Transformer-Based STCF DTOF Particle Identification

Banghui Li, Teng Li, Xingtao Huang

Shandong University

2026.8.19

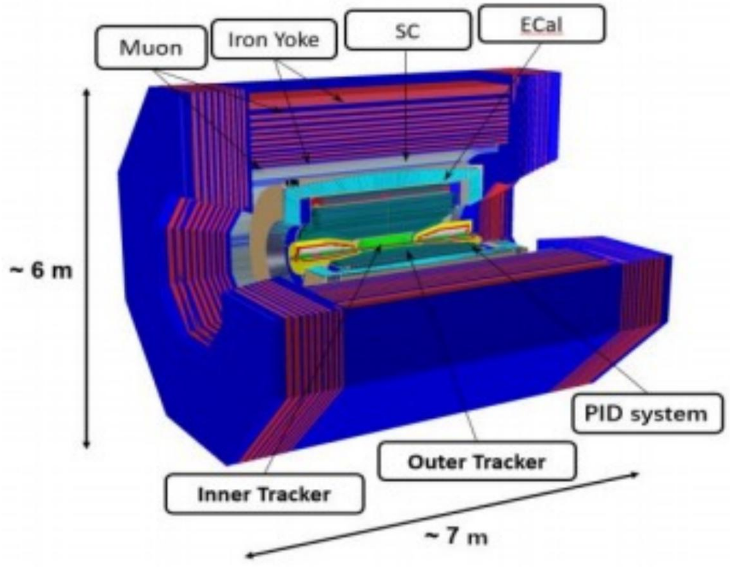
Outline

- Super Tau Charm Facility
- Single-Particle PID
- Multi-Particle PID
- Summary

Super Tau Charm Facility

The Super Tau Charm Facility (STCF) proposed in China is a new-generation facility of electron positron collider

- the peak luminosity above 0.5×10^{35} $cm^{-2}s^{-1}$
- center-of-mass energies covering 2-7 GeV
- potential for further upgrading to improve the peak luminosity and realize beam polarization in the future



Broad Physics at tau-charm Energy Region

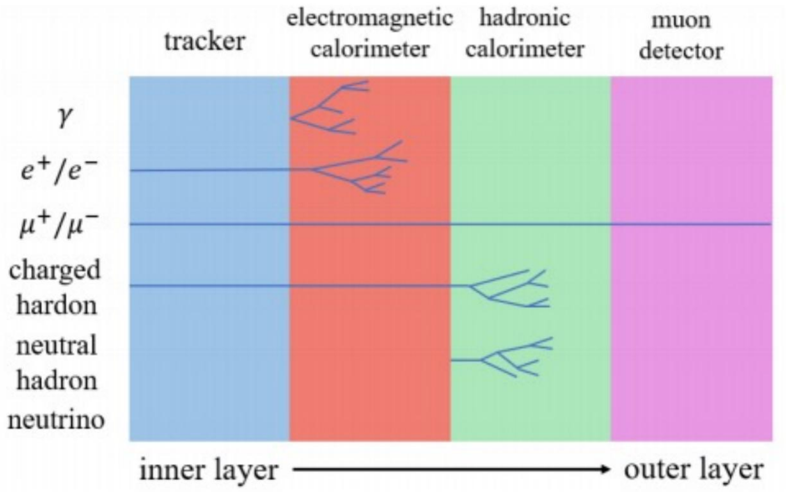
- Rich of physics with c quark and τ leptons
- Important playground for study of QCD, hadron physics
- Search for new physics beyond the Standard Model

From the interaction point outward, the STCF detector consists of a tracking system(**ITK** and **MDC**), a particle identification (**PID**) system, an electromagnetic calorimeter (**EMC**), a superconducting solenoid (**SCS**) and a muon detector (**MUD**).

Particle Identification System

The PID is one of the most fundamental tools in various physics studies. The PID for the full momentum range is essential for charm physics studies and fragmentation function studies.

- The identification of hadrons in the low momentum range is achieved through measurements ofthe specific energy loss rate (**dE/dx**) by the **MDC**.
- The identification of leptons and neutral particles is provided by the **EMC** and the **MUD**.
- To enhance PID and charged hadrons in the high momentum range, the PID system of the STCF is designed.



The PID system uses two different Cherenkov detector technologies:

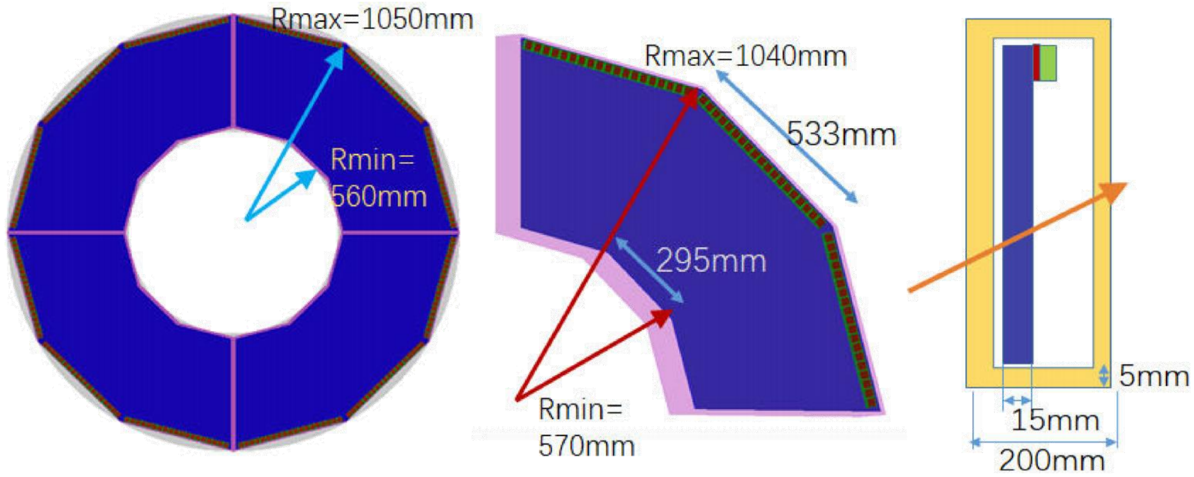
- a Ringing Imaging Cherenkov detector (**RICH**) in the barrel
- a time-of-flight detector based on the detection of the internal total-reflected Cherenkov light (**DTOF**) in the endcap

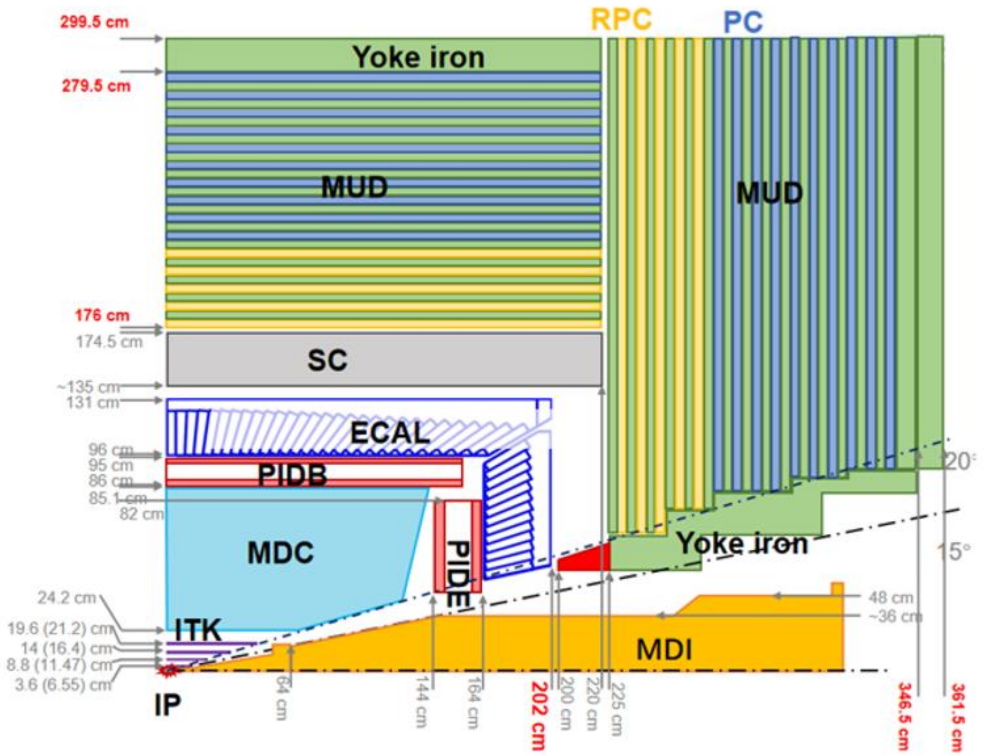
For p < 2 GeV/c, when the π/K misidentification rate is < 2%, the corresponding identification efficiency should be > 97%.

