Machine Learning Applications in JUNO

Wenjie Wu*

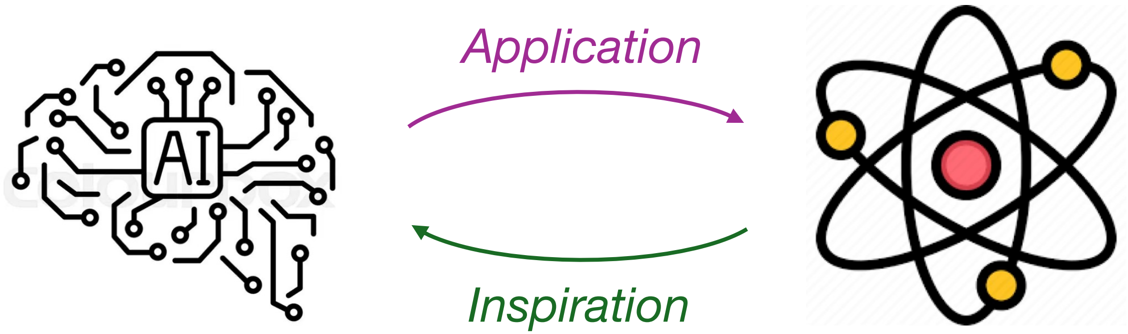
on behalf of the JUNO Collaboration

2026/8/17 @ SYSU, Shenzhen

Quantum Computing and Machine Learning Workshop

AI/ML in Particle Physics

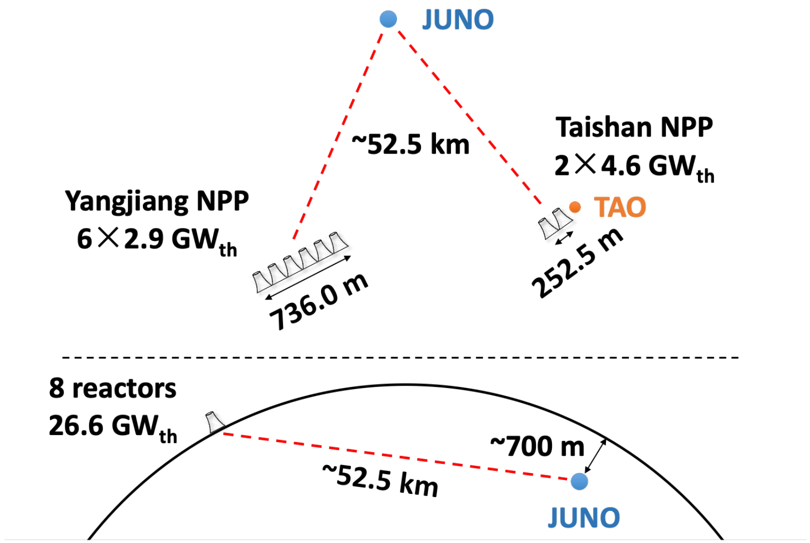
- The ongoing revolution in the filed of AI/ML has significantly influenced the particle physics community
- The applications include but not limited to
    - Operation of accelerators and detector systems: beam monitoring, trigger, anomaly detection, automatic shift, etc
    - Improving sensitivity by extracting more information from data: simulation and reconstruction
    - Improving Monte Carlo calculations for lattice QCD
    - ……

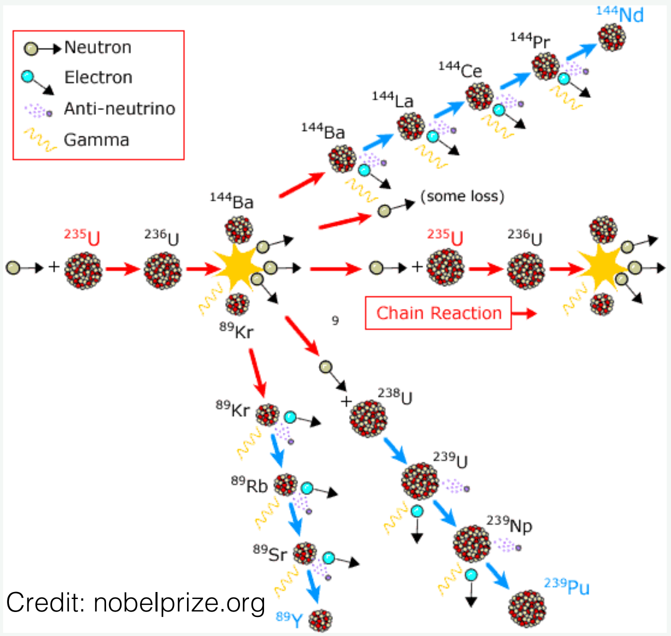


- Non-uniform detector geometries
- Sparse data with high dynamic range
…

JUNO Experiment

- The **J**iangmen **U**nderground **N**eutrino **O**bservatory (JUNO) is a large multi-purpose experiment recently started data-taking





Credit: nobelprize.org

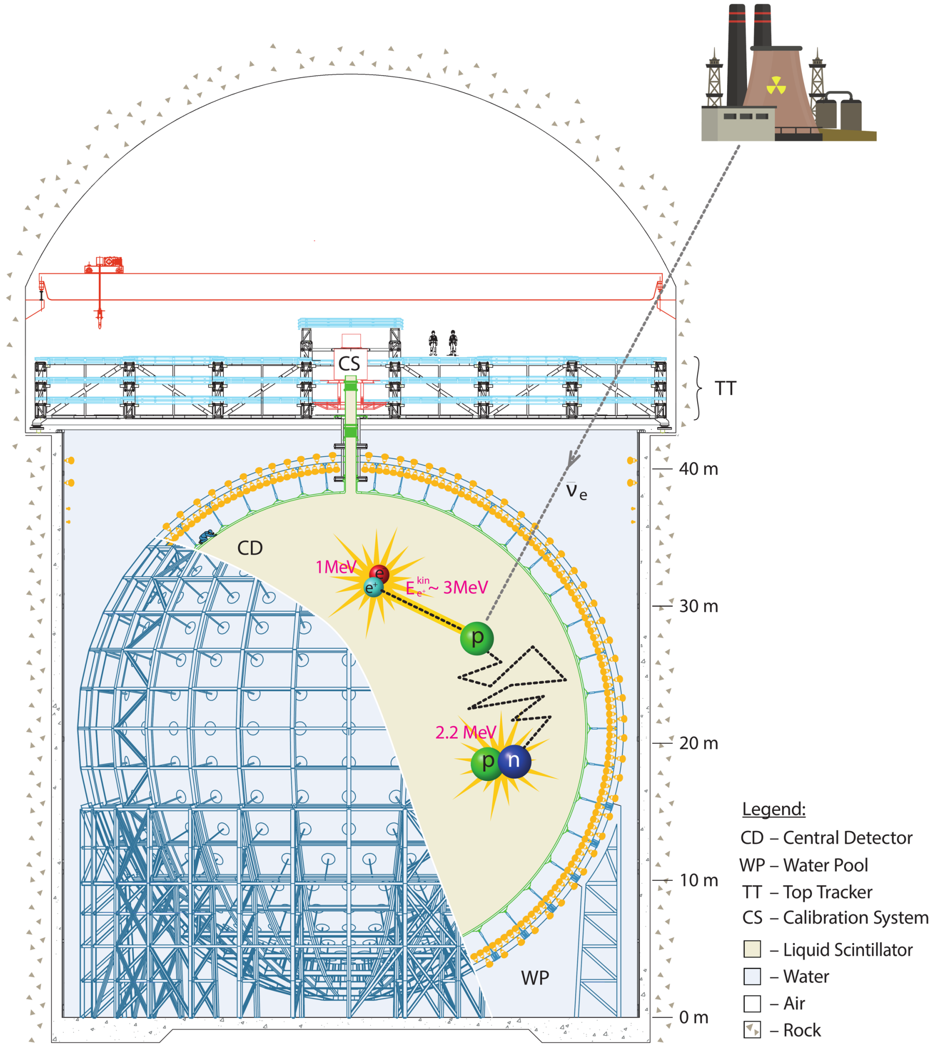
$n \rightarrow p + e^- + \bar{\nu}_e$

✓ ~52.5 km from two major nuclear power plants (1 GW$_{th}$ ~ 2×10^{20} $\bar{\nu}_e$/s)

✓ First commissioned “next-generation” neutrino experiment targeting ν mass ordering

✓ Unprecedented energy resolution of 3% at 1 MeV

JUNO detector



- Key features of the detector
  - High liquid scintillator transparency
  - High PMT coverage
  - High PMT quantum efficiency

