Advanced Particle Flow

- **Development of advanced particle flow and pattern recognition algorithms in PandoraPFA**
- **Application to LHC, LC and neutrino experiments**

S.Green, J.S.Marshall,
M. Weber, A. Sailer, O. Viazlo
M. Ruan, D. Yu, K. Shpak, **V. Boudry**

Advanced Particle Flow

The AIDA 2020 Advanced Particle Flow group consists of the following institutions:

1) Cambridge - Pandora pattern recognition development for liquid argon time projection chamber (LArTPC) experiments such as MicroBooNE, SBND, ProtoDUNE, DUNE and support for Pandora pattern recognition at the linear collider experiments.

2) CERN - Developing and improving the Pandora linear collider pattern recognition aimed at CLIC.

3) LLR - Development of ArborPFA for the e^+e^- collider experiments.

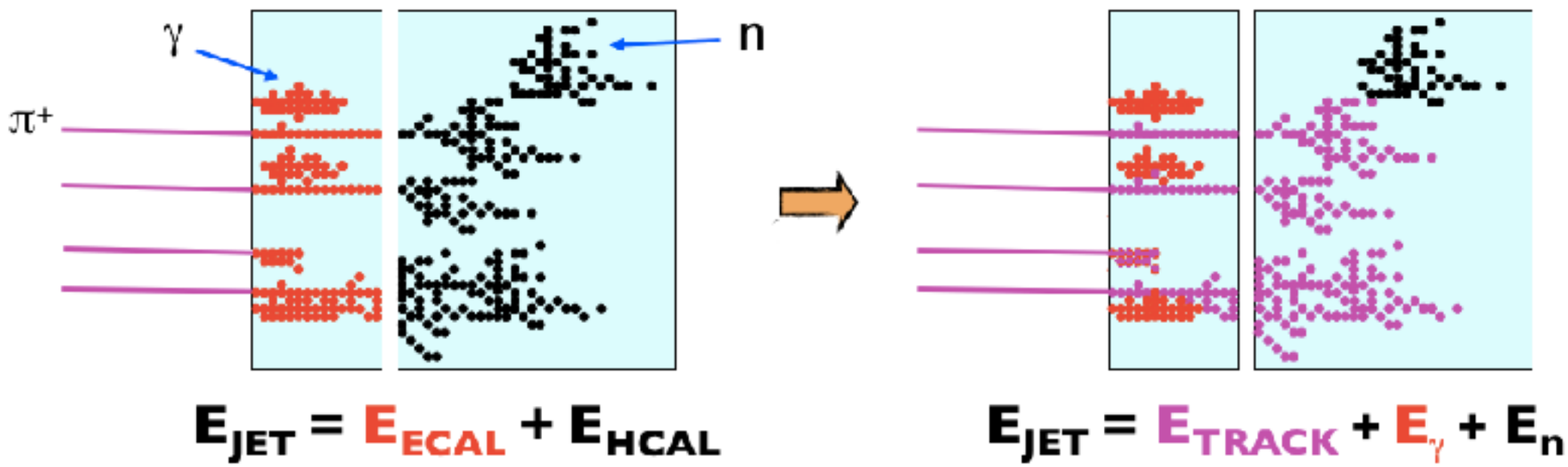
UNIVERSITY OF CAMBRIDGE

Institute of High Energy Physics
Chinese Academy of Sciences

PFA

Aim: separate hadronic decay of bosons

- Principle: reconstructing all the final state particles - different sub-detectors suitable for different kind of particles.
  - Final physics objects recognised with high efficiency and purity
  - Jets: 63% charged + 27%photon + 10% neutral hadron
  - Key performance: Jet Energy Resolution (JER) + ParticleID (Tau's, BSM)



Cambridge: MicroBooNE

A paper describing the Pandora LArTPC reconstruction was published in January this year.

The European Physical Journal C
January 2018, 78:82 | Cite as

The Pandora multi-algorithm approach to automated pattern recognition of cosmic-ray muon and neutrino events in the MicroBooNE detector

Authors | Authors and affiliations

R. Acciarri, C. Adams, R. An, J. Anthony, J. Asaadi, M. Auger, L. Bagby, S. Balasubramanian, B. Baller, C. Barnes, G. Barr, M. Bass, F. Bay, M. Bishai, A. Blake, show 132 more

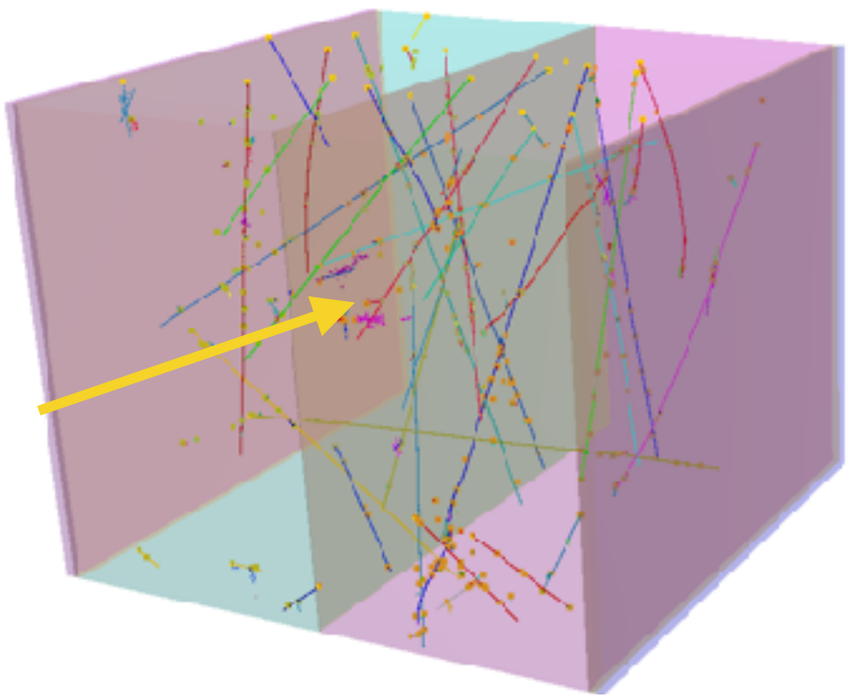
Open Access | Regular Article - Experimental Physics
First Online: 29 January 2018

Abstract

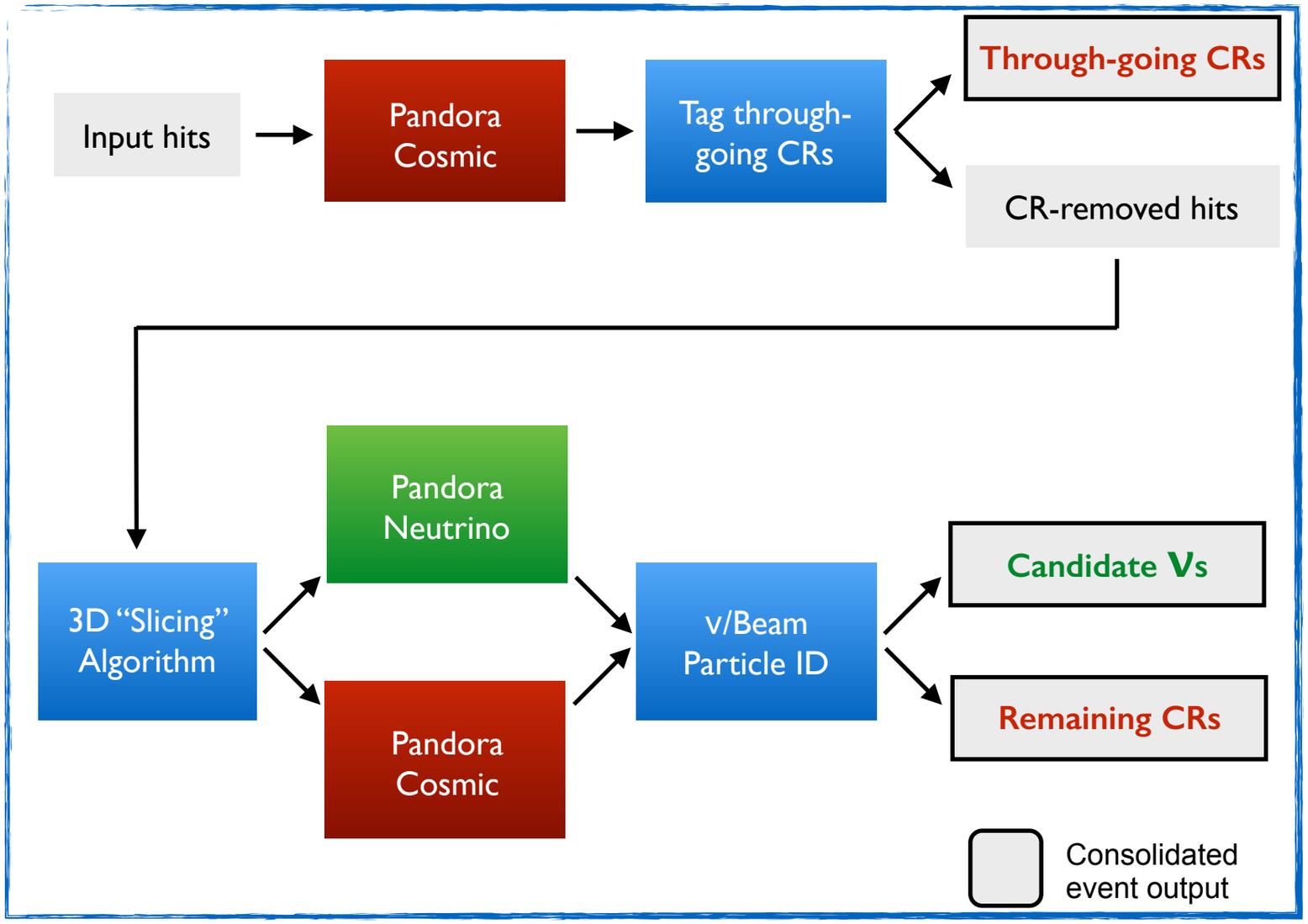
The development and operation of liquid-argon time-projection chambers for neutrino physics has created a need for new approaches to pattern recognition in order to fully exploit the imaging capabilities offered by this technology. Whereas the human brain can excel at identifying features in the recorded events, it is a significant challenge to develop an automated,

