Particle Identification using Boosted decision tree in semi-digital hadron calorimeter

Bing Liu[1,2], Haijun Yang[1], Imad Laktineh[2] ,Guillaume Garillot[2]

[1]SJTU,[2]IPNL

Oct. 15, 2018

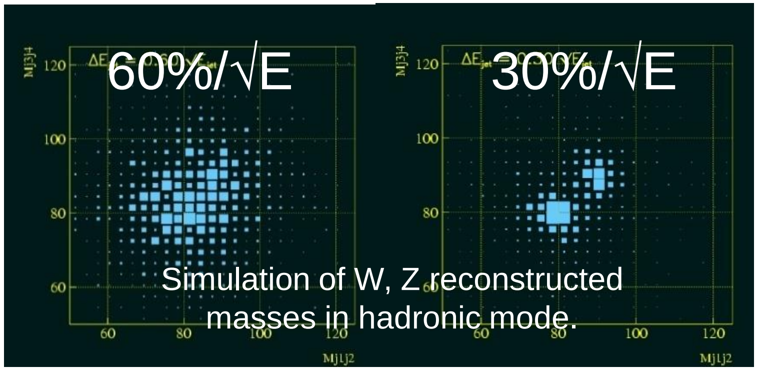
IPNL CALICE 上海交通大學 SHANGHAI JIAO TONG UNIVERSITY

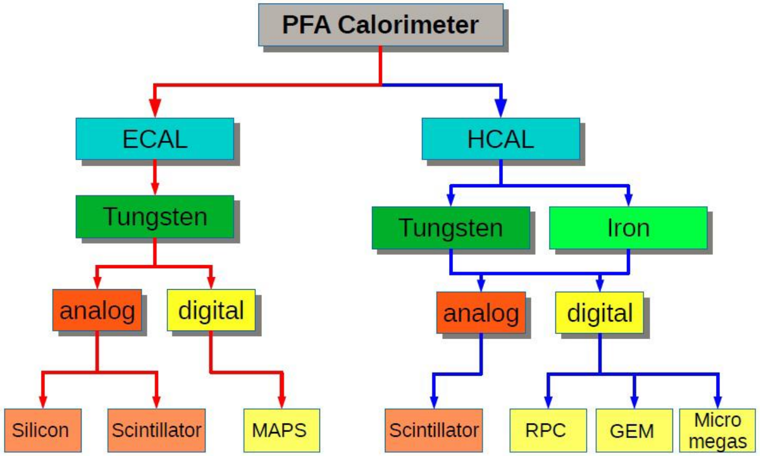
Outline

- The introduction of SDHCAL concept
  - PFA and Imaging Calorimeter
  - SDHCAL technological prototype
- Particle Identification using BDT
  - BDT using MC training
  - BDT using DATA training
- Conclusion

PFA and Imaging Calorimeter

For future colliders, jet energy resolution will be a determinant factor of understanding high energy physics.





$$\sigma_{jet}^2 = \sigma_{charged}^2 + \sigma_{EM}^2 + \sigma_{hadronic}^2 + \sigma_{confusion}^2$$

To improve on the jet energy resolution PFA is a promising solution to reduce the confusion term → high granularity Calorimeters

SDHCAL Concept

Ultra-granular HCAL can provide a powerful tool for the PFA leading to an excellent Jet energy resolution.

It is based on two points:

1- Gaseous Detector

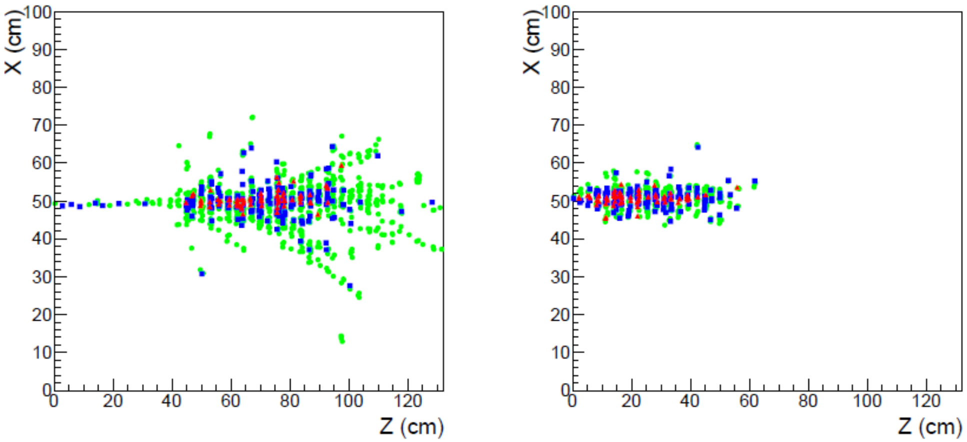
Gaseous detectors like GRPC are homogenous, cost-effective, and allow high longitudinal and transverse segmentation.

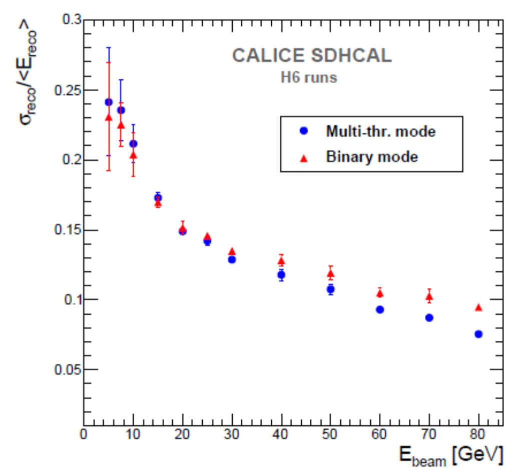
2- Embedded electronics Readout

A simple binary readout leads to a very good energy resolution

However, at high energy the shower core is very dense and saturation shows up

→ Multi-thresholds readout(Semi-digital) improves on energy resolution at energies> 30 GeV

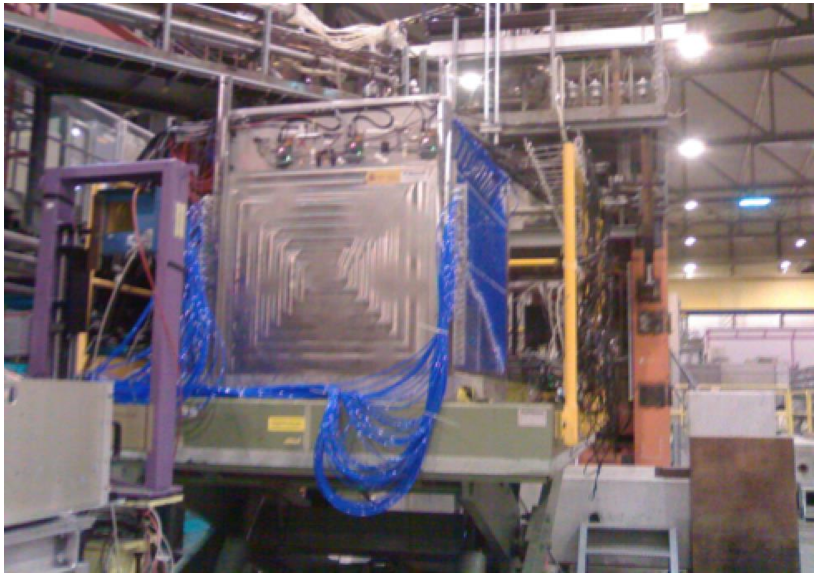




SDHCAL technological prototype

- Total Size:1.0x1.0x1.4m^3
- Total Layers: 48
- Total Channel(pads):440000
- Power consumption:$10\mu W/channel$

