

# Calorimetry at a Future Linear Collider

Steven Green  
of Emmanuel College

A dissertation submitted to the University of Cambridge  
for the degree of Doctor of Philosophy



# Abstract

This thesis describes the optimisation of the calorimeter design for collider experiments at the future Compact Linear Collider (CLIC) and the International Linear Collider (ILC). The detector design of these experiments is built around high-granularity Particle Flow Calorimetry that, in contrast to traditional calorimetry, uses the energy measurements for charged particles from the tracking detectors. This can only be realised if calorimetric energy deposits from charged particles can be separated from those of neutral particles. This is made possible with fine granularity calorimeters and sophisticated pattern recognition software, which is provided by the PandoraPFA algorithm. This thesis presents results on Particle Flow calorimetry performance for a number of detector configurations. To obtain these results a new calibration procedure was developed and applied to the detector simulation and reconstruction to ensure optimal performance was achieved for each detector configuration considered.

This thesis also describes the development of a software compensation technique that vastly improves the intrinsic energy resolution of a Particle Flow Calorimetry detector. This technique is implemented within the PandoraPFA framework and demonstrates the gains that can be made by fully exploiting the information provided by the fine granularity calorimeters envisaged at a future linear collider.

A study of the sensitivity of the CLIC experiment to anomalous gauge couplings that effect vector boson scattering processes is presented. These anomalous couplings provide insight into possible beyond standard model physics. This study, which utilises the excellent jet energy resolution from Particle Flow Calorimetry, was performed at centre-of-mass energies of 1.4 TeV and 3 TeV with integrated luminosities of  $1.5\text{ab}^{-1}$

and  $2\text{ab}^{-1}$  respectively. The precision achievable at CLIC is shown to be approximately one to two orders of magnitude better than that currently offered by the LHC.

Finally, a study into various technology options for the CLIC vertex detector is described.

## Declaration

This dissertation is the result of my own work, except where explicit reference is made to the work of others, and has not been submitted for another qualification to this or any other university. This dissertation does not exceed the word limit for the respective Degree Committee.

Steven Green



## Acknowledgements

Of the many people who deserve thanks, some are particularly prominent, such as my supervisor. . .





# Contents

<b>1</b>	<b>Energy Estimators</b>	<b>1</b>
1.1	Motivation . . . . .	1
1.2	Calibration in the Particle Flow Paradigm . . . . .	4
1.2.1	Overview of the Calibration Procedure . . . . .	5
1.2.2	MIP Scale Determination in the Digitiser . . . . .	6
1.2.3	Digitisation Implementation . . . . .	6
1.2.3.1	ECal Digitisation Implementation . . . . .	6
1.2.3.2	HCal Digitisation Implementation . . . . .	8
1.2.3.3	HCal Ring Digitisation Implementation . . . . .	10
1.2.4	MIP Scale Determination in PandoraPFA . . . . .	11
1.2.5	Electromagnetic Scale in PandoraPFA . . . . .	12
1.2.6	Hadronic Scale in PandoraPFA . . . . .	14
1.2.7	Summary . . . . .	17
1.3	Novel Energy Estimators . . . . .	18
1.3.1	HCal Hit Energy Truncation . . . . .	19
1.3.1.1	Legacy Energy Corrections . . . . .	20
1.3.1.2	Impact on Single Particle Energy Resolution . . . . .	20
1.3.1.3	Impact on Jet Energy Resolution . . . . .	21
1.3.2	Software Compensation . . . . .	22
1.3.2.1	Impact on Single Particle Energy Resolution . . . . .	26
1.3.2.2	Impact on Jet Energy Resolution . . . . .	27
1.3.3	Summary . . . . .	30
1.4	Timing Cuts . . . . .	30
1.4.1	Impact on Single Particle Energy Resolution . . . . .	31
1.4.2	Impact on Jet Energy Resolution . . . . .	32
1.4.3	Summary . . . . .	33
	<b>Bibliography</b>	<b>37</b>



*“Writing in English is the most ingenious torture  
ever devised for sins committed in previous lives.”*

— James Joyce



# Chapter 1

## Energy Estimators

*“There, sir! that is the perfection of vessels!”*

— Jules Verne, 1828–1905

### 1.1 Motivation

This section outlines a procedure for calibrating the Monte-Carlo (MC) response of the linear collider detector simulations with a focus on converting the detector response into accurate energy measurements, "energy estimators", for particles showering in the calorimeters. In the particle flow paradigm, all neutral particle energies are measured using the calorimeters, which makes accurate energy estimators crucial for determining detector performance. Additionally, comparisons of particle shower energy estimators and charged particle track momenta govern the event reconstruction in PandoraPFA during the reclustering stage, which further emphasises the importance of reliable energy estimators.

The goal of a calorimeter is to measure the energy of particles that shower within it. Particle showers are a cascade of secondary particles that are produced as a high energy particle interacts with a dense material. The energy deposits produced by a showering particle in the calorimeter are known as hits. The number of hits created by a particle shower in a calorimeter depends upon the size and shape of the particle shower and the segmentation of the calorimeter. The energy of the showering particle,  $E_{Cluster}$ , is determined by grouping these energy deposits together into clusters and summing their

energy

$$E_{Cluster} = \sum_{ECal \text{ hits}, i} E_{ECal}^i + \sum_{HCal \text{ hits}, i} E_{HCal}^i, \quad (1.1)$$

where  $E_{ECal}^i$  is the energy of ECal hit  $i$  and  $E_{HCal}^i$  is the energy HCal hit  $i$ . In this example, the energy deposits made by the showering particle are assumed to be split across an ECal and a HCal, therefore, the sum runs over the hits in both calorimeters. This naive energy estimator will act as a starting point for the development of more sophisticated procedures aimed at improving detector performance.

Sampling calorimeters will be used at the linear collider experiments. These calorimeters are comprised of alternating layers of active and absorber materials [1]. The absorber layers initiate particle showers and propagate their growth, while the active layers produce a signal that is proportional to the energy deposited within them. The signal produced in the active layers is measured by sampling calorimeters and used to estimate the energy deposited in the absorber layers. This estimation is made by assuming the energy deposited across a calorimeter hit, that is one active and one absorber layer, is uniform. Working under this assumption, the total calorimeter hit energy is proportional to the active layer hit energy. This estimation procedure is loosely referred to as digitisation and, in this way, the cluster energy estimator introduced above can be written as

$$E_{Cluster} = \sum_{ECal \text{ hits}, i} \epsilon_{ECal}^i \alpha_{ECal} + \sum_{HCal \text{ hits}, i} \epsilon_{HCal}^i \alpha_{HCal}, \quad (1.2)$$

where  $\alpha_{ECal}$  and  $\alpha_{HCal}$  are digitisation constants for the ECal and HCal respectively,  $\epsilon_{ECal}^i$  is the ECal active layer hit energy for hit  $i$  and  $\epsilon_{HCal}^i$  is the HCal active layer hit energy for hit  $i$ . The first stage of the calibration procedure presented in this chapter covers the determination of these digitisation constants.

Once the basic energy estimator has been calibrated, it is possible to apply more advanced procedures designed to give a compensating calorimeter response [2]. A compensating calorimeter produces an identical response to a particle shower irrespective of whether the particle shower is electromagnetic or hadronic in nature. The primary cause of the difference in the response of a calorimeter to electromagnetic and hadronic showers is the undetectable energy component that is found in hadronic showers. These undetectable energy components are energy deposits produced from a showering particle that do not produce a signal in the calorimeters. Hadronic showers contain this undetectable component due to a combination of effects such as neutrons stopping within the calorimeter

and nuclear binding energy losses. Typically, this leads to calorimeters having a weaker response to hadronic showers than to electromagnetic showers.

There are two distinct routes available for achieving a compensating response from a calorimeter: The first is hardware compensation [3], whereby calorimeters are constructed using materials that yield extra energy in response to hadronic showers, and the second is software compensation [4], whereby the uncompensated calorimetric energies for hadronic showers are modified at the software level.

A novel example of hardware compensation is the ZEUS calorimeter [3]. The ZEUS calorimeter was constructed using uranium as the absorber material. In response to neutral hadrons the uranium undergoes fission producing extra energy that increases the hadronic response of the calorimeter. The amount of uranium was carefully chosen to achieve a fully compensating calorimeter response, i.e. identical calorimeter response to electromagnetic and hadronic showers. While hardware compensation is possible for the linear collider calorimeters, restrictions on calorimeter construction and the use of a large amount of radioactive material are highly undesirable.

The linear collider lends itself to software compensation as the fine segmentation of the calorimeters and precise reconstruction of individual particles makes identification of hadronic showers, and modifying their energies, feasible. A basic form of software compensation included in the linear collider reconstruction is the modification of the electromagnetic cluster energy estimator to

$$E_{EM\ Cluster} = \sum_{ECal\ hits, i} E_{ECal}^i \beta_{ECal}^{EM} + \sum_{HCal\ hits, i} E_{HCal}^i \beta_{HCal}^{EM} , \quad (1.3)$$

and the hadronic cluster energy to

$$E_{Had\ Cluster} = \sum_{ECal\ hits, i} E_{ECal}^i \beta_{ECal}^{Had} + \sum_{HCal\ hits, i} E_{HCal}^i \beta_{HCal}^{Had} , \quad (1.4)$$

where the  $\beta$ s are scaling factors that are applied to the energy of clusters of calorimeter hits associated with electromagnetic and hadronic clusters in the ECal and HCal. This simple scaling of energies compensates the response of the calorimeters, which leads to better detector performance. Determination of these energy scale setting constants is the second stage of the calibration procedure that is presented in this chapter.

While this scaling of energies improves detector performance, it does not account for any changes to the  $\beta$  scaling factors as a function of the total energy deposited. An

energy dependence in the scaling factors is expected as the mechanisms governing the propagation of hadronic showers are sensitive to the shower energy [5]. To account for this, more sophisticated software techniques have been developed that vary the calorimeter cluster energy estimator as a function of energy to achieve a compensating response across a wider range of energies. These techniques make use of the fine segmentation of the linear collider calorimeters to identify hadronic showers. These techniques also address the problem of spuriously high energy calorimeter hits, which are caused by Landau fluctuations [6]. Landau fluctuations originate from high energy knock-on electrons appearing within particle showers [7] and can lead to overestimates of the particle shower energy if they occur in the active layers of a sampling calorimeter.

## 1.2 Calibration in the Particle Flow Paradigm

Calibration of the linear collider detector simulation primarily related to two processors in the software framework; the digitiser, which performs the digitisation process for sampling calorimeters, and PandoraPFA. The input to the digitiser is the active layer calorimeter hit energies and the output is the combined, active and absorber layer, calorimeter hit energies. The combined hit energies are then used by PandoraPFA for event reconstruction. Calibration of the digitiser involves determining the digitisation constants,  $\alpha$ s, and the minimum ionising particle (MIP) scale, the average energy response on a per hit level, using the active layer calorimeter hit energies. Similarly, calibration of PandoraPFA requires setting the scaling factors,  $\beta$ , and the MIP response using the combined calorimeter hit energies.

The  $\alpha$  and  $\beta$  constants are determined by tuning the mean of reconstructed energy distributions. A number of cuts are applied when populating these reconstructed energy distributions that ensure the relevant reconstructed energy is being tuned. The application of these cuts means that linear scaling of the  $\alpha$  and  $\beta$  constants does not lead to a linear shift in the mean of the reconstructed energy distributions. Therefore, when calibrating the  $\alpha$  and  $\beta$  constants an iterative approach is taken; the next iteration of the calibration constant is determined by repeating the reconstruction using the current iteration of the constant and adjusting the constant based on the mean of the reconstructed energy distribution.

