Identification of atmospheric neutrino's flavor in JUNO with machine learning

Wing Yan Ma[1], Fanrui Zeng[1], Jiaxi Liu[2], Xinhai He[2], Zhen Liu[2], Wuming Luo[2], Hongye Duyang[1], Teng Li[1], Yongpeng Zhang[2]

[1] Shandong University

[2] Institute of High Energy Physics, CAS

NPML 2024 @ Zurich

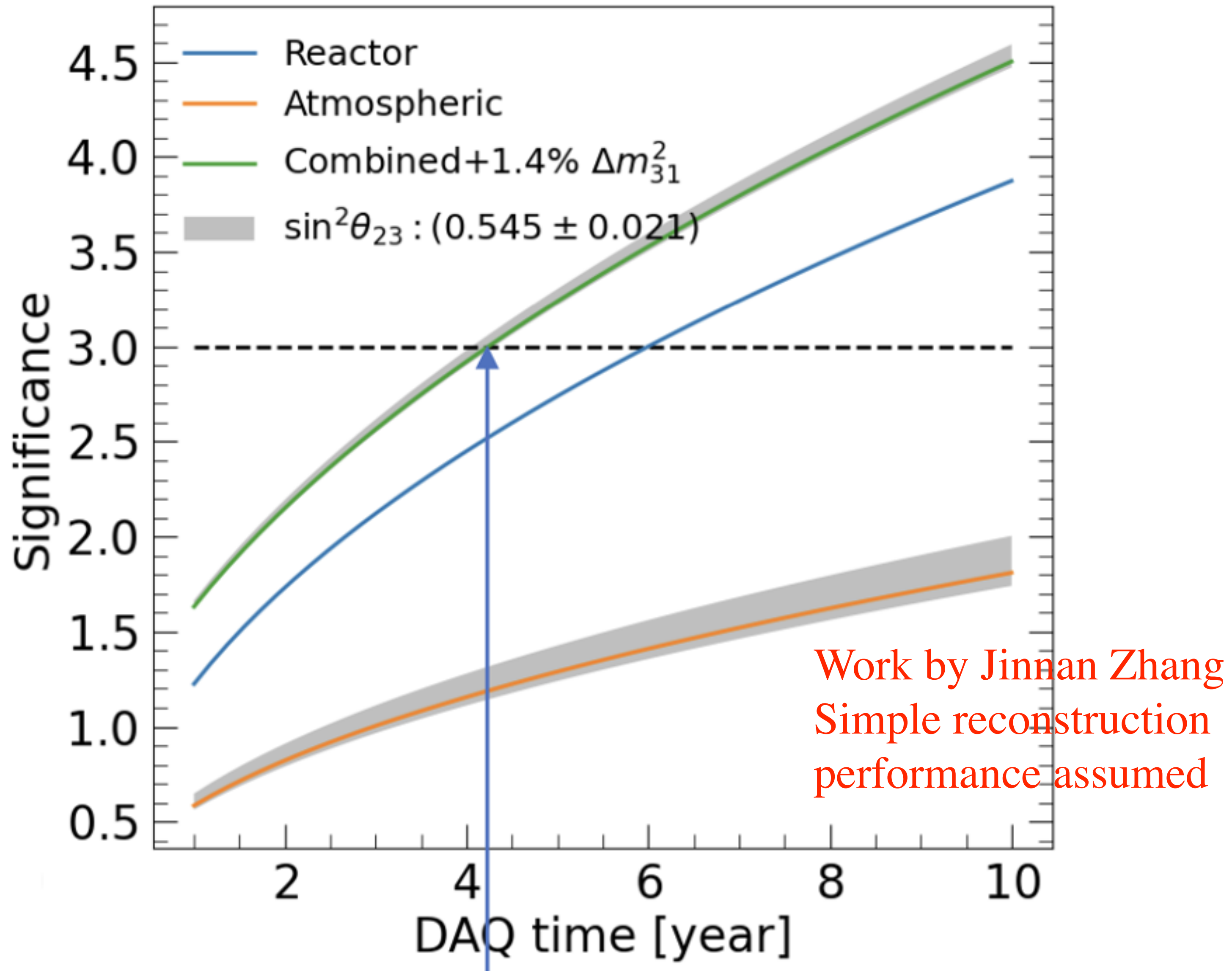
2024/06/25

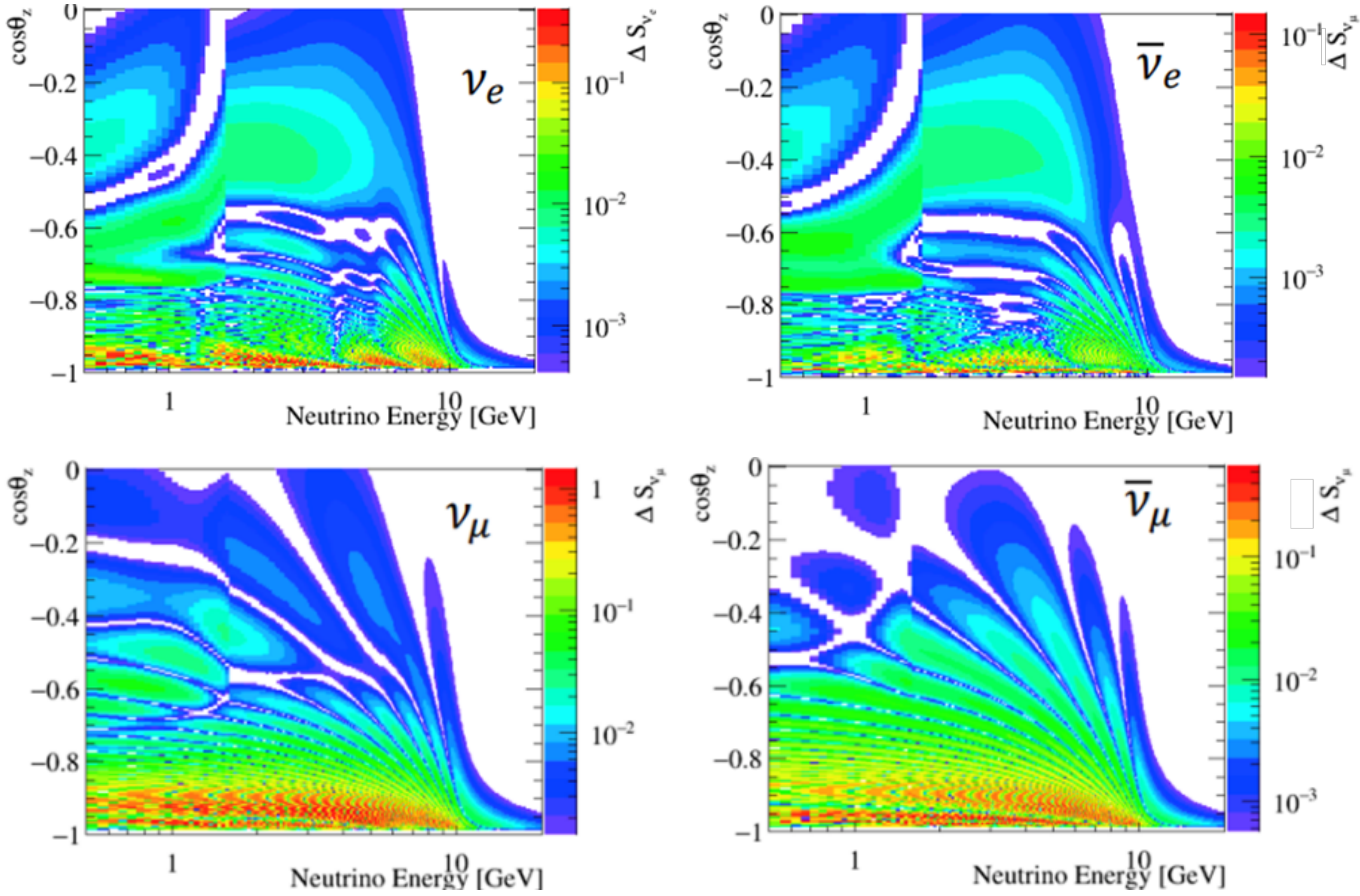
Outline

- Motivation
- Methodology
- Strategies and different ML models considered
- Model performance
- Summary

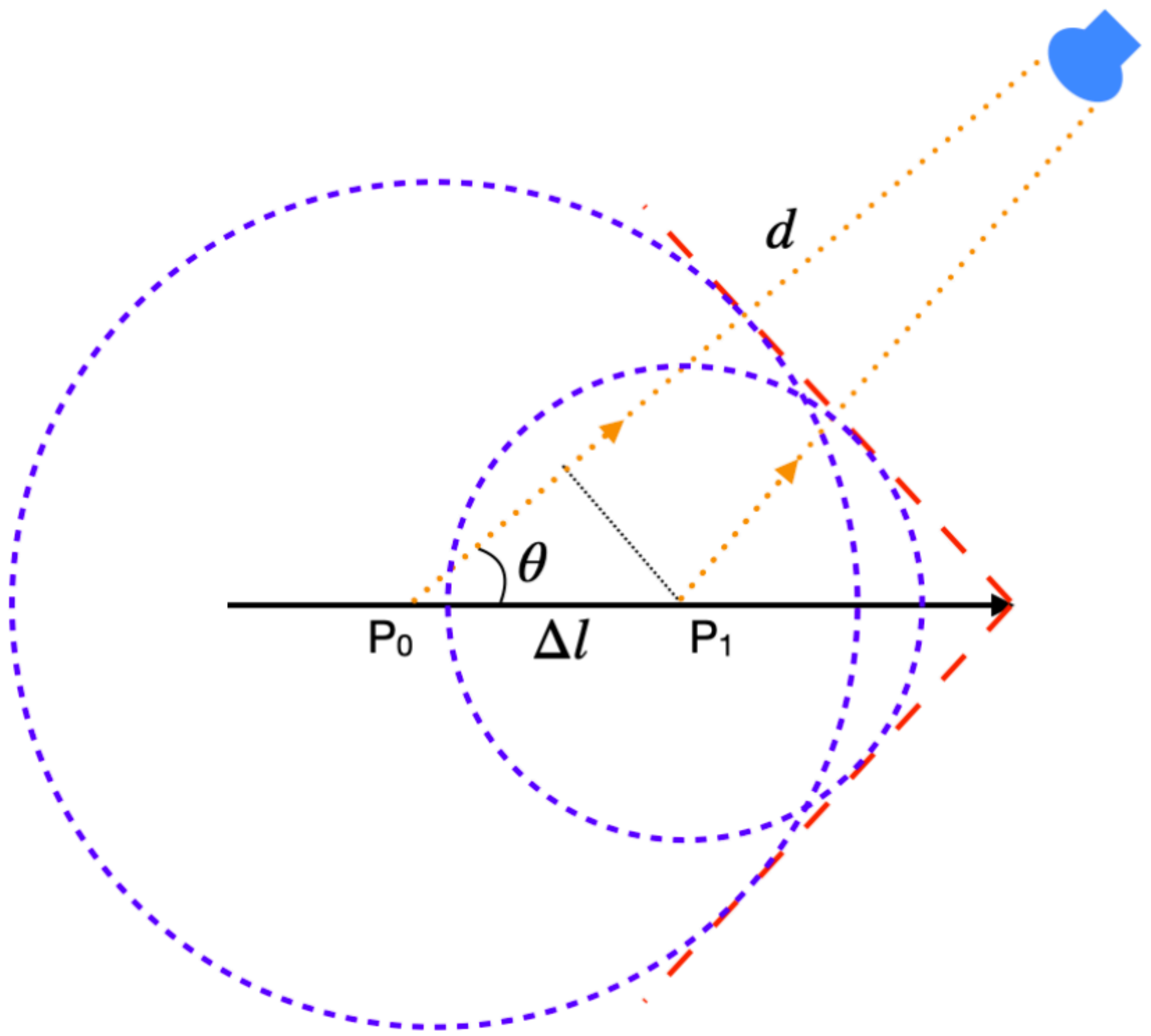
Motivation

- NMO sensitivity can be enhanced by studying neutrino oscillations in GeV region
- To study ν_{atm} oscillations one needs to reconstruct neutrinos' direction/energy/**flavor (particle type)**
- Different neutrino flavor exhibits different oscillation probabilities between two neutrino mass order
  - Signal Charged-Current (CC) vs Background Neutral-Current (NC)
  - Muon (anti)neutrinos vs electron (anti)neutrinos $\overset{(-)}{\nu}_{\mu}/\overset{(-)}{\nu}_{e}$
  - Neutrinos vs Antineutrinos $\nu/\bar{\nu}$
- We demonstrate the capability of our ML approach in performing PID for atmospheric neutrinos

