中国科学院高能物理研究所
Institute of High Energy Physics
Chinese Academy of Sciences

# Pattern recognition at CEPC AHCAL Prototype using test beam data at CERN

**Xin Xia, Yuzhi Che**

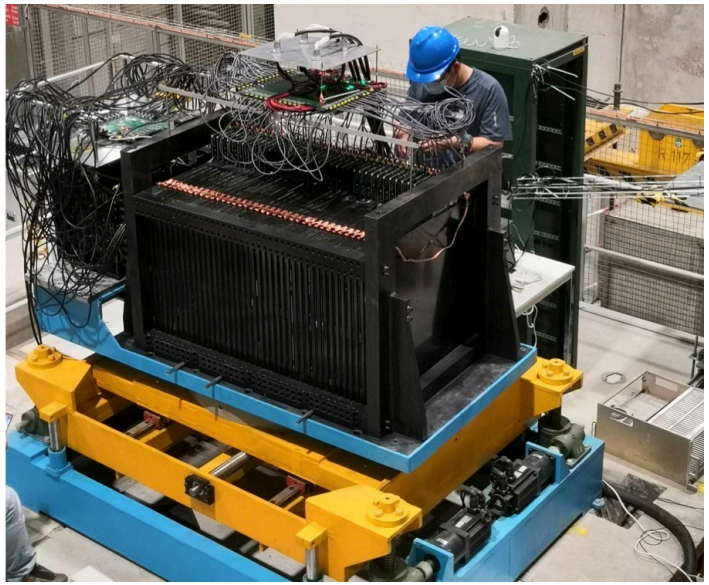
**2023/9/6**

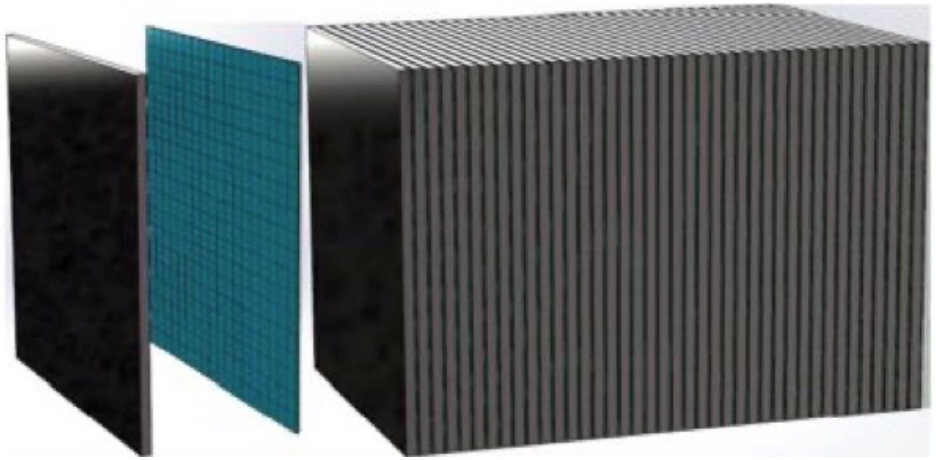
# Outline

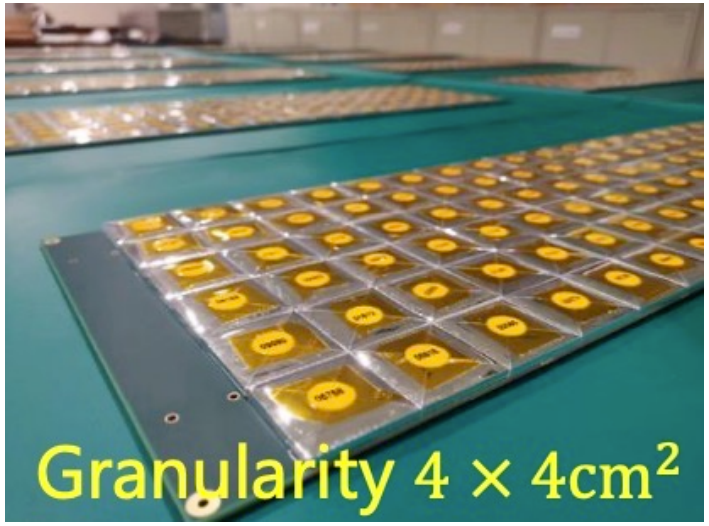
- **Introduction**
  - AHCAL Prototype Configuration
  - Test Beam Data Set
  - Motivation
- **How to classify different data components?**
- **Data Component analysis**
- **Component Fraction and Comparison**
- **Energy Performance**

# AHCAL Prototype Configuration

- **CEPC AHCAL prototype commissioned**
  - Active material: scintillator tiles
  - Absorber: stainless steel
  - Transverse size 72×72 $\mathrm{cm}^2$, 40 longitudinal layers (~4.6$\lambda$0)
  - Granularity $4 \times 4\ \mathrm{cm}^2$ , 12960 readout channels
  - ~5 ton in weight
  - Developed during 2018 – 2022







# Test Beam Data Set

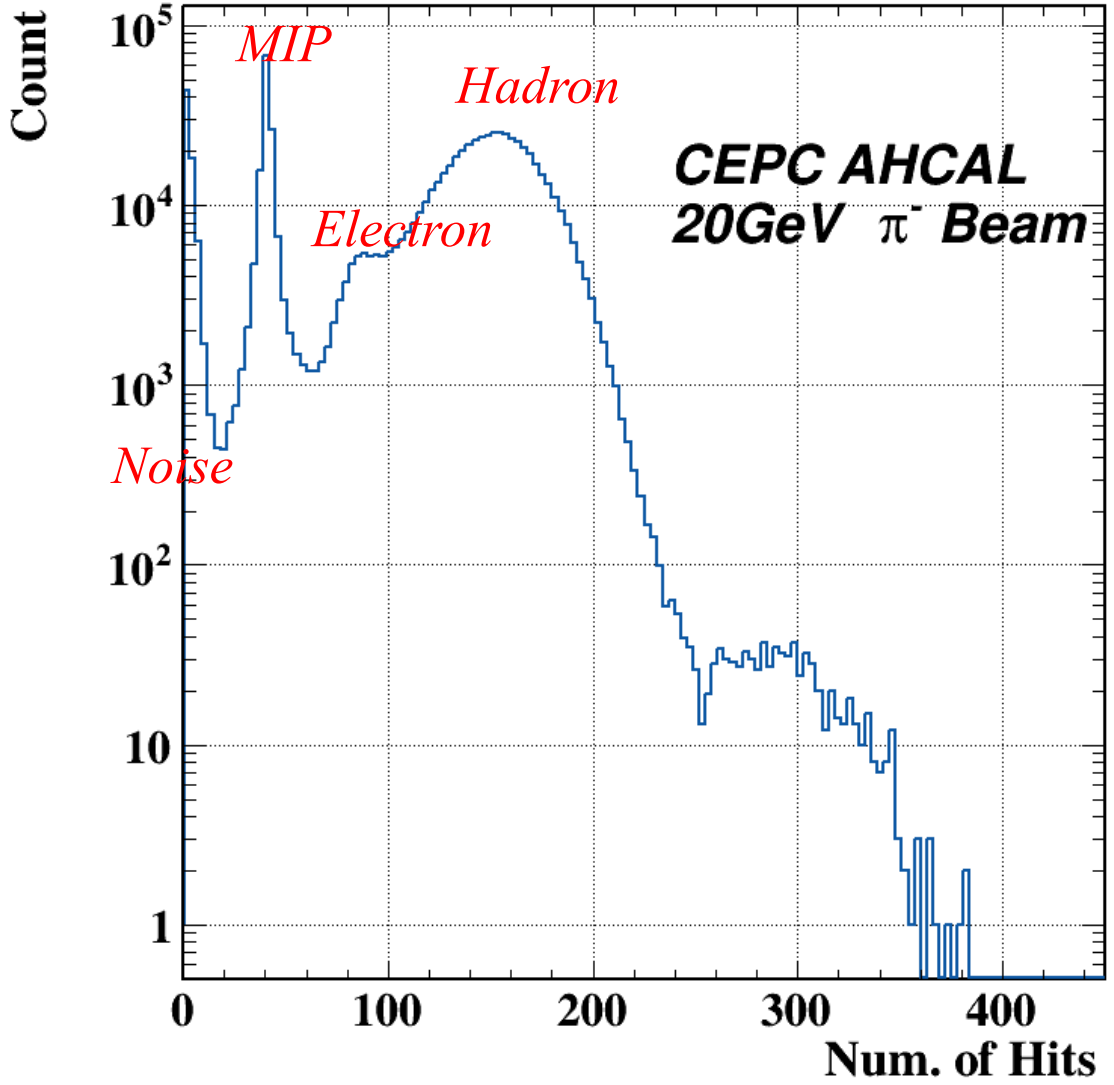
- Beam test at CERN at SPS H2 beamline at 2023/05
  - Data Taking:
    - ~ 11 billion pion from 10GeV to 120GeV
    - ~ 4 billion electron from 10GeV to 250GeV

| AHCAL Data List of Pion | | |
|---|---|---|
| Momentum(GeV) | Total Runs | Number of Events |
| 10 | 6 | 1027814 |
| 15 | 4 | 953774 |
| 20 | 4 | 749791 |
| 30 | 4 | 1081725 |
| 40 | 5 | 1038984 |
| 50 | 5 | 1258699 |
| 60 | 6 | 1034938 |
| 70 | 5 | 1058822 |
| 80 | 7 | 1073061 |
| 100 | 7 | 1286608 |
| 120 | 6 | 1036400 |

| AHCAL Data List of Electron | | |
|---|---|---|
| Momentum(GeV) | Total Runs | Number of Events |
| 10 | 2 | 238139 |
| 20 | 1 | 573677 |
| 30 | 2 | 556570 |
| 40 | 1 | 421484 |
| 50 | 2 | 743810 |
| 60 | 1 | 377014 |
| 70 | 3 | 580935 |
| 80 | 1 | 273278 |
| 100 | 1 | 289511 |
| 120 | 1 | 242049 |
| 150 | 2 | 415991 |
| 250 | 1 | 121850 |

# Motivation

- **The data contains:**
  - Noise,
  - MIP, Electron, Hadron
  - Multiple injection events
  - …
- **In order to:**
  - Fully understand data components
  - Further quantify the detector respons
- **Pattern Recognition:**
  - Event classification
  - Component Fraction Estimation



