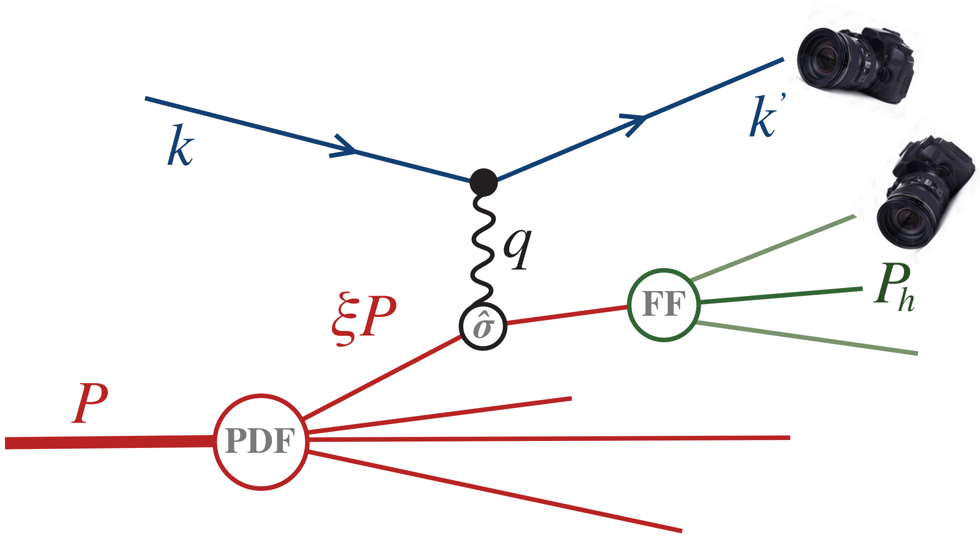
This work is supported by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics, under contract DE-AC02-06CH11357.

Argonne NATIONAL LABORATORY

EXPERIMENTAL CONSIDERATIONS

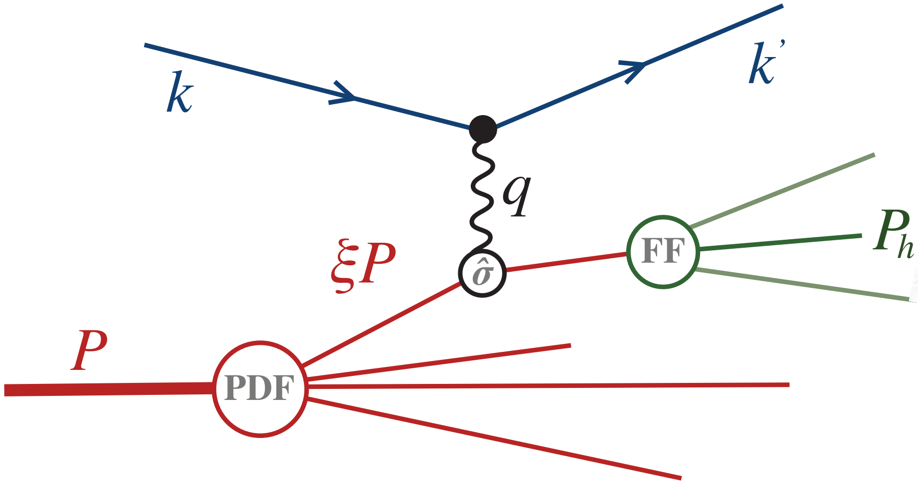
TOWARD EVENT-LEVEL ANALYSIS OF SEMI-INCLUSIVE DEEP INELASTIC SCATTERING



SYLVESTER JOOSTEN
sjoosten@anl.gov

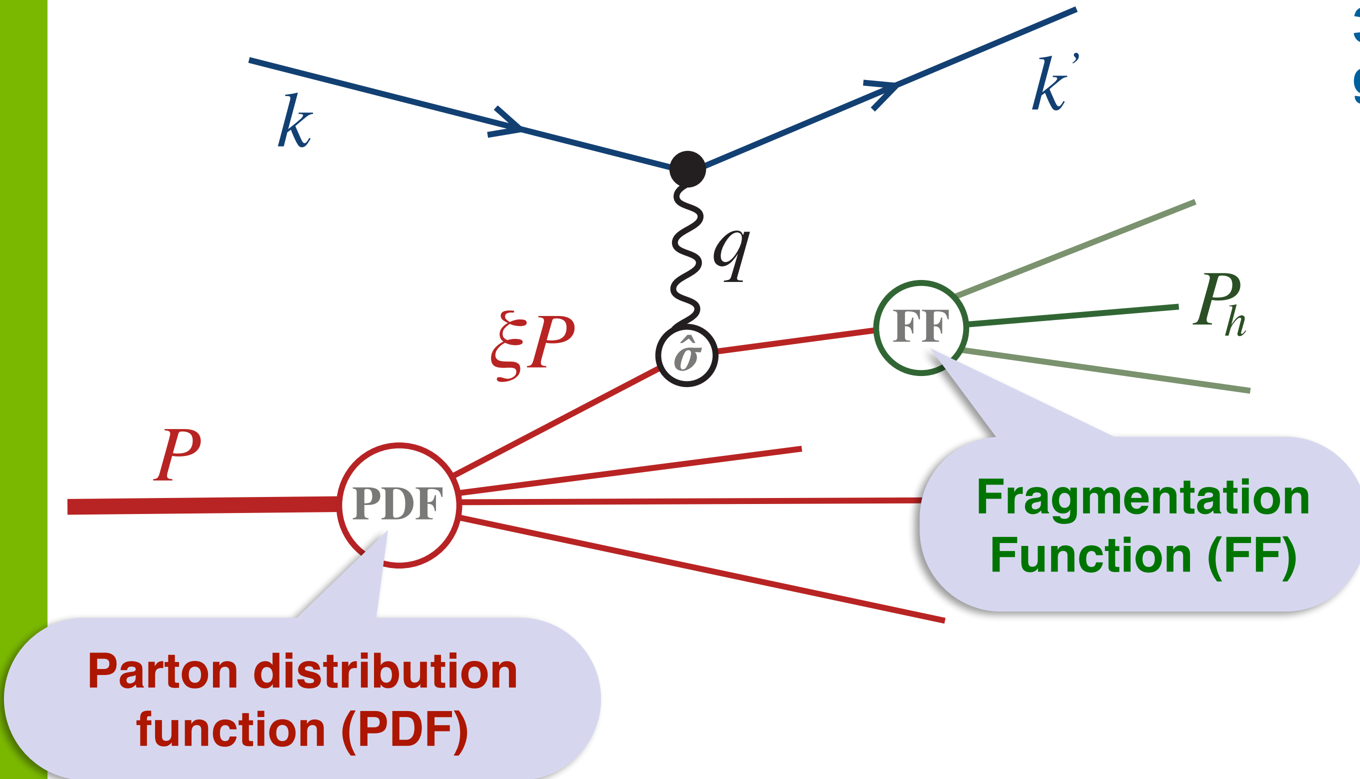
This work is supported by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics, under contract DE-AC02-06CH11357.

