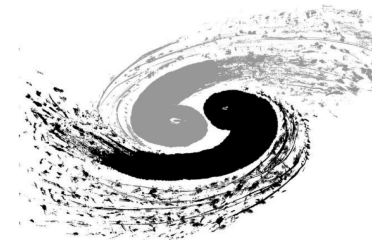


ML4Jets 2026
Vienna, Austria

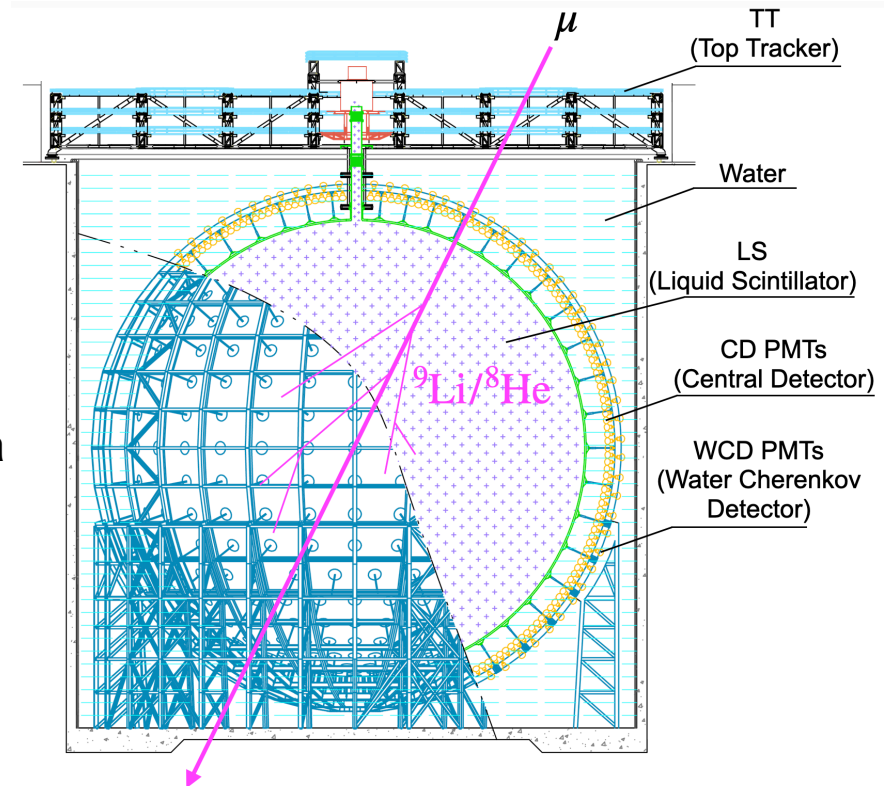


Amber: Self-Attention Muon Track Reconstruction in JUNO

Xiaoying Lu (IHEP China)
on behalf of the JUNO Collaboration
2026/9/15

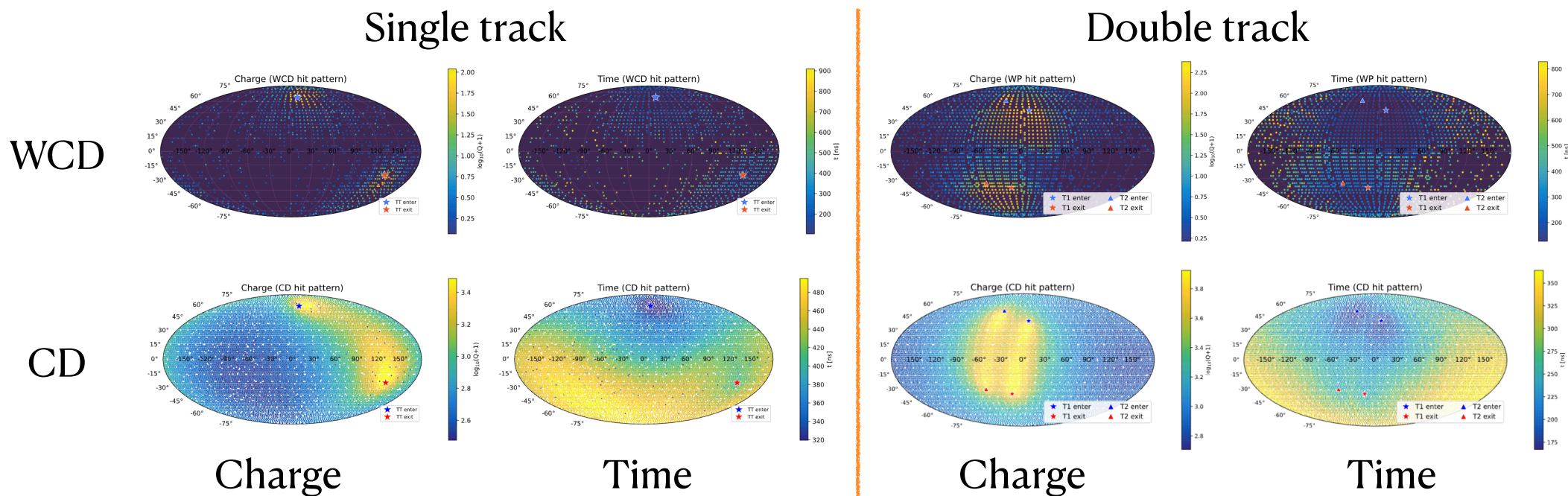
Motivation and physical picture

- JUNO: 20-kton liquid scintillator detector determining the Neutrino Mass Ordering
- Cosmic-muon-produced ${}^9\text{Li}/{}^8\text{He}$ isotopes mimic the Inverse Beta Decay signal, limiting this physics sensitivity
- Muon track reconstruction for background suppression
 - TT: Multiple layers of plastic scintillator strips -> precisely detect subset muon tracks
