

PROSPECTS FOR HIGGS BOSON & TOP QUARK MEASURMENTS AND APPLICATIONS OF CMOS MAPS FOR DIGITAL CALORIMETRY AT FUTURE LINEAR COLLIDERS

A. Winter

*Thesis submitted for the degree of
Doctor of Philosophy*



Particle Physics Group,
School of Physics and Astronomy,
University of Birmingham.

March 21, 2018

ABSTRACT

X was measured, we showed that $Y \neq Z$ and that $M_H = 126 \text{ GeV}/c^2$.

DECLARATION OF AUTHORS CONTRIBUTION

I did this, and that, and some of the other.

ACKNOWLEDGEMENTS

I would like to thank bla, and bla ...

Motto or dedication

Contents

1	Introduction	1
2	Experiments	5
2.1	ILC	6
2.1.1	Energy Staging	7
2.1.2	Beam Production, Acceleration and Focusing	8
2.1.3	Positron Production	9
2.2	CLIC	10
2.2.1	Energy Staging	12
2.2.2	Acceleration Technology	13
2.3	Linear Collider Analysis Framework	14
2.3.1	Event Generation	14
2.3.2	Pandora Particle Flow Algorithm	15
2.4	Detectors	16
2.4.1	ILD	16
2.4.1.1	Vertexing	18
2.4.1.2	Tracking	19
2.4.1.3	Calorimetry	21
2.4.1.4	ECAL	25
2.4.1.5	HCAL	26
2.4.1.6	Muon Detection	27
2.4.2	CLIC ILD	28
3	Theory	29
3.1	The Standard Model	29
3.2	The Higgs Boson and the Origin of Mass	33
3.3	Higgs Measurements at CLIC	34
3.3.1	Higgsstrahlung	35
3.3.2	Model Independent Extraction of Higgs Couplings	36
3.4	Top Quark Physics	42
4	Higgs Analysis	46
4.1	Introduction	47
4.2	Event Reconstruction	48
4.2.1	Lepton Identification	49
4.2.2	Jet Finding	50
4.3	Flavour Tagging	51

4.4	Event Selection	51
4.5	Conclusion	53
5	Top Physics	56
5.1	Introduction	56
5.2	Event Reconstruction	58
5.2.1	Lepton Finding	58
5.2.2	Fat Jet Finding	61
5.2.2.1	Jet Association	65
5.2.3	S' Reconstruction	71
5.2.3.1	Transverse/Longitudinal Association	71
5.2.3.2	Analytic Mass Constraint	72
5.2.3.3	Collinearity	76
5.2.3.4	Kinematic Fitting	76
5.2.4	Flavour Tagging	76
5.3	Calculating AFB	78
5.4	Event Selection	78
5.4.1	NSubjettiness	79
5.4.2	Subjet Angular Distributions	79
5.4.3	Jet Multiplicity	79
5.5	Quality Cuts	79
5.5.1	Differential Measurement	79
5.6	Conclusions	79
6	DECALStudies	80
6.1	DigiMAPs	80
6.2	Design Optimization	81
6.3	Hardware studies	81
7	Conclusion	82

List of Tables

2.1	Parameters for the CLIC energy stages. The power consumptions for the 1.5 and 3 TeV stages are from the CDR; depending on the details of the upgrade they can change at the percent level [1].	11
2.2	Properties of the CLIC vertex detector assuming three pairs of layers [2]	19
3.1	Generic size of Higgs coupling modifications from the Standard Model (SM) values when all new pariles are $M \sim 1\text{TeV}$ and mixing angle satisfy precision electroweak fits. The Decoupling MSSM numbers assume $\tan \beta = 3.2$ and a stop mass of 1 TeV with $X_t = 0$ for the κ_γ prediction [3]	42
5.1	Top Pair Decay Modes	57
5.2	Top Pair Decay Modes	57
5.3	Methods used for identifying which fat jet corresponds to the hadronically decaying top	67

List of Figures

2.1	The ILC Experiment	6
2.2	Schematic Of The ILC	8
2.3	Superconducting Cavities For The ILC	9
2.4	The CLIC Experiment	10
2.5	Cross Sections For Super Symmetric Processes	13
2.6	ILD Detector	17
2.7	ILD Vertex Detector	19
2.8	Silicon Tracking Systems For ILD	21
2.9	ECAL Structure	26
2.10	HCAL Structure	27
3.1	Particles of the Standard Model	30
3.2	Cross Sections For Higgs Production Mechanisms	35
3.3	The Higgstrahlung Process	36
3.4	Reconstructed recoil mass from Higgsstrahlung process	37
3.5	Expected statistical uncertainties for Higgs measurements at 350GeV at CLIC assuming unpolarised beams	38
3.6	Expected statistical uncertainties for Higgs measurements at 1.4TeV and 3TeV at CLIC assuming unpolarised beams	39
3.7	Expected precision on model independent measurements of the Higgs couplings	39
3.8	Expected precision on model dependent measurements of the Higgs couplings at CLIC	40
3.9	Expected precision on model dependent measurements of the Higgs couplings at CMS	40
3.10	Dominant top production mechanism at electron positron colliders . .	43
3.11	Expected precision on CP conserving electroweak form factors at future colliders	45
4.1	Higgs Cross Sections	48
4.2	Signal Feynmann Diagram	48
4.3	Samples Used	49
4.4	B-Tagging Purity vs Efficiency	54
4.5	Classifier BDT response	54
4.6	Samples Used	55
5.1	Semileptonic $t\bar{t}$ decay	56
5.2	Charge Tagging Efficiency	59

5.3	Angular dependence of lepton finding for particles vs antiparticles	60
5.4	Lepton Momentum Vs Angle	61
5.5	Lepton efficiency for $ee \rightarrow H\nu\nu, H \rightarrow WW \rightarrow qql\nu$	61
5.6	Charge Tagging Efficiency After 20GeV Lepton Momentum Cut	62
5.7	Separation between W and b jet from top decay	62
5.8	Performance of jet finding algorithms	64
5.9	Performance of Valencia algorithm for high energy events	64
5.10	Comparison of reconstructed top decay angle to truth	65
5.11	Reconstructed fat jet mass	65
5.12	Reconstructed vs true top decay angles with truth level cuts	67
5.13	Reconstructed $\text{Cos}\theta$ distribution for various jet association methods .	68
5.14	Mean and width from fitting $\Delta\text{Cos}\theta_{True-Reco}$ to a gaussian	69
5.15	Efficiency for reconstructing the hadronically decaying top in the correct $\text{Cos}\theta$ bin	70
5.16	Expected s' spectrum for $t\bar{t}$ at 1.4 TeV	70
5.17	Angular energy distribution of initial state photons	72
5.18	Reconstructed s' vs true s' for Transverse/Longitudinal Association Method	73
5.19	Reconstructed s' vs true s' for mass constraint method	74
5.20	Mass of reconstructed top when using mass constraint method	74
5.21	Collinearity of $t\bar{t}$ pair	75
5.22	Reconstructed s' vs true s' for collinearity method	76
5.23	B-Tagging Purity vs Efficiency	77
5.24	B-Tagging performance	78

DEFINITIONS OF ACRONYMS

HL-LHC High Luminosity Large Hadron Collider

LHC Large Hadron Collider

Superconducting collider occupying the 27 km ring at CERN.

QCD Quantum Chromodynamics

SM Standard Model

BSM Beyond the Standard Model

HL-LHC High Luminosity Large Hadron Collider

FCC Future Circular Collider

CLIC Compact Linear Collider

ILC International Linear Collider

ILD International Large Detector

SiD Silicon Detector

CMOS Complimentary Metal-Oxide Semiconductor

MAPS Monolithic Active Pixel Sensors

EM Electromagnetic

ECAL Electromagnetic Calorimeter

HCAL Hadronic Calorimeter

DECAL Digital Electromagnetic Calorimeter

CMS Compact Muon Solenoid

ATLAS A Toroidal LHC Apparatus

CERN European Organisation for Nuclear Research

TDR Technical Design Report

CDR Conceptual Design Report

IP Interaction Point

BDS Beam Delivery System

BS Beamstrahlung

TPC Time Projection Chamber

SIT Silicon Internal Tracker

SET Silicon External Tracker

FTD Forward Tracking Detector

ETD Endcap Tracking Detector

RF Radio Frequency

RF Radio Frequency

PFA Particle Flow Algorithm

PFO Particle Flow Object

CHAPTER 1

Introduction

With the expected shutdown of the High Luminosity Large Hadron Collider (HL-LHC) in 2038, and the long time scales associated with the construction of any new colliding facility (~ 10 years), the time for physicists to agree on what experiments should follow in the post Large Hadron Collider (LHC) era is rapidly approaching, with initial decisions expected to take place in the early 2020s. However, following the discovery of a Higgs Boson at the LHC [4, 5], with properties in agreement with those predicted by the SM, the particle physics community is left in a situation where there is no definitive course of action through which new physics phenomena might be discovered. There are still many open questions remaining; one clear example being the origin of dark matter, which has been observed to make up $\sim 27\%$ of the universe. Despite being examined through multiple astrophysical observations such as gravitational lensing or galaxy rotation curves [6], there is still no proven particle physics explanation for what it is made from or how it is produced. Other notable examples include the matter-antimatter asymmetry of the universe which is yet to

be explained by the levels of CP violation measured in the SM and the Higgs hierarchy problem / fine tuning problem where it is expected that a precise cancellation of quantum corrections is needed to be able to simultaneously explain the difference in strength between the weak and gravitational forces while allowing for the measured value of the Higgs mass. Currently there is no clear direction for how we might solve these mysteries; as such there are two main approaches that may be taken. The first is to continue to push the boundaries of the “energy frontier” and look for new physics phenomena at higher energy scales that are not predicted by the SM. In this scenario the natural option is to build a circular hadron collider, much like the LHC. While hadron collisions result in more complex interactions due to the fact they possess substructure, they are well suited for high energy collisions due to their high masses which reduce the amount of synchotron radiation emitted (radiation produced from accelerating a charged body through an Electromagnetic (EM) field (see (1.1)) when accelerating them in a circular path.

$$P = \frac{e^4}{6\pi\varepsilon_0 m^4 c^5} E^2 B^2. \quad (1.1)$$

Where P is power, e is elementary charge, E is particle energy, B is magnetic field, m is mass and all other symbols have their usual meaning.

Pushing the energy frontier has the appeal that it allows direct detection of particles at new energy scales and is supported by the fact that many Beyond the Standard Model (BSM) models rely on new particles appearing in the >multiTeV energy range e.g. supersymmetry, however it does have drawbacks and risks. Due to the composite structure of hadrons they provide collision energies that are often significantly below the provided beam energy and that are challenging to measure. This limits the type of measurement that can be performed as the initial state of the interaction is poorly defined and so all measurements must rely on measurement of the final state particles. This increases the effect of uncertainties introduced by detector acceptances and resolutions and makes it highly challenging to identify particles that can't be measured by the detector e.g. neutrino/ dark matter candidates.

Due to fragmentation of the hadrons, there are also significant Quantum Chromodynamics (QCD) background jets produced which can dominate over potential new signal channels. While these do make measurements more challenging, the real risk with pushing the energy frontier is that the constraints on the scale of at which new physics might be observed are currently very poor. This makes it challenging to choose what collision energy any future collider should operate at as choosing too low an energy could result in no new phenomena being seen.

The second option is to advance in the “precision frontier” and search for small deviations from the SM or harder to detect processes. In this case the more natural choice is to use a lepton collider as the fact the interaction is an annihilation of fundamental particles rather than composite particles means that the initial conditions of the interaction can be known to a high precision determined entirely by the quality of the colliding beams. For leptons it is also possible to produce polarised beams which opens up a new range of potential measurements when examining interactions that couple differently to left and right handed particles. Doing this, areas of the SM that are less well measured such as the Higgs and Top quark sectors can be probed for evidence of physics beyond the SM. The worst case scenario for a lepton collider is to simply reinforce the SM without seeing any new phenomena, however even in this case the new levels of precision on many of the SM parameters will be beneficial for constraining BSM theories and reducing systematic uncertainties on measurements being made at other future colliders. While allowing for precision measurements of the SM, lepton colliders do also provide opportunities for both direct and indirect discoveries of new physics through channels that are either unavailable at hadron colliders or that are challenging due to the QCD backgrounds. The main draw back of colliding leptons is that currently the only viable option is to use electrons and positrons (though there is effort underway to use muons [7]) which have extremely low masses and so produce considerable levels of synchotron radiation (10^{16} as much as protons) when used in a circular collider. The usual solution to this is to use a linear collider instead. This prevents losses from synchotron radiation, however it limits the maximum collision energy that can be achieved as the path over which

the particles can be accelerated is limited to the length of the accelerator which is itself limited by the increasing cost of extending the footprint of the machine. It is worth noting however, that for leptons the collision uses the full beam energy each time and so higher energy interactions can be produced from lower energy beams than for hadrons.

Considerable work has already been carried out into designing both high energy and high precision colliders. On the high energy side is the Future Circular Collider (FCC), a 100 TeV circular proton collider proposed as a project for European Organisation for Nuclear Research (CERN). It is possible to also use this device as an e^+e^- collider operating above the Higgs threshold so as to act as a “Higgs factory”. On the precision side there are multiple proposed projects, however the most mature of these are the linear electron-positron colliders: Compact Linear Collider (CLIC) [1] and International Linear Collider (ILC) [8]. The ILC is a 500 GeV collider proposed as a joint endeavour between the Japanese government and the international community while CLIC is a multi-TeV machine being proposed by CERN. Due to the large cost of these devices it unlikely that CERN would build both FCC and CLIC together.

The focus of this thesis will be on the prospects of the proposed high precision colliders. In particular we discuss the prospects for measuring properties of the Higgs Boson and Top quark at CLIC, which are both relatively less precisely measured when compared to the other particles of the SM, while also examining a novel design for a digital calorimeter based on Complimentary Metal-Oxide Semiconductor (CMOS) Monolithic Active Pixel Sensors (MAPS) technology for use in future detectors as an extremely high granularity Electromagnetic Calorimeter (ECAL).

CHAPTER 2

Experiments

There are many possible designs for future lepton colliders [9, 10] however here we focus on the two most developed projects, CLIC and ILC. Both projects are linear colliders which propose using electron-positron collisions and were founded over twenty years ago, though ILC is currently the more mature design of the two. We will also discuss the detectors proposed for both experiments. ILC currently has two detector concepts being developed, International Large Detector (ILD) and Silicon Detector (SiD), which will be operated in a 'push-pull' scheme in which both detectors are placed on a single platform that is periodically moved to alternate which detector is placed in the path of the beams. This is necessary as there is only one interaction point at a linear collider. Having two detectors has the advantage that any results made with one detector can be verified with the second to help reduce any systematic bias from either machine, however it comes with the penalty that each detector will only be able to take data half of the time and the process of moving the detectors in and out is lengthy ($\sim 3\text{-}4$ days) resulting in reduced time

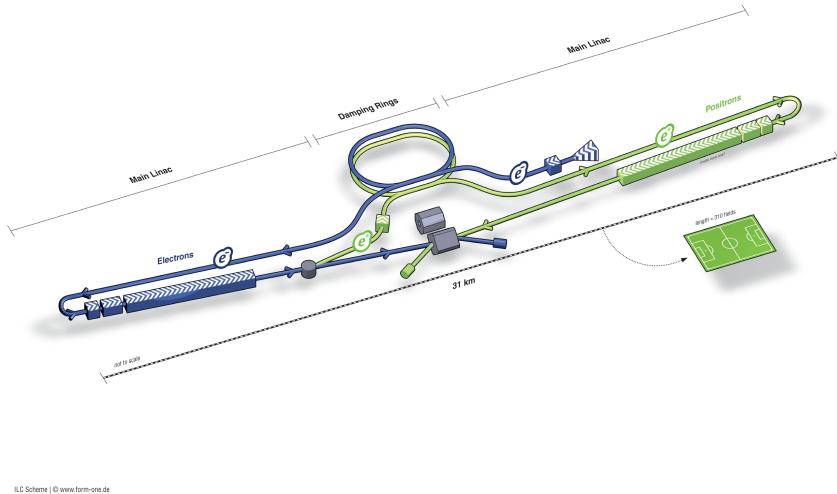


Figure 2.1: The ILC Collider (from [11])

for data taking for the experiment. CLIC intends to operate with only one detector, a variation of the ILD developed for ILC, adapted for the different beam conditions present at CLIC.

2.1 ILC

The ILC (Figure 2.1) is a proposed experiment consisting of a 31km e^+e^- collider to be built in Kitakami in the northern region of Japan. The current construction schedule predicts the experiment will be finished in the mid 2020s with a cost of the order £6 billion and will run for approximately 20 years. However, until funding is secured for the experiment this is just an estimate. The ILC Technical Design Report (TDR) [11] was released in 2013 and gives a full description of the experiments' baseline design. While the TDR is highly detailed, because the experiment is still under development it is possible that some of the information contained within it will become outdated and change before construction takes place. For simplicity any figures given in this section can be assumed to be taken from the TDR unless otherwise stated.

2.1.1 Energy Staging

The ILC will first be built with a maximum collision energy capability of 500GeV but with the potential for a later upgrade to 1TeV which would require doubling the length of the machine to 62km. The decision of whether the 1TeV upgrade is necessary will largely be determined by the results of the LHC experiments; if any new physics is discovered above 500GeV then the 1TeV upgrade could be essential to characterise them. Assuming the 1TeV upgrade is realised the energy staging will be as described below.

The first three years will involve the ILC running at an energy of 250GeV and taking 250fb^{-1} of data. The main aim at this stage will be to measure the Higgs mass and ZH cross section from the Higgsstrahlung process as described in chapter 3 to allow model independent measurements of the Higgs couplings to be performed.

For the following three years, the collider will then run at 500GeV and will accrue a further 500fb^{-1} of data. The main aims here will be to measure the H-WW coupling, the total Higgs width and the absolute Higgs couplings to fermions. At this energy, measurements of top physics will also be possible including the top forward backward asymmetry. Outside of the Higgs sector, the top quark is perhaps the least well measured of the standard model particles and so provides another area in which to look for deviations from the standard model predictions.

After this there will be an upgrade to 1TeV followed by another three years of data taking accumulating 1000fb^{-1} of data. The aim of running at this high energy will be to search for new particles such as dark matter candidates and supersymmetric particles while improving upon the precision of the measurements performed at the lower energies. If one of these (or something entirely new) has already been discovered at the LHC then the choice of 1TeV might be scaled down to somewhere between 500GeV and 1TeV to match the scale of the newly discovered physics.

After this the collider will undergo a high luminosity upgrade and will run at the same energies for the same time periods for another 9 years but instead accruing 900,

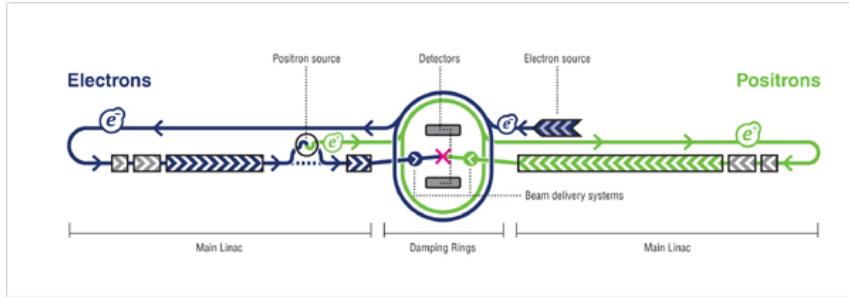


Figure 2.2: A simplified schematic of the ILC

1100 and 1500fb^{-1} at the respective energies. This will allow for a further increase in the precision of all measurements taken during the lower luminosity run. While the TDR proposes the above run scheme for the ILC there is still debate about what energies should be used with arguments being made for running at 90GeV (the Z mass) to gain precision measurements of the Z boson, 350GeV (the top production threshold) to better measure the properties of the top quark or to simply only run at 250GeV to provide precision Higgs measurements for minimal cost.

