ILD Optimisation Studies

Steven Green, Boruo Xu, John Marshall, Mark Thomson

Context

- Primary aim is to summarise the optimisation studies performed for ILD using PandoraPFA and jet energy resolution as the primary metric for determining detector performance.
- These studies will most likely be the last studies performed using Mokka for the detector model simulation thanks to the development of DD4HEP.
- In these studies a careful calibration procedure was applied for each detector model. An overview of this procedure and how it interplays with the training of photon likelihood data will be given.
- Software compensation (c.f. Lan's Talk) in the scheme of detector optimisation studies will be introduced and the gain in performance will be discussed.

Calibration

Calibration

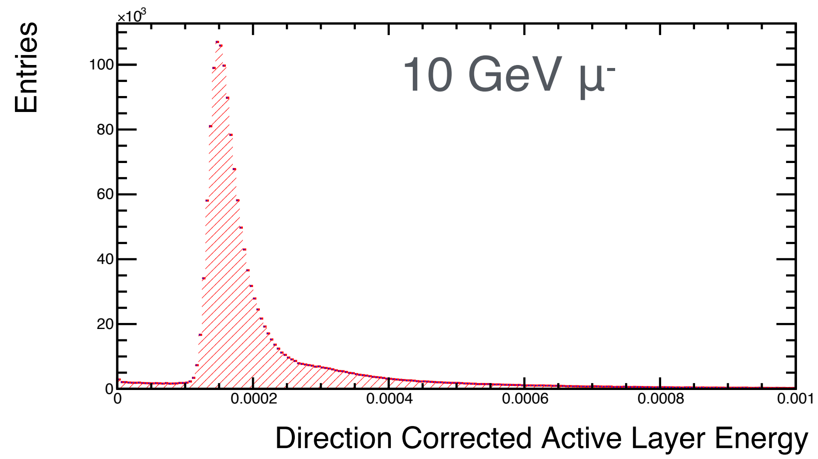
- What does the calibration do?
  1. To ensure that the raw energy estimators for calorimeter hits are accurate. This is essential as these calorimeter hits form PFOs and so will determine physics performance of the detector.
  2. To ensure that the MIP scale in the digitiser and PandoraPFA are accurately set. This is important for determining a physical energy scale for the application of cuts at both stages.
  3. To ensure that the electromagnetic and hadronic energy scales are properly distinguished at PFO level. This is essential as electromagnetic and hadronic showers deposit energy within the detector differently and so shouldn't be treated using the same energy estimators.

- How does the procedure work?
  - By using the PandoraAnalysis toolkit. The toolkit contains a series of executables designed to run over the output root files produced from the PandoraAnalysis processor, it is possible to extract the relevant calibration parameters used in the steering of the digitiser and PandoraPFA processors.
  - The calibration procedure is run on samples (~50,000 events) of 10 GeV μ^-, 10 GeV γ and 20 GeV K^0_L.
  - This is an iterative procedure and so you may need to run the executable more than once.
  - A brief overview of the outputs will be given here, but for full details please see the documentation built with the PandoraAnalysis package www.github.com/PandoraPFA/LCPandoraAnalysis.

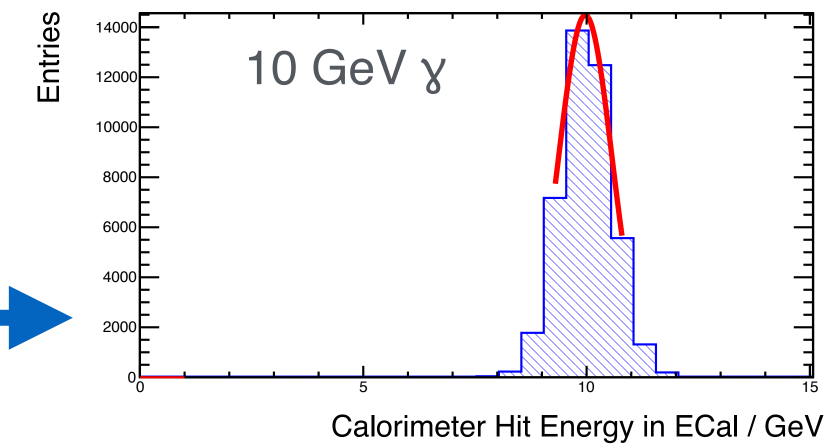
Calibration - Overview

UNIVERSITY OF CAMBRIDGE

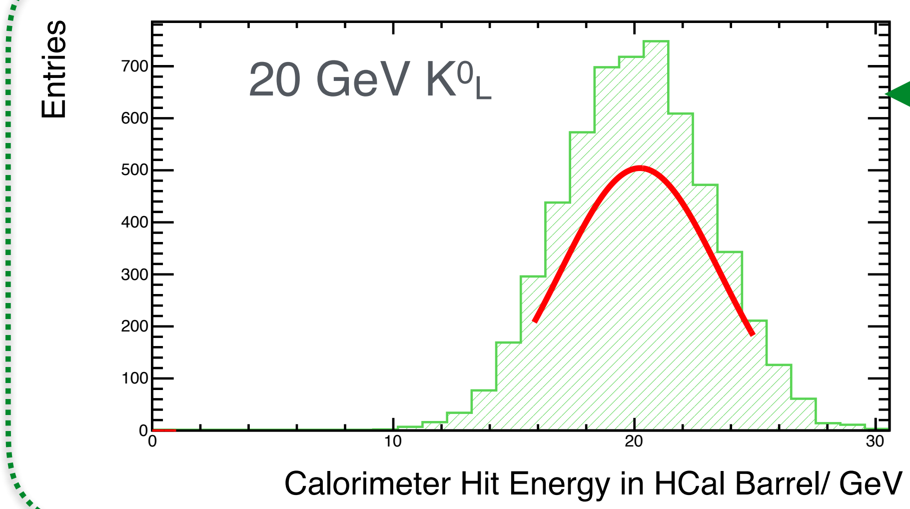
1) Set MIP Scale Digitiser



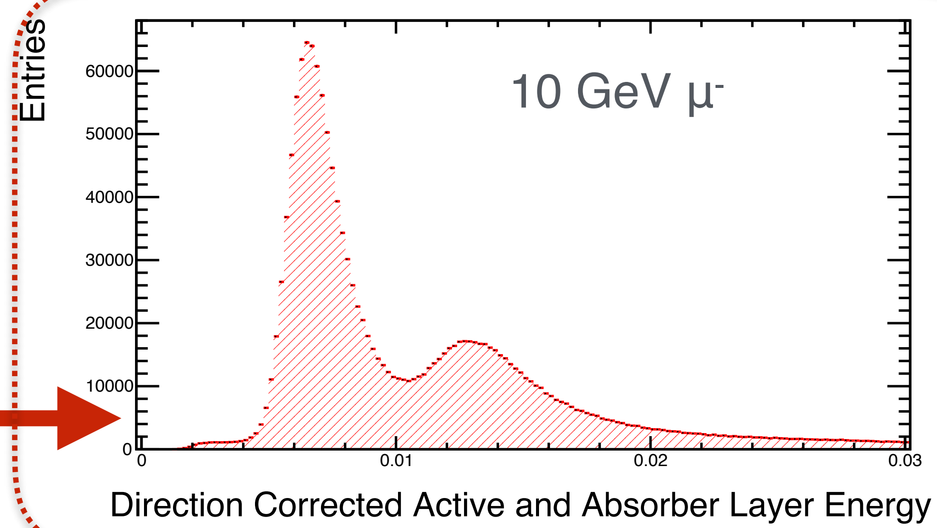
2) Set Digitsation Constants in ECal



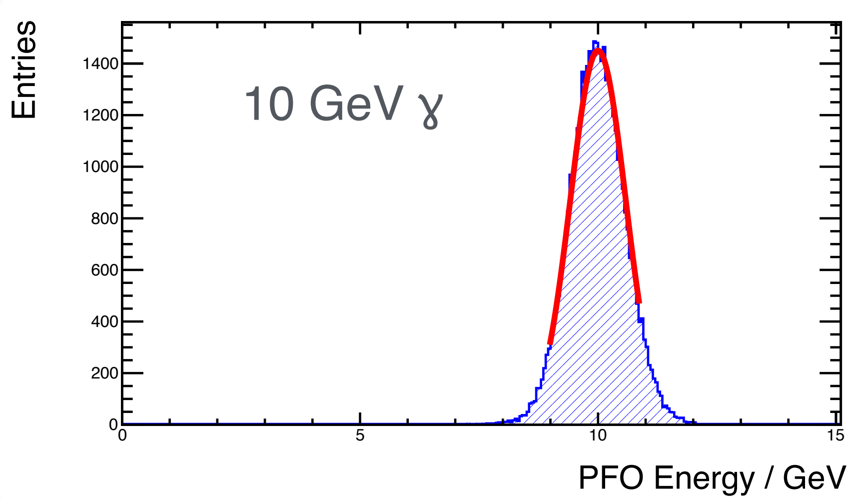
3) Set Digitsation Constants in HCal



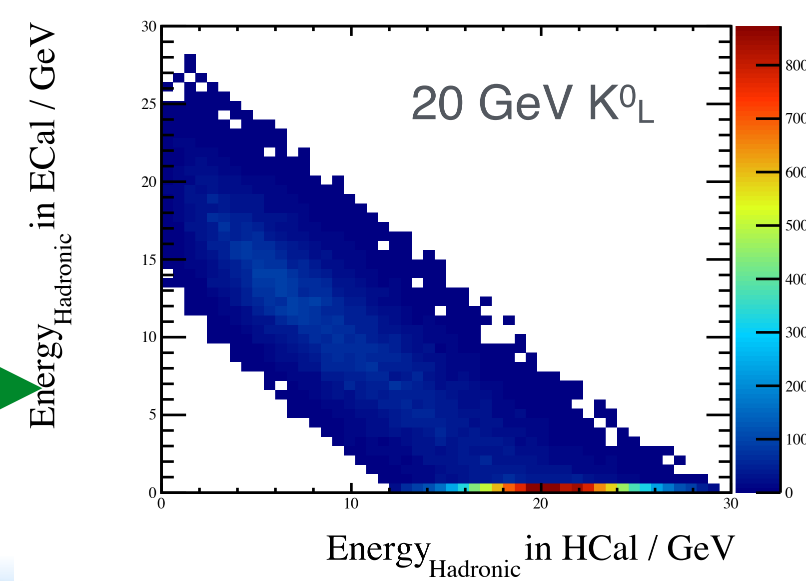
4) Set MIP Scale in PandoraPFA



5) Set Electromagnetic Energy Scale in PandoraPFA

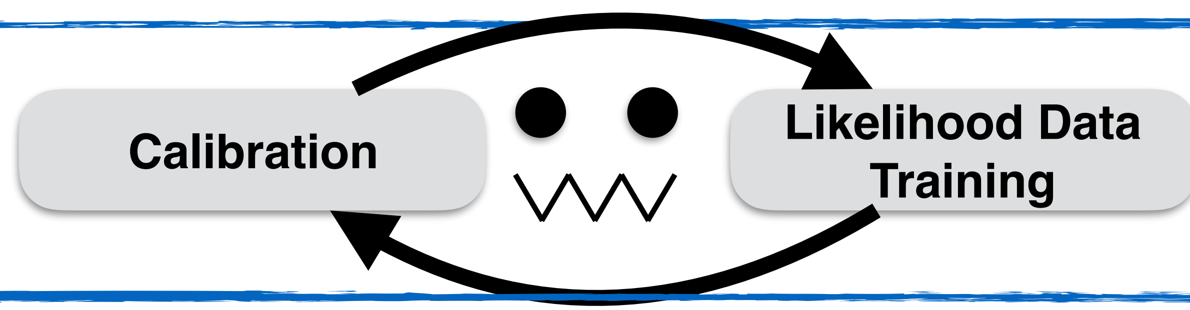


6) Set Hadronic Energy Scale in PandoraPFA



Calibration - Photon Likelihood Data

- The best performance observed in PandoraPFA used likelihood data for the identification of photons (PandoraSettingsDefault.xml).
- There's no guarantee that the default likelihood data, trained for the ILD ECal, will be applicable to all detector models.
  - It is essential to retrain photon likelihood data when changing the detector model or reconstruction settings in the ECal.
- As the likelihood data uses PFOs to train on and the calibration uses the likelihood data to produce PFOs there is room for a vicious circle…



- We avoid this issue by iterating once over the calibration procedure.
  - First we calibrate using a non-standard reconstruction (PandoraSettingsMuon.xml), which uses no likelihood data.
  - Then we train the likelihood data on this output.
  - Then we calibrate using the standard reconstruction (PandoraSettingsDefault.xml), which uses the newly trained likelihood data.
- This works as the PFOs used in calibration of the ECal are 10 GeV ɣ, which have a simple topology, so we wouldn't expect the likelihood data to significantly change the PFO output.

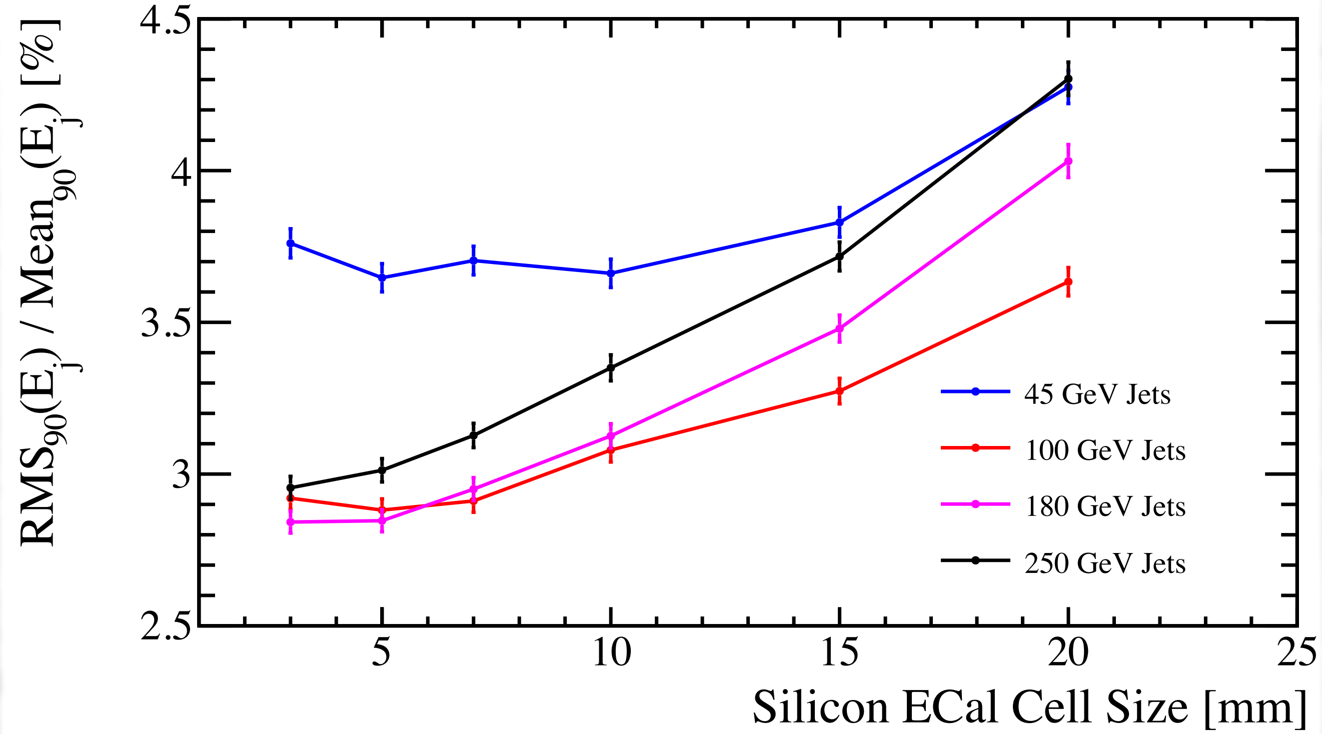
This procedure is used for the following ECal optimisation studies

