PID Studies with dN/dx and Time-of-Flight for CEPC

Guang Zhao, Dian Yu, Yue Chang, Houqian Ding, Mingyi Dong, Yunyun Fan, Huirong Qi, Linghui Wu, Jinxian Zhang

The 2025 International Workshop on the CEPC
11/10/2025, Guangzhou

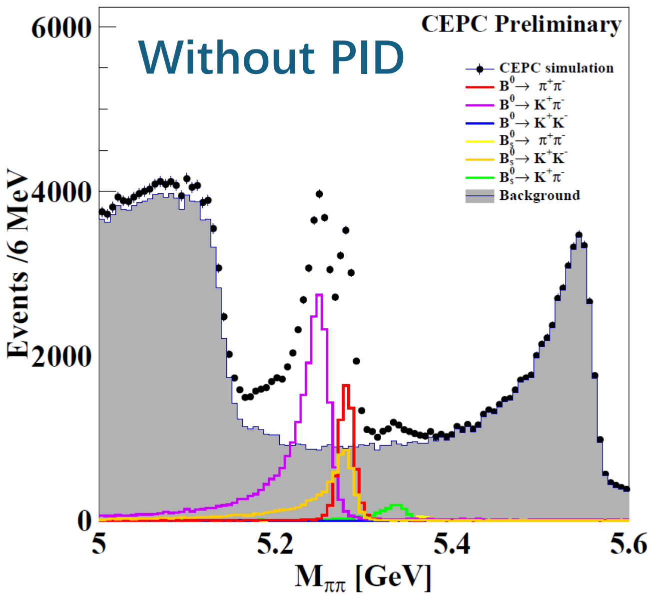
Post-TDR efforts

- ✓ **dN/dx reconstruction with deep learning in TPC (arXiv: 2510.10628)**
- ✓ **Time-of-flight in ITK (arXiv: 2507.18164)**

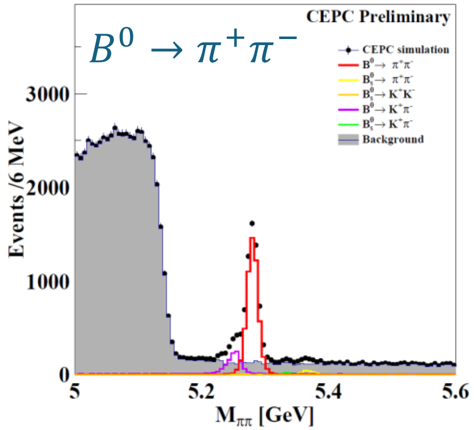
Motivation: Particle identification

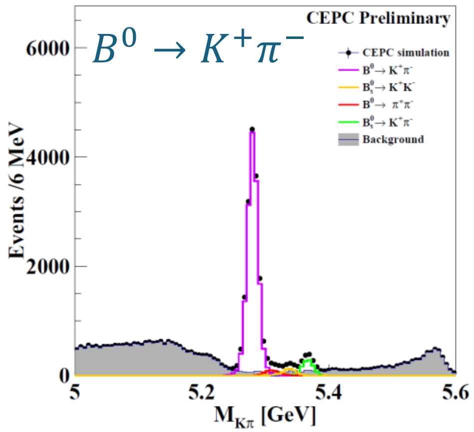
- **PID is essential for high energy physics experiments**
  - Suppressing combinatorics
  - Distinguishing between same topology final-states
  - Adding valuable additional information for flavor tagging of jets
  - …

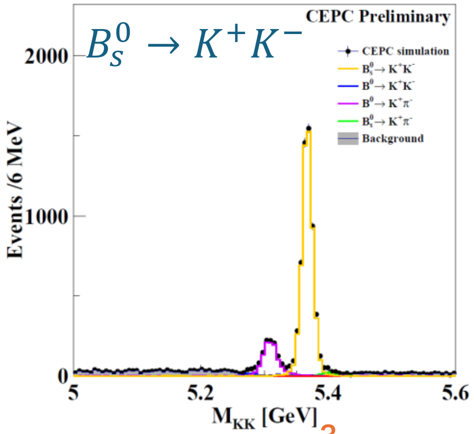
Benchmark channel:
$\mathbf{B^0_{(s)} \rightarrow h^+h'^-}$



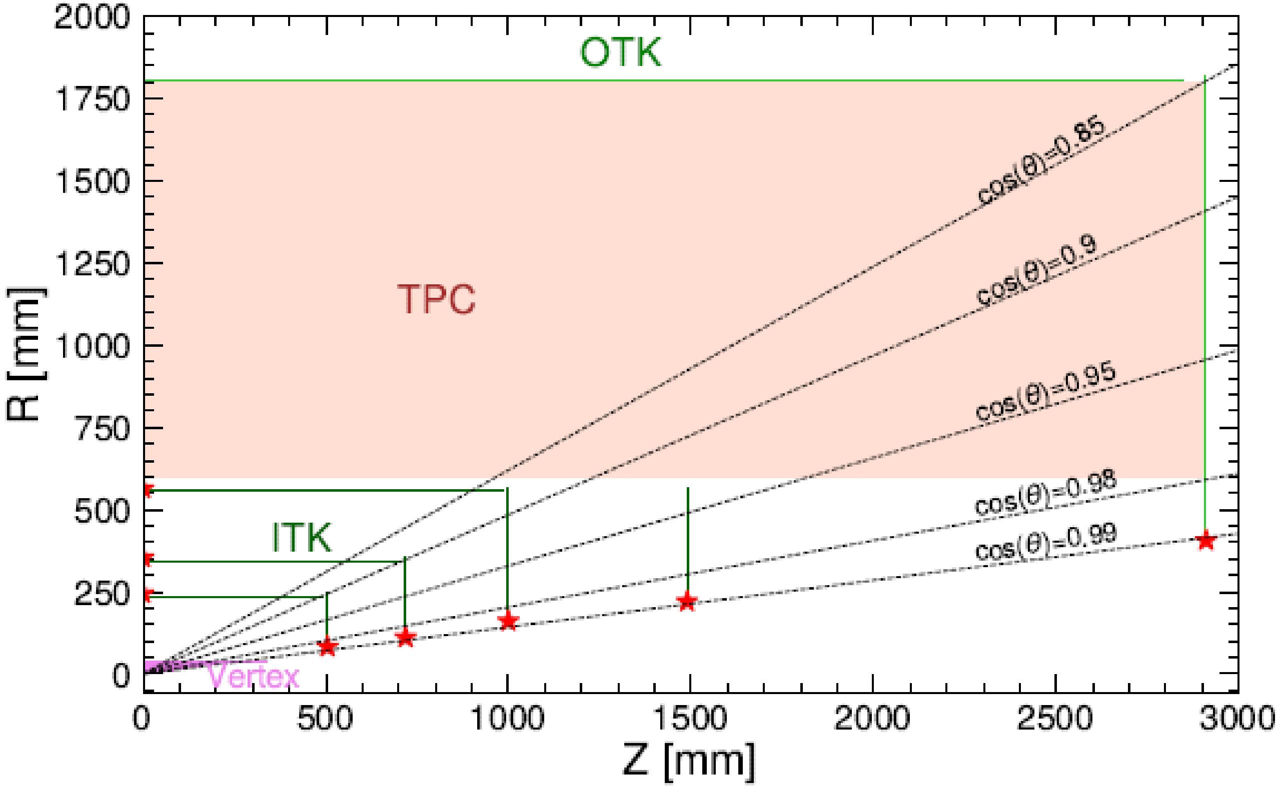
With PID







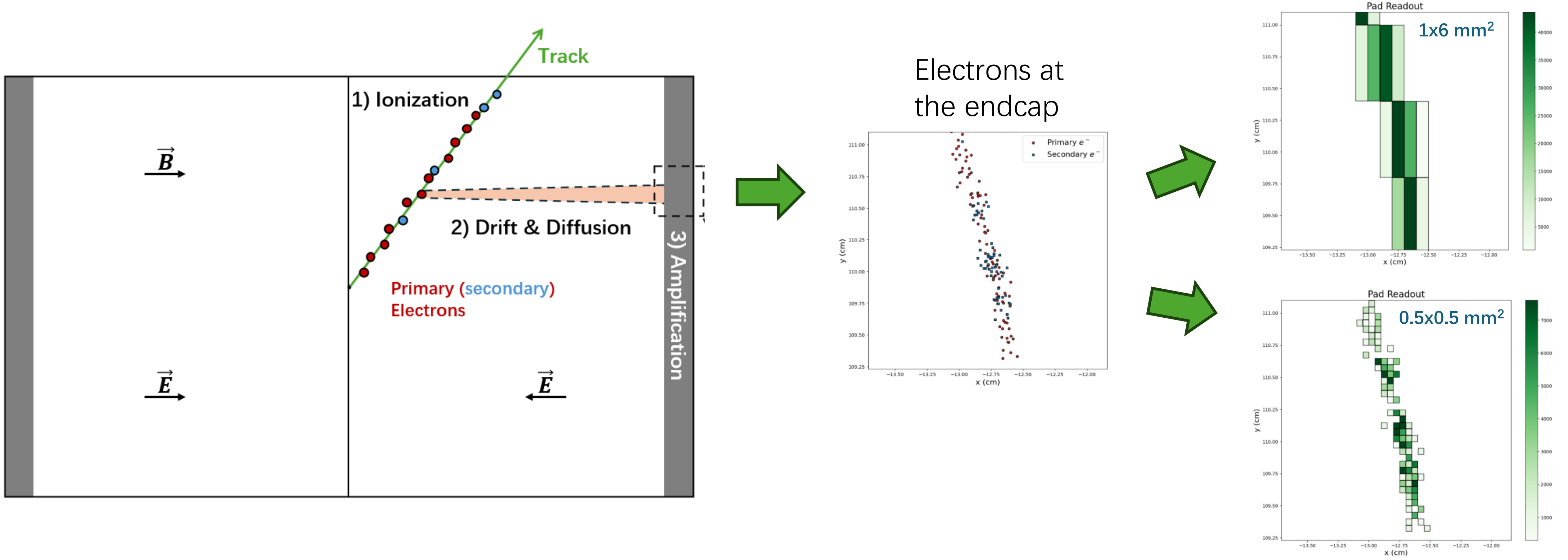
CEPC hadron ID system



- **TPC:**
  - Large volume:
    - L = 5.8 m
    - 0.6 m < R < 1.8 m
  - High granularity: 0.5 x 0.5 mm^2
- **OTK:**
  - Strip AC-LGAD: σ = 50 ps
- **ITK (post-TDR):**
  - Outermost layer
  - Pixel AC-LGAD: σ = 30 ps

