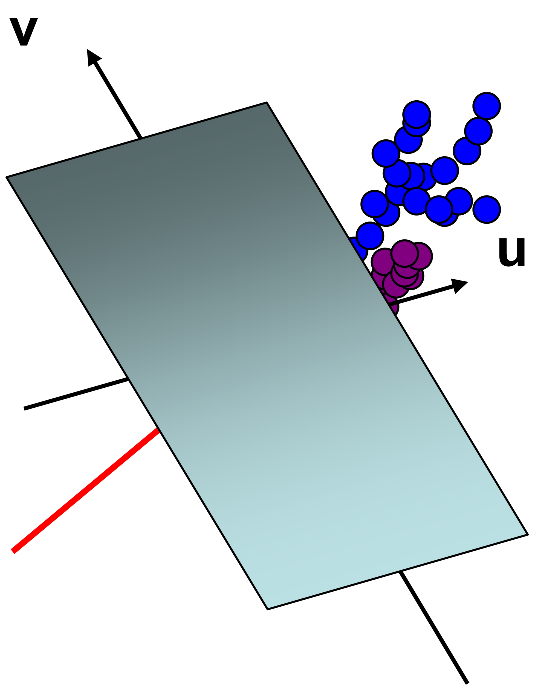
Particle Flow : QED ?

Mark Thomson
University of Cambridge



This Talk:

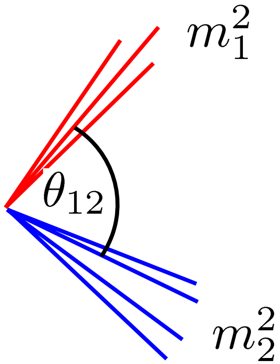
❶ Introductory comments
❷ PandoraPFA: status
❸ PandoraPFA: the algorithm
❹ Performance (old vs. new)
❺ Detector studies
❻ PerfectPFA
❼ Deficiencies, random comments, outlook, and conclusions

❶ Introductory Comments

What is the ILC goal for PFA ?

★ **For a pair of jets have:**

$$m^2 = m_1^2 + m_2^2 + 2E_1E_2(1 - \beta_1\beta_2\cos\theta_{12})$$



★ **Assuming a single jet energy resolution of**

$$\sigma_E/E = \alpha(E)/\sqrt{E(\mathrm{GeV})}$$

⇨ $$\sigma_m/m \approx \alpha(E_j)/\sqrt{E_{jj}(\mathrm{GeV})}$$

+ term due to θ_{12} uncertainty

★ **For a Gauge boson mass resolution of order** $\Gamma_{\mathrm{W/Z}}$

$$\frac{\sigma_m}{m} \approx \frac{2.5}{91.2} \approx \frac{2.1}{80.3} \approx 0.027$$

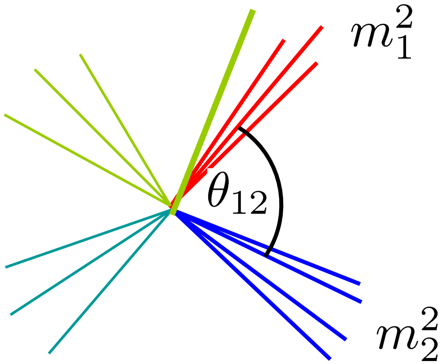
⇨ $$\alpha(E_j) < 0.027\sqrt{E_{jj}(\mathrm{GeV})}$$

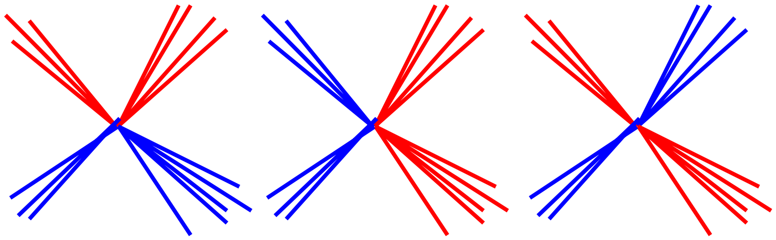
E_{jj}/GeV	$\alpha(E_j)$
91	< 26 %
200	< 38 %
360	< 51 %
500	< 60 %

★ Don't take exact numbers too seriously
★ Just note that the jet energy resolution requirements depend on energy

Other effects:

- ★ jet angle measurement
- ★ jet finding
  - • not perfect – will degrade mass resolution
- ★ jet pairing
  - • multiple jet-pair combinations





Relative importance of these effects ?

❷ PandoraPFA Status

★ Last results reported at Valencia meeting

★ Much recent progress:

- ValGrind – highlighted a few other minor issues
- Reclustering significantly improved +
  fixed a couple of "features"
- Added new "photonRecovery" code
- Added new "fragmentRemoval" code
- Version finalised last week and now in CVS repository (v01-01)

➡ **Much improved performance…**

The results presented here use the CVS version of PandoraPFA

❸ PandoraPFA: the Algorithm

- ★ ECAL/HCAL reconstruction and PFA performed in a single algorithm
- ★ Keep things fairly generic algorithm
  - ▪ applicable to multiple detector concepts
- ★ Use tracking information to help ECAL/HCAL clustering

- ★ This is a fairly sophisticated algorithm : ~9000 lines of code

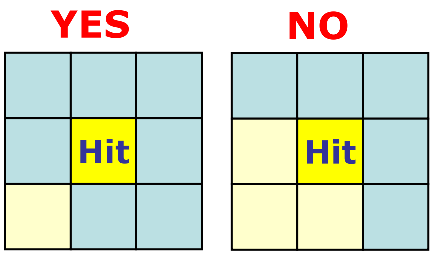
Eight Main Stages (was six):

i. Preparation
ii. Loose clustering in ECAL and HCAL
iii. Topological linking of clearly associated clusters
iv. Courser grouping of clusters
v. Iterative reclustering
vi. Photon Recovery (NEW)
vii. Fragment Removal (NEW)
viii. Formation of final Particle Flow Objects (reconstructed particles)

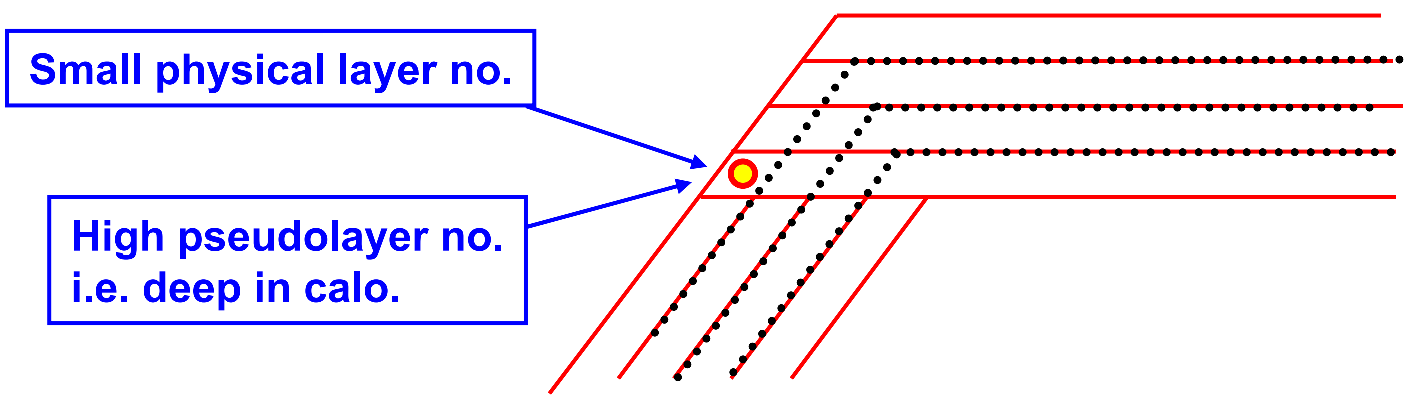
Order inter-changable (v and vi)

Preparation I: Extended Hits

- Create internal **ExtendedCaloHits** from **CaloHits**
- **ExtendedCaloHits** contain extra info:
  - pointer to original hit
  - pseudoLayer (see below)
  - measure of isolation for other hits
  - is it MIP like
  - actual layer (decoded from **CellID**)
  - Pixel Size (from GEAR) – **hits are now self describing**

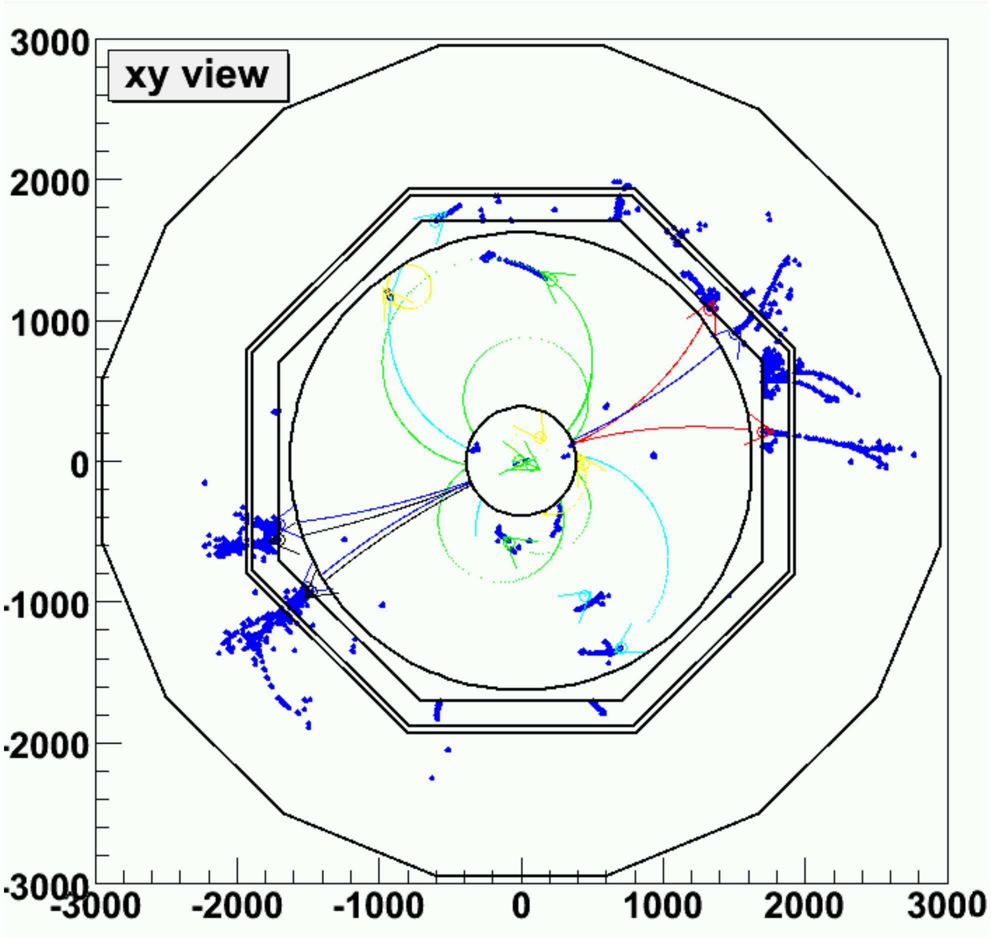


- Arrange hits into **PSEUDOLAYERS**
  - **i.e.** order hits in increasing depth within calorimeter
  - PseudoLayers follow detector geometry



