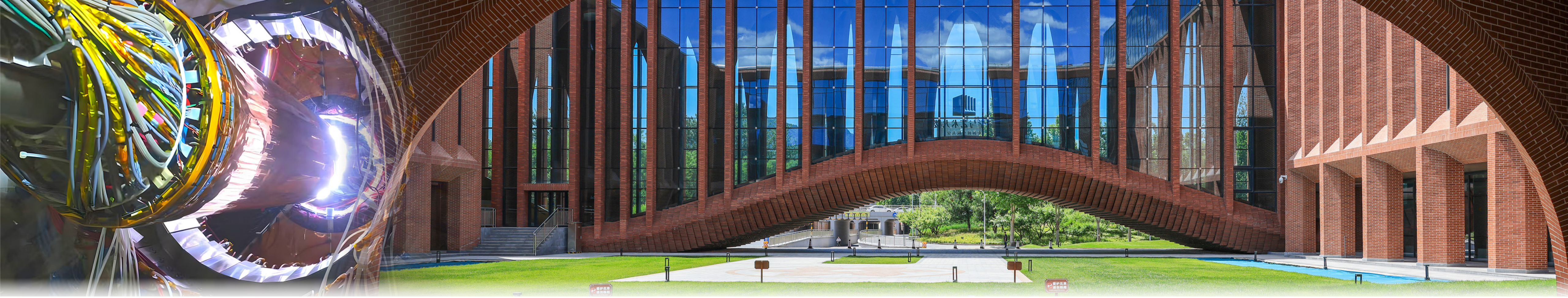


Vision / Language Calorimeter

深度学习驱动的电磁量能器上反中子重建

Yangu Li (李彦谷)
on behalf of the development team
University of Chinese Academy of Sciences

August 22, 2025

First of all...

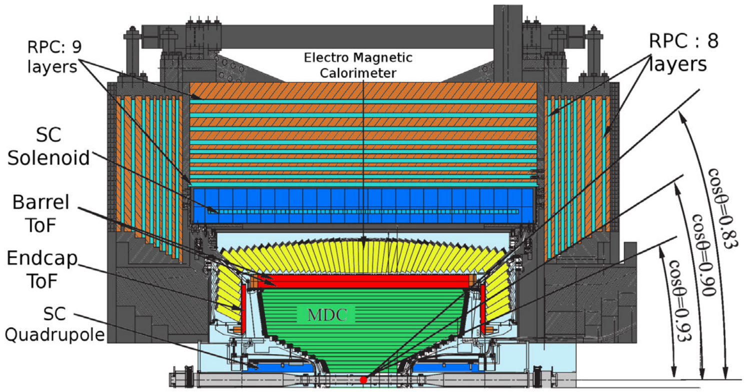
Why bother reconstruct hadrons in an electromagnetic calorimeter when there are dedicated hadronic calorimeters...?

The reason

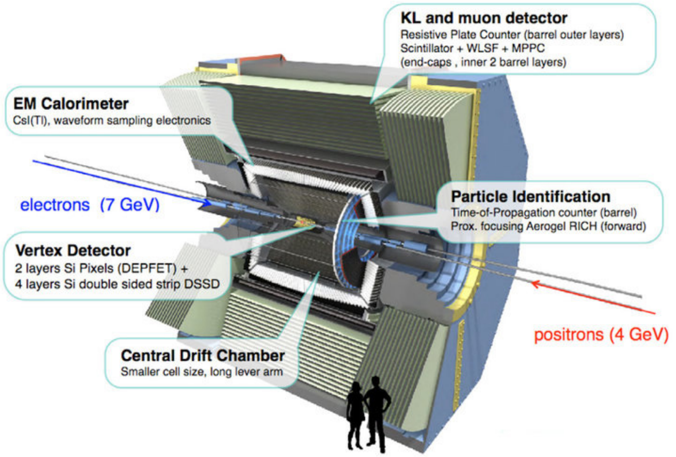
- **Long-lived neutral hadrons $(n,\ K_L^0)$ are important probes for colliders at τ-charm region.**
  - Participated in hyperon & charmed hadron decays, light hadron spectrum, exotic hadron states, etc.
  - e.g., about 1/3 of $\bar{\Lambda}_c^-$ decays contain $\bar{n}$, in which 20% are still unknown (PRD **108**, L031101)

- **However, most τ-charm facilities have no dedicated hadronic calorimeter.**
  - Detection rely on electromagnetic calorimeter (ECAL, EMC)

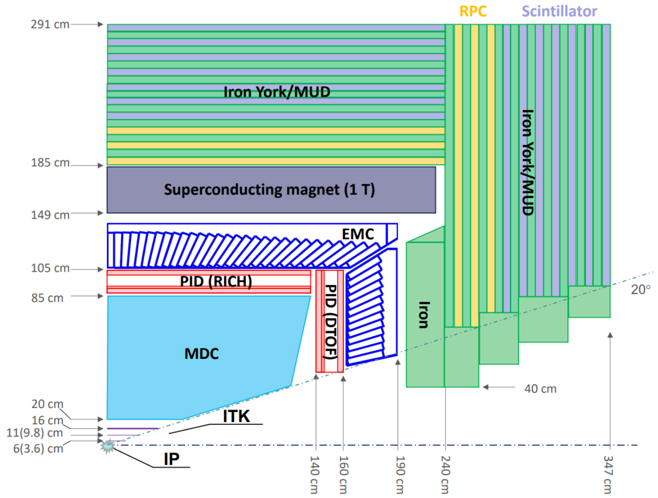
BESIII



Belle II

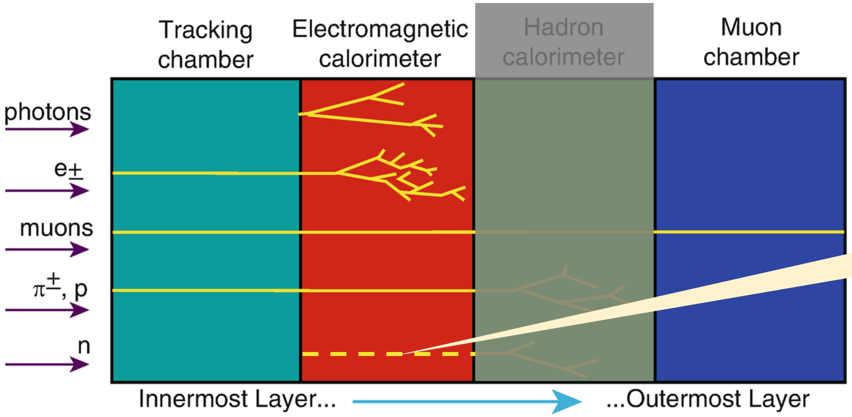


STCF 超级陶粲装置
Super Tau-Charm Facility



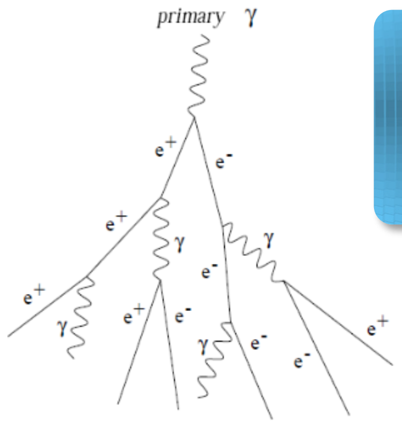
The issue

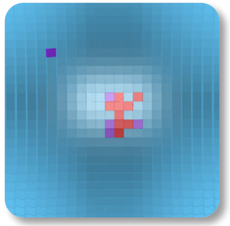
- **Direct reconstruction of neutral hadron in EMC is very challenging.**
  - EMC's size & material prevent full deposition of hadronic showers



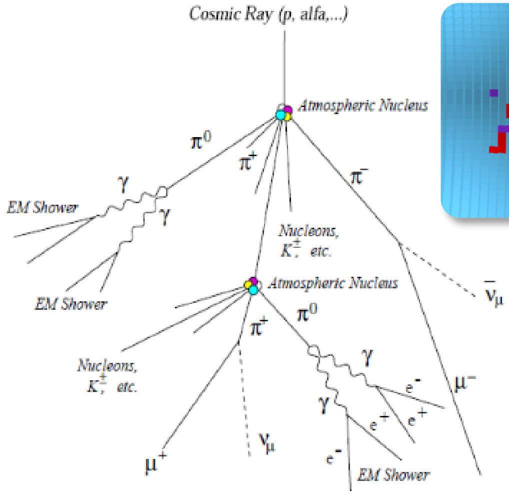
Only deposit a small amount of energy in it.

