– DRAFT VERSION –
LHC Data Challenge
Establishing a Bound on the Higgs Boson Mass

Adam Garay Jalloh, Dhruva Gowda Storz, Mark Fingerhuth, Samuel Natarajan, Henry Rice, and Annabel Wolf

Maastricht Science Programme, Maastricht University,
Maastricht, The Netherlands
`{a.garayjalloh,d.gowdastorz,`
`m.fingerhuth,s.natarajan,h.rice,`
`annabel.wolf}@student.maastrichtuniversity.nl`

Abstract. The abstract should summarize the contents of the paper and should contain at least 70 and at most 150 words. It should be written using the *abstract* environment.

Keywords: We would like to encourage you to list your keywords within the abstract section

1 Introduction

The Large Hadron Collider, located in CERN, Geneva, is the largest and most powerful particle accelerator in the world. It consists of a ring, about 27 kilometres in circumference, with a number of accelerating structures to increase the energy of the particles along the way. Within the loop there are two channels surrounded by superconducting magnets which allow the charged particles to be accelerated in opposite directions. The LHC is used by an international team of scientists to investigate the nature of the universe. The main objectives are to understand the shortcomings of the standard model; to explain the origin of mass using the Higgs field, to discover the reason for the missing anti-matter of the universe, and to understand the nature of dark matter and dark energy. Furthermore, supersymmetry is investigated, because this might unify the four fundamental forces by including gravity in the Standard Model.

This report discusses data collected from the ATLAS detector in 2012, which was released publicly in June 2016. It will give an overview of the most important background processes and explain how a signal is distinguished from the background. Furthermore, the statistical analysis of potential new physics is discussed.

1.1 The LHC and ATLAS

ATLAS, A Toroidal LHC Apparatus, is one of the four main particle detectors located along the LHC, where proton collision experiments are conducted. Two high-energy particle beams, each with 2808 bunches of 10_{11} protons, travel through the accelerator close to the speed of light in an ultrahigh vacuum (10_{13} atm) and at a temperature of -271.3C in opposite direction. The two beams are aligned to collide inside the ATLAS detector, which uses a large array of detectors organised in cylindrical shells to measure the properties of the particles passing through them.

Each layer of the detector reconstructs different particles. The first layer consists of tracking devices, which show the tracks of electrically charged particles through the trails they leave by ionizing matter. The particles momentum can be determined by measuring the curvature of the trajectory. Tracking devices either consist of a gaseous chamber or a semiconductor detector. In gaseous chambers, electrons or ions are collected on electrodes under a strong electric field, whereas semiconductor detectors are based on the fact that electron-hole pairs are created when particles pass through a reverse-biased semiconductor. The next layer of the detector is an electromagnetic calorimeter. This absorbs all electrons and photons, which have electromagnetic properties. Hadrons begin to lose energy in this stage, but are not stopped until they reach the hadronic calorimeter. The final layer of the detector catches muons. See Fig. 1 for a schematic overview. (Aad et al., 2012)

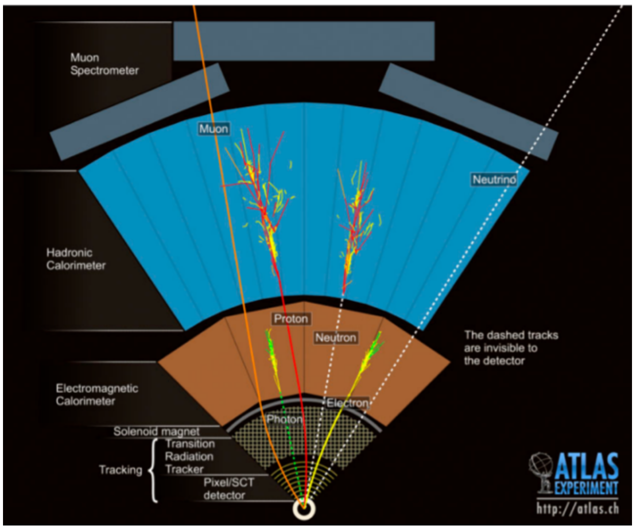

Fig. 1. Overview of the different layer of the ATLAS detector. Retrieved from: http://atlas-opendata.web.cern.ch/atlas-opendata/webanalysis/

1.2 The Standard Model

The Standard Model is a rigorously tested mathematical model that describes the properties and interactions of elementary particles. It is a quantum field theory, relating The standard model is composed of three generations of quarks and leptons, four gauge bosons and the Higgs boson (see Fig. 2). The formulation of The mathematical formula used to make calculations for the standard model is known as the Lagrangian equation. The Standard Model has resulted in the prediction and subsequent discovery of the Higgs Boson; the mode of interaction between energy and space-time, resulting in particles with mass.

