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.

October 12, 2017

ABSTRACT

X was measured, we showed that $Y \neq Z$ and that $M_{\mathrm{H}} = 126$ 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

List of Tables

List of Figures

DEFINITIONS OF ACRONYMS

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

CMOS Complimentary Metal-Oxide Semiconductor

MAPS Monlithic Active Pixel Sensors

The Higgs discovery is cited [1]

CHAPTER 1

INTRODUCTION

Following the discovery of a Higgs Boson at the Large Hadron Collider (LHC) [1, 2], with properties in agreement with those predicted by the Standard Model (SM), the particle physics community is left in a situation where there is no definitive course of action through which new phsyics might be discovered. There are many open questions remaining such as the nature of dark matter or the origins of matter-antimatter asymmetry, but no clear direction for how to answer them. As such there are two main approaches that may be taken- the first would be to continue to push the boundaries of the "energy frontier" (following the approach of the LHC) and look for new physics at higher energy scales that is not predicted by the SM but is predicted by many Beyond the Standard Model (BSM) models such as super-symmetry. The second option is to advance in the "precision frontier" to search for small deviations from the SM and harder to detect processes. Both approaches come with their own advantages and disadvantages. Going to higher energies is more likely to allow direct detection of new particles and is supported by the fact that the

majority of BSM models predict that new physics effects should exist at energies beyond what has been explored so far, however the problem is that the scale at which new physics should appear is unknown. This makes designing a future high energy collider difficult as without a clear idea of what energy is needed it is possible the collision energy will be below the new physics scale and so no new phenomenum will be observed. Pushing the precision frontier has the drawback that it may only serve to reinforce confidence in the SM and has a smaller chance of direct discovery of new physics, however even in this worst case scenario it will still provide precise measurements of SM properties which are beneficial both for constraining BSM models and for reducing uncertainties on future measurements at high energy colliders.

The choice in which the community decides to proceed will greatly influence the design options for the next generation of particle colliders. For high energy measurments the obvious choice will be a circular hadron collider similar to the LHC. The circular design allows for particles to be accelerated over as large a time as necessary without the need for a large collider which facilitates reaching higher energies. It also allows for collisions to occur at multiple sites simultaneously allowing for more experiments to take place at the facility. For a circular design, hadrons (most likely protons due to their stability and charge) are needed as their high mass minimises energy loss from synchotron radiation. This is radiation produced by a charged particle undergoing a transverse acceleration and is described by equation (1.1):

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

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.

While protons allow for high energy collisions, they are not suitable for precision measurements as their composite nature means that the initial four monentum and quantum numbers of the collision cannot be known. This means that all information must be extracted purely from a collisions decay products which are subject to uncertainties from detector resolutions/acceptances and the decay products visibility

e.g. neutrinos/dark matter cannot be detected. For precision physics the better choice is to collide electron-positron pairs. This is an annihilation reaction where all quantum numbers will cancel out and the initial energy and momentum are only limited by the quality of the colliding beams. This allows conservation laws to be used to infer the properties of missing particles such as neutrinos or new particles. For colliding electrons, a circular collider is no longer feasible as the electrons low mass result in energy being lost through synchotron radiation at 10^{16} times the rate of protons. Instead, lepton colliders are traditionally built as linear colliders. This prevents synchotron radiation occuring but limits the maximum collision energy achievable as the path over which a particle can be accelerated is limited to just one length of the collider and it is expensive and impractical to produce longer colliders. That being said, because a lepton collider provides an annihilation interaction rather than a parton interaction it is possible to build a lepton collider with a lower beam energy but still have a higher average collision energy.

With the High Luminosity Large Hadron Collider (HL-LHC) expected to finish taking data in the 2030s and the long construction times associated with super colliders, a decision on what form the next generation of colliders should take is expected to occur by the early 2020s. 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 propsed as a project for CERN. On the precision side there are multiple proposed projects (REFERENCE SMALLER PROJECTS), however the most mature of these are the linear electron-positron colliders: Compact Linear Collider (CLIC) **??** and International Linear Collider (ILC) **??**. The ILC is a 500 GeV collider proposed by the Japanese government while CLIC is a multi-TeV machine being proposed by CERN. Due to the large cost of these devices it unlikely that CERN will 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 relatvely poorly measured when com-

pared to other standard model particles, while also examining a novel design for a digital electromagnetic calorimeter based on Complimentary Metal-Oxide Semiconductor (CMOS) Monlithic Active Pixel Sensors (MAPS) technology for use in future detectors.

CHAPTER 2

Experiments

Leave description of DECAL concept to DECAL chapter but mention it as one of the ECAL alternatives here

There are many possible designs for future lepton colliders [3, 4] however here we will focus on the two most developed projects, CLIC and ILC. Both projects 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, the **ILD!** (**ILD!**) and the **SiD!** (**SiD!**). Both of these detectors were originally designed for use at ILC as general purpose devices but are currently being adapted for use at CLIC. Because the optimization of the detectors for use in CLIC has only recently begun, the designs for both versions of the detectors are still similar enough that we will neglect describing both here but will instead describe just the ILC versions of both detectors.