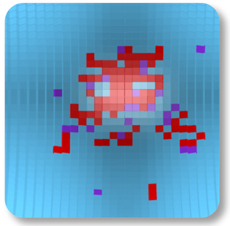
  - Momentum is unknown
    - Sizable energy leakage
  - Position is imprecise
    - Hit clusters are less centralized than photons
  - Identification is not perfect
    - Can be confused with photon / beam background / detector noise
  - Monte-Carlo simulation is imprecise
    - Up to ~10% discrepancy from data (NIMA **1033**, 166672)





EM shower (↑) vs. hadronic shower (↓)

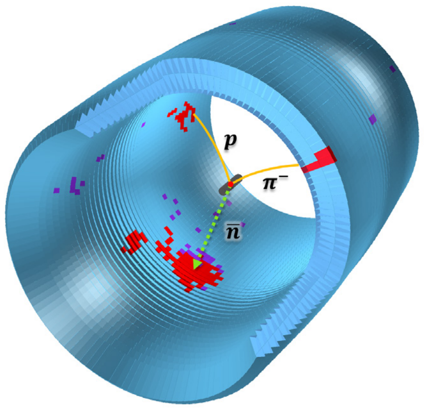




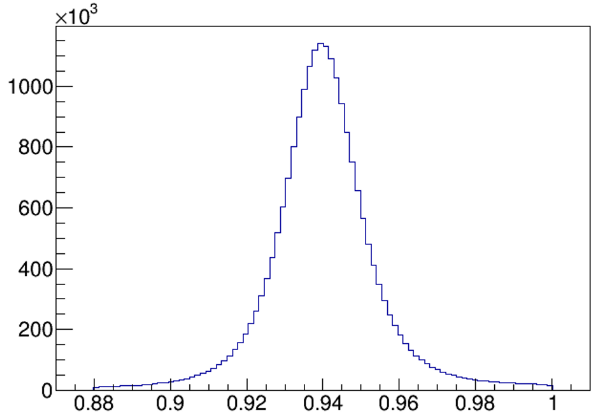
Conventional solutions

- **Common solution: recoil**
  - Calculate 4-momentum of neutral hadron using energy-momentum conservation
  - Only works when all other particles in the event are reconstructable
  - Can't deal with:
    - Semi-leptonic decays containing neutrinos e.g., $\Lambda_c^+ \to n e^+ \nu_e$
    - Radiative decays containing photons e.g., $\Lambda \to n\gamma$
    - Decays containing multiple neutral hadrons e.g., $e^+e^- \to n\bar{n}$

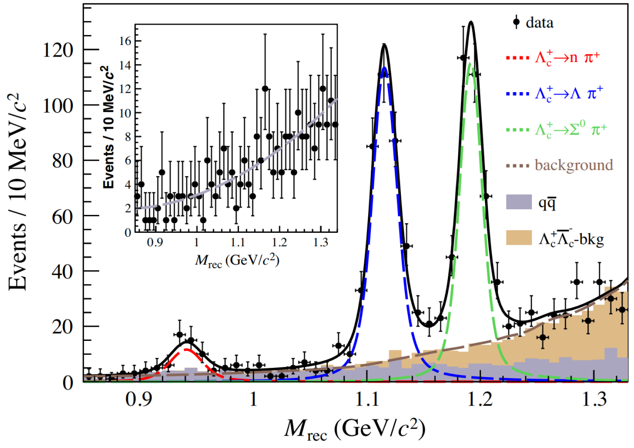
Event display of $J/\psi \to p\bar{n}\pi^-$



Invariant mass spectrum for recoil $\bar{n}$



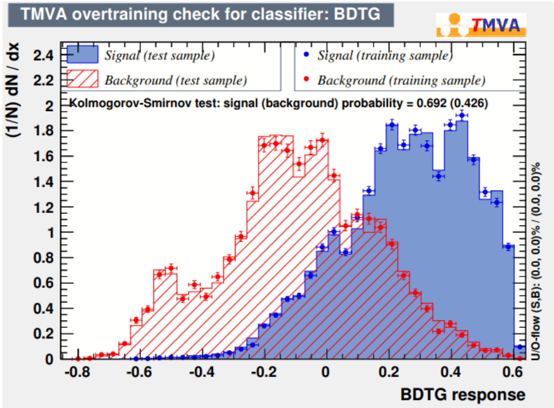
Study of $\Lambda_c^+ \to n\pi^+$ at BESIII
PRL **128**, 142001 (2022)



Conventional solutions

- **Specific solutions**
  - Require much effort, unable to generalize

Study of $\Lambda \to n\gamma$ at BESIII
PRL **129**, 212002 (2022)



Use BDT to identify EMC showers from neutron / photon

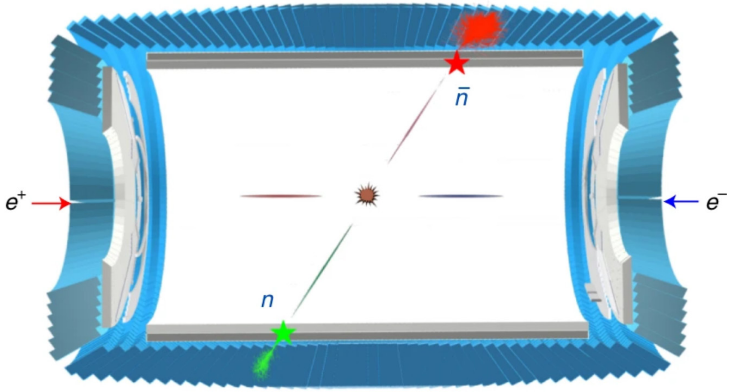
Study of $D^+ \to K_L^0 e^+ \nu_e$ at BESIII
PRD **92**, 112008 (2015)

$$U_{\rm miss} \equiv E_{\rm miss} - c|\vec{p}_{\rm miss}|, \equiv \mathbf{0}$$

$$E_{\rm miss} = E_{\rm tot} - E_{\rm tag} - E_{K_L^0} - E_e,$$
$$\vec{p}_{\rm miss} = \vec{p}_{\rm tot} - \vec{p}_{\rm tag} - \vec{p}_{K_L^0} - \vec{p}_e;$$

Fix K_L^0 position and solve kinematic equations

Study of $e^+e^- \to n\bar{n}$ at BESIII
Nature Phys. **17**, 1200-1204 (2021)



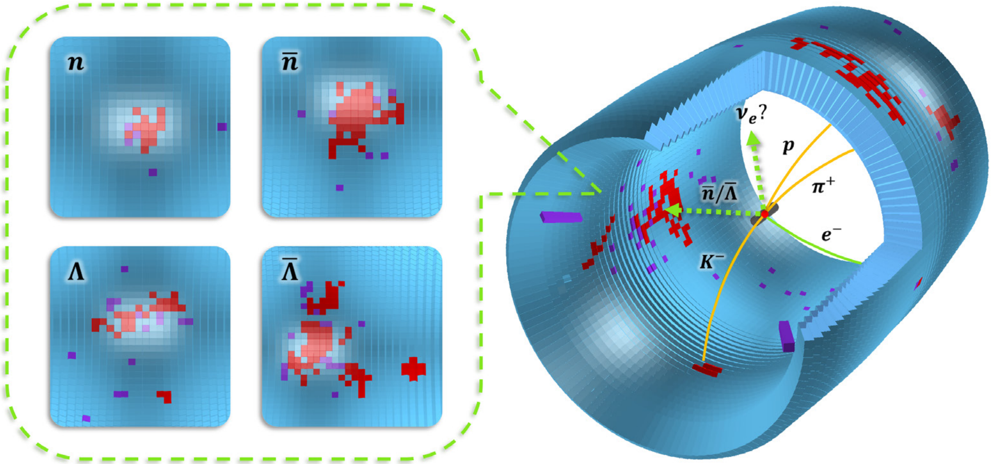
Use time-of-flight detector to calculate neutron momentum

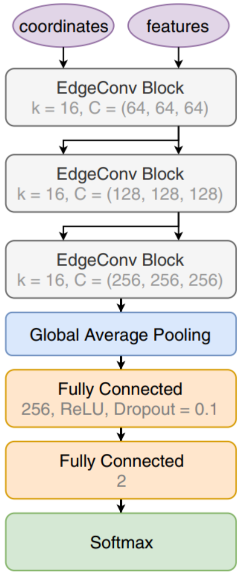
Can we use deep learning to recognize the hadronic shower pattern?