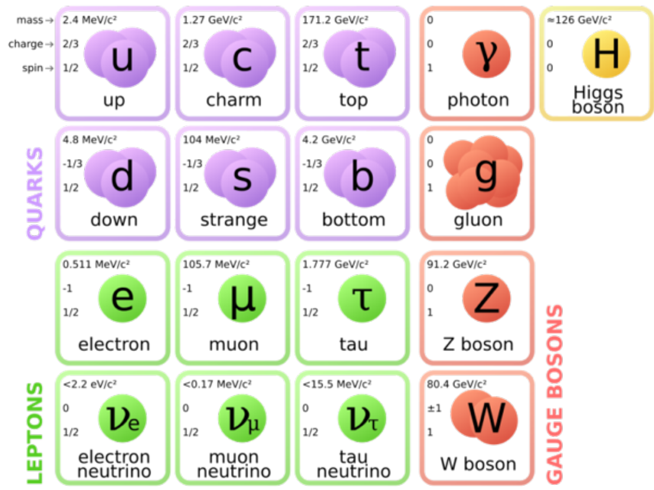

Fig. 2. The fermions and bosons described by the Standard Model

Every elementary particle can be said to have a corresponding field residing within space-time. The motion of particles can be described by ripples in their respective fields. Each of these fields is considered to be a quantum mechanical operator that has the ability to create particles out of the ground-state or vacuum (therefore, is able to violate the law of conservation of mass and momentum). The creation of particles is comparable to ripples in the field that move through space-time. (Raby & Slansky, 1997) There are four fundamental quantum fields for mass particles (quarks and leptons), electroweak bosons, gluons, and the Higgs boson.

The up and down quarks make up the first generation of particles and are the lightest and most stable. All stable matter in the universe consists of first generation particles. Second (c and s) and third (t and b) generations are less stable and heavier. Quarks come in different colours and only mix in such ways as to form colourless objects.

Leptons are also arranged in three generations (1st column 1st generation, etc.). The electron, muon and tau are electrically charged and have mass, but the neutrinos are neutral and have very little mass.

Weak, electromagnetic and strong forces result from the exchange of force-carrier particles, which are all bosons. The strong force is carried by the gluon, electromagnetic force by photon and weak force by W and Z bosons.

The Higgs boson is one of the most recent discoveries in the field of particle physics and It has been postulated that its interaction with the Higgs field creates a friction on the particle. This friction limits the particles energy so that it can no longer travel at the speed of light, and consequently creates mass. (Melnikov & Zanderighi, 2010)

Mathematics behind the measurements

Protons are composed of quarks and gluons which are released in "jets" at the point of collision. The particles of interest, which scientists intend to study, are usually present directly after the collisions. These particles then decay into other particles (usually more stable ones with lighter mass) with the mediation of force carriers, or Bosons. This process can be best described with a feynman diagram.

Importantly, charge and momentum are conserved during the decay process, allowing us to extract the invariant mass of the original particle. The decay process can happen several times before anything is measured in the detector. Furthermore, several decay processes can produce the same stable particles. Considering these factors, it can be difficult to isolate the target decay process when there are millions of events happening in each proton bunch. In this sense there is a considerable quantity of background "noise" which must be filtered out of the data to correctly identify processes of interest.

Background signals

When scientists first developed the standard model, many of the predicted elementary particles had not yet been observed or studied. When these particles slowly began to be understood and tested against the standard model, the questions surrounding them dissipated as well and became factual interpretation of large sums of data. Once predictions have been validated sufficiently, the particles are accepted as part of the standard model and become so-called "background processes". In other words, the particles which are well established and understood are no longer the target of measurements.

