Vrije Universiteit Brussel

The top quark pair cross-section at the threshold at the FCC-ee

Author:
Mathias Mancini
Academic year:
2018-2019

Promotor:
Freya Blekman

Proefschrift ingediend met het oog op het behalen van de academische graad Master in de Fysica en Astronomie

June 7, 2019

Abstract

An analysis study of the the semi-leptonic $t\bar{t}$ process at the FCC-ee is performed. Samples for $t\bar{t}$ events and backgrounds were simulated including the detector response, using a fast simulation approach and subjected to a succession of kinematic cuts to obtain data consisting primarily of signal. A selection efficiency of 10 % and 81 % purity was chosen after selection. The selection process was studied in function of $\sqrt{s}$ and the simulated top quark mass. The efficiency to select $t\bar{t}$ events in the semi-leptonic channel is relatively independent of these parameters and around 10 %. The statistical uncertainty on the cross-section improves with higher $\sqrt{s}$ and stabilises in the continuum. The analysis aims to contribute to the effort of optimising the FCC-ee top quark program.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The object of study in this thesis is the production of a top quark pair at future lepton colliders with a centre of mass energy in the range $\sqrt{s} = 340 - 350$ GeV. The top quark has never been produced yet at a lepton collider due to the extremely high energies required. This is bound to change as at the time of writing global strategy meetings are taking place to decide which accelerators will succeed to the Large Hadron Collider. The study presented here concerns the Future Circular Collider facility.

Chapter 2 introduces the Standard Model as well as describe the importance of the top quark. The various accelerator projects are discussed with a focus on the Future Circular Collider in Chapter 3. The analysis starting from Monte-Carlo event generation to simulation of the detector is explained in Chapter 4. Multiple statistical methods developed in this thesis are compared to a more complex approach in Chapter 5. In Chapter 6 the direct measurements of the top quark and W boson mass are studied for different simulated top masses and centre of mass energies.

Chapters 2 and 3 are based on literature study for the thesis. Chapter 4 explains how I created data samples from scratch running the appropriate software and decided on event selection. For Chapters 5 and 6 I came up with a statistical method to sort the data and produce new histograms for predictions. In essence, I had to go along every step of the way from creating data to treating it and finally analysing it.

Throughout this entire thesis the natural unit system will be used: $\hbar = c = 1$. As a consequence all units are expressed as powers of energy using GeV. The basic unit being 1 eV $= 1.602 \cdot 10^{-19}$ J, is defined as the energy gained by an electron in a potential difference of 1 V in vacuum. Exceptions are made for some quantities to provide a more intuitive value.

Chapter 2

Theory

The Standard Model of particle physics describes the elementary particle content and their interactions is presented in section 2.1. Although it is commonly accepted to work for the particles we observe, the Standard Model faces some serious problems, these are discussed in section 2.2.

This thesis studies one of the particles in the Standard Model, the top quark. Section 2.3 focuses on the properties and the role of this particle. Finally, section 2.4 is dedicated to the production of top quark pairs at a lepton collider.

Most of the concepts presented below are based on [1, 2].

2.1 The Standard Model

Our universe is described at its smallest scales by the Standard Model (SM). This theory characterises the interactions and properties of the elemental particles in the mathematical framework of quantum field theory. The SM is one of the most successful theories to date with excellent predictions in a large range of high energy physics experiments [1]. For example, the mass of the top quark and W boson matched the predictions very well [3].

The forces described by the SM are the electromagnetic force, the weak force and the strong force. These are carried by particles called mediator bosons, which have spin 1. The quantum numbers that characterise elementary particles are mass, spin, electrical charge, weak hypercharge and colour charge. In table 2.1 these mediator bosons are presented along with the measured masses. The electromagnetic force is responsible for the attraction or repulsion of charged particles and is mediated by the photon. The weak force allows for the radioactive decay of atoms and is mediated by the massive $W^{\pm}$ and Z bosons. The strong interaction is the glue for the structure of the nucleus inside an atom, which inspired the name gluon for its mediator boson.

Table 2.1: An overview of the mediator bosons and their respective interactions in the Standard Model with the measured mass from [4].

Interaction	**Mediator**	**Symbol**	**Mass (GeV)**
Electromagnetic	Photon	γ	0
Weak	W boson	W^+, W^-	80.379 ± 0.012
	Z boson	Z^0	91.187 ± 0.0021
Strong	Gluon	g	0

Fermions are particles which have half-integer spin and obey the exclusion principle, stating that two fermions with the same quantum numbers cannot occupy the same space at the same time.

Matter consists of twelve fermions. These fermions are subdivided into two groups, the quarks which are sensitive to the strong force and the leptons which are not. The particles can be separated into three generations, which share the same quantum numbers but differ only in mass, see Table 2.2. Take for example the electron e^-, a first generation lepton, and the muon μ^- as a second generation lepton. The muon is essentially the same particle as an electron but heavier by a factor $\mathrm{m}_\mu \sim 200\,\mathrm{m}_e$.

All these fermions have anti-particles that share the same properties but have the opposite electric charge. They are denoted by a bar over the symbol, so for quarks and neutrinos the anti-quarks and anti-neutrinos are respectively $\bar{t}, \bar{\nu}_e$. The leptons are marked by a plus sign in the superscript instead of the usual minus sign, e^+, μ^+. It is important to note that since the neutrinos have no electrical or colour charge, they do not interact via the electromagnetic force. As indicated in Table 2.2, quarks have fractional charges. Another property of quarks is called colour confinement, which states that only colourless particles can propagate freely. Since quarks have a colour charge, they can only propagate as composite particles called hadrons. A combination of two up-quarks and one down-quark will result in a proton with a positive charge, two down-quarks and one up quark form the neutron. These nucleons are with the electrons the building blocks for all the elements present in the periodic table.

Lastly there is the Brout-Englert-Higgs (usually called the Higgs) boson, a scalar particle with spin-0. The existence of this particle allows fermions and the W, Z bosons to acquire mass through a process called spontaneous symmetry breaking. In essence the Higgs field exists in a potential which has the special property that it has multiple lowest energy states. The action of "choosing" a minimum breaks the so-called electroweak symmetry. As a consequence of this process, the massive

Table 2.2: An overview of the fermions in the Standard Model. Note the charge has no units in the natural unit system, the values are multiples of the charge of an electron.

Group	**Electrical charge**	1^{st} **generation**	2^{nd} **generation**	3^{rd} **generation**
Quarks	$+2/3$	up u	charm c	top t
	$-1/3$	down d	strange s	bottom b
Leptons	-1	electron e^-	muon μ^-	tau τ^-
	0	electron neutrino ν_e	muon neutrino ν_μ	tau neutrino ν_τ

H boson generates masses for other particles and new interactions. This particle was discovered in 2012 at the LHC by the CMS [5] and ATLAS [6] experiments.

2.2 Shortcomings of the Standard Model

The Standard Model is an incomplete theory, it faces several problems that remain unsolved to this day. Some of the questions that have not yet been answered are listed below.

The orbital velocity of the stars in distant galaxies stabilises after a certain radius, instead of the predicted decrease with $r^{-1/2}$. This phenomenon cannot be explained without the presence of another massive component which we cannot observe. This is known as the dark matter problem. These observations since the 1930s have gathered a lot of evidence for physics beyond the Standard Model, as no particle in the SM can explain this.

Two of the forces described by the SM have been unified under what is called the Electroweak theory. Naturally one wonders if it is possible to also include the strong interaction in the same picture. Such theories have emerged since the 1970s under the collective name of Grand Unified Theories (GUT) [7]. The existence of a GUT would mean that at higher energy scales all forces have the same coupling. This would imply that at some point in the early universe a symmetry was spontaneously broken in order to separate into three distinct interactions. Some configurations of Supersymmetry, an extension to the SM, allow to unify the gauge couplings of the three SM interactions at the cut-off energy scale $\Lambda_{\mathrm{GUT}} \sim 10^{16}$ GeV [8]. To give some perspective, the Planck scale $\Lambda_{\mathrm{Planck}} \sim 10^{19}$ GeV gives the energy at which the SM and general relativity are expected to break down.

The Higgs boson has been measured to have a mass of $\mathrm{m_H} = 125.18 \pm 0.16$ GeV [4].

In the SM the mass of the Higgs boson depends on quantum loop correction terms. This becomes a problem if the Standard Model is part of a larger theory that is valid up to high energy scales such as for Grand Unified Theory $\Lambda_{\mathrm{GUT}} \sim 10^{16}$ GeV or the Planck scale at $\Lambda_{\mathrm{Planck}} \sim 10^{19}$ GeV. The corrections from the loop diagrams scale as Λ^2, meaning the parameters of the model need to be fine-tuned very precisely to cancel out. Assuming the SM is consistent up to the GUT scale, the mass of the Higgs boson would have to be tuned down to $(10^{16})^2 = 10^{32}$ orders of magnitude. This unnatural amount of fine-tuning is called the Hierarchy problem. The large top quark mass is at the heart of this problem since it brings the largest contribution to the Higgs boson mass.

2.3 The top quark

The top quark was observed for the first time in 1995 at the Tevatron collider [9] by the CDF [10] and D0 [11] experiments. Its discovery confirmed a prediction made more than a decade earlier when the bottom quark was first detected.

The top quark is the heaviest particle in the SM, the combined measurements from the Tevatron analyses and the LHC experiments results in the value $m_{\mathrm{t}} = 173.0 \pm 0.4$ GeV [4]. From the measured top quark decay width $\Gamma_{\mathrm{t}} = 1.41^{+0.19}_{-0.15}$ GeV [4] we can find the lifetime: $\tau_{\mathrm{t}} = \Gamma_{\mathrm{t}}^{-1} \sim 5 \cdot 10^{-25}$ s. As a result of this very short lifetime the top quark does not have time to hadronise, instead it decays via the weak interaction.

The top quark decays almost exclusively via the production of a bottom quark and W boson. The main reason for this is the large value of the CKM matrix element $|V_{tb}| = 1.019 \pm 0.025$ [12], which quantifies the probability of a top quark to decay into a bottom quark via the weak interaction. This vertex is illustrated in Figure 2.1.

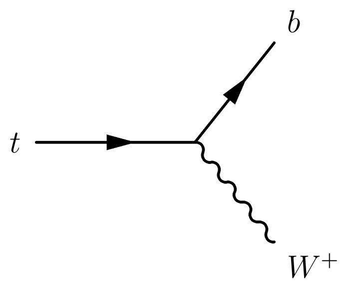


Figure 2.1: Feynman diagram showing the main decay channel of the top quark into a bottom quark and a W boson.

An important aspect of the top quark mass measurement is that it allows to conduct consistency checks for the SM. In principle any free parameter of the SM can be used for such tests, but the very large mass of the top quark leads to bigger correction terms. One of these tests is the simultaneous measurement of the W, H bosons and top quark masses using the electroweak mixing angle θ_W. Figure 2.2 shows the direct measured masses in green against the indirect, predicted, masses in grey and blue. The blue contours have correction terms for the top quark and the W boson while the grey one also takes the H boson into account. More precise direct measurements of the top quark, W and H boson masses can greatly improve this test for the SM.

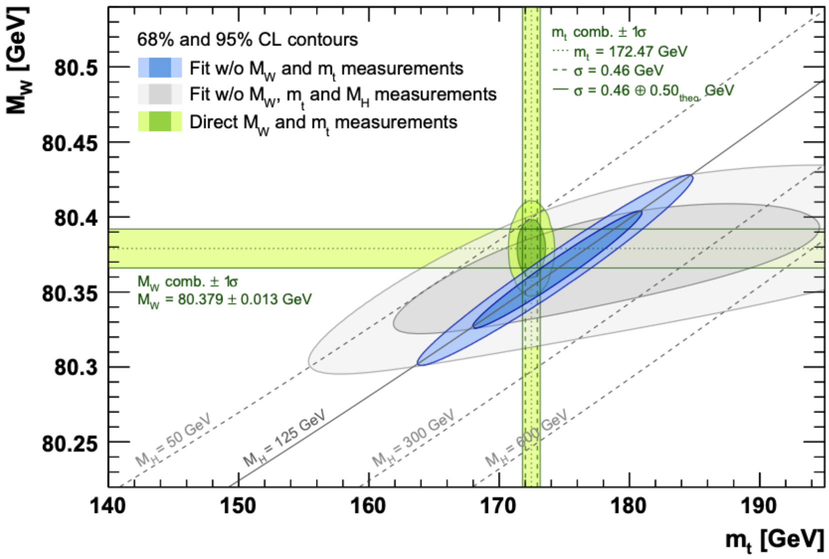


Figure 2.2: Comparison of the direct measured masses of the W boson and top quark (green) and the indirect measurements excluding (grey) and including the H boson mass (blue). [13]

2.4 The $t\bar{t}$ process in lepton colliders

At lepton colliders, the main production mechanism for a top quark and an anti-top quark is pair creation via the s channel as shown in Figure 2.3.

Production of top quarks at lepton colliders is an interesting prospect for multiple reasons. First off, no top quarks have been produced yet via this mechanism due to the high collision energy required. Lepton colliders also offer the possibility to know the exact centre of mass energy $\sqrt{s}$ for each collision. This allows to study properties such as the cross-section, a measure of the frequency at which a certain final state occurs, at different $\sqrt{s}$ values. The cross-section for a top quark-antiquark pair at the threshold $\sqrt{s} \simeq 2\mathrm{m_t}$ is particularly compelling, as the shape of this function depends on four SM free parameters.

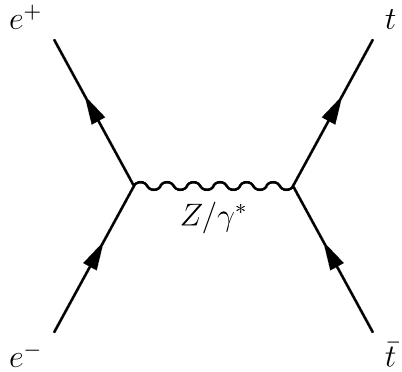


Figure 2.3: Feynman diagram showing the main process for $t\bar{t}$ events at lepton colliders.

The projected threshold is shown in Figure 2.4 with arrows indicating the influence of each parameter. The top quark mass $\mathrm{m_t}$ changes the location of the peak, the width Γ_t broadens the peak and the Yukawa coupling $\mathrm{y_t}$ and strong coupling α_s affect the height of the continuum region past $\sqrt{s} > 2\mathrm{m_t}$. The black curve indicates the latest theory prediction at NNNLO [14] while the other curves show the expected values for different scenarios at the Future Circular Collider, see section 3.2. The green curve shows the effect of Initial State Radiation (ISR) on the cross-section. ISR is the radiation of one or more photons before the collision by one of the electrons lowering the resulting energy. The blue and red curve represent the FCC-ee expected measurements with and without ISR respectively. The red curve is used for the analysis.

