Software compensation and particle flow

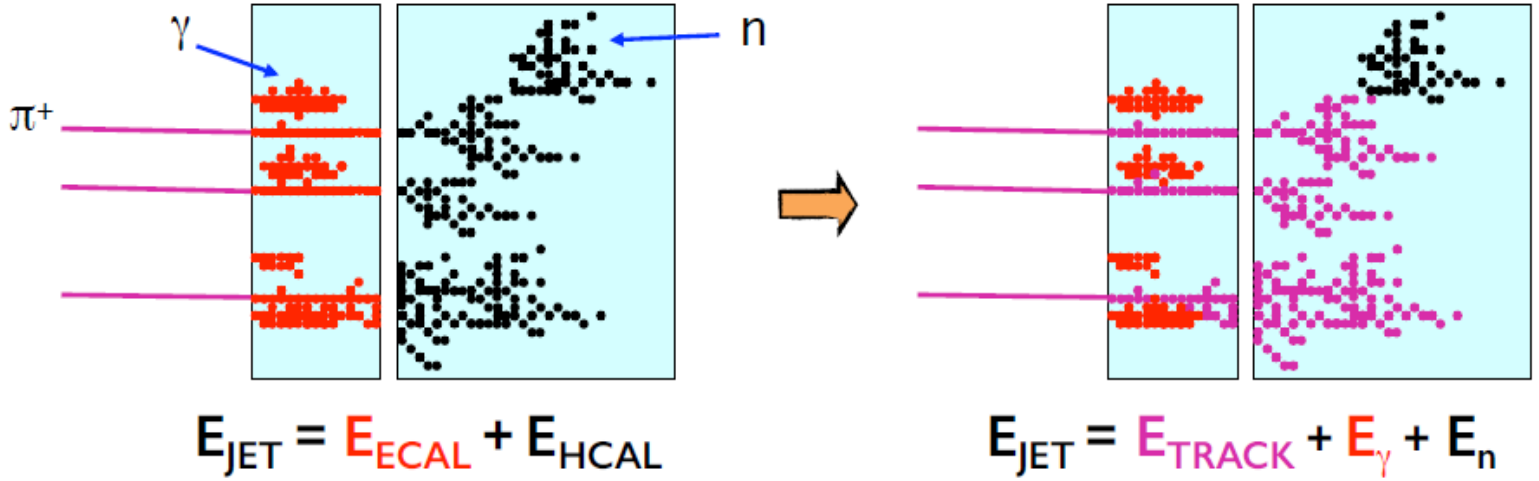
Boruo Xu, on behalf of
CALICE collaboration
Beijing, TIPP 2017

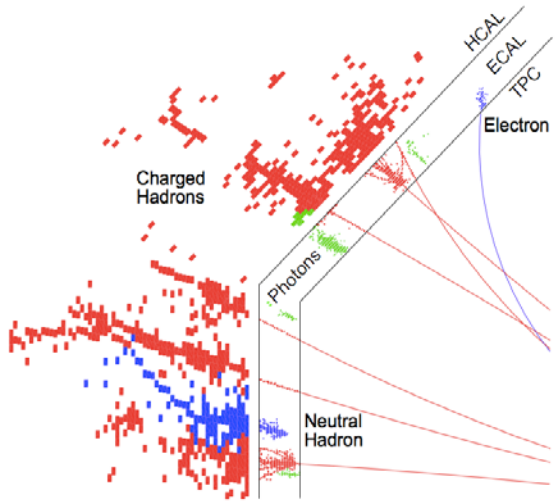
Content

- Background information
- Software compensation
  - Implementation with single particle data from test beam
  - Implementation in PandoraPFA for simulation
- Single hadron and jet energy resolutions improvements
  - Transverse cell size optimisation study
  - Software compensation for semi-digital HCAL

Particle Flow approach to calorimetry

- High-granularity Particle Flow: reconstruct individual particles.



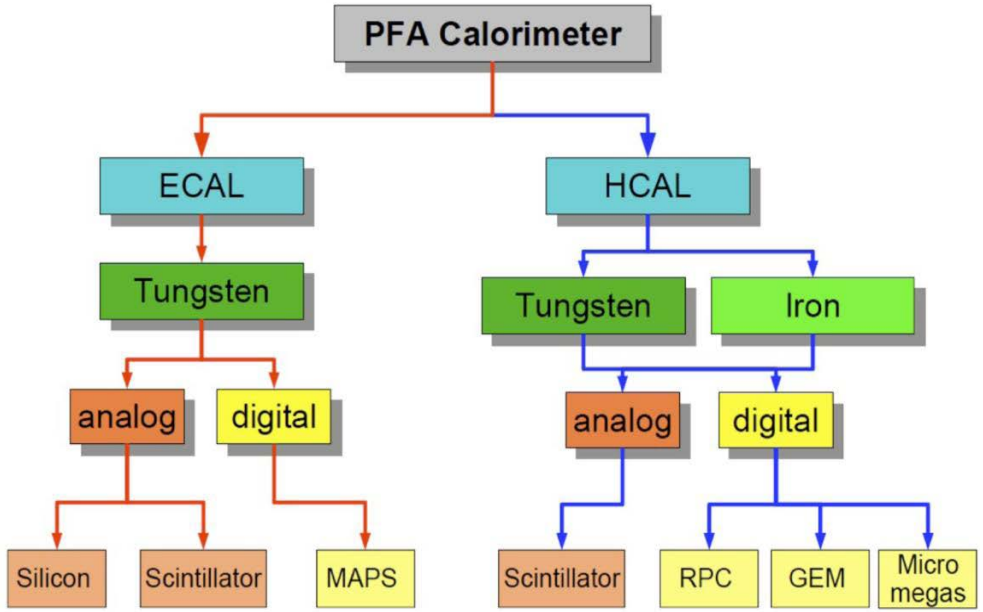


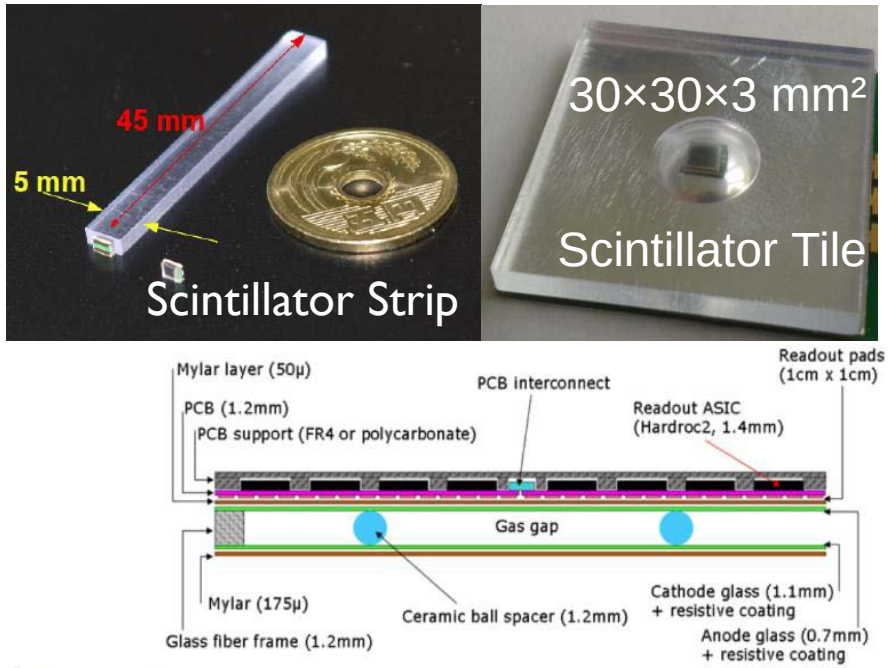
- ~3% jet energy resolution required for WW/ZZ separation using 4 quarks channel, achieved in simulation studies with PandoraPFA Eur.Phys.J. C75 (2015) no.9, 439
- Driving factor for high-granularity calorimeter design of future collider detectors

- Charged particle momentum measured in tracker (60% jet energy)
- Photon energies measured in ECAL: $\sigma E/E < 20\% /\sqrt{E}$ (30% jet energy)
- Only neutral hadron energies measured in HCAL (10% jet energy)
- Much improved resolution.

CALICE collaboration

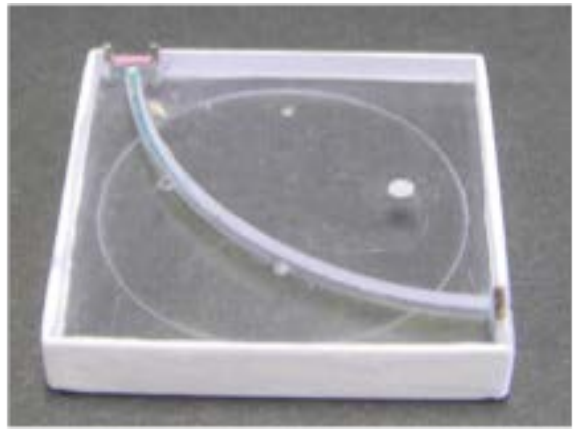
- CALICE collaboration: 55 institutes in 19 countries, ~ 350 members
- Research and development of high-granularity calorimeters for future lepton colliders
- A rich program exploring full spectrum of imaging calorimeter technologies https://twiki.cern.ch/twiki/bin/view/CALICE/WebHome



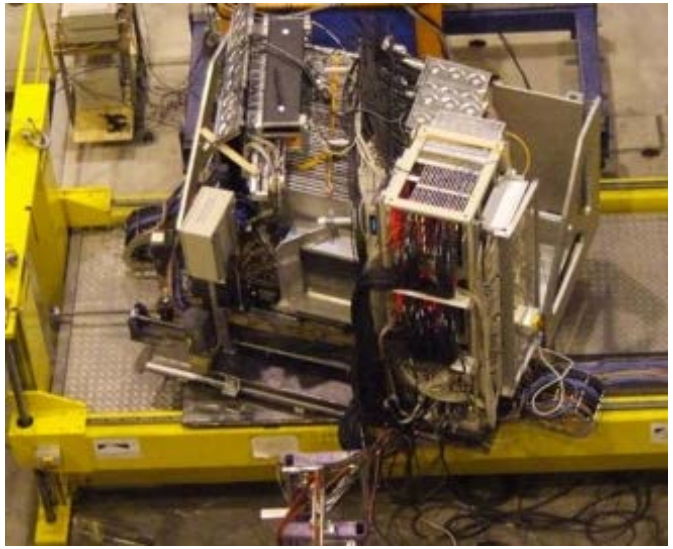


CALICE calorimeter prototypes

- A prototype to demonstrate high-granularity analogue hadronic calorimeters (AHCAL) based on plastic scintillators with SiPM readout
- Tested with two absorbers: Steel, 21mm/layer Tungsten, 10mm/layer
  - Exploring more compact layout for higher energies at CLIC