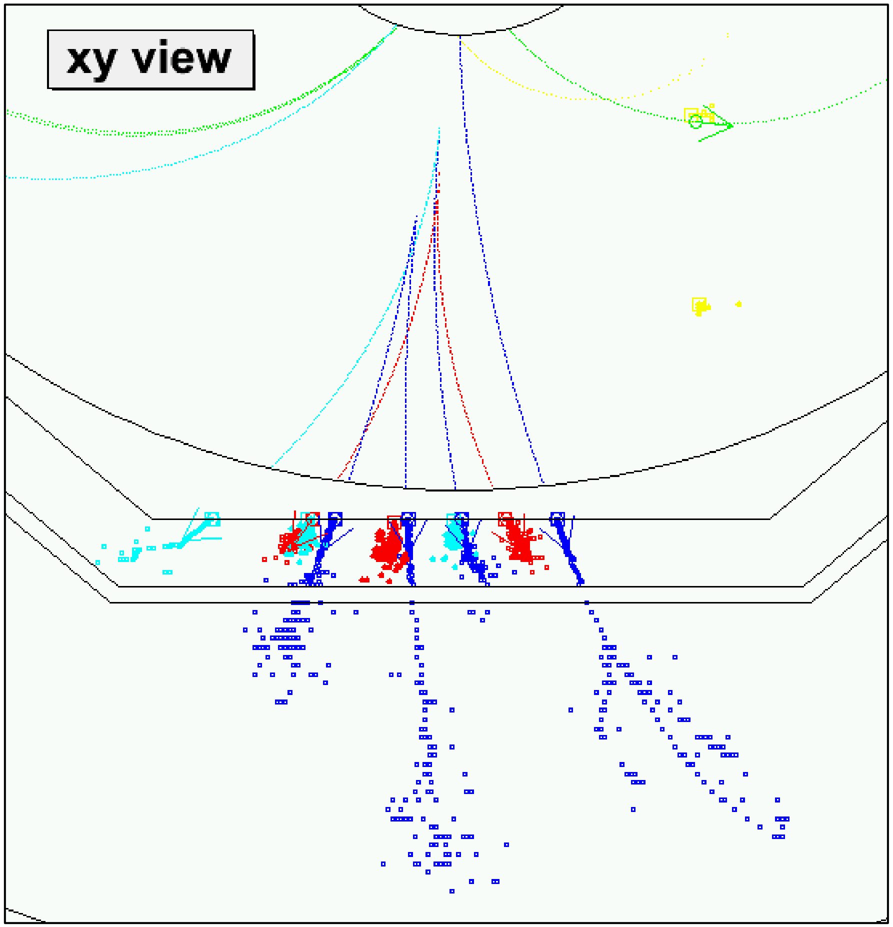
Preparation II: Isolation

- ★ Divide hits into isolated and non-isolated
- ★ Only cluster **non-isolated** hits
- ★ "Cleaner"/Faster clustering
- ★ Significant effect for scintillator HCAL



- ✦ Removal of isolated hits degrades HCAL resolution
- ✦ <u>e.g. LDC</u>
  50 %/√E/GeV → 60 %/√E/GeV

Preparation III: Tracking



- ★Use MARLIN TrackCheater
- ★Tracks formed from MC Hits in TPC/FTD/VTX
- ★ HelixFit (Alexei R) ⇒ track params
- ★ Cuts (primary tracks):
  - ♦ $|d_0| < 5$ mm
  - ♦ $|z_0| < 5$ mm
  - ♦ >4 non-Si hits

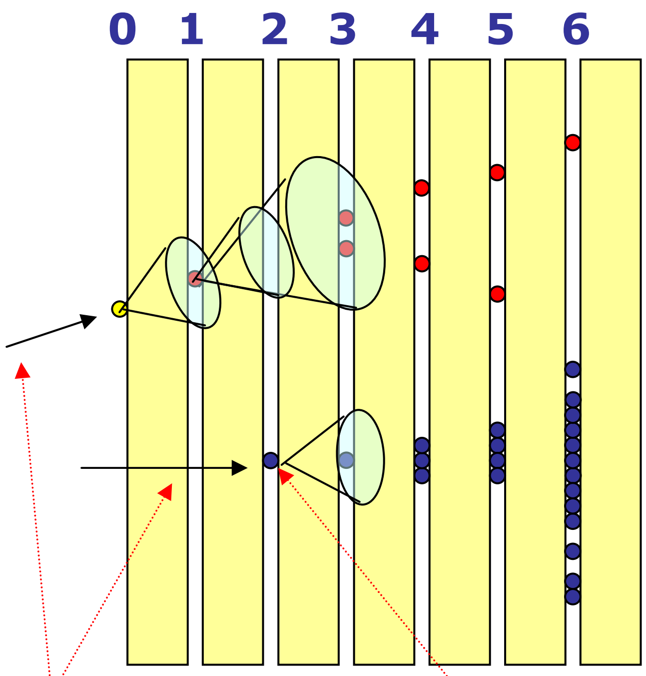
+ V_0 and Kink finding:

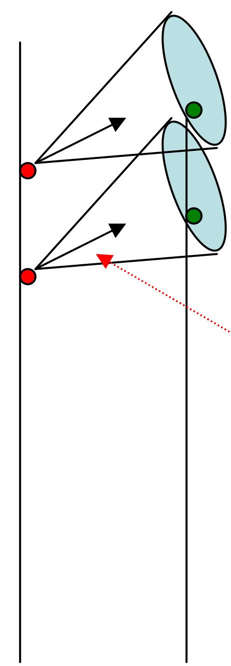
- ✦Track resolution better than cluster
- ✦ Improves PFA performance by ~2 %

Will soon move to fully reconstructed tracks (LDCTracking)

ii) ECAL/HCAL Clustering

- ★ Start at inner layers and work outward
- ★ Tracks can be used to "seed" clusters
- ★ Associate hits with existing Clusters
- ★ If no association made form new Cluster
- ★ Simple cone based algorithm





Simple cone algorithm based on current direction + additional N pixels

Cones based on either:
initial PC direction or
current PC direction

Initial cluster direction

Unmatched hits seeds new cluster

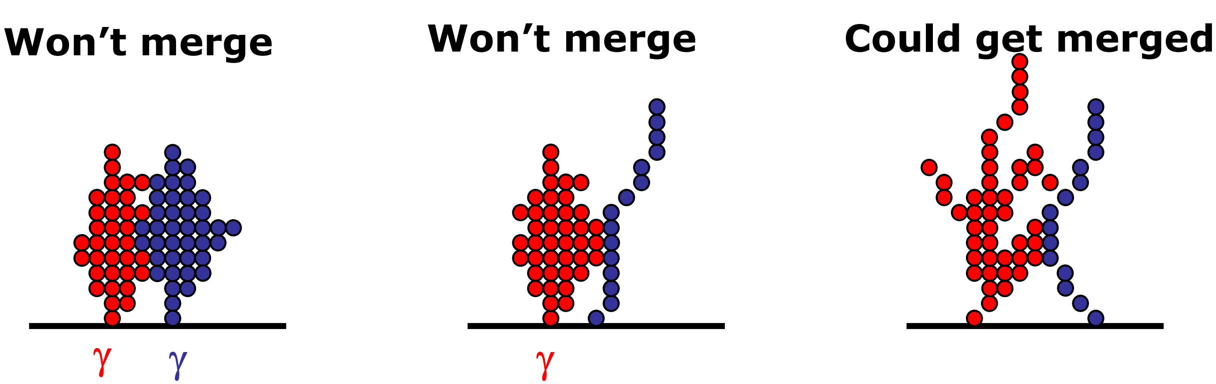
Parameters:
- cone angle
- additional pixels

iii) Topological Cluster Association

- By design, clustering errs on side of caution
  i.e. clusters tend to be split
- **Philosophy:** easier to put things together than split them up
- Clusters are then associated together in two stages:
  - 1) Tight cluster association – clear topologies
  - 2) Loose cluster association – fix what's been missed

Photon ID

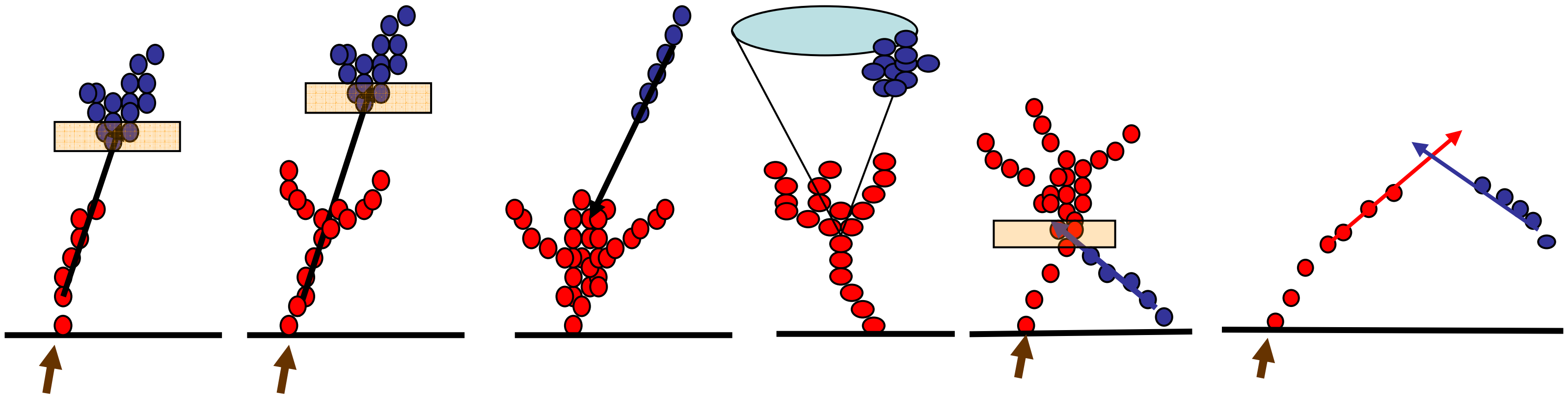
- Photon ID plays important role
- Simple "cut-based" photon ID applied to all clusters
- Clusters tagged as photons are immune from association procedure – just left alone



★ Clusters associated using a number of topological rules

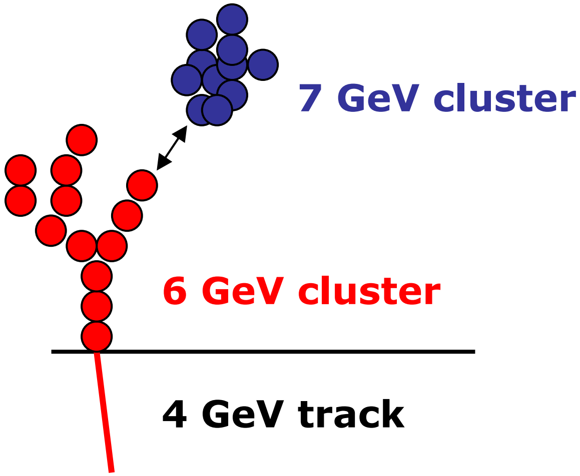
Clear Associations:

- Join clusters which are clearly associated making use of high granularity + tracking capability: very few mistakes



Less clear associations:

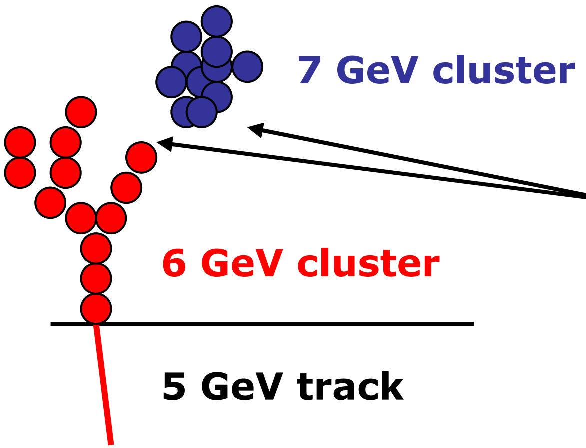
e.g. Proximity



Use E/p consistency to veto clear mistakes

iv) Cluster Association Part II

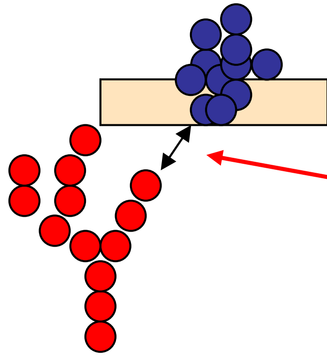
- Have made very clear cluster associations
- Now try "cruder" association strategies
- BUT first associate tracks to clusters (temporary association)
- Use track/cluster energies to "veto" associations, e.g.



This cluster association would be forbidden if $|E_1 + E_2 - p| > 3\,\sigma_E$

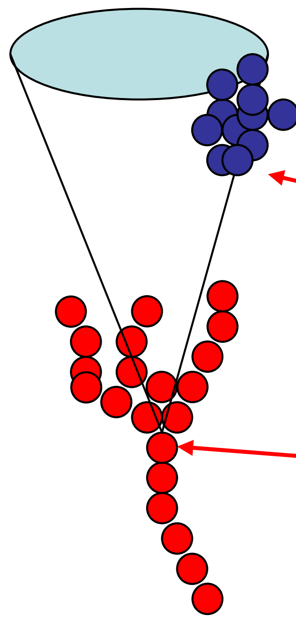
Provides some protection against silly mistakes

Proximity



Distance between hits : limited to first pseudo-layers of cluster

Shower Cone



Associated if fraction of hits in cone > some value

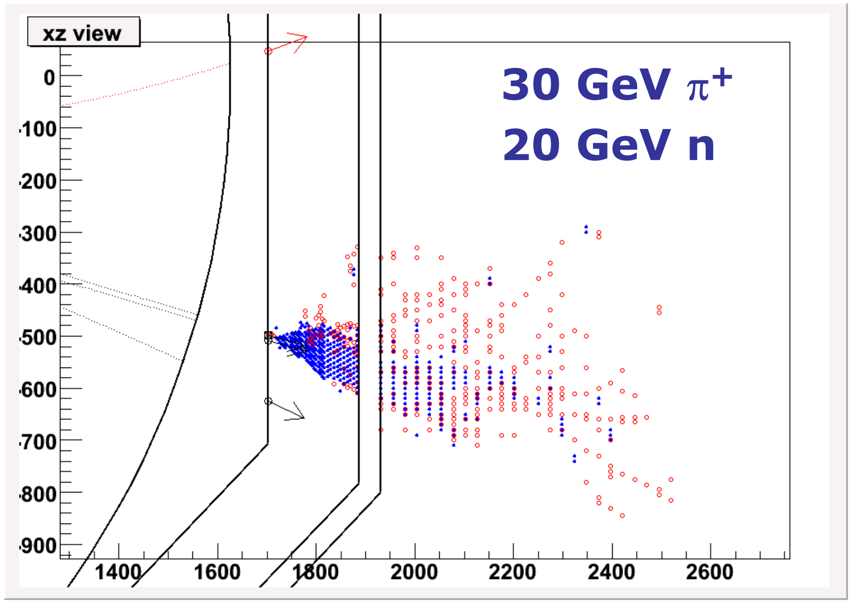
Shower start identified

+Track-Driven Shower Cone

Apply looser cuts if have low E cluster associated to high E track

v) Iterative Reclustering

- ★ Upto this point, in most cases performance is good – but some difficult cases…



- ★ At some point hit the limit of "pure" particle flow
  - ♦ just can't resolve neutral hadron in hadronic shower

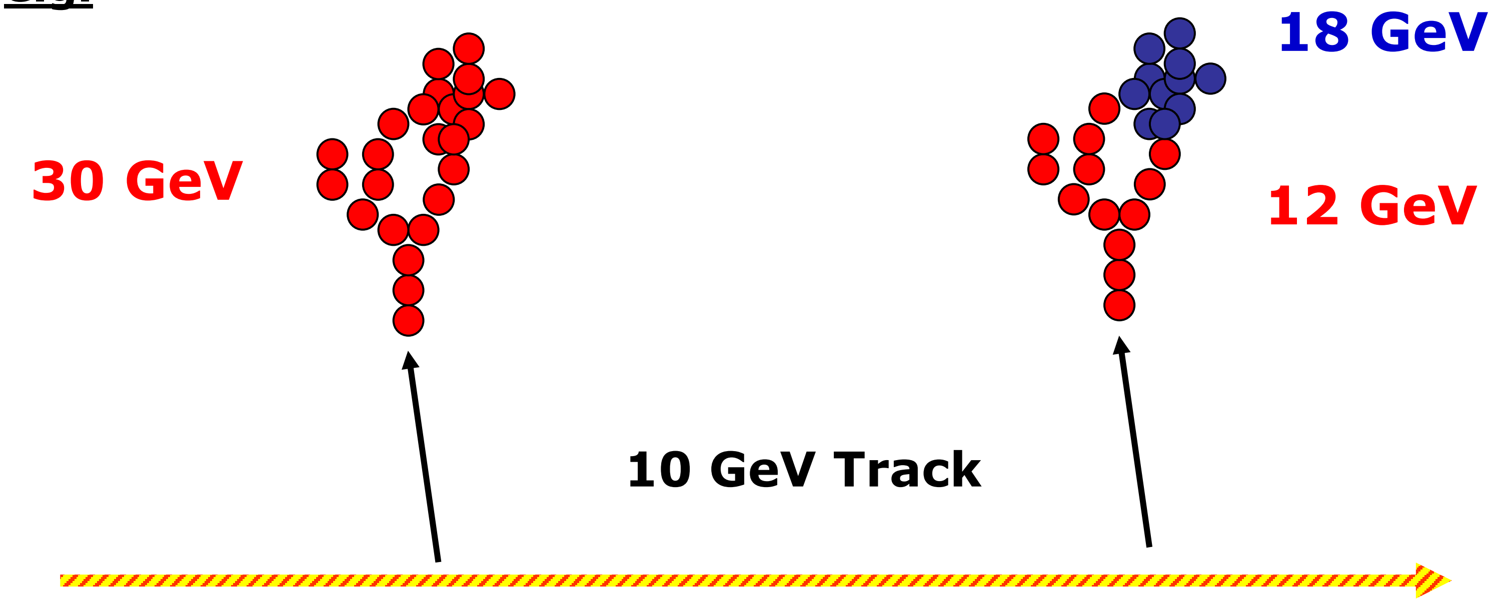
The ONLY(?) way to address this is "statistically"

→

e.g. if have 30 GeV track pointing to 20 GeV cluster
SOMETHING IS WRONG

★ **If track momentum and cluster energy inconsistent : RECLUSTER**

e.g.



Change clustering parameters until cluster splits and get sensible track-cluster match

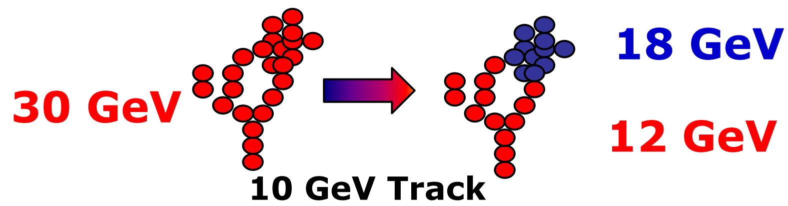
NOTE: **NOT FULL PFA as clustering driven by track momentum**

This is very important for higher energy jets

Iterative Reclustering Strategies

① Cluster splitting

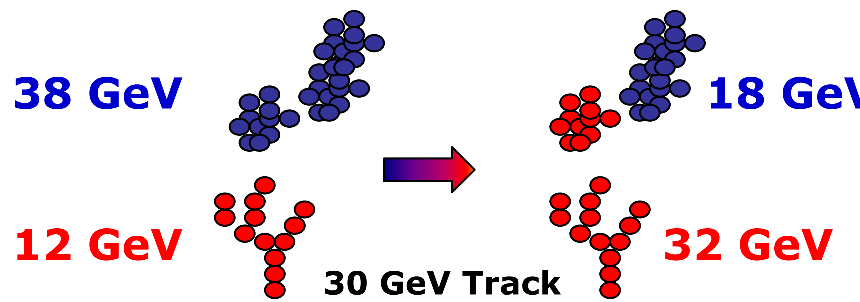
Reapply **entire** clustering algorithm to **hits** in "dubious" cluster. Iteratively reduce cone angle until cluster splits to give acceptable energy match to track



★ Could plug in alternative clustering (to some extent this is now done)

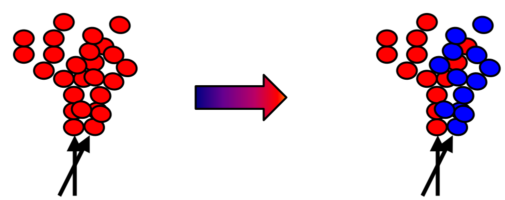
② Cluster merging with splitting

Look for clusters to add to a track to get sensible energy association. If necessary iteratively split up clusters to get good match.



③ Track association ambiguities

In dense environment may have multiple tracks matched to same cluster. Apply above techniques to get ok energy match.

