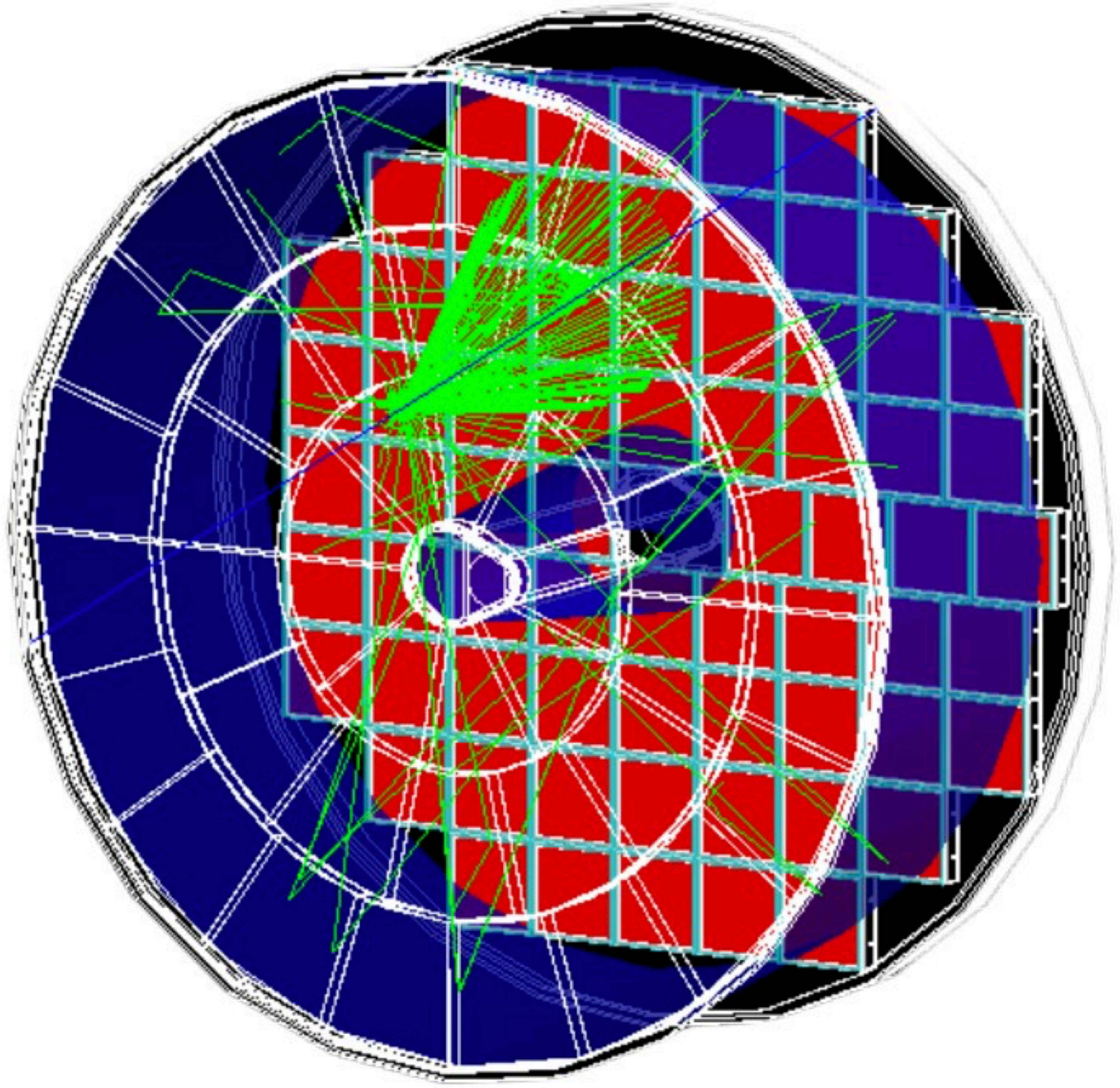
Machine Learning for the pfRICH Particle Identification subsystem



