APRIL — a novel algorithm of particle reconstruction for the ILC

Bo Li

Institut de Physique des 2 Infinis de Lyon (IP2I)

LCWS 2019, Sendai
October 30, 2019

Outline

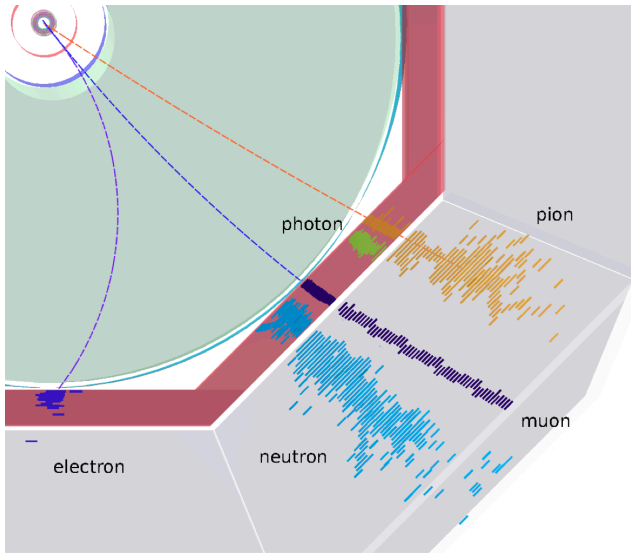
1. Introduction: Particle flow calorimetry and the ILD
2. The particle flow algorithm of APRIL
3. Summary

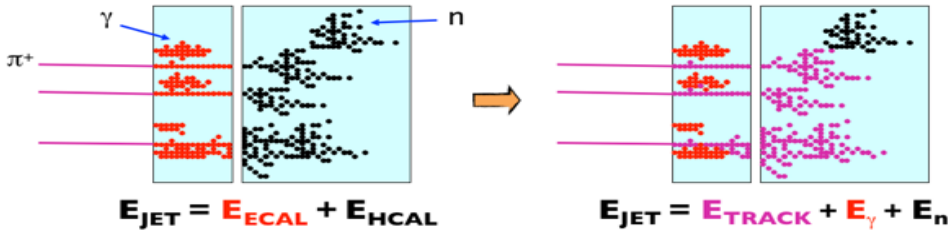
Outline

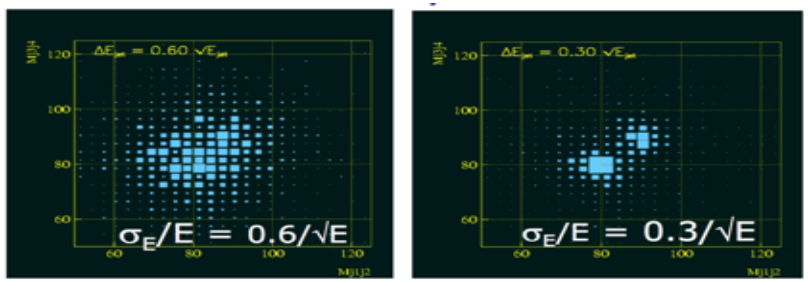
1. Introduction: Particle flow calorimetry and the ILD
2. The particle flow algorithm of APRIL
3. Summary

