# Something something something physics

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 lumi-

nosities 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.

Andy Buckley

# Acknowledgements

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

# Preface

This thesis describes my research on various aspects of the LHCb particle physics program, centred around the LHCb detector and LHC accelerator at CERN in Geneva.

For this example, I'll just mention Chapter **??** and Chapter **??**.

x

# Contents

*"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

In the particle flow paradigm measured energy deposits fall into two distinct classes (i) calorimetric energy deposits and (ii) charged particle tracks. Calorimetric energy deposits are the essential building blocks for the application of the pattern recognition aspect of particle flow. The separation of energy deposits from charged and neutral particles in the calorimeters is crucial for achieving good energy resolutions, which is only possible if the energy estimators for those energy deposits are accurate.

The ultimate goal of the calibration procedure outlined in this chapter is to obtain the best energy estimator for particles showering in the calorimeters. The energy of a cluster of calorimeter hits is:

$$E_{Cluster} = \sum_{ECal\ hits,\ i} E^i_{ECal} + \sum_{HCal\ hits,\ i} E^i_{HCal} \, , \qquad (1.1)$$

where $E^i_{ECal}$ is the energy of a calorimeter hit $i$ in the ECal, $E^i_{HCal}$ is the energy of a calorimeter hit $i$ in the HCal and $\sum$ runs over all hits in a given calorimeter. This

is a naive energy estimator, but it is a starting point for the development of more sophisticated procedures aimed at improving detector performance. However, before going further the calorimeter hit energies must be determined, which for a sampling calorimeter is a non-trivial task. Sampling calorimeters, such as those found at the linear collider, are comprised of alternating layers of active and absorber materials [15]; The absorber layers are designed to initiate and propagate particle shower development, while the active layers are designed to give a response that is proportional to the energy deposited within them. As the response of a sampling calorimeter gives a measure of the energy deposited in the active layers, the total energy deposited in the calorimeter has to be estimated based on the active layer energies. It is assumed that the energy deposited the active and absorber layers of a sampling calorimeter are the proportional to each other, which makes the total hit energy is proportional to the active layer hit energy. This estimation procedure is loosely referred to as digitisation. In this way the cluster energy estimator introduced above can be written as:

$$E_{Cluster} = \sum_{ECal\ hits,\ i} \epsilon^i_{ECal} \alpha_{ECal} + \sum_{HCal\ hits,\ i} \epsilon^i_{HCal} \alpha_{HCal} \ , \qquad (1.2)$$

where $\epsilon^i_{ECal}$ is the energy response in the active medium for a calorimeter hit $i$ in the ECal, where $\epsilon^i_{HCal}$ is the energy response in the active medium for a calorimeter hit $i$ in the HCal and $\sum$ runs over all hits in a given calorimeter. 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. A compensating calorimeter produces an identical response to electromagnetically and hadronically showering particles. Depending on the properties of the materials used, calorimeters can have different responses to electromagnetic and hadronic particle showers. The primary cause of this different response is the invisible energy component found in hadronic showers. As this component cannot be measured by the calorimeter, the response of a calorimeter is typically lower to hadronically showering particles than to electromagnetically showering particles. The invisible component exists due to a combination of various effects, such as neutrons stopping within the calorimeter and nuclear binding energy losses. If left unchecked, this difference would

lead to a systematic loss of energy for hadronic showers that would harm detector performance.

There are two distinct routes available for negating this unwanted effect and achieving a compensating response from a calorimeter: The first is hardware compensation, whereby calorimeters are constructed using materials that yield extra energy in response to hadronic showers and the second is software compensation, whereby the uncompensated calorimetric energies are corrected based on whether the shower is hadronic. 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:

$$E_{EM\ Cluster} = \sum_{ECal\ hits,\ i} E^i_{ECal} \beta^{EM}_{ECal} + \sum_{HCal\ hits,\ i} E^i_{HCal} \beta^{EM}_{HCal} \qquad (1.3)$$

and the hadronic cluster energy to:

$$E_{Had\ Cluster} = \sum_{ECal\ hits,\ i} E^i_{ECal} \beta^{Had}_{ECal} + \sum_{HCal\ hits,\ i} E^i_{HCal} \beta^{Had}_{HCal} \qquad (1.4)$$

where the $\beta$s are constant scaling factors that are applied to the calorimeter hit energies associated to electromagnetic and hadronic clusters in the ECal and HCal. This simple scaling of energies compensates the response of the calorimeters at the fixed energy where the $\beta$s are calibrated, 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 energy. As such, more sophisticated software techniques, which makes use of the fine segmentation of the linear collider calorimeters, 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. In addition to obtaining a compensating calorimeter response,

these techniques address the problem of spuriously high energy calorimeter hits that if left unchecked would damage the reconstruction. These high energy hits are caused by Landau fluctuations [18], which originate from high energy knock-on electrons appearing within particle showers [8]. In this chapter, following a description of the calibration procedure outlined above, is an explanation of these novel energy estimators and the impact they have on detector performance.

This chapter concludes with a study determining the impact on detector performance of timing cuts that are applied to the calorimeter hits. These cuts form part of the software trigger that will be used at the linear collider experiment. Details regarding how all the detector performance metrics used in this chapter are calculated can be found in section 2.2.

**Hardware Compensation**

A novel example of hardware compensation would be the ZEUS calorimeter [14]. 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.

**Calibration and detector optimisation**

Optimising the detector at a future linear collider will be crucial to exploit the full physics potential available to it. An extensive optimisation of the calorimeters was performed and the results can be found in chapter 2.1. For each detector model considered in that study, the calibration procedure outlined in section 1.2 was applied to ensure optimal energy estimators were used to quantify the detector performance. This made unbiased comparison between detector models performance possible and ensured reliability in the conclusions drawn from that study.

## 1.2 Calibration in the particle flow paradigm

The calibration procedure described in the following section discusses how the digitisation constants, $\alpha$, and the scaling factors $\beta$ introduced in section 1.1 are determined. In addition, two further operations are included in the general calibration procedure designed to ensure optimal detector performance. These are minimum ionising particle (MIP) scale setting and retraining of likelihood data used in the reconstruction.

The MIP scale is the average energy response, on a per hit level, of a calorimeter when a normally incident MIP passes through it. This scale must be set in the digitisation processor as the response of the readout technology is simulated in this processor, which includes a maximum readout value set in units of MIPs. The digitiser also applies a minimum threshold on the active layer hit energy, in units of MIPs, for calorimeter hit to be created. PandoraPFA uses the MIP scale to place further threshold cuts on the hit energy that must be exceeded for a calorimeter hit to be used in the reconstruction. Both of these thresholds are designed to veto noise that would be present in a real detector. Although noise is not applied in any simulations presented in this work, these cuts are applied to better reflect the performance of a real detector. The MIP scale in the digitiser and PandoraPFA, while intrinsically linked, have to be calculated separately as the digitiser requires the MIP scale to be defined using the active layer hit energies while, PandoraPFA requires the definition from the full hit, the active and absorber layer, energies. It may be expected that the digitisation MIP scale could be related to the PandoraPFA MIP scale by the digitisation constants, however, this is not necessarily the case as additional cuts are applied to the full calorimeter hit energies that are not applied to the active layer hit energies e.g. timing cuts. Therefore, the conservative approach of independently recalculating the MIP scale for PandoraPFA was taken.

The likelihood data that is used in the reconstruction consists of a series of probability density functions (PDFs), which describe a number of topological properties of clusters of calorimeter hits. This data is used to determine whether a cluster of hits corresponds to an electromagnetically showering particle. The topological properties used in these PDFs are related solely to the ECal, which means that every time the ECal is changed from that of the nominal ILD detector this likelihood data must be retrained. Therefore, it is natural to included this retraining a description of the calibration procedure that will be used for calorimeter optimisation.

To calibrate the $\alpha$ and $\beta$ constants reconstructed energy distributions, of either clusters of calorimeter hits for the $\alpha$ or PFO energies for the $\beta$, are produced and their mean value tuned by varying these constants. However, when filling these distributions a number of cuts are applied to the data to ensure the correct constant 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 whereby the calibration is repeated several times, each time updating the calibration constants, until a certain accuracy in the mean of these reconstructed energy distributions is achieved.

**Ordering of Calibration Procedure**

The calibration procedure is broadly split into four separate operations: determination of digitisation constants, determination of scaling factor constants, MIP scale setting and retraining of likelihood data. The ordering of each of these calibration steps is crucial as it is possible to get interference between the different stages if applied in an arbitrary order. With that in mind, the procedure order to minimise interference is as follows:

1. MIP scale setting in the digitiser. No dependancies on other parameters.

2. Digitisation constant, $\alpha$, setting. This relies upon the MIP scale in the digitiser being set.

3. MIP scale setting in the PandoraPFA. This relies upon correct calibration of calorimeter hit energies, therefore, the $\alpha$s must be calibrated first.

4. Scaling factor constant, $\beta$, setting. This relies upon correct calibration of calorimeter hit energies and correct MIP scale setting in PandoraPFA.

5. If required, retrain likelihood data and then repeat the first four steps using the updated likelihood data. This relies upon the PFO energies being correct, therefore, the $\beta$s must be calibrated first.

## 1.2.1 Digitisation Implementation

Digitisation. in this sense, refers to the process whereby an estimation made of the energy deposited in a calorimeter hit, in both the active and absorber layers, based on the energy deposited in the sensitive element, the active layers. The relationship between the energy deposited in the active layers and the absorber layers of a calorimeter is linear as the energy deposited in both layers is proportional to the number of charged particle tracks, or $\gamma$s depending on the material choices, passing through them. This works under the assumption that the density of charged particle tracks across a calorimeter hit within a particle shower is uniform. The ratio of the calorimeter hit energy to the energy deposited in the active layers is hereby called the digitisation constant. 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.

**ECal Digitisation Implementation**

The procedure for determining the digitisation constants in the ECal involves simulation of single $\gamma$ events at energy $E_{MC} = 10$ GeV. $\gamma$ events are ideal for calibration of the ECal as $\gamma$s are, at this energy, largely contained within the ECal, as shown in figure 1.1a. 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 $\gamma$. The polar angle cut veto events where the $\gamma$ goes down the beam pipe. $\gamma$ conversions 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 $\gamma$ events is shown in figure 1.1b.

The calibration of the digitisation in the ECal is an iterative procedure, which begins with the simulation of single $\gamma$ events using a trial calibration, with digitisation constant in the ECal $\alpha_{ECal}^0$. 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 1.1b. For an ideal calorimeter this distribution should be Gaussian, as was described in chapter 2, therefore, a Gaussian fit is applied to this distribution and the mean, $E_{Fit}$, extracted. To remove the effect of any outliers in this distribution, the fit is applied to
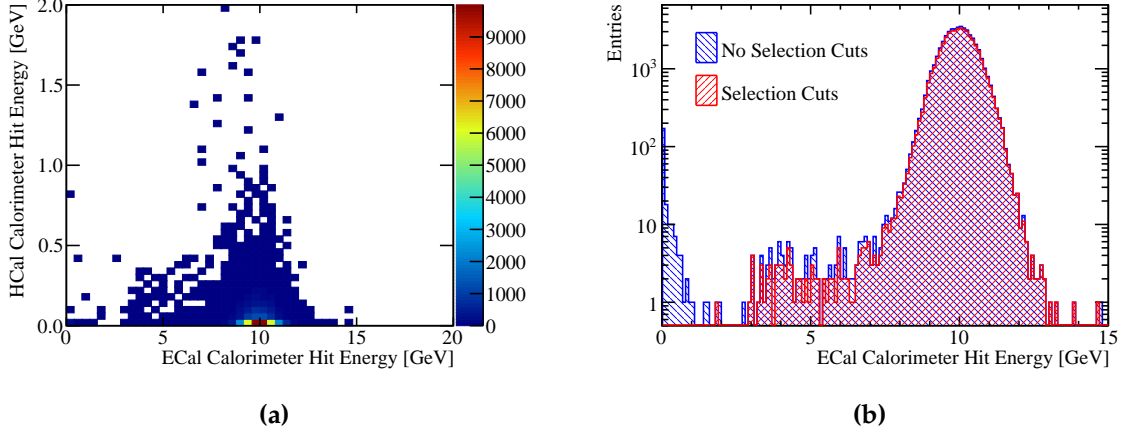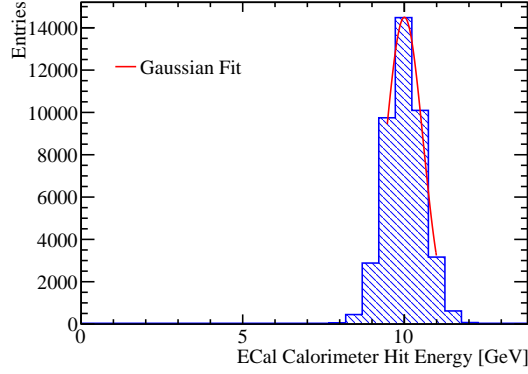
**Figure 1.1:** (a) The sum of calorimeter hit energies in ECal and HCal for 10 GeV $\gamma$ events. (b) The sum of the ECal calorimeter hit energies for 10 GeV $\gamma$ events with and without the selection cuts.

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 1.2. In the case of ideal calibration the mean of this fit, $E_{\text{Fit}}$, would be equal $E_{MC}$. It is assumed that any difference between the two is due to the calibration and to correct for this the digitisation constant from the trial calibration, $\alpha^0_{\text{ECal}}$, is rescaled by the ratio of the $E_{MC}$ to $E_{\text{Fit}}$:

$$\alpha^0_{\text{ECal}} \rightarrow \alpha_{\text{ECal}} = \alpha^0_{\text{ECal}} \times \frac{E_{MC}}{E_{Fit}} \, , \tag{1.5}$$

This procedure is then repeated until the $E_{\text{Fit}}$ falls within a specified tolerance. 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. This tolerance is somewhat loose, however, it is tight enough to ensure successful application of PFA. It should also 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.

**Figure 1.2:** Gaussian fit to sum of the ECal calorimeter hit energies for 10 GeV $\gamma$ events with selection cuts. The coarse binning reflects the tolerance on the digitisation constant calibration.

**HCal Digitisation Implementation**

The calibration for the digitisation in the HCal proceeds in a similar manor to that described for the ECal with a few key differences. This calibration uses $K_L^0$ events at $E_{MC} = 20$ GeV as these neutral hadrons will deposit the bulk of their energy in the HCal. The higher energy, with respect to the ECal digitisation, is used to create larger particle showers that sample deeper into the HCal. As the $K_L^0$s must pass through the ECal before arriving at the HCal and, as the ECal contains $\approx 1\lambda_I$, some of the $K_L^0$s begin showering in the ECal, as shown by figure 1.3a. Such events are unsuitable for calibration of the HCal digitisation constants as rescaling $\alpha_{HCal}^0$ would not lead to a linear rescaling in $E_{Fit}$. These events are vetoed in the even selection by requiring events deposit less than less than 5% of $E_{MC}$ outside of the HCal. In addition to this, the last layer of the HCal where energy is deposited is required to be in the innermost 90% of the HCal. This cut vetoes events that shower late in the HCal and deposit a significant amount of energy in the uninstrumented coil region of the detector. 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 1.3b. There are two HCal digitisation constants used in the detector simulation, one applied for the barrel and another for the endcap. This is to account 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 manor, 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

are conservative to account for the transverse profile of the hadronic showers and ensure that they are confined to the relevant sub-detector. One further difference to the ECal digitisation procedure is that the target reconstructed energy for the $K_L^0$ samples is the kinetic energy as opposed to the total energy. As the majority of the neutral hadrons appearing in jets are neutrons and their accessible energy is their kinetic energy, calibrating to the kinetic energy should give the best performance for jet reconstruction.
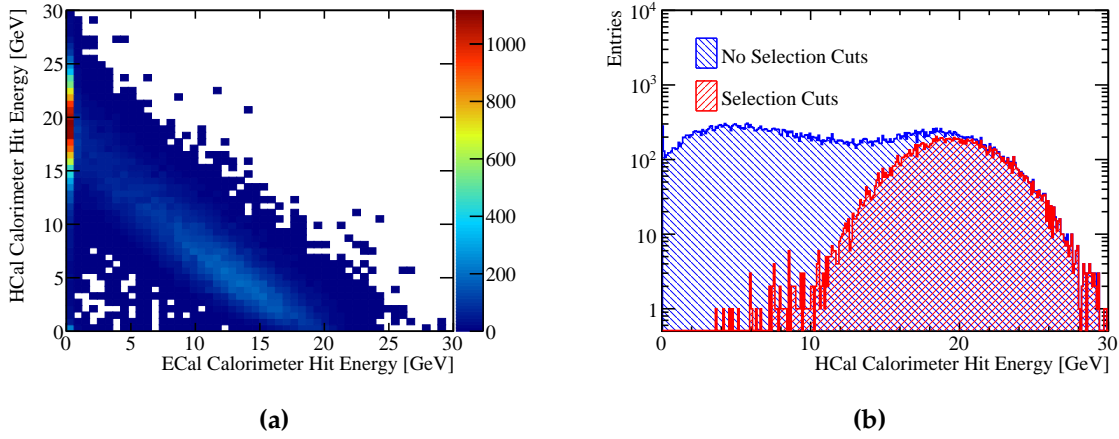


**Figure 1.3:** (a) Sum of calorimeter hit energies in ECal and HCal for 20 GeV $K_L^0$ events. (b) Sum of the HCal calorimeter hit energies for a 20 GeV $K_L^0$ events with and without the selection cuts.

Using these cuts the calibration procedure for the digitisation of the HCal barrel and endcap proceeds in the same manor as was described for the ECal, the details of which can be found in section 1.2.1. Examples of the Gaussian fits applied to the sum of the calorimeter hit energies in the HCal barrel and endcap can be found in figure 1.4.

**HCal Ring Digitisation Implementation**

The HCal ring, illustrated in figure 1.5, also has an independent digitisation constant to account for any difference in the hadronic shower development between the ring, barrel and endcap. The procedure used to calibrate this constant has to differs from that presented in section 1.2.1 as it is unfeasible, due to the depth of the ring, to produce events that are wholly contained within it. Fortunately, the size of the HCal ring means it plays a minimal role in the reconstruction, so precise calibration is not crucial. To

**(a)** HCal barrel.
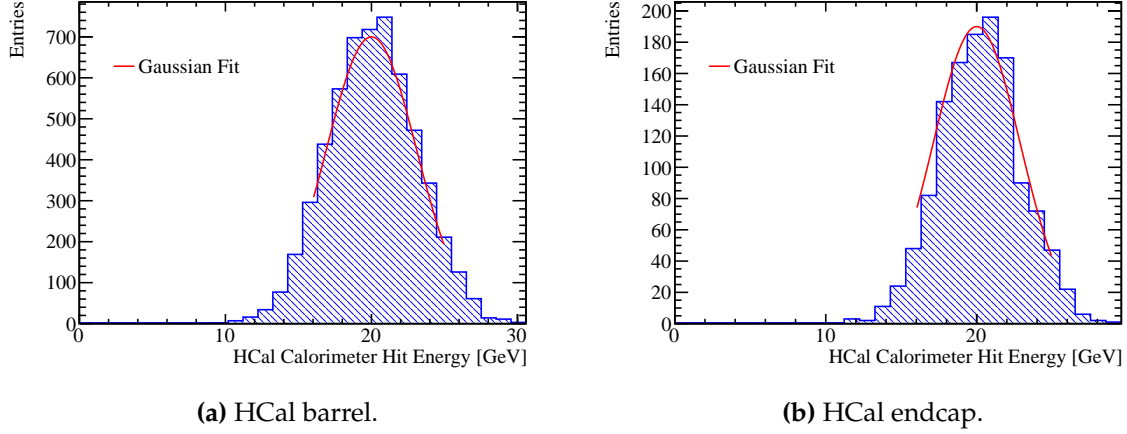
**(b)** HCal endcap.

**Figure 1.4:** Gaussian fit to sum of the HCal calorimeter hit energies for 20 GeV $K_L^0$ events with selection cuts.

ensure that the calibration is approximately correct for the HCal ring, $\alpha_{\text{HCal ring}}$ is assumed to equal $\alpha_{\text{HCal endcap}}$ multiplied by several factors designed to accounts for changes in the active layer thickness, absorber layer thickness and the MIP response between the HCal endcap and ring. In detail:

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

where $\theta$ is the incident angle of the incoming particle to the calorimeter determined using the 20 GeV $K_L^0$ events, $L^{Active}$ is the active layer thickness and $L^{Absorber}$ 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 can be found in section 1.2.2.

## 1.2.2  MIP Scale Determination

The digitiser MIP scale was defined as the, non-zero, peak in the distribution of the active layer calorimeter hit energies for normally incident 10 GeV $\mu^-$ [8], as shown in figure 1.6. This distribution is made for each calorimeter where the MIP scale needs
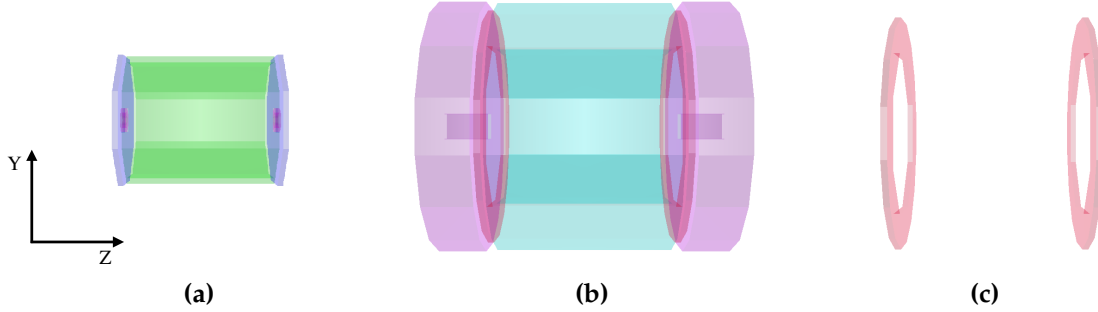
**Figure 1.5:** A PandoraPFA event display showing the nominal ILD calorimeters. (a) the ECal, (b) the full HCal and (c) the HCal ring.

to be determined. As the average energy deposited per hit in a given sub-detector is relevant for setting the MIP scale, as opposed to the total energy deposited in a sub-detector, no selection cuts are required. The sample of $\mu^-$ events used for setting the MIP scale were generated isotropically. When populating the active layer energy distribution a direction correction factor of $\cos(\theta)$, where $\theta$ is the incident angle of the $\mu^-$ to the calorimeter hit, was applied to the hit energy to generate the effect of having the $\mu^-$ enter the calorimeter at normal incidence. The MIP scale was determined separately for the ECal, HCal barrel, HCal endcap and HCal ring, however, only a single HCal MIP scale, taken as the HCal barrel, was required by the digitisation processor. The HCal endcap and ring MIP scales were calculated for the purposes of the HCal ring digitisation described in section 1.2.1. No MIP scale setting was required in the digitisation processor for the muon chamber.

