Lepton non-universality at the LHC

Author: Koen van den Brandt
Supervisor: Prof. Dr. Nicolo de Groot

August 22, 2017

Contents

1 Introduction

The fundamental physics processes at the smallest known distances are described by the Standard Model of particle physics. It encompasses the fermions, the fundamental building blocks of all known matter: 6 quarks (found inside protons and neutrons) and 6 leptons (electron (e), muon (μ) and tau (τ) and 3 corresponding neutrinos). Three different forces, each of which is carried by boson, act on the matter particles : the strong nuclear interaction carried by gluons, the weak nuclear interaction carried by W and Z bosons and the electromagnetic interaction carried by photons. The fourth force, gravity is not included in the Standard Model. The final piece of the Standard Model is the Higgs particle. This neutral particle has spin 0 and is needed to explain the mass of the other particles. Its own mass is not predicted by the theory. In July 2012 the particle has been observed by the ATLAS detector in The European Organisation for Nuclear Research(CERN) in Geneva, Switzerland[1]. The particles of the Standard Model are shown in Figure 1

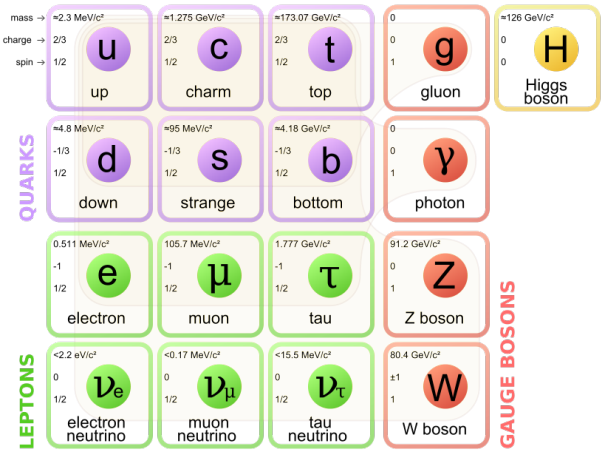


Figure 1: Overview of the particles in the Standard Model[2]

An essential element of the Standard Model is lepton universality. This means that the strength of force is the same for all three charged leptons, electron, muon and tau. Interestingly enough there have been hints from the LEP experiments[3] and more recently from the BaBar, Belle and LHCb experiments[4] that the W couples slightly stronger to the tau lepton than to the other. The problem with this measurement was the error, which was roughly 2%. This is not accurate enough to claim a discovery. However, it could be a sign that something is wrong with the Standard Model. Therefore new research needs to be done into this subject, and this thesis does exactly that.
The research in this thesis consist of two parts. One part which has taken place at CERN under supervision of Panagiotis Gkountoumis. At CERN I have helped with the upgrade of the of the muon detectors of the ATLAS experiment at CERN. The other part has taken place at the Radboud University in Nijmegen under supervision of Prof. Dr. Nicolo de Groot. Here I have analysed collision data from the ATLAS-detector at CERN. With this data the lepton universality will be tested.

2 The LHC and the ATLAS detector

This section has been written extensively in many reports, so my version of this section is a summary of a similar section in the PhD thesis of Hartger Weits[5].

2.1 LHC

The Large Hadron Collider[6] (LHC) is a particle collider which accelerates proton beams to very close to the speed of light, to a centre-of-mass energy of 14 TeV and a peak luminosity of $\mathcal{L} = 10^{34} cm^{-2} s^{-1}$, before colliding them. It has a circumference of 27 km, which makes it the largest single machine in the world. The LHC can also accelerate heavy ions. This make heavy ion collisions or heavy ion-proton collisions possible, but this happens at lower energies than the earlier mentioned proton-proton collisions. The LHC cannot accelerate the particles to this high energies by itself. It needs a few pre-accelerators to get the particles to a desired energy to be further accelerated in the LHC. A schematic overview of this process is shown in Figure 2. First, hydrogen gets stripped of its electrons and is inserted into the Linac2. Here the protons get accelerated to roughly 50MeV. These protons are further accelerated in successive cyclotrons called Proton Synchrotron Booster (50MeV to 1.4 GeV), Proton Synchrotron (1.4 GeV to 25 GeV) and the Super Proton Synchrotron (25 GeV to 450 GeV). After the protons leave the Super Proton Synchrotron, they have enough energy to be further accelerated in the LHC to the maximum energy of 7 TeV. The protons travel in the LHC in bunches where half are accelerated clockwise and the other half is accelerated anti-clockwise. In 2012, the proton beam consisted of roughly 1374 bunches, each containing $1,6 \times 10^{11}$ protons, with a spacing between the bunches of 50 nanoseconds. The tubes of the clockwise-bunches and the anti-clockwise-bunches are crossed at four different collision points. All collisions happen at these collision points, where detectors are placed. These detectors are: A Large Ion Collider Experiment (ALICE), A Toroidal LHC Apparatus (ATLAS) , the Compact Muon Solenoid (CMS), and the Large Hadron Collider beauty (LHCb) experiment. The LHC has 2 extra experiments that are not placed on the collision points. TOTal Elastic and diffractive cross sections Measurement (TOTEM), located along the beampipe on both sides of the CMS detector, and the Large Hadron Collider forward (LHCf) experiment, located along the beampipe on both sides of the ATLAS detector. Only the ATLAS detector will be used to use data, so this detector will be evaluated in detail in the next section.

2.2 ATLAS

The ATLAS detector, shown in Figure 3, is the largest of the experiments at the LHC, and also the largest collider-detector in the world. It has a length of 46 meters, a height and width of 25 meters and a weight of 7000 tonnes. The detector consists of different parts with each a different purpose in the detector. To differentiate where the detector parts are placed, barrel- and endcap region will be introduced. The barrel region is the region in a cylinder, or "barrel", around the beampipe. This barrel has a front- and a back-end. The regions at these ends are called the endcap region.

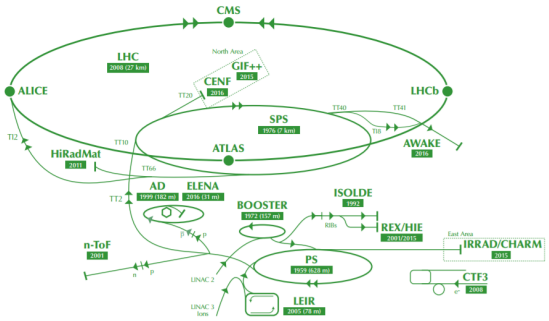


Figure 2: Schematic overview of the LHC and its pre-accelerators and detectors [7]

2.2.1 The Magnet System

The magnet system consists of 3 different magnets. A solenoid magnet encircling the inner detector, which provides a 2 Tesla magnetic field along the beampipe, a barrel toroid and two end-cap toroids which produce toroidal magnetic fields for the central and end-cap regions of 0.5T and 1T respectively. These magnetic fields change the trajectory of charged particles moving through the detector.

2.2.2 The Inner Detector

The inner detector tracks the path travelled by charged particles, travelling at relativistic speed, by measuring the trail of ionized particles that is left behind by this particle due to electromagnetic interaction with the atoms of the inner detector. The detector consists of three parts: the pixel detector, the semiconductor tracker (SCT) and the transition radiation tracker (TRT). The pixel detector, consisting of silicon pixels of $50 \times 400 \mu m$, is the closest to the collision point and is the most precise of the inner detector parts. If a charged particle passes the pixel detector, the ionization will create electron-hole pairs which can be detected. The pixels that are hit can help to determine the path of the particle. The SCT operates in the same way, but uses strips of $80 \mu m \times 12 cm$ instead of pixels. The TRT is a "straw tube" tracker, and it is the outermost of the tracking systems of the inner detector. It consists of gas filled drift tubes with a gold-plated tungsten wire in the centre. When an ionized particle passes the TRT, it ionises the gas in the drift tubes. The free electron that gets created drifts towards the wire, where it gets amplified until the signal can be read [8].

2.2.3 The Calorimeters

The Calorimeters are the next closest to the interaction point. These measure the energy of both charged and neutral particles. The energy of the particle is determined by absorbing the energy of the particle by a high-density metal. The inner part of the calorimeters consists of electromagnetic calorimeters, which are specially designed to measure electrons and photons. The part further out is designed to detect hadrons that pass through the electromagnetic calorimeters.

2.2.4 The Muon System

The muon system is placed at the outer part of the detector. This is the ideal location for the muon systems because the muon is the only charged particle that can pass the inner detector and the calorimeters without being absorbed. The charged particles measured in these muon systems must thus be muons. The muon chambers are positioned in three cylindrical layers in the barrel region, and three chambers in the endcap regions. It consists of four different measuring technologies. These are the Monitored Drift Tubes (MDT), Cathode Strip Chambers (CSC), Resistive Plate Chambers(RPC), and Thin Gap Chambers (TGC). All technologies use the same method of ionising gas and amplifying the electron signal that is created, like the TRT in Section 2.2.2. The difference between the technologies is the layout of the detecting parts. For the muon system, the RPC and the TGC can do really fast, but not so precise measurements. Therefore the RPC and TGC are used for the trigger of the event. On the other hand, the MDT and CSC are much slower, but are way more precise, which makes them ideal for tracking measurements.