DTOF

The DTOF consists of two identical endcap discs positioned at $\sim \pm 1400$ mm away from the collision point along the beam direction. Each disc is made up of several quadrantal sectors, with an inner radius of $\sim$ 560 mm and an outer radius of$\sim$ 1050 mm.

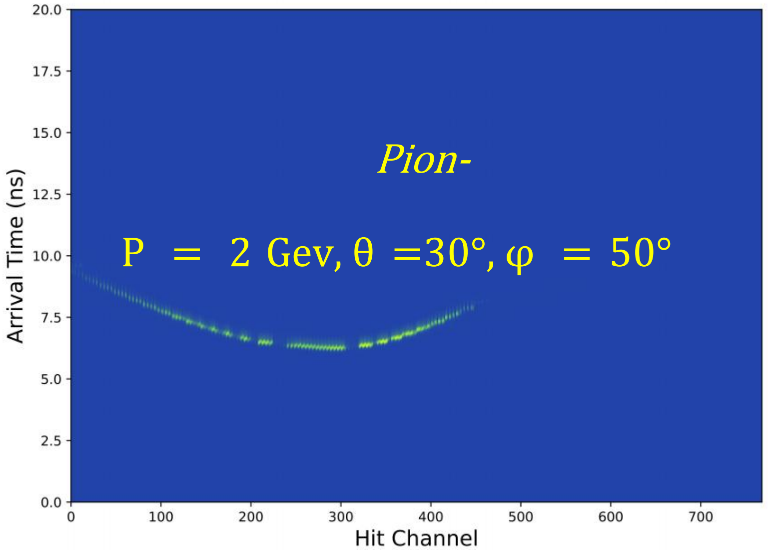
- covering in polar angles of $\sim 22^\circ - 36^\circ$
- synthetic fused silica radiator
- Photoelectric Detection: Multi-Anode PMT





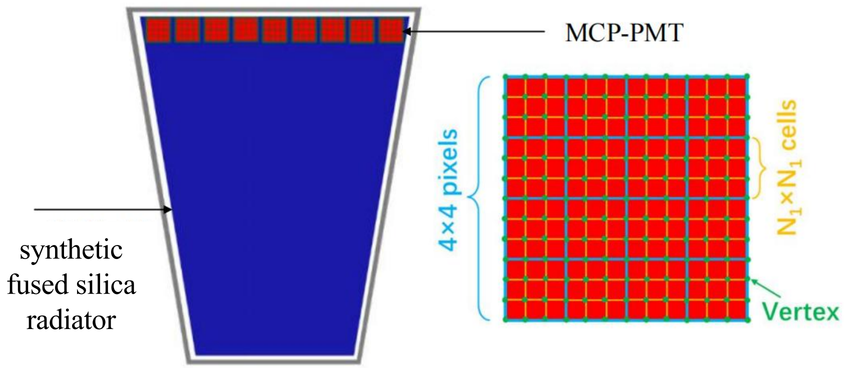
Schematic layout of the STCF detector concept

Convolutional Neural Network



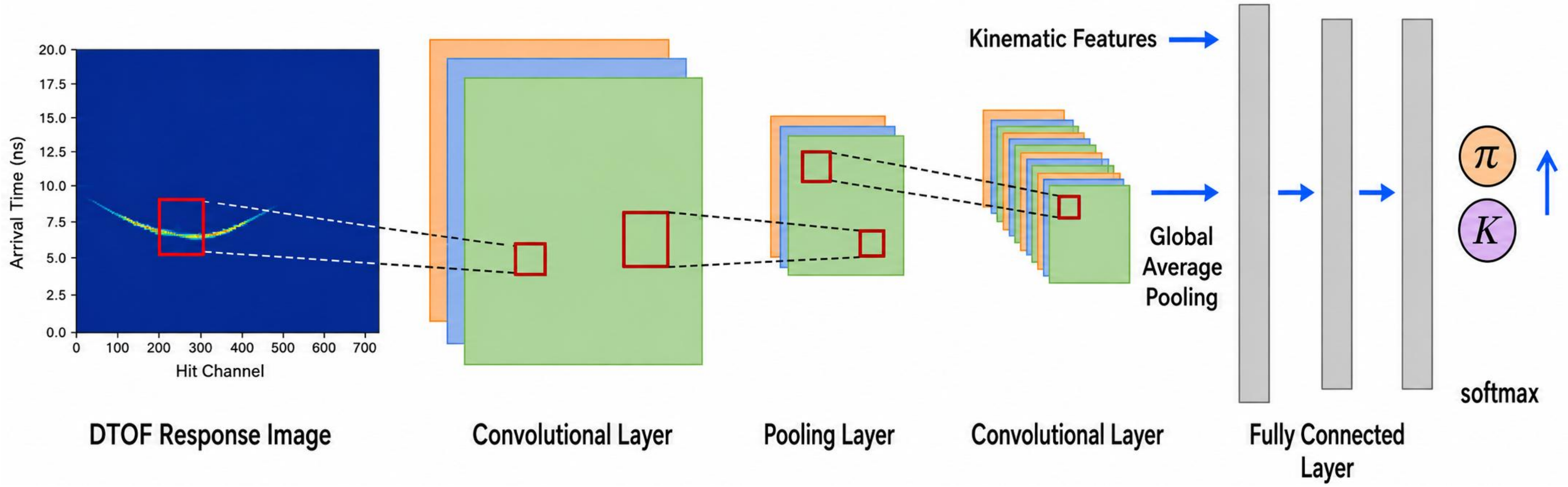
Using the raw response of the DTOF detector, a two-dimensional pixel map is constructed:

- X-label: the hit position of Cherenkov collected photon by PMT (spatial coordinates X and Y are merged into 768 one-dimensional channels)
- Y-label: the arrival time of Cherenkov collected photon by PMT
- Value: the number of photons within this bin



- 0 ≤ 4*Channel_x + Channel_y ≤ 768 (191*4+4)
- 5 ≤ Time ≤ 20 ns (Time resolution： 50 ps)
- Bin number: Channel_bin * Time_bin = 192 * 300

The Structure of CNN



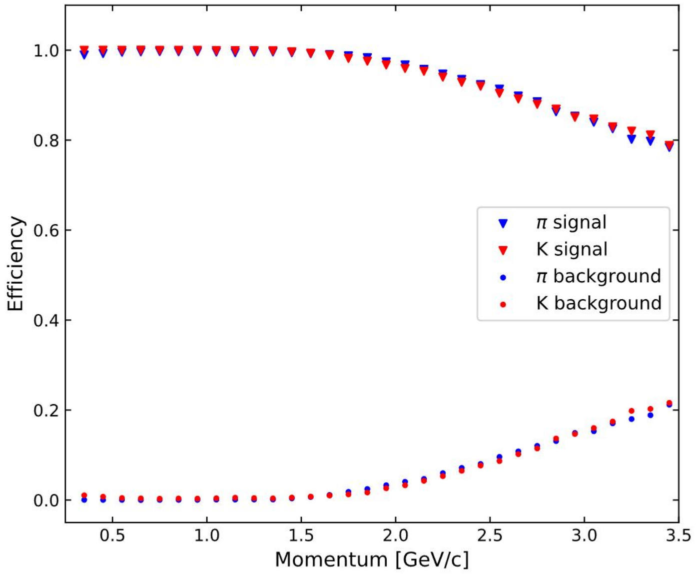
Single-particle π/K MC samples are simulated with OSCAR (2.7.0) offline software

