Calorimetry at a Future Linear Collider

Steven Green
of Emmanuel College

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

Abstract

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

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

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

*"Writing in English is the most ingenious torture
ever devised for sins committed in previous lives."*
   — James Joyce

Chapter 1

Calorimeter Optimisation Studies

"The simple believes everything, but the prudent gives thought to his steps."

  — Proverbs 14:15

1.1 Introduction

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

1.2 Electromagnetic Calorimeter Optimisation

The purpose of an electromagnetic calorimeter (ECal) is to measure the energy deposits from electromagnetic showers. The nominal ILD ECal, summarised in table 1.1, is a silicon-tungsten sampling calorimeter. It contains 29 readout layers and 24 radiation lengths (X_0), which is sufficient to contain all but the highest energy electromagnetic showers. The absorber thickness of the last nine layers is twice that of the first 20 layers to reduce the number of readout channels and cost of the calorimeter. The high

longitudinal sampling frequency is crucial for the pattern recognition aspect of particle flow calorimetry, especially in the region where particle showers start developing.

Parameter	Default Value
Cell Size	5×5 mm^2 square cells
Number of Layers	29 readout layers
Active Material Choice	Silicon or Scintillator
Active Material Thickness	0.5 mm (Silicon) or 2 mm (Scintillator)
Absorber Material Choice	Tungsten
Absorber Material Thickness	20 layers of 2.1 mm followed by 9 layers of 4.2 mm

Table 1.1: The configuration of the silicon and scintillator ECal options for the ILD detector model [1].

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

- Cell size: This is a vital aspect of the detector in the particle flow paradigm as smaller cell sizes leads to better separation between nearby showering particles, which helps to minimise the effect of confusion. Modifying the cell size should have little effect on the intrinsic energy resolution of the detector.

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

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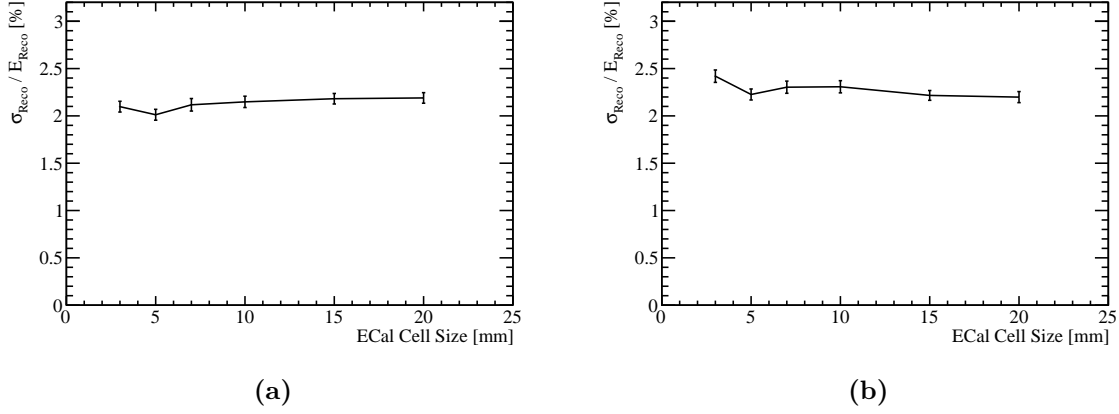
1.2.1 ECal Cell Size

Different detector models were considered where the cell size in the ECal was varied about the nominal value of 5×5 mm^2 square cells. The granularities considered were 3×3 mm^2, 5×5 mm^2, 7×7 mm^2, 10×10 mm^2, 15×15 mm^2 and 20×20 mm^2 square cells for both the silicon and scintillator active material options.

The energy resolution, using 100 GeV photons, as a function of the ECal cell size is shown in figure 1.2a for the silicon option and in figure 1.2b for the scintillator option. At this energy, the photons will be largely contained within the ECal and the reported energy resolution reflects solely the performance of the ECal. For both the silicon and scintillator ECal options the energy resolution does not depend strongly on the ECal cell size. This is to be expected as there is no change in the number of layers, which is the main factor in determining the energy resolution of a sampling calorimeter.

The only statistically significant variation observed occurs for the scintillator ECal option. A degradation in the energy resolution of $\sim 10\%$ is observed when reducing the ECal cell size from 5×5 mm^2 to 3×3 mm^2. The most likely cause is the "dead" region in the active material, which represents the readout multi pixel photon counter (MPPC) [2]. The MPPC occupies a fixed area of the cell, irrespective of cell size, and so the dead region of the cell fractionally increases as cell size is reduced. The larger this dead region, the worse the sampling of the electromagnetic showers in the ECal and the worse the resolution. While this effect will be present in all scintillator ECal options, it will only be significant for the small cell sizes when the dead region is fractionally the largest.

The ability to separate nearby electromagnetic particle showers within a calorimeter is limited by the Moliére radius of the absorber material and the cell size. The Moliére radius controls the width of the electromagnetic shower, while the cell size controls how the transverse shower profile is sampled. By reducing the cell size, it becomes easier to resolve nearby electromagnetic showers, which in turn reduces the effect of confusion. Therefore, it is expected that the jet energy resolution will be sensitive to the ECal cell size, even though the intrinsic energy resolution is not. The jet energy resolution as a function of ECal cell size is shown in figure 1.2a for the silicon option and figure 1.2b for the scintillator option. There is a strong dependance on the ECal cell size, with smaller cell sizes leading to lower values of the jet energy resolution; the jet energy resolution for 250 GeV jets for both ECal options goes from $\sim 3.0\%$ to $\sim 4.3\%$ when the ECal cell size goes from 3×3 mm^2 to 20×20 mm^2. The origin of this trend is best illustrated by

Figure 1.1: The energy resolution as a function of ECal cell size for 100 GeV photons using the nominal ILD detector model with (a) the silicon and (b) the scintillator ECal option.

considering the intrinsic energy resolution and confusion contributions to the jet energy resolution. These contributions are shown as a function of ECal cell size for 45 and 250 GeV jets in figure 1.3. It is clear from these contributions that the intrinsic energy resolution of the detector does not change when varying the cell size, which agrees with both prior expectations of calorimeter behaviour and the single particle energy resolution study. As expected, it can be seen that the trend in jet energy resolution as a function of the ECal cell size is being driven purely by changes to the confusion contribution and, in particular, the confusion caused by the reconstruction of photons.
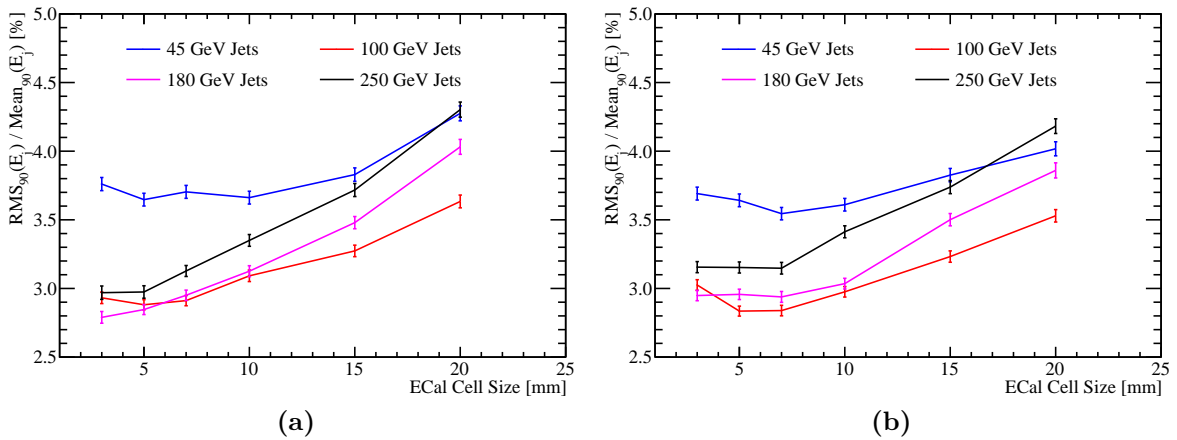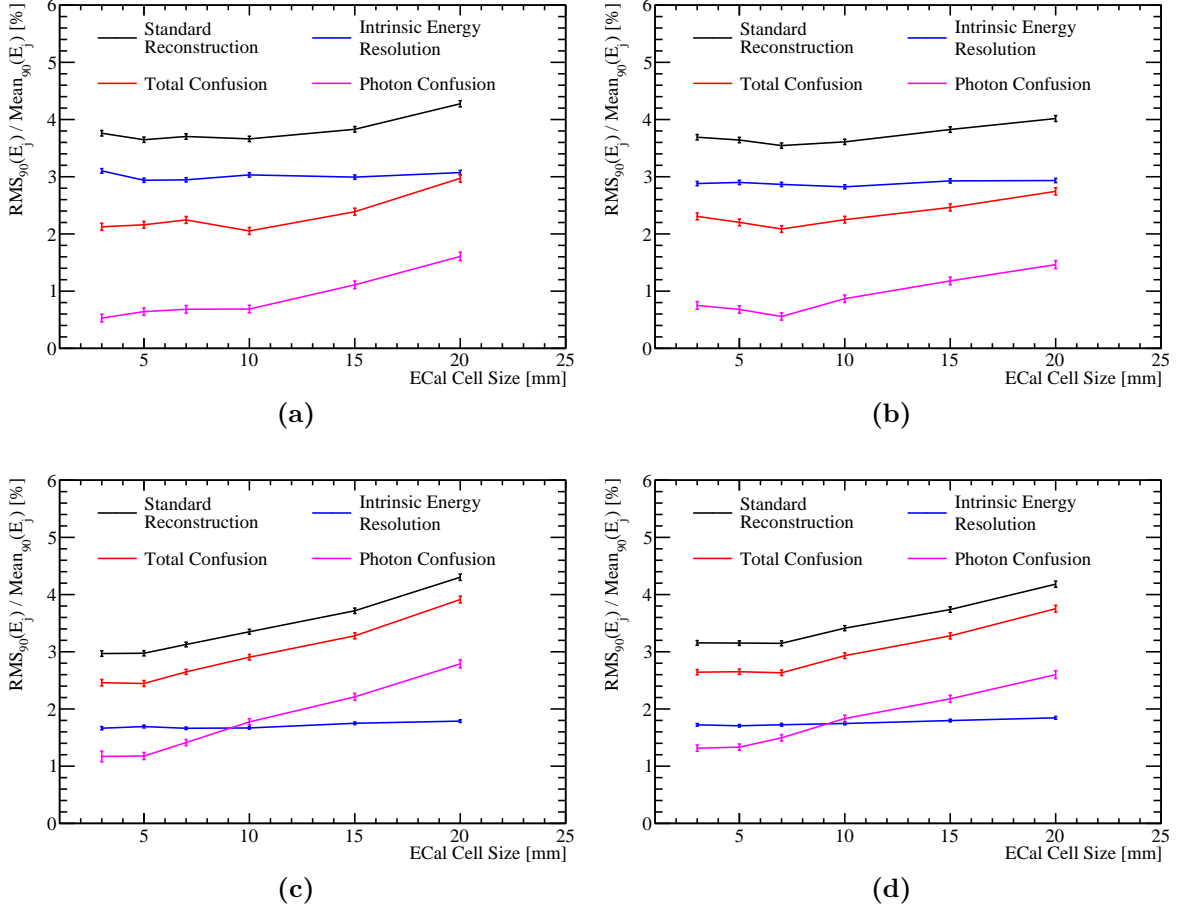

Figure 1.2: The fractional jet energy resolution as a function of ECal cell size for various jet energies using the nominal ILD detector model with (a) the silicon and (b) the scintillator ECal option.

Figure 1.3: Contributions to the jet energy resolution shown as function of ECal cell size using the nominal ILD detector model for (a) the silicon ECal option and 45 GeV jets, (b) the scintillator ECal option and 45 GeV jets, (c) the silicon ECal option and 250 GeV jets and (d) the scintillator ECal option and 250 GeV jets. The black curves correspond to the standard reconstruction, the blue curves to the intrinsic energy resolution contribution to the jet energy resolution, the red curves to the confusion contribution to the jet energy resolution and the magenta curves to the confusion contribution to the jet energy resolution related solely to photon reconstruction

It is clear that the ECal cell size is extremely important for jet energy measurements, although it has little bearing on the intrinsic energy resolution of the ECal. Separation of the hadronic decays of the W and Z bosons, i.e. $\sigma_E/E \lesssim 3.8\%$ [3], can be achieved across the jet energy range considered here using a maximum ECal cell size of 15×15 mm^2. However, as reducing the ECal cell size further continues to benefit the jet energy resolution, minimising the ECal cell size is desirable.

1.2.2 ECal Longitudinal Sampling Frequency

The detector performance was simulated where the number of layers in the ECal was varied, while keeping the total material budget (X_0) approximately constant. This study was performed for both the silicon and scintillator active material options. In all cases tungsten was used for the ECal absorber material and the active layer thicknesses were not changed from those used in the nominal ILD ECal summarised in table 1.1. The different ECal layouts considered are summarised in table 1.2.

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

Table 1.2: The longitudinal structure of the ECal models considered in the optimisation study. The radiation length of tungsten absorber is 3.504 mm [4]. Note that a presampler layer contributes one extra layer to the cumulative number of layers.

The energy resolution, for 100 GeV photons, as a function of the number of layers in the ECal is shown in figure 1.4a for the silicon option and in figure 1.4b for the scintillator option. When the number of layers is increased σ_E/E decreases, which is expected because the energy resolution for a sampling calorimeter is $\propto 1/\sqrt{E \times N_{Layers}}$, where E is the reconstructed energy and N_{Layers} is the number of layers in the calorimeter.

When the number of layers in the ECal is increased, the intrinsic energy resolution benefits; the intrinsic energy resolution of the ECal improves by $\sim 25\%$ in both ECal options when increasing the number of layers from 16 to 30. This has the knock-on effect of reducing the confusion contribution to the jet energy resolution, which can be seen in figures 1.5a and 1.5b for the silicon and scintillator ECal options respectively. In both cases, the jet energy resolution was found to improve when the number of layers in the ECal was increased; the jet energy resolution goes from ~ 4.4 to $\sim 3.6\%$ for the silicon option and from ~ 4.1 to $\sim 3.6\%$ for the scintillator option when increasing the number of layers from 16 to 30. The magnitude of the change in jet energy resolution is dependent upon the jet energy, with a stronger dependancy being observed for low
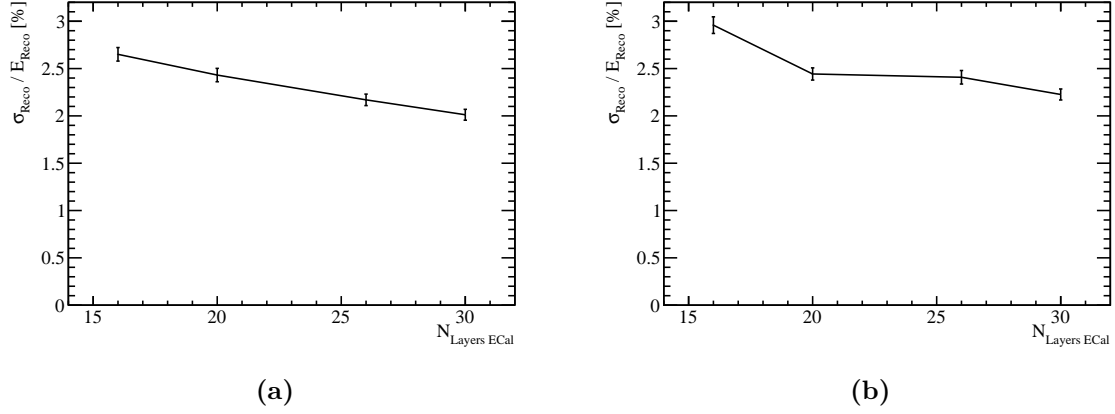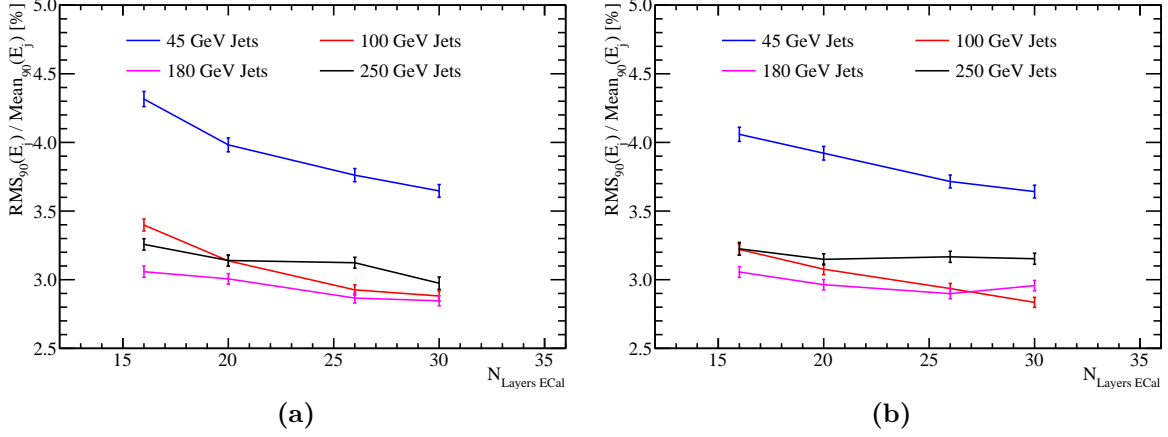
Figure 1.4: The energy resolution as a function of number of layers in the ECal for 100 GeV photons using the nominal ILD detector model with (a) the silicon and (b) the scintillator ECal option.

energy jets. This is expected from the stochastic contribution to the energy resolution for a sampling calorimeter. For high jet energies, changing the number of layers in the ECal does not significantly affect the jet energy resolution because the jet energy resolution is dominated by confusion. For low jet energies, the stochastic contribution to the energy resolution is bigger making it possible to resolve the changes to it when varying the number of layers in the ECal.

The decomposition of the jet energy resolution into the intrinsic energy resolution and confusion contributions for 45 and 250 GeV jets are shown, for both the silicon and scintillator ECal options, in figure 1.6. As expected, the improvement to the intrinsic energy resolution seen when increasing the number of layers in the ECal leads to the knock-on effect of lowering the confusion. However, significantly the magnitude of the change to the intrinsic energy resolution and confusion contributions to the jet energy resolution when varying the number of layers in the ECal are comparable in size. This shows that pattern recognition is as important for detector performance in the particle flow paradigm than intrinsic energy resolution.

1.2.3 ECal Active Material

In sections 1.2.1 and 1.2.2 the performance of the ECal was reported for both the silicon and scintillator options and to a large extent the performance of the two options was similar, but not identical:
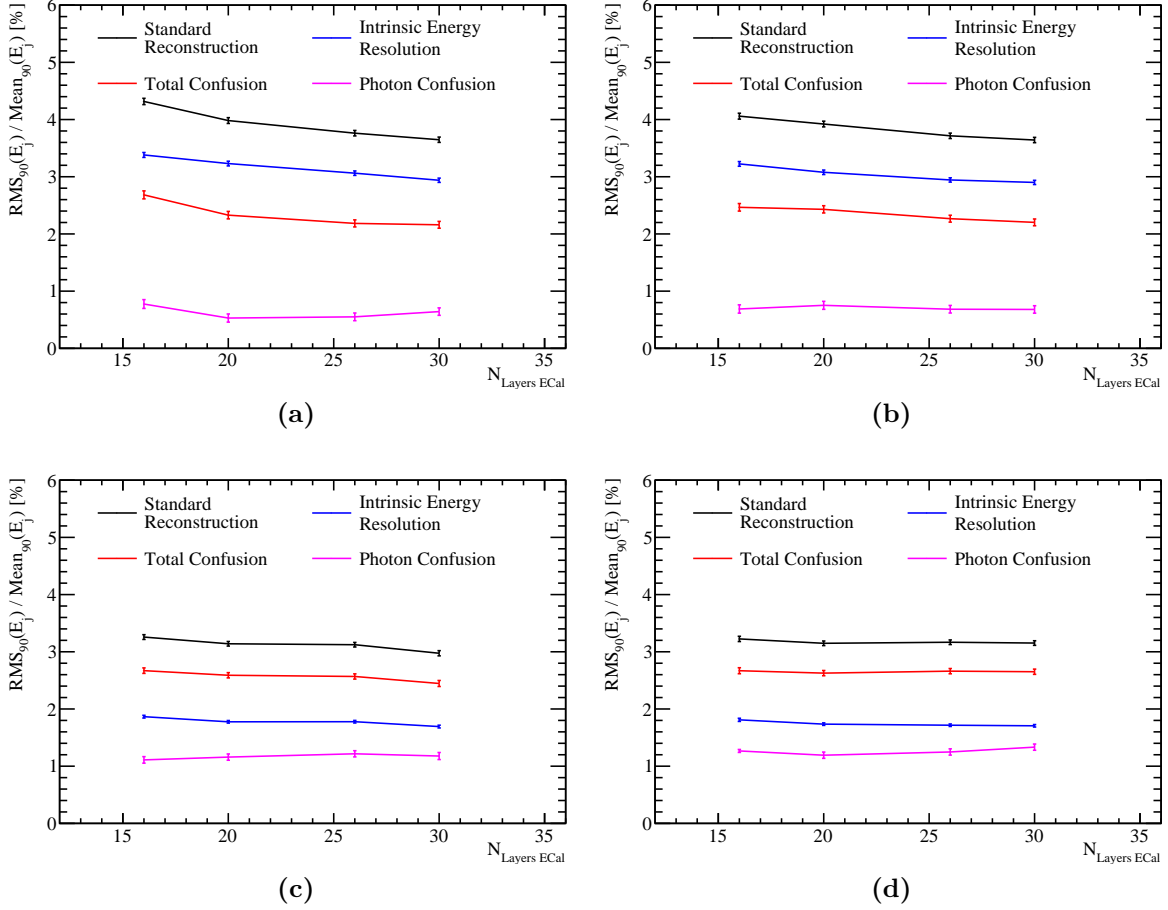
Figure 1.5: The jet energy resolution as a function of number of layers in the ECal for various jet energies using the nominal ILD detector model with (a) the silicon and (b) the scintillator ECal option.