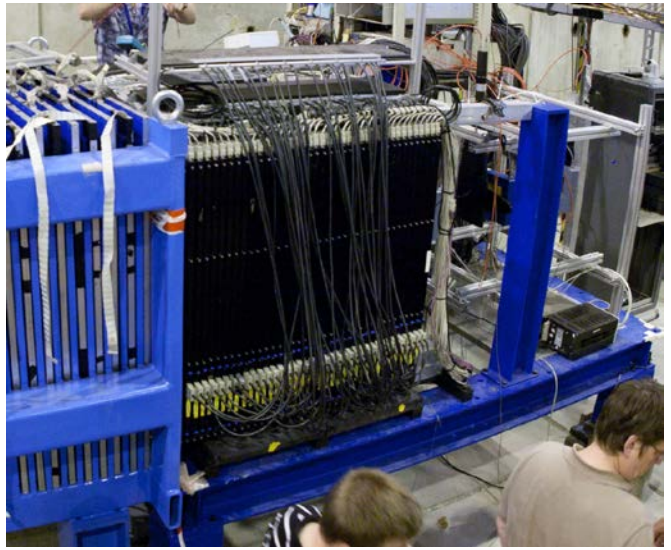


- The basic element of the active layers: a 3 x 3 cm^2 scintillator tile, with individual SiPM readout
- The full system: The first large-scale use of SiPMs, ~ 8000 in prototype, started taking data in 2006 - 38 layers



Steel

Tungsten

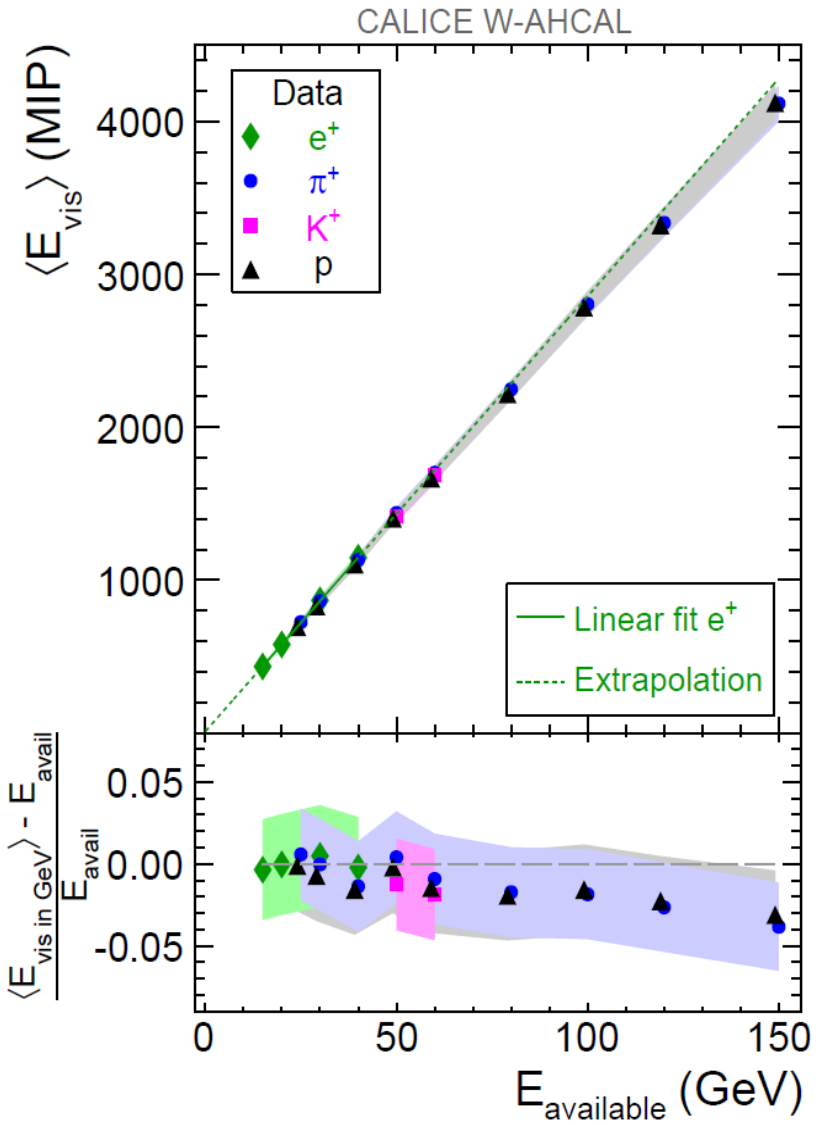


Compensation calorimeter

- Non-compensating calorimeter: response for electromagnetic (EM) showers is typically higher than hadronic shower (e/h > 1)
  - Hadronic showers contain EM sub-showers from photons, and inaccessible processes of hadronic sub-showers, i.e. nuclear target recoil, binding energy loss.
- Difference in response of EM and hadronic sub-shower degrades energy resolution.
  - The fraction of EM and hadronic sub-showers fluctuates strongly
- Compensating calorimeter: response for electromagnetic showers is identical to hadronic shower (e/h = 1) → improves energy reconstruction and hadron resolution
  - Specific detector material design, or
  - Software compensation

Compensating calorimeter: material design

- Specific detector material design for compensating calorimeter
- With Tungsten absorbers, the (analogue hadronic) calorimeter is very close to intrinsic compensation.
- Responses for EM and hadronic sub-showers are similar
- No improvement from software compensation is expected
- JINST 10 (2015) no.12, P12006



Calorimeter: software compensation

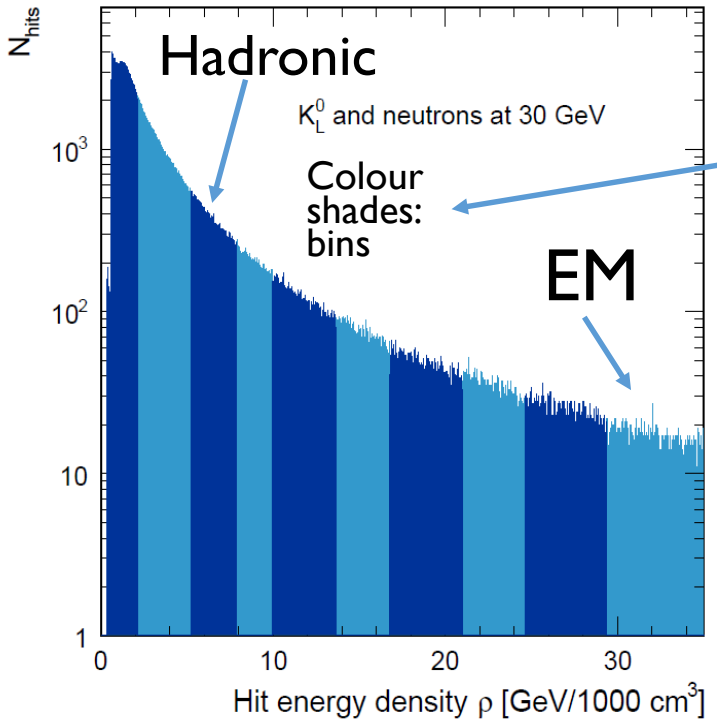
- Non-compensating calorimeter: response for electromagnetic showers is typically higher than hadronic shower (e/h > 1)

- Software compensation (SC): use shower density to correct for different response to EM and purely hadronic sub-showers
  - does not use prior knowledge of the beam energy - the uncorrected shower energy / cluster energy is used to select the software compensation weights for a given event.
- Non-compensation $E_{RAW} = \sum_{hits} E_{ECAL} + \sum_{hits} E_{HCAL} \times 1$
- Software compensation: $E_{SC} = \sum_{hits} E_{ECAL} + \sum_{hits} E_{HCAL} \times \omega$

Weight is a function of shower density and E_{RAW}

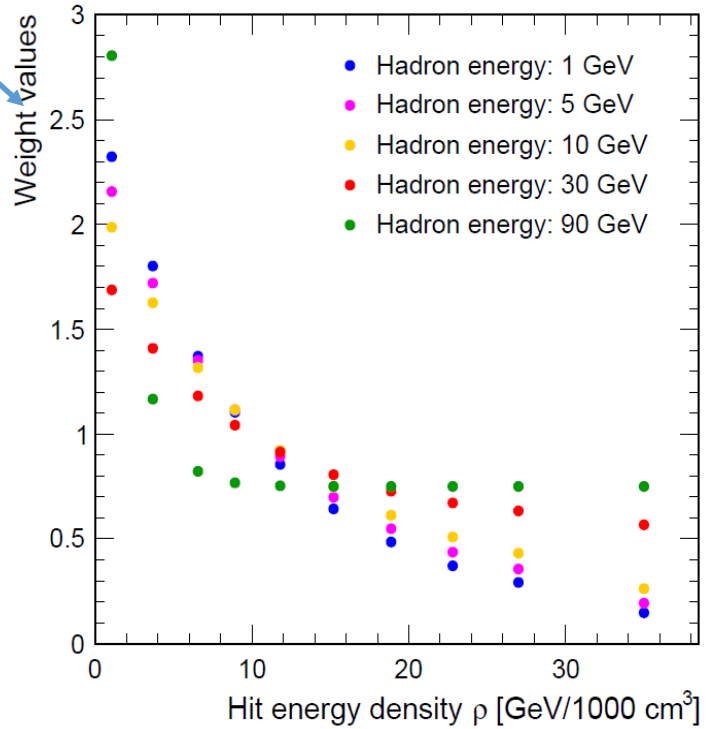
Software compensation: logic

- Identify high hit-density (ρ) as EM sub-showers, low hit-density as hadronic sub-showers
  - Implemented in PandoraPFA in simulation, but the concept is taken from the CALICE study and tested with real data



