Particle flow: a calorimeter reconstruction exercise with Monte Carlo and CALICE test beam events

Katja Krüger, Ambra Provenza, Felix Sefkow, Huong Lan Tran

Introduction

This note is a guide through an exercise with CALICE simulated and real data. Together with the short introductory lecture, it should convey an impression on the following topics:

- technologies for highly segmented imaging calorimeters;
- topology of electromagnetic and hadronic showers, fluctuations;
- particle flow algorithms at work;
- multi-particle event reconstruction.

The exercise uses graphical event displays and interactive reconstruction algorithms, it involves some simple statistical analysis, but does not require specific computing skills. The analysis can be done using spread sheets, or prepared ROOT tools which are straightforward to use.

A) Shower topologies and fluctuations

The goal of part A is to observe the energy dependence and fluctuations of electromagnetic and hadronic showers, to understand their differences and to compare the shower evolution in different media.

A1. Electromagnetic showers

Tools: CED event display [1], event-by-event longitudinal shower profile

Data: electron test beam events in the CALICE calorimeter at 10 GeV and 50 GeV beam energy

Task

- Scan about 15 events at both energies. How does the electromagnetic shower changes at higher energies?

Instructions

The software used by the CALICE collaboration is based on the ILC software [2] (Marlin, LCIO, LCCD, etc). To be able to use it, you have to set some environment variables. This can be easily done by issuing the following on the command line:

```
source /afs/desy.de/group/flc/pool/provenza/EDIT_new/FirstPart/init_calice.sh
```

Now you can start Marlin

```
cd /afs/desy.de/group/flc/pool/provenza/EDIT_new/FirstPart/mysteer/Steel
./start.sh electron 10 (electron 50)
```

An example of the event display for a shower produced by a 10 GeV positron in the CALICE HCAL and Tail Catcher is shown in Fig 1. Please note that at any point in time, you can enable the “help” for the event display by either right-clicking on the event display window, or by typing “h”. Be aware of the fact that the “layers” that appear in

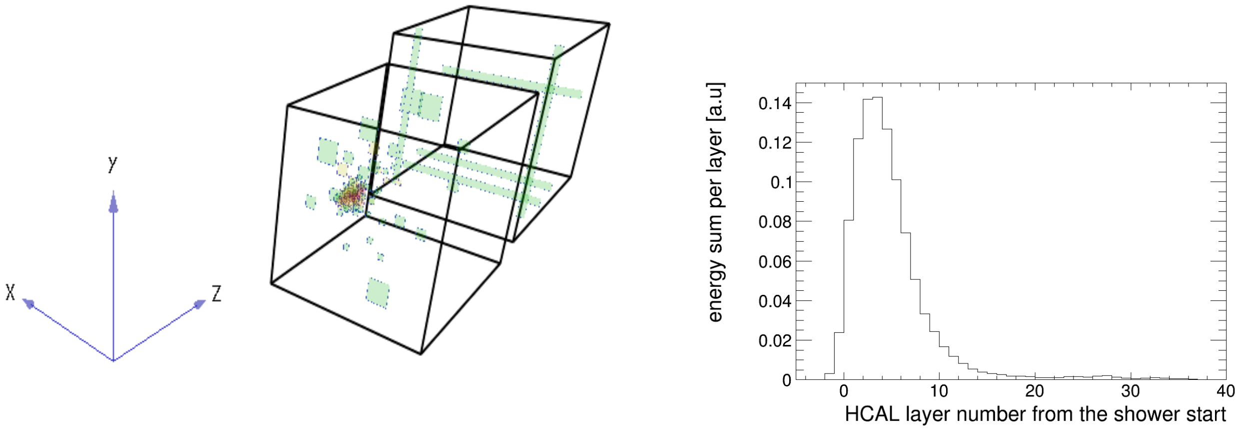

Figure 1: *a) An example of the event display for a shower produced by a 10 GeV positron in the CALICE HCAL and Tail Catcher; b) Longitudinal profile in the HCAL produced by 10 GeV positrons*

the help window do not refer to the HCAL physical layers, but to internal level associated to CED commands.

For a description of the color coding of the displayed HCAL cells, please see [3]. You can rotate and zoom into the detector using the mouse, and you can move it along the z-axis using the "up" and "down" keys.

