The particle identification using BDT in the Semi-Digital Hadronic calorimeter

Speaker: Bing Liu,
On behalf of the CALICE SDHCAL group

CHEF2019, Fukuoka, Nov29, 2019
Outline

- Semi-Digital Hadronic calorimeter (SDHCAL)
- Particle identification using Boosted decision tree (BDT)
- Conclusions
**Particle Flow Algorithm** : an algorithm that tries to reconstruct individual final state particle from the information recorded in detector.

- It can lead to measure jet energy accurately, and consequently improve the energy reconstruction and event identification
- High granularity needed

\[ E_{JET} = E_{ECAL} + E_{HCAL} \]
\[ E_{JET} = E_{TRACK} + E_\gamma + E_n \]
Advantages of Resistive Plate Chambers:

- High efficiency
- Good Homogeneity
- Well contained avalanches
- Not expensive, robust, ...

SDHCAL prototype — Main features

Detector: **GRPC** (glass resistive plate chamber, 1 X 1 m²) operating in avalanche mode

1 x 1 cm² pads.

**Semi-Digital readout, 2bits — 3 thresholds**

→ Counts how many and which pads have a signal larger than one of three thresholds

Advantages of Resistive Plate Chambers:

- High efficiency
- Good Homogeneity
- Well contained avalanches
- Not expensive, robust, ...

Energy resolution
SDHCAL prototype — Main features

Embedded electronics:
- PCB separated from the GRPC by a mylar layer (50 μm)
  - **Bottom**: 1 × 1 cm² pads, one layer 9216 pads.
  - **Top**: 144 HARDROC2 (HAdronic Rpc Detector ReadOut Chip)

Absorber:
- Stainless steel (20mm)
- Surface planarity < 1mm
- Thickness tolerance 50 μm
10500 ASICs were tested and calibrated using a dedicated ASICs layout: 93%. SDHCAL prototype construction: 310 PCBs were produced, cabled, and tested. They were assembled by sets of six to make 1m² ASUs. 50 detectors were built and assembled with their electronics into cassettes. Self-supporting mechanical structure. Full assembly took place at CERN. 170 DIF, 20 DCC were built and tested. DAQ system using both USB and HTML protocol was developed and used. Since 2011, SDHCAL prototype Size: 1m x 1m x 1.4m. No. of layers: 48. Cell size: 1cm x 1cm. No. of channels: ≈440K. TB: 2012, 2015, 2016, 2017 and 2018 with ECAL.
Hadronic shower

Hits associated to:
1st threshold (110 fC)
2nd threshold (5 pC)
3rd threshold (15 pC)
$E_{reco} = \alpha N_1 + \beta N_2 + \gamma N_3$

$\rightarrow N_1 = \text{Nb. of hits with 1st threshold} < \text{signal} < \text{2nd threshold}$
$\rightarrow N_2 = \text{Nb. of hits with 2nd threshold} < \text{signal} < \text{3rd threshold}$
$\rightarrow N_3 = \text{Nb. of hits with} \text{signal} > \text{3rd threshold}$
$\rightarrow N_{\text{tot}} = N_1 + N_2 + N_3$

$\alpha, \beta, \gamma$ are parameterized as quadratic functions of $N_{\text{tot}}$
$\rightarrow$ They can be computed by minimizing a $\chi^2$

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

SDHCAL with 3-threshold results in better energy resolution than binary mode for $E_{beam} > 30$ GeV
◆ Tool: Standard cuts, BDT
Standard cut: shower begins after 5 layer

◆ TMVA package

◆ 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)
Strategy I: with MC samples training

Training phase

- **Signal**: 10000 simulated pion events for per energy point (10, 20, 30, 40, 50, 60, 70 and 80 GeV)
- **Background1**: 10000 simulated electron events for per energy point (10, 20, 30, 40, 50, 60, 70 and 80 GeV)
- **Background2**: ≈ 7500 simulated muon events for per energy point (10, 20, 30, 40, 50, 60, 70 and 80 GeV)

Test phase: the same as training phase but with independent samples
PID: Performance

◆ Good $\pi/e$ and $\pi/muon$ separation

◆ High pion efficiency exceeding 99% with electron and muon rejection rate of the same level (>99%)
**PID: beam data validation**

SPS 2015

- **electron** 10, 20, 30, 40 and 50 GeV
- **Pion** 10, 20, 30, 40, 50, 60, 70, 80 GeV
- **Muon** 110 GeV
**Strategy II: with data samples training**

**Training phase**

- **Signal:** 10000 simulated pion events for per energy point (10, 20, 30, 40, 50, 60, 70 and 80 GeV)

- **Background1:** 25000 beam electron events for energy 10, 20, 30, 40, 50 GeV

- **Background2:** ≈25000 beam muon events for energy 110 GeV

**Test phase:** the same as training phase but with independent samples
Good separation power and in agreement with strategy I
In the low energy, BDT saves many events
The improvements on statistics

BDT with MC Training

BDT with DATA Training
Summary

- Technological prototype running successfully since 2012
- Many results validate the potential of the concept
- PID with BDT is reliable: Good pion efficiency with high electron and muon rejection rate
- Good improvement on statistics comparing with standard method
Thanks for your attention !