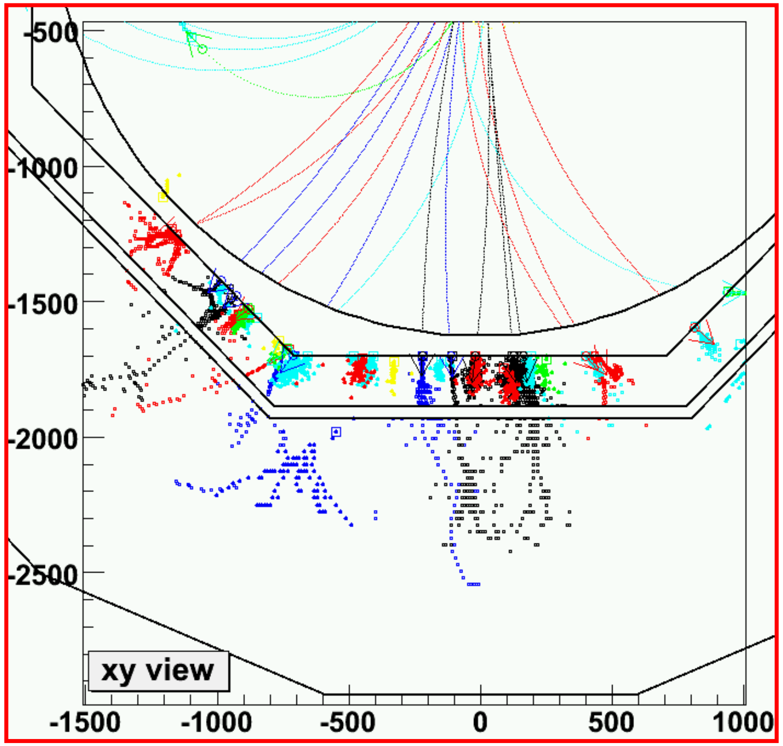
Status of Particle Flow with PandoraPFA

Mark Thomson
University of Cambridge



This Talk:

➊ **Brief reminder of basic algorithm**
➋ **Statistical Reclustering (NEW)**
➌ **Updated Results**
going beyond the Z
➍ **Remaining problems**
➎ **Outlook**

➊ PandoraPFA : brief overview

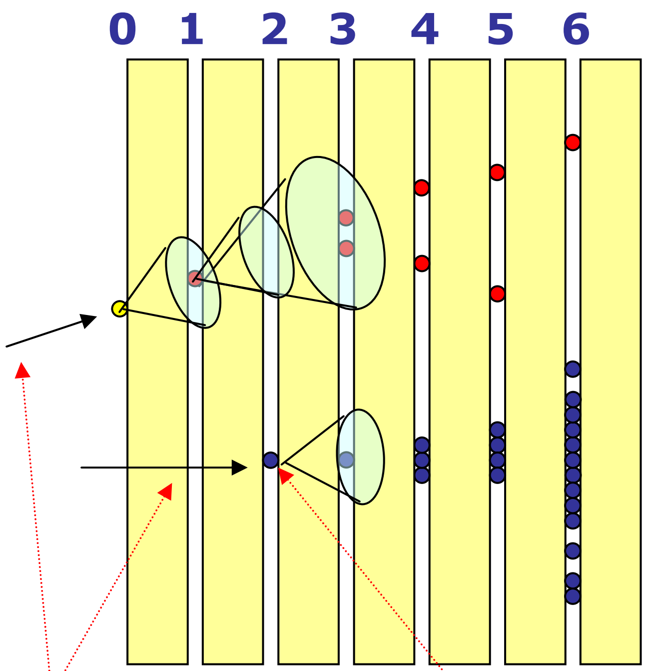
- ★ **ECAL/HCAL reconstruction and PFA performed in a single algorithm**
- ★ **Keep things fairly generic algorithm**
  - ★ **applicable to multiple detector concepts**
- ★ **Use tracking information to help ECAL/HCAL clustering**

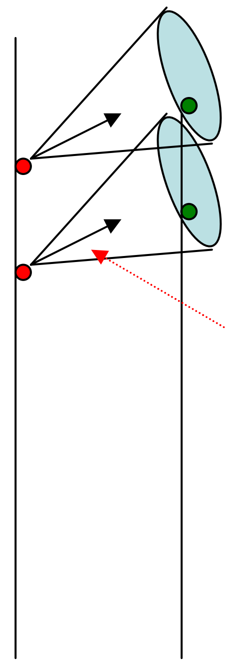
Five Main Stages:

i. **Loose clustering in ECAL and HCAL**
ii. **Topological linking of clearly associated clusters**
iii. **Courser grouping of clusters**
iv. **Statistical reclustering**
v. **Formation of final Particle Flow Objects (reconstructed particles)**

i) ECAL/HCAL Clustering

- ★ Start at inner layers and work outward
- ★ Associate Hits with existing Clusters
- ★ Step back **N** layers until associated
- ★ Then try to associate with hits in current layer
- ★ If no association made form new Cluster
- ★ + tracks used to seed clusters