D.H Dongwi, C-J. Naïm and L. Rhode

Artificial Intelligence for the Electron Ion Collider (2025)
October, 28, 2025

Stony Brook University

AI4EIC

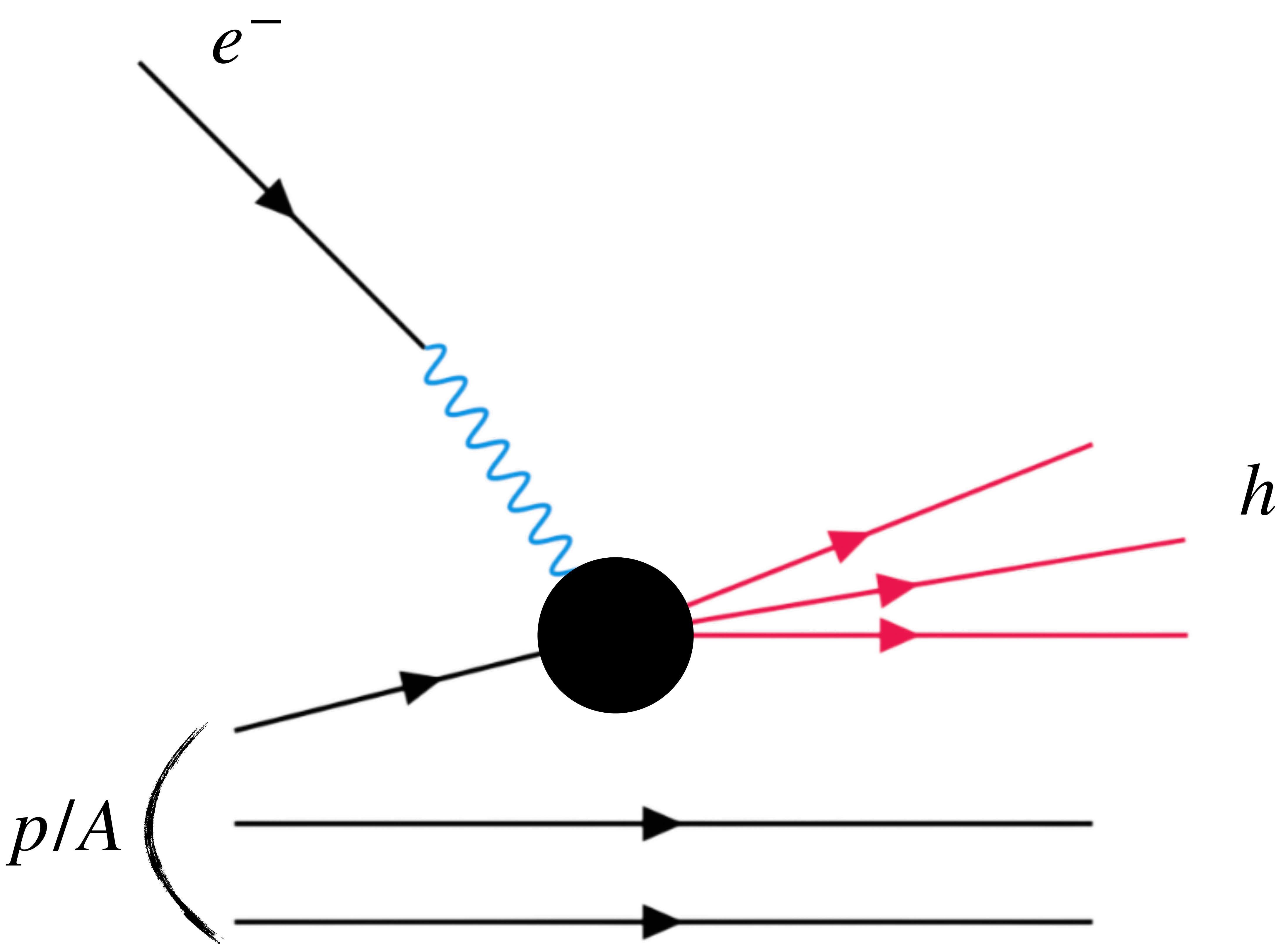
Center for Frontiers in Nuclear Science

The ePIC Detector at the EIC

pfRICH subsystem in backward region

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

Physics Motivations at the EIC



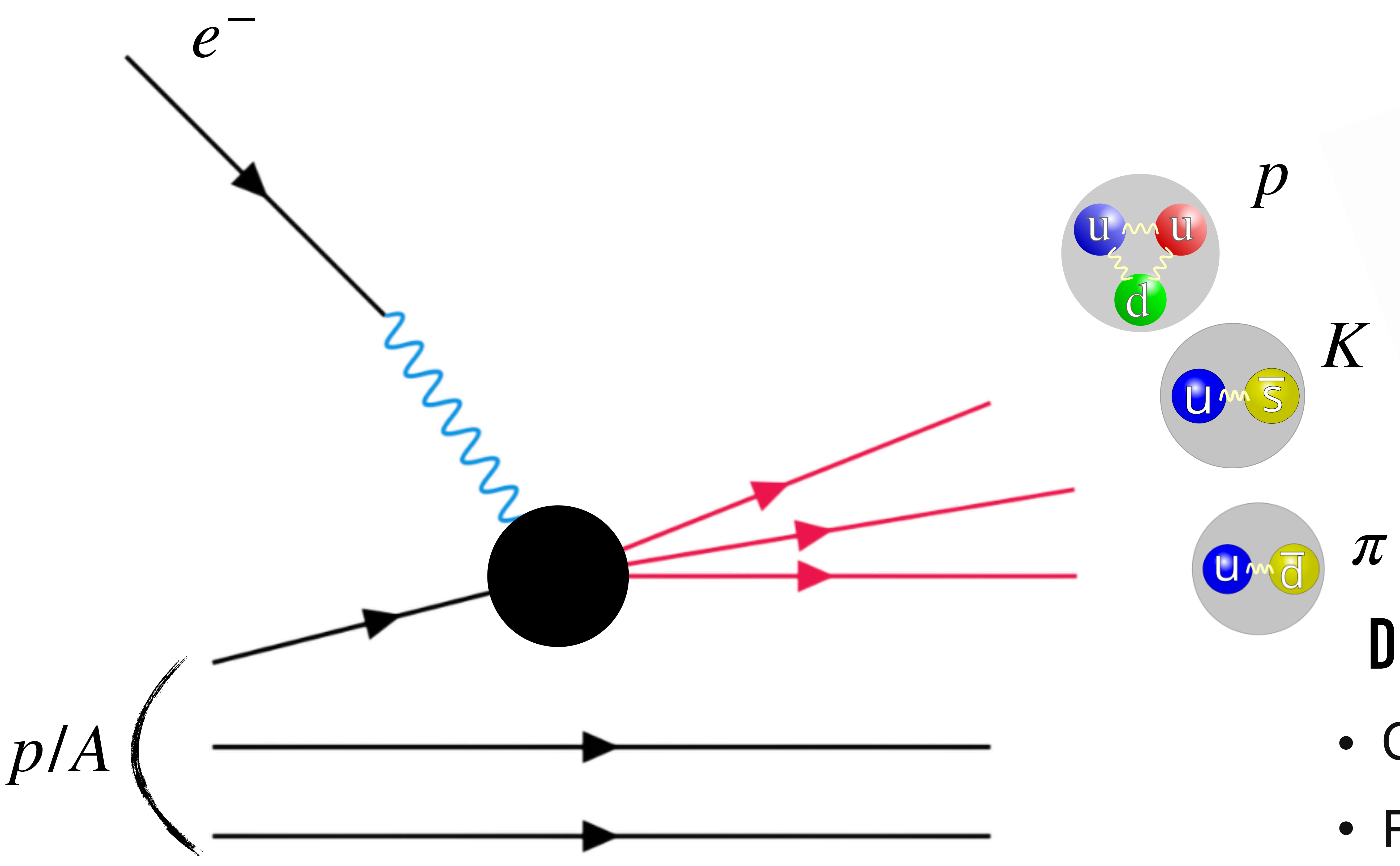
- **Semi-Inclusive Deep Inelastic Scattering**
- Production of hadrons in final-state
- Provide information on:

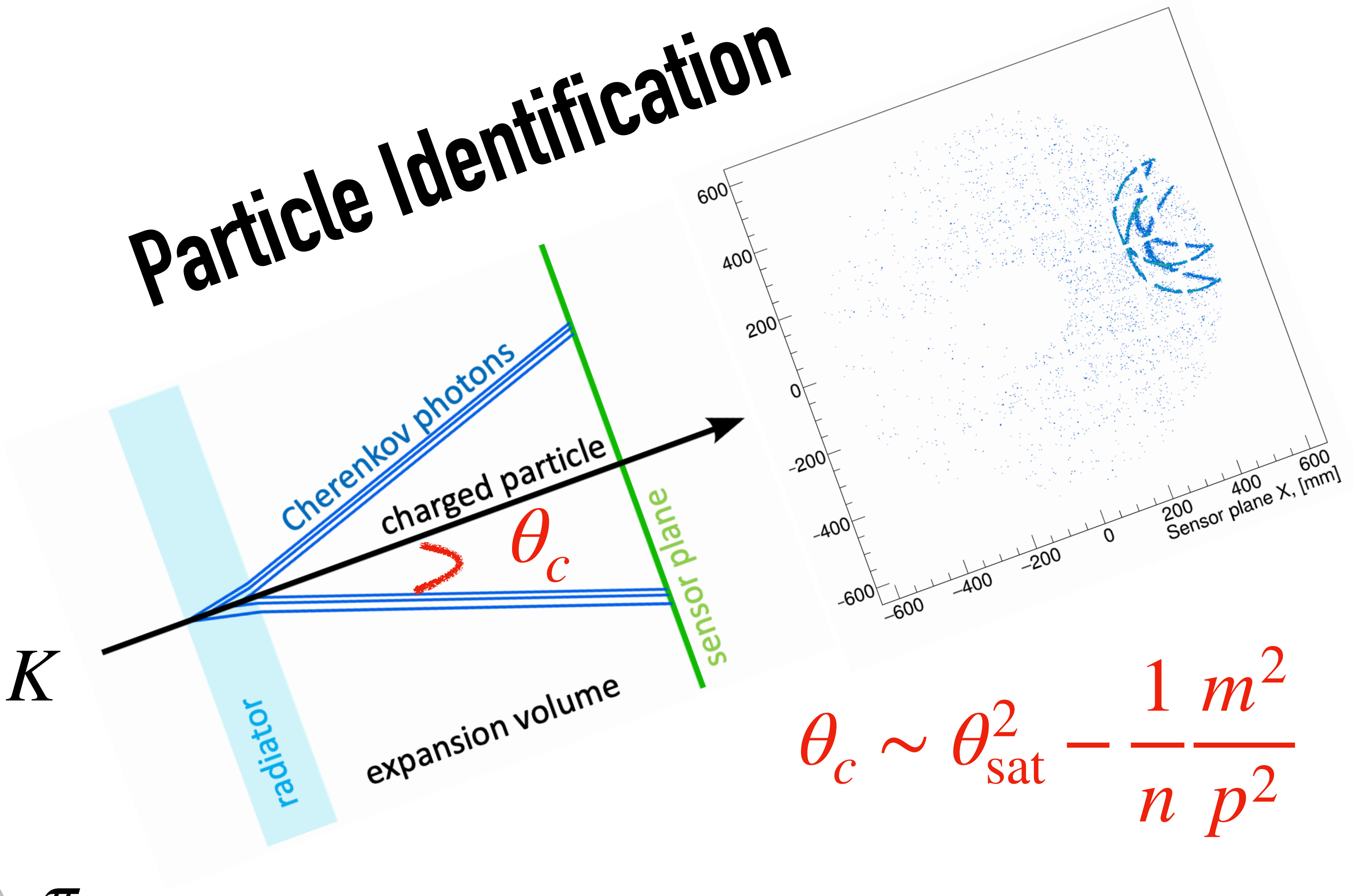
→ the ***fragmentation process (hadronization)***

→ the ***hadronic structure***

Particle Identification detectors are crucial

The pfRICH Concept





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

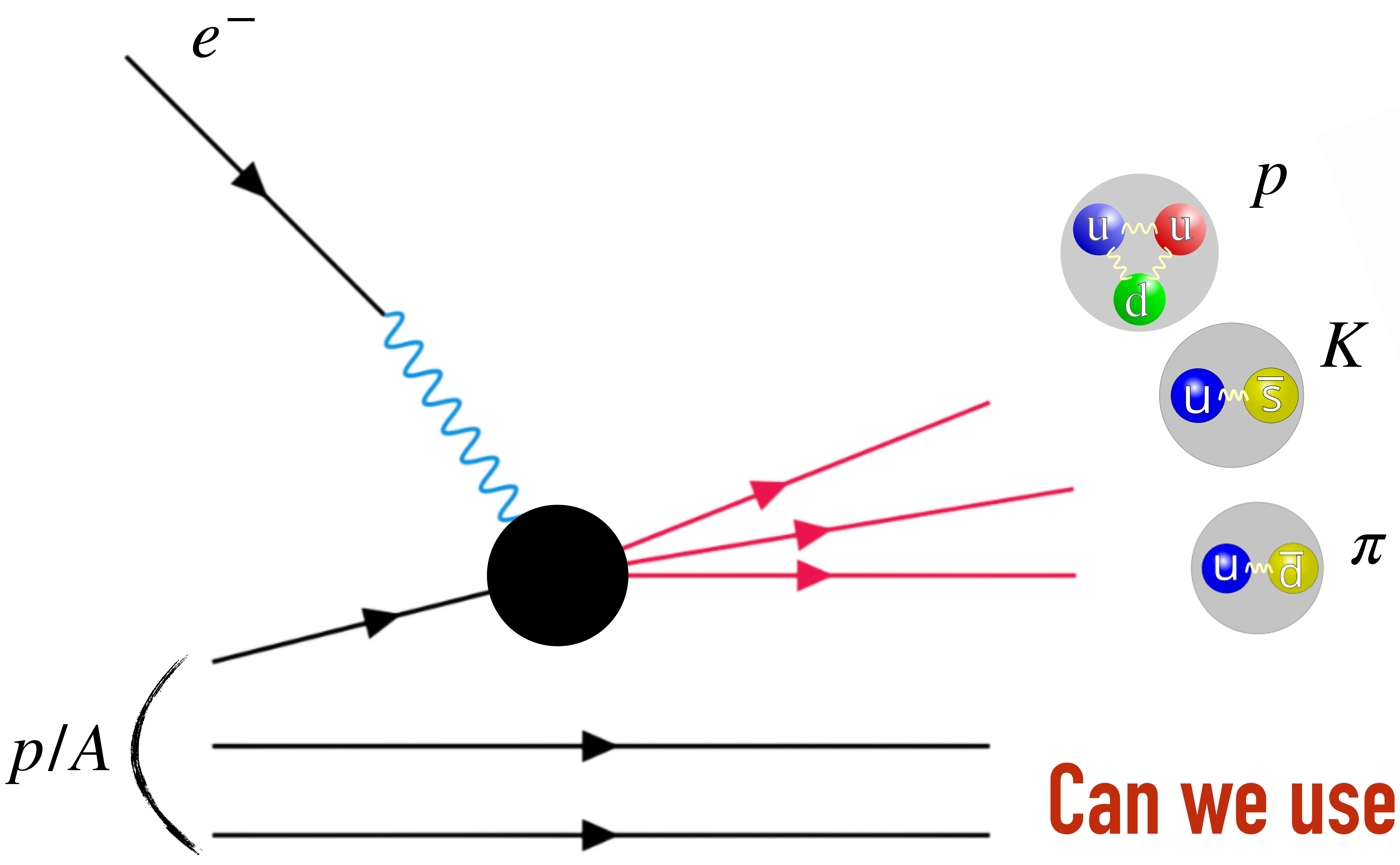
The pfRICH will provide > 3σ π/K separations for momentum up to 7 GeV/c for –3.5 < η < -1.5

