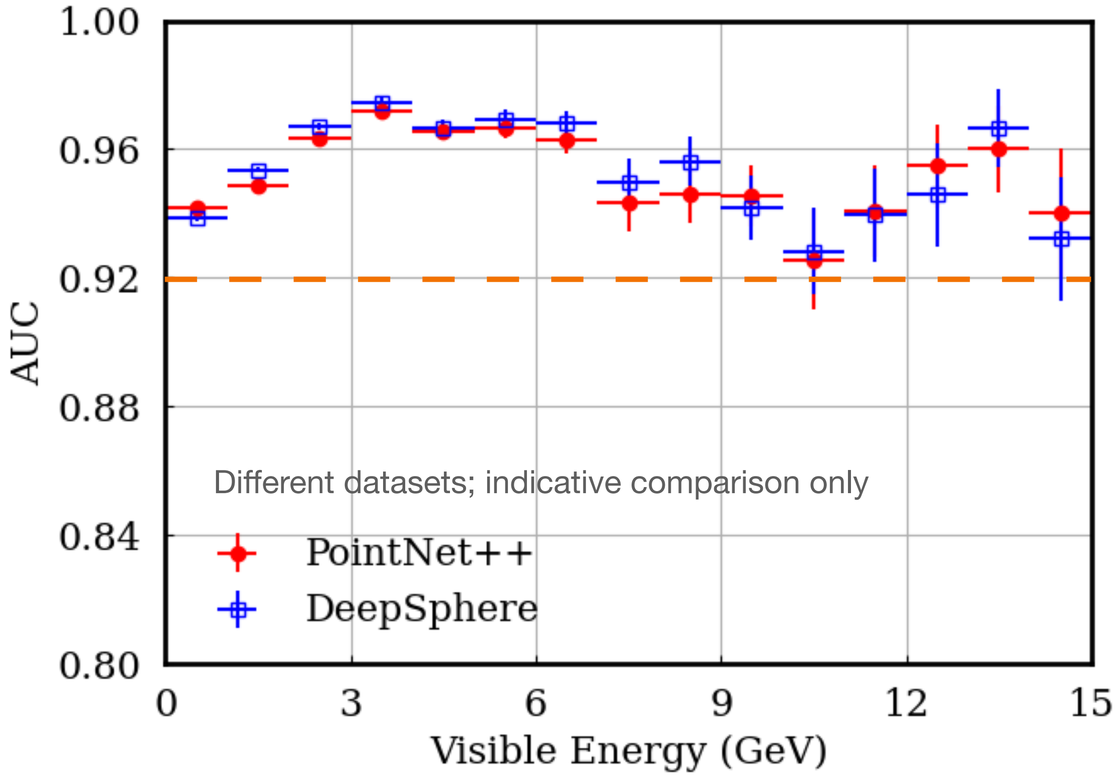


Secondary triggers

Improved 3-class PID



Other ML methods in JUNO



Adding secondary triggers improves 3-class PID across the visible-energy range.

Summary and future work

Completed:

- A Transformer-based framework was developed for atmospheric neutrino PID in JUNO by treating calib T/Q pairs as a time-ordered sequence of 1 ns hit frames.
- The model gives good performance for 3-class PID.
- Information from secondary triggers improves the performance on 3-class PID.
  - Secondary triggers also were utilized for 5-class PID.

Todo:

- Adapt the model to real data by investigating and addressing the reported discrepancies in detector response between data and MC.
- Atmospheric-neutrino analysis: learn VOLOR, incorporate my reconstruction results, and calculate JUNO's NMO sensitivity.

Thanks for your attention!

Backup

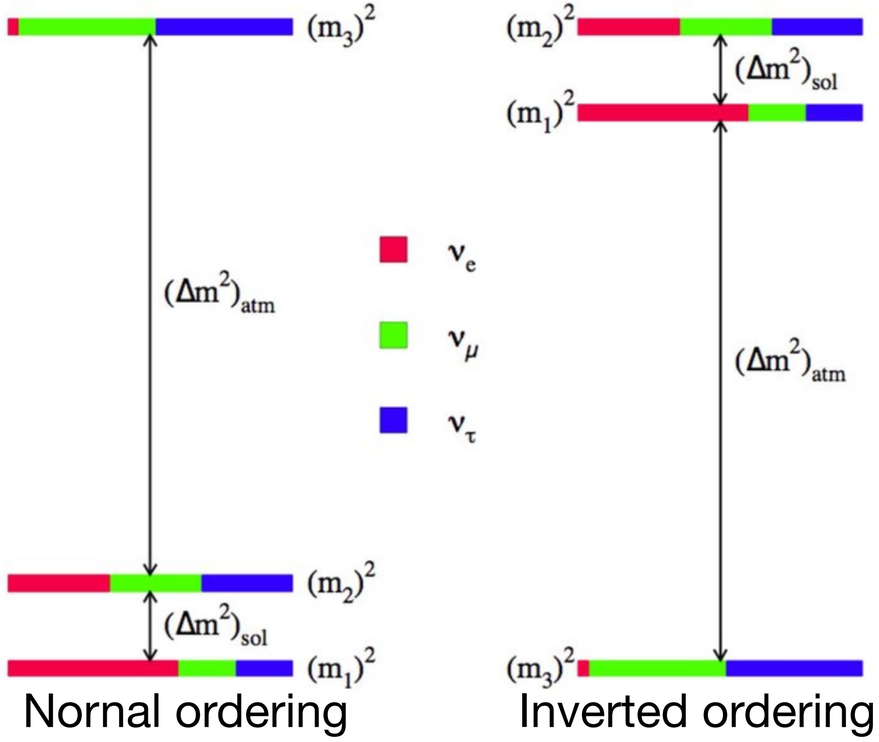
Background

Mass ordering

Neutrino Mass Ordering (NMO) has two hypotheses:

- Normal ordering, $m_3 > m_2 > m_1$;
- Inverted ordering, $m_2 > m_1 > m_3$.

Determining the mass ordering is important for interpreting neutrino oscillations and for measurements of **CP violation**.



Background

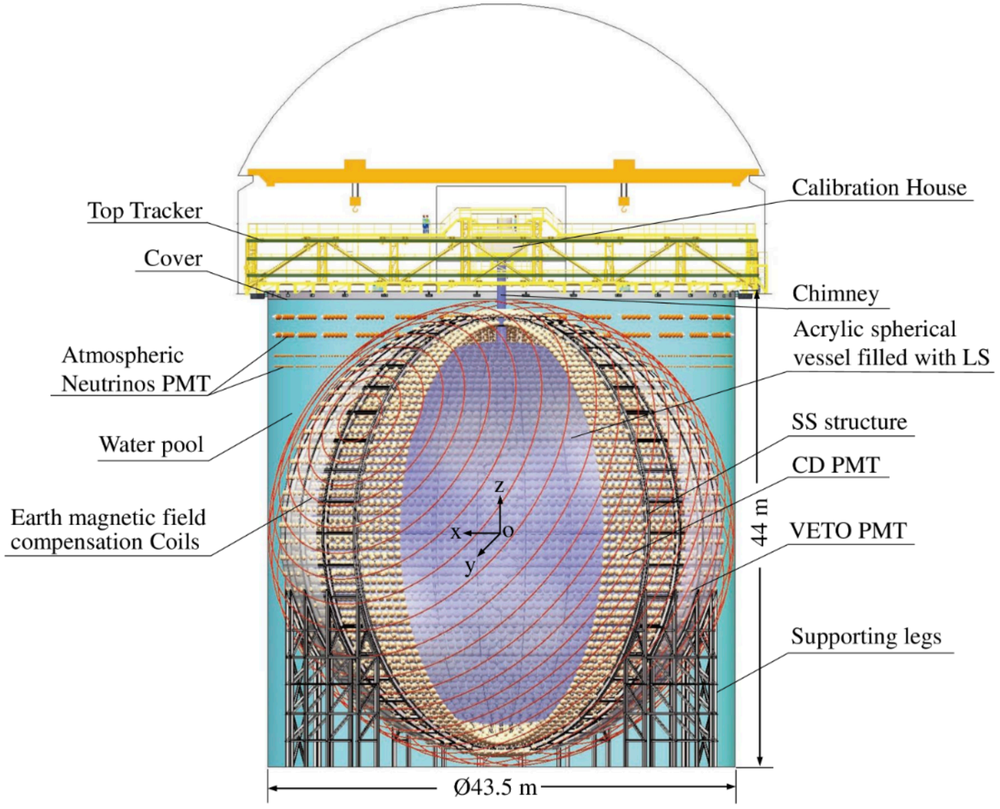
Introduction to JUNO

JUNO is located 700 meters underground in southern China, ~53 km away from Taishan and Yangjiang Nuclear Power Plants. Its primary goal is to determine the NMO through precise measurements of reactor antineutrino oscillations.

JUNO combines excellent **energy resolution (~3% at 1 MeV)** with high neutron-tagging efficiency, low radioactive background, and a low energy threshold (~0.2 MeV).