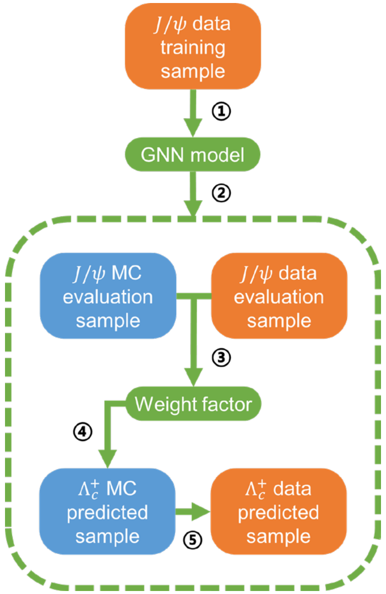
Our first attempt

Nature Commun. **16**, 681 (2025)

- **Neutron identification with Graph Neural Network (GNN)**
  - Aim to measure the sub-dominant Λ_c^+ semi-leptonic decay $\Lambda_c^+ \to \boldsymbol{n} e^+ \nu_e$
  - Decades of study history, dozens of theoretical predictions wait to be confirmed
  - Need to suppress major background $\Lambda_c^+ \to \boldsymbol{\Lambda(n\pi^0)} e^+ \nu_e$ efficiently





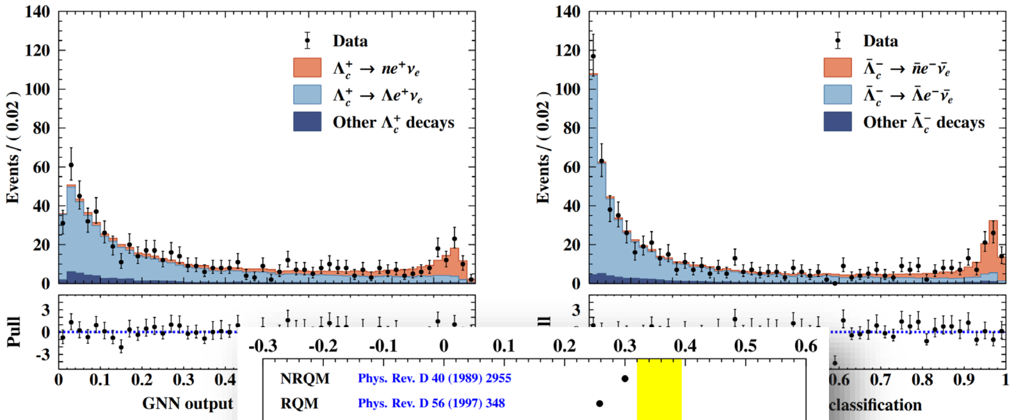


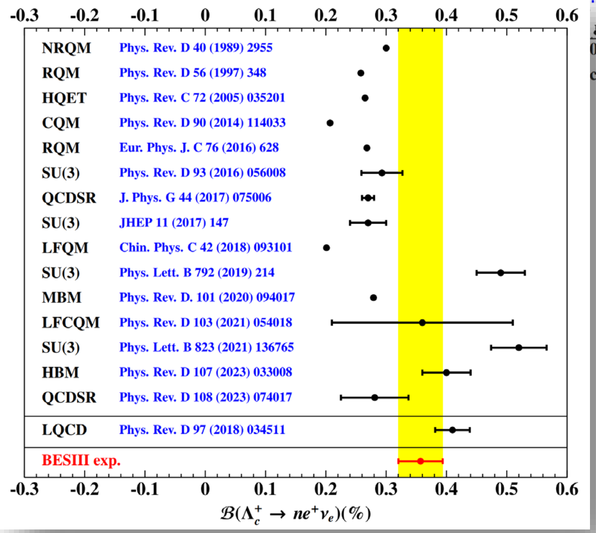
- **Train a GNN-based n/Λ classifier**
  - Organize EMC showers as a point cloud
  - Use data-driven methods for GNN calibration, physics results validation & systematic uncertainty quantification

Our first attempt

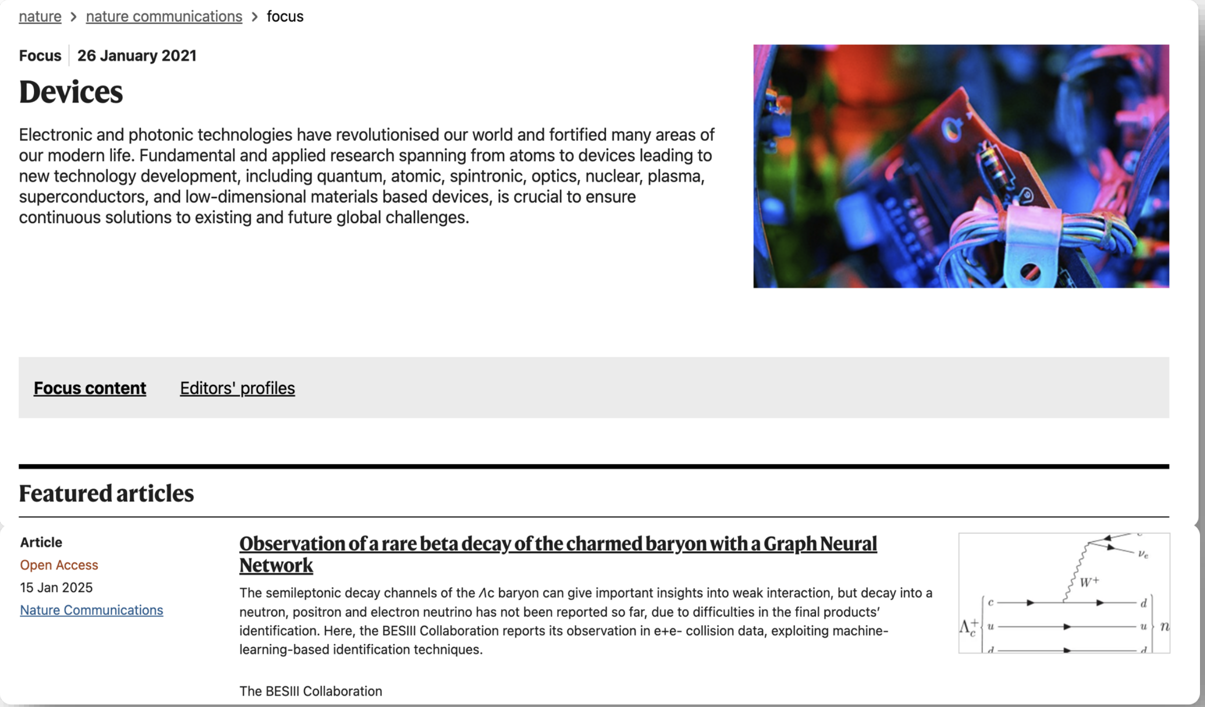
Nature Commun. **16**, 681 (2025)

- **Achieve first observation of $\Lambda_c^+ \to n e^+ \nu_e$**
  - Significance improved from $< 3\sigma$ to $> 10\sigma$
  - Precision capable to examine theoretical models





Selected as Editors' Highlights



Moving forward

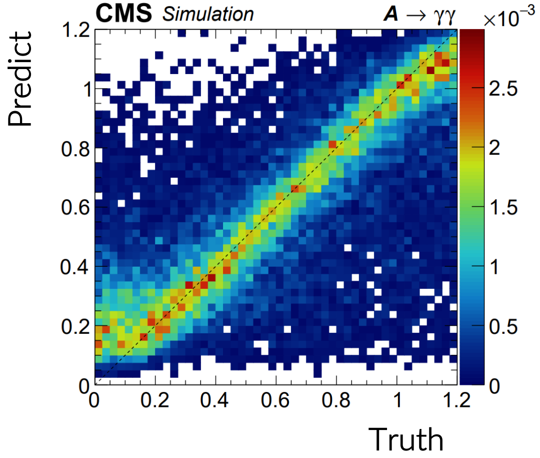
- **Full neutron reconstruction beyond identification is desired.**
  - More meaningful physics results (e.g., form factors) require knowing the neutron momentum

- **We tried predicting neutron momentum with GNN, but failed**
  - A regression task with much higher difficulty

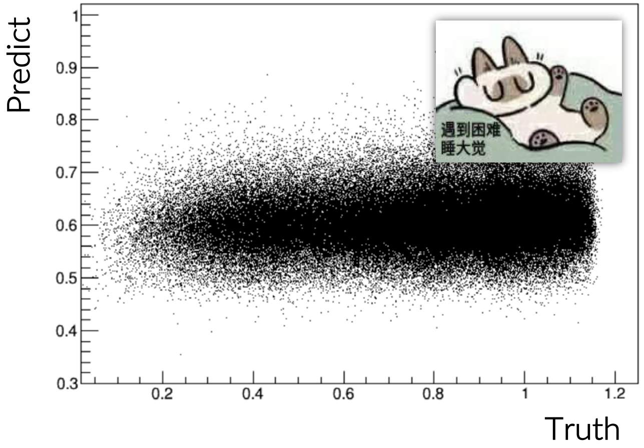
Does the limitation come from detector, or our deep learning technique?

We should seek help from computer scientists.