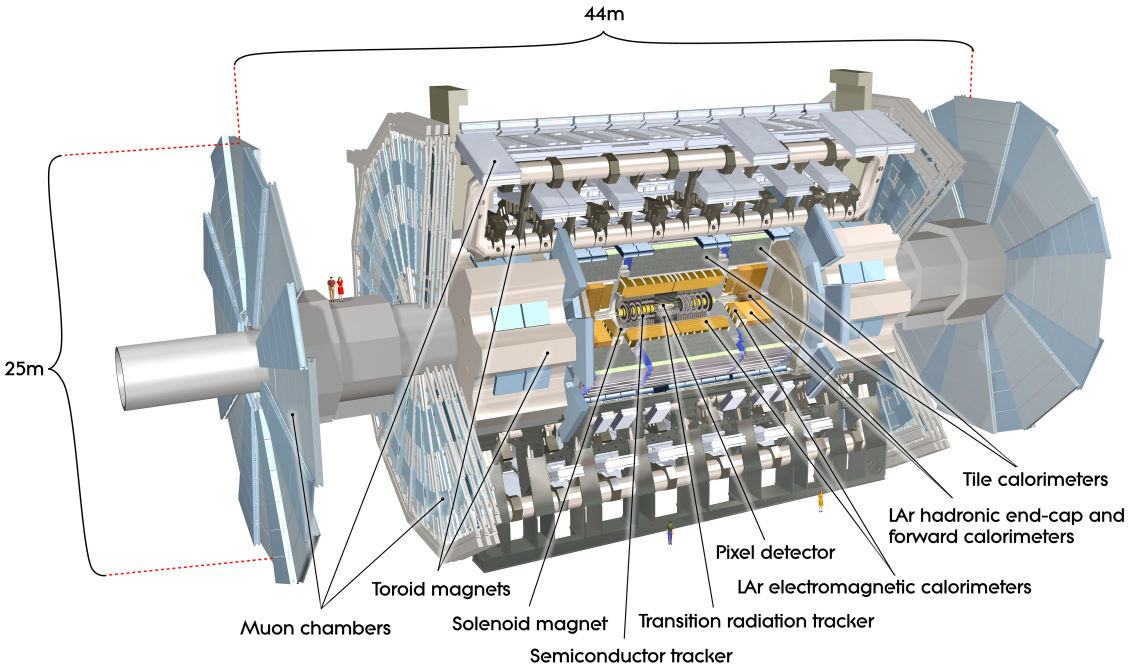


Figure 3: An image of the ATLAS detector. Two people are shown to give a sense of how large the ATLAS detector really is[9].

2.2.5 Trigger

The ATLAS trigger system consist of three subsequent filters: level-1, level-2 and event filters. The level-1 filter uses hardware to determine if a collision event is worth analysing. This process can take up to $2.5\mu s$ which reduces the event rate to roughly $75kHz$. The

event trigger is set up to let events pass with at least 1 high p_T electron, muon or photon or the event has a large transverse energy. The level-2 and the event filter access more detector information for a final data-rate up to $400Hz$ with an event size of roughly 1.6 Mb.

3 My Work at Atlas

For my internship, I had the possibility to not only do data-analysis, but also experience how it is to work at the place where the data is taken. Because of this, I worked for 9 weeks at CERN to help with the upgrade of the ATLAS-detector.

3.1 NSW and FELIX

I worked on the New Small Wheel (NSW) upgrade of the ATLAS "Small wheel" muon chambers. These "Small Wheel" muon chambers are located at the endcap of the inner detector, shown in Figure 4. The current small wheels will not be able to handle the ever increasing collision rate of the LHC machine. Therefore a faster detector with better granularity has been developed; the New Small Wheels. This upgrade is planned during the second long shutdown in 2018[10]

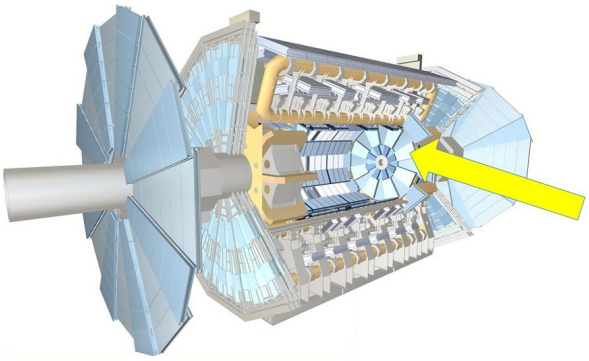


Figure 4: The position of the Small Wheel muon detectors[11].

With the NSW upgrade, ATLAS is implementing a new unified optical link technology for connecting detector and front end electronics. The front end electronics are the electronic systems located on the detectors, which forward the detector data to the readout systems. The new technology is called "Front End LInk eXchange"(FELIX)[12].

3.2 My Work

My part of the the project consisted of testing the FELIX software and setup, and make sure the setup worked as intended for future test uses. This was done under the supervision of Panagiotis Gkountoumis, a PhD student from the Nation Technical University of Athens.

Figure 5: A photo of the computer with the Felix card in it

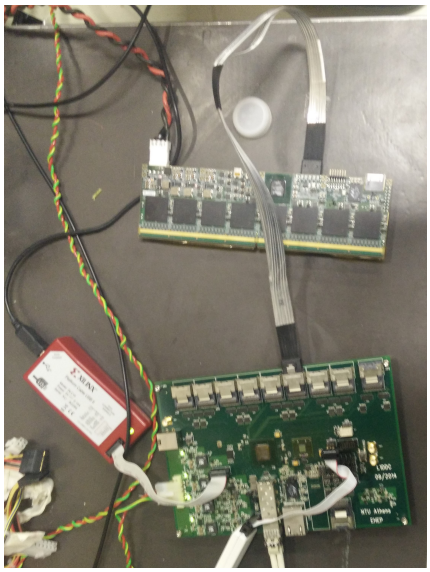


Figure 6: A photo of the setup used in one of the experiments. The bottom board is the L1DDC, the top board is the MMFE8, and the red box is the programming device for the GBTX(on the L1DDC). The connector at the bottom of the L1DDC is connected to the back of the computer shown in Figure 5

3.2.1 Setup

The setup used for the tests consisted of a computer with a FELIX-card, software that was provided by the FELIX-group at CERN, a set of commands that activates the software, a Level 1 Data Driver Card (L1DDC) with a radiation tolerant chip with bidirectional high speed optical link connector on it, called a GBTX, a connecting device to connect the GBTX to the computer, and a MicroMega Front-End board (MMFE8, 8 because the board contains 8 Venites MicroMegas (VMM's)). The FELIX-card has 4 ports, each containing a set number of E-links which can be individually programmed. The MMFE8 is the the electronic device that will be connected to the Wire-chamber detectors (also called Micromegas) of the NSW to collect and forward the data from this detector. A photo of the computer with the FELIX card can be seen in Figure 5, and a photo of the L1DDC, the MMFE8 and the programming device for the GBTX can be seen in Figure 6.

3.2.2 Software

The software that I used was a list of commands, prepared by the person that worked with the FELIX before me. It contained commands to do all the operations necessary to do the tests that I needed to do. The commands were:

- **flx-info**: This command gives information about the status of the FELIX card in the computer. Different sub-commands can give you different information, for example if the clock of the card has been locked, of which output channels of the card were used.
- **flx-throughput**: This command gives the combined data flow of all the ports of the Felix card.
- **flx-init**: This command initialises the driver of the Felix-card. This command can also set the frequency of the clock of the card.
- **fel**: This command sends out a specific amount of generated data using an either an internal or an external emulator. The internal emulator sends generated data through an internal loopback inside the Felix card, where the external emulator sends its generated data out through the ports of the Felix card.
- **fdaq**: This command uses the same emulators as **fel** and saves the generated data in a file. The time over which the data is generated can also be specified.
- **fcheck**: This command can check a data file, e.g. the data file created by **fdaq** on specific criteria.

Alongside these commands, two extra programs were needed to do the tests.

- **E-link configuration tool** : This program is used to enable or disable the specific receiving or emitting E-links of the ports of the Felix card. This tool also gives the possibility to set the encryption of the data for each E-link and choosing the kind of data that is generated by the emulators. The interface of the E-link configuration tool can be seen in Figure 7.
- **GBTX-programmer tool**: This program can change the values of "registers" of the GBTX-chip. These registers give define the function of the GBTX-chip. The GBTX-manual[13] can be used to find the function of specific values of each registers.