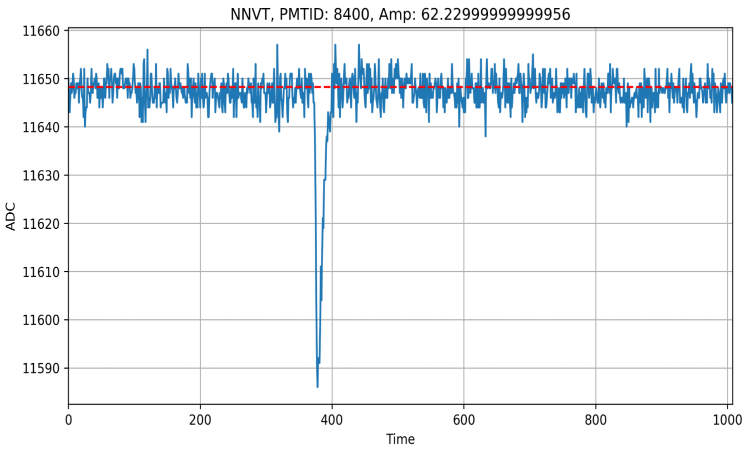
High light yield: ~1600 PE/MeV for ^{68}Ge

- **17,596 20-inch PMTs** serve as the main light-detection device
- **25,600 3-inch PMTs** as the dual-calorimetry system
- **2,400 20-inch PMTs** in water pool for cosmic muon veto

PMT signal is the only portal to physics

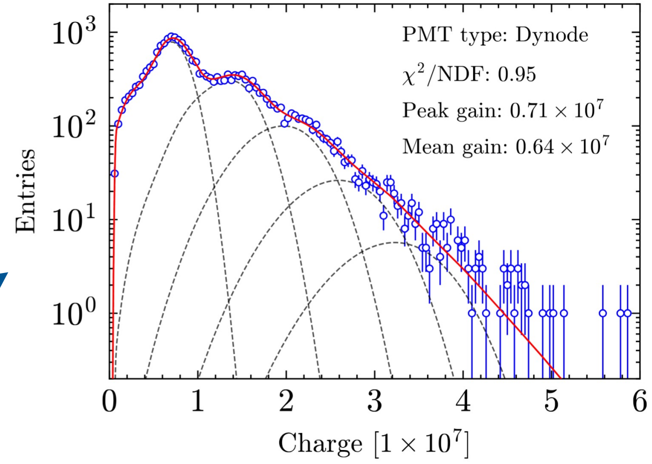
What data we got from the JUNO detector

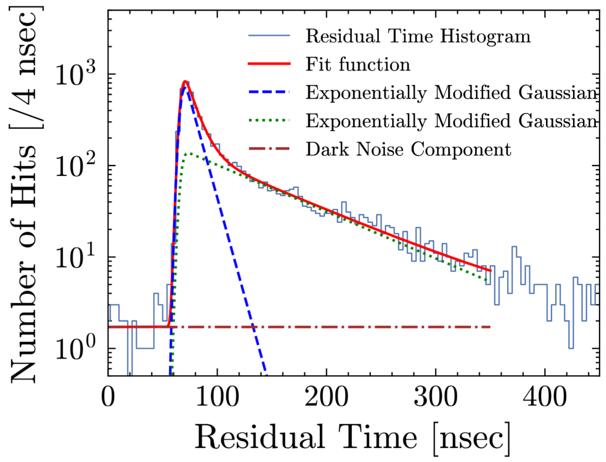
A typical data processing pipeline:
waveform → charge/time → vertex, energy, track, etc



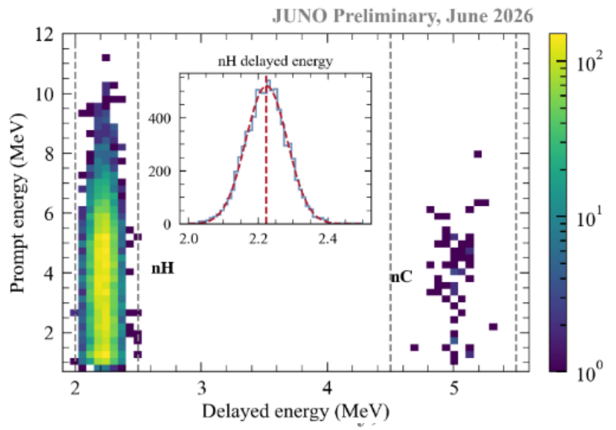
charge

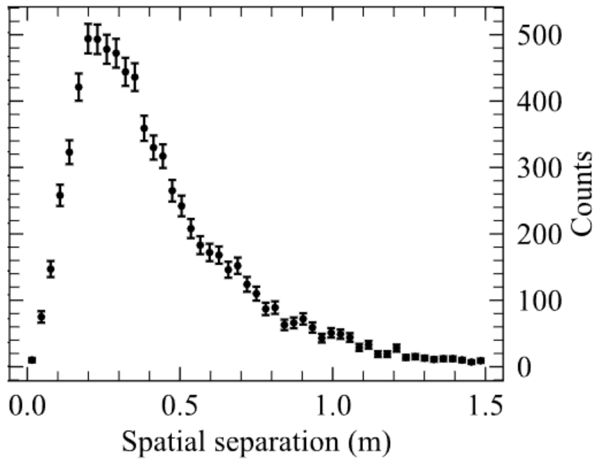
time





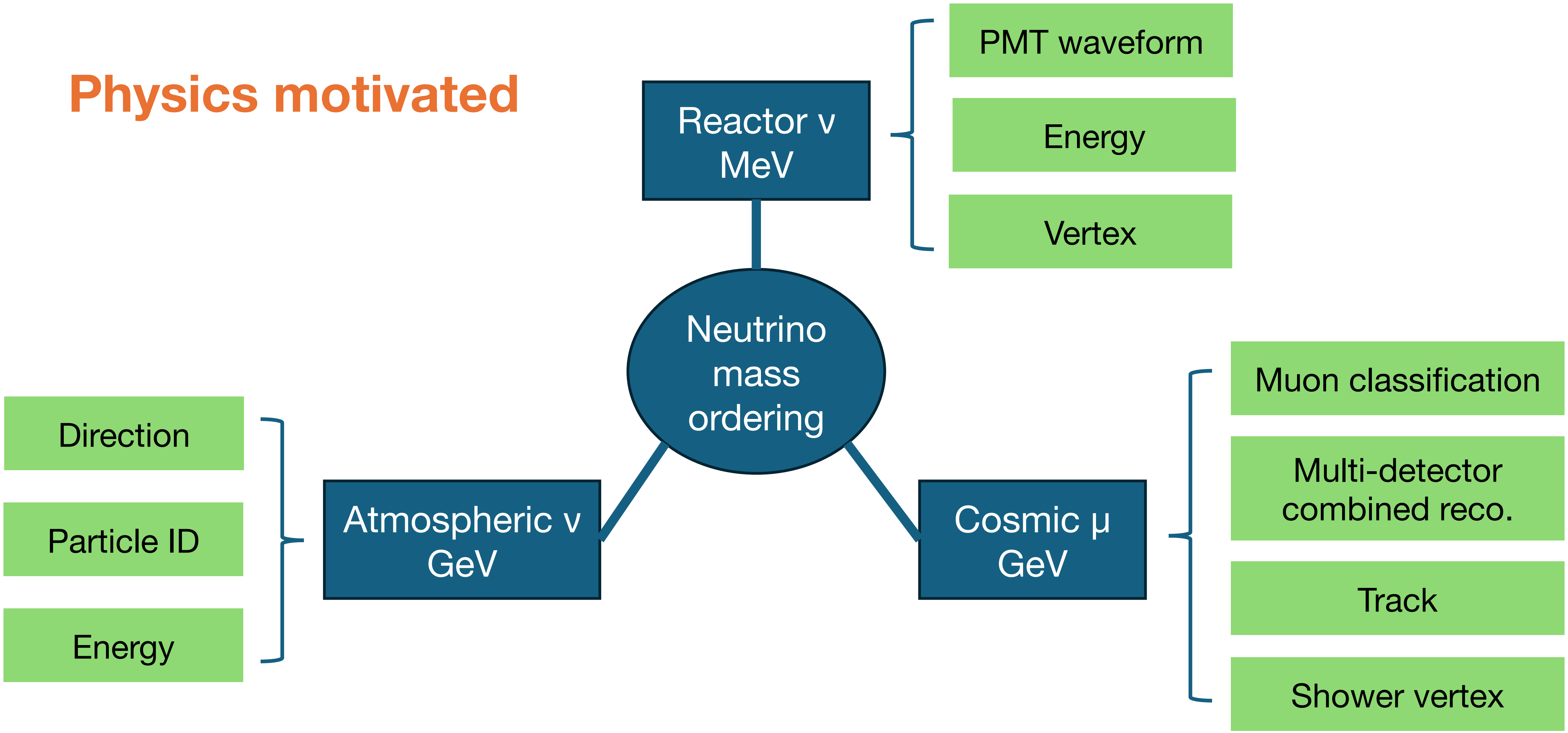
Selected events energy and spatial distributions





