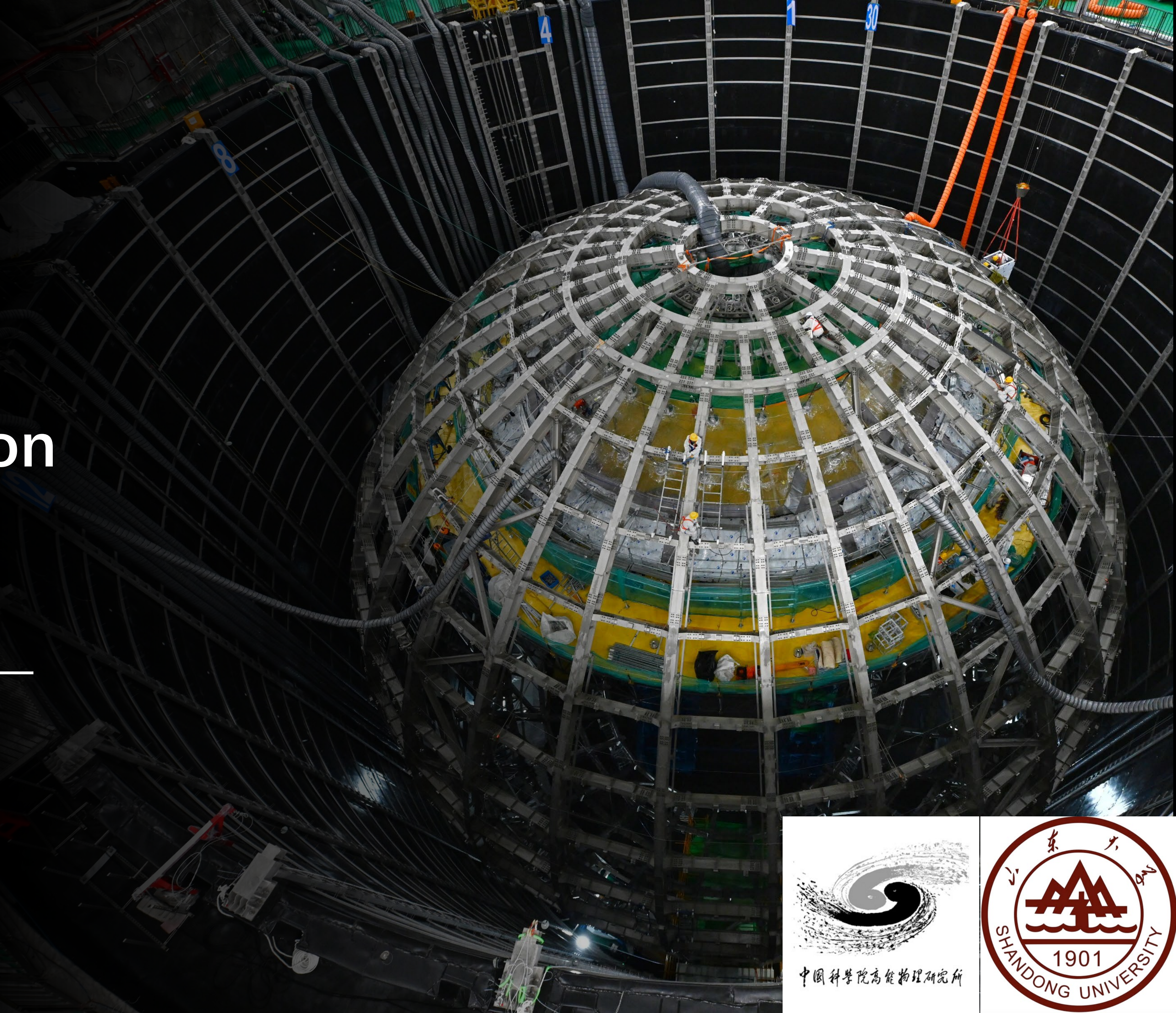
The directionality reconstruction and particle identification for atmospheric neutrinos in JUNO



刘佳熙[1], **刘震**[1], 罗武鸣[1], 张永鹏[1]

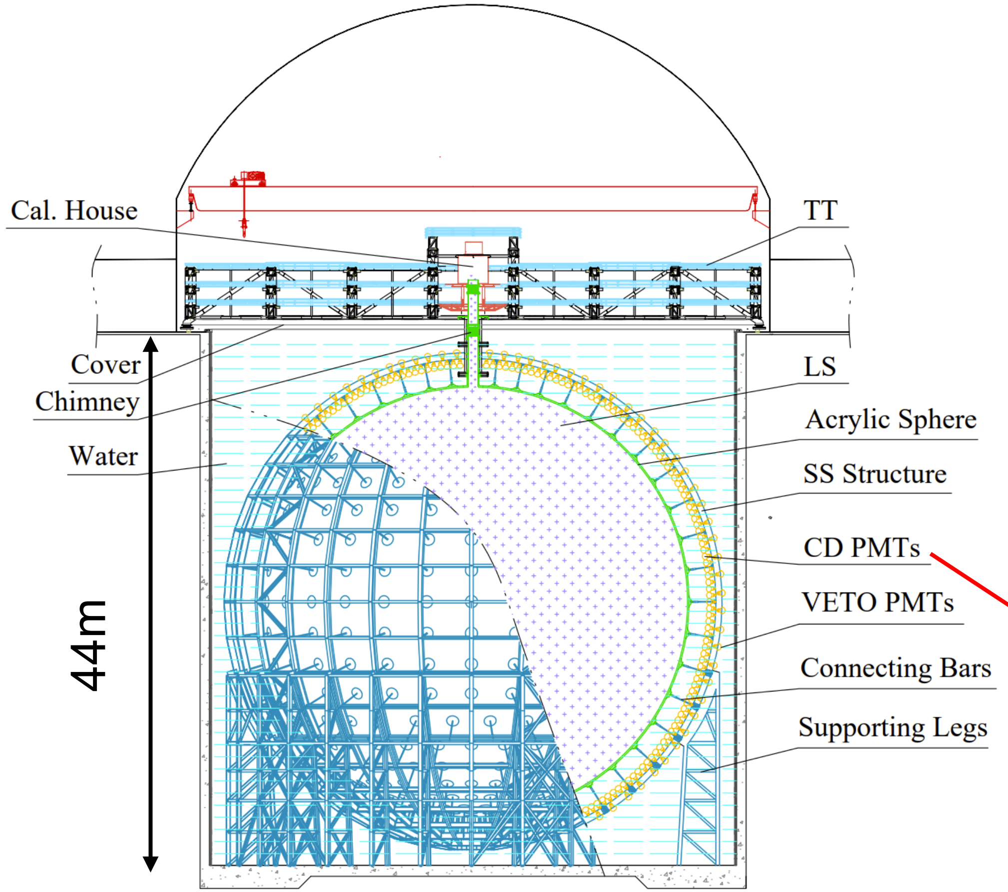
杜杨洪岳[2], 李腾[2], 杨泽坤[2]

[1]高能物理研究所（IHEP）

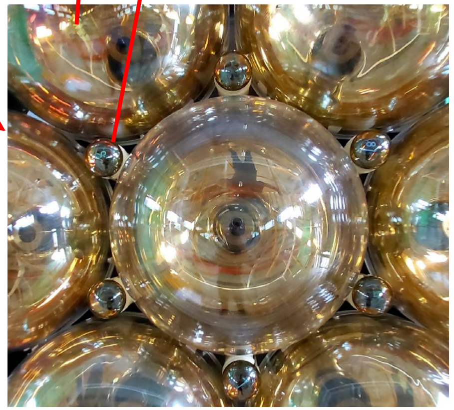
[2]山东大学

杭州, 2023年5月

JUNO Experiment



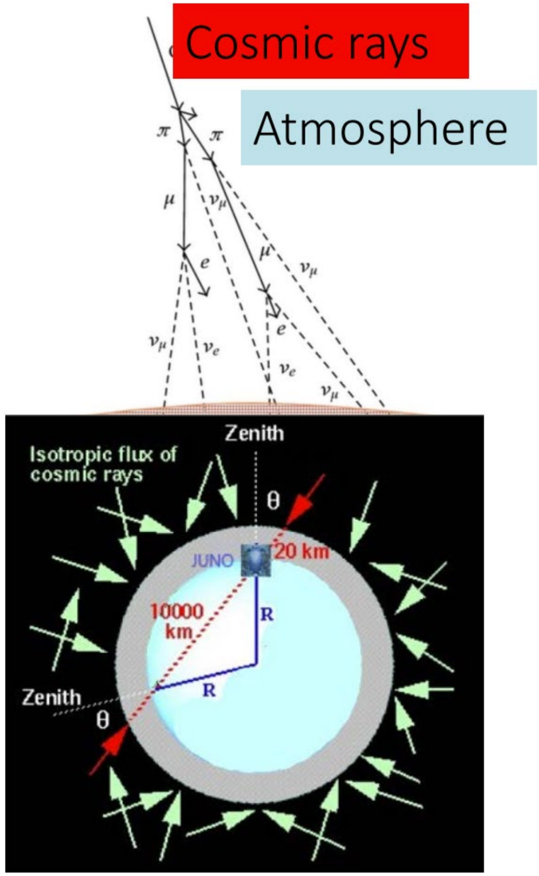
- JUNO is a next-generation large liquid-scintillator detector
- The Central Detector is instrumented with
  - 17’612 20-inch Large-PMTs (LPMTs)
  - 25’600 3-inch Small-PMTs (SPMTs)



- The PMT system has large optical coverage (>75%)
- It acts like a “video camera”, which captures the temporal evolution signals originated from incident particles.

