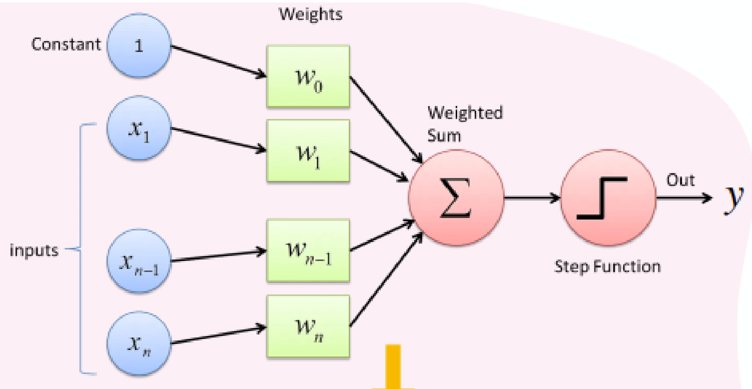
Reconstruction Based on GNN

Cen Mo, Fan Hu, Fuyudi Zhang, Liang Li

Neural Network

- Neural network is only a **function** that maps input to output.
- Machine learning: use computer to find the **BEST function** for our tasks.

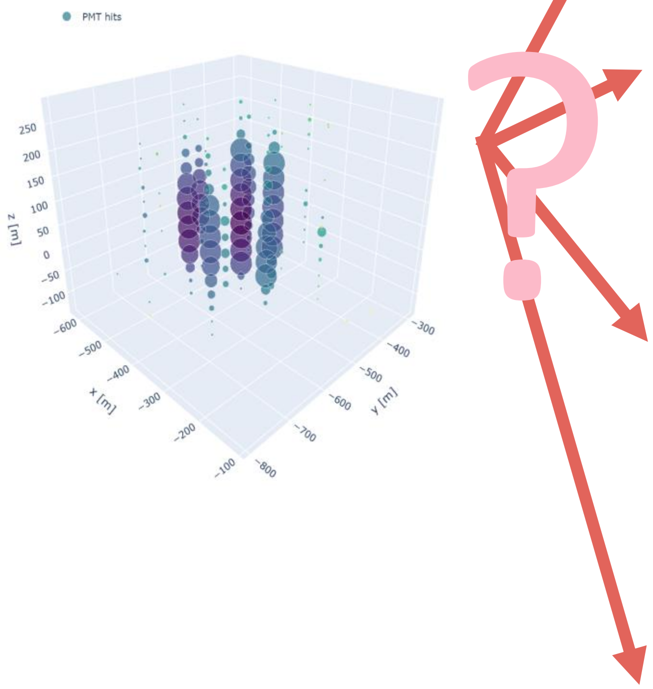


$$\sigma\left(\Sigma\begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m1} & w_{m2} & \cdots & w_{mn}\end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n\end{bmatrix}\right) = y$$

$$y = f_W(x)$$

- Multilayer Perceptron **(MLP)**
- Boosted Decision Tree **(BDT)**
- Convolutional Neural Network **(CNN)**
- Recurrent Neural Network **(RNN)**
- Graph Neural Network **(GNN)**

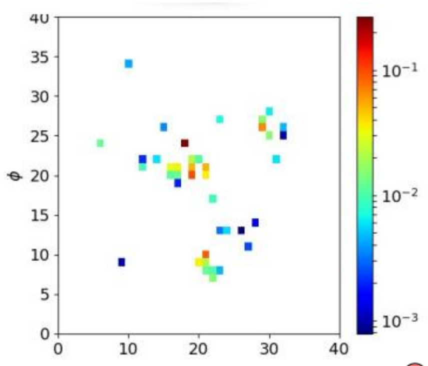
Data Representation



Vector with fixed length: $[\mathrm{DOM}_1, \mathrm{DOM}_2, \ldots, \mathrm{DOM}_{24220}]$

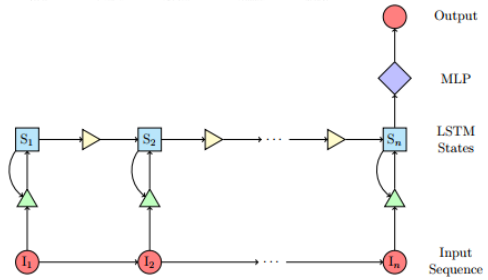
- Multilayer Perceptron **(MLP)** or Boosted Decision Tree **(BDT)**

Image



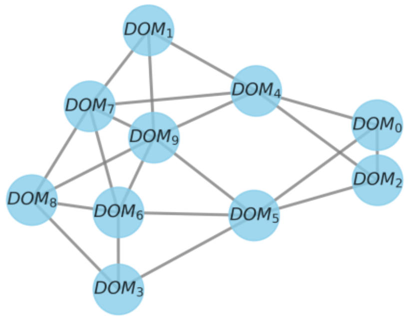
- Convolutional Neural Network **(CNN)**

Ordered **Sequence**



- Recurrent Neural Network **(RNN)**

Graph



- Graph Neural Network **(GNN)**

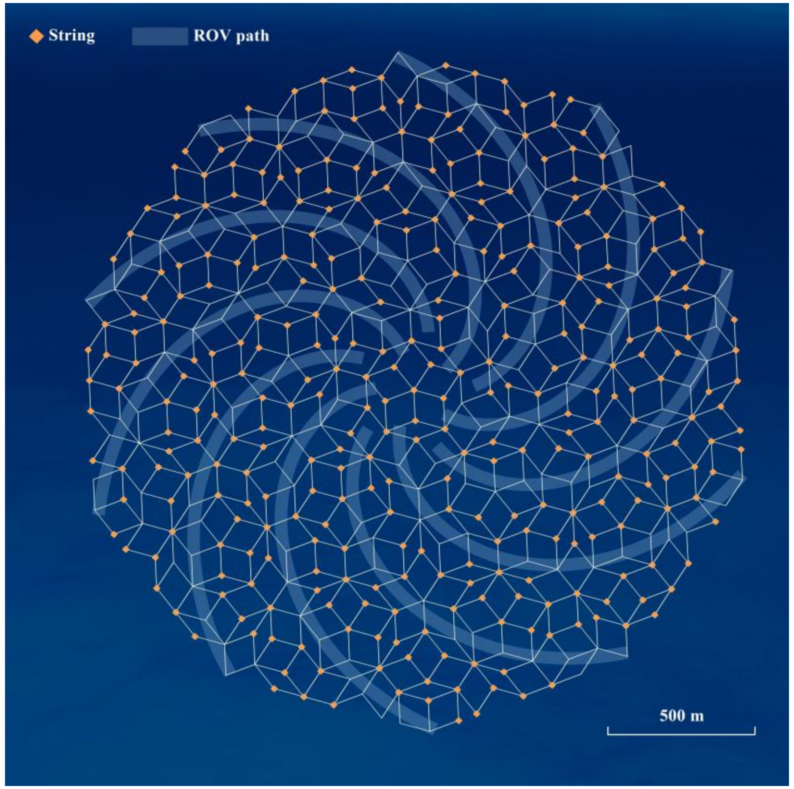
Data Representation

Neutrino telescope:

- Irregular detector geometry
- Sparse signal

Compared GNN and SSCNN (arxiv:1706.01307) performance:

- GNN outperforms SSCNN in terms of angular resolution in track-like events.



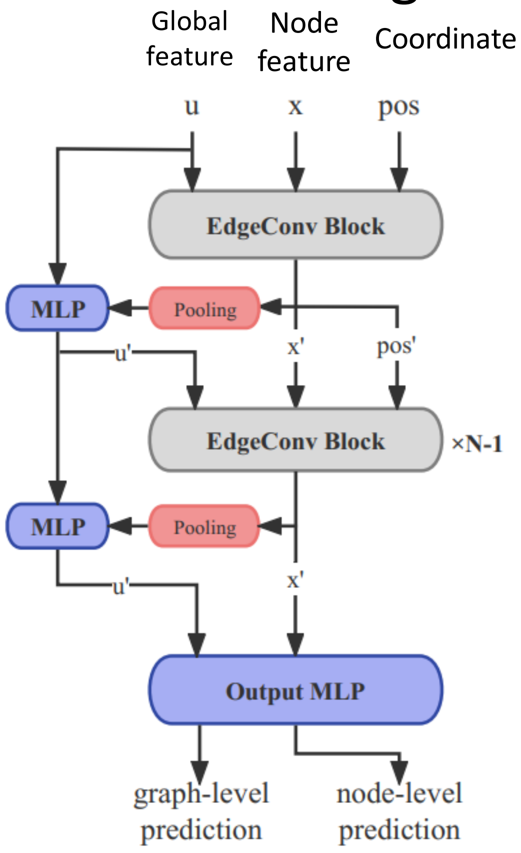
Top view of TRIDENT

Use point cloud to represent neutrino events:

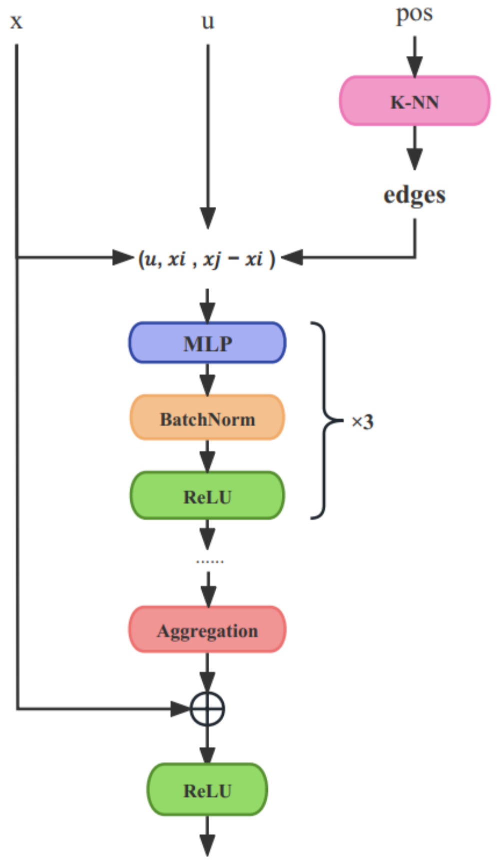
- Triggered DOMs → Nodes of point cloud
- Location of DOMs → Coordinate of nodes, pos_i.
- DOM-measured time and charge → Features of nodes, x_i.