2.1.2 Beam Production, Acceleration and Focusing

A simplified schematic of the ILC accelerating structure is shown above in diagram Figure 2.2 while a summary of the key beam parameters is shown in Table 2.1(ADD TABLE FROM TDR). The first stage of the acceleration process is the production of electrons. This is done using the photoelectric effect by firing photons onto a GaAs target to produce photoelectrons. These electrons then enter a 3.2km long damping ring which accelerates the beam up to 15GeV. The primary purpose of the damping ring is to produce a homogeneous beam of electrons with uniform energy and momentum. After the damping ring the electrons enter into a two stage bunch compressor which separates the electron beam into ~ 1300 bunches, each containing 2×10^{10} electrons, with each bunch being separated by 554ns and a maximum beam pulse length of $\sim 1.6\text{ms}$. The overall intended collision rate of these pulses is 5Hz, which means that the duration for collisions is less than 1% of the collision rate. This has important consequences for the detector design as it means the detectors



Figure 2.3: A 1.3GHz Superconducting Niobium Radio Frequency Cavity [11]

have a large period of time in which to relax after events. As the detectors do not need to be operating for 99% of the time, it is considerably easier to cool them meaning the material budget for the cooling systems within them can be greatly reduced. Following the bunch compression the electrons enter the main 11km linac where they are accelerated up to the nominal beam energy using 7,400 1.3GHz superconducting niobium Radio Frequency (RF) cavities (see Figure 2.3)

The RF cavities are kept at a temperature of 2K and act to produce an average accelerating gradient of up to 31.5MV/m (14.7MV/m for the 250GeV stage.) The final stage before the collision is the Beam Delivery System (BDS) which primarily acts to compress the beam into a ribbon shape with a cross-section of 7.7 x 729.0 nm while also handling the beam monitoring. The ribbon shape is primarily chosen to reduce Beamstrahlung (BS) radiation from beam interactions while giving a small enough cross-section that a high instantaneous luminosity can be achieved. Following the BDS the beam finally enters the detector and collides with the opposing positron beam at a crossing angle of 14mrad then exits into the beam dump system which quenches what is left of the beam.

2.1.3 Positron Production

Positrons are produced at the ILC by tapping off energy from the electron beam after it has been accelerated by the main linac. The electron beam is passed through an undulator which causes the electrons to emit synchrotron radiation in the form of 10-30MeV photons by forcing the beam to take a rapidly varying path in the transverse plane to it's direction of motion. The resulting photons are then separated from the electron beam and are collided with a Titanium alloy target to produce electron

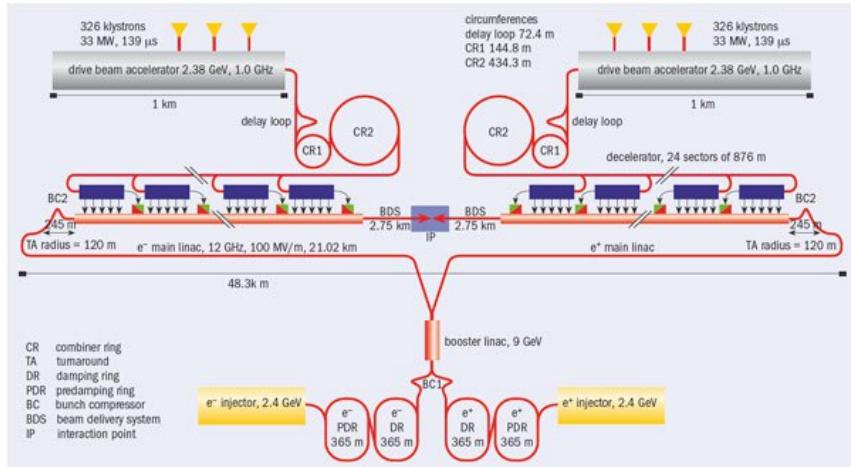


Figure 2.4: The CLIC Collider. Image taken from CLIC Conceptual Design Report[12]

positron pairs. The electrons and positrons are then separated and the electrons are dumped while the positrons are then passed into a damping ring and undergo all the same stages of acceleration and shaping as described above for the electrons.

2.2 CLIC

CLIC is an experiment based at CERN which proposes the building of a 42km accelerator at the main CERN site in Geneva (Figure 2.4.) Despite being named as compact, CLIC is actually longer than the initial 500GeV ILC. The reason for this naming is that CLIC has a much higher accelerating gradient (100MeV/m) compared to ILC and so provides a much higher energy per length. The expected build date for CLIC is still relatively uncertain though is likely to be no earlier than 2030 as the accelerating technology required for CLIC is less developed than that used by ILC. This difference in the maturity of the two experiments can be seen from the fact that the ILC has released its TDR while the most comprehensive document for the CLIC project is still its Conceptual Design Report (CDR) [12]. Updates on this document have been provided in the New Baseline Report [1] released in 2016 and any details specified here can be assumed to come from these two documents.

Table 2.1: Parameters for the CLIC energy stages. The power consumptions for the 1.5 and 3 TeV stages are from the CDR; depending on the details of the upgrade they can change at the percent level [1].

Parameter	Symbol	Unit	Stage 1	Stage 2	Stage 3
Centre-of-mass energy	\sqrt{s}	GeV	380	1500	3000
Repetition frequency	f_{rep}	Hz	50	50	50
Number of bunches per train	n_b		352	312	312
Bunch separation	Δt	ns	0.5	0.5	0.5
Pulse length	τ_{RF}	ns	244	244	244
Accelerating gradient	G	MV/m	72	72/100	72/100
Total luminosity	\mathcal{L}	$10^{34} \text{ cm}^{-2}\text{s}^{-1}$	1.5	3.7	5.9
Luminosity above 99% of \sqrt{s}	$\mathcal{L}_{0.01}$	$10^{34} \text{ cm}^{-2}\text{s}^{-1}$	0.9	1.4	2
Main tunnel length		km	11.4	29.0	50.1
Number of particles per bunch	N	10^9	5.2	3.7	3.7
Bunch length	σ_z	μm	70	44	44
IP beam size	σ_x/σ_y	nm	149/2.9	$\sim 60/1.5$	$\sim 40/1$
Normalised emittance (end of linac)	ϵ_x/ϵ_y	nm	920/20	660/20	660/20
Normalised emittance (at IP)	ϵ_x/ϵ_y	nm	950/30	—	—
Estimated power consumption	P_{wall}	MW	252	364	589

Overall the design for CLIC is relatively similar in layout to the ILC but with a few changes. Positron production at CLIC is done completely independently from the main electron beam, though they are still produced via the same mechanism as before. The BDS still compresses the beam into a ribbon shape to give it a small cross-section and reduced BS, however the aspect ratio is reduced compared to at ILC. This results in larger contributions from beam photon radiations at CLIC. The collision rate at CLIC is significantly higher as it aims to be a high luminosity device—the collision rate will be 50Hz with 354 bunches per pulse with a separation of just 0.5ns. This means that CLIC will have a significantly higher duty cycle which will make cooling of the detectors harder and will give the detectors less time to relax after events. A summary of the beam parameters for CLIC is shown in Table 2.1, however the most significant differences are in the energy staging and acceleration technology used at CLIC.

2.2.1 Energy Staging

CLIC will operate at three energy stages- 380GeV, 1.5TeV and 3TeV collecting $500fb^{-1}$, $1.5ab^{-1}$ and $2ab^{-1}$ of data respectively. During the running of the 380 GeV energy stage, construction of the 1.5 TeV structure will be carried out (and so on for the 1.4TeV and 3TeV scales) so as to reduce the delay between operation at successive energy stages.

The 380GeV energy scale is chosen as it is above the $t\bar{t}$ production threshold and provides a significant cross section for many channels involving the top quark. This stage is also supplemented by a series of 10 measurements around the $t\bar{t}$ threshold taking $10fb^{-1}$ each with the aim of measuring the top mass and width from the line shape of the $t\bar{t}$ production cross section at threshold. The 380 GeV stage will also be used to provide measurements of the higgs boson similar to those performed at ILC during it's two lower energy stages.

The 1.5TeV energy stage provides the ability to further study the top and higgs in more detail with several new channels becoming significant e.g top yukawa coupling, higgs self coupling while the 3 TeV stage pushes the energy frontier allowing the possibility of direct detection of new physics at the multi-TeV scale. The choice of 3 TeV is based upon certain models of supersymmetry which predict new particles to exist at this energy.

For clarification it should be stated that for many years the proposed scheme for CLIC was actually to operate at 500 GeV, 1.4 TeV and 3 TeV. These were updated to provide better precision on measurements of the top quark during the lowest energy stage ($t\bar{t}$ production threshold is ~ 350 GeV) and improved precision on the Higgs self-coupling during the second stage. It is important to be aware of these changes as the studies presented in chapters 4 and 5 were both carried out at 1.4TeV assuming the original energy staging, however it is expected that there will be negligible impact on the findings of these studies from changing the energy to 1.5TeV.

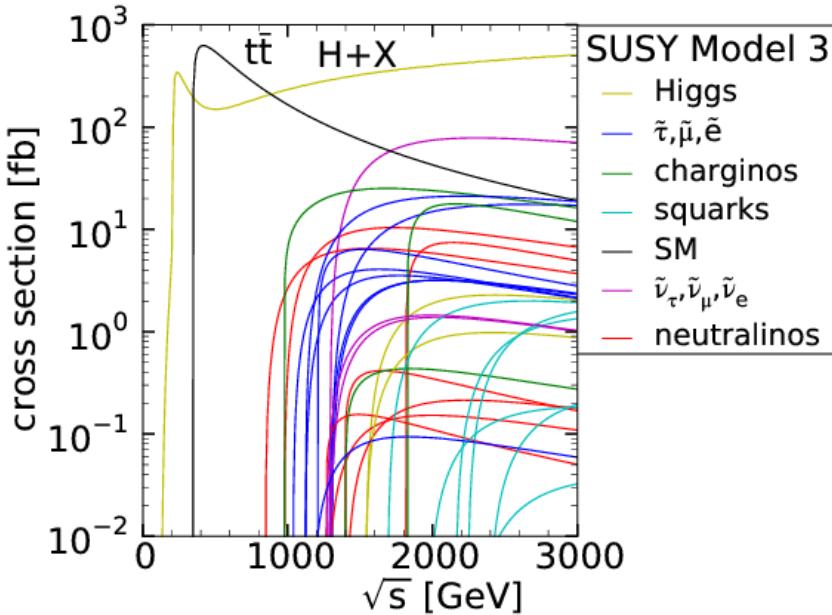


Figure 2.5: Cross sections for production of various super symmetric particles at an e^+e^- collider as a function of centre of mass energy.

2.2.2 Acceleration Technology

Unlike ILC, the acceleration technology will not be superconducting and will use two beams of electrons—referred to as the main beam and the drive beam—rather than just one main accelerated beam. The drive beam is accelerated using standard accelerating technology (klystrons) as in ILC to accelerate bunches of electrons to 2.75GeV. These bunches then enter a series of delay/control rings which are designed such that the electrons within them get combined with the new electrons being added from the drive beam accelerator to build up a large number of low energy electrons which combined carry a large amount of energy. The energy from this beam is then used to drive the main beam. This is done by rapidly decelerating the drive beam electrons down to 10% of their initial energy and using the resulting RF produced to accelerate the smaller number of electrons in the main beam resulting in a rapid acceleration. The main beam is then used to supply collisions. Overall the result is that the machine is simply acting as a novel form of transformer, converting a high current, low energy beam of electrons into a low current, high energy beam. This approach allows for very high accelerating gradients but has the disadvantage

that in approximately 1% of events the sudden input of energy from the drive beam can cause electrical breakdowns in the main drive beam cavity, which disrupt the alignment and structure of the main beam making them unsuitable for use.

2.3 Linear Collider Analysis Framework

A common framework known as ILCSoft which is used for event simulation, reconstruction and analysis has been developed for both ILC and CLIC and allows for sharing of techniques between the two experiments. Here we will provide an overview of the key packages used.

2.3.1 Event Generation

Event generation is performed using an external package called WHIZARD [13]. WHIZARD itself handles most of the event generation such as the calculation of hard matrix elements, phase space considerations, accounting for interference between processes, however for certain aspects it relies on additional packages. The most relevant of these are tau decays which are handled by TAUOLA[14] and hadronization which is handled by PYTHIA[15]. Unfortunately no other hadronization package is available within WHIZARD which makes it challenging for evaluating systematic uncertainties arrising from how jets are modelled. The output from WHIZARD is a series of four momenta for all the particles produced in a collision. These are then passed to a package called MOKKA whcih acts as an interface to GEANT4[16]. Within MOKKA the interaction of the particles with the detector is modelled and a series of energy deposits are recorded for the various subdetector components. These are then finally passed on to the ILCSoft reconstruction package MARLIN in which digitisation of the hits and track reconstruction occur to produce realistic outputs from the detector.

2.3.2 Pandora Particle Flow Algorithm

Pandora[17] is an advanced Particle Flow Algorithm used at linear colliders which allows an increased level of precision from detector measurements. The underlying principle behind particle flow algorithms is to try and use the most precise detector component for performing the energy measurement where possible. Typical values for energy resolutions for a charged particle in the main detector components are $10^{-4} \times E^2$ in the tracker, $0.15 \times \sqrt{E}$ in the ECAL and $0.55 \times \sqrt{E}$ in the Hadronic Calorimeter (HCAL). For a typical jet the composition will usually be $\sim 60\%$ charged hadrons, 30% photons and 10% hadrons. Traditionally for measuring the energy in a jet one would simply sum the deposits in both calorimeters resulting in a relatively poor energy resolution of $\sim 60\%/\sqrt{E}$ due to the large component being measured in the HCAL. If one can measure the charge hadron component in the tracker instead, this performance can be vastly improved to $\sim 20\%/\sqrt{E}$. In order to be able to reach this performance, accurate association of tracks with deposits in the calorimeters is crucial. This is achieved by having a high granularity calorimeter and a high spatial resolution for the tracker. In practice, even with a well designed detector, the particle flow algorithm can still fail to reconstruct the correct energy due to ambiguities. For example, if a photon enters the calorimeter near to a charged hadron, it is possible that the two will not be resolved and the energy identified from just using the track will neglect the contribution from the photon. Energy can also be overestimated in cases where a charged hadron showers in such a way that it looks like two separate calorimeter deposits which results in part of the shower being identified as a neutral hadron and the other fragment being associated with the track. Effects such as these are commonly referred to as “confusion”. One of the main design aims of the detectors will be to try and minimise these confusion effects.

2.4 Detectors

The ILC has been designed with the intention of having two unique detectors so that results can be validated by cross-checking between the two detectors. However, because ILC is a linear collider it is only feasible to have one interaction point and as a result the beam time will have to be shared between the detectors. This will be done using a 'push-pull' design in which both detectors are placed on a single platform at the interaction point which can be moved back and forth to position the desired detector in the path of the beams. While having two detectors is certainly desirable as it allows the gathering of two independent sets of results for the collider and allows the continued taking of results when one of the detectors requires maintenance, it also has disadvantages as it means an increase in the dead time of the machine (as swapping the detectors is a slow process taking several days which will be done multiple times a year) and an increase in the cost of the experiment. As a result the possibility of using only one detector is still being considered as a potential option. The possibility of splitting the main beam and having two IPs is also being proposed so that both detectors could be used simultaneously however this would be expensive as extra tunnels would have to be built to accomodate this and there would also be a reduction in the beam quality as splitting the beam would produce synchotron radiation. The studies presented in this thesis are based on simulations of only one of these detectors, ILD, and as such we will not give details of the alternative: SiD.

2.4.1 ILD

The ILD (shown in Figure 2.6) is a general purpose detector which is cylindrical in design with radius 8m and length 14m. The different sub-detectors are arranged in a concentric manner in the main barrel of the detector, and are positioned with the vertexing technology closest to the beamline, followed by trackers, then electromagnetic and hadronic calorimeters, then the magnetic field coils which supply a 3.5T B-field and finally muon tail catchers. The detector has two endcaps at each end of

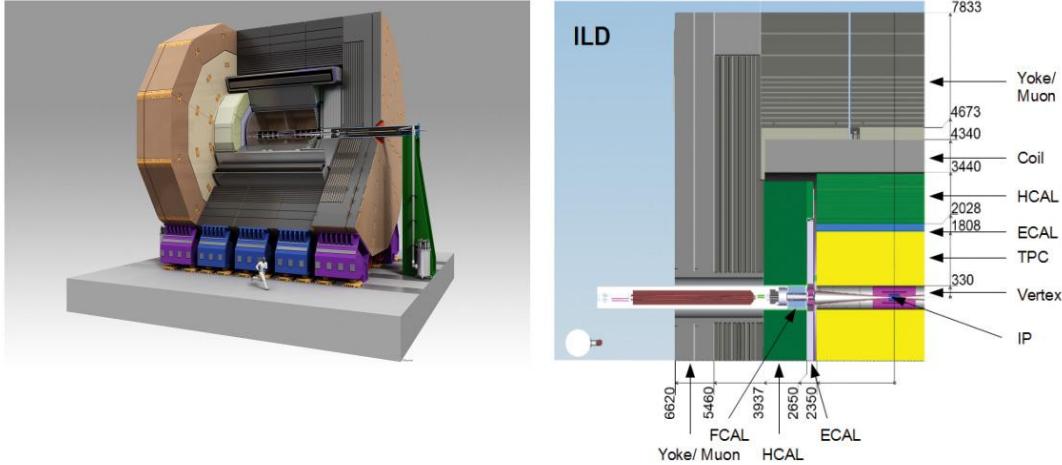


Figure 2.6: The International Large Detector Concept (left). Schematic of the ILD showing the key components in a one-quarter view of a vertical section of the detector (right). [8]

the barrel creating a hermetic seal.

In order to provide precision measurements of the various processes proposed in the ILC and CLIC physics schemes, there are several strict requirements imposed upon the performance of the detector:

- **Momentum Resolution:** $\sigma_{p_t}^2/p_t^2 \sim 2 \times 10^{-5} GeV^{-1}$, key for precision Higgs recoil mass measurements
- **Jet Energy Resolution:** $\sigma_E/E \sim 3-4\%$, allows separation of hadronic W/Z decays
- **Impact Parameter Resolution:** $\sigma_b < 5 \oplus \frac{10}{p \sin \theta^{\frac{3}{2}}} \mu m$, allows accurate flavour tagging for short lived particles
- **Hermetic Coverage:** Needed for processes with a strong angular dependence or missing energy component

Detailed specifications for the detector can be found in the ILD Letter of Intent [2]. Here we will give a brief overview of the key components, their functions, and the methods used for making the most of the information they provide.

2.4.1.1 Vertexing

The vertexing technology is used to gain information about heavier particles such as b-quarks which have very short lifetimes ($\sim 10^{-12}$ s) and so decay close to the beamline before they can reach the trackers or calorimeters. As such, the vertexers are placed extremely close to the beamline and work by looking for displaced vertices from the initial Interaction Point (IP) which correspond to the point at which the heavy flavour paricles decayed. Due to their proximity to the beam line it is always necessary for the vertex detectors to be radiation hard as they are exposed to stray high energy particles from the beam. The vertexers also act as trackers for short lived particles that fail to reach the main trackers and so are required to be highly granular to separate particles that have had very little time to spread out since the IP. The design for the vertex detectors is yet to be finalised as there are numerous competing technologies under consideration, but the target performance is to achieve a track impact parameter resolution of

$$\sigma_b < 5 \oplus \frac{10}{p \sin \theta^{\frac{3}{2}}} \mu m \quad (2.1)$$

where p is the track momentum in GeV, θ is the angle between the track and the vertex detector plane and the first and second terms describe contributions from the transverse impact parameter resolution and multiple scattering effects respectively. In practice it is found that to achieve this impact parameter resolution a spatial resolution of at least $3\mu m$ is required. As well as achieving a sufficient impact parameter resolution, the vertexing detectors are also required to have sufficient granularity and low enough occupancy rates to allow separation of individual tracks passing through the detector. On top of these requirements for the vertexing and tracking performance, the design for the vertexer must also avoid any negative impact on later parts of the detector. In particular, the material budget of the whole detector system is limited to be less than one radiation length to avoid unwanted production of electromagnetic showers prior to the ECAL. The detector layout used

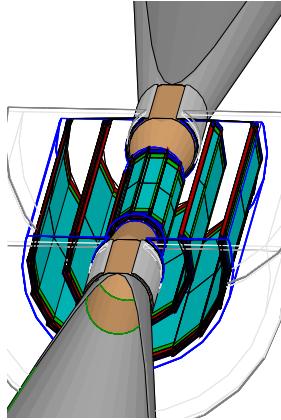


Figure 2.7: Proposed vertex detector geometry for ILD [2]

Table 2.2: Properties of the CLIC vertex detector assuming three pairs of layers [2]

layer	radius [mm]	ladder length [mm]	read-out time [μs]
1	16.0	125.0	25-50
2	18.0	125.0	25-50
3	37.0	250.0	100-200
4	39.0	250.0	100-200
5	58.0	250.0	100-200
6	60.0	250.0	100-200

for the baseline studies in the ILD TDR assumes six layers of $50\mu m$ silicon pixels arranged in pairs. The layout and details of the structure are shown in more detail in Figure 2.7 and Table 2.2.