dN/dx reconstruction with deep learning in TPC

Ionization measurement in TPCs



dE/dx (traditional method):

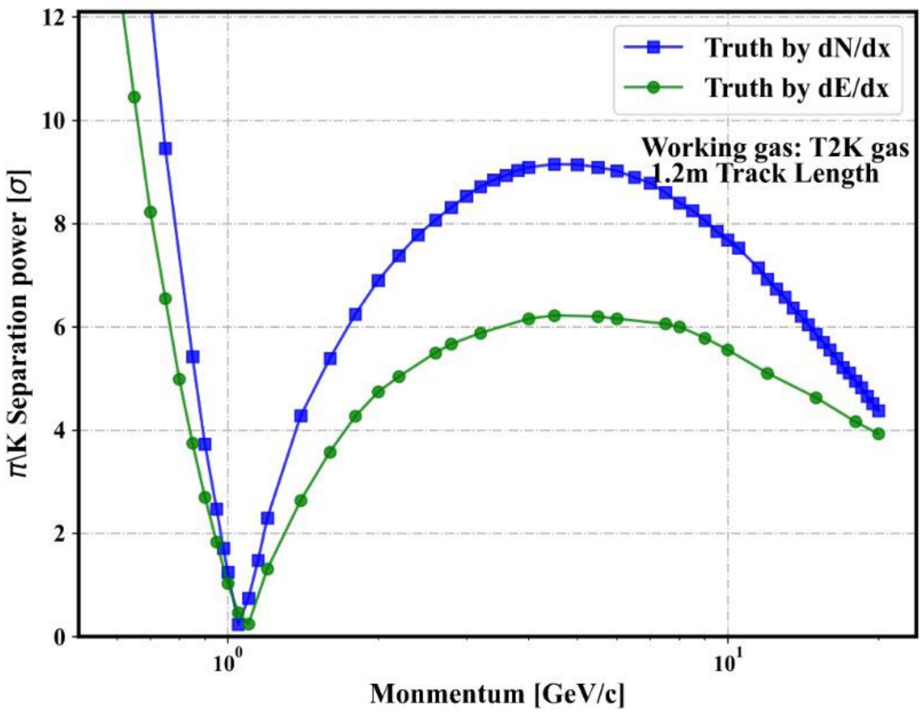
- **Method:** Total energy loss measurement by integrating the energies in large pads
- **Characteristics:**
  - Large fluctuations from energy measurements, amplification, secondary ionizations, etc

dN/dx or cluster counting ("ideal" method):

- **Method:** Number of primary ionization cluster measurement, **requiring high granularity readout**
- **Characteristics:**
  - **Small fluctuation (resolution potentially improved by a factor of 2)**

PID evaluation: Particle separation power

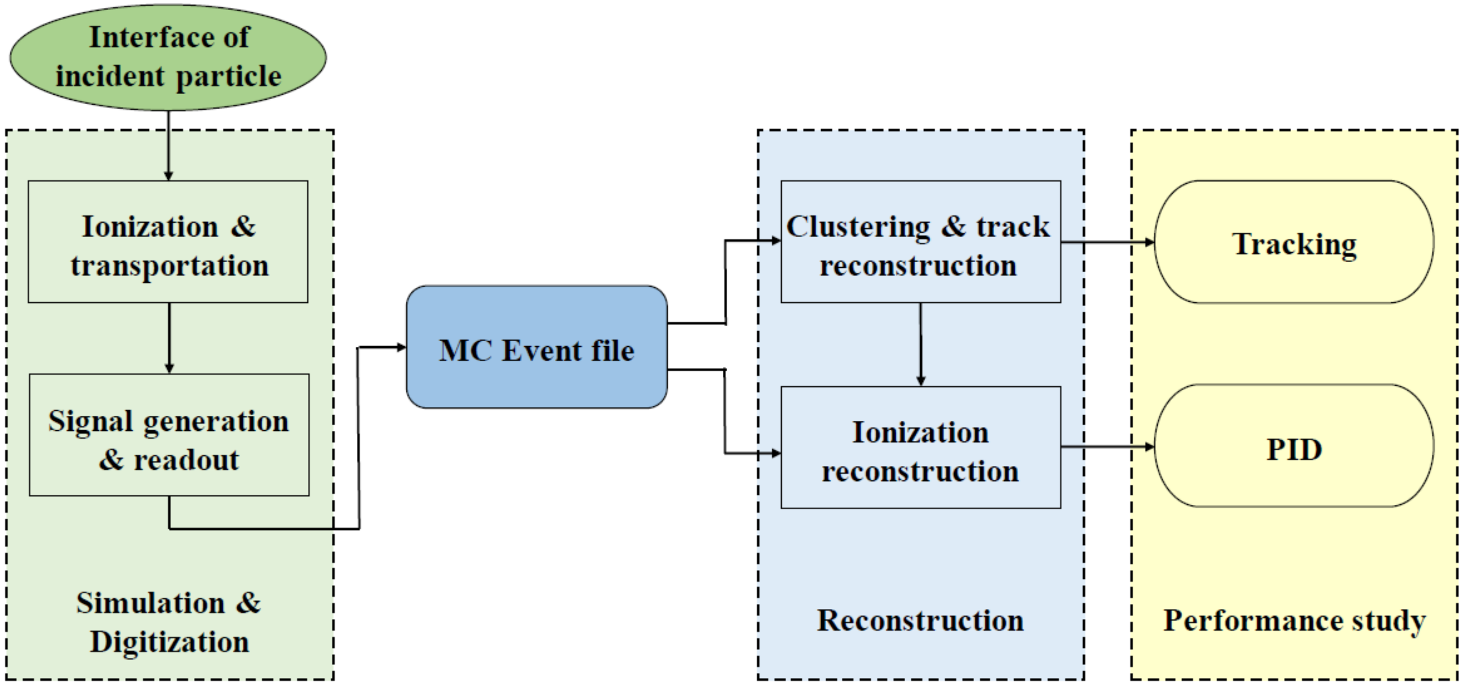
- **Definition:** $\frac{\text{separation}}{\text{resolution}} = \frac{|\mu_A - \mu_B|}{\sqrt{\frac{\sigma_A^2 + \sigma_B^2}{2}}}$
  - Very important for physics
  - Resolution is NOT important