What physically differentiates the signal from the background is that the signal must be equal to, or in excess of five sigma difference between itself and the background. This enforces a scientific standard for uncertainty in measurements and prevents improperly validated speculation.

When searching for particular signals (for example, two bosons decaying into four leptons) one must consider that many other decay process produce the same stable particles. In order to differentiate the ZZ signal from these background events, ATLAS strategically looks at data in regions which produce the same leptons, and then extrapolates that to the signal region.

As the Large Hadron Collider accellerates particles close to the speed of light, it leaves the territory of classical mechanics and enters that of relativity. This manifests mathematically and provides new tools to calculate and predict important features of a collision. One of the most elementary changes brought about by this change is the use of 4-vectors instead of 3-vectors (Griffiths, 1987). 3-vectors as used in classical mechanics only have three components which represent the x, the y and the z axes. 4 vectors on the other hand, while still representing the same thing as its counterpart 3-vector, make use of time as their 0th component. The time component is multiplied by the speed of light to give it the same units as the other components. 4 vectors are useful as the momentum 4 vector makes it relatively straightforward to do momentum calculations in collisions. The momentum 4-vector has energy as its 0th component, which proves extremely useful due to Einsteins famous equation

$$E^2 = (pc)^2 + (mc^2)^2 \tag{1}$$

Using this, we can see that at relativistic speeds, the momentum, mass and energy are closely related. Using this relationship and energy/momentum conservation laws, one can reconstruct collision events from the fragments encountered by the detectors more easily.

In order to reconstruct events, we first and foremost need data. What data we need depends on the situation we are looking at. The most common data made use of at CERN is the invariant mass, the missing transverse energy, and the reconstructed mass.

The Invariant Mass or relativistic mass needs to be introduced as mass changes as a particle accelerates. The invariant mass is the mass of a body at rest (Helliwell, 2010). The mass at a certain velocity can then be calculated by transforming the rest mass by a factor of gamma. In a decay, invariant mass is conserved, so it can be used to reconstruct the mass of the decaying particle. It is assumed in most cases that the total mass of the fragments of a decay is equal to the mass of the decaying particle. Off-shell particles and relativistic masses provide exceptions to this, but a histogram of the total invariant mass often results in a bell curve with a peak at the mass of the decaying particle.

Sometimes, one encounters a case where the momenta of the fragments dont add up to that of the colliding particles. This manifests itself as missing transverse energy. This phenomenon is due to the production of particles that cannot be detected by the detectors. A prime example is the neutrino, which is capable of passing through large bodies of matter without interacting with it. Neutrino production is observed quite often in several phenomenon such as in W decays. By calculating exactly how much energy is missing in the collision, one can use the missing energy as a value for the invariant mass of the produced neutrino and carry out calculations to recreate the original decay from this information. Since the actual mass of the neutrino is not known, we can calculate a reconstructed mass for the decaying particle (the W boson in this case). This mass is known as the reconstructed mass or m. As an example to illustrate this, the formula (*Energy measurement of W- and Z- bosons*, 13ADAD) to calculate the reconstructed transverse mass of a W boson is as follows

$$m_T \quad = \quad \sqrt{2p_T^l p_T^v(1 - cos\Delta\phi_{lv})} \tag{2}$$

where p_T^l represents the momentum of the lepton, p_T^v represents the momentum of the neutrino (which is assumed to be the missing energy), and $\Delta\phi_{lv}$ represents the angle between the lepton and the neutrinos trajectory.

2 Computational Methods

Particle physics is highly dependent on competent programming and software engineering to make sense of the quantum world that scientists seek to explore. Programming is a vastly adaptable and versatile tool that enables data collection, storage, management, modelling, analysis, plotting and fitting.

This project utilized Ubuntu, which is a Linux operating system, for managing the large sums of data from ATLAS in a more hands-on, customizable environment. It was necessary to install an OS with a C++ compiler in order to run a piece of code that processed all the events in a number of selected datasets. This processed data could then be graphed in Histograms allowing for further analytical processing. The programming environment terminal was used, which is a command-line window that compiles C++. C++ is a useful language in this scenario as it allows for classes to inherit properties from their base class. For example, the class Z Boson extends the base class Boson and therefore inherits an integer only spin.