2.4.1.2 Tracking

Tracking in ILD is performed by multiple subsystems. We have already discussed the vertexing systems which act as trackers for low transverse momentum and short lived particles, however the majority of the tracking is performed by a large Time Projection Chamber (TPC).

Tracking at the ILD uses a large TPC. This is a large gas filled cylinder extending from $r=395$ mm to $r=1739$ mm with an electric field applied across it and readout electronics at each end of the cylinder on the endcaps. As particles pass through

the gas, they ionize it producing charged particles. The electric field then causes these particles to drift to each end of the detector where they are collected by the electronics. By measuring the position and time at which the charged particles arrive, the track of the original ionizing particle can be reconstructed. A magnetic field is also generated across the chamber to deflect the charged particles so that the momentum and charge of the particle can be estimated. The magnetic field used in the ILD is a 3.5T coil placed outside the calorimeters to minimize the material budget in front of the calorimeters. The use of a TPC provides several benefits over alternative technologies such as silicon tracking (the technology used in the SiD tracker.) Because the ionization occurs across the whole track, it is possible to reconstruct the particles path from numerous spacial points to provide a precise measurement. This is not the case for a silicon tracker where the number of data points is proportional to the number of tracking layers present, however this is compensated for by the fact silicon trackers typically have a higher spatial resolution on each point ($\sim 1\mu m$) compared to TPCs ($\sim 1mm$.) TPCs also benefit from having a low material budget compared to silicon trackers. In ILD the gas used will be Ar:CH₄:CO₂ (95:3:2) which gives a material budget of $\sim 0.04(0.15)X_0$ radially(longitudinally.) The choice of readout technology is yet to be finalised with several options being pursued (Micro-Pattern Gas Detectors, MicroMegas and GEM(REFERENCES FROM LoI PAGE 70!)) however in all cases it is expected that there will be 10^{-6} channels of dimension $\sim 1 \times 6mm^2$. This system will allow a single point resolution of $<100\mu m$ (0.5 mm) and two hit resolution of 2 mm (6 mm) in the x-y (r-z) planes, and a resolution of 5% on dE/dx.

The TPC is supplemented by a series of silicon based tracking systems which act to provide high spatial resolution points at the entrance and exit of the TPC which yields an improved momentum resolution and an improved performance from the Particle Flow Algorithm (PFA)s, provide time stamping for bunch tagging and assist in calibration of the TPC. These additional subdetector systems come in four parts. In the barrel region, between the vertexer and the TPC lies the Silicon Internal Tracker (SIT) which provides two high spatial resolution points at $r= 165$ mm and

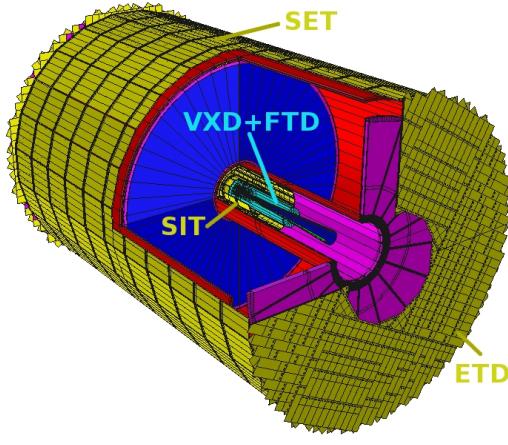


Figure 2.8: Silicon Tracking Systems For ILD [2]

$r=309$ mm, while between the TPC and the ECAL lies the Silicon External Tracker (SET) which provides a single spatial point at $r=1844$ mm. Both of these systems are based on double sided silicon microstrips and provide a resolution of $\sim 50\mu\text{m}$. The Forward Tracking Detector (FTD) covers the very forward region of the detector down to 0.15 radians and consists of 7 disks positioned in the innermost tracking region, the first three using silicon pixels and the end 4 using silicon microstrips. The Endcap Tracking Detector (ETD) is similar in structure to the SET but is positioned outside the TPC endcaps rather than the barrel to provide high spatial resolution for particles exiting the tracker into the endcap calorimeters. The poistioning of all these subdetector systems can be seen in Figure 2.8.

The combined performance of the vertexer, TPC and silicon tracking systems is a momentum resolution of $\sigma_{p_t}^2/p_t^2 < 2 \times 10^{-5} \text{GeV}^{-1}$ and a tracking coverage reaching to as low as $\cos \theta < 0.996$.

2.4.1.3 Calorimetry

The function of calorimeters is to measure the energy of particles by passing them through a medium in which they will deposit some of their energy. As the way that particles interact with other matter is determined by the type of particle involved, the calorimeters are usually split into two sections, the ECAL and HCAL, that are

designed to interact with electromagnetic particles (electron, photons) and hadrons respectively. As we will later be presenting work on a proposed novel design for a Digital Electromagnetic Calorimeter (DECAL) it is pertinent to discuss in greater detail the relevant processes and terminology involved in electromagentic calorimetry to understand what issues there are with current ECAL technologies and how the DECAL might improve upon them.

When an electron interacts with matter it will typically radiate a photon via BS. This photon then can then decay into an electron-positron pair which will in turn radiate further photons. This cascade process results in the formation of what is referred to as an electromagnetic shower. The shower will continue to develop until the energy of the shower particles reaches a critical value, E_C , at which the energy of the particles becomes similar to the ionization energy of the material it is traversing and ionization begins to dominate over beamsstrahlung. The development of the electromagnetic shower can be characterised using several parameters. The most commonly used of these is the the radiation length, χ_0 , which is defined as the distance an electron can travel through a material before it's energy has reduced by a factor of $1/e$ via beamsstrahlung. The interaction can be expressed as a function of a materials nuclear parameters (REFERENCE):

$$\frac{1}{\chi_0} = 4\alpha r_e^2 \frac{N_A}{A} Z^2 \ln(183/Z^{-\frac{1}{3}}) \quad (2.2)$$

For the purposes of designing a detector, perhaps the most relevant parameters are those related to the size of the showers as these determine the dimensions needed from the calorimeter. The longitudinal detector requirements are decided by the rate of energy loss for a particle which is given by the Bethe-Bloch equation which can be simplified to: (REFERENCE)

$$\frac{dE}{dx} = E_0 b \frac{(bx)^{a-1} e^(-bx)}{\Gamma(a)} \quad (2.3)$$

Where x is the material depth in units of χ_0 , E_0 is the initial energy of the particle, and a and b are properties of the absorbing material. The exponential term means that it is typically not possible to capture 100% of the energy in a shower, instead an acceptable amount of loss must be decided and the detector designed accordingly. For example, a typical energy scale for CLIC would be ~ 100 GeV. For a working point of 5% loss a calorimeter depth of $\sim 17 \chi_0$ is required, while for an improved performance of just 1% loss a depth of $\sim 20 \chi_0$. The transverse profile of the shower is described by the Moliere radius, the radius in which 90% of a particles energy will be deposited:

$$R_M = \frac{21MeV}{E_c} \chi_0 \quad (2.4)$$

In general, it is necessary to have a Moliere radius that is smaller than the typical separation of particles produced in collisions so as to avoid overlapping of showers. For ILD this is especially true where Pandora PFA relies on accurate association of tracks to calorimeter deposits which is only possible if the deposits from nearby particles can be distinguished.

For ILD the ECAL and HCAL are both sampling calorimeters. This means that the structure is divided into layers of two alternating materials known as the absorber and active material. The absorber is typically a thick piece of high Z material that acts to initiate an EM shower. The active material is then a thin low Z material that is easily ionizable and so acts to collect charge deposited from the shower. The active layer will then be instrumented to collect and readout the charge deposited within it. In order to reconstruct the energy of the initial particle that produced the shower, one would ideally just sum the deposits from each of the active layers, however in reality there will also be some energy deposited within the absorbing material that must be accounted for. This is done by scaling the energy deposited by the expected ratio of the energy deposited in the active layer to the total energy deposited in the active and absorbing layers combined. The scale factors will usually be determined as part of a calibration procedure for the detector in which muons

are passed through each layer. The application of these scale factors introduces an uncertainty in the reconstructed energy as they represent an average scale correction, whereas the actual ratio of the energy deposited in the active and absorbing layers will be determined by additional factors that can't be easily measured. One example would be the path taken by the particle which can change the relative distance travelled by the particle in the active and absorbing layers.

The overall performance of a calorimeter is given by the energy resolution. This represents the quadrature sum of all sources of uncertainty in the energy reconstruction and it is usually expressed as follows:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \quad (2.5)$$

Where a, b and c are typically referred to the stochastic, noise and leakage terms. The energy dependence of the background and leakage terms are straightforward to understand. Noise typically arises from the electronics used for collecting and reading out the hits in the active layers. This means that it is independent of the energy of the incident particles energy and so the absolute uncertainty doesn't vary with E. The leakage term accounts for the amount of energy lost from the calorimeter not being deep enough to contain the shower. One can see from (2.3) that the energy lost scales with the incident particles energy.

The stochastic term is slightly more complicated as it represents a combination of effects. The first of these is the intrinsic resolution of the detector which is determined by the physics of how an EM shower develops. The number of particles produced in a shower (N) is proportional to the energy of the incident particle (E), however the formation of bremsstrahlung photons and electron-positron pairs is a quantum mechanical process and so is inherently statistical. As a result N will follow a Poisson distribution and so the uncertainty on it will be $1/\sqrt{N}$. As N is proportional to E, this means there is an inherent uncertainty in the energy proportional to $1/\sqrt{E}$. There are also further statistical contributions that arise

from using a sampling approach. For low energy particles produced in the absorber there is a chance that they be absorbed before making it to the active layer and so will not be accounted for in the scale factors. The uncertainty associated with this can be described by $\sqrt{E_c x/E}$. This factor is further added to by the effect mentioned above where x will vary from particle to particle depending on the path it takes through the detector. Because the energy deposited in a material as a function of the material depth is described by a landau distribution, uncertainties from varying path lengths are often referred to as landau fluctuations with the form σ_{landau}/\sqrt{e} .

2.4.1.4 ECAL

The ILD ECAL is a highly granular calorimeter poitioned at $r=1847\text{mm}$ which consists of 30 active layers separated by layers of absorbing material. Tungsten is chosen for the absorber due to it's short radiation length, $\chi_0=0.35\text{cm}$. The first 20 absorber layers are $0.6\chi_0(2.1\text{mm})$ thick while the later layers are $1.2\chi_0$ to contain higher energy EM showers while maintaining a compact design. The structure of the ECAL is shown in Figure 2.9. The active material will consist of $5\text{mm}\times 5\text{mm}$ pitch silicon pixels and yields a resolution of $(16.6 \pm 0.1)/\sqrt{(E(\text{GeV})} \oplus (1.1 \pm 0.1)\%$. While this is currently the default used in simulations for physics analyses, the choice of active material is yet to be finalised. A variation that uses $10\times 45\text{mm}$ silicon scintillator strips which would be rotated by 90° in each successive layer to produce an effective cell size of $10\times 10\text{mm}$ with photomultipliers attached to each strip for readout also exists. The energy resolution for this form of the detector has been measured to be $\sigma_E/E = 14/\sqrt{E} \oplus 2\%$, however the pixel version is typically favoured due to it's simpler design which doesn't require additional processing to produce the desired granularity.

Later on (see Section 6) we will discuss our work on developing an alternative form of the silicon pixel technology with ultra high granularity $50\times 50\mu\text{m}$ pixels which acts as a digital machine and purely counts the number of particles absorbed in the

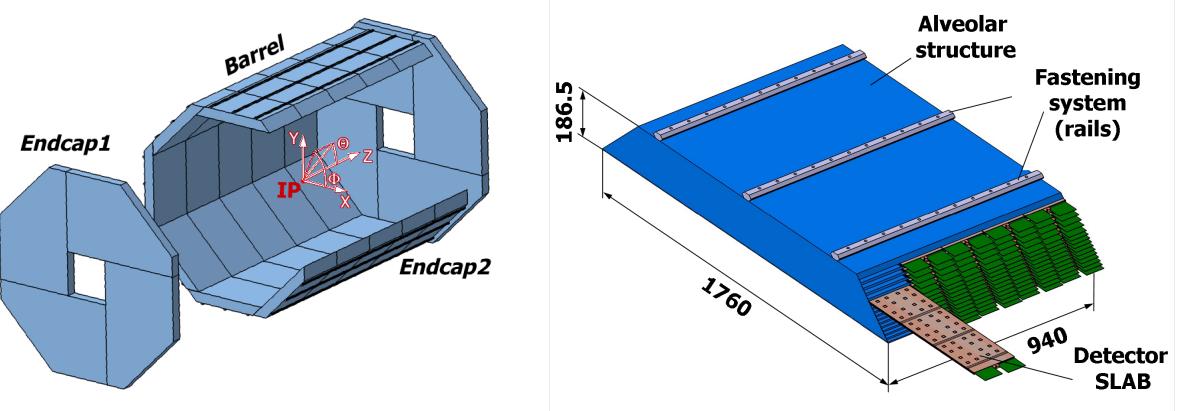


Figure 2.9: The Overall ILD Structure (left) and one individual module (right). The ECAL is made up 40 modules, each containing 30 detector slabs. The modules are combined into groups of 5 referred to as a stave which extend along the full length of the barrel. There are then 8 of these staves arranged in a circle to create the circumference of the barrel [2].

active medium from the showering in the absorber and deduces the energy of the original particle from this. This form of the technology has already begun to be studied [18]. It is expected to be cheaper than the standard silicon pixel technology as it is based on CMOS technology which is already mass produced commercially, and has the potential for improved performance over its analogue counterpart due to reduced sensitivity to Landau fluctuations.

2.4.1.5 HCAL

The HCAL is immediately outside the ECAL at $r=2058\text{mm}$ and has a similar overall modular structure to the ECAL as shown in Figure 2.10. Each module consists of 48 stainless steel absorber plates of thickness 20mm interspersed with 3mm silicon scintillators with a transverse segmentation of 30mmx30mm. This gives the HCAL a total depth of $\sim 5 \lambda_I$ (where λ_I is the nuclear interaction length, the equivalent of χ_0 for hadronic showers, which is typically much longer than χ_0) and an energy resolution of $49\%/\sqrt{E}$.

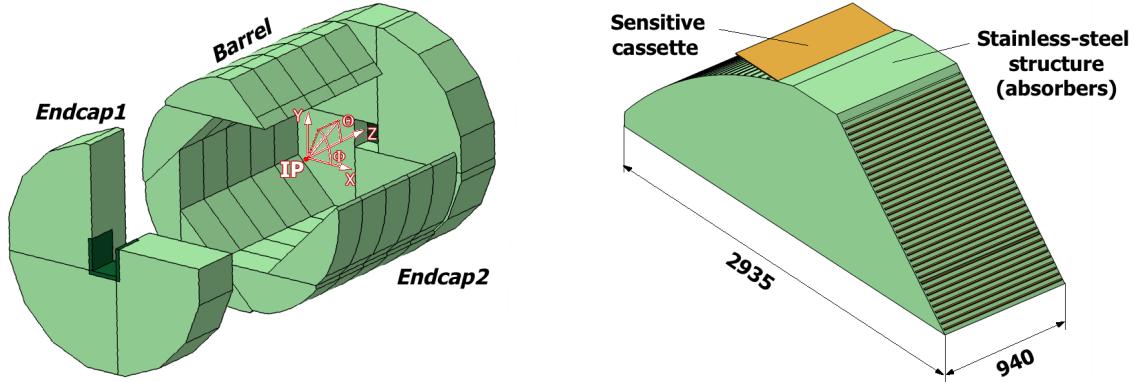


Figure 2.10: The Overall ILD HCAL Structure (left) and one individual module (right). The HCAL is made up 40 modules, each containing 30 detector slabs. The modules are combined into groups of 5 referred to as a stave which extend along the full length of the barrel. There are then 8 of these staves arranged in a circle to create the circumference of the barrel [2].

2.4.1.6 Muon Detection

Muon detection is perhaps the easiest process to perform at the ILC. Because the signal events at the ILC are typically clean with few high energy particles, few particles other than muons are capable of penetrating through the inner detector layers and the coil generating the magnetic field. As a result the muon detectors are produced by instrumenting the return yolk ($r=4424$) that already surrounds the detector to contain the magnetic field. The number of muons produced in an event is also relatively small which means that the cell size for the muon detectors can be moderately large without the risk of multiple occupancy. The instrumentation is done by placing 10 layers of resistive plate chambers into the return yolk with strip sizes of the order 3-4cm. This system is sufficient for accurately detecting muons and contributing to the measurement of their momentum. This system provides $\sim 100\%$ efficiency for identifying muons with momentum $> 3\text{GeV}$. Below this the muons do not have enough penetrating power to traverse through the yolk. This identification performance can be extended down to 1.5 GeV when information from the calorimeters is included as well.

2.4.2 CLIC ILD

At CLIC the detector designs were originally based on the two ILC detectors, ILD and SiD, but with a few changes to adapt for the different experimental conditions at CLIC. In the case of ILD, due to the large beam related backgrounds the vertex detectors were moved to be 15mm further from the IP to avoid pixel saturation. To account for the higher energy jets produced in interactions, the HCAL depth was extended to $7.5 \lambda_I$ to reduce leakage out of the back of the detector. To avoid increasing the radius of the magnetic solenoid (one of the main driving costs of the whole detector) the choice of absorber material in the HCAL was switched to tungsten to provide the increased interaction length over the same depth as in the original steel design. In the barrels, because the thickness does not affect the solenoid radius, the absorber was left as steel. To improve the charge identification of higher energy tracks, the magnetic field strength was changed to be 4T which was found to still be achievable using the original ILD solenoid design. Further details on the CLIC version of ILD can be found in the CLIC CDR[12].

This version of the ILD detector is used for the analyses presented in Chapters 4 and 5. Since these analyses have been conducted, CLIC has recently produced a new unified detector design that will be used for future studies. Overall the design is similar to that of ILD but with a deeper ECAL to allow for higher energy photon containment and the tracker has been changed from a TPC to an all silicon tracker. As this version is not used in the studies presented here we will not give a detailed account of the detector but more information is available in [19]. Overall the impact of the change in detector design is expected to be negligible for the studies presented here.

CHAPTER 3

Theory

Within this thesis two analyses are presented describing prospective measurements looking at the $H \rightarrow WW$ branching ratio and the forward-backward asymmetry in $t\bar{t}$ production at CLIC during the 1.4 TeV stage. As such it is important to first understand the physics behind these measurements and examine their significance in the context of the physics programme of CLIC and the wider state of particle physics.

3.1 The Standard Model

The SM is a quantum field theory representing our best current description of fundamental particles and the interactions between them. It consists of twelve spin- $\frac{1}{2}$ fermions (and their corresponding antiparticles), five spin 1 gauge bosons and one spin 0 scalar boson (as shown in Figure 3.1) where the interactions of the model

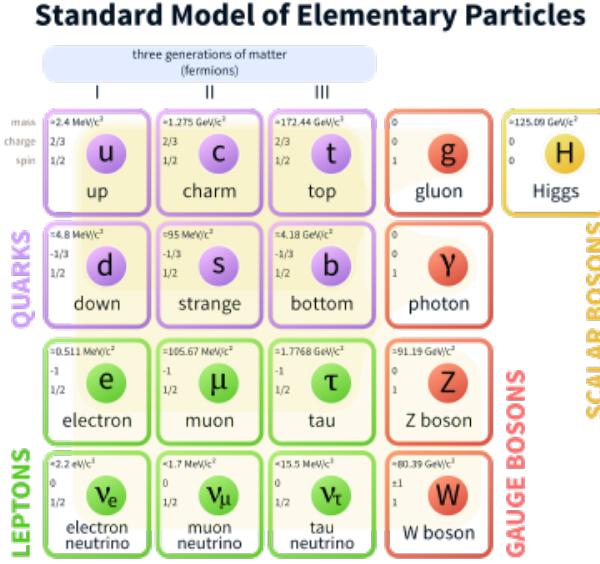


Figure 3.1: Particles of the Standard Model [20]

are described by an $SU(3)_C \oplus SU(2)_L \oplus SU(1)_Y$ local gauge symmetry. The model describes pointlike particles which interact via the strong, weak and electromagnetic forces. No gravitational interactions are described within the model.