Determining the MIP scale is included in the calibration procedure as it is used by PandoraPFA in the identification of muons and for applying energy thresholds designed



to limit the impact of noise. This scale is also used in the digitiser for simulating electrical noise, saturation effects in scintillator readout technologies and for applying noise vetoing energy thresholds [8]. The MIP scale characterised in the digitiser and PandoraPFA by the, non-zero, peak in the distribution of the active layer and combined calorimeter hit energies for normally incident 10 GeV  $\mu^-$  respectively [7]. In the linear collider detector simulation, several realistic effects are simulated by the digitiser including saturation effects, energy thresholds, timing cuts and electrical noise. Application of these effects at this point in the software chain means that the active layer hit energies are not subject to them, while the post digitisation combined calorimeter hit energies are. Consequently, the MIP scale in PandoraPFA cannot be obtained from the digitiser MIP scale, instead both have to be independently determined.

Although this overall procedure is referred to as calibration, it is actually setting the Monte-Carlo (MC) response. In a real detector, calibration would follow the setting of the MC response and would involve calibrating the real data against this clearly defined MC response.

### 1.2.1 Overview of the Calibration Procedure

The calibration procedure is split into four separate operations: determination of digitisation constants ( $\alpha$ s) in the digitiser; determination of scaling factor constants ( $\beta$ s) in PandoraPFA; MIP scale setting in the digitiser; and MIP scale setting in PandoraPFA. Calibration of the digitiser, digitisation constants and MIP scale, uses calorimetric energy measurements prior to any reconstruction, while calibration of PandoraPFA, scale factors and MIP scale, uses fully reconstructed particle flow objects (PFOs). As reconstructed PFOs are created using calorimetric energy measurements that have been digitised, it is wise to calibrate the digitiser before PandoraPFA, therefore, the calibration procedure is applied in the following order:

1. Setting the MIP response in the digitiser.
2. Setting the digitisation constants,  $\alpha$ s, in the digitiser.
3. Setting the MIP response in PandoraPFA.
4. Setting the scaling factors,  $\beta$ s, in PandoraPFA.

### 1.2.2 MIP Scale Determination in the Digitiser

The MIP scale in the digitiser was determined by simulating 10 GeV  $\mu^-$  events and creating a distribution of active layer calorimeter hit energies for each calorimeter in the detector simulation. When populating these distribution, a direction correction factor of  $\cos(\theta)$ , where  $\theta$  is the incident angle of the  $\mu^-$  to the calorimeter hit, was applied to account for the path length of the MIP through the active medium of the calorimeter. This correction effectively makes the muons intersect the calorimeters at normal incidence. No selection cuts were applied to the sample of muon events.

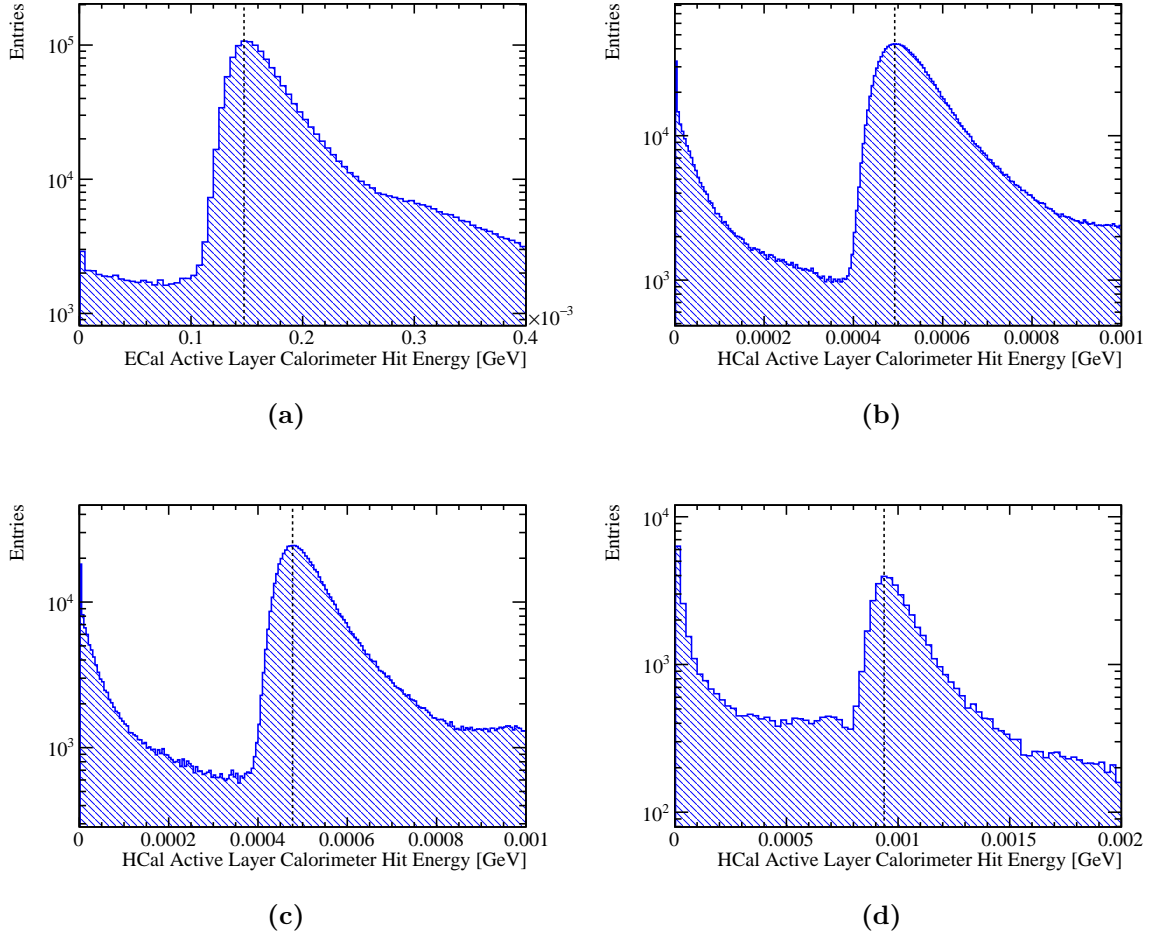
The MIP scale was determined separately for the ECal, HCal barrel, HCal endcap and HCal ring as can be seen in figure 1.1. In the digitiser software only a single HCal MIP scale, taken as the HCal barrel, was required. The HCal endcap and ring MIP scales were calculated for the purposes of the HCal ring digitisation.

### 1.2.3 Digitisation Implementation

This section discusses how the digitisation constants,  $\alpha_s$ , are determined. The digitisation constant for a given calorimeter depends upon several factors such as the material properties of the active and absorber layers, the magnetic field strength and energy losses occurring within the gaps in the detector. Therefore, each calorimeter in the ILD detector model has a distinct constant that must be independently determined.

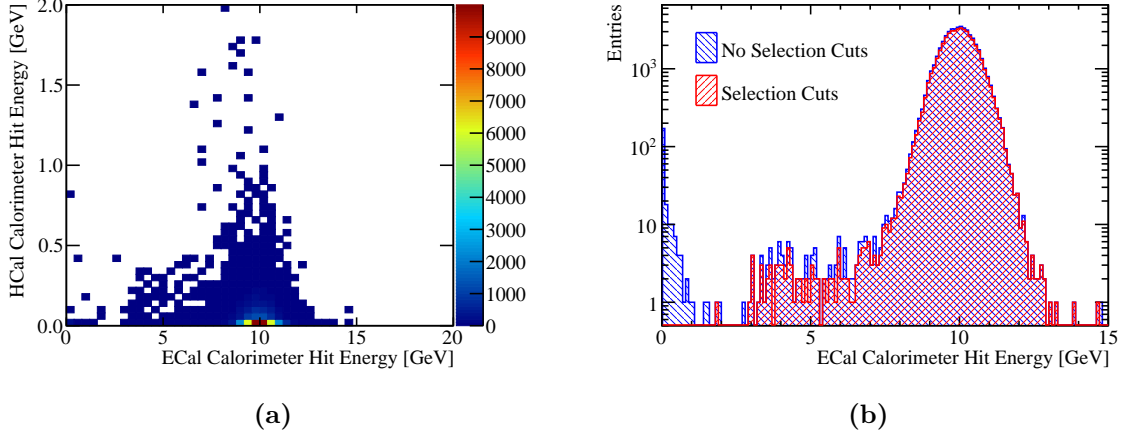
#### 1.2.3.1 ECal Digitisation Implementation

The procedure for determining the digitisation constants in the ECal involves simulation of single photons at an energy  $E_{MC} = 10$  GeV. Single photons at this energy are largely contained within the ECal, as shown in figure 1.2a. This makes them ideal for isolating the ECal digitisation calibration from that of the HCal digitisation calibration. Events are only used for calibrating the ECal digitisation if they are confined to the ECal. To that extent, cuts are applied ensuring that the sum of the reconstructed energy found outside the ECal is less than 1% of  $E_{MC}$  and that the  $\cos(\theta) < 0.95$ , where  $\theta$  is the polar angle of the photon. Photons that convert are also vetoed in this event sample at MC level. The impact of these cuts on the sum of ECal hit energies for the  $E_{MC} = 10$  GeV photons is shown in figure 1.2b.



**Figure 1.1:** The active layer calorimeter hit energy distributions for (a) the ECal, (b) the HCal barrel, (c) the HCal endcap and (d) the HCal ring for 10 GeV  $\mu^-$  events. The hit energies were corrected to account for the path length of the muons through the active medium of the calorimeter. The vertical black dotted lines indicate the position of the peak in each of these distributions that is used for defining the MIP scale in the digitisation processor.

The calibration of the digitisation in the ECal is an iterative procedure, which begins with the simulation of single photons using a trial calibration,  $\alpha_{\text{ECal}}^0$ . Next the distribution of the sum of calorimeter hit energies within the ECal is produced for events passing the selection cuts, as shown in figure 1.2b. For an ideal calorimeter this distribution should be Gaussian, as described in chapter ??, therefore, a Gaussian fit is applied to this distribution and the mean,  $E_{\text{Fit}}$ , extracted. To remove the effect of any outliers in this distribution, the fit is applied to the range of data with the smallest root mean square that contains at least 90 % of the data. An example of such a fit is shown in figure 1.3. In the case of ideal calibration, the mean of this fit,  $E_{\text{Fit}}$ , would be equal  $E_{MC}$ .



**Figure 1.2:** (a) The sum of calorimeter hit energies in ECal and HCal for 10 GeV photons. (b) The sum of the ECal calorimeter hit energies for 10 GeV photons with and without the selection cuts.

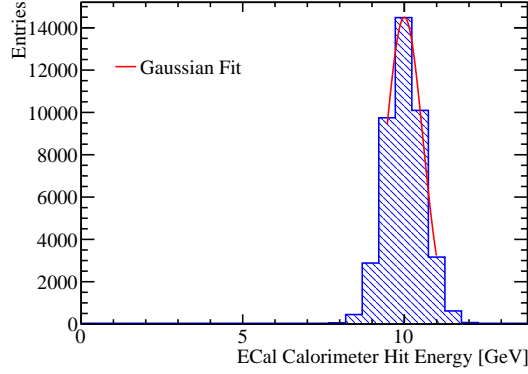
It is assumed that any difference between the two is due to the calibration, therefore, to correct this the digitisation constant from the trial calibration,  $\alpha_{\text{ECal}}^0$ , is rescaled by the ratio of the  $E_{MC}$  to  $E_{\text{Fit}}$

$$\alpha_{\text{ECal}}^0 \rightarrow \alpha_{\text{ECal}} = \alpha_{\text{ECal}}^0 \times \frac{E_{MC}}{E_{\text{Fit}}} . \quad (1.5)$$

This procedure is then repeated until the  $E_{\text{Fit}}$  falls within a specified tolerance of  $E_{MC}$ . The tolerance applied here was  $|E_{\text{Fit}} - E_{MC}| < E_{MC} \times 5\%$ . The binning used for the fitted histogram is chosen such that the bin width is equal to the desired tolerance on  $E_{\text{Fit}}$  e.g.  $E_{MC} \times 5\% = 0.5$  GeV. It should be emphasised that the PFO energies used for downstream analyses have the electromagnetic and hadronic energy scale corrections applied, which are calibrated to a much tighter accuracy.