As discussed in section 2.3, the top quark decays via the weak interaction, regardless of the initial particles colliding. Most frequently the top quarks will produce two W bosons and two bottom quarks. The final state to look for in the detectors will hence depend on the W decays. The W boson decays about 2/3 of the time to a quark-antiquark pair and about 1/3 of the time to lepton and neutrino. As a result there are three main final states [16]:

- All-hadronic: BR($t\bar{t} \rightarrow bW^+\bar{b}W^- \rightarrow bq\bar{q}\bar{b}q\bar{q}$) $\simeq 45.7\%$
- Semi-leptonic: BR($t\bar{t} \rightarrow bW^+\bar{b}W^- \rightarrow bq\bar{q}\bar{b}l^-\bar{\nu}_l$) $\simeq 43.8\%$
- Di-leptonic: BR($t\bar{t} \rightarrow bW^+\bar{b}W^- \rightarrow bl^+\nu_l\bar{b}l^-\bar{\nu}_l$) $\simeq 10.5\%$

These three final states have a very different topology and require different analysis strategies. The semi-leptonic channel is very interesting with its single neutrino, which makes it easy to reconstruct using the missing energy. This mode also has a good signal to background ratio in general. The di-leptonic decay is easy to identify and has a very high signal to background ratio. The two neutrinos however make event reconstruction more difficult and the branching ratio is quite low. Finally the all-hadronic channel with six jets remains. This final state is very dependent on the capabilities of the detector to separate b quark jets from light quark jets and suffers from a worse signal to background ratio. After considering the options, the

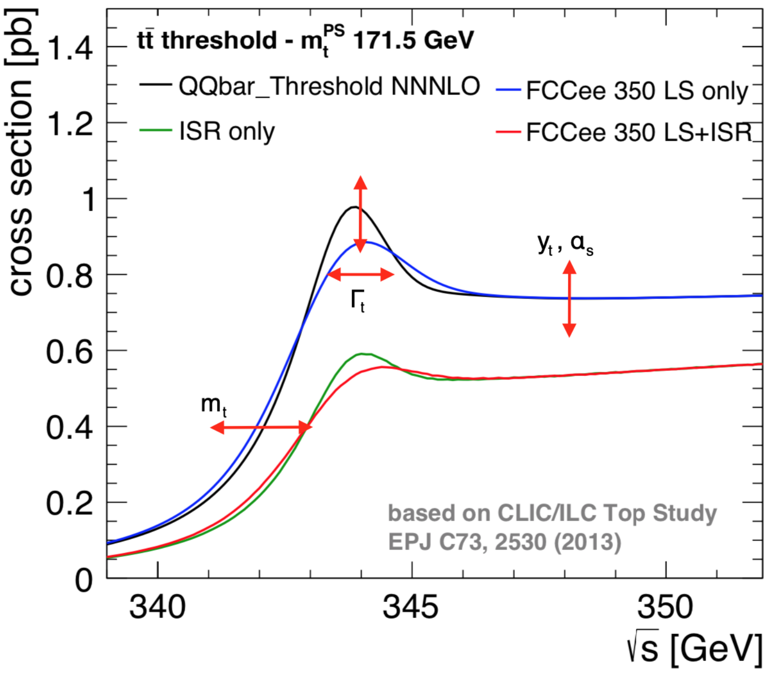


Figure 2.4: cross-section of the $t\bar{t}$ process at the FCC-ee. Modified from [14, 15]

semi-leptonic channel was chosen for the study in this thesis.

Chapter 3

Experimental setup

The Future Circular Collider is a proposed collider that includes multiple sub-projects. To be able to decide if these are worthwhile building, it is necessary to study the physics potential of such machines. The analysis in this thesis is one such study. A brief summary of the various proposed machines is given in section 3.1. An overview of the goals and design of the FCC project are described in section 3.2. More details are given for the lepton collider FCC-ee in section 3.3. The detector facilities and their properties are then discussed. A short summary of the software used for the works is included in section 3.4.

The information concerning the FCC is based on the recently released Conceptual Design Reports, in particular volume 1 and 2 [17, 18].

3.1 Future lepton colliders

Four future lepton colliders are currently being proposed. The Future Circular Collider (FCC) is planned to have multiple phases, as both a lepton and hadron collider. For the lepton program it would run at multiple centre of mass energies ranging from the Z pole at $\sqrt{s} = 90$ GeV to the top quark pair threshold $\sqrt{s} \sim 365$ GeV. The Circular Electron Positron Collider (CEPC) offers a similar lepton program but would stop at the HZ production $\sqrt{s} = 240$ GeV. An upgrade is under consideration to add the top quark to the program.

Two of these are linear colliders, starting with the International Linear Collider (ILC). This project would only run at the HZ production threshold with the main design. Finally the Compact Linear Collider (CLiC) is planned to run at very high energies starting at $\sqrt{s} \sim 380$ GeV up to 3 TeV. The projected luminosities L, a measure of the instantaneous amount of collisions, for the various accelerators are shown in Figure 3.1. This quantity is not to be confused with the integrated luminosity L_{int}, which is directly related to the amount of events observed.

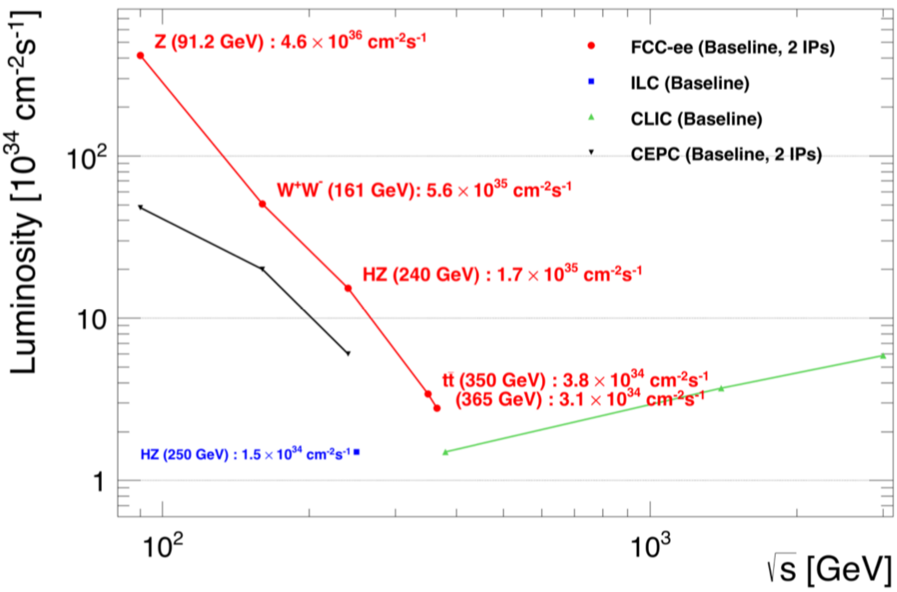


Figure 3.1: The expected luminosities to be delivered in function of the centre of mass energy $\sqrt{s}$ for four international lepton colliders. In red the FCC-ee, in blue the International Linear Collider (ILC), in green the Compact Linear Collider (CLiC) and in black the Circular Electron Positron Collider (CEPC). [18]

3.2 The Future Circular Collider

The Future Circular Collider is an accelerator project proposed as a successor to the current Large Hadron Collider. The collider would be hosted in a 100 km tunnel under Geneva as shown in Figure 3.2. The design of the accelerator would provide high luminosity collisions at various centre of mass energies. As with the Large Electron Positron Collider (LEP) [19] and the LHC, the tunnel and components can be reused to accelerate and collide different kinds of particles. The FCC can be set up as a lepton collider, the FCC-ee, as a machine for high precision measurements. Similarly as for the LHC, protons or more generally hadrons can be collided, this is the FCC-hh project which could reach energies up to 100 TeV. Potentially the collider could be used for lepton-hadron collisions as the FCC-eh in order to further study the substructure of the protons. The work of this thesis concerns the program for the FCC-ee.

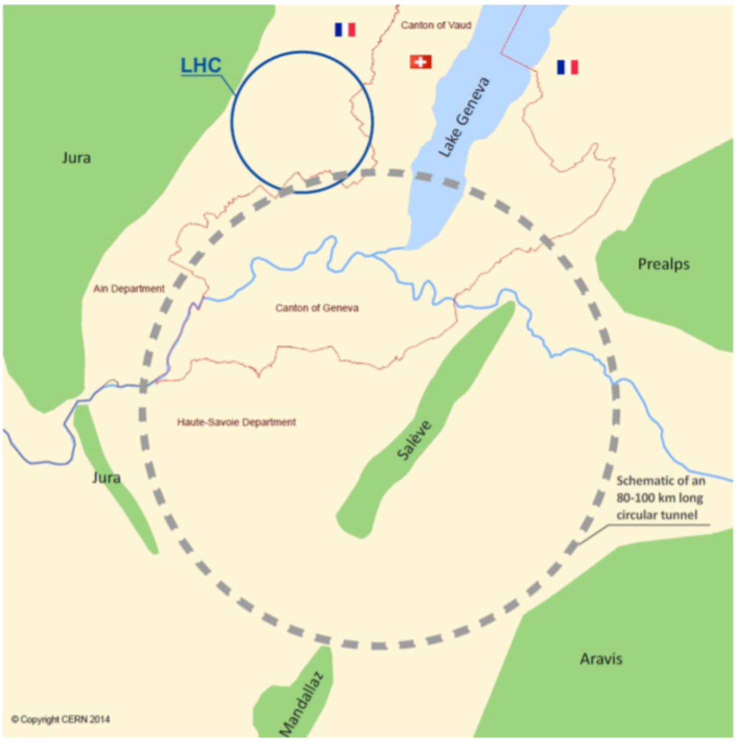


Figure 3.2: Footprint of the Future Circular Collider tunnel in the Geneva area at CERN. [20]

3.3 The FCC-ee

The FCC-ee is designed as the highest luminosity and energy limit reaching circular electron-positron collider. Equipped with two general purpose detectors it will conduct high precision measurements for the Z, W and Higgs bosons and the top quark, see Figure 3.1. Lepton colliders are the ideal machines to study particular phenomena in great detail. The beam collides with a known energy and produces much less background compared to hadron colliders. As a result of the high amount of statistics, the FCC-ee will be able to look for signs of physics beyond the SM, these will appear as small deviations in consistency checks. For instance the very high statistics it will provide are ideal to observe flavour changing neutral currents or weakly coupled particles such as right-handed neutrinos [17].

3.3.1 Design and program

The FCC-ee is designed as a double-ring collider with a circumference of 97.76 km. The rings will host a large number of bunches, a collection of electrons or positrons grouped together, and cross at two interaction points (IP). The bunches are accelerated at two RF sections after injection via a top-up scheme, which means the beams are refilled during operation when the current is low enough. A schematic of the accelerator is shown in Figure 3.3.

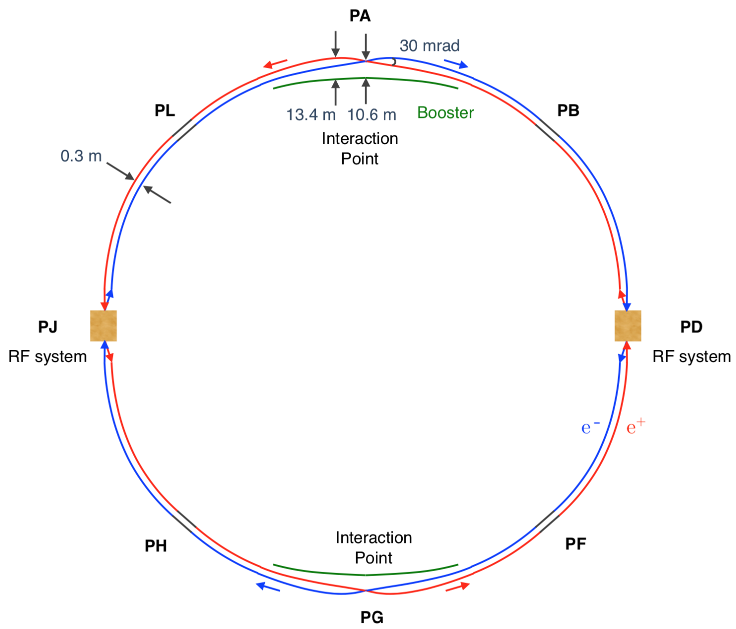


Figure 3.3: Schematic of the layout for the FCC-ee rings. The blue line indicates the electron beam and the red line the positron beam. The two detectors will be placed at points PA and PG. Modified from [18]

The FCC-ee would operate for 15 years in total, increasing the energy every few years. At first the Z pole will be studied, then W pairs will be produced, followed by the HZ process and finally the $t\bar{t}$ threshold during the last five years. A short overview of the program can be found in Table 3.1. The integrated luminosity goal L_{int} and the total events give an idea of the statistics the accelerator will produce.

Table 3.1: This table gives an overview of the program for the FCC-ee. A slash (/) separates the properties if multiple phases are planned for the same final state. For the Z pole the last two years will ramp up the luminosity. The $t\bar{t}$ process will first have a threshold scan then a longer continuum phase. Based on [21]

Quantity	**Z**	**WW**	**HZ**	**t$\bar{\text{t}}$**
$\sqrt{s}$ (GeV)	88,91,94	157,163	240	340-350,360
Luminosity goal (ab^{-1})	150	10	5	0.2/1.5
Run time (year)	2/2	2	3	1/4
Total events	$5 \cdot 10^{12}$	10^8	10^6	10^6

The top quark pair production will have in total five years of run time, one for a scan at threshold energies and four at the continuum level. Note that the exact energies for the scan are not finalised at the time of writing and could change in the future. For each energy point scan for the top quark threshold in this thesis we will assume we have the integrated luminosity [18, 21]:

$$L_{\text{int}} = 0.2\,\text{ab}^{-1} \tag{3.1}$$

3.3.2 Detectors

The detector used in this thesis is the International Large Detector (ILD), originally intended for the ILC. Two other complementary detector designs are also being studied for the FCC-ee, the CLiC-like Detector (CLD) and the International Detector for Electron-positron Accelerators (IDEA). Note that the properties of the CLD and ILD should be quite similar. The CLD is in fact a modified version of the ILD adapted to the specifics for the FCC-ee.

The ILD is designed as a multi-purpose detector, it uses a pixel vertex detector followed by a hybrid tracking system consisting of silicon layers and a time projection chamber (TPC). In the outer edges are the calorimeters, all of this is placed inside a large 2T solenoid magnet. Finally the muon system is located outside of the iron yoke for the magnet. Figure 3.4 shows an impression of the ILD. Overall the detector is quite similar to the Compact Muon Solenoid at the LHC.

