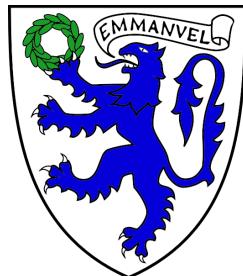


Calorimetry at a Future Linear Collider

Steven Green
of Emmanuel College



A dissertation submitted to the University of Cambridge
for the degree of Doctor of Philosophy
August 2017

Abstract

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

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

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

and $2ab^{-1}$ 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.

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

Declaration

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

Steven Green

Acknowledgements

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

Contents

1 Particle Flow Calorimetry for Future Linear Colliders	1
1.1 Particle Flow Calorimetry	1
1.2 International Large Detector	4
1.2.1 Overview	4
1.2.2 Vertex Detector	5
1.2.3 Time Projection Chamber	6
1.2.4 Supplemental Silicon Tracking System	8
1.2.5 Electromagnetic Calorimeter	10
1.2.6 Hadronic Calorimeter	11
1.2.7 Solenoid, Yoke and Muon System	13
1.2.8 Forward Calorimetry	14
1.3 Simulation	16
1.4 CLIC_ILD	17
1.5 Particle Flow Reconstruction	18
1.5.1 PandoraPFA	19
1.6 Performance	21
1.6.1 Jet Energy Resolution	22
1.6.2 Decomposition of the Jet Energy Resolution	25
1.6.3 Single Particle Energy Resolution	27
1.7 Summary of ILD Detector Performance	28
2 Energy Estimators	33
2.1 Motivation	33
2.2 Calibration in the Particle Flow Paradigm	36
2.2.1 Overview of the Calibration Procedure	37
2.2.2 MIP Scale Determination in the Digitiser	38
2.2.3 Digitisation Implementation	38
2.2.3.1 ECal Digitisation Implementation	38

2.2.3.2	HCal Digitisation Implementation	40
2.2.3.3	HCal Ring Digitisation Implementation	42
2.2.4	MIP Scale Determination in PandoraPFA	43
2.2.5	Electromagnetic Scale in PandoraPFA	44
2.2.6	Hadronic Scale in PandoraPFA	46
2.2.7	Summary	49
2.3	Novel Energy Estimators	50
2.3.1	HCal Hit Energy Truncation	51
2.3.1.1	Legacy Energy Corrections	52
2.3.1.2	Impact on Single Particle Energy Resolution	52
2.3.1.3	Impact on Jet Energy Resolution	53
2.3.2	Software Compensation	55
2.3.2.1	Impact on Single Particle Energy Resolution	59
2.3.2.2	Impact on Jet Energy Resolution	60
2.3.3	Summary	62
2.4	Timing Cuts	63
2.4.1	Impact on Single Particle Energy Resolution	64
2.4.2	Impact on Jet Energy Resolution	65
2.4.3	Summary	66
	Bibliography	69

“I thank and praise you, God of my ancestors: You have given me wisdom and power”
— Daniel 2:23a

Chapter 1

Particle Flow Calorimetry for Future Linear Colliders

“How much better to get wisdom than gold, to get insight rather than silver!”

— Proverbs 16:16

Particle flow calorimetry can provide extremely good jet energy resolutions at a future linear collider. Jet energy resolution is crucial at the linear collider as many of the interesting processes will be characterised by multi-jet final states. Many of these multi-jet final states will be produced from the hadronic decays of W and Z bosons and one of the key goals of the future linear collider is to be able to separate these decays. However, separation of these decays can be achieved only by placing a tight requirement on the jet energy resolution; $\sigma_E/E \lesssim 3.5\%$ for 50-500 GeV jets at the ILC and up to 1.5 TeV at CLIC [1]. The use of particle flow calorimetry will also be highly beneficial for quantifying final states of interest that contain charged leptons and missing momentum.

1.1 Particle Flow Calorimetry

The premise of particle flow calorimetry is to use the sub-detector that offers the best energy resolution to measure the energy of any given particle, which corresponds to energy measurements being made in the ECal for photons, the HCal for neutral hadrons and, crucially, the tracker for charged particles. The starker contrast of this approach to that

of traditional calorimetry occurs in the measurement of the energy of charged particles. In particle flow calorimetry the energy of a charged particle is measured using the curvature of the path it traverses as it bends in a magnetic field, while in traditional calorimetry the energy would be measured using the calorimeters, predominantly the hadronic calorimeter (HCal). The tracker energy resolution for a single charged particle of energy E_{X^\pm} is typically $10^{-4} \times E_{X^\pm}^2$, while for the HCal it is $\sim 0.55 \times \sqrt{E_{X^\pm}}$ [1]. The energy resolution offered by the tracker is significantly better than that offered by the HCal for energies up to $\sim \mathcal{O}(300 \text{ GeV})$. This means that particle flow calorimetry has the potential to offer a much better energy resolution for charged particles below $\sim \mathcal{O}(300 \text{ GeV})$, than that of the traditional calorimetry approach. Particle flow calorimetry offers gains in performance for collision energies well beyond 300 GeV as the average long-lived particle energy for physics processes of interest is typically much less than 300 GeV. Furthermore, it also leads to a significant improvement in the measurement of jet energies as, after the decay of short-lived particles, approximately 60% of the energy of a jet is carried in the form of charged particles. The measurement of jet energies in the particle flow paradigm is summarised in table 1.1. The benefits to the energy resolution for both charged particles and jets offered by the particle flow approach to calorimetry is the driving factor behind why it is planned for use at the linear collider experiments.

Jet Component	Detector	Energy Fraction	Energy Resolution
Charged Particles (X^\pm)	Tracker	$\sim 0.6E_j$	$10^{-4} \times E_{X^\pm}^2$
Photons (γ)	ECal	$\sim 0.3E_j$	$0.15 \times \sqrt{E_\gamma}$
Neutral Hadrons (X^0)	HCal	$\sim 0.1E_j$	$0.55 \times \sqrt{E_{X^0}}$

Table 1.1: The approximate jet fractions and energy resolutions for charged particles (X^\pm) of energy E_{X^\pm} , photons (γ) of energy E_γ and neutral hadrons (X^0) of energy E_{X^0} . The energy resolution for photons and neutral hadrons reflects the performance of a linear collider like ECal and HCal respectively. Taken from [1].

Particle flow calorimetry is challenging to put into practice as it requires a precise reconstruction for all long-lived particles within a detector. Charged particle energy measurements are made using the curvature of the track they traverse as they bend in the magnetic field, but they also produce calorimetric energy deposits, as shown in figure 1.1. If both of these energy measurements are used, 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 known as "confusion", degrade the energy resolution of a particle flow detector. Therefore, 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 [2]. 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 energy measurements from the optimal sub-detectors.

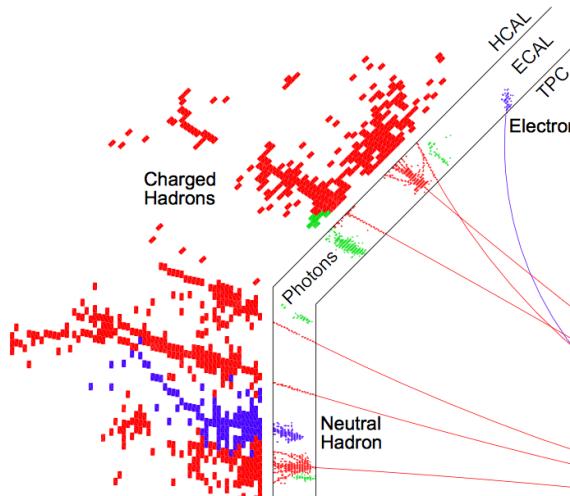


Figure 1.1: A typical simulated 250 GeV jet in the CLIC_ILD detector, with labels identifying constituent particles. Image taken from [3].

1.2 International Large Detector

The current detector concepts for the linear collider experiments have been designed to make particle flow calorimetry possible. While there are a number of different concepts that are under consideration for both the ILC and CLIC, one of the most prominent, and the focus of this work, is the International Large Detector (ILD). The ILD detector, shown in figure 1.2, achieves very high spatial resolution for all sub-detector systems thanks to its highly segmented calorimeters and central tracking system, both of which are encompassed within a 3.5 T magnetic field. PandoraPFA [1,3] provides the sophisticated pattern recognition software that is required for particle flow calorimetry.

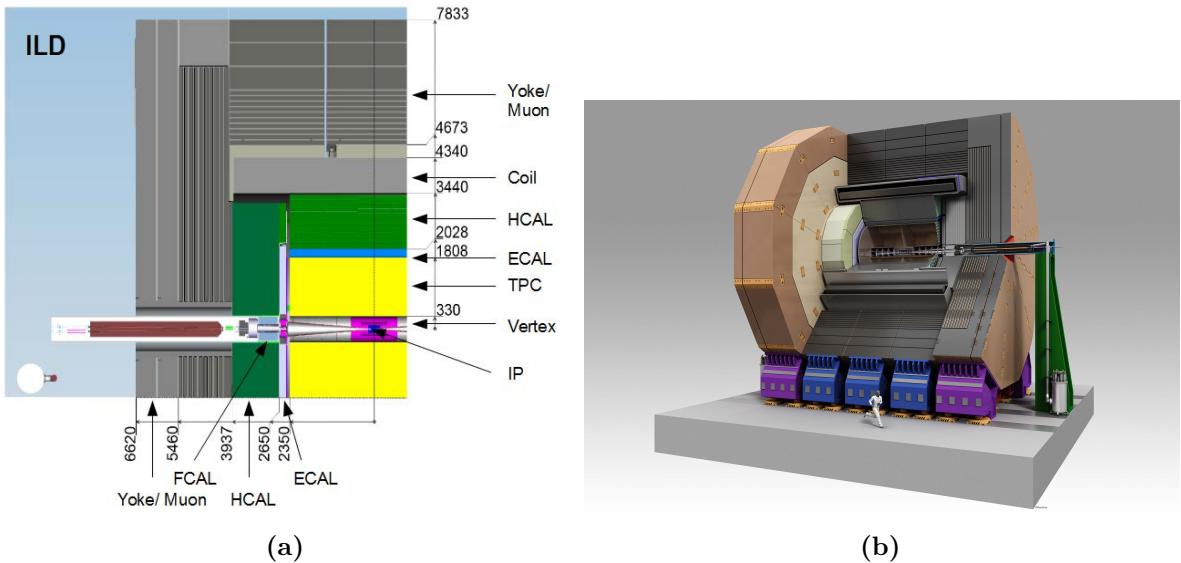


Figure 1.2: (a) Quadrant view of the ILD detector concept. The interaction point is in the lower right corner of the picture. Dimensions are in mm. (b) An artistic view of the ILD detector concept. Figures taken from [4].

1.2.1 Overview

The tracking system for the ILD detector consists of a vertex detector, a Time Projection Chamber (TPC) and a number of supplementary silicon detectors. The vertex detector is designed to give precise information about displaced vertices with respect to the impact point (IP), which is crucial for the study of short lived particles such as the D and B mesons. The vertex detector is located close to the IP and surrounding it is the TPC, which is the central tracker for ILD. The TPC provides detailed measurements of the

trajectory of charged particle tracks passing through it, up to 224 measurements per track. This information is used for determining the curvature of the charged particle track and hence the momentum of the charged particle that traversed it. Finally, the purpose of the supplementary silicon detectors is to provide additional, high precision, spatial measurements to aid track fitting and extend coverage of the detector down to low polar angles.

The calorimetric system for the ILD detector is comprised of an electromagnetic calorimeter (ECal), a hadronic calorimeter (HCal) and a number of forward calorimeters (FCal). The primary function of the ECal is to induce electromagnetic particles to shower within it and to measure the energy of these particle showers. Similarly, the HCal is designed to induce and measure the energy of hadronic particle showers. The ECal surrounds the tracking system in the ILD detector and is itself surrounded by the HCal. The function of FCal is to extend the coverage of the calorimeter system to low polar angles and to provide measurements of the luminosity of the colliding e^\pm beams.

The outermost elements of the ILD detector are the solenoid, iron yoke and muon system. The solenoid generates a magnetic field of 3.5 T, which is essential for determining the energy of charged particles in the particle flow paradigm. The iron yoke is used to return the magnetic field generated by the solenoid. The yoke is instrumented by the muon system to provide additional information, which supplements the calorimetric energy measurements made by the ILD calorimeters.

1.2.2 Vertex Detector

The main goal of the ILD vertex detector is to achieve a resolution on the impact parameter of charged particle tracks of

$$\sigma_b < 5 \oplus \frac{10}{p \sin(\theta)^{3/2}} \mu\text{m}, \quad (1.1)$$

where σ_b is the resolution on the track impact parameter, p is the momentum of the track and θ is the angle between the track and the vertex detector plane. The first term in this parameterisation is the transverse impact parameters resolution and the second is a multiple-scattering term. This makes precisely tagging secondary vertices from charm and bottom mesons possible. Typically these mesons have relatively short proper lifetimes, τ , such that $c\tau \approx \mathcal{O}(300 \mu\text{m})$. To achieve this impact parameter resolution, a spatial resolution of better than 3 μm is required near the impact point (IP). Furthermore, a

low material budget of less than 0.15 % of a radiation length per layer is required to ensure that few electromagnetic showers are initiated within the vertex detector. A low pixel occupancy is essential for determining the trajectory of individual tracks in the detector. Furthermore, consideration will have to be given to the mechanical structure of the detector, power consumption and cooling.

There are a number of different pixel technology options under consideration for the vertex detector for the ILD detector. This is an active area of ongoing research and development for the linear collider collaboration. The current design of the vertex detector consists of three concentric layers of double-sided ladders with the first layer containing 10 ladders, the second 11 ladders and the third 17 ladders as shown in figure 1.3. Every ladder has two silicon pixel sensors on each side and the ladder thickness is approximately 2 mm. The radii covered by the detector range from 16 mm to 60 mm from the IP.

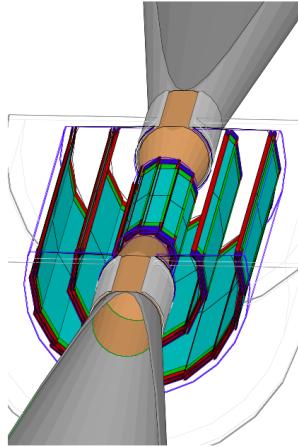


Figure 1.3: Vertex detector design for ILD. Figures taken from [5].

1.2.3 Time Projection Chamber

The central tracking system for the ILD detector is a TPC, which is shown in figure 1.4. The TPC consists of a cylindrical gas volume with a central electrode providing an axial electric field. When a charged particle passes through the TPC, it ionises the gas and the ionised molecules drift in the axial electric field. The direction of the electric field is chosen such that the electrons drift towards the endplates where they are collected. The position of the ionisation point can then be calculated using the drift time of the electrons in the TPC. Combining these TPC hits together makes reconstruction of the

charged particle track possible. TPCs have an advantage over silicon tracking in that they continuously track any charged particle passing through them, while silicon detectors are only sensitive within each silicon layer. This compensates for the poor single point resolution that TPCs have in comparison to silicon detectors and makes TPCs a viable option for the ILD detector. Furthermore, TPCs have a very low material budget. This benefits calorimetry as it minimises energy losses prior to the particle energy entering the calorimeters, which means the calorimetric energy deposits give a better reflection of the true particle energy.

The ILD TPC has a point resolution of better than $100\ \mu\text{m}$ and a double hit resolution in ϕ of less than 2 mm. The gas used for the TPC will be Ar:CH₄:CO₂ (95:3:2) [5]. Several readout technology options designed to measure the ionisation current are currently under development. For all potential options it is envisaged that the readout pads would be $\approx 1 \times 6\text{mm}^2$ giving a total of approximately 10^6 pads on each TPC endplate.

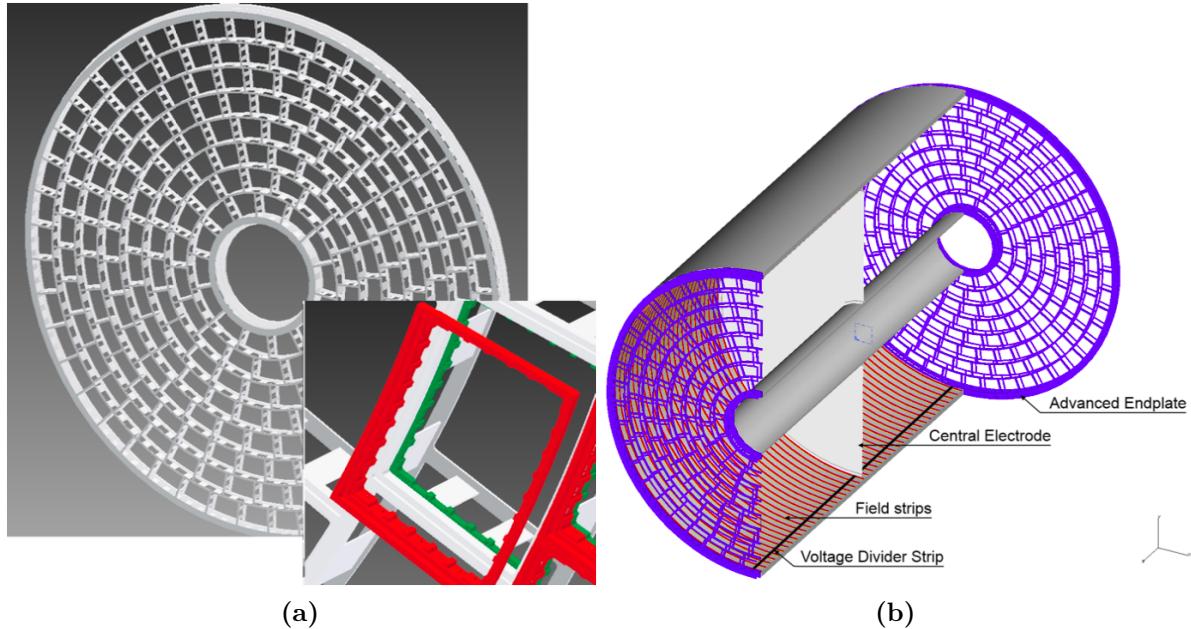


Figure 1.4: (a) Drawing of the proposed end-plate for the TPC. In the insert a back frame, which is designed to support the readout modules, is shown. (b) Conceptual sketch of the TPC system showing the main parts of the TPC (not to scale). The central electrode generates the axial electric field, the endplates collect the ionisation electrons, the field strips help to maintain a uniform electric field across the TPC and the voltage divider strips maintains the voltage difference between the anode and cathode. The field strips are held at fixed voltages such that they replicate the electric field produced by the electrodes. This reinforcing of the electric field configuration minimises non-uniformities in the electric field. The field cage of the TPC is not shown. Figures taken from [4]

1.2.4 Supplemental Silicon Tracking System

There are four components that make up the supplemental silicon tracking system in the ILD detector, shown in figure 1.5, which are:

- Silicon Inner Tracker (SIT) and Silicon External Tracker (SET). These are both barrel components, which are positioned immediately inside and outside the TPC. The SIT helps form associations between hits in the vertex detector and the TPC, while the SET helps with extrapolation of TPC tracks into the calorimeter.
- Endplate of the TPC (ETD). This sensor is identical to the SET, but is positioned in front of the ECal endcap calorimeter. The ETD extends the coverage of the supplemental silicon tracking system envelope.
- Forward tracker (FTD). This detector consists of seven silicon disks that extend the coverage of the tracking down to small angles that are not covered by the TPC.

