

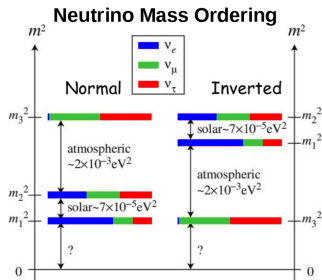
AI 在 JUNO 分析中的应用

from PMT Waveforms to Differentiable Workflows

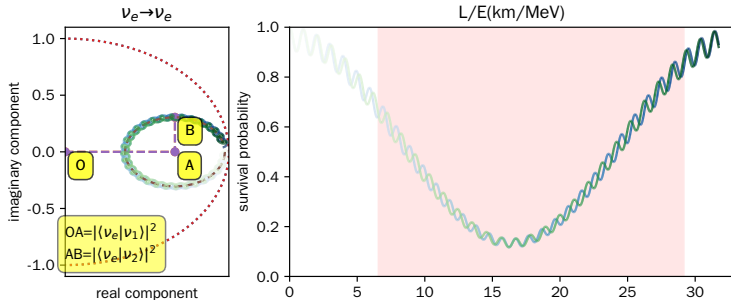
续本达 (Benda Xu)

中国科学院大学 (代表 JUNO 合作组)

2026-07-16 CPS-HEP @ Guangzhou

ν mass ordering (NMO): sign of Δm_{31}^2 

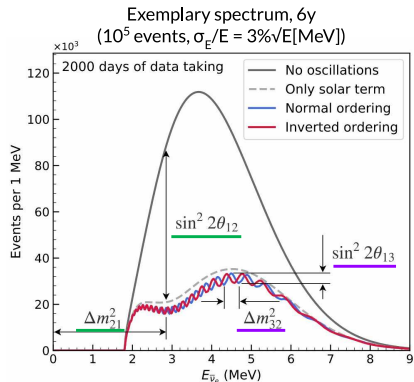
- $\Delta m_{21}^2 = m_2^2 - m_1^2 = 7.5 \times 10^{-5} \text{ eV}^2$
- $\Delta m_{31}^2 = m_3^2 - m_1^2 = 2.4 \times 10^{-3} \text{ eV}^2$
- normal order: $m_3^2 > m_1^2$;
 - inverted order: $m_3^2 < m_1^2$
- imprints subtle difference on energy spectra.



Challenges towards the JUNO Physics Goal

- Neutrino Mass Ordering (NMO) via reactor $\bar{\nu}_e$ spectrum.
- Requires an unprecedented **3% energy resolution @ 1 MeV**.
- Imprints subtle differences on energy spectra.

$$\vec{w}_{\text{PMT}} \rightarrow \{\gamma_{\text{vis.}}\} \rightarrow E, \vec{x}, t \rightarrow \mathcal{L}_{\text{phys}}$$



Challenges towards the JUNO Physics Goal

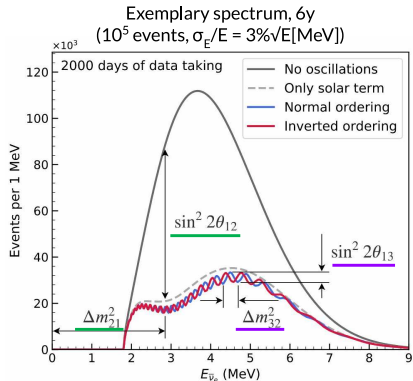
- Neutrino Mass Ordering (NMO) via reactor $\bar{\nu}_e$ spectrum.
- Requires an unprecedented **3% energy resolution @ 1 MeV**.
- Imprints subtle differences on energy spectra.

$$\vec{w}_{\text{PMT}} \rightarrow \{\gamma_{\text{vis.}}\} \rightarrow E, \vec{x}, t \rightarrow \mathcal{L}_{\text{phys}}$$

AI @ JUNO

Maximize the **information throughput** from PMT raw waveform to NMO.

$$2\text{PB} / \text{year} \rightarrow 1\text{bit} / 6\text{years}$$

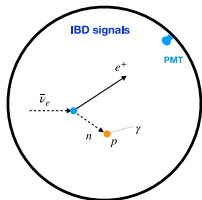


JUNO software and analysis have undergone phase transitions driven by AI.