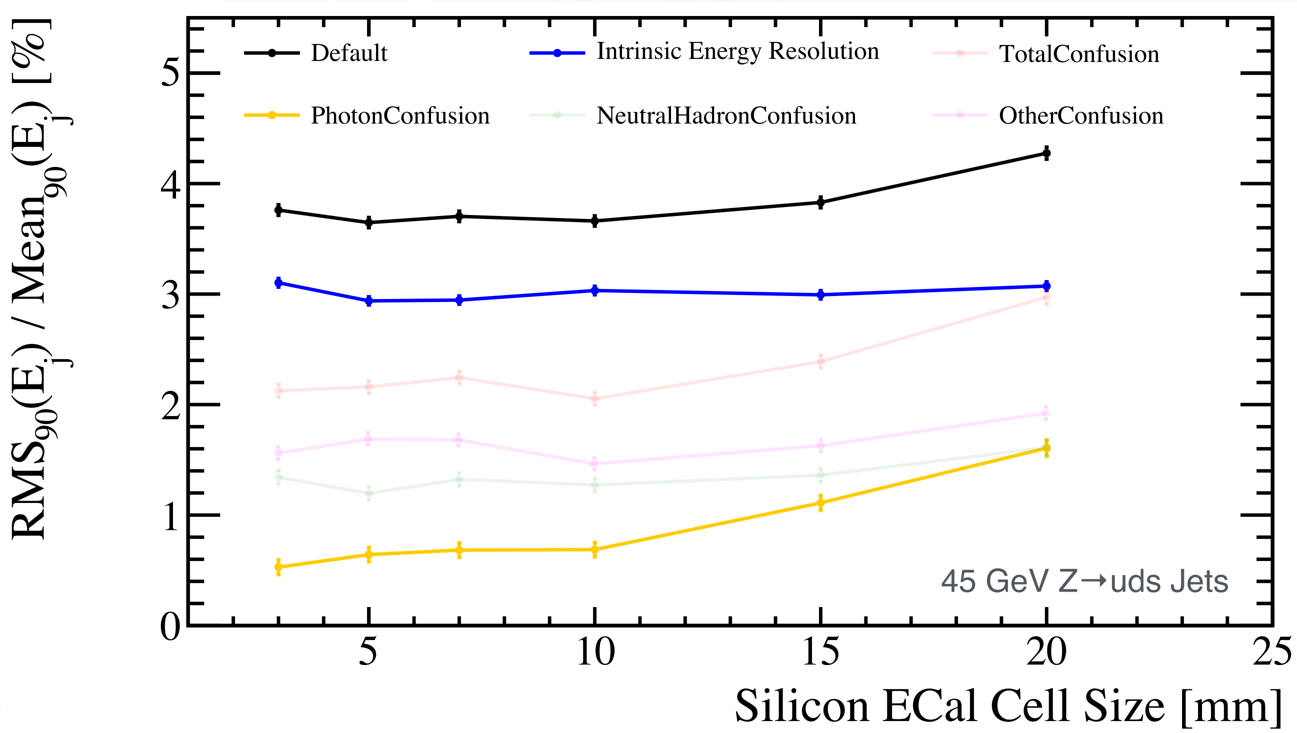
Optimisation Studies

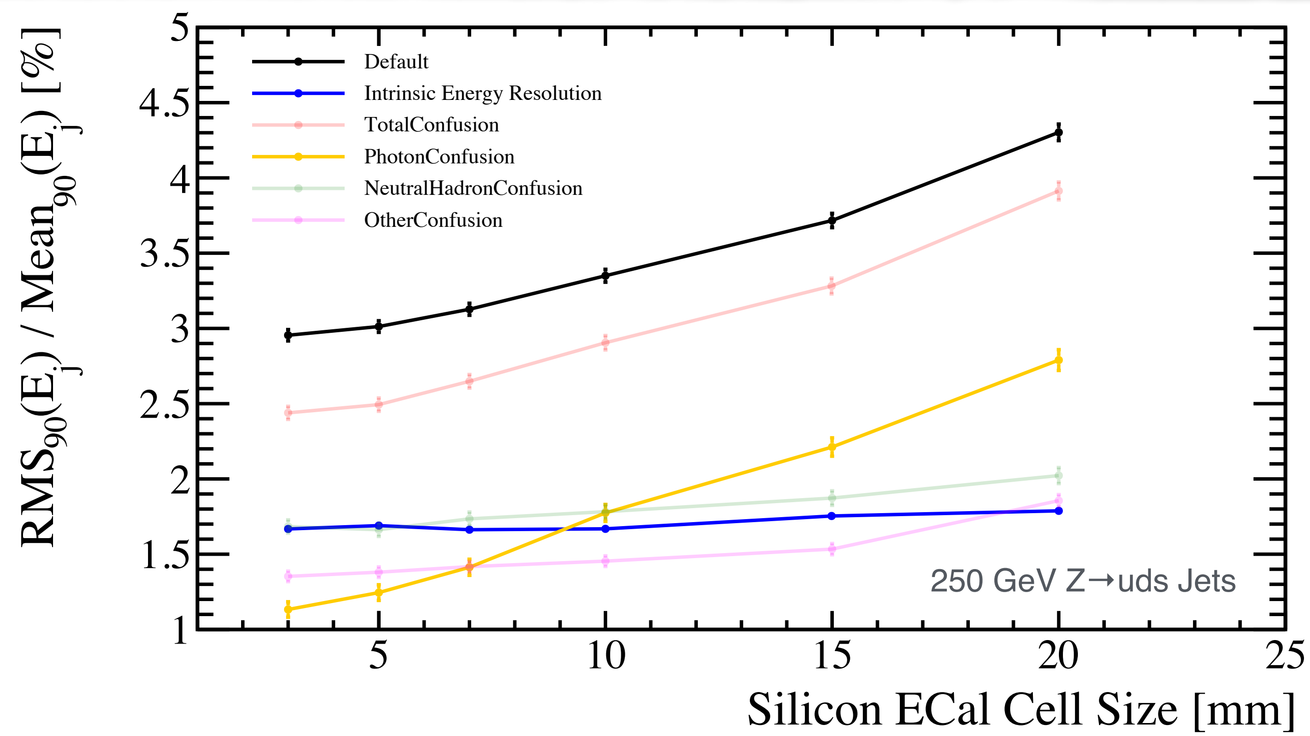
Silicon ECal Transverse Granularity Optimisation

- Jet energy resolution significantly benefits when the cell size is reduced for the silicon ECal option.
- By cheating various parts of the pattern recognition we can see that the improved performance when the ECal cell size is reduce comes primarily from a reduction in photon confusion.
- The intrinsic energy resolution is invariant to changes in ECal cell size as is expected.

HCal Timing Cuts : 100 ns
ECal Timing Cuts : 100 ns
HCal Hadronic Cell Truncation: 1 GeV (Optimal for Default HCal)
Software : ilcsoft_v01-17-07, including PandoraPFA v02-00-00
Digitiser : ILDCaloDigi, realistic ECal and HCal digitisation options enabled
Calibration : PandoraAnalysis toolkit v01-00-00



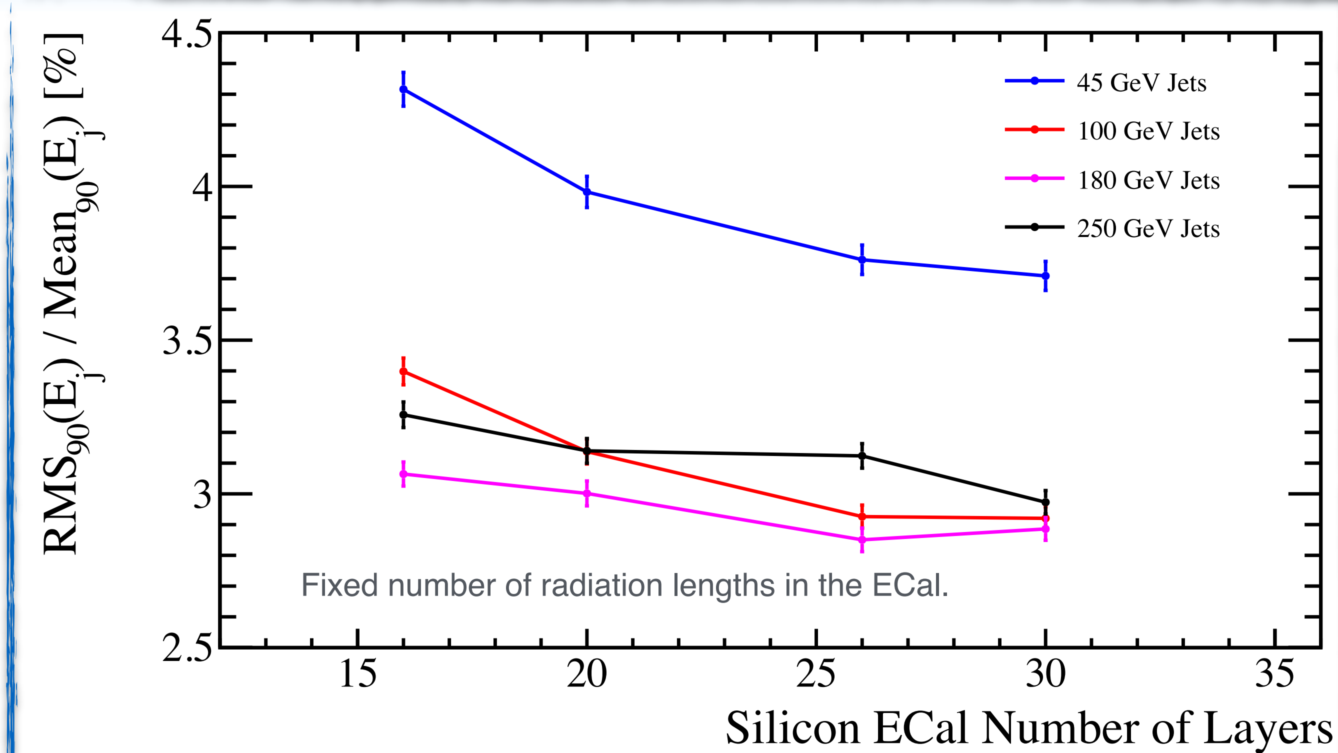


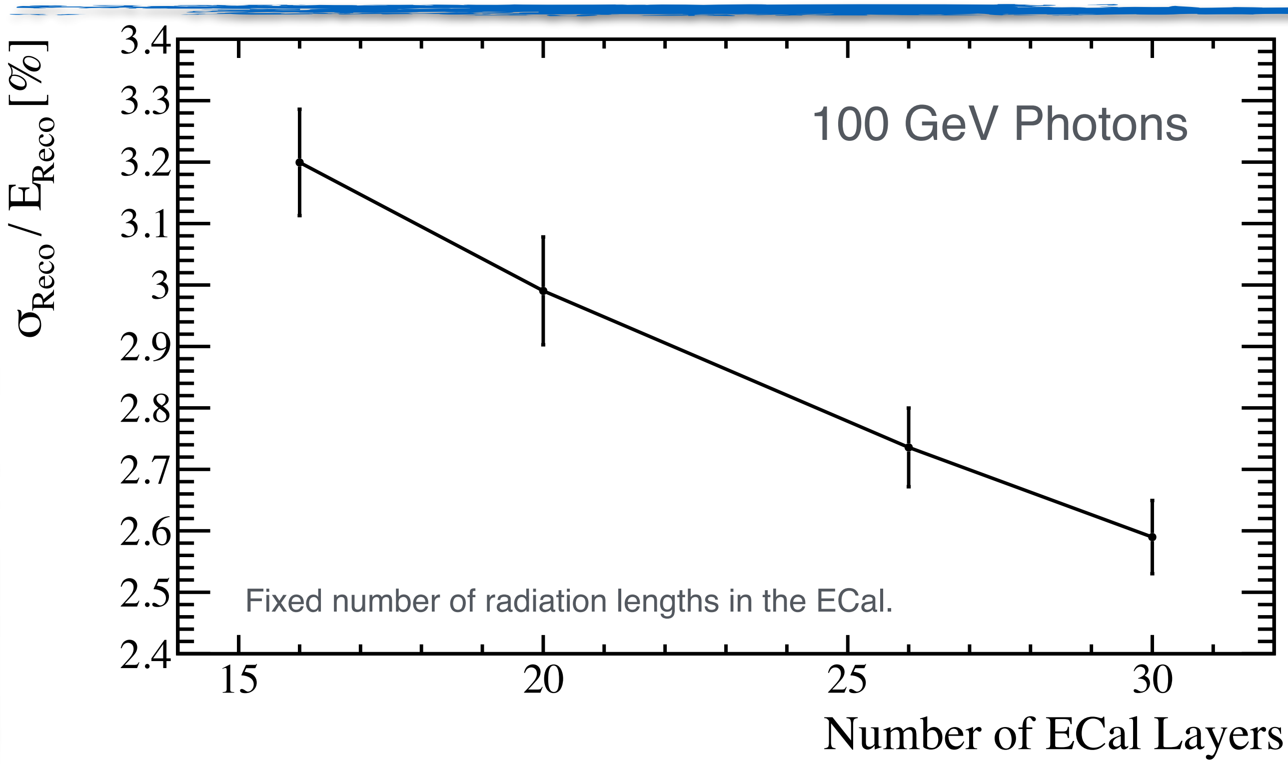


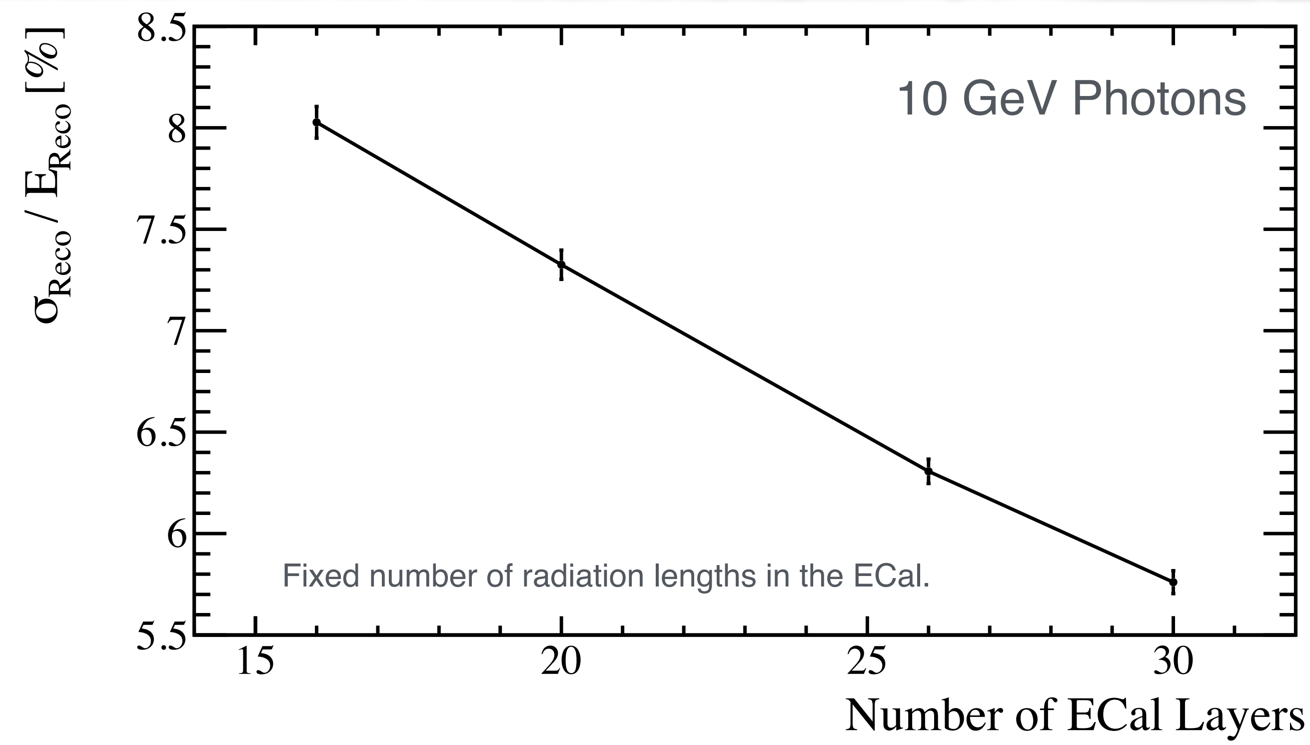
Silicon ECal Longitudinal Granularity Optimisation

- Jet energy resolution and single particle energy resolution benefit from increasing the number of layers for the silicon ECal option.
- The most significant gain in jet energy resolution is found for low energy jets where we are dominated by the intrinsic energy resolution of the calorimeters.