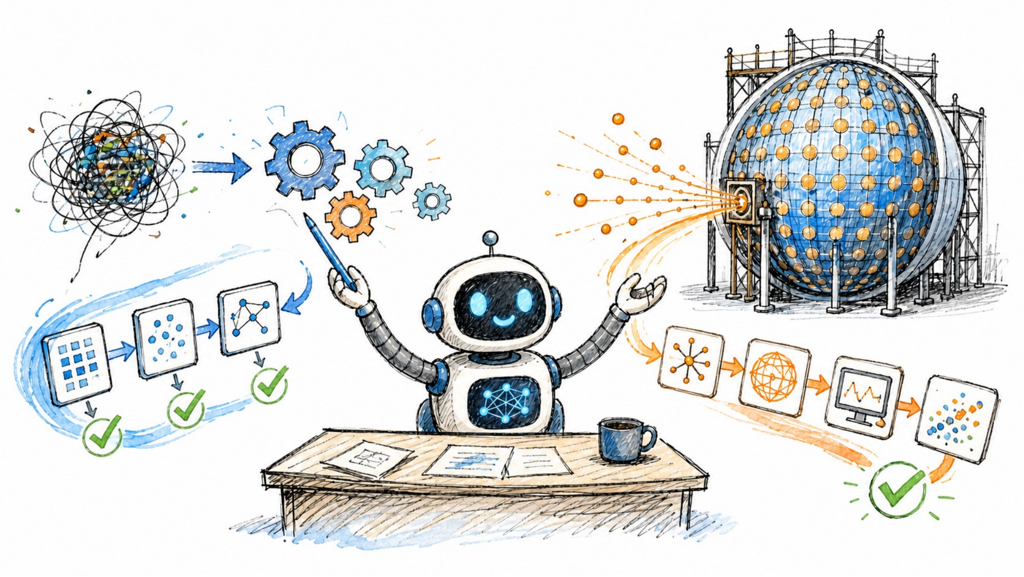
What AI can help for JUNO

Physics motivated



What AI can help for JUNO

Workflow motivated

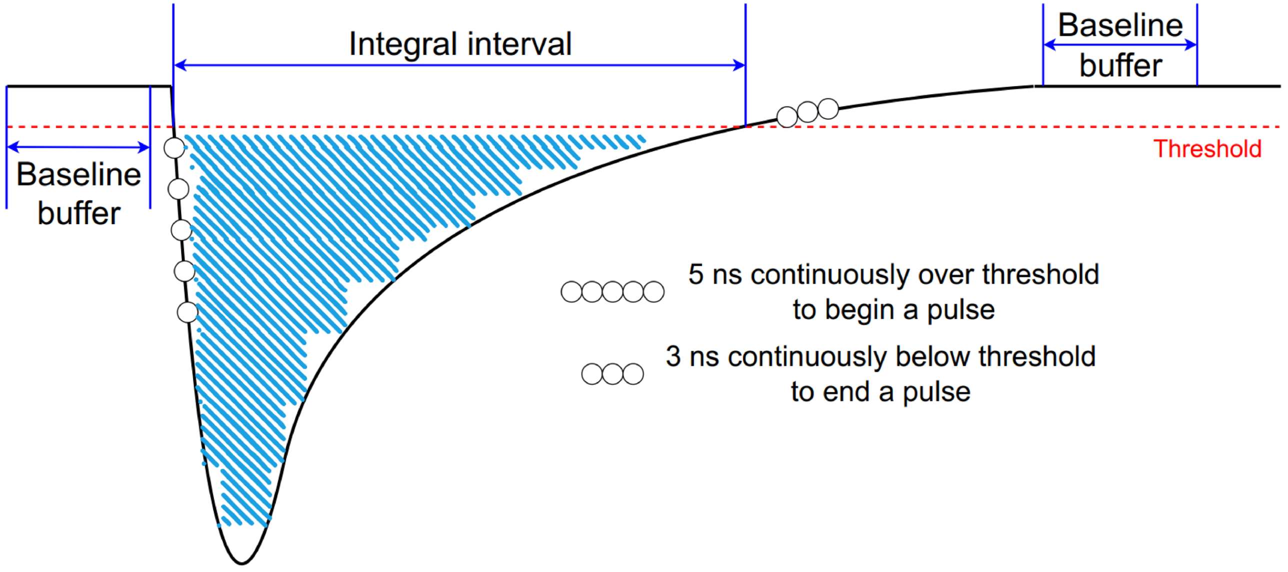


LLM Agents

- Code optimization / automation
- Calib. & Reco. campaign
- Physics analysis workflow tuning
- Software development / Data production

"Agentic AI for software development in JUNO and CEPC" by Tao Lin, Aug. 19

PMT Waveform Reconstruction



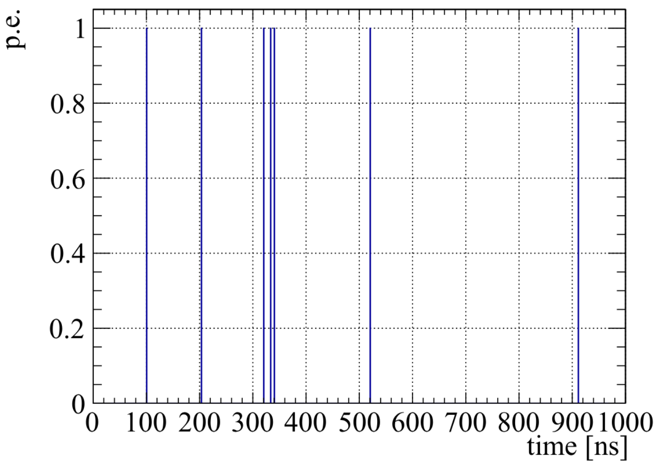
- Simple integration to calculate charge
- Linear fit of the rising edge to obtain hit time

- **Pros: fast and robust, suitable for online processing**
- **Cons: limited charge resolution, difficult to separate piled up hits**

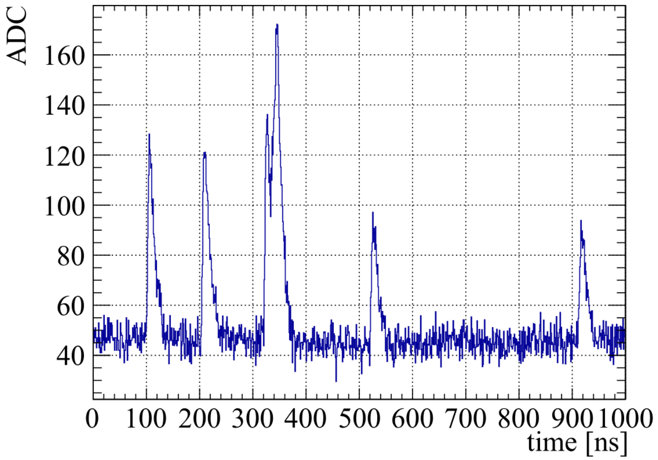
PMT Waveform Reconstruction

- **Input**: pre-processed PMT waveform within 420 ns signal window
- **Model**: customized RawNet
- **Output**: $\{p_k\}$ the probability for predicting $(k=0, 1, \ldots \geq 9)$ PEs
- **It directly counts the number of photons ~ no charge smearing**

Truth PE



Simulated Waveform

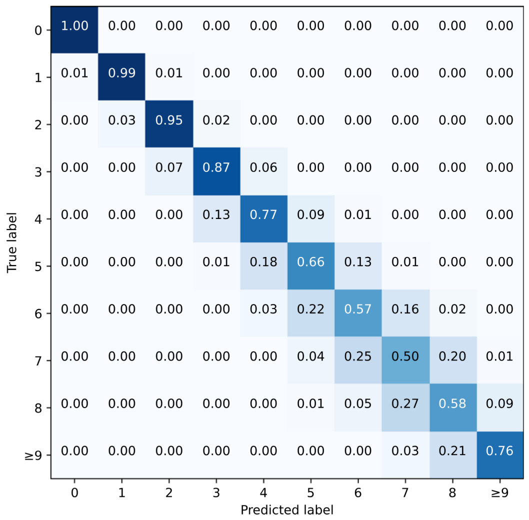


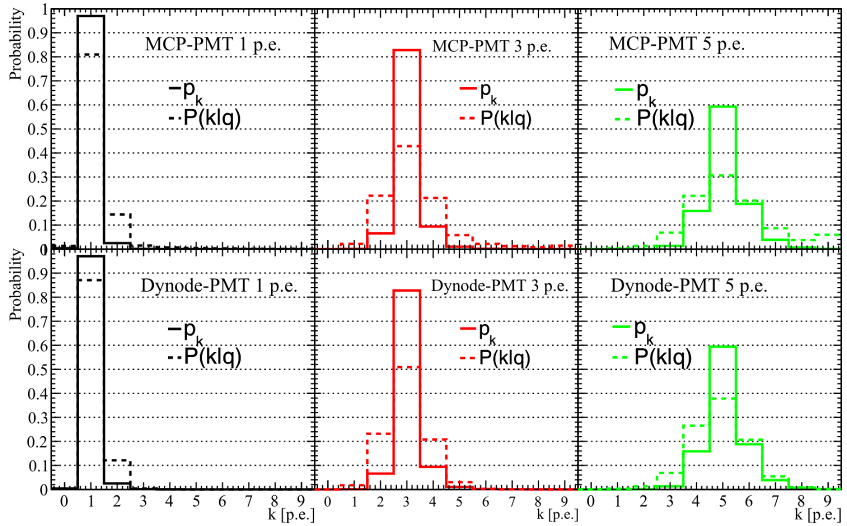
Layer	Input	Output shape
Strided -conv	Conv(3,3,128) BN LeakyReLU	(128, 140)
Res block	$\left\{ \begin{array}{c} Conv(3,1,128) \\ BN \\ LeakyReLU \\ Conv(3,1,128) \\ BN \\ LeakyReLU \\ MaxPool(3) \end{array} \right\} \times 2$	(128, 46)
Res block	$\left\{ \begin{array}{c} Conv(3,1,256) \\ BN \\ LeakyReLU \\ Conv(3,1,256) \\ BN \\ LeakyReLU \\ MaxPool(3) \end{array} \right\} \times 2$	(256, 1)
GRU	GRU(1024)	(1024,)
Speaker embedding	FC(128)	(128,)
Output	FC(10)	(10,)