the first technological prototype among a family of prototypes of high-granularity calorimeters



developed by the CALICE collaboration

SDHCAL technological prototype

$(0.12\lambda_I, 1.14X_0)$

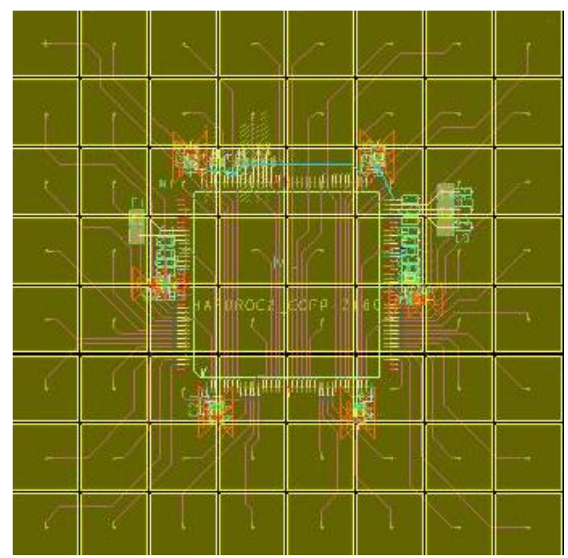
Stainless steel Absorber(15mm)

Stainless steel wall(2.5mm)

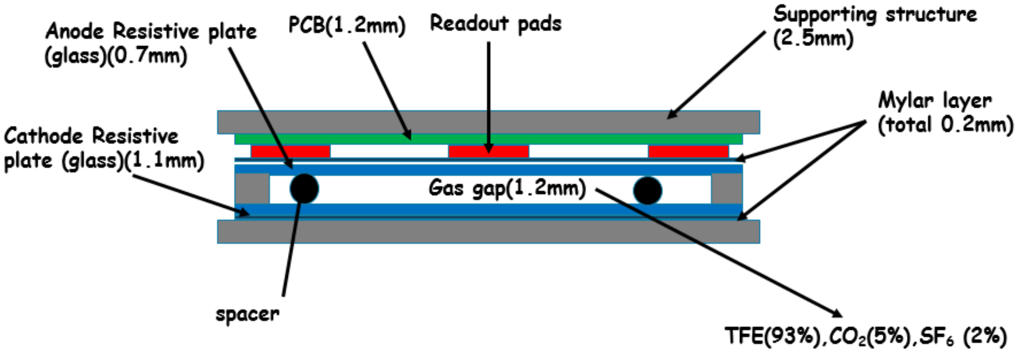
GRPC(6mm≈ $0\ \lambda_I, X_0$)

Stainless steel wall(2.5mm)

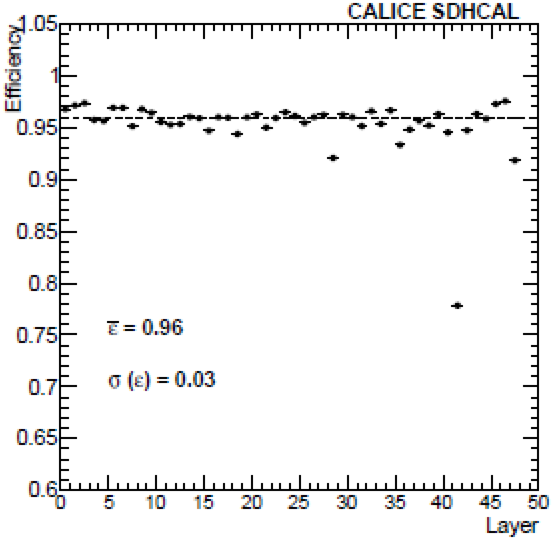
ASIC HARDROC(64 channel)
three-threshold (Semi-digital)
110fC,5pC,15pC

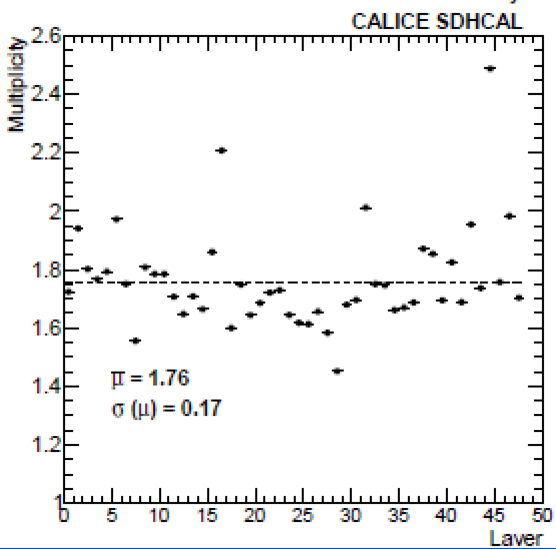


SDHCAL based on RPC



A schematic of GRPC (not to scale)





- **GRPC advantages:**
  - homogenous, cost-effective, negligible dead zone
  - allow high longitudinal and transverse segmentation

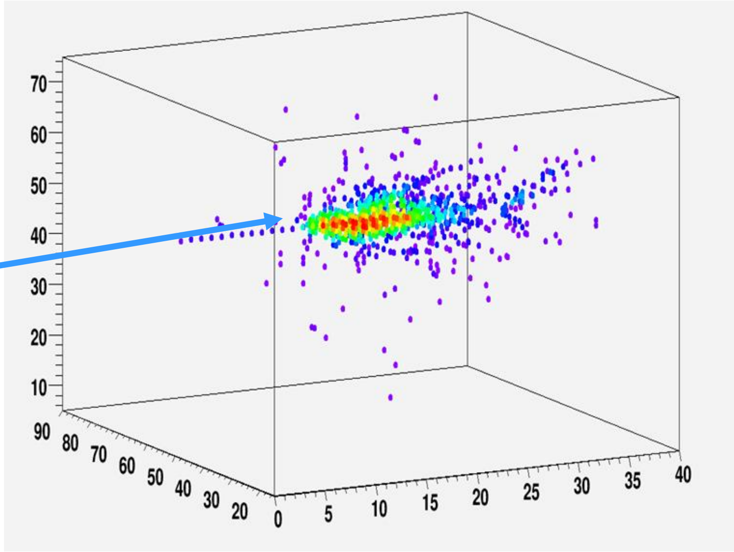
Particle identification based on BDT

Particle identification using BDT in SDHCAL

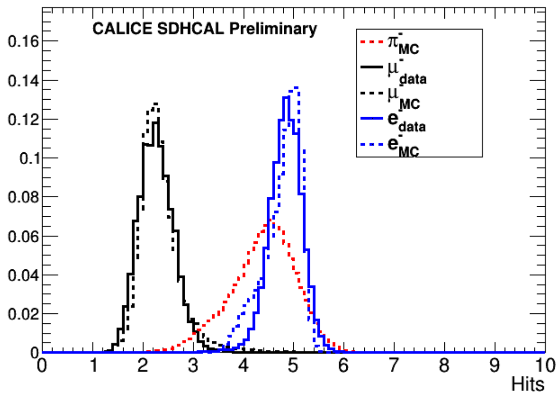
MC samples training

- PID
  - Application
    - Event selection
    - Better estimation in energy reconstruction
- Tool: Standard cuts, BDT
- TMVA of root, Scikit-learn
- BDT 6 var Input:

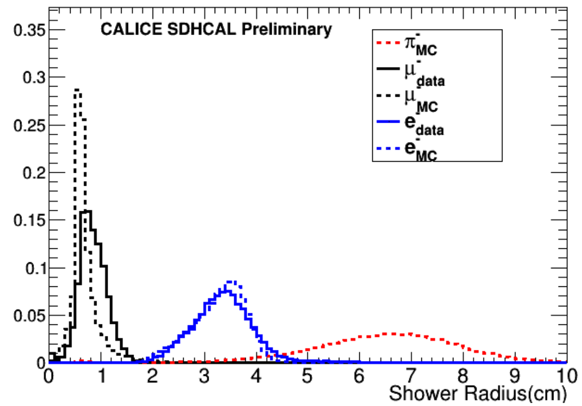
1. First layer of the shower(Begin)
2. Number of tracks in the shower (TrackMultiplicity)
3. Ratio of shower layers over total fired layers(NInteractinglayer/Nlayers)
4. Shower density(Density)
5. Shower radius(Radius)
6. Maximum shower position(Length)



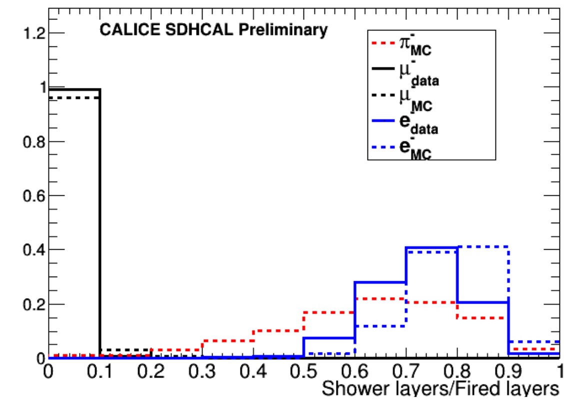
BDT Input variables



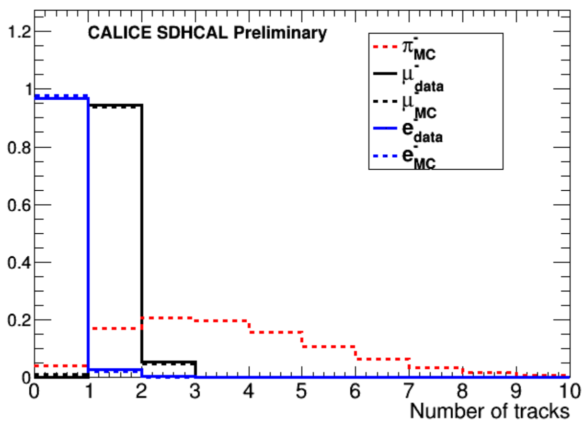
Density



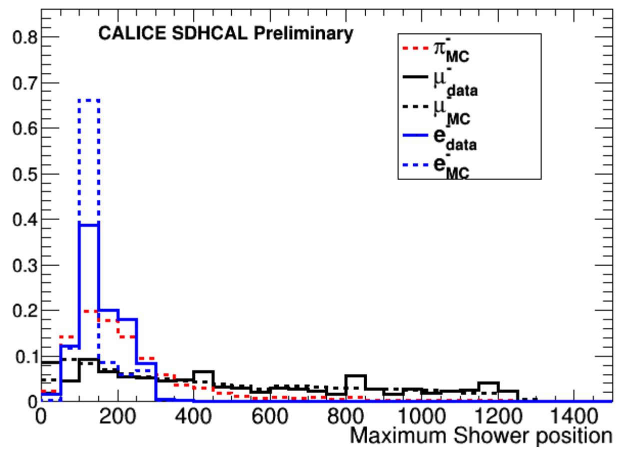
Radius



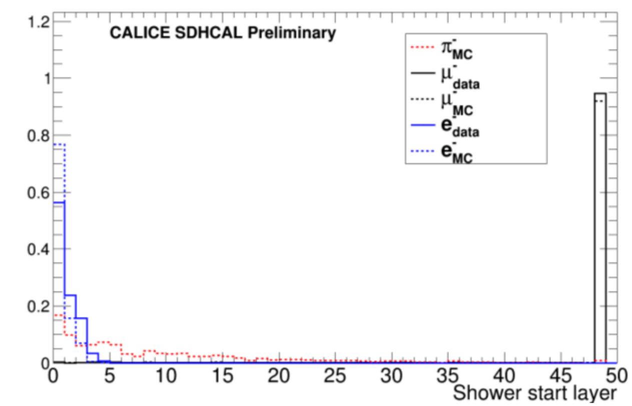
Shower layers / fired layers



Trackmultiplicty



Shower Length



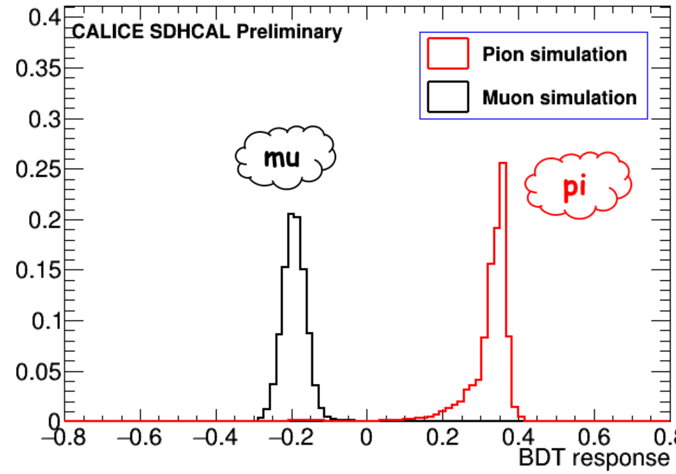
Begin

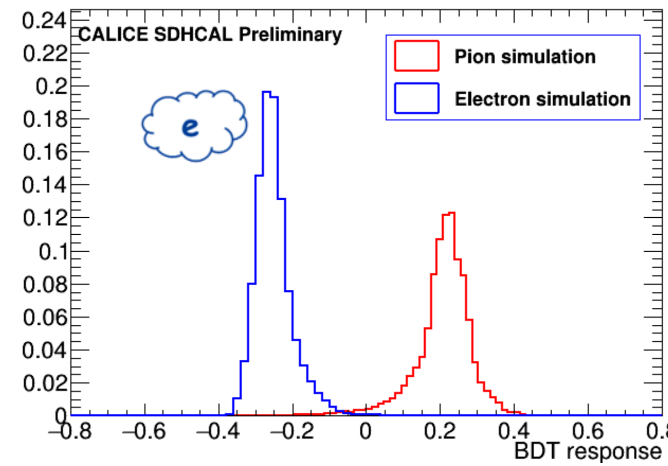
Training and Test

MC samples training

- TMVA of root, Methods: BDT 6var
- Training and Test
  - Signal: 160000 pion events with energy 10,20,30,40,50,60,70 and 80GeV
  - Background:160000 electron events with energy 10,20,30,40,50,60,70 and 80GeV
  - Background: ≈120000 muon events with energy 10,20,30,40,50,60,70 and 80GeV

