Neutron reconstruction at tau-charm facilities with Deep Learning

Yangu Li (李彦谷)
Peking University & University of Chinese Academy of Sciences

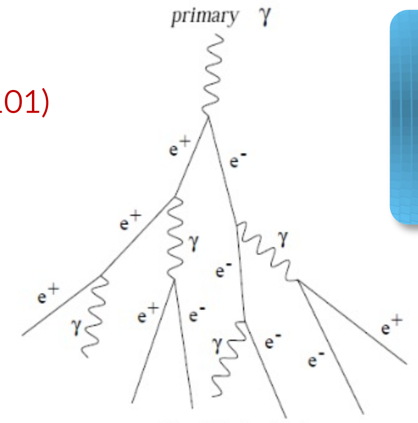
FTCF 2024, Guangzhou
November 21, 2024

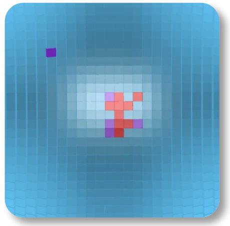
Introduction

- **Long-lived neutral hadrons $(n,\ K_L^0)$ are important probes for physics at τ-charm region**
  - e.g., about 1/3 of $\bar{\Lambda}_c^-$ decays contain $\bar{n}$ where 20% of them are still unknown (PRD **108**, L031101)

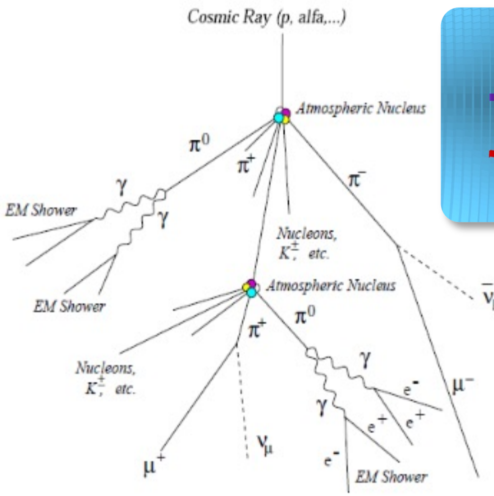
- **However, τ-charm facilities have no dedicated hadronic calorimeter**
  - Detection relies on electromagnetic calorimeter (ECAL, EMC)
  - Its size & material prevent full deposition of hadronic showers

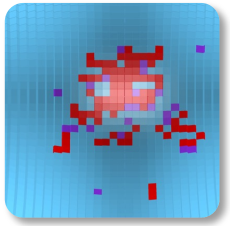
- **Direct reconstruction of neutral hadron is very challenging**
  - Momentum is unknown
    - Sizable energy leakage
  - Position isn't always known
    - Clusters are less centralized than photons
  - Identification is not perfect
    - Can be confused with photon / beam background / detector noises
  - MC simulation is imprecise
    - Up to $\sim 10\%$ discrepancy from data (NIMA **1033**, 166672)





EM shower (↑) vs. hadronic shower (↓)

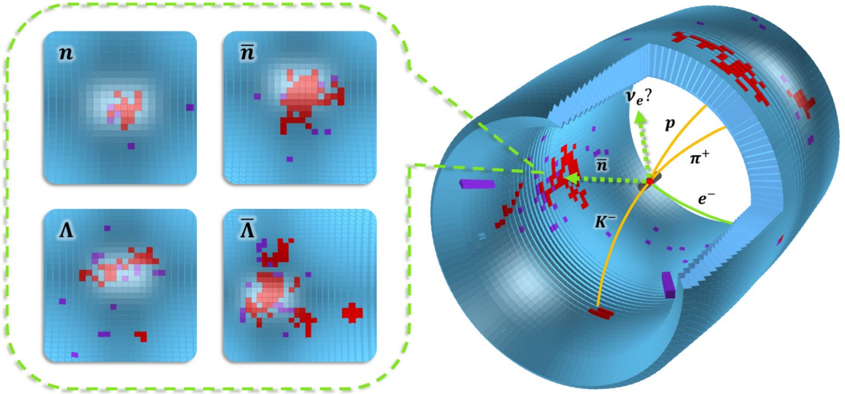




Our works

Observation of a rare beta decay of the charmed baryon with a Graph Neural Network

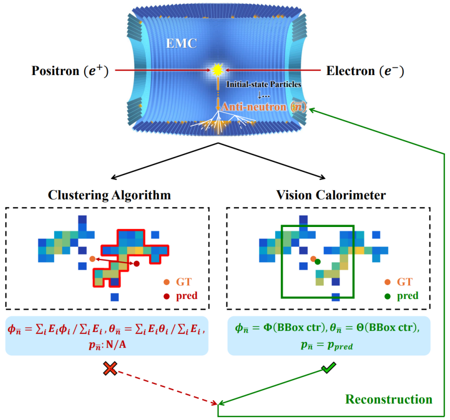
arXiv: [2410.13515](2410.13515), accepted by Nature Commun.



- Practical solution for neutron identification
- Enables an important charm study at BESIII

Vision Calorimeter for Anti-neutron Reconstruction: A Baseline

arXiv: [2408.10599](2408.10599), submitted to AAAI 2025 Conference



- Precise anti-neutron position measurement
- First realization of momentum prediction ability

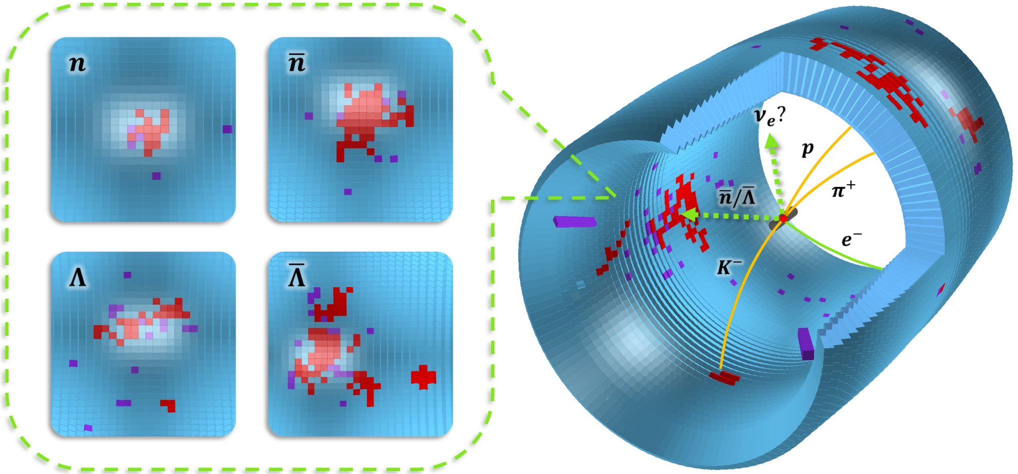
The studies are based on BESIII data,
but are also applicable to STCF with similar EMC designs.