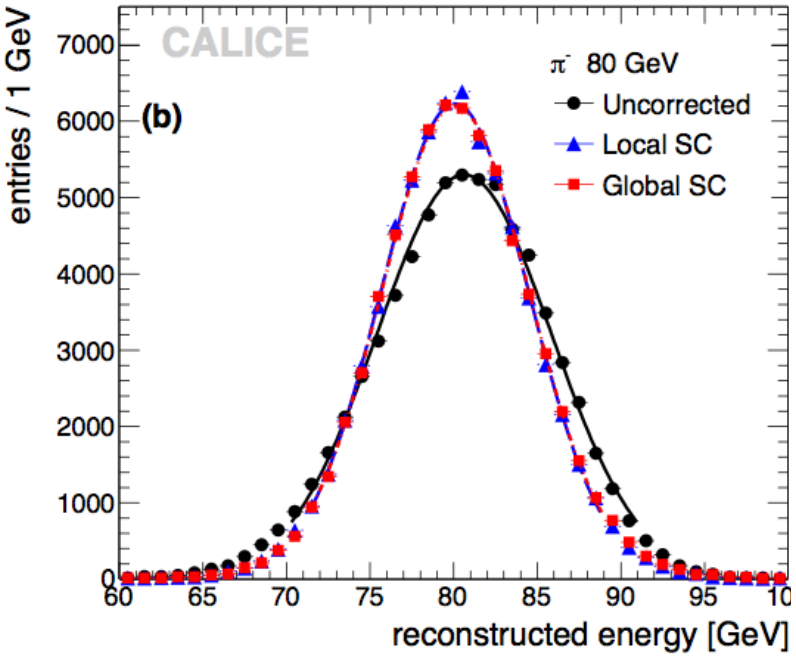
$$E = \sum_{hits} E_{ECAL} + \sum_{bins\,i} E^{i}_{HCAL} \times \omega(\rho_i)$$

- Decrease EM sub-showers weight
- Increase hadronic sub-showers weight

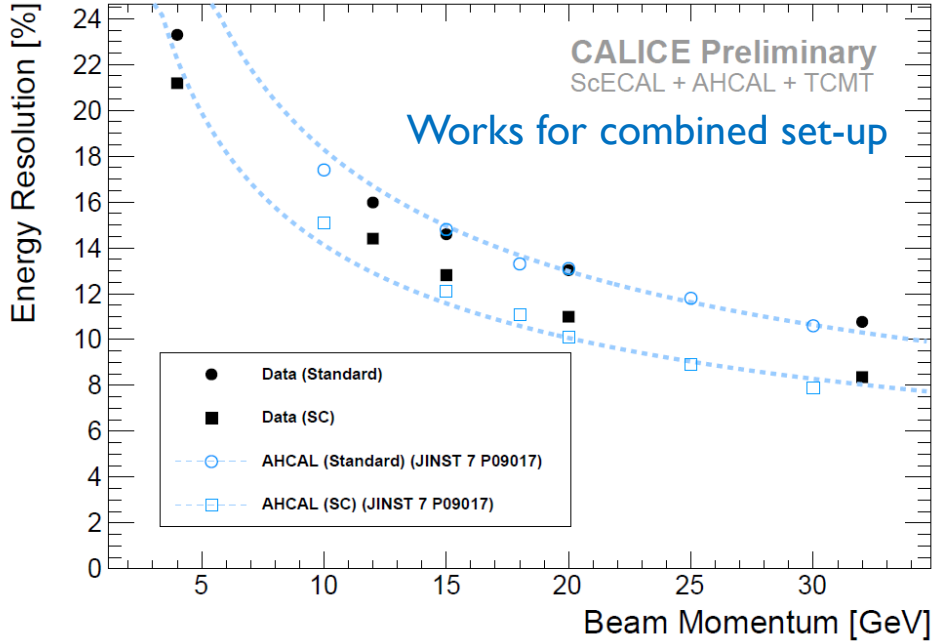


Software compensation with test beam data

- Software compensation developed and tested with test beam data
- Steel analogue calorimeter is intrinsically non-compensating ($e/\pi^- \sim 1.2$), but high granularity allows software compensation
  - JINST 7, P09017 (2012)



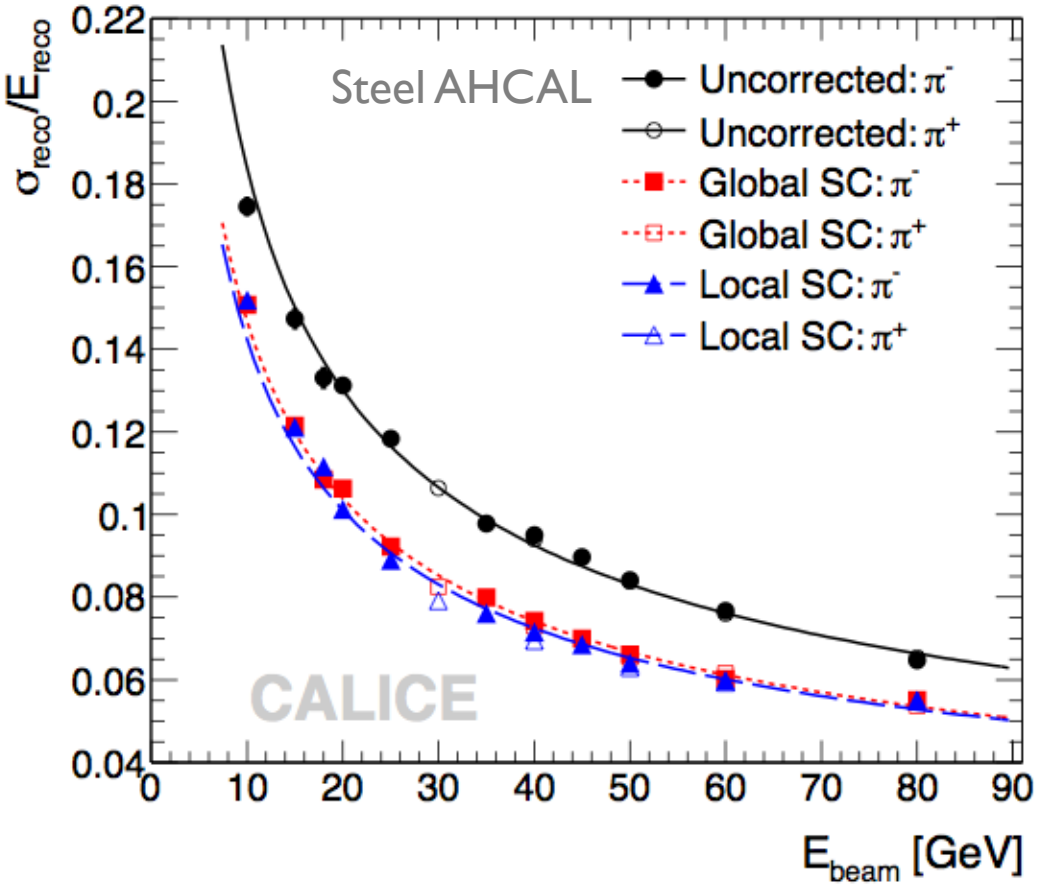
CALICE Analysis Note CAN-056



- The response of individual cells is calibrated using muons
- Particle energy reconstructed from the sum of amplitudes of all detector hits, with appropriate MIP → energy calibration

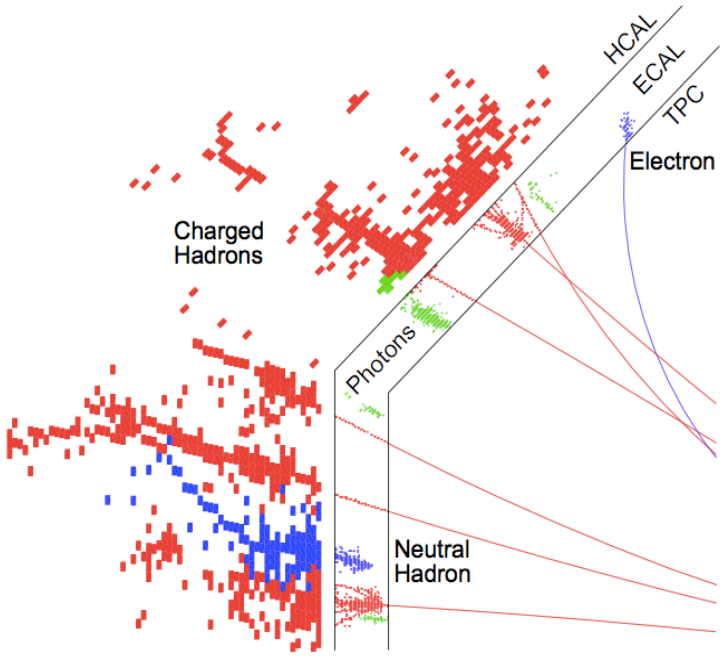
Software compensation with test beam data

- Software compensation improves single particle energy resolution of non-compensating calorimeter, tested with data
- Resolution (AHCAL):
- Steel: ~ 58%/√E w/o SC,
  ~ 45%/ √E with SC
- Tungsten: ~ 58%/√E
  - Tungsten comparable to steel w/o SC, due to coarser sampling of EM sub-showers in the W-AHCAL: 3 X_0 / layer rather than 1 X_0 with steel



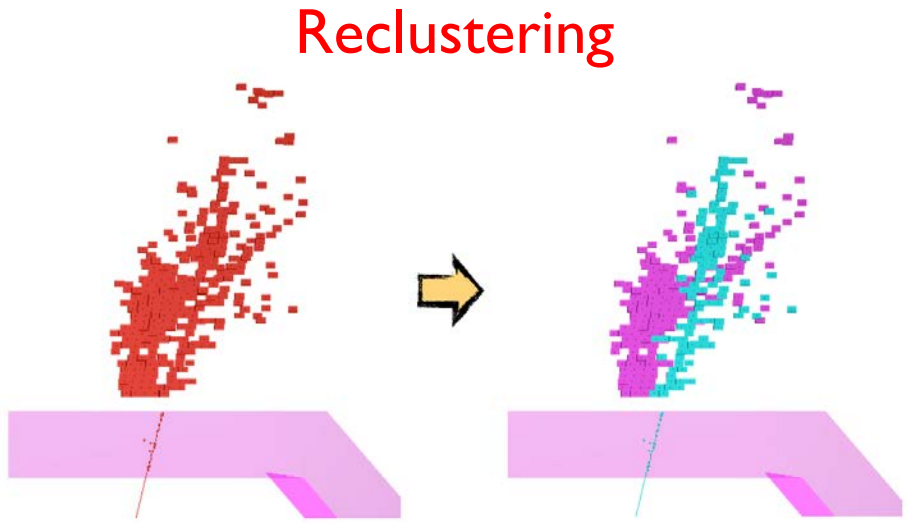
Software compensation: particle flow

- Developed and tested on real data, now implemented in PandoraPFA for simulation
- Software compensation improves:
  - 1) Single hadron energy resolution
  - 2) Jet energy resolution
    - better energy resolution → better pattern recognition (next slide)
- Software compensation implemented in cluster energy estimation in PandoraPFA: 1)
- Full use of software compensation throughout pattern recognition reconstruction in PandoraPFA: 1) and 2)