Cambridge: LArTPC

- A consolidated event reconstruction has been developed for use in LArTPC experiments, which simplifies the use of Pandora outputs in downstream analyses.
- The output is one reconstructed beam particle and N reconstructed cosmic-ray muons.
- This is a generic approach that can be applied for all LArTPC detectors (MicroBooNE, ProtoDUNE, DUNEFD, etc.)!

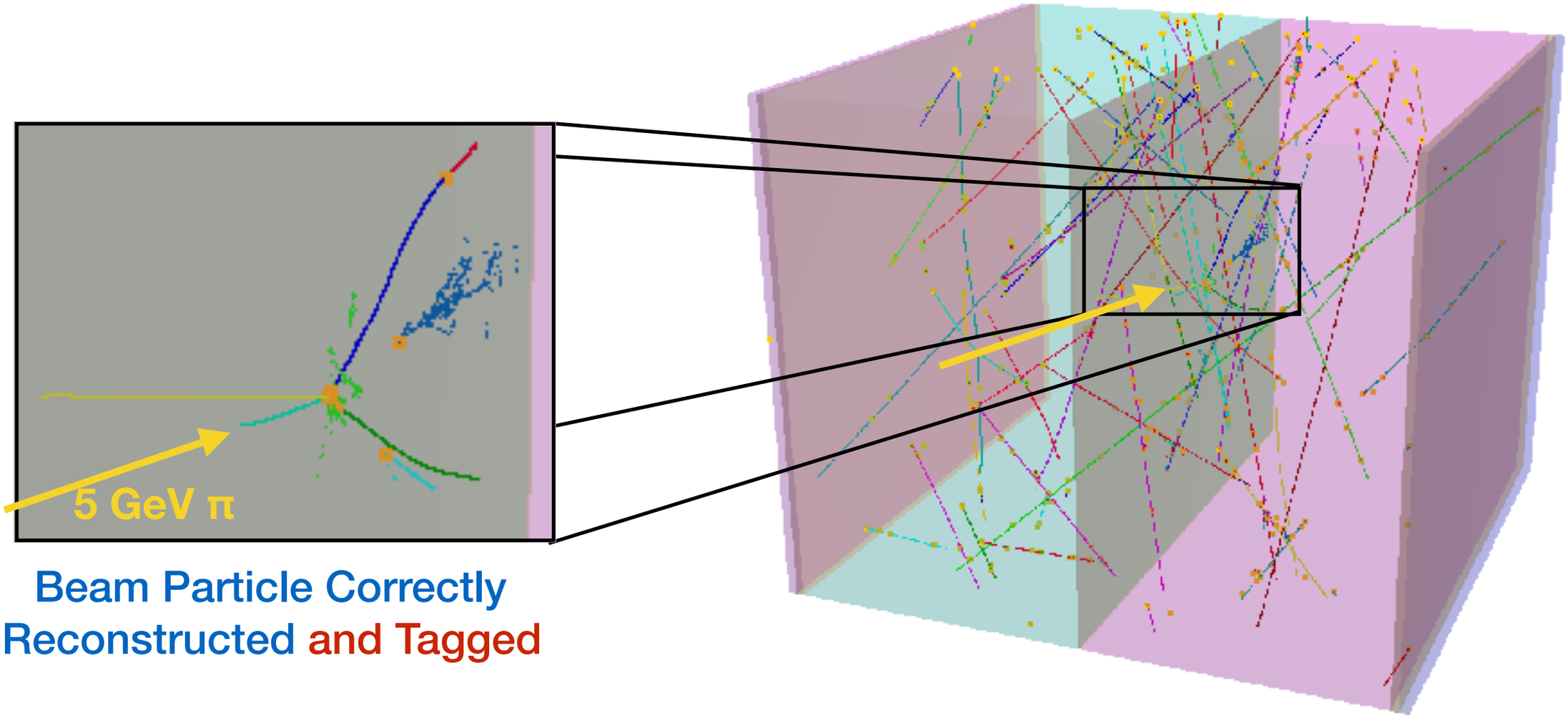


ProtoDUNE 7 GeV Beam + Cosmic Simu & Reconstructed Event



Cambridge: LArTPC

- A cut based beam particle ID tool has developed for use at ProtoDUNE.
- This made it into the MCC10 sample production and we are now analysing both the high (Particle Level) and low level reconstruction (Cluster Level) of these samples to further refine the LArTPC reconstruction.



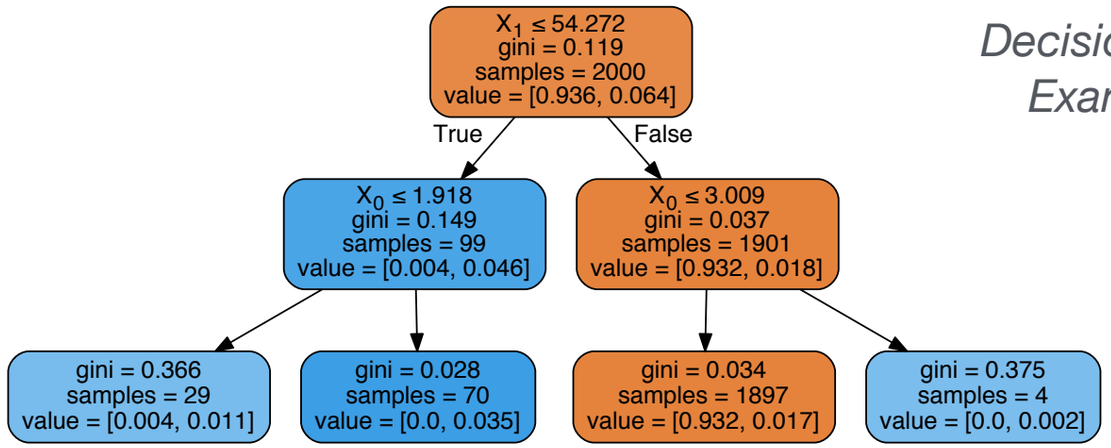
Beam Particle Correctly Reconstructed and Tagged

