SDHCAL evolution : Recent developments and algorithmic improvements for the APRIL Particle Flow

Tanguy Pasquier (speaker), Gérald Grenier, Imad Laktineh, Bo Li, Rémi Été

IP2I/Univ Lyon 1

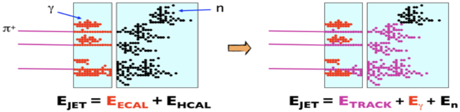
ECFA ECR Workshop
October 8, 2024

Contents

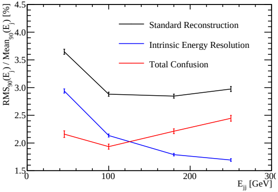
Particle flow calorimetry

Particle Flow Algorithm (PFA)

- PFA is the approach chosen for future colliders
- Use optimal sub-detector for jet energy estimation : tracker ($\sim 60\%$), ECAL ($\sim 30\%$), HCAL ($\sim 10\%$).
- Separate energy depositions from close-by particles : high granularity is key point



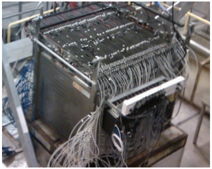
Extensive studies have been done with ILD detector option 1 (AHCAL, ILD baseline) and PandoraPFA algorithm.
At higher jet energy ($E \gtrsim 100$ GeV), dominant contribution to resolution is confusion.
See Steven Green, Cambridge University Thesis 2017

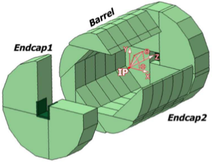


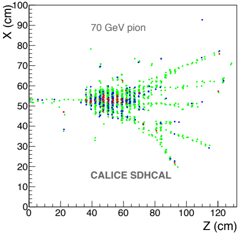
Semi-Digital HCAL

SDHCAL energy reconstruction

$$E_{reco} = \alpha_1 N_1 + \alpha_2 N_2 + \alpha_3 N_3$$







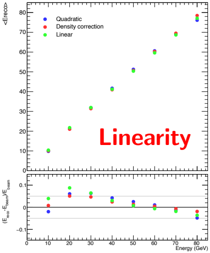
Options for E_{reco}

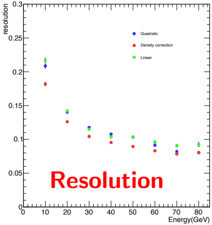
Thresholds(pC): 0.11, 5, 15

Quadratic $\alpha_i = a_i N_{hit}^2 + b_i N_{hit} + c_i$ (TB default)

Linear α_i constant (Pandora default).

Density $\alpha_i =$ above $\times$ f(N neighbour hits)



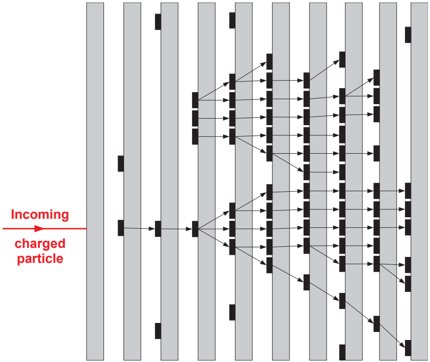


The APRIL algorithm

- Based on the Arbor concept and implemented in PandoraSDK

APRIL : Algorithm for Particle Reconstruction at ILC from Lyon.

- Based on graph theory : Reconstruct the showers as spatial trees
- ⇒ APRIL usable as long as you have the positions of the hits



Contents

Angular corrections

- Recent work on ILD Option 2 (SDHCAL and Videau Geometry) to improve performances
- Linear reconstruction of single KLong samples → HCAL calibration procedure
- E_{reco} too low → Angular corrections needed (φ, θ)
- Low energy range still needed improvements after correction