### 1.2.3.2 HCal Digitisation Implementation

The calibration for the digitisation in the HCal proceeds in a similar manner to that described for the ECal with a few key differences. This calibration uses simulated MC long-lived neutral kaons ( $K_L^0$ s) at  $E_{MC} = 20$  GeV. The higher energy, with respect to the ECal digitisation, results in particle showers that sample deeper into the HCal. The  $K_L^0$ s must pass through the ECal, which contains one  $\lambda_I$ , before arriving at the HCal. Consequently, approximately 15% of these events begin showering in the ECal, as can be

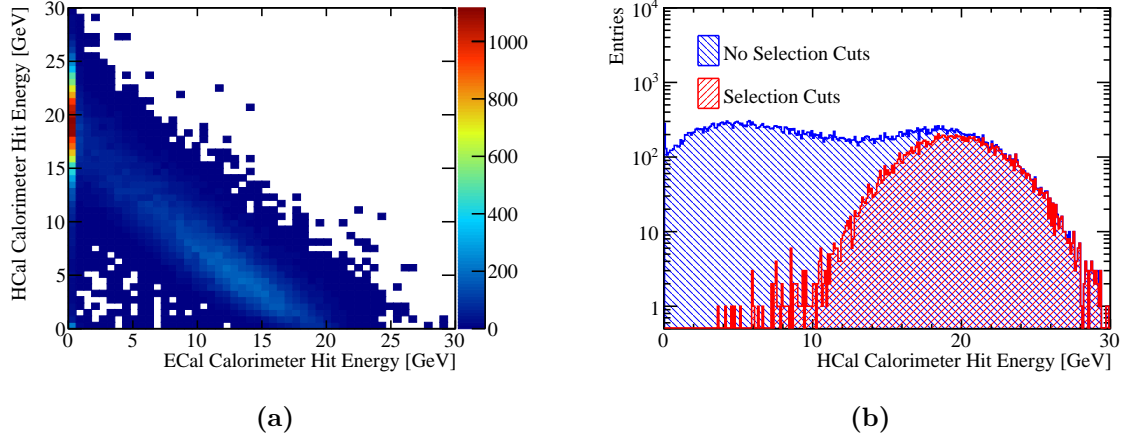


**Figure 1.3:** Gaussian fit to sum of the ECal calorimeter hit energies for 10 GeV photons with selection cuts. The coarse binning reflects the tolerance on the digitisation constant calibration.

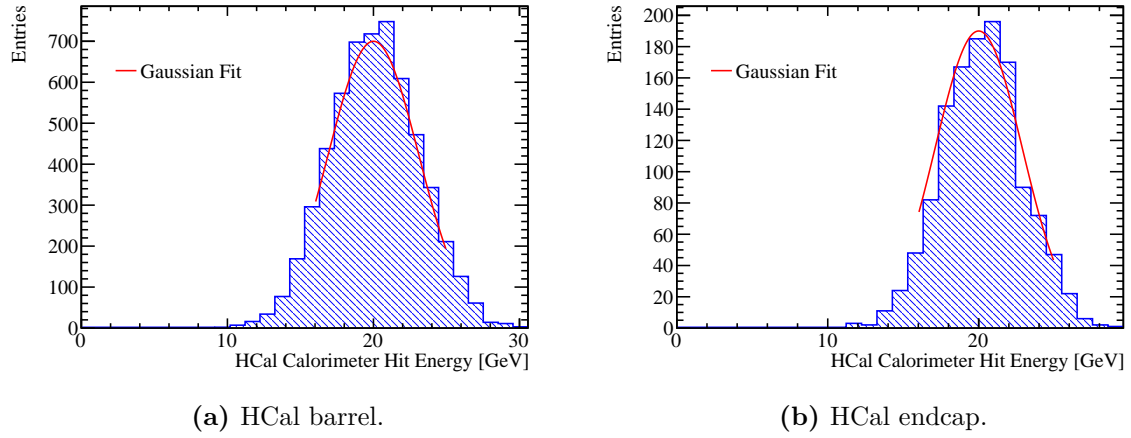
seen in figure 1.4a. Only events that deposit less than 5% of their energy in the ECal are used for calibrating the HCal digitisation constants. Furthermore, events that are not contained in the HCal are removed by requiring the last layer of the HCal where energy is deposited is required to be in the innermost 90% of the HCal. The impact of these cuts on the sum of HCal calorimeter hit energies for the  $E_{MC} = 20$  GeV  $K_L^0$  events is shown in figure 1.4b.

There are two HCal digitisation constants used in the detector simulation, one applied for the barrel and another for the endcap. The use of two digitisation constants accounts for differences in hadronic shower dynamics between the two, such as differing magnetic field configurations in the barrel and endcap. Both parameters are calibrated in the same manner, but have different cuts on  $\theta$ , the polar angle of the  $K_L^0$ . For the barrel region of the HCal events are selected if  $0.2 < \cos(\theta) < 0.6$ , while for the endcap events are selected if  $0.8 < \cos(\theta) < 0.9$ . These angular cuts account for the transverse profile of the hadronic showers and ensure that the showers are largely confined to the relevant sub-detector. As the majority of the neutral hadrons appearing in jets are neutrons and their accessible energy is the kinetic energy as opposed to the total energy, the target reconstructed energy for these  $K_L^0$  samples is the kinetic energy.

After applying the above  $K_L^0$  selection cuts, the calibration procedure for the digitisation of the HCal barrel and endcap proceeds in the same manner as was described for the ECal. An example of the Gaussian fits applied to the sum of the calorimeter hit energies in the HCal barrel and endcap are shown in figure 1.5.



**Figure 1.4:** (a) Sum of calorimeter hit energies in ECal and HCal for 20 GeV  $K_L^0$  events. (b) Sum of the HCal calorimeter hit energies for a 20 GeV  $K_L^0$  events with and without the selection cuts.



**Figure 1.5:** Gaussian fit to sum of the HCal calorimeter hit energies for 20 GeV  $K_L^0$  events with selection cuts.

### 1.2.3.3 HCal Ring Digitisation Implementation

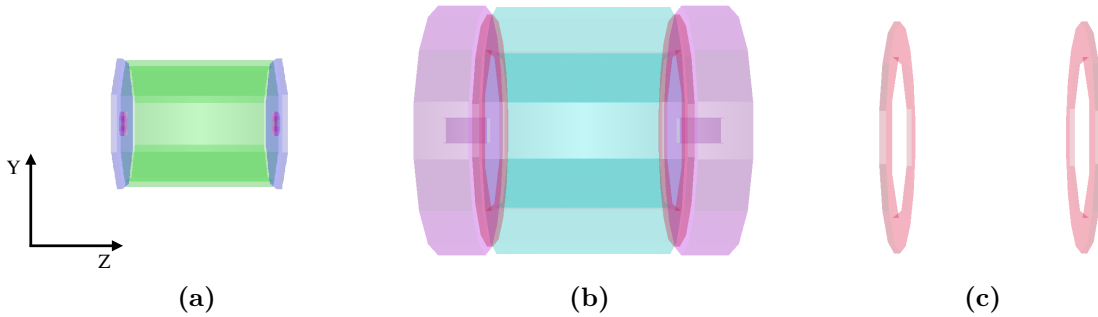
The HCal ring, as illustrated in figure 1.6, is a hadronic calorimeter that surrounds the ECal endcap and is sandwiched between the HCal barrel and endcap. This calorimeter is required to ensure hermetic coverage of the hadronic calorimeter system across the barrel/endcap cross-over region [9].

The HCal ring has an independent digitisation constant to account for any difference in the hadronic shower development between the ring, barrel and endcap. Due to

the thickness of the HCal ring, particle showers are never fully contained in it, so a different approach to calibration is required. To ensure that the HCal ring calibration is approximately correct,  $\alpha_{\text{HCal ring}}$  is assumed to equal  $\alpha_{\text{HCal endcap}}$  multiplied by several factors designed to account for differences in the active layer thickness, absorber layer thickness and the MIP response between the HCal endcap and ring. In detail

$$\alpha_{\text{HCal ring}} = \alpha_{\text{HCal endcap}} \times \frac{\langle \cos(\theta_{\text{endcap}}) \rangle}{\langle \cos(\theta_{\text{ring}}) \rangle} \times \frac{P_{\text{endcap}}}{P_{\text{ring}}} \times \frac{L_{\text{endcap}}^{\text{Absorber}}}{L_{\text{ring}}^{\text{Absorber}}} \times \frac{L_{\text{ring}}^{\text{Active}}}{L_{\text{endcap}}^{\text{Active}}}, \quad (1.6)$$

where  $\theta$  is the incident angle of the incoming particle to the calorimeter determined using the 20 GeV  $K_L^0$ s,  $L^{\text{Active}}$  is the active layer thickness and  $L^{\text{Absorber}}$  is the absorber layer thickness.  $P$  is the position of the MIP peak in the distribution of active layer hit energies, which has been corrected so that the MIP appears to enter the calorimeter at normal incidence, and is determined using 10 GeV  $\mu^-$  events. Details on how  $P$  is determined are given in section 1.2.2.



**Figure 1.6:** A PandoraPFA event display showing the nominal ILD calorimeters. (a) the ECal, (b) the full HCal and (c) the HCal ring, which covers the barrel/endcap cross-over region.

### 1.2.4 MIP Scale Determination in PandoraPFA

The MIP scale in PandoraPFA is set by simulating 10 GeV  $\mu^-$  events and creating the distribution of combined calorimeter hit energies. The MIP scale in PandoraPFA must to be determined for the calorimeters and, in contrast to the digitiser, the muon chamber. Consequently, an additional distribution showing the calorimeter hit energy for the muon chamber must be constructed at this stage of the calibration. As was done for the digitiser, a direction correction factor was applied to the hit energies account for the

path length of the MIP through the active medium of the calorimeter and no selection cuts were applied.

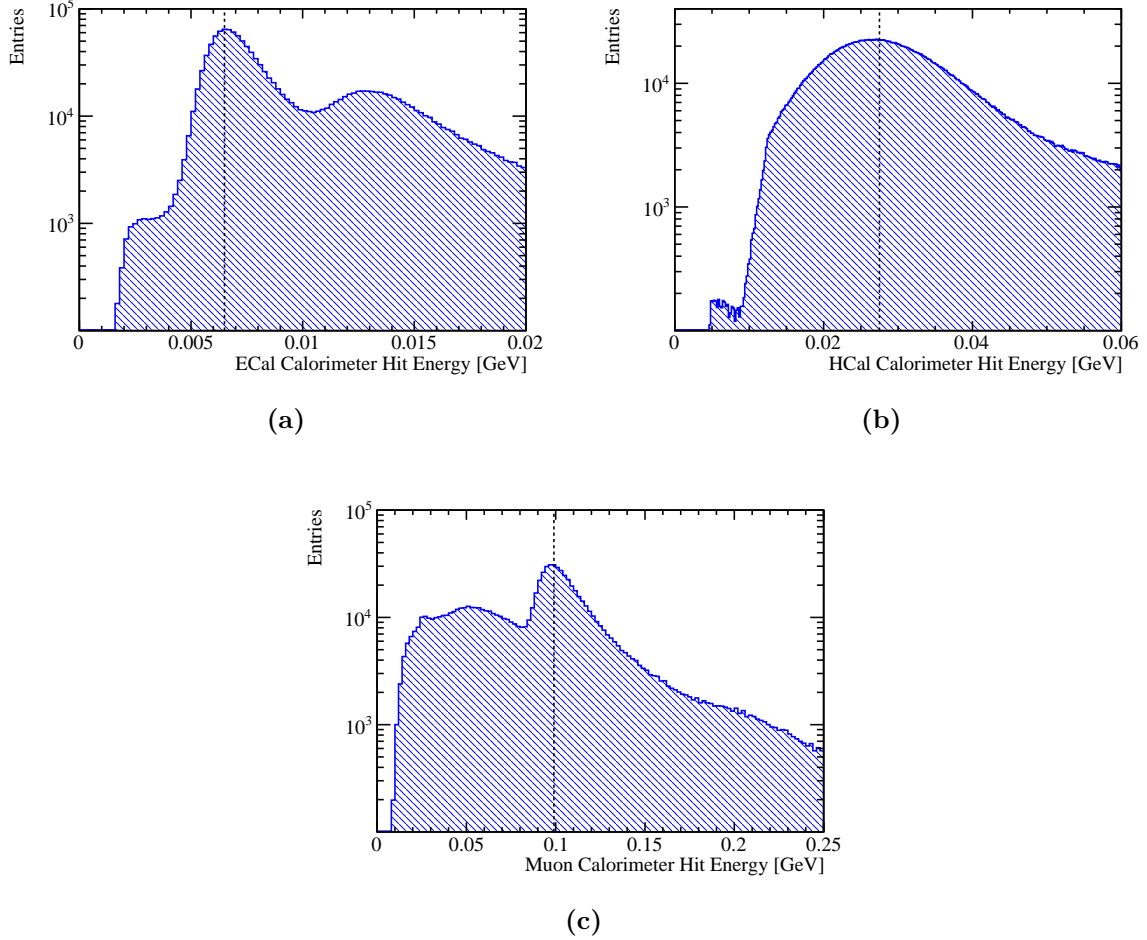
Examples of the distributions used to set the MIP scale in PandoraPFA can be found in figure 1.7. Due to the energy thresholds applied in the digitiser, there are few populated low calorimeter hit energy bins in these distributions. The double peak structure observed in the ECal calorimeter hit energy distribution is expected given the ECal absorber material thickness doubling in the back 10 layers of the ECal. The MIP peaks used for defining the MIP scale in PandoraPFA, figure 1.7, are broader than those used for determining MIP scale setting in the digitiser, 1.1, as the realistic effects applied by the digitiser are only present in the combined calorimeter hit energy distributions.