Particle flow calorimetry

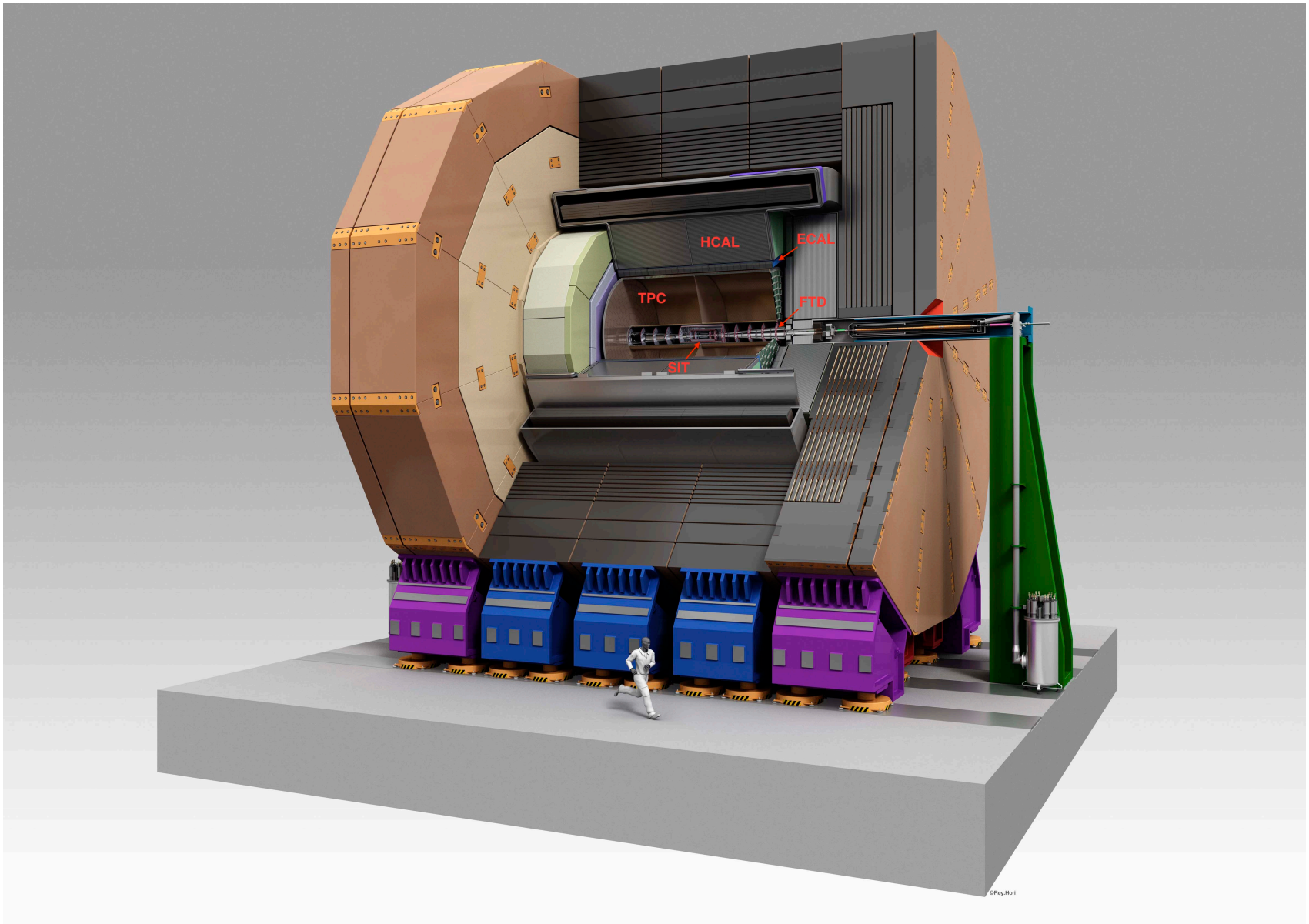
- Particle flow calorimetry: attempt to reconstruct visible final state particles from the information recorded by the detector.
- First result of Z mass resolution at 91 GeV with the Particle Flow Algorithm for the TESLA: 3.1 GeV [1] (ALEPH: 6 GeV).
- Jet energy resolution requirement at ILC: $\sigma_E/E = 3 \sim 4\%$ in the energy range from 50 to 500 GeV [2].
- Algorithms: Pandora [3], Arbor [4], Garlic [5] ...







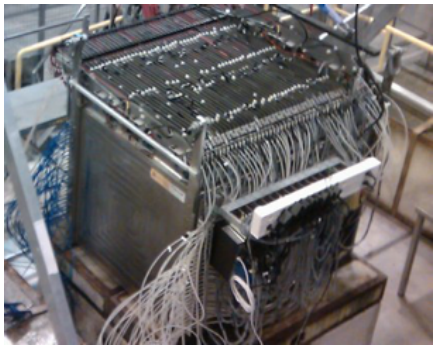
International Large Detector [6]