dN/dx significantly enhances PID performance and represents a breakthrough

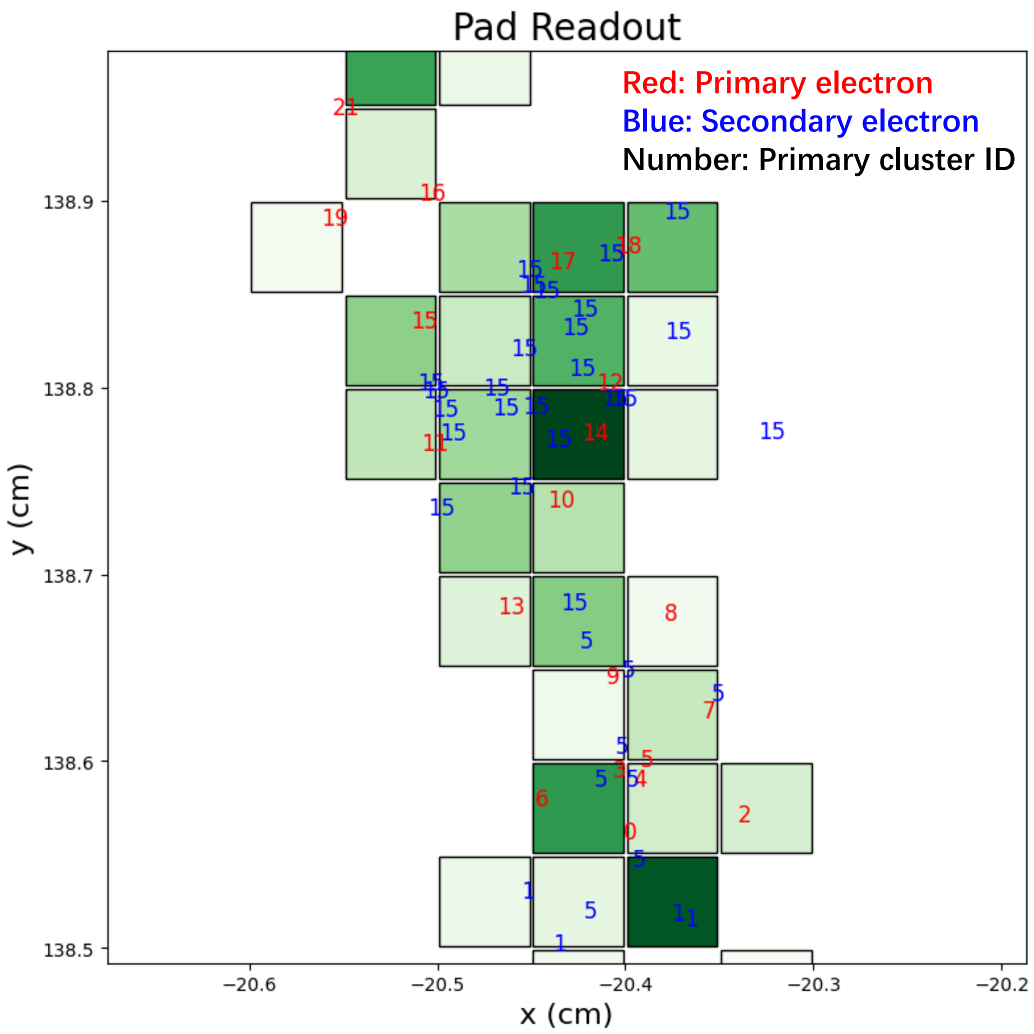
dN/dx is proposed in future detector designs (e.g., IDEA/ILD)

TPC simulation framework



- **Precise simulation:**
  - Full geometry
  - Gas mixture: $Ar/CF_4/iC_4H_{10}$ (95:3:2)
  - Magnetic field: B = 3 T
  - Ionization: Heed
- **Parameterized digitization:**
  - Transport: Drift and diffusion
  - Amplification & readout (from experiment input):
    - Gas gain: ~2000; Noise: ~100e$^-$/ch.; Eff. width: ~100 um

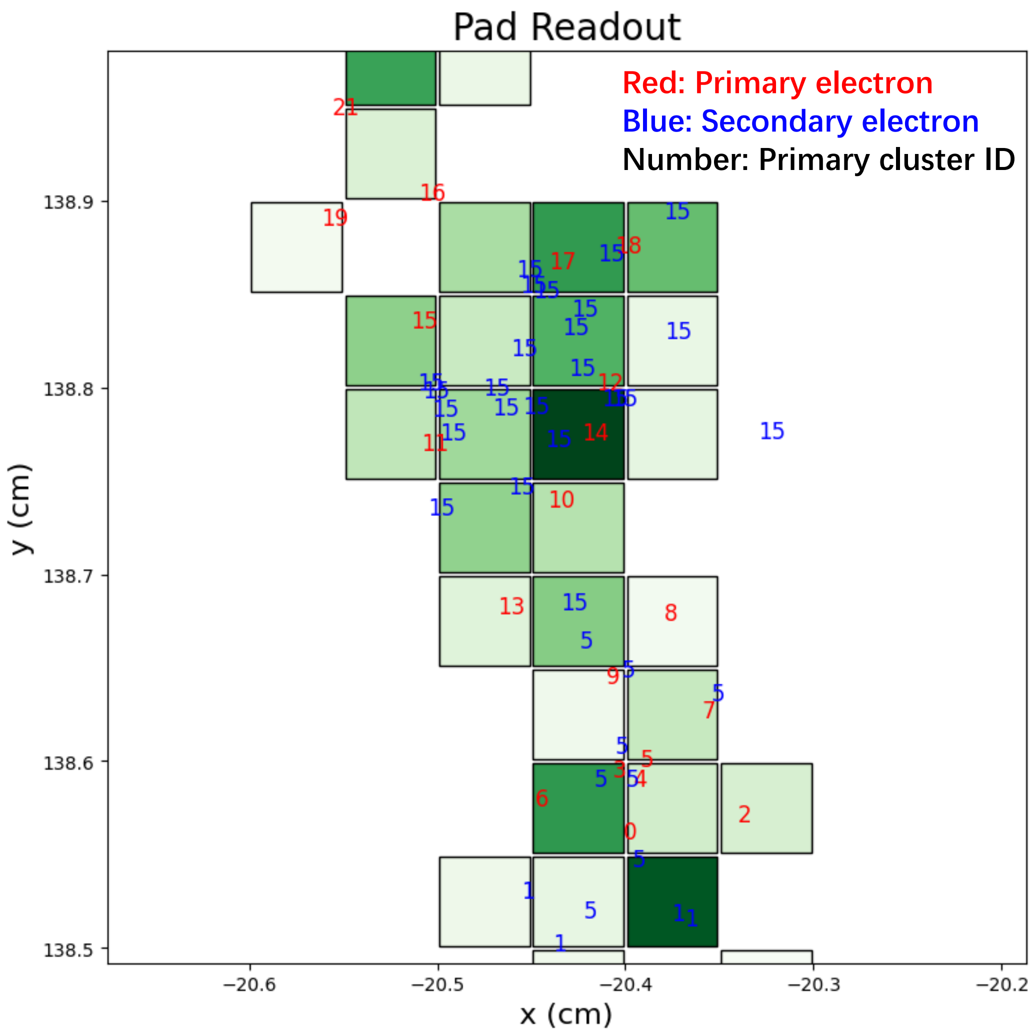
dN/dx reconstruction



Readout information:

- Charge/timing in each pad

dN/dx reconstruction



Readout information:

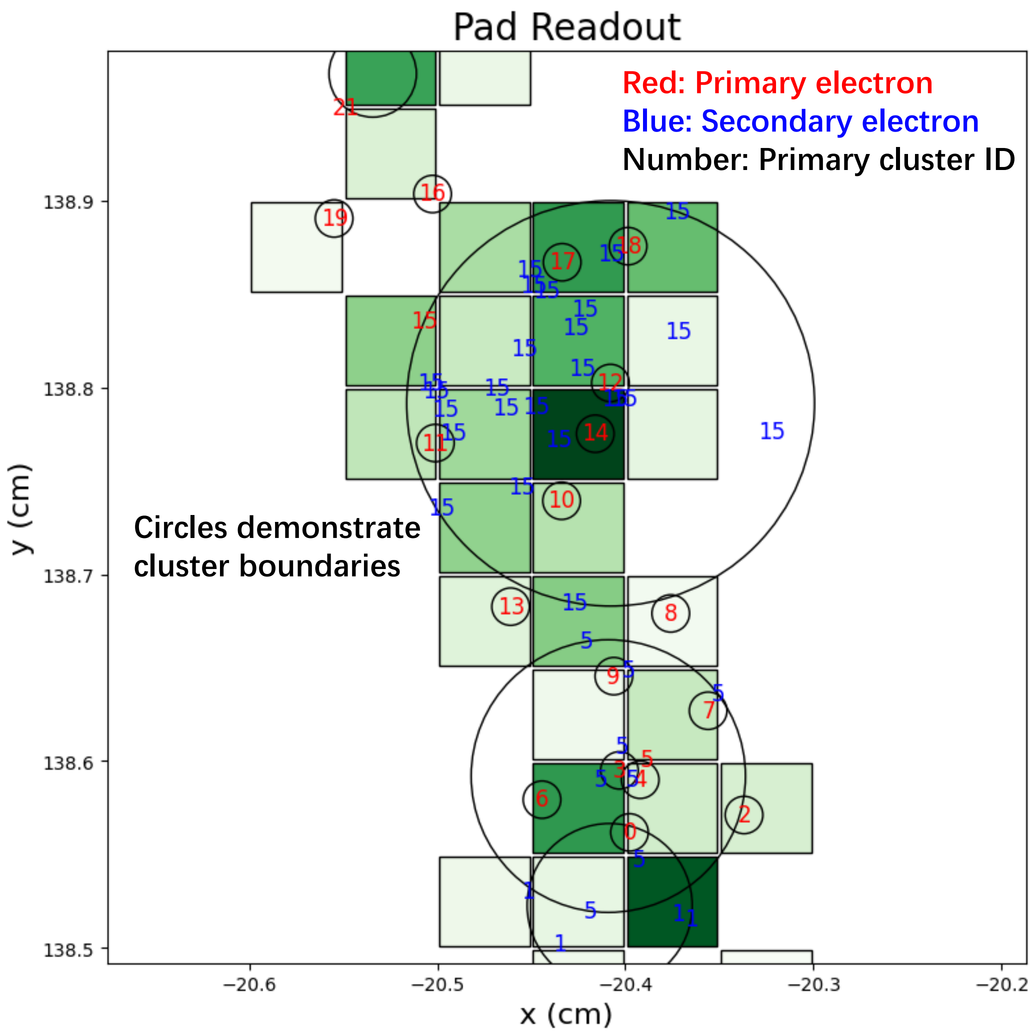
- Charge/timing in each pad

Reconstruction:

- Determine the number of primary electrons from 2D readout pads

  (or mitigating the impact of secondary electrons)

dN/dx reconstruction: Challenges



Cluster 1, 5, 15 overlap with other clusters

Challenges:

- Max. drift length 2.9m ➔ max. diffusion 550um ➔ Overlapped clusters ➔ Very difficult to use position locality for clustering

Baseline method: Truncated mean

- **Most used method:**
  - Reject measurements M_i with highest values ➔ most related to the secondaries
  - Calculate mean of remaining samples

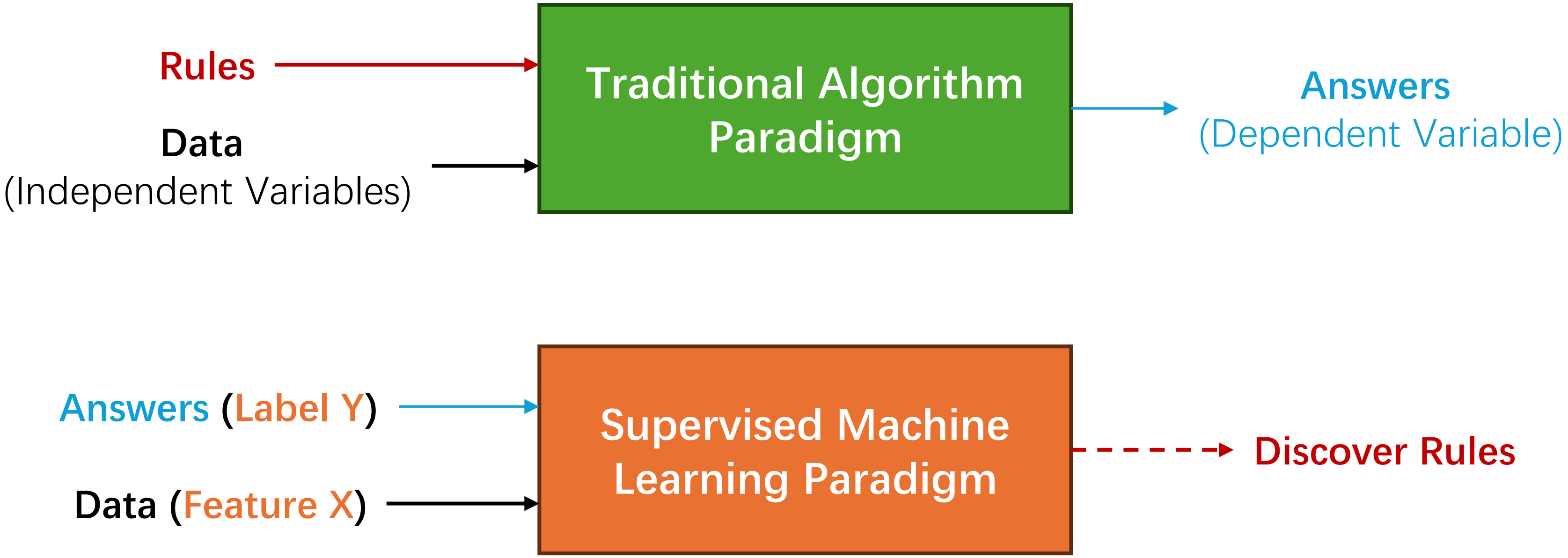
Average over the lowest n measurements:

$$\langle M \rangle_\alpha = \frac{1}{n} \sum_{i=1}^{n} M_i$$

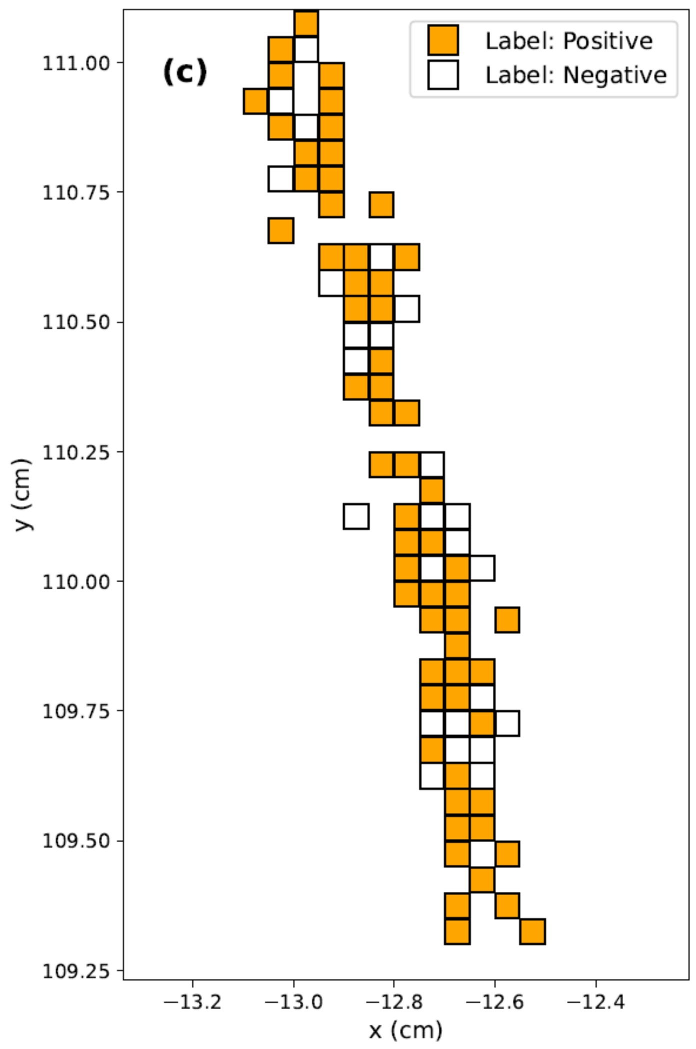
where $M_i \leq M_{i+1}$ for $i = 1, \ldots, N-1$ and $\alpha = n/N$ is a fraction.

- **Measurement $\boldsymbol{M}$ refers to:**
  - **charge** for dE/dx
  - **number of activated pads** for dN/dx

Deep-learning-based method

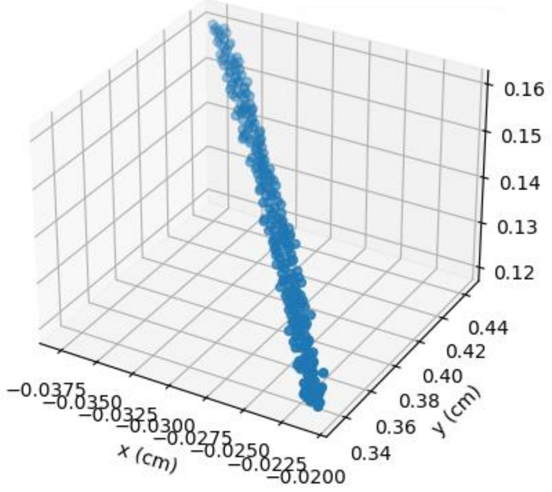


Problem definition



- **Method: Supervised learning**
  - Pad labeling:
    - Positive: More than 1 primary e^-
    - Negative: No primary e^-
  - Perform binary classification

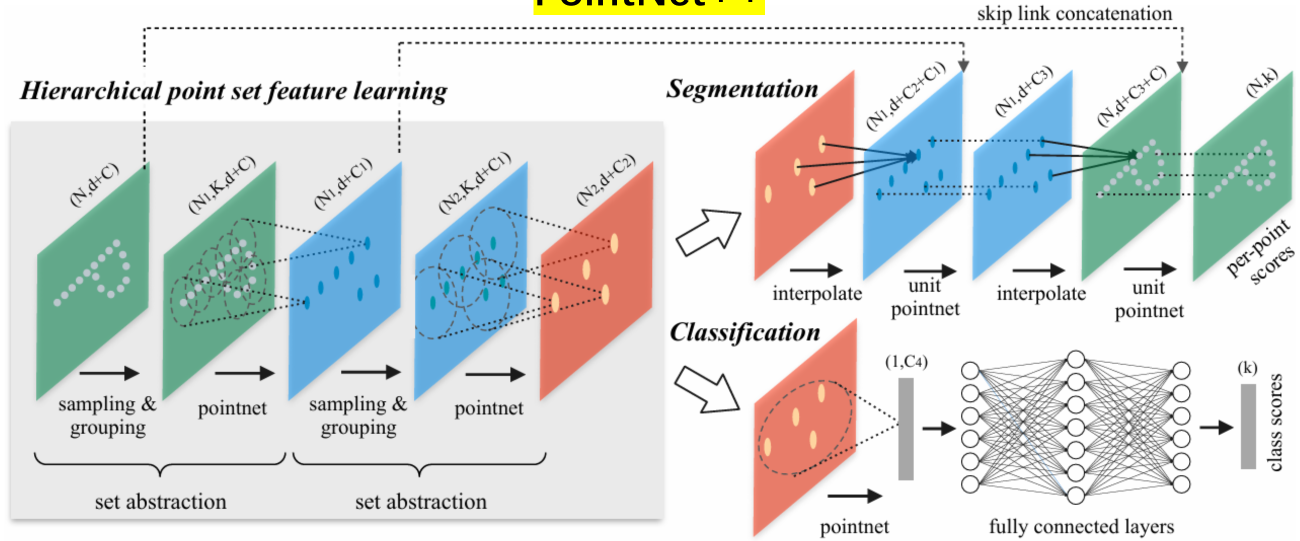
- **Data structure: Point cloud**
  - 2D position + 1D timing ➔ 3D spatial coordinates
    ➔ Point-cloud-based methods