Simple cone algorithm based on current direction + additional N pixels

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

Initial cluster direction

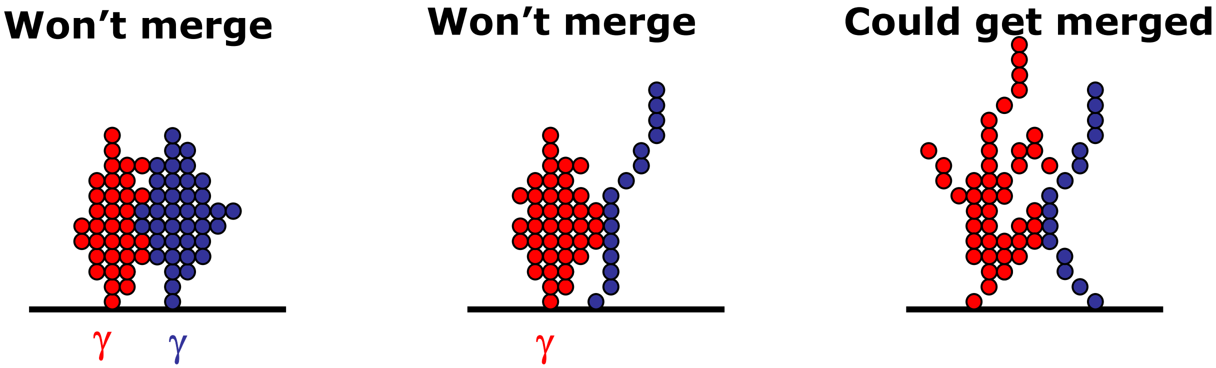
Unmatched hits seeds new cluster

ii) Cluster Association Part I

- 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 – catches what's been missed but rather crude

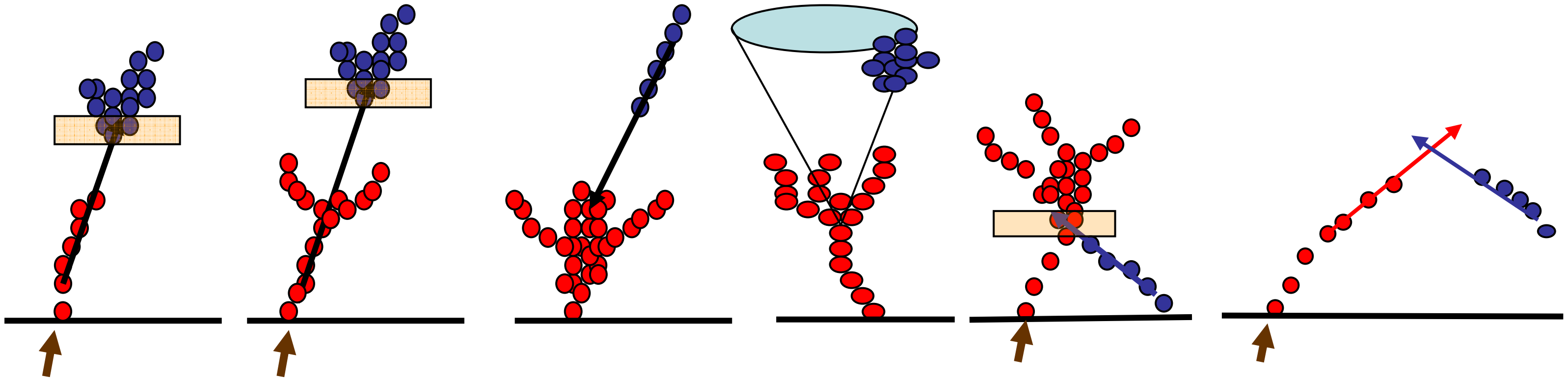
Photon ID

- 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



Cluster Association Part I cont.

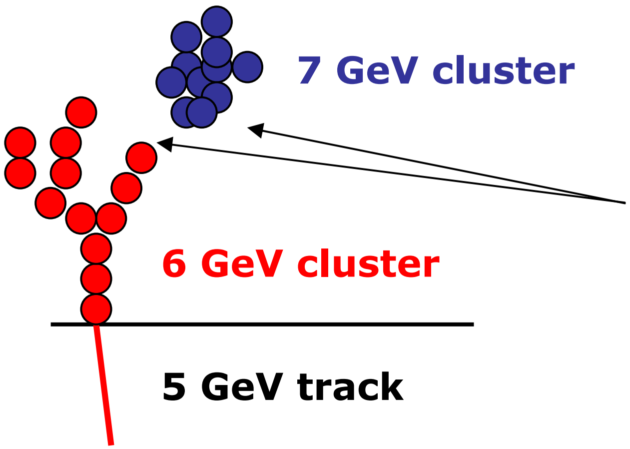
Join clusters which are clearly associated
make use of high granularity + tracking capability



Only clear associations – almost no mistakes

iii) 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.



