Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Reconstruction and Identification of Atmospheric Neutrino Events at JUNO Using Machine-Learning Methods

Milo Charavet[1] – On behalf of the JUNO Collaboration

[1] Univ. Hamburg – Inst. für Experimentalphysik

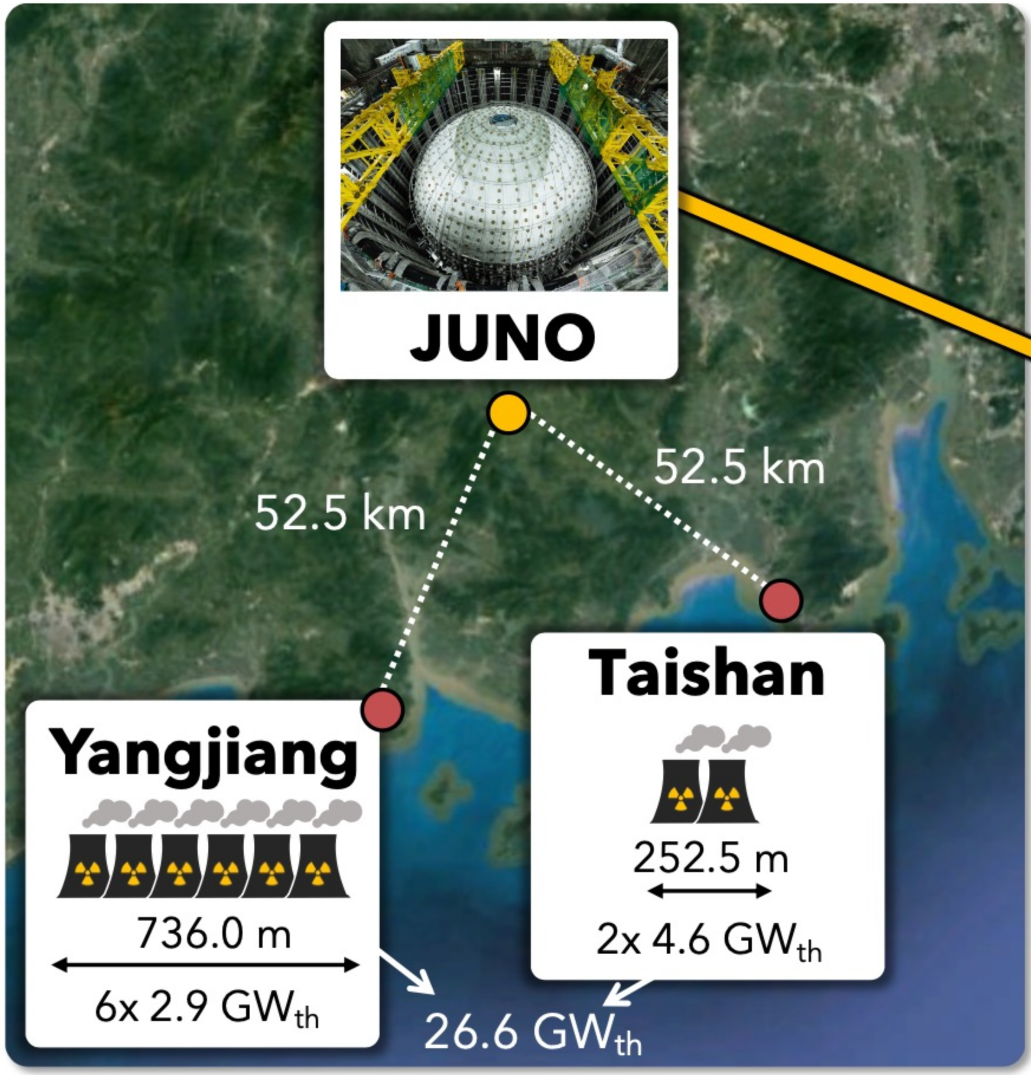
NPML 2025– University of Tokyo, 27/10/2025

Table of Contents

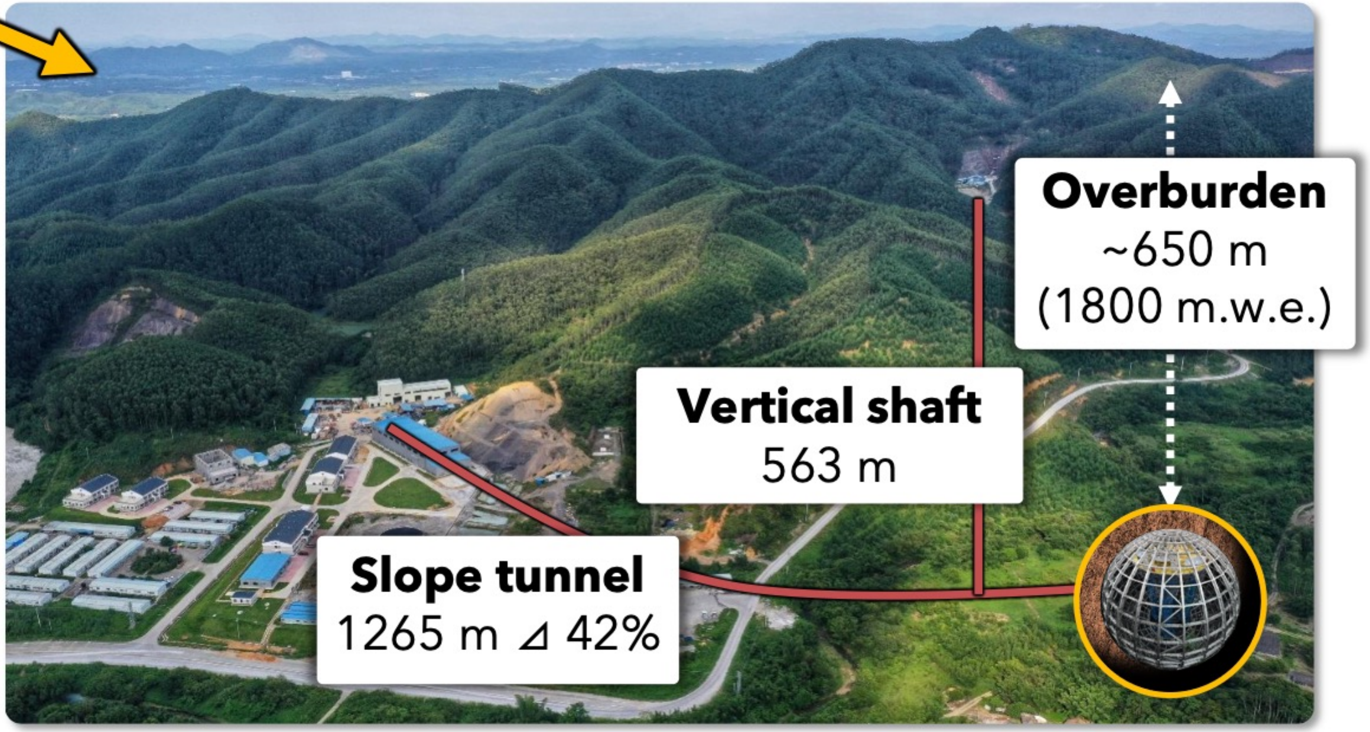
Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Table of Contents

JUNO Overview



- Jiangmen **U**nderground **N**eutrino **O**bservatory: next-generation multipurpose Liquid Scintillator (LS) detector with 20 kton target mass
- Locates at baseline 52.5 km from nuclear reactors



JUNO Overview

Central Detector (CD)

- 20 kton of liquid scintillator
- 17612 20-inches large-PMTs and 25600 3-inches small-PMTs ensure a 78% photo-coverage
- Earth's magnetic field compensation coils
- Unprecedented energy resolution: ~3% @ 1 MeV

Water Cherenkov Detector (WCD)

- 35 kton of high pure water as shield
- 2400 20-inches large-PMTs for active veto

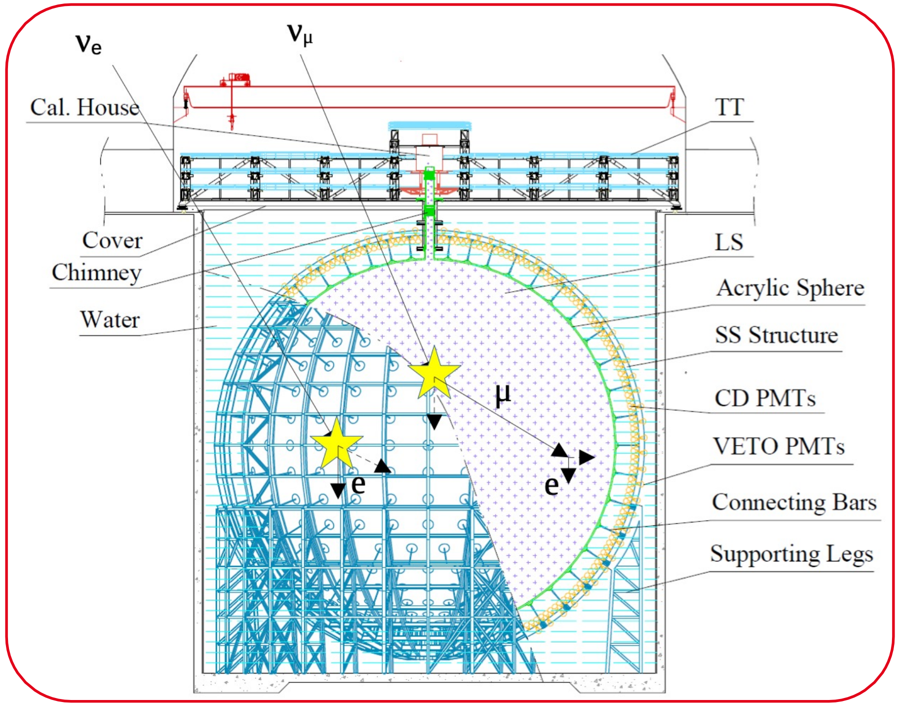
Top Tracker (TT)

- 3 plastic scintillator layers (cover ~30% of muons)

Calibration system

- More than 6 sources + laser

PPNP 123 (2022): 103927



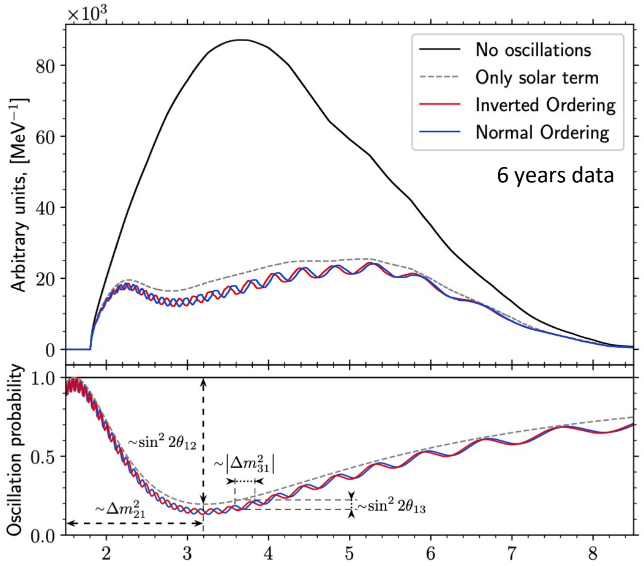
JUNO Physics

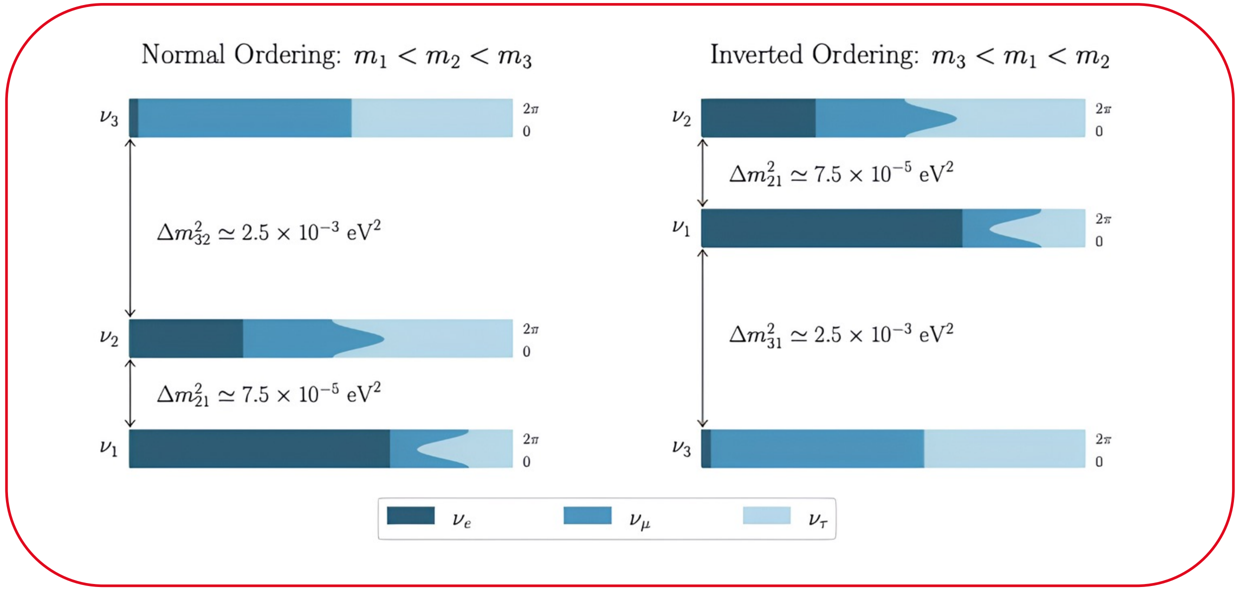
- Neutrino Mass Ordering (NMO) measurements
  - Reactor will determine NMO with 3σ significance in 6 years of data taking
  - Atmospheric neutrino: combined analysis with reactor further improve NMO sensitivity

- Precision measurement of oscillation parameters
  - for $\sin^2\theta_{12}$, Δm^2_{21} and $|\Delta m^2_{32}|$ => world leading precision in 100 days and precision <0.5% in 6 years

- Rich physics program with neutrinos from several sources
  - Solar neutrinos
  - Supernovae
  - Geo-neutrinos
  - BSM physics

  ...