3.2.3 Testing

The testing of the software and the setup consisted of creating a setup where either the generated or received data can be checked. To check the generated data, loop-backs can be used. Different kind of loop-backs were used to test the data and the decoding of the system. The internal loop-back is the easiest one to check, since it only requires you use the internal emulator for your tests. The external loop-back can be done in different ways. One

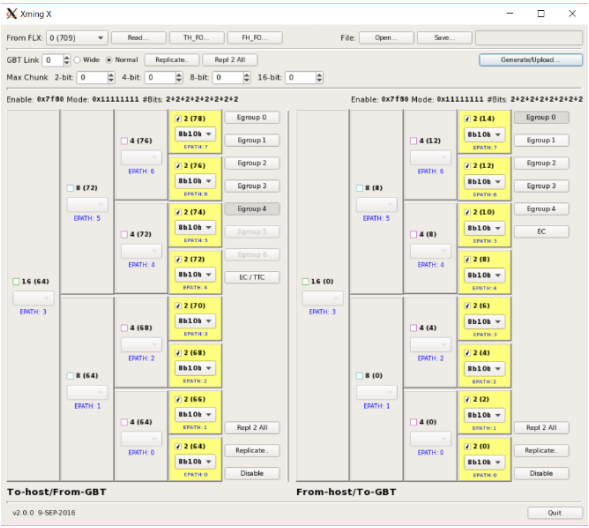


Figure 7: The interface of the E-link configuration tool. Each block represents a specific E-link of the specific ports (named GBT-link in the configuration tool).

can send data through one of the ports of the Felix-card, and receive it in another. This can be done by enabling only the transmitting side of one of the ports, and only enabling the receiving side on the other port via the E-link configuration tool.
An other possibility is programming the GBTX on the L1DDC to return the incoming data, which gives a PC→L1DDC→PC loop-back. This can be done by setting changing the values of specific registers, using the GBTX-programmer tool. The same tactic can be used to create a loop-back also including the MMFE8 in the cycle. This way you get a loop-back which goes like PC→ L1DDC→ MFE8→ L1DDC→ PC.
Using the above mentioned loop-backs it is possible to check how good your setup is working, since every loop-back should give the same data file. Also, if you know that the setup is working, you can use the E-link configuration tool to change different settings of the E-links, resulting in tests of the programming of these settings.

To check the received data, it is possible to program the GBTX on the MMFE8 to create a known pattern. If the setup and software are working as intended, the pattern should be clearly distinguishable in the saved data-file.

3.2.4 My contribution

As the testing went on, some remarkable results were found for the specific version of the FELIX-card used.
Firstly, a problem was found where the order of the physical ports of the FELIX-card did not correspond with the order of the ports in the software. In our case, the order from top to bottom was 3-2-0-1, instead of 3-2-1-0. This lead to a lot of confusion initially.
Secondly, it was found that the 0^{th} port of the FELIX card had to be enabled for the other ports of the FELIX card to work. More specifically, if an E-link is disabled in the 0^{th} port, that specific E-link will not work in any of the other ports. Before this problem was

diagnosed, it was impossible to test the external loop-backs because the cable to create a loop-back was initially not put in the 0^{th} port. This made all attempts to try a external loop-back result in failure.
It was also found that connections on the L1DDC correspond to specific E-links. This means if you disable the E-link corresponding to a connection, this connection will not receive or transmit data. The 8 connections on the L1DDC were connected to 8 specific E-links. This mapping made it possible to disable all E-links but these 8 specific E-links, removing most "broken" E-links. Broken E-links sometimes give off a random signal, crashing the software. If any the 8 specific E-links are broken, it is now a lot easier to find this E-link, since the possible broken E-links are reduced from 40 to 8.
Lastly it was found that the computer the external loop-backs all were working as intended and, apart from the broken E-links, the setup was working as intended. Also, the computer was able to receive the signal generated on the MMFE8. Unfortunately the data received did not correspond to the data that was supposed to be sent. In our setup, we were sending a counter, but from the received data files no counter could be distinguished. This was the problem that occupied me for the longest part of my stay. Together with Panagiotis Gkountoumis, we had checked every possible step of process, but we could not find the problem. When I left CERN, this problem was still to be solved. The solution for the problem came a few weeks later, where it was found that our software which generated the data was not encoding its data properly, thus leading to wrong data after decoding.

4 Theory

The theory of this thesis will be rather short. It won't go into the deep mathematics of the processes that are analysed in this thesis because partly this is too advanced for a bachelor thesis, and partly because it is not necessary to understand the research that has been done in this thesis.

4.1 Elementary particles

This thesis will consist of the analysis of elementary particles. These elementary particles are created by the interaction of the quarks which make up the protons. For this analysis, only the leptons, quarks, and the W- an Z-bosons analysed.

4.1.1 Leptons

The leptons are the green particles shown in Figure 1. These leptons are half-integer spin particles which are split in two groups. One of them is the charged leptons, which consist of the electron(e), muon (μ) and tau (τ) particles. The other group is the neutral group, which consists of the electron-, muon- and tau-neutrino (ν_e ν_μ ν_τ). Of these particles, only the electron is stable, but for our analysis purpose, the muon has a decay time which is long enough for it to travel through the whole detector, and so it can be measured. The tau particle has a decay time which is a lot shorter than that of the mu particle. Therefore, no tau particles can be measured by the detector, since they will have decayed before they reach the end of the inner detector. The tau particle will however decay into either a

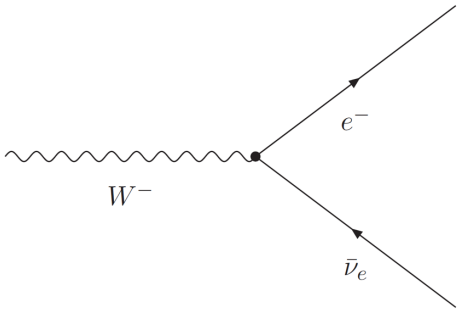


Figure 8: A Feynman diagram of a possible decay of the W^- particle. The electron could also have been a μ or τ

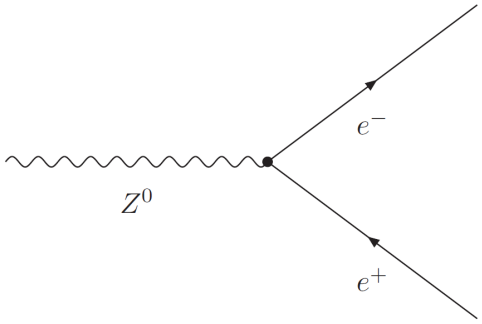


Figure 9: A Feynman diagram of a possible decay of the Z^0 particle. The electron could also have been a μ or τ

$\mu+\nu_\mu+\nu_\tau$, or $e+\nu_e+\nu_\tau$, making it possible to detect the tau particles through analysing the muon and electron data from the detector.

4.1.2 Quarks

Quarks are the elementary particles that stick together to form all the known hadronic particles we know. These quarks, shown in the purple in Figure 1, consist of the up-type quarks (up, charm, top) and the down-type quarks (down, strange, bottom). Each quark has a colour charge which is independent of the type of quark. The only restriction to the colour charge is that the combination of quarks that makes up a particle has to be colour neutral. Therefore, a single quark can not exist in free space since it has a colour charge.

4.1.3 W-boson

The W-boson and the Z-boson are the particle responsible for the weak interaction. The W-boson has a mass of roughly $81GeV/c^2$, and can either have a positive or negative charge and isospin, creating a W^+ and a W^- variant of the W-boson. The W-boson can decay in two different manners. First, and most likely, the W-boson can decay hadronically Secondly, and very important for this analysis, the W-boson can decay lepotonically. This means the W-boson can decay in one of the three leptons and its corresponding neutrino. If Lepton universality holds, the decay of the W-boson to either e, μ, or τ should be equally probable. A Feynman diagram of the leptonic W-decay is shown in Figure 8.

4.1.4 Z-boson

The Z-boson is the other force carrier of the weak interaction. The Z-boson has a mass of roughly 91 GeV/c^2, but is not charged like its brother. Thats why the Z-boson is shown in most literature as Z^0. The Z^0 decays into a fermion and it's antiparticle. The Feynman diagram of this decay can be seen in Figure 9.

4.2 Definitions

In this thesis, some variables will be used which might not be known to the reader. Here all variables used will be explained.

4.2.1 Transverse Momentum

The Transverse momentum (Pt) is the momentum measured in the transverse plane. This is the plane perpendicular to the direction of the beampipe. Because the 2 protons that are collided only have a momentum in the direction of the beampipe, the Pt of both protons, and thus the resulting Pt of the collision, is zero. Also, because the momentum of the proton (in the direction of the beampipe) is randomly distributed over the quarks and gluons in the proton, it is not known what part of the proton-momentum the quarks or gluons carry. Since the collisions happen because of these quarks- and gluon interactions, it is not known what the initial momentum of the collision is. That's why the momentum in the direction of the beampipe is not used as a measuring variable, but the Pt is.

4.2.2 Missing Transverse Momentum

The missing transverse momentum (Etmiss) can be calculated from assuming that the total Pt should be zero. If now the addition of all Pt measurements from a collision result in a non-zero total Pt, this lacking Pt can be ascribed to one or several neutrinos in the sample. The neutrinos do not get measured by the detector, so the Pt of these particles can be seen as "missing", which will result in a Etmiss vector.