Software compensation: pattern recognition

- In high density jets (high energies), limit of "pure" pattern recognition based particle flow may be reached
  - Cannot cleanly resolve neutral hadrons in hadronic showers.
- Address the problem "statistically"; recluster if significant discrepancy between cluster energy and its associated-track momentum.
  - Alter clustering parameters, or change clustering algorithm entirely, until cluster splits to achieve sensible track-cluster association
- Better hadron energy estimation→ better reclustering → better reconstruction

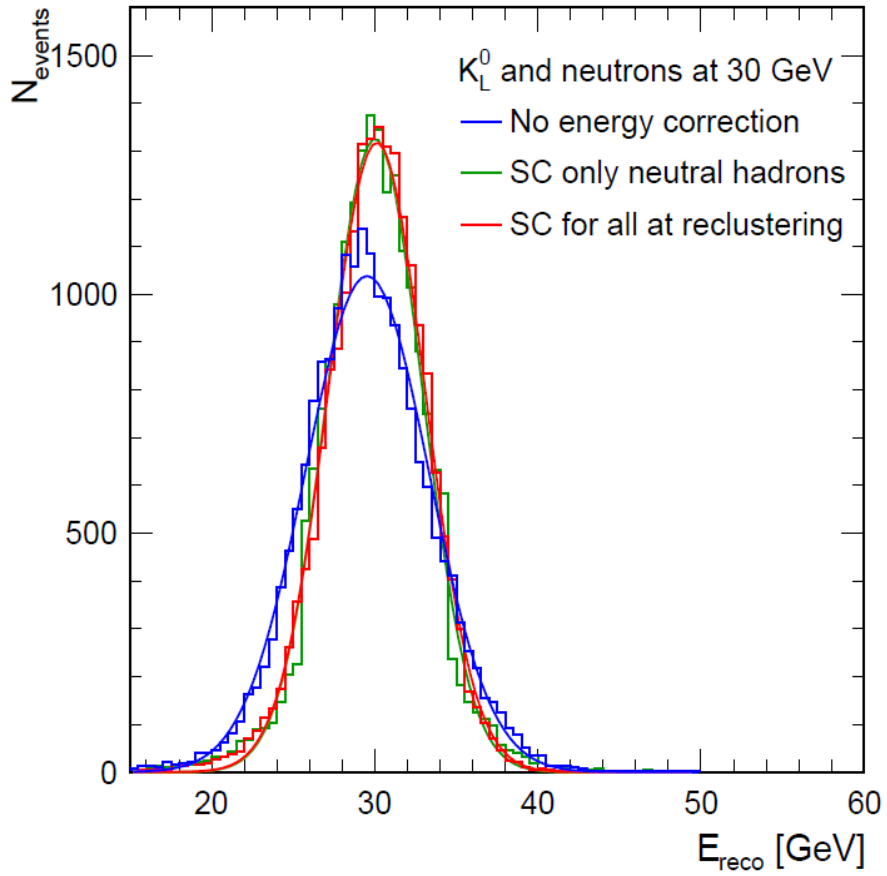


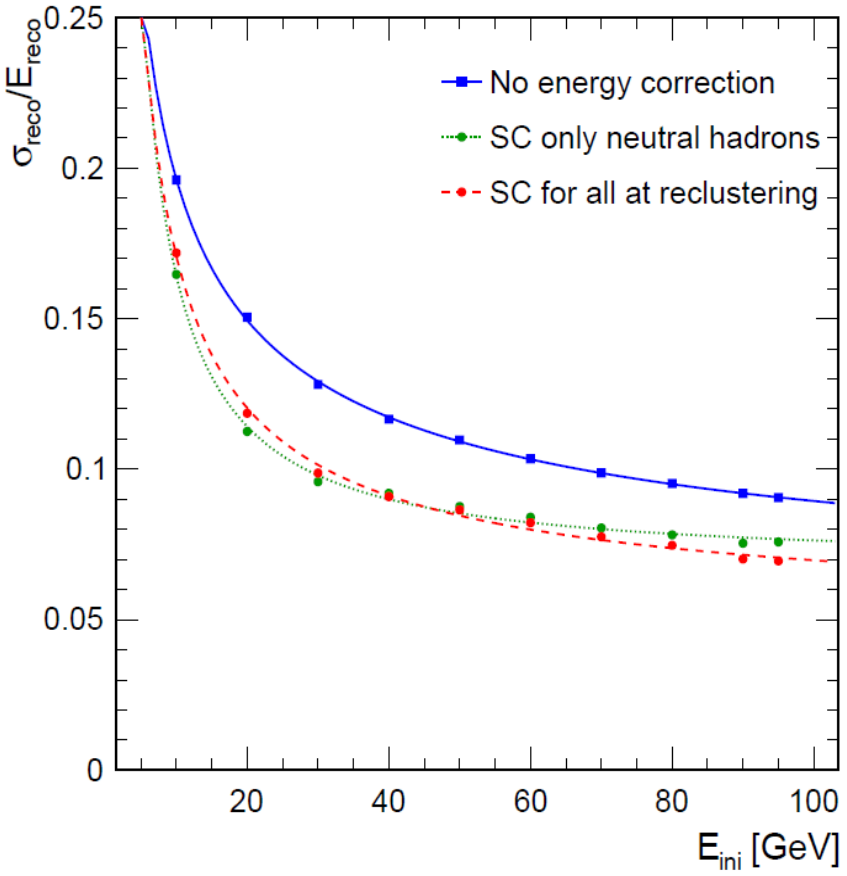
topological track-cluster association, large discrepancy between cluster energy and its associated-track momentum.

After reclustering, cluster splits to achieve sensible track-cluster association

Software compensation: single hadron resolution

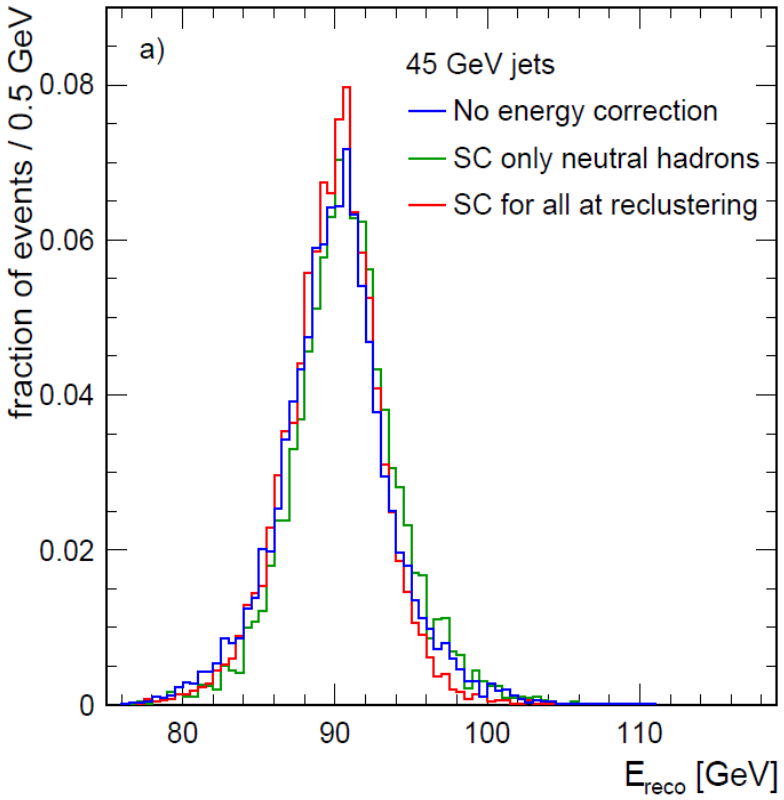
- Software compensation improves: 1) Single hadron energy resolution
- Narrower energy distributions