PPNP 123 (2022): 103927

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

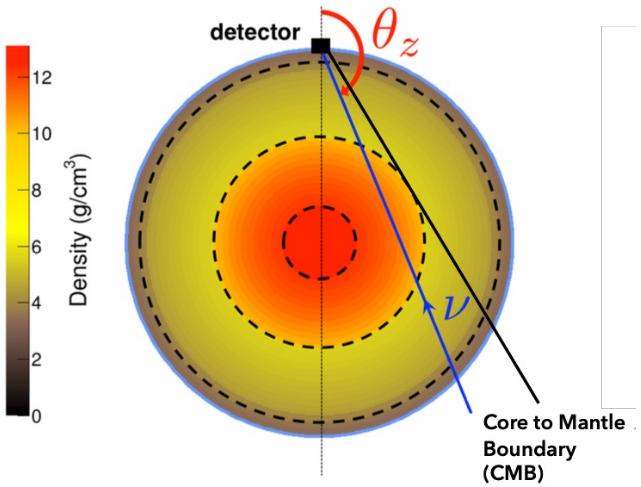
Table of Contents

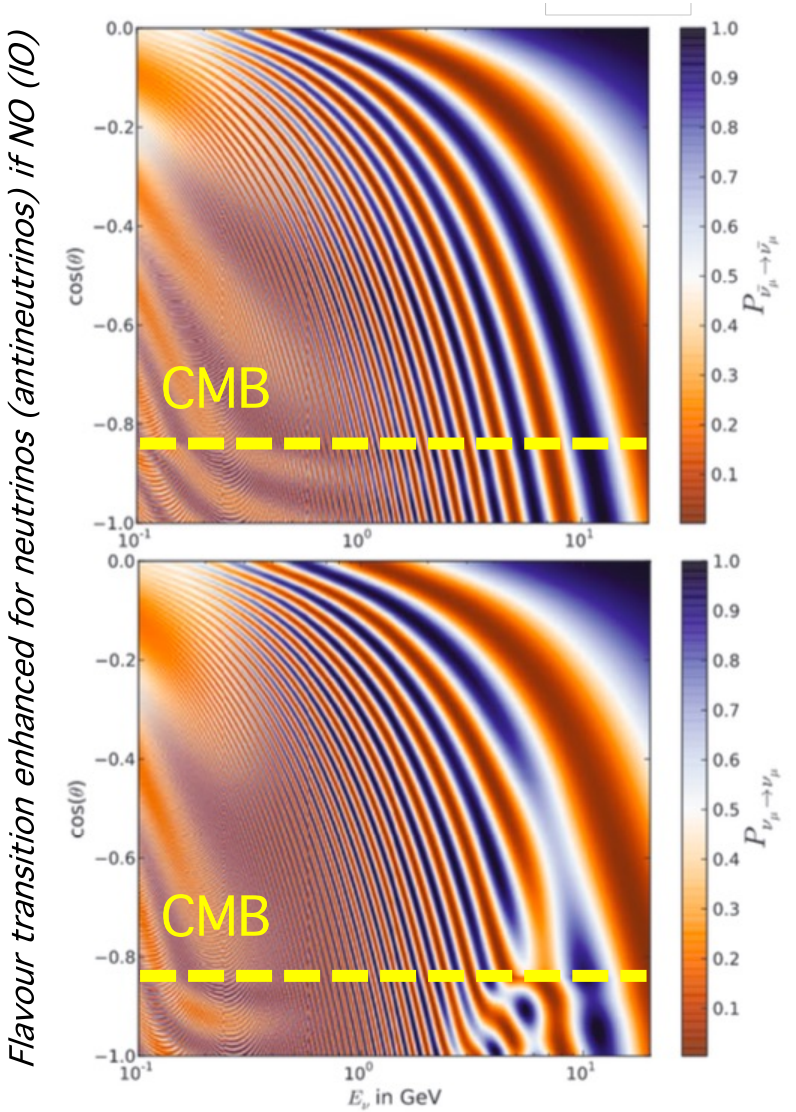
Table of Contents

Atmospheric neutrino

- Neutrino oscillation probabilities are function of neutrino energy, path length traversed, flavour identity and density of the medium
- Multi-GeV neutrinos undergo matter effect (MSW effect) while passing Earth's matter
- Atmospheric multi-GeV neutrino and antineutrino passing through Earth offer complementary channel to measure NMO

- Neutrino oscillation probability $P = f(L/E), L \sim \cos\theta$, depends on the neutrino energy, incident zenith angle θ and flavour
- **Neutrino energy, flavour and incoming angle need to be reconstructed.**

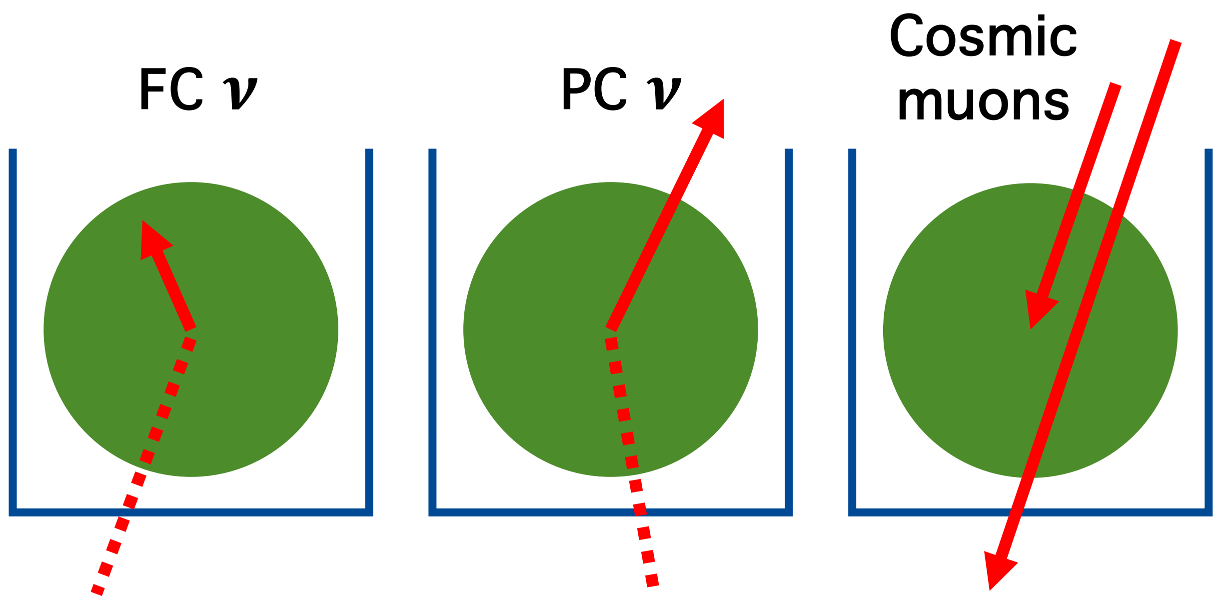


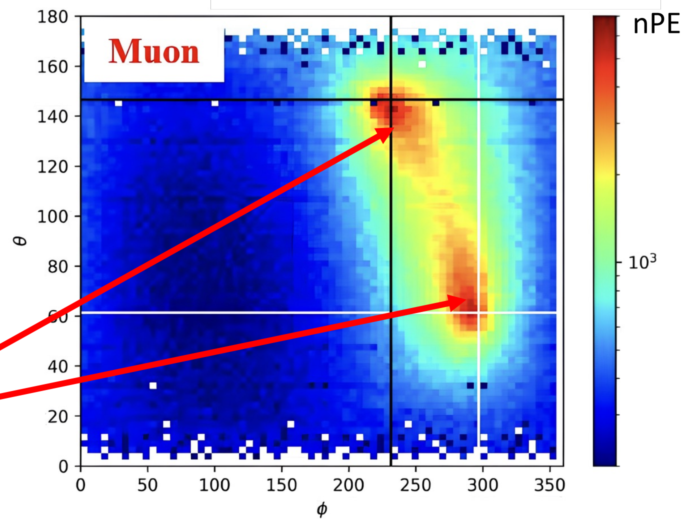


Atmospheric neutrino vs Cosmic muons

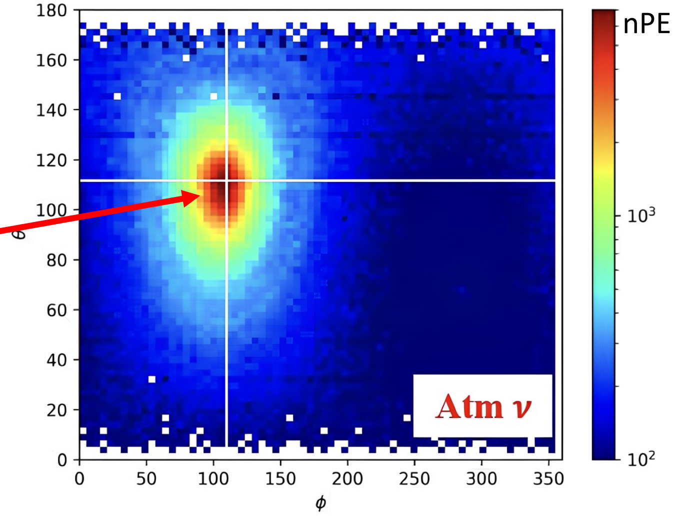
- Around 650 m rock overburden suppress muon background
- Expected muon rate ~ 5 Hz ; Neutrino interactions in JUNO LS ~ 10/day
- Correlation between CD and WP and TT is used to reject muons
- Remaining muons can be removed using PMT features (charge and time patterns)

FC : Fully Contained / PC : Partially Contained





Two red area correspond to an entry and an exit points of muon

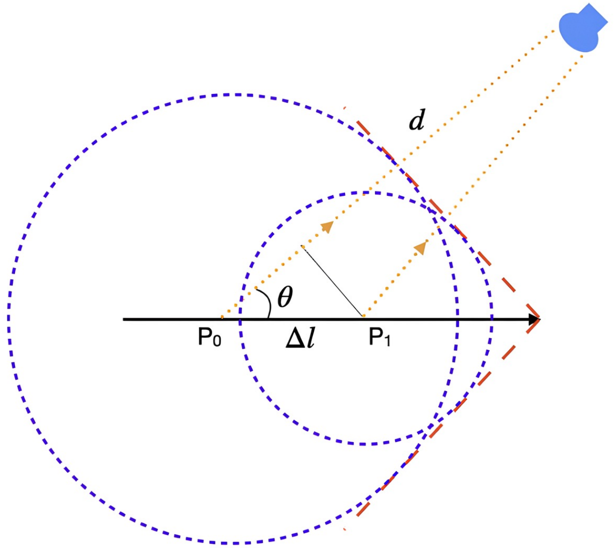


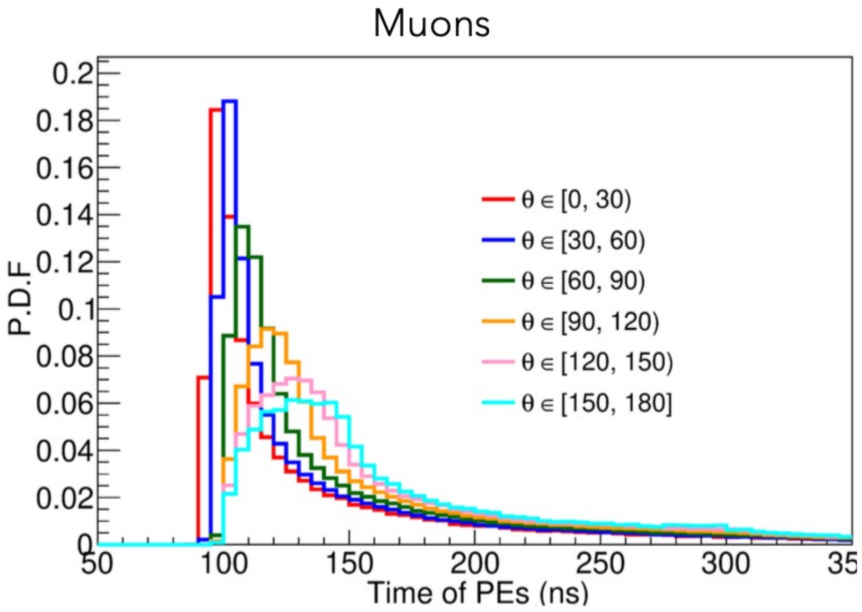
FC atmospheric neutrino only have one high nPE patch

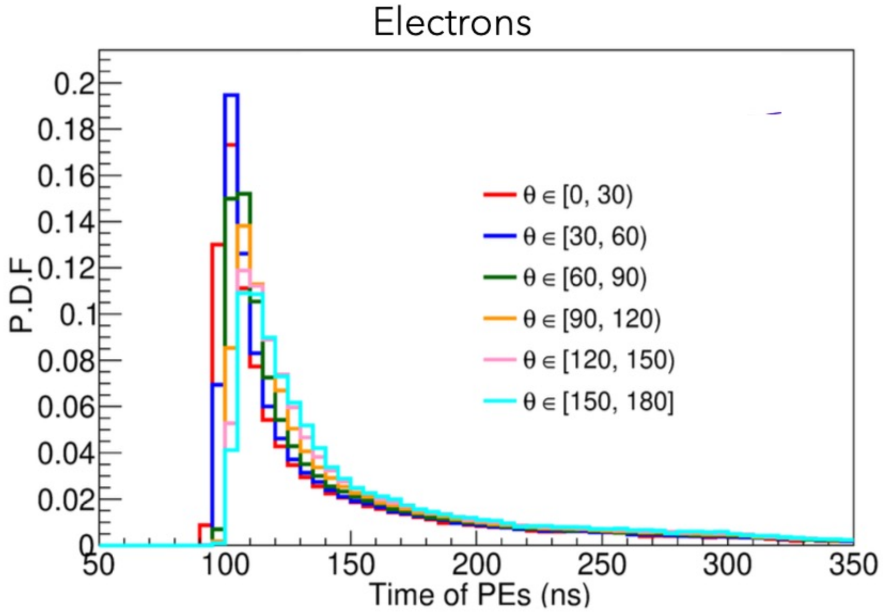
UHH Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Scintillation light at the detector

- Detected light is superposition of photon emissions along the particle track
- No direct track information
- Isotropic scintillation dominating over directional Cherenkov light
- Waveform received by each PMT will depend on the angle w.r.t. to the track direction, the track starting and stopping point and the particle type (inducing different dE/dx)







arXiv:2503.21353

Table of Contents

Table of Contents

Systematic Uncertainties

$$N\left(E_{\text{rec}}, \cos\theta_{\text{rec}}\right) \sim \int_{E_v}\int_{\cos\theta_v} \mathrm{d}E_v\, \mathrm{d}\cos\theta_v\, \Phi\left(E_v, \cos\theta_v\right) P_{\text{osc}}\left(E_v, \cos\theta_v\right) \sigma\left(E_v\right) R_{\text{det}}\left(E_v, \cos\theta_v, E_{\text{rec}}, \cos\theta_{\text{rec}}\right)$$

Number of events (Observables)

Neutrino Flux
Honda flux with normalization and shape uncertainties

Oscillation Probabilities
Oscillation parameter uncertainties from PDG

Cross-sections
Different generators and models (GENIE, NuWro, GiBUU, NEUT) And ReWeight

External data constraints (MiNERvA, T2K, NOvA…)

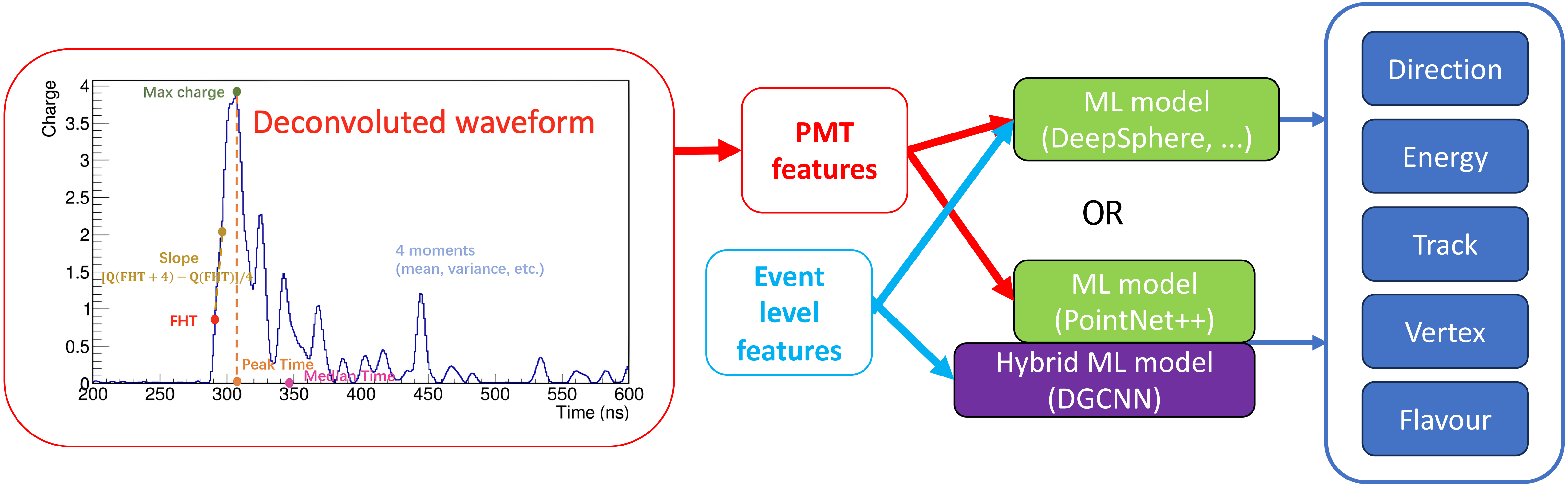
Detector responses
Cross-check with different ML models and non-ML methods

- A dedicated group is established to study the neutrino interactions for JUNO's atmospheric neutrino studies

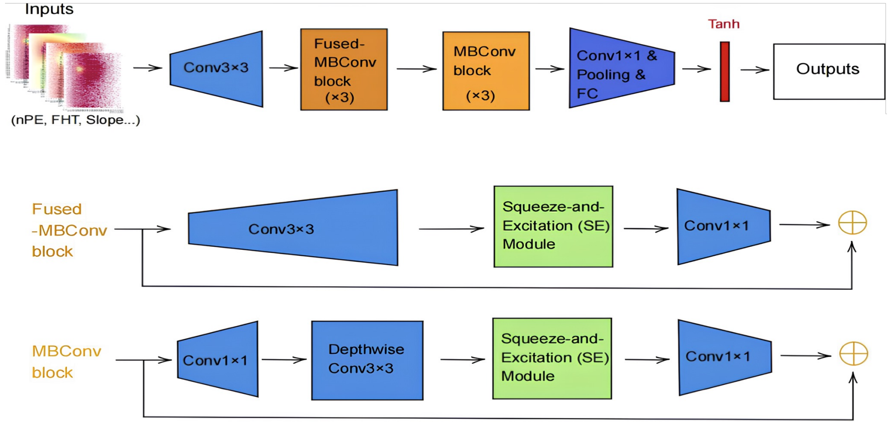
From Q. Yan talk at NuFact 2025

Methodology

- Event reconstruction with **D**eep-learning and **W**aveform **IN**formation (**EDWIN**)
- Current machine learning approach to find neutrino directions from feature patterns through training
- Investigation on both PMT level features and event level features (neutron multiplicity, n-capture, ...)