QCD at the Femtoscale in the Era of Big Data
INT 24-2 - June 2024

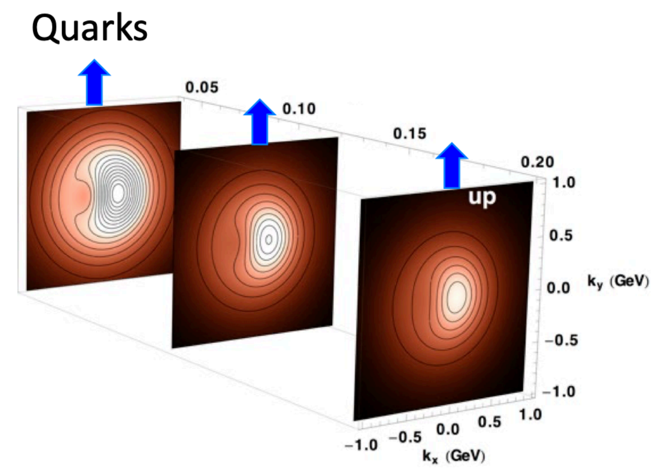


WHAT ARE WE TRYING TO MEASURE?

SEMI-INCLUSIVE DEEP-INELASTIC SCATTERING



3-D imaging of quarks and gluons in momentum space



Measuring an identified hadron in coincidence with the scattered electron

Degrees of freedom:

Continuous:

- Electron: x, Q^2, ϕ
- Hadron $z, P_h^{\perp}, \phi_h$

Discrete:

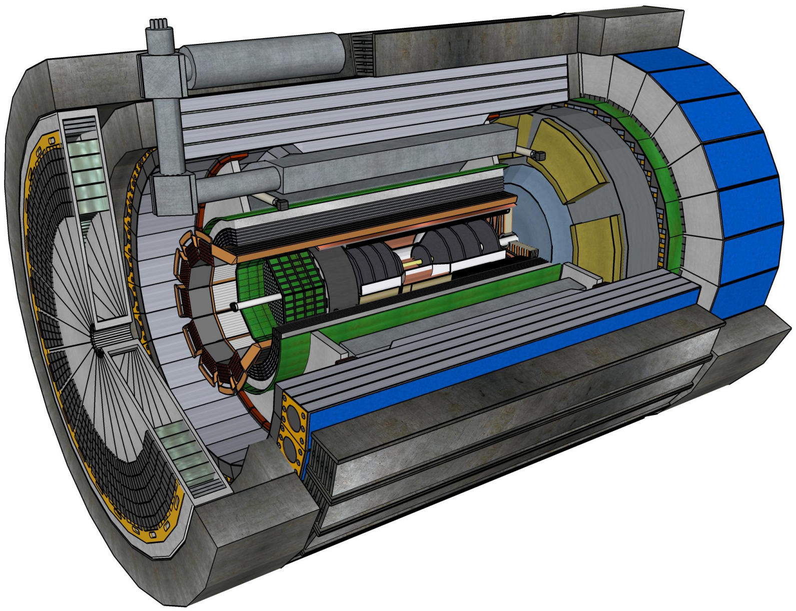
- Electron spin and proton spin
- Hadron flavor (π, K, p/n)

WHAT DO WE NEED TO MEASURE FOR THIS?