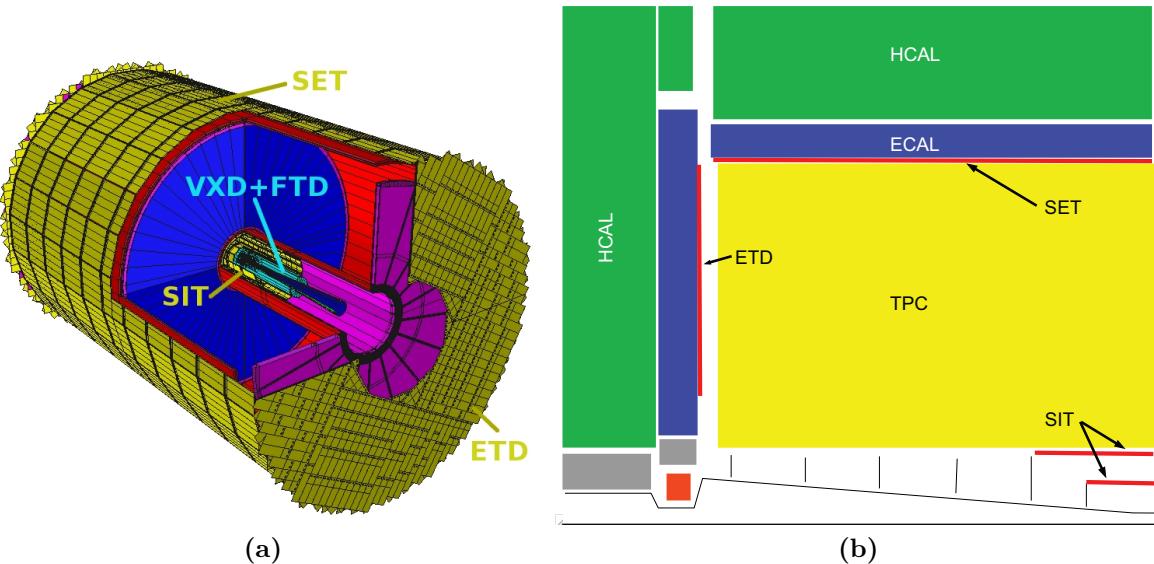


Figure 1.5: (a) A 3D detailed GEANT4 simulation description of the silicon system. (b) A quadrant view of the ILD silicon envelope made of the four components SIT, SET, ETD and FTD as included in the full MOKKA simulation. Figures taken from [4].

The coverage of the SIT, SET, ETD and FTD is given in table 1.2. These detectors are designed to give high precision space points that can be used in track fitting. Furthermore, the ETD and SET are of particular use for extrapolating the charged particle tracks into the calorimeters. This is key for particle flow calorimetry, which relies upon correct

Tracking System	Coverage [$\cos\theta$]
SIT	0.910
SET	0.789
ETD	0.799 - 0.985
FTD	0.802 - 0.996

Table 1.2: Coverage of the supplementary silicon tracking systems in the ILD detector. In this table θ is the polar angle with respect to the beam direction. Taken from [4].

association of charged particle tracks and clusters of calorimeter hits. Analogously to the vertex detector, these detectors require low material budget and low occupancy. The FTD, due to its proximity to the beam axis, is particularly prone to high occupancies.

The SIT, SET and ETD are silicon pixel sensors with $50 \mu\text{m}$ pitch embedded in $200 \mu\text{m}$ thick silicon. The FTD consists of seven silicon tracking disks, the first two being pixel detectors and the remaining five being strip detectors. The pixel detector disks are formed of 16 petals, as shown in figure 1.6. Within these petals the pixel size varies from $26 \times 29 \mu\text{m}^2$ to $26 \times 67 \mu\text{m}^2$. Strip detectors are used for the outermost tracking disks as the occupancy considerations do not demand a high granularity detector i.e. a pixel detector. These detector disks will have a pitch of $50 \mu\text{m}$. The active sensor and readout ASIC design for each of these detectors is an active area of development for the linear collider.

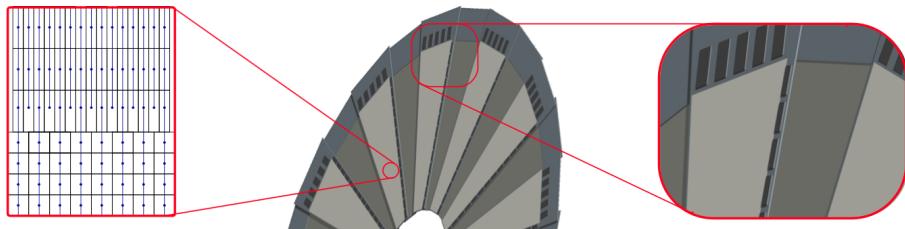


Figure 1.6: A half-disk for the FTD showing the petal concept. The rightmost zoom image shows a detail of the end-of-petal area that houses the read-out electronics. The leftmost image shows the region at $R = 8 \text{ cm}$ where both the column width and the R -dimension of the pixels changes. Figures taken from [4].

1.2.5 Electromagnetic Calorimeter

The nominal ILD detector contains a finely segmented electromagnetic sampling calorimeter (ECal). The ILD ECal has been specifically designed with particle flow calorimetry in mind. To that extent the spatial resolution of particle showers within the ECal takes as much, if not more, precedence than the energy resolution.

There are a number of design requirements for the ECal:

- The ECal must be compact in size to reduce the overall cost of the detector.
- Fine segmentation of the ECal is required so that nearby particle showers can be separated. This is an essential requirement for particle flow calorimetry.
- Electromagnetic showers should be contained within the ECal.

Based on these requirements tungsten is used as the absorber material for the ILD ECal as it has a small radiation length (X_0), a small Moli  re radius and a large ratio of radiation length to nuclear interaction length. A comparison of these properties for other ECal absorber material candidates is shown in table 1.3. The small radiation length in tungsten allows for a large number of radiation lengths, $\approx 24 X_0$, to be compacted within a relatively short distance, ≈ 20 cm, in nominal ILD ECal. This is sufficient for containing all but the highest energy electromagnetic showers. The small Moli  re radius in tungsten will lead to compact electromagnetic showers. This makes separation of nearby showers easier. Finally, the large ratio of the radiation length to the nuclear interaction length in tungsten will lead to greater longitudinal separation between electromagnetic and hadronic showers, again making shower identification easier.

Material	λ_I (cm)	X_0 (cm)	ρ_M (cm)	λ_I/X_0
Fe	16.8	1.76	1.69	9.5
Cu	15.1	1.43	1.52	10.6
W	9.6	0.35	0.93	27.4
Pb	17.1	0.56	1.00	30.5

Table 1.3: Comparison of the nuclear interaction length λ_I , radiation length X_0 and Moli  re radius for iron, copper, tungsten and lead. Table taken from [1].

The active material in the nominal ILD ECal is silicon, however, a scintillator strip option is also being considered. Figure 1.7 shows a cross section through a layer of the

ECal for both of these options. It contains a total of 30 longitudinal readout layers, which is sufficient to provide a good energy resolution. The tungsten thickness for the innermost 20 layers is 2.1 mm, while for the final 10 layers it is 4.2 mm. This configuration of absorber material thickness is chosen to reduce the number of readout channels and hence the cost, while maintaining a high sampling rate for particle showers at the start of the ECal. It should be noted that this ECal offers no gains in terms of energy resolutions in comparison to pre-existing particle collider experiments, as shown in table 1.4. This is the case because the focus of this calorimeter is split between imaging the particle showers and recording their energy as opposed to purely focusing on the energy measurement. Each of the ECal layers is divided up into square cells, of 5 mm side length, which makes separation of nearby particle showers possible. This cell size was chosen as a balance between being able to resolve nearby particle showers and reducing the overall cost of the calorimeter, which scales with the number of readout channels. An optimisation study of the various ECal parameters for the ILD detector can be found in section ??.

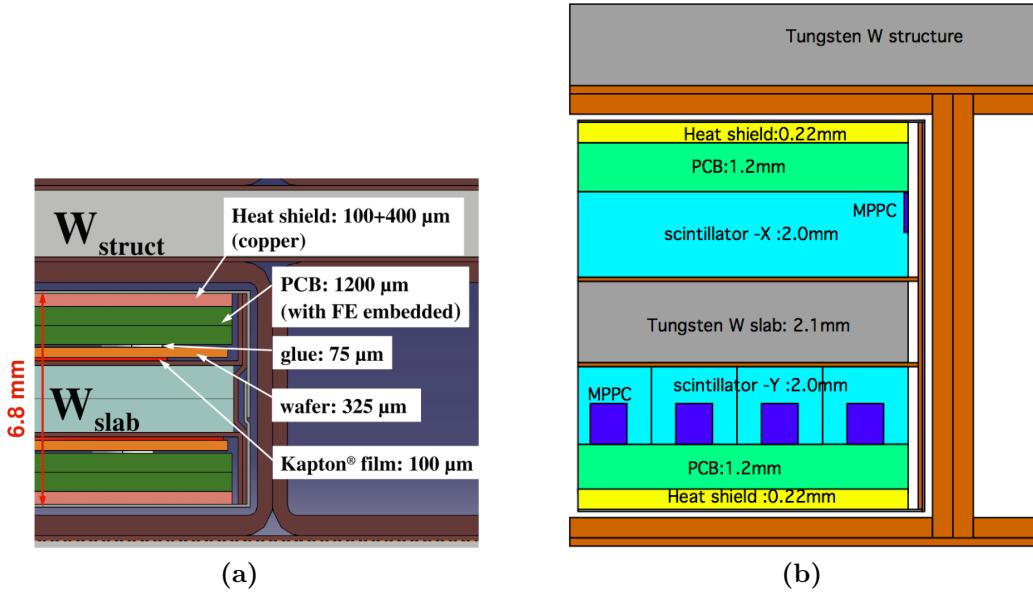


Figure 1.7: Cross section through ECal layer for (a) silicon and (b) scintillator option. Figures taken from [4].

1.2.6 Hadronic Calorimeter

A finely segmented hadronic sampling calorimeter (HCal) is used in the nominal ILD detector. The design requirements for the ILD HCal mirror those of the ECal, which can

Experiment	ECal Energy Resolution $\frac{\sigma_E}{E}$
CMS [6]	$\sim \frac{2.8\%}{\sqrt{E(\text{GeV})}} \oplus 0.3\% \oplus \frac{12\%}{E(\text{GeV})}$
ATLAS [7]	$\sim \frac{10.1\%}{\sqrt{E(\text{GeV})}} \oplus 0.1\%$
LHCb [8]	$\sim \frac{9\%}{\sqrt{E(\text{GeV})}} \oplus 0.8\%$
ILC (ILD Silicon Option) [4]	$\sim \frac{16.6\%}{\sqrt{E(\text{GeV})}} \oplus 1.1\%$

Table 1.4: Comparison of the ECal energy resolutions for various experiments.

be found in section 1.2.5, with one exception; the HCal is designed to contain hadronic showers as opposed to electromagnetic showers. Steel is used as the absorber material for the HCal as it has durable mechanical properties that allow the HCal to be constructed without the need for auxiliary supports. When required, auxiliary supports create dead regions in the detector that would harm performance. Furthermore, steel is relatively inexpensive and has a small nuclear interaction length, meaning it is possible to achieve a compact calorimeter design at low cost. The nominal ILD HCal contains approximately $6\lambda_I$, which when combined with the $1\lambda_I$ in the ECal is enough to contain the majority of hadronic showers at ILC like energies.

The active material in the nominal ILD HCal is scintillator. In total, the HCal contains 48 readout layers, which provides an extremely good energy resolution. This can be seen when comparing the HCal energy resolution between different experiments, as shown in table 1.5. An individual layer in the HCal is comprised of 20 mm of steel absorber material with 3 mm of scintillator active material. Each layer in the HCal is segmented into square cells of side length 30 mm. This cell size was chosen as a balance between reducing the cost of the detector, which is proportional to the number of readout channels, and achieving the required spatial resolution to make particle flow calorimetry possible. The segmentation of the ILD HCal gives excellent spatial resolution and sufficiently good energy resolution to make the use of particle flow calorimetry a reality. An optimisation study of the various HCal parameters for the ILD detector can be found in section ??.

The ILD HCal is intrinsically non-compensating, which means that it has a different response to electromagnetic and hadronic showers. The origin of this different response is the fundamentally different mechanisms governing the propagation of electromagnetic and hadronic showers. One key difference between the mechanisms is that hadronic showers have an invisible energy component, which occurs due to effects such as neutrons

coming to rest in the detector and nuclear biding energy losses [9]. In general, this leads to a lower response from a calorimeter to a hadronic shower than an electromagnetic shower. A number of different software techniques have been developed for the linear collider experiment that attempt to correct this non-compensating response. For more details see chapter 2. The ILD ECal has a compensating response due to the use of tungsten as the absorber material [10], therefore, no additional treatment of energies is required.

Experiment	HCal Energy Resolution $\frac{\sigma_E}{E}$
CMS [11]	$\sim \frac{90\%}{\sqrt{E(\text{GeV})}} \oplus 4.8\%$
ATLAS [12]	$\sim \frac{52.1\%}{\sqrt{E(\text{GeV})}} \oplus 3.0\% \oplus \frac{1.6\%}{E(\text{GeV})}$
LHCb [8]	$\sim \frac{69\%}{\sqrt{E(\text{GeV})}} \oplus 9.0\%$
ILC (ILD Silicon Option) [4]	$\sim \frac{43.3\%}{\sqrt{E(\text{GeV})}} \oplus 1.8\%$

Table 1.5: Comparison of the HCal energy resolutions for various experiments.

1.2.7 Solenoid, Yoke and Muon System

Surrounding the ILD calorimeter system is the solenoid that generates a 3.5 T magnetic field. The magnetic field produced by the coil is crucial for bending charged particles so that their momentum can be determined from the curvature of the path they traverse. Furthermore, the bending of charged particles leads to greater separation of calorimetric energy deposits between charged and neutral particles, which will reduce the effects of confusion when using particle flow calorimetry.

The magnetic field in the ILD detector is returned by an iron yoke that surrounds the solenoid. Iron is chosen for the yoke material as it has a very large permeability. The yoke is instrumented by a muon system in the barrel and forward regions of the detector. The goal of this instrumentation is to identify muons escaping the calorimeters and to act as a tail catcher for the calorimeters. The muon system consists of 10 layers, spaced 140 mm apart, followed by 2 and 3 layers spaced 600 mm apart in the barrel and endcap regions of the detector respectively, as shown in figure 1.8. There is also an additional sensitive layer for the barrel region placed immediately outside the HCal to help with association energy deposits between the calorimeters and the yoke. As the majority of particles at ILC like energies will be contained within the calorimeters, the

energy and spatial resolution of the muon system are not critical to performance. It is for that reason that the number of layers is lower and the layer thicknesses wider in the yoke than in the calorimeters. The nominal ILD model uses 30 mm wide and 1 m long scintillator strips as the readout technology for the yoke.

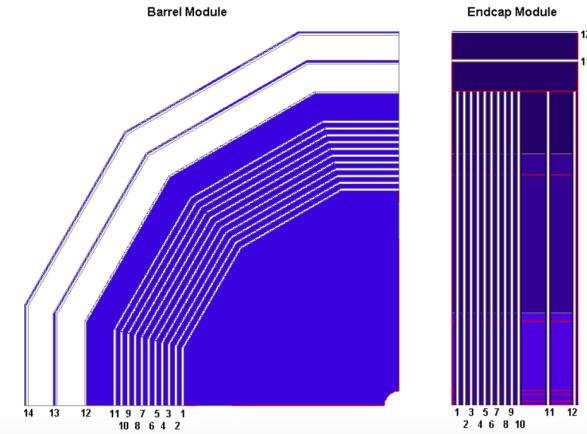


Figure 1.8: The sensitive layers of the ILD muon system. Figure taken from [4].

1.2.8 Forward Calorimetry

Forward calorimetry in the ILD detector consists of three additional sampling calorimeters:

- The LumiCal, which is located within the octagonal hole in the ECal endcap. This will give a precise measurement of the luminosity of the linear collider beam. The LumiCal uses Bhabha scattering, $e^+e^- \rightarrow e^+e^-(\gamma)$, as a gauge process for the luminosity measurement. Using this approach the luminosity can be measured with precision of less than 10^{-3} at $\sqrt{s} = 500$ GeV [5].
- The LHCAL, which is positioned within the square hole of the HCal endcap. This hadronic calorimeter is designed to extend the coverage of the HCal down to small polar angles.
- The BeamCal, which is located just in front of the final focusing quadrupole. This calorimeter will perform a bunch-by-bunch estimate of the luminosity based on the energy deposited in the calorimeter.

The layout of these calorimeters is shown in figure 1.9 and their coverage is summarised in table 1.6. Each of the forward calorimeters will have to deal with high occupancies due to the presence of background processes, e.g. beamstrahlung, which makes fast readout

crucial. Furthermore, the BeamCal experiences a large flux of low energy electrons due to its proximity to the beam pipe, which results in a large radiation dose. This makes radiation hard sensors essential for the BeamCal.

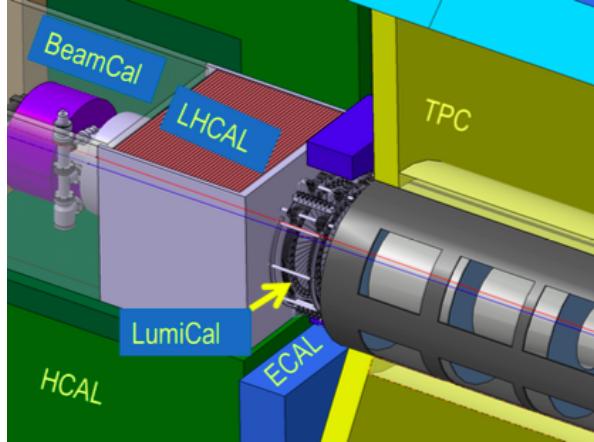


Figure 1.9: The very forward region of the ILD detector. LumiCal, BeamCal and LHCAL are carried by the support tube for the final focusing quadrupole, QD0, and the beam pipe. Figure taken from [4].

Forward Calorimeter	Polar Angle Coverage [mrad]
LumiCal	31 – 77
LHCAL	~29 – 122
BeamCal	5 – 40

Table 1.6: Coverage of the forward calorimeters in the ILD detector.

Each of these forward calorimeters is constructed using tungsten as the absorber material. The small Moli  re radius of tungsten ensures that narrow electromagnetic showers are formed within them, which makes separation and identification of showering particles easier.

The layout of these calorimeters is as follows:

- The LumiCal is a silicon tungsten sampling calorimeter that contains 30 readout layers. This gives the LumiCal a total depth of $\approx 24X_0$.
- The LHCAL is also a silicon tungsten sampling calorimeter, which contains 40 readout layers. The total depth of the LHCAL is $\approx 4\lambda_I$.
- The BeamCal is a tungsten based sampling calorimeter. The sensitive detector material for the BeamCal is an ongoing area of research as, due to the extremely high

occupancy from the beam induced backgrounds, a very fast readout is required. The exact layer configuration of the BeamCal will depend upon the choice of sensitive detector material and hence is yet to be specified.

The segmentation within the layers, i.e. the cell size, in these forward calorimeters is yet to be fully optimised.