- The intrinsic energy resolution of the silicon ECal option is better than that of the scintillator option at very high energies. For 500 GeV photons the intrinsic energy resolution is $\sim 25\%$ better for the silicon option. Section **??** contains a comparison between the photon energy resolution for the two ECal options, which clearly illustrates this. The most likely origin of the differing energy resolutions is the implementation of Birks' law [5] for scintillator active materials, which states

$$d\mathcal{L}/dx \propto \frac{dE/dx}{1 + k_B dE/dx} \ , \tag{1.1}$$

where $d\mathcal{L}/dx$ is the scintillation light yield per unit path length, dE/dx is the energy deposited per unit path length and k_B is a material property constant. For large energy deposits per unit length, such as those found in high energy photons, the light yield saturates causing a degradation in the energy resolution. When comparing the photon energy resolution for the silicon and scintillator ILD ECal options, which can be found in section **??**, the saturation effect starts to degrade the energy resolution for the scintillator option around 50 GeV. However, the degradation in energy resolution is relatively small event up to 100 GeV.

- The "dead" region due to the presence of the MPPC in the simulation of the scintillator ECal option degrades performance of the detector for small transverse granularities, see figure 1.1.

Figure 1.6: Contributions to the jet energy resolution shown as function of number of layers in the ECal using the nominal ILD detector model for (a) the silicon ECal option and 45 GeV jets, (b) the scintillator ECal option and 45 GeV jets, (c) the silicon ECal option and 250 GeV jets and (d) the scintillator ECal option and 250 GeV jets. The black curves correspond to the standard reconstruction, the blue curves to the intrinsic energy resolution contribution to the jet energy resolution, the red curves to the confusion contribution to the jet energy resolution and the magenta curves to the confusion contribution to the jet energy resolution related solely to photon reconstruction.

In summary, the performance of the two options, in terms of energy and jet energy resolution, at ILC-like energies is comparable. However, the silicon option is preferred when manufacture and implementation of the two models is compared. While constructing silicon wafers to fit a 5×5 mm^2 square cell size is achievable, this would be extremely challenging for scintillator tiles. To resolve this in actuality, the scintillator ECal option would have to use 5×45 mm^2 scintillator strips that are arranged in alternating directions in each ECal layer [1]. By combining information from neighbouring layers it becomes possible to approach an effective 5×5 mm^2 square cell size.

1.3 Hadronic Calorimeter Optimisation

The purpose of an hadroinc calorimeter (HCal) is to measures the energy deposits from hadronic showers. The HCal in the default ILD detector model, summarised in table 1.3, is approximately 6 nuclear interaction lengths (λ_I) deep. The ECal contributes approximately one λ_I giving a total of $\approx 7\lambda_I$, which is sufficient to contain jets at ILC like energies. The longitudinal structure of this model consists of 48 readout layers each containing a 3 mm active layer of scintillator and a 20 mm absorber layer of iron.

Parameter	Default Value
Cell Size	30×30 mm^2 square cells
Number of Layers	48 readout layers
Active Material Choice	Scintillator
Active Material Thickness	3 mm
Absorber Material Choice	Steel
Absorber Material Thickness	20 mm

Table 1.3: The configuration of the HCal in the nominal ILD detector model [1].

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

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

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

- Number of readout layers: The number of layers in the HCal are varied, however, the thickness of those layers match those of the nominal ILD HCal design. This means the total depth of the HCal in λ_I is changing. It is expected that this study will determine the effect of leakage of energy out of the back of the HCal.

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

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

- Absorber material choice: Two options have been considered: steel and tungsten. This choice affects the growth and propagation of hadronic showers.

1.3.1 HCal Cell Size

The HCal cell size is an important detector parameter in the application of particle flow calorimetry. Smaller HCal cell sizes will lead to a finer spatial resolution that can be used to better separate charged and neutral particle calorimetric energy deposits. On the other hand, this will also lead to an increase in the number of readout channels that will raise the cost of the calorimeter. Therefore, it is highly desirable to achieve the optimal physics performance using the largest cell size possible. The nominal ILD HCal has a 30 mm square cell size and in this study the following cell sizes were considered; 10×10 mm^2, 20×20 mm^2, 30×30 mm^2, 40×40 mm^2, 50×50 mm^2 and 100×100 mm^2.

In the nominal ILD detector, 50 GeV long-lived neutral kaons (K$_L^0$s) will deposit $\sim 65\%$ of their energy in the HCal and $\sim 35\%$ in the ECal. As 50 GeV K$_L^0$s deposit the bulk of their energy in the HCal, they are appropriate to use when determining

the performance of the HCal. However, it should be emphasised that the K_L^0 energy resolutions represent the intrinsic energy resolution of the whole ILD detector and not purely that of the HCal.

Figure 1.7 shows the energy resolution for 50 GeV K_L^0s as a function of cell size. As expected, the hadronic energy resolution does not strongly depend on the HCal cell size. The only statistically significant variation in energy resolution is observed for the 100×100 mm^2 HCal cell size. For this model the energy resolution gets gets worse by $\sim 8\%$ in comparison to the other models considered. The most likely cause of this is a reduction in the effectiveness of the HCal hit energy truncation, which is described in section **??**. The reduced effectiveness is expected because the precision used when obtaining the optimal energy truncation becomes worse as HCal cell size diverges from the nominal value.
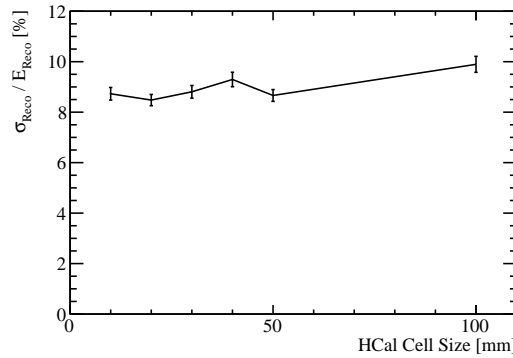
Figure 1.7: The energy resolution as a function of HCal cell size for 50 GeV K_L^0 events using the nominal ILD detector model.

A smaller HCal cell size will lead to better separation of charged and neutral hadron calorimetric energy deposits, therefore, it is expected that the confusion contribution to the jet energy resolution will be reduced by using smaller HCal cell sizes. Figure 1.8 shows the jet energy resolution as a function of cell size in the HCal. At low jet energies there is no strong dependency of the jet energy resolution on the HCal cell size, which is as expected from the K_L^0 energy resolution study. For high energy jets there is a clear dependence, with lower HCal cell sizes leading to better jet energy resolutions; the jet energy resolution for 250 GeV jets goes from $\sim 2.7\%$ to $\sim 3.5\%$ when the HCal cell size is increased from 10×10 mm^2 to 100×100 mm^2. Examining the different contributions to the jet energy resolution, shown in figure 1.9 it can be seen that the intrinsic energy resolution contribution does not depend on the HCal cell size; it is the confusion contribution that drives the overall trend in the jet energy resolution.

This is particularly clear at high jet energies where the confusion contribution to the jet energy resolution dominates that of the intrinsic energy resolution contribution. At high jet energies smaller HCal cell sizes leads to a reduction in the effect of confusion; the confusion contribution to the jet energy resolution is reduced by $\sim 25\%$ when reducing the HCal cell size from 100×100 mm^2 to 10×10 mm^2. At low jet energies the trend is less clear, as the confusion contribution is less dominant. Nevertheless, a reduction in the effect of confusion with decreasing cell size is still visible for all but the smallest HCal cell size. The most likely cause of the increase in confusion for the smallest HCal cell size at low energies is the tuning of the PandoraPFA algorithms to the nominal ILD HCal cell size. For both the 45 and 250 GeV jets, the photon confusion does not depend on the HCal cell size. This indicates that changes to the confusion term seen when varying the HCal cell size are related solely to the reconstruction of hadrons.
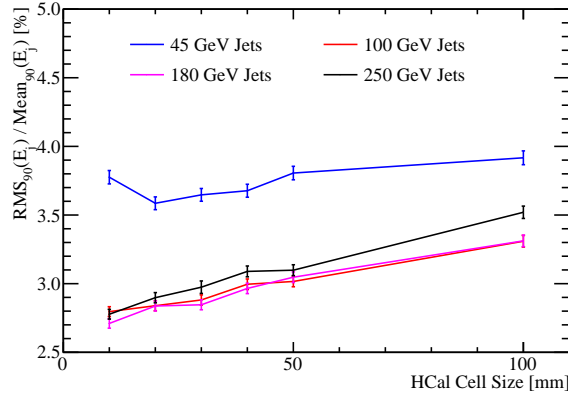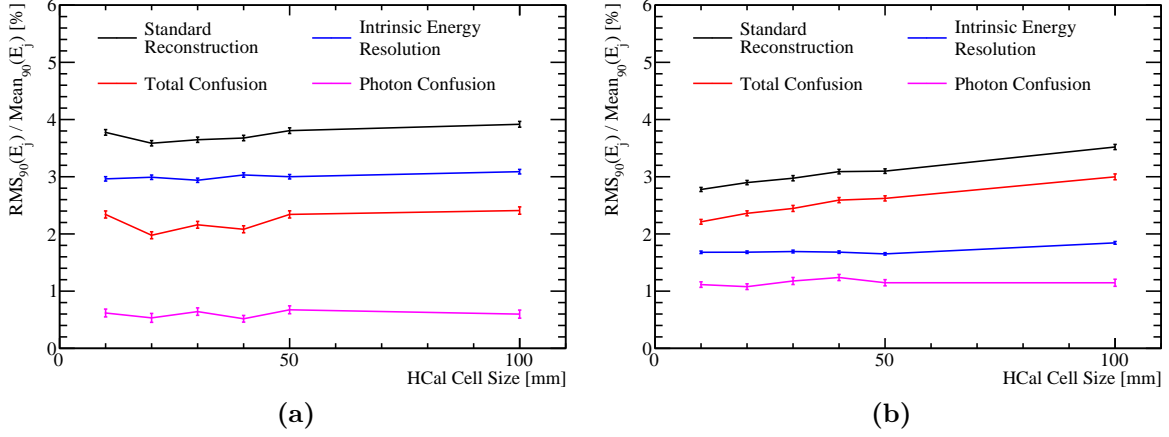

Figure 1.8: The jet energy resolution as a function of HCal cell size for various jet energies using the nominal ILD detector model.

A comparison of the results from the ECal and HCal cell size optimisation studies shows that the jet energy resolution has a stronger dependency on the ECal cell size than on the HCal cell size; increasing the nominal ECal cell size by a factor of three makes the jet energy resolution for 250 GeV jets worse by $\sim 20\%$, while increasing the nominal HCal cell size by the same factor makes the jet energy resolution worse by $\sim 12\%$. This is to be expected as in the particle flow paradigm $\approx 30\%$ of jet energy is recorded in the ECal, while only $\approx 10\%$ is recorded in the HCal. Consequently, the potential effect of double counting and omitting energy deposits, i.e. confusion, is greater in the ECal than the HCal. Therefore, minimising confusion in the ECal is expected to be more crucial for the overall jet energy resolution, which is what is observed. Furthermore, as PandoraPFA groups calorimeter hits together using a cone clustering approach, identifying the start of a particle shower is key for determining how calorimeter hits are grouped together deeper

Figure 1.9: Contributions to the jet energy resolution shown as function of HCal cell size using the nominal ILD detector model for (a) 45 GeV jets and (b) 250 GeV jets. The black curves correspond to the standard reconstruction, the blue curves to the intrinsic energy resolution contribution to the jet energy resolution, the red curves to the confusion contribution to the jet energy resolution and the magenta curves to the confusion contribution to the jet energy resolution related solely to photon reconstruction.

into the calorimeters. In effect, this means the grouping of calorimeter hits in the HCal depends upon information gathered in the ECal. Therefore, if the ECal performance is sufficiently good, even with coarse HCal cell sizes, excellent performance can be achieved.

In summary, the confusion contribution to the jet energy resolution falls as the HCal cell size is reduced, while the intrinsic energy resolution of the detector is largely unaffected. As this dependancy is relatively weak, even the use of 100×100 mm^2 HCal cell sizes would be enough to allow for separation of the hadronic decays of W and Z bosons, i.e. $\sigma_E/E \lesssim 3.8\%$ [3], at ILC like energies. However, there are benefits to having smaller HCal cell size; the jet energy resolution is reduced from $\sim 3.5\%$ to $\sim 2.8\%$ for 250 GeV jets when decreasing the HCal cell size from 100×100 mm^2 to 10×10 mm^2.

1.3.2 HCal Number of Layers

In this study, the total number of layers in the HCal was varied. In contrast to the longitudinal sampling frequency study, the active and absorber layer thicknesses in the HCal were not altered. Changing the number of layers in this way leads to a change in the total thickness of the calorimeter. This study is sensitive to the effects, if any, of leakage of energy out of the back of the calorimeters. The manufacturing cost of the HCal

is proportional to the number of readout channels and layers. Therefore, minimising the number of layers, while retaining excellent physics performance is important. Here detector models were simulated with a HCal containing 36, 42, 48 (nominal), 54 and 60 layers.

It is expected that the energy resolution of the detector will improve when the number of layers in the HCal is increased since fewer events should suffer from the effects of leakage. Any improvements seen by increasing the number of layers in the HCal is expected only up to the point where the majority of hadronic showers are fully contained by the calorimeters. The energy resolution as a function of number of layers in the HCal for 50 GeV K_L^0 is shown in figure 1.10. The energy resolution becomes worse as the number of layers in the HCal is reduced below 48 layers, while above this point additional layers do not change the energy resolution. This indicates that the majority of hadronic showers at this energy are fully contained by a 48 layer HCal. As reducing the number of HCal layers to 36 only causes a small degradation, $\sim 10\%$, in the neutral hadron energy resolution, it is feasible to consider reducing the number of layers in the ILD HCal.
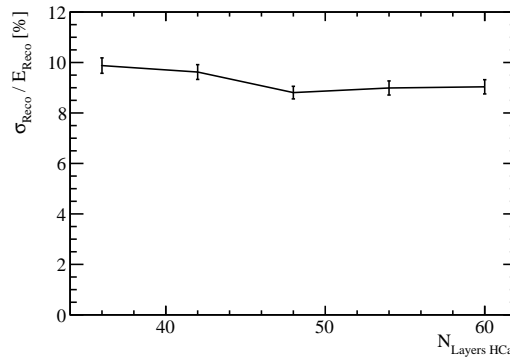
Figure 1.10: The energy resolution as a function of number of layers in the HCal for 50 GeV K_L^0 events using the nominal ILD detector model.

Figure 1.11 shows the jet energy resolution as a function of the number of layers in the HCal. For low energy jets, where intrinsic energy resolution dominates, the jet energy resolution is does not depend on the number of layers in the HCal. At high jet energies, where confusion dominates, increasing the number of layers in the HCal improves the jet energy resolution; the jet energy resolution is goes from $\sim 3.4\%$ to $\sim 3.0\%$ for 250 GeV jets when increasing the number of HCal layers from 36 to 48. The origin of these trends is leakage of energy out of the back of the calorimeters, which becomes more problematic as the number of layers in the HCal is reduced and the jet energy increases.

Figure 1.12 shows the jet energy resolution contributions as a function of the number of layers in the HCal. These results appear somewhat counterintuitive in that the intrinsic energy resolution of the detector does not seem to depend on the number of layers in the HCal even for high energy jets. However, this is expected given only 10% of jet energy is carried in the form of neutral hadrons and the neutral hadron energy resolution, for 50 GeV hadrons, is only weakly dependent on the number of HCal layers. Leakage does have an effect on the intrinsic jet energy resolution, however, the use of RMS_{90} obscures part of this by excluding events where leakage is significant. The fractional decrease in RMS_{90} for the intrinsic jet energy distribution when increasing the number of HCal layers from 36 to 60 is $\sim 4\%$, however, the change in the full RMS is $\sim 23\%$. Figure 1.12 also shows that the confusion contribution is far more sensitive to the number of layers in the HCal than the intrinsic energy resolution. This sensitivity originates from the reclustering stage of the reconstruction in events where leakage has occurred. In these events, when PandoraPFA compares the momentum of a charged particle track to the cluster of calorimeter hits that it produces, there will be a disparity. To resolve the disparity, PandoraPFA will associate other calorimeter energy deposits that were not produced by the charged particle to the track to compensate for the leaked energy, which produces confusion. As photons are largely contained within the ECal at these energies, the photon confusion contribution to the jet energy resolution has no dependence on the number of layers in the HCal.
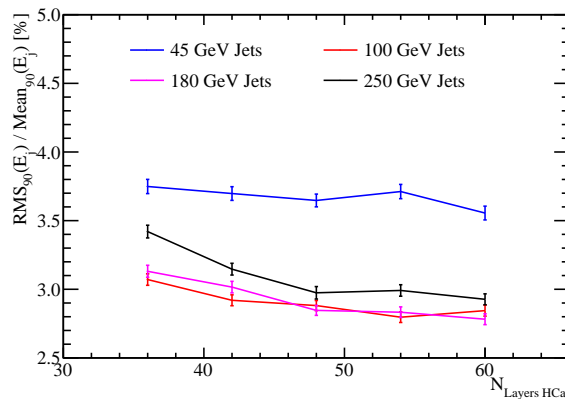
Figure 1.11: The jet energy resolution as a function of number of layers in the HCal for various jet energies using the nominal ILD detector model.
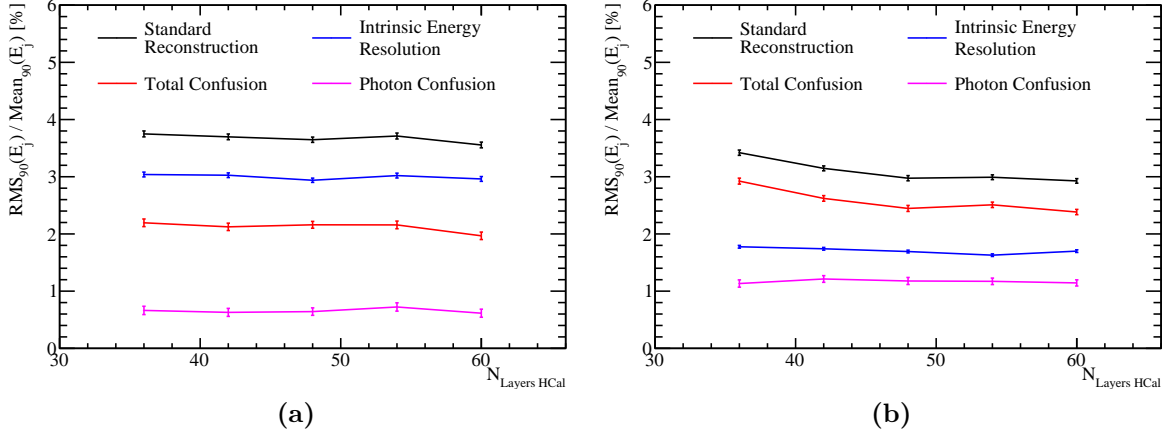
In summary, even if the number of layers in the HCal were reduced by 25%, the jet energy resolution would be sufficient for separating the hadronic decays of the W and Z bosons at ILC energies, i.e. $\sigma_E/E \lesssim 3.8\%$ [3]. Although, the effects of leakage do make the jet energy resolution worse for ILC like energies, once the number of layers in the

Figure 1.12: Contributions to the jet energy resolution shown as function of number of layers in the HCal using the nominal ILD detector model for (a) 45 GeV jets and (b) 250 GeV jets. The black curves correspond to the standard reconstruction, the blue curves to the intrinsic energy resolution contribution to the jet energy resolution, the red curves to the confusion contribution to the jet energy resolution and the magenta curves to the confusion contribution to the jet energy resolution related solely to photon reconstruction.

HCal is reduced from 48 layers, therefore, it is desirable to have a minimum of 48 layers in the ILD HCal.

1.3.3 HCal Longitudinal Sampling Frequency

Several detector models were simulated where the longitudinal sampling frequency in the HCal was modified. The longitudinal sampling frequency was altered by changing the number of layers in the HCal, while simultaneously changing the active and absorber layer thicknesses, to maintain the total number of nuclear interaction lengths. For each model considered, the absorber material was steel, containing a total of 5.72 λ_I, and the active material was scintillator, containing a total of 0.19 λ_I. The ratio of the active to absorber layers thicknesses (the sampling fraction) in these models is the same as in the nominal ILD HCal. A summary of the detector models considered is given in table 1.4.

Figure 1.13 shows the energy resolution for 50 GeV K_L^0 as a function of number of layers in the HCal. As the number of layers in the HCal is increased, the energy resolution improves. This is because increasing the number of layers in a sampling calorimeter, while leaving the total material budget unchanged, will lead to greater sampling of particles

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

Table 1.4: Longitudinal configuration of the HCal in the detector models considered.

showering within it and a reduction the stochastic contribution to the energy resolution. The energy resolution is less pronounced than the naive expectation of $1/N_{\text{HCal}}$, where N_{HCal} is the number of layers in the HCal, because this relationship only holds for the energy resolution of a single sampling calorimeter and these results are for the full ILD detector, including the $\approx 1\lambda_I$ in the ECal. Furthermore, the $1/N_{\text{HCal}}$ functional form neglects a number of effects, such as instrumentation defects and electrical noise, that should be included when parameterising the energy resolution [6].
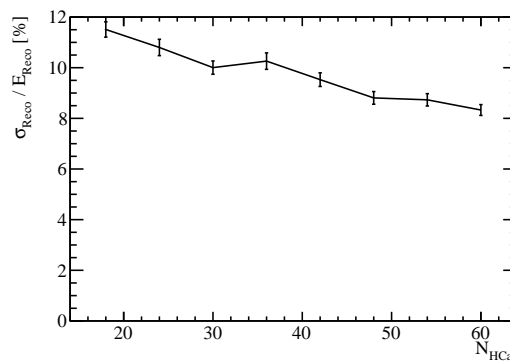
Figure 1.13: The energy resolution as a function of the longitudinal sampling frequency in the HCal for 50 GeV K_L^0 events using the nominal ILD detector model.

Figure 1.14 shows the jet energy resolution as a function of the longitudinal sampling frequency in the HCal. Increasing the number of layers in the HCal leads to an improvement in the HCal; when the number of layers in the HCal is increased from 18 to 60 the jet energy resolution for 250 GeV jets improves by $\sim 17\%$.