    Mixed Background
  - Ntraining ：Ntest=1 ：1



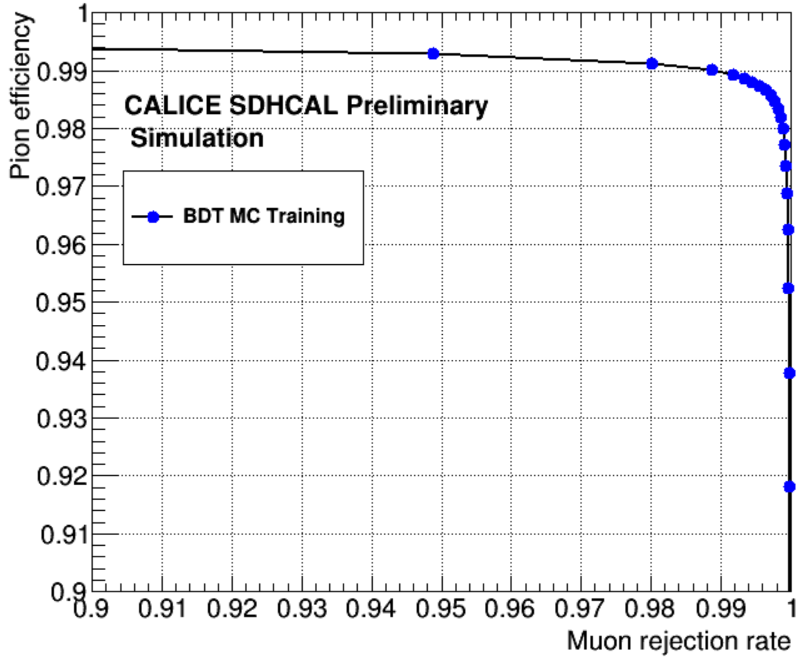


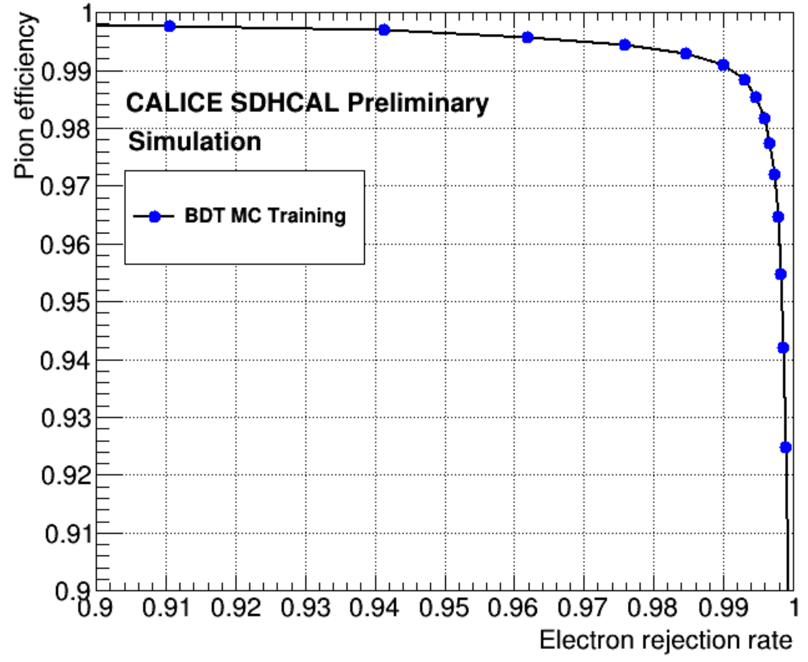
Strong separation power in pi/e and pi/muon

Pion eff vs Bkg rejection rate

MC samples training

- Good pi/e and pi/muon separation
- High pion efficiency exceeding 99% with electron and muon rejection of the same level (>99%)

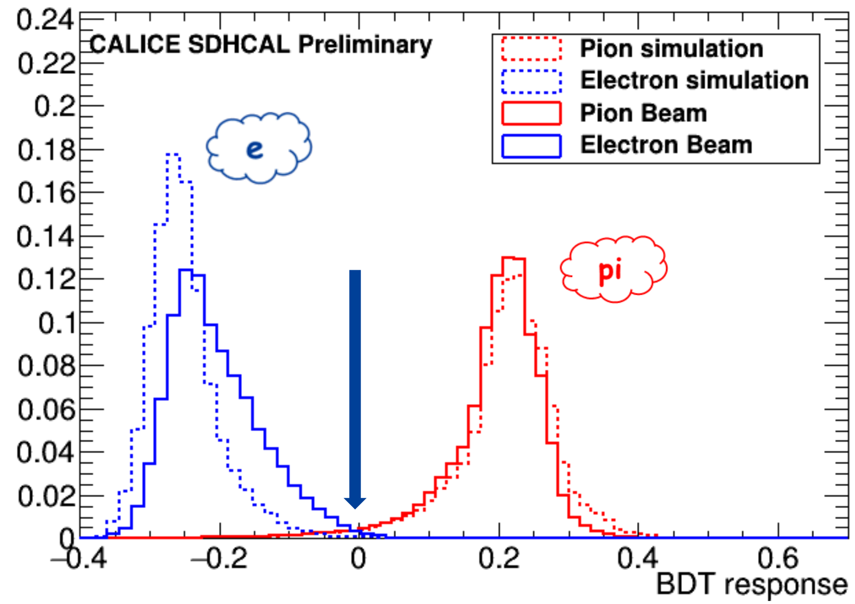


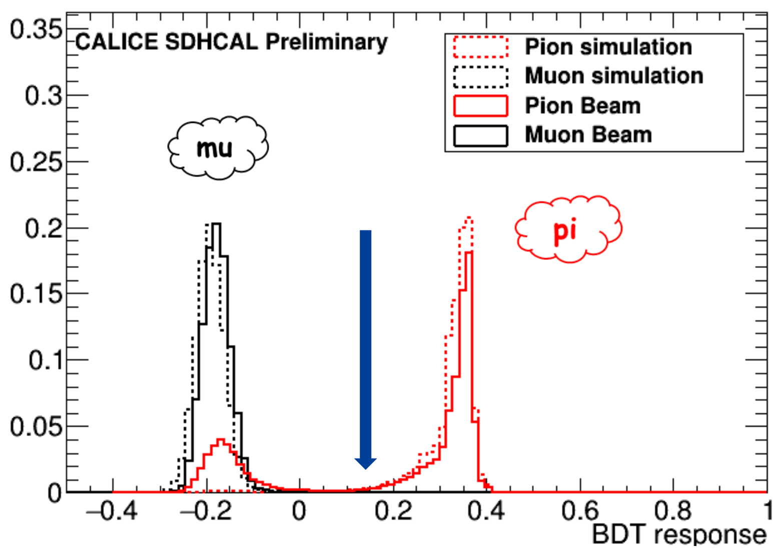


Particle identification using BDT

@Beam data validation

- **Beam data**
  SPS 2015
  - electron 10,20,30,40 and 50 GeV
  - Pion 10,20,30,40,50,60,70,80GeV
  - Muon 110 GeV