The JUNO detector consists of three main parts:

- Center detector (CD) containing 20k tons of liquid scintillator (LS);
- Water pool (WP) that also serves as a water Cherenkov detector (WCD); —> veto muons, shield radiation
- Top tracker (TT). —> detect muon direction



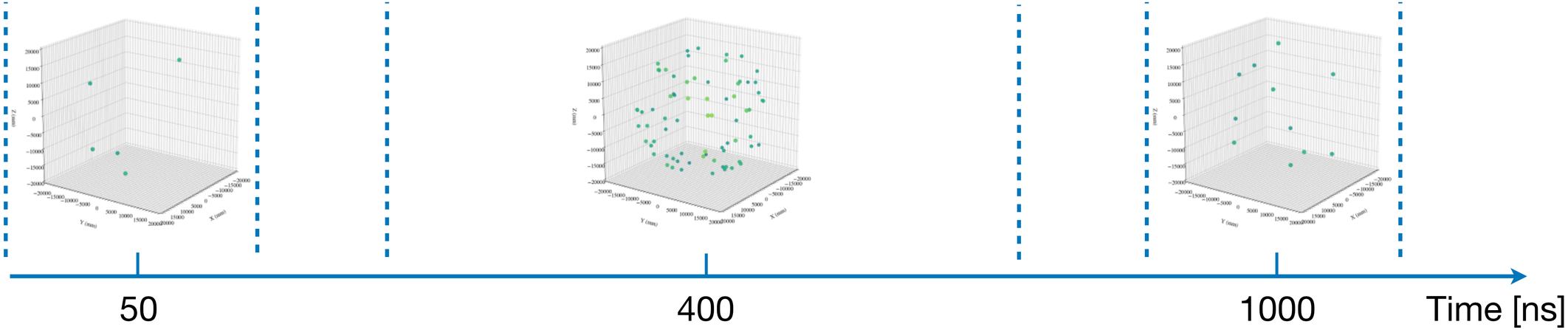
Drawing of JUNO Detector

Model Architecture

Step 1

PMT id	3827	4390	…	13	78	201	…	15155	…	524	743	…
Time	56	56	…	400	400	400	…	400	…	1000	1000	…
Charge	1.28	1.04	…	376.26	300.97	339.39	…	0.13	…	1.63	1.27	…

Frames



The Model treats a event as a video. It divides the whole sequence of raw hits to 1ns long frames:

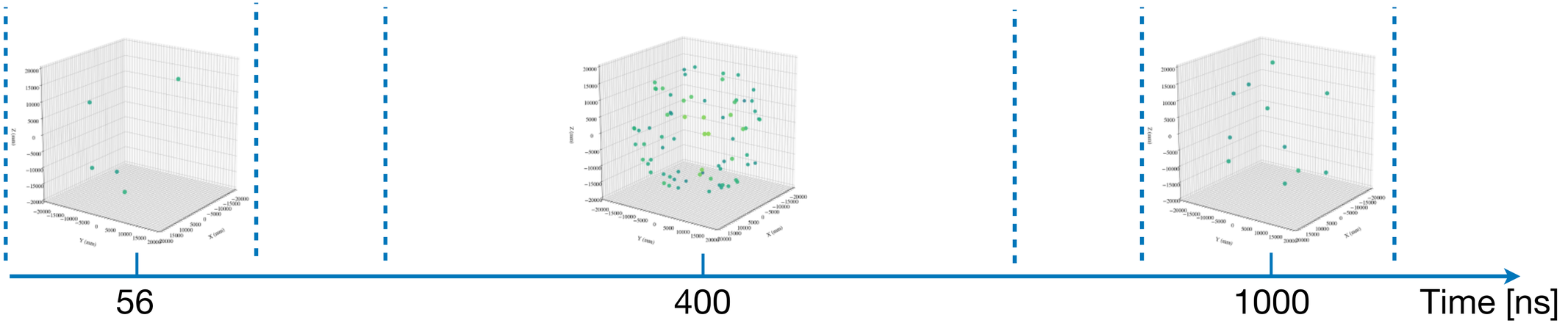
1. Select raw hits from [50, 1000) ns to focus on the reaction.
2. A hit feature includes PMT (x, y, z), charge and time, linearly projected into a abstract vector.
3. Group the hits by time into frames.

Model Architecture

Hits

PMT id	3827	4390	…	13	78	201	…	15155	…	524	743	…
Time	56	56	…	400	400	400	…	400	…	1000	1000	…
Charge	1.28	1.04	…	376.26	300.97	339.39	…	0.13	…	1.63	1.27	…

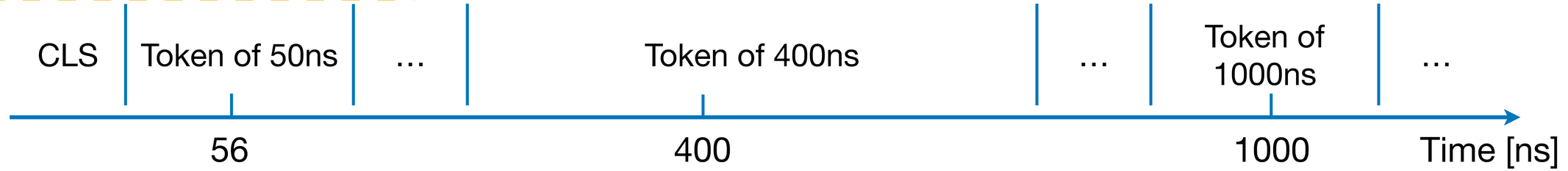
Frames



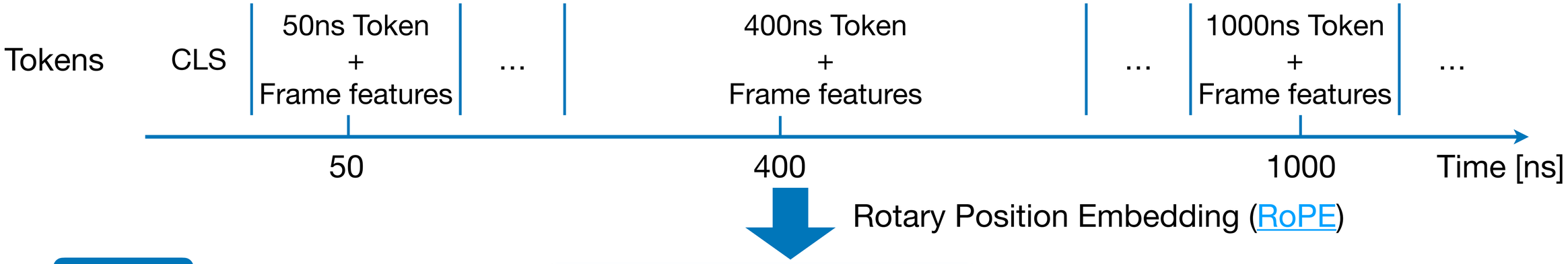
Treat each event as a temporal sequence of 1 ns frames. Hits within each frame are compressed into one token by attention pooling. This greatly **reduces the sequence length while preserving the temporal structure** of the event.

Attention pooling *Giving different weights to hits for pooling*

Tokens



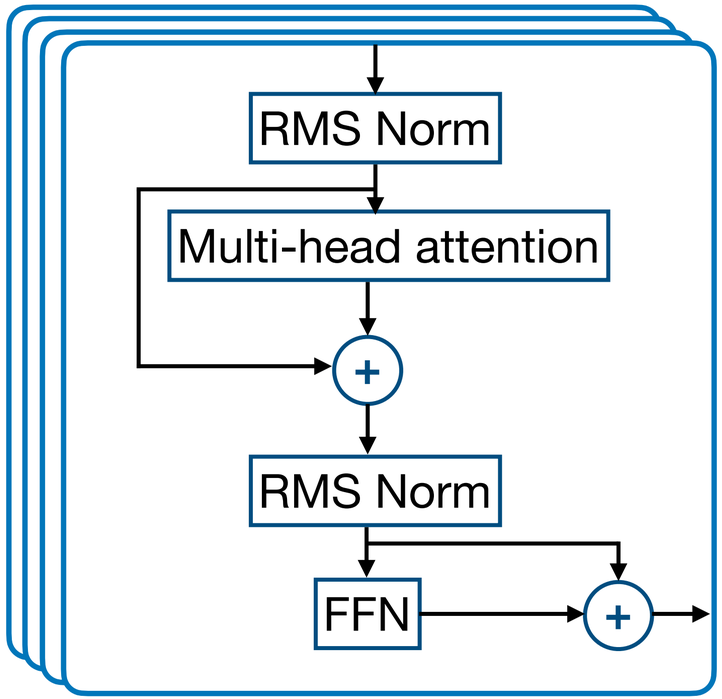
Model Architecture



Rotary Position Embedding (RoPE)

Step 3

1. Connect frame-level features: (qsum, nhit, nPMT, charge center) with the frames' token.
2. Embed time into frames' tokens by RoPE.
3. Sequence of tokens are fed to 4 self-attention blocks.
4. After self-attention, task-specific head gives 3-class logits, directions or energy.



