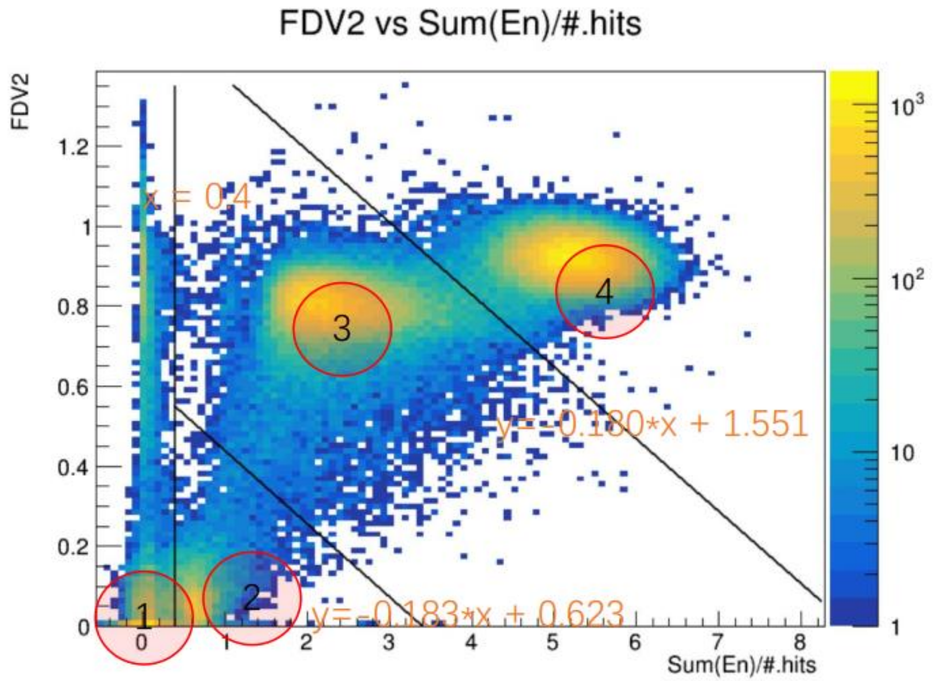
BDT-based PIDTool

--Zhen Wang, Jiyuan Chen

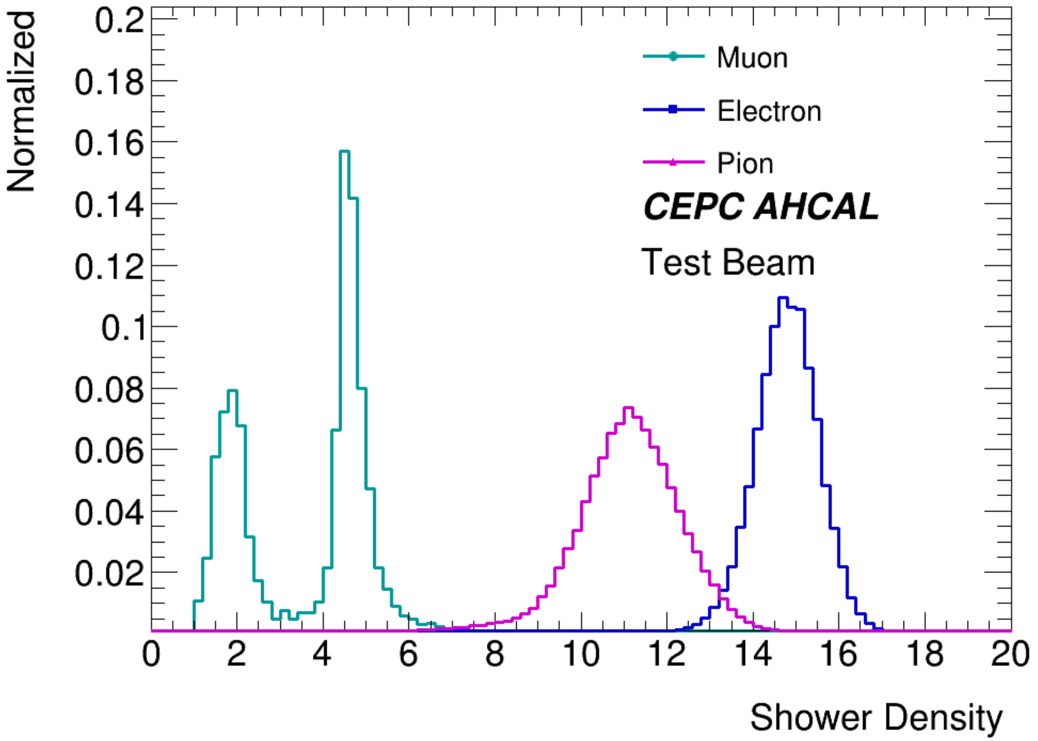
Pure samples

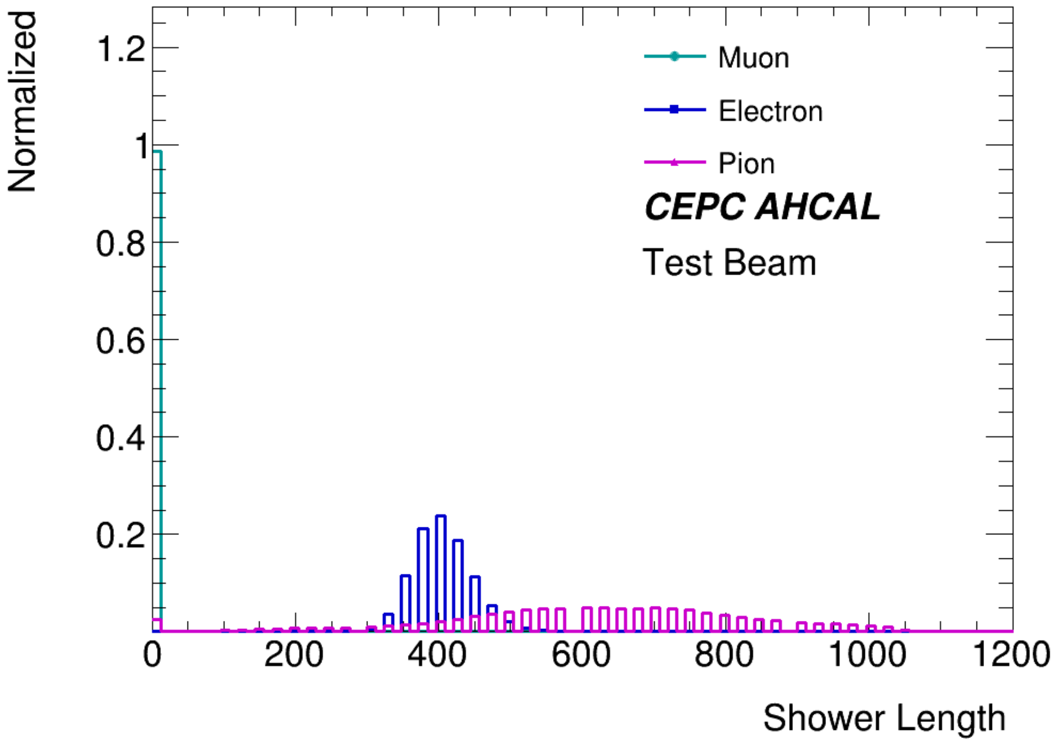
- Provided by Xin Xia.
- Clear signs of distinguishing power



- **if (y<0)**
  no hits in Event;
  ParticleID = 0;
- **if (y>0 && x<0.4)**
  noise
  ParticleID = 1
- **if (y>0 && x<0.4)**
  muon
  ParticleID = 2
- **if (y>0 && x>0.4 && y<-0.183*x + 0.623)**
  pi+
  ParticleID = 3
- **if (y>0 && x>0.4 && y>-0.183*x + 0.623) && y<-0.180*x + 1.551)**
  e+
  ParticleID = 4