### 1.2.5 Electromagnetic Scale in PandoraPFA

Setting the electromagnetic scale in PandoraPFA is performed by examining the energies of fully reconstructed particles. The reconstruction was performed using the combined calorimeter hit energies that were set by the digitiser and having applied the noise vetoing MIP cuts.

The electromagnetic scale in the ECal,  $\beta_{ECal}^{EM}$ , is determined using simulated photons at  $E_{MC} = 10$  GeV. To ensure that the events used for this part of the calibration are largely confined to the ECal, a cut requiring less than 1% of the reconstructed energy to be found outside the ECal is applied. Furthermore, only events reconstructed as a single photon are used veto conversions. The impact of the selection cuts on the electromagnetic energy measured in the ECal for 10 GeV photons is shown in figure 1.8a. The peak at zero electromagnetic energy in the ECal is due to events traveling down the beam pipe and conversions. In photon conversion events the calorimetric energy deposits made by the  $e^\pm$  will be associated to charged particle tracks and therefore, in the particle flow paradigm, the reconstructed PFO energy will be measured using the charged particle tracks. In this case the energy measured using the calorimeters will be reported as zero, as although there are calorimetric energy deposits, they are not used when determining the PFO energies. The tail of events with low electromagnetic energy in the ECal occurs primarily due to pattern recognition failures in photon conversion events. In these events a small fraction of the calorimetric energy deposits made by the  $e^\pm$  are not associated to charged particle tracks and instead are reconstructed as separate photons with a reconstructed energy much less than  $E_{MC}$ .





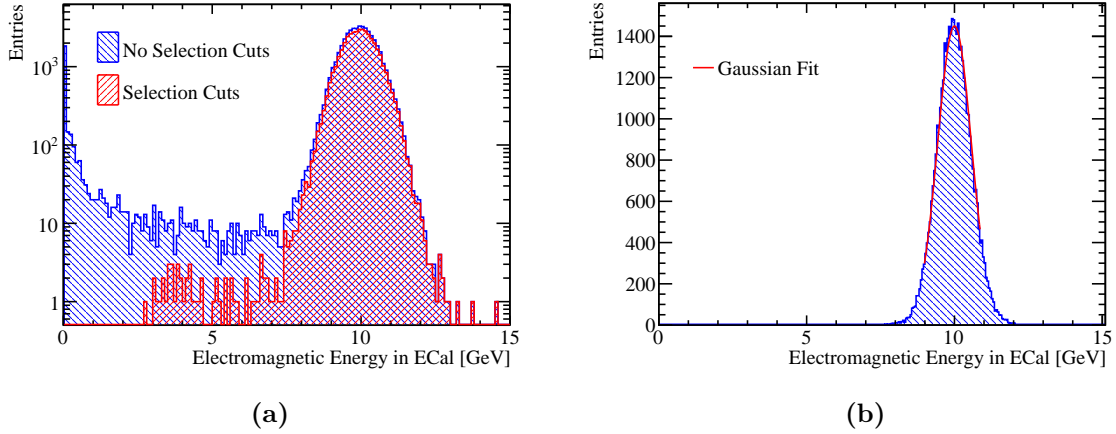
**Figure 1.7:** The combined calorimeter hit energy distributions for (a) the ECal, (b) the HCal and (c) the muon chamber for 10 GeV  $\mu^-$  events. These hit energies were corrected to account for the path length of the muons through the active medium of the calorimeter. The vertical black dotted lines indicate the position of the peak in each of these distributions that is used for defining the MIP scale in PandoraPFA.

The fitting procedure follows that used for the ECal digitisation, described in section 1.2.3.1, whereby a trial calibration for the electromagnetic energy scale in the ECal,  $\beta_{ECal}^{EM0}$ , is first assumed. The initial trial calibration is approximate and will be iteratively updated until it converges to within a chosen tolerance. Using the trial calibration the photon events are reconstructed and the distribution of the electromagnetic energy in the ECal created. A Gaussian fit is then applied to this distribution in the range with the smallest root mean square containing at least 90 % of the data. The mean of the

fitted Gaussian,  $E_{\text{Fit}}$ , is then used to scale  $\beta_{\text{ECal}}^{EM0}$  in the following way

$$\beta_{\text{ECal}}^{EM0} \rightarrow \beta_{\text{ECal}}^{EM} = \beta_{\text{ECal}}^{EM0} \times \frac{E_{\text{MC}}}{E_{\text{Fit}}} . \quad (1.7)$$

An example distribution and fit used in the calibration of the nominal ILD detector model can be found in figure 1.8b. This procedure is repeated using the updated  $\beta_{\text{ECal}}^{EM}$  until  $E_{\text{Fit}}$  falls within a specified tolerance. The tolerance applied here was  $|E_{\text{Fit}} - E_{\text{MC}}| < E_{\text{MC}} \times 0.5\%$ . The binning for the fitted histogram is chosen such that the bin width is equal to the desired target tolerance on  $E_{\text{Fit}}$  e.g.  $E_{\text{MC}} \times 0.5\% = 0.05$  GeV. This tolerance is tighter than was applied for the digitisation as it is these and only these energies that are used in downstream analyses.

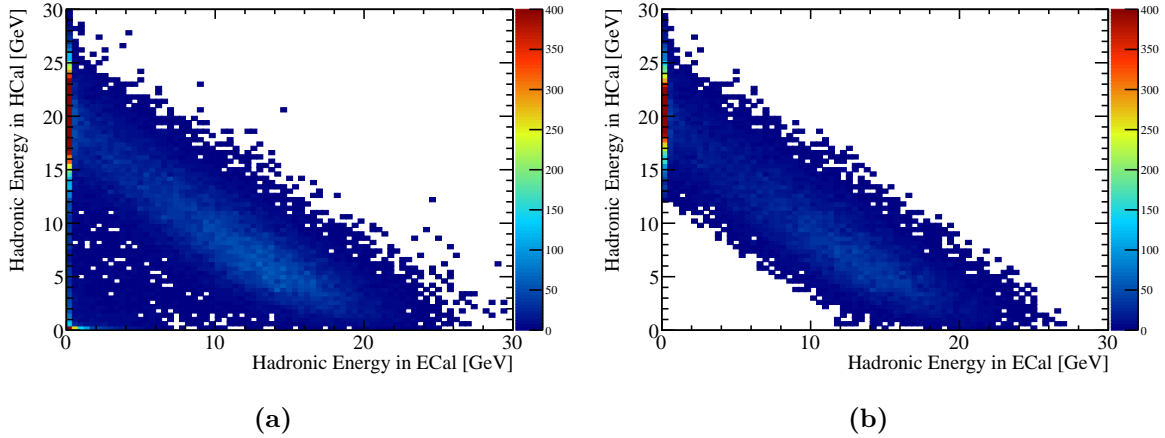


**Figure 1.8:** (a) The sum of the electromagnetic energy measured in the ECal for simulated 10 GeV photons with and without the selection cuts. (b) Gaussian fit to sum of the electromagnetic energy deposited in the ECal for simulated 10 GeV photons with selection cuts.

### 1.2.6 Hadronic Scale in PandoraPFA

The hadronic energy scale factors for the ECal,  $\beta_{\text{ECal}}^{\text{Had}}$ , and HCal,  $\beta_{\text{HCal}}^{\text{Had}}$ , are determined using simulated  $K_L^0$  events at  $E_{\text{MC}} = 20$  GeV. As the ECal contains approximately one nuclear interaction length, a non-negligible amount of hadronic energy will be deposited in the ECal, which makes the hadronic scale in the ECal,  $\beta_{\text{ECal}}^{\text{Had}}$ , important for detector performance. The hadronic scale in the ECal and HCal are simultaneously set as it is unfeasible to create a large sample of 20 GeV  $K_L^0$ s that are fully contained within the ECal.

For the reasons outlined in section 1.2.3.2, the target reconstructed energy for the sample of  $K_L^0$ s used for setting the hadronic energy scale is the kinetic energy,  $E_K$ , as opposed to the total energy. To ensure the events used are not affected by leakage of energy out of the back of the HCal, a cut is applied that vetoes events where energy is deposited in the outermost 10% of the HCal. In addition, a cut requiring a single neutral hadron to be reconstructed is applied to veto events with reconstruction failures. Finally, it is required that the total hadronic energy measured within the calorimeters falls within three  $\sigma$  of the kinetic energy of the  $K_L^0$ , where  $\sigma$  is taken to be  $55\% \times \sqrt{E_K}$  GeV. This definition for  $\sigma$  is approximately the energy resolution for neutral hadrons using the nominal ILD HCal [9]. This cut ensures that when fitting the two dimensional distribution of hadronic energy measured in the ECal and HCal, outliers do not skew the fit. The impact of these selection cuts can be seen in figure 1.9.



**Figure 1.9:** The distribution of hadronic energy measured in the ECal and HCal for 20 GeV  $K_L^0$  events (a) without selection cuts and (b) with selection cuts.

