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

Ö. Oğul Öncel
Matrikelnummer 5089050

Universität zu Köln
Mathematisch-Naturwissenschaftliche Fakultät
Institut für Kernphysik

Erstgutachter: Prof. Dr. Jan Jolie
Zweitgutachter: PD Dr. Markus Cristinziani

Köln im Mai 2018

Abstract

The dominant mode of top-quark production in proton-proton collisions is the pair production. The single top quark production is an electroweak process with a small rate, providing an additional source for top-quark physics. Associated production of single top quarks are among the rare processes predicted by the Standard Model.

A search for associated production of a top quark with a W and a Z gauge boson in the trileptonic final state is presented in this thesis. The study has been done using the data collected with the ATLAS detector at the Large Hadron Collider during the years 2015 and 2016. This dataset corresponds to proton-proton collisions at a center-of-mass energy of $\sqrt{s}$ = 13 TeV and a total integrated luminosity of 36.1 fb^{-1}. Multivariate techniques are used to enhance signal and background separation.

An upper limit of 6.6 (Expected: $4.6^{+7.6}_{-3.3}$) times the Standard Model prediction on the cross section at the 95% confidence level has been set using the $\mathrm{CL_s}$ method. No significant deviation from the Standard Model has been observed.

Zusammenfassung

Top Quarks werden hauptsächlich via Top-Quark-Paarproduktion erzeugt. Seltener sind einzelne Top Quarks, welche als elektroschwacher Prozess innerhalb des Standardmodells mit einer kleinen Rate produziert werden. Trotzdem bieten diese eine weitere Möglichkeit das Top Quark zu studieren. Die vorliegende arbeit präsentiert eine Suche nach einzelnen Top Quarks in Assoziation mit einem W- und einem Z-Boson im trileptonischen Zerfallskanal. Die Studie nutzt Daten des ATLAS Detektors am Large Hadron Collider. Die Daten stammen aus den Jahren 2015 und 2016 und entsprechen Proton-Proton Kollisionen bei einer Schwerpunktsenergie von 13 TeV mit einer Luminosität von 36.1 fb^{-1}. Multivariate Methoden werden genutzt, um das Signal vom Untergrund möglichst effektiv zu unterscheiden. Ein Limit von 6.6 (Erwartetes Limit: $4.6^{+7.6}_{-3.3}$) konnte auf die Vorhersage des Standardmodells mit einem Konfidenzlevel von 95% gesetzt werden. Es wurde keine Abweichung vom Standardmodell gefunden werden.

Contents

CHAPTER 1

Introduction

Elementary particle physics aims to classify, understand and describe the most fundamental building blocks of the universe, their structure and interactions with each other at the smallest scales. During the 20th Century a theoretical framework has emerged thanks to advances made in theoretical models either guided or verified by experimental efforts. This model is called the Standard Model (SM) [1–3]. Although extremely successful in predicting a large range of natural phenomena to a high precision, the Standard Model fails to give a complete picture of nature. Some examples of unexplained phenomena in SM are gravity, the matter-antimatter asymmetry in the universe, neutrino mass and dark matter.

Various theoretical extensions to the Standard Model attempting to explain these unsolved problems have been proposed. These models are known as Beyond Standard Model theories. So far none of them have been experimentally verified. Experimental particle physics can search for direct evidence of predictions made by those theories. Another approach is to make precision measurements of the Standard Model predicted processes. A high enough precision might reveal a discrepancy between the Standard Model prediction and point to new physics.

The top quark, with its disproportionately high mass and short decay time below the hadronization scale, is a natural candidate for such searches. The most common production mode of top quarks is the pair-production, where a top and anti-top quark are produced together. Another production mode is single top production. Here a single top quark is produced in association with a quark, boson or a combination of them. In this thesis associated production of a single top quark with a W and a Z boson is studied. This is a rare process that has not been measured yet. The measurement has been done in the trilepton final state in order to suppress high rate hadronic background processes in the final states with few or no leptons. The tetralepton final state has not been used due to very low production rate of the tWZ process.

Such rare measurements are very challenging for they are often obscured by irreducible background contaminations mimicking the signal. To facilitate signal and background separation, machine learning methods are often applied. In this thesis an optimization study using Boosted Decision Trees has been conducted.

The organisation of chapters is as follows: In Chapter 2 a brief overview of the Standard Model of particle physics is given. As particle of interest, top quark and associated production channels are also described. Chapter 3 explains the experimental setup and instruments used to collect the data used in this thesis. The Large Hadron Collider, the ATLAS detector and its subunits are described. Since most of the analysis is done with simulated samples, a section on Monte-Carlo simulations is also included. In Chapter 4 the inputs and setup of the analysis are given. The event and object pre-selections are discussed. The selection of analysis regions is motivated and Boosted Decision Trees are introduced. Chapter 5 describes the studies done to improve the selections made in previous chapter. First a full event reconstruction for $tWZ \rightarrow 3\ell$ process is done. After that, a new set of variables, made possible by full event reconstruction has been used to train a Boosted Decision Tree. A further optimization study has been conducted on selected training parameters. Chapter 6 concludes the analysis description. First, the considered systematic uncertainties are introduced. Then a fit is performed to extract the exclusion upper limit. Chapter 7 summarizes the work done and discusses possible ways of improvement for the future.