HCal Timing Cuts : 100 ns
ECal Timing Cuts : 100 ns
HCal Hadronic Cell Truncation: 1 GeV (Optimal for Default HCal)
Software : ilcsoft_v01-17-07, including PandoraPFA v02-00-00
Digitiser : ILDCaloDigi, realistic ECal and HCal digitisation options enabled
Calibration : PandoraAnalysis toolkit v01-00-00



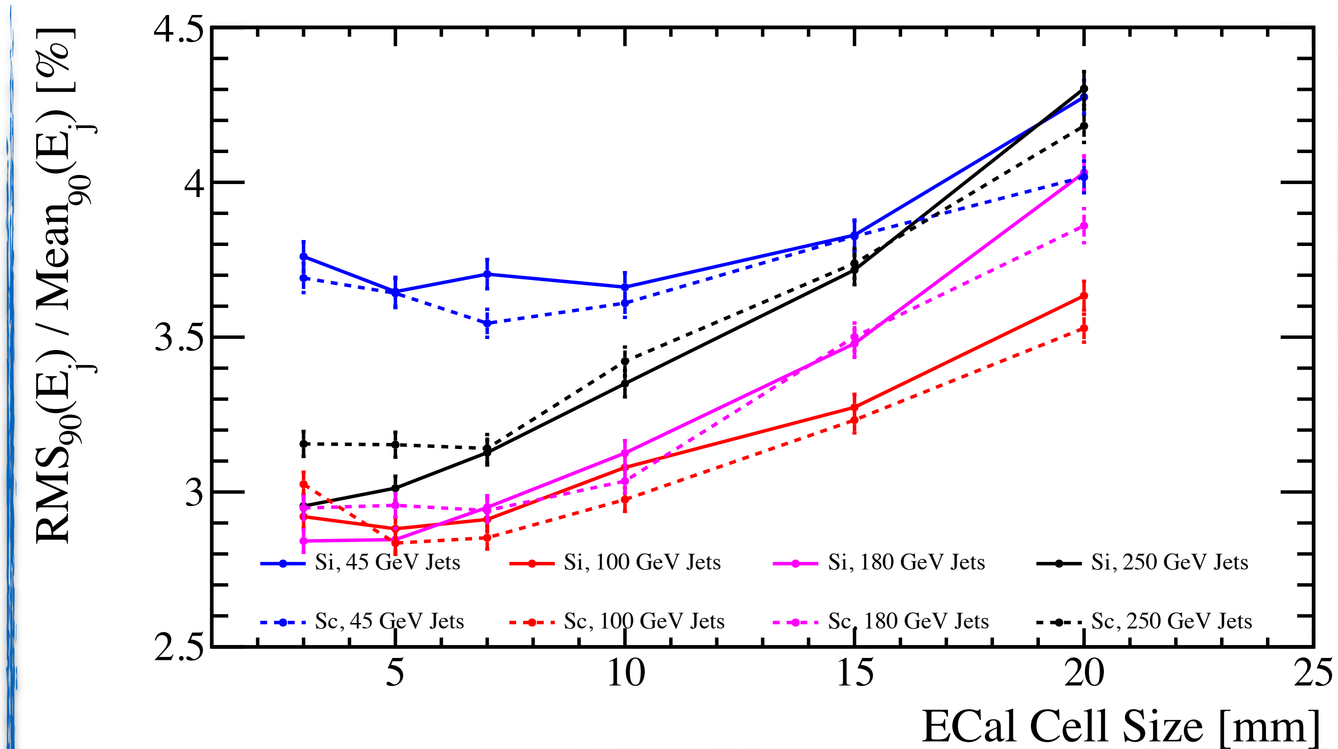


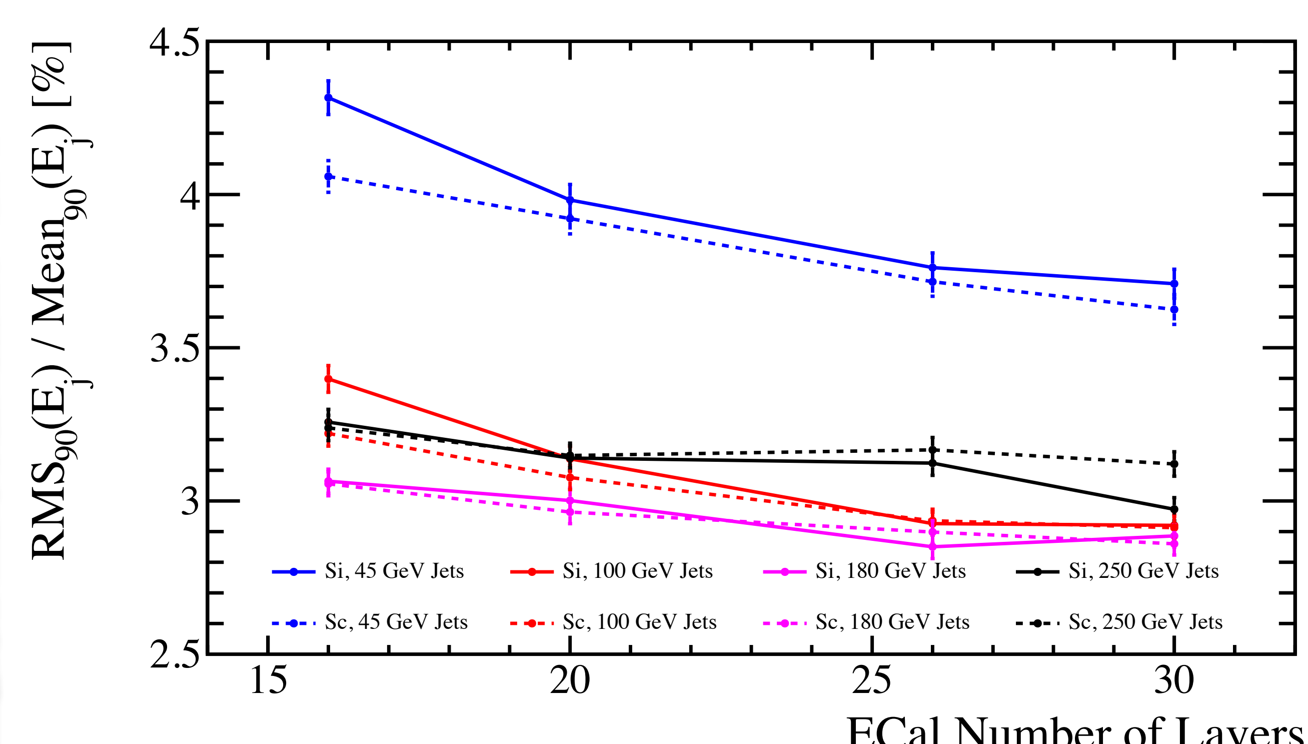


Silicon vs Scintillator ECal

- Now we consider a comparison of the silicon and scintillator ECal models.
- To begin with we look at the effects of varying the longitudinal and transverse granularities of the two ECal models using the jet energy resolution as the figure of merit.
- As you can see there is similar performance in terms of jet energy resolution for both the silicon and scintillator ECal models for the various detector models considered.
- There is no clear preferred option here based on this data.

HCal Timing Cuts : 100 ns
ECal Timing Cuts : 100 ns
HCal Hadronic Cell Truncation: 1 GeV (Optimal for Default HCal)
Software : ilcsoft_v01-17-07, including PandoraPFA v02-00-00
Digitiser : ILDCaloDigi, realistic ECal and HCal digitisation options enabled
Calibration : PandoraAnalysis toolkit v01-00-00



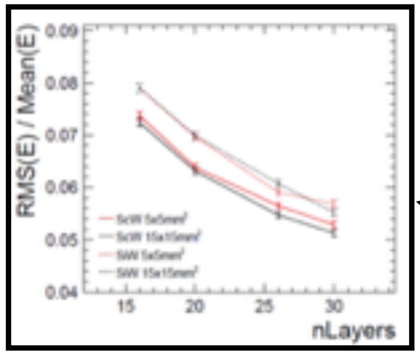


Silicon vs Scintillator ECal

- The dependancy of the single particle energy resolution on the longitudinal segmentation of the ECal was examined for both low and high energy photon events.

- For the 10 GeV samples we find the scintillator ECal offers better energy resolution. This is consistent with the results previously shown for ECal optimisation studies and is explained by the larger sampling fraction in the scintillator ECal.

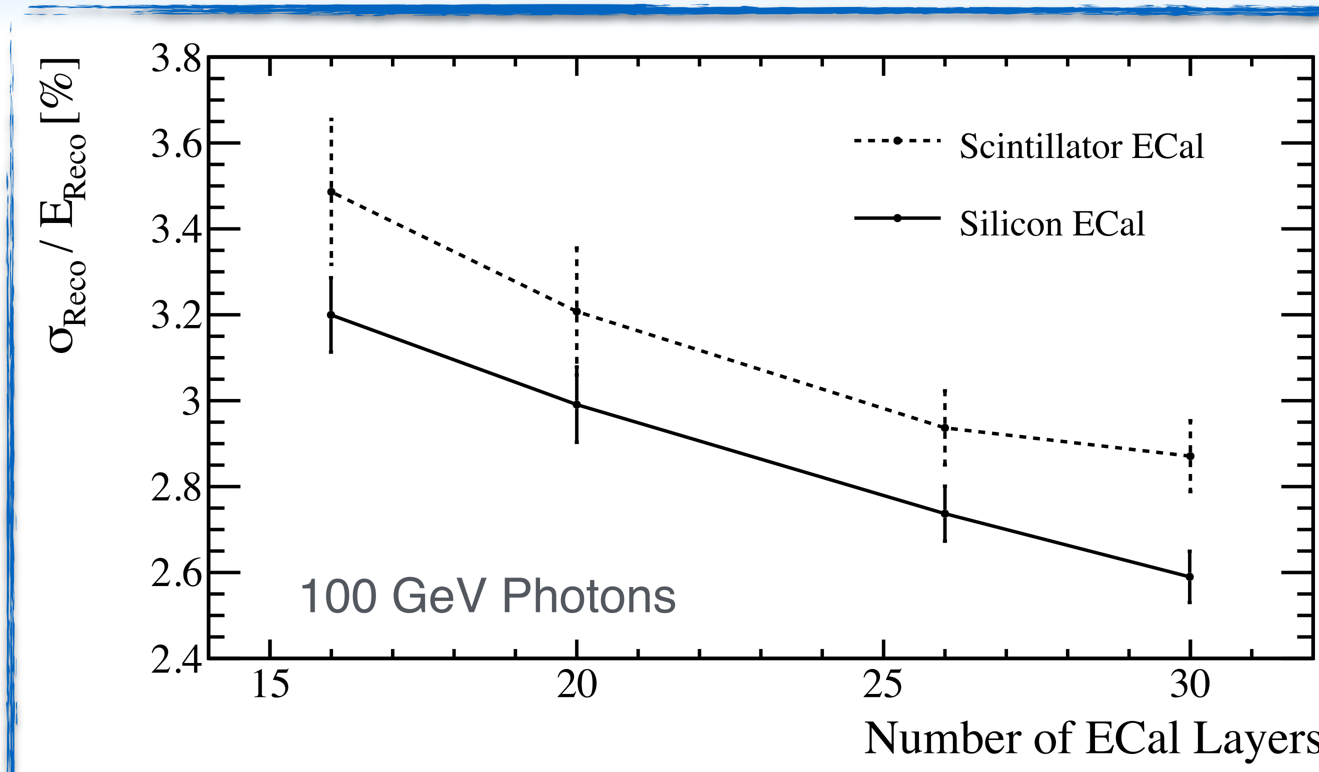
Thanks to J. Marshall

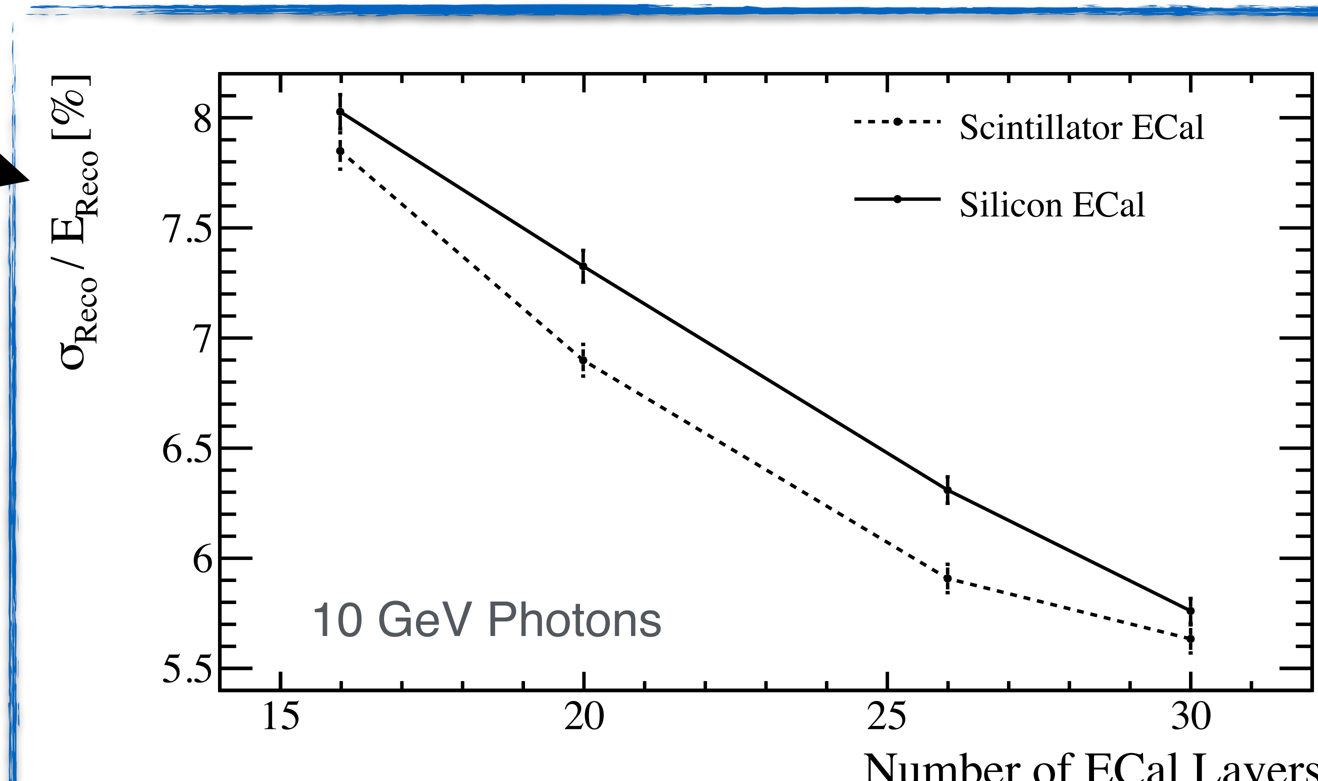


Consistent Results

- However, at high energies (not considered in previous studies) we find that silicon offers better energy resolution. This reverse in trend at high energies may be due to new realistic digitisation effects not simulated in the previous studies.