A similar procedure was employed for calculation of the MIP peak in PandoraPFA. One important difference was the distribution used for setting the MIP scale in PandoraPFA is the distribution of calorimeter hit energies, i.e. the energy in the active and absorber layers of a hit, and not just the active layer energies. Examples of the distributions used to set the MIP scale in PandoraPFA can be found in figure 1.7. There are few populated low calorimeter hit energy bins in this these distributions due to cuts applied in the digitiser on the active layer energy. The double peak structure observed in the ECal calorimeter hit energy distribution is due to the ECal absorber material thickness doubling in the back 10 layers of the ECal. Further differences between the MIP scale setting in the digitiser and PandoraPFA are as follows: The MIP scale setting in PandoraPFA combines the HCal sub-detectors, the barrel, endcap and ring, together when creating the calorimeter hit energy distributions; PandoraPFA

**(a)**

**(b)**

**(c)**

**(d)**
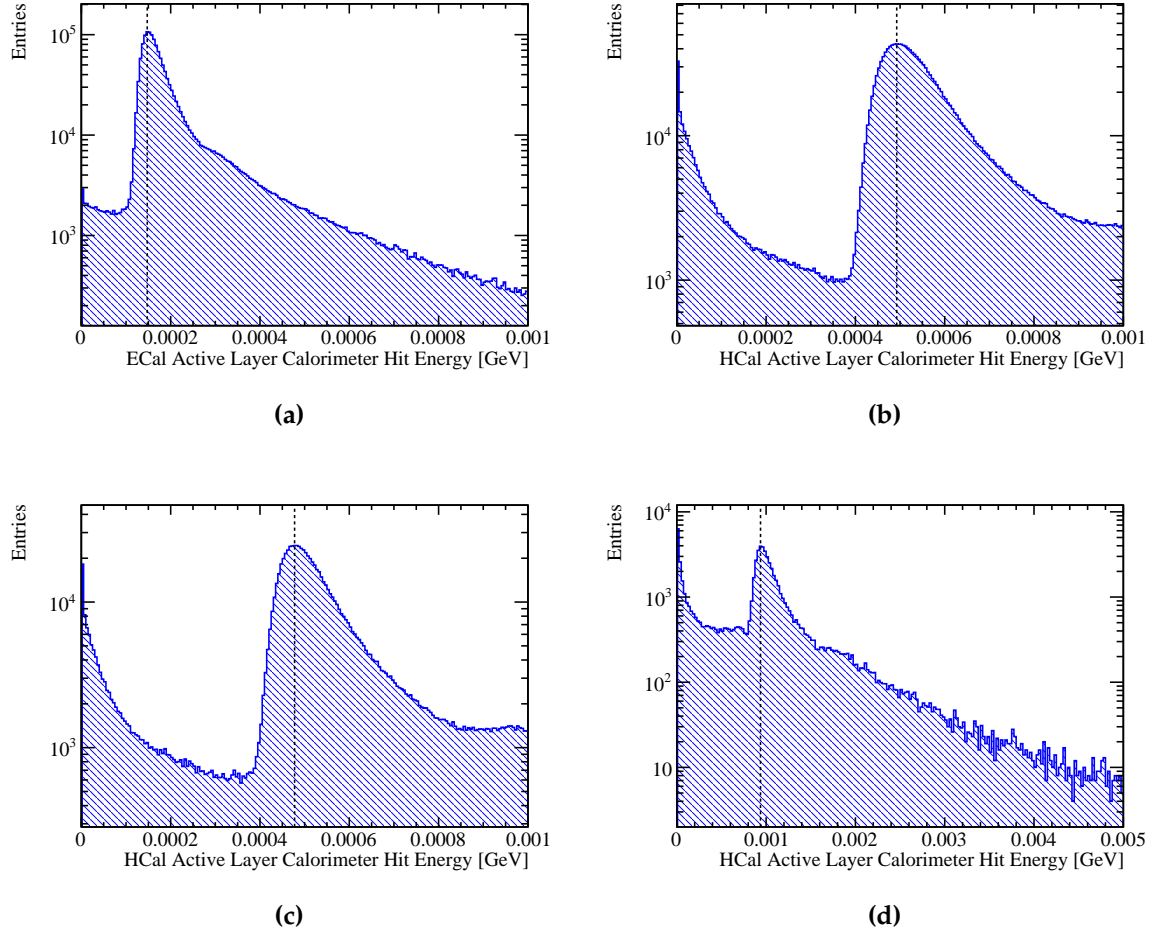
**Figure 1.6:** The active layer calorimeter hit energy distributions for (a) the ECal, (b) the HCal barrel, (c) the HCal endcap and (d) the HCal ring for 10 GeV $\mu^-$ events.

requires the MIP scale to be set in the muon chamber unlike the digitisation processor and, therefore, the muon chamber hit energy distribution must also be created.

### 1.2.3 Electromagnetic and Hadronic Scale Setting

**Electromagnetic scale setting**

The electromagnetic scale in the ECal, $\beta_{ECal}^{EM}$, is determined using $\gamma$ events at $E_{MC} = 10$ GeV. $\gamma$ events are ideal for the setting of the electromagnetic scale as they procedure electromagnetic showers and are primarily confined to the ECal at the energy considered, which was shown in figure 1.1a. A cut requiring less than 1% of the reconstructed

**(a)**



**(b)**



**(c)**
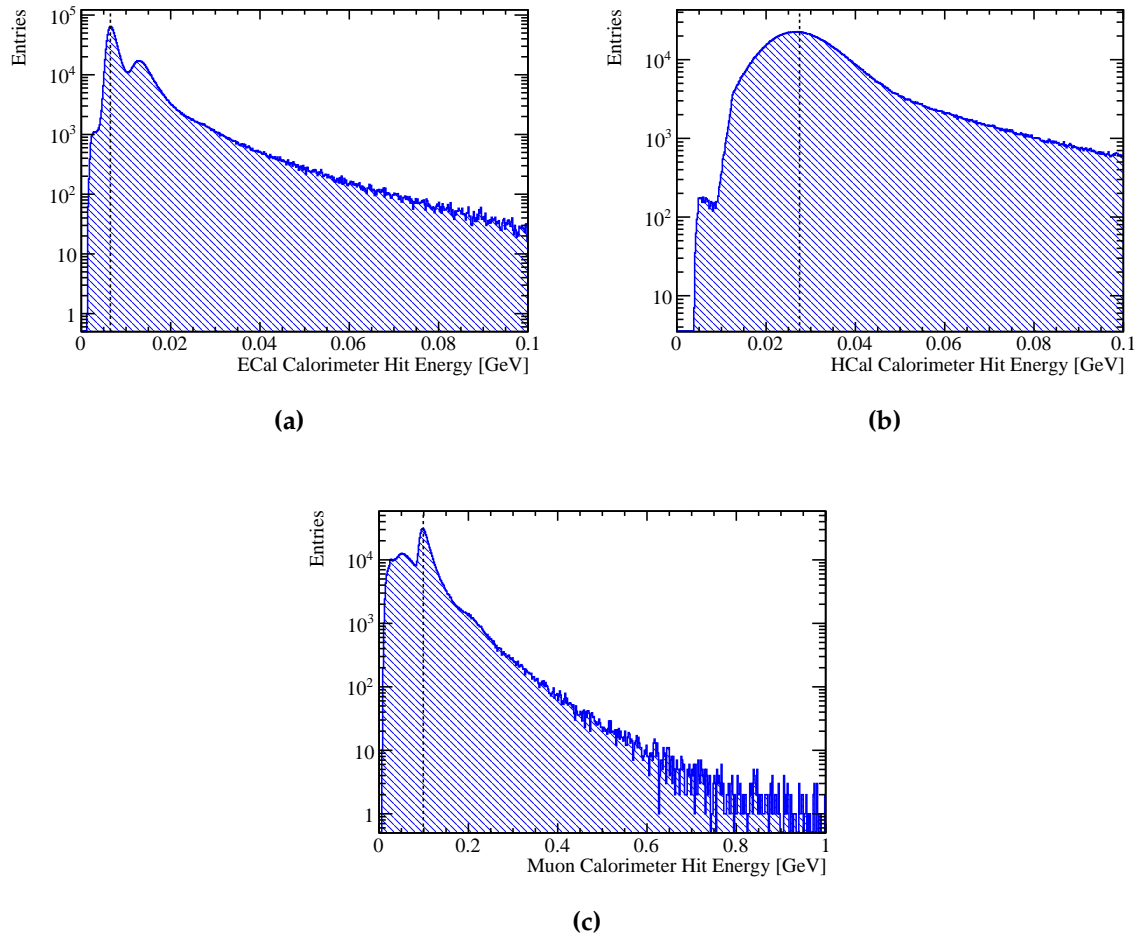
**Figure 1.7:** The calorimeter hit energy distributions for (a) the ECal, (b) the HCal and (c) the muon chamber for 10 GeV $\mu^-$ events.

energy to be found outside the ECal is applied to ensure that only events where the bulk of the energy is deposited within the ECal are used for this part of the calibration. Furthermore, a cut requiring a single $\gamma$ be reconstructed are added to veto events with pattern recognition failures. $\gamma$ conversions are excluded at MC level to ensure energy measurements used in the calibration arise from the calorimeters and not the charged particle tracks. The impact of these cuts on the electromagnetic energy measured in the ECal for 10 GeV $\gamma$ events is shown in figure 1.8a.

The fitting procedure follows that used for the ECal digitisation, described in section 1.2.1, whereby a trial calibration for the electromagnetic energy scale in the ECal, $\beta_{ECal}^{EM0}$, is assumed and the single $\gamma$ events simulated. The distribution of the electromagnetic energy in the ECal is created and a Gaussian fit applied to the range of data with the smallest root mean square containing at least 90 % of the data. The mean of this fit, $E_{\text{Fit}}$, is then used to scale $\beta_{ECal}^{EM0}$ in the following way:

$$\beta_{ECal}^{EM0} \to \beta_{ECal}^{EM} = \beta_{ECal}^{EM0} \times \frac{E_{MC}}{E_{Fit}} \, , \tag{1.7}$$

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 $\beta_{ECal}^{EM}$ 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 energies that are used in downstream analyses.

The electromagnetic scale in the HCal, $\beta_{HCal}^{EM}$, is chosen to be equal to the hadronic scale in the HCal, $\beta_{HCal}^{Had}$. The details of the determination of $\beta_{HCal}^{Had}$ can be found in section 1.2.3. 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.

**Hadronic scale setting**

The hadronic scale in the ECal, $\beta_{ECal}^{Had}$, is important to detector performance as a non-negligible amount of hadronic energy will be measured in the ECal. As the ECal
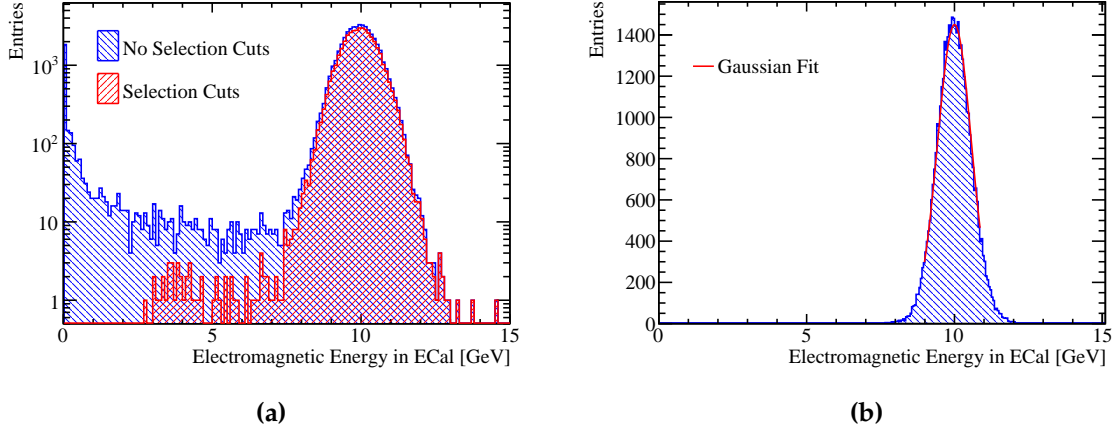
**Figure 1.8:** (a) The sum of the electromagnetic energy measured in the ECal for 10 GeV $\gamma$ events with and without the selection cuts. (b) Gaussian fit to sum of the electromagnetic energy deposited in the ECal for 10 GeV $\gamma$ events with selection cuts. The fine binning reflects the tolerance on the electromagnetic scale calibration constant in the ECal.

contains $\approx 1\lambda_I$, the hadronic scale in the ECal cannot be independently set as it is unfeasible to create a large sample of 20 GeV $K_L^0$ events that are fully contained within it. Therefore, the hadronic scale in the ECal and HCal have to be set simultaneously. Cuts are applied to select the $K_L^0$ events that are appropriate to use for the hadronic scale calibration. To ensure the events do not suffer from the effect of leakage out of the back of the HCal, a cut is applied requiring the last layer in which energy is deposited in the HCal to be in the first 90% of the HCal. In addition to this, a cut requiring a single neutral hadron to be reconstructed is applied to veto events with reconstruction failures. Finally, the total hadronic energy measured in ECal and HCal, $E_{ECal}^{Had} + E_{HCal}^{Had}$, must fall within $3\,\sigma$ of the desired hadronic energy distribution, $E_{ECal}^{Had} + E_{HCal}^{Had} = 20\text{GeV} - m_{K_L^0} = E_K$. Where $\sigma$ is defined to be $55\% \times \sqrt{E} = 2.46\text{GeV}$ for 20 GeV $K_L^0$. This definition for $\sigma$ is derived from the energy resolution for neutral hadrons using the nominal ILD HCal [5]. 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. Again, the target reconstructed energy for this sample is the kinetic energy and not the total energy of the $K_L^0$ for the reasons outlined in section 1.2.1. The impact of cuts are illustrated in figure 1.9.

This part of the calibration procedure is again iterative and begins by assuming trial values, $\beta_{ECal}^{Had0}$ and $\beta_{ECal}^{Had0}$, for the hadronic scale calibration factors $\beta_{ECal}^{Had}$ and $\beta_{ECal}^{Had}$.

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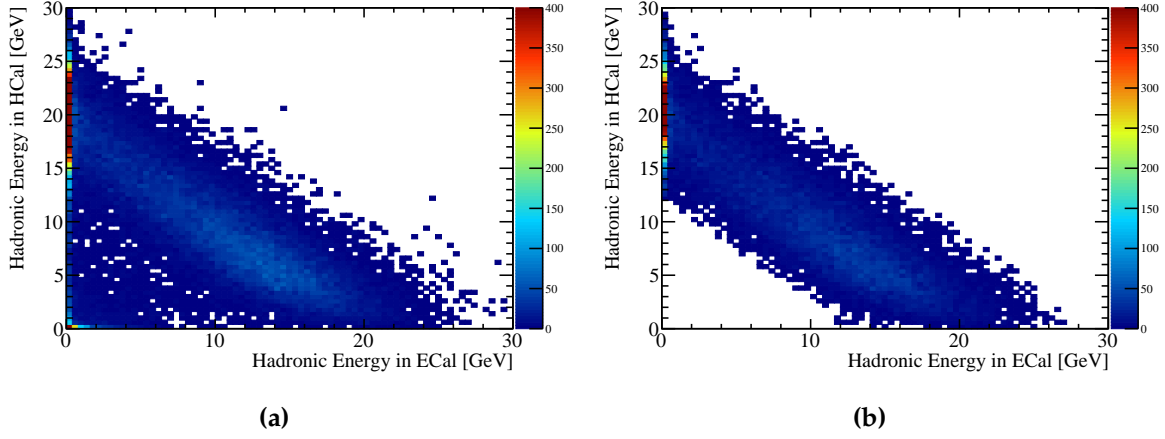
**Figure 1.9:** The distribution of hadronic energy measured in the ECal and HCal for 20 GeV $K_L^0$ events (a) without selection cuts and (b) with selection cuts.

The 20 GeV $K_L^0$ events are then simulated and reconstructed. Following that a linear fit to the distribution of $E_{ECal}^{Had}$ against $E_{HCal}^{Had}$ for 20 GeV $K_L^0$ events passing the selection cuts is applied. The fit is performed by minimising $\chi^2$, which is defined as:

$$\chi^2(\delta_{ECal}^{Had}, \delta_{HCal}^{Had}) = \sum_i \frac{x_i}{\sigma_{x_i}} \, , \qquad (1.8)$$

where $x_i$ is the perpendicular distance from $E_{ECal}^{Had}$ and $E_{HCal}^{Had}$ for event $i$ to the line $E_{HCal}^{Had} = \delta_{HCal}^{Had} - E_{ECal}^{Had} \frac{\delta_{HCal}^{Had}}{\delta_{ECal}^{Had}}$. The definition of $x_i$ is given in equation 1.9, but best illustrated by considering figure 1.10. $\sigma_{x_i}$ is the uncertainty on $x_i$, which is calculated by propagating the uncertainties on $E_{ECal}^{Had}$ and $E_{HCal}^{Had}$, which are assumed to be $\sigma_{E_{E/HCal}^{Had}} = 55\% \times \sqrt{E_{E/HCal}^{Had}}$, into the expression for $x_i$. The result of this propagation of errors is given in equation 1.10. The sum runs over all events, $i$, passing the selection cuts.

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

$$\sigma_i = \frac{(\sigma_{E_{HCal}^{Had}} \delta_{ECal}^{Had})^2 + (\sigma_{E_{ECal}^{Had}} \delta_{HCal}^{Had})^2}{\sqrt{(\delta_{ECal}^{Had})^2 + (\delta_{HCal}^{Had})^2}} \, , \tag{1.10}$$
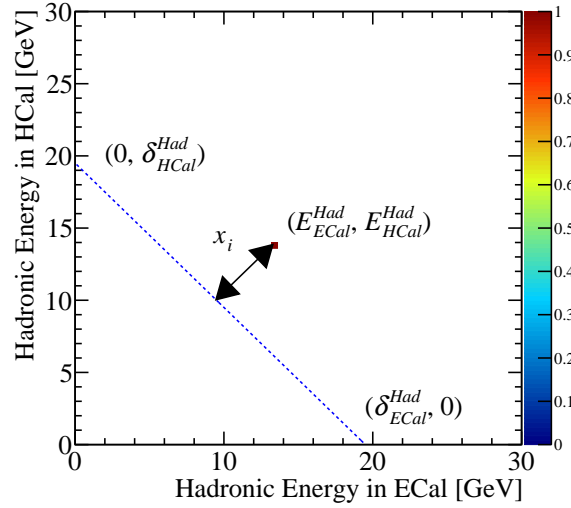


**Figure 1.10:** An example showing the definition of $x_i$, the variable used for the calculation of $\chi^2$ for the setting of the hadronic energy scale. For an event that has been measured with hadronic energy $E_{ECal}^{Had}$ in the ECal and $E_{HCal}^{Had}$ in the HCal, the geometric interpretation of $x_i$ is shown. The blue dotted line is defined as $E_{HCal}^{Had} = \delta_{HCal}^{Had} - E_{ECal}^{Had} \frac{\delta_{HCal}^{Had}}{\delta_{ECal}^{Had}}$.

The minimisation of $\chi^2$ is done by stepping over a range of $\delta_{ECal}^{Had}$ and $\delta_{HCal}^{Had}$ centred about the ideal value of $20\,\text{GeV} - m_{K_L^0}$ in search for the minimum $\chi^2$. Once the minima in $\chi^2$ is found the trial calibration factors $\beta_{ECal}^{Had0}$ and $\beta_{ECal}^{Had0}$ are rescaled to correct for any deviation from the desired fit as follows:

$$\beta_{ECal}^{Had0} \to \beta_{ECal}^{Had} = \beta_{ECal}^{Had0} \times \frac{E_K}{\Delta_{ECal}^{Had}} \tag{1.11}$$

$$\beta_{HCal}^{Had0} \to \beta_{HCal}^{Had} = \beta_{HCal}^{Had0} \times \frac{E_K}{\Delta_{HCal}^{Had}} \, , \tag{1.12}$$

where $\Delta_{ECal}^{Had}$ and $\Delta_{ECal}^{Had}$ are the values of $\delta_{ECal}^{Had}$ and $\delta_{ECal}^{Had}$ giving the minimum $\chi^2$. The step sizes used for minimising $\chi^2$ with respect to $\delta_{ECal}^{Had}$ and $\delta_{ECal}^{Had}$ is chosen such that a single step corresponds to the target final tolerance on $\delta^{Had}$ i.e. $|\delta_{E/HCal}^{Had} - E_{MC}| <$

$E_{\text{MC}} \times 0.5\% \approx 0.1\text{GeV}$. This procedure is then repeated until $\Delta_{ECal}^{Had}$ and $\Delta_{ECal}^{Had}$ both fall within a given tolerance, which in this case it taken to be $|\Delta_{E/HCal}^{Had} - E_{\text{MC}}| < E_{\text{MC}} \times 0.5\% \approx 0.1\text{GeV}$.