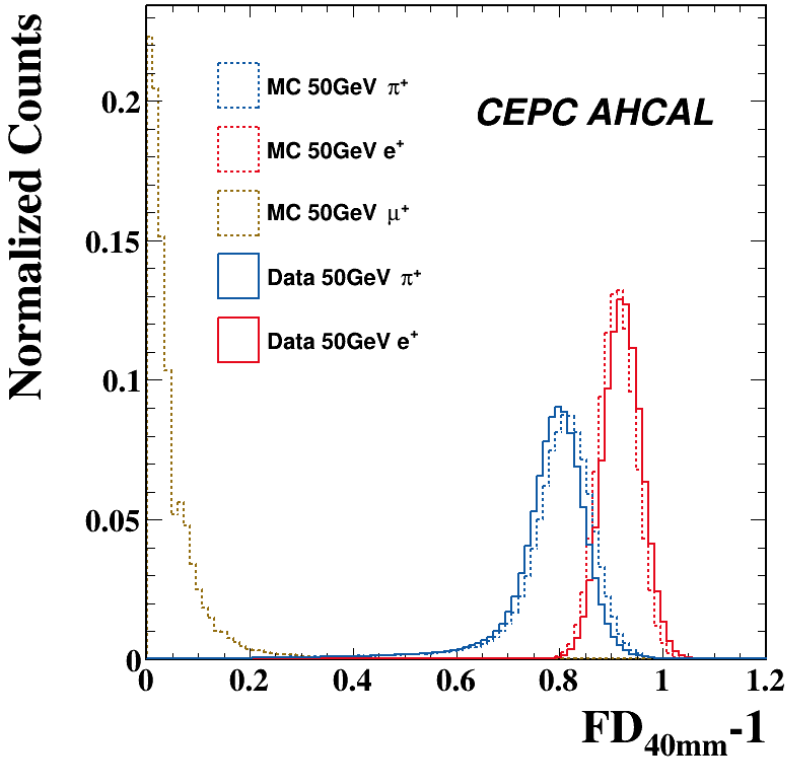
# Outline

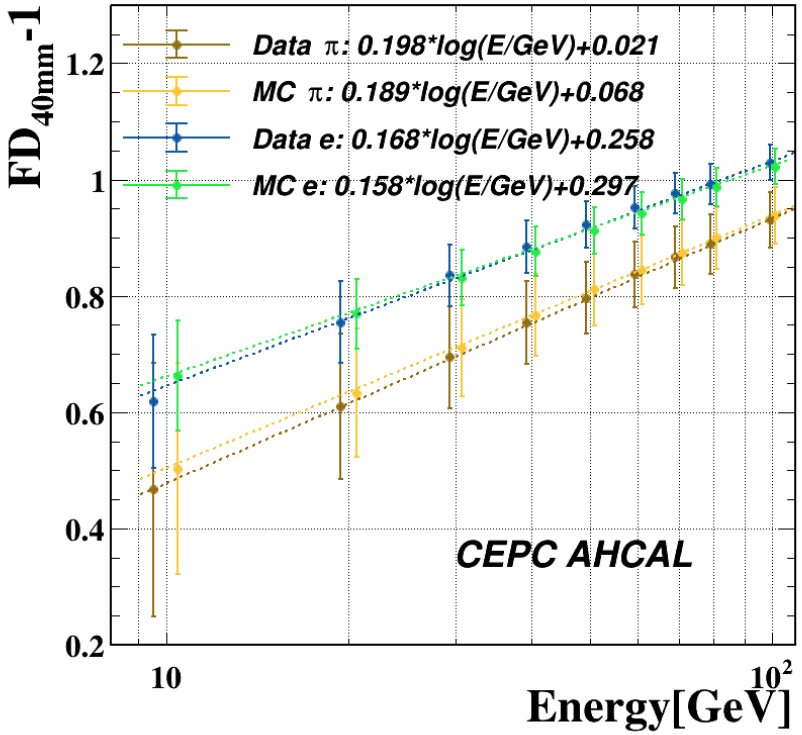
- Introduction
- **How to classify different data components?**
  - Fractal Dimension
  - Average Hit Energy
- Data Component analysis
- Component Fraction and Comparison
- Energy Performance

# Variables for PID

[1] M. Ruan et al, PRL 112, 012001

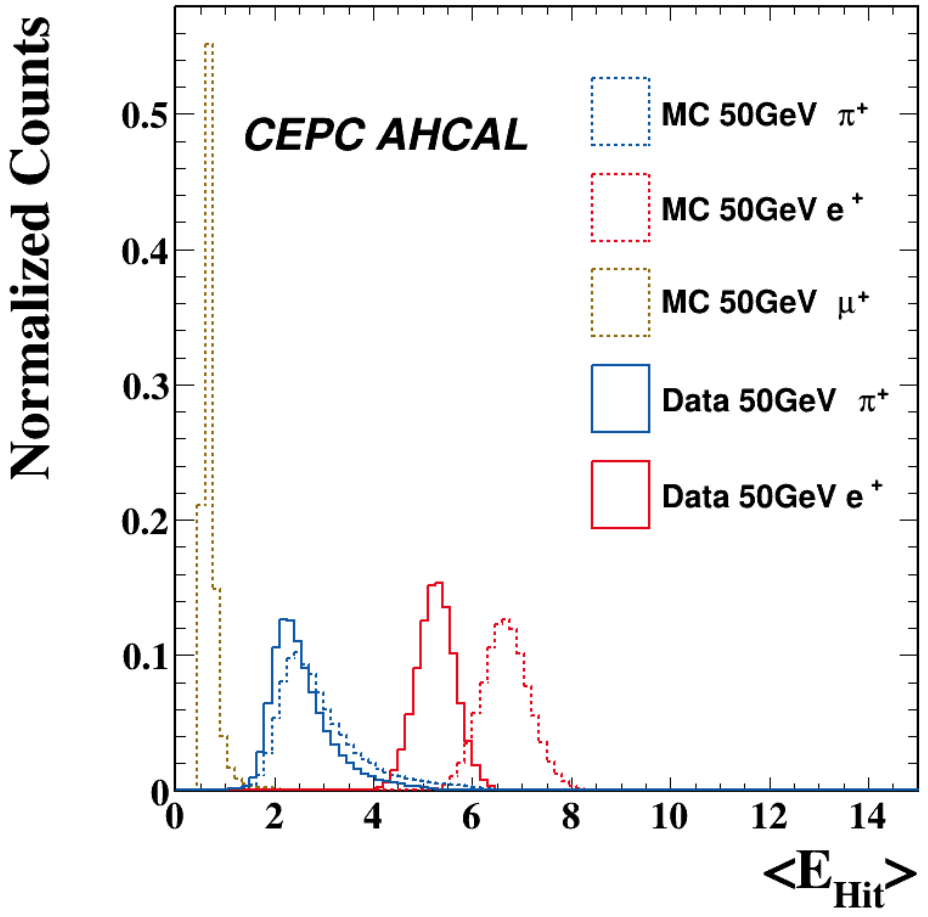
- **Fractal Dimension**[1] dedicated for high granularity calorimeter
  - $FD_{40mm} = \left\langle \frac{\log(R_{\alpha,40mm})}{\log(\alpha)} \right\rangle + 1$, $R_{\alpha,40mm} = \frac{N_{40mm}}{N_\alpha}$, $N_{40mm}$ is total number of hits, $N_\alpha$ is the number of hits at scale $\alpha$
  - MC/Data consistent



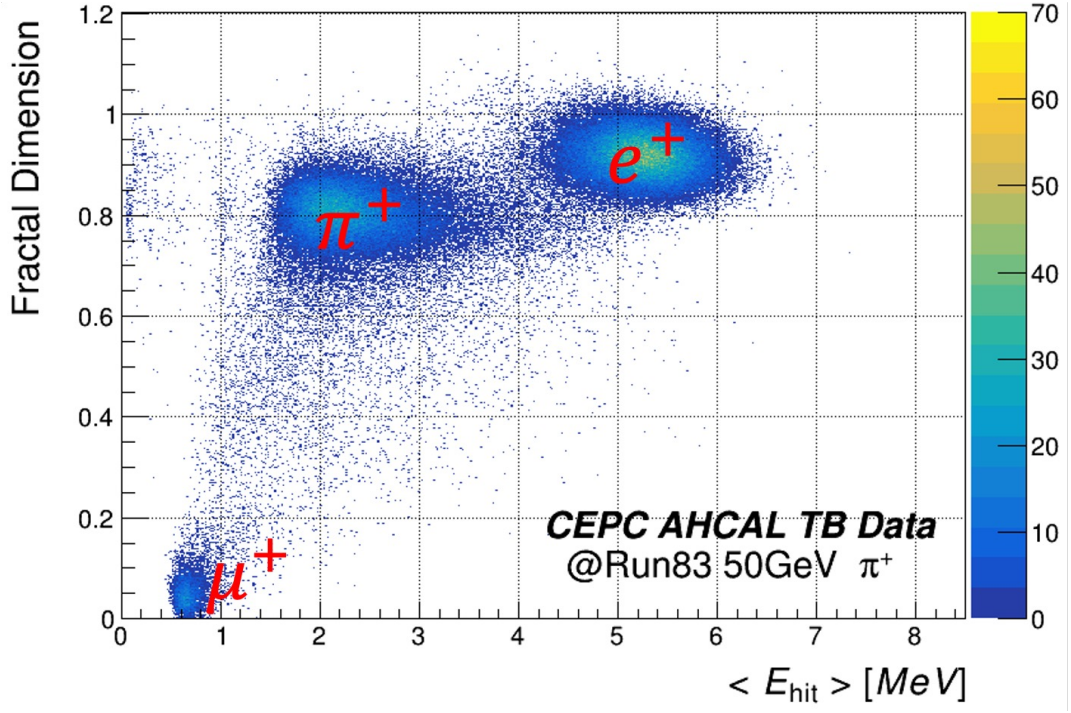
# Variables for PID

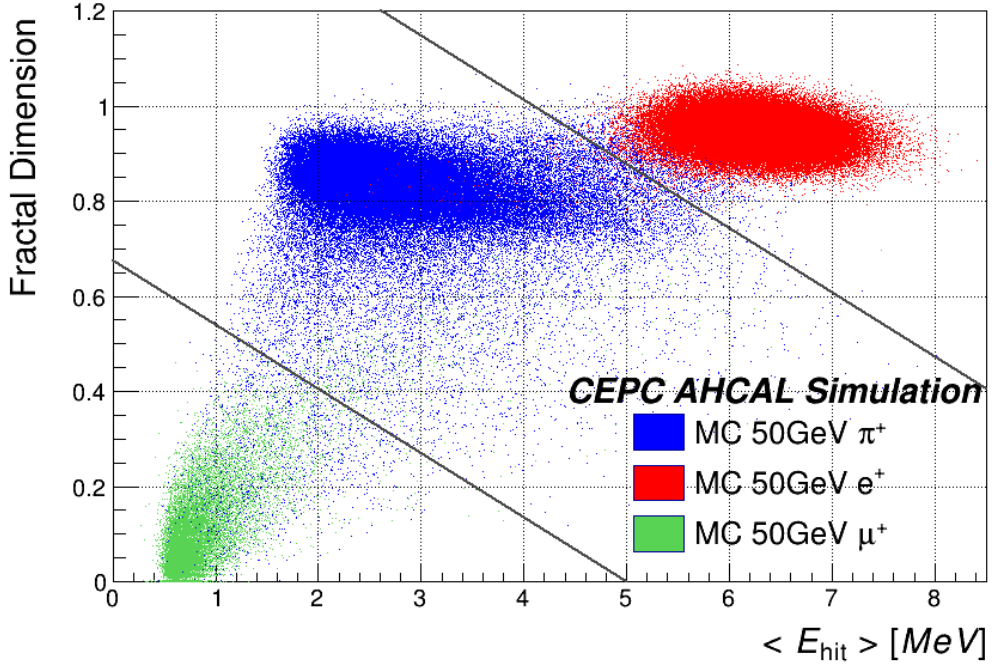
- **Average Hit Energy** :
  - Definition : $Event\ Energy / Num.of\ Hits$
  - MC/Data compare :
    - Distribution consistent
    - Value inconsistent