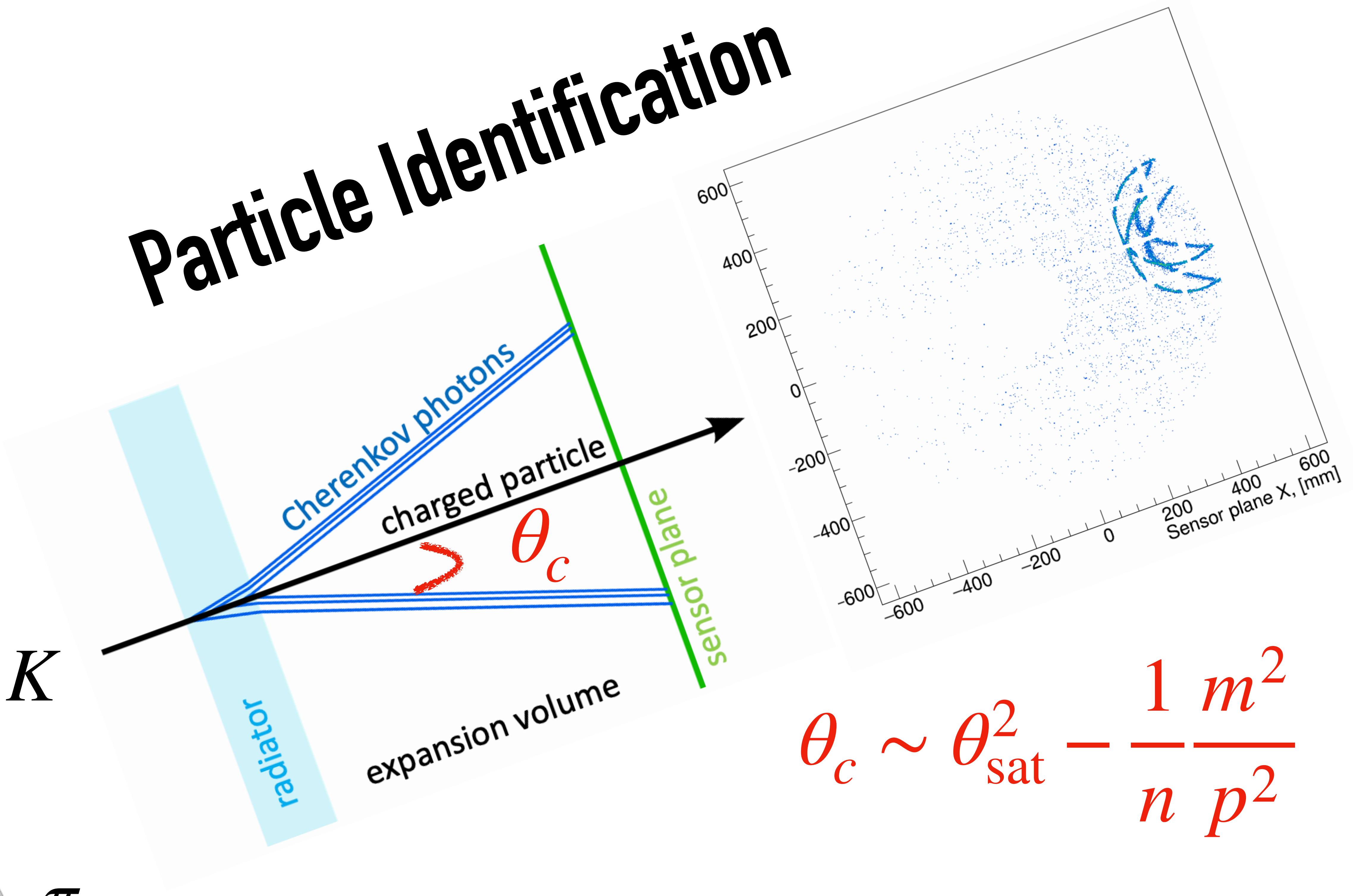
Detection Principle

- 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**

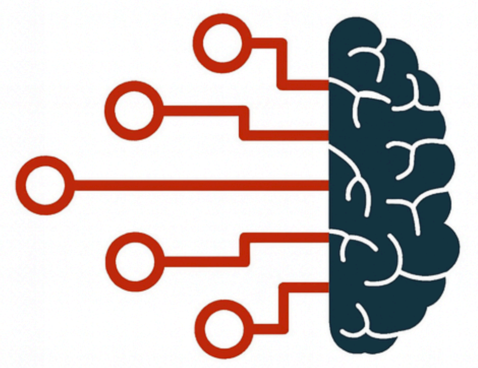
The pfRICH Concept





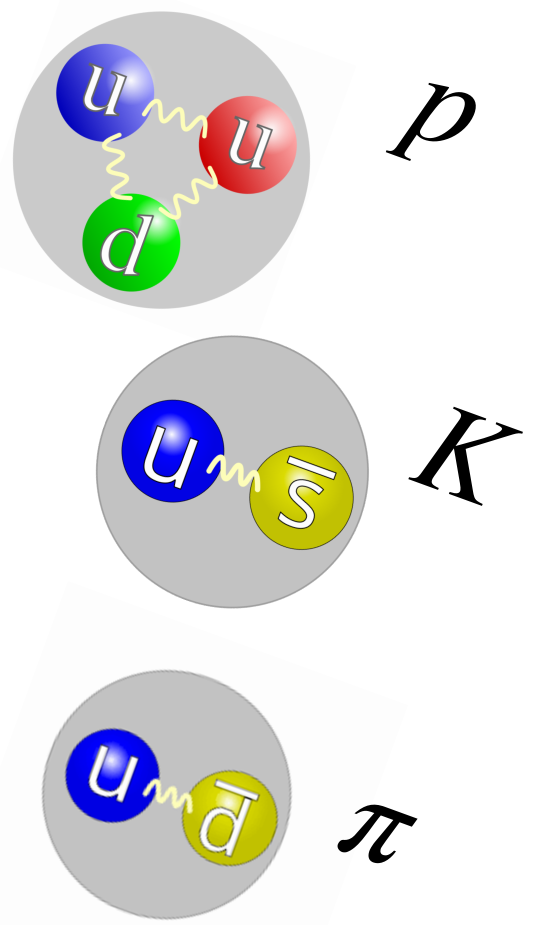
$$\theta_c \sim \theta_{\text{sat}}^2 - \frac{1}{n}\frac{m^2}{p^2}$$

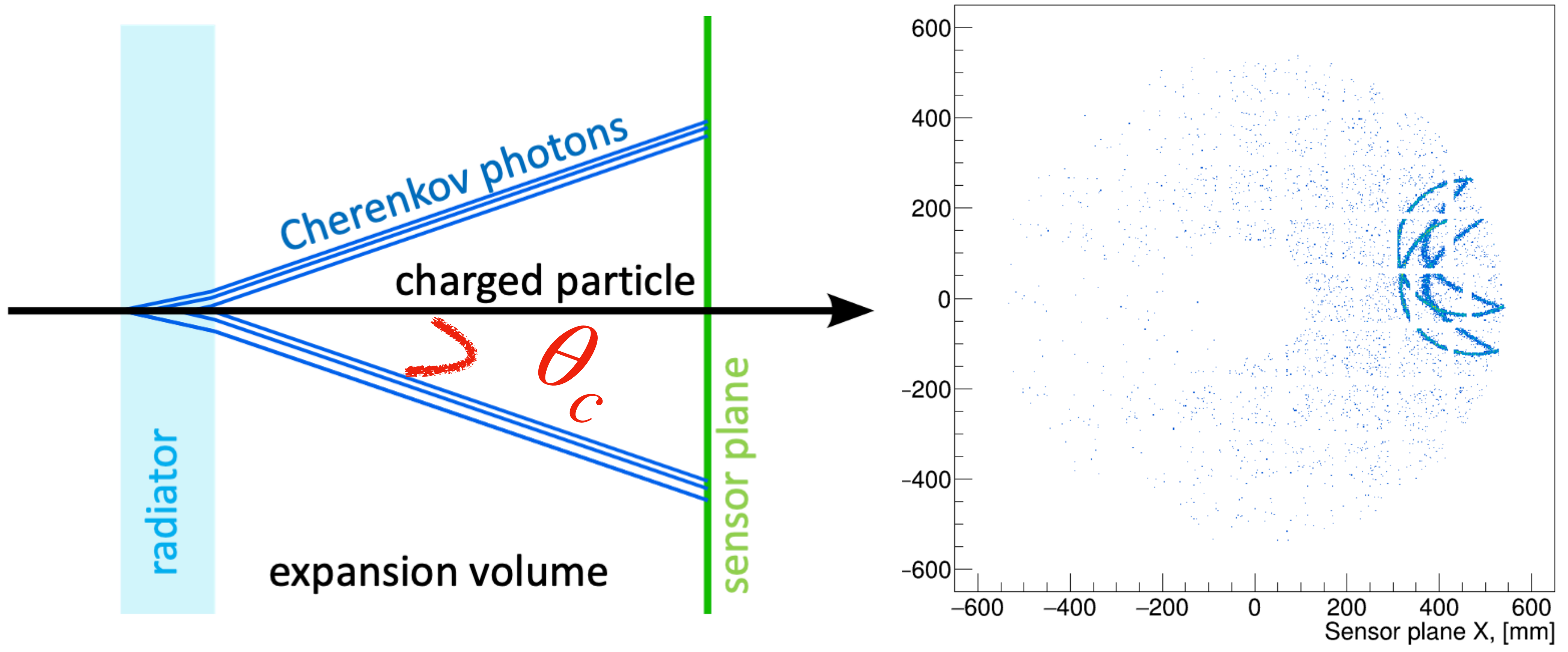
Can we use machine learning to improve particle identification?