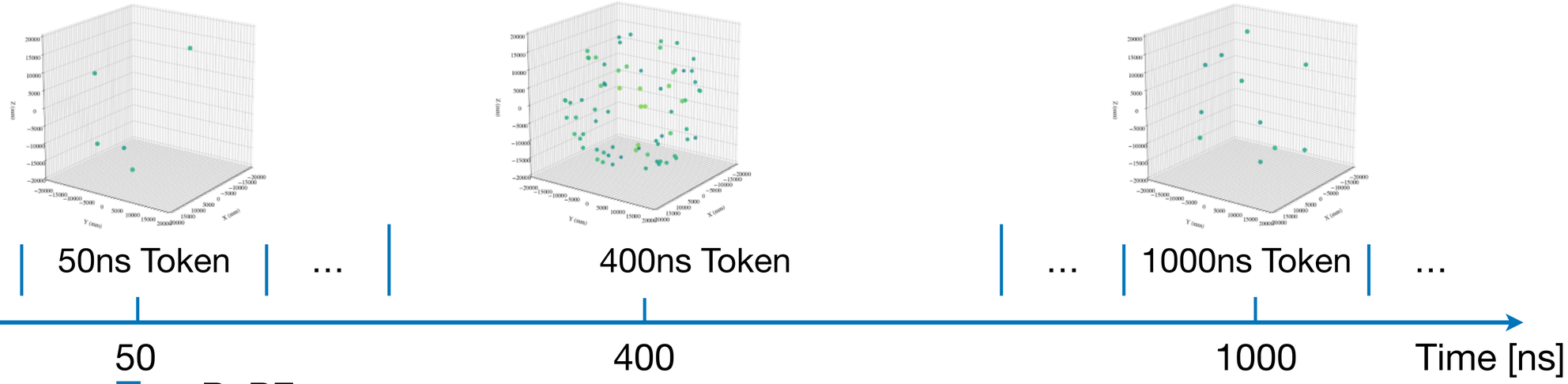
+ Residual connection

FFN: Feed-Forward Network

Task-specific head

Logits, direction or energy

Model Architecture



+ RoPE

4 Attention Blocks

Task-specific head

PID, direction or energy

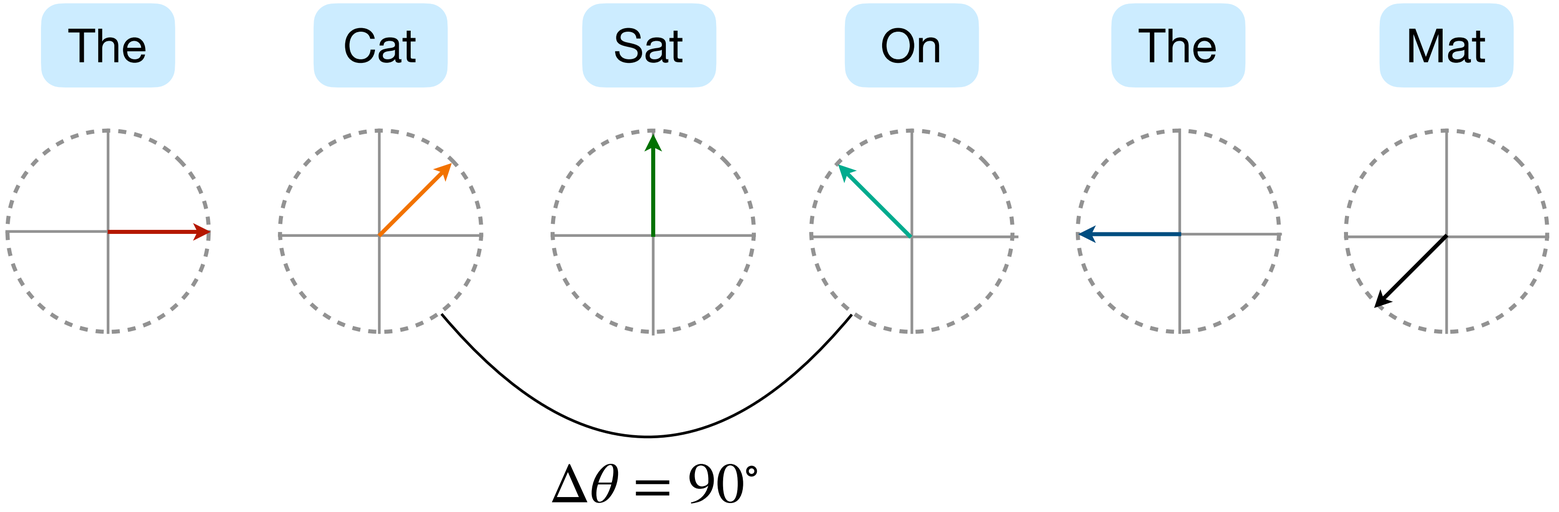
Event-level Transformer

Frame tokens are combined with frame-level information, and temporal order is encoded with rotary positional encoding **(RoPE)**.

Multi-head self-attention then learns correlations between different stages of the event to understand **how the event evolves**.

The final event representation is passed to a task-specific prediction head.

Model Architecture



RoPE encodes token positions through rotations.

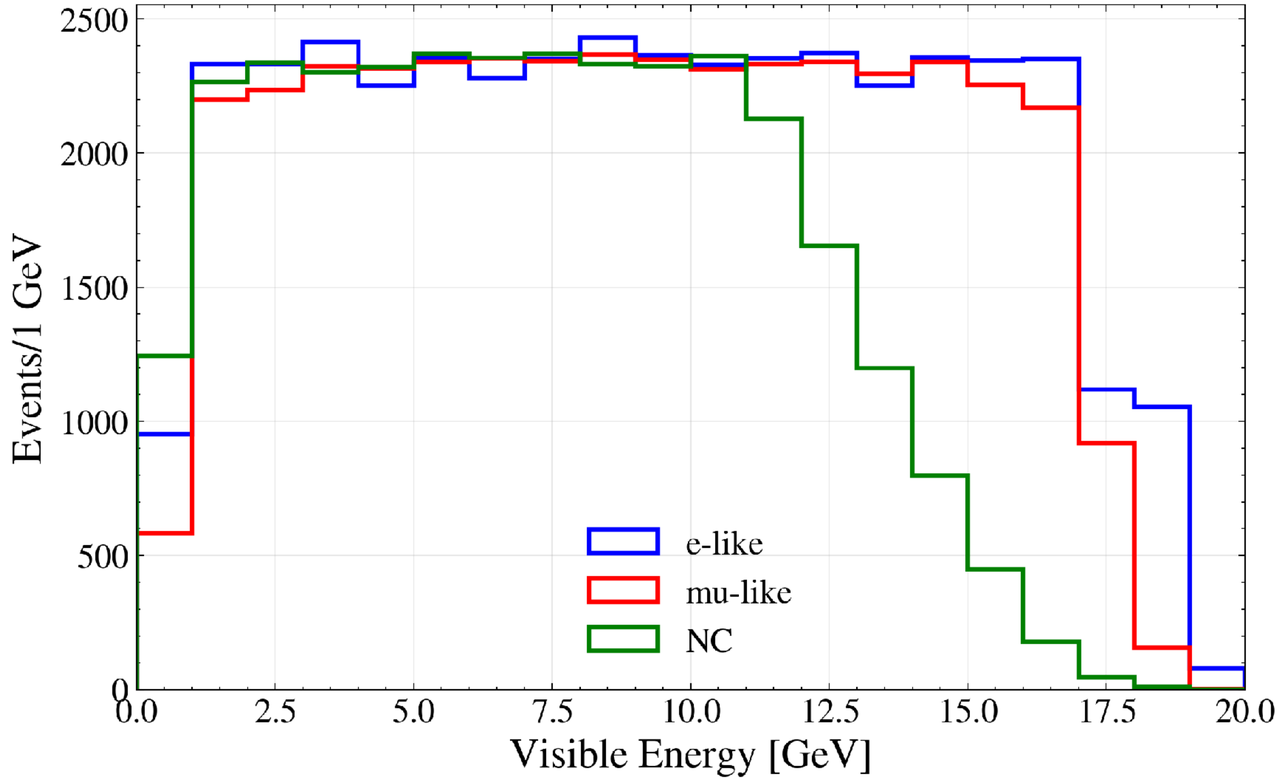
The relative rotation angles are **preserved in attention dot products**, enabling relative positional awareness. ($\vec{q} \cdot \vec{k} = |\vec{q}||\vec{k}|\cos\Delta\theta$)