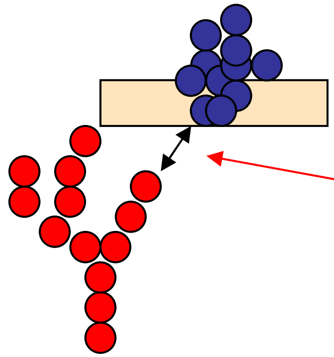
This cluster association would be forbidden if $|E_1 + E_2 - p| > 3\,\sigma_E$

Provides some protection against "silly" mistakes

★ Cluster reconstruction and PFA not independent

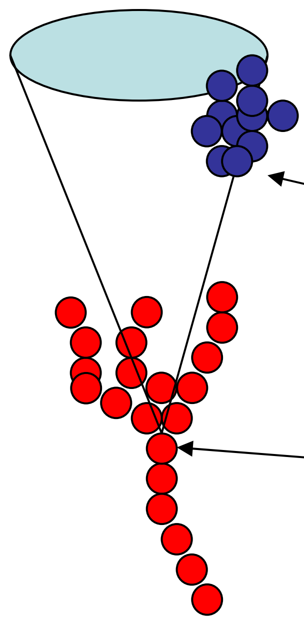
Course Cluster Association

Proximity



Distance between hits -limited to first layers

Shower Cone



Associated if fraction of hits in cone > some value

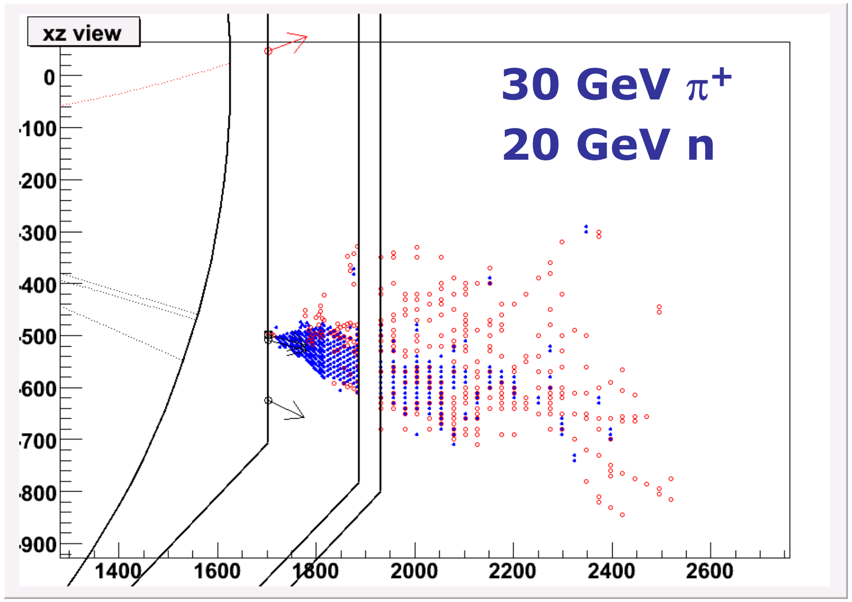
Shower start identified

+Track-Driven Shower Cone

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

❷ Statistical Reclustering

- ★ **Generally performance is good at Z** **e.g. see Cambridge slides**
- ★ **But things get much harder at higher energies**
- ★ **At some point the algorithm fails…**



e.g. overlapping hadronic showers

- ★ **hard to see how pure clustering can resolve these two clusters**

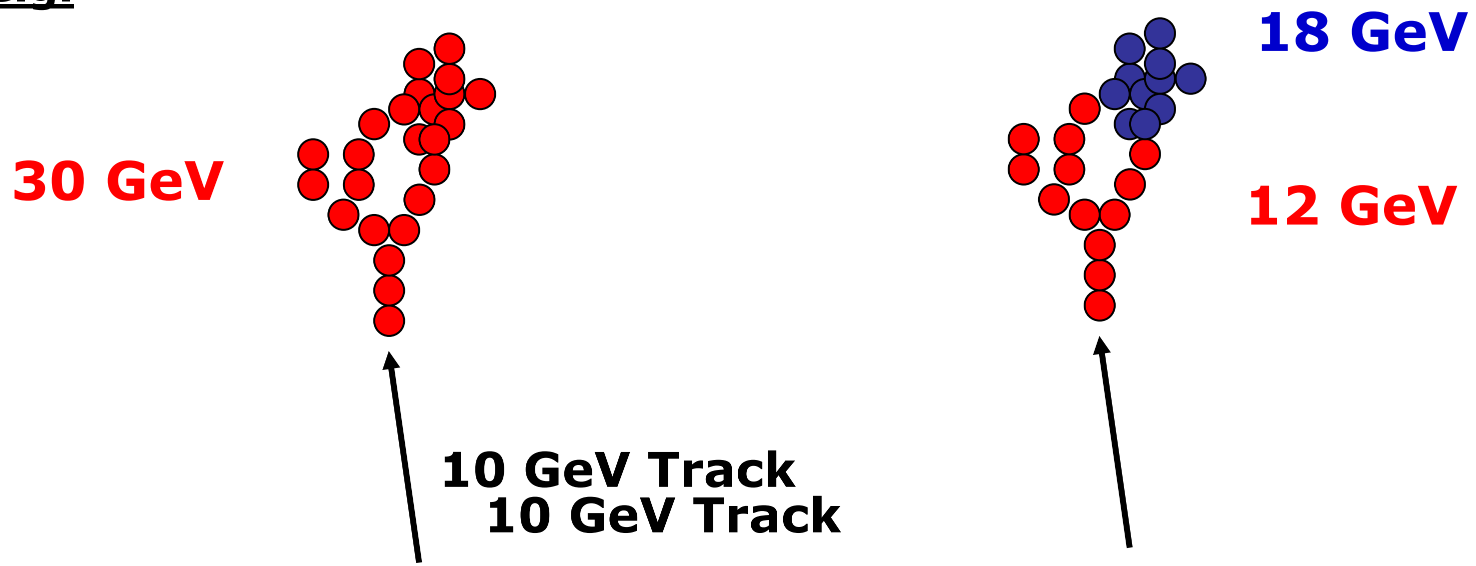
- ★ **Ulitmately will hit the limit of "pure" particle flow**
  - ♦ **just can't resolve neutral hadron in hadronic shower**