 - WCD PMTs: Detects Cherenkov light -> muon hit patterns
 - CD PMTs: Detects scintillation light -> high-multiplicity muon hit patterns



Features

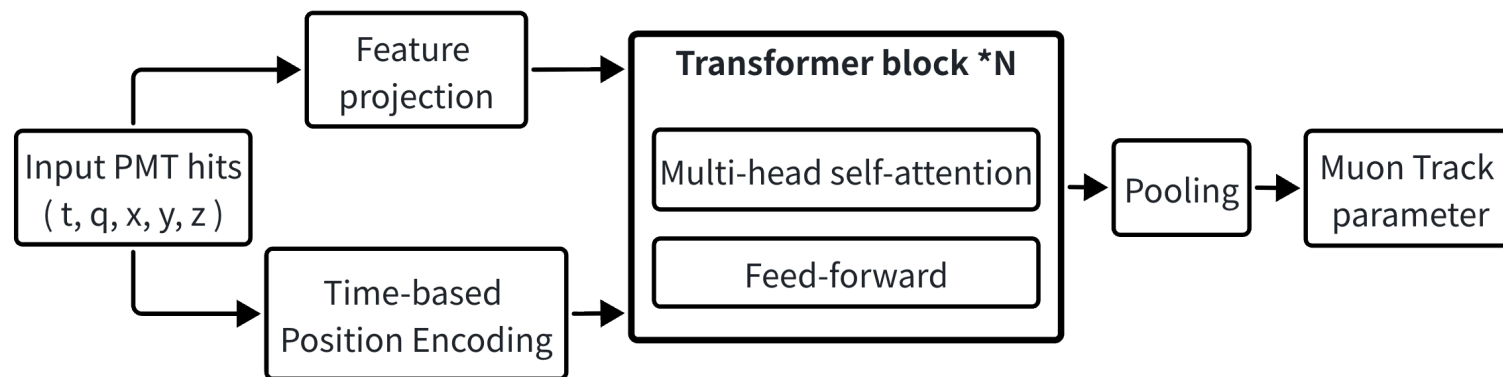
- Muon is observed as a detector-wide pattern of PMT hits carrying spatial, timing and charge information



Amber model architecture

(Attention Mechanism Based Event Reconstruction)

- Model the correlations among PMT hits
- Learning long-range dependency
- Output track parameters



Amber overview

	WCD PMTs Cherenkov light	CD PMTs Scintillation light
Amber-C Classification	 Done	 To do
Amber-S Single-track reconstruction	 Done Basic reconstruction	 Done
Amber-D Double-track reconstruction	 Done Complex topology	 To do