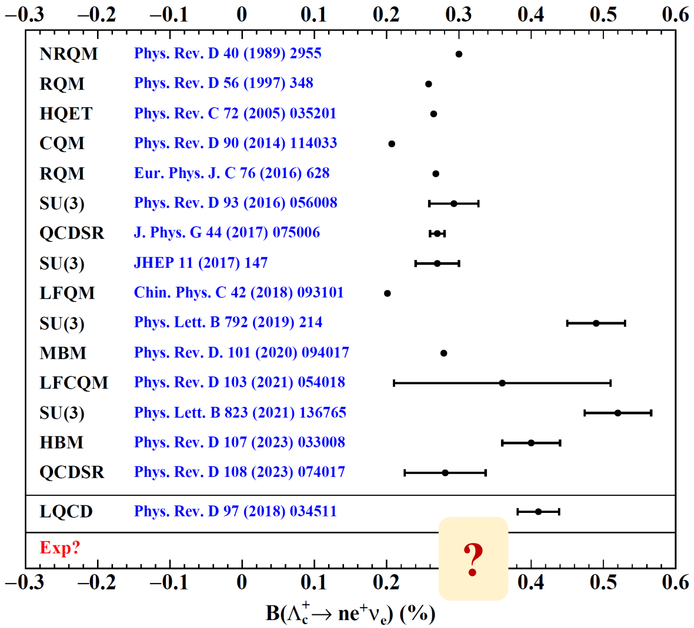
Neutron identification via Graph Neural Network

Task overview

- **The physics goal:** measure $\Lambda_c^+ \to n e^+ \nu_e$
  - The second-most-dominant Λ_c^+ semi-leptonic decay is still unobserved
  - Tons of theoretical predictions wait to be confirmed & calibrated

- **Experimental challenge**
  - Simultaneous existence of neutron & neutrino
  - Dominant background from $\Lambda_c^+ \to \Lambda(n\pi^0)e^+\nu_e$, yield ~10x signal
  - Need very powerful tool to identify neutron from photon & noises



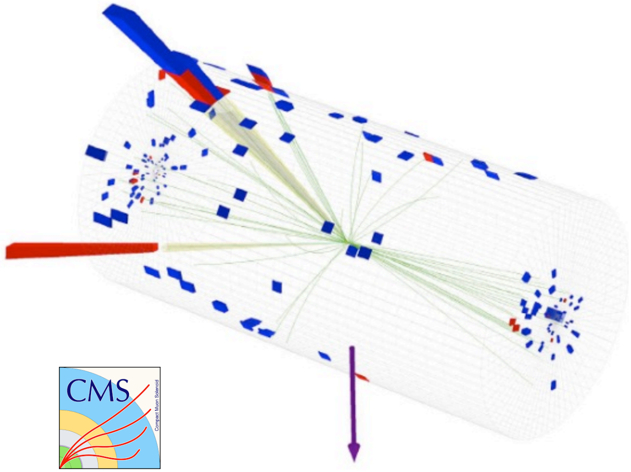


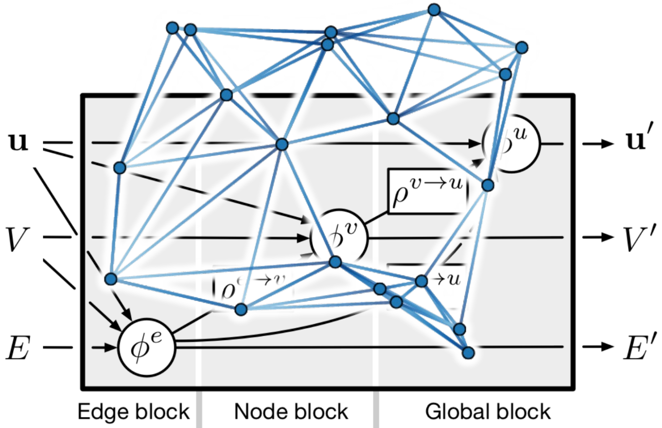
Distinction in n/Λ EMC patterns can be seen from eyes.

Deep learning may recognize such distinctions in a smart & flexible way.

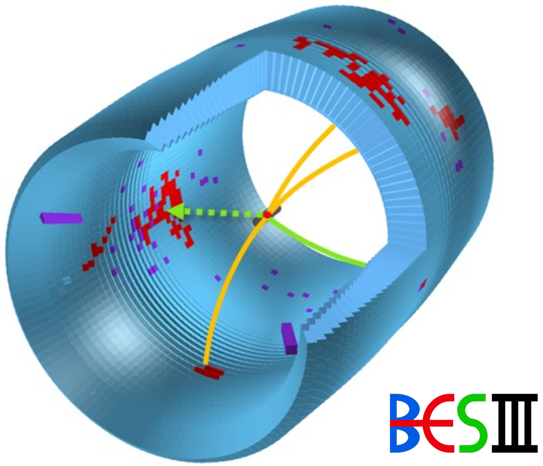
Why Graph Neural Network (GNN)

- **Our task parallels jet tagging in LHC experiments at a new energy scale**
  - EMC cells are in regular grids, but users can only access reconstructed shower objects
  - The graph-based model ParticleNet beats image/sequence-based ones in jet tagging (PRD **101**, 056019)
  - GNN can represent more arbitrary relations between data objects as nodes & edges in a graph