Network Architecture

- TridentNet is built based on EdgeConv block: modified from EdgeConv block used in ParticleNet (arxiv:1902.08570).
- Both graph-level and node-level target can be predicted.



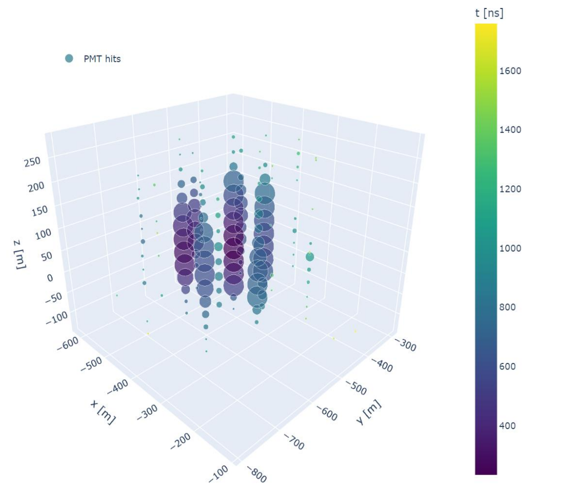
Network Structure



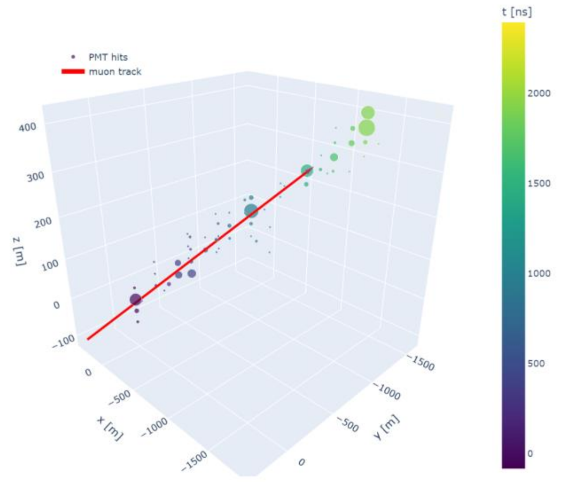
EdgeConv Block

Contents of Tasks

- Reconstruction of ν_e/**cascade** events
  - Direction Reconstruction
  - Energy Reconstruction



- Reconstruction of ν_μ/**track** events
  - Direction Reconstruction
  - Energy Reconstruction

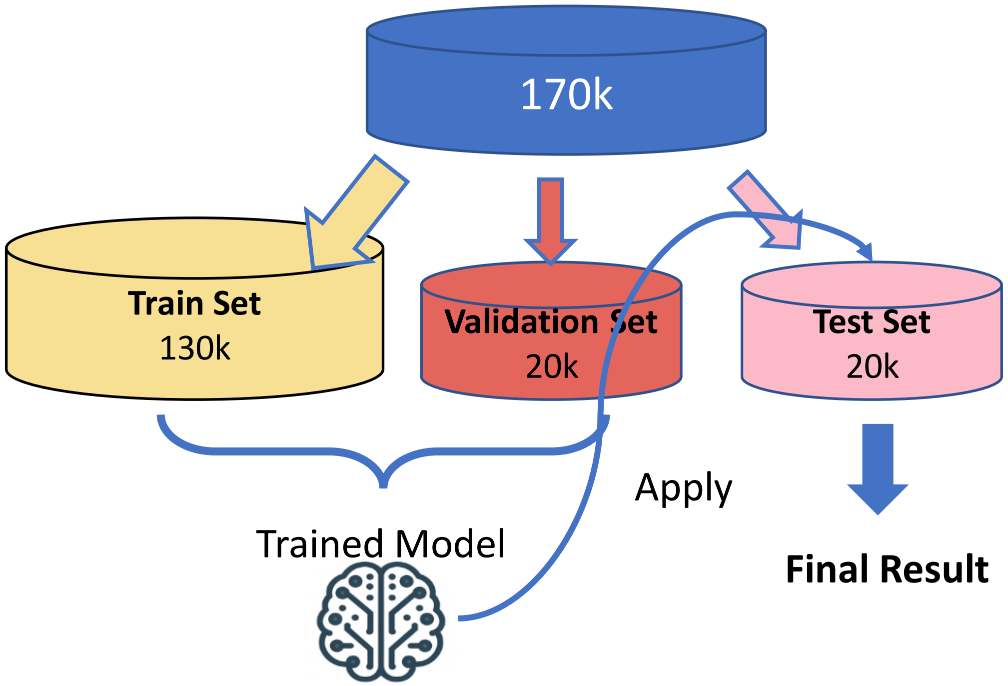


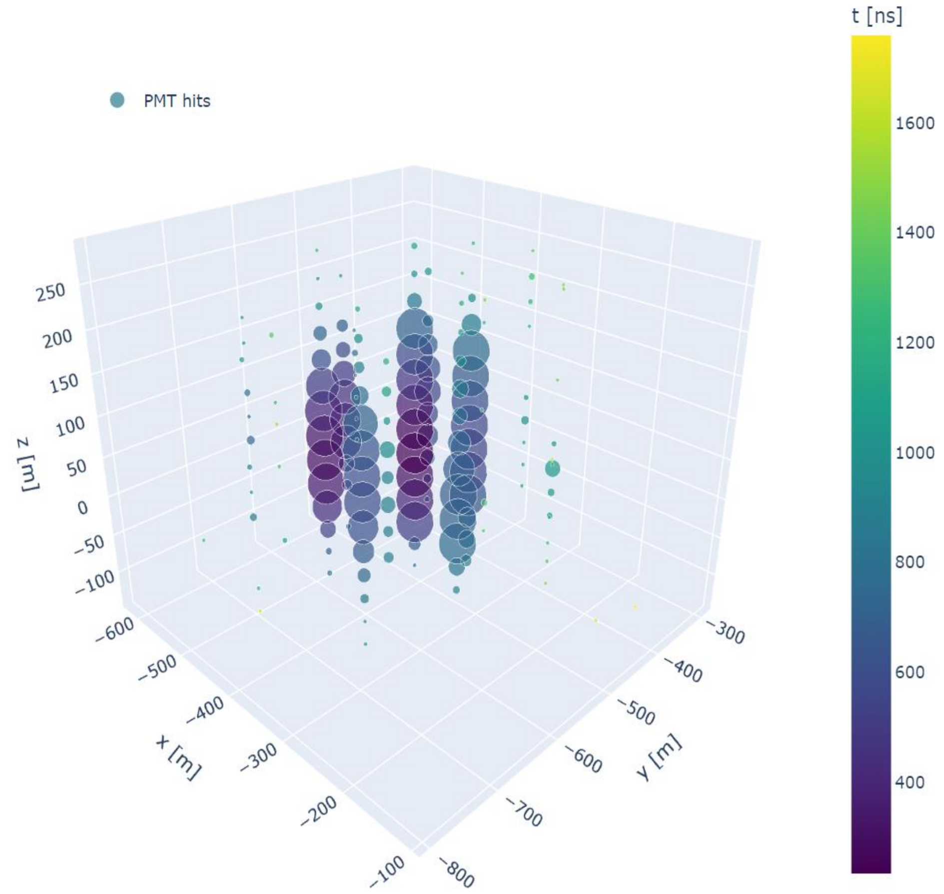
Cascade Reconstruction

ν_e direction reconstruction

Configuration for Direction Reconstruction

- ν_e energy: 100TeV
- Sample size:





Cascade Reconstruction

ν_e direction reconstruction

- Input feature of $\mathrm{DOM_i}$:

$[n_x, n_y, n_z]$

$$\mathrm{DOM_i}\!: [hit_0, hit_1 \dots hit_N] \quad (hit_i = [\boxed{x_i, y_i, z_i, n_x, n_y, n_z,} t_i])$$

$$\mathrm{Node_i}\!: Histogram([t_0, t_1 \dots t_N], bin\ width = 5ns) + \boxed{[x, y, z] + [n_x, n_y, n_z]}$$