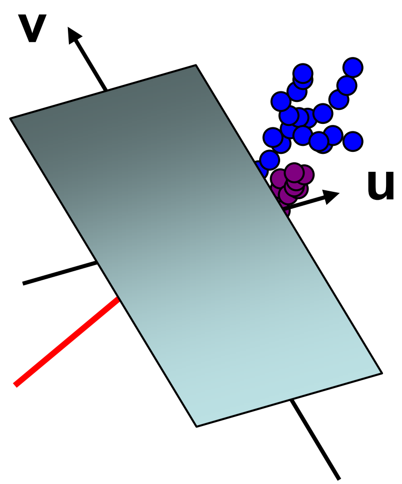


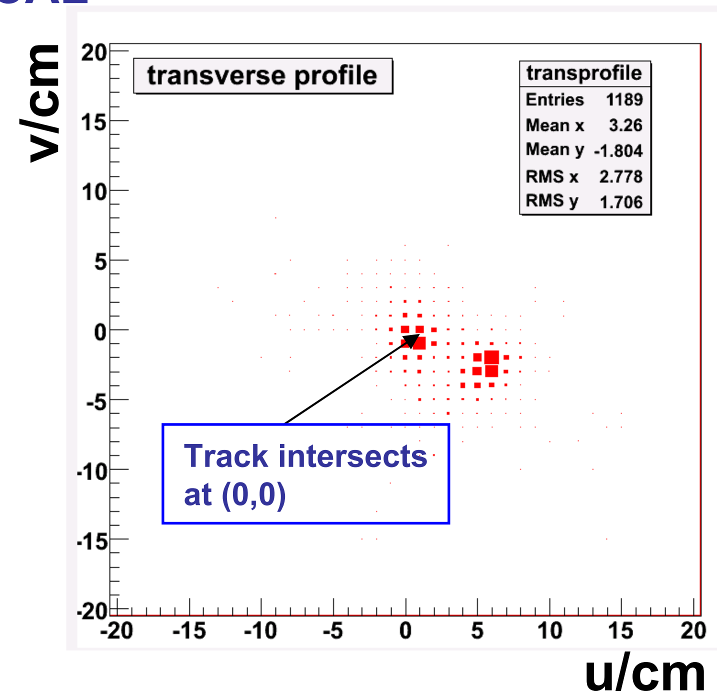
④ "Nuclear Option"

★ If none of above works – kill track and rely on clusters alone (NOT USED)

vi) Photon Recovery (new in v01-01)

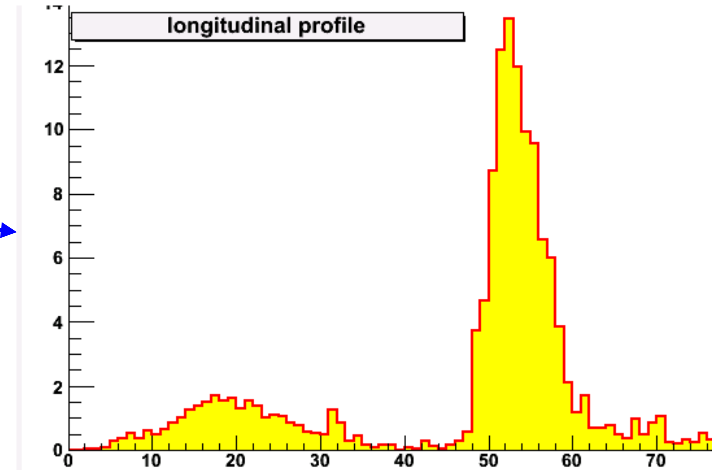
- Using cone clustering photons close to early showering charged hadrons can be merged into a single cluster.
- Introduce a new (very simple) algorithm to recover these
- For each cluster associated with a track:
  - project ECAL hits onto plane perpendicular to radial vector to point where track intersects ECAL
  - search for peaks…



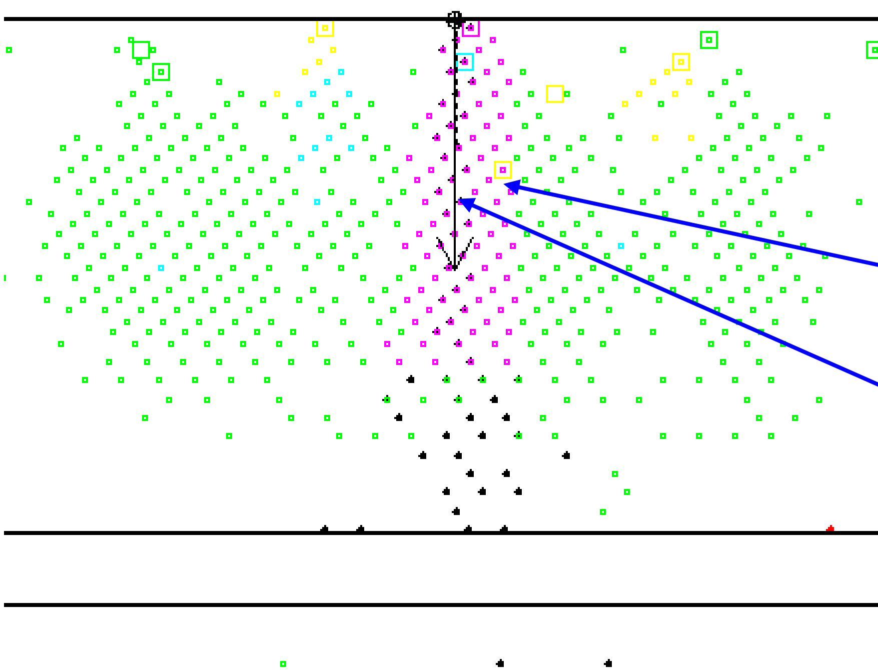


- If there is an isolated peak not associated with "track peak" make new photon cluster if track energy and remaining cluster energy still statistically compatible with track momentum

- Also look for photons where only a single peak is found 
- Implemented by looking at longitudinal profile of "shower"



- Currently very simple cuts – no attempt to identify EM shower profile

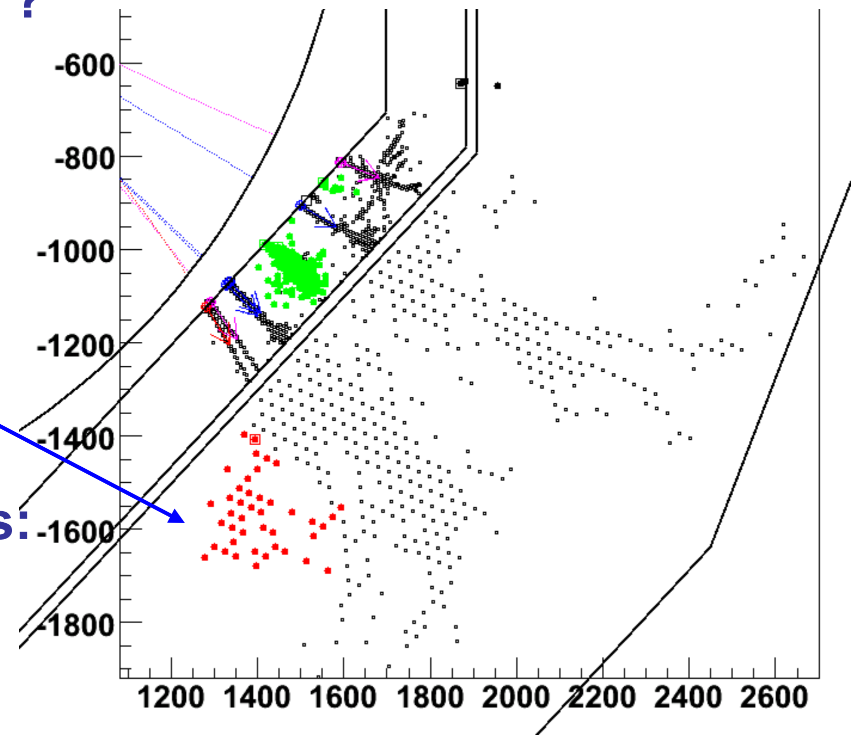


★ Enables resolution of overlapping MIPs and photons, e.g.

vii) Fragment Removal (new in v01-01)

★ Final stage in PFA is to exam “neutral hadron” candidates
- true neutral PFO ?
- fragment from another cluster ?

e.g. is this a neutron/K_L ?



Procedure:

- Look at each non-photon neutral cluster
- Construct some “useful” variables:
  - contact layers
  - cone fraction
  - cluster direction
- Apply very simple (non-optimised) cuts
- Don’t do anything too dumb – i.e. if associating fragment to another results in significant track momentum – cluster energy mismatch.

④ Previous performance

Figures of Merit:

rms_{90}

★ Find smallest region containing 90 % of events

★ Determine rms in this region

E_{JET}	$\sigma_E/E = \alpha\sqrt{(E/GeV)}$ $\lvert\cos\theta\rvert<0.8$
45 GeV	0.30
100 GeV	0.37
180 GeV	0.57
250 GeV	0.75

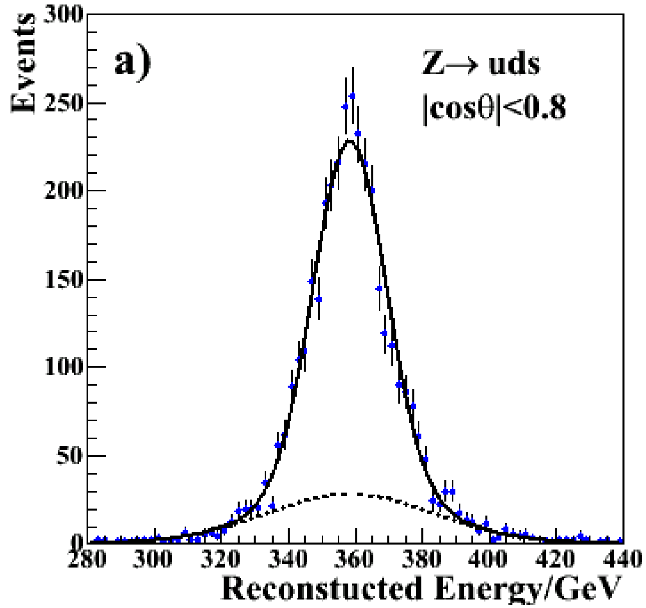
♦ For jet energies < 100 GeV performance was OK

♦ But >100 GeV performance "worryingly" poor

σ_{75}

★ Fit sum of two Gaussians with same mean. The narrower one is constrained to contain 75% of events

★ Quote σ of narrow Gaussian

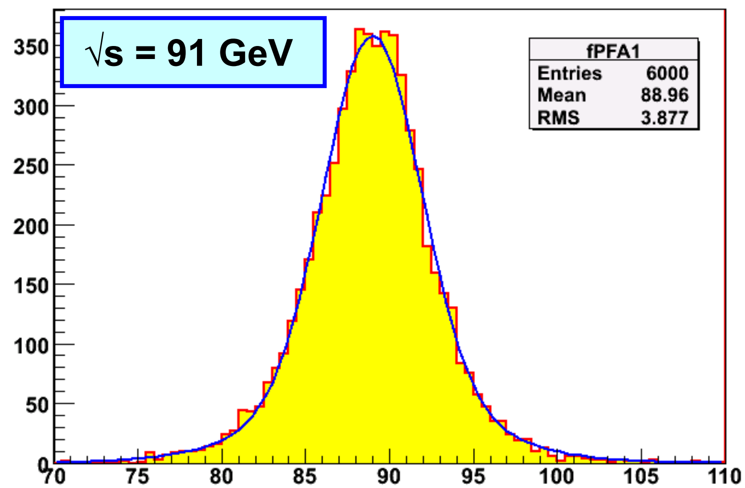


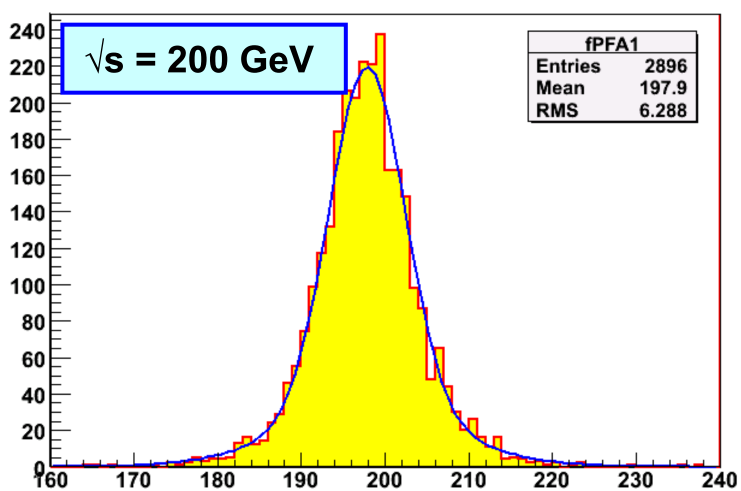
It is found that $rms_{90} \approx \sigma_{75}$