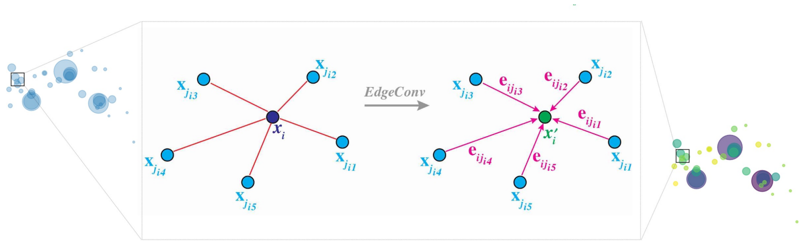


arXiv: 2007.13681



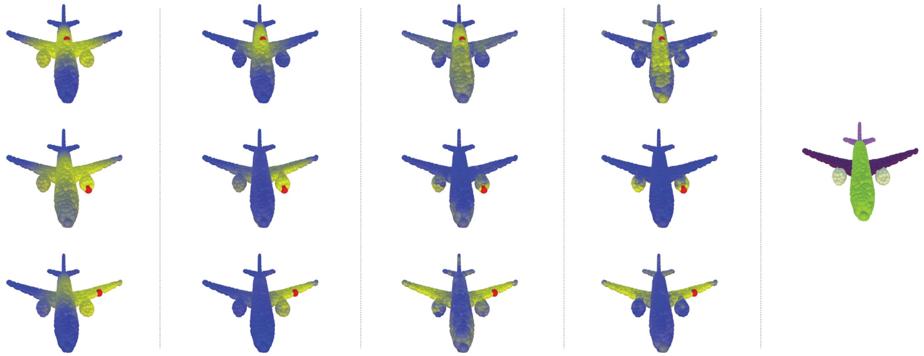
Our GNN toolkit

- **Event representation – point cloud**
  - Unordered, permutation-invariant set of particle showers
  - Each shower carries spatial coordinates + additional features
    - Energy, time, number of hits, cluster expansion moments…
  - Symmetry-preserving, high expressiveness, low computational cost

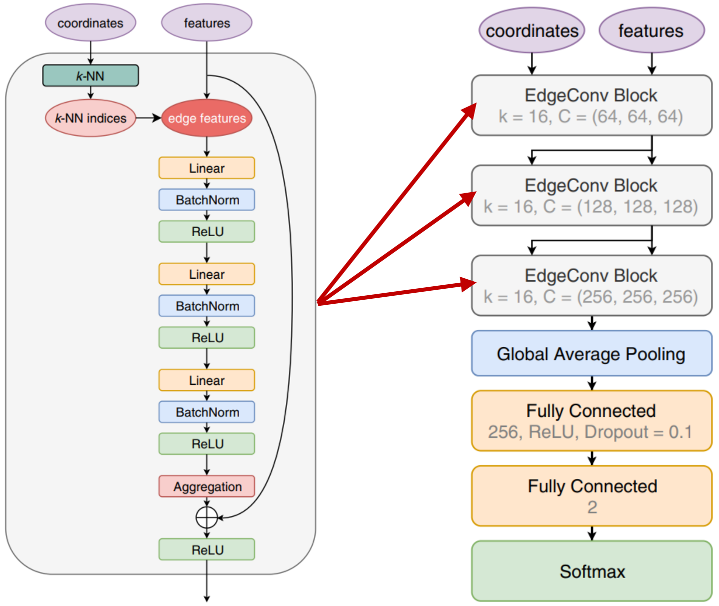


$e_{ij} = \mathrm{MLP}(x_i, x_j)$ $\quad x_i' = \mathrm{mean}_j\, e_{ij}$

- **Model structure – Dynamic Graph CNN** (arXiv: 2007.13681)
  - Build “edge features” between k-nearest neighboring points
  - Design a symmetric “convolution” function on edges
  - Dynamically update the graph after each convolution block