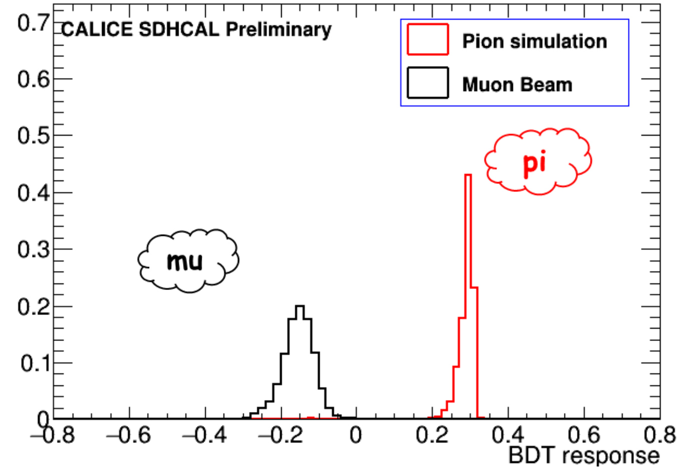


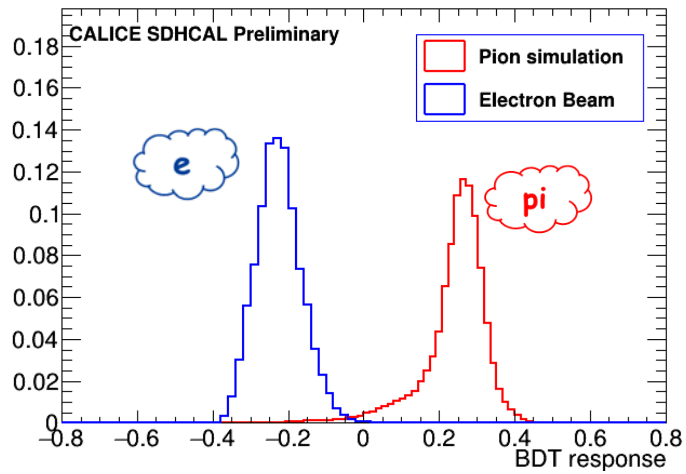
The beam data also show the performance of pi-e and pi-mu separation are good .

Particle identification using BDT in SDHCAL

DATA samples training

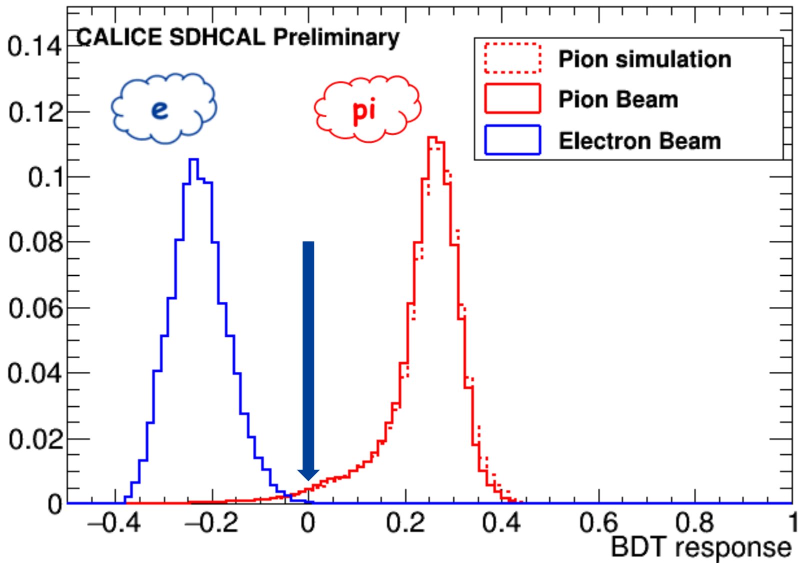
- **Training and Test**
  - Signal: 160000 pion events with energy 10,20,30,40,50,60,70 and 80GeV
  - Background:50000 beam electron events with energy 10,20,30,40,50GeV
  - Background: ≈50000 beam muon events with energy 110 GeV
  - Ntraining : Ntest=1 : 1

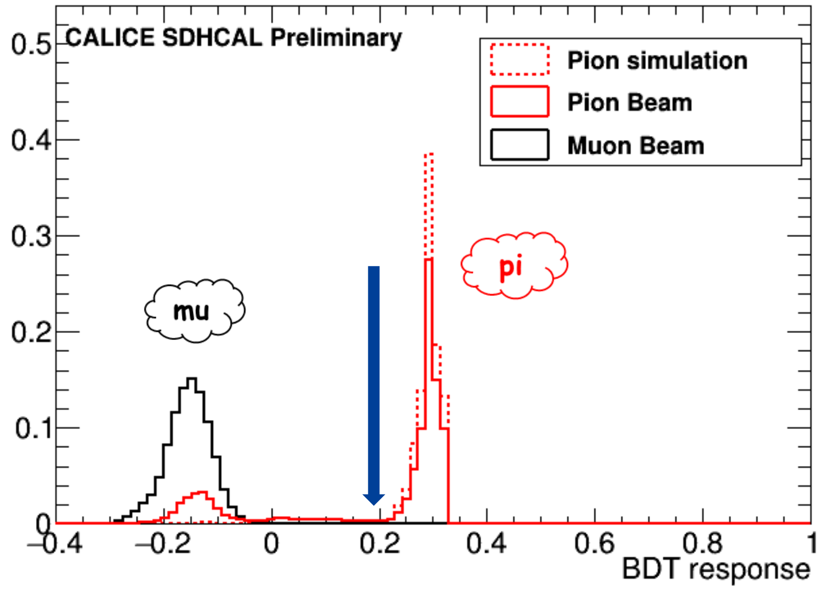




Particle identification using BDT

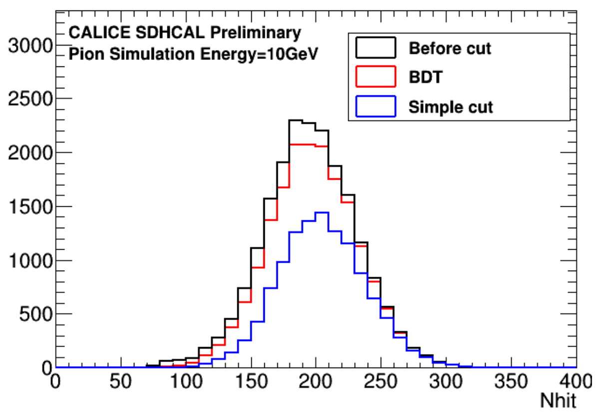
@Beam data validation

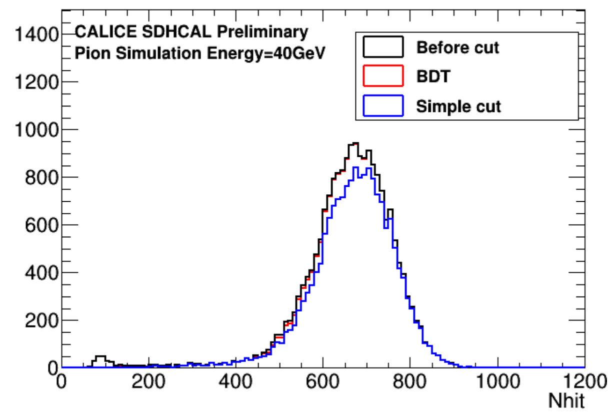


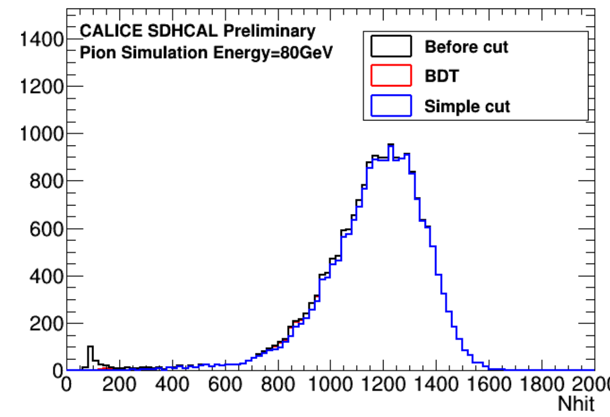


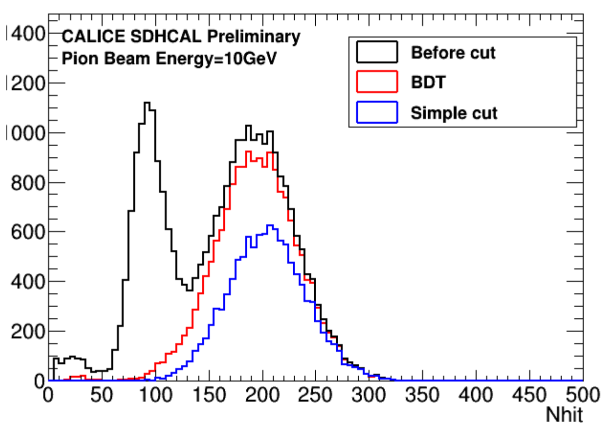
Good separation power and in agreement with previous MC training process