HCal Timing Cuts : 100 ns
ECal Timing Cuts : 100 ns
HCal Hadronic Cell Truncation: 1 GeV (Optimal for Default HCal)
Software : ilcsoft_v01-17-07, including PandoraPFA v02-00-00
Digitiser : ILDCaloDigi, realistic ECal and HCal digitisation options enabled
Calibration : PandoraAnalysis toolkit v01-00-00



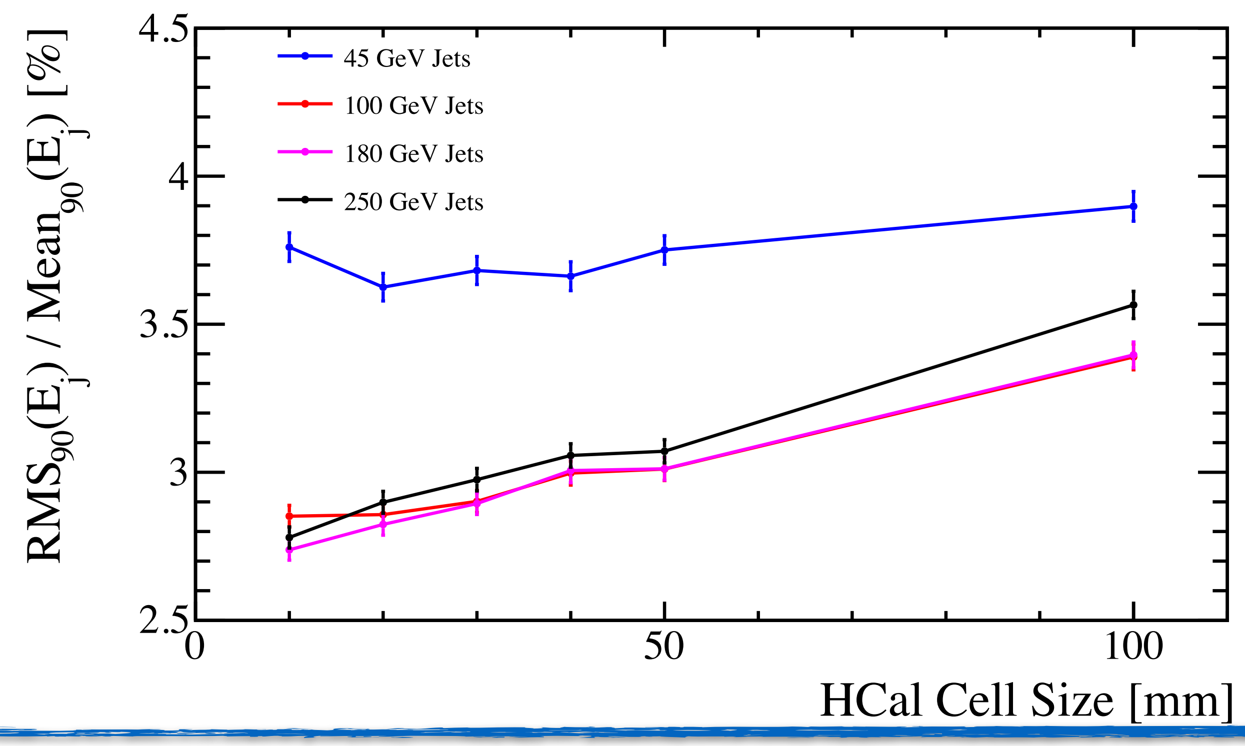


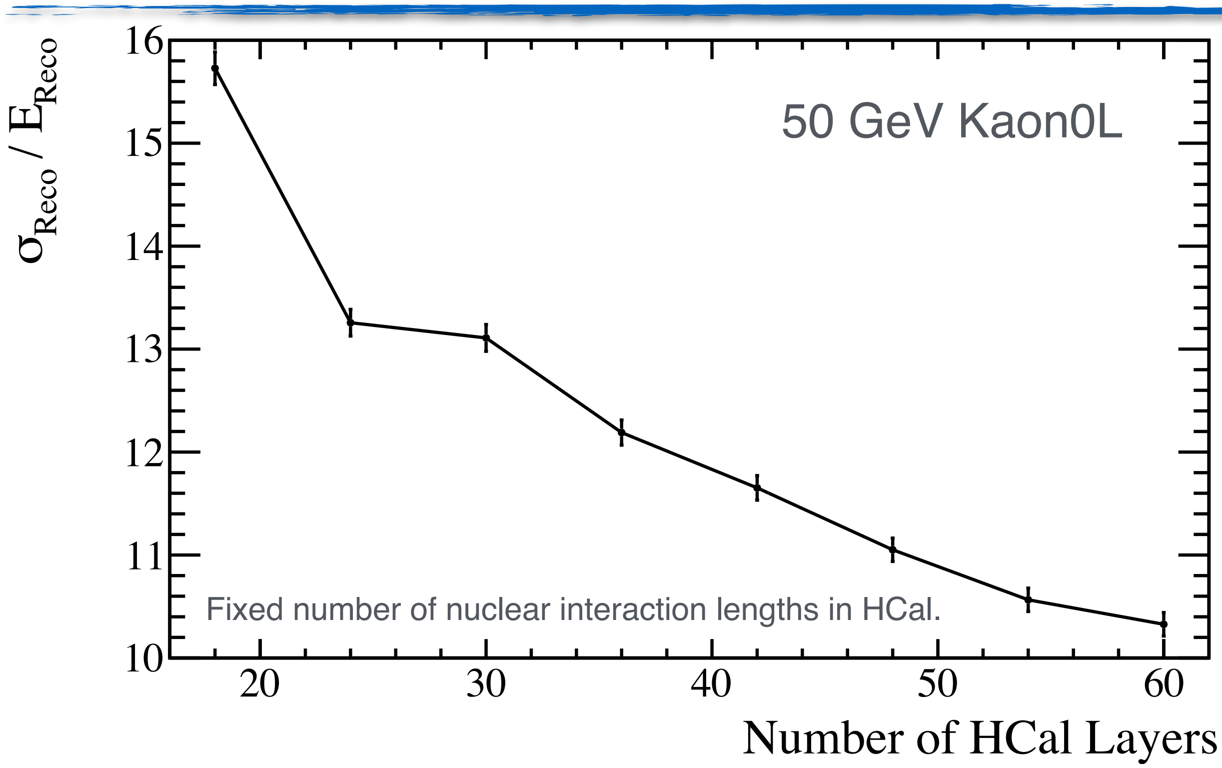
HCal Optimisation

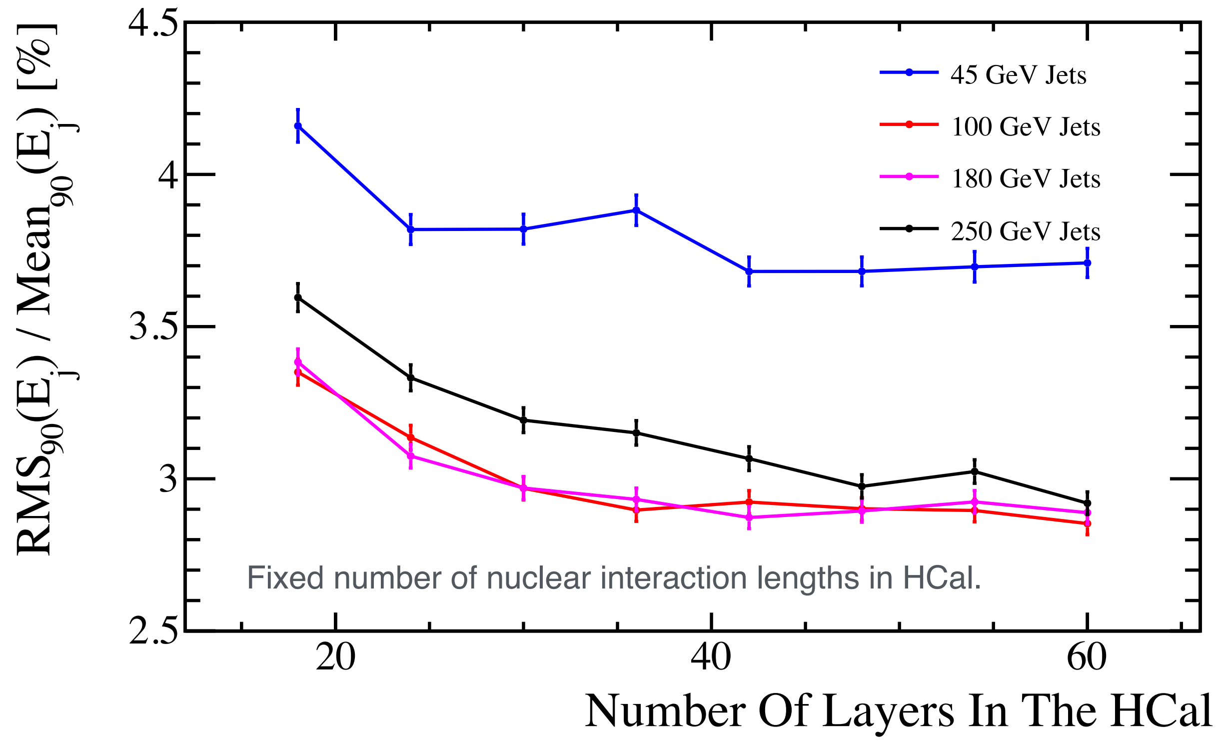
- This updated study, combined with the HCal optimisation studies presented at LCWS15, represents a complete unified optimisation study for the ILD calorimeters.
- A summary of the key results for the HCal optimisation studies shown at LCWS15 is presented here.

See backup slides for full set of results.

HCal Timing Cuts : 100 ns
ECal Timing Cuts : 100 ns
HCal Hadronic Cell Truncation: Optimised for each detector model
Software : ilcsoft_v01-17-07, including PandoraPFA v02-00-00
Digitiser : ILDCaloDigi, realistic ECal and HCal digitisation options enabled
Calibration : PandoraAnalysis toolkit v01-00-00



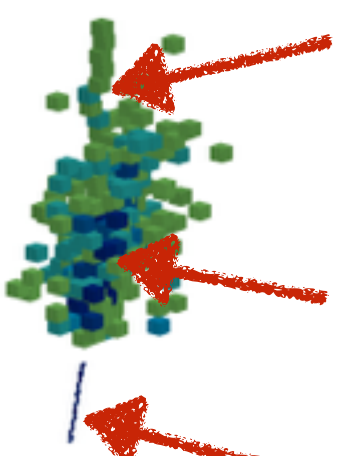


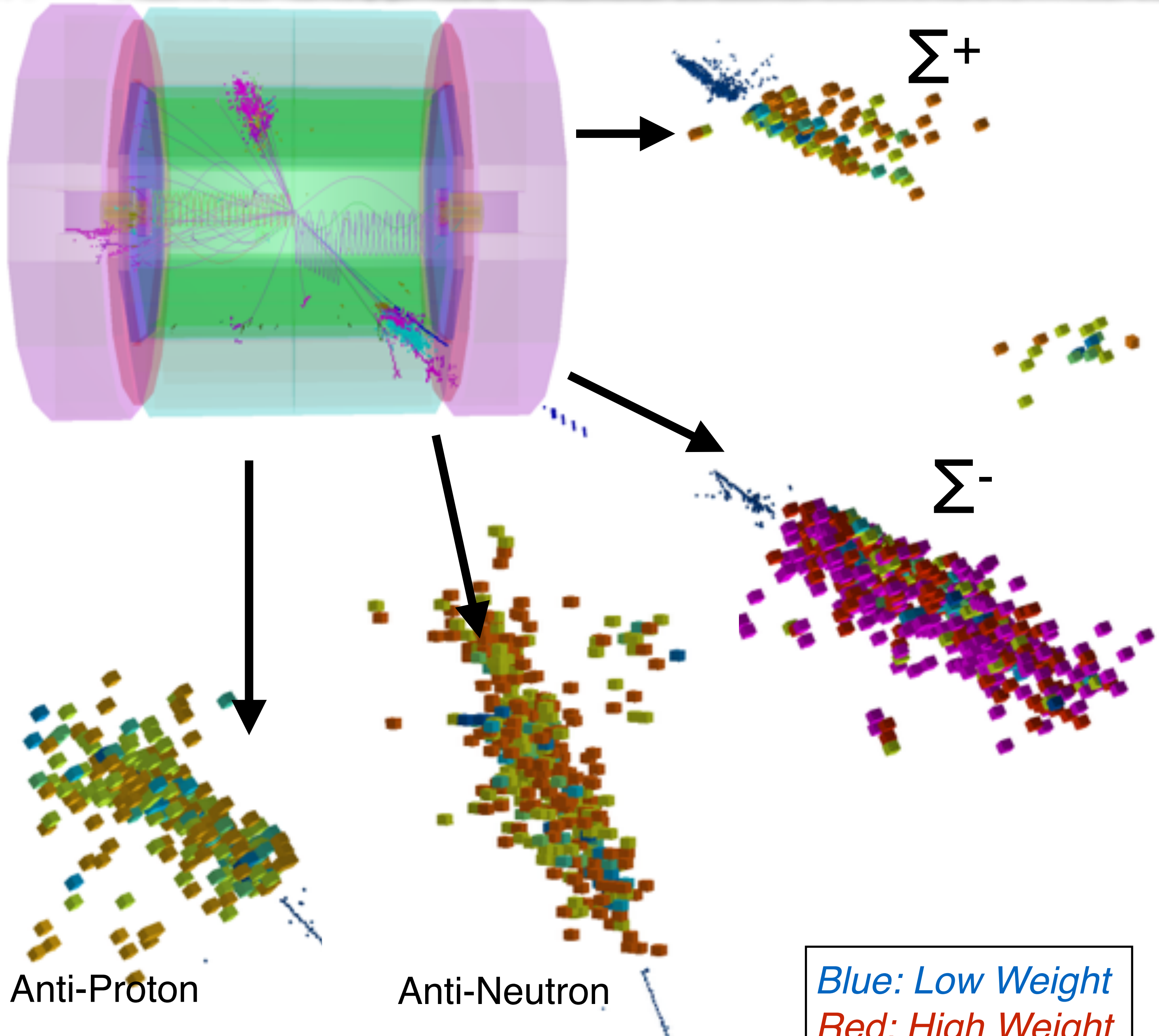


Software Compensation

Software Compensation

- Goal: Improve the energy estimators for hadronic clusters via a reweighting technique based on hit energy density.
- This is to compensate for the "invisible" energy component of hadronic showers (from low energy neutrons, nuclear binding energy losses etc.).
- In the current implementation software compensation is applied to clusters in the HCal only.





Event display where colour indicates the weight applied in software compensation. Event is a pair of 250 GeV Z_uds jets.

Software Compensation - Training and Calibration

- Software compensation works by reweighting the energy of individual calorimeter hits based on it's energy density (ρ).
- The reweighting uses a series of parameters (p_{ij}), which must be trained on the detector model being used.

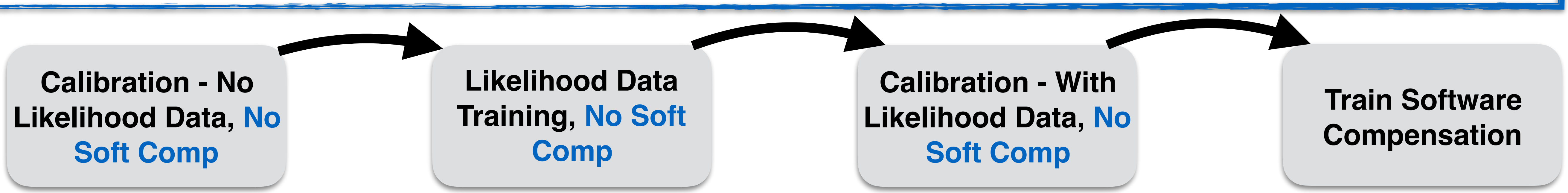
$$\text{Weight} = p_1 \times \exp(p_2 \times \rho) + p_3$$

$$p_1 = p_{11} + p_{12} \times E_{\text{Cluster}} + p_{13} \times E_{\text{Cluster}}^2$$

$$p_2 = p_{21} + p_{22} \times E_{\text{Cluster}} + p_{23} \times E_{\text{Cluster}}^2$$

$$p_3 = \frac{p_{31}}{p_{32} + \exp(p_{33} \times E_{\text{Cluster}})}$$

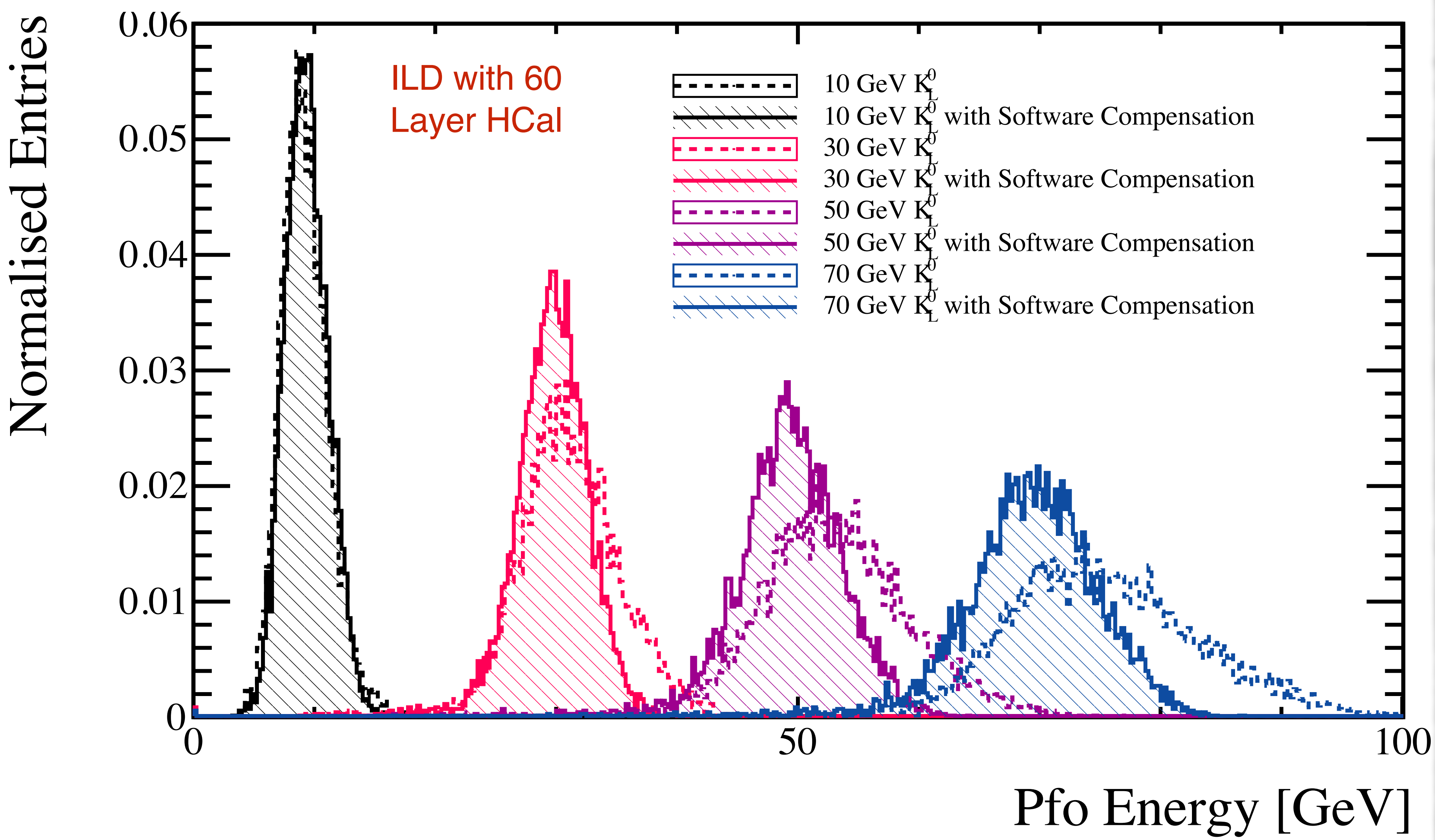
- These parameters are determined by performing a fit of the PFO cluster energy to the MC energy for samples of K^0_L of varying energies. In this case the samples were 10 GeV - 100 GeV in steps of 10 GeV, 10,000 events in each sample.
- The parameters are only guaranteed to be physical up to the highest energy used in training, so software compensation is only applied to cluster with energy less than this value i.e. < 100 GeV.
- In the calibration hierarchy it is sufficient to train software compensation as a final step once the standard calibration procedure has been performed with software compensation disabled.