Inspired by DeepSeek-OCR [1]:

- Reduce training time for large CD PMT hits
- Improve reconstruction performance

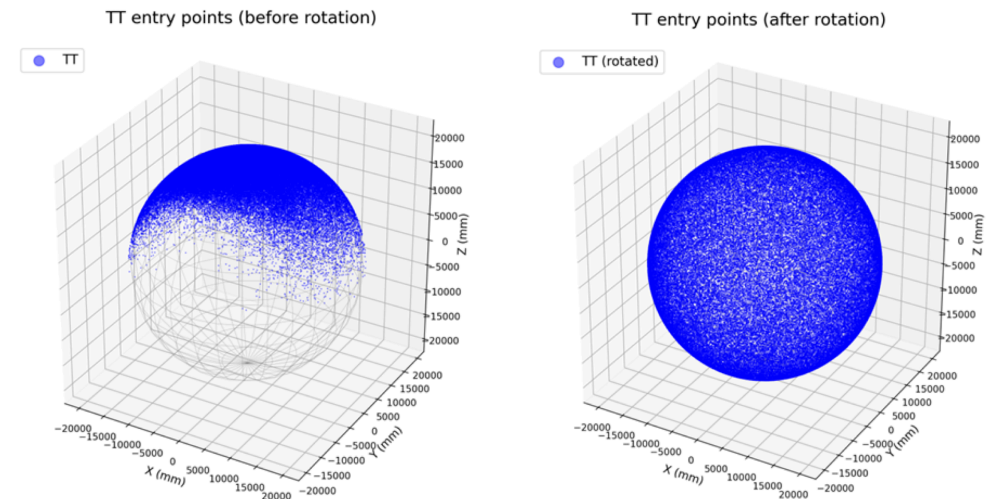
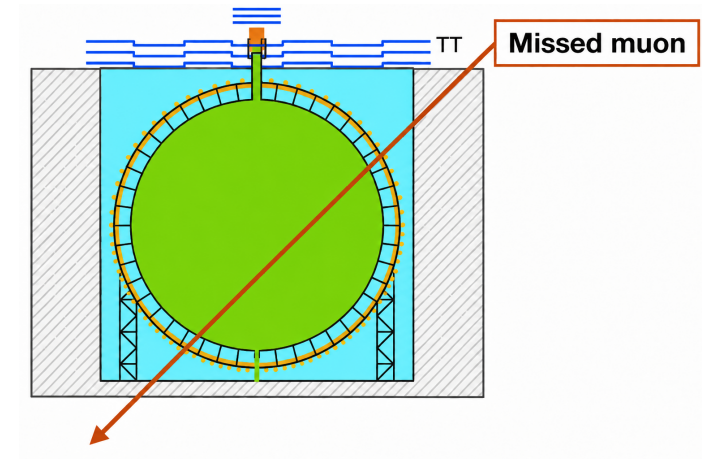
[1] arXiv:2510.18234

Data-driven method

- Learn directly from real detector response
- Label: Precise TT tracks from real data
- Limited coverage of TT => Augmentation
 - Random 3D rotation

$$\{\mathbf{r}_i, \mathbf{p}_{in}, \mathbf{p}_{out}\} \xrightarrow{R \in SO(3)} \{R\mathbf{r}_i, R\mathbf{p}_{in}, R\mathbf{p}_{out}\}$$

- Same rotation applied to PMT positions and track labels
- New random 3D rotation for every sample, every epoch



Before rotation

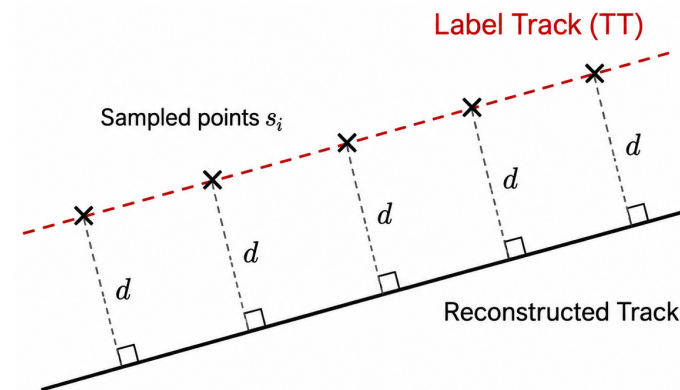
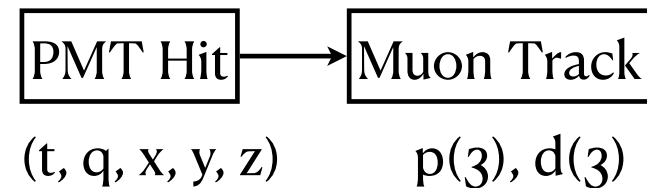
After rotation

