UNIVERSITY OF CAPE TOWN

MASTERS DISSERTATION

A search for tWZ production in the trilepton channel using Run 2 data from the ATLAS experiment

Author:
Benjamin WARREN

Supervisor:
Dr. James KEAVENEY

A thesis submitted in fulfillment of the requirements for the degree of Master of Science

in the

UCT-ATLAS Group
Department of Physics

June 17, 2021

Declaration of Authorship

I, Benjamin WARREN, declare that this thesis titled, "A search for tWZ production in the trilepton channel using Run 2 data from the ATLAS experiment" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: Signed by candidate

Date: 12 March 2021

"The more important fundamental laws and facts of physical science have all been discovered, and these are now so firmly established that the possibility of their ever being supplanted in consequence of new discoveries is exceedingly remote...our future discoveries must be looked for in the sixth place of decimals."

Albert Abraham Michelson

UNIVERSITY OF CAPE TOWN

Abstract

Faculty of Science

Department of Physics

Master of Science

A search for tWZ production in the trilepton channel using Run 2 data from the ATLAS experiment

by Benjamin WARREN

This dissertation describes an analysis of events containing three leptons from 136 fb^{-1} of proton-proton collision data, with a centre of mass energy of 13 TeV, recorded by the ATLAS detector between 2016 and 2018. The aim of the analysis was to search for evidence of a top quark produced in association with a W boson and a Z boson (tWZ). An event selection scheme was developed using simulation to broadly suppress background events and to select signal events. Events were separated into mutually-exclusive regions of phase space to increase the ratio of signal to backgrounds and to calibrate the modelling of the backgrounds. Background events were further suppressed through the use of the Gradient Boosted Decision Tree (GBDT) machine learning algorithm. First, a hadronically-decaying W boson candidate was identified using a GBDT; this was used to suppress WZ background events. Then, an event-level GBDT was used to suppress all background events. A maximum likelihood fit was used to estimate the signal strength μ of tWZ production, where nuisance parameters were assigned to theoretical and experimental systematic uncertainties. The work presented here forms the basis of an official ATLAS experiment analysis, thus, the signal region was blinded to avoid potential biases in the future development of the official ATLAS analysis. The best fit value of μ resulting from a fit to a modified Asimov dataset was $\hat{\mu} = 2.08^{+1.48}_{-1.45}$. This corresponds to an expected significance of $Z_{\mu}^{exp} = 0.72\,\sigma$. An expected upper limit $\mu_{up}^{exp} = 2.77^{+2.39}_{-1.28}$ was also determined from the fit to the modified Asimov dataset. Thus, this analysis has the potential to put the strongest ever constraint on tWZ production, but does not have the potential to observe tWZ production as predicted in the Standard Model. These constraints are limited by statistical uncertainties.

Acknowledgements

I would like to thank my supervisor, Dr. James Keaveney, for giving me the opportunity to do this research. I have learnt an incredible amount and I greatly appreciate all of the time and effort you have put into making this research project possible. Thank you to my co-supervisor, Dr. Sahal Yacoob, for your support and many insights on this research project.

Thank you to the University of Cape Town for funding this research though masters research scholarships, and thank you to SA-CERN for funding a trip to CERN when such trips were possible.

I would like to also thank the PhD students in the UCT-ATLAS group, Kevin and Ryan, for offering guidance and support whenever it was asked of them. Thank you to my fellow masters students, Cameron, Jake, and Senzo, for staying in contact during this masters and making it an enjoyable experience from afar.

Thank you to my parents, Colleen and Cliff, and the rest of my family for your love and support on this journey. Thank you to my friends for countless runs, coffees, and hikes throughout the duration of this masters.

Contents

Declaration of Authorship i

Abstract iii

Acknowledgements iv

1 Introduction 1

2 Theoretical Overview 2
2.1 The Standard Model . . . 2
2.1.1 Forces and Gauge bosons . . . 2
2.1.2 Quarks and leptons . . . 3
2.1.3 Cross sections . . . 4
2.2 Motivation for the search for tWZ production in the trilepton channel . . . 4

3 Introduction to Gradient Boosted Decision Trees 7
3.1 Supervised learning . . . 7
3.2 Gradient Tree Boosting . . . 7

4 The LHC and the ATLAS detector 10
4.1 The LHC . . . 10
4.2 The ATLAS detector . . . 12
4.2.1 Coordinate system . . . 13
4.2.2 Detector components . . . 14
4.2.3 Object definitions . . . 17

5 A search for tWZ production in the trilepton channel 19
5.1 Data and simulation . . . 19
5.1.1 Data . . . 19
5.1.2 Simulation . . . 20
5.1.3 Selections and regions . . . 23
5.2 Hadronically-decaying W boson identification . . . 33
5.2.1 Labelling dijets . . . 33
5.2.2 Features . . . 37
5.2.3 $t\bar{t}$ training and event re-weighting . . . 39
5.2.4 Max W_{had} GBDT score . . . 39
5.3 Event-level GBDT . . . 43
5.3.1 Labels and features . . . 43

5.3.2 Re-weighting of signal events 47
5.3.3 Event GBDT Score 47
5.4 Determining an expected upper limit on tWZ production . . . 49
5.4.1 Systematic uncertainties 49
5.4.2 Results 50
5.4.3 Studying the expected significance and expected upper limit at higher luminosity 55
5.5 Summary and Outlook 55

6 Conclusion **57**

Bibliography **58**

Chapter 1

Introduction

The Standard Model (SM) of particle physics describes the current understanding of the fundamental components of matter and three of the four fundamental forces. It has thus far stood up to every experimental test to which it has been subjected, and with the discovery of the Higgs Boson [1, 2] in 2012, the last major piece of the SM has been confirmed. However, even with the success of the SM, it is unable to address some unanswered questions about matter and forces in the universe. Gravity is not included in the SM, there is no explanation for why the Higgs boson mass seems to arise from very precise cancellations of different contributions [3], and there is no explanation for dark matter [4]. The current goal of particle physics is to search for solutions to these problems by testing the predictions of the SM in the hope of finding hints of Beyond the SM (BSM) physics that could provide solutions to these problems.

The Large Hadron Collider (LHC) is capable of colliding particles at high energies that have never been attained in a lab environment, and is capable of producing collisions at rates far higher than previously achieved. Due to its high energy and high collision rate capabilities, the LHC can produce a significant number of top quarks. This is of interest because as the heaviest elementary particle, the top quark serves as a unique probe in searches for BSM physics at high energies [5].

This dissertation presents a search for the rare and unobserved production process of a top quark produced in association with a W boson and a Z boson. This is done with the goal of constraining the electroweak couplings of the top quark since these serve as interesting probes of BSM physics [5]. This dissertation is organised as follows: Chapter 2 presents a brief overview of the SM and the motivation for the search for tWZ production in the trilepton channel, Chapter 3 presents an introduction to the Gradient Boosted Decision Trees as they plays an essential role in Chapter 5 in suppressing background production processes, Chapter 4 presents a brief overview of the LHC and the ATLAS detector, Chapter 5 presents the search for tWZ production and an outlook on the potential for tWZ observation in the future, and Chapter 6 presents the conclusions of this dissertation.

Chapter 2

Theoretical Overview

The first section of this chapter provides an overview of the SM and draws from a comprehensive description in [6] unless stated otherwise. The second section in this chapter describes the motivation for the search for tWZ production in the trilepton channel.

2.1 The Standard Model

Particle physics is concerned with the fundamental components of matter and the forces between them. The current understanding of particle physics is the SM, which is a model that describes all matter as being composed of elementary particles, where the forces between elementary particles are described by the exchange of elementary particles. In the SM, there are three types of elementary particles: quarks, leptons, and mediators. These interact via three of the four fundamental forces: the strong, weak, and electromagnetic forces. Gravity is not included in the SM.

2.1.1 Forces and Gauge bosons

According to the SM, forces are described by a Quantum Field Theory (QFT) which describes forces as occurring via the exchange of spin-1 force-carrying bosons; these are called gauge bosons. Electromagnetism is described by the theory of Quantum Electrodynamics (QED) where the interactions between electromagnetically-charged particles are mediated by the exchange of photons; the strong force is described by the theory of Quantum Chromodynamics (QCD) where the interactions between color charge carrying particles are mediated by the exchange of gluons; and lastly, and interactions due to the weak force are mediated by charged W bosons, or the neutral Z bosons. In Table 2.1, the three fundamental forces included in the SM are shown with the approximate relative strengths of the forces for two fundamental particles at a distance of 1 fm, the bosons that mediate the forces, the spins of these bosons, and the masses of the bosons. The relative strengths of forces shown for the associated gauge bosons are only indicative as the strengths of the forces depending on the distance and energy scale being considered.

Force	Strength	Boson	Spin	Mass (GeV)
Strong	1	Gluon g	1	0
Electromagnetism	10^{-3}	Photon γ	1	0
Weak	10^{-8}	$W^{\pm}$ and Z	1	80.4 and 91.2

TABLE 2.1: This table summarizes the three fundamental forces included in the SM. For the respective forces the relative strengths, bosons, boson spins, and mass of bosons are shown.

In the 1960s, a unified picture of the electromagnetic and weak interactions was developed by Glashow, Salam, and Weinberg (GSW) [7, 8, 9]. This is referred to the GSW electroweak model and interactions involving either the electromagnetic or weak force (forces mediated by the photon, W boson, or Z boson) are referred to as electroweak interactions.

2.1.2 Quarks and leptons

Fermions are half-integer spin particles that obey the Pauli exclusion principle [10]. In the SM, there are twelve fundamental spin-$\frac{1}{2}$ fermions; these being the six quarks and six leptons. The quarks and leptons are broken up into three generations of matter that experience the same forces and differ only in their masses, with each generation getting progressively heavier. The first generation is made up of the up quark (u), down quark (d), electron (e) and electron neutrino (ν_e); the second generation is made up of the charm quark (c), strange quark (s), muon (μ), and muon neutrino (ν_μ); the third generation is made of up of the top quark (t), bottom quark (b), tau (τ), and tau neutrino (ν_τ).

Quarks are the only elementary particles that experience all three of the forces incorporated in the SM. They are the only elementary particles that carry the QCD equivalent of electric charge, called color charge. Furthermore, due to these QCD interactions, quarks are never observed on their own, but rather in colorless bound states called hadrons, where the hadrons can be classified as mesons ($q\bar{q}$), baryons (qqq) or antibaryons ($\bar{q}\bar{q}\bar{q}$). Examples of baryons are protons and neutrons, where the proton is made of two up quarks and a down quark, and the neutron is made of two down quarks and an up quark.

The leptons are split into charged leptons and neutral leptons, where the neutral leptons are referred to as neutrinos. The charged leptons experience the electromagnetic and weak forces, while the neutrinos only experience the weak force. The neutrino masses are sufficiently small that they have yet to be determined and instead have upper limits set.

2.1.3 Cross sections

One way to test predictions of the SM is through measurements of particle interaction cross sections, where the cross section, σ, is an expression of the underlying quantum mechanical probability that an interaction will occur. It is of interest because it takes into account the Feynman diagrams contributing to the production process of the elementary particles, where a Feynman diagram is a representation of transitions between states in QFT. The cross section of a production process provides insight into the fundamental physics involved in the interactions between the elementary particles. Cross sections are measured in units of area and are usually quoted in fb or pb, where f = femto = 10^{-15}, p = pico = 10^{-12}, and b = barn = 10^{-28} m^2. Using the cross section and the integrated luminosity $\mathcal{L}$ (described in section 4.1), one can calculate the expected number of events, N, due to a production process using

$$N = \sigma \mathcal{L}$$

This allows one to test the SM predictions using the number of events recorded by a detector.

2.2 Motivation for the search for tWZ production in the trilepton channel

The top quark is the heaviest known elementary particle. It was first observed in 1995 at the Fermilab Tevatron by the CDF [11] and DØ [12] collaborations. The electroweak couplings of the top quark are sensitive to BSM physics, such as that explored in the framework of SM Effective Field Theory (SMEFT) [5]. Therefore, if measurements of production process cross sections (involving the electroweak couplings of the top quark) deviate from SM expectations, this may imply BSM physics at the LHC. The tWZ production process cross section can be used to constrain the electroweak couplings of the top quark. The only searches for tWZ production thus far have been two masters theses [13, 14] which were not able to observe tWZ production. Therefore, tWZ production remains unobserved.

At Next-to-Leading Order (NLO), tWZ production has Feynman diagrams that overlap with those of a top quark antiquark pair produced in association with a Z boson (referred to as $t\bar{t}Z$ production). There are several methods used to deal with this overlap in order to avoid double counting these Feynman diagrams. In this analysis a diagram removal scheme in which quantum interference effects are neglected (referred to as DR1) was used for tWZ production [15]. This leads to additional uncertainty on the signal modelling.

In order to reduce the number of background events in the search for tWZ production, a trilepton decay channel was chosen. This was chosen because in the case of less than three leptons, hadronic background processes dominate to such an extent as to make it impossible to make a measurement of

tWZ production in these channels. The trilepton and four lepton channels were considered for further study and the trilepton channel was chosen because it offered manageable background production processes and offered reasonable statistical power.

In Figure 2.1, a Feynman diagram of tWZ production in the trilepton channel is shown, where from an initial state of a bottom quark and a gluon, a top quark, W boson, and Z boson are produced. The Z boson is required to decay into a pair of opposite sign same flavour (OSSF) leptons, while the W bosons (both the W boson from the top quark decay and the one produced in association with the top quark) could decay in the leptonic or hadronic channels. The cross section for tWZ production, calculated using the DR1 diagram removal scheme, multiplied by the branching fraction of a leptonically-decaying Z boson is $\sigma_{tWZ} \cdot Br(Z \to l^+ l^-) = 16.06$ fb [15].

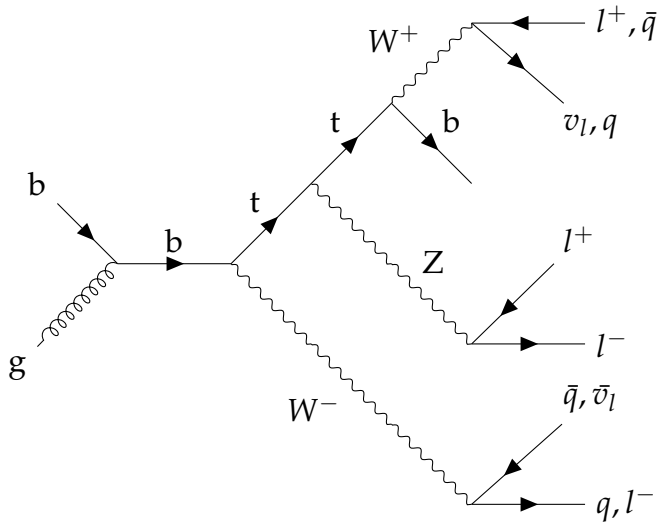


FIGURE 2.1: An example Feynman Diagram is shown for tWZ production in the trilepton channel.

In Figure 2.2, the current SM production cross section measurements for run 2 data at $\sqrt{s} = 13$ TeV from the ATLAS experiment are shown in purple. The cross sections are shown on the y-axis, and different production processes are shown on the x-axis. Production processes with a cross section in the fb range, such as $WZjj$ and $ZZjj$, are shown in the bottom right corner. These production processes have cross sections that are comparable to the cross section of tWZ production and their measurements show that measurements of cross sections in the range of tWZ production are possible with the run 2 ATLAS experiment dataset.

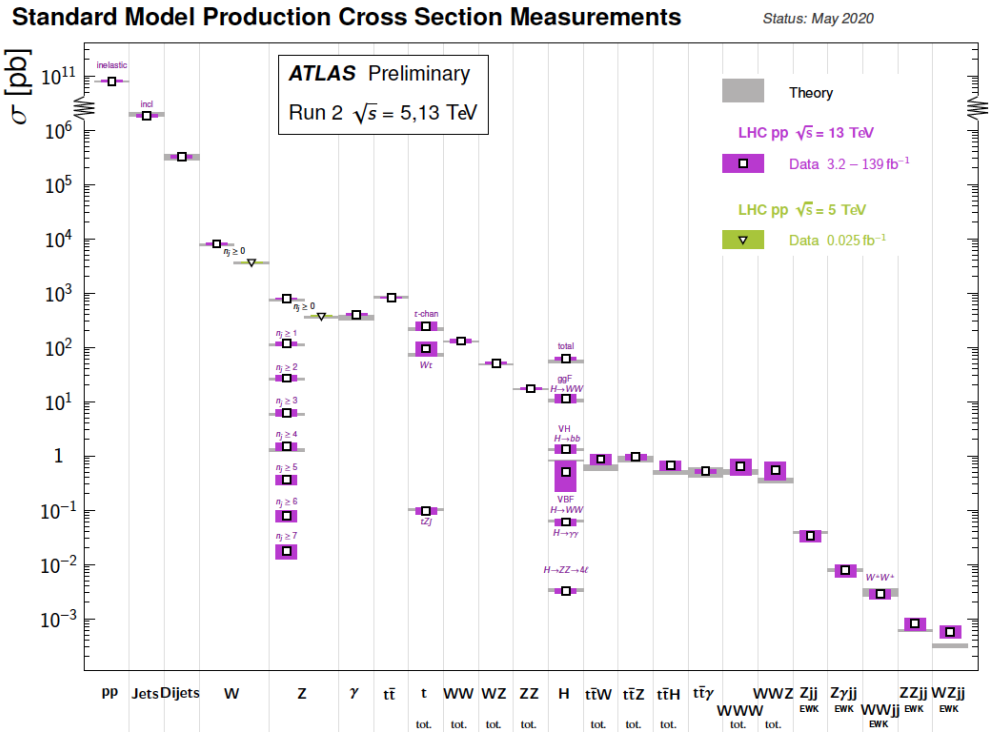


FIGURE 2.2: This figure shows the SM production processes cross sections measured by the ATLAS experiment as of May 2020. The measurements using data with $\sqrt{s} = 13$ TeV are shown in purple. This figure is taken from [16]

Chapter 3

Introduction to Gradient Boosted Decision Trees

The field of machine learning is concerned with the question of how to construct computer programs that automatically improve with experience [17]. A Gradient Boosted Decision Tree (GBDT) is a machine learning algorithm widely used in particle physics research. In Chapter 5, GBDTs are used to construct high level variables for the purpose of suppressing background production processes. Due to the key role GBDTs play in this analysis, they are briefly described in this chapter.

3.1 Supervised learning

A GBDT is a type of supervised learner i.e. a model that results from the training of a machine learning algorithm on labelled examples. It establishes a relationship between features, x_i, and labels, y_i, of an example such that a supervised learner can produce predicted labels, $\hat{y}_i$, given only the features. This relationship is represented by a function that uses the features to predict the labels, $f(x_i\,; \alpha, ...) = \hat{y}_i$, and the training of the supervised learner amounts to the learning of the function parameters such that algorithm is able to accurately predict the labels given only the features.

During each stage of the training, the learner predicts labels – using the function it has at that instance in the training – and compares these to the labels of the example. The performance of the learner at each stage is quantified by an objective function, $\mathcal{L}$, where the value of the objective function is used to update the function parameters at the next instance in the training. The training process is repeated until a condition on the number of training iterations is met, or the value of the objective function does not change for successive iterations of training.

3.2 Gradient Tree Boosting

A decision tree is a type of supervised learning algorithm that is capable of approximating discrete-valued functions [17]. It is a weak supervised learner

i.e. it is a learner that produces inaccurate predictions. Decision trees classify examples by sorting them down a tree structure, from a root node down branches to internal nodes, and finally to the leaf nodes. In Figure 3.1, an example of the structure of a decision tree is shown. The root and internal nodes of the tree correspond to features, and a branch from a node corresponds to a selection on that feature. Branches can only move down the tree i.e. a feature may only have a selection placed on it once in a sequence of nodes in a tree. The leaf nodes are the final nodes in the tree with no branches extending from them downwards. A leaf node corresponds to a prediction of the decision tree given the selections on the features made in the particular sequence leading to the leaf.

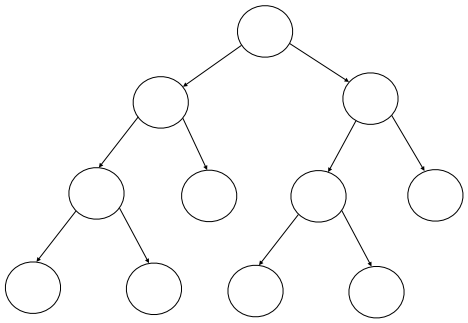


FIGURE 3.1: An example of the structure of a decision tree is shown with the root node at the top that branches down to the internal nodes and ends at the leaf nodes

Boosting is the process of turning a group of weak learners into a single accurate learner [18]. There are many types of boosting methods, but the one that concerns GBDTs is boosting by successively adding weak learners which are each assigned a weight parameter, α. The predictions resulting from this type of boosting are given by

$$\hat{y}_i^t = \sum_{i=j}^{t} \alpha_j \, h_j(\boldsymbol{x_i})$$

where $\hat{y}_i^t$ is the prediction for example i at iteration t, α_j is the weight parameter assigned to a decision tree at iteration j, and $h_j(\boldsymbol{x_i})$ is the value of a decision tree for example i at training iteration j [19].