Determining the hadronic scale in PandoraPFA is an iterative process and begins by assuming trial values,  $\beta_{ECal}^{Had0}$  and  $\beta_{HCal}^{Had0}$ , for the hadronic scale calibration factors  $\beta_{ECal}^{Had}$  and  $\beta_{HCal}^{Had}$ . The  $K_L^0$  events are first reconstructed using the trial scale factors. Then a linear fit is applied to the two dimensional distribution of the reconstructed hadronic energies measured in the ECal and HCal for events passing the selection cuts. The best fit is obtained by minimising  $\chi^2$  with respect to variables describing a linear fit to the distribution. In this case,  $\chi^2$  is defined as

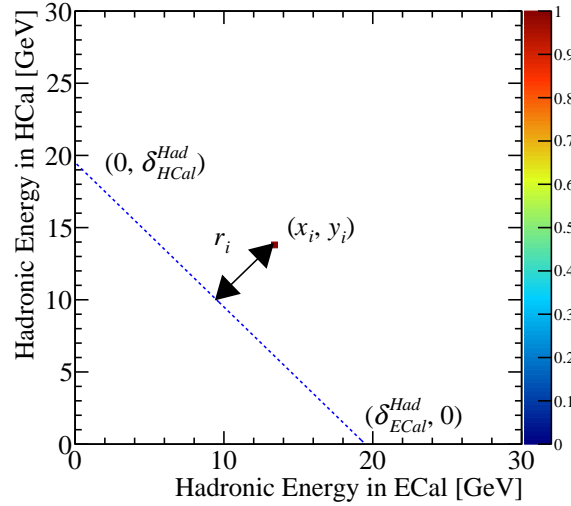
$$\chi^2(\delta_{ECal}^{Had}, \delta_{HCal}^{Had}) = \sum_i \left( \frac{r_i}{\sigma_{r_i}} \right)^2, \quad (1.8)$$

where  $r_i$  is the perpendicular distance in the two dimensional plane of hadronic energies measured in the ECal and HCal from the point  $(x_i, y_i)$  to a straight line passing through the points  $(\delta_{ECal}^{Had}, 0)$  and  $(0, \delta_{HCal}^{Had})$ . In this definition,  $x_i$  and  $y_i$  are the hadronic energies measured in the ECal and HCal respectively for event  $i$ . The variables  $\delta_{ECal}^{Had}$  and  $\delta_{HCal}^{Had}$  describe a linear fit to the hadronic energy distribution, which are to be varied when minimising  $\chi^2$ . The explicit definition of  $r_i$  is given in equation 1.9 and illustrated in figure 1.10. The uncertainty on  $r_i$  is given by  $\sigma_{r_i}$ , which is explicitly defined in equation 1.10. This uncertainty is calculated by propagating the uncertainties on  $x_i$  and  $y_i$ , which are assumed to be  $\sigma_{x_i/y_i} = 55\% \times \sqrt{x_i/y_i}$ , into the expression for  $r_i$ . The sum runs over all events,  $i$ , passing the selection cuts.

$$r_i = \frac{y_i \delta_{ECal}^{Had} + x_i \delta_{HCal}^{Had} - \delta_{ECal}^{Had} \delta_{HCal}^{Had}}{\sqrt{(\delta_{ECal}^{Had})^2 + (\delta_{HCal}^{Had})^2}}, \quad (1.9)$$

$$\sigma_i = \frac{(\sigma_{y_i} \delta_{ECal}^{Had})^2 + (\sigma_{x_i} \delta_{HCal}^{Had})^2}{\sqrt{(\delta_{ECal}^{Had})^2 + (\delta_{HCal}^{Had})^2}}. \quad (1.10)$$

The minimisation of  $\chi^2$  is done by stepping over a range of  $\delta_{ECal}^{Had}$  and  $\delta_{HCal}^{Had}$  centred about



**Figure 1.10:** An example showing the definition of  $r_i$ . The blue dotted line corresponds to  $y_i = \delta_{HCal}^{Had} - x_i \delta_{HCal}^{Had} / \delta_{ECal}^{Had}$ .

the ideal value of  $E_K$  in search for the minimum  $\chi^2$ . Once the minima in  $\chi^2$  is found the trial calibration factors  $\beta_{ECal}^{Had0}$  and  $\beta_{HCal}^{Had0}$  are rescaled to correct for any deviation from

the desired fit as follows

$$\beta_{ECal}^{Had0} \rightarrow \beta_{ECal}^{Had} = \beta_{ECal}^{Had0} \times \frac{E_K}{\Delta_{ECal}^{Had}} , \quad (1.11)$$

$$\beta_{HCal}^{Had0} \rightarrow \beta_{HCal}^{Had} = \beta_{HCal}^{Had0} \times \frac{E_K}{\Delta_{HCal}^{Had}} , \quad (1.12)$$

where  $\Delta_{ECal}^{Had}$  and  $\Delta_{HCal}^{Had}$  are the values of  $\delta_{ECal}^{Had}$  and  $\delta_{HCal}^{Had}$  giving the minimum  $\chi^2$ . The step size used for minimising  $\chi^2$  with respect to  $\delta_{ECal}^{Had}$  and  $\delta_{HCal}^{Had}$  was chosen such that a single step would correspond to the final tolerance on  $\delta^{Had}$ , which in this case is  $\approx 0.1$  GeV. This procedure is then repeated using the updated hadronic scaling factors until  $\Delta_{ECal}^{Had}$  and  $\Delta_{HCal}^{Had}$  both fall within a specified final tolerance, which in this case it taken to be  $|\Delta_{E/HCal}^{Had} - E_K| < E_K \times 0.5\% \approx 0.1$  GeV.

The electromagnetic scale in the HCal,  $\beta_{HCal}^{EM}$ , is chosen to be equal to the hadronic scale in the HCal,  $\beta_{HCal}^{Had}$ . For the ILC and CLIC,  $\beta_{HCal}^{EM}$  is not a critical parameter in the reconstruction as photons are largely contained within the ECal meaning little to no electromagnetic energy is measured in the HCal.

Setting the hadronic scale in PandoraPFA ensures that the energy estimators for neutral hadrons are accurate at 20 GeV, however, this is not true for all energies. The undetectable energy component of a hadronic shower varies as a function of particle shower energy [5], which means the response of a calorimeter to neutral hadrons non-linear with the hadron energy. This is an inherent limitation of this calibration procedure that will be addressed by the development of more sophisticated energy estimators in subsequent chapters.

### 1.2.7 Summary

The procedure for setting the MC response in the linear collider detector simulation has been outlined. This procedure ensures that when modifying the detector simulation, the response of the detector will yield reliable energy estimators for particles showering in the calorimeter. For completion, after this calibration procedure has been applied, retraining of the likelihood data used by specific algorithms in PandoraPFA for the reconstruction of photons can be performed. The calibration procedure presented here ensures reliability when examining the performance of the linear collider detector simulation.

### 1.3 Novel Energy Estimators

This section describes two novel energy estimators that are introduced with a view to improving the energy resolution for hadronic showers. Two techniques will be discussed: HCal hit energy truncation, which focuses on limiting the impact of Landau fluctuations; and software compensation, which focuses on obtaining a compensating calorimeter response. Both of these techniques are implemented by introducing weights,  $\omega^i$ , to calorimetric energy deposits made by showering particles in the HCal. The energy of a showering particle,  $E_{Cluster}$ , is determined by grouping together a clusters of calorimeter hits and summing their energies. When weights are applied to HCal hits this energy estimator becomes

$$E_{Cluster} = \sum_{ECal\ hits, i} E_{ECal}^i + \sum_{HCal\ hits, i} E_{HCal}^i \omega^i(\rho^i) . \quad (1.13)$$

Weights are only applied to calorimeter hits in the HCal as these techniques modify the energy of hadronic showers, which are primarily contained within the HCal. The weights,  $\omega^i$ , vary a function of the energy density of the calorimeter hit,  $\rho^i = E_{HCal}^i/V$  where  $V$  is the physical volume of a calorimeter hit in the HCal, which includes the both the active and absorber layer thicknesses.

Although the exact weights depend on the implementation of the technique, a general feature is that at large  $E_{HCal}^i$  the weight is less than one. This limits the impact of spuriously high energy hits caused by Landau fluctuations. The energy loss probability distribution function for scintillator detectors, such as the ILD HCal, is given by a Landau function, which is essentially a Gaussian function with a high energy tail [6]. Energy deposits from the high energy tail of this distribution, which are known as Landau fluctuations, account for high energy knock-on electrons that appear within particle showers [7]. As Landau fluctuations deposit a disproportionately large amount of energy with respect to the bulk of the particle shower, they can lead to overestimates of the particle shower energy.

The energy loss probability distribution function for  $n$  particles passing through a calorimeter hit is given by the convolution of  $n$  Landau functions, which by the central limits theorem will tend to a Gaussian as  $n$  becomes large. Consequently, as the average number of particles passing through a calorimeter hit increases, the high energy tail in the energy loss probability distribution function for the hit becomes less pronounced and the impact of Landau fluctuations decreases. This means that the impact of Landau

fluctuations on energy measurements is dictated by the density of a particles within a particle shower and the transverse segmentation, or cell size, of the calorimeter in use. If the transverse segmentation, or cell size, of a calorimeter decreases, the average number of particles passing through each hit decreases and the impact of Landau fluctuations increases. Any technique used for minimising the impact of Landau fluctuations will be sensitive to the transverse segmentation of the calorimeters in use.

### 1.3.1 HCal Hit Energy Truncation

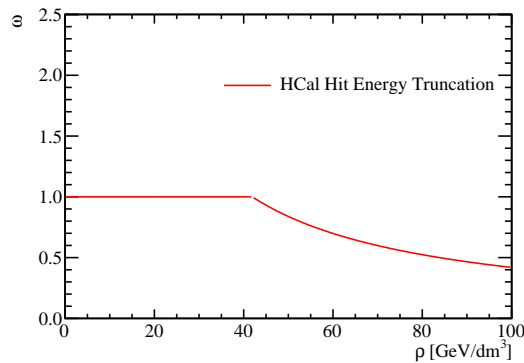
The first technique to be examined is a simple truncation of the hadronic energy,  $E$ , recorded in any given HCal hit

$$E \rightarrow E' = \begin{cases} E & \text{if } E < \kappa , \\ \kappa & \text{otherwise ,} \end{cases} \quad (1.14)$$

where  $\kappa$  is the value of the truncation. This improves the energy estimators for hadronic clusters by limiting the impact of Landau fluctuations. In terms of  $\omega$  introduced in equation 1.13 the truncation corresponds to

$$\omega(\rho) = \begin{cases} 1 & \text{if } \rho \times V < \kappa , \\ \frac{\kappa}{\rho \times V} & \text{otherwise .} \end{cases} \quad (1.15)$$

This weight as a function of hit energy density is shown in figure 1.11.



**Figure 1.11:** The weights,  $\omega$ , used in the HCal hit energy truncation as a function of the energy density of the HCal hit,  $\rho$ . The truncation shown here corresponds to a 1 GeV truncation in the nominal ILD HCal.

### 1.3.1.1 Legacy Energy Corrections

Alongside the HCal hit energy truncation, PandoraPFA also applied two other energy corrections designed at limiting the impact of Landau fluctuations. They are:

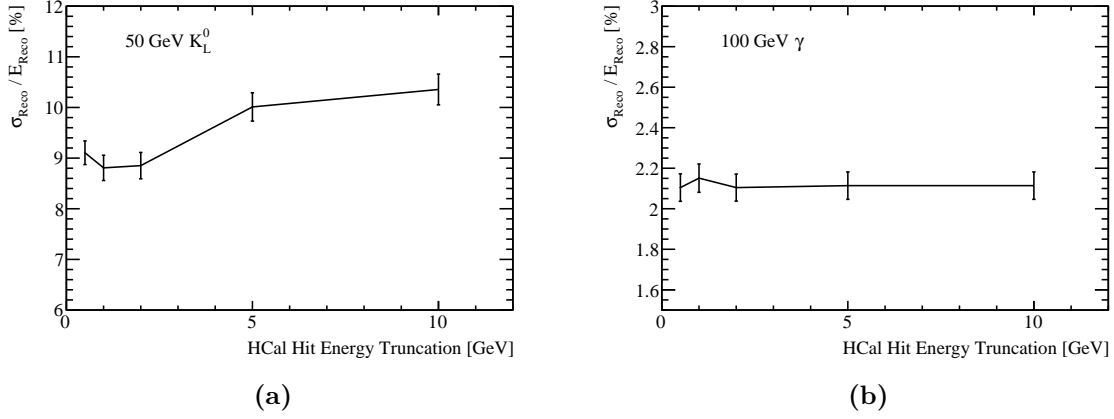
- **Clean Clusters.** This algorithm checks to see whether the energy measured within a calorimeter hit is anomalously high. Anomalously high energy hits are defined as hits where the energy contained within the hit is greater than 10% of the energy of the cluster that the hit has been associated to. If a hit is deemed to have an anomalously high energy and if this energy is above a threshold (0.5 GeV) the hit energy used by PandoraPFA is modified. The updated hit energy is taken as the average hit energy in the calorimeter layers immediately before and after the layer containing the high energy hit.
- **Scale Hot Hadrons.** This algorithm calculates the average number of MIP equivalent particles passing through each calorimeter hit in a cluster. If this number is larger than a given value, default 15 MIPs per hit, the cluster energy is rescaled to give a lower average number of MIPs per hit, default is 5 MIPs per hit.