Figure 1.15 shows that both the intrinsic energy resolution and confusion improve with increasing longitudinal sampling frequency. For 250 GeV jets, when increasing the number of HCal layers from 18 to 60 the intrinsic energy resolution contribution goes from $\sim 1.9\%$ to $\sim 1.6\%$ and the confusion contribution goes from $\sim 3.0\%$ to $\sim 2.4\%$. The twofold improvement is expected because increasing the longitudinal sampling frequency improves the intrinsic energy resolution of a sampling calorimeter, which has the knock-on effect of lowering the confusion. The resulting reduction in confusion is due to the improved precision obtained when comparing the momenta of charged particle tracks and the energy of clusters of calorimeter hits. These comparisons are used to guide event reconstruction in PandoraPFA, therefore, if the precision of these comparisons is improved, the confusion is reduced as described in section **??**.
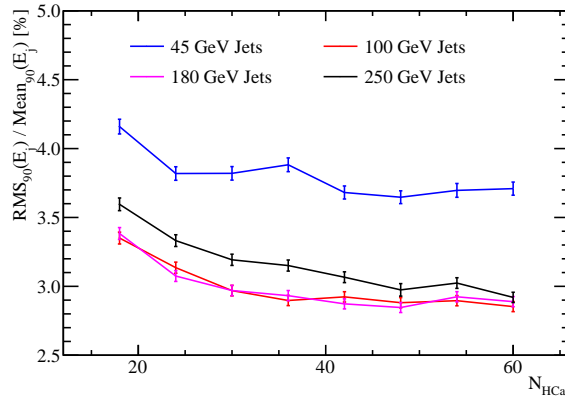
Figure 1.14: The jet energy resolution as a function of longitudinal sampling frequency in the HCal for various jet energies using the nominal ILD detector model.

It is clear that a larger number of layers in the HCal benefits both the intrinsic energy resolution of the ILD detector as well as reducing the confusion contribution to the jet energy resolution. As there are few physics analyses that rely on the identification and categorisation of individual neutral hadrons, but there are many that rely on identification and categorisation of photons, the intrinsic energy resolution of the HCal is less crucial from a physics perspective than that of the ECal. However, these studies show the HCal has a crucial role to play in jet reconstruction in the particle flow paradigm. To achieve a jet energy resolution of $\sigma_E/E \lesssim 3.8\%$ [3], which is required to separate the W and Z hadronic decays, the ILD detector will require a minimum of 42 layers in the HCal. This longitudinal sampling frequency is required particularly for low energy jets where the energy resolution is dominated by the intrinsic energy resolution of the detector.
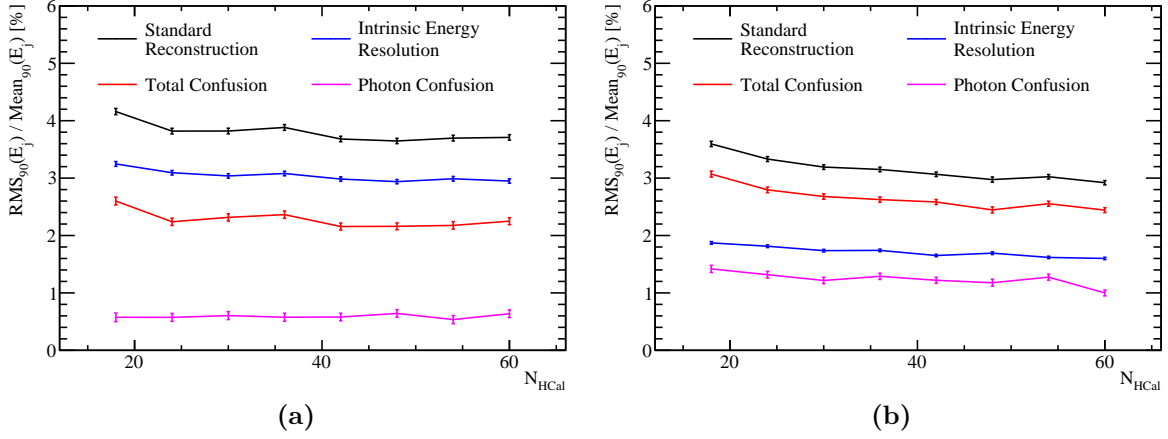
Figure 1.15: Contributions to the jet energy resolution shown as function of the longitudinal sampling frequency in the HCal using the nominal ILD detector model for (a) 45 GeV jets and (b) 250 GeV jets. The black curves correspond to the standard reconstruction, the blue curves to the intrinsic energy resolution contribution to the jet energy resolution, the red curves to the confusion contribution to the jet energy resolution and the magenta curves to the confusion contribution to the jet energy resolution related solely to photon reconstruction.

1.3.4 HCal Sampling Fraction

The performance of the ILD detector was studied for different ratios of active to absorber later thicknesses in the HCal. In the nominal detector model, the active scintillator layer thickness is 3 mm, while the absorber layer thickness is 20 mm giving a sampling fraction of 0.15. HCal models were simulated where this ratio was changed from 0.05 to 0.25 in steps of 0.05, while retaining the same number of interaction lengths.

No performance changes in the energy resolution for 50 GeV K_L^0s or the jet energy resolution for 91, 200, 360 and 500 GeV Z→uds di-jet events were observed when varying the ratio of active to absorber later thicknesses. Based on these simulations, there is no suggestion that varying this ratio has any statistically significant effect on the physics performance. Although this study indicates that thinning the active layer thickness would not change performance, hardware effects must also be considered to determine whether these conclusions hold true in a real detector. A study into the effects of the readout electronics is required before changing the active layer thicknesses to determine whether a MIP signal can be clearly distinguished when changing the sampling fraction.

1.3.5 HCal Absorber Material

The nominal choice of HCal absorber material is steel with tungsten providing a feasible alternative [7]. Although tungsten is more expensive than steel, it contains a larger number of nuclear interaction lengths per unit length. Therefore, using tungsten as the absorber material would allow for a reduction in the size of the HCal, while retaining the same number of nuclear interaction lengths. Reducing the depth of the calorimeter would decrease the size of the solenoid required, which would offset some of the additional cost of tungsten.

Table 1.5 shows the configuration for the steel and tungsten HCal options that were used in the full ILD simulation. To isolate the effects of changing the absorber material, the total depth, in nuclear interaction lengths, was kept constant when comparing the two options. Furthermore, the sampling fraction was also held constant. A number of different physics lists exist within GEANT4 for the modelling of hadronic showers. The default model for high energy physics calorimetry is the QGSP_BERT physics list. This uses the quark-gluon string model [8] with the precompound model of nuclear evaporation [9] (QGSP) for high energy interactions and the Bertini (BERT) cascade model [10] for intermediate energy interactions. For the study of absorber materials both the QGSP_BERT and the QGSP_BERT_HP physics lists were used. The QGSP_BERT_HP list uses the high precision neutron package (NeutronHP) to deal with the transportation of neutrons from below 20 MeV to thermal energies. This added detail is necessary for accurate modelling of hadronic showers in tungsten [11].

Parameter	Steel HCal Option	Tungsten HCal Option
Cell Size	30×30 mm^2 square cells	30×30 mm^2 square cells
Number of Layers	48 readout layers	48 readout layers
Absorber Material Thickness [mm]	20.0	12.0
Active Material Choice	Scintillator	Scintillator
Active Material Thickness [mm]	3.0	1.8

Table 1.5: The configuration of the steel and tungsten HCal options [1].

One of the dominant processes governing the energy deposition of hadronic showers in calorimeters is spallation [12]. Spallation begins with the collision of a high energy incident particle with nucleons in the calorimeter absorber material. This collision creates a cascade of high energy hadronic particles, e.g. protons, neutrons and pions, are

produced within the nucleus. If these energies are large enough, some of these particles may escape the nucleus and form secondary particles in the hadronic shower. After this initial collision, the nuclei of the absorbing material are left in an excited state. Assuming the excited nuclei are sufficiently stable that they will not undergo fission, they will return to a stable state by ejecting energy in the form of particles in a process called evaporation. Evaporation of neutrons, which is the dominant form of evaporation, significantly delays the growth of hadronic showers as after the evaporation process some of these neutrons participate in neutron capture [11]. Neutron capture involves an absorber nuclei capturing a neutron and then emitting a photon as it returns to a stable state. The time taken for the neutron capture mechanism to proceed is limited by the lifetime of the unstable nuclei [13], which typically makes neutron capture one of the slowest mechanisms by which hadronic showers can propagate. The number of evaporation neutrons released in a hadronic shower increases with the atomic number, Z, of the absorber material of a calorimeter increases. This is because of the increase in neutron content of the absorber material nuclei [11]. As the number of evaporation neutrons increases, more neutron capture processes are initiated, which results in a longer hadronic shower development time. Figure 1.16 shows the shower development times for hadronic showers in the tungsten (Z=74) and steel (iron, Z=26) HCal options and, as expected, the shower development time is greater for tungsten.
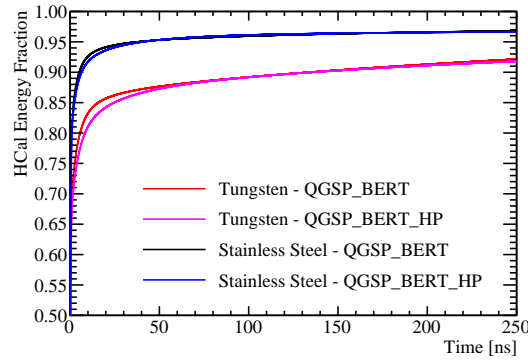
Figure 1.16: The fraction of the total calorimetric energy deposited in the HCal as a function of time for 25 GeV K_L^0 events using the steel and tungsten HCal options. Results are shown for both the QGSP_BERT and QGSP_BERT_HP physics lists. The calorimeter hit times have been corrected for straight line time of flight to the impact point.

Table 1.6 shows the energy resolution for 50 GeV K_L^0s obtained using the nominal ILD detector model with various HCal absorber materials and GEANT4 physics lists. In comparison to steel, tungsten option offers a $\sim 8\%$ improvement in the energy resolution

for 50 GeV neutral hadrons (using the QGSP_BERT_HP physics list). This can be attributed to differences in the nuclear structure of the two materials, which will lead to different developments of the hadronic showers within them. For example, the energy losses to nuclear binding energies are smaller in tungsten than steel, as the target nucleons are less stable than in iron, therefore, less energy is needed to liberate them. This will lead to a larger signal for tungsten and a reduction in the energy resolution in comparison to steel. The results of table 1.6 also indicate that the addition of the high precision neutron package was not important for this study.

HCal Option	Energy Resolution [%]
Steel, QGSP_BERT	8.8 ± 0.2
Steel, QGSP_BERT_HP	9.0 ± 0.3
Tungsten, QGSP_BERT	8.3 ± 0.2
Tungsten, QGSP_BERT_HP	8.3 ± 0.2

Table 1.6: The energy resolution for 50 GeV K_L^0s obtained using the nominal ILD detector with various HCal absorber materials and GEANT4 physics lists. A 100 ns timing cut was applied to the steel and tungsten HCal options in these simulations.

It should be emphasised that the HCal hit energy truncation, as described in chapter **??**, used for the tungsten and steel HCal options differs because tungsten contains a larger number of radiation lengths per nuclear interaction length than steel does. As the HCal primarily measures hadronic showers, one may naively expect the number of radiation lengths in the HCal to be irrelevant, given both options have the same number of nuclear interaction lengths. However, this is not the case because all hadronic showers have an electromagnetic component generated by the decays of hadrons to photons, e.g $\pi^0 \to \gamma\gamma$ and $\eta \to \gamma\gamma$. This leads to hadronic showers depositing more energy per calorimeter hit in tungsten than in steel and makes retuning the HCal hit energy truncation a necessity. As expected, the truncation used for tungsten, 5 GeV, is larger than for steel, 1 GeV, because of the increased average hit energy.

Table 1.7 shows the jet energy resolutions for selected jet energies obtained using the nominal ILD detector with various HCal absorber materials and GEANT4 physics lists. These results indicate that steel outperforms tungsten as the HCal absorber material. The magnitude of the improvement offered using steel grows as the jet energy increases; the jet energy resolution is $\sim 3\%$ better for the steel option for 45 GeV jets, while for 250 GeV jets the improvement is $\sim 11\%$. The intrinsic energy resolution and confusion

contributions to the jet energy resolution for 45 and 250 GeV jets are shown in table 1.8. The intrinsic energy resolution contribution to the jet energy resolution is almost identical for the two HCal options, which is expected because the K_L^0 energy resolution was only slightly better for the tungsten option. The tungsten option is unlikely to give a significantly better intrinsic jet energy resolution because only the small fraction of jet energy associated with neutral hadrons is measured in the HCal. The confusion contribution to the jet energy resolution is larger for tungsten than for steel; for 250 GeV jets the confusion contribution is $\sim 3.4\%$ in tungsten and only $\sim 3.0\%$ in steel. The larger confusion contribution is expected for the tungsten option because hadronic showers are generally wider in tungsten. The transverse profile of hadronic showers in the two HCal options is illustrated in figure 1.17, which shows the normalised distribution of the energy weighted transverse distance from the shower axis to the calorimeter hits for 50 GeV hadronic showers for both the steel and tungsten HCal options. Increasing the average hadronic shower width makes resolving individual particle showers in a dense jet environment more challenging, which means more calorimetric energy deposits will be incorrectly clustered together. This in turn results in incorrect associations being made between calorimetric energy deposits and charged particle tracks i.e. an increased confusion contribution. Again, the use of the QGSP_BERT_HP physics list, as opposed to QGSP_BERT, made a minimal impact on these results.

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

Table 1.7: The jet energy resolution for selected jet energies obtained using the nominal ILD detector with various HCal absorber materials and GEANT4 physics lists. A 100 ns timing cut was applied to the steel and tungsten HCal options in these simulations.

The impact of the choice of HCal absorber material is small on both the neutral hadron energy resolution and intrinsic jet energy resolution, however, the steel HCal option outperforms the tungsten option in terms of pattern recognition confusion. When examining the mechanical properties of steel and tungsten, it is clear that steel has a significant advantage over tungsten in terms of rigidity [14]. This means that fewer

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

Table 1.8: The contributions to the jet energy resolution obtained using the nominal ILD detector with various HCal absorber materials and GEANT4 physics lists. A 100 ns timing cut was applied to the steel and tungsten HCal options in these simulations.
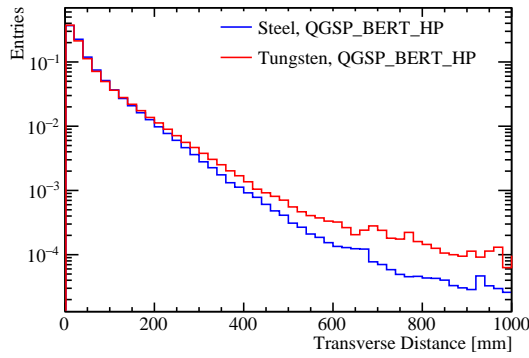
Figure 1.17: The normalised distribution of the energy weighted transverse distance of the calorimeter hits from a 50 GeV hadronic shower to the shower axis. The blue and red lines show the energy weighted transverse distance obtained using a steel and tungsten HCal absorber material in the ILD detector respectively. The simulations used the QGSP_BERT_HP physics list.

support structures would be required for the calorimeter leading to less dead material and better performance, which makes steel the preferred option.

1.4 Global Detector Parameters

The overall detector size and the magnetic field strength are major cost drivers for the ILD detector. Both will affect the jet energy resolution and studies showing their impact on detector performance are presented here.

1.4.1 The Magnetic Field Strength

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

Detector models were simulated where the magnetic field was varied from 1.0 to 5.0 T in steps of 0.5 T and the resulting jet energy resolutions are shown in figure 1.18. The larger the magnetic field strength, the better the jet energy resolution. Increasing the magnetic field strength from 1.0 to 5.0 T improves the jet energy resolution for 250 GeV jets by $\sim 25\%$. The higher the jet energy, the stronger the dependence of the jet energy resolution on the magnetic field strength.
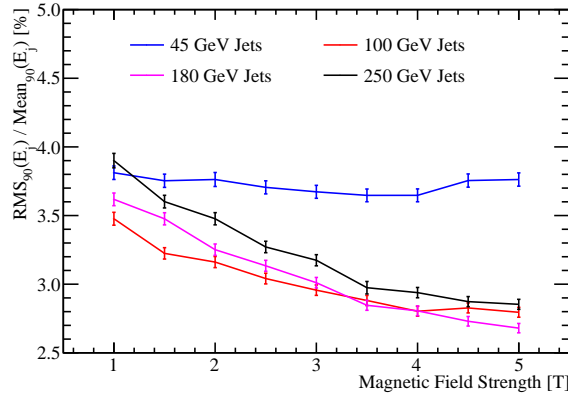
Figure 1.18: The jet energy resolution using the nominal ILD detector as a function of the magnetic field strength for various jet energies.

Figure 1.19 shows the breakdown of the jet energy resolution into the various contributions. As expected, there is a reduction in the confusion contribution with increasing magnetic field strength. Furthermore, there is a reduction in intrinsic energy resolution with increasing magnetic field strength for low energy jets. This is most likely due to particles being directed into the forward region of the detector. When a charged particle passes through a magnetic field it will traverse a helix. The radius of curvature, R, of

that helix is given by

$$R = \frac{p_t}{qB} \, , \tag{1.2}$$

where p_t is the transverse momentum of the charged particle with respect to the magnetic field, q is the particle charged and B is the magnetic field strength. As the magnetic field strength increases, the radius of curvature for charged particles will decrease. This leads to more particles being directed towards the forward region of the detector and a worsening intrinsic energy resolution because tracking coverage in the forward region of the detector is worse than in the barrel region [1]. This is illustrated in figure 1.20, which shows that as the magnetic field increases fewer charged particles are reconstructed. For high jet energies, low transverse momentum charged particles will still get directed to the forward regions of the detector, however, these contribute fractionally less energy to the total reconstructed energy. Therefore, the trend of worsening intrinsic energy resolution with increasing magnetic field strength is less pronounced as the jet energy grows.

At high jet energies, reducing the magnetic field strength appears to degrade the jet energy resolution; the intrinsic jet energy resolution for 250 GeV jets goes from $\sim 1.8\%$ to $\sim 2.1\%$ when reducing the magnetic field strength from 5.0 T to 1.0 T. This is due to an artefact in the definition of the intrinsic jet energy resolution. The intrinsic energy resolution is highly non-trivial to determine; Monte-Carlo (MC) information is used to make all associations between charged particle tracks and clusters of calorimeter hits as follows:

1. Each calorimeter hit is associated to the MC particle that deposits the largest amount of energy in that hit;

2. Clusters of calorimeter hits formed by the same MC particle are clustered together;

3. Each charged particle track is associated to the MC particle that produced it;

4. Clusters of calorimeter hits are associated to charged particle tracks if they are made by the same MC particle.

This procedure assumes that only one MC particle deposits energy per calorimeter hit. If multiple MC particles deposit energy in the same calorimeter hit, this assumption breaks down and errors will be made when associating charged particle tracks to calorimetric energy deposits. These errors cause the same double counting and omission of energy deposits as confusion does, however, this confusion-like effect is smaller because MC

particles depositing energy in the same calorimeter hit in fine granularity calorimeters are rare. As the overlap of particle showers within the calorimeter grows, as it does for high energy jets when reducing the magnetic field, the intrinsic energy resolution appears to get worse because of this confusion-like effect. Because this confusion-like effect is second order small in comparison to changes in the confusion contribution, the dependence of the detector performance on the magnetic field strength can be confidently quantified.
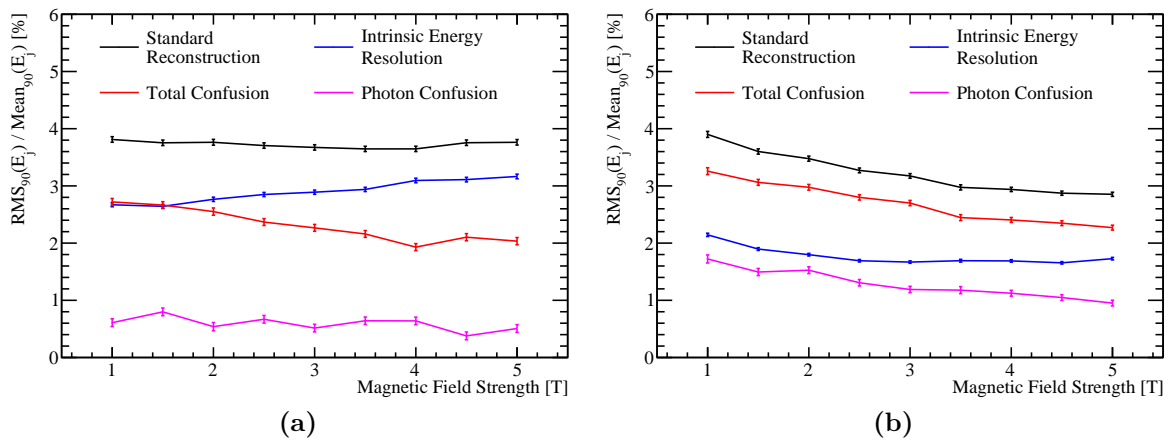
Figure 1.19: Contributions to the jet energy resolution shown as function of the magnetic field strength using the nominal ILD detector model for (a) 45 GeV jets and (b) 250 GeV jets. The black curves correspond to the standard reconstruction, the blue curves to the intrinsic energy resolution contribution to the jet energy resolution, the red curves to the confusion contribution to the jet energy resolution and the magenta curves to the confusion contribution to the jet energy resolution related solely to photon reconstruction.
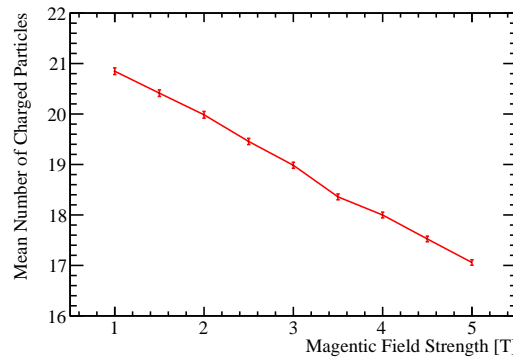
Figure 1.20: The mean number of reconstructed charged particles as a function of the magnetic field strength for 91 GeV Z→uds di-jet events. The nominal ILD detector model was used and the pattern recognition has been fully cheated using the MC information.

In summary, increasing the magnetic field strength is beneficial to the jet energy resolution because it reduces confusion from associating tracks to calorimetric energy deposits from charged particles. The intrinsic energy resolution is also dependent upon the magnetic field strength, however, the effect is second order small in comparison to the confusion. Although the nominal field of 3.5 T gives good performance, increasing the field strength is a clear way of making gains in overall detector performance.

1.4.2 Inner ECal Radius

The impact on the jet energy resolution of the overall size of the detector was studied by simulating detector models where the ECal inner radius was altered. The ECal inner radii considered were 1208, 1408, 1608, 1808 (nominal) and 2008 mm.