CHAPTER 2

Theoretical Overview

In this chapter theoretical knowledge that is used in this work is briefly described. Firstly a general introduction to the Standard Model of particle physics is given where the particles and their interactions are explained. In the second part, properties of top quark and associated production with two heavy gauge bosons studied in this work are introduced.

In elementary particle physics, energies are often quoted in terms of electron volts [4]. Due to very high energy scales typical values are of the order of MeV(10^6), GeV(10^9) or TeV(10^{12}). Momenta are measured in GeV/c and masses are in GeV/c^2. In this thesis natural units in which $\hbar = c = 1$ are used [5]. Thus, masses and momenta are quoted in GeV. Barn(b) is the common unit used in particle physics to denote the cross-sectional area. 1 barn is equal to 10^{-28}m^2 [4]. The typical cross-sectional areas encountered in current elementary particle physics measurements are in the range of picobarn(10^{-12}) and femtobarn(10^{-15}).

2.1 The Standard Model of Particle Physics

The Standard Model [6] is a renormalizable quantum field theory which obeys the principle of local gauge invariance. It is the current best theoretical framework that categorizes and describes the elementary particles and two out of three fundamental forces characterizing their interactions, excluding gravity [7]. These two forces are the electroweak and strong force.

The particles in the SM are divided into two categories dictated by their spin-statistics. The half-spin particles obeying the Pauli exclusion principle are called fermions and integer-spin particles are called bosons.

Fermions are further divided into two groups depending on their interaction with the strong force. Those that can interact with the strong force are called quarks and those that cannot are leptons, therefore quarks are able to interact with all of the forces.

Quarks come in six different types, or flavours. They are grouped into three generations. A generation is composed of two quarks having fractional electric charges: Q=$\frac{2}{3}$ also known as up-type and Q=-$\frac{1}{3}$ as down-type quark [4]. The inter-generational hierarchy is ordered through particle masses with 3$^{\text{rd}}$ generation having the highest mass. Therefore the natural sequence of decays proceed from higher generations to lighter ones, creating a disproportional abundance of lower generation quarks. Each quark has also its own anti-particle, differing just by having the negative of some of the original particle's quantum numbers. Quarks can combine to form two types of composite bound states that are called mesons and baryons. Mesons are composed of one quark and one anti-quark whereas baryons are composed of three quarks or three anti-quarks. Mesons and baryons are collectively named hadrons.

The group of fermions that do not participate in the strong interaction are called leptons. Leptons come in two types: charged and uncharged. Each charged lepton has a corresponding uncharged partner which is its neutrino. Due to their lack of charge and mass neutrinos are among the hardest of all elementary particles to detect. A charged lepton and its neutrino together make up a generation. There are 6 leptons in 3 generations making up a total of 12 leptons, when considering also their corresponding antiparticles.

Interactions between particles are mediated by spin-1 gauge bosons. There exist three force-carriers for the electroweak force: The massless and chargeless photon, the chargeless but massive Z, charged and massive $W^{\pm}$. The gauge boson associated with the strong force is the gluon. The "charge" of strong force is called "color" and unlike photon which does not carry electric charge, gluons do carry colour charge and therefore self-interact.

The strength of each interaction is determined through its coupling constant, which is a function of the energy scale at which interactions happen. An important phenomenon related to couplings is the *quark confinement*. In the strong interaction the coupling constant characterizes the interaction of quarks and gluons. The coupling constant is small at high energies or short distances. But at low energies or large distances it increases. Thus, for very short distances such as within the hadron radius, quarks roam almost free due to extremely weak interaction coupling, this is called *asymptotic freedom*. However, when the distance increases the coupling becomes stronger and stronger, eventually creating an energy high enough for new quarks to be created. These quarks immediately bond and form hadrons. This process is called hadronization. Due to the hadronization, in particle detectors quarks and gluons are observed through collimated hadronic particle showers. Such showers are called jets.

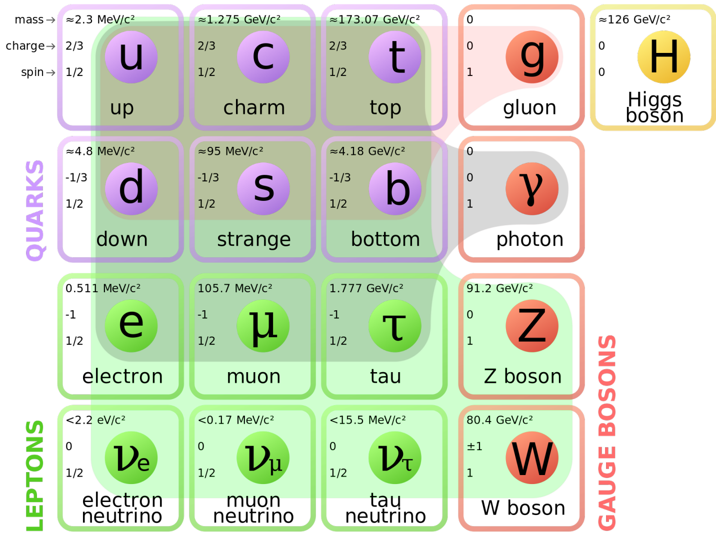


Figure 2.1: Chart showinging the Standard Model particles categorized according to their types. Illustration taken from [8].