Average

- Histogram is further modified (as a **normalization**):

$$\mathrm{Node_i} = [x, y, z, n_x, n_y, n_z, \ln(Histogram + 1)]$$

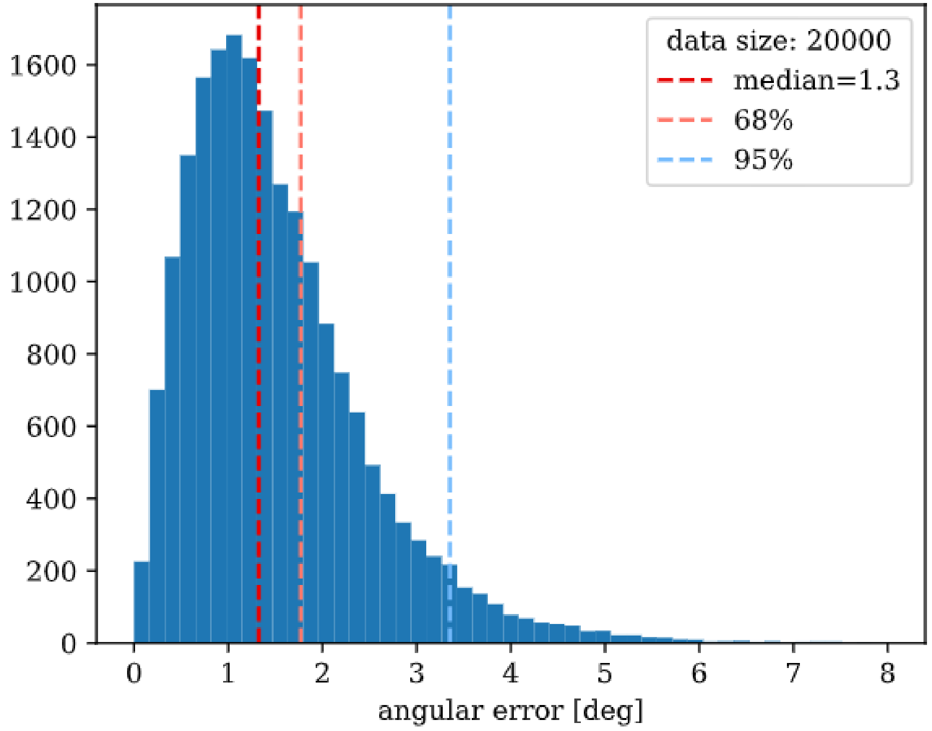
- Network is trained to predict $\hat{n}_\nu$ with MSE loss:

$$Loss = \left| \frac{\overrightarrow{output}}{|output|} - \hat{n}_\nu \right|^2$$

Cascade Reconstruction

ν_e direction reconstruction

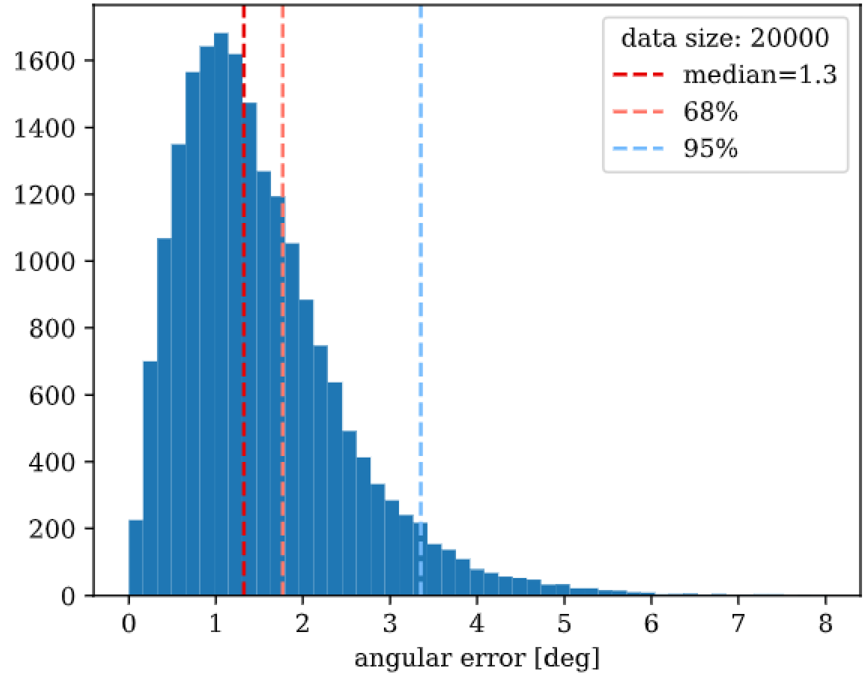
- Resolution reaches **1.3 degrees**.



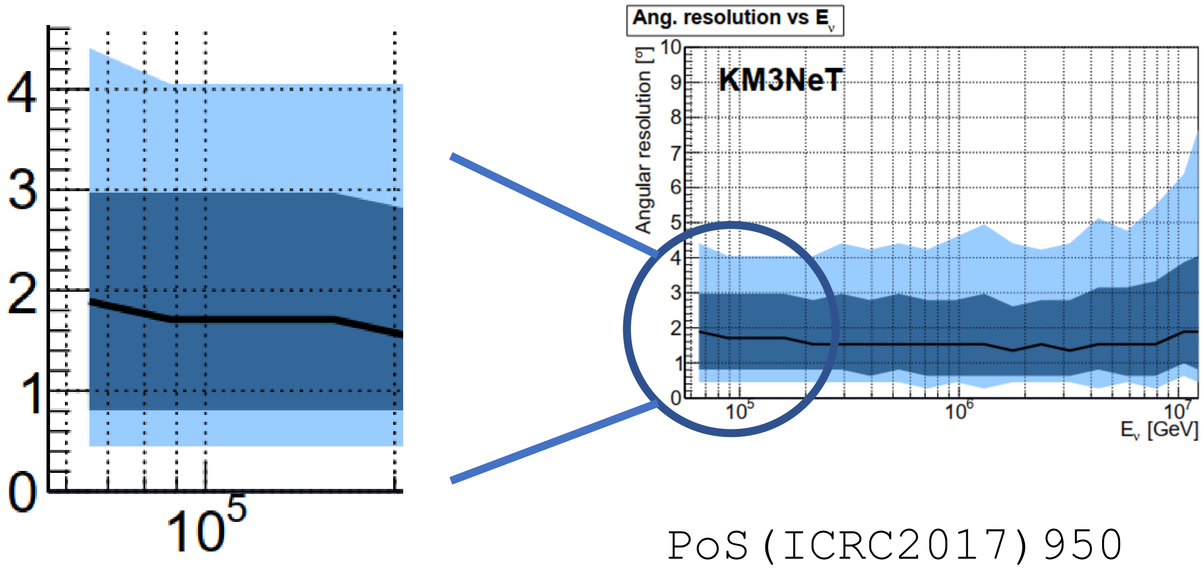
Cascade Reconstruction

ν_e direction reconstruction

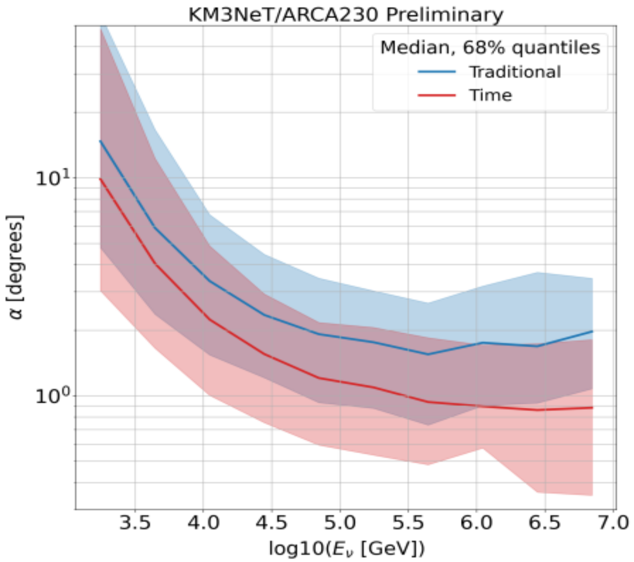
- Resolution reaches **1.3 degrees**.



KM3NeT Results (likelihood)



PoS(ICRC2017)950



PoS(ICRC2023)1074

With pre-selection

Cascade Reconstruction

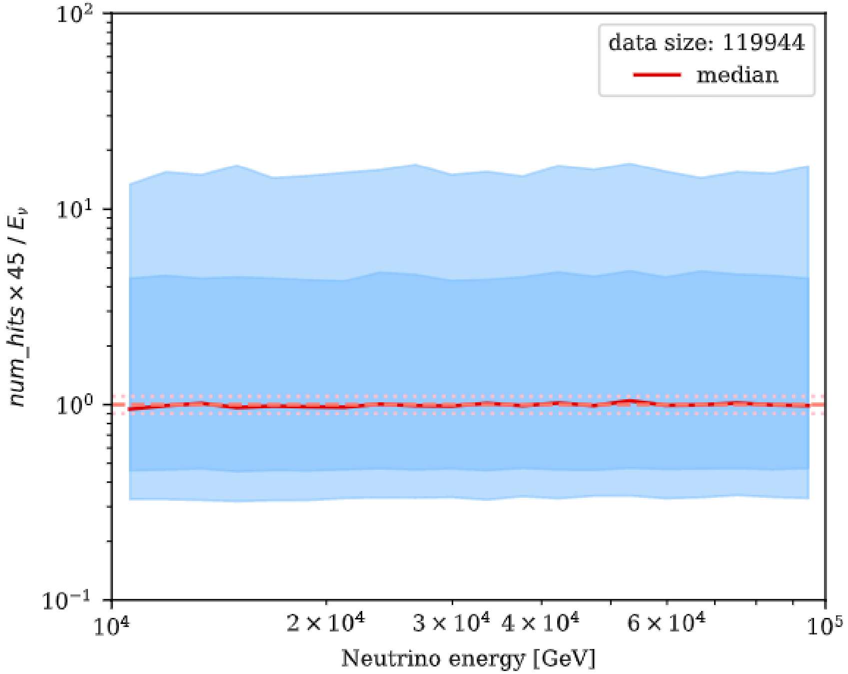
- **ν_e energy reconstruction**

Configuration for Energy Reconstruction

- ν_e energy: 10TeV ~ 100TeV
- Sample size: 150k samples are splitted into:

$$train : validation : test = 120k : 15k : 15k$$

Linearity between num_hits & Energy

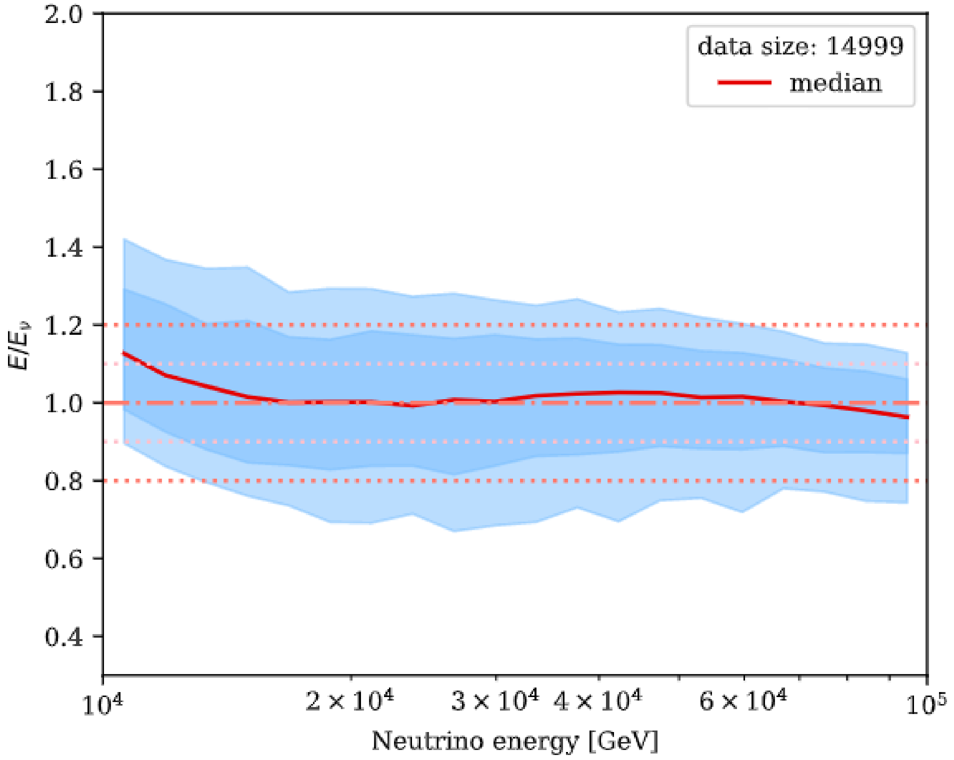


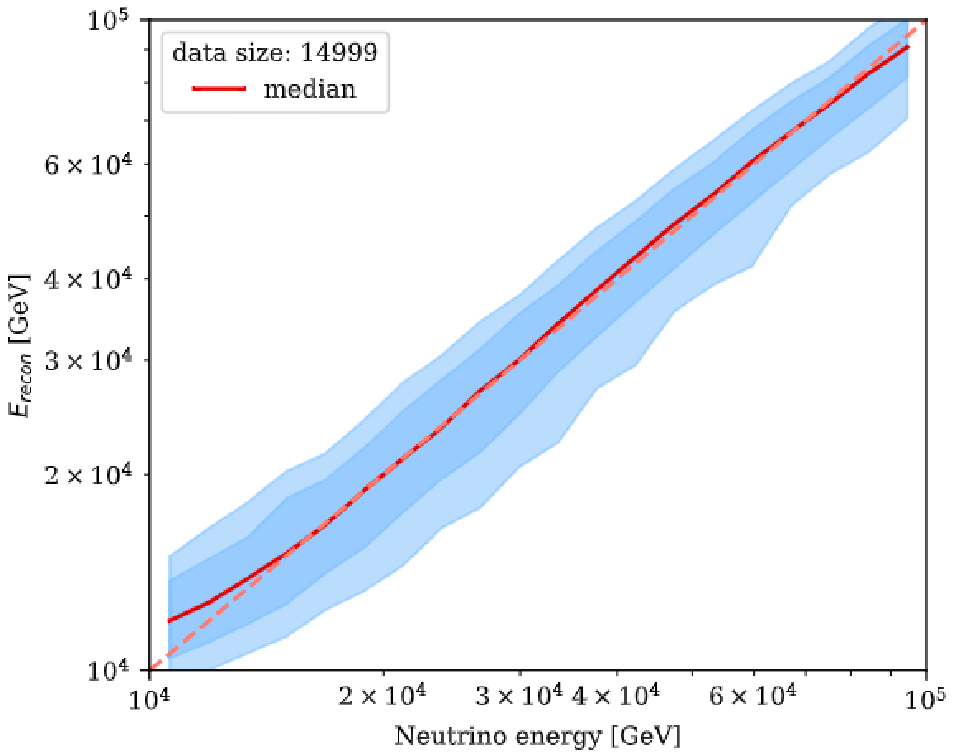
Cascade Reconstruction

- Train GNN with:

$$\log_{10} E = GNN(graph) + \log_{10}(\text{num\_hits} \times 45.78\,)$$
$$Loss = (\log_{10} E \;- \log_{10} E_{truth})^2$$

- Energy resolution is around 10% for high energy event.

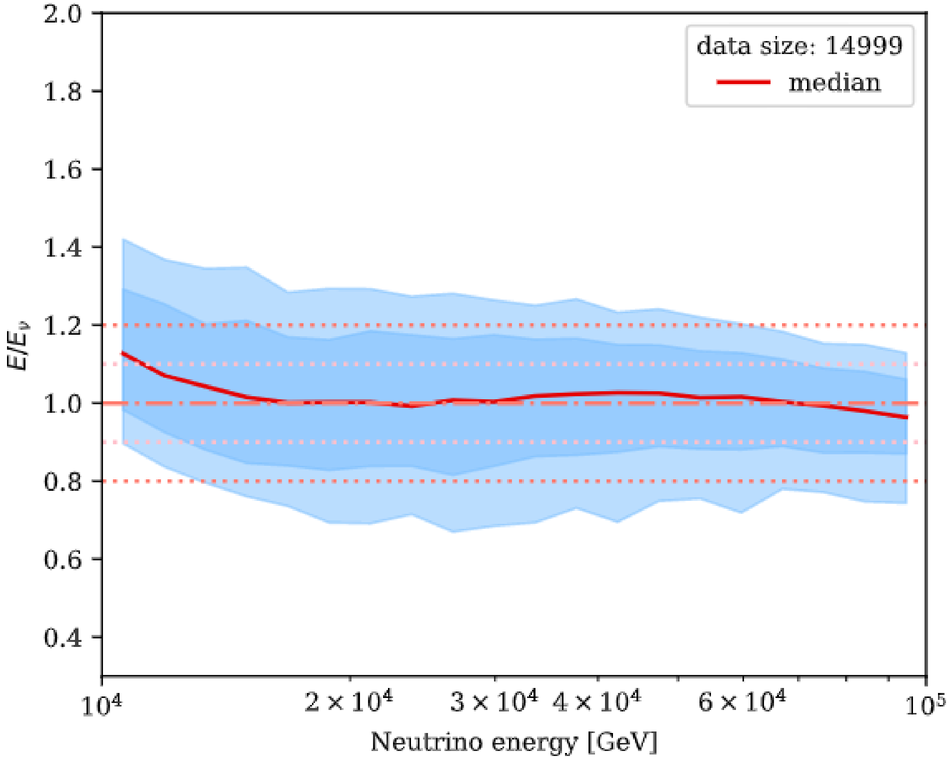


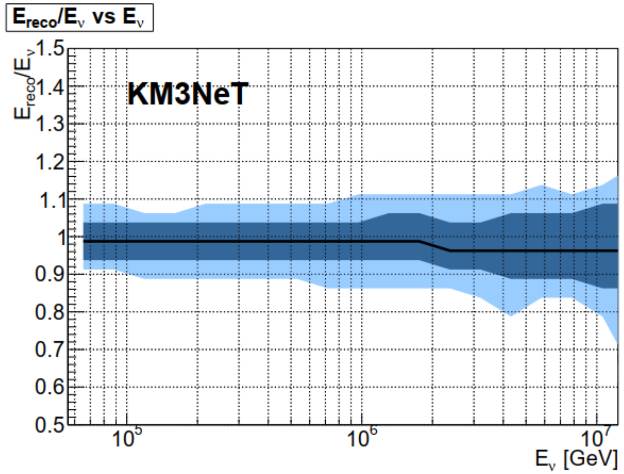


Cascade Reconstruction

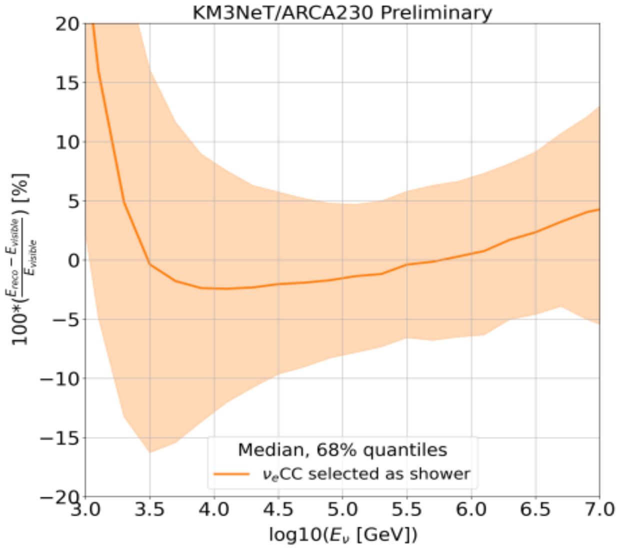
- Comparison

KM3NeT Results (likelihood)