Apart from the event display window, a canvas will appear, displaying the energy sum deposited in the HCAL versus the HCAL layer number, on an event by event basis. From this histogram you can read the number of the layer which contains the maximum energy deposition, and the corresponding energy sum.

In the event display, the layer where the shower start is shown in violet. You can disable this option by typing "@".

To proceed to the next event, just press ENTER. At the end don't forget to close the CED window.

The output root file can be found here:

```
cd /afs/desy.de/group/flc/pool/provenza/EDIT_new/FirstPart/Analysis/electron
root -l ShowerInfo_ePlus_10GeV_350118_fewEvents.root
```

You can find the root file for around 60000 events (90000 for electron 50 GeV) here:

```
/afs/desy.de/group/flc/pool/provenza/EDIT_new/FirstPart/Analysis/electron/ShowerInfo_ePlus_10GeV_350
```

This contains information about the layer where the shower start, the energy with the maximum energy, the value of the energy in this layer, the energy sum per layer, the total energy sum per events.

Look at the distribution typing, for example

```
root -l ShowerInfo_ePlus_10GeV_350118.root
tree → Draw(“startLayer″)
```

Make a comparison between these variables for electron 10 GeV and electron 50 GeV, look at the correlation. To do that you can run the script

```
cd /afs/desy.de/group/flc/pool/provenza/EDIT_new/FirstPart/Analysis/macros
root -l comparison_electron.C
```

Here a selection on the shower start is done: only events for which the shower start in layer between 1 and 35 are selected. This is to don't take into account events for which the layer where the shower starts is not properly determined.

A) 2. Electromagnetic shower and Hadronic shower

Compare the 10 GeV electromagnetic showers with 10 GeV hadron showers.

To run the event display for the hadronic shower the same software in part A1.1 is used.

Tools: as above in part A.1

Data: 2 data sets (electron 10 GeV and pion 10 GeV)

Task: Scan about 15 events. What are the main differences between the two different types of showers?

Like before, to run the event display you can do:

```
cd /afs/desy.de/group/flc/pool/provenza/EDIT_new/FirstPart/mysteer/Steel
./start.sh pion 10
```

The output root file can be found here:

```
cd /afs/desy.de/group/flc/pool/provenza/EDIT_new/FirstPart/Analysis/pion/Steel
```

```
root -l ShowerInfo_pion_10GeV_330332_fewEvents.root
```

the output root file for the all events can be found in the same directory:

ShowerInfo_pion_10GeV_330332.root

Compare the info stored in the root file for electron 10 GeV and pion 10 GeV using root.

Than you can run the script

```
cd /afs/desy.de/group/flc/pool/provenza/EDIT_new/FirstPart/Analysis/macros
root -l comparison_electronVSpion.C
```

Also here a selection on the layer where the shower starts is done.

A) 3. Hadronic showers

Tools: as above in part A.1

Data: 3 data sets (20 and 80 GeV pions in Fe-HCAL and 20 GeV pions in W-HCAL

Task:

1. Scan about 50 events.

2. Analysis Look at the distributions of the different quantities stored in the root file, compare the 20 GeV pion in Steel with the 80 GeV pion in Steel and that compare the 20 GeV pion in Steel with the 20 GeV pion in W.

- measure the hadronic interaction length for iron and tungsten
- test its energy dependence

[optional]

- determine mean and r.m.s. of the energy sums

[optional]

- determine mean and width of the shower max, measured from the shower start

[optional]

Instructions

For the analysis, the same software is used as for the electromagnetic showers (see instructions in A.1.).

The necessary steering files can be found here:

```
cd /afs/desy.de/group/flc/pool/provenza/EDIT_new/FirstPart/mysteer/Steel/
```

To run it just do

```
./start.sh particle energy
```

(Similar path and command to use the Tungsten HCAL data)

This will create a root file with all the quantities needed. The output file can be found here:

```
cd /afs/desy.de/group/flc/pool/provenza/EDIT_new/FirstPart/Analysis/pion/Steel
ShowerInfo_pion_20GeV_330771_fewEvents.root
```

(the analogue file for Tungsten data is in the directory ../Tungsten/) To obtain what is required (and for a larger number of events) uses the two root files

ShowerInfo_pion_20GeV_330771.root (ShowerInfo_pion_330962_80GeV_moreEvents.root)

and run the macro

```
cd /afs/desy.de/group/flc/pool/provenza/EDIT_new/FirstPart/Analysis/macros
root -l comparison_pion.C
```

For an example of the event display, please see Figure 2.