.

The only spin-0 particle in the SM is the Higgs boson. The Higgs boson is understood as the carrier for the Higgs field, whose interaction strength with particles are proportional to the particles' masses. The so-called Higgs mechanism [9–11] allows for the presence of mass for the particles in the SM, explaining how these particles acquire their mass. In Figure 2.1 particles in SM are shown, categorized according to their types.

2.2 The Top Quark

The top quark is the heaviest known elementary particle with a mass of approximately 173 ± 0.4 GeV [7]. It was discovered in 1995 at Fermilab in USA by the CDF and DØ collaborations [12, 13] . Because of its large mass, the top quark is unstable and has a very fast decay time. The typical top-quark decay time is around 5×10^{-25} seconds, which is an order of magnitude faster than the hadronization timescale of 3×10^{-24} seconds. Because of this the top quark provides a unique opportunity to directly probe the properties of quarks through its decay products.

Top quarks almost always decay to a W boson and a b-quark. This makes understanding of b-quarks critical for the study of top quarks. Bottom quarks are the second heaviest quarks therefore they also decay fast, but still hadronize. However, b-hadrons have relatively long lifetimes compared to other hadrons of their mass scale due to suppressed weak decay channels. In particle detectors, b-hadrons emerging from a b-quark travel a certain distance before they decay.

Top quarks most commonly emerge in pairs. Top anti-top quark pairs ($t\bar{t}$) can be produced via gluon fusion or quark annihilation. The secondary mode of top-quark production is the single top quark production. Single top quark production always happens in association with quarks or bosons through electroweak interaction. Associated production of top quark is interesting to measure the weak coupling of the top quark. It was first discovered in association with a b-quark, in 2009 at Fermilab by the CDF and DØ collaborations [14]. Some examples of these processes are $t\bar{t}Z$ [15], $t\bar{t}W$ [15], tW [16] and tZ [17]. A chart of SM production cross section measurements of Standard Model processes is given in Figure 2.2.

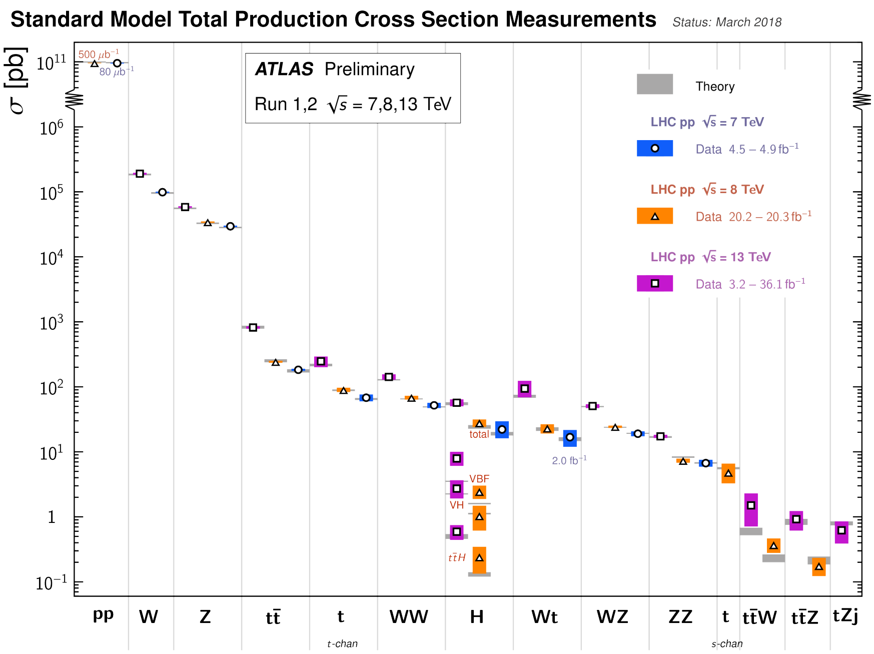


Figure 2.2: Status of SM cross section measurements by the ATLAS Collaboration as of March 2018 [18].

2.3 Associated Production of Top Quark with a W and a Z Boson

In this thesis associated production of a single top quark with a W and a Z boson, the tWZ process is studied. Some Feynman diagrams of tWZ are shown in Figure 2.3. No measurement of this process has been performed so far. Due to similar final state particles, all previously mentioned associated production modes are natural backgrounds to tWZ production and vice versa. Thus, a better understanding of the tWZ process is expected to contribute to improvement of several related analyses.