In most cases the events that occur in the LHC are very unpredictable processes and often only part of the entire process is recorded by the particle detectors. For this reason, Monte-Carlo simulation methods are used to model and infer the outcomes of the full range of processes. These simulate the conditions and parameters that are conducive for new-physics processes and shows scientists what to expect when the actual data is collected.

2.1 Statistics

New events are observed continuously, but in order to identify whether these new effects are significant or not to the Standard Model a key requirement is to propose a null hypothesis, stating that there are no fluctuations in the data set, and an alternative hypothesis, that there has been a fluctuations in the data set. In the interest of finding new particles the null hypothesis must be rejected, hence the event selected will be found in the tail of the distribution. In other words significance is the probability that by mistake one rejects the null hypothesis. Typically a 5 value is used to assess the confidence level, p-value, of the data collection. In high energy physics (HED) the selection of a steep confidence level is done to distinguish whether it is a measurement (5), an observation (a p-value of 4 is needed) or if there is evidence (a p-value of 3 is needed) for the event taking place if more time were invested in the researched of this physical effect. (http://www.hep.caltech.edu/ fcp/statistics/lectures0802/L0802B.pdf)

To perform more meaningful significance estimates a blind analysis approach is usually always conducted, in essence designing the analysis before a glimpse is even taken into the results. The reason is to handle the data in this manner is to already know the specific cuts to make so reducing the possibility of the wanted or another interesting signal which in the future might be an interesting background. The cuts are picked following the extensive data provided by MonteCarlo simulations, the already known background processes and furthermore QED behaviors involved with the physical variable under consideration which may have led to a systematic error.

The level of significance, how well the cuts were carried out, is given by the score function. The score function is defined as the sum of the background events, B, and the score function, S, will be equal to the total number of events, N. Hence:

$$S = N - B \tag{3}$$

Therefore the uncertainty of S will be the following:

[....]

3 Background Processes/Signals

3.1 Z Decay

Being the heavy analog of the photon, the neutral Z boson has a mass of $m_Z \cong 91.2GeV$ and a cross-section of $\Gamma_Z \cong 2.5GeV$. It decays either hadronically or leptonically with a very short half life of $\tau = 1/\Gamma \cong 10^{-25}sec$. This analysis will focus on the leptonic decay mode into two oppositely charged leptons of same flavour as shown in the Feynman diagram in Fig. 3.

For the single Z analysis the following search criteria were employed:

- Two good leptons with $p_T > 25$ GeV
- Leptons must have opposite charge
- Leptons were required to have same flavour

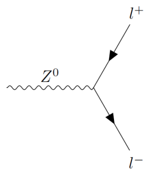

Fig. 3. Feynman diagram for Z^0 decay

$Z \quad \rightarrow \quad e^- \quad + \quad e^+$

The Zee process is one of the main background processes when trying to extract events with two electrons in their final state since it has a relatively high cross section. In order to filter specifically for Zee processes the two good leptons were required to be of lepton type 11, which stands for electrons. The results of running this search on the Zee Monte-Carlo simulation are shown below. Fig. 4 plots the frequency distribution against mass (GeV) and a distinct peak at ∼90 GeV can be seen, which is a good estimate for the Z mass.

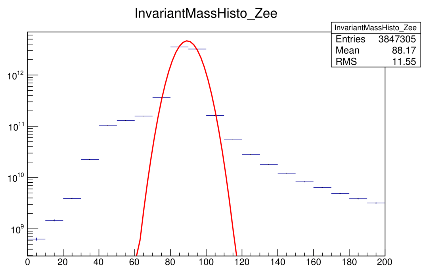

Fig. 4. Invariant mass histogram for Zee decay

The analysis of background processes also requires the familiarization with plots of different parameter spaces such as plotting the frequency distribution against missing transverse energy. This plot, visualized in Fig. 5, exhibits a linear decrease in MET until 200 GeV and then contains a peak centered around roughly 250 GeV.

Furthermore, the number of jets was plotted against frequency and is demonstrated in Fig. 6. It can be concluded, that small jets have a much higher prevalence for Zee processes than large number of jets. All in all, the conducted Z analysis revealed distinct patterns in the frequency plots of MET, invariant mass and number of jets that can be exploited in subsequent searches by introducing cutoff values for the parameters.