The fermions of the model can be classified into two families- leptons and quarks- according to how they interact. The quark family consists of the up(u), down(d), charm(c), strange(s), top(t) and bottom(b) quarks, all of which are capable of interacting via the strong, weak and electromagnetic force. The lepton family, consisting of the electron(e), muon(μ), tau(τ), electron neutrino(ν_e), muon neutrino(ν_μ) and tau neutrino(ν_{τ}), are defined by the fact they carry no colour charge and so are incapable of interacting via the strong force, however they are still all capable of interacting via the weak force and the $e/\mu/\tau$ can interact electromagnetically. The gauge bosons are the mediators for the three fundamental forces of the model. The photon is a massless boson that mediates the electromagnetic force by coupling to particles with electrical charge. The gluon is also massless and mediates the strong force by coupling to particles with colour charge. The gluon is unique amongst the gauge bosons in that it is the only boson that carries the charge to which it couples (i.e. it is coloured) and so couples to itself. One direct consequence of this is that it is impossible to form a stable coloured state due to colour confinement and so quarks

are only observed in net-colourless states called hadrons. When a quark is produced in an interaction, it will typically undergo a process known as hadronization in which the quark will bind to quarks/antiquarks spontaneously produced from the vacuum to form quark-antiquark pairs known as mesons, or triplets of quarks or antiquarks known as baryons. The only exception to this is the top quark which will typically decay in a shorter timescale than is needed for hadronization to occur. The final three gauge bosons are the Z , W^+ and W^- which are all massive and mediate the weak interaction via their coupling to weak isospin.

Much like the fermions can be separated into quarks and leptons according to the way they interact, the underlying symmetry of the SM of $SU(3)_C \oplus SU(2)_L \oplus SU(1)_Y$ can be decomposed into separate parts according to the interactions that the symmetries describe. The $SU(3)_C$ group represents transformations of the colour state of a system and so describes interactions involving the strong force. These interactions are commonly referred to as QCD. The $SU(2)_L \oplus SU(1)_Y$ symmetry represents electroweak theory- a unified description of the weak and electromagnetic interactions. In this description, fermions can be thought of as consisting of left and right handed fields, where the left handed components transform as doublets under $SU(2)$ transformations while the right handed components only transform as singlets. The result of this is that the weak interaction only acts on the left handed field components. Hence the weak force only couples to left(right) handed particles (antiparticles.)

One of the most interesting features of electroweak theory occurs when considering the effect of gauge transformations on the Lagrangian of the system. In quantum field theory, fermions can be described by a dirac field with the following Lagrangian:

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial\mu\psi - m\bar{\psi}\psi \quad (3.1)$$

Applying a global phase transition of the form:

$$\psi \rightarrow \psi' = e^{i\alpha} \psi \quad (3.2)$$

will leave the Lagrangian unchanged due to the fact $e^{i\alpha}\psi e^{-i\alpha}\psi = 1$. However, in the case of local gauge transformations where $e^{i\alpha}\psi \rightarrow e^{i\alpha(\mathbf{x})}\psi$, i.e. the phase has a local space-time dependence, then eq(3.1) is no longer invariant:

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi - \bar{\psi}(\mathbf{x})\gamma^\mu\partial_\mu\alpha(\mathbf{x}) \quad (3.3)$$

In order to restore the invariance, the derivative ∂_μ must be replaced with the covariant derivative D_μ which is of the form:

$$D_\mu = \partial_\mu + ieA_\mu \quad (3.4)$$

where A_μ is a gauge field which transforms as:

$$A_\mu \rightarrow A'_\mu = A_\mu - \frac{1}{e}\partial_\mu\alpha(\mathbf{x}) \quad (3.5)$$

In electroweak theory the gauge fields required are found to consist of three weak isospin fields, W_1, W_2 and W_3 , coming from the $SU(2)$ group and one weak hypercharge field, B , from $U(1)$. The interesting result of this is the prediction that the bosons associated with these fields and the fermions they interact with should be massless, however this is experimentally found to be false as the bosons of the weak force, Z and W , have masses of 91.876 ± 0.0021 GeV and 80.385 ± 0.015 GeV respectively. Furthermore, in electroweak theory it can be shown that the presence of massive electroweak bosons results in unphysical predictions in the SM e.g. violation of unitarity when calculating the amplitude of $WW \rightarrow WW$ scattering [21]. These problems can be fixed via consideration of the final particle within the SM, the Higgs boson.

3.2 The Higgs Boson and the Origin of Mass

To solve the problems seen in the electroweak sector, Brout, Englert and Higgs [22][23] proposed that mass terms could be generated within the SM via the addition of a complex, scalar doublet of the group $SU(2)$ possessing four degrees of freedom:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \quad (3.6)$$

with potential:

$$V(\phi) = \mu^2 \phi^\dagger \phi + \frac{\lambda^2}{2} (\phi^\dagger \phi)^2 \quad (3.7)$$

The Higgs field is found to interact with the W_1, W_2, W_3 and B gauge fields. In the case that $\mu^2 < 0$, due to the Higgs field acquiring a non-zero expectation value, the $SU(2)_L \oplus SU(1)_Y$ symmetry is found to break leaving only a $U(1)_{em}$ symmetry corresponding to a massless photon. Of the four degrees of freedom associated with the Higgs field, the interaction of the field with the W and B gauge fields results in three massive gauge bosons corresponding to the measured Z and W^\pm masses, where the physically observed bosons actually represent mixtures of the underlying gauge fields:

$$\gamma = \cos \theta_W B + \sin \theta_W W_3 \quad (3.8)$$

$$Z = \cos \theta_W W_3 - \sin \theta_W B \quad (3.9)$$

$$W^\pm = \frac{1}{\sqrt{2}} (W_1 \mp i W_2) \quad (3.10)$$

Where θ_W is the weak mixing angle.

The last remaining degree of freedom of the Higgs field corresponds to the Higgs boson itself. The mass of the Higgs boson can be determined to be $m_H = \sqrt{2\lambda}\nu$, where λ is the Higgs self coupling parameter and ν is the vacuum expectation value for the Higgs field. While ν can be calculated within the standard model, λ is a free parameter and so the mass of the Higgs is not derivable. Experimentally it is found to be $\sim 125\text{GeV}$.

While the mass of the Higgs is of interest as it represents a free parameter in the standard model, there are many more properties of the Higgs that are important to measure. In particular the way in which the Higgs boson couples to other particles is well predicted within the SM and is expected to vary between various BSM models. Within the SM the coupling of the Higgs to fermions and bosons is different but depends on mass in each case:

$$g_{Hf\bar{f}} = \frac{M_f}{\nu} \quad g_{HBB} = \frac{2M_B^2}{\nu} \quad (3.11)$$

Due to this clear mass dependence, a fit of the coupling to each fermion as a function of the fermions mass represents a powerful way of testing the SM. The mass dependence on the Higgs couplings also presents a new way to perform indirect searches for new physics involving as yet unseen massive particles by looking at the branching ratio of Higgs decays to invisible decay products and the total Higgs decay width. This is of particluar interest in searches for dark matter which is known to interact gravitationally and so must possess mass.

3.3 Higgs Measurements at CLIC

The CLIC physics programme places a large focus on characterising the Higgs boson as it presents a new and relatively less well measured sector of the SM to explore. In

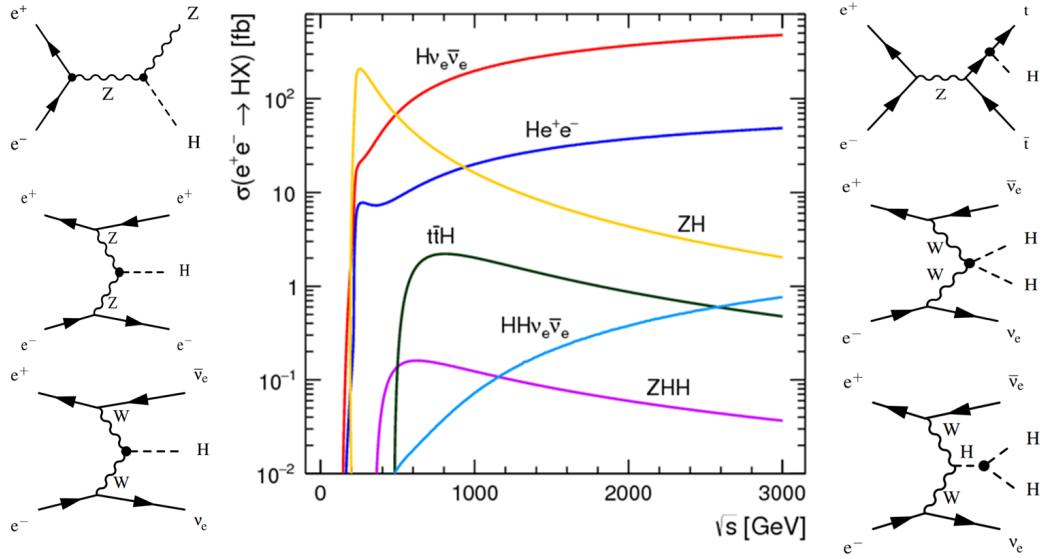


Figure 3.2: Cross Sections For Higgs Production Mechanisms [24]

particluar it will aim to measure the mass, width, and couplings of the Higgs in a model independent manner. Electron positron collisions provide access to numerous Higgs production mechanisms which can be seen in Figure 3.2. Due to the strong energy dependence on many of the cross sections on energy, different processes will be of interest at each of the three energy stages operated at CLIC. At 380GeV the focus will predominantly be on measuring the Higgsstrahlung (ZH) process in which a Z boson radiates a Higgs, while at higher energies vector boson fusion ($H\nu\bar{\nu}, He^+e^-$) dominates and new processes such as di-Higgs production become accessible. A summary of all the results from current Higgs studies performed by CLIC is available in [24].

3.3.1 Higgsstrahlung

One of the key aims of the experiment will be to examine the Higgsstrahlung process shown in Figure 3.3. In this process, if the four momentum of the Z boson can be measured to high precision, then because the initial conditions of the collision are well known, one can determine the mass of the particle it is recoiling against ($m_{rec}^2 = s + m_z^2 - 2E_z^2$) and infer the presence of the Higgs. This allows properties

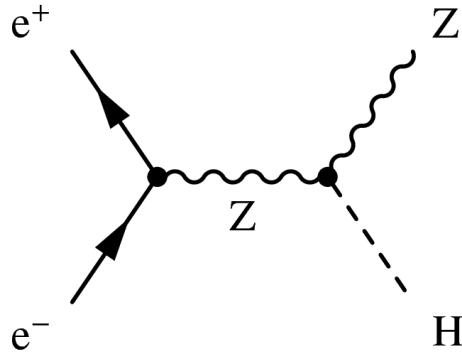


Figure 3.3: The Higgstrahlung Process

such as the Higgs mass, cross-section and coupling to the Z to be measured without actually ever measuring the decay products of the Higgs boson, which in turn allows the measurements to be model independent as few assumptions must be made about the properties of the Higgs. This method is not possible at hadron colliders such as the LHC where, even though the Higgsstrahlung process still occurs, the four momentum of the colliding particles can never be known as precisely due to their composite nature. Using the clean signal from cases where the Z decays to a pair of muons or electrons it is possible to measure the recoil mass to high precision and thus determine the mass of the Higgs to $\Delta m_H = 110 \text{ MeV}$ (see Figure 3.4) using data from the low energy stage only. This value can be further improved to $\Delta m_H = 44 \text{ MeV}$ when including direct measurement results from the $ee \rightarrow H\nu\bar{\nu}, H \rightarrow b\bar{b}$ channel at 3 TeV. Despite giving a poorer resolution on the Z four momentum, the $Z \rightarrow qq$ higgsstrahlung channel is also considered due to its larger cross section. Using this channel a limit of $BR(H \rightarrow \text{invis.}) < 0.97\%$ at 90% C.L. can be set.

3.3.2 Model Independent Extraction of Higgs Couplings

While the Higgsstrahlung alone allows the mass and branching ratios of the Higgs to be determined, it is further possible to extract the absolute width of the Higgs, Γ_H , by measuring the rates of several different Higgs processes and combining them in the right ratio. One such scheme proposed for doing this is shown in (3.12) [25]:

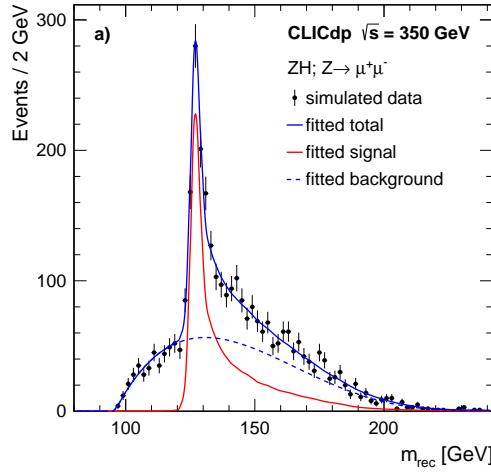


Figure 3.4: Reconstructed recoil mass from Higgsstrahlung process [24]

$$\Gamma_H = \frac{X_1^2 X_3^2}{X_4^2 X_2} \quad (3.12)$$

where

$$X_1 = \sigma_{ZH} \propto g_{HZZ}^2 \quad (3.13)$$

$$X_2 = \sigma_{H\nu\bar{\nu}} \times BR(H \rightarrow WW^*) \propto \frac{g_{HWW}^4}{\Gamma_H} \quad (3.14)$$

$$X_3 = \sigma_{H\nu\bar{\nu}} \times BR(H \rightarrow b\bar{b}) \propto \frac{g_{HWW}^2 g_{Hbb}^2}{\Gamma_H} \quad (3.15)$$

$$X_4 = \sigma_{ZH} \times BR(H \rightarrow b\bar{b}) \propto \frac{g_{HZZ}^2 g_{Hbb}^2}{\Gamma_H} \quad (3.16)$$

With the exception of X_1 , the choice of variables used is not unique (e.g. one could replace the production mechanism in X_1 and X_2 with ZZ-fusion rather than WW-fusion,) however the combination shown here is expected to give the highest precision on Γ_H due to the large cross-section associated with WW-fusion and the

Channel	Measurement	Observable	Statistical precision	
			350 GeV 500 fb ⁻¹	350 GeV 500 fb ⁻¹
ZH	Recoil mass distribution	m_H	110 MeV	
ZH	$\sigma(ZH) \times BR(H \rightarrow \text{invisible})$	Γ_{inv}	0.6%	
ZH	$\sigma(ZH) \times BR(Z \rightarrow l^+l^-)$	g_{HZZ}^2	3.8%	
ZH	$\sigma(ZH) \times BR(Z \rightarrow q\bar{q})$	g_{HZZ}^2	1.8%	
ZH	$\sigma(ZH) \times BR(H \rightarrow b\bar{b})$	$g_{HZZ}^2 g_{Hbb}^2 / \Gamma_H$	0.84%	
ZH	$\sigma(ZH) \times BR(H \rightarrow c\bar{c})$	$g_{HZZ}^2 g_{Hcc}^2 / \Gamma_H$	10.3%	
ZH	$\sigma(ZH) \times BR(H \rightarrow gg)$		4.5%	
ZH	$\sigma(ZH) \times BR(H \rightarrow \tau^+\tau^-)$	$g_{HZZ}^2 g_{H\tau\tau}^2 / \Gamma_H$	6.2%	
ZH	$\sigma(ZH) \times BR(H \rightarrow WW^*)$	$g_{HZZ}^2 g_{HWW}^2 / \Gamma_H$	5.1%	
Hv _e ̄v _e	$\sigma(Hv_e \bar{v}_e) \times BR(H \rightarrow b\bar{b})$	$g_{HWW}^2 g_{Hbb}^2 / \Gamma_H$	1.9%	
Hv _e ̄v _e	$\sigma(Hv_e \bar{v}_e) \times BR(H \rightarrow c\bar{c})$	$g_{HWW}^2 g_{Hcc}^2 / \Gamma_H$	14.3%	
Hv _e ̄v _e	$\sigma(Hv_e \bar{v}_e) \times BR(H \rightarrow gg)$		5.7%	

Figure 3.5: Expected statistical uncertainties for Higgs measurements at 350GeV at CLIC assuming unpolarised beams [24]

high branching ratio of $H \rightarrow b\bar{b}$ ($\sim 65\%$). In chapter 4 we will present our research on the precision to which X_2 can be measured during the 1.4 TeV run at CLIC. Currently at the LHC the standard process for extracting couplings from the equivalent measurements of $X_{2,3\&4}$ is to multiply through by the standard model value of the Higgs width [26]. This type of measurement is referred to as ‘model-dependent’ as the values determined for the Higgs couplings carry the implicit assumption that the standard model is correct in its prediction of the Higgs width. At CLIC, because the width can be measured experimentally there is no need to make this assumption and so the couplings are measured in a “model-independent” way. The unique ability of linear colliders to perform model-independent measurements is one of the largest driving factors for constructing and using them as a so called “Higgs-Factory”. One limiting factor for the model-independent measurements of the couplings is that they are always ultimately dependent on the precision to which the ZH cross section can be measured (predicted to be $\Delta h_{HZZ} = 0.8\%$) as this quantity is always needed in the ratio used to extract Γ_H .