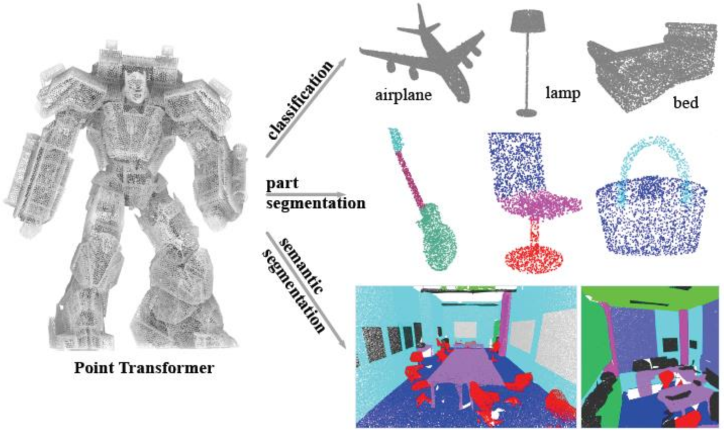
Point-based methods

- **PointNet++ [CR Qi, et. al., NIPS 2017]**
  - Hierarchical structured: Encoder-decoder with skip links
  - Basic building block for feature aggregation: MLP + pooling

PointNet++



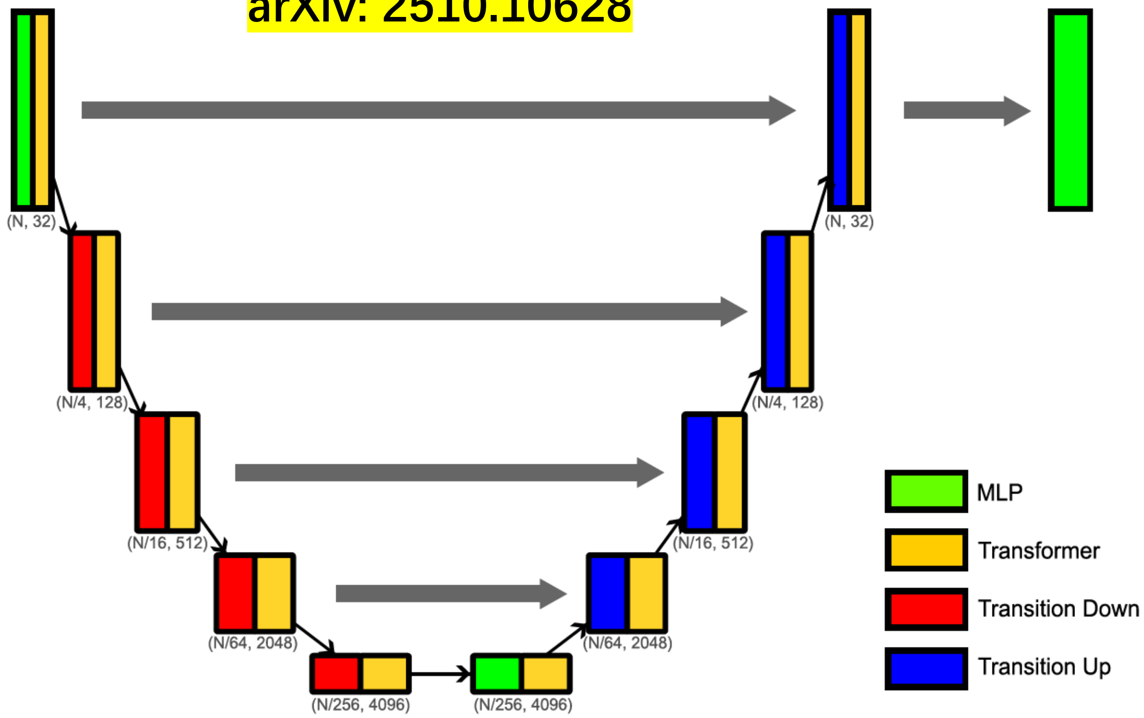
Point Transformer



- **Point Transformer [H. Zhao, et. al., IEEE/CVF 2021]**
  - Replace the MLP + pooling with self-attention blocks

GraphPT: Point Transformer on Graphs

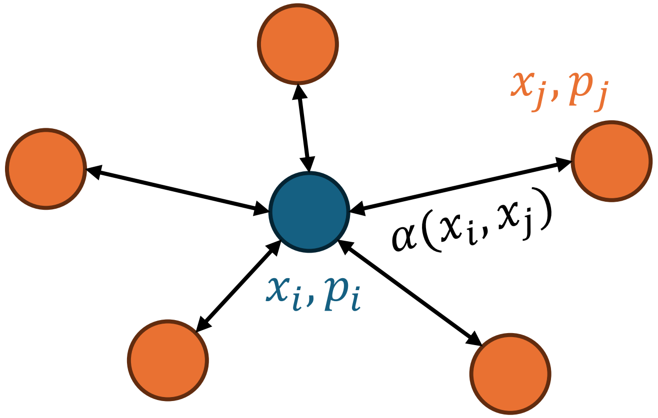
arXiv: 2510.10628



- **Design principles:**
  - PT built upon GNNs
  - Self-attention via message passing
- **Backbone structure:**
  - U-Net-based hierarchical architecture
    - Encoder-decoder with skips
    - MLP + softmax layer outputs [0, 1]

GraphPT: Point Transformer on Graphs

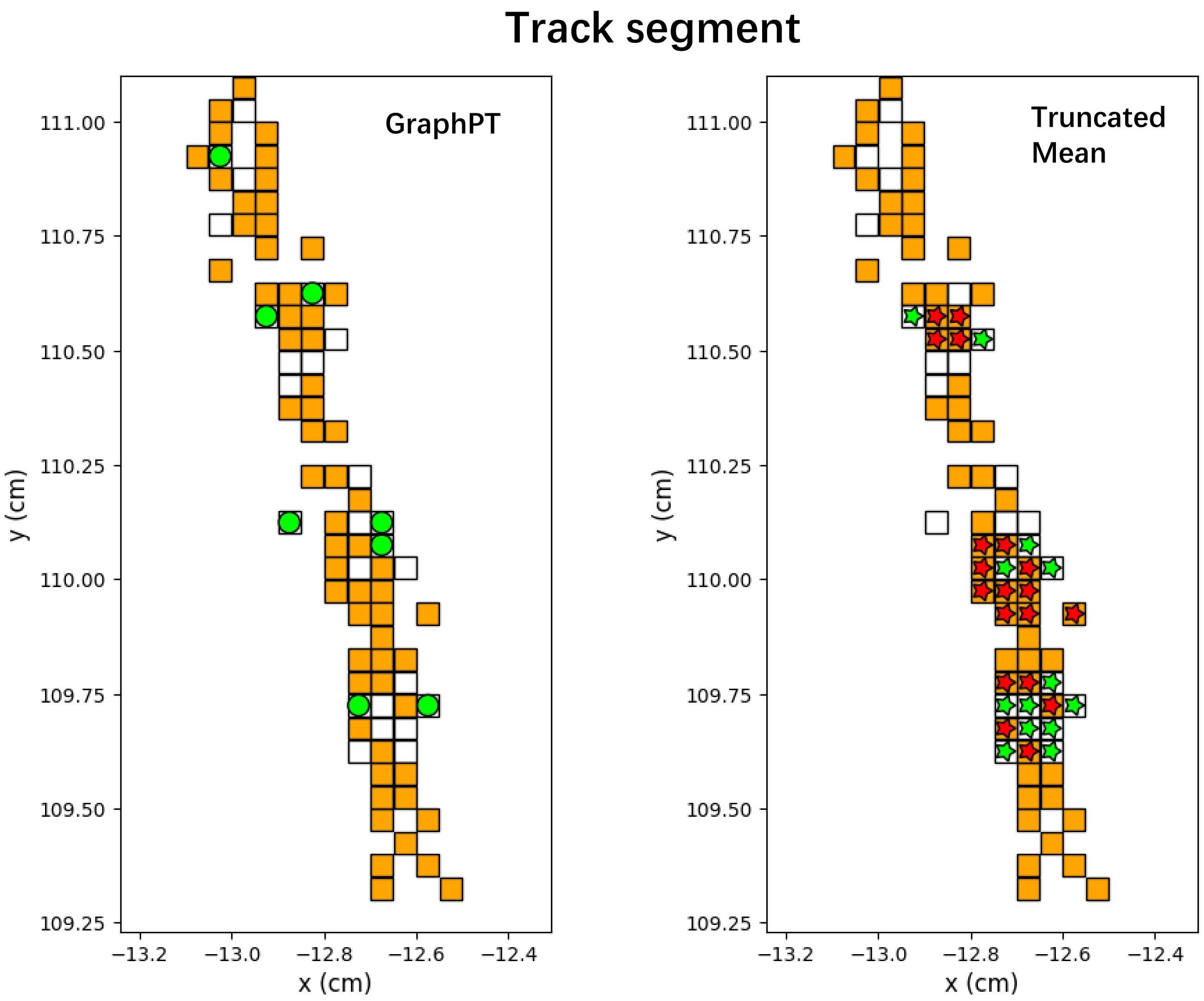
arXiv: 2510.10628



x: (Q, T)
p: 3D position

- **Self-attention mechanisms:**
  - Message-passing:
    - $x_i' = \beta(x_i) + \Sigma_{j\in N(i)}\alpha(x_i, x_j)\beta(x_j)$
  - Subtract operator (from PT):
    - $\alpha(x_i, x_j) = \text{softmax}(\delta(\phi(x_i) - \psi(x_j) + \theta(p_i - p_j)))$
  - Multi-head dot-product operator (**this work**):
    - $\alpha(x_i, x_j) = \text{softmax}\left(\frac{\phi(x_i)^{\mathrm{T}}(\psi(x_j)+\theta(p_i-p_j))}{\sqrt{d_{\text{out}}}}\right)$