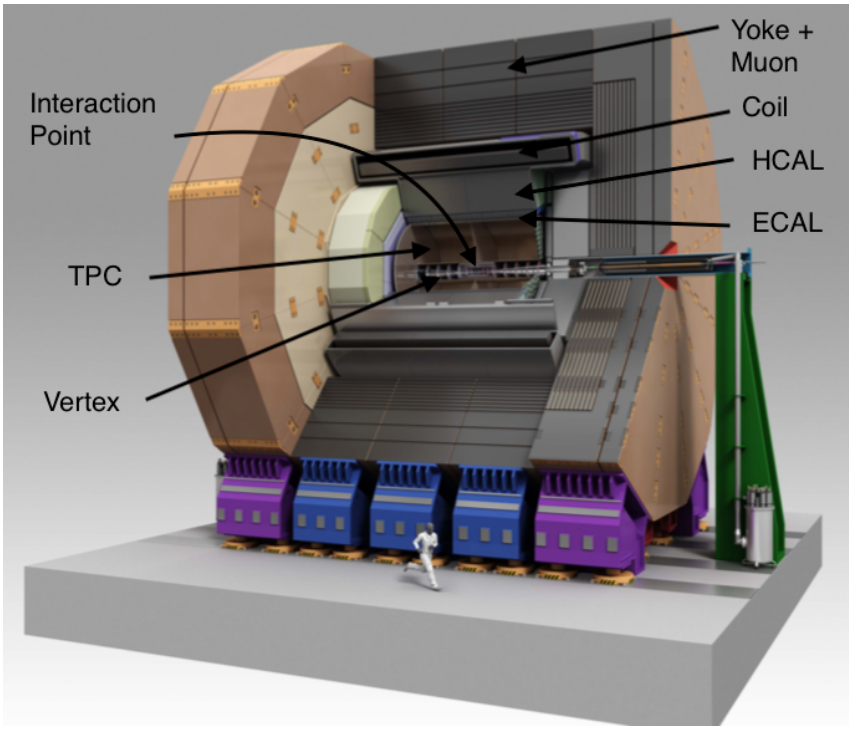


Figure 3.4: Schematic overview of the ILD. Modified from [22]

The characteristics used for the analysis are listed below [23]:

- Time Projection Chamber and silicon tracker: $\Delta\left(\frac{1}{p}\right) < 5 \cdot 10^{-5}$ GeV^{-1}
- Silicon-Tungsten ECAL: $\Delta E/E = 15\%/\sqrt{E}$
- High impact parameter resolution: $\sigma(rz) = 5\mu\text{m} \oplus 10/p\sin^{3/2}\theta$
- Jet flavour tagging: see Figure 3.5

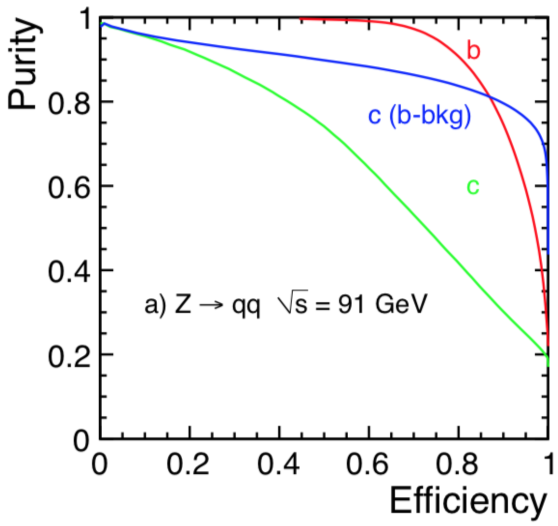


Figure 3.5: Jet flavour tagging of the ILD detector. [22]

The CLD is designed in the same fashion as the ILD. The dimensions and technologies going into the subdetectors are slightly different however. The IDEA detector is specifically developed for the FCC-ee. The entire concept for the detector is pushing the limits of current technology. IDEA relies on a drift chamber system with silicon sensors for the tracking system. Unlike the conventional 2T solenoid magnet for the ILD and CLD, the IDEA magnet system would be placed in front of the calorimeters. More information and technical details can be found in the FCC-ee CDR [18].

3.3.3 Important detector features

The important factors for the $t\bar{t}$ process reconstruction are good momentum resolution and electromagnetic calorimeter resolution for the leptons. Good discrimination for jet flavour tagging is very useful since there are always two b quarks in the final state, which would not be the case for many backgrounds. A good impact parameter resolution is also particularly useful, that data is at the core of most jet b-tagging algorithms. This is achieved by having a very good vertex resolution, as the bottom quarks have a relatively long lifetime and can travel large distances compared to other particles before the hadronisation process starts. This is particularly useful for H bosons and top quarks that often decay into bottom quarks.

3.4 Simulation tools

Various programs and frameworks have been created in order to help the FCC community simulate data for predictions. These tools are essential to help physicists collaborate on projects and avoid replicating the same work. Below is a summary of each of the programs used for this thesis.

3.4.1 FCCSW framework

The FCCSW framework [24] is a set of packages and tools designed to help FCC analyses. It encompasses the simulation, reconstruction and analysis of Monte Carlo samples. At first a selected process with a chosen initial and final state is generated using the Monte Carlo method. This method revolves around using random number generators and probability distribution functions for the various properties of the particles. It also uses the matrix elements of the physics process that one is interested in. The next step is the simulation of the process as perceived through a detector. This means the generated event values become smeared according to the detector resolutions and simulate hits on the subdetector components.

The FCCSW framework is relatively young and a lot of effort is put into its development. As a result of this, an earlier version had to be run in order to ensure stable operation.

3.4.2 Pythia8

Pythia8 [26] is a widely used Monte Carlo (MC) event generator for LO processes. It can compute simple scattering processes and complex multiparticle final states. The physics aspect combines derivations through theory and phenomenological models. It also automatically computes the cross-sections and hadronises the quarks into jets. The FCCSW has embedded Pythia8 and stores the MC samples into the FCC event data model.

3.4.3 PAPAS

Papas is the parameterized particle simulation [27] for the FCC-ee, it allows to simulate the detector and is integrated into the FCCSW. Papas is a fast simulation tool that is intended to be used for analysis benchmarks at very high speed. It is particularly useful to iterate what designs for a detector will work best and extract the optimum configuration.

Papas uses an enhanced version of the Particle Flow [25] algorithm originally developed by the CMS Collaboration for reconstruction. Reconstruction is the process of using the detector's data together to find out which particles have passed through what parts of the various subdetector systems. In principle, the same reconstruction program could run on a simulation the same as for real data in a running experiment. For example, an electron will pass through a tracker system and produce an electromagnetic shower in the ECAL, combining the information from the electronics will produce a four-vector with the energy and momentum of the electron. The algorithm used to combine this information back into a four-vector is called Particle Flow.

To recover the jet four-momentum from a particle shower Papas uses a jet finding algorithm such as FastJet [28, 29]. FastJet is a sequential algorithm that combines the particles in the shower like pions, leptons and photons that are closest to each other until a stopping criterion is reached. Muons pass through the entire detector and leave hits in the tracking system, these are used to reconstruct the original four-momentum. An additional feature of Papas is that it can display events with tracks and energy clusters in the detector.

Chapter 4

Event generation and selection

This Chapter focuses on the practical side of the sample generation. At first the topology of the events to look out for is discussed in section 4.1. The main sources of backgrounds that can produce the same signature are also considered. Section 4.2 describes the steps and details to generate the Monte Carlo samples. The process of eliminating the background events through successive cuts on certain variables is explained in section 4.3.

4.1 Decay channels and background for the $t\bar{t}$ process

The $t\bar{t}$ process is studied in this thesis for the semi-leptonic decay channel. At the reconstruction level this means that the final state signature contains:

- Four jets, of which two are b-jets
- One high energy isolated lepton
- Missing energy, attributed to a neutrino

It is important to note that the W boson can also decay into a τ lepton, which has a short lifetime. The τ can decay into hadrons or into another charged lepton and neutrinos, in that case the event is still considered for selection. From this point on the term lepton will refer only to electrons and muons, unless stated otherwise.

The main sources of background that can generate the same final states are the $t\bar{t}$ all-hadronic and di-leptonic decay channels as well as WW, HZ and ZZ production. The Feynman diagram for the signal is shown in Figure 4.1 and the backgrounds are shown in Figure 4.2. The background processes, except for WW, mostly have cross-sections in the same order of magnitude:

- $t\bar{t}$ production: $\sigma_{t\bar{t}}(\sqrt{s} = 340 - 350\text{GeV}) = 83 - 550$ fb
- WW production: $\sigma_{\text{WW}}(\sqrt{s} = 350\text{GeV}) = 5000$ fb
- HZ production: $\sigma_{\text{HZ}}(\sqrt{s} = 350\text{GeV}) = 130$ fb
- ZZ production: $\sigma_{\text{ZZ}}(\sqrt{s} = 350\text{GeV}) = 500$ fb

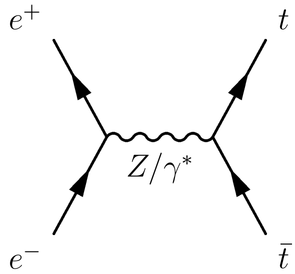
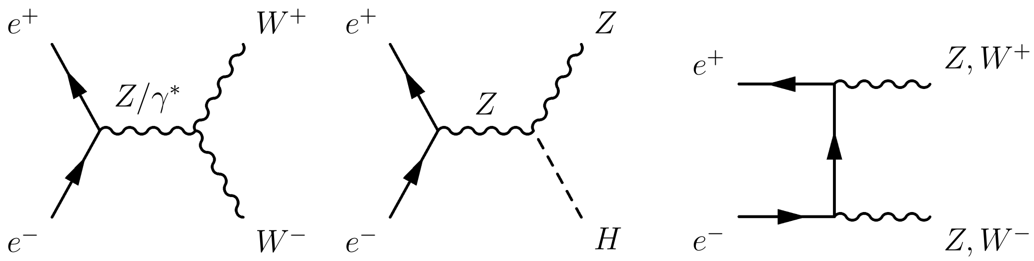


Figure 4.1: $t\bar{t}$ production



(a) WW production (b) HZ production (c) WW & ZZ production

Figure 4.2: Representative Feynman diagrams for the main sources of background.

4.2 Event generation

The events are generated using Pythia8 [26] inside the FCCSW framework at tree level for both signal and backgrounds. For the purpose of the analysis on the efficiency of the selection procedure explained in sections 4.3 and 5.1, Monte Carlo (MC) samples are generated at multiple values of $\sqrt{s}$ and m_t. The values used for the samples are,

$$\text{m}_\text{t} = [171.8; 172.6; 173.0; 173.4; 174.2] \text{ GeV}, \tag{4.1}$$

$$\sqrt{s} = [340, 342, 344, 346, 350] \text{ GeV}. \tag{4.2}$$

The values were chosen in order to cover the top quark pair threshold similar to the FCC-ee program, see Table 3.1. The top quark masses are centered around the combined measurements from [4], and deviations are taken with $\pm 1\,\sigma$ and $\pm 3\,\sigma$.

The background samples are assumed to be constant with the top quark mass, in fact top quarks can only appear in the loops of Feynman diagrams for these processes. The effect of the mass change is also expected to be much smaller than the corrections from for example higher order MC simulations. The $t\bar{t}$ semi-leptonic samples are created by considering full $e^-e^+ \to t\bar{t}$ events, then identified by recovering the so-called Monte Carlo truth. This means that after an event is saved, the Monte Carlo simulated data from the real process is used to verify that the top quarks have indeed decayed into $bW^+\bar{b}W^- \to bq\bar{q}\bar{b}l^-\bar{\nu}_l$. As a consequence of this method the signal events efficiency is lower than immediately generating the semi-leptonic channel. For all the backgrounds the events are generated in their corresponding decay channels.

All of the samples are generated with 100.000 events as goal, that number can in practice be lower due to failed runs. Creating the MC events with Pythia8 is very fast and takes about a minute for 2500 events, while the analysis requires roughly one hour of work with eight worker nodes in a cluster. The samples are then scaled to the number of events N using the integrated luminosity expected at the FCC-ee and the predicted cross-section for the process,

$$N = L_{\text{int}}^{\text{goal}} \cdot \sigma. \tag{4.3}$$

The events are scaled such that if the FCC-ee runs as stated in the program from the CDR [18], the event count measured should be similar.

The cross-section values are obtained by modifying the reference red curve from Figure 2.4 [15]. For a different top quark mass the entire curve is displaced to higher energies with a shift of $2 \cdot \Delta \mathrm{m_t}$ from the reference mass $\mathrm{m_t} = 171.5$. The missing data at lower energies is linearly extrapolated using the first two cross-section values and set to zero if it becomes negative. The modified cross-sections can be visualised in Figure 4.3, with in blue the reference curve.

After creating the samples with Pythia8 the LO cross-section was extracted and can be found in Figure 4.4. Since all the background processes occur at energies in the continuum region, the cross-sections stay relatively constant against the centre of mass energy. The pink curve shows top quark pair production which increases significantly. As one can notice, the shape of the curve is quite different from the latest threshold prediction, see Figure 2.4. The theoretical corrections for these measurements have been calculated up to NNLO at the time of writing, this explains the large discrepancy in cross-sections.

Finally all the samples have been generated with the following settings:

- PartonLevel:ISR = on

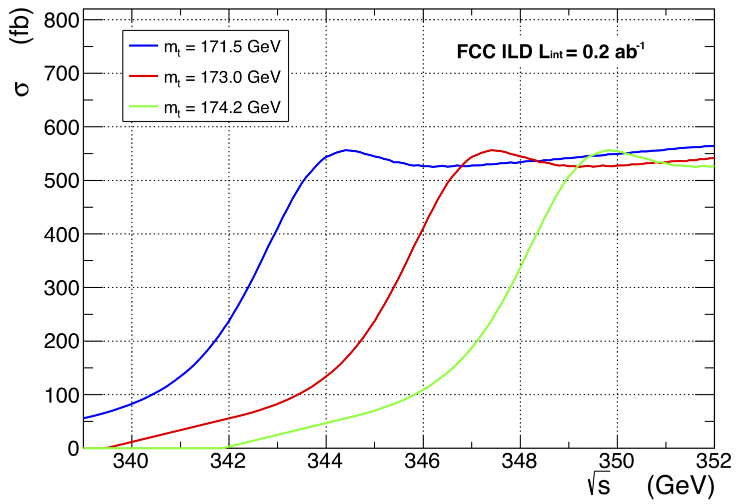


Figure 4.3: Cross-section curves for different top quark masses after modification.

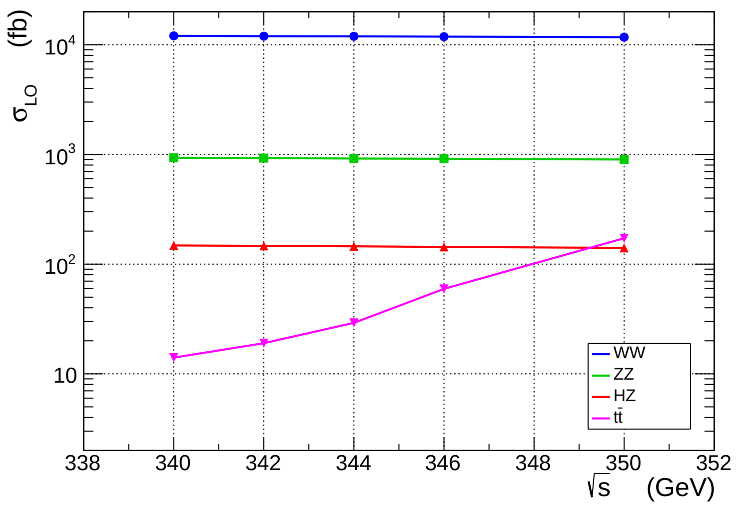


Figure 4.4: cross-section at LO from the generated MC samples.

- PartonLevel:FSR = on
- HadronLevel:Hadronize = on

4.3 Event pre-selection

All of the background processes have different topologies from the signal, that is to say they do not produce exactly the same signature as the signal. These can end up producing the same topology but at lower efficiencies if a selection is made optimised for the signal. Another important aspect is that the backgrounds behave differently in terms of kinematics. The momentum and energy distributions will in general only overlap partially.

The pre-selection of events starts with the single isolated lepton. The reconstructed leptons considered are electrons and muons, for which the energy exceeds at least,

$$\mathrm{E_{lepton}} > 10\ \mathrm{GeV}. \tag{4.4}$$

The isolation of the lepton is defined as the sum of the energy of all the hadrons and photons within a cone with $\Delta R < 0.3$ rad, where ΔR is defined as the spatial angle. To pass this step the isolation must be below the threshold value,

$$\mathrm{Isolation} < 15\ \mathrm{GeV}. \tag{4.5}$$

The isolation requirements remove leptons produced through hadronisation processes of quarks, essentially removing leptons from jets as candidates. The settings for this cut have not yet been optimised but are planned for further studies [23].

The event is also required to have at least four particles reconstructed in Papas which are then turned into jets with FastJet [28, 29]. The events that pass this first criterion are saved for further analysis, see Table 4.1.