- The hadronic scale in PandoraPFA now differs from that specified in the calibration, but it is much better!

Software Compensation - Application

As a test the versatility of the software compensation procedure and the code in PandoraPFA, I applied software compensation to the ILD detector, but with a 60 layer HCal (~7λ_I) and ran some high energy jet energy resolutions studies to make it more relevant for CLIC.

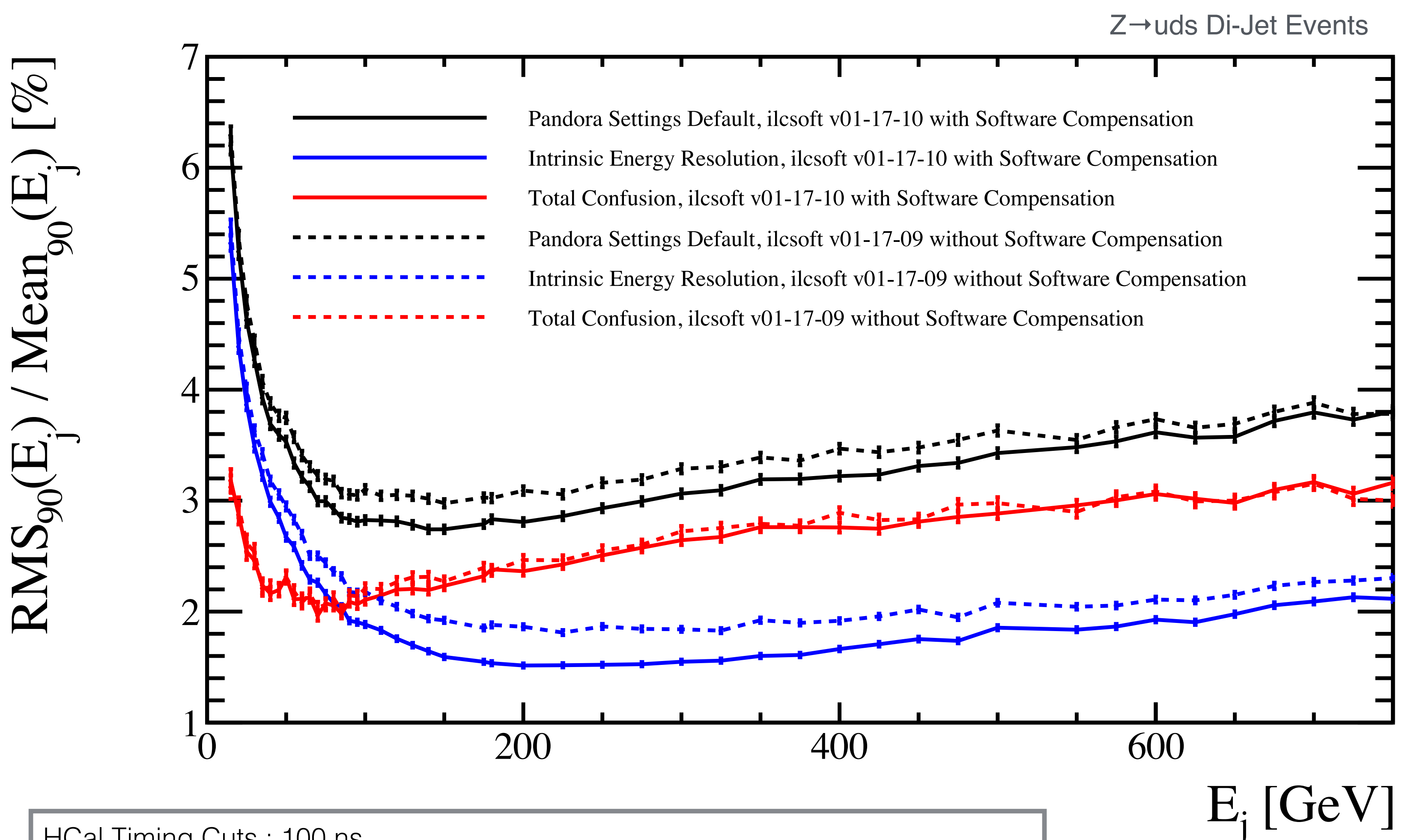


We can see that software compensation improves the energy resolution of K^0_L significantly.

HCal Timing Cuts : 100 ns
ECal Timing Cuts : 100 ns
HCal Hadronic Cell Truncation : Not Applied
Software : ilcsoft_v01-17-10, including PandoraPFA v02-05-00
Digitiser : ILDCaloDigi, realistic ECal and HCal digitisation options enabled
Calibration : PandoraAnalysis toolkit v01-02-00

Software Compensation - Jet Energy Resolution

- Software compensation significantly improves the jet energy resolution at semi-low jet energies in comparison to the previous best performance.

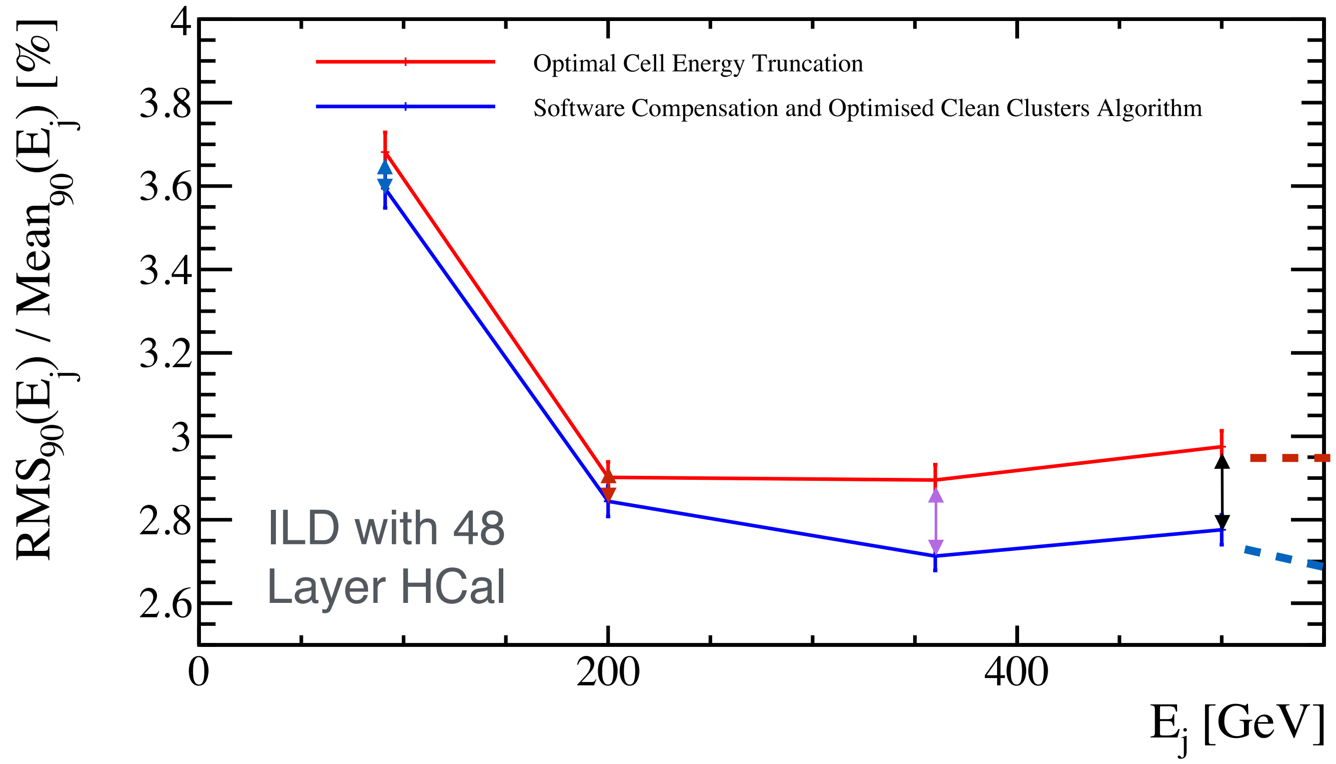


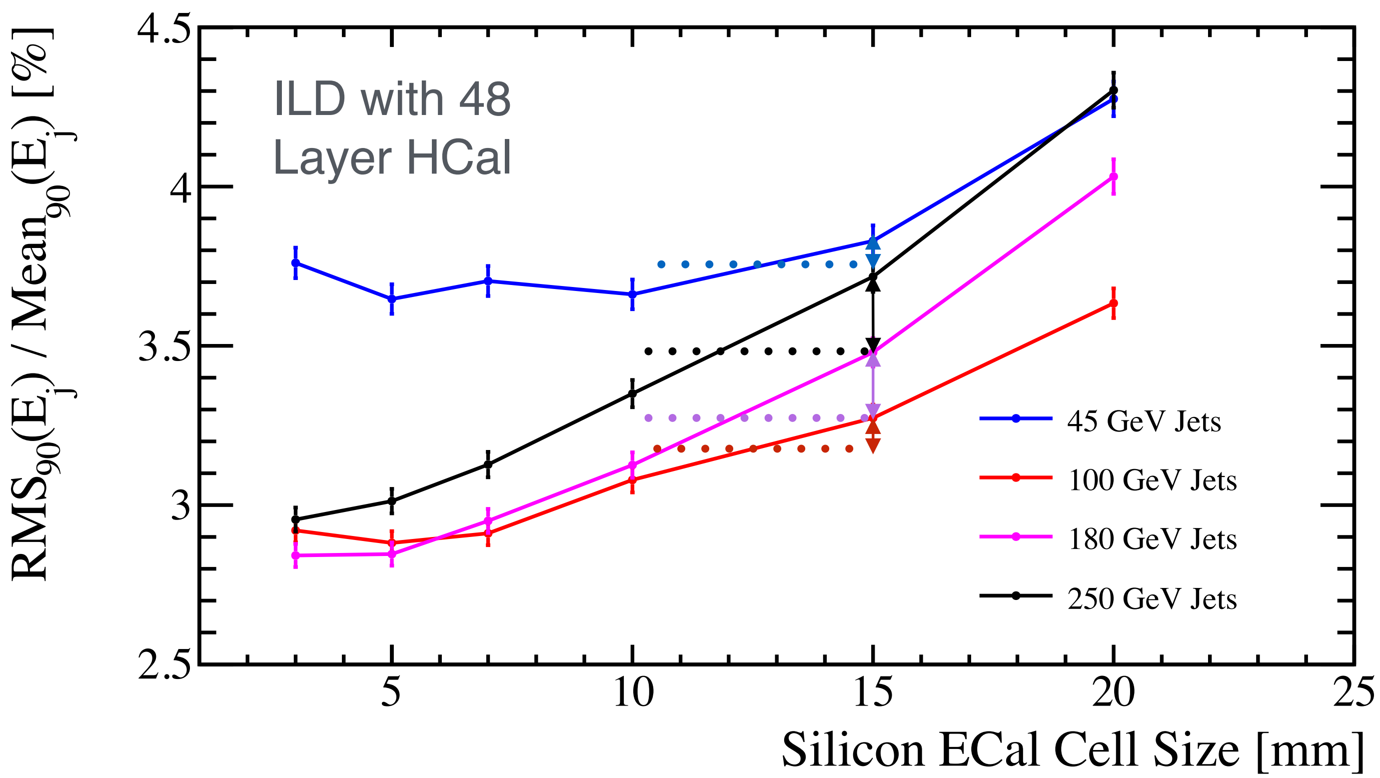
HCal Timing Cuts : 100 ns
ECal Timing Cuts : 100 ns
HCal Hadronic Cell Truncation : Not Applied
Software : Mixed, see legend
Digitiser : ILDCaloDigi, realistic ECal and HCal digitisation options enabled
Calibration : PandoraAnalysis toolkit v01-02-00

ILD with 60 Layer HCal

- The dominant change comes from an improvement to the intrinsic energy resolution.
- A secondary improvement comes from a reduction in confusion as pattern recognition becomes easier with better energy estimators.
- The improvement is smaller at higher energies, which is most likely because software compensation is only trained on samples less than 100 GeV.
- The full potential of software compensation hasn't been fully achieved by a long way!

Software Compensation - In Optimisation Studies Context





HCal Timing Cuts : 100 ns
ECal Timing Cuts : 100 ns
HCal Hadronic Cell Truncation: 1 GeV (Optimal for Default HCal) for optimal cell energy truncation graph and Not Applied for software compensation and optimised clean clusters graph
Software : ilcsoft_v01-17-07, including PandoraPFA v02-00-00
Digitiser : ILDCaloDigi, realistic ECal and HCal digitisation options enabled
Calibration : PandoraAnalysis toolkit v01-00-00

Previous Best Performance (with optimal hadronic energy truncation)

Optimisation of Clean Clusters and Software Compensation. (New Best Performance)

- The gains made by software compensation are comparable to the changes observed when we vary certain detector parameters.
- Takeaway message :
  - We need to have a unified approach in detector optimisation to both software and hardware to get a true measure of physics potential.

Conclusions

- The optimisation studies using jet energy resolution as a metric of detector performance (simulated with Mokka) have been performed.
- The calibration for each detector model was carefully applied allowing us to have confidence in the conclusions we draw from these studies.
- Software compensation is a powerful tool for improving detector performance. It's application shows us that we need to have a unified approach in detector optimisation to both software and hardware to get a true measure of physics potential.

Thank you for your attention!