Classification visualization



Color code:

- Orange pad: Ground truth positive
- Write pad: Ground truth negative
- Green marker: True negative ☺
- Red marker: False negative ☹

Marker:

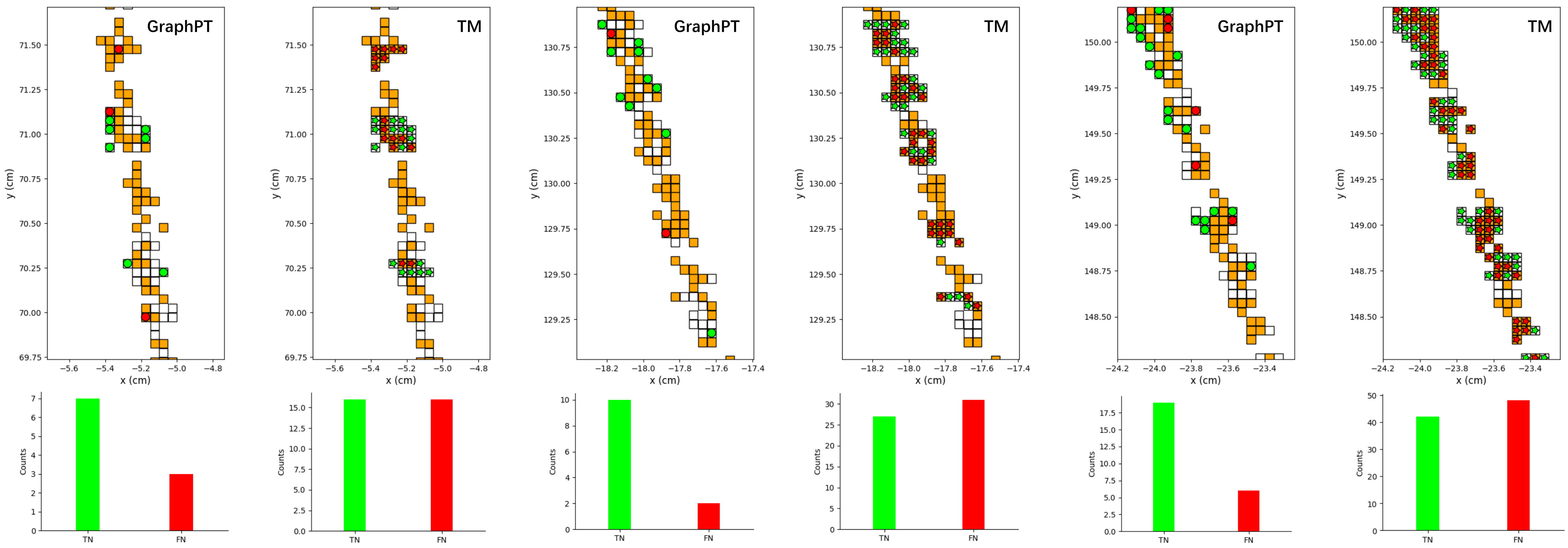
- Circle: Classified as negative by GraphPT
- Star: Classified as negative by Truncated Mean

Conclusions:

- Less false negative predictions by GraphPT
- Most signals are preserved, leading to higher signal efficiencies

More classification visualization

Track segment 2 | Track segment 3 | Track segment 4



Classification performance

- **Metrics:**
  - Accuracy = $\frac{\mathbf{TP+TN}}{\mathbf{TP+TN+FP+FN}}$
  - F1 score = $\mathbf{2} \times \frac{\boldsymbol{P \times R}}{\boldsymbol{P+R}}$, where $\boldsymbol{P} = \frac{\mathbf{TP}}{\mathbf{TP+FP}}$ and $\boldsymbol{R} = \frac{\mathbf{TP}}{\mathbf{TP+FN}}$

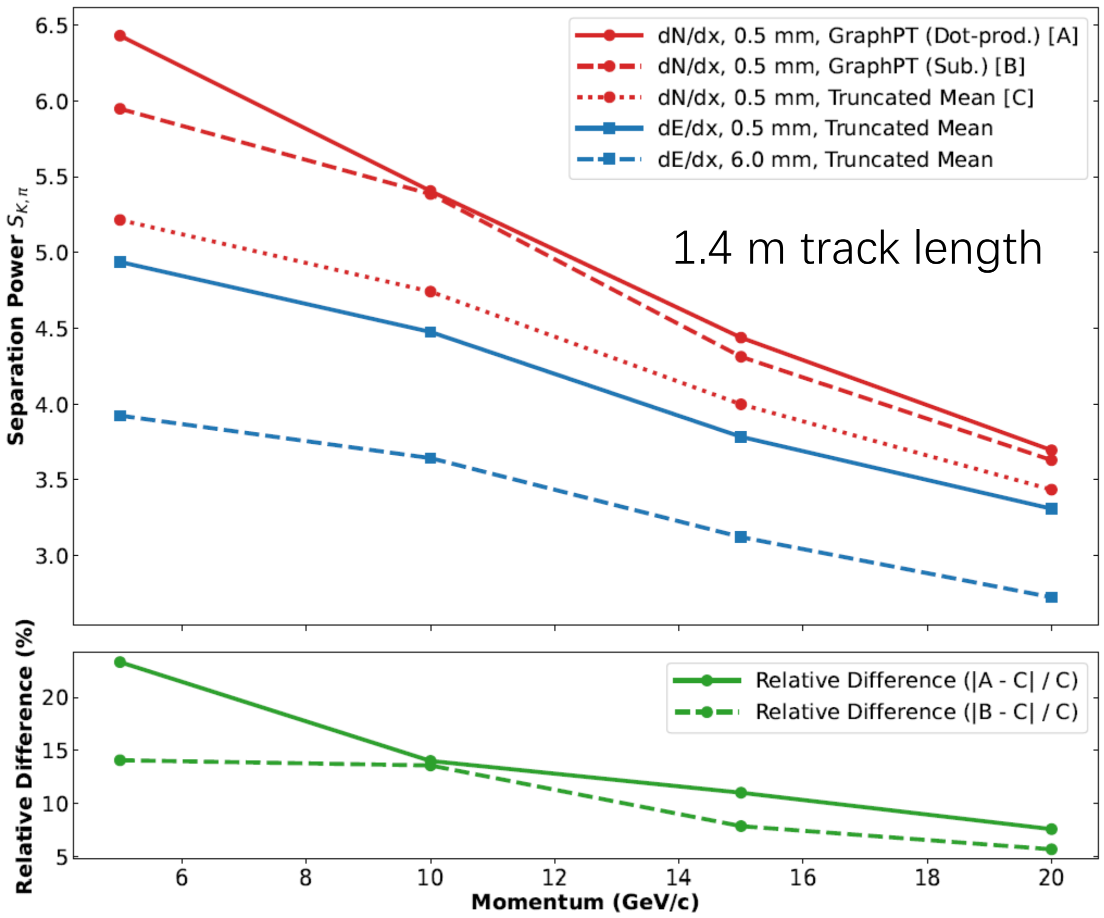
- **Conclusions:**
  - GraphPT has lower precision but much higher recall ➔ Much improved signal efficiency: **~60% improvement**
  - GraphPT has much better accuracy and F1 score: **~20% improvement**
  - The GraphPT with dot-product attention achieve the overall best classification power

Table 3. Classification metrics.