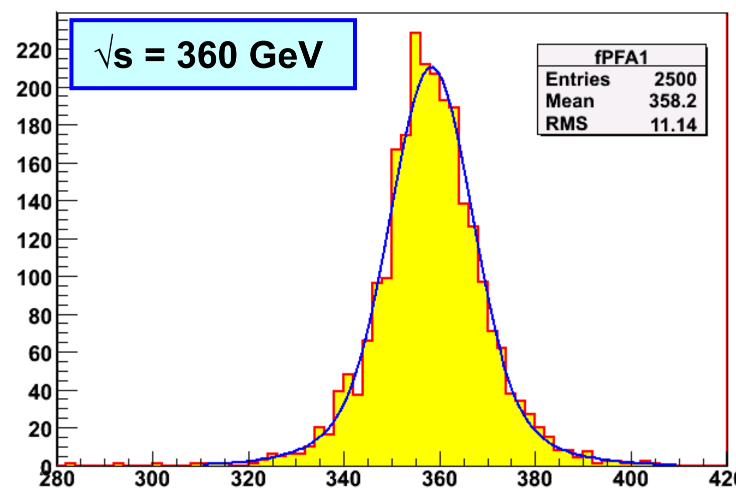
❹ Current performance

Details:

- ★ PandoraPFA v01-01
- ★ LDC00Sc
- ★ B = 4 T
- ★ 3x3 cm HCAL (63 layers)
- ★ Z → uds (no ISR)
- ★ TrackCheater







Current performance

rms90

E_{JET}	$\sigma_E/E = \alpha\sqrt{(E/GeV)}$ $\lvert\cos\theta\rvert<0.7$
45 GeV	0.295
100 GeV	0.305
180 GeV	0.418
250 GeV	0.534

For jet energies < 100 GeV ILC goal reached !!!

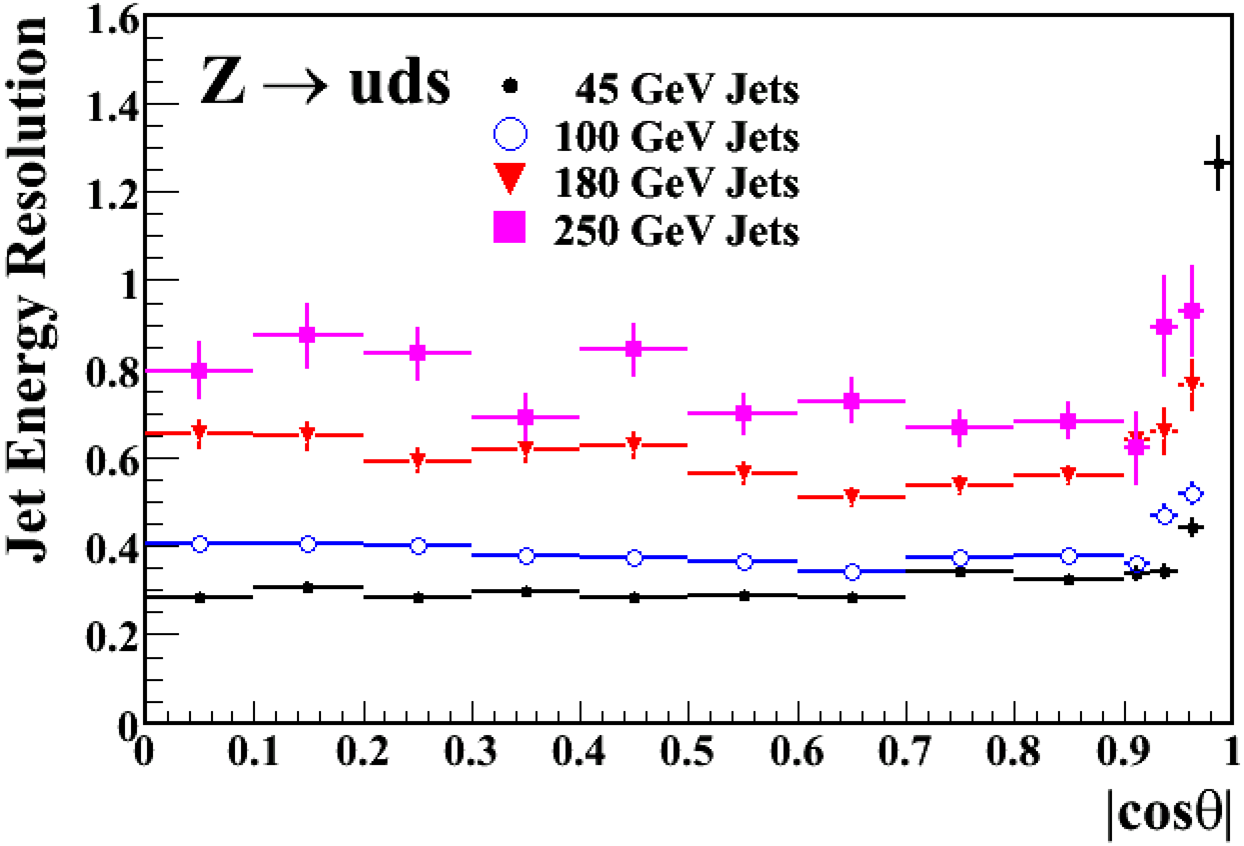
For jet energies ~ 200 GeV close to 40 %/√E(GeV) !!

Opinion:

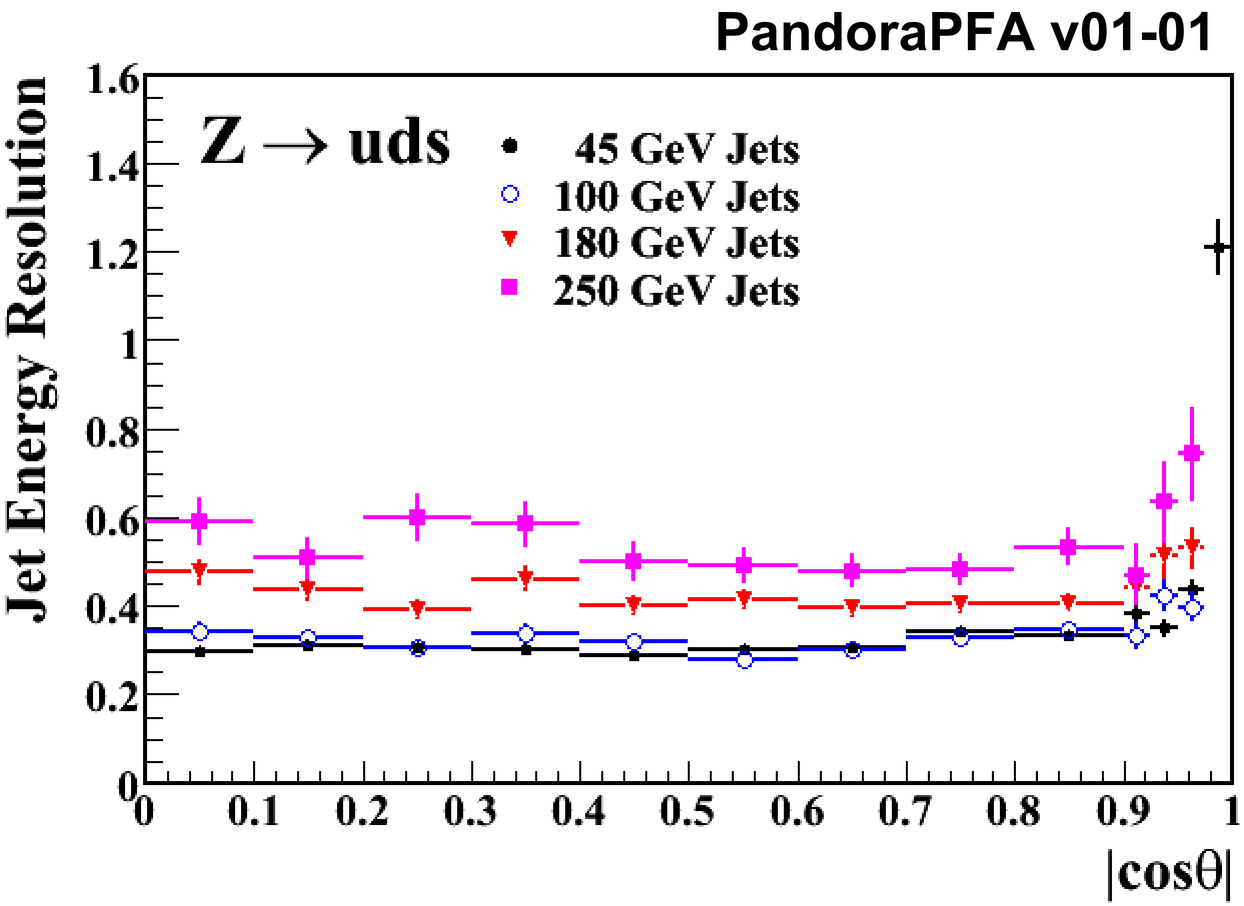
- ★ **There is no doubt in my mind that PFA can deliver the required ILC jet energy performance*.**
- ★ **It is already there for 100 GeV jets - QED**
- ★ **The current code is not perfect (see later), things will get better**

***this is not a statement have made before - please feel free to quote me on this**

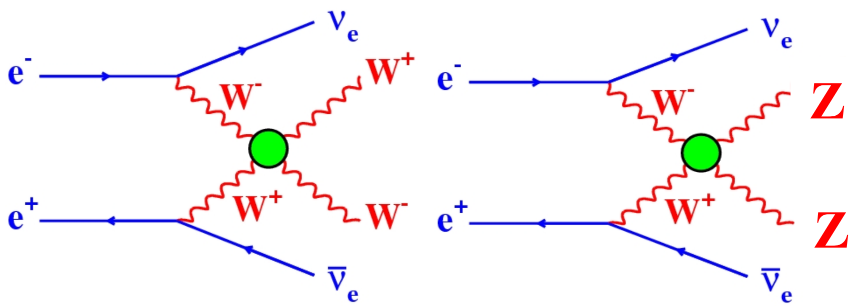
Angular Dependence (OLD)