Domain	AI method	Role	Status
Cosmic μ	ML reconstruction / classification	μ background control	Data
Waveforms	CNN, diffusion, transformer	PE couting and timing	Validation
Reconstruction	Deep inverse models	Vertex/energy reconstruction	R&D
Physics-informed AI	Differentiable programming	Unified engine for simulation, calibration and reconstruction	R&D
Workflow	LLM agents	Code optimization / automation	Data

Reconstruction as inverse problems

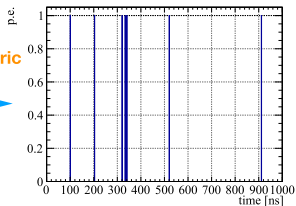
Charged particles deposit energy and produce scintillation light



Photoelectric effects



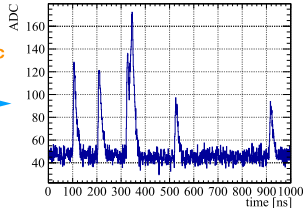
PE count and time



Electronic effects



Waveform



Event reconstruction



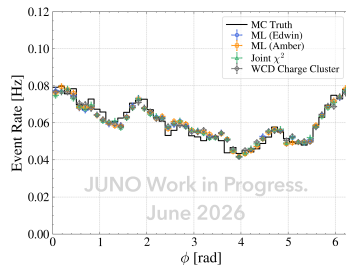
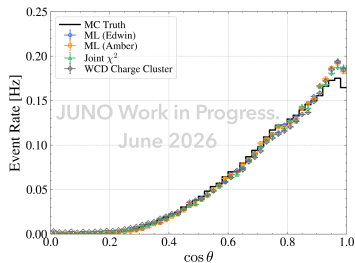
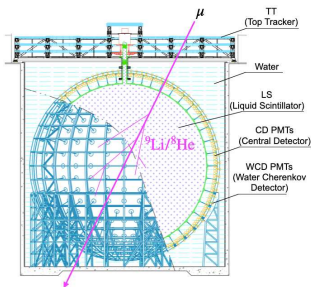
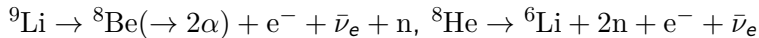
Event energy, vertex etc.

Waveform reconstruction



PE count and time and other waveform features

Waveform



EDWIN Event reconstruction based on Deep-learning and Waveform Information

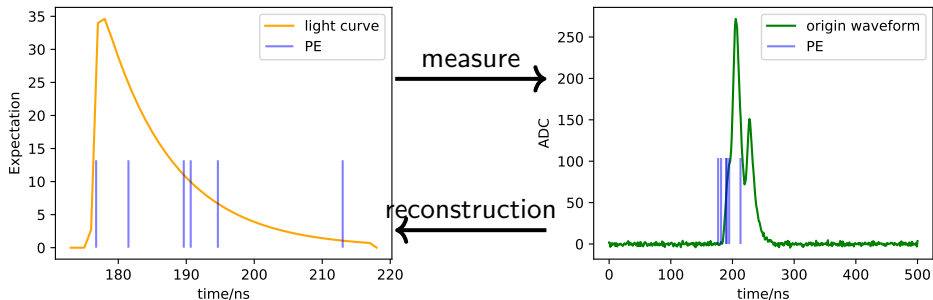
- Spherical graphical neural network.

Amber Attention Mechanism-Based Event Reconstructor

- Transformer.
- 1° angular and $<0.5\text{m}$ positional resolution.

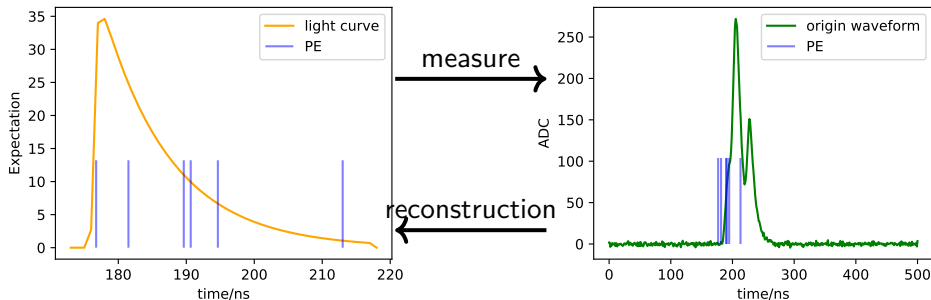
Given a **light curve** $\lambda\phi(t - t_0)$ received **photoelectron** (PE), their times being $\vec{z} := (t_1, t_2, \dots, t_N)$. **waveform** $\vec{w}$ is a time series.