Figure 1.21 shows the dependence of the jet energy resolution on the ECal inner radius. Increasing the ECal inner radius increases the separation between particles as they enter the calorimeters, which reduced the effect of confusion and improves the jet energy resolution. As confusion is more dominant at higher energies, the benefits to using a larger ECal radius grow with increasing jet energy; increasing the ECal inner radius from 1208 mm to 2008 mm improves the jet energy resolution by $\sim 9\%$ for 45 GeV jets, but by $\sim 25\%$ for 250 GeV jets. Figure 1.22 shows the decomposition of the jet energy resolution into its different component. These results explicitly show a reduction in confusion with increasing ECal inner radius; the confusion contribution goes from $\sim 3.4\%$ to $\sim 2.4\%$ when increasing the ECal inner radius from 1208 mm to 2008 mm. The intrinsic energy resolution of the detectors shows no strong dependence on the inner ECal radius. The apparent degradation in intrinsic energy resolution at low ECal inner radii is an artefact of the association of a single MC particle per calorimeter cell when running the cheated pattern recognition as explained in section 1.4.1. The dominant effect driving the jet energy resolution is, as expected, the confusion.

In conclusion, increasing the ECal inner radius benefits the jet energy resolution because it increases the separation between particles as they enter the calorimeter, which reduces confusion.
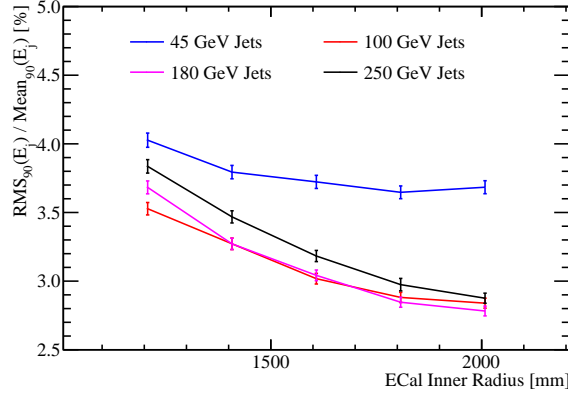
Figure 1.21: The jet energy resolution using the nominal ILD detector as a function of the ECal inner radius for various jet energies.
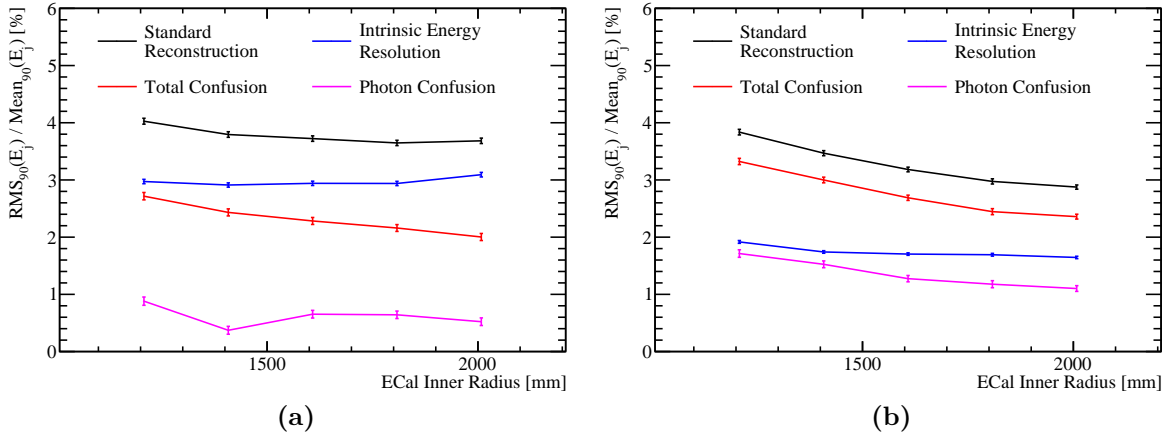
Figure 1.22: Contributions to the jet energy resolution shown as function of the ECal inner radius using the nominal ILD detector model for (a) 45 GeV jets and (b) 250 GeV jets. The black curves correspond to the standard reconstruction, the blue curves to the intrinsic energy resolution contribution to the jet energy resolution, the red curves to the confusion contribution to the jet energy resolution and the magenta curves to the confusion contribution to the jet energy resolution related solely to photon reconstruction.

1.5 Summary

The effect of varying the configuration of the calorimeters, the magnetic field strength and the overall detector size on the single particle and jet energy resolutions were presented in this chapter. For both the ECal and the HCal, the dominant factor determining the intrinsic energy resolution was the longitudinal sampling frequency. However, the jet energy resolution had the strongest sensitivity to the ECal cell size, which shows that spatial recognition is more important when using particle flow calorimetry than

intrinsic energy resolution. The HCal cell size was found to be less significant than the ECal cell size for determining the jet energy resolution because separation of nearby particle showers in the HCal uses the spatial information gathered in the ECal. In the particle flow paradigm, fine segmentation in the ECal can compensate for the coarser HCal granularities. The jet energy resolution also showed a strong dependence on the magnetic field strength and the overall detector size. Increasing both the magnetic field and overall detector size leads to greater separation of nearby particle showers in the calorimeters, which reduces the effect of confusion.

Chapter 2

The Sensitivity of CLIC to Anomalous Gauge Couplings through Vector Boson Scattering

"Kids, you tried your best, and you failed miserably. The lesson is, never try."

— Homer Simpson

2.1 Motivation

Vector boson scattering is the interaction of the form VV → VV where V is any of the electroweak gauge bosons W^+, W^-, Z or γ. This is an interesting process to study because it provides understanding of how the Standard Model Higgs is able to unitarise the otherwise unbounded cross-section for longitudinal massive gauge boson scattering. Vector boson scattering also provides insights into beyond standard model physics that impacts the electroweak sector by probing potential anomalous triple and quartic gauge couplings.

Triple and quartic gauge couplings lead to interactions of the form V → VV and VV → VV respectively. In the Standard Model there are five allowed vertices, shown in figure 2.1, which arise from the kinematic term $\mathcal{L}_{kin} = -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}W_{\mu\nu}W^{\mu\nu}$ in the Standard Model Lagrangian.
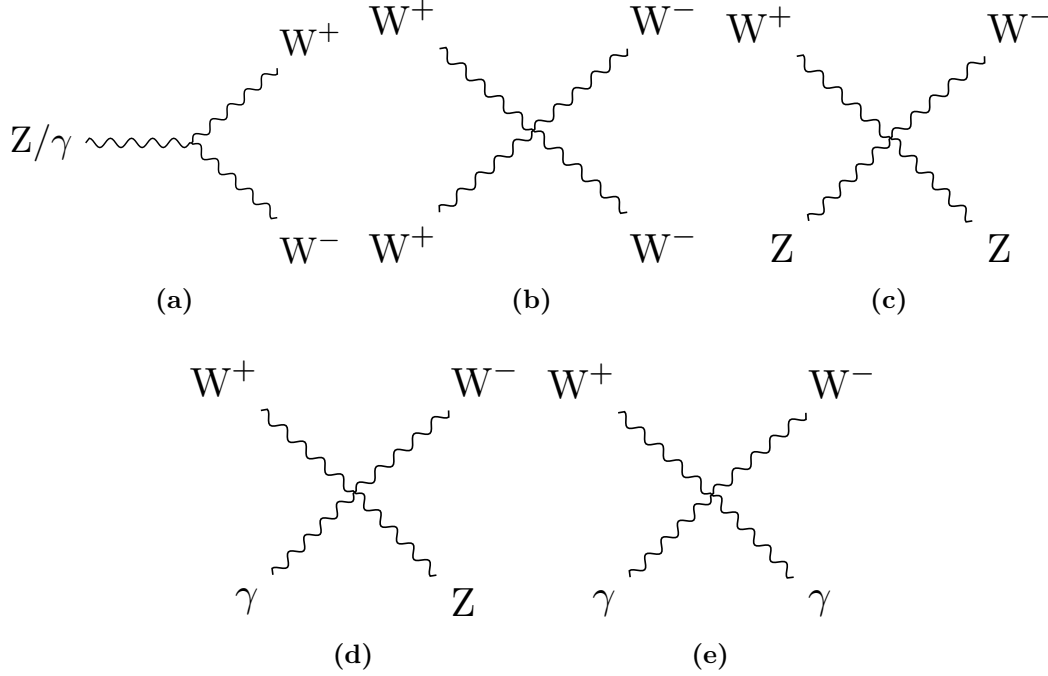
Figure 2.1: Triple and quartic gauge boson vertices in the Standard Model.

Anomalous triple and quartic gauge couplings are introduced as parameters in effective field theories (EFTs). These couplings either modify the Standard Model triple and quartic gauge boson vertices or introduce new triple and quartic vertices that were previously forbidden. EFTs are a mathematical construct designed to introduce new physics in a manner that builds upon the Standard Model. They work under the assumption that new physics exists at an energy scale, Λ, that is much higher than the energy scales currently accessible to modern day particle physics experiments. In the limit $\Lambda \to \infty$, the Standard Model is reproduced as the new physics becomes kinematically inaccessible. Such theories are model independent, giving them a wide span in the search for new physics. A classic example of an EFT theory is the Fermi theory for beta decay [15]. At energies much below the mass of the W boson, the weak interaction occurring when a neutron decays into a proton, electron and anti-neutrino can be treated as a four-point vertex with quartic coupling strength G_F, the Fermi Coupling constant as shown in figure 2.2.

The study presented in this chapter examines the anomalous quartic gauge couplings α_4 and α_5 through vector boson scattering process. The anomalous gauge couplings that are to be examined are introduced as part of an EFT that is described in chapter **??**.
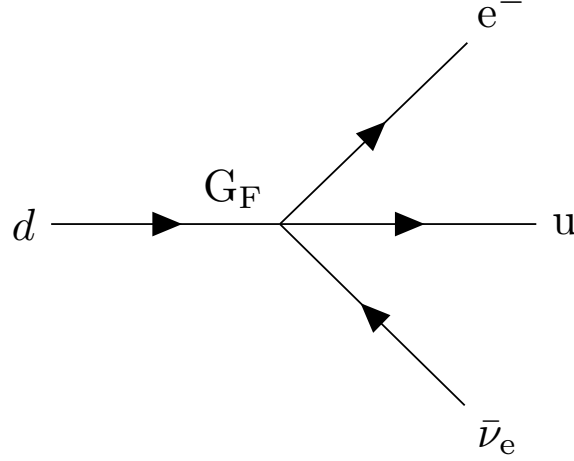
**Figure 2.2:** Four-point vertex proposed for explanation of beta decay by Fermi.

These couplings appear in the Lagrangian through the following terms

$$\alpha_4[\mathrm{Tr}(V^\mu V_\mu)]^2 \quad \text{and} \quad \alpha_5 \mathrm{Tr}(V^\mu V_\nu)]\mathrm{Tr}(V^\nu V_\mu)] \,, \tag{2.1}$$

where V_μ corresponds, in a carefully chosen gauge, to a linear combination of the massive gauge bosons W^+, W^- and Z. These terms modify the Standard Model vertices $W^+W^- \to W^+W^-$ and $W^+W^- \to ZZ$ as well as introducing the new vertex $ZZ \to ZZ$. The anomalous gauge couplings α_4 and α_5 can be studied in vector boson scattering processes such as those shown in figure 2.3.

CLIC is designed for precision measurements in e^+e^- collisions at high energies and it is ideal for a study of vector boson scattering. The application of Particle Flow Calorimetry with fine granularity calorimeters gives CLIC excellent jet energy resolution, which allows it to clearly characterise multi-jet final states and final states containing missing energy in the form of neutrinos. The excellent jet energy resolution also allows for accurate separation of W and Z bosons through di-jet invariant mass, which will be invaluable for event selection.

The cross-sections for vector boson scattering processes are sufficiently large at the proposed running energies for CLIC to give large signal sample sizes. A study of anomalous gauge boson couplings at CLIC has the potential to give results several orders of magnitude better than the complementary studies performed at the LHC because of the reduction in hadronic backgrounds and increased cross-section for vector boson
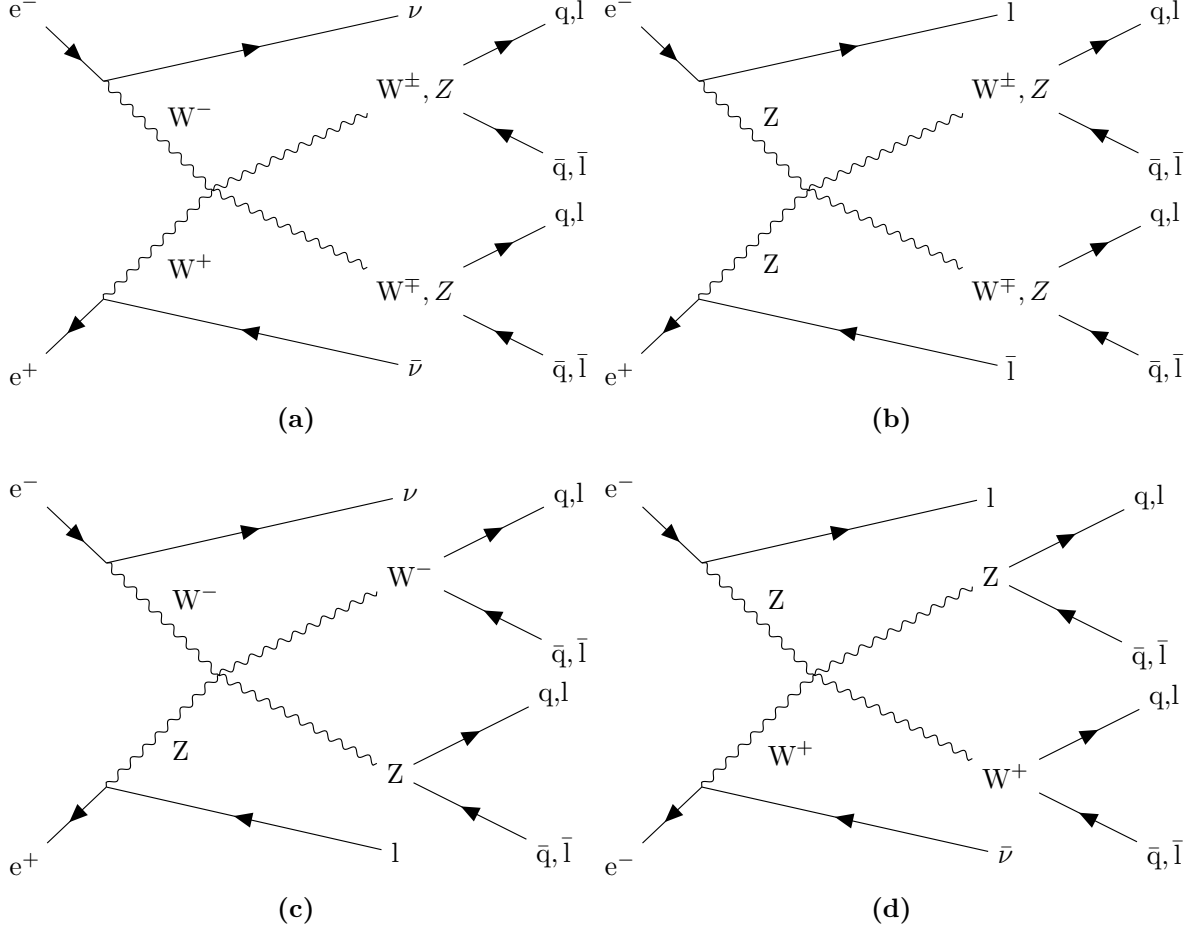
Figure 2.3: Example of vector boson scattering Feynman diagrams showing sensitivity to quartic gauge boson self-interaction vertices. The processes shown are relevant for CLIC. In these diagrams q represents the u, d, s, c and b quarks; l represents e^-, μ^- and τ^- leptons; and ν represents the ν_e, ν_μ and ν_τ neutrinos.

scattering processes [16]. The above reasons make a strong case for performing a vector boson scattering analysis at CLIC.

The branching fractions for the hadronic decays of both the $W^\pm$ and Z bosons is of the order of 70% [17], therefore, the signal final states for the analysis presented in this chapter are vector boson scattering processes where the outgoing bosons decay purely hadronically: $\nu\nu$qqqq, νlqqqq and llqqqq.

2.2 Event Generation, Simulation and Reconstruction

Events were generated using Whizard [18, 19] version 1.95. Due to the presence of beamstrahlung photons in the CLIC beam, events were generated for collisions of e^+e^-, $e^+\gamma$, γe^- and $\gamma\gamma$. The energy spectra used for all particles involved in these collisions took into account the effects of radiation in the form of beamstrahlung photons and the intrinsic energy spread of the CLIC beam. Furthermore, events involving the interaction between the electromagnetic field of the beam particles involving quasi-real photon mediators with low momenta, described by the Weizsacker-Williams approximation or the Equivalent Photon Approximation (EPA), were generated using Whizard and included in this analysis. Fragmentation and hadronisation was implemented using PYTHIA 6.4 [20] that was tuned for OPAL e^+e^- collision data recorded at LEP [21]. The decays of tau leptons was simulated using TAUOLA [22]. The full list of events used in this analysis, along with their Standard Model cross-section for $\sqrt{s} = 1.4$ TeV can be found in table 2.1. The samples comprise all final states that are relevant, either as signal or background processes, for an analysis involving the purely hadronic decay channels from the vector boson scattering process:

- Final states from the purely hadronic decay channels of the vector boson scattering process. These states are expected to show sensitivity to the anomalous couplings α_4 and α_5: $e^+e^- \to \nu\nu qqqq$, $e^+e^- \to \nu lqqqq$ and $e^+e^- \to llqqqq$

- Final states with four primary quarks arising from e^+e^- interactions: $e^+e^- \to qqqq$.

- Final states with two primary quarks arising from e^+e^- interactions: $e^+e^- \to \nu\nu qq$, $e^+e^- \to \nu lqq$, $e^+e^- \to llqq$ and $e^+e^- \to qq$.

- Final states with four primary quarks arising from the interactions of either e^+ or e^- with a beamstrahlung photon: $e^-\gamma_{BS} \to e^- qqqq$, $e^+\gamma_{BS} \to e^+ qqqq$, $e^-\gamma_{BS} \to \nu_e qqqq$ and $e^+\gamma_{BS} \to \bar{\nu}_e qqqq$.

- Final states with four primary quarks arising from the interactions of either e^+ or e^- with the electromagnetic field of the opposing beam particle. These cross-sections are calculated using the EPA approximation, which represents the electromagnetic field of the opposing beam particle as a series of photons, so the final states appear as interactions of e^+ or e^- with photons: $e^-\gamma_{EPA} \to e^- qqqq$, $e^+\gamma_{EPA} \to e^+ qqqq$, $e^-\gamma_{EPA} \to \nu_e qqqq$ and $e^+\gamma_{EPA} \to \bar{\nu}_e qqqq$.

- Final states with four primary quarks arising from the interaction of the electromagnetic fields of opposing beam particles using the EPA approximation: $\gamma_{\mathrm{EPA}}\gamma_{\mathrm{EPA}} \to$ qqqq.

- Final states with four primary quarks arising arising from the interaction of the electromagnetic field of either e^+ or e^- using the EPA approximation with a beamstrahlung photon: $\gamma_{\mathrm{EPA}}\gamma_{\mathrm{BS}} \to$ qqqq or $\gamma_{\mathrm{BS}}\gamma_{\mathrm{EPA}} \to$ qqqq.

- Final states with four primary quarks arising from the interaction of two beamstrahlung photons: $\gamma_{\mathrm{BS}}\gamma_{\mathrm{BS}} \to$ qqqq.

In the above list q represents u, $\bar{\mathrm{u}}$, d, $\bar{\mathrm{d}}$, s, $\bar{\mathrm{s}}$, c, $\bar{\mathrm{c}}$, b or $\bar{\mathrm{b}}$; l represents $e^{\pm}$, $\mu^{\pm}$ or $\tau^{\pm}$; and ν represents ν_{e}, $\overline{\nu}_{\mathrm{e}}$, ν_{μ}, $\overline{\nu}_{\mu}$, ν_{τ} and $\overline{\nu}_{\tau}$.

Monte-Carlo (MC) samples were simulated using the CLID_ILD detector model [2]. Further details of this detector model can be found in chapter **??**. The simulation was performed in MOKKA [23], which is a GEANT4 [24] wrapper providing detailed geometric descriptions of detector concepts for the linear collider. Events were reconstructed using the MARLIN [25] c++ framework, designed for reconstruction at the linear collider. PandoraPFA [3, 26] was used to apply Particle Flow Calorimetry in the reconstruction, the full details of which can be found in chapter **??**.

The effect of the $\gamma\gamma \to hadrons$ backgrounds, discussed in section **??**, were incorporated in the analysis by overlaying $\gamma\gamma \to hadrons$ events onto the signal and background event samples. The overlaid backgrounds were added prior to reconstruction so that their impact on the reconstruction was fully accounted for. For each physics event of interest, $\gamma\gamma \to hadrons$ background events equivalent to 60 bunch crossings (BXs) are included. As readout time windows are applied in detector readout, 60 BXs is sufficient for accounting for the $\gamma\gamma \to hadrons$ backgrounds. These backgrounds occur in a time window of -5 ns to 25 ns around the physics event and the BXs are separated by 0.5 ns, to mimic the CLIC bunch train structure. The number of background events overlaid per BX is drawn from a Poisson distribution with a mean of 1.3 (3.2) events per bunch crossing at $\sqrt{s} = 1.4$ (3) TeV [14].

Detector readout is simulated using a readout time window, of 10 ns on all detectors apart from the TPC and HCal barrel. All hits are kept in the TPC, while a 100 ns time window is used in the HCal barrel to account for the additional time it takes hadronic showers to develop in tungsten. All readout times are corrected for straight time-of-flight

Final State	Cross Section [fb]
$e^+e^- \to \nu\nu qqqq$	24.7
$e^+e^- \to \nu lqqqq$	110.4
$e^+e^- \to llqqqq$	62.1
$e^+e^- \to qqqq$	1245.1
$e^+e^- \to \nu\nu qq$	787.7
$e^+e^- \to \nu lqq$	4309.7
$e^+e^- \to llqq$	2725.8
$e^+e^- \to qq$	4009.5
$e^-\gamma_{\mathrm{EPA}} \to e^-qqqq$	287.1
$e^-\gamma_{\mathrm{BS}} \to e^-qqqq$	1160.7
$e^+\gamma_{\mathrm{EPA}} \to e^+qqqq$	286.9
$e^+\gamma_{\mathrm{BS}} \to e^+qqqq$	1156.3
$e^-\gamma_{\mathrm{EPA}} \to \nu_e qqqq$	32.6
$e^-\gamma_{\mathrm{BS}} \to \nu_e qqqq$	136.9
$e^+\gamma_{\mathrm{EPA}} \to \overline{\nu}_e qqqq$	32.6
$e^+\gamma_{\mathrm{BS}} \to \overline{\nu}_e qqqq$	136.4
$\gamma_{\mathrm{EPA}}\gamma_{\mathrm{EPA}} \to qqqq$	753.0
$\gamma_{\mathrm{EPA}}\gamma_{\mathrm{BS}} \to qqqq$	4034.8
$\gamma_{\mathrm{BS}}\gamma_{\mathrm{EPA}} \to qqqq$	4018.7
$\gamma_{\mathrm{BS}}\gamma_{\mathrm{BS}} \to qqqq$	21406.2

Table 2.1: Cross sections of signal and background processes for $\sqrt{s} = 1.4$ TeV. In the above table q represents u, $\bar{\mathrm{u}}$, d, $\bar{\mathrm{d}}$, s, $\bar{\mathrm{s}}$, c, $\bar{\mathrm{c}}$, b or $\bar{\mathrm{b}}$; l represents $e^\pm$, $\mu^\pm$ or $\tau^\pm$; and ν represents ν_e, $\overline{\nu}_e$, ν_μ, $\overline{\nu}_\mu$, ν_τ and $\overline{\nu}_\tau$. The EPA and BS subscript on the incoming photon indicates whether the photon is generated from the equivalent photon approximation or beamstrahlung.

to the impact point (IP). Any hits that have are measured outside of these windows are not used in the reconstruction.