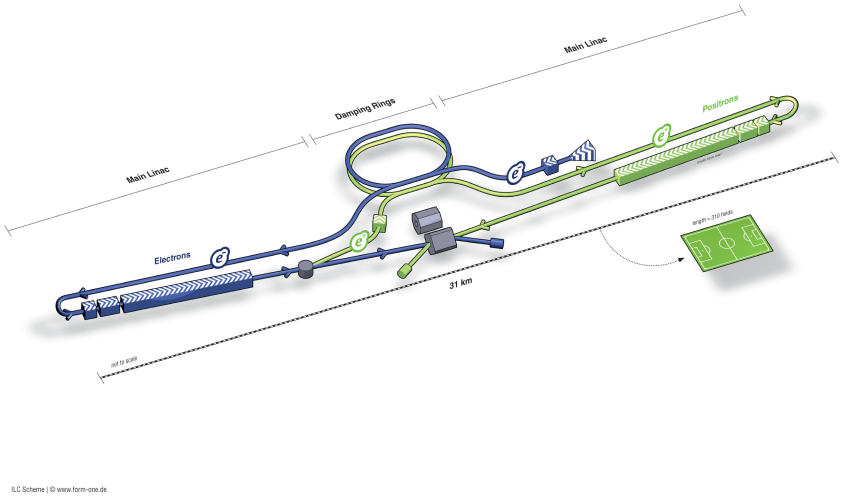

Figure 2.1: The ILC Collider (from [5])

2.1 ILC

The ILC (Figure 3.1) 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 **TDR!** (**TDR!**) [5] 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 is likely to change in the future. 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 experiment; if any

new particles are discovered above 500GeV then the 1TeV upgrade will 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 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 described above. At this energy the experiment will have little sensitivity to the Higgs-WW coupling.

For the following three years, the collider will run at 500GeV and will accrue 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 Yukawa coupling. Outside of the Higgs, 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 the upgrade to 1TeV followed by another three years of running 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. 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 mass of the newly discovered particle.

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, 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 and 350GeV (the top production threshold) to better measure the properties of the top quark.

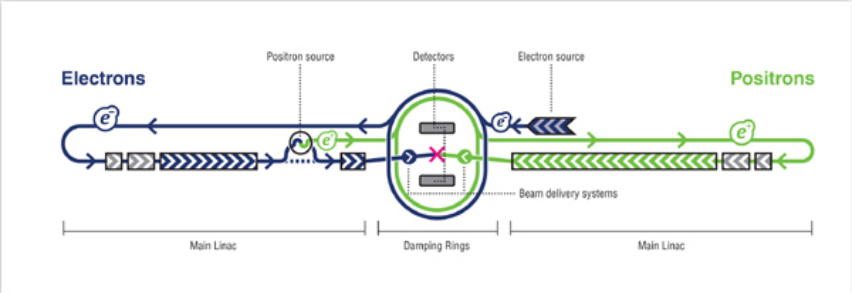

Figure 2.2: A simplified schematic of the ILC

2.1.2 Beam Production, Acceleration and Focusing

A simplified schematic of the ILC accelerating structure is shown above in diagram Figure 3.2. 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 giving a beam pulse length of 730μs. 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 have a large period of time in which to relax after events.As the detectors do not need to be on 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 **RF!** (**RF!**) cavities (see Figure 3.3)

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 **BDS!** (**BDS!**) which primarily acts to compress

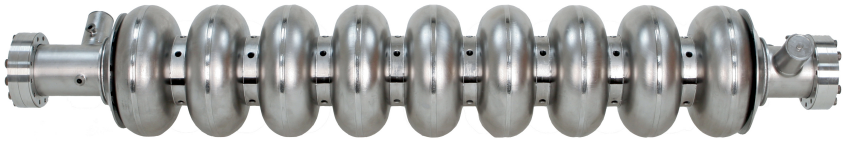

Figure 2.3: A 1.3GHz Superconducting Niobium Radio Frequency Cavity [5]

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 acts to reduce the **ISR!** (**ISR!**) radiation described earlier (see Section **??**) while giving a small enough cross-section that the **IP!** (**IP!**) of the collision can be well known. Following the **BDS!** the beam finally enters the detector and collides with the 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. The resulting photons are then separated from the electron beam and are collided with a Titanium alloy target to produce electron positron pairs. The electrons and positrons are then separated- 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 the electrons underwent before arriving at the **IP!**.