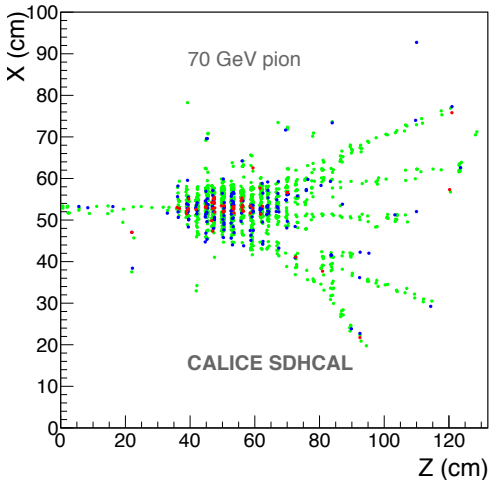


Semi-Digital HCAL

Two high granularity HCAL options at ILD

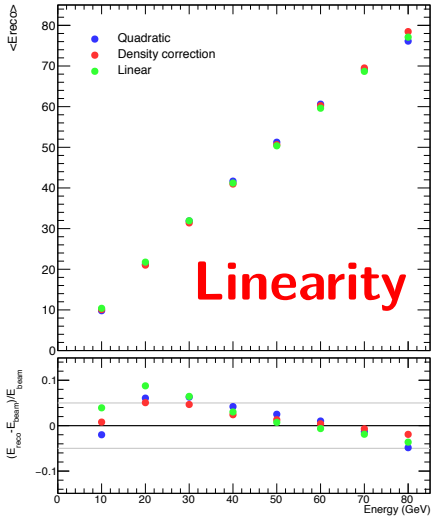
- Analog HCAL (AHCAL)
- Semi-Digital HCAL (SDHCAL [7]), implemented in ILD_l5_o2_v02.

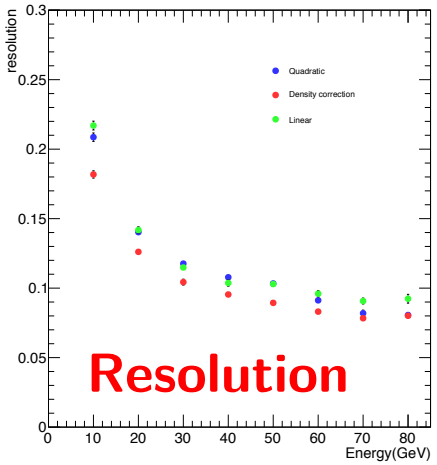




SDHCAL prototype

- 48 layers, $6\lambda_I$
- GRPC $(1 \times 1m^2)$
- Pad: 1×1 cm^2
- Thresholds(pC): 0.11, 5, 15
- Power-pulsing
- Self-supporting structure.



