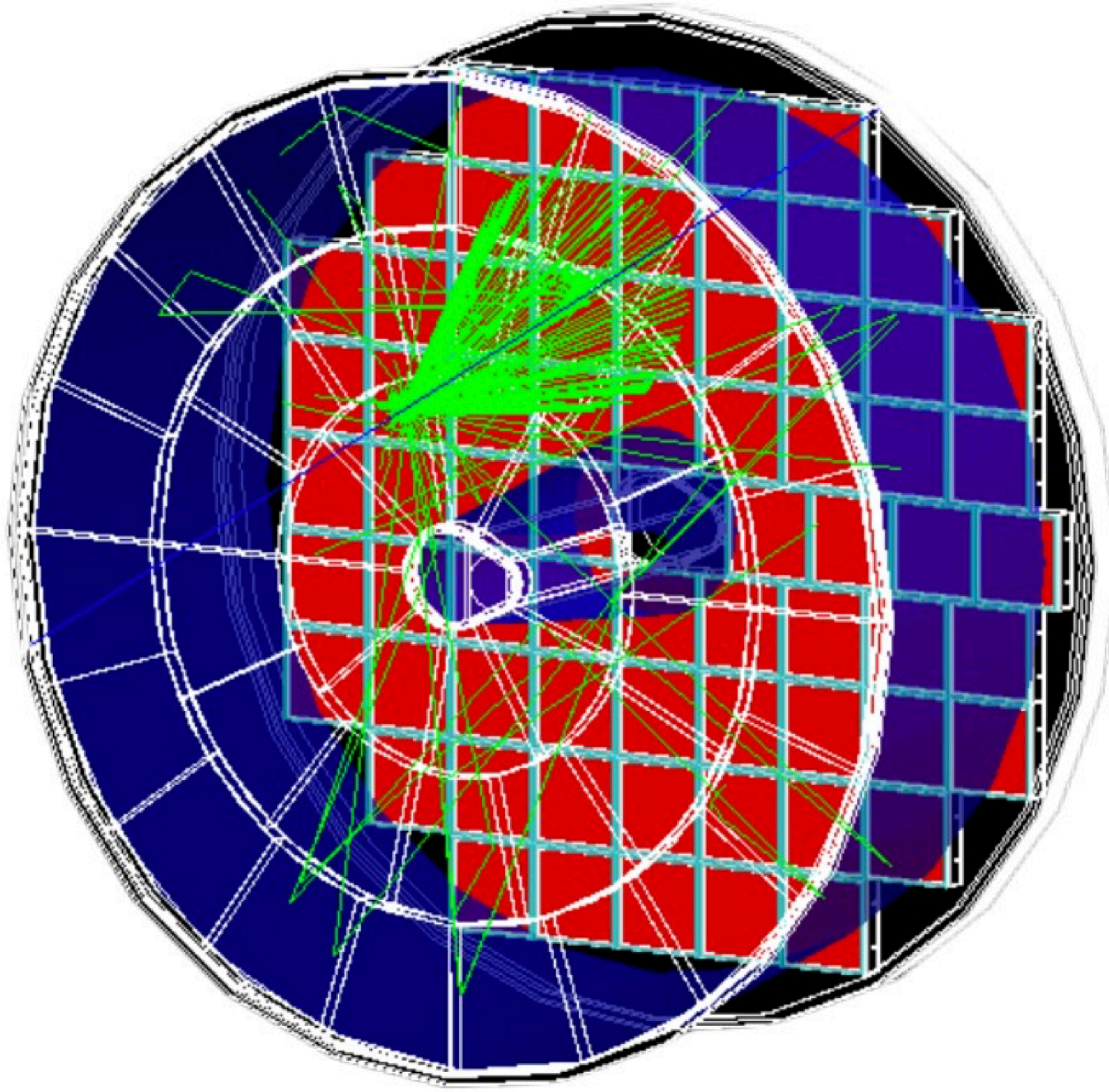
Enhancing PID separation at ePIC using machine learning



Charles Joseph Naïm

January 23, 2025

Stony Brook University

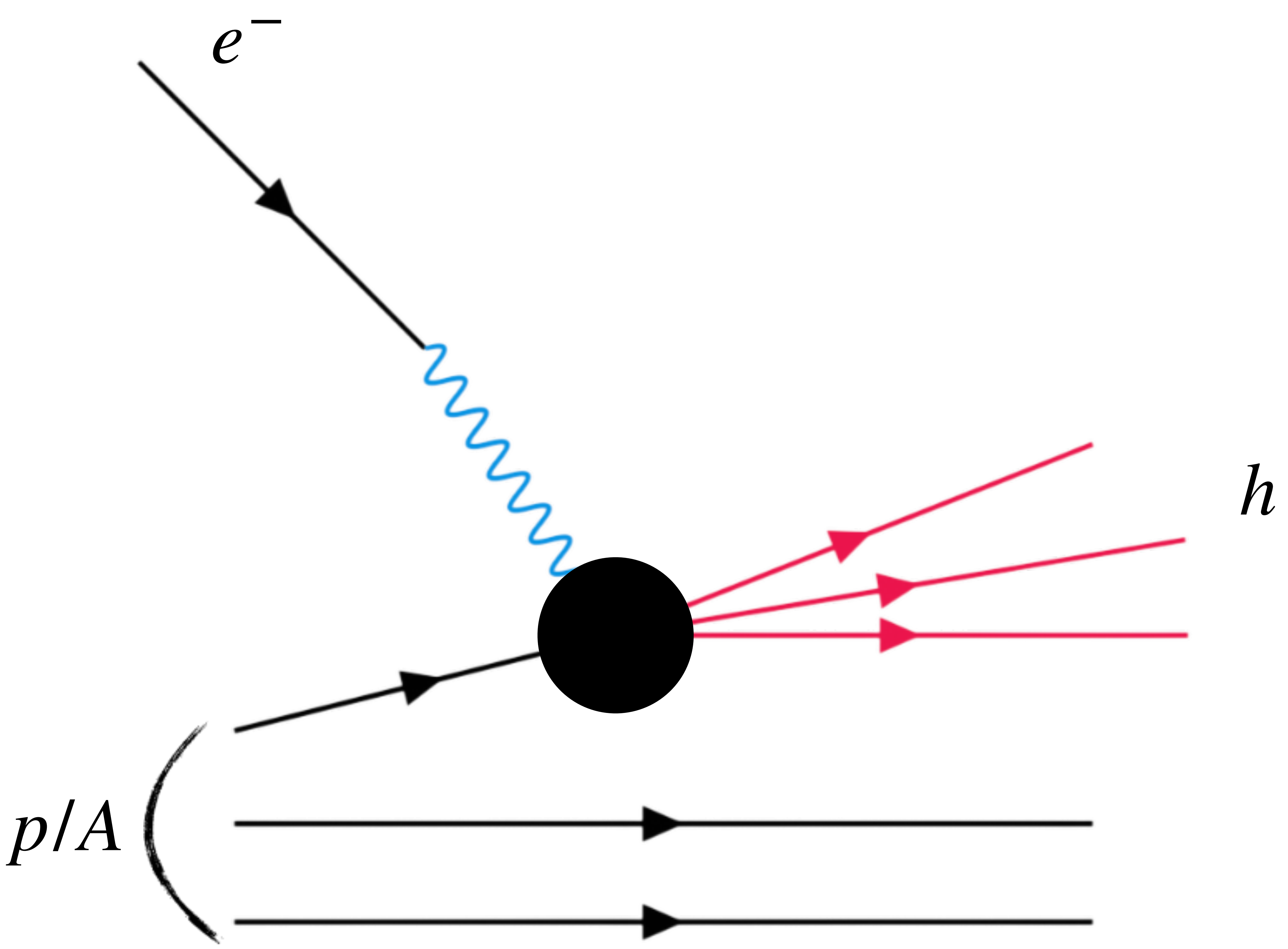
Center for Frontiers in Nuclear Science

The ePIC Detector at the EIC

pfRICH subsystem in backward region

- **A compact central detector** with several subsystems
- **Hermetic coverage**: $-3.5 < \eta < 3.5$ (tracking, calorimetry, **particle identification**)

Physics Motivations at the EIC



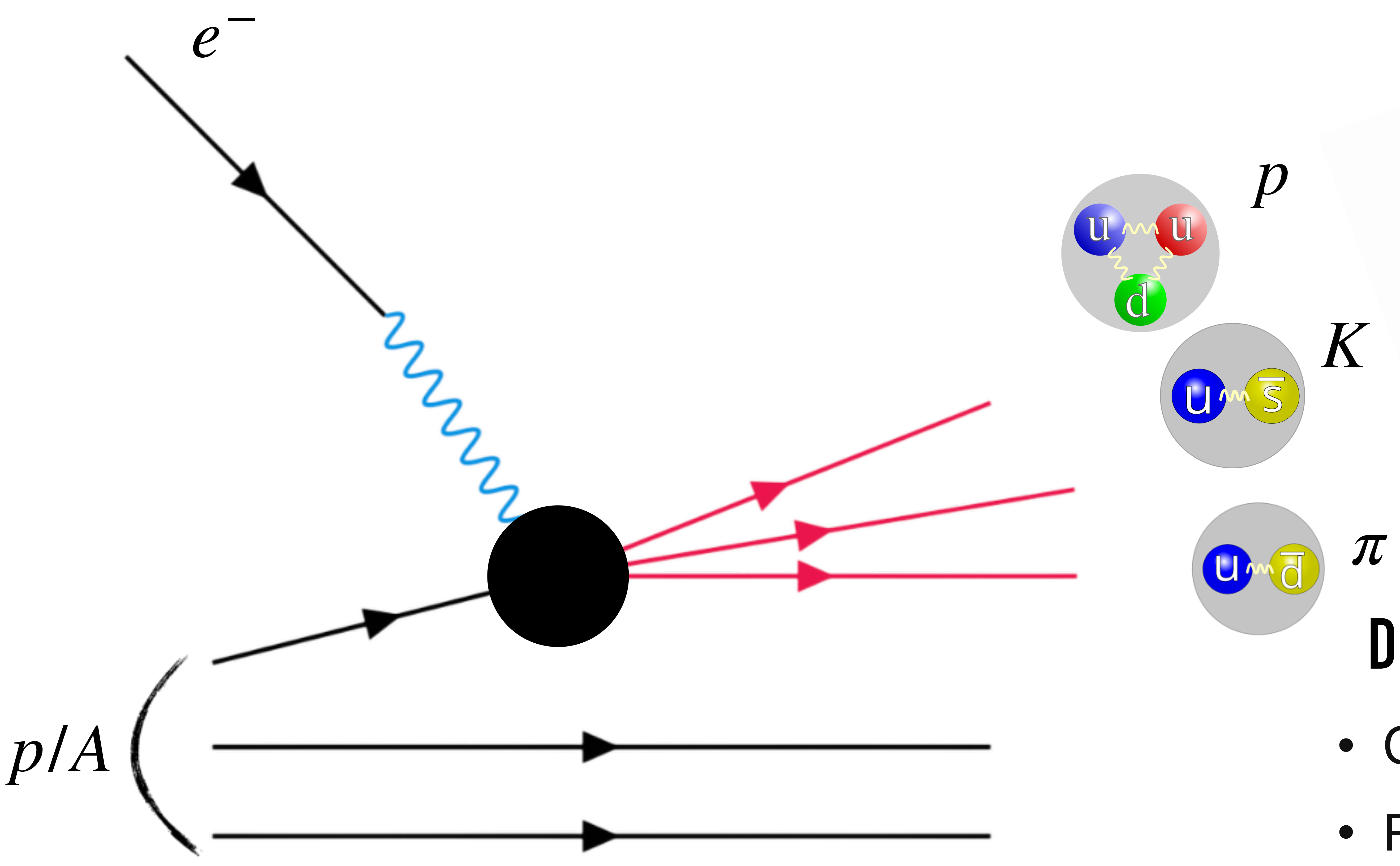
- **Semi-Inclusive Deep Inelastic Scattering**
- Production of hadrons in final-state
- Provide information on:

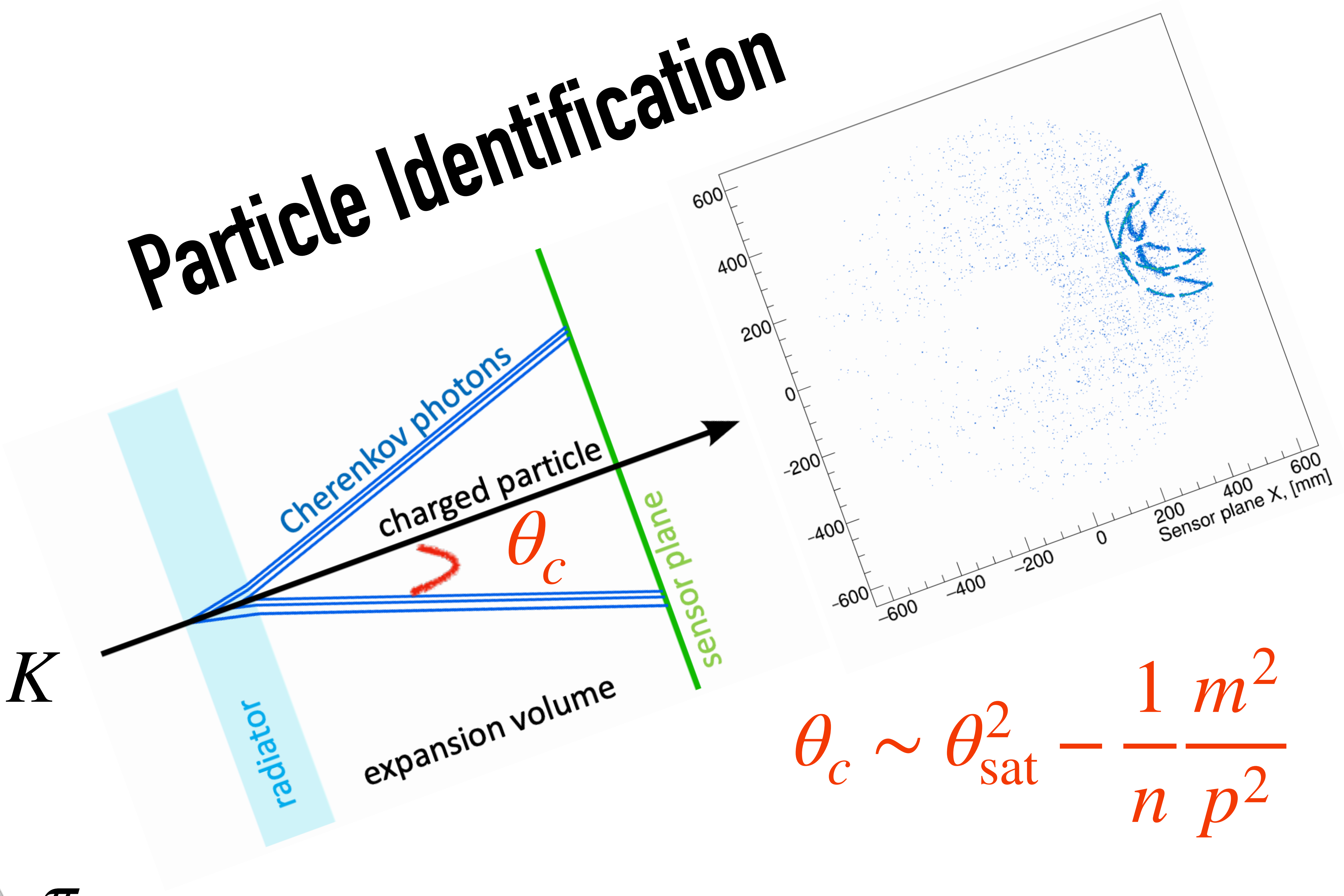
→ the ***fragmentation process (hadronization)***

→ the ***hadronic structure***

Particle Identification detectors are crucial

The pfRICH Concept





$$\theta_c \sim \theta_{\text{sat}}^2 - \frac{1}{n}\frac{m^2}{p^2}$$

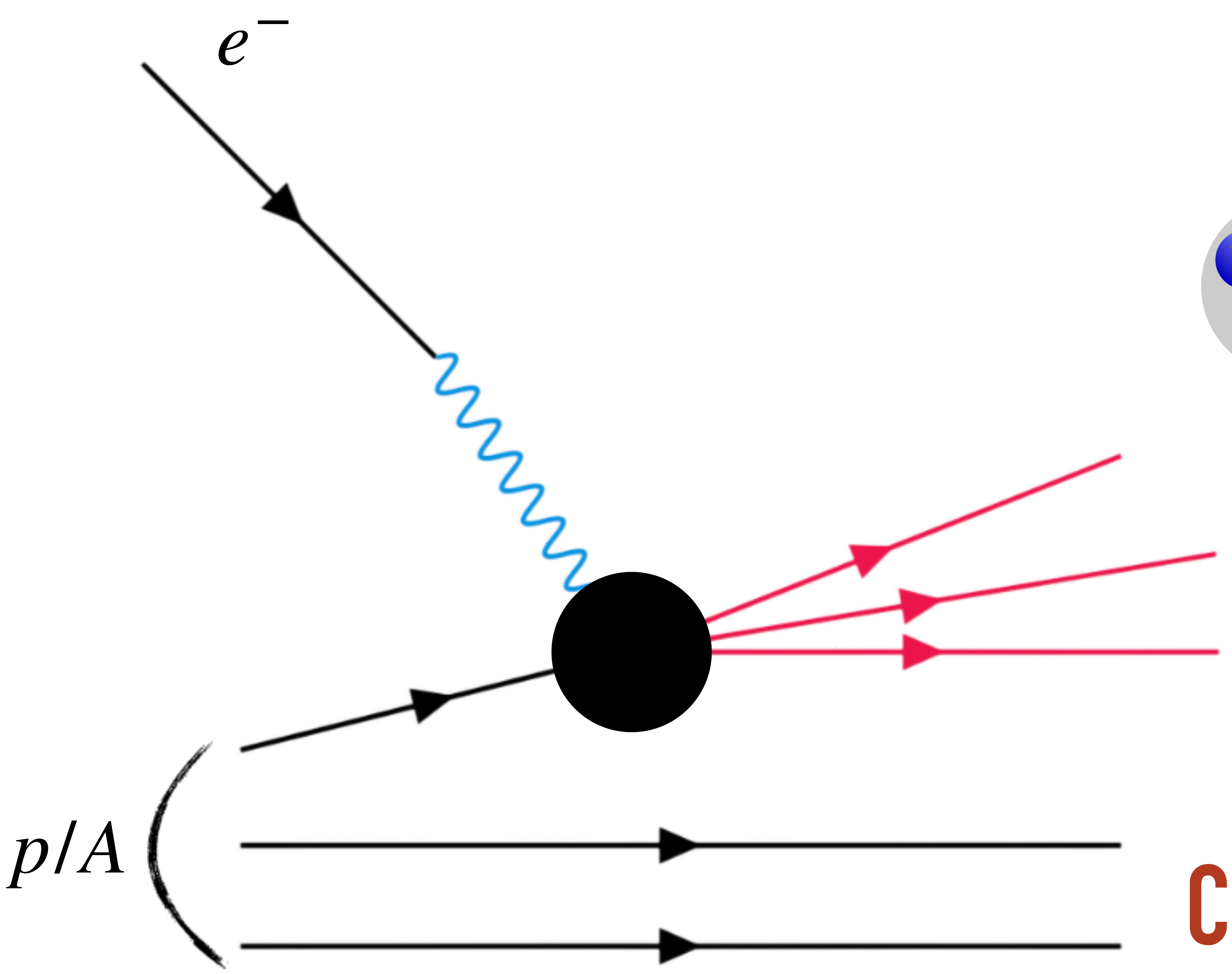
The pfRICH will provide > 3σ π/K separations for momentum up to 7 GeV/c for –3.5 < η < -1.5

