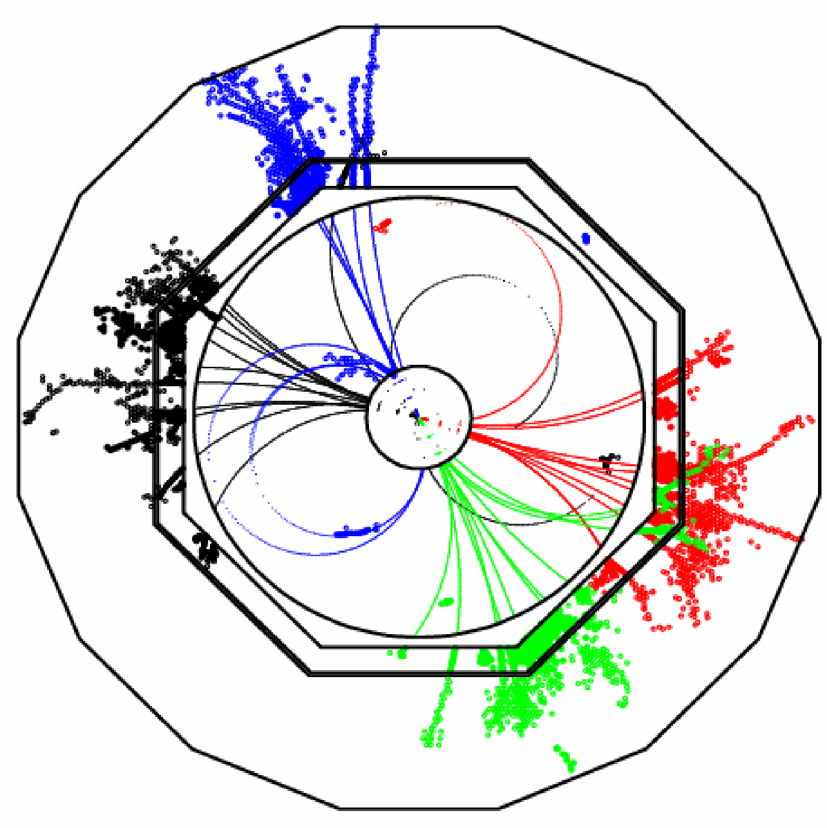
Particle Flow Calorimetry

Mark Thomson
University of Cambridge



This Talk:

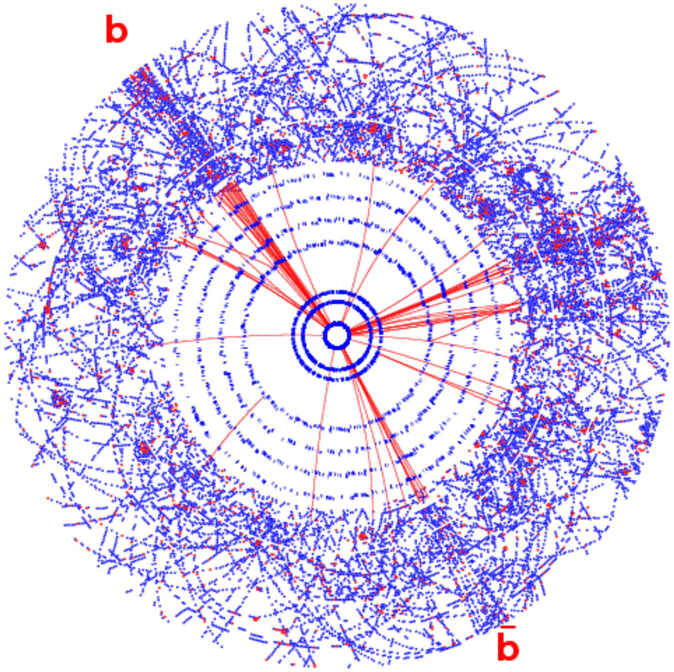
❶ e^+e^- Collider Physics ↔ Calorimetry
❷ The Particle Flow Paradigm
❸ Calorimetry in the ILD Detector Concepts
❹ The PandoraPFA Particle Flow Algorithm
❺ Understanding Particle Flow
❻ Optimisation of a P. Flow detector
❼ Potential at CLiC
❽ Conclusions

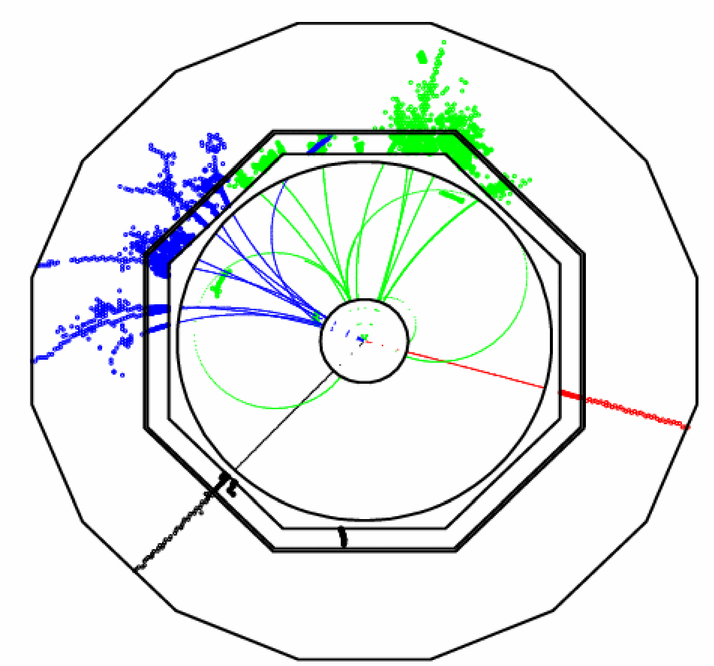
➊ e^+e^- Physics ↔ Calorimetry

★ **Electron-positron colliders provide clean environment for precision physics**

The LHC $\mathrm{pp} \rightarrow \mathrm{H} + X$

The ILC $\mathrm{e^+e^-} \rightarrow \mathrm{HZ}$





★ **A detector at a future lepton collider (ILC/CLiC) will be designed to take full advantage of this clean environment**

★ **Very different detector design requirements c.f. LHC**

e.g. ILC Physics

ILC PHYSICS:

Precision Studies/Measurements

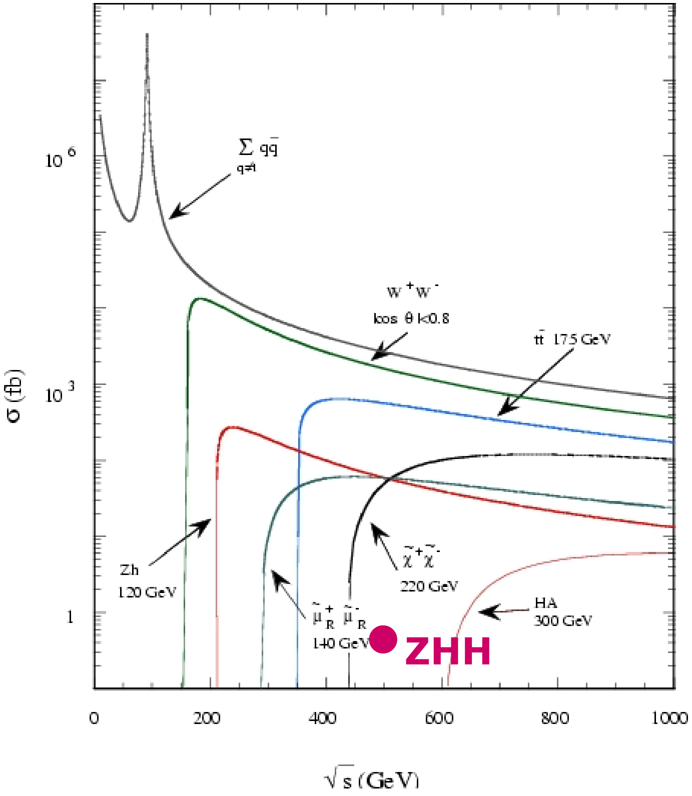
- ★ Higgs sector
- ★ SUSY particle spectrum (if there)
- ★ SM particles (e.g. W-boson, top)
- ★ and much more...

Physics characterised by:

★ High Multiplicity final states
often 6/8 jets

★ Small cross-sections, e.g.

$$\sigma(e^+e^- \to \mathrm{ZHH}) = 0.3\,\mathrm{fb}$$



✸ **Require High Luminosity – i.e. the ILC/CLiC**

✸ **Detector optimized for precision measurements in difficult multi-jet environment**

Compare with LEP

At LEP:

★ Signal dominates: $e^+e^- \rightarrow Z$ and $e^+e^- \rightarrow W^+W^-$

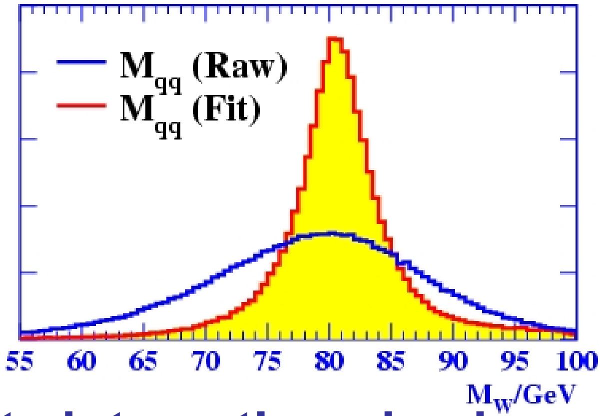
backgrounds not too problematic

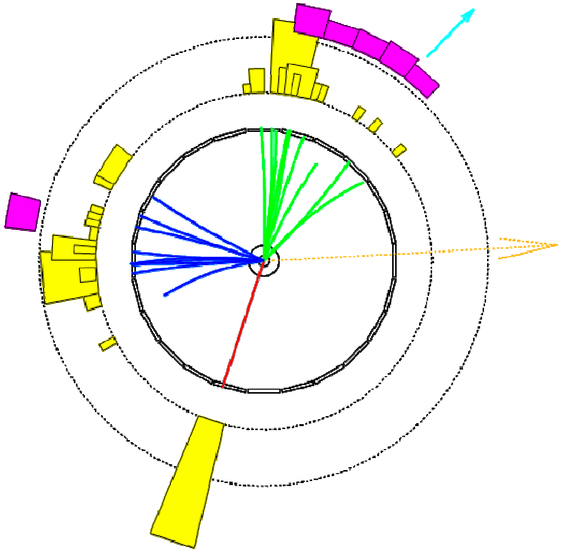
★ Even for W mass measurement, **jet energy resolution not too important**

Kinematic Fits

$$\sum E_i = \sqrt{s}$$

$$\sum \vec{p}_i = 0$$





At the ILC:

★ Backgrounds dominate interesting physics

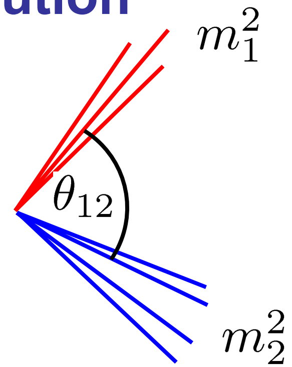
★ Kinematic fitting much less useful: Beamsstrahlung + many final states with > 1 neutrino

✸ Physics performance depends critically on the detector performance (not true at LEP)

✸ Places stringent requirements on the ILC detector

ILC Calorimetry Goals

- ★ Aim for jet energy resolution giving di-jet mass resolution similar to Gauge boson widths



- ★ For a pair of jets have:

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

- ★ For di-jet mass resolution of order $\Gamma_{\mathrm{W/Z}}$

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

⇒ $\boxed{\sigma_{E_j}/E_j < 3.8\%}$

+ term due to θ_{12} uncertainty

- ★ Assuming a single jet energy resolution of normal form

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

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

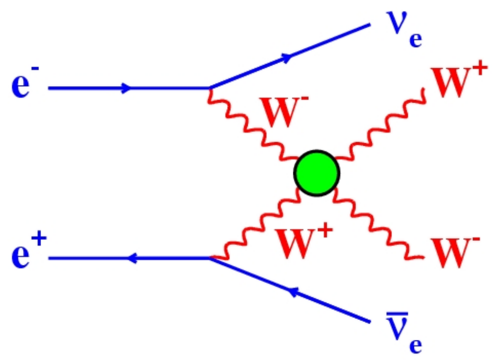
⇒ $\boxed{\alpha(E_j) < 0.027\sqrt{E_{jj}(\mathrm{GeV})}}$

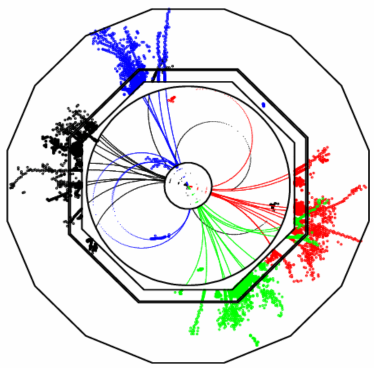
E_{jj}/GeV	$\alpha(E_{jj})$
100	< 27 %
200	< 38 %

- ★ Typical di-jet energies at **ILC** (100-300 GeV) suggests jet energy resolution goal of $\sigma_E/E < 0.30/\sqrt{E_{jj}(\mathrm{GeV})}$

Why is this important ?