https://arxiv.org/abs/2104.09864

Dataset

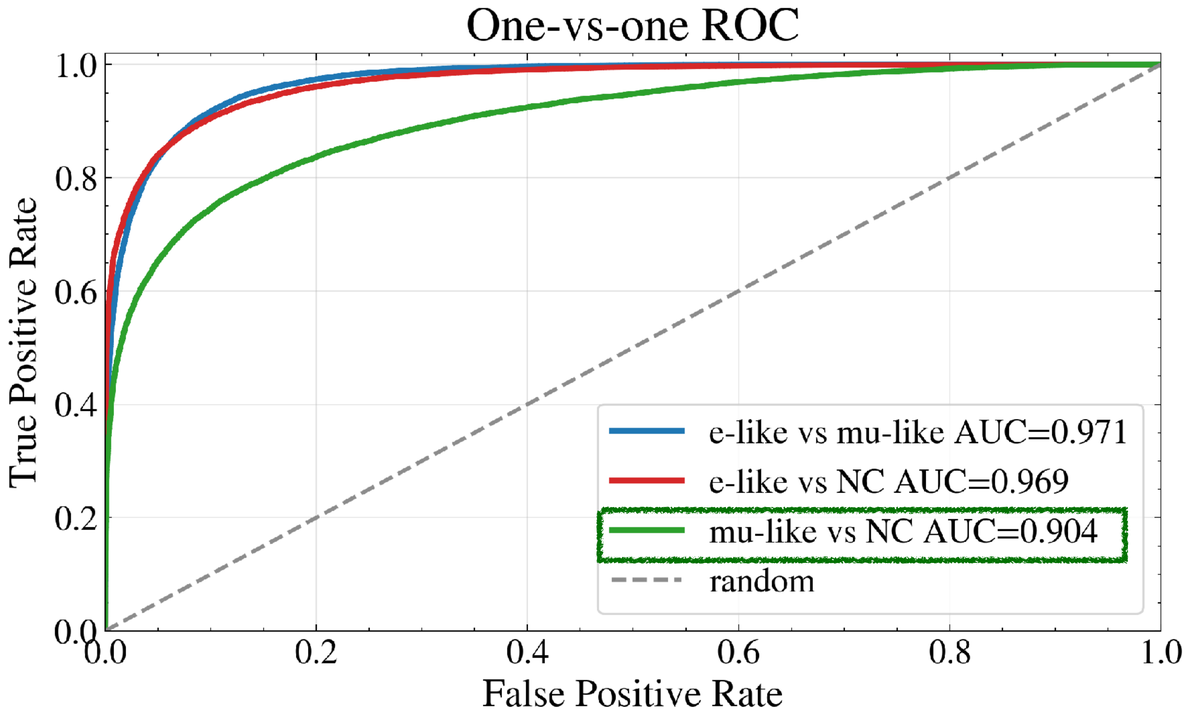
JUNO official J24.1.2 version simulation: fully contained events, 0.1–20 GeV

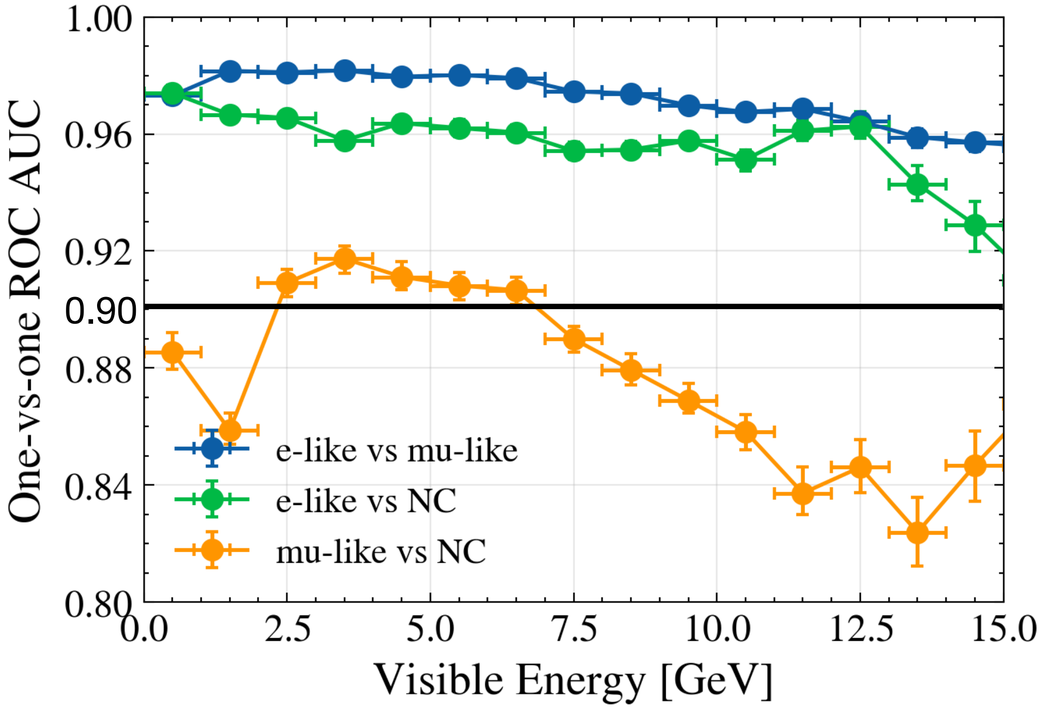
- ~110k training events; ~90k validation/testing events
- Training distribution shown below; The **visible energy distribution is basically flat**, and the statistics are roughly balanced to avoid introducing statistical imbalance during training.



3-class PID

AUC: **A**rea **U**nder the ROC **C**urve. The higher the AUC, the better the discrimination ability.



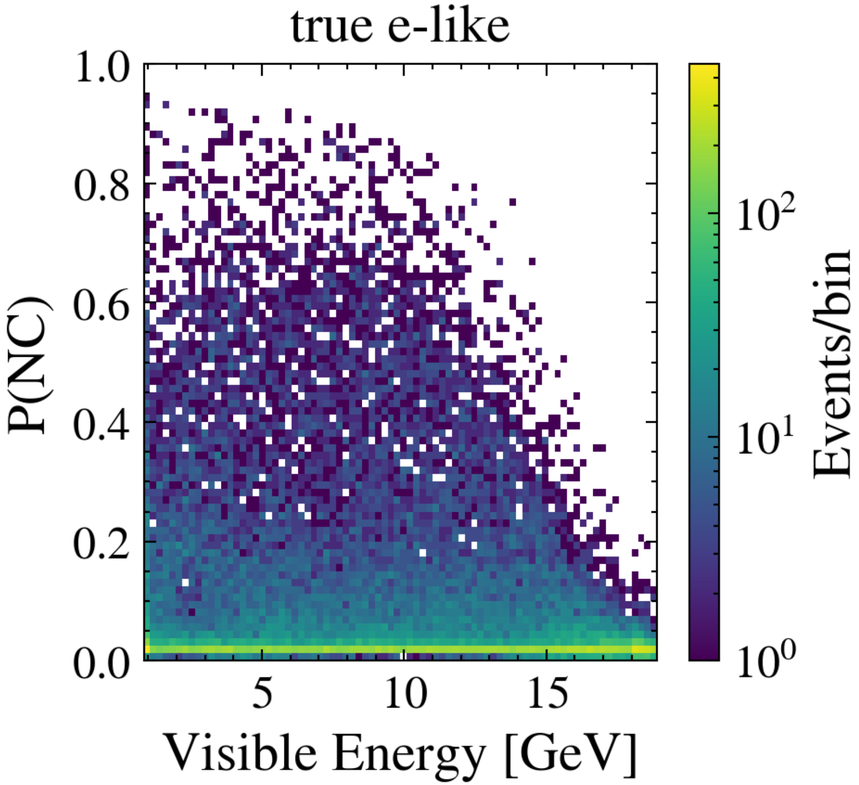


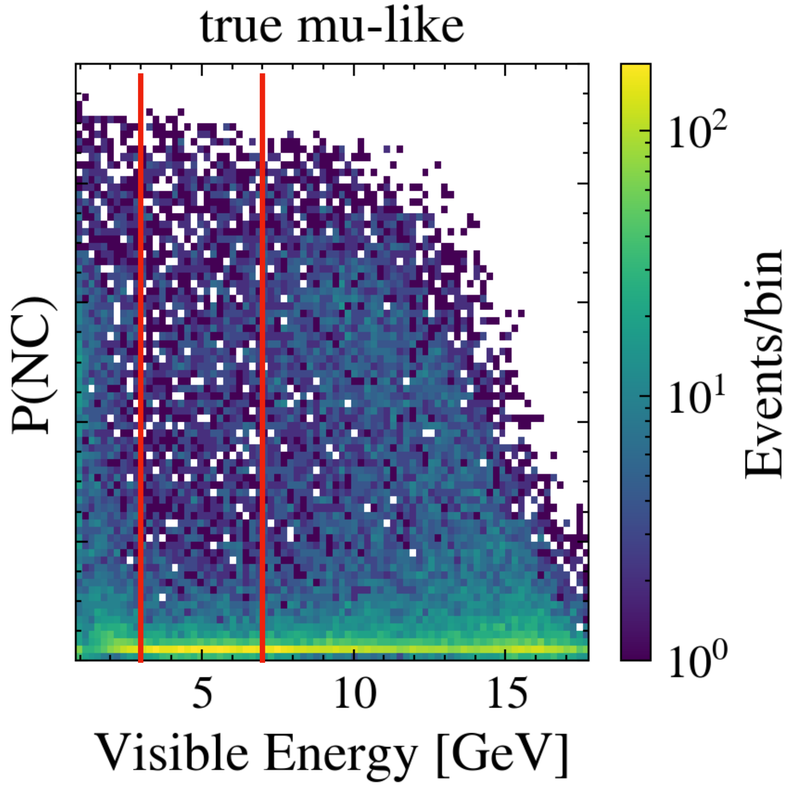
- Bad performance of classifying between mu-like and NC.