Motivation

JUNO's main physics goal is the determination of neutrino mass ordering.



Major flavors:
$\nu_\mu, \overline{\nu}_\mu, \nu_e, \overline{\nu}_e$

To enhance its sensitivity to the mass ordering, JUNO will combine the measurements of

- reactor anti-neutrinos at low energies
- **atmospheric neutrinos at high energies (GeV level)**

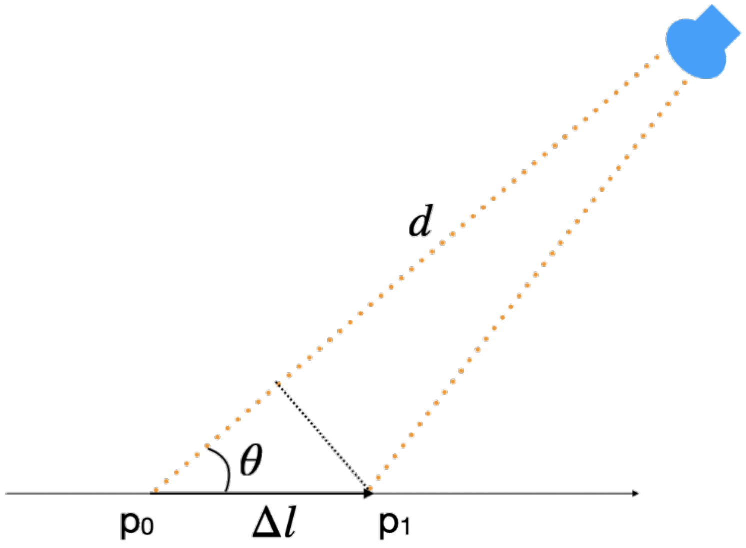
Event reconstruction and particle identification for atmospheric neutrinos are challenging works for liquid scintillator detectors like JUNO:

- particle trajectory is not directly visible
- Cherenkov light yield is negligible
- interactions within the detector are complex

For the first time ever, machine learning methods have made these tasks achievable.

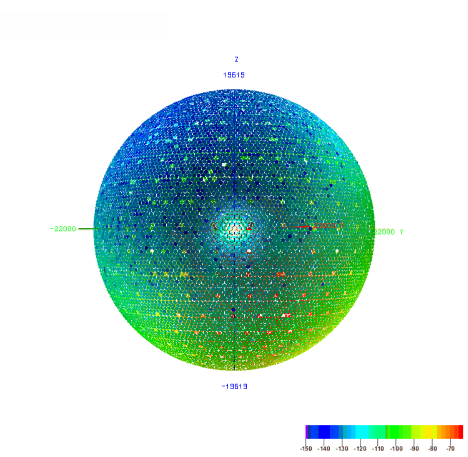
Methodology for the directionality reconstruction and particle identification of atmospheric neutrinos

- The scintillation light received by a PMT is the superposition of light from many points on particle tracks inside the detector.
- The directional information of the incident $\boldsymbol{v}$ and topological information characterizing event interaction types are reflected in the PMT waveforms.
- Features extracted from waveforms will be used as inputs to machine learning models.

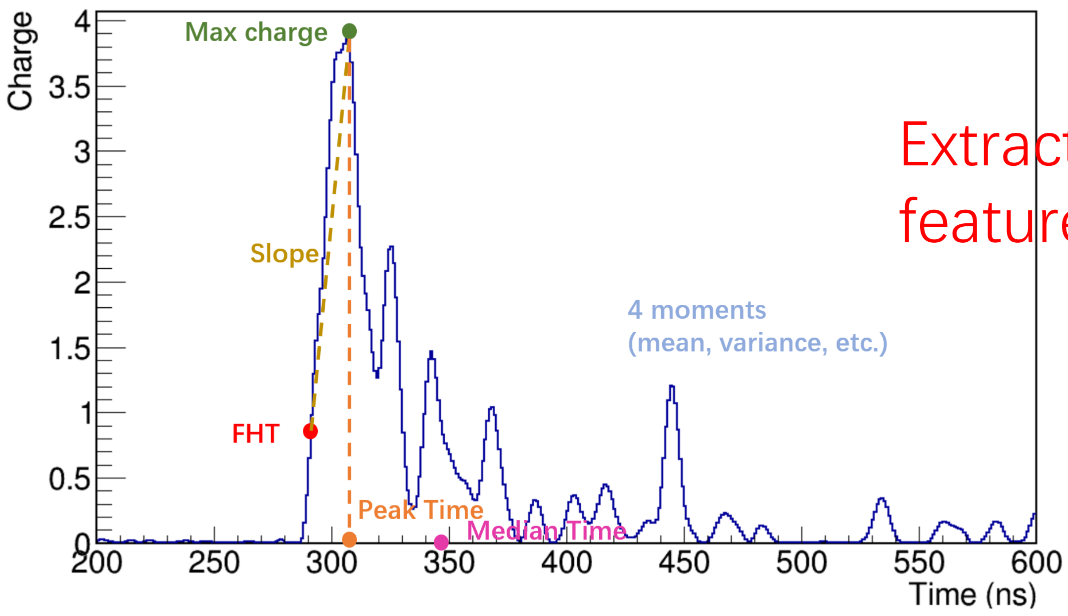


$$\frac{\mathrm{d}l}{\mathrm{d}t} \propto \frac{v}{|1-n\beta\cos\theta|}$$

PMT waveform feature extraction



Animated schematic of the time evolution of PMT signals in JUNO central detector (produced by a 1.67GeV atmospheric υ_μ event)



Waveform signal from one of the PMTs (first readout, with waveform reconstruction: deconvolution and noise removing)

Extracted feature

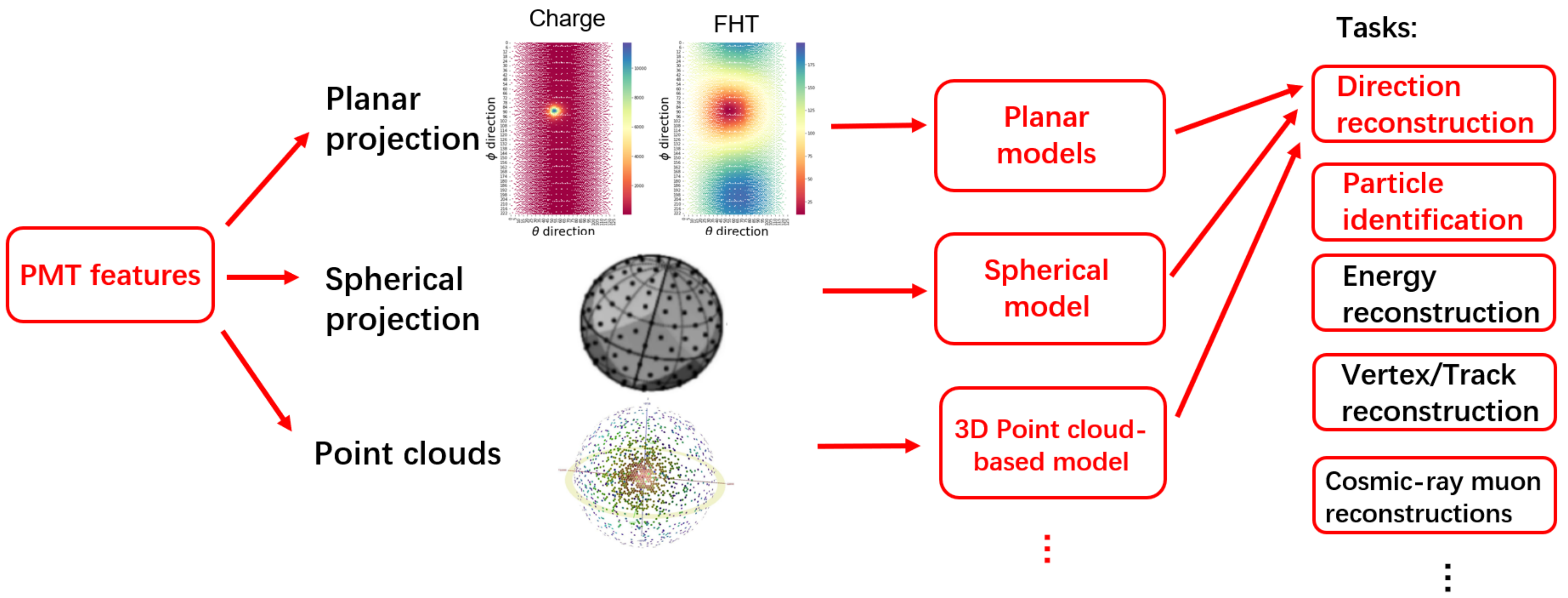
- FHT: First hit time.
- Total charge: The total charge in the first readout time window.
- Slope: max charge divided by peak time.
- Charge ratio: Charge in the first 4ns divided by the total.
- Max charge, Peak Time