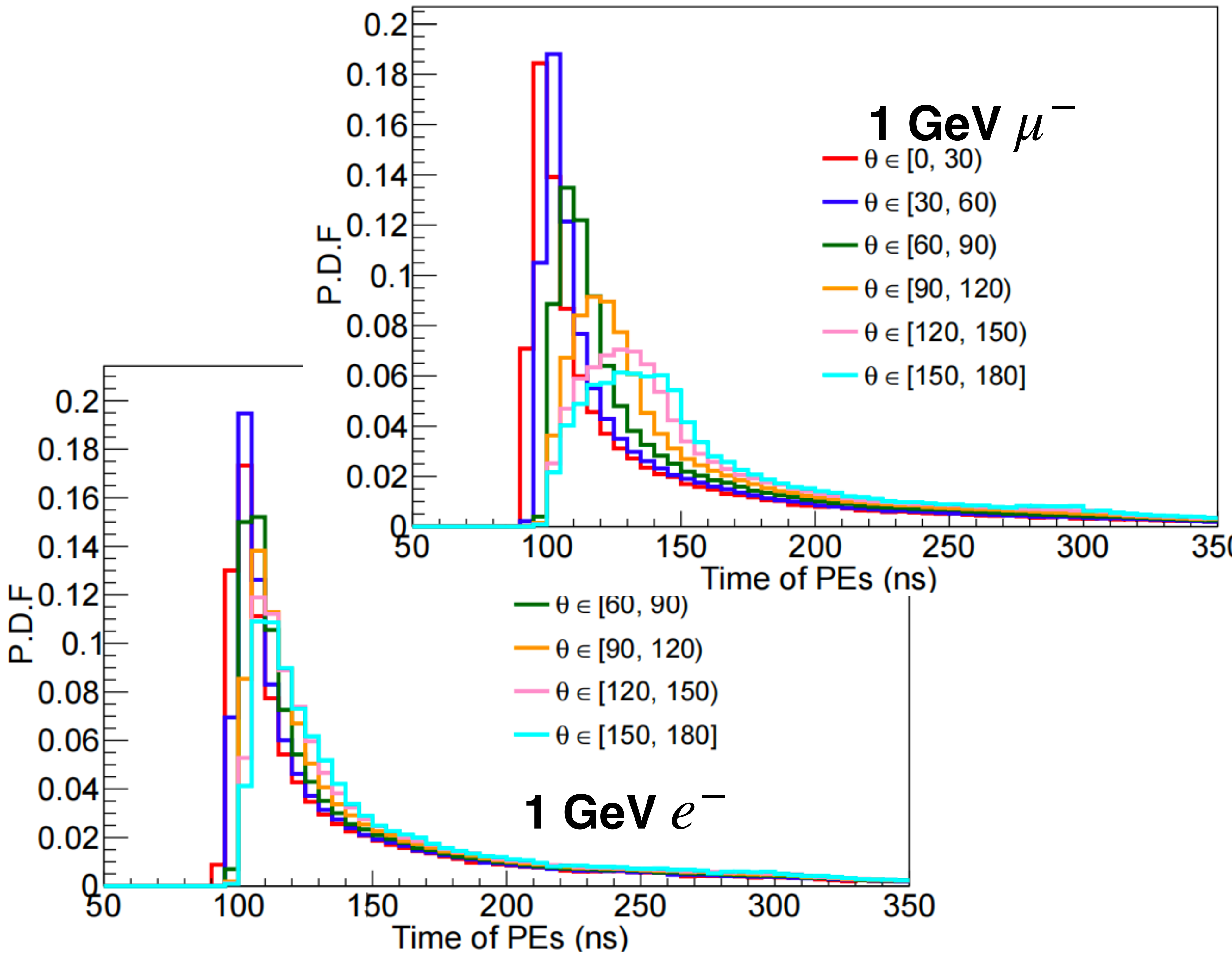




Scintillation light at the detector



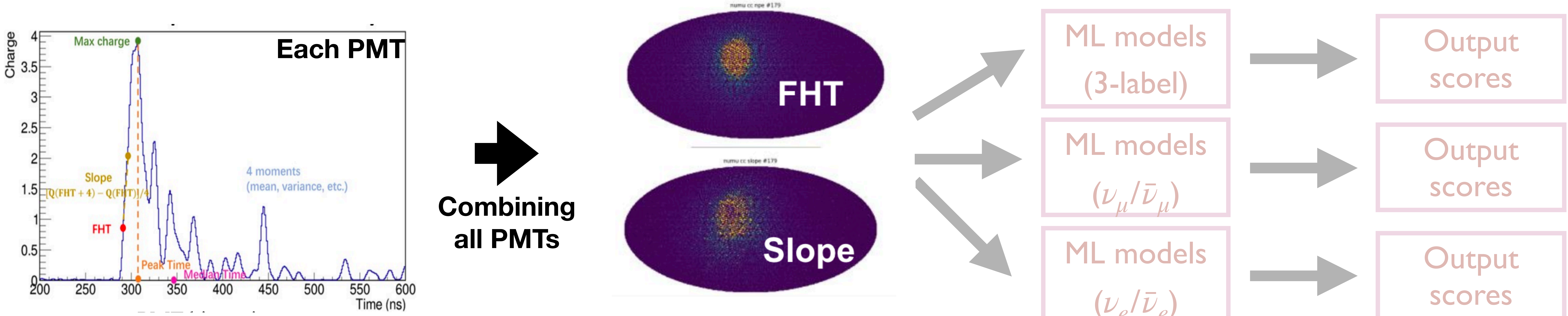
- Light seen by PMTs of an LS detector is a superposition of light generated from many points along the track
- Shape of light curve received by each PMT depends on :
  - Angle w.r.t. track direction θ
  - Track starting and stopping position
  - Particle type - different dE/dx
- Typical LS detectors are designed for low-energy neutrinos - ν_{atm} **oscillations measurements using LS detectors is challenging**



Methodology

Directly feeding full waveform from all PMTs are computationally expensive - features that reflects the waveforms are extracted to reduce data volume

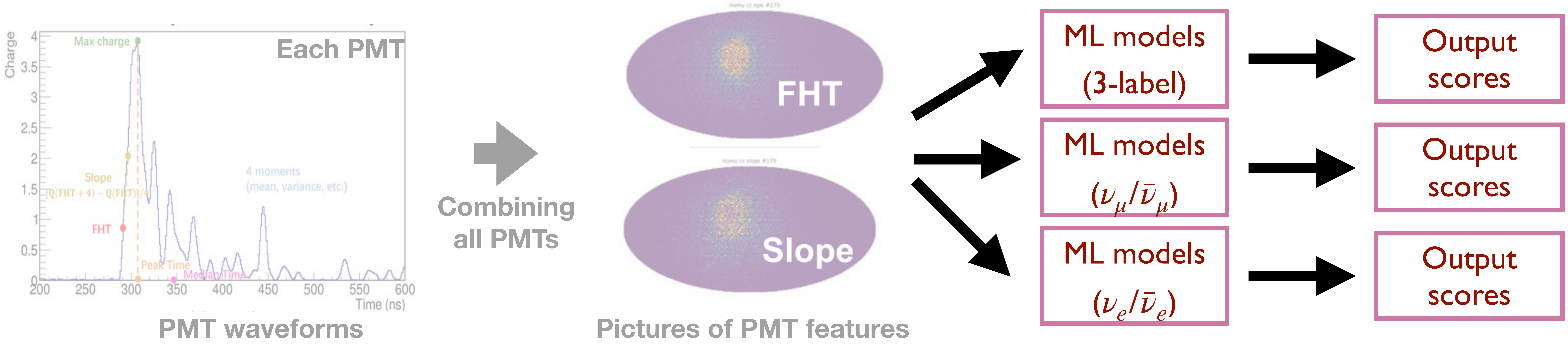
- FHT: time of first photon arriving at a PMT
- Slope: average slope of curve at the first 4 ns
- Peak time, peak charge, total charge
- **Other features such as median time and four moments (mean, std, skewness, kurtosis)**



PMT waveforms **Pictures of PMT features**

Methodology

- The output of the ML models is a set of scores associated for each category for a given event
- By default, the ML models will assign the category with the highest score for each event
- **2-step approach**: 3-label classification (NC, $\overset{(-)}{\nu}_{\mu}$, $\overset{(-)}{\nu}_{e}$) followed by $\nu/\bar{\nu}$ classification, expect the ML models can each learn to specifically perform one classification tasks, either 3-label or 2-label