4.2.3 Transverse Mass

The Transverse Mass is a Lorentz invariant quantity used to describe particles in particle physics. The transverse mass of a particle in the LHC is given by Equation 1, where θ is the angle between the two Pt-vectors. In our analysis, the $Pt_{lepton1}$ is the Pt of the lepton, and $Pt_{lepton2}$ is the Etmiss of the event.

$$W^2_{\text{tmass}} = 2 * Pt_{lepton1} * Pt_{lepton2}(1 - cos(\theta)) \tag{1}$$

4.2.4 d0

The d0 of the lepton is the impact parameter of the collision. The d0 is calculated as the point of closest approach of the lepton track to the interaction point of the detector. The lepton d0 can be used to distinguish particles that are created directly by proton-proton interactions, and particles that are formed by the decay of other particles. These last particles will have a larger d0 parameter because they will get created only after its "mother-particle" has decayed, which takes some time. In this time, this "mother-particle" will have moved away from the interaction point, thus increasing the d0 of the created lepton.

4.2.5 Jet

A jet is a phenomena in particle physics where a quark-anti-quark pair is created which move in different directions. Because there is a potential between the quarks, the moving apart of the quarks makes in energetically more favourable to create another quark-anti-quark pair from the vacuum to annihilate the initial quark-anti-quark pair. This creates a meson, and another quark-anti-quark pair with lower energy. For this new pair the same process happens, creating another meson. This process continues until the energy of the remaining quark-anti-quark pair energy is negligibly small. All these created mesons will be travelling in the same direction. This combination of particles will be called a jet.

4.2.6 Jet_MV1

The Jet_MV1 is defined as the weight from a combination of tagging algorithms based on a Multi-Variate technique. This quantity is created to improve the light-flavoured-jet rejection and to increase the range of b-jet tagging efficiency for which the algorithms can be applied[14].

4.3 Created variables

To improve our analysis, some new variables were created from a combination of variables already available

4.3.1 MuonAngle

The MuonAngle is defined in Equation 2, and can be seen as the difference in angle between the lepton vector and the Etmiss vector. The reasoning behind creating this variable was that if you have $W \to \mu$ decay, the neutrino created will most go in opposite direction of the muon, resulting in a large MuonAngle close to π. For $W \to \tau$ decay, three neutrinos will be created instead of one. One of these muons is created when the W-boson decays to a τ and a neutrino. This neutrino will go in the opposite direction of the τ. The other neutrinos are created when the τ decays to a muon (only accounting for the leptonic decay of the τ). These neutrino will go in the direction of the muon. The resulting direction of the Etmiss, calculated by adding the vectors of all the neutrinos, won't be in the opposite direction of the muon. This leads to an average smaller MuonAngle for the $W \to \tau$ decay compared to the $W \to \mu$ decay.

$$MuonAngle = |\phi_{Etmiss} - \phi_{lepton}|, MuonAngle \in [0, \pi] \quad (2)$$

$$AnglePhi = |\phi_{jet} - \phi_{Etmiss}|, AnglePhi \in [0, \pi] \quad (3)$$

4.3.2 AnglePhi

The AnglePhi is defined in Equation 3 for events with 1 jet. For events with more jets, the AnglePhi can be calculated for each jet individually. This variable has been created to check for either a pion, created in a jet, which gives a "fake" muon-signal, or a wrongly measured jet-energy which will result in a "fake" Etmiss. Both events result in a small AnglePhi.

5 Data and Simulated Data

All data used for the analysis comes from the ATLAS Open Data website[15]. The "real" data sets are comprised of real data recorded with the ATLAS detector in 2012. The original set of data was recorded in period D of 2012, and has an integrated luminosity of approximately $1fb^{-1}$ ($1fb^{-1} \sim 10^{12}$ proton-proton collisions[16]) , and a centre-of-mass energy of 8 TeV. The real data will be referred to as data for the rest of this thesis

The simulated data, commonly named Monte Carlo (MC), is required to do the collision analysis. The MC data is simulated according to our current knowledge of particle physics. The simulation is a process of four steps.
The events get generated, using programs that use theoretical calculations, phenomenological models and experimental inputs to calculate the hadronic final states of the proton-proton collisions.
The ATLAS detector needs to be simulated to account for interaction of the generated particles inside the detector.
The detector response is derived from the particle interaction and it is written in a format compatible with the real output of the detector In addition, because of the high rate of collisions in the LHC, digested signals from several simulated events can be piled-up to crate samples with a realistic experimental background. This step is called "Digitization".
Lastly the particle trajectories and energies are reconstructed from the the detector. These final reconstructed samples are given to the physicists.
The MC data is simulated using programs, but not all datasets are simulated with the same program. The programs used are Pythia, PowHeg and Sherpa. A list of which dataset is simulated by which program can be found on the ATLAS Open Data website under in Software Book GitBook repository, section "Dataset Details" [15].

6 Analysis

The analysis done in this thesis consists of a combination of python and C++ code, used to analyse the data and MC-samples. The data and MC-samples are taken from the ATLAS open data website[15]. The python analysis is done on the data and MC-samples. The samples are in the ROOT format[17], and therefore PyRoot is used to analyse the samples. The C++ codes will use analysed data to test the lepton-universality of the data.

6.1 Python Analysis

The first part of the analysis has been taken form the Atlas open data website. This analysis was titled as WAnalysis, and contained the standard analysis code, plus some specific standard code to make this analysis a specific analysis for W-decay.

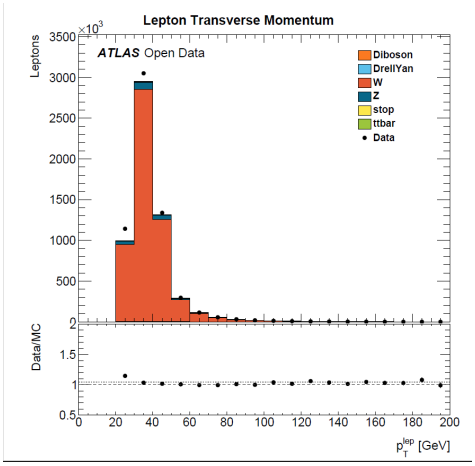
6.1.1 Standard W-Analysis

The standard analysis part contained code to create histograms for the different variables in the rootfiles, and corresponding code to fill these histograms with the data contained in the different rootfiles. These rootfiles consist of 44 MC files, and 2 files containing measured

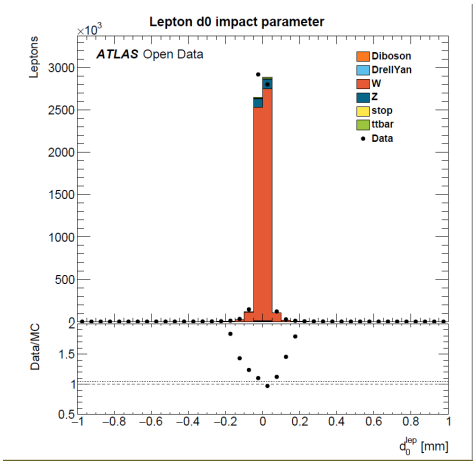
data. These data files were split in Muon data and Electron data. The MC is split into six different categories, showing what part of the simulated data comes from which source. These categories are: Diboson, DrellYan, W, Z, stop and ttbar. The code uses these files to show the Data and the MC-samples in the same histogram, where the MC-samples are scaled appropriately to accurately describe the data. There will be a difference between the data and the MC-samples in each bin. This difference is called background. This background can have different origins. The conditions of the data-taking can be different from the conditions which are programmed in the MC. Also, the MC can not be complete. This means certain processes are not modulated in the MC, or the processes that give the background are currently unknown to physicists. In the histograms, the data divided by the MC is also shown to show how accurate the MC-simulation is. If the Data/MC=1, this means the data is perfectly described by the simulations. The W-Analysis provided contained only a few lines to distinguish the W-decay from the rest of the data-samples. The analysis required 1 lepton in the event, since we want to analyse the leptonic W decay(W to 1 lepton and a lepton neutrino, as stated in section 4.1.3). This lepton had to have a Pt larger than 25 GeV and a Good muon track. The event was also required to have a Wtmass >30 GeV and a Etmiss >30 GeV. For the leptonic W decay, one can assume that the mass of the W (80GeV) is converted half into the Pt of the lepton, and the half into the Pt of the neutrino. If you now cut on 25 for the Pt of the lepton, and 30 on the Pt of the neutrino (Etmiss), you will reduce the non-leptonic W-decay samples from both data and MC by a lot, while barely affecting the total number of leptonic W-decays. Since the Etmiss and Pt cut of the lepton are already specified, the WTmass cut (from Eqn.1) is used to remove the events where the lepton and Etmiss have the same direction. The events removed by this cut could never come from leptonic W-decay, so the cut will only remove unwanted events. The standard W-analysis will result a histogram of the combined MC-samples and the data. A few of these histograms are shown in Figure 10.
The histograms shown in Figure 10 are only 4 out of the 20 possible histograms that are created. It is chosen to go for these specific variables, because these are used most in the analysis. The histograms contain the data (black dots) and the different MC-samples which are grouped to describe a certain process. The Data/MC ratio is also shown for every bin of the histogram.