To reduce the background while maintaining a large part of the signal more cuts are necessary. To this end the next variables are defined:

$$\mathrm{m_{4j}} = \text{Invariant mass of the sum of the four jets,} \tag{4.6}$$

$$\mathrm{m_{2j}^{min}} = \text{Minimum invariant mass of the sum of any two jets} \tag{4.7}$$

$$\mathrm{m_{2j}^{2nd\ min}} = \text{Second minimum invariant mass of the sum of any two jets} \tag{4.8}$$

$$\not{E}^{\mathrm{reco}} = \text{Missing reconstructed energy} \tag{4.9}$$

First the invariant mass of the four jets is considered in Figure 4.5. The WW mode dominates the spectrum and has a peak located around the W boson mass. From this distribution with only the previous cut, it is clear that a large amount of background can be eliminated as it is mostly spread across evenly. In order to preserve the signal the following two cuts are chosen:

$$\mathrm{m_{4j}} > 150\ \mathrm{GeV}, \tag{4.10}$$

$$\mathrm{m_{4j}} < 270\ \mathrm{GeV}. \tag{4.11}$$

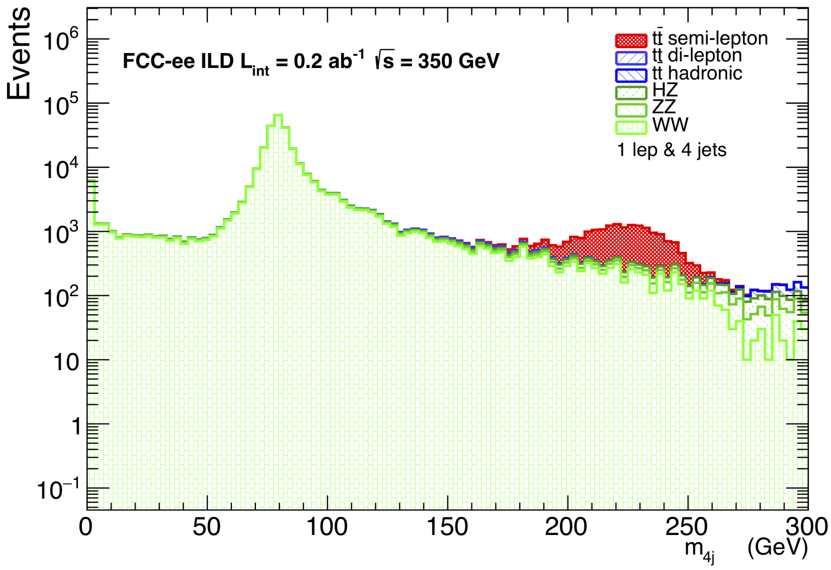


Figure 4.5: Invariant mass of the sum of the four jets for all samples. The semi-leptonic $t\bar{t}$ (red) decay are located in a short energy range $\sqrt{s} \in [150, 270]$ GeV.

These two cuts together eliminate more than 93 % of the background events.

The invariant mass of a couple of jets is computed, this procedure is repeated for all possible combinations of two jets. The minimum in this list is defined as $\mathrm{m}_{2\mathrm{j}}^{\mathrm{min}}$ and the second minimum as $\mathrm{m}_{2\mathrm{j}}^{\mathrm{2nd\ min}}$. The minimum invariant mass of any two jets is considered next in Figure 4.6.

The distributions from the backgrounds are very clearly skewed towards lower energies compared to the semi-leptonic $t\bar{t}$ decay. Therefore the cut,

$$\mathrm{m}_{2\mathrm{j}}^{\mathrm{min}} > 10\,\mathrm{GeV}, \tag{4.12}$$

is chosen. The value for the cut could potentially be slightly raised without a significant loss.

The next variable in the list is the second minimum for the mass of any two jets. Figure 4.7 shows a similar trend as for the previous quantity. The choice for the cut is set to,

$$\mathrm{m}_{2\mathrm{j}}^{\mathrm{2nd\ min}} > 20\,\mathrm{GeV}. \tag{4.13}$$

The isolated lepton energy is then used for the pre-selection, see Figure 4.8. Here the background continues to much higher energies compared to the signal. There are fewer particles in the final states for the WW, HZ and ZZ processes, therefore each particle can have a larger share of the collision energy. The chosen cut on the lepton energy is,

$$\mathrm{E}_{\mathrm{lepton}} < 100\,\mathrm{GeV}. \tag{4.14}$$

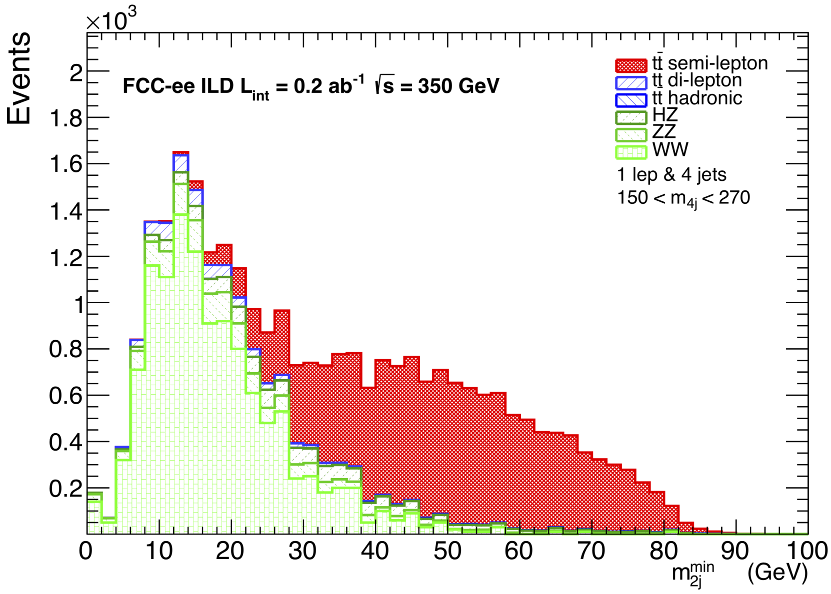


Figure 4.6: Minimum invariant mass of any two jet combination for all samples. All the background processes tend to have distributions with much lower energies than the signal. All previous cuts have been applied.

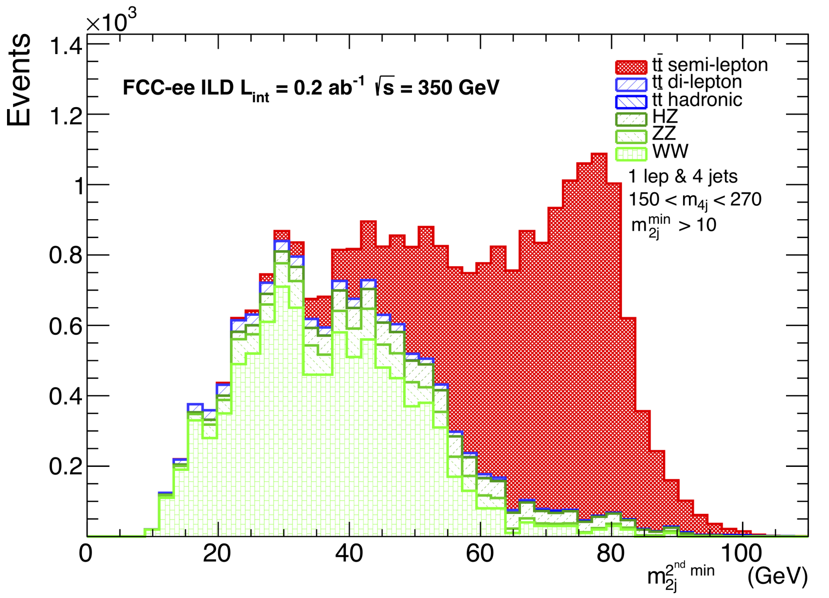


Figure 4.7: Second minimum invariant mass of any two jet combination for all samples. The backgrounds have much lower values compared to the semi-leptonic decay mode. All previous cuts have been applied.

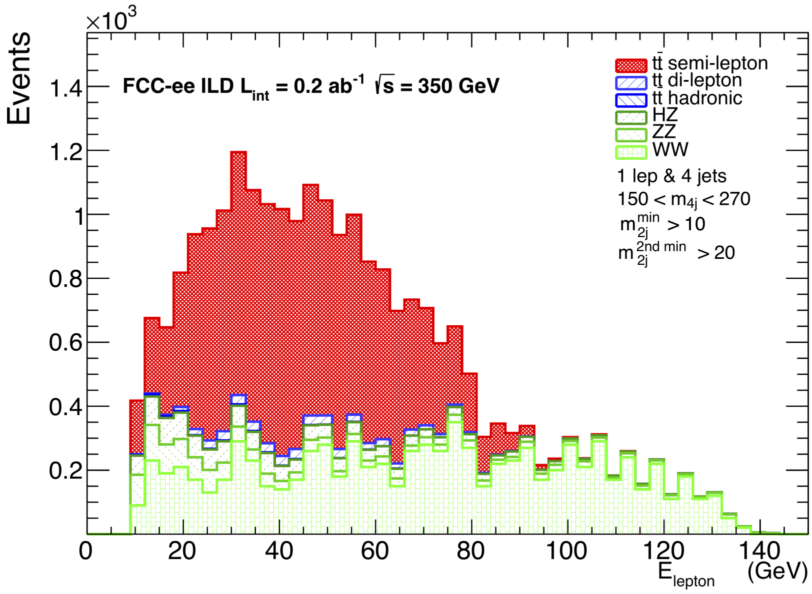


Figure 4.8: Lepton energy spectrum for all samples. The background distributions run to much higher energies than the signal. All previous cuts have been applied.

Finally consider the missing reconstructed energy. The distributions are shown in Figure 4.9. For this variable the distributions are very similar, this is to be expected as the remaining largest sources of backgrounds also produce neutrinos. The gain to be expected from this cut is much lower compared to the previous ones. The value for the selection is set to:

$$\not{E}^{\text{reco}} > 20\,\text{GeV}. \tag{4.15}$$

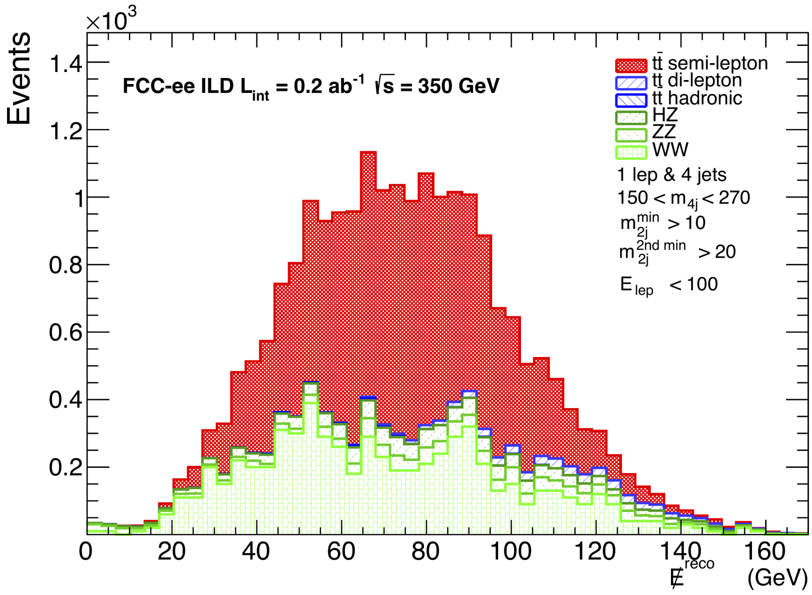


Figure 4.9: Missing reconstructed energy distribution for all samples. The spectra are quite similar in this case. All previous cuts have been applied.

The results of these cuts can be found in Table 4.1 for the yields. From this it is obvious that very little signal is lost, while all of the background samples are very strongly reduced. The largest source that passes the cuts is the WW boson production, which has more yield than the other backgrounds combined. This is logical as the signature of the signal also depends on two W bosons and it is the process with the largest cross-section to begin with. At this stage the data is ready for the analysis.

The efficiencies after the cuts are shown in Table 4.2. The signal passes with roughly 27 %, this low number is caused by the way of generating the semi-leptonic decay channel events, see section 4.2. The backgrounds with the largest efficiencies are the HZ process and the di-leptonic decay channel. The other three backgrounds are suppressed very well with less than 1 % remaining.

Table 4.1: Pre-selection cut flow table with the yields for all of the data samples. The scaled events correspond to the FCC-ee planned luminosity goal per year $L_{\text{int}} = 0.2\ \text{ab}^{-1}$. All of the cuts are applied in succession.

cuts	$t\bar{t}$ semi-lep	$t\bar{t}$ di-lep	$t\bar{t}$ had	HZ	ZZ	WW
Generated	95000	100000	90000	100000	95000	100000
Scaled	46234	11083	48240	26000	100000	1000000
1lep&4ptcles	12914 ± 79	1572 ± 13	941 ± 22	2964 ± 27	5923 ± 78	303420 ± 1741
m_{4j}>150	12877 ± 79	646 ± 8	941 ± 22	1984 ± 22	2607 ± 52	13310 ± 364
m_{4j}<270	12856 ± 79	646 ± 8	42 ± 4	1396 ± 19	1607 ± 41	11890 ± 344
m_{2j}^{min}>10	12854 ± 79	550 ± 7	42 ± 4	1336 ± 18	1332 ± 37	9510 ± 308
$m_{2j}^{\text{2nd min}}$>20	12853 ± 79	476 ± 7	42 ± 4	1311 ± 18	1256 ± 36	8560 ± 292
E_{lepton}<100	12844 ± 79	476 ± 7	42 ± 4	1258 ± 18	1120 ± 34	6480 ± 254
$\not{E}^{\text{reco}}$>20	12822 ± 78	476 ± 7	41 ± 4	1222 ± 17	1010 ± 32	6370 ± 252

Table 4.2: Pre-selection cut flow table with the efficiencies for all of the data samples. The scaled events correspond to the FCC-ee planned luminosity goal per year $L_{\text{int}} = 0.2\ \text{ab}^{-1}$. All of the cuts are applied in succession.

cuts	$t\bar{t}$ semi-lep	$t\bar{t}$ di-lep	$t\bar{t}$ had	HZ	ZZ	WW
Generated	95000	100000	90000	100000	95000	100000
Scaled	46234	11083	48240	26000	100000	1000000
1 lep & 4 ptcles	$0.2793^{+0.0021}_{-0.0021}$	$0.1418^{+0.0033}_{-0.0033}$	$0.0195^{+0.0006}_{-0.0006}$	$0.1140^{+0.0020}_{-0.0020}$	$0.0592^{+0.0008}_{-0.0007}$	$0.3034^{+0.0005}_{-0.0005}$
$m_{4j} > 150$	$0.2785^{+0.0021}_{-0.0021}$	$0.0583^{+0.0023}_{-0.0022}$	$0.0195^{+0.0006}_{-0.0006}$	$0.0763^{+0.0017}_{-0.0016}$	$0.0261^{+0.0005}_{-0.0005}$	$0.0133^{+0.0001}_{-0.0001}$
$m_{4j} < 270$	$0.2781^{+0.0021}_{-0.0021}$	$0.0583^{+0.0023}_{-0.0022}$	$0.0009^{+0.0001}_{-0.0001}$	$0.0537^{+0.0014}_{-0.0014}$	$0.0161^{+0.0004}_{-0.0004}$	$0.0119^{+0.0001}_{-0.0001}$
$m_{2j}^{\text{min}} > 10$	$0.2780^{+0.0021}_{-0.0021}$	$0.0496^{+0.0021}_{-0.0020}$	$0.0009^{+0.0001}_{-0.0001}$	$0.0514^{+0.0014}_{-0.0014}$	$0.0133^{+0.0004}_{-0.0004}$	$0.0095^{+0.0001}_{-0.0001}$
$m_{2j}^{\text{2nd min}} > 20$	$0.2780^{+0.0021}_{-0.0021}$	$0.0429^{+0.0020}_{-0.0019}$	$0.0009^{+0.0001}_{-0.0001}$	$0.0504^{+0.0014}_{-0.0013}$	$0.0126^{+0.0004}_{-0.0003}$	$0.0086^{+0.0001}_{-0.0001}$
$E_{\text{lepton}} < 100$	$0.2778^{+0.0021}_{-0.0021}$	$0.0429^{+0.0020}_{-0.0019}$	$0.0009^{+0.0001}_{-0.0001}$	$0.0484^{+0.0013}_{-0.0013}$	$0.0112^{+0.0003}_{-0.0003}$	$0.0065^{+0.0001}_{-0.0001}$
$\not{E}^{\text{reco}} > 20$	$0.2773^{+0.0021}_{-0.0021}$	$0.0429^{+0.0020}_{-0.0019}$	$0.0008^{+0.0001}_{-0.0001}$	$0.0470^{+0.0013}_{-0.0013}$	$0.0101^{+0.0003}_{-0.0003}$	$0.0064^{+0.0001}_{-0.0001}$

Chapter 5

Statistical method comparison

In this chapter, three different statistical tests are compared to isolate a purer sample of top quarks. The choice of the tests is explained in section 5.1. Section 5.2 is dedicated to determining reference values for the tests. The results for the tests are discussed in section 5.3.