ProtoDUNE 5 GeV Beam + Cosmic Reconstructed Events

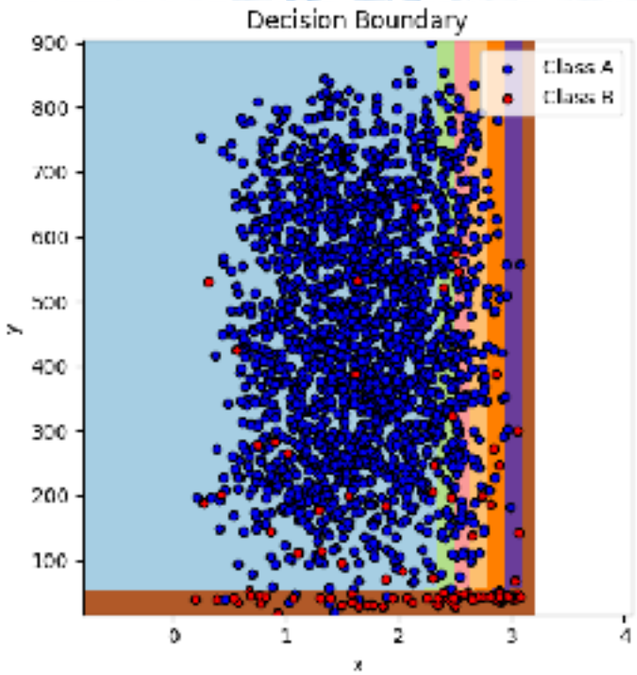
All Remaining Particles Correctly Reconstructed and Tagged as Cosmic-Rays

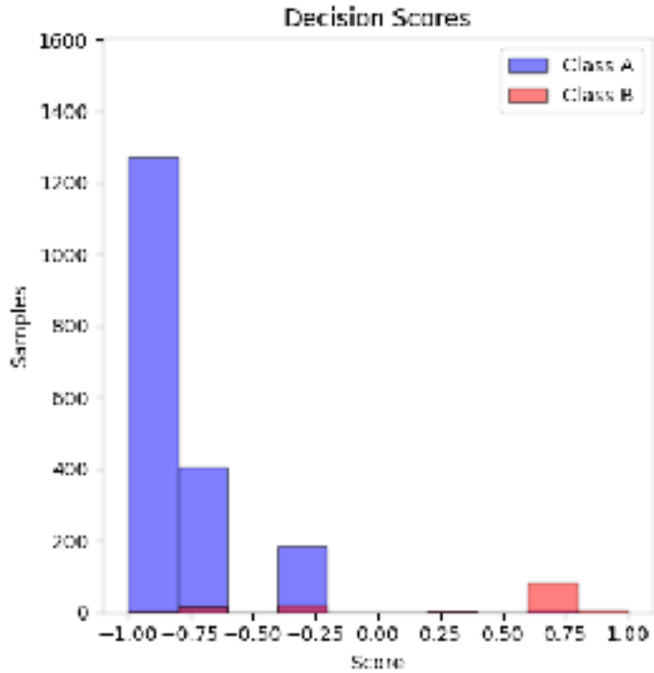
Cambridge: LArTPC

- We are aiming to extend the use of Multivariate techniques inside Pandora.
- Method is to make the application of machine learning techniques more generic, which will allow us to use more varied machine learning techniques (alongside Likelihood and Support Vector Machines already in use).
- Machine learning models will be trained using the scikit-learn python package and used in Pandora.
- The immediate physics goal driving this work is to improve the Beam Particle ID for use at ProtoDUNE by replacing the cut based approach with a BDT.



Decision Tree Example

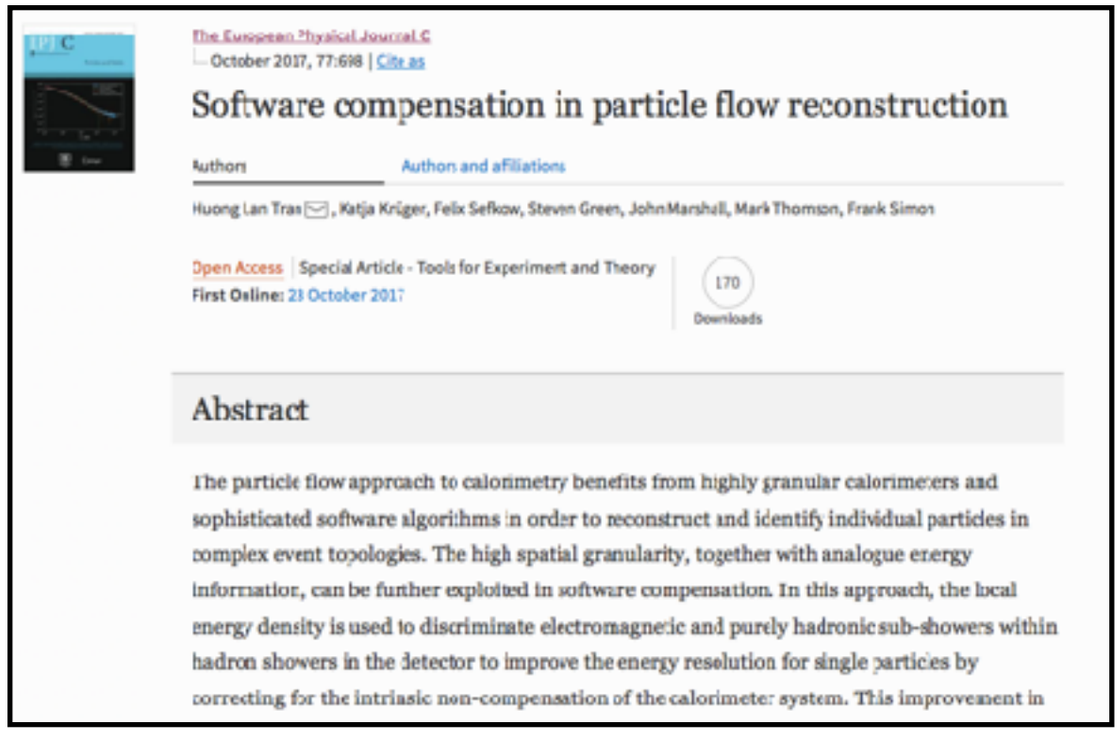




Distributions of classification score for signal (Class A, background cosmic rays and Class B, target beam particle.

Cambridge/DESY/MPPI: Linear Collider

- There is continued exploitation of the Pandora pattern recognition at the linear collider.
- A paper describing a novel software compensation technique applied inside Pandora was published in October 2017.
- This technique compensates for the invisible energy component found in hadronic showers by reweighing the energy of calorimeter hits based on their energy density.
- Many thanks to F. Simon and H.L.Tran for their efforts getting this to publication!



The European Physical Journal C
October 2017, 77:698 | Cite as

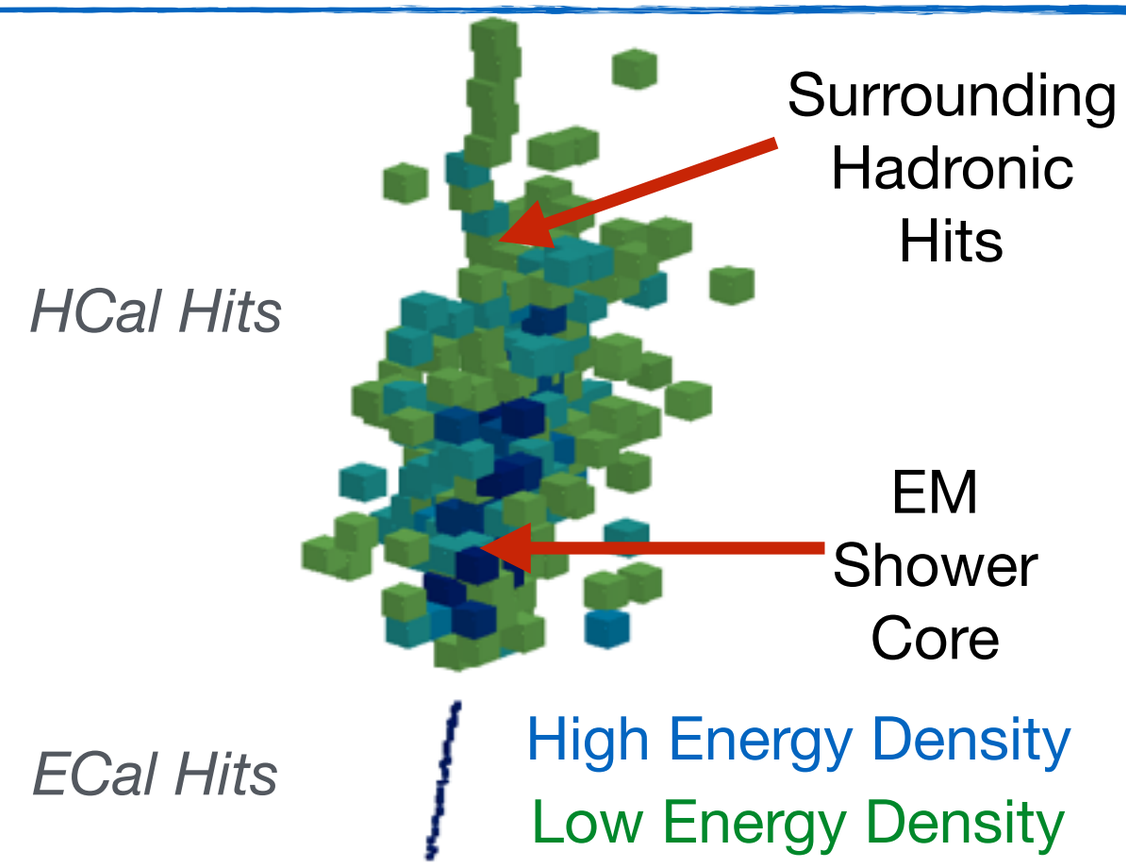
Software compensation in particle flow reconstruction