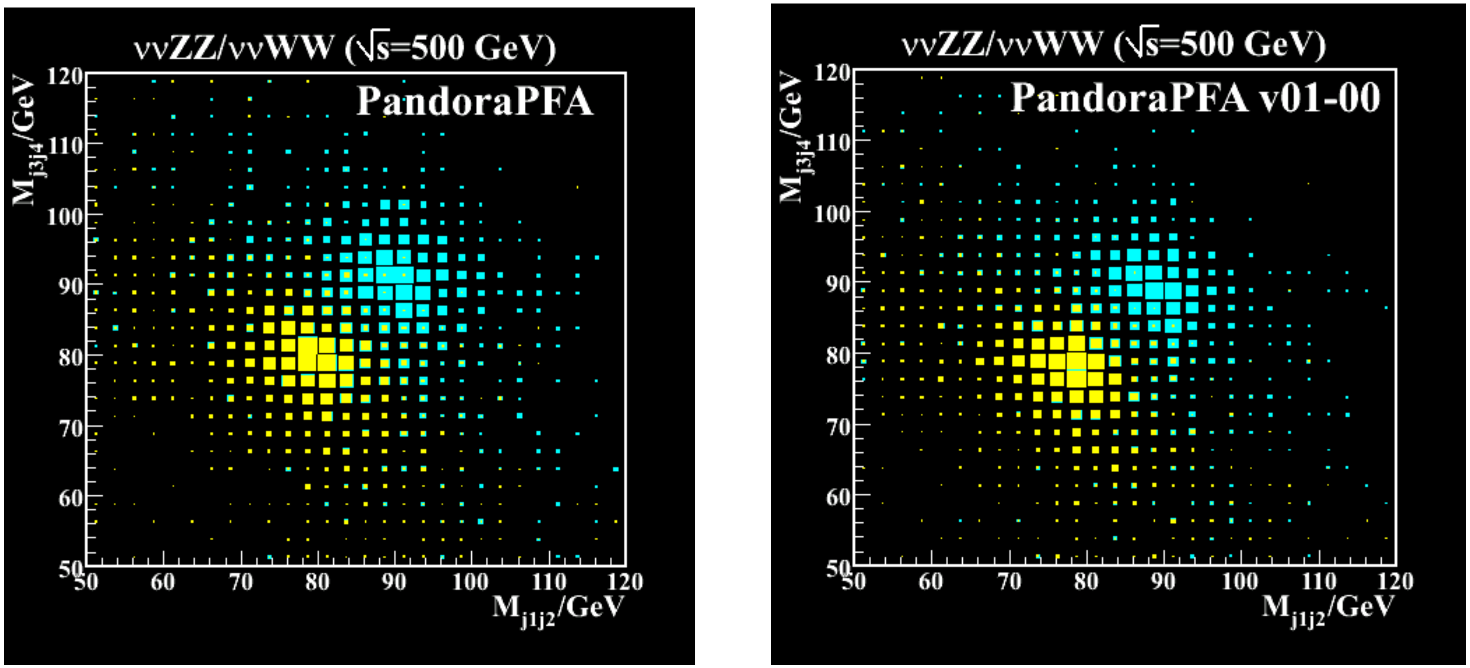


Angular Dependence (NEW)



for completeness…





Visible improvement in WW/ZZ separation **(will return to this later)**

❺ Detector Optimisation Studies

★ From point of view of detector design – what do we want to know ?

Optimise performance vs. cost

★ Main questions (the major cost drivers):

- Size : performance vs. radius
- Granularity (longitudinal/transverse): ECAL and HCAL
- B-field : performance vs. B

★ To answer them use **MC simulation + PFA algorithm**

- Need a good MC simulation
- Need realistic PFA algorithm
  (want/need results from multiple algorithms)

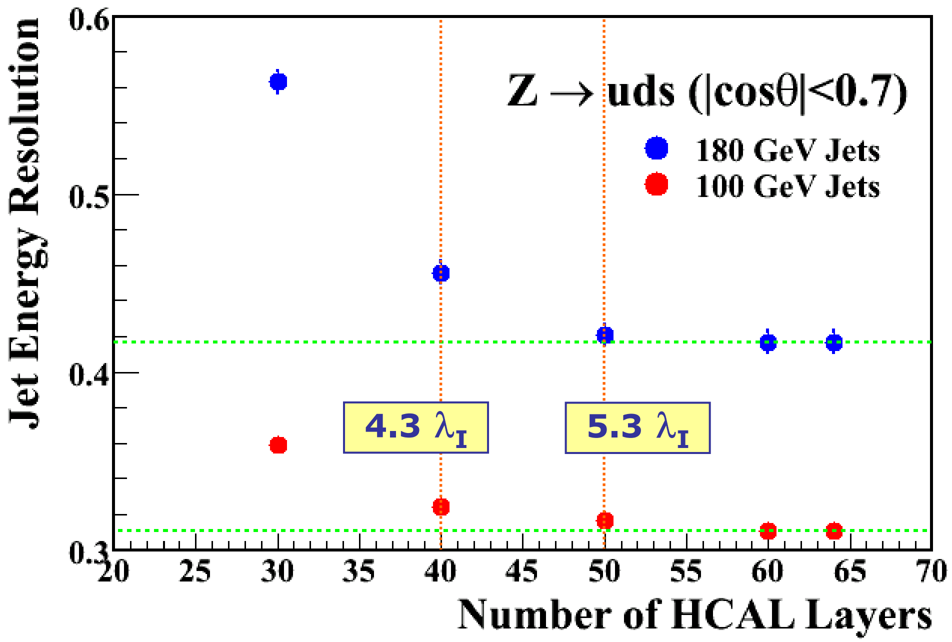
Caveat Emptor

★ These studies are interesting but not clear how seriously they should be taken

- how much is due to the detector
- how much due to imperfect algorithm

HCAL Depth and Transverse segmentation

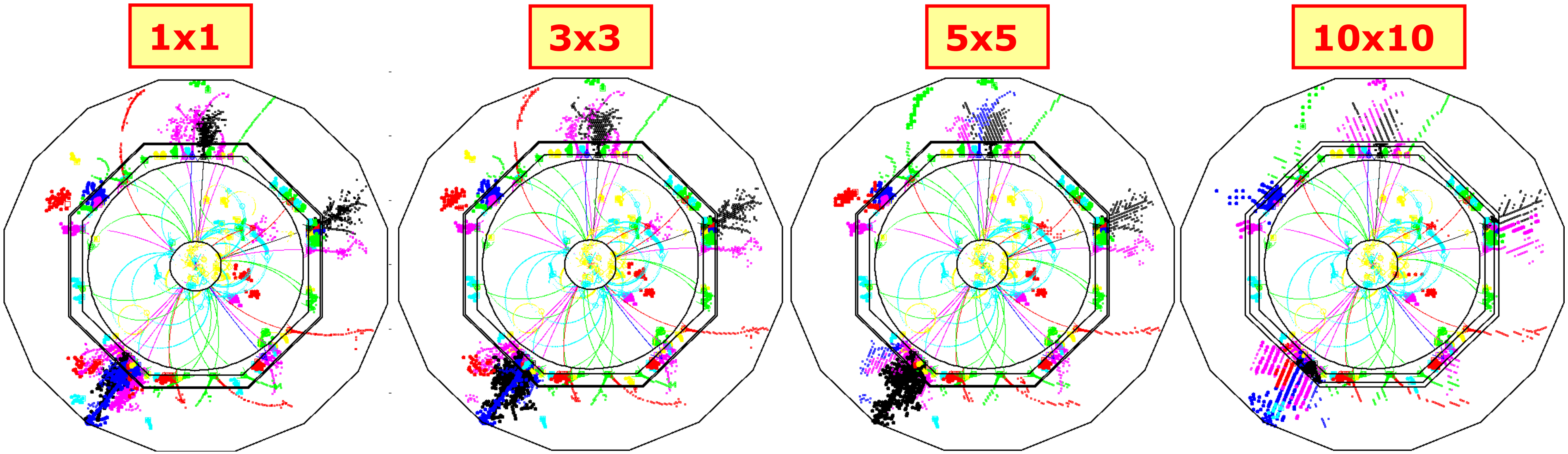
★ Investigated HCAL Depth (interaction lengths)
- Generated Z→uds events with a large HCAL (63 layers)
  - approx 7 λ_I
- In PandoraPFA introduced a configuration variable to truncate the HCAL to arbitrary depth
- Takes account of hexadecagonal geometry

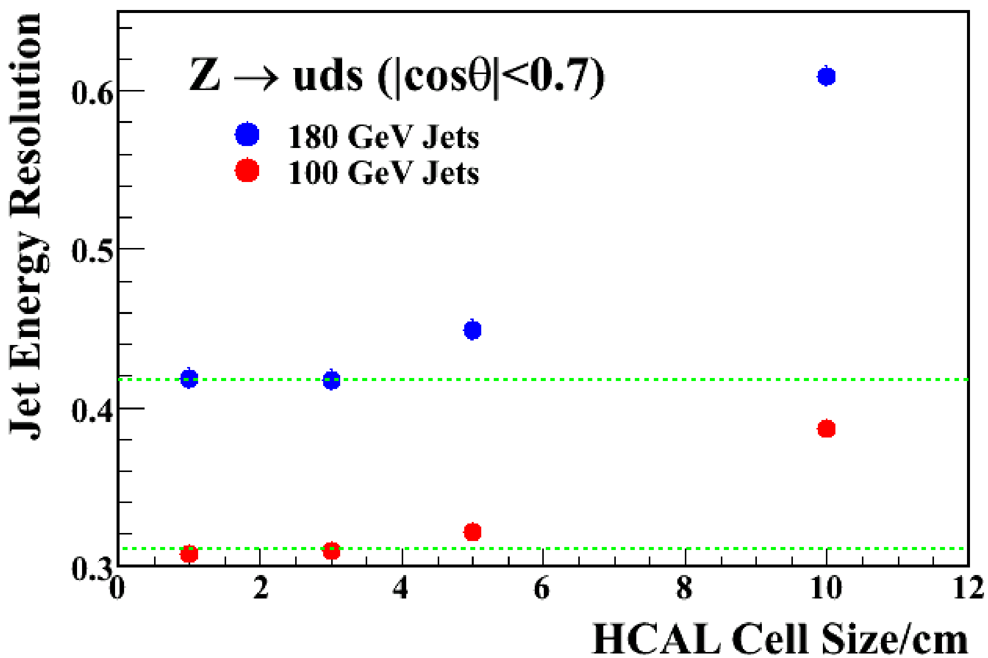


♦ HCAL leakage is significant for high energy
♦ Argues for ~ 5 λ_I HCAL

NOTE: no attempt to account for leakage – i.e. using muon hits - this is a worse case