6.1.2 Specific W-Analysis

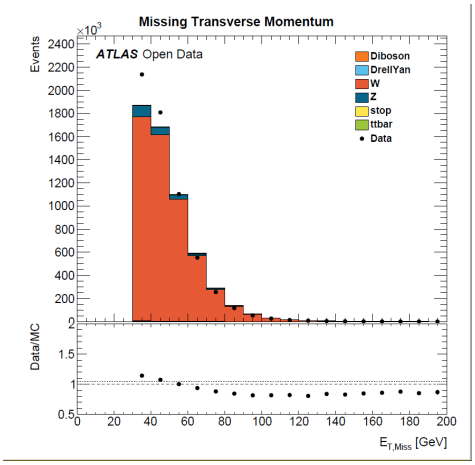
For the specific analysis of the lepton non-universality, the W$\rightarrow \mu$ and W$\rightarrow \tau \rightarrow \mu$ decay is analysed. The tricky part of this analysis comes from the fact that the $\tau \rightarrow \mu$ decay only accounts for 17% of the total W$\rightarrow \tau$ decays. This in combination with the 25 GeV Pt cut results in a maximum yield of 3-4% of the total W$\rightarrow \tau$ decay. Therefore, for a statistically significant result, filters need to be added to reduce the contribution of all other processes, while keeping the W$\rightarrow \tau$ contribution as high as possible.
First a filter is added that removes all events without muons, which removes all events with only electrons in it. Also, the legend of the histograms was changed to more accurately describe the W- and Z-parts of the legend in Figure 10. Lastly, the boundaries of the lepton d0 impact parameter have been changed, and the amount of bins have been increased to give a more accurate result of this distribution. These changes have been applied to all



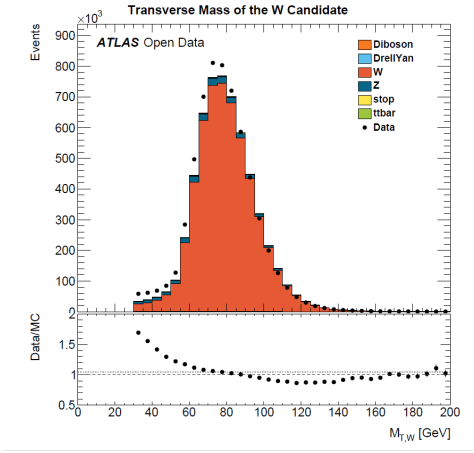
(a) Pt of the Lepton



(b) D0 of the lepton



(c) Etmiss of the system

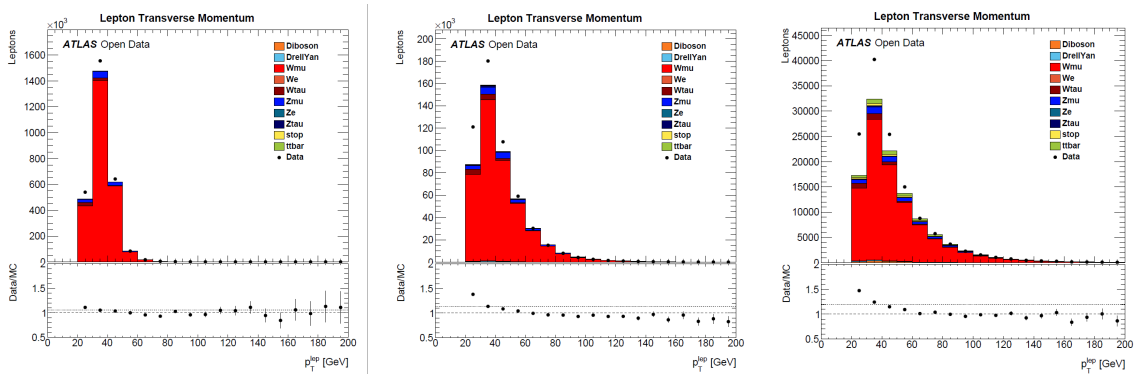


(d) WTmass of the system

Figure 10: 4 out of the 20 different variables that are plotted in histograms with the standard W-analysis of the ATLAS Open Data code. The definitions of these variables are given in Section 4.2

the upcoming histograms. The idea came up to look at different jet samples, just to see if different jet samples would give different results. The different jet samples are shown in Figure 11. It was concluded that the 2+ jet samples should be excluded from the analysis. This was because the 2 jet sample contained to much s-top and t-tbar, and the 3+ jet samples do not contain enough data for a decent statistical error. Also, because we are looking at the decay of a single W-particle which decays into a single lepton and a neutrino, we can add a filter which removes all events with more than 1 lepton. This filter will mainly remove Z-decay events, because they decay into two leptons. The filter for the leptons is made to remove all events which contain 2 leptons, both with a Pt larger than 10 GeV. The Z-boson mass is 91 GeV, so the leptons created from the decay will on average have an energy of 45 GeV, but could theoretically have all possible combinations which add up to 91 GeV. The upper bound is put at 10 GeV instead of the 25 GeV, as used in the standard analysis, to remove the most likely combinations of Pt's of the leptons from the Z-decay

from the sample. A comparison with and without the 2-lepton veto can be seen in Figure 12 for the the 1 jet sample. It can be seen that the 2-lepton veto removes predominantly Z-decays from the sample.

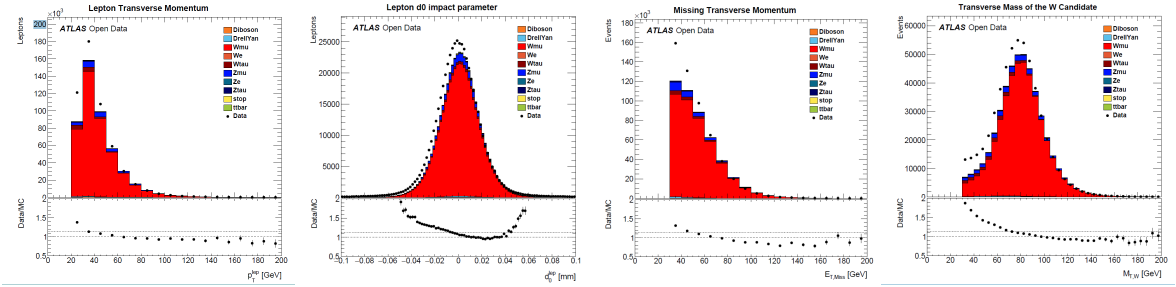


(a) WTmass of the Lepton of the 0 jet sample

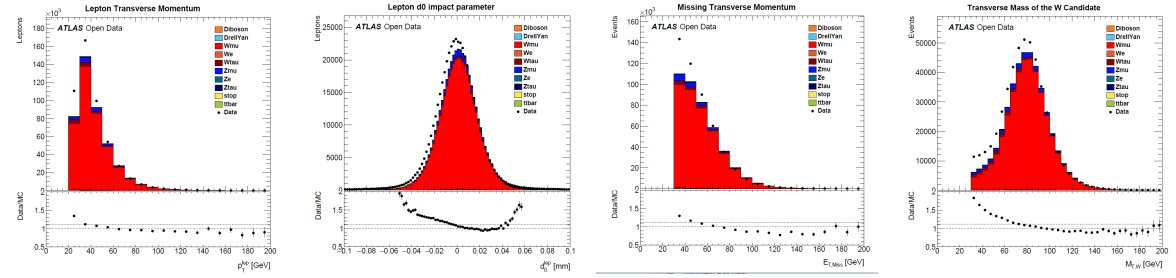
(b) WTmass of the lepton of the 1 jet sample

(c) WTmass of the system of the 2 jet sample

Figure 11: A comparison of the Pt of the lepton for different jet samples



(a) Variables of the system without 2-lepton veto



(b) Variables of the system with 2-lepton veto

Figure 12: A comparison of the data samples with and without 2-lepton veto for the 1-jet sample

6.2 Improving W $\rightarrow \tau$ ratio

Since there are a lot less W$\rightarrow \tau$ decays than W$\rightarrow \mu$ decays, it is preferred to increase the W $\rightarrow \tau$ ratio to reduce the error in the final calculation. This is done by adding additional