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

e.g. if have 30 GeV track pointing to 50 GeV cluster SOMETHING IS WRONG

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

<u>e.g.</u>



Change clustering parameters until split cluster + get sensible track-cluster match

<u>NOTE:</u> THIS IS NO LONGER "FULL PFA" as clustering is driven by track momentum

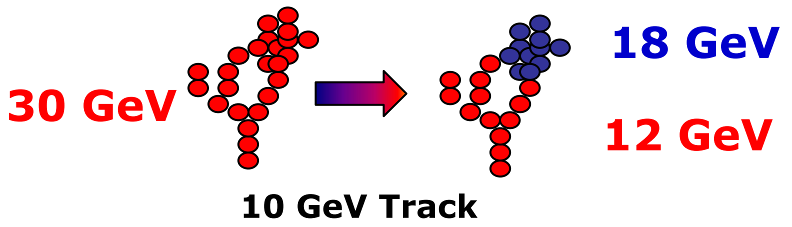
★ **If can't find a sensible reclustering use the ultimate sanction i.e. do not use track information**

Iterative Reclusting

❶ Cluster splitting

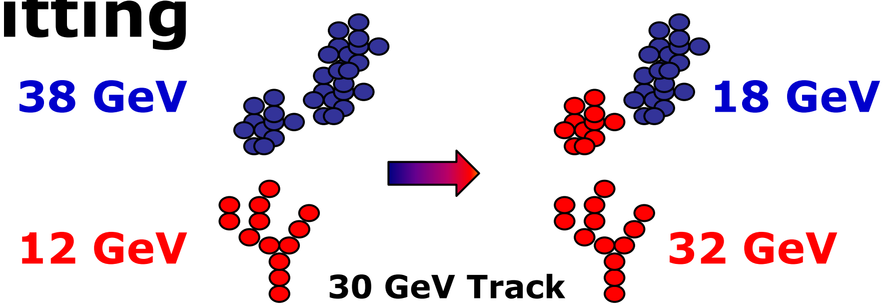
Reapply entire clustering algorithm to hits in "dubious" cluster. Iteratively reduce cone angle until cluster splits to give acceptable energy match to track

- ★ Could plug in alternative clustering



❷ 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.



❸ Track association ambiguities

In dense enviroment may have multiple tracks matched to same cluster. Apply above techniques to get a good energy match.

❸ "Nuclear Option"

- ★ If none of above works – kill track and rely on clusters alone

❸ Current Performance (as of 15/6/06)

