MSci Major Project Report

Royal Holloway, University of London

Department of Physics

Measuring the mass of the top quark using simulated events at ATLAS

Author:
Sophie Hatia

Supervisor:
Prof. Veronique Boisvert

March 2022

Contents

Abstract

The top quark is the heaviest known elementary particle in the Standard Model. It is the only quark to decay before hadronisation occurs, allowing a unique insight into the properties of bare quarks, a crucial area of knowledge for studying the Standard Model and beyond. The mass of the top quark is not predicted by the Standard Model and therefore must be determined experimentally. It is incredibly useful in probing the electroweak symmetry breaking as well as other areas of beyond the standard model research. This project aims to reconstruct the top quark using simulated data of $t\bar{t}$ decay final state products and compare the reconstructed invariant mass to the true top quark invariant mass for each simulated event.

1 Introduction

The top quark is the heaviest elementary particle in the Standard Model. Its high mass gives it interesting and unique properties that allow a wider insight into the Standard Model. The mass is a key parameter in understanding the electroweak consistency and symmetry and its coupling to the Higgs boson allows for further studies into it. This project aims to discuss the top quark and its mass and to reconstruct the top quark using the data of final state products of simulated $t\bar{t}$ decay events and to compare the invariant mass of the reconstructed top quark to the true top invariant mass for each decay event. This projected is a coding based project conducted in C++ within a unix environment. As the $t\bar{t}$ decay events are simulated using a Monte Carlo simulation method, the final state products are in the form that they would be in an experimental run. Therefore, they are reconstructed from the trajectory of the products. So to analyse the events, plots of the different parameters for each final state product in each event are created. However, as these events are simulated rather than actual experimental results, the true information or the true parameter values of each particle are available. So it is possible to compare the values of the true parameters to the parameter values of the reconstructed particles to analyse the methods and accuracy of the reconstruction.

2 A Brief Introduction to the Standard Model

The Standard Model is the theory that explains and classifies all of the known elementary particles, as well as the four fundamental forces through the particles assigned to them. The four fundamental forces of the universe are the strong force, the weak force, the electromagnetic force and the gravitational force [1]. The strong force is, as the name would suggest, the strongest force in the Standard Model as it deals with the binding of the fundamental particles in order to form larger particles of matter. The weak force, also known as the weak nuclear interaction, is responsible for many fundamental

particle interactions, including particle decay. Electromagnetism acts on charged particles and deals with the attraction of oppositely charged particles and repulsion of like charged particles, such as the positively charge nucleus of an atom and the negatively charged electrons. At high energies the electromagnetic and weak forces combine to form the electroweak force. The final fundamental force is the gravitational force. This is the weakest fundamental force and is always attractive. Its strength increases with the size and the mass of an object, whereas for subatomic levels the effects of gravity are negligible compared to the other fundamental forces. Each force is represented by the exchange of a spin-1 particle, which carries the force, known as a gauge boson. The particle associated with the electromagnetic force is a virtual photon, while the exchange particle associated with the strong force is the gluon. Both of these particles are massless. The weak force is carried by vector bosons, which are known as the W or Z bosons. The W boson can have either a positive or negative charge and is used in weak charged-current interactions (for example the nuclear β-decay). The Z boson, on the other hand, is exclusively neutral and is used in the weak neutral-current interaction. The gravitational force is theorised to have a gauge boson named the graviton, however it has not yet been proven or observed [2][1].

The fundamental particles of the Standard Model are split into two categories: quarks, which experience the strong, weak and electromagnetic forces, and leptons, which do not experience the weak and electromagnetic forces only. Each particle in the Standard Model has an associated anti-particle, which is opposite in all of its physical charges but with the same mass. There are six flavours of quarks (up, down, charm, strange, top, bottom), which are split up into two group types: up-type (with a charge of $+\frac{2}{3}$) and down-type quarks (with a charge of $-\frac{1}{3}$). The up type quarks are the up, charm and top quarks and the down type quarks are the down, strange and bottom quarks. Quarks have an increasing range of masses, with the up quark having the lowest mass ($1.7 - 3.3$ MeV) and the bottom quark having the largest mass (4.81 GeV). Quarks cannot exist alone, they must be bound to either two other quarks (to form a baryon) or to an antiquark (to form a meson). Leptons are also split into two groups: the charged leptons and the neutrinos. There are three different types of charged leptons: the electron, the muon and the tau, which are all negatively charged, while their associated anti-particles are positively charged. The charged leptons increase in mass, meaning that the tau lepton has the highest mass (1777 MeV) and the electron has the lowest mass (0.511 MeV). Within the SM, Neutrinos are assumed to be massless and chargless and do not experience the electromagnetic force. Though it has been shown that the neutrinos are massive and have a mass of < 0.120 eV. The three different neutrino types are related to each of the charged lepton types and there are also associate antiparticles for each type [2][1].

3 LHC and ATLAS

The Large Hadron Collider (LHC) located at CERN is the largest operational proton-proton collider currently and has so far reached a centre-of-mass energy of 13 TeV [2]. It has a circumfrence of 27 km with a luminosity of 10^{34} $\text{cm}^{-2}\text{s}^{-1}$ and each beam has an energy of 6.5 TeV. A schematic diagram of the LHC site is shown in 1. Particles are first accelerated through a Linear accelerator, then move into a Proton Synchrotron Booster as a beam and are injected into a Proton Synchrotron, which is still used as a beam source for lower energy experiments. For high energy experiments, such as those that are conducted at the LHC, the beam is injected into the Super Proton Synchrotron, which is also used for a range of lower energy experiments, before then being injected into the LHC itself. Around the collider there are four beam intersection points and located at each of them are the ALICE, CMS, LHC-b and ATLAS experiments[2]. The protons are accelerated using superconducting magnets, which use liquid nitrogen and liquid helium to cool the magnets to 1.9 K.

The ATLAS detector, also known as the "A Toroidal LHC Apparatus" is the largest particle detector built. It has a width of 44 m, a diameter of 22 m and a weight of 7000 tonnes [3]. It is different from other detectors as it contains a straw tracker as well as a silicon pixel and a silicon strip sensor in the inner tracker. This straw tracker has transition radiation capabilities that allow the detector to discriminate between electrons and pions. Due to the superconducting air-core toroids that make up the system of magnets for the muon chambers, ATLAS differs from some other detectors that may use a second solenoid magnet [4].

The ATLAS experiment is designed with layers so that it can detect many different types of particles. It contains an electromagnetic calorimeter, a hadronic calorimeter, a solenoid and a muon chamber [3]. A schematic diagram of the layout of the ATLAS detectors is shown in 2. As can be seen in the diagram, the ATLAS detector contains a muon detectors surrounding a Hadronic calorimeter (HCAL), which surrounds an Electromagnetic calorimeter (ECAL). This surrounds a solenoid encasing a tracking chamber within it.

4 Theory of the Top Quark

4.1 The Top Quark

The top quark is the heaviest elementary particle known in the Standard Model. It is an up type quark, meaning that it has a charge of $+\frac{2}{3}e$ and is from the third family of elementary particles indicating that it and its associated down type quark (the bottom quark) are the heaviest of the elementary particles. The top and antitop quarks were discovered in 1995 at the Tevatron, a proton antiproton

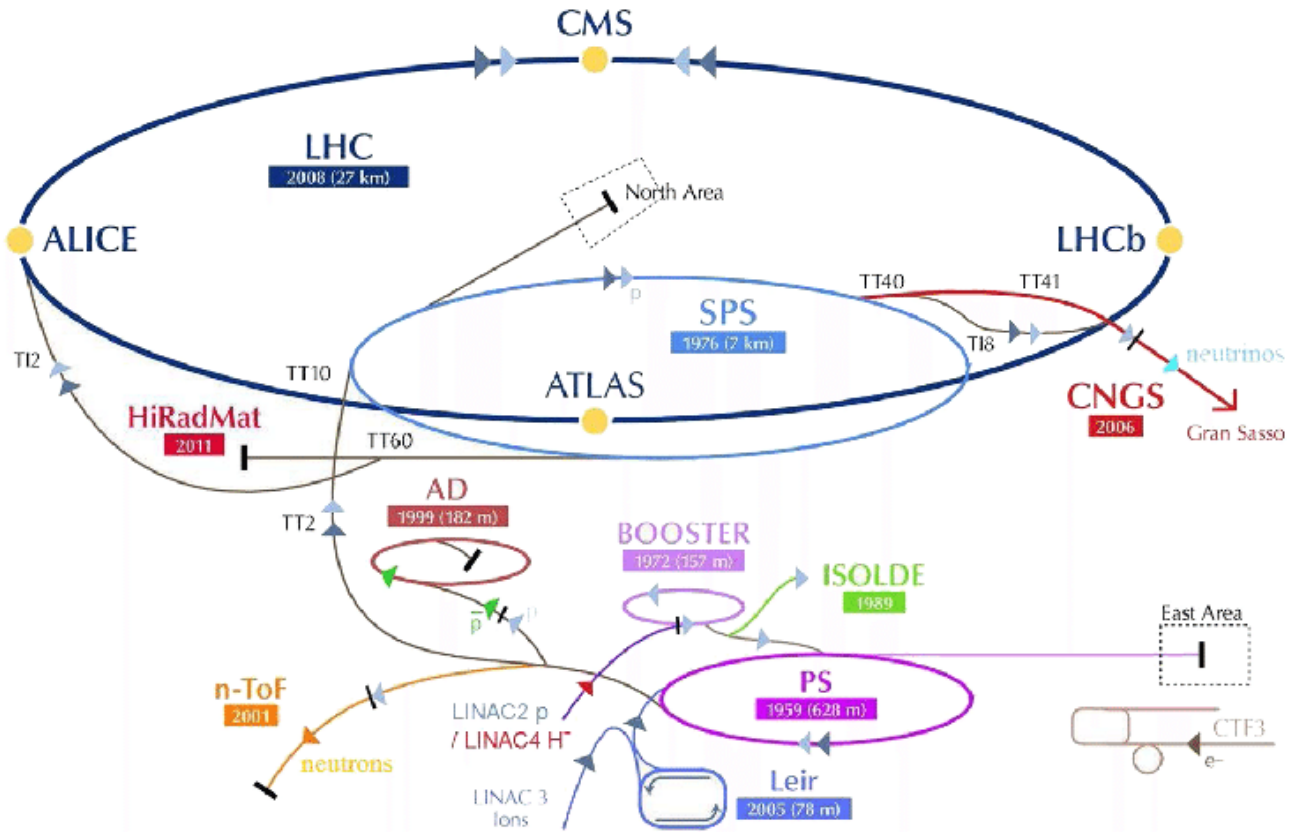

Figure 1: Schematic layout of the LHC at CERN [5]

collider in Chicago, by the Collider Detector at Fermilab (CDF) and the *Dø* experiments. Due to its heavy mass, it decays before hadronisation, making it the only quark to do so. This allows an unique insight into the properties of bare quarks, which are crucial in understanding both the Standard Model and beyond it [7].