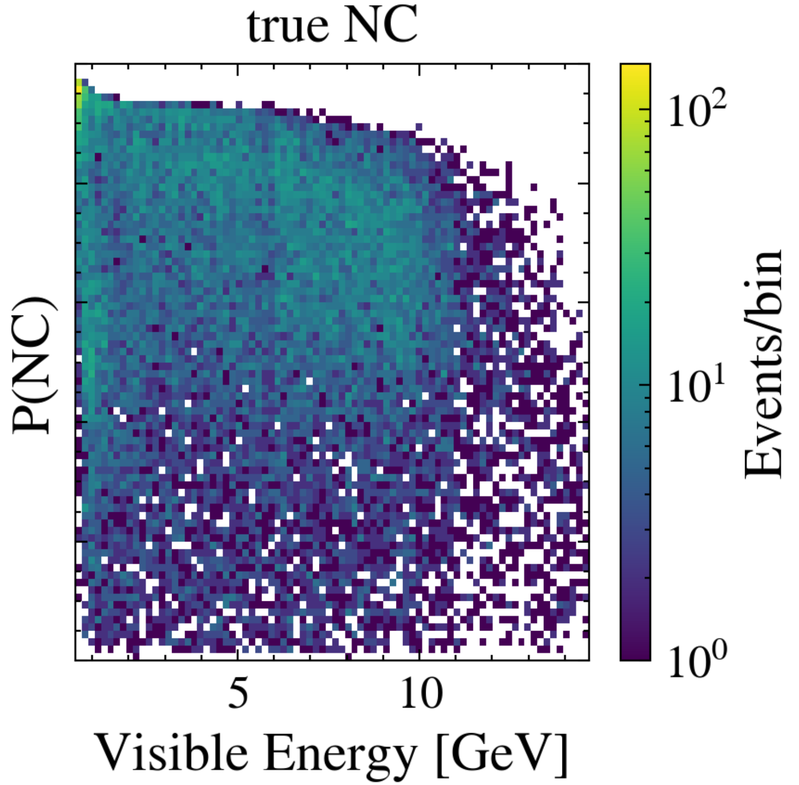
3-class PID: Evis Dependence

The model gives:

- High P(NC) to mu-like events with Evis > 7 GeV;
- Low P(NC) to mu-like events with Evis between 2 and 7 GeV;
- High P(NC) to mu-like events with Evis < 2 GeV.



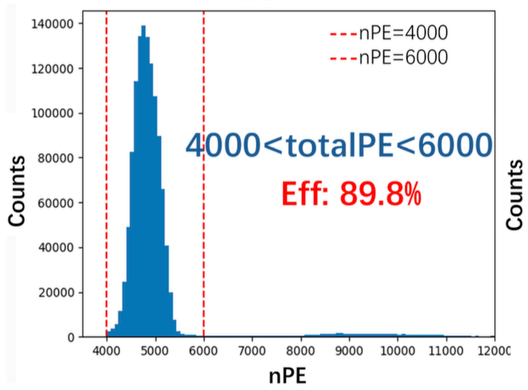


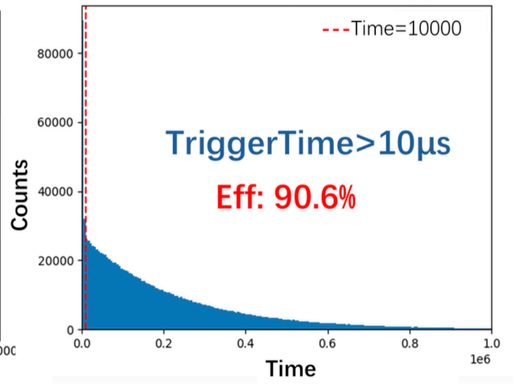


5-class PID: Secondary Triggers

- Selection: TriggerTime > 10us after primary trigger, 4000 < totalPE < 6000. The selection efficiency for neutron capture is 87.9%.
- Each secondary trigger includes 9 features: totalPE, peak charge, peak time, nhits, nPMTs, charge center xyz, full width at half maximum.

nPE and TriggerTime distributions of neutron capture triggers

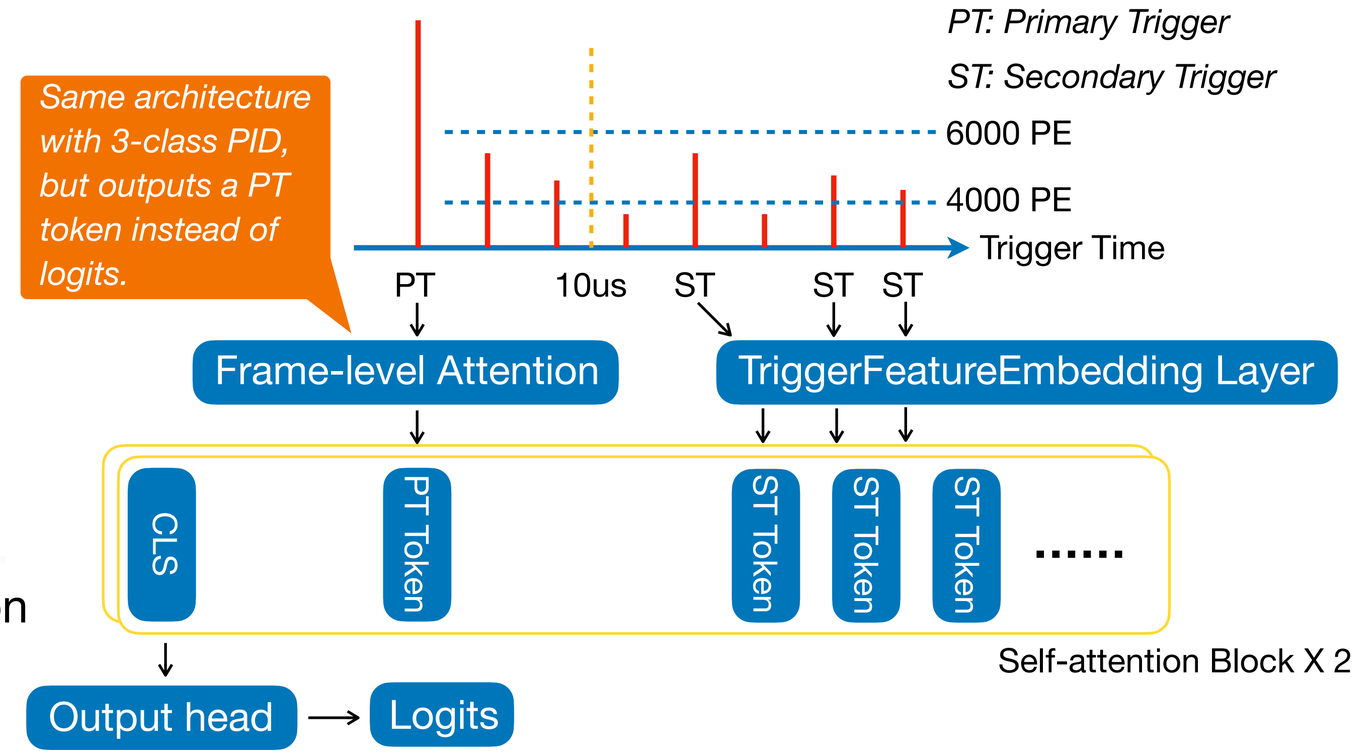




corresponds to 2–2.7 MeV (2.2 MeV gamma by neutron capture)

reduces contamination from the primary trigger

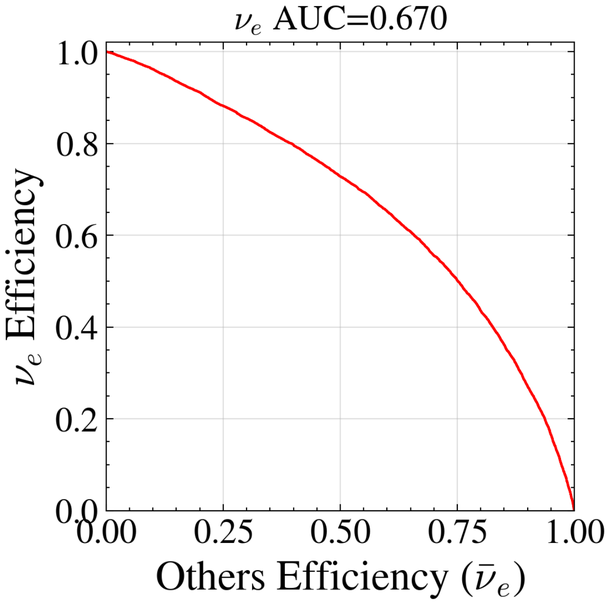
Same architecture with 3-class PID, but outputs a PT token instead of logits.