5.1 Methods

The next step in the analysis is defining a statistical test to find out which events are good candidates for the $t\bar{t}$ semi-leptonic decay. The test chosen verifies that the invariant masses of the decay products come from top quarks and W bosons. This is achieved by summing over χ^2 terms. The χ^2-test is chosen for its simplicity and efficiency, it is also a common choice in top quark physics and can be used to compare results to other detectors and colliders.

In this thesis multiple test variables are compared, these are defined as:

$$\chi_3^2 = \left(\frac{\mathrm{m}_{\mathrm{t}}^{\mathrm{had}} - \mathrm{m}_{\mathrm{t,reco}}^{\mathrm{had}}}{\sigma_{\mathrm{t}}^{\mathrm{had}}}\right)^2 + \left(\frac{\mathrm{m}_{\mathrm{t}}^{\mathrm{lep}} - \mathrm{m}_{\mathrm{t,reco}}^{\mathrm{lep}}}{\sigma_{\mathrm{t}}^{\mathrm{lep}}}\right)^2 + \left(\frac{\mathrm{m}_{\mathrm{W}}^{\mathrm{had}} - \mathrm{m}_{\mathrm{W,reco}}^{\mathrm{had}}}{\sigma_{\mathrm{W}}^{\mathrm{had}}}\right)^2 \tag{5.1}$$

$$\chi_4^2 = \chi_3^2 + \left(\frac{\mathrm{m}_{\mathrm{W}}^{\mathrm{lep}} - \mathrm{m}_{\mathrm{W,reco}}^{\mathrm{lep}}}{\sigma_{\mathrm{W}}^{\mathrm{lep}}}\right)^2 \tag{5.2}$$

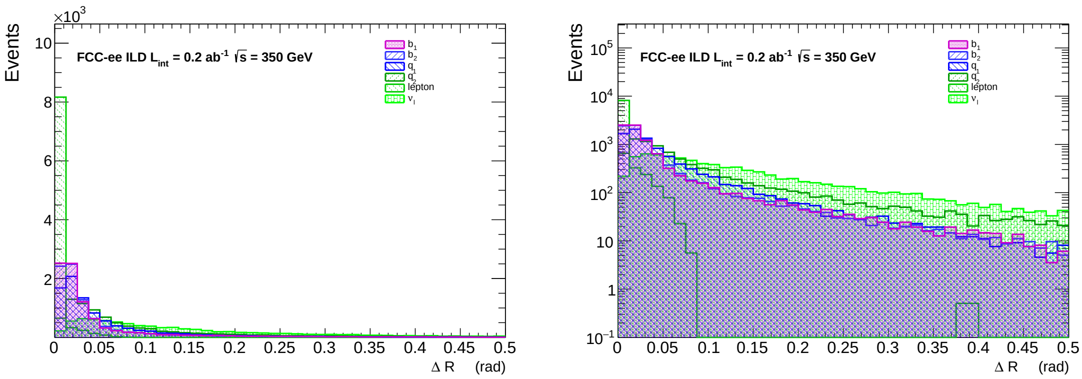
$$\chi_5^2 = \chi_4^2 + \left(\frac{(\mathrm{m}_{\mathrm{t}}^{\mathrm{had}} + \mathrm{m}_{\mathrm{t}}^{\mathrm{lep}}) - (\mathrm{m}_{\mathrm{t,reco}}^{\mathrm{had}} + \mathrm{m}_{\mathrm{t,reco}}^{\mathrm{lep}})}{\sigma_{\mathrm{t}}^{\mathrm{had}} + \sigma_{\mathrm{t}}^{\mathrm{lep}}}\right)^2 \tag{5.3}$$

In these equations, the subscripts *reco* indicate the values obtained by taking the data from one event. The other parameters are reference values, which are determined in section 5.2. The χ^2 terms are computed for each event by trying all of the possible jet combinations to reconstruct the W bosons and top quarks. The lowest overall $\chi^2_{3,4,5}$ combination is saved.

It is important to note that a multitude of methods, outside of the scope of this thesis, can be used to reconstruct candidate events. The template method and matrix element methods are described in [16]. Another approach uses the fact that the decay mode can be treated as a nonlinear set of equations with eight unknowns and eight constraints [23]. One can also use tools such as artificial neural networks or boosted decision trees. The comparison of alternative methods to the sum of χ^2 term used in this thesis is beyond the scope of this work but should be useful in the future.

5.2 Determination of the χ^2 parameters

For the χ^2 terms eight reference values are necessary. These are obtained by taking the best possible events after reconstruction using the Monte Carlo collision data only. The reconstructed particles and jets are compared to the MC particles. The best possible combination is chosen by comparing the spatial angles ΔR and using the particle ID's from the MC data. If all of the MC particles' four vectors are situated within a cone of $\Delta R < 0.1$ rad of the reconstructed particles and jets, the event is saved. The missing energy-momentum four-vector is assumed to be the neutrino. The distribution for the angles is given in Figure 5.1



(a) (b)

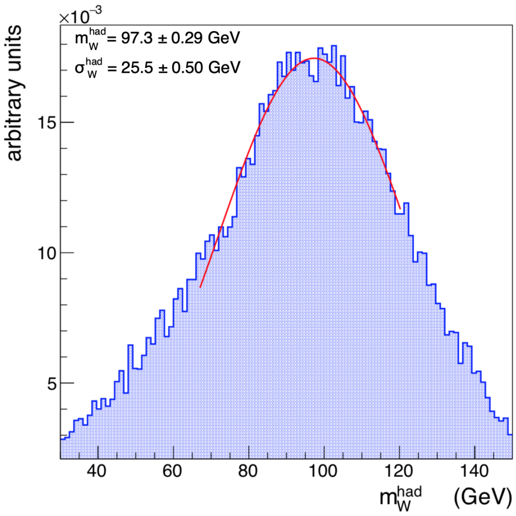
Figure 5.1: Distribution of the angles between the reconstructed particles and jets against the real MC event data. The lepton angle is very small, indicating very high resolution. The neutrino has the largest tail, this is to be expected as not only the neutrino high energy can lead to missing energy. The good events require $\Delta R < 0.1$ for all particles.

In the distribution, the lepton stands out with a very high peak near zero. The jets, indicated with $b_{1,2}$ and $q_{1,2}$ for the original quarks from the collision, are mostly located towards zero and mostly die down by $\Delta R \sim 0.15$ rad. The neutrino distribution is skewed towards larger angles and will eliminate the largest amount of events. The selection efficiency for the good combinations is,

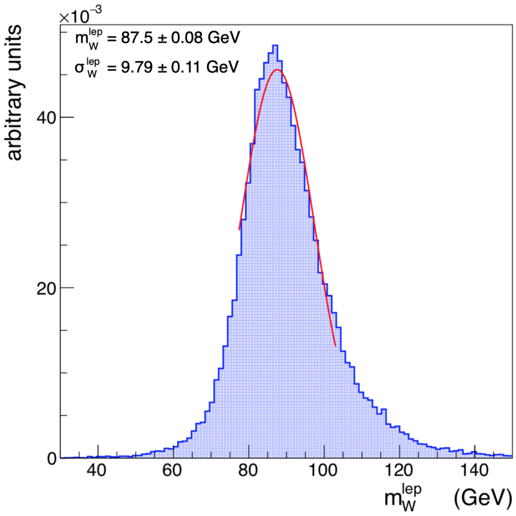
$$\varepsilon_{GC} = 0.0508^{+0.000221}_{-0.000220}. \tag{5.4}$$

This value is quite low and leaves few events for the fit. As a consequence a larger sample is necessary for this purpose, it is not really a problem since it will not affect the analysis efficiency.

After this requirement is met, the mass distributions are created for the two top quarks and W bosons. The histograms are fitted with a Gaussian function in the range centered around the mean with a single root-mean-square deviation to fit the peak. The W boson masses are shown in Figure 5.2. As can be seen the central values are quite different from the expected $\mathrm{m_W} = 80.379 \pm 0.012$ GeV [4], they remain within uncertainties as the spread is very large. The leptonic distribution is much narrower compared to the hadronic W.



(a) Hadronic W boson mass distribution for good combination reconstructed events.

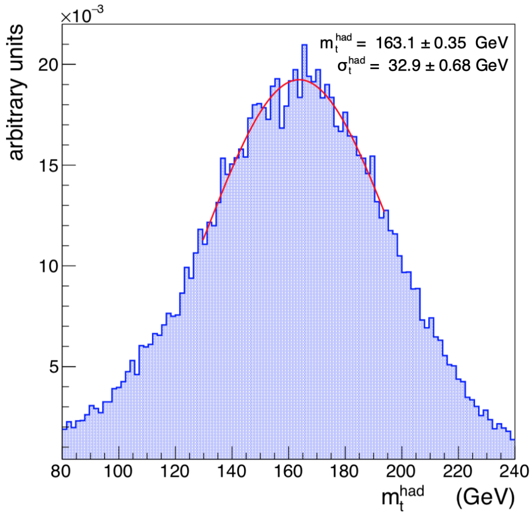


(b) Leptonic W boson mass distribution for good combination reconstructed events.

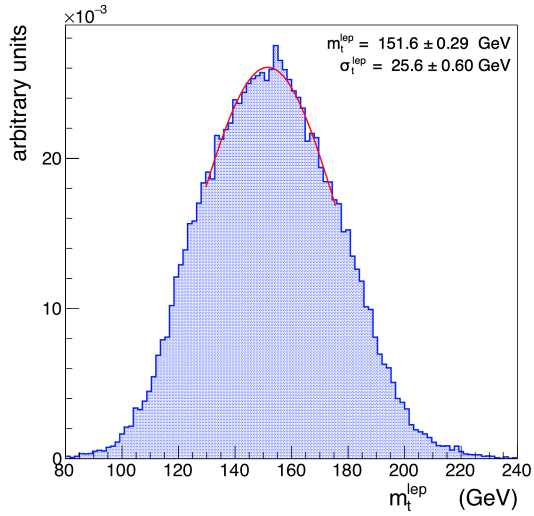
Figure 5.2

The reconstructed top quark masses are shown in Figure 5.3. The same trend as for the W bosons can be noticed, the lepton distribution appears to be narrower than the hadrons. The effect is less striking however because a b-jet with a large spread is included in the reconstruction. The fitted masses are lower on average than the expected value $m_{\mathrm{t}} = 173.0 \pm 0.4$ GeV, but they are consistent due to the large uncertainties.

The fitted parameters are summarised in Table 5.1. The quantity $\chi^2/\#\mathrm{dof}$ gives a value for the goodness of fit. A value closer to 1 indicates that the fit is good. All of



(a) Hadronic top quark mass distribution for good combination reconstructed events.



(b) Leptonic top quark mass distribution for good combination reconstructed events.

Figure 5.3

the fits are quite good except for the leptonic W. It is logical that this distribution performs the worst, as the neutrino is the hardest particle to reconstruct. It is also quite visible that the distribution of the leptonic W boson is different from a Gaussian distribution in Figure 5.2b.

Table 5.1: Resolution parameters recovered for the χ^2 terms.

Distribution	μ **(GeV)**	σ **(GeV)**	$\chi^2/\#$**dof**
W_{had}	97.3 ± 0.29	25.5± 0.50	1.12
W_{lep}	87.5 ± 0.08	9.79 ± 0.11	8.44
t_{had}	163.1± 0.35	32.9± 0.68	1.70
t_{lep}	151.6± 0.29	25.6± 0.60	0.89

5.3 Results

The test variables are compared by calculating the background efficiency against the signal efficiency. This is done by computing both values for a lot of different $\chi^2_{3,4,5}$ cuts. The produced curve is called a ROC curve. The name comes from the term Receiver Operating Characteristic curve, which shows the ability from an operator to discriminate true from false positives. In this case, a true positive is a semi-leptonic decay and a false positive any of the background processes. The goal is to produce a curve that approaches the top left corner i.e. full signal efficiency and no background. The ROC curve for the tested variables is shown in Figure 5.4. In this case, the variable χ_3^2 performs the best for higher backgrounds, up to $\sim 10\%$ the curves are relatively similar.

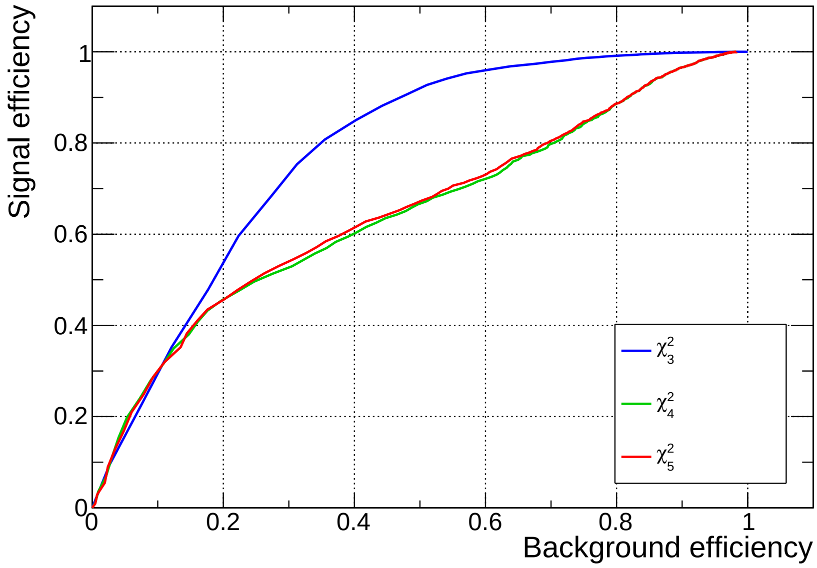


Figure 5.4: ROC curve for the different statiscal test variables. The blue line representing χ_3^2 is closest to the ideal point with all signal and no background.

