APRIL : a particle flow algorithm for future e^+e^- colliders

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

IP2I/Univ Lyon 1

2024 CEPC International Workshop EU Edition, Marseille
April 9, 2024

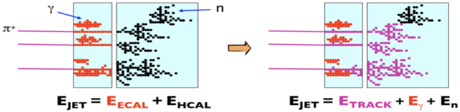
Plan

Contents

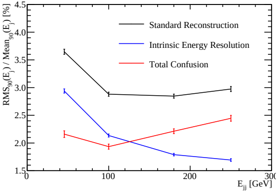
Particle flow calorimetry

Particle Flow Algorithm (PFA)

- PFA is the approach chosen for ILC and CEPC
- 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



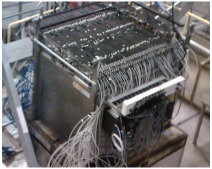
Extensive studies have been done with ILD detector option 1 (AHCAL, ILD baseline) and PandoraPFA algorithm.
At higher jet energy ($E \gtrsim 100$ GeV), dominant contribution to resolution is confusion.
See Steven Green, Cambridge University Thesis 2017

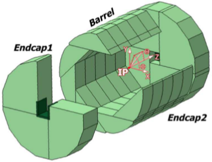


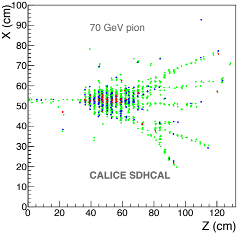
Semi-Digital HCAL

SDHCAL energy reconstruction

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







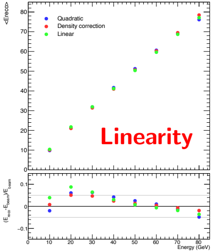
Options for E_{reco}

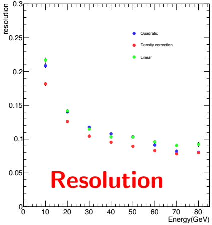
Thresholds(pC): 0.11, 5, 15

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)





Jets in CEPC

Jets reconstruction in CEPC

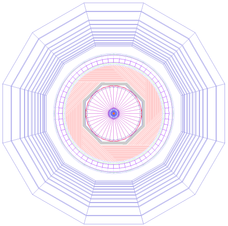
Baseline detectors Si-WECAL and SDHCAL.

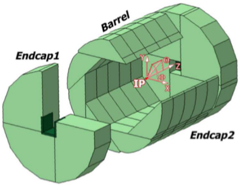
Geometry ILD option 2, "Videau" Geometry.

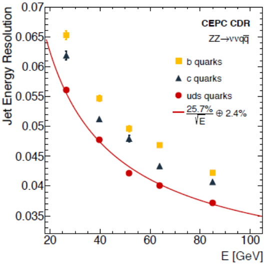
PFA ARBOR

Jet algorithm Durham.

Contrary to PandoraPFA jet studies with ILD, CEPC jet studies uses jet algorithm.

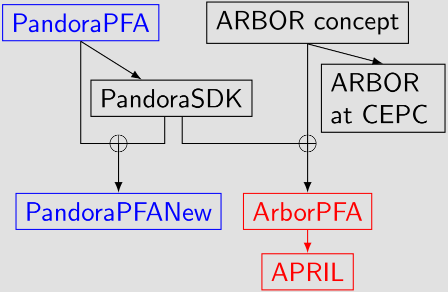






PFA history

ILD PFA reconstruction



PFA strategy

Both PandoraPFA and ArborPFA, construct many small clusters then merge them.

- APRIL $\simeq$ ARBOR concept + PandoraSDK algorithms

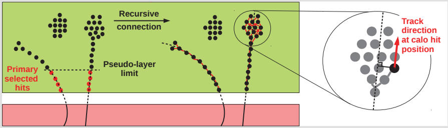
Contents

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.



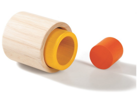
- Perform Arbor Clustering with all hits.
- → Clusters containing track cluster define charged clusters.
- Arbor parameters set to avoid making big clusters.
- → Some hits remain unclustered.
- Nearby hits merging : remaining unclustered hits are clustered with mlpack DBSCAN (efficient Nearest Neighbour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

Arbor clustering

Graph theory : a shower is an oriented tree.