Ideal effect:
Positive correlated mapping

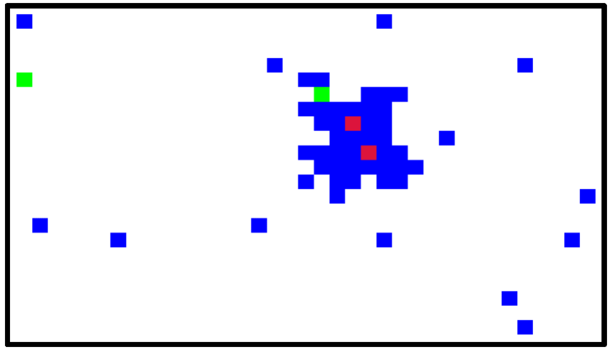


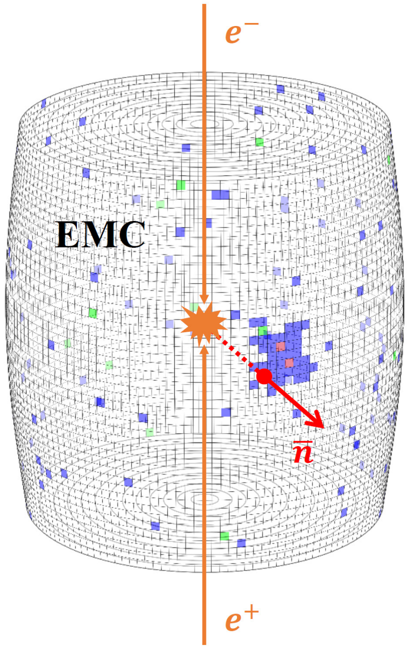
Our result:
Lazily assign all outputs near average to yield local minimum of loss



Methodology

- **Two representations of EMC hit map**





$E = 5.37\text{e} - 4\ \text{GeV}, \phi = -51.01°, \theta = 20.93°$

$E = 2.11\text{e} - 2\ \text{GeV}, \phi = -13.53°, \theta = 88.45°$

$E = 9.85\text{e} - 1\ \text{GeV}, \phi = 7.55°, \theta = 100.81°$

⋮

$E = 1.18\text{e} - 3\ \text{GeV}, \phi = 67.57°, \theta = 35.24°$

$E = 7.21\text{e} - 4\ \text{GeV}, \phi = -1.59°, \theta = 144.76°$

As image:
Vision Calorimeter
(ViC)
arXiv: 2408.10599

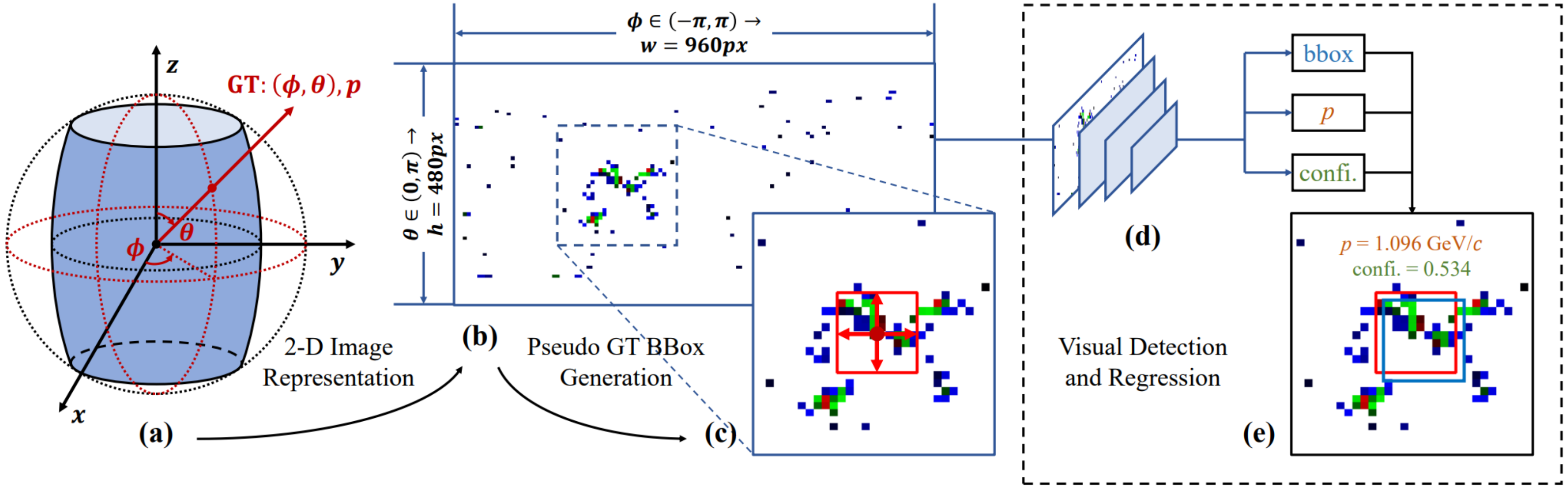
As sequence:
Language Modeling Calorimeter
(LMC)
paper in submission

Vision Calorimeter

Vision Calorimeter

- **Take advantage of object detection approach in computer vision.**
  - Represent EMC hits on a 2D image
  - Find the position of $\bar{n}$ within a binding box
  - Predict its confidence score, class and momentum as downstream tasks

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



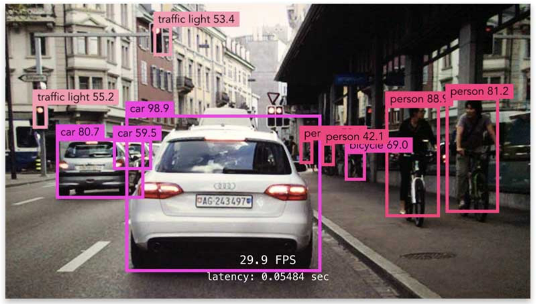
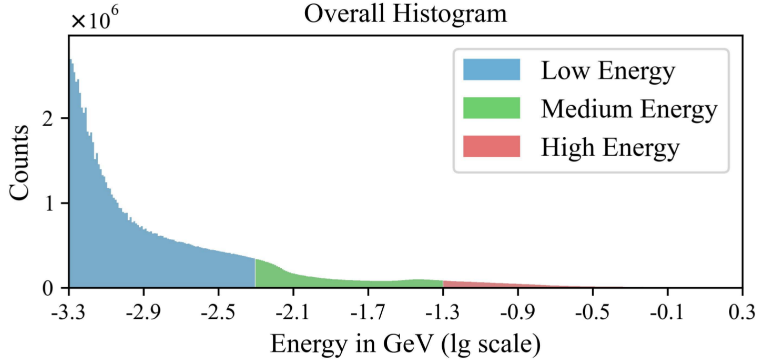


Image quantification

- **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	

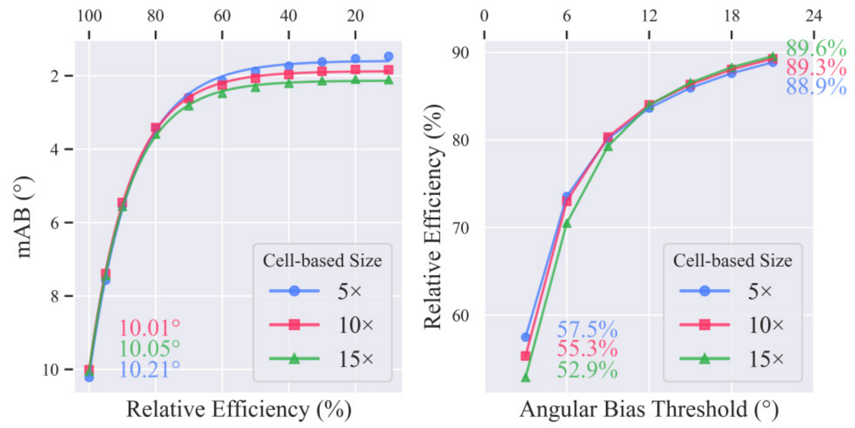
- **Colors**
  - EMC deposited energy range is 0.5 MeV ~ 2 GeV
  - Take log scale: $[10^{-3.3}, 10^{0.3}]$
  - Divide low, medium and high measures to fill blue, green and red channels
  - Add a -30db Gaussian noise to address the sparsity of EMC hits



Bounding box

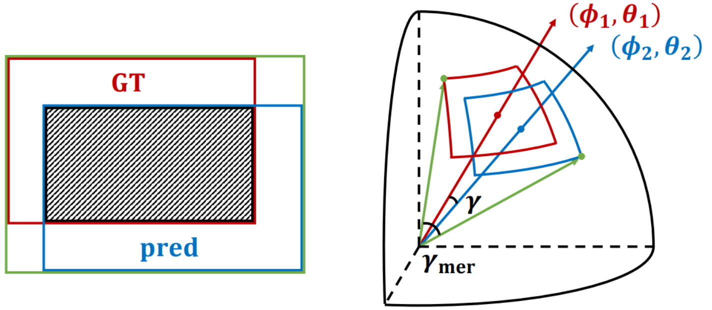
- **Why not use point-wise position regression?**
  - Bounding box (BBox) prediction can better exploit contextual information
  - Superior performance in experiments
  - Need to generate pseudo BBox around $\bar{n}$ incident position

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



- **Definition of loss**
  - Standard loss in object detection is IOU
    - $IOU = S(GT \cap Pred)/S(GT \cup Pred)$
  - We design a more center-oriented version