The training of a GBDT amounts to minimizing the objective function

$$\mathcal{L}^{(t)} = \sum_{i=1}^{N} l(y_i, \hat{y_i}^t) + \Omega(h_t)$$

for a set of N training examples [20]. The objective function is made up of a loss term, l, and a regularization term, Ω. The loss term takes the labels y_i and compares these to the predicted labels at iteration t, $\hat{y}_i^t$. This loss function provides a value from this comparison which reflects the current ability

of the GBDT to accurately predict labels i.e. a low value of the loss function reflects an accurate GBDT and a high value reflects an inaccurate GBDT. The regularization term, $\Omega(h_t)$, is used to penalize the complexity of the generated decision tree at this training iteration.

It is not enough to just improve the objective function using boosting, one also needs to efficiently choose the decision tree to add to the sum of previously predicted labels, such that the training can occur in a reasonable amount of time. To do this, the gradient-based optimization technique is used, where the loss term in the objective function is expanded through to first and second order gradients so that a new decision tree can efficiently be chosen. This allows the algorithm to train in a much shorter amount of time.

The GBDT algorithm is a supervised learner that makes use of gradient-based optimization and boosting [21]. The gradient-based optimization amounts to expanding the loss term in the objective function and the boosting in the case of GBDTs amounts to adaptively combining many decision trees. In Figure 3.2, an illustration of the structure of a GBDT is shown where for an example with features, x, several decision trees give scores of $h_j(x)$ which are each assigned a weight α_j. These are then combined in order to produce a predicted label for the example.

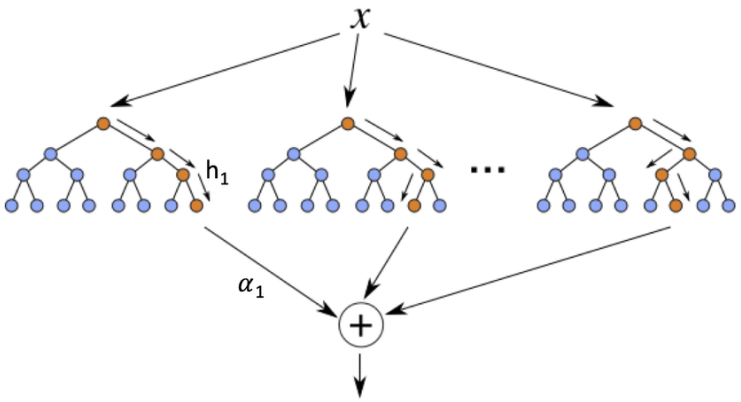


FIGURE 3.2: A figure illustrating the structure of a GBDT is shown. This figure is taken from [19]

There are many software libraries that implement GBDTs, each have the same underlying implementation but have different benefits and limitations when it comes to training models and predicting labels. For this analysis, the Scikit-learn [22] GBDT implementation was used. Scikit-learn is a general purpose python-based library built for predictive data analysis with several popular machine learning algorithm implementations. It was chosen for this analysis because of its ability to predict labels within a parallelized function, which was not possible with other libraries such as XGBoost [20].

Chapter 4

The LHC and the ATLAS detector

This chapter describes the experimental environment in which this analysis is performed. It contains descriptions of the Large Hadron Collider (LHC) and the ATLAS detector. The LHC was built to collide protons and heavy ions at unprecedented high energies and rates. The ATLAS detector was built to measure the byproducts of the proton-proton (pp) collisions produced by the LHC. This chapter will deal with the description of the LHC insofar as it concerns the pp collisions relevant to the ATLAS detector, and it will also deal with aspects of the ATLAS detector relevant to this analysis.

4.1 The LHC

The LHC [23] is a 27km superconducting accelerator built to collide 10^{11} protons 40 million times per second. At its peak, the LHC will collide protons at a centre-of-mass energy of 14 TeV, and it will reach a peak instantaneous luminosity of $10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ for the general purpose experiments – ATLAS and CMS. The LHC has been running since 2008 [24] and since then has had two successful data-taking periods – referred to as Runs. It recorded pp collision data from 2009 until 2013 (referred to as Run 1) and again from 2015 until 2018 (referred to as Run 2).

In Figure 4.1, the European Organization for Nuclear Research (CERN) accelerator complex is shown. The process of colliding protons starts with the production of these protons using a hydrogen source injected into a a duoplasmatron [25]. The protons emerging from the duoplasmatron are then passed to a linear accelerator called Linac2. From Linac2 they are sent through a chain of three accelerator rings of increasing diameter, which are capable of accelerating the protons to higher and higher energies. These three accelerator rings are: the Proton Synchrotron Booster (PSB), the Proton Synchrotron (PS), and the Super Proton Synchrotron (SPS). After the SPS, the protons are injected into the main LHC ring in two separate beams of protons where they circulate in opposite directions around the collider in parallel evacuated tubes.

The two beams of protons are broken up into bunches of protons that are separated in the main collider by 25 ns intervals. In 2018, the maximum number of colliding bunch pairs was 2544 [27]. During run 2, the LHC collided

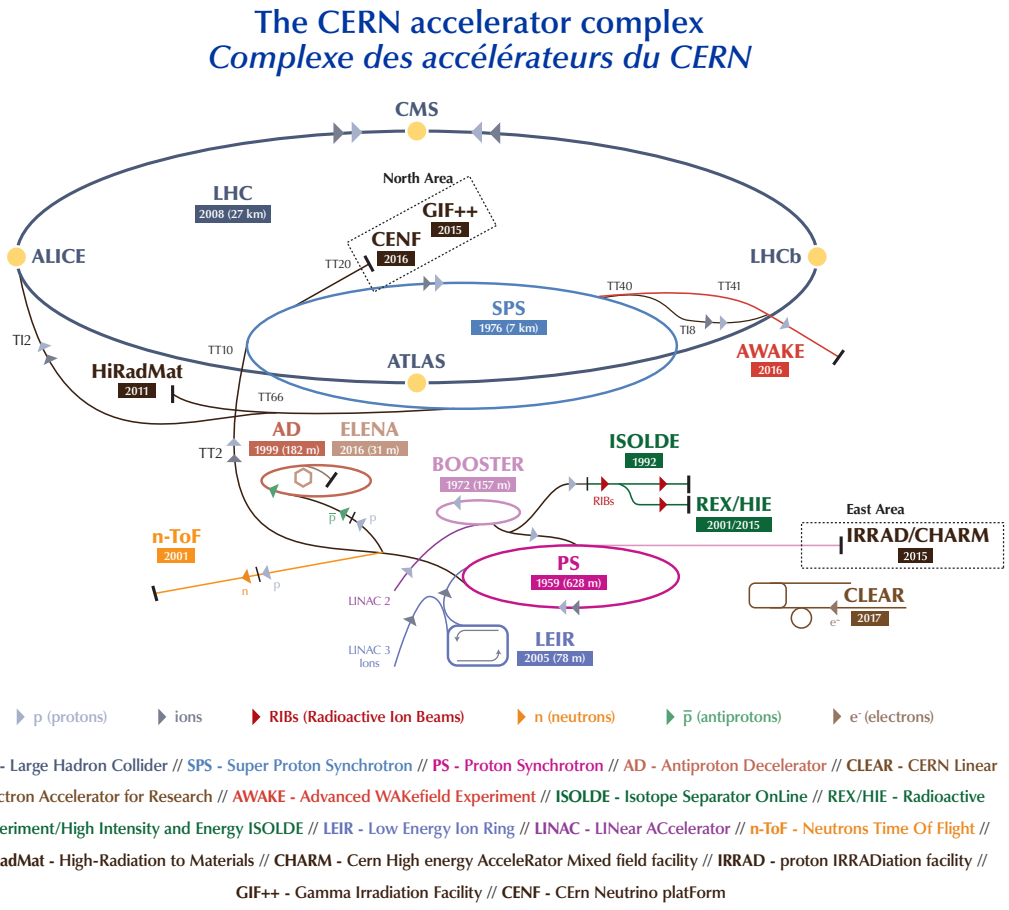


FIGURE 4.1: The CERN accelerator complex is shown with Linac2, the PSB (labelled as Booster), PS, SPS, and the main LHC ring. This image is taken from [26]

the proton bunches at a center-of-mass energy of $\sqrt{s} = 13$ TeV, where the centre-of-mass energy is calculated using the square root of the Mandelstam variable s [6] in the centre-of-mass frame

$$s = (p_1 + p_2)^2$$

where p_1 and p_2 are the four-momentum of protons in each bunch.

Luminosity

The instantaneous luminosity, L, gives the number of pp collisions per unit area per unit time, and is given by the equation:

$$L = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi \epsilon_n \beta^*} F$$

where N_b is the number of protons per bunch, n_b is the number of bunches per ring, f_{rev} is the revolution frequency, γ_r is the relativistic gamma factor, ϵ_n is the normalised transverse beam emittance, β^* the beta-function at the interaction point, and F is the geometric luminosity reduction factor due to the crossing angle at the interaction point [25].

The integrated luminosity, $\mathcal{L}$, is the instantaneous luminosity integrated over a period of time:

$$\mathcal{L} = \int L \, dt$$

The integrated luminosity is a measure for the amount of data recorded over a period of time, and using the integrated luminosity the number of expected events for a given process X, produced from the pp collisions with cross section $\sigma(pp \to X)$, can be calculated. This is done using the equation:

$$N_X = \sigma(pp \to X) \, \mathcal{L}$$

Because of this, the integrated luminosity is given in units of the inverse cross section fb^{-1}.

4.2 The ATLAS detector

The ATLAS detector [28] is one of two general purpose experiments at the LHC designed to collect pp collision data. Given the high luminosity and high energy of the LHC, the goal of the ATLAS detector is to perform high precision tests of Quantum Chromodynamics, electroweak interactions, and flavor physics.

The ATLAS detector is the largest particle physics detector, by volume, ever built. It stands at a height of 25 m, a length of 44m, and weighs approximately 7000 tonnes. In Figure 4.2, a cut-away view of the ATLAS detector is shown with several key sub-components labels. These are: the inner detector, the

calorimeters, the muon spectrometer, and the trigger and data acquisition systems.

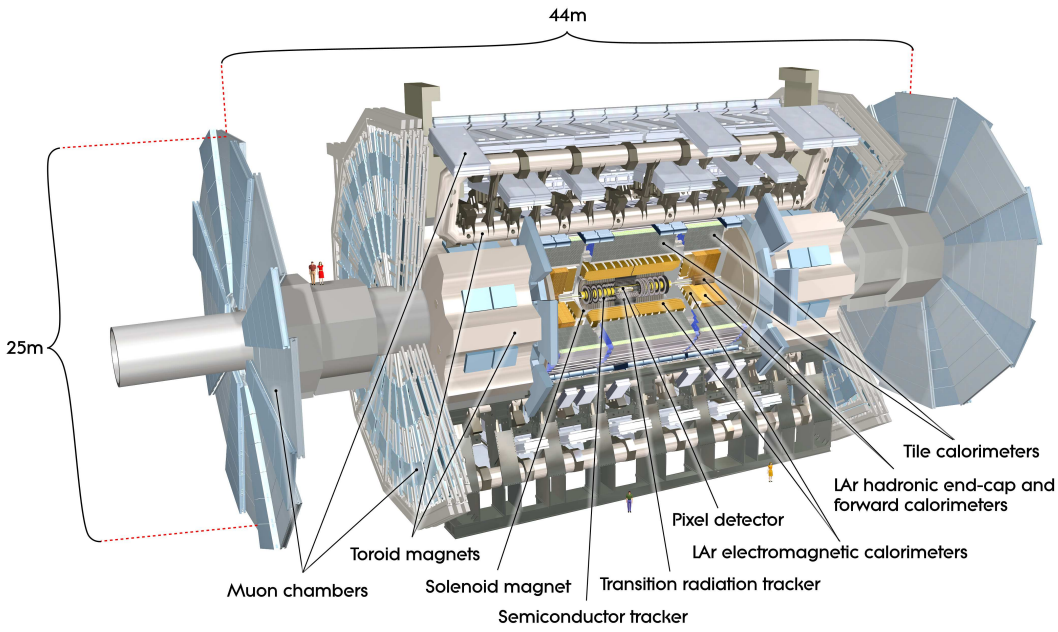


FIGURE 4.2: A cut-away view of the inside of the ATLAS detector taken from [28]

4.2.1 Coordinate system

The coordinate system is used to describe the coverage of the sub-components. For this coordinate system, the origin of the coordinate system is taken as the nominal interaction point. The z-axis points along the beam direction, and the x-y plane is transverse to the beam direction. In the x-y plane, the positive x-axis points from the interaction point to the centre of the LHC ring and the positive y-axis points upwards.

In the x-y plane several key quantities are defined. These are labelled as the transverse components of these quantities since they are the components perpendicular the to the beam direction. These quantities are: the transverse momentum p_T , the transverse energy E_T , and the missing transverse energy E_T^{miss}.

The azimuthal angle ϕ is measured around the z-axis, and the polar angle θ measured from the nominal interaction point on the beam axis. Using θ, the pseudorapidity η is defined as

$$\eta = -\ln(\tan(\theta/2))$$

When dealing with the ATLAS detector and defining the coverage of the sub-components, one works in the $\eta - \phi$ space. In this space, a distance metric, ΔR, is defined as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

4.2.2 Detector components

Inner Detector

The Inner Detector (ID) measures the direction, momentum, and charge of electrically-charged particles produced in each collision. In Figure 4.3, a cut-away view of the three separate but complementary sub-systems of the ID are shown. These being the pixel detectors (pixels), silicon microstrip (SCT) detectors and the transition radiation tracker (TRT).

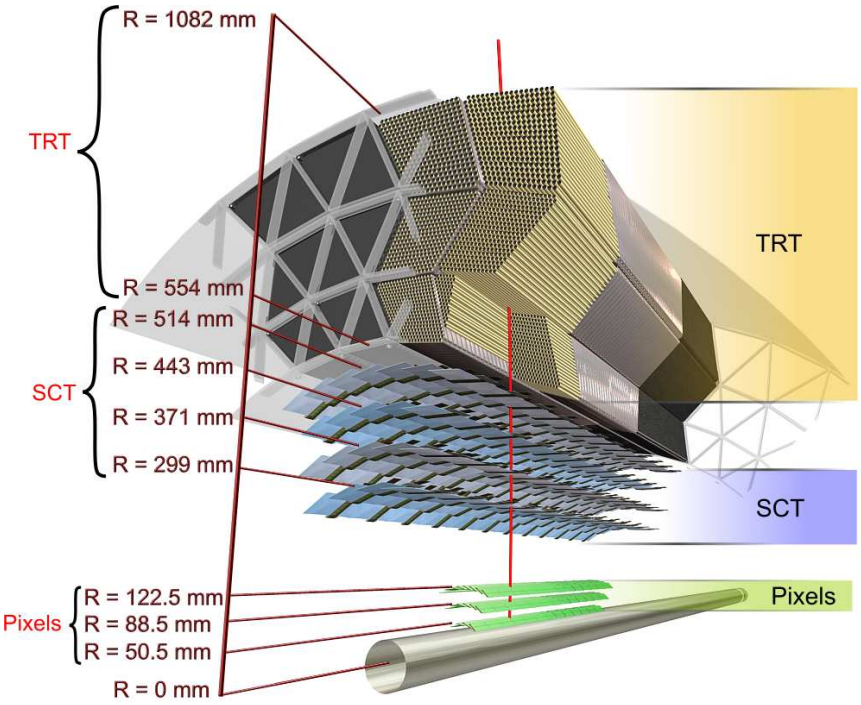


FIGURE 4.3: A cut-away view of the three sub-systems of the inner detector (the pixel, SCT, and TRT) are shown. The image is taken from [28].

All the components of the ID are positioned in a 2T solenoidal magnetic field, which allows one to determine the momentum and sign of charged particles that interact with the ID. The pixels and SCT layers provide high-resolution pattern recognition at the inner part of the ID, while in the outer part of the ID, the TRT layer provides enhanced pattern recognition, improved momentum resolution above $|\eta| < 2.0$, and electron identification complementary to the calorimeters.

Calorimeters

The ATLAS calorimeter system – shown in Figure 4.4 – covers the range $|\eta| < 4.9$ and comprises a high granularity liquid-argon (LAr) electromagnetic calorimeters and hadronic calorimeters. The high granularity LAr electromagnetic calorimeters are used for precision measurements of the energies of electrons and photons, while the hadronic calorimeters are used for jet reconstruction and E^T_{miss} measurements.

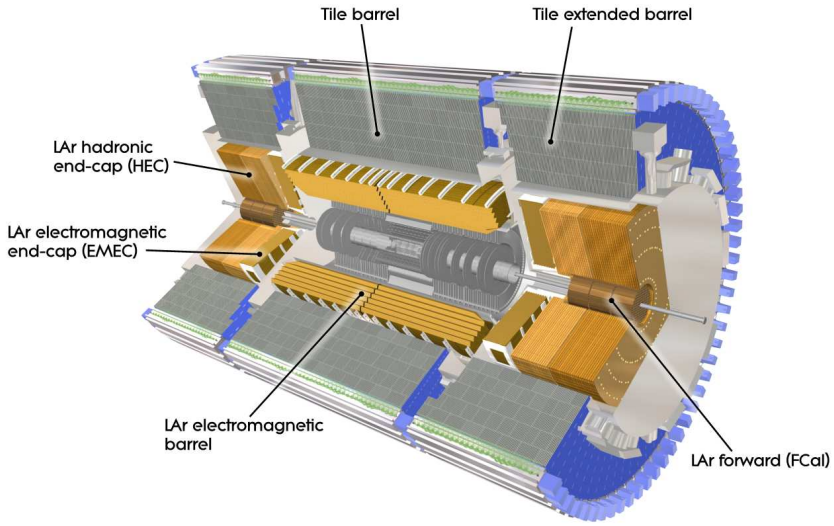


FIGURE 4.4: A cut-away view of the calorimeter system is shown. It comprises high granularity LAr electromagnetic calorimeters (shown as the LAr electromagnetic barrel and EMEC) and hadronic calorimeters (shown as the tile barrel, tile extended barrel, HEC, and FCal). The image is taken from [28].

The LAr electromagnetic calorimeters comprise a barrel component – covering $|\eta| < 1.475$ – and two end-cap (EMEC) components – covering $1.375 < |\eta| < 3.2$. The hadronic calorimeters comprise the tile calorimeter, the LAr hadronic end-cap (HEC) calorimeter, and the LAr forward calorimeter (FCal). The tile calorimeter is situated directly outside the LAr electromagnetic calorimeter and has a barrel that covers $|\eta| < 1.0$ and two extended barrels that cover $0.8 < |\eta| < 1.7$. The HEC calorimeter consists of two separate wheels per end-cap, which sit just outside the EMECs, and extend out to $|\eta| = 3.2$. The FCals are situated in the same cryostats as the EMECs and the HECs, but cover $3.1 < |\eta| < 4.9$.

Muon spectrometer

The muon spectrometer is designed to detect muons and to measure their momenta. It sits outside the other sub-components since muons only minimally interact with the calorimeters. It is capable of measuring the momenta of muons in the range $|\eta| < 2.7$, and triggering on muons in the region $|\eta| < 2.4$. In Figure 4.5, the components of the muon spectrometer are shown. The muon spectrometer uses three toroidal magnets, a system responsible for precision measurement of muon tracks, and a trigger system. The system responsible for precision measurement of muon tracks consists of Monitored Drift Tubes (MDT) and Cathode Strip Chambers (CSC), while the trigger system consists of Resistive Plate Chambers (RPC) and Thin Gap Chambers (TGC).

The magnetic field for the muon spectrometer is generated by the three toroidal magnets, which are used to determine the momentum and sign of the muons.

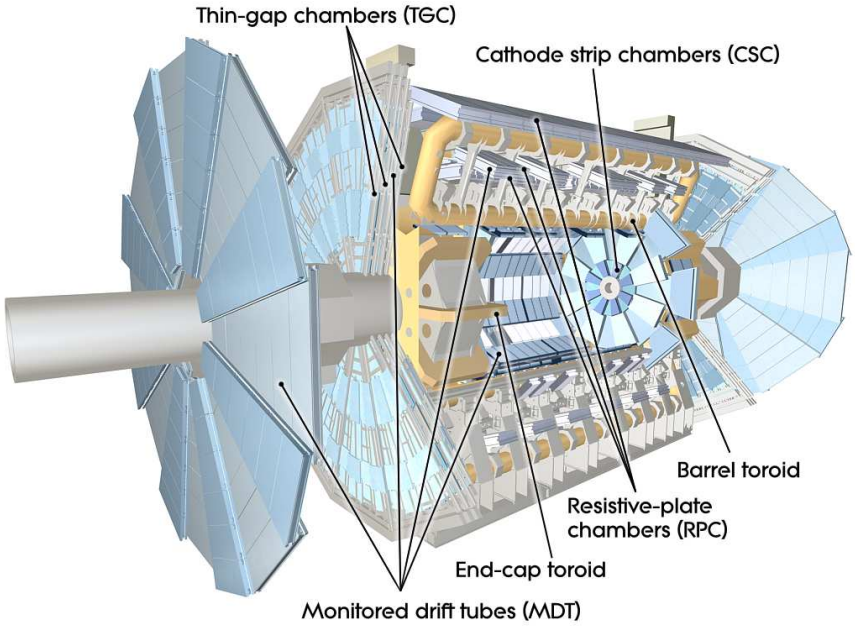


FIGURE 4.5: A cut-away view of the muon spectrometer is shown. It comprises the three toroidal magnets, the systems responsible for precision measurement of muon tracks (MDTs and CSCs), and the trigger system for muons (RPCs and TGCs). The image is taken from [28].