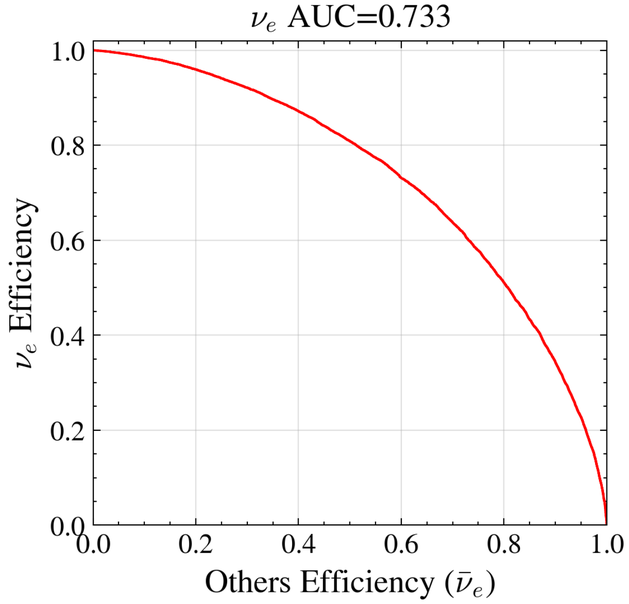
5-class PID: Secondary Triggers

- Selection: TriggerTime > 10us after primary trigger, 4000 < totalPE < 6000. The selection efficiency for neutron capture is 87.9%.
- Each secondary trigger includes 9 features: totalPE, peak charge, peak time, nhits, nPMTs, charge center xyz, full width at half maximum.