✶ Analogue scintillator tile HCAL : change tile size 1x1 ➞ 10x10 mm^2

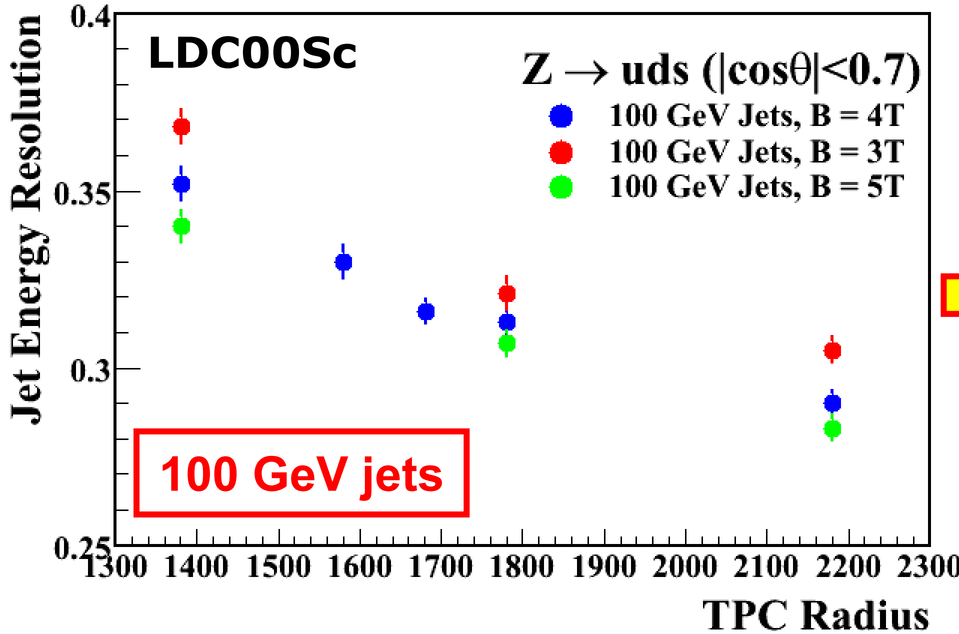




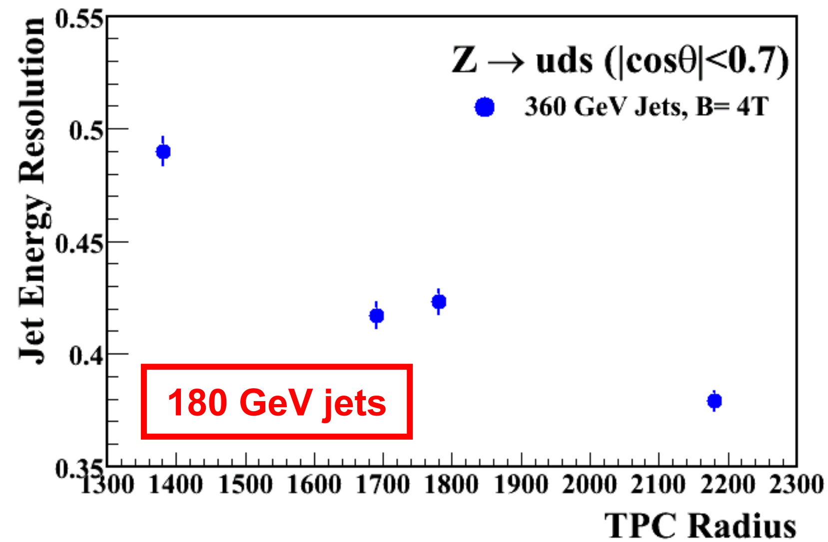
"Preliminary Conclusions"

- 3x3 cm^2 cell size ☺
- No advantage ➞ 1x1 cm^2
  - physics ?
  - algorithm artefact ?
- 5x5 cm^2 degrades PFA
  - Does not exclude coarser granularity deep in HCAL

Radius vs Field



Radius more important than B-field

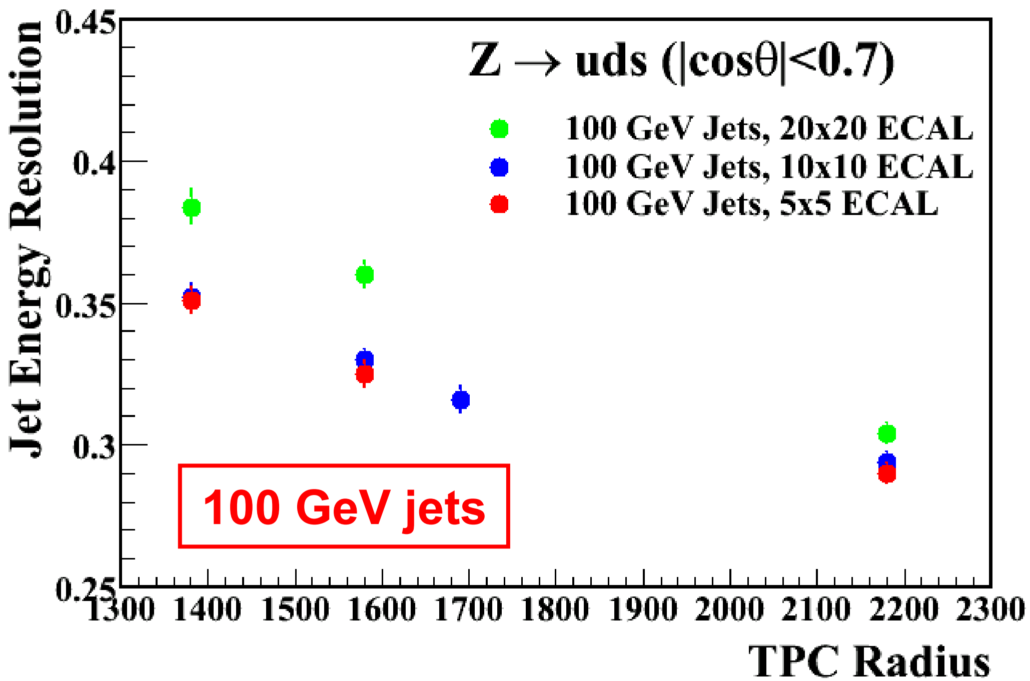


Radius more important at higher energies

B-field studies on the way…

ECAL Transverse Granularity

- Use Mokka to generate Z → uds events @ 200 GeV with different ECAL segmentation: 5x5, 10x10, 20x20 [mm^2]



<u>With PandoraPFA</u>

- 20x20 segmentation looks too coarse
- For 100 GeV jets, not a big gain going from 10x10 → 5x5mm^2

[for these jet energies the contributions from confusion inside the ECAL is relatively small – need]

❻ PandoraPerfectPFA

- ★ Recently added PerfectPFA option in Pandora (not yet in CVS)
  - <parameter name="PerfectPFA" type="int"> 1 </parameter>
- ★ Uses MC information to create the ProtoClusters
- ★ The rest of the algorithm is the same
- ★ Although very fresh, can already learn something…
- ★ Process same events/same analysis and compare PFA to perfect PFA
  - Note in these studies the tracks are the same "TrackCheater"

i) How close to being "Perfect" is PandoraPFA?

E_{JET}	$\sigma_E/E = \alpha\sqrt{(E/GeV)}$ $\lvert\cos\theta\rvert<0.7$	
	PerfectPandora	PandoraPFA
100 GeV	0.220	0.305
180 GeV	0.305	0.418

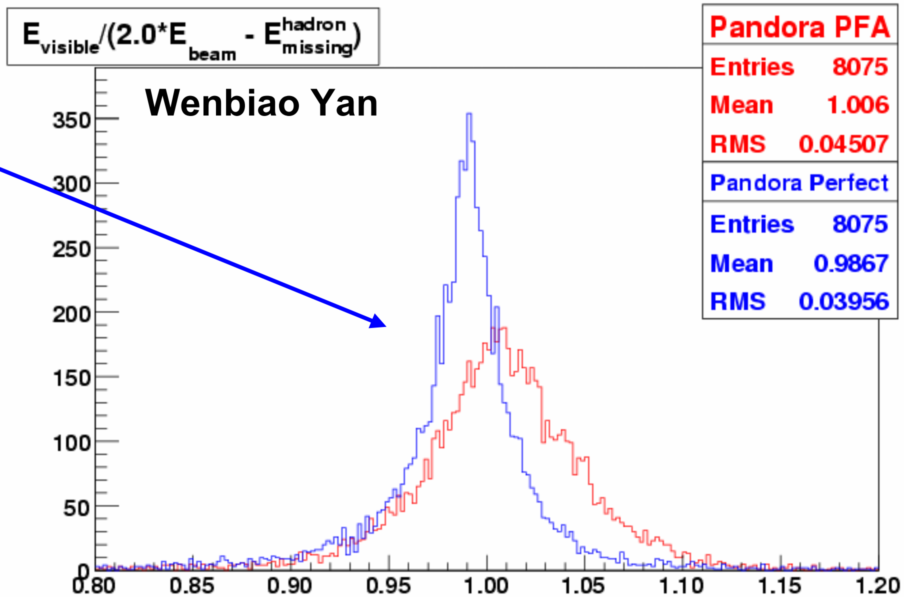
Still someway to go even for low energy jets – needs study

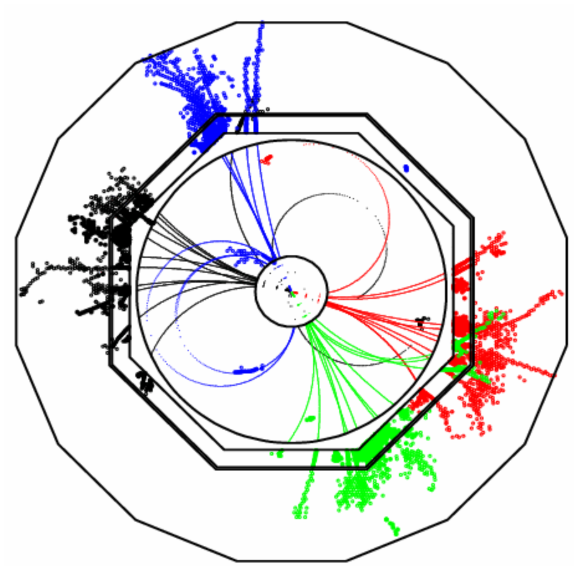
ii) PFA impact in a real physics process

e.g. $e^+e^- \rightarrow \nu\bar{\nu}W^+W^- \rightarrow \nu\bar{\nu}qqqq$ $\sqrt{s} = 800\,\text{GeV}$

★ First compare visible energy from PFA with expected (i.e. after removing neutrinos/forward tracks+clusters)

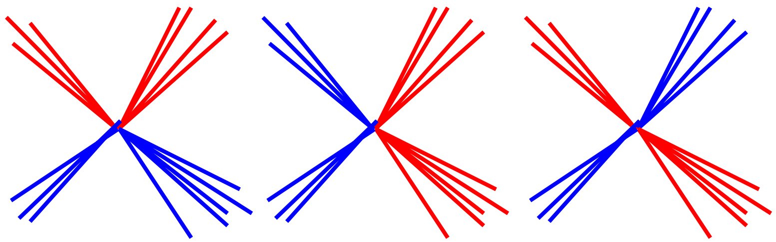
♦ PerfectPFA gives better energy resolution than PandoraPFA (as expected)