These algorithms, with the HCal hit truncation, form the "legacy" energy corrections that are used by PandoraPFA when performing the event reconstruction.

### 1.3.1.2 Impact on Single Particle Energy Resolution

Figure 1.12a shows the energy resolution for neutral hadrons as a function of the HCal hit energy truncation. The optimal truncation for the ILD detector model simulation was 1 GeV and, using this truncation, a neutral hadron energy resolution of  $\sim 8.8\%$  was obtained. This energy resolution is approximately  $\sim 1.6\%$  lower than that obtained without the use of a truncation. Smaller energy truncations begin to truncate the energy of calorimeter hits produced in typical hadronic shower development, while larger truncations allow for a larger impact from Landau fluctuations. Both of these effects result in worsening neutral hadron energy resolutions. For completeness the  $\gamma$  energy resolutions as a function of HCal hit energy truncation are shown in figure 1.12b. As expected the  $\gamma$  energy resolution is unaffected by the HCal hit energy truncation as the  $\gamma$ s are largely contained within the ECal.



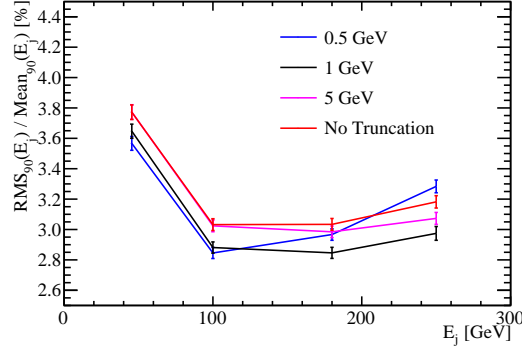


**Figure 1.12:** The energy resolution as a function of HCal cell truncation for (a) 50 GeV  $K_L^0$  events and (b) 100 GeV  $\gamma$  events using the nominal ILD detector model.

### 1.3.1.3 Impact on Jet Energy Resolution

Figure 1.13 shows the jet energy resolution as a function of jet energy for selected values of the HCal hit energy truncation. The trends in this plot are complex as the optimal HCal hit energy truncation varies with the jet energy. For 45.5 GeV jets, the best jet energy resolution,  $\sim 3.55\%$ , is obtained using a 0.5 GeV truncation, while for 180 GeV jets, the best jet energy resolution,  $\sim 2.85\%$ , is obtained using a 1 GeV truncation. This is expected as at low jet energies the average number of particles passing through each active calorimeter hit will be small, meaning the impact of Landau fluctuations is large and that to limit them a small truncation is needed. As the jet energy increases more particles on average pass through each calorimeter hit and the impact of Landau fluctuations decreases.

It is clear that a 1 GeV HCal hit energy truncation is beneficial for the performance of the nominal ILD detector model as the jet energy resolution is reduced by roughly  $\sim 0.15\%$  across the jet energy range from 45.5 GeV to 250 GeV. As the HCal hit truncation technique offers significant performance gains, it is used for the calorimeter optimisation studies presented in chapter ???. Included in those studies is an optimisation of the HCal cell size. Increasing the HCal cell size will increase the average number of particles passing through each calorimeter hit, which in turn reduces the impact of Landau fluctuations and vice versa. For all detector models considered where the HCal cell size was varied, the HCal hit energy truncation was reoptimised to account for the changing impact of Landau fluctuations. For detector models with a HCal cell size of



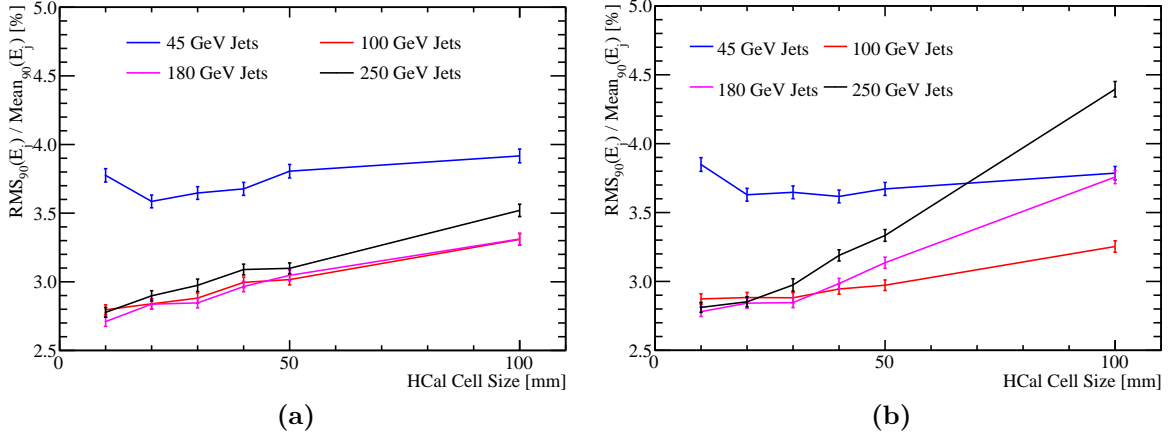
**Figure 1.13:** The jet energy resolution as a function of jet energy for various HCal hit energy truncations. The results shown use the nominal ILD detector model, which contains  $30 \times 30 \text{ mm}^2$  square scintillator tiles in the HCal.

10, 20, 30, 40, 50 and 100 mm the reoptimised truncations were 0.5, 0.75, 1, 1.5, 2 and 5 GeV respectively. Furthermore, the average particle density in a HCal hit will also be sensitive to the properties of the absorber material used in the calorimeters, therefore, the HCal hit energy truncation was also reoptimised in the HCal absorber material study. The optimal truncation for tungsten HCal options was 5 GeV, while for all other detector models considered it was 1 GeV.

Understanding the affect of the HCal hit energy truncation is crucial when performing optimisation studies. This can be seen in figure 1.14, which shows the results of the HCal cell size optimisation study when using a 1 GeV truncation and when optimising the truncation for each detector model. By applying a uniform HCal hit energy truncation the importance of the HCal cell size to particle flow calorimetry is vastly overinflated. For example, if the HCal cell size is increased from 10 mm to 100 mm the jet energy resolution for 250 GeV jets goes from  $\sim 2.8\%$  to  $\sim 4.5\%$  for the flat 1 GeV truncation, but only  $\sim 3.5\%$  when using an optimised truncation. As the jet energy and HCal cell size increase, the flat 1 GeV truncation throws away a larger fraction of typical hadronic shower energy measurements, which causes the jet energy resolution to degrade rapidly.

### 1.3.2 Software Compensation

Particle showers that are produced when a hadron interacts with a calorimeter contain two components [5]; an electromagnetic shower core, which originates from the production and decay of  $\pi^0$ s and  $\eta$ s, and a hadronic shower component originating from other interacting



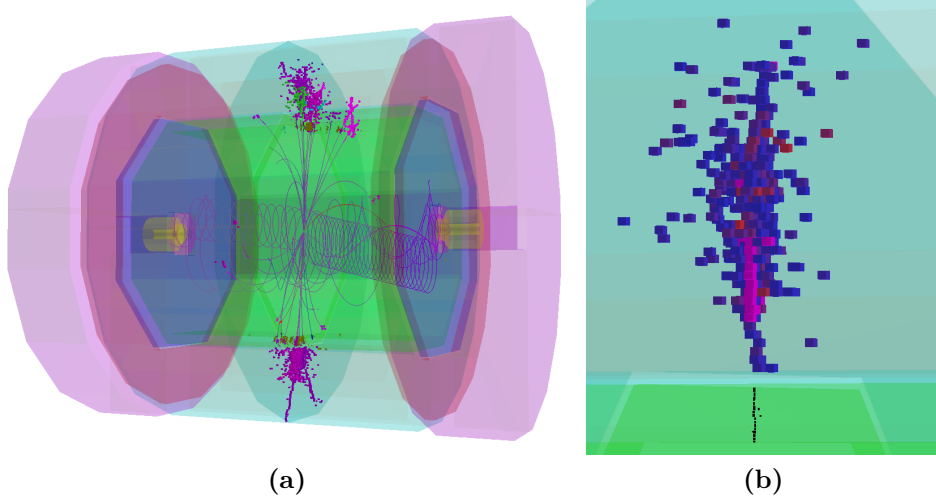
**Figure 1.14:** The jet energy resolution as a function of HCal cell size in the ILD detector model using a HCal hit energy truncation that is (a) optimised and (b) fixed at 1 GeV.

and decaying particles. By identifying each of these components in the reconstruction, it is possible to modify their energies to give a compensating calorimeter response. This technique known as software compensation.

Software compensation achieves a compensating calorimeter response by applying weights, as introduced in equation 1.13, that modify the energy of calorimeter hits in the HCal. These weights increase the energy found in the hadronic hits to compensate for the undetectable energy component found in hadronic showers. Additionally, these weights is to reduce the energy of spuriously high energy hits to minimise the impact of Landau fluctuations. The weights vary as a function of the calorimeter hit energy density,  $\rho^i$ , and the uncompensated energy of the particle shower,  $E_{Raw}$ , where

$$E_{Raw} = \sum_{ECal \text{ hits}, i} E_{ECal}^i + \sum_{HCal \text{ hits}, i} E_{HCal}^i. \quad (1.16)$$

The electromagnetic and hadronic components of a hadronic particle shower are treated differently in this approach by applying wights that are sensitive to the energy density of the calorimeter hits. High energy density hits are likely to be part of the electromagnetic core, while low energy density hits are likely to be part of satellite hadronic hits around the electromagnetic shower core [10]. By tailoring the weights as a function of the energy density, a compensating calorimeter response can be obtained. Figure 1.15 shows the electromagnetic and hadronic shower components, determined by the energy density of the calorimeter hits, for a hadronic shower in a 500 GeV  $Z \rightarrow uds$  di-jet event.



**Figure 1.15:** An event display for a 500 GeV  $Z \rightarrow uds$  di-jet event reconstructed using the nominal ILD detector. (a) The full event environment. (b) A single hadronic cluster from the same event where shading indicates the energy density in the HCal. High energy density cells are coloured red, while lower energy density cells are coloured blue. All ECal hits are shaded black. The high energy density electromagnetic core of the selected hadronic cluster is clearly visible.

The software compensation weights also depend on  $E_{Raw}$  to account for the sensitivity of the hit energy density distribution on the total particle shower energy. For hadronic showers, the fraction of the total energy carried in the electromagnetic core increases as the total shower energy increases [5], therefore a dependency of the weights on  $E_{Raw}$  is needed to obtain a compensating calorimeter response across a wide range of energies.