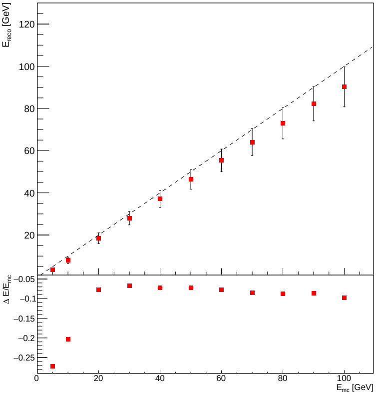


Figure: Before correction

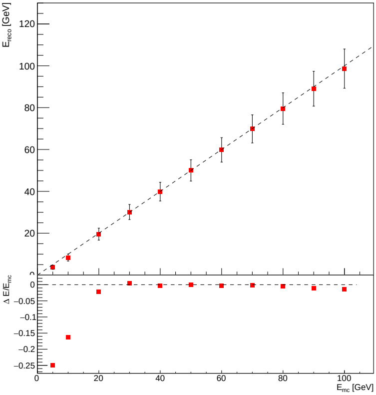
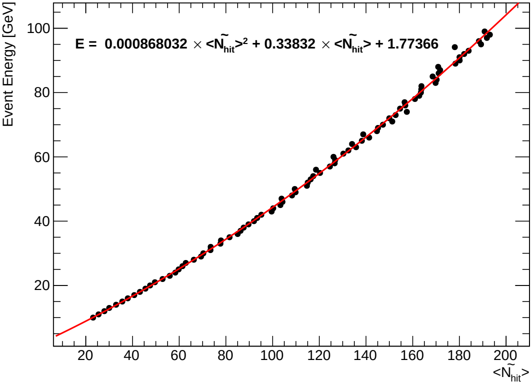


Figure: After correction

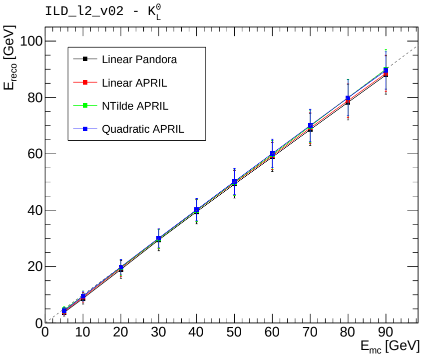
New reconstruction method

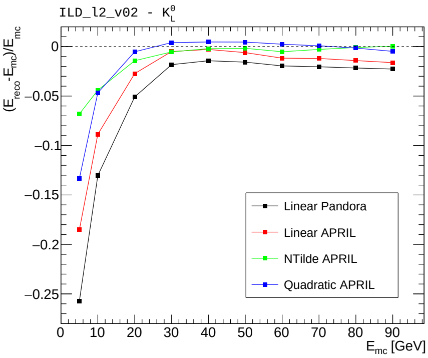
- Quadratic method : works well but implies a lot of parameters + difficult to constrain them
- We define : $\widetilde{N}_{hit} = N_3 + \beta \times (\alpha \times N_1 + N_2)$
- Look at correlations between N_1, N_2, N_3 to extract parameters
- $E =$ quadratic function of $\widetilde{N}_{hit}$
- 5 parameters in total



Results

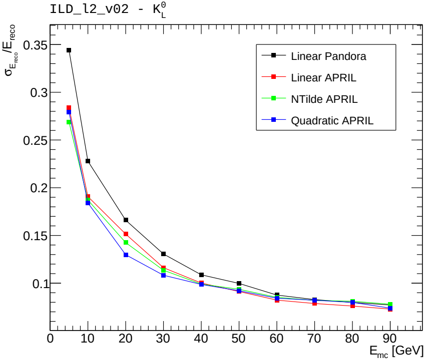
- Achieve good linearity with the different methods for $E \geq 20$ GeV
- New method gives the best linearity on average

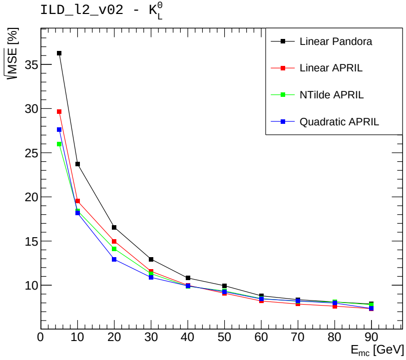




Results

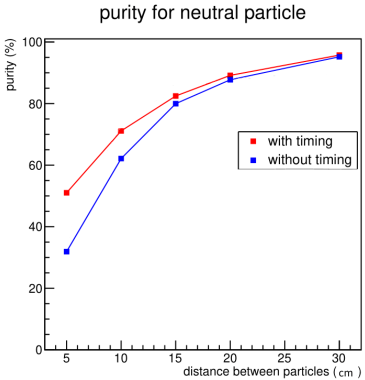
- MSE $= \left(\frac{\sigma_{E_{\text{reco}}}}{E_{\text{reco}}}\right)^2 + \left(\frac{E_{\text{reco}} - E_{\text{mc}}}{E_{\text{mc}}}\right)^2$
- Best resolution and MSE with NTilde method for $E \leq 10$ GeV
- Quadratic still better from 10 GeV to 40 GeV
- All methods tend to converge for higher energies