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 3.4.) Despite being named as compact, CLIC is actually longer than the initial 500GeV ILC. The reason for this

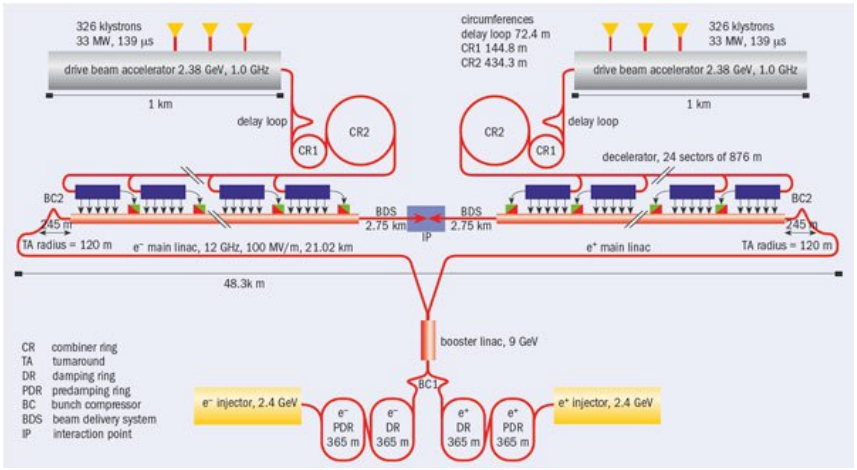

Figure 2.4: The CLIC Collider. Image taken from CLIC Conceptual Desgin Report[6]

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 technology required for CLIC is less developed than for 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 **CDR!** (**CDR!**) [6]. This is a much less detailed document with more preliminary figures primarily aimed at justifying the physics case for why CLIC might be built rather than detailing the intricate design of the machine, though again any figures reported in this section can be assumed to be from this document.

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 to give it a small cross-section but the beam is no longer shaped into a ribbon shape- this is why **ISR!** is a more significant problem 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

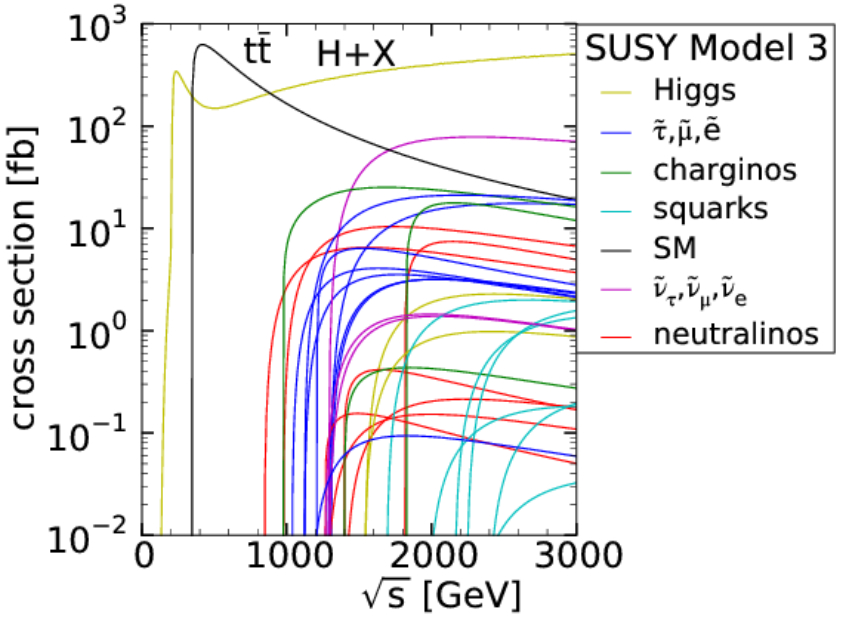

Figure 2.5: Cross sections for production of various super symmetric particles at an e^+e^- collider as a function of centre of mass energy.

detectors less time to relax after events. The most significant differences however are the energy staging and the acceleration technology used at CLIC.

2.2.1 Energy Staging

CLIC will operate at three energy stages- 350GeV, 1.4TeV and 3TeV collecting $500fb^{-1}$, $1.5ab^{-1}$ and $2ab^{-1}$ of data respectively. During the running of the 350GeV energy stage, construction of the 1.4TeV will be carried out (and so on for the 1.4TeV and 3TeV scales) so as to reduce the waiting time between successive energy stages.

The 350GeV energy scale will aim to measure the properties of the Higgs Boson and top quark in a similar manner to the 250GeV and 500GeV energy scales at ILC while the 1.4TeV and 3TeV energy scales will be searching for new BSM physics. The choice of 1.4TeV and 3TeV are based upon the predictions of one version of supersymmetry shown in Figure 3.5.

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 have a large 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 lost energy (emitted as photons) to accelerate the smaller number of electrons in the main beam resulting in a sudden rapid acceleration. The main beam is then used for the collisions. 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 accelerator, which disrupt the alignment and structure of the main beam.