Method	Accuracy	Precision	Recall	F1-Score
Truncated Mean	0.601	0.743	0.574	0.648
GraphPT (Sub.)	0.698	0.689	0.960	0.802
GraphPT (Dot-prod.)	**0.707**	**0.702**	**0.941**	**0.804**

PID performances

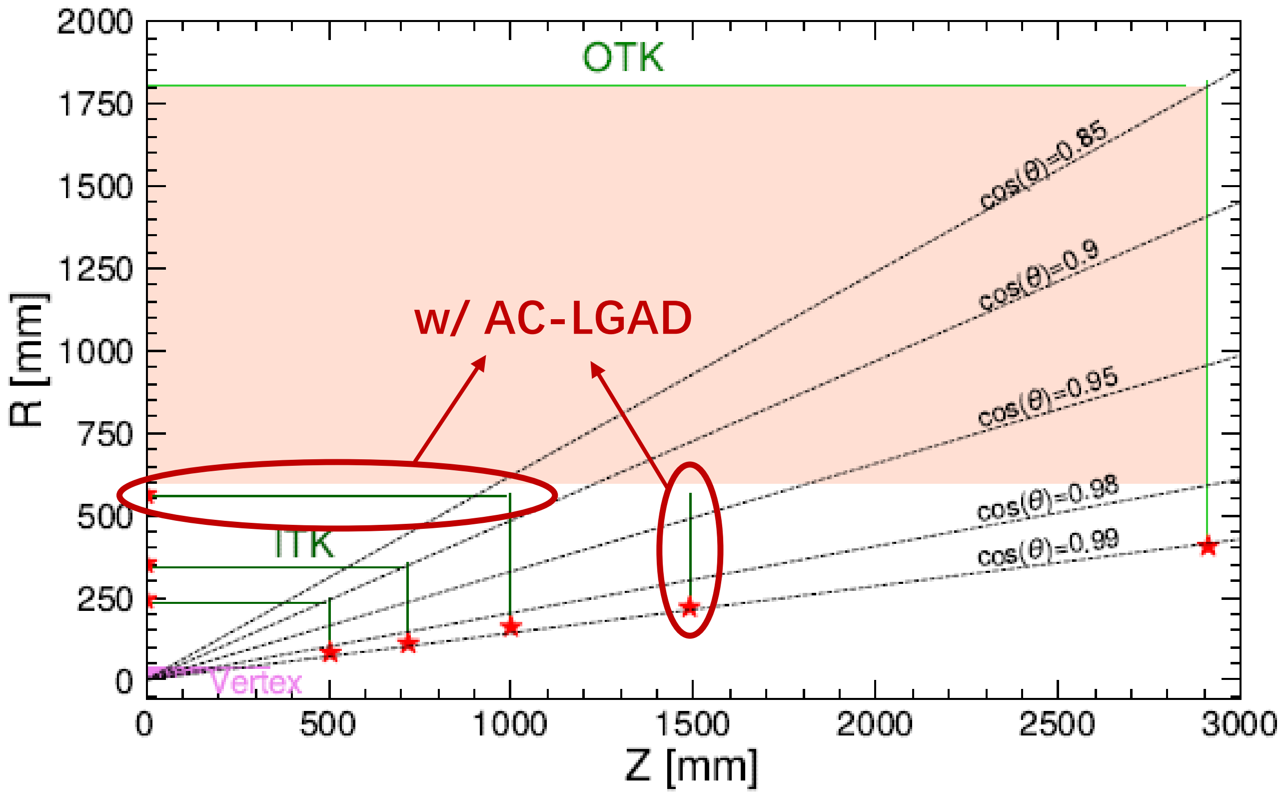
K/π separation power



- **Metrics:**
  - K/π separation power within 5-20 GeV/c
- **dN/dx (vs. truncated mean):**
  - **GraphPT (sub.): 5-15% improvement**
  - **GraphPT (dot-prod.): 10-20% improvement**
- **Overall:**
  - All dN/dx results outperform dE/dx
  - Up to 50% improvement compared with traditional 6mm-pad dE/dx

Time-of-flight in ITK

ToF in ITK

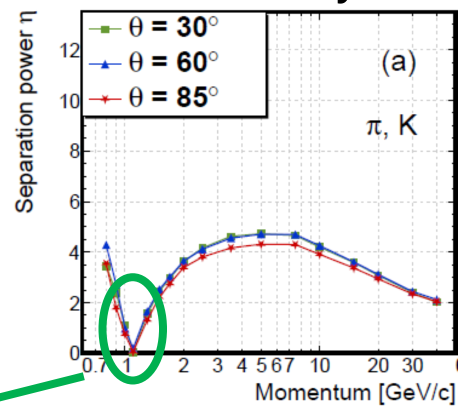


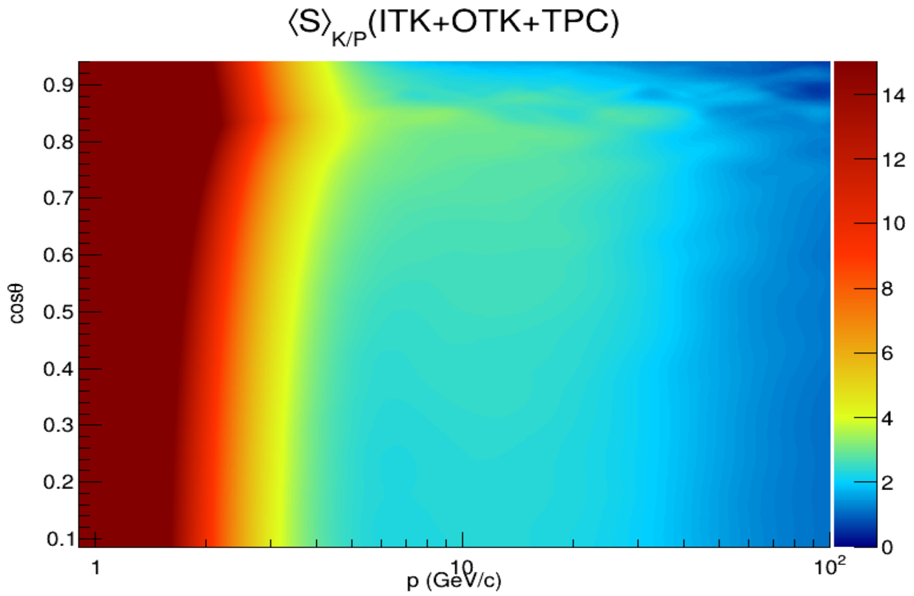
- **Besides OTK, add ToF at the outermost layer of ITK:**
  - Pixel AC-LGAD: σ = 30 ps
- **P_T coverage:**
  - OTK: > 0.81 GeV/c
  - **ITK: > 0.25 GeV/c**

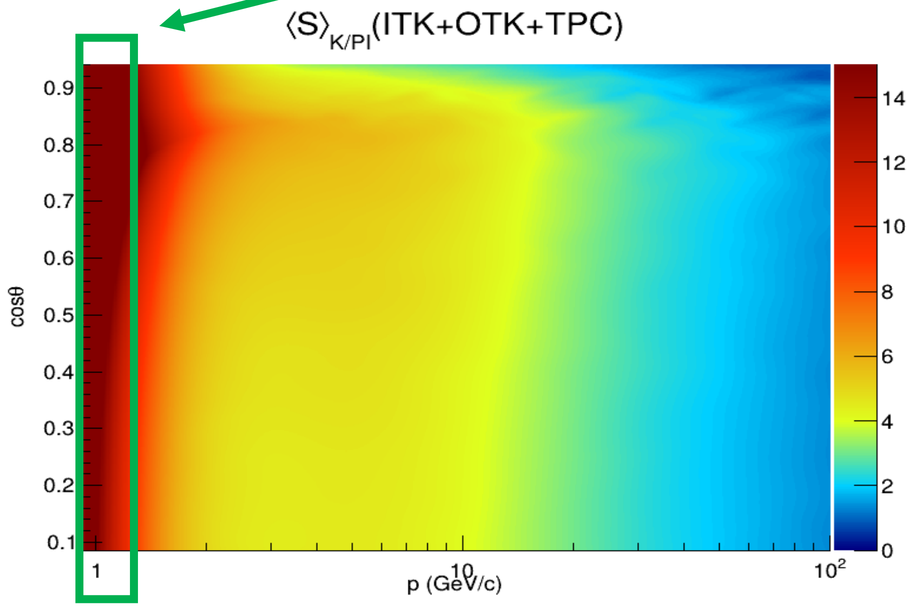
Combined separation power

- **Definition:** $S_{comb} = \sqrt{S_{TPC}^2 + S_{OTK}^2 + S_{ITK}^2}$
  - Note: dN/dx by truncated mean
- **The low-momentum gap in separation power by TPC is covered by ToF**

TPC only



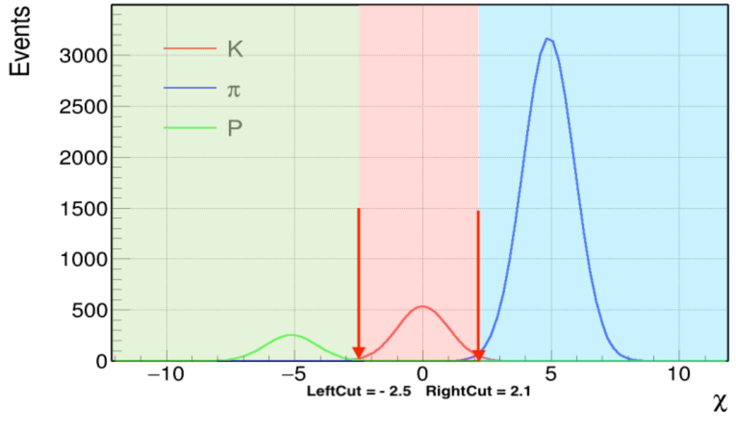




Unified PID methodology

Calculate combined $\chi^2 = \chi^2_{TPC} + \chi^2_{OTK} + \chi^2_{ITK}$

where $\chi_D = \frac{\mu_{D,meas} - \mu_{D,exp}}{\sigma_D}\ (D = TPC, OTK, ITK)$



PID: According to χ-region defined by LeftCut and RightCut.