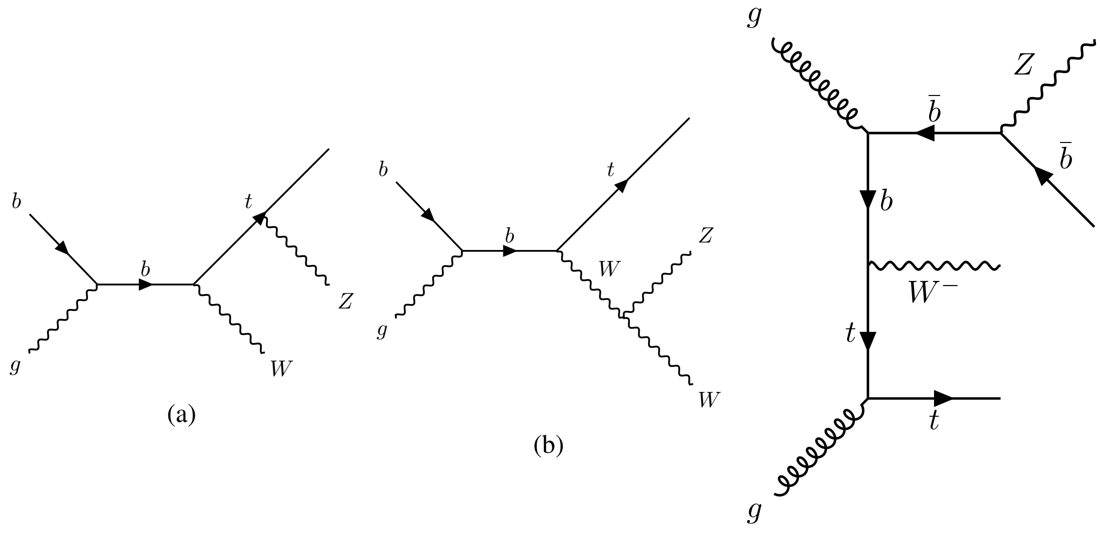


Figure 2.3: Example Feynman diagrams for tWZ production.

In particle physics measurements hadronic multi-jet backgrounds often dominate. To cope with this, measurements are required to have a certain number of leptons in the final states, suppressing the major jet-decay modes of particles. In this study, the channel with exactly three charged leptons in the final state is used. These charged leptons can either be electrons or muons i.e $\ell = e, \mu$. Tau leptons either decay hadronically and are reconstructed as jets or they can decay leptonically and reconstructed as leptons. This special treatment of taus are due to the short lifetime of the order of 10^{-13} seconds, making their identification difficult [19].

The dileptonic channel, despite its higher production rate, is not chosen because of previously explained hadronic backgrounds. The tetralepton channel on the other hand is not preferred because of its small cross section. The cross sections at Next-to-Leading-Order (NLO) for different decay modes are shown in Table 2.1. The total cross section has been derived from simulation sample since currently there is no theoretical calculation available for the tWZ process.

tWZ@NLO	Inclusive	$Z \to \ell^+\ell^-$	2ℓ	3ℓ	4ℓ
σ(fb)	150	10.5	5.9	3.9	0.7

Table 2.1: Branching fractions of tWZ process for different multileptonic final states in NLO precision. Numbers are rounded.

In $tWZ \to 3\ell$ signal process considered in this thesis, Z boson is only allowed to decay to two charged leptons. There are two W bosons, one emerging from the production vertex and one decaying from the top quark. One of these two W bosons can decay leptonically, producing the remaining charged lepton and a neutrino. The hadronically decaying W boson decays to two quarks. These two W boson decay modes are complementary. Thus, following final state particles are allowed:

- $Z\to \ell^+\ell^-$ where $\ell = e, \mu$.
- Top$\to$ Wb where W can decay leptonically i.e. $W \to \ell\nu$ or hadronically i.e. $W \to q\bar{q}^{'}$.
- W decaying leptonically i.e. $W \to \ell\nu$ or hadronically i.e. $W \to q\bar{q}^{'}$.

CHAPTER 3

The ATLAS Experiment

ATLAS is one of the four high energy proton collision experiments within the European Organization for Nuclear Research, known as CERN (*fr. Conseil Européen pour la Recherche Nucléaire*), which was the name of the initial committee that founded the organisation. CERN is located around the Franco-Swiss border near the city of Geneva. Each of these experiments are associated with a detector that is collecting data provided by the collisions created with the *Large Hadron Collider* (LHC), a circular particle accelerator. The abbreviation ATLAS stands for ***A T**oroidal **LHC A**pparatu**S***. The name ATLAS also refers to titan Atlas in ancient greek mythology. According to this myth, after the olympian gods defeated the titans in the war, Zeus punished Atlas with the eternal duty of carrying the sky on his shoulders. This posture of Atlas carrying a globe over his shoulders had been a common theme in classical sculpture and reflected in the logo of the ATLAS experiment.

The construction of the ATLAS detector was completed in 2008 and its performance is measured to be very close to original design targets. Together with the Compact Muon Solenoid (CMS) it is one of the two general-purpose detectors served by LHC, enabling both experiments to validate and cross-check measurements made by one another. On July 4 2012, a joint statement by both collaborations announced [20] the discovery of the long anticipated Higgs boson, completing the list of particles predicted by the SM.