$$\lambda, t_0 \rightarrow \vec{z} \rightarrow \vec{w}$$



Given a **light curve** $\lambda\phi(t - t_0)$ received **photoelectron** (PE), their times being $\vec{z} := (t_1, t_2, \dots, t_N)$. **waveform** $\vec{w}$ is a time series.

$$\lambda, t_0 \rightarrow \vec{z} \rightarrow \vec{w}$$



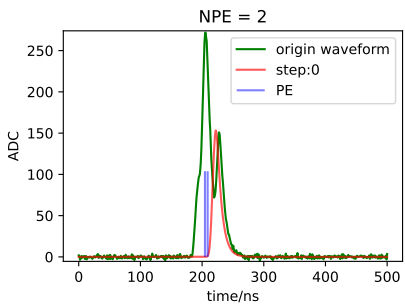
$$\lambda, t_0 \xleftarrow{\text{fit}} \vec{z} \xleftarrow{\text{analyze}} \vec{w}$$

- overlapping pulses, cannot measure $\vec{z}$ precisely. Big challenge.

AI-enabled PE counting and timing extraction

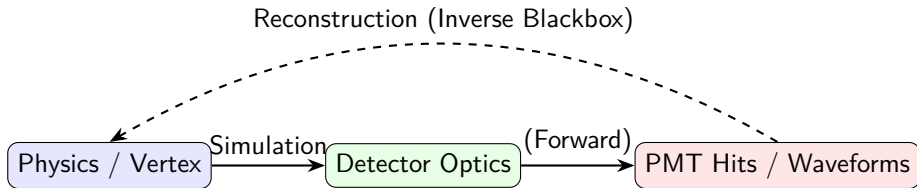
NPE: uncertain count of elements in $\vec{z}$.

$$\begin{aligned} p(\vec{w}|t_0, \lambda) &= \underbrace{p(\vec{w}|\emptyset)p(\emptyset|t_0, \lambda)}_{\text{no integral}} \\ &+ \underbrace{\int_{\vec{z} \in T} p(\vec{w}|\vec{z})p(\vec{z}|t_0, \lambda)d\vec{z}}_{\text{single integral}} \\ &+ \underbrace{\iint_{\vec{z} \in T^2} p(\vec{w}|\vec{z})p(\vec{z}|t_0, \lambda)d\vec{z}}_{\text{double integral}} \\ &+ \underbrace{\iiint_{\vec{z} \in T^3} p(\vec{w}|\vec{z})p(\vec{z}|t_0, \lambda)d\vec{z}}_{\text{triple integral}} \dots \end{aligned}$$



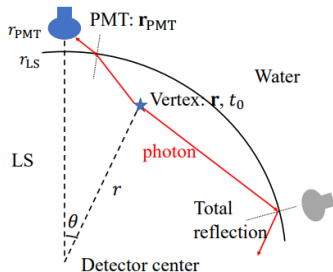
Algorithm	Advantage	Limit
Convolutional NN	Feature extraction, speedy	No uncertainty
Conditional diffusion	Probabilistic uncertainty	Not fast
Transformer-UNet	Multi-scale attention, break pile-up	Systematics validation

Open the Black Boxes

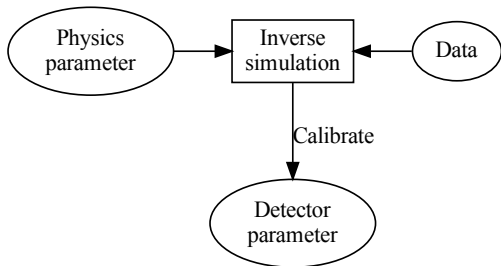
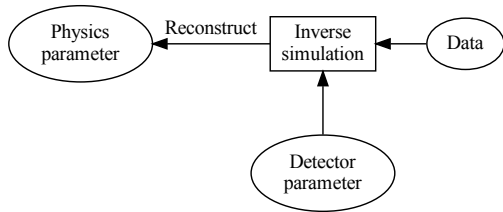
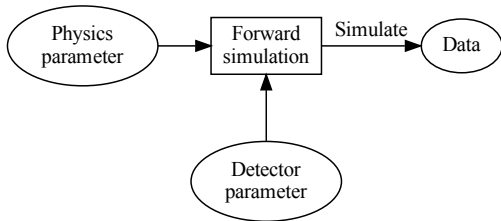


- Deep learning acted as a black box inverse model that may violate causality.