PoS(ICRC2017)950



PoS(ICRC2023)1074
With pre-selection

Track Reconstruction

- **ν_μ Direction reconstruction**

Configuration for Energy Reconstruction

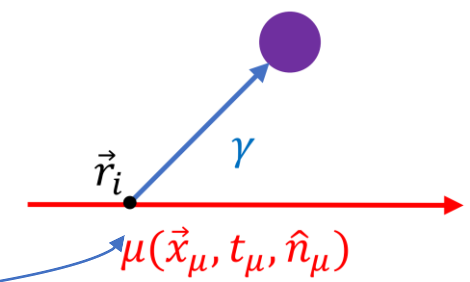
- ν_μ energy: 1TeV ~ 1PeV
- Sample size:

$$train : validation : test = 900k : 70k : 100k$$

Track Reconstruction

ν_μ Direction reconstruction

- Input features: location $\vec{D}_i$, first photon arrival time T_i and number of photo hits n_i



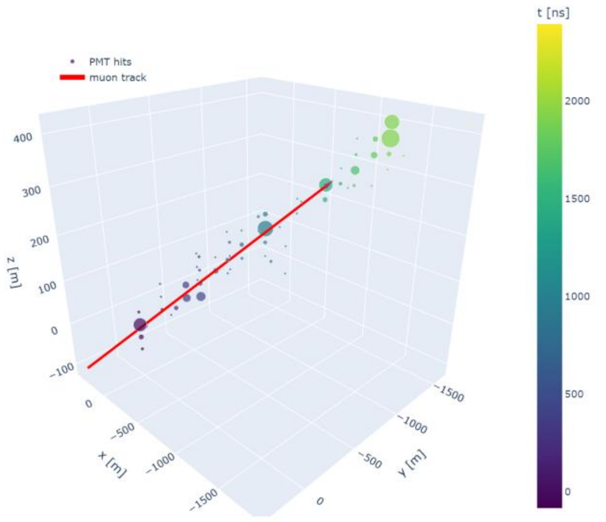
μ emits photon and triggers $\mathrm{DOM_i}$

- To make full use of the geometric feature of track-like events, the network is trained to predict $\vec{r}_i$ for each $\mathrm{DOM_i}$.

- Loss function: mean square error (MSE) with weight proportional to n_i:

$$Loss = \Sigma_i n_i \times \left|\overrightarrow{output}_i - \vec{r}_i\right|^2 / \Sigma_i n_i$$

- Linear fit on the predicted $\vec{r}_i'$ then reconstructs $\hat{n}_\mu$.

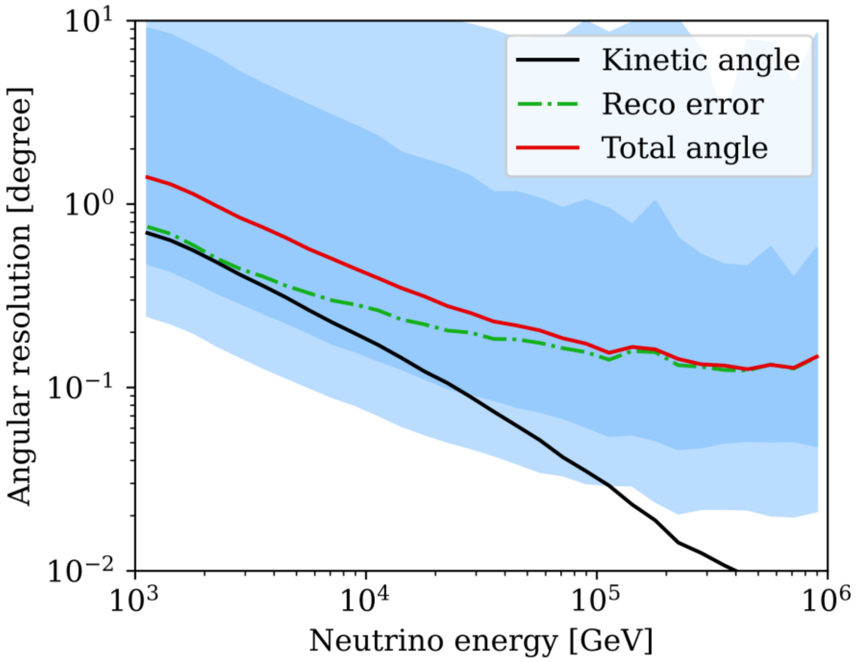


Track-like event display

Track Reconstruction

Direction reconstruction

- Model is trained on events with track length > 500m.
- Median angular error decreases from 1 degree to 0.1 degree as the energy of ν_μ increases – similar to the result of likelihood method.

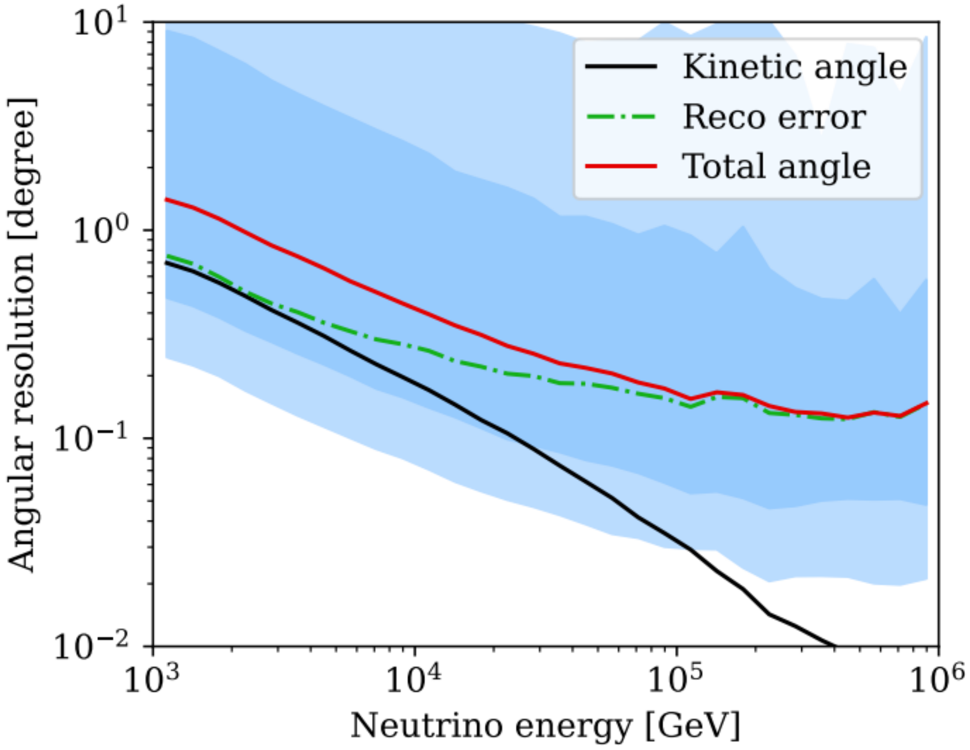


- Kinetic angle = $< \vec{n}_\mu, \quad \vec{n}_\nu >$
- Reco error = $< \vec{n}_\mu, \vec{n}_{recon} >$
- Total angle = $< \vec{n}_\nu, \vec{n}_{recon} >$