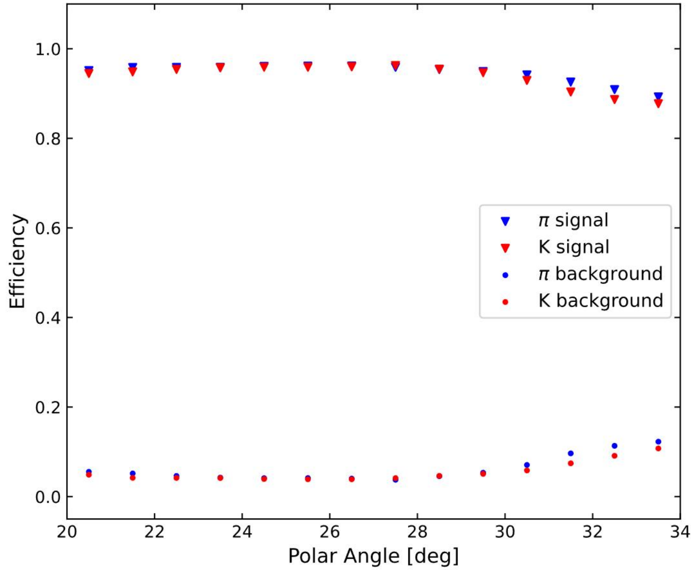
- Momentum: 0.3–3.5 GeV/c
- Polar angle: 20°–39°
- Events per particle type: $1 * 10^7$

- Select single-particle data hitting one DTOF sector
- Inputs: 2D image + 3D hit position & momentum (obtained by extrapolation from the tracking detector)
- EfficientNetV2-S is used as the baseline model and further optimized

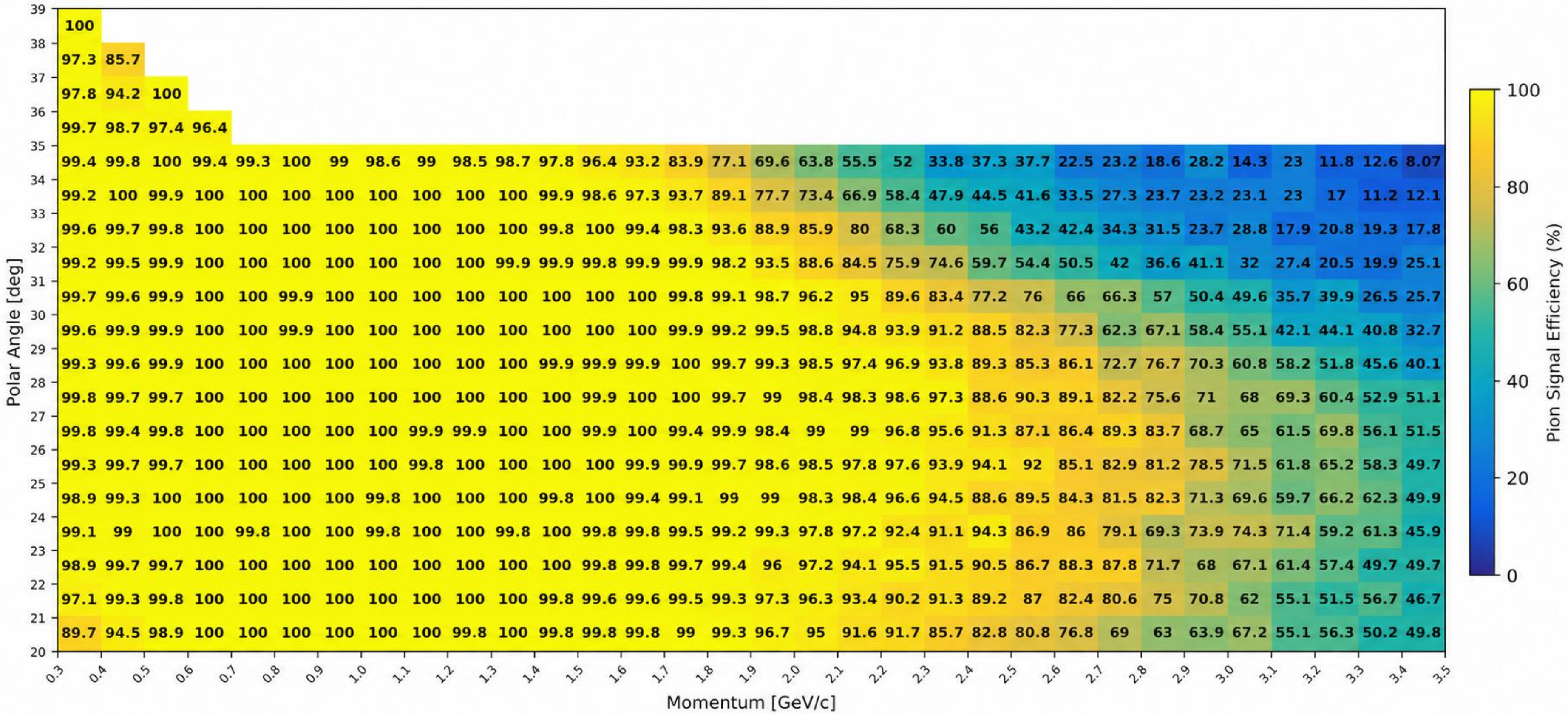
The Performance of CNN

Signal efficiency and background misidentification rate for pions and kaons in different momentum and polar-angle regions



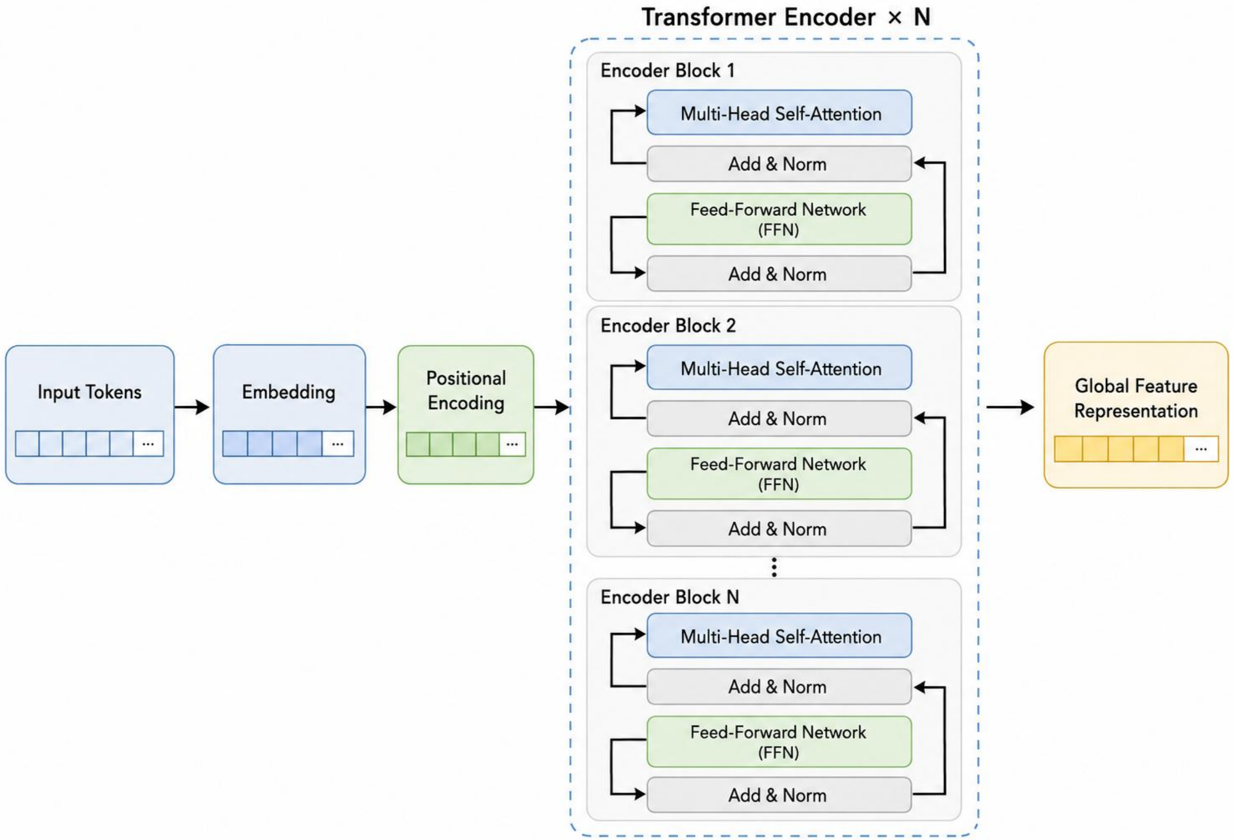


Using kaons as the background and controlling their misidentification rate below 2%, the pion signal-efficiency distribution is shown below:



- Image features + kinematic features
- Efficiency decreases at high momentum and large polar angle

Transformer



Method	Input format	Feature extraction	Main characteristics
CNN	2D pixel map	Local convolution kernels	Effective at extracting local spatial textures
Transformer	Token sequence / point-cloud features	Self-Attention	Effective at modeling global dependencies

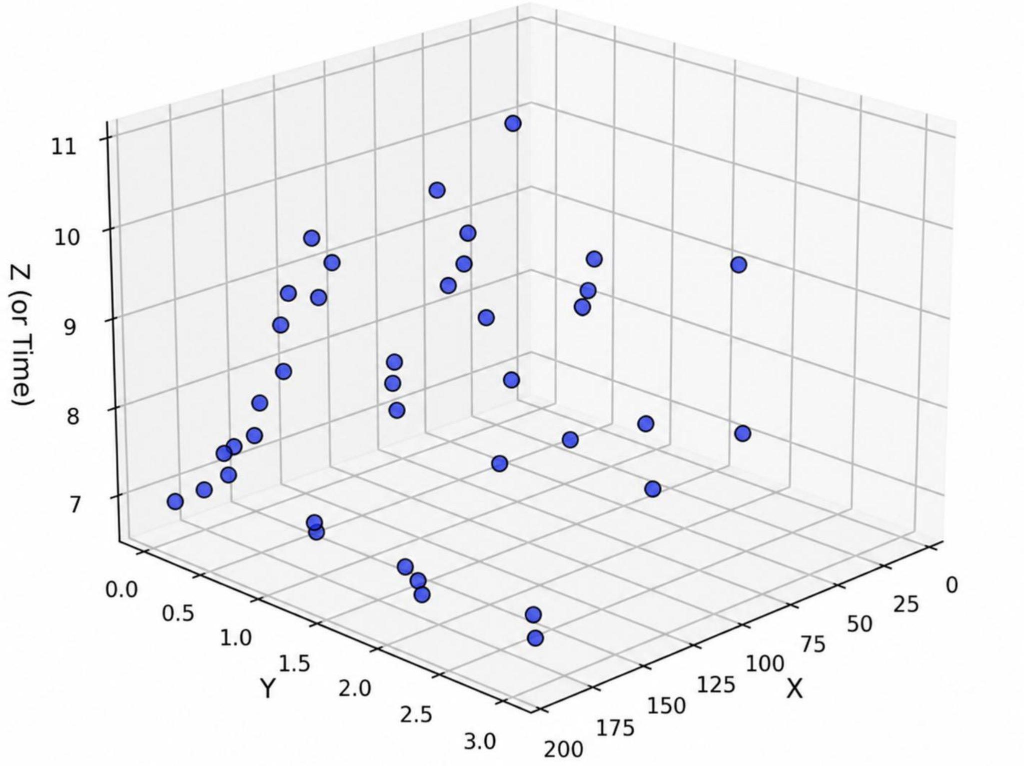
Transformer Architecture

- Input features are represented as a token sequence.
- Self-attention models global correlations.
- Multiple encoder layers extract high-level features.

Advantages over CNN

- CNN emphasize local patterns, whereas Transformer capture global dependencies.
- More flexible for sparse and irregular point-cloud inputs.
- Well suited for capturing the overall spatiotemporal structure of DTOF photon hits.

Point-cloud features



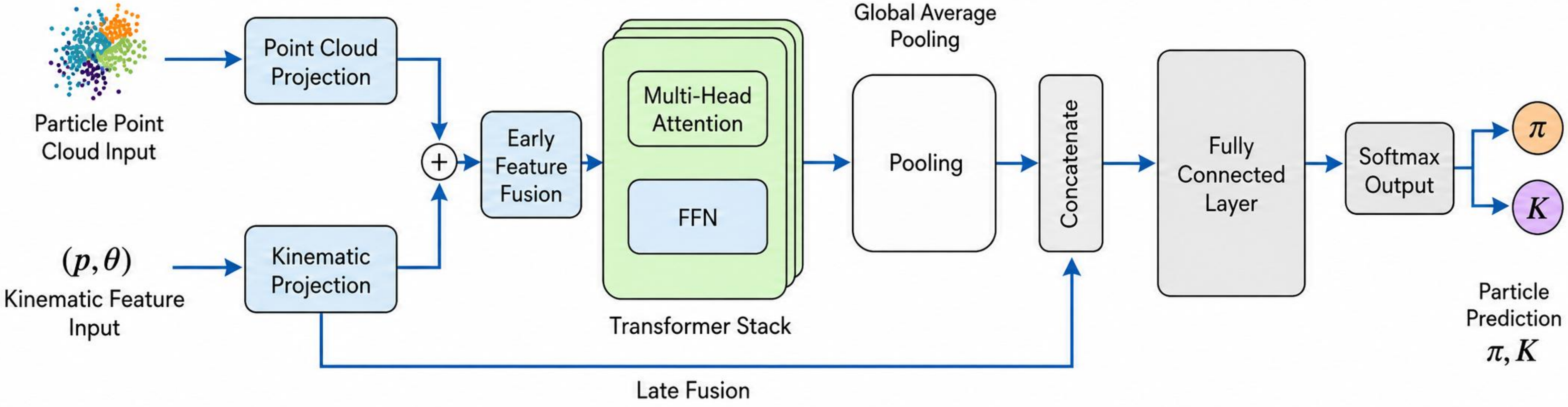
Using the raw response of the DTOF detector, three-dimensional point-cloud features are constructed:

- X-axis: PMT readout channel of the detected Cherenkov photon (Channel_x)
- Y-axis: PMT readout channel of the detected Cherenkov photon (Channel_y)
- *Z*-axis: arrival time of the Cherenkov photon detected by the PMT

Point-cloud features are normalized to map numerical values into [0, 1], which helps the neural network extract local spatial patterns.

- X channel: 192 channels along X, $X_{norm} = \frac{X}{191}$
- Y channel: 4 channels along Y, $Y_{norm} = \frac{Y}{3}$
- Time window: $Time_{norm} = \frac{Time - 5.0}{15.0}$

The Structure of Transformer

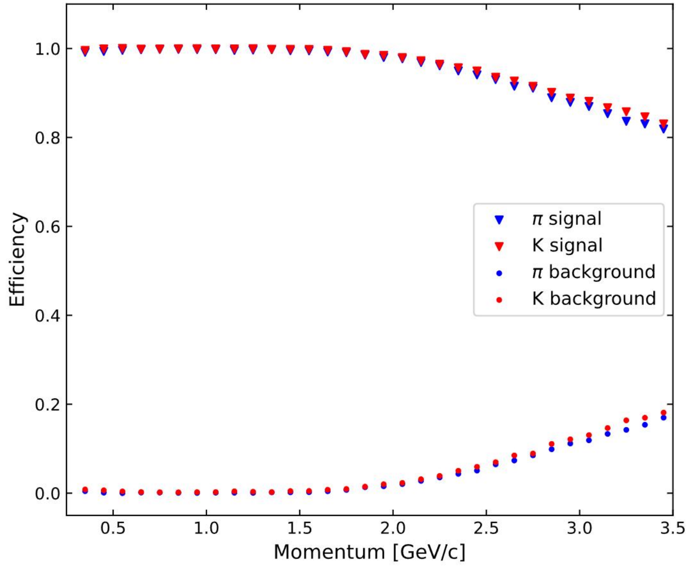


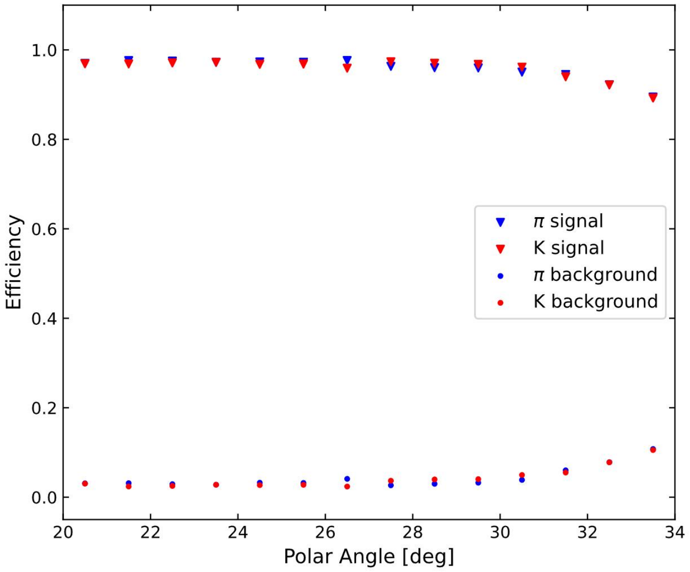
$$Attention(Q,K,V) = softmax(\frac{QK^T}{\sqrt{d_k}})V$$