Features after importance check

Machine learning models

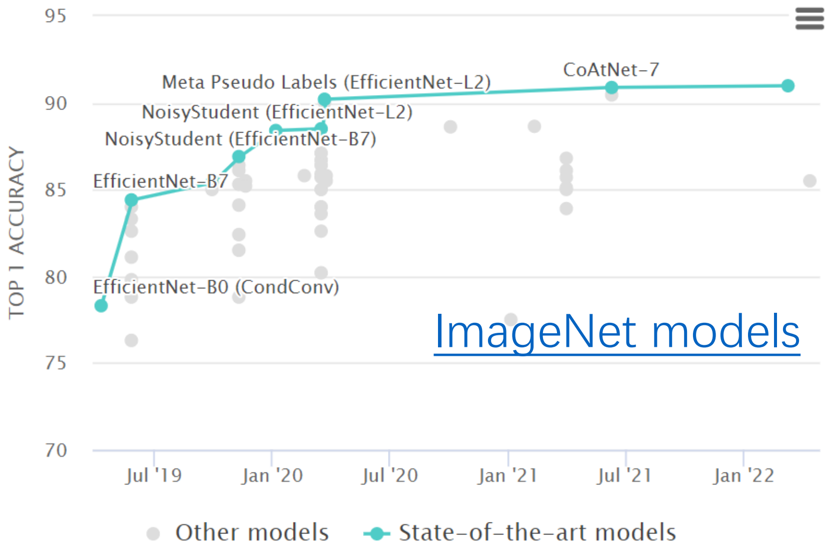
3 categories of machine learning methods:

- Planar-image-based method: EfficientNetV2
- Spherical-image-based method: DeepSphere
- 3D-based method: PointNet++



Introduction to models: EfficientNetV2

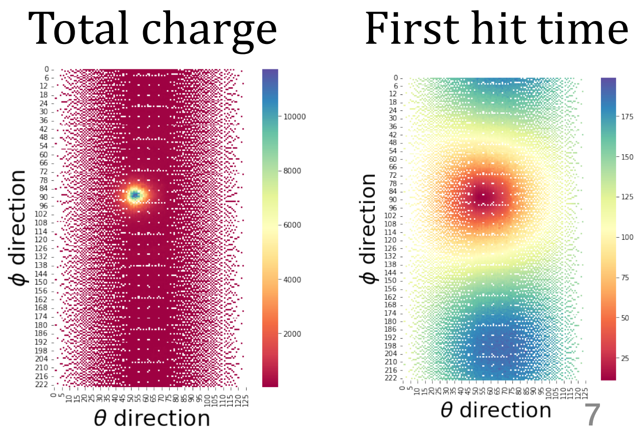
- **The state-of-the-art performance among CNNs;**
- **Smaller model size and fast training.**



EfficientNetV2: arXiv:2104.00298

Model input: 2D grid

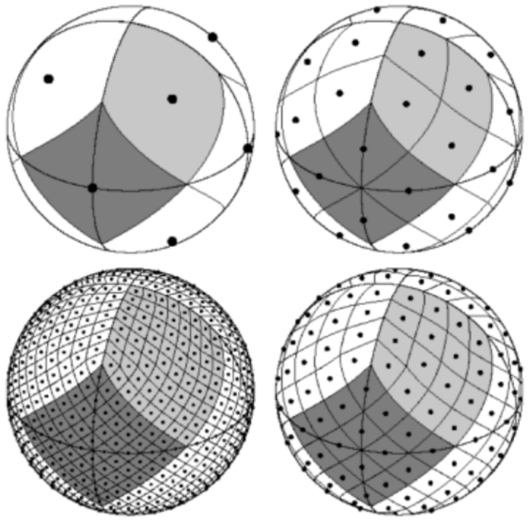
- The PMT map is projected onto a 2D θ−ϕ grid (according to PMT spherical coordinates);
- The grid size of 128 × 224 for LPMT and 256×256 for LPMT+SPMT is chosen to ensure each grid cell corresponds to at most one PMT.

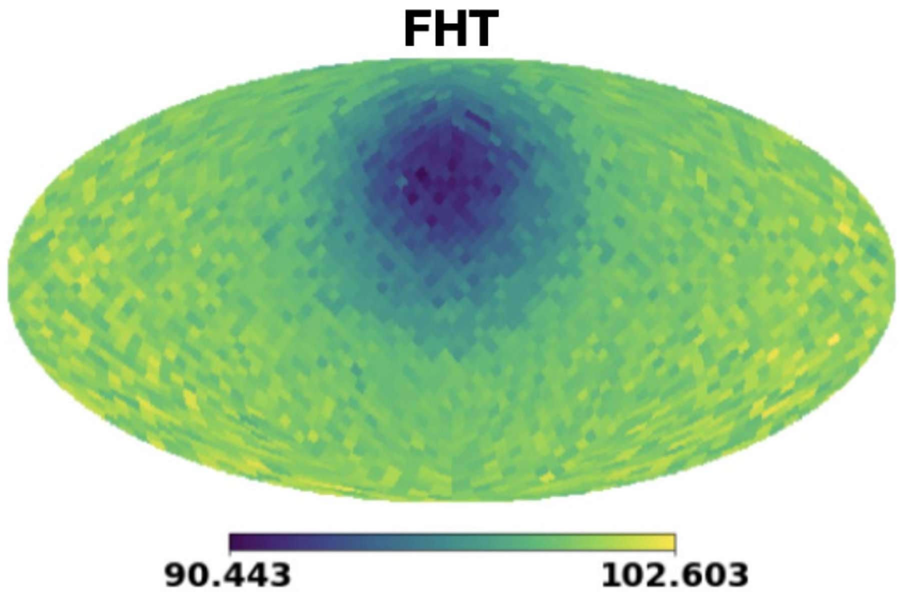


Introduction to models: DeepSphere

DeepSphere: a popular tool processing spherical data originally developed for cosmology studies.

- Maintain rotation covariance;
- Avoid distortions caused by projection to a planar surface.





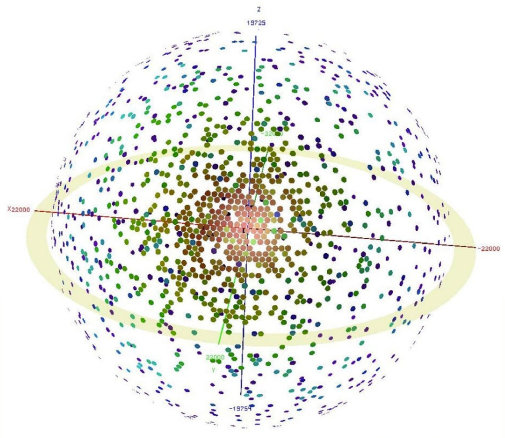
Chose N_{side}=16 (3072 pixels)