This chapter provides information about the LHC, the ATLAS detector, how the collected information is processed, and important quantitites used in the physics analyses.

3.1 Large Hadron Collider

The LHC is currently the most powerful particle accelerator in the world. It has a 27 km circumference and consists of two separate and vacuum beam pipes that run parallel to each other. Protons are accelerated in these beam pipes as groups of discrete packages that are called *bunches*. Bunches have 7 meters of space between each other and contain more than 100 billion protons each. Current technology allows for a maximum number of 2556 bunches to be circulated simultaneously.

The two beam pipes intersect at four *bunch crossing point*s, where the detectors are located, allowing protons circulating in opposite directions to collide. The bunch structure means that protons cannot be brought into collision one at a time. During the collision of billions of protons in two proton beams, multiple collisions occur in the detector region. This is called *pile-up*. Pile-up is undesired for they make identifying the particles belonging to the same collision challenging.

Maximizing the bunch crossings will increase the pile-up rate but are important for having higher

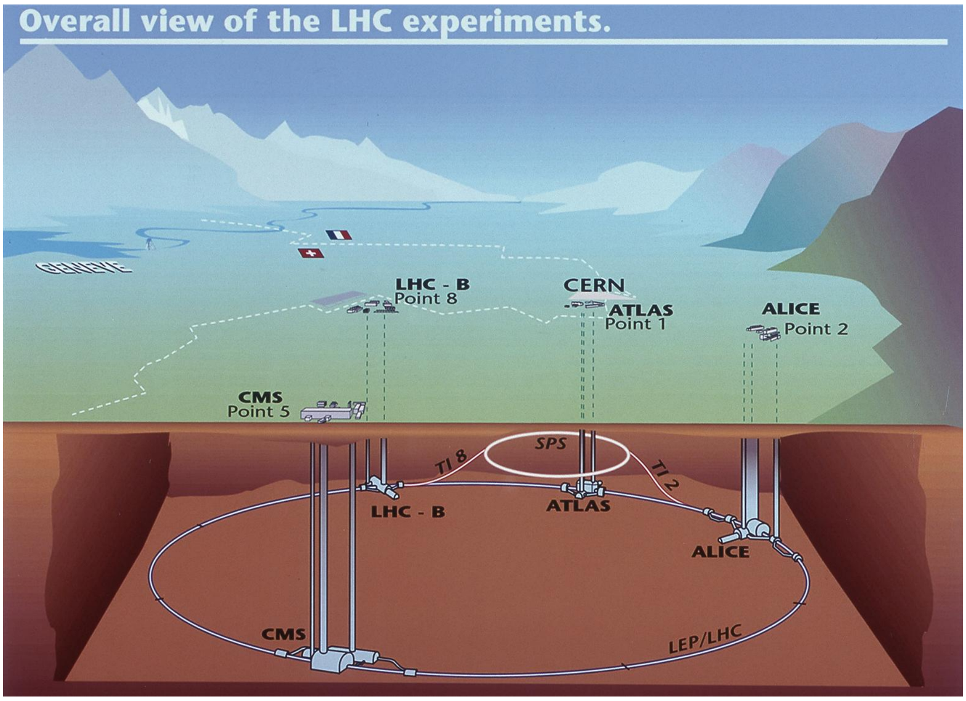


Figure 3.1: Schematic display of the LHC collider and detectors. Copyright: CERN.

collision rates. The number of proton-proton crossings per unit area per unit time is called *luminosity*:

$$\mathcal{L} = \frac{N_{pp}}{A \cdot t}. \tag{3.1}$$

The luminosity is independent of the process of interest, and is a property of the collider. The LHC has a maximum design luminosity of $\mathcal{L} = 10^{34}\text{cm}^{-2}\text{s}^{-1}$. The expected number of events for a specific process in a collider is calculated by multiplying the production cross-section of that process with the Luminosity.

$$\mathcal{L} \,.\, \sigma_x = \frac{N(pp \to X)}{t}. \tag{3.2}$$

In experimental particle physics one often works with data collected over a period of time. The luminosity for that total amount of time is called total integrated luminosity:

$$\int \mathcal{L}\, dt = L. \tag{3.3}$$

Another important quantity in the characterization of a particle accelerator is the total available centre-of-mass energy. The superiority of particle colliders lie in this number, because in fixed target experiments the final centre-of-mass energy available is proportional to the square root of the initial energy. But since colliding particles are already in the centre-of-mass frame, the available energy is the direct sum of the energies of the two particles. The trade-off is a reduced collision rate. The LHC has a design centre-of-mass energy of 14 TeV. Currently the highest centre-of-mass energy reached is 13 TeV.

At high energies, the composite nature of protons cannot be ignored. Protons are made up of quarks and gluons, collectively named partons. During the collisions interactions take place between the partons. The acceleration of partons, in analogy with bremsstrahlung due to acceleration of photons, can cause gluon radiation. If a parton radiates a gluon before the interaction it is called an Initial State Radiation (ISR). If the radiation occurs after the interaction of partons it is called a Final State Radiation (FSR).