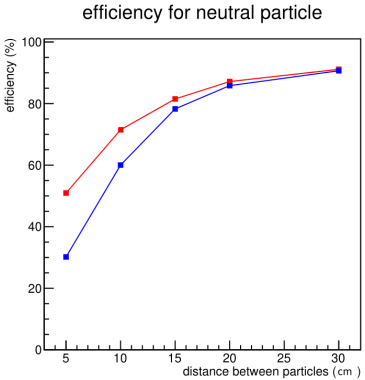




PFA with timing

- SDHCAL should allow timing segmentation < 100 ps thanks to MRPC
- Possibility to follow the particles in "real time"
- Previous studies : timing improves separation
- **Goal** : Add timing to APRIL





Different applications

- Delete non-causal connectors between hits $(\beta > 1)$
- Late neutrons tagging to treat them separately
- Identify the seeds of the showers and count them
- Hit ordering by time instead of radius
- Each of the above can be added with a PandoraSDK algorithm

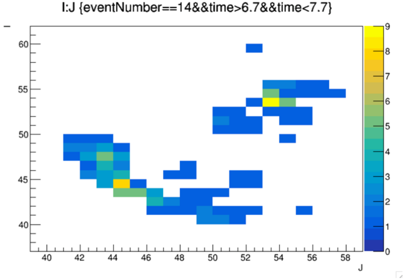


Figure: 1 ns interval

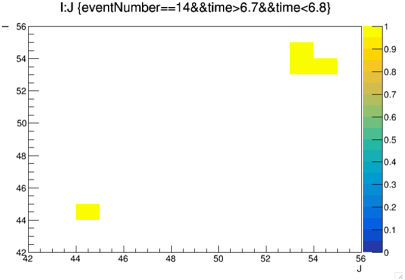


Figure: 100 ps interval

Contents

Summary and outlook

Summary

- Angular corrections of the SDHCAL have been added
- A new way of computing the SDHCAL energy has been found and gives encouraging results on samples of single KLong
- Timing inclusion in APRIL has started

Other activities

- Official inclusion of a "Marlin" processor able to run APRIL PFA and the SDHCAL plugins is underway
- Implementation of split cluster procedure (AMSTER) has started

Outlook

- Reconstruct jet events to compare the different methods
- Fully include timing in APRIL

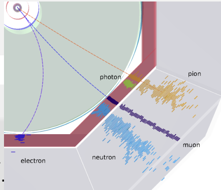
Thank you for your attention !

Backup

Other lab activities

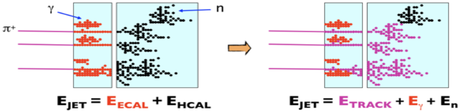
- Areas of interest for future colliders:
  - Monolithic CMOS sensors (Tracking system)
  - Full simulation software development (Geometry, reconstruction)
  - Crystal ECAL detector

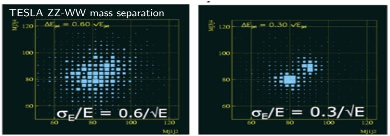
Particle flow calorimetry



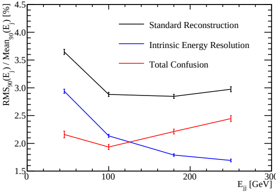
Particle Flow Algorithm (PFA)

- ILC/CEPC physics program requires W/Z$\rightarrow q\bar{q}$ mass separation.
- $\Rightarrow$ jets resolution $[50, 500]$ GeV better than $\sim 3-4$ % $\sim 30\%/\sqrt{E}$.
- Use optimal sub-detector for jet energy estimation :
  tracker ($\sim 60\%$), ECAL ($\sim 30\%$), HCAL ($\sim 10\%$).
- Separate energy depositions from close-by particles : high granularity is key point





Extensive studies have been done with ILD detector option 1 and PandoraPFA algorithm.
At higher jet energy (E$\gtrsim$100 GeV), dominant contribution to resolution is confusion.
See Steven Green, Cambridge University Thesis 2017

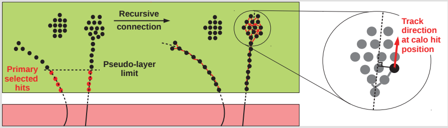


The APRIL algorithm

APRIL : Algorithm for Particle Reconstruction at ILC from Lyon.

The clustering strategy

- start from tracks (track driven clustering) extrapolate tracks in calorimeters → cluster hits close to the tracks.