To compare the variables on equal footing the same background efficiency is chosen:

$$\text{Background efficiency} = 10\,\% \tag{5.5}$$

The signal efficiencies are very similar at this value and might seem like a bad choice, but there is another argument for a low background efficiency. At this point it is possible to compute the measured cross-section and its statistical error for the $t\bar{t}$ process,

$$\sigma = \frac{N_{\text{total}} - N_{\text{BG}}}{\varepsilon \cdot L_{\text{int}}}. \tag{5.6}$$

Here N_{total} refers to the total event passing the cut, N_{BG} is the number of background events passing the cut. The ε represents the efficiency of the selection. The analysis concerns only the semi-leptonic decay mode so the right-hand side also has to be divided by the branching ratio. The statistical uncertainty can be calculated, it is assumed the efficiency and luminosity are known perfectly and the uncertainty for the event count is taken as Poissonian.

$$\left(\frac{\Delta\sigma}{\sigma}\right)_{\text{stat}} = \sigma^{-1} \cdot \left(\sqrt{\frac{N_{\text{total}}}{(\varepsilon L_{\text{int}}\text{BR})^2} + \frac{N_{\text{BG}}}{(\varepsilon L_{\text{int}}\text{BR})^2}}\right). \tag{5.7}$$

The statistical errors on the cross-sections are plotted as a function of the χ^2 cuts in Figure 5.5. The errors starts off at a high value then decreases to eventually reach a constant value. Note that the cuts for $\chi^2_{3,4,5}$ are different for the same background efficiency.

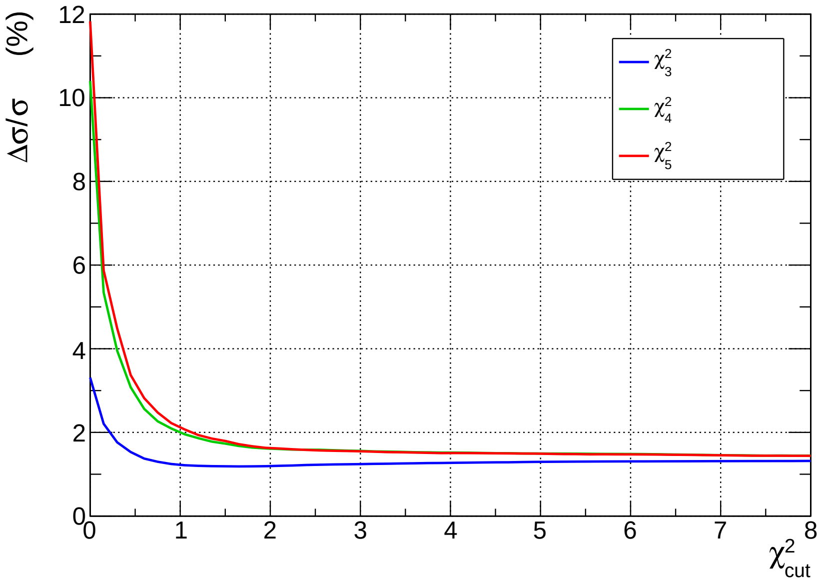


Figure 5.5: Statistical error on the cross-section measurement for the $t\bar{t}$ semi-leptonic process.

The results for the χ^2 comparison are in Table 5.2. Purity is defined as the signal event count divided by the total count that passed the cuts. In the end, χ^2_3 was chosen as it has the best ROC curve with the cut:

$$\chi^2_3 < 0.45. \tag{5.8}$$

Table 5.2: Comparison of the performance of the different test variables.

Variable	**Cut**	**Signal efficiency** (%)	**Purity** (%)	$(\Delta\sigma/\sigma)_{\text{stat}}$ (%)
χ^2_3	0.45	29.47	80.3	1.81
χ^2_4	1.35	29.93	80.3	1.92
χ^2_5	1.50	29.97	80.3	1.91

The yields are shown in Table 5.3. The signal yields are all above 4000 events, while the total backgrounds yield is about 900 events. The largest background source is still the WW process. Note that for $\chi^2_{4,5}$ the two other decay modes of the $t\bar{t}$ process are almost completely eliminated.

Table 5.3: Cut flow table for the yield of the different test variables. The background efficiency is set to 10% after the pre-selection.

cuts	$t\bar{t}$ semi-lep	$t\bar{t}$ di-lep	$t\bar{t}$ had	HZ	ZZ	WW
Scaled	46234	11083	48240	26000	100000	1000000
Pre-selection	12822 ± 78	476 ± 7	41 ± 4	1222 ± 17	1010 ± 32	6370 ± 252
$\chi_3^2 < 0.45$	4520 ± 46	99 ± 3	15 ± 2	203 ± 7	95 ± 10	690 ± 83
$\chi_4^2 < 1.35$	4049 ± 44	3 ± 0	2 ± 1	131 ± 5	54 ± 7	800 ± 89
$\chi_5^2 < 1.5$	4096 ± 44	3 ± 0	3 ± 1	134 ± 5	56 ± 7	810 ± 90

The final efficiencies of the selection with respect to the scaled events are shown in Figure 5.4. After the selection about $8.5 - 10$ % of the signal remains. All of the backgrounds are reduced to below 1 %, which is a very good result. The largest contribution is the same as after the pre-selection, the HZ process.

Table 5.4: Cut flow table for the efficiency of the different test variables. The background efficiency is set to 10% after the pre-selection.

cuts	$t\bar{t}$ semi-lep	$t\bar{t}$ di-lep	$t\bar{t}$ had	HZ	ZZ	WW
Scaled	46234	11083	48240	26000	100000	1000000
Pre-selection	$0.2773^{+0.0021}_{-0.0021}$	$0.0429^{+0.0020}_{-0.0019}$	$0.0008^{+0.0001}_{-0.0001}$	$0.0470^{+0.0013}_{-0.0013}$	$0.0101^{+0.0003}_{-0.0003}$	$0.0064^{+0.0001}_{-0.0001}$
$\chi_3^2 < 0.45$	$0.0978^{+0.0014}_{-0.0014}$	$0.0089^{+0.0009}_{-0.0009}$	$0.0003^{+0.0001}_{-0.0001}$	$0.0078^{+0.0006}_{-0.0005}$	$0.0009^{+0.0001}_{-0.0001}$	$0.0007^{+0.0000}_{-0.0000}$
$\chi_4^2 < 1.35$	$0.0876^{+0.0013}_{-0.0013}$	$0.0003^{+0.0002}_{-0.0001}$	$0.0000^{+0.0000}_{-0.0000}$	$0.0050^{+0.0005}_{-0.0004}$	$0.0005^{+0.0001}_{-0.0001}$	$0.0008^{+0.0000}_{-0.0000}$
$\chi_5^2 < 1.5$	$0.0886^{+0.0013}_{-0.0013}$	$0.0003^{+0.0002}_{-0.0001}$	$0.0001^{+0.0000}_{-0.0000}$	$0.0052^{+0.0005}_{-0.0004}$	$0.0006^{+0.0001}_{-0.0001}$	$0.0008^{+0.0000}_{-0.0000}$

Chapter 6

Effects of top mass and centre of mass energies on measurements

6.1 Top quark and W boson Measurement projections

After the procedure of selecting candidate events out of all of the samples, the data can be used to recover the mass measurements for the top quarks and W bosons. First, let's take a look at the χ_3^2 distribution after the pre-selection. For this distribution a distinction was made between bad combinations (BC) and good combinations (GC). The MC information was used to verify that $\Delta R < 0.1$ rad for all particles. Figure 6.1. The efficiency on GC events out of the remaining semi-leptonic sample is,

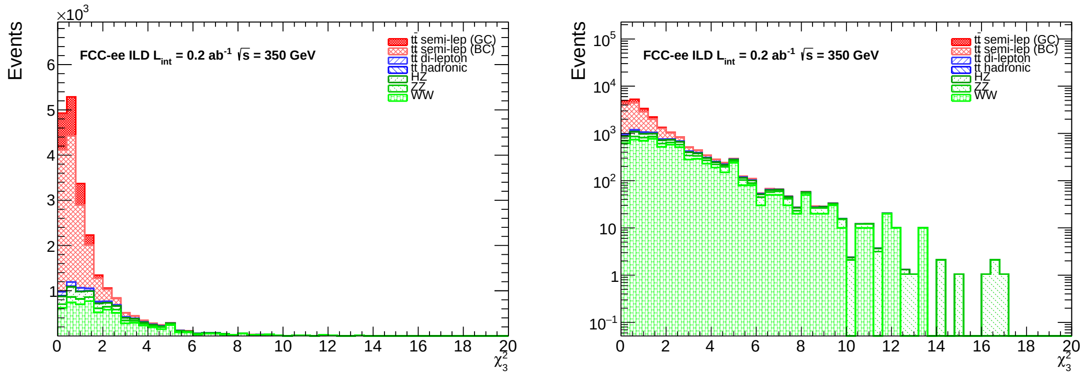
$$\varepsilon_{GC}^{\text{before cut}} = 0.1928^{+0.0034}_{-0.0034}. \tag{6.1}$$

The efficiency for GC events after the cut on χ_3^2 is,

$$\varepsilon_{GC} = 0.1918^{+0.0059}_{-0.0058}. \tag{6.2}$$

It is quite logical to have fairly similar results here as the cut is made on a very small value of χ_3^2. The distribution in Figure 6.1 for BC and GC are similar, but the BC events have a larger tail than the GC.

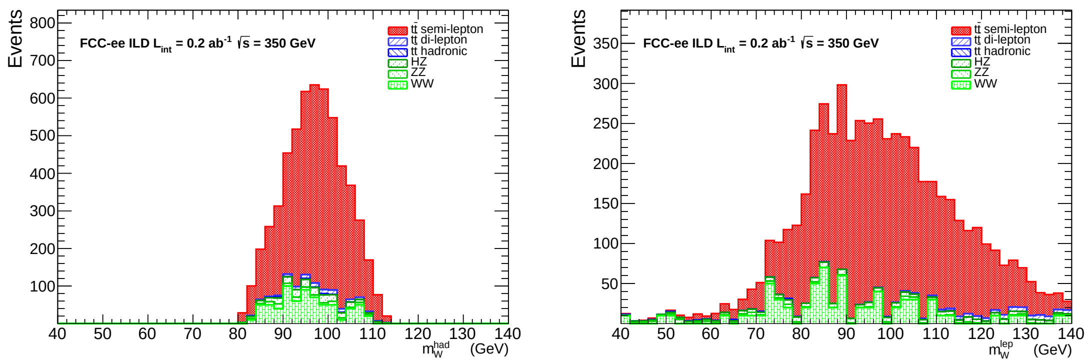
The reconstructed W boson masses are shown in Figure 6.2a and 6.2b. The hadronic W boson consists of a relatively narrow spectrum after the selection centered around $\sim$ 96 GeV. This value is close to the resolution parameters obtained in the χ^2 analysis in Table 5.1. On the other hand, the distribution for the leptonic W boson has a large spread. The histogram has multiple peaks and dips between values of 80 to 90 GeV, these seem to be caused by the background fluctuations. The signal



(a) (b)

Figure 6.1: χ^2_3 distribution after the pre-selection. Good combinations events are marked in red and bad combinations in pink.

and background also have a large tail towards higher mass. It is logical for this distribution to be less precise as the test statistic χ^2_3 does not include the leptonic W boson term, see 5.1.

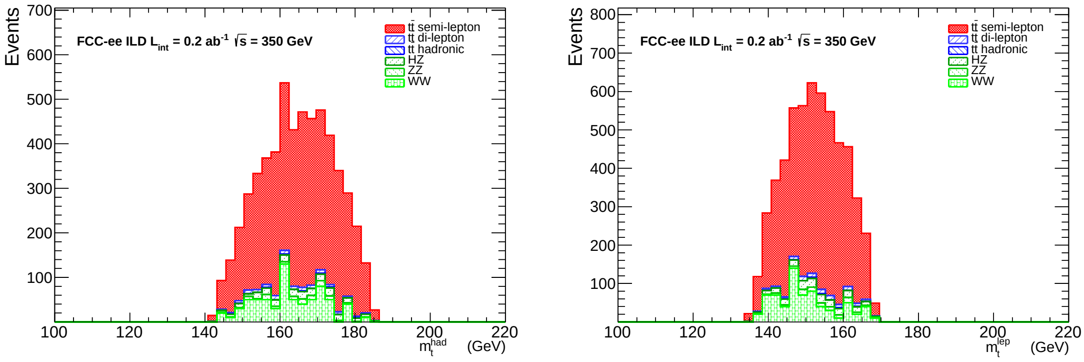


(a) Reconstructed hadronic W boson mass. (b) Reconstructed leptonic W boson mass.

Figure 6.2

Figures 6.3a and 6.3b show the reconstructed top quark masses. The hadronic mass spectrum has a slightly larger spread by ~ 10 GeV. The hadronic distribution peaks overall close to the expected top quark mass of 173 GeV but contains peaks and dips near that value due to the inconsistent backgrounds. The leptonic spectrum is clean in comparison but is centered at a lower value of 150 GeV, which corresponds with the resolution in Table 5.1.

The figures above show that the choice for χ^2_3 with a hard cut produces wide distributions for the invariant masses, but with sharp cut-offs at the edges. The only exception is the leptonic W boson which is not included in the test variable, so it is



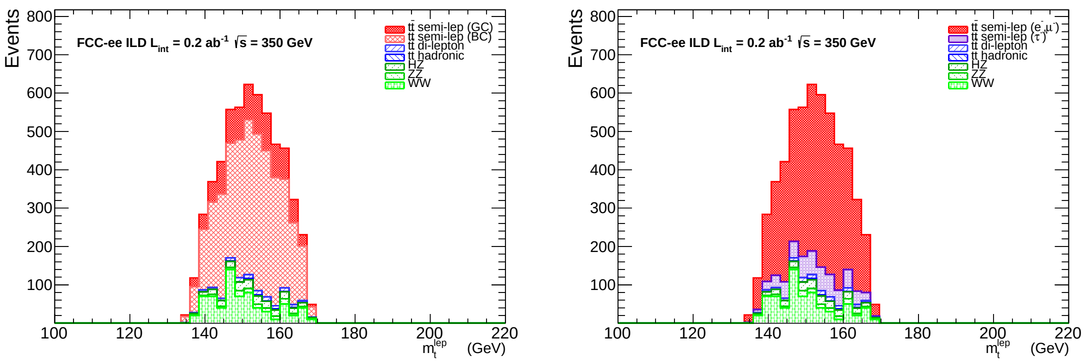
(a) Reconstructed hadronic top quark mass. (b) Reconstructed leptonic top quark mass.

Figure 6.3

consistent with the expectations. The terms that are included in χ_3^2 remain within the resolutions from Table 5.1 for $\mu \pm 1\,\sigma$.

All of the distributions above can be separated into GC and GC events as for the χ_3^2. It is also possible to discriminate events which have decayed from a τ^- lepton. The top quark in the leptonic mode is used as a demonstration in Figures 6.4a and 6.4b. The efficiency of events that have decayed via τ^- leptons after all the cuts is,

$$\varepsilon_\tau = 0.1190^{+0.0048}_{-0.0047}. \tag{6.3}$$



(a) Reconstructed leptonic top quark mass. Good combination (GC) are in red and bad combination in pink (BC). (b) Reconstructed leptonic top quark mass. Decays via τ^- are in purple and decays via e^-, μ^- in red.

Figure 6.4

6.2 Effect of top mass and centre of mass energies

6.2.1 Measurements

The effect of multiple values of m_{t} and $\sqrt{s}$ on the invariant masses is now studied. The distributions are not stacked in any of the plots that follow.