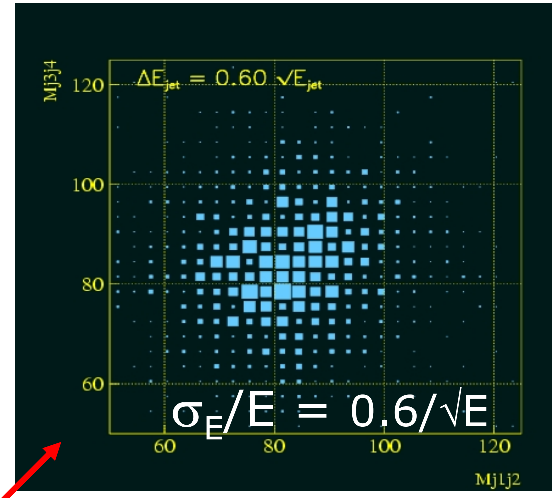
★ **Direct impact on physics sensitivity, e.g. "WW-scattering"**

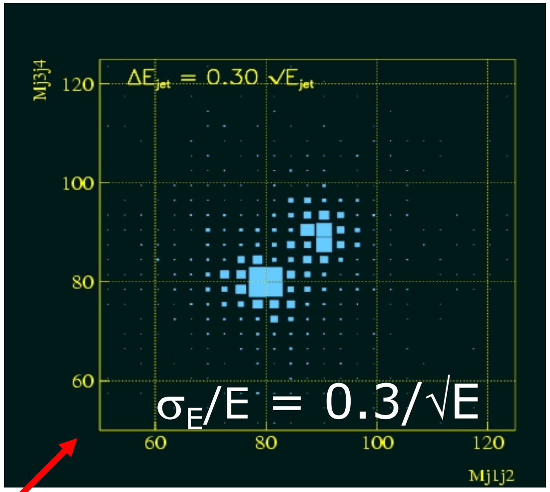




If the Higgs mechanism is not responsible for EWSB then WW fusion processes important

$e^+e^- \to \nu\nu WW \to \nu\nu qqqq$,
$e^+e^- \to \nu\nu ZZ \to \nu\nu qqqq$





Reconstruction of two di-jet masses allows discrimination of WW and ZZ final states

Best at LEP (ALEPH):
$\sigma_E/E = 0.6\,(1+|\cos\theta_{Jet}|)\,/\sqrt{E(\mathrm{GeV})}$

ILC GOAL:
$\sigma_E/E = 0.3/\sqrt{E(\mathrm{GeV})}$

★**Want** $\sigma_E/E < 0.30/\sqrt{E(\mathrm{GeV})}$

or more correctly

$\sigma_E/E < 3.8\%$

Remember this number

★**Very hard (may not be possible) to achieve this with a traditional approach to calorimetry**

Limited by typical HCAL resolution of > 50%/√E(GeV)

⇒ **a new approach to calorimetry**

Particle Flow

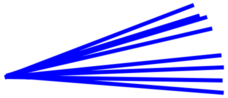
Dual Readout

Within the ILC community, widely believed to be most promising approach

❷ The Particle Flow Paradigm

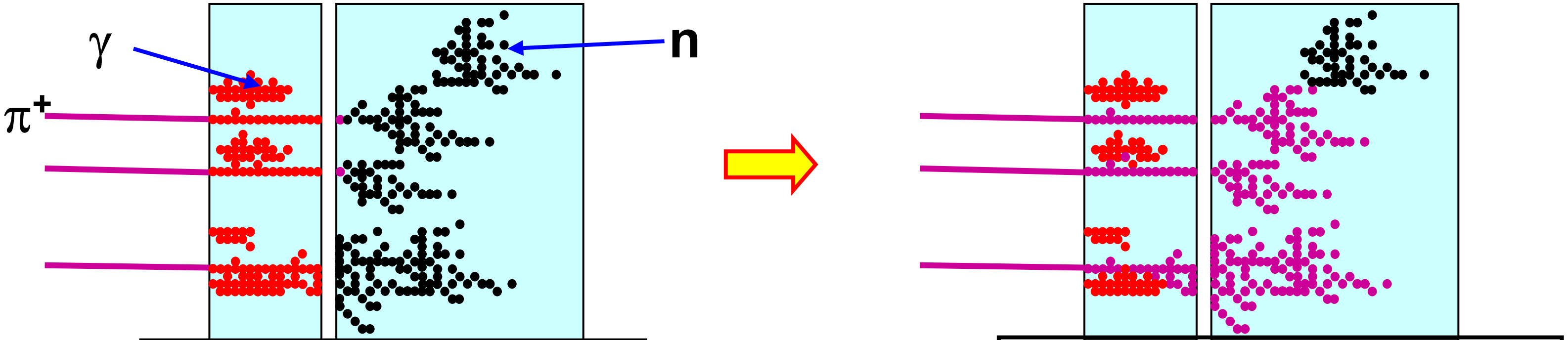
★ **In a typical jet :**

- 60 % of jet energy in charged hadrons
- 30 % in photons (mainly from $\pi^0 \rightarrow \gamma\gamma$)
- 10 % in neutral hadrons (mainly n and K_L)



★ **Traditional calorimetric approach:**

- Measure all components of jet energy in ECAL/HCAL !
- ~70 % of energy measured in HCAL: $\sigma_{\mathrm{E}}/\mathrm{E} \approx 60\,\%/\sqrt{\mathrm{E(GeV)}}$
- Intrinsically "poor" HCAL resolution limits jet energy resolution



$\mathrm{E_{JET} = E_{ECAL} + E_{HCAL}}$

$\mathrm{E_{JET} = E_{TRACK} + E_{\gamma} + E_{n}}$

★ **Particle Flow Calorimetry paradigm:**

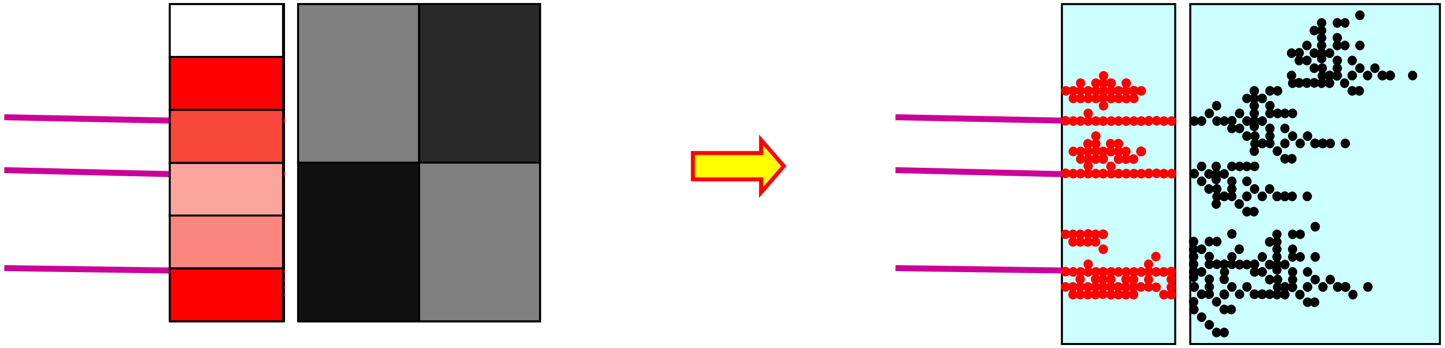
- charged particles measured in tracker (essentially perfectly)
- Photons in ECAL: $\sigma_{\mathrm{E}}/\mathrm{E} < 20\,\%/\sqrt{\mathrm{E(GeV)}}$
- Neutral hadrons (ONLY) in HCAL
- Only 10 % of jet energy from HCAL ⇨ much improved resolution

Particle Flow Calorimetry

Hardware:

★ Need to be able to resolve energy deposits from different particles

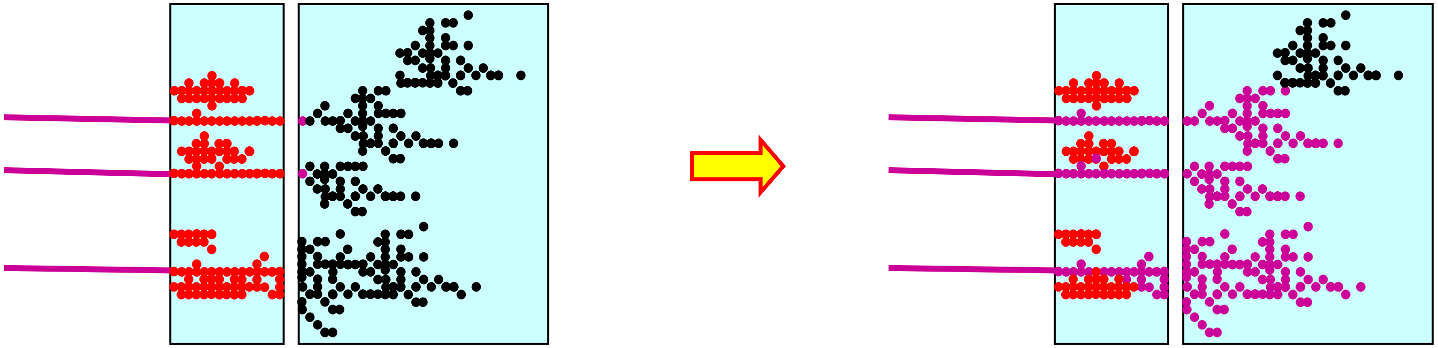
→ Highly granular detectors (as studied in CALICE)



Software:

★ Need to be able to identify energy deposits from each individual particle !

→ Sophisticated reconstruction software



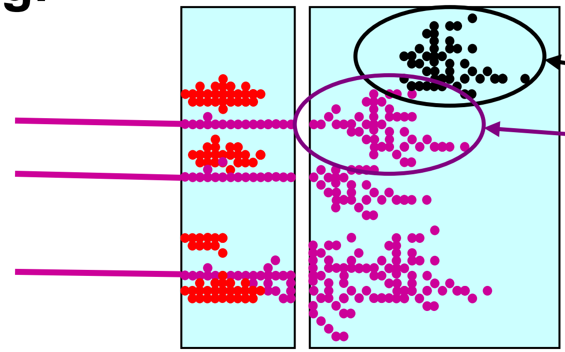
★ **Particle Flow Calorimetry = HARDWARE + SOFTWARE**

Particle Flow Algorithms (PFA)

Reconstruction of a Particle Flow Calorimeter:

- ★ **Avoid double counting of energy from same particle**
- ★ **Separate energy deposits from different particles**

e.g.



If these hits are clustered together with these, lose energy deposit from this neutral hadron (now part of track particle) and ruin energy measurement for this jet.

Level of mistakes, "confusion", determines jet energy resolution not the intrinsic calorimetric performance of ECAL/HCAL

sounds easy….