In practice it is expected that an 11 parameter global fit to multiple variations of these measurements will be performed at each stage of operation to extract the Higgs width and its couplings to both fermions and bosons. The relevant inputs

Channel	Measurement	Observable	Statistical precision	
			1.4 TeV 1.5 ab ⁻¹	3 TeV 2.0 ab ⁻¹
Hv _e ̄v _e	H → b̄b mass distribution	m_H	47 MeV	44 MeV
Hv _e ̄v _e	$\sigma(Hv_e\bar{v}_e) \times BR(H \rightarrow b\bar{b})$	$g_{HWW}^2 g_{Hbb}^2 / \Gamma_H$	0.4 %	0.3 %
Hv _e ̄v _e	$\sigma(Hv_e\bar{v}_e) \times BR(H \rightarrow c\bar{c})$	$g_{HWW}^2 g_{Hcc}^2 / \Gamma_H$	6.1 %	6.9 %
Hv _e ̄v _e	$\sigma(Hv_e\bar{v}_e) \times BR(H \rightarrow gg)$		5.0 %	4.3 %
Hv _e ̄v _e	$\sigma(Hv_e\bar{v}_e) \times BR(H \rightarrow \tau^+\tau^-)$	$g_{HWW}^2 g_{H\tau\tau}^2 / \Gamma_H$	4.2 %	4.4 %
Hv _e ̄v _e	$\sigma(Hv_e\bar{v}_e) \times BR(H \rightarrow \mu^+\mu^-)$	$g_{HWW}^2 g_{H\mu\mu}^2 / \Gamma_H$	38 %	25 %
Hv _e ̄v _e	$\sigma(Hv_e\bar{v}_e) \times BR(H \rightarrow \gamma\gamma)$		15 %	10 %*
Hv _e ̄v _e	$\sigma(Hv_e\bar{v}_e) \times BR(H \rightarrow Z\gamma)$		42 %	30 %*
Hv _e ̄v _e	$\sigma(Hv_e\bar{v}_e) \times BR(H \rightarrow WW^*)$	g_{HWW}^4 / Γ_H	1.0 %	0.7 %*
Hv _e ̄v _e	$\sigma(Hv_e\bar{v}_e) \times BR(H \rightarrow ZZ^*)$	$g_{HWW}^2 g_{HZZ}^2 / \Gamma_H$	5.6 %	3.9 %*
He ⁺ e ⁻	$\sigma(He^+e^-) \times BR(H \rightarrow b\bar{b})$	$g_{HZZ}^2 g_{Hbb}^2 / \Gamma_H$	1.8 %	2.3 %*
t̄tH	$\sigma(t\bar{t}H) \times BR(H \rightarrow b\bar{b})$	$g_{Htt}^2 g_{Hbb}^2 / \Gamma_H$	8.4 %	—
HHv _e ̄v _e	$\sigma(HHv_e\bar{v}_e)$	λ	32 %	16 %
HHv _e ̄v _e	with -80 % e ⁻ polarisation	λ	24 %	12 %

Figure 3.6: Expected statistical uncertainties for Higgs measurements at 1.4TeV and 3TeV at CLIC assuming unpolarised beams [24]

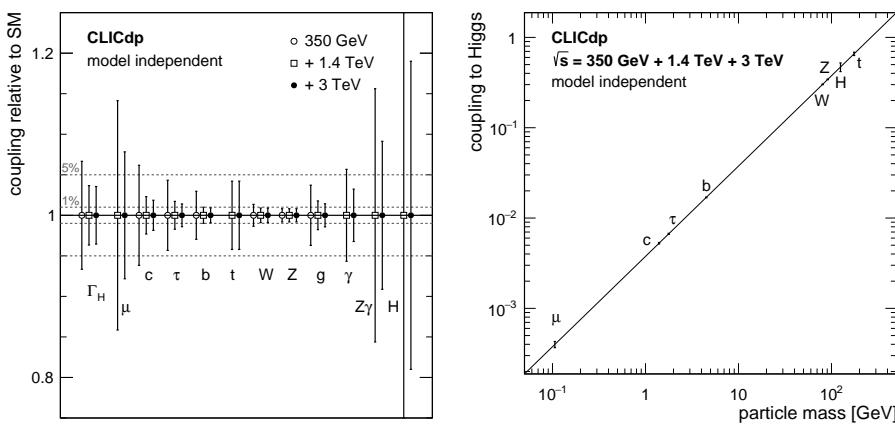


Figure 3.7: Expected precision on model independent measurements of the Higgs couplings [24]

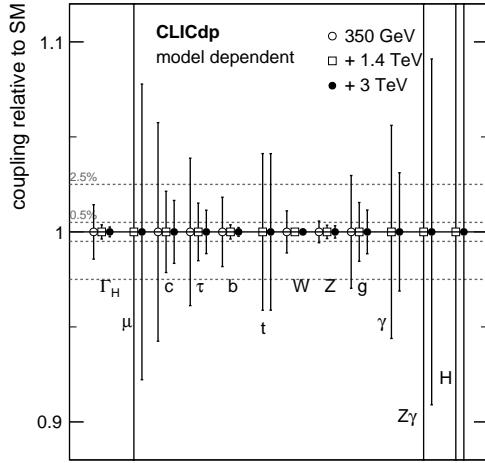


Figure 3.8: Expected precision on model dependent measurements of the Higgs couplings at CLIC [24]

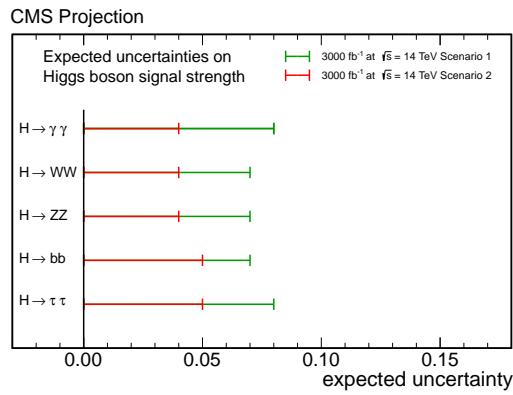


Figure 3.9: Expected precision on model dependent measurements of the Higgs couplings at CMS for the HL-LHC. Scenario 1 represents a case where the systematic and theoretical uncertainties remain at their current levels. In scenario 2 the theoretical uncertainty is scaled by a factor of a half and the systematic uncertainties are scaled by the square root of the integrated luminosity [27].

for these fits are shown in Table 3.5 and 3.6 while the results of the fits are shown in Figure 3.7

For context it is also important to compare these results to what can be expected from current leading experiments such as ATLAS and CMS at the LHC. Because the Higgs width can not be explicitly calculated at Hadron colliders, it is best to compare the model dependent version of the CLIC analysis with those predicted by ATLAS and CMS. In this situation, because the precision of the couplings is no longer limited by the precision on g_{HZZ} , the predicted precision for CLIC is seen to improve considerably. One can see from Figure 3.8 and 3.9 that in many cases CLIC is expected to provide an order of magnitude improvement over what can be achieved at the LHC with many of the key parameters associated with the Higgs being measured to sub percent precision.

Ultimately the aim of performing precision measurements is to allow the validation or rejection of theoretical models. While the results seen so far at the LHC suggest that the observed Higgs Boson is that of the SM, there are numerous alternative theories that predict a Higgs like particle with properties similar to what has been observed but which differ to a degree not yet measureable by current experiments. The details of these theories will not be expanded upon within this thesis, however the deviations expected in the Higgs couplings of these theories relative to the SM are shown in Table 3.1. These values should only be taken as a rough guideline for the precision required to discover/reject the theories as they are based on the assumption that new physics occurs at a specific scale (in this case 1 TeV,) however it is clear that the level of precision required to provide sensitivity to these models is expected to be greater than what will be possible with the LHC but could be within the scope of the proposed CLIC physics programme.

Model	κ_V	κ_b	κ_γ
Singlet Mixing	~6%	~6%	~6%
2HDM	~1%	~10%	~1%
Decoupling MSSM	~-0.0013%	~1.6%	~-4%
Composite	~-3%	~-(3-9)%	~-9%
Top Partner	~-2%	~2%	~+1%

Table 3.1: Generic size of Higgs coupling modifications from the SM values when all new pariles are $M \sim 1\text{TeV}$ and mixing angle satisfy precision electroweak fits. The Decoupling MSSM numbers assume $\tan \beta = 3.2$ and a stop mass of 1 TeV with $X_t = 0$ for the κ_γ prediction [3]

3.4 Top Quark Physics

The top quark is currently the heaviest particle within the SM and is the only quark that decays before undergoing hadronization. Due to it's high mass, top interactions are often seen as good channels for looking for BSM physics with a characteristic energy scale beyond what has currently been discovered. Due to it's high mass, the top is also the fermion with the highest coupling to the Higgs making it a good candidate for finding deviations from the SM within the Higgs sector. As such, the physics programme for CLIC places a strong focus on measuring the top quarks properties with the lowest energy stage of operation featuring a dedicated top threshold scan aiming to provide precision measurments of the top mass and width. The dominant production mechanism for top production is through the s-channel: $e^+e^- \rightarrow \gamma/Z \rightarrow t\bar{t}$ process shown in Figure 3.10. Using this process the properties of the $t\bar{t}\gamma$ and $t\bar{t}Z$ vertices can be measured. Examining these can provide sensitivity to contributions from BSM effects such as the existence of extra bosons (e.g. Z' [28]) which could provide an additional production channel, modifying the behaviour at the vertex. The $t\bar{t}X$ vertex can be written as below [29]:

$$\Gamma_\mu^{t\bar{t}X}(s, q, \bar{q}) = ie\{\gamma_\mu(F_{1V}^X(s) + \gamma_5 F_{1A}^X(s)) - \frac{\sigma_{\mu\nu}}{2m_t}(q + \bar{q})^\nu(iF_{2V}^X(s) + \gamma_5 F_{2V}^X(s))\} \quad (3.17)$$

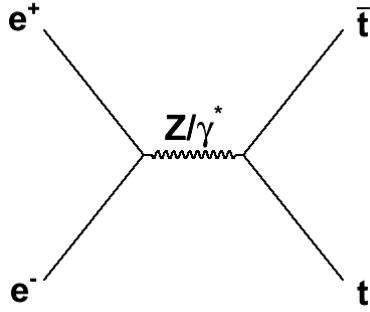


Figure 3.10: Dominant top production mechanism at electron positron collider

Where $X = \gamma/Z$, q and \bar{q} are the four momenta of the top and anti top, s is Mandelstam , γ_μ and γ_5 are the Dirac matrices corresponding to vector and axial-vector currents respectively, $\sigma_{\mu\nu} = \frac{i}{2}(\gamma_\mu\gamma_\nu - \gamma_\nu\gamma_\mu)$ allows for describing the scattering and F are the electroweak form factors. Within the SM, at tree level all the F_2 are zero as is F_{1A}^γ leaving only the following non-zero form factors:

$$F_{1V}^\gamma = \frac{2}{3} \quad (3.18)$$

$$F_{1V}^Z = \frac{1}{4 \sin \theta_W \cos \theta_W} \left(1 - \frac{8}{3} \sin \theta_W\right) \quad (3.19)$$

$$F_{1A}^Z = \frac{1}{4 \sin \theta_W \cos \theta_W} \quad (3.20)$$

where θ_W is the weak mixing angle. If higher order corrections are present it is possible for the zero valued form factors to become non-zero. Combinations of these factors can be related to physical observables which can be measured at CLIC. The couplings of the bosons to quarks with left or right handed helicity can be expressed as:

$$g_L^X = F_{1V}^X - F_{1A}^X \quad g_R^X = F_{1V}^X + F_{1A}^X \quad (3.21)$$

The most relevant variables are perhaps the total cross section and the forward backward asymmetry (A_{FB}) as these are the prospective measurements that are presented later within this thesis and will be discussed in more detail in Chapter 5. These variables are found to be dependent on the helicity of the incoming electrons [30] and so are more easily expressed in terms of the alternative form factors:

$$F_{ij}^L = -F_{ij}^\gamma + \left(\frac{-\frac{1}{2} + \sin \theta_W^2}{\sin \theta_W \cos \theta_W} \right) \left(\frac{s}{s - m_Z^2} \right) - F_{ij}^Z \quad (3.22)$$

$$F_{ij}^R = -F_{ij}^\gamma + \left(\frac{\sin \theta_W^2}{\sin \theta_W \cos \theta_W} \right) \left(\frac{s}{s - m_Z^2} \right) - F_{ij}^Z \quad (3.23)$$

Where L,R represent the polarization of the electron, i=1,2 and j=V,A. In this notation, for an electron polarization P, the total $ee \rightarrow Z/\gamma \rightarrow tt$ cross section and A_{FB} can be expressed as:

$$\sigma_P = \frac{8\pi\alpha(s)^2}{s} \beta \left\{ \left(1 + \frac{1}{2\gamma^2} \right) (F_{1V}^P)^2 + (\beta F_{1A}^P)^2 + 3F_{1V}^P F_{2V}^P + \left(1 + \frac{1}{2\gamma^2} \right) (F_{2V}^P)^2 \right\} \quad (3.24)$$

$$A_{FB}(P) = \mp \frac{12\pi\alpha(s)^2 \beta^2}{s} \frac{F_{1A}^P (F_{1V}^P + F_{2V}^P)}{\sigma_P} \quad (3.25)$$

where $\alpha(s)$ is the electromagnetic coupling, γ and β are the Lorentz factor and velocity of the top, and for (3.25), the + and – refer to the $P = R$ and $P = L$ cases respectively. If one performed just one measurement of the cross section and A_{FB} it is clear that it would not be possible to extract the form factors from these variables alone as there are more variables than there would be constraints. However, because the cross section and A_{FB} vary with β , γ and P , then by performing the measurement at multiple energies and making use of the fact CLIC can be operated with different beam polarizations, it becomes possible to extract all the relevant

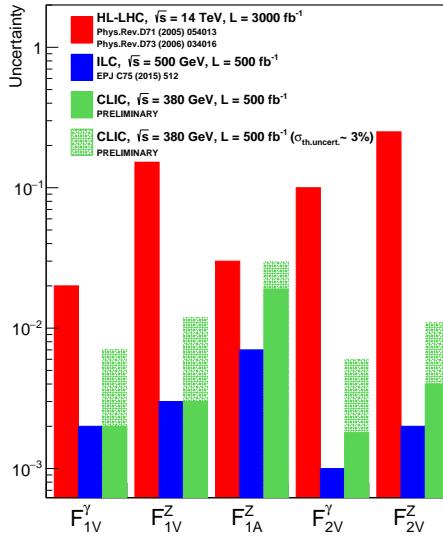


Figure 3.11: Expected precision on CP conserving electroweak form factors at future colliders [1]

couplings. The only exceptions to this are the F_{2A}^X factors which do not effect these two variables and so must be measured using alternative methods.

The predicted uncertainty to which the couplings are expected to be measured to at CLIC based on generator level studies, as well as the equivalent results for ILC and HL-LHC is shown in Figure 3.11. The expected precision from performing these measurements at a lepton collider is an order of magnitude better than what can be expected from hadron colliders. Overall there will be more tops produced in a hadron collider, however the production mechanisms are often more complicated making it harder to extract the couplings from a theory perspective, and it is also harder to successfully identify tops (which typically decay to at least one jet) in an environment that contains QCD jets from beam remnants compared to at lepton colliders where there is minimal QCD background within an event.

CHAPTER 4

Higgs Analysis

Reinforce context for measurement as part of the higgs width measurement

Take everything from analysis note!!

lepton finding

jet finding- higgs and W mass plots

btagging

describe BDTs

final selection and uncertainty

Impact of this on the overall higgs measurements at CLIC & compare to Higgs to
qqqq channel

4.1 Introduction

One of the key aims for the CLIC experiment will be to perform a model independent measurement of the Higgs total width. Any deviations from the value predicted by the Standard Model for this would be clear evidence that there is new physics beyond the Standard Model involving particles that can interact with the Higgs. The size of any deviations could also give us an indication of the scale at which new physics occurs helping to guide the design of future high energy colliders. No current experiment has the capability of performing this measurement and so it is essential that it is measured at an electron-positron collider. The determination of the Higgs total width is dependent on the measurement of four quantities e.g. as described in Ref.[25]. Here we will consider the measurement of one of these quantities, $\sigma_{H\nu\nu} \times \text{Br}(H \rightarrow WW^*)$. As can be seen in Fig.4.1, this measurement is made possible during the 1.4TeV CLIC run as the cross section for the WW fusion process is the dominant Higgs production mechanism at this energy. This observable has already been studied at 1.4TeV using the $WW \rightarrow qqqq$ channel [?], yielding an expected statistical precision of 1.4% for the nominal integrated luminosity of 1.5ab^{-1} . Here we will look at the complementary $WW \rightarrow qql\nu$ channel (shown in Fig.4.2), where $l=e,\mu$, with the intention of combining our results with the existing measurement to estimate the overall precision achievable at CLIC.

The signal and background processes we have examined in our analysis, along with their cross sections and sample production IDs, are summarised in Table. 4.3. In all cases the detector model used is CLICILD_CDR, a variant of the ILD detector designed for ILC [12]. The main backgrounds of note are: the $ee \rightarrow qql\nu$ process (dominated by $e^+e^- \rightarrow W^+W^-$) as it has a very similar topology to our signal process and so is expected to be the most difficult to exclude; and the $ee \rightarrow H(WW^* \rightarrow qqqq)\nu\nu$ process as our ability to correctly classify these events will determine how easily our results can be combined with measurements using this channel.

ⁱSample 2022 is the $ee \rightarrow H\nu\nu$ inclusive sample. The relevant events were extracted from this main sample

ⁱⁱThe cross-section for these events were scaled by a factor of 2 to account for interactions

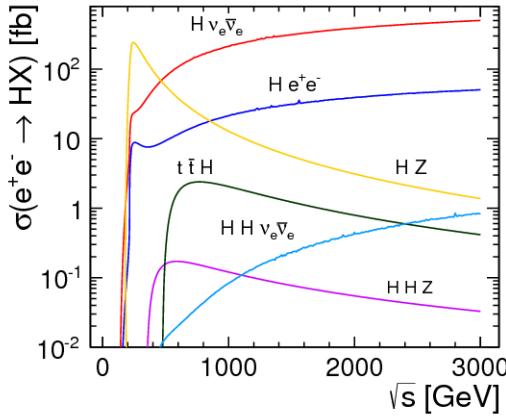


Figure 4.1: Cross sections for Higgs production mechanisms in an e^+e^- collider as a function of centre-of-mass energy [31]. For energies above 500GeV, Higgs production is dominated by the WW-fusion process(red.)

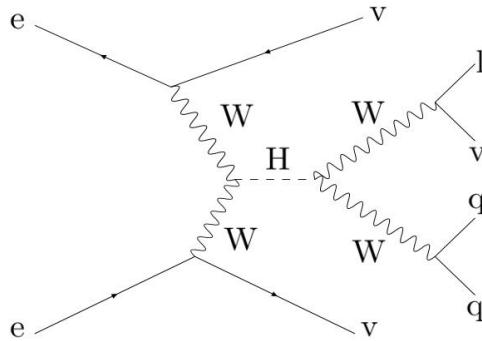


Figure 4.2: Higgs production via WW fusion with the Higgs decaying to a WW pair, which then decays semileptonically.

4.2 Event Reconstruction

Reconstruction of the signal events was performed using ILCsoft v01-17-06 and was carried out in two main stages as described below. The first stage was to identify the isolated lepton associated with the leptonic W boson decay. The second stage involved removing this isolated lepton and resolving the remaining particles into two jets that were associated with the two quarks produced by the hadronically decaying W boson. Using the two jets, the W boson could then be reconstructed and combined

occurring with both the electron and positron. In the case of beamsstrahlung events simulated using GUINEA-PIG [33], a further scaling of 0.75 was applied to account for the lower luminosity of $e\gamma$ events

Process	Cross Section(fb)	Production ID[32]	Events Used
Signal: ee \rightarrow H(WW* \rightarrow qqll)ll	18.9	2022 ⁱ	70000
ee \rightarrow H(WW* \rightarrow qqqq)ll	25.6	2022 ⁱ	140000
ee \rightarrow H* \rightarrow Other	199.6	2022 ⁱ	750000
ee \rightarrow qq	4009.5	2091	500000
ee \rightarrow qqqq	1328.1	2163	300000
e γ \rightarrow eqqq (γ from EPA)	32308	2515 ⁱⁱ	500000
γ e \rightarrow eqqq (γ from BS)	56043	2527 ⁱⁱ	500000
ee \rightarrow qqll	787.7	3243	500000
ee \rightarrow qql ν	2725.8	3246	400000
ee \rightarrow qql ν	4309.7	3249	1000000

Figure 4.3: Samples used for the analysis

with the isolated lepton to reconstruct the Higgs boson. The Higgs candidate we reconstruct will not be complete due to the lepton neutrino produced from the W decay. However, the observed properties will still be sufficient for discriminating between signal and background events.

4.2.1 Lepton Identification

Two different methods were used for identifying leptons. Our primary method for particle identification is to assume that the highest energy electron or muon (as identified by PandoraPFA [17]) corresponds to the isolated lepton from the leptonically decaying W boson. This method was found to have an efficiency and purity of 93% and 96% for identifying the isolated lepton.

The second method used a series of cuts to select the isolated lepton. The first stage of this was to group the particles in the event into four jets. This was done using FastJet [34] to implement the kt-algorithm using the E-scheme for recombination with an R-parameter of 0.4. We then required that the energy of the isolated lepton (electron or muon) constituted more than 35% of the visible energy of the jet it

was contained within. For electrons it was then required that at least 90% of the total energy of the particle was deposited in the ECAL, and the ratio of energy to momentum for the particle was between 0.75 and 1.25. For muons it was required that less than 35% of the total energy of the particle was deposited in the ECAL, and the ratio of energy to momentum should be between 0.01 and 0.60. This method yielded an efficiency of 91% and a purity 74%. Although this approach is not as performant as the first method, it allows more than one lepton to be selected. As a result it is useful for discriminating between signal and background processes as requirements can be placed on the number of leptons identified by this selection.

In summary, the first method is used to select a single isolated lepton, which is then used for reconstruction, while the number of lepton candidates selected by the second method is used as a discriminating variable to distinguish between signal and background processes.

4.2.2 Jet Finding

Following the lepton finding, the remaining particles (not including the isolated lepton) are forced into two jets to reconstruct the properties of the two quarks produced from the hadronic W decay. This was carried out using the same jet finding algorithm as was used for the lepton identification. The optimization of the R-parameter was performed by using Monte Carlo information to obtain what mass we would measure for the reconstructed Higgs for various values of R, when using Monte Carlo truth kinematic information of the lepton neutrino in our reconstruction. An acceptably small bias in the reconstructed mass was found for an R value of 0.4, indicating we were successfully reconstructing the quark pair.