Amber-S

- Training data
 - WCD PMT hit feature: (t, q, x, y, z)
 - Label: TT tracks

- Loss function

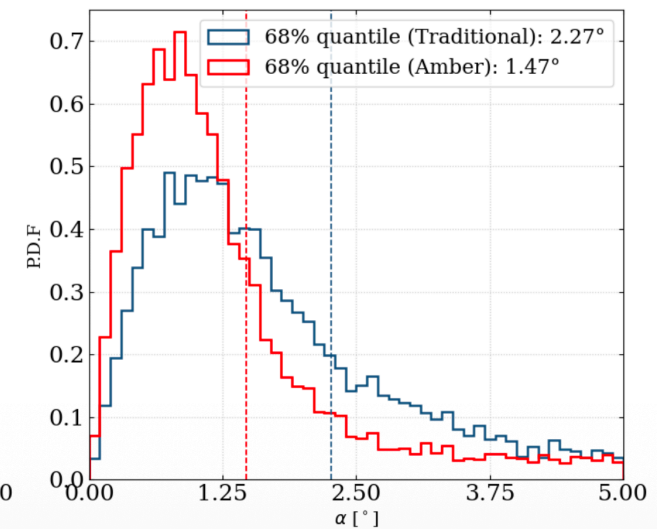
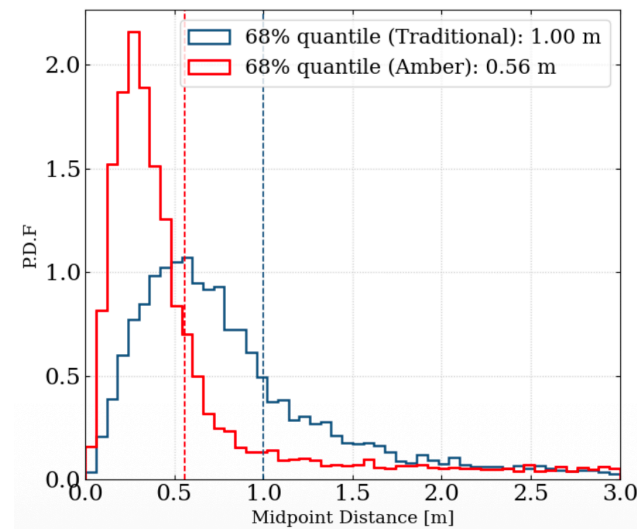
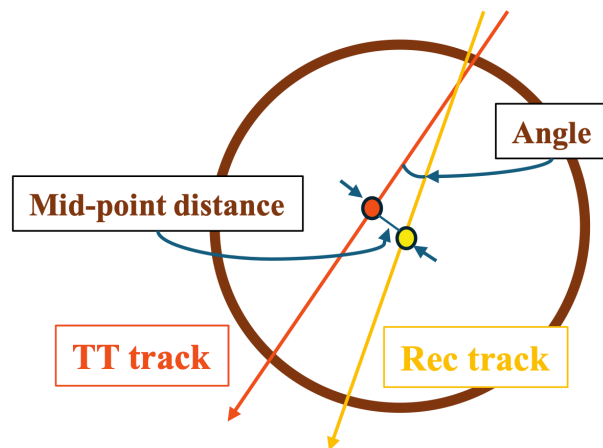
$$\mathcal{L}_S = \frac{1}{N_s} \sum_{i=1}^{N_s} \frac{\text{distance} (s_i, \text{line}_{\text{pred}})}{S_0}$$