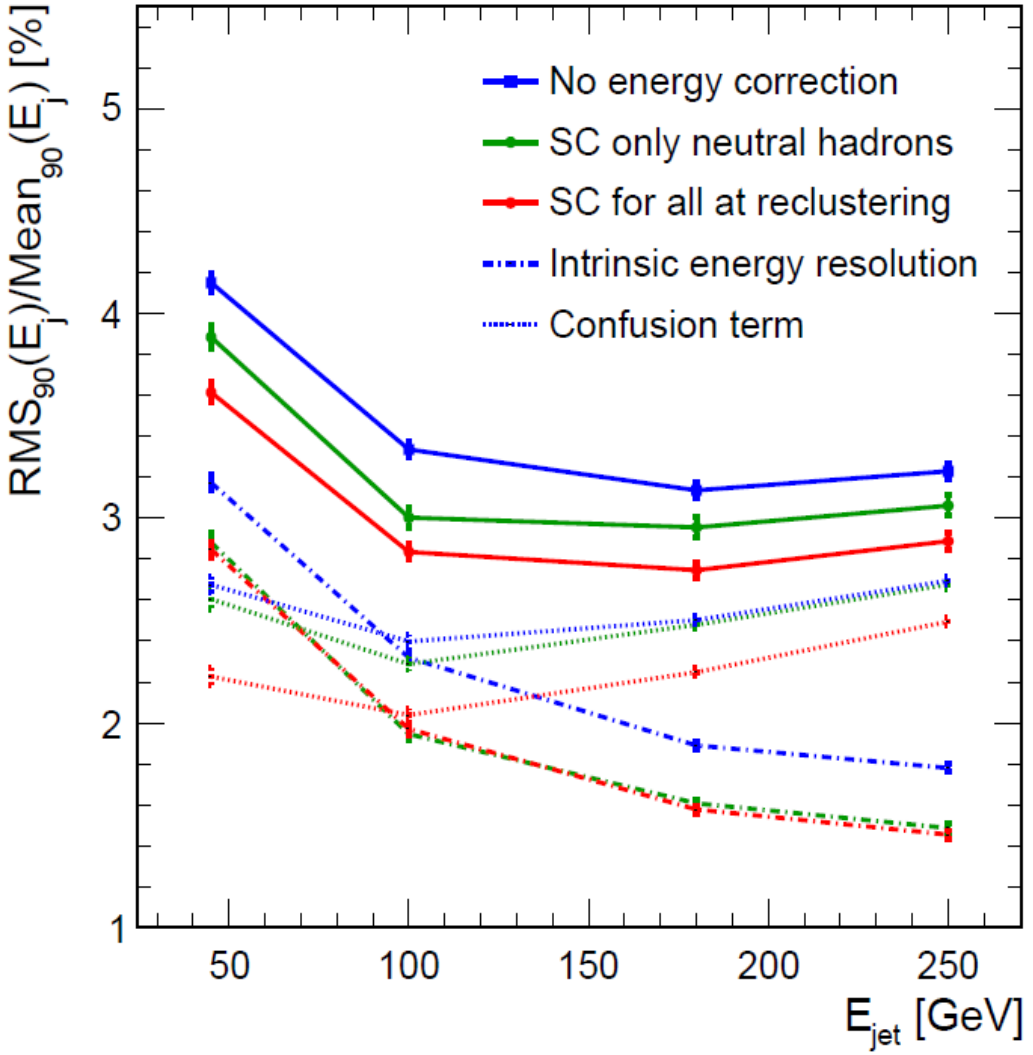




Software compensation: jet energy resolution

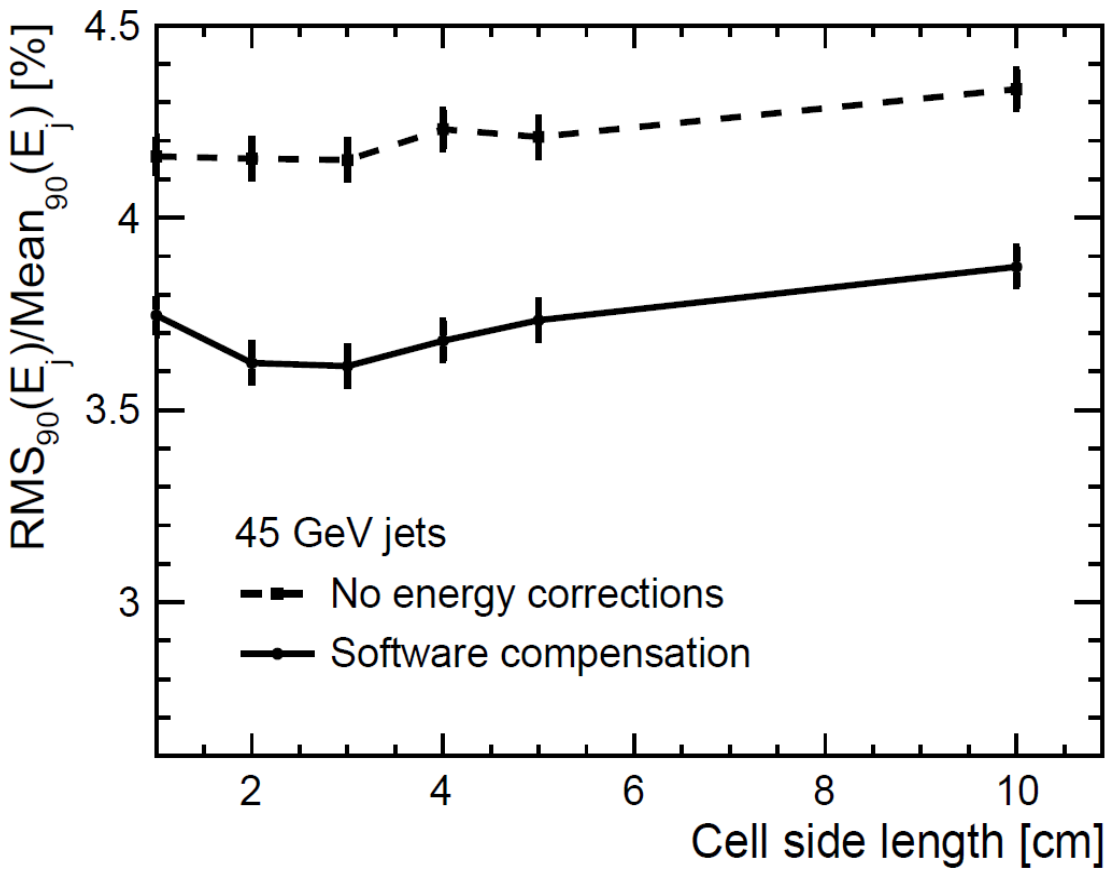
- Software compensation improves:
  2) Jet energy resolution (JER)
- Narrower distribution → better intrinsic JER
- Better cluster energy → better reconstruction (reclustering) → fewer mistakes

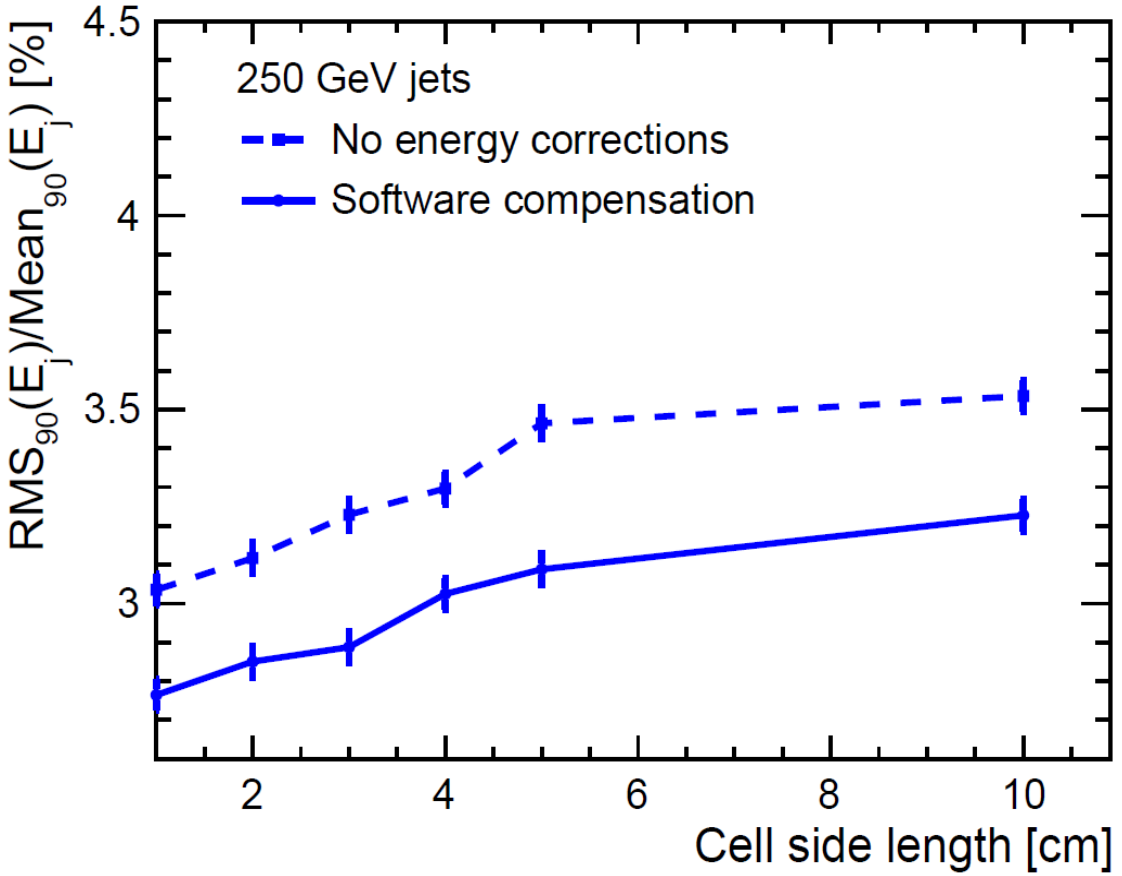




Software compensation: transverse granularity

- Software compensation offers best jet energy resolution, used in optimisation studies: here AHCAL square cell size
  - Same trend w/ and w/c software compensation





Software compensation: semi-digital calorimeter

- Software compensation technique can be used for semi-digital calorimeter.
- Energy calculation of semi-digital calorimeter can be reformulated to be mathematically similar to software compensation in AHCAL
- Details in upcoming paper

Total energy is simple summation of three thresholds N_1, N_2, N_3

Can rewrite as

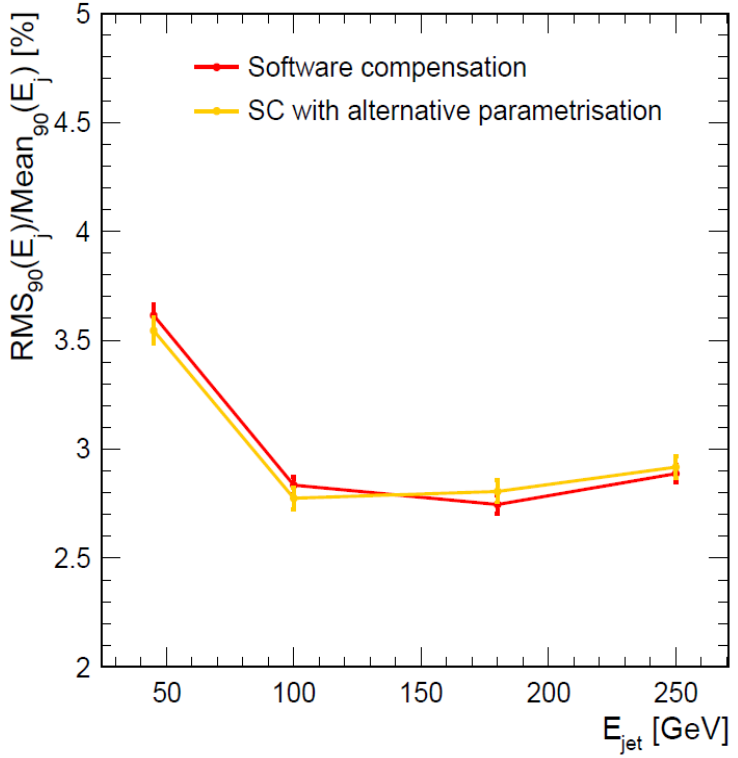
$$E_{SD} = \sum_{bins} \alpha_i . N_i$$

$$E_{SD} = \sum_{hits} \alpha_i . \frac{E_j}{E_j} = \sum_{hits} \omega_j . E_j \text{ with } \omega_j = \frac{\alpha_i}{E_j}$$