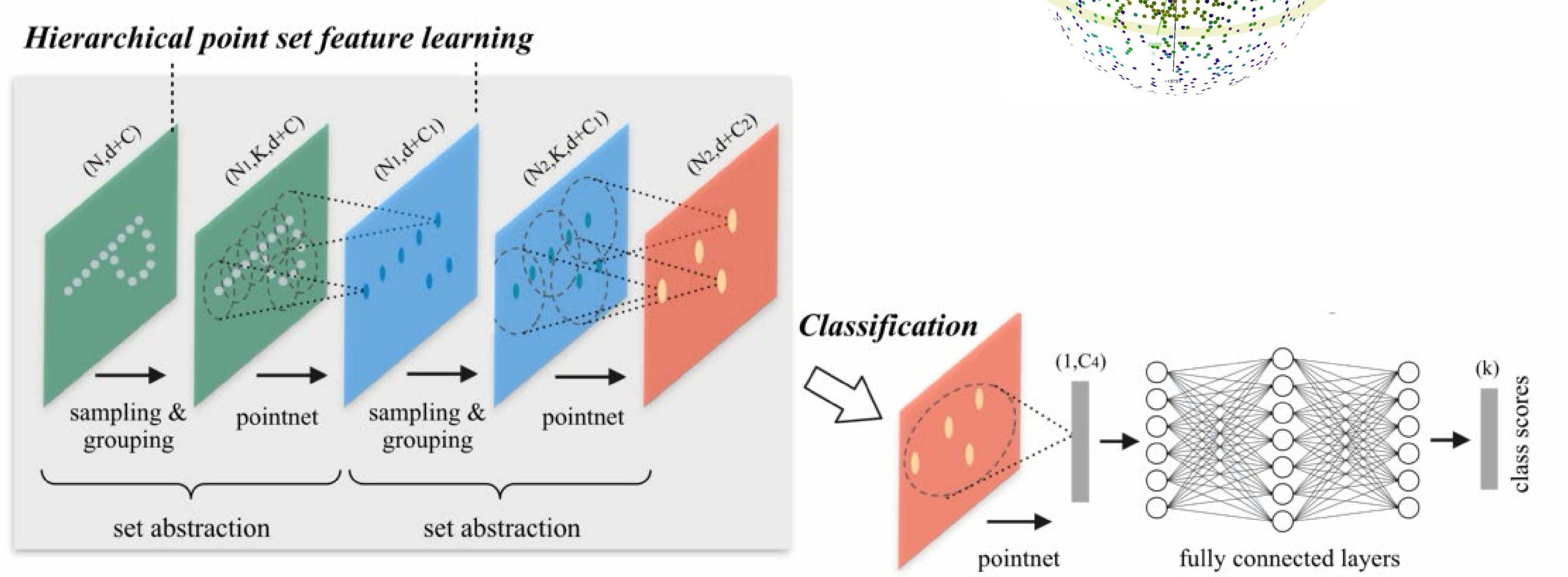
- 3072 pixels for JUNO signal
- If more than one PMTs are grouped into one pixel, information is merged:
  - FHT: the earliest;
  - charge: the sum;
  - Slope and other: the average.

Introduction to models: PointNet++

Directly taking 3D point clouds as input

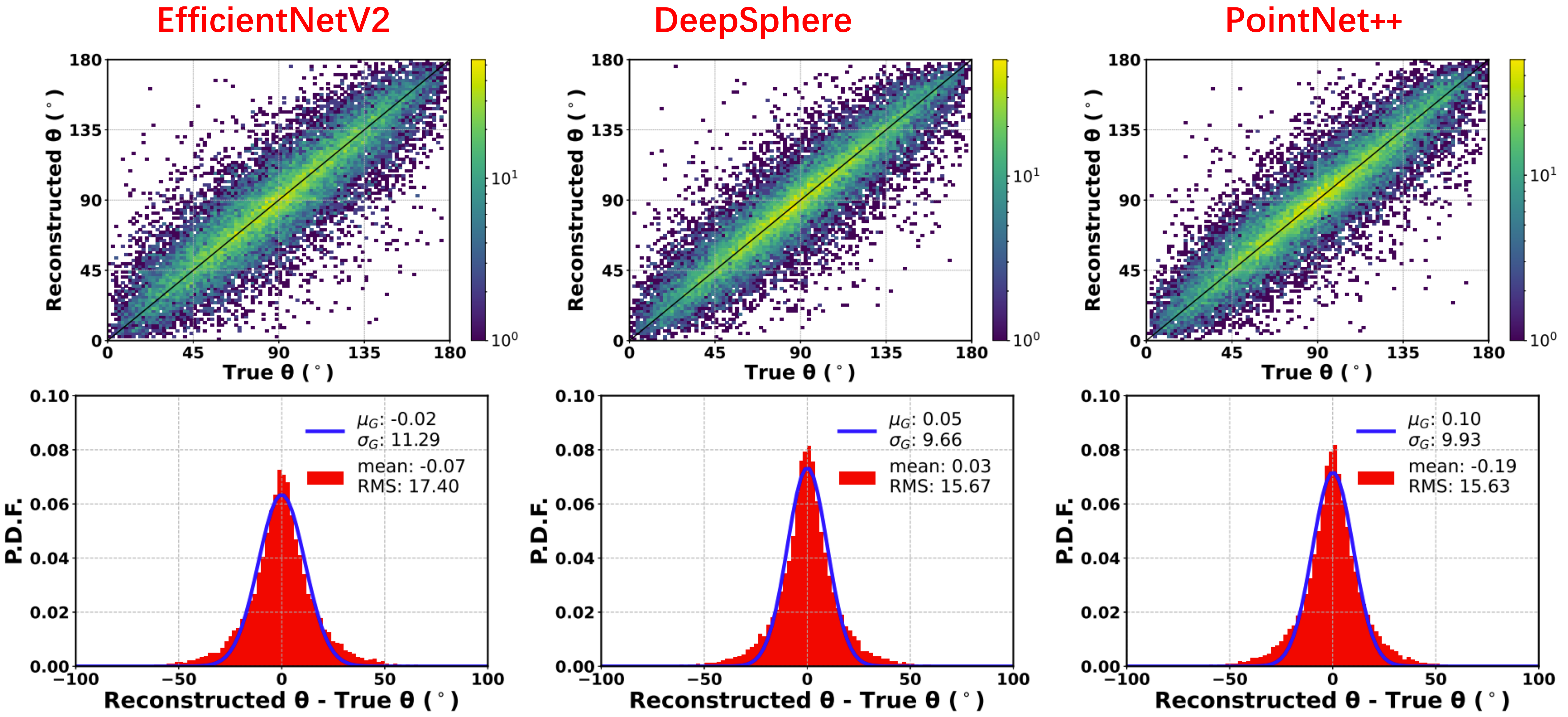
→ JUNO signal more resembles point clouds.





PointNet++ (taken from [C. Qi et al.(2017).])

ν_μ zenith angle (θ) reconstruction results

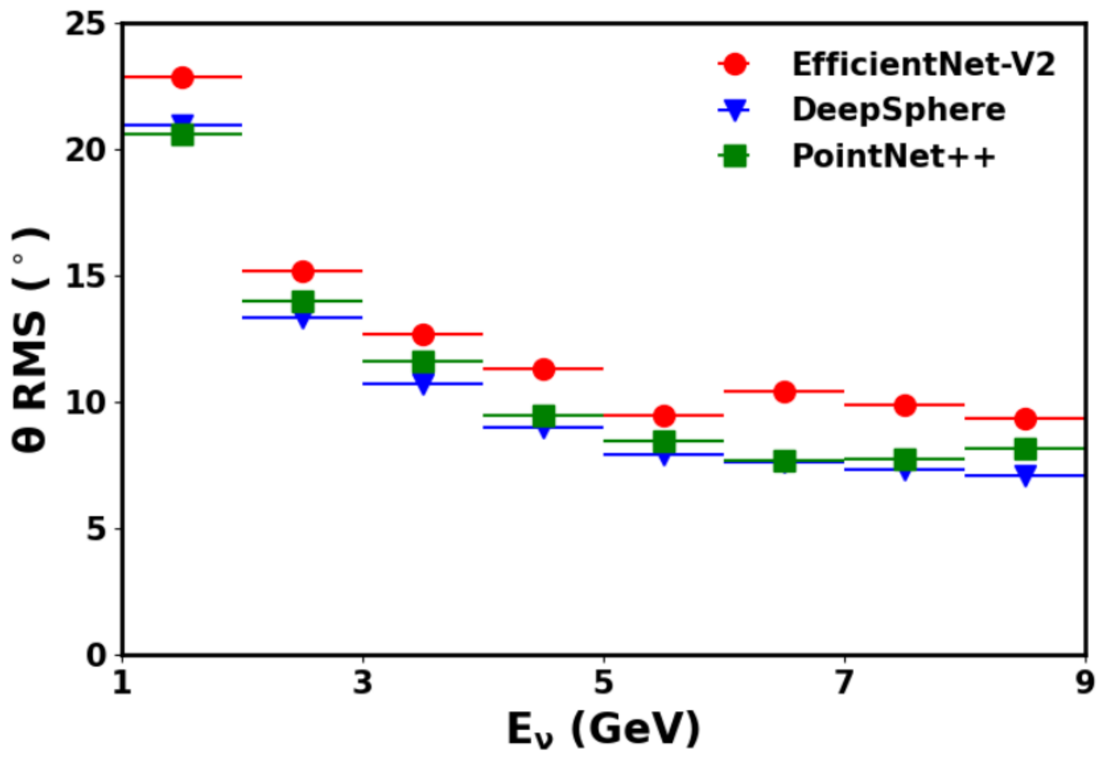


(Unpublished results)