W/O Secondary triggers



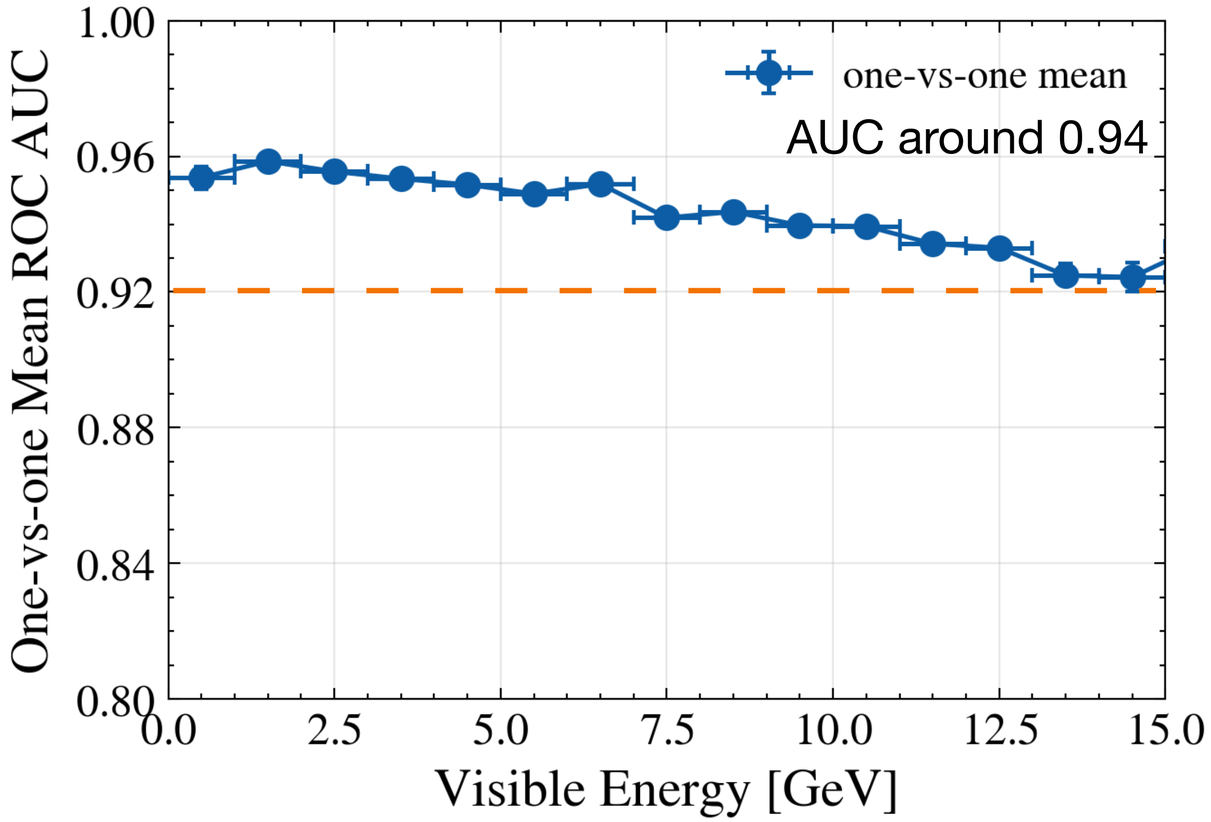
W/ Secondary triggers



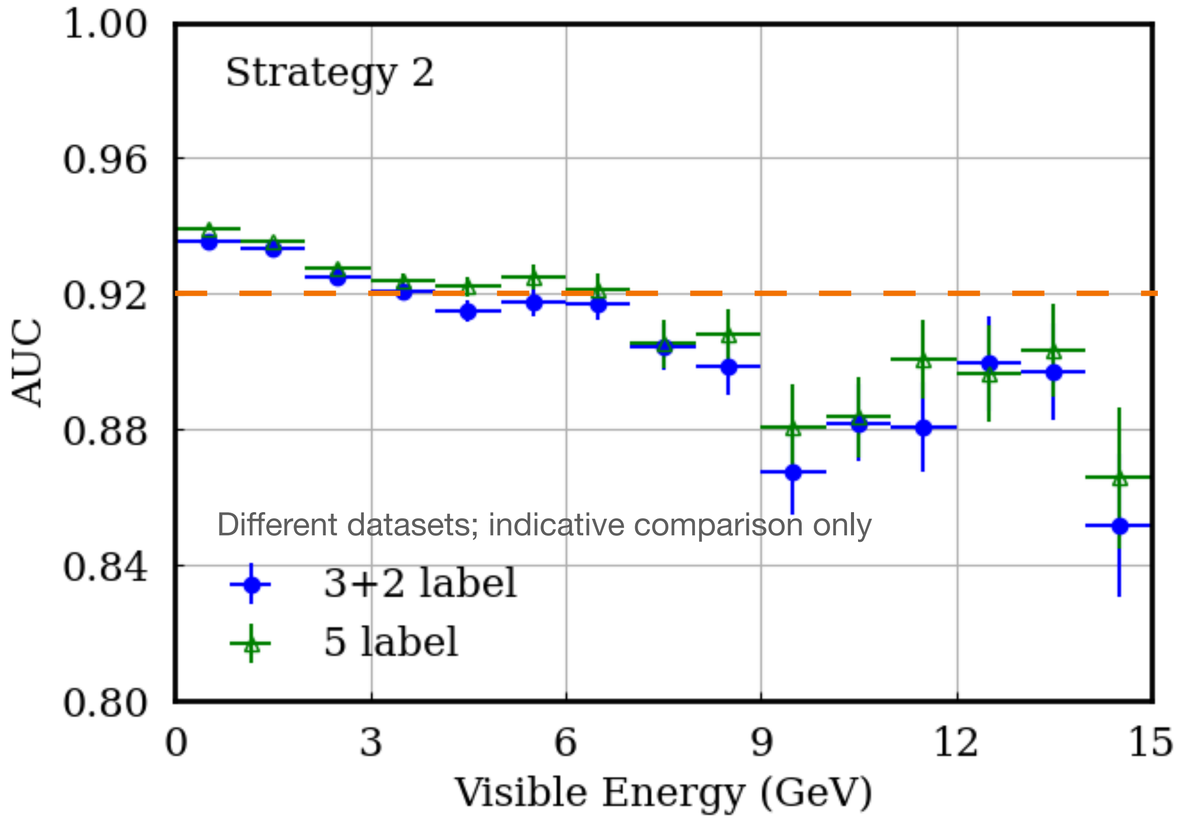
Secondary triggers

5-label: $\nu_e/\bar{\nu}_e/\nu_\mu/\bar{\nu}_\mu/NC$

My work

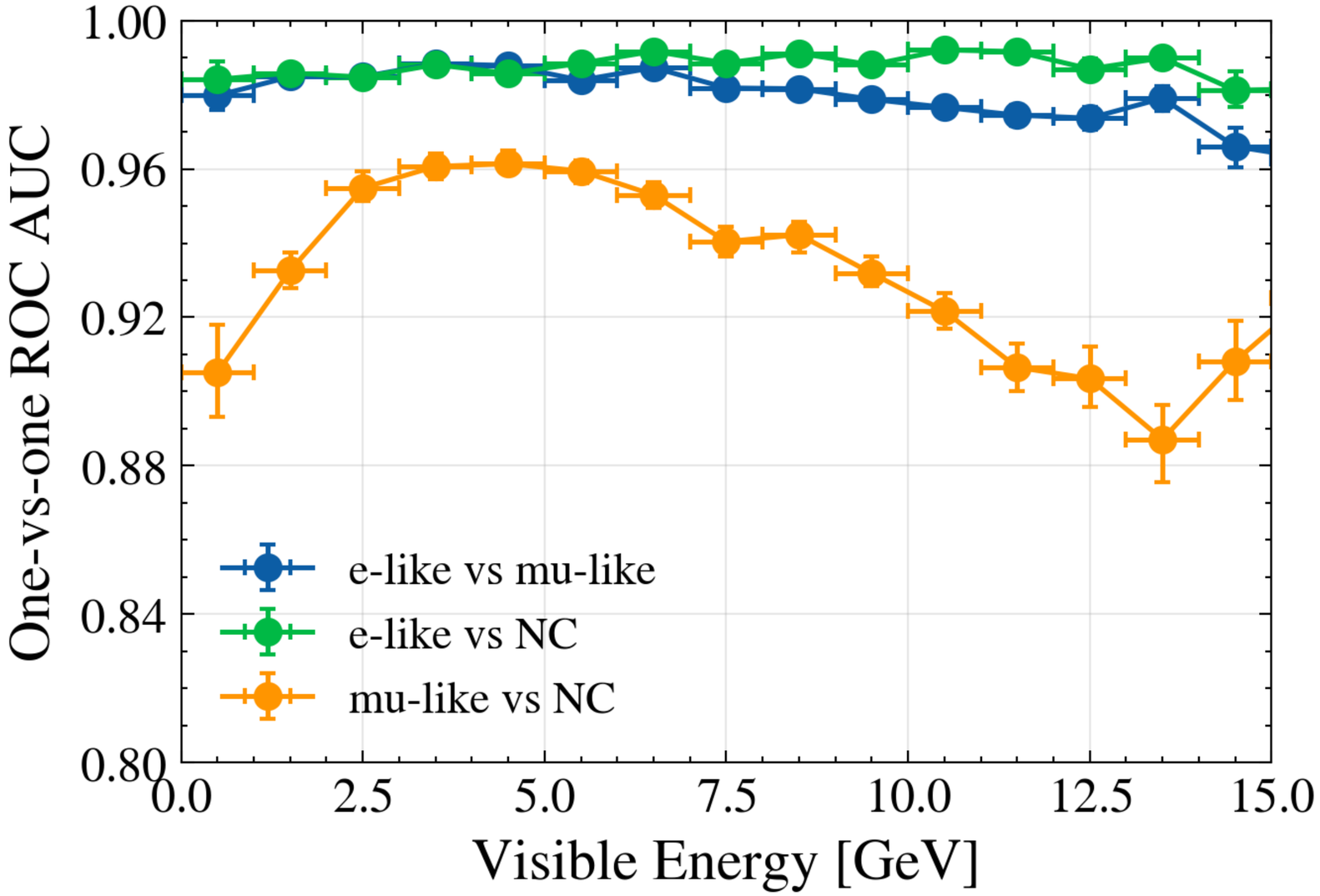


Other ML methods in JUNO



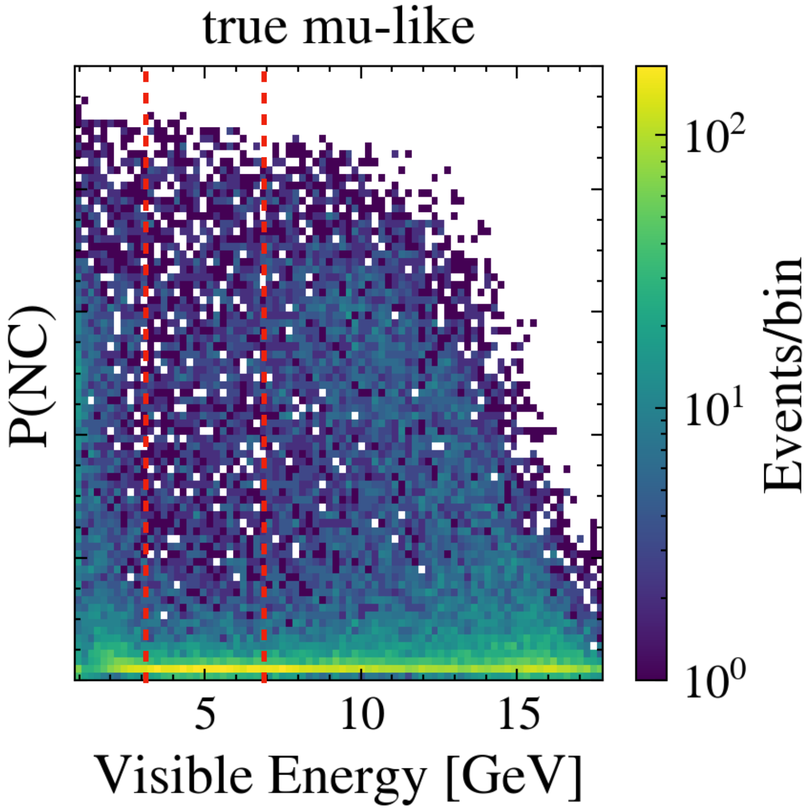
My model shows higher AUC values and a milder decline with increasing visible energy.

5-class PID: Secondary triggers



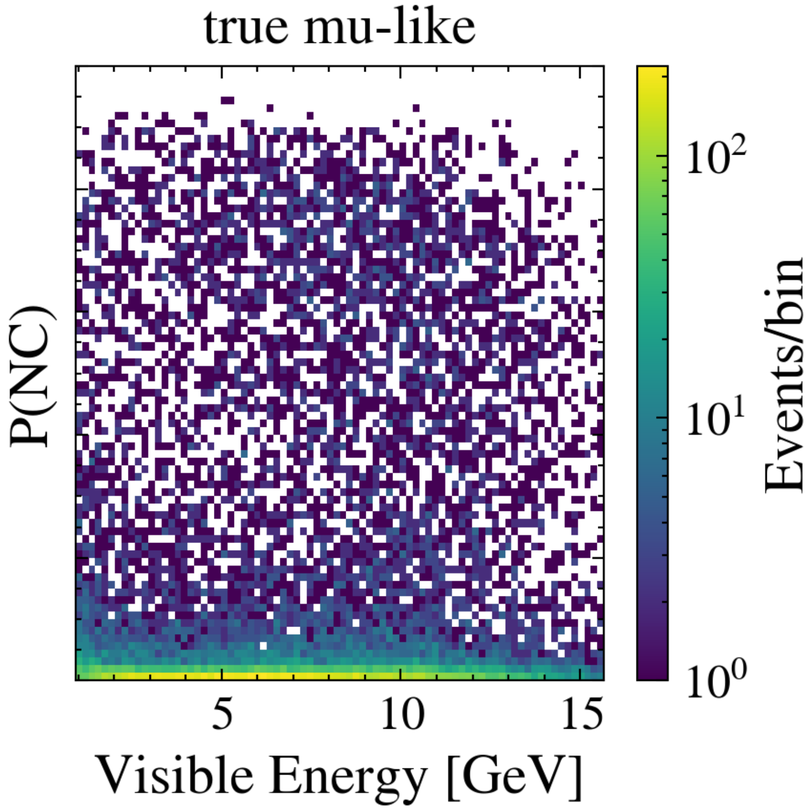
Energy-dependence check

Primary trigger only



Energy-dependent P(NC)