First, the variation of centre of mass energy is considered for the fixed mass,

$$m_{\mathrm{t}} = 173\,\mathrm{GeV}. \tag{6.4}$$

The χ_3^2 distribution is shown in Figure 6.5. The spectra are all quite similar with lower events at lower energies. This behaviour is expected as the cross-section rises as the energy grows. The ratios show a small rising trend at lower energies until $\chi_3^2 = 4$, then the ratios fluctuate a lot. The last fluctuations are due to the low event count and can be suppressed by using more data or simply be ignored as it is past the cut.

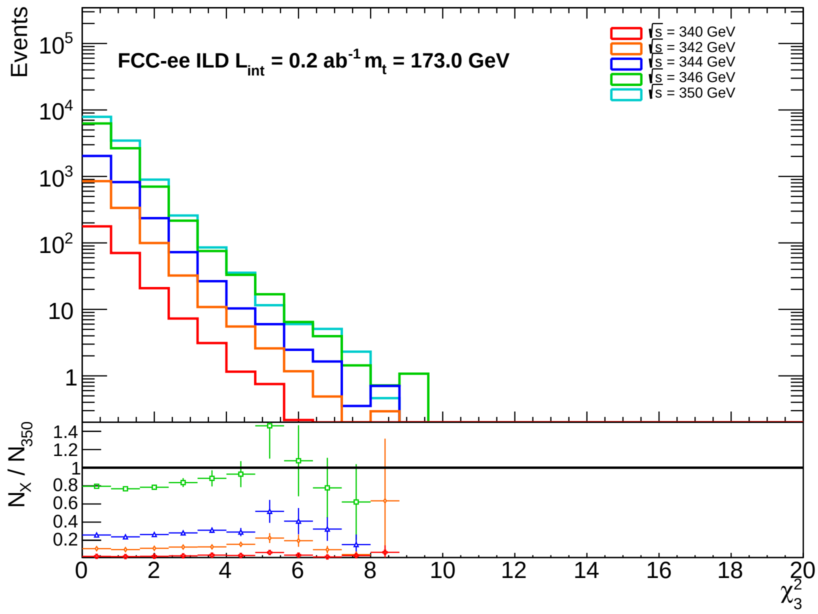


Figure 6.5: χ_3^2 distribution for fixed top quark mass $m_{\mathrm{t}} = 173$ GeV at different centre of mass energies. The bottom panel shows the ratio between the histograms with the energies below $\sqrt{s} = 350$ GeV.

The reconstructed top quark mass in the hadronic mode is shown in Figure 6.6. Here all of the distributions have very similar shapes. Small fluctuations can be observed for the different energies which grow towards the edges. This is caused by the low statistics away from the peak.

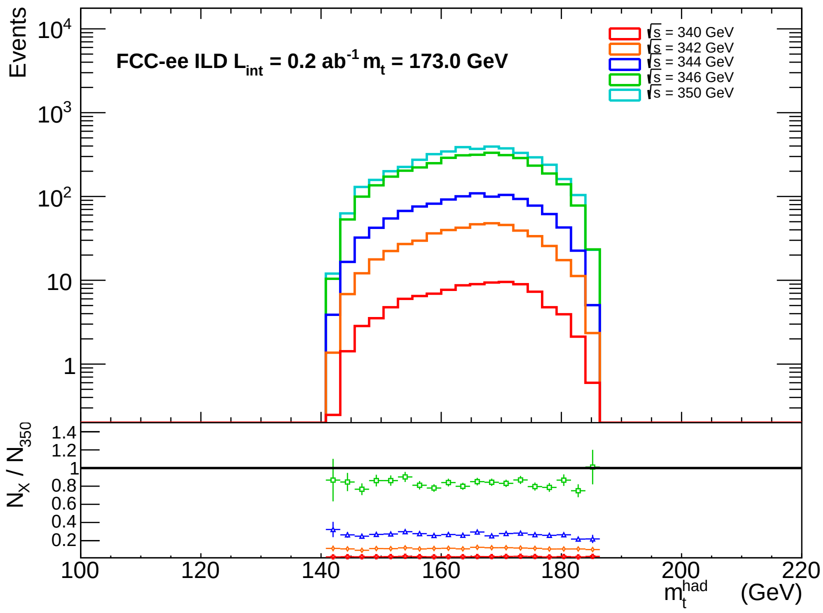


Figure 6.6: Invariant hadronic top quark mass at different centre of mass energies. The bottom panel shows the ratio between the histograms with the energies below $\sqrt{s} = 350$ GeV.

The leptonic top quark mass is presented in Figure 6.7. The same trend can be observed as for the hadronic top quark reconstruction.

Overall no conclusive distinction can be made from these figures at different energies besides the expected lower event count at lower $\sqrt{s}$.

Let's consider the fixed centre of mass energy,

$$\sqrt{s} = 350 \text{ GeV}. \tag{6.5}$$

The χ_3^2 distribution in Figure 6.8 shows very similar behaviour for all masses. The fractions increase with larger values of χ_3^2, this is due to the lower event count.

The top quark hadronic mass shows very stable behaviour in the fractions in Figure 6.9. Small oscillations seem to arise but these are all within uncertainties of zero so no conclusion can be taken.

Figure 6.10 shows the leptonic invariant mass of the top quarks. Here a larger discrepancy can be observed at the edges but it still isn't significant.

For fixed centre of mass energies no deviations in the distributions can be found.

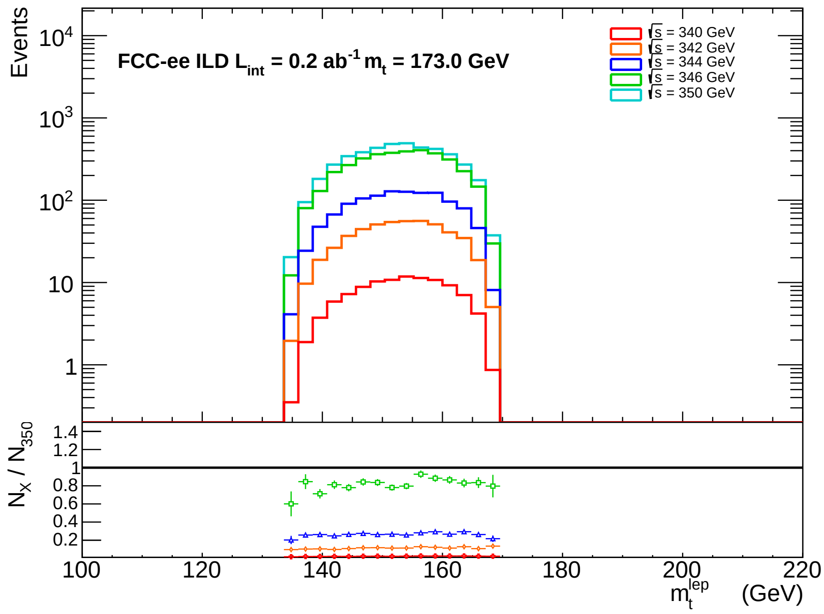


Figure 6.7: Invariant leptonic top quark mass at different centre of mass energies. The bottom panel shows the ratio between the histograms with the energies below $\sqrt{s} = 350$ GeV.

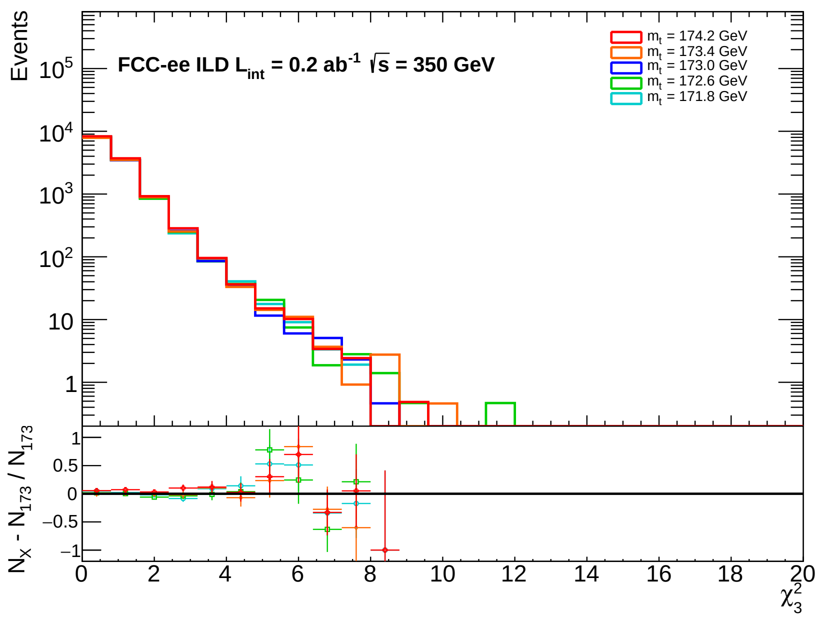


Figure 6.8: χ^2_3 distribution for different MC simulated top quark masses. The bottom panel shows the fraction between the histograms with the other masses.

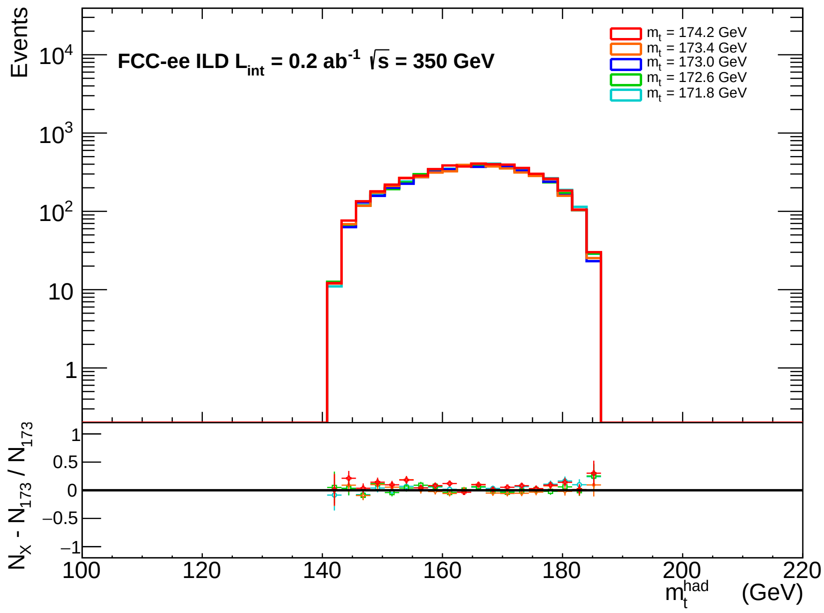


Figure 6.9: Invariant hadronic top quark mass at different MC simulated top quark masses. The bottom panel shows the fraction between the histograms with the other masses.

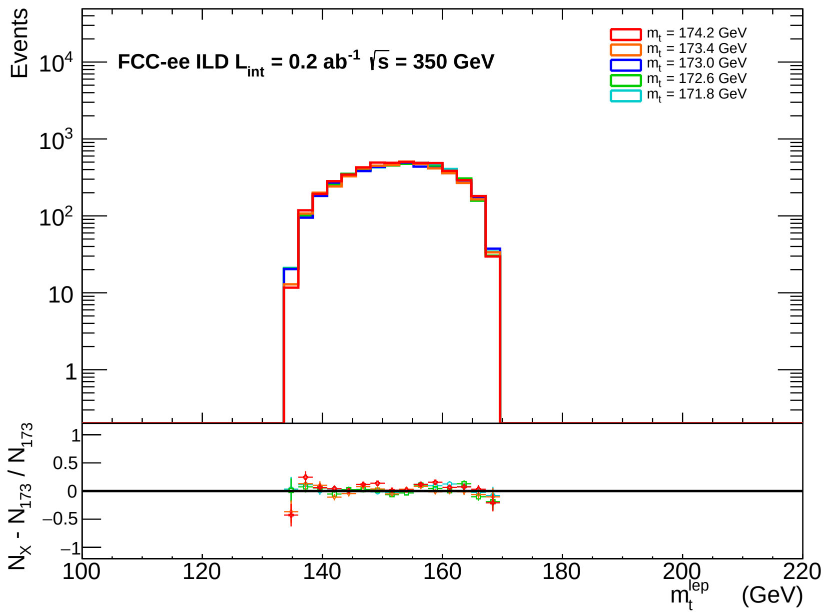


Figure 6.10: Invariant leptonic top quark mass at different MC simulated top quark masses. The bottom panel shows the fraction between the histograms with the other masses.

6.2.2 Statistics

In this section, the influence of the top quark mass and centre of mass energies is considered for the statistical variables concerning the analysis.

Table 6.1 gives the yield after the selection for the $t\bar{t}$ semi-leptonic samples. The yield is low for lower energies and rises until $\sqrt{s} = 344$ GeV, at that point it slowly decreases again. This is to be expected as the cross-section peaks near this value in Figure 2.4 and remains roughly stable in the continuum. The trend concerning the top quark mass effect is not clear as all values remain within uncertainty of each other for the same $\sqrt{s}$.

Table 6.1: Yield of the $t\bar{t}$ semi-leptonic samples after being scaled to the integrated luminosity $L_{\text{int}} = 0.2$ ab^{-1} for different values of $\sqrt{s}$ and m_{t}. The units for the energy and mass are given in GeV.

$\sqrt{s}$ \ m_{t}	171.8	172.6	173.0	173.4	174.2
340	587 ± 5	265 ± 2	104 ± 1	0 ± 0	0 ± 0
342	1520 ± 14	683 ± 6	505 ± 4	338 ± 3	26 ± 0
344	4177 ± 41	1866 ± 18	1183 ± 11	799 ± 7	422 ± 4
346	4690 ± 46	4635 ± 46	3653 ± 36	2337 ± 23	966 ± 9
350	4562 ± 46	4463 ± 45	4404 ± 45	4431 ± 45	4721 ± 47

In Table 6.2 the efficiency of the selection on the signal is compared. Here for all energies below $\sqrt{s} < 350$ GeV the values remain within uncertainties of $\varepsilon = 10$ %. At lower energies the selection seems to perform slightly better than at higher energies.

The purity of the data is compared in Table 6.3. The backgrounds are assumed to be constant with the top quark mass and are therefore using the same samples for each row. As explained in section 4.2, all of the background processes are located in the continuum region and have fairly stable cross-sections. The signal process cross-section is at the threshold, so it is expected to have better purity as the signal increases. This trend is clearly visible in the table, with more than doubling of the purity between the lower energies and the higher values.

The statistical uncertainty on the cross-section measurements are compared in Table 6.4. In this table it is very obvious that the uncertainty becomes better for higher $\sqrt{s}$. The top quark mass does not seem to influence this value as the changes are very small.

In conclusion, the centre of mass energy has a significant impact for some analysis variables, while the top quark mass seems to have little effect.

Table 6.2: Efficiency of the $t\bar{t}$ semi-leptonic samples after being scaled to the integrated luminosity $L_{\mathrm{int}} = 0.2\ \mathrm{ab}^{-1}$ for different values of $\sqrt{s}$ and m_{t}. The units for the energy and mass are given in GeV.