Example Reconstruction

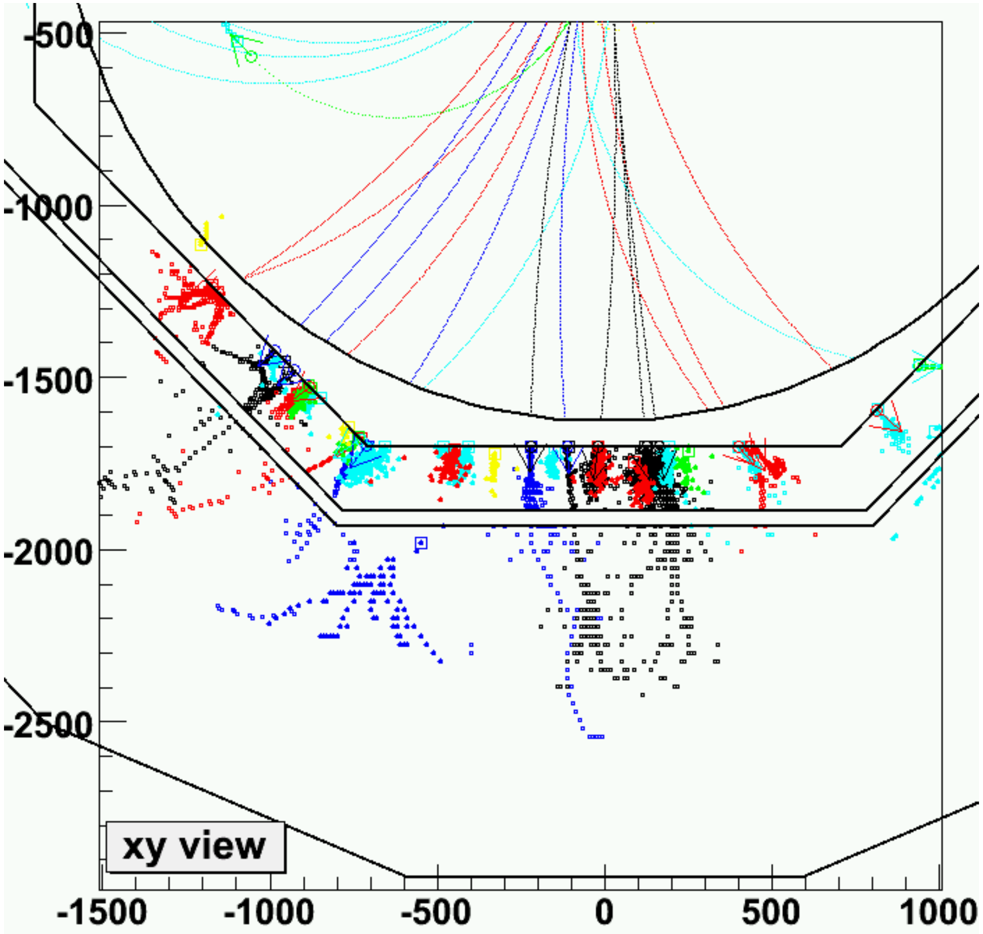
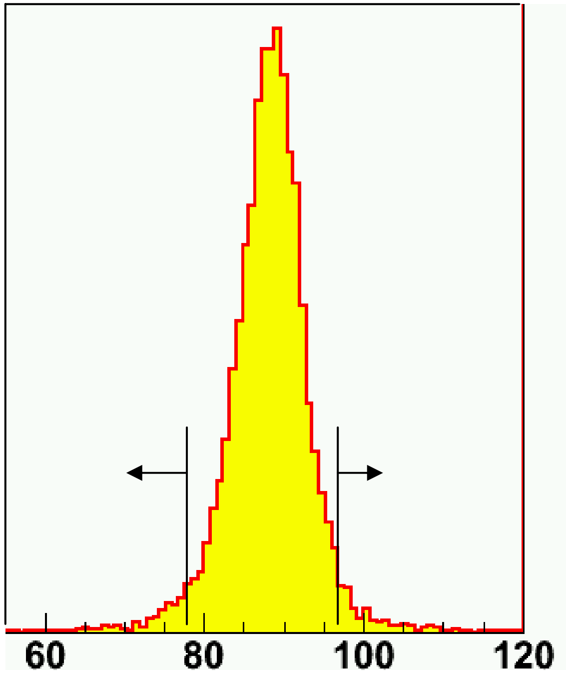


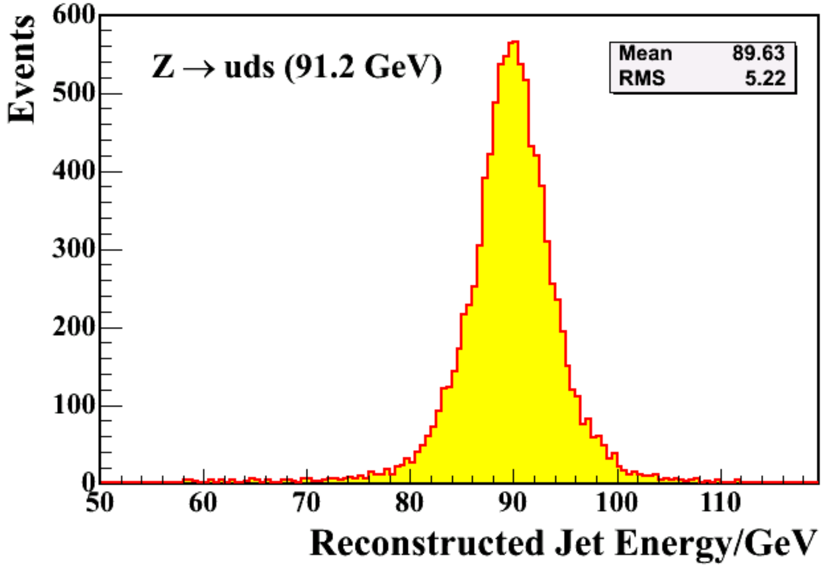
Figure of Merit:



★ **Find smallest region containing 90 % of events**

★ **Determine rms in this region**

PFA Results (Z →uds)