Eur. Phys. J. C (2025) 85:69

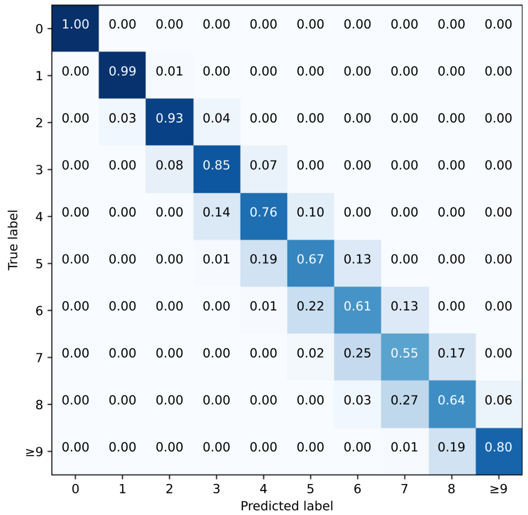
PMT Waveform Reconstruction

MCP-PMT





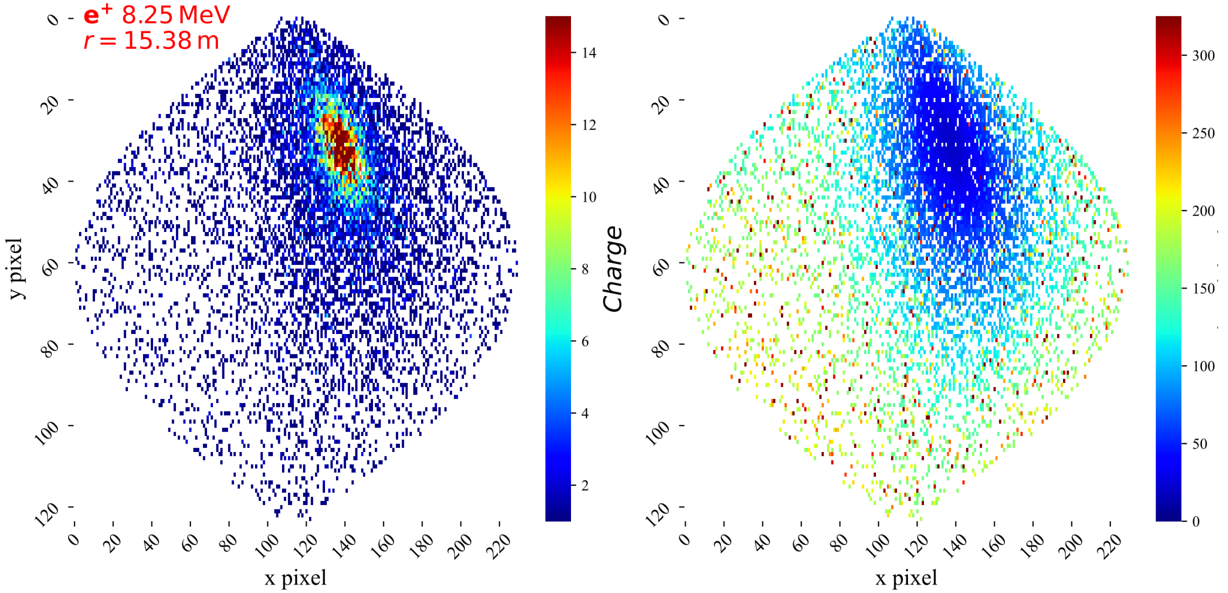
Dynode-PMT

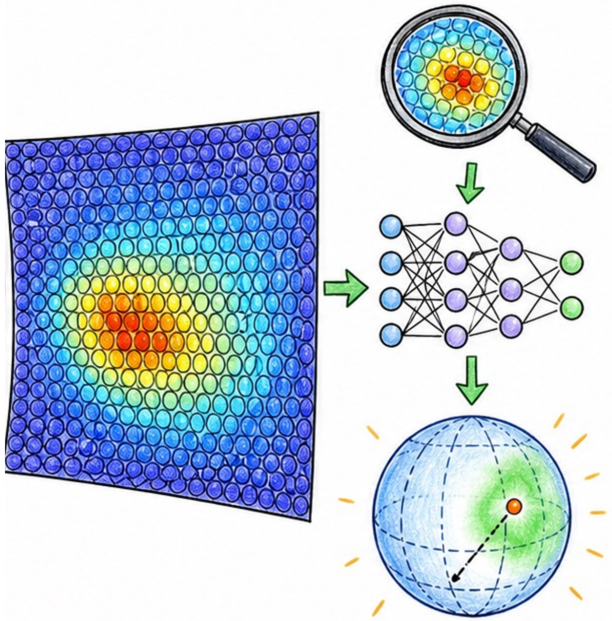


- 99% (95%, 87%) accuracy for 1PE (2PEs, 3PEs)
- Accuracy decreases rapidly as nPEs increases
- Smaller nPE spread compared to traditional method
- Previous study (*Eur. Phys. J. C (2025) 85:69*) based on MC
  - **Application to real data needs more development**
  - **MC/data consistency, computing resource requirements!**

Vertex and Energy Reconstruction

- Large number of PMTs installed on a sphere
  - Each PMT as a pixel ↦ **JUNO as a camera**
  - Ensemble of PMTs charge/time form an image
- Image is highly vertex and energy dependent
- Vertex/energy reconstruction ↦ **Image recognition**