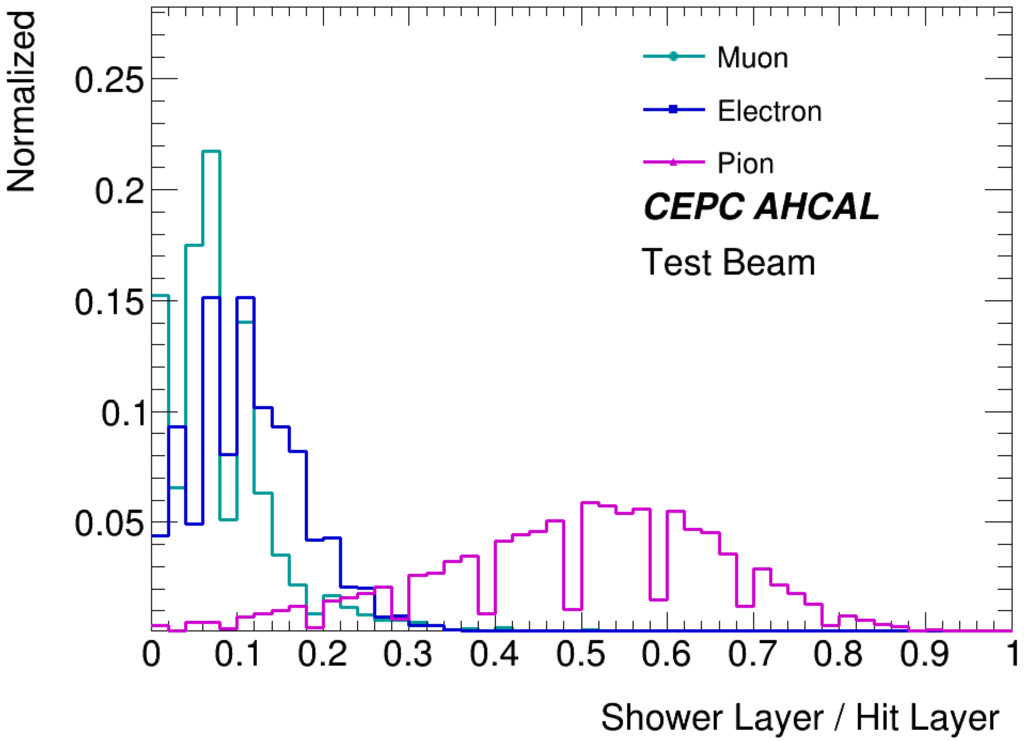
BDT Variables

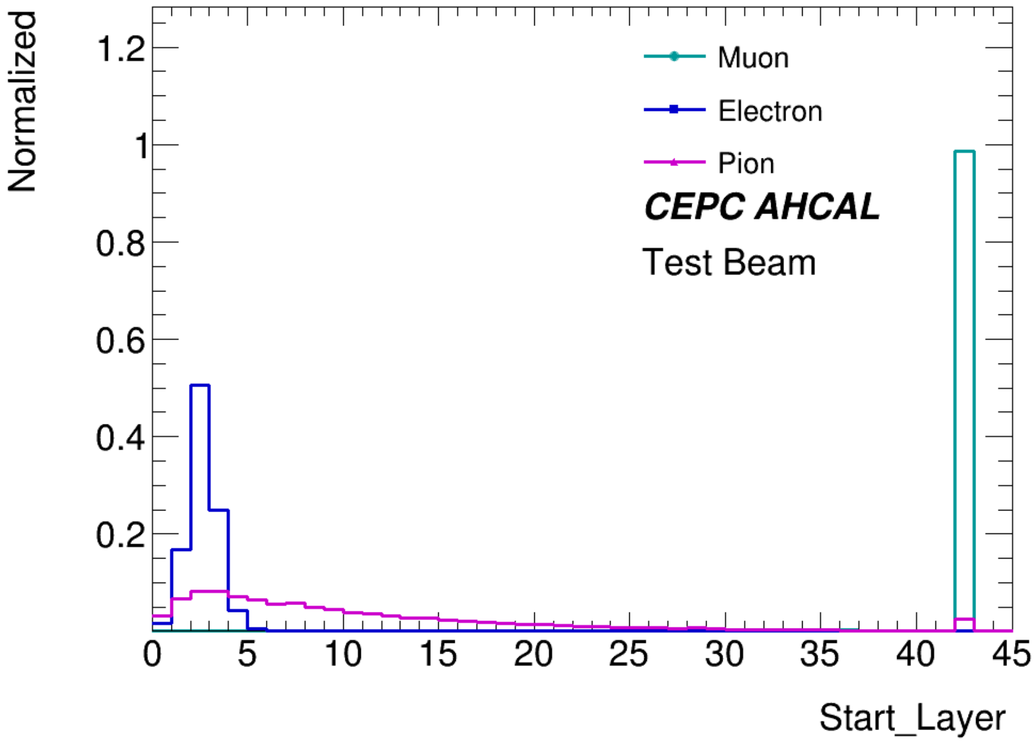




Pure test beam