- ★ **PFA performance depends on detailed reconstruction**
- ★ **Relatively new, still developing ideas**
- ★ **Studies need to be based on a sophisticated detector simulations**
- ★ **Need a "strawman" detector design to study potential…**

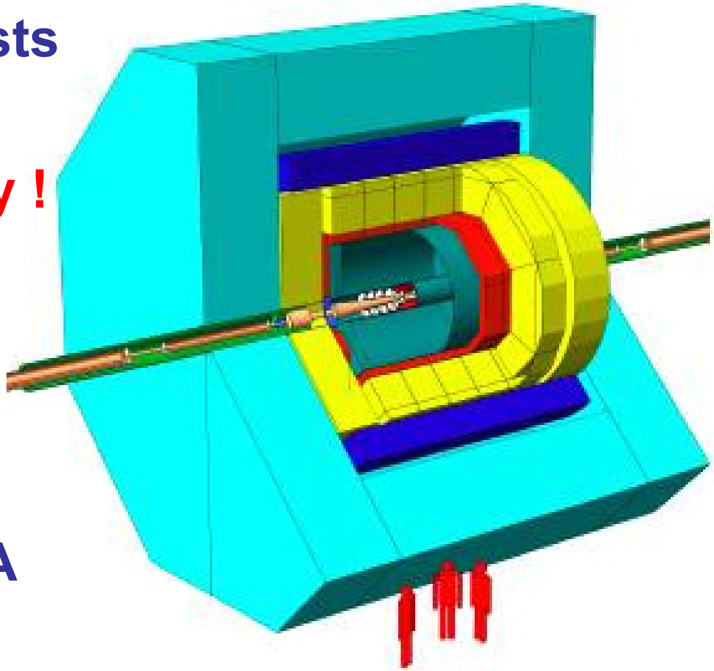
❸ The ILD Detector Concept*

NOTE:

- ★Particle flow reconstruction involves "whole detector"
- ★To study potential performance need a detector model, tracking, calorimeters, …

ILC Detector Concepts:

- ★Here performance of particle flow calorimetry shown in the context of the ILD detector concept for the ILC
- ★ Detailed GEANT 4 detector model exists
- ★ A potential design for an ILC detector
- ★ Designed for Particle Flow calorimetry !

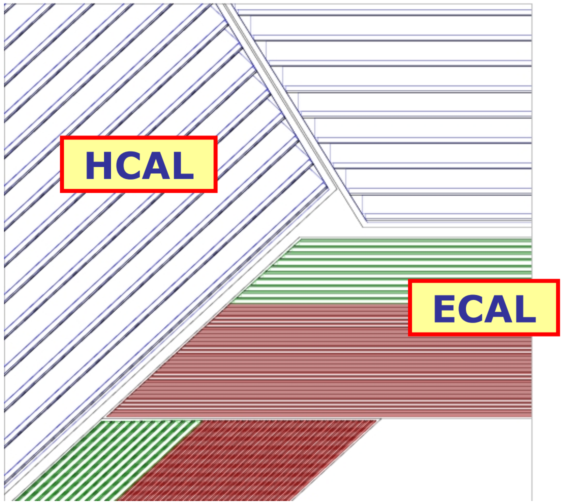


ILD Main Features:

- ▪ Large TPC central tracker (R=1.8 m)
- ▪ CMS like solenoid (B = 3.5 T)
- ▪ ECAL and HCAL inside solenoid
- ▪ ECAL/HCAL highly segmented for PFA

ILD calorimetry concept*

Very high longitudinal and transverse segmentation



ECAL:

- SiW sampling calorimeter
- Tungsten: X_0/λ_{had} = 1/25, $R_{Mol.}$ ~ 9mm
  - → Narrow EM showers
  - → longitudinal sep. of EM/had. showers
- longitudinal segmentation: 30 layers
- transverse segmentation: 5x5 mm^2 pixels

HCAL:

- Steel-Scintillator sampling calorimeter
- longitudinal segmentation: 48 layers (6 interaction lengths)
- transverse segmentation: 3x3 cm^2 scintillator tiles

Comments:

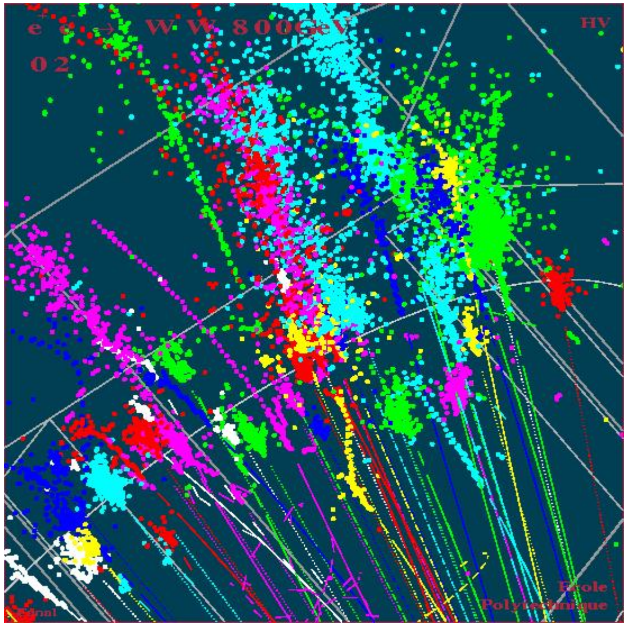
★ Technologically feasible (although not cheap)

★ Ongoing test beam studies (CALICE collaboration)

*Other ILD calorimetry options being actively studied, e.g. RPC DHCAL, Scintillator strip ECAL

Calorimeter Reconstruction

- ★ High granularity calorimeters – **very different** to previous detectors (except LEP lumi. calorimeters)
- ★ "Tracking calorimeter" – requires a new approach to ECAL/HCAL reconstruction



Particle Flow Reconstruction

- ★ PFA calorimetric performance = **HARDWARE** + **SOFTWARE**
- ★ **Performance will depend on the software algorithm**

→ difficult to evaluate full potential

$\sigma_E/E = f\,(\text{software})$

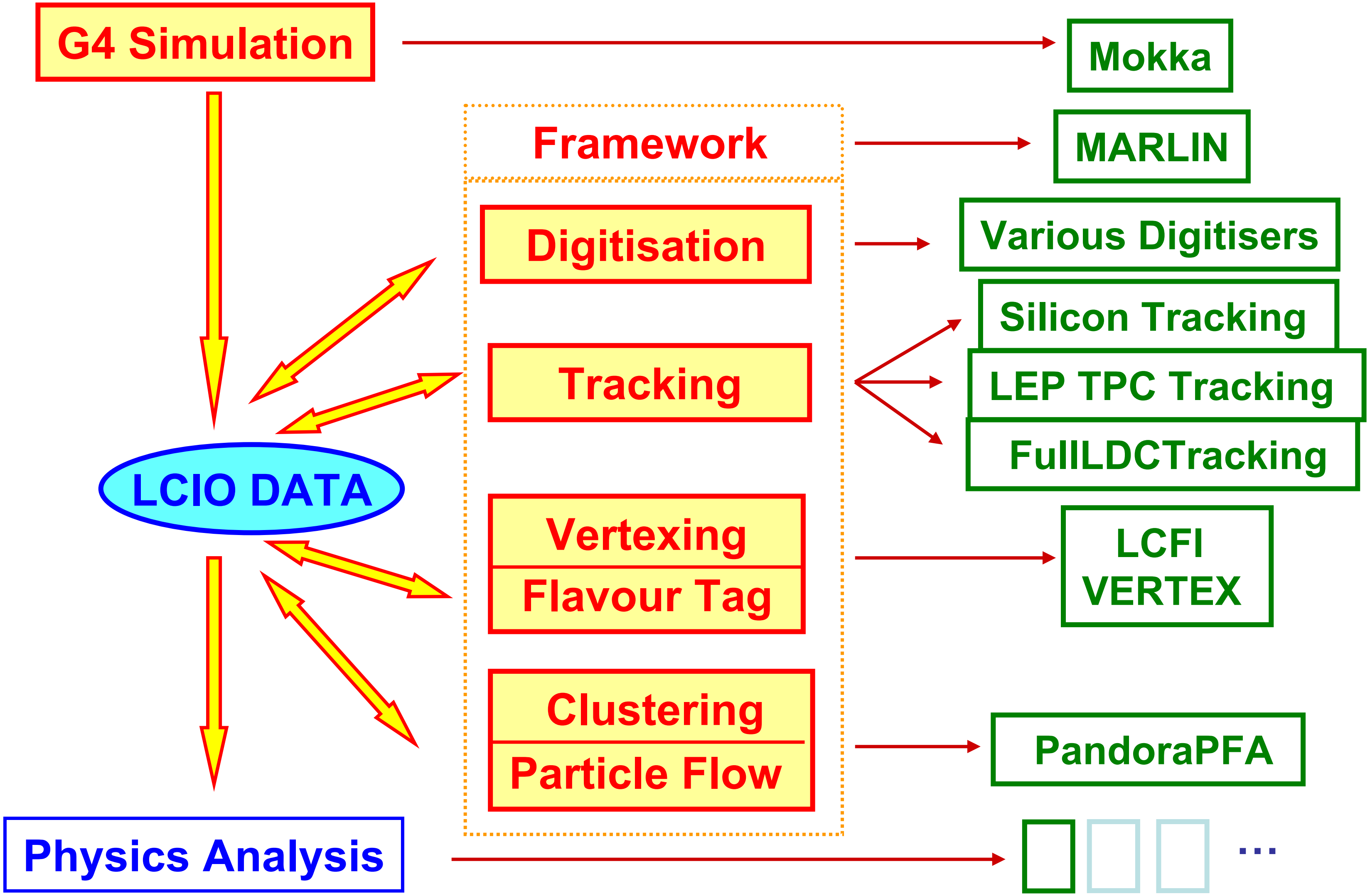
- ★ There are no short cuts, fast simulation doesn't help…

To evaluate Particle Flow Calorimetry at ILC:
need realistic reconstruction chain
>10 years before start of ILC !!!

- ★ But, as a result of a great deal of work within ILC detector community, **we already have a first version…**

ILD Reconstruction Framework (C++)

★ Everything exists – level of sophistication ~LEP experiment



G4 Simulation

Framework

Digitisation

Tracking

LCIO DATA

Vertexing

Flavour Tag

Clustering

Particle Flow

Physics Analysis

Mokka

MARLIN

Various Digitisers

Silicon Tracking

LEP TPC Tracking

FullLDCTracking

LCFI VERTEX

PandoraPFA

…

❹ Particle Flow Reconstruction: PandoraPFA

- Need "realistic" Particle Flow to evaluate potential of method (again no shortcuts)
- New paradigm – nobody really knows how to approach this
- So where are we now ?
- Significant effort in context of ILC detector design (~4 groups developing PFA reconstruction worldwide)

For this talk concentrate on: PandoraPFA

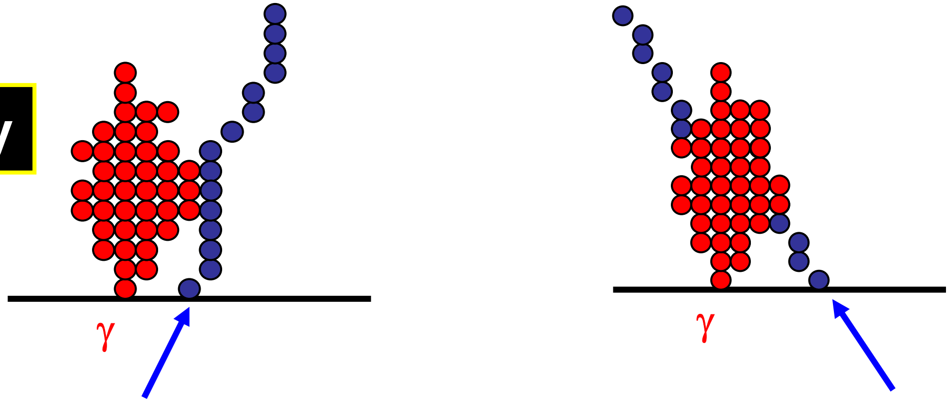
- This is still work-in-Progress – currently it gives the best performance
- Will give an overview of the algorithm to highlight how particle flow reconstruction works

PFA : Basic issues

- ★ **Separate energy deposits** from different particles
- ★ **Avoid double counting of energy** from same particle
- ★ **Mistakes** drive particle flow jet energy resolution

e.g.

- ★ Need to separate “tracks” (charged hadrons) from photons



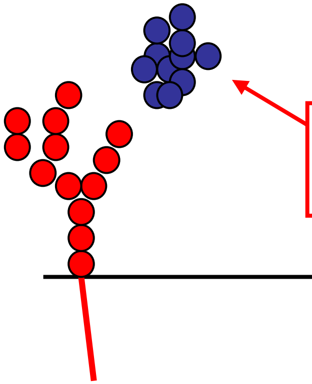
hardware

granularity

software

PFlow Algorithm

- ★ Need to separate neutral hadrons from charged hadrons



Isolated neutral hadron or fragment from shower ?