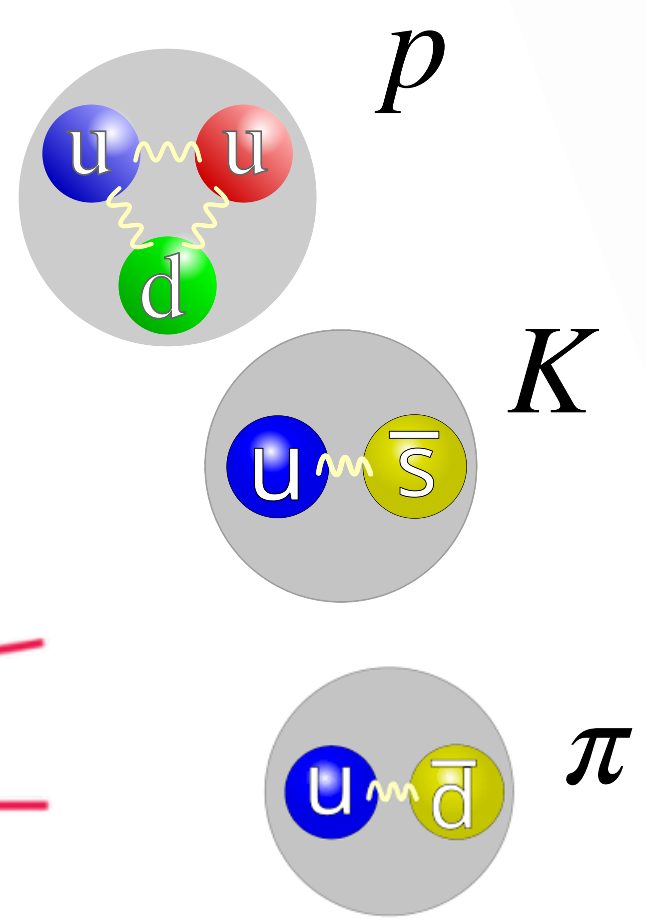
Detection Principle

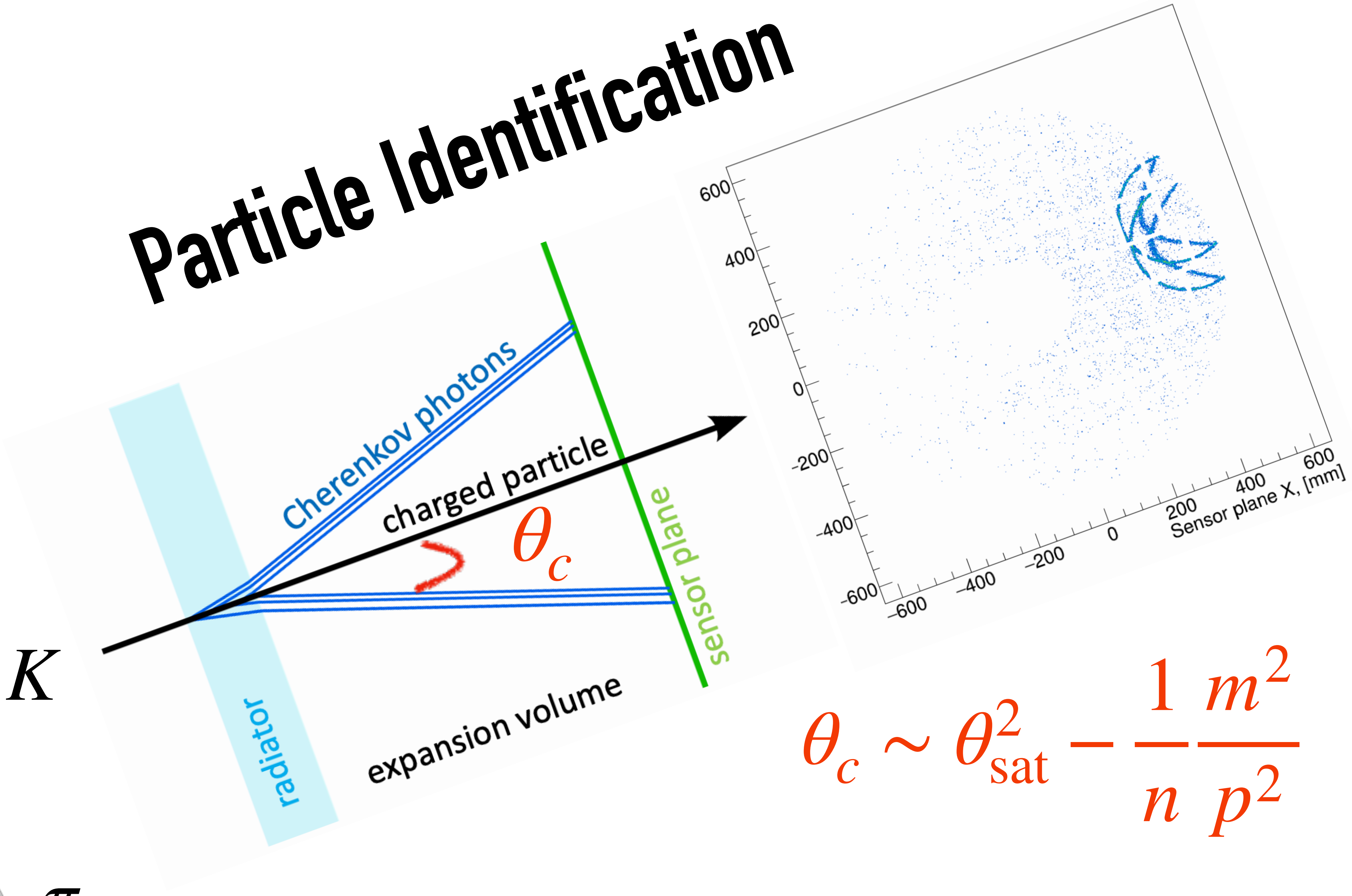
- Charged particle → emits Cherenkov photons at angle θ_c
- Photons project onto photodetectors → form a **ring**

  → **Ring radius** $\propto \tan\theta_c$
- Measuring ring size → deduce θ_c → **particle mass**

The pfRICH Concept



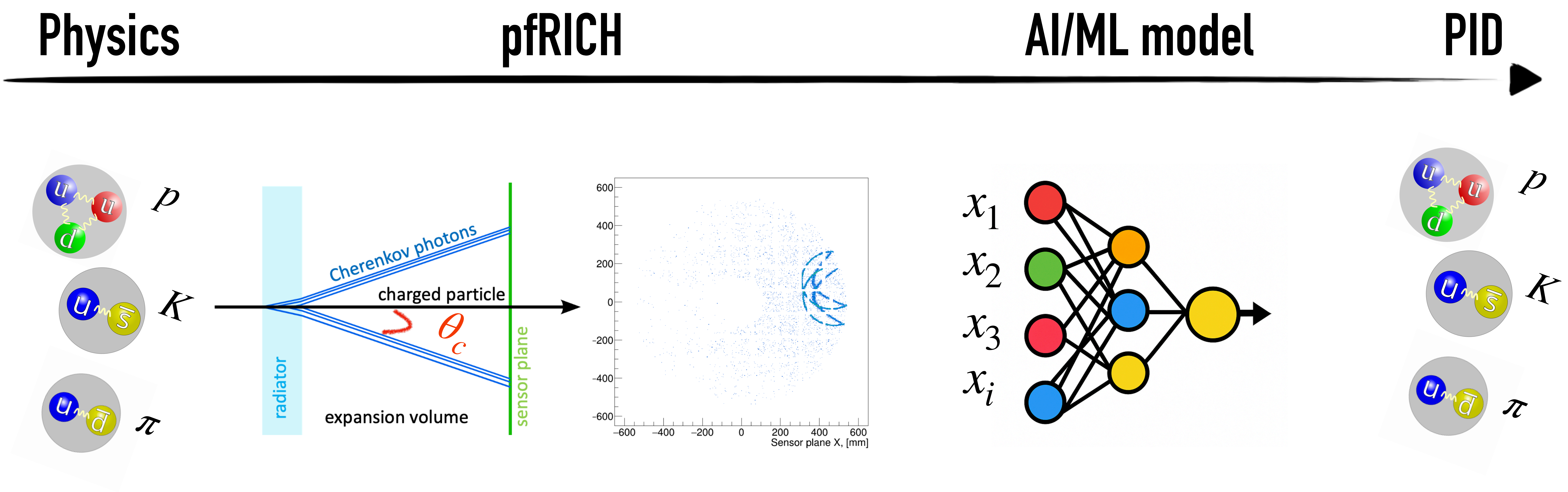




$$\theta_c \sim \theta_{\mathrm{sat}}^2 - \frac{1}{n}\frac{m^2}{p^2}$$

Can we use machine learning to improve particle identification?

The Approach

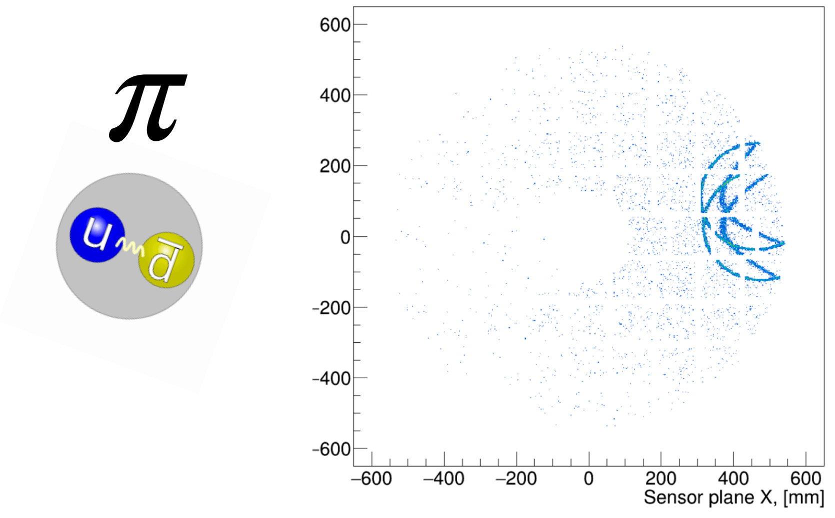


Can we use machine learning to improve particle identification?: Yes!