samples are used here.

BDT Variables

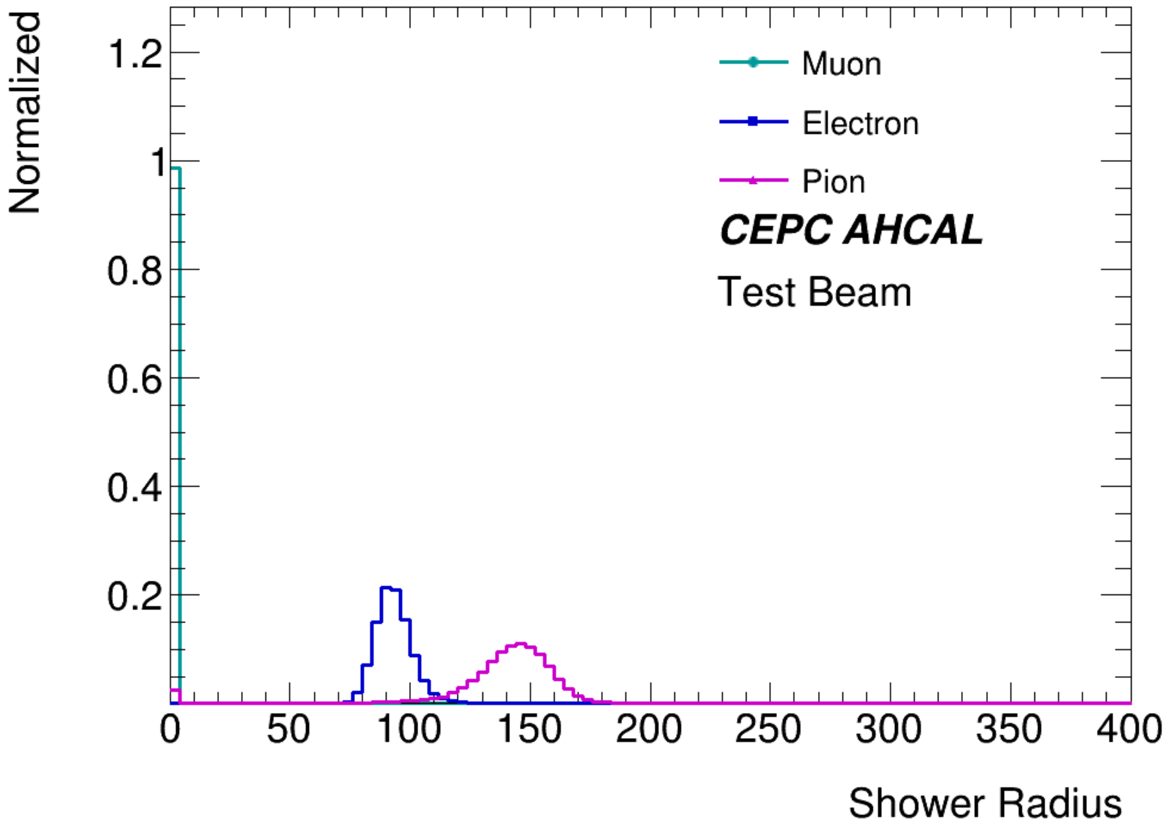




Pure test beam samples are used here.