1.3 Simulation

Detector model simulation for all studies presented in this work was performed using MOKKA [13], a GEANT4 [14, 15] wrapper providing detailed geometric descriptions of detector concepts for the linear collider. The MOKKA simulation of the ILD detector includes the following [4]:

- The vertex detector is simulated using silicon as the sensitive material. Support material and the cryostat are also included.
- The supplementary silicon tracking systems are included. Again, material has been added to the simulation to represent the support material for these systems. Furthermore, an estimation has been made of the material budget for power and readout cables from the vertex detector, SIT and FTD and material has been added to the simulation to represent these. The material added to represent the power and readout cables comes in the form of an aluminium cylinder running inside the TPC field cage and a cone around the beam pipe.
- The TPC is simulated as a cylindrical volume of a gas mixture surrounded by a field cage. A conservative estimate of the endplate is included in the simulation to account for the support structure, electronics and cooling pipes for the TPC.
- As well as including the silicon tungsten sampling calorimeter, the simulation of the ILD ECal contains additional material to represent the instrumented region of the sensor and a heat shield as shown in figure 1.7.
- Simulation of the ILD HCal has a number of realistic features including detailed modelling of the electronics, detector gaps and the implementation of Birk's law [16] for the scintillator sensitive detector elements.
- The muon system, which is the instrumentation of the iron yoke, uses scintillator as the active material in the simulation. A square cell size of side length 30 mm is

assumed. This is in contrast to the nominal ILD model, but as the tail-catcher plays a minimal role in event reconstruction at ILC like energies this difference should have negligible impact.

- The forward calorimeters, the LumiCal, LHCAL and BeamCal, are all included in the simulation. Tungsten is used as the absorber material for each of the calorimeters. The LumiCal and LHCAL use a silicon readout material, while the BeamCal uses a diamond readout.

The simulation and reconstruction of the large event samples used in the studies presented in the work was performed using the ILCDIRAC [17, 18] grid production tools.

1.4 CLIC_ILD

The increased collision energy of the proposed CLIC accelerator means the use of the nominal ILD detector model would be inappropriate. Therefore, a new detector model, CLIC_ILD [19, 20], based upon the nominal ILD detector model was created to cope with the experimental conditions found at the CLIC experiment. The main differences between the nominal ILD detector and CLIC_ILD are:

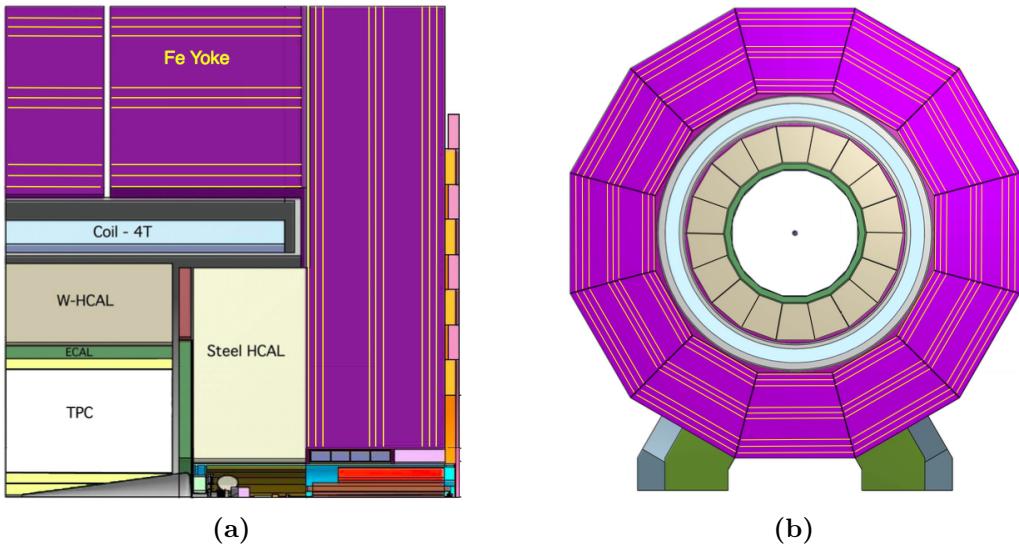


Figure 1.10: (a) Longitudinal (top quadrant) and (b) transverse cross section of the CLIC_ILD detector. Figures taken from [19].

- The higher energies found at the CLIC experiment lead to more intense beam induced backgrounds, which is especially problematic for detectors close to the IP

where the occupancies will be extremely high. For this reason the inner vertex detector in CLIC_ILD is moved 15 mm further out from the IP.

- The HCal thickness is increased from $6 \lambda_I$ to $7.5 \lambda_I$. This ensures that higher energy hadronic showers found at the CLIC experiment are contained within the calorimeters.
- The HCal absorber material for the barrel is tungsten as opposed to steel. This reduces the overall thickness of the HCal and keeps the coil size, one of the driving cost factors for the detectors, similar for the nominal ILD and CLIC_ILD detectors. Steel is used as the absorber material for the HCal endcaps as there are no spatial requirements relating to the coil size and this will lower the detector cost. Furthermore, the shower development time in steel is faster than in tungsten. This makes effective time stamping of energy deposits easier, which is crucial for the CLIC experiment for vetoing beam induced backgrounds.
- The magnetic field strength in the CLIC_ILD detector is increased to 4 T. This was found to benefit the reconstruction, particularly at high energies, as it leads to greater separation of charged particle tracks. Furthermore, it was possible to achieve this increase in field strength using the nominal ILD coil design.
- The CLIC_ILD detector contains masking, graphite layers placed in front of the BeamCal, to prevent particles produced by the beam-induced interactions from backscattering into the main detector. It is the increased collision energy that makes backscattering of particles a more problematic effect for the CLIC experiment than it is for the ILC experiment.

The CLIC_ILD detector will be used for the study of anomalous gauge couplings presented in chapter ??.

1.5 Particle Flow Reconstruction

Particle flow calorimetry relies upon correct associations being made between calorimetric energy deposits and charged particle tracks. Even with a finely segmented detector, such as the ILD detector described in section 1.2, correctly making these associations is a highly non-trivial task and must be done using advanced pattern recognition software. This is provided by the PandoraPFA particle flow algorithm [2]. PandoraPFA is applied

in the linear collider reconstruction using MARLIN [21], a c++ framework specifically designed for the linear collider.

1.5.1 PandoraPFA

PandoraPFA’s inputs are calorimeter hits and charged particles tracks and it produces as output reconstructed particles known as particle flow objects (PFOs). The pattern recognition in PandoraPFA is applied in eight main stages as described in the original PandoraPFA paper [1] and the CLIC focused PandoraPFA paper [3]:

1. Track selection. The input track collections are examined to determine whether V^0 decays or kinks are present. Two charged tracks originating from a point displaced from the IP indicates the presence of a V^0 decay. A charged particle decaying into a single charged particle and a number of neutral ones indicates a kink. This information will be propagated in the reconstruction to the final PFO creation stage.
2. Calorimeter hit treatment. The treatment of calorimeter hits by PandoraPFA is of paramount importance to the work presented in chapters 2 and ???. Therefore, full details of the calorimeter hit selection procedure are presented here. This selection procedure is broken down into several steps:
 - The various collection of, post digitisation, calorimeter hits are passed into the Pandora framework and converted into Pandora calorimeter hits.
 - To minimise any dependancy on the detector geometry each calorimeter hit is assigned to a pseudo-layer, which is representative of the hits position in the calorimeter. All further topological association algorithms work using the pseudo-layer definition, illustrated in figure 1.11.
 - A minimum ionising particle equivalent energy cut is applied to the calorimeter hits. If a calorimeter hit contains less than 0.5 (0.3) of the energy of a normally incident MIP passing through the calorimeter cell in the ECal (HCal) then it is not used in the reconstruction.
 - If a calorimeter hit is sufficiently far away from other hits it is flagged as an isolated hit. Such hits are most likely due to low energy neutrons produced in hadronic showers which can travel a significant distance from the original shower before depositing energy. Due to the distance they travel, these hits are very difficult to associate to the correct particle shower. Furthermore, as such

hits are unlikely to be the seed for a particle shower, they are not used by the initial clustering algorithm.

- Any calorimeter hit that contains an energy consistent with a MIP signal and where one Pandora calorimeter hit at most exists in the neighbouring hits within the same layer is flagged as a MIP consistent hit. This information is used in the identification of MIPs in the reconstruction.
 - The energy contribution for each calorimeter hit ultimately depends on whether the cluster the calorimeter hit has been associated to is deemed to have originated from an electromagnetic or hadronic particle shower. Different scale factors are applied to the energy for electromagnetic and hadronic showers to account for the non-compensating response of the calorimeters. These scale factors are used throughout the reconstruction, including the final reconstructed particle energy, once the particle shower type has been identified. For energy comparisons prior to the shower type being identified, the uncorrected calorimeter hit energy is used. Further details on how these calibration constants are determined can be found in chapter 2.
3. Clustering. This begins by using the projection of the charged particle tracks onto the front face of the ECal as seeds for the initial clustering phase. Calorimeter hits are looped over on a per layer basis, working from the inner to the outer pseudo-layer, and if they fall within a cone of fixed dimensions surrounding a cluster direction they are associated to the cluster. If no association can be made to any pre-existing calorimeter hit clusters then the calorimeter hit is used to seed a new cluster.
 4. Topological cluster merging. The initial clustering algorithm is designed to be conservative to avoid mixing together energy deposits from several particles. The fragments produced by the initial clustering are then merged together by various algorithms whose logic is determined by a number of well-motivated topological rules, such as those shown figure 1.12.
 5. Statistical re-clustering. Comparisons between the cluster energy and any associated track momenta are made to determine whether they are consistent. If a large discrepancy is observed then statistical re-clustering is initiated. This involves running a number of differently configured algorithms to change the cluster configuration to determine if a new optimal configuration of tracks and clusters can be found.

This step relies upon the reported cluster energies being accurate. To ensure this is the case, a well defined calibration procedure is applied for all detector models considered in this work, for more details see chapter 2. At this point in the reconstruction, the energy resolution of the calorimeters impacts the way that the pattern recognition is performed. The better the energy resolution of the calorimeters, the fewer the number of mistakes that are made when pairing up clusters of calorimeter hits to charged particle tracks.

6. Photon identification and recovery. Topological likelihood data is used to identify clusters of calorimeter hits that are consistent with photons. This is possible due to the clear transverse and longitudinal profiles observed for electromagnetic showers.
7. Fragment removal. Neutral clusters originating from a nearby charged particle cluster are identified and merged back into the parent charged particle cluster. These algorithms take into account the changes in the compatibility of the track and cluster associations when merging any neutral clusters into charged clusters.
8. Formation of particle flow objects. Finally, reconstructed particles are produced. The energy for charged particles is taken from the track momenta, while neutral particle energies are taken from the calorimeter cluster measurements. Furthermore, the different electromagnetic and hadronic scales are applied to the output neutral particle energies depending on whether the neutral cluster is consistent with a photon.

The application of the pattern recognition algorithms in PandoraPFA when combined with a highly segmented detector make particle flow calorimetry a reality. In turn this provides excellent jet energy resolution for studying many interesting physics processes at the linear collider experiments.

1.6 Performance

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 the traditional calorimetric approach.

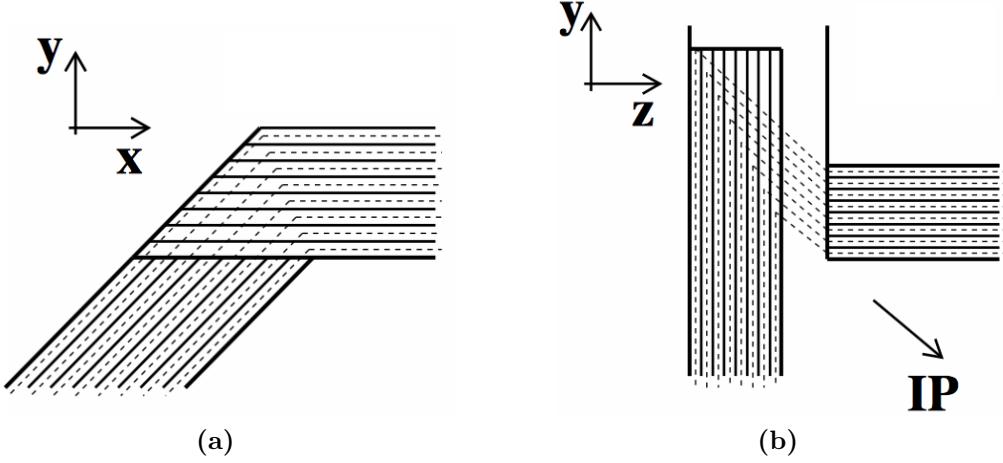


Figure 1.11: Schematic showing the definition of the pseudo-layer assignment for calorimeter hits. The solid lines indicate the positions of the physics ECal layers and the dashed lines show the definition of the virtual pseudo-layers. (a) The xy -view shows the ILD ECal stave structure. (b) The xz view shows a possible layout for the ECal barrel/endcap overlap region. The pseudo-layers are defined using projection back to the IP. Figures taken from [1].

As many physics processes of interest at the linear collider involve multi-jet final states [22], good jet energy resolution is a crucial aspect of detector performance. As shown in chapter ??, the sensitivity of the linear collider experiments to areas of new physics can be 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, a key metric for describing detector performance is the jet energy resolution. Jet energy resolution in particular can benefit from the application of particle flow calorimetry because $\approx 70\%$ of the energy of jets is carried in the form of charged particles. As particle flow calorimetry 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.

1.6.1 Jet Energy Resolution

The jet energy resolution in these studies was determined through the simulation of off-mass shell Z boson events decaying to light quarks (u, d, s). PYTHIA version 6.4 [23], which had been trained on fragmentation data from the OPAL experiment [24], was used to generate these events. The decay of tau leptons appearing in the events was simulated

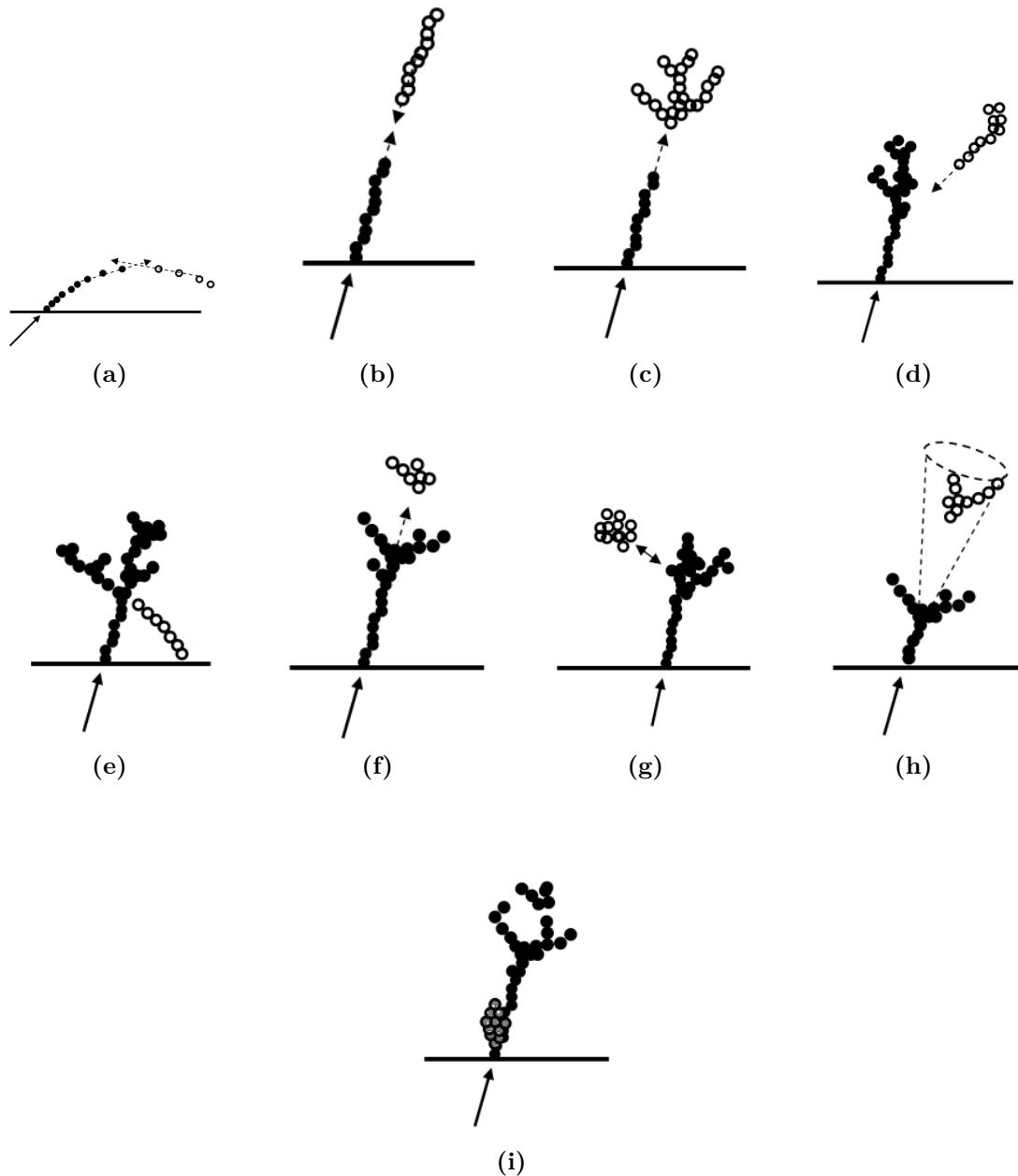


Figure 1.12: The main topological rules for cluster merging: (a) looping track segments; (b) track segments with gaps; (c) track segments pointing to hadronic showers; (d) track-like neutral clusters pointing back to a hadronic shower; (e) back-scattered tracks from hadronic showers; (f) neutral clusters which are close to a charged cluster; (g) a neutral cluster near a charged cluster; (h) cone association; and (i) recovery of photons which overlap with a track segment. In each case the arrow indicates the track, the filled points represent the hits in the associated cluster and the open points represent hits in the neutral cluster. Figures taken from [1].

using TAUOLA [25]. Detector simulation and event reconstruction was carried out as described in sections 1.3 and 1.5 respectively.

As the Z boson in these events is produced at rest, the typical decays form two mono-energetic jets that are produced back-to-back as shown in figure 1.13. Only events where $|\cos(\theta)| < 0.7$, where θ is the polar angle of the quarks, are used in the jet energy resolution calculation. This ensures that little energy is lost down the beam axis. Using these events, the jet energy resolution was calculated as follows:

$$\frac{\text{RMS}_{90}(E_j)}{\text{Mean}_{90}(E_j)} = \frac{\text{RMS}_{90}(E_{jj})}{\text{Mean}_{90}(E_{jj})} \times \sqrt{2}, \quad (1.2)$$

where E_{jj} is the total reconstructed energy. The variables $\text{Mean}_{90}(E_{jj})$ and $\text{RMS}_{90}(E_{jj})$ are the mean and root mean squared (RMS) of the E_{jj} distribution respectively. They 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 [1]. 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. When the distribution of reconstructed jet energies is truncated to the narrowest range that contains 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 1.14. In this example $\text{RMS}(E_{jj})$, the RMS calculated using the full range, is 5.8 GeV, while $\text{RMS}_{90}(E_{jj})$, the RMS using the reduced range, is 4.1 GeV. This corresponds to a reduction in the jet 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.