Compare with software compensation

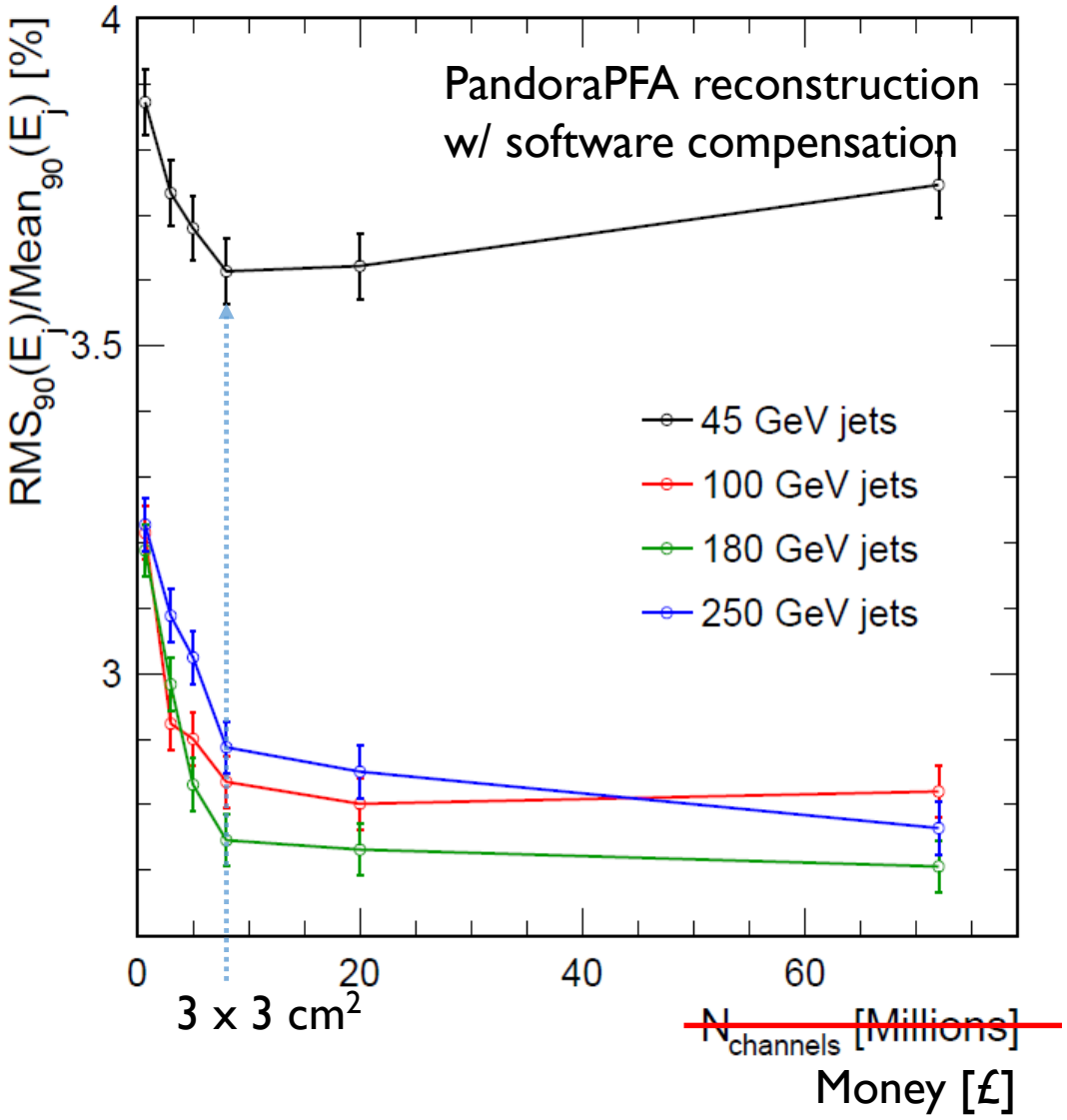
$$\mathrm{E_{SC}} = \sum_{\text{hits}} \mathrm{E_{ECAL}} + \sum_{\text{bin } i} (\mathrm{E^i_{HCAL}} \times \omega(\rho_i))$$

Software compensation bins (analogue calorimeter) ~ thresholds in SD calorimeter



Software compensation: Conclusion

- Significant gain in jet energy resolution over a wide jet energy range
  - Best performance seen in ILD detector simulation
- Does not significantly alter view on transverse granularity optimisation
  - Default ILD 3 x 3 cm^2 cell size is still a very reasonable choice
- Available in PandoraPFA v02-09-01 onwards (including)
  - Includes semi-digital reconstruction scheme
  - Installed in ILCsoft v01-17-10 onwards (including)
- A paper to be submitted to EPJC imminently



Thank you

Software compensation in Particle Flow reconstruction

Huong Lan Tran*, Katja Krüger, Felix Sefkow

Deutsches Elektronen-Synchrotron DESY, Hamburg, Germany

Steven Green, John Marshall, Mark Thomson

Cavendish Laboratory, Cambridge, United Kingdom

Frank Simon

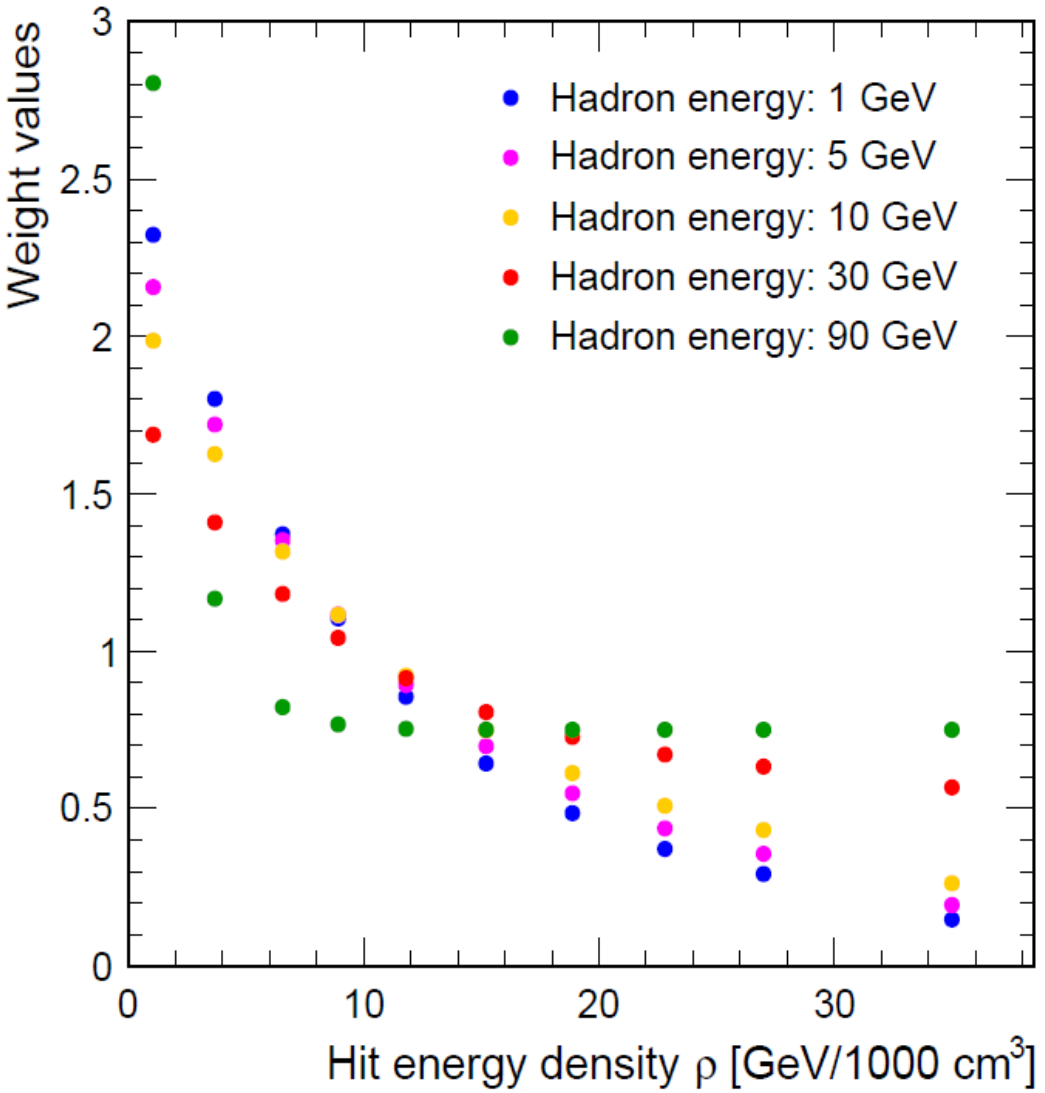
Max-Planck-Institut für Physik, Munich, Germany

Backup

Software compensation: weight

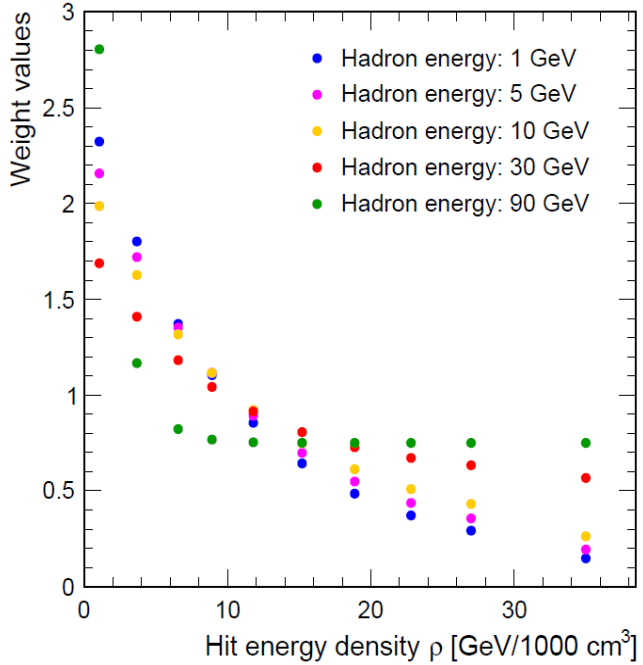
- $E_{SC} = \sum_{hits} E_{ECAL} + \sum_{bins,i} E^{i}_{HCAL} \times \omega(\rho_i)$
- $\omega(\rho) = p_1 \exp(p_2 \rho) + p_3$