### 1.2.4 Retraining photon likelihood data

The likelihood data used by PandoraPFA in the identification of electromagnetic showers is trained using off-shell mass Z boson (Z′) events at 500 GeV decaying into light quarks (u, d, s). It is necessary to retrain this data only when varying the ECal as electromagnetic showers are largely contained within the ECal at the energies being considered and the likelihood data only uses measurements made in the ECal.

As this data uses post digitisation hits it is important to ensure that a fully calibrated detector is used when retraining the likelihood data. Therefore, it is necessary to run the calibration procedure, as described in section **??**, before retaining. However, as the reconstruction uses likelihood data the calibration procedure must be performed twice. Initially the calibration procedure is performed where PandoraPFA is run without the inclusion of this likelihood data, using PandoraSettingsMuon.xml. Then the likelihood data is retrained using the results of the first calibration pass and then the retrained likelihood data is used in the second pass of the calibration procedure.

## 1.3 Novel Energy Estimators

This section focuses upon novel energy estimators, introduced in section 1.1. Both techniques modify the cluster energy estimator by introducing weights, $\omega^i$, to energy deposits in the HCal that vary as a function of energy:

$$E_{Cluster} = \sum_{\text{ECal hits, } i} E_{ECal}^i + \sum_{\text{HCal hits, } i} E_{HCal}^i \omega^i, \tag{1.13}$$

As hadronic showers are largely contained in the HCal, weights are only applied to HCal hits. The weights $\omega^i$ vary a function of the hit energy density $\rho^i = \frac{E_{HCal}^i}{V}$, where $V$ is the HCal hit volume, and, while the exact weight values depend on the implementation of the technique, a general feature present in both is that at large $E_{HCal}^i$

the weight is less than one. This reduces the energy of spuriously high energy cells to limit the impact of Landau fluctuations. The metrics used for quantifying detector performance when using these novel energy estimators are thoroughly defined in chapter 2.

### 1.3.1 HCal Hit Energy Truncation

**Application**

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

$$\omega(\rho) = 1$$
$$= \frac{\Lambda}{V \times \rho} \tag{1.14}$$

where $\Lambda$ is the value of the truncation and $V$ is the volume of a HCal hit. This weight as a function of hit energy density is shown in figure 1.11.

**Results: Energy Resolution**

The application of these weights can lead to improvements in the energy resolution for neutral hadrons, which can be seen in figure 1.12a. However, a great deal of care has to be given to the truncation energy so that hits from typical hadronic shower development are not truncated, while the spuriously high energy hits are. Figure 1.12a indicates that a 1 GeV truncation is sufficient for dealing with the Landau fluctuations. For hit energy truncations greater than this the energy resolution worsens as the effect Landau fluctuations are not accounted for, while for hit energy truncations below 1 GeV the truncation is too aggressive and hits from typical hadronic shower development are truncated. For completion 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 invariant to the HCal hit energy truncation as the $\gamma$s are largely contained within the ECal.
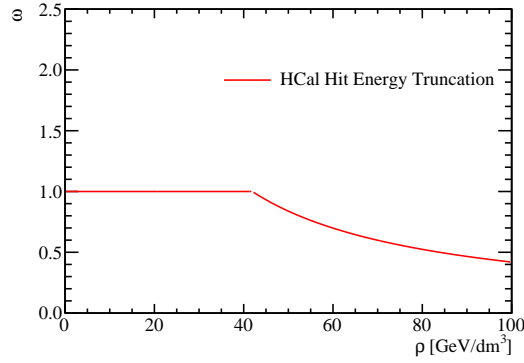


**Figure 1.11:** The weights, $\omega$, used in the HCal hit energy truncation as a function of the energy density of the HCal hit, $\rho$. The truncation shown here corresponds to a 1 GeV truncation in the nominal ILD HCal.



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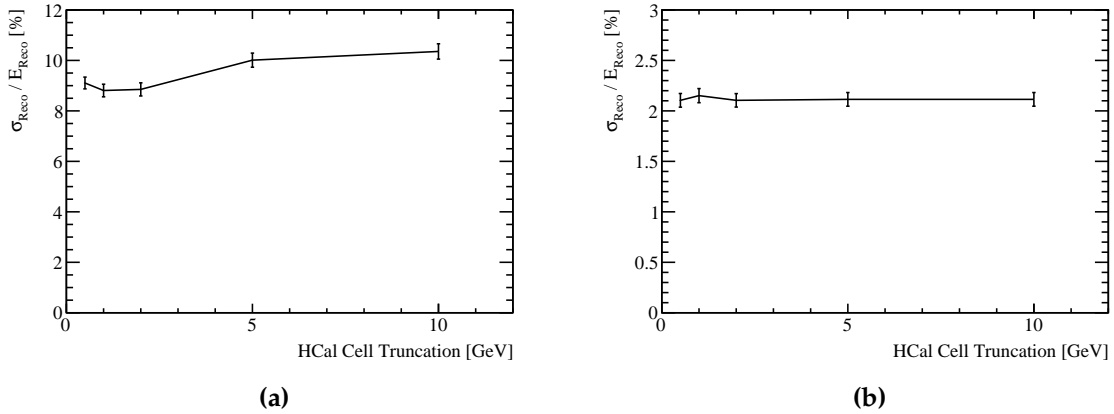
**Figure 1.12:** The energy resolution as a function of HCal cell truncation for (a) 50 GeV $K_L^0$ events and (b) 100 GeV $\gamma$ events using the nominal ILD detector model.

## Results: Jet Energy Resolution

The improved energy resolution for careful choice of truncation can be seen in the jet energy resolution. Figure 1.13 shows the jet energy resolution as a function of the jet energy for various values of the HCal hit energy truncation. The trends in this plot are complex as the optimal cell truncation varies with the jet energy. At low energies a 0.5 GeV truncation gives the best performance, however, when the jet energies reach

$\approx$ 180 GeV a 1-2 GeV truncation gives the best performance. This is to be expected based on the Landau fluctuations. The Landau distribution, governing the energy deposition for hadronic showers, is essentially a Gaussian with a high energy tail and as the jet energy increases, the mean of the Gaussian increases and the definition of hit energies falling in the high energy tail changes.
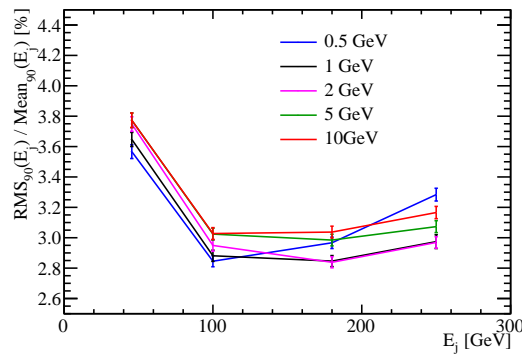


**Figure 1.13:** The jet energy resolution as a function of jet energy for various hadronic cell truncations. The results shown use the nominal ILD detector model.

While it is challenging to determine the optimal performance for a given detector model it is clear that applying an appropriate truncation produces significant improvement in detector performance. Therefore, for the optimisation studies presented in chapter 2 this form of novel energy estimator is applied. The optimal truncation for each detector model considered in that study was determined by performing the reconstruction using range of HCal hit energy truncations and quoting the optimal performance. The HCal hit energy truncations considered in the optimisation study were 0.5, 0.75, 1, 1.5, 2, 5, 10, and $10^6$ GeV (semi-infinite). For the HCal cell size study the truncation used for the 10, 20, 30, 40, 50 and 100 mm cell size detector was 0.5, 0.75, 1, 1.5, 2 and 5 GeV respectively, for the tungsten HCal options the truncation used was 5 GeV and for all other options the truncation used was 1 GeV. This optimisation has a significant impact on detector optimisation, which can be seen by comparing the jet energy resolutions obtained when using the optimised cell truncation and a fixed 1 GeV truncation, shown in figure 1.14. Without this optimisation of hit energy truncation the significance of the HCal cell size is overinflated and could lead to a misinformed detector design choice.
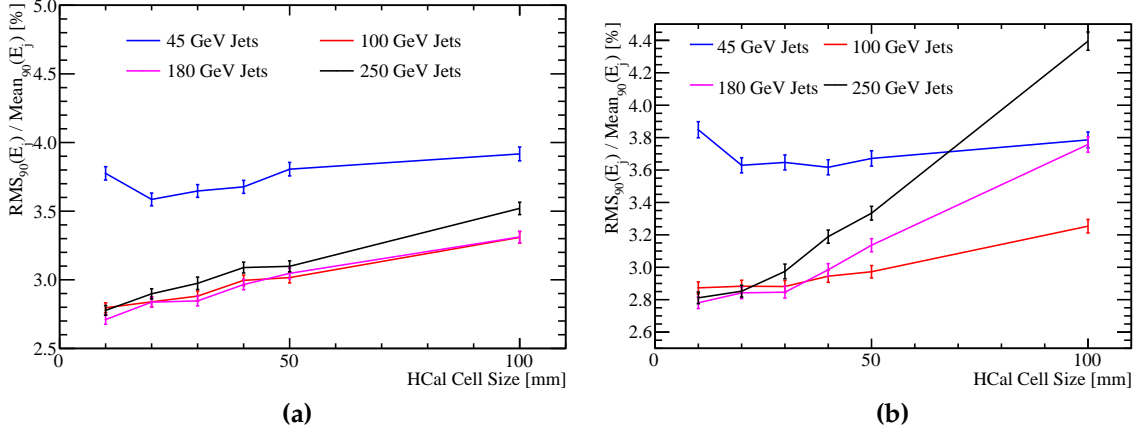
**Figure 1.14:** The jet energy resolution as a function of HCal cell size using a HCal hit energy truncation that is (a) optimised and (b) fixed at 1 GeV.

## 1.3.2  Software Compensation

### Application

A particle shower produced when a hadron passing through a calorimeter has two components [22]; 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 all other interacting and decaying particles in in the shower. By identifying each of these components in the reconstruction, it is possible to increase the energy of the hadronic hits to give a compensating response and decrease the energy of spuriously high energy hits that arise from Landau fluctuations. The challenge of applying this approach is to identify whether a hit is likely to be hadronic or electromagnetic in nature. This is done based on the energy density of the hit, with high energy densities likely to be part of the electromagnetic core and the low energy densities likely to be part of the satellite hadronic hits of the shower. An event display showing the energy density of a hadronic shower, where the electromagnetic core can be clearly seen, of a hadronic cluster in a 500 GeV Z→uds di-jet event can be found in figure 1.15.

An additional layer of sophistication in this approach is the weights that are applied vary as a function of the uncompensated cluster energy, $E_{\text{Raw}}$, as well as the hit energy density, $\rho^i$. This is to account for the changing energy densities found in hadronic showers as the total shower energy varies. For example, the fraction of hits in a hadronic shower that are electromagnetic in nature increases as the total energy of the

**(a)** **(b)**
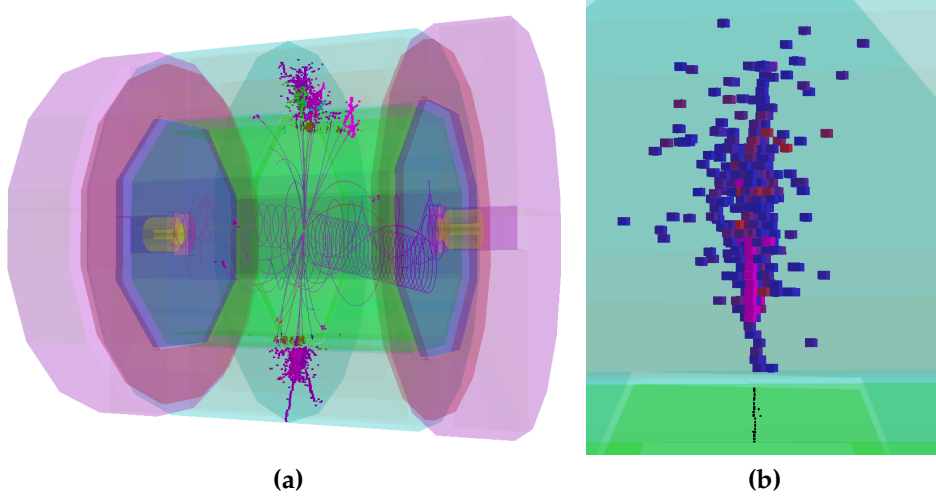
**Figure 1.15:** An event display for a 500 GeV Z→uds di-jet event reconstructed using the nominal ILD detector. (a) The full event environment. (b) A single hadronic cluster from the same event where shading indicates the energy density in the HCal. High energy density cells are coloured red, while 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.

hadronic shower increase [22]. Therefore, for a high energy hadronic shower fewer hits are purely hadronic in nature and so fewer will require weights that are larger that one. The highly segmented calorimeters used at the linear collider give excellent resolution on individual particle showers. This enables precise mapping of the calorimeter hits to different shower components as described above, which allows the linear collider to employ this software compensation technique with greater effectiveness than was possible for previous collider experiments.

In terms of the $\omega$ weights introduced in equation 1.13, the weights [6] in this technique corresponds to:

$$\omega(E_{\text{Raw}}, \rho) = p_1(E_{\text{Raw}}) \times exp(p_2(E_{\text{Raw}}) \times \rho) + p_3(E_{\text{Raw}}) \, , \tag{1.15}$$

$$p_1 = p_{11} + p_{12} \times E_{\text{Raw}} + p_{13} \times E_{\text{Raw}}^2 \, , \tag{1.16}$$

$$p_2 = p_{21} + p_{22} \times E_{\text{Raw}} + p_{23} \times E_{\text{Raw}}^2 \, , \tag{1.17}$$

$$p_3 = \frac{p_{31}}{p_{32} + exp(p_{33} \times E_{\text{Raw}})} \, , \tag{1.18}$$

where $p_{ij}$ are trained parameters. The parameters $p_{ij}$ are determined by performing a $\chi^2$ fit of $E_{Cluster}$ using software compensation to the MC energy for samples of $K_L^0$ ranging from 10 to 100 GeV in steps of 10 GeV. Using the fitted parameters, $p_1$, $p_2$ and $p_3$ as a function of $E_{Raw}$ and $\omega(E_{Raw}, \rho)$ as a function of $\rho$ for various $E_{Raw}$ are shown in figure 1.16 and 1.17 respectively. 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 invisible energy component found in hadronic showers. Furthermore, the weights vary as a function of the hadronic shower raw energy to account for the change in fraction of the shower energy appearing in the electromagnetic shower core.



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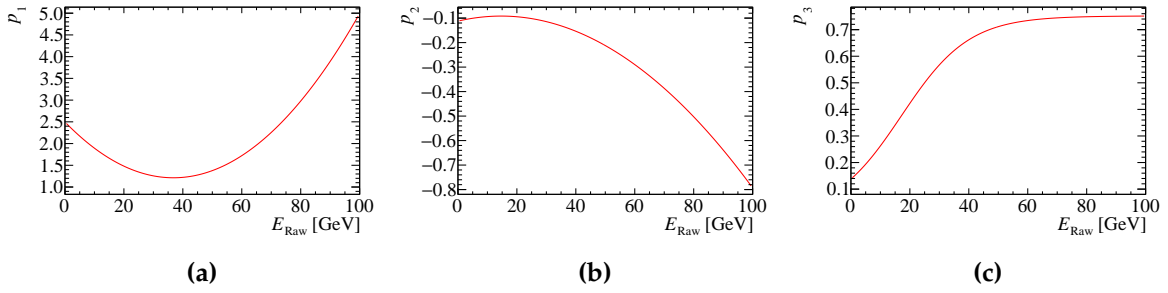
**Figure 1.16:** Parameters used in software compensation weight determination as a function of $E_{Raw}$.
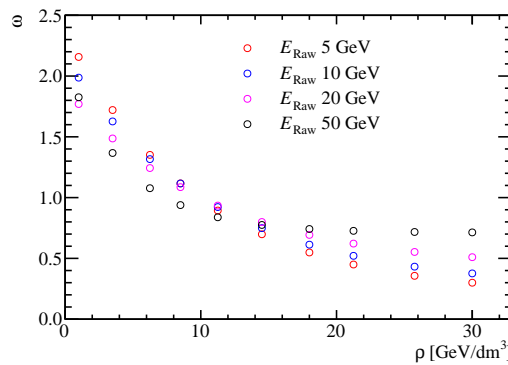


**Figure 1.17:** The software compensation weight applied to a calorimeter hit as a function of calorimeter hit energy density for various cluster energies.

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, from incorrect association of charged particle tracks to calorimeter hit clusters, is reduced.

As software compensation only modifies the energy of HCal hits there is freedom to apply further energy corrections to the ECal hits. Application of the Clean Clusters energy correction logic, described in section 1.3.2, to the ECal hits alongside software compensation gave further improvements 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 is trained using a maximum $K_L^0$ energy of 100 GeV, there-fore, it is only applied to clusters where $E_{\text{Raw}} < 100$ 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 for the use case considered here. This is the case as the ILD detector was designed for use at the ILC, which has a maximum running energy of 500 GeV.

**Context: Legacy Energy Corrections**

Before examining the impact of software compensation on detector performance is it necessary to address the 'legacy' energy corrections that are used as default in PandoraPFA. The three energy correction that were in use prior to the development of software compensation are:

- **HCal cell truncation**, the details of which can be found in section 1.3.1.

- **Clean Clusters**. This algorithm checks to see whether the energy measured within a calorimeter hit is anomalously high. Anomalously high energy cells are defined as cells where the energy contained within the cell is greater than 10% of the energy of the cluster that the cell has been associated to. If a cell is deemed to have an anomalously high energy and if this energy is above a threshold (0.5 GeV) the cell energy used by PandoraPFA is modified. The updated cell energy is taken as the average cell energy in the calorimeter layers immediately before and after the layer containing the high energy cell.

- **Scale Hot Hadrons**. This algorithm calculates the average number of MIP equivalent particles passing through each calorimeter cell in a cluster. If this number is larger than a given value, default 15 MIPs per cell, the cluster energy is rescaled to give a lower average number of MIPs per hit, default is 5 MIPs per hit.

Each of these energy corrections help to deal with the effects of spuriously high energy cells the origin of which is described in section 1.3.1. However, the algorithms are simplistic and software compensation is expected to give far better results than these 'legacy' options. The optimisation studies presented in section 2 use all three of these legacy options simultaneously, which was the default behaviour for PandoraPFA when the studies were undertaken. The new default behaviour in PandoraPFA is to use software compensation.

### Results: Energy Resolution

The energy resolution as a function of the MC energy for single $K_L^0$ events is shown in figure 1.18 using various energy correction settings. When comparing the energy resolution given by software compensation to that obtained using no energy corrections, it can be seen that software compensation offers a gain of $\approx 2\%$ in energy resolution across the energy range considered. The uniformity of this improvement is encouraging, indicating software compensation is achieving a compensating calorimeter response across this wide range of energies.

Comparing the performance of software compensation to the legacy corrections 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 $E_{K_L^0} \approx 50$ GeV where the performance of software compensation and the legacy corrections are comparable. By removing the cell truncation from the legacy options it is clear that the changes in energy resolution when using the legacy options are being driven by the cell truncation. This makes the trend in energy resolution observed using the legacy corrections clear as at low $K_L^0$ energies very few cells 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 cells that are not spuriously high leading to a worsening energy resolution. Between these two extremes, $E_{K_L^0} \approx 50$ GeV, the truncation works ideally and improvement in energy resolution using the legacy corrections is the largest.
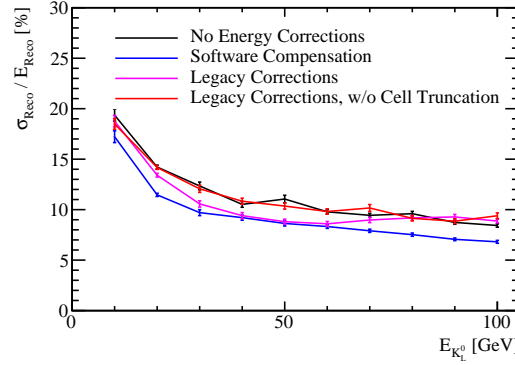
**Figure 1.18:** The energy resolution as a function of the MC energy for single $K_L^0$ events using various energy correction settings. The detector model used was the nominal ILD detector model.

**Results: 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. These effects are illustrated by examining the jet energy resolution as a function of jet energy, which is shown in figure 1.19. Again it is clear that software compensation is extremely beneficial to the detector performance.
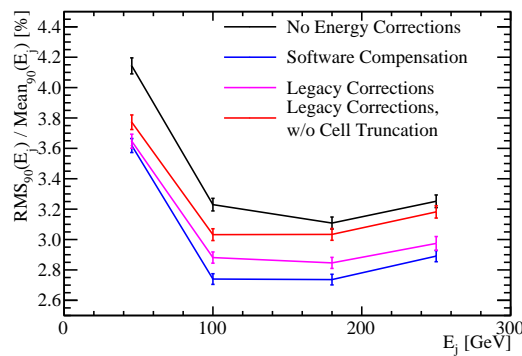


**Figure 1.19:** The jet energy resolution as a function of the jet energy for a variety of different energy correction options. These results were produced for the nominal ILD detector model.

Software compensation gives a significant reduction in the jet energy resolution in comparison to using no energy corrections. It also reduces the jet energy resolution in comparison to using the legacy corrections.

Further light can be shed on these trends by examining the contribution to the jet energy resolutions from the intrinsic energy resolution 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. Unlike the single particle study there is no jet energy for which the cell truncation matches the performance obtained using software compensation. This is due to the fact that the energy resolution when using the cell truncation is only comparable to the energy resolution using software compensation for a narrow range of hadronic cluster energies. As the jet contains a broad spectrum of hadronic cluster energies the performance obtained when using the cell truncation will always be worse than when using software compensation. When comparing the jet energy resolution for the legacy corrections is again apparent that the term driving the jet energy resolution is the cell truncation.