Store optimal LeftCut and RightCut into a lookup table

Efficiency: ϵ

Purity: p

Optimize LeftCut and RightCut within (p, |cosθ|) bins subject to max. $\epsilon \times p$

Kaon ID results

Catagory	efficiency	purity	eff*pur
a	0.9932	0.1427	0.1417
b	0.9904	0.2573	0.2548
c	0.9680	0.8743	0.8463
d	0.9938	0.9241	0.9184

Control sample: $Z \to q\bar{q}$

```
a: TPC
b: TPC + OTK
c: TPC + OTK + ITK
d: piecewise TPC + OTK + ITK
   (TPC+OTK+ITK, p < 1GeV/c
    TPC+OTK    , 1GeV/c < p < 3GeV/c
    TPC        , p > 3GeV/c)
```

Remarks:

- a: Limited PID in the TPC around 1 GeV/c; the sample is imbalanced (dominated by pions)
- b: The OTK improves PID but is not efficient enough due to the large radius
- c: The ITK improves PID significantly, due to improvements from low momentum particles
- d: A piecewise combine strategy can exploit the available information more effectively

Conclusions

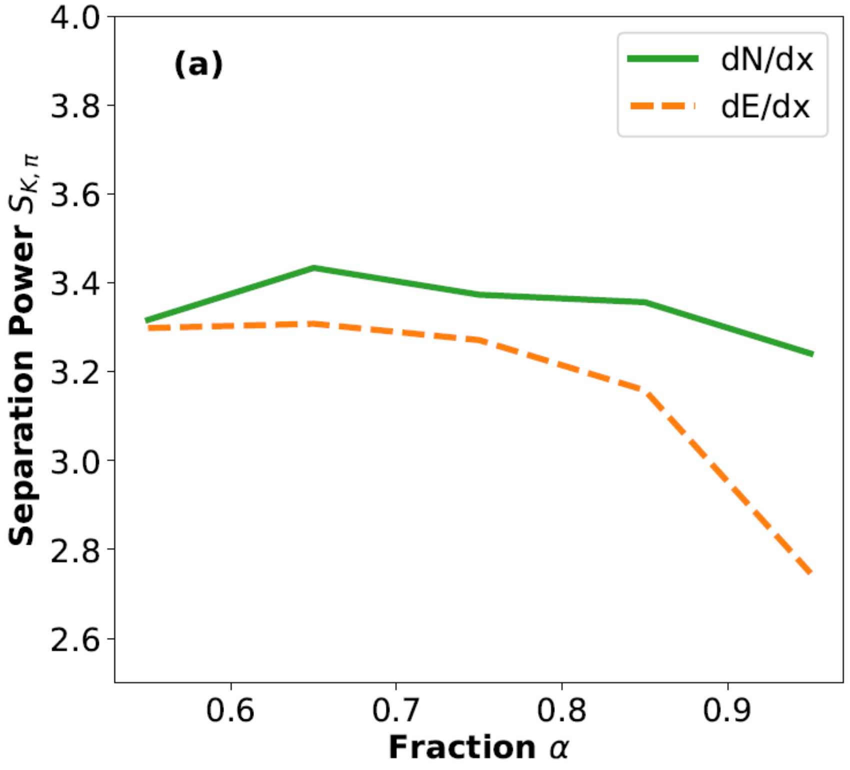
- **Hadron identification is essential for CEPC physics. Both dN/dx and time-of-flight have made significant progress.**

- **A point-cloud transformer-based dN/dx reconstruction method has been developed. Key improvements over traditional methods include:**
  - Signal efficiency: ~60% improvement
  - Accuracy/F1 score: ~20% improvement
  - K/π separation power: ~10-20% improvement

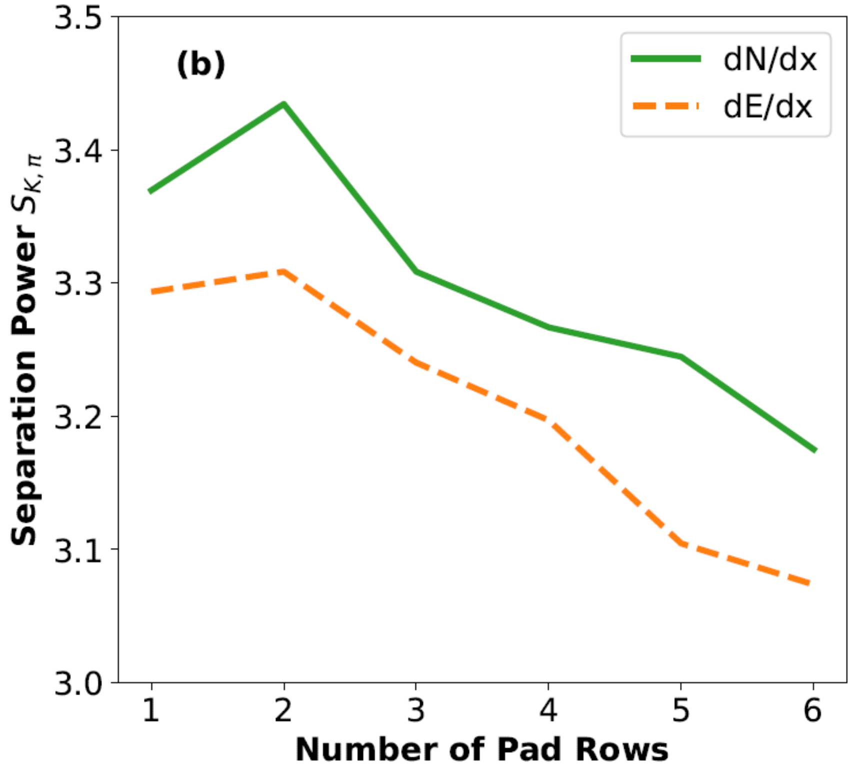
- **Time-of-flight in the ITK can significantly enhance PID, especially for low-momentum particles, achieving an efficiency x purity of ~91.8%**

Backup slides

Optimization of the truncated mean



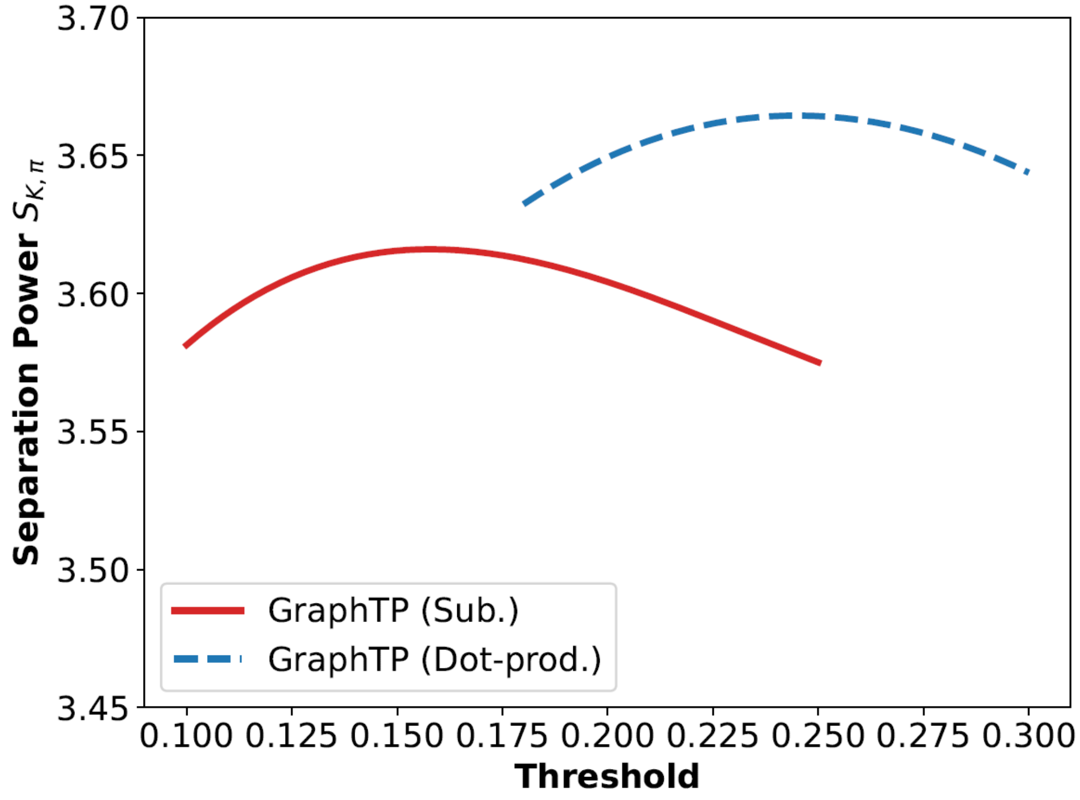
Optimal value: 0.65



Optimal value: 2

Combine layers to reduce fluctuations for small pads

Parameter optimization for PID



- ✓ **Figure of merit: K/π separation power**
- ✓ **Variable: Threshold τ for dN/dx calculation**
  - ✓ $dN/dx = \frac{1}{L}\sum_{i=1}^{N} \mathbf{1}_{\{p_i>\tau\}}$
- ✓ **Optimal thresholds:**
  - ✓ Subtraction operator: 0.157
  - ✓ Dot-product operator: 0.261

Kaon ID results

```
----- PID Summary -----
Default sigma:          1
Efficiency (K):         0.92768
Purity (K):             0.863539
Efficiency * Purity:    0.801088
```

With $p > 4$ GeV/c and $|\cos\theta < 0.95|$ cuts