The precise form of the weights used for this technique are [10]

$$\omega(E_{Raw}, \rho) = p_1(E_{Raw}) \times \exp(p_2(E_{Raw}) \times \rho) + p_3(E_{Raw}) \quad (1.17)$$

$$p_1(E_{Raw}) = p_{11} + p_{12} \times E_{Raw} + p_{13} \times E_{Raw}^2 \quad (1.18)$$

$$p_2(E_{Raw}) = p_{21} + p_{22} \times E_{Raw} + p_{23} \times E_{Raw}^2 \quad (1.19)$$

$$p_3(E_{Raw}) = \frac{p_{31}}{p_{32} + \exp(p_{33} \times E_{Raw})}, \quad (1.20)$$

where  $p_{\alpha\beta}$  are constants and

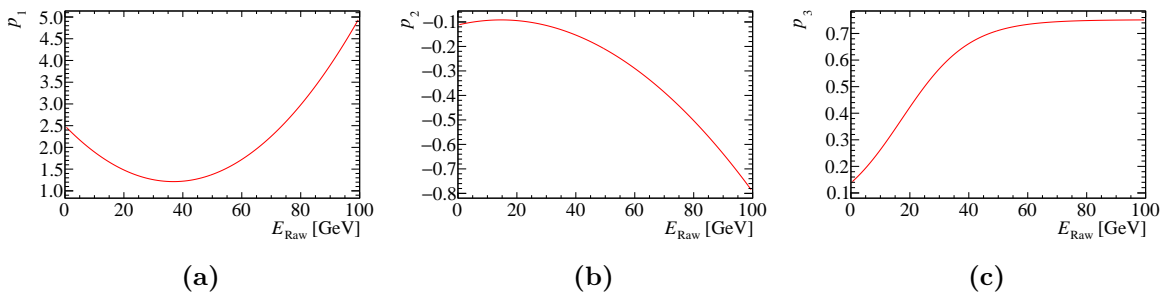
$$E_{Raw} = \sum_{ECal \text{ hits}, i} E_{ECal}^i + \sum_{HCal \text{ hits}, i} E_{HCal}^i. \quad (1.21)$$

The parameters  $p_{\alpha\beta}$  were determined by minimising  $\chi^2(p_{\alpha\beta})$  where

$$\chi^2(p_{\alpha\beta}) = \sum_{Events} \left( \frac{(E_{SC,Cluster}(p_{\alpha\beta}) - E_{MC})}{0.5 \times \sqrt{E_{MC}}} \right)^2 \quad (1.22)$$

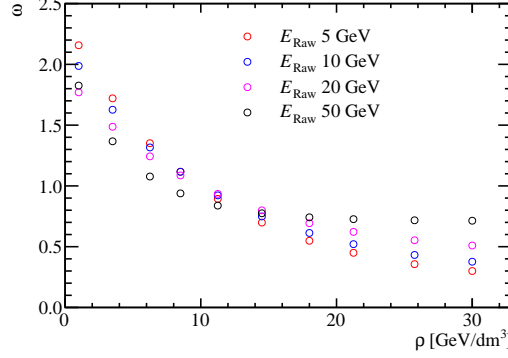
where the sum runs over single  $K_L^0$  events that ranged in energy from 10 to 100 GeV in steps of 10 GeV. At each energy the same number of events was used to avoid biasing to particular energies. In each event,  $E_{SC,Cluster}$  is the software compensated energy estimator for the reconstructed event and  $E_{MC}$  is the MC energy of the  $K_L^0$ . Normalising the deviation of  $E_{SC,Cluster}$  from  $E_{MC}$  by the stochastic term in the HCal energy resolution,  $\sim 50\% \times \sqrt{E}$ , made sure events of different MC energy contributed the same weight to  $\chi^2$ .

Using the fitted parameters obtained for the nominal ILD detector,  $p_1$ ,  $p_2$  and  $p_3$  as a function of  $E_{Raw}$  is shown in figure 1.16 and  $\omega(E_{Raw}, \rho)$  as a function of  $\rho$  for selected values of  $E_{Raw}$  is shown in figure 1.17. Figure 1.17 shows that the high energy density hits are being reduced in energy to compensate for the effects of Landau fluctuations, while the low energy density hits are being increased in weight to compensate for the undetectable energy component found in hadronic showers. Furthermore, the weights vary as a function of the raw hadronic shower energy to account for the changing energy density topology of hadronic showers with increasing shower energy.



**Figure 1.16:** Parameters used in software compensation weight determination as a function of  $E_{Raw}$ .

This technique is applied in the PandoraPFA framework in the form of an energy correction function, which means whenever the energy of a cluster of hits is considered by PandoraPFA the software compensated energy is used. Applying software compensation in this way benefits the detector energy resolution in two ways; firstly, the intrinsic energy resolution of the detector improves and secondly, the confusion contribution to the energy resolution is reduced.



**Figure 1.17:** The software compensation weight applied to a calorimeter hit as a function of calorimeter hit energy density for various cluster energies.

As software compensation only modifies the energy of HCal hits there is freedom to apply further energy corrections to the ECal hits. Applying the "Clean Clusters" logic, described in section 1.3.1.1, to the ECal hits alongside software compensation was found to be beneficial to the jet energy resolution. Therefore, the application of software compensation within PandoraPFA implicitly involves the application of the "Clean Clusters" logic to the ECal hits.

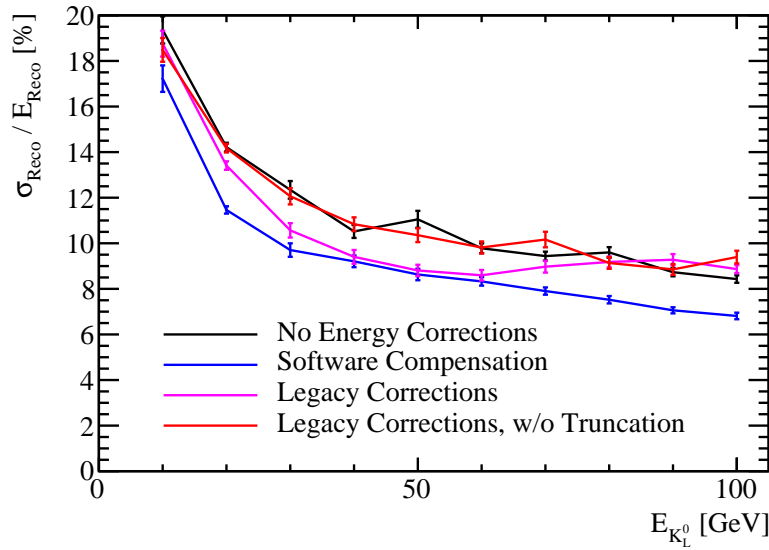
Software compensation was tuned using a maximum  $K_L^0$  energy of 100 GeV, therefore, it is only applied to clusters where  $E_{Raw} < 100$  GeV as sensible behaviour outside this range cannot be ensured. While it would be possible to modify the energy range of the training sample to go to higher energies, hadronic clusters with energy greater than 100 GeV will be rare at the ILC like energies considered here.

### 1.3.2.1 Impact on Single Particle Energy Resolution

Figure 1.18 shows the energy resolution as a function of MC energy for single  $K_L^0$  events obtained using various energy correction configurations in PandoraPFA. When comparing the energy resolution given by software compensation to that obtained using no energy corrections, it can be seen that software compensation offers an improvement in the energy resolution of  $\sim 15\%$  across the energy range considered. The uniformity of this improvement is encouraging, indicating that software compensation is achieving a compensating calorimeter response across this wide range of energies.

Comparing the performance of software compensation to the legacy corrections, described in section 1.3.1.1, it can be seen that software compensation gives a better

energy resolution across almost the entire range of energies considered. The only exception to this is around  $E_{K_L^0} \sim 50$  GeV where the performance of software compensation and the legacy corrections are comparable. By removing the hit truncation from the legacy options it is clear that the changes in energy resolution when using the legacy options are being driven by the hit truncation. This makes the trend in energy resolution observed using the legacy corrections clear as, at low  $K_L^0$  energies, very few hits are affected by the truncation so the performance is comparable to not using any energy corrections. At high  $K_L^0$  energies, the truncation is too aggressive and removes energy from hits that are not spuriously high leading to a worsening energy resolution. Between these two extremes,  $E_{K_L^0} \sim 50$  GeV, the truncation works ideally and the improvement in energy resolution when using the legacy corrections is the largest.

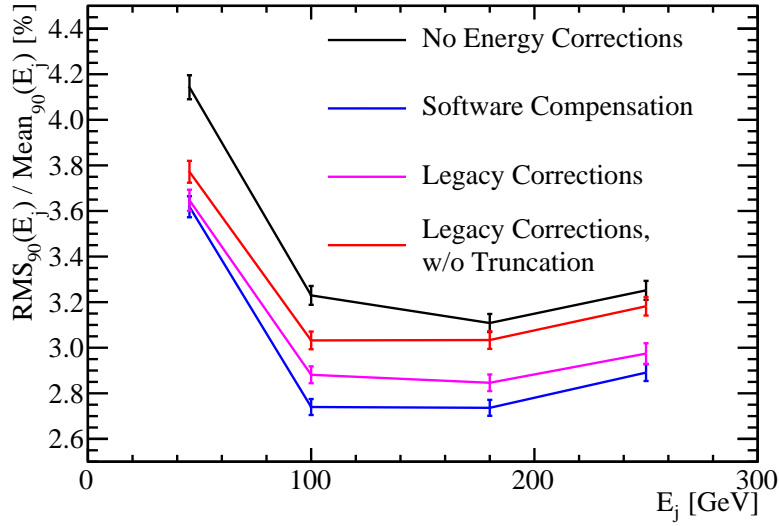


**Figure 1.18:** The energy resolution as a function of the MC energy for single  $K_L^0$  events using various energy correction settings. The black line represents no energy corrections, the blue line represents software compensation, the magenta line represents the legacy energy corrections and the red line represents the legacy corrections without the HCal hit energy truncation. The nominal ILD detector model was used in these simulations.

### 1.3.2.2 Impact on Jet Energy Resolution

The improvements in the intrinsic energy resolution of the detector observed when using software compensation will propagate into the reconstruction of jets. Figure 1.19 shows the jet energy resolution as a function of jet energy when using selected energy correction

configurations in PandoraPFA. It can be seen that software compensation is extremely beneficial to the detector performance as applying it improves the jet energy resolution by  $\sim 15\%$  across the energy range considered in comparison to using no energy corrections. Furthermore, software compensation offers an improvement in the jet energy resolution of the order of 5% for jet energies  $\gtrsim 100$  GeV in comparison to the legacy corrections, which prior to the development of software compensation had given the best jet energy resolutions.



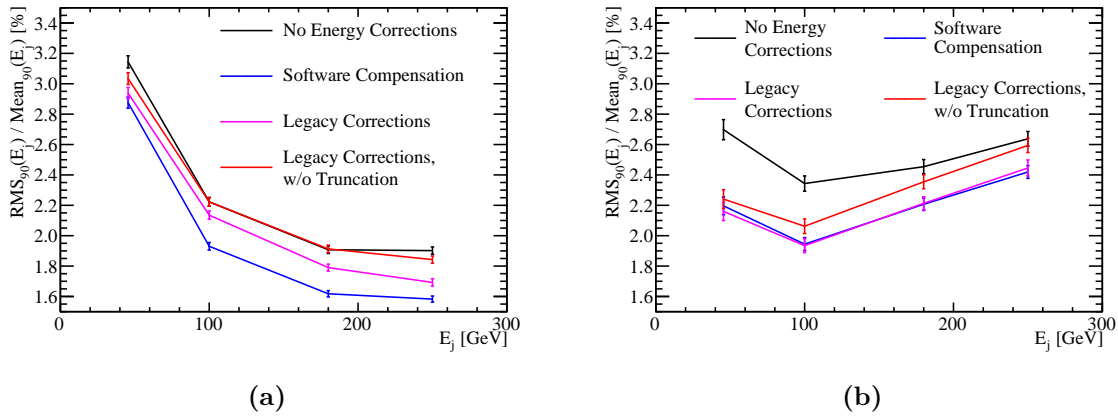
**Figure 1.19:** The jet energy resolution as a function of the jet energy for a variety of different energy correction options. The black line represents no energy corrections, the blue line represents software compensation, the magenta line represents the legacy energy corrections and the red line represents the legacy corrections without the HCal hit energy truncation. The nominal ILD detector model was used in these simulations.

Further light can be shed on these trends by examining the contributions to the jet energy resolution from the intrinsic energy resolution of the detector and the pattern recognition confusion, which are shown in figure 1.20. The intrinsic energy resolution contribution shows that software compensation is significantly better than all other energy corrections options, which is to be expected from the energy resolution studies presented in section 1.3.2.1. When compared to the legacy energy corrections, software compensation improves the intrinsic energy resolution by up to 12% across the energy range considered, with the largest improvement occurring for 100 GeV jets. As jets contain a broad spectrum of hadronic cluster energies, there is no jet energy for which the intrinsic energy resolution of the detector is comparable between the legacy corrections



and software compensation. The confusion contributions to the jet energy resolution when using software compensation and the legacy corrections are almost identical. This indicates that the improvement seen in the jet energy resolution when comparing software compensation to the legacy corrections, shown in figure 1.19, is being driven by the intrinsic energy resolution.