The confusion contribution 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, shown in figure 1.19, when using software compensation as opposed to the legacy corrections is being driven by improvements to the intrinsic energy resolution. At low jet energies the Clean Clusters and Scale Hot Hadrons energy corrections are beneficial at reducing the confusion contribution, while the cell truncation is largely redundant. For high jet energies jets this trend is reversed. As the use of the Clean Clusters and Scale Hot Hadrons energy corrections do not alter the intrinsic energy resolution of the detector it is apparently that these energy corrections are purposed to account for failures in the pattern recognition that occur largely at low jet energies. On the other hand the cell truncation and software compensation techniques aim to improve the energy resolution of the hadronic clusters, which has a knock-on effect of improving the track cluster associations made in the pattern recognition. These corrections work across all energy ranges, but have a greater impact at high energies. By extracting the Clean Clusters logic, which is the driving term reducing the confusion contribution to the jet energy resolution at low jet energies, and embedding it within the software compensation energy correction, it is possible to achieve exceptional jet energy resolutions that will extend the physics reach of the linear collider detector.

**(a)**



**(b)**
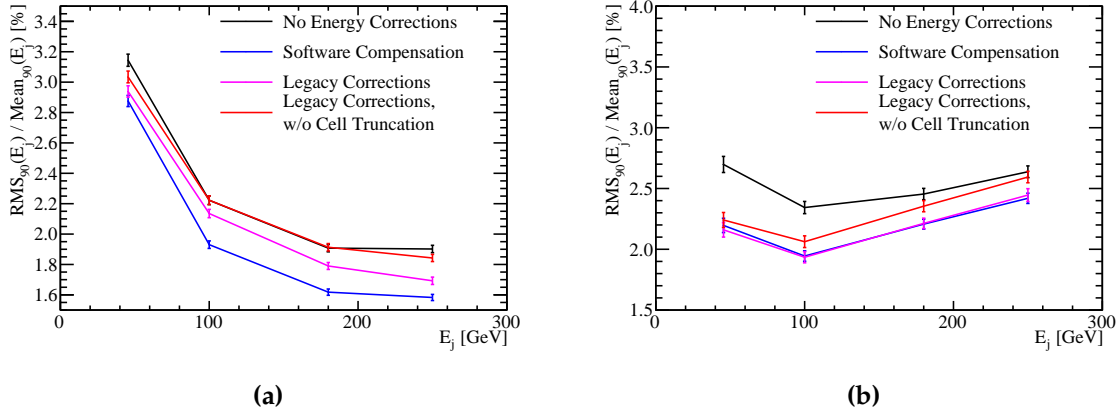
**Figure 1.20:** The contributions to the jet energy resolution as a function of the jet energy for a variety of different energy correction options. (a) is the intrinsic energy resolution of the detector and (b) is the total confusion term. The quadrature sum of both yields the standard reconstruction performance. These results were produced for the nominal ILD detector model.

## 1.4 Timing Cuts

The ILC and CLIC will operate using a trigger-less readout approach whereby the recorded data for each sub-detector is readout between collisions of $e^+$ and $e^-$ bunches. The train structure for the ILC and CLIC at maximum operating energy is shown in table 1.1. Event selection will proceed through the application of a software trigger. This involves the identification of 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 this interaction. Timing cuts placed on the calorimeter hits are 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. As the size of the time window around the hard interaction changes the amount of time particle showers have to develop varies and this will affect the performance of the detector.

For all choices of time window considered in this study the calibration procedure was reapplied. This means 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.

|  | 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 [5, 19].

## Results: 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 in figure 1.21 for the nominal ILD detector model . The timing cut makes little difference to the energy resolution of the $\gamma$ events, however, there is a significant decrease in the energy resolution for the neutral hadrons. This is to be expected as electromagnetic showers develop far more rapidly than their hadronic counterparts [22], which can be seen in figure 1.22. This is expected hadronic showers often involve intermediate states that must decay to continue the propagation of the shower and as these states have non-zero lifetimes they slow 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.



(a)



(b)
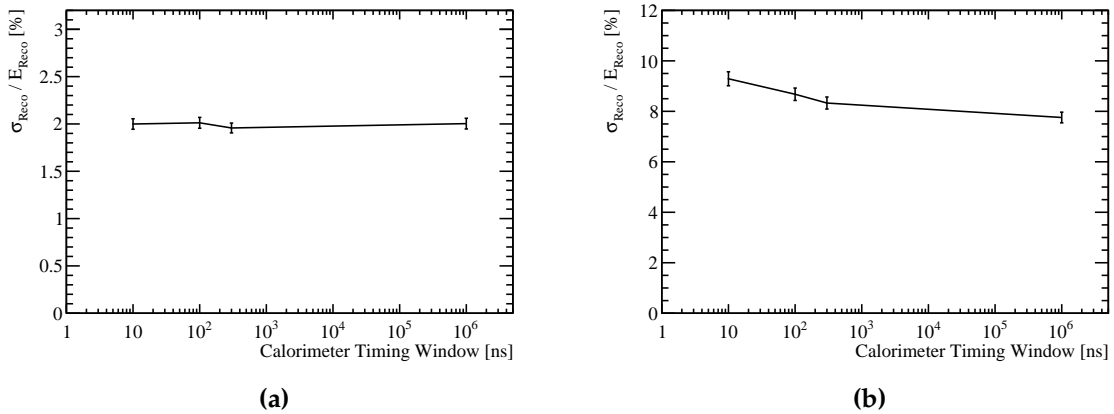
**Figure 1.21:** The energy resolution as a function of calorimeter timing window for (a) 100 GeV $\gamma$ events and (b) 50 GeV $K_L^0$ events using the nominal ILD detector model.
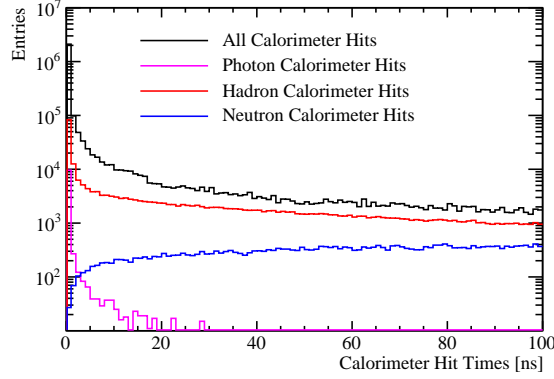
**Figure 1.22:** The distribution of the time of the calorimeter hits, corrected for time of flight to the impact point, for 91 GeV Z→uds di-jet events.

## Results: 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. As the magnitude of the changes to the jet energy resolution when varying the time window size are small in comparison to the absolute resolutions, this exception will most likely be due to a fluctuation in either the event sample used or in the reapplication of the calibration procedure.
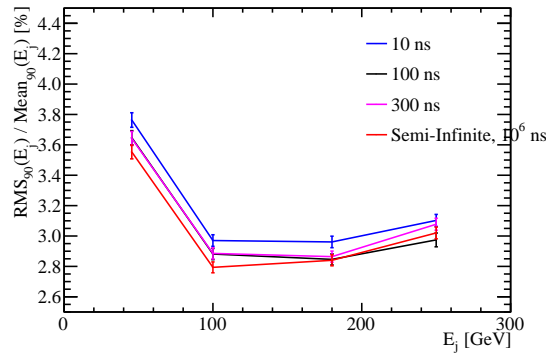


**Figure 1.23:** The jet energy resolution as a function of jet energy for various calorimeter timing cuts. The nominal ILD detector model was used for this study.

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.

For results presented in this chapter and the optimisation studies found in chapter 2 a 100 ns timing window was applied across all 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 [7]. Furthermore, it adds additional realism to the detector simulation in comparison to omitting the effect of the calorimeter time window. The categorisation of changes to the detector performance when varying the calorimeter timing window presented here can be used to discern the impact of changing the timing window used for the optimisation studies at a later date if so desired.

# Chapter 2

# Calorimeter Optimisation Studies

*"The simple believes everything, but the prudent gives thought to his steps."*
— Proverbs 14:15

## 2.1 Introduction

The fundamental principle of particle flow calorimetry is to measure the energy of a particle passing through a detector in whichever sub-detector offers the best energy resolution. For particle collider experiments, this involves measuring the momenta of charged particles using the curvature of the track they create in the detector. This offers extremely good energy resolution in comparison to calorimetric energy measurements. As neutral particles produce no tracks, their energies must be measured using calorimetric energy deposits.

The application of particle flow calorimetry is extremely challenging as charged particles create both tracks and calorimetric energy deposits. If both of these energy measurements were included, the energy of all charged particles would be double counted. Therefore, to avoid this, any calorimetric energy deposits originating from charged particles are not included in the final energy measurement. However, this methodology makes it possible to double count and omit energy measurements if the origin of a calorimetric energy deposit is misidentified. For example:

- If a calorimetric energy deposit, made by a charged particle, is not associated to a track, the calorimetric energy deposit will be double counted: Firstly, when the

track energy is accounted for and secondly, when the calorimetric energy deposit is incorrectly reported as the energy of a neutral particle.

- If a calorimetric energy deposit, made by a neutral particle, is incorrectly associated to a track, that calorimetric energy deposit is not accounted for.

These effects, collectively know as "confusion", degrade the energy resolution of a particle flow detector. Therefore, it it is crucial to make correct associations between charged particle tracks and their calorimetric energy deposits to minimise the effect of confusion. These associations can only be successfully made if the calorimeters used have fine segmentation, such as those found at the linear collider experiment, so that it becomes possible to separate the energy deposits from nearby showering particles. Even with this segmentation, making the association of charged particle tracks to calorimetric energy deposits is highly non-trivial. At the linear collider experiment, these associations are made using sophisticated pattern recognition algorithms, provided by PandoraPFA. The fine segmentation of the linear collider calorimeters allows PandoraPFA to reconstruct the four-momenta of all particles passing through the detector and to report the energy of all reconstructed particles using the energy measurement from the optimal sub-detector. For additional details on PandoraPFA see chapter **??**.

This chapter considered optimisation of the calorimeters used at the linear collider, with focus placed on obtaining the best energy resolution for jets. Parameters such as the number of layers, cell size and material choices for the calorimeters are investigated. This chapter concludes with an optimisation of several global detector parameters such as the magnetic field strength and the inner radius of the ECal. These parameters are not calorimeter specific, but affect the jet energy resolution obtained from particle flow.

## 2.2 Jet Energy Resolution

As many physics processes of interest at the linear collider involve multi-jet final states [4], good jet energy resolution is a crucial a aspect of detector performance. As shown in chapter **??**, the sensitivity of the linear collider experiment to areas of new physics can determined using reconstructed jet energies. Furthermore, parameters derived from the energy measurements of jets are extremely useful for identification of

physics channels of interest. Therefore, the primary metric used for this optimisation study is the jet energy resolution. Jet energy resolution in particular can benefit from the application of particle flow calorimetry as $\approx 70\%$ of the energy of jets are carried in the form of charged particles. As particle flow aims to measure the energy of charged particles using the tracker, it has the potential to offer extremely large benefits when measuring jet energies in comparison to the traditional calorimetric approach.

### 2.2.1 Jet Energy Resolution Metrics

The jet energy resolution in these studies was determined through the simulation of off-shell mass Z boson events decaying to light quarks (u, d, s). In these events, the Z boson is produced at rest, which means the typical decays form two mono-energetic jets that are produced back to back as shown in figure 2.1. Only events where $|\cos(\theta)| < 0.7$, where $\theta$ is the polar angle of the quarks, are used in the jet energy resolution calculation. This ensures little energy is lost down the beam axis. Using these events the jet energy resolution is calculated as follows:

$$\frac{\text{RMS}_{90}(E_j)}{\text{Mean}_{90}(E_j)} = \frac{\text{RMS}_{90}(E_{jj})}{\text{Mean}_{90}(E_{jj})} \times \sqrt{2} \,, \tag{2.1}$$

where $E_{jj}$ is the total reconstructed energy. $\text{Mean}_{90}(E_{jj})$ and $\text{RMS}_{90}(E_{jj})$ are the mean and root mean squared (RMS) of the $E_{jj}$ distribution. These variables are calculated across the range of $E_{jj}$ with the smallest RMS containing at least 90% of the data. This definition is used to remove the effect of outliers in the distribution [21]. If all associations between charged particle tracks and calorimeter clusters were correctly made, the reconstructed jet energy distribution would be Gaussian. However, the effect of confusion on certain events will distort this distribution and broaden the tails significantly. If the full range were to be used in the jet energy resolution calculation, the effect of these tails is overinflated. If the distribution of reconstructed jet energies is truncated to the narrowest range of the data containing at least 90% of the data, the effect of these tails can be negated. This removes events where confusion is dominant, which makes the jet energy resolution metric far more robust and representative of the bulk of the data.

An example of the application of this metric can be found in figure 2.2. In this example $\text{RMS}(E_j)$, the RMS calculated using the full range, is 5.8 GeV, while $\text{RMS}_{90}(E_j)$, the RMS using the reduced range, is 4.1 GeV. This corresponds to a reduction in the jet
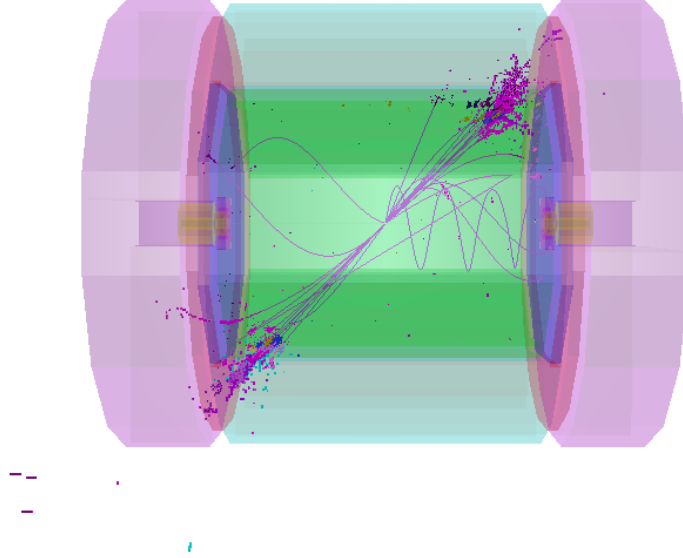
**Figure 2.1:** 500 GeV di-jet Z→uds event display for nominal ILD detector.
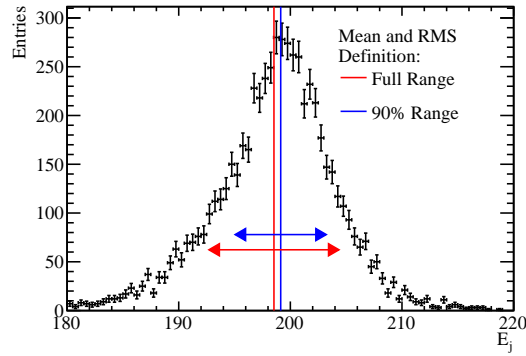


**Figure 2.2:** Definition of jet energy resolution. Reconstructed jet energy for 200 GeV di-jet Z→uds events for nominal ILD detector.

energy resolution from 4.1% to 2.9%, which clearly shows an overemphasis of the tails of the distribution when using the full jet energy range.

In the subsequent analysis a range of di-jet energies were considered ranging from the Z mass, 91 GeV, to the nominal running energy of the ILC, 500 GeV. Each event sample contained 10,000 events generated isotropically so that, given the polar angle cut, approximately 7,000 events contribute to the jet energy resolution calculation.

## 2.2.2 Jet Energy Resolution Decompositions

It is possible to gain additional insight into the detector performance by cheating the pattern recognition, the clustering of calorimeter hits together and the creation of track to cluster associations, using MC information. Cheating the pattern recognition removes the effect of confusion as it ensures no errors are made when either clustering calorimeter hits together or when associating charged particle tracks to those calorimeter clusters. This allows the detector performance to be deconstructed into two terms; one related exclusively to the intrinsic energy resolution of the detector and another related to the pattern recognition confusion. The additional information this provides is extremely useful for characterising changes to the overall detector performance.

The intrinsic energy resolution contribution to the jet energy resolution is determined by fully cheating the pattern recognition; in this case all confusion is negated. The total confusion is defined as the quadrature difference between the jet energy resolution using the standard reconstruction and this fully cheated reconstruction. Furthermore, it is possible to cheat the pattern recognition associated with individual types of particles. This is particularly useful for studies related to the ECal as, by cheating the photon pattern recognition, it is possible to isolate the confusion associated with photons. The photon confusion is defined as the quadrature difference between the jet energy resolution using the standard reconstruction and the reconstruction where the photon pattern recognition is cheated. Examples of the calculation of the various confusion terms defined above are given in table 2.1.

| Reconstruction | Jet Energy Resolution [%] |
|---|---:|
| Standard Reconstruction (No MC Information) | $a = 2.97 \pm 0.05$ |
| Cheating Entire Reconstruction | $b = 1.69 \pm 0.02$ |
| Confusion | $\sqrt{a^2 - b^2} = 2.45 \pm 0.05$ |
| Cheating Photon Reconstruction | $c = 2.73 \pm 0.04$ |
| Photon Confusion | $\sqrt{a^2 - c^2} = 1.18 \pm 0.06$ |

**Table 2.1:** Example calculation of the confusion contributions to the jet energy resolution. These jet energy resolutions are for 250 GeV jets using the nominal ILD detector model.

A common feature that is observed in these calibration studies is that as the intrinsic energy resolution of a calorimeter improves, the effect of confusion is reduced. This

occurs as a better energy resolution means more precise comparisons can be made between the energy of a cluster of calorimeter hits and the momentum of any charged particle tracks associated to it. Comparisons such as these are made by PandoraPFA to determine whether the track cluster associations that have been made are consistent. If a large discrepancy is observed between the cluster energy and track momenta, the clustering of calorimeter hits is modified until a consistent association can be made. For more details on this comparison see chapter **??**. This consistency check vastly reduces the number of errors made when clustering calorimeter hits and associating charged particle tracks to those clusters i.e. the confusion. Therefore, improving the precision of this consistency check, by improving the energy resolution, reduces the effect of confusion.

### 2.2.3 Single Particle Energy Resolution

The energy resolution for individual particles is crucial for a number of physics studies of interest to the linear collider, such as $\gamma$ energy resolutions in the study of anomalous triple and quartic gauge couplings [2, 11, 12]. Therefore, $\gamma$ and $K_L^0$ energy resolutions, alongside the jet energy resolution, will be considered in these optimisation studies. As both $\gamma$ and $K_L^0$ are uncharged, their energy measurements will be made using the calorimeters as opposed to the tracker. $\gamma$s are a natural choice of particle to consider as they are particularly relevant for several physics studies and, as they are largely contained within the ECal, they will be highly sensitive to changes in the ECal performance. $K_L^0$s were used as, analogously to $\gamma$s and the ECal, their energies are primarily measured using the HCal. In general, neutral hadron energy resolutions are less crucial to physics studies, however, they do make crucial contribution to the jet energy resolution that should not be overlooked. The reported $\gamma$ energy resolutions were determined using events containing a single 100 GeV $\gamma$, while the $K_L^0$ energy resolutions were determined using events containing a single 50 GeV $K_L^0$. These energies were chosen to be as large as possible, to maximise sampling of the calorimeter response, while minimising the affect of leakage of energy from the ECal to the HCal for the $\gamma$ events and leakage of energy out of the rear of the HCal for the $K_L^0$ events.

For these single particle samples, the energy resolution is defined using a Gaussian fit to the reconstructed energy distributions. To aid convergence of the fit, the fit was applied to the narrowest range of the reconstructed energy distribution containing at

least 75% of the data. The single particle energy resolution is defined as the standard deviation divided by the mean of the fitted Gaussian. For each energy resolution calculation, a total of 10,000 events were used to populate the reconstructed energy distribution. For clarity, a cut of $|\cos(\theta)| < 0.7$ was applied to veto events where particles travelled down the beam pipe or where they passed through the barrel/endcap overlap region. An example of the reconstructed energy distributions for 100 GeV $\gamma$s and 50 GeV $K_L^0$s, alongside the Gaussian fits used to determine the energy resolutions, are shown in figure 2.3. The errors quoted on single particle energy resolutions are determined by propagating the errors reported from the Gaussian fit into the resolution calculation.



(a)                                        (b)
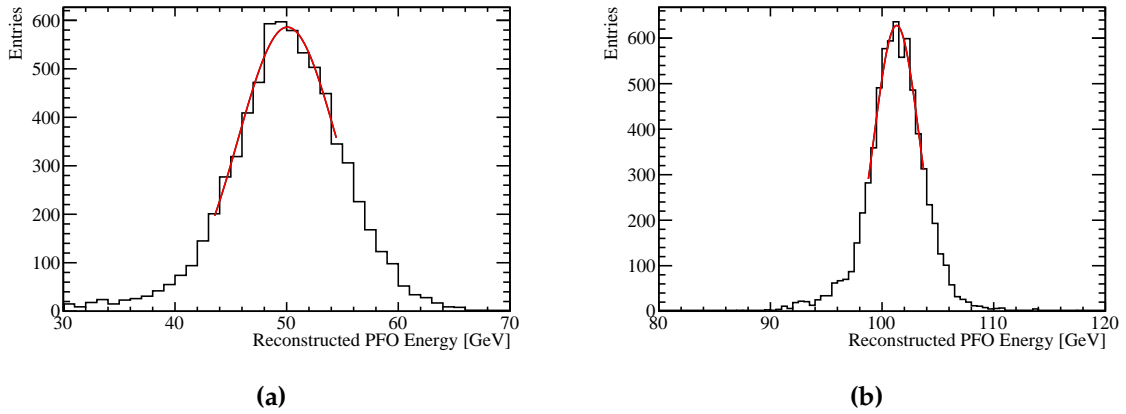
**Figure 2.3:** The reconstructed energy distribution for (a) 50 GeV $K_L^0$ and (b) 100 GeV $\gamma$ events. The red line shows a Gaussian fit used to parameterise the detector performance. The fit was applied to the truncated range of the reconstructed PFO energy distribution containing at least 75% of the data with the narrowest RMS. The nominal ILD model was used in this simulation.