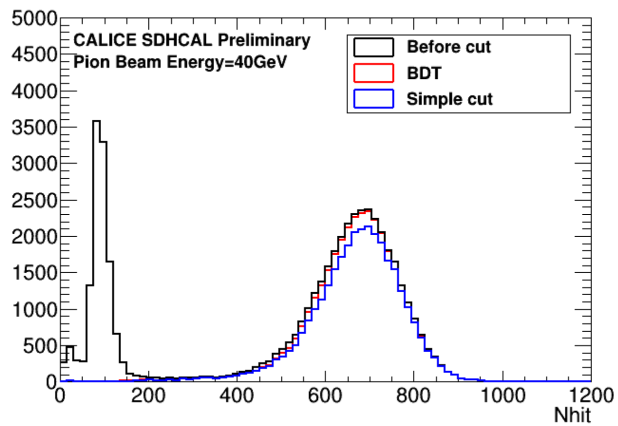
Comparison with standard selection

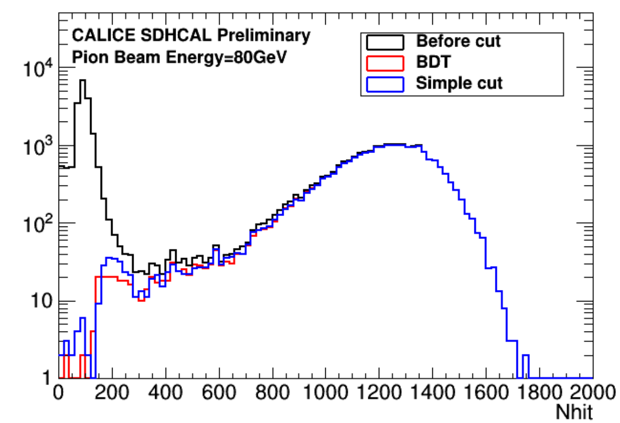












In the low energy, Using BDT save many events

Electron rejection	Shower start ≥ 5 or $N_{\mathrm{layer}} > 30$
Muon rejection	$\frac{N_{\mathrm{hit}}}{N_{\mathrm{layer}}} > 2.2$
Radiative muon rejection	$\frac{N_{\mathrm{layer}} \backslash \mathrm{RMS} > 5cm}{N_{\mathrm{layer}}} > 20\%$
Neutral rejection	$N_{\mathrm{hit} \in \mathrm{First\ 5\ layers}} \geq 4$

Energy reconstruction

- **Energy reconstruction formula:**

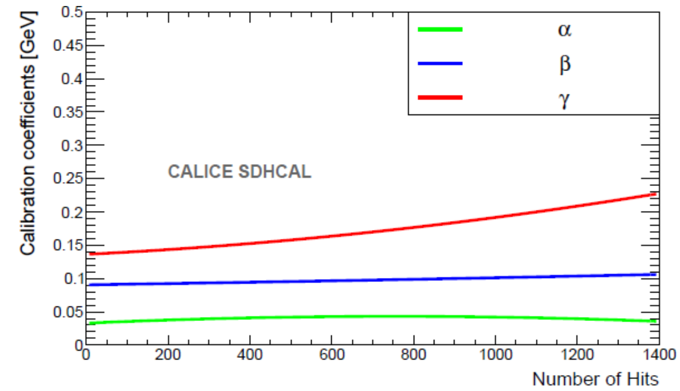
$$E_{reco} = \alpha N_1 + \beta N_2 + \gamma N_3$$

α, β, γ are parameterized as functions of total number of hits(N1+N2+N3)

$$\alpha = \alpha_1 + \alpha_2 N_{total} + \alpha_3 N_{total}^2$$

$$\beta = \beta_1 + \beta_2 N_{total} + \beta_3 N_{total}^2$$

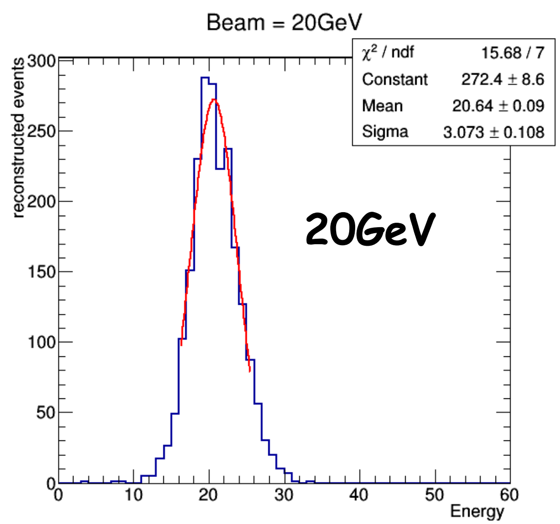
$$\gamma = \gamma_1 + \gamma_2 N_{total} + \gamma_3 N_{total}^2$$