2.3 Detectors

The ILC has been designed with the intention that it will have 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 us to get two independent sets of results for the collider and allows us to still take results when one of the detectors requires maintenance, it

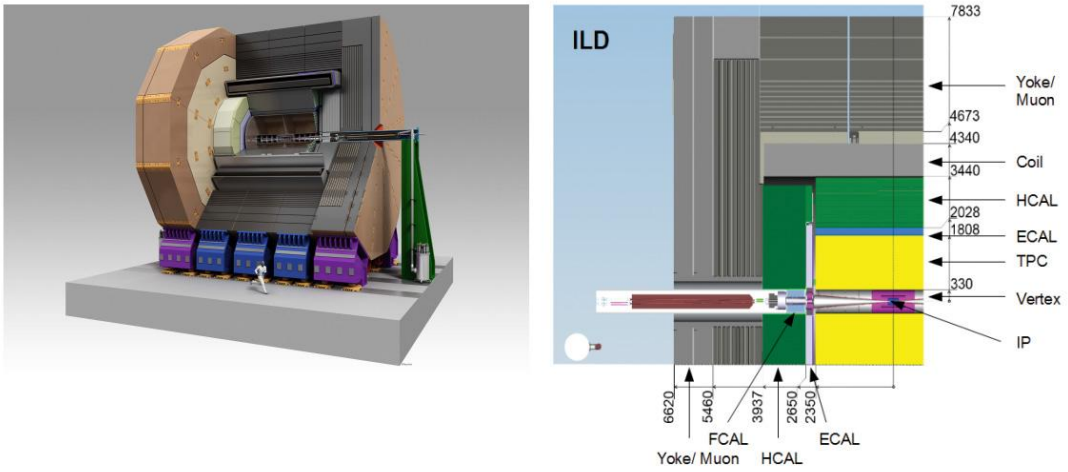

Figure 2.6: The International Large Detector Concept (left). Schematic of the ILD showing the key components in a one-quarter view of a vertical section of the detector (right). [5]

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.

2.3.1 ILD

The ILD (shown in Figure 3.6) 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 and finally muon tail catchers. The detector has two endcaps at each end of the barrel creating a hermetic

seal. Because many of the physics processes at ILC involve Z and W bosons, one of the main performance requirements is the ability to distinguish between the two particles which means achieving a 3-4% uncertainty in the energy of 100GeV jets and a $\delta p/p^2$ (where p is the momentum of a charged particle) of $5 \times 10^{-5}(GeV/c)^{-1}$. All specifications for the detector can be found in the **ILD!** Letter of Intent [7]. Here we will give a brief overview of the key components and their functions.

2.3.1.1 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 **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 very 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 it is expected to consist of either 3 or 5 cylinders of sensors starting at a radius of $r = 15$mm from the beamline.

2.3.1.2 Tracking

Tracking at the **ILD!** uses a **TPC!** (**TPC!**). This is a large gas filled cylinder with an electric field across it and readout electronics at each end of the cylinder. 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. To gain extra precision on the entry and exit points of the **TPC!**, the chamber has two silicon detector layers referred to as the **SIT!** (**SIT!**) and **SET!** (**SET!**) positioned immediately before (r=165mm) and after (r=1833mm) the **TPC!** which give two high spatial resolution points for the entrance and exit points of particles. These high spatial resolution points are particularly useful for reconstructing individual particles within jets using the Pandora **PFA!** (**PFA!**) (see Section 3.4.2.)

2.3.1.3 Calorimetry

The function of calorimeters is to measure the energy of particles. The ILD uses sampling calorimeters which work by having alternating layers of a dense absorber material that destroy the incoming particle causing it to shower into lower energy particles and active sensor materials which collect the low energy particles and convert them into an electrical signal. The calorimeters are split into electromagnetic and hadronic sections in which the absorbing material is chosen to interact mainly with particles through electromagnetic or strong interactions respectively. In practice this means the **ECAL!** (**ECAL!**) mainly detects electrons and photons while the **HCAL!** (**HCAL!**) mainly detects hadrons such as pions. Neutral hadrons are recorded exclusively in the **HCAL!**.

The ILD ECAL is a highly granular calorimeter poitioned at r=1847mm and consists of 30 active layers separated by layers of tungsten which acts as the absorbing material. The structure of the **ECAL!** is shown in Figure 3.7. The choice of active material is yet to be made though the two leading techologies are silicon scintillators or pixels. The scintillator form of the technology uses 10x45mm strips which would be rotated by 90^o in each successive layer to produce an effective cell size of

10x10mm with photomultipliers attached to each strip for readout. This form of the technology is considerably cheaper but has a lower performance and relies on algorithms accurately correlating hits in successive layers to produce the effective 10x10mm cell size. The pixel form of the technology simply uses 5mm or 10mm square silicon pixels directly connected to the readout electonics. This is more expensive but produces more consistent results.