## 2.3  Nominal Detector Performance

Before addressing optimisation of the calorimeters for use at the linear collider experiment, it is necessary to properly quantify the behaviour of the nominal linear collider detector model. The nominal detector model used in these simulation studies will be the ILD detector model. For more details on this detector model see chapter **??**.

The reconstructed energy distributions for particles whose energies are measured using calorimeters will be Gaussian. This is the case for sampling calorimeters as

the active material in each calorimeter hit essentially counts the number of charged particle tracks passing through it, or possible the number of photons for scintillator options. An estimation of the total energy deposited in a calorimeter hit, including the absorber material, can be made based upon this number of tracks or photons. For more details on how this estimation is made see chapter 1. Finally, the energy of the entire particle shower is estimated by grouping together calorimeter hits and summing their energy. As each calorimeter hit energy is an independent random measurement the particle shower energy will, by the central limit theorem, have a Gaussian distribution.

As each calorimeter hit involves counting a number of objects, charged particle tracks or photons, Poisson statistics governing the distribution of calorimeter hit energies. If the mean of the distribution of the energy of a cluster of calorimeter hits is $\lambda = N$, where $N$ is the mean number of objects the are measured in the calorimeters, the standard deviation of that distribution is $\sigma = \sqrt{\lambda} = \sqrt{N}$ and the energy resolution is $\frac{\sigma}{\lambda} = \frac{1}{\sqrt{N}}$. As the total shower energy, $E_{Reco}$, is proportional to $N$ the energy resolution for a particle shower in an ideal calorimeter is proportional to $\frac{1}{\sqrt{N}} = \frac{1}{\sqrt{E_{Reco}}}$. Therefore, the energy resolution as a function of energy for an ideal calorimeter is $\frac{\sigma_{Reco}}{E_{Reco}} = \frac{a}{\sqrt{E_{Reco}}}$. In reality, it is typical to express the energy resolution of a calorimeter in the following form:

$$\frac{\sigma_{Reco}}{E_{Reco}} = \frac{a}{\sqrt{E_{Reco}}} \oplus b \oplus \frac{c}{E_{Reco}} \, , \qquad (2.2)$$

where the $b$ term is a constant term that accounts for a variety of instrumental effects that do not depend on energy, e.g. mechanical imperfections, and the $c$ term accounts for electrical noise [15]. $\oplus$ denotes the quadrature sum.

Prototypes of the various ILD calorimeter options have been constructed and validated using test beam measurements. For the ECal, the energy resolution was parameterised using test beam measurements as $\frac{16.6}{\sqrt{E_{Reco}}} \oplus 1.1\%$ for the silicon option and $\frac{12.9}{\sqrt{E_{Reco}}} \oplus 1.2\%$ for the scintillator option [5]. The electrical noise was deemed sufficiently small that the $c$ term in the parameterisation could be neglected in both cases. These results were determined using an e$^-$ test beam with energies ranging up to $\approx 40$ GeV. This parametrisation is compared to the full ILD detector simulation in figures 2.4a and 2.4b for the silicon and scintillator ECal options respectively. The test

beam parameterisation of the energy resolution for the silicon ECal option is almost identical to the energy resolution observed in the full simulation. At very high energies, $\approx 500$ GeV, the ECal is no longer sufficient to fully contain the $\gamma$s and so leakage of energy into the HCal leads to a minor degradation in the energy resolution. This accounts for the worse energy resolution seen in the full simulation when compared to an extrapolation of the test beam parameterisation at high energies. The test beam parameterisation of the energy resolution for the scintillator ECal option is significantly better than that observed in the full simulation. This difference is most likely due to an imperfect implementation of the scintillator ECal within the full detector simulation. The $\gamma$ energy resolutions seen in the full ILD simulation are similar for the silicon and scintillator ECal options.

Similarly, the energy resolution using test beam was parameterised as $\frac{57.6}{\sqrt{E_{Reco}}} \oplus 1.6\%$ for the nominal ILD HCal [6]. A comparison between this test beam parameterisation and the full ILD simulation, using the silicon ECal option, is shown in figure 2.4c. The test beam measurements were made using $\pi^{\pm}$s with energies ranging from 10 to 80 GeV, while the full ILD simulation used $K_L^0$s ranging from 10 to 100 GeV. For determining the test beam parameterisation, only showers starting in the HCal were considered to minimise the effect of errors associated with the ECal calibration, while all showers were considered in the full simulation. The deviation between the test beam parameterisation and the full simulation, which grows at as the $K_L^0$ energy increases, is most likely due to the treatment of energy deposits leaking out of the back of the HCal. A tail catcher was used in the test beam analysis that had a similar structure to the HCal, but with a much wider average absorber thickness. Energy deposits in this tail catcher were calibrated in a similar fashion to the HCal giving a very good energy resolution for these energy deposits. In the full simulation a muon chamber acts as the tail catcher, however, the calibration applied to it is less sophisticated than that applied to the test beam tail catcher data, which gives a worse energy resolution for those hits. Furthermore, energy deposits in the uninstrumented solenoid region in the full simulation could not be accounted for.

The jet energy resolutions as a function of jet energy using the full ILD simulation are shown in figure 2.4d. Alongside this, the intrinsic energy resolution and confusion contributions to the jet energy resolution are also presented. The jet energy resolution at low energies is dominated by the intrinsic energy resolution of the detector, while at high energies it is dominated by the effect of confusion. This is to be expected as the intrinsic energy resolution of the calorimeters is proportional $\frac{1}{\sqrt{E_j}}$, where $E_j$ is the jet

energy, which means that it will dominate the energy resolution at low energies. On the other hand, confusion, which is a measure of how easy it is to group calorimeter hits together into clusters and associate those clusters to charged particle tracks, will grow with increasing jet energy as the event topology becomes more dense. The total jet energy resolution for the ILD detector are sufficiently small, $\frac{\sigma_{E_j}}{E_j} \lesssim 3.8\%$ [5, 19, 21], across the energy range considered to make separation of the hadronic decays of the W and Z bosons possible, which is one of the key requirements for the future linear collider

## 2.4 Electromagnetic Calorimeter Optimisation

The ECal primarily measures the energy deposits from electromagnetic showers. The nominal ILD detector model ECal, summarised in table 2.2, is a silicon-tungsten sampling calorimeter. It contains 29 readout layers and 24 radiation lengths ($X_0$), which is sufficient to contain all but the highest energy electromagnetic showers. The absorber thickness of the last nine layers is twice that of the first 20 layers to reduce the number of readout channels and cost of the calorimeter. This high sampling rate is crucial for the pattern recognition aspect of particle flow calorimetry, especially in the region where particle showers start developing, as shown in section 2.4.1.

| Parameter | Default Value |
|---|---|
| Cell Size | $5 \times 5\text{mm}^2$ square cells |
| Number of Layers | 29 readout layers |
| Active Material Choice | Silicon or Scintillator |
| Active Material Thickness | 0.5 mm (Silicon) or 2 mm (Scintillator) |
| Absorber Material Choice | Tungsten |
| Absorber Material Thickness | 20 layers of 2.1 mm followed by 9 layers of 4.2 mm |

**Table 2.2:** The configuration of the ECal in the nominal ILD detector model. The parameters are given for the nominal silicon model as well as the alternative scintillator option.

The calorimeter performance was simulated for a number of detector models where the following detector parameters were varied:

- Cell size: This is a vital aspect of the detector in the particle flow paradigm as smaller cell sizes leads to better separation between nearby showering particles,
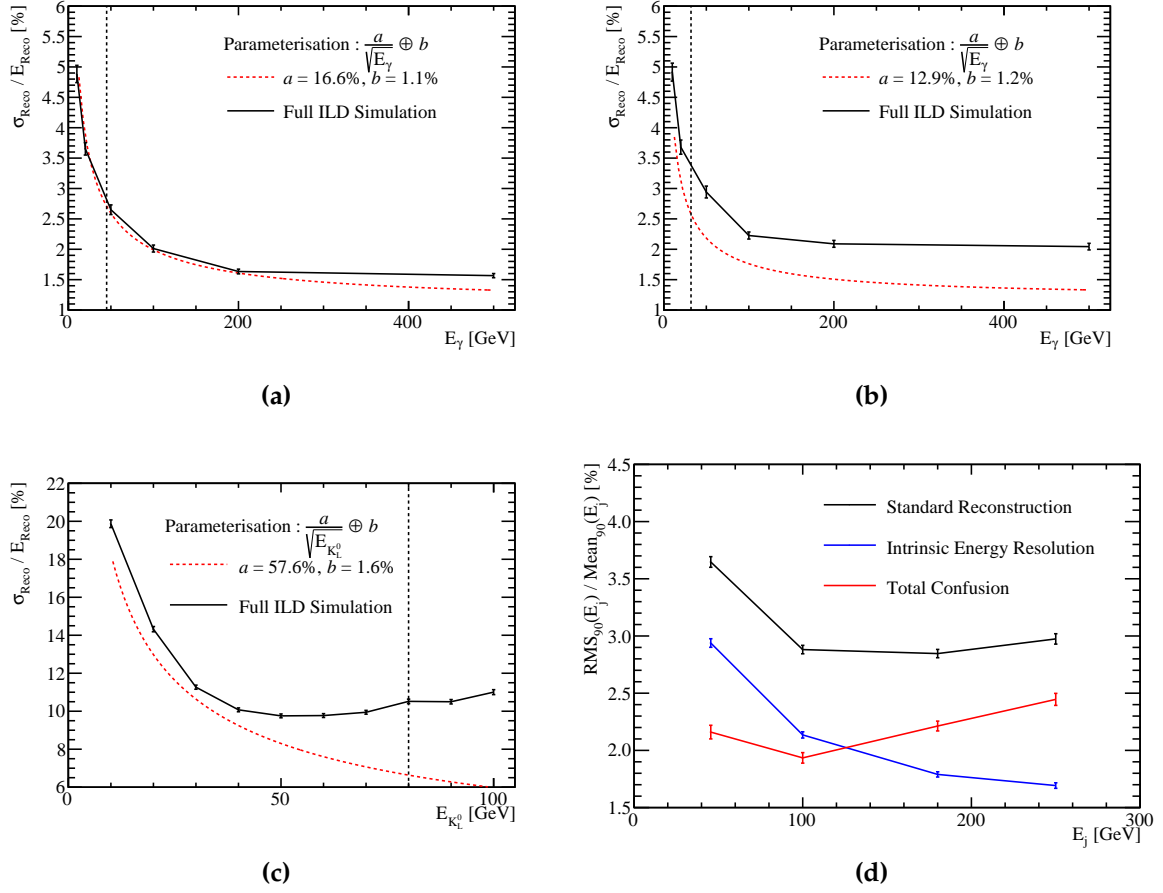
**Figure 2.4:** (a) The energy resolution as a function of $\gamma$ energy for the silicon ECal option. The black markers indicate the energy resolutions for the full ILD simulation and the red dotted line shows the test beam parameterisation of the ECal energy resolution. (b) The energy resolution as a function of $\gamma$ energy for the scintillator ECal option. The black markers indicate the energy resolutions for the full ILD simulation and the red dotted line shows the test beam parameterisation of the ECal energy resolution. (c) The energy resolution as a function of neutral hadron energy. The black markers indicate the energy resolutions for the full ILD simulation, with the silicon ECal option, which was determined using $K_L^0$s. The red dotted line shows the test beam parameterisation of the HCal energy resolution, which was determined using $\pi^{\pm}$s. (d) The jet energy resolution ($\mathrm{RMS}_{90}$) as a function of jet energy using the nominal ILD model, with the silicon ECal option. The intrinsic energy resolution and confusion contributions these the jet energy resolutions are also presented. The black dotted vertical line on the single particle energy resolutions shows the highest energy particles used in the test beam measurements.

which helps to minimise the effect of confusion. Modifying the cell size should have little effect on the intrinsic energy resolution of the detector.

- Sampling frequency: The sampling frequency in the ECal was varied by changing the number of layers in the ECal while simultaneously changing the thicknesses of the layers such that the total depth, in radiation lengths, was held constant. Increasing the number of layers in a sampling calorimeter means any particles showering within it are sampled more thoroughly. This leads to a reduction in the stochastic contribution to the energy resolution, as discussed in section 2.3. Therefore, varying the number of layers is expected to change in intrinsic energy resolution of the calorimeter.

- Active material choice: The options under consideration for the active material choice are silicon or scintillator. As well as providing different intrinsic energy resolutions the readout mechanics of these two options are significantly different. There is no clear prior knowledge as to which should provide better performance.

### 2.4.1 ECal Cell Size

A number of different detector models were considered where the cell size in the ECal was varied about the nominal value of $5 \times 5\text{mm}^2$ square cells. The granularities considered were $3 \times 3\text{mm}^2$, $5 \times 5\text{mm}^2$, $7 \times 7\text{mm}^2$, $10 \times 10\text{mm}^2$, $15 \times 15\text{mm}^2$ and $20 \times 20\text{mm}^2$ square cells for both the silicon and scintillator active material options.

The energy resolution, using 100 GeV $\gamma$ events, as a function of the ECal cell size is shown in figure 2.6a for the silicon option and figure 2.6b for the scintillator option. As, at these energies, the $\gamma$s will be largely contained within the ECal, these results solely reflect the performance of the ECal. For both the silicon and scintillator ECal options the energy resolution does not depend strongly on the ECal cell size. This is to be expected as the number of layers, which is the main factor in determining the energy resolution of a sampling calorimeter, does not change when modifying the cell size. There are minor fluctuations in the energy resolution when varying the cell size, but these are mostly likely due to fluctuations in the energy response from calibration procedure applied to each detector models. For the scintillator ECal option there is a significant degradation in the energy resolution for the $3 \times 3\text{mm}^2$ cell size model. The most likely cause of this is a "dead" region in the active material, which represents the readout multi pixel photon counter (MPPC) [3]. The MPPC occupies a fixed area of the cell, irrespective of cell size, and so the dead region of the cell fractionally increases as cell size is reduced. The larger this dead region, the worse the sampling

of the electromagnetic showers in the ECal and the worse the resolution. While this effect will be present in all scintillator ECal options, it will only be significant for the small cell sizes when the dead region is fractionally the largest. This explains why degradation is seen only for the smallest cell size considered in the scintillator ECal option.
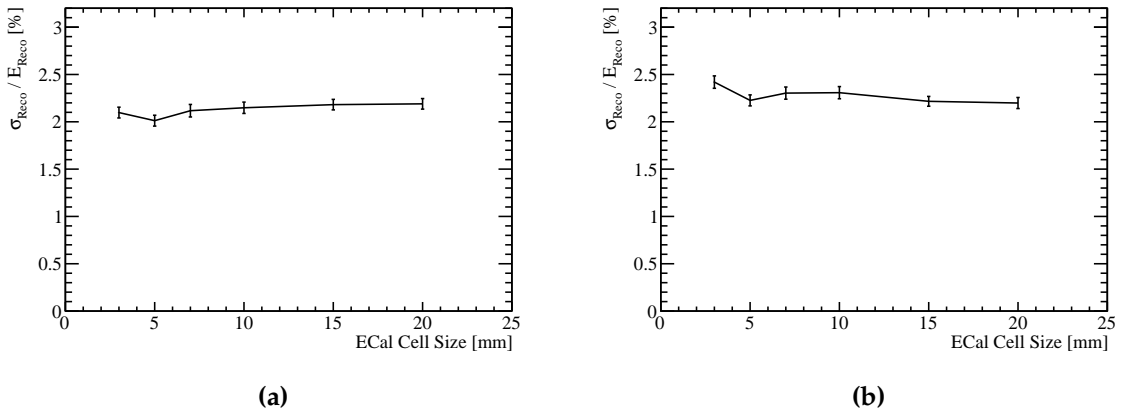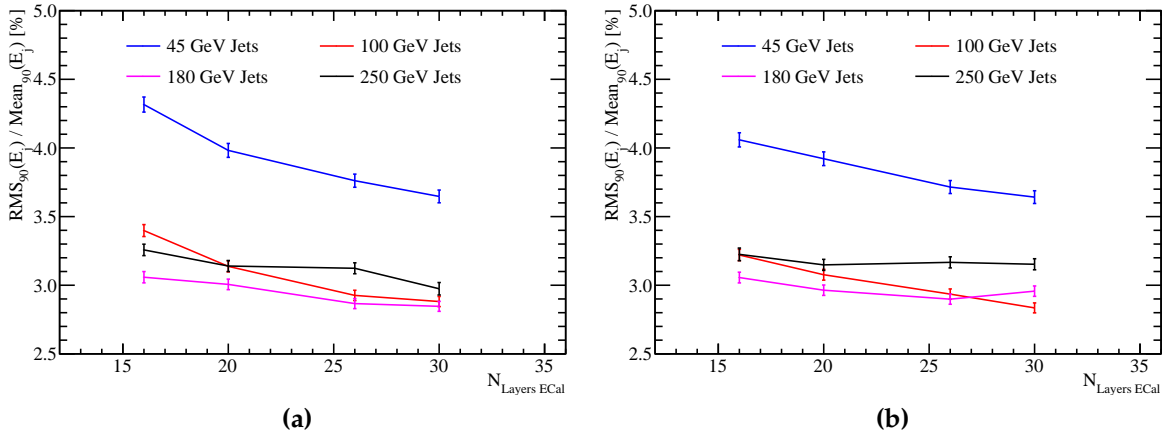


**Figure 2.5:** The energy resolution as a function of ECal cell size for 100 GeV $\gamma$s using the nominal ILD detector model with (a) the silicon and (b) the scintillator ECal option.

The separation of nearby particle showers within the calorimeter is fundamentally limited by the cell size. The smaller the cell size the easier it is to separate nearby particle showers and the better this separation the smaller the affect from confusion is. Therefore, it is expected that the jet energy resolution be sensitive to the ECal cell size even though the intrinsic energy resolution is not. The jet energy resolution as a function of ECal cell size is shown in figure 2.6a for the silicon option and figure 2.6b for the scintillator option. As expected there is a very strong dependancy on the ECal cell size, with smaller cell sizes leading to lower values of the jet energy resolution. The origin of this trend is best illustrated by considering the intrinsic energy resolution and confusion contributions to the jet energy resolution. These contributions are shown as a function of ECal cell size for 45 and 250 GeV jets, for both the silicon and scintillator ECal options, in figure 2.7. It is clear from these contributions that the intrinsic energy resolution of the detector does not change when varying the cell size, which agrees with both prior expectations of calorimeter behaviour and the single particle energy resolution study. The minor fluctuations seen in the energy resolution for the single particle study are washed out when considering the intrinsic energy resolution for jets, as only 30% of jet energy is carried in the form of $\gamma$s. Furthermore, it can be seen that the trend in jet energy resolution as a function of the ECal cell size is being driven

purely by changes to the confusion contribution and, in particular, the confusion caused by the reconstruction of $\gamma$s. This is exactly what is to expected given the ECal primarily measures $\gamma$s and shows that the performance of the ECal when varying the cell size is well understood.
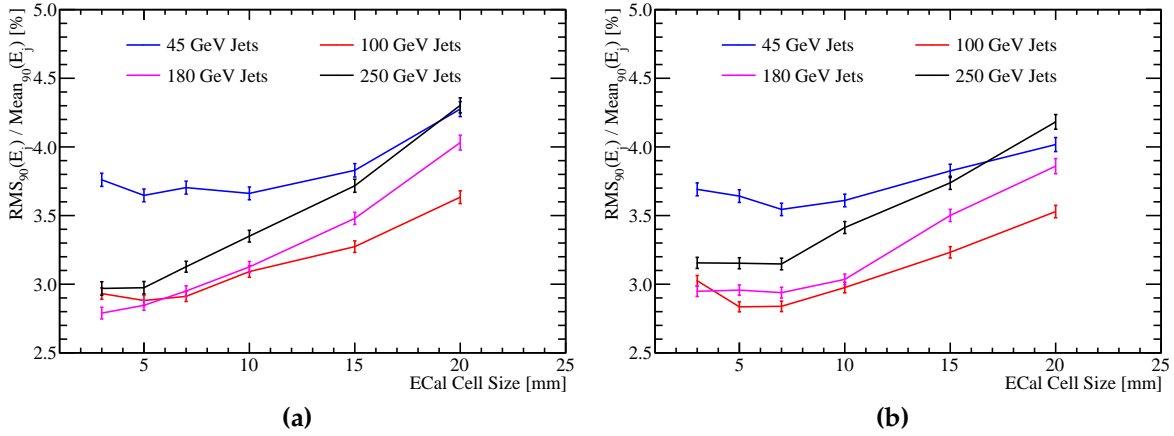


**Figure 2.6:** The jet energy resolution as a function of ECal cell size for various jet energies using the nominal ILD detector model with (a) the silicon and (b) the scintillator ECal option.

It is clear that the ECal cell size is extremely important for jet energy measurements, although it has little bearing on the intrinsic energy resolution of the ECal. To ensure separation of hadronic decays of W and Z bosons is possible at ILC like energies, an ECal cell size of least $15 \times 15\text{mm}^2$ is crucial. However, as reducing the ECal cell size beyond this leads to further improves the jet energy resolution choosing the smallest ECal cell size is desirable.