4.3 Flavour Tagging

Flavour tagging of events was performed using LCFIPlus v00-05-02 [35]. Three neural nets were trained to identify u/d/s, b and c quarks respectively with 50,000 $e\bar{e} \rightarrow Z\nu\nu$, $Z \rightarrow qq$ events used for each neural net. Application of these neural nets returned two parameters for jets within the event that quantify the probability of the jet being either a b-jet or c-jet. For this analysis, identifying b-jets is more useful for discriminating against the relevant backgrounds. Performance of the b-tagging was evaluated by applying the neural nets to a sample of 150,000 events containing an equal number of $Z \rightarrow$ light, c and b quarks. It can be seen from Fig.4.4 that a purity of 90% can be achieved while still retaining an efficiency of 80%.

4.4 Event Selection

Event selection was performed using the TMVA package [36] to produce a Boosted Decision Tree (BDT). The BDT used 7×10^4 signal events and 4×10^6 background events, split evenly between training and testing samples. A collection of 19 variables is used for the training: mass of the reconstructed Higgs and W bosons; energy of the W boson; missing energy and transverse momentum; number of isolated leptons selected; PID of lepton; transverse momentum of lepton; angle of lepton and W boson relative to the beam axis; magnitude of minor thrust value; number of particle flow objects (PFOs) in the two jets; average angle of the two jets relative to the beam axis; kt jet resolution parameter y_{12} ; number of tightly selected PFOs in the event; angular separation of the isolated lepton and the W boson; minimum angular separation and transverse momentum of the lepton relative to either jet, and the combined b-tag value for both jets. A set of loose pre-selection cuts were also applied before the training to remove events that were clearly background. The cuts used were: energy of the W boson $\geq 591\text{GeV}$, Mass of the W boson $\geq 231\text{GeV}$, Mass of the reconstructed Higgs $\geq 306\text{GeV}$ and 667GeV ; total missing energy $\geq 1400\text{GeV}$. The input signal and background distributions for every input variable after application

of these cuts can be seen in the appendix, and the resulting BDT classifier output can be seen in Fig.4.5.

Figure4.5 shows that there is a high degree of separation achieved between signal and background events. The optimal BDT cut for maximising the signal to background ratio was determined to be at 0.21 and the effect of the pre-selection cuts and applying this BDT cut on the signal and background processes can be seen in Fig. 4.6. The resulting significance ($S/\sqrt{S+B}$) after these cuts has been calculated to be 77 giving a statistical uncertainty of 1.3% on σ_{Br} for an integrated luminosity of 1.5ab^{-1} . This value is similar to that observed for the $\text{WW}\rightarrow\text{qqqq}$ final state, as expected. By neglecting the case where the isolated lepton is a τ , we have reduced our statistics to two thirds that of the hadronic channel which inherently limits the precision that can be achieved. However, the backgrounds for the hadronic channel are much larger, making them harder to remove which leads to a reduced precision. Looking in detail at the backgrounds after our selection, we can see that many of the backgrounds have been almost completely removed leaving only $\text{ee}\rightarrow\text{H}(\rightarrow\text{other})\nu\nu$ and $\text{ee}\rightarrow\text{qql}\nu$ as the dominant backgrounds. This is to be expected as these events most closely mimic our signal, which is mainly distinguished by its large missing energy. In the case of $\text{H}\rightarrow\text{other}$ events it was determined that 26% of the remaining events came from $\text{H}\rightarrow\tau^+\tau^-$ processes with a further 25% from $\text{H}\rightarrow\text{WW}^*$ processes with one or more of the Ws decaying to a τ . As such, attempts were made to veto τ events by rejecting events in which one or more hadronically decaying τ was explicitly identified using a Tau Finder [37]. However, the number of τ s misidentified in the signal channel was determined to be too high to veto the τ events without significantly increasing the overall statistical uncertainty on σ_{Br} and so τ identification is not used in the final selection. The efficiency for selecting $\text{WW}^*\rightarrow\text{qqqq}$ events in the $\text{WW}^*\rightarrow\text{qql}\nu$ channel has been calculated to be 1.8% which should be sufficiently low that a straightforward combination of the uncertainties determined by both channels can be made. However, further investigation must be done to confirm this. The efficiency of the $\text{H}\rightarrow\text{WW}\rightarrow\text{qql}\nu$ events in the $\text{H}\rightarrow\text{WW}\rightarrow\text{qqqq}$ channel has yet to be confirmed, therefore a final combined result has not yet been

performed. Where $H \rightarrow WW$ candidates are identified by both selections, attributing them to a final state on the basis of the predicted purities, would simplify the calculation. Systematic uncertainties have not been described here but are expected to be dominated by the uncertainty on the measured $WW \rightarrow qql\nu$ branching ratio of 1.1% [38].

4.5 Conclusion

In summary, we have performed a full analysis of the $ee \rightarrow H(WW^*)\nu\nu$, $WW^* \rightarrow qql\nu$ decay channel using a large set of backgrounds with the aim of measuring the $H \rightarrow WW^*$ branching ratio. A 19 variable BDT was used to select signal events where the final state charged lepton is either an electron or a muon, and to remove background which was found to be dominated by $ee \rightarrow H(\rightarrow \text{Other})\nu\nu$ and $ee \rightarrow qql\nu$ in the final selection. The resulting statistical uncertainty was found to be:

$$\delta\sigma_{H\nu\nu} \times \text{BR}(H \rightarrow WW^*) = 1.3\%$$

The efficiency for incorrectly selecting $ee \rightarrow H(WW^*)\nu\nu$, with $WW^* \rightarrow qqqq$, in the $WW^* \rightarrow qql\nu$ channel, was found to be 1.8%. The correlated overlap in selections developed for the $WW^* \rightarrow qqqq$ and $WW^* \rightarrow qqlv$ final states would be taken into account when combining the individual results.

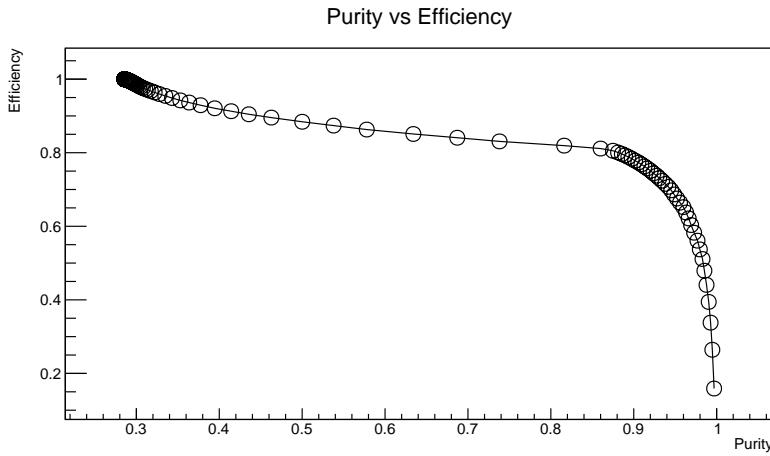


Figure 4.4: Purity vs efficiency for identifying b-jets, obtained from a sample of $Z \rightarrow$ light, c and b quark events simulated at $\sqrt{s} = 1.4\text{TeV}$

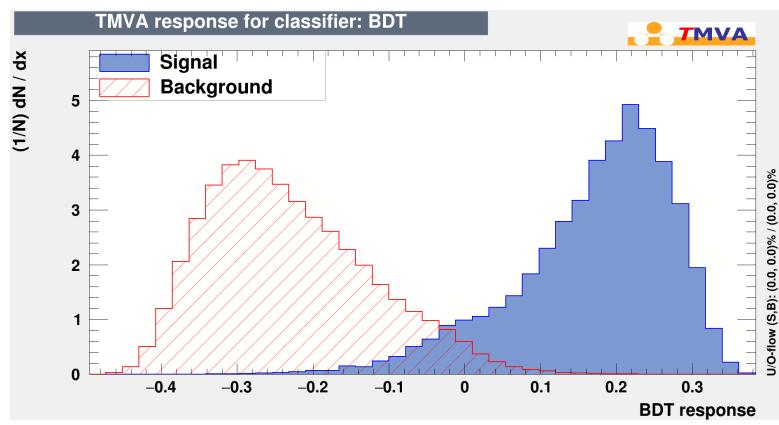


Figure 4.5: BDT response for signal and background events after TMVA classification

Process	Cross Section(fb)	Pre-selection Eff(%)	BDT Cut Eff(%)	Events After
Signal	18.9	99.99	42.65	
$ee \rightarrow H(WW^* \rightarrow qqqq)\nu\nu$	25.6	99.96	1.79	
$ee \rightarrow H(\rightarrow \text{Other})\nu\nu$	199.6	99.62	1.26	
$ee \rightarrow qq$	4009.5	76.95	0.01	
$ee \rightarrow qqqq$	1328.1	36.03	0.01	
$e\gamma \rightarrow eqq (\gamma \text{ from EPA})$	32308	67.00	0.01	
$\gamma e \rightarrow eqq (\gamma \text{ from BS})$	56043	95.84	0.01	
$ee \rightarrow qq\nu\nu$	787.7	96.59	0.07	
$ee \rightarrow qql\bar{l}$	2725.8	89.75	0.01	
$ee \rightarrow qql\nu$	4309.7	66.44	0.07	

Figure 4.6: Efficiency for all processes following pre-selection and BDT response cuts and the number of events expected to satisfy these requirements, for an integrated luminosity of 1.5ab^{-1} .

CHAPTER 5

Top Physics

5.1 Introduction

Define what the top asymmetry is, measurement at previous experiments(tevatron, LHC, b asymmetry at lep), significance of it in terms of EW form factors

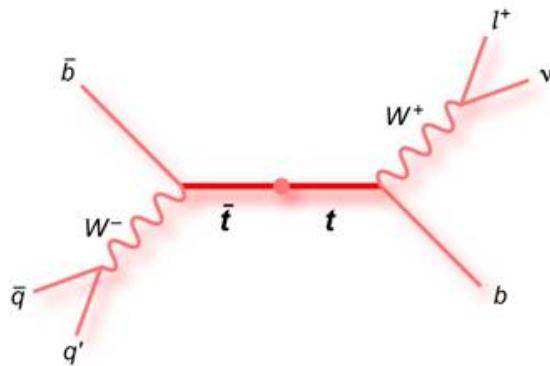


Figure 5.1: Semileptonic $t\bar{t}$ decay

Decay Mode	Branching Fraction
$t\bar{t} \rightarrow WbWb \rightarrow qqbqqb$	
$t\bar{t} \rightarrow WbWb \rightarrow qqbql\nu b$	
$t\bar{t} \rightarrow WbWb \rightarrow l\nu bl\nu b$	
$t\bar{t} \rightarrow Other$	

Table 5.1: Top Pair Decay Modes

Process	Cross Section(fb)	Production ID	Events Used
Signal: $ee \rightarrow t\bar{t} \rightarrow qqqql\nu$	xx	6589,6592,6634,6637	xx

Table 5.2: Top Pair Decay Modes

As tops decay almost exclusively to a W and b they are typically described in terms of the resulting decay modes of the Ws. The dominant decay modes are described in table 5.1. Here we will look at measuring A_{FB}^t using the semileptonic $t\bar{t}$ decay channel (see figure 5.1) in which one. This decay mode is ideal as the lepton from the leptonically decaying top provides the ability to charge tag the tops while the fully hadronic decay allows an accurate measurement of the production angle of the tops, both of which are necessary for measuring A_{FB}^t to high precision. The dominant signal and background processed examined by this analysis, as well as their cross sections and production ID numbers are shown in table 5.2. All samples are simulated using the CLICILD_CDR detector model. This is a variation of the ILD detector model developed for use at the ILC. The samples also include an overlay of $\gamma\gamma \rightarrow$ hadron events from beamstrahlung based on a 30 ns window around the generated physics events. The dominant backgrounds are expected to be from alternative $t\bar{t}$ decays (fully hadronic decay modes and semileptonic decays containing taus) which will have very similar topologies to the fact they will both contain a hadronically decaying top.

5.2 Event Reconstruction

Reconstruction of the signal events is performed using ILCSoft v01-17-10 and consists of three main stages. The first stage is to identify isolated leptons arising from the leptonically decaying top. These leptons are then removed and the remaining PFOs are resolved into two large radius “fat jets”. These two fat jets must then be associated with either the b jet produced by the leptonically decaying top or to the combination of three jets arrising from the hadronically decaying top. A kinematic fitter is then used to reconstruct the neutrino and any ISR/Beamstrahlung photons present in the event.

5.2.1 Lepton Finding

Lepton finding is the first stage of reconstruction performed in each event. Due to the fact the measurement of A_{FB}^t is entirely reliant on using the lepton charge to distinguish between top and antitop decays, it is essential that a high efficiency and purity are achieved and that there is no angular dependence on the performance. For this analysis lepton finding is done in two steps. Firstly, lepton candidates with energy $\gtrsim 10$ GeV are identified using the particle ID provided by the Pandora Particle Flow Algorithm [17]. Only muons and electrons are examined due to the fact tau leptons require different reconstruction techniques to identify and are typically reconstructed with significantly lower efficiency. This first stage removes $\gtrsim 90\%$ of fake candidates with negligible impact on efficiency. The second stage of lepton selection is to examine how isolated each of the candidates are. This is evaluated by resolving all PFOs in the event into five jets, then for each lepton candidate, measuring the energy of the candidate relative to the jet it was been associated with. For this process the ee kt algorithm was chosen for the jet finding to ensure that all lepton candidates are always placed within a jet. The lepton candidate found to have the highest ratio of $E_{Candidate}/E_{Jet}$ is then declared to be the isolated lepton arrising from the letonically decaying top. In the event that no lepton is

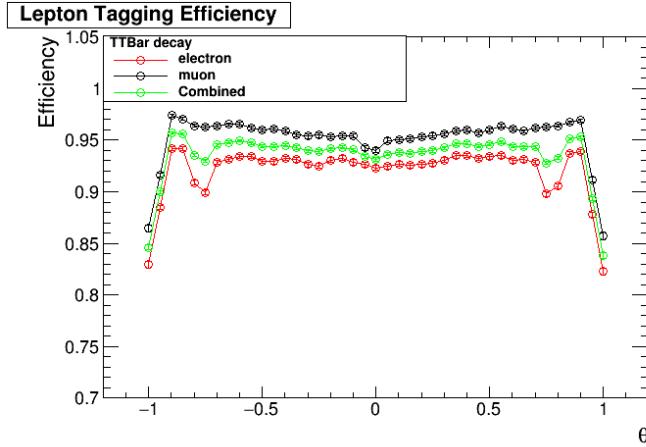


Figure 5.2: Efficiency for identifying leptons with the correct charge as a function of angle

selected by the first step, the restrictions on the particle ID and energy are relaxed and the lepton is selected purely based on which PFO is the most isolated according to step two. This method ensures that there is always exactly one lepton selected per event. The net efficiency with which this method selects a candidate with the correct charge is found to be 93% for electrons and 96% for muons.

As well as understanding the net efficiency for finding leptons it is also important to examine the angular dependence of the efficiency to ensure there is no bias that could effect the measurement of A_{FB}^t . Figure 5.2 shows how the efficiency varies with angle. The efficiency is seen to rapidly decline for $|Cos\theta| > 0.9$ due to detector acceptances. A decrease in efficiency is also seen for electrons at angles corresponding to the transition point between the ECAL barrel and endcaps. This effect is not seen for muons as they are also reconstructed using the muon detectors placed at a larger radius. Overall the efficiency is seen to be consistently worse for electrons than muons. This is to be expected as muons produce easily recognisable signatures in the detector due to the fact they typically penetrate through the tracker,ECAL,HCAL and muon sytems whereas electrons only leave deposits in the tracker and ECAL. In the case that tracks are lost during reconstruction or are wrongly associated to other PFOs it is then possible for photons to wrongly be labelled as electrons and vice versa leading to a higher fake rate for electrons.

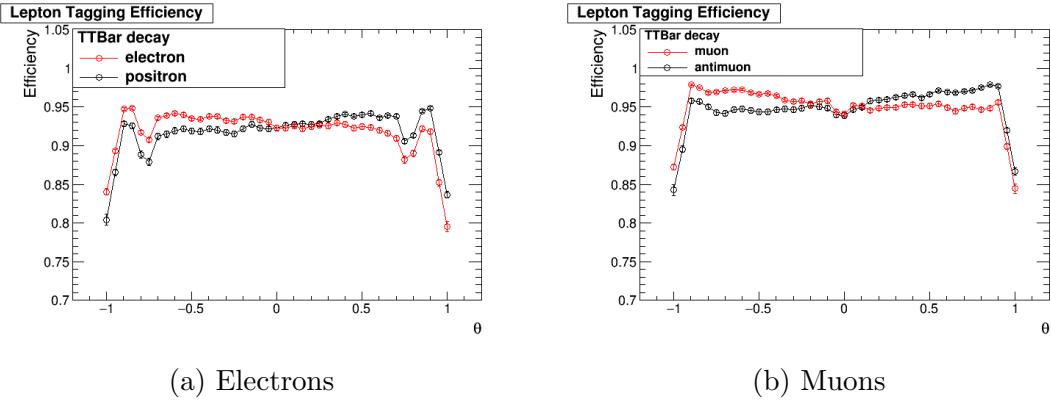


Figure 5.3: Angular dependence of lepton finding for particles vs antiparticles

As well as checking the angular dependence of the charge tagging efficiency it is also key to examine the charge dependence of the lepton finding to make sure there is no preference for identifying particles over antiparticles. The angular dependence of the charge tagging efficiency for particles vs antiparticles is shown in figures 5.3. An asymmetry in the performance is observed in both electrons and muons.

It arises from the underlying asymmetry in the production of particles vs antiparticles due to forward-backward asymmetries. The forward backward asymmetry mean that tops are preferencialy produced in one direction while antitops are produced more often in othe opposite direction, however due to charge conservation this also means that the W bosons and leptons are produced asymmetrically too. Because the collisions are taking place well above the top pair production threshold, the W bosons will gain a large boost forcing them to travel in the same direction as the inital top. The polarization of the W also means that the lepton will also be preferentially produced along the same diection as the W and can only be produced in the opposite direction with a lower energy. Overall this means that leptons are produced with higher energy in one direction and lower energy in the opposite direction while for antileptons this directional dependence is reversed. The effect is shown in 5.4 where it is seen that positrons are produced with higher energy in the forward direction($\cos\theta > 0$) than the backward direction. It is known that the efficiency for reconstructing leptons at CLIC increases with energy and so the fact the energy and angle at which leptons are produced are correlated results in the asymmetric

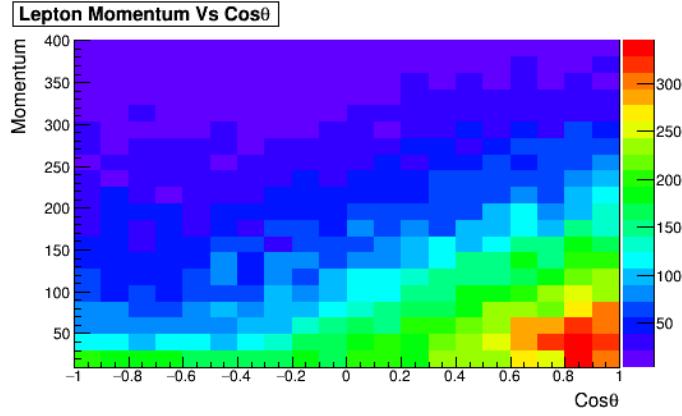


Figure 5.4: Correlation between lepton momentum and angle for positrons only

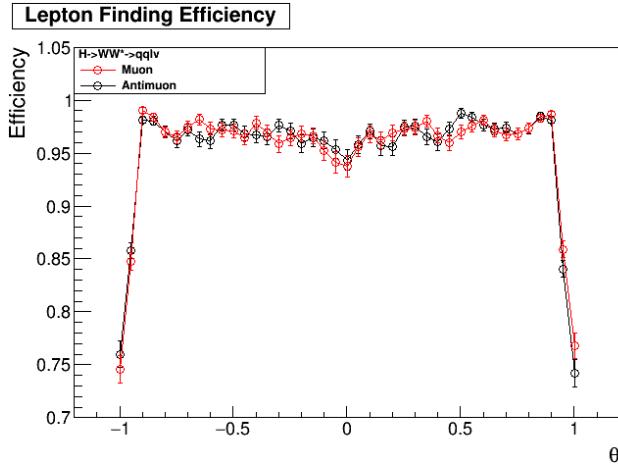


Figure 5.5: Charge tagging efficiency for $ee \rightarrow H\nu\nu, H \rightarrow WW \rightarrow qql\nu$. The efficiency is seen to be symmetric for particles and antiparticles when they are produced with the same initial angular distribution.

angular efficiency for correctly reconstructing the lepton. Further evidence for this theory is shown in figures 5.5 and 5.6 which show that the asymmetry disappears when either the production mode for the leptons is symmetric or when low energy leptons are not included.