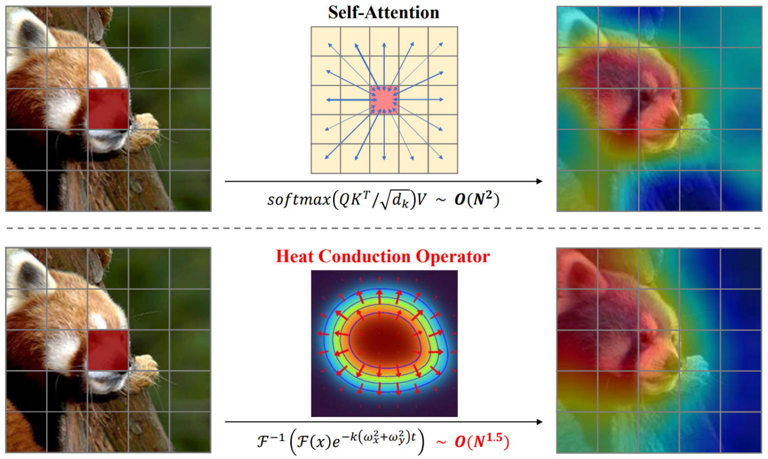
$$\mathcal{L}_{CO} = 1 - IoU + \alpha \cdot \frac{(\cos\gamma - 1)^2}{(\cos\gamma_{\text{mer}} - 1)^2}$$

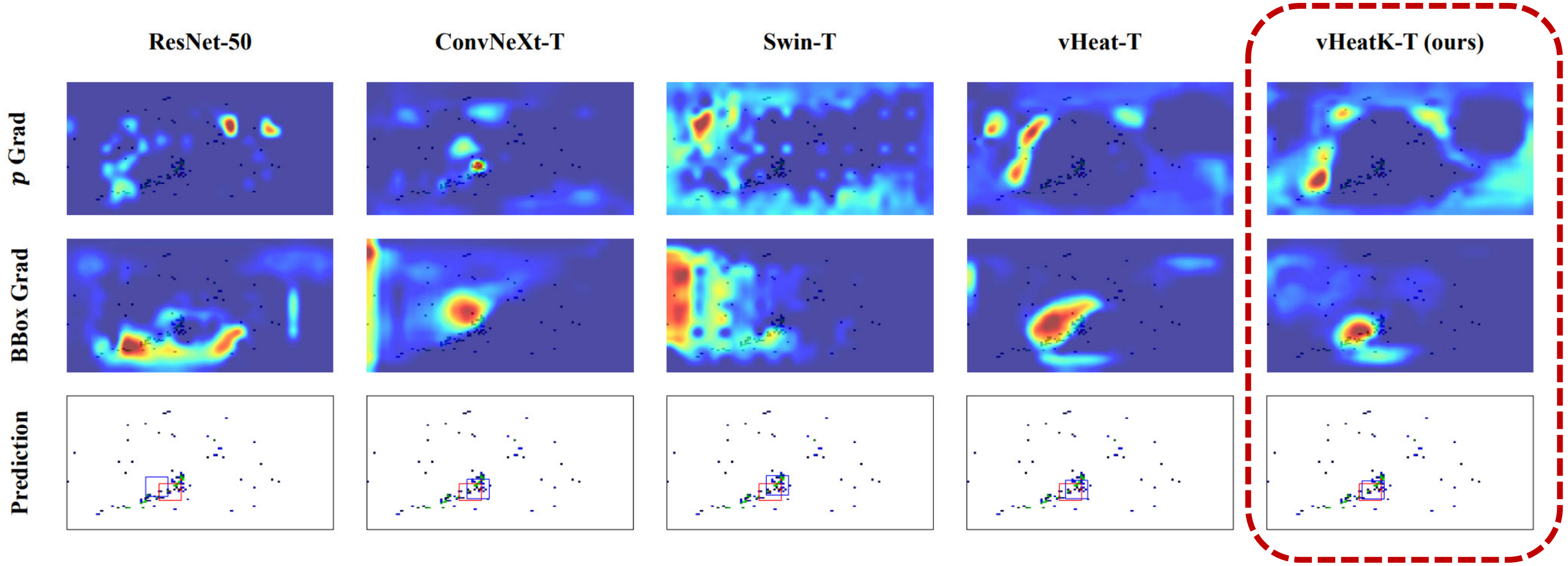


(a) (b)

Network architecture

- **Backbone: Swin-Transformer → vHeat**
  - A computer vision model inspired by physics law
  - Use heat conduction operator to propagate visual information
    - Enjoy global receptive fields and $\mathcal{O}(N^{1.5})$ complexity
    - Analog to hadronic shower production

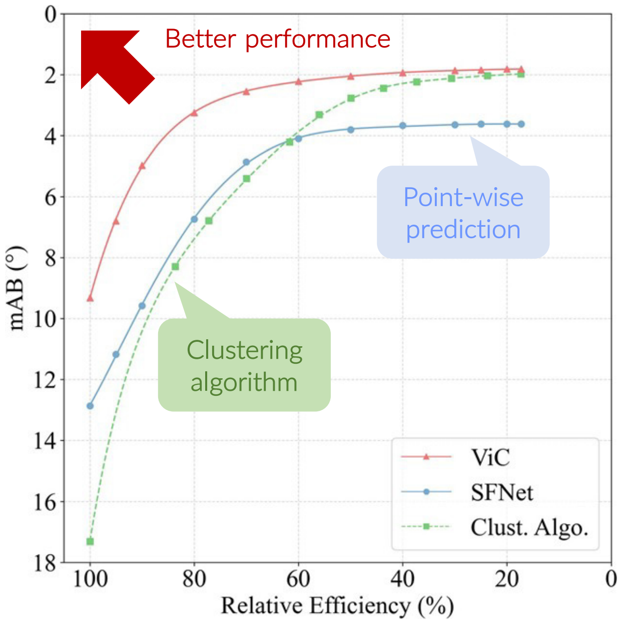


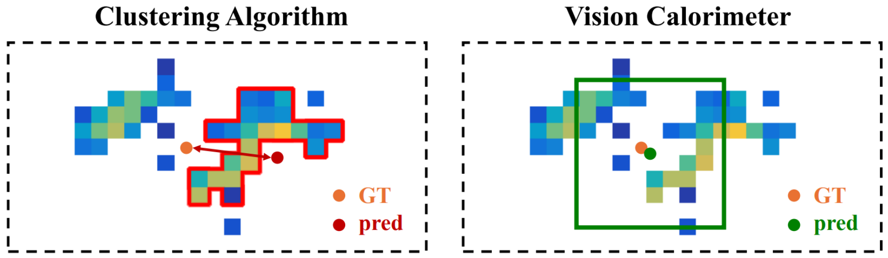


Combine global attention in momentum prediction & local attention in position prediction

Performance of ViC

- **Dataset**
  - 1 million anti-neutron images taken from BESIII data
- **In position measurement**
  - Compared with conventional clustering algorithm, ViC improves the precision by 80% at full efficiency ($17.4^\circ \rightarrow 9.3^\circ$)
  - This precision can be doubled at 90% efficiency
  - Upper limit is EMC cell granularity
- **How comes the improvement?**
  - Conventional clustering algorithm may split a discontinues hadronic shower
    - Usually caused by multiple scattering
    - Only the most energetic one is considered
  - ViC can better handle such scenarios



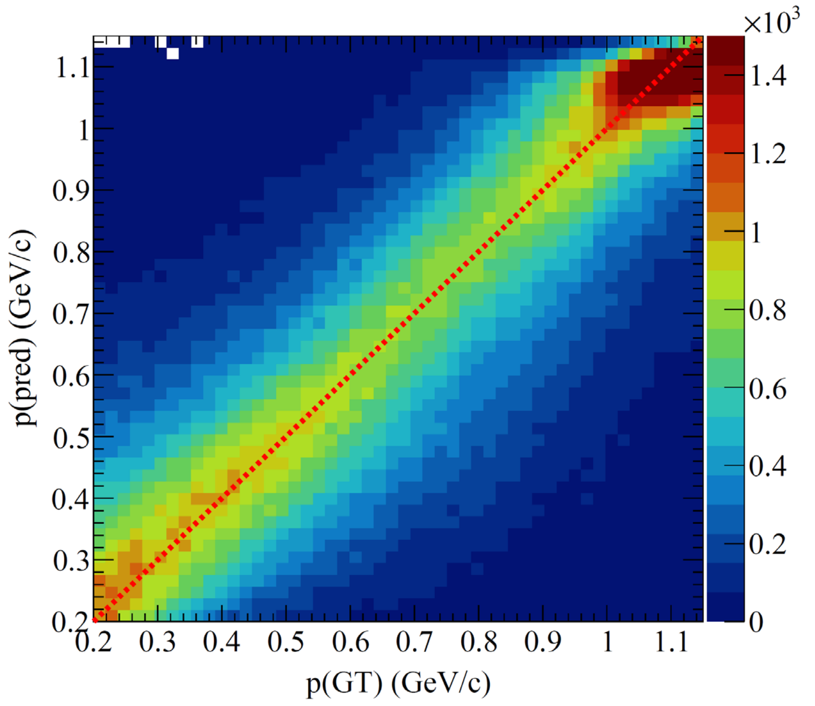


$\phi_{\bar{n}} = \sum_i E_i\phi_i / \sum_i E_i, \theta_{\bar{n}} = \sum_i E_i\theta_i / \sum_i E_i,$
$p_{\bar{n}}$: N/A

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

Performance of ViC

- **In momentum measurement**
  - No conventional solution at all
  - ViC firstly realize such capability in EMC
  - Resolution ~15% @ 1 GeV, 30% @ 500 MeV
    - Even better than dedicated HCALs in sub-GeV region ($\sim 80\%/\sqrt{E}$)