# FD-<$E_{Hit}$> Diagram

- MC/Data Compare
  - Distribution consistent
  - <$E_{Hit}$> Value inconsistent

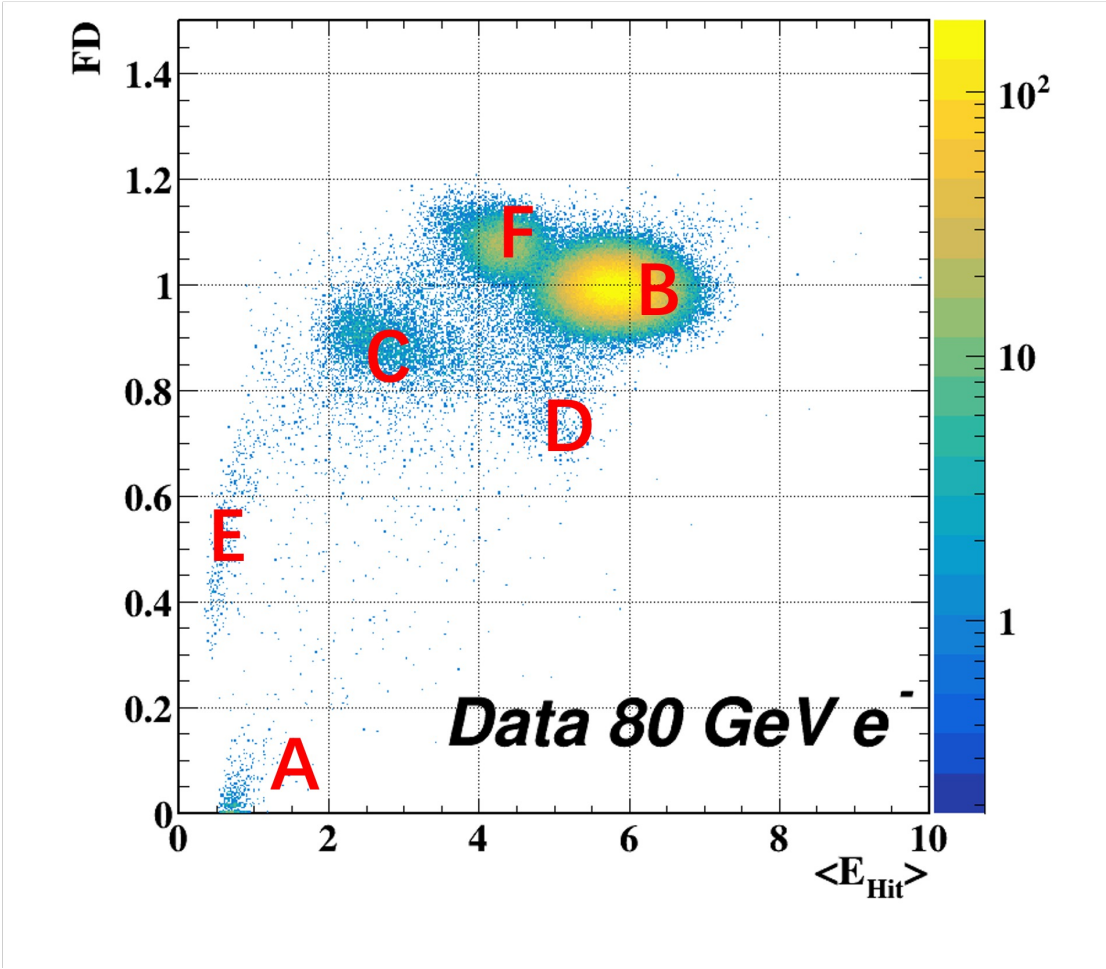


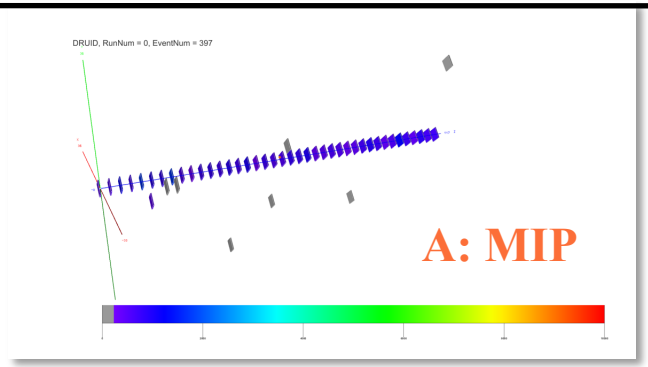



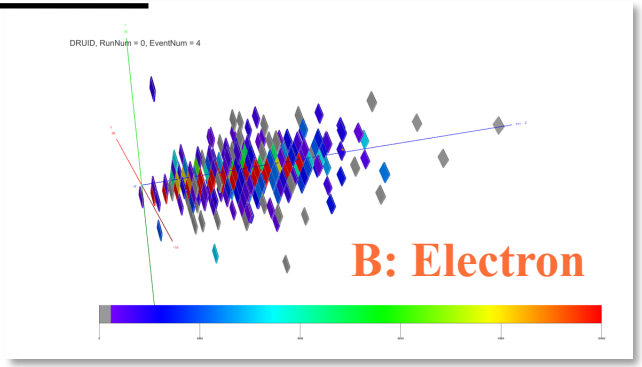
# Outline

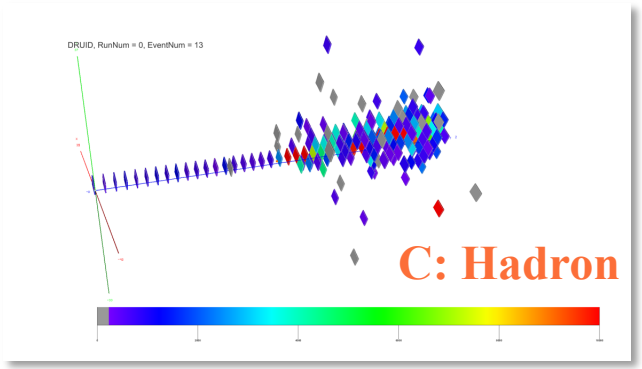
- Introduction
- How to classify different data components?
- **Data Component analysis**
- Component Fraction and Comparison
- Energy Performance

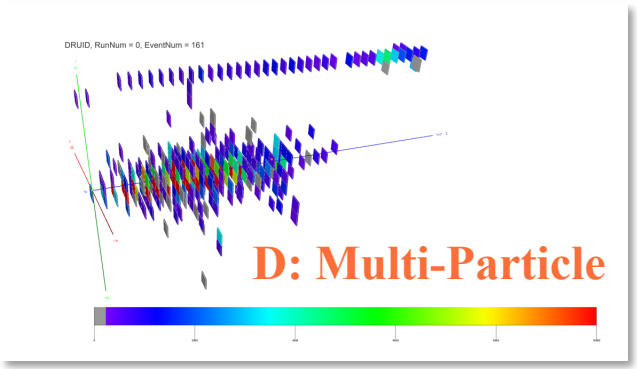
# Data Overview: 80 GeV electron as an Example