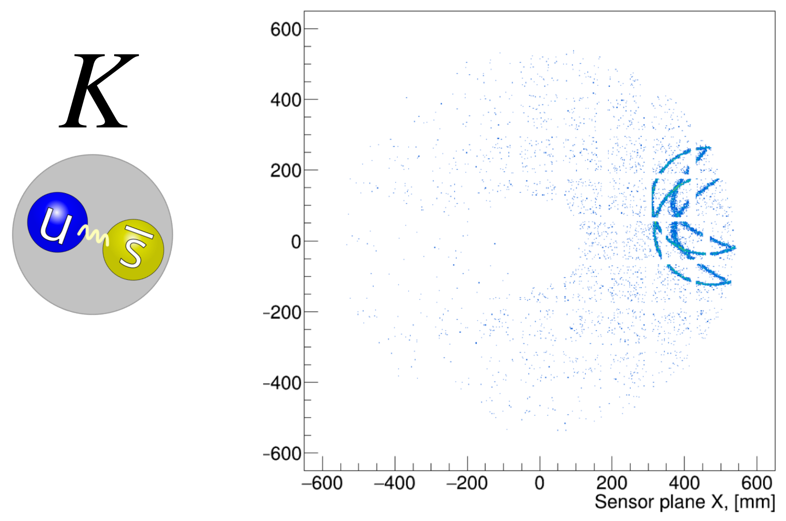
An ideal use case for AI/ML, since the signal is well defined and fully understood

Model Training

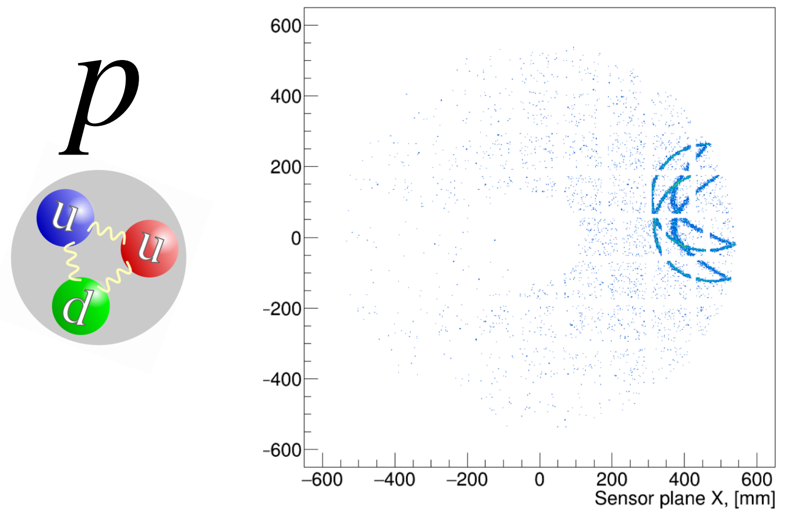
π



K

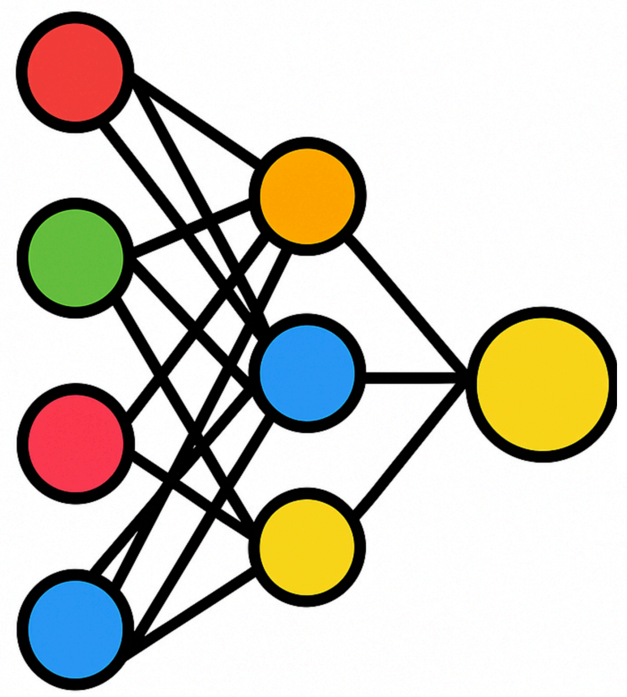


p



Training

x_1
x_2
x_3
x_i



AI/ML model

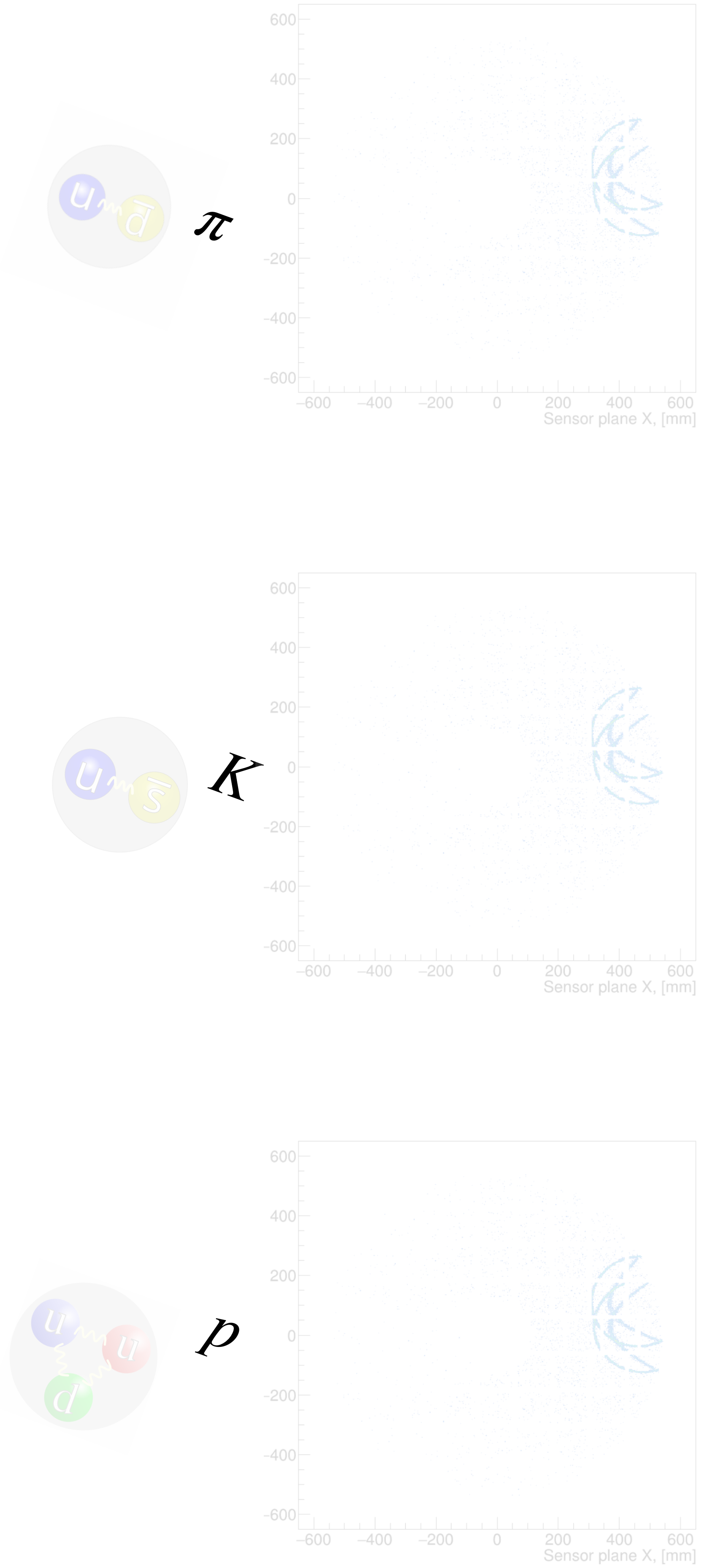
$$f\left(\sum_i w_i x_i + b\right)$$

- Pattern Recognition
- XGBoost - Gradient-boosted

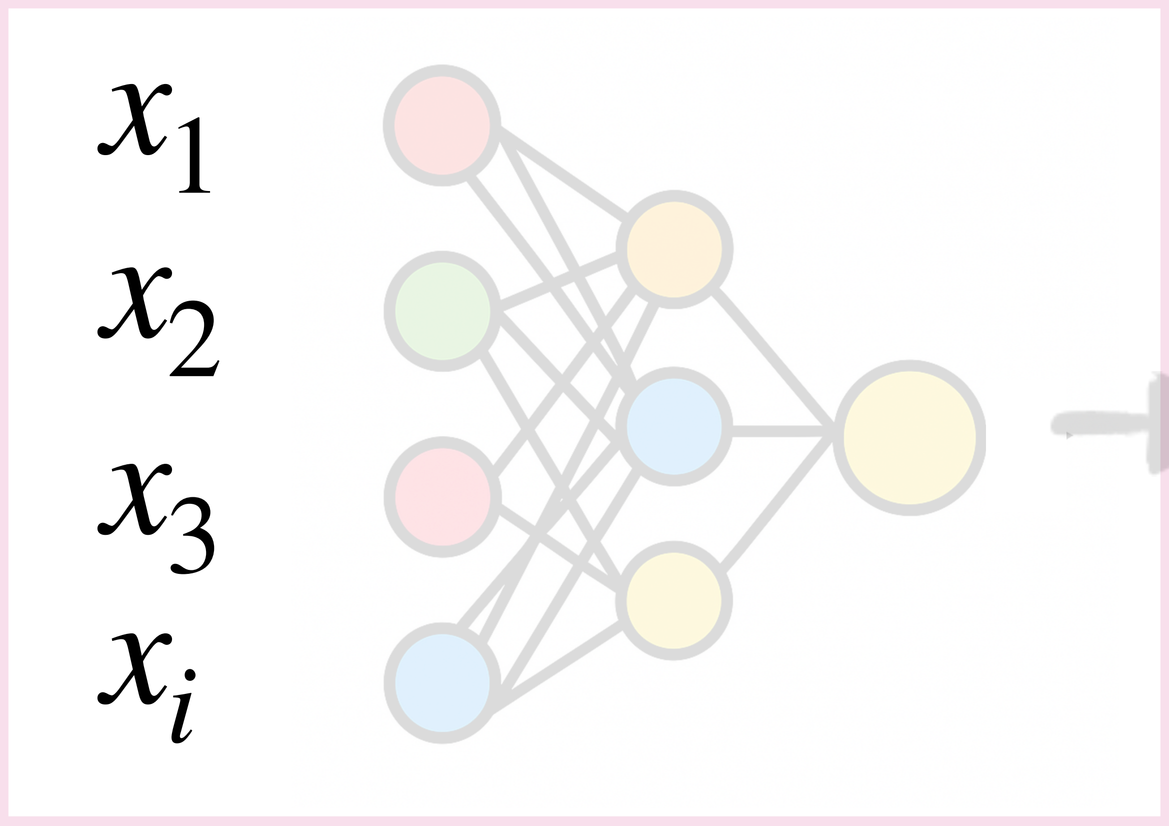
Standalone ePIC pfRICH GEANT4:

- Timing, hits position, momentum …
- More (good) data → better training

Model Inference



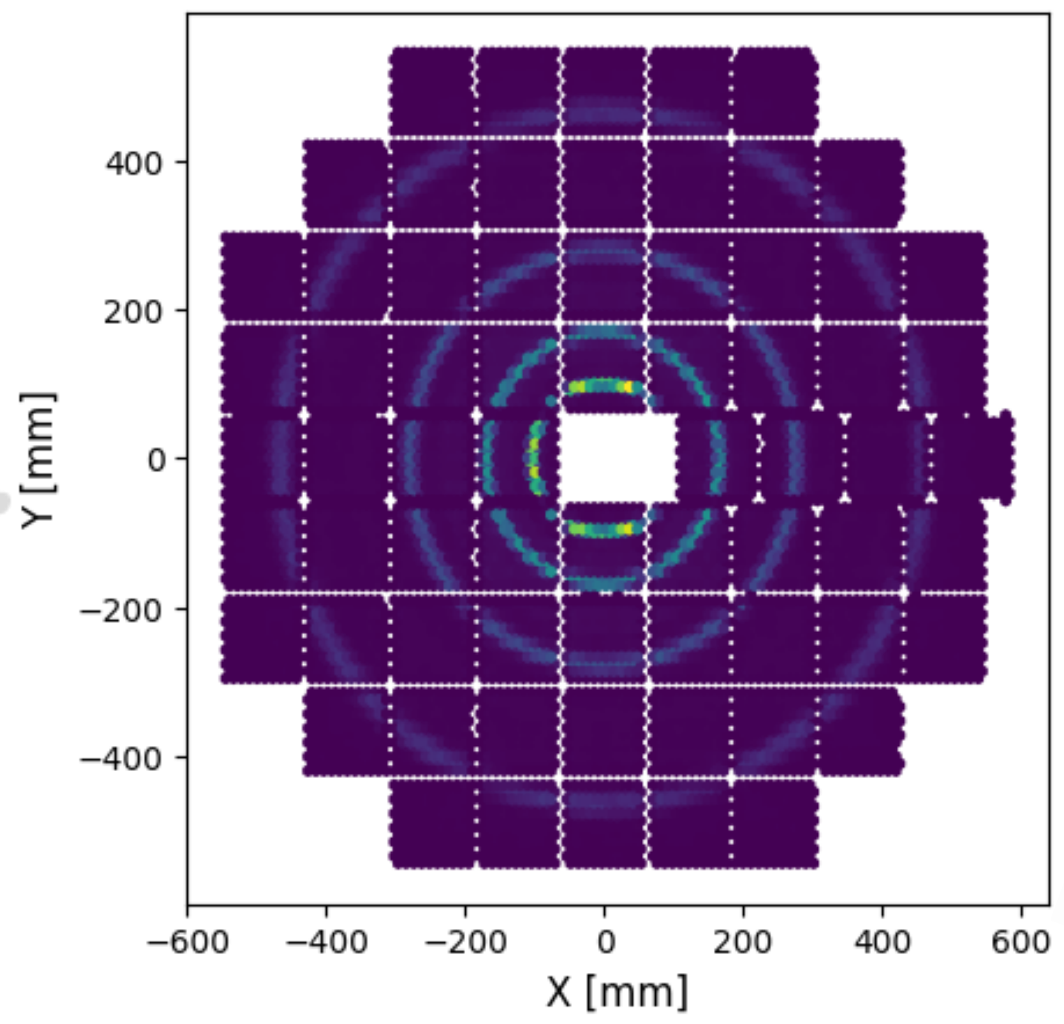
Training

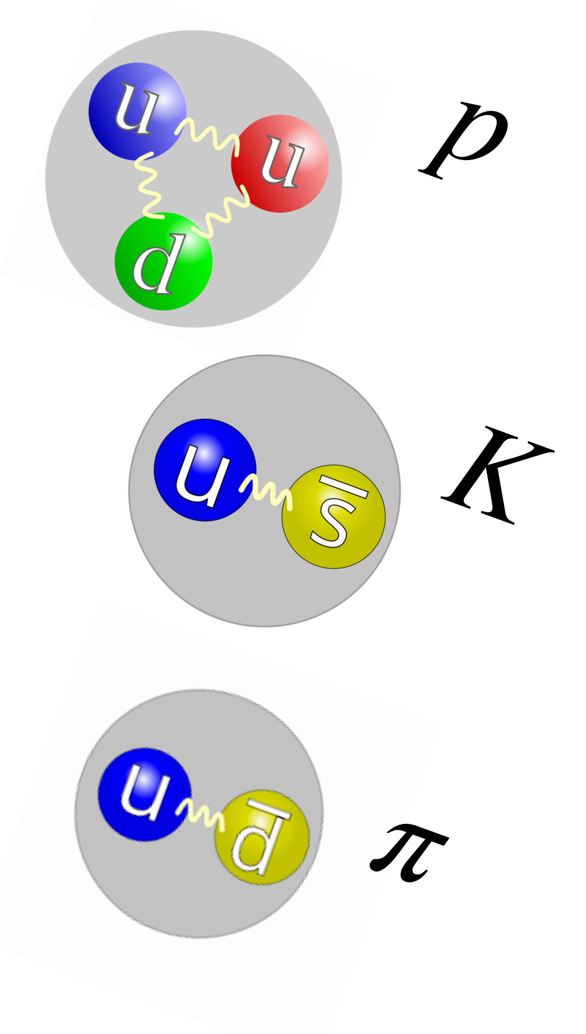


AI/ML model

$$f\left(\sum_i w_i x_i + b\right)$$

Data





Particle Identification

Model Steps

Simulation

→ Generate particle events with Cherenkov photons

Data Preprocessing

→ Convert hit patterns to 64 x 64 images

CNN Encoder

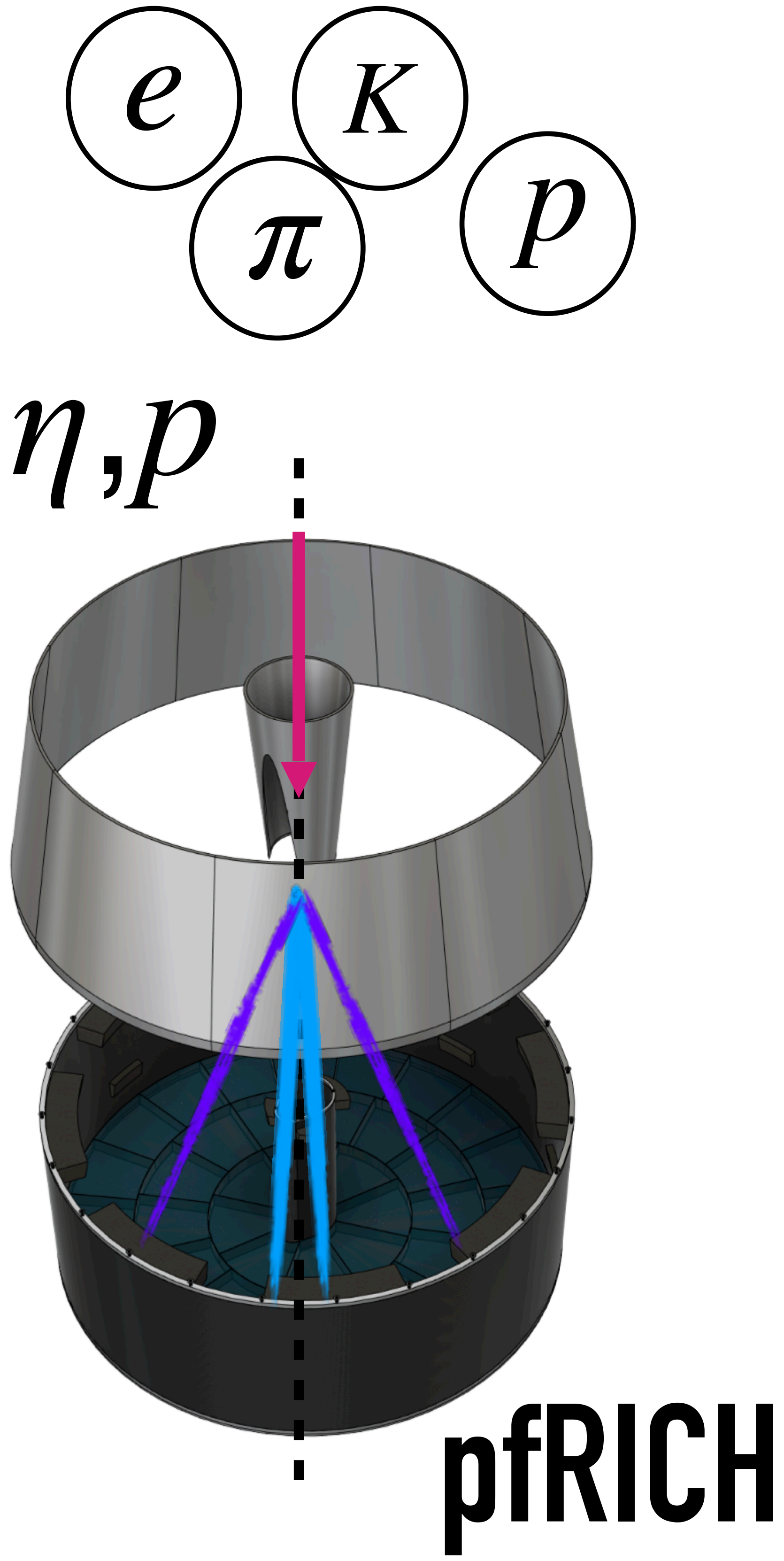
→ Extract spatial features from ring patterns