The large barrel toroid provides the magnetic field over the range of $|\eta| < 1.4$, and the end-cap toroids provide the magnetic field over the range of $1.6 < |\eta| < 2.7$. In the transition region ($1.4 < |\eta| < 1.6$), magnetic field comes from by a combination of barrel and end-cap fields.

Precision measurement of the muon tracks is provided by the MDT, which covers $|\eta| < 2.7$, except in the innermost end-cap layer where their coverage is limited to $|\eta| < 2.0$. In the forward region of $2 < |\eta| < 2.7$, the CSC is used for the precision measurements since it is able to deal with the higher rates and challenging background conditions.

The trigger system for the muon spectrometer covers $|\eta| < 2.4$ and provides discrimination on p_T, bunch-crossing identification, fast and coarse tracking information, second coordinate measurement, and robustness towards random hits. It is made of the RPC in the barrel region, and the TGC in the end-cap regions. The RPC covers $|\eta| < 1.05$ and the TGC covers $1.05 < |\eta| < 2.4$.

Trigger and data acquisition system

The recording of event data from the ATLAS detector is limited to $\sim$ 200 Hz based on current technology and resource limits. In order to reduce the amount of data down to this rate, a trigger and data acquisition system (DAQ) is used. The trigger system can be broken up into the Level-1 (L1) trigger and the High Level Trigger (HLT), where the HLT consists of the Level-2 (L2) trigger and the event filter. The DAQ is responsible for the flow of data

between the different triggers and is also responsible for the monitoring of data-taking.

The L1 trigger reduces the rate of data to 75kHz using calorimeter and muon triggers to search for high-p_T muons, electrons, photons, jets, τ-leptons decaying into hadrons, and events with large E_T^{miss}. In order to do this, the L1 trigger uses information from the RPC and TGC for high-p_T muons, and the calorimeter system for electromagnetic clusters, jets, τ-leptons, and E_{miss}^T. The L2 trigger uses Regions of Interest (RoIs), which are regions that the L1 trigger has identified as possible trigger objects within an event. It uses the coordinates, energy, and type of signature information from the RoIs to limit the amount of data transferred from the readout of the detector to the event filter. The event filter is an offline trigger that selects events at a rate which can be recorded for subsequent offline analysis ($\sim$ 200 Hz).

The DAQ receives and buffers the event data from the readout electronics of the detector, and controls the flow of data through the various triggers. In addition to controlling movement of data down the trigger selection chain, the DAQ is also responsible for the configuration, control and monitoring of the ATLAS detector during data-taking.

4.2.3 Object definitions

Final state objects from the the ATLAS detector are reconstructed from the sub-detectors using different procedures. In the search for tWZ production in the trilepton channel, the objects used are leptons (electrons and muons), jets, and b-tagged jets. Their definitions are briefly described below based on a more comprehensive description in [28] unless stated otherwise.

Leptons

Muons are reconstructed using a combination of three strategies. The first is based solely on the muon spectrometer data in the region of $|\eta| \geq 2.7$, the second uses information from a muon spectrometer track and an ID track in the region of $|\eta| \geq 2.5$, and the third uses information from an ID track with a muon spectrometer segment.

For the reconstruction of electrons, a cluster is taken from the electromagnetic calorimeter and a loosely matching track is searched for among all tracks in the ID. From the electromagnetic calorimeter cluster, shower-shape variables are calculated for all electron candidates, and cuts are placed on the shower-shape variables, information from the the chosen track, and the combined reconstruction in order to identify isolated high-p_T electrons. For electron candidates there are three possible cut criteria corresponding to different efficiencies when selecting electrons. These are loose, medium and tight criteria. In this analysis, a tight cut criteria on electrons was used.

Jets

Due to hadronisation, quarks and gluons produced in pp collisions produce a jet of hadrons. These jets of hadrons are referred to jets [6]. To identify jets in the ATLAS detector, jet clustering algorithms are used. In this analysis a sequential clustering algorithm called the anti-k_T algorithm [29] is used with the radius parameter set to $R = 0.4$. Once these jets are created using the anti-k_T algorithm, they are calibrated to account for effects such as pileup, non-compensating calorimeter response, and differences between data and simulation.

B-tagged jets

Jets containing a hadron originating from a bottom quark are identified using a deep-learning neural network, DL1r, that includes discriminating variables constructed by a recurrent neural network (RNN) [30]. This was first used in [31] and the following description is based on the description contained in this reference. The DL1r b-tagging algorithm uses distinctive features of b-hadrons such as the impact parameters of tracks and the displaced vertices reconstructed in the inner detector. The inputs of the DL1r b-tagging algorithm also make use of a RNN to exploit the spatial and kinematic correlations between tracks originating from the same b-hadron. The DL1r b-tagging algorithm has a multidimensional output corresponding to the probabilities for a jet to originate from a bottom quark, a charm quark, or some other quark. Operating points for DL1r are chosen based on a selection on the neural network output distribution and provide specific b-tagging efficiencies. The operating point chosen for this analysis had an efficiency of 77%, and was chosen to provide maximal signal sensitivity.

Chapter 5

A search for tWZ production in the trilepton channel using Run 2 data of the ATLAS experiment

In this chapter, the search for tWZ production in the trilepton channel using 136 fb^{-1} of Run 2 data from the ATLAS experiment is presented. The first section of this chapter deals with the ATLAS data and the simulation of the SM predictions for the ATLAS data. The second section deals with the use of a GBDT – described in Chapter 3 – to identify hadronically-decaying W bosons in order to suppress the WZ background production process. The third section deals with an event-level GBDT that is used to suppress all background production processes. The fourth section deals with a maximum likelihood fit to a modified Asimov dataset in order estimate the experimental sensitivity of the analysis. Lastly, the fifth section of this chapter deals with a summary of the results of this search and presents an outlook for future searches for tWZ production.

5.1 Data and simulation

5.1.1 Data

In this analysis, ATLAS data with $\sqrt{s} = 13$ TeV collected from 2016 to 2018 during Run 2 was used. This dataset will be referred to as the data throughout. The data used in this analysis corresponds to a total integrated luminosity of 136 fb^{-1}, compared to the possible 139 fb^{-1} of data from the full Run 2 dataset. The data from 2015 was not used and is not expected to make a significant contribution to the search for tWZ production as it only accounts for 2% of the full Run 2 dataset [27]. The 136 fb^{-1} of data that was used passed the "Good for Physics" criteria. This means that all data used in this analysis was recorded while the LHC declared stable beam conditions and all components of the ATLAS detector were functioning normally.

5.1.2 Simulation

All SM production processes that significantly contribute to the event samples in this analysis are simulated and compared to the data. The simulation of production processes, and their decays to stable particles, is done using event generators that use Monte Carlo (MC) methods to generate simulated samples of events. These simulated events are then passed to GEANT4 [32] which simulates the response of the ATLAS detector to these stable particles.

MC event samples

An event generator is a tool for simulating events and a parton shower simulator is interfaced with an event generator to account for the successive emission of particles off the initial- and final-state partons of the hard process. In this case, the event generators used were MadGraph5_aMC@NLO [33] (abbreviated to MadGraph5), Sherpa [34], and Powheg [35]. The samples modelled with the Sherpa event generator used the dedicated Sherpa parton shower simulator, while the MadGraph5 and Powheg event generators were interfaced with the Pythia8 [36] parton shower simulator. The details of the simulated samples used in this analysis are described below for the respective production processes. The event generators and parton shower simulators used for each sample are summarized in Table 5.1

There is a non-negligible background expected from events containing a fake lepton. A fake lepton is defined as a selected lepton candidate that originated from either the mis-identification of a hadron or from the semi-leptonic decay of a b or c hadron within a jet. A data-driven estimation of this background was beyond the scope of this analysis. Instead this background was estimated using the $t\bar{t}$ sample, where all selected events must contain a fake lepton.

- ***tWZ* production**

  The production of a top quark in association with a W boson and a Z boson (tWZ) is modelled by one sample, which was generated using the MadGraph5 event generator at NLO in QCD, interfaced with the Pythia8 parton shower simulator. The sample was generated with the DR1 diagram removal scheme to suppress the overlap with $t\bar{t}Z$ production [15]. The sample was also required to have a leptonically-decaying Z boson.

- ***ttZ* production**

  The production of top quark antiquark pair produced in association with a Z boson ($t\bar{t}Z$) is modelled by three samples that correspond to different leptonic decays of the Z Boson ($Z \to ee$, $Z \to \mu\mu$, and $Z \to \tau\tau$). These samples were generated using the MadGraph5 event generator at NLO in QCD, interfaced with the Pythia8 parton shower simulator.

- ***WZ* production**

The production of a W boson produced in association with a Z boson (WZ) is modelled by two samples corresponding to the QCD and electroweak production mechanisms. The WZ production process is grouped separately to other diboson production processes since it is the process that contributes significantly in the trilepton channel. Both samples were generated using the Sherpa event generator, and interfaced with its dedicated parton shower simulator.

- **tZq production**

  The production of a top quark in association with a Z boson and a quark (tZq) is modelled by one sample and was generated using the MadGraph5 event generator at NLO in QCD, interfaced with the Pythia8 parton shower simulator. In the trilepton channel, the W boson from the top quark and the Z boson decay leptonically.

  Due to the complexities of simulating this sample the four flavour scheme calculation (corresponding to a cross section 28.94 fb) was used to produce the sample. However, the five flavour scheme calculation (corresponding to a cross section of 101.7 fb) was used to normalize the expected number of events.

- **Triboson production**

  The production of triboson events is modelled by four samples which correspond to the following boson combinations: WWW, WWZ, WZZ, and ZZZ. The samples were all generated using the Sherpa event generator and were interfaced with its dedicated parton shower simulator.

- **Diboson production**

  The production of diboson events is modelled by three samples: an electroweak vector boson fusion (VBF_{EWK}) sample where the bosons decay to four leptons, and two ZZ samples, $ZZgg$ and $ZZqq$, which are produced by two different initial states ($gg \rightarrow ZZ$ and $qq \rightarrow ZZ$) leading to ZZ. The samples were all generated using the Sherpa event generator, interfaced with its dedicated parton shower simulator.

- **$t\bar{t}$ production**

  The production of a top quark antiquark pair ($t\bar{t}$) is modelled by one sample which was generated using the Powheg event generator at NLO in QCD, interfaced with the Pythia8 parton shower simulator. In this sample, both W bosons from the decay of the two top quarks are required to decay leptonically.

- **$t\bar{t}H$ production**

  The production of a top quark antiquark pair produced in association with a Higgs (H) boson ($t\bar{t}H$) is modelled by one sample, which was generated using the Powheg event generator at NLO in QCD, interfaced with the Pythia8 parton shower simulator. In this sample, it is

required that just one of the W bosons, from the top quarks, decays leptonically.

- ***$t\bar{t}W$* production**

  The production of a top quark antiquark pair produced in association with a W boson ($t\bar{t}W$) is modelled by one sample which was generated using the Sherpa event generator at NLO in QCD, interfaced with its dedicated parton shower simulator.

Production process	Sample(s)	Event generator	Parton shower
tWZ	tWZ	MadGraph5	Pythia8
WZ	WZ	Sherpa	Sherpa
	WZ_{EWK}	Sherpa	Sherpa
$t\bar{t}Z$	$t\bar{t}Zee$	MadGraph5	Pythia8
	$t\bar{t}Z\mu\mu$	MadGraph5	Pythia8
	$t\bar{t}Z\tau\tau$	MadGraph5	Pythia8
tZq	tZq	MadGraph5	Pythia8
Diboson	VBF_{EWK}	Sherpa	Sherpa
	$ZZgg$	Sherpa	Sherpa
	$ZZqq$	Sherpa	Sherpa
Triboson	WWW	Sherpa	Sherpa
	WWZ	Sherpa	Sherpa
	WZZ	Sherpa	Sherpa
	ZZZ	Sherpa	Sherpa
$t\bar{t}$	$t\bar{t}$	Powheg	Pythia8
$t\bar{t}H$	$t\bar{t}H$	Powheg	Pythia8
$t\bar{t}W$	$t\bar{t}W$	Sherpa	Sherpa

TABLE 5.1: This table summarizes all of the simulated samples used, as well as the event generators and parton shower simulators used for each.

Event weights and calibrations

Event weights are applied to the events in the samples so that histograms of any distribution constructed from the samples match the expected distributions in data, after the simulation of the response of the ATLAS detector. The event weights must also be adjusted to take into account the different detector conditions during different data taking periods, as well as the known differences between simulation and data. These adjustments come in the form of calibration weights that are applied to the event weights. The calibrations applied are those associated with the pile-up, global lepton triggers, lepton scale factors, jet vertex tagger, and b-tagging algorithm.

Large tWZ event weights

Before calibrations were applied, a tiny fraction of events in the tWZ sample were found to have erroneously large event weights. These large weights could not be motivated on physical grounds and are believed to arise from a technical error in sample merging [37]. Therefore, the events in the tWZ sample with these large weights were removed.

5.1.3 Selections and regions

All events were required to satisfy trigger, object, and event selections which aim to select tWZ events with a high efficiency while largely rejecting background events. For all events satisfying object and event selections, further selections were applied in order to select regions dominated by either tWZ or background events. A region that is dominated by tWZ events is referred to as a signal region (SR), and a region dominated by one of the backgrounds is referred to as a control region (CR).

High level trigger selections

An event is considered for further study if the event passes a single-electron High Level Trigger (HLT) or a single-muon HLT. The single-electron HLTs select a calorimeter cluster matched to a track. The selected electrons must satisfy identification criteria – based on a multivariate technique using a likelihood (lh) discriminant – and an E_T selection. The single-muon HLTs select muons by matching tracks reconstructed in the muon spectrometer and the ID. The selected muons must satisfy identification criteria and a p_T selection.

Data collected in 2016, 2017, and 2018, and the corresponding simulation, had to satisfy one of three single-electron HLTs or one of two single-muon HLTs. The single-electron triggers correspond to a tight identification criteria with an $E_T > 26$ GeV, a medium identification criteria with an $E_T > 60$ GeV, and a loose identification criteria with an $E_T > 140$ GeV. The single-muon triggers correspond to a medium identification criteria with a $p_T > 26$ GeV selection, and no identification criteria but a $p_T > 50$ GeV selection.

Object and event selections

The object selections described below are summarized in Table 5.2. Selections were applied to the p_T of the leading, next-to-leading (NL), and next-to-next-to-leading (NNL) leptons. The leading lepton refers to the lepton with the highest p_T value, the NL lepton refers to the lepton with the second highest p_T value, and the NNL lepton refers to the lepton with the third highest p_T value. The leading, NL, and NNL leptons were required to have p_T values greater than 30, 20, and 10 GeV respectively. Selections were also applied to the η values of the leptons. Each muon (μ) was required to have an $|\eta_\mu| \leq 2.5$, and each electron (e) was required to have an $|\eta_e| \leq 2.47$ – excluding the region between $1.37 \leq |\eta_e| \leq 1.52$ since this is a transition region between the barrel and the end-caps with a relatively large amount of inactive material.

All jets were required to have a $p_T > 30$ GeV, an $|\eta_{jet}| < 2.5$, and a jet vertex tagger (jvt) score greater than 0.5. The jvt above 0.5 was required to reject jets originating from pileup interactions. The jvt requirement rejects jets originating from pileup interactions by using information about the jet shapes and topological jet correlations. An operating point for the $DL1r$ b-tagging algorithm with an efficiency of 77% was chosen.

Object selections
Lepton selections
Leading lepton $p_T > 30$ GeV
NL lepton $p_T > 20$ GeV
NNL lepton $p_T > 10$ GeV
$\|\eta_\mu\| \leq 2.5$
$\|\eta_e\| \leq 2.47$ excluding $1.37 \leq \|\eta_e\| \leq 1.52$
Jet selections
all jet $p_T > 30$ GeV
$\|\eta_{jet}\| < 2.5$
$jvt > 0.5$
DL1r eff = 77%

TABLE 5.2: A summary of all object selections, with the lepton selections is shown in the top half of the table, and the jet selections are shown in the bottom half of the table.

The event selections described below are summarized in Table 5.3. Exactly three leptons were required in every event, where the three leptons could be any combination of electrons and muons. To suppress events containing Υ and J/Ψ particles, which are not included in the simulation, every OSSF dilepton pair in an event was required to have an invariant mass greater than 10 GeV – this being greater than the mass of these two particles. Exactly one Z candidate per event was required, where the Z candidates are defined as OSSF dilepton pairs that were within a 10 GeV range of the Z boson mass. Furthermore, at least two jets per event, and at least one b-tagged jet per event were also required.

Regions

The SR and CR selections described below are summarized at the bottom of Table 5.3. The regions were defined according to selections on the total number of jets and the number of b-tagged jets contained within the total number of jets. The combination of the total number of jets and the number of b-tagged jets contained withing the total number of jets is referred to as the jet multiplicity. The tWZ SR was chosen as the region that gave the the highest significance score, $(s/\sqrt{N})$, where s is the number of expected signal events in simulation that satisfy the SR selections, and N is the total number of expected events for simulated signal and background production processes satisfying the SR selections. This metric for significance was used

Event selections		
event selections		
Exactly 3 leptons		
$\geq$ 2 jets		
$\geq$ 1 b-tagged jet		
all $m_{OSSF_{ll}} > 10$ GeV		
1 Z Candidate ($\lvert m_{OSSF_{ll}} - m_Z \rvert < 10$ GeV)		
Regions		
tWZ SR	WZ CR	$t\bar{t}Z$ CR
$\geq$3 jets	2 jets	$\geq$ 2 jets
1 b-tagged jet	1 b-tagged jet	$\geq$ 2 b-tagged jets

TABLE 5.3: A summary of all event selections and the region selections is shown. The event selections are shown in the top half of the table, and the region selections are shown in the bottom half of the table.

to select a SR that was the most sensitive to the observation of tWZ production.

The CRs are chosen to be dominant in the largest background production processes, WZ and $t\bar{t}Z$. The WZ CR is the region that gave the the highest significance score ($b_{WZ}/\sqrt{N}$), where b_{WZ} is the number of WZ events in simulation that satisfy the WZ CR selections, and N is the total number of events for signal and background production processes satisfying the WZ CR selections. The same metric was used for choosing the $t\bar{t}Z$ CR, but instead of the expected number of WZ events, the expected number $t\bar{t}Z$ events, $b_{t\bar{t}Z}$, was used as the numerator, and the square root of total number of events for signal and background production processes satisfying the $t\bar{t}Z$ CR selections was taken as the denominator. The CRs were chosen to constrain these backgrounds in the fit – described in section 5.4 – and to probe the modelling of the background production processes in simulation through comparison of simulation to data.

In Figure 5.1, two-dimensional plots of the jet multiplicities for the SR and CRs are shown with the significance scores shown by a colour axis on the right of each plot. The jet multiplicity chosen for the SR was one with greater than or equal to three jets, and exactly one b-tagged jet ($\geq 3j\,1b$). A jet multiplicity of exactly two jets, and exactly one b-tagged jet ($2j\,1b$) was chosen as the CR for WZ production, and a jet multiplicity of greater than or equal to two jets and greater than or equal to two b-tagged jets ($\geq 2j\ \geq 2b$) was chosen as the CR for $t\bar{t}Z$ production.

In Table 5.4, the simulated number of events for each sample is shown for

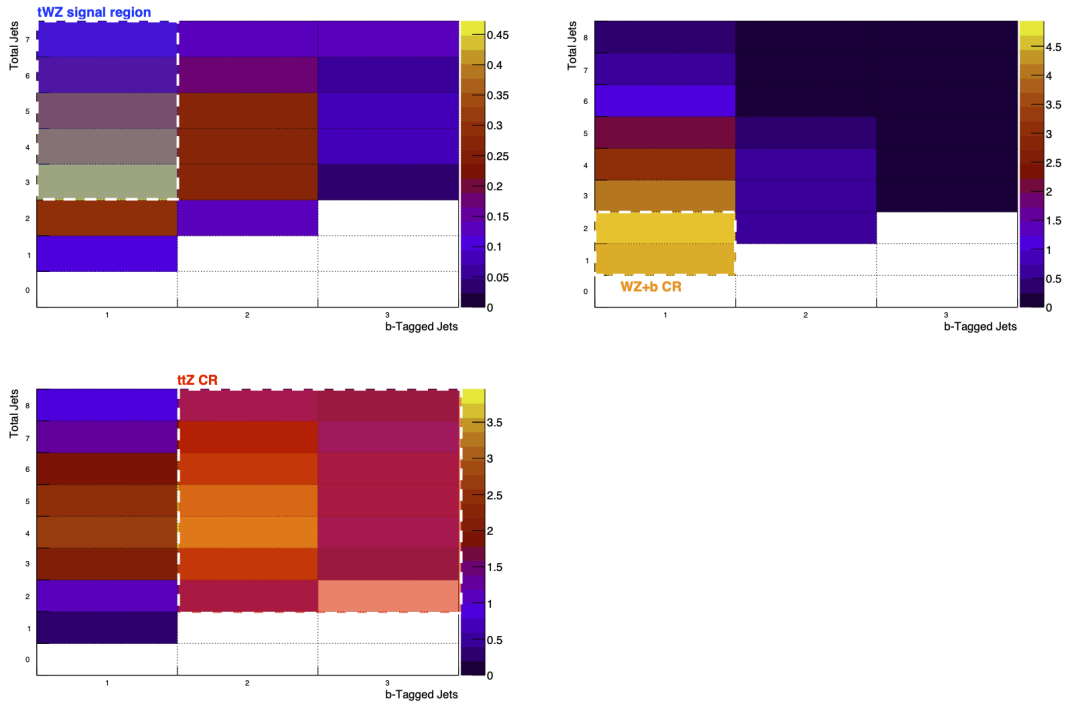


FIGURE 5.1: Two-dimensional plots of the jet multiplicities for the SR and CRs are shown with the significance scores shown by a colour axis on the right of each plot.

each region. The total number of events for simulation is shown for each region. In the CRs, the number of events in data is also shown. The work presented here forms the basis of an official ATLAS experiment analysis. Therefore, the SR remains blinded to data to avoid potential biases in the future development of the official ATLAS analysis. According to the total number of events for simulation and data, the simulation in the CRs seems to model the data well within uncertainty. Furthermore, the production processes that these regions are chosen for are in fact produced in their greatest quantities in these regions.