2.3 Modelling of Anomalous Gauge Couplings

The samples that were sensitive to the anomalous gauge couplings α_4 and α_5 were generated using Whizard version 1.97, instead of the previously quoted version 1.95. This change was required as version 1.97 contained a unitarisation scheme that ensured

cross-sections for processes involving longitudinal gauge boson scattering did not violate unitarity at the energies considered here.

Two alternative methods exist for modelling the sensitivity of the vector boson scattering process to the anomalous gauge couplings α_4 and α_5. The first is to generate multiple samples with different values of α_4 and α_5 and the second is to generate a single sample with $\alpha_4 = 0$ and $\alpha_5 = 0$ and reweight that sample. The latter approach was taken in this analysis as the former approach is impractical when considering a fine sampling of the α_4 and α_5 space.

Event weights, w, are calculated according to the ratio of the matrix elements, M, for the particular event configuration [27]

$$w(\alpha_4, \alpha_5) = \frac{|M(event, \alpha_4, \alpha_5)|^2}{|M(event, 0, 0)|^2}. \tag{2.2}$$

Figure 2.4 shows the dependence of the event weights on α_4 and α_5 for four individual $\nu\nu$qqqq final state events, generated for $\sqrt{s} = 1.4$ TeV.

Only final states involving contributions from massive gauge boson quartic vertices require reweighting. Whizard was used to evaluate the cross-sections for all final states shown in table 2.1 with $\alpha_4 = \alpha_5 = 0$ and with $\alpha_4 = \alpha_5 = 0.05$. Only the three final states shown in table 2.2 were found to have a dependency on α_4 and α_5.

Final State	Cross Section [fb] ($\alpha_4 = \alpha_5 = 0.00$)	Cross Section [fb] ($\alpha_4 = \alpha_5 = 0.05$)	Percentage Change[%]
$e^+e^- \rightarrow \nu\nu$qqqq	24.7	34.6	+40.1
$e^+e^- \rightarrow \nu$lqqqq	115.3	113.0	-2.0
$e^+e^- \rightarrow$ llqqqq	62.1	68.6	+10.5

Table 2.2: Cross sections for selected processes showing the effect of the anomalous gauge couplings α_4 and α_5 for $\sqrt{s} = 1.4$ TeV.

To maximise the sensitivity to the anomalous gauge couplings, the $\nu\nu$qqqq final state is used to define signal in this analysis. The lνqqqq and llqqqq final states are treated as backgrounds that are invariant to changes in α_4 and α_5 because they have a much reduced sensitivity to the anomalous gauge couplings in comparison to the $\nu\nu$qqqq final state. Furthermore, the lνqqqq and llqqqq final states can be easily vetoed during event selection because of the presence of the primary lepton, which means the sensitivity of

(a)

(b)

(c)

(d)
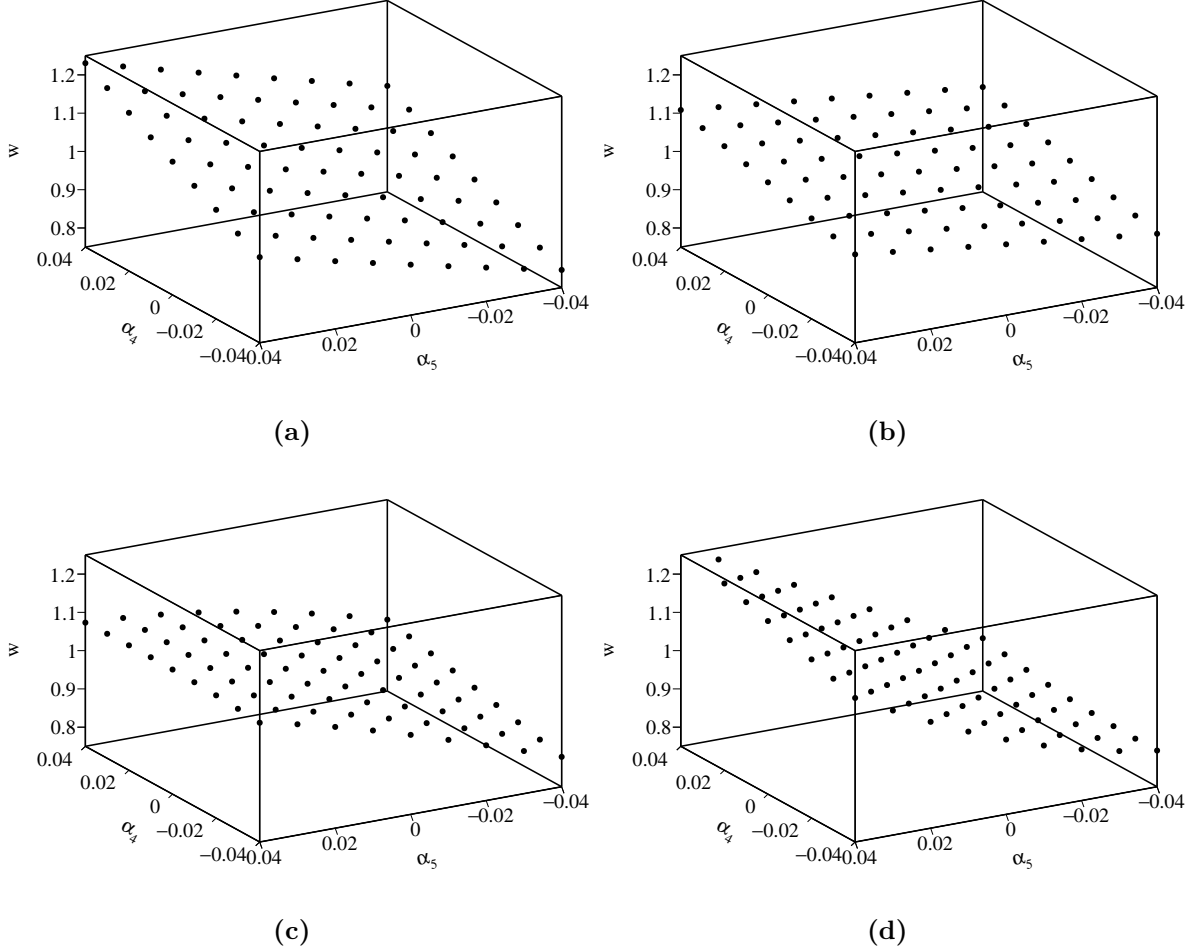
Figure 2.4: The event weights, w, determined by the generator as a function of the anomalous couplings α_4 and α_5 for a selection of $\nu\nu$qqqq final state events for $\sqrt{s} = 1.4$ TeV.

these states to the anomalous gauge couplings will have a negligible effect on the results from this study.

Use of the unitarisation scheme in Whizard 1.97, which is needed to ensure cross-sections do not violate unitarity when studying anomalous gauge couplings at CLIC like energies, requires a unit CKM matrix [27]. The impact of this requirement was examined by comparing several reconstructed and MC distributions for $\nu\nu$qqqq final state events generated with Whizard using a Standard Model and unit CKM matrix. No significant differences were observed, which indicates that enforcing a unit CKM matrix when generating the $\nu\nu$qqqq final state samples did not significantly affect this analysis.

2.4 Data Analysis

The following section contains a description of how the variables used throughout the anomalous gauge coupling sensitivity study are determined.

2.4.1 Limiting Beam Related Backgrounds

During the reconstruction, after the inner detector tracks have been reconstructed, the CLICTrackSelection processor is applied, which vetoes poorly reconstructed and fake tracks by applying simple quality cuts to the number of hits in the tracking sub-detectors. The CLICTrackSelection processors also reject tracks where the time of arrival at the calorimeter differs by more than 50 ns between a straight line of flight and a helix fit to the track. Applying this cut ensures that associations made between charged particles tracks and calorimetric energy deposits are consistent.

Following the reconstruction, the CLICPfoSelector processor is applied to remove reconstructed particle flow objects (PFOs) that originate from beam related backgrounds. This processor applies cuts on the p_T and timing information of the PFOs, which vary as a function of position in the detector and the PFO type to target regions of the detector where backgrounds are more prominent, e.g. low p_T for $\gamma\gamma \rightarrow hadrons$ events. Three configurations of the CLICPfoSelector have been developed for the CLIC environment and were considered in this analysis. They are, in order of increasing background rejection, the Loose, Default and Tight selections [26].

2.4.2 Jet Finding

After the application of the CLICPfoSelector, the MarlinFastJet processor, a wrapper for the FastJet [28] processor, was used to cluster each event into four jets. These jets are then paired up to form two candidate bosons working under the assumption that the correct pairing is achieved when the difference between the invariant masses of the candidate bosons is a minima. In the case of the signal final state, $\nu\nu$qqqq, it is assumed that the four jets and two candidate bosons map onto the four primary quarks and two outgoing bosons in the vector boson scattering process. The jet clustering was performed using the longitudinally invariant k_t jet algorithm in exclusive mode. The longitudinally invariant k_t algorithm proceeds as follows

1. Determine the k_t distance, d_{ij}, for each pair of particles, i and j, and the beam, d_{iB}, distance for each particle, i. These distances are defined as

$$d_{ij} = \min(p_{ti}^2, p_{tj}^2)\Delta R_{ij}^2/R^2 \ , \tag{2.3}$$

$$d_{iB} = p_{ti}^2 \ , \tag{2.4}$$

   where $\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, p_{ti} is the transverse momentum of particle i, y_i is the rapidity of particle i, ϕ_i is the azimuthal angle of the direction of travel of particle i and R is a configurable parameter that typically is of the order of 1.

2. Find the minimum distance, $d_{\min}$, of all the k_t and beam distances. If the minimum occurs for a k_t distance, particles i and j are merged, summing their 4-momenta. If the beam distance is the minima, particle i was declared to be part of the "beam" jet and the particle is removed from the list of particles and not included in the final jet output.

3. Repeat until the desired number of jets is created. Alternatively, in inclusive mode this would be repeated until no particles are left in the event.

Two other clustering algorithms were considered, however, they were found to be inappropriate for the experimental conditions at CLIC. These alternative algorithm choices are applied in the same manner as the longitudinally invariant k_t algorithm, however, they differ in the definition of d_{ij} and d_{iB}. Figure 2.5 shows the distribution of the invariant mass of the candidate bosons for $\sqrt{s} = 1.4$ TeV $\nu\nu$qqqq final state events for each of the jet algorithms considered. The candidate boson masses are determined by forcing the events into 4 jets and then pairing the jet pairs to form candidate bosons. The jet pairing configuration is determined by pairing jets such that the mass differences between the two candidate bosons is a minimum.

The first alternative jet algorithm considered was the k_t algorithm for e^+e^- colliders, the e^+e^- k_t or Durham algorithm. In this algorithm d_{iB} is not used and

$$d_{ij} = 2\min(E_i^2, E_j^2)(1 - cos\theta_{ij}) \ , \tag{2.5}$$

where θ_{ij} is the opening angle of particles i and j and E_i is the energy of particle i. In the collinear limit d_{ij} corresponds to the relative transverse momenta of the particles. The major failure of this algorithm when applied to CLIC is the absence of d_{iB}, which leads to large numbers of beam related background particles being associated to jets. As figure 2.5 shows, the invariant mass of the paired jets, which should peak around
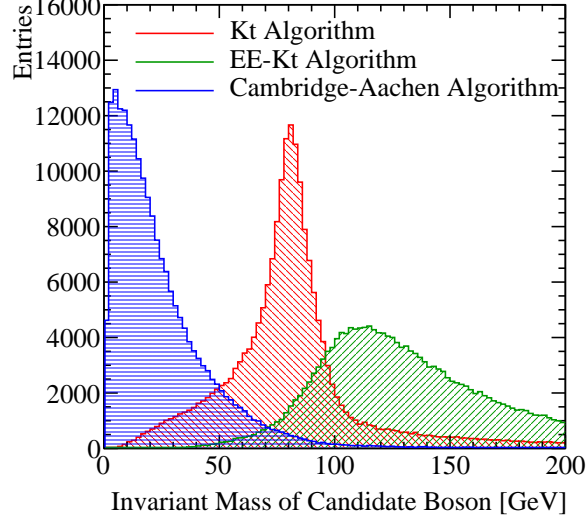
Figure 2.5: The reconstructed masses for different choices of jet algorithm for $\sqrt{s} = 1.4$ TeV
$\nu\nu$qqqq final state events. These samples should be dominated by vector boson
scattering involving pairs of outgoing W bosons and so it is expected that a peak
at the W boson mass, $m_W = 80.385 \pm 0.015$ GeV [17], should be observed. In
the case of the k_t algorithm and the $e^+e^- k_t$ algorithm an R parameter of 0.7 was
used. All distributions show raw number of events.

the W and Z boson masses, is much larger than expected, due to the presence of these
backgrounds. Also this algorithm is not invariant to boosts along the beam direction
meaning that it is inappropriate for use at CLIC given the beam induced backgrounds
modify the nominal collision kinematics.

The second alternative jet algorithm considered was the Cambridge-Aachen jet
algorithm where

$$d_{ij} = \Delta R_{ij}^2 / R^2 \ , \tag{2.6}$$

$$d_{iB} = 1 \ . \tag{2.7}$$

This algorithm performs poorly as it does not account for the transverse momentum or
the energy of the particles being clustered. In essence, this is a cone clustering algorithm
with a cone radius defined through $\Delta R_{ij} = R$, which even for large R was found to
discard too much energy in the event to be useful for this analysis. This can be seen in
figure 2.5 where the invariant mass of the paired jets is much lower than expected. This
algorithm is appropriate for events that contain highly boosted jets, however, at CLIC
the jets are too disperse for this algorithm to be successful.

2.4.2.1 Optimal Jet Finding Algorithm

Optimisation of the jet finding procedure was performed on both the PFO selection and the value of the R parameter used in the longitudinally invariant k_t algorithm. The optimisation procedure involved performing the sensitivity study, described in section 2.6, using solely the $\nu\nu$qqqq signal final state. This methodology ensures that the optimisation was done with respect to the physics of interest without having to perform the jet reconstruction for the large number of background events for each jet algorithm configuration considered.

Table 2.3 shows the one σ confidence limits on the measurement of α_4 and α_5 obtained using the $\nu\nu$qqqq signal final state only at $\sqrt{s} = 1.4$ TeV for different jet algorithm configurations. These confidence limits represent the idealised sensitivity of the CLIC experiment to the anomalous gauge couplings. Once the effects of backgrounds and event selection are included in the analysis, these confidence limits will increase in size.

| R Parameter | PFO Selection | | |
	Tight Selected PFOs	Selected PFOs	Loose Selected PFOs
0.7	$-0.0039 < \alpha_4 < 0.0051$	$-0.0035 < \alpha_4 < 0.0047$	$-0.0037 < \alpha_4 < 0.0047$
	$-0.0027 < \alpha_5 < 0.0031$	$-0.0025 < \alpha_5 < 0.0031$	$-0.0024 < \alpha_5 < 0.0028$
0.9	$-0.0036 < \alpha_4 < 0.0047$	$-0.0035 < \alpha_4 < 0.0045$	$-0.0035 < \alpha_4 < 0.0045$
	$-0.0026 < \alpha_5 < 0.0031$	$-0.0023 < \alpha_5 < 0.0027$	$-0.0022 < \alpha_5 < 0.0027$
1.1	$-0.0036 < \alpha_4 < 0.0047$	$-0.0036 < \alpha_4 < 0.0048$	$-0.0036 < \alpha_4 < 0.0046$
	$-0.0026 < \alpha_5 < 0.0031$	$-0.0025 < \alpha_5 < 0.0029$	$-0.0024 < \alpha_5 < 0.0028$

Table 2.3: One σ confidence limits on the measurement of α_4 and α_4 obtained using the $\nu\nu$qqqq signal final state only at $\sqrt{s} = 1.4$ TeV for different jet algorithm configurations.

The configuration for the jet algorithm for the $\sqrt{s} = 1.4$ TeV analysis was chosen as selected PFOs with an R parameter of 0.9. While the loose PFO selection gives a marginally better performance, the selected PFO selection was preferred to minimise the effects of the $\gamma\gamma \rightarrow hadrons$ background. Figure 2.6a shows confidence contours, given a null hypothesis of $\alpha_4 = \alpha_5 = 0$, for the selected PFO and R parameter of 0.9 jet algorithm configuration for $\sqrt{s} = 1.4$ TeV. Figures 2.6b and 2.6c show the one dimensional χ^2 distribution for α_4 and α_5, assuming $\alpha_5 = 0$ and $\alpha_4 = 0$ respectively, for the same configuration.

(a)

(b)

(c)
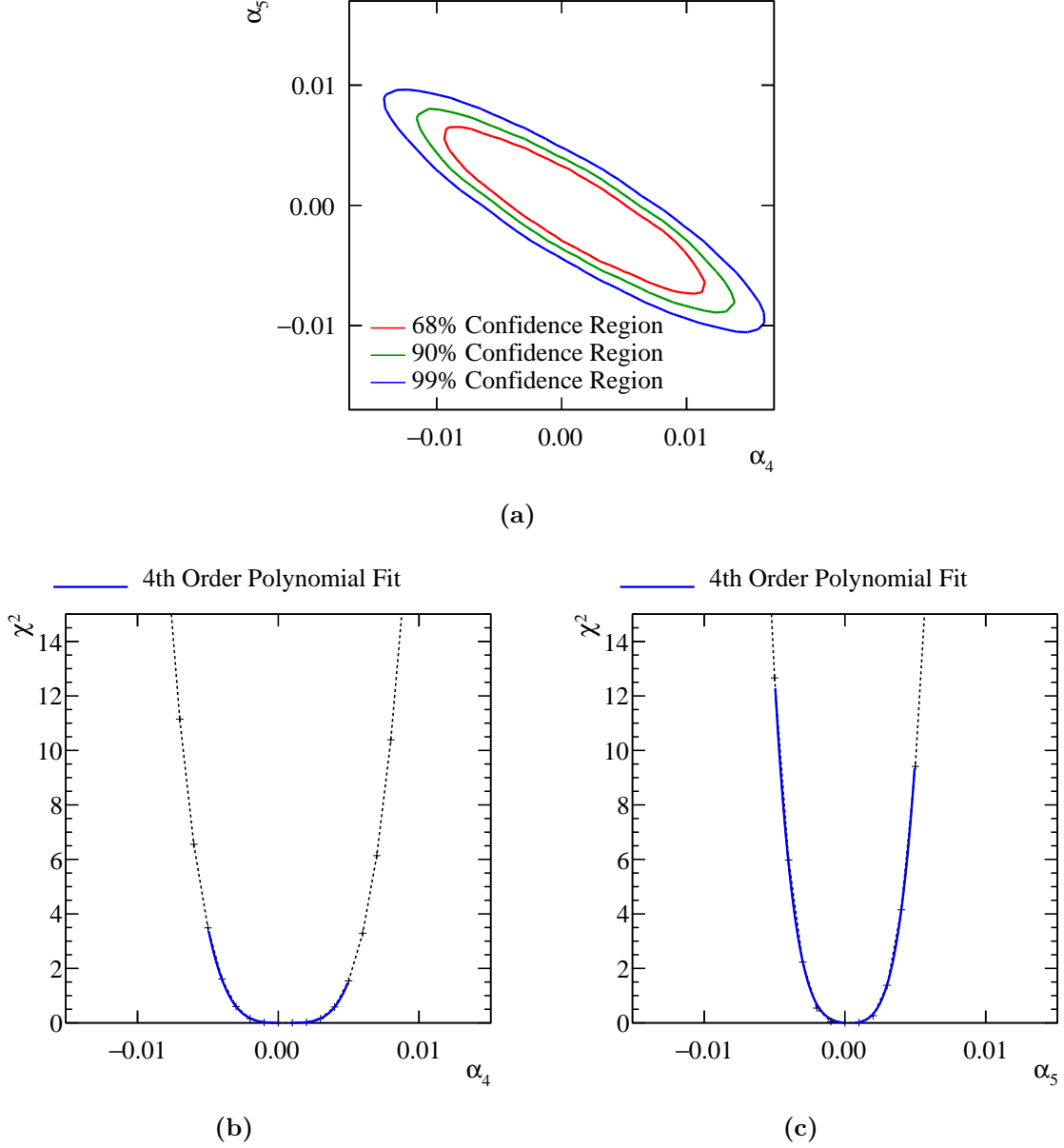
Figure 2.6: χ^2 sensitivity distributions from a fit to M_{VV} for the signal qqqq$\nu\nu$ final state only for $\sqrt{s} = 1.4$ TeV. These results use the optimal jet algorithm configuration of selected PFOs and an R parameter of 0.9 in the k_t algorithm. (a) χ^2 sensitivity contours in α_4 and α_5 space. (b) χ^2 as a function of α_4 assuming $\alpha_5 = 0$. (c) χ^2 as a function of α_5 assuming $\alpha_4 = 0$. All distributions are normalised to an integrated luminosity of $\mathcal{L}_{int} = 1.5$ ab^{-1}.

2.4.3 Lepton Finding

An isolated lepton finder [29] was included in the analysis chain to reject background final states containing primary leptons. Leptons produced via hadronisation are unlikely to be flagged as isolated because all hadronisation products are boosted along the direction of the parent quark. This means isolated leptons are likely to correspond to primary leptons, which makes the number of isolated leptons a powerful discriminating variable to use in event selection.