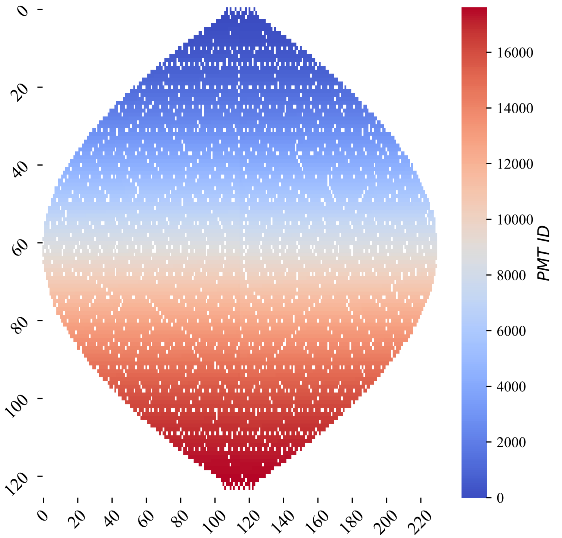


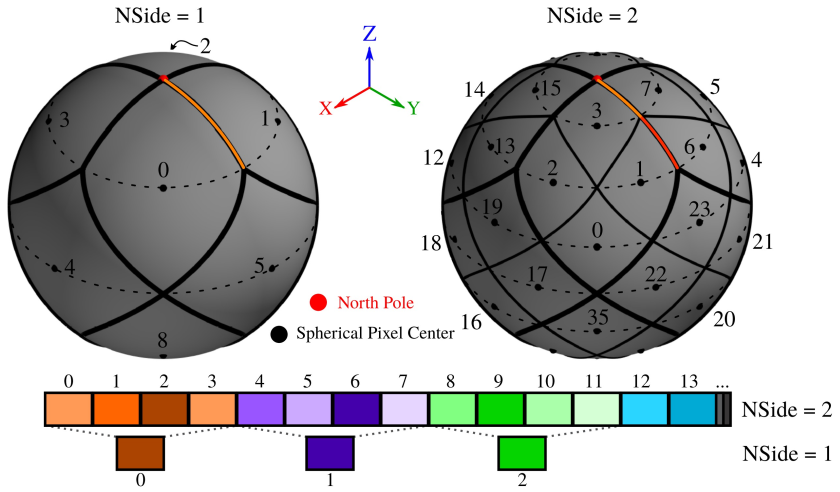


Nucl.Sci.Tech. 33 (2022) 7, 93

Vertex and Energy Reconstruction

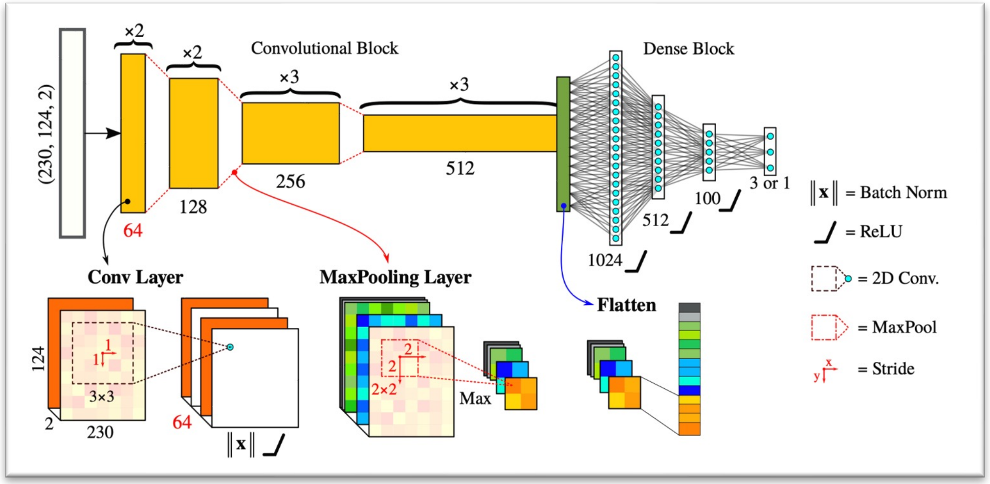
- Large number of PMTs installed on a sphere
- How to construct the inputs to ML models
  - Method 1: projection to 2D plane ↦ Plane CNN
  - Method 2: HEALPix ↦ plane/spherical CNN
  - Method 3: 3D models such as PointNet++/Transformer

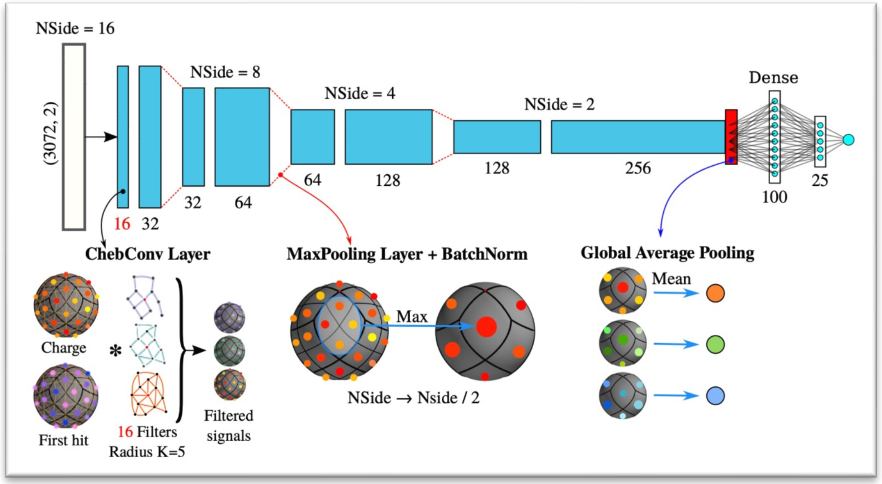


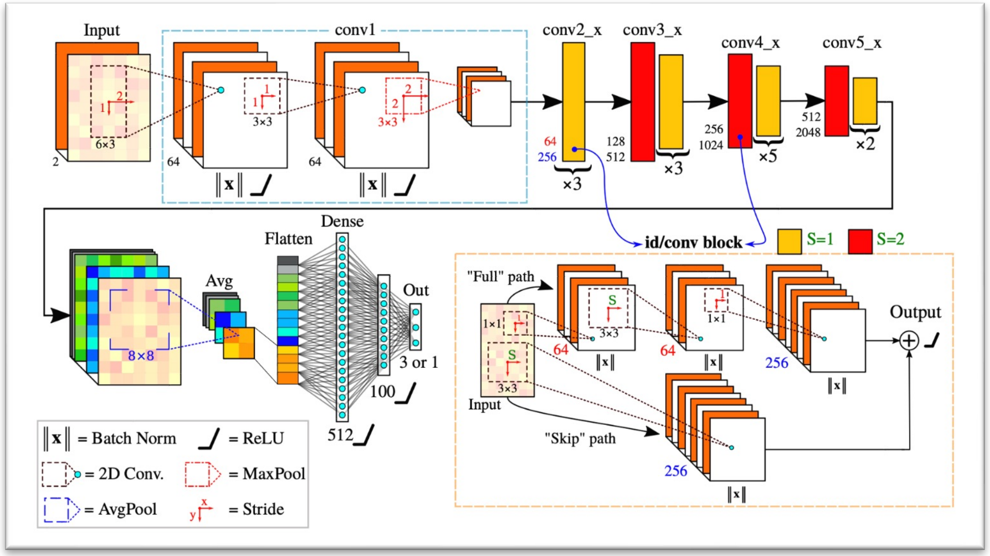


NIMA 1010 (2021) 165527

Vertex and Energy Reconstruction





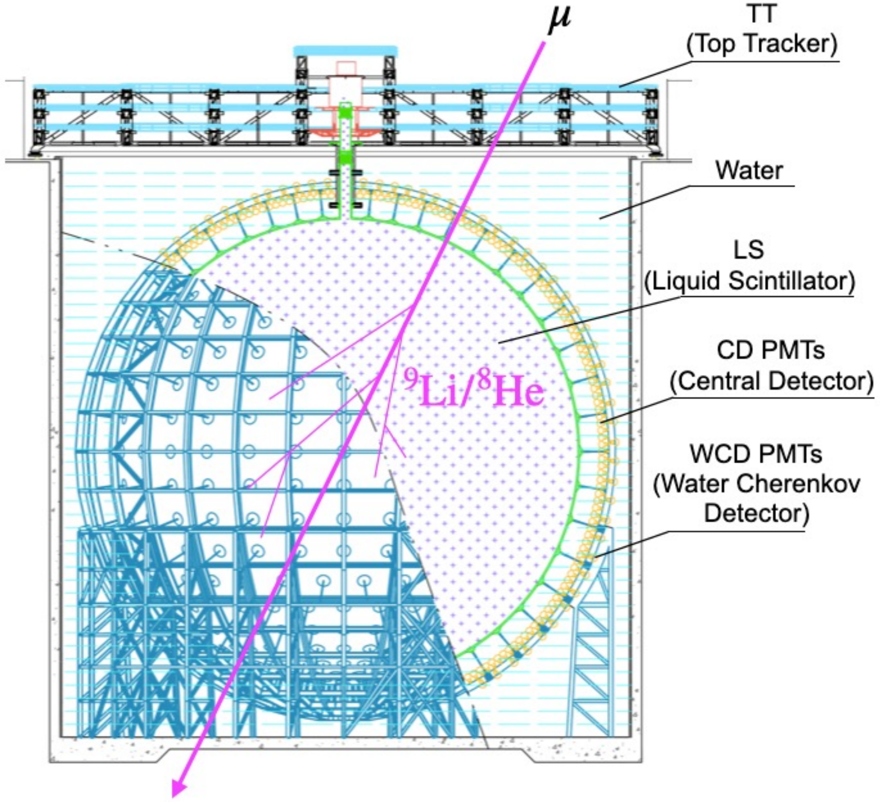


- Several ML methods have been investigated for event vertex and energy reconstruction on MC simulation data, and achieved good performance
- Need more studies on the applicability of the models trained on the simulation to the real data

Muon Track Reconstruction

- ^{9}Li/^{8}He Background
  - $^{9}\mathrm{Li}/^{8}\mathrm{He} \xrightarrow{\beta^{-}n} e^{-} + n +$ daughter nuclei $\Longrightarrow$ prompt β + delayed n capture Mimic IBD signals
  - Muon-SPN correlations provide direct guidance to suppress this cosmogenic bkg.