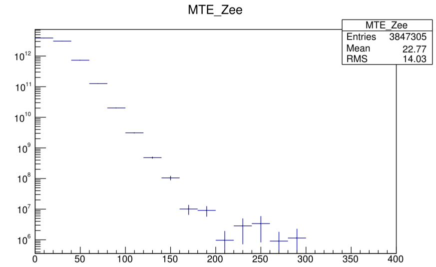

Fig. 5. Missing transverse energy histogram for Zee process

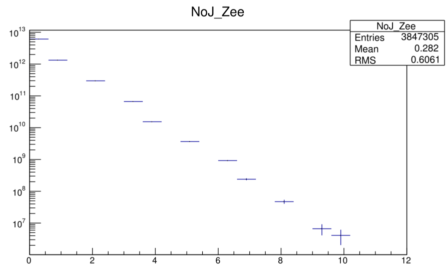

Fig. 6. Histogram for the number of jets for the Zee process

$Z \quad \rightarrow \quad \mu^- \quad + \quad \mu^+$

Since the Z boson can decay into any two leptons of the same flavour, the decay into two muons (μ), each with a mass of 105.7 MeV, also needs to be considered. Compared to the previous search, only the lepton type was changed to 13, signifying muons instead of electrons. The search was conducted on the Zmumu MC data set and the invariant mass was again plotted against the frequency distribution and the resulting histogram is shown in Fig. 7. The Gaussian fit clearly shows a peak centered around ~90 GeV which constitutes another good approximation of the Z mass. It should be noted that Fig. 4 and 7 look very similar, implying that the two processes have roughly the same cross section. [can I say that?]

Fig. 8 shows the MTE histogram, which demonstrates that there is much less MTE in Zmumu processes compared to Zee processes. After a small peak at 30 GeV the MTE frequency rapidly approaches zero and all MTE is found below 100 GeV. Since there are no neutrinos involved in the Zmumu decay process, no MTE is expected and, hence, the entire MET in Zee and Zmumu processes is considered fake MTE. This special type of MTE is attributed to wrong measurements and errors such as incorrectly classifying jet energy as MTE. A distinct feature of fake MTE is its fast convergence to zero and the resulting small tails in the MTE histogram as seen in Fig. 8.

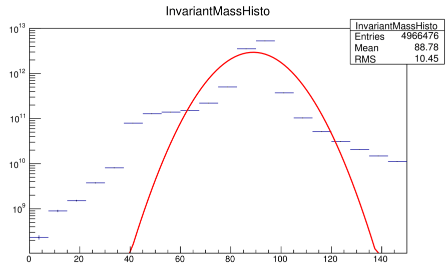

Fig. 7. Invariant mass histogram for Zmumu decay

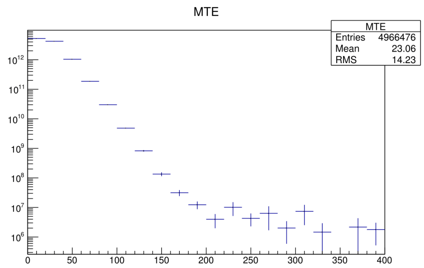

Fig. 8. Missing transverse energy histogram for Zmumu process

When incorrectly measured jet energy is indeed the cause of the MTE distribution obtained in Fig. 8, it expected that the MTE vector will mostly point in the direction of the most energetic jet. Therefore, when calculating the difference in azimuthal angle between the MTE vector and the most energetic jet, by definition given by the 0^{th} entry in the jet_phi vector, a decreasing frequency distribution is expected with increasing radians. This prediction is verified by the histogram shown in Fig. 9. Oscillations are visible in the frequency distribution but no distinct linear increase with radians can be observed. To verify these result the azimuthal angular differences between the second most energetic jets and the MTE vectors were computed and are demonstrated in Fig. 7. Again no increase with radians is observed which implies that the fake MTE shown in Fig. 8 is probably due to other reasons than incorrectly measured jet energy.

$Z \quad \rightarrow \quad \tau^- \quad + \quad \tau^+$

Besides decaying into electrons and muons, the Z boson can also decay into two tau leptons. However, since tau leptons are very heavy compared to electrons and muons, their cross section and contribution to the background will be relatively low. Furthermore, searching for tau leptons requires much more sophisticated reconstruction methods and it is therefore much more difficult than searching electrons and muons in the final states. Due to all these reasons this decay mode was not searched and analysed further in this paper.