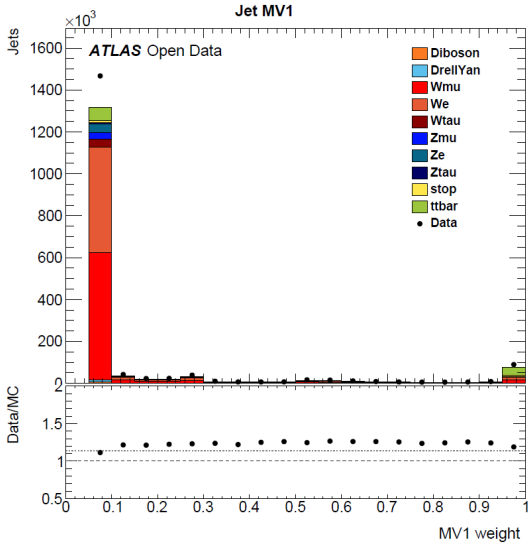


Figure 13: The Jet_MV1 of the 1 jet sample of the initial unaltered W-Analysis

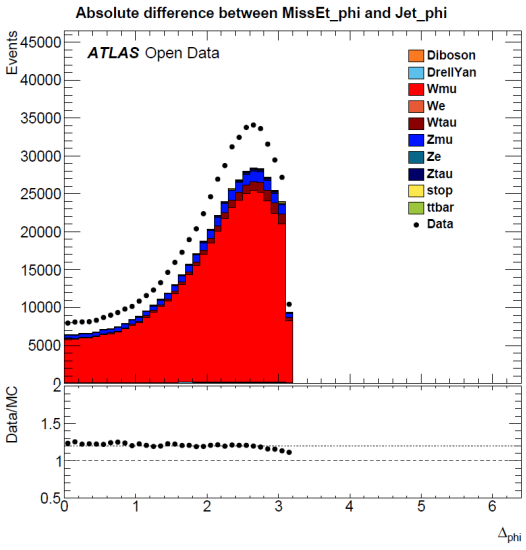


Figure 14: The AnglePhi of the 1 jet sample of the W-Analysis, including Jet_MV1 cut, 2 lepton veto, and electron veto

cuts on certain variables. The cut-value is determined by looking at the histograms for all the variables, and looking for regions where the W $\rightarrow \tau$ decays are negligible, but other kind of decays or background are still present in this region. The version of the W-analysis used will be stated in the caption of the figure.

6.2.1 Jet_MV1

When looking closely at all the different histograms, the Jet_MV1 gave us an option to remove some b-quark data, since the Jet_MV1 is a b-tagging algorithm. In this histogram, seen in Figure 13, everything above Jet_MV1 $= 0.5$ contains little to none of W$\rightarrow \tau$ decay, but does contain unwanted b-jet "junk". Therefore it is chosen to remove all data above the Jet_MV1 $= 0.5$ threshold.

6.2.2 AnglePhi

When looking at the AnglePhi in Figure 14, it can be seen that for AnglePhi ≤ 0.7, the bins do not contain any Wtau data. They contain only Wmu, Zmu and background. Therefore the events with AnglePhi ≤ 0.7 can safely be removed from the total database. Removing this part increases the relative $W \rightarrow \tau$ ratio of the sample.

WTmass >	Etmiss >	$\frac{\#W\rightarrow \tau \text{ events}}{\#W\rightarrow \mu \text{ events}}$
0	0	0.055417
10	10	0.049335
10	20	0.047755
10	30	0.04618
10	40	0.045222
20	10	0.043888
20	20	0.04253
20	30	0.041073
30	10	0.03841
30	20	0.03729
30	30	0.038227
40	10	0.03254

Table 1: The number of W$\rightarrow \tau$ events divided by the number of W$\rightarrow \mu$ events, for different cuts on the WTmass and the Etmiss

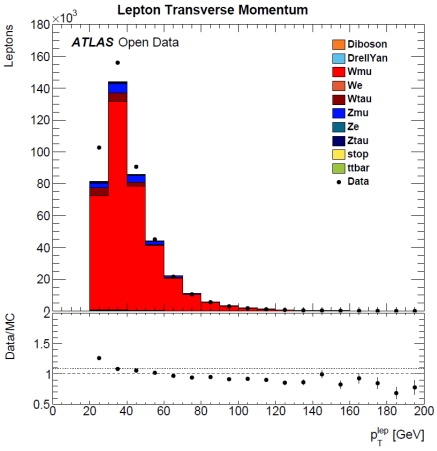
Pt >	$\frac{\#W\rightarrow \tau \text{ events}}{\#W\rightarrow \mu \text{ events}}$
25	0.053124
50	0.026221
100	0.019452

Table 2: The number of W$\rightarrow \tau$ events divided by the number of W$\rightarrow \mu$ events, for different cuts of the lepton Pt

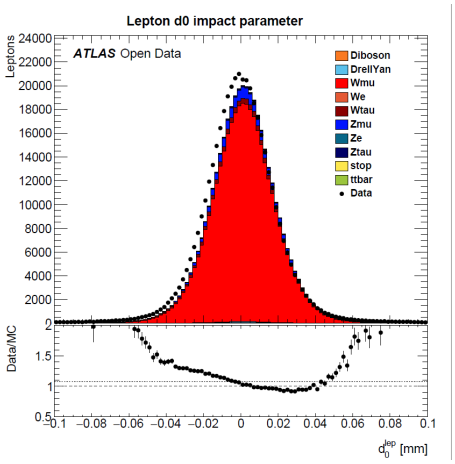
6.2.3 Varying Cut-values

To further improve the W $\rightarrow \tau$ ratio, the values of the cuts made in the analysis have been analysed. This was done by changing the cut-values of the WTmass and the Etmiss, since the WTmass depends on the Etmiss, as well as varying the Pt-cut of the lepton. The results of this varying can be seen in Table 1 and Table 2 respectively. The $\frac{\#W\rightarrow \tau \text{ events}}{\#W\rightarrow \mu \text{ events}}$ value used for the "0-cut" of the Etmiss/WTmass is the value in the analysis before the cut is implemented. The value of the Pt-cut starts at 25 because the ATLAS open data has been prepared in such a way that all the events without at least one lepton with Pt>25 GeV have been removed.
From these tables we find that the lower we put the cut-value, the better the $\frac{\#W\rightarrow \tau \text{ events}}{\#W\rightarrow \mu \text{ events}}$ ratio. Unfortunately, the cuts allow us to remove a lot of the background and the unwanted events. Therefore the ideal configuration of no WTmass- and Etmiss-cuts at all is not possible, since this will result in way to much background and unwanted events. The same argument goes for the WTmass/Etmiss>10 cut. The Pt has been set as low as possible, since the muon that we want comes from a $\tau \rightarrow \mu\nu_\mu\nu_\tau$ decay, which has an energy of roughly 50 GeV. This means the muon on average has one third of the energy of this τ-particle, which is almost 17 GeV. As can be seen from Table 2, the τ-ratio goes down drastically if you increase the Pt-cut value. Therefore, the ideal combination of cuts consists of Pt>25, WTmass>10 and Etmiss>30.

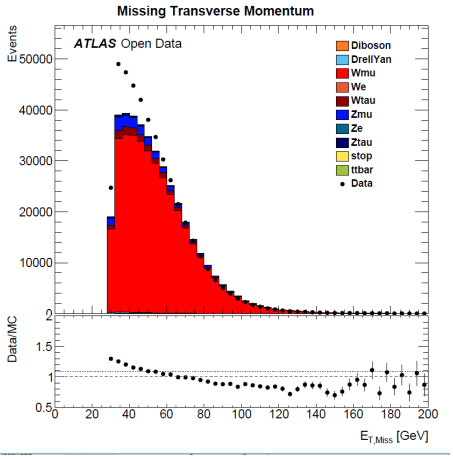
Combining the Jet_MV1 and AnglePhi cuts with the specific W-analysis and changing the WTmass- and Etmiss cuts as mentioned above, leads to the optimal conditions for analysing the W $\rightarrow \tau$ over W $\rightarrow \mu$ ratio. The histogram is shown in Figure 15.



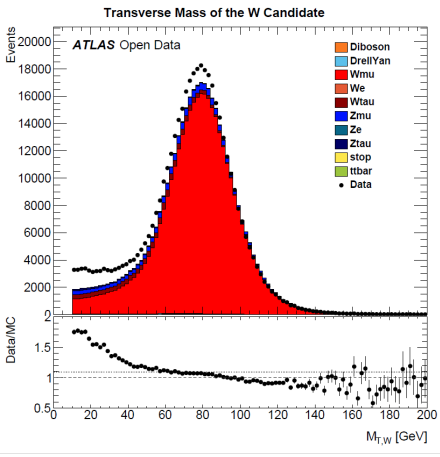
(a) Pt of the Lepton



(b) D0 of the lepton



(c) Etmiss of the system



(d) WTmass of the system

Figure 15: The histograms of the variables after running the final version of the W-analysis on the 1 jet sample, using all the methods to increase the Wtau over Wmu ratio of the sample

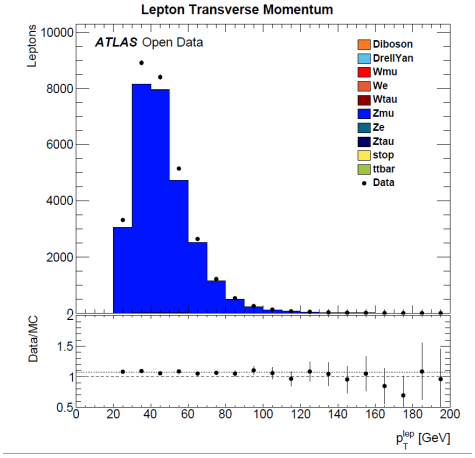
6.3 Removing the No-W-event background

To optimise the the W $\rightarrow \tau$ over W $\rightarrow \mu$ ratio, the events that are not W $\rightarrow \tau$ or W $\rightarrow \mu$ should be removed from both the data and the MC. For this, each different case of event has to be analysed individually.