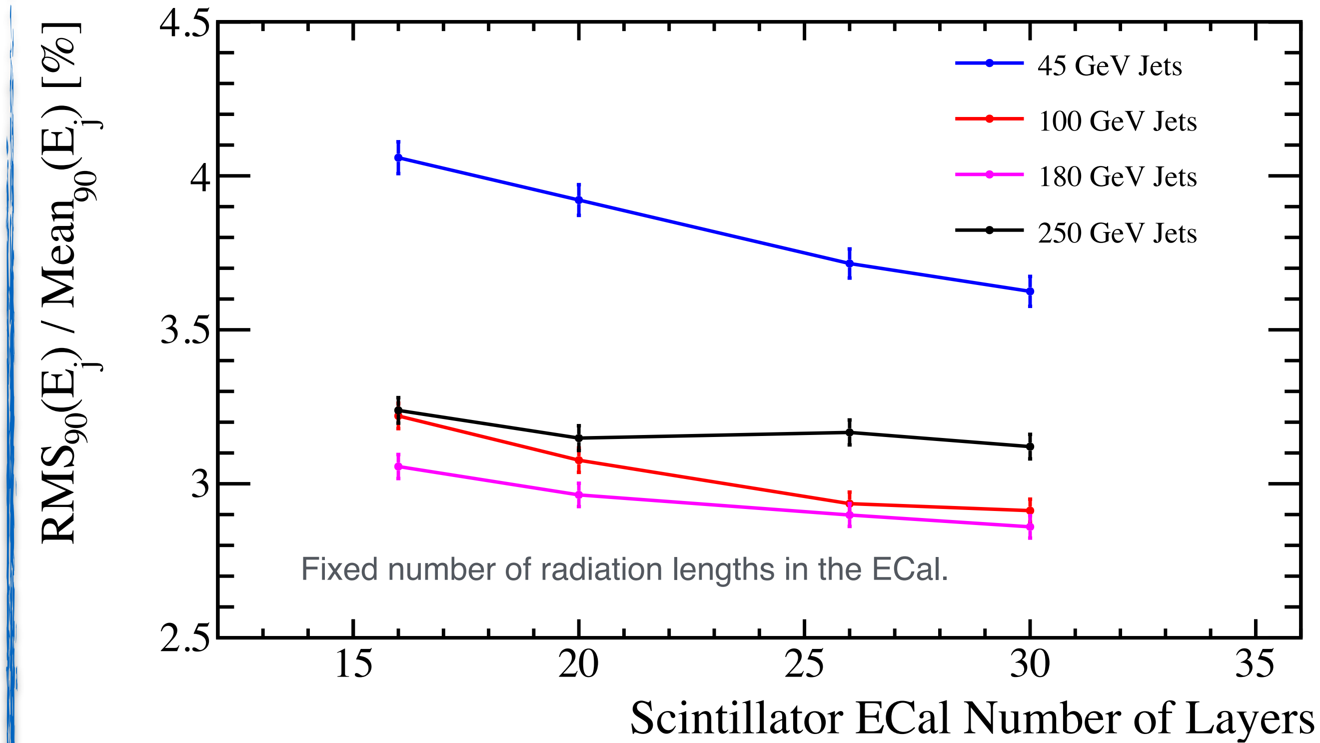
Backup

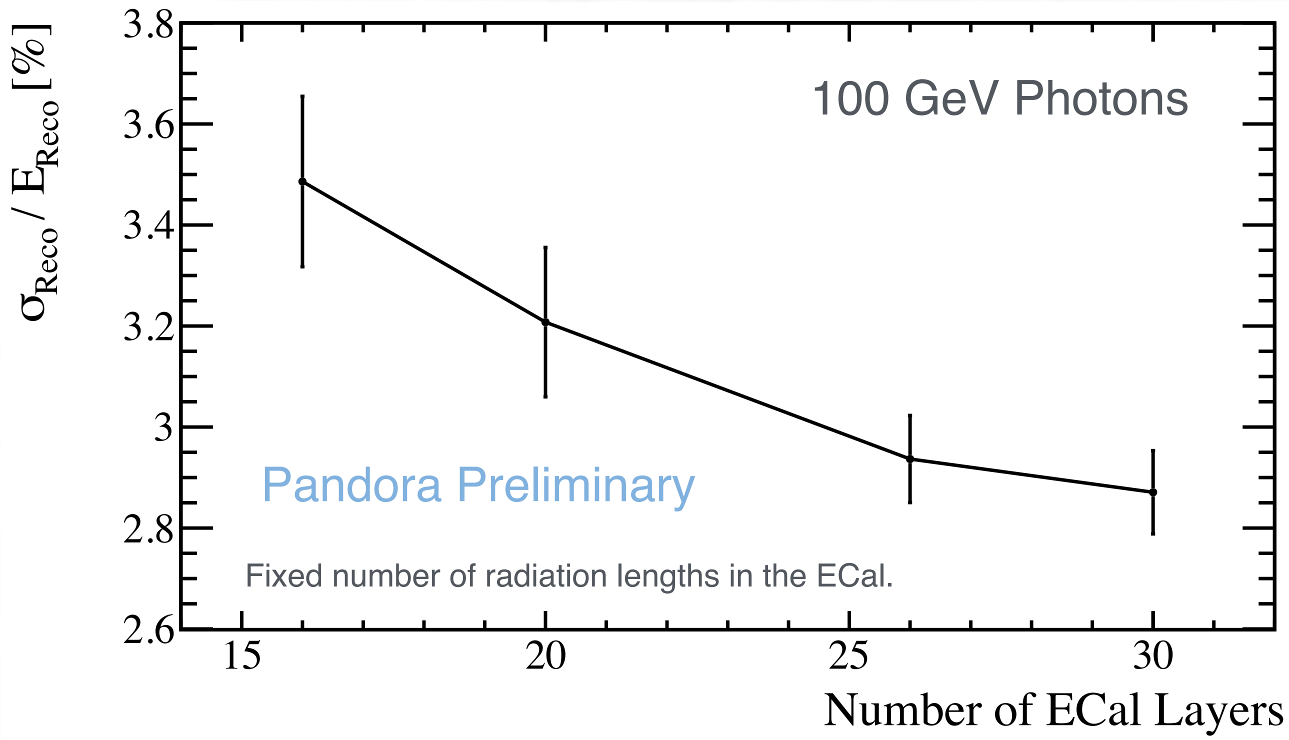
Backup - Calorimeter Optimisation Studies

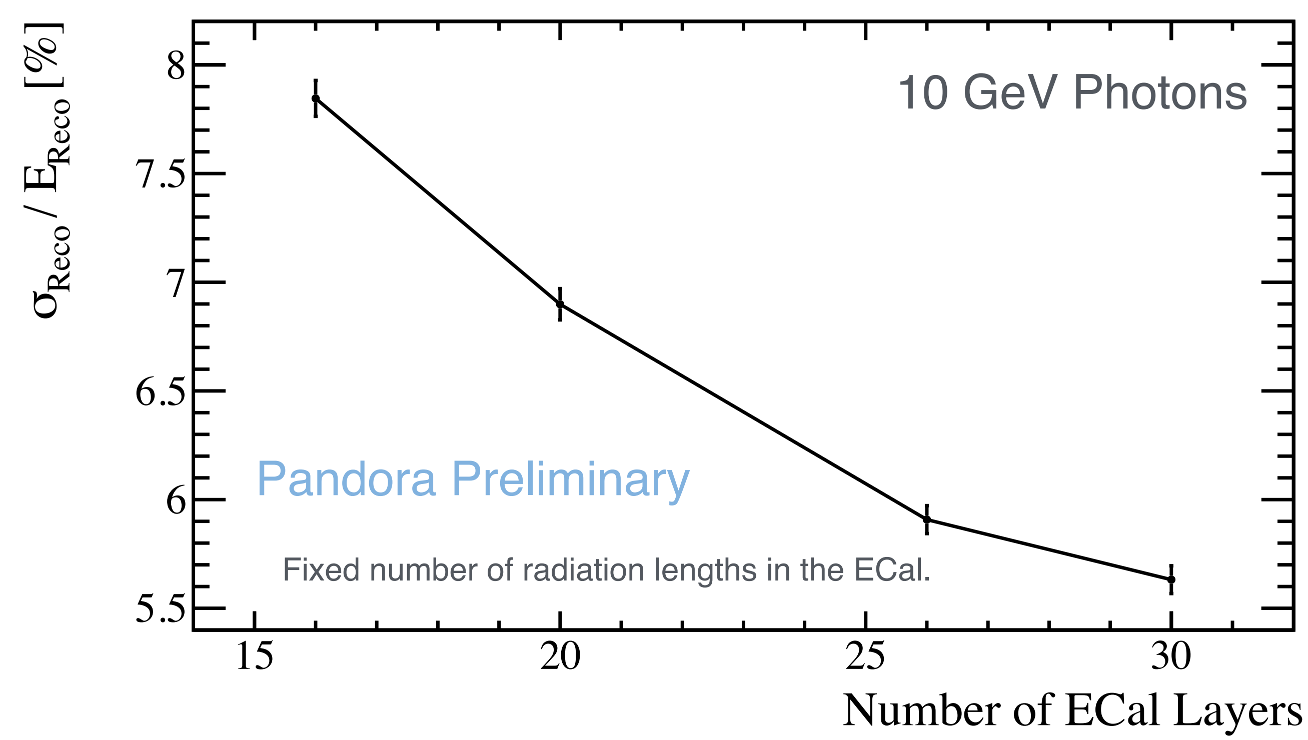
Scintillator ECal Longitudinal Granularity Optimisation

- Jet energy resolution and single particle energy resolution benefit from increasing the number of layers for the scintillator ECal option.
- Again, the most significant gain in jet energy resolution is found for low energy jets where we are dominated by the intrinsic energy resolution of the calorimeters.

HCal Timing Cuts : 100 ns
ECal Timing Cuts : 100 ns
HCal Hadronic Cell Truncation: 1 GeV (Optimal for Default HCal)
Software : ilcsoft_v01-17-07, including PandoraPFA v02-00-00
Digitiser : ILDCaloDigi, realistic ECal and HCal digitisation options enabled
Calibration : PandoraAnalysis toolkit v01-00-00



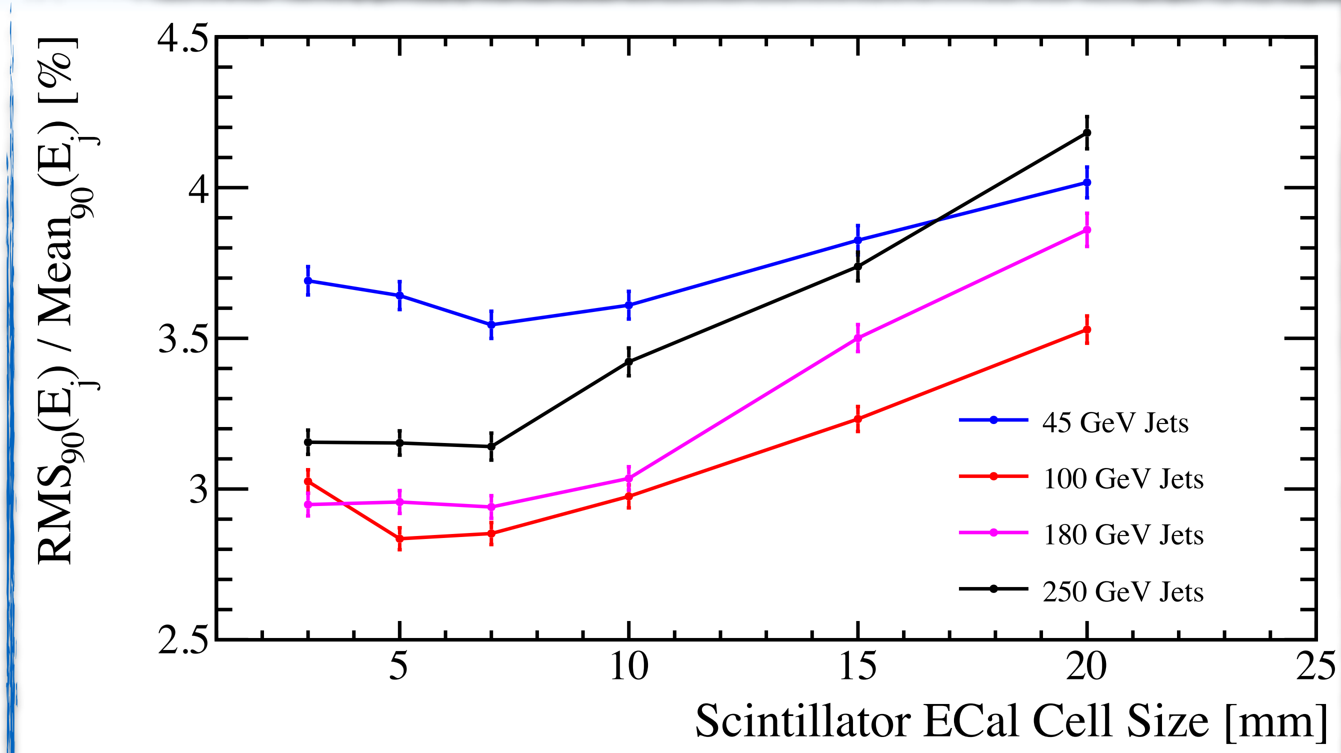


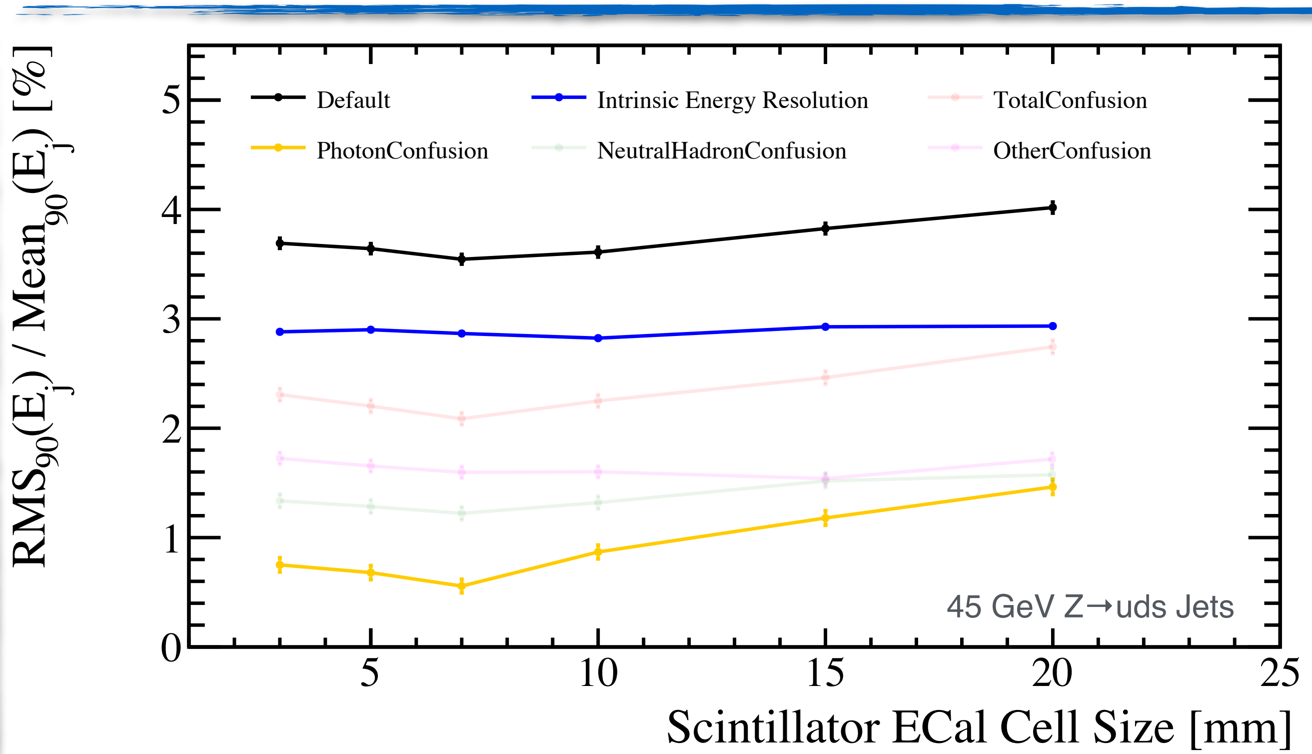


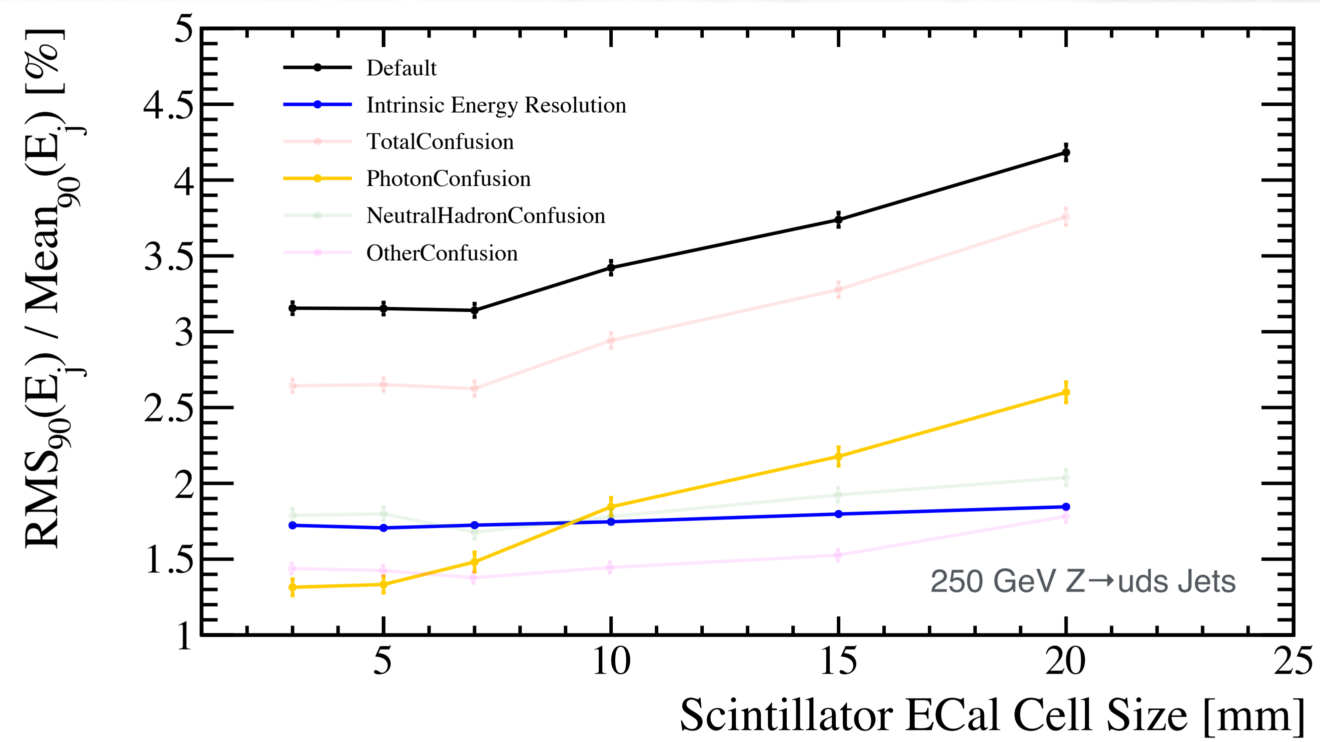
Scintillator ECal Transverse Granularity Optimisation

- Jet energy resolution significantly benefits when the cell size is reduced for the scintillator ECal option.
- Once again we see the improved performance observed by reducing in the ECal cell size is primarily from a reduction in photon confusion.
- Again, the intrinsic energy resolution is invariant to changes in ECal cell size as is expected.