In Figure 5.2, the number of events for each production process and the total number of events for simulation and data are shown by a histogram with each bin corresponding to a different region. Simulation is shown with different production processes stacked on top of one another, with the non-dominant background production processes grouped together into a class labelled as "other". At the bottom of the histogram the ratio of the total number of data events divided by the total number of events in simulation is shown. This is not shown in the SR because the SR is blind to data. In this figure, the $t\bar{t}Z$ CR is shown with an abundance of $t\bar{t}Z$ events with minimal WZ, tWZ, and other background events, while an abundance of tZq events is observed. In the WZ CR, an abundance of WZ events is observed with minimal $t\bar{t}Z$, tWZ, and other background events, while an abundance of tZq events is observed. From this histogram, one can see that tZq production is abundant in both regions but is particularly prevalent in the WZ CR. This

histogram also shows that simulation and data agree in the $t\bar{t}Z$ and WZ CRs – as can be seen by the ratio of Data/Pred. around 1. This shows that the simulation successfully models the data.

Process	Sample(s)	tWZ SR	$t\bar{t}Z$ CR	WZ CR
tWZ	tWZ	27.57 ± 0.91	14.54 ± 0.51	13.07 ± 0.47
WZ	WZ	130.56 ± 26.44	24.24 ± 4.95	153.23 ± 31.06
	WZ_{EWK}	6.20 ± 1.28	0.65 ± 0.16	5.72 ± 1.19
$t\bar{t}Z$	$t\bar{t}Zee$	75.50 ± 15.29	83.67 ± 16.94	18.54 ± 3.78
	$t\bar{t}Z\mu\mu$	100.11 ± 20.27	113.38 ± 22.95	24.02 ± 4.88
	$t\bar{t}Z\tau\tau$	0.63 ± 0.16	0.89 ± 0.22	0.33 ± 0.09
tZq	tZq	140.79 ± 28.59	160.07 ± 32.44	268.09 ± 54.29
Diboson	VBF_{EWK}	1.15 ± 0.07	0.09 ± 0.02	0.75 ± 0.06
	$ZZgg$	0.86 ± 0.18	0.15 ± 0.03	1.40 ± 0.29
	$ZZqq$	20.84 ± 4.22	5.02 ± 1.03	28.34 ± 5.76
Tribsoson	WWW	$6.00 \pm 2.45 \times 10^{-6}$	$6.00 \pm 2.45 \times 10^{-6}$	0.01 ± 0.03
	WWZ	0.12 ± 0.02	0.01 ± 0.01	0.09 ± 0.02
	WZZ	0.05 ± 0.01	$0.02 \pm 0.01 \times 10^{-1}$	0.04 ± 0.01
	ZZZ	$0.02 \pm 0.01 \times 10^{-1}$	$0.01 \pm 0.01 \times 10^{-1}$	$0.04 \pm 0.01 \times 10^{-1}$
$t\bar{t}$	$t\bar{t}$	37.05 ± 7.74	16.77 ± 3.62	67.37 ± 13.88
$t\bar{t}H$	$t\bar{t}H$	2.71 ± 0.55	4.01 ± 0.81	1.23 ± 0.25
$t\bar{t}W$	$t\bar{t}W$	3.37 ± 0.27	9.04 ± 0.40	4.58 ± 0.23
	Total	547.51 ± 56.05	432.53 ± 52.97	586.79 ± 67.14
	Data	——	411	623

TABLE 5.4: The number of expected events for each simulated sample is shown in the SR and CRs. The total number of predicted events in simulation is shown in each region at the bottom of the table alongside the number of events in data in the CRs.

The CRs are also used to check the modeling of the jet and lepton kinematics – η, ϕ, and p_T – and certain event level variables – the number of jets, and the number of b-tagged jets. These variables are used to construct higher level variables that are then used in the search for tWZ production. Therefore, it is essential to check that the data is well modelled by simulation before using the kinematics to construct these high level variables.

In Figure 5.3, the kinematics of the leading, NL, and NNL leptons are shown for simulation and data in the $t\bar{t}Z$ CR. In Figure 5.4, the same is shown in the WZ CR. There is generally good agreement between simulation and data for all the object kinematics as shown by the Data/Pred. ratio in each plot. One notable exception is the NNL lepton η in the WZ CR shown in the bottom left of Figure 5.4 where there is tension between data and simulation at low values of η. This is not expected to have a significant effect on the results

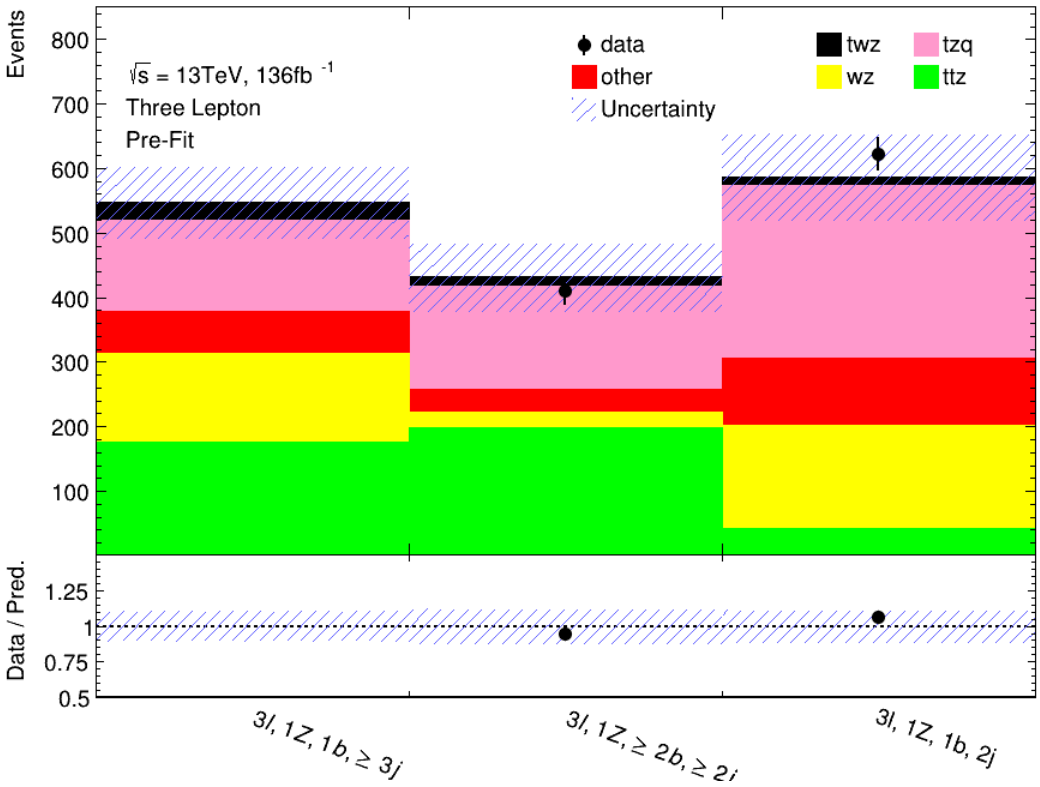


FIGURE 5.2: The number of events in the SR and CRs are shown for simulation and data in the upper panel. Simulation is shown with different production processes stacked on top of one another in different colours, and data in the CRs is shown by the black points. The agreement between simulation and data is shown by the Data/Pred. ratio in the lower panel.

of this analysis due to the small number of events in this bin. However, this could be evidence that the estimation of backgrounds from fake leptons entirely from simulation is an inadequate approach.

In Figure 5.5, the kinematics of the leading and NL jets are shown for simulation and data in the $t\bar{t}Z$ CR. The number of jets, and the number of b-tagged jets are also shown for simulation and data in the $t\bar{t}Z$ CR. In Figure 5.6 the same variables are shown in the WZ CR. Generally good agreement is seen between simulation and data in both CRs for these variables.

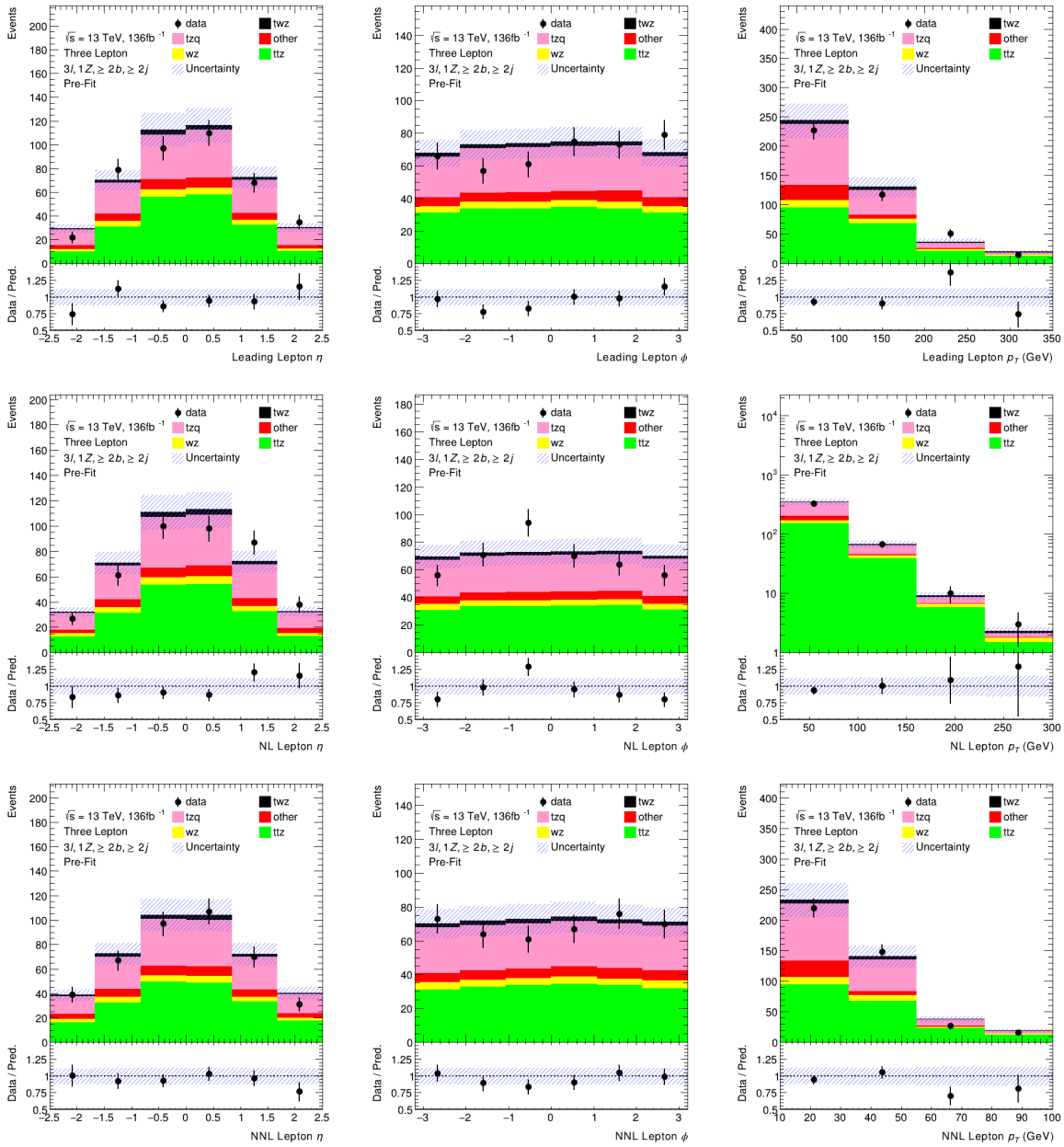


FIGURE 5.3: A comparison of simulation and data for the object kinematics – η, ϕ, and p_T – is shown for the leading, NL, and NNL leptons in the $t\bar{t}Z$ CR

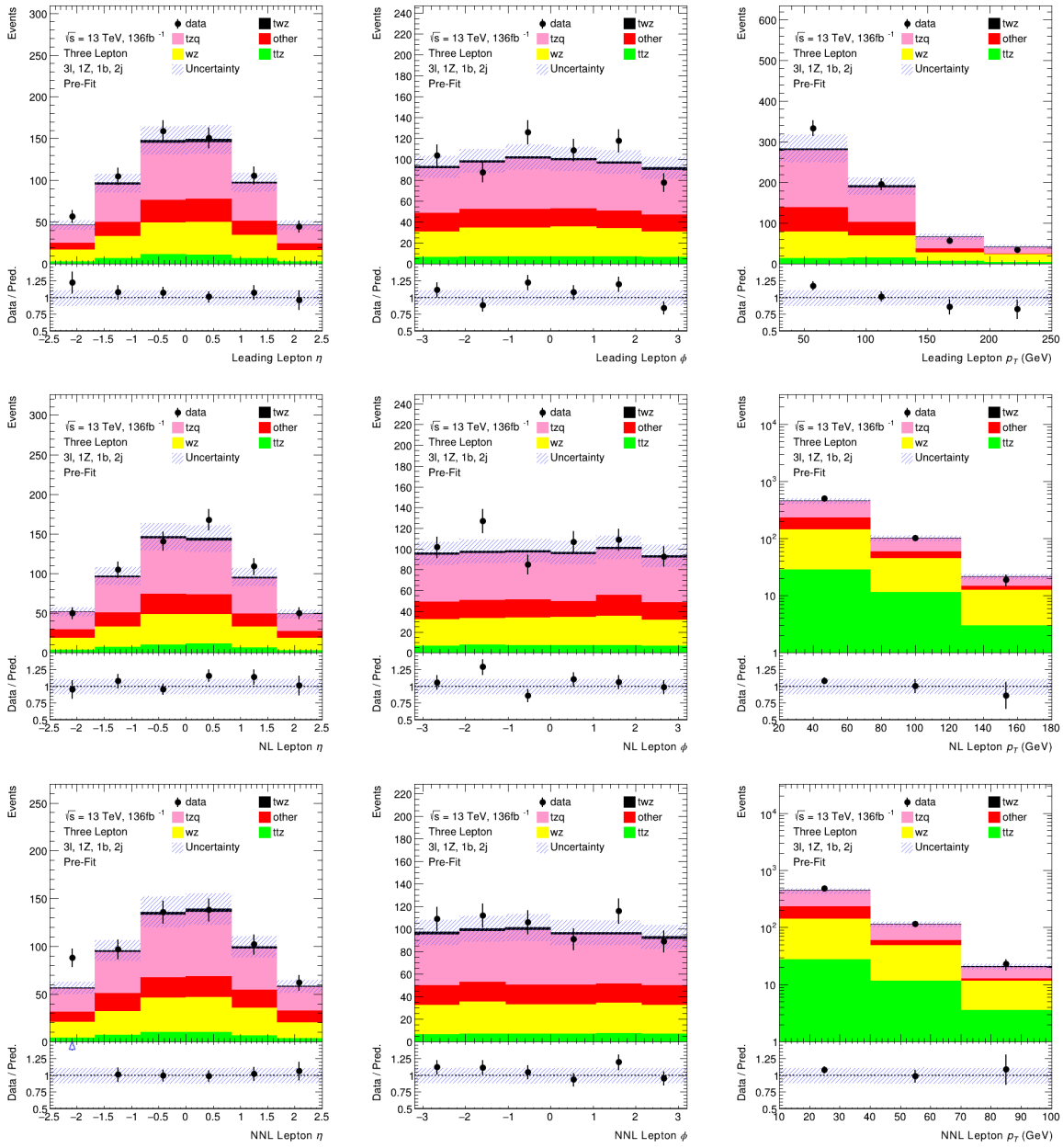


FIGURE 5.4: A comparison of simulation and data for the object kinematics – η, ϕ, and p_T – is shown for the leading, NL, and NNL leptons in the WZ CR

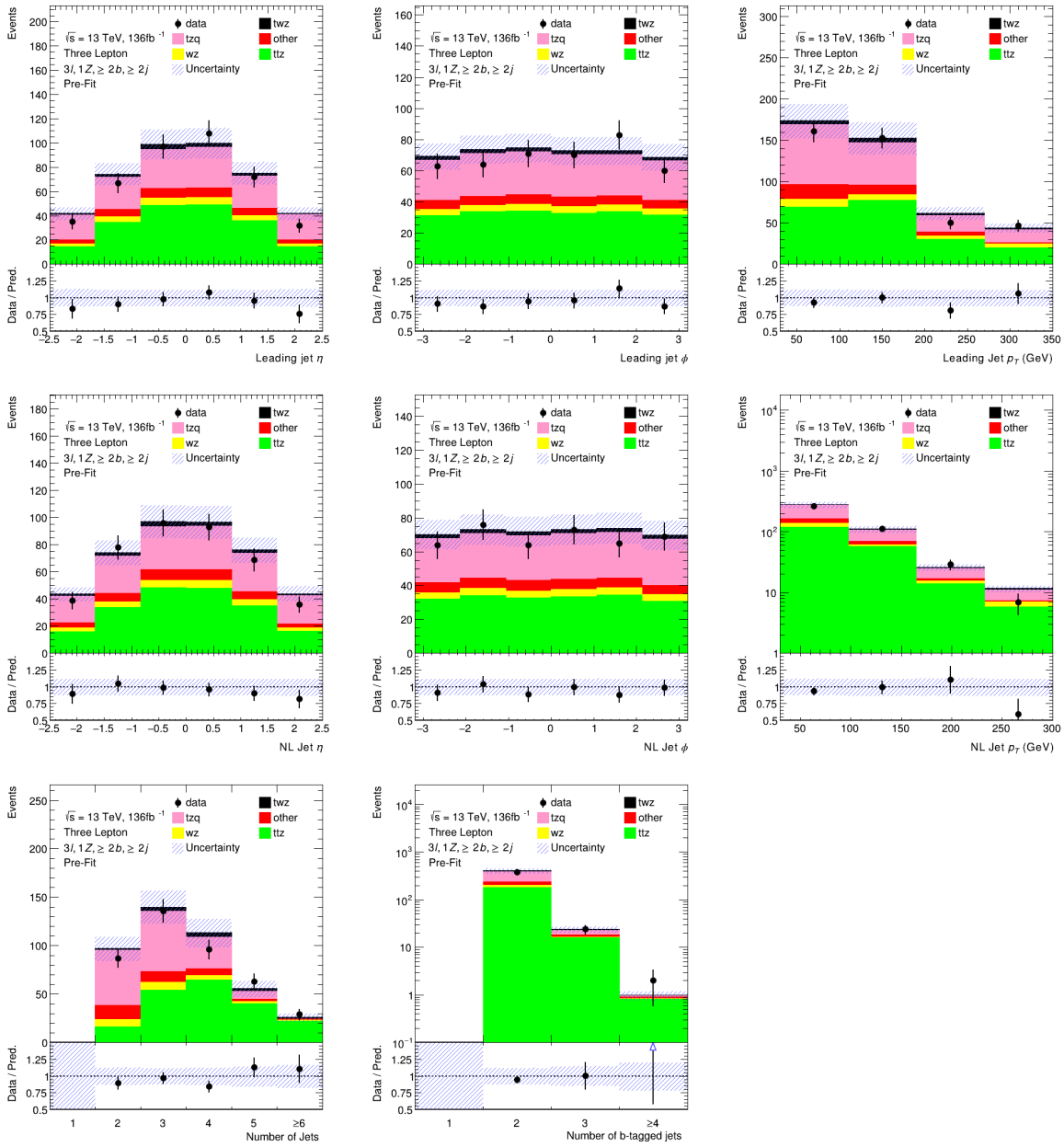


FIGURE 5.5: A comparison of simulation and data for the object kinematics – η, ϕ, and p_T – is shown for the leading, and NL jets in the $t\bar{t}Z$ CR