## 2.4.2 ECal Sampling Frequency

The detector performance was simulated where the number of layers in the ECal was varied, while keeping the total material budget ($X_0$) approximately constant. This study was performed for both the silicon and scintillator active material options. In all cases tungsten was used for the ECal absorber material and the active layer thicknesses were not changed from those used in the nominal ILD ECal summarised in table 2.2. The different layouts for the ECals considered here are summarised in table 2.3.
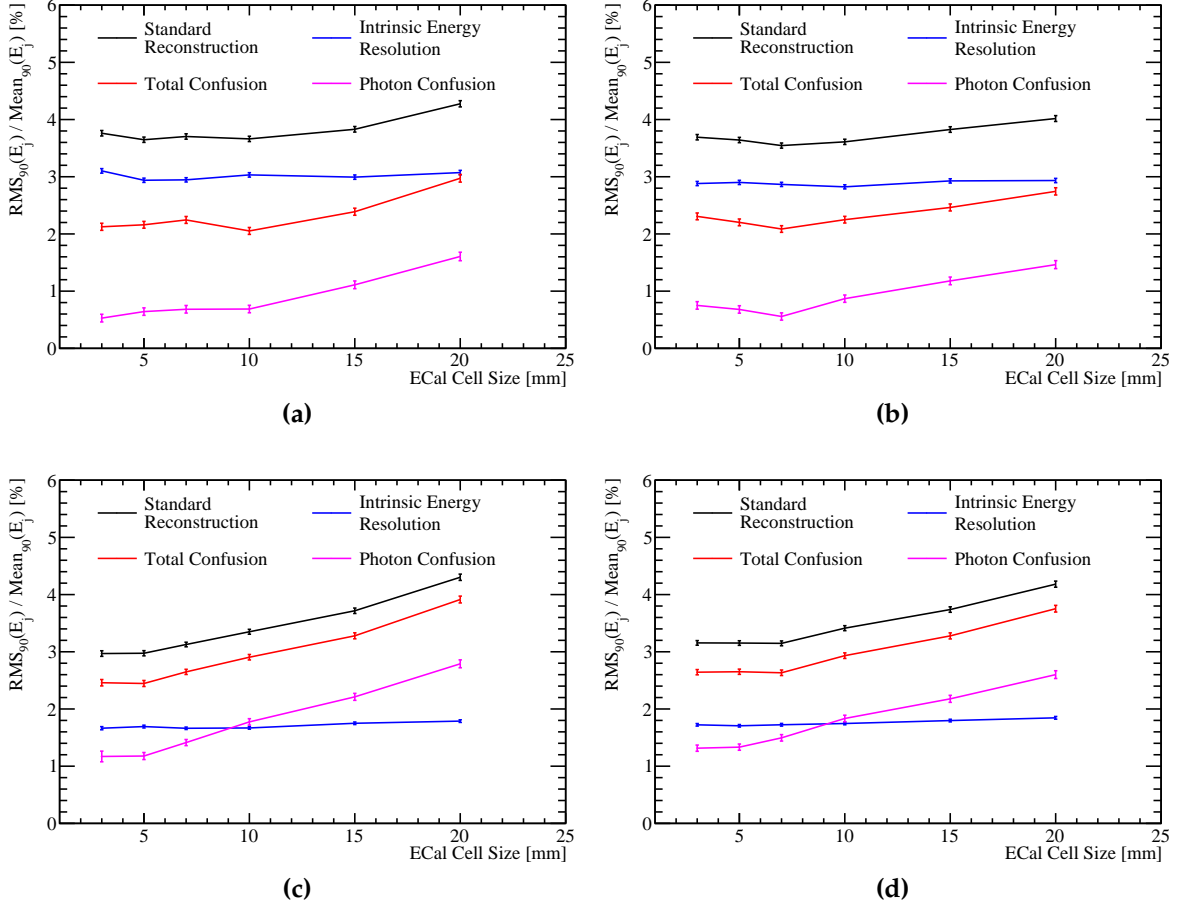
**Figure 2.7:** The contributions to the jet energy resolution as a function of ECal cell size using the nominal ILD detector model for (a) the silicon ECal option and 45 GeV jets, (b) the scintillator ECal option and 45 GeV jets, (c) the silicon ECal option and 250 GeV jets and (d) the scintillator ECal option and 250 GeV jets. The black curves correspond to the standard reconstruction, the blue curves to the intrinsic energy resolution contribution to the jet energy resolution, the red curves to the confusion contribution to the jet energy resolution and the magenta curves to the confusion contribution to the jet energy resolution related solely to $\gamma$ reconstruction

The energy resolution, for 100 GeV $\gamma$ events, as a function of the number of layers in the ECal is shown in figure 2.8a for the silicon option and in figure 2.8b for the scintillator option. As the number of layers is reduced the energy resolution increases, which is expected as more layers leads to greater sampling of the particle shower and a reduction in the stochastic contribution to the energy resolution.

When the number of layers in the ECal is increased, the intrinsic energy resolution of the ECal improves. This has the knock-on effect of reducing the confusion contribution to the jet energy resolution, which can be seen in figures 2.9a and 2.9b for the

| Total Number of Layers $N_{\text{Layers ECal}}$ | $N_{Layers}$ Region 1 | Absorber Thickness Region 1 [mm] | $N_{Layers}$ Region 2 | Absorber Thickness Region 2 [mm] | Total Thickness [$X_0$] |
|---|---|---|---|---|---|
| 30 | 20 | 2.10 | 9 | 4.20 | 22.77 |
| 26 | 17 | 2.40 | 8 | 4.80 | 22.60 |
| 20 | 13 | 3.15 | 6 | 6.30 | 22.47 |
| 16 | 10 | 4.00 | 5 | 8.00 | 22.31 |

**Table 2.3:** The longitudinal structure of the ECal models considered in the optimisation study. The radiation length of tungsten absorber is 3.504mm [20]. Note that a presampler layer contributes one extra layer to the cumulative number of layers.



**(a)**　　　　　　　　　　　　**(b)**
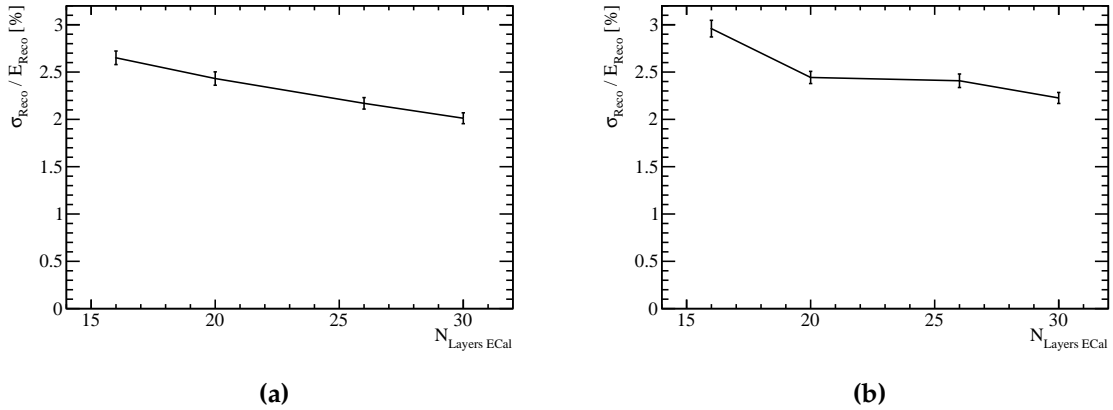
**Figure 2.8:** The energy resolution as a function of number of layers in the ECal for 100 GeV $\gamma$s using the nominal ILD detector model with (a) the silicon and (b) the scintillator ECal option.

silicon and scintillator ECal options respectively. In both cases the jet energy resolution was found to improve when the number of layers in the ECal was increased. The magnitude of the change in jet energy resolution was dependent upon the jet energy, with a stronger dependancy being observed for low energy jets. This is expected from the stochastic contribution to the energy resolution for a sampling calorimeter, which is $\propto \frac{1}{\sqrt{E \times N_{Layers}}}$ where $E$ is the reconstructed energy and $N_{Layers}$ is the number of layers in the calorimeter. At high jet energies, the energy resolution in the ECal is small and changes to the stochastic term that occur when varying the number of layers are too fine to be resolved using jet energy resolution. While at low jet energies the stochastic term is larger making it possible to resolve the changes to it when varying the number of layers in the ECal. The jet energy resolution is less sensitive than the

single $\gamma$ energy resolution to changes in the number of ECal layers as only $\approx 30\%$ of jet energy is carried in the form of $\gamma$s. The decomposition of the jet energy resolution into the intrinsic energy resolution and confusion contributions for 45 and 250 GeV jets are shown, for both the silicon and scintillator ECal options, in figure 2.10.

As expected, the improvement to the intrinsic energy resolution seen when increasing the number of layers in the ECal leads to the knock-on effect of lowering the confusion. However, significantly the magnitude of the change to the intrinsic energy resolution and confusion contributions to the jet energy resolution when varying the number of layers in the ECal are comparable in size. This shows that pattern recognition is as important for detector performance in the particle flow paradigm than intrinsic energy resolution.



**(a)**            **(b)**

**Figure 2.9:** The jet energy resolution as a function of number of layers in the ECal for various jet energies using the nominal ILD detector model with (a) the silicon and (b) the scintillator ECal option.

Increasing the number of layers in the ECal is beneficial to the intrinsic energy resolution and the jet energy resolution, particularly for low jet energies. Separation of the W and Z hadronic decays should be possible for ILC like energies given there are at least 26 layers in the ECal, however, it is desirable to have as large a number of layers as possible to benefit single $\gamma$ energy resolution.
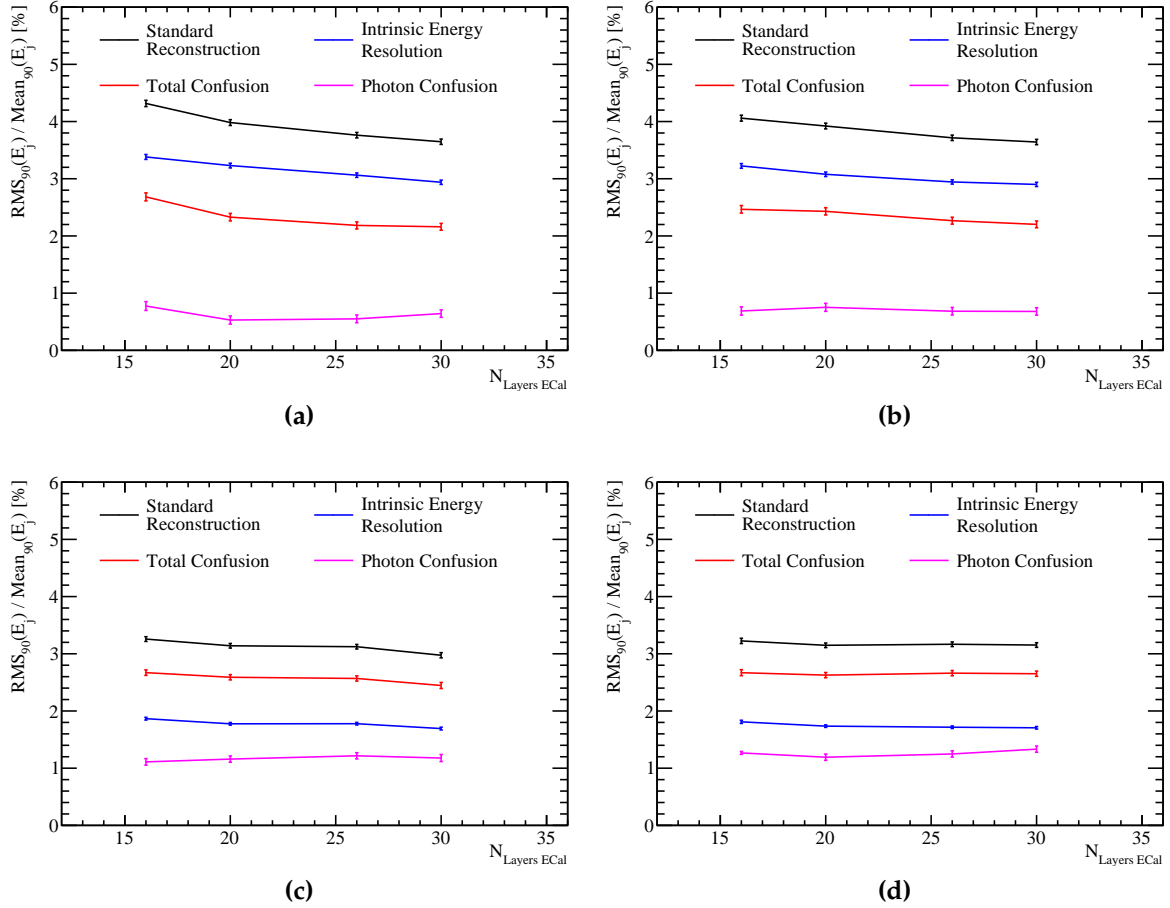
**Figure 2.10:** The contributions to the jet energy resolution as a function of number of layers in the ECal using the nominal ILD detector model for (a) the silicon ECal option and 45 GeV jets, (b) the scintillator ECal option and 45 GeV jets, (c) the silicon ECal option and 250 GeV jets and (d) the scintillator ECal option and 250 GeV jets. The black curves correspond to the standard reconstruction, the blue curves to the intrinsic energy resolution contribution to the jet energy resolution, the red curves to the confusion contribution to the jet energy resolution and the magenta curves to the confusion contribution to the jet energy resolution related solely to $\gamma$ reconstruction.

## 2.4.3 ECal Active Material

In sections 2.4.1 and 2.4.2 the performance of the ECal was reported for both the silicon and scintillator options and to a large extent the performance of the two options was similar, but not identical:

- The intrinsic energy resolution of the silicon ECal option is better than that of a scintillator option for high energies, see figures 2.4a and 2.4b. This is most likely

due to the implementation of Birks' law [9] for scintillator active materials., which states:

$$\frac{d\mathcal{L}}{dx} \propto \times \frac{dE/dx}{1 + k_B dE/dx} \, ,\qquad (2.3)$$

where $\frac{d\mathcal{L}}{dx}$ is the light yield per unit path length, $dE/dx$ is the energy deposited per unit path length and $k_B$ is a material property constant. For large energy deposits per unit length, such as those found in high energy $\gamma$ events, the light yield saturates causing a degradation in the energy resolution. Based on a comparison with the silicon ECal option performance, this effect starts to degrade the energy resolution for the scintillator option around 50 GeV. However, the degradation in energy resolution up to 100 GeV is relatively small.

- The "dead" region due to the presence of the MPPC in the simulation of the scintillator ECal option degrades performance of the detector for small transverse granularities, see figure 2.5.

In summary, the performance of the two options, in terms of energy and jet energy resolution, are similar, meaning no clear option is preferred. However, the silicon option is preferred when manufacture and implementation of the two models is compared. While constructing silicon wafers to fit a $5 \times 5\text{mm}^2$ square cell size is achievable, this would be extremely challenging for scintillator tiles. To resolve this in actuality, the scintillator ECal option would have to use $5 \times 45\text{mm}^2$ scintillator strips that are arranged in alternating directions in each ECal layer. By combining information from neighbouring layers it becomes possible to effectively achieve a $5 \times 5\text{mm}^2$ square cell size.

## 2.5 Hadronic Calorimeter Optimisation

The HCal primarily measures energy deposits from hadronic showers. The HCal in the default ILD detector model, summarised in table 2.4, is approximately 6 nuclear interaction lengths ($\lambda_I$) deep. The ECal contributes approximately one $\lambda_I$ giving a total of $\approx 7\lambda_I$, which is sufficient to contain jets at ILC like energies. The longitudinal structure of this model consists of 48 readout layers each containing a 3 mm active layer of scintillator and a 20 mm absorber layer of iron.

| Parameter | Default Value |
|---|---|
| Cell Size | $30 \times 30\text{mm}^2$ square cells |
| Number of Layers | 48 readout layers |
| Active Material Choice | Scintillator |
| Active Material Thickness | 3 mm |
| Absorber Material Choice | Steel |
| Absorber Material Thickness | 20 mm |

**Table 2.4:** The configuration of the HCal in the nominal ILD detector model.

There are several readout approaches under consideration for the HCal including fully analogue, fully digital and semi-digital. The analogue readout measures the energy within each HCal cell using a continuous spectrum of measurements, while the digital readout only produces a response if the energy deposited within a calorimeter cell is above a given threshold. The semi-digital approach mirrors that of the digital approach, but has three responses each with a different energy threshold. While the energy resolution for digital calorimeters is not as good as that of analogue calorimeters, it is possible to construct smaller cell sizes using a digital readout. In traditional calorimetry, a digital calorimeter would give a worse jet energy resolution than the analogue equivalent, however, that is not necessarily the case in particle flow calorimetry. If a digital calorimeter could be realised with a much small cell size than the analogue equivalent, then the affect of confusion in the digital calorimeter may be reduced so much that it compensates for any loss to intrinsic energy resolution. In the following studies optimisation of the analogue HCal is presented as this is the readout approach used in the nominal ILD detector model.

A number of options were simulated where the following parameters in the HCal were varied:

- Cell size: This is crucial for successful application particle flow calorimetry for making associations between clusters of calorimeter hits and charged particle tracks. It is expected that the intrinsic energy resolution be invariant to changes in the HCal cell size.

- Number of readout layers i.e. varying total depth of HCal: The number of layers in the HCal are varied, however, the thickness of those layers match those of the nominal ILD HCal design. This means the total depth of the HCal in $\lambda_I$ is

changing. It is expected that this study will determine the effect of leakage of energy out of the back of the HCal.

- Sampling frequency: This involves changing the of readout layers in the HCal while simultaneously changing the thicknesses of the active and absorber layers to keep the total number of $\lambda_I$ in the HCal constant. As this modifies the sampling of particle showers in the HCal, this will affect the intrinsic energy resolution of the HCal.

- Sampling fraction: This is the ratio of the active medium thickness to the absorber medium thickness. This controls how particle showers within the calorimeter are sampled. In this study the total depth of the HCal in $\lambda_I$ is held constant between detector models.

- Absorber material choice: Two options have been considered: steel and tungsten. This choice dictates the growth and propagation of hadronic showers and so plays a crucial role in calorimetry.

## 2.5.1 HCal Cell Size

The HCal cell size is an important detector parameter in the application of particle flow calorimetry. Smaller HCal cell sizes will lead to a finer spatial resolution that can be used to better separate charged and neutral particle calorimetric energy deposits. On the other hand, this will also lead to an increase in the number of readout channels that will raise the cost of the calorimeter. Therefore, it is highly desirable to achieve the optimal physics performance using the largest cell size possible. The nominal ILD HCal has a 30 mm square cell size and in this study the following square cell sizes were considered; 10 mm, 20 mm, 30 mm, 40 mm, 50 mm and 100 mm.

The energy resolution for 50 GeV $K_L^0$ events as a function of cell size in the HCal is shown in figure 2.11. These $K_L^0$ samples will deposit energy primarily within the HCal, making them appropriate events to consider when determining the performance of the HCal. However, a non-neglibile amount of energy will also be deposited within the ECal. Therefore, these energy resolutions represent the intrinsic energy resolution of the ILD detector as a whole and not purely that of the HCal. It is clear that there is no strong dependency on the energy resolution of the $K_L^0$ as a function of the HCal cell size. There are small fluctuations in the energy resolution that are most likely due to

variations in applying the calibration procedure and the tuning of the optimal HCal cell truncation, described in section 1.3.1. The precision on the optimal cell truncation is worse for large HCal cell sizes as the truncations considered in the optimisation are focused around the optimal truncation for the nominal ILD HCal.
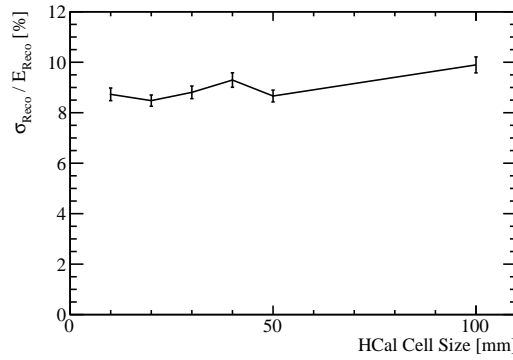


**Figure 2.11:** The energy resolution as a function of HCal cell size for 50 GeV $K^0_L$ events using the nominal ILD detector model.

As a smaller HCal cell size will lead to better separation of charged and neutral hadron calorimetric energy deposits, it is expected that the confusion contribution to the jet energy resolution will be reduced by using smaller HCal cell sizes. The jet energy resolution as a function of cell size in the HCal shown in figure 2.12. At low jet energies there is no strong dependency of the jet energy resolution on the HCal cell size, which is as expected from the $K^0_L$ energy resolution study. For high energy jets there is a clear dependence, with lower HCal cell sizes leading to better jet energy resolutions. Examining the different contributions to the jet energy resolution, shown in figure 2.13 it can be seen that the intrinsic energy resolution contribution is largely invariant to changes in the HCal cell size. Instead it is the confusion contribution that drives the overall trend in the jet energy resolution. This is particularly clear at high jet energies where the confusion contribution to the jet energy resolution dominates that of the intrinsic energy resolution contribution. At high jet energies smaller HCal cell sizes clearly lead to a reduction in the effect of confusion. At low jet energies the trend in confusion is less clear, as the confusion contribution is less dominant, but a reduction in the effect of confusion with decreasing cell size is still visible for all but the smallest HCal cell size. The most likely cause of the increase in confusion for the smallest HCal cell size at low energies is tuning of the PandoraPFA algorithms to the nominal ILD HCal cell size. Furthermore, for both energies the photon confusion is largely invariant to changes in the HCal cell size. This indicates that the changes in confusion

when varying the HCal cell size are due to pattern recognition improvements related to hadrons.
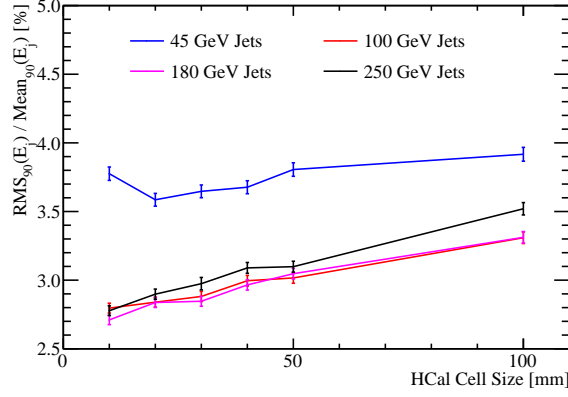


**Figure 2.12:** The jet energy resolution as a function of HCal cell size for various jet energies using the nominal ILD detector model.
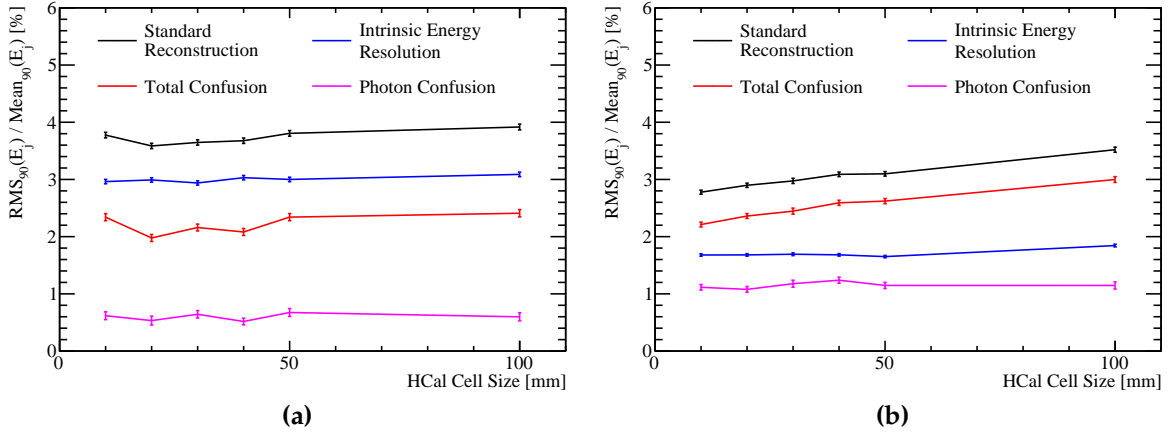


**Figure 2.13:** The contributions to the jet energy resolution as a function of HCal cell size using the nominal ILD detector model for (a) 45 GeV jets and (b) 250 GeV jets. The black curves correspond to the standard reconstruction, the blue curves to the intrinsic energy resolution contribution to the jet energy resolution, the red curves to the confusion contribution to the jet energy resolution and the magenta curves to the confusion contribution to the jet energy resolution related solely to $\gamma$ reconstruction.