The main production modes of the top quark for hadron hadron collisions occur through the strong interaction for top antitop production. These are shown in Figure 3. The Feynman diagram labelled (a) shows the production of a top quark through through gluon gluon fusion for the s channel, in which two gluons fuse , producing a further gluon, which decays into a top quark and an antitop quark. Figure 3(b) shows the production via gluon gluon fusion for the u channel, in which one of the gluons decays into a top and antitop quark, the top quark then couples to the other gluon producing another top quark achieving the same final products as the s channel. Figure 3(c) shows the production via gluon gluon fusion in the t channel. This channel is very similar to the u channel, though the top and antitop quarks are produced parallel to each other rather than being perpendicular at an angle to one another. The final Feynman diagram in Figure 3(d) shows the production of the top and antitop quarks via quark antiquark anihilation. This occurs when a quark and its respective antiquark meet and anihilate, producing a gluon, which then decays into a top quark and an antitop quark. These

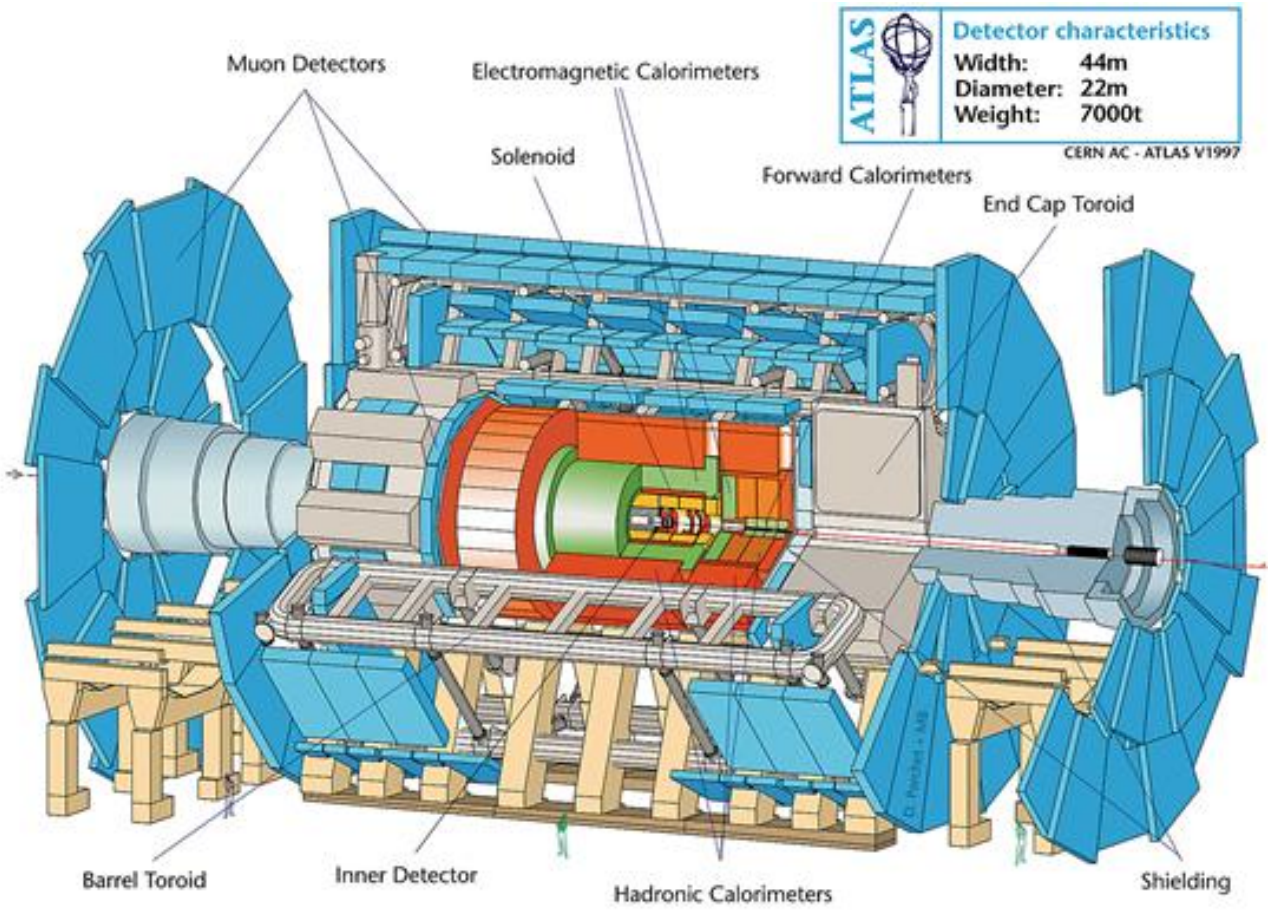

Figure 2: A schematic view of the ATLAS detector at CERN. [6]

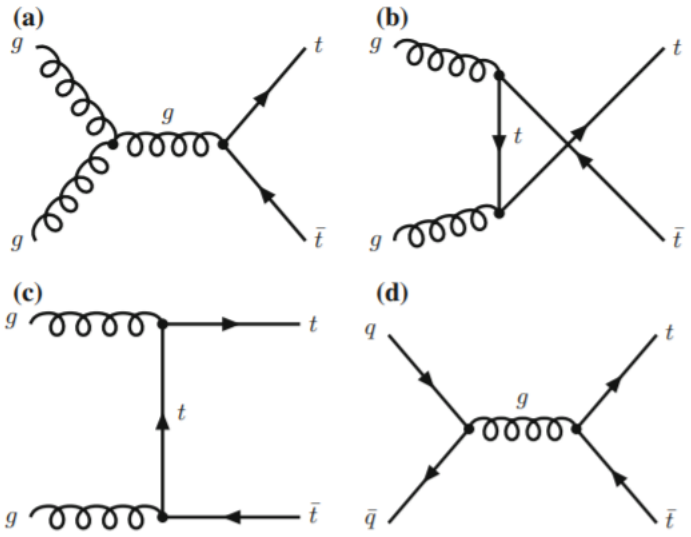

Figure 3: Feynman Diagrams showing the production of a top and antitop quark via gluon gluon fusion for the s channel (a), gluon gluon fusion for the u channel (b), gluon gluon fusion for the t channel (c) and quark antiquark anihilation (d) for hadron hadron collisions [7]

diagrams are at leading order, meaning that there are no loops within the decays. the required energy to produce a $t\bar{t}$ pair is approximately doube the mass of the mass of the top quark. At the LHC 90% of $t\bar{t}$ pairs are produced through gluon gluon fusion. This is because at 8 TeV and $\sqrt{s} = 7$, the luminosity of gluon gluon pairs is higher than that of quarks at the corresponding momentum fractions [7].

4.2 Electro-Weak consistency of the Standard Model

The electroweak model (also known as the Glashow-Salam-Weinberg model) is the unification between the electromagnetic and weak interactions, meaning that the two forces were part of on larger unified whole within the Standard Model. In 1961 the first paper on this unification model was published by Glashow and it was required that neutral interactions and a subsequent paper by Weinberg and Salam in 1968 modified this model to include a physical mechanism that granted the force carriers (the W and Z vector bosons) mass. in 1983 the W and Z bosons were discovered, providing more evidence to this theory and all successive experiments in particle physics have proven this theory to be correct [8].

Shortly after the W boson was first proposed as the associated force carrier for the weak interaction in 1958, Bludman first postulated the idea of neutral currents. It was proposed that the W boson had an electrically neutral partner the Z boson, which would couple to any fermion as demonstrated by Figure 4, where the identity of this fermions is maintained in the interaction. These weak neutral currents were very difficult to observe as they would be saturated with electromagnetic effects, which would also be neutral currents as the photon has zero electric charge. However, in 1973 the first

evidence for the weak neutral currents was observed at CERN via neutrino interactions of the form

$$\nu_\mu + N \rightarrow \nu_\mu + x \tag{1}$$

$$\bar{\nu}_\mu + N \rightarrow \bar{\nu}_\mu + X, \tag{2}$$

where N is a nucleon and X is a hadron. As there were no electrically charged leptons in the final products of the interactions, this proved that the interactions must be conducted via a weak neutral current thus implying that the weak and electromagnetic interactions were part of a larger unified whole [8].

The vertices of the coupling of the W boson to the fermions is shown in Figure5. However, as the W boson is charged it will need to couple to the photon thus a $WW\gamma$ vertex is needed. Therefore a Yang-Mills theory is needed as it allows for vertices with three gauge particles. A Yang-Mills theory is a generalisation of Maxwell's theory in which the gauge symmetry group is non-abelian, meaning that there exists at least one pair of elements within the group that do not commute. A theory was proposed by Glashow, Salam and Weinberg which had a non-abelian symmetry $SU(2) \bigotimes U(1)$ with 3 gauge bosons, $W_\mu^1, W_\mu^2, W_\mu^3 = \overrightarrow{W_\mu}$, associated with the $SU(2)$ group and one gauge boson, B_μ associated with the $U(1)$ group. This theory has come to be known as the electroweak unification theory. This non-abelian theory means that the W and B bosons can interact with each other along with the quarks and leptons, giving the electroweak theory additional vertex rules [8].

However, for this symmetry, the lowest energy state does not reflect the full symmetry of the theory. When implementing the Higgs mechanism into the electroweak theory, it is assumed that the Higgs wavefunction is a doublet under the $SU(2)$ group. When this is applied, it can be seen that the wavefunction for the $U(1)$ group (B_μ) has a current coming from both the left and right handed fermions as well as the Higgs current. However, the $SU(2)$ group wavefunctions ($\overrightarrow{W_\mu}$) only possesses the surrent arising from the left hand fermion as well as the Higgs current. Therefore, the electroweak symmetry is broken [8].