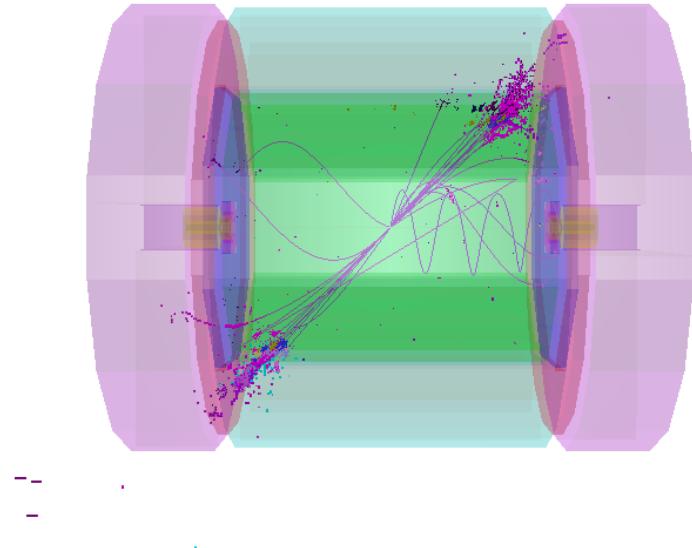


Figure 1.13: 500 GeV di-jet $Z \rightarrow u\bar{d}s$ event display for nominal ILD detector.

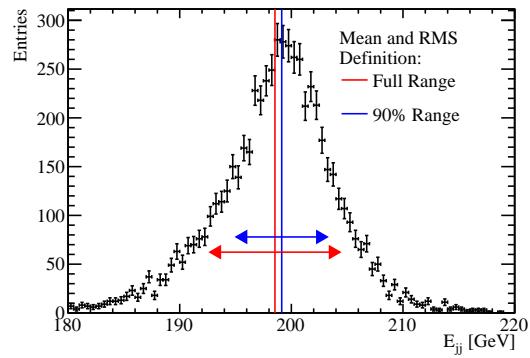


Figure 1.14: Definition of jet energy resolution. Reconstructed jet energy for 200 GeV di-jet $Z \rightarrow u\bar{d}s$ events for nominal ILD detector. The solid vertical line shows the mean of the distribution and the horizontal arrows indicate the mean \pm the root mean square (RMS) of the distribution. The red and blue lines show the mean and RMS calculated using the full range and 90% of the full range with the smallest RMS respectively.

1.6.2 Decomposition of the Jet Energy Resolution

It is possible to gain additional insight into the detector performance by cheating the pattern recognition. Cheating the pattern recognition involves using Monte-Carlo (MC) information to correctly cluster calorimeter hits together and associate them to charged

particle tracks. This has the effect of removing confusion from the reconstruction. By comparing the detector performance obtained from the standard and cheated reconstructions, it is possible to decompose the detector performance 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 1.7.

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 1.7: 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 and are calculated using the range of jet energies with the smallest RMS containing at least 90% of the data.

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 1. 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.

1.6.3 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 photon energy resolutions in the study of anomalous triple and quartic gauge couplings [26–28]. Therefore, photon and K_L^0 energy resolutions, alongside the jet energy resolution, will be considered in these optimisation studies. As both photon and K_L^0 are uncharged, their energy measurements will be made using the calorimeters as opposed to the tracker. Photons 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 photons 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 photon energy resolutions were determined using events containing a single 100 GeV photon, 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 effect of energy leakage from the ECal to the HCal for the photons and leakage of energy out of the rear of the HCal for the K_L^0 events.

The energy resolution for these single particle samples is determined using a Gaussian fit to the reconstructed energy distributions. To aid convergence, 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 photons and 50 GeV K_L^0 s, alongside the

Gaussian fits used to determine the energy resolutions, are shown in figure 1.15. The errors quoted on single particle energy resolutions are determined by propagating the errors reported from the Gaussian fit into the resolution calculation.

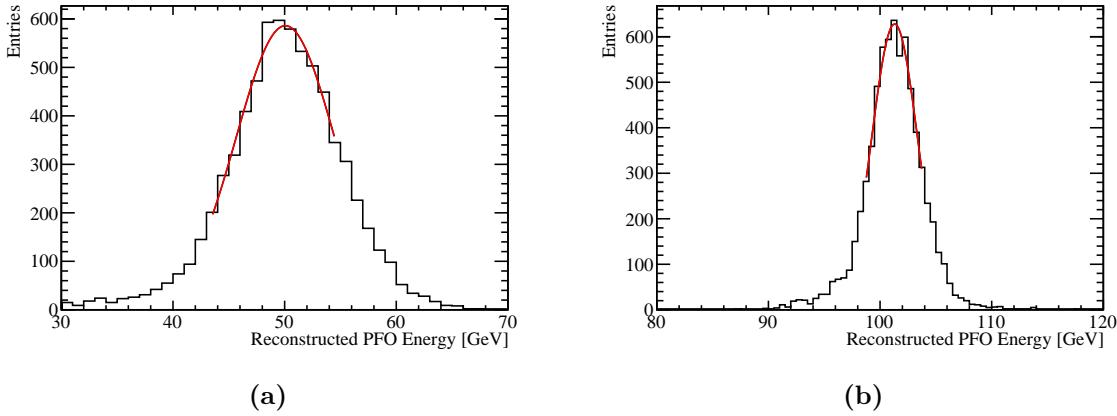


Figure 1.15: The reconstructed energy distribution for (a) 50 GeV K_L^0 and (b) 100 GeV photons. The red line shows a Gaussian fit used to parameterise the detector performance. The fit was applied to the truncated range of the reconstructed PFO energy distribution containing at least 75% of the data with the narrowest RMS. The nominal ILD model was used in this simulation.

1.7 Summary of ILD Detector Performance

The following section outlines the nominal ILD detector performance using the metrics outlined in section 1.6.

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 possibly 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 2. Finally, the energy of the entire particle shower is estimated by grouping together calorimeter hits and summing their energy. As each calorimeter hit's energy is an independent random measurement the particle shower energy will, by the central limit theorem, have a Gaussian distribution.

The energy of a calorimeter hit is obtained by counting the number of charged particle tracks or photons found in the active material of the calorimeter, therefore, Poisson statistics govern 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 that are measured in the calorimeters, the standard deviation of that distribution is $\sigma = \sqrt{\lambda} = \sqrt{N}$ and the energy resolution is $\sigma/\lambda = 1/\sqrt{N}$. As the total shower energy, E_{Reco} , is proportional to N , the energy resolution for a particle shower in an ideal calorimeter is $\sigma_{Reco}/E_{Reco} = a/\sqrt{E_{Reco}}$. In reality, it is typical to express the energy resolution of a calorimeter in the following form

$$\frac{\sigma_{Reco}}{E_{Reco}} = \frac{a}{\sqrt{E_{Reco}}} \oplus b \oplus \frac{c}{E_{Reco}}, \quad (1.3)$$

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 [29]. Here, \oplus denotes the quadrature sum of variables.

Prototypes of the various ILD calorimeter options have been constructed and validated using test beam measurements. The energy resolution of the ILD ECal, determined from test beam measurements, was parameterised as $16.6\%/\sqrt{E_{Reco}(\text{GeV})} \oplus 1.1\%$ for the silicon option and $12.9\%/\sqrt{E_{Reco}(\text{GeV})} \oplus 1.2\%$ for the scintillator option [4]. 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 ≈ 40 GeV. This parametrisation is compared to the full ILD detector simulation in figures 1.16a and 1.16b 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, ≈ 500 GeV, the ECal is no longer sufficient to fully contain the photons and so leakage of energy into the HCal leads to a minor degradation in the simulated 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, which is most likely due to an imperfect implementation of the scintillator ECal within the full detector simulation. The photon energy resolutions seen in the full ILD simulation are similar for the silicon and scintillator ECal options.

Similarly, the energy resolution, determined from test beam measurements, for the nominal ILD HCal was parameterised as $57.6\%/\sqrt{E_{Reco}(\text{GeV})} \oplus 1.6\%$ [30]. A comparison between this test beam parameterisation and the full ILD simulation, using the silicon ECal option, is shown in figure 1.16c. The test beam measurements were made using π^\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. The deviation between the test beam parameterisation and the full ILD simulation, which grows 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. In the test beam studies, to minimise the effect of leakage, events were only considered if the particle showers started developing at the front of the HCal. In the full simulation studies, all particle showers were used, which means some energy will have leaked out of the back of the calorimeters and been deposited in the uninstrumented solenoid region of the detector, resulting in a degradation in the energy resolution.

Figure 1.16d shows the jet energy resolution as a function of jet energy for the full ILD simulation. 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 because the intrinsic energy resolution of the calorimeters is approximately proportional to $1/\sqrt{E_{Reco}}$. On the other hand, confusion grows with energy because increasing energy leads to more dense event topologies, which makes pattern recognition more challenging. The total jet energy resolution for the ILD detector are sufficiently small, $\sigma_{E_j}/E_j \lesssim 3.8\%$ [1, 4, 19], 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.

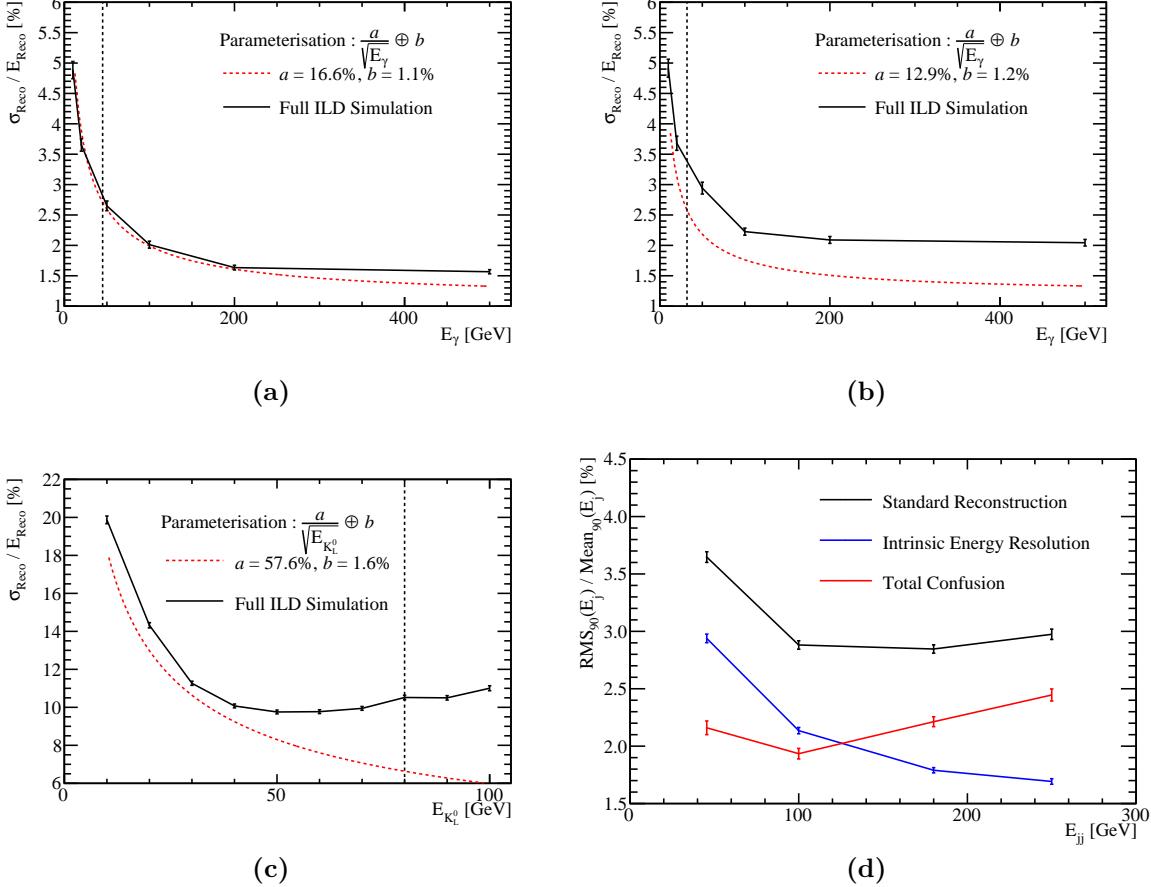


Figure 1.16: (a) The energy resolution as a function of photon energy for the silicon ECal option. The black markers indicate the energy resolutions for the full ILD simulation and the red dotted line shows the test beam parameterisation of the ECal energy resolution. (b) The energy resolution as a function of photon 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 π^\pm s. (d) The jet energy resolution (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.

Chapter 2

Energy Estimators

“He reveals the deep things of darkness and brings utter darkness into the light.”

— Job 12:22

2.1 Motivation

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

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

summing their energy

$$E_{Cluster} = \sum_{ECal\ hits,\ i} E_{ECal}^i + \sum_{HCal\ hits,\ i} E_{HCal}^i , \quad (2.1)$$

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

The linear collider detector concepts employ highly-granular sampling calorimeters [4, 19]. These calorimeters are comprised of alternating layers of active and absorber materials [29]. The absorber layers initiate particle showers and propagate their growth, while the active layers produce a signal that is proportional to the energy deposited within them. The signal produced in the active layers is measured by sampling calorimeters and used to estimate the energy deposited in the absorber layers. This estimation is made by assuming the energy deposited across a calorimeter hit, that is one active and one absorber layer, is uniform. Working under this assumption, the total calorimeter hit energy is proportional to the active layer hit energy. This estimation procedure is loosely referred to as digitisation and, in this way, the cluster energy estimator introduced above can be written as

$$E_{Cluster} = \sum_{ECal\ hits,\ i} \epsilon_{ECal}^i \alpha_{ECal} + \sum_{HCal\ hits,\ i} \epsilon_{HCal}^i \alpha_{HCal} , \quad (2.2)$$

where α_{ECal} and α_{HCal} are digitisation constants for the ECal and HCal respectively, ϵ_{ECal}^i is the ECal active layer hit energy for hit i and ϵ_{HCal}^i is the HCal active layer hit energy for hit i . The first stage of the calibration procedure presented in this chapter covers the determination of these digitisation constants, which convert the raw analogue-to-digital converter (ADC) response to a hit energy.

Once the basic energy estimator has been calibrated, it is possible to apply more advanced procedures designed to give a compensating calorimeter response [1]. A compensating calorimeter produces an identical response to a particle shower irrespective of whether the particle shower is electromagnetic or hadronic in nature. The primary cause of the difference in the response of a calorimeter to electromagnetic and hadronic showers is the undetectable energy component that is found in hadronic showers. These undetectable energy components are energy deposits produced from a showering particle that do

not produce a signal in the calorimeters. Hadronic showers contain this undetectable component due to a combination of effects such as neutrons stopping within the calorimeter and nuclear binding energy losses. Typically, this leads to calorimeters having a weaker response to hadronic showers than to electromagnetic showers.

There are two distinct routes available for achieving a compensating response from a calorimeter: the first is hardware compensation [31], whereby calorimeters are constructed using materials that yield extra energy in response to hadronic showers; and the second is software compensation [9], whereby the uncompensated calorimetric energies for hadronic showers are modified at the software level.

A novel example of hardware compensation is the ZEUS calorimeter [31]. 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. Although hardware compensation is possible for the linear collider's calorimeters, it would require a large amount of radioactive material, which is highly undesirable.

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

$$E_{EM\ Cluster} = \sum_{ECal\ hits,\ i} E_{ECal}^i \beta_{ECal}^{EM} + \sum_{HCal\ hits,\ i} E_{HCal}^i \beta_{HCal}^{EM}, \quad (2.3)$$

and the hadronic cluster energy to

$$E_{Had\ Cluster} = \sum_{ECal\ hits,\ i} E_{ECal}^i \beta_{ECal}^{Had} + \sum_{HCal\ hits,\ i} E_{HCal}^i \beta_{HCal}^{Had}, \quad (2.4)$$

where the β s are scaling factors that are applied to the energy of clusters of calorimeter hits associated with electromagnetic and hadronic clusters in the ECal and HCal. This simple scaling of energies achieves a compensating calorimeter response, 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 β scaling factors as a function of the total energy deposited. An energy dependence in the scaling factors is expected as the mechanisms governing the propagation of hadronic showers are sensitive to the shower energy [32]. To account for this, more sophisticated software techniques have been developed that vary the calorimeter cluster energy estimator as a function of energy in order to achieve a compensating response across a wider range of energies. These techniques make use of the fine segmentation of the linear collider’s calorimeters to identify hadronic showers. These techniques also address the problem of spuriously high energy calorimeter hits, which are caused by Landau fluctuations [33]. Landau fluctuations originate from high energy knock-on electrons appearing within particle showers [34] and can lead to overestimates of the particle shower energy if they occur in the active layers of a sampling calorimeter.

2.2 Calibration in the Particle Flow Paradigm

Calibration of the linear collider detector simulation is performed by two processors in the software framework; the digitiser, which performs the digitisation process for sampling calorimeters, and PandoraPFA. The input to the digitiser is the active layer calorimeter response (ADC values) and the output is the combined calorimeter hit energies i.e. the active and absorber layer energies. The hit energies are then used by PandoraPFA for event reconstruction. Calibration of the digitiser involves determining the digitisation constants ($\alpha_{E\text{Cal}}$ and $\alpha_{H\text{Cal}}$) and the minimum ionising particle (MIP) scale, which is the average energy response for a MIP on a per hit basis. Similarly, calibration of PandoraPFA requires setting the scaling factors, β , and the MIP response using the combined calorimeter hit energies.

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

Determining the MIP scale is included as part of the calibration procedure as it is used by PandoraPFA in the identification of muons and for applying energy thresholds designed to limit the impact of noise. This energy scale is also used by the digitiser when simulating electrical noise, saturation effects in scintillator readout technologies and for applying noise vetoing energy thresholds [35].

The non-zero peak in the distribution of calorimeter hit energies for 10 GeV μ^- s entering the calorimeters at normal incidence is used to define the MIP scale in PandoraPFA, whereas in the digitiser ADC values are used [34]. In the linear collider detector simulation, several realistic effects are simulated by the digitiser including saturation effects, energy thresholds, timing cuts and electrical noise. Application of these effects at this point in the software chain means that the active layer hit energies are not subject to them, while the post digitisation combined calorimeter hit energies are. Consequently, the MIP scale in PandoraPFA cannot be obtained from the digitiser MIP scale, instead both have to be independently determined.

Although this overall procedure is referred to as calibration, strictly speaking this is not correct. Calibration sets the detector response to real data, while this procedure sets the simulated detector response to Monte-Carlo (MC) events. In a real detector, calibration would follow the setting of the simulated detector response to MC events so that simulations can be used to guide the calibration process.

2.2.1 Overview of the Calibration Procedure

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

1. Setting the MIP response in the digitiser.
2. Setting the digitisation constants, α_s , in the digitiser.

3. Setting the MIP response in PandoraPFA.
4. Setting the scaling factors, β_s , in PandoraPFA.

2.2.2 MIP Scale Determination in the Digitiser

The MIP scale in the digitiser was determined by simulating 10 GeV μ^- events and creating a distribution of active layer calorimeter hit energies for each calorimeter in the detector simulation. When populating these distributions, a direction correction factor of $\cos(\theta)$, where θ is the incident angle of the μ^- to the calorimeter cell, was applied to account for the path length of the MIP through the active medium of the calorimeter. This converts the individual response to a normal-incident equivalent response. No selection cuts were applied to the sample of muon events.

Figure 2.1 shows the distribution used to determine the MIP peak in the ECal, HCal barrel, HCal endcap and HCal ring. In the digitiser software only a single HCal MIP scale, taken as the HCal barrel, is implemented.