$\sqrt{s}$ \ m_{t}	171.8	172.6	173.0	173.4	174.2
340	$0.1034^{+0.0041}_{-0.0040}$	$0.1031^{+0.0061}_{-0.0058}$	$0.1011^{+0.0098}_{-0.0090}$	$0.0000^{+0.0000}_{-0.0000}$	$0.0000^{+0.0000}_{-0.0000}$
342	$0.1039^{+0.0026}_{-0.0025}$	$0.1023^{+0.0038}_{-0.0036}$	$0.1033^{+0.0044}_{-0.0043}$	$0.1011^{+0.0053}_{-0.0051}$	$0.1020^{+0.0203}_{-0.0172}$
344	$0.0994^{+0.0015}_{-0.0015}$	$0.1011^{+0.0022}_{-0.0022}$	$0.1007^{+0.0028}_{-0.0027}$	$0.1010^{+0.0034}_{-0.0033}$	$0.1026^{+0.0048}_{-0.0046}$
346	$0.1003^{+0.0014}_{-0.0014}$	$0.1003^{+0.0014}_{-0.0014}$	$0.1016^{+0.0016}_{-0.0016}$	$0.0995^{+0.0020}_{-0.0019}$	$0.1012^{+0.0031}_{-0.0030}$
350	$0.0954^{+0.0014}_{-0.0013}$	$0.0955^{+0.0014}_{-0.0014}$	$0.0953^{+0.0014}_{-0.0014}$	$0.0963^{+0.0014}_{-0.0014}$	$0.0972^{+0.0014}_{-0.0013}$

Table 6.3: Purity of the $t\bar{t}$ semi-leptonic samples after being scaled to the integrated luminosity $L_{\mathrm{int}} = 0.2\ \mathrm{ab}^{-1}$ for different values of $\sqrt{s}$ and m_{t}. The units for the energy and mass are given in GeV.

$\sqrt{s}$ \ m_{t}	171.8	172.6	173.0	173.4	174.2
340	$0.3166^{+0.0109}_{-0.0107}$	$0.1730^{+0.0099}_{-0.0094}$	$0.0760^{+0.0075}_{-0.0068}$	$0.0000^{+0.0008}_{-0.0000}$	$0.0000^{+0.0008}_{-0.0000}$
342	$0.5801^{+0.0096}_{-0.0097}$	$0.3834^{+0.0116}_{-0.0114}$	$0.3148^{+0.0117}_{-0.0115}$	$0.2352^{+0.0114}_{-0.0110}$	$0.0236^{+0.0050}_{-0.0041}$
344	$0.8028^{+0.0055}_{-0.0056}$	$0.6453^{+0.0088}_{-0.0089}$	$0.5356^{+0.0106}_{-0.0106}$	$0.4379^{+0.0116}_{-0.0116}$	$0.2915^{+0.0121}_{-0.0118}$
346	$0.8101^{+0.0051}_{-0.0052}$	$0.8082^{+0.0051}_{-0.0053}$	$0.7686^{+0.0061}_{-0.0062}$	$0.6800^{+0.0079}_{-0.0080}$	$0.4678^{+0.0110}_{-0.0110}$
350	$0.7962^{+0.0053}_{-0.0054}$	$0.7926^{+0.0054}_{-0.0055}$	$0.7904^{+0.0054}_{-0.0055}$	$0.7914^{+0.0054}_{-0.0055}$	$0.8017^{+0.0051}_{-0.0052}$

6.2.3 cross-section measurements

The measured cross-section can be computed using the values in the tables from section 6.2.2 using equation (5.6). The cross-sections are computed for the central top quark mass $\mathrm{m_t} = 173$ GeV. The red curve shows the measured value against the latest predictions [15] in Figure 6.11. The measurements are very close to the predicted value which is a good result. This result is to be expected as the reference values to produce the red curve was the latest prediction.

Table 6.4: Statistical uncertainty on the cross-section measurement $(\Delta\sigma/\sigma)_{\text{stat}}$ (%) of the $t\bar{t}$ semi-leptonic samples after being scaled to the integrated luminosity $L_{\text{int}} = 0.2$ ab^{-1} for different values of $\sqrt{s}$ and m_{t}. The units for the energy and mass are given in GeV.

$\sqrt{s}$ \ m_{t}	171.8	172.6	173.0	173.4	174.2
340	9.51	19.94	49.28	-	-
342	4.01	7.85	10.29	14.9	177.23
344	1.89	3.35	4.81	6.68	11.78
346	1.77	1.78	2.09	2.88	5.82
350	1.82	1.85	1.86	1.86	1.78

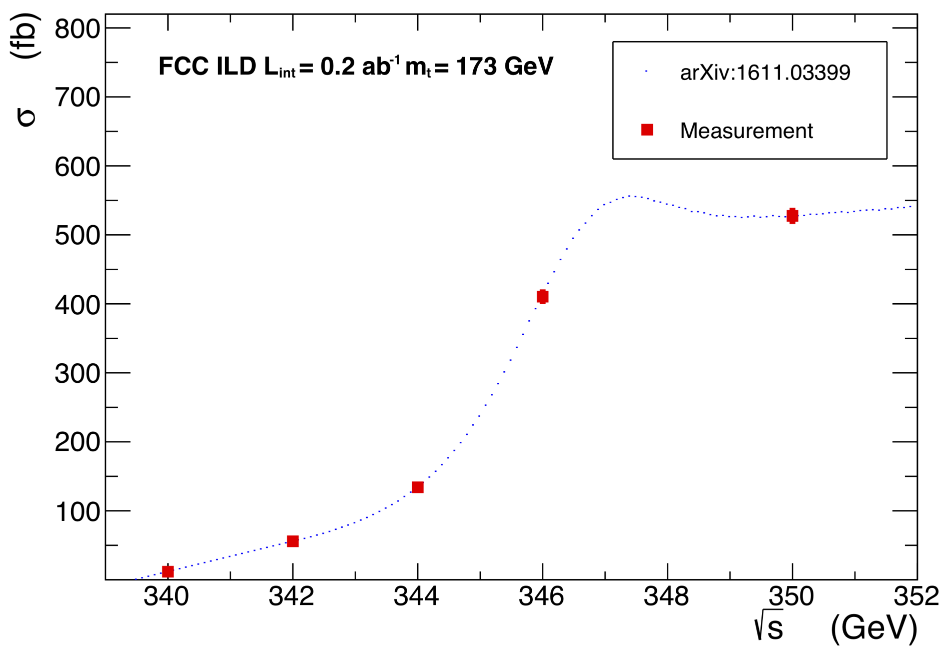


Figure 6.11: Measured cross-section compared to the predicted cross-section for the $t\bar{t}$ process at the FCC-ee [15].

Conclusions and Outlook

A study of the $t\bar{t}$ threshold is performed for the Future Circular Collider-ee.

Multiple statistical tests based on χ^2 terms were compared to obtain a purer sample of top quark events decaying into the semi-leptonic mode. The resolution of the FCC-ee ILD detector was analysed and used to determine which test variable is optimal. In the end, a selection is made for which 10 % of the signal events pass with a purity of 81 %. The statistical uncertainty on the measured cross-section was also determined and is in the order of $1.8 - 7.7$ %.

The samples after the selection process were then used to reconstruct the hadronic and leptonic W boson and top quark masses. The results obtained were consistent with the resolutions for the ILD detector with the statistical test chosen. The effect of setting a different top quark mass in the simulated data was looked into but produced no significant deviations. Changing the centre of mass energies did not affect the reconstructed mass distributions but modified the behaviour of the selection. A lesser performance of the selection and test was observed at lower energies, producing a lower purity sample. This is consistent with the cross-section of the $t\bar{t}$ producing less signal while the background processes remain constant.

The event pre-selection process should be studied further for some of the cuts, notably the lepton isolation and the minimum invariant mass of any two jets. The most important aspect to improve is the statistical test after the pre-selection. In this thesis a common χ^2 based test for top quark physics was used. The obtained results can vastly be improved by using more intricate methods, such as the ones discussed in section 5.1.

After improving the selection mechanism itself, the process should be studied for higher order samples at NLO or NNLO. The analysis so far only considered the effects of $\sqrt{s}$ and $\mathrm{m_t}$. This should be completed with the other variables that influence the threshold cross-section shape i.e. the top quark width Γ_t and the Yukawa coupling y_t. Concluding these analyses could help improve the selection for which energies $\sqrt{s}$ the FCC-ee are included in the top quark program.

This work is focused on the semi-leptonic decay channel of the $t\bar{t}$ process. The other decay modes should be studied in the same manner, which will require an entirely different analysis. Combining the results will provide good predictions for the threshold cross-sections and accuracy potential at the FCC-ee. Additionally, the various detector designs should be compared.

Summary

The Future Circular Collider is an ambitious project under consideration to succeed the Large Hadron Collider. This could be the first lepton collider to produce top quark pairs with centre of mass energies of $\sqrt{s} = 350$ GeV. In this thesis the semi-leptonic decay channel of the two top quark final state is studied. The detector used for the work is the International Linear Detector (ILD).

The main backgrounds which could interfere with the measurements were generated along with the expected signal. The most important sources of background were the WW and HZ processes. A succession of restrictions on the kinematic distribution of the events was set in place in order to obtain an estimate of how samples could be isolated which mainly contained the signal and reduced background. A few different test variables were developed and studied and compared in their performance to reduce the number of background events in the final data sample. The event selection resulted in a 10 % efficiency on the signal and a purity of 81 %, which is defined as the ratio of signal over total events after selection. The statistical uncertainty on the cross-section measurements was determined to be of the order of $1.8 - 7.7$ % for the collision energies $\sqrt{s} = 340 - 350$ GeV.

The events passing the selection were used to study the top quark and w boson mass distribution. The resolution prediction of the ILD detector, one of the proposed detector scenarios, was determined for the chosen selection. The influence of the simulated top quark mass and centre of mass energy was studied. Changing these variables did not influence the measurement but modified the performance of the selection. A better purity of the sample is observed at higher collision energy as more top quark pairs are produced. This is expected as the background processes remain roughly constant. The effect of $\sqrt{s}$ and the top quark mass on the efficiency is very small, only a small decrease observed for the highest collision energy. These results will help determine which parameters are important or not for the FCC top quark program. It will allow to make estimations of the top quark mass, width and coupling to the Higgs boson at the $\sqrt{s} = 2\ \mathrm{m_t}$ threshold, which is an important input to the assessment of the performance of future colliders.

Acknowledgements

First I would like to thank my mother, my brother and sister-in-law for the continuous support during the last months. Taking on this large project was not an easy task and I am very thankful they were there for me when I needed it most.

I would also like to thank my promotor Freya Blekman for believing in me and pushing me when necessary. I appreciate all the opportunities you have given me in the past few years. I am very happy I could pursue all of the internships, exchanges and physics schools in and mostly outside of Belgium, knowing I had support at home.

Finally, I would like to thank my friends Sander, Dieder, Florian, Medina and Robin. You guys are the best and I am very glad to have people like you in my life. Thank you for listening to my rants and helping me stand back up.

Bibliography

[1] M. Thomson, "Modern Particle Physics", Cambridge University Press (2013), 554 p.

[2] F. Mandl and G. Shaw, "Quantum Field theory", Wiley (2010), 471 p.

[3] *The LEP Electroweak Working Group,* http://lepewwg.web.cern.ch/LEPEWWG

[4] M. Tanabashi *et al.* [Particle Data Group], "Review of Particle Physics," Phys. Rev. D **98** (2018) no.3, 030001. doi:10.1103/PhysRevD.98.030001

[5] S. Chatrchyan *et al.* [CMS Collaboration], "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC," Phys. Lett. B **716** (2012) 30 doi:10.1016/j.physletb.2012.08.021 [arXiv:1207.7235 [hep-ex]].

[6] G. Aad *et al.* [ATLAS Collaboration], "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC," Phys. Lett. B **716** (2012) 1 doi:10.1016/j.physletb.2012.08.020 [arXiv:1207.7214 [hep-ex]].

[7] H. Georgi and S. L. Glashow, Phys. Rev. Lett. **32** (1974) 438.

[8] D. I. Kazakov, "Beyond the standard model: In search of supersymmetry" hep-ph/0012288.

[9] *The Tevatron Collider,* http://www-bdnew.fnal.gov/tevatron

[10] *The CDF experiment,* http://www-cdf.fnal.gov/

[11] *The D0 experiment,* http://www-d0.fnal.gov/

[12] Particle Data Group, "CKM Quark-Mixing Matrix", http://pdg.lbl.gov/2018/reviews/rpp2018-rev-ckm-matrix.pdf

[13] J. Haller, A. Hoecker, R. Kogler, K. Mönig, T. Peiffer and J. Stelzer, "Update of the global electroweak fit and constraints on two-Higgs-doublet models," Eur. Phys. J. C **78** (2018) no.8, 675 doi:10.1140/epjc/s10052-018-6131-3 [arXiv:1803.01853 [hep-ph]].

[14] K. Seidel, F. Simon, M. Tesar and S. Poss, "Top quark mass measurements at and above threshold at CLiC," Eur. Phys. J. C **73** (2013) no.8, 2530 doi:10.1140/epjc/s10052-013-2530-7 [arXiv:1303.3758 [hep-ex]].

[15] F. Simon, "Impact of Theory Uncertainties on the Precision of the Top Quark Mass in a Threshold Scan at Future e+e- Colliders," PoS ICHEP **2016** (2017) 872 doi:10.22323/1.282.0872 [arXiv:1611.03399 [hep-ex]].

[16] Particle Data Group, "Top Quark", http://pdg.lbl.gov/2018/reviews/rpp2018-rev-top-quark.pdf

[17] Future Circular Collider Study. Volume 1: Physics Opportunities. Conceptual Design Report, preprint edited by M. Mangano et al. CERN accelerator reports, CERN-ACC-2018-0056, Geneva, December 2018. Submitted to Eur. Phys. J. C

[18] Future Circular Collider Study. Volume 2: The Lepton Collider (FCC-ee) Conceptual Design Report, preprint edited by M. Benedikt et al. CERN accelerator reports, CERN-ACC-2018-0057, Geneva, December 2018. Submitted to Eur. Phys. J. ST.

[19] T. Taylor and D. Treille, "The Large Electron Positron Collider (LEP): Probing the Standard Model," Adv. Ser. Direct. High Energy Phys. **27** (2017) 217. doi:10.1142/9789814749145_0007

[20] Beyond the Large Hadron Collider: A First Look at Cryogenics for CERN Future Circular Colliders - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Footprint-of-the-100-km-FCC-accelerator-in-the-Geneva-basin_fig1_282555395 [accessed 16 May, 2019]

[21] P. Janot, "Future Circular Colliders, Phase 1: The FCC-ee", 103rd Plenary ECFA meeting, CERN, 16 November 2018

[22] T. Behnke *et al.*, "The International Linear Collider Technical Design Report - Volume 4: Detectors," arXiv:1306.6329 [physics.ins-det].

[23] N. Foppiani, "Precision measurement of the top quark mass at threshold with the FCC-ee", CERN Summer student project report, 2016

[24] *The FCCSW framework*, http://hep-fcc.github.io/FCCSW/

[25] A. M. Sirunyan *et al.* [CMS Collaboration], "Particle-flow reconstruction and global event description with the CMS detector," JINST **12** (2017) no.10, P10003 doi:10.1088/1748-0221/12/10/P10003 [arXiv:1706.04965 [physics.ins-det]].

[26] T. Sjöstrand, S. Mrenna and P. Skands, JHEP05 (2006) 026, Comput. Phys. Comm. **178** (2008) 852.

[27] *PArametrized PArticle Simulation,* https://github.com/HEP-FCC/papas

[28] M. Cacciari, G. P. Salam and G. Soyez, "FastJet user manual," Eur. Phys. J. C **72** (2012) 1896 [arXiv:1111.6097 [hep-ph]].

[29] M. Cacciari and G. P. Salam, "Dispelling the N^3 myth for the k_t jet-finder," Phys. Lett. B **641** (2006) 57 [hep-ph/0512210].