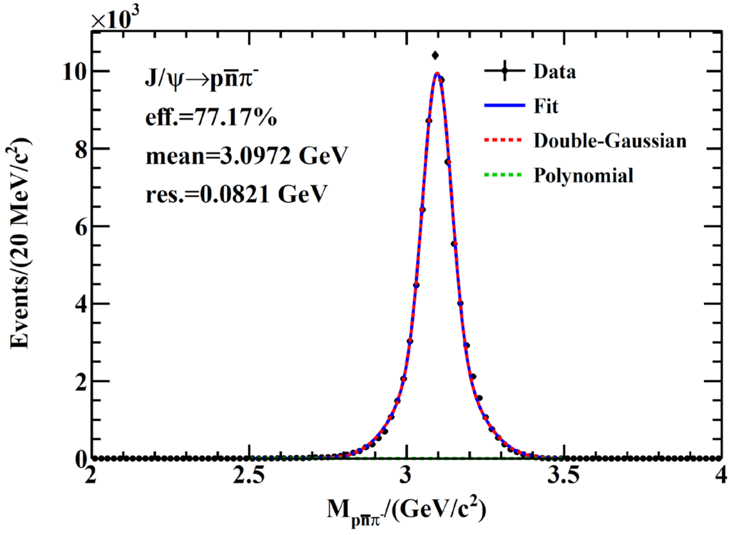


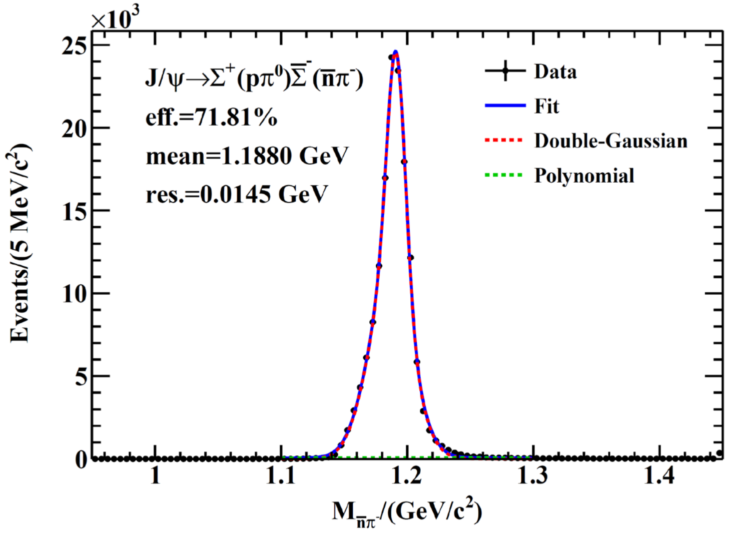
- **In classification**
  - ViC is capable to identify $\bar{n}$ & $\bar{\Lambda}$ (though not optimized)
  - Position & momentum measurements also compatible for $\bar{\Lambda}$ case

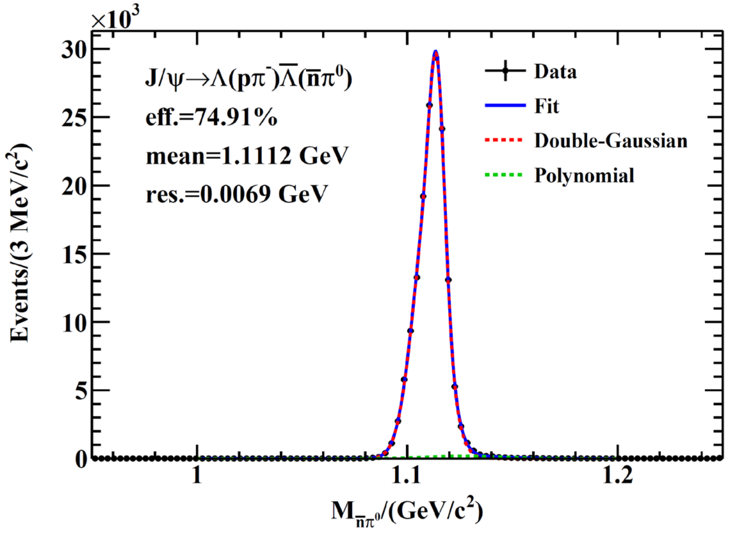
		↓ mAB (°)	↓ mAE (GeV/c)	↓ mRE (%)	↑ Corr.	↑ Acc. (%)
S.F.S.	$\bar{n}$	16.34	0.1546	28.17	0.5733	95.38
	$\bar{\Lambda}$	20.15	0.1421	36.93	0.5389	54.04
	avg.	18.24	0.1483	32.55	0.6390†	74.71
ViC	$\bar{n}$	10.16	0.1414	25.52	0.6365	93.14
	$\bar{\Lambda}$	15.10	0.1285	33.60	0.5469	73.82
	avg.	**12.63**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

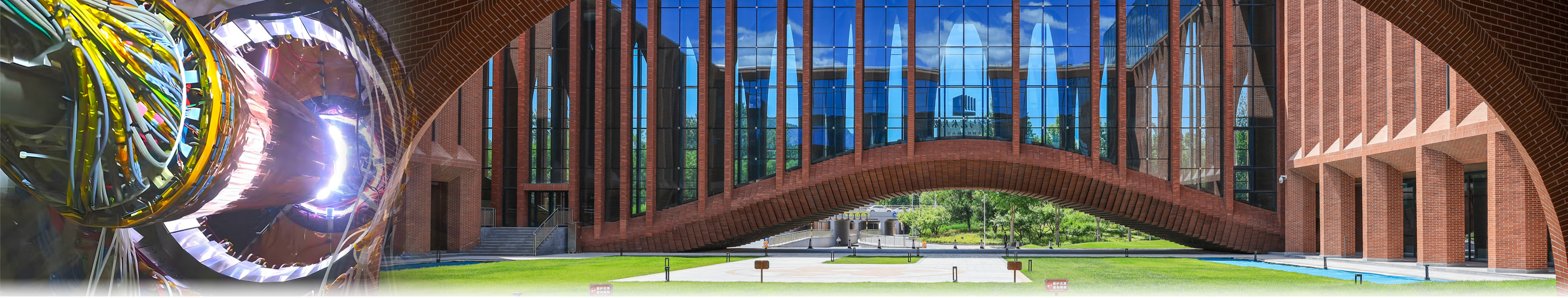
Performance of ViC

- **In physics measurements**
  - Test on toy analysis tasks
    - Reconstruct J/ψ invariant mass in $J/\psi \to p\bar{\boldsymbol{n}}\pi^-$
    - Reconstruct $\bar{\Sigma}^-$ invariant mass in $J/\psi \to \Sigma^+(p\pi^0)\bar{\Sigma}^-(\bar{\boldsymbol{n}}\pi^-)$
    - Reconstruct $\bar{\Lambda}$ invariant mass in $J/\psi \to \Lambda(p\pi^-)\bar{\Lambda}(\bar{\boldsymbol{n}}\pi^0)$
  - Obtain unbiased & well-resolved resonance mass peaks
    - Ability of generalization in different decay scenarios
    - Potential of application in real physics analyses







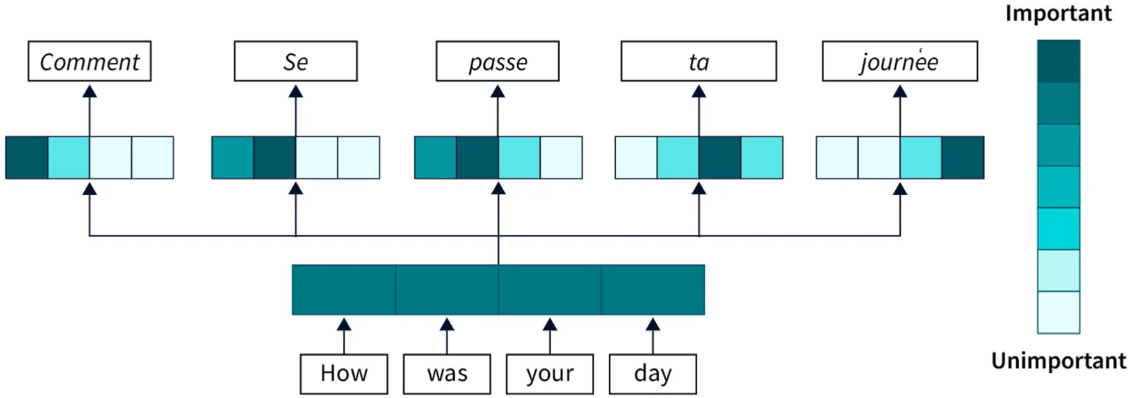


Language Modeling Calorimeter

Language Modeling Calorimeter

- **Take advantage of powerful Transformer architecture**
  - The fundamental operator of Large Language Models

- **But attention is NOT all we need**
  - EMC hit map is very unlike nature language texts
  - How to construct the token sequence properly?
  - How to design the pre-training task efficiently?