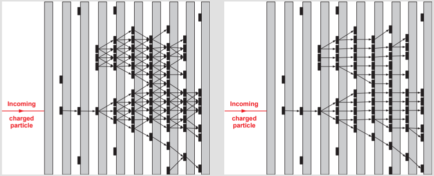
Orientation

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

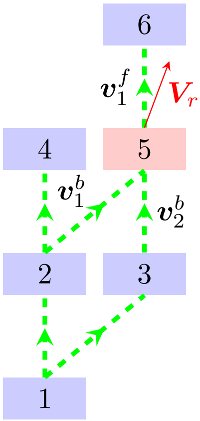


Arbor

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



Connector cleaning



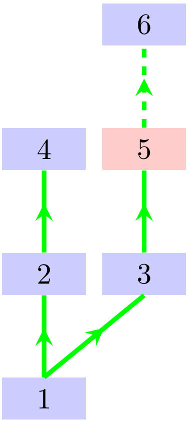
Reference direction

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

- 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)



Results

- Event samples: $e^+e^- \to q\bar{q}$, where $q = u, d, s$ $(|\cos\theta_q| < 0.7)$
- With ILD option 2 large (SDHCAL), PandoraSDK, ILCSoft, linear reconstruction
- Jet energy resolution, JER $= \frac{\mathrm{RMS}_{90}(\mathrm{E}_j)}{\mathrm{Mean}_{90}(\mathrm{E}_j)} = \sqrt{2} \cdot \frac{\mathrm{RMS}_{90}(\mathrm{E}_{jj})}{\mathrm{mean}_{90}(\mathrm{E}_{jj})}$
- JER at 91.2 GeV: APRIL: 4.2%; Pandora: 4.1%; Perfect PFA: 3.25%

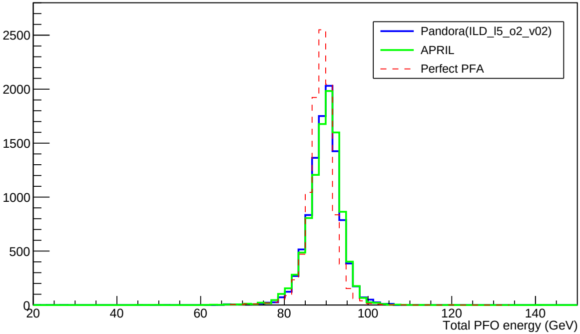
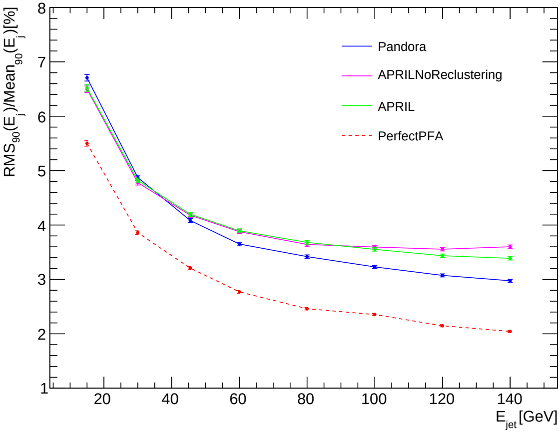


Figure: The energy of reconstructed PFO at $E_{\mathrm{CM}} = 91.2$ GeV.

Results (continued)



- (PandoraPFA without reclustering) $\lesssim$ APRIL performance $<$ (PandoraPFA with reclustering)
- APRIL has no reclustering $\rightarrow$ tried simple approach.

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

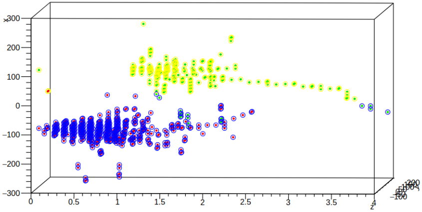


Figure: Event of a π^- (red) and a K^0 (green) in two clusters (blue and yellow)

SDHCAL corrections

- Recent work on ILD Option 2 (SDHCAL and Videau Geometry) to improve performances
- SDHCAL needed theta and phi angle correction $\rightarrow E_{\text{rec}}$ too low
- E_{rec} is now corrected with both angles
- Created an official SDHCALContent repository for all SDHCAL related plugins ▸ Git repo
- Separating detector (SDHCAL, ILD option 2) from PFA (APRIL).