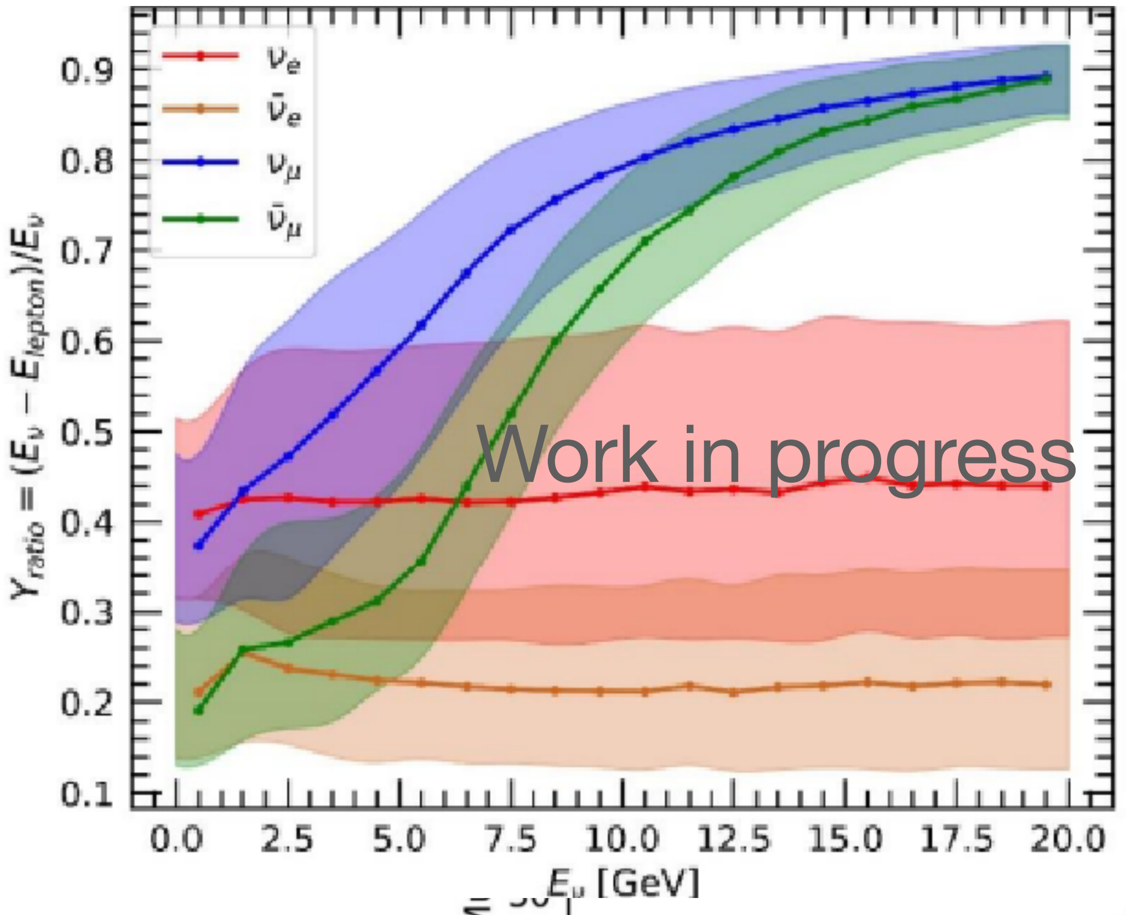


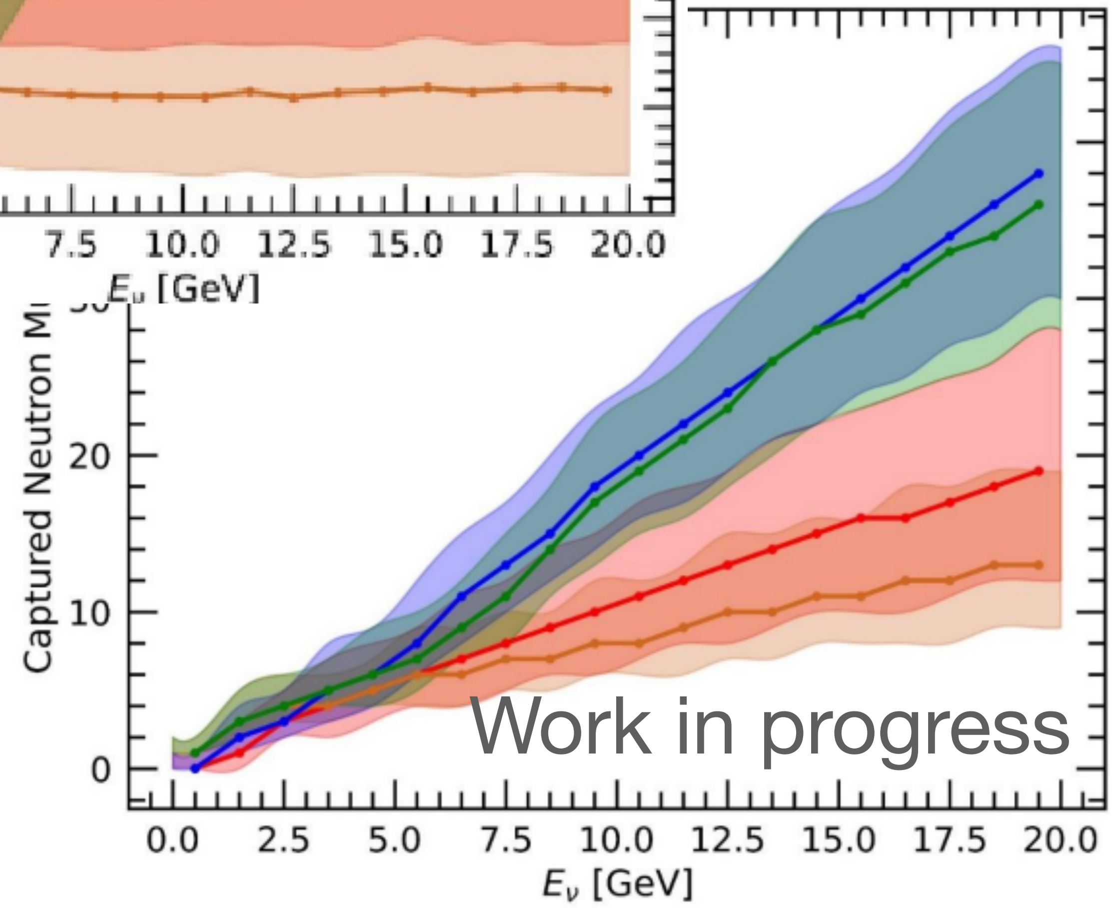
Utilising neutron capture information

E.g. CCQE interactions

$$\begin{aligned}\nu_e + n &\rightarrow e^- + p\\ \bar{\nu}_e + p &\rightarrow e^+ + n\\ \nu_\mu + n &\rightarrow \mu^- + p\\ \bar{\nu}_\mu + p &\rightarrow \mu^+ + n\end{aligned}$$

- The difference between each CC interactions are also reflected by the final state hadrons from ν interactions
- Final state neutrons are captured by hydrogens in LS and emit a 2.2 MeV in ~ 200 μs, create delayed triggers after primary interactions
- Such events can be selected from delayed trigger with high efficiency
- The difference between $\nu/\bar{\nu}$ interactions can also be reflected by the hadronic energy fraction variable $Y_{ratio} = (E_\nu - E_{lepton})/E_\nu$, reflected by observables such as neutron multiplicity
- Expect to provide additional power especially for $\nu_e/\bar{\nu}_e$

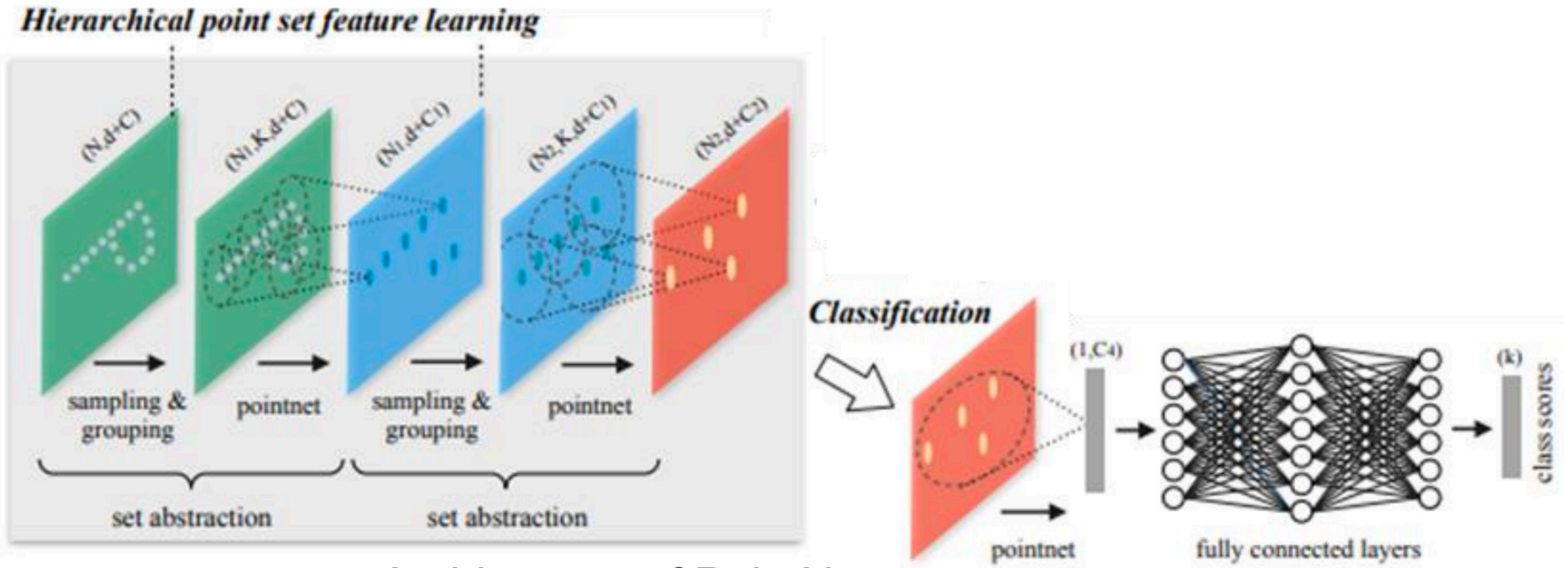




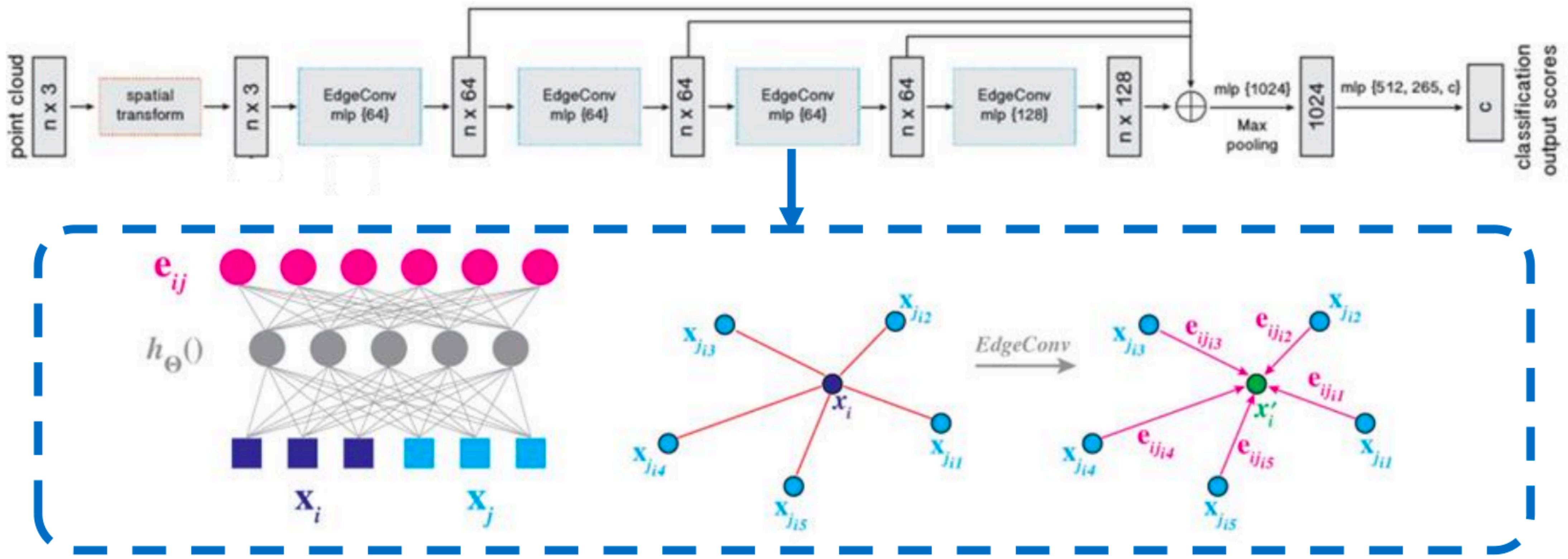
Two strategies (1)