$E = 5.37\text{e}-4\ \text{GeV}, \phi = -51.01°, \theta = 20.93°$

$E = 2.11\text{e}-2\ \text{GeV}, \phi = -13.53°, \theta = 88.45°$

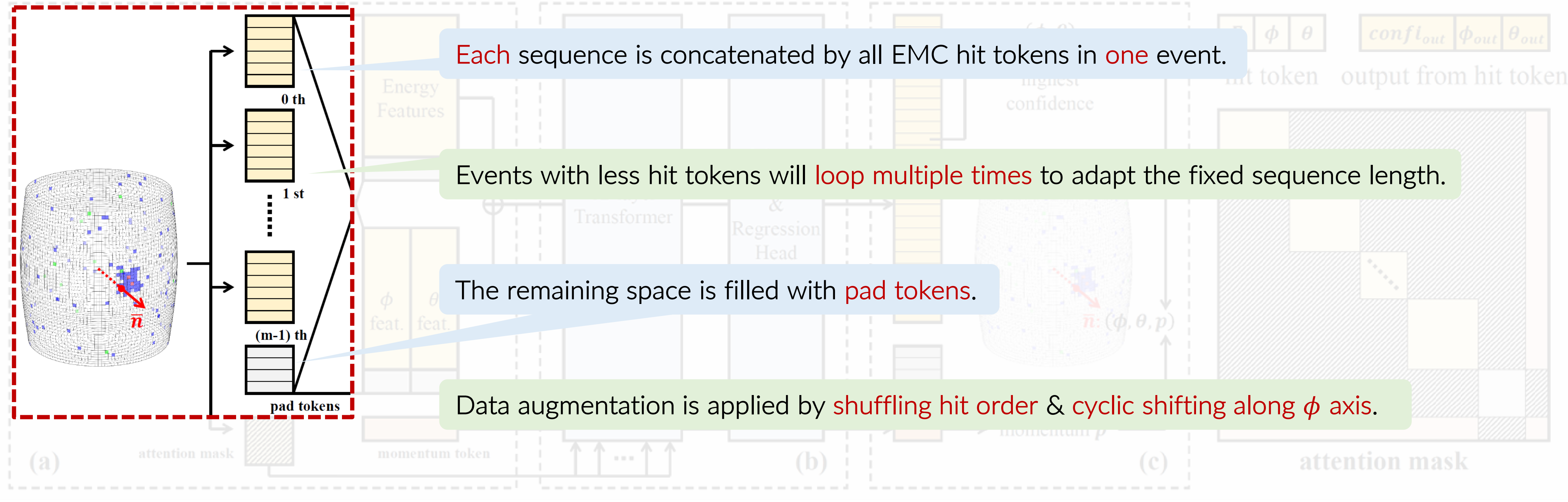
$E = 9.85\text{e}-1\ \text{GeV}, \phi = 7.55°, \theta = 100.81°$

⋮

$E = 1.18\text{e}-3\ \text{GeV}, \phi = 67.57°, \theta = 35.24°$

$E = 7.21\text{e}-4\ \text{GeV}, \phi = -1.59°, \theta = 144.76°$

Architecture of LMC



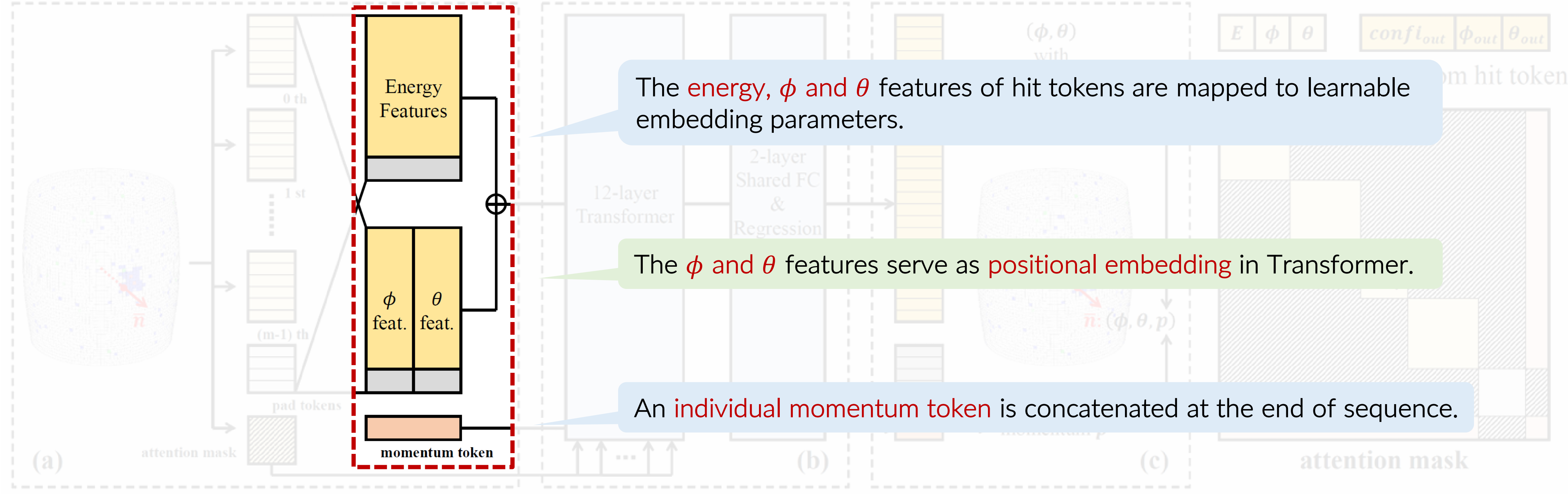
Each sequence is concatenated by all EMC hit tokens in one event.

Events with less hit tokens will loop multiple times to adapt the fixed sequence length.

The remaining space is filled with pad tokens.

Data augmentation is applied by shuffling hit order & cyclic shifting along ϕ axis.

Architecture of LMC



The energy, ϕ and θ features of hit tokens are mapped to learnable embedding parameters.

The ϕ and θ features serve as positional embedding in Transformer.

An individual momentum token is concatenated at the end of sequence.

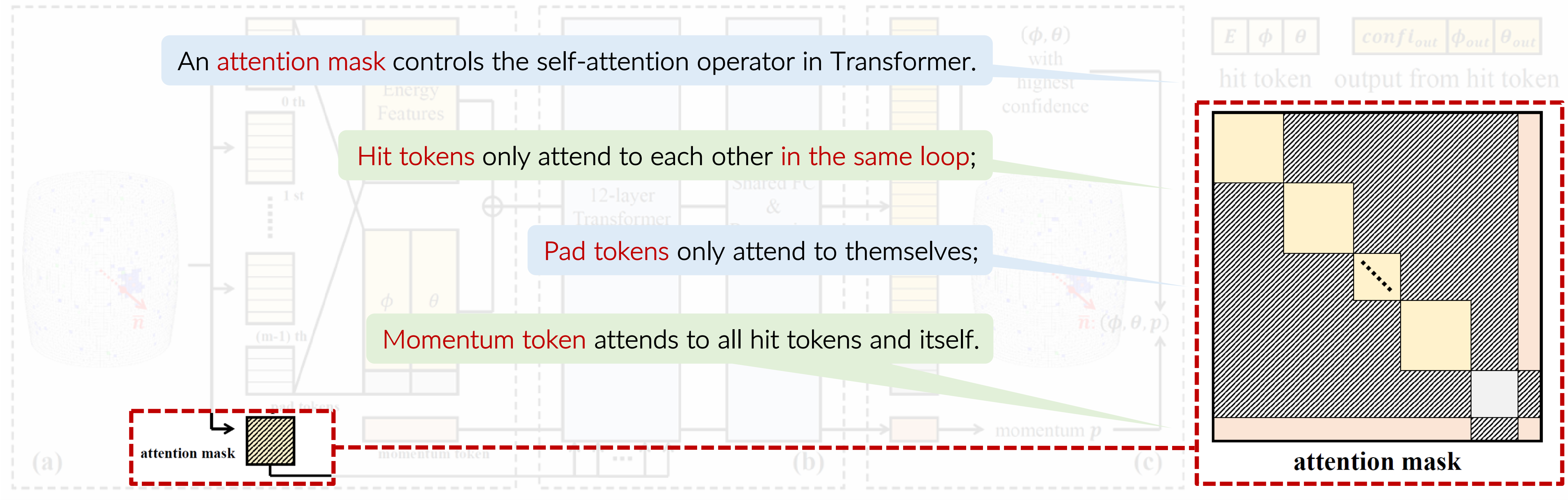
Architecture of LMC

An attention mask controls the self-attention operator in Transformer.