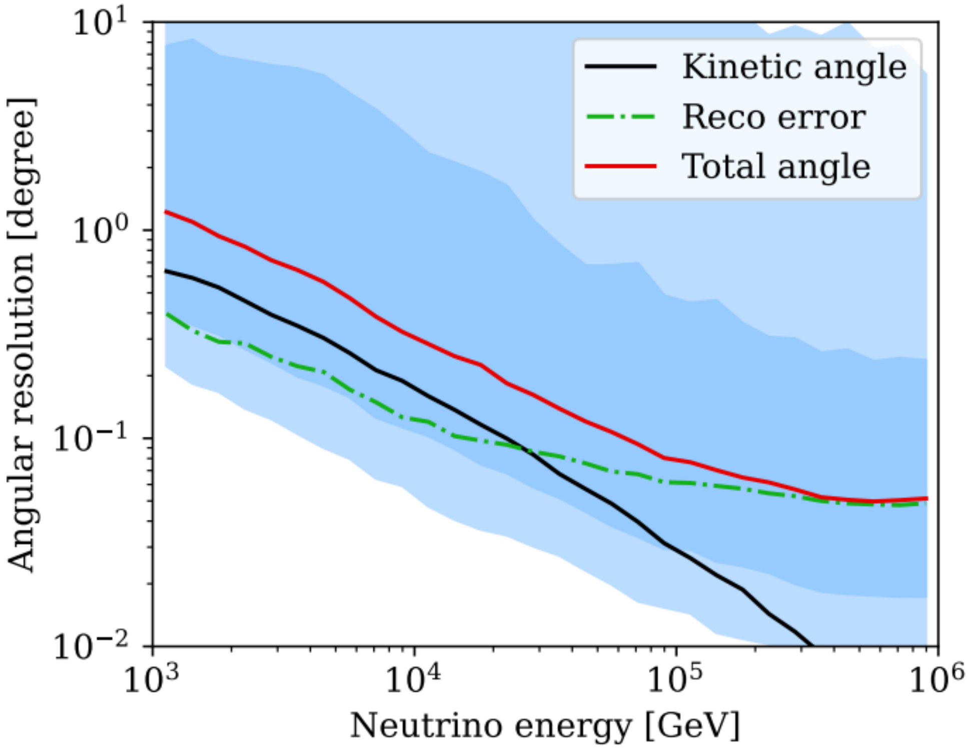
Track Reconstruction

Comparison

GNN



Fan's



Track-like Events Reconstruction

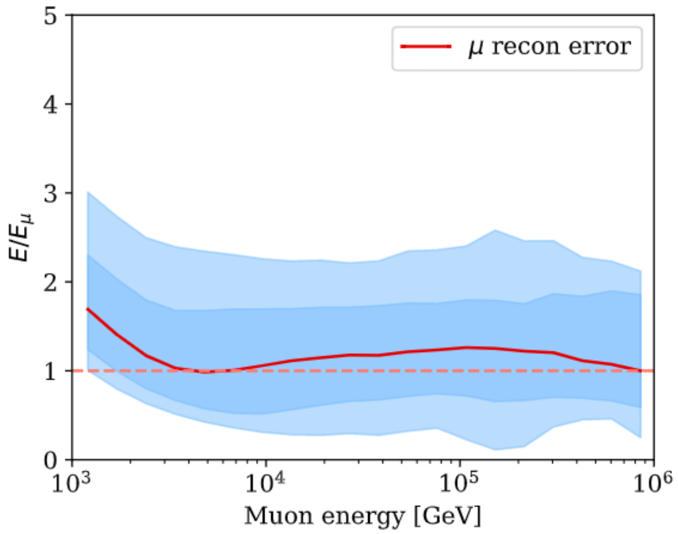
Energy reconstruction

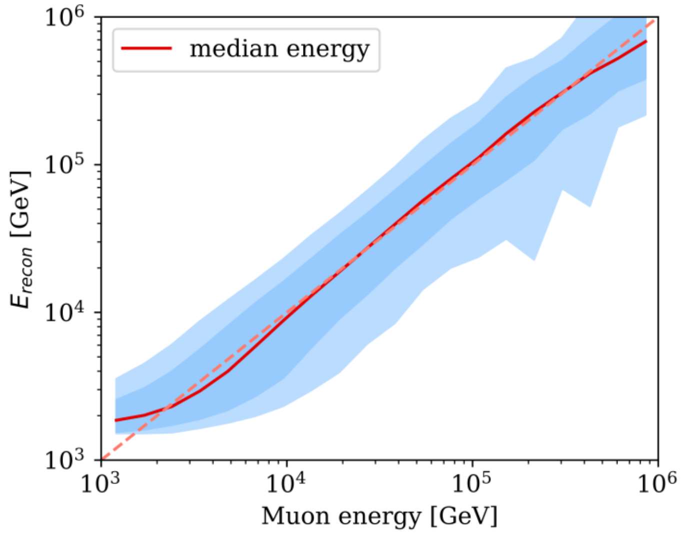
- Same input features as the direction reconstruction.
- Network is trained with MSE loss to predict $\log_{10} E_\mu$. Weight $w = \log_{10} E_\mu - 2.5$ is applied:

$$Loss = w\left(output - \log_{10} E_\mu\right)^2$$

- A shift term, $b = 0.15$ is added to outputs of the model:

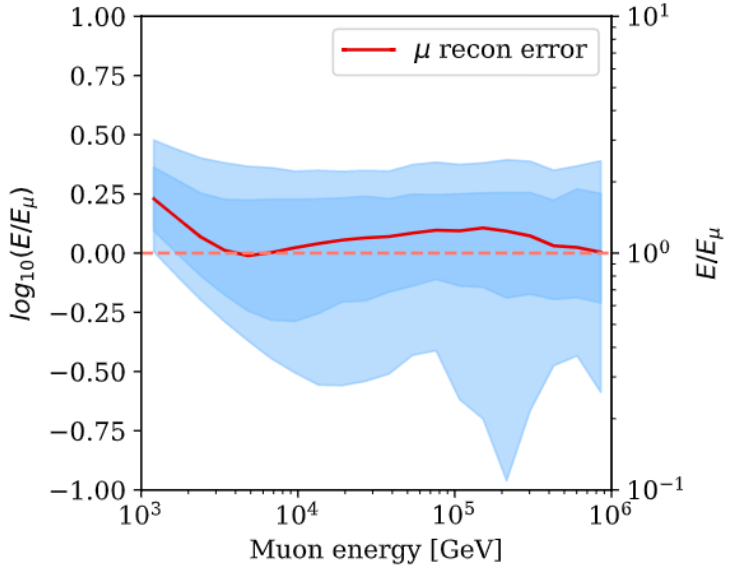
$$E_{recon} = 10^{\mathrm{output+b}}$$

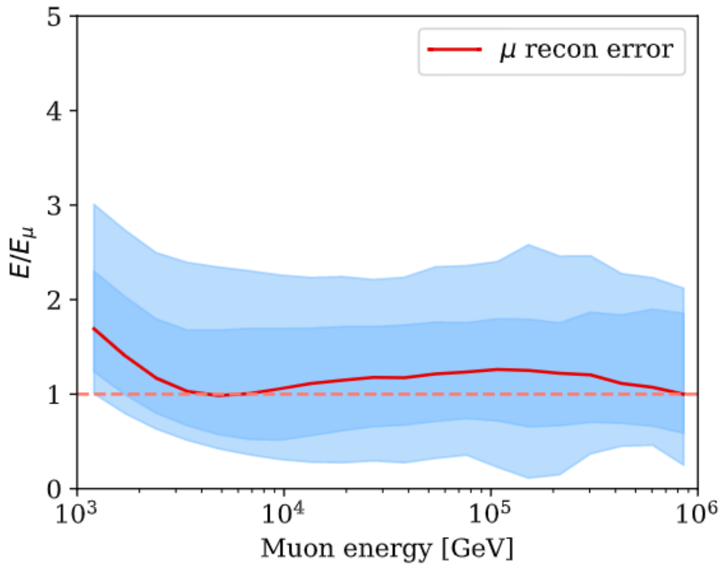




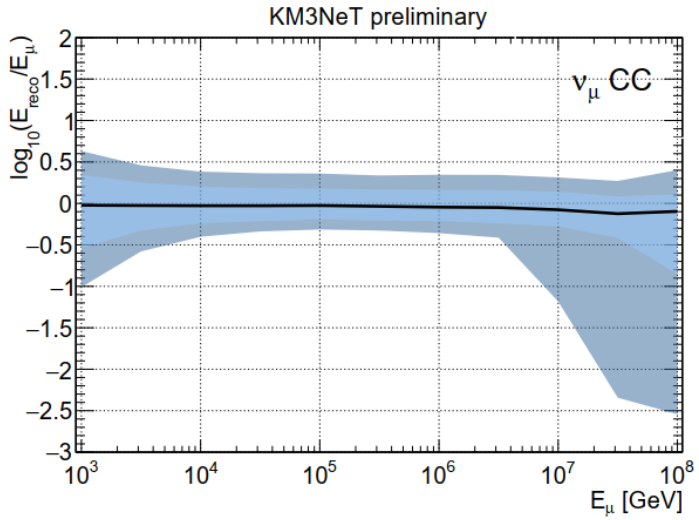
Track Reconstruction

Comparison

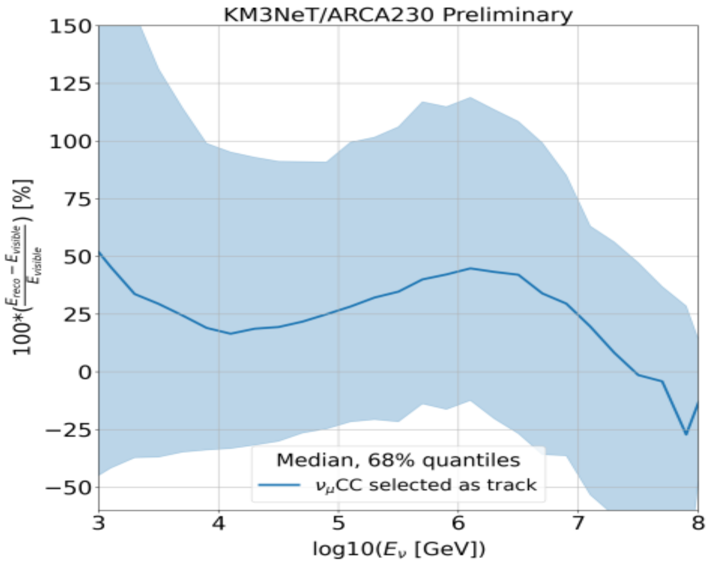




KM3NeT Results (likelihood)



PoS(ICRC2017)950



PoS(ICRC2023)1074

With pre-selection

Summary and Outlook

Task	Resolution (GNN)	Resolution (KM3NeT)
Cascade direction	**1.3 degrees**	2 degrees (w/o time) ~1 degree (with time)
Cascade energy	**10% (high E)**	5~10%
Track direction	**~0.1 (high E)**	0.1 (high E)
Track energy	**100%**	100%

Summary and Outlook

- Simulated neutrino events in TRIDENT are represented as point clouds and are reconstructed by TridentNet.
- GNN demonstrates high accuracy in reconstructing ν_e and ν_μ .

- **Improvement** of neutrino reconstruction will be further studied.
- Future research will try to enhance the method's robustness against experimental **uncertainties and noise**.

Thanks for listening!