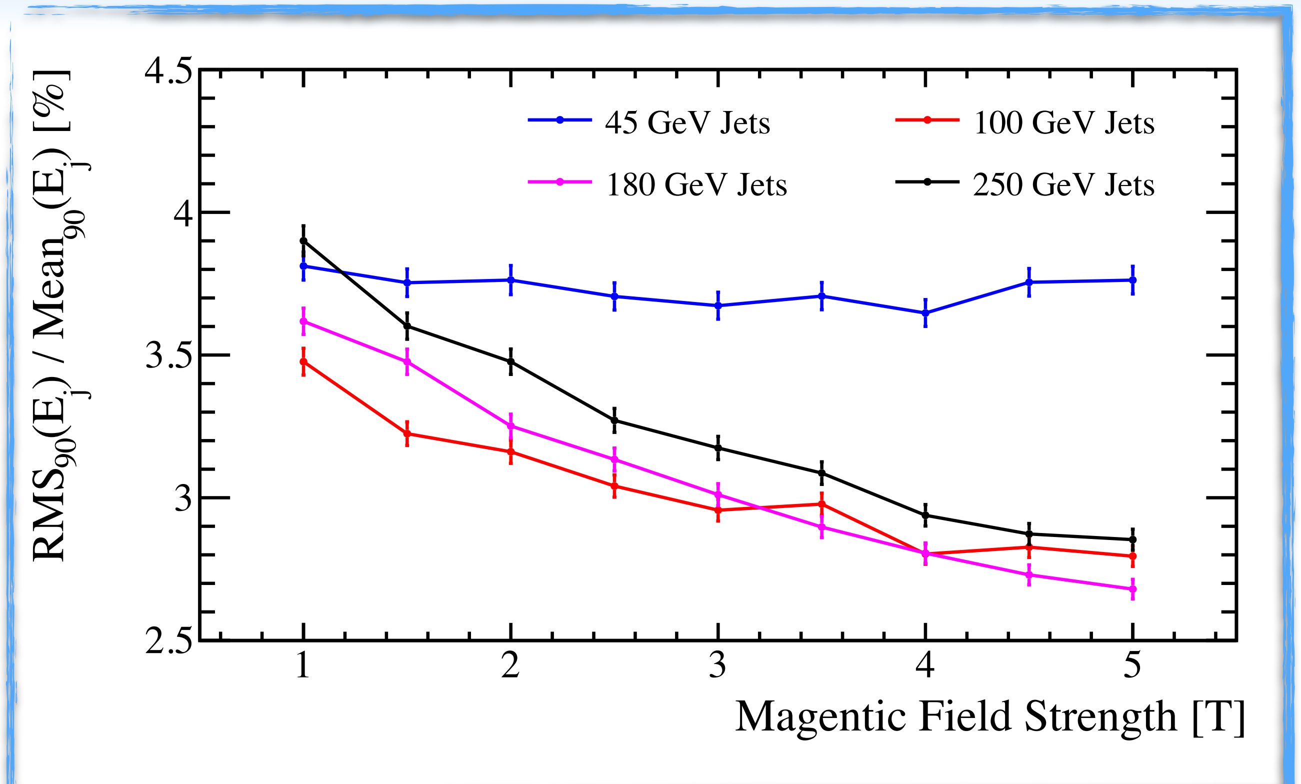
HCal Timing Cuts : 100 ns
ECal Timing Cuts : 100 ns
HCal Hadronic Cell Truncation: 1 GeV (Optimal for Default HCal)
Software : ilcsoft_v01-17-07, including PandoraPFA v02-00-00
Digitiser : ILDCaloDigi, realistic ECal and HCal digitisation options enabled
Calibration : PandoraAnalysis toolkit v01-00-00

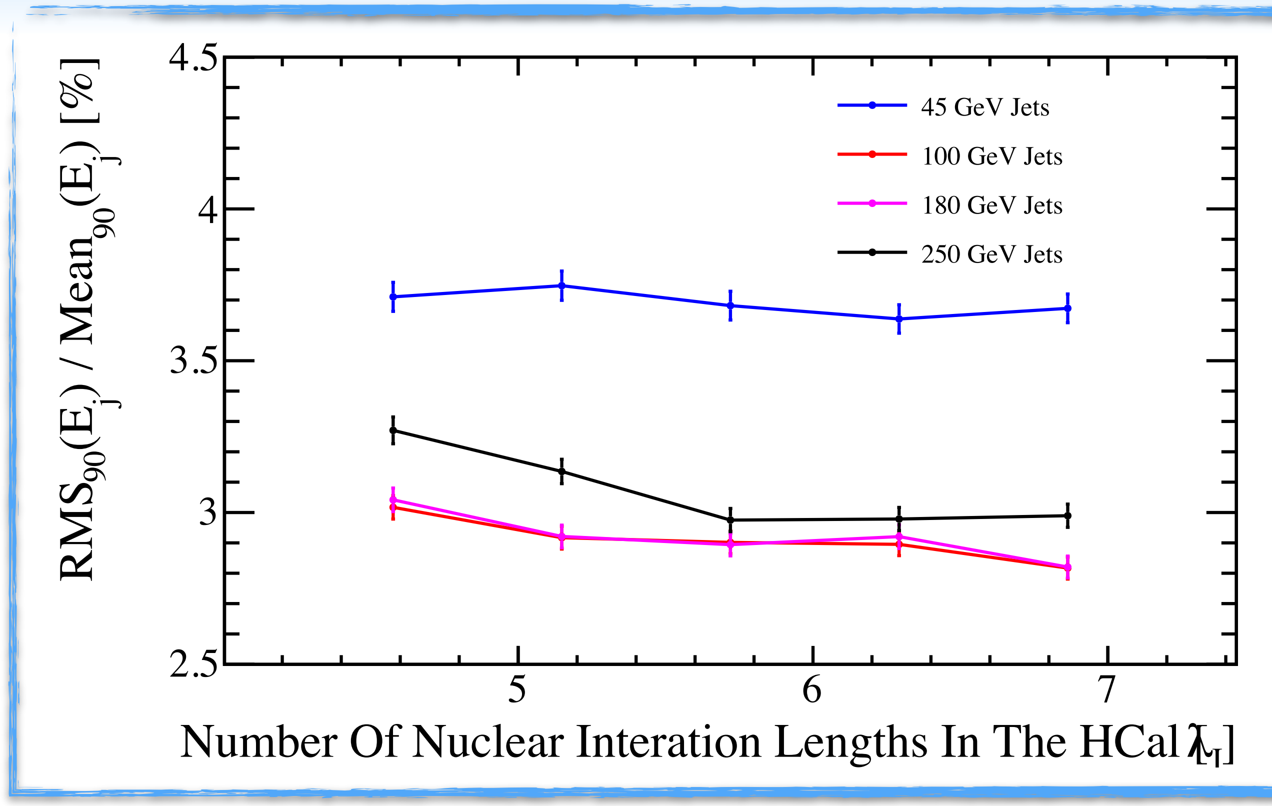


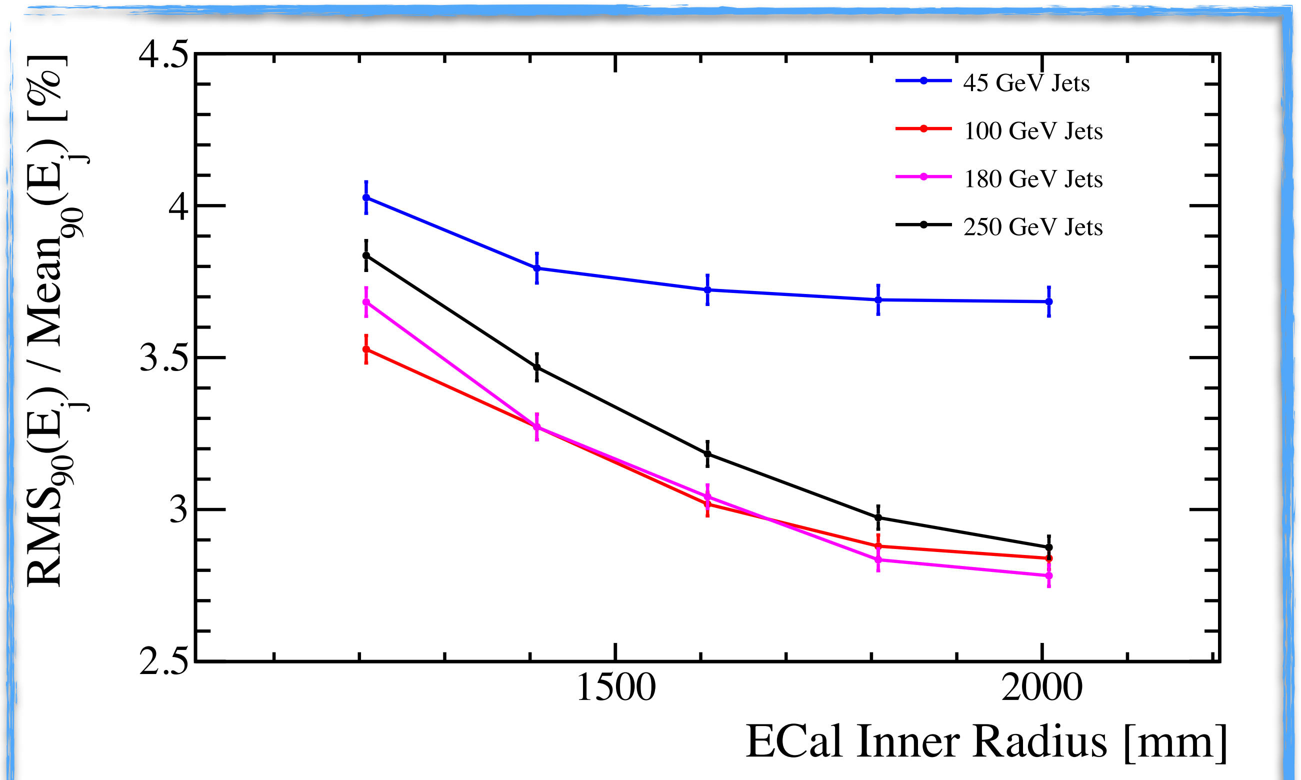


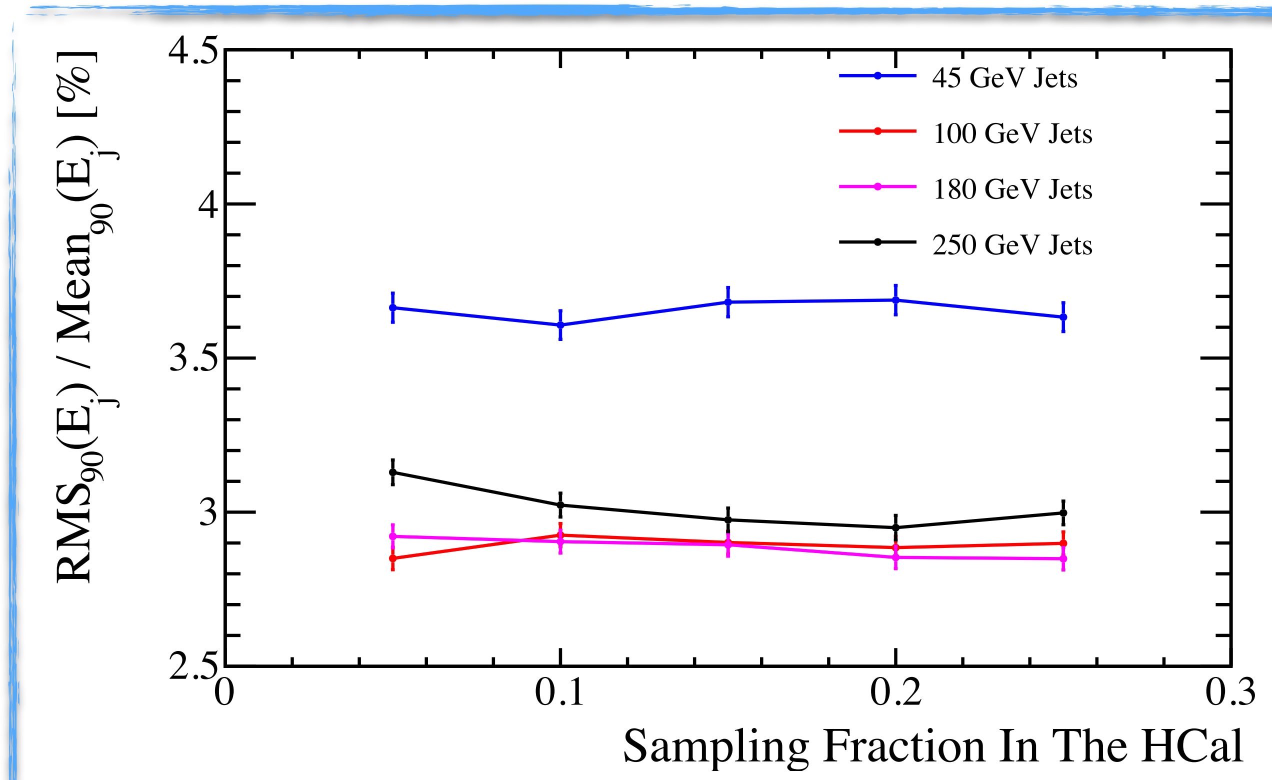


HCal Timing Cuts : 100 ns
ECal Timing Cuts : 100 ns
HCal Hadronic Cell Truncation: Optimised for each detector model (1 GeV)
Software : ilcsoft_v01-17-07, including PandoraPFA v02-00-00
Digitiser : ILDCaloDigi, realistic ECal and HCal digitisation options enabled
Calibration : PandoraAnalysis toolkit v01-00-00





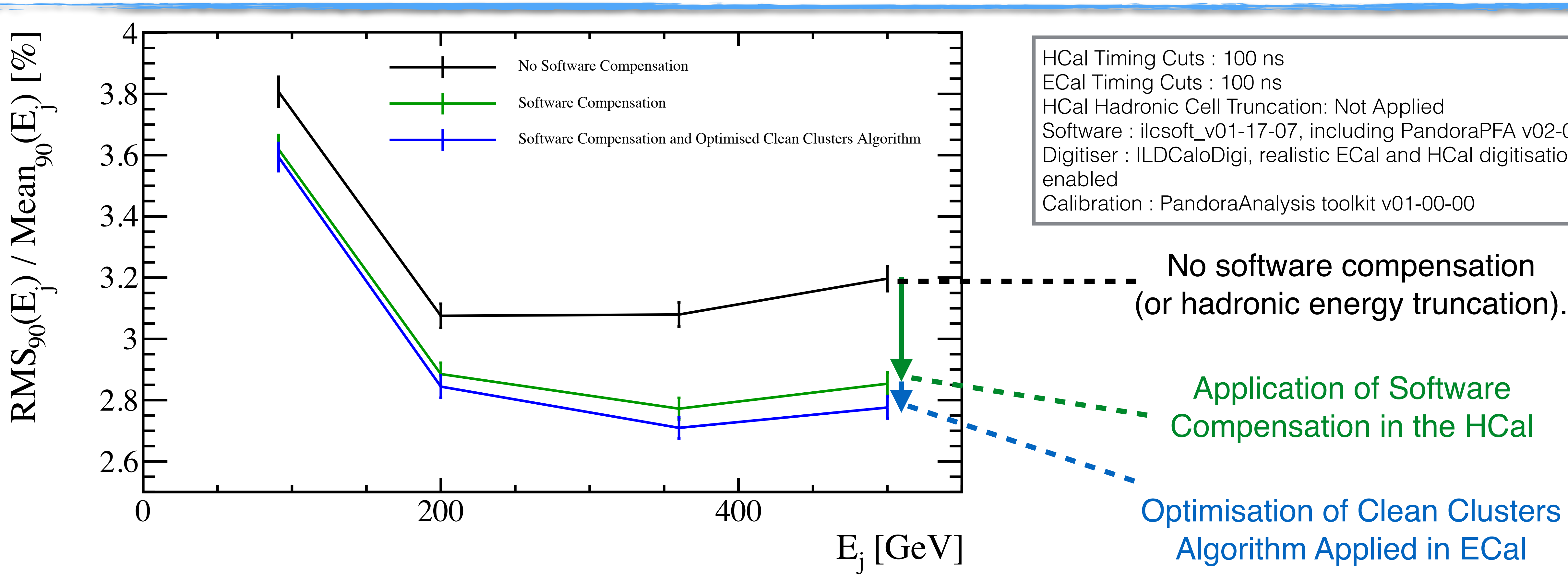




Backup - Software Compensation, Previous Talk Slides

Software Compensation

- The application of software compensation improves performance significantly.
- For the ECal the CleanClusters and the ScaleHotHadrons algorithms are applied. These are designed to improve the hadronic energy estimators by taking account of anomalously high energy calorimeter hits.
- The parameters of the CleanClusters algorithms were optimised alongside the development of software compensation.
- This optimisation also yielded significant improvements to the jet energy resolution.