PandoraPFA Overview

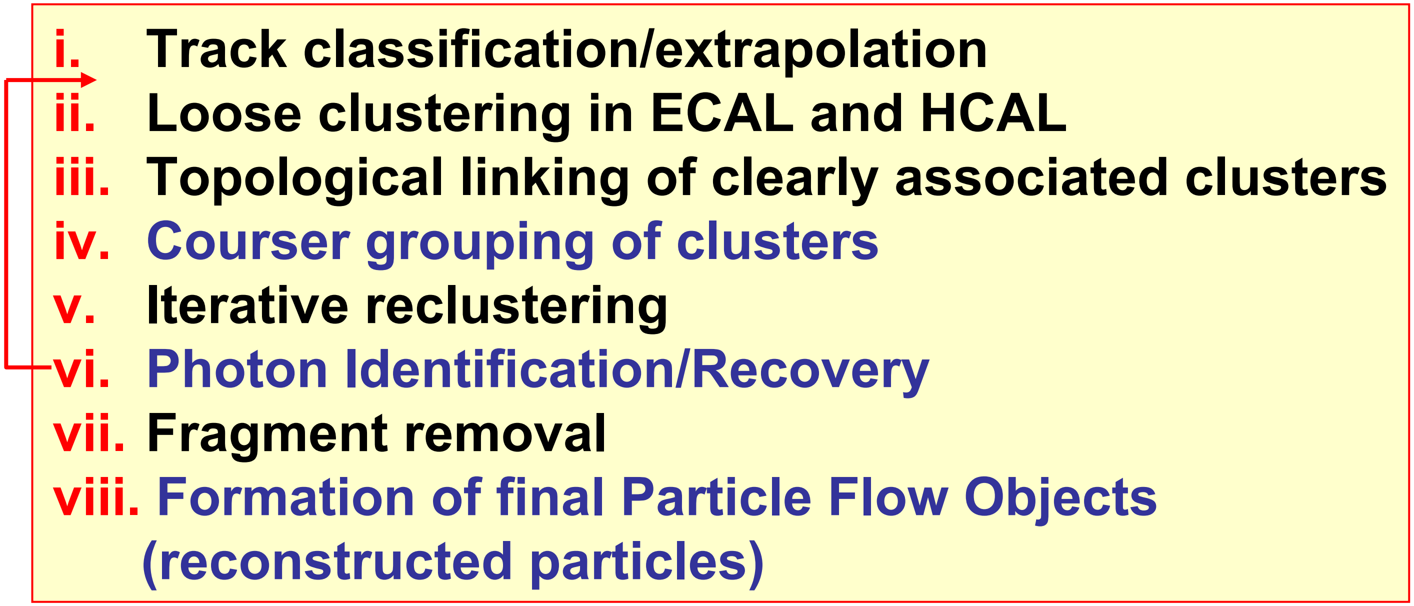
★ ECAL/HCAL reconstruction and PFA performed in a single algorithm

★ Fairly generic algorithm
  - applicable to multiple detector concepts

★ Use tracking information to help ECAL/HCAL clustering

★ This is a sophisticated algorithm : 10^4 lines of code

Eight Main Stages:



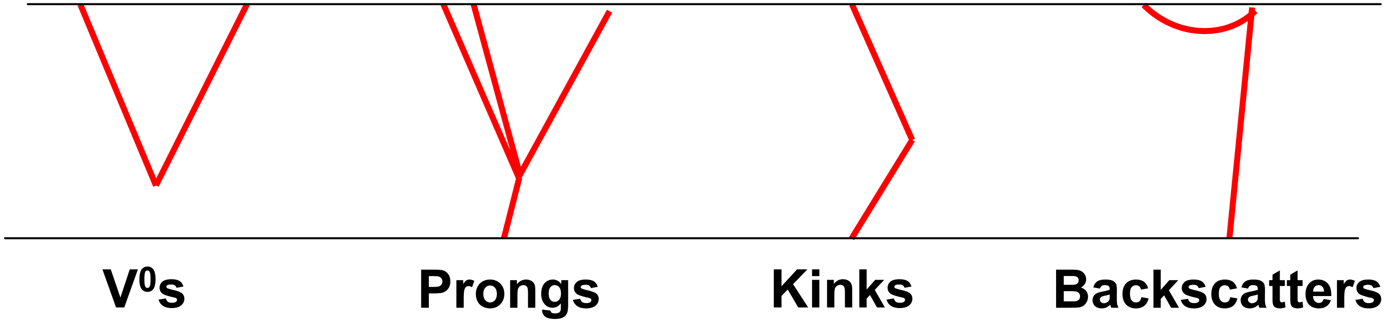
i. Track classification/extrapolation
ii. Loose clustering in ECAL and HCAL
iii. Topological linking of clearly associated clusters
iv. Courser grouping of clusters
v. Iterative reclustering
vi. Photon Identification/Recovery
vii. Fragment removal
viii. Formation of final Particle Flow Objects (reconstructed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

i) Tracking

- ★ The use of optimal use of tracking information in PFA is essential
- ★ Non trivial for looping tracks (even in a TPC)
- ★ Matching of tracks to endcap clusters is non-trivial
- ★ Use of track information is a major part of PandoraPFA
- ★ Big effort to use as many tracks in the event as possible
  - helps particularly for lower energy jets
  - motivation I : better energy resolution
  - motivation II : correct measurement of direction

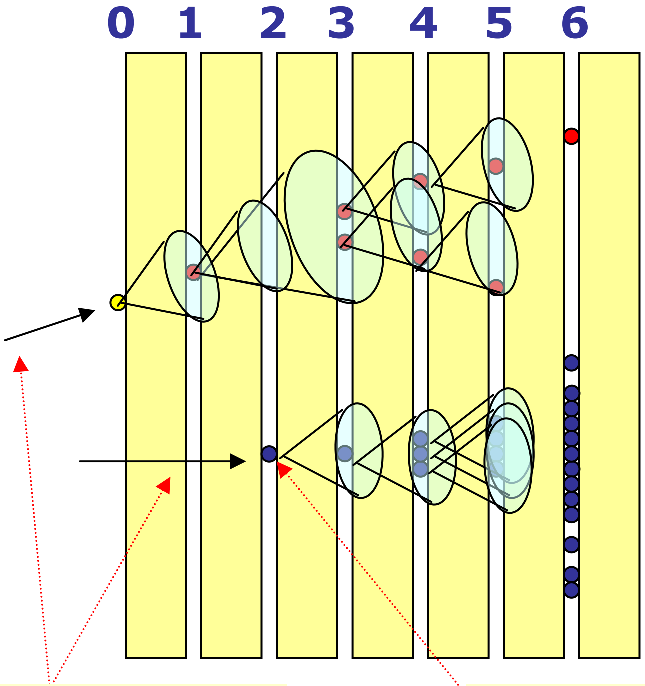
★ **TPC-oriented:** take advantage of pattern recognition capability

★ From fully reconstructed tracks identify:



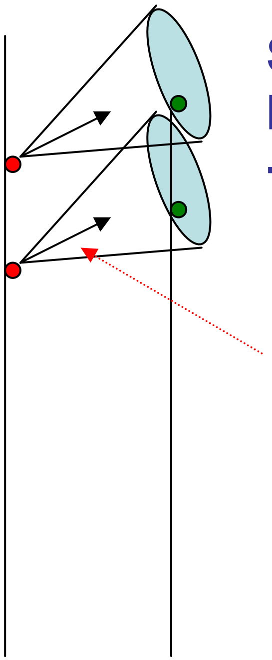
ii) ECAL/HCAL Clustering

- ★ Tracks used to “seed” clusters
- ★ Start at inner layers and work outward
- ★ Associate hits with existing Clusters
- ★ If no association made form new Cluster
- ★ Very simple cone based algorithm



Initial cluster direction

Unmatched hits seeds new cluster



Simple cone algorithm based on current direction + additional N pixels

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

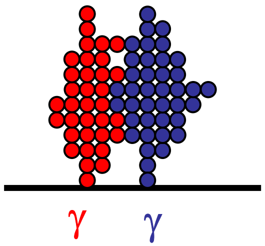
Parameters:
- cone angle
- additional pixels

iii) Topological Cluster Association

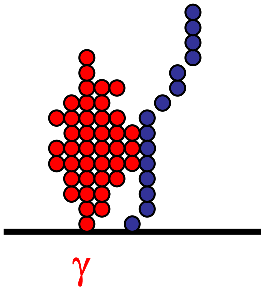
- By design, clustering errs on side of caution
  i.e. clusters tend to be split
- **Philosophy:** easier to put things together than split them up
- Clusters are then associated together in two stages:
  - 1) Tight cluster association – clear topologies
  - 2) Loose cluster association – fix what's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

Photon ID

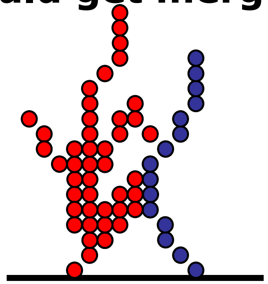
- Photon ID plays important role
- Simple "cut-based" photon ID applied to all clusters
- Clusters tagged as photons are immune from association procedure – just left alone



Won't merge



Won't merge

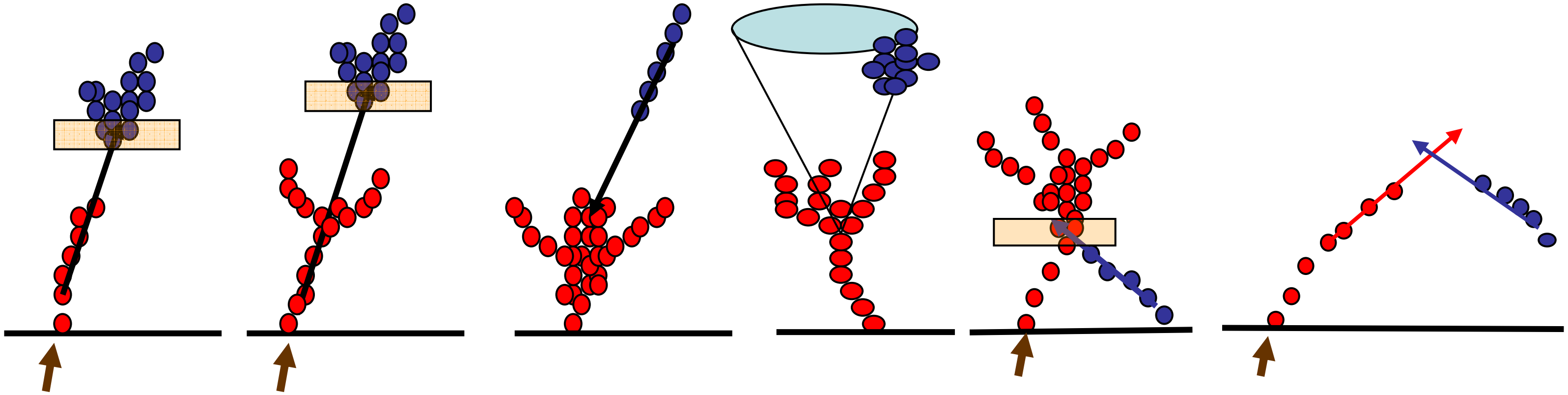


Could get merged

★ Clusters associated using a number of topological rules

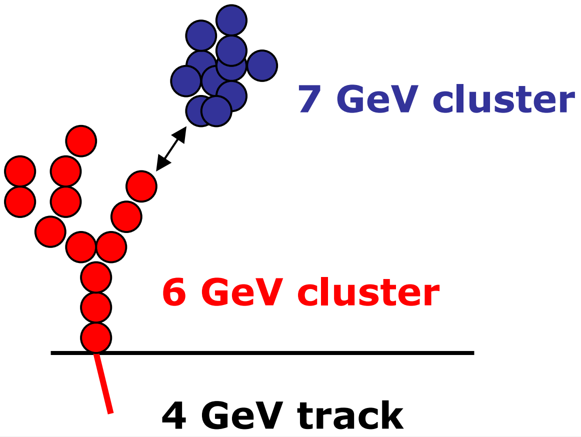
Clear Associations:

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



Less clear associations:

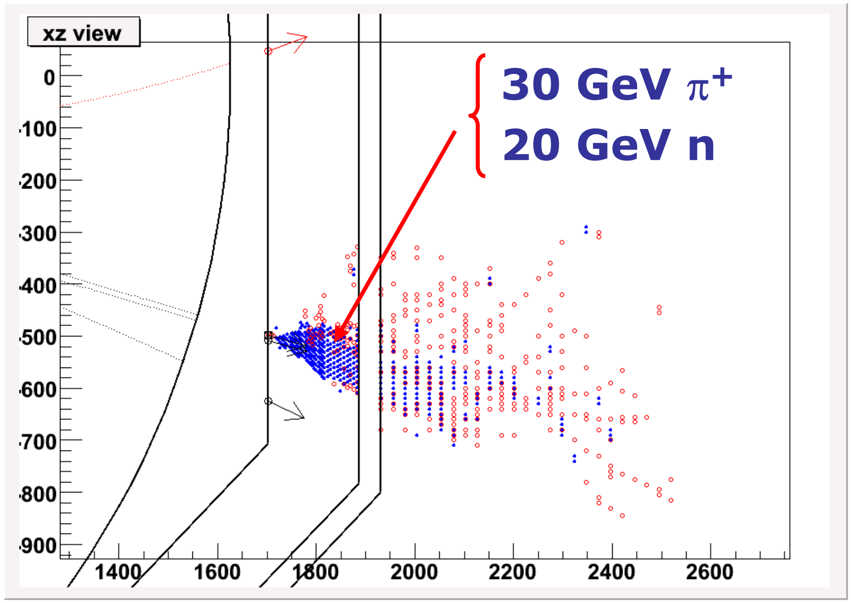
e.g. Proximity



Use E/p consistency to veto clear mistakes

v) Iterative Reclustering

- ★ To this point, performance is good – but some difficult cases…



- ★ At some point reach the limit of “pure” particle flow
  - ♦ i.e. can’t cleanly resolve neutral hadron in hadronic shower

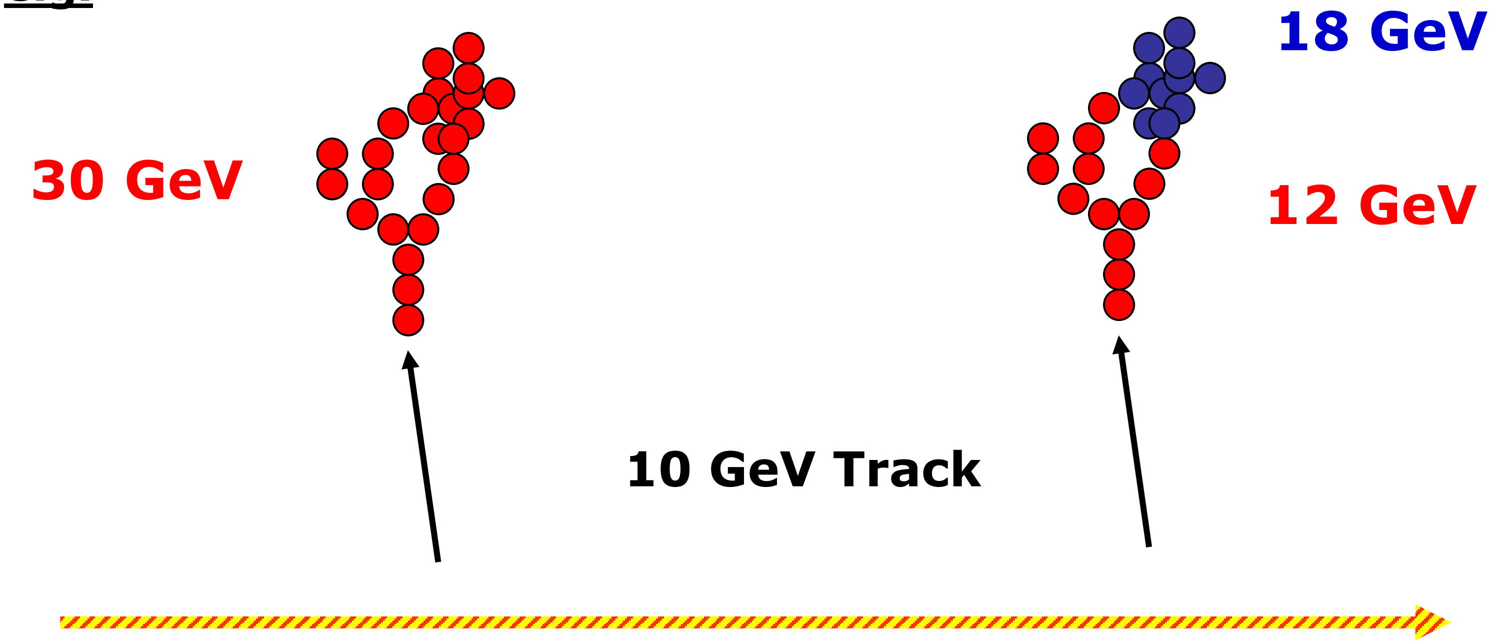
The ONLY(?) way to address this is “statistically”

→

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

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

e.g.



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

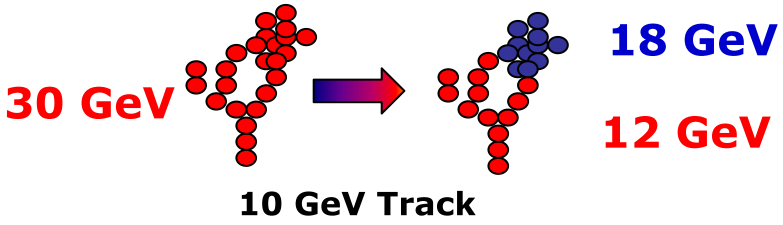
NOTE: clustering driven by track momentum (but not subtraction)

This is very important for higher energy jets

Iterative Reclustering Strategies

① Cluster splitting

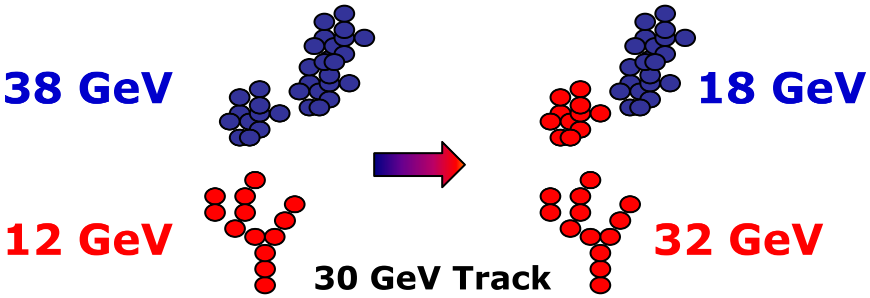
Reapply entire clustering algorithm to hits in "dubious" cluster. Iteratively reduce cone angle until cluster splits to give acceptable energy match to track



★ + plug in alternative clustering algorithms

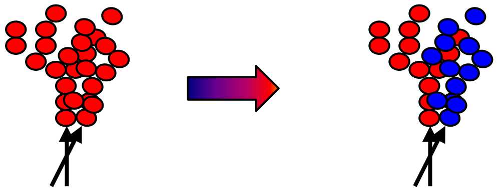
② Cluster merging with splitting

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



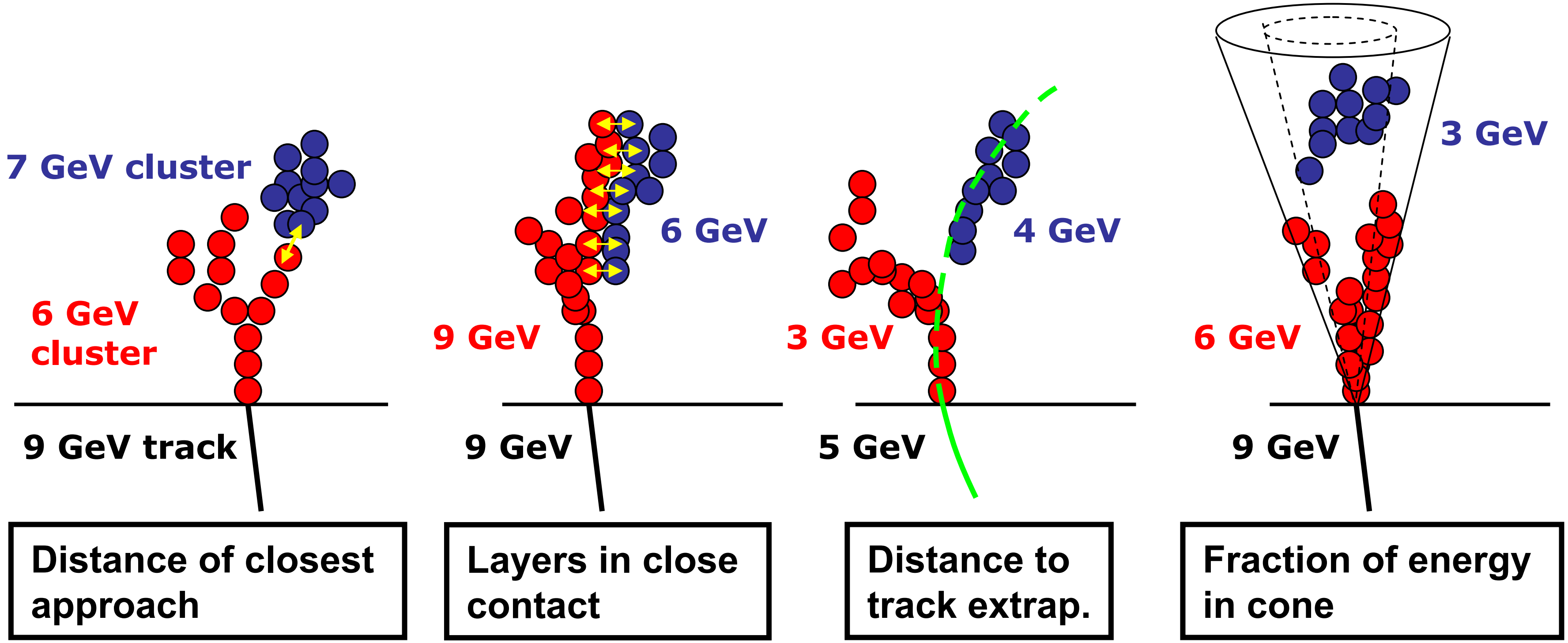
③ Track association ambiguities

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



viii) Fragment removal : basic idea

- Look for "evidence" that a cluster is associated with another



Distance of closest approach

Layers in close contact

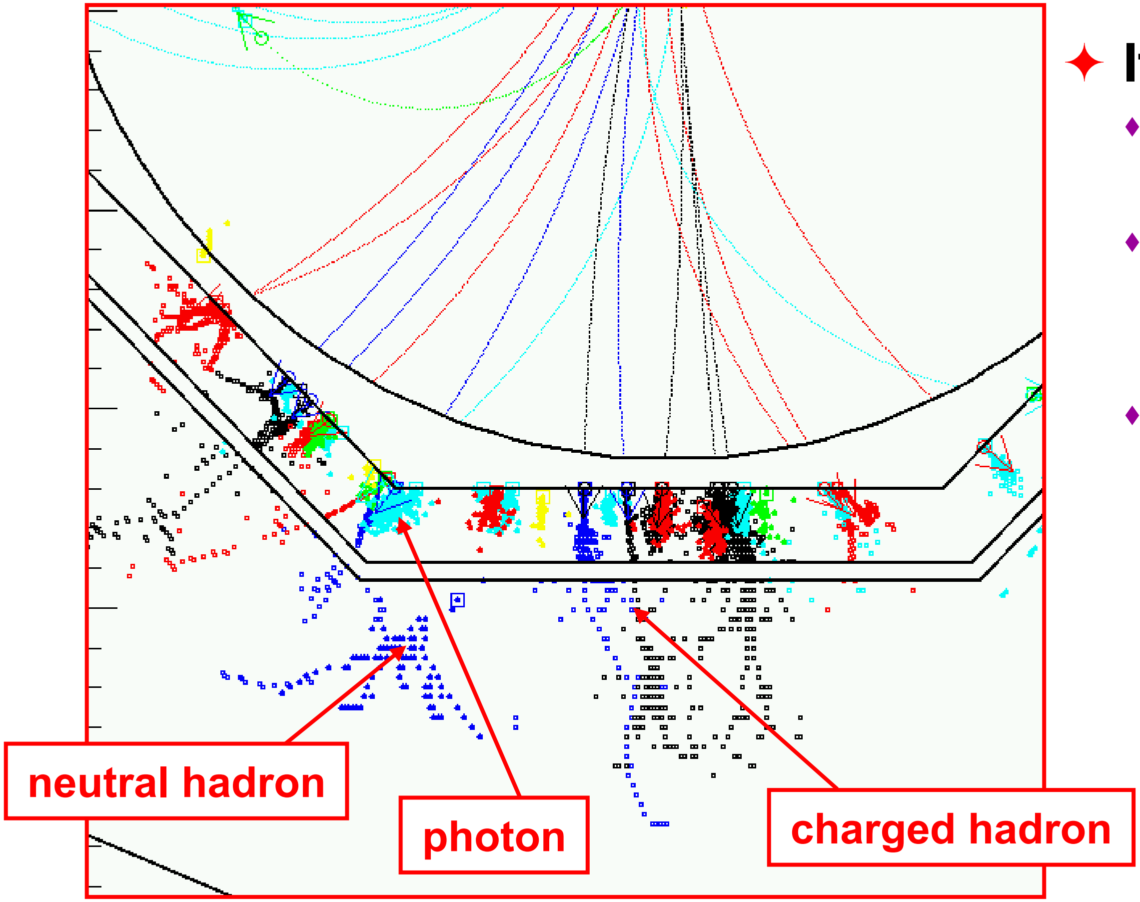
Distance to track extrap.