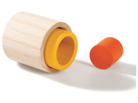
- Perform Arbor Clustering with all hits.
- → Clusters containing track cluster define charged clusters.
- Arbor parameters set to avoid making big clusters.
- → Some hits remain unclustered.
- Nearby hits merging : remaining unclustered hits are clustered with mlpack DBSCAN (efficient Nearest Neighbour clustering)
- If $E_{track} > E_{cluster}$, merge nearby cluster.

Arbor clustering

Graph theory : a shower is an oriented tree.

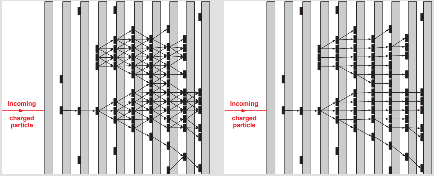
Orientation

- Rearrange hits in virtual nested cylinders (= pseudo layers)
- Count them from the inside.
- Forward direction = increase pseudo layer number.

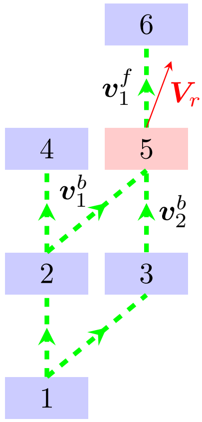


Arbor

1. Connect all neighbouring hits (use mlpack NeighborSearch).
2. Clean connectors = keep max one backward connection per hit.



Connector cleaning



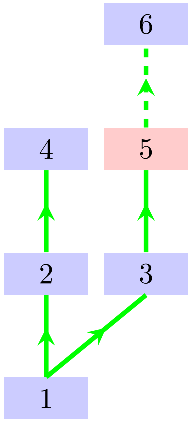
Reference direction

$$\mathbf{V}_r = w_b \times \sum_i \mathbf{v}_i^b + w_f \times \sum_j \mathbf{v}_j^f$$

- Can use different depth (max distance in pseudolayer) and weight in forward-backward direction.
- Define angle θ with respect to $\mathbf{V}_r$ and distance d between hits (in λ_I or X_0 unit).
- Keep connector with smallest

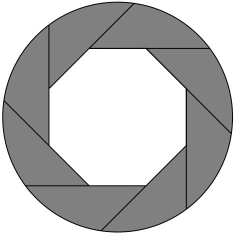
$$\kappa = \theta^{p_\theta} \times d^{p_d}$$

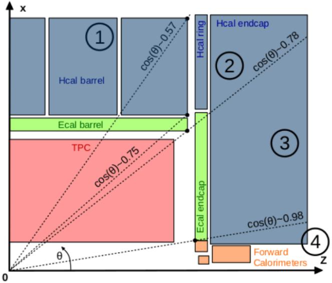
(if $\theta = 0$, smallest d)



Angle corrections

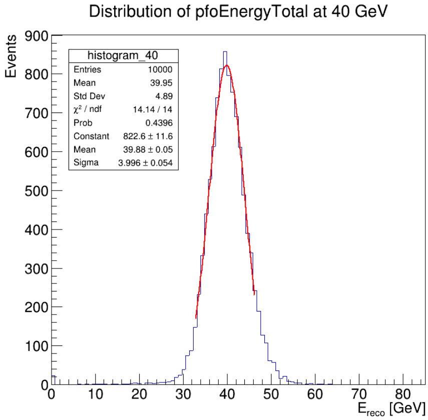
- **Goal** : implement angle corrections
- Purely geometric corrections
  - $N_{\text{hit}}^{\text{new}} = N_{\text{hit}} \times \text{Effect}$
  - Effect $\frac{1}{\cos\theta}$ for endcap
  - Effect $\frac{1}{\sin\theta}$ for barrel
  - Effect $\frac{1}{\cos\varphi}$ for barrel only
  - Videau geometry taken in consideration
- Created SDHCALContent for all SDHCAL related plugins ▸ Git repo
- Separating detector (SDHCAL, ILD option 2) from PFA (APRIL).