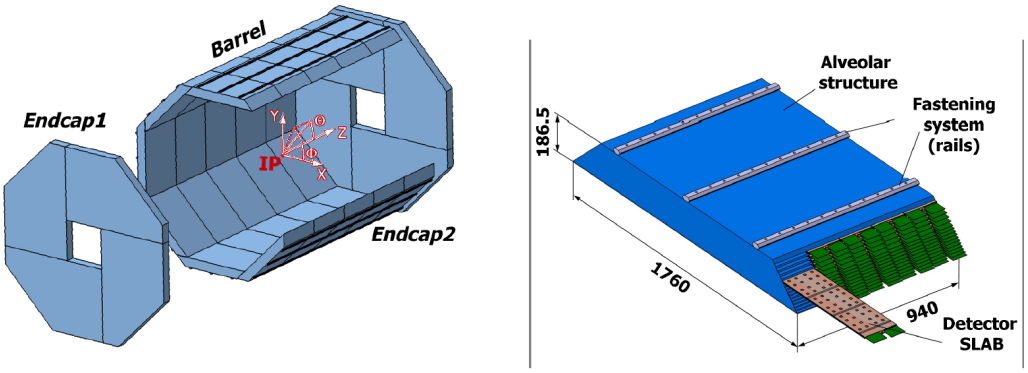

Figure 2.7: The Overall ILD Structure (left) and one individual module (right).The ECAL is made up 40 modules, each containing 30 detector slabs. The modules are combined into groups of 5 referred to as a stave whih extend along the full length of the barrel. There are then 8 of these stave arranged in a circle to create the circumference of the barrel [7].

Later on (see Section **??**) we will discuss our work on developing an alternative form of the silicon pixel technology with a cell size of 50x50μm which acts as a digital machine and purely counts the number of particles absorbed in the 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 [8]. It is expected to be cheaper than the standard silicon pixel tehnology and has already been shown to produce no significant decrease in performance.

The **HCAL!** is immediately outside the ECAL at r=2058mm and has a similar overall modular structure. The HCAL uses stainless steel as an absorbing medium combined with scintillators and Silicon PhotoMultipliers. Again both digital and analogue variations are available with the analogue using 3x3cm cells and the digital 1x1cm. The relative performance of each technology is still being evaluated to see if there is any degredation in the performance when using the digital variation.

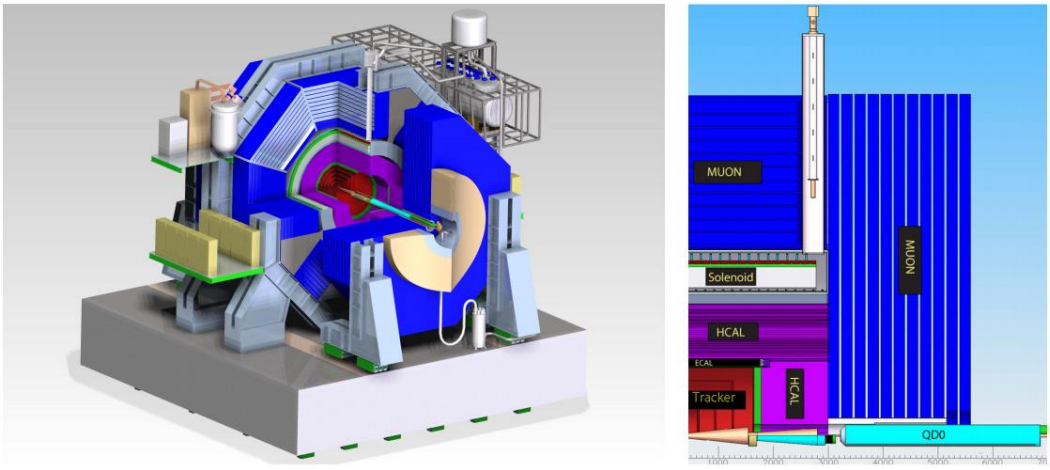

Figure 2.8: The Silicon Detector Concept (left). Schematic of the SiD showing the key components in a one-quarter view of a vertical section of the detector (right). [5]

2.3.1.4 Muon Detection

Muon detection is perhaps the easiest process to perform at the ILC. Because the signal events at the ILC are 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 detetors 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.

2.3.2 SiD

The **SiD!** (Figure 3.8) is overall quite similar to ILD with the two main differences being that SiD uses a stronger 5T magnet and a different form of tracking. Tracking

at the SiD uses silicon strips rather than a TPC. This results in a better performance but is considerably more expensive. As this is the only signifcant difference from ILD we will not go into any further detail about the subdetector technologies used for SiD.