Authors Author and affiliations
Huong Lan Tran, Katja Krüger, Felix Sefkow, Steven Green, John Marshall, Mark Thomson, Frank Simon

Open Access | Special Article - Tools for Experiment and Theory
First Online: 23 October 2017
170 Downloads

Abstract

The particle flow approach to calorimetry benefits from highly granular calorimeters and sophisticated software algorithms in order to reconstruct and identify individual particles in complex event topologies. The high spatial granularity, together with analogue energy information, can be further exploited in software compensation. In this approach, the local energy density is used to discriminate electromagnetic and purely hadronic sub-showers within hadron showers in the detector to improve the energy resolution for single particles by correcting for the intrinsic non-compensation of the calorimeter system. This improvement in

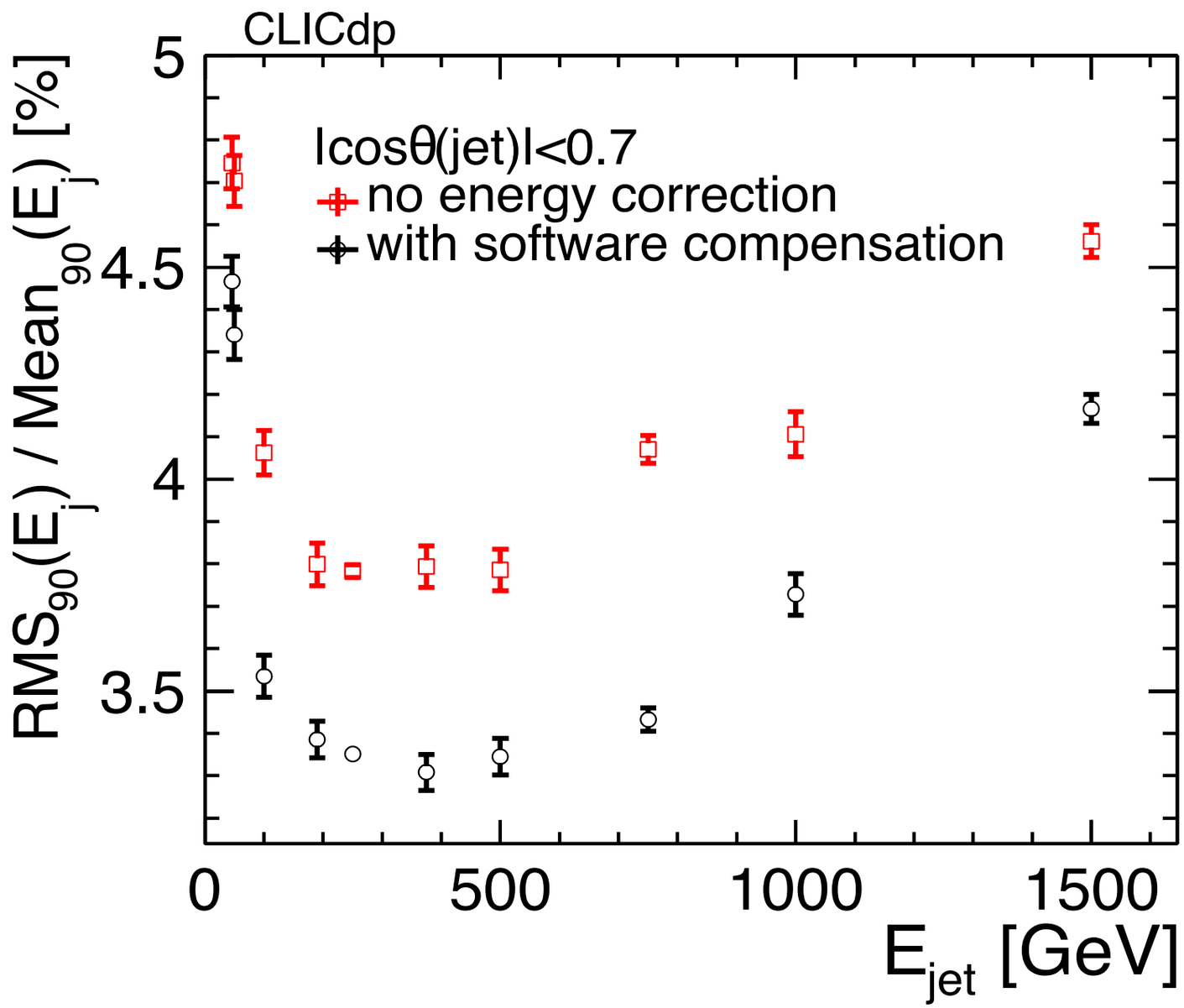


PandoraPFA at CLIC and FCCee

CERN LCD group

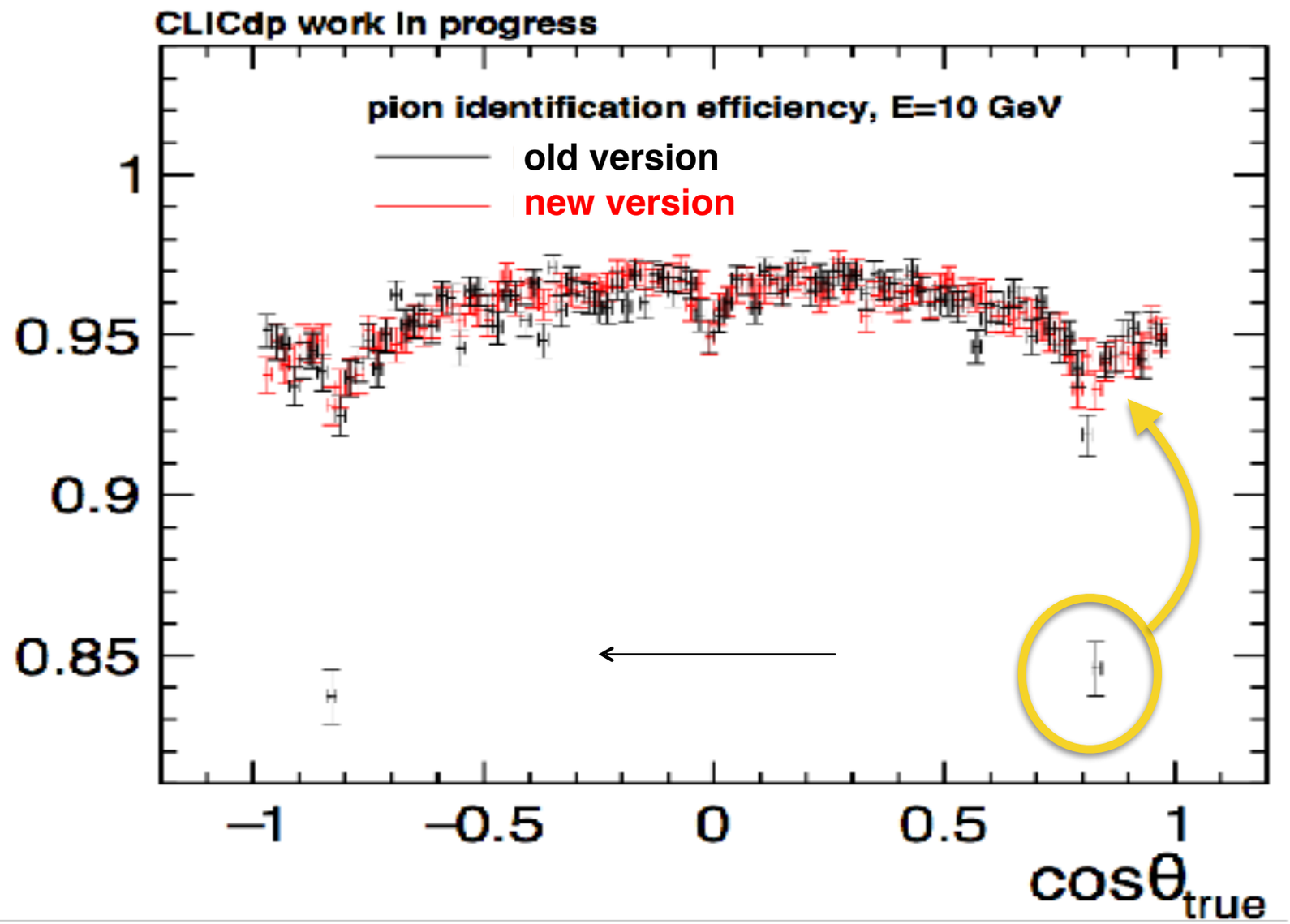
Software Compensation at CLIC

Software compensation weight calibration integrated within Pandora
Weights calibrated using single neutrons and K^0_L's over a large energy range
Software compensation improves the energy response and resolution of hadronic clusters → on average jet energy resolution improves by around 10 %