2.2.3 Digitisation Implementation

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

2.2.3.1 ECal Digitisation Implementation

The procedure for determining the digitisation constants in the ECal involves simulation of single photons at an energy $E_{MC} = 10$ GeV. Single photons at this energy are largely contained within the ECal, as shown in figure 2.2a. This makes them ideal for isolating the ECal digitisation calibration from that of the HCal digitisation calibration. Events are only used for calibrating the ECal digitisation if they are confined to the ECal. To that extent, cuts are applied ensuring that the sum of the reconstructed energy found outside the ECal is less than 1% of E_{MC} and that the $\cos(\theta) < 0.95$, where θ is the polar angle of the photon. Photons that convert are also vetoed in this event sample at MC

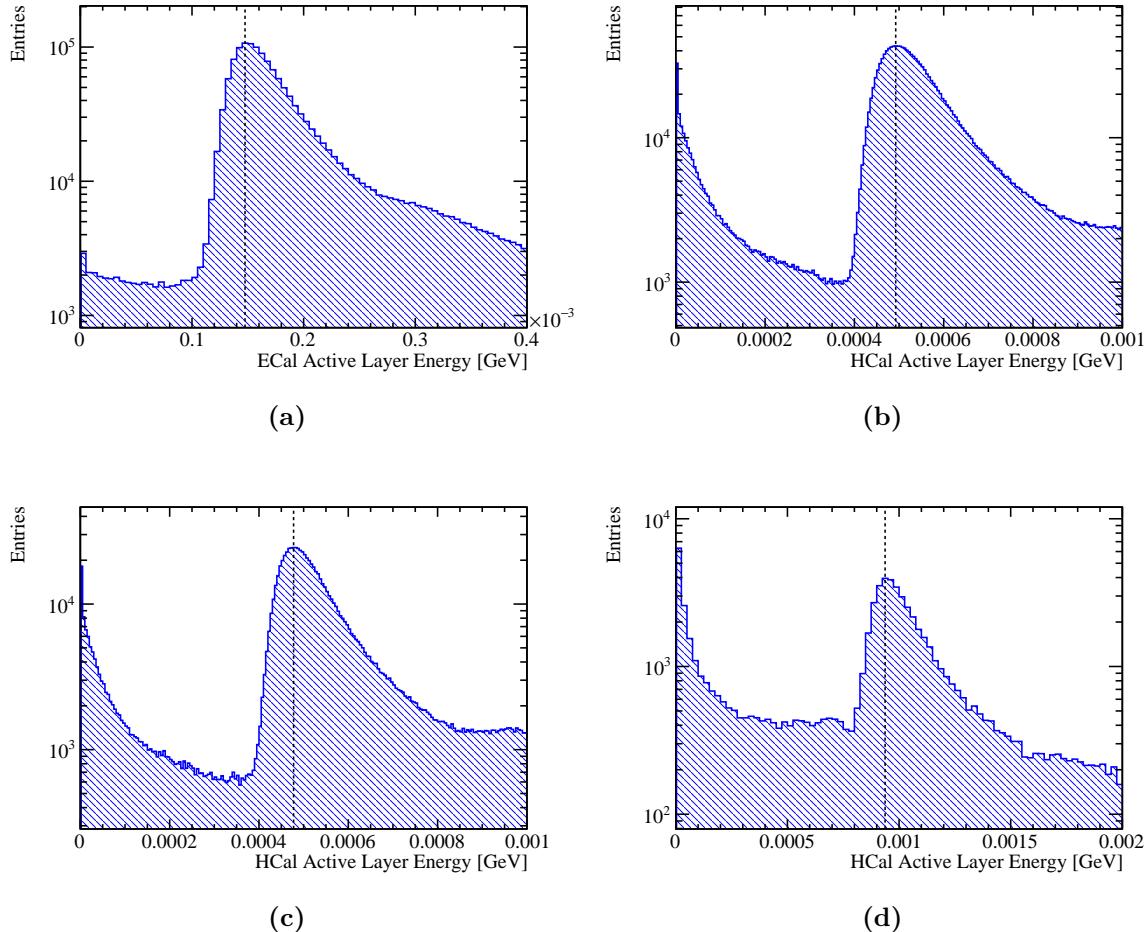


Figure 2.1: The active layer calorimeter hit energy distributions for (a) the ECal, (b) the HCal barrel, (c) the HCal endcap and (d) the HCal ring for 10 GeV μ^- events. The hit energies were corrected to account for the path length of the muons through the active medium of the calorimeter. The vertical black dotted lines indicate the position of the peak in each of these distributions that is used for defining the MIP scale in the digitisation processor.

level. The impact of these cuts on the sum of ECal hit energies for the $E_{MC} = 10$ GeV photons is shown in figure 2.2b.

The calibration of the digitisation in the ECal is an iterative procedure which begins with the simulation of single photons using a trial calibration, α_{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 2.2b. For an ideal calorimeter this distribution should be Gaussian, as described in chapter ??, therefore, a Gaussian fit is applied to this distribution and the mean, E_{Fit} , extracted. To remove the effect of any outliers

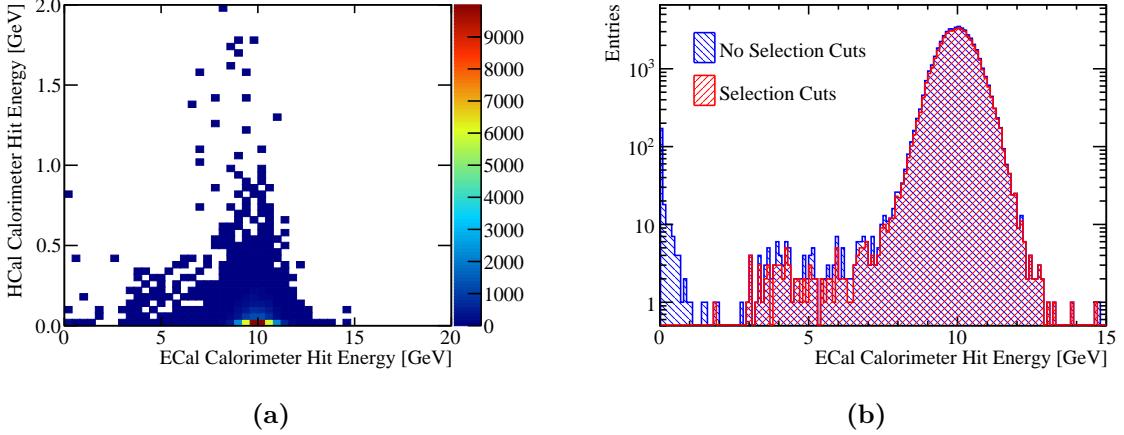


Figure 2.2: (a) The sum of calorimeter hit energies in ECal and HCal for 10 GeV photons.
(b) The sum of the ECal calorimeter hit energies for 10 GeV photons with and without the selection cuts.

in this distribution, the fit is applied to the range of data with the smallest root mean square that contains at least 90 % of the data. An example of such a fit is shown in figure 2.3. In the case of ideal calibration, the mean of this fit, E_{Fit} , would be equal E_{MC} . It is assumed that any difference between the two is due to the calibration, therefore, to correct this, the digitisation constant from the trial calibration, α_{ECal}^0 , is rescaled by the ratio of the E_{MC} to E_{Fit}

$$\alpha_{\text{ECal}}^0 \rightarrow \alpha_{\text{ECal}} = \alpha_{\text{ECal}}^0 \times \frac{E_{MC}}{E_{\text{Fit}}} . \quad (2.5)$$

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

2.2.3.2 HCal Digitisation Implementation

The calibration for the digitisation in the HCal proceeds in a similar manner to that described for the ECal with a few key differences. This calibration uses simulated MC long-lived neutral kaons (K_L^0) at $E_{MC} = 20 \text{ GeV}$. The higher energy, with respect to

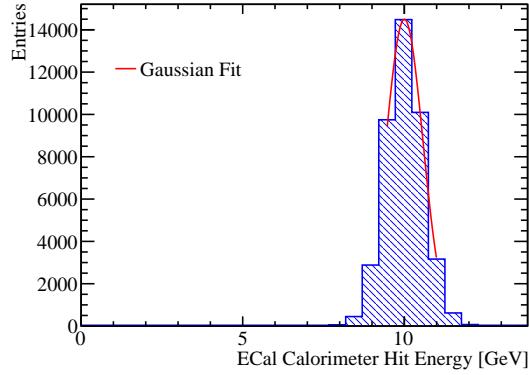


Figure 2.3: Gaussian fit to sum of the ECal calorimeter hit energies for 10 GeV photons with selection cuts. The coarse binning reflects the tolerance on the digitisation constant calibration.

the ECal digitisation, results in particle showers that sample deeper into the HCal. The K_L^0 s must pass through the ECal, which contains one λ_I , before arriving at the HCal. Consequently, approximately 15% of these events begin showering in the ECal, as can be seen in figure 2.4a. Only events that deposit less than 5% of their energy in the ECal are used for calibrating the HCal digitisation constants. Furthermore, events that are not contained in the HCal are removed by requiring the last layer of the HCal where energy is deposited to be in the innermost 90% of the HCal. The impact of these cuts on the sum of HCal calorimeter hit energies for the $E_{MC} = 20$ GeV K_L^0 events is shown in figure 2.4b.

There are two HCal digitisation constants used in the detector simulation, one applied for the barrel and another for the endcap. The use of two digitisation constants accounts for differences in hadronic shower dynamics between the two, such as differing magnetic field configurations in the barrel and endcap. Both parameters are calibrated in the same manner, but have different cuts on θ , the polar angle of the K_L^0 . For the barrel region of the HCal events are selected if $0.2 < \cos(\theta) < 0.6$, while for the endcap events are selected if $0.8 < \cos(\theta) < 0.9$. These angular cuts account for the transverse profile of the hadronic showers and ensure that the showers are largely confined to the relevant sub-detector. The target reconstructed energy for these K_L^0 s is the kinetic energy because many of the neutral hadrons appearing in jets are neutrons and their accessible energy is the kinetic energy [32] as opposed to the total energy.

After applying the above K_L^0 selection cuts, the calibration procedure for the digitisation of the HCal barrel and endcap proceeds in the same manner as was described

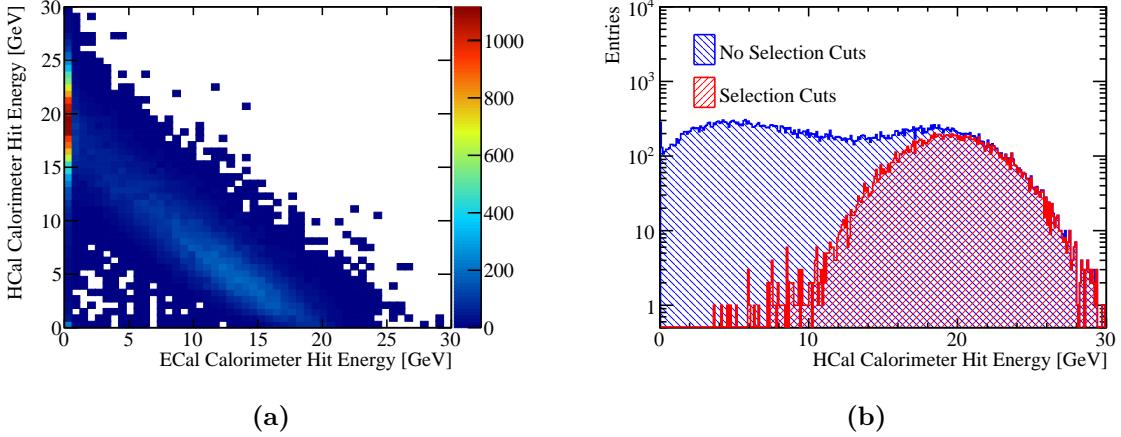


Figure 2.4: (a) Sum of calorimeter hit energies in ECal and HCal for 20 GeV K_L^0 events.
(b) Sum of the HCal calorimeter hit energies for a 20 GeV K_L^0 events with and without the selection cuts.

for the ECal. An example of the Gaussian fits applied to the sum of the calorimeter hit energies in the HCal barrel and endcap are shown in figure 2.5.

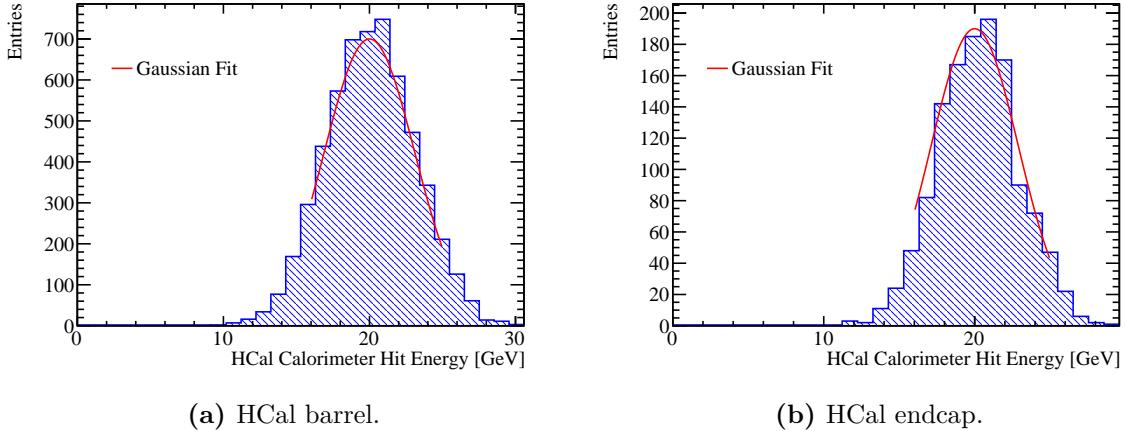


Figure 2.5: Gaussian fit to sum of the HCal calorimeter hit energies for 20 GeV K_L^0 events with selection cuts.

2.2.3.3 HCal Ring Digitisation Implementation

The HCal ring, as illustrated in figure 2.6, is a hadronic calorimeter that surrounds the ECal endcap and it is sandwiched between the HCal barrel and endcap. This calorimeter

is required to ensure hermetic coverage of the hadronic calorimeter system across the barrel/endcap cross-over region [4].

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

$$\alpha_{\text{HCal ring}} = \alpha_{\text{HCal endcap}} \times \frac{\langle \cos(\theta_{\text{endcap}}) \rangle}{\langle \cos(\theta_{\text{ring}}) \rangle} \times \frac{P_{\text{endcap}}}{P_{\text{ring}}} \times \frac{L_{\text{endcap}}^{\text{Absorber}}}{L_{\text{ring}}^{\text{Absorber}}} \times \frac{L_{\text{ring}}^{\text{Active}}}{L_{\text{endcap}}^{\text{Active}}}, \quad (2.6)$$

where θ is the incident angle of the incoming particle to the calorimeter determined using the 20 GeV K_L^0 s, 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 μ^- events.

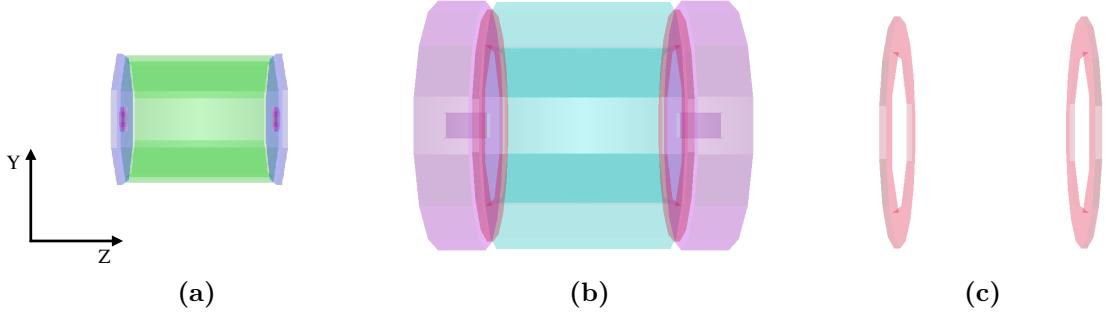


Figure 2.6: A PandoraPFA event display showing the nominal ILD calorimeters. (a) the ECal, (b) the full HCal and (c) the HCal ring, which covers the barrel/endcap cross-over region.

2.2.4 MIP Scale Determination in PandoraPFA

The MIP scale in PandoraPFA is set by simulating 10 GeV μ^- events and creating the distribution of combined calorimeter hit energies. The MIP scale in PandoraPFA must be determined for the calorimeters and, in contrast to the digitiser, the muon chamber. Consequently, an additional distribution showing the calorimeter hit energy for the muon

chamber must be constructed at this stage of the calibration. As was done for the digitiser, a direction correction factor was applied to the hit energies to account for the path length of the MIP through the active medium of the calorimeter and no selection cuts were applied.

Examples of the distributions used to set the MIP scale in PandoraPFA can be found in figure 2.7. Due to the energy thresholds applied in the digitiser, there are fewer populated bins with low hit energies. The double peak structure observed in the ECal calorimeter hit energy distribution is expected given the ECal absorber material thickness doubling in the back 10 layers of the ECal. The MIP peaks used for defining the MIP scale in PandoraPFA, figure 2.7, are broader than those used for determining MIP scale setting in the digitiser, 2.1, as the realistic effects applied by the digitiser are only present in the combined calorimeter hit energy distributions.

2.2.5 Electromagnetic Scale in PandoraPFA

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

The electromagnetic scale in the ECal, β_{ECal}^{EM} , is determined using simulated photons at $E_{MC} = 10$ GeV. To ensure that the events used for this part of the calibration are largely confined to the ECal, a cut requiring less than 1% of the reconstructed energy to be found outside the ECal is applied. Furthermore, only events reconstructed as a single photon are used to veto conversions. The impact of the selection cuts on the electromagnetic energy measured in the ECal for 10 GeV photons is shown in figure 2.8a. The peak at zero electromagnetic energy in the ECal is due to events traveling down the beam pipe and photon conversions. In photon conversion events, the calorimetric energy deposits made by the e^\pm are associated to charged particle tracks. In this case, the energy measured using the calorimeters will be reported as zero because the charged particle tracks are used to determine the reconstructed particle energies. The tail of events with low electromagnetic energy in the ECal occurs primarily due to pattern recognition failures in photon conversion events. In these events a small fraction of the calorimetric energy deposits made by the e^\pm are not associated to charged particle tracks and instead are reconstructed as separate photons with a reconstructed energy much less than E_{MC} .