2.4 Linear Collider Analysis Framework

Before discussing the work we have done this year, it is useful to have a brief overview of the software used for linear collider analysis. Both ILC and CLIC use the same core set of packages referred to as ILCSoft. This shared framework makes it easy for work done for one experiment to be ported for use in the other which is beneficial for both experiments as it reduces the amount of work needed for each and helps prevent duplication of efforts. Here we will give an overview of the most commonly used packages in ILCSoft.

2.4.1 Mokka

Mokka is a Geant4 based simulation package available in ILCSoft. Mokka allows users to take ".stdhep" files created from an event generator such as PYTHIA or WHIZARD, simulates how a detector will respond in the presence of the events and outputs info such as calorimeter and tracker hits into a ".slcio" (the standard linear collider file format.) Standard detector geometries are stored in a database online from which Mokka extracts them at runtime. The geometry is modular in design with components such as the ECAL and TPC described separately. Users are able to customize the detector for simulation studies by replacing the standard detector components with an alternative version. The modules can even be customized manually in the Mokka steering file e.g. the number of layers in the ECAL or the thickness of the sensitive layer in the HCAL can be changed. Mokka will then automatically try to resize the customized component to make it fit in the appropriate

space within the larger detector design. As well as accepting ".stdhep" files Mokka also has an inbuilt "particle gun" which can be used to fire particles with any energy at any position within the detector. This tool is particularly useful for studying a particular component of the detector without worrying about the effect of any of the components between it and the **IP!**.

2.4.2 Pandora Particle Flow Algorithm

Pandora is an advanced Particle Flow Algorithm used at linear colliders which allows an increased level of precision from detector measurements by combining information from all the different detector components. The main aim of a **PFA!** is to associate calorimeter hits with individual tracker hits using topology then use the tracker information to calculate the particles' energy, as there is less uncertainty on this than on the measurement by the calorimeter. If there is ambiguity when associating tracks with hits (e.g. in jets) then the **PFA!** will instead use the calorimeter to determine the particles energy. Because we ideally want to use the track measurements for every particle the performance of the **PFA!** is boosted by having a high spatial resolution measurement at the end of the tracker (e.g. the **SET!** in **ILD!**) and having a high granularity calorimeter at the start of the **ECAL!** as this reduces the ambiguity when associating tracks with calroimeter hits. The optimization for **PFA!** has been one the main influencing factors for the design of the detectors at ILC and CLIC. In ILCSoft Pandora is implemented in Marlin (see below) and is used to convert the hits output by Mokka into reconstructed particles referred to as **PFOs!** (**PFOs!**).

2.4.3 Marlin

Marlin is the main analysis tool in ILCSoft. It is a modular framework which applies a series of processors to an event where each processor can be built separately to perform a seperate task. For example one processor might perform jet finding while another might carry out lepton finding. Marlin is controlled with xml steering files

containing a list of the processors to be used, the lcio files on which Marlin will act, the detector geometry used by Mokka and the parameters which the processors need e.g. energy cuts or jet radius. The main benefit of the Marlin system is the ability to "plug and play" by which we mean processors can be added/removed and their parameters can be changed in the Marlin steering file to change the analysis at any time without having to recompile any code. Processors from one analyis can also be used in another analysis simply by including them in the steering file with the idea being that eventually there will be a large database of processors which users can pick and choose from without having to construct their own. The ILCSoft installation already comes with standard processors that link in to FastJet (jet finding), LCFIPLUS (flavour tagging), Root and Pandora.

A typical analyis from start to finish might proceed as follows:

- Generate event with Whizard, output is stdhep file
- Simulate event in detector with Mokka, output is lcio file with raw detector hits
- Apply Pandora PFA using Marlin, output is lcio file containing reconstructed particles
- Perform Analysis using Marlin, output is Root ntuple containing parameters like missing energy, number of jets etc

CHAPTER 3

Theory

Moriond talk should be useful for deciding content here!!

Physics program for colliders- three main areas will be higgs top and BSM physics

3.1 Higgs

Plot of Higgs production cross sections vs energy

3.1.1 HiggsStrahlung

Model independent measurements Mass Width Couplings

3.2 Top

Mass Width AFB EW Couplings

3.3 BSM

Predictions for SUSY

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 is 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.[9]. Here we will consider the measurement of one of these quanties, $\sigma_{H\nu\nu}$ x 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→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→qqlν channel (shown in Fig.4.2), where l=e,μ, with the intention of combining our results with the existing measurement to estimate the overall precision acheivable 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 CLIC_ILD_CDR, a variant of the ILD detector designed for ILC [6]. The main backgrounds of note are: the ee→qqlν 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→ H(WW* →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.