Improvement in track-cluster matching

Track and cluster matched if first cluster hits spatially close to impact point of track on calorimeter → in **previous** Pandora versions in **barrel-endcap transition region** matching failed for late starting calorimeter showers leading to **inefficiency**



Modification of track-cluster match calculation → **recover performance** in **new version**

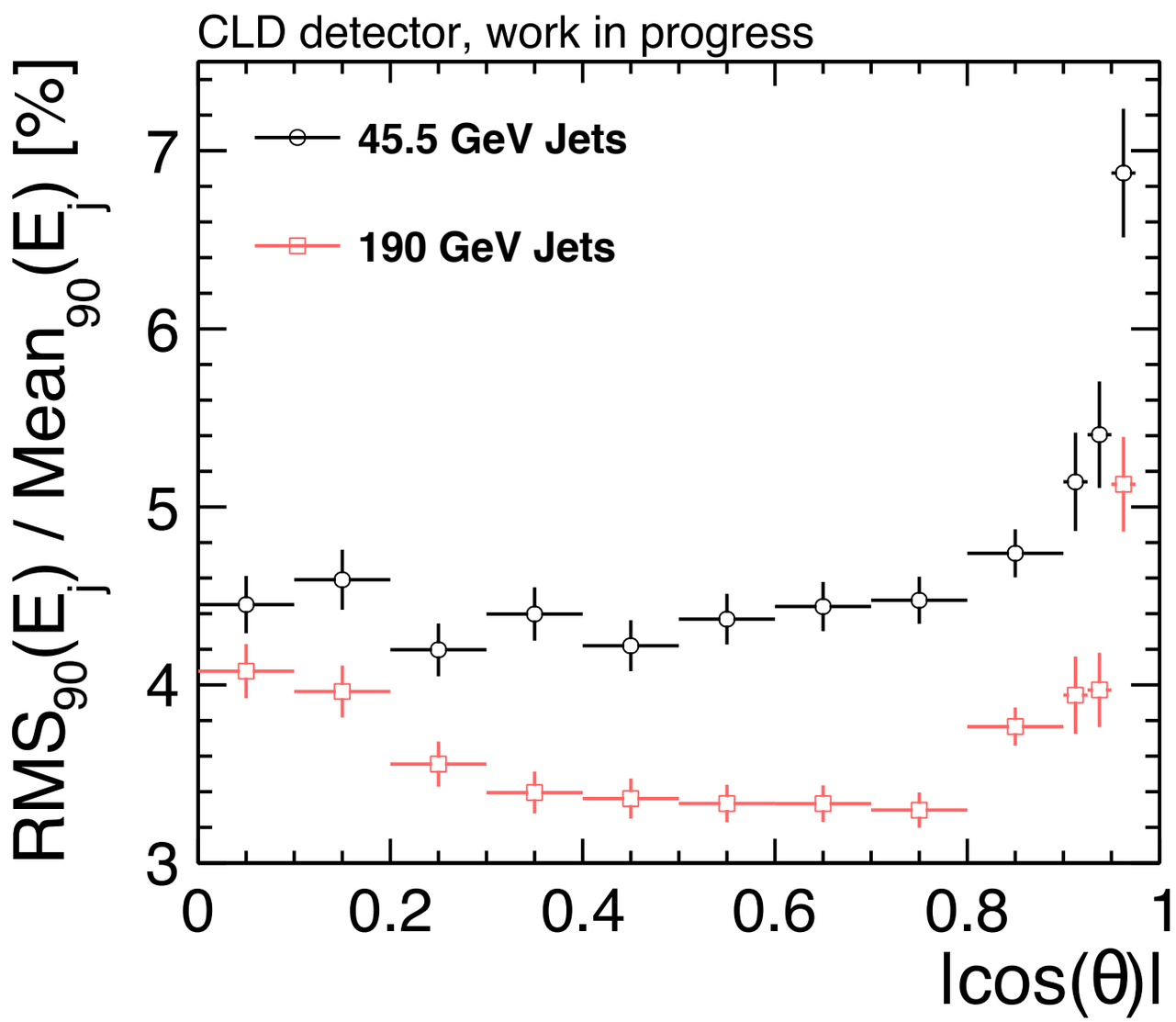
PandoraPFA at FCCee

Pandora pattern recognition used in the CERN Future Circular Collider (FCC) Study in the context of the electron-positron option

- The main difference between the **CLD** detector and the **CLIC detector** concept are the larger calorimeter radius and the smaller magnetic field

The performance of pandora pattern recognition performs very similar at CLIC and FCC

Jet energy resolutions between 3 and 5 % are reached in both cases



Advanced Particle Flow

Development of ArborPFA e+e- collider experiments.
LLR & IHEP (associated to AIDA)

1. Usage of ARBOR for Detector optimisation
    1. CEPC CDR
    2. Performances in separation:
2. Lepton ID : LICH

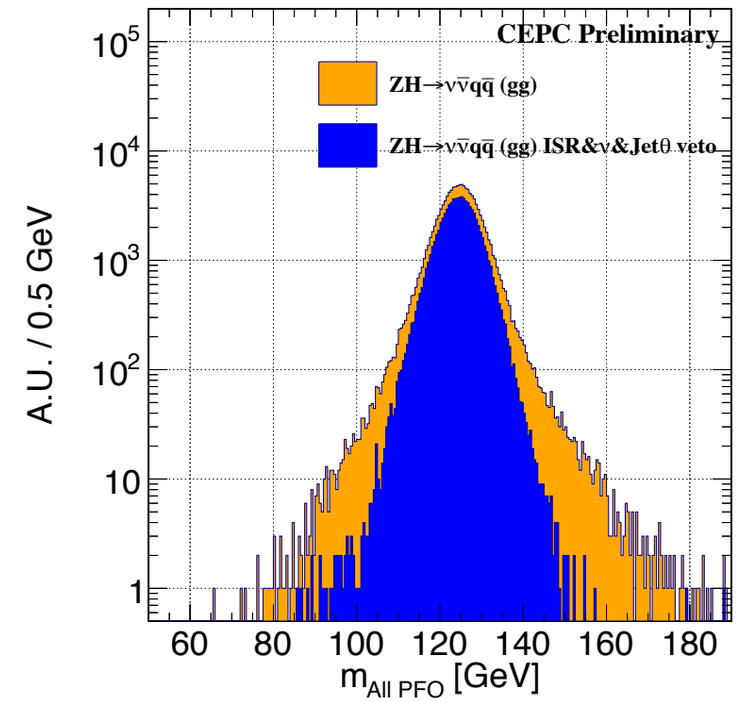
Results of PhD defences by
Dan Yu (IHEP & LLR, Feb. 2018),
K. Shpak (LLR, Jan. 2018)