Total internal reflection



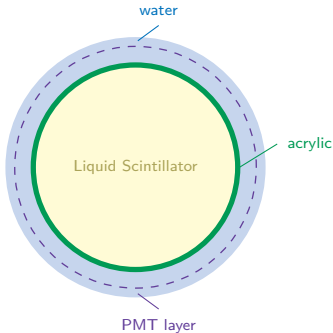
The inversed differentiable simulation



Physics-informed differentiable programming as Inverse simulation

Make the **forward Monte Carlo simulation differentiable**, to solve the inverse problem directly using gradient descent over a physical likelihood.

Physics-informed Hybrid Differentiable Programming Architecture



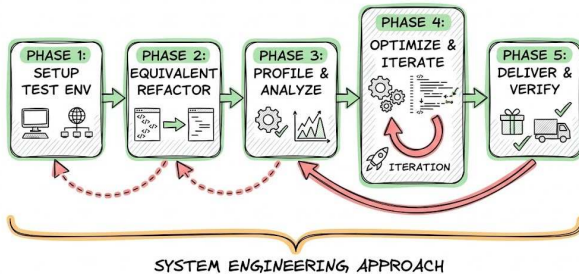
- Gradients $\nabla \mathcal{L}$ flow through the optical model $\rightarrow$ event reconstruction and detector calibration.
- Outer Layer (Surrogate): A conditional neural spline flow learned from Monte Carlo to map boundary states to PMT timing densities.
- Inner Layer (Math): The Radiative Transfer Equation. Neumann series solves photon ballistic and single-scatter paths analytically.

$$\underbrace{\frac{1}{c} \frac{\partial L}{\partial t} + \hat{s} \cdot \nabla L}_{\text{transport}} + \underbrace{\mu_t L}_{\text{loss}} = \underbrace{S(\mathbf{r}, t, \hat{s})}_{\text{source}} + \underbrace{\mu_s \int f(\hat{s} \cdot \hat{s}') L(\mathbf{r}, t, \hat{s}') d\hat{s}'}_{\text{in-scatter}}$$

$$\mu_t = \mu_a + \mu_s, \quad \mu_a = 1/L_{\text{abs}}, \quad \mu_s = 1/L_{\text{scat}}.$$

J. Quant. Spectrosc. Radiat. Transfer 326, 109126 (2024), arXiv:2602.24129.

LLM Agents for Physics Analysis (Dr. Sai)

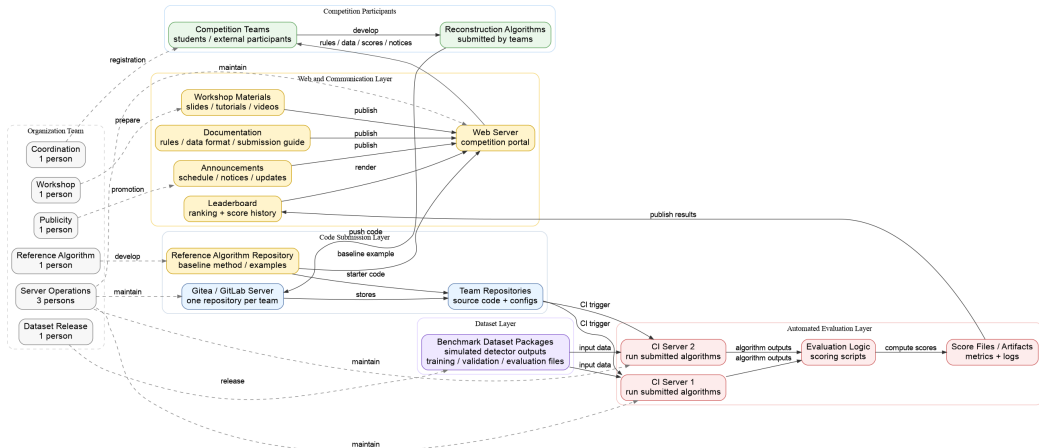


- AI for data → AI for scientists.
- Dr. Sai framework as autonomous agent integrating LLMs with JUNO/HEP environments (ROOT, SNIPEr, HTCondor).

- 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.

Connecting with the Global AI Ecosystem

- A strong global trend towards **Neutrino Open Data**.
 - e.g. Neutrino Physics and Machine Learning 2026 @ UC Irvine.
- AI/CS researchers long for standardized ground-truth labeled datasets.
 - Ghost Hunter: a JUNO-like benchmark since 2019.