4.3 The Top Mass

The mass of the top quark cannot be calculated from the SM and must be measured experimentally, though it does derive its mass from the Higgs mechanism as other particles do. One mass scheme used to derive to top mass is the pole mass scheme, in which the renormalised pole mass (m_t^{pole}) coincides with the pole of the top quark propagator. This corresponds to measuring the top mass alone in isolation, meaning that it is done without external interactions taking place. This is the most commonly used scheme to determine the top mass. The running top mass has been extracted by some

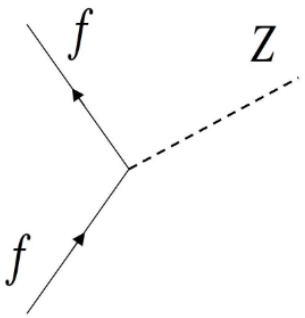

Figure 4: Bludman's proposed coupling for the Z boson [8]

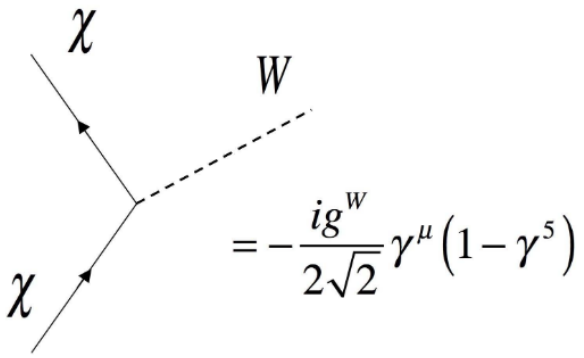

Figure 5: The general form of the vertex between the W boson and the fermion [8]

groups using the modified minima subtraction scheme ($\overline{MS}$), which uses the total production cross section for the top quark pair $t\bar{t}$. These two mass schemes can be connected using the perturbative series

$$m_t^{pole}(k) = m_t^{\overline{MS}}(\mu) \cdot \Big[1 + \Sigma_{n=1}^{k} c_n \Big(\frac{\mu}{m_t^{\overline{MS}}}\Big) \alpha_S^n\Big], \tag{3}$$

where the value of k is the order of the system and μ is the renormalisation scale. The pole mass is well defined for all fixed orders in perturbation theory but cannot be extended to be used for all orders i.e. $k \to \infty$, which is why the modified minima mass scheme is used. A current value of the top mass obtained through the pole mass scheme is $m_t^{pole} = 171.1 \pm 0.4$ (stat.) ± 0.9 (syst.) GeV, while through the modified minima mass scheme it is $m_t^{\overline{MS}} = 162.9 \pm 0.5$ (stat.) ± 1.0 (syst.) GeV [**?**] [**?**].

The mass of the top quark is a very important parameter within the Standard Model. It induces large radiative corrections to the electroweak coupling at low energies as a direct result of the electroweak symmetry breaking. It is likely that the top quark is the most sensitive probe of the electroweak symmetry breaking therefore, in many extensions of the Standard Model it plays a key role. The mass of the top quark gives rise to theories that could be used to sole the hierarchy problem of the Standard Model, in which there is a large discrepance between the electroweak and Planck scale for example, the Higgs mass is far lighter than the Planck mass (also known as the Grand Unification energy or heavy neutrino mass scale) [7]. This is interesting as due to the due to the large quantum contributions to the square of the Higgs mass, the mass of the Higgs itself would be extremely large [9]. In theories that attempt to solve this problem the coupling between the Higgs boson and the top quark is very important [**?**].

Investigations into the mass of the top quark also proved to be useful in determining the theoretical mass of the Higgs boson. Figure 6 shows the contours of the 65% and 95% confidence levels that were obtained using fits of the fixed variables M_W and m_t. The horizontal bands show the 1σ regions of the measurements of M_W and m_t. The green circle in the centre is the direct measurements of M_W and m_t, the larger grey band shows the fits excluding the measurements for the mass of the Higgs, M_H, and the narrower blue band shows the fits including these measurements. The grey diagonal lines show the values of M_W and m_t for different predicted values of M_H. As can be seen, the only diagonal line to pass through all contours and regions of measurements is the line denoting the values for if the mass of the Higgs was $M_H = 125$ GeV, highlighting that this is the true value of the Higgs boson (denoted by the fact that this line is unbroken rather than being dashed as the others are) [10].

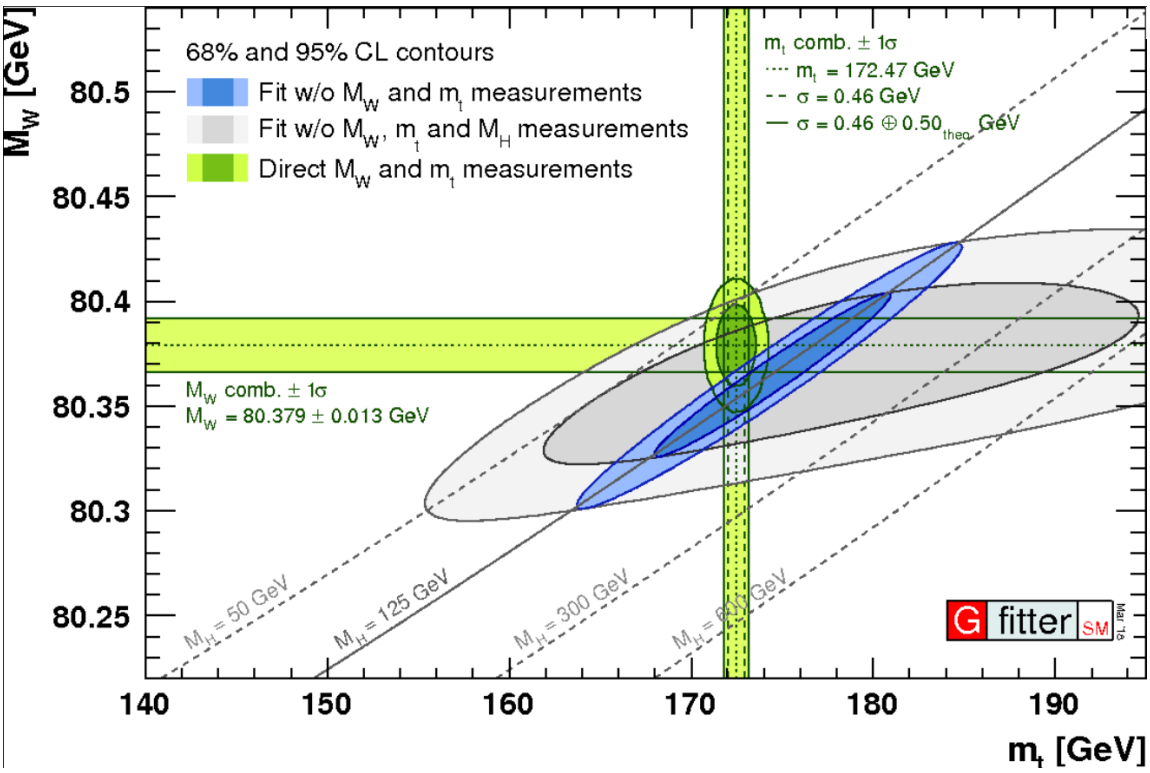

Figure 6: Plot of the contours of the 65% and 95% confidence levels obtained from fits of the variables M_W and m_t, where the horizontal bands are the 1σ regions for the values of M_W and m_t. The blue and grey bands show the fits including and exlcuding the data for M_H respectively [10].

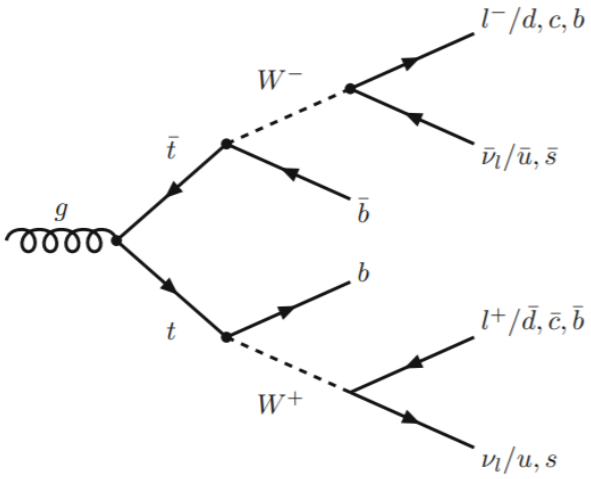

Figure 7: Feynman diagram showing the $t\bar{t}$ decays [7]

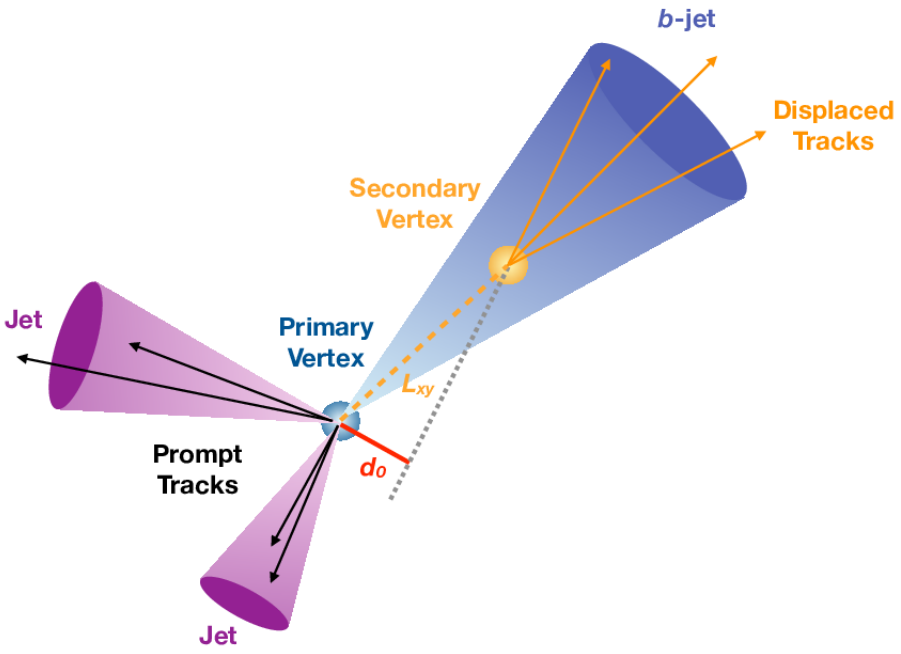

Figure 8: Diagram demonstrating the primary and secondary vertices of a b jet [11]

4.4 $t\bar{t}$ events

Figure 7 shows a Feynman Diagram of the decay of a $t\bar{t}$ production and the different possible products. As can be seen the top quark decays into a W boson and a b-quark. It can also be seen that the W boson decays into either two hadronic jets or into a lepton and its corresponding neutrino. The b-quark then hadronises into a b jet. The products that the W boson decays into determines the type of channel that this decay is known as. These decay channels are known as: the dilepton channel (in which both W bosons decay to a lepton and its corresponding neutrino), the full hadronic channel (in which the W bosons decay into a pair of hadronic jets) and the semi-leptonic channel (also known as the lepton + jets channel), in which one of the W bosons decays into a pair of hadronic jets and the other W boson decays into a lepton and its corresponding neutrino.The dilepton channel contributes approximately 10% of the total $t\bar{t}$ decays whereas, the full hadronic and semi-leptonic channels each contribute approximately 45% to the total [7].

The dilepton channel provides the cleanest signature of all three decay channels despite it's low branching ratio. It is easy to separate from the main background processes due to the high p_T from the two neutrinos and the two energetic b jets. It is also easy to set triggers on this channel (energy conditions that the ATLAS detector imposes on events in order for them to be recorded) due to the

high energy leptons. The main backgound for this channel is the Drell-Yan process, which, due to the difficulty of accurately predicting p_T tails, is often extracted from the data by looking at the p_T tails around the Z boson mass peak. However, the main challenge that arises from this channel (aside from the small branching ratio and lack of statistics) is that the information including the four-momenta of the neutrinos cannot be measured due to them passing through the detector [12].