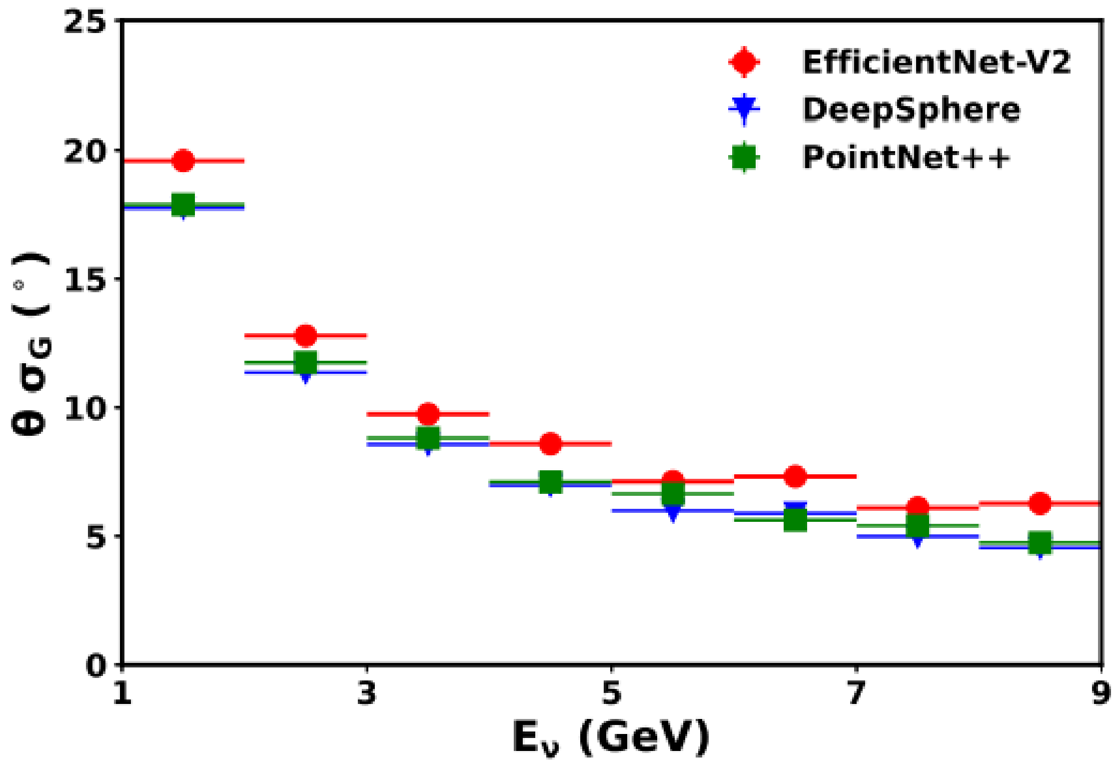
υ_μ θ reconstruction performance: comparison between different models

Resolution improves with the increasing of energy.

RMS



σ_G

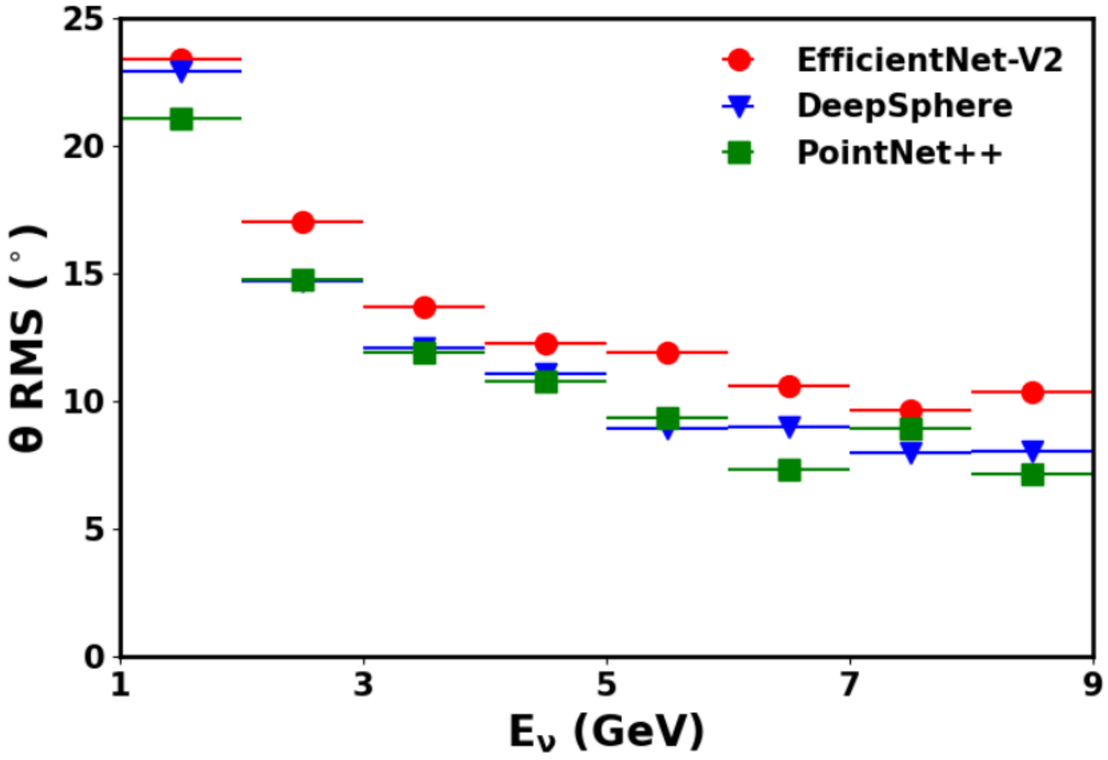


(Unpublished results)

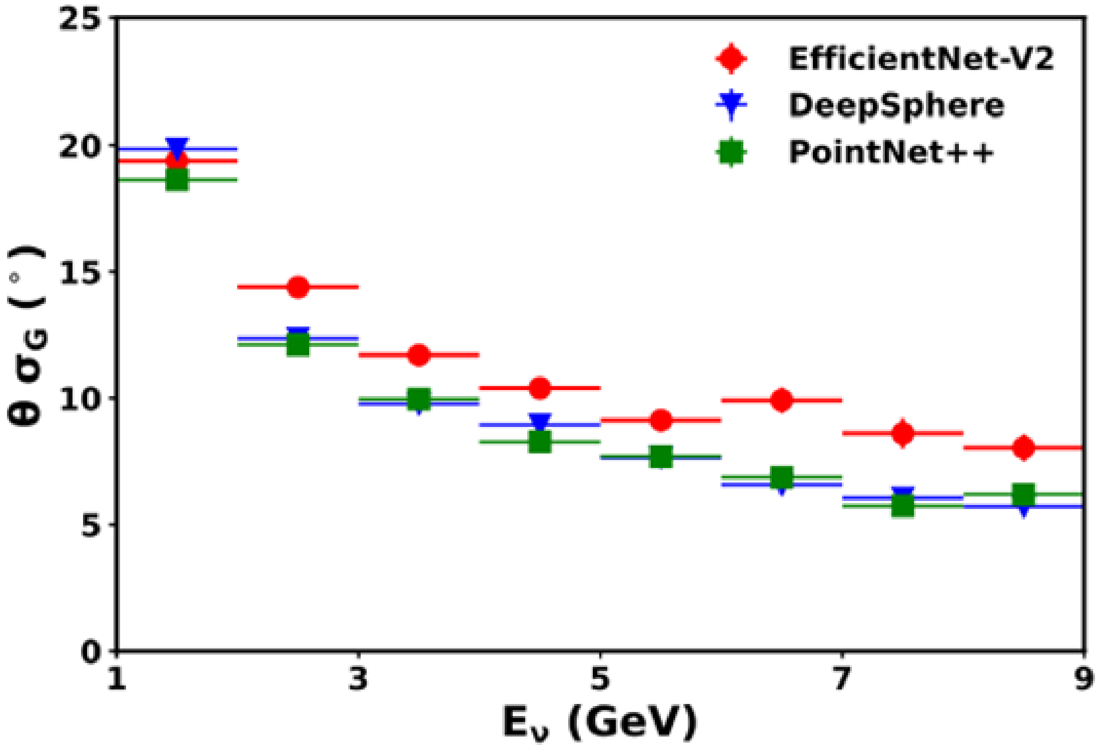
υ_e θ reconstruction performance: comparison between different models

Resolution improves with the increasing of energy.