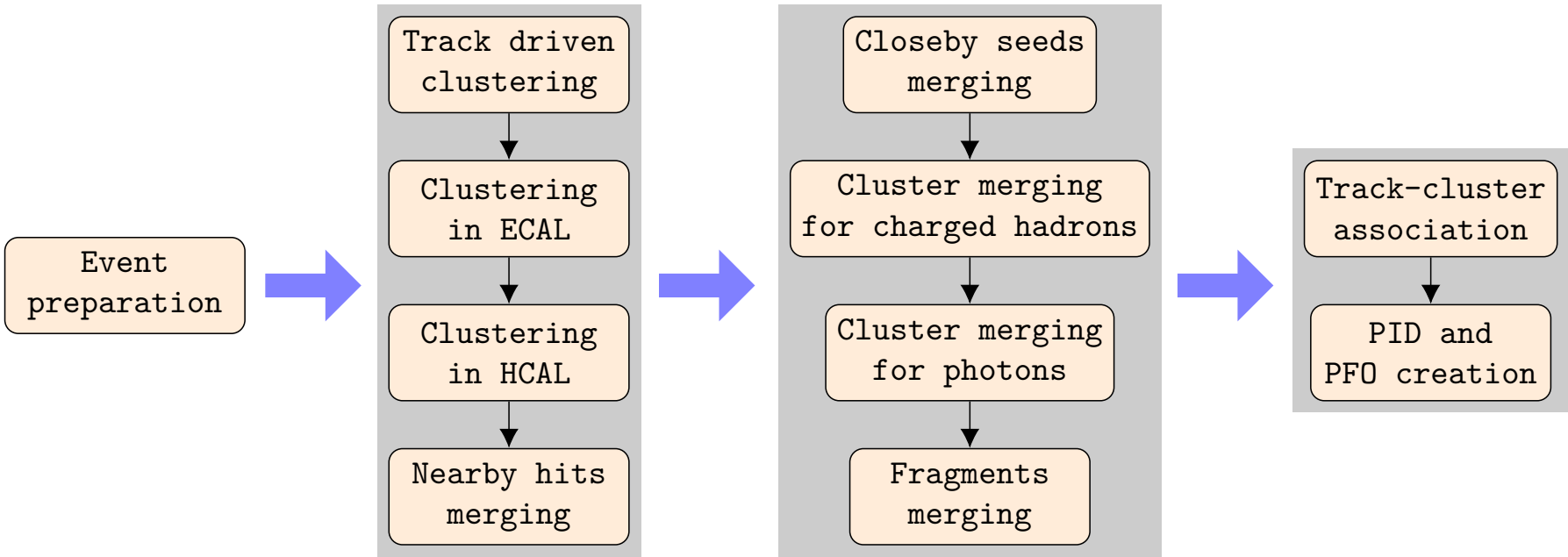


Outline

1. Introduction: Particle flow calorimetry and the ILD
2. The particle flow algorithm of APRIL
3. Summary

Design of algorithm

- The name of this PFA: APRIL (Algorithm of Particle Reconstruction for the ILC developed in Lyon)
- Jet energy composition: charged particles: $\sim 60\%$; photons: $\sim 30\%$; neutral hadrons: $\sim 10\%$.



Related software

- Algorithms developed by using the **PandoraSDK** [8]
  - Multi-algorithm approach
  - Objects: track, hit, cluster, PFO
- **ILCSoft** (https://github.com/iLCSoft)
  - **Marlin** [9]
  - Geometry: ILD detector model (ILD_l5_o2_v02) implemented in **lcgeo**, which is based on **DD4hep** [10]
  - Digitizers of SDHCAL: SimDigital in **MarlinReco**)
  - **LCCalibration**: automated energy calibration for calorimeters at ILC (https://github.com/iLCSoft/LCCalibration).
- **mlpack** [11]: NeighborSearch, DBSCAN(Density-based spatial clustering of applications with noise).

Clustering

- Arbor [4] : use it as the algorithm for clustering the hits in calorimeter with tree topology.
- Nearby hits are linked by *connector*. The nearby hits are searched by the NeighborSearch (and NearbySearch) in mlpack [11].
- *Reference direction*

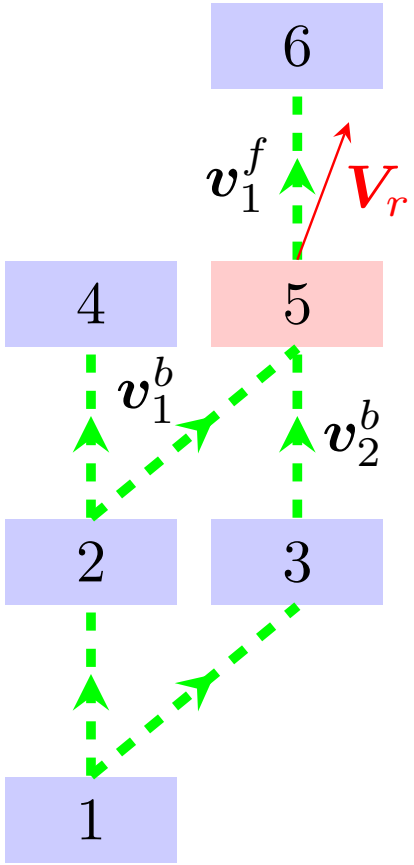
$$\mathbf{V}_r = w_b \times \sum_i \mathbf{v}_i^b + w_f \times \sum_j \mathbf{v}_j^f \tag{1}$$

- *Connector order*

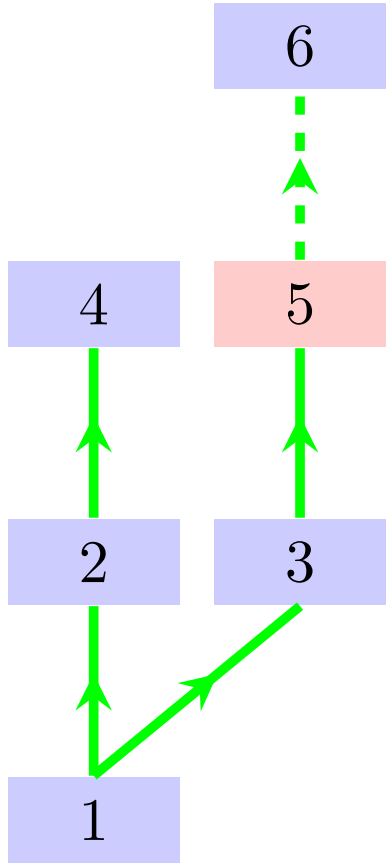
$$\kappa = \theta^{p_\theta} \times d^{p_d} \tag{2}$$

- Ambiguity: connector order at small angle, e.g., $\theta = 0$.
- Hits which are not clustered are dealt with by DBSCAN in mlpack.
- To restrain the error in clustering, the parameters are set to avoid forming big clusters.