Fraction of energy in cone

- Convert to a numerical evidence score **E**
- Compare to another score "required evidence" for matching, **R**, based on change in E/p chi-squared, location in ECAL/HCAL etc.
- If **E > R** then clusters are merged
- Rather *ad hoc* but works well – but works well

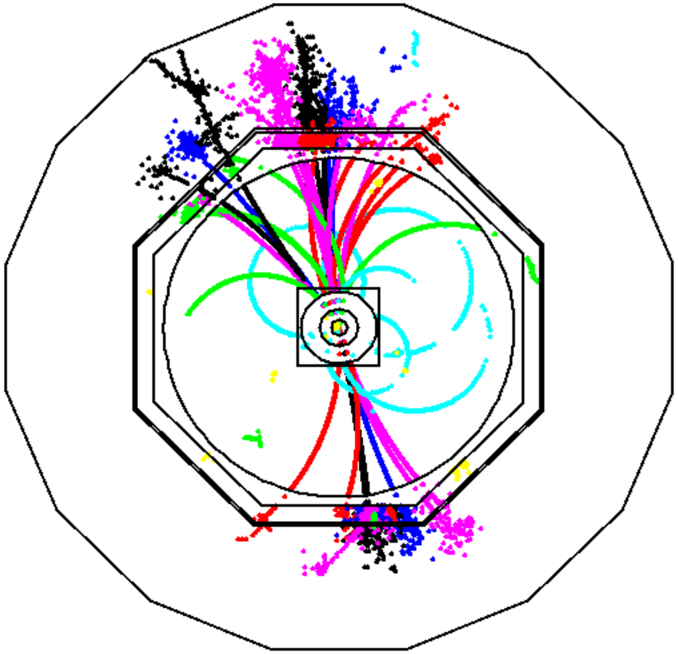
Putting it all together…

100 GeV Jet



- **If it all works…**
  - **Reconstruct the individual particles in the event.**
  - **Calorimeter energy resolution not critical: most energy in form of tracks.**
  - **Level of mistakes in associating hits with particles, dominates jet energy resolution.**

❺ Current performance



- ★Benchmark performance using $Z \rightarrow u\bar{u}$ and $Z \rightarrow d\bar{d}$ events (clean, no neutrinos)
- ★Test at for different energies with Z decays at rest
- ★OPAL tune of Pythia fragmentation
- ★Full reconstruction (track + calo) using no Monte Carlo "cheat" information

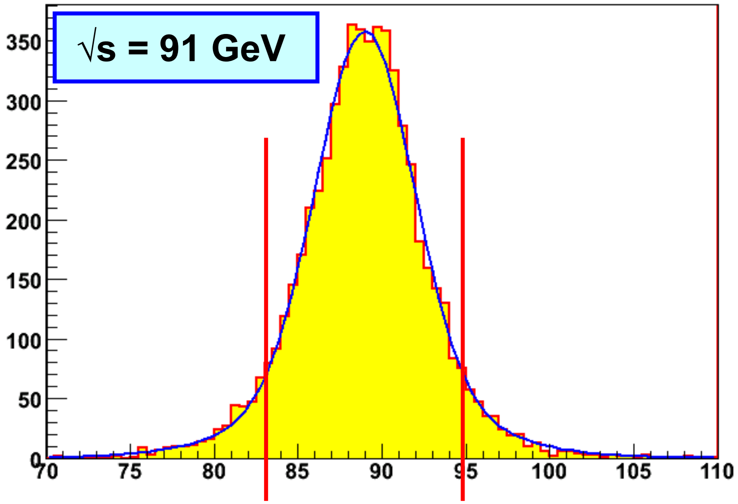
NOTE:

- Quoting rms of reconstructed energy distribution is misleading
- Particle Flow occasionally goes very wrong → tails dominate rms
- Conventional to measure performance using **rms90** which is relatively insensitive to tails

Figures of Merit:

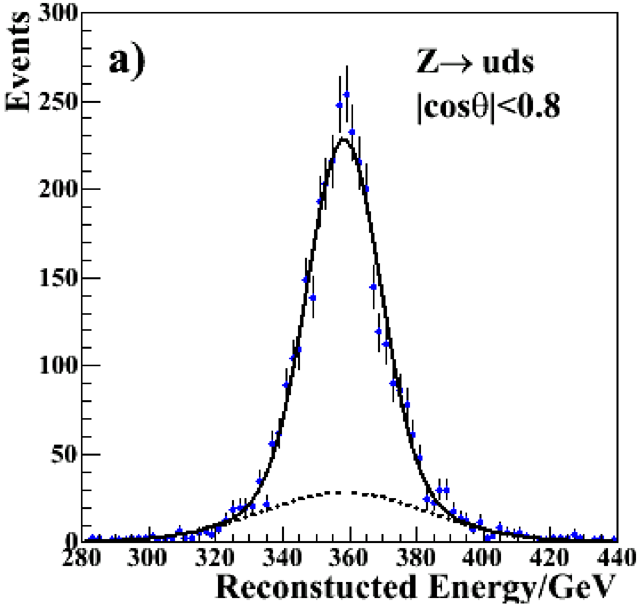
rms_{90}

- Find smallest region containing 90 % of events
- Determine rms in this region



σ_{80}

- Fit sum of two Gaussians with same mean. The narrower one is constrained to contain 80 % of events
- Quote σ of narrow Gaussian



It turns out that $\mathrm{rms}_{90} \approx \sigma_{80}$

(need care when comparing to Gaussian resolution)

Performance (ILD) $Z \rightarrow d\bar{d}, Z \rightarrow u\bar{u}$

rms90 **PandoraPFA v03-β**

E_{JET}	$\sigma_E/E = \alpha/\sqrt{E_{jj}}$ $\lvert\cos\theta\rvert<0.7$	σ_E/E_j
45 GeV	23.8 %	3.5 %
100 GeV	29.1 %	2.9 %
180 GeV	37.7 %	2.8 %
250 GeV	45.6 %	2.9 %

- Full G4 simulation
- Full reconstruction

★ Particle flow achieves ILC goal of $\sigma_E/E_j < 3.8\ \%$

★ For lower energy jets Particle Flow gives unprecedented levels of performance, e.g. @ 45 GeV : 3.5% c.f. ~10% (ALEPH)

★ "Calorimetric" performance (α) degrades for higher energy jets

★ Current PFA code is not perfect – lower limit on performance

Proof of principle:

PARTICLE FLOW CALORIMETRY WORKS

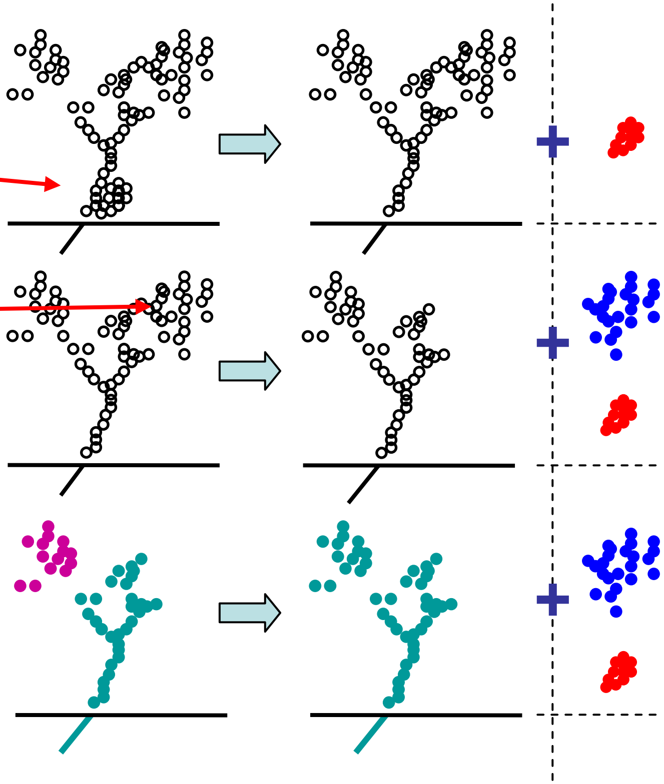
At least in simulation

Understanding PFA Performance

What drives Particle Flow performance ?

- ★ Try to use various "Perfect PFA" algorithms to pin down main performance drivers (resolution, confusion, …)
- ★ Use MC to "cheat" various aspects of **Particle Flow**

PandoraPFA options:



- **PerfectPhotonClustering**
  hits from photons clustered using MC info and removed from main algorithm
- **PerfectNeutralHadronClustering**
  hits from neutral hadrons clustered using MC info…
- **PerfectFragmentRemoval**
  after PandoraPFA clustering "fragments" from charged tracks identified from MC and added to charged track cluster
- **PerfectPFA**
  perfect clustering and matching to tracks

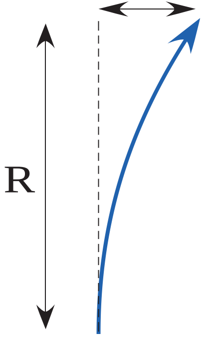
Contribution	σ_E/E			
	45 GeV	100 GeV	180 GeV	250 GeV
Calo. Resolution	3.1 %	2.1 %	1.5 %	1.3 %
Leakage	0.1 %	0.5 %	0.8 %	1.0 %
Tracking	0.7 %	0.7 %	1.0 %	0.7 %
Photons “missed”	0.4 %	1.2 %	1.4 %	1.8 %
Neutrals “missed”	1.0 %	1.6 %	1.7 %	1.8 %
Charged Frags.	1.2 %	0.7 %	0.4 %	0.0 %
“Other”	0.8 %	0.8 %	1.2 %	1.2 %

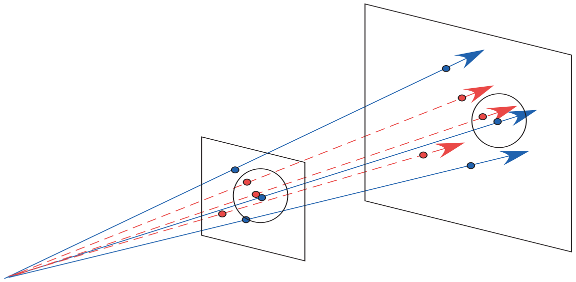
Comments:

- ★ For 45 GeV jets, jet energy resolution dominated by ECAL/HCAL resolution
- ★ Track reco. not a large contribution (Reco ≈ CheatedTracking)
- ★ “Satellite” neutral fragments not a large contribution
    - ▪ efficiently identified
- ★ Leakage only becomes significant for high energies
- ★ Missed neutral hadrons dominant confusion effect
- ★ Missed photons, important at higher energies

❻ Optimisation of a Particle Flow detector

- ★ Particle Flow Calorimetry lives or dies on ability to separate energy deposits from individual particles.
  - **Large detector** – spatially separate particles
  - **High B-field** – separate charged/neutrals
  - **High granularity ECAL/HCAL** – resolve particles