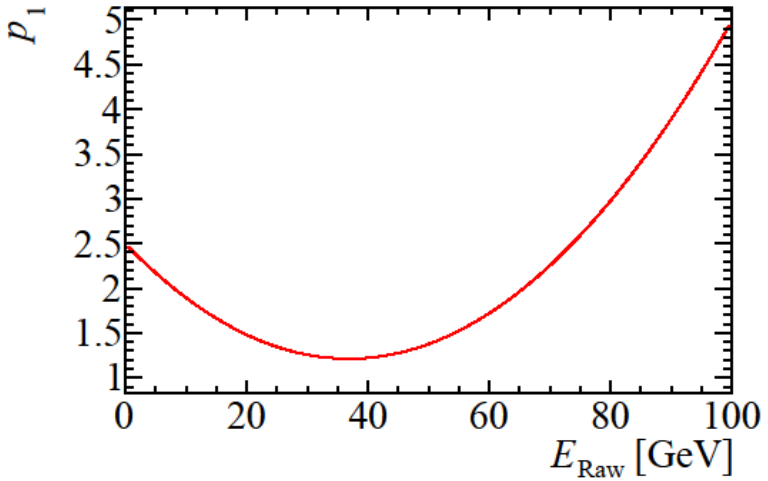
- At high ρ, ω tends to a constant, as shower dominated by EM sub-showers
- At low ρ, ω increases for hadronic sub-showers
- Coefficient p is a function of E_{RAW}
- $P_2 < 0$

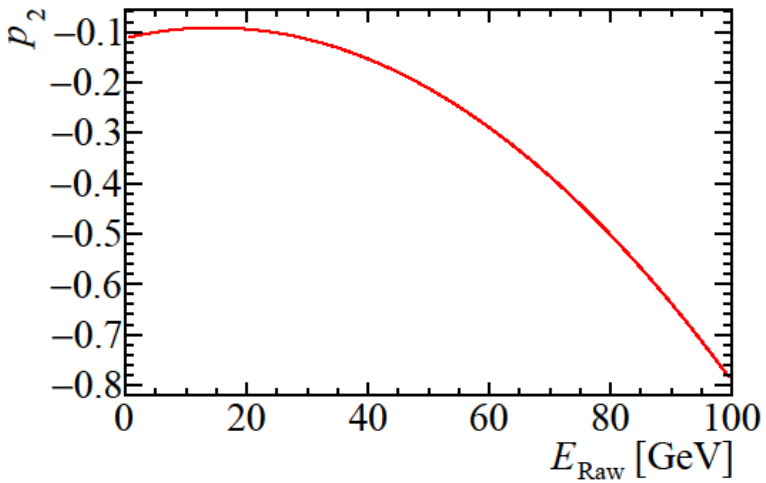


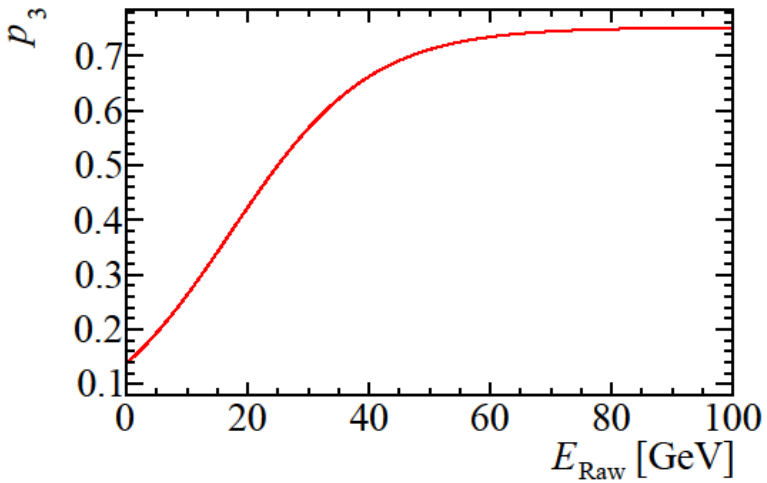
Software compensation: weight

- $E_{SC} = \sum_{hits} E_{ECAL} + \sum_{bins,i} E^{i}_{HCAL} \times \omega(\rho_i)$
- $\omega(\rho) = p_1 \exp(p_2 \rho) + p_3$
  - $p_1 = p_{10} + p_{11} \times E_{sum} + p_{12} \times E^2_{sum}$
  - $p_2 = p_{20} + p_{21} \times E_{sum} + p_{22} \times E^2_{sum}$
  - $p_3 = \frac{p_{30}}{p_{31} + e^{p_{32} \times E_{sum}}}$.