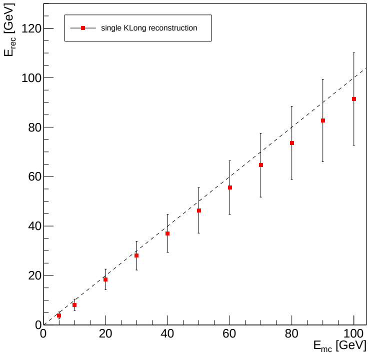


Figure: E_{rec} before correction

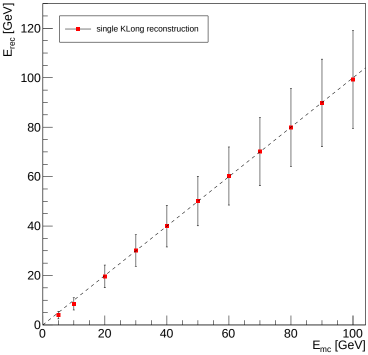
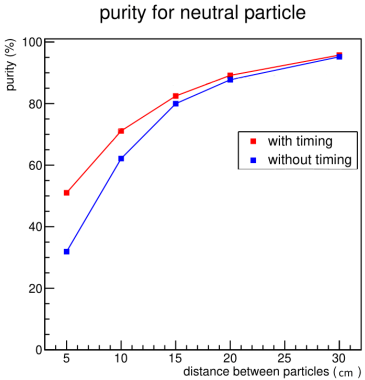


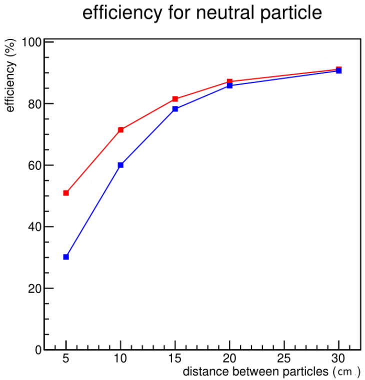
Figure: E_{rec} after correction

Contents

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
- Pseudo layers ordered in timing and not in space
- Each of the above can be added with a PandoraSDK algorithm

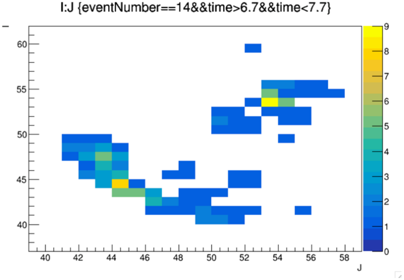


Figure: 1 ns resolution

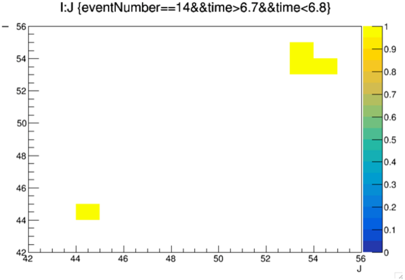


Figure: 100 ps resolution

Contents

Summary

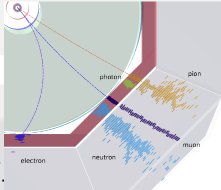
- A particle flow algorithm implementing the ARBOR approach has been developed in the PandoraSDK framework.
- Competing with PandoraPFA at low and intermediate jet energies in SDHCAL.
- Work started for cluster cutting algorithm for APRIL is expected to improve the reconstruction at higher jet energies.
- Next steps
  - Implement split cluster procedure (AMSTER)
  - Fully include timing in APRIL

Thank you for your attention !