The full hadronic channel has the largest branching fraction of all the decay modes at 46%. Another advantage is that the energy of the jets can be measured only using the calorimeter, for which the coverage is generally broader than that of the spectrometer for collider detectors. This allows the maximum acceptance to the signal. One can also in principle reconstruct the entire final state kinematics. However, as it is extremely difficult to determine the charge or flavour of a quark from the jets produced it is very difficult to accurately match up the jets produced from the same particle. In the same vein, it is very difficult to distinguish the top quark from the antitop quark so many measurements are rendered unfeasible for this decay mode [12].

The semi-leptonic channel, by comparison, has many advantages: the single energetic lepton provides a simple and convenient sample for creating triggers; the channel occurs in approximately 44% of top antitop decay events, which is a good compromise between the clear signal from the dilepton channel and the high number of events for the full hardonic channel and the kinematics of the top antitop quark final state can be reconstructed due to the fact that the only component of the neutrino that is not measured is the z component. This means that by constraining the neutrino and its lepton to the W boson mass and requiring that the two reconstructed top quarks have the same mass, the kinematics can be reconstructed. As the advantages of this channel fix some of the disadvantages of the other two channels, the simulated $t\bar{t}$ events decay using the semi-leptonic channel [12].

As both b quarks hadronise into jets and one of the W bosons decay into a pair of jets, the events that are necessary to look at in this project are events which contain at least four jets, where exactly two of those are b-tagged jets. It can also be seen that if the charge of the W boson is negative an anti b quark accompanies it and if the charge of the W boson is positive a b quark accompanies it. The charge of the W boson will also affect the charge of the lepton as if the W boson has a positive charge, so too will the lepton and vice versa. Therefore it is possible to determine whether a b jet came from a b or anti b quark by using the charge of the lepton, which will be discussed further in Section 5.3.

As the mass of a b quark is so high it is much larger than its decay products, which therefore mean that the products have higher transverse momentum (p_T), causing the b jets to be wider than average jets and to have higher invariant masses. These properties can be measured and analysed to identify

b-tagged jets. Another property of b quarks is that it hadronises into a "B" hadron, containing a b quark and a an antiquark as well as hadronising into other particles such as a pion. This "B" hadron has a longer decay life time than the other hadrons so after the t or $\bar{t}$ quark decays into a b quark it travels a distance before decaying again. However, this decay lifetime still remains within the detector before decaying. This results in a primary and secondary vertex within the detector, where the primary vertex is the decay from the t quark to the b quark and the secondary vertex is the decay from the b quark to the b jet as can be seen in Fig 8 [11].

5 Reconstructing the Top Mass

5.1 Kinematics of the top decay products

This project is run on simulated data from a Monte Carlo simulation. This means that the uncertainties that arise in the values calculated arise from the uncertainty in the Monte Carlo simulation. The expectation of a Monte Carlo simulation is given by

$$E[M] = \int_x M(x)p(x)dx, \tag{4}$$

where $M(x)$ is a real valued function defined over x and $p(x)$ is the probability. Therefore, the variance of a Monte Carlo simulation is given by

$$V[M] = E[M^2] - E[M]^2. \tag{5}$$

[13]

As this project runs on simulated data, there are the reconstructed particles and jets, which are the objects that have travelled through the detector and have been detected and reconstructed into particles, as well as the true input data of the objects from the simulation. Therefore, it is possible to compare the data of the particles that were reconstructed and the true particles before they decayed or hadronised. This is useful in ensuring that the correct jets are matched up in order to create a reconstructed top quark that matches with the true top quarks of the decay. This is done through comparing the difference in invariant mass, transverse momentum (p_T), and the ΔR.

The ΔR of a pair of particles is the angular separation (or angular distance) between the particles and expresses the extent to which two particles travel along the the same direction as each other. It is defined as

$$\Delta R = \sqrt{(\Delta\eta)^2 - (\Delta\phi)^2}, \tag{6}$$

where $\Delta\eta$ is the pseudorapidity and $\Delta\phi$ is the azimuthal angle measured with respect to the x axis.

The pseudorapidity is an approximation of the rapidity Δy and is related to the polar angle θ through the equation defined as $\eta = -lntan(\frac{\theta}{2})$, where θ is measured along the beam line within the LHC [14].

The total number of events within the data set are 38259 events, this means that the total number of leptons within the data set is also 38259 as there is one lepton per event. The total number of jets, however, is much larger as it is 183464 jets as for most events there are at least four jets, meaning that it would be expected that there would be at least four times the number of events. The total number of b jets is 56987, which is interesting as, though it is larger than the total number of events, it would be expected that there would be approximately double the number of events. However, there is slightly less than double the number of events, implying that not every b jet was accurately identified as a b jet by the b-tagging algorithm for some of the events. Therefore, the total number of non b jets within the data set is 126477. This is interesting as looking purely at the Feynman diagram for this decay mode, it would be assumed that there would be equal numbers of b jets to non b jets as there are four jets within the decay. However, there are far more non b jets to b jets, implying that there are other jets that arise from other places.

Figure 11 shows the minimum ΔR between the true b quark and the reconstructed b jets. As can be seen, there is a large peak at around 0 and then a logarithmic decrease after the peak so that there are very few events in which the ΔR between the reconstructed b jet and true b quark is greater than 0.2 . This is the expected and desired distribution of the ΔR plot for these particles, implying that the b jets have little angular separation to the associated true b quarks. This is further shown by the mean ΔR, 0.04389 ± 0.001809, which is very small and has a small uncertainty.

In contrast Figure 12 shows the distribution of the ΔR between the reconstructed leptons and true leptons. It can be seen that there is a sharp peak at approximately 0 and a rapid decrease past 0.01 so that there are essentially no events with leptons that have a ΔR greater than 0.02 to the true lepton. This is further shown as the mean minimum ΔR of the true and reconstructed leptons is $0.001539 \pm 2.071 \cdot 10^{-5}$, which is much smaller than that of the b jets, highlighting that the reconstructed leptons are less scattered and are more likely to travel in the same direction as the true leptons. This is to be expected as per decay event there is only one lepton, which decays from the leptonic W boson, therefore the reconstructed lepton should be a much more exact match than for the b jets where it is necessary to ensure that the right b jet is matched to the relevant b quark.

Figure 9 shows the relative p_T of the reconstructed leptons to the true leptons. The relative p_T of a reconstructed particle to its true particle is calculated using the equation

$$\frac{P_{T_{reco}} - P_{T_{true}}}{P_{T_{true}}}, \tag{7}$$

where $P_{T_{reco}}$ is the transverse momentum of the recovered particle or jet and $P_{T_{true}}$ is the transverse

	Leptons	Jets
Smallest p_T	27.00 GeV	25.00 Gev
Largest p_T	714.95 GeV	1323.49 GeV

Table 1: Highest and lowest p_T for the leptons versus the jets.

momentum of the true particle.

As can be seen in the plot, there is a large peak at 0 GeV, implying that the majority of the leptons do not lose much momentum due to radiation within the detector. There is a sharp decrease in the number of events with relative p_Ts of a value larger than 0.1 GeV, again highlighting that for most events the leptons gain or lose very little momentum when travelling through the detector. However, it can also be seen that the distribution is slightly skewed in the negative direction, suggesting that if the leptons have a momentum change within the detector, it will be a loss of momentum rather than an increase. This is further demonstrated as the mean relative p_T of the leptons within the detector is -0.02101 ± 0.0004999 GeV, which conveys that the average momentum change is a very small loss in momentum through the detector.

In contrast, the relative p_T distribution for the b jets is shown in Figure 10. As there are two b quarks in the $t\bar{t}$ decay being simulated, it is necessary to ensure that the b jet being analysed is the correct b jet for the specific b quark for the purely hadronic top decay. Therefore, a "best case scenario" condition is imposed upon the b jets so that the jets analysed have a ΔR of less than 0.4 with the true b or $\bar{b}$ quark as 0.4 is roughly the width of a jet.

As can be seen in the figure, there is a peak around 0 GeV, implying that the highest number of b jets have a very small energy change when hadronising to a jet within the detector. However, it is also clear that there is a negative skew within the data as there is also a large tail below 0 GeV, showing that a lot of b jets lose energy when hadronising and travelling through the beam pipe. In contrast, there is a sharp decrease in events with a positive relative p_T, suggesting that very few b jets gain energy when hadronising in the detector. This is further shown as the mean relative p_T is -0.1287 ± 0.00121 Gev, which is larger than the mean p_T of the leptons, again demonstrating that the b jets lose more energy when travelling through the detector.

Table 1 shows the highest and lowest values of the p_T for the leptons versus the jets. As can be seen the p_T range of the jets is much higher at it ranges from 25.00 Gev to 1323.49 GeV, whereas the p_T for the leptons ranges from 27.00 GeV to 714.95 GeV. This is to be expected as there are far more jets produced in the events as leptons as there is only one lepton produced in the event whereas the total number of jets in the event is at least four.

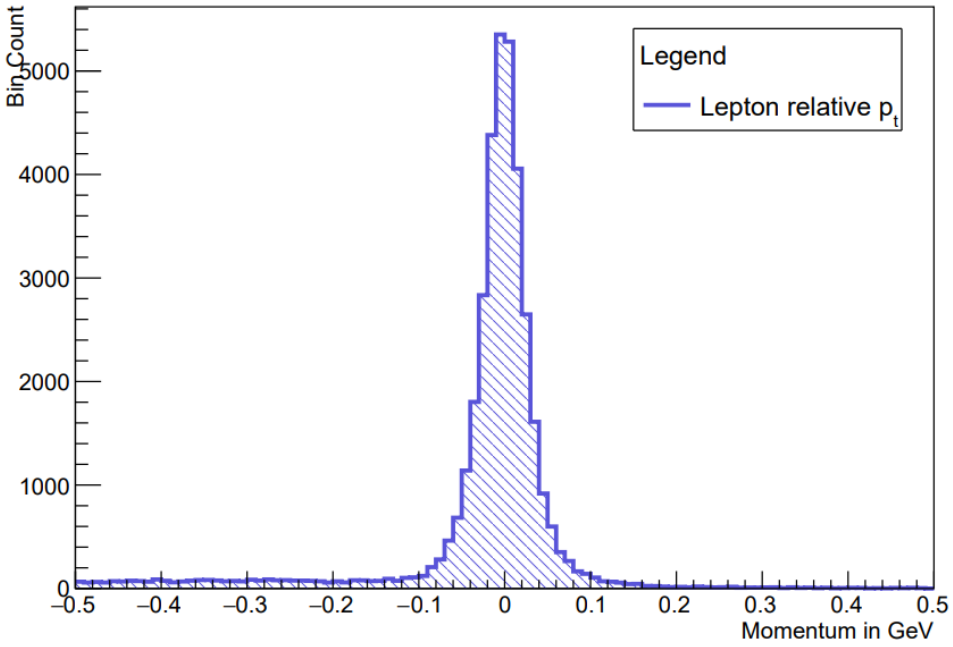

Figure 9: Plot of the relative p_T of the reconstructed leptons versus the true leptons