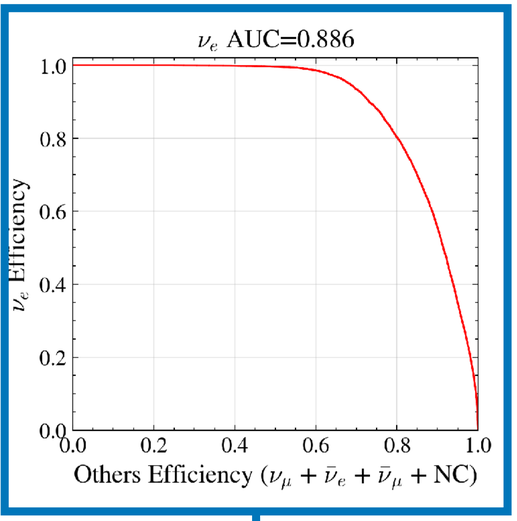
Primary + secondary triggers

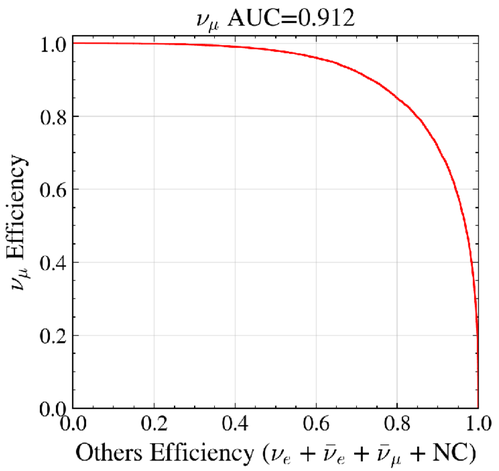


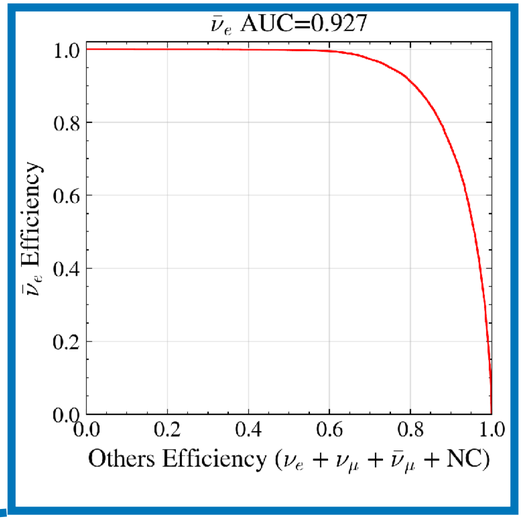
Pattern less clear in this sample

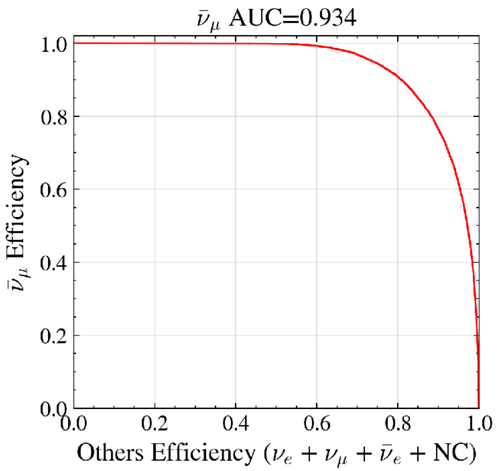
5-class PID

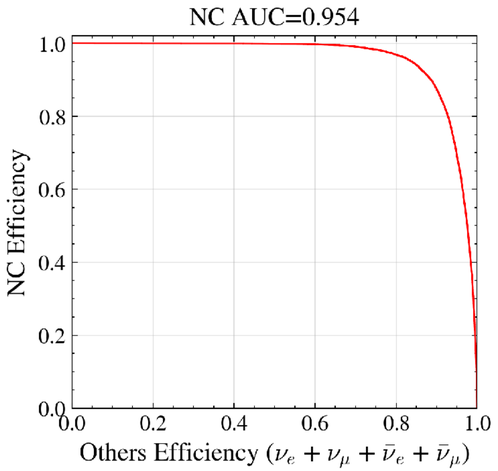
$\nu_e/\bar{\nu}_e/\nu_\mu/\bar{\nu}_\mu/NC$

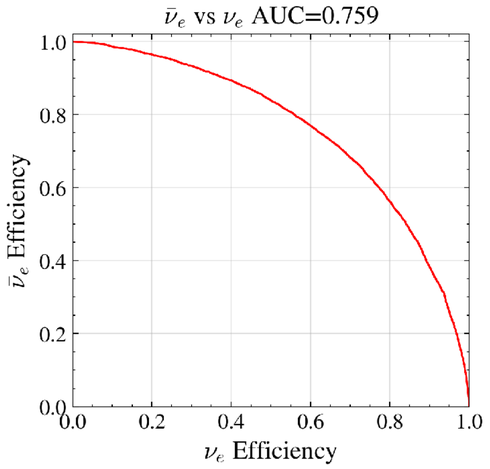


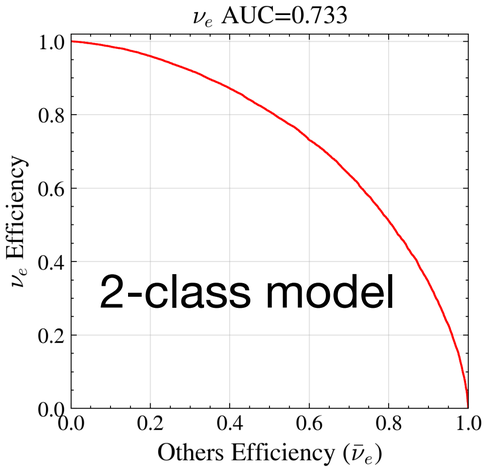








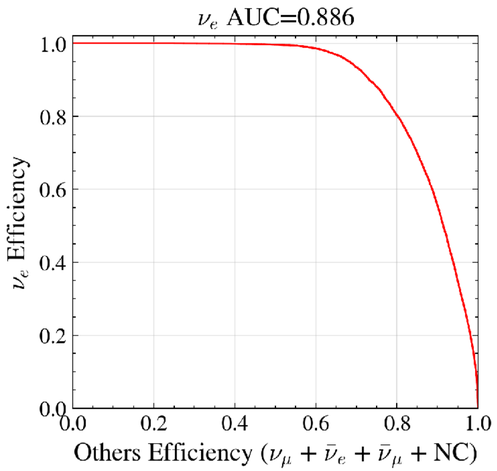


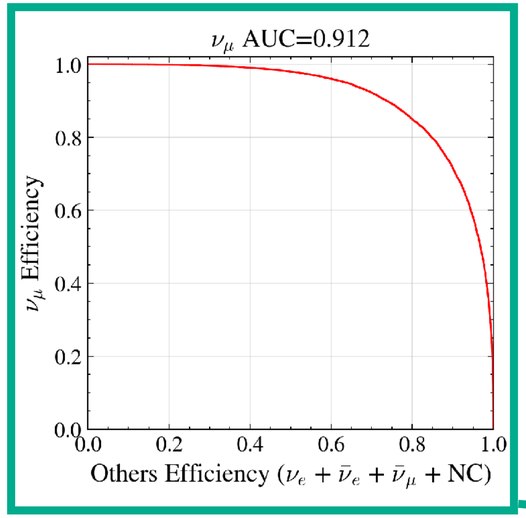


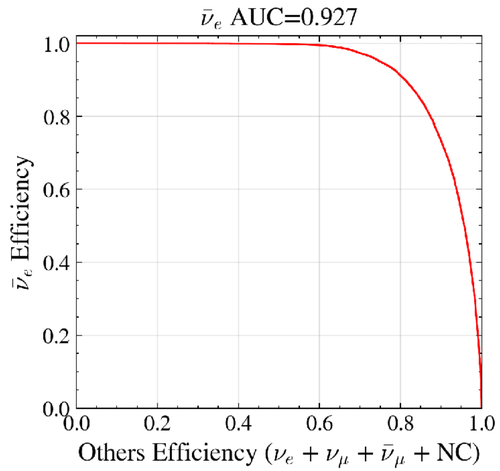
2-class model

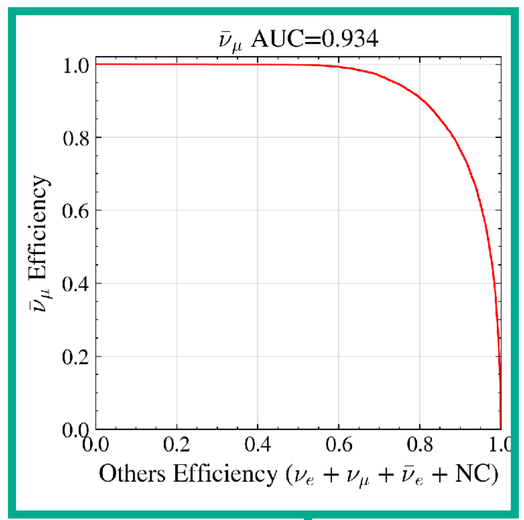
5-class PID

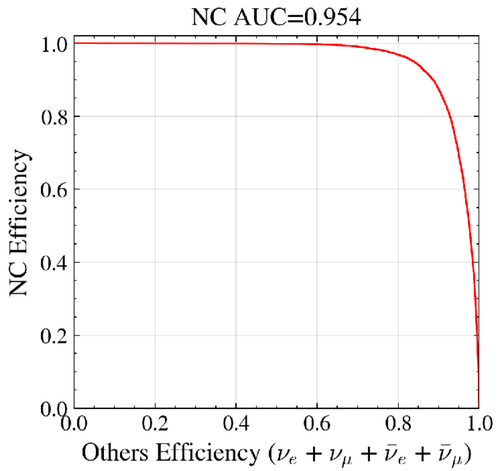
$\nu_e/\bar{\nu}_e/\nu_\mu/\bar{\nu}_\mu/NC$

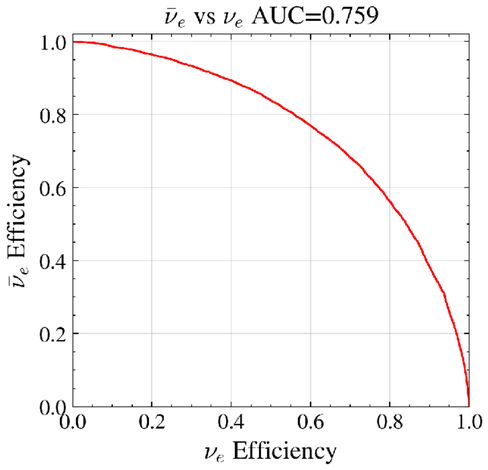


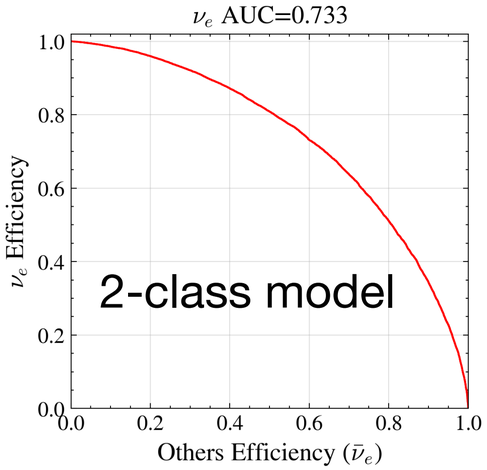












5-class PID is better than 2 class PID