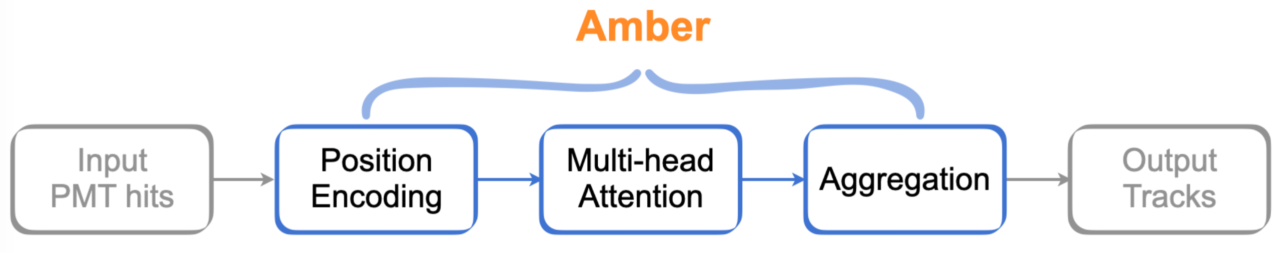
- Muon reconstruction
  - **CD & WCD:** provide independent or joint muon track reconstruction
  - **TT:** provides high-precision tracks (0.2° angular resolution) as an internal reference
- ML-based methods
  - **Data-driven training with TT reference tracks**



Muon Track Reconstruction

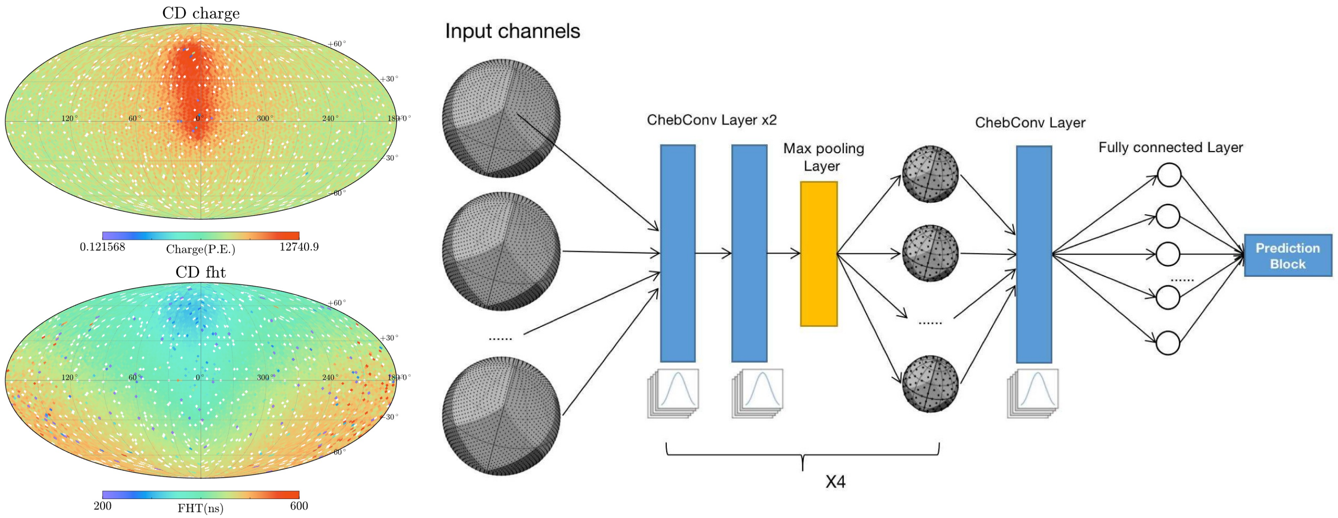
- Attention Mechanism-Based Event Reconstructor (Amber)

  - Transformer-based model
  - **Input: fired PMT (t, q, x, y, z)**



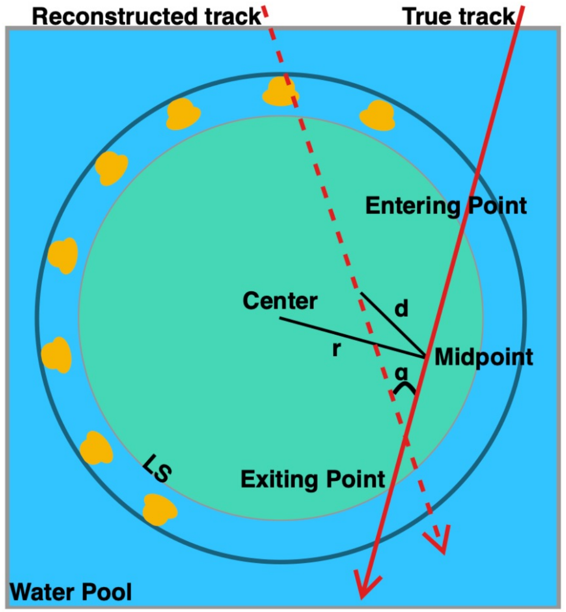
- Event reconstruction based on Deep-learning and Waveform INformation (Edwin)

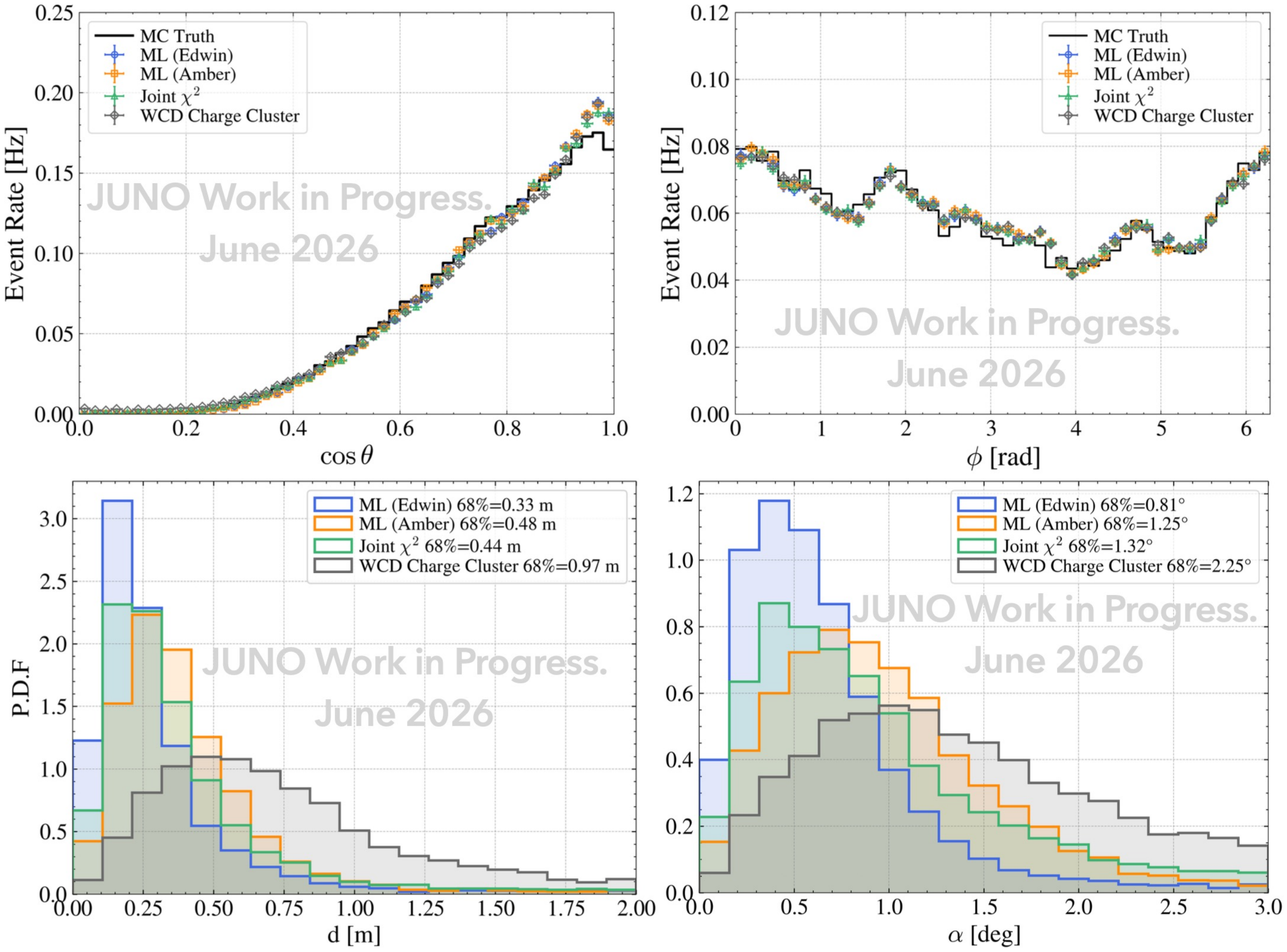
  - HEALPix-based spherical GNN with rotation-equivariant graph convolutions
  - **Input: extracted spatial features from PMT waveforms** (only (fht, q) for now)



Muon Track Reconstruction

- **Angular Spectra:**
  The reconstructed direction
- **TT Reference Tracks Comparison:**
  Common benchmark for evaluating angular and spatial reconstruction residuals.
- Good agreement among independent reconstruction frameworks