Degrees of freedom:
Continuous:
- Electron: x, Q^2, ϕ
- Hadron $z, P_h^\perp, \phi_h$

Discrete:
- Electron spin and proton spin
- Hadron flavor (π, K, p/n)

- Electron variables (x, Q^2, ϕ):
  - Event vertex’
  - 3-momentum (preferably through a direct measurement)
  - Electron identification (are we sure this is an electron?)
  - Scattered electron selection (are we sure this is the scattered electron?)
- Hadron variables $(z, P_h^\perp, \phi_h)$, flavor
  - Reconstructed hard scattering (i.e. measured scattered electron)
  - Hadron vertex (should be the same as the event vertex)
  - 3-momentum
  - Positive hadron identification (how sure are we that this is a π, K, p/n?)
- Beam spin states:
  - Polarimetry to determine the polarization percentage

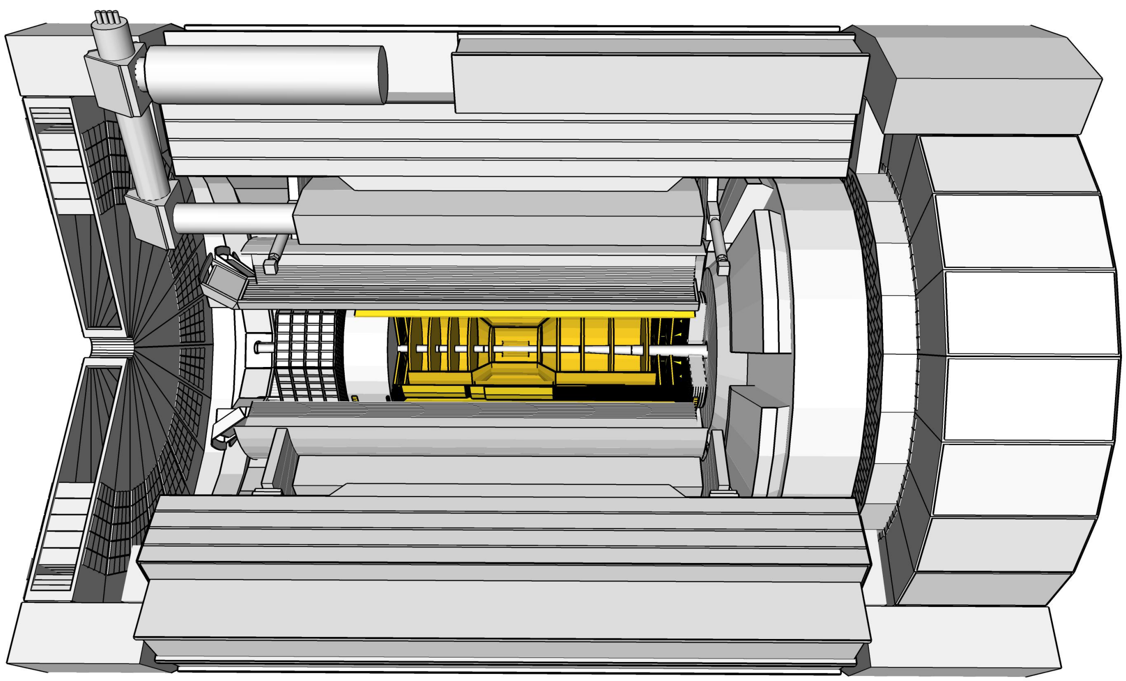


HOW DO WE DO THIS MEASUREMENT?

Particles: e-, e+, π+, π-, K+, K-, p, μ-, μ+

CHARGED PARTICLE MOMENTUM AND VERTEX

Tracking detectors in a solenoidal magnetic field



Principle:

- Charged particles curl in the magnetic field of the detector and leave hits in the tracking detector
- The radius of the particle trajectory fully constrains its momentum

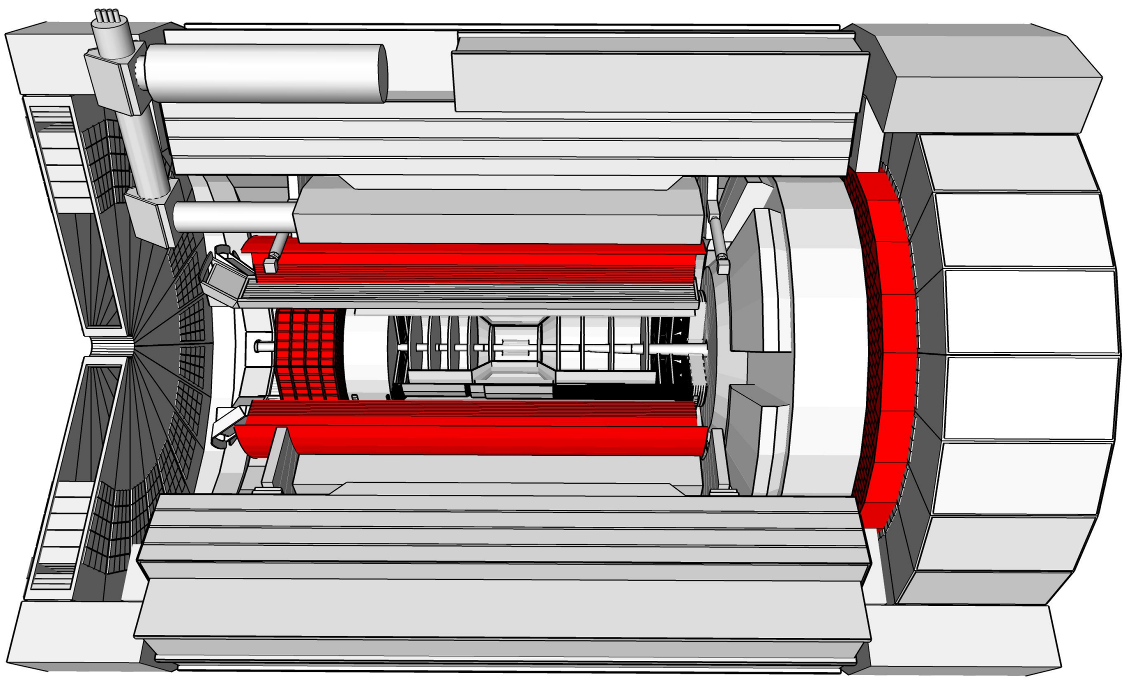
Considerations:

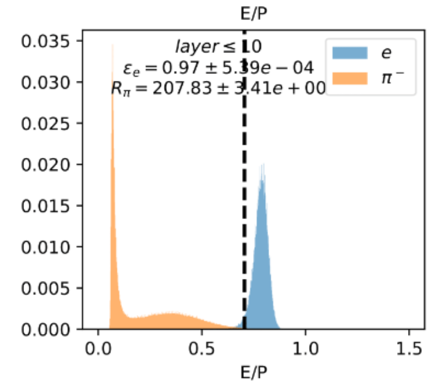
- Neutral particles do not leave hits in the tracking detectors
- Sensitive to material effects (multiple scattering and radiation in detector material worsens resolution)

Particles: e-, e+, π+, π-, (K+, K-, p, μ-, μ+)

ELECTRON-PION SEPARATION

Mostly carried by the electromagnetic calorimeters





Principle:

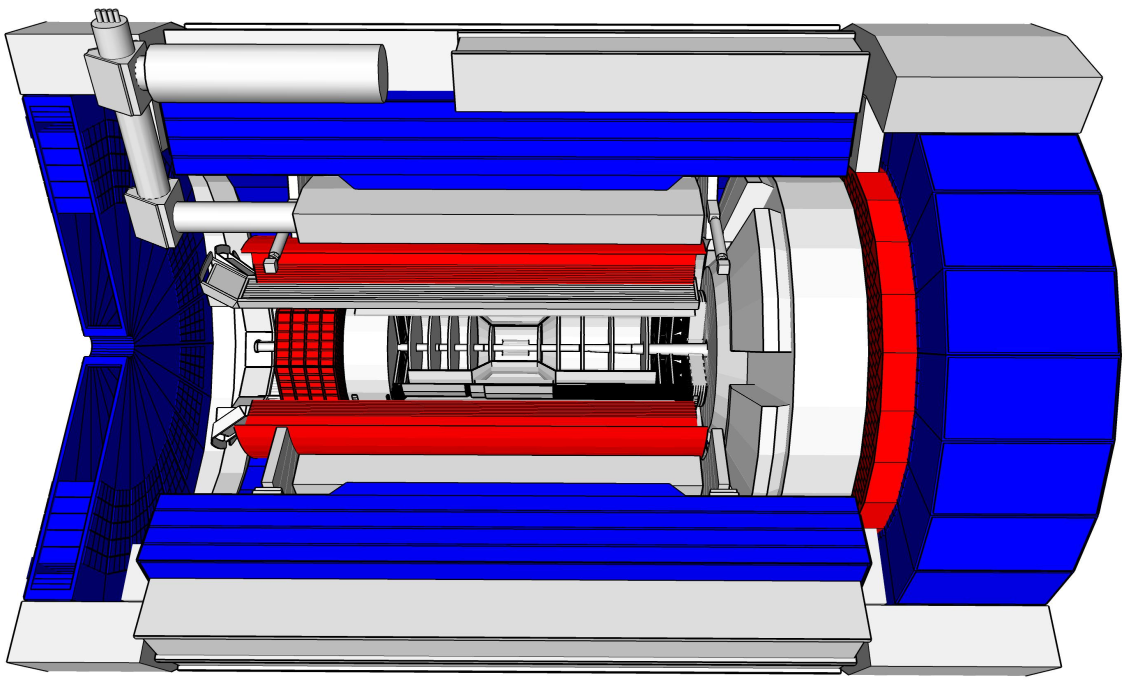
- Electrons lose all their energy in the electromagnetic calorimeters, while pions only sometimes loose all their energy.
- Electron showers and pion showers look different.

Considerations:

- Electron shower tails overlap pion shower tails: source of irreducible contamination
- Very high amounts of pion rejection (10^3 - 10^4) needed to achieve reasonable levels of background

NEUTRAL PARTICLE MEASUREMENTS

Combined EM and Hadronic Calorimeters



Principle:

- Neutrals travel in a straight line through the detector
- Neutrals deposit their energy in the combined calorimeter system
- Can use known event vertex and calorimeter information to reconstruct 3-momentum
- Photons (and pi0) fully measured by EM calorimeters