The isolated lepton finder determines whether a PFO is an electron or muon by first checking that the PFO has a single charged particle track associated to it. If that is the case, the calorimetric energy deposits of the PFO are examined to see if they are consistent with what is expected for an electron or muon. If they are consistent with expectations, the properties of the charged particle track are examined to determine whether the track originates from the IP. If the PFO is deemed to have originated from the IP, isolation checks, which examine the energy deposited in the calorimeters within a cone surrounding the PFO, are applied to determine whether the particles belongs to a jet. If the PFO does not appear to belong to a jet then it is counted as an isolated lepton. The fraction of events rejected by the lepton finder is summarised in table 2.4.

Final State	$\epsilon_{\text{Lepton Finding}}$
$e^+e^- \rightarrow \nu\nu qqqq$	99.7
$e^+e^- \rightarrow \nu lqqqq$	48.9

Table 2.4: The fraction of events rejected by of isolated lepton finding for $\sqrt{s} = 1.4$ TeV for the $\nu\nu qqqq$ and l$\nu qqqq$ final states.

2.4.4 Discriminant Variables

The next stage of the analysis involved the calculation of a number of event-based variables that were found to be useful for this analysis. The variables that were calculated are as follows

- **Particle level** variables:
    - Number of PFOs in each jet;
    - Energy of the highest energy PFO;

- Energy of the highest energy electron;

- Cosine of the polar angle of the highest energy track;

- The number of isolated leptons found using the isolated lepton finder.

- **Candidate boson** variables:

  - Energy of the candidate bosons;

  - Invariant mass of the candidate bosons;

  - Acolinearity of the candidate boson pair, which is defined as 180 degrees minus the opening angle of the pair of bosons in the rest frame of the detector.

- **Event based** variables:

  - The invariant mass of the visible system, M_{VV};

  - The vector sum of the transverse momentum of all PFOs in the event;

  - Sphericity, defined through the sphericity tensor S^{ab};

$$S^{ab} = \frac{\Sigma_i p_i^\alpha p_j^\alpha}{\Sigma_{i,\alpha=1,2,3} |p_i^\alpha|^2} \tag{2.8}$$

Where p_i are the components of the momenta of PFO i in the rest frame of the detector and the sum Σ_i runs over all particles in the event. Sphericity is defined as $S = \frac{3}{2}(\lambda_2 + \lambda_3)$, where λ_i are the eigenvalues of the sphericity tensor defined such $\lambda_1 \geq \lambda_2 \geq \lambda_3$. This provides a measure of how spherical the reconstructed event topology is with isotropic events having $S \approx 1$, while two jet events have $S \approx 0$.

- **Jet clustering parameters** variables:

  - The y_{ij} variables where $i = 3, 4$ and $j = i+1$. These are the smallest k_t distance found when combining j jets into i jets.

2.4.5 Jet Energy Resolution at CLIC

The importance of the jet energy resolution, which is extensively discussed in chapters **??** and 1, should be emphasised at this point. Many of the discriminant variables that are calculated for this analysis are dependant upon the jet energy resolution. In particular,
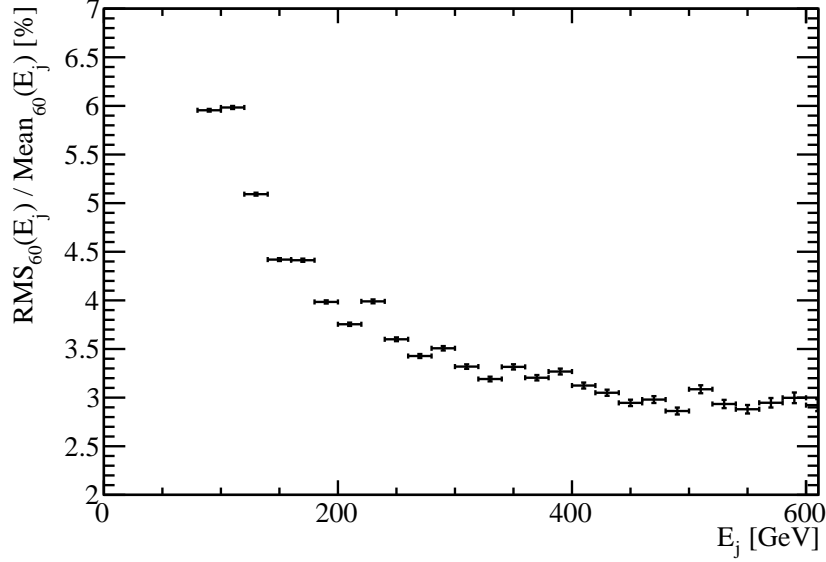
Figure 2.7: The jet energy resolution as a function of the jet energy for the $\nu\nu$qqqq final state for $\sqrt{s} = 1.4$ TeV.

all variables related to the candidate bosons that are formed from pairing up jets, are dependent upon the measurement of jet energies.

Figure 2.7 shows the jet energy resolution as a function of the MC jet energy for the $\nu\nu$qqqq event sample used in the $\sqrt{s} = 1.4$ TeV analysis. The MC jet energy was obtained by pairing up quarks appearing in the final state to the reconstructed jets. The events were then binned in terms of their MC jet energy and the jet energy resolution calculated for each bin. When calculating the jet energy resolution, a narrower range of jet energies was used in compared to previous studies, 60% of the data with narrowest RMS as opposed to 90%, to minimise the effects of jet finding and beam-induced backgrounds. The jet energy resolutions reported here are worse than those quoted in earlier chapters. This is to be expected given the effects of jet finding and beam-induced backgrounds.

2.5 Event Selection

This section discusses the event selection procedure. The goal of this procedure is to isolate the $\nu\nu$qqqq final state from the background final states, i.e. those containing two and four primary quarks. The procedure consists of a set of preselection cuts followed by the application of a multivariate analysis (MVA). All event numbers have

been normalised, prior to event selection, to an integrated luminosity of $\mathcal{L}_{int} = 1.5$ ab^{-1} for the $\sqrt{s} = 1.4$ TeV analysis and $\mathcal{L}_{int} = 2$ ab^{-1} for the $\sqrt{s} = 3$ TeV analysis.

2.5.1 Preselection

A refined selection of the $\nu\nu$qqqq signal final state is achieved using a MVA, however, to ensure efficiency in the training and application of that MVA a number of simple preselection cuts were developed to veto obvious background final states prior to the application of the MVA. Preselection cuts were applied to the transverse momentum of the system and the number of isolated leptons found in the event. The raw distributions of these variables is shown in figure 2.8 and based on these distributions the following cuts were applied

- Transverse momentum of system > 100 GeV. This cut is effective due to the presence of missing energy in the form of neutrinos in the signal final state.

- Number of isolated leptons in system $= 0$. This cut is effective as the signal final state does not contain leptons, while numerous background final states do.

The impact of these preselection cuts can be found in table 2.5, which can be found on page 56.

2.5.2 Multivariate analysis

Having established the preselection cuts, a MVA was applied using the TMVA toolkit [30], to refine the event selection. The signal and background final state samples were separated into two equally sized samples; one sample was used to independently train the MVA and the other sample was used in the subsequent analysis.

The performance of several MVA classifiers was examined to determine the optimal classifier for this analysis. The MVA classifiers considered were [30]:

- **Boosted Decision Tree (BDT)**. Decision trees are formed by the sequential application of cuts that split the data into multiple classes. After the application of the final cut, the remaining classes are used to classify whether the input event corresponds to signal or background. Boosting a decision tree involves the use of several decision trees. A single classifier output is obtained from a weighted average
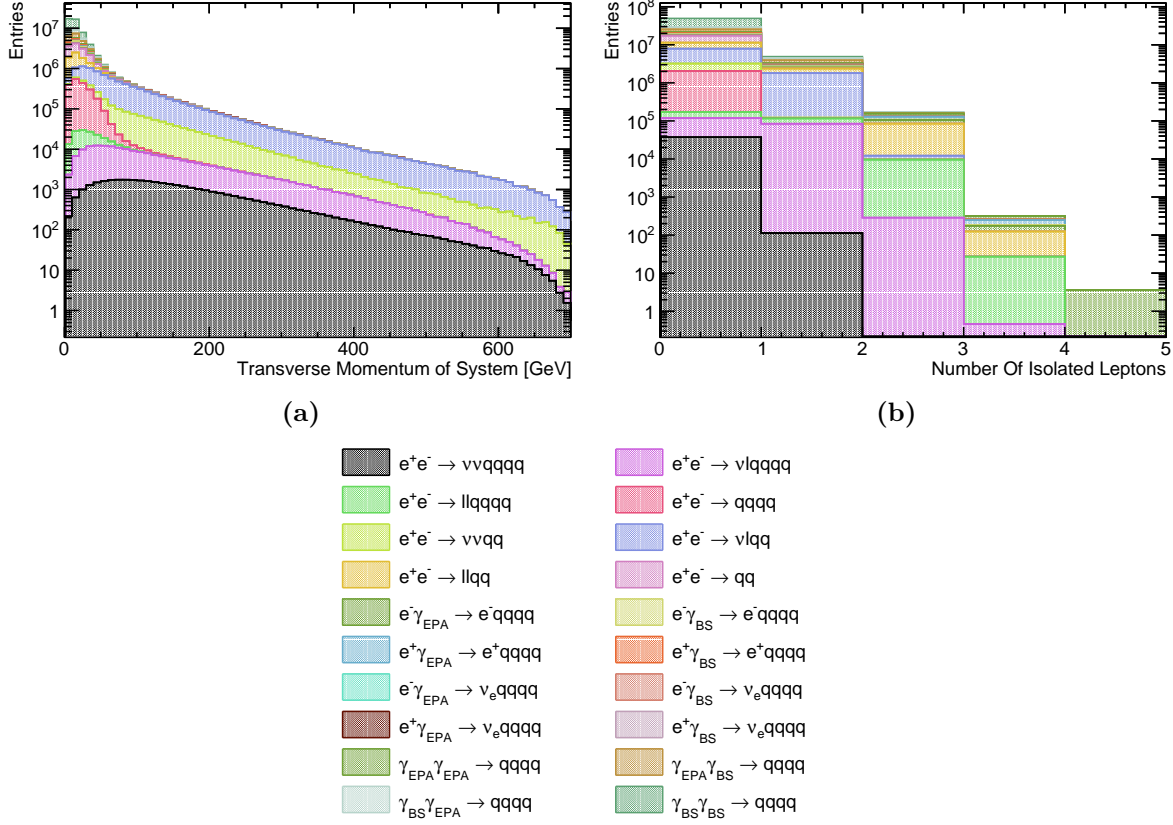
Figure 2.8: Distribution of the preselection cut variables for $\sqrt{s} = 1.4$ TeV: (a) the transverse momentum of the visible system; and (b) the number of isolated leptons in the system. All distributions are normalised to an integrated luminosity of $\mathcal{L}_{int} = 1.5 \text{ ab}^{-1}$.

of the individual decision trees. The cuts applied in the decision tree are determined using the training sample.

- **k-Nearest Neighbour (KNN)**. For a given input event, the k closest neighbours from the training sample are found. The classifier for that input event is determined as the fraction of those k events that belong to the signal sample. Distances in this classifier are defined as the Euclidean distance between events in the n-dimensional space of the variables used for training the classifier. Weights are applied when calculating the distances to account for the differing widths of the input variable distributions. The value of k used in this analysis was 20.

- **Multilayer Perceptron (MLP)**. This is an example of a neutral network. Neutral networks consist of an interconnected series of neurons each with a different response to a set of input signals. The signal for the first layer of neurons in this case are

the event variables used to train the MVA. The input signal proceeds to travel through several layers of neurons. The number of neurons in a given layer is reduced as the number of layers passed through increases until two neurons are left, one corresponding to signal and the other background. The neuron giving the larger response in the final layer determines the event classifier. The training sample is used to determine the response of each neurons in the network.

- **Fisher and H-Matrix Discriminants**. These procedures involve the calculation of a hyperplane in n-dimensional space that maximally separates signal and background events in the training sample. The location of an input event in that n-dimensional space with respect to that hyperplane determines the classifier for the event. The hyperplane is determined by maximising the differences between the means of the input event variables normalised by a measure of their spread. Both the Fisher and H-Matrix discriminants search for the hyperplane in n-dimensional space, however, the Fisher discriminant begins this procedure by transforming the input variables into a variable space with zero linear correlations.

- **Likelihood**. The likelihood is determined using the probability density function (PDF) for each of the input variables. PDFs are determined using the training sample for both signal and background events. For a given event, the likelihood is given by the product of the probability of obtaining each of the input variables for that event. The signal and background likelihoods are calculated using the signal and background PDFs respectively and the ratio of the signal likelihood to the sum of the signal and background likelihoods gives the event classifier.

The input variables used for these MVA classifiers were:

- Number of PFOs in each jet;

- Energy of the highest energy PFO;

- Energy of the highest energy electron;

- Cosine of the polar angle of the highest energy track;

- Energy of the candidate bosons;

- Invariant mass of the candidate bosons;

- Acolinearity of the candidate boson pair;

- The vector sum of the transverse momentum of all PFOs in the event;
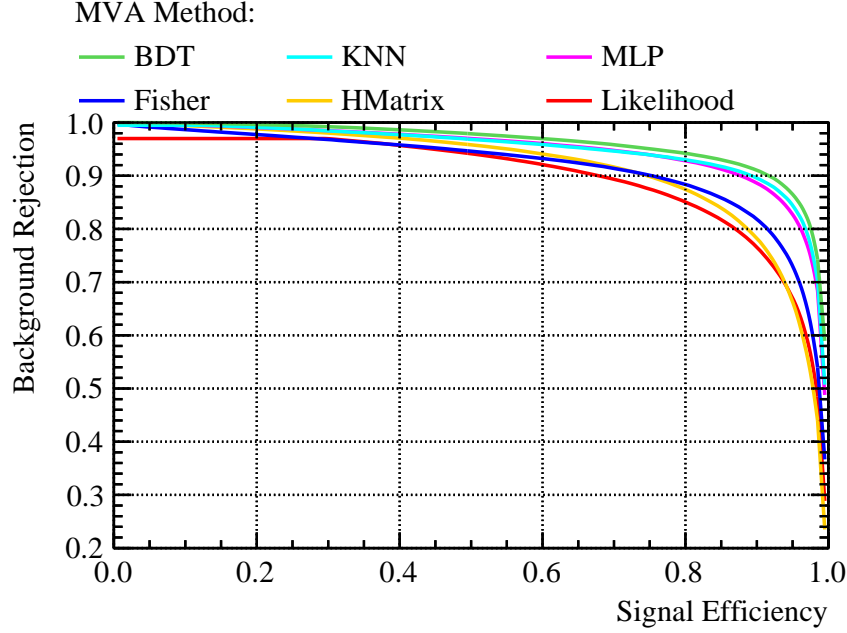
Figure 2.9: Background rejection as a function of signal efficiency for a variety of MVA options for $\sqrt{s} = 1.4$ TeV.

- The sphericity of the event;

- The derived jet clustering parameter variables $-\log_{10}(y_{ij})$ where y_{ij} are jet clustering parameters, $i = 3, 4$ and $j = i + 1$.

Figure 2.9 shows the background rejection, which is equivalent to one minus the background efficiency, as a function of signal efficiency for various MVA classifiers. Efficiency is defined as the fraction of events classified as signal by the MVA. The efficiencies reported by TMVA are calculated after the application of the preselection cuts, which are described in section 2.5.1.

The classifier giving the optimal performance in terms of signal efficiency and background rejection was the BDT. The performance of the BDT was optimised further by varying the number of trees used and the depth of the trees. An optimal significance, $S/\sqrt{(S + B)}$, where S and B are the number of signal and background events passing the preselection respectively, of 52.7 was obtained using the optimised BDT.

2.5.3 Event Selection Summary

The event selection is summarised using the distribution of the invariant mass of the candidate bosons, which for the signal final state should peak around the W mass. This distribution is shown in figure 2.10 with: no event selection; with the preselection cuts applied; and with both preselections cuts and MVA applied. The event selection efficiencies are also summarised in table 2.5.

As expected the dominant background processes after the MVA is applied are those that have the same topology as the signal process, i.e. four primary quarks with missing energy. Two smaller sources of background are also present: two jet events with missing energy that are confused with four jet events with missing energy and events where a lepton is not properly reconstructed causing the event to look like four jets and missing energy.

2.6 Anomalous Coupling Fitting Methodology

This section describes the procedure used for constructing the χ^2 surface and the subsequent confidence contours used to determine the sensitivity of CLIC to the anomalous gauge couplings α_4 and α_5.

2.6.1 Sensitive Distribution

The sensitivity of CLIC to the anomalous gauge couplings will be determined through the use of a χ^2 fit. Three variables showing sensitivity to the anomalous gauge couplings were considered for use in the χ^2 fit:

- M_{VV}. The invariant mass of the visible system;

- $\cos\theta^*_{Bosons}$. The angle between the boost direction and the back-to-back candidate bosons in the rest frame of the visible system;

- $\cos\theta^*_{Jets}$. The angle between the boost direction and the back-to-back jets in the rest frame of the candidate bosons. As each event contains two candidate bosons, there are two $\cos\theta^*_{Jets}$ variables per event.

(a)

(b)

(c)
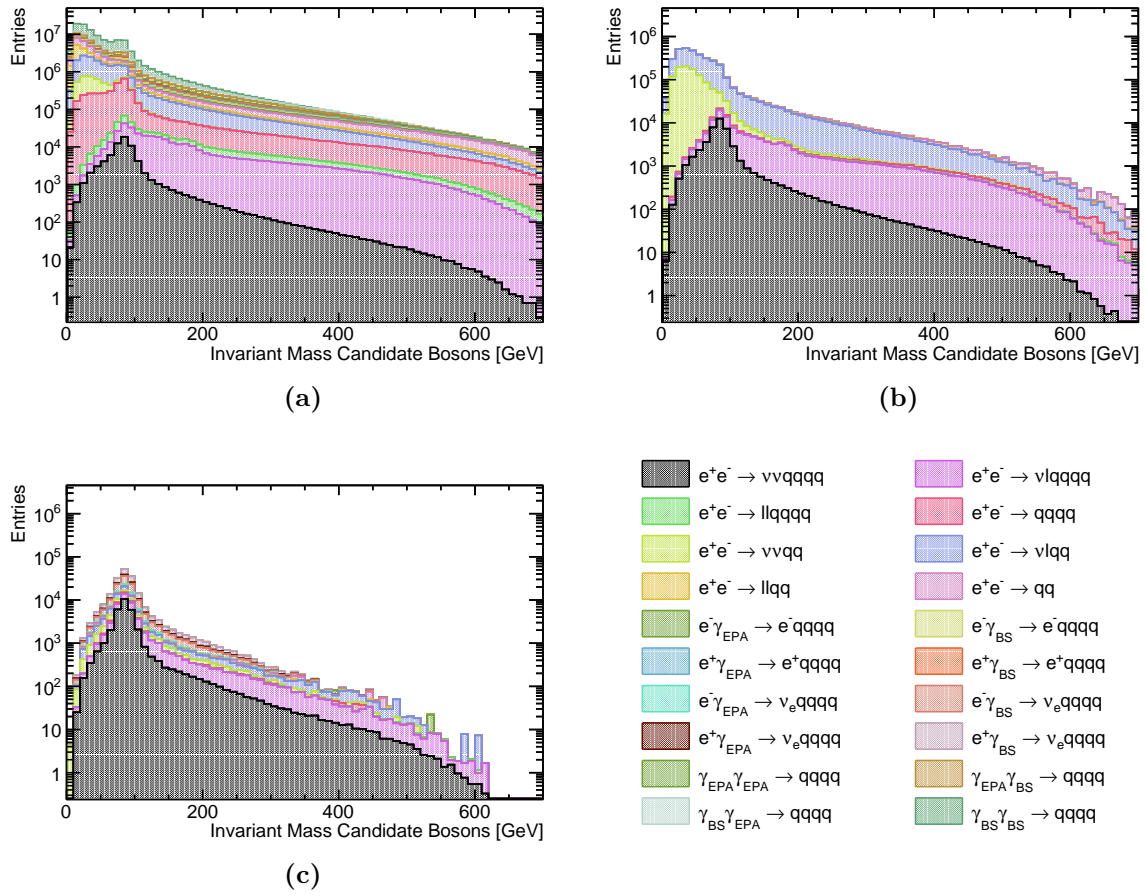
Figure 2.10: Impact of preselection and MVA on the reconstructed invariant mass of the
candidate bosons for $\sqrt{s} = 1.4$ TeV: (a) no cuts; (b) after preselection; and
(c) after preselection and MVA. All distributions correspond to an integrated
luminosity of $\mathcal{L}_{int} = 1.5$ ab^{-1}.

Final State	ϵ_{presel}	ϵ_{BDT}	N_{BDT}
$e^+e^- \to \nu\nu qqqq$	64.1%	44.5%	16,470
$e^+e^- \to \nu l qqqq$	26.1%	5.2%	8,582
$e^+e^- \to ll qqqq$	0.8%	0.1%	100
$e^+e^- \to qqqq$	0.3%	0.1%	1,698
$e^+e^- \to \nu\nu qq$	43.4%	0.5%	5,351
$e^+e^- \to \nu l qq$	19.1%	0.1%	9,319
$e^+e^- \to ll qq$	0.1%	-	234
$e^+e^- \to qq$	0.6%	-	1,586
$e^-\gamma_{\text{EPA}} \to e^- qqqq$	0.2%	-	48
$e^-\gamma_{\text{BS}} \to e^- qqqq$	0.1%	-	42
$e^+\gamma_{\text{EPA}} \to e^+ qqqq$	0.3%	-	19
$e^+\gamma_{\text{BS}} \to e^+ qqqq$	-	-	65
$e^-\gamma_{\text{EPA}} \to \nu_e qqqq$	26.0%	9.0%	4,421
$e^-\gamma_{\text{BS}} \to \nu_e qqqq$	36.1%	15.0%	23,150
$e^+\gamma_{\text{EPA}} \to \bar{\nu}_e qqqq$	25.9%	9.2%	4,495
$e^+\gamma_{\text{BS}} \to \bar{\nu}_e qqqq$	36.4%	15.3%	23,410
$\gamma_{\text{EPA}}\gamma_{\text{EPA}} \to qqqq$	0.2%	-	81
$\gamma_{\text{EPA}}\gamma_{\text{BS}} \to qqqq$	0.1%	-	55
$\gamma_{\text{BS}}\gamma_{\text{EPA}} \to qqqq$	-	-	53
$\gamma_{\text{BS}}\gamma_{\text{BS}} \to qqqq$	-	-	0

Table 2.5: Event selection efficiencies for $\sqrt{s} = 1.4$ TeV. In the above table, ϵ_{presel} denotes the number of events passing the preselection as a fraction of the total number of events, while ϵ_{BDT} denotes the number of events passing both the preselection and the BDT as a fraction of the total number of events. The EPA and BS subscript on the incoming photon indicates whether the photon is generated from the equivalent photon approximation or beamstrahlung. Entries with a dash indicate an efficiency of less than 0.1%. The event numbers correspond to an integrated luminosity of $\mathcal{L}_{int} = 1.5$ ab^{-1}.