For measuring the hadronic interaction length for iron and tungsten, you need to know how to calculate the nuclear interaction length for compound materials. This is done similar to the radiation length case.

$$\frac{1}{X_0} = \sum_j \frac{w_j}{X_j}$$

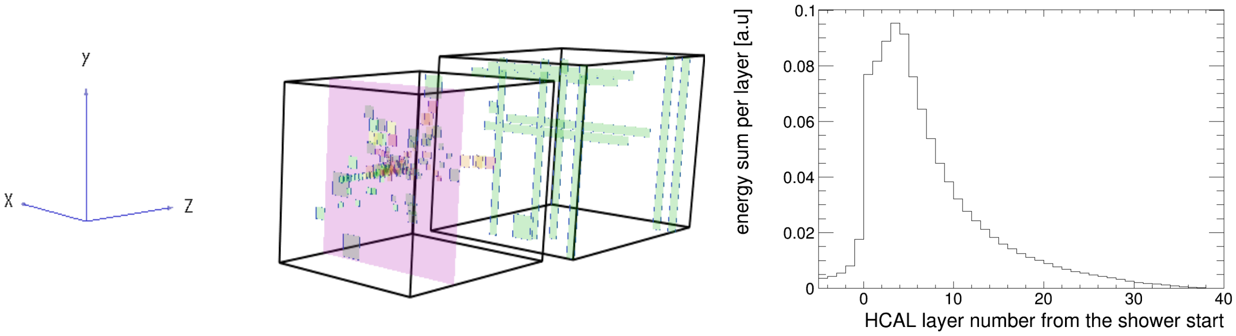

Figure 2: *a) An example of the event display of a shower produced by a 10 GeV pion in the CALICE ECAL, HCAL and Tail Catcher. Here the HCAL layer in which the shower starting point is found is highlighting; b) Longitudinal profile in the HCAL produced by 10 GeV pion*

with w_j: fractions length if the interaction length X_0 is in mm.

The official PDG values of the interaction lengths of the different absorber materials can be found at [7], by clicking on the corresponding element. One Fe-HCAL layer contained 16 mm Fe absorber, and another 4 mm steel cassette for holding the scintillator tiles. In the W-HCAL case, one layer contained 10 mm W, plus the 4 mm Fe for the cassette. For simplicity, the interaction length due to the scintillator tiles and to the other components in the detector can be neglected.

To study its energy dependence you need to look at the distribution of the layer where the shower start, expressed in interaction length.

You can find the macro in

```
cd /afs/desy.de/group/flc/pool/provenza/EDIT_new/FirstPart/Analysis/macros
root -l interactionLength_pion.C
```

How do you express the layers in terms of the interaction length?

B) Particle flow algorithm at work

The goal of part B is to understand the principles of particle flow reconstruction. In the interactive program the different steps of the algorithm - track and cluster reconstruction, cluster association and re-clustering - can be executed sequentially. Their results can be monitored in the event display, using color codes and pop-up information boxes.

The exercise uses complex Monte Carlo di-jet events generated at 91 GeV and 500 GeV center-of-mass energy and simulated in a 4π collider detector. It starts with event-by-event analysis of difficult reconstruction situations, and proceeds to a quantitative performance evaluation, comparing with purely calorimetric measurements at the different jet energies.

Tools: Pandora monitoring program [5], TEve event display

Data: MC di-jet events at a center of mass energy of 91 GeV and 500 GeV, selected with polar angles in the barrel region

Taks:

1. Explore event display, display options, color coding
2. Study display in different reconstruction stages MC truth, different particle types
    - tracks, hits, clusters, particle hypotheses
    - particle flow objects
    - look for different particles, try to find K_s^0, π_0, neutrons
3. Study event energy reconstruction
    - look simple and for difficult situations with close-by showers
    - study the confusion matrix for a few events, i.e. which kind of true energy deposition is reconstructed as what type of particle flow objects
4. Scan 20 events at each cms energy, and for outliers try to find out what went wrong or was difficult