Amber-S performance

- Accurately reconstructs single through-going muons from WCD hit features

	Mid-point distance	Angle
Traditional	1.00 m	2.27 deg
Amber-S	0.56 m	1.47 deg



Amber-D

- Single track + Single track $\rightarrow$ paired Double track

- Label uses TT; feature uses hit pattern

- $q = q_1 + q_2$; $t = \min(t_1, t_2)$

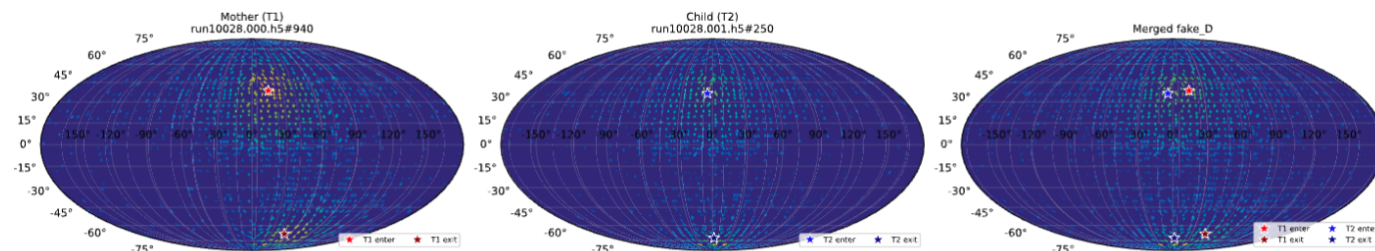
- **60 times** the MC data

- 3D rotation augmentation

- Loss function

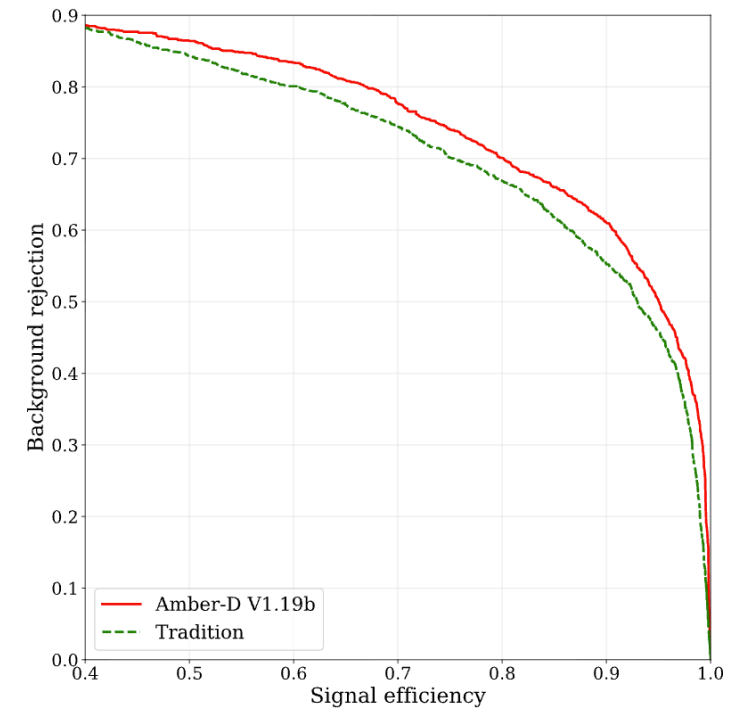
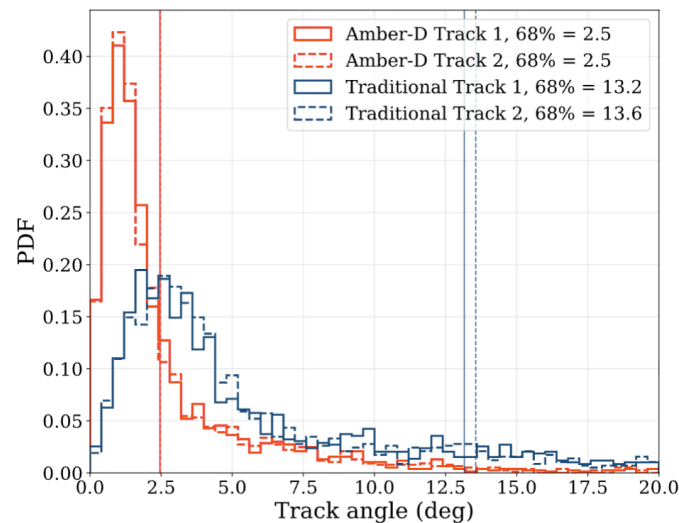
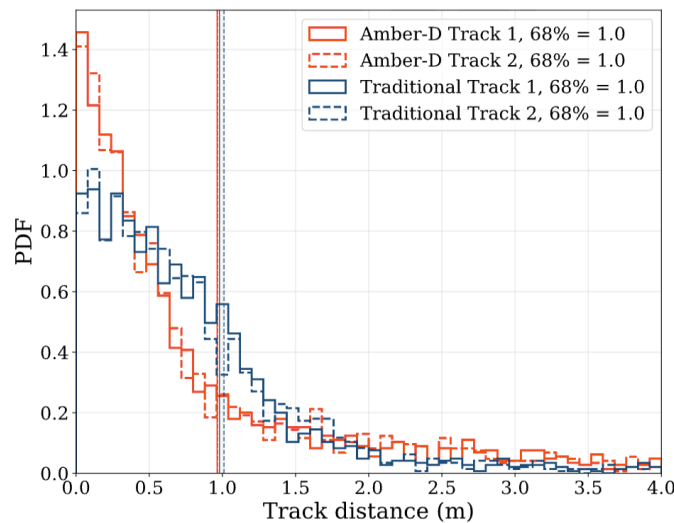
$$\mathcal{L}_D = w_{\text{mid}} \cdot \min(\mathcal{L}_A, \mathcal{L}_B) + w_{\text{dir}} \cdot \mathcal{L}_{\text{dir}}$$