Figure 2.11 shows the distribution of these variables for the $\nu\nu qqqq$ final state for selected values of the anomalous gauge couplings α_4 and α_5. A χ^2 fit to each of these variables was applied to obtain confidence limits on the sensitivity of CLIC to the anomalous gauge couplings, as described in section 2.6.2. The distributions used for the χ^2 fit contained signal and background events that passed event selection. Table 2.6 shows the one σ confidence limits on the measurement of α_4 and α_5 obtained using

each of the variables considered. The M_{VV} distribution shows the greatest sensitive to the anomalous gauge couplings; therefore, it will be used by all subsequent χ^2 fits when reporting sensitivities. This distribution shows the greatest sensitivity of the variables considered because the couplings primarily affect events with large values of M_{VV} and there are relatively few of these events.

Sensitive Variable	One σ Confidence Limits
M_{VV}	$-0.0082 < \alpha_4 < 0.0116$
	$-0.0055 < \alpha_5 < 0.0078$
$\cos\theta^*_{Bosons}$	$-0.0111 < \alpha_4 < 0.0155$
	$-0.0082 < \alpha_5 < 0.0110$
$\cos\theta^*_{Jets}$	$-0.0100 < \alpha_4 < 0.0142$
	$-0.0070 < \alpha_5 < 0.0098$

Table 2.6: One σ confidence limits on the measurement of α_4 and α_5 obtained at $\sqrt{s} = 1.4$ TeV. These sensitivities include the affect from backgrounds and event selection.

2.6.2 χ^2 Surface and Confidence Limit Definition

A χ^2 surface was used to determine confidence limits on the anomalous gauge couplings given the null hypothesis that $\alpha_4 = \alpha_5 = 0$. This surface is defined as

$$\chi^2 = \sum_i \frac{(O_i - E_i)^2}{E_i}, \tag{2.9}$$

where O_i is the observed, $\alpha_4 = \alpha_5 = 0$, and E_i the expected, $\alpha_4 \neq 0$ and $\alpha_5 \neq 0$, bin content for bin i in the distribution of interest. The summation Σ_i runs over bins in the distribution of interest.

When applying the χ^2 fit to the M_{VV} distribution, the distribution was binned using 13 bins as shown in figure 2.12. The first bin spanned the invariant mass range between 0 GeV and 200 GeV, this was followed by 11 bins of width 100 GeV ranging from 200 GeV to 1300 GeV and finally the last bin contained all invariant masses above 1300 GeV. The expanded bin widths at the tails of the distribution were chosen to ensure the bin contents were sufficiently large to give a reliable estimate the likelihood function using the χ^2 parameter. This choice of bin width also ensured the bin contents were sufficiently large to minimise fluctuations arising from individual events with large weights. When applying

(a)

(b)

(c)
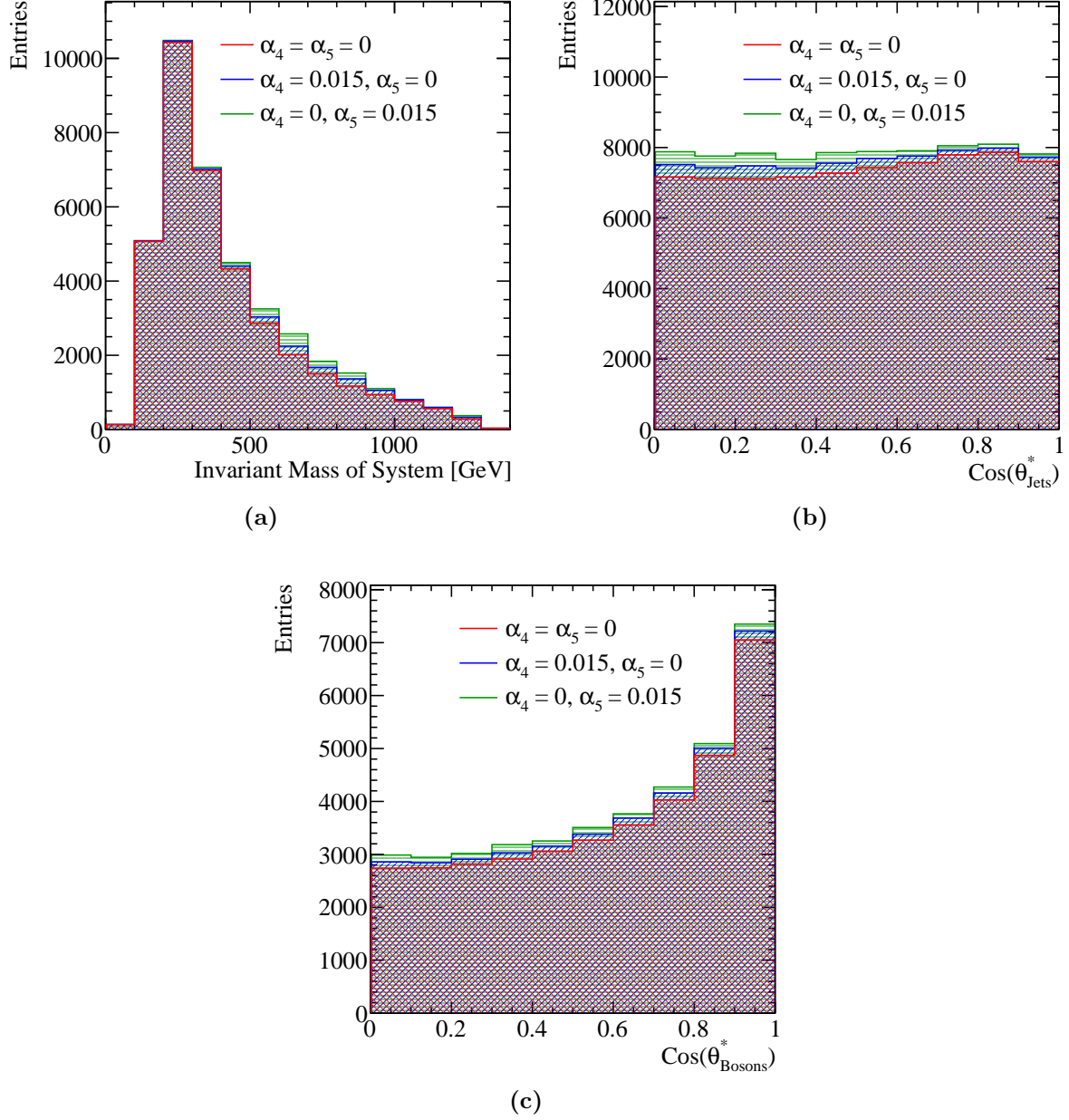
Figure 2.11: The distributions of (a) M_{VV}, (b) $\cos\theta^*_{Jets}$ and (c) $\cos\theta^*_{Bosons}$ for selected values
of the anomalous gauge couplings α_4 and α_5 for the $\nu\nu$qqqq final state for
$\sqrt{s} = 1.4$ TeV. The jet algorithm used was the longitudinally invariant k_t
algorithm with an R parameter of 0.9 and Selected PFOs. All distributions are
normalised to an integrated luminosity of $\mathcal{L}_{int} = 1.5$ ab^{-1}.

the χ^2 fit to distributions of the $\cos\theta^*_{Bosons}$ and $\cos\theta^*_{Jets}$ variables, the distributions were
binned using 10 bins ranging from zero to one. As there are two $\cos\theta^*_{Jets}$ variables
per event, the χ^2 fit was applied to a two dimensional distribution of $\cos\theta^*_{Jets}$, where

a distinction between the two $\cos\theta^*_{Jets}$ variables was made based on the energy of the candidate bosons. The use of a two dimensional distribution in the χ^2 fit was needed to account for any correlation between the two $\cos\theta^*_{Jets}$ variables.

Confidence limits describing the sensitivity of the CLIC experiment to the anomalous gauge couplings were found by examining the χ^2 surface in the space of α_4 and α_5. Deviations from the minima of this surface, which by construction occurs at $\alpha_4 = \alpha_5 = 0$, yield confidence limits that indicate the probability of observing a particular value of α_4 and α_5 given the null hypothesis that $\alpha_4 = \alpha_5 = 0$. The confidence limits reported in subsequent sections, 68%, 90% and 99%, are defined using fixed deviations from the minima of χ^2 surface ($\Delta\chi^2$) of 2.28, 4.61 and 9.21 respectively.

Confidence limits on the individual parameters α_4 and α_5 were determined by setting the corresponding coupling term to zero and examining the remaining one dimensional χ^2 distribution. A fourth order polynomial was fitted to the minima of this distribution and the one sigma confidence limit defined using $\Delta\chi^2 = 1$. The definition of a one sigma confidence limit accounts for changes in the number of degrees of freedom in the fit, therefore, it changes when fixing the corresponding coupling term to zero.
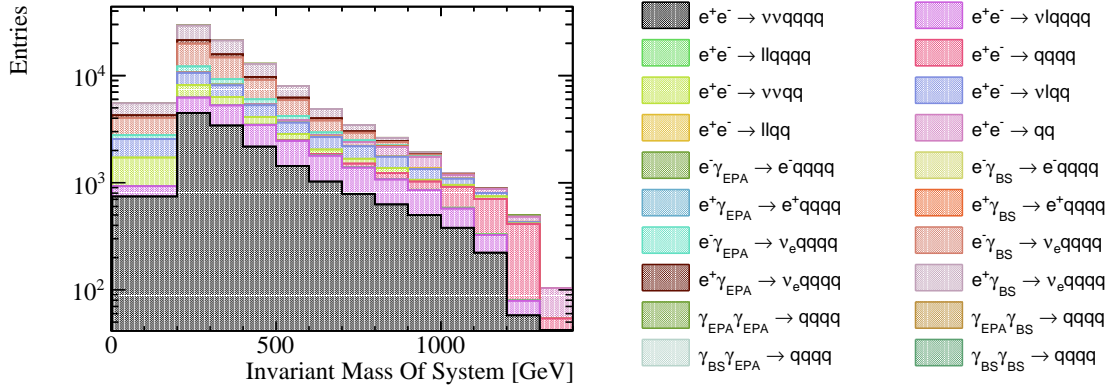
Figure 2.12: The distribution of the invariant mass of the system for both signal and background finals states that is used in the χ^2 fit for $\sqrt{s} = 1.4$ TeV. The distribution includes effect of event selection and corresponds to an integrated luminosity of $\mathcal{L}_{int} = 1.5$ ab^{-1}.

2.6.3 Event Weight Interpolation Scheme

In order to obtain a smooth χ^2 surface a fine sampling of the event weights in the α_4 and α_5 space is required, however, it is unfeasible to generate a finely sampled grid of event

weights on an event by event basis because event generation is highly CPU intensive. To resolve this issue, an interpolation scheme was applied to determine the event weights within a sampled region of the α_4 and α_5 space. This allows for an infinite sampling of the event weights in the space of α_4 and α_5 without having to call the generator an infinite number of times.

A bicubic interpolation scheme, cubic interpolation along the two dimensions, was applied to the event weights produced by the generator. This procedure is best illustrated by figure 2.13, which shows the interpolated event weight surface superimposed with the raw event weights from the generator for four $\nu\nu$qqqq events for $\sqrt{s} = 1.4$ TeV. This interpolation scheme produces a smooth and continuous surface that can be used for generating a smooth χ^2 surface.

2.7 Results

The sensitivity of the CLIC experiment to the anomalous gauge couplings α_4 and α_5 for $\sqrt{s} = 1.4$ TeV is shown in figure 2.14a. This result shows the sensitivity after the application of preselection and MVA purposed to remove the included background channels. These contours yield the one σ confidence limits for CLIC operating for $\sqrt{s} = 1.4$ TeV of

$$-0.0082 < \alpha_4 < 0.0116, \tag{2.10}$$

$$-0.0055 < \alpha_5 < 0.0078. \tag{2.11}$$

2.7.1 Systematic Uncertainties

A source of systematic error in this experiment is the uncertainty on the cross-sections for the signal and background processes. Based on the event selection summary shown in table 2.5, the dominant source of background in this analysis comes from the $e^{\pm}\gamma_{BS} \rightarrow \nu_e$qqqq processes. Therefore, uncertainties on the cross-section for these processes, as well as the signal process $e^+e^- \rightarrow \nu\nu$qqqq, will be considered.

The uncertainty on the cross-section for a given process is included in the χ^2 definition through the use of a nuisance parameter. This procedure allows the cross-section for a process to fluctuate, however, the magnitude of the fluctuation, r, is moderated by an
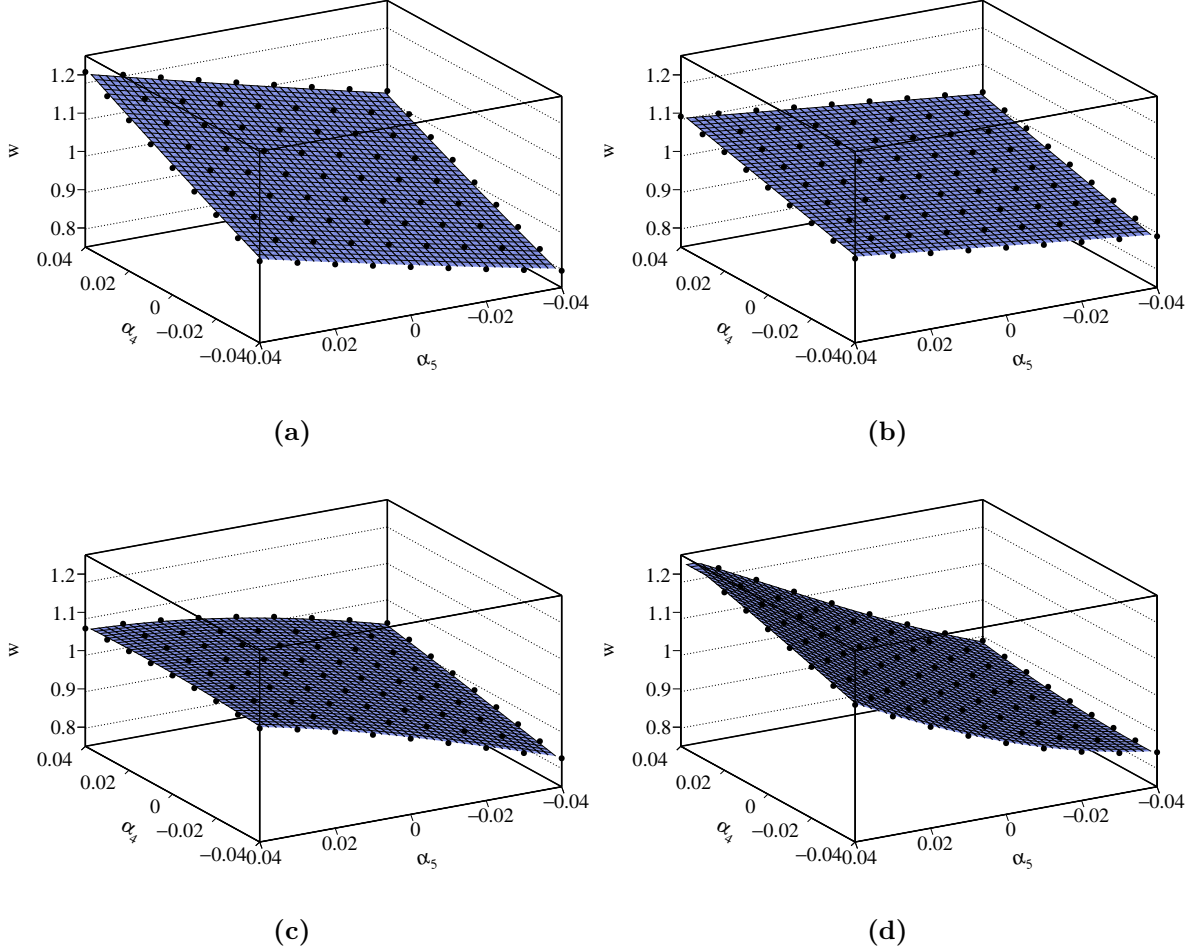
Figure 2.13: The event weight, w, as a function of the anomalous couplings α_4 and α_5 for a selection of $\sqrt{s} = 1.4$ TeV $\nu\nu$qqqq final state events. The black circles show the event weight produced from the generator and the blue surface is determined using bicubic interpolation between these points.

additional penalty term in the χ^2 as follows

$$\chi^2(r) = \sum_i \frac{(O_i - E_i(r))^2}{E_i(r)} + \frac{(r-1)^2}{\sigma_r^2}, \tag{2.12}$$

where O_i is the observed, $\alpha_4 = \alpha_5 = 0$, bin content for bin i in the distribution of M_{VV} with no background fluctuations and $E_i(r)$ is the expected, $\alpha_4 \neq 0$ and $\alpha_5 \neq 0$, bin content for bin i in the distribution of M_{VV} where the cross-section for the process of interest has been fluctuated by the factor r. The sum $\sum_i$ runs over the bins in the M_{VV} distribution. The σ_r variable is the width of the distribution of r, which indicates
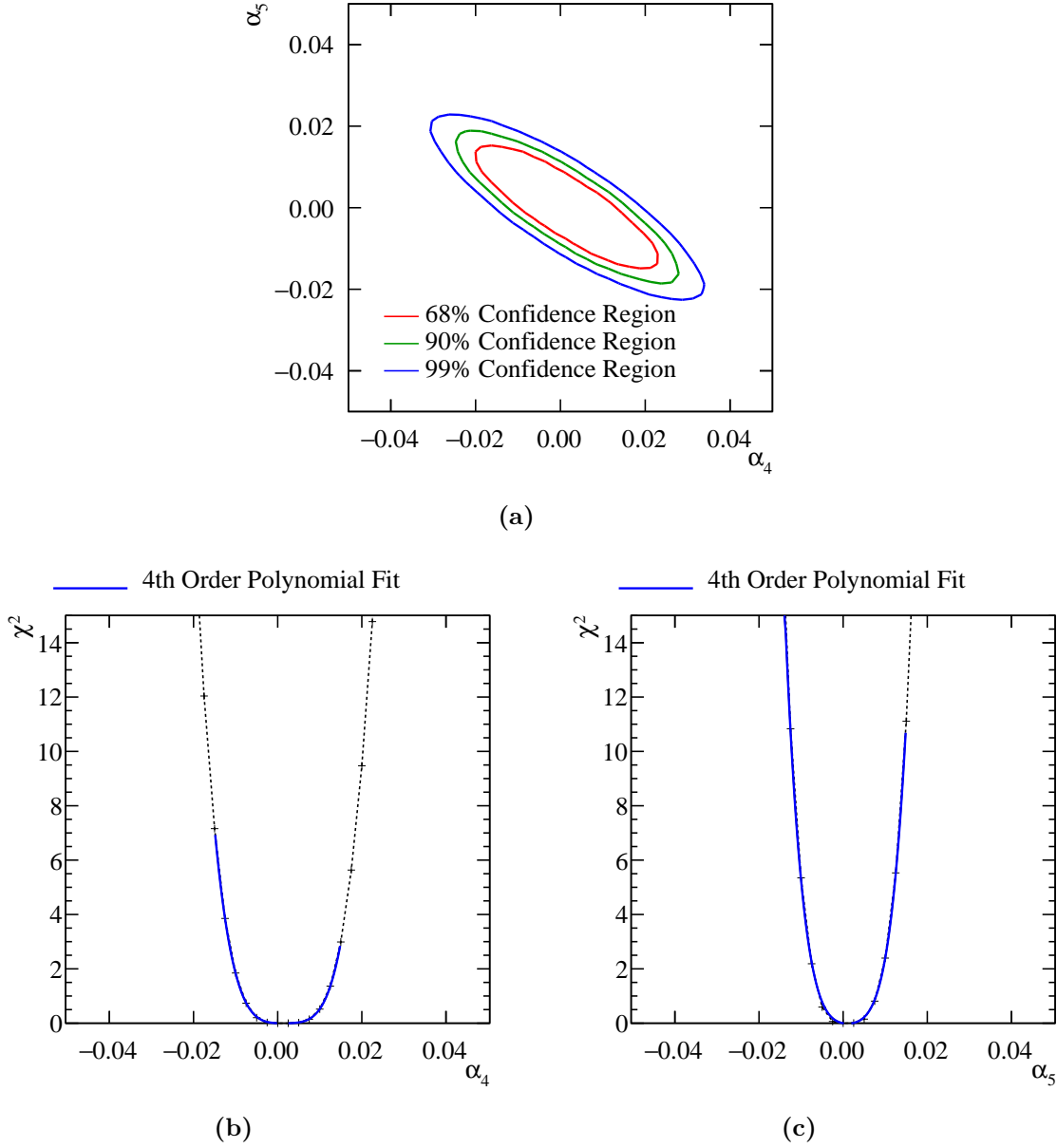
Figure 2.14: χ^2 sensitivity distributions from a fit to M_{VV} for $\sqrt{s} = 1.4$ TeV. Results include the effect of backgrounds after the application of a series of preselection cuts and MVA. (a) χ^2 sensitivity contours in α_4 and α_5 space. (b) χ^2 as a function of α_4 assuming $\alpha_5 = 0$. (c) χ^2 as a function of α_5 assuming $\alpha_4 = 0$.

the uncertainty on the measurement of the cross-section of interest. A χ^2 surface is constructed in the space of α_4 and α_5 by minimising $\chi^2(r)$ at each point.

The 68% confidence region is shown with the inclusion of a nuisance parameter for the signal process $e^+e^- \rightarrow \nu\nu qqqq$ and the dominant background processes $e^{\pm}\gamma_{BS} \rightarrow \nu_e qqqq$

in figures 2.15a and 2.15b respectively. Minimal changes in sensitivity are observed when allowing the signal and dominant backgrounds to fluctuate. This can be understood by considering the shape of the M_{VV} distribution for the signal and dominant background processes, which is shown in figure 2.16. These distribution shows that anomalous couplings primarily affect events with large invariant masses, while both the signal and dominant backgrounds peak at low invariant masses. Therefore, by fluctuating the cross-section for the signal and dominant background processes, it is not possible to gain a significantly better match between the observed and expected bin contents in the M_{VV} distribution. This is encouraging as despite the $e^{\pm}\gamma_{BS} \to \nu_e qqqq$ backgrounds dominating the χ^2 fit that determines the sensitivity of CLIC to the anomalous gauge couplings, precise knowledge of their cross-section is not crucial. As the uncertainty on these cross-sections does not significantly affect the confidence regions, no cross-section uncertainties are accounted for when reporting the sensitivity of CLIC to the anomalous gauge couplings elsewhere in this analysis.