Clustering (continued)



(a) Reference direction



(b) After clean

Cluster merging

- For cluster merging, the **order of cluster connection is proposed** for cluster merging in analog to $\kappa = \theta^{p_\theta} \times d^{p_d}$ in hit clustering, and the geometrical properties of cluster are utilized
  - Distances: COG distance; closest distance approach;
  - Angles: cluster axis; direction between clusters.
- The energy criteria for cluster merging
  - $\chi = (E_c - p_t)/\sigma_{E_c}$
  - σ_{E_c}
    - ★ HCAL: $0.55/\sqrt{E_c}$ for hadrons.
    - ★ ECAL: $0.15/\sqrt{E_c}$?
- Cluster merging is still under optimisation.

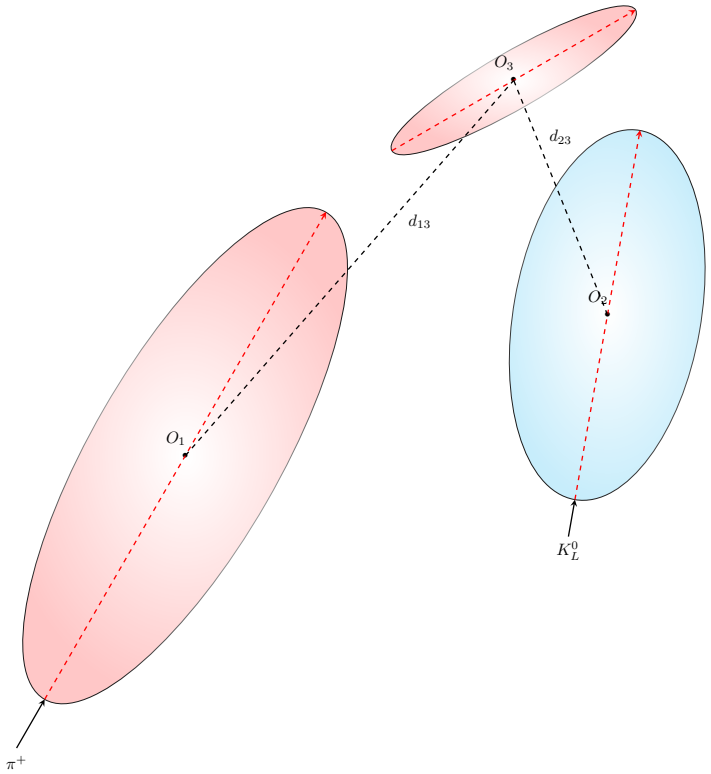


Figure: Merging between charged and neutral cluster.

Cluster properties

(a) Photon cluster (b) Charged cluster

The center of gravity of a cluster

$$\boldsymbol{o} = \frac{1}{E}\sum_k e_k \boldsymbol{r}_k \tag{3}$$

The cluster axis is computed from the eigen vector of inertial tensor

$$I_{ij} = e_k \sum_k (\boldsymbol{r}_k^2 \delta_{ij} - \boldsymbol{r}_k^{(i)} \boldsymbol{r}_k^{(j)}) \tag{4}$$

Distance of clusters axes

- For two clusters, 1 and 2, the equation of their axes can be represented by

$$\boldsymbol{y}_i = \boldsymbol{o}_i + \lambda_i \boldsymbol{a}_i \tag{5}$$

  where $i = 1, 2$; $\boldsymbol{o}_i$ is the center of gravity as a reference to the points on the axis; $\boldsymbol{a}_i$ is the direction of axis; λ_i is a free parameter which determines the position $\boldsymbol{y}_i$ on the axis. The distance of the two axes (skew lines) is given by

$$d = |\boldsymbol{n} \cdot (\boldsymbol{o}_2 - \boldsymbol{o}_1)| \tag{6}$$

  in which, $\boldsymbol{n} = (\boldsymbol{a}_1 \times \boldsymbol{a}_2)/|\boldsymbol{a}_1 \times \boldsymbol{a}_2|$.