5.2.2 Fat Jet Finding

Jet reconstruction was performed using the FastJet package [34]. Due to the high energy of the collisions relative to the top mass, the tops produced are highly boosted and produce highly collimated decay products (see figure 5.7.) This means it is

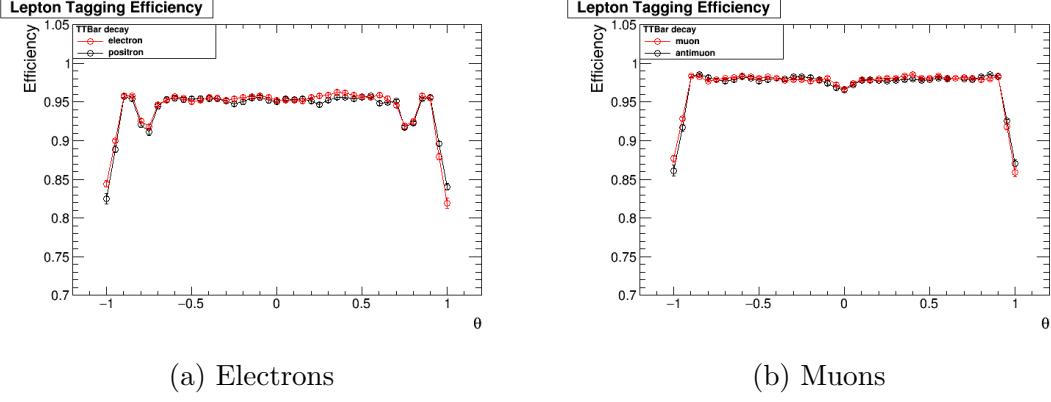


Figure 5.6: Charge tagging efficiency after 20 GeV lepton momentum cut. The efficiency is seen to be symmetric for leptons with momentum $\gtrsim 20$ GeV/

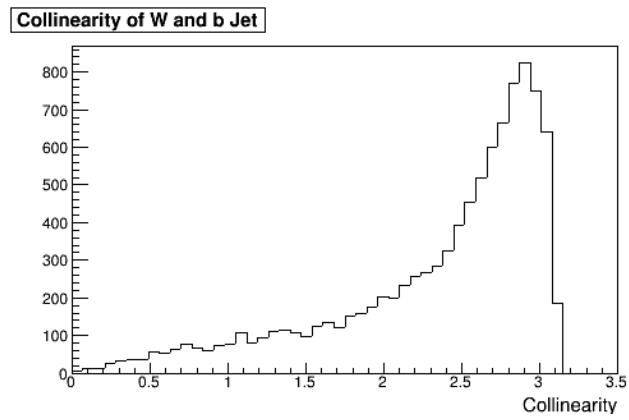


Figure 5.7: Separation between W and b jets from top decay. The pair are typically too collimated to allow the b-jet and the pair of jets from the W decay to be successfully resolved into three distinct jets

typically not possible to resolve the decay products from the hadronically decaying top into three jets corresponding to the b-jet and light quark jets from the W decay. As a result an alternative approach to jet reconstruction is considered based on the concept of fat jets, an approach already being used at the LHC[?]. Fat jets are large radius jets and are used to cluster groups of jets that can't be accurately resolved individually into one larger jet. For the purpose of this analysis the events are clustered into fat jets which should correspond to the b-jet from the leptonically decaying top and to the whole set of decay products from the hadronically decaying top. The mass and substructure variables (see 5.4) of these fat jets can then be used to distinguish genuine top events from backgrounds. Two jet algorithms were considered for reconstructing the fat jets- the kt algorithm [?] and Valencia algorithm [?]. The kt algorithm is already extensively used at hadron colliders while the Valencia algorithm is a newer algorithm designed for future lepton colliders that offers improved performance in handling beam backgrounds. The performance of both algorithms is shown in figure 5.8. For both algorithms it is seen that at higher R the resolution on the top mass gets worse while for lower R sub peaks start to appear in the mass distribution corresponding to partial reconstructions of the top (either W Boson or single quark). The kt algorithm is seen to produce a consistently broader distribution in the top mass. Placing a cut on the collision energy of $E_t < 1.2$ TeV reveals that these lower mass peaks only occur for lower collision energies where the top decay products will be less collimated and so the fat jet finding can merge components from both the hadronic and leptonic tops into each jet. This analysis will be focusing on reconstructing the most boosted tops. As a result the Valencia algorithm is preferred due to its better mass resolution. Performance for less boosted top decays might be improved by examining the performance of a more conventional jet analysis looking to resolve all four individual quarks whenever the fat jet finding produces jets outside the top mass window. This possibility is discussed later in section 5.5. Here the Valencia algorithm with $R=1.5$, $\beta=1$ and $\gamma=1$ is chosen as the optimal jet reconstruction method to provide a balance between mass resolution and the frequency of partial reconstructions.

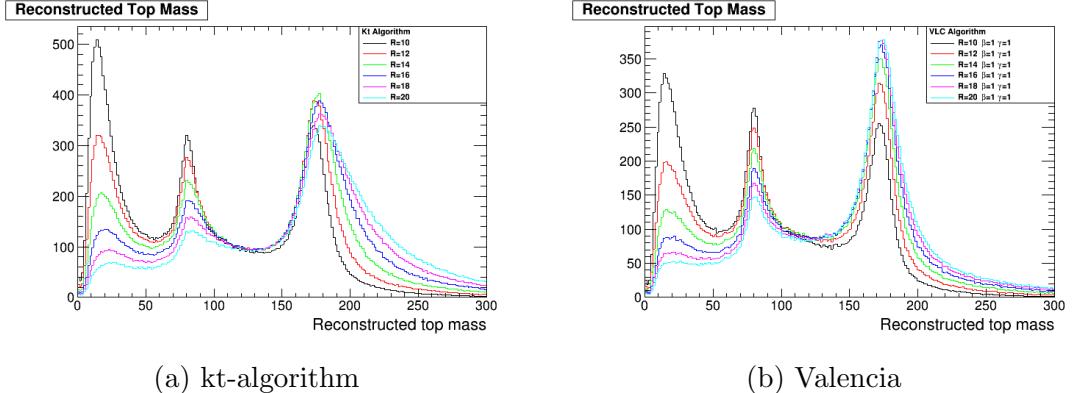


Figure 5.8: Performance of both jet finding algorithms for various parameter settings. The k_t algorithm is seen to produce a broader distribution in the top mass peak so the Valencia algorithm is preferred. For both methods it is seen that a lower R results in the development of peaks from partial reconstruction of the top jet (W Boson or single quark) while a larger R produces a broader peak at the top mass. A balance is found between producing a narrow top mass width while minimising sub peaks by selecting a radius of $R=1.5$

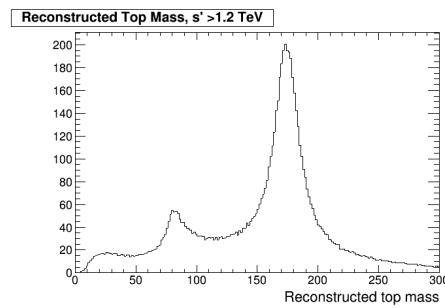


Figure 5.9: Reconstructed top mass for the Valencia algorithm in events close to the nominal collision energy (E_c ; 1.2 TeV)

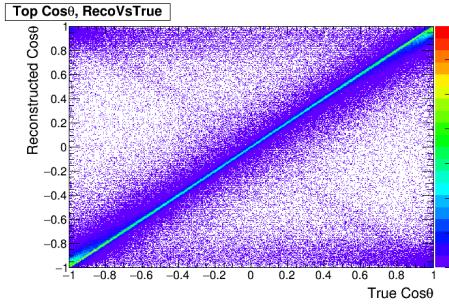
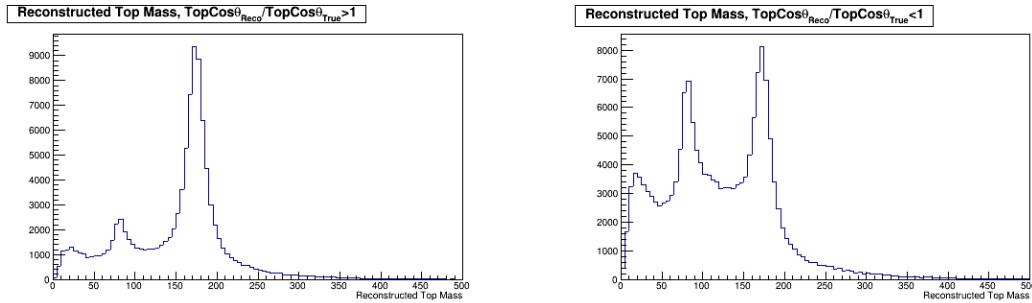


Figure 5.10: Comparison of reconstructed top decay angle to truth. A strong correlation is seen over most of the range, however this starts to break down for large angles of $|Cos\theta| > 0.9$ where non-negligible off diagonal contributions are seen.



(a) Fat jet mass when $|\frac{Cos\theta_{Reco}}{Cos\theta_{True}}| > 1$, on (b) Fat jet mass when $|\frac{Cos\theta_{Reco}}{Cos\theta_{True}}| < 1$, off diagonal regions of fig 5.10.

Figure 5.11: Reconstructed fat jet mass. In the regions where $|\frac{Cos\theta_{Reco}}{Cos\theta_{True}}| > 1$ (upper right and lower left quadrants of fig 5.10) the reconstructed fat jet matches the top mass, while in the regions corresponding to the off diagonal regions of fig 5.10) the mass is not consistent

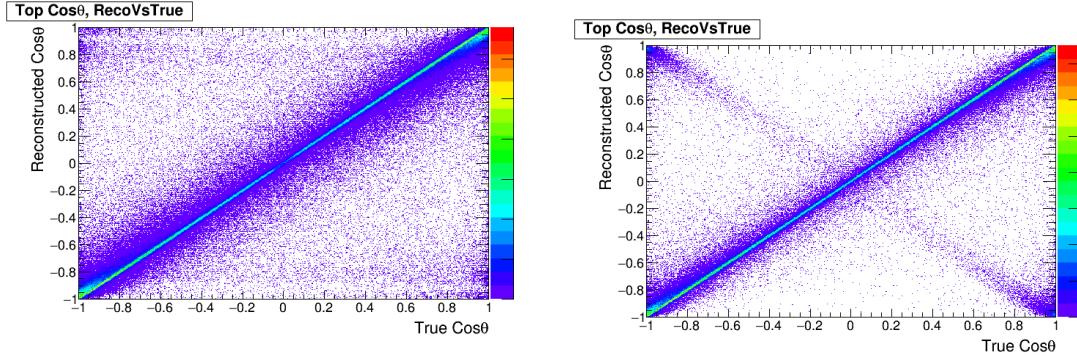
5.2.2.1 Jet Association

After the fat jet finding has been performed, the two reconstructed jets must then be associated as either coming from the hadronically decaying top or from the b jet from the leptonically decaying top. The default method for this was to associate the highest energy fat jet to the hadronically decaying top as, due to the neutrino not being reconstructed and the lepton already being removed, the remaining decay products from the leptonically decaying top should typically have considerably less energy. The performance of this method can be examined by comparing the reconstructed decay angle relative to the true value(see 5.10). While the performance over most of the range studied is good, for $|Cos\theta| > 0.9$ the correlation between the true and reconstructed angles breaks down and off diagonal elements start to appear.

Performance in these forward regions is typically poor due to detector acceptances which result in losses down the beam line. In cases where parts of the hadronic top decay are not able to be reconstructed, using the fat jets energy to perform the jet association no longer becomes a reliable method. Evidence that misreconstruction is the source of these off diagonal elements is presented in figure 5.11 where it is clear that the fat jets in the off diagonal regions are not reconstructed with a consistent mass. When the jets are not fully reconstructed, it is more likely that the wrong jet is assigned to be from the hadronic top. When the wrong jet is selected the reconstructed angle will be approximately π radians off the true value as the tops are predominantly produced back to back. This explanation is further supported by the results shown in figure 5.12a which show that the off diagonal elements can be removed when a cut is placed on the angle between the reconstructed top and the b jet from the leptonic top decay indicating that these elements are definitely coming from selecting the wrong jet. As well as the π radian flips from selecting the wrong jet, there are also additional off diagonal contributions seen which arise from poor reconstruction of the fat jets. This typically happens when the tops are not produced back to back due to ISR/Beamstrahlung. When this happens, during the fat jet reconstruction it is possible for contributions from both true fat jets to be mixed e.g instead of grouping the 3 jets from the hadronic jets together only two of them are grouped together and the third is grouped with the lone bjet from the leptonic top. When this mismatching happens the hadronic top is no longer fully reconstructed and so the angle measured for the top decay has little correlation with the true value. Figure 5.12b shows that these remaining off diagonal elements disappear when a cut is placed on the separation of the tops at truth level.

In order to avoid the problems close to the beam line multiple alternative jet association methods were devised- see 5.3.

The relative effectiveness of these methods were evaluated in three ways shown in figures 5.13, 5.14 and 5.15 respectively. The first method was to look at the overall distribution of $\text{Cos}\theta$ produced by each method compared to the distribution at truth level as this is what will be used to extract A_{fb}^t . All the methods agree well with



(a) Cut placed on angle between reconstructed top and true b jet from leptonic decay, $\Delta\text{Cos}\theta_{\text{Reco}-\text{Bjet}} > 0.1$
(b) Cut placed on collinearity between top pair at truth level, separation < 3 radians

Figure 5.12: Reconstructed vs true top decay angles with truth level cuts to explain the off diagonal elements seen in 5.10.

Fat Jet Selection Method	Description
Lepton	The hadronically decaying top is deemed to be the fat jet with the greatest separation from the isolated lepton
B tag	The hadronically decaying top is deemed to be the fat jet with the greatest separation from the jet with the highest b tag (see 5.2.4 for details on how flavour tagging is performed)
Energy	Select the fat jet with the highest energy to be the hadronically decaying top
Multiplicity	Recluster both fat jets into N “micro jets” (see 5.4.3 for methodology) The hadronically decaying top should have a higher number of micro jets found within it
Mass	The hadronically decaying top is deemed to be the fat jet with the greatest mass
Top Mass	Select the fat jet whose mass is closest to the top mass as the hadronically decaying top
Democratic	A combination of the lepton, energy and mass methods. Each method votes for which fat jet it thinks is the hadronically decaying top. The fat jet with the most votes is then selected as the hadronically decaying top

Table 5.3: Methods used for identifying which fat jet corresponds to the hadronically decaying top

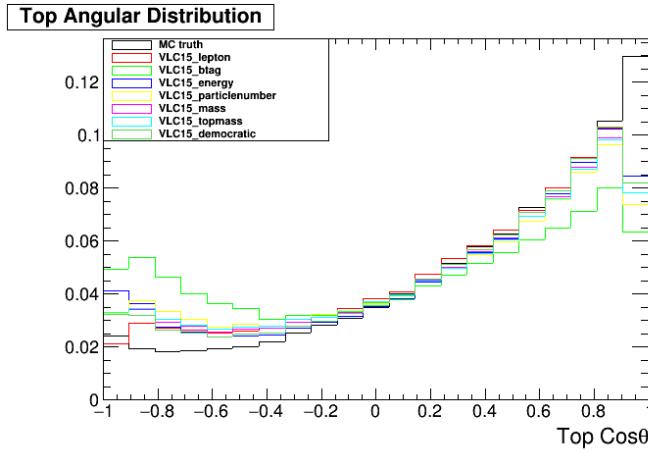


Figure 5.13: Reconstructed $\text{Cos}\theta$ distribution for various jet association methods. The expected distribution from truth level information is unclued for reference

the true distribution in the central region of the detector but diverge in the high $| \text{Cos}\theta |$ region. This is mainly caused by the effect described above. Close to the beam line the jets aren't fully reconstructed, the jet association fails and the b jet from the leptonic side is selected rather than the hadronic top jet. This causes migrations from the forward region to the backward regions producing a deficit in the forward region and an excess in the backward region. Migrations do occur in the opposite direction too for the same reason, however because the top forward-backward asymmetry means that more tops are produced in the forward region to begin with, the net migration is from forward to backward. The migrations are not always a shift of π radians as one might expect. Instead the migrations occur from very close to the beam line to a broader range in the opposite direction. This is due to the fact that ISR/Beamstrahlung can mean the top pair aren't produced exactly back to back in the lab frame and because the b-jet produced by the leptonic decay is not exactly collinear with the top decay axis. Comparing the methods we see that all the methods show similar levels of migration except for the bttag method which shows the highest migration. This is attributed to the fact the highest btagged jet can sometimes be from the hadronic side even in events that are well reconstructed, and so the jet association will fail in more events than the other methods which only fail for events close to the beam line.

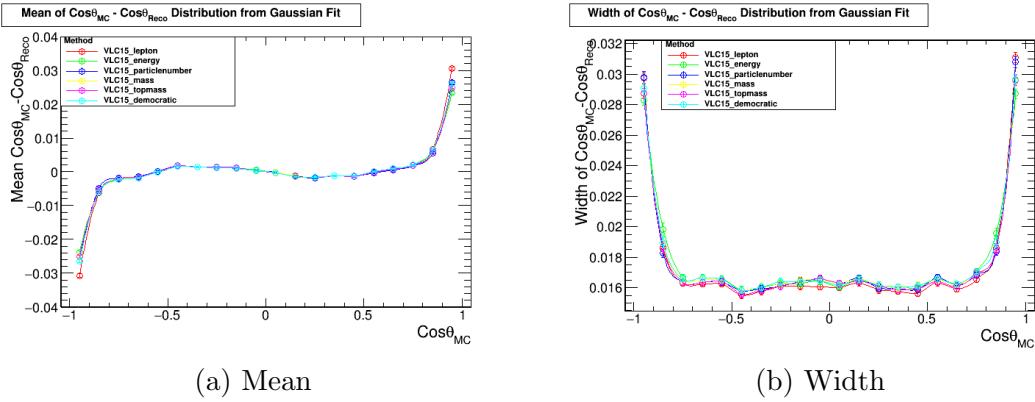


Figure 5.14: Mean and width from fitting $\Delta\text{Cos}\theta_{\text{True-Reco}}$ to a gaussian. Mean: migrations close to $|\text{Cos}\theta| > 0.9$ result in a bias in the mean. The migrations cause a shift of roughly π radians resulting in the bias being in the opposite direction for each end of the range. Width: migrations close to $|\text{Cos}\theta| > 0.9$ cause a broadening in the resolution of the reconstructed $\text{Cos}\theta$

The second method was to measure the difference between the reconstructed and MC(true) $\text{Cos}\theta$ per event and fit this with a gaussian. The variation in the width and mean of these distributions were plotted against the true $\text{Cos}\theta$ and are shown in fig 5.14. The effects of migration at high $\text{Cos}\theta$ is more pronounced in these plots where in the width we can see that the resolution on $\text{Cos}\theta$ gets much worse in the forward regions and the mean shows a pull in opposite directions in these regions proving the migrations do indeed occur in both directions with the same. Unfortunately there is little discrimination seen between the methods except for showing that there are slightly larger migrations when using the btag method.