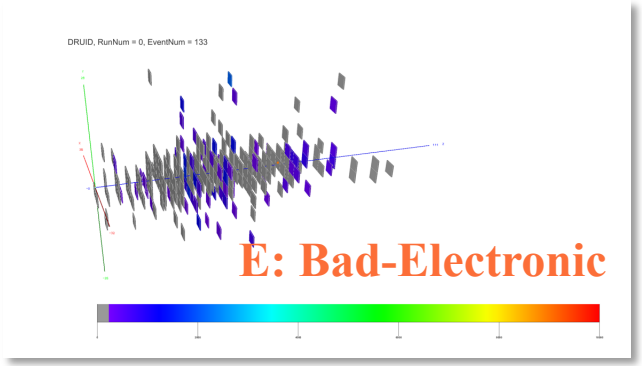


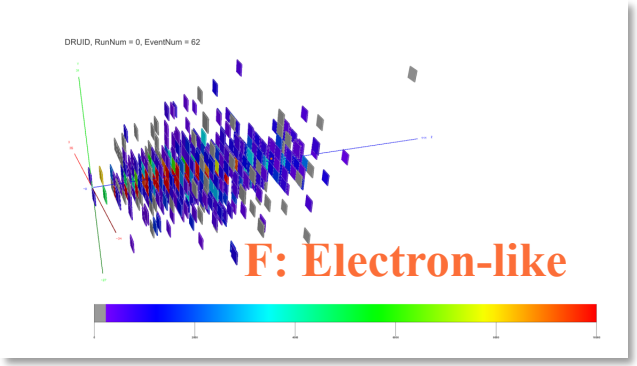



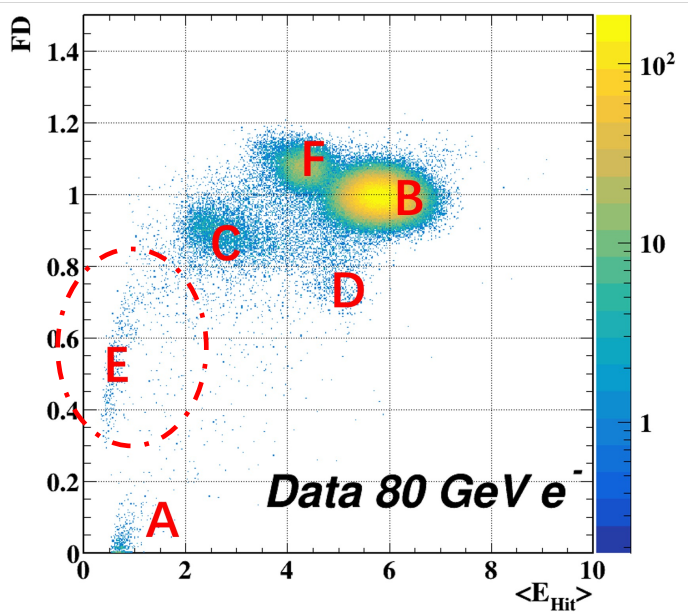



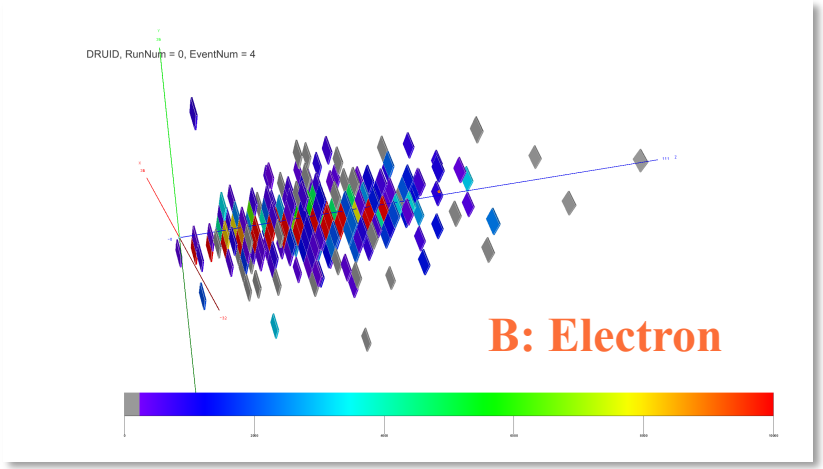



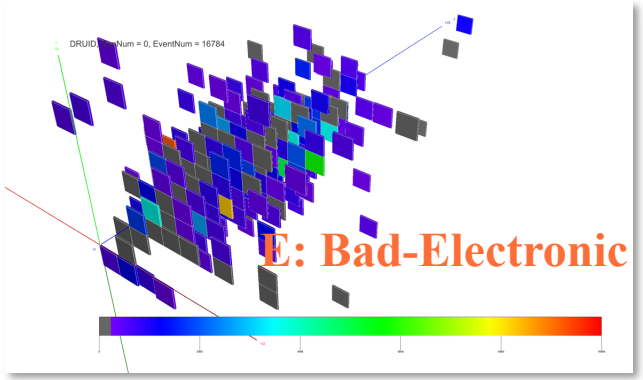
# Component Highlight: E

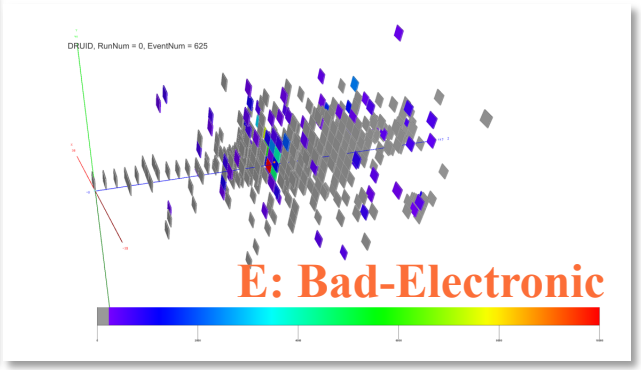
**Characteristics:**

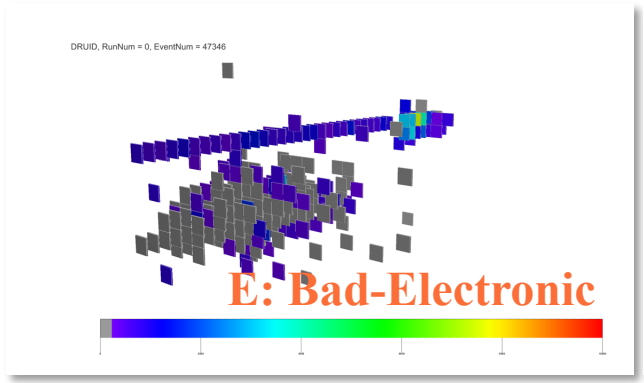
- Shape: Shower Like
- Energy: More low energy ( < 0.5 MIP) hits in inner region of shower
- Found in:
  - Electron + pion beams
  - SPS H8 (2022) + SPS H2 (2023)