Neighboring in feature space → Neighboring in semantics

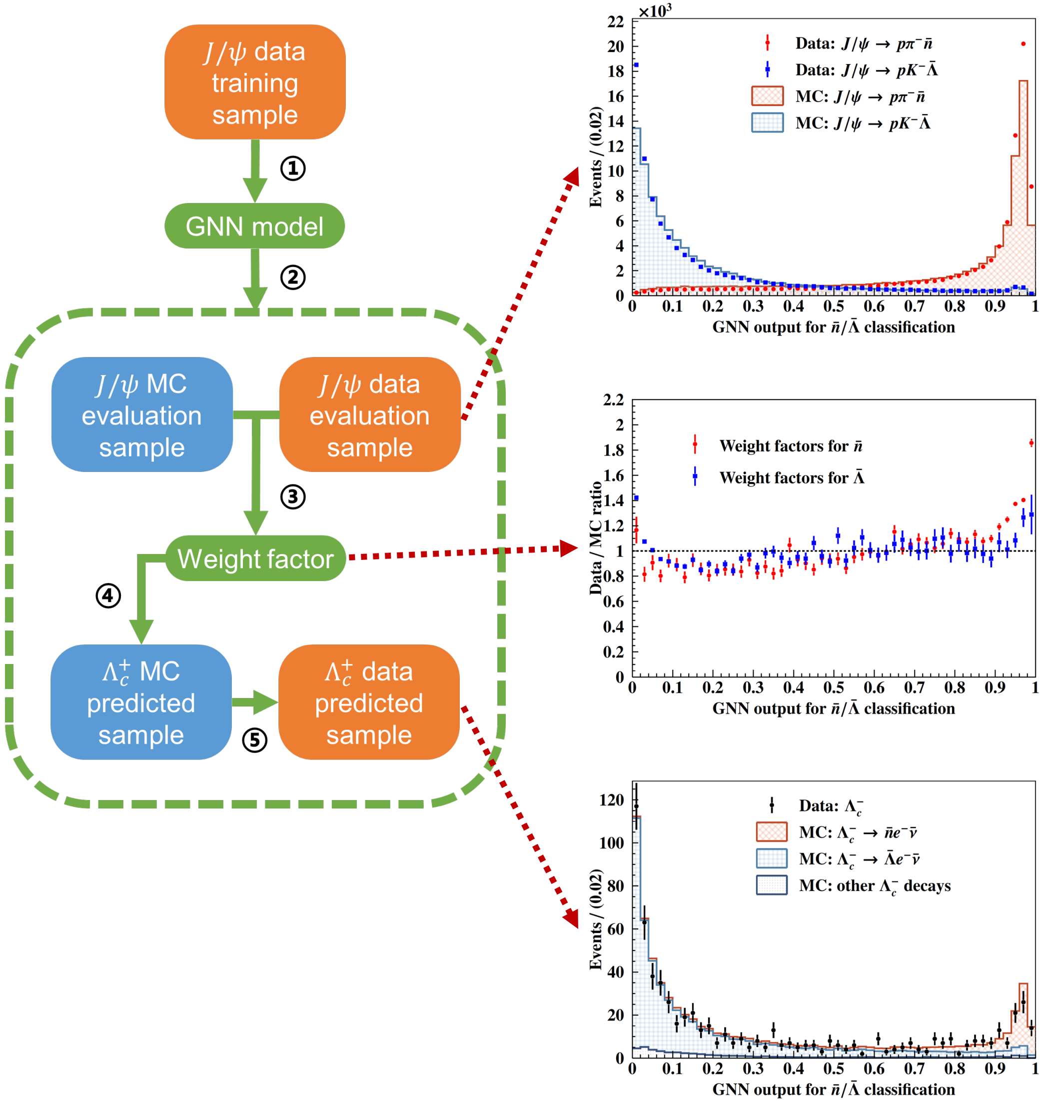


Data-driven analysis pipeline

- **It's easy to train n/Λ classifier with MC samples and apply to data**
  - But the MC simulation is imprecise!

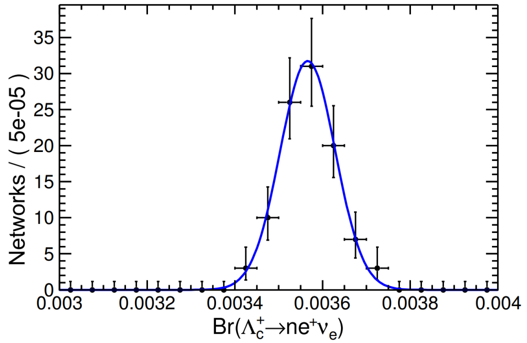
- **Utilize control samples from 10 billion J/ψ events at BESIII**
  1. Train GNN model with $J/\psi \to \bar{p}n\pi^+$ vs. $\bar{p}\Lambda K^+$ events from real data
  2. Use the model to predict J/ψ data, J/ψ MC, Λ_c^+ data and Λ_c^+ MC samples
  3. Weight the GNN responses between J/ψ data & MC
  4. Correct Λ_c^+ MC shape using bin-by-bin weight factors
  5. Fit to Λ_c^+ data with Λ_c^+ MC shapes

Conjugate channels are treated separately,
as anti-neutron may annihilate with EMC while neutron won't

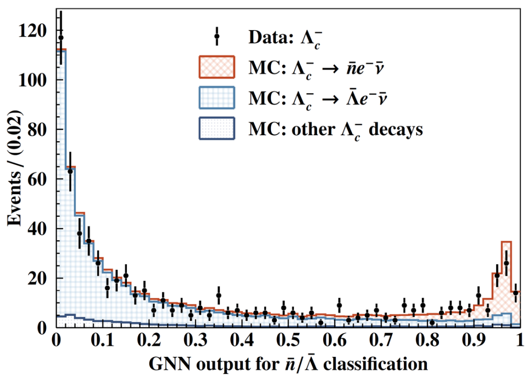


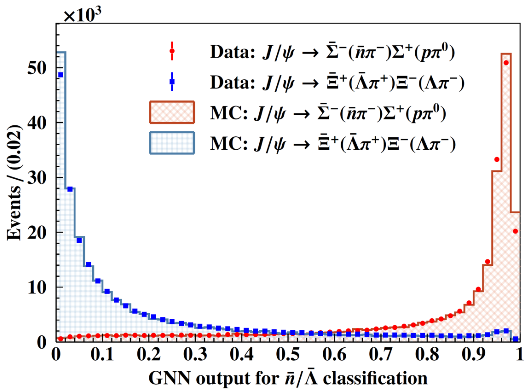
Quantify ML-related systematic uncertainties

- **Model uncertainty**
  - Lack of knowledge about best model configurations
  - Incorporate randomness in training
    - Data processing sequence, network initialization, dropout…
  - Ensemble different models in prediction
    - Physics results form a Gaussian distribution