1. Point cloud-based model: PointNet++, DGCNN

- Features extracted from primary triggers are fed into the PointNet++
- For neutron capture candidates, taking 3D point clouds $N \times$ [x, y, z] as inputs to a separate DGCNN model, capable of recovering neighborhood topology of point clouds with edge information
- Preserves multiplicity and spacial distributions of neutrons, minimise the information loss

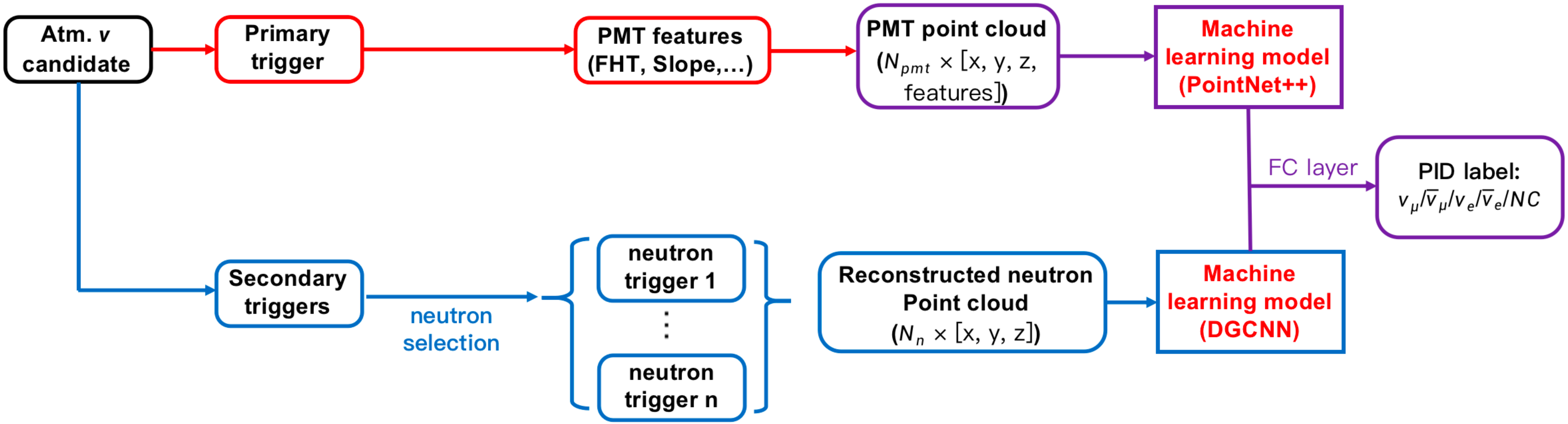


Architecture of PointNet++



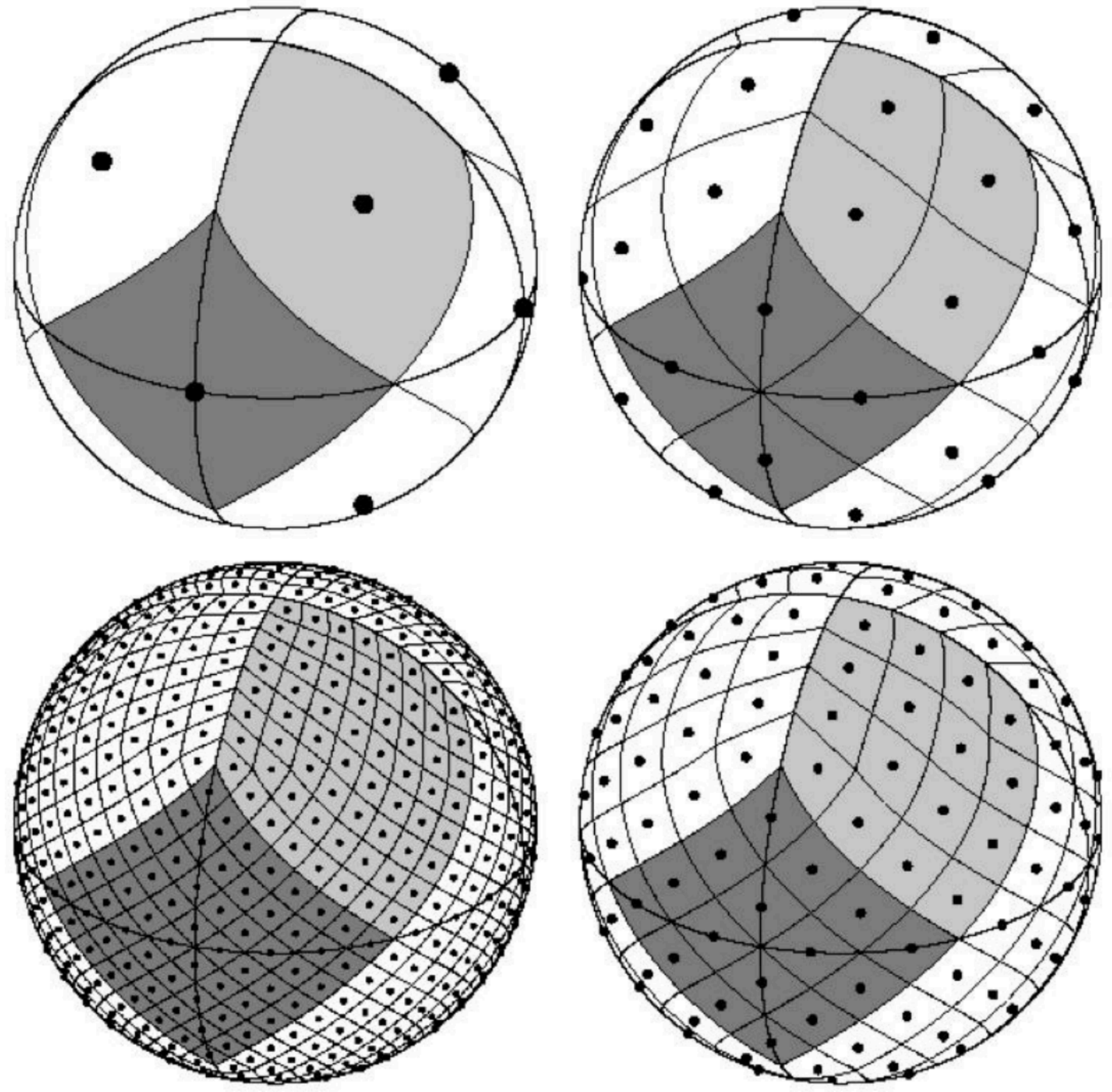
Architecture of DGCNN

Two strategies (1)



- DGCNN is used to extract features from the reconstructed neutron information, concatenate with PointNet++ model with a FC layer for final output

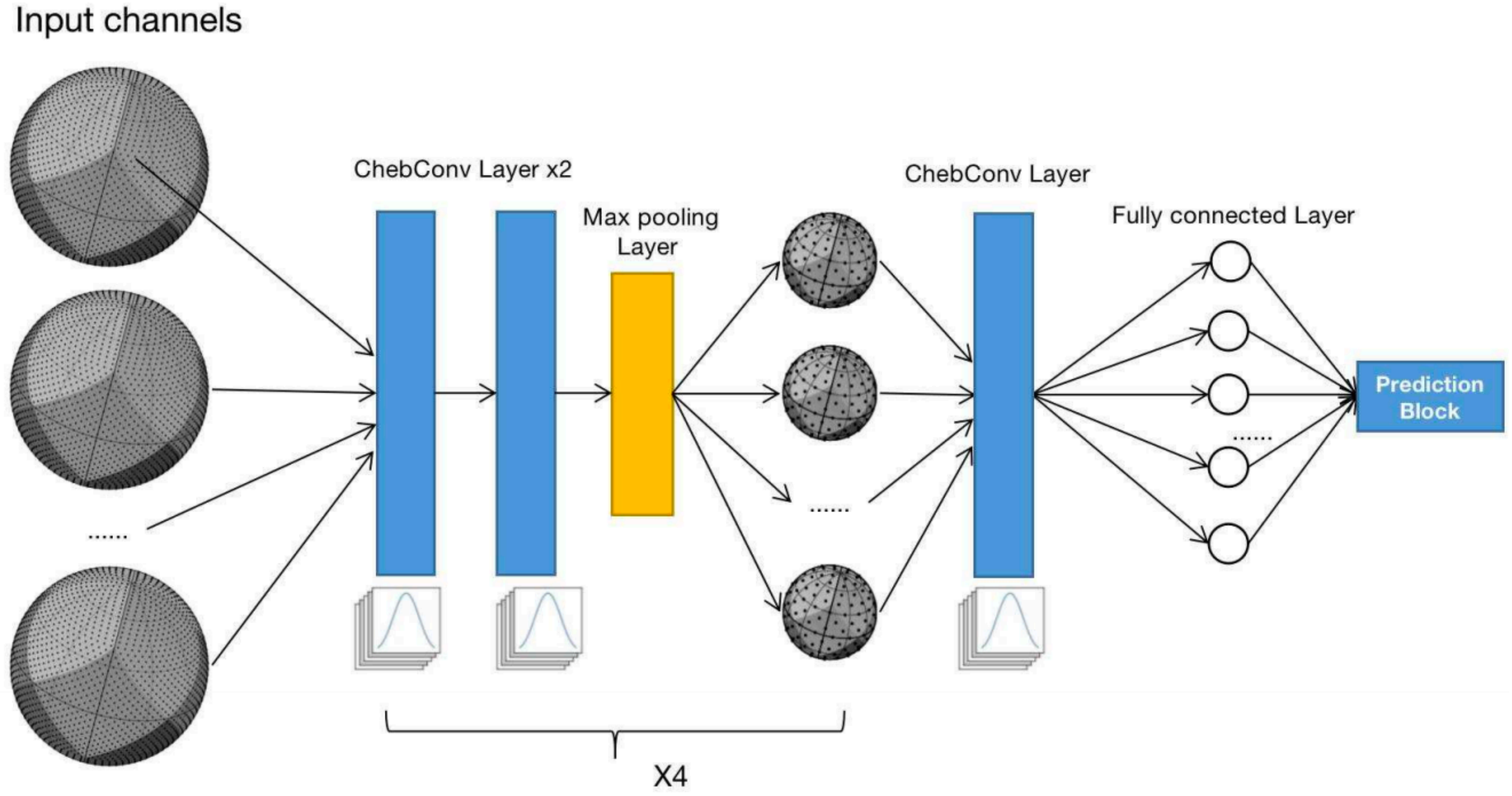
Two strategies (2)



2. Spherical CNN: DeepSphere

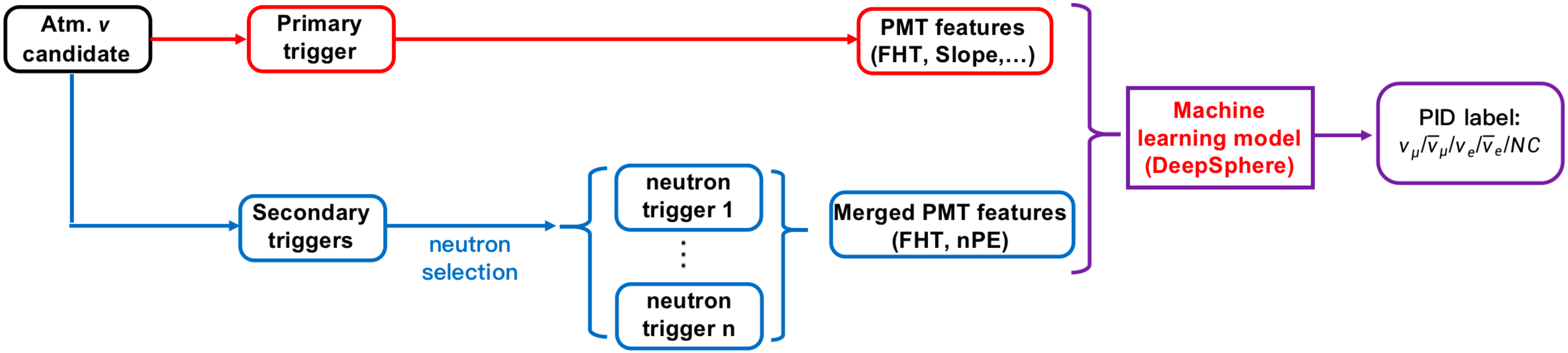
- Graph-CNN: developed for processing spherical data originally developed for cosmology studies
- Maintain rotation covariance, Avoid distortions caused by projection to a planar surface

.