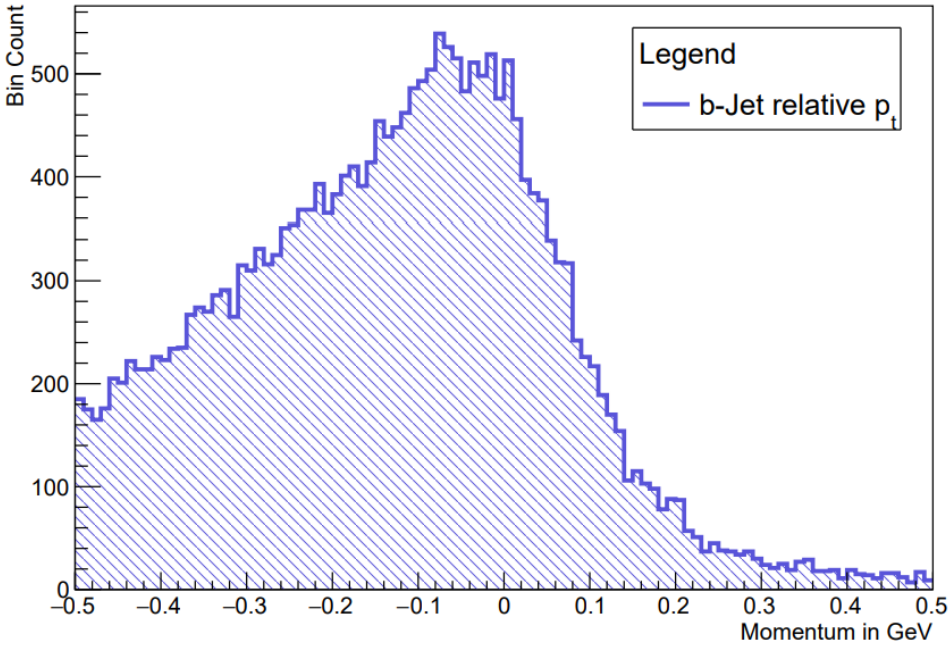

Figure 10: Plot of the relative p_T of the reconstructed b jets compared to the true b quarks

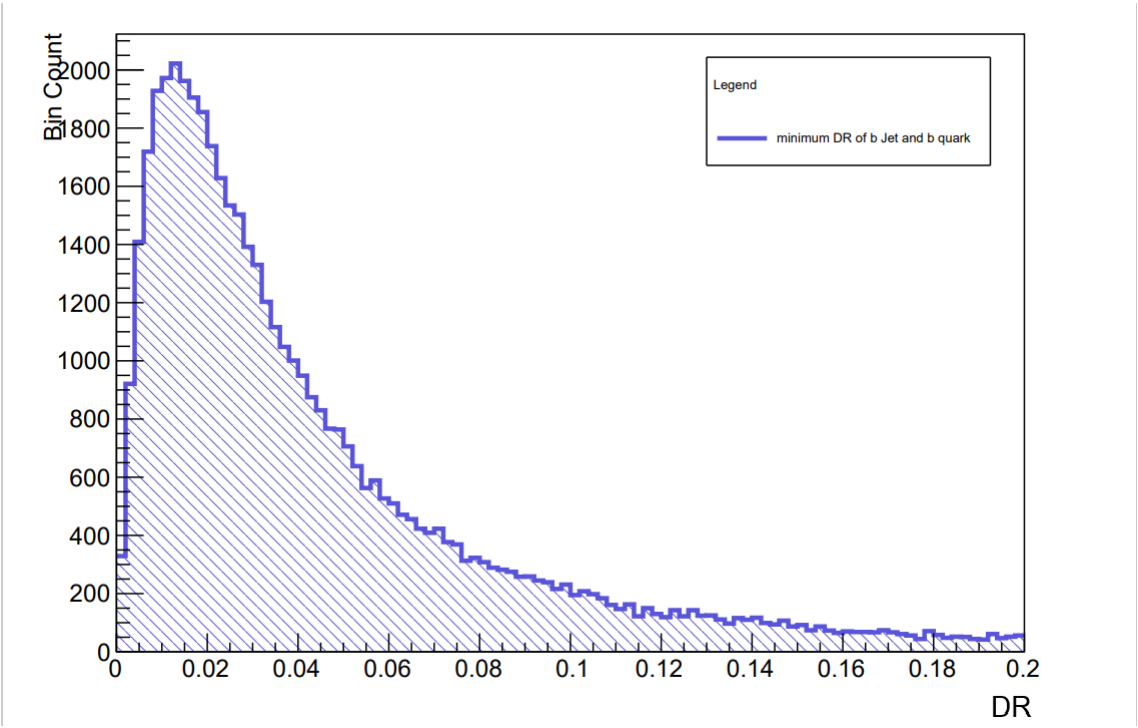

Figure 11: Plot of the ΔR of the reconstructed b jets and the true b quarks

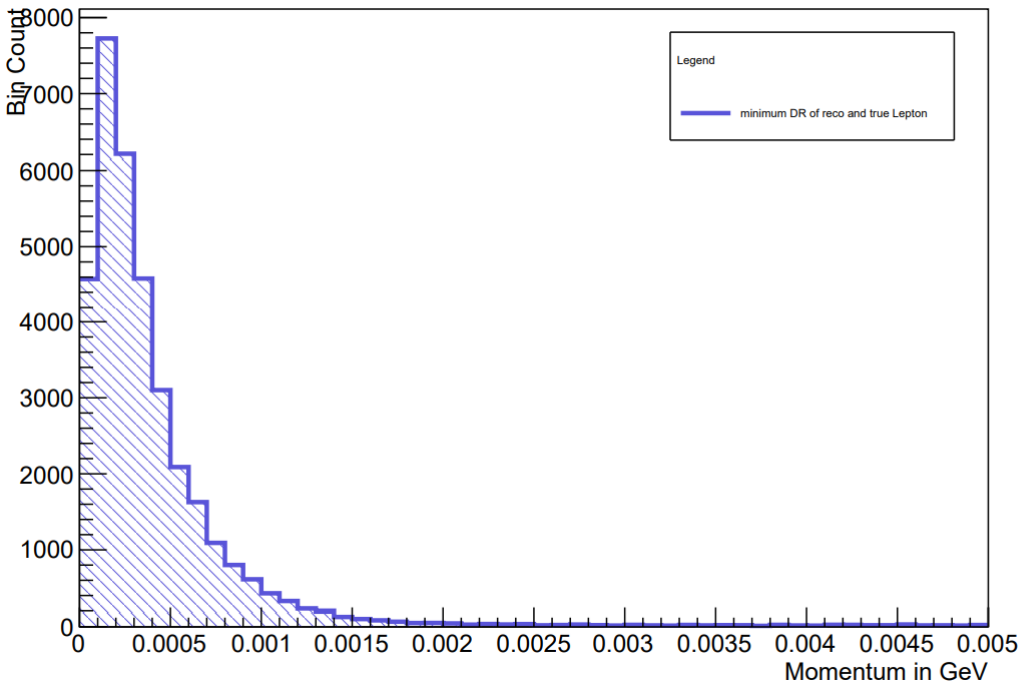

Figure 12: Plot of the ΔR between the reconstructed leptons and the true leptons

5.2 Selecting the W-boson

As there are multiple non b-tagged jets in each event, it is important to identify which jets arise from the W boson and which arise from radiation from other particles. Two methods were employed to do this: one in which the two jets with the highest p_T are combined and the invariant mass is taken and another in which the jets are combined in every possible way and the invariant mass of each is taken. The difference between this invariant mass and 80 (the true invariant mass of a W boson) is calculated and the jet pair with the smallest absolute difference is chosen. As each method required reconstructing a jet pair and taking the invariant mass difference, only events with more than one non b-tagged jet were considered. Furthermore, the method of finding the jet pair with the closest invariant mass to 80 GeV required finding the invariant mass difference for each jet pair therefore, the number of combinations of jets increases exponentially as the number of non b-tagged jets increases. This means that for events with four non b-tagged jets or higher, the number of jet pair combinations are too many to calculate the invariant mass difference for every jet pair and so these events are not included in the reconstructions for this method.

The ΔR of each method and the true W boson is then found and plotted. These histograms are shown in Figure 13, where the method of the highest p_T jets is shown in red and the method of the jetpair with the invariant mass closest to 80 GeV is shown in blue. As can be seen from the plot the distributions of each method are very similar, however, the method of combining the highest p_T jets has many more entries and will therefore give a more accurate mean for the Δ R. The mean ΔR of the method using the highest p_T jets is 1.318 ± 0.005354 with an RMS of 1.03 ± 0.003786 and the mean ΔR of the method of finding the jet pair with the closest invariant mass to 80 GeV is 1.31 ± 0.007698 with an RMS of 1.153 ± 0.005443. Therefore, as the two methods both produce very similar means, where the uncertainties close enough to the difference so that it can be seen as insignificant, the method of using the highest p_T jets was chosen as it gave a more accurate mean due to the higher number of events included in the calculation as well as having a lower RMS with a smaller uncertainty.

5.3 Associating the reconstructed W-boson with the correct b-jet

As both the top and anti top quarks decay into b quarks, it is necessary to correctly select the b quark that is associated with the hadronic W boson. This can by done by using two methods: selecting the b jet with the smallest ΔR to the reconstructed W boson or selecting the b jet with the largest ΔR to the reconstructed lepton. This is because the b jet would generally be closer to the W boson as they would have decayed from the same top quark whereas, the b jet would be further away from the lepton as they decayed from separate top quarks. For the first method, the Delta R of each b jet

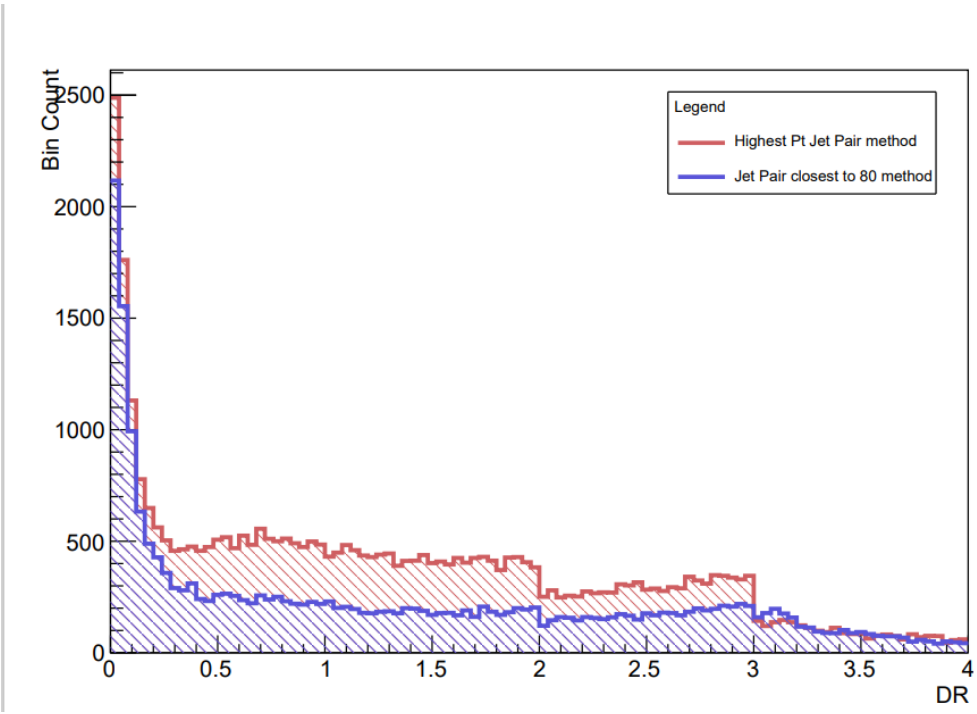

Figure 13: Overlayed distributions of the ΔR between the true W boson and the reconstructed boson via two different methods. The method in which the reconstructed W is composed of the jets with the highest p_T is shown in red and the method in which the jet pair with an invariant mass closest to the invariant mass of a W boson (80 GeV) was chosen is shown in blue.