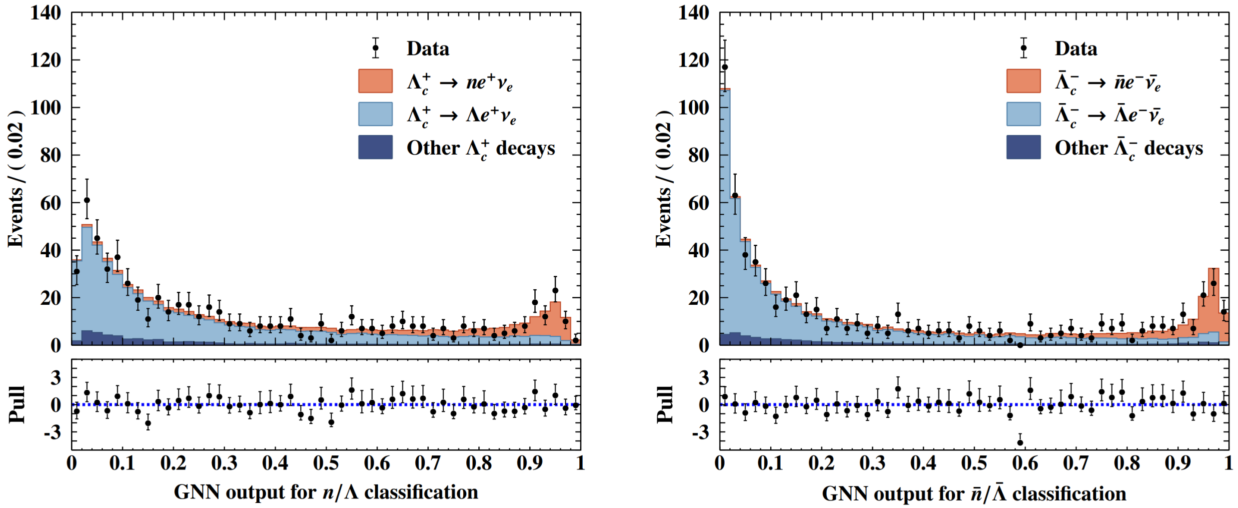


- **Domain shift**
  - Residual diffs between J/ψ & Λ_c^+ datasets (kinematics, BKG environment…)
  - Replace Λ_c^+ datasets with another independent control sample: $J/\psi \to \Sigma^+(n\pi^+)\bar{\Sigma}^-(\bar{p}\pi^0)$ vs. $J/\psi \to \Xi^-(\Lambda\pi^-)\bar{\Xi}^+(\bar{\Lambda}\pi^+)$
  - Data & MC still well consistent in large statistics after correction

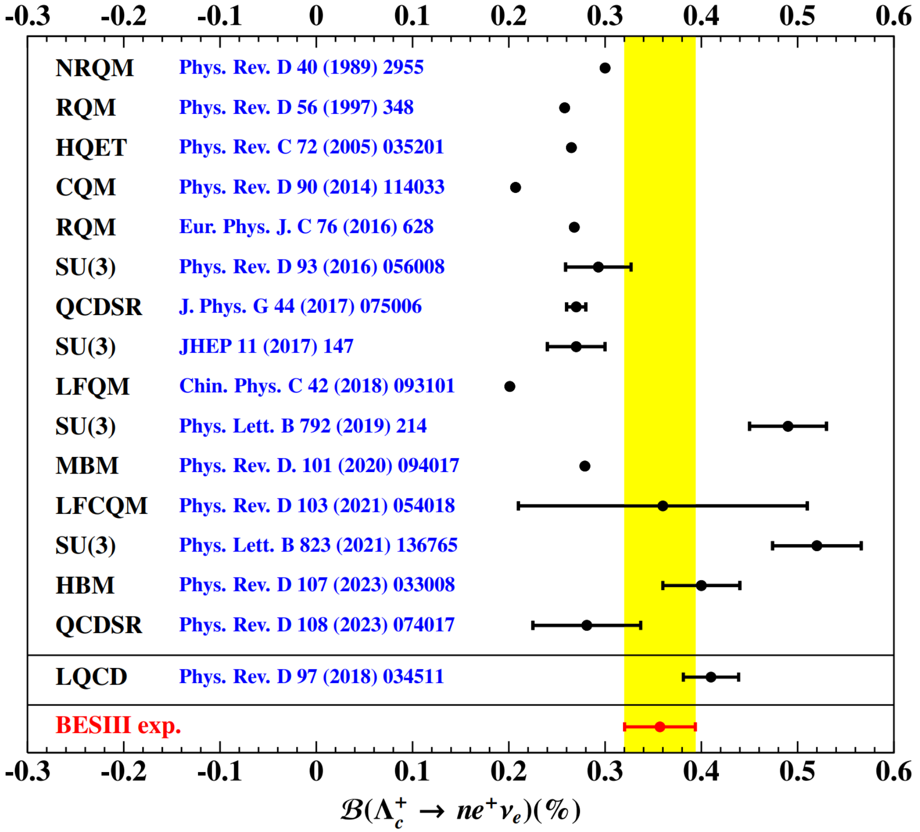




Physics outcomes



- **First observation of $\Lambda_c^+ \to n e^+ \nu_e$ with over 10σ significance**
  - $\mathcal{B}(\Lambda_c^+ \to n e^+ \nu_e) = \left(0.357 \pm 0.034_{\text{stat.}} \pm 0.014_{\text{syst.}}\right)\%$
- **First measurement of $|V_{cd}|$ from charmed baryon decays**
  - $|V_{cd}| = 0.208 \pm 0.011_{\text{exp.}} \pm 0.007_{\text{LQCD}} \pm 0.001_{\tau(\Lambda_c^+)}$



The study is based on 4.5 fb^{-1} data at BESIII.
One-year run at STCF can improve BR and $|V_{cd}|$ precisions to ~3% (systematics dominant)

One step forward

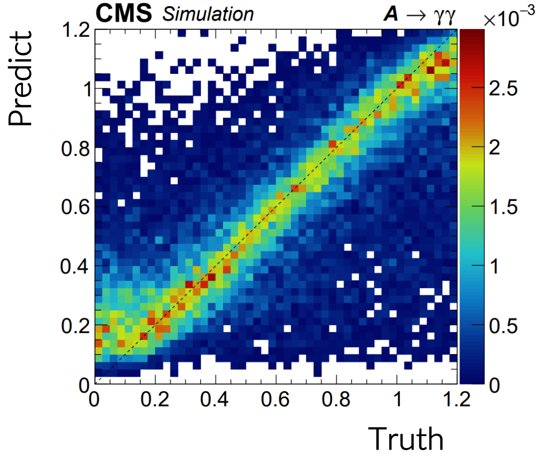
- **Full neutron reconstruction beyond identification is desired.**
  - More meaningful physics results (e.g., form factors) require knowing the neutron momentum

- **We tried predicting neutron momentum with GNN, but failed**
  - A regression task that doesn't perfectly fit our toolkit

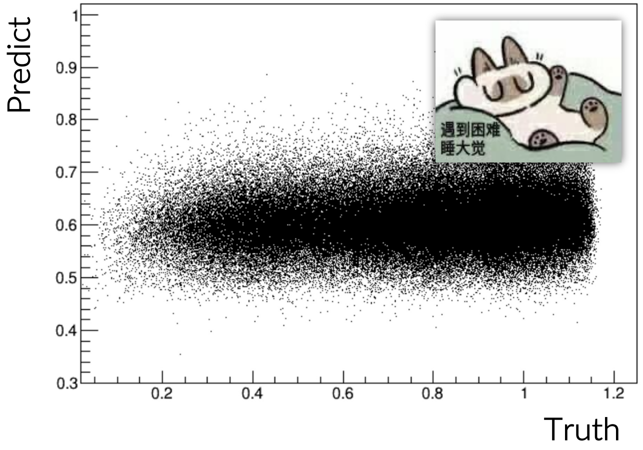
Does the limitation come from detector, or our deep learning technique?