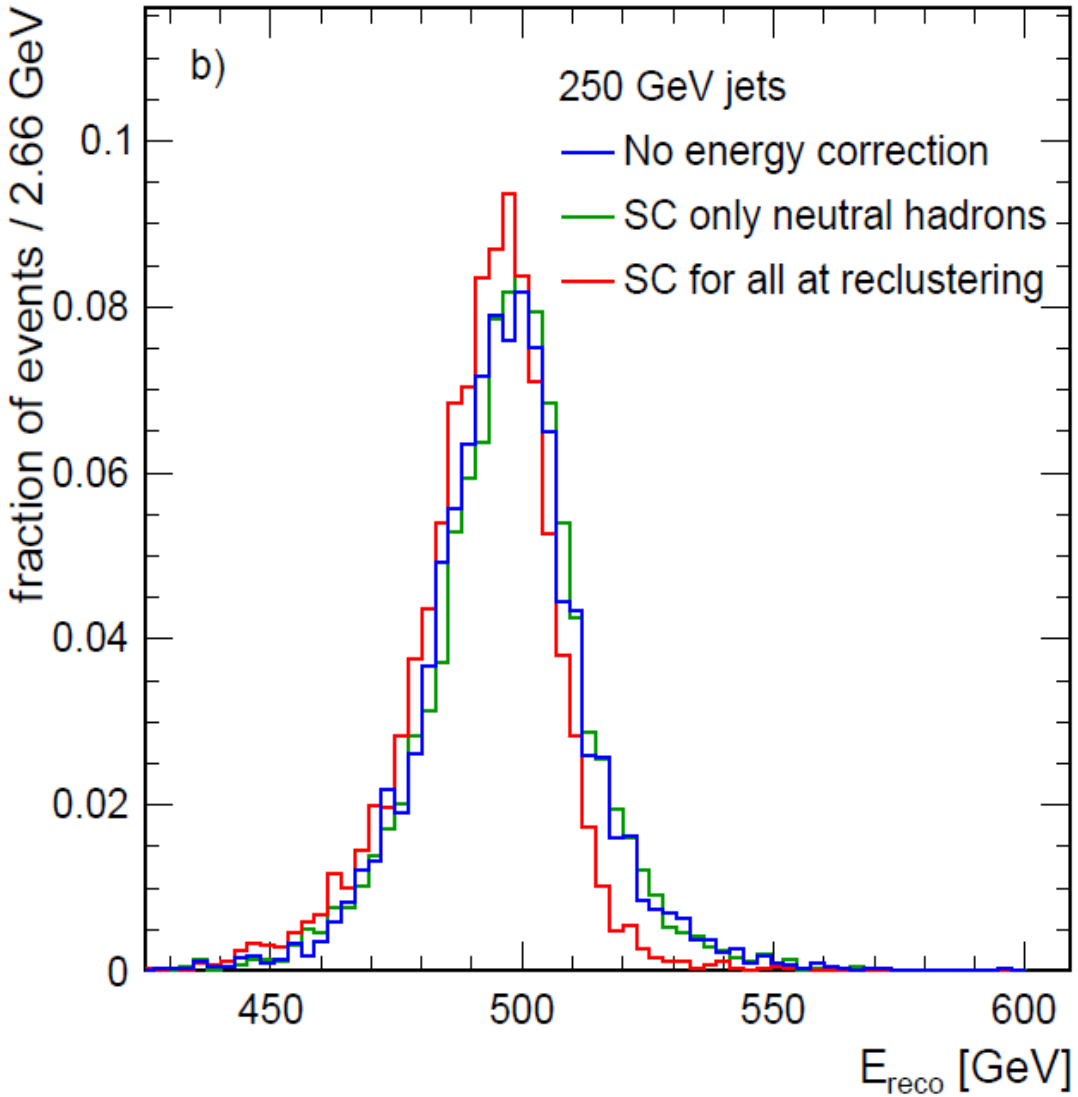


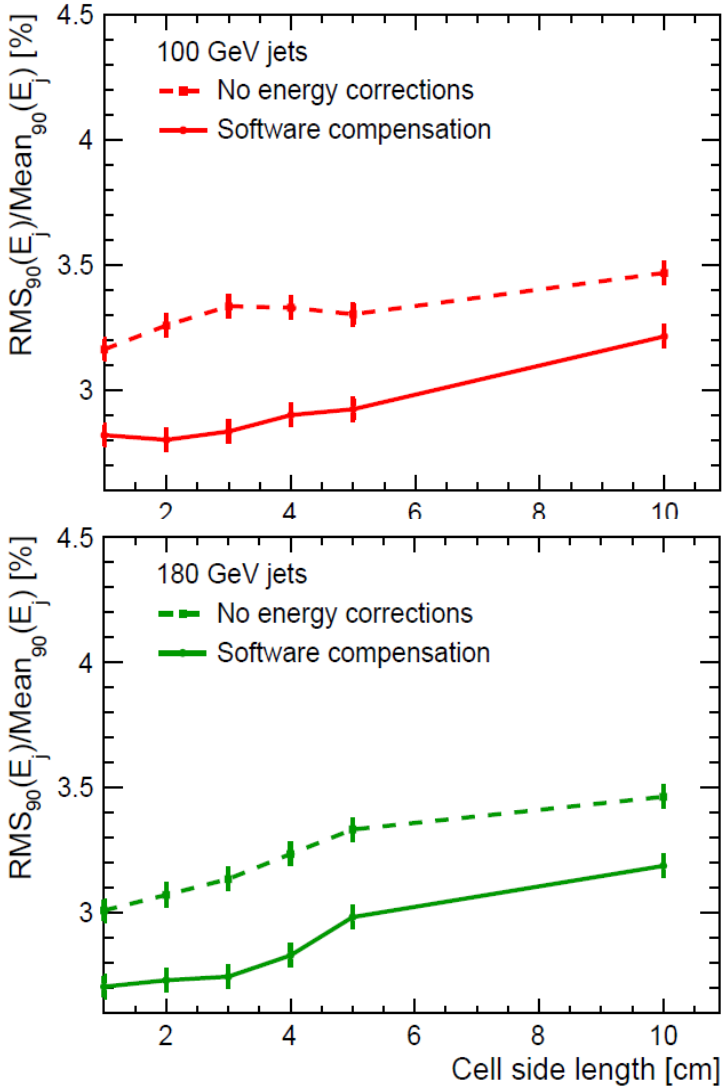






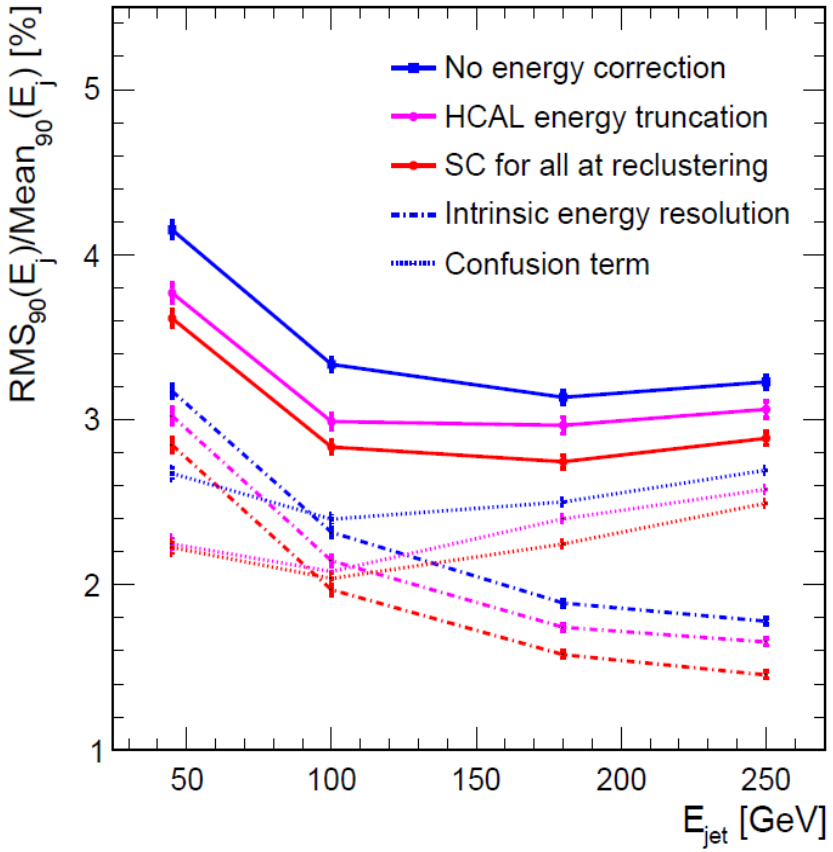
Software compensation for high energy jets

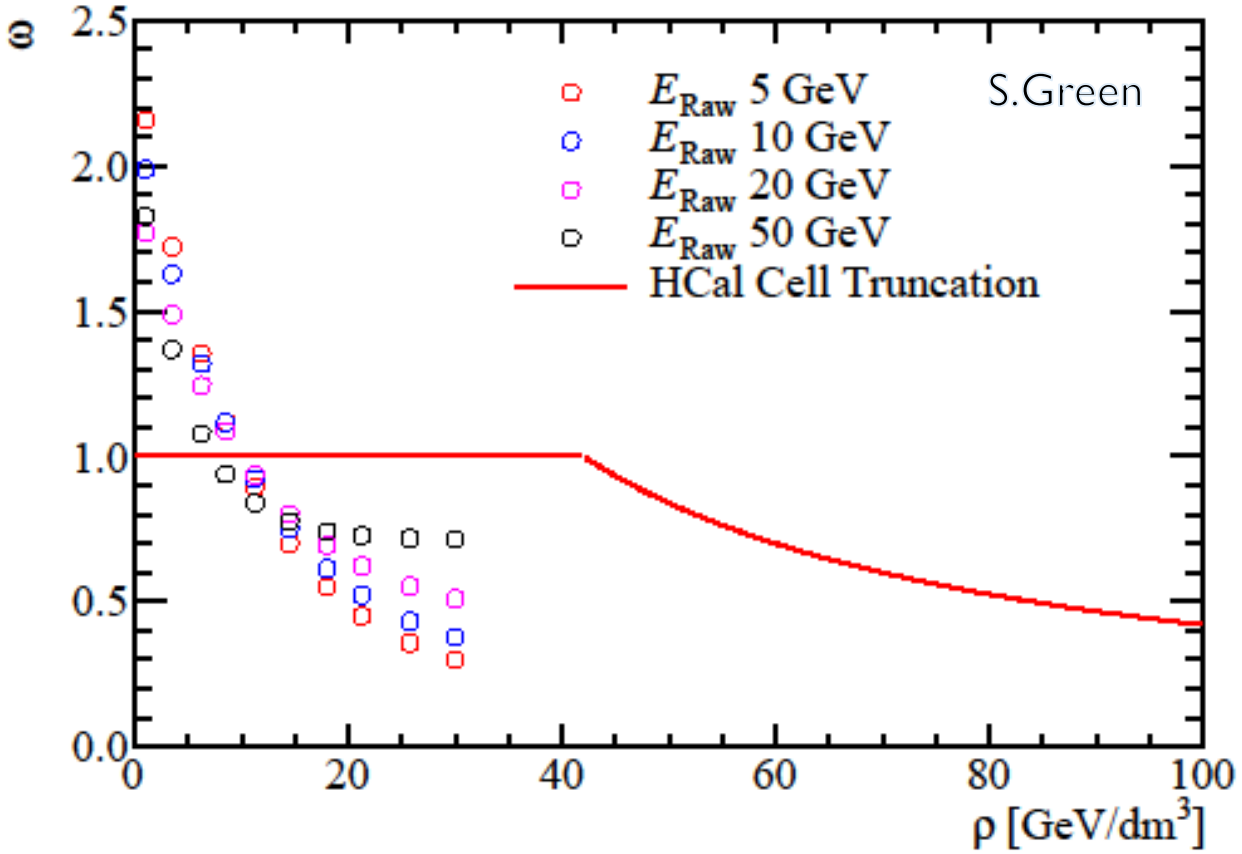




Earlier work on SC in PandoraPFA

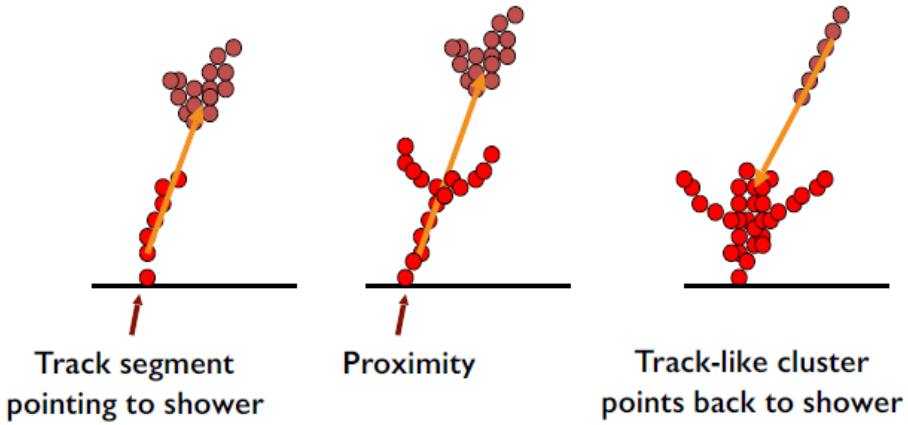
- HCAL cell truncation is a rudimentary software compensation, aided with two other energy corrections algorithms: CleanClusters and ScaleHotHadrons





PandoraPFA Track-cluster association

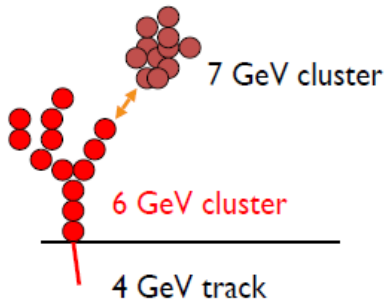
- Clear Associations using cluster mip-segments:
  - Local straight-line fits are performed to hits identified as mip-like and backwards/forward projections are used to identify associations. Tight matching criteria are applied.



https://github.com/PandoraPFA/Documentation/blob/master/Pandora_LC_Reconstruction.pdf

- Less clear associations:

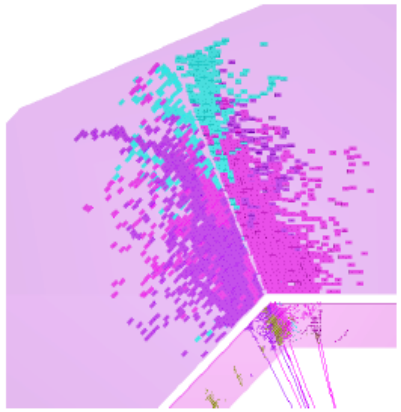
e.g. Small fragments removed, based on proximity to charged hadron clusters



Use E/p consistency to veto clear mistakes

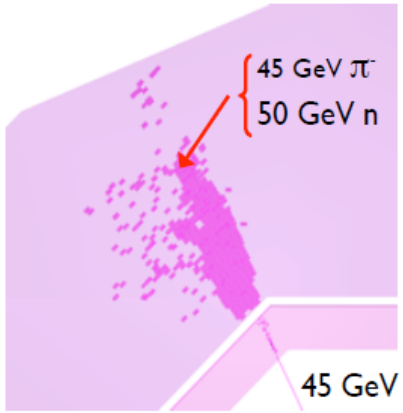
PandoraPFA reclustering

- At some point, in high density jets (high energies), reach limit of "pure" particle flow.
- Cannot cleanly resolve neutral hadrons in hadronic showers.
- Use information from track-cluster associations to identify pattern-recognition problems:



After topological association

Compare E/p values to find problems



Find n absorbed into π^- cluster

e.g. 45GeV track associated to 95GeV cluster:

identify and address clustering problem

- Address the problem "statistically"; if we identify significant discrepancy between energy of a cluster and momentum of its associated track, choose to recluster.
- Alter clustering parameters, or change clustering algorithm entirely, until cluster splits in such a way that we obtain sensible track-cluster associations.

https://github.com/PandoraPFA/Documentation/blob/master/Pandora_LC_Reconstruction.pdf

Software version and configuration

- **Detector model**: ILD_o1_v06
- **Reconstruction software**: ilcsoft_v01-17-07 combined with PandoraPFA version v02-09-01:
  - PandoraSDK v02-03-01
  - LCContent v02-04-01 including software compensation in LCPlugins and hits information registration for software compensation weight training in LCUtility
  - PandoraMonitoring v02-03-00
- **Digitiser**: ILDCaloDigi with realistic options for ECAL and HCAL
- **Calibration constants** optimised using PandoraAnalysis toolkit
- **Timing cut**: 100 *ns*