Institute of High Energy Physics
Chinese Academy of Sciences

Performances estimation

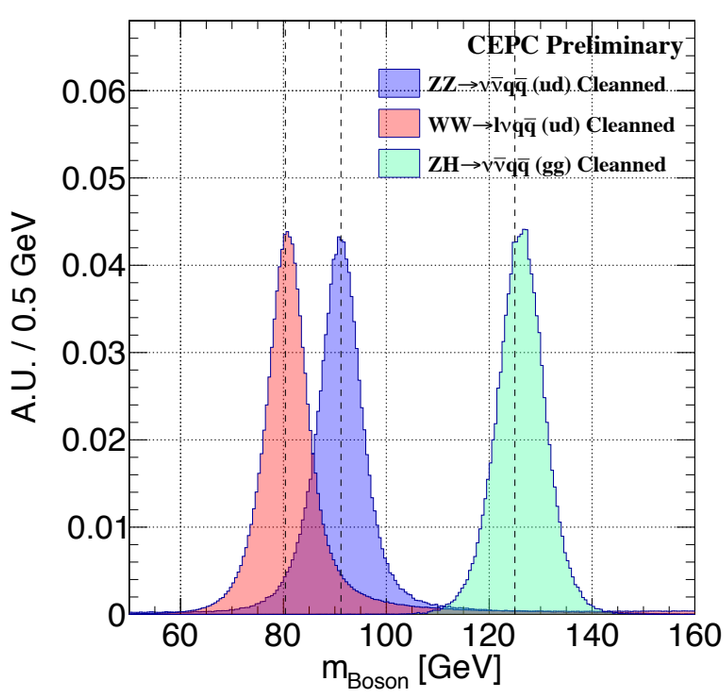
Boson Mass Resolution (BMR)

- Higgs Boson Mass Resolution $\nu\nu gg$ channel @ 250 GeV
  - Focus on influence from algorithm/detector
  - Rejection: ISR photons (Pt(ISR)< 1GeV), neutrinos from Higgs(Pt(ν)< 1GeV), jets shooting to the endcaps (|cos θ| < 0.85)
- Efficiency ~ 65%

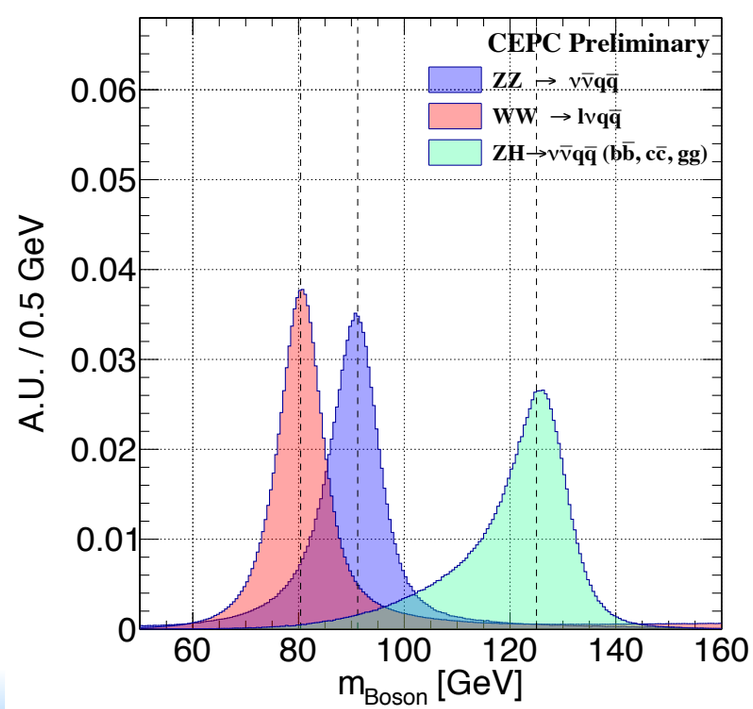


Di-jet W, Z and H masses:

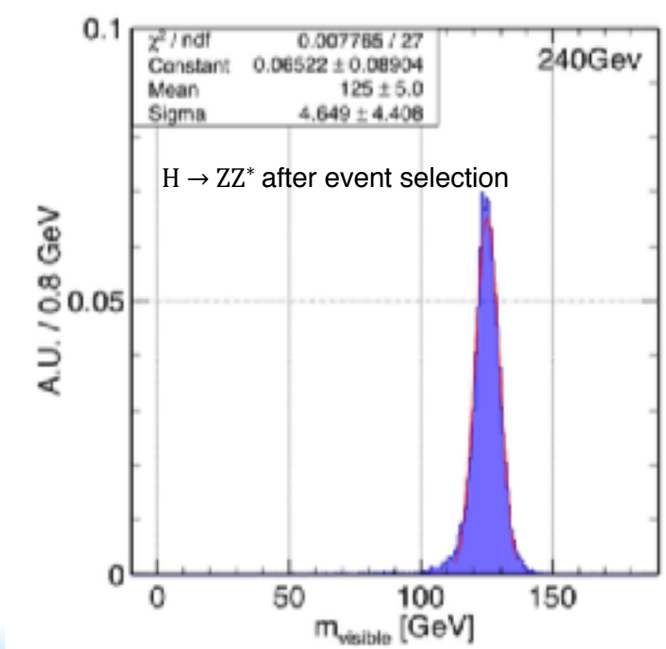
vv/ud final state (cleaned)



inclusive qq final state (separable)

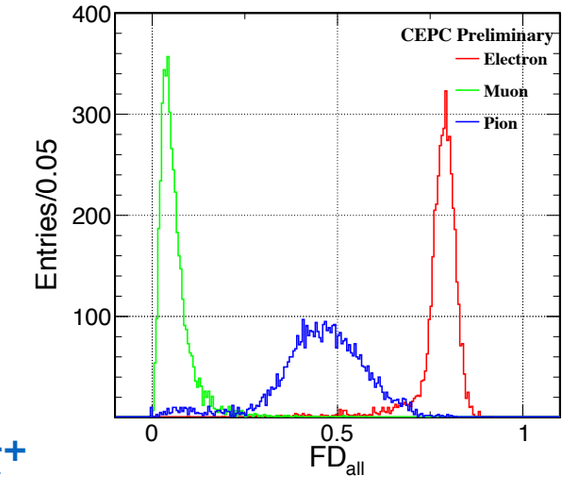


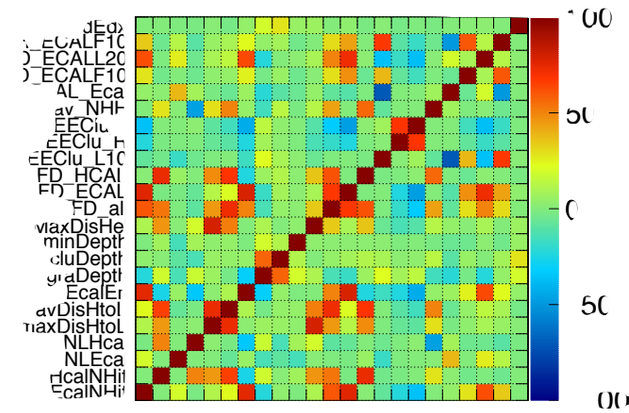
Validated @ 240 GeV on $\nu\nu H$, $\nu\nu H$, $H \to bb, cc, gg, \mu\mu, \gamma\gamma, ZZ*andWW*$



LICH (Lepton Identification for Calorimeter with High granularity)

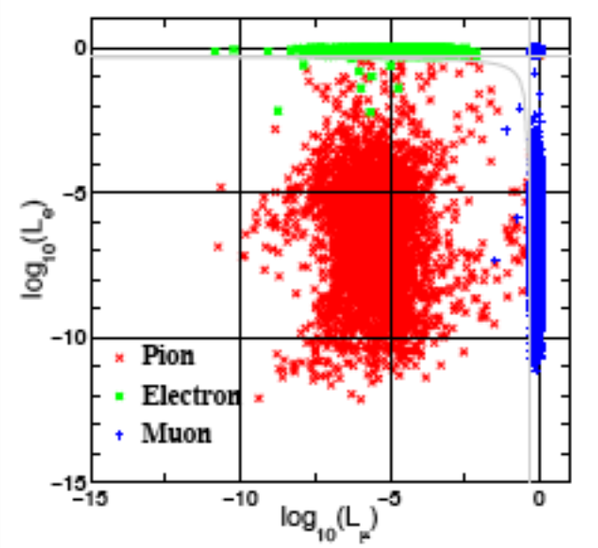
- Input: 24 variables from reconstructed charged particle
  - incl. dE/dx, Fractal dimension, shower shapes and position, …
- TMVA on samples: Single particle: e^-, μ^-,π^+ (1 GeV – 120 GeV)
- Output: likelihood



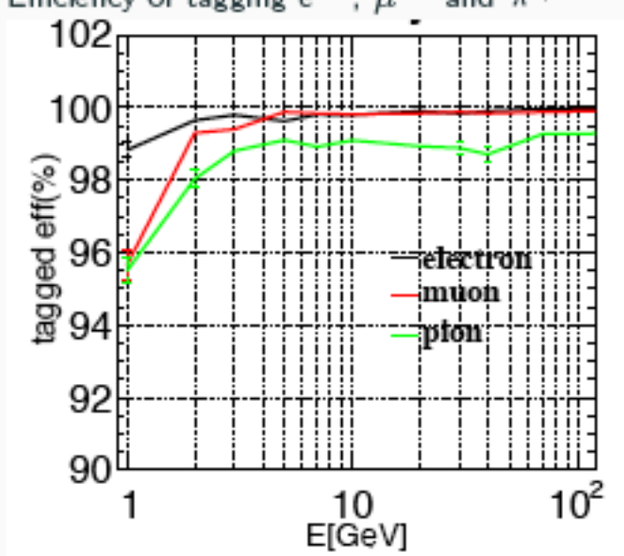


Integrated to ILCSOFT & applied for preparation of CEPC CDR
Phys. J. C77 no. 9, (2017) 591, arXiv:1701.07542 [physics.ins-det].