and the W boson, reconstructed using the previously chosen method (as discussed in section 5.2), was calculated and the minimum value for the event was selected. The b jet associated with this minimum value was then selected and the ΔR between this b jet and the true b quark was calculated. For the second method the ΔR of each b jet with the reconstructed lepton was calculated and the minimum value for each event was found and the b jet associated with this value was then selected. The ΔR of this b jet with the true b quark was then calculated also. The plots for each method are shown in Figure 14, where the method of determining the largest ΔR to the lepton is show in blue and the method of determining the smallest ΔR to the reconstructed W boson is shown in red.

As can be seen in the graphs the distributions of both methods are very similar and both have the expected distribution as both have a large peak around 0. This implies that both methods are fairly accurate for determining the correct b jet for the system as the desired ΔR is 0 as it indicates that there is no distance between the reconstructed jet and the true b quark. As both methods have such similar distributions, it is clear that both methods are able to find the correct b jets that match to the desired b quark with near identical accuracy. However, the method in which the b jet with the largest ΔR to the lepton has a slightly larger peak at zero and can be seen to have slightly fewer events with a ΔR greater than 0.4, implying that it is slightly better at determining which b jet is associated with the relevant b quark. This is further demonstrated by the means of the two methods, 1.035 ± 0.006674 for the method choosing the b jet closest to the reconstructed W boson and 0.9574 ± 0.006547 for the method choosing the b jet furthest from the reconstructed lepton. These highlight that though the two methods produce considerably similar results as the two mean ΔRs are within approximately 0.08 of each other, the method of selecting the b jet by comparing it to the lepton has a lower mean and a marginally smaller uncertainty. Therefore they imply that the method of comparing the b jet to the lepton is slightly more accurate. The RMS of for the W boson method is 1.225 ± 0.00472, which is larger than the RMS for the method involving the lepton (1.201 ± 0.004629) which also has a smaller uncertainty as well, again indicating that the lepton comparison method is slightly more accurate. This is consistent with the expectations as the recovered lepton is much more accurate to the true lepton by way of it being a single object that was reconstructed by the simulation rather than a combination of jets selected through imposed conditions.

Figure 19 shows the distributions of the difference in invariant mass between the true b quark and the chosen b jet. As can be seen from the figure, there is a large peak between approximately -5 and 0 GeV, conveying that most of the b jets chosen have invariant masses very close to the true invariant mass of the b quark. There is a tail after the peak implying there are a number of jets that have masses fairly significantly larger than the true b quark. However, the tail does not extend very far past -30

GeV, highlighting that the difference in invariant masses for both methods are reasonably accurate in selecting the relevant b jet to match the hadronic W boson for a certain decay event. It can be seen that for the method in which the b jet with the smallest ΔR to the reconstructed W boson has a slightly larger peak around 0 GeV while it has fewer events that correspond to the larger invariant mass differences. This implies that this method is slightly better at identifying and selecting the jets that correspond to the correct b quark for the purely hadronic top decay. This is further shown as the mean invariant mass difference for the method of comparing the b jet to the W boson is -4.128 ± 0.02667 GeV with an RMS of 4.916 ± 0.01886 GeV, whereas the mean invariant mass difference for the lepton comparison method is -4.327 ± 0.02829 GeV with an RMS of 5.213 ± 0.02. Both the mean and RMS of the W boson comparison method are lower than those of the lepton comparison method and have slightly smaller uncertainties. This is in contrast to the values of the mean ΔR for the two methods wherein the method with the smaller values was the method of comparing the b jet to the lepton. However, much like the values of the ΔR the values of the mean and RMS for each method are very similar and the differences between them are very small (0.199 for the means and 0.297 for the RMS). Therefore, it is evident that both methods are almost equivalent in terms of the accuracy in identifying the desired b jet as the distributions and values of the means of the ΔR and invariant mass difference for each method are both so similar and the method with the lowest mean and RMS is different for each value.

It would be expected the the number of entries for these histograms would equal the number of events as there is one b quark associated with the hadronic W boson in each event. However, the number of entries is fewer than expected as the program used to determine if a b jet is not always accurate in identifying b-tagged jets and can assume that a jet that is b-tagged is not, meaning that that event is displayed to contain no b jets. One reason for this is that the distance a b quark travels before it decays is not always fixed and thus the secondary vertex for a certain decay maybe so close to the primary vertex that the algorithm is unable to detect that they are separate.

5.4 Comparing the reconstructed Top Mass to the true Top Mass

As the distributions of the ΔR of the b quark are so similar for each method, the combination of the W and b jets for each method of determining the correct b jets are found. The W bosons found through combining the jets with the highest p_T are first combined with the b jets found through the smallest ΔR to the W boson. The invariant mass of this top quark is then found and the difference between this and the invariant mass of the true top quark is taken, as well as the ΔR between them. The W boson is then combined with the b jets found through the method of determining the largest

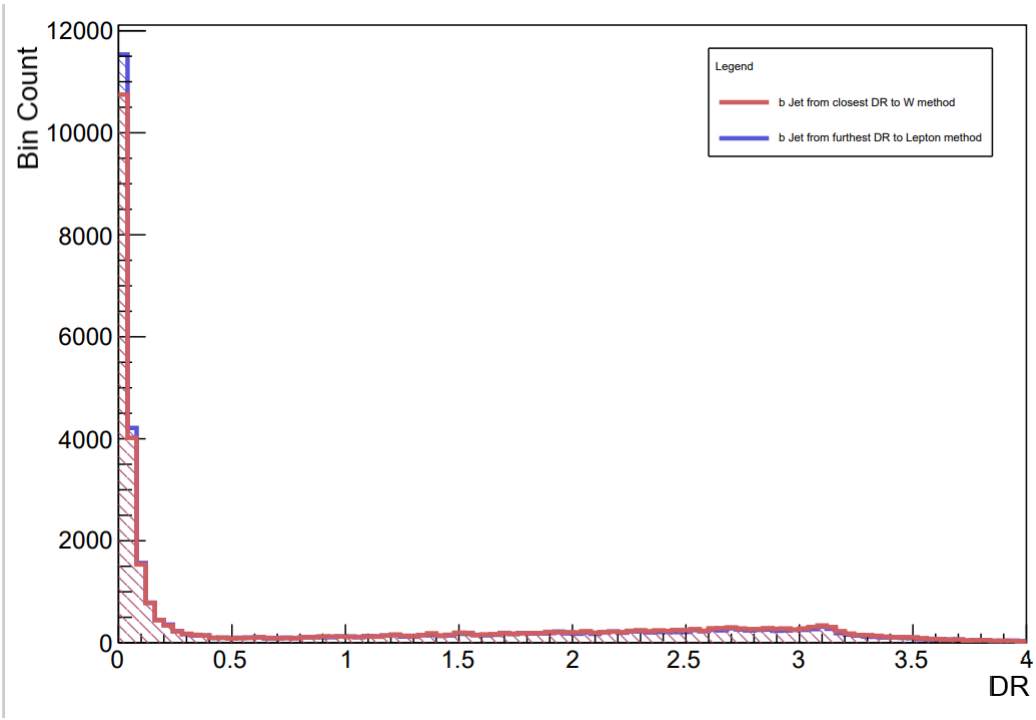

Figure 14: Overlayed distiributions of the ΔR between the true b quark and the reconstructed b jet via two different methods. The method in which the correct b jet is chosen by selecting the b jet that has the largest ΔR to the Lepton is shown in blue and the method in which the correct b jet is chosen by selecting the b jet with the smallest ΔR to the reconstructed W boson is shown in red.

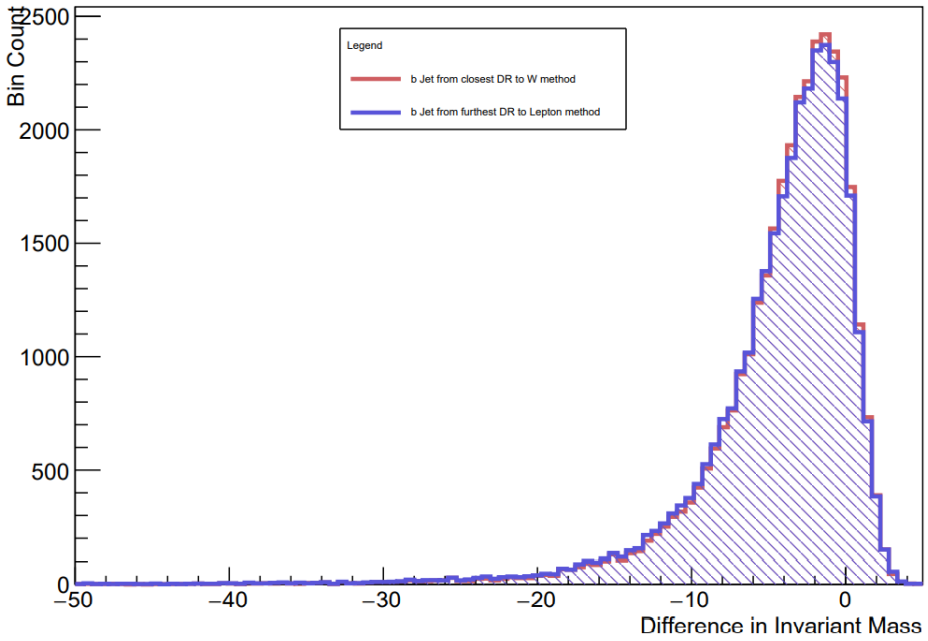

Figure 15: Overlayed distributions of the difference in invariant mass between the selected reconstructed b jet and the true b quark for two different selection methods. The method in which the b jet with the smallest ΔR to the W boson is shown in red and the method in which the b jet with the largest ΔR to the lepton is selected is shown in blue.

ΔR to the lepton and the invariant mass is then calculated and compared to the invariant mass of the true top quark and the ΔR between them is calculated.