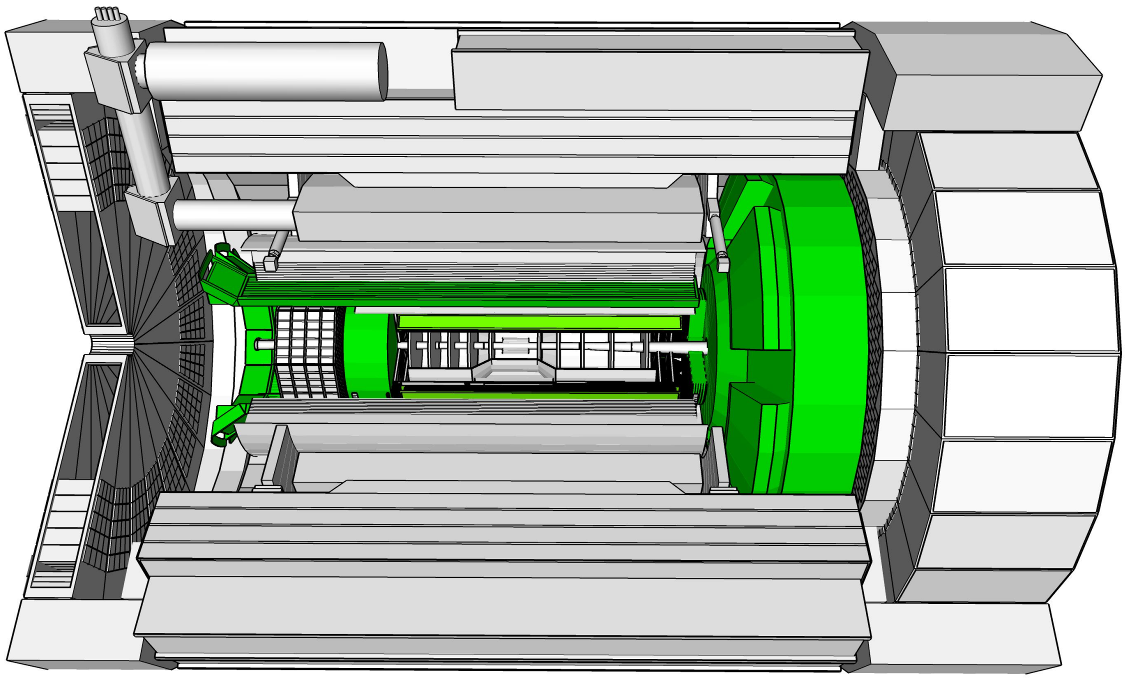
Considerations:

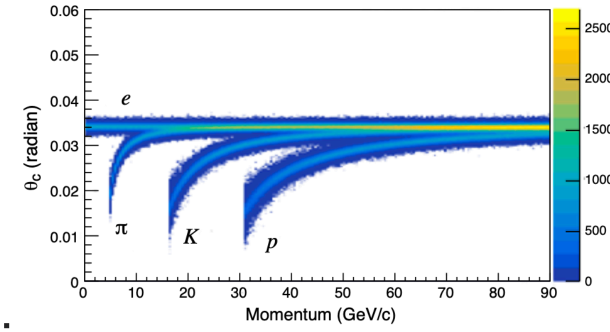
- Hadronic calorimeter in some cases only sees the tail of the energy distribution (e.g. magnet material in-between)
- Worse position resolution (intrinsic loss of precision in calorimeter showers)
- Difficulty separating particles close to each other.

Particles: π+, π-, K+, K-, p, (e-, e+, μ-, μ+)

CHARGED HADRON PID

Mostly through Cherenkov detectors



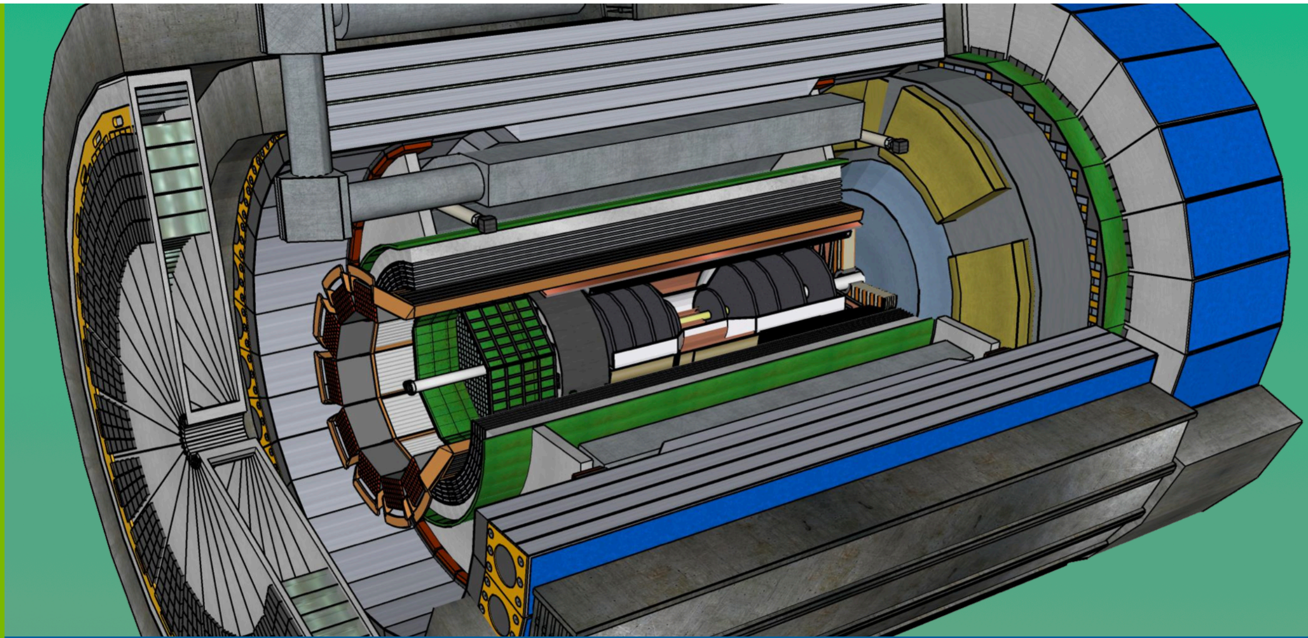


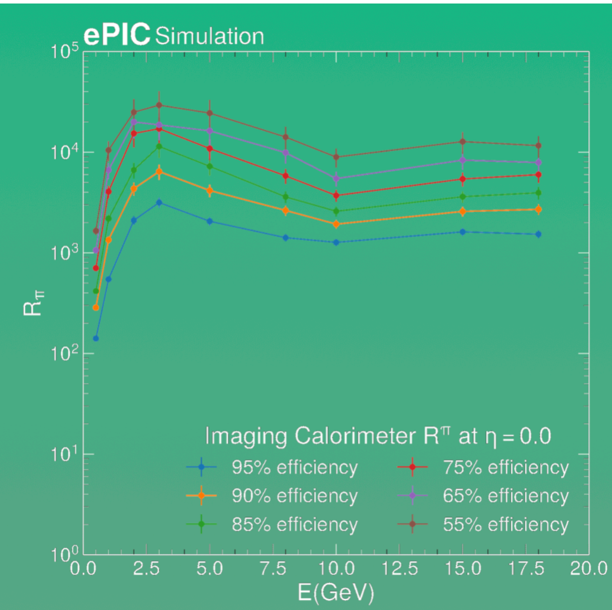
Principle:

- Charged particles traveling through a medium at a velocity faster than the speed of light in this medium emit Cherenkov radiation in a cone
- The opening angle of this cone is related to the particle's velocity
- Combining this velocity with the momentum from tracking detectors gives a handle on mass (and hence PID)

Considerations:

- Very sensitive to tracking resolution
- Only sensitive in limited momentum range
- Assigning detected Cherenkov rings to reconstructed trajectories non-trivial
- Non-zero probability of misidentification

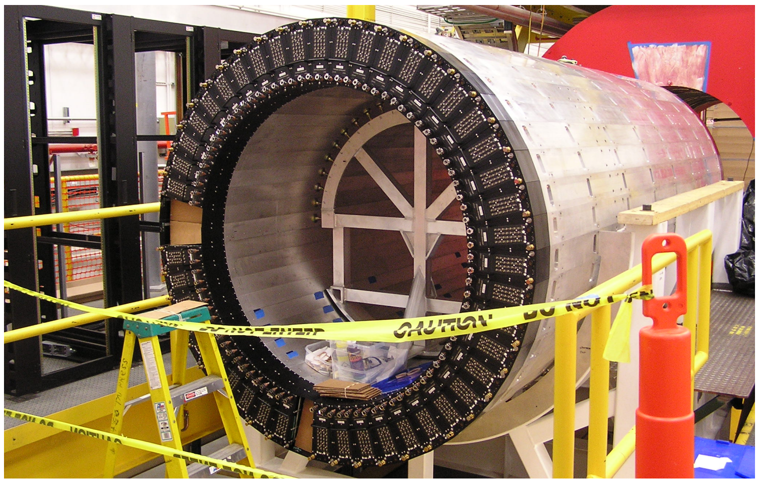




THE PROBABILISTIC NATURE OF PID

THE EPIC BARREL IMAGING CALORIMETER

Optimized for electron-pion separation by combining a high-performance sampling calorimeter with inexpensive silicon sensors for shower profiling



- Start from mature layered Pb/ScFi technology with side-readout (same as the GlueX calorimeter) for state-of-the-art sampling calorimeter performance

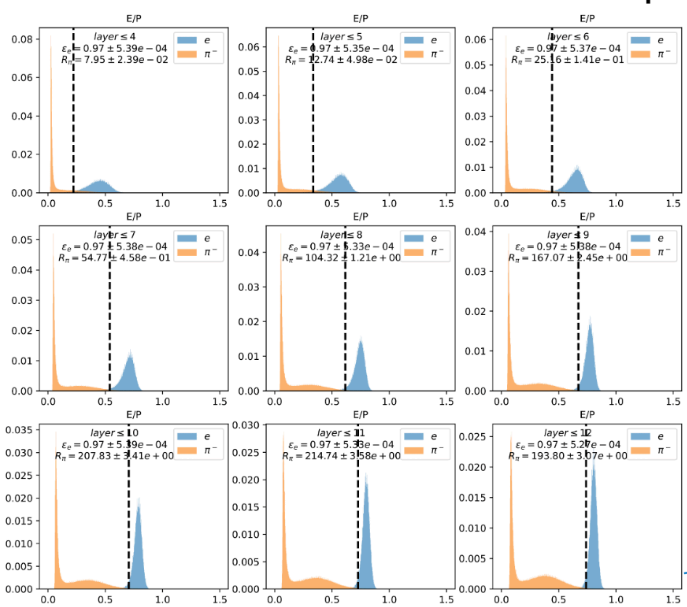
- Insert layers of monolithic AstroPix sensors (inexpensive ultra-low-power silicon sensor developed for NASA) in the first half of the calorimeter to capture a 3-D image of the developing shower

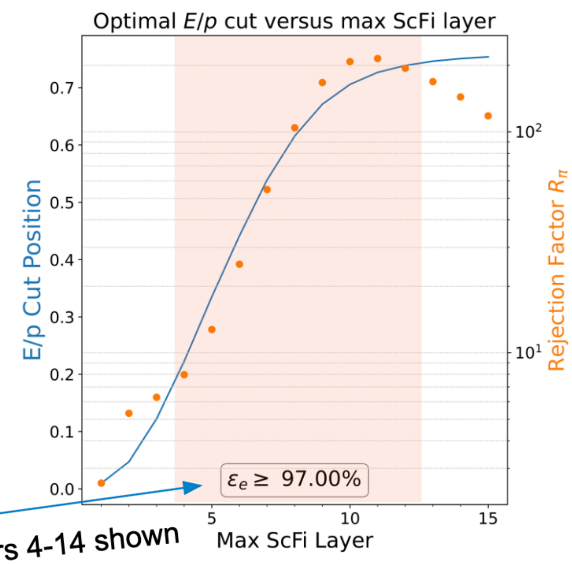
E/Π SEPARATION - EXAMPLE METHOD

Steps:

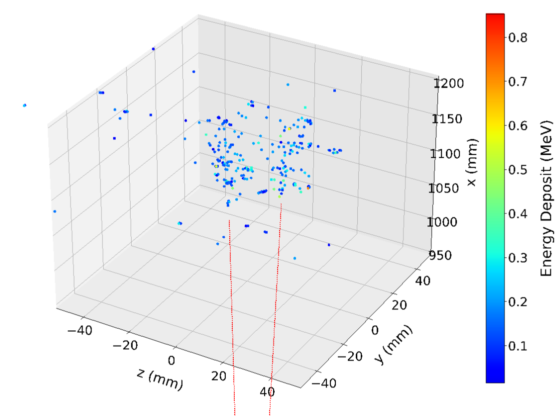
1. **Optimized cut on E/p** from different depth of Pb/ScFi layers at very high electron efficiency
2. **AI to leverage 3-D shower evolution** “pictures” to optimally classify electrons and pions (our proof-of-concept used CNNs, now working to move to GNNs or PointNet-like architectures

Example for 2 GeV e/π

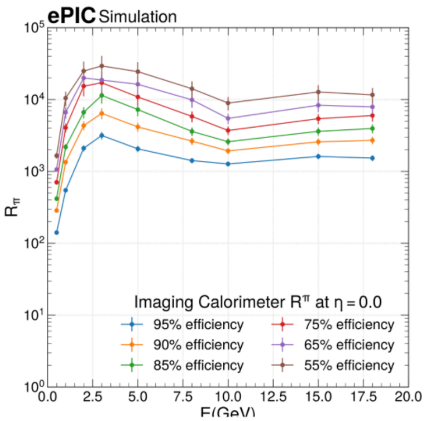




ScFi layers 4-14 shown

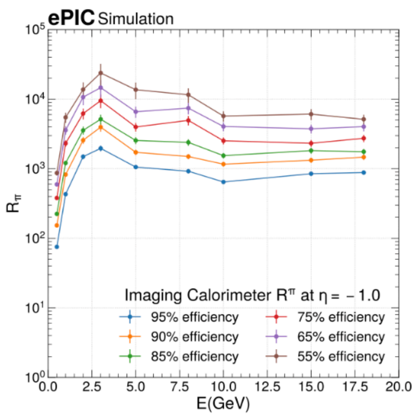


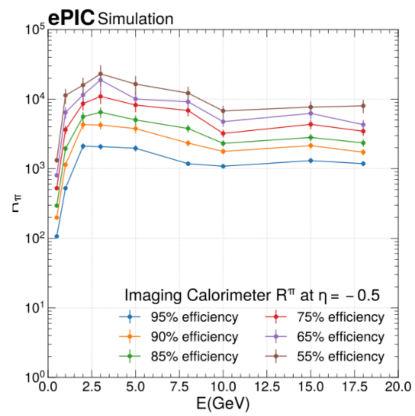
E/Π SEPARATION - EXAMPLE RESULTS

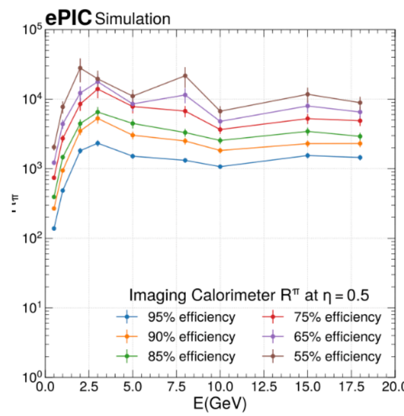


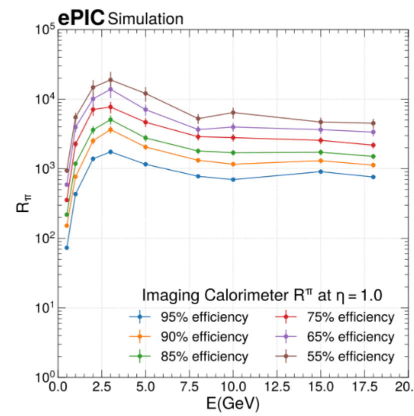
e/π separation - η, energy and efficiency dependence

- Results depend strongly on electron efficiency
- **For desired 95% efficiency for all η regions we are ≥ 10^3 above ~ 1.5 GeV**
- Responses at different energies and η have been folded into the purity studies