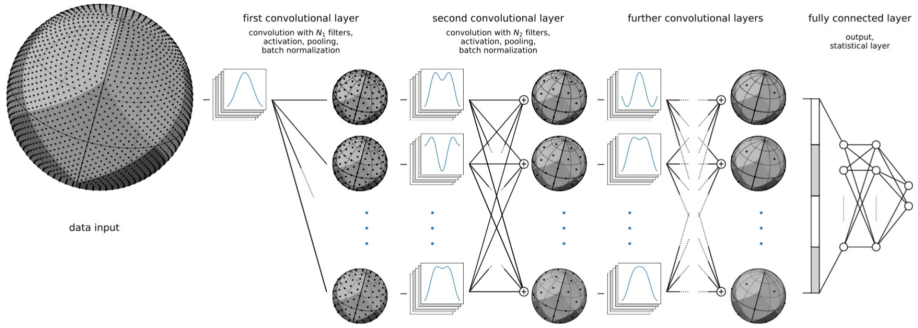


Machine Learning Model



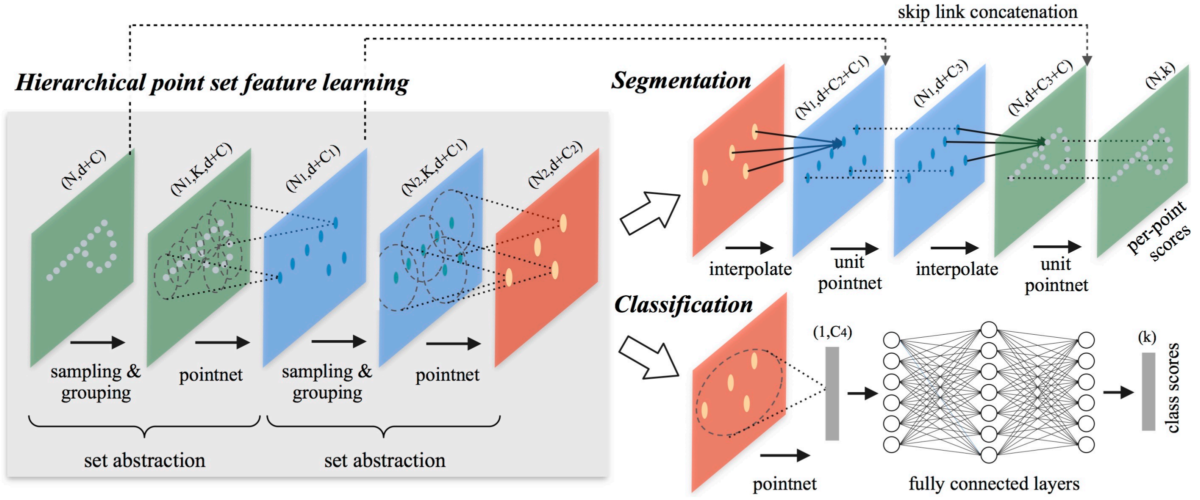
EfficientNetV2

- CNN model adapted for spherical data by projecting it onto a 2D grid
- High performance for short training time
- PMTs are seen as pixels, with each feature projected from the sphere to the planar surface



DeepSphere

- Graph-CNN: developed for processing spherical data originally developed for cosmology studies
- Maintains rotation co-variance and avoid distortions caused by projection to planar surface

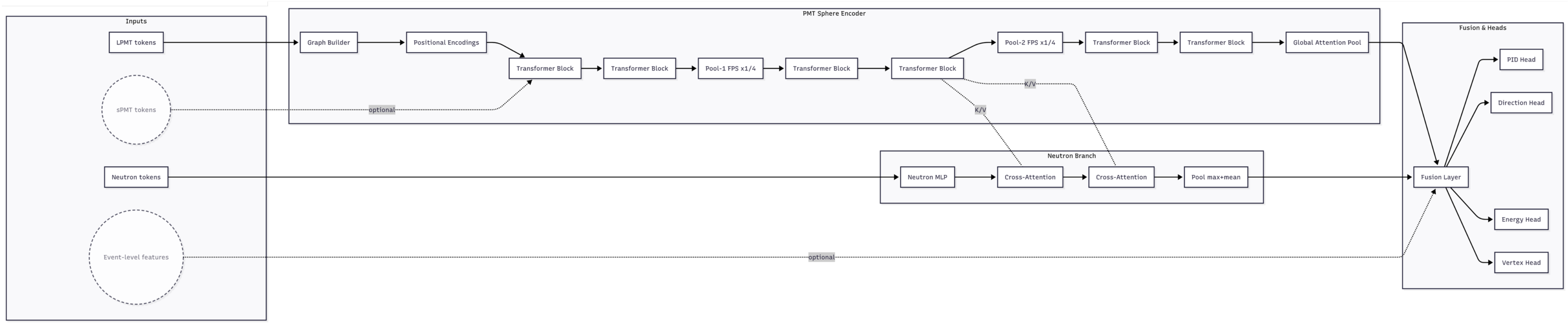


PointNet++

- Directly taking 3D point clouds as inputs
- Captures both global and local features.
- Detector signal more resemble point clouds
- Minimise information loss during projection

Original Model

- **ORION** (**O**n-sphere **R**econstruction with **I**nteracting t**O**kens & **N**eutrons) is a physics-aware spherical Transformer for JUNO atmospheric ν

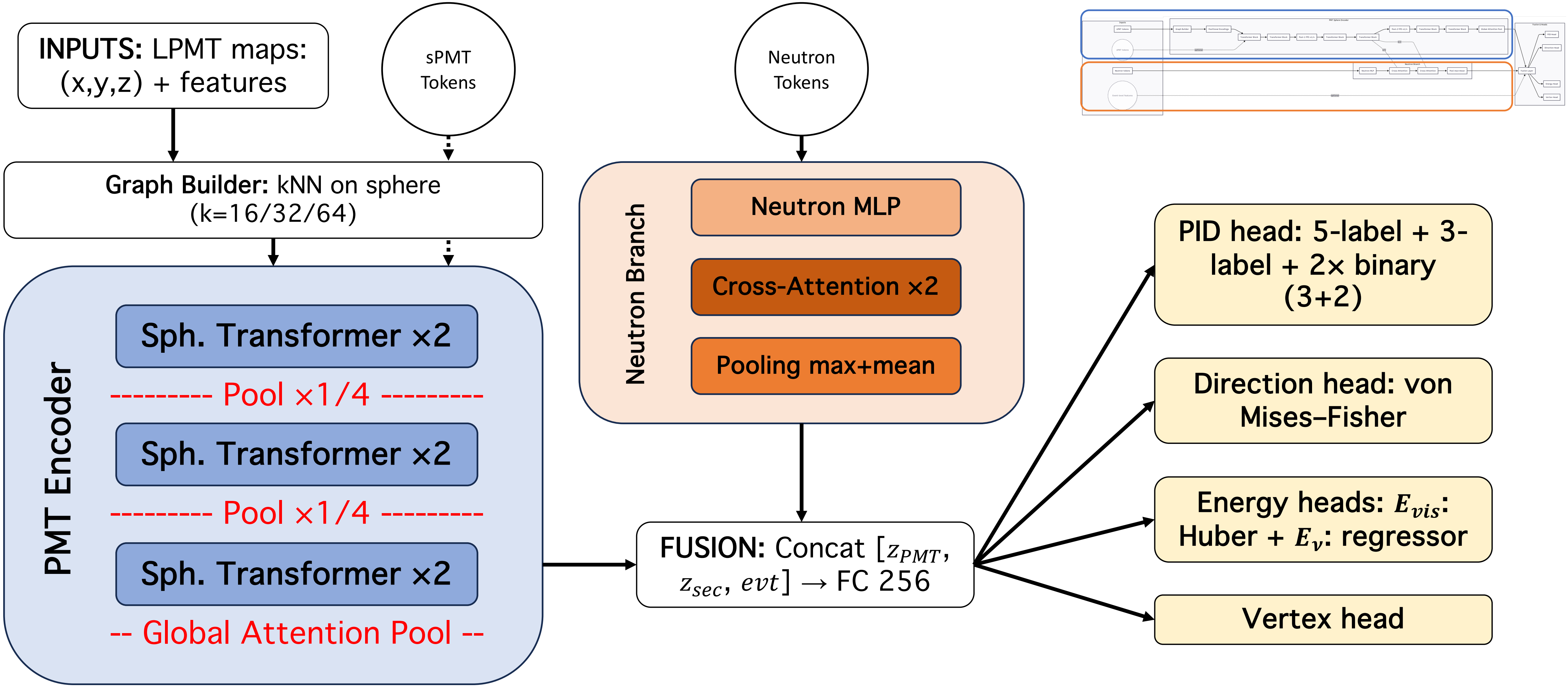


- Physics-aware on the sphere. ORION treats the PMT array as a spherical graph and uses attention guided by detector geometry and timing – no 2D projection

- Fuse delays as tokens. Delayed activity (neutrons/Michels) is handled as separate tokens that interact with the prompt PMT field, rather than being merged or tacked on at the end

- One model, many tasks. A single fused representation drives PID, direction, and energy/vertex reconstruction, trained jointly for consistency and robustness

Simplified Architecture



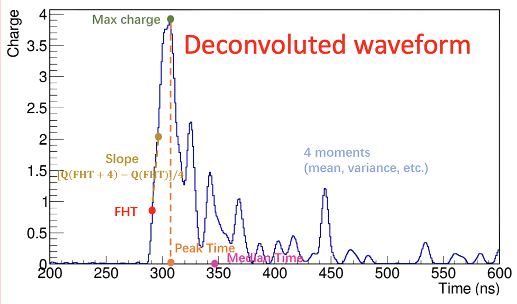
Updated Workflow

- **ORION** (**O**n-sphere **R**econstruction with **I**nteracting t**O**kens & **N**eutrons) is a physics-aware spherical Transformer for JUNO atmospheric ν

Pipeline

Deconvolution

gain-aware thresholds, time-offset unify



EDWIN per PMT features

- nPE
- FHT
- Slope/nPE
- Slope4/nPE
- MaxSlope
- 0–4ns ratio
- PeakCharge
- PeakTime
- mean/median/std/skew/kurt
- PMT type; coords (x,y,z).

Waveform AE / VQ-VAE

8–16D embedding per PMT

Event file (HDF5/NPZ)

- LPMT matrix (17612×C)
- secondary tokens (neutrons; Michels later)
- optional event-level scalars

Training

Build multi-scale kNN graphs; SO(3) rotations, PMT dropout, timing jitter

Training

Output

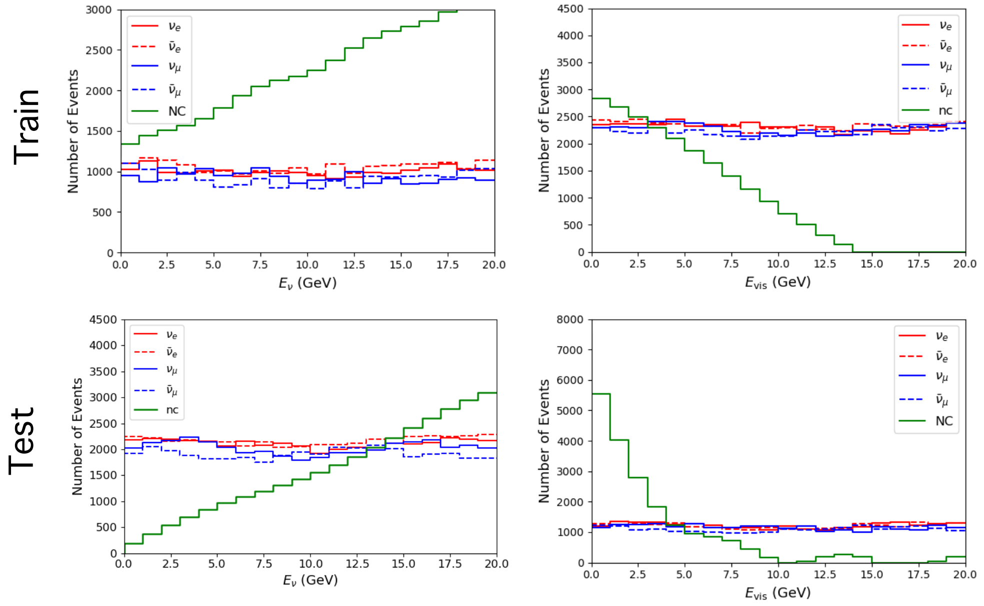
Table of Contents

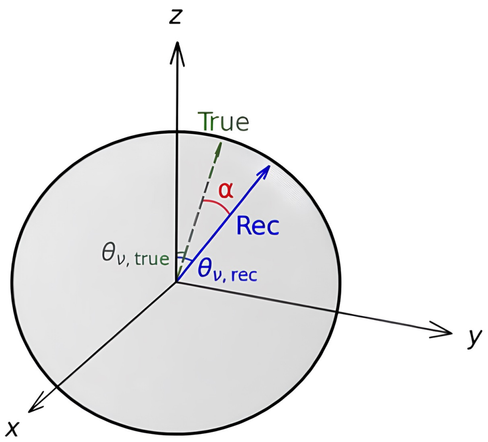
Table of Contents

Performance of direction reconstruction

MC data sample

- Statistic:
  - Training sample:
    - ~40k for each flavour ($\nu_\mu - CC$, $\nu_\mu - CC$, $\overline{\nu_\mu} - CC$, $\overline{\nu_e} - CC$)
    - Flat visible energy distribution [0,20] GeV (+ inclusion of additional NC) = each flavour has similar statistics
    - (avoid biases by energy shape dependence and uneven statistics among flavours for PID)
  - Testing sample: ~20k for each flavour





[Phys. Rev. D 112, 012018 (2025)]