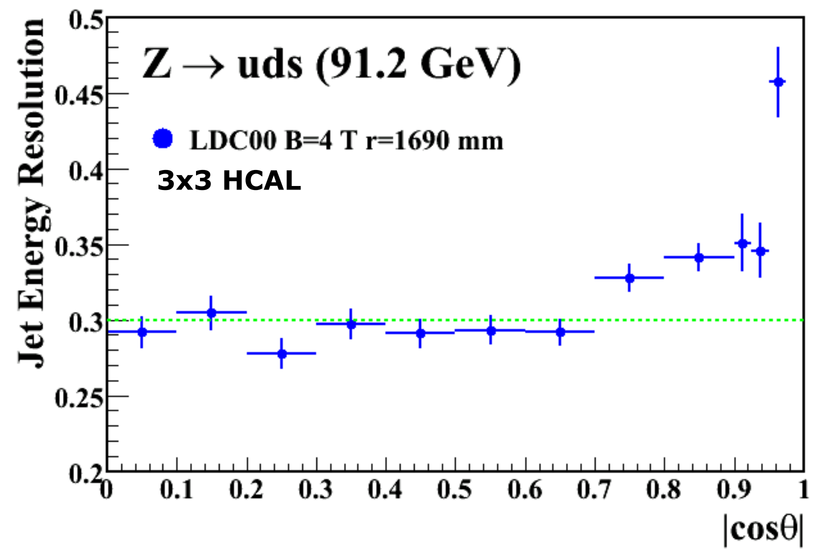


LDC00

|cosθ|	$\sigma_E/E = \alpha\sqrt{(E/GeV)}$
all	33.4±0.3%
<0.9	30.5±0.3 %
<0.7	29.2±0.4 %

ILC GOAL OF 30 % ACHIEVED !

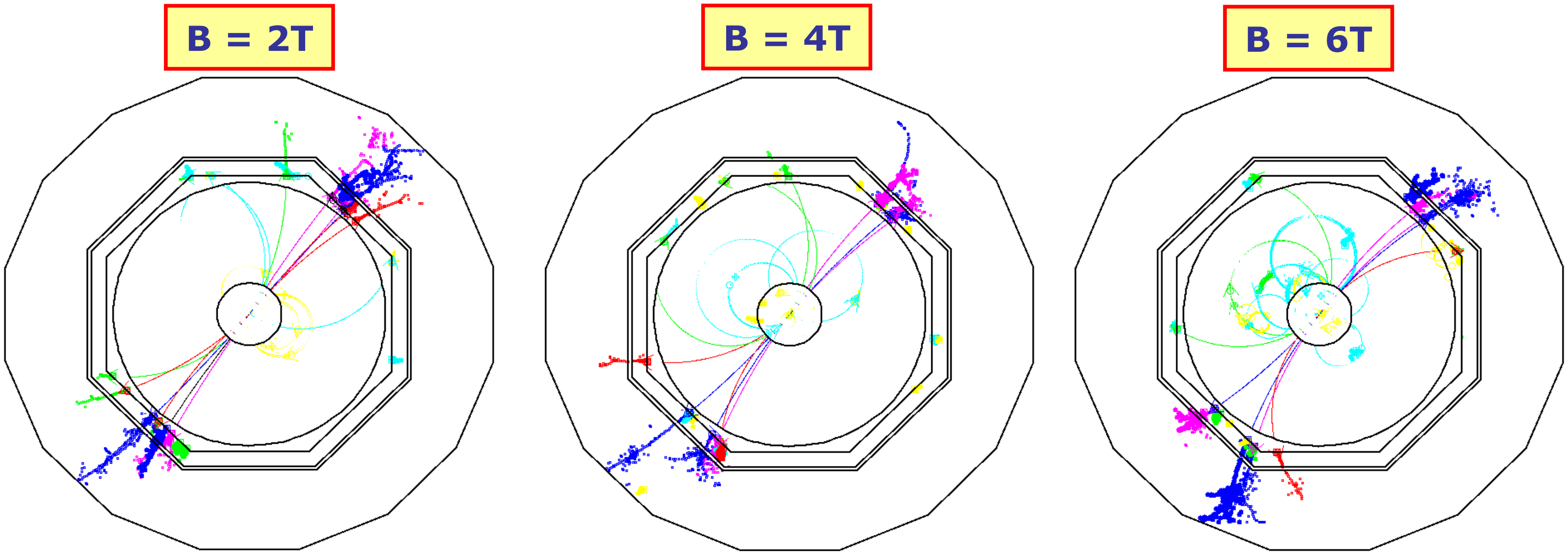
- ★ BUT only for Z at 91.2 GeV
- ★ In barrel essentially "perfect"
- ★ Endcap issues



Good perfomance at Z : now test algorithm on other detectors/samples

e.g. B-Field

LDC00 Detector (≈ TESLA TDR) – same event different B



B-Field	$\sigma_E/E = \alpha\sqrt{(E/GeV)}$	
	All angles	$\lvert\cos\theta\rvert<0.7$
2 Tesla	34.1±0.3%	30.8±0.4 %
4 Tesla	33.4±0.3 %	29.2±0.4 %
6 Tesla	34.4±0.3 %	29.7±0.4 %