3.2 ATLAS Detector

3.2.1 Topology

The ATLAS detector has a cylindrical shape with a length of 44 meters and a diameter of 25 meters weighing around 7000 tonnes [21]. A schematic view of the detector is shown in Figure 3.2. ATLAS has a right-handed coordinate system. The origin is defined as the collision point. The x-axis points towards the accelerator center and the y-axis points upwards. The z-axis is parallel to the beam pipe, which passes through the central axis of the detector. Common coordinates used when working with ATLAS data are cylindrical coordinates. Here, ϕ is the azimuthal angle measured in the transverse plane to the beam axis. Instead of the polar angle θ, a Lorentz invariant transformation, called pseudo-rapidity η, is used, where

$$\eta = -\ln \tan\frac{\theta}{2}. \tag{3.4}$$

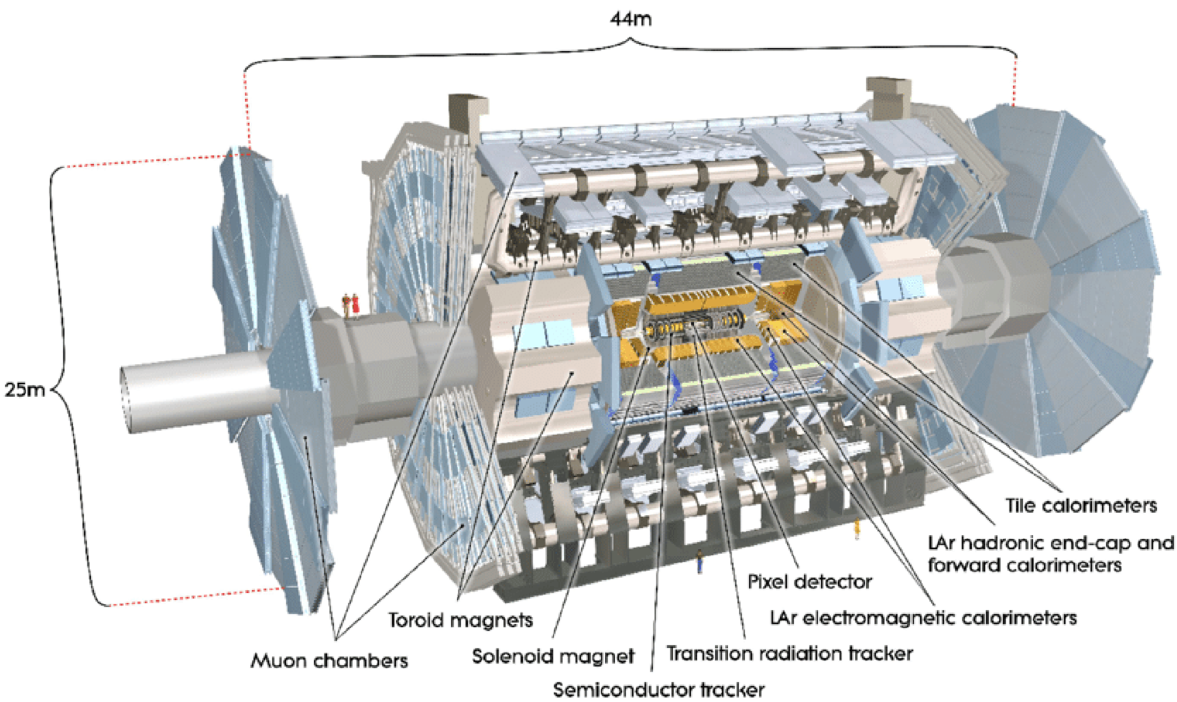


Figure 3.2: A schematic view of the ATLAS detector with separate detector units identified. Copyright: CERN.

The distance between two objects with coordinates (η_1, ϕ_1) and (η_2, ϕ_2) is defined as,

$$\Delta R = \sqrt{(\eta_1 - \eta_2)^2 + (\phi_1 - \phi_2)^2}. \tag{3.5}$$

An important quantity that is defined in the transverse plane is the transverse component of momentum, which is symbolized as p_{T}.

3.2.2 Detector Subunits

The beam pipe is enclosed with concentric layers of different detector units to help detecting different particles with high efficiency. There is a cylindrical barrel part and endcaps that allow detection up to higher $|\eta|$.

Inner Detector

The inner detector [22] is the first layer enclosing the beam pipe. It consists of three parts: The pixel detector, the semi-conductor tracker and the transition radiation tracker. In total the inner detector occupies the space up to a radius of 1.2 m starting few centimeters from the beam axis. The main function of the inner detector is to detect, measure the momenta, and reconstruct the charged particle tracks. For the charge and momentum measurement, the inner detector is equipped with a solenoidal magnetic field with field lines parallel to the beam axis. Because of that, charged particles will bend in the detector and their charge and momentum can be inferred from the direction and degree of bending respectively. For the track reconstruction, the hits created by energy deposited in the layers of the detector material are used.