- Transformer-based backbone with task-specific optimization
- Point-cloud and particle kinematic features as model inputs
- Attention-based modeling of correlations among photon hits
- Early and late feature fusion for enhanced physics-informed learning

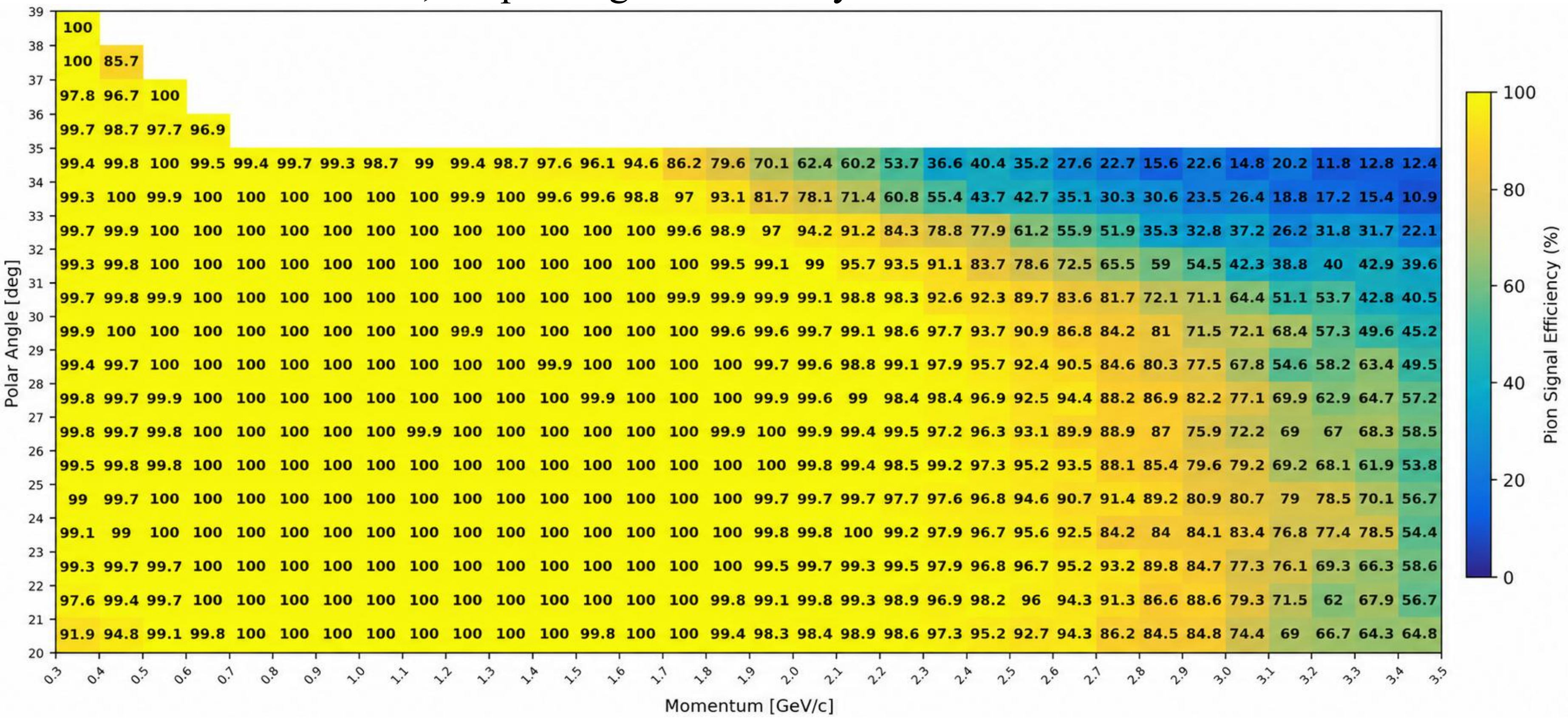
The Performance of Transformer

Signal efficiency and background misidentification rate for pions and kaons in different momentum and polar-angle regions





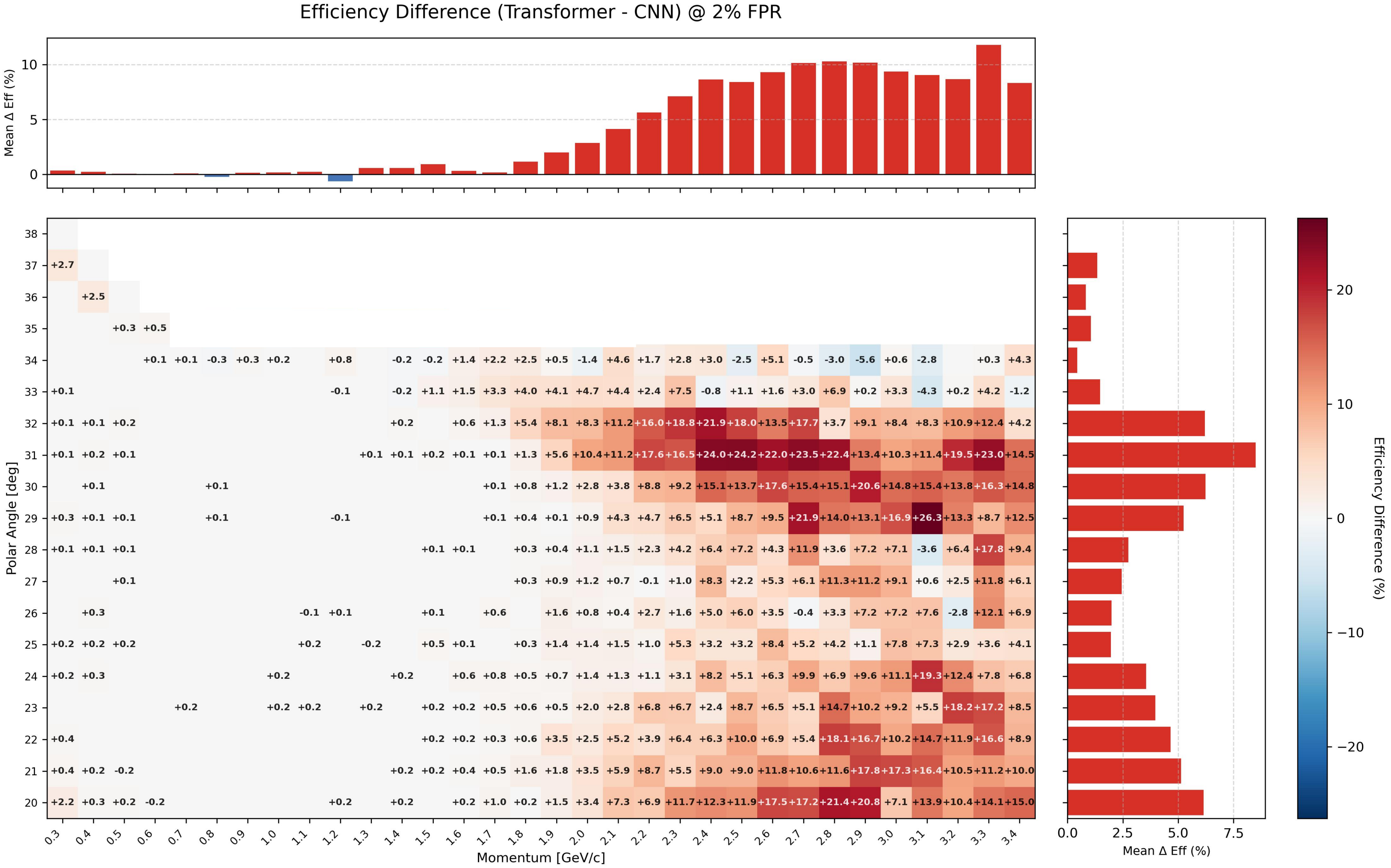
Using kaons as the background and controlling their misidentification rate below 2%, the pion signal-efficiency distribution is shown below:



- Transformer and CNN models use the same Monte Carlo (MC) samples
- Image features + kinematic features
- Efficiency decreases at high momentum and large polar angle

Performance Comparison

Comparison of pion-identification efficiency between Transformer and CNN in different momentum and polar-angle regions.