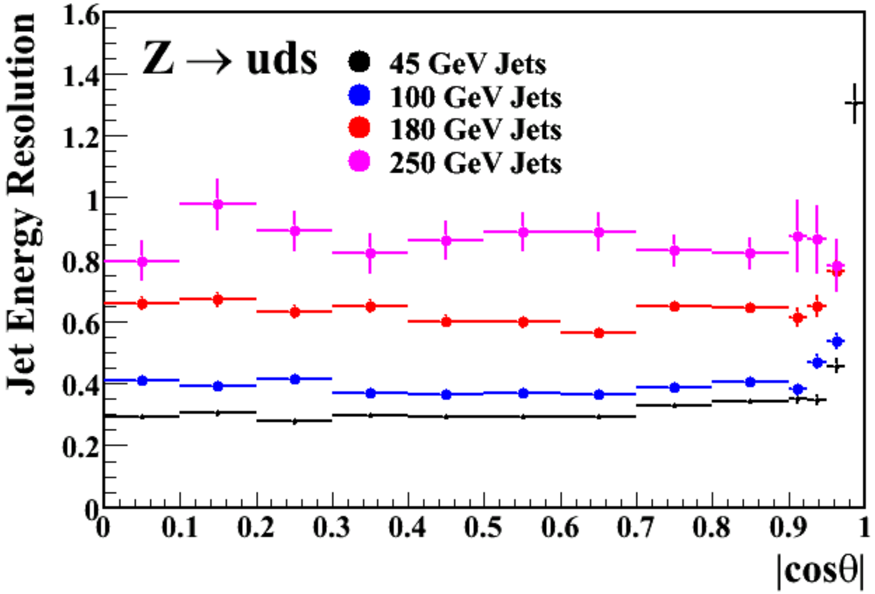
Only weak B-field dependence

★ **BUT still Z at 91.2 GeV**

Jet Energy Dependence

★ Look at $Z \rightarrow uds$ at $\sqrt{s} > 91.2$ GeV
★ LDC00 detector model

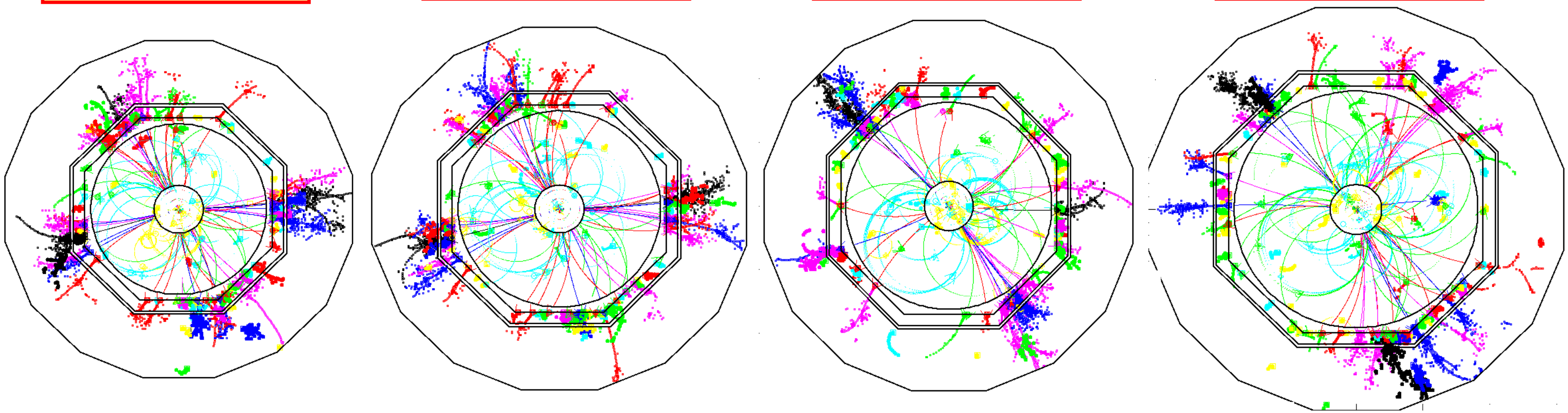
E_{JET}	$\sigma_E/E = \alpha\sqrt{(E/GeV)}$ All angles	$\sigma_E/E = \alpha\sqrt{(E/GeV)}$ $\lvert\cos\theta\rvert<0.7$
45 GeV	33.4±0.3%	29.2±0.4 %
100 GeV	42.0±0.3 %	38.4±0.5 %
180 GeV	71.7±0.3 %	63.8±0.4 %
250 GeV	90.7±2.0 %	87.2±2.5 %



★ Rapid degradation of performance with increasing jet energy
★ However, for 100 GeV jets not bad
★ At ILC typically interested in 6 fermion final states
★ Current performance probably OK for physics studies at $\sqrt{s}$ =500 GeV
★ Probably not yet good enough for $\sqrt{s}$ = 1 TeV