Procedure

- Simulate and reconstruct samples of KLong going straight into the barrel ($\theta = \varphi = \frac{\pi}{2}$)
- For each energy, we look at the `pfoEnergyTotal` distribution
- Distributions fitted with Gaussian function in a RMS90 range (smallest interval with $\geq$ 90% of the events)
- Mean value and standard deviation of the Gaussian function treated as the reconstructed energy and its resolution



α_i parameters computation

- Different methods tried
  - Classical method : χ^2 minimization
  - Solving system of equations with GSL
  - New reconstruction method
- For quadratic : $\alpha_i = a_i N_{hit}^2 + b_i N_{hit} + c_i$ with $N_{hit} = N_1 + N_2 + N_3$

Classical method : χ^2 minimization

- $\chi^2 = \sum_{i=1}^{N} \frac{(E_{\text{mc,i}} - E_{\text{reco,i}})^2}{E_{\text{mc,i}}}$, N : Number of events used
- Minimization with TMinuit and MIGRAD (ROOT)

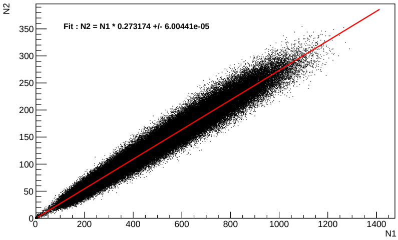
Solving system of equations with GSL

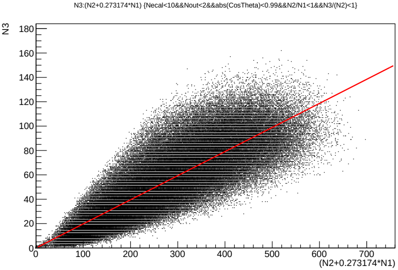
- 9 equations with 9 unknowns and E_{reco} known
- Compute the N means for each of the 9 energies : $< N_1 >, < N_{hit} N_1 >, \ldots$
- Solving with GSL (GNU Scientific Library)

- Cons :
  - A lot of parameters to adjust (9 for the quadratic formula)
  - Difficult to constrain the parameters

New reconstruction method

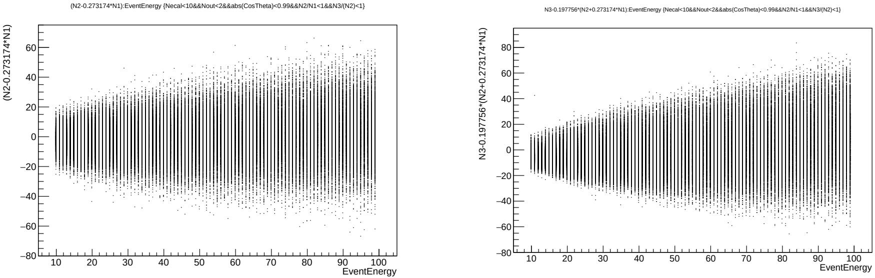
- Initial observation : N_2 behaves approximately as a fraction of N_1, and N_3 behaves approximately as a fraction of the combination of N_1 and N_2
- $N_2 = 0.273174 \times N_1$
- $N_3 = 0.197756 \times (0.273174 \times N_1 + N_2)$





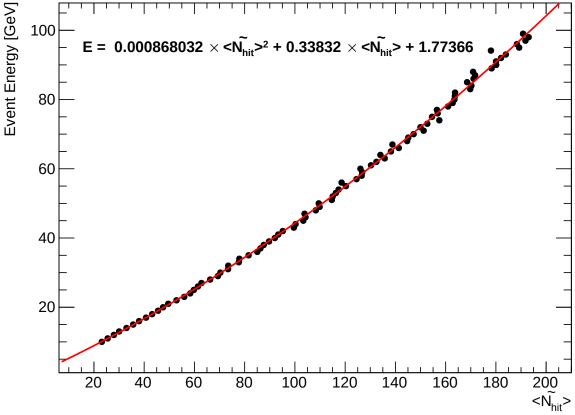
Checking the coefficients

- Averages around zero when we look at the difference between the different terms as a function of energy
- Relations remain correct no matter what the energy is



New formula

- We define : $\tilde{N_{hit}} = N_3 + 0.197756 \times (0.273174 \times N_1 + N_2)$
- We plot $E = f(< \tilde{N_{hit}} >)$
- Best results for $f(< \tilde{N_{hit}} >)$ as quadratic



Reclustering / Cluster cutting

- Reclustering : Break the cluster and restart clustering with different parameters (Pandora strategy)
- Cluster cutting : cut the cluster and remove some parts

AMSTER algorithm

- Based on graph theory with Minimum Spanning Tree (MST)
- Cut the connections with bigger weights
- Implementation started and tested on simple cases

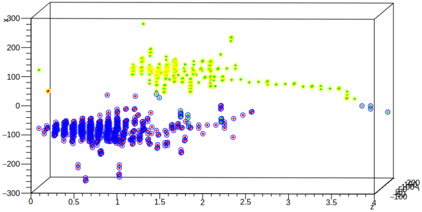


Figure: Event of a π^- (red) and a K^0 (green) in two clusters (blue and yellow)