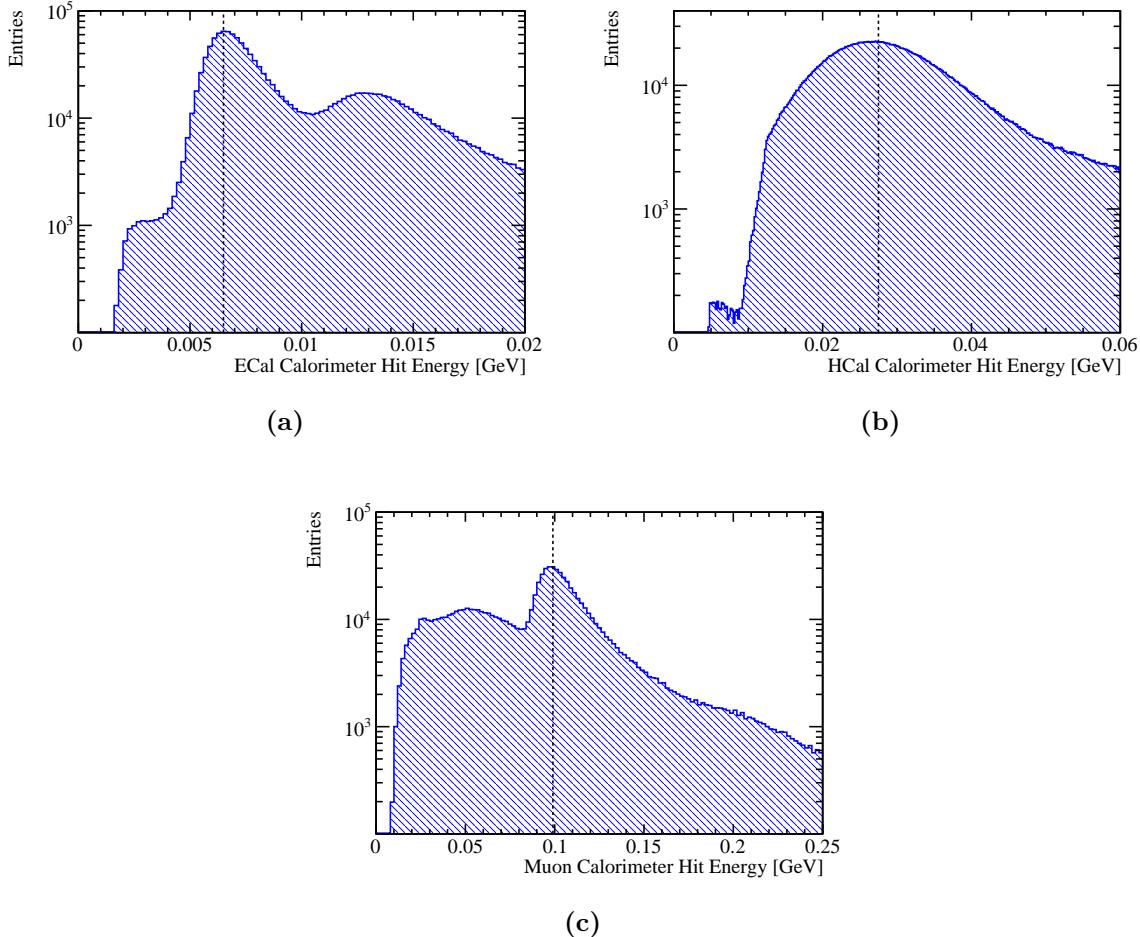


Figure 2.7: The combined calorimeter hit energy distributions for (a) the ECal, (b) the HCal and (c) the muon chamber for 10 GeV μ^- events. These hit energies were corrected to account for the path length of the muons through the active medium of the calorimeter. The vertical black dotted lines indicate the position of the peak in each of these distributions that is used for defining the MIP scale in PandoraPFA.

The fitting procedure follows the procedure used for the ECal digitisation, described in section 2.2.3.1, whereby a trial calibration for the electromagnetic energy scale in the ECal, $\beta_{E\text{Cal}}^{EM0}$, is first assumed. The initial trial calibration is approximate and is iteratively updated until it converges to within a chosen tolerance. Using the trial calibration, the photons are reconstructed and the distribution of the electromagnetic energy in the ECal created. A Gaussian fit is then applied to this distribution in the range with the smallest root mean square containing at least 90 % of the data. The mean of the fitted Gaussian,

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

$$\beta_{E\text{Cal}}^{EM0} \rightarrow \beta_{E\text{Cal}}^{EM} = \beta_{E\text{Cal}}^{EM0} \times \frac{E_{MC}}{E_{\text{Fit}}} . \quad (2.7)$$

An example distribution and fit used in the calibration of the nominal ILD detector model can be found in figure 2.8b. This procedure is repeated using the updated $\beta_{E\text{Cal}}^{EM}$ until E_{Fit} falls within a specified tolerance. The tolerance applied here was $|E_{\text{Fit}} - E_{MC}| < E_{MC} \times 0.5\%$. The binning for the fitted histogram is chosen such that the bin width is equal to the desired target tolerance on E_{Fit} , e.g. $E_{MC} \times 0.5\% = 0.05 \text{ GeV}$. This tolerance is tighter than was applied for the digitisation as it is these and only these energies that are used in downstream analyses.

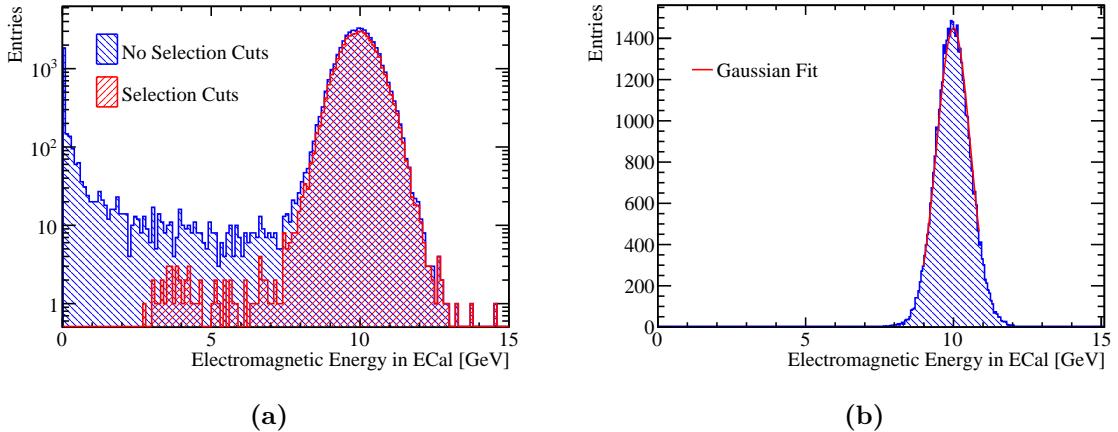


Figure 2.8: (a) The sum of the electromagnetic energy measured in the ECal for simulated 10 GeV photons with and without the selection cuts. (b) Gaussian fit to sum of the electromagnetic energy deposited in the ECal for simulated 10 GeV photons with selection cuts.

2.2.6 Hadronic Scale in PandoraPFA

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

For the reasons outlined in section 2.2.3.2, the target reconstructed energy for the sample of K_L^0 s used for setting the hadronic energy scale is the kinetic energy, E_K , as opposed to the total energy. To ensure the events used are not affected by leakage of energy out of the back of the HCal, a cut is applied that vetoes events where energy is deposited in the outermost 10% of the HCal. In addition, a cut requiring a single neutral hadron to be reconstructed is applied to veto events with reconstruction failures and decays in the tracker. Finally, it is required that the total hadronic energy measured within the calorimeters falls within three σ of the kinetic energy of the K_L^0 , where σ is taken to be $55\% \times \sqrt{E_K}$ GeV. This definition for σ is approximately the energy resolution for neutral hadrons using the nominal ILD HCal [4]. This cut ensures that when fitting the two dimensional distribution of hadronic energy measured in the ECal and HCal, outliers do not skew the fit. The impact of these selection cuts can be seen in figure 2.9.

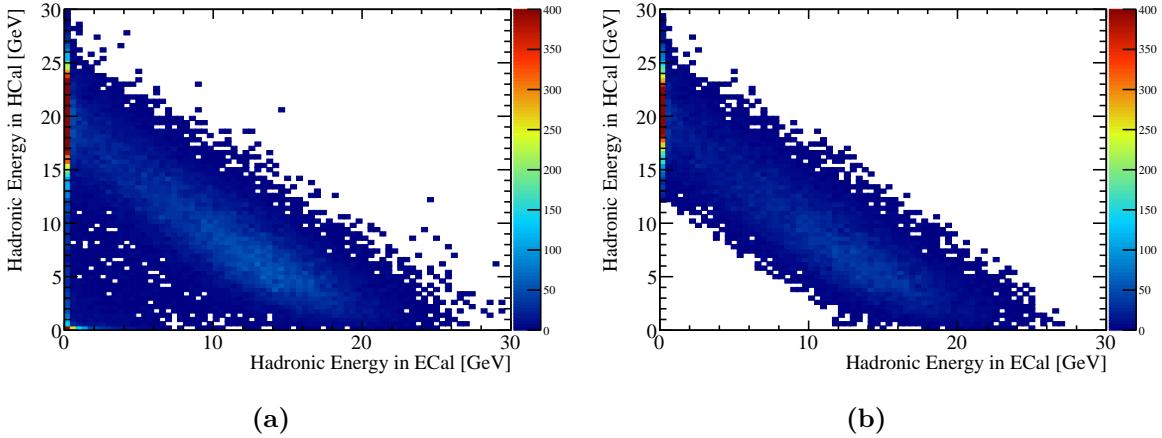


Figure 2.9: The distribution of hadronic energy measured in the ECal and HCal for 20 GeV K_L^0 events (a) without selection cuts and (b) with selection cuts.

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

$$\chi^2(\delta_{ECal}^{Had}, \delta_{HCal}^{Had}) = \sum_i \left(\frac{r_i}{\sigma_{r_i}} \right)^2, \quad (2.8)$$

where r_i is the perpendicular distance in the two dimensional plane of hadronic energies measured in the ECal and HCal from the point (x_i, y_i) to a straight line passing through the points $(\delta_{E\text{Cal}}^{\text{Had}}, 0)$ and $(0, \delta_{H\text{Cal}}^{\text{Had}})$. In this definition, x_i and y_i are the hadronic energies measured in the ECal and HCal respectively for event i . The variables $\delta_{E\text{Cal}}^{\text{Had}}$ and $\delta_{H\text{Cal}}^{\text{Had}}$ describe a linear fit to the hadronic energy distribution, which are to be varied when minimising χ^2 . The explicit definition of r_i is given in equation 2.9 and illustrated in figure 2.10. The uncertainty on r_i is given by σ_{r_i} , which is explicitly defined in equation 2.10. This uncertainty is calculated by propagating the uncertainties on x_i and y_i , which are assumed to be $\sigma_{x_i/y_i} = 55\% \times \sqrt{x_i/y_i}$, into the expression for r_i . The sum runs over all events, i , passing the selection cuts.

$$r_i = \frac{y_i \delta_{E\text{Cal}}^{\text{Had}} + x_i \delta_{H\text{Cal}}^{\text{Had}} - \delta_{E\text{Cal}}^{\text{Had}} \delta_{H\text{Cal}}^{\text{Had}}}{\sqrt{(\delta_{E\text{Cal}}^{\text{Had}})^2 + (\delta_{H\text{Cal}}^{\text{Had}})^2}} , \quad (2.9)$$

$$\sigma_i = \frac{(\sigma_{y_i} \delta_{E\text{Cal}}^{\text{Had}})^2 + (\sigma_{x_i} \delta_{H\text{Cal}}^{\text{Had}})^2}{\sqrt{(\delta_{E\text{Cal}}^{\text{Had}})^2 + (\delta_{H\text{Cal}}^{\text{Had}})^2}} . \quad (2.10)$$

The minimisation of χ^2 is done by stepping over a range of $\delta_{E\text{Cal}}^{\text{Had}}$ and $\delta_{H\text{Cal}}^{\text{Had}}$ centred about

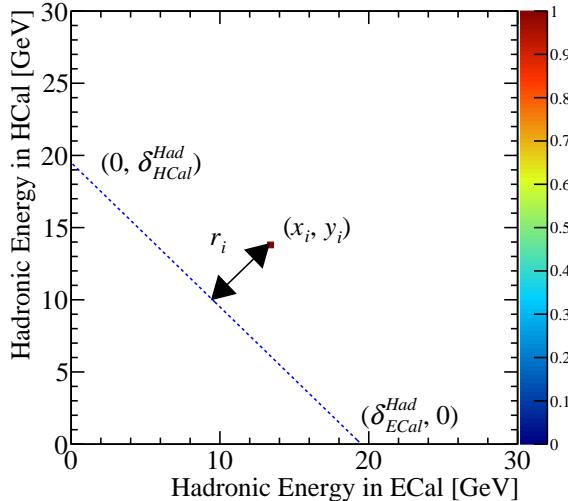


Figure 2.10: An example showing the definition of r_i . The blue dotted line corresponds to $y_i = \delta_{H\text{Cal}}^{\text{Had}} - x_i \delta_{H\text{Cal}}^{\text{Had}} / \delta_{E\text{Cal}}^{\text{Had}}$.

the ideal value of E_K in search for the minimum χ^2 . Once the minima in χ^2 is found the trial calibration factors $\beta_{E\text{Cal}}^{\text{Had}0}$ and $\beta_{H\text{Cal}}^{\text{Had}0}$ are rescaled to correct for any deviation from

the desired fit as follows

$$\beta_{E\text{Cal}}^{\text{Had}0} \rightarrow \beta_{E\text{Cal}}^{\text{Had}} = \beta_{E\text{Cal}}^{\text{Had}0} \times \frac{E_K}{\Delta_{E\text{Cal}}^{\text{Had}}} , \quad (2.11)$$

$$\beta_{H\text{Cal}}^{\text{Had}0} \rightarrow \beta_{H\text{Cal}}^{\text{Had}} = \beta_{H\text{Cal}}^{\text{Had}0} \times \frac{E_K}{\Delta_{H\text{Cal}}^{\text{Had}}} , \quad (2.12)$$

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

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

Setting the hadronic scale in PandoraPFA ensures that the energy estimators for neutral hadrons are accurate at 20 GeV, however, this is not true for all energies. The undetectable energy component of a hadronic shower varies as a function of particle shower energy [32]. This means the response of a calorimeter to neutral hadrons is non-linear with the hadron energy. This is an inherent limitation of this calibration procedure that will be addressed by the development of more sophisticated energy estimators in subsequent chapters.

2.2.7 Summary

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

2.3 Novel Energy Estimators

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

$$E_{Cluster} = \sum_{ECal\ hits,\ i} E_{ECal}^i + \sum_{HCal\ hits,\ i} E_{HCal}^i \omega^i(\rho^i) . \quad (2.13)$$

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

Although the exact weights depend on the implementation of the technique, a general feature is that at large E_{HCal}^i the weight is less than one. This limits the impact of spuriously high energy hits caused by Landau fluctuations. The energy loss probability distribution function for scintillator detectors, such as the ILD HCal, is given by a Landau function [33]. Energy deposits from the high energy tail of this distribution, which are known as Landau fluctuations, account for high energy knock-on electrons that appear within particle showers [34]. As Landau fluctuations deposit a disproportionately large amount of energy with respect to the bulk of the particle shower, they can lead to overestimates of the particle shower energy.

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

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

2.3.1 HCal Hit Energy Truncation

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

$$E \rightarrow E' = \begin{cases} E & \text{if } E < \kappa , \\ \kappa & \text{otherwise ,} \end{cases} \quad (2.14)$$

where κ is the value of the truncation. This improves the energy estimators for hadronic clusters by limiting the impact of Landau fluctuations. In terms of ω introduced in equation 2.13 the truncation corresponds to

$$\omega(\rho) = \begin{cases} 1 & \text{if } \rho \times V < \kappa , \\ \frac{\kappa}{\rho \times V} & \text{otherwise .} \end{cases} \quad (2.15)$$

This weight as a function of hit energy density is shown in figure 2.11.

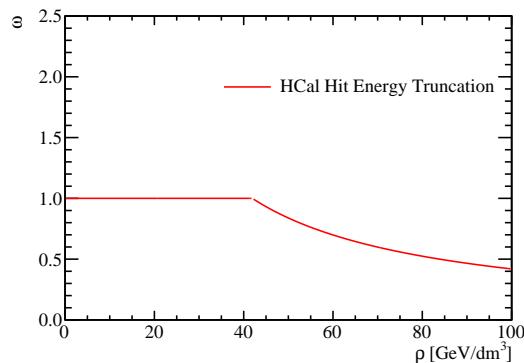


Figure 2.11: The weights, ω , used in the HCal hit energy truncation as a function of the energy density of the HCal hit, ρ . The truncation shown here corresponds to a 1 GeV truncation in the nominal ILD HCal.

2.3.1.1 Legacy Energy Corrections

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

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

In the reconstruction, these corrections are applied to each cluster of calorimeter hits, irrespective of the location of that cluster in the detector. These algorithms, with the HCal hit truncation, form the "legacy" energy corrections that are used by PandoraPFA when performing the event reconstruction.

2.3.1.2 Impact on Single Particle Energy Resolution

Figure 2.12a shows the energy resolution for neutral hadrons as a function of the HCal hit energy truncation. The optimal truncation for the ILD detector model simulation was 1 GeV and, using this truncation, a neutral hadron energy resolution of $\sim 8.8\% = 62\%/\sqrt{E(\text{GeV})}$ was obtained for $E = 50 \text{ GeV } K_L^0$ events. In comparison, the neutral hadron energy resolution for $E = 50 \text{ GeV } K_L^0$ events obtained without a truncation was $\sim 10.4\% = 74\%/\sqrt{E(\text{GeV})}$. Smaller energy truncations begin to truncate the energy of calorimeter hits produced in typical hadronic shower development, while larger truncations allow for a larger impact from Landau fluctuations. Both of these effects result in worsening neutral hadron energy resolutions. For completeness the photon energy resolutions as a function of HCal hit energy truncation are shown in figure 2.12b.

As expected the photon energy resolution is unaffected by the HCal hit energy truncation as the photons are largely contained within the ECal.

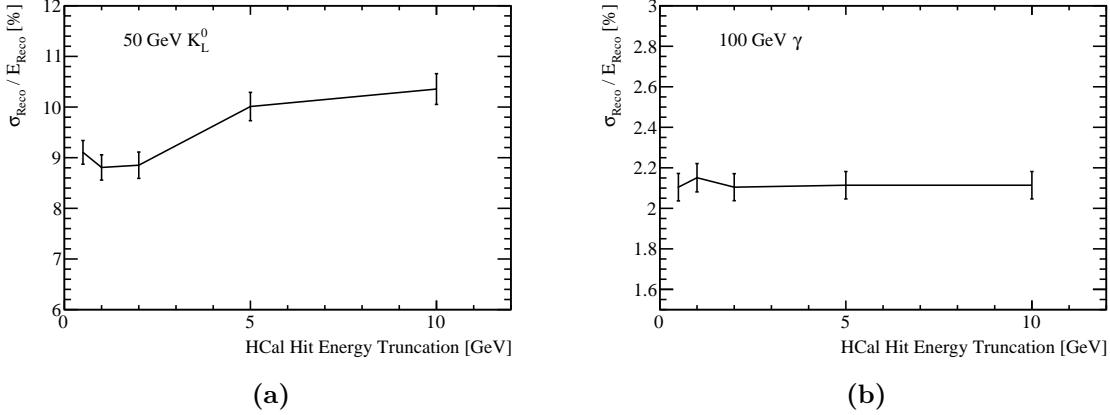


Figure 2.12: The energy resolution as a function of HCal cell truncation for (a) 50 GeV K_L^0 events and (b) 100 GeV photons using the nominal ILD detector model.

2.3.1.3 Impact on Jet Energy Resolution

Figure 2.13 shows the jet energy resolution as a function of jet energy for selected values of the HCal hit energy truncation. The trends in this plot are complex as the optimal HCal hit energy truncation varies with the jet energy. For 45.5 GeV jets, the best jet energy resolution, $\sim 3.6\%$, is obtained using a 0.5 GeV truncation, while for 180 GeV jets, the best jet energy resolution, $\sim 2.9\%$, is obtained using a 1 GeV truncation. This is expected because at low jet energies the average number of particles passing through each calorimeter hit will be small. Therefore, the impact of Landau fluctuations is large and, to limit them, a low truncation energy is needed. As the jet energy increases, more particles on average pass through each calorimeter hit and the impact of Landau fluctuations decreases.