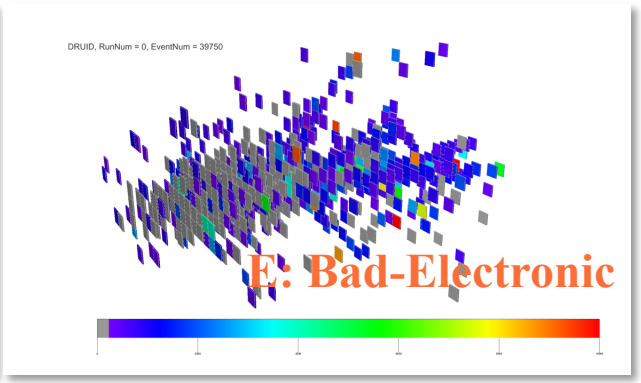


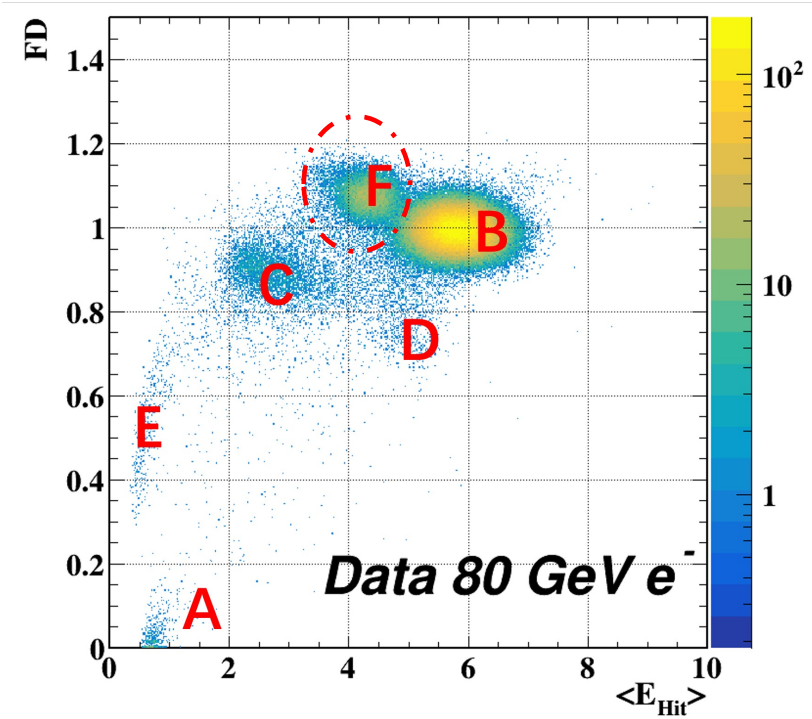



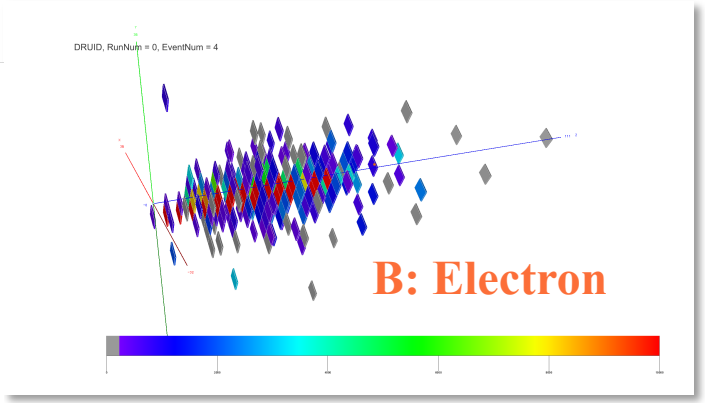



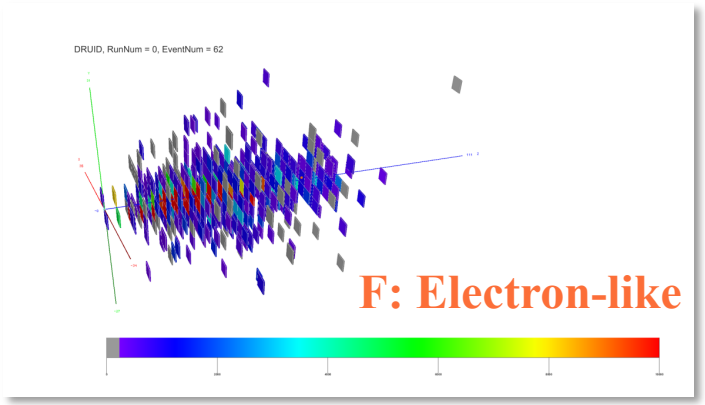
# Component Highlight: F

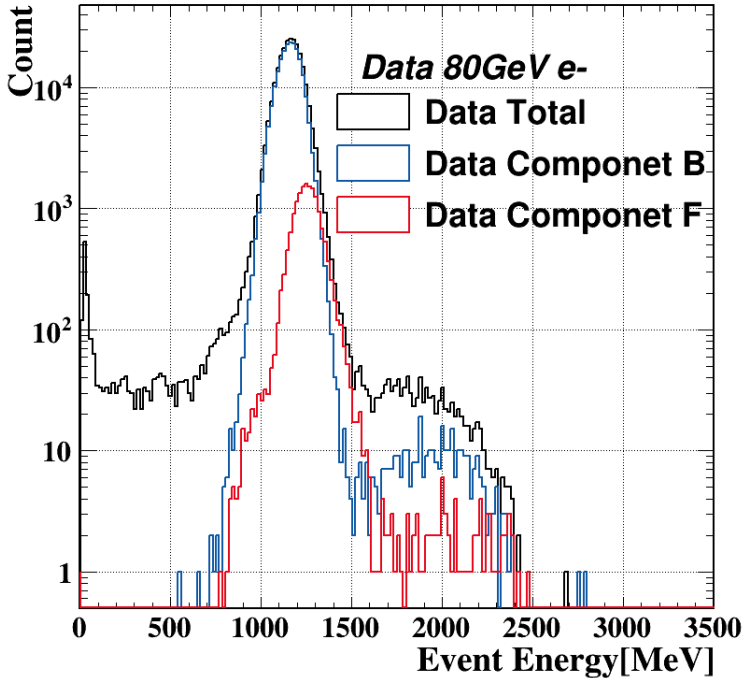
**Characteristics:**

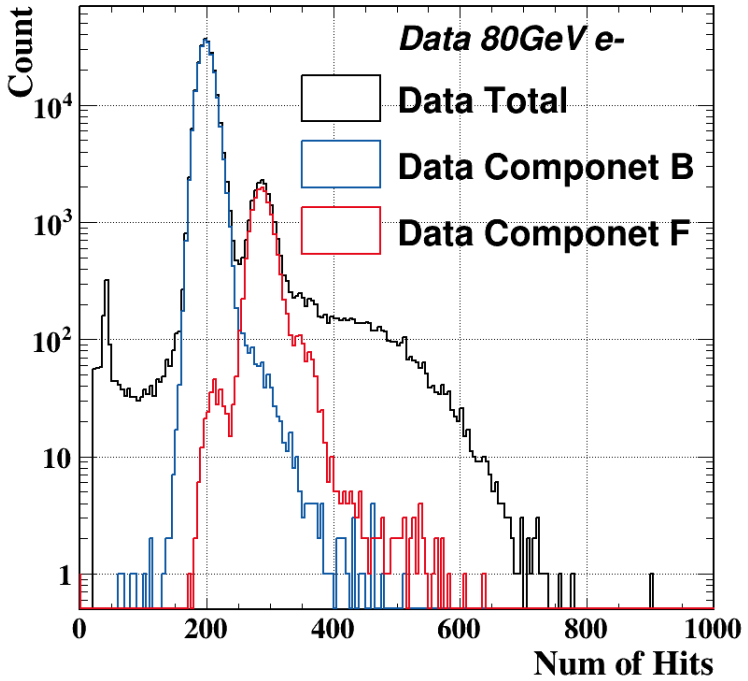
- Shape: EM. like
- More hits and higher energy deposition
- Only in electron beam taking at SPS H2 at 2023



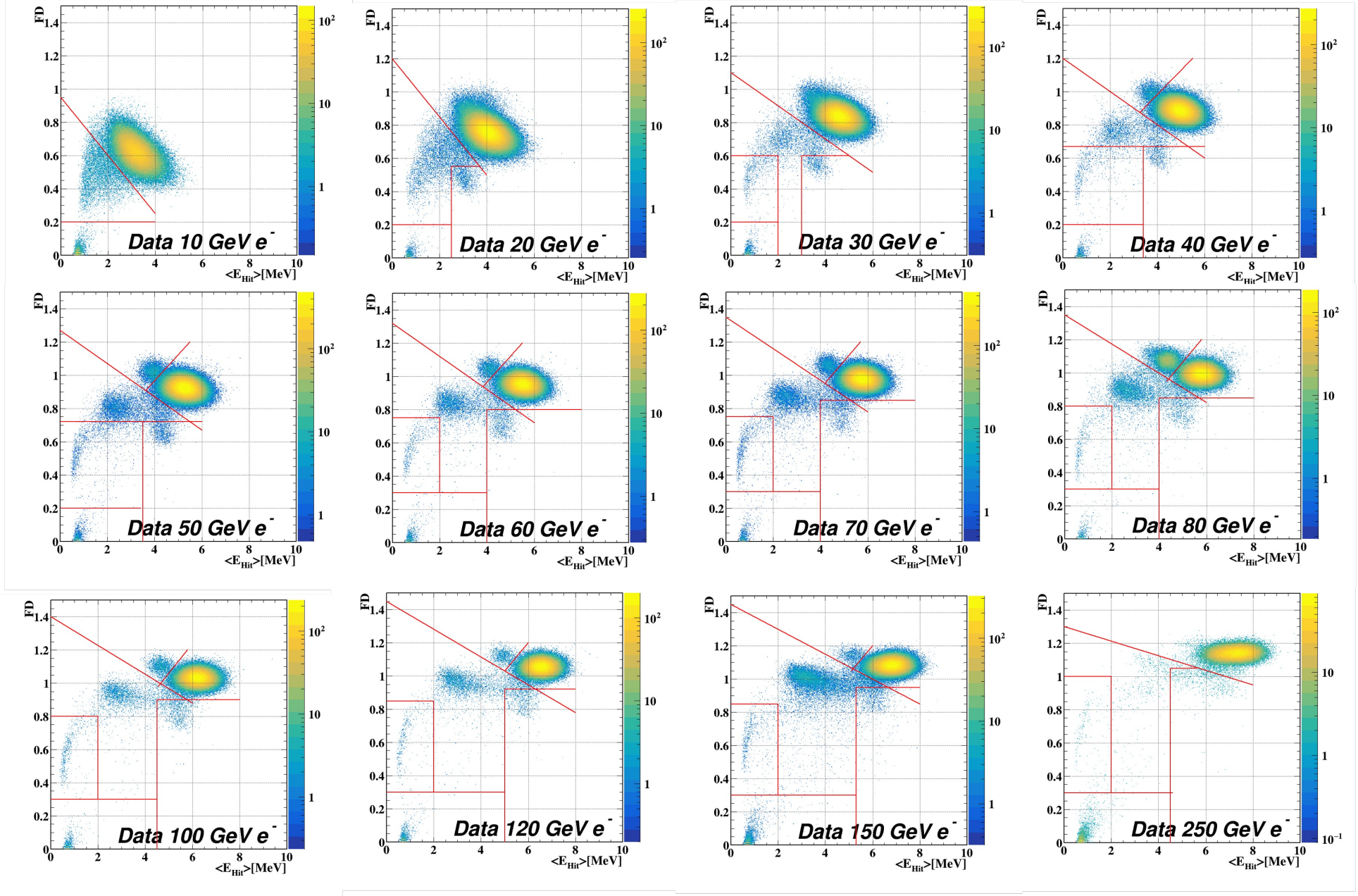



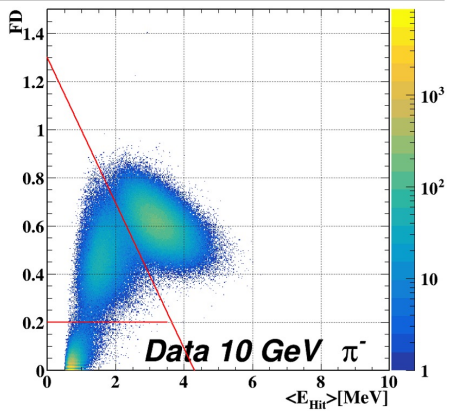



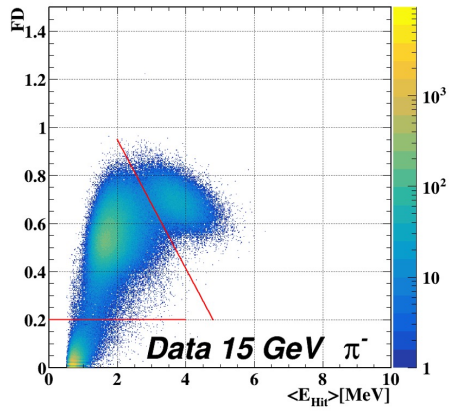



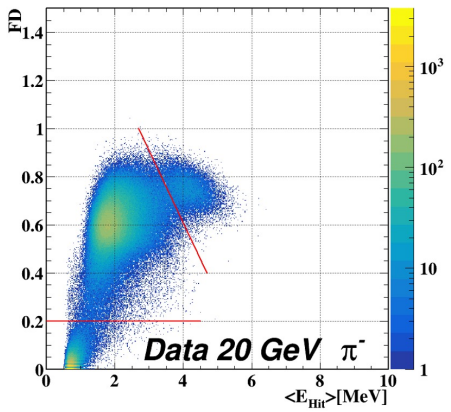
# Evolute versus Incident Energy: Electron Beam

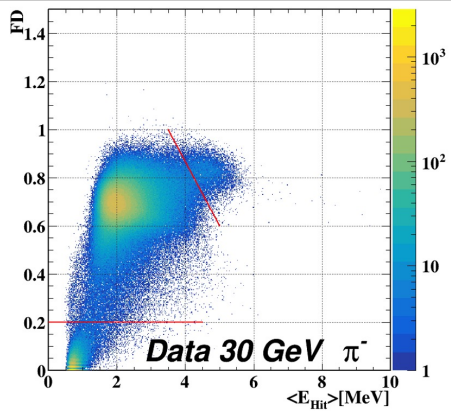


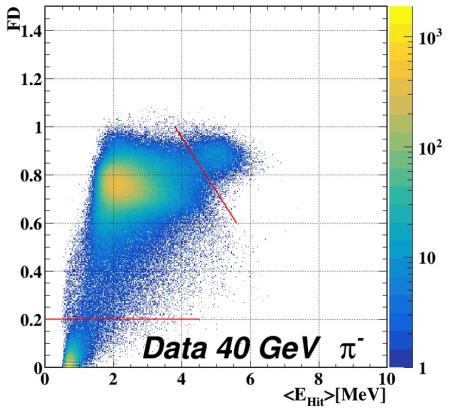
# Evolute versus Incident Energy: Pion Beam