- Permutation-invariant midpoint loss over two track assignments



Amber-D performance

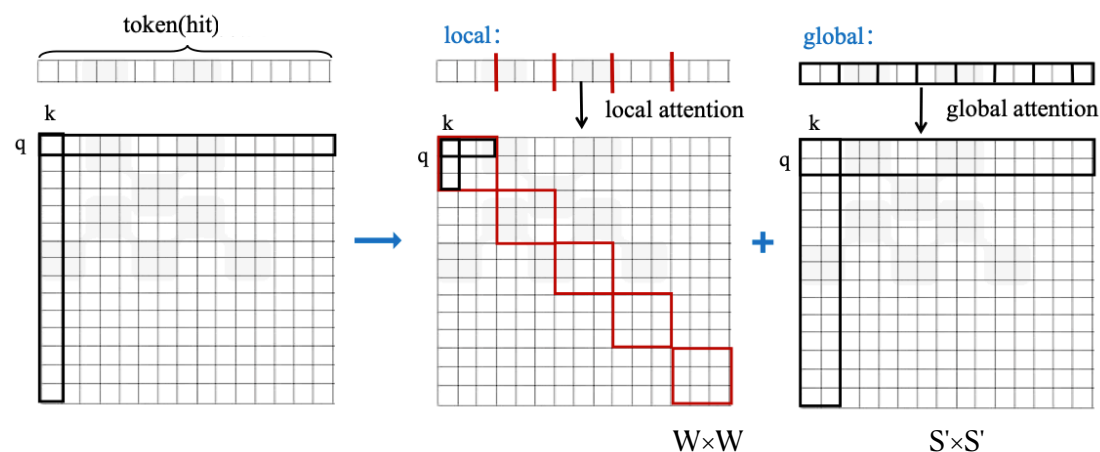
- Better agreement with TT track geometry
 - Fewer total track-distance pairing modes
- Improved spallation neutron suppression



Amber-OCR

- Introduce OCR to the CD model
- Partition tokens into windows
- Attention $\rightarrow$ local attention + convolution + global attention
 $O(S^2)$ $O(S \cdot W)$ 3×3 $O(S'^2)$