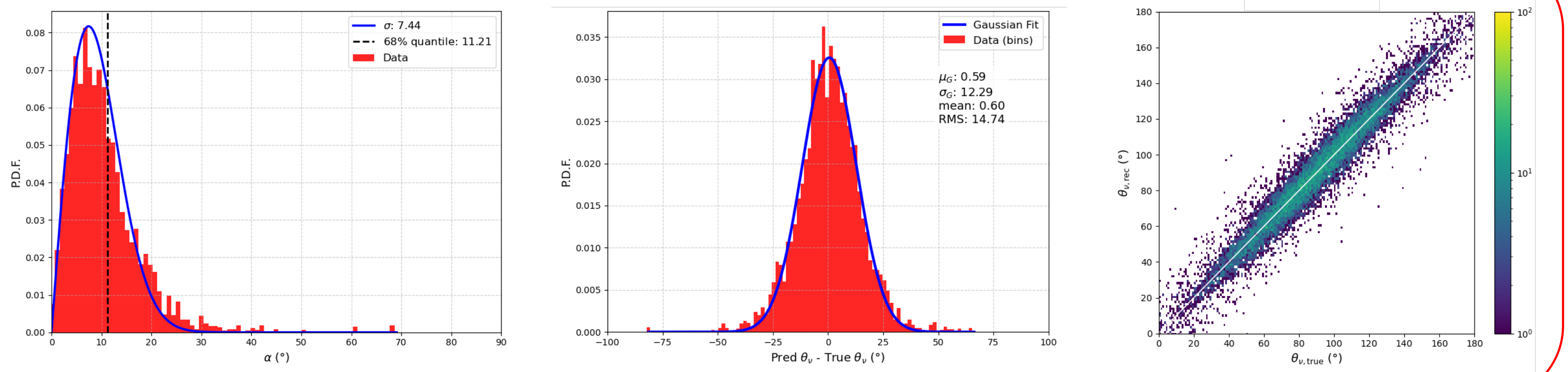
Directionality

- Angle between the true and the reconstructed directional vector is defined as $\alpha \in [0{,}180]°$ => performance quantified by 68% quantile
- Model performance for direction reconstruction evaluated over resolutions σ_α and σ_{θ_ν}

Performance of direction reconstruction

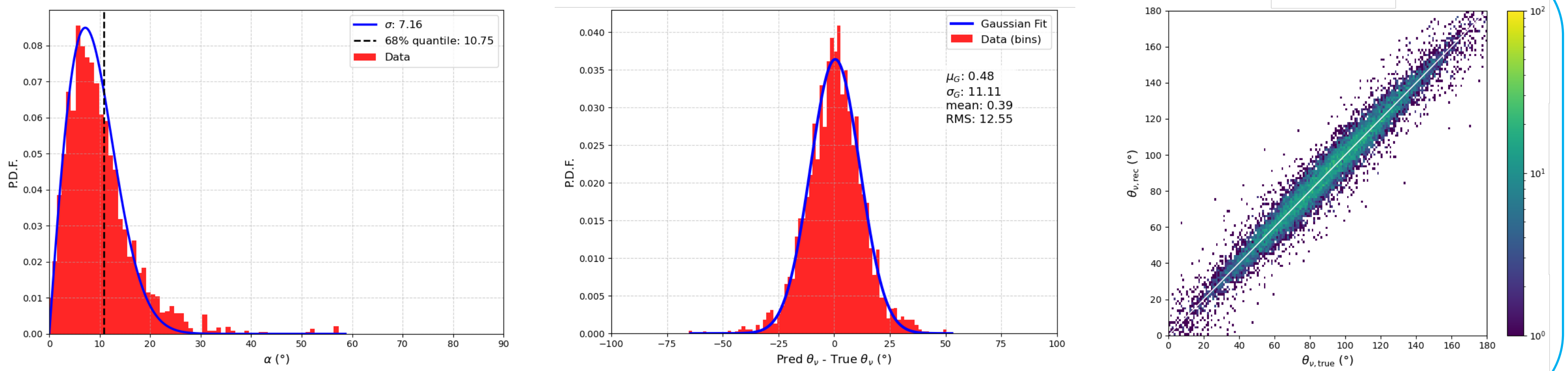
- **ORION**

$\nu_e/\overline{\nu_e}$
(4-5 GeV)

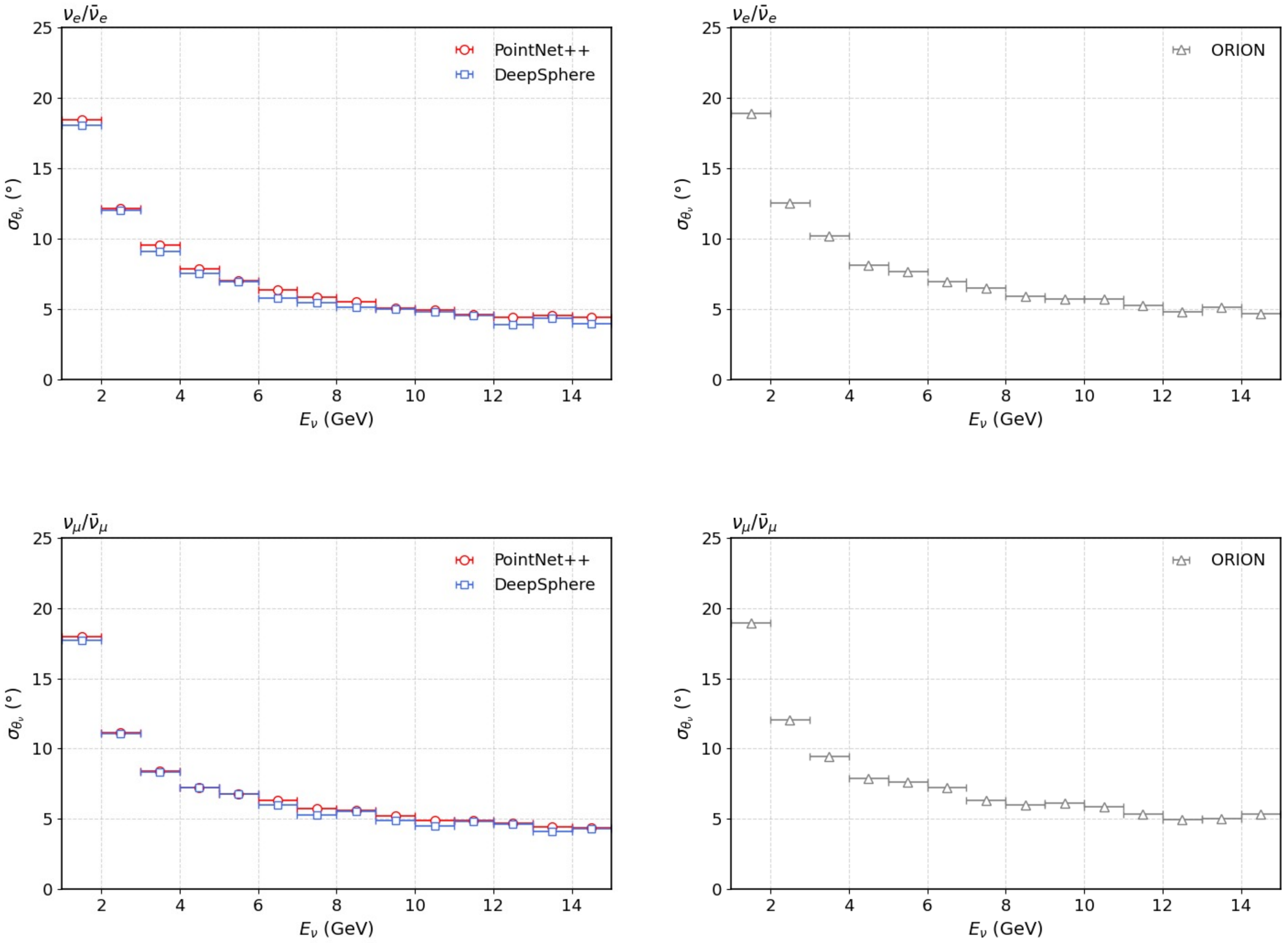


- $\delta\theta_\nu$ distribution within each 1GeV E_ν bin is approximately Gaussian.

$\nu_\mu/\overline{\nu_\mu}$
(4-5 GeV)

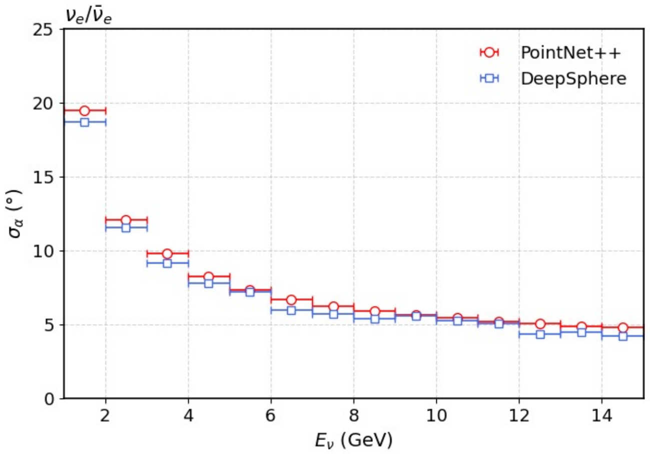


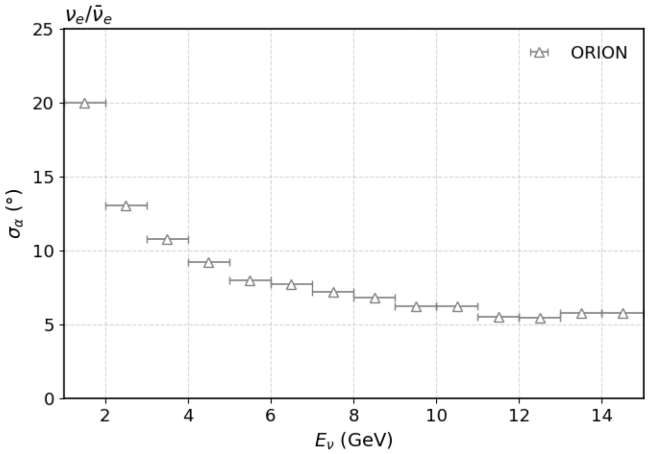
Performance of direction reconstruction

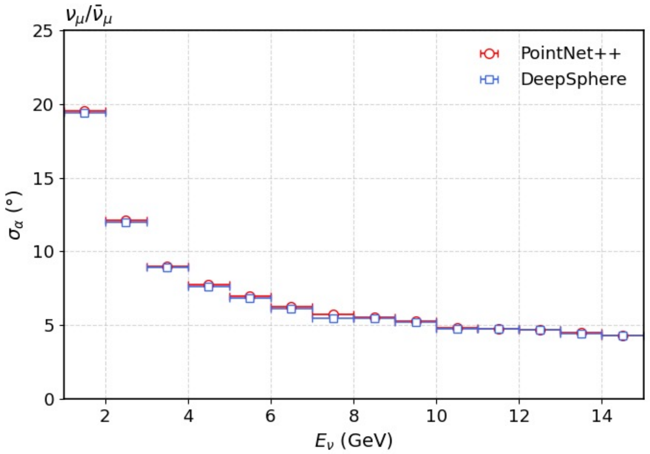


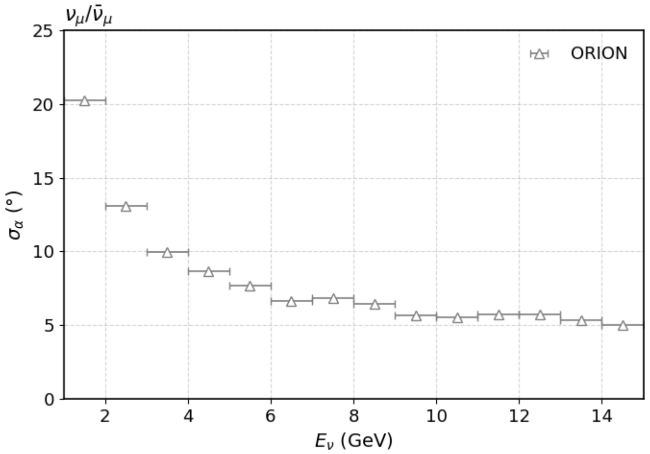
- θ_ν resolution
- Scintillation light from both leptons and hadrons are capable to directly reconstructing atmospheric neutrino direction.
- Angular resolution results with ORION are consistent with PointNet++ and DeepSphere for both flavour
- Slightly better result with PointNet++ and DeepSphere: Expect to improve ORION performances with tuning of Graph/attention hyperparams, ...

Performance of direction reconstruction









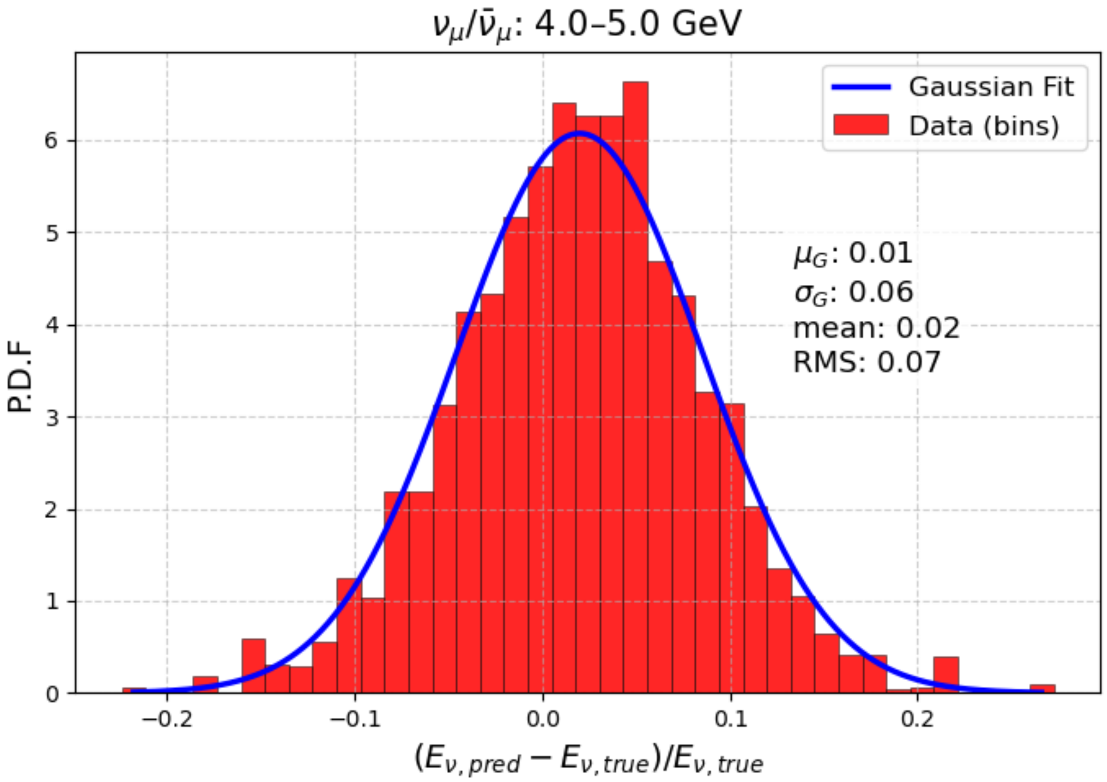
- α resolution
- Scintillation light from both leptons and hadrons are capable to directly reconstructing atmospheric neutrino direction.
- Angular resolution results with ORION are consistent with PointNet++ and DeepSphere for both flavour
- Slightly better result with PointNet++ and DeepSphere: Expect to improve ORION performances with tuning of Graph/attention hyperparams, ...