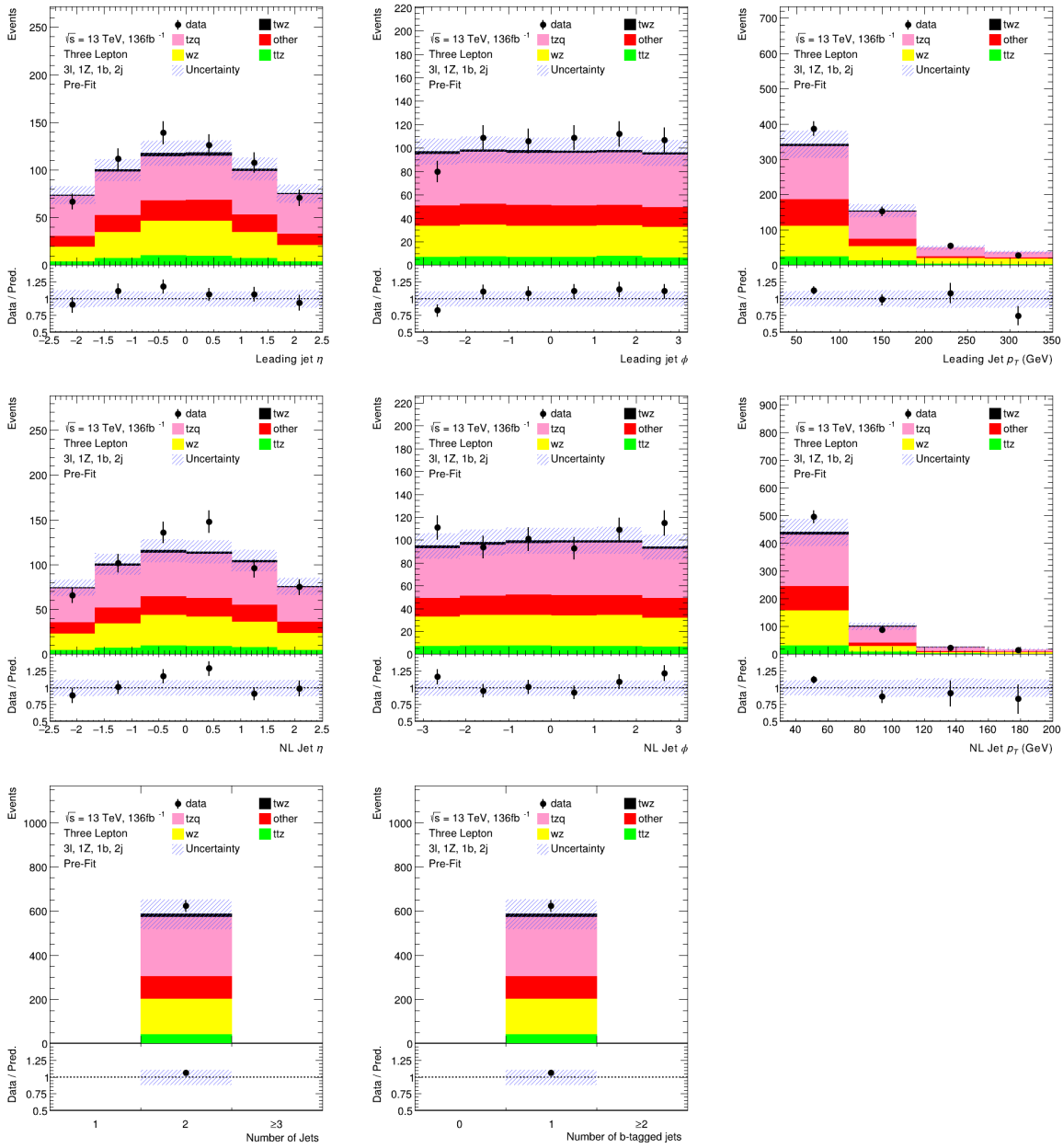


FIGURE 5.6: A comparison of simulation and data for the object kinematics – η, ϕ, and p_T – is shown for the leading, and NL jets in the WZ CR

5.2 Hadronically-decaying W boson identification

In a previous search for tWZ production in the trilepton channel using 36.1 fb^{-1} of $\sqrt{s} = 13$ TeV pp collision data from the ATLAS detector [13], the leading source of uncertainty was due to the uncertainty on the size of the background from WZ production. This analysis aimed to minimize the uncertainty due to this background by using the presence of a hadronically-decaying W boson (W_{had}) to distinguish between tWZ and WZ production. The presence of a W_{had} was used because in events with three leptons originating from tWZ production, there will always be a W_{had} (either coming from the W produced in association with the top quark or from the W coming from the decay of the top quark). However, there will never be a W_{had} in events with three leptons originating from WZ production. In Figures 5.7a and 5.7b, two Feynman diagrams show the two possible decay channels of tWZ to the three leptons, where a W_{had} must be present. However, in Figure 5.7c, WZ production in the trilepton channel can be seen with no W_{had} present.

The identification of W_{had} in an event was achieved using a GBDT – described in Chapter 3 – trained on dijets from a simulated $t\bar{t}$ sample with a single lepton selection. To train the GBDT, labels and features for signal and background dijets were needed, where a dijet is the combination of two jets in an event, a signal dijet is a dijet found to originate a W_{had} in an event, and a background dijet is any other dijet combination in an event.

5.2.1 Labelling dijets

To label dijets as signal or background dijets in simulation, the true particle from which the dijets originate must be known. The information about the origin of the dijets is called the truth information i.e. it is the simulation before the simulated response of the ATLAS detector. When making comparisons between simulation and data, the simulated response of the detector to the dijets must be taken into account, where the information about the simulated response of the detector is referred to as reconstruction, or "reco", information. To label dijets, truth information and reco information must be matched, i.e., in order to label a dijet as a signal dijet, it must be matched to a W boson found in the truth information record from the MC event generator.

In Figure 5.8, a flow chart depicting the procedure for labelling signal and background dijets is shown. First, all the dijet combinations in an event are constructed. Then, to associate a dijet with a W boson, the spatial separation between a W boson, found in the simulation truth information, and a dijet, formed using reco information, was used as a metric to quantify how likely the dijet is to originate from a true W_{had}. The spatial separation is given by:

$$\Delta R_{Truth,Reco} = \sqrt{(\eta_{Truth} - \eta_{Reco})^2 + (\phi_{Truth} - \phi_{Reco})^2}$$

where $\eta_{Truth} - \eta_{Reco}$ is the difference in the η of the W boson and a dijet, and $\phi_{Truth} - \phi_{Reco}$ is the difference in the ϕ of the W boson and a dijet. There

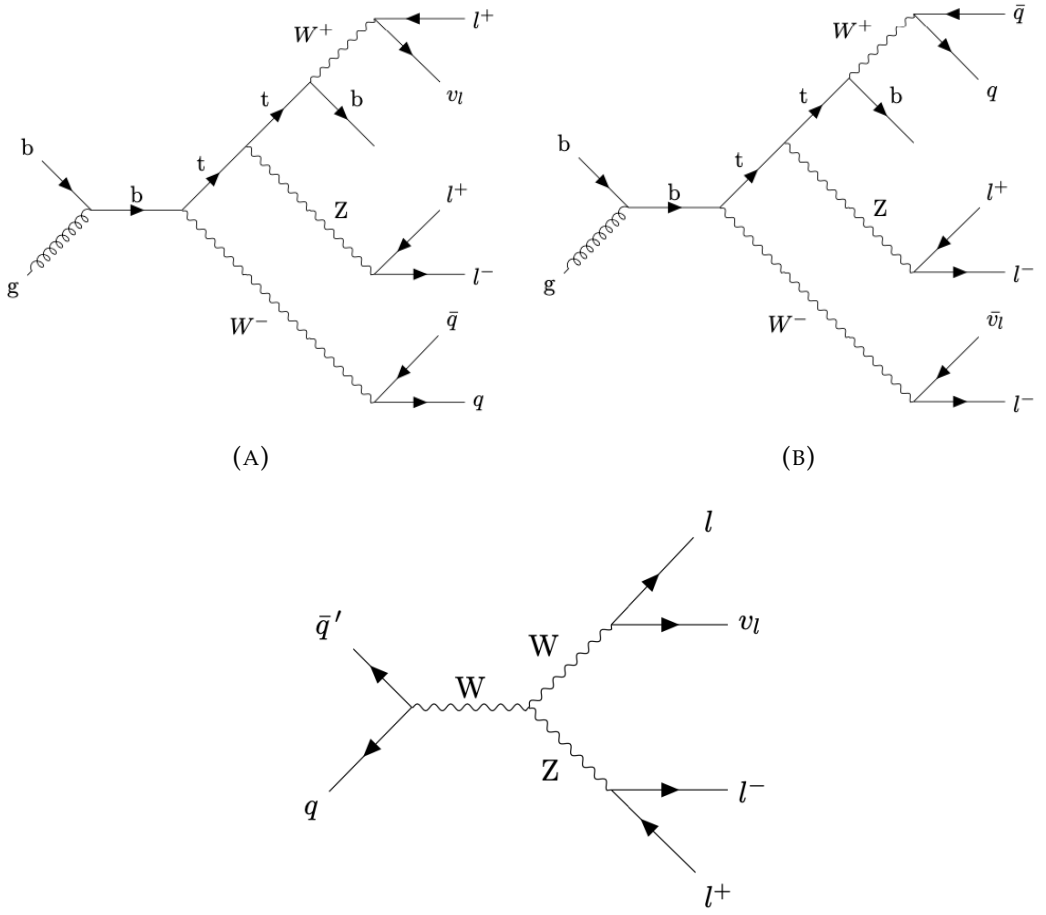


FIGURE 5.7: The top two Feynman diagrams show the two possible decays of tWZ production with a W_{had} present. The bottom Feynman diagram shows the decay of WZ production with no W_{had} present.

can only be one W_{had} per event in tWZ production, therefore, the dijet with the smallest $\Delta R_{Truth,Reco}$ in an event was taken as a candidate signal dijet, and all other dijet combinations were labelled as background dijets. To label candidate signal dijets as signal dijets, a candidate signal dijet was required to have a $\Delta R_{Truth,Reco}$ below a certain cut value. This was done to remove the best matching dijets that actually don't originate from a W boson, since it could be the case that the best matching dijet is not a W_{had}. If a candidate signal dijet satisfied this cut value on $\Delta R_{Truth,Reco}$, it was labelled as a signal dijet.

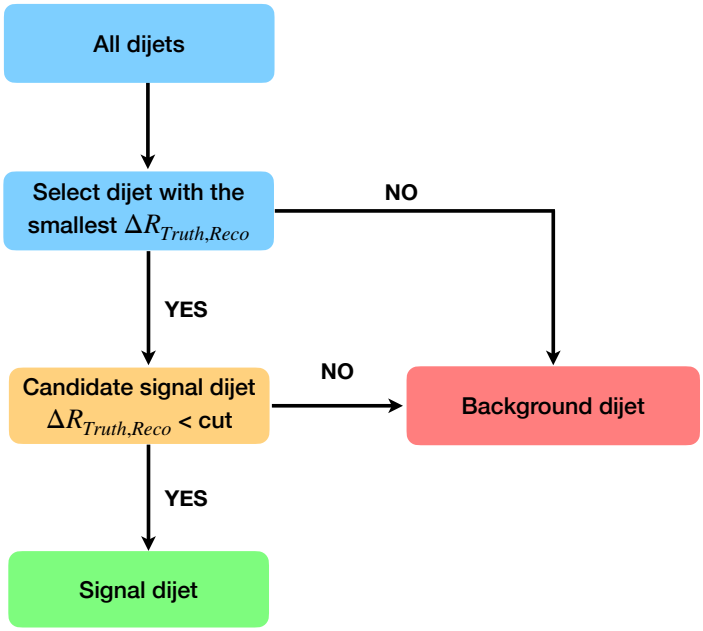
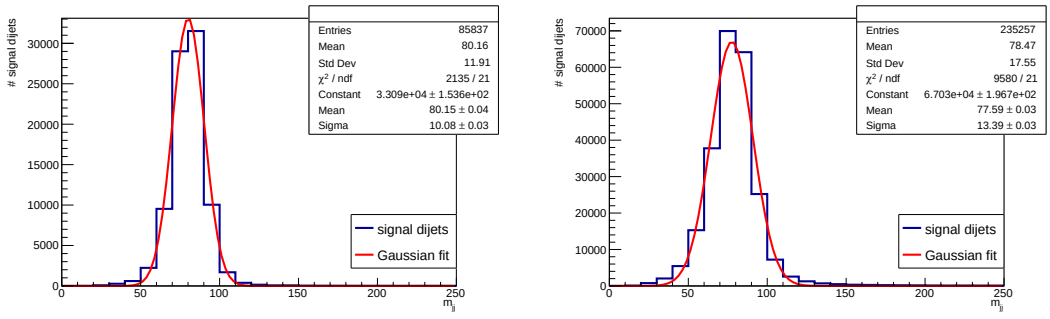


FIGURE 5.8: This flow chart details the procedure for taking all dijets and labelling them as either signal or background dijets.

To determine the optimal cut value on $\Delta R_{Truth,Reco}$, the effect of varying the cut value was studied. In order to study the effect of different cut values, a measure of the purity of the signal dijet class was needed, where purity is taken to be the percentage of W_{had}'s labelled as signal dijets that are actually W_{had}'s. For a pure class of signal dijets, one expects the invariant mass, m_{jj}, distribution of the signal dijets to follow a sharply peaked gaussian distribution with its peak near the mass of the W boson. One would expect a fit to the m_{jj} distribution to have a small standard deviation, $\sigma_{Truth,Reco}$, since a pure signal dijet class would have dijets that all come from W bosons. If background dijets are erroneously added to this signal dijet class, one expects the m_{jj} distribution to have a larger $\sigma_{Truth,Reco}$ value, due to the increased component coming from dijets not originating from a W boson. Given this expected increase in $\sigma_{Truth,Reco}$ for an impure signal dijet class, $\sigma_{Truth,Reco}$ was taken as a measure of the purity of the signal dijet class.

In Figure 5.9, two examples are shown of Gaussian fits (in red) to the m_{jj} distributions of the signal dijets (in blue) for cuts of $\Delta R_{Truth,Reco} < 0.05$ and $\Delta R_{Truth,Reco} < 0.45$. The signal dijets are taken from a $t\bar{t}$ sample with a single lepton selection. In Figure 5.10, the increase in $\sigma_{Truth,Reco}$ – shown in Figures 5.9a and 5.9b as the sigma values in the top right block – as a function of the cut applied to define the signal dijet class can be seen. Figure 5.10 shows that as the cut on $\Delta R_{Truth,Reco}$ increases, $\sigma_{Truth,Reco}$ increases dramatically up until a cut of around 0.5, and then increases slowly for larger values of the cut value. This highlights the fact that as the cut increases, the signal dijet class becomes more contaminated with erroneously added background dijets. Based on Figure 5.10, a cut of $\Delta R_{Truth,Reco} \leq 0.05$ was chosen because it is as tight a cut as possible that still provides a sufficient number of signal dijets for training the GBDT, has a mean value close to the mass of the W boson, and has a small $\sigma_{Truth,Reco}$.

This method is inherently imprecise and it was the case that some signal dijets were labelled as background dijets and vice versa. This was not addressed further, but a method considered to combat this was to find an exclusion region – a range in $\Delta R_{Truth,Reco}$ – where dijets would be excluded from the training of the GBDT if they were in this region. This was not done as it was not found to have a significant effect on the final sensitivity to tWZ production, due to the fact that WZ production was found to no longer be the leading source of uncertainty in the measurement of tWZ production after the inclusion of the W_{had} GBDT as will be discussed in section 5.4.



(A) $\Delta R_{Truth,Reco} < 0.05$ (B) $\Delta R_{Truth,Reco} < 0.45$

FIGURE 5.9: Two examples of the m_{jj} distributions for signal dijets (in blue) with a gaussian fit (in red) to the distributions are shown for all signal dijets. The signal dijets are taken from a $t\bar{t}$ sample with a single lepton selection. In Figure 5.9a the m_{jj} distribution is shown for signal dijets satisfying a cut of $\Delta R_{Truth,Reco} < 0.05$, and in Figure 5.9b the m_{jj} distribution is shown for signal dijets satisfying a cut of $\Delta R_{Truth,Reco} < 0.45$.

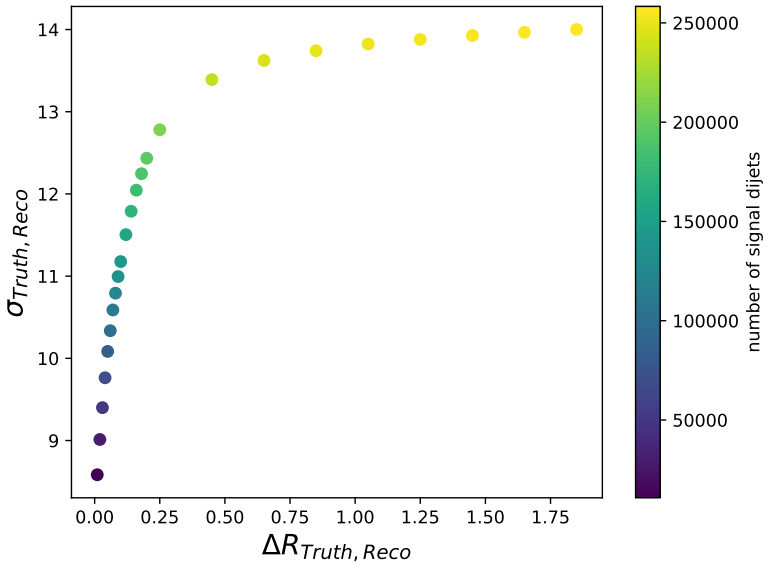


FIGURE 5.10: The $\sigma_{Truth,Reco}$ of the fits to the m_{jj} distributions, as a function of the cut on $\Delta R_{Truth,Reco}$, is shown with the number of signal dijets indicated by the colour axis

5.2.2 Features

For the training of the GBDT, only three features were found to have a significant effect when identifying W_{had}'s. This is because there are only a few degrees of freedom given by the two four-vectors and identification criteria of the dijets. In Figure 5.11, normalized distributions of the chosen features for signal and background dijets are shown. These plots show a clear distinction between signal and background dijets. The three features used for the training were:

- **The separation of the jets – ΔR_{jj}**

  The separation of the jets is given by $\Delta R_{jj} = \sqrt{(\eta_{j1} - \eta_{j2})^2 + (\phi_{j1} - \phi_{j2})^2}$ where $\eta_{j1} - \eta_{j2}$ is the difference in η of the jets, and $\phi_{j1} - \phi_{j2}$ is the difference in ϕ of the jets. Signal dijets should be closer together than background dijet combinations because the signal dijets are decaying from the same W boson. In Figure 5.11a, a normalized distribution of this feature is shown for signal and background dijets, where the signal dijet distribution of this feature peaks close to 1, and the background dijet distribution is more uniformly distributed with a lower peak that is close to 3.

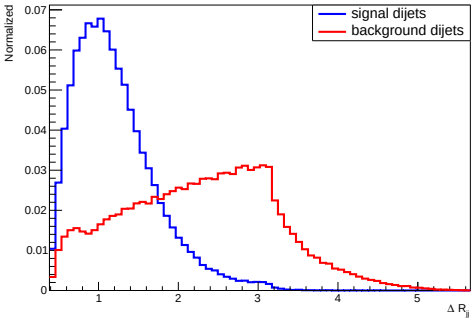
- **The invariant mass of the dijet system – m_{jj}**

  The invariant mass of the dijet system is given by $m_{jj} = \sqrt{E_{jj}^2 - p_{jj}^2}$, where E_{jj} is the energy of the dijet system, and p_{jj} is the momentum of the dijet system. Signal dijets should have an m_{jj} distribution around the W boson mass, whereas background dijets should not be sharply

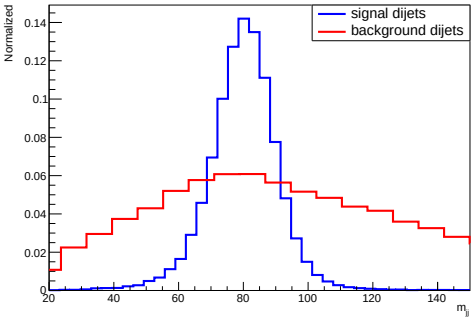
peaked. In Figure 5.11b, a normalized distribution of this feature is shown for signal and background dijets, where the signal dijet distribution peaks around the W boson mass, and the background dijets are more uniformly distributed.

- **Dijet b-tagging information** – *Oneb*

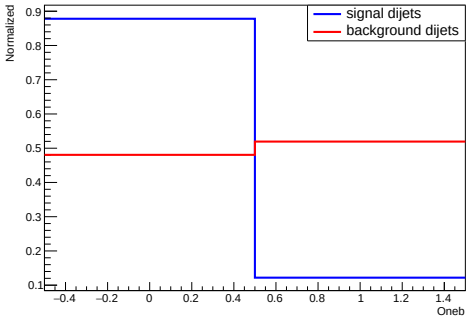
  This feature is constructed to show if at least one of the jets in a dijet system is b-tagged. It is 1 if at least one of the jets in the dijet system is b-tagged and 0 otherwise. Signal dijets should not have a b-tagged jet, whereas a background dijet could, but does not have to, have a b-tagged jet. In Figure 5.11c, a normalized distribution of this feature is shown for signal and background dijets, where the signal dijet distribution has most dijets with no b-tagged jets, and the background dijet distribution has close to an equal number of dijets with and without at least one b-tagged jet.



(A) ΔR_{jj}



(B) m_{jj}



(C) *Oneb*

FIGURE 5.11: Normalized distributions of the GBDT features for signal dijets (in blue) and background dijets (in red) are shown. The dijets used are from a $t\bar{t}$ sample with a single lepton selection. Figure 5.11a shows the ΔR_{jj} distribution, Figure 5.11b shows the m_{jj} distribution, and Figure 5.11c shows the *Oneb* distribution

5.2.3 $t\bar{t}$ training and event re-weighting

Any simulated events that a GBDT is trained on have to be discarded because the GBDT is biased towards these events. Furthermore, an event-level GBDT is used in section 5.3, which means that any events used for the the training of the event-level GBDT must also be discarded. Therefore, the training of these two GBDTs would require a large percentage of the simulated events to be discarded for the fit. Specifically this would mean that a large percentage of tWZ events would have to be discarded, which is problematic because there are only a limited number of simulated tWZ events. To combat this problem, dijets from a $t\bar{t}$ sample with a single lepton selection were used to train the W_{had} GBDT.