(a) (b)
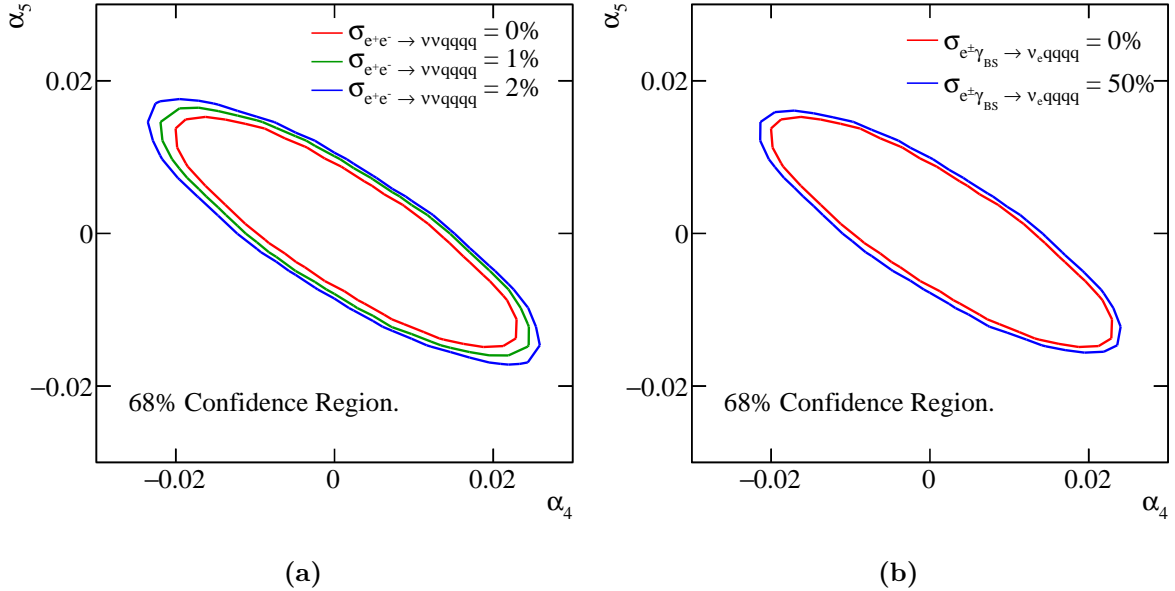
Figure 2.15: The 68% confidence region including the effect of uncertainties in the cross-section for (a) the signal process $e^+e^- \to \nu\nu qqqq$ and (b) the dominant background processes $e^{\pm}\gamma_{BS} \to \nu_e qqqq$.
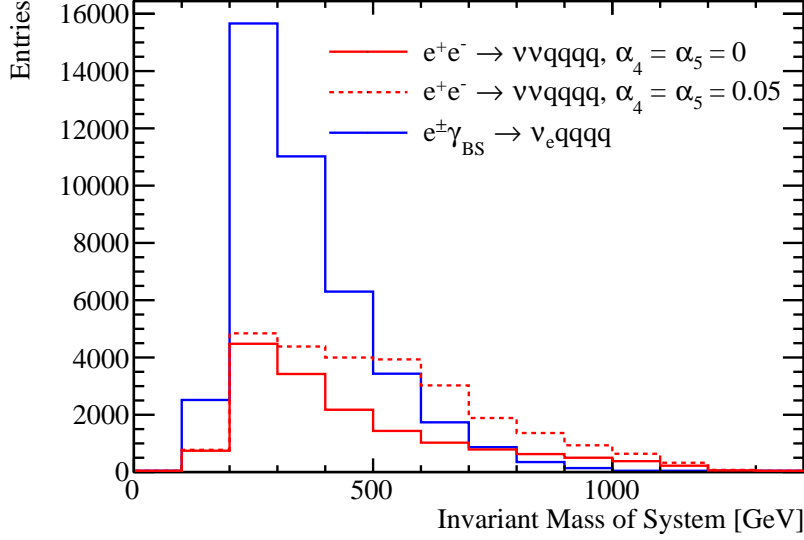
Figure 2.16: Distributions of M_{VV} for the $e^+e^- \to \nu\nu$qqqq signal process, with and without the effect from anomalous couplings, and the combined dominant background processes $e^{\pm}\gamma_{BS} \to \nu_e$qqqq. All distributions include the effect of event selection and correspond to an integrated luminosity of $\mathcal{L}_{int} = 1.5$ ab^{-1}.

2.8 Sensitivity for $\sqrt{s} = 3$ TeV

The anomalous gauge coupling sensitivity study described in this chapter was repeated for CLIC operating for $\sqrt{s} = 3$ TeV. As this analysis largely mirrors that of the $\sqrt{s} = 1.4$ TeV analysis, this section focuses on the differences between the two analyses.

The signal and background final states for the $\sqrt{s} = 3$ TeV analysis were identical to those used for the $\sqrt{s} = 1.4$ TeV analysis. Cross sections for these processes for $\sqrt{s} = 3$ TeV are given in table 2.7. The data analysis and event selection procedures used for $\sqrt{s} = 3$ TeV mirrored those used for $\sqrt{s} = 1.4$ TeV.

Jet finding was performed using the longitudinally invariant k_t algorithm as described in section 2.4.2. The jet algorithm configuration was optimised using the sensitivity of CLIC to the anomalous gauge couplings using pure signal only, as described in section 2.4.2.1. The optimal jet algorithm configuration for $\sqrt{s} = 3$ TeV used tight selected PFOs and an R parameter of 1.1. As the cross-section for the $\gamma\gamma \to hadrons$ increases with energy, the effect of these background is more problematic for $\sqrt{s} = 3$ TeV than for $\sqrt{s} = 1.4$ TeV [26]. Therefore, it is to be expected that the optimal PFO selection for

Final State	Cross Section [fb]
$e^+e^- \to \nu\nu qqqq$	71.5
$e^+e^- \to \nu lqqqq$	106.6
$e^+e^- \to llqqqq$	169.3
$e^+e^- \to qqqq$	546.5
$e^+e^- \to \nu\nu qq$	1317.5
$e^+e^- \to \nu lqq$	5560.9
$e^+e^- \to llqq$	3319.6
$e^+e^- \to qq$	2948.9
$e^-\gamma_{\text{EPA}} \to e^- qqqq$	287.8
$e^-\gamma_{\text{BS}} \to e^- qqqq$	1268.6
$e^+\gamma_{\text{EPA}} \to e^+ qqqq$	287.8
$e^+\gamma_{\text{BS}} \to e^+ qqqq$	1267.3
$e^-\gamma_{\text{EPA}} \to \nu_e qqqq$	54.2
$e^-\gamma_{\text{BS}} \to \nu_e qqqq$	262.5
$e^+\gamma_{\text{EPA}} \to \overline{\nu}_e qqqq$	54.2
$e^+\gamma_{\text{BS}} \to \overline{\nu}_e qqqq$	262.3
$\gamma_{\text{EPA}}\gamma_{\text{EPA}} \to qqqq$	402.7
$\gamma_{\text{EPA}}\gamma_{\text{BS}} \to qqqq$	2423.1
$\gamma_{\text{BS}}\gamma_{\text{EPA}} \to qqqq$	2420.6
$\gamma_{\text{BS}}\gamma_{\text{BS}} \to qqqq$	13050.3

Table 2.7: Cross sections of signal and background processes for $\sqrt{s} = 3$ TeV. In the above table q represents u, $\bar{\text{u}}$, d, $\bar{\text{d}}$, s, $\bar{\text{s}}$, c, $\bar{\text{c}}$, b or $\bar{\text{b}}$; l represents $e^\pm$, $\mu^\pm$ or $\tau^\pm$; and ν represents ν_e, $\overline{\nu}_e$, ν_μ, $\overline{\nu}_\mu$, ν_τ and $\overline{\nu}_\tau$. The EPA and BS subscript on the incoming photon indicates whether the photon is generated from the equivalent photon approximation or beamstrahlung.

$\sqrt{s} = 3$ TeV, tight selected PFOs, is more aggressive at vetoing these backgrounds than for $\sqrt{s} = 1.4$ TeV, selected PFOs, which is what is observed.

As opposed to training the MVA using 50% of the signal and background events, as was done for the $\sqrt{s} = 1.4$ TeV analysis, the $\sqrt{s} = 3$ TeV analysis trained the MVA using 10% of the signal and background events. This modification prevented those events with very large event weights from dominating the χ^2 fit and producing exaggerated sensitivities. The sensitivity to the anomalous gauge couplings grows with increasing centre of mass energy, therefore, for $\sqrt{s} = 1.4$ TeV very large event weights were not an

issue. The sample sizes for all signal and background processes was sufficiently large that training on 10% of the total sample was sufficient to achieve good MVA performance. Event selection for the $\sqrt{s} = 3$ TeV analysis is summarised in table 2.8.

Final State	ϵ_{presel}	ϵ_{BDT}	N_{BDT}
$e^+e^- \to \nu\nu qqqq$	74.4%	46.0%	65,740
$e^+e^- \to \nu lqqqq$	40.0%	12.0%	25,660
$e^+e^- \to llqqqq$	7.5%	1.1%	3,570
$e^+e^- \to qqqq$	3.7%	0.3%	3,224
$e^+e^- \to \nu\nu qq$	50.5%	1.2%	30,510
$e^+e^- \to \nu lqq$	32.0%	0.4%	48,320
$e^+e^- \to llqq$	1.4%	-	1,028
$e^+e^- \to qq$	1.4%	0.1%	3,268
$e^-\gamma_{\text{EPA}} \to e^- qqqq$	6.6%	0.8%	4,736
$e^-\gamma_{\text{BS}} \to e^- qqqq$	4.6%	0.7%	13,660
$e^+\gamma_{\text{EPA}} \to e^+ qqqq$	6.5%	0.8%	4,686
$e^+\gamma_{\text{BS}} \to e^+ qqqq$	4.7%	0.7%	13,310
$e^-\gamma_{\text{EPA}} \to \nu_e qqqq$	45.6%	17.2%	18,610
$e^-\gamma_{\text{BS}} \to \nu_e qqqq$	55.9%	26.7%	110,900
$e^+\gamma_{\text{EPA}} \to \bar{\nu}_e qqqq$	45.9%	17.3%	18,750
$e^+\gamma_{\text{BS}} \to \bar{\nu}_e qqqq$	56.5%	27.4%	113,700
$\gamma_{\text{EPA}}\gamma_{\text{EPA}} \to qqqq$	5.3%	0.7%	5,531
$\gamma_{\text{EPA}}\gamma_{\text{BS}} \to qqqq$	3.5%	0.4%	16,640
$\gamma_{\text{BS}}\gamma_{\text{EPA}} \to qqqq$	3.5%	0.4%	15,900
$\gamma_{\text{BS}}\gamma_{\text{BS}} \to qqqq$	0.6%	-	4,124

Table 2.8: Event selection efficiencies for $\sqrt{s} = 3$ TeV. In the above table, ϵ_{presel} denotes the number of events passing the preselection as a fraction of the total number of events, while ϵ_{BDT} denotes the number of events passing both the preselection and the BDT as a fraction of the total number of events. The EPA and BS subscript on the incoming photon indicates whether the photon is generated from the equivalent photon approximation or beamstrahlung. Entries with a dash indicate an efficiency of less than 0.1%. The event numbers correspond to an integrated luminosity of $\mathcal{L}_{int} = 2$ ab^{-1}.

Due to the increased sensitivity of the signal sample, event weights were sampled with greater frequency in the space of α_4 and α_5 for $\sqrt{s} = 3$ TeV than for $\sqrt{s} = 1.4$ TeV analysis. Bicubic interpolation was again used to make a continuous surface for the event

weights. These event weight surfaces were then used to construct the M_{VV} distribution and the χ^2 surface used to determine the reported sensitivities. Figure 2.17 shows an example of the event weights extracted from the generator and the interpolated surface used to define the χ^2 surface as a function of α_4 and α_5 for a selected $\nu\nu$qqqq event for $\sqrt{s} = 3$ TeV.
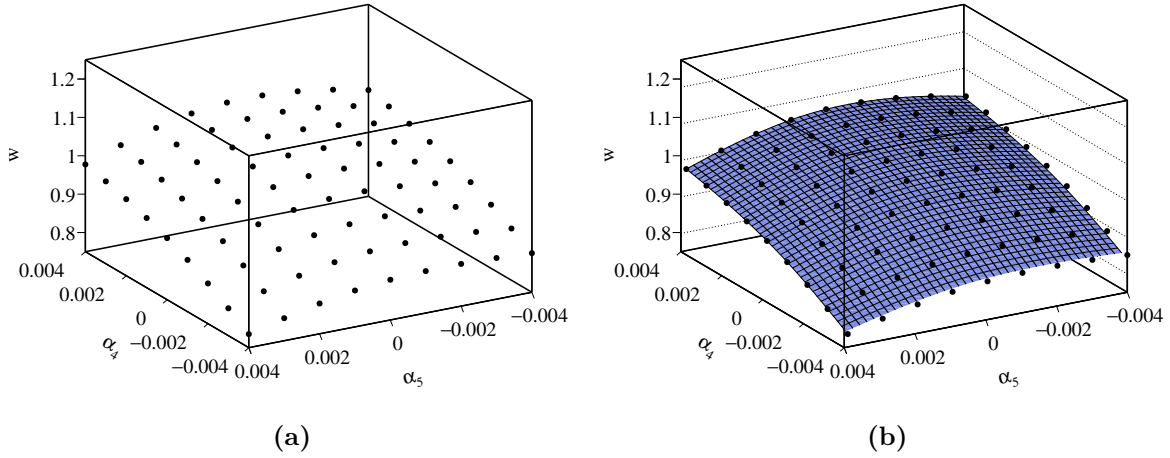
Figure 2.17: The event weights, w, as a function of the anomalous couplings α_4 and α_5 for a selected $\nu\nu$qqqq final state events for $\sqrt{s} = 3$ TeV. These weights are calculated using (a) the generator and (b) bicubic interpolation.

A χ^2 was applied to the distribution of M_{VV} to determine the sensitivity of CLIC to the anomalous gauge couplings α_4 and α_5 for $\sqrt{s} = 3$ TeV. The M_{VV} distribution used for the fit had an increased number of bins with respect to the $\sqrt{s} = 1.4$ TeV analysis; the first bin spanned the invariant mass range between 0 GeV and 200 GeV, this was followed by 27 bins of width 100 GeV ranging from 200 GeV to 1300 GeV and finally the last bin contained all invariant masses above 2900 GeV. Figure 2.18 shows the M_{VV} distribution for signal and background processes for $\sqrt{s} = 3$ TeV that was used in the χ^2 fit.

The sensitivity of the CLIC experiment to the anomalous gauge couplings α_4 and α_5 for $\sqrt{s} = 3$ TeV is shown in figure 2.19a. This result shows the sensitivity after the application of preselection and MVA, described in sections 2.5.1 and 2.5.2, purposed to remove the included background channels. These contours yield the one σ confidence limit on the measurement of

$$-0.0010 < \alpha_4 < 0.0011, \tag{2.13}$$

$$-0.0007 < \alpha_5 < 0.0007. \tag{2.14}$$

Figure 2.18: The distribution of the invariant mass of the system for both signal and background finals states that is used in the χ^2 fit for $\sqrt{s} = 3$ TeV. The distribution includes effect of event selection and corresponds to an integrated luminosity of $\mathcal{L}_{int} = 2$ ab^{-1}.

Figure 2.20 shows how the 68% confidence region for the $\sqrt{s} = 3$ TeV analysis varies with the uncertainty in the cross-section for the signal, $e^+e^- \rightarrow \nu\nu qqqq$, and dominant background processes, $e^{\pm}\gamma_{BS} \rightarrow \nu_e qqqq$. These contours were produced using a nuisance parameter as discussed in section 2.7.1. Once again, these systematic uncertainties have a small effect on the reported sensitivity of CLIC to the anomalous gauge couplings due to the shape of the M_{VV} distribution.

(a)

(b)

(c)

Figure 2.19: χ^2 sensitivity distributions from a fit to M_{VV} for $\sqrt{s} = 3$ TeV. Results include
the effect of backgrounds after the application of a series of preselection cuts
and MVA. (a) χ^2 sensitivity contours in α_4 and α_5 space. (b) χ^2 as a function
of α_4 assuming $\alpha_5 = 0$. (c) χ^2 as a function of α_5 assuming $\alpha_4 = 0$.
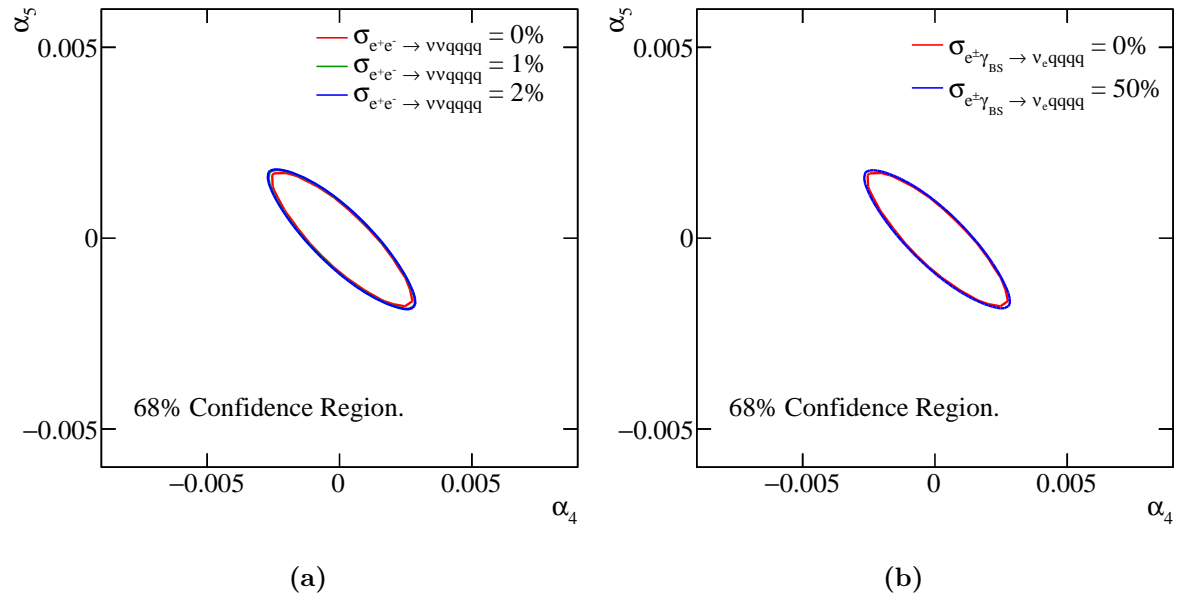
Figure 2.20: The 68% confidence region including the effect of uncertainties in the cross-section for (a) the signal process $e^+e^- \to \nu\nu qqqq$ and (b) the dominant background processes $e^{\pm}\gamma_{BS} \to \nu_e qqqq$.

Colophon

This thesis was made in LaTeX 2_ε using the "hepthesis" class [31].

Bibliography

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

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

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

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

[5] J. B. Birks. Scintillations from Organic Crystals: Specific Fluorescence and Relative Response to Different Radiations. *Proc. Phys. Soc.*, A64:874–877, 1951.

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

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

[8] G. Folger and J. P. Wellisch. String parton models in GEANT4. *eConf*, C0303241:MOMT007, 2003.

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

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

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

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

[13] A. Caldwell, L. Hervas, J. A. Parsons, F. Sciulli, W. Sippach, and L. Wai. Measurement of the time development of particle showers in a uranium scintillator calorimeter. *Nucl. Instrum. Meth.*, A330:389–404, 1993.

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

[15] E. Fermi. An attempt of a theory of beta radiation. 1. *Z. Phys.*, 88:161–177, 1934.

[16] Georges Aad et al. Evidence for Electroweak Production of $W^{\pm}W^{\pm}jj$ in pp Collisions at $\sqrt{s} = 8$ TeV with the ATLAS Detector. *Phys. Rev. Lett.*, 113(14):141803, 2014.

[17] J. Beringer et al. Review of Particle Physics (RPP). *Phys. Rev.*, D86:010001, 2012.

[18] Wolfgang Kilian, Thorsten Ohl, and Jurgen Reuter. WHIZARD: Simulating Multi-Particle Processes at LHC and ILC. *Eur. Phys. J.*, C71:1742, 2011.

[19] Mauro Moretti, Thorsten Ohl, and Jurgen Reuter. O'Mega: An Optimizing matrix element generator. 2001.

[20] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 05:026, 2006.

[21] G. Alexander et al. A Comparison of b and u d s quark jets to gluon jets. *Z. Phys.*, C69:543–560, 1996.

[22] Z. Was. TAUOLA the library for tau lepton decay, and KKMC / KORALB / KORALZ /... status report. *Nucl. Phys. Proc. Suppl.*, 98:96–102, 2001. [,96(2000)].

[23] P. Mora de Freitas and H. Videau. Detector simulation with MOKKA / GEANT4: Present and future. In *Linear colliders. Proceedings, International Workshop on physics and experiments with future electron-positron linear colliders, LCWS 2002, Seogwipo, Jeju Island, Korea, August 26-30, 2002*, pages 623–627, 2002.

[24] S. Agostinelli et al. GEANT4: A Simulation toolkit. *Nucl. Instrum. Meth.*, A506:250–303, 2003.

[25] F. Gaede. Marlin and LCCD: Software tools for the ILC. *Nucl. Instrum. Meth.*, A559:177–180, 2006.

[26] J. S. Marshall, A. Münnich, and M. A. Thomson. Performance of Particle Flow Calorimetry at CLIC. *Nucl. Instrum. Meth.*, A700:153–162, 2013.

[27] W. Kilian. WHIZARD 1.0: A generic Monte-Carlo integration and event generation package for multi-particle processes.

[28] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. FastJet User Manual. *Eur. Phys. J.*, C72:1896, 2012.

[29] O. Wendt, F. Gaede, and T. Kramer. Event Reconstruction with MarlinReco at the ILC. *Pramana*, 69:1109–1114, 2007.

[30] Andreas Hoecker, Peter Speckmayer, Joerg Stelzer, Jan Therhaag, Eckhard von Toerne, and Helge Voss. TMVA: Toolkit for Multivariate Data Analysis. *PoS*, ACAT:040, 2007.

[31] Andy Buckley. The hepthesis LaTeX class.