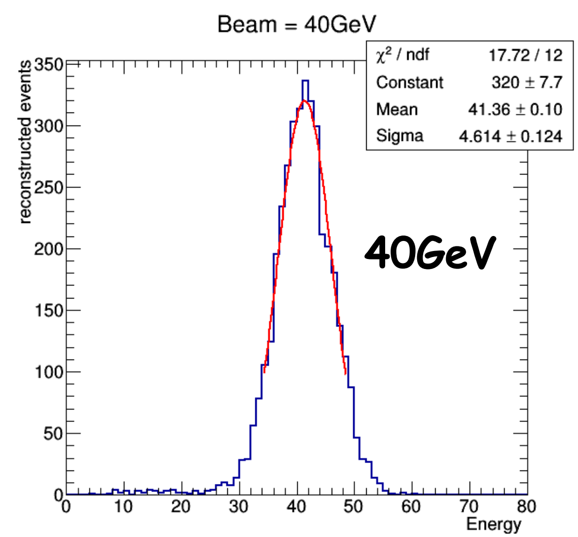


- **optimizer**

$$\chi^2 = \sum_{i=1}^{N} \frac{(E_{beam}^i - E_{reco}^i)^2}{\sigma_i^2}$$

$$\sigma_i = \sqrt{E_{beam}^i}.$$

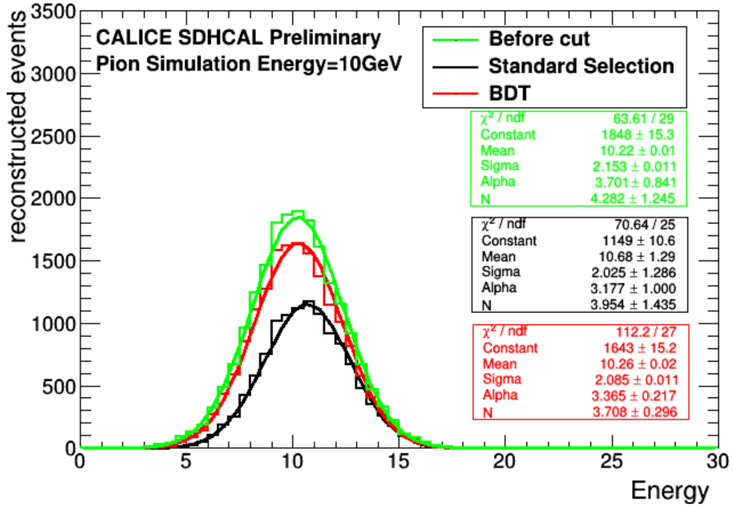


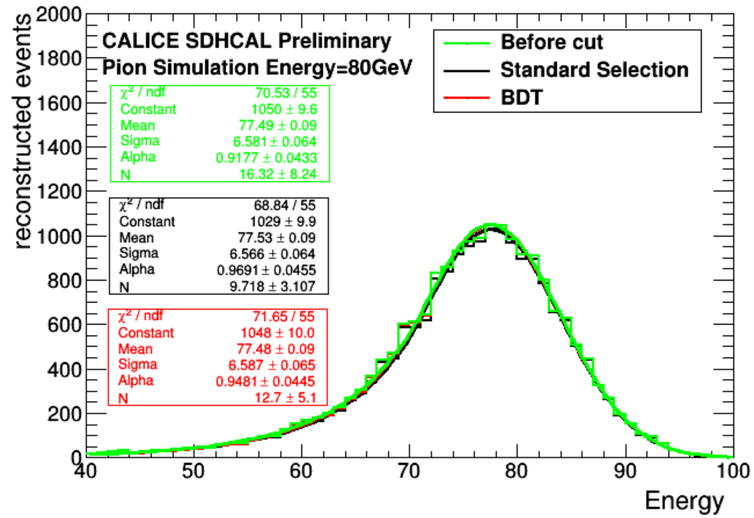


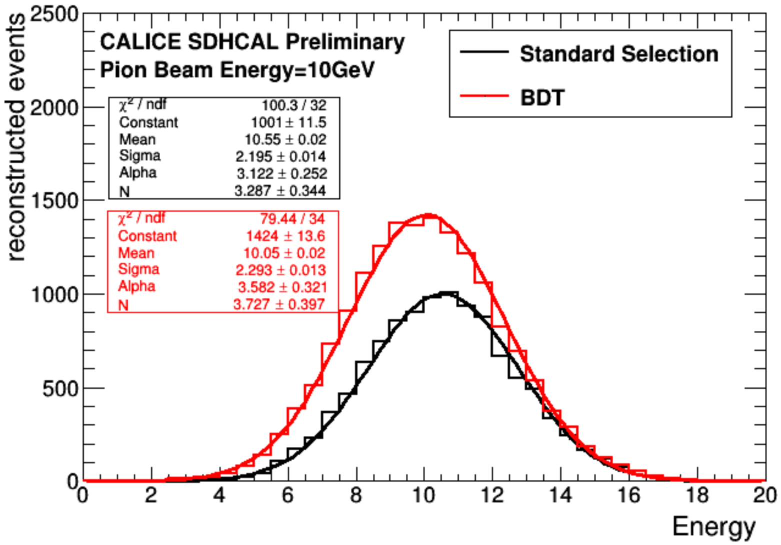
The distributions are fitted with a Gaussian Function in a 1.5σ range around the mean