Performance of energy reconstruction

- ORION is aiming to have two energy heads trained jointly:
  - Visible energy E_{vis} - directly from the PMT representation.
  - Neutrino energy E_{ν} - a “physics-guided” head that takes the shared representation plus the PID logits and the neutron/Michel summary

- Each head uses heteroscedastic Huber loss and the multi-task weighting is learned

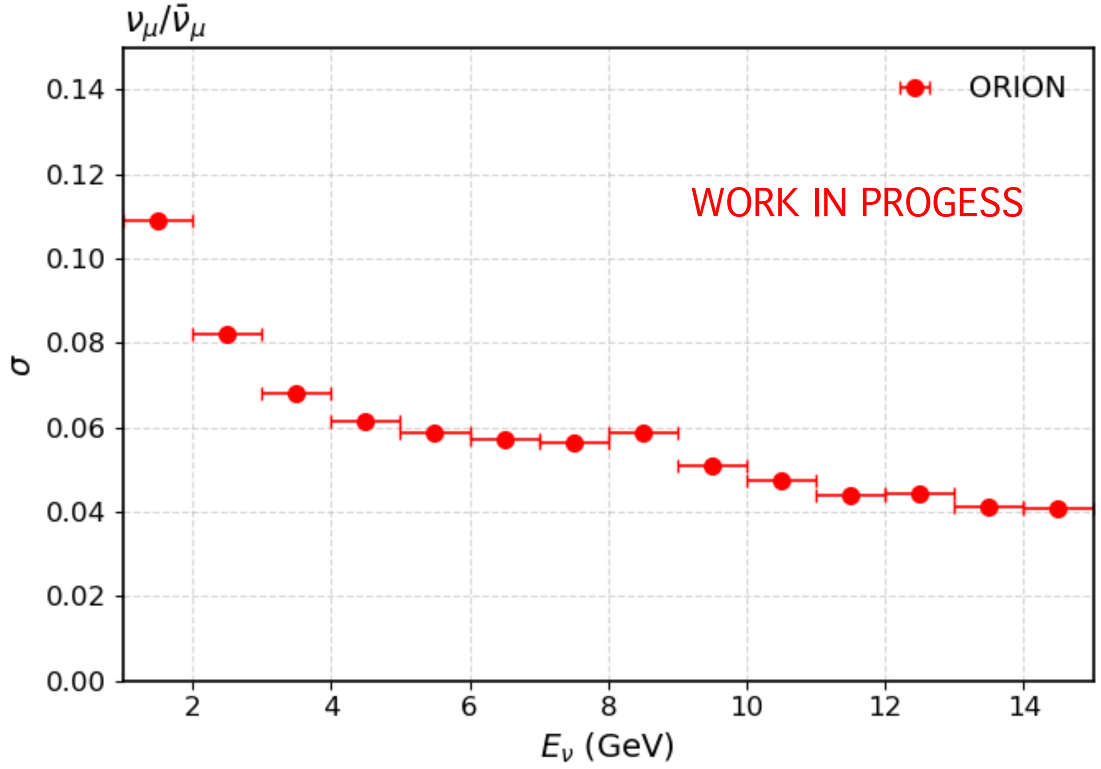
- For now, we report FC-CC events: we output E_{ν} as the primary estimate

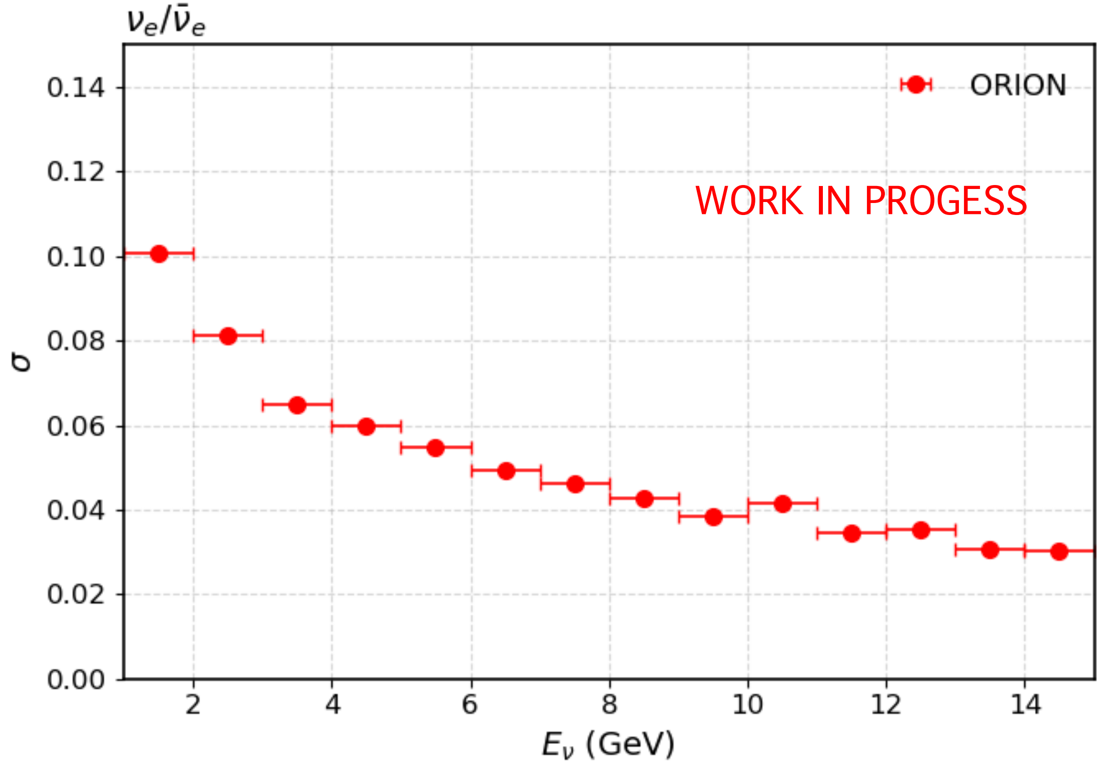


- $\left(E_{(\nu,pred)} - E_{(\nu,true)}\right)/E_{(\nu,true)}$ fitted with Gaussian

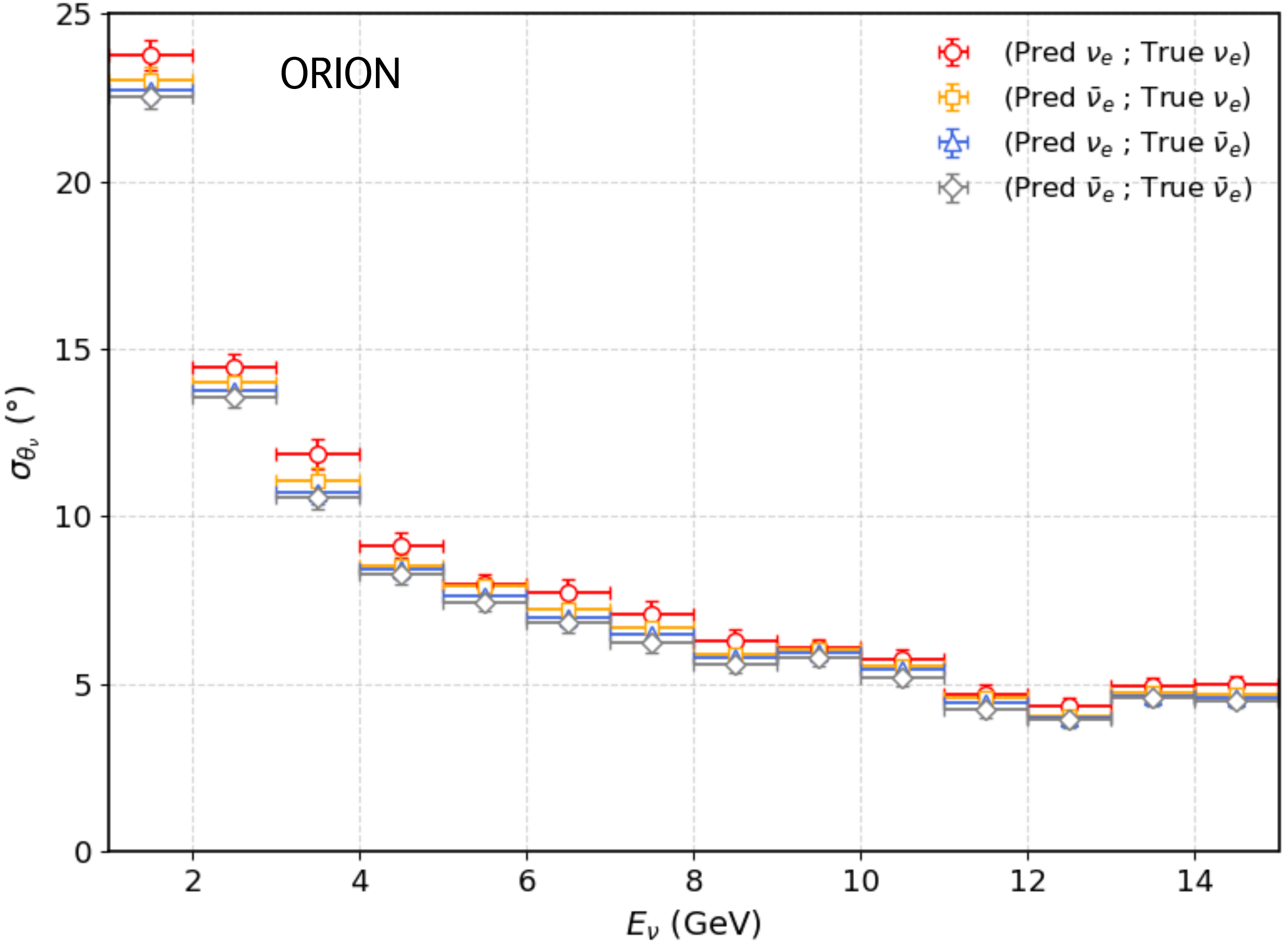
Performance of energy reconstruction

- Preliminary results
- For E_ν > 3GeV:
  - Better than 6% resolution for electron neutrinos (<5% for both PointNet++ and DeepSphere)
  - Better than 6% resolution for muon neutrinos (comparable to PointNet++ and DeepSphere)





Importance of PID



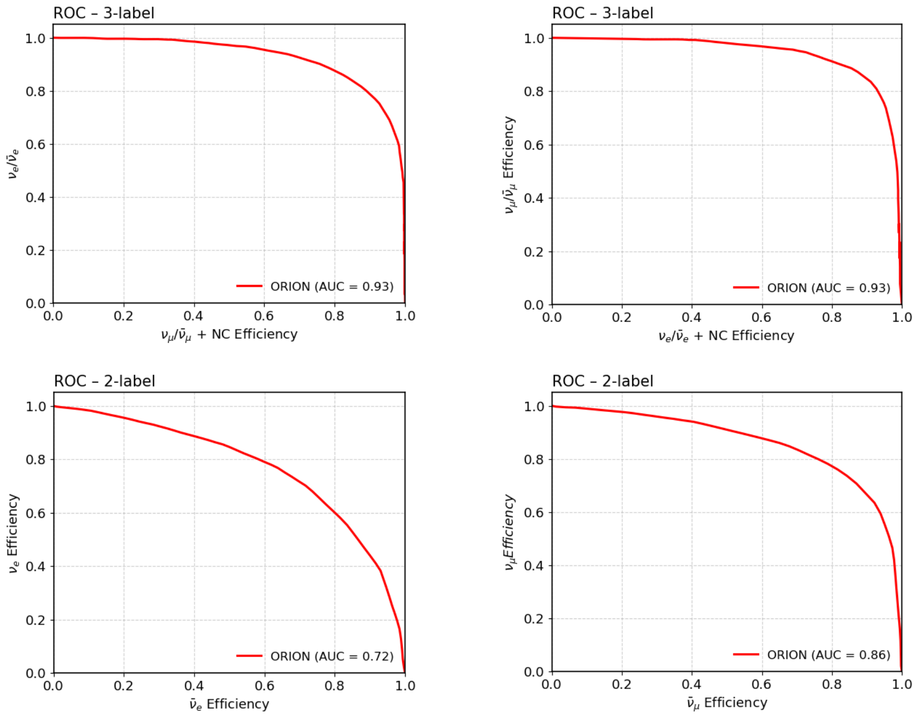
- Why PID is also important for directionality and energy reconstruction ?
- Strong correlations as $\nu/\bar{\nu}$ have different behaviour in direction and energy reconstruction
- Overall, $\bar{\nu}$-like events have better resolution

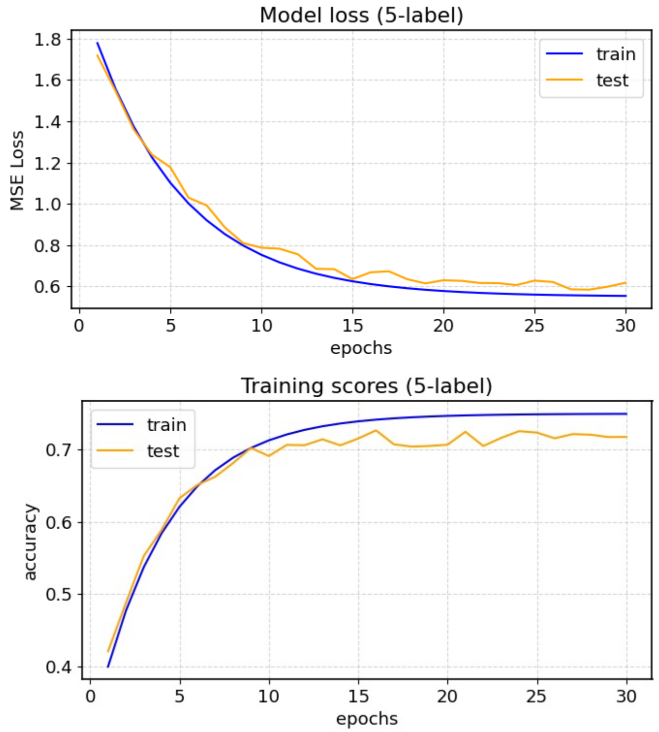
Performance of PID

- Event topology information are reflected in the PMT waveforms so we use PMT waveforms to classify: μ-like, e-like and NC-like
- Motivation: Direction/energy reconstruction and PID are correlated. The events identified as $\bar{\nu}$-like transfer less energy to hadrons (less hadronic interaction) in general, inducing a better resolution.
- 2 cross-check alternatives: 3-label + 2-label / 5-label