RMS



σ_G

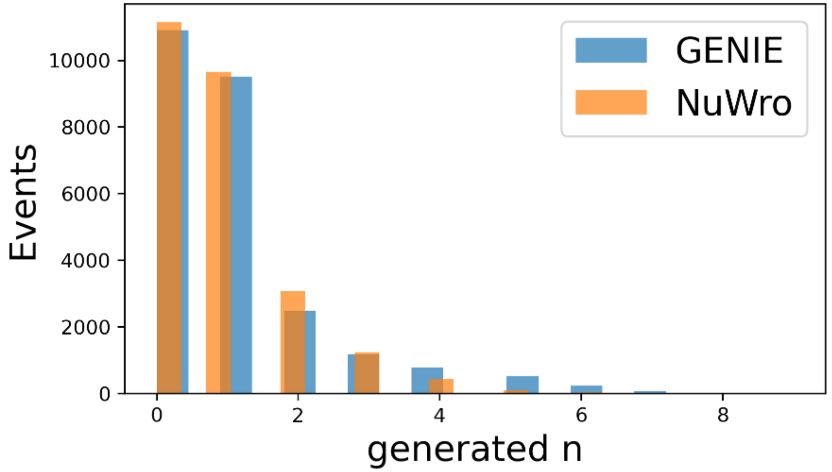


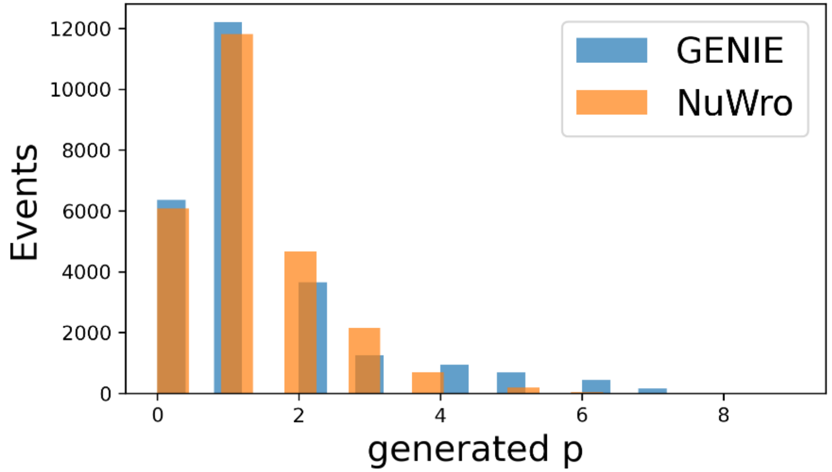
(Unpublished results)

Validation with different event generators

The models were trained with GENIE sample. To check model robustness and estimate systematic uncertainties, different generators (GENIE and NuWro) are used for validation.

Number of neutron or proton generated at neutrino interaction vertices for GENIE and NuWro event generators (from MC truth, for ν_μ CC events):

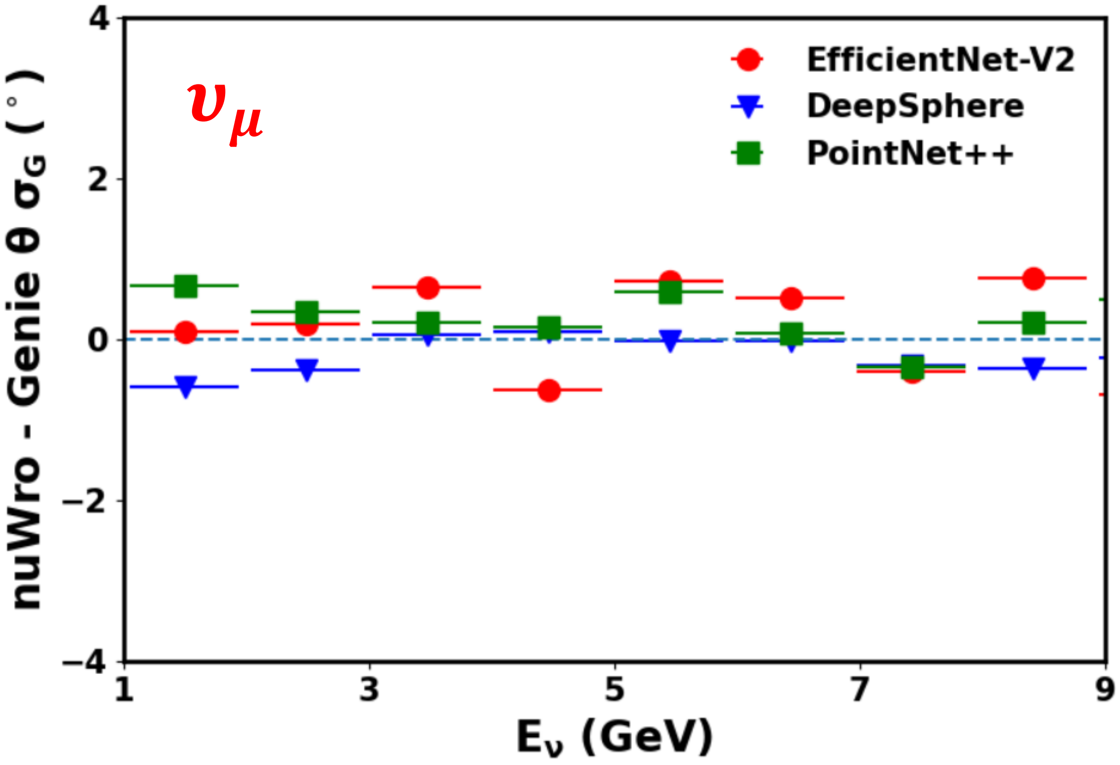


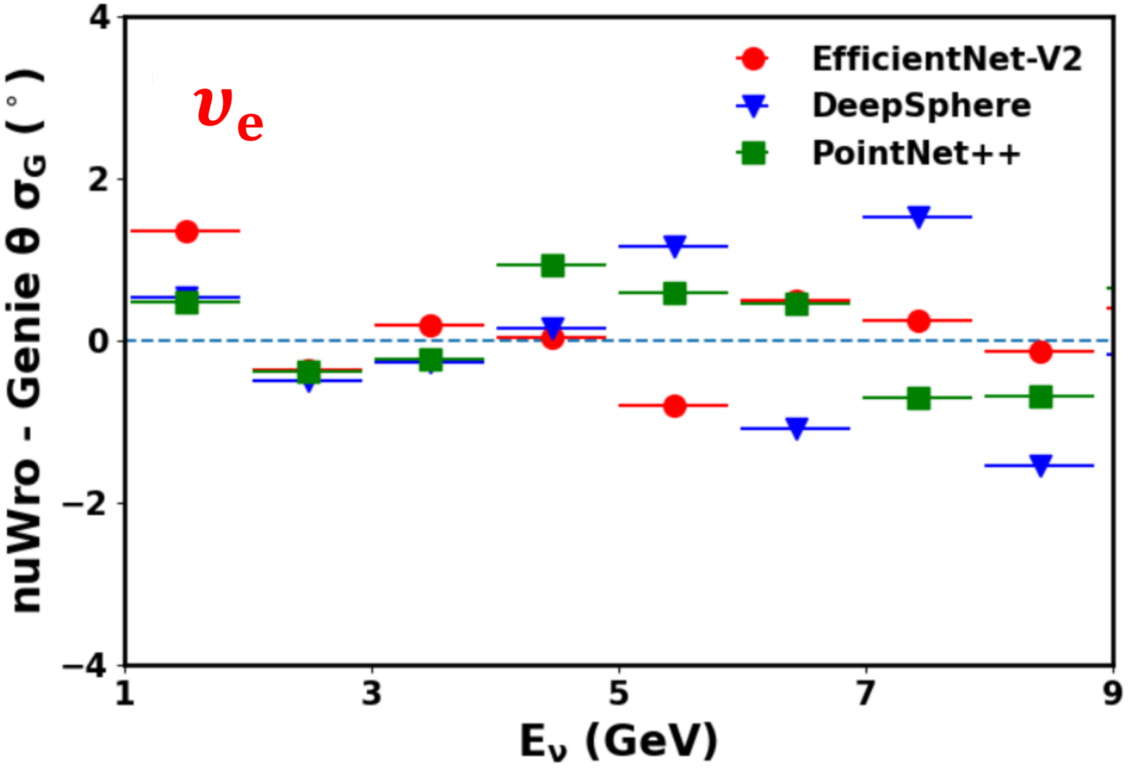


In general, NuWro predicts less neutron and proton than GENIE.