Might expect "figure-of-merit":

$$\frac{BR^2}{\sigma}$$

Separation of charge/neutrals

Calorimeter granularity/$R_{Moliere}$

- ★ Argues for: large + high granularity + ⇧ B
- ★ Cost considerations: small + lower granularity + ⇩ B

Optimise detector parameters using PandoraPFA

Interpretation: observing effects of detector + imperfect software

e.g. Radius vs. B-field

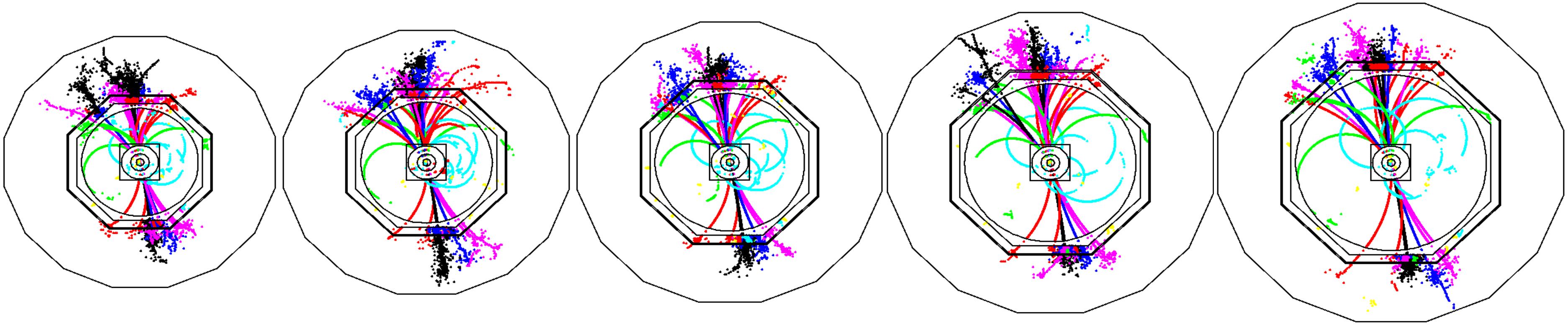
Cost drivers:

- For Particle flow, ECAL and HCAL inside Solenoid
- Calorimeters and solenoid are the main cost drivers of an ILC detector optimised for particle flow
- Cost of calorimeters scales with active area
- Cost of solenoid scales with stored energy, (very approx.)

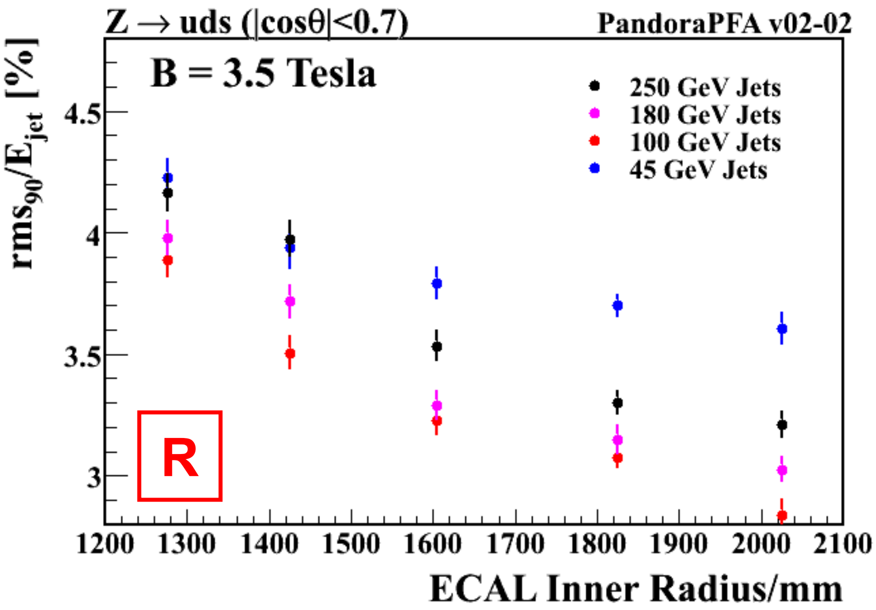
$$\$\$\$ \propto \left(B^2R^2L\right)^{0.66}$$

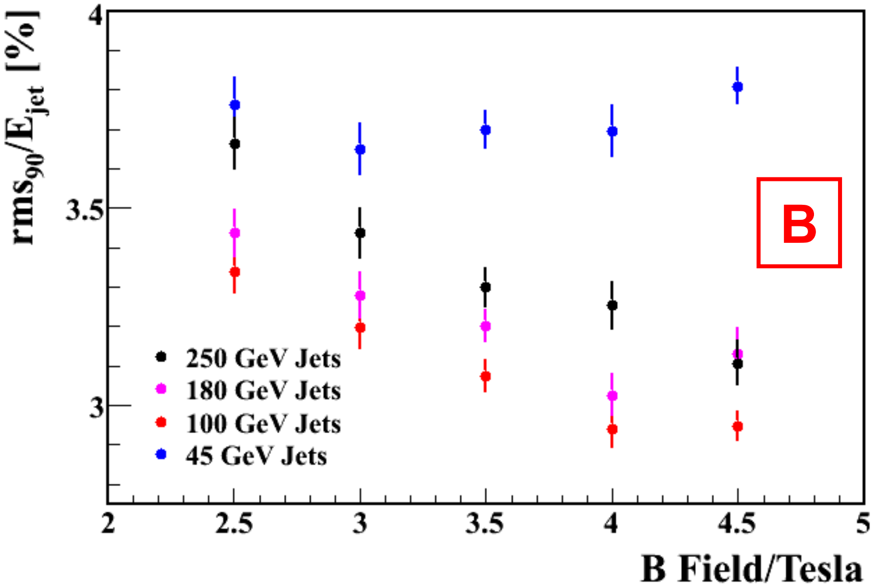
★TPC radius and B-field play major role in total detector cost

  - Study jet energy resolution as a function of **B** and **R**



e.g. vary ECAL Inner Radius





★ **Empirically (for current algorithm) find**

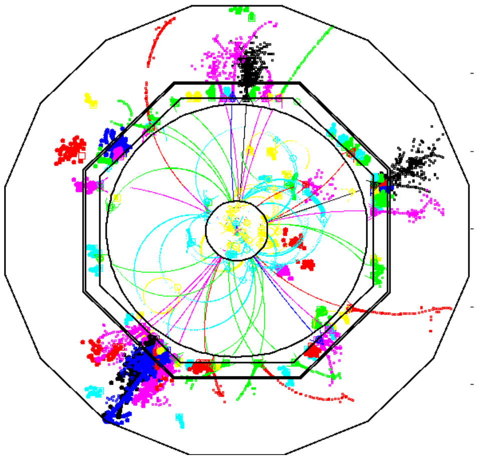
$$\frac{\sigma_E}{E} = \underbrace{\frac{0.021}{\sqrt{E}}}_{\text{Resolution}} \oplus \underbrace{0.01}_{\text{Tracking/Leakage/Fragments}} \oplus \underbrace{0.02\left(\frac{R}{1825}\right)^{\boxed{-1.0}}\left(\frac{B}{3.5}\right)^{\boxed{-0.35}}\left(\frac{E}{100}\right)^{+0.4}}_{\text{Confusion}}$$

- **As expected, larger + higher field gives best performance**
- **R more important than B**
  ⇨ **motivates choice of ILD detector concept parameters**

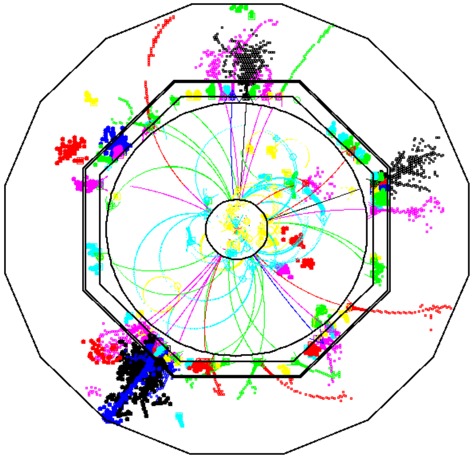
(SiD concept team investigating small, high B-field option)

How important is segmentation ?

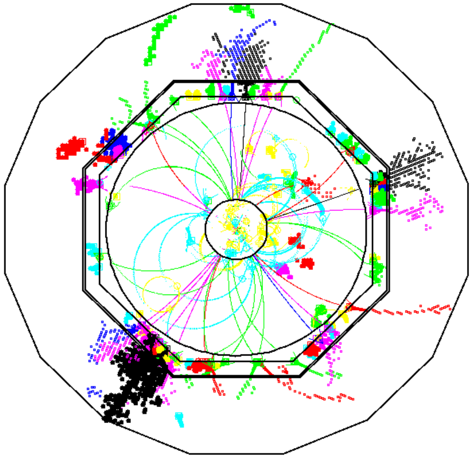
1x1



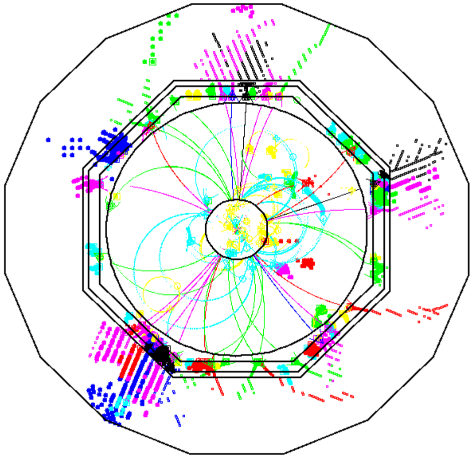
3x3

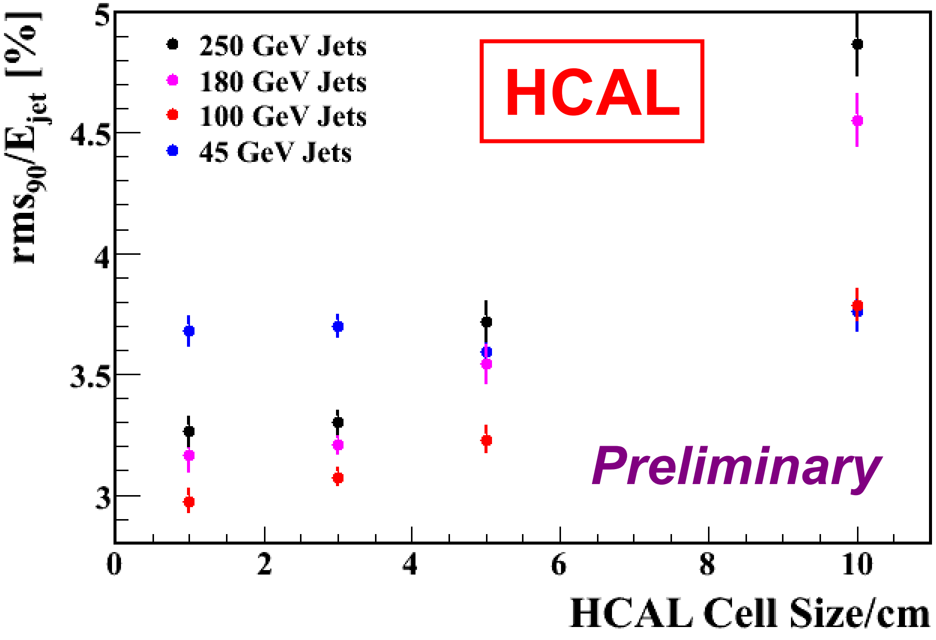


5x5



10x10

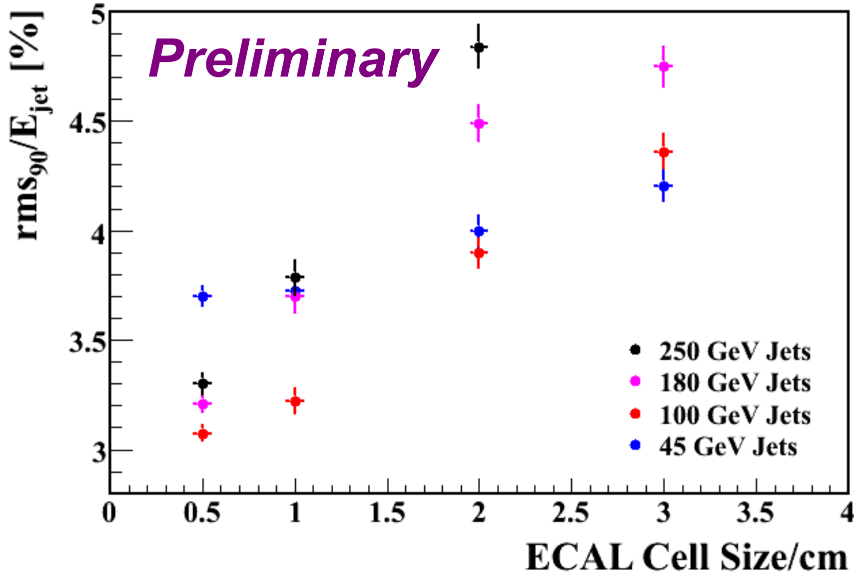




- $3\times3\text{cm}^2$ looks reasonable
- Hint of gain going to $1\times1\text{cm}^2$
- Significant degradation for larger tile sizes, e.g. $5\times5\text{cm}^2$

and ECAL Segmentation ?