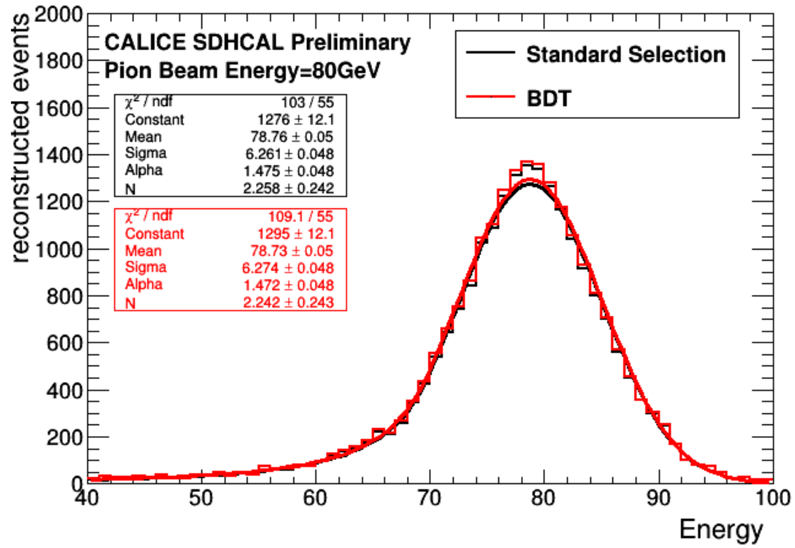
Comparison with standard selection

energy reconstruction

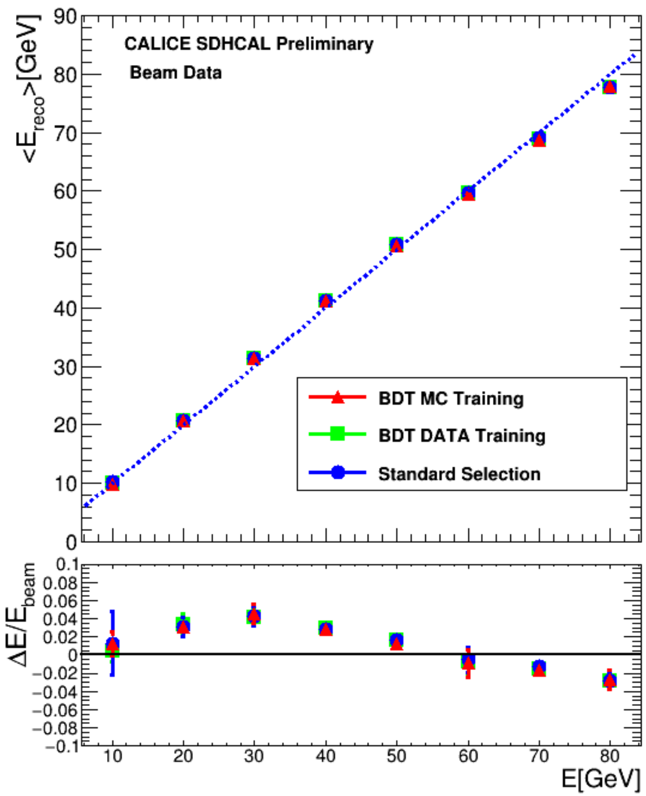


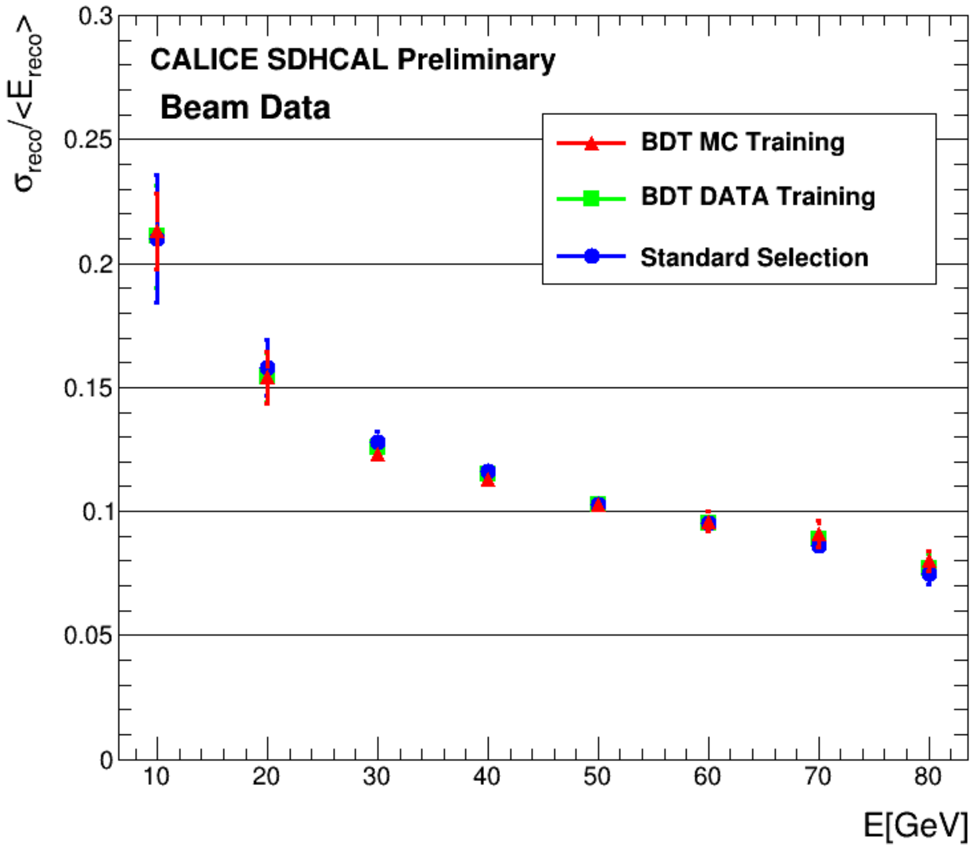






Particle identification using BDT in SDHCAL





Conclusion

◆SDHCAL

- SDHCAL provide a powerful high granularity tool for the PFA
- leading to an excellent energy resolution.

◆Particle identification using BDT in SDHCAL

- PID with BDT is reliable: Good pion efficiency with high electron and muon rejection rate
- Good resolution and linearity in agreement with standard method, and we got the improvement at 10GeV

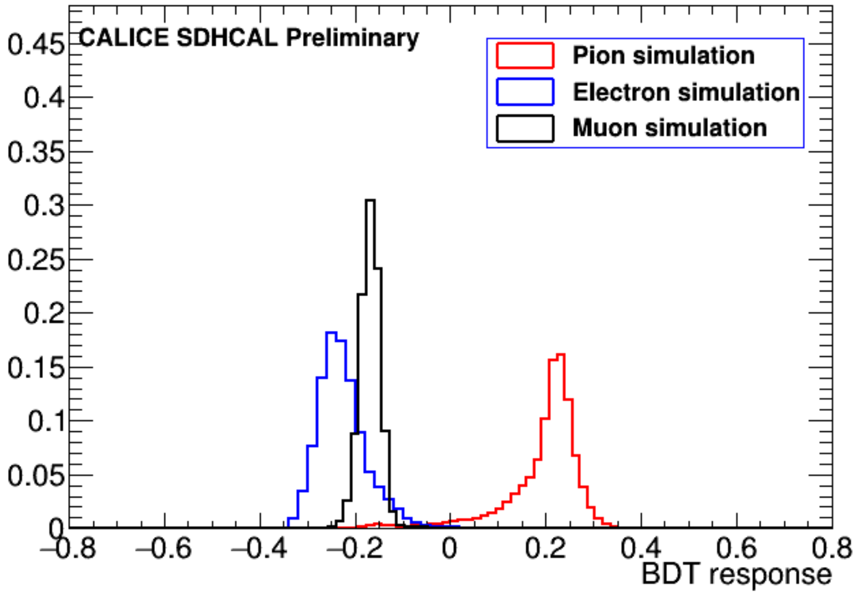
Back up

Training and Test

MC samples training @ Mixed

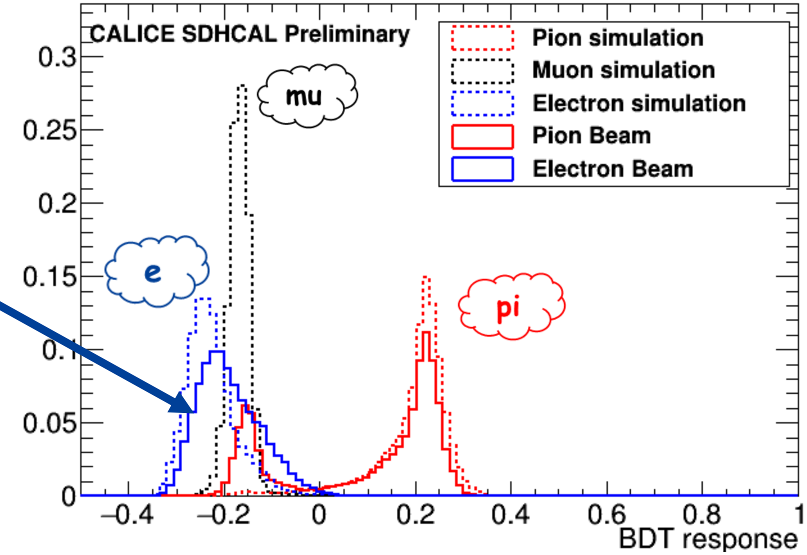
- Training and Test
  - Signal: 160000 pion events with energy 10,20,30,40,50,60,70 and 80GeV
  - Background:160000 electron events with energy 10,20,30,40,50,60,70 and 80GeV
  - Background: ≈120000 muon events with energy 10,20,30,40,50,60,70 and 80GeV

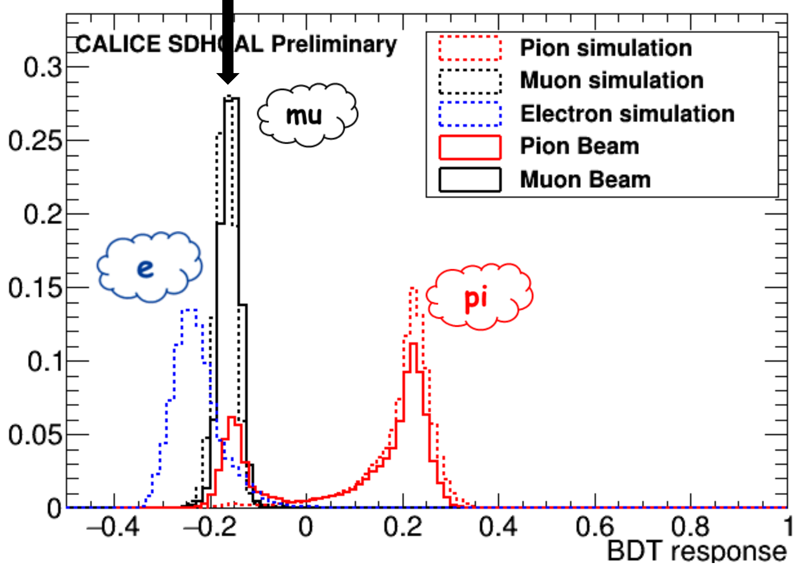
    Mixed Background
  - Ntraining : Ntest=1 : 1



Particle identification using BDT
@Beam data validation

- **Beam data**
  - electron 10,20,30,40 and 50 GeV
  - Pion 10,20,30,40,50,60,70,80GeV
  - Muon 110 GeV





the performance between
pi- and (e-+mu)separation are good .