Calorimeters

The calorimeters occupy the middle space between the muon spectrometer and inner detector. They consist of an electromagnetic calorimeter (ECAL) which is closer to the beam axis and a hadronic calorimeter (HCAL) on the outer part. Both calorimeters detect particles via absorbing their energies.

The ECAL is able to detect electromagnetic showers created by photons or electrons, that lose their energies as they interact with the detector material. Electrons create these showers through bremsstrahlung and photons through pair production.

Hadrons pass through the ECAL with minimal energy deposition in the calorimeter. They are detected by the HCAL, as they create hadronic particle showers and deposit their energies. Hadronic particle showers take longer to develop relative to electromagnetic showers, and because of that, the HCAL occupies a larger volume. The HCAL is not as precise as ECAL in particle detection.

Muon Spectrometer

Muons are able to pass through the inner detector and calorimeters. To detect them the outermost layer of the ATLAS detector is equipped with the so-called muon spectrometer. Muons can be detected with very high efficiency because they are the only particles detected in the muon spectrometer. The muon spectrometer also uses a toroidal magnetic field to determine momenta and charge information of muons.

Particle Detection

Through one or combination of the above units; electrons, muons, photons, charged and neutral hadrons are observable in the ATLAS detector. The detector response to these particles are shown in Figure 3.3. Short lifetime renders detection of many other particles such as quarks, W and Z bosons impossible. Those can be indirectly observed through their long-lived decay products. This is called reconstruction.

Neutrinos are not detectable at the ATLAS detector because they interact very weakly. They are quantified via amount of missing transverse energy. Missing transverse energy is defined through total transverse momentum imbalance in the detector and represented with the symbol $E_{\mathrm{T}}^{\mathrm{miss}}$. The initial transverse momentum in the detector is zero. After the collision, the vectorial sum of all particles do not

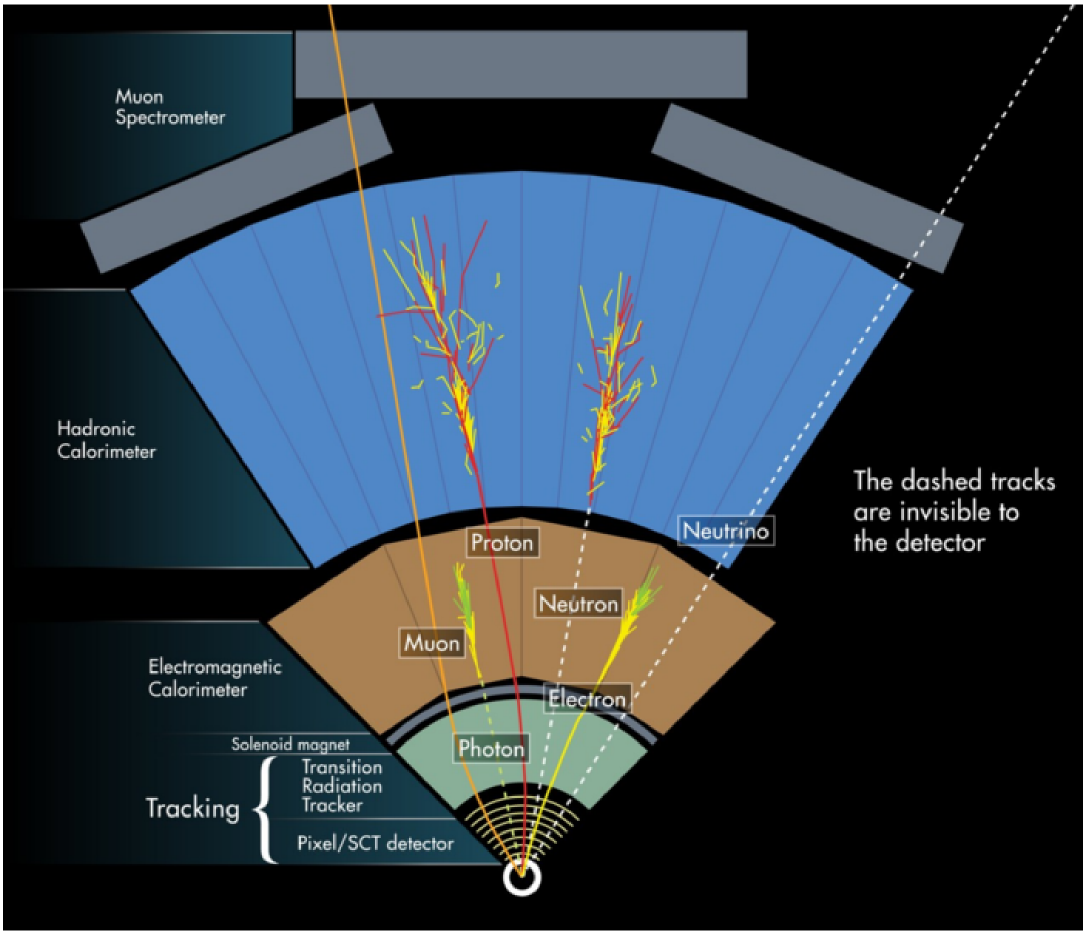


Figure 3.3: Schematic view of the ATLAS detector showing how particles interact with different detector units. Copyright: CERN.

sum up to zero, seemingly violating the conservation of momentum. This missing amount is taken out of the detector volume through undetected neutrinos. Thus, $E_{\mathrm{T}}^{\mathrm{miss}}$ is an indirect measure of neutrinos.