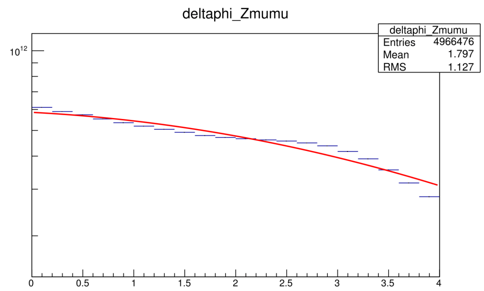

Fig. 9. Distribution of angular difference in azimuthal angle plotted against radians

3.2 ZZ Decay

The decay of two Z bosons is one of the main background processes when searching for the Higgs boson. The decay of ZZ into four good leptons with pT ¿ 10 GeV will be looked at with the following search criteria:

[search criteria]

[Feynman diagram]

The two Z candidates were required to be built from lepton pairs of opposite charge and same flavour. This minimizes the total deviation of the Z candidates from the mass of the Z boson. The maximal total deviation for the ZZ analysis is set to 20 GeV:

$$|massZcandidate1 - massZ| + |massZcandidate2 - massZ| < 20GeV. \quad (4)$$

Since the mass of a single Z is equal to ∼90.2 GeV (reference), the invariant mass of ZZ is always greater than ∼180 GeV. This is confirmed in Fig. 10. As it can be seen from the landau fit, the frequency is greatest at ∼220 GeV and drops after this point.

The number of jets were also plotted, and resulted in a very low number which can be neglected (see Fig. 12). This result makes sense, because no neutrinos are produced in the ZZ to 4 leptons process.

3.3 W Analysis

The spin-1 W^+ and W^- are the only charged bosons in the Standard Model and together with the neutral Z boson they mediate the weak interaction. W bosons have a mass of $m_W \cong 80.4GeV$ and a cross-section of $\Gamma_Z \cong 2.1GeV$. With a branching fraction of $\sim 67\%$ the W boson mostly decays hadronically, however,

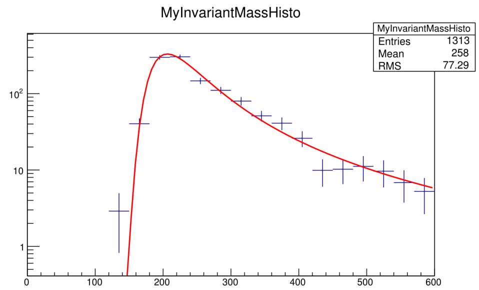

Fig. 10. Invariant mass histogram for the ZZ decay

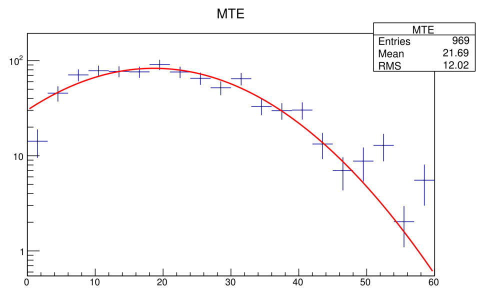

Fig. 11. Missing transverse energy for the ZZ process

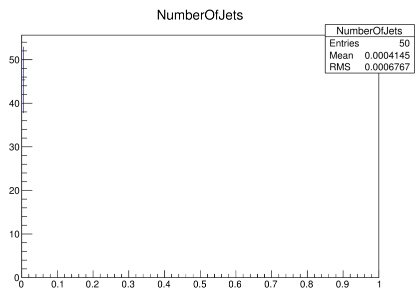

Fig. 12. Histogram for the number of jets for the ZZ decay

this analysis will only focus on the leptonic decay with $\sim 11\%$ branching fraction for each lepton generation (Olive et al., 2014). Depending on its charge the W boson decays into either a positive lepton and a neutrino or a negative lepton and an anti-neutrino as depicted in the Feynman diagrams below (Fig. 14 and 13). Similar to the Z, the W bosons have a very short half life of $\tau = 1/\Gamma \cong 10^{-25} sec.$

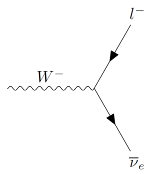

Fig. 13. Feynman diagram for W^- decay

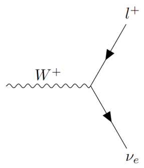

Fig. 14. Feynman diagram for W^+ decay

Contrarily to the Z analysis, no invariant mass histograms can be obtained from the W analysis since only one lepton can be detected and measured in the final state whilst the neutrino escapes the detectors without measurement. Hence, the only way to approximate the momentum of the neutrino is to attribute all the $\not{E}_T$ to the neutrino and based on this assumption the $m_t(W)$ can be calculated.