[optional]

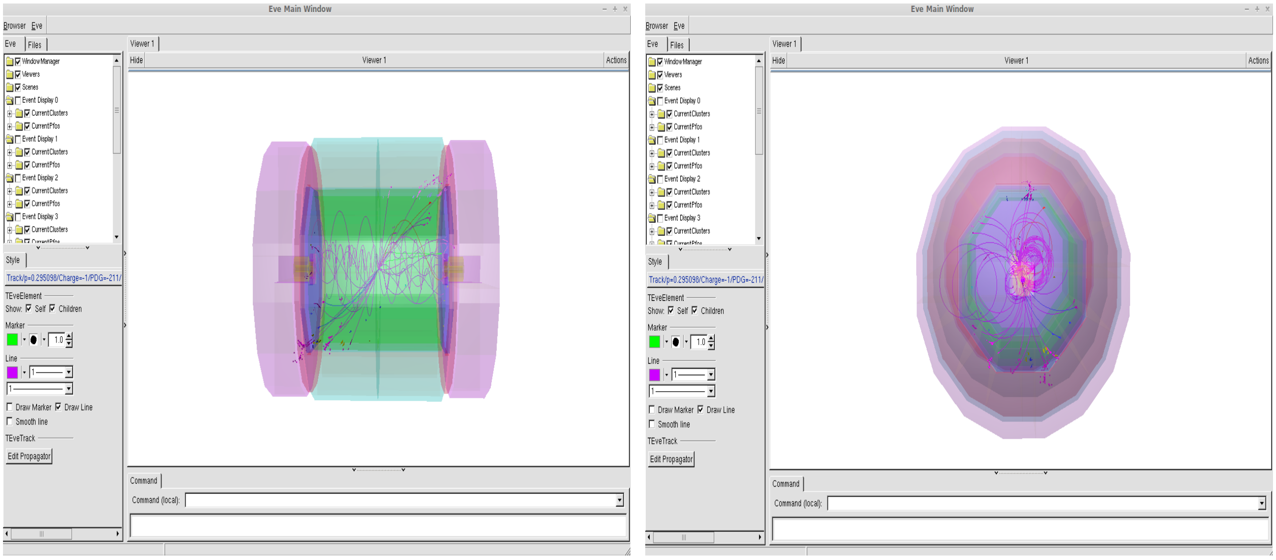

Figure 3: *Example of an event display at 91 GeV:left - (yz) plane, right - (xy) plane.*

5. Extract an average confusion matrix for 10 events at each energy

[optional]

Instructions

Also in this case the inizialization of the ilc software is required:

```
source /afs/desy.de/group/flc/pool/provenza/MyIlcSoft/init_ilcsoft.sh
```

The needed steering files can be found in the subdirectory MyConfig and you can run Marlin like

```
cd /afs/desy.de/group/flc/pool/provenza/MyIlcSoft/MyConfig
./myMarlin1706 JER_5x5_30x30_91GeV.xml (JER_5x5_30x30_500GeV.xml)
```

In the same directory the output root file is created ($ILD\_o1\_v06\_uds91(500)\_00\_0\_Default.root$). Here you can have a look at the energy sum reconstructed only with calorimeter and the energy sum reconstructed with the particle flow algorithm (PfoEnergyTotal and energyTotalPureClusters). An example of such an event display at 91 GeV is shown in figure 3.

Please note that you can change the angle of observation by moving the detector with the mouse and you can obtain information about the particle just pointing with the mouse on the track.

About the used colour coding:

- **Cluster**
  - magenta: clusters linked to a track
  - blue: clusters without a track
  - yellow: electromagnetic clusters
- **Current PFOs**: colour coding according to the reconstructed PDG. For the color of the tracks in the event display and the PDF indices see the Table A.1 in the Appendix A.

Regarding the exercise 3, the confusion matrix is printed for each event. This shows the amount of energy (in GeV) reconstructed as true energy and not. The second time the same quantities are expressed in percentage.

At the the end the confusion matrix for all the events is printed.

Furthermore the "confusion covariance" is printed.