Comparison between different neutrino generators

Difference of σ_G between NuWro and GENIE results:



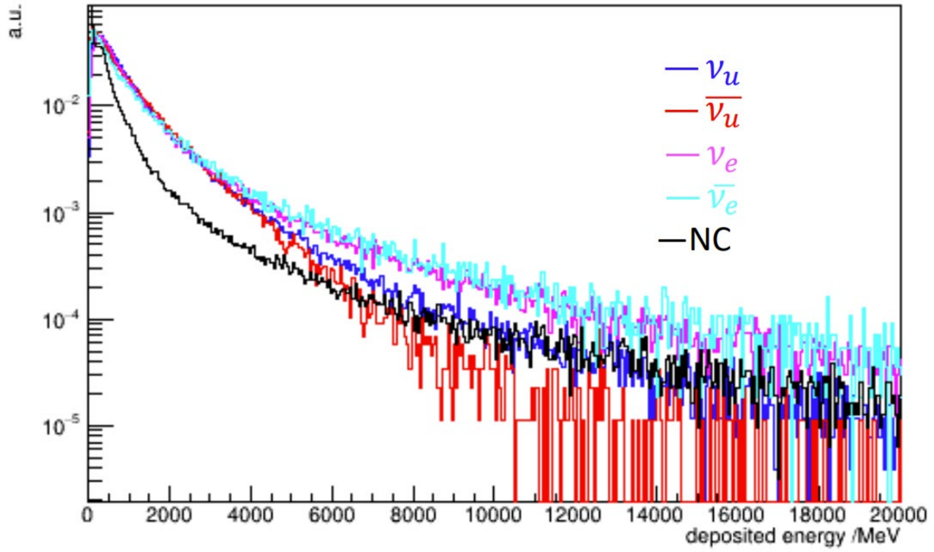


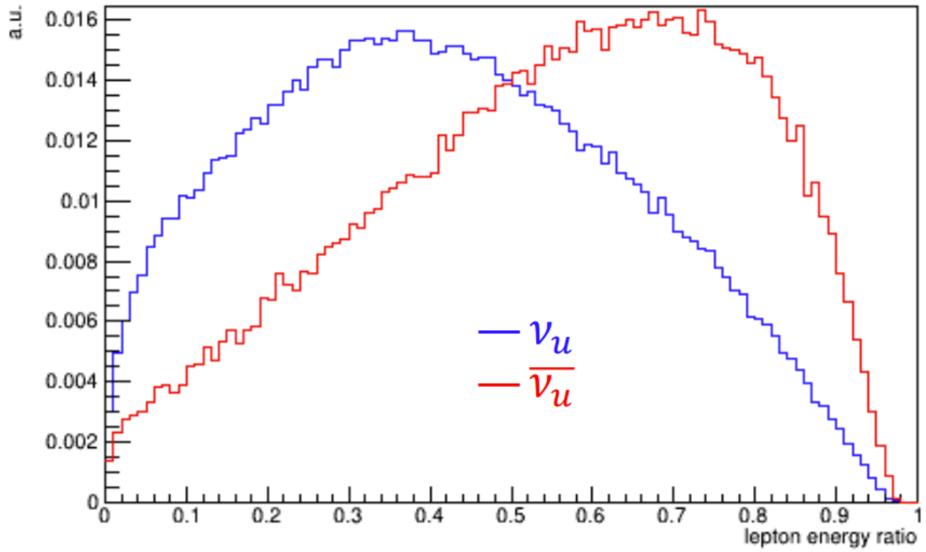
(Unpublished results)

Event level features for particle identification of atmospheric neutrinos

To distinguish between ν and anti-ν, event level features are necessary, which include:

- Neutron multiplicity
- Neutron positions
- Distance from the deposit center to interaction vertex
- Lepton energy ratio for $\nu_u/\overline{\nu_u}$
- Michel electrons, Isotopes, etc.

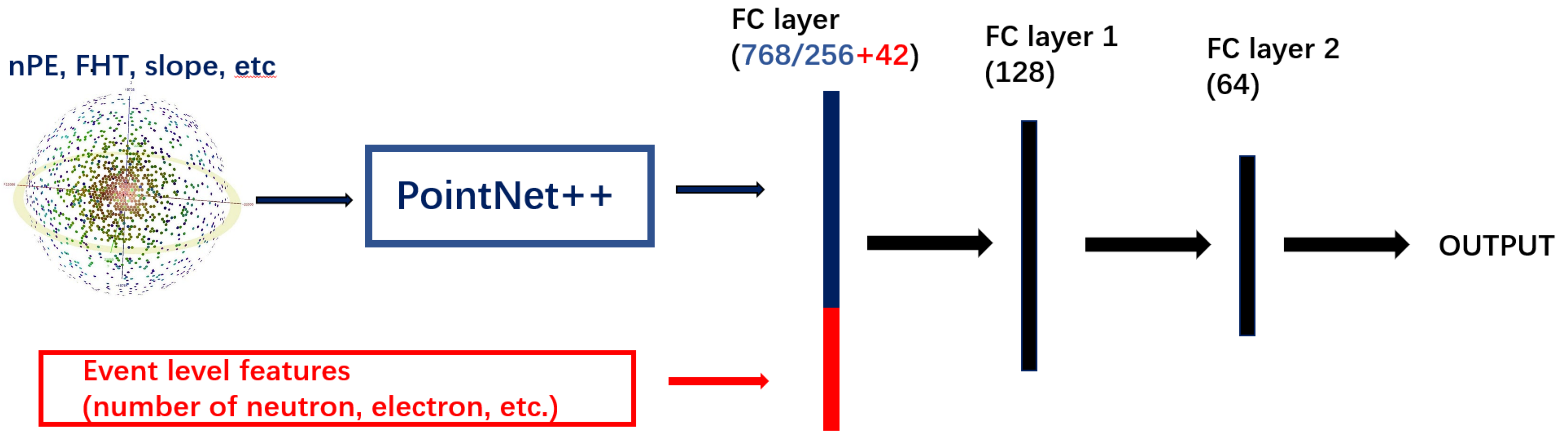




Particle Identification for atmospheric ν with Machine Learning methods

A 5-label classifier is developed to identify ν_μ -CC, $\bar{\nu}_\mu$-CC, ν_e-CC, $\bar{\nu}_e$-CC and $NC\ \nu$