For the single W analysis the following search criteria were employed:

- Exactly one good lepton with $p_T > 25$ GeV
- $\not{E}_T > 30$ GeV
- $m_t(W) > 30$ GeV

The obtained histogram of m_t is shown in Fig. 15 and the gaussian fit shows a clear peak close to the W mass given in the literature. Fig. 16 shows the W analysis results that was obtained from all events and visualizes the accuracy of the MC simulation when overlayed with the actual data from the Eγ-channel. Furthermore, it demonstrates that single top decay is often misinterpreted as the decay of a W boson.

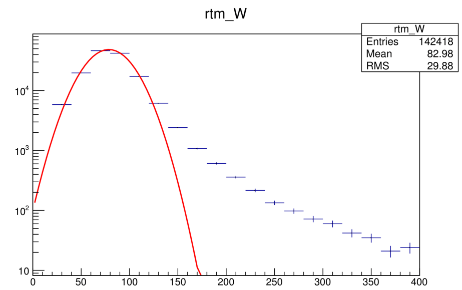

Fig. 15. Histogram of m_t for the W decay

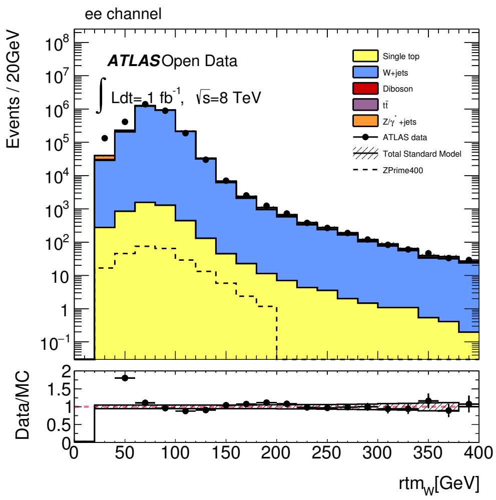

Fig. 16. Distribution of m_t over all events

The $\not{E}_T$ histogram is shown in Fig. **??** and allows for the same conclusions regarding the single top decay and the accuracy of the MC simulations as Fig. 16. In contrast to the Z decay, the $\not{E}_T$ is considered 'real' since the undetectable neutrino in the final state carries energy.

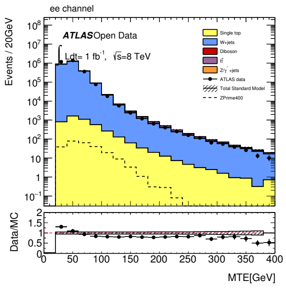

Fig. 17. Missing transverse energy for the W process

3.4 Z' search

References

Aad, G., Abajyan, T., Abbott, B., Abdallah, J., Khalek, S. A., Abdelalim, A., ... others (2012). Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc. *Physics Letters B*, *716*(1), 1–29.

Energy measurement of w- and z- bosons. (13ADAD, Aug). Retrieved from http://physics.stackexchange.com/questions/74099/energy-measurement-of-w-and-z-bosons

Griffiths, D. (1987). Introduction to elementary particles. *John Willey & Sons Inc.*

Helliwell, T. M. (2010). *Special relativity.* Univ Science Books.

Melnikov, K., & Zanderighi, G. (2010). W+ 3 jet production at the lhc as a signal or background. *Physical Review D*, *81*(7), 074025.

Olive, K. A., et al. (2014). Review of Particle Physics. *Chin. Phys.*, *C38*, 090001. doi: 10.1088/1674-1137/38/9/090001

Raby, S., & Slansky, R. (1997). Neutrino masses: How to add them to the standard model. *Los Alamos Science*(25), 64–71.