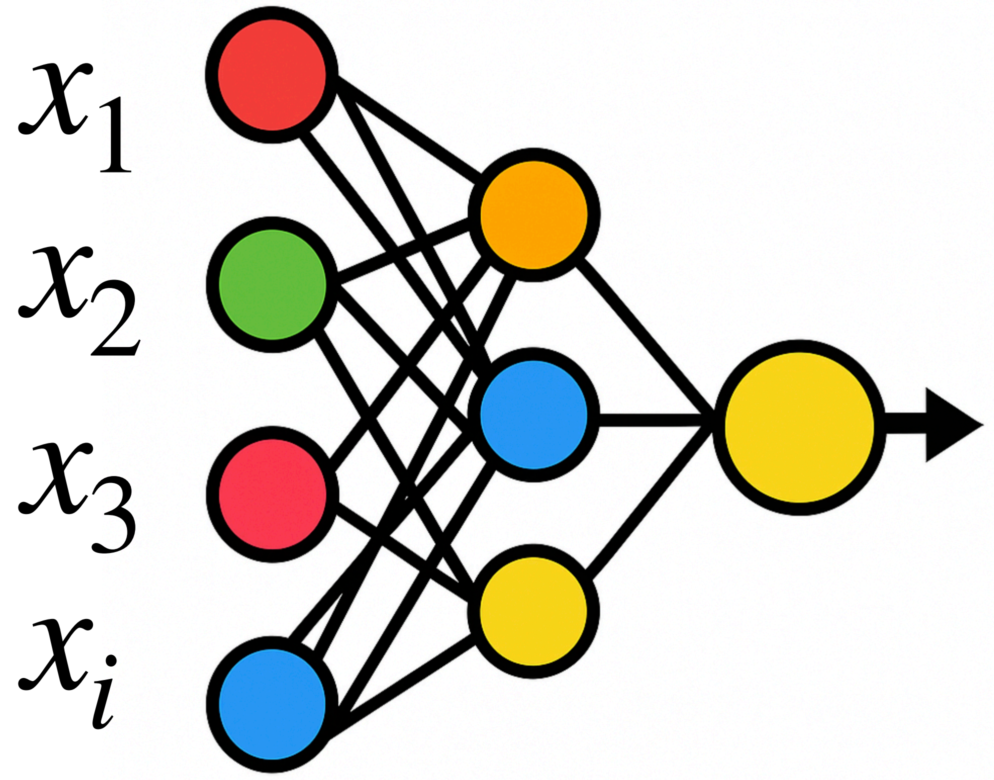


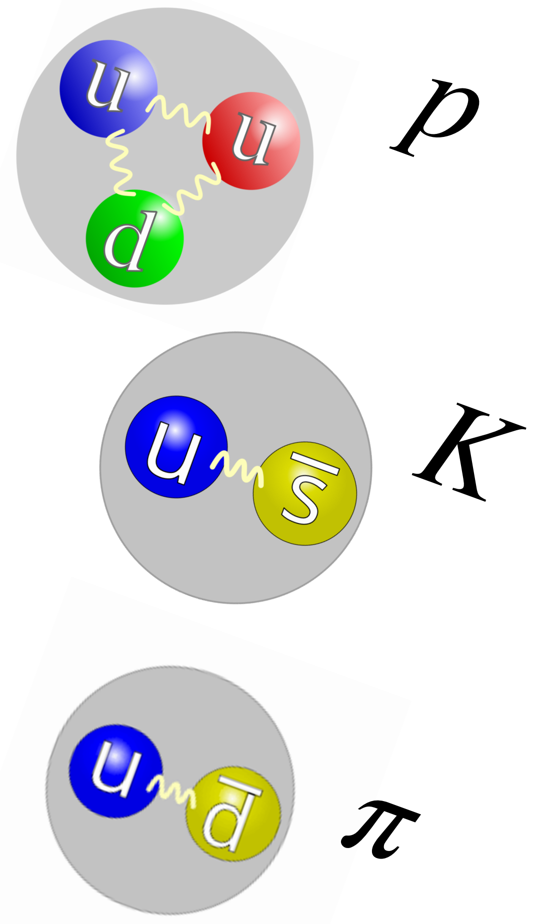
The Approach

Physics → pfRICH → AI/ML model → PID







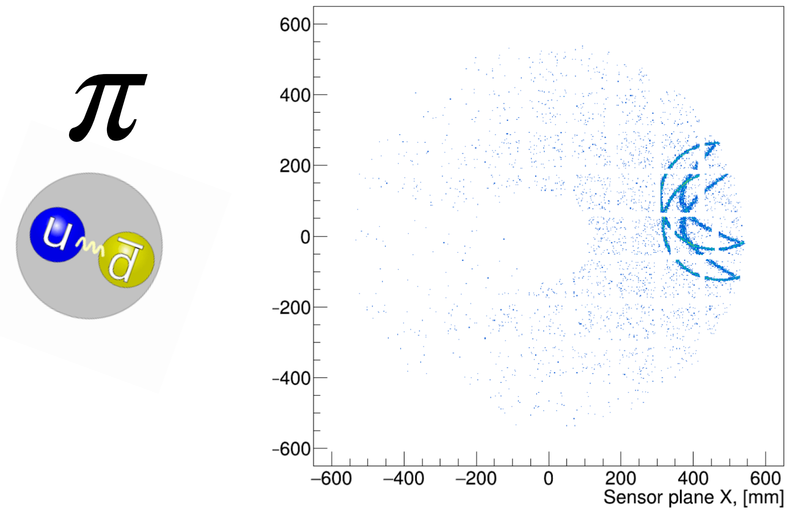


Can we use machine learning to improve particle identification?: Yes!

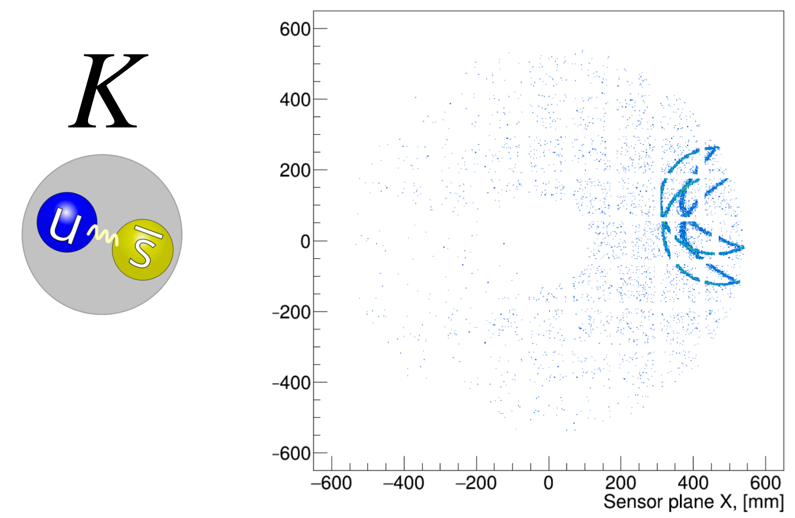
An ideal use case for AI/ML, since the signal is well defined and fully understood