- The cluster is used for fragment (*i.e.* the relatively small and far cluster form the main cluster) merging.

PFO creation

- Track-cluster assocaition: position, direction and energy are considered.
- PID
  - γ, $\pi^{\pm}$, neutal 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 $E_{\mathrm{CM}} = 91.2\ \mathrm{GeV}$.

Results

- Event samples: $e^+e^- \to q\bar{q}$, where $q = u, d, s$ $(|\cos\theta_q| < 0.7)$
- Detector model: ILD_l5_o2_v02
- 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%

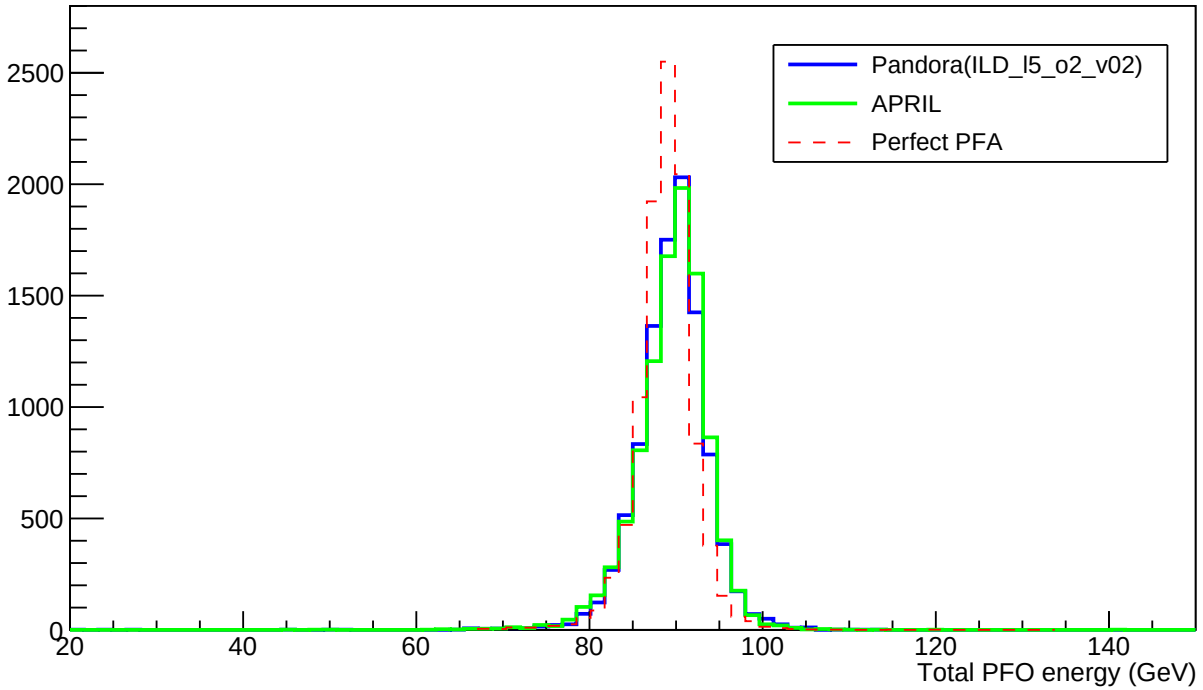


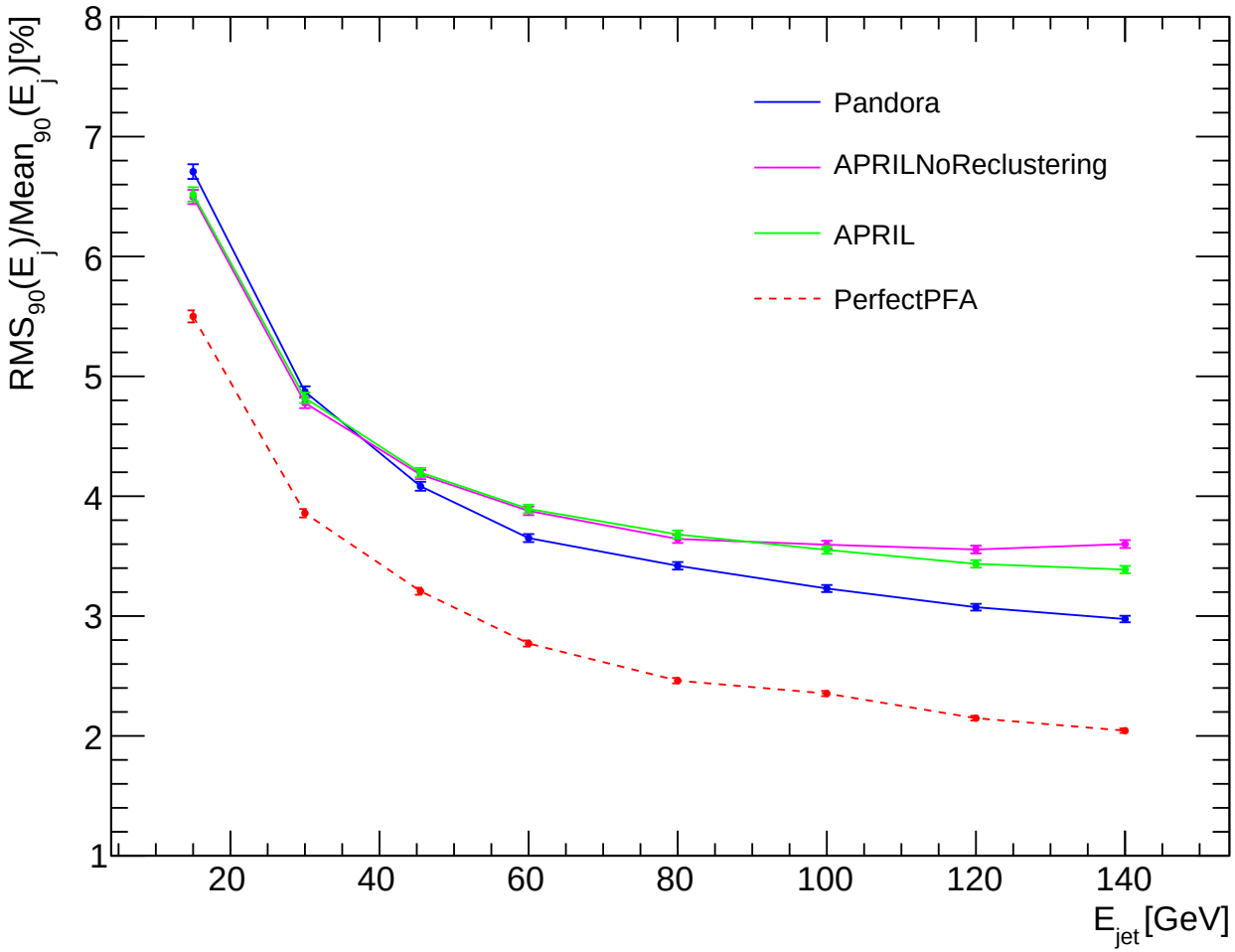
Figure: The energy of reconstructed PFO at $E_{\mathrm{CM}} = 91.2$ GeV.

Results (continued)

- Error estimation of each step at $E_{\mathrm{CM}} = 91.2\ \mathrm{GeV}$

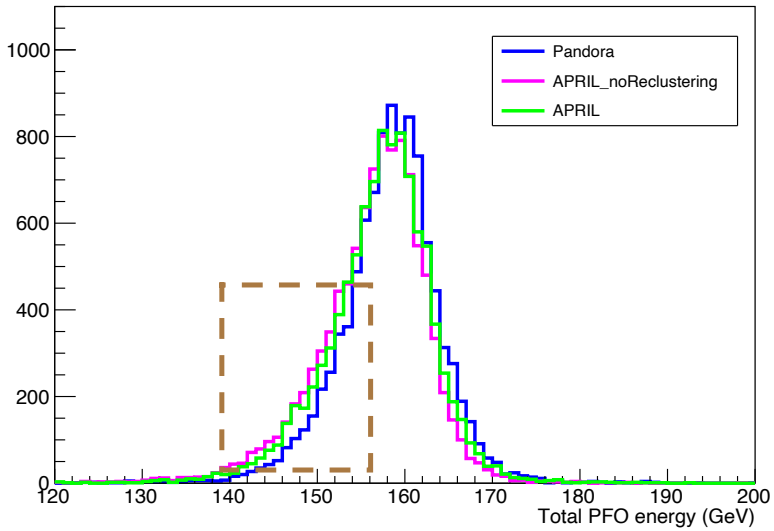
stage	error contribution
Clustering	$\sim 0.05\%$
Nearby hits merging	$\sim 0.15\%$
Cluster merging	$\sim 0.30\%$
Track-cluster association and PFO creation	$\sim 0.30\%$