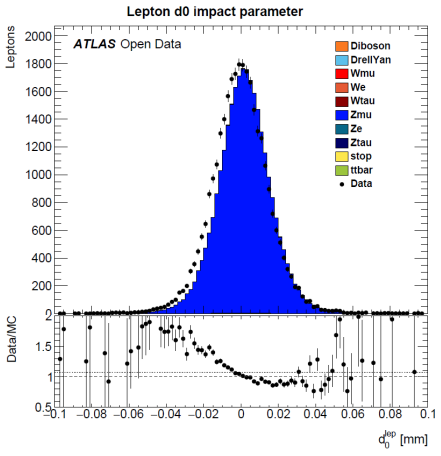
6.3.1 $Z \rightarrow \mu\mu$ decay

A sample has been created, which mimics the $Z \rightarrow \mu\mu$ events which pass the W-analysis, shown as Zmu in Figure 15. This has been done by selecting the events which contain two muons, and then randomly removing one of the muons by adding the removed muons energy and angle to the Etmiss. This leaves us with events with one muon and a new value for Etmiss. On this combination we can again apply our W-analysis criteria, which leaves us with the histogram shown in Figure 16. If we now compare Figure 16 with the Zmu

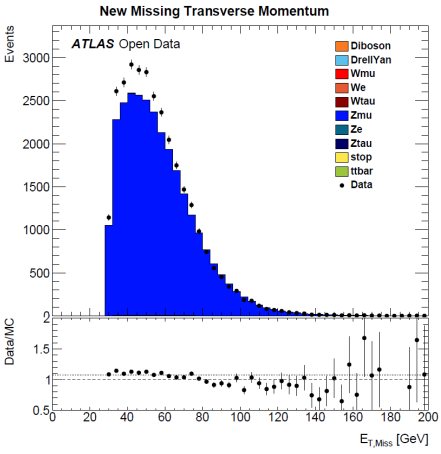
part of Figure 15, we can see an almost identical distribution of the Z-distributions. So, we have created a histogram that is a good representation of the Z$\rightarrow \mu\mu$ distribution in our sample. This sample allows us to validate the MC data. The criteria on the Z-decay leave very little room for background, so in theory the MC should fit the data very closely in this sample. This good correspondence between Data and MC in this sample can clearly seen from Figure 16. From the Figure 16b can be seen that the Data and MC distributions do not align perfectly in the d0-sample. Because we know this sample should be aligned very well, we conclude that the d0 is not modelled correctly. Therefore, a smearing factor can be calculated which aligns the Data with the MC. This smearing factor will not be shown in the data, but will be used to determine the error on the d0-resolution in Section 6.4.3.



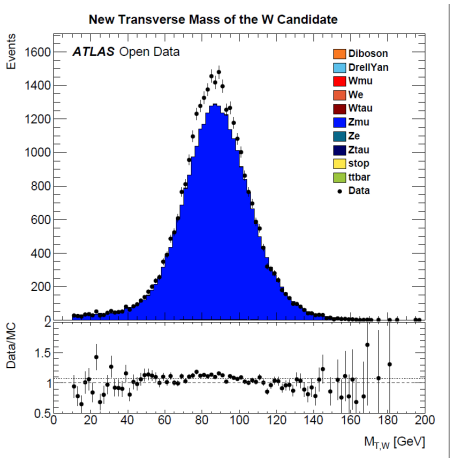
(a) Pt of the Lepton



(b) D0 of the lepton



(c) Etmiss of the system



(d) WTmass of the system

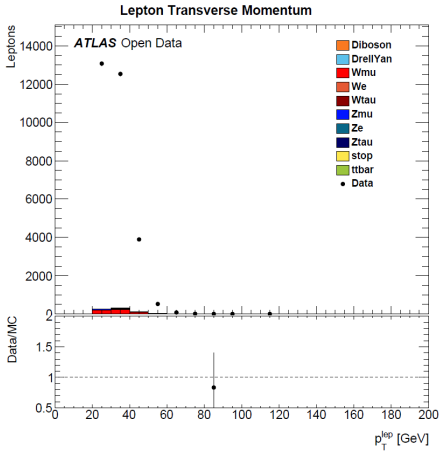
Figure 16: The histograms of the variables after running the Z pseudo W analysis on the 1 jet sample

6.3.2 $Z \rightarrow \tau\tau$ decay

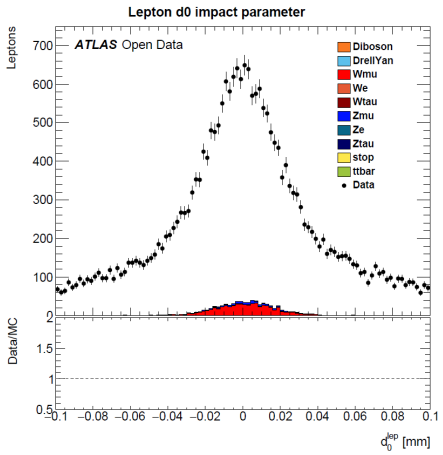
An attempt has been made to also recreate a sample for Z to tau data in Figure 15, by asking for a sample with at least 1 electron, and 1 or more muons. First, a Z peak had been recreated, using the collinear approximation[18], to estimate the Pt of the neutrinos, creating a new Etmiss. Then, the Pts of the muon and the electron were added for x, y and z direction. From these variables, the new total Pt, Etmiss and WTmass were calculated. Unfortunately, the data files were split in muon data and electron data. Therefore, asking for both a muon and an electron made it so that very little data came through our analysis, while its content did hardly contain any $Z \rightarrow \tau\tau$ data. After this, it was decided that the contribution of Z-tau was small enough compared to Z to mu and the background, that for this analysis it could be neglected.

6.3.3 Remaining Background

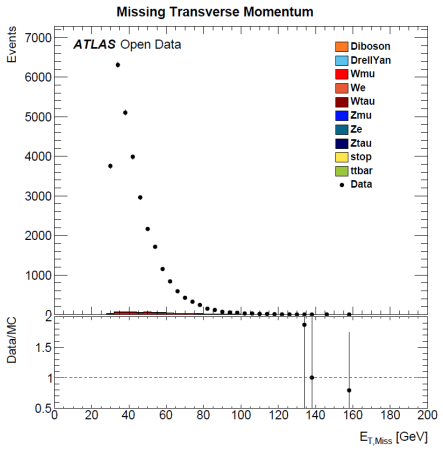
The background is the difference between simulated data and the measured data. If we assume that the "background-distribution shape" is the same for all histograms of the same variable, so independent of filter, it is possible to create a histogram with almost solely background. This can be done by taking the analysis code of the analysis that made Figure 15, and reverse the isolation requirements. This leads to the figure shown in Figure 17. It is now possible to remove the last few simulation events, using a C++ code, leaving only the background. This background can now be scaled up to the size of Figure 15. This scaling can be done by calculating a scale-factor, by comparing the background samples from Figure 17 and Figure 15. This leaves us with a scale-factor of 6.953. The scale factor was calculated using a C++ code made by Nicolo de Groot.



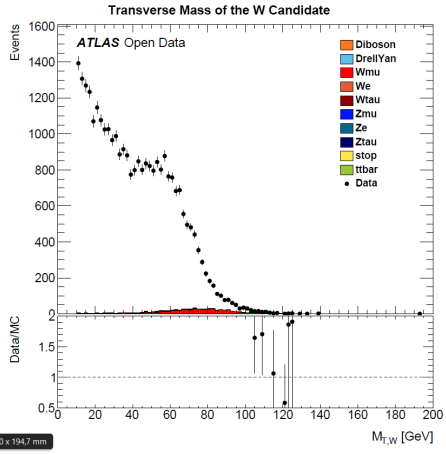
(a) Pt of the Lepton



(b) D0 of the lepton



(c) Etmiss of the system



(d) WTmass of the system

Figure 17: The histograms of the variables after running the background analysis on the 1 jet sample

6.4 C++ Analysis

The whole C++ code has been written by Nicolo de Groot. This section is an overview of how the code works and what has been done with this code.