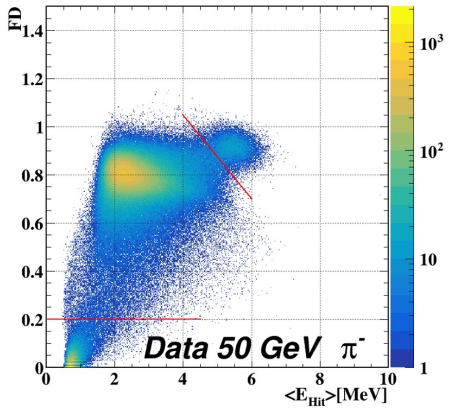


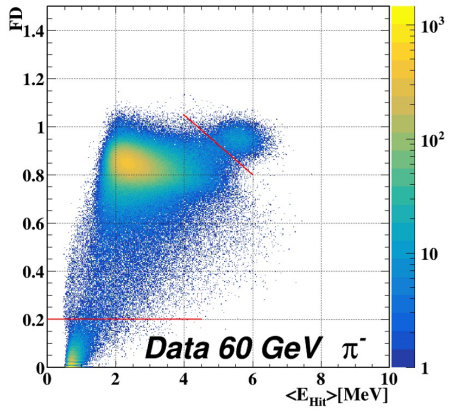



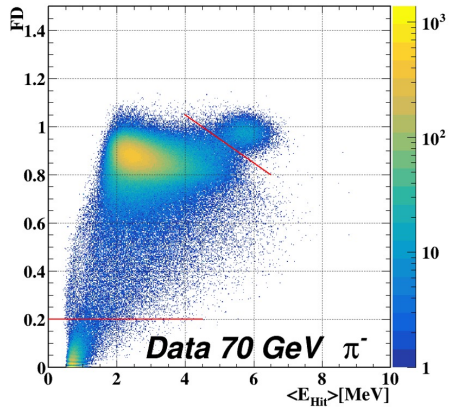



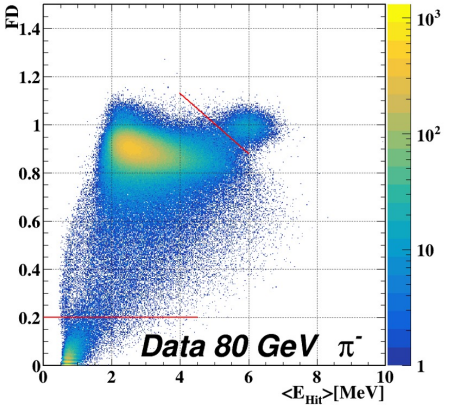



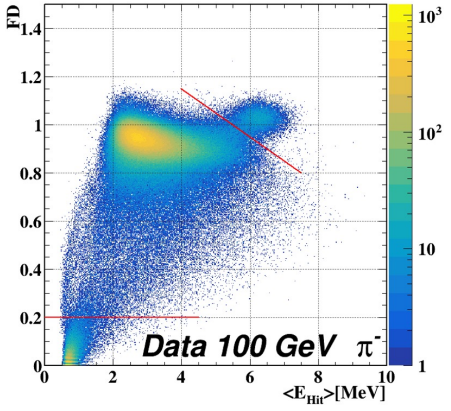



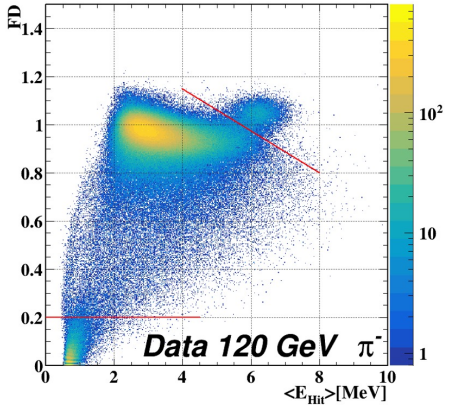



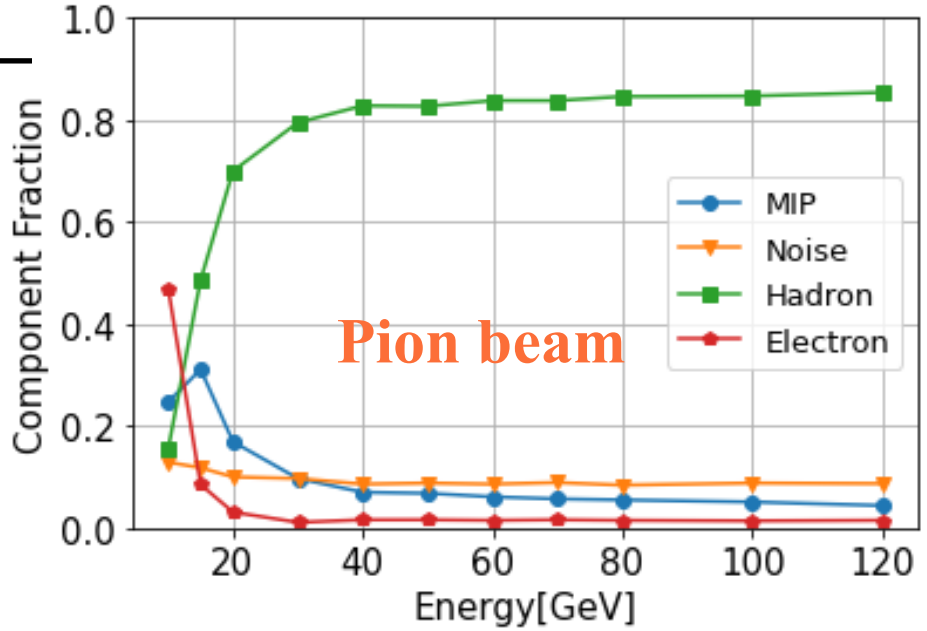



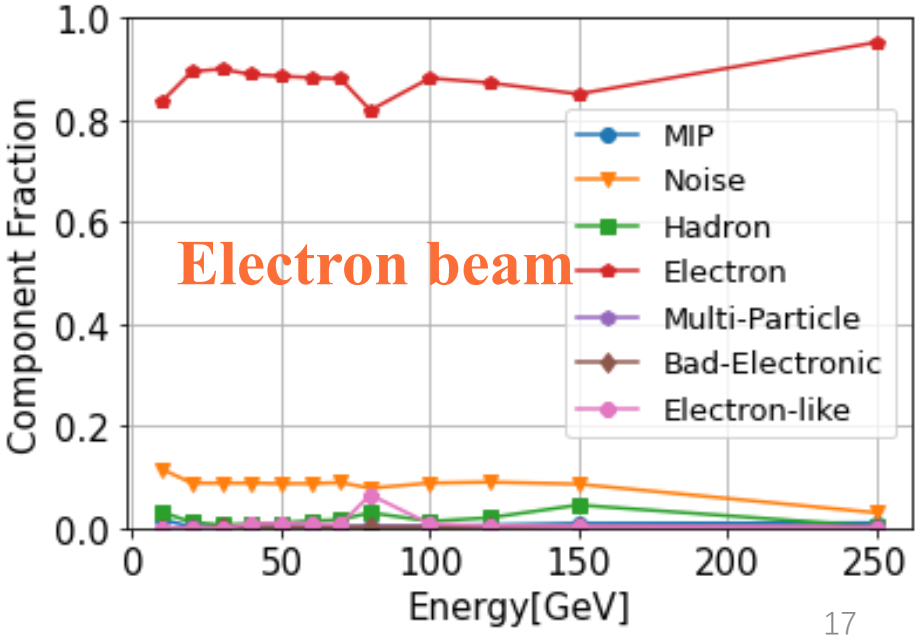
# Outline

- Introduction
- How to classify different data components?
- Data Component analysis
- **Component Fraction and Comparison**
- Energy Performance

# Component Fractions

- Beam purity:
  - Electron beam: purity > 80%
  - Pion beam:
    - increases with increasing energy
    - purity > 80%  when Energy > 30 GeV
    - Lots of electron events mixed in low-energy
- Noise is the dominant factor affecting beam purity.



# ANN Analysis

See Siyuan's Talk

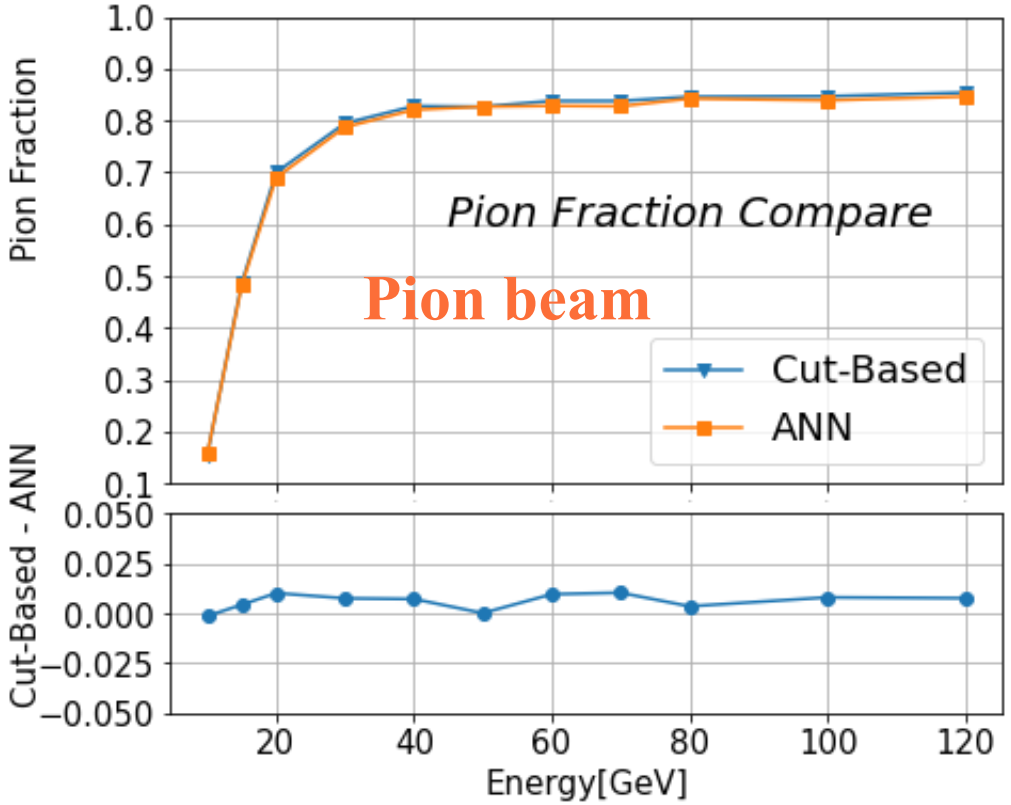
- New PID tool based on ANN developed by Siyuan
- Train set
  - MC
  - Data preselection by Cherenkov and FD-<$E_{Hit}$>