The Histograms for the ΔR of the true top quark and the reconstructed top quark for each method can be seen if Figure 16, where the values for the top quark reconstructed using the b jet found through the smallest ΔR to the W boson is shown in red and the values for the top quark reconstructed using the b quark found through the largest ΔR to the lepton is shown in red. Much like the distributions for the ΔR between the b quark and jet, these two distributions are very similar, which again proves that these two methods are fairly accurate methods of determining the correct b quark. However, as can be seen in the plot there is a larger peak at 0 for the b jet method of determining the smallest ΔR to the W boson, implying that this method is slightly more accurate. This is highlighted as the mean value for the W boson comparison method is 1.233 +- 0.005048 with an RMS of 0.9723 +- 0.003569 , whereas the lepton comparison method has a mean of 1.255 +- 0.005084 and an RMS of 0.9791 +- 0.003595. The mean for the W boson comparison method is slightly lower than that of the lepton comparison method, again highlighting that this method is slightly more accurate in determining the right jets to reconstruct a accurate top quark.

Figure 17 shows the distributions for the difference in invariant mass for the true top quark against the reconstructed top quark for each method of selecting the b jet. The method in which the b jet is chosen by selecting the jet with the smallest ΔR to the W boson is shown in red, while the method in which the b jet is selected by finding the jet with the largest ΔR to the lepton is shown in blue. As can be seen in the graph, there is a large peak at approximately 0 GeV, which highlights that the majority of the reconstructed top quarks produced by both methods have invariant masses very similar to the invariant mass of the true top quark. The peak has a very steep decline for differences above 0 GeV, implying that there were very few events that produced a recovered top quark with a smaller invariant mass than the true top quark invariant mass. It can also be seen that there is a fairly large tail for values that are less than 0 GeV, again conveying that the methods produce a high number of reconstructed top quarks with an invariant mass higher than the true invariant mass. This also implies that the methods are not entirely accurate when selecting the correct jets to reconstruct the top quark as the jets have much higher invariant masses than are desired. The method of selecting the b jet via finding the smallest ΔR to the W boson has a slightly larger peak around zero and also has a slightly smaller tail past 0 GeV. This implies that this method is slightly better at identifying the correct jets to combine in order to reconstruct the top quark. This is further shown by the mean difference in the data. The W boson comparison method has a mean invariant mass difference of -111 ± 0.9595 GeV with and RMS of 176.2 ± 0.6785 GeV, whereas the lepton comparison method

has a mean invariant mass difference of -123 ± 0.9894 GeV with an RMS of 181.5 ± 0.6996 GeV. As both the mean and the RMS for the W boson comparison method is larger and have slightly smaller uncertainties than that of the lepton comparison method, it is again conveyed that this method is a more accurate method for determining the jets required to reconstruct the top quark, which is in contrast to the results for the ΔR distribution.

However, for the ΔR distribution, both methods have a range of ΔR of [0, 4], which is much larger than the general ΔR range of [0, 0.4]. This indicates that though the ΔR between the b jet and the b quark as well as the ΔR of the reconstructed and true W bosons are fairly accurate and low this does not necessarily mean that they are the correct jets to combine to create an accurate representation of the top quark. The mean invariant mass differences are quite large for both methods, which is not to be expected from the distribution or the desired results. This is most likely due, however, to the large negative tail after 0 GeV, which extends all the way until -1000 GeV. Though there are only a few events with such high invariant mass differences, they are high enough that the mean value will be increased by them. This again suggests that these methods are not completely accurate for selecting the correct jets to combine for to reconstruct the top quark. This can be rectified by creating a "Best Case Scenario" top quark, in which the conditions placed on the chosen W bosons and b quarks so that it is also required that the ΔR of them be less than 0.4 as that is the general width of a jet. This is discussed further in section 5.5.

5.5 Finding the Best Case Scenario Top Quark

Figure 18 shows the "Best Case Scenario" plot of a reconstructed Top mass, in which the reconstructed W bosons and chosen reconstructed b jets must have a ΔR of less than 0.4 (the width of a jet) to be selected so that a top quark can be reconstructed. The W boson is reconstructed using the method of combining the non b-tagged jets with the highest p_T as discussed in Section 5.2 and is compared to the true W boson, where the ΔR is found and constrained to be less than 0.4. The b jet is chosen by taking the b jet with the smallest ΔR to the true b or $\bar{b}$ quark and constraining the selection of these jets so that the ΔR must be less than 0.4, much like the W boson. As can be seen in the plot, the majority of the events give a Δ R of around 0 to 1.5, which is a much better distribution than the other methods of reconstructing the top quark using no truth information as the range of angles between the true top quark and the reconstructed top quark is much smaller, implying that these jets are much more likely to be the correct jets. However, this distribution of ΔR values still continues beyond 0.4, which suggests that even within this "best case scenario" with each jet constrained to be within the width of a jet from the true particles, not every chosen jet was relevant to the decay event

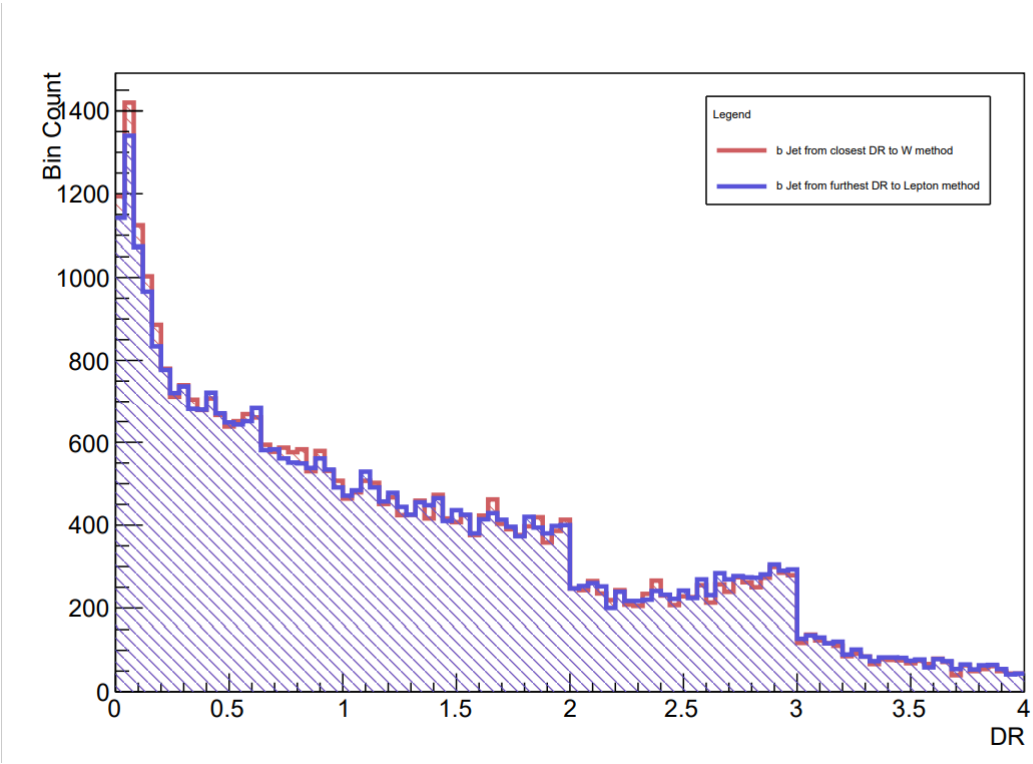

Figure 16: Overlayed distributions of the ΔR between the true top quark and the reconstructed top quark using two different methods of selecting the b jet. The method in which the correct b jet is chosen by selecting the b jet that has the largest ΔR to the Lepton is shown in blue and the method in which the correct b jet is chosen by selecting the b jet with the smallest Delta R to the reconstructed W boson is shown in red.

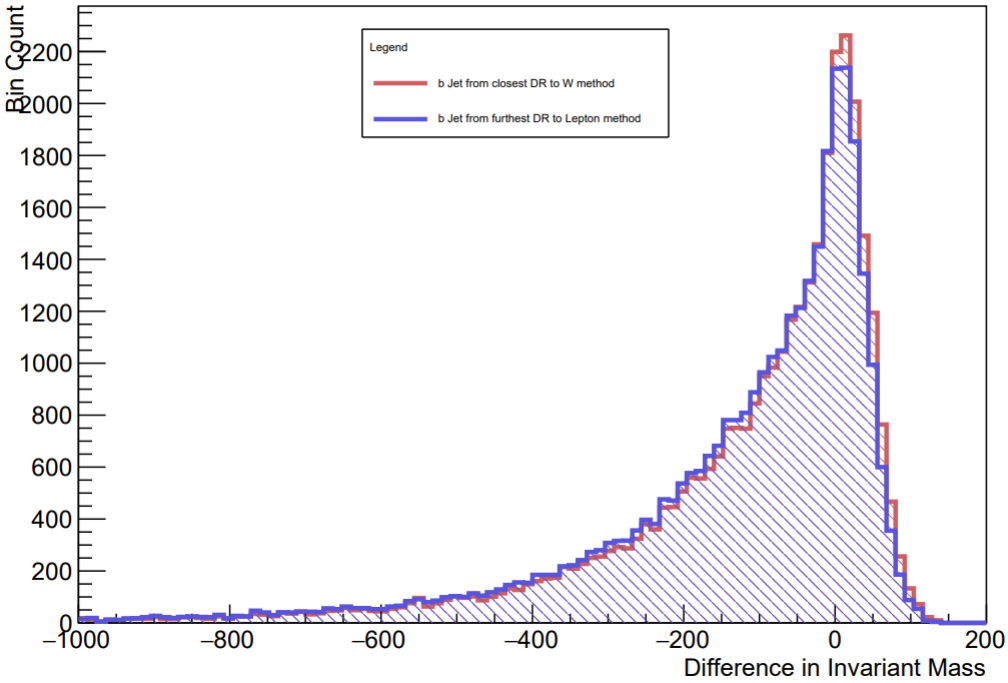

Figure 17: Overlayed distributions of the difference in invariant mass of the true top quark and reconstructed top quark for two different methods of selecting the b jet. The method in which the b jet is chosen by finding the largest ΔR to the reconstructed lepton is shown in blue and the method in which the b jet is chosen by selecting the smallest ΔR to the reconstructed W boson is shown in red.

and could come from either radiation or the decay of the leptonic top quark. The mean of this plot is 0.4594 ± 0.005073 with an RMS of 0.553 ± 0.003587. This mean value is only slightly larger than the desired value of 0.4. This shows that this method is much better at accurately identifying the correct jets in order to reconstruct an accurate top quark. This is to be expected as it uses the information of the true b quark and W boson, meaning that it is not a fully experimental reconstruction as it uses data only available due to the simulation.