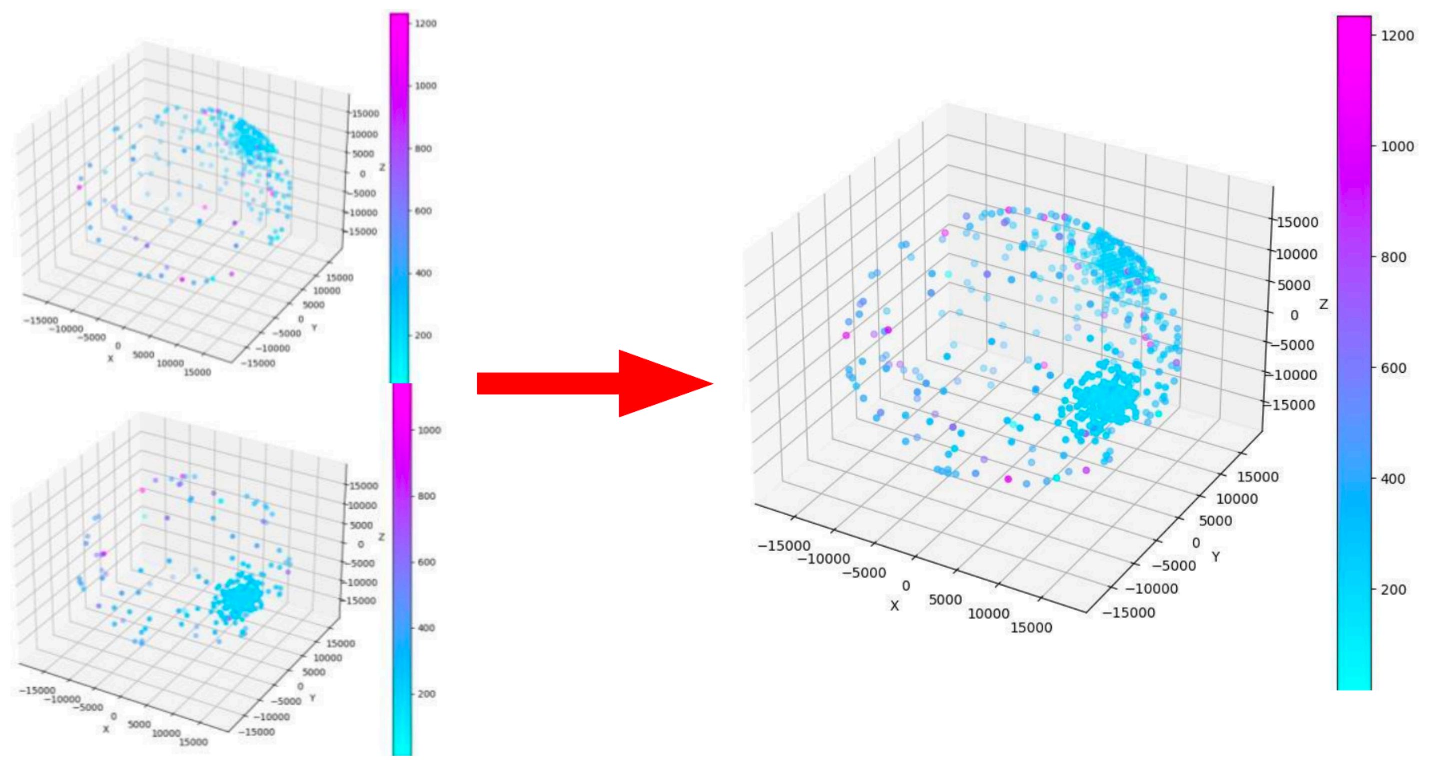


- Use Healpix sampling to define vertices
  - Equally divide the sphere into 12 parts
  - Further divide each part into N_{side} parts ($N_{side} = 2^n$)
  - Chose $N_{side} = 32$ total number of pixels: 12288
  - **If more than one PMTs are in one pixel, info is merged**

Two strategies (2)

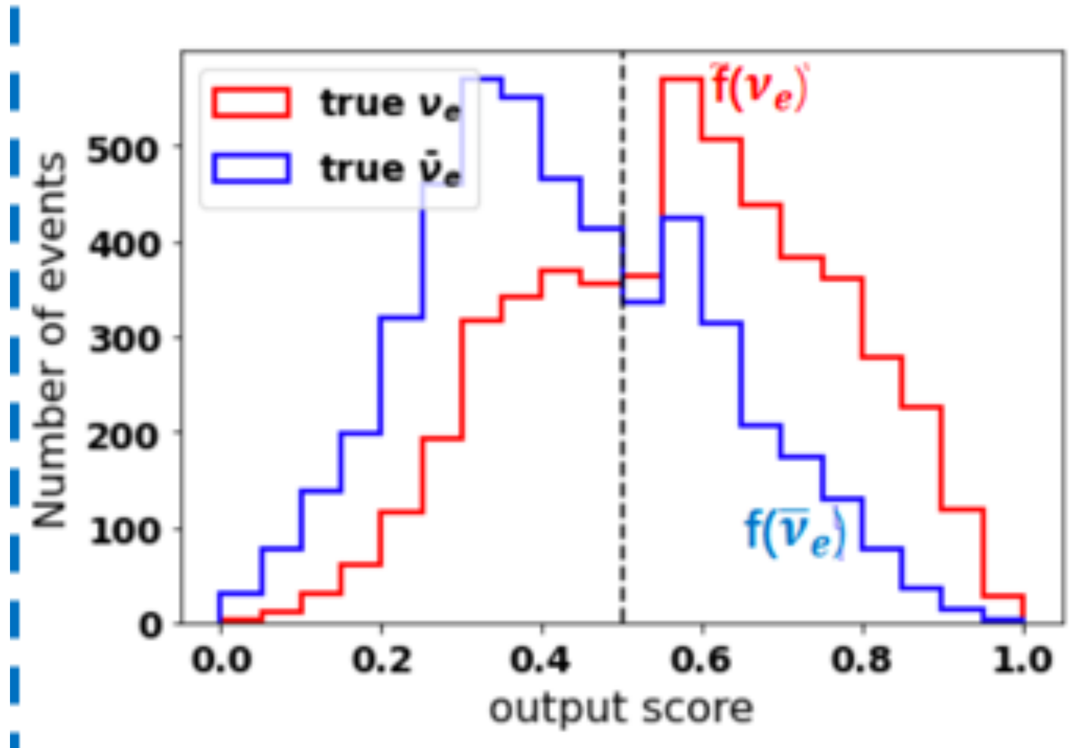


- Multiple neutron-candidate triggers are merged into one
- FHT and nPE are extracted and feed into model together with primary trigger features



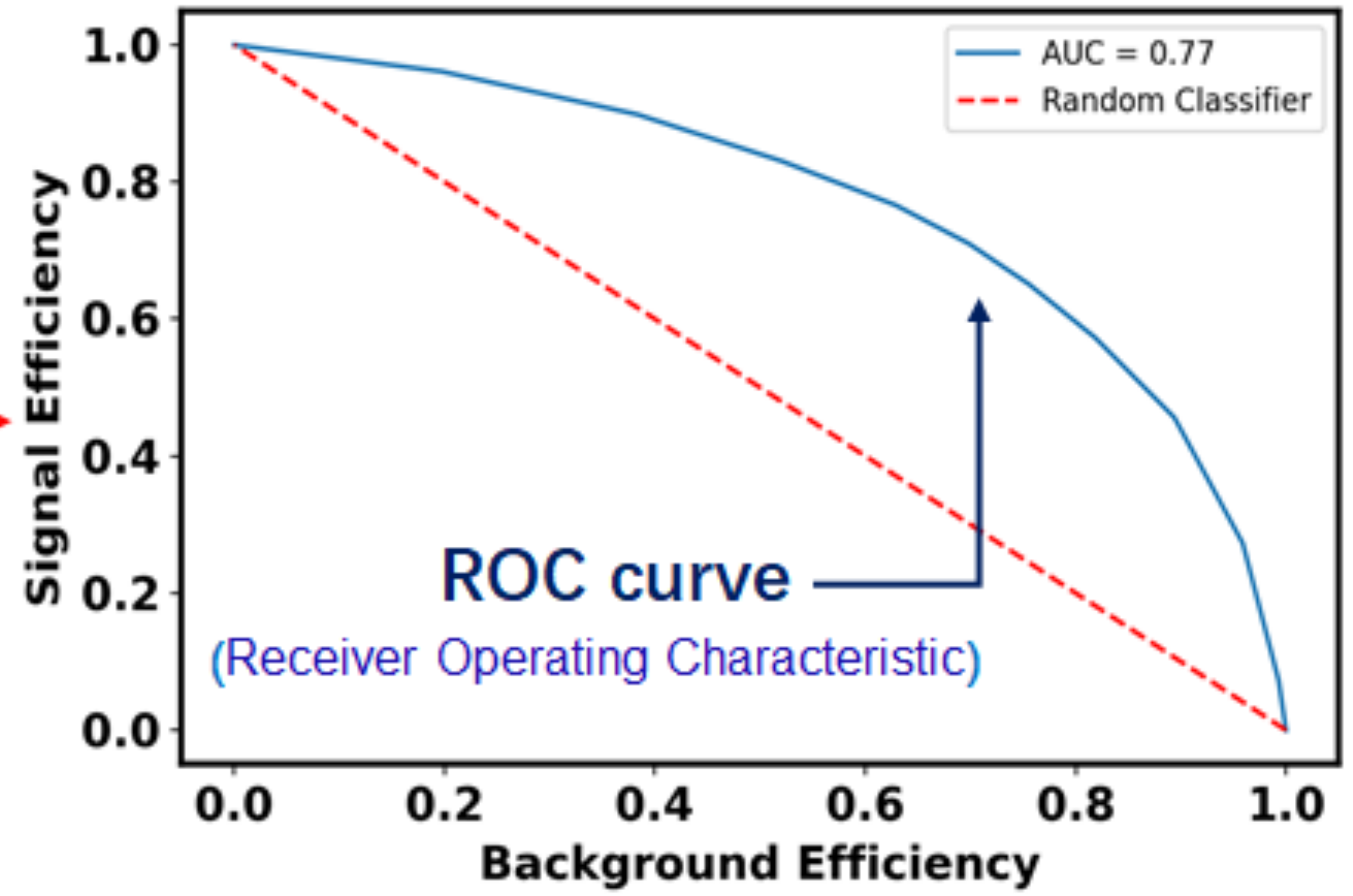
Evaluating model performance

- Training sample consist of ~25k events for all 5 categories considered (ν_μ-CC, $\bar{\nu}_\mu$-CC, ν_e-CC, $\bar{\nu}_e$-CC, NC), with flat neutrino energy spectrum [0, 20] GeV to avoid bias in model training
- Testing sample consist of ~5k events for all 5 categories