3-label

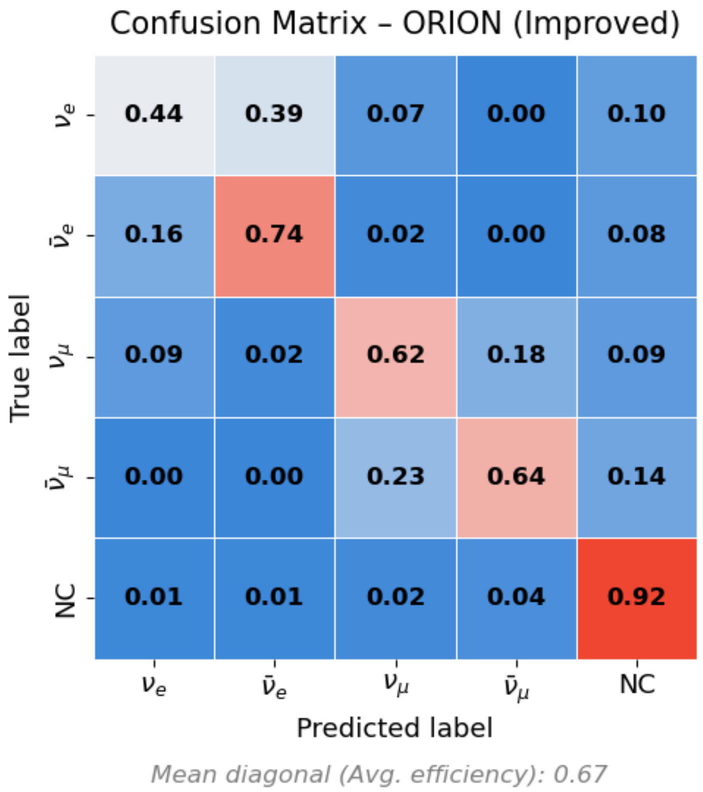
2-label

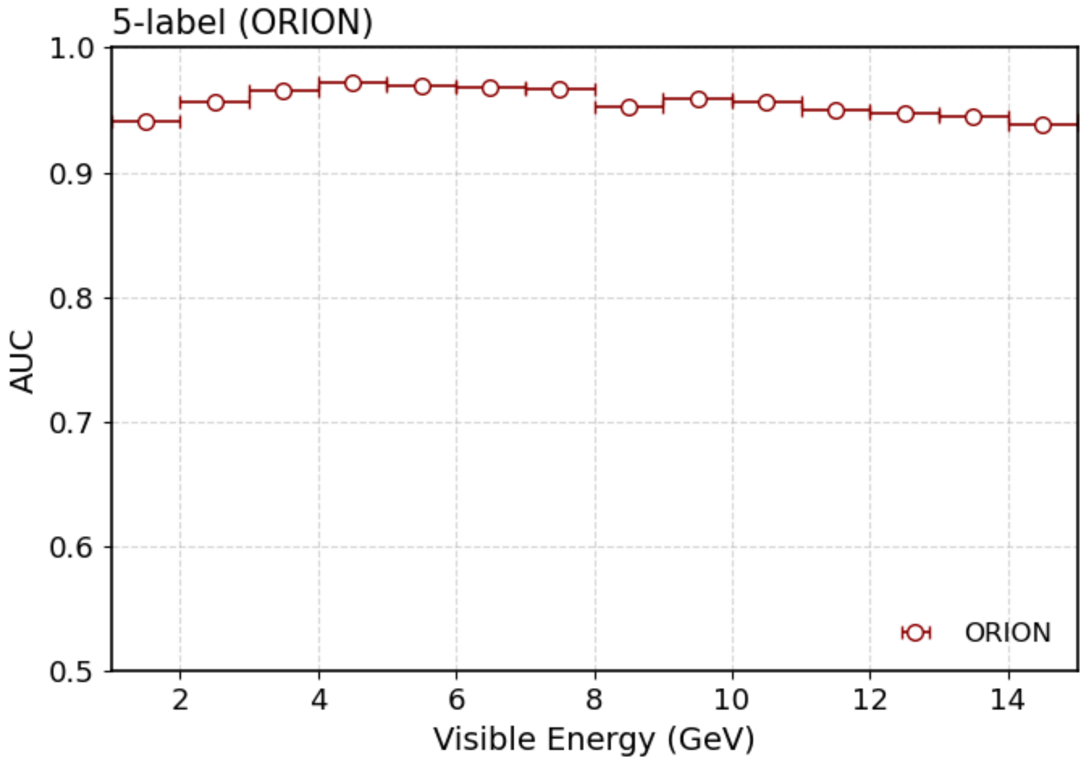




Performance of PID

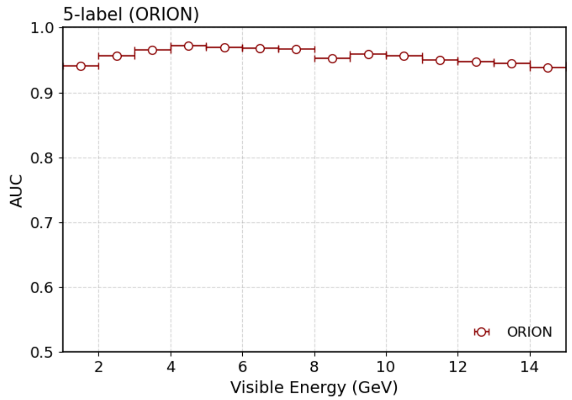
- Reminder: Only fully contained events
- 5-level classification is using both PMT features from prompt trigger and delayed trigger information
- For the oscillation analysis the score can be tuned depending on the requirement

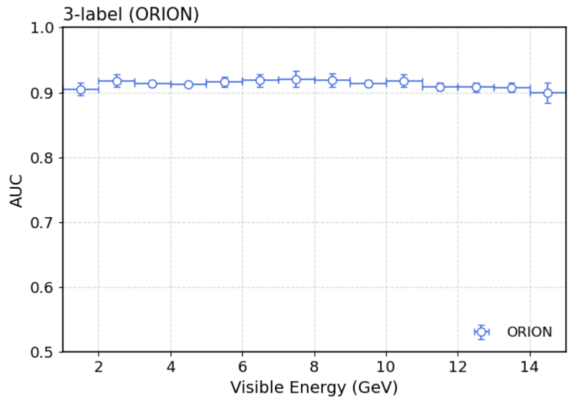




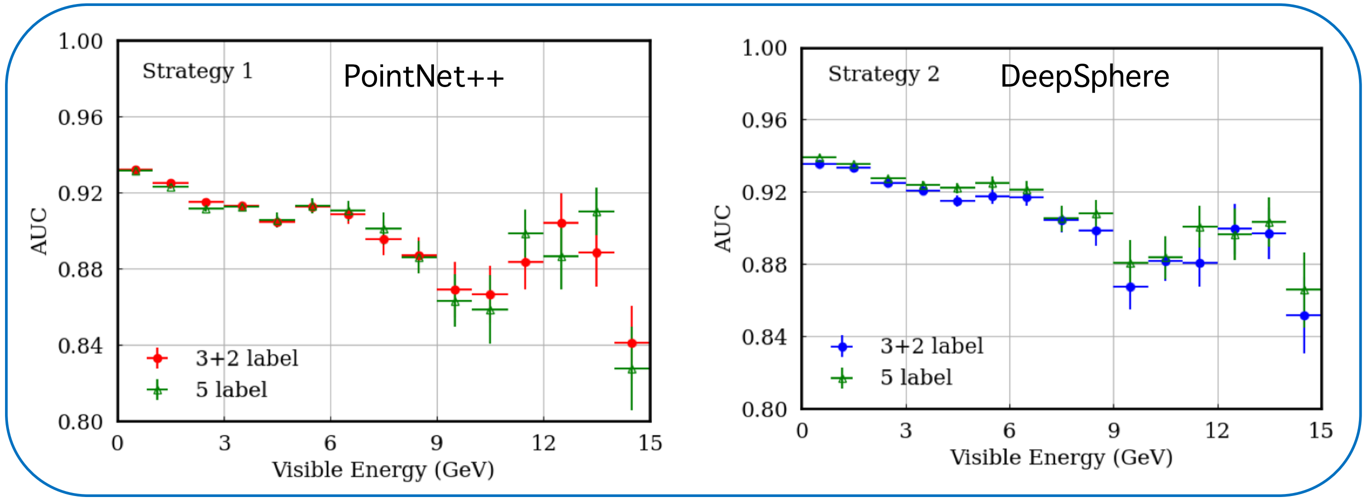
Performance of PID

- Reminder: Only fully contained events
- 5-level classification is using both PMT features from prompt trigger and delayed trigger information
- For the oscillation analysis the score can be tuned depending on the requirement





arXiv:2503.21353

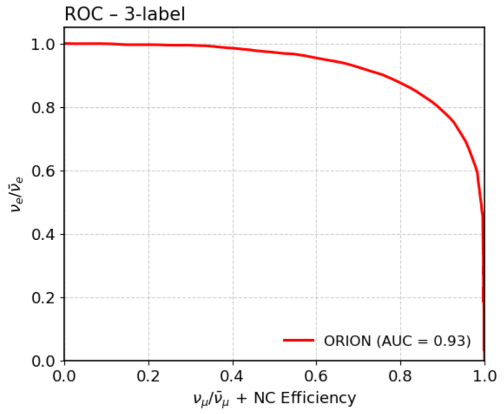


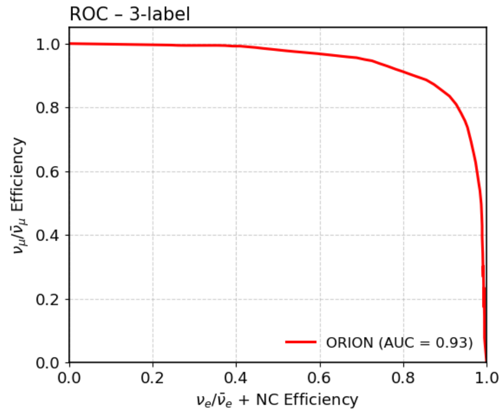
- The results are consistent between the 3 models.
- Strong difference for high energetic events because sample is different (Honda flux + ratio of neutrino different)

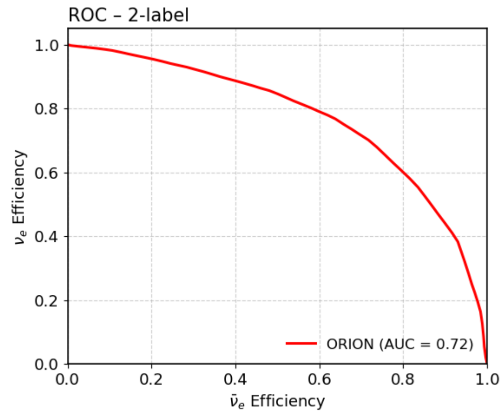
Performance of PID

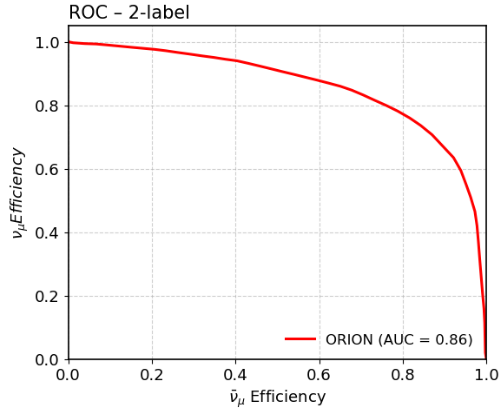
- Reminder: Only fully contained events
- 5-level classification is using both PMT features from prompt trigger and delayed trigger information
- For the oscillation analysis the score can be tuned depending on the requirement

arXiv:2503.21353









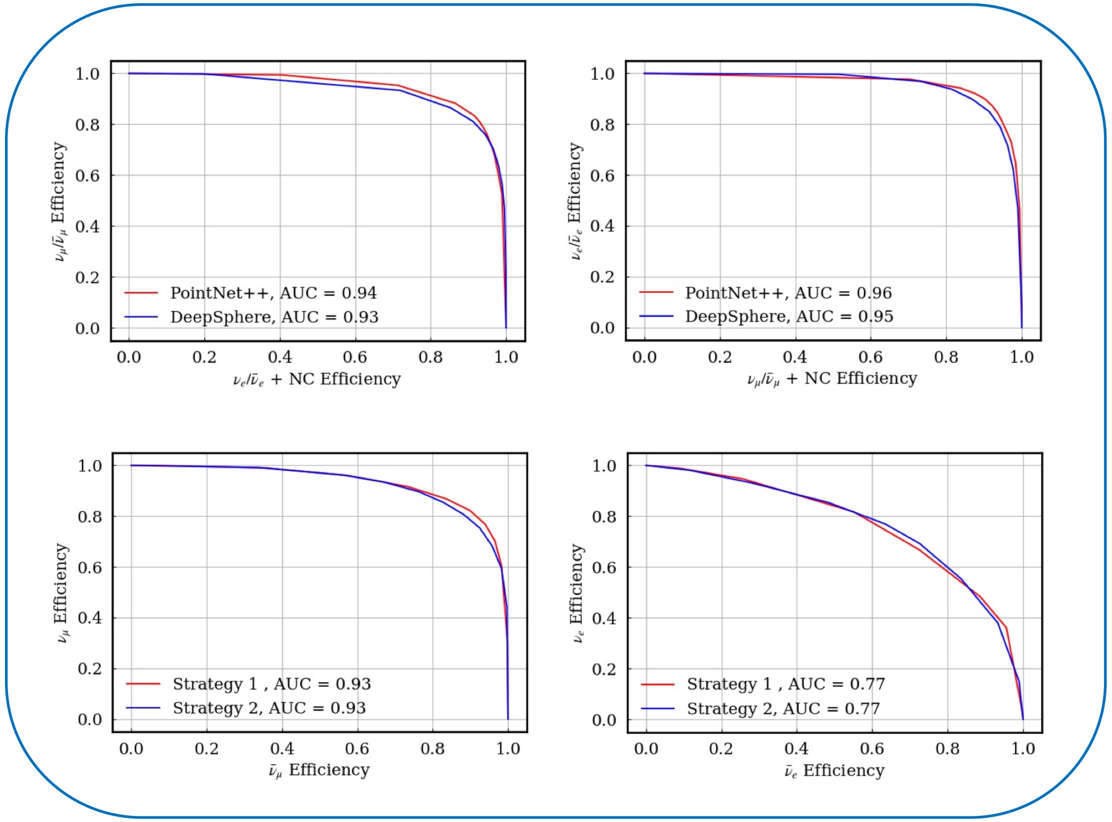


Table of Contents

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Table of Contents

Summary

➢ Improvements in detector response

	Optimistic estimate	Recent Improvements (Machine Learning)
Directionality	$\sigma_{\theta_\nu} = 10°$ $\sigma_{\theta_\mu} = 1°$	$\sigma_{\theta_\nu} < 10°$ for all ML models based on PMT features
Energy	$\sigma_{E_{vis}} = 1\%/\sqrt{E_{vis}}$	E_ν is reconstructed instead of E_{vis} (FC events): resolution < 6% for E_ν > 3GeV (even better for $\nu_e/\overline{\nu}_e$ with PointNet++ and DeepSphere)
Classification	NC/CC-ν_e/CC-ν_μ: 100%	70-95% efficiency for all ML models using PMT features based on primary and secondary triggers
Sensitivity	To be updated	

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Summary

- In this talk, we presented a machine learning approach for the reconstruction of atmospheric neutrinos in JUNO

- **JUNO aimed at precision oscillation physics and the neutrino mass ordering** - Atmospheric neutrinos provide :
  - Broad L/E and matter-effect leverage,
  - Independent cross-check to reactor results
  - Control samples (cosmics, Michels, spallation n) to validate reconstruction and systematics

- Here, we introduced ORION, a physics-aware spherical ML model that processes PMTs natively on the sphere and exploits delayed activity; competitive on energy, direction, and PID with other ML models for cross-check

- Outlook: JUNO has started taking data!
  - Real data integration & calibration
  - Robustness/systematics (optical/electronics, generators, θ-bias checks)
  - Reproducibility via shared baselines and standard metrics

Thank you for your attention!