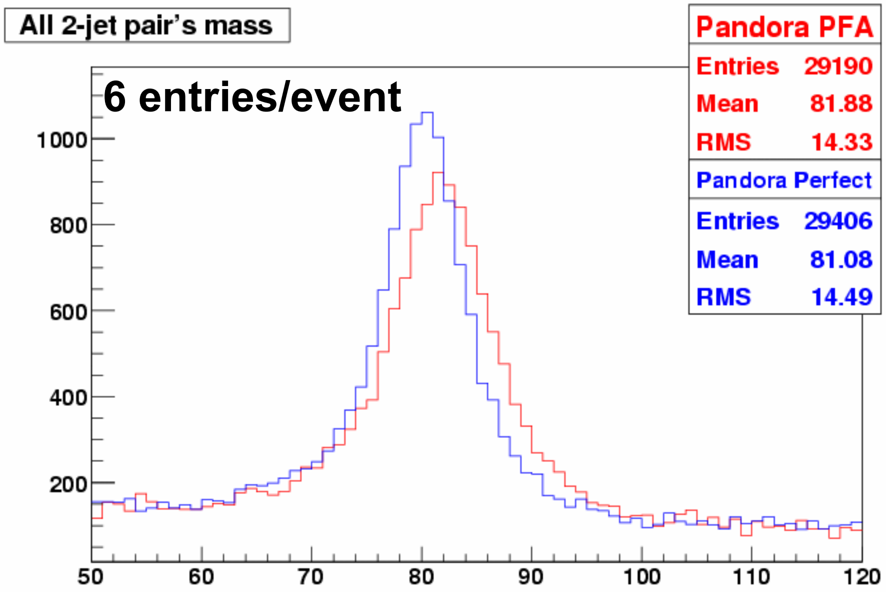


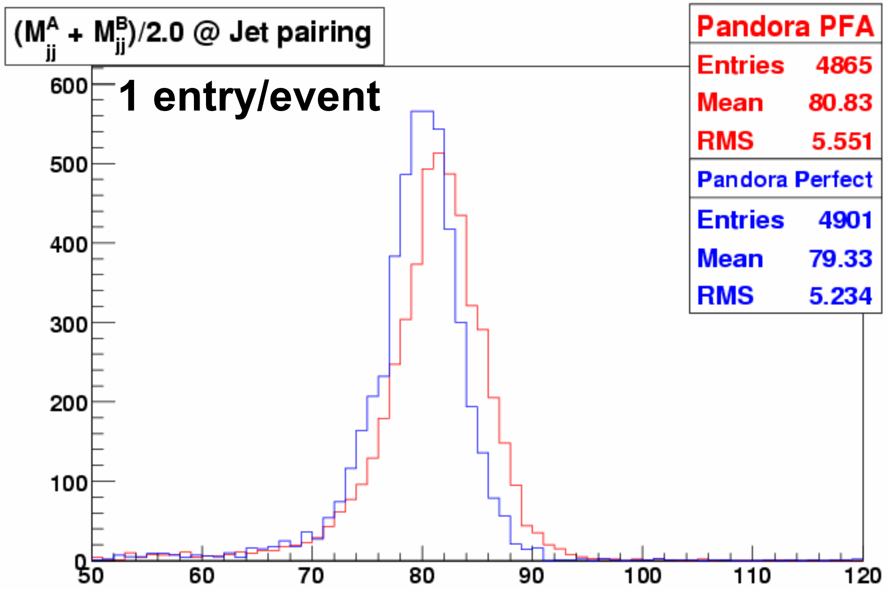
★ Does this difference make it through to a physics analysis (i.e. after jet finding/ jet pairing) ?

★Force event into 4 jets (Durham)
★Plot masses of the 2 Ws formed from the 3 possible jet-pairings



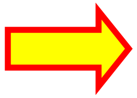
HERE: PandoraPFA ~ PerfectPFA





★Choose pairing with smallest mass difference
★Plot average mass of the 2 Ws

HERE: PandoraPFA ~ PerfectPFA



"Physics Ready PFA"

❼ Deficiencies, Random Comments, Outlook and Conclusions

Deficiencies:

★ PandoraPFA has evolved solely with the aim of improving performance … never overly concerned with niceties…

★ Very little has been optimised:
- Photon ID – very crude
- Photon Recovery – very crude
- Fragment Removal – very crude

Plenty of room for improvement

Random Comments:

★ PFA = much more than clustering

★ In developing code – learnt importance of:
- extreme care - do make any unnecessary mistakes.
- use of track momentum – cluster energy to spot to PFA errors

★ Calibration – not trivial. Must check with single particles

★ HCAL energy resolution vs. reclustering ? Improved resolution would help resolve PFA errors…

Outlook:

★ Aim for improved version for **LCWS07…**

★ User feedback very important

- it is possible that there are still hidden bugs
- help identify deficiencies, e.g. Predrag's talk

Conclusions:

★ PandoraPFA is not perfect

★ Things can only get better

★ Nevertheless, I believe it has been demonstrated that PFA can give ILC performance goals at $\sqrt{s}$ = 500 GeV and $\sqrt{s}$ = 1 TeV

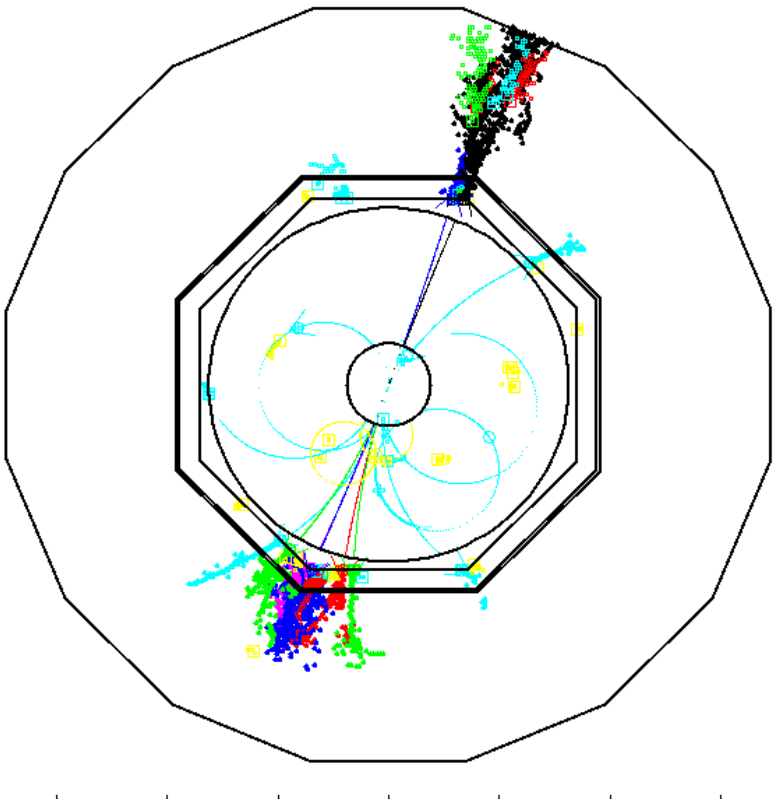
E_{JET}	$\sigma_E/E = \alpha\sqrt{(E/GeV)}$ $\lvert\cos\theta\rvert<0.7$
45 GeV	0.295
100 GeV	0.305
180 GeV	0.418
250 GeV	0.534

PFA - QED ?

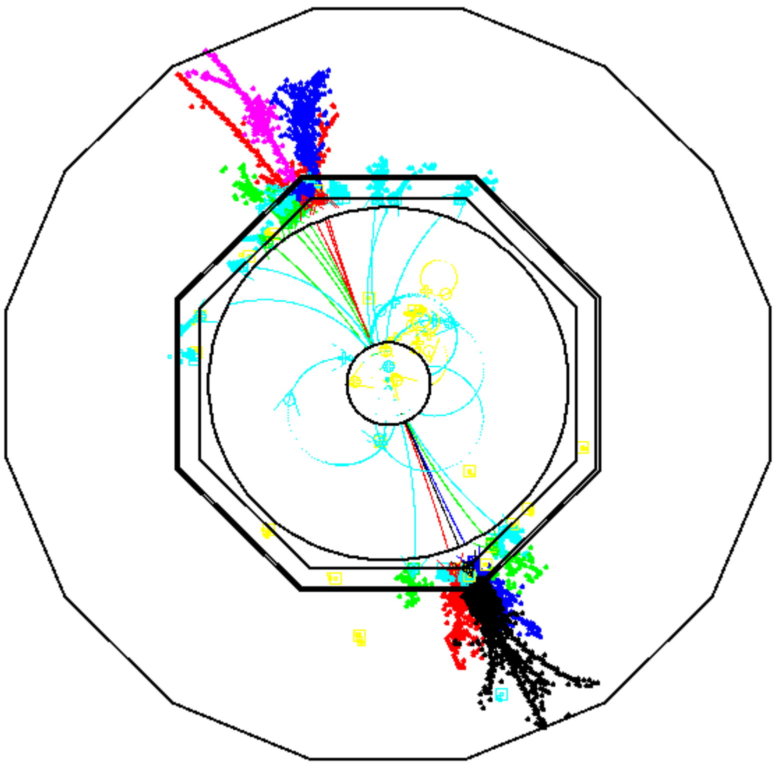
Backup : When PandoraPFA Goes Bad

- A few di-jet events at $\sqrt{s}$ = 360 GeV

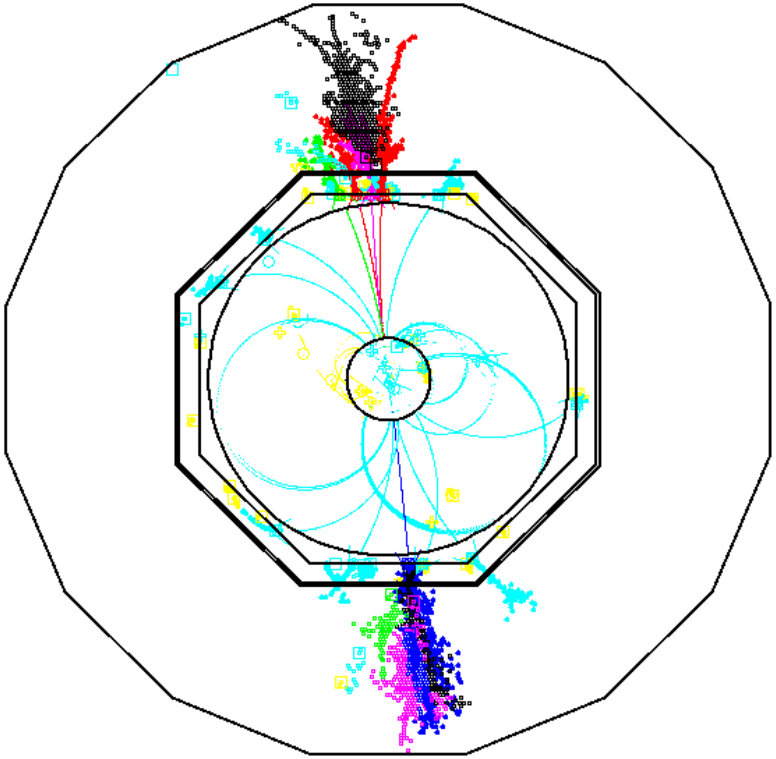
E_{reco} = 337 GeV



E_{reco} = 338 GeV



E_{reco} = 382 GeV



E_{reco} = 391 GeV