The $t\bar{t}$ sample with a single lepton selection was chosen because this sample with this selection is not used in the analysis, and it has an abundance of signal dijets. However, because this production process is different to tWZ production, the kinematics of the signal and background dijets as well as the features for the dijets were checked to make sure that they were not different from the dijets found in the mix of signal and background events used in the analysis. No significant differences were found between the $t\bar{t}$ dijets and the dijets from the mix of signal and background events. Thus, dijets from the $t\bar{t}$ sample with the single lepton selection were used for the training of the W_{had} GBDT.

In this sample, there are significantly more background dijets used for training because there is only one signal dijet per event, but there can be multiple background dijets per event. This resulted in an imbalanced training dataset which produced a GBDT that mistakenly classified all dijets as background dijets, since the GBDT was able to quickly minimized its objective function by doing so. In order to correctly classify signal dijets, all dijets had to be re-weighted such that the sum of weights of all the signal dijet weights was equal to the sum of weights of all the background dijet weights. By equally weighting the sum of signal dijet weights and the sum of background weights, the objective function was minimized by correctly classifying the signal dijets as well as the background dijets.

5.2.4 Max W_{had} GBDT score

In Figure 5.12, a histogram of the W_{had} GBDT score – where the score is the probability of a dijet being a signal dijet – is shown for a test dataset comprising an equal number of signal and background dijets. The histogram shows a clear distinction between the signal and background dijets with the former peaking at 1 and the later at 0.

The W_{had} GBDT score is used to construct an event-level variable that can be used to identify W_{had}'s in an event, so that tWZ events can be distinguished from WZ events. This variable is called the Max W_{had} GBDT score and is constructed by taking the maximum scoring dijet per event. Events that have a W_{had} can be identified because the Max W_{had} GBDT score should be high for

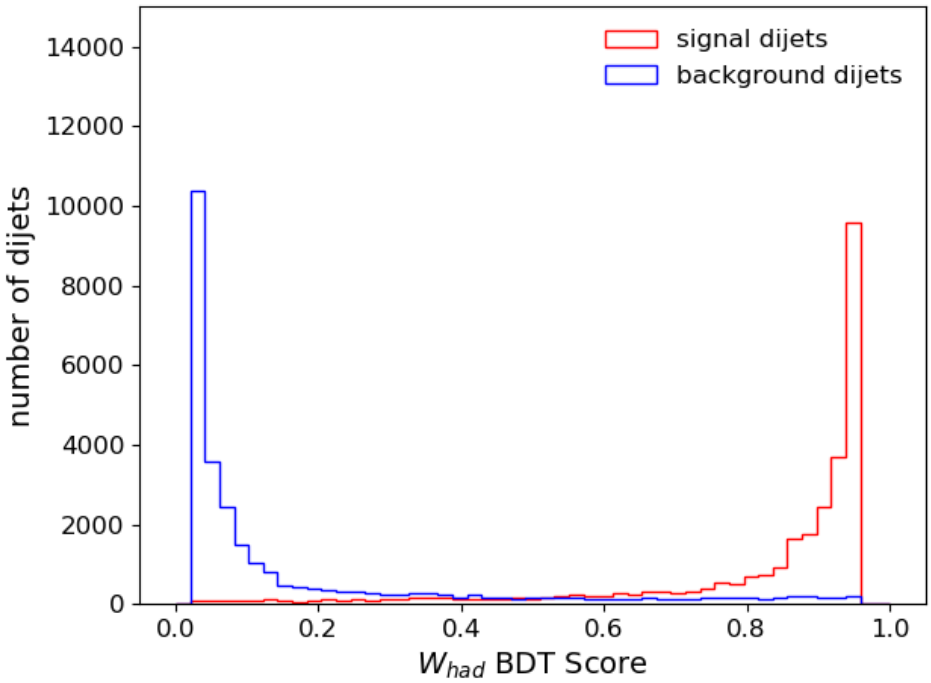


FIGURE 5.12: W_{had} GBDT scores are shown for an equal number of signal dijets (in blue) and background dijets (in red).

an event with a W_{had} present and low for an event without a W_{had} present. In Figure 5.13, the Max W_{had} GBDT score is shown in the SR and CRs for simulation and data. In the SR, tWZ events, normalized to the total number of background events, are shown by a dotted line. The normalized tWZ events show that this variable discriminates well between signal and background, since the tWZ events are placed predominantly to the right of the histogram since they should have a W_{had} present. In the CRs, there is good agreement between data and simulation, except in the last bin of the WZ CR. This disagreement in the WZ CR is under investigation, but because there are few events in this bin, it is not expected to have a significant impact on the analysis. This could also be evidence that the estimation of backgrounds from fake leptons entirely from simulation is an inadequate approach.

In Figure 5.14, the features of the Max W_{had} GBDT score are shown for simulation and data in the $t\bar{t}Z$ CR (on the top row), and the WZ CR (on the bottom row). These are used to check the modeling of the simulation for these features. In both CRs, there is good agreement between data and simulation as shown by the Data/Pred. at the bottom the histograms. However, there is disagreement in the second last bin of the Max Scoring W_{had} m_{jj} variable. This disagreement in the WZ CR is under investigation, but because there are few events in this bin, it is not expected to have a significant impact on the analysis. This could also be evidence that the estimation of backgrounds from fake leptons entirely from simulation is an inadequate approach.

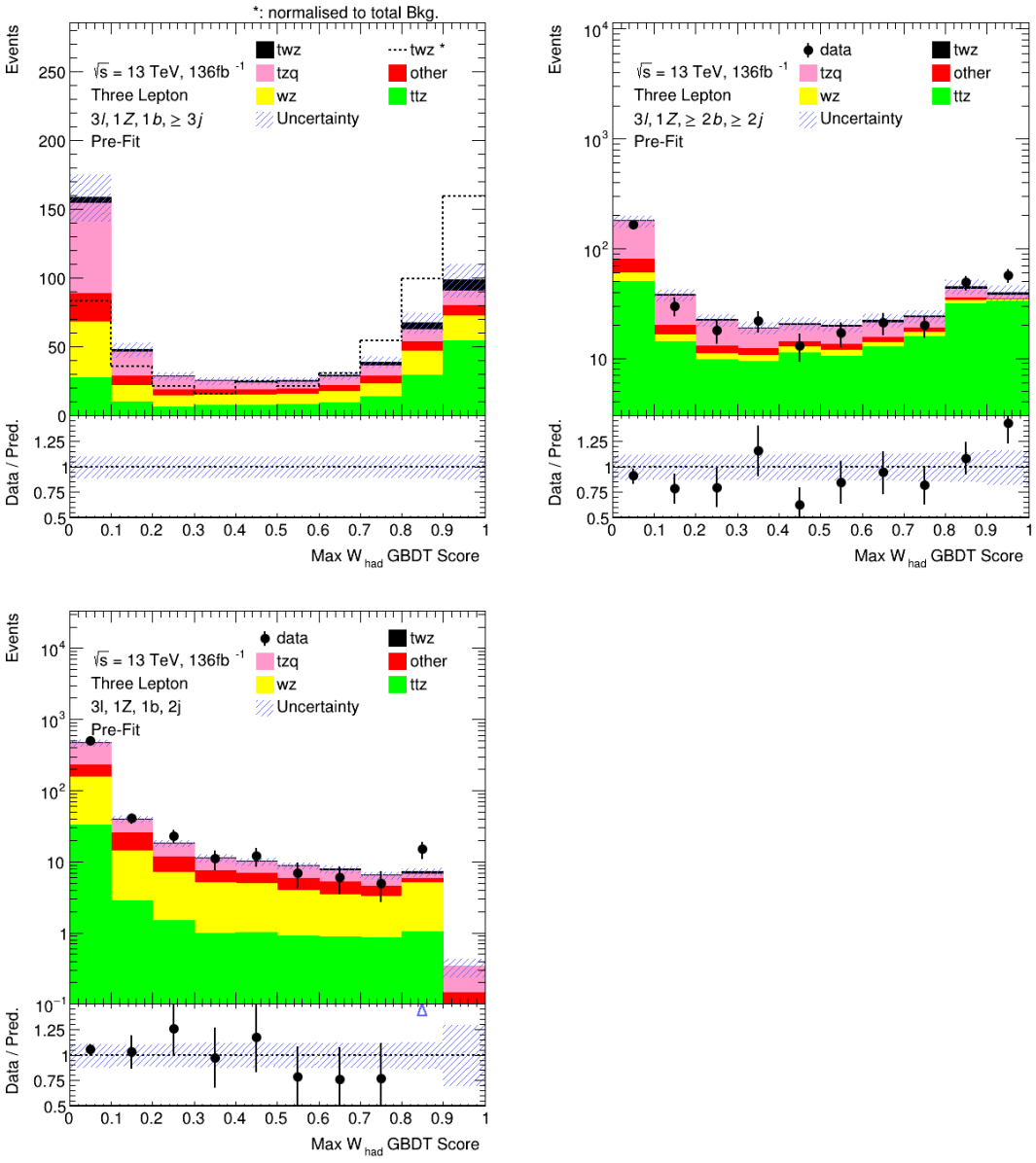


FIGURE 5.13: The Max W_{had} GBDT score is shown for simulation in the SR and CRs, and for data in the CRs. A distribution of tWZ events, normalized to the total number of background events, is shown in the SR by a dotted line.

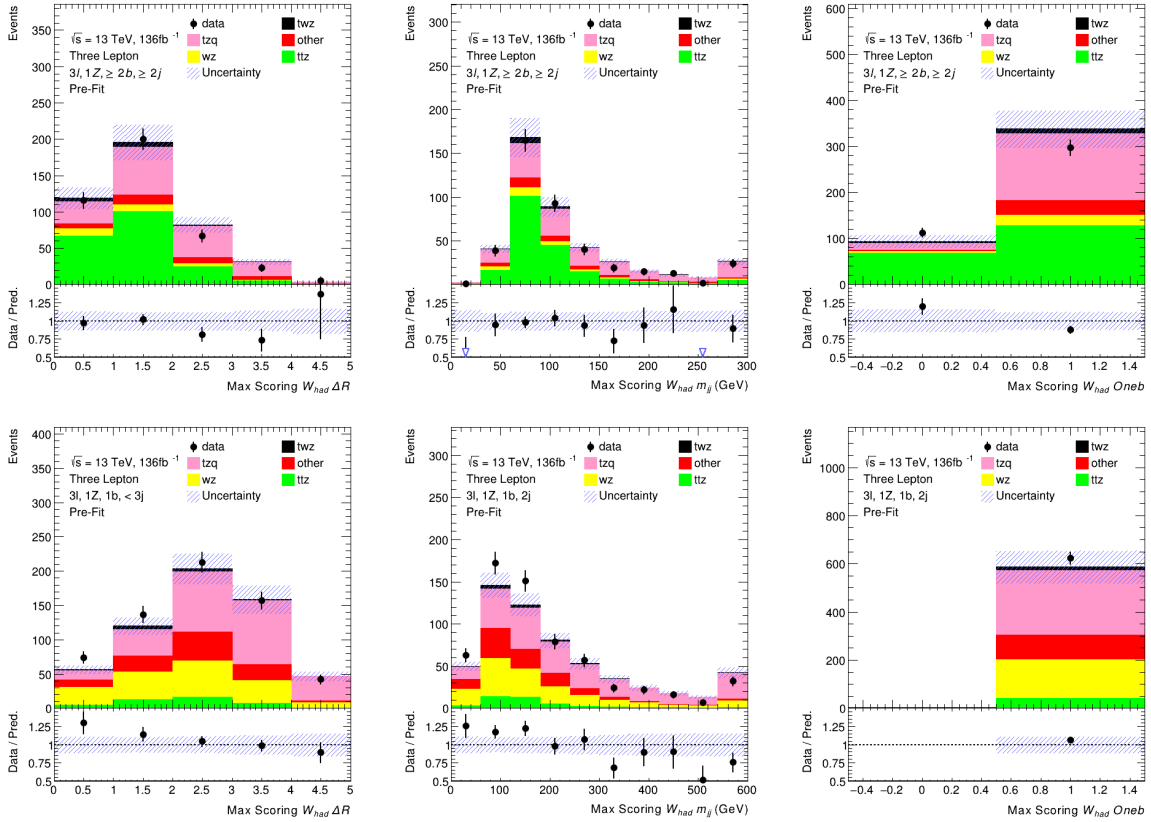


FIGURE 5.14: A comparison of data and simulation for W_{had} GBDT input features – ΔR_{jj}, m_{jj}, and $Oneb$ – is shown for the maximum scoring dijet in the $t\bar{t}Z$ CR – shown on the top row – and the WZ CR – shown on the bottom row.

5.3 Event-level GBDT

To distinguish between tWZ production and background production processes, a GBDT was trained on half of all the simulated events. The GBDT makes use of combinations of features to distinguish between tWZ and multiple backgrounds simultaneously, since one feature may be useful for distinguishing between tWZ and one particular background, while another feature may be useful for distinguishing between tWZ and another background.

5.3.1 Labels and features

The events from tWZ production were labelled as the signal class and all background events were labelled as the background class. To account for the expected number of signal and background events found in the detector, each event was assigned a weight in the training that was equal to its simulated event weight. Additional re-weighting of the signal class was applied to account for the imbalance in the signal and background classes due to the rarity of tWZ events, this is dealt with in section 5.3.2.

In Figure 5.15, normalized distributions of the features used for training are shown for the half of all simulated events used for training the event-level GBDT, with all the object and event selections applied. tWZ events are shown in red, $t\bar{t}Z$ events in dark blue, WZ events in dark green, tZq events in purple, and the combination of all the other smaller backgrounds shown in yellow.

- **Max W_{had} GBDT score**

  This feature is the event-level discriminator outlined in the section 5.2.4. It is used to distinguish between tWZ and WZ production. The normalized distributions for the different production processes are shown in Figure 5.15a. This feature does not help distinguish between tWZ and $t\bar{t}Z$, because $t\bar{t}Z$ also has a W_{had} in the trilepton channel, but it does help distinguish between the other backgrounds since a W_{had} is not expected in tZq or any of the other smaller background events.

- **Min $m_{\text{Non-Z l},b_i}$**

  In the event selections, exactly one Z boson is required in all events. This leaves one lepton that does not come from the Z boson. This lepton is called the Non-Z lepton. This feature is the minimum invariant mass of the vectorial sum of the Non-Z lepton and a b-tagged jet per event. This feature is used to distinguish between tWZ and $t\bar{t}Z$, since one of the top quarks in $t\bar{t}Z$ has to decay leptonically in the three lepton channel. Therefore, the minimum vectorial sum of a lepton and b-tagged jet originating from $t\bar{t}Z$ should be more sharply peaked since these two objects decay from the same top quark. In Figure 5.15b, a normalized histogram of this variable is plotted for all the different production processes. The difference between tWZ and these backgrounds

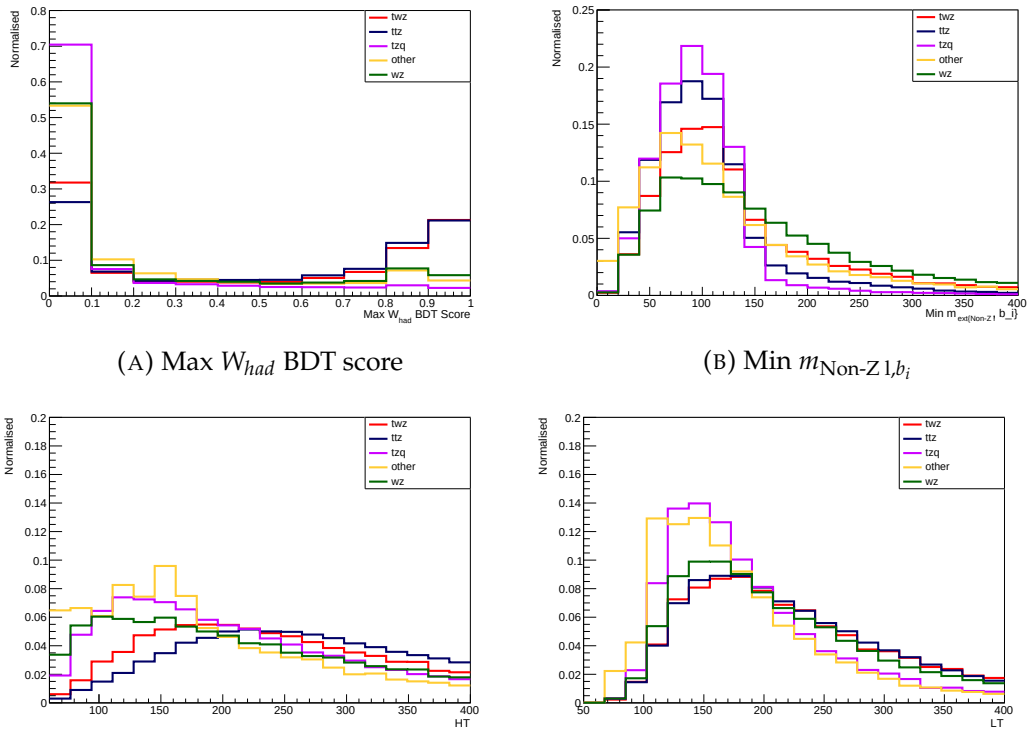
can be seen since tWZ has a right shifted peak and a longer tail compared to $t\bar{t}Z$ and tZq, whereas WZ and the other production processes have a distribution that is similar to tWZ events.

- H_T

  This feature is the scalar sum of jet p_T's. It does not help distinguish between tWZ and $t\bar{t}Z$, but does help distinguish between tWZ and several other backgrounds. In Figure 5.15c, normalized histogram of this variable is plotted for all the different production processes. The difference between tWZ and its backgrounds can be seen where the distributions of tWZ and $t\bar{t}Z$ are shifted to the right compared to the other backgrounds which are more prominent at low values of H_T.

- L_T

  This feature is the scalar sum of lepton p_T's. It is used to distinguish between tWZ, tZq, and combination of smaller background processes. In Figure 5.15d, a normalized histogram of this variable is plotted for all the different production processes, where the tZq, and other backgrounds are more sharply peaked at lower values of L_T compared to tWZ, $t\bar{t}Z$, and WZ.



(A) Max W_{had} BDT score

(B) Min $m_{\text{Non-Z}\,l,b_i}$

(C) H_T

(D) L_T

FIGURE 5.15: Normalized distribution of event GBDT features for tWZ and background events are shown. (a) shows the Max W_{had} GBDT score feature, (b) shows the Min $m_{\text{Non-Z}\,l,b_i}$ feature, (c) shows the H_T feature, and (d) shows L_T feature

In Figure 5.16, the event-level GBDT features are shown for simulation and data in the $t\bar{t}Z$ CR. In Figure 5.17 the same is shown in the WZ CR. In both cases there is generally good agreement, except in the Max W_{had} GBDT score in the WZ CR, as discussed previously, and in the LT feature in the WZ CR at high LT. This is under investigation, but due to the small number of entries in the bins, these discrepancies are not expected to have a significant effect. This could also be evidence that the estimation of backgrounds from fake leptons entirely from simulation is an inadequate approach.