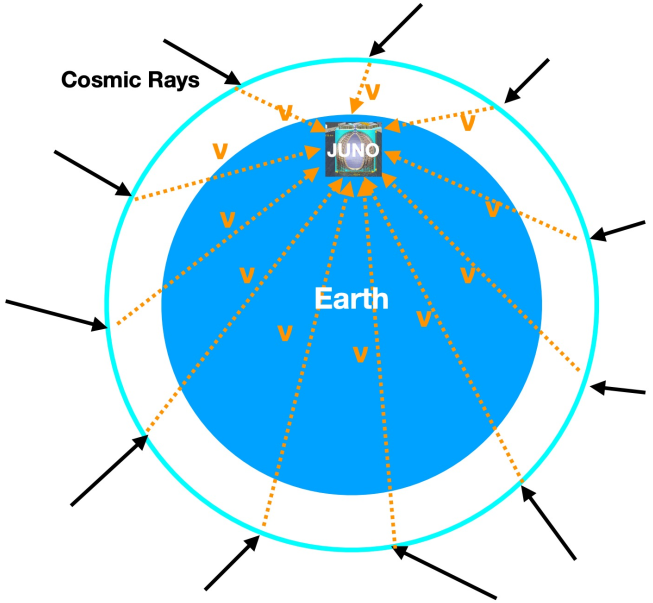


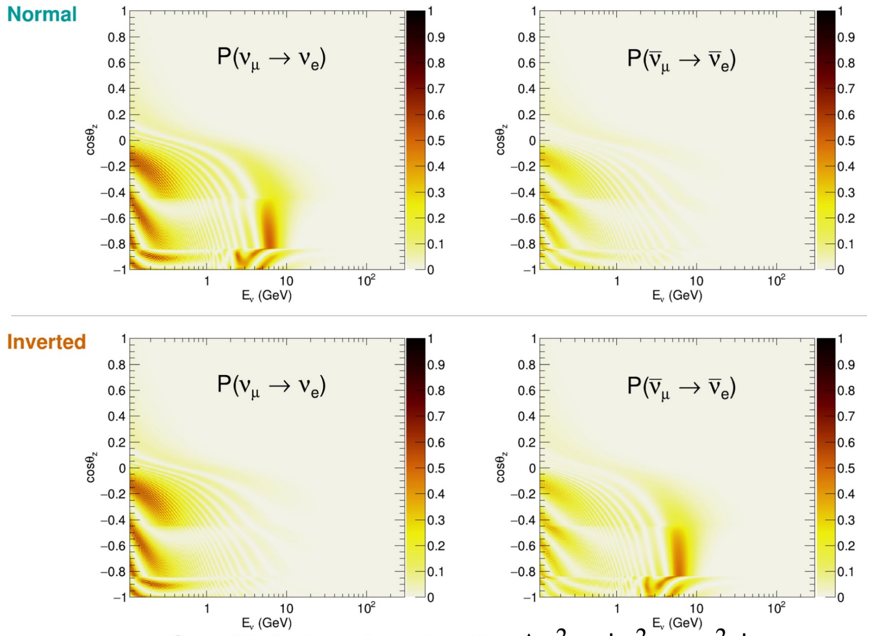


Atmospheric Neutrino Measurement

- Atmospheric ν provide independent sensitivity to NMO via matter effects
- Atmospheric ν is a background to other physics such as DSNB, etc
- Very challenging task in a LS detector
  - Neutrino directionality
  - Flavor identification

Oscillation probability P(L/E), L~cosθ

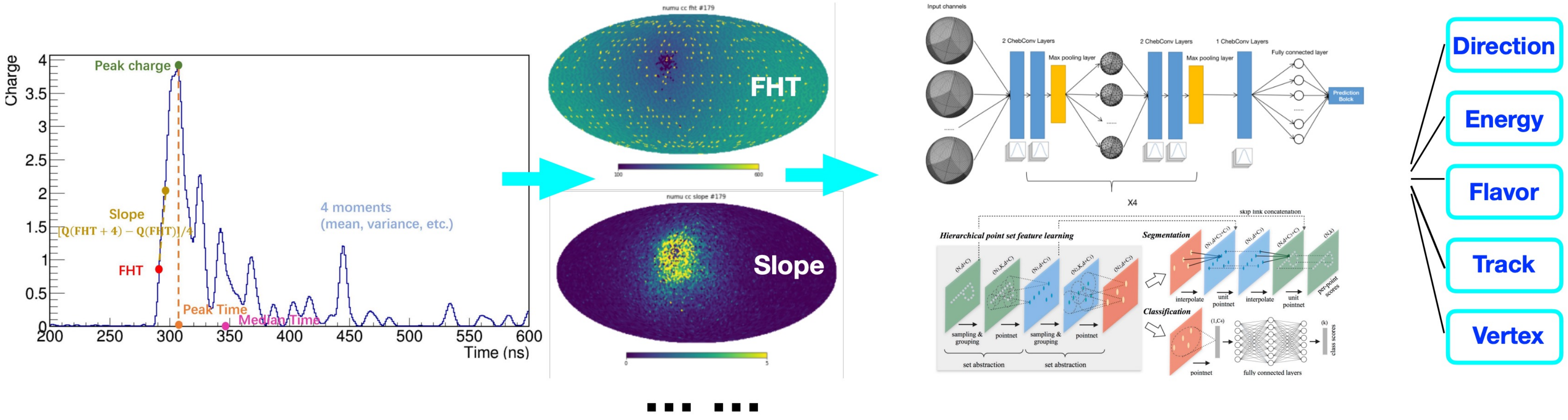




Sensitivity benchmarked by $\Delta\chi^2 = |\chi^2_{NO} - \chi^2_{IO}|$

Atmospheric Neutrino Measurement

- The success of muon track reconstruction in real data provides evidence that PMTs at different angles w.r.t the track see different shapes of nPE(t)
- PMT waveforms depend on track direction, track starting and stopping points, track dE/dx, etc



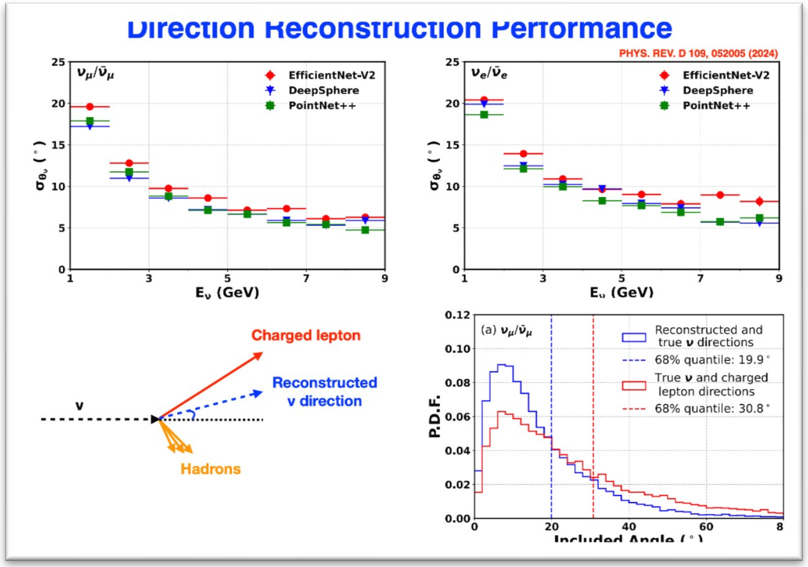
PMT Waveforms (After deconvolution and noise-removing)