Confusion Matrix

Truth \ Pred	ν_e-like	$\bar{\nu}_e$-like
ν_e	A=$\int_{cut}^{1} f(\nu_e)$	B=$\int_{0}^{cut} f(\nu_e)$
$\bar{\nu}_e$	C=$\int_{cut}^{1} f(\bar{\nu}_e)$	D=$\int_{0}^{cut} f(\bar{\nu}_e)$

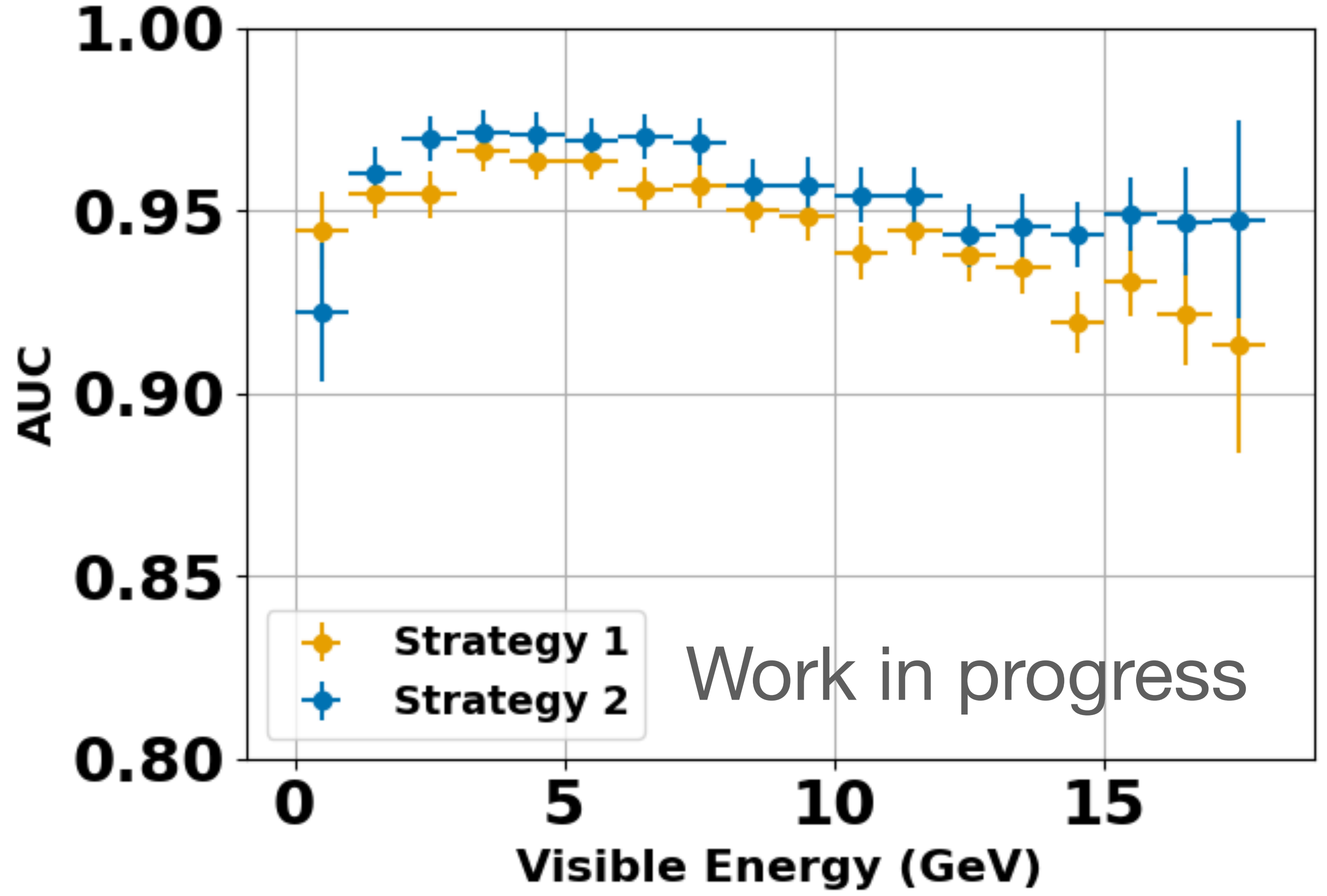


- **AUC ROC** is used to assess models' performances (optimisation of signal efficiency/background efficiency)
  - Does not depend on the choice of score cut
  - Not affected by class-imbalance in the dataset

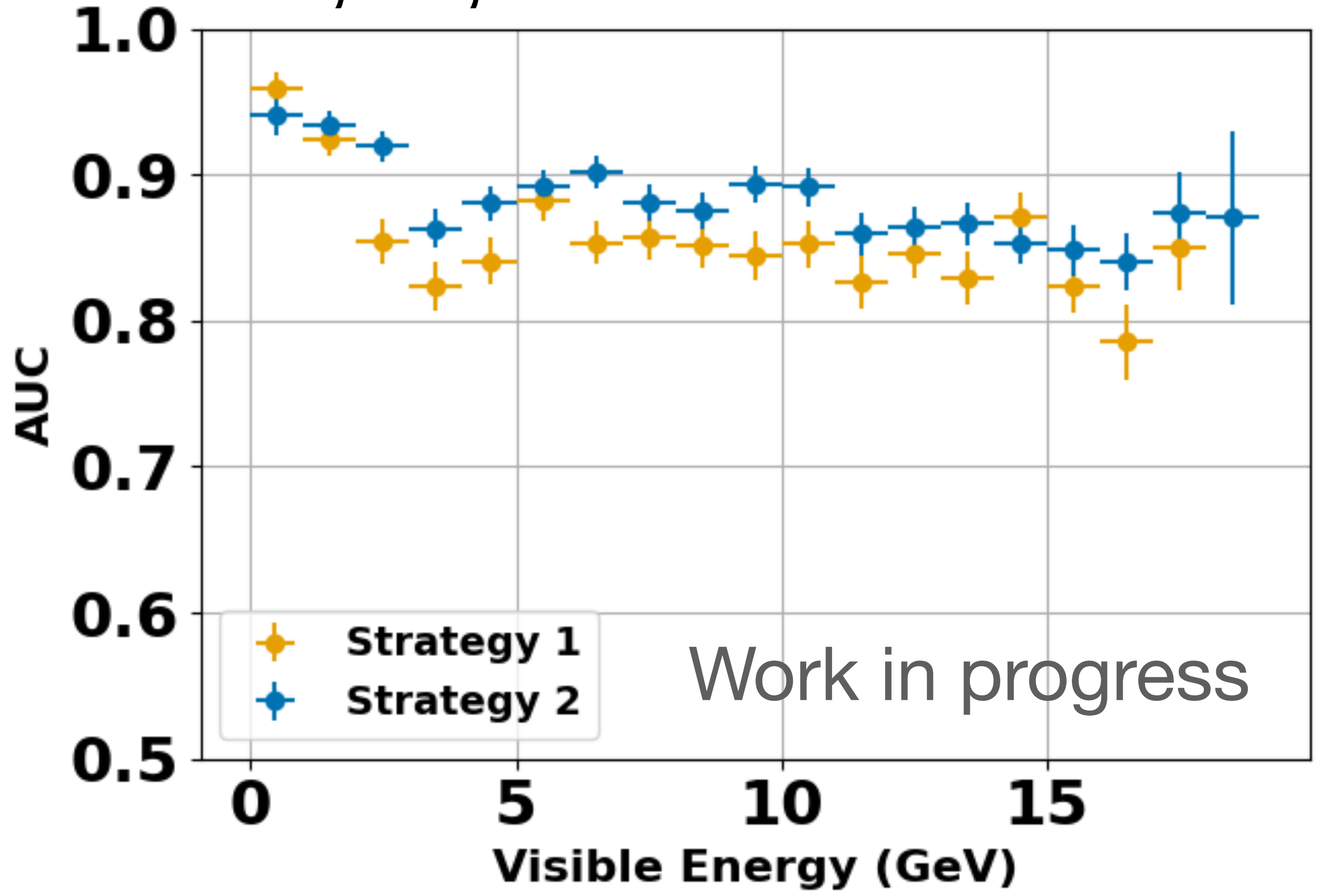
Results

- We observe the AUC scores as a function of visible energy
- Results are consistent between the two strategies for all classification tasks
- For 3-label classification, AUC scores are calculated for each label (“one-vs-rest”) and averaged to get the mean AUC score for each energy bin