Here the following quantities (in percentage) are computed:

- **photon/neutral reconstructed as charged**:

  We can define:

  $$(confusionA)_i = \left(\frac{E_{(\gamma RecoAsTrack)} + E_{(neutralRecoAsTrack)}}{\sum E}\right)_i \cdot 100$$

  with: i event, $E_{(\gamma RecoAsTrack)_i}$ amount of energy of photon reconstructed as charged particles for the event i, $E_{(neutralRecoAsTrack)_i}$ amount of energy of neutral hadrons reconstructed as charged particles for the event i, $(\sum E)_i$ total energy for that event; the total confusion is :

  $$confusionA = \frac{\sum_i (confusionA)_i}{N}$$

where N is the total number of events.

- **rms**:

  $rmsA = (\frac{\sum (confusionA)_i^2}{N} - (confusionA)^2)^{1/2}$

- **charged reconstructed as neutral/photon**:

  We can define, like we did before, the following quantities:
  $(confusionB)_i = (\frac{E_{(TrackRecoAsNeutral)} + E_{(TrackRecoAs\gamma)}}{\sum E})_i \cdot 100$
  with: i event, $E_{(TrackRecoAsNeutral)}$ amount of energy of charged particles reconstructed as neutral hadrons for the event i, $E_{(TrackRecoAs\gamma)}$ amount of energy of charged particles reconstructed as photon for the event i, $\sum E)_i$ total energy for that event;
  Also in this case the total confusion is:
  $confusionB = \frac{\sum_i (confusionB)_i}{N}$

- **rms** is computed like before

  $rmsB = (\frac{\sum (confusionB)_i^2}{N} - (confusionB)^2)^{1/2}$

- **covariance**:

  $covariance = \frac{\sum ((confusionA)_i \cdot (confusionB)_i)}{N} - [(confusionA) \cdot (confusionB)]$

- **correlation coefficient**:

  $correlation\ coefficient = \frac{covariance}{rmsA \cdot rmsB}$

- **rms total confusion**:

  $(rmsA)^2 + (rmsB)^2 + 2 \cdot correlation\ coefficient \cdot rmsA \cdot rmsB$

- **rms net confusion**:

  $(rmsA)^2 + (rmsB)^2 - 2 \cdot correlation\ coefficient \cdot rmsA \cdot rmsB$

To compare the energy sum using the standard calorimeter approach and using the particle flow approach, for the 2 selected energies, please see:

```
cd /afs/desy.de/group/flc/pool/provenza/EDIT_new/SecondPart/Analysis/macro
root -l plots.C
```

References

[1] CED event display: `http://ilcsoft.desy.de/portal/software\_packages/ced/`

[2] ILC software: `http://ilcsoft.desy.de/portal/software\_packages/`

[3] `http://www-flc.desy.de/hcal/calice\_soft/pro\_test/doc/ADDONPROCS/html/classCALICE\_1\_1EventDisplayProcessor.html`

[4] ROOT: `http://root.cern.ch/drupal/`

[5] Pandora PFA algorithm: `http://www.hep.phy.cam.ac.uk/twiki/bin/view/Main/PandoraPFA`

[6] `http://pdg.lbl.gov/2010/reviews/rpp2010-rev-passage-particles-matter.pdf`

[7] `http://pdg.lbl.gov/2010/AtomicNuclearProperties/index.html`

[8] `http://pdg.lbl.gov/2010/reviews/rpp2010-rev-monte-carlo-numbering.pdf`

Appendix A

Tracks color

particle	**PDG particle code**	**Color**
γ	22	dark yellow
e^-	11	light blue
e^+	-11	light red
μ^-	13	blue
μ^+	-13	red
τ^-	15	dark blue
τ^+	-15	dark red
ν_e	12	dark blue
$\bar{\nu_e}$	-12	dark red
ν_τ	16	dark blue
$\bar{\nu_\tau}$	-16	dark red
ν_μ	14	dark blue
$\bar{\nu_\mu}$	-14	dark red
π^+	211	magenta
π^-	-211	violet
π^0	111	light green
λ	3122	dark green
$\bar{\lambda}$	-3122	dark green
K^+	321	dark green
K^-	-321	dark green
K_s	310	light green
K_L	130	green
σ^-	3112	green
σ^+	3222	green
$proton$	2212	orange
$neutron$	2112	cyan

Table A.1: *PDG particle's code and color.*