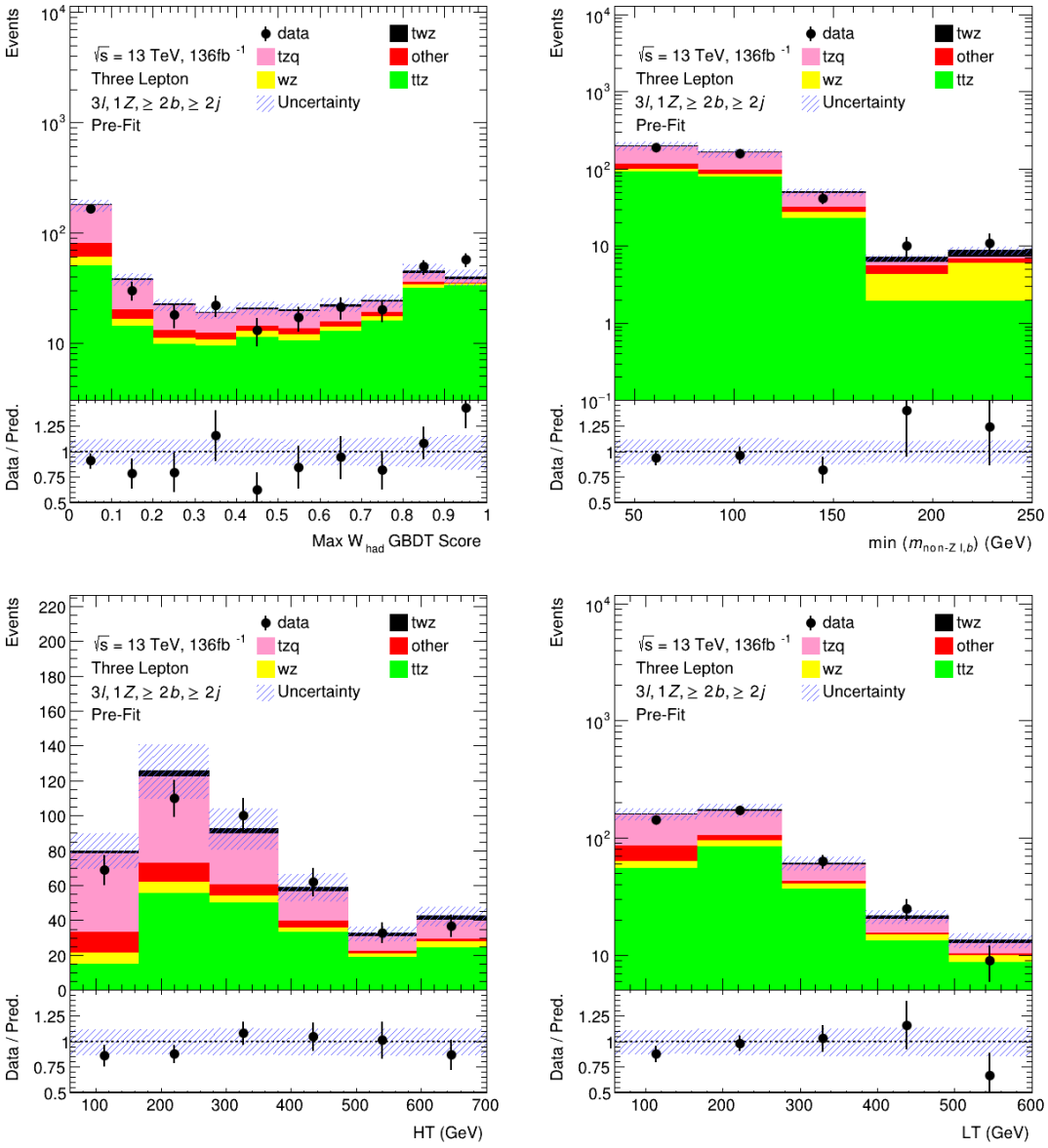


FIGURE 5.16: A comparison of data and simulation for the event-level GBDT input features – Max W_{had} GBDT score, Min $m_{\text{Non-Z}\,l,b_i}$, H_T, and L_T – is shown in the $t\bar{t}Z$ CR

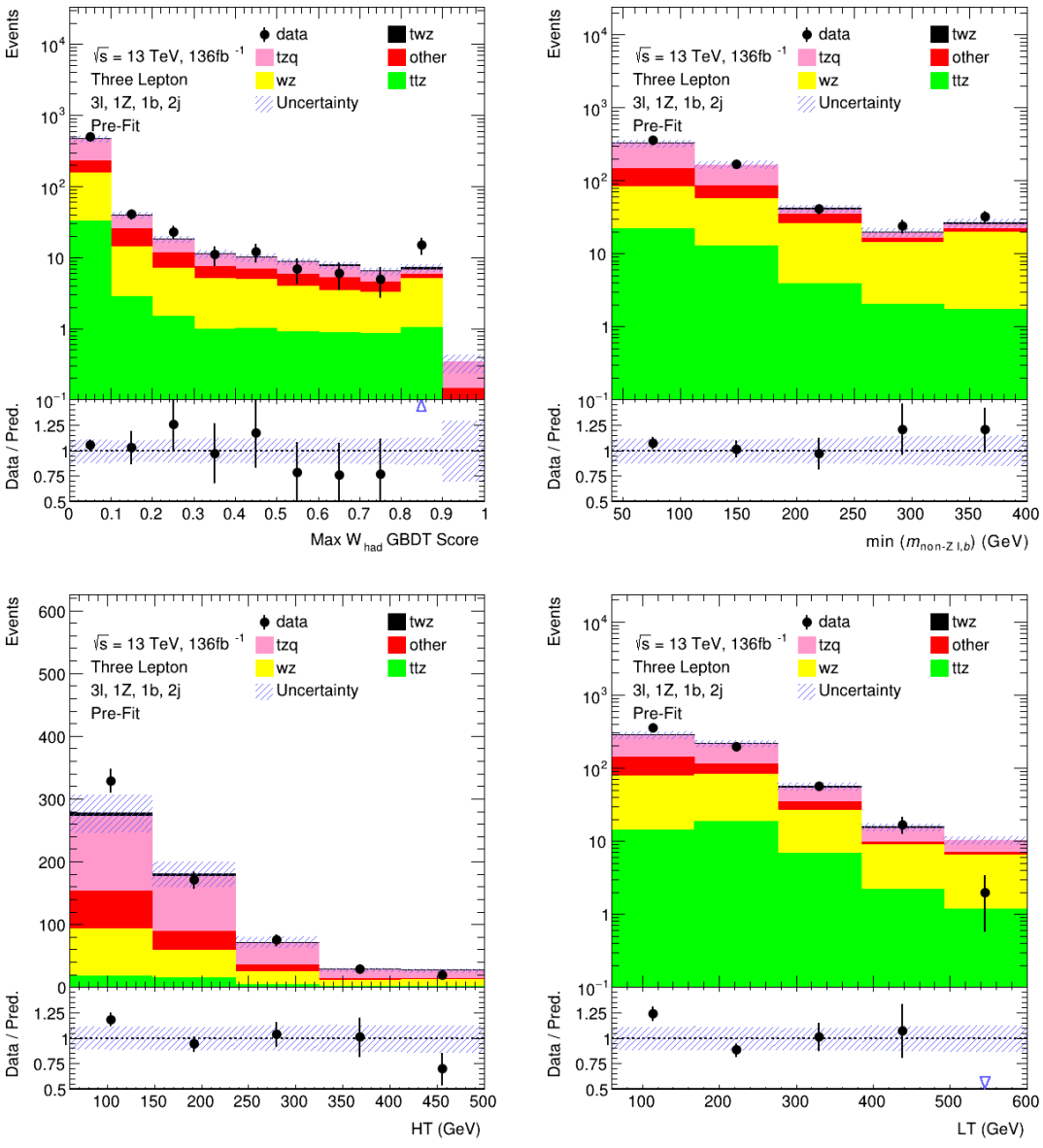


FIGURE 5.17: A comparison of data and simulation for the event GBDT input features – Max W_{had} GBDT score, Min $m_{\text{Non-Z l},b_i}$, HT, and LT – is shown in the WZ CR

5.3.2 Re-weighting of signal events

In the case of training a dataset with significantly more background events, the objective function of a GBDT will be minimized by trivially classifying all events as background. Since this is the case when training the event-level GBDT, each tWZ event is up-weighted in the training by multiplying its weight by the ratio, a, of the sum of all the background simulated weights to the sum of all the old signal simulation weights

$$w_s^{new} = w_s^{old} \cdot a$$

where $a = \frac{\sum w_b}{\sum w_s^{old}}$ This forces the GBDT to focus more on tWZ events during training, because now the sum of signal weights equals the sum of background weights.

$$\sum w_s^{new} = \sum w_b$$

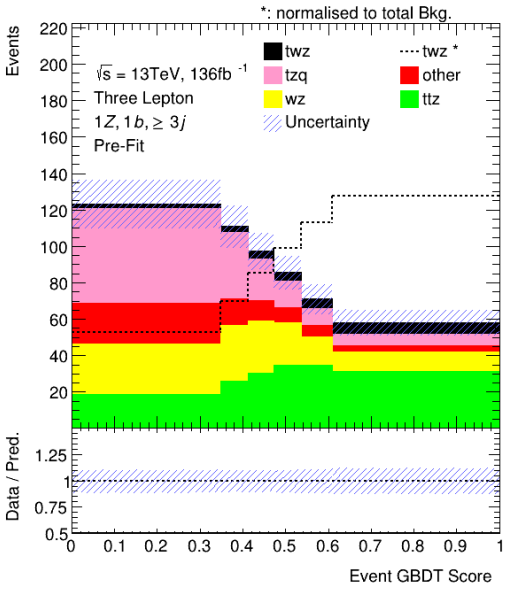
This means that now the GBDT minimizes the objective function by correctly classifying the rare tWZ events.

In addition to the re-weighting of signal event weights, all events with negative MC weights had to be removed before training and before the signal re-weighting was applied. This was done because of a technical limitation in the scikit-learn library. This means that the events used for training are not entirely representative of what is expected in data from the ATLAS detector. This is an unavoidable issue when using the sci-kit learn library and the effect is not expected to be significant due to the relatively small number of negatively weighted events compared to the total number of events.

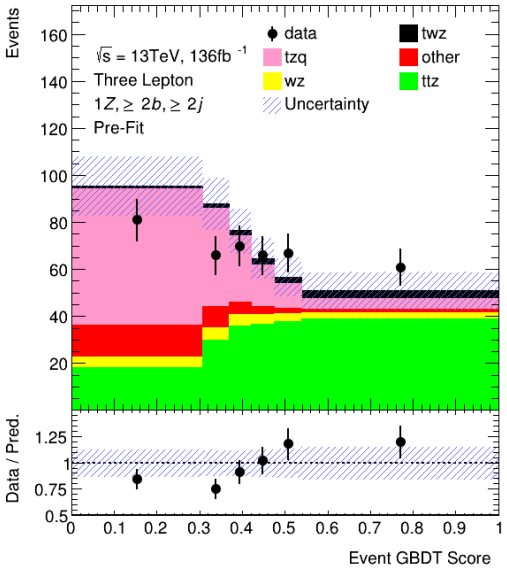
5.3.3 Event GBDT Score

In Figure 5.18, the Event GBDT Score is shown for simulation and data in the SR and CRs, where the histograms have been auto-binned to minimize uncertainty due to statistical fluctuations. In this case, the "TransfoD" autobinning algorithm was used with both parameters set to 3. The autobinning algorithm starts from the bin with largest output and merges bins until a certain fraction of signal and background events is obtained [38].

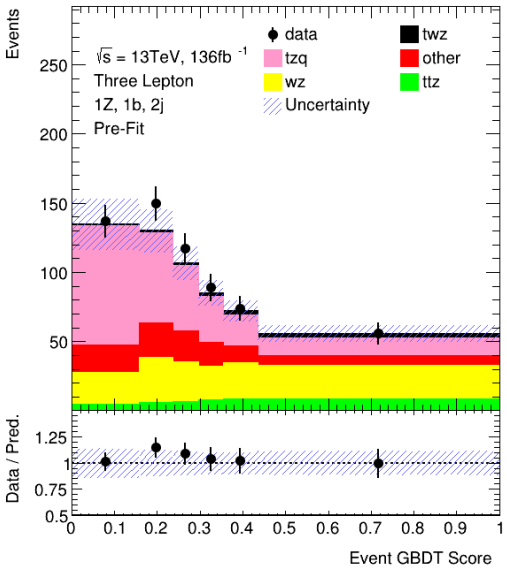
In Figure 5.18a, the Event GBDT Score is shown in the SR and is blind to data. A distribution of tWZ events, normalized to the total number of background events, is shown by a dotted line. This distribution shows that the Event GBDT Score discriminates well between signal and backgrounds. In Figures 5.18b and 5.18c, the Event GBDT Score is shown in the $t\bar{t}Z$ and WZ CRs. There is good agreement in the WZ CR, however, in the $t\bar{t}Z$ CR there is disagreement in the bin from 0.3 to 0.4. This could be evidence that the estimation of backgrounds from fake leptons entirely from simulation is an inadequate approach.



(A) tWZ SR



(B) $t\bar{t}Z$ CR



(C) WZ CR

FIGURE 5.18: The Event GBDT Score is shown for simulation and data in the $t\bar{t}Z$ CR (b) and the WZ CR (c). The tWZ SR is blind to data and a distribution of tWZ events, normalized to the total number of background events, is shown by a dotted line. In all regions the histograms have been auto-binned using the "TransfoD" algorithm.

5.4 Determining an expected upper limit on tWZ production

The procedure for estimating the parameters of a model given some data is referred to as parameter estimation [39]. Parameter estimation is used to determine a parameter of interest, μ, in a hypothesised model while taking into account systematic uncertainties that are represented by nuisance parameters, $\vec{\theta}$. The parameter fitting technique used in the search for tWZ production is the the method of maximum likelihood. This is a method that uses a finite sample of data to maximize a likelihood function, $L(\mu, \vec{\theta})$, by determining estimated values of the parameter of interest, $\hat{\mu}$, and the nuisance parameters, $\vec{\hat{\theta}}$.

In the search for tWZ production, the likelihood function is the product of Poisson probabilities for all the bins of the Event GBDT Score histograms in the SR and CRs, where the prior knowledge of the nuisance parameters are expressed via explicit constraint terms in the likelihood function. The parameter of interest in this fit is the signal strength of tWZ production, which is defined as the ratio of the observed cross section to the predicted SM cross section of tWZ production: $\mu = \frac{\sigma_{tWZ,obs}}{\sigma_{tWZ,SM}}$.

In this section, an expected upper limit on the signal strength of tWZ production is determined from a fit to a modified Asimov dataset. An Asimov dataset is a dataset that exactly matches the SM prediction of signal and background in every bin (i.e $\mu = 1$) and the nuisance parameters equal their nominal pre-fit values [40]. For a modified Asimov dataset, first a background-only fit is performed to the CRs using ATLAS pp collision data. This provides a set of fitted values for the nuisance parameters that affects the background predictions in the CRs. These fitted nuisance parameter values are then used to construct a modified Asimov dataset in the SR, where no data is used in the SR for the fit. This is done so that the experimental sensitivity of the analysis setup can be more accurately determined while avoiding any potential bias introduced by knowledge of the data in the SR.

5.4.1 Systematic uncertainties

Various sources of systematic uncertainty are included in the fit to the likelihood function as nuisance parameters. In this analysis, the fit includes only a subset of all the possible systematic uncertainties. These include theory uncertainties on the background production process cross sections, experimental uncertainties, and uncertainties due to the statistical fluctuations of simulation in bins. The included systematic uncertainties are expected to be the dominant systematic uncertainties and the systematic uncertainties that were not included are not expected to have a large impact. Furthermore, the statistical uncertainties are the dominant source of uncertainty in this analysis.

The theory systematic uncertainties are assigned to the $t\bar{t}H$, $t\bar{t}Z$, WZ, $t\bar{t}$, tZq, and ZZ cross sections. These were all allowed to vary by 20%, which is higher than the calculated theoretical uncertainties on the inclusive cross sections. This is set higher than the calculated uncertainties in order to increase the chances of producing a good fit to the data, to ensure that these uncertainties are not underestimated, and to account for the uncertainties on the acceptances of these production processes in the SR and CRs.

Experimental uncertainties are assigned to the luminosity and the calibrations applied to the simulation in section 5.1.2. These are based on ATLAS recommendations and are as follows:

- **Luminosity** The systematic uncertainty due to the luminosity is set to 3%.
- **Pileup** To account for the difference between the pile-up distributions in simulation and data, a systematic uncertainty related to the calibration used to adjust the simulated pileup to the data pileup, is applied. This uncertainty is obtained by re-scaling the average number of additional proton-proton collisions per bunch crossing.
- **Lepton Triggers** To account for the difference between the performance of the lepton triggers in simulation and data, a systematic uncertainty related to the calibration used to adjust the simulated performance of the lepton triggers to the data lepton trigger performance, is applied.
- **Jet vertex tagger** To account for the difference between the performance of the jet vertex tagger (jvt) in simulation and data, a systematic uncertainty related to the calibration used to adjust the simulated performance of the jvt to the data jvt performance, is applied.
- **b-tagging** To account for the difference between the performance of the b-tagging algorithm in simulation and data, systematic uncertainties related to the calibrations used to adjust the simulated performance of the b-tagging algorithm are applied. These are applied separately for jets and b-tagged jets, and are labelled bTagL and bTagB respectively.

Lastly, the uncertainties due to the statistical fluctuations of simulation in certain bins was included in the fit. These are referred to as gammas, γ's. In this fit, γ nuisance parameters were added for each bin with a statistical uncertainty greater than a predefined threshold of 0.001.

5.4.2 Results

The TRExFitter framework [41] for binned template profile likelihood fits, which uses the HistFactory tool [42], is used to perform the maximum likelihood fit. The fit to a modified Asimov dataset in the TRExFitter framework is performed using the histograms of the event GBDT score in the SR and CRs, with the autobinning setting turned on to minimize statistical fluctuations.

The expected signal strength resulting from the fit to the modified Asimov dataset is:

$$\hat{\mu} = 2.08^{+1.48}_{-1.45}$$

The best fit value is consistent with $\mu = 1$, as expected from a fit to a modified Asimov dataset. In Figure 5.19, the post-fit histograms of the event GBDT score are shown for simulation and data, where the estimates of nuisance parameters extracted from the background-only fit to the CRs are used. Comparing these to the pre-fit plots in Figure 5.18, one sees that the uncertainty bands have narrowed due to the use of the nuisance parameters estimated from the background-only fit in the CRs. In Figure 5.19c, there is good agreement between simulation and data in the WZ CR even given the narrowing of uncertainty bands. However, in Figure 5.19b, there is poor agreement between simulation and data in the bin from 0.3 to 0.4 in the $t\bar{t}Z$ CR. This is evidence that perhaps a fit including more systematic uncertainties is needed to produce a good fit to the data, and that the estimation of backgrounds from fake leptons entirely from simulation is an inadequate approach.

The significance, Z, of a measurement is defined such that a Gaussian distributed variable found Z standard deviations, σ, above its mean has a upper tail probability equal to the p-value. That is

$$Z = \Phi^{-1}(1 - p)$$

where $\Phi^{-1}(1 - p)$ is the quantile of the standard Gaussian, and p is the p-value [40]. The significance is reported in units of σ, and quantifies the deviation of the best-fit value of μ from 0, assuming the SM hypothesis.

The expected significance of μ was found to be

$$Z_{\mu}^{exp} = 0.72\,\sigma$$

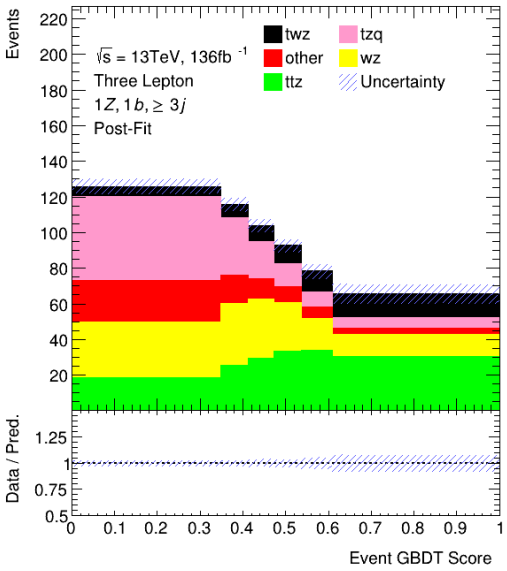
The particle physics community has tended to regard rejection of the background-only hypothesis with a significance of at least $Z = 3\,\sigma$ as appropriate for evidence, and at least $Z = 5\,\sigma$ as appropriate to constitute a discovery. Therefore, given the current analysis setup, one does not expect to be able to measure tWZ production at the level necessary to constitute evidence for this production process.

Furthermore, when one does not expect to observe the production process, the result of the fit is reported by giving an upper limit on the parameter of interest. Since this is the case here, this is what is done. The upper limit at a confidence level of $1 - \beta$ is defined as

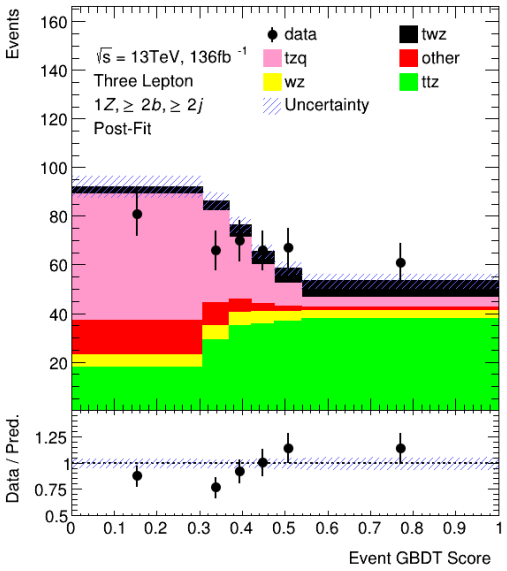
$$\mu_{up} = \hat{\mu}_{obs} + \sigma_{\hat{\mu}}^{up}\Phi^{-1}(1 - \beta)$$

where $\sigma_{\hat{\mu}}^{up}$ is the upper uncertainty on μ and $\Phi^{-1}(1 - \beta)$ is the quantile of a standard Gaussian at a confidence level of $1 - \beta$.