It is clear that a 1 GeV HCal hit energy truncation is beneficial for the performance of the nominal ILD detector model since the jet energy resolution is reduced by roughly $\sim 0.15\%$ across the jet energy range from 45.5 GeV to 250 GeV. As the HCal hit truncation technique offers significant performance gains, it is used for the calorimeter optimisation studies presented in chapter ???. These studies include optimisation of the HCal cell size. Increasing the HCal cell size will increase the average number of particles passing through each calorimeter hit, which in turn reduces the impact of Landau fluctuations and vice

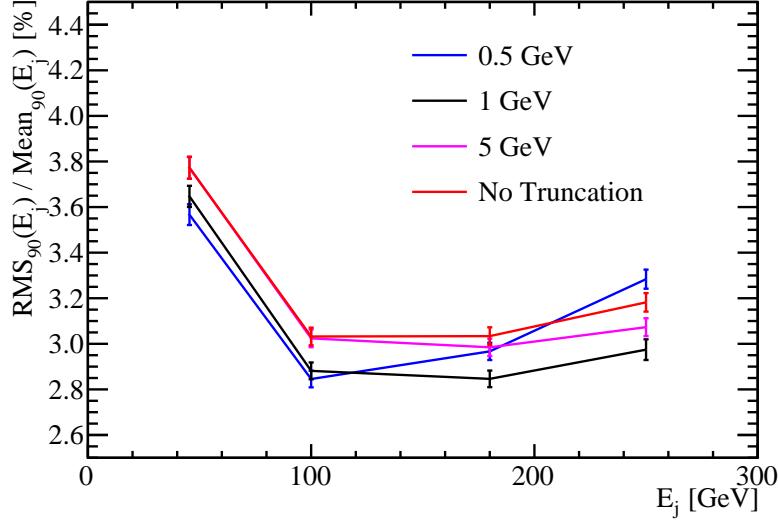


Figure 2.13: The jet energy resolution as a function of jet energy for various HCal hit energy truncations. The results shown use the nominal ILD detector model, which contains $30 \times 30 \text{ mm}^2$ square scintillator tiles in the HCal.

versa. For all detector models considered where the HCal cell size was varied, the HCal hit energy truncation was re-optimised to account for the changing impact of Landau fluctuations. For detector models with a HCal cell size of $10 \times 10 \text{ mm}^2$, $20 \times 20 \text{ mm}^2$, $30 \times 30 \text{ mm}^2$, $40 \times 40 \text{ mm}^2$, $50 \times 50 \text{ mm}^2$ and $100 \times 100 \text{ mm}^2$ the reoptimised truncation values were 0.5, 0.75, 1, 1.5, 2 and 5 GeV respectively. Furthermore, the average particle density in a HCal hit will also be sensitive to the properties of the absorber material used in the calorimeters, therefore, the HCal hit energy truncation was also reoptimised in the HCal absorber material study. The optimal truncation energy cut for the $30 \times 30 \text{ mm}^2$ cell size tungsten HCal option was 5 GeV, while for all other detector models considered it was 1 GeV. The cause of increased truncation energy cut for tungsten is discussed in section ??.

Understanding the effect of the HCal hit energy truncation is crucial when performing optimisation studies. This can be seen in figure 2.14, which shows the results of the HCal cell size optimisation study when using a 1 GeV truncation and when optimising the truncation for each detector model. By applying a uniform HCal hit energy truncation the importance of the HCal cell size to particle flow calorimetry is vastly overinflated. For example, if the HCal cell size is increased from 10 mm to 100 mm the jet energy resolution for 250 GeV jets goes from $\sim 2.8\%$ to $\sim 4.5\%$ for the flat 1 GeV truncation, but only $\sim 3.5\%$ when using an optimised truncation. As the jet energy and HCal cell

size increase, the flat 1 GeV truncation throws away a larger fraction of typical hadronic shower energy measurements, which causes the jet energy resolution to degrade rapidly.

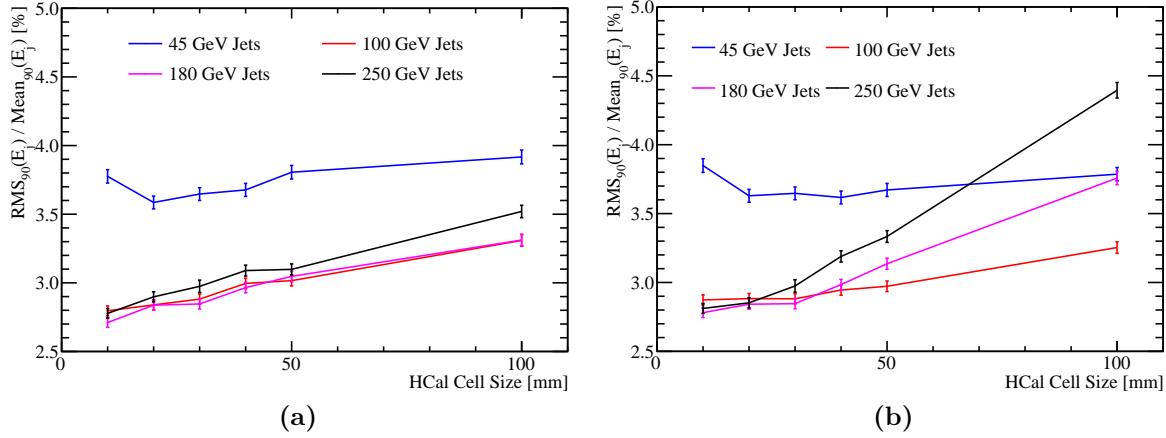


Figure 2.14: The jet energy resolution as a function of HCal cell size in the ILD detector model using a HCal hit energy truncation that is (a) optimised and (b) fixed at 1 GeV.

2.3.2 Software Compensation

Particle showers that are produced when a hadron interacts with a calorimeter contain two components [32]; an electromagnetic shower core (which originates from the production and decay of π^0 s and η s) and a hadronic shower component originating from other interacting and decaying particles. By identifying each of these components in the reconstruction, it is possible to modify their energies to give a compensating calorimeter response. This technique is known as software compensation.

Software compensation achieves a compensating calorimeter response by applying weights, as introduced in equation 2.13, that modify the energy of calorimeter hits in the HCal. These weights increase the energy found in the hadronic hits to compensate for the undetectable energy component found in hadronic showers. Additionally, these weights reduce the energy of spuriously high energy hits to minimise the impact of Landau fluctuations. The weights vary as a function of the calorimeter hit energy density, ρ^i , and the uncompensated energy of the particle shower, E_{Raw} , where

$$E_{Raw} = \sum_{ECal\ hits,\ i} E_{ECal}^i + \sum_{HCal\ hits,\ i} E_{HCal}^i . \quad (2.16)$$

The electromagnetic and hadronic components of a hadronic particle shower are treated differently in this approach by applying weights that are sensitive to the energy density of the calorimeter hits. Hits with large energy densities are likely to be part of the electromagnetic core, while low energy density hits are likely to be part of satellite hadronic hits around the electromagnetic shower core [30]. By tailoring the weights as a function of the energy density, a compensating calorimeter response can be obtained. Figure 2.15 shows the electromagnetic and hadronic shower components, determined by the energy density of the calorimeter hits, for a hadronic shower in a 500 GeV $Z \rightarrow u\bar{d}s\bar{s}$ di-jet event.

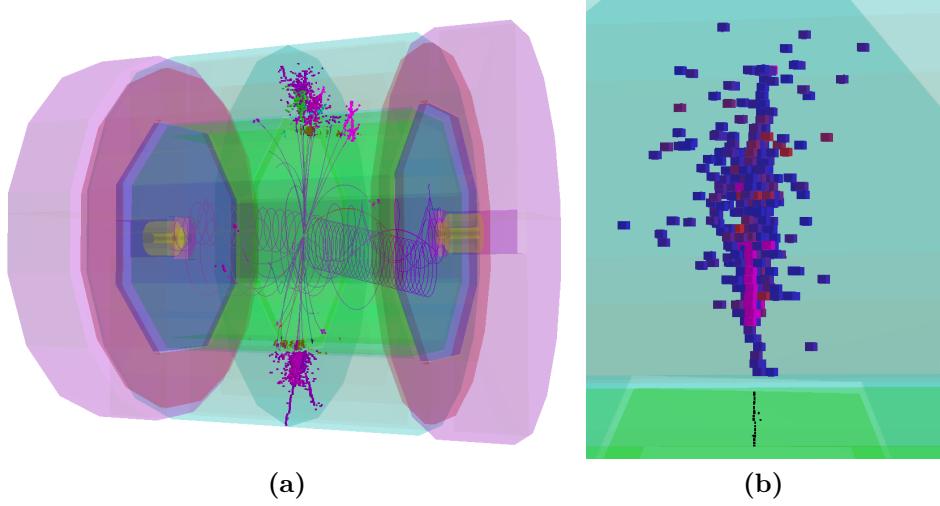


Figure 2.15: An event display for a 500 GeV $Z \rightarrow u\bar{d}s\bar{s}$ di-jet event reconstructed using the nominal ILD detector. (a) The full event environment. (b) A single hadronic cluster from the same event where shading indicates the energy density in the HCal. High energy density cells are coloured red, while lower energy density cells are coloured blue. All ECal hits are shaded black. The high energy density electromagnetic core of the selected hadronic cluster is clearly visible.

The software compensation weights also depend on E_{Raw} , the total raw cluster energy, to account for the sensitivity of the hit energy density distribution on the total particle shower energy. For hadronic showers, the fraction of the total energy carried in the electromagnetic core increases as the total shower energy increases [32], therefore a dependency of the weights on E_{Raw} is needed to obtain a compensating calorimeter response across a wide range of energies.

The precise form of the weights used in this technique are [30]

$$\omega(E_{Raw}, \rho) = p_1(E_{Raw}) \times \exp(p_2(E_{Raw}) \times \rho) + p_3(E_{Raw}) , \quad (2.17)$$

$$(2.18)$$

with

$$p_1(E_{Raw}) = p_{11} + p_{12} \times E_{Raw} + p_{13} \times E_{Raw}^2 \quad (2.19)$$

$$p_2(E_{Raw}) = p_{21} + p_{22} \times E_{Raw} + p_{23} \times E_{Raw}^2 \quad (2.20)$$

$$p_3(E_{Raw}) = \frac{p_{31}}{p_{32} + \exp(p_{33} \times E_{Raw})} , \quad (2.21)$$

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

$$E_{Raw} = \sum_{ECal\ hits,\ i} E_{ECal}^i + \sum_{HCal\ hits,\ i} E_{HCal}^i . \quad (2.22)$$

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

$$\chi^2(p_{\alpha\beta}) = \sum_{Events} \left(\frac{(E_{Cluster}^{SC}(p_{\alpha\beta}) - E_{MC})}{0.5 \times \sqrt{E_{MC}}} \right)^2 \quad (2.23)$$

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

Figure 2.17 shows ω as a function of ρ for selected values of E_{Raw} and figure 2.16 shows p_1 , p_2 and p_3 as a function of E_{Raw} . These weights shown in figures 2.17 and 2.16 were obtained by training the software compensation technique on samples simulated using the nominal ILD detector model. Figure 2.17 shows that the high energy density hits are being reduced in energy to compensate for the effects of Landau fluctuations, while the low energy density hits are being increased in weight to compensate for the undetectable energy component found in hadronic showers. Furthermore, the weights vary as a function of the raw hadronic shower energy to account for the changing energy density topology of hadronic showers with increasing shower energy.

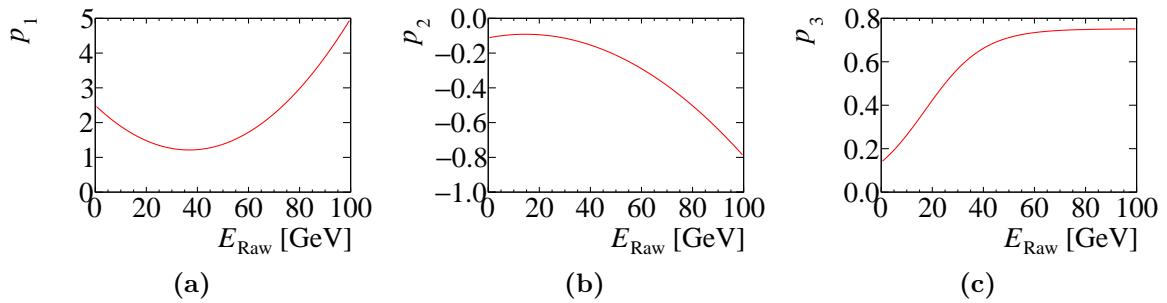


Figure 2.16: The software compensation parameters (a) p_1 , (b) p_2 and (c) p_3 as a function of E_{Raw} , the total raw cluster energy. These weights were obtained by training the software compensation technique on samples simulated using the nominal ILD detector model.

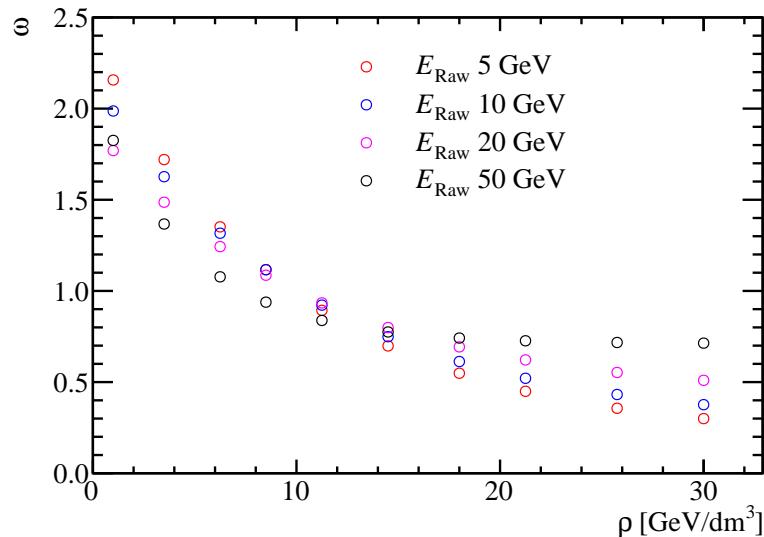


Figure 2.17: The software compensation weight applied to a calorimeter hit as a function of calorimeter hit energy density for various cluster energies.

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

As software compensation only modifies the energy of HCal hits there is freedom to apply further energy corrections to the ECal hits. Applying the "Clean Clusters"

logic, described in section 2.3.1.1, to the ECal hits alongside software compensation was found to be beneficial to the jet energy resolution. Therefore, the application of software compensation within PandoraPFA implicitly involves the application of the "Clean Clusters" logic to the ECal hits.

Software compensation was tuned using a maximum K_L^0 energy of 100 GeV, therefore, it is only applied to clusters where $E_{Raw} < 100$ GeV; sensible behaviour outside this range cannot be ensured. While it would be possible to modify the energy range of the training sample to go to higher energies, hadronic clusters with energy greater than 100 GeV will be rare at the ILC-like energies, i.e. $\sqrt{s} \leq 500$ GeV, considered here.

2.3.2.1 Impact on Single Particle Energy Resolution

Figure 2.18 shows the energy resolution as a function of MC energy for single K_L^0 events obtained using the various energy correction configurations in PandoraPFA. When comparing the energy resolution given by software compensation to that obtained using no energy corrections, it can be seen that software compensation offers an improvement in the energy resolution of $\sim 15\%$ across the energy range considered. The uniformity of this improvement is encouraging, indicating that software compensation is achieving a compensating calorimeter response across this wide range of energies.

Comparing the performance of software compensation to the legacy corrections, described in section 2.3.1.1, it can be seen that software compensation gives a better energy resolution across almost the entire range of energies considered. The only exception to this is around $E_{K_L^0} \sim 50$ GeV where the performance of software compensation and the legacy corrections are comparable. By removing the hit truncation from the legacy options it is clear that the changes in energy resolution when using the legacy options are being driven by the hit truncation. This makes the trend in energy resolution observed using the legacy corrections clear as, at low K_L^0 energies, very few hits are affected by the truncation so the performance is comparable to not using any energy corrections. At high K_L^0 energies, the truncation is too aggressive and removes energy from hits that are not spuriously high leading to a worsening energy resolution. Between these two extremes, $E_{K_L^0} \sim 50$ GeV, the truncation works ideally and the improvement in energy resolution when using the legacy corrections is the largest.

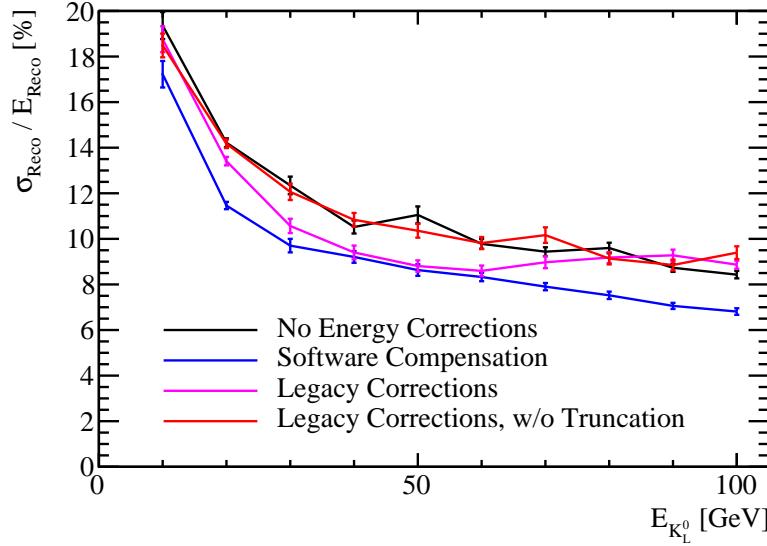


Figure 2.18: The energy resolution as a function of the MC energy for single K_L^0 events using various energy correction settings. The black line represents no energy corrections, the blue line represents software compensation, the magenta line represents the legacy energy corrections and the red line represents the legacy corrections without the HCal hit energy truncation. The nominal ILD detector model was used in these simulations.

2.3.2.2 Impact on Jet Energy Resolution

The improvements in the intrinsic energy resolution of the detector observed when using software compensation will propagate into the reconstruction of jets. Figure 2.19 shows the jet energy resolution as a function of jet energy when using selected energy correction configurations in PandoraPFA. It can be seen that software compensation improves the jet energy resolution by $\sim 15\%$ across the energy range considered in comparison to using no energy corrections. Furthermore, software compensation offers an improvement in the jet energy resolution of the order of 5% for jet energies $\gtrsim 100$ GeV in comparison to the legacy corrections, which prior to the development of software compensation had given the best jet energy resolutions.

Figure 2.20 shows the intrinsic energy resolution and confusion contributions to the jet energy resolution as a function of jet energy when using selected energy correction configurations in PandoraPFA. 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 2.3.2.1.