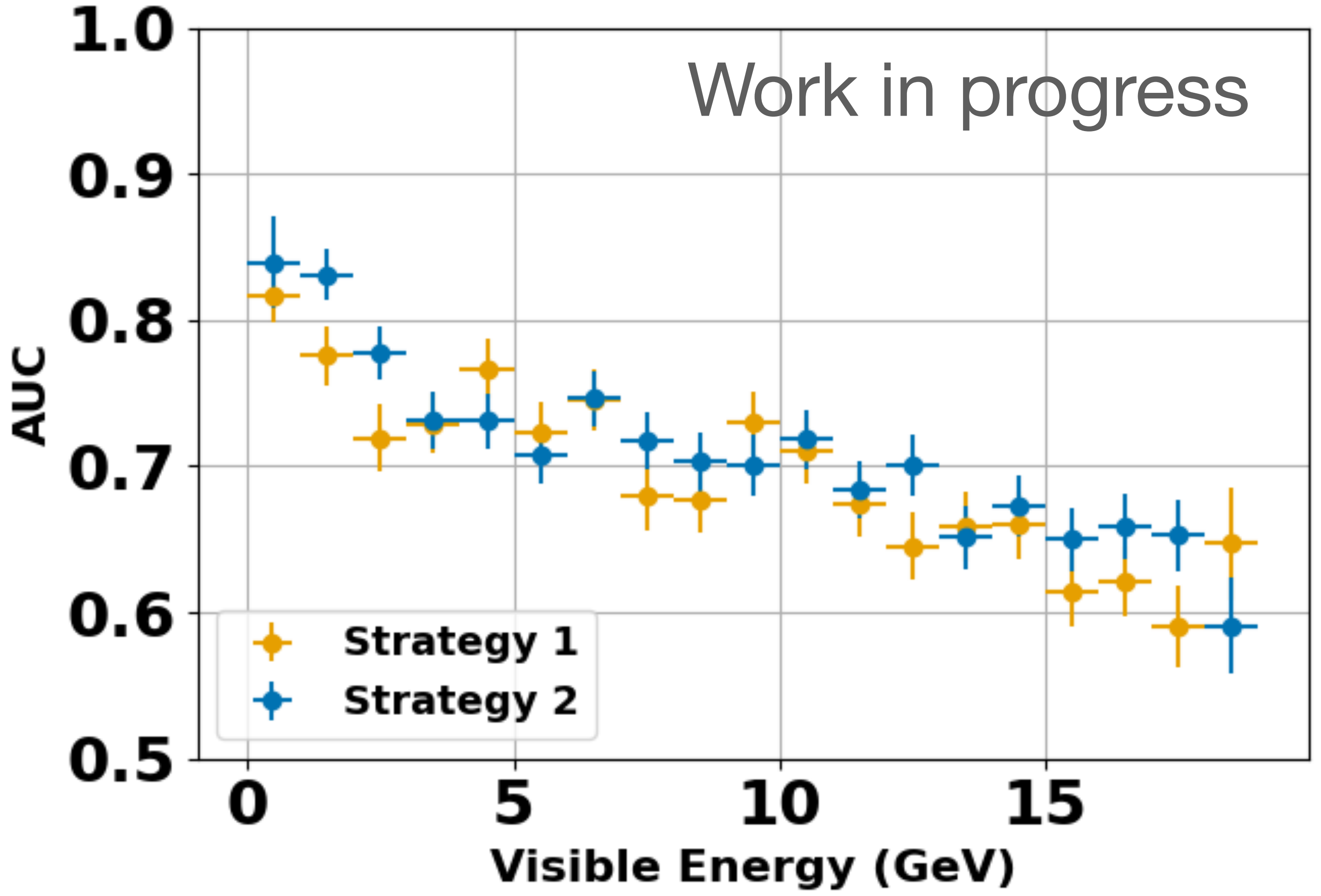
3-label classification



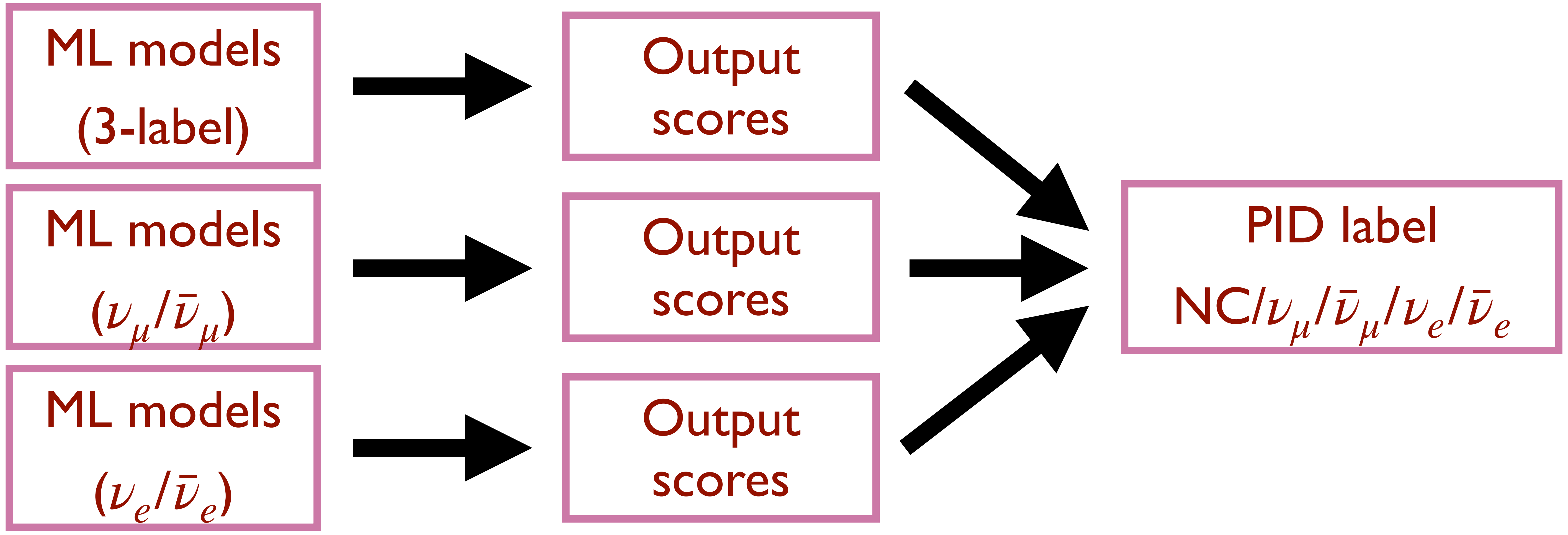
$\nu_\mu/\bar{\nu}_\mu$ classification

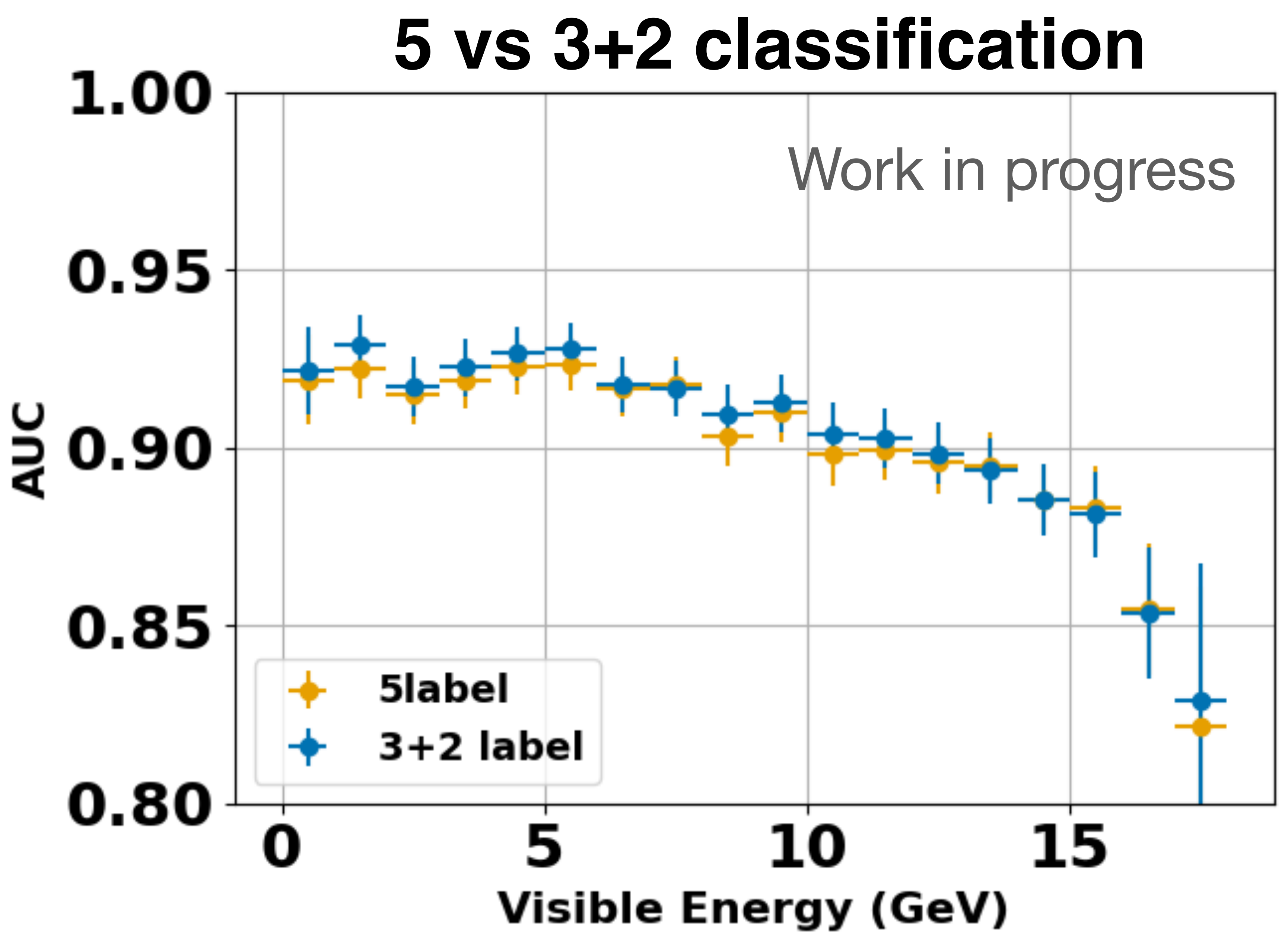


$\nu_e/\bar{\nu}_e$ classification



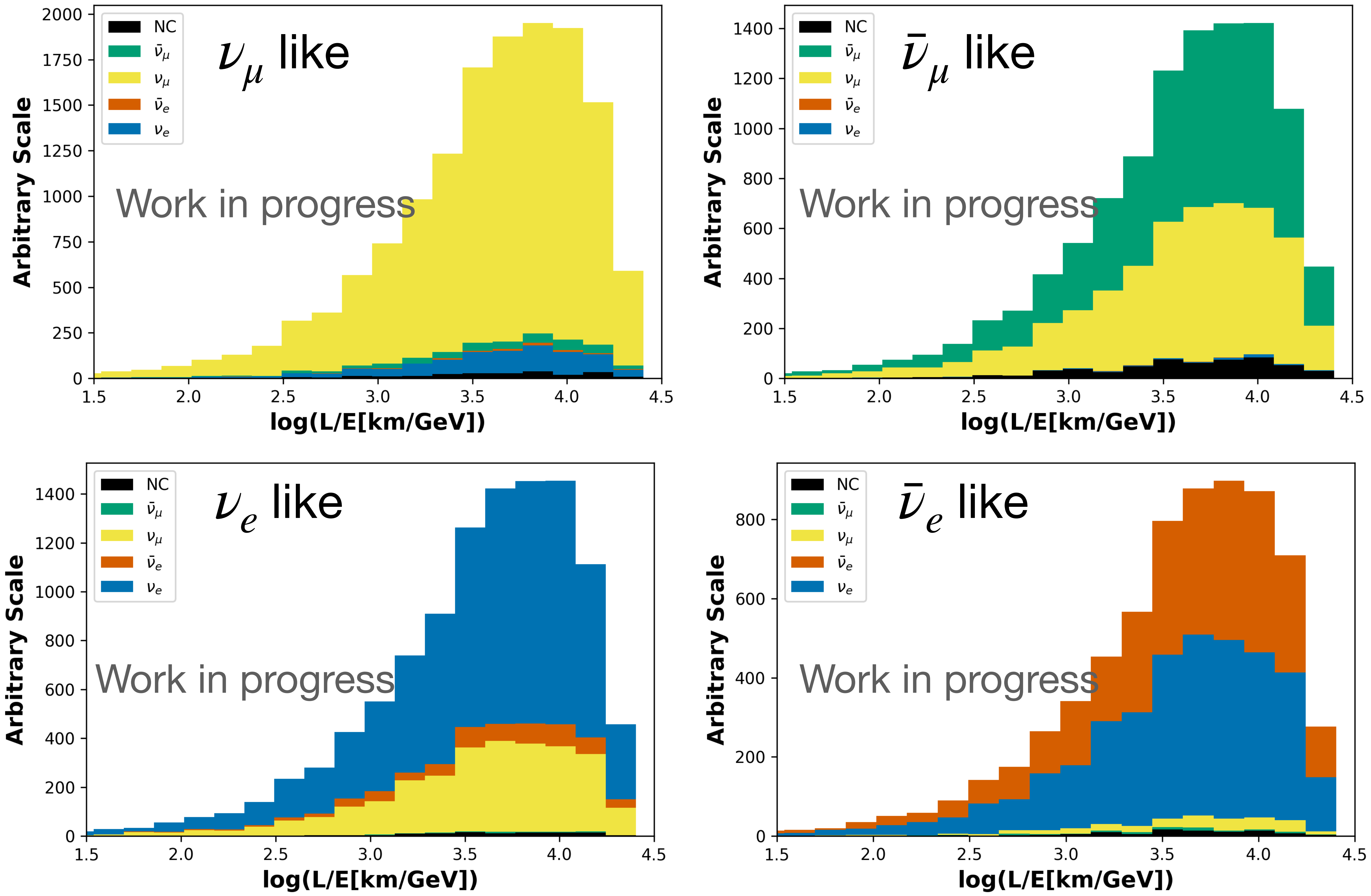
Results





- Combining output scores from each model gives 5-label PID → can be compared with a model which directly performs a 5-label classification
- Agreement suggest that the models considered are capable of directly classifying the 5 categories