We should seek help from ML experts.

Ideal effect:
Positive correlated mapping



Our result:
Assign all outputs near average
to yield local minimum for loss function

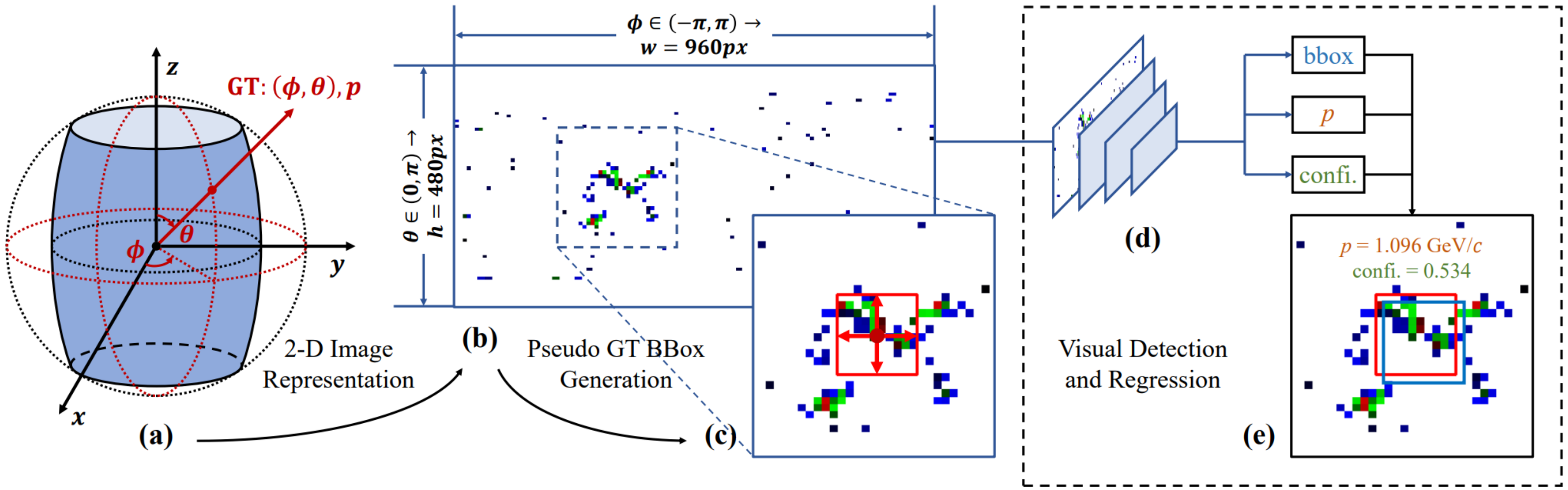


Anti-neutron reconstruction via Visual Object Detector

Vision Calorimeter (ViC)

- **An object detection approach**
  - Represent EMC hits on a 2D image
  - Find the position of $\bar{n}$ within a binding box
  - Determine its confidence score, class and incident momentum

A comprehensive reconstruction with particle type, position and momentum measurements.



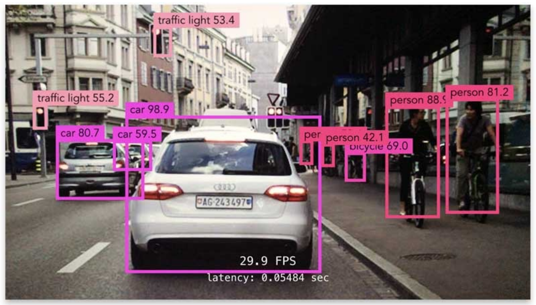
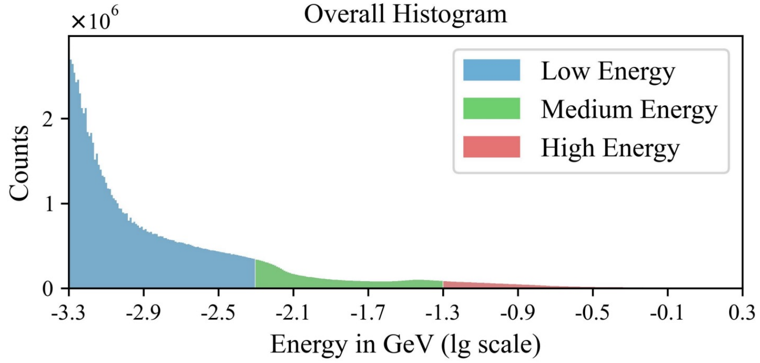


Image quantification

- **Pixels**
  - BESIII barrel EMC has 44 rings × 120 cells, end-cap EMC has 6 rings × [96, 96, 80, 80, 64, 64] cells
  - Set image size with 960 × 480 pixels
    - 960 is the least common multiple of (120, 96, 80, 64)
  - Define position-varied cell height according to their center positions

layers	cells	w (pixels)	h (pixels)	note
2	-	30	8	empty
2	-	24	8	
3	-	20	7	
2	64	15	6	end-cap
2	80	12	6	
2	96	10	5	
1	-	10	5	empty
5	120	8	5	barrel
4	120	8	6	
5	120	8	7	
16	120	8	8	
5	120	8	7	
4	120	8	6	
5	120	8	5	
1	-	10	5	empty
2	96	10	5	end-cap
2	80	12	6	
2	64	15	6	
3	-	20	7	empty
2	-	24	8	
2	-	30	8	