- ★ Investigate **10×10mm², 20×20mm²** and **30×30mm²**
  - ▪ Note: retuned PandoraPFA clustering parameters



- ★ Performance is a strong function of pixel size
- ★ High ECAL segmentation is vital for PFA

Caveat:

- ▪ Remember results are algorithm dependent
- ▪ Could reflect flaw in reconstruction

★**Nevertheless:** highly segmented HCAL/ECAL clearly essential

- ★ Particle Flow can deliver **ILC** jet energy goals
- ★ Whole detector concepts studied, and (partially) optimised

  e.g. **ILD**
- ★ What about Particle Flow for higher energy machines ?

❼ Potential at CLiC (many questions)

- ★Traditional calorimetry $\sigma_E/E \approx 60\%/\sqrt{E/\mathrm{GeV}}$
- ★Does not degrade significantly with energy (leakage)
- ★Particle flow gives much better performance at "low" energies
  - ▪ very promising for ILC

rms90 **PandoraPFA v03-β**

E_{JET}	$\sigma_E/E = \alpha/\sqrt{E_{jj}}$ $\|\cos\theta\|<0.7$	σ_E/E_j
45 GeV	23.8 %	3.5 %
100 GeV	29.1 %	2.9 %
180 GeV	37.7 %	2.8 %
250 GeV	45.6 %	2.9 %

$Z \rightarrow d\bar{d}$, $Z \rightarrow u\bar{u}$

What about a machine like CLiC ?

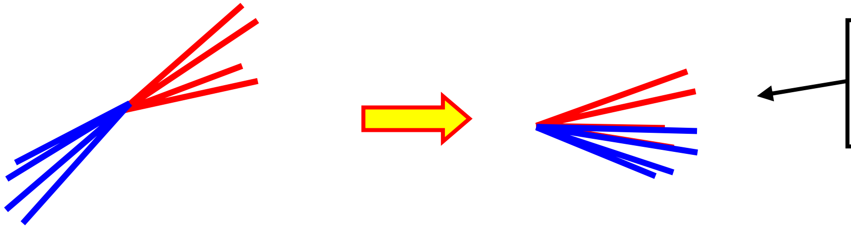
- ★For 1 TeV jets, particle flow will not give $\sigma_E/E < 60\%/\sqrt{E/\mathrm{GeV}}$ (probably substantially worse)
- ★But not interested in 1 TeV jets: most interesting physics likely to be 6, 8, ... fermion final states
- ★For 0.5 TeV jets, particle flow likely to be comparable to a traditional calorimetric approach

\+ a PFlow calorimeter still has good calorimetric resolution

can design algorithm to move away from p.flow at higher energies

Physics Considerations

- Whether particle flow is appropriate for a **multi-TeV e^+e^-** collider needs detailed study
- Will depend on physics program, e.g.
  - lower energy operation (staged?) / threshold scans
  - what matters at higher energies, W/Z separation or new particle mass resolution ?
  - How important is mass resolution for boosted di-jets



Particle flow reco. might help here

Many Questions

Potential to get the answers:

- Have the tools, simulation/reconstruction
- With increased cooperation between ILC and CLiC could make rapid progress…

❽ Summary/Outlook

Summary

- ★ Great deal of effort (worldwide) in the design of the ILC detectors
- ★ Widely believed that calorimetry and, in particular, jet energy resolution drives the ILC detector design
- ★ Also believed that it is likely that PFA is the key to achieving ILC goal
- ★ Calorimetry at the ILC = HARDWARE + SOFTWARE (new paradigm)
- ★ It is difficult to disentangle detector/algorithm….
- ★ Can only address question with "realistic algorithms"
  - i.e. serious reconstruction 10+ years before ILC turn-on
- ★ With **PandoraPFA** have reached the ILC "goal"

Now have a proof of principle of Particle Flow Calorimetry

⇨ **Unprecedented Jet Energy Resolution**

Outlook

- Beginning to understand Particle Flow Calorimetry
  vital to understand ultimate potential
- Starting to investigate high energy potential

Good time to tie in with CLiC detector/physics studies

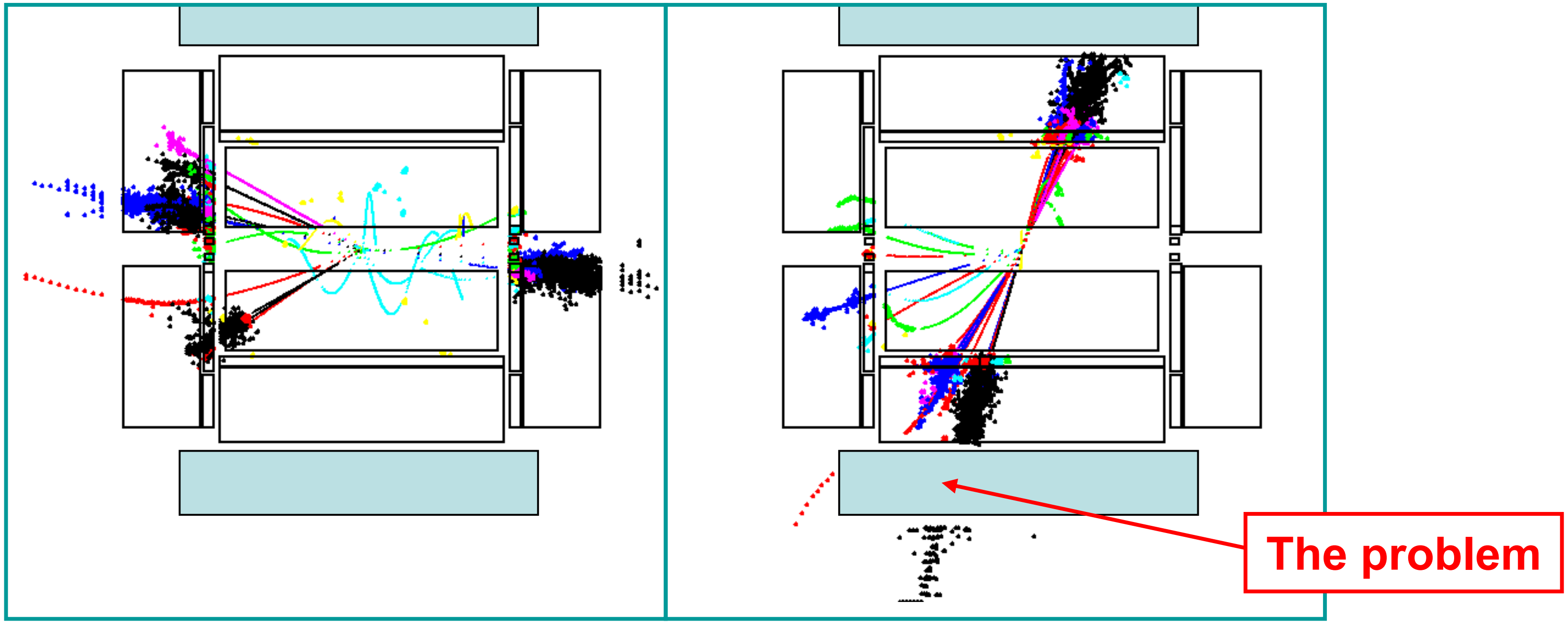
End

❸ Optimisation Studies: ① HCAL Depth

Two interesting questions:

- ★ How important is HCAL leakage ?
  - vary number of HCAL layers
- ★ What can be recovered using MUON chambers as a "Tail catcher"
  - PandoraPFA now includes MUON chamber reco.
  - Switched off in default version
  - Simple standalone clustering (cone based)
  - Fairly simple matching to CALO clusters (apply energy/momentum veto)
  - Simple energy estimator (digital) + some estimate for loss in coil

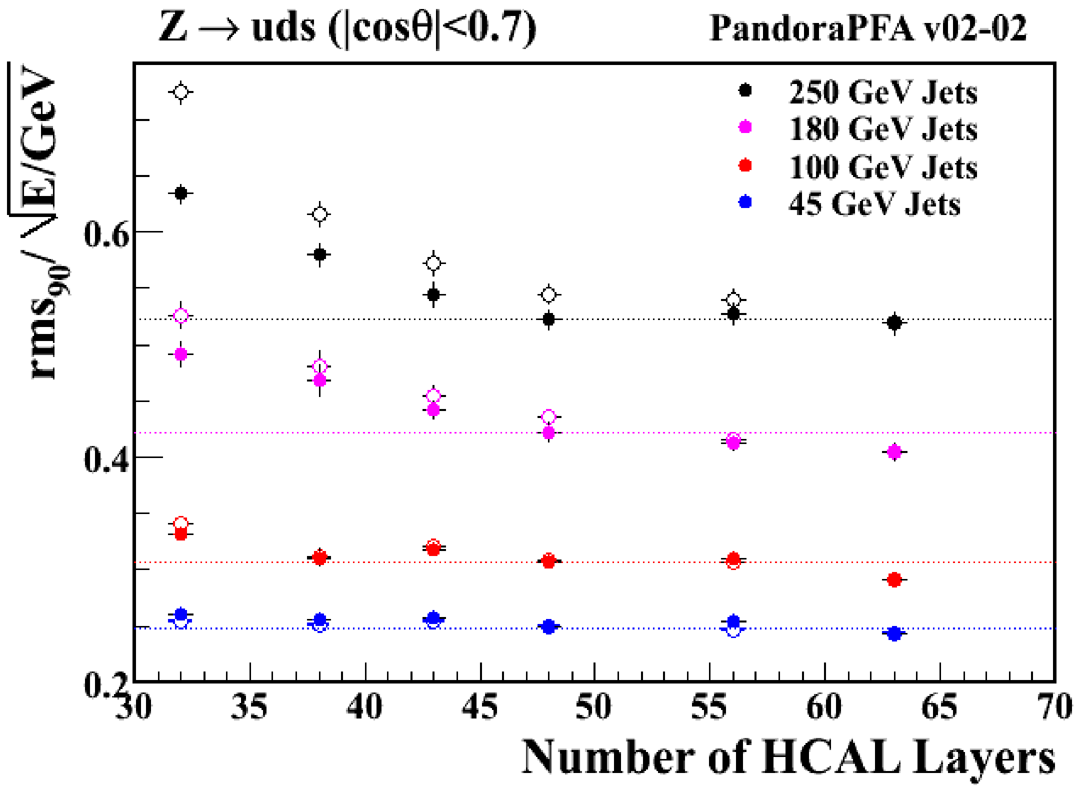
e.g.



The problem

HCAL Depth Results

- **Open circles** = no use of muon chambers as a "tail-catcher"
- **Solid circles** = including "tail-catcher"



HCAL Layers	λ_I	
	HCAL	+ECAL
32	4.0	4.8
38	4.7	5.5
43	5.4	6.2
48	6.0	6.8
63	7.9	8.7

ECAL : λ_I = 0.8
HCAL : λ_I includes scintillator

★ Little motivation for going beyond a 48 layer (6 λ_I) HCAL

★ Depends on Hadron Shower simulation

★ "Tail-catcher": corrects ~50% effect of leakage, limited by thick solenoid

For 1 TeV machine "reasonable range" ~ 40 – 48 layers (5 λ_I - 6 λ_I)