Similarly, Figure 19 shows a plot of the difference in invariant mass between the true top quark and the "best case scenario" reconstructed top quark. As can be seen in the plot, there is a large peak around zero, implying that for the majority of events the right jets were chosen to reconstruct a top quark very similar to the true top quark. However, there is a negative tail after 0 that extends until -200 GeV, which implies that some of the events create a top quark with a mass much greater than the true top mass. This again implies that even with the constraints on the jets to ensure that the jets are as closely matched to the true particles as possible, not all of the chosen jets are relevant and do not necessarily match up to the top decay. The mean of this plot is -9.199 ± 0.5088 GeV with an RMS value of 53.81 ± 0.3598 GeV. This is much closer to the desired value of 0 GeV than the previous methods of reconstructing the top quark and has a much smaller RMS value. This highlights that this method of constraining the selection of the jets is much more accurate at getting rid of the jets that do not contribute to the creation of the top quark. However, these numbers are still larger than desire, showing that this method - though it is much better - is still not an entirely accurate means of identifying the correct jets.

6 Conclusions

Within this project the top quark and its mass were discussed as well as the reasons that it is such an important parameter to the Standard Model. Unlike other elementary particles, the mass of the top quark cannot be calculated from the SM and must be measured experimentally. It induces large radiative corrections to the electroweak coupling at low energies and is likely to be the most sensitive probe of the electroweak symmetry breaking. This mass also gives rise to theories that could be used to solve the hierarchy problem of the Standard Model, something that many theories are attempting to resolve.

This project also reconstructed the top quark through the combination of the final state products of the simulated $t\bar{t}$ decay events and compared the resulting top invariant mass to the true invariant mass of the simulated top. The ΔR between this reconstructed top quark and the true top quark was also calculated and analysed. In order to reconstruct the top quark different methods to combine and

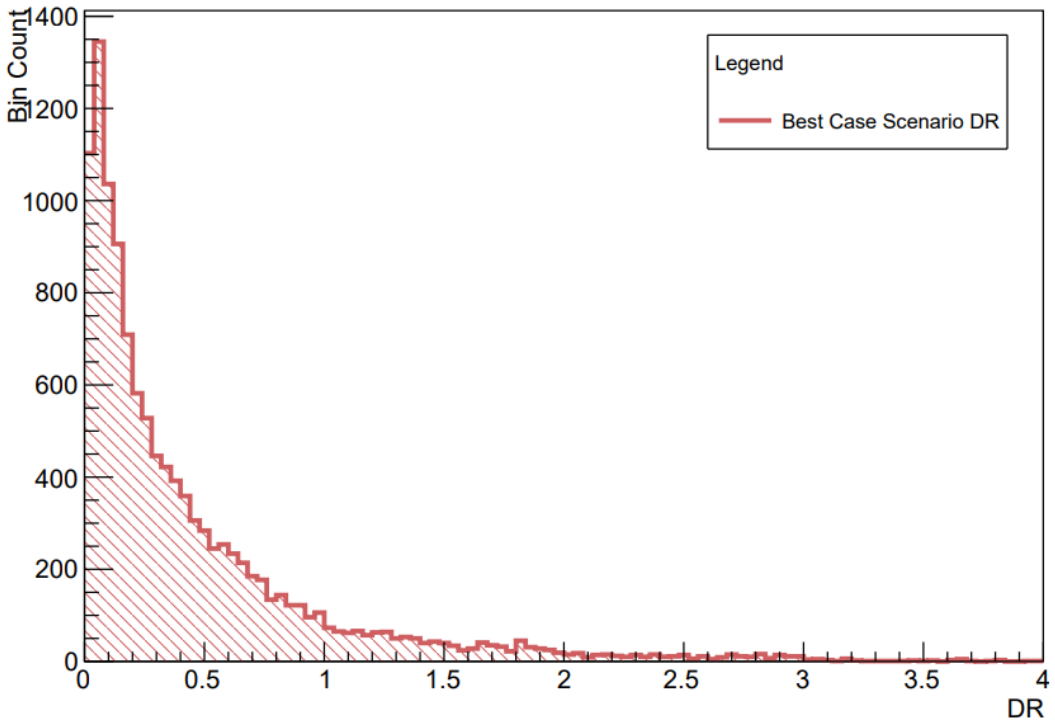

Figure 18: Distribution of the ΔR between the true top quark and the best case scenario reconstructed top quark

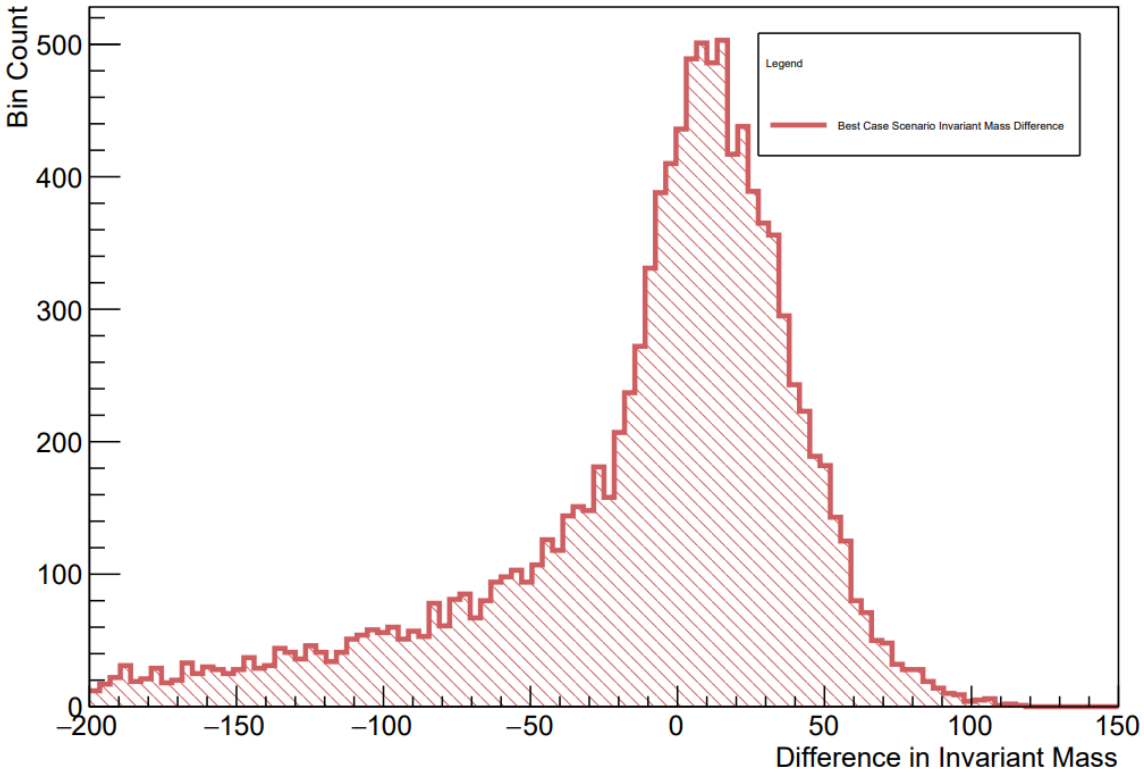

Figure 19: Distribution of the difference in invariant mass of the true top quark and the best case scenario top quark

select the correct jets were used. In order to reconstruct the W boson, the jets were either selected by choosing the two highest p_T jets and combining them or combining each possible combination of jets and comparing which jetpair had the smallest difference in invariant mass to the true W boson, 80 GeV. The latter method only considered the cases where there were two or three non b-tagged jets as past three there were two many possible combinations of jets to consider for the code. This made this method slightly less accurate as the other method was able to consider all events where there were more than one non b-tagged jet recorded. This was shown in the means and distributions of the ΔR as the method of combining the highest p_T jets gave a slightly lower mean with a slightly better distribution. Therefore, it was sensible to continue using this method further on into the project. However, the method of selecting which b jet to use was less clear as both distributions for the different methods were extremely similar and the method with the lower mean value was different depending on whether the ΔR or the invariant mass difference was being considered. This meant that it was not obvious which method to choose so they were both used further on into the project. The reconstruction of the top mass gave fairly good results, with the distributions for the ΔR having a high peak at 0 and a logarithmic decline for higher angles and the distributions for the invariant mass differences peaking at 0 GeV. However, these distributions had many events that gave values much higher and larger means than was expected implying that the methods of reconstructing the top quarks and selecting the jets was not entirely accurate. Therefore, a "best case scenario" was undertaken in which the reconstructed particles were constrained to have a ΔR of less than 0.4 (the width of a jet) with the respective true particles. These distributions were closer to the expected and desired results, showing that these method of constraining the choices using the truth information allowed for better selection of the jets. However, even with these "best case scenario" conditions in place the distributions were not entirely as expected as there were still events that gave values of ΔR and the invariant mass difference that were much higher than the desired values, implying that even using the information from the true particles to select the jets it is difficult to reconstruct the top quarks 100% accurately.

References

[1] Thomson M. Modern particle physics / Mark Thomson.; 2013.

[2] Martin BR, Shaw G. Particle Physics (Third Edition); 2009.

[3] Krause M. CERN : how we found the Higgs boson; 2014.

[4] Hance M. The Large Hadron Collider and the ATLAS Detector. 2013:17-39.

[5] Nie Y, Schmidt R, Chetvertkova V, Rosell-Tarragó G, Burkart F, Wollmann D. Numerical simulations of energy deposition caused by 50 MeV - 50 TeV proton beams in copper and graphite targets. Physical Review Accelerators and Beams. 2017;20(8).

[6] CERN AC. Layout of ATLAS. Dessin reptesentant le detecteur ATLAS. 1998. Available from: https://cds.cern.ch/record/39038.

[7] Kieseler J. In: Introduction to Top Quark Production and Decay in Proton-Proton Collisions. Cham: Springer International Publishing; 2016. p. 7-26. Available from: https://doi.org/10.1007/978-3-319-40005-1_2.

[8] Mann R. Electroweak Symmetry Breaking. In: An Introduction to Particle Physics and the Standard Model. CRC Press; 2010. p. 459-88.

[9] Chung Y. Flavorful composite Higgs model: Connecting the B anomalies with the hierarchy problem. Physical review D. 2021;104(11).

[10] Gfitter. Results for the Global Electroweak Standard Model Fit; 2008. Available from: http://project-gfitter.web.cern.ch/project-gfitter/Standard_Model/.

[11] Aad G, Abbott B, Abbott DC, Abed Abud A, Abeling K, Abhayasinghe DK, et al. Configuration and performance of the ATLAS b-jet triggers in Run 2. vol. 81; 2021.

[12] Lannon K, Margaroli F, Neu C. Measurements of the production, decay and properties of the top quark: A review. European Physical Journal C. 2012;72(8):1-22.

[13] Zhang J. Modern Monte Carlo methods for efficient uncertainty quantification and propagation: A survey. Wiley Interdisciplinary Reviews: Computational Statistics. 2021;13(5):1-42.

[14] Aad G, Abajyan T, Abbott B, Abdallah J, Abdel Khalek S, Abdelalim AA, et al. Performance of jet substructure techniques for large-R jets in proton-proton collisions at s=7 TeV using the ATLAS detector. Journal of High Energy Physics. 2013;2013(9).