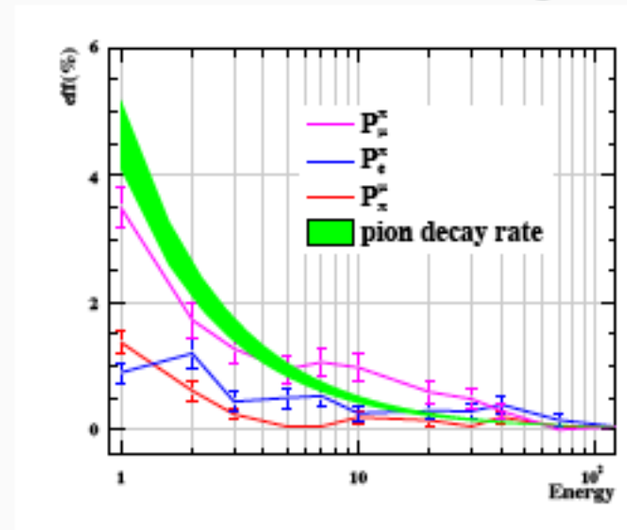
μ-e likeness for e^-, μ^- and π^+ at 40GeV



Efficiency of tagging e^-, μ^- and π^+



misidentification rate at different energies



Migration Matrix at 40GeV (LICH)

Type	e^- *like*	μ^- *like*	π^+ *like*
e^-	99.71 ± 0.08	< 0.07	0.21 ± 0.07
μ^-	< 0.07	99.87 ± 0.08	0.05 ± 0.05
π^+	0.14 ± 0.05	0.35 ± 0.08	99.26 ± 0.12

Migration Matrix for ALEPH PID ($>$ 2GeV)(*Eur.Phys.J.C20:401-430,2001*)

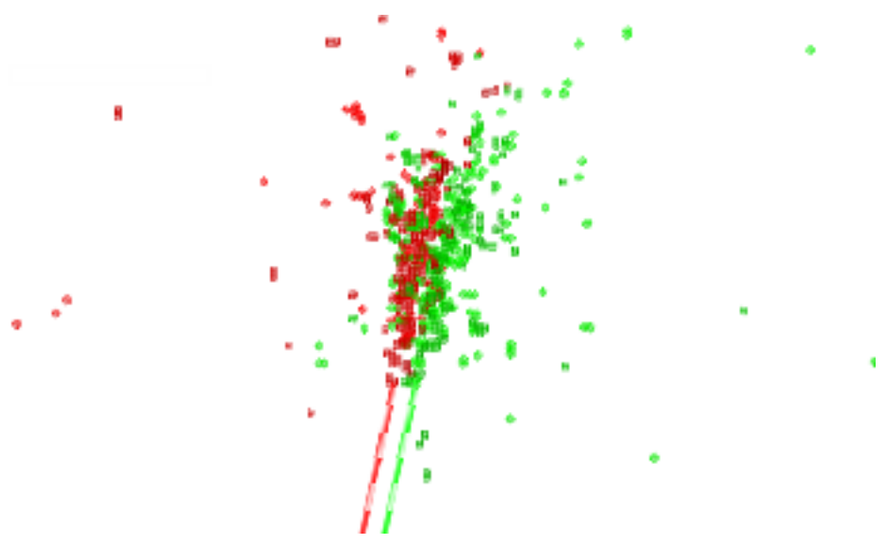
Type	e^- *like*	μ^- *like*	π^+ *like*	undefined
e^-	99.57 ± 0.07	< 0.01	0.32 ± 0.0	0.09 ± 0.04
μ^-	< 0.01	99.11 ± 0.08	0.88 ± 0.08	0.01 ± 0.01
π^+	0.71 ± 0.04	0.72 ± 0.04	98.45 ± 0.06	0.12 ± 0.03

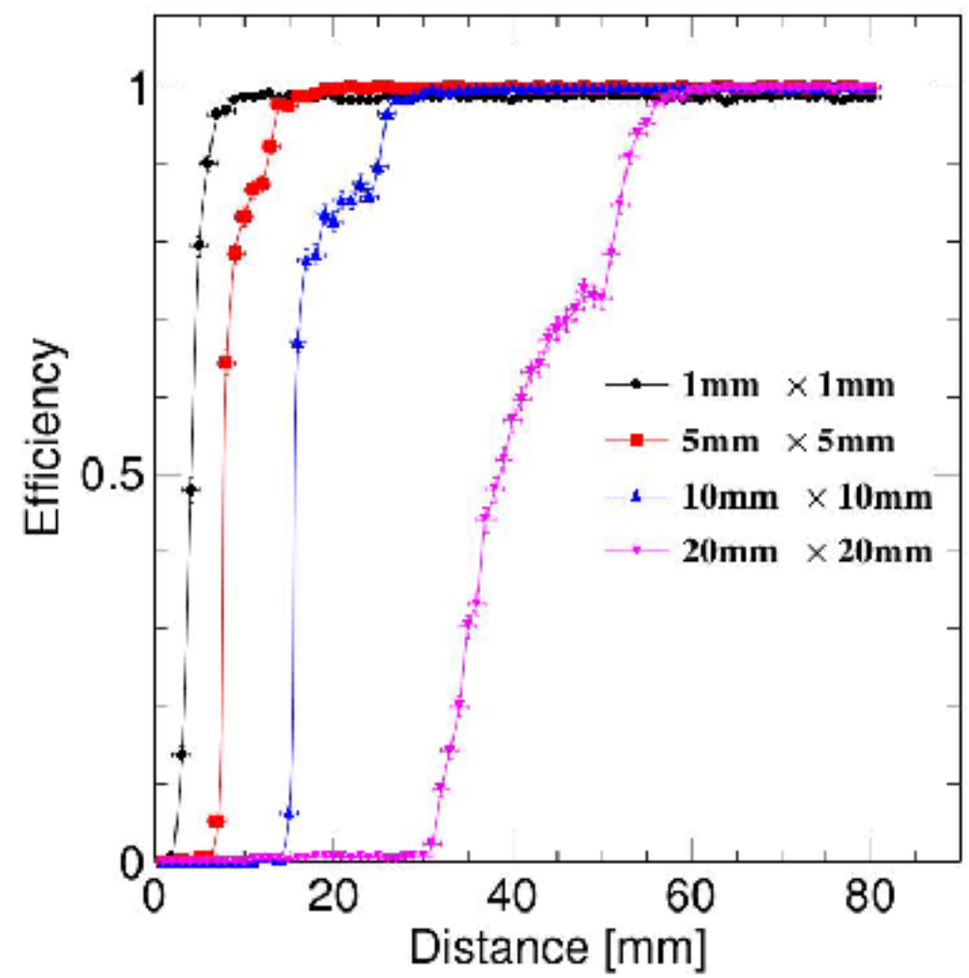
Separation Studies

For detector optimisation

Sample: di-photon system

Efficiency: both the reconstructed photon has more than 1/3 and less than 2/3 of all the deposit energy