The "Clean Clusters" and "Scale Hot Hadrons" energy corrections, i.e. the legacy corrections without the HCal hit energy truncation, benefits the pattern recognition by reducing the confusion contribution. The confusion contribution is reduced by  $\sim 18\%$  for 45.5 GeV jets using these energy corrections, however, as the jet energy increases the magnitude of this improvement decreases, such that at 250 GeV jets no improvement is seen. These corrections do not significantly affect the intrinsic energy resolution of the detector. As these corrections benefit pattern recognition, selected aspects of their logic is applied to ECal hits in the software compensation energy correction as previously discussed.



**Figure 1.20:** The contributions to the jet energy resolution as a function of the jet energy for a variety of different energy correction options. The jet energy resolution contributions presented are (a) the intrinsic energy resolution of the detector and (b) the total confusion contribution. The jet energy resolution obtained using the standard reconstruction is given by the quadrature sum of the intrinsic energy resolution and total confusion contributions. The black line represents no energy corrections, the blue line represents software compensation, the magenta line represents the legacy energy corrections and the red line represents the legacy corrections without the HCal hit energy truncation. The nominal ILD detector model was used in these simulations.

### 1.3.3 Summary

The effect on both single particle and jet energy resolution of the HCal hit energy truncation and software compensation have been examined. Although relatively simplistic, the HCal hit energy truncation was found to be beneficial for detector performance by limiting the impact of Landau fluctuations. The more sophisticated software compensation procedure was found to be highly effective at producing a compensating calorimeter response across a wide range of energies, which translated into excellent performance in terms of jet energy resolution.

## 1.4 Timing Cuts

The linear collider will operate using trigger-less readout whereby the recorded data for each sub-detector is read out between collisions of the  $e^+$  and  $e^-$  bunches. The bunch train structure for ILC and CLIC is compared in table 1.1. Event selection will proceed through the application of a software trigger. This involves the identification of any hard interactions, prior to full event reconstruction, and only putting data into the event reconstruction if it is measured within a chosen time window about these interactions. The recorded time of a calorimeter hit, which is cut on to make the time window for the software trigger, is corrected for straight time-of-flight to the IP. This ensures that the amount of time particle showers have to develop in the calorimeters is independent of their position in the detector. As the width of this time window changes, the amount of time particle showers have to develop changes, which will affect the performance of the detector.

At CLIC, the application of a software trigger is challenging due to the 0.5 ns bunch separation. At CLIC, the integration time of the calorimeters will span many bunch crossings and, due to the intense beam-induced backgrounds, identifying energy deposits produced from the hard interaction of interest is difficult. By placing tight timing constraints on the energy deposits made in the CLIC calorimeters, it is possible to minimise the impact of the beam-induced backgrounds. As well as minimising the impact of the backgrounds, these tight timing requirements will also truncate the shower development time for hard interaction of interest. Understanding the impact of these timing requirements on physics performance is vital to the success of the CLIC experiment. Timing cuts are not as challenging at the ILC because the bunch separation is much larger,

meaning the calorimeters could be read out between bunches, and the beam-induced backgrounds are much smaller.

	ILC 500 GeV	CLIC 3 TeV
Electrons per bunch [ $10^{10}$ ]	2.0	0.37
Bunches per train	2820	312
Train repetition rate [Hz]	5	50
Bunch separation [ns]	308	0.5

**Table 1.1:** The train structure for 500 GeV ILC and 3 TeV CLIC [9, 11].

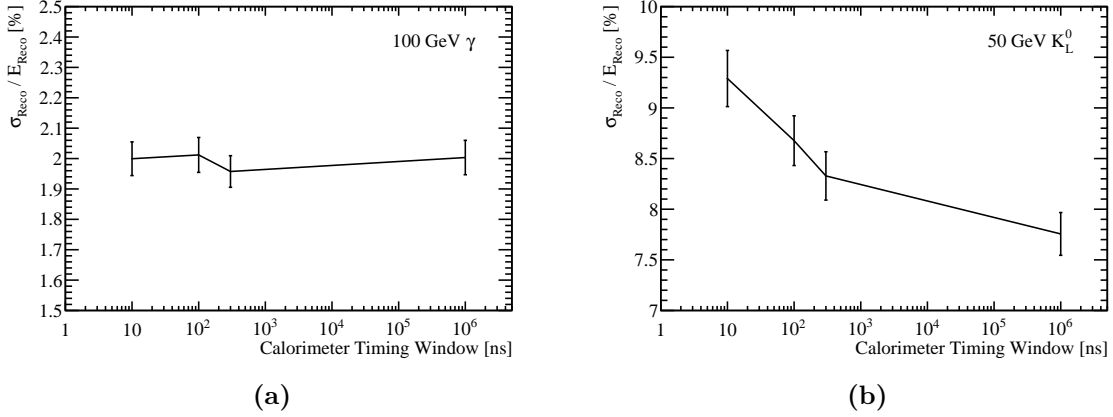
For all choices of time window considered in this study the calibration procedure described in section 1.2 was reapplied. This ensures that the mean of the reconstructed energy distributions will be invariant to changes in the calorimeter timing window as the calibration procedure compensates for any energy losses incurred by truncating the particle shower development time.

For results presented in this chapter and the optimisation studies found in chapter ?? a 100 ns timing window was applied to all detector models considered. As the choice of timing window has yet to be finalised for the linear collider this value was chosen as it represents something that could be achieved using the readout technology options presently available [12].

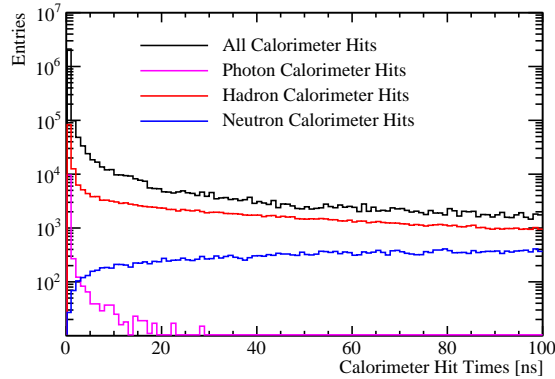
### 1.4.1 Impact on Single Particle Energy Resolution

The energy resolution for 100 GeV  $\gamma$  and 50 GeV  $K_L^0$  events as a function of the timing window applied to the calorimeter hits is shown, for the nominal ILD detector, in figure 1.21. The timing cut makes little difference to the energy resolution of the  $\gamma$  events, however, the energy resolution for neutral hadrons gets significantly worse as the time window is reduced. The neutral hadron energy resolution becomes worse by almost 20 % when the time window is reduced from  $10^6$  ns to 10 ns. These trends are to be expected as electromagnetic showers develop far more rapidly than their hadronic counterparts [5]. This can be seen in figure 1.22, which shows the distribution of the measurement time of calorimeter hits, corrected for time-of-flight, for selected shower components within 91 GeV  $Z \rightarrow uds$  di-jet events. Hadronic showers develop slowly as they often involve intermediate states that must decay to continue the propagation of

the shower. If a narrow calorimeter timing window is used, energy measurements from the hadronic shower will be lost and the energy resolution will degrade, which is what is observed.



**Figure 1.21:** The energy resolution as a function of calorimeter timing window for (a) 100 GeV  $\gamma$  events and (b) 50 GeV  $K_L^0$  events using the nominal ILD detector model.

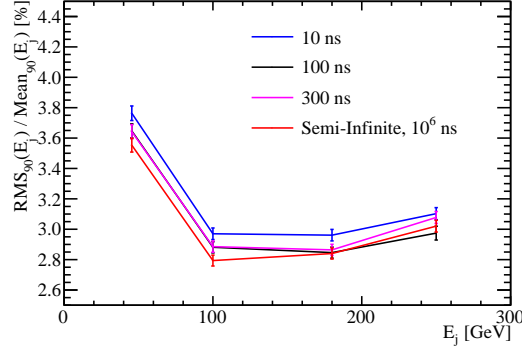


**Figure 1.22:** The distribution of the time of the calorimeter hits, corrected for time of flight to the impact point, for 91 GeV  $Z \rightarrow uds$  di-jet events. Electromagnetic shower energy deposits, i.e. photon calorimeter hits, are deposited very rapidly, while hadronic shower energy deposits are deposited over a much longer time period.

### 1.4.2 Impact on Jet Energy Resolution

The jet energy resolution as a function of the jet energy for selected calorimeter time windows is shown in figure 1.23. As expected, the jet energy resolution will also be affected by the reduced neutral hadron energy resolution when the calorimeter timing

window is reduced. The sole exception to this is the 250 GeV jets for the 100 ns time window whereby the jet energy resolution is slightly better than when using the 300 ns and semi-infinite time windows, although this is not statistically significant.



**Figure 1.23:** The jet energy resolution as a function of jet energy for various calorimeter timing cuts. The nominal ILD detector model was used in these simulations.

The time window applied to the calorimeter hits affects both the neutral hadron and jet energy resolutions with a larger timing window leading to better resolutions. It can be seen that applying an aggressive choice of time window, such as 10 ns, the jet energy resolution is degraded as many of the hadronic showers being sampled do not have time to fully develop. However, even using a 10 ns timing cut the jet energy resolutions are still sufficiently low to give excellent detector performance. Both the single particle and jet energy resolutions indicate that the majority of hadronic showers will have fully developed within 100 ns and that there are little gains to be made by extending the size of this window.

### 1.4.3 Summary

Simulations were performed where the size of the calorimeter hit time window used in the software trigger at the linear collider experiments was varied to determine what impact it had on physics performance. The energy resolution for electromagnetic showers did not change significantly when varying the size of time window, however, the neutral hadron energy resolution became worse as the size of the time window was reduced. The jet energy resolution was also sensitive to the size of the time window used, however, the trend was far weaker than that seen for neutral hadrons. There were no major benefits to increasing the time window beyond 100 ns indicating that this time is sufficient for the bulk of hadronic showers to develop.



# Colophon

This thesis was made in L<sup>A</sup>T<sub>E</sub>X 2<sub>ε</sub> using the “hepthesis” class [\[13\]](#).





# Bibliography

- [1] C. W. Fabjan and F. Gianotti. Calorimetry for particle physics. *Rev. Mod. Phys.*, 75:1243–1286, 2003.
- [2] M. A. Thomson. Particle Flow Calorimetry and the PandoraPFA Algorithm. *Nucl. Instrum. Meth.*, A611:25–40, 2009.
- [3] M. Derrick, D. Gacek, N. Hill, B. Musgrave, R. Noland, E. Petereit, J. Repond, R. Stanek, and K. Sugano. Design and construction of the ZEUS barrel calorimeter. *Nucl. Instrum. Meth.*, A309:77–100, 1991.
- [4] Huong Lan Tran, Katja Krüger, Felix Sefkow, Steven Green, John Marshall, Mark Thomson, and Frank Simon. Software compensation in Particle Flow reconstruction. 2017.
- [5] R. Wigmans. Calorimetry: Energy measurement in particle physics. *Int. Ser. Monogr. Phys.*, 107:1–726, 2000.
- [6] L. Landau. On the energy loss of fast particles by ionization. *J. Phys.(USSR)*, 8:201–205, 1944.
- [7] H. Bichsel, Donald E. Groom, and S. R. Klein. Passage of particles through matter. 2004.
- [8] Oskar Hartbrich. AHCAL Digitisation. CALICE Spring Meeting 2015 & LCWS 2015, Tsukuba (Japan), 19 Apr 2015 - 24 Apr 2015, Apr 2015.
- [9] Halina Abramowicz et al. The International Linear Collider Technical Design Report - Volume 4: Detectors. 2013.
- [10] C. Adloff et al. Hadronic energy resolution of a highly granular scintillator-steel hadron calorimeter using software compensation techniques. *JINST*, 7:P09017, 2012.
- [11] Lucie Linssen, Akiya Miyamoto, Marcel Stanitzki, and Harry Weerts. Physics and

Detectors at CLIC: CLIC Conceptual Design Report. 2012.

- [12] C. Adloff et al. The Time Structure of Hadronic Showers in highly granular Calorimeters with Tungsten and Steel Absorbers. *JINST*, 9:P07022, 2014.
- [13] Andy Buckley. The hepthesis L<sup>A</sup>T<sub>E</sub>X class.