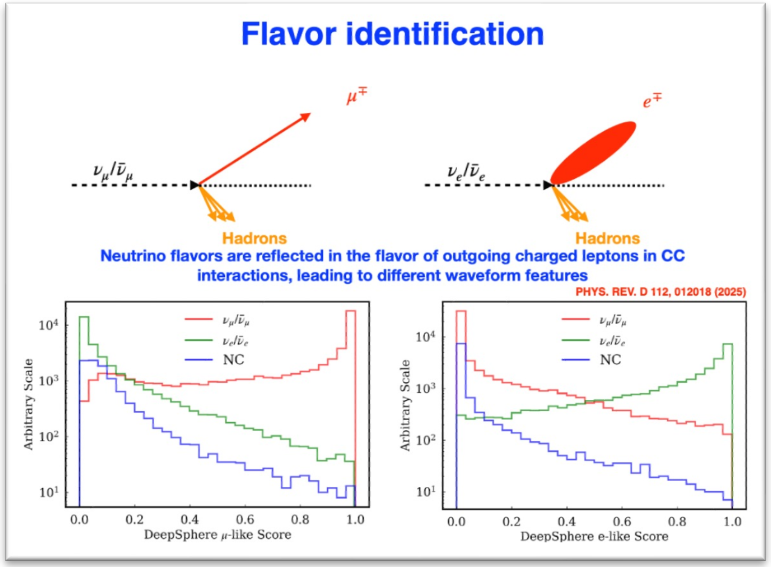
Pictures of PMT Features

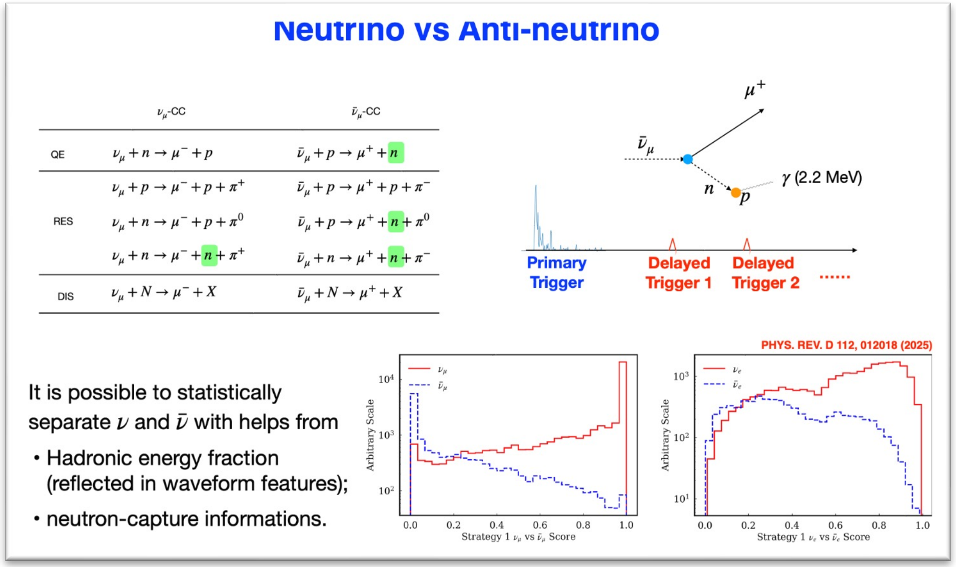
Machine Learning Models

Outputs

Atmospheric Neutrino Measurement

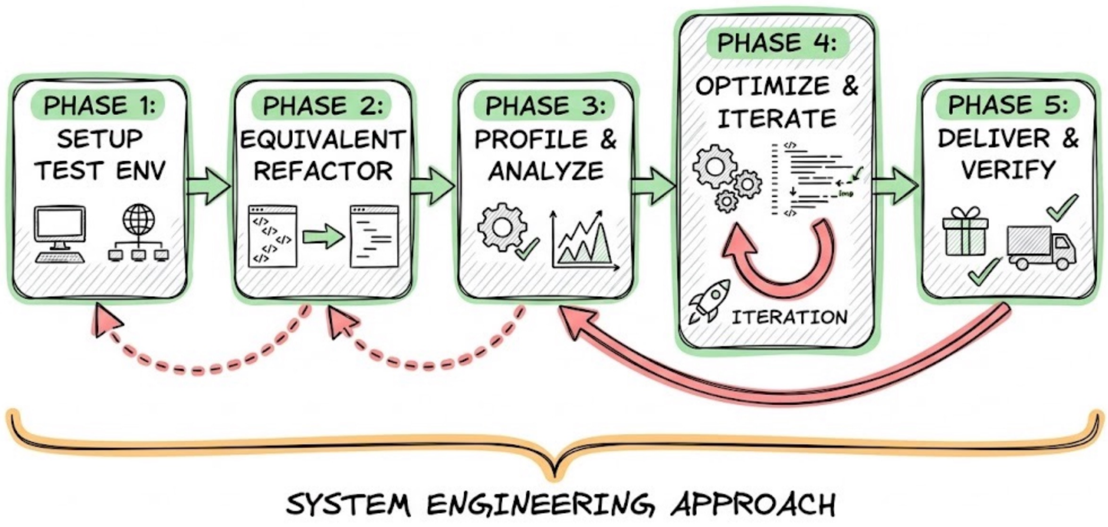






- The performance is promising on MC simulation
- Stay tuned for the real data performance on reconstruction and PID of atmospheric neutrinos in a LS detector!

LLM Agents for Scientific Workflow (Dr. Sai)

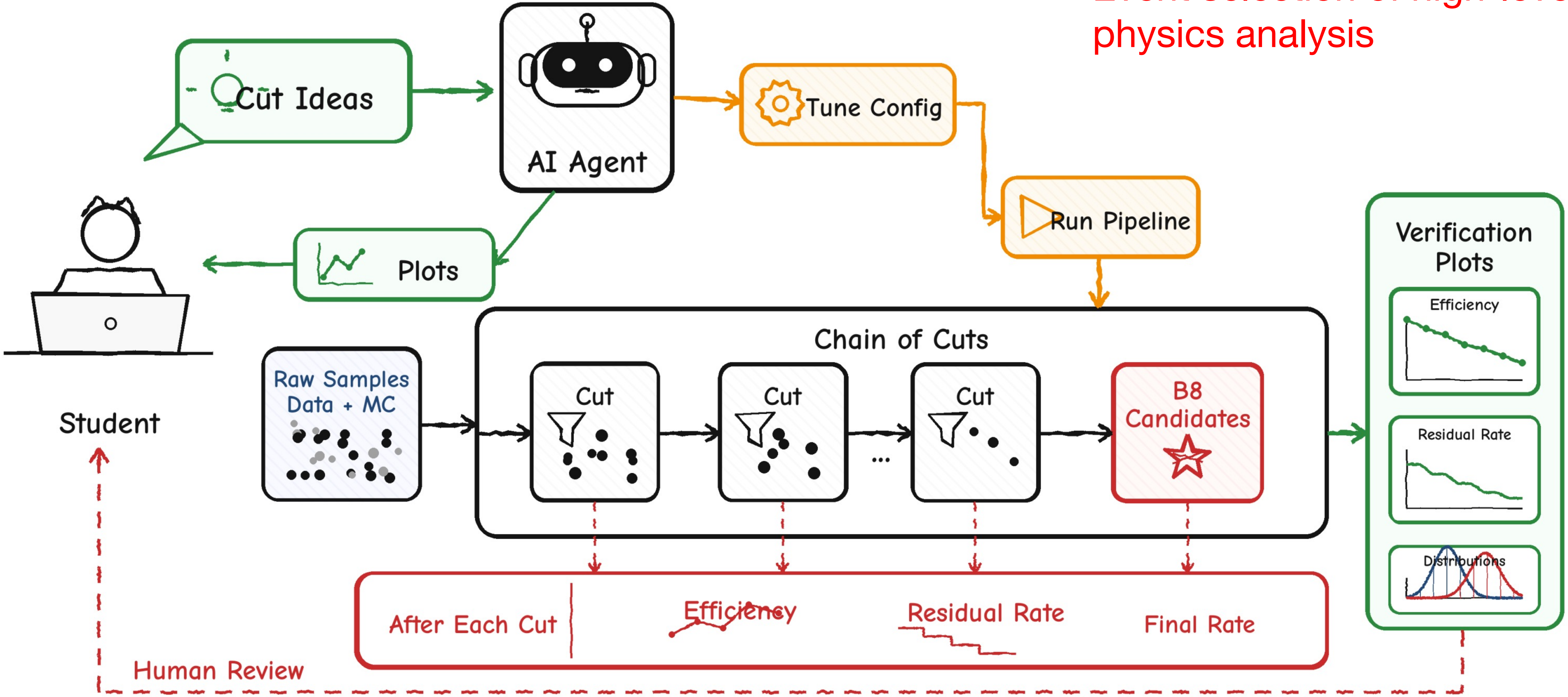


- Dr. Sai pioneers autonomous physics discovery at BESIII (and beyond) through agentic AI
- It integrates LLMs with JUNO/HEP environments (ROOT, SNiPER, HTCondor)

- Optimize the heavily CPU-intensive likelihood reconstruction *OMILREC*
- Demonstrated 5x speedup via agent-guided C++ profiling and refactoring, reducing the human bottleneck in physics optimization

LLM Agents for Scientific Workflow

Event selection of high-level physics analysis



Summary

- With $O(10^4)$ PMTs, and large-scale dataset, JUNO provides a good playground for machine learning applications
- A few examples were presented
  - MeV region: PMT waveform/vertex/energy reco.
  - GeV region: muon and atmospheric neutrino reco.
  - Agentic workflow in physics analysis and data production
- More efforts are ongoing to adapt MC-based experience to real data
- ML can enhance the capability of the detector, yielding faster, better results

Thank you!

Backup

Calibration System

- 1D (ACU, z-axis): ^{137}Cs, ^{54}Mn, ^{40}K, ^{68}Ge, ^{60}Co, Am-C, Am-Be, laser
- 2D (CLS, z-y plane) : Am-C, ^{68}Ge, ^{137}Cs
- 2D (GT, boundary) : Am-Be

Source	Energy	System
	Gamma Sources	
^{68}Ge	0.511 $\times$ 2 MeV	ACU
^{137}Cs	0.662 MeV	ACU
^{54}Mn	0.835 MeV	ACU
^{40}K	1.460 MeV	ACU
^{60}Co	1.173+1.333 MeV	ACU
	Neutron Sources	
^{241}Am–^{13}C	neutron + 6.13 MeV (^{16}O*) (n,γ)p 2.223 MeV (n,γ)^{12}C 4.94 MeV (n,γ)^{56}Fe 7.63 MeV, etc.	ACU CLS
^{241}Am–^{9}Be	neutron + 4.43 MeV (^{12}C*) (n,γ)p 2.22 MeV	GT
	Optical Calibration	
Laser	Optical pulses (420 nm and 266 nm)	ACU