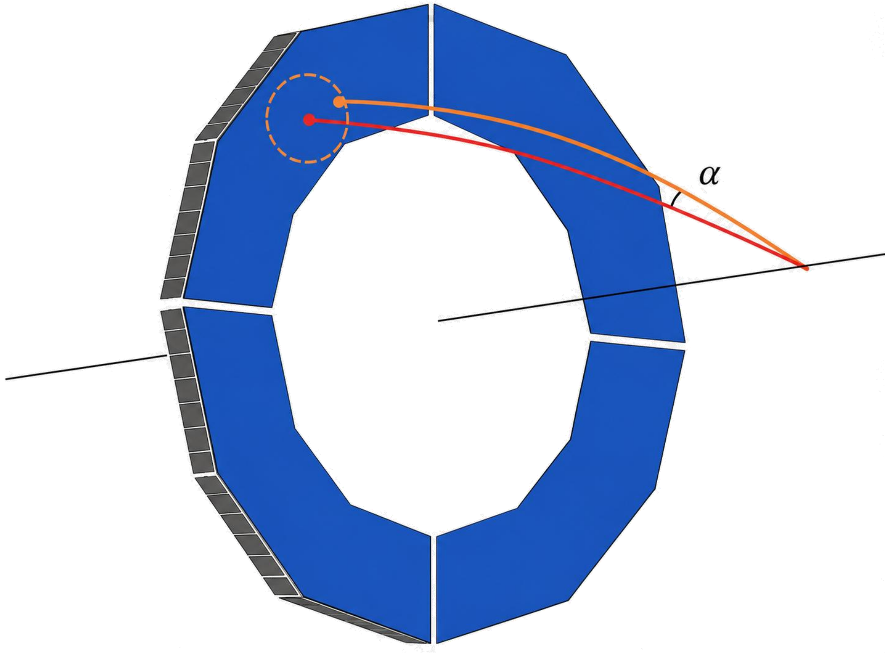
- ΔEff = Eff(Transformer) − Eff(CNN)

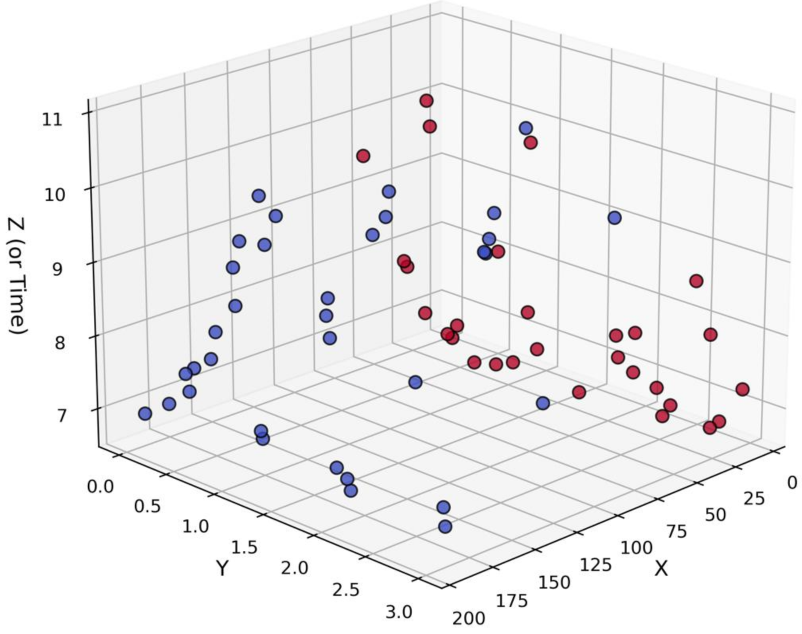
- Red: Transformer better
- Blue: CNN better

- Low momentum: comparable performance
- High momentum: Transformer advantage

- At high momentum, π/K separation relies more on the global photon-hit distribution