Model Training

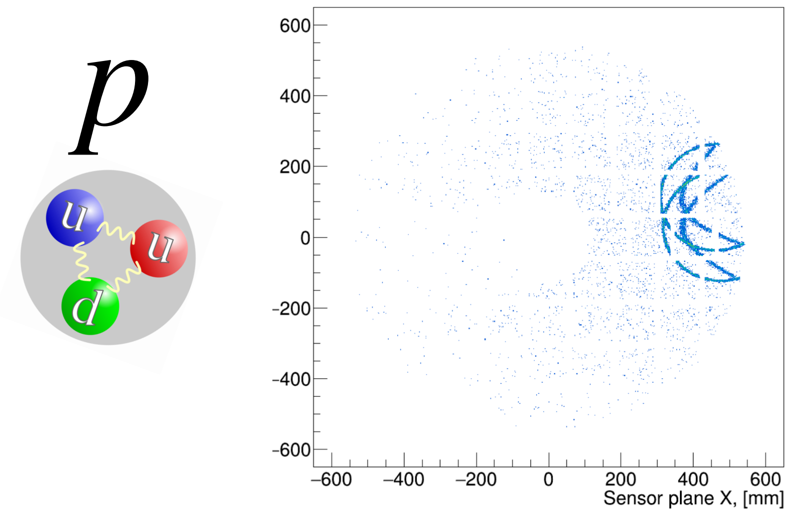
π



K

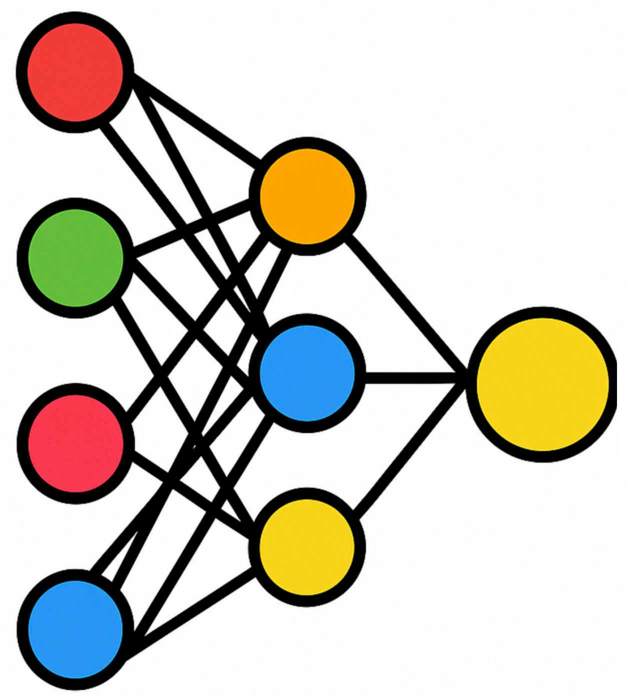


p



Training

x_1
x_2
x_3
x_i



AI/ML model

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

XGBoost

Gradient-boosted hybrid

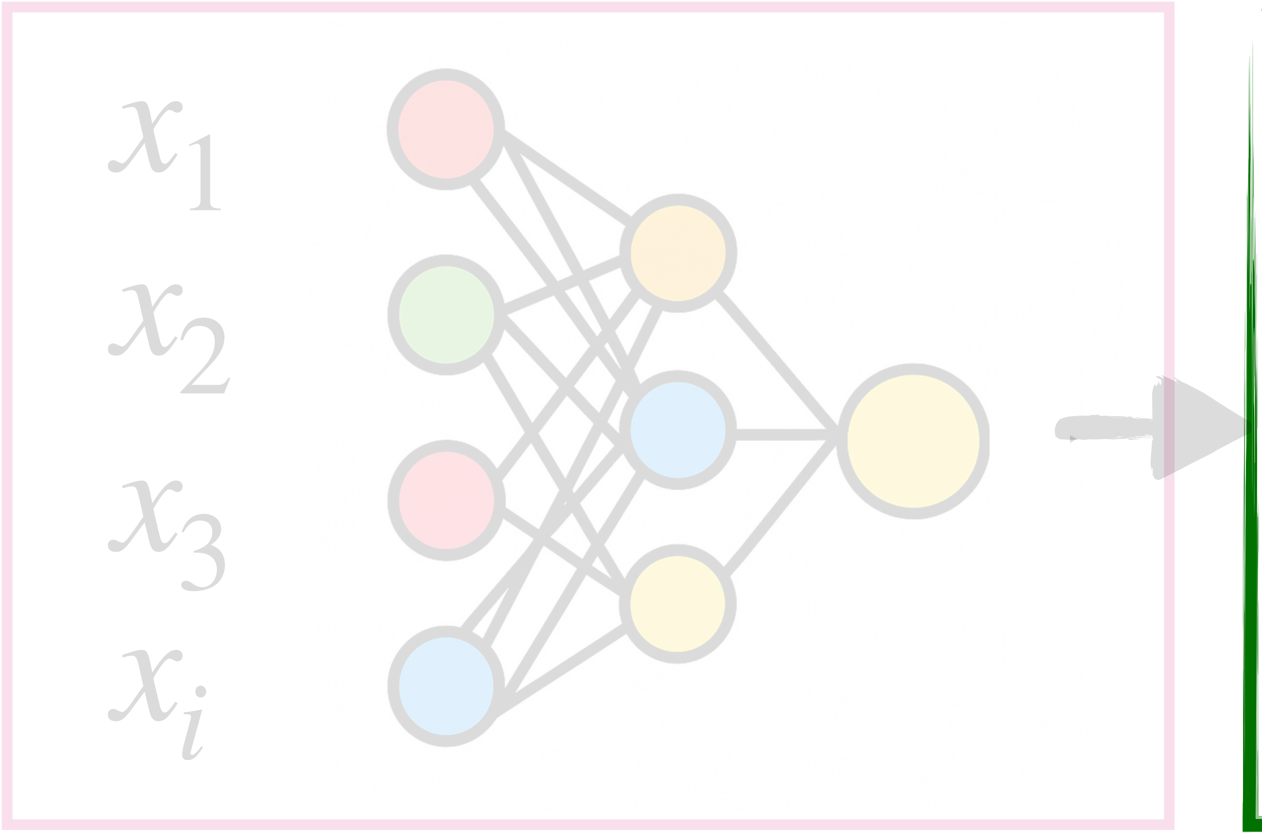
Diffusion model (ongoing)

Standalone ePIC pfRICH GEANT4:

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

Model Inference

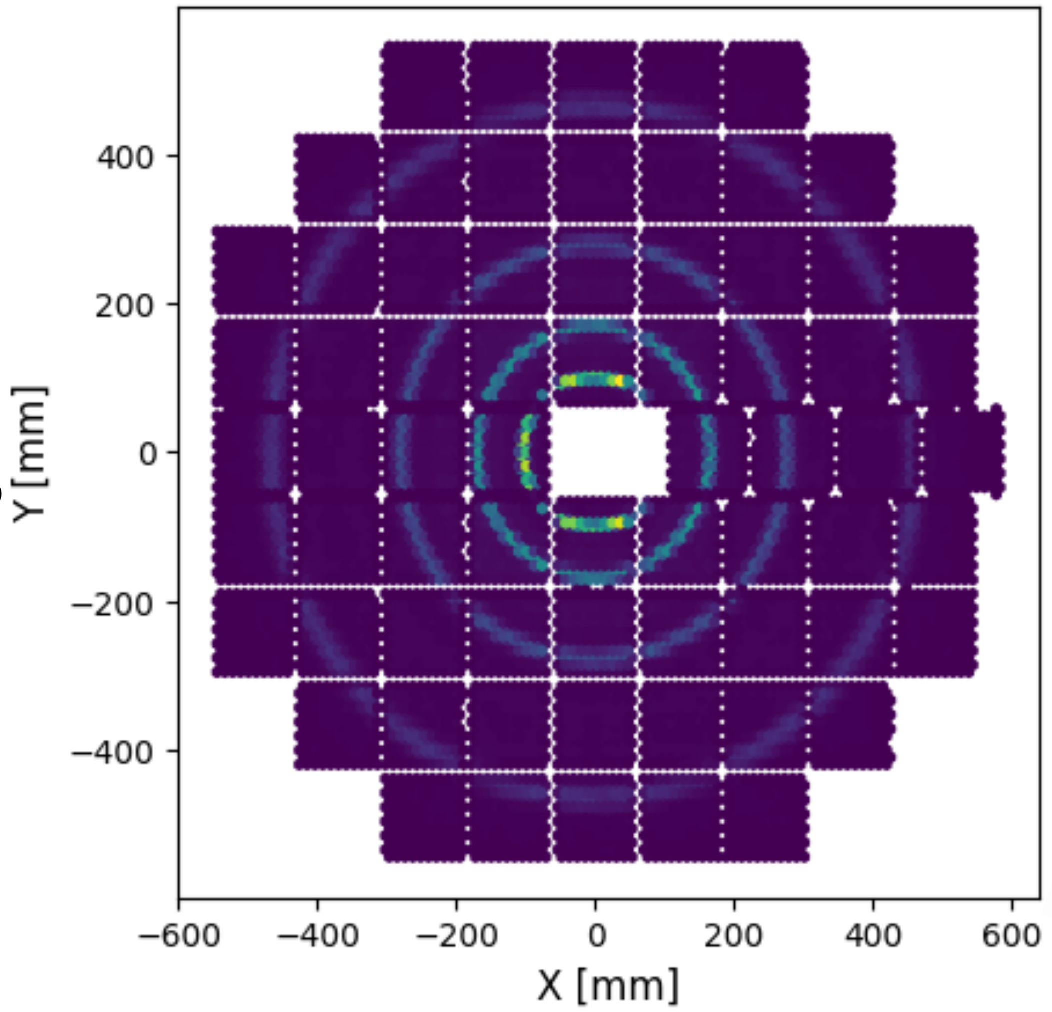
Training

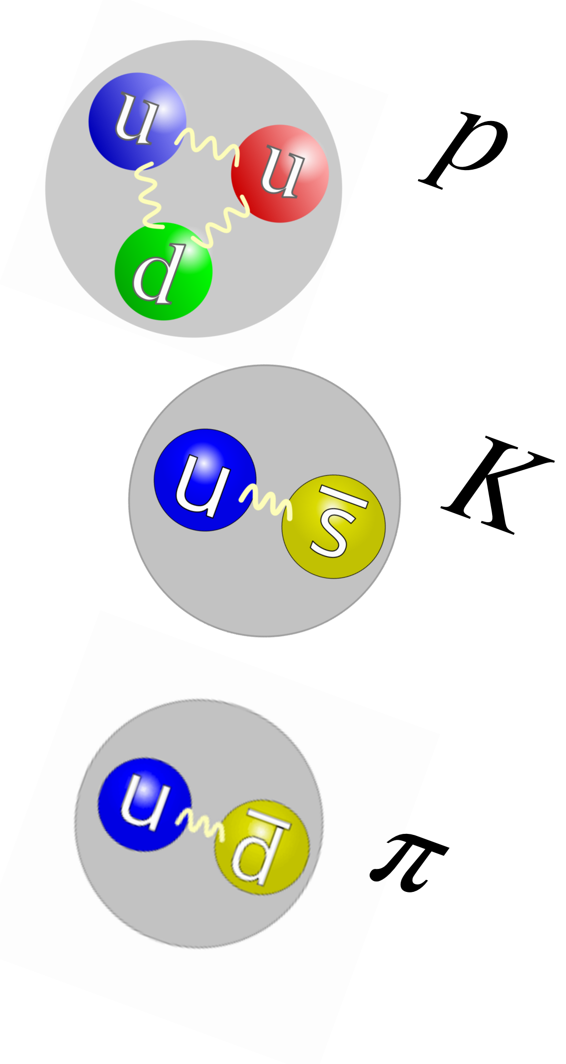


AI/ML model

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

Data

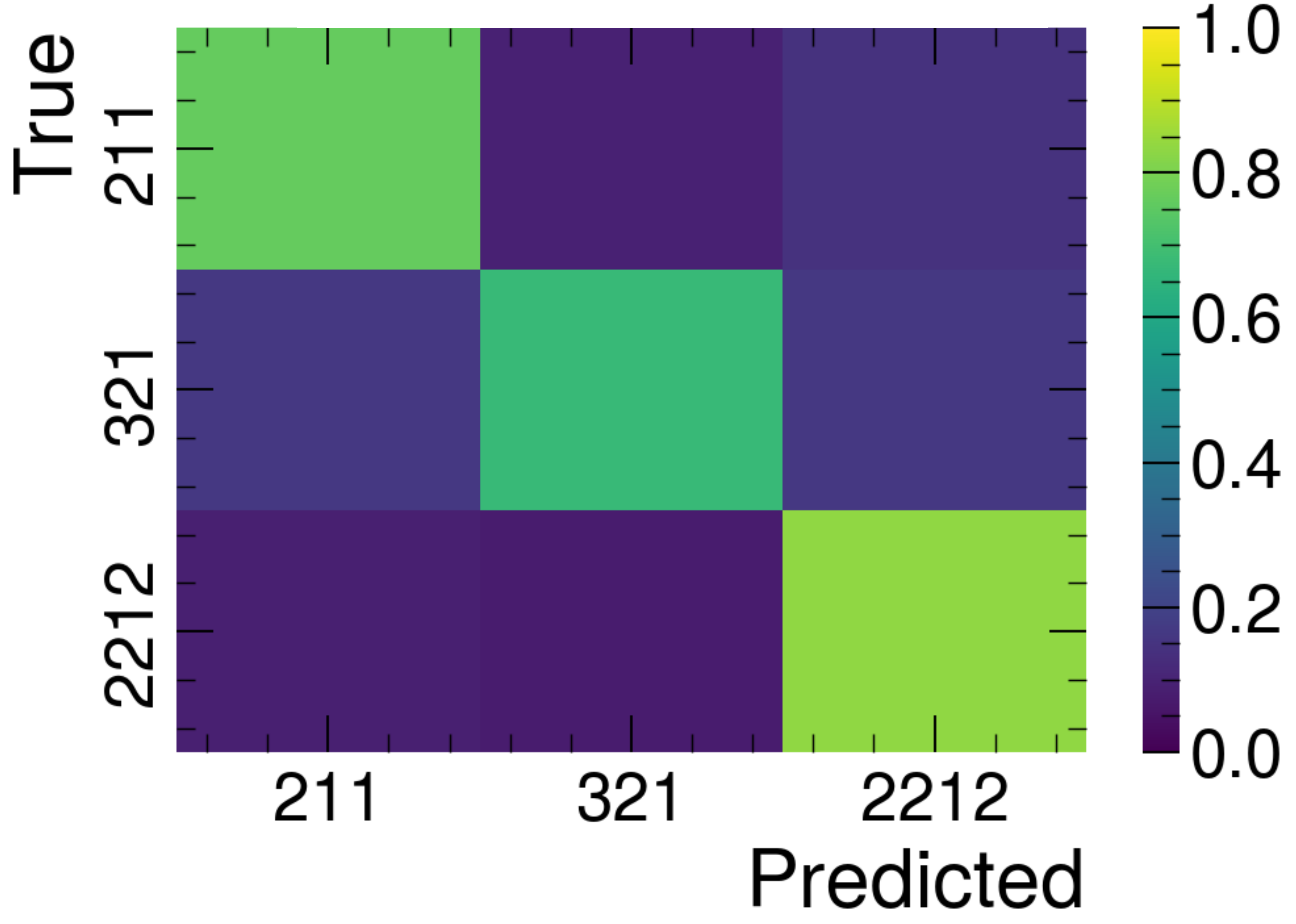




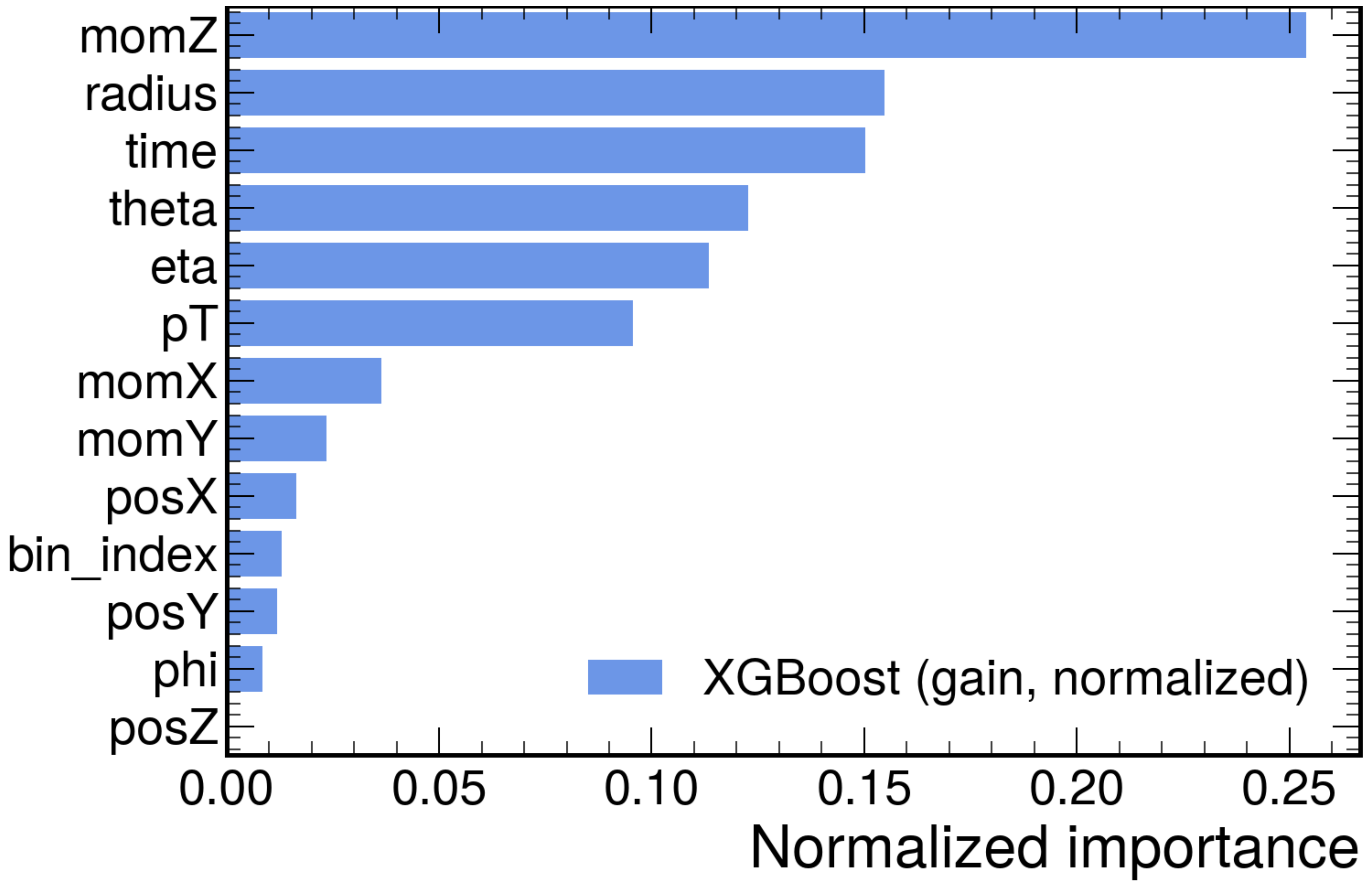
Particle Identification

Results

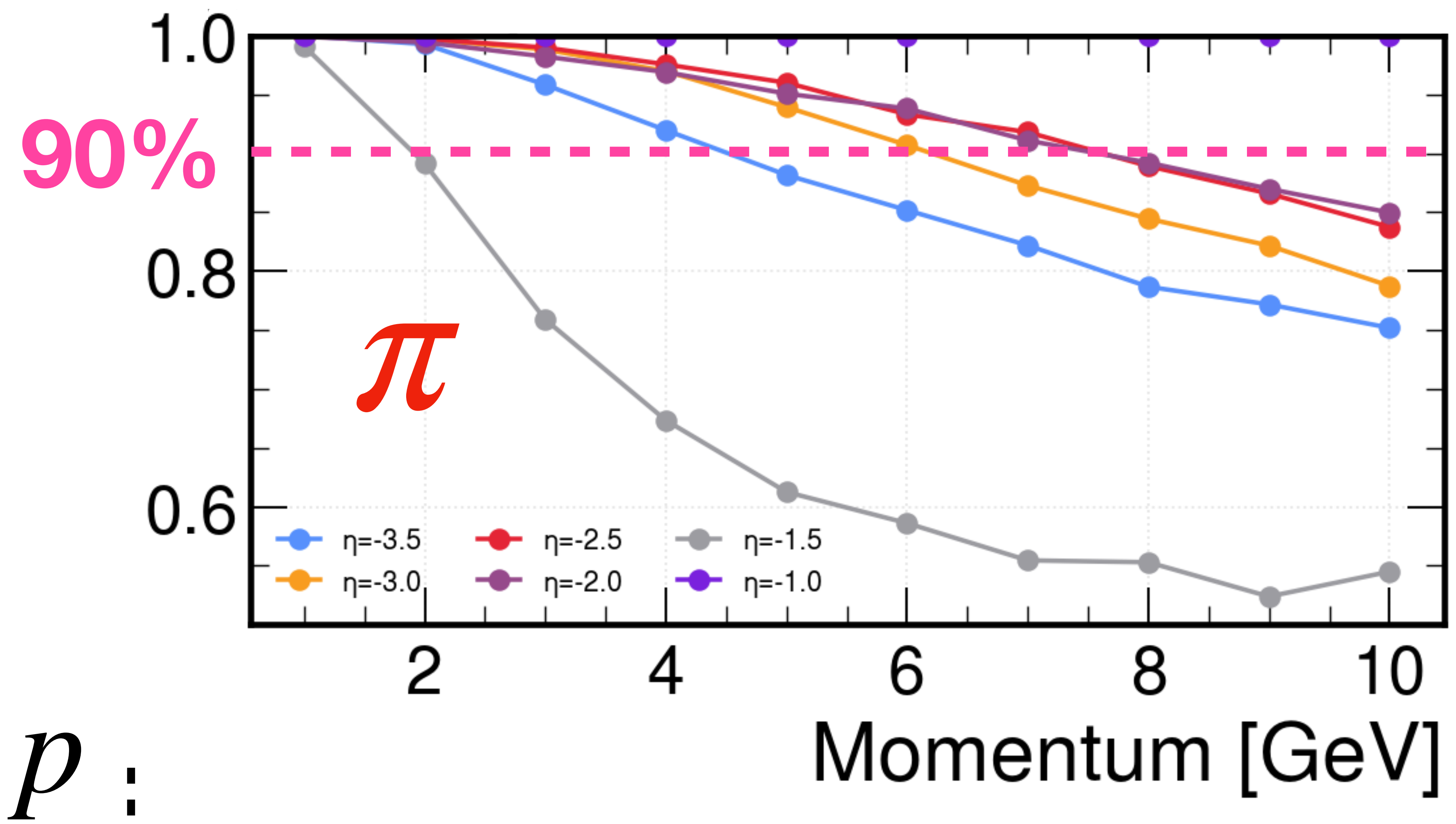
Confusion matrix

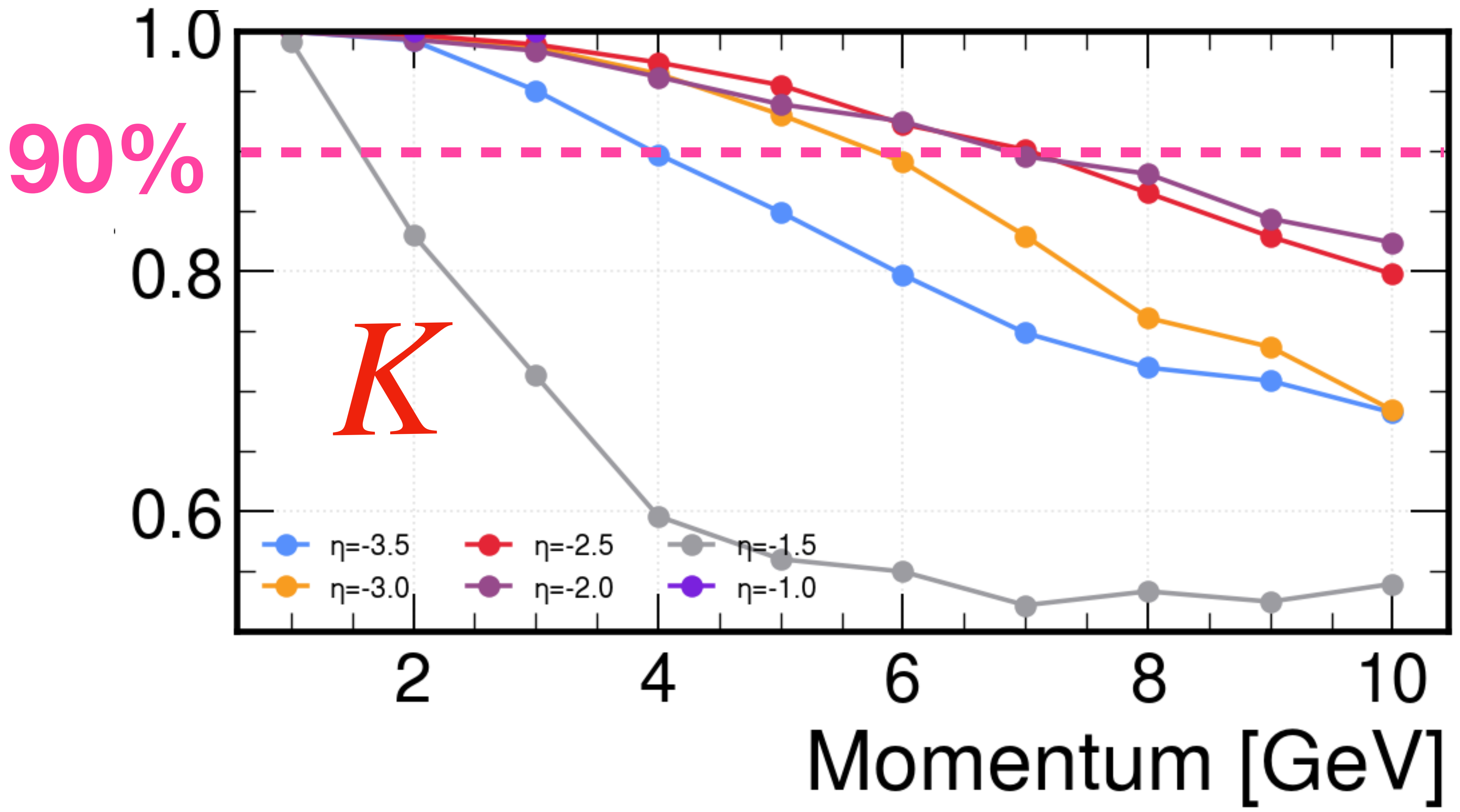


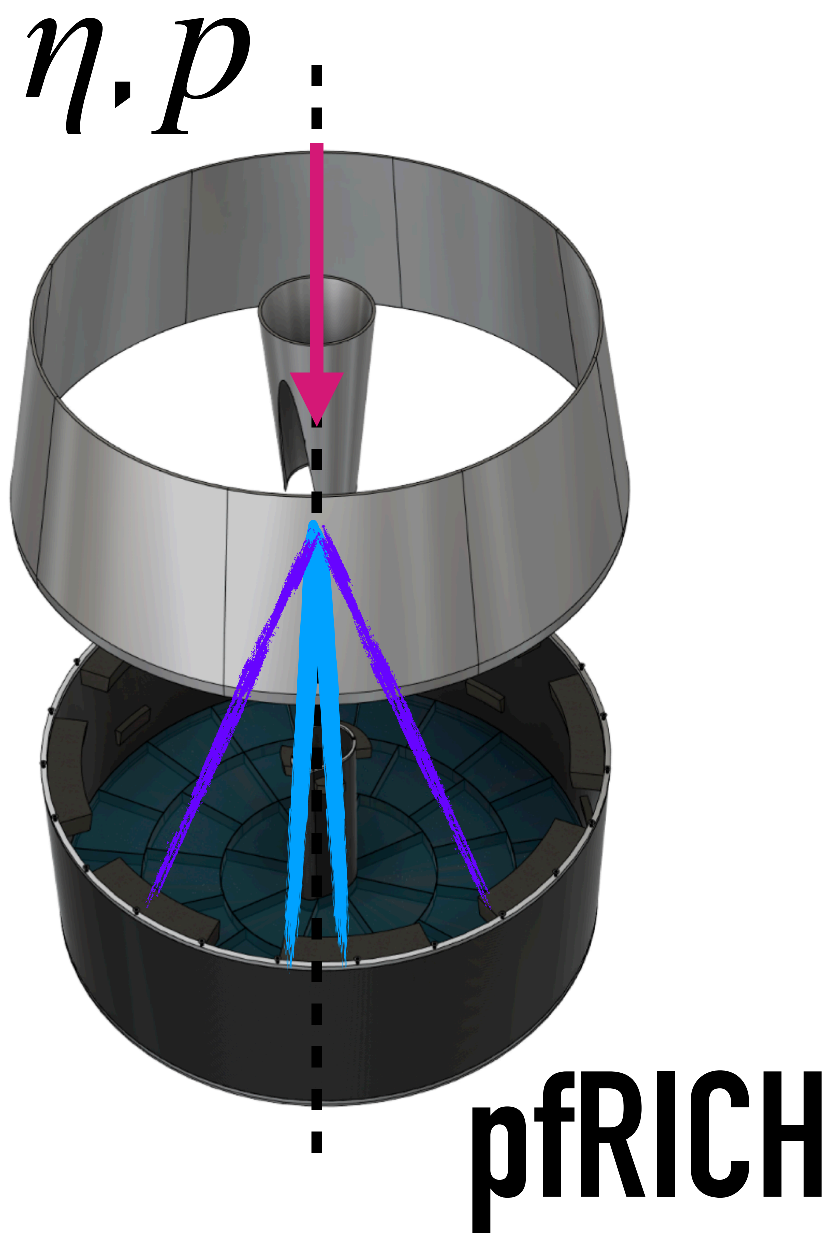