6.4.1 Method

To find the $\frac{W \to \tau}{W \to \mu}$ ratio, a binned χ^2 fit was applied to one or more histograms of variables that show different distributions for muons and taus. There are several candidate distributions; the lepton pt, the WTmass, the MuonAngle and the d0 value. These variables are histogrammed for data, W$\to \mu$, W$\to \tau$ and the backgrounds, and a template fit is performed where the χ^2 is the sum over all histogram bins of the squared difference between data and Monte Carlo divided by its error. The total W cross-section, the $\frac{W \to \tau}{W \to \mu}$ ratio and the normalisation of the QCD fakes are kept as free parameters. The other normalisations are taken from MC and varied within their error. Positive and negatively charged muons

are fit separately to benefit from the fact that the W production shows an asymmetry and the background does not. We examine all candidate distributions on there statistical power and on their goodness of fit.

6.4.2 Jet and Variable selection

For both Jet-samples different variables are tried, to see which of them gives the best fit with a decent χ^2. The used variables were WTmass, Pt, MuonAngle(only for the 1 jet sample), d0 and $|d0|$. These were the variables considered the best for distinguishing tau-decays.
It turns our that for the 0 jet sample it is not possible to find a suitable variable. None of the variables has significant discrimination power. This is probably due to the very small amount of taus which survive the pre-selection cuts for 0 jets. The fits typically have a bad χ^2 , which indicate that the modelling in the MC is not correct.
The situation is better in the case of the 1 jet sample where we start off with a larger fractions of taus. Most distributions have only weak statistical power to constrain the ratio, but the $|d0|$ is actually reasonably effective. It also provides a good fit with a χ^2 /dof of 5.3/7. Therefore the fit is only done on the $|d0|$ values of the events.

6.4.3 Uncertainties

There is some sensitivity to the binning of $|d0|$. We choose our baseline binning such that we have bins with clear difference between tau and mu signal and exclude regions which are dominated by fakes. We evaluate a binning uncertainty by repeating the fit with a larger number of bins and with a different range.

- For the final analysis file, a set number of bins are used for the $|d0|$ value, on which the analysis was done. The average difference of different binning has been taken to be the binning systematic error.

- The impact parameter is not perfectly modelled in MC. We evaluate this effect by comparing d0 from Z0 decays, which have very little background, between data and MC. It turns out that the d0 distribution can be made to match almost perfectly by shifting it with 18 micron and smearing it with 40 micron. We take the difference between the smeared and un-smeared results as the uncertainty due to the d0 resolution.

- The largest systematic uncertainty comes from the fakes modelling. The spectrum for the fakes is extracted from inverting the isolation cut and the normalisation is left floating in the fit. The ratio of + over - is not exactly symmetric, but smaller that for W production. We vary the relative +/- contributions by 15% of their nominal value to estimate the uncertainty.

- The Jet and Muon energy scale errors are taken as the difference in the $\frac{\mathrm{W}\to\tau}{\mathrm{W}\to\mu}$ ratio, by varying the Jet-momentum cuts by 2.5% and the muon momentum cuts by 0.5%.

- The Z0 cross-section is known to 2%. We vary it with this number in the fit to find the associated systematic uncertainty. Likewise the ratio of W+ over W- is known

to 0.4% and varied accordingly in the fit.

Doing the error-measurements as mentioned will lead to the systematic uncertainties shown in Table 3.

source	uncertainty
Binning systematic	0.015
Fake modelling	0.097
resolution smearing	0.085
Jet energy scale	0.011
Muon energy scale	0.008
Z cross-section	0.002
W+/W- ratio	0.004

Table 3: The systematic uncertainties of the analysis

The resulting systematic error on the $\frac{W\to\tau}{W\to\mu}$ ratio is determined by quadratic adding of all the systematic errors of Table 3. This results in a systematic error of ± 0.131.

6.4.4 Results

Combining all the systematic errors from Table 3 and the statistic error, calculated by the fit code, the resulting ratio is:

$$\frac{W\to\tau}{W\to\mu} = 1.027 \pm 0.056(\text{stat}) \pm 0.131(\text{syst})$$

7 Conclusion & Outlook

In the end, the $\frac{W\to\tau}{W\to\mu}$ ratio that was calculated has a relatively large error. Therefore it is not possible to confirm or to dismiss the notion of lepton-non-universality from this thesis. The main problem of this analysis was the fact that the Monte Carlo simulated data was not optimal. The simulated data was incomplete, and the part that was simulated seemed to have some errors in them, as could be seen in the d0 modelling in Figure 16b. The MC thus gave us a lot of unwanted background, and an error in the resolution. These two facts account for most of the systematic error of this analysis. Also, the the data that we used was prepared by ATLAS. This meant that some data cuts were made in the preparation process of making this data ready for public use. Some of these cuts could be bad for this research, especially the cut which requires the event to have at least one lepton with a Pt $>$25 GeV. Since we expect to have an average momentum of roughly 17 GeV, this cut drastically reduces the number of τ-decays in the analysis. Further, the ATLAS open data website already states that this software and these data sets are for pure educational purpose. They said precision has been traded for simplicity of use. This means more precise software, and less prepared data sets are at ATLAS, which could both improve the measurement. Lastly, some more filters might have been added to reduce some of the background that was not modelled. Unfortunately due to time constrain, this could not be tested enough and was therefore not done.

The fact that this research did not reach the precision that was aimed for is unfortunate, but there are certainly ways to improve this measurement. As mentioned above, the MC could be improved by adding the events that the ATLAS did not simulate in their open data set and use the newest versions of the simulation algorithms, which both will increase the precision of the simulation. Likewise, the software and data sets could be improved by not using the "user-friendly"software which does not have the optimal precision. Also, newer data sets could be used. The ATLAS detector has had an upgrade after 2012, which should improve the D0 measurements. This should at least increase the resolution and reduce the smearing, removing a big part of our systematic error. Also, more data could be used. Currently, ATLAS has collected over $20fb^{-1}$, which should result in a statistical error of less that 1%. Further, more specific filters could be added to remove specific decays that are supposedly in our background. Lastly, there has been talk about a low-intensity run at CERN, which would create data with a lot less background than normal data. With less intensity comes less data, but since the requested data-taking will roughly equal the amount of data used for this thesis, and the background is the biggest problem at the moment, it will certainly increase the systematic error of the measurement.

8 Bibliography

[1] Georges 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.*, B716:1–29, 2012.

[2] Wikipedia. The standard model wikipedia page, 2016.

[3] S. Schael et al. Electroweak Measurements in Electron-Positron Collisions at W-Boson-Pair Energies at LEP. *Phys. Rept.*, 532:119–244, 2013.

[4] The Heavy Flavor Averaging Group. Hflav: Semileptonic b decay parameters, 2014.

[5] Hartger Weits. *Looking For Lepton Flavour Violation With The ATLAS Deteector, A Search For $Z \rightarrow \tau l$ Decays*. PhD thesis, Radboud University Nijmegen, 2016.

[6] Lyndon Evans and Philip Bryant. Lhc machine. *Journal of Instrumentation*, 3(08):S08001, 2008.

[7] CERN. LHC Guide. LHC Guide, Mar 2017.

[8] A Vogel. ATLAS Transition Radiation Tracker (TRT): Straw Tube Gaseous Detectors at High Rates. Technical Report ATL-INDET-PROC-2013-005, CERN, Geneva, Apr 2013.

[9] G.Aad et al. The atlas experiment at the cern large hadron collider. *Journal of Instrumentation*, 3(08):S08003, 2008.

[10] T Kawamoto, S Vlachos, L Pontecorvo, J Dubbert, G Mikenberg, P Iengo, C Dallapiccola, C Amelung, L Levinson, R Richter, and D Lellouch. New Small Wheel Technical Design Report. Technical Report CERN-LHCC-2013-006. ATLAS-TDR-020, CERN, Jun 2013. ATLAS New Small Wheel Technical Design Report.

[11] T. Kawamoto. Atlas muon small wheels for atlas phase-1 upgrade. CERN, 2012.

[12] J Anderson, A Borga, H Boterenbrood, H Chen, K Chen, G Drake, D Francis, B Gorini, F Lanni, G Lehmann Miotto, L Levinson, J Narevicius, C Plessl, A Roich, S Ryu, F Schreuder, J Schumacher, W Vandelli, J Vermeulen, and J Zhang. Felix: a high-throughput network approach for interfacing to front end electronics for atlas upgrades. *Journal of Physics: Conference Series*, 664(8):082050, 2015.

[13] K. Wyllie P. Moreira, J. Christiansen. Gbtx manual v0.15. CERN, 2016.

[14] ATLAS Collaboration. Performance of b -jet identification in the atlas experiment. *Journal of Instrumentation*, 11(04):P04008, 2016.

[15] ATLAS. The atlas open data website, 2016.

[16] CERN. Cern writing guidelines: inverse femtobarn, 2017.

[17] CERN. Root format website. CERN, 2017.

[18] A. Elagin, P. Murat, A. Pranko, and A. Safonov. A New Mass Reconstruction Technique for Resonances Decaying to di-tau. *Nucl. Instrum. Meth.*, A654:481–489, 2011.