BDT Variables

Added by Jiyuan Chen



Shower radius:

Between the shower start layer and end layers, the radius could be calculated as the RMS of $r = \sqrt\{x^2 + y^2\}$, x and y are position for each hit in the events.

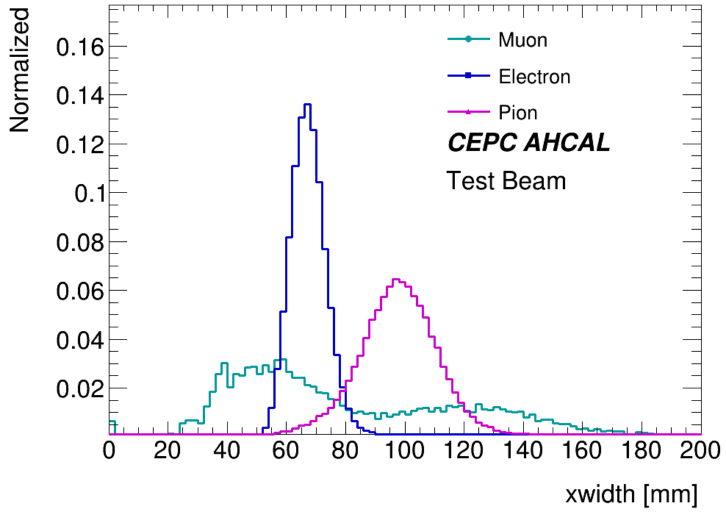
Pure test beam samples are used here.

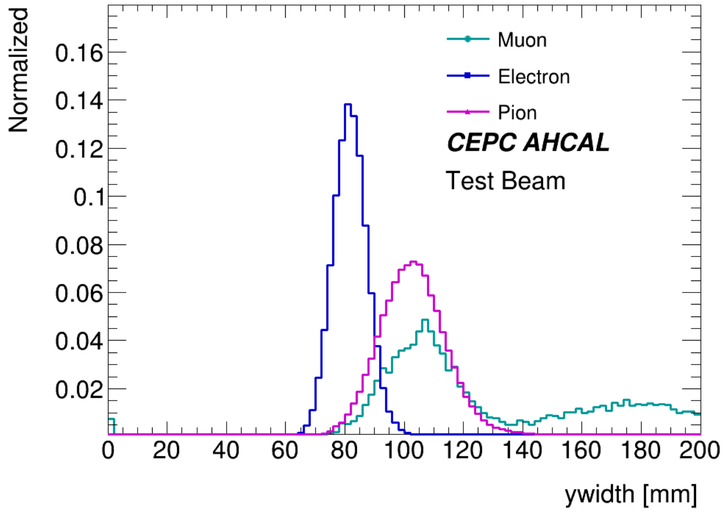
BDT Variables

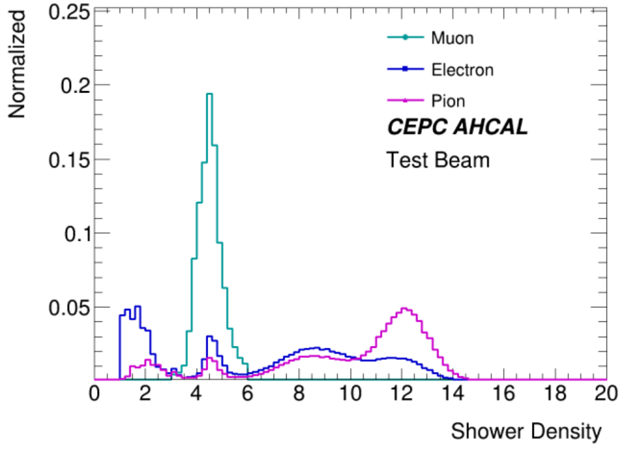
Conclusion:

Clear single-feature distribution

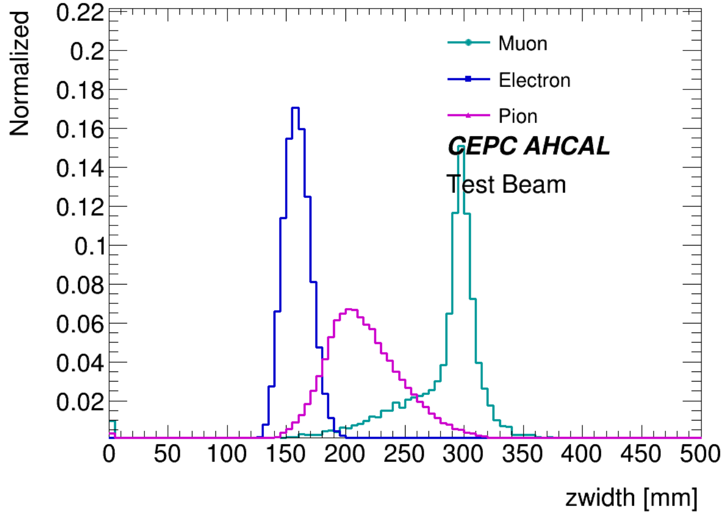
No more double peaks

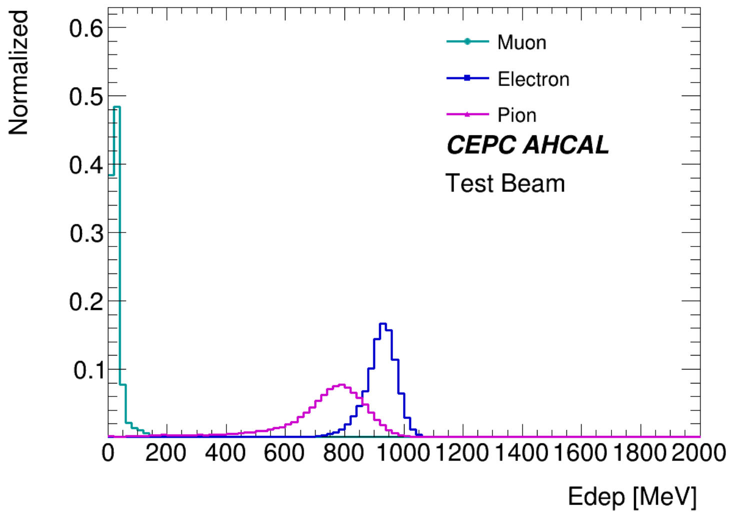




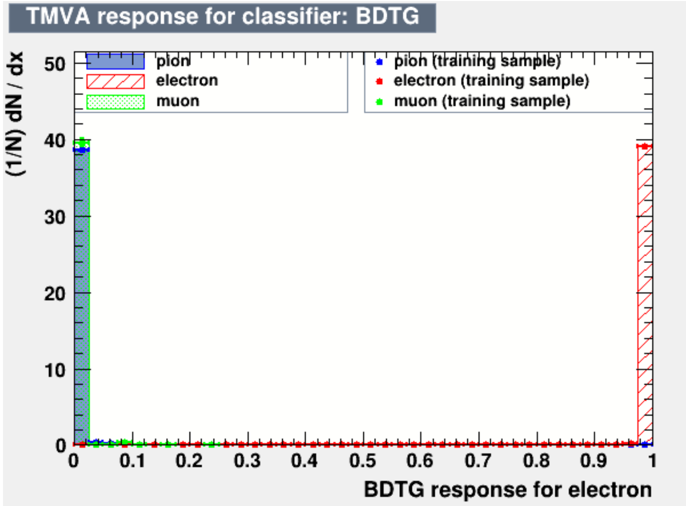


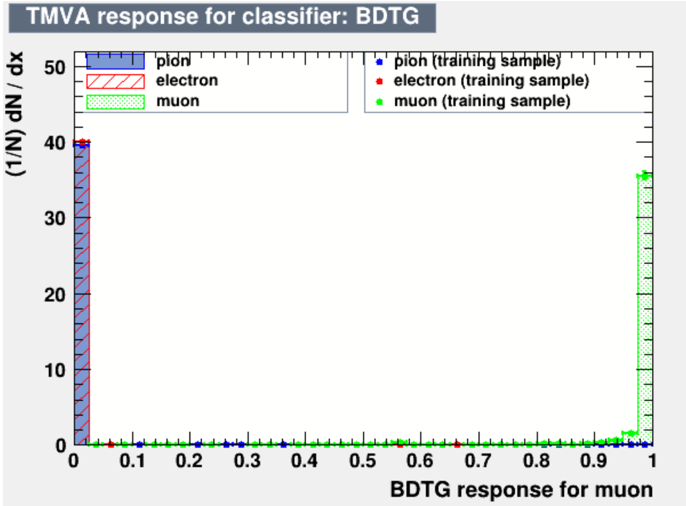
Original samples

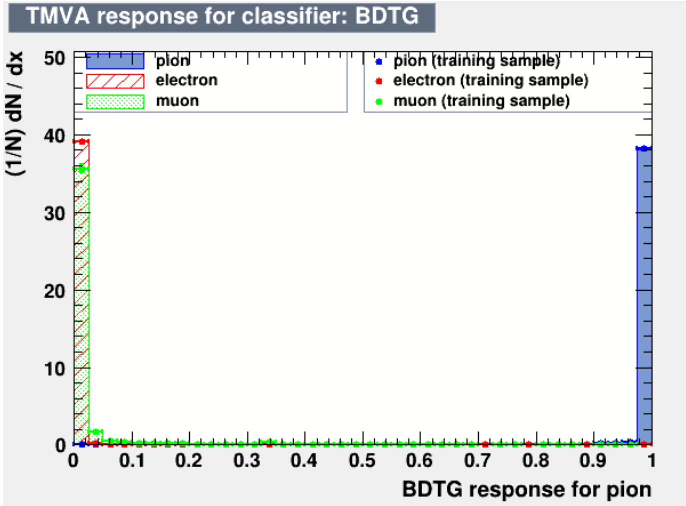


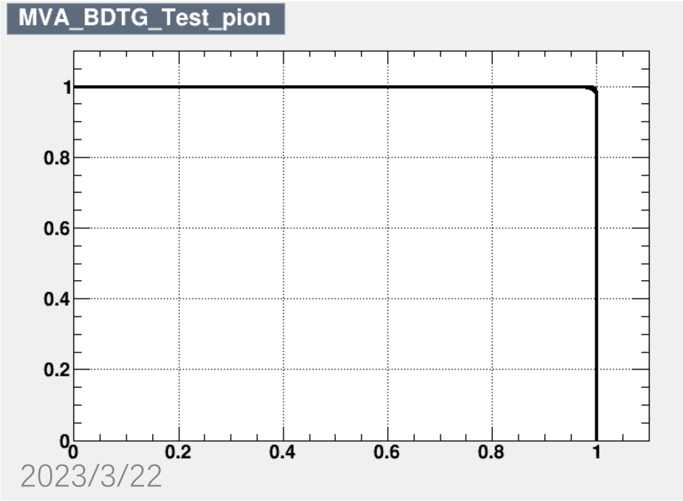


MultiClass BDTG



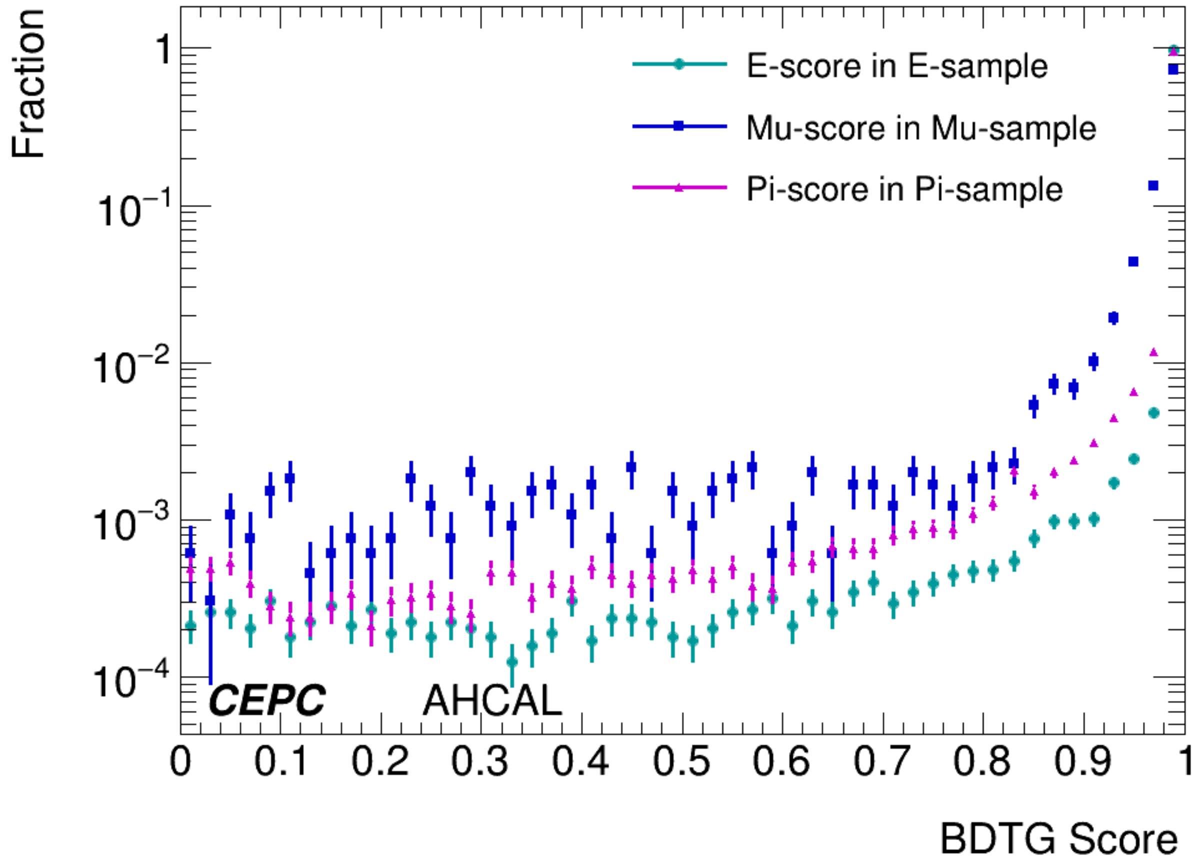






Conclusion:
Multiclass BDTG could be powerful in distinguishing different showers
Can be used to tag all of the samples

Application of Multiclass BDTG



BDTG Scores in corresponding samples

BDT_e_plus
BDT_mu_plus
BDT_pi_plus

```
root [2] Calib_Hit->Show(0)
======> EVENT:0
 BDT_pi_plus       = 7.25975e-05
 BDT_e_plus        = 0.999927
 BDT_mu_plus       = 1.16402e-16
```

After tagging, three extra branches are added in the tree demonstrating the possibility for the shower to be all three particles.

Different showers could be distinguished from the others

Plans

- Extra cuts could be applied on shower density to remove empty(noise) events
- Currently only mono-energy samples are obtained. Need samples for more energy points.
- Could be used to tag all of the test beam samples
- Extra information could be used for downstream analysis.

Backup

Shower Radius

- Since the directions of the incident particles are **perpendicular to the surface of HCAL**, we can use a very simple definition of “**shower radius**.”
- Definition: For each event, the layers where the shower **begins** and **ends** should be defined at first. Between them, we obtain the x and y values of the hits, and calculate $r = \sqrt{x^2 + y^2}$ for each of them; using all of the r values, r_{RMS} can be calculated.

Results from Simulation

- These three kinds of particles in our simulation samples can be very well distinguished.