Back-up slides

JUNO Timeline

Scheme from A. Garfagnini @ TAUP 2025

01/2025
Start of TT Installation

Water Phase

Mixed Phase

Full detector calibration

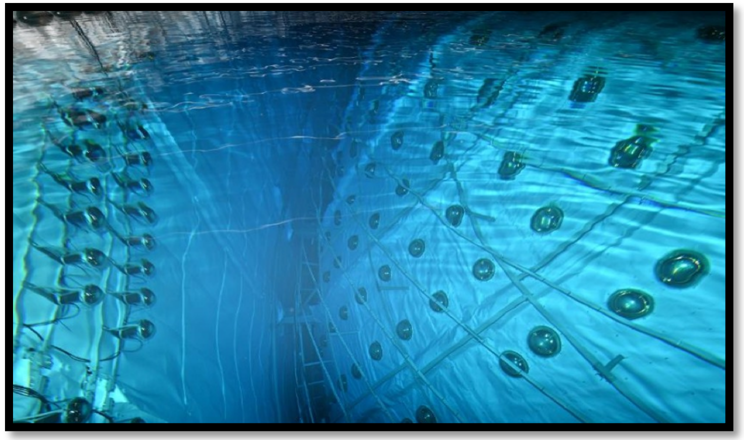
18/12/2024
Start of Water Filling

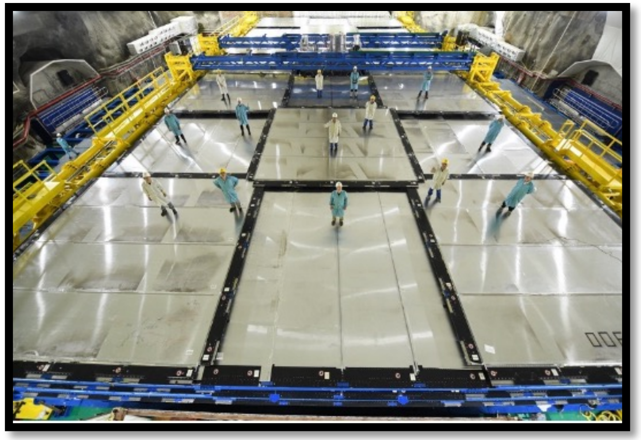
02/02/2025
Start of Water Filling

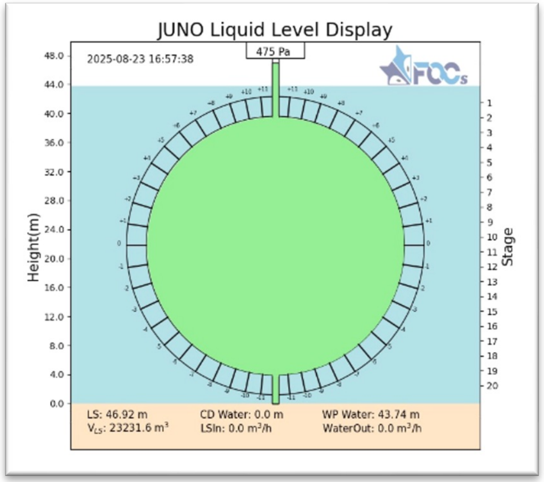
08/02/2025
Start of LS Filling

22/08/2025
End of LS Filling

26/08/2025
Start of Data Taking

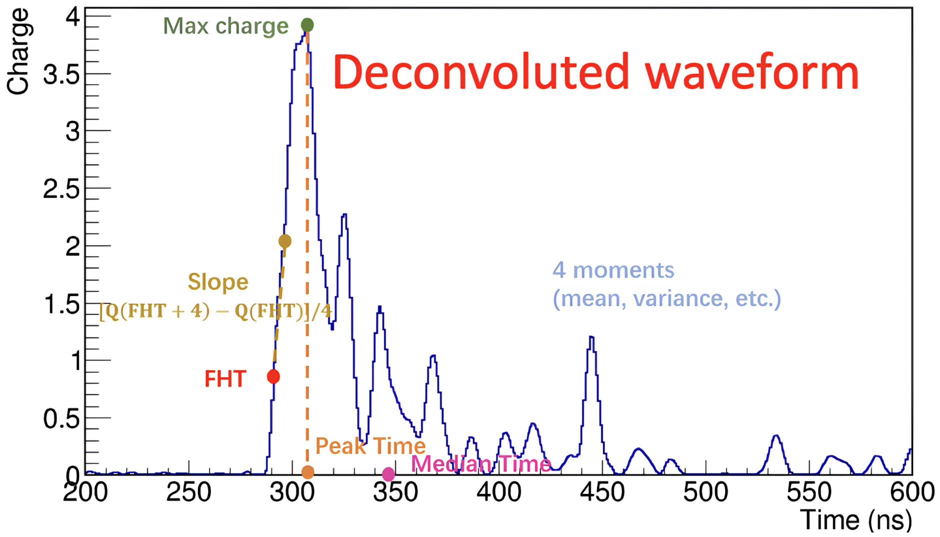






Methodology

- Event reconstruction with **D**eep-learning and **W**aveform **IN**formation (**EDWIN**)
- Current machine learning approach to find neutrino directions from feature patterns through training
- Investigation on both PMT level features and event level features (neutron multiplicity, n-capture, …)



- First Hit Time
- Total charge: The total number of PEs before electronic effects
- Charge ratio: Charges in the first 4ns divided by the total
- Slope: Describes the average slope in the first 4ns.
- nPE
- nPE slope
- nPE ratio
- Max charge, Peak Time

New strategy

- **ORION** (**O**n-sphere **R**econstruction with **I**nteracting t**O**kens & **N**eutrons) is a physics-aware spherical Transformer for JUNO atmospheric ν

- Native sphere + global attention with physics-biased edges (geodesic, bearing, Δt after TOF, PMT type).
- Learned waveform embedding (8–16D) per PMT in addition to EDWIN features.
- Tokenized secondaries (neutrons, Michels) fused by cross-attention instead of late concatenation / merge.
- Multi-task training with angle-aware vMF and heteroscedastic energy losses.

- Attention with physic-bias

$$\alpha_{ij}^{(h)} = \frac{q_i^{(h)} \cdot k_j^{(h)}}{\sqrt{d}} + g_\phi\left(\text{geoij}\,,\ \text{bearij}\,, \Delta t_{ij},\ \text{type}\,_{ij} \right), \quad z_i = \sum_j \text{softmax}\, j(\alpha ij) v_j$$

geodesic distance on the sphere

relative bearing

$(FHT_i - FHT_j)$ after a coarse TOF correction

- vMF direction loss

$$\mathcal{L}_{\text{dir}} = -\log C_3(\kappa) + \kappa\left(\mu^\top \widehat{\mu}\right), \|\widehat{\mu}\| = 1$$

Predict unit vector $\hat{\mu}$ and concentration κ for a von Mises–Fisher likelihood in $\mathbb{S}^2$

New strategy

"ORION's core operator is attention with a physics bias on the PMT sphere. For PMT token i, we compute a query vector q_{ij} for neighbor PMT j, a key k_j and value v_j. The standard attention logit is the scaled dot product $(q_i \cdot k_j)/\sqrt{d}$, where d is the head dimension. We add a small physics bias b_{ij} computed by a tiny MLP (multilayer perceptron) from simple scalars: the geodesic distance on the sphere $d_{\text{geo}}(i,j)$, the relative bearing ϕ_{ij}, the TOF-corrected timing residual $\Delta t_{ij}^{(10\text{ F})}$, and the PMT types (LPMT vs sPMT).

The full logit is

$$\alpha_{ij} = \frac{q_i \cdot k_j}{\sqrt{d}} + \underbrace{\mathrm{MLP}\left(d_{\text{geo}}(i,j), \phi_{ij}, \Delta t^{(\text{TOF})} ij,\ \text{type } i,\ \text{type } j\right)} \text{ physics bias bi } j.$$

We mask to a k-nearest-neighbors (kNN) graph by geodesic distance so attention is local and spherical. The softmax over neighbors turns logits into weights wij, and the head output is $\sum_j w_{ij} v_j$. Each Transformer block is: Pre-LayerNorm $\rightarrow$ Multi-Head Attention (MHA) + physics bias $\rightarrow$ residual $\rightarrow$ LayerNorm $\rightarrow$ GEGLU FFN $\rightarrow$ residual, where GEGLU is a gated GELU feed-forward that improves expressivity.

Delayed activity (neutron captures and Michel electrons) is handled with secondary tokens. Each capture becomes a token with features (capture time Δt, delayed charge $\sum \text{PE}$, optional position (x, y, z), quality flags). These tokens perform cross-attention into PMT tokens: the capture is the query, PMTs are keys/values. If positions are missing, we mask the geometry part of the bias so timing/charge still contribute. This keeps multiplicity and per-capture relations intact instead of compressing them away.

For direction, we use the von Mises-Fisher (vMF) likelihood on the unit sphere (spherical analogue of a Gaussian). The model outputs a unit vector $\widehat{\mu}$ and a concentration $\widehat{\kappa} \geq 0$ (confidence). The per-event negative log-likelihood is

$$\mathcal{L}_{\text{v}MF} = -\log C_3(\hat{\kappa}) - \hat{\kappa}\hat{\mu} \cdot \mathbf{u},$$

with u the true direction and $C_3(\hat{\kappa}) = \widehat{\kappa}/(4\pi \sinh \widehat{\kappa})$. Intuitively, higher $\widehat{\kappa}$ means a narrower cone (more confidence).

For energies we use heteroscedastic Huber losses: each head predicts a mean and a log-variance (event-wise uncertainty). Huber is robust to outliers; the learned variance down-weights ambiguous events. We also add a light penalty discouraging unphysical $E_\nu < E_{\text{vis.}}$

All tasks are combined via uncertainty weighting (learned loss scales per head) so no single objective dominates training. Abbreviations: MHA=multi-head attention, MLP=fully connected stack, GEGLU=gated GELU feed-forward, TOF=time-of-flight correction."

Architecture

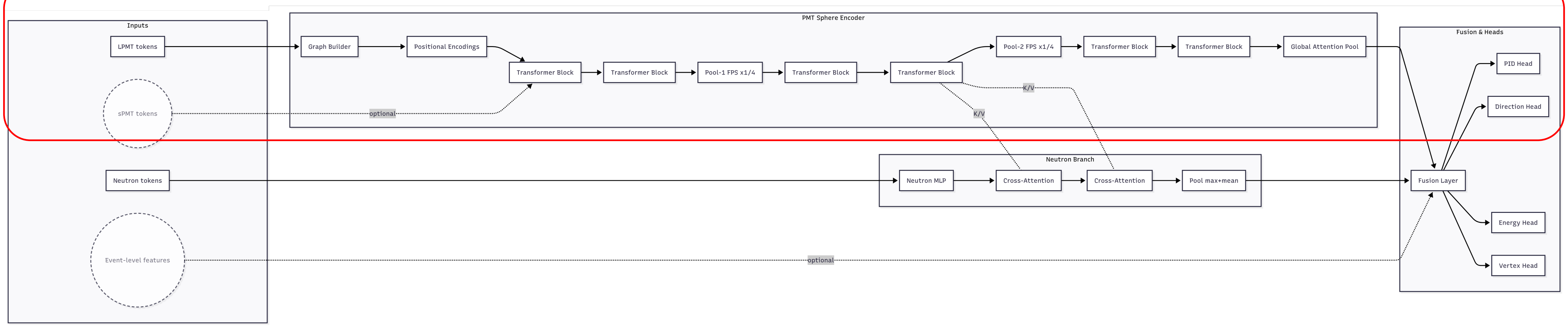
INPUTS

- PMT tokens: active LPMTs $N = 3 - 8k$; features = (coords + EDWIN + embedding).
- Graph builder: kNN on sphere at $k = 16/32/64$; edge attrs = {geodesic, bearing, Δt(TOF), type}.
- Positional encodings $Y_{\ell m}$ ($\ell \leq 4$, 25d) + 16 Laplacian eigenvectors → Linear(32).

ENCODER

- Stage-1: 2× Transformer blocks, width 96, heads 8 → Pool ×¼.
- Stage-2: 2× blocks, width 128, heads 8 → Pool ×¼.
- Stage-3: 2× blocks, width 192, heads 8 → Global attention pooling → $z_{PMT} \in \mathbb{R}^{192}$.

Transformer blocks



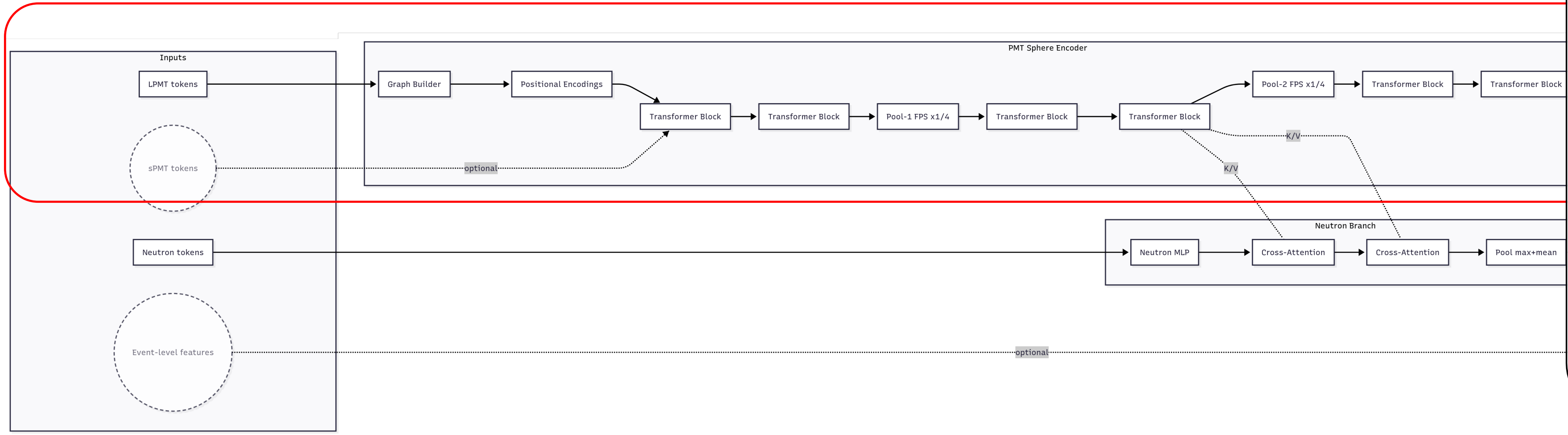
Architecture

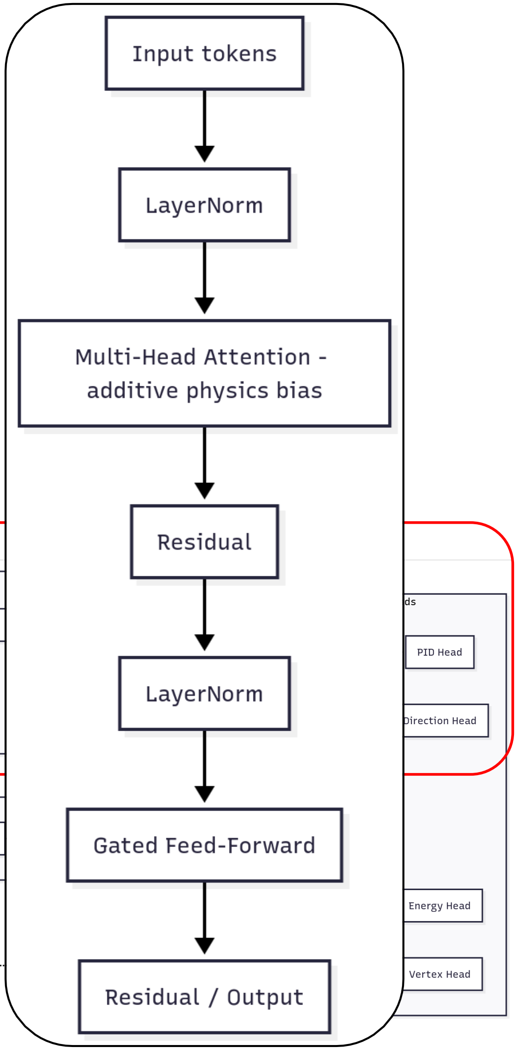
INPUTS

- PMT tokens: active LPMTs $N = 3 - 8k$; features = (coords + EDWIN + embedding).
- Graph builder: kNN on sphere at $k = 16/32/64$; edge attrs = {geodesic, bearing, Δt(TOF), type}.
- Positional encodings: $Y_{\ell m}$ (ℓ≤4, 25d) + 16 Laplacian eigenvectors → Linear(32).

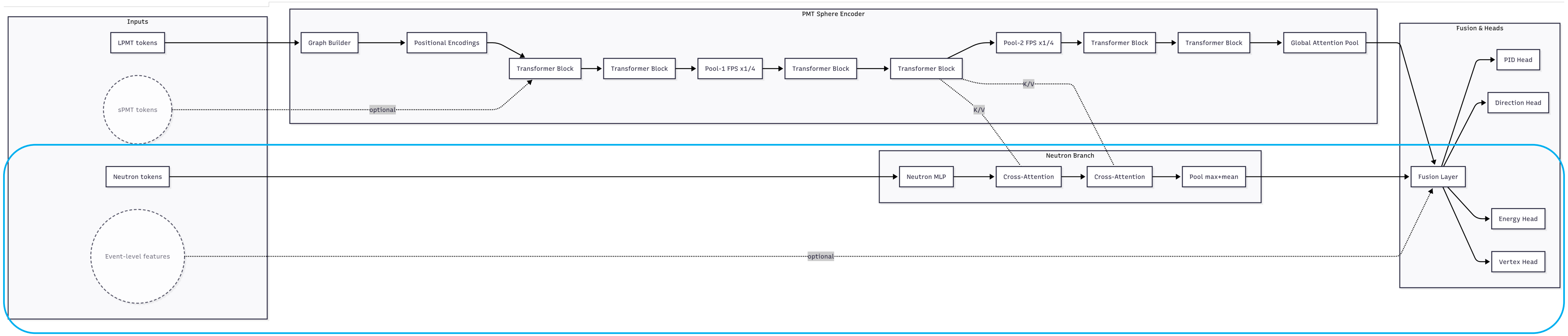
ENCODER

- Stage-1: 2× Transformer blocks, width 96, heads 8 → Pool ×¼.
- Stage-2: 2× blocks, width 128, heads 8 → Pool ×¼.
- Stage-3: 2× blocks, width 192, heads 8 → Global attention pooling → $z_{PMT} \in \mathbb{R}^{192}$.