Symbol	Value	Meaning
S	17612	CD hits
W	512	window size
S'	1954	tokens after conv

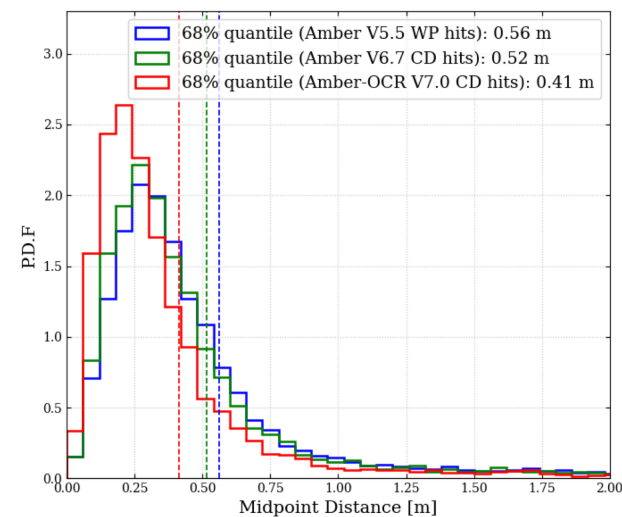
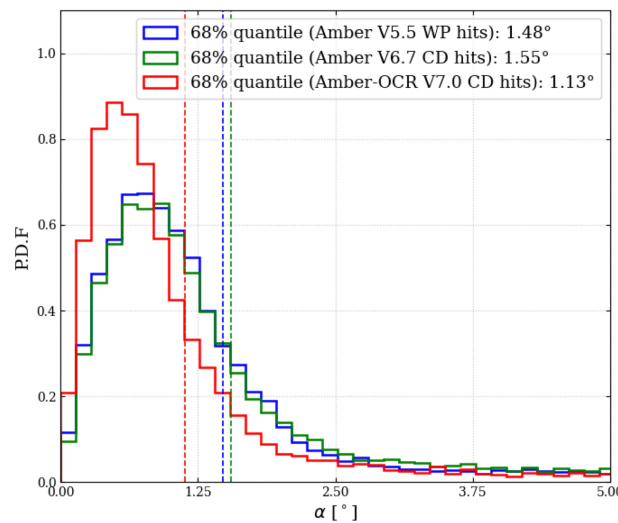
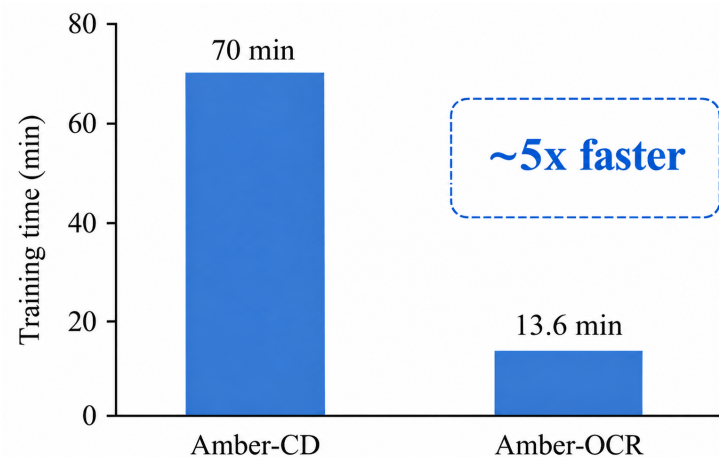


Amber-OCR performance

- CD-based Single-Muon Reconstruction
 - Faster training
 - Higher reconstruction performance

Model	Angle	Mid-point distance
Amber-WCD	1.48 deg	0.56 m
Amber-CD	Un-optimized	
Amber-OCR	1.13 deg	0.41 m