Feature importance

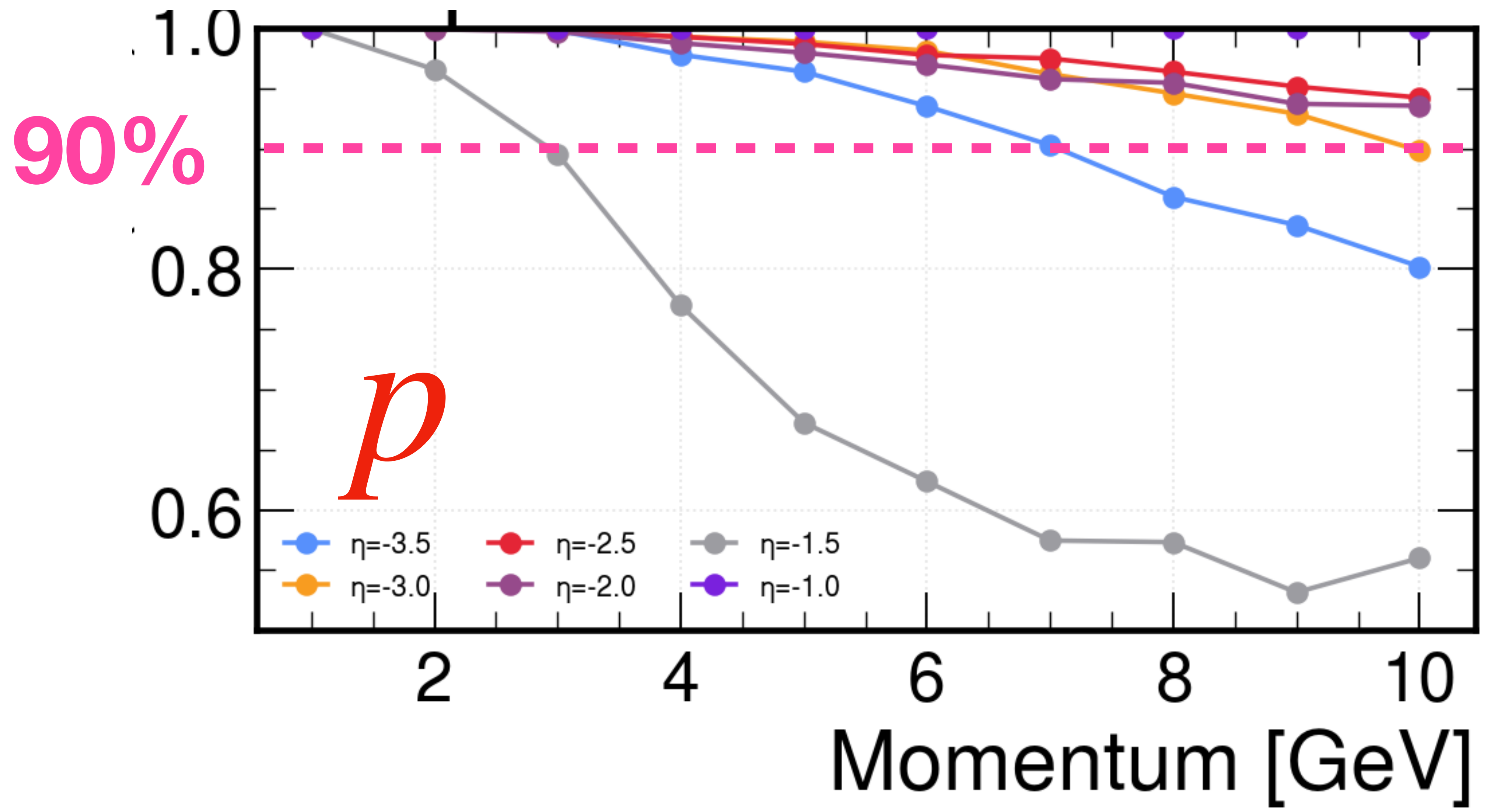


Separation Efficiency





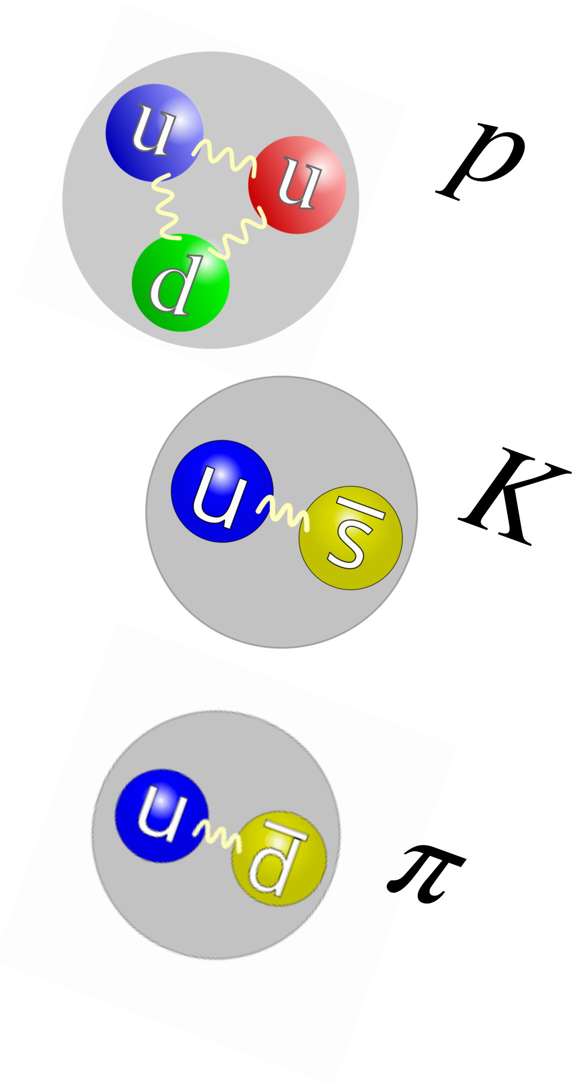




Promising results!

Generalize the Approach

Physics → Subsystems → AI/ML model → PID



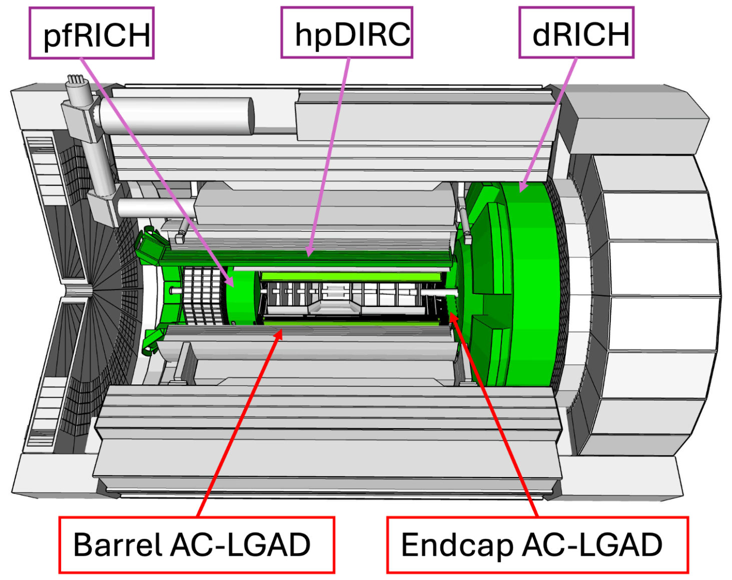
pfRICH

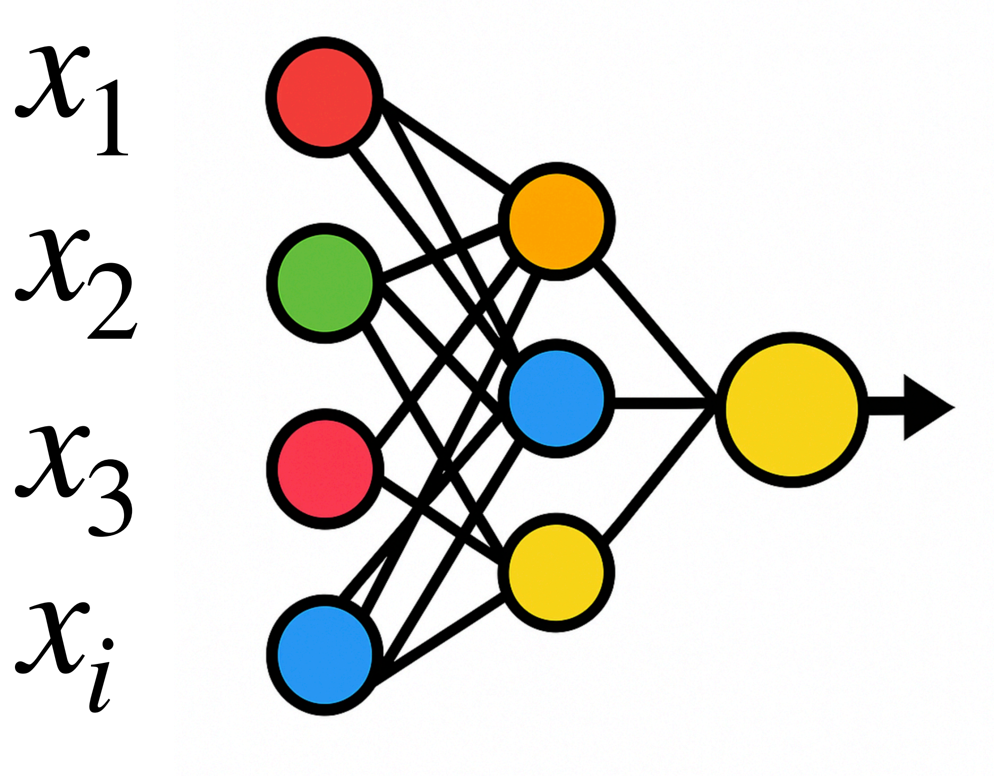
hpDIRC

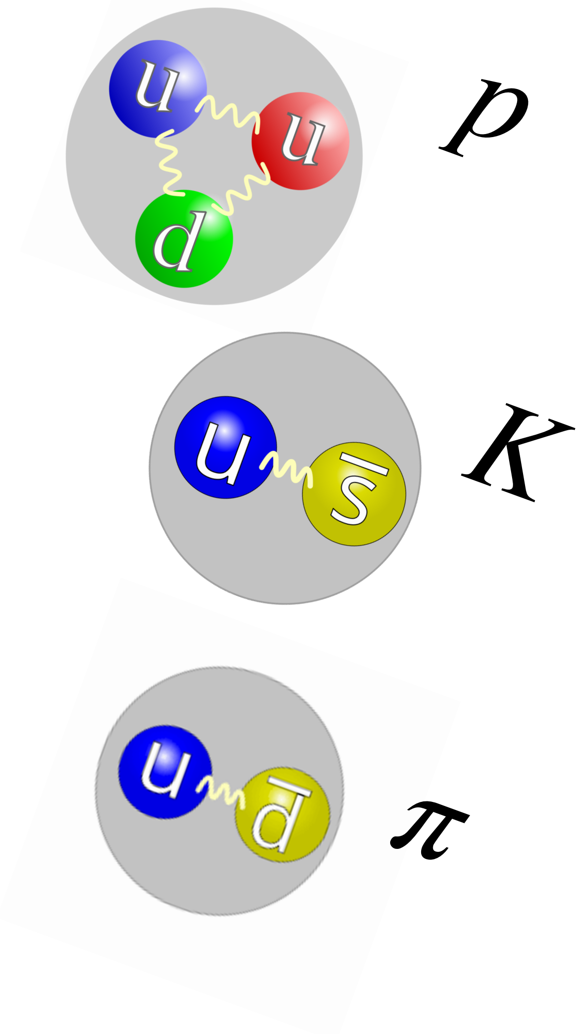
dDRICH

...

ePIC







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

The more information available, the better the model will perform

Conclusion

- **Promising separation** observed — clear distinction between particle species
- Overall, a solid step toward robust and reliable particle identification
- AI/ML methods can be generalized to all ePIC subsystems

See Dmitry's talk

Next steps:

- Optimize PID parameters + diffusion model (work in progress)
- Use data produced using NVIDIA OptiX and GDML-based detector geometries
- Generate true SIDIS events within the ePIC software

See Gabor's talk

Backup Slides

Diffusion model (work ongoing)