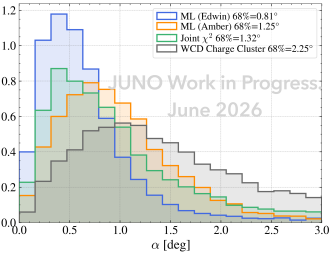
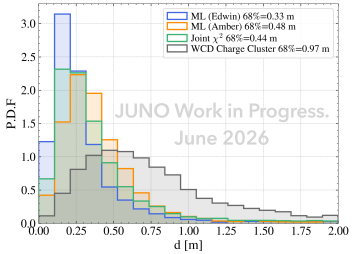
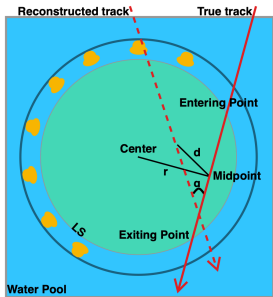


- ① AI is already entering JUNO analysis through μ reconstruction and code optimization.
- ② At R&D frontier, PMT waveform models are moving from CNN regression to probabilistic and sequence-aware inverse modeling.
- ③ Differentiable programming aims to unify reconstruction, calibration and simulation in a single physics likelihood.

AI is redefining the workflow of neutrino physics.

In the past few years, :

- ① Deep Learning (Data-Driven)
 - CNNs, Transformers, and Unet architectures for direct waveform and vertex reconstruction.
- ② Differentiable Physics (Physics-Informed)
 - Unifying Reconstruction-Calibration-Simulation (RCS) via JAX and Differentiable Ray-Tracing.
- ③ Agentic Workflows (LLM-Driven)
 - Autonomous AI agents (Dr. Sai) accelerating high-level physics analysis and optimization.
- ④ Open Science Ecosystem Emerging
 - Standard benchmarks connecting Particle Physics to Computer Science.



Deep Learning for Waveform Analysis

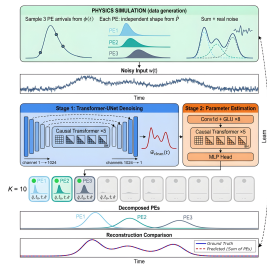
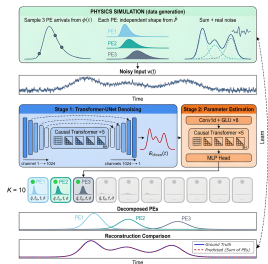


图: Multi-PE Inhomogeneous Poisson Process

图: Transformer-UNet Architecture

- Physics-driven MC training set. We encode the generative process (scintillation time profile $\otimes$ transit time spread) into the training data.
- Two-Stage Transformer
 - ① Waveform denoising (learning real non-Gaussian noise).
 - ② Interpretable parameter estimation (recovering exact PE sequence).
- Achieves superior precision under heavy pile-up compared to traditional deconvolution.

Source term S

Scintillation emitted *isotropically* from vertex V :

$$S = \frac{Y_{\text{eff}}}{4\pi} f_{\text{scint}}(t) \delta^3(\mathbf{r} - V)$$

f_{scint} : time PDF (bi-exponential $\otimes$ Gaussian).

Key property

The in-scatter term is a **linear** operator on $L \Rightarrow$ the solution is a convergent **Neumann series**.

In-scatter term

Photons in direction $\hat{s}'$ redirected into $\hat{s}$ with probability $f(\hat{s} \cdot \hat{s}')$:

- Rayleigh: $f \propto 1 + \cos^2 \vartheta$
- Re-emission (isotropic): $f = 1/(4\pi)$

$$|\nu_\alpha(t)\rangle = \sum_{i=1,2,3} U_{\alpha i}^* |\nu_i(t)\rangle$$

Mixing of different ν mass eigenstates leads to oscillation:

$$\langle \nu_e | \nu_e(t) \rangle = |\langle \nu_e | \nu_1 \rangle|^2 + e^{-i\frac{\Delta m_{21}^2 L}{2E}} |\langle \nu_e | \nu_2 \rangle|^2 + e^{-i\frac{\Delta m_{31}^2 L}{2E}} |\langle \nu_e | \nu_3 \rangle|^2$$

describes the **survival probability** of ν_e as $P(\nu_e \rightarrow \nu_e) = |\langle \nu_e | \nu_e(t) \rangle|^2$.