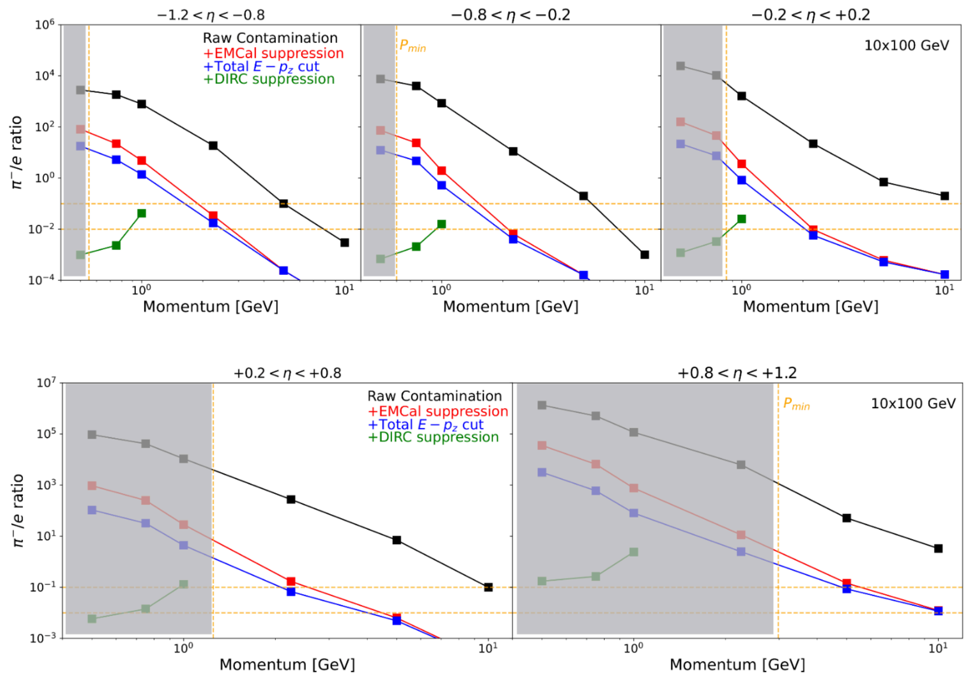








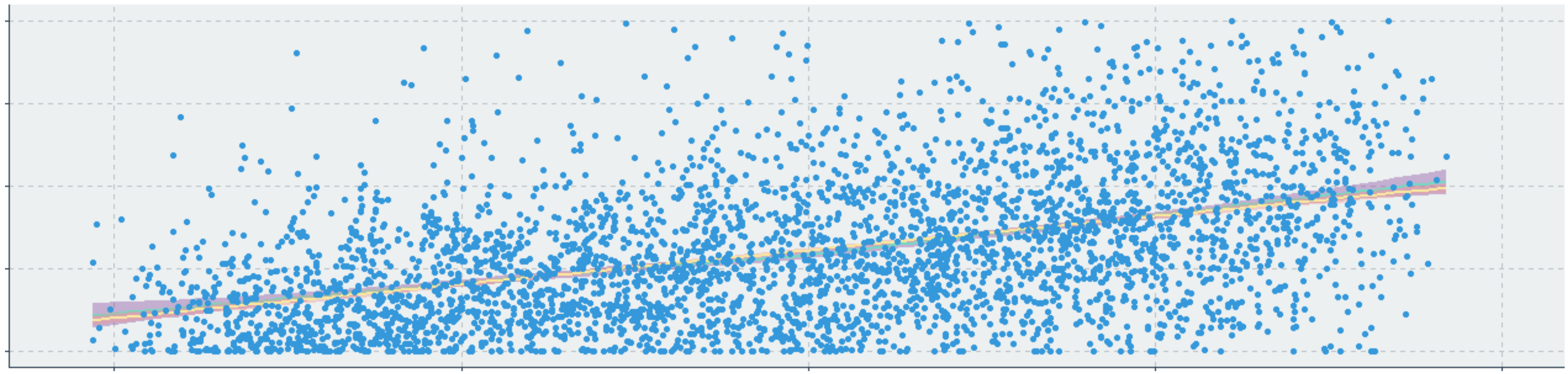
E/Π SEPARATION - EXAMPLE RESULTS



Challenging goal: at least 90% electron purity everywhere

However, this means there are regions were 10% of our “electrons” are really pions!

Not all of these will be problematic (i.e. reconstruct as the most likely primary electron), but some effects unavoidable.



TOWARDS AN UNBINNED APPROACH

TYPES OF EFFECTS

What did we measure that we shouldn't have measured?

1. Misidentification: We detected an electron but really this was a pion (or anything else)
2. Ambiguity:
   a. We detected an electron, but this is not a scattered electron (e.g. a decay electron from a neutral pion).
   b. We detected a scattered electron but it is not from a DIS event (e.g. elastically scattered electron)
   c. Same as b but it's really from accelerator background

Give events an associated weight?

Subtract estimated (or measured) fraction of events) → negative weights?

TYPES OF EFFECTS

What did we miss (or almost miss)?

1. Misidentification: We classified our electron as a pion
2. Acceptance:
   a. Our electron fell in a gap between sensors so we did not see it. Hopefully there wasn't anything else that looked like an electron!
3. Smearing:
   a. Radiative effects, e.g. Our electron radiated a hard photon which changed its four-momentum. Now it does not look like a scattered electron anymore
   b. Our detector did not reconstruct the correct four-vector for the electron (e.g. our tracking algorithm gives us a lower momentum by adding some erroneous detector hits to our track).

Classically: Acceptance correction. Need to ensure detector model does this properly.

Classically: unfolding. Need to ensure detector model does this properly

SYSTEMATIC UNCERTAINTIES

- Uncertainty on the normalization (“luminosity”)
  - We measure counts/# potential interactions. There is an uncertainty on the denominator
- Uncertainty due to exact size of background contributions
  - E.g. How certain are we of the probabilities assigned to our particle identification?
- Uncertainty due to changing detector performance over time
  - Maybe some detector started to degrade due to radiation?
- Uncertainty due to knowledge of detector limitations.
- Uncertainties related to doing corrections on binned data
  - E.g. model uncertainty for unfolding, etc.
- Uncertainty due to limitations in our detector modeling
- … Many more depending on the exact experimental setup!

Bottom line: many systematic uncertainties *could* require us to process the workflow multiple times

Classically: Scale uncertainty
Event-level: ???

Classically: Change background model and look at impact on results
Event-level: Same?

Classically: Evaluate key metrics over time. Try to correct (weigh) data, residual effect becomes uncertainty.

Classically: Fluctuate model and look at impact on results.
Event-level: Same?

DISCUSSION