Better



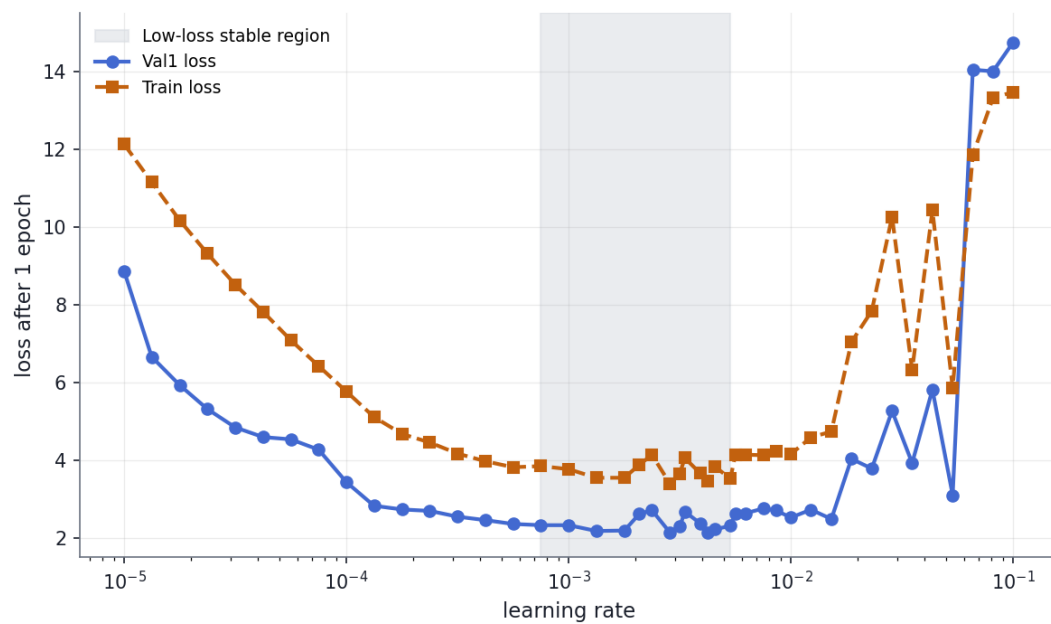
Summary

- Fully data-driven training strategy using real TT-labeled events
 - Amber-S and Amber-D
 - Best-performing WCD-based reconstruction method
 - Adopted in JUNO's official reconstruction
 - Used in JUNO physics results presented at Neutrino 2026
 - Amber-OCR shows promising performance on high-multiplicity CD PMT inputs
 - $\sim 5 \times$ faster training
 - Further improved reconstruction accuracy
 - Extending Amber to more reconstruction tasks and advanced architectures
- Thanks!

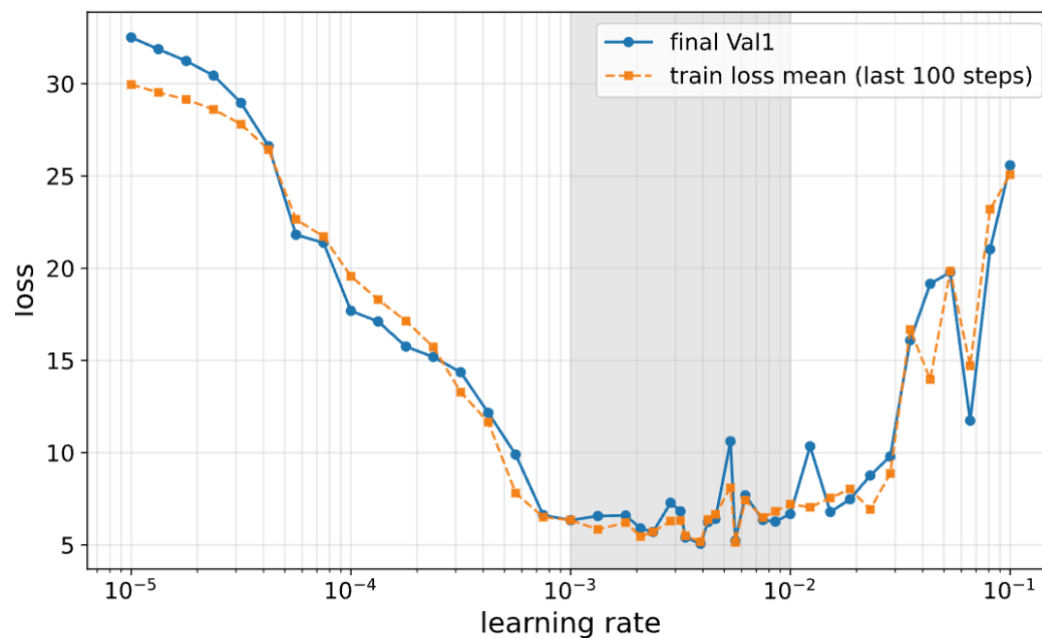
Back Up

Learning-rate sweep

- Stable optimization over a broad LR range



Amber-S



Amber-D