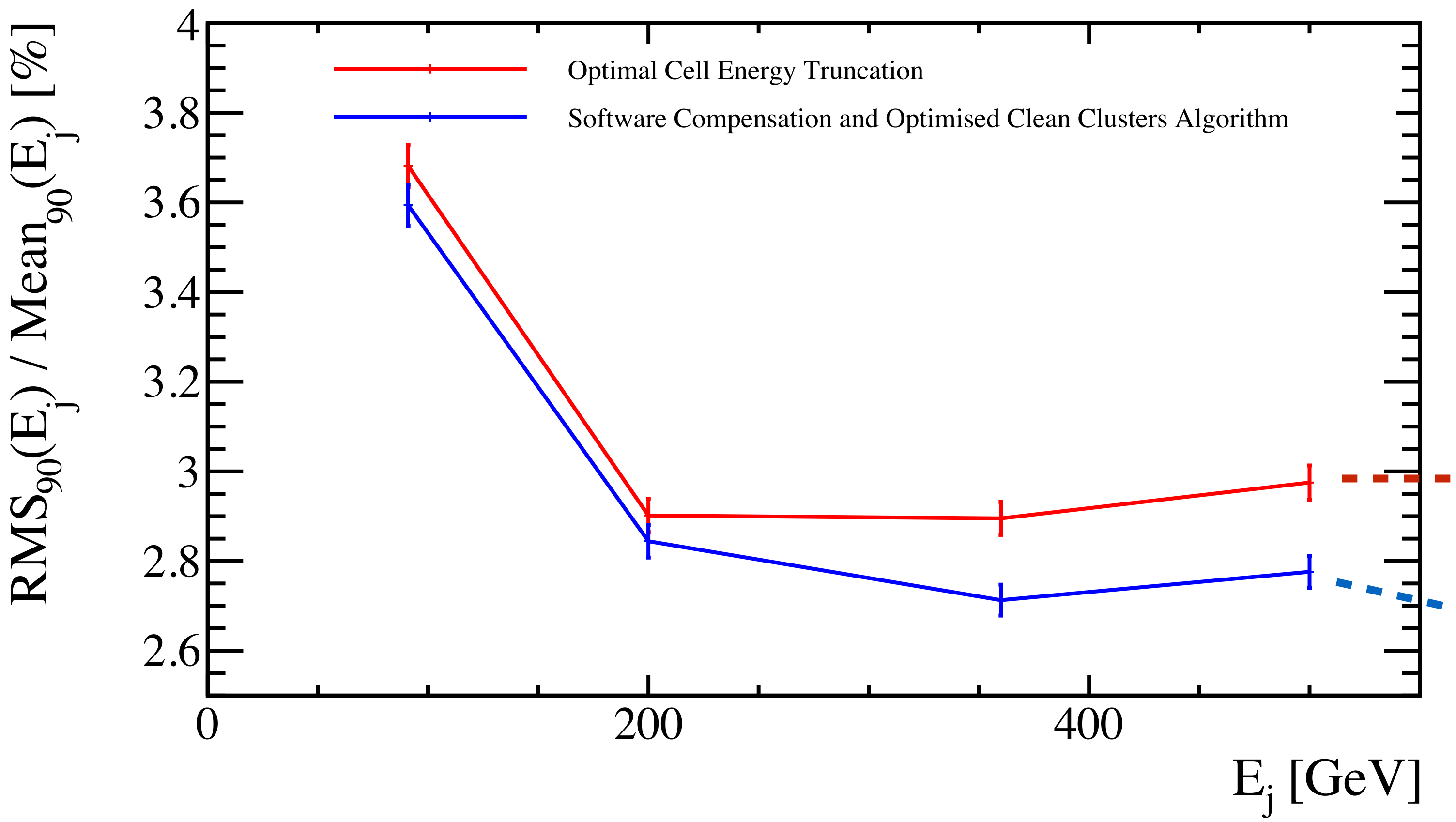


HCal Timing Cuts : 100 ns
ECal Timing Cuts : 100 ns
HCal Hadronic Cell Truncation: Not Applied
Software : ilcsoft_v01-17-07, including PandoraPFA v02-00-00
Digitiser : ILDCaloDigi, realistic ECal and HCal digitisation options enabled
Calibration : PandoraAnalysis toolkit v01-00-00

Software Compensation

- There is a clear improvement when we compare the jet energy resolution using software compensation and the optimised CleanClusters parameters in comparison to the previous best performance.
- The previous best performance applied a hadronic energy truncation for HCal hits and the unoptimised parameters in the CleanClusters algorithm. The hadronic energy truncation for HCal hits is a naive form of software compensation.

For full details please see 'Software Compensation' talk by Lan Tran Huong.



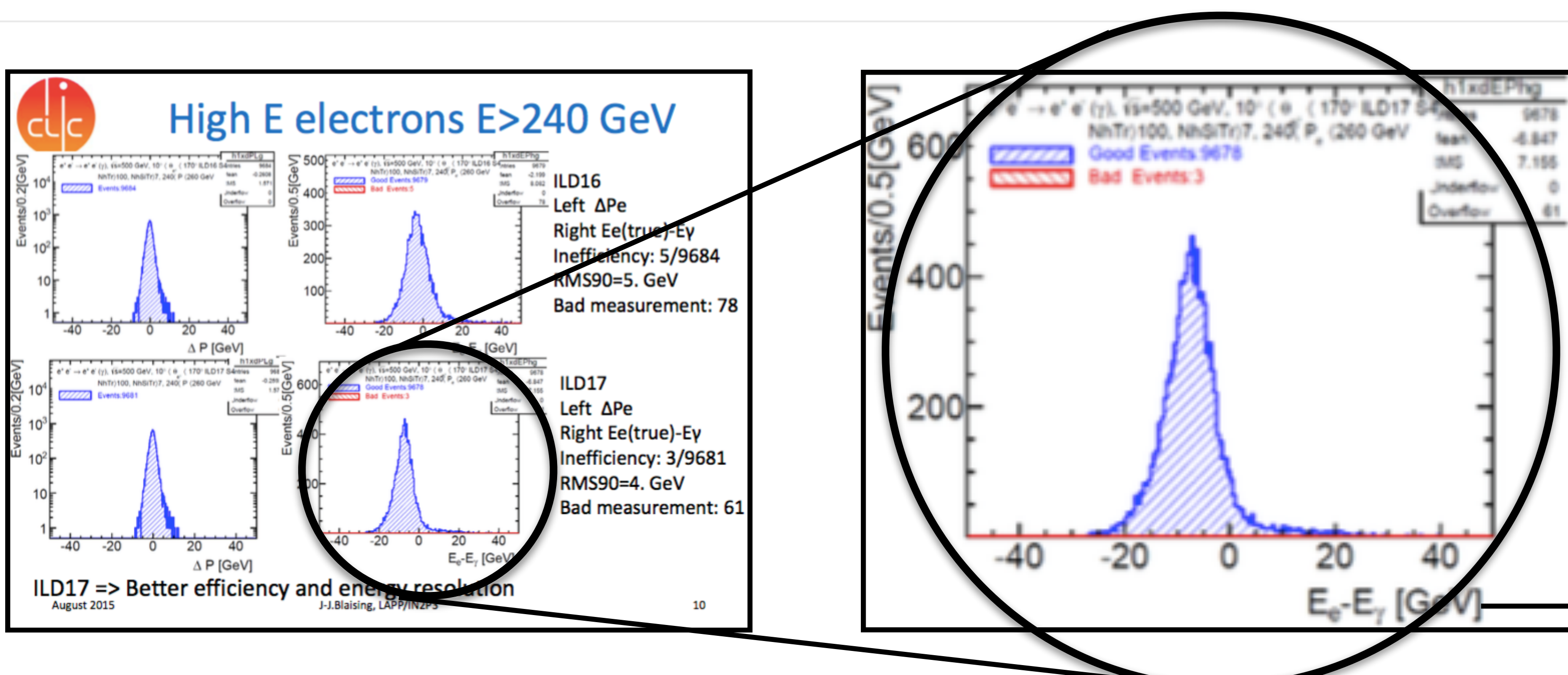
HCal Timing Cuts : 100 ns
ECal Timing Cuts : 100 ns
HCal Hadronic Cell Truncation: 1 GeV (Optimal for Default HCal) for optimal cell energy truncation graph and Not Applied for software compensation and optimised clean clusters graph
Software : ilcsoft_v01-17-07, including PandoraPFA v02-00-00
Digitiser : ILDCaloDigi, realistic ECal and HCal digitisation options enabled
Calibration : PandoraAnalysis toolkit v01-00-00

Previous Best Performance (with optimal hadronic energy truncation)

Optimisation of Clean Clusters and Software Compensation. (New Best Performance)

Backup - High Energy Photons

Thanks to J-J Blaising



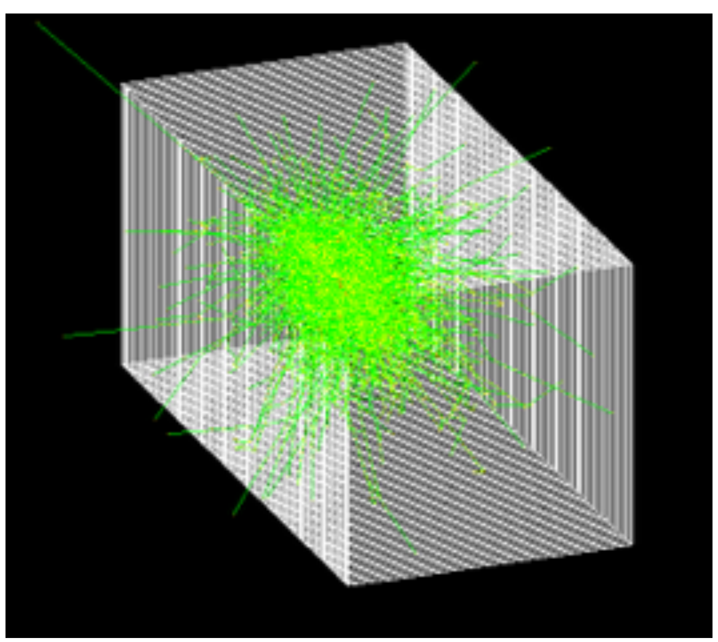
* The apparent miscalibration of the ECal for high energy (240GeV and above) electron/photon clusters.

For full talk see: https://indico.cern.ch/event/401048/contributions/1842291/attachments/1139940/1632536/mtg-1508yy-jjb-v3.pdf

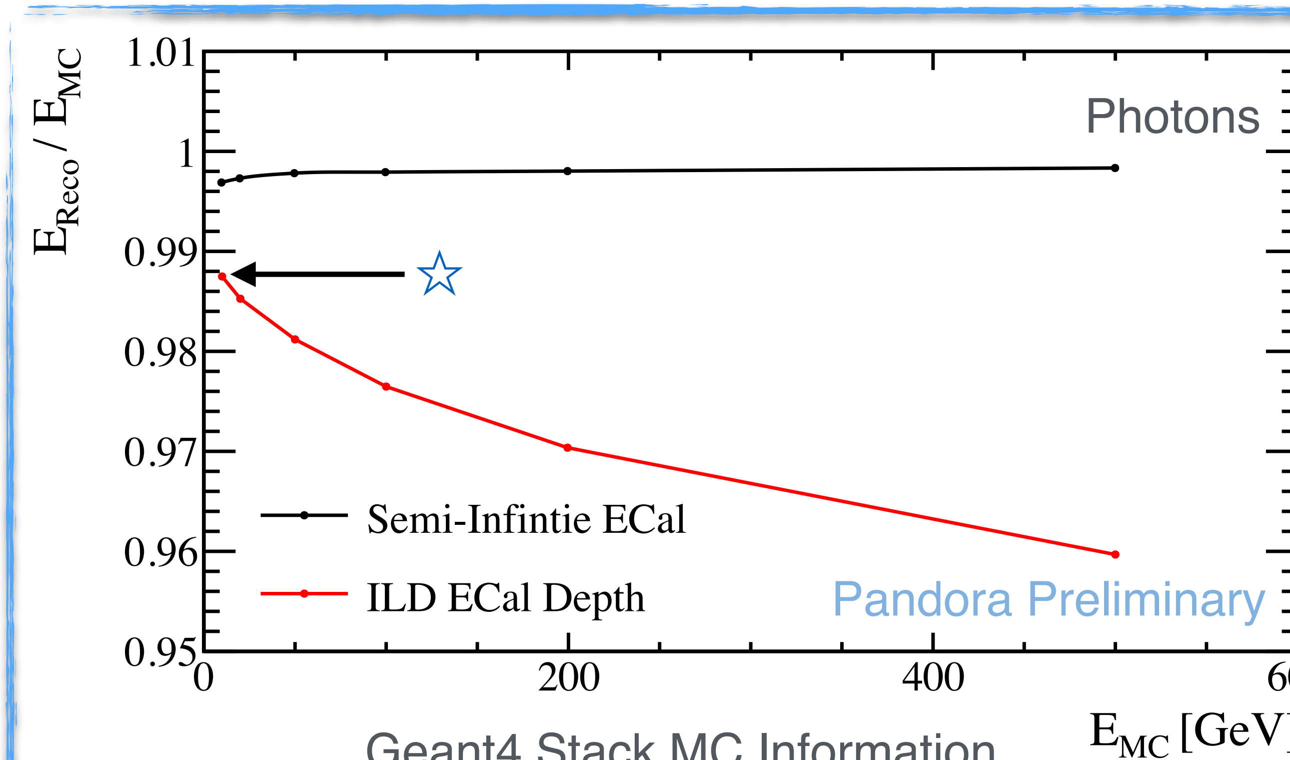
Track Energy - Cluster Energy
→Cluster Energy too large

Calibration

- Current calibration procedure assumes that at 10 GeV, a photon will be fully contained within the ILD ECal (24 X_0).
- To test this hypothesis a Geant4 calorimeter stack was built using the same materials at those specified in the Mokka (Si) ECal simulation.
- Two calorimeter stacks were simulated. The first had an identical longitudinal profile to that found in the ILD ECal (24 X_0) and the second one was semi-infinite (120 X_0).
- Several photons of different energies were fired through these stacks and their total (active+absorber) deposited energy recorded.
- It was found that approximately 98.8% (☆) of the energy is contained within an ILD ECal like stack for 10 GeV photons (in companions to almost 100% for the semi-infinite stack).
- If this isn't accounted for the calibration procedure for the full ILD reconstruction (Mokka + Marlin) it will scale the overall ECal energy by ~1.012, which could account for part of the high cluster energies being observed.



500 GeV Photon showering in the Geant4 calorimeter stack. Note: the full simulate used a much wider calorimeter to reduce transverse leakage.



Calibration

- A possible reason for high energy electromagnetic (EM) clusters are having their energy overestimated is the fact that the EM energy scale is not fine tuned for the HCal in the current calibration scheme.
- This is a pragmatic choice as even at the highest energies photons are largely contained within the ECal as the 2D calorimeter hit plot below shows.
- By setting the EM scale in the ECal precisely (as is currently done) you are able to achieve sub-percent level accuracy in photon/electron reconstructed energies.
- However, as you go to high energy photons/electrons the EM energy deposits in the HCal become more significant and the calibration of the EM scale in the HCal will be called into question.
- Currently the calibration procedure in Pandora set EM scale in the HCal to be the same as the hadronic scale in the HCal working under the (largely correct assumption) that there is a negligible amount of EM energy ever reading the HCal.