- for fully contained events
- both PMT and event level features are used

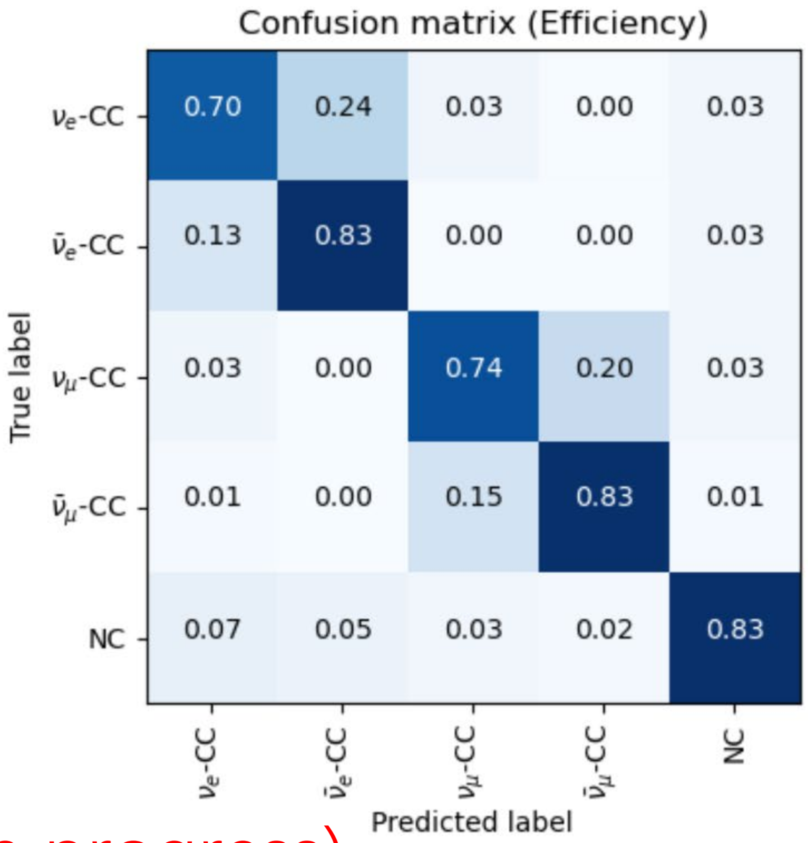


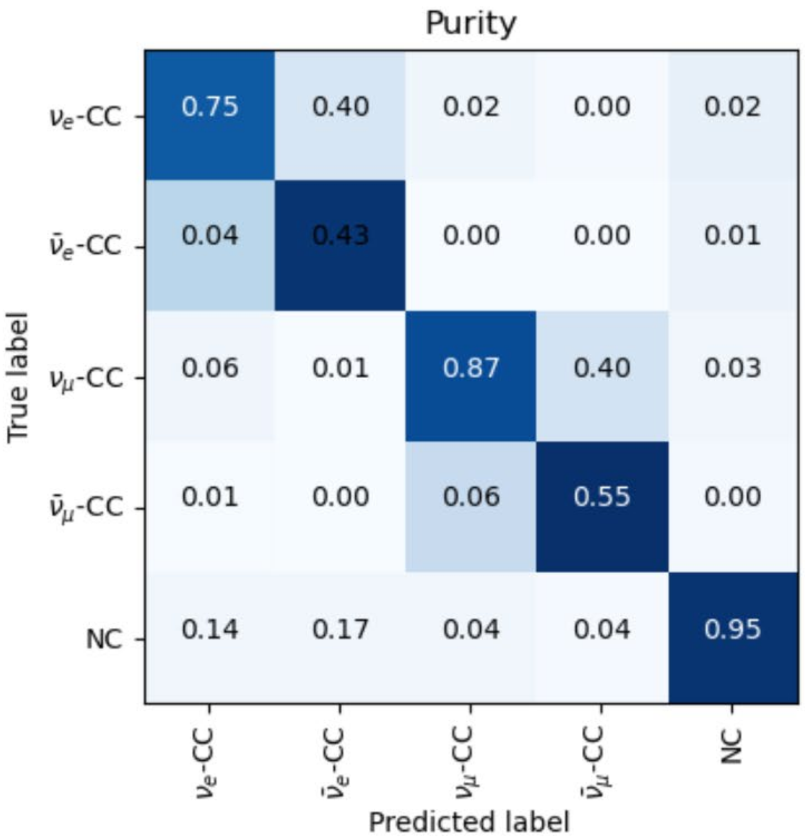
5-label classifier results

Balanced data sample is used for trainning; Purity is corrected with true event ratio

$$\text{Efficiency} = \frac{N^i_{true}}{\sum_i N^i_{true}}$$

$$\text{Purity} = \frac{N^i_{pred} * Ratio^i}{\sum_i N^i_{pred} * Ratio^i}$$

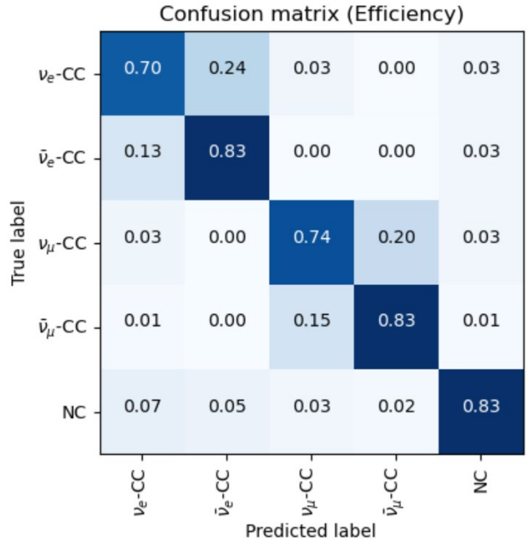


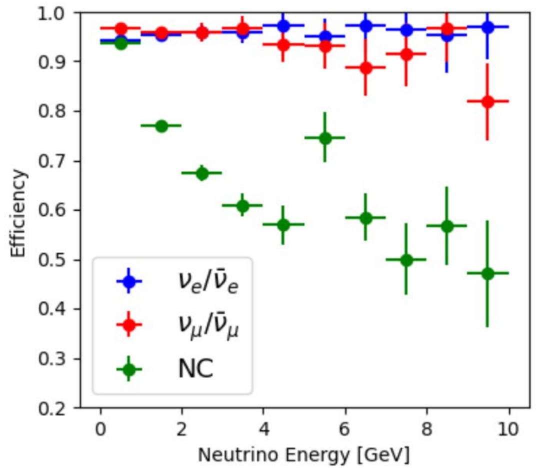


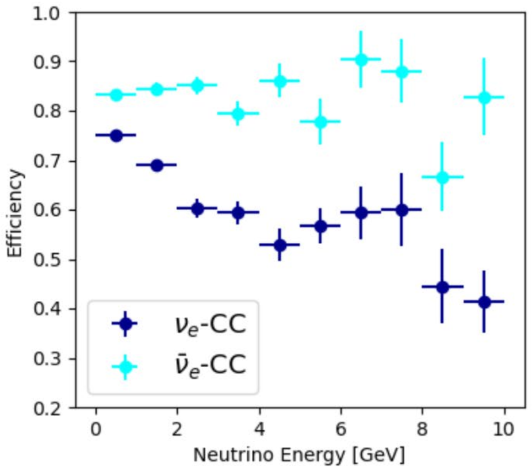
(Work in progress)

Validation with different event generators

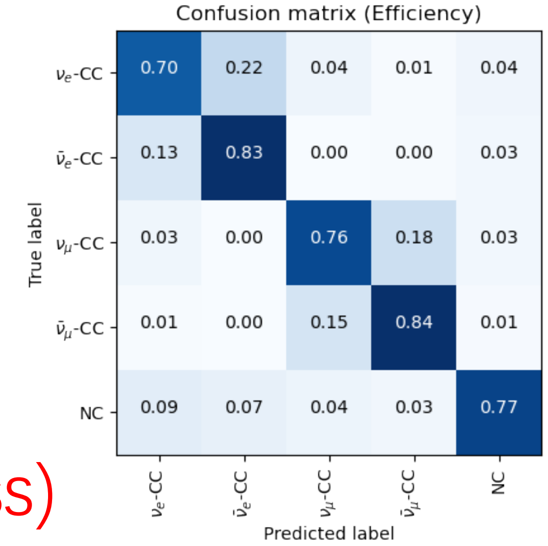
Validation with GENIE sample

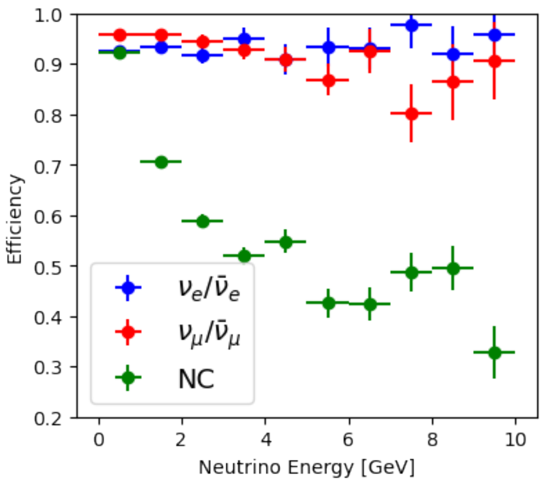


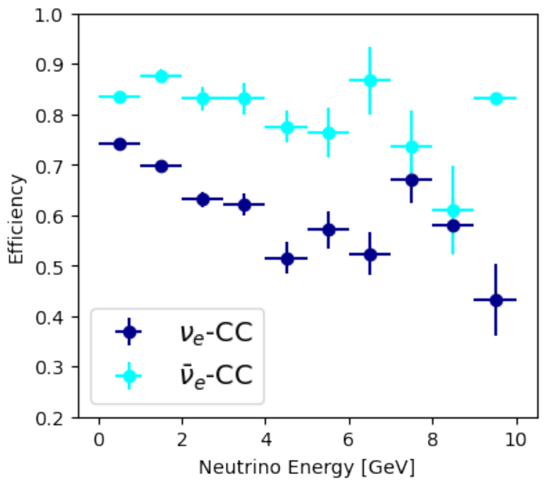




Validation with NuWro sample







(Work in progress)

总结

1. 本报告展示了一套针大型液闪探测器中高能(GeV)事例的多用途的机器学习算法，这套算法可以用来实现对大气中微子事例的入射方向的重建和对大气中微子事例类型的鉴别。
2. 对大气中微子事例入射方向的重建：
   - 在国际上第一次实现了基于大型液体闪烁体探测器的大气中微子方向重建；
   - 结果的可靠性通过使用不同的机器学习模型和不同类型的事例产生子进行了检验。
3. 对大气中微子事例类型的鉴别,
   - 初期的研究表明，通过结合PMT的波形特征和事例级别的特征（例如中子多重度等）作为机器学习的输入，可以实现对中微子味道和正反的高效率、高纯度鉴别。