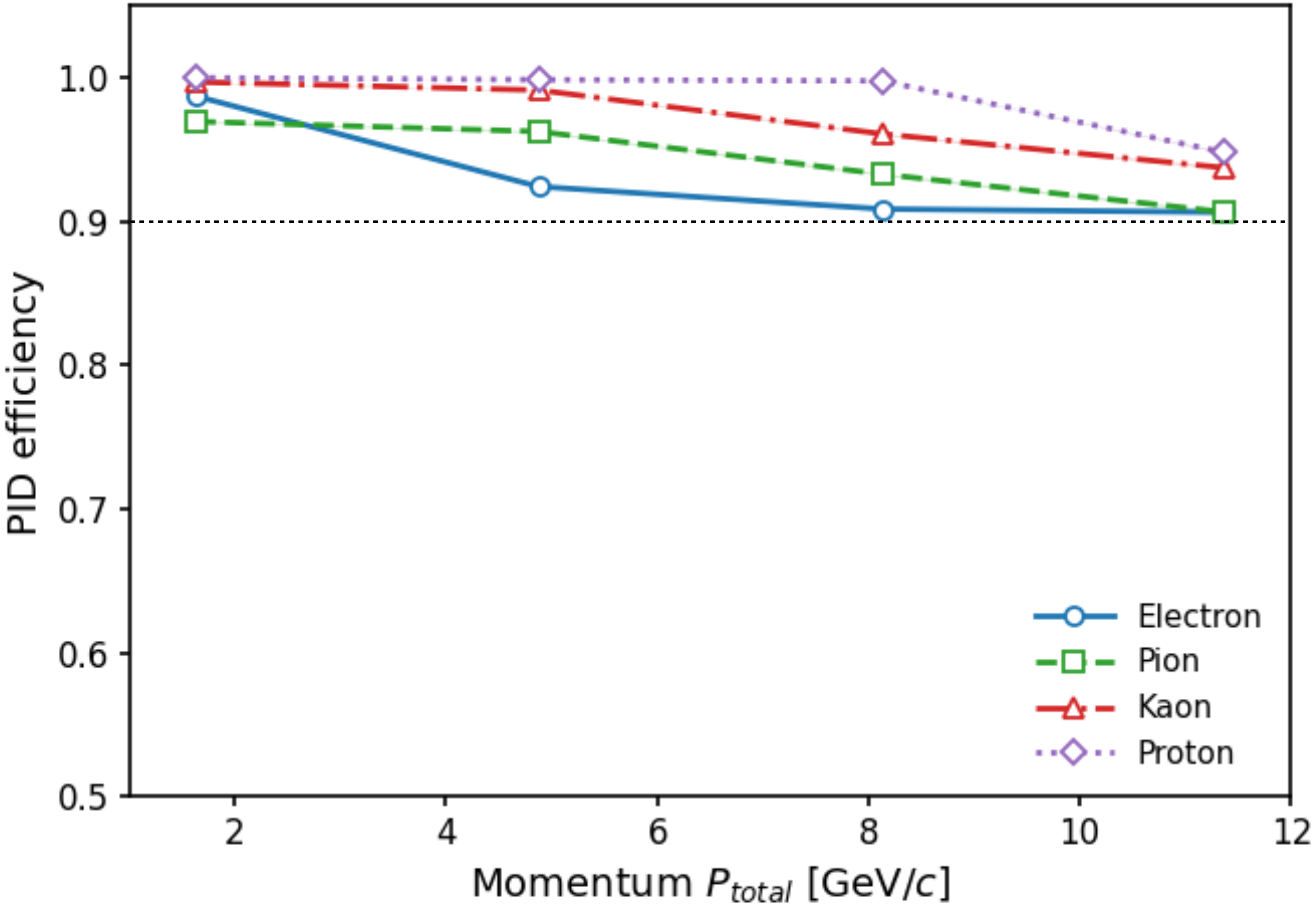
XGBoost Classifier

→ Combine CNN features with kinematics for PID

Hybrid model

Separation Efficiency

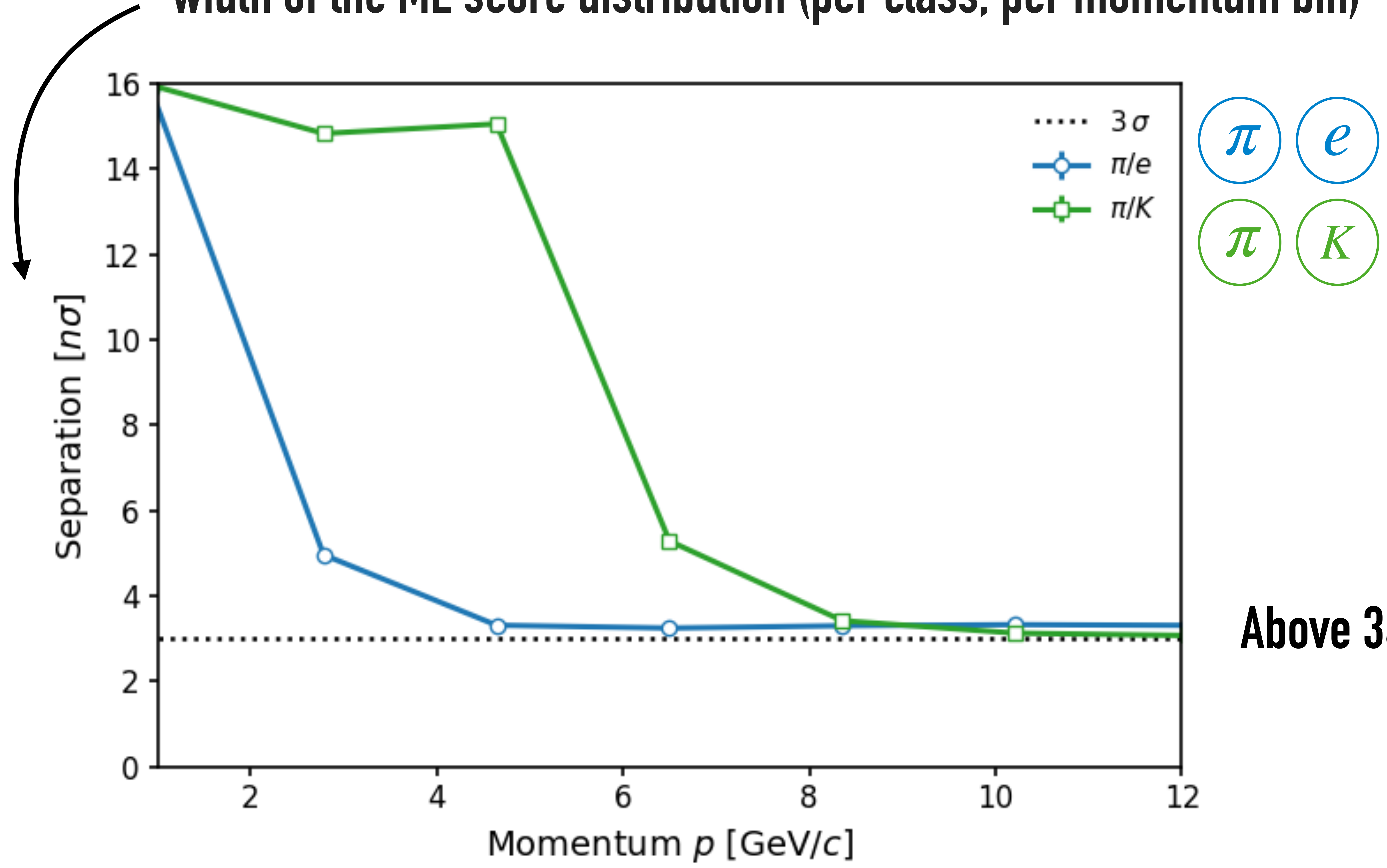




Above 90%

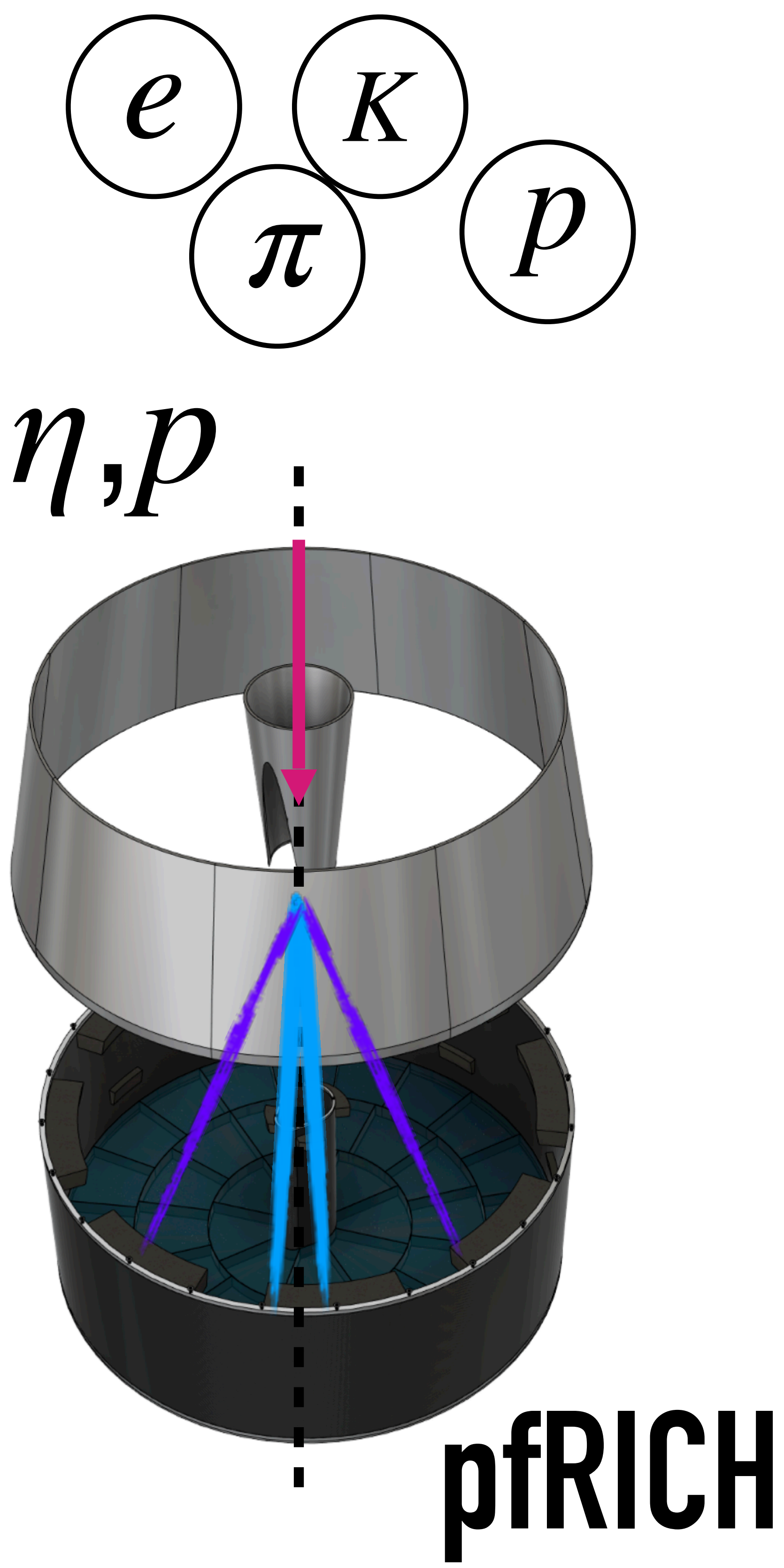
PID Separation

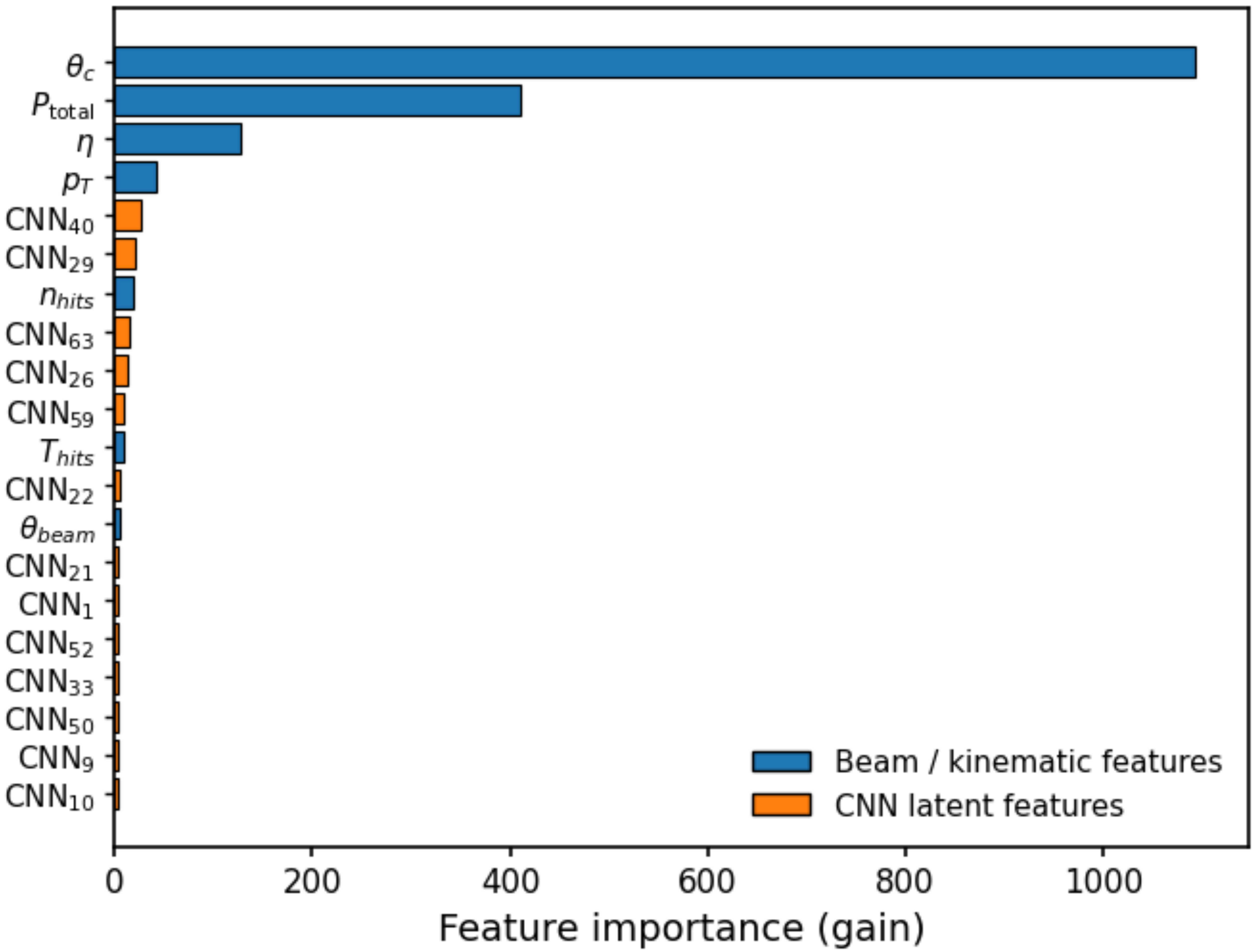
width of the ML score distribution (per class, per momentum bin)