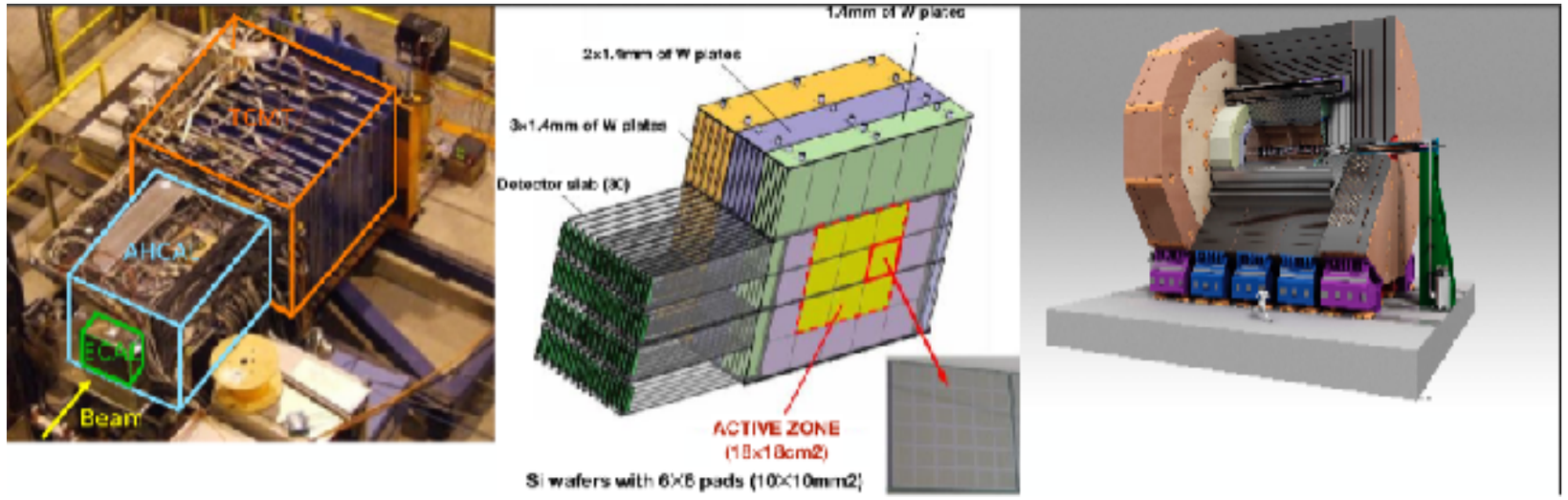


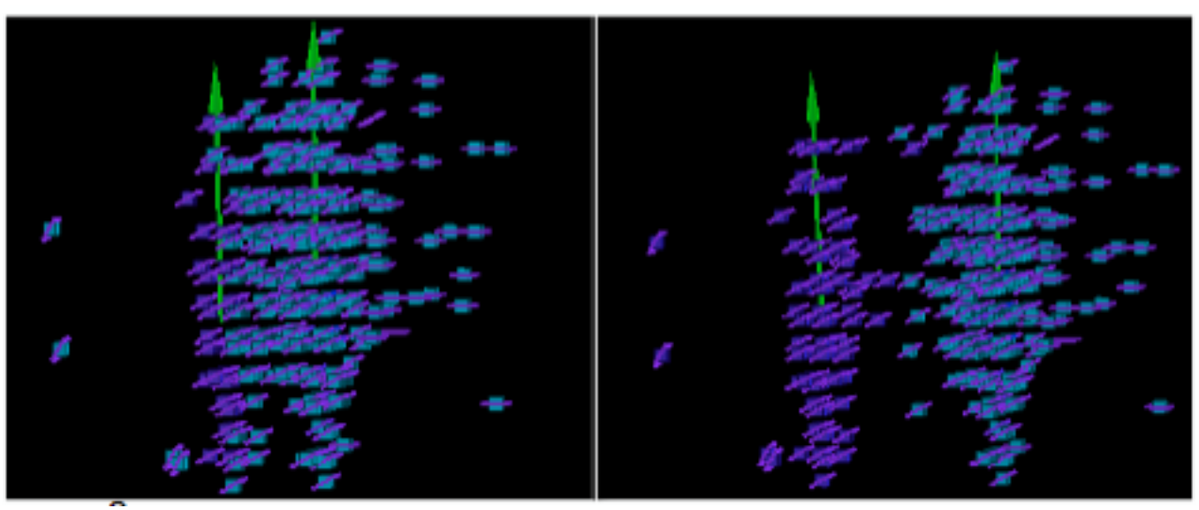


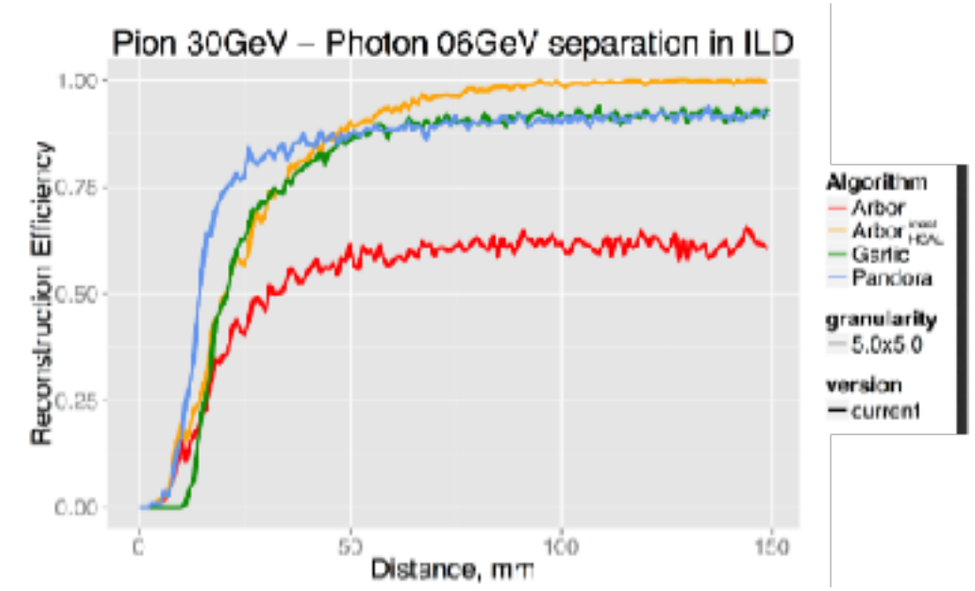
Separation studies (K. Shpak)

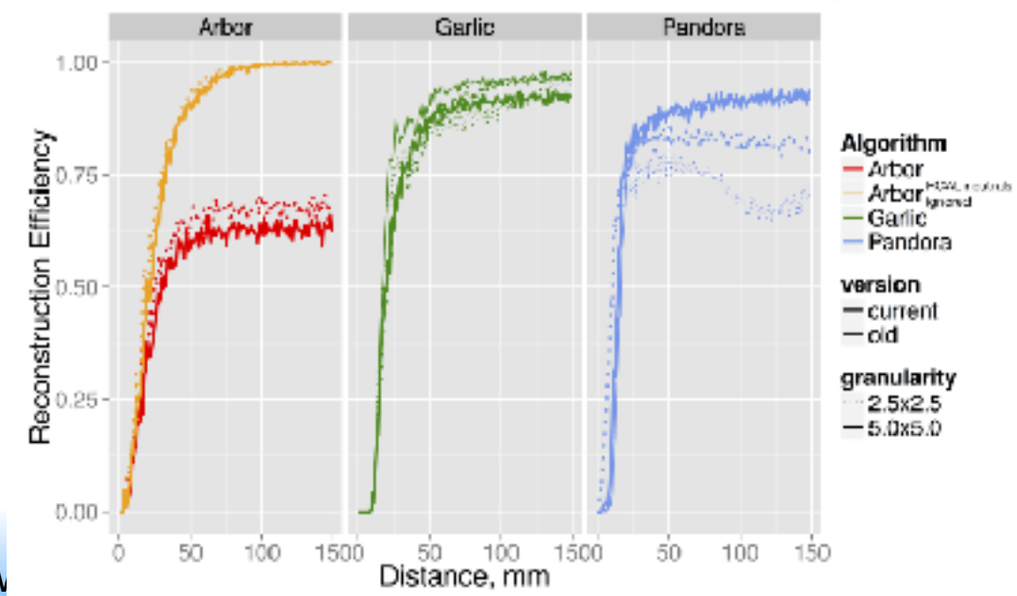
- "raw performances"
  - Efficiency vs separation distance
  - EM vs EM (e / γ)
  - EM vs π
- PandoraPFA, Arbor, GARLIC
  - Beam test data vs ILD simulation

CAN-057: (Calice Analysis Note)









Conclusions

- PFA algorithms are improving continuously
- Development of Advanced Particle Flow software has resulted in a number of publications over the past year.
- Larger application ranges:
  - ℓAr + all e^+e^- colliders (ILC, CLIC, FCCee & CEPC)
  - LHC experiments still facing computing challenge
- Pandora is now setup to exploit the full benefits of multivariate techniques in PFA
- and automatic calibration (see work from R. Été)

ProtoDune Geometry