[i]Sample 2022 is the ee→H$\nu\nu$ inclusive sample. The relevant events were extracted from this main sample

[ii]The cross-section for these events were scaled by a factor of 2 to account for interactions

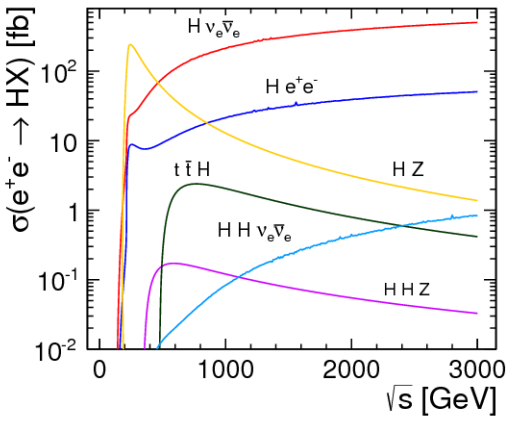

Figure 4.1: Cross sections for Higgs production mechanisms in an e^+e^- collider as a function of centre-of-mass energy [10]. For energies above 500GeV, Higgs production is dominated by the WW-fusion process(red.)

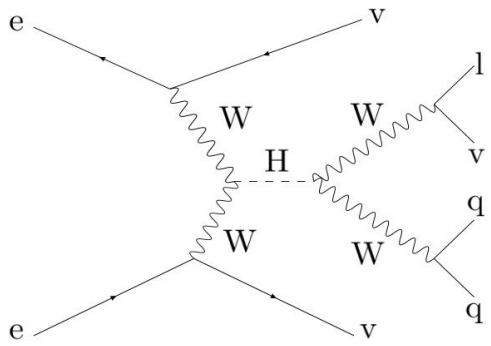

Figure 4.2: Higgs production via WW fusion with the Higgs decaying to a WW pair, which then decays semileptonically.

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

occuring with both the electron and positron. In the case of beamsstrahlung events simulated using GUINEA-PIG [12], a further scaling of 0.75 was applied to account for the lower luminosity of $e\gamma$ events

Process	Cross Section(fb)	Production ID[11]	Events Used
Signal: ee→ H(WW* →qqlν)$\nu\nu$	18.9	2022[i]	70000
ee→ H(WW* →qqqq)$\nu\nu$	25.6	2022[i]	140000
ee→ H* → Other	199.6	2022[i]	750000
ee→qq	4009.5	2091	500000
ee→qqqq	1328.1	2163	300000
eγ→eqq (γ from EPA)	32308	2515[ii]	500000
γe→eqq (γ from BS)	56043	2527[ii]	500000
ee→qq$\nu\nu$	787.7	3243	500000
ee→qqll	2725.8	3246	400000
ee→qqlν	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 [13]) 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 identifiying 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 [14] 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 [15]. Three neural nets were trained to identify u/d/s, b and c quarks respectively with 50,000 ee→ Z$\nu\nu$, Z→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→ 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 [16] 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 ¡591GeV, Mass of the W boson ¡231GeV, Mass of the reconstructed Higgs ¡306GeV and 667GeV¡ total missing energy ¡1400GeV. 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 WW→qqqq final state, as expected. By neglecting the case where the isolated lepton it a τ, we have reduced our statistics to two thirds that of the hadronic channel which inherently limits the precision that can be acheived. 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 ee→ H(→ other)$\nu\nu$ and ee→qqlν 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 H→other events it was determined that 26% of the remaining events came from H→ $\tau^+\tau^-$ processes with a further 25% from H→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 [17]. However, the number of τs misidentified in the signal channel was determined to be too high to veto the τ events without significantly increasing the overall statisitcal uncertainty on σ.Br and so τ identification is not used in the final selection. The efficiency for selecting WW* →qqqq events in the WW* →qqlν 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 H→WW→qqlν events in the H→WW→qqqq channel has yet to be confirmed, therefore a final combined result has not yet been

performed. Where H→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→qqlν branching ratio of 1.1% [18].

4.5 Conclusion

In summary, we have performed a full analysis of the ee→ H(WW*)$\nu\nu$, WW* →qqlν decay channel using a large set of backgrounds with the aim of measuring the H→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→ H(→ Other)$\nu\nu$ and ee→qqlν in the final selection. The resulting statistical uncertainty was found to be:

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

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

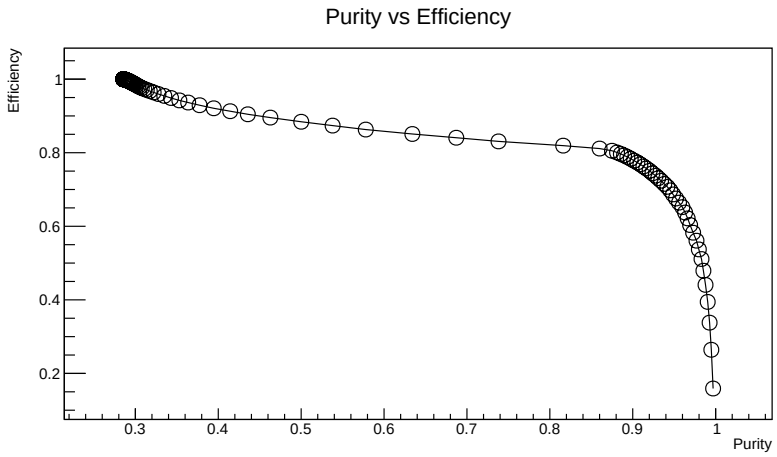

Figure 4.4: Purity vs efficiency for identifying b-jets, obtained from a sample of Z→ light, c and b quark events simulated at $\sqrt{s}$ =1.4TeV