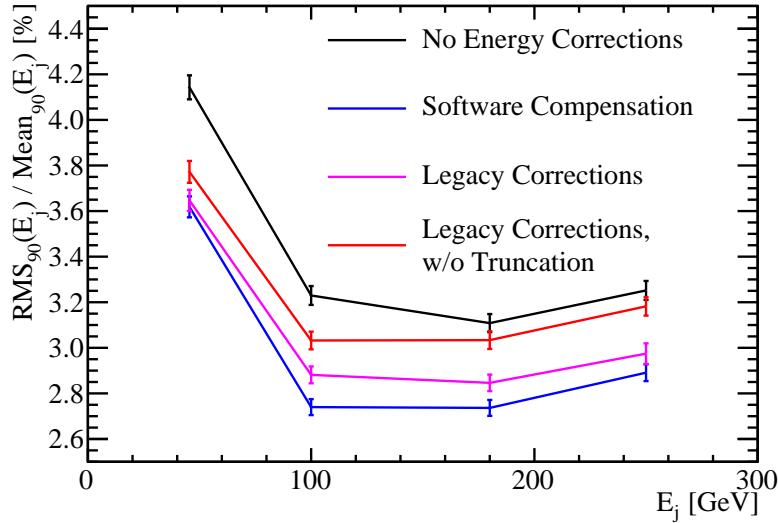


Figure 2.19: The jet energy resolution as a function of the jet energy for a variety of different energy correction options. The black line represents no energy corrections, the blue line represents software compensation, the magenta line represents the legacy energy corrections and the red line represents the legacy corrections without the HCAL hit energy truncation. The nominal ILD detector model was used in these simulations.

When compared to the legacy energy corrections, software compensation improves the intrinsic energy resolution by up to 12% across the energy range considered, with the largest improvement occurring for 100 GeV jets. As jets contain a broad spectrum of hadronic cluster energies, there is no jet energy for which the intrinsic energy resolution of the detector is comparable between the legacy corrections and software compensation. The confusion contributions to the jet energy resolution when using software compensation and the legacy corrections are almost identical. This indicates that the improvement seen in the jet energy resolution when comparing software compensation to the legacy corrections, shown in figure 2.19, is being driven by the intrinsic energy resolution.

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

logic is applied to ECal hits in the software compensation energy correction as previously discussed.

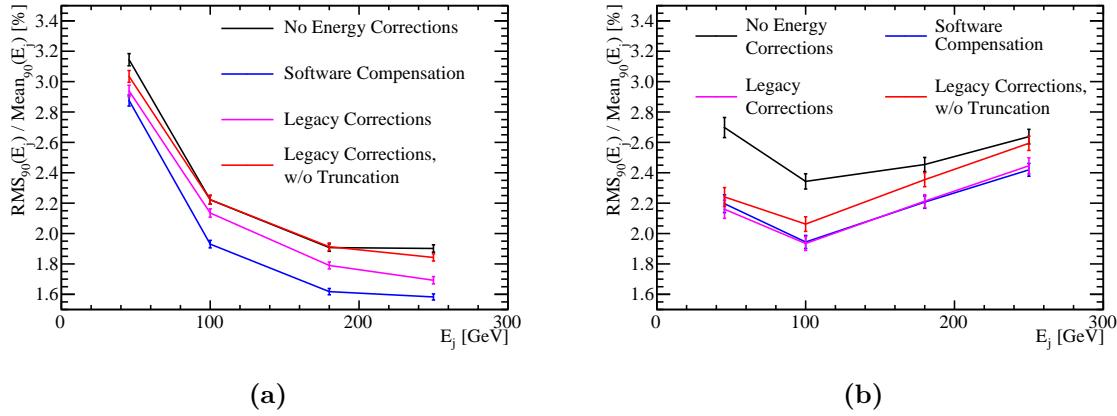


Figure 2.20: The contributions to the jet energy resolution as a function of the jet energy for a variety of different energy correction options. The jet energy resolution contributions presented are (a) the intrinsic energy resolution of the detector and (b) the total confusion contribution. The jet energy resolution obtained using the standard reconstruction is given by the quadrature sum of the intrinsic energy resolution and total confusion contributions. The black line represents no energy corrections, the blue line represents software compensation, the magenta line represents the legacy energy corrections and the red line represents the legacy corrections without the HCal hit energy truncation. The nominal ILD detector model was used in these simulations.

2.3.3 Summary

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

2.4 Timing Cuts

The linear collider will operate using trigger-less readout whereby the recorded data for each sub-detector is read out between collisions of the e^+ and e^- bunches. The bunch train structure for ILC and CLIC is compared in table 2.1. Event selection will proceed through the application of a software trigger. This involves the identification of any hard interactions, prior to full event reconstruction, and only putting data into the event reconstruction if it is measured within a chosen time window about these interactions. The recorded time of a calorimeter hit, which is cut on to make the time window for the software trigger, is corrected for straight time-of-flight to the IP. This ensures that the amount of time particle showers have to develop in the calorimeters is independent of their position in the detector. The energy resolution of a calorimeter is sensitive to the choice of time window applied because energy measurements made outside the time window are rejected. Therefore, the overall detector performance will be sensitive to the choice of time window used.

At CLIC, the application of a software trigger is challenging because of the 0.5 ns bunch separation. The small bunch separation means the integration time of the calorimeters will span many bunch crossings. When this is combined with the intense beam-induced backgrounds, identification of energy deposits produced from a hard interaction of interest becomes difficult. By placing tight timing constraints on the energy deposits made in the CLIC calorimeters, it is possible to minimise the impact of the beam-induced backgrounds. As well as minimising the impact of the backgrounds, these tight timing requirements will also change how particle showers from the hard interaction of interest are sampled. Understanding the impact of these timing requirements on physics performance is vital to the success of the CLIC experiment. Application of the software trigger at the ILC is less challenging than at CLIC because the bunch separation is much larger, meaning the calorimeters could be read out between bunches, and the beam-induced backgrounds are much smaller.

For all choices of time window considered in this study the calibration procedure described in section 2.2 was reapplied. This ensures that the mean of the reconstructed energy distributions will not depend on the calorimeter timing window because the calibration will compensate for energy losses incurred by rejecting energy measurements made outside the time window.

	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 2.1: The train structure for 500 GeV ILC and 3 TeV CLIC [4, 19].

For the results presented in this chapter and the optimisation studies found in chapter ??, a 100 ns timing window was applied to all detector models considered. This value was chosen as it reflects particle shower development time [19] and could be reasonably achieved using readout technology options presently available [36].

2.4.1 Impact on Single Particle Energy Resolution

Figure 2.21 shows the energy resolution of the nominal ILD detector for 100 GeV photons and 50 GeV K_L^0 s as a function of the timing window applied to the calorimeter hits. The timing cut makes little difference to the energy resolution of photons, however, the energy resolution for neutral hadrons gets significantly worse as the time window is reduced. The neutral hadron energy resolution becomes worse by almost 20% when the time window is reduced from 10^6 ns to 10 ns. These trends are to be expected because electromagnetic showers develop far more rapidly than their hadronic counterparts [32]. This can be seen from figure 2.22, which shows the distribution of the measurement time of calorimeter hits, corrected for time-of-flight, for selected shower components for 91 GeV $Z \rightarrow u\bar{d}s$ events. Hadronic showers develop more slowly as they often involve intermediate states that must decay to continue the propagation of the shower.

If a narrow calorimeter timing window is used, energy measurements from the hadronic shower will be lost and the energy resolution will degrade, which is what is observed. On the other hand, electromagnetic showers develop so rapidly that even the 10 ns time window does not reject many energy measurements.

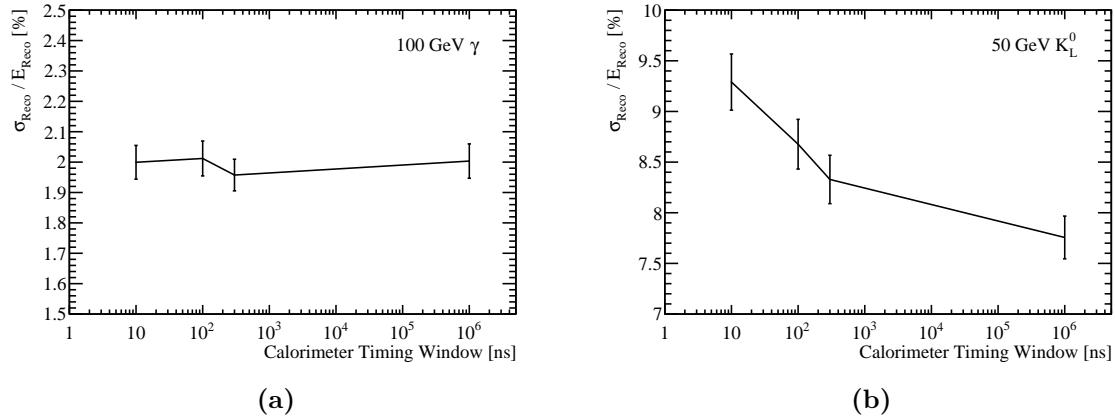


Figure 2.21: The energy resolution as a function of calorimeter timing window for (a) 100 GeV photons and (b) 50 GeV K_L^0 events using the nominal ILD detector model.

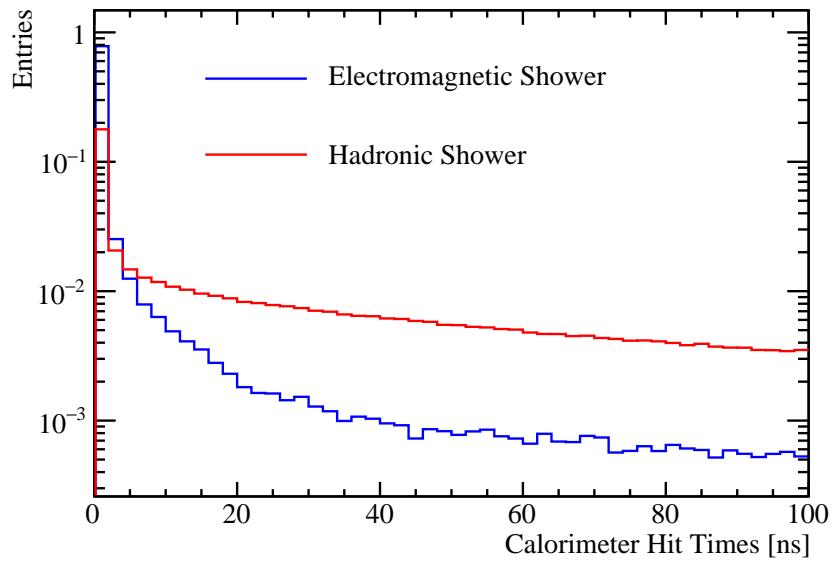


Figure 2.22: The normalised distribution of the time of the electromagnetic and hadronic shower calorimeter hits, corrected for time of flight to the impact point, for 91 GeV $Z \rightarrow u\bar{d}s$ events. Electromagnetic shower energy deposits are deposited very rapidly, while hadronic shower energy deposits are deposited over a much longer time period.

2.4.2 Impact on Jet Energy Resolution

Figure 2.23 shows the jet energy resolution as a function of the jet energy for selected calorimeter time windows. As expected, the jet energy resolution becomes worse when the calorimeter timing window is reduced.

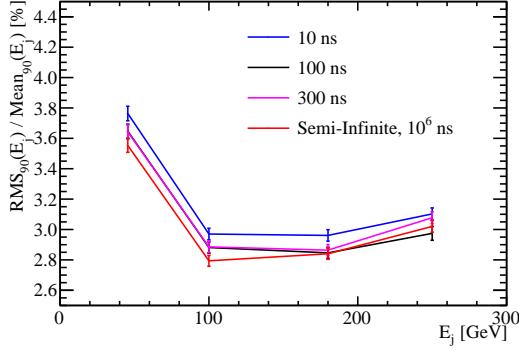


Figure 2.23: The jet energy resolution as a function of jet energy for various calorimeter timing cuts. The nominal ILD detector model was used in these simulations.

The time window applied to the calorimeter hits affects both the neutral hadron and jet energy resolutions with a larger timing window leading to better resolutions. It can be seen that by applying an aggressive choice of time window, such as 10 ns, the jet energy resolution would be degraded because many of the hadronic showers are not fully sampled. 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 at the energies considered will be fully sampled using a 100 ns time window and that there is little to be gained by increasing this time window further.

2.4.3 Summary

Simulations were performed to study the impact of the calorimeter hit time window used in the software trigger at the linear collider experiments. The energy resolution for electromagnetic showers did not change significantly when varying the size of time window, however, the neutral hadron energy resolution becomes worse as the size of the time window is reduced. The jet energy resolution is also sensitive to the size of the time window used, however, the trend was far weaker than that seen for neutral hadrons because only 10% of the jet energy is carried in the form of neutral hadrons. Increasing the time window beyond 100 ns did not have any significant benefit indicating that the majority of hadronic shows are fully sampled in this time.

Colophon

This thesis was made in L^AT_EX 2 _{ϵ} using the “heptesis” class [37]. Feynman diagrams were produced using TikZ-Feynman [38].

Bibliography

- [1] M. A. Thomson, *Particle Flow Calorimetry and the PandoraPFA Algorithm*, Nucl. Instrum. Meth. **A611**, 25 (2009), arXiv:0907.3577.
- [2] J. S. Marshall and M. A. Thomson, *The Pandora Software Development Kit for Pattern Recognition*, Eur. Phys. J. **C75**, 439 (2015), arXiv:1506.05348.
- [3] J. S. Marshall, A. Münnich, and M. A. Thomson, *Performance of Particle Flow Calorimetry at CLIC*, Nucl. Instrum. Meth. **A700**, 153 (2013), arXiv:1209.4039.
- [4] H. Abramowicz *et al.*, *The International Linear Collider Technical Design Report - Volume 4: Detectors*, ILC-REPORT-2013-040 (2013), arXiv:1306.6329.
- [5] Linear Collider ILD Concept Group, T. Abe *et al.*, *The International Large Detector: Letter of Intent*, FERMILAB-LOI-2010-03 (2010), arXiv:1006.3396.
- [6] CMS Collaboration, S. Chatrchyan *et al.*, *Energy Calibration and Resolution of the CMS Electromagnetic Calorimeter in pp Collisions at $\sqrt{s} = 7 \text{ TeV}$* , JINST **8**, P09009 (2013), arXiv:1306.2016.
- [7] ATLAS Electromagnetic Barrel Calorimeter Collaboration, M. Aharrouche *et al.*, *Energy linearity and resolution of the ATLAS electromagnetic barrel calorimeter in an electron test-beam*, Nucl. Instrum. Meth. **A568**, 601 (2006), arXiv:physics/0608012.
- [8] LHCb Collaboration, P. Perret, *First Years of Running for the LHCb Calorimeter System*, PoS **TIPP2014**, 030 (2014), arXiv:1407.4289.
- [9] H. L. Tran *et al.*, *Software compensation in Particle Flow reconstruction*, DESY-17-083 (2017), arXiv:1705.10363.
- [10] CALICE collaboration, 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**, P12006 (2015), arXiv:1509.00617.

- [11] H. S. Budd, *CMS central hadron calorimeter*, Nucl. Phys. Proc. Suppl. **54B**, 191 (1997), arXiv:hep-ex/0102032.
- [12] ATLAS Collaboration, A. Airapetian *et al.*, *ATLAS calorimeter performance Technical Design Report*, CERN-LHCC-96-40 (1996).
- [13] P. Mora de Freitas and H. Videau, *Detector simulation with MOKKA / GEANT4: Present and future*, in *Proceedings, International Workshop on physics and experiments with future electron-positron linear colliders (LCWS)*, LC-TOOL-2003-010 (2002).
- [14] GEANT4 Collaboration, S. Agostinelli *et al.*, *GEANT4: A Simulation toolkit*, Nucl. Instrum. Meth. **A506**, 250 (2003).
- [15] GEANT4 Collaboration, J. Allison *et al.*, *GEANT4 developments and applications*, IEEE Trans. Nucl. Sci. **53**, 270 (2006).
- [16] J. B. Birks, *Scintillations from Organic Crystals: Specific Fluorescence and Relative Response to Different Radiations*, Proc. Phys. Soc. **A64**, 874 (1951).
- [17] CLICdp Collaboration, C. Grefe, S. Poss, A. Sailer, and A. Tsaregorodtsev, *IL-CDIRAC, a DIRAC extension for the Linear Collider community*, J. Phys. Conf. Ser. **513**, 032077 (2014).
- [18] A. Tsaregorodtsev *et al.*, *DIRAC: A community grid solution*, J. Phys. Conf. Ser. **119**, 062048 (2008).
- [19] L. Linssen, A. Miyamoto, M. Stanitzki, and H. Weerts, *Physics and Detectors at CLIC: CLIC Conceptual Design Report*, CERN-2012-003 (2012), arXiv:1202.5940.
- [20] N. Alipour Tehrani *et al.*, *CLICdet: The post-CDR CLIC detector model*, CLICdp-Note-2017-001 (2017).
- [21] F. Gaede, *Marlin and LCCD: Software tools for the ILC*, Nucl. Instrum. Meth. **A559**, 177 (2006).
- [22] H. Abramowicz *et al.*, *Higgs Physics at the CLIC Electron-Positron Linear Collider*, CLICDP-PUB-2016-001 (2016), arXiv:1608.07538.
- [23] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **05**, 026 (2006), arXiv:hep-ph/0603175.
- [24] OPAL Collaboration, G. Alexander *et al.*, *A Comparison of b and u d s quark jets*

- to gluon jets*, Z. Phys. **C69**, 543 (1996).
- [25] Z. Was, *TAUOLA the library for tau lepton decay, and KKMC / KORALB / KORALZ /... status report*, Nucl. Phys. Proc. Suppl. **98**, 96 (2001), arXiv:hep-ph/0011305.
- [26] CMS Collaboration, S. Chatrchyan *et al.*, *Measurement of the $W\gamma$ and $Z\gamma$ inclusive cross sections in pp collisions at $\sqrt{s} = 7$ TeV and limits on anomalous triple gauge boson couplings*, Phys. Rev. **D89**, 092005 (2014), arXiv:1308.6832.
- [27] ATLAS Collaboration, G. Aad *et al.*, *Measurement of W^+W^- production in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector and limits on anomalous WWZ and $WW\bar{t}\gamma$ couplings*, Phys. Rev. **D87**, 112001 (2013), arXiv:1210.2979.
- [28] CMS Collaboration, S. 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**, 032008 (2014), arXiv:1404.4619.
- [29] C. W. Fabjan and F. Gianotti, *Calorimetry for particle physics*, Rev. Mod. Phys. **75**, 1243 (2003).
- [30] CALICE Collaboration, C. Adloff *et al.*, *Hadronic energy resolution of a highly granular scintillator-steel hadron calorimeter using software compensation techniques*, JINST **7**, P09017 (2012), arXiv:1207.4210.
- [31] M. Derrick *et al.*, *Design and construction of the ZEUS barrel calorimeter*, Nucl. Instrum. Meth. **A309**, 77 (1991).
- [32] R. Wigmans, *Calorimetry: Energy measurement in particle physics*, Int. Ser. Monogr. Phys. **107**, 1 (2000).
- [33] L. Landau, *On the energy loss of fast particles by ionization*, J. Phys. (USSR) **8**, 201 (1944).
- [34] Particle Data Group, S. Eidelman *et al.*, *Review of particle physics*, Phys. Lett. **B592**, 1 (2004).
- [35] O. Hartbrich, *AHCAL Digitisation*, CALICE Spring Meeting 2015 & LCWS 2015, PUBDB-2015-05964 (2015).
- [36] CALICE Collaboration, C. Adloff *et al.*, *The Time Structure of Hadronic Showers in highly granular Calorimeters with Tungsten and Steel Absorbers*, JINST **9**, P07022

- (2014), arXiv:1404.6454.
- [37] A. Buckley, *The hepthesis L^AT_EX class*.
- [38] J. Ellis, *TikZ-Feynman: Feynman diagrams with TikZ*, Comput. Phys. Commun. **210**, 103 (2017), arXiv:1601.05437.