A comparison of the results from the ECal and HCal cell size optimisation studies shows that the jet energy resolution has a much stronger dependency on the ECal cell size that the HCal cell size. This is to be expected as in the particle flow paradigm $\approx 30\%$ of jet energy is recorded in the ECal, while only $\approx 10\%$ is recorded in the HCal meaning the potential effect of double counting and omitting energy deposits, i.e.

confusion, is greater in the ECal than the HCal. Therefore, minimising confusion in the ECal should be more crucial for the overall jet energy resolution, which is what is observed. Furthermore, as PandoraPFA groups calorimeter hits together using a cone clustering approach, identifying the start of a particle shower is key for determining how calorimeters hits are grouped together deeper into the calorimeters. In effect, this means the grouping of calorimeter hits in the HCal heavily relies upon information gathered in the ECal. Therefore, if the ECal performance is sufficiently good, even with corse HCal cell sizes excellent performance can be achieved.

In summary, the confusion contribution to the jet energy resolution falls as the HCal cell size is reduced, while the intrinsic energy resolution of the detector is largely invariant to changes in the HCal cell size. As this dependancy is relatively weak, even the use of 100 mm square HCal cell sizes would be enough to allow for separation of the hadronic decays of W and Z bosons at ILC like energies. However, as jet energy resolution does improve with decreasing cell sizes it is desirable to have as small a HCal cell size as possible.

## 2.5.2  HCal Number of Layers

A number of detector models were simulated where the total number of layers in the HCal was varied. In contrast to the sampling frequency study, the thickness of the HCal layers were not modified. Changing the number of layers in this way leads to a change in the total thickness of the calorimeter. It is expected that this study will determine what effects, if any, leakage of energy out of the back of the calorimeter will have on detector performance. The cost of the HCal is proportional to the number of readout channels, which is proportional to the number of layers used in the HCal. Therefore, minimising the number of layers, while retaining excellent physics performance is vital. For this study detector models were simulated with a HCal containing 36, 42, 48 (nominal), 54 and 60 layers.

It is expected that the energy resolution of the detector will improve when the number of layers in the HCal is increased as fewer events should suffer from the effects of leakage. Any improvements seen by increasing the number of layers in the HCal is expected only up to a point, as beyond this point the majority of hadronic showers will be fully contained making additional HCal layers redundant. The energy resolution as a function of number of layers in the HCal for 50 GeV $K_L^0$ is shown in figure 2.14.

The energy resolution becomes worse as the number of layers in the HCal is reduced below 48 layers, while above this point additional layers do not change the energy resolution. This indicates that the majority of hadronic showers at this energy are fully contained by a 48 layer HCal. However, as reducing the number of HCal layers to 36 only causes a small degradation in the neutral hadron energy resolution, it is feasible for the ILD to consider using a thinner HCal design.
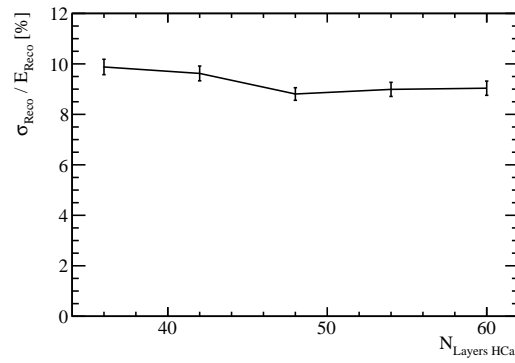


**Figure 2.14:** The energy resolution as a function of number of layers in the HCal for 50 GeV $K_L^0$ events using the nominal ILD detector model.

The trend in the 50 GeV $K_L^0$ energy resolution observed when varying the number of layers in the HCal is unlikely to be seen in the jet energy resolution as the trend weak and only a small fraction of jet energy is measured within the HCal. However, the confusion contribution to the jet energy resolution is expected to change as errors will be introduced, in the association of charged particle tracks to clusters of calorimeters hits, if part of the cluster energy has leaked from the back of the calorimeter. For example, if a particle shower from a charged particle suffers heavily from leakage there will be a large disparity between the track momentum and the energy it deposits within the calorimeter. In that case, PandoraPFA will be overly aggressive in associating other calorimeter energy deposits to this track to account for the energy that has leaked out of the calorimeter, which will increase the confusion. The jet energy resolution as a function of the number of layers in the HCal is shown in figure 2.15. As expected for low energy jets, where intrinsic energy resolution dominates, the jet energy resolution is invariant to the number of layers, while for high jet energies, where confusion dominates, a larger number of layers benefits the jet energy resolution. When examining the different contributions to the jet energy resolution, shown in figure 2.16, it becomes clear that the confusion contribution that drives the observed trends. The intrinsic energy resolution is largely invariant to changes to the number of

HCal layers as only a small fraction of jet energy is recorded in the HCal. Furthermore, the photon confusion is invariant to changes in the number of HCal layers, indicating that the change in the confusion contribution are originating from pattern recognition involving hadrons.
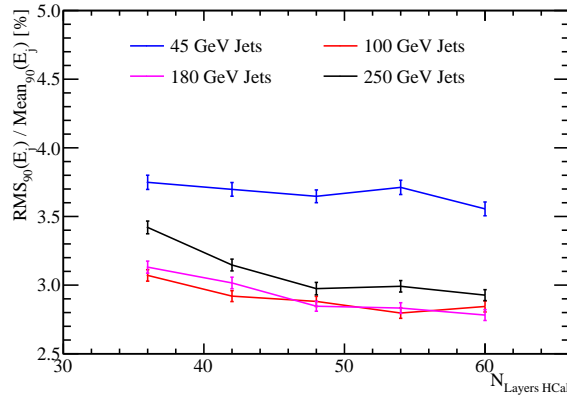


**Figure 2.15:** The jet energy resolution as a function of number of layers in the HCal for various jet energies using the nominal ILD detector model.
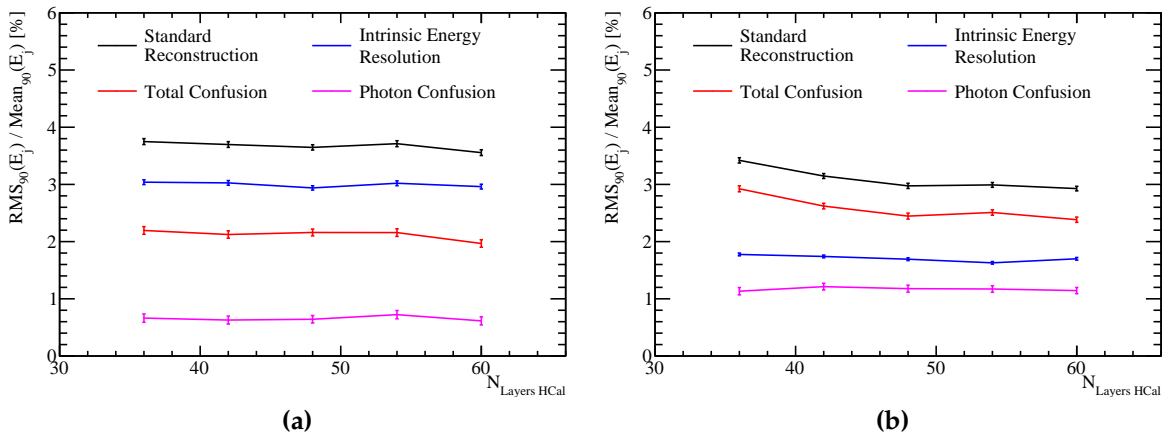


**Figure 2.16:** The contributions to the jet energy resolution as a function of number of layers in the HCal using the nominal ILD detector model for (a) 45 GeV jets and (b) 250 GeV jets. The black curves correspond to the standard reconstruction, the blue curves to the intrinsic energy resolution contribution to the jet energy resolution, the red curves to the confusion contribution to the jet energy resolution and the magenta curves to the confusion contribution to the jet energy resolution related solely to $\gamma$ reconstruction.

In summary, even if the number of layers in the HCal were reduced by a factor of 25% the jet energy resolution would be sufficient for separating the hadronic decays of the W and Z bosons at ILC energies. However, it is clear that leakage of energy from

the back of the HCal would negatively affect events at ILC like energies if the number of layers is reduced from the nominal value of 48 layers, therefore, it is desirable to have a minimum of 48 layers in the HCal.

### 2.5.3 HCal Sampling Frequency

The sampling frequency in the HCal was by varying the number of readout layers in the HCal, while simultaneously varying the active and absorber layer thicknesses such that the total number of nuclear interaction lengths in the HCal is constant. For each model considered, the absorber material was steel, containing a total of 5.72 $\lambda_I$, while the active material was scintillator, containing a total of 0.19 $\lambda_I$. Furthermore, the ratio of the active to absorber layers thicknesses (the sampling fraction) was not changed. A summary of the detector models considered in this study can be found in table 2.5.

| Number $N_{\text{Layers HCal}}$ | Absorber Thickness [mm] | Active Thickness [mm] |
|---|---|---|
| 60 | 16.00 | 2.40 |
| 54 | 17.78 | 2.67 |
| 48 | 20.00 | 3.00 |
| 42 | 22.86 | 3.43 |
| 36 | 26.67 | 4.00 |
| 30 | 32.00 | 4.80 |
| 24 | 40.00 | 6.00 |
| 18 | 53.33 | 8.00 |

**Table 2.5:** Cell size layout of various HCal models considered.

Increasing the number of layers in the HCal will increase the number of times a showering particle is sampled, which in turn will reduce the stochastic term in the energy resolution for a sampling calorimeter. Therefore, it is expected that the energy resolution will improve with increasing number of layers in the HCal. This is shown for 50 GeV $K_L^0$ in figure 2.17. Although there are fluctuations in the energy resolution, due to both fitting the Gaussian to extract the energy resolution and in the calibration of the detector, there is a clear trend showing that the energy resolution of the calorimeter is strongly dependent upon the sampling frequency in the HCal. The

trend in the energy resolution does not exactly follow a $\frac{1}{\sqrt{N_{\text{Readout Layers HCal}}}}$ relationship, but this is to be expected as this relationship only holds for the energy resolution of the HCal. On the other hand, these results are for the whole ILD detector, including the $\approx 1\lambda_I$ in the ECal. Furthermore, this functional form neglects the constant term in the energy resolution described in section 2.3.
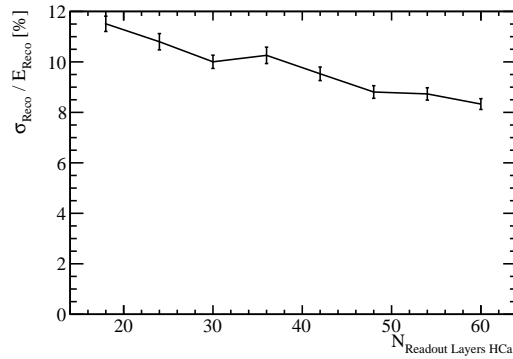


**Figure 2.17:** The energy resolution as a function of sampling frequency in the HCal for 50 GeV $K_L^0$ events using the nominal ILD detector model.

As changing the sampling frequency of a calorimeter alters the intrinsic energy resolution, it is expected that this will produce a twofold effect on the jet energy resolution whereby the intrinsic energy improvements have the knock-on effect of lowering confusion, as described in section 2.4.2, for changes to the number of layers in the ECal. The jet energy resolution as a function of sampling frequency in the HCal is shown in figure **??**. As expected there is an improvement in jet energy resolution at all energies due as both the intrinsic energy resolution and confusion are reduced with increasing sampling frequency, which can be seen when examining the different contributions to the jet energy resolution, shown in figure 2.19.

It is clear that a larger number of layers in the HCal benefits both the intrinsic energy resolution of the ILD detector as well as reducing the confusion contribution to the jet energy resolution. As there are few physics analyses that rely on the identification and categorisation of individual neutral hadrons, but there are many that rely on identification and categorisation of $\gamma$s, the intrinsic energy resolution of the HCal is less crucial from a physics perspective than that of the ECal. However, these studies show the HCal still has a crucial role to play in jet reconstruction in the particle flow paradigm and therefore cannot be neglected. To achieve a jet energy resolution of $\frac{\sigma_E}{E} \lesssim 3.8\%$, which is required to separate the W and Z hadronic decays, the ILD detector will require a minimum of 42 layers in the HCal. This sampling frequency is required
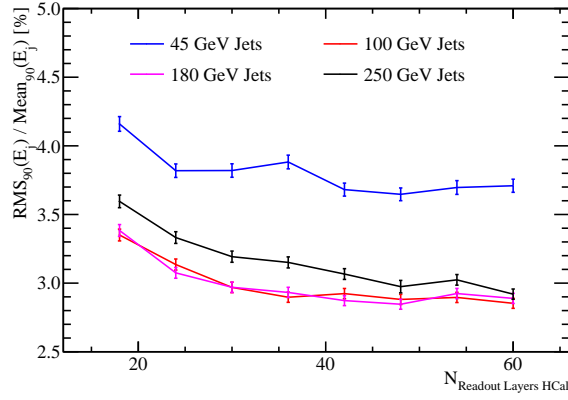
**Figure 2.18:** The jet energy resolution as a function of sampling frequency in the HCal for various jet energies using the nominal ILD detector model.



**(a)**

**(b)**
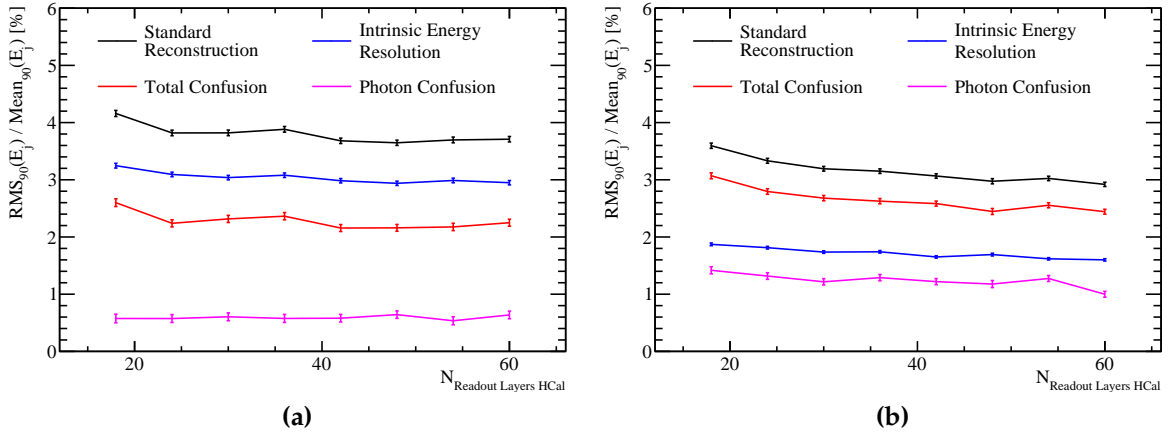
**Figure 2.19:** The contributions to the jet energy resolution as a function of sampling frequency in the HCal using the nominal ILD detector model for (a) 45 GeV jets and (b) 250 GeV jets. The black curves correspond to the standard reconstruction, the blue curves to the intrinsic energy resolution contribution to the jet energy resolution, the red curves to the confusion contribution to the jet energy resolution and the magenta curves to the confusion contribution to the jet energy resolution related solely to $\gamma$ reconstruction.

in particular for lower energy jets where the energy resolution is dominated by the intrinsic energy resolution of the detector, while at higher energies the resolution is more than good enough to achieve the separation of the decay channels.

### 2.5.4 HCal Sampling Fraction

The ILD detector performance was simulated where the ratio of the active to absorber layer thicknesses in the HCal were varied. In the nominal detector model the active scintillator layer thickness is 3 mm, while the absorber layer thickness is 20 mm giving a sampling fraction of 0.15. HCal models were simulated where this ratio was changed from 0.05 to 0.25 in steps of 0.05, while retaining the same number of interaction lengths in the absorber and active layers as is found in the nominal HCal model. If the active layer thickness becomes excessively small then it is possible that any signal produced in that layer would be insufficient to accurately estimate the energy deposited within the surrounding absorber layers. However, no performance differences were observed for any of these detector models when considering the energy resolution for 50 GeV $K_L^0$ events or the jet energy resolution for the 91, 200, 360 and 500 GeV Z→uds di-jet events. This indicates the particle showers are sufficiently well sampled in all detector models considered and that thinning the active layer of the ILD HCal down to $\sim 1$ mm would not harm the detector performance.

### 2.5.5 HCal Absorber Material

The nominal choice of absorber material is steel with tungsten providing a feasible alternative [13]. Although tungsten is more expensive than steel, it contains a larger number of nuclear interaction lengths per unit length. Therefore, using tungsten as opposed to steel as the absorber material would reduce the size of the HCal, while retaining the same number of nuclear interaction lengths. Reducing the depth of the calorimeter would decrease the size of the solenoid required, which would offset some of the additional cost if tungsten were used as the absorber material.

| Parameter | Steel HCal Option | Tungsten HCal Option |
|---|---|---|
| Cell Size | $30 \times 30\text{mm}^2$ square cells | $30 \times 30\text{mm}^2$ square cells |
| Number of Layers | 48 readout layers | 48 readout layers |
| Absorber Material Thickness [mm] | 20.0 | 12.0 |
| Active Material Choice | Scintillator | Scintillator |
| Active Material Thickness [mm] | 3.0 | 1.8 |

**Table 2.6:** The configuration of the stainless steel and tungsten HCal options.

The configuration for the stainless steel and tungsten HCal options that were used in the full ILD simulation can be found in table 2.6. To isolate the effects of changing the absorber material, the total depth, in nuclear interaction lengths, was kept constant when comparing these two options. Furthermore, the sampling fraction was also held constant. The interaction of hadrons with the absorber material within the detector is simulated by GEANT4. A number of different physics lists exist within GEANT4 for the modelling of hadronic showers. The default model for high energy physics calorimetry is the QGSP_BERT physics list, which uses the quark-gluon string model [16] with the precompound model of nuclear evaporation [1] (QGSP) for high energy interactions and the Bertini (BERT) cascade model [17] for intermediate energy interactions. For this study both the QGSP_BERT and the QGSP_BERT_HP physics lists were used. The QGSP_BERT_HP list uses the high precision neutron package (NeutronHP) to deal with the transportation of neutrons from below 20 MeV to thermal energies. This added detail was thought to be necessary for a study involving tungsten due to the expected increase in shower development.

One of the dominant processes governing the energy deposition of hadronic showers in calorimeters is spallation [22]. Spallation begins with the collision of a high energy incident particle with an atomic nuclei from the calorimeter absorbing material. This collision creates an internuclear cascade where a shower of high energy hadronic particles, e.g. protons, neutrons and pions, are produced within the nucleus. If these energies are large enough, some of these particles may escape the nucleus and form secondary particles in the hadronic shower. After this initial collision, the nuclei of the absorbing material are left in an excited state. Assuming the excited nuclei are sufficiently stable that they will not undergo fission, they will return to a stable state by ejecting energy in the form of particles in a process called evaporation. Evaporation of neutrons, which is the dominant form of evaporation, significantly delays the growth of hadronic showers as some of these neutrons undergo neutron capture [7]. Neutron capture involves an absorber nuclei capturing a neutron and then emitting a $\gamma$ as it returns to a stable state. Therefore, the time it take for the neutron capture mechanism to proceed is limited by the lifetime of the unstable nuclei. This makes neutron capture one of the slowest mechanisms by which hadronic showers can propagate. As absorber materials with a large atomic number, Z, have a larger number of neutrons, it is expected that there will be an increase in the number of evaporation neutrons within hadronic showers developing in such materials. In turn this will lead to more neutron capture processes and a longer development time for the hadronic showers.

This is what is observed when considering the shower development times using the tungsten (Z=74) and steel (iron, Z=26) HCal options as seen in figure 2.20.
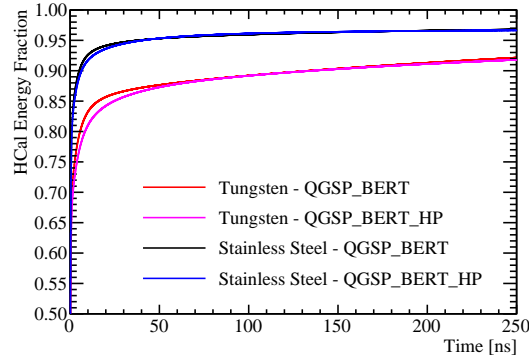


**Figure 2.20:** The fraction of the total calorimetric energy deposited in the HCal as a function of time for 25 GeV $K_L^0$ events using the steel and tungsten HCal options. Results are shown for both the QGSP_BERT and QGSP_BERT_HP physics lists. The calorimeter hit times have been corrected for straight line time of flight to the impact point.