Email: mo_cen@sjtu.edu.cn

Hyper-Parameters

- **Learning rate**:

  Initial learning Rate = 0.003

  lr scheduler: ReduceLROnPlateau(factor=0.5, patience=5)

- **optimizer**:

  Adam(betas=(0.9, 0.999), eps=1e-8, weight_decay=0)

Likelihood Reconstruction of ν_e

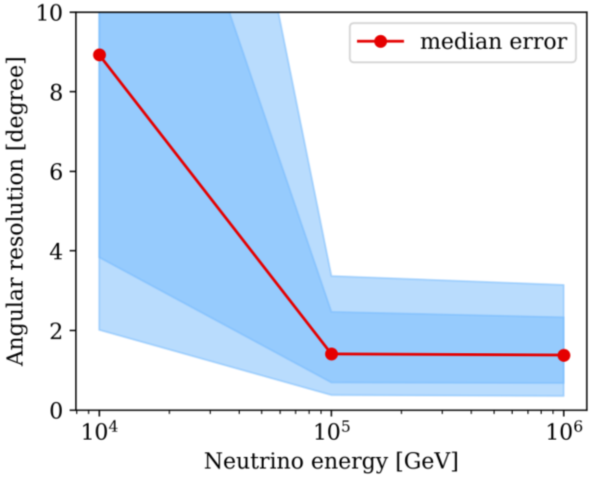
$$\log \mathscr{L} = \sum_{empties} \log[P_i^{nohit}] + \sum_{hitPMTs} \log[1 - P_i^{nohit}]$$

$$P_i^{nohit} = \exp[-\mu_{sig}(r_i, z_i, a_i, E_S) - R_{bg} \cdot T]$$

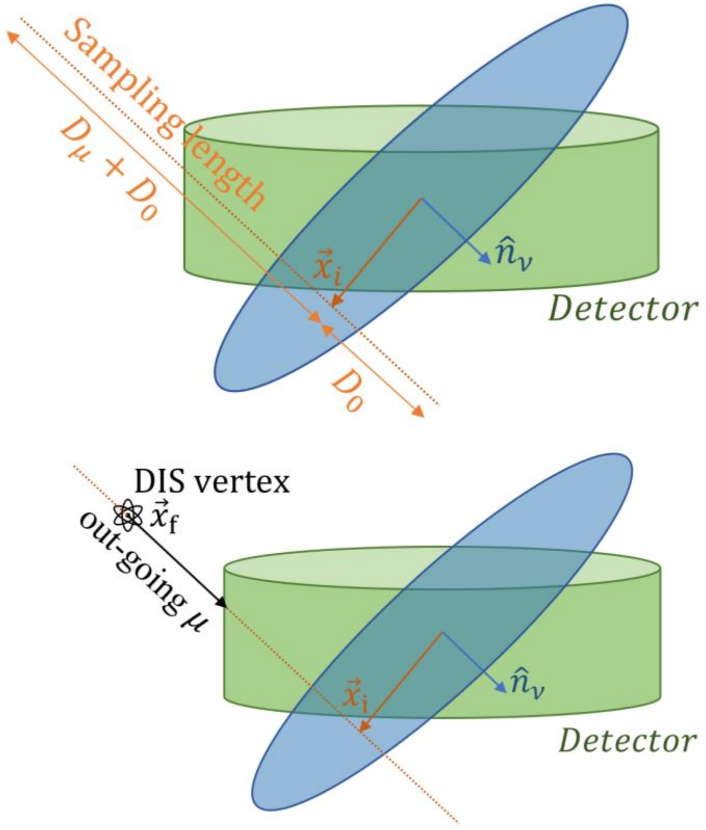
(equation 5.2) are considered. The expected number of hits from the shower μ_{sig} is evaluated using interpolation of a three-dimensional histogram depending on r_i, z_i and a_i at a shower energy of 1 PeV (figure 8). The expected number of hits at different shower energies is calculated using the fact that the number of emitted photons scales linearly with the shower energy E_S.

- Former GNN result on samples with other energy (by linear scaling num_photons):

$$n' = n \times \frac{100TeV}{E_\nu}$$



ν_μ Vertex Sampling



Effective Area of ν_μ

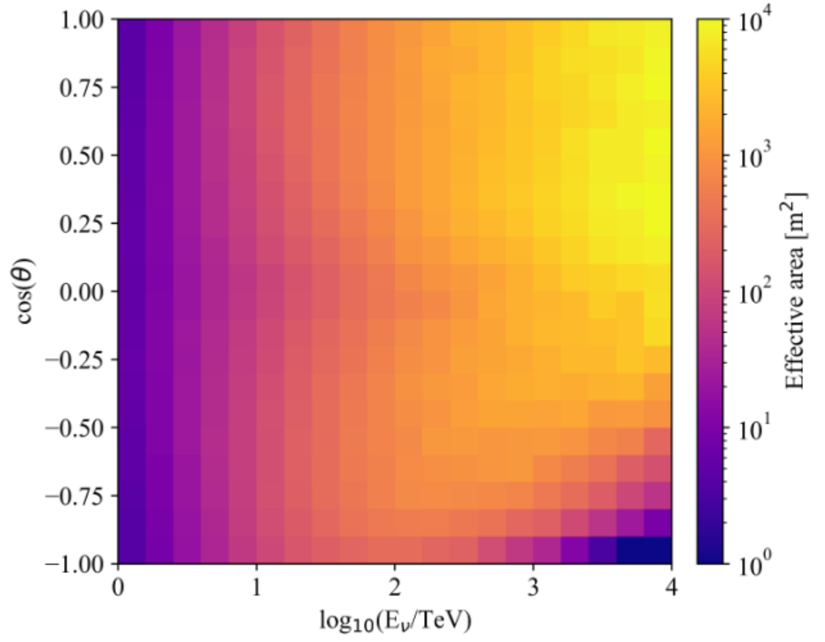
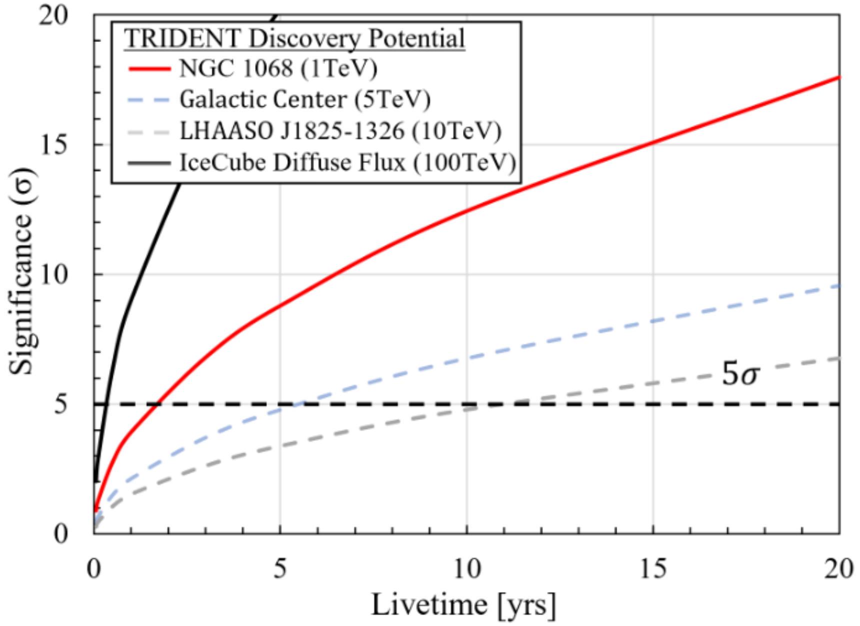


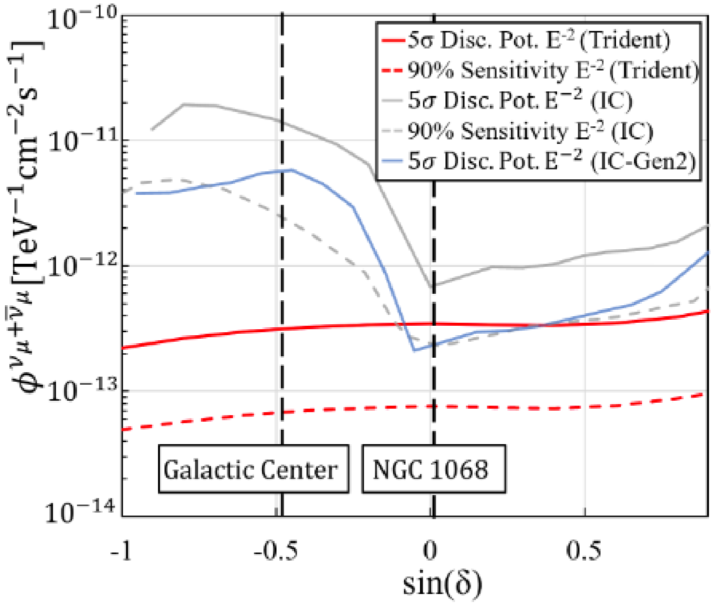
Figure 15: Effective areas at event reconstruction level for ν_μ track events as a function of primary neutrino energy and zenith angle in TRIDENT. At an energy of $\sim$ 100 TeV, the effective area for up-going events is expected to reach 7×10^2 m^2. Only events with anglular error less than 6 degree are selected to evaluate the effective area.