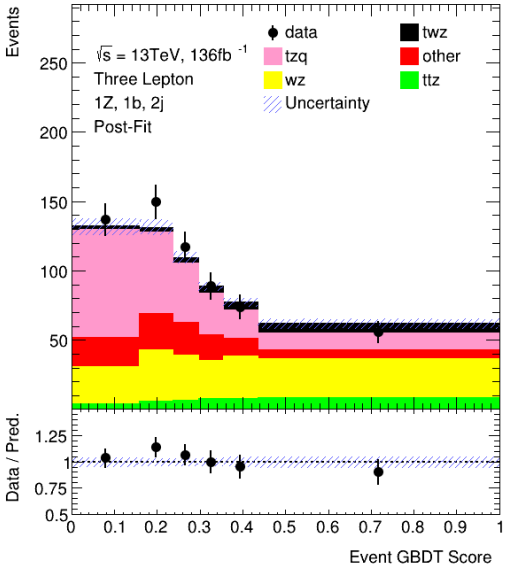
In TRExFitter, the upper limit is calculated using the CL_s procedure [43]. By



(A) tWZ SR



(B) $t\bar{t}Z$ CR



(C) WZ CR

FIGURE 5.19: The post-fit event GBDT score histograms are shown with with the post-fit model resulting in narrowed uncertainty bands.

convention, a confidence level of 95% ($\beta = 0.05$) is chosen such that the interval [0,μ_{up}] will cover μ with a probability of at least 95%. This amounts to $\Phi^{-1}(0.95) = 1.645$. The expected upper limit determined from the fit to the modified Asimov dataset was found to be

$$\mu_{up}^{exp} = 2.77^{+2.39}_{-1.28}$$

Impact of nuisance parameters

The effect of the nuisance parameters is quantified by the impact, $\Delta\mu$, they have on the estimation of μ. The impact is determined by the difference between the best-fit value of μ, and values obtained where each nuisance parameter has a pre-fit or post-fit uncertainty assigned to it i.e. $\hat{\theta} \pm x$, where x is equal to either the pre-fit or post-fit uncertainty. To assess the pre-fit impact of a nuisance parameter, the parameter is set with $x = \Delta\theta$, which is the pre-fit uncertainty of a parameter. In this case, predefined pre-fit uncertainties are only associated with the background cross sections. For other nuisance parameters, the pre-fit uncertainty is set to $\Delta\theta = 1$, where $\Delta\theta = 1$ describes the uncertainty of θ as used for constrained nuisance parameters that add a unit Gaussian term to the likelihood. To assess the post-fit impact of a nuisance parameter, the parameter is set with $x = \Delta\hat{\theta}$, which is the uncertainty on the estimated parameter. In Figure 5.20, the nuisance parameters are ranked by decreasing impact. The pre-fit impacts are shown by the the empty blue and cyan rectangles, and the post-fit impacts are shown by the filled rectangles. On the upper axis, the impact $\Delta\mu$ is shown, and on the lower axis the nuisance parameter pull, $\frac{(\hat{\theta}-\theta_0)}{\Delta\theta}$,is shown, where the pull is comparing the best-fit value to the nominal pre-fit value, θ_0. The nuisance parameter pulls are consistent with 0, as expected from a fit to a modified Asimov dataset.

According to Figure 5.20, the leading source of uncertainty is found to be the uncertainty due to the $t\bar{t}Z$ production cross section. This shows that the Max W_{had} GBDT score variable, which was constructed to lower the uncertainty due to the WZ production cross section, is effective, because the WZ cross section is no longer the leading source of uncertainty.

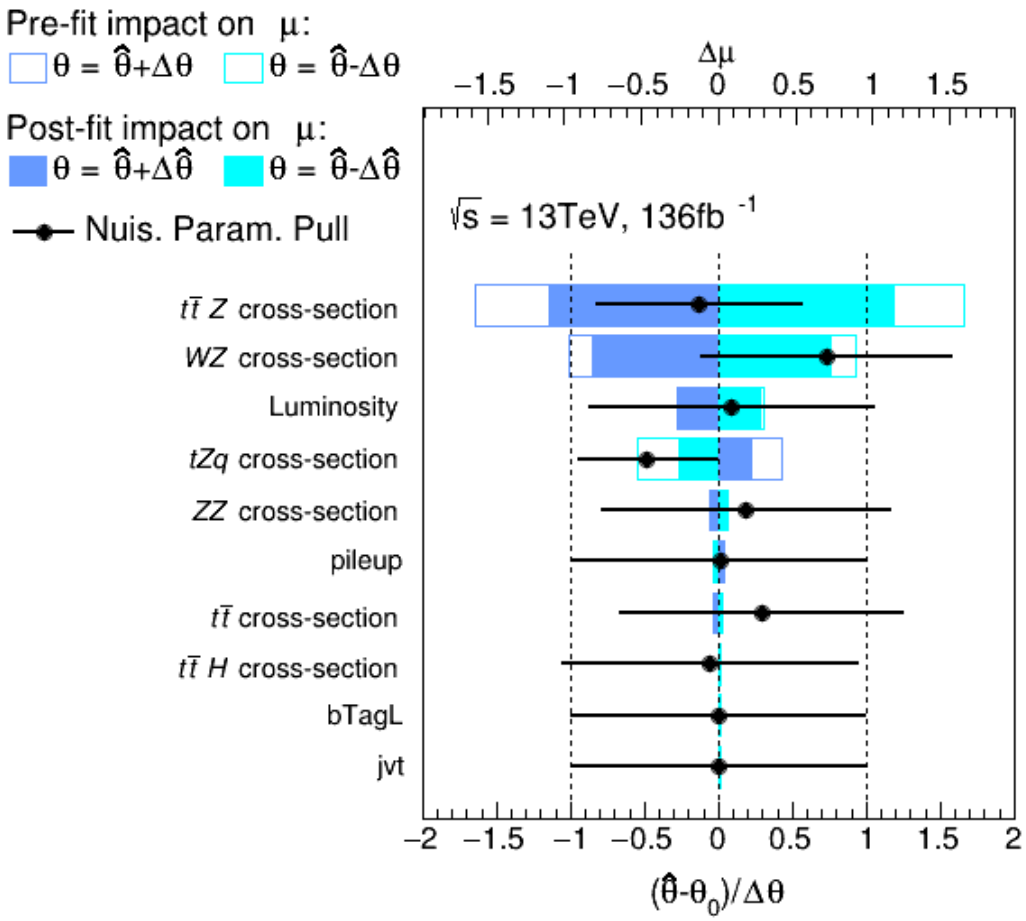
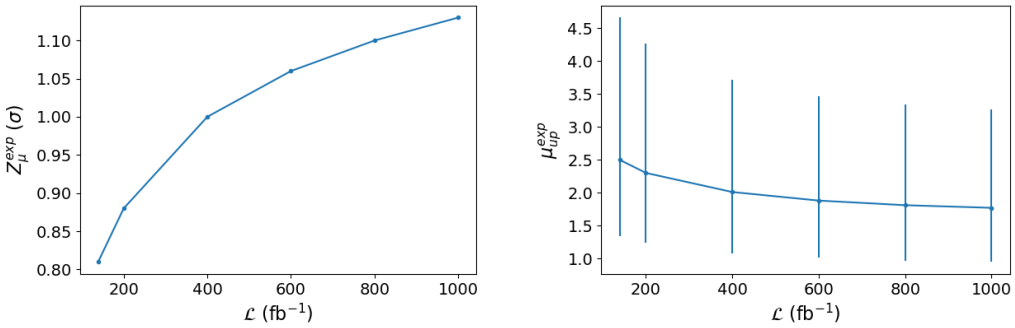


FIGURE 5.20: A ranking plot of the impact of the nuisance parameters is shown. The pre-fit and post-fit impacts are shown by empty and coloured rectangles, and the nuisance parameter pull values are shown by the black points. On the upper axis the impact, $\Delta\mu$, is shown and on the lower axis the nuisance parameter pull is shown.

5.4.3 Studying the expected significance and expected upper limit at higher luminosity

In Figure 5.21, plots of Z_μ^{exp} and μ_{up}^{exp} are shown for fits to Asimov datasets where the simulation has been scaled to a higher integrated luminosity, $\mathcal{L}$, for each fit. A modified Asimov dataset, like that used for the fit in this analysis, was not used since there is 136 fb^{-1} of data that could be used to determine the estimates of the nuisance parameters. Furthermore, the auto-binning algorithm was turned off to study the sensitivity at higher luminosity because this significantly altered the expected significance since this changed the optimal binning.

Fits to Asimov datasets with simulation scaled to higher luminosity were done in order to determine if this analysis is statistically limited. In Figure 5.21a, Z_μ^{exp} increases with increased luminosity, showing that a more accurate measurement could be achieved with more data. In Figure 5.21b, μ_{up}^{exp} and the uncertainties on μ_{up}^{exp} decrease with increased luminosity. This shows that the current analysis is strongly statistically limited.



(A) Expected significance (B) Expected upper limit

FIGURE 5.21: Plots of Z_μ^{exp} (a) and μ_{up}^{exp} (b) are shown for increasing $\mathcal{L}$

5.5 Summary and Outlook

In this chapter, the search for tWZ production in the trilepton channel was presented using 136 fb^{-1} of $\sqrt{s} = 13$ TeV pp collision data of Run 2 data from the ATLAS experiment. In a maximum likelihood fit to a modified Asimov dataset using an event GBDT score variable in the SR and CRs, a signal strength of $\hat{\mu} = 2.08^{+1.48}_{-1.45}$ was determined with a corresponding expected significance of $Z_\mu^{exp} = 0.72\,\sigma$. From the fit, an expected upper limit on μ of $\mu_{up}^{exp} = 2.77^{+2.39}_{-1.28}$ was also determined. These results are only preliminary as a comprehensive list of systematic uncertainties need to be added, and disagreements in CRs need to be investigated.

In a previous search for tWZ production in the trilepton channel [13], an expected upper limit on μ of $\mu_{up}^{exp} = 4.6^{+7.6}_{-3.3}$ was determined. Therefore, the current analysis has reduced the expected upper limit on the signal strength, meaning that one expects the current analysis to be able to put a tighter constraint on the the signal strength of tWZ production. This can be attributed to the reduced statistical uncertainties from the larger dataset and the better suppression of WZ production, where the latter is due to the inclusion of the Max W_{had} GBDT score variable developed for this analysis with the purpose of suppressing this background. Furthermore, the dominant systematic uncertainty in the analysis is now the uncertainty associated with the $t\bar{t}Z$ production cross section, as can be seen in Figure 5.20. Therefore, this analysis would benefit from a method to suppress this background. However, as shown in section 5.4.3, this analysis is strongly statistically limited and this is the dominant factor when determining the expected upper limit.

There are several areas that were not investigated in this analysis. The four lepton channel was not explored in this analysis and could offer a promising channel in the search for tWZ production as was shown in [14] where the trilepton and four lepton channels were combined to obtain a higher expected significance of approximately 4.5σ for comparable scales of luminosity. However, in this previous analysis, the cross section used for tWZ production was 227.9 fb which is approximately 40% larger than that used in this analysis (which is approximately 160.8 fb when the branching ratio of the leptonic decay of the Z is not used) and could be the reason for this higher expected significance. Furthermore, a data-driven estimation of the background due to fake leptons was beyond the scope of this analysis and could be the reason that there is poor agreement between simulation and data in some of the CR figures.

Chapter 6

Conclusion

The SM describes the current understanding of the fundamental components of matter and three of the four fundamental forces. It has thus far stood up to every experimental test to which it has been subjected. However, there are still many phenomena that are not described by the SM. The goal of ATLAS experiment at the LHC is to test the predictions of the SM in the hope of finding hints of BSM physics that could explain these phenomena.

In this dissertation, the search for tWZ production in the trilepton channel was presented using 136 fb^{-1} of $\sqrt{s} = 13$ TeV pp collision data taken from the Run 2 data of the ATLAS experiment. To this end, an overview of the SM and a motivation for this search were presented, an overview of GBDTs was provided, a description of the LHC and the ATLAS detector followed, and lastly, the details of this search were presented.

In a fit to a modified Asimov dataset using an Event GBDT score variable in the SR and CRs, a signal strength of $\hat{\mu} = 2.08^{+1.48}_{-1.45}$ was determined with a corresponding expected significance of $Z_{\mu}^{exp} = 0.72\,\sigma$. An expected upper limit on μ of $\mu_{up}^{exp} = 2.77^{+2.39}_{-1.28}$ was also determined. These values show that this analysis has the ability to put the tightest constraint yet on the cross section of tWZ production. However, these results are only preliminary as disagreements in the CRs need to be investigated further, and a comprehensive list of systematic uncertainties need to be included.

This analysis was strongly statistically limited, and the leading systematic uncertainty was due to the $t\bar{t}Z$ production cross section. Therefore, this analysis would benefit from more data in future runs of the LHC, and it would benefit from a method to distinguish between tWZ and $t\bar{t}Z$ production.

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". In: *Phys. Lett. B* 716 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

[2] S. Chatrchyan et al. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". In: *Physics Letters B* 716.1 (2012), 30–61. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2012.08.021. URL: http://dx.doi.org/10.1016/j.physletb.2012.08.021.

[3] Gian Francesco Giudice. "Naturally Speaking: The Naturalness Criterion and Physics at the LHC". In: *Perspectives on LHC Physics* (2008), 155–178. DOI: 10.1142/9789812779762_0010. URL: http://dx.doi.org/10.1142/9789812779762_0010.

[4] Vera C. Rubin. "One Hundred Years of Rotating Galaxies". In: *Publications of the Astronomical Society of the Pacific* 112.772 (2000), pp. 747–750. DOI: 10.1086/316573. URL: https://doi.org/10.1086/316573.

[5] Fabio Maltoni, Luca Mantani, and Ken Mimasu. "Top-quark electroweak interactions at high energy". In: *Journal of High Energy Physics* 2019.10 (2019). ISSN: 1029-8479. DOI: 10.1007/jhep10(2019)004. URL: http://dx.doi.org/10.1007/JHEP10(2019)004.

[6] Mark Thomson. *Modern particle physics*. New York: Cambridge University Press, 2013. ISBN: 978-1-107-03426-6.

[7] Sheldon L. Glashow. "Partial-symmetries of weak interactions". In: *Nuclear Physics* 22.4 (1961), pp. 579–588. ISSN: 0029-5582. DOI: https://doi.org/10.1016/0029-5582(61)90469-2. URL: https://www.sciencedirect.com/science/article/pii/0029558261904692.

[8] A. Salam and J.C. Ward. "Electromagnetic and weak interactions". In: *Physics Letters* 13.2 (1964), pp. 168–171. ISSN: 0031-9163. DOI: https://doi.org/10.1016/0031-9163(64)90711-5. URL: https://www.sciencedirect.com/science/article/pii/0031916364907115.

[9] Steven Weinberg. "A Model of Leptons". In: *Phys. Rev. Lett.* 19 (21 1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264. URL: https://link.aps.org/doi/10.1103/PhysRevLett.19.1264.

[10] Ramamurti Shankar. *Principles of Quantum Mechanics*. 2nd. Springer, 1994.

[11] F. Abe et al. "Observation of Top Quark Production in $\bar{p}p$ Collisions with the Collider Detector at Fermilab". In: *Physical Review Letters* 74.14 (1995), 2626–2631. ISSN: 1079-7114. DOI: 10.1103/physrevlett.74.2626. URL: http://dx.doi.org/10.1103/PhysRevLett.74.2626.

[12] S. Abachi et al. "Observation of the Top Quark". In: *Physical Review Letters* 74.14 (1995), 2632–2637. ISSN: 1079-7114. DOI: 10.1103/physrevlett.74.2632. URL: http://dx.doi.org/10.1103/PhysRevLett.74.2632.

[13] O. Ogul Oncel. "Search for Single Top Quark Production in Association with a W and a Z Boson in the 3 Lepton Final State with the ATLAS Experiment at 13 TeV". Presented 04 Jun 2018. 2018. URL: https://cds.cern.ch/record/2625170.

[14] Tommy Tschida. "Cross section measurements of processes with a single top quark and two vector bosons with the CMS experiment". MA thesis. Vienna, Tech. U., 2020. DOI: 10.34726/hss.2020.79420.

[15] Maria Moreno Llacer et al. *Modelling of rare top quark processes at $\sqrt{s} = 13$ TeV in ATLAS*. Tech. rep. ATL-COM-PHYS-2020-089. Geneva: CERN, 2020. URL: https://cds.cern.ch/record/2708924.

[16] *Standard Model Summary Plots Spring 2020*. Tech. rep. ATL-PHYS-PUB-2020-010. Geneva: CERN, 2020. URL: http://cds.cern.ch/record/2718937.

[17] T.M. Mitchell. *Machine Learning*. McGraw-Hill International Editions. McGraw-Hill, 1997. ISBN: 9780071154673. URL: https://books.google.co.za/books?id=EoYBngEACAAJ.

[18] Robert E. Schapire. "The Strength of Weak Learnability". English (US). In: *Machine Learning* 5.2 (June 1990), pp. 197–227. ISSN: 0885-6125. DOI: 10.1023/A:1022648800760.

[19] Mahsa Shoaran et al. "Energy-Efficient Classification for Resource-Constrained Biomedical Applications". In: *IEEE Journal on Emerging and Selected Topics in Circuits and Systems* PP (June 2018), pp. 1–1. DOI: 10.1109/JETCAS.2018.2844733.

[20] Tianqi Chen and Carlos Guestrin. "XGBoost". In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2016). DOI: 10.1145/2939672.2939785. URL: http://dx.doi.org/10.1145/2939672.2939785.

[21] Jerome H. Friedman. "Greedy function approximation: A gradient boosting machine." In: *The Annals of Statistics* 29.5 (2001), pp. 1189 –1232. DOI: 10.1214/aos/1013203451. URL: https://doi.org/10.1214/aos/1013203451.

[22] F. Pedregosa et al. "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.

[23] Oliver Sim Brüning et al. *LHC Design Report*. CERN Yellow Reports: Monographs. Geneva: CERN, 2004. DOI: 10.5170/CERN-2004-003-V-1. URL: http://cds.cern.ch/record/782076.

[24] *CERN Timeline*. URL: https://timeline.web.cern.ch/timeline-header/93 (visited on 01/27/2021).

[25] Michael Benedikt et al. *LHC Design Report*. CERN Yellow Reports: Monographs. Geneva: CERN, 2004. DOI: 10.5170/CERN-2004-003-V-3. URL: http://cds.cern.ch/record/823808.

[26] Esma Mobs. "The CERN accelerator complex - August 2018. Complexe des accélérateurs du CERN - Août 2018". In: (2018). General Photo. URL: https://cds.cern.ch/record/2636343.

[27] *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*. Tech. rep. ATLAS-CONF-2019-021. Geneva: CERN, 2019. URL: http://cds.cern.ch/record/2677054.

[28] The ATLAS Collaboration et al. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *Journal of Instrumentation* 3.08 (2008), S08003–S08003. DOI: 10.1088/1748-0221/3/08/s08003. URL: https://doi.org/10.1088/1748-0221/3/08/s08003.

[29] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. "The anti-k_t jet clustering algorithm". In: *Journal of High Energy Physics* 2008.04 (2008), 063–063. ISSN: 1029-8479. DOI: 10.1088/1126-6708/2008/04/063. URL: http://dx.doi.org/10.1088/1126-6708/2008/04/063.

[30] *Identification of Jets Containing b-Hadrons with Recurrent Neural Networks at the ATLAS Experiment*. Tech. rep. ATL-PHYS-PUB-2017-003. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2255226.

[31] G. Aad et al. "Search for new resonances in mass distributions of jet pairs using 139 fb^{-1} of pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector". In: *Journal of High Energy Physics* 2020.3 (2020). ISSN: 1029-8479. DOI: 10.1007/jhep03(2020)145. URL: http://dx.doi.org/10.1007/JHEP03(2020)145.

[32] S. Agostinelli et al. "Geant4—a simulation toolkit". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(03)01368-8. URL: https://www.sciencedirect.com/science/article/pii/S0168900203013688.

[33] J. Alwall et al. "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations". In: *Journal of High Energy Physics* 2014.7 (2014). ISSN: 1029-8479. DOI: 10.1007/jhep07(2014)079. URL: http://dx.doi.org/10.1007/JHEP07(2014)079.

[34] Enrico Bothmann et al. "Event generation with Sherpa 2.2". In: *SciPost Physics* 7.3 (2019). ISSN: 2542-4653. DOI: 10.21468/scipostphys.7.3.034. URL: http://dx.doi.org/10.21468/SciPostPhys.7.3.034.

[35] C. Oleari. "The POWHEG BOX". In: *Nuclear Physics B - Proceedings Supplements* 205-206 (2010), 36–41. ISSN: 0920-5632. DOI: 10.1016/j.nuclphysbps.2010.08.016. URL: http://dx.doi.org/10.1016/j.nuclphysbps.2010.08.016.

[36] Torbjörn Sjöstrand et al. "An introduction to PYTHIA 8.2". In: *Computer Physics Communications* 191 (2015), 159–177. ISSN: 0010-4655. DOI: 10.1016/j.cpc.2015.01.024. URL: http://dx.doi.org/10.1016/j.cpc.2015.01.024.

[37] Florian Fischer. *tWZ DR1/DR2 Monte Carlo samples*. Tech. rep. Ludwig Maximilian University of Munich, Dec. 2019.

[38] Thomas Philippe Calvet. "Search for the production of a Higgs boson in association with top quarks and decaying into a b-quark pair and b-jet identification with the ATLAS experiment at LHC". Presented 08 Nov 2017. 2017. URL: https://cds.cern.ch/record/2296985.

[39] G. Cowan. *Statistical Data Analysis*. Oxford science publications. Clarendon Press, 1998. ISBN: 9780198501558. URL: https://books.google.co.za/books?id=ff8ZyW0nlJAC.

[40] Glen Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *The European Physical Journal C* 71.2 (2011). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-011-1554-0. URL: http://dx.doi.org/10.1140/epjc/s10052-011-1554-0.

[41] Michele Pinamonti and Loic Valery. *TRExFitter*. URL: https://trexfitter-docs.web.cern.ch/trexfitter-docs/.

[42] Kyle Cranmer et al. *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*. Tech. rep. CERN-OPEN-2012-016. New York: New York U., 2012. URL: http://cds.cern.ch/record/1456844.

[43] Thomas Junk. "Confidence level computation for combining searches with small statistics". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 434.2-3 (1999), 435–443. ISSN: 0168-9002. DOI: 10.1016/s0168-9002(99)00498-2. URL: http://dx.doi.org/10.1016/S0168-9002(99)00498-2.