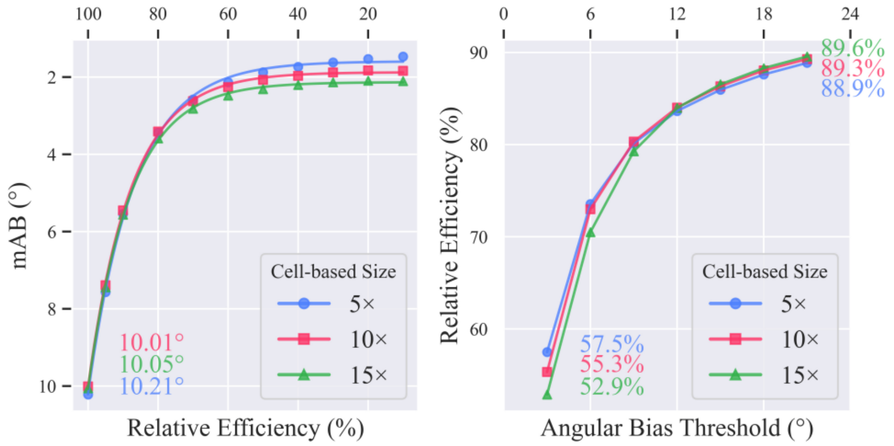
- **Colors**
  - EMC deposited energy range is 0.5 MeV ~ 2 GeV
  - Take log scale: $[10^{-3.3}, 10^{0.3}]$
  - Divide low, medium and high measures to fill blue, green and red channels
  - Add a -30db Gaussian noise due to the sparsity of on-fire EMC cells



Bounding box generation

- **Why not just make point-wise position prediction?**
  - Bounding box (BBox) prediction can better exploit contextual information
  - Superior performance in afterward tests
  - Need to generate pseudo BBox around $\bar{n}$ incident position

- **Choice of BBox width**
  - Smaller size → higher precision upper limit
  - Larger size → more available contextual information
  - Best performance at 10× cell-based size



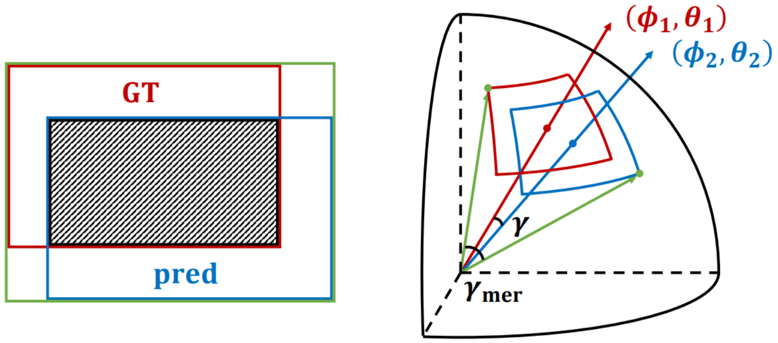
The performance of incident position prediction with different pseudo GT BBox sizes. **Left**: mAB at different relative efficiency levels; **Right**: relative efficiency with different angular bias thresholds.

Network configuration

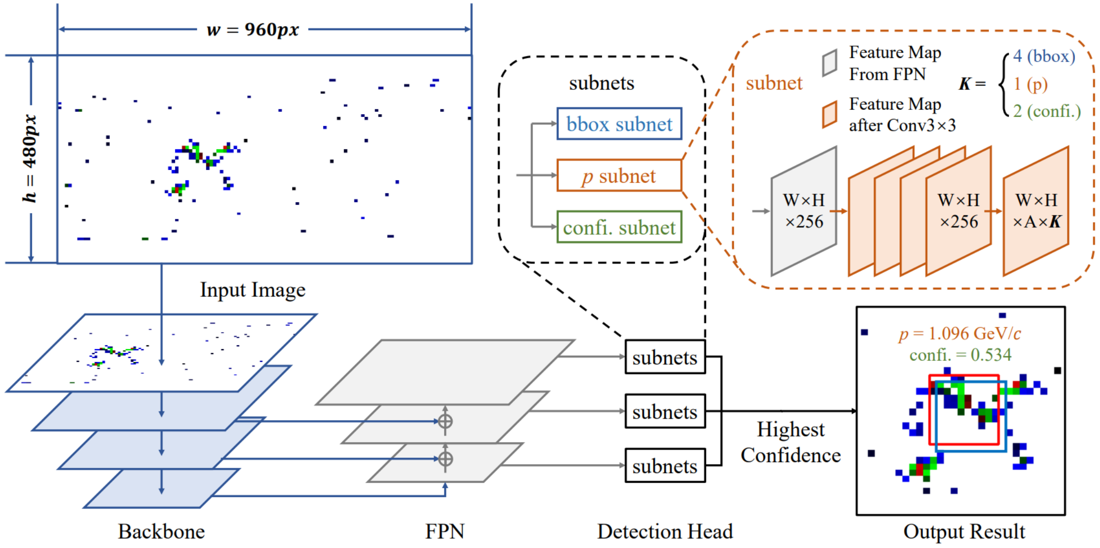
- **Architecture**
  - Backbone: Swin-Transformer pre-trained on ImageNet
  - Detection head: RetinaNet

- **Loss function**
  - Conventional choice in object detection is IOU
    - $IOU = S(GT \cap Pred)/S(GT \cup Pred)$
  - We design a more center-oriented version

$$\mathcal{L}_{CO} = 1 - IoU + \alpha \cdot \frac{(\cos\gamma - 1)^2}{(\cos\gamma_{\text{mer}} - 1)^2}$$



(a) (b)