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under GA no 101004761

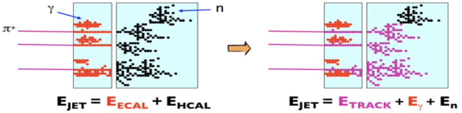
Backup

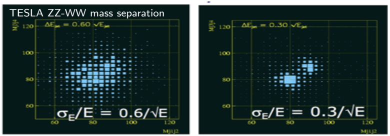
Particle flow calorimetry



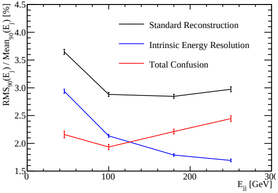
Particle Flow Algorithm (PFA)

- ILC/CEPC physics program requires W/Z→ $q\bar{q}$ mass separation.
- ⇒ 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



Charged cluster merging

Arbor like merging

- Connect charged cluster with all neighbour neutral clusters.
- Select neutral to charged backward connection with parameter order $\kappa = \theta^{p_\theta} \times d^{p_d}$
- Merge only if χ increase, staying below a maximum.
  - $\chi = (E_c - p_t)/\sigma_{E_c}$
  - σ_{E_c}
    - HCAL: $0.55/\sqrt{E_c}$ for hadrons.
    - ECAL: $0.15/\sqrt{E_c}$
- Cluster merging is still under optimisation.

Cluster geometry

- Cluster properties used to compute θ and d.
- For hit k, $e_k = \alpha_{threshold}$.
- The center of gravity of a cluster (COG)

$$\boldsymbol{o} = \frac{1}{\sum_k e_k} \sum_k e_k \boldsymbol{r}_k$$

- The cluster axis is computed from the eigen vector of inertial tensor (Principal Component Analysis)

$$I_{ij} = \sum_k e_k (\boldsymbol{r}_k^2 \delta_{ij} - \boldsymbol{r}_k^{(i)} \boldsymbol{r}_k^{(j)})$$

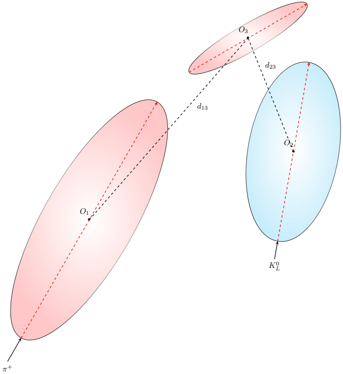
Merging clusters distances

For 2 big clusters

- d Distance of closest approach between the 2 axes.
- θ Angle between the 2 axes.

For a small cluster

- d Distance between the 2 COG.
- θ Angle between the 2 COG directions (from origin).



PFO creation

- Remaining neutral clusters to track association: add neutral cluster to existing or create new charged clusters, use position, direction and energy matching.
- PID
  - γ, $\pi^{\pm}$, neutral hadron; More particle categories (such as muon, electron) is to be considered.
  - Shower profile, energy deposition and track information are used.

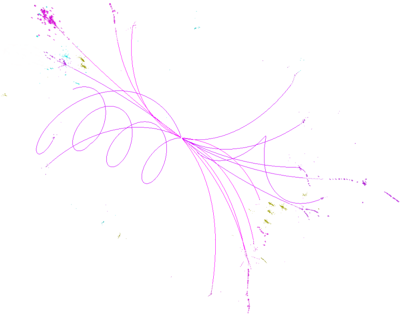


Figure: The reconstructed PFOs in an event of $E_{\mathrm{CM}} = 91.2\ \mathrm{GeV}$.

Simple "reclustering" approach

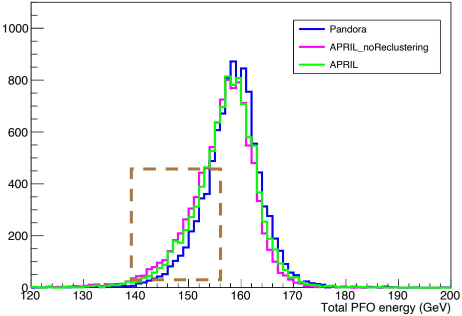


Figure: 160 GeV

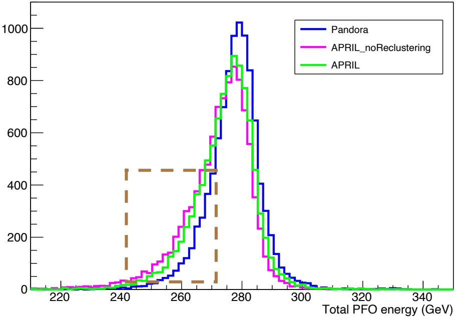


Figure: 280 GeV

- If $E_c > p_t$, remove hits from cluster until $E_c \simeq p_t$ and make a neutral hadron cluster with the removed hits.

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).