From Single-Particle PID to Dual-Particle Events

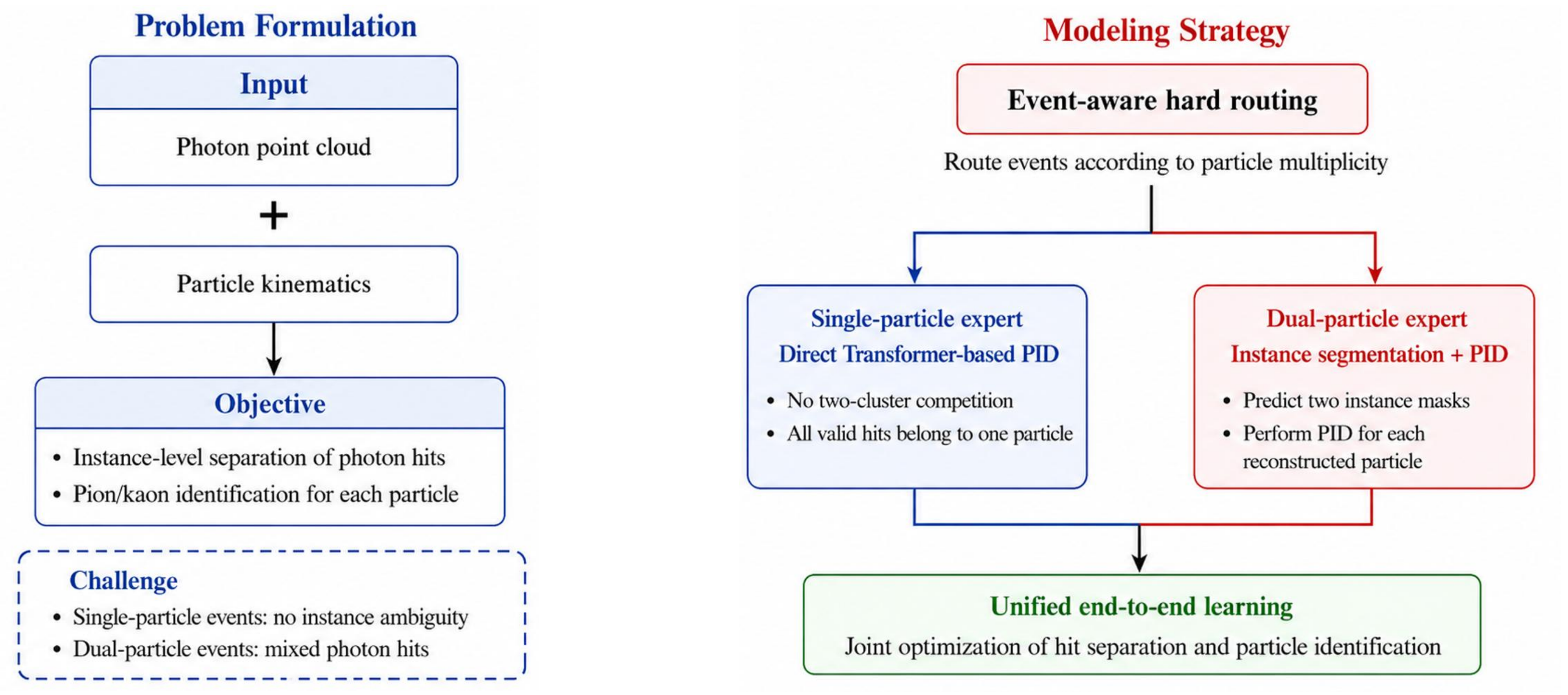




Photon hits from two particles are mixed in the same DTOF sector

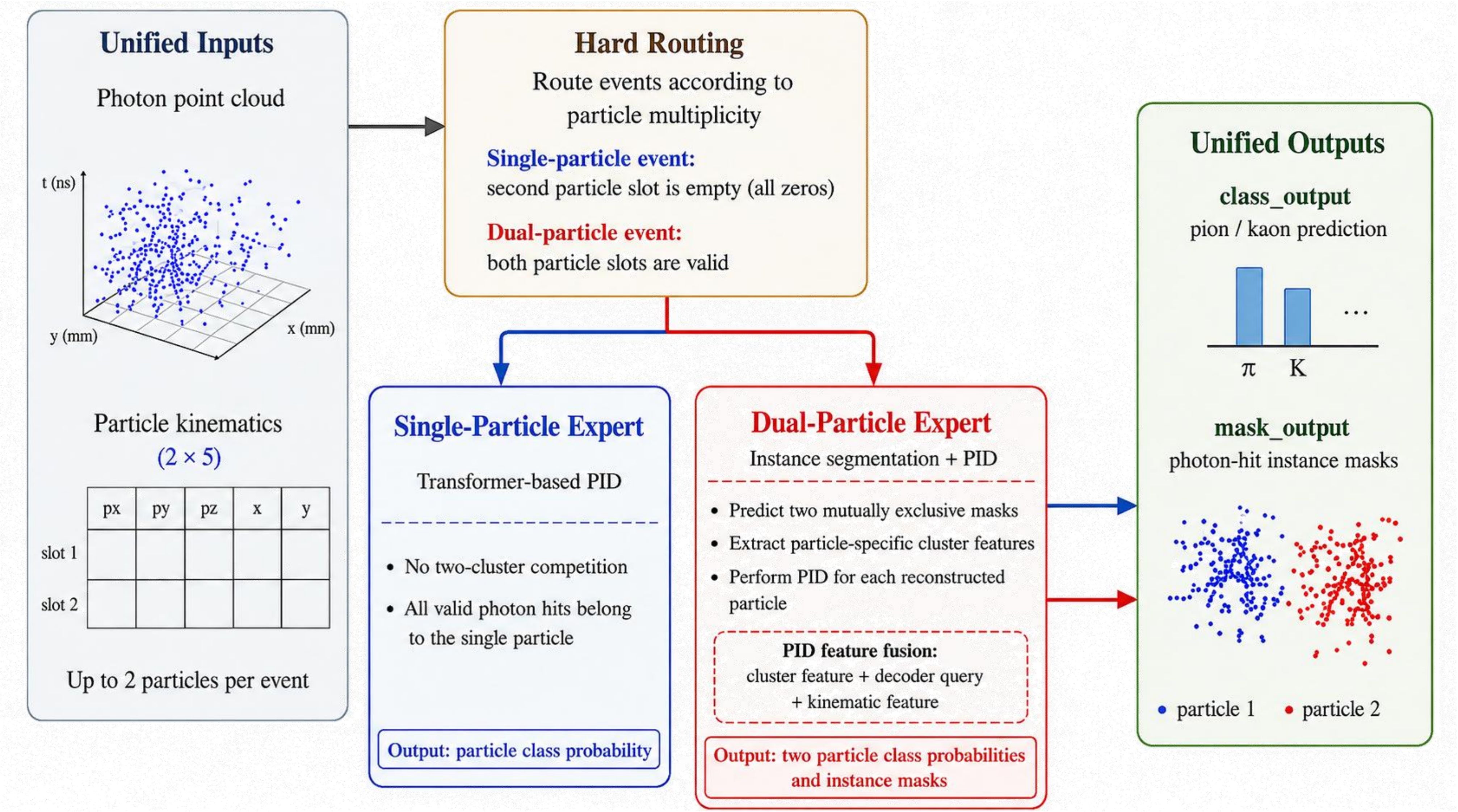
- Previous work focuses on single-particle PID
- In realistic DTOF events, more than one charged particle may enter the same sector
- Photon hits from different particles may overlap in space and time, making PID more challenging
- Multi-particle DTOF analysis requires PID, and instance-level separation of photon hits

Problem Formulation and Modeling Strategy



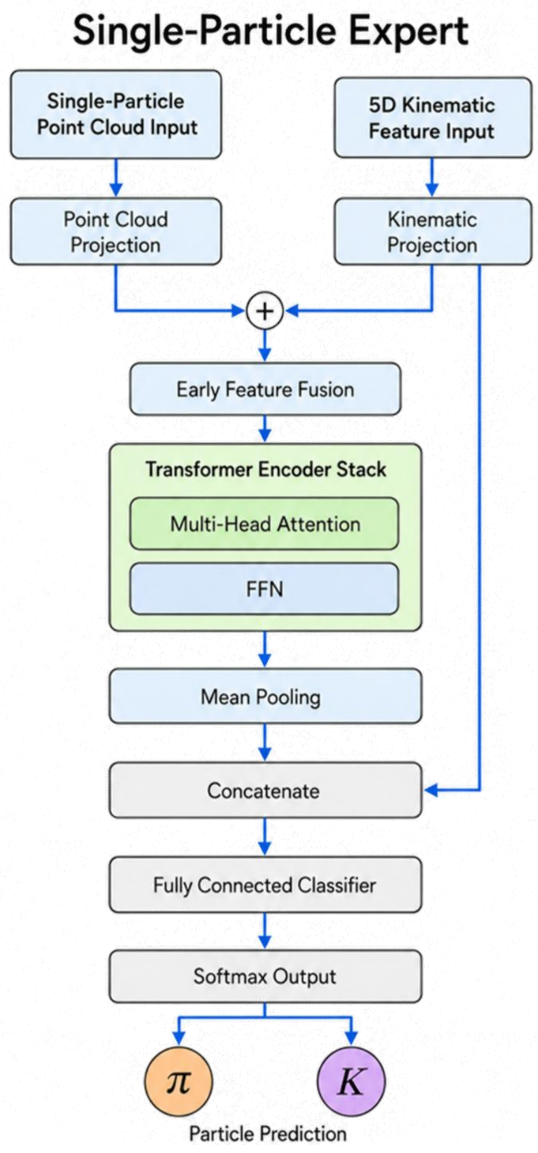
Single- and dual-particle events are handled by specialized experts within a unified framework.

Hard-Routed Dual-Expert Architecture

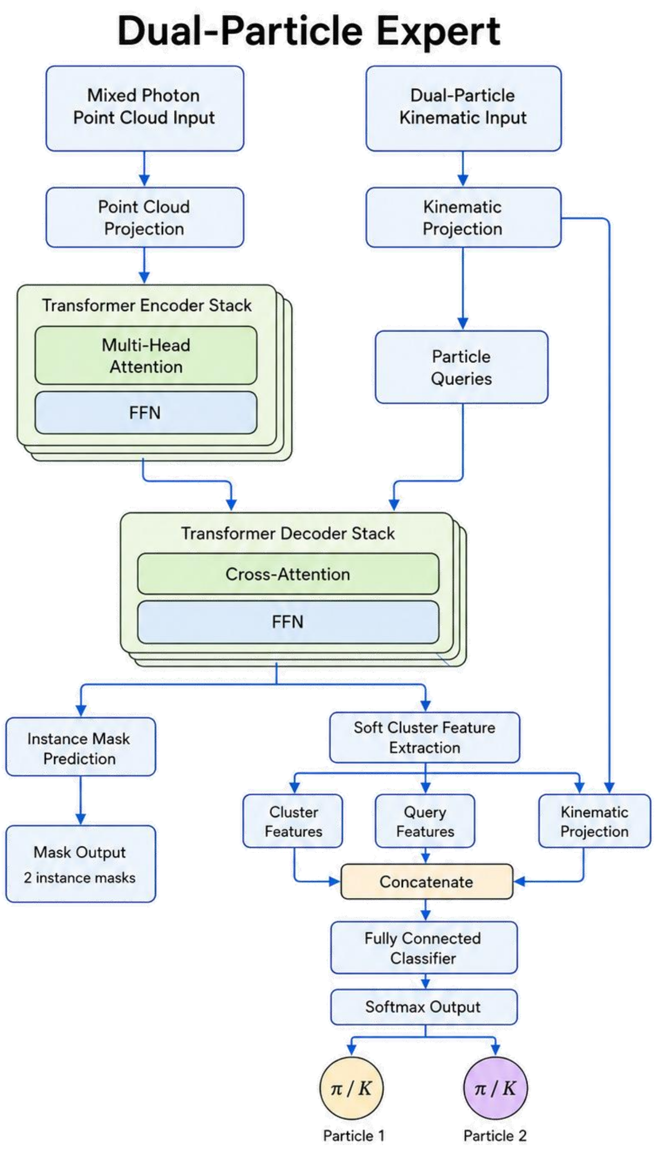


Single- and dual-particle events are handled by specialized experts within a unified end-to-end framework.