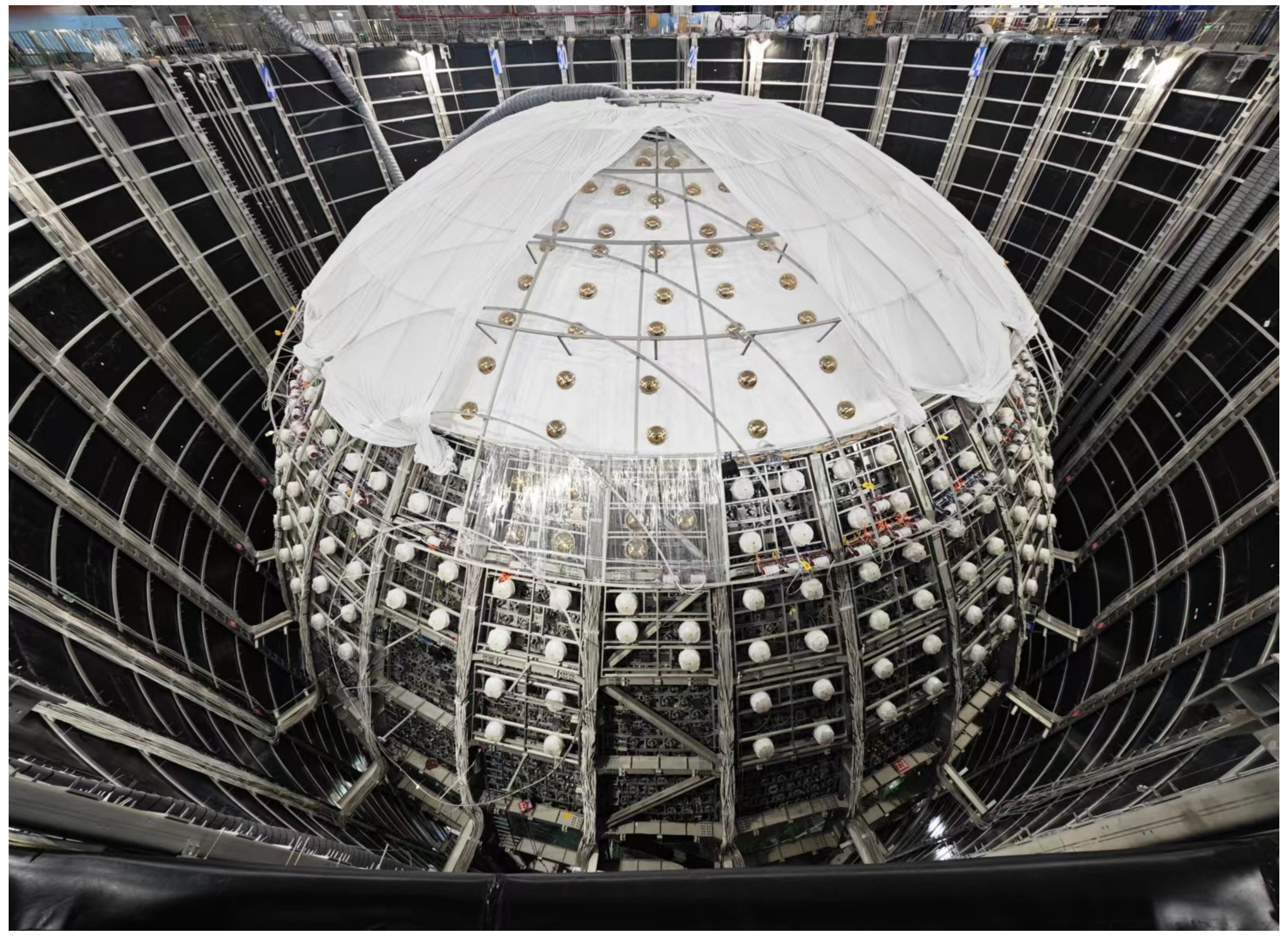
Selected sample



- Efficiencies and purities can be tuned to obtain an optimised sample for NMO analysis

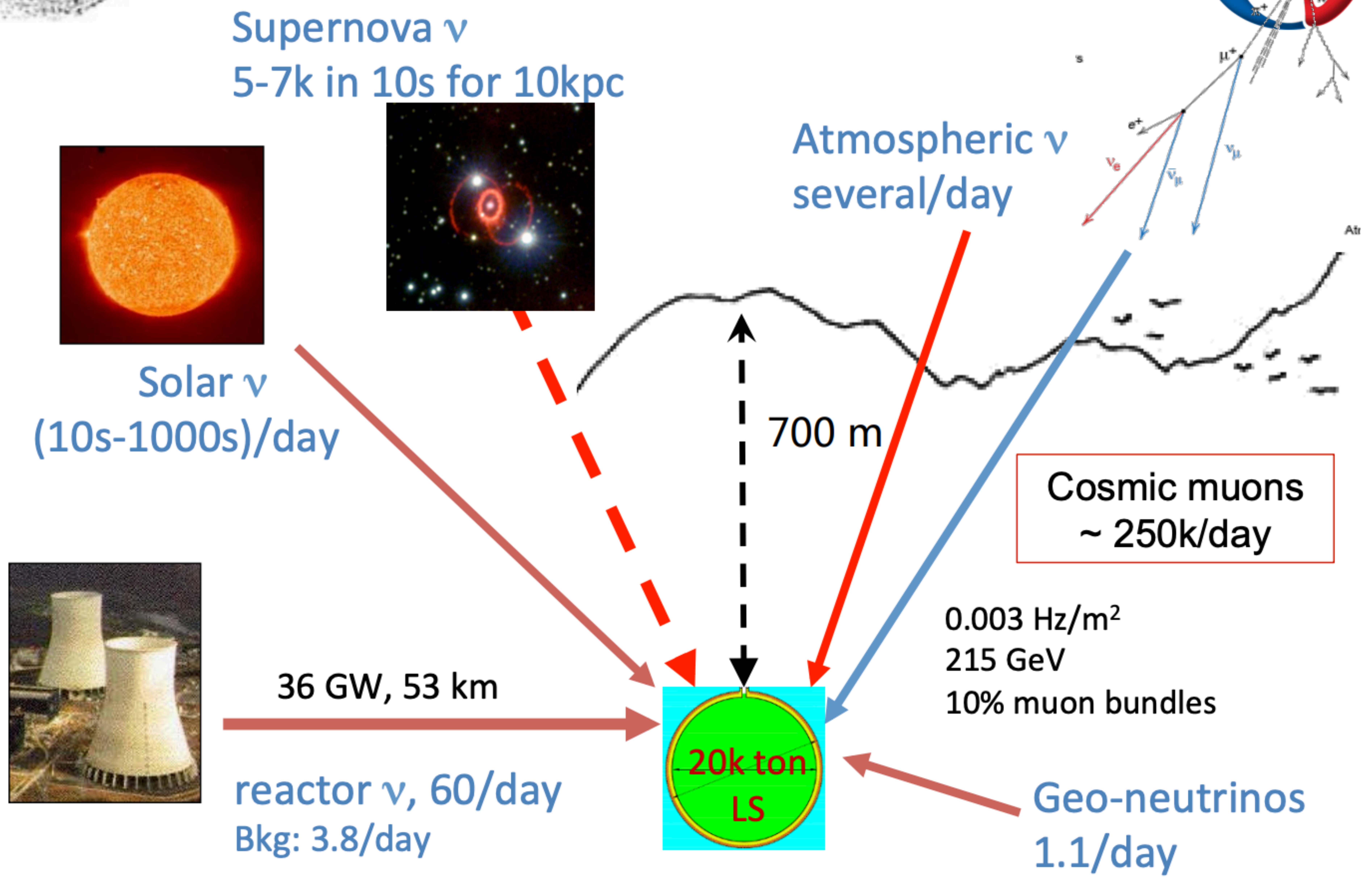
Summary

- A novel method of reconstructing atmospheric neutrino events for LS detector is presented
- Two strategies with different ML models are developed to validate the reconstruction method
- Using JUNO MC samples, variables that are crucial to physics analyses such as **direction, energy, particle identification** of atmospheric neutrinos can be reconstructed with good resolution

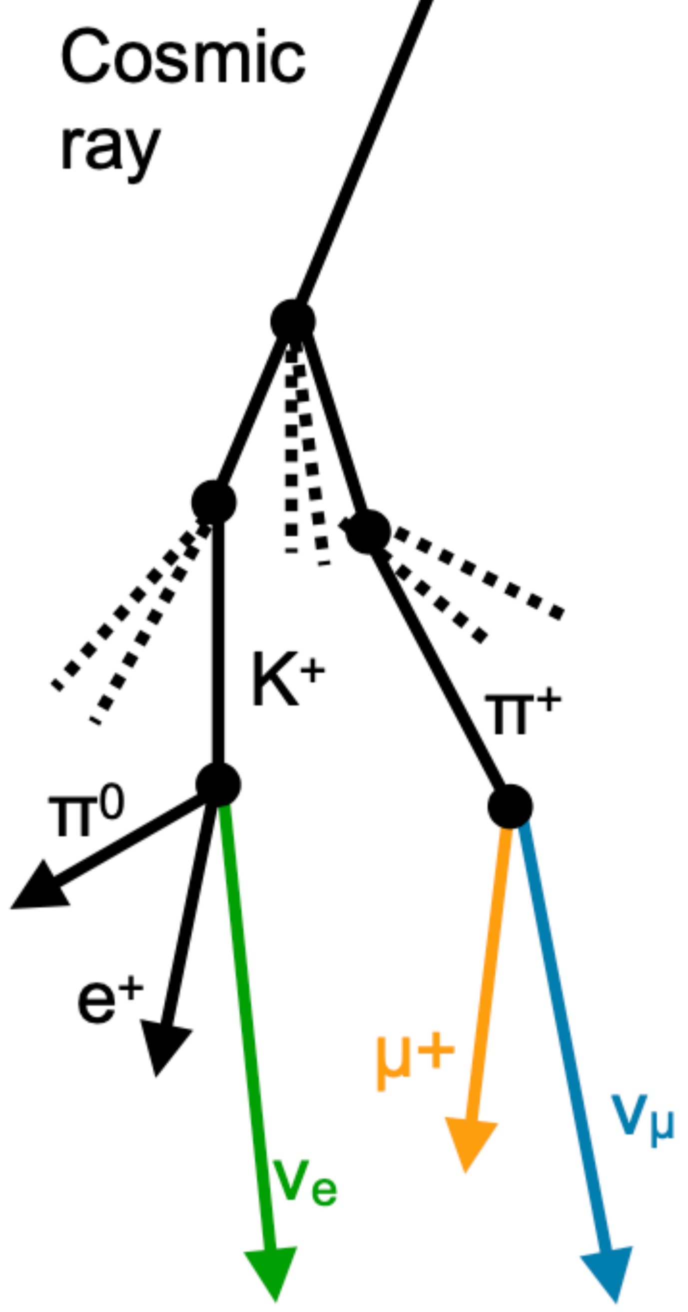


Backup

JUNO Event Rates after selection



Atmospheric Neutrinos



- Large flux of atmospheric neutrinos (ν_{atm}) produced by cosmic ray interactions
- Isotropic with different baseline (L) and energy (E)
- Natural source of neutrinos in GeV region