Hit tokens only attend to each other in the same loop;

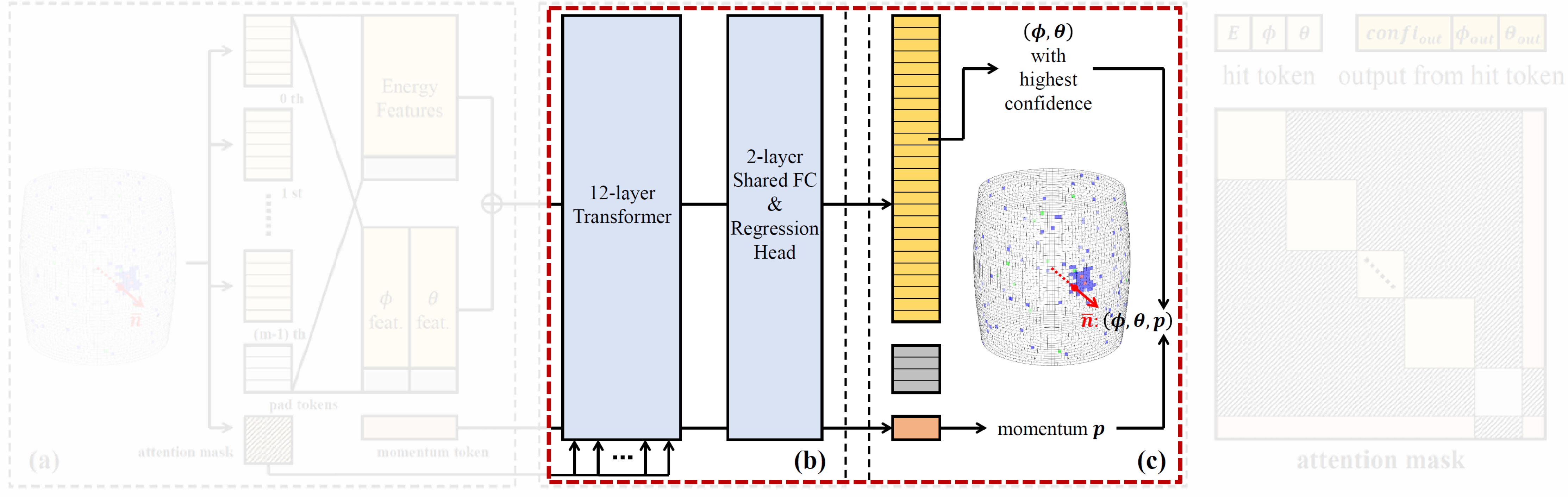
Pad tokens only attend to themselves;

Momentum token attends to all hit tokens and itself.



Architecture of LMC

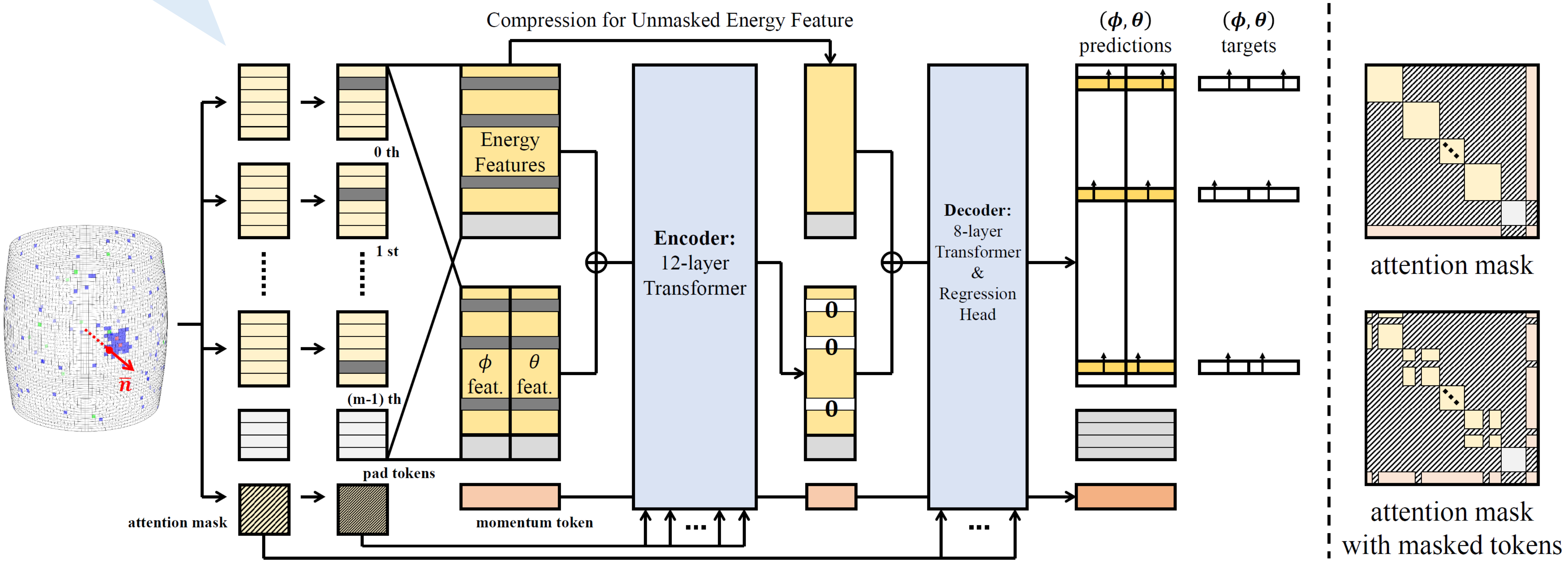
Transformer layers perform feature extraction,
Detection head predicts position and momentum of anti-neutron.



Architecture of LMC

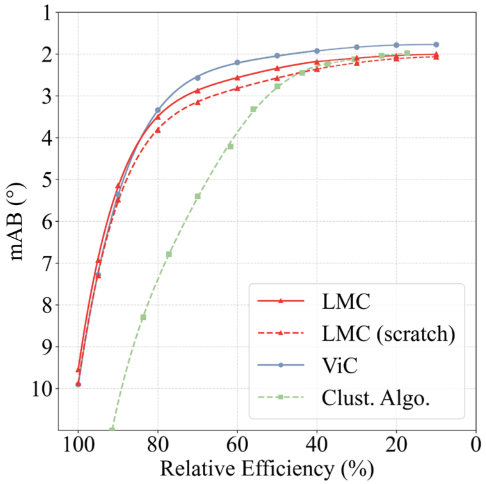
For pre-training task,
we mask out some high-energy hit tokens.

Model is required to regress the (ϕ, θ) of masked hit tokens according to their energy prompts.

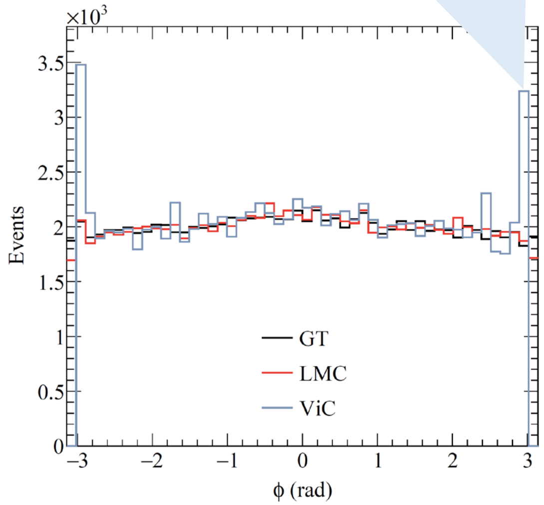


Performance of LMC

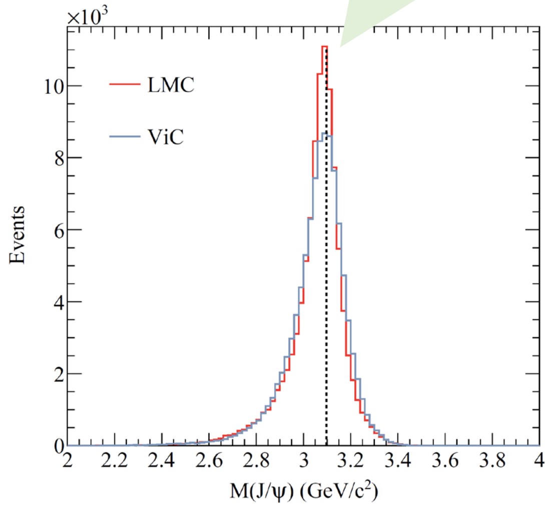
- **Overall comparable with ViC**
  - Superior in some metrices



Edge effect is eliminated due to data augmentation



Better resolution in resonance mass peak



- **Larger potential of improvement**
  - Room for optimizations via better pre-training
  - Able to combine information from multiple detectors
    - e.g., timing information from Time-of-flight detector

Summary

- **Neutral hadron reconstruction remains a great challenge at τ-charm facilities.**
  - Rare & sparse energy deposit in non-dedicated detector

- **We propose two deep learning models to reconstruct anti-neutron in an electromagnetic calorimeter.**
  - Vision Calorimeter (ViC) based on visual object detector
  - Language Modeling Calorimeter (LMC) based on pre-trained Transformer

- **The models show promising performances for application.**
  - Outperform conventional method by far in position measurement
  - Firstly realize ability of momentum measurement
  - Capable to run in real physics analysis scenarios

Thanks for your attention!