$e^+e^- \rightarrow t\bar{t} \rightarrow$ 6 jets at $\sqrt{s}$=500 GeV

r_{TPC} = 1380 mm | r_{TPC} = 1580 mm | r_{TPC} = 1690 mm | r_{TPC} = 1890 mm



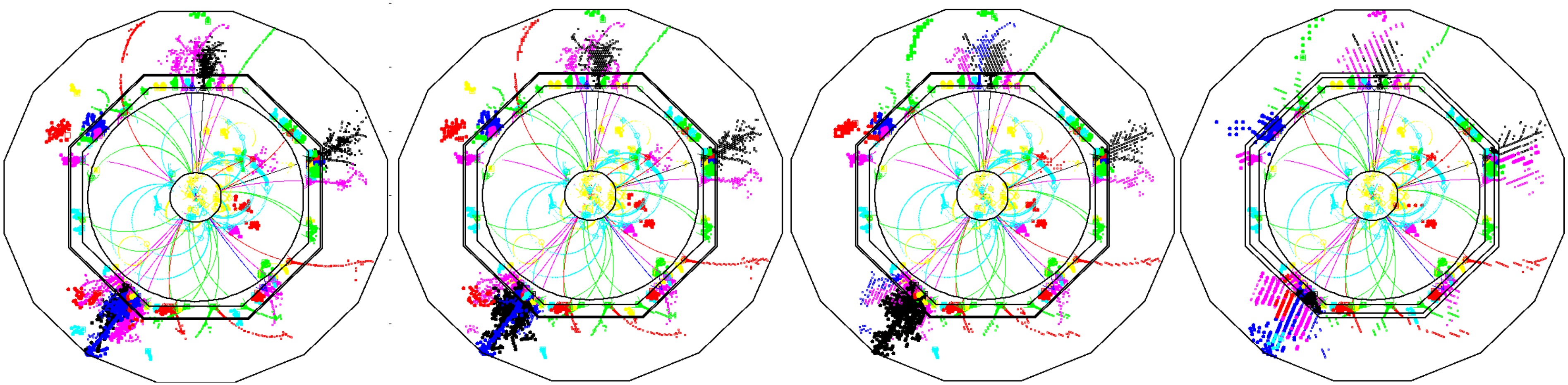
★compare raw resolutions; + add in ν; + add energy lost in forward region

Detector Model	$\sigma_E/E = \alpha\sqrt{(E/GeV)}$		
	E_{RECO}	$+ E_{\nu}$	$+E_{FWD}$
LDC01Sc r_{tpc} = 1380mm	89 ± 2 %	61 ± 1 %	56 ± 1 %
LDC01Sc r_{tpc} = 1580mm	83 ± 2 %	56 ± 1 %	52 ± 1 %
LDC00Sc r_{tpc} = 1690mm	76 ± 2 %	48 ± 1 %	45 ± 1 %
LDC00Sc r_{tpc} = 1890mm	75 ± 2 %	46 ± 1 %	42 ± 1 %

- ★ Fairly strong dependence of performance on Radius
- ★ Discontinuity in going from LDC00 → LDC01 (alg. tuned on LDC00)
- ★ "$+E_{FWD}$" : imperfect accounting of lost energy in FWD region

HCAL Granularity

1x1 3x3 5x5 10x10



Detector Model	$\sigma_{Evis}/E = \alpha\sqrt{(E/GeV)}$		
	Z @91 GeV	tt@500 GeV	Z@500GeV
LDC00Sc 1cm x 1cm	31.4 ± 0.3 %	42 ± 1 %	81 ± 2 %
LDC00Sc 3cm x 3cm	30.6 ± 0.3 %	45 ± 1 %	88 ± 2 %
LDC00Sc 5cm x 5cm	31.3 ± 0.3 %	48 ± 1 %	94 ± 2 %
LDC00Sc 10cm x 10cm	33.7 ± 0.3 %	56 ± 1 %	114 ± 2 %

- ★ **10x10 too coarse (can be seen clearly from display)**
- ★ **Finer granularity helps somewhat at higher energies – why ?**

❹ Remaining Issues

Outstanding questions regarding high energy performance:

- ★ **Above Z at 91.2 GeV : leakage out of HCAL is clearly important**
  - ★ **need to study how important**
    - ★ **need MC samples with extended HCAL (e.g. 40→100 layers)**
  - ★ **how much can be recovered with muon chambers**
    - ★ **need muon chambers in simulation**
- ★ **Energy calibration : currently imperfect**
  - ★ **clear problems with EM scale for clusters > 50 GeV**
- ★ **Photon ID : current cut based selection could be improved**
- ★ **Algorithmic Improvements : still a few features where algorithm fails to spot something obvious**
- ★ **Tuning : current algorithm semi-tuned for 45 GeV jets**
  - ★ **Currently try to optimise parameters for > 100 GeV jets**

Of the above issues – leakage from HCAL is the most pressing as it conceals true high energy performance.

❺ Outlook – towards a first release

- Happy with basic approach in PandoraPFA
- Performance is good for jets below 100 GeV
  - useable for physics studies at $\sqrt{s}$ = 500 GeV
- Still a few things to address, but nothing major…
  - … except maybe photon ID
- Need to address HCAL leakage to progress much further
- Need to tune/optimise for LDC !
  - have currently failed to get decent peformance from LDC detector (algorithm or …?)
- Code is a mess – many unused/obsolete functions still exist

PLEDGE

First version to be released in MarlinReco before end of July 2006