The final method of comparison was to measure the efficiency with which the hadronic top was measured within the correct $\text{Cos}\theta$ bin as a function of the true $\text{Cos}\theta$. For this study a bin width of 0.1 in $\text{Cos}\theta$ was used. The results are shown in figure 5.15. Here there is a clearer separation in the performance of the different methods. B-tagging is seen to provide the worst efficiency while the energy and democratic methods provide the highest level of performance. The mass based selections provide slightly lower performance than the energy/democratic methods. This is likely explained by the fact they are less robust when the jets are not fully reconstructed. Missing a small section of the jet via acceptance losses/reconstruction inefficiencies can have a large impact on the reconstructed

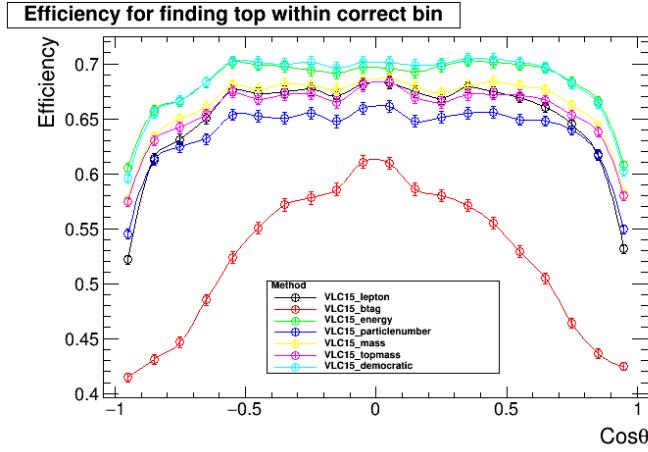


Figure 5.15: Efficiency for reconstructing the hadronically decaying top in the correct $\text{Cos}\theta$ bin

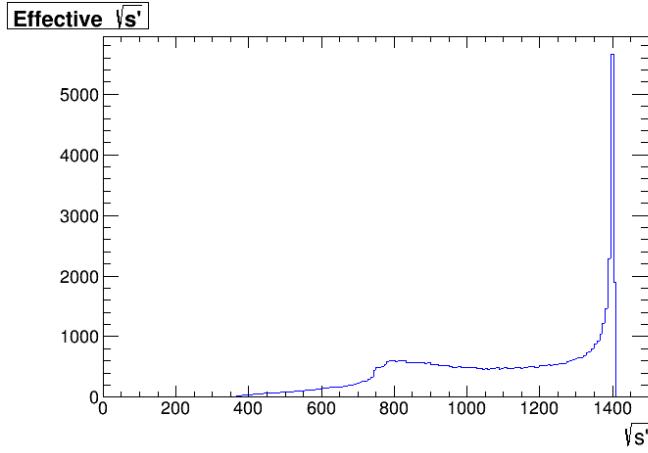


Figure 5.16: Expected s' spectrum for $t\bar{t}$ at 1.4 TeV

mass, however in the case of energy, if we naively assume that the energy is split evenly between the 6 final state particles, then we would expect that the energy of the hadronic fat jet would be three times that of the b-jet from the leptonic top and so considerable energy losses must occur before the wrong jet is selected. Due to their higher bin by bin efficiency, the energy and democratic methods are the best methods to use. Due to it's simplicity the energy method is then chosen as the preferred method for the rest of the analysis.

5.2.3 s' Reconstruction

Following the reconstruction of the lepton and hadronically decaying top it is already possible to calculate A_{FB}^t , however there is still benefit to first reconstructing the effective centre-of-mass energy of the events (along with the neutrino and any photons produced too.) This allows the calculation of A_{FB}^t to be determined in the $t\bar{t}$ rest frame where it is predicted to be up to 50% bigger [?], and also allows a differential measurement of A_{FB}^t to be performed. The expected $\sqrt{s'}$ spectrum for $t\bar{t}$ production at 1.4 TeV is shown in figure 5.16. Here it is seen that there is a large tail to the energy spectrum which can be taken advantage of to measure A_{FB}^t over a large range of energies. This differential measurement provides more greater power for discriminating between different phyhsics models than a single A_{FB}^t measurement. If s' can not be reconstructed per event, A_{FB}^t would have to either be measured as an integral over the full s' range or be measured just around the peak energy where there are only small s' corrections ($E \gtrsim 1200$ GeV), however this would mean disregarding INSERT PERCENTAGE of events produced during the 1.4 TeV run. As well as directly effecting the ways in which we can measure A_{FB}^t , reconstructing s' typically involves reconstructing the neutrino and photon contributions in the event. Having the information about these objects provides further information that can be used to reconstruct the leptonic top and help distinguish signal events from similar backgrounds.

In order to reconstruct s' , multiple methods were attempted with varying complexity:

5.2.3.1 Transverse/Longitudinal Association

The simplest method attempted was to assume that all missing momentum in the transverse direction is attributed to the neutrino, while all longitudinal momentum comes from photon contributions. These assumptions are motivated by the results from figure 5.17 which show that the photons produced are predominantly collinear



Figure 5.17: Angular energy distribution of initial state photons

with the beam.

This allows full reconstruction of the neutrino and photon objects and so s' can be reconstructed by looking at the sum of the energy of the reocnstructed objects (excluding the photon as it is from ISR or Beamstrahlung.) A comparison of the reconstructed s' to the true s' spectrum is shown in figure 5.18

5.2.3.2 Analytic Mass Constraint

The second method attempted is an adaptation of the first method that makes use of the high efficiency with which the lepton is reconstructed to improve the performance. It starts in the same manor by assuming that all transverse missing momentum comes from the neutrino, however the missing longitudinal momentum is then divided beween the neutrino and photon. This is done by constraining the z component of the neutrino momentum by insisting the combination of the lepton



Figure 5.18: Reconstructed s' vs true s' for Transverse/Longitudinal Association Method



Figure 5.19: Reconstructed s' vs true s' for mass constraint method



Figure 5.20: Mass of reconstructed top when using mass constraint method



Figure 5.21: Collinearity of $t\bar{t}$ pair

and neutrino four momenta reproduces the W mass. The details of the necessary calculations are described in more detail in (PUT AN APPENDIX IN), but the key detail is that there are four possible solutions for the neutrino momentum. To decide the most suitable solution the W is combined with the b jet and the solution found to give an invariant mass closest to the top mass was chosen to be best. The resulting reconstructed top mass and s' reconstruction performance are shown in figures 5.19 and 5.20. The reconstructed s' is seen to agree well with the true value for energies over $\tilde{900}$ GeV, however below this the performance is considerably worse.



Figure 5.22: Reconstructed s' vs true s' for collinearity method

5.2.3.3 Collinearity

5.2.3.4 Kinematic Fitting

5.2.4 Flavour Tagging

Flavour tagging was performed using LCFIPlus v00-05-02[35]. LCFIPlus makes use of three BDTs dedicated to searching for u/d/s (light), b and c quarks respectively, to provide a b-tag and c-tag indicating the probability of a jet containing a b or c quark. As the signal process contains two b jets, only the results of the b-tag are considered here. The BDTs were trained using 50,000 $ee \rightarrow Z\nu\nu, Z \rightarrow qq$ events each. The base performance of the BDTs was assessed using a further 150,000 $ee \rightarrow Z\nu\nu, Z \rightarrow qq$ events containing an even mixture of bb, cc and light quarks to measure the efficiency and purity that could be obtained. The results of this test (shown in 5.23) indicate that in the case of $Z \rightarrow qq$ events high efficiencies and

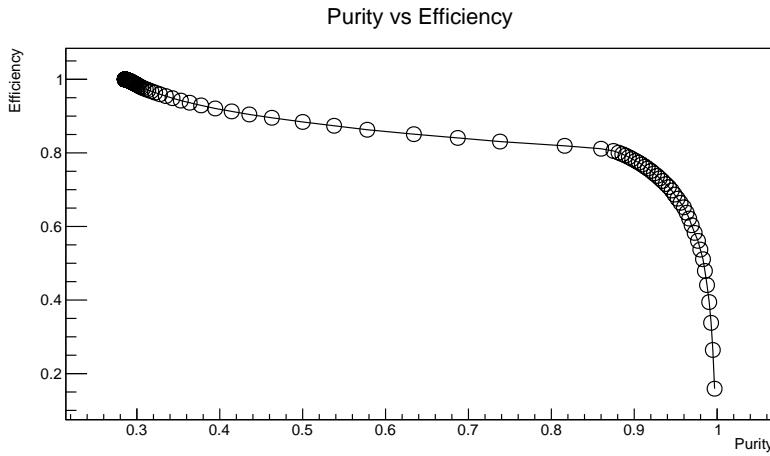


Figure 5.23: Purity vs efficiency for identifying b-jets, obtained from a sample of $Z \rightarrow \text{light}$, c and b quark events simulated at $\sqrt{s} = 1.4$ TeV

puritys of $\tilde{85}\%$ can be achieved simultaneously. Before we apply the flavour tagging to our analysis we first recluster our events into four jets to try and capture the bjets separately from the light quark jets. This is done within the LCFIPlus package which uses the Durham algorithm by default. Ideally the BDTs would also be retrained using top events rather than Z, however due to limited sample sizes this was not a realistic option. The performance of the btagging for semileptonic top events was evaluated by comparing the highest and second highest b-tags assigned to any of the four jets in signal events to those in backgrounds. The results of this comparison are seen in figure 5.24. It is clear that the btagging is consistently successful in finding the first b jet, but is less reliable for finding a second b jet. This is expected due to the topology of the event. The b jet produced by the leptonically decaying top should be well isolated from everything but the lepton which is identified and removed with high efficiency whereas the bjet from hadronically decaying top will be close to two other jets meaning the jet finder is less likely to accurately associate the PFOs in that region to the correct initial quark. As a result the b jet from the leptonic side should be consistently reconstructed and tagged whereas the the b jet on the hadronic side will be less consistent as the efficiency for reconstructing the jet correctly is much lower. Despite the poorer performance of the second highest b-tag, both variables provide clear potential as discriminating variables for removing background.

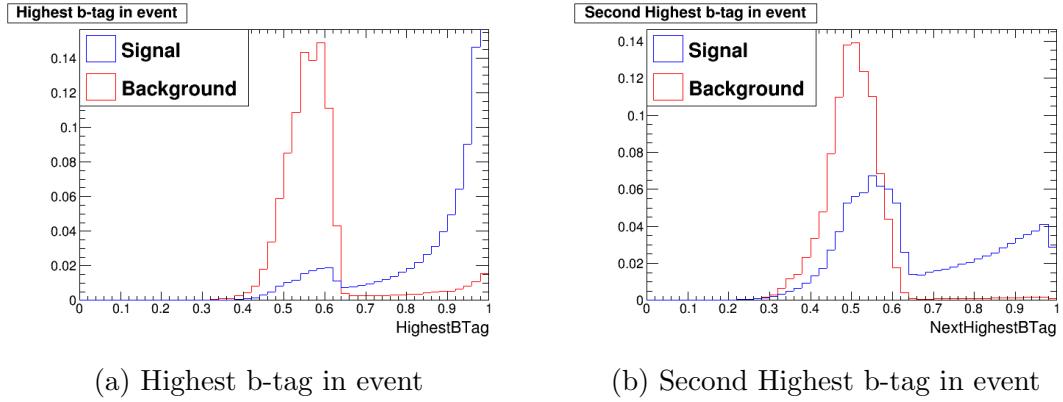


Figure 5.24: B-Tagging performance

5.3 Calculating AFB

Measuring AFB from theta distributions vs counting- benefit that theta fit isn't effected by acceptance cutoff Precision expected from just signal, no event selection, after preselection, after bdt

5.4 Event Selection

Split into two regions- low and high energy due to different topology from s'

Using bdt based selection to maximise performance

Describe jet substructure variables (or any other confusing ones...)

5.4.1 NSubjettiness

5.4.2 Subjet Angular Distributions

5.4.3 Jet Multiplicity

5.5 Quality Cuts

Preselection cuts BDT variables Acceptance cuts

Large table showing efficiency of all cuts on all samples and the final expected number of events for full luminosity

Overall signal efficiency/purity/significance

5.5.1 Differential Measurement

Differential version of at least one of the plots- probably just signal before selection and then signal+background after selection with background subtraction. Explain that most precise point will be for $s' \gtrsim 1200$ as there is the cleanest signal- highly boosted jets. Suggest that lower s' results could benefit from a different reconstruction approach based on a resolved 4 jet analysis- probably deserves its own dedicated study though we can do if we have time. Would also likely benefit from having different variables used for the low s' BDT.

5.6 Conclusions

We measured the Afb to x precision....

CHAPTER 6

DECALStudies

General concept- energy proportional to nParticles

Explain potential benefits- cheape, same tech as inner detectors, granularity could improve particle flow

Description of Mokka & Geant4

Detector changes implemented for simulations- smaller pixel size, active silicon replaced with thinner active silicon layer and passive silicon layer

6.1 DigiMAPs

Various effects on resolution Charge Spread Background Noise Signal Noise Clustering Threshold

6.2 Design Optimization

Aspect Ratio for resolution Linearity Leakage= do studies based on change in radiator- tungsten vs lead

6.3 Hardware studies

Assuming this ever works....

CHAPTER 7

Conclusion

REFERENCES

- [1] M. J. Boland *et al.*, “Updated baseline for a staged Compact Linear Collider,” 2016.
- [2] T. Abe *et al.*, “The International Large Detector: Letter of Intent,” 2010.
- [3] S. Dawson *et al.*, “Working Group Report: Higgs Boson,” in *Proceedings, 2013 Community Summer Study on the Future of U.S. Particle Physics: Snowmass on the Mississippi (CSS2013): Minneapolis, MN, USA, July 29-August 6, 2013*, 2013.
- [4] G. Aad *et al.*, “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC,” *Phys. Lett. B*, 2012.
- [5] S. Chatrchyan *et al.*, “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC,” *Phys. Lett.*, vol. B716, pp. 30–61, 2012.
- [6] V. Trimble, “Existence and Nature of Dark Matter in the Universe,” *Ann. Rev. Astron. Astrophys.*, vol. 25, pp. 425–472, 1987.
- [7] M. Bogomilov *et al.*, “Lattice design and expected performance of the Muon Ionization Cooling Experiment demonstration of ionization cooling,” *Phys. Rev. Accel. Beams*, vol. 20, no. 6, p. 063501, 2017.
- [8] T. Behnke, J. E. Brau, B. Foster, J. Fuster, M. Harrison, J. M. Paterson, M. Peskin, M. Stanitzki, N. Walker, and H. Yamamoto, “The International Linear Collider Technical Design Report - Volume 1: Executive Summary,” 2013.
- [9] R. Lipton, “Muon Collider: Plans, Progress and Challenges,” in *Particles and fields. Proceedings, Meeting of the Division of the American Physical Society, DPF 2011, Providence, USA, August 9-13, 2011*, 2012.

- [10] M. Koratzinos, “FCC-ee accelerator parameters, performance and limitations,” in *International Conference on High Energy Physics 2014 (ICHEP 2014) Valencia, Spain, July 2-9, 2014*, 2014.
- [11] T. Behnke, J. E. Brau, B. Foster, J. Fuster, M. Harrison, J. M. Paterson, M. Peskin, M. Stanitzki, N. Walker, and H. Yamamoto, “The International Linear Collider Technical Design Report - Volume 1: Executive Summary,” *arXiv:1306.6327 [physics.acc-ph]*, 2013.
- [12] P. Lebrun, L. Linssen, A. Lucaci-Timoce, D. Schulte, F. Simon, S. Stapnes, N. Toge, H. Weerts, and J. Wells, “The CLIC Programme: Towards a Staged e+e- Linear Collider Exploring the Terascale : CLIC Conceptual Design Report,” *arXiv:1209.2543 [physics.ins-det]*, Sept. 2012.
- [13] W. Kilian, T. Ohl, and J. Reuter, “WHIZARD: Simulating Multi-Particle Processes at LHC and ILC,” *Eur. Phys. J.*, vol. C71, p. 1742, 2011.
- [14] S. Jadach, J. H. Kuhn, and Z. Was, “TAUOLA: A Library of Monte Carlo programs to simulate decays of polarized tau leptons,” *Comput. Phys. Commun.*, vol. 64, pp. 275–299, 1990.
- [15] T. Sjostrand, S. Mrenna, and P. Z. Skands, “PYTHIA 6.4 Physics and Manual,” *JHEP*, vol. 05, p. 026, 2006.
- [16] 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*, pp. 623–627, 2002.
- [17] M. Thomson, “Particle flow calorimetry and the pandorapfa algorithm,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 611, no. 1, pp. 25 – 40, 2009.
- [18] J. A. Ballin, R. Coath, J. P. Crooks, P. D. Dauncey, A.-M. Magnan, Y. Mikami, O. D. Miller, M. Noy, V. Rajovic, M. Stanitzki, K. D. Stefanov, R. Turchetta, M. Tyndel, E. G. Villani, N. K. Watson, J. A. Wilson, and Z. Zhang, “Design and performance of a CMOS study sensor for a binary readout electromagnetic calorimeter,” *Journal of Instrumentation*, vol. 6, p. 5009, May 2011.
- [19] F. Pitters, “The CLIC Detector Concept,” 2018.
- [20] W. contributors, “Standard model — wikipedia, the free encyclopedia,” 2018. [Online; accessed 8-February-2018].
- [21] M. Szleper, “The Higgs boson and the physics of WW scattering before and after Higgs discovery,” 2014.

- [22] P. W. Higgs, “Broken symmetries and the masses of gauge bosons,” *Phys. Rev. Lett.*, vol. 13, pp. 508–509, Oct 1964.
- [23] F. Englert and R. Brout, “Broken symmetry and the mass of gauge vector mesons,” *Phys. Rev. Lett.*, vol. 13, pp. 321–323, Aug 1964.
- [24] H. Abramowicz *et al.*, “Higgs physics at the CLIC electronpositron linear collider,” *Eur. Phys. J.*, vol. C77, no. 7, p. 475, 2017.
- [25] C. Durig, K. Fujii, J. List, and J. Tian, “Model Independent Determination of HWW coupling and Higgs total width at ILC,” in *International Workshop on Future Linear Colliders (LCWS13) Tokyo, Japan, November 11-15, 2013*, 2014.
- [26] “Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV,” Tech. Rep. ATLAS-CONF-2015-044, CERN, Geneva, Sep 2015.
- [27] “Projected Performance of an Upgraded CMS Detector at the LHC and HL-LHC: Contribution to the Snowmass Process,” in *Proceedings, 2013 Community Summer Study on the Future of U.S. Particle Physics: Snowmass on the Mississippi (CSS2013): Minneapolis, MN, USA, July 29-August 6, 2013*, 2013.
- [28] P. Langacker, “The Physics of Heavy Z' Gauge Bosons,” *Rev. Mod. Phys.*, vol. 81, pp. 1199–1228, 2009.
- [29] M. S. Amjad *et al.*, “A precise characterisation of the top quark electro-weak vertices at the ILC,” *Eur. Phys. J.*, vol. C75, no. 10, p. 512, 2015.
- [30] C. R. Schmidt, “Top quark production and decay at next-to-leading order in $e^+ e^-$ annihilation,” *Phys. Rev.*, vol. D54, pp. 3250–3265, 1996.
- [31] F. Simon, “Higgs Physics at future Linear Colliders - A Case for precise Vertexing,” *PoS*, vol. Vertex2013, p. 019, 2014.
- [32] “Monte carlo samples for clic higgs studies.”
- [33] D. Schulte, “Beam-Beam Simulations with GUINEA-PIG,” Mar 1999.
- [34] M. Cacciari, G. P. Salam, and G. Soyez, “FastJet User Manual,” *Eur. Phys. J.*, vol. C72, p. 1896, 2012.
- [35] T. Suehara and T. Tanabe, “LCFIPlus: A Framework for Jet Analysis in Linear Collider Studies,” *Nucl. Instrum. Meth.*, vol. A808, pp. 109–116, 2016.
- [36] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, H. Voss, M. Backes, T. Carli, O. Cohen, A. Christov, D. Dannheim, K. Danielowski, S. Henrot-Versille, M. Jachowski, K. Kraszewski, A. Krasznahorkay, Jr., M. Kruk, Y. Mahalalel, R. Ospanov, X. Prudent, A. Robert, D. Schouten, F. Tegenfeldt, A. Voigt, K. Voss, M. Wolter, and A. Zemla, “TMVA - Toolkit for Multivariate Data Analysis,” *ArXiv Physics e-prints*, Mar. 2007.

- [37] A. Muennich, “TauFinder: A Reconstruction Algorithm for Tau Leptons at Linear Colliders,” Oct 2010.
- [38] K. A. Olive *et al.*, “Review of Particle Physics,” *Chin. Phys.*, vol. C38, p. 090001, 2014.