Single-Particle and Dual-Particle Experts



Direct PID

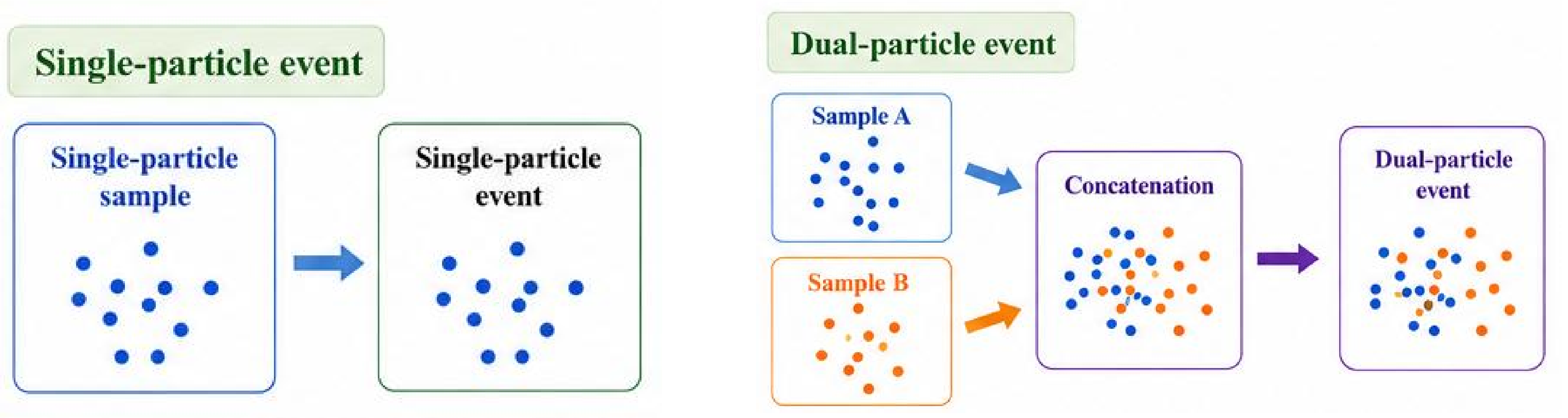


Instance segmentation + PID

Why Dual Experts?

- Topology-aware specialization
- Reduced task interference

Dataset Design

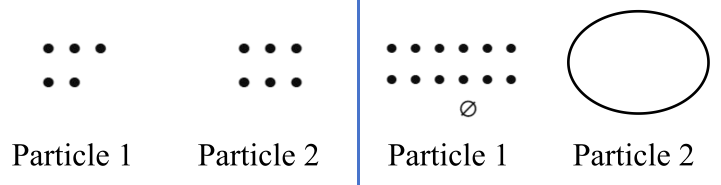


Single-particle π/K MC samples are simulated with OSCAR (2.7.0) offline software

- Momentum: 0.3–3.5 GeV/c
- Polar angle: 20°–39°
- Events per particle type: $1 * 10^7$
- Select single-particle data hitting one DTOF sector

- Single-particle : Dual-particle = 1 : 1
- Single-particle : $\pi: K = 1:1$
- Dual-particle : $\pi - \pi: K - \pi: K - K = 1:1:1$

Multi-Particle Performance

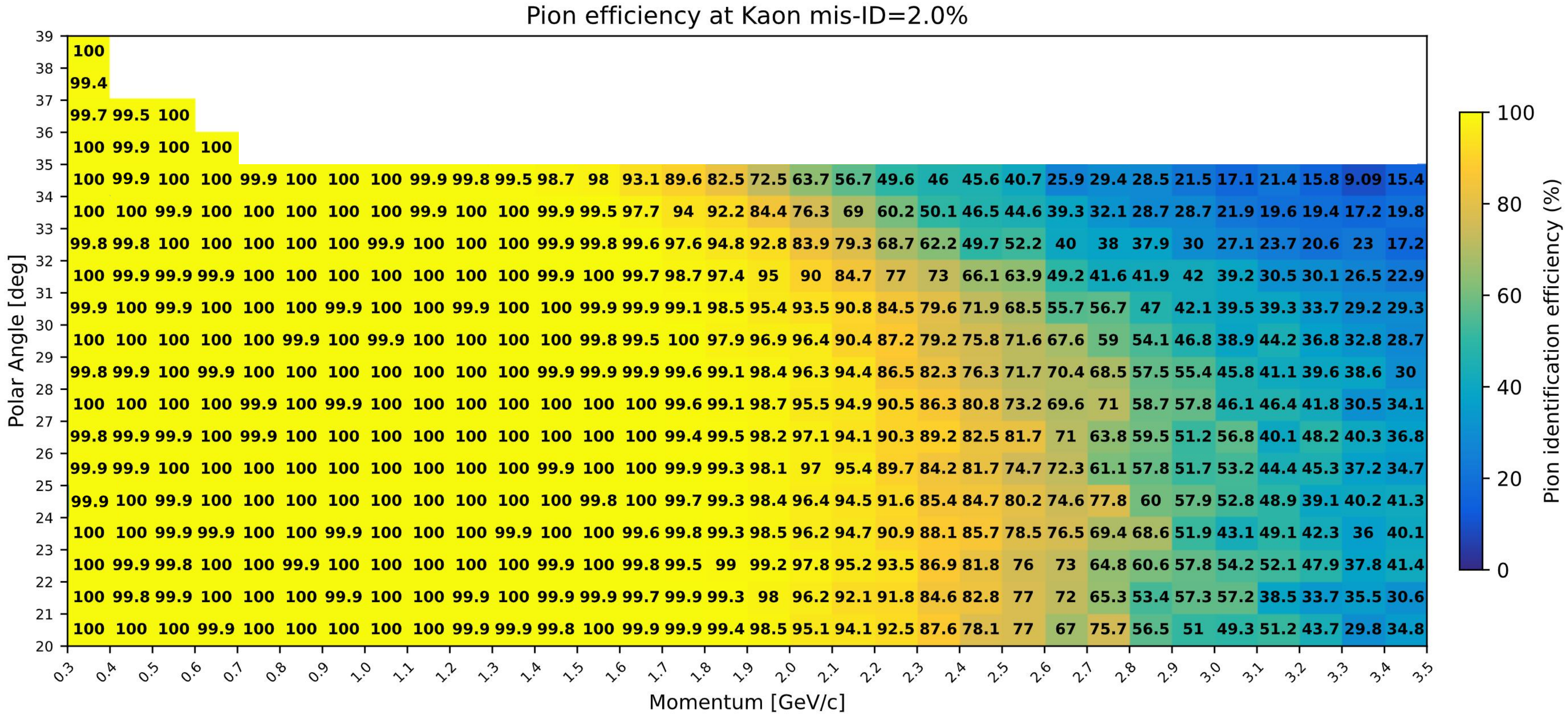


A dual-particle event is considered collapsed if either predicted particle cluster contains no photon hits after hard assignment.

$$R_{collapse} = 0.0188\%$$

$$R_{success} = 1 - R_{collapse} = 99.9812\%$$

Particle-Level Performance (test)	
Dual-Particle Photon Assignment Accuracy	$ACC_{photon} = 96.79\%$
Single-particle	$ACC_{single} = 94.50\%$
Dual-particle	$ACC_{dual} = 92.00\%$
Overall end-to-end	$ACC_{e2e} = 92.83\%$



The extremely low cluster-collapse rate (0.0188%) indicates robust instance reconstruction; the remaining performance limitation mainly arises from particle identification.

Summary

Single-Particle PID

- Joint learning from point-cloud and kinematic features
- Global photon-hit correlations modeled by self-attention
- Improved PID performance over CNN, especially at high momentum

Multi-Particle PID

- Joint photon-hit separation and particle identification
- Hard-routed experts for different event topologies
- Direct PID for single particles; instance segmentation + PID for dual particles

Outlook

- Validate the multi-particle framework using directly simulated multi-particle MC events

Quantum Computing and Machine Learning Workshop 2026

Thanks