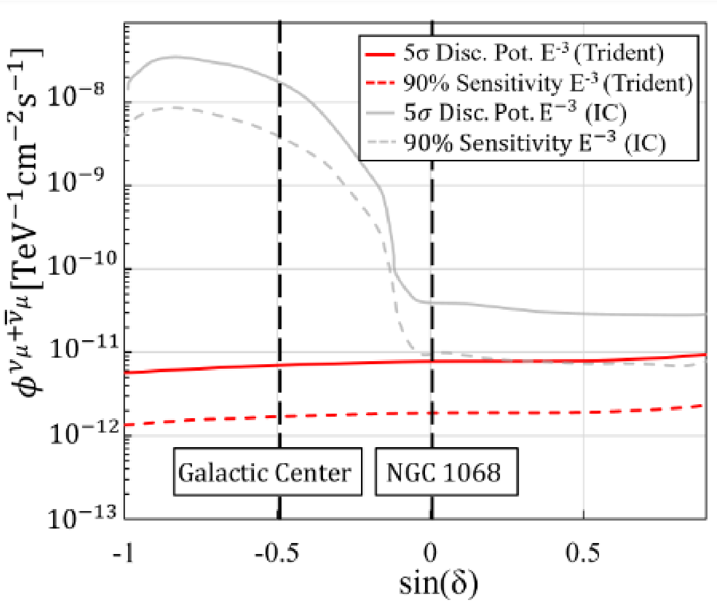
arXiv:2207.04519

Significance & Sensitivity

arXiv:2207.04519







Comparison with Likelihood Method

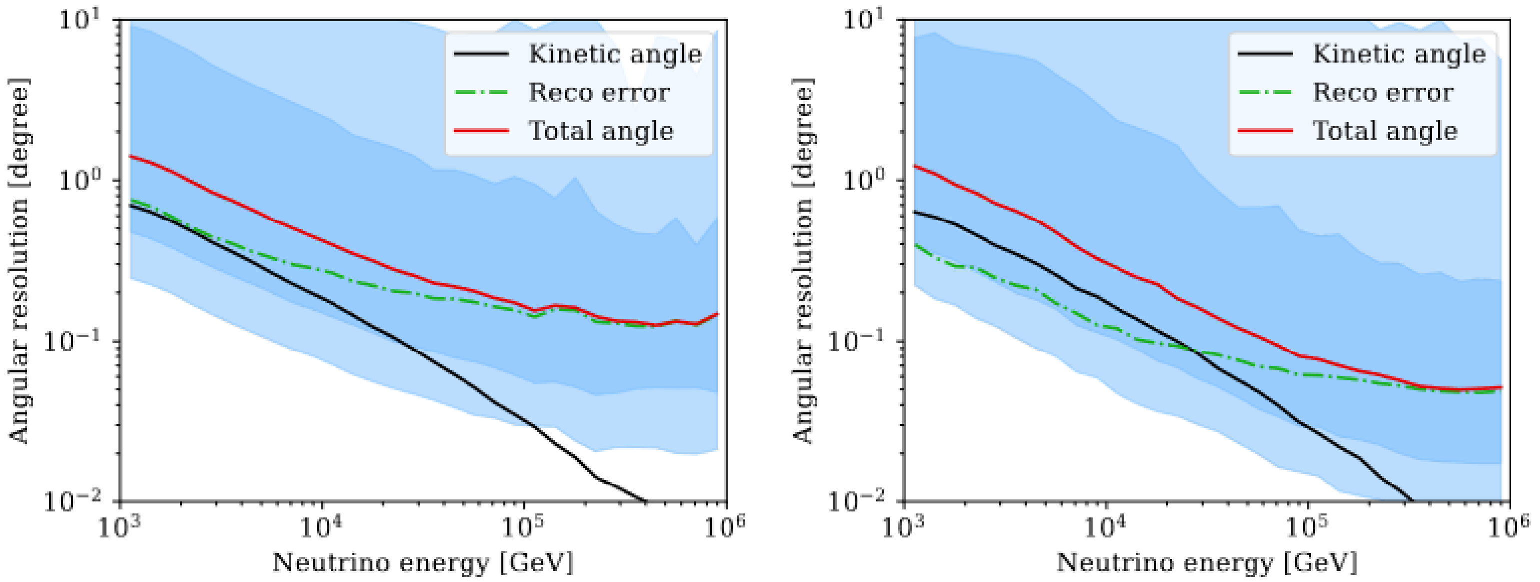
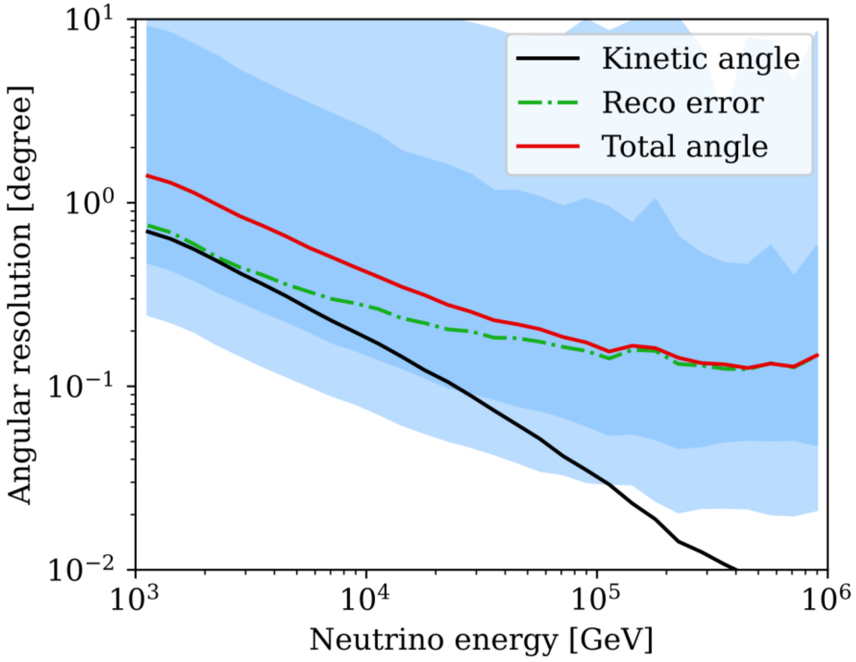


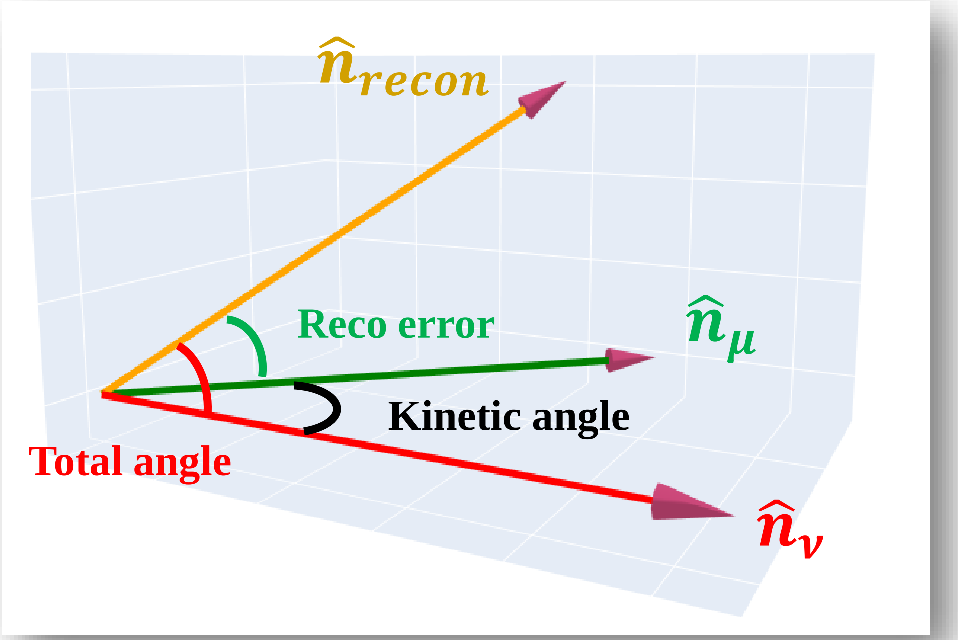
Figure 5: Median angular error of GNN (left) and likelihood method (right) depend on energy of ν_μ. The median angle between the reconstructed track and the true direction of μ and ν_μ is visualized by the green and red lines, respectively. Color bands exhibits the 68% and 90% quantiles. Black lines are the median angle between direction of μ and ν_μ.

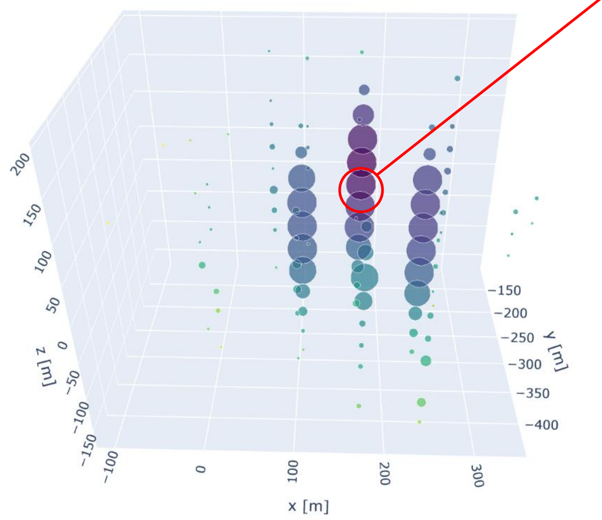
Track-like Events Reconstruction

Direction reconstruction

- Model is trained on events with track length > 500m.
- Median angular error decreases from 1 degree to 0.1 degree as the energy of ν_μ increases.







$$[x, y, z]_{DOM} \ \& [t_{hit1}, t_{hit2} \dots t_{hitn}]$$

$$[x, y, z]_{DOM} \ \& \min(t_{hit1}, t_{hit2} \dots t_{hitn})$$

&