Results (continued)

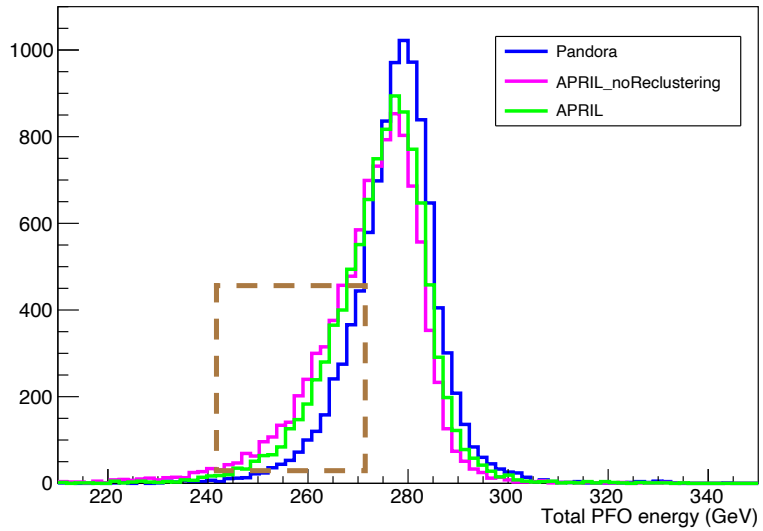


- A simple reclustering algorithm is tried for APRIL.

Simple reclustering



(a) 160 GeV



(b) 280 GeV

Outline

1. Introduction: Particle flow calorimetry and the ILD
2. The particle flow algorithm of APRIL
3. Summary

Summary

- A particle flow algorithm is deveoped in the framework of up-to-date ILCSoft.
- We proposed a cluster merging approach by constructing the cluster connection order from the computation of cluster geometrical properties.
- The code has great improvement from its first version, and the current result is quite close to our expectation.
- Plans
  - Optimisation: cluster merging, track-cluster association and PID.
  - Implementation: reclustering.
  - Far future: machine learning for PFA.

References I

[1] TESLA.

Technical Design Report, Part IV, A Detector for TESLA.

2001.

[2] J. S. Marshall and M. A. Thomson.

Pandora Particle Flow Algorithm.

In *Proceedings, International Conference on Calorimetry for the High Energy Frontier (CHEF 2013): Paris, France, April 22-25, 2013*, pages 305–315, 2013.

[3] M. A. Thomson.

Particle Flow Calorimetry and the PandoraPFA Algorithm.

Nucl. Instrum. Meth., A611:25–40, 2009.

[4] Manqi Ruan and Henri Videau.

Arbor, a new approach of the Particle Flow Algorithm.

In *Proceedings, International Conference on Calorimetry for the High Energy Frontier (CHEF 2013): Paris, France, April 22-25, 2013*, pages 316–324, 2013.

References II

[5] D. Jeans, J. C. Brient, and M. Reinhard.
GARLIC: GAmma Reconstruction at a LInear Collider experiment.
JINST, 7:P06003, 2012.

[6] Halina Abramowicz et al.
The International Linear Collider Technical Design Report - Volume 4: Detectors.
2013.

[7] V. Buridon et al.
First results of the CALICE SDHCAL technological prototype.
JINST, 11(04):P04001, 2016.

[8] J. S. Marshall and M. A. Thomson.
The Pandora Software Development Kit for Pattern Recognition.
Eur. Phys. J., C75(9):439, 2015.

[9] F. Gaede.
Marlin and LCCD: Software tools for the ILC.
Nucl. Instrum. Meth., A559:177–180, 2006.

References III

[10] Markus Frank, F. Gaede, C. Grefe, and P. Mato.

DD4hep: A Detector Description Toolkit for High Energy Physics Experiments.

J. Phys. Conf. Ser., 513:022010, 2014.

[11] Ryan R. Curtin, Marcus Edel, Mikhail Lozhnikov, Yannis Mentekidis, Sumedh Ghaisas, and Shangtong Zhang.

mlpack 3: a fast, flexible machine learning library.

Journal of Open Source Software, 3:726, 2018.