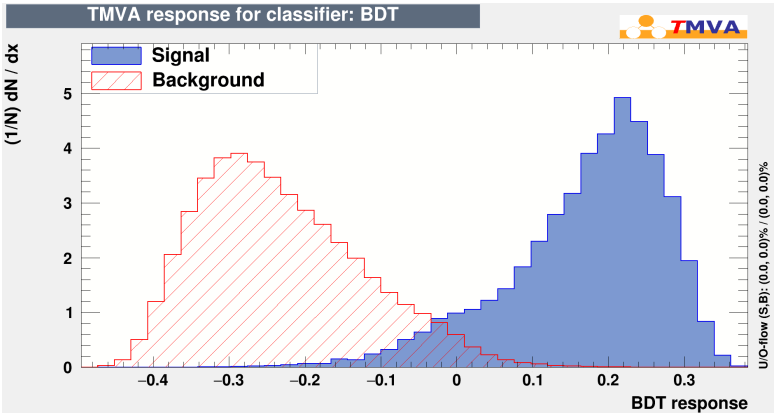

Figure 4.5: BDT response for signal and background events after TMVA classification

Process	Cross Section(fb)	Pre-selection Eff(%)	BDT Cut Eff(%)	Events After
Signal	18.9	99.99	42.65	
ee→ H(WW* →qqqq)$\nu\nu$	25.6	99.96	1.79	
ee→ H(→ Other)$\nu\nu$	199.6	99.62	1.26	
ee→qq	4009.5	76.95	¡0.01	
ee→qqqq	1328.1	36.03	¡0.01	
eγ→eqq (γ from EPA)	32308	67.00	¡0.01	
γe→eqq (γ from BS)	56043	95.84	¡0.01	
ee→qq$\nu\nu$	787.7	96.59	0.07	
ee→qqll	2725.8	89.75	¡0.01	
ee→qqlν	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 Tagging

Generic top tagger, explain overall approach of using 6 BDTs

Overall analysis method

If we end up with cut on the energy, start with full energy range analysis to highlight problems

Explain it will be tested in terms of AFB. Explain importance of AFB measurement

5.1 lepton finding

Method- based on using Pandora ID and jet isolation

Lepton finding efficiency Asymmetry Gamma correction for electrons Charge tagging efficiency QbyP

5.2 jet finding

Concept of fat jets reconstruction efficiency in terms of theta Compare alogrithms in terms of top mass and angle for reconstruction Compare methods picking the correct

5.2.1 fat jet properties

Multiplicity- systematic uncertainty Angular relations nSubjettiness

5.3 ISR effects

Photon energy vs angle Inability to reconstruct leptonic top Boost introduced to $t\bar{t}$ system Imact on reconstrcuted angle

5.4 event selection

No need to re-explain BDTs but should explain why and how we use 6 bdts Final selection results for a particular working point

5.5 AFB determination

Methods for calculating it and variation between them- Effect of ISR Effect of BDT selection Background modelling

Potentially adding cuts on acceptance- need to add in extrapolation??

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] 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.

[2] 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.

[3] 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.

[4] 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.

[5] 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.

[6] 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.

[7] T. Abe *et al.*, "The International Large Detector: Letter of Intent," 2010.

[8] 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.

[9] 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.

[10] F. Simon, "Higgs Physics at future Linear Colliders - A Case for precise Vertexing," *PoS*, vol. Vertex2013, p. 019, 2014.

[11] "Monte carlo samples for clic higgs studies."

[12] D. Schulte, "Beam-Beam Simulations with GUINEA-PIG," Mar 1999.

[13] 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.

[14] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet User Manual," *Eur. Phys. J.*, vol. C72, p. 1896, 2012.

[15] T. Suehara and T. Tanabe, "LCFIPlus: A Framework for Jet Analysis in Linear Collider Studies," *Nucl. Instrum. Meth.*, vol. A808, pp. 109–116, 2016.

[16] 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.

[17] A. Muennich, "TauFinder: A Reconstruction Algorithm for Tau Leptons at Linear Colliders," Oct 2010.

[18] K. A. Olive *et al.*, "Review of Particle Physics," *Chin. Phys.*, vol. C38, p. 090001, 2014.

APPENDIX A

FIRST APPENDIX

Tables of datapoints...