The energy resolution for 50 GeV $K_L^0$ events using the nominal ILD detector model with a stainless steel and tungsten HCal absorber material can be found in table 2.7. The tungsten HCal option offers a slight improvement over steel in the energy resolution for these samples. This will be caused by differences in the atomic structure of the two materials, which will lead to different developments of the hadronic showers within them. One example of this would be that energy losses to nuclear binding energies will be smaller in tungsten than steel (iron) as the atomic nuclei for tungsten is less stable and so less energy is needed to liberate nucleons in the shower development. This will lead to a larger signal in the tungsten calorimeter and a reduction in the energy resolution in comparison to steel. These results also indicated that the addition of the high precision neutron package did not alter the detector performance significantly for either option.

It should be emphasised that in determining these results, the HCal cell truncation, as described in chapter 1, was separately tuned for the tungsten option. This had to be done as the average cell energy is greater when using tungsten, as opposed to steel, for the HCal absorber material because tungsten has a larger number of radiation lengths per unit length than steel does. As the HCal primarily measures hadronic showers one may naively expect the number of radiation lengths in the HCal to be irrelevant to performance, given both options have the same number of nuclear interaction lengths,

| HCal Option | Energy Resolution [%] |
|---|---|
| Stainless Steel, QGSP_BERT | $8.8 \pm 0.2$ |
| Stainless Steel, QGSP_BERT_HP | $9.0 \pm 0.3$ |
| Tungsten, QGSP_BERT | $8.3 \pm 0.2$ |
| Tungsten, QGSP_BERT_HP | $8.3 \pm 0.2$ |

**Table 2.7:** The energy resolution using the nominal ILD detector with various HCal options determined using 50 GeV $K_L^0$ events.

but this is not the case as all hadronic showers have an electromagnetic component, from the decays of $\pi^0 \rightarrow \gamma\gamma$. Therefore, these shower components deposit more energy per unit length in the HCal, which raises the average cell energy.

The jet energy resolutions for various jet energies are shown in table 2.8 for the various HCal options considered. These results indicate that steel outperforms tungsten as the absorber material for the HCal. Furthermore, as the jet energy increases, the magnitude of the difference in jet energy resolutions between the two options grows. This indicates the differences in jet energy resolution between the two options is driven by the confusion contribution since this contribution grows with increasing jet energy. Furthermore, as the $K_L^0$ energy resolution was only slightly better for the tungsten option it is expected that the intrinsic energy resolution contribution to the jet energy resolution will not vary significantly between the two as only a small fraction of jet energy is measured in the HCal. The intrinsic energy resolution and confusion contributions to the jet energy resolution for 45 and 250 GeV jets are shown in table 2.9. As expected, the intrinsic energy resolution contribution to the jet energy resolution is nearly identical between the various options. The confusion contribution to the jet energy resolution is larger for the tungsten HCal option than for the steel HCal option. This will be due to the PandoraPFA algorithms being tuned for HCal cell dimensions for the steel HCal option, while the cells for the tungsten option are thinner by a factor of approximately $\frac{\lambda_I^{Steel}}{\lambda_I^{Tungsten}} \approx 1.7$, where $\lambda_I^x$ is the distance of one radiation length in material $x$. It is unfeasible to tune all of the PandoraPFA algorithms to each detector geometry, however, the breakdowns of the jet energy resolution indicate that even if it were possible to obtain the same confusion contributions for both options, the tungsten option would offer no advantage to the steel option in terms of the intrinsic jet energy resolution. Once again, it was noted that the use of the QGSP_BERT_HP physics list, as opposed to QGSP_BERT, made a minimal impact on these results.

| HCal Option | Jet Energy Resolution [%] | | | |
|---|---|---|---|---|
| | 45 GeV | 100 GeV | 180 GeV | 250 GeV |
| Stainless Steel, QGSP_BERT | $3.65 \pm 0.05$ | $2.88 \pm 0.04$ | $2.85 \pm 0.04$ | $2.97 \pm 0.05$ |
| Stainless Steel, QGSP_BERT_HP | $3.67 \pm 0.05$ | $2.92 \pm 0.04$ | $2.86 \pm 0.04$ | $3.03 \pm 0.04$ |
| Tungsten, QGSP_BERT | $3.78 \pm 0.05$ | $3.12 \pm 0.04$ | $3.15 \pm 0.04$ | $3.43 \pm 0.04|$ |
| Tungsten, QGSP_BERT_HP | $3.80 \pm 0.05$ | $3.08 \pm 0.04$ | $3.24 \pm 0.04$ | $3.41 \pm 0.04$ |

**Table 2.8:** The jet energy resolution using the nominal ILD detector with various HCal options for various jet energies.

| HCal Option | Jet Energy Resolution [%] | | | |
|---|---|---|---|---|
| | 45 GeV | | 250 GeV | |
| | Intrinsic | Confusion | Intrinsic | Confusion |
| Stainless Steel, QGSP_BERT | $2.93 \pm 0.04$ | $2.16 \pm 0.06$ | $1.69 \pm 0.02$ | $2.45 \pm 0.05$ |
| Stainless Steel, QGSP_BERT_HP | $2.98 \pm 0.04$ | $2.15 \pm 0.06$ | $1.65 \pm 0.02$ | $2.53 \pm 0.04$ |
| Tungsten, QGSP_BERT | $2.97 \pm 0.04$ | $2.34 \pm 0.06$ | $1.65 \pm 0.02$ | $3.01 \pm 0.05$ |
| Tungsten, QGSP_BERT_HP | $2.92 \pm 0.04$ | $2.42 \pm 0.06$ | $1.65 \pm 0.02$ | $2.99 \pm 0.05$ |

**Table 2.9:** The contributions to the jet energy resolution using the nominal ILD detector with various HCal options for 45 and 250 GeV jet energies.

In conclusion, there are no large differences in the intrinsic energy resolution of the ILD detector simulation, for either neutral hadrons or jets, when changing the HCal absorber material from steel to tungsten. The steel option HCal outperforms the tungsten option in terms of pattern recognition confusion, when using the default PandoraPFA settings, although this could be addressed should it become clear that tungsten were a preferred option. However, when examining the mechanical properties of steel and tungsten, it is clear that steel has a significant advantage over tungsten in terms of rigidity [19]. This means that fewer support structures would be required for the calorimeter leading to less dead material and better performance, which makes steel the more preferred option of the two.

# 2.6 Global Detector Parameters

The detector geometry and the magnetic field strength are major cost drivers for the detector. Both will affect the jet energy resolution and, for completeness, are considered here.

## 2.6.1 The Magnetic Field Strength

In the particle flow paradigm the momentum of charged particles is obtained through the curvature of the track they traverse as they bend in the magnetic field. Therefore, the magnetic field is an integral element for the successful application of particle flow calorimetry. Furthermore, the magnetic field deflects charged particles away from neutral ones in jets. The stronger the magnetic field, the larger the separation between the calorimetric energy deposits made by charged and neutral particles from jets. This, in turn, makes it easier to associate charged particle tracks to the correct calorimetric energy deposits. Therefore, it is expected that a stronger magnetic field will lead to better jet energy resolutions through a reduction of the confusion contribution to the jet energy resolution.

Detector models were simulated where the magnetic field was varied from 1 to 5 T in steps of 0.5 T and the jet energy resolutions for these detectors is shown in figure 2.21. For high energy jets there is a strong trend whereby a stronger magnetic field lead to a better jet energy resolution. On the other hand, at low jet energies the jet energy resolution is almost invariant to changes in the magnetic field.

The jet energy resolution breakdowns into the various contributions are shown in figure 2.22 and as expected there is clear reduction in the confusion contribution to the jet energy resolution with increasing magnetic field strength. Furthermore, there is a reduction in intrinsic energy resolution with increasing magnetic field strength for low energy jets. At low energies, the radius of curvature of the helix charged particles traverse will be small. If the radius for a given particle is small enough, it will not make it into the calorimeters. In this case, only if the track produced from this particle passes tight selection cuts, designed to ensure the track originates from the IP, will the track be used in the reconstruction. Therefore, energy can and will be lost from events where particles are stuck within the tracker. Given the radii of curvature is inversely proportional to the magnetic field strength, the larger the magnetic field strength the

more tracks will be confined to the tracker as is shown by figure 2.23. The more tracks that are confined to the tracker, the worse the intrinsic energy resolution becomes as inevitably some tracks fail the quality cuts. At high jet energies, low transverse momentum charged particles will still get trapped within the tracker, however, these contribute fractionally little energy to the total reconstructed energy. Therefore, the trend of worsening intrinsic energy resolution with increasing magnetic field strength is less pronounced as the jet energy grows. At very low magnetic field strengths and high jet energies, the intrinsic energy resolution actually degrades, due to an artefact in the determination of the intrinsic energy resolution. The intrinsic energy resolution is determined by associating a single MC particle to each calorimeter cell. At low magnetic field strengths and high jet energies, many of the MC particles will have overlapping energy deposits wihtin the calorimeter cells and so associating a single MC particle per cell is inaccurate. This leads to imperfect association of charged particle tracks to calorimetric energy deposits, which worsens the intrinsic energy resolution. However, as this effect is second order small in comparison to changes in the confusion contribution the overall dependancy of detector performance on the magnetic field strength can be confidently quantified.
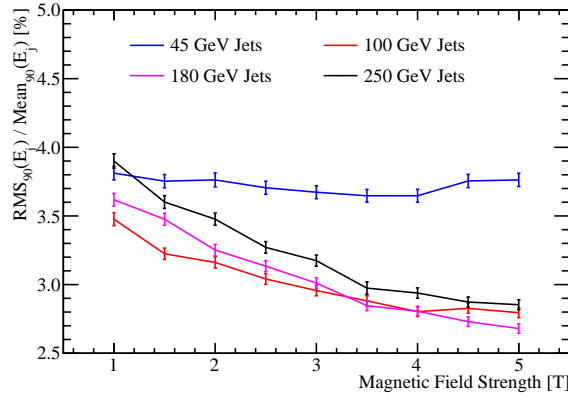


**Figure 2.21:** The jet energy resolution using the nominal ILD detector as a function of the magnetic field strength for various jet energies.

In summary, increasing the magnetic field strength is beneficial to detector performance as it reduces confusion from associating tracks to calorimetric energy deposits from charged particles. While there is a reduction in the intrinsic energy resolution for low transverse momentum jets with increasing magnetic field strength, this effect is largely offset by the change in confusion. While the nominal field of 3.5 T gives good performance increasing the field strength is a clear way of making gains in detector performance.
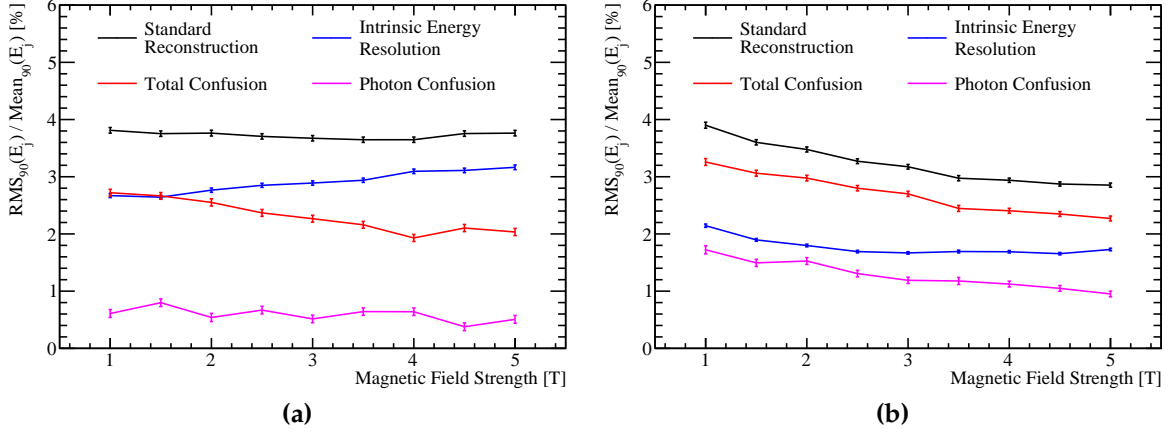
**Figure 2.22:** The contributions to the jet energy resolution as a function of the magnetic field strength using the nominal ILD detector model for (a) 45 GeV jets and (b) 250 GeV jets. The black curves correspond to the standard reconstruction, the blue curves to the intrinsic energy resolution contribution to the jet energy resolution, the red curves to the confusion contribution to the jet energy resolution and the magenta curves to the confusion contribution to the jet energy resolution related solely to $\gamma$ reconstruction.
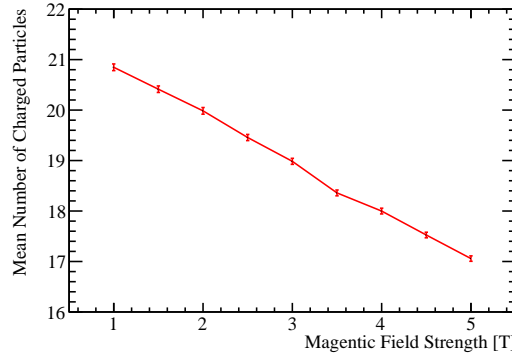


**Figure 2.23:** The mean number of reconstructed charged particles as a function of the magnetic field strength for 91 GeV Z→uds di-jet events. The nominal ILD detector model was used and the pattern recognition has been fully cheated using the MC information.

## 2.6.2 Inner ECal Radius

Detector models were considered where the ECal inner radius was set to 1208, 1408, 1608, 1808 (nominal) and 2008 mm. All other detector parameters are identical to those found in the nominal ILD detector model.

Figure 2.24 shows the dependence of the jet energy resolution on the ECal inner radius. The simulations indicate that a large ECal inner radius was highly beneficial to detector performance, which is due to the increase in the time it takes for particles to reach the calorimeters. The longer it takes for the particles to reach the calorimeters the more the charged particles will bend and the larger separation between calorimetric energy deposits from charged and neutral particles. This larger separation reduces the confusion when associating calorimetric energy deposits to tracks and so improves the detector performance. This conclusion is backed up by the decomposition of the jet energy resolution for the low and high energy jets, shown in figure 2.25, which explicitly show a reduction in confusion with increasing ECal inner radius. The intrinsic energy resolution of the detectors showed no strong dependence on the inner ECal radius. There is a small degradation in intrinsic energy resolution at low ECal inner radii due to the association of a single MC particle per calorimeter cell when running the cheated pattern recognition as explained in section 2.6.1. Again, this effect has little bearing on the final conclusions as the change in intrinsic energy resolution across the detector models is a second order effect.
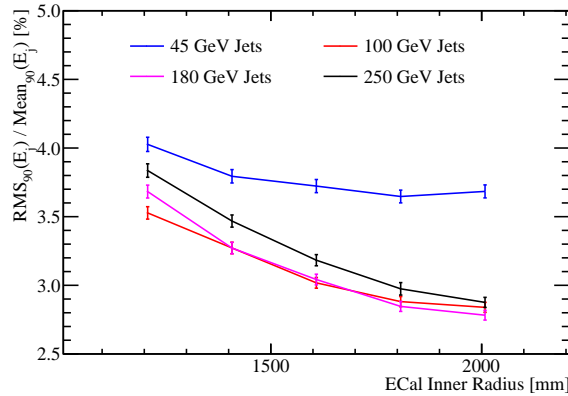


**Figure 2.24:** The jet energy resolution using the nominal ILD detector as a function of the ECal inner radius for various jet energies.

In conclusion, increasing the ECal inner radius benefits the jet energy resolution significantly. This trend is driven by changes to the confusion in associating calorimetric energy deposits to charged particle tracks, with a larger ECal inner radius producing a reduction in the confusion as separation of charged and neutral particle energy deposits increases.
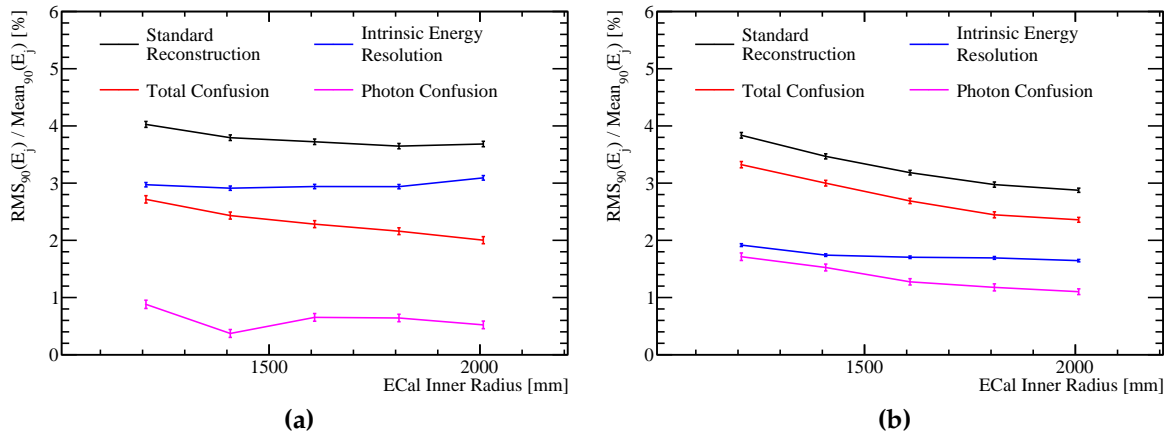
**Figure 2.25:** The contributions to the jet energy resolution as a function of the ECal inner radius using the nominal ILD detector model for (a) 45 GeV jets and (b) 250 GeV jets. The black curves correspond to the standard reconstruction, the blue curves to the intrinsic energy resolution contribution to the jet energy resolution, the red curves to the confusion contribution to the jet energy resolution and the magenta curves to the confusion contribution to the jet energy resolution related solely to $\gamma$ reconstruction.

# Colophon

This thesis was made in LaTeX $2_\varepsilon$ using the "hepthesis" class [10].

# Bibliography

[1] *Geant4 Physics Reference Manual, Section IV, Chapter 28.*

[2] Georges Aad et al. Measurement of $W^+W^-$ production in pp collisions at $\sqrt{s}$=7âĂĽâĂĽTeV with the ATLAS detector and limits on anomalous WWZ and WWÎş couplings. *Phys. Rev.*, D87(11):112001, 2013, 1210.2979. [Erratum: Phys. Rev.D88,no.7,079906(2013)].

[3] Toshinori Abe et al. The International Large Detector: Letter of Intent. 2010, 1006.3396.

[4] H. Abramowicz et al. Higgs Physics at the CLIC Electron-Positron Linear Collider. 2016, 1608.07538.

[5] Halina Abramowicz et al. The International Linear Collider Technical Design Report - Volume 4: Detectors. 2013, 1306.6329.

[6] C. Adloff et al. Hadronic energy resolution of a highly granular scintillator-steel hadron calorimeter using software compensation techniques. *JINST*, 7:P09017, 2012, 1207.4210.

[7] C. Adloff et al. The Time Structure of Hadronic Showers in highly granular Calorimeters with Tungsten and Steel Absorbers. *JINST*, 9:P07022, 2014, 1404.6454.

[8] H. Bichsel, Donald E. Groom, and S. R. Klein. Passage of particles through matter. 2004.

[9] J. B. Birks. *The theory and practice of scintillation counting*. Pergamon Press Oxford, 1964.

[10] Andy Buckley. The hepthesis LaTeX class.

[11] Serguei Chatrchyan et al. Measurement of the $W\gamma$ and $Z\gamma$ inclusive cross sections

in $pp$ collisions at $\sqrt{s} = 7$âĂĽâĂĽTeV and limits on anomalous triple gauge boson couplings. *Phys. Rev.*, D89(9):092005, 2014, 1308.6832.

[12] Serguei Chatrchyan et al. Search for $WW\gamma$ and $WZ\gamma$ production and constraints on anomalous quartic gauge couplings in $pp$ collisions at $\sqrt{s} = 8$ TeV. *Phys. Rev.*, D90(3):032008, 2014, 1404.4619.

[13] M. Chefdeville et al. Shower development of particles with momenta from 15 GeV to 150 GeV in the CALICE scintillator-tungsten hadronic calorimeter. *JINST*, 10(12):P12006, 2015, 1509.00617.

[14] 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.

[15] C. W. Fabjan and F. Gianotti. Calorimetry for particle physics. *Rev. Mod. Phys.*, 75:1243–1286, 2003.

[16] G. Folger and J. P. Wellisch. String parton models in GEANT4. *eConf*, C0303241:MOMT007, 2003, nucl-th/0306007.

[17] M. P. Guthrie, R. G. Alsmiller, and H. W. Bertini. Calculation of the capture of negative pions in light elements and comparison with experiments pertaining to cancer radiotherapy. *Nucl. Instrum. Meth.*, 66:29–36, 1968.

[18] L. Landau. On the energy loss of fast particles by ionization. *J. Phys.(USSR)*, 8:201–205, 1944.

[19] Lucie Linssen, Akiya Miyamoto, Marcel Stanitzki, and Harry Weerts. Physics and Detectors at CLIC: CLIC Conceptual Design Report. 2012, 1202.5940.

[20] C. Patrignani et al. Review of Particle Physics. *Chin. Phys.*, C40(10):100001, 2016.

[21] M. A. Thomson. Particle Flow Calorimetry and the PandoraPFA Algorithm. *Nucl. Instrum. Meth.*, A611:25–40, 2009, 0907.3577.

[22] R. Wigmans. Calorimetry: Energy measurement in particle physics. *Int. Ser. Monogr. Phys.*, 107:1–726, 2000.

# List of figures

# List of tables