Above 3σ

Feature importance



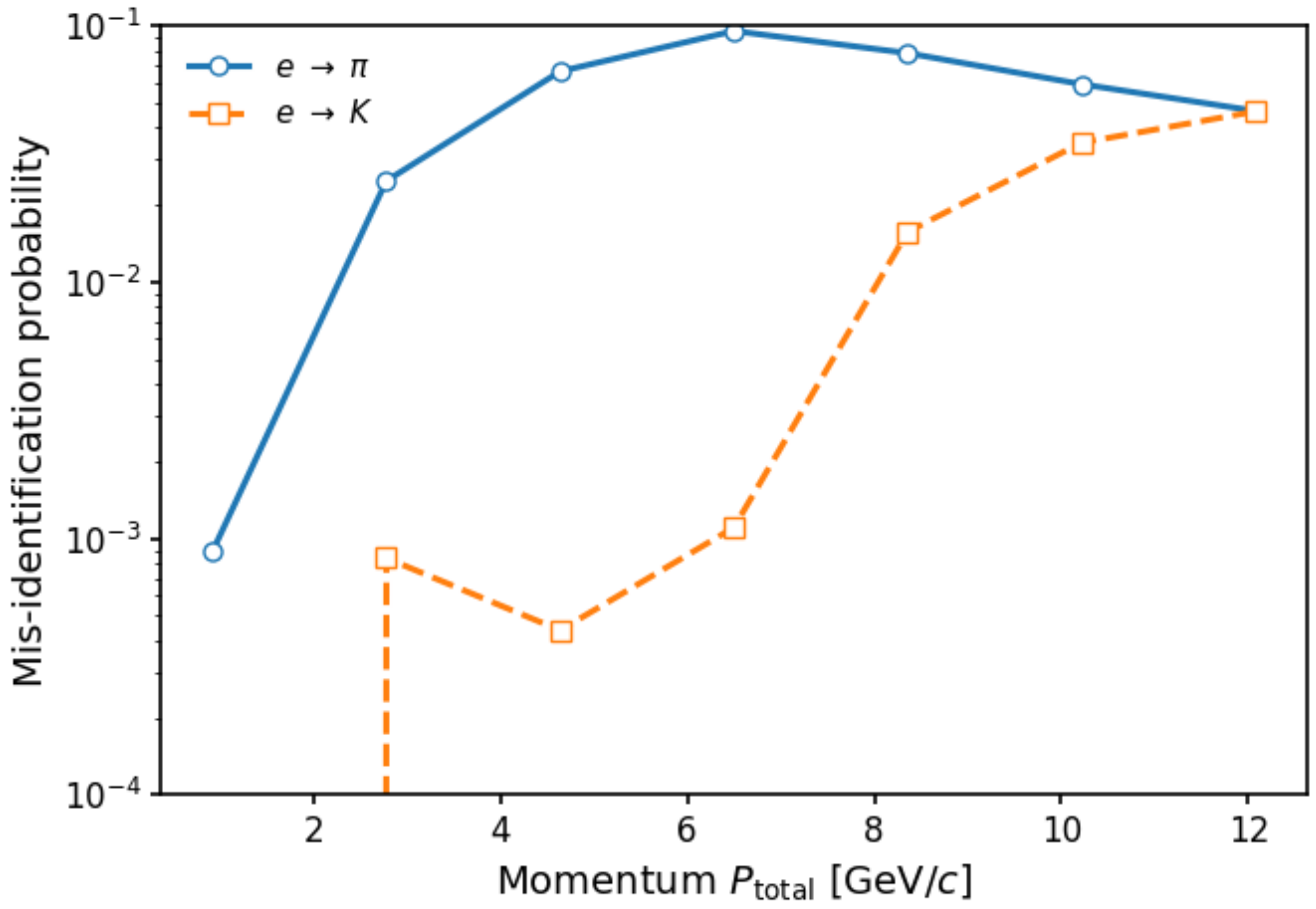


Kinematics info. is dominant

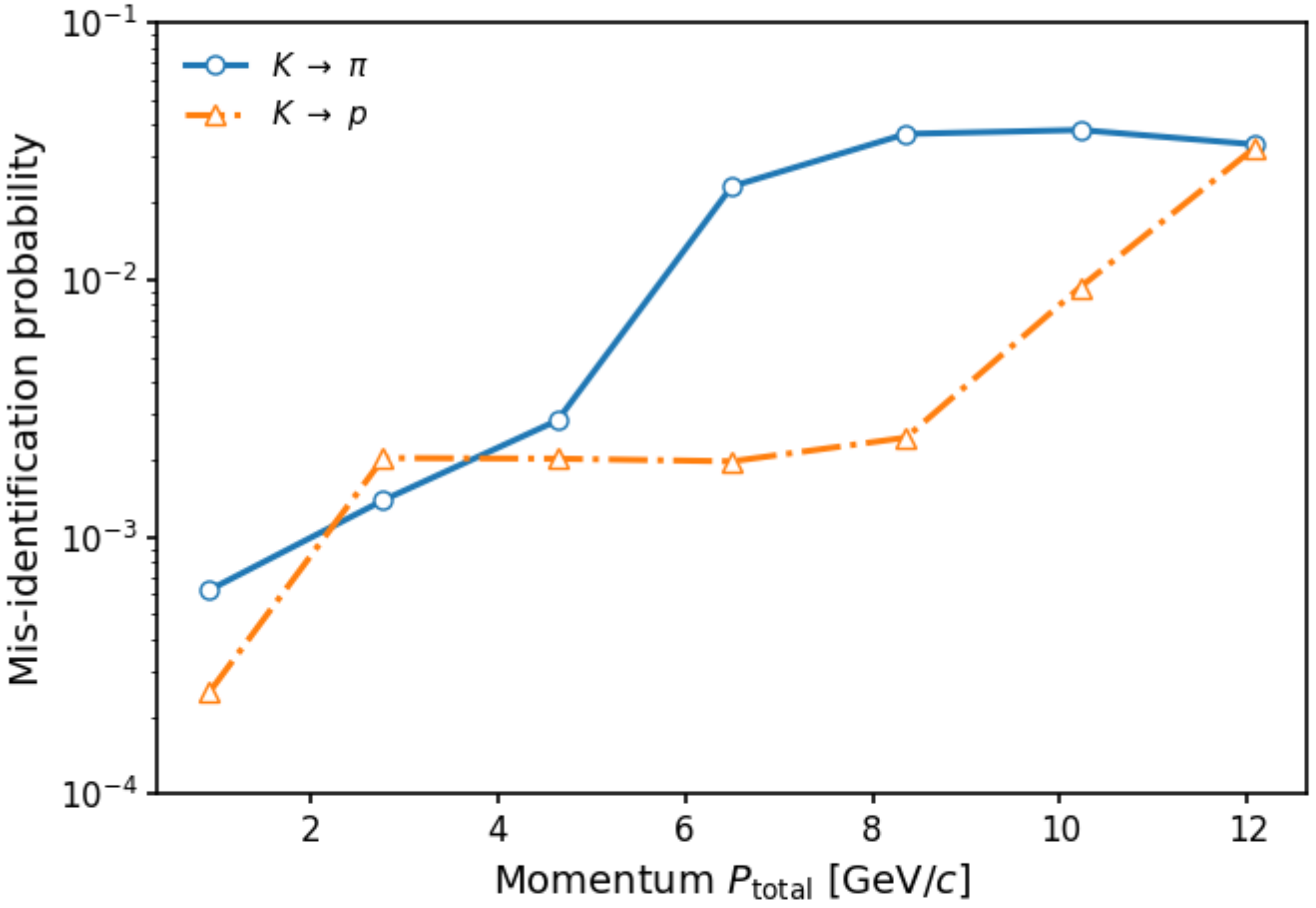
CNN info plays a non-neglieable role

Misidentification

Electron

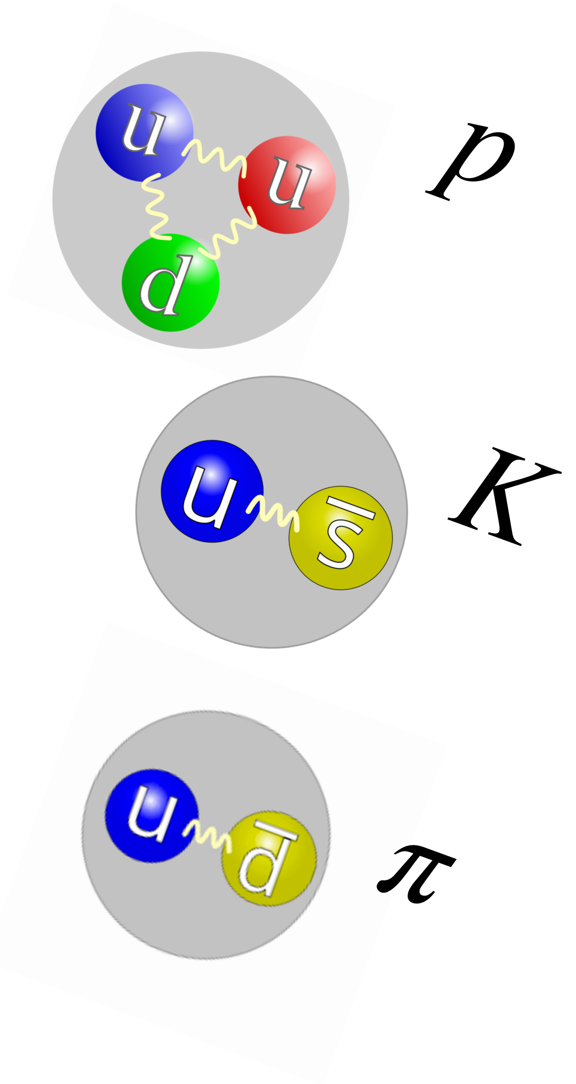


Kaon



Generalize the Approach

Physics | Subsystems | AI/ML model | PID



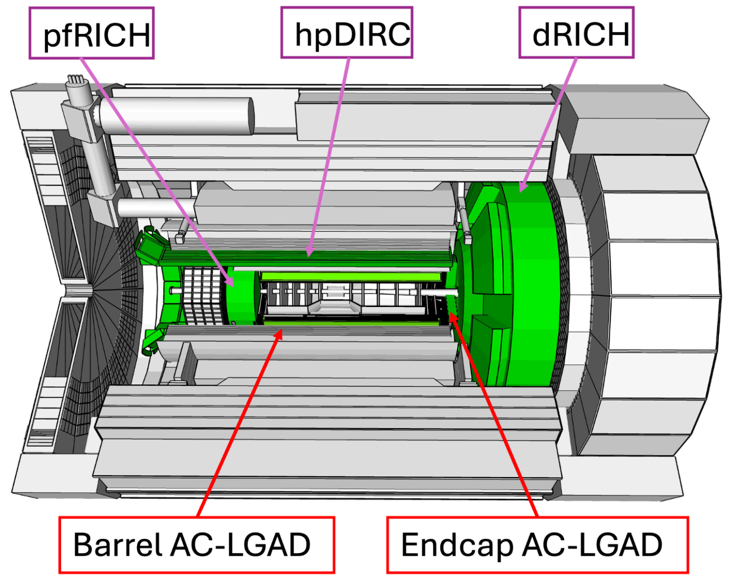
pfRICH

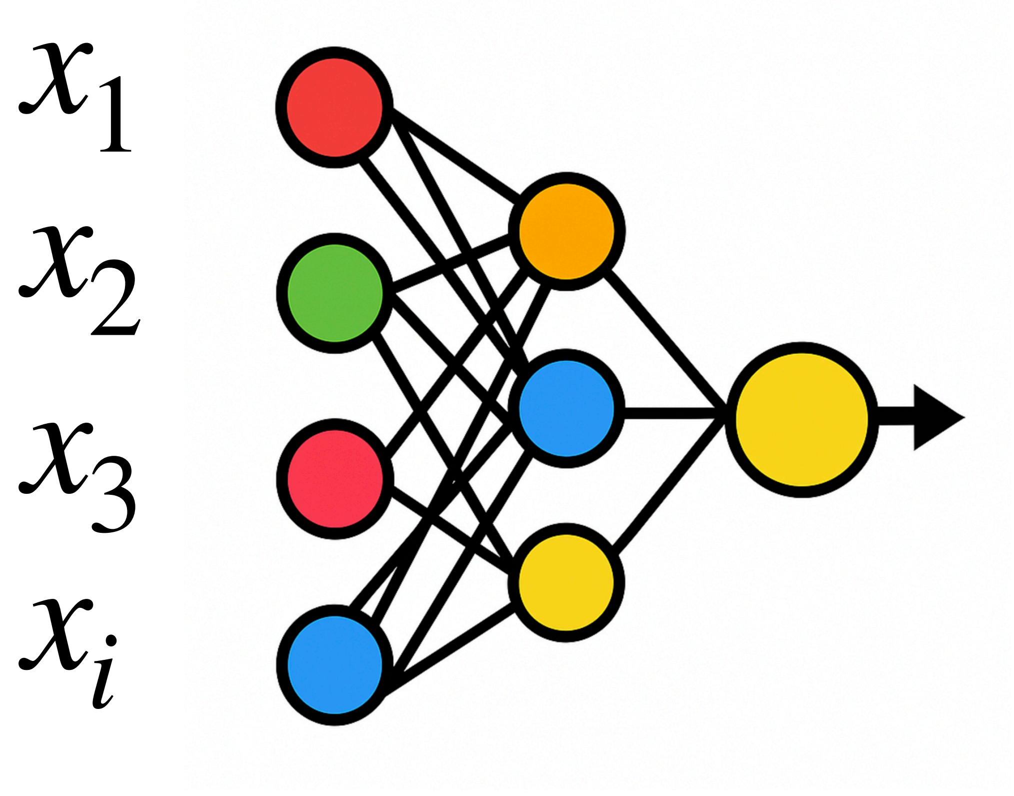
hpDIRC

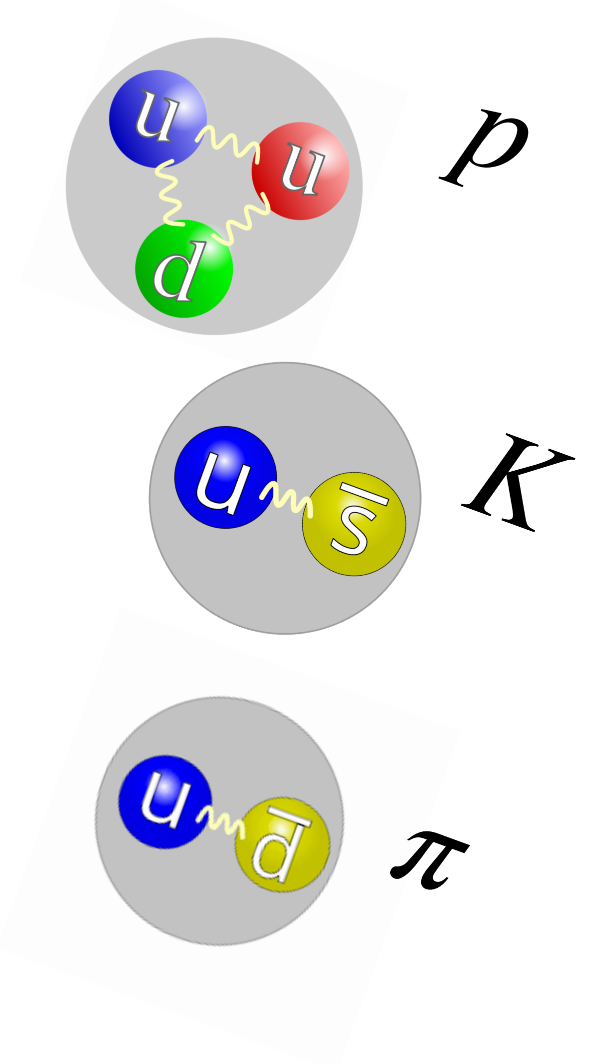
dDRICH

...

ePIC







This method can be extended to <u>all ePIC PID subsystems</u>

Ongoing

The more information available, the better the model will perform

Conclusion

ePIC is a state-of-the-art detector

It requires state-of-the-art computational tools

Detector performance is the best entry point for impactful AI/ML applications

The hybrid ML model for pfRICH is robust

Its integration into the ePIC software is ongoing

The approach is readily extendable to other PID systems