Architecture



- Today (realistic): ~11 features per neutron (t, ΣPE, Δt, cluster size/spread, early-charge ratio, flags; position if available).
- Future (ideal): 22 neutron + 20 Michel features—drop-in richer tokens.
- Masking: geometry terms in the attention bias are dropped when positions are missing/coarse.

- Token MLP: 12→64→96.
- Cross-attention ×2: Q = secondaries (96), K/V = Stage-2 PMTs (128) → pool across tokens → $z_{sec} \in \mathbb{R}^{96}$.
- Fusion: $[z_{PMT}(192) \| z_{sec}(96) \| evt] \rightarrow$ FC 256 (GELU).

OUTPUT

Architecture

- Bayesian optimization plan (Optuna)
- Objective (multi-task): maximize macro-AUC (5-way PID) and minimize visible energy and mean opening-angle, with a pruner at ~35% of epochs.

Hyperparameter	Module / Symbol	Search space	Default	Rationale / effect
Learning rate	all	log-uniform [1e{-}5, 5e{-}3]	1,00E-03	Controls convergence; interacts with OneCycle/cosine scheduler.
Batch size	data loader	{16, 24, 32, 48, 64}	32	Memory vs. stability; larger with AMP if GPU allows.
Dropout	all blocks/heads	uniform [0.00, 0.30]	0.10	Regularizes attention & FFN; reduces overfit at high E.
Stage blocks	PMT stages	{(1,1,1), (2,2,2)}	(2,2,2)	Depth vs. speed; (2,2,2) is default used here.
Widths C	PMT stages	(80-112,112-160,160-224)	(96, 128, 192)	Capacity allocation across scales.
Heads H	PMT/cross-attn	{6, 8}	8	More heads → finer angular partitions, higher cost.
FFN expansion r	PMT/cross-attn	{2, 3}	2	GEGLU width; r=2 is good trade-off.
kNN per scale	Graph	k_1 in {12,16,20}, k_2 in {24,32,40}, k_3 in {48,64,80}	16/32/64	Neighborhood size; too small misses context, too big adds noise.
Cross-attn depth	Secondary branch	{1, 2}	2	More depth helps neutrino/antineutrino with many captures.
Token-MLP width	Secondary MLP	{64, 96, 128}	96	Encodes neutron/Michel features.
Waveform emb dim	PMT features	{0, 8, 16}	8	0 = ablation; >0 captures residual timing info.
Label smoothing	PID	{0.0, 0.05}	0.05	Helps robust multi-class calibration.
Focal γ (binary)	PID (3+2 binaries)	{0.0, 1.0, 2.0}	0.0	Use >0 if class imbalance in neutrino/antineutrino.
Scheduler	all	{OneCycle, Cosine}	Cosine	Both work; Cosine simpler for BO.
vMF κ cap	Direction	{None, clip(κ≤100)}	clip	Stabilizes training at early epochs.

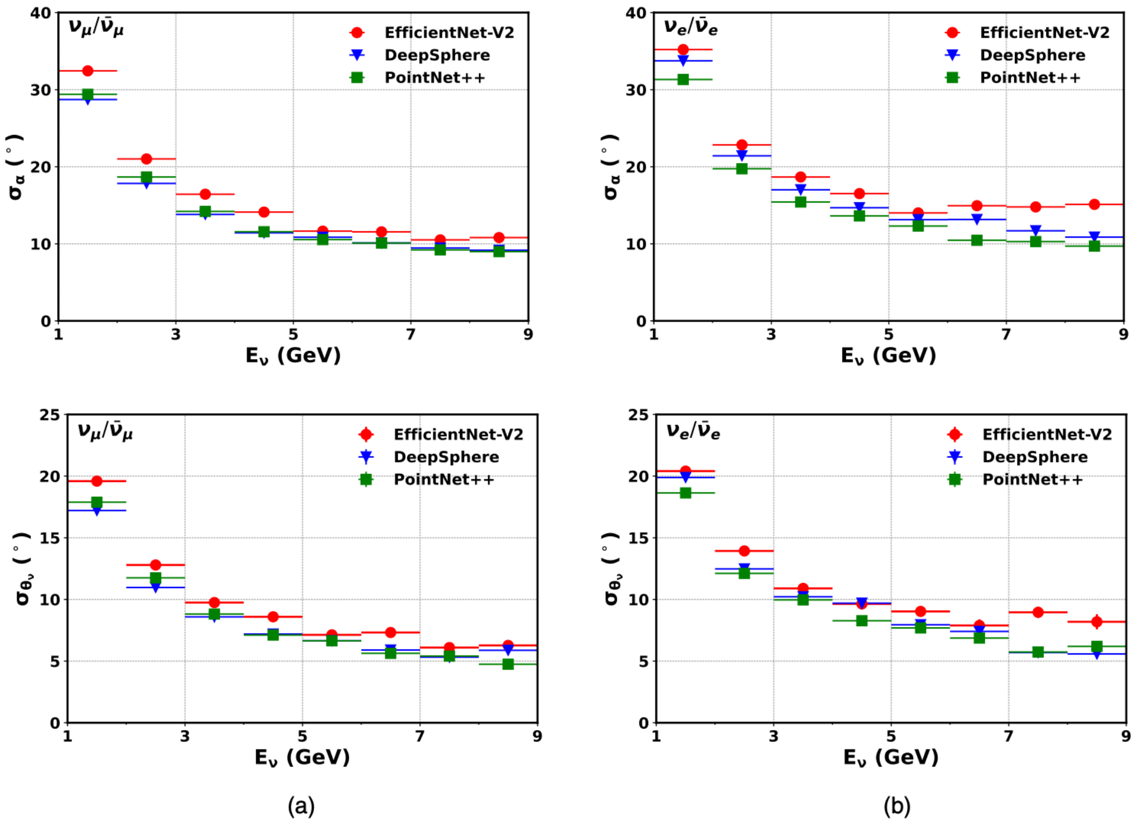
Architecture

- Parameter inventory (ORION, this configuration)

Layer / block	Width (C)	Heads	Depth	Params
Input embedding (features+pos → 96)	96	–	1	**6.7 K**
Stage-1 Transformer block ×2	96	8	2	**184.3 K**
Pool ×¼	–	–	1	–
Stage-2 Transformer block ×2	128	8	2	**327.7 K**
Pool ×¼	–	–	1	–
Stage-3 Transformer block ×2	192	8	2	**737.3 K**
Global attn pool → z_PMT(192)	–	–	1	–
Secondary token MLP (12→64→96)	96	–	1	**7.1 K**
Cross-attention block ×2 (Q=96, K/V=128)	96	4	2	**196.6 K**
Global pool → z_sec(96)	–	–	1	–
Fusion FC (301→256)	256	–	1	**77.3 K**
Shared hidden for classifiers (256→128)	128	–	1	**32.8 K**
PID heads (5-way + 3-way + 2 + 2)	–	–	–	**1.5 K**
Direction head (256→128→(û,κ))	–	–	–	**33.3 K**
Energy heads (E_vis, E_ν ; each 256→128→2)	–	–	–	**66.0 K**
Vertex head (256→128→3)	–	–	–	**33.2 K**
TOTAL (no sPMTs)	–	–	–	**≈ 1.71 M**

Performance of direction reconstruction

- For $E_\nu > 3$ GeV, angular resolution is better than 10° for all ML models and for both flavour
- θ_ν resolution results are coherent with ORION
- α resolution is different (sample different: Honda flux vs. flat flux)

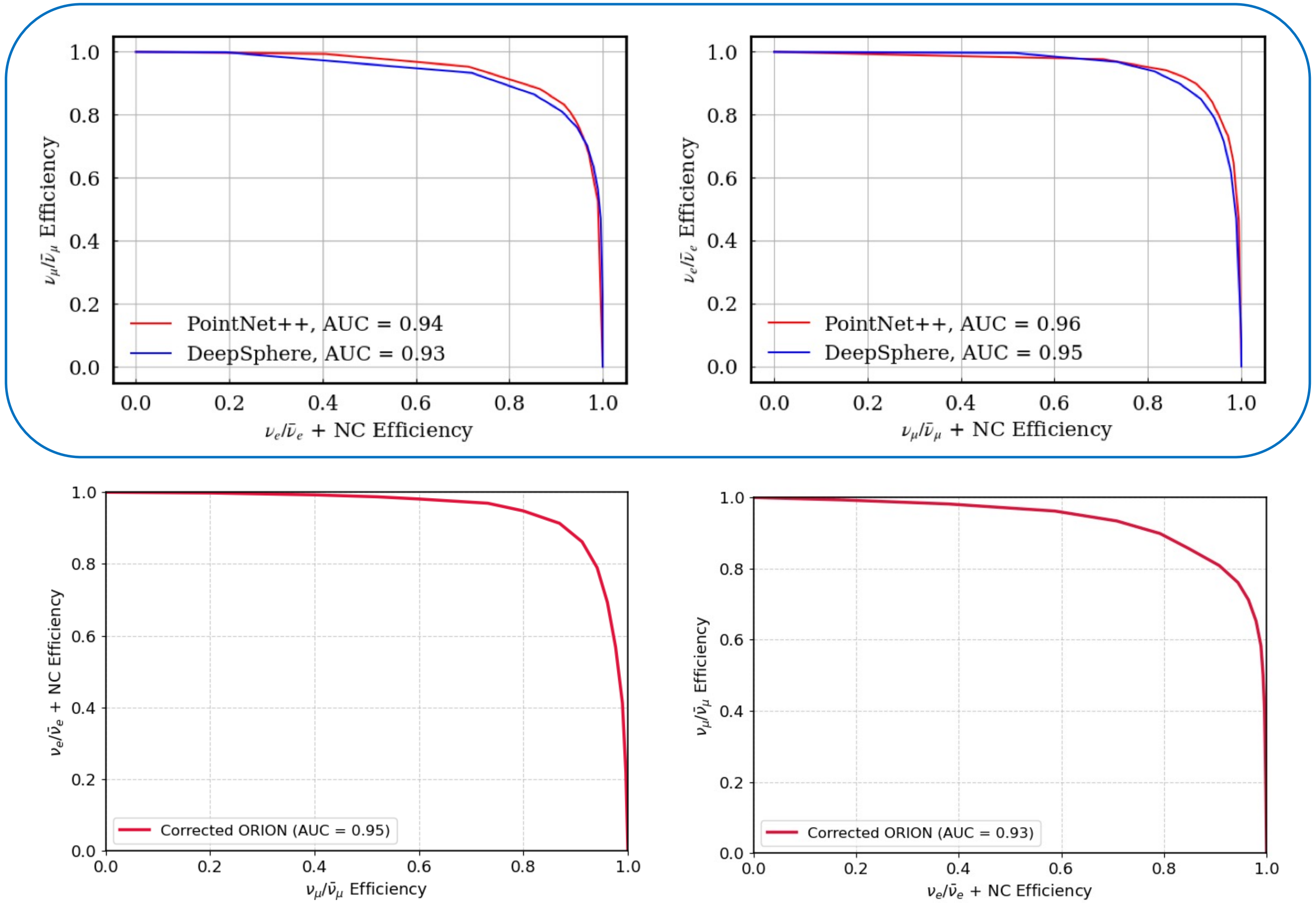


Phys. Rev. D 112, 012018 (2025)

Performance of PID

- ROC curves of the 3-label identification ML models using events across all energies

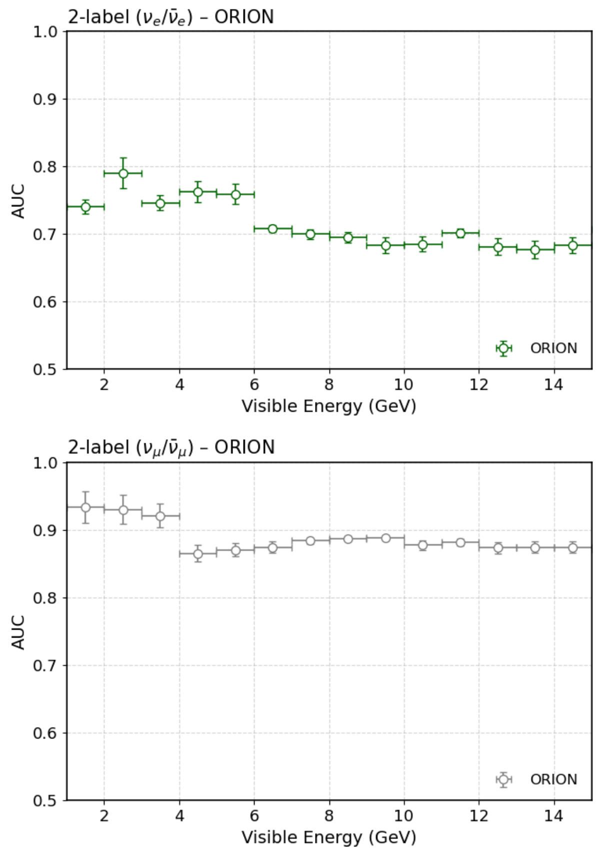
[arXiv:2503.21353](https://arxiv.org/abs/2503.21353)

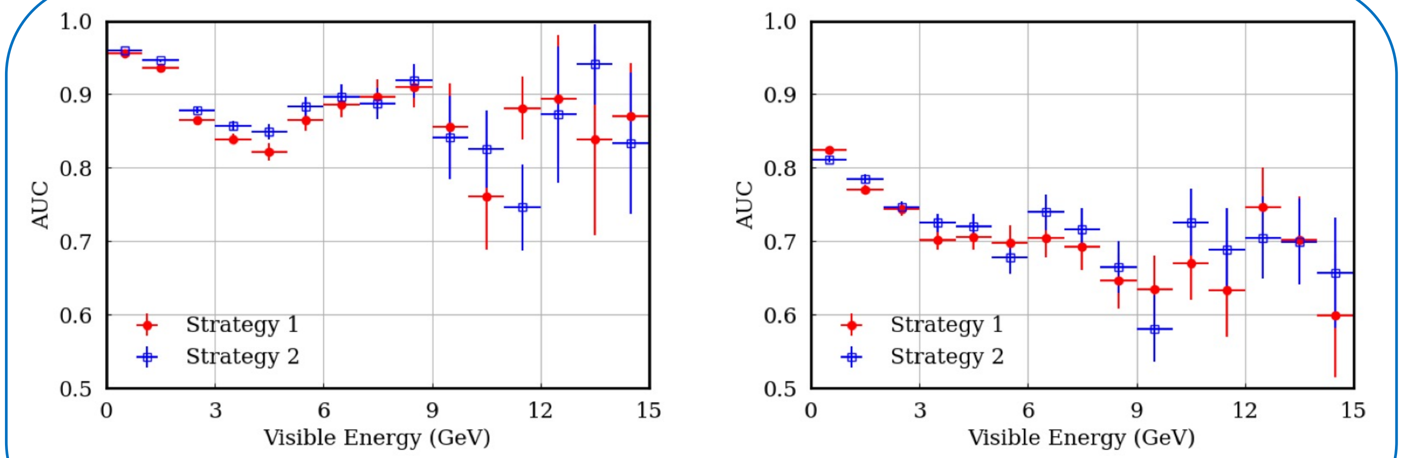


Performance of PID

- Reminder: Only fully contained events
- 5-level classification is using both PMT features from prompt trigger and delayed trigger information
- For the oscillation analysis the score can be tuned depending on the requirement

arXiv:2503.21353





(a) $\nu_\mu/\bar{\nu}_\mu$ identification

(b) $\nu_e/\bar{\nu}_e$ identification

- The results are consistent between the 3 models
- (1: PointNet++ ; 2: DeepSphere)
- Difference because of Honda flux + ratio of neutrino different