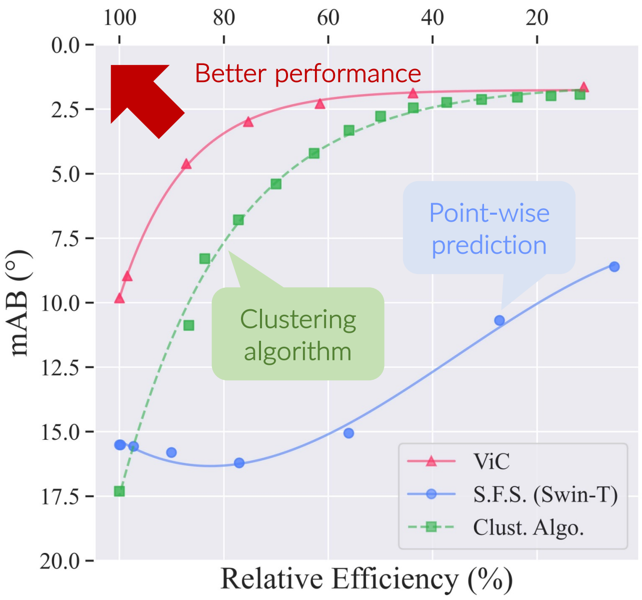


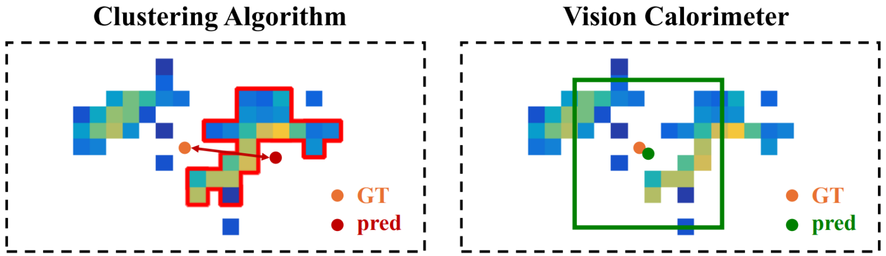
Performance of ViC (I)

- **Dataset**
  - 1 million $J/\psi \to p\bar{n}\pi^-$ events taken from BESIII data
  - Plan to extend to 10 million events

- **In position measurement**
  - Compared with conventional clustering algorithm, ViC improves the precision by 75% at full efficiency
    - From 17.4° to 9.9°
  - ViC can double this precision at 90% efficiency
    - A near-practical performance of 5°

- **How comes the improvement?**
  - Conventional clustering algorithm would split a discontinues hadronic shower
    - Usually caused by multiple scattering
    - Only the most energetic one is considered
  - ViC can better handle such scenarios





$\phi_{\bar{n}} = \sum_i E_i\phi_i / \sum_i E_i, \theta_{\bar{n}} = \sum_i E_i\theta_i / \sum_i E_i,$ $p_{\bar{n}}$: N/A

$\phi_{\bar{n}} = \Phi(\text{BBox ctr}), \theta_{\bar{n}} = \Theta(\text{BBox ctr}),$ $p_{\bar{n}} = p_{pred}$

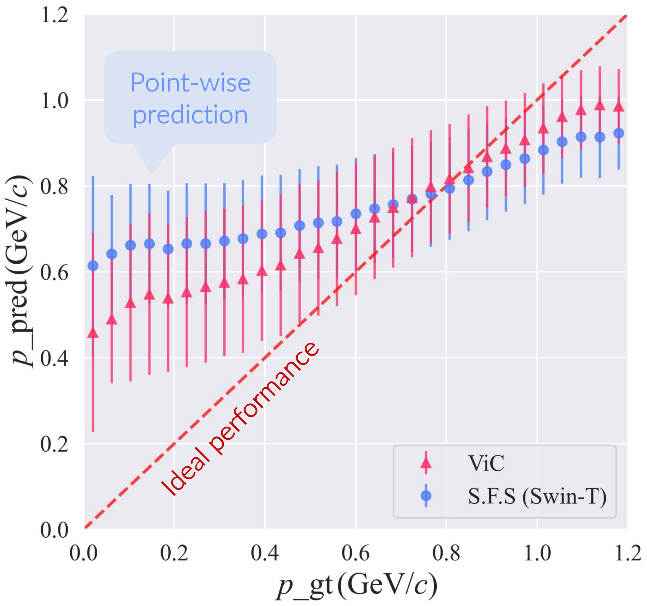
Performance of ViC (II)

- **In momentum measurement**
  - No conventional methods so far
  - ViC firstly realize such capability in EMC
  - Resolution ~25% in sub-GeV region
    - Even better than dedicated HCALs in this region ($\sim 80\%/\sqrt{E}$) ?!

- **In classification**
  - ViC is capable to identify $\bar{n}$ & $\bar{\Lambda}$ (though not optimized)
  - Position & momentum measurements also compatible for $\bar{\Lambda}$ case

Reconstruction performance of $\bar{n}$ and $\bar{\Lambda}$. † indicates that the corresponding correlation value is not an average but is recalculated across all testing samples.

		↓ mAB (°)	↓ mAE (GeV/c)	↓ mRE (%)	↑ Corr.	↑ Acc. (%)
S.F.S.	$\bar{n}$	16.34	0.1546	28.17	0.5733	95.38
S.F.S.	$\bar{\Lambda}$	20.15	0.1421	36.93	0.5389	54.04
S.F.S.	avg.	18.24	0.1483	32.55	0.6390†	74.71
ViC	$\bar{n}$	10.16	0.1414	25.52	0.6365	93.14
ViC	$\bar{\Lambda}$	15.10	0.1285	33.60	0.5469	73.82
ViC	avg.	**12.63**	**0.1349**	**29.56**	**0.6785†**	**83.48**



ViC shows potential to develop an universal neutral hadron reconstruction algorithm!

Summary & outlook

- **Neutral hadron reconstruction is challenging at τ-charm facilities.**
  - Information recorded in detector is rare & sparse

- **Deep learning could be a key to fully exploit such information.**
  - A practical solution for neutron identification with Graph Neural Network
  - A baseline model for anti-neutron reconstruction with Visual Object Detector

- **Prospects in STCF**
  - EMC maintains BESIII spec but with faster time response
    - Allow identification & momentum measurement via time-of-flight method
    - 300 ps time resolution can offer 3σ n/γ separation & 8% momentum precision for a neutron @ $1\ \mathrm{GeV}/c$
  - MUD serves as auxiliary detector for neutral hadron

More data & more information is always welcomed for deep learning.

Thanks for your attention!