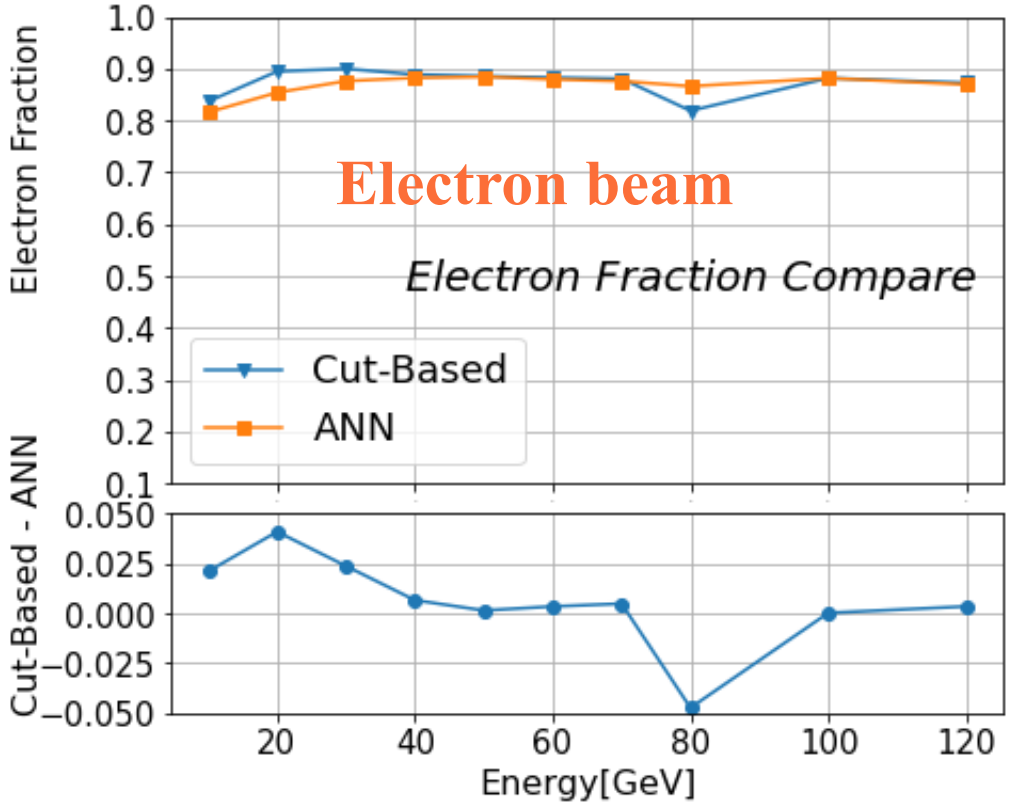
| Energy ID | 5GeV | 8GeV | 10GeV | 30GeV | 50GeV | 60GeV | 80GeV | 100GeV | 120GeV |
|---|---|---|---|---|---|---|---|---|---|
| Muon | 10k (MC) | 10k (MC) | 10k (MC) | 10k (MC) | 10k (MC) | 10k (MC) | 10k (MC) | 10k (MC) | 10k (MC) |
| Electron | 10k (Ckv.) | - | 10k (Ckv.) | 10k (Ckv.) | 10k (FD) | 10k (FD) | 10k (FD) | 10k (FD) | 10k (FD) |
| Pion | 10k (Ckv.) | 10k (Ckv.) | 10k (Ckv.) | 10k (Ckv. &FD) | 10k (Ckv. &FD) | 10k (Ckv. &FD) | 10k (Ckv. &FD) | 10k (Ckv. &FD) | 10k (Ckv. &FD) |

# Comparison with ANN

- The difference between cut-based PID and ANN
  - within $\pm 1\%$ at pion beam
  - within $\pm 5\%$ at electron beam

See Siyuan's Talk



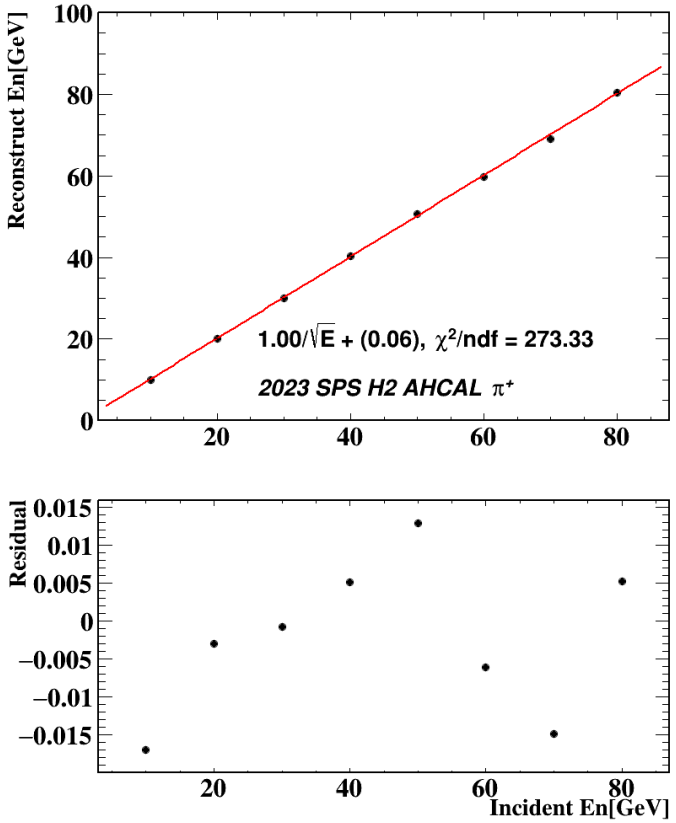



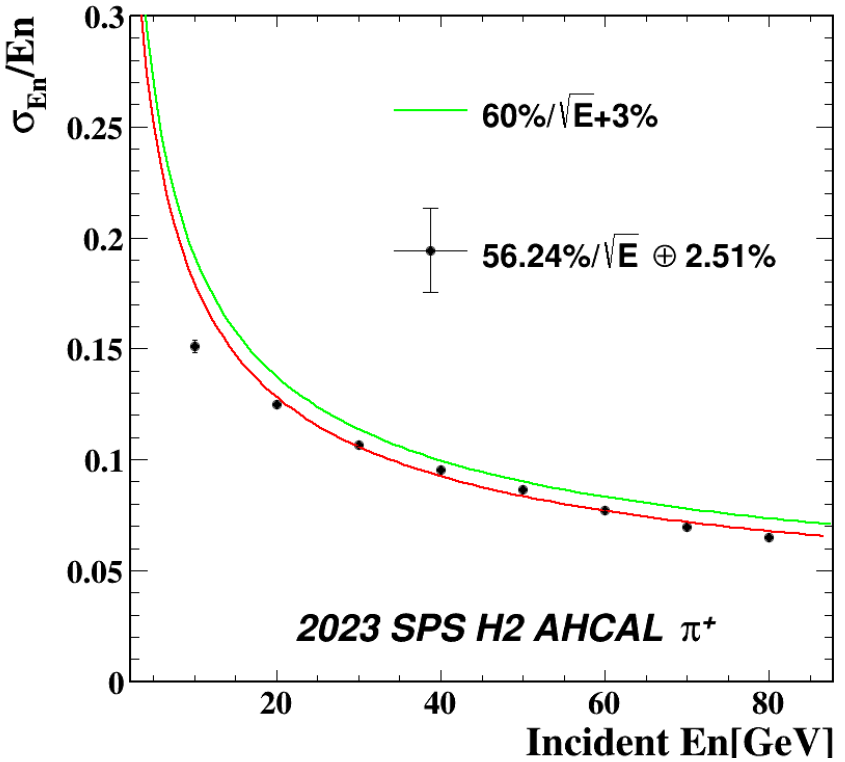
# Outline

- Introduction
- How to classify different data components?
- Data Component analysis
- Component Fraction and Comparison
- **Energy Performance**

# Energy performance

- Based on the pattern recognition algorithm
  - Quantify the energy performance of hadronic shower
    - Linearity : 1.00 * E + 0.006, within $\pm 1.5\%$
    - Resolution : $\frac{56.24\%}{\sqrt{E}} \oplus 2.51\%$



# Summary

- Different data components are observed using Fractal dimension and average hit energy.
- The corresponding component fractions are estimated by artificial selection.
  - Electron beam: purity > 80 %
  - Pion beam:
    - E > 30 GeV: purity > 80 %
    - E < 30 GeV: 15% < purity < 80%, electron events dominant
  - Consistent with ANN result.
- Based on referred pattern recognition:
  - Energy resolution of hadronic shower: $56.24\%/\sqrt{E} \oplus 2.51\%$

# Thank you !

Xin Xia

2023/9/6

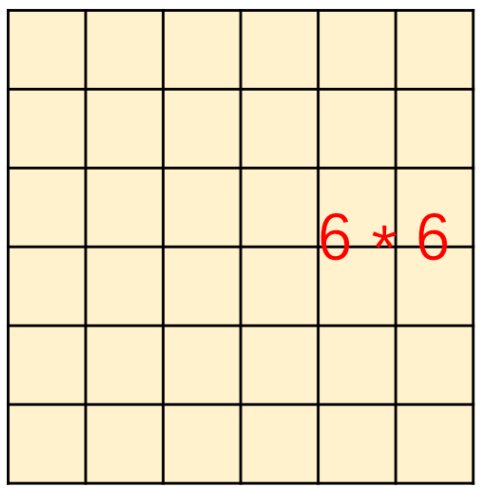
# Backup !

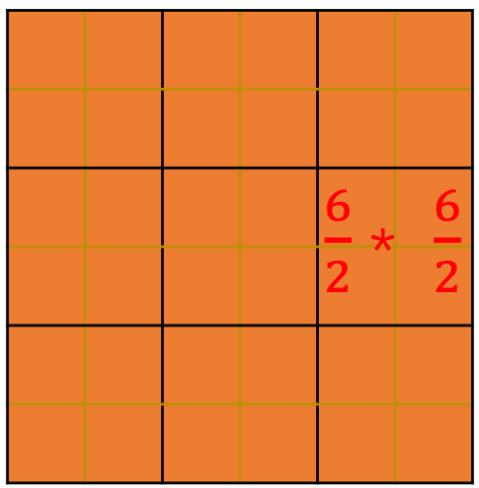
Xin Xia

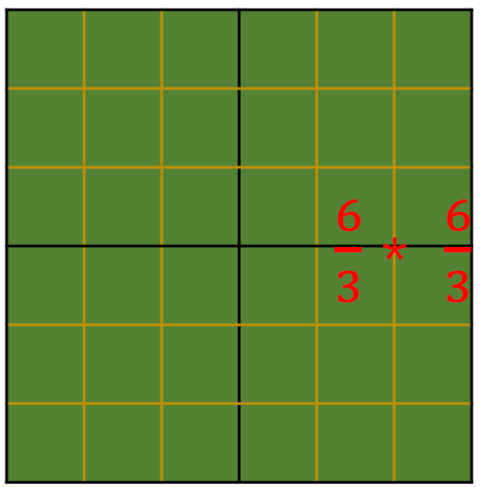
2023/9/6

For AHcal:

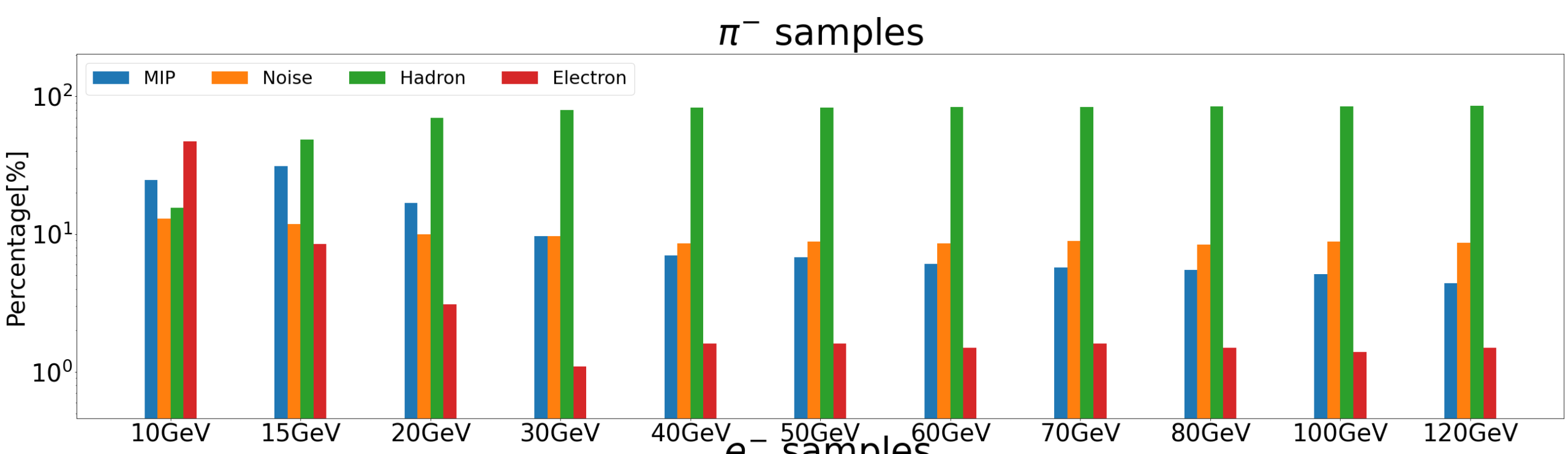
- Cell number: 18*18*40;
- Cell size: 40.2996mm*40.2996mm*26mm
- For each Hit, whose position is (Pos_x, Pos_y, Pos_z), define a new code($\frac{\text{Pos\_x}}{40.2996}$ , $\frac{\text{Pos\_y}}{40.2996}$ , $\frac{\text{Pos\_z}}{26}$ )= (x, y, z)
- We define Ratio_x[10] = Ratio_y [10] = (2,3,4,5,6,7,8,9,10,20); Ratio_z = 1;
- So new cell number for each layer is ($int(\frac{x}{Ratio\_x[i]}) * int(\frac{y}{Ratio\_y[i]})$)
- FD_V2 =+ $\frac{1}{10} * \frac{\log(\frac{Origin\ Hit\ Number}{New\ Hit\ Number})}{\log(Ratio[i])}$

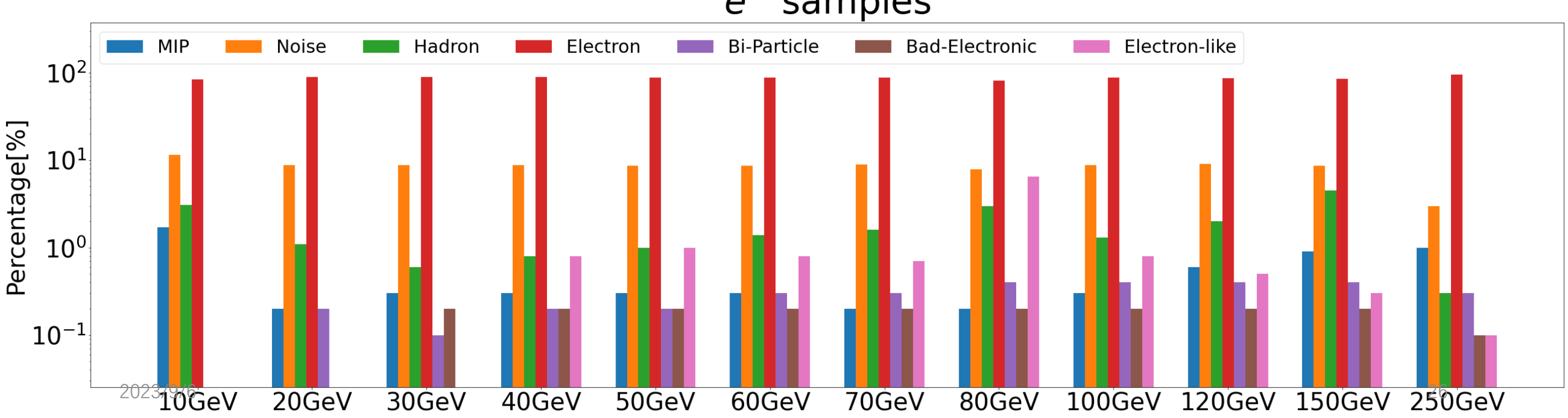


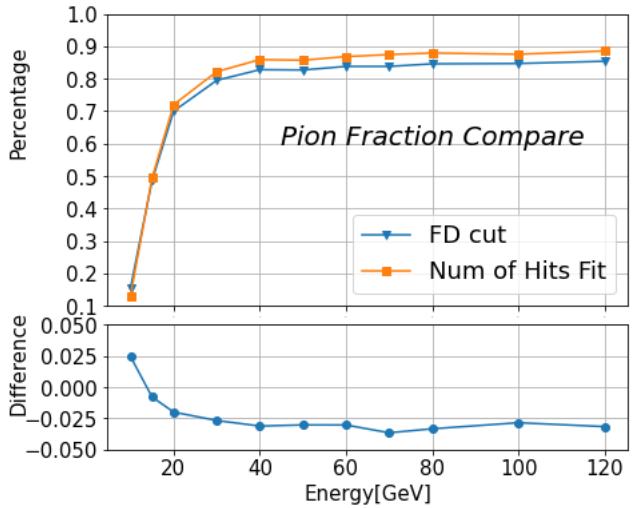



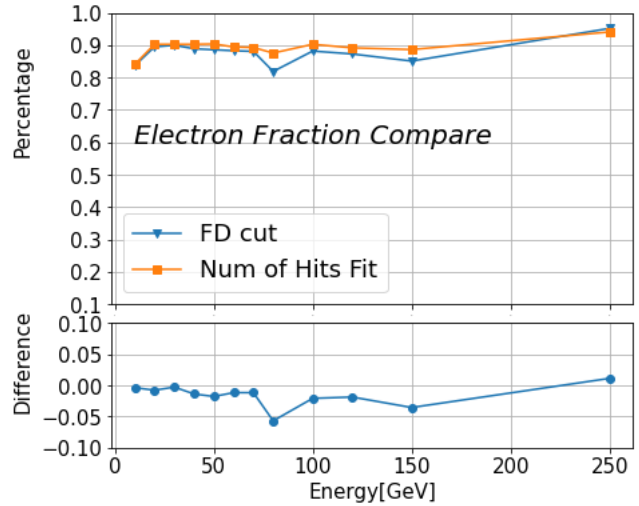



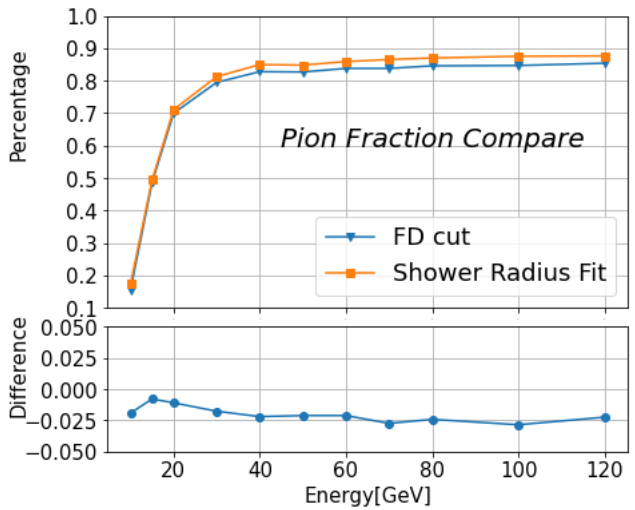

DOI: 10.1088/1742-6596/490/1/012227
DOI: 10.1103/PhysRevLett.112.012001
DOI: 10.1088/1742-6596/368/1/012038

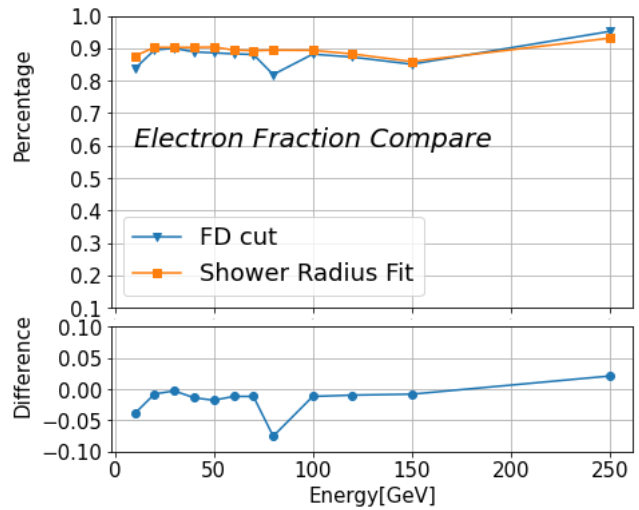

$\pi^-$ samples
MIP
Noise
Hadron
Electron
Percentage[%]
$10^2$
$10^1$
$10^0$
10GeV
15GeV
20GeV
30GeV
40GeV
50GeV
60GeV
70GeV
80GeV
100GeV
120GeV


$e^-$ samples
MIP
Noise
Hadron
Electron
Bi-Particle
Bad-Electronic
Electron-like
Percentage[%]
$10^2$
$10^1$
$10^0$
$10^{-1}$
10GeV
20GeV
30GeV
40GeV
50GeV
60GeV
70GeV
80GeV
100GeV
120GeV
150GeV
250GeV


Pion Fraction Compare
FD cut
Num of Hits Fit
Percentage
Difference
Energy[GeV]


Electron Fraction Compare
FD cut
Num of Hits Fit
Percentage
Difference
Energy[GeV]


Pion Fraction Compare
FD cut
Shower Radius Fit
Percentage
Difference
Energy[GeV]


Electron Fraction Compare
FD cut
Shower Radius Fit
Percentage
Difference
Energy[GeV]