Bottom quarks are central to top quark physics analyses for they almost always decay to a b-quark, as discussed in Section 2.2. To classify whether a jet is emerging from a bottom quark or not is an important subfield in itself. This identification of b-jets is called *b-tagging*. The tool used in this thesis is MV2c10 algorithm [23], which exploits the characteristic decay time and invariant mass of b-hadrons for the identification of jets originating from b-quarks. b-quark is the second heaviest quark after the top-quark. Thus, it also decays relatively faster but slow enough to hadronize. The resulting b-hadron formed has a decay time of around 10^{-12} seconds. This is a time interval such that the b-hadron travels few millimeters before it initiates hadron showering. Because of that, the reconstruction of its tracks allows for the discrimination between other jets which emerge from the *primary vertex* (PV) i.e collision point and b-jets coming from the *secondary vertex* (SV) which is a displaced vertex relative to the primary vertex. This principle of b-tagging is illustrated in Figure 3.4.

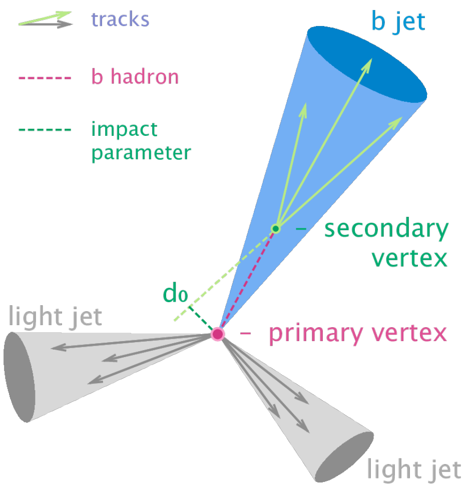


Figure 3.4: Figure demonstrating the principle of b-tagging based on discriminating between the primary vertex and the displaced secondary vertex, using the characteristic b-hadron lifetime. Illustration taken from [24].

3.2.3 Trigger and Data Acquisition

The large amount of data accumulated during the LHC runs makes it impossible to store all information. The data is filtered in several steps, so as to remove uninteresting and common processes, while passing possible candidate events that might signal at new physics. ATLAS uses three levels of filtering to succeed in that. These levels are called *triggers*. The Level-1 trigger uses a limited set of information from the calorimeters and the muon spectrometer to decide whether to pass the event to the next trigger or not, within two microseconds. Less than 100 000 events from 40 million bunch crossings per second pass the first level. Level-2 triggers employ more information from detectors to decide if event should be passed to last stage. Few thousands events per second pass Level-2. Final Level-3 trigger uses full detector information to analyse the event and on average 1000 events per second pass this final stage and are recorded for offline analyses.

3.3 Monte Carlo Simulations

Monte-Carlo (MC) simulations are a way to model physics processes and their interaction with the detectors. A particle process is modeled by random sampling to approximate its probability distribution function.

Since a process can only be observed through some detector simulated particle process cannot be directly compared to measured data. The detector response also needs to be simulated. Detector simulation takes the final state particles produced in the process simulation as input and simulates each units' response to these particles. These include known engineering defects or inefficiencies such as cracks, or post-detection processing of digital data.

Contrary to real data, in a MC simulation the sequence of all particle decays including their parent and children particles can be traced and recorded together with their four-vectors. This type of information is called *MC truth information*. Truth information can be used to investigate the accuracy of particle reconstruction in analyses or to understand the processes better such as finding the origin of a certain particle of interest.

CHAPTER 4

Analysis Setup and Strategy

This chapter describes the tools and approach used in the analysis done in the next chapter. In the first section the dataset and the background processes are introduced. In the second and third sections object definitions and event pre-selection are discussed. In the fourth section the analysis strategy and approach are explained.

4.1 Backgrounds and Dataset

4.1.1 Backgrounds

The simulation samples used in this analysis are summarized in the Table 4.1 and detailed sample information is provided on Appendix B.2. Given that events with three leptons are selected, in this study following background processes are considered:

- Z + jets/γ: Two separate background processes are included for the associated Z boson production. In the Z +jets sample Z bosons are produced together with jets. In the Z +γ process photons instead of jets are considered. In both cases the Z boson decays to two charged leptons.
- Diboson: WZ and ZZ diboson processes are considered.
- Top-quark pair production: Top-quark pair production ($t\bar{t}$) background is considered in dileptonic final states. Thus, both top quarks decay semi-leptonically.
- Associated production of $t\bar{t}$ with bosons: Different associated production modes of $t\bar{t}$ process with one or two bosons are included. These are associated production with Z boson ($t\bar{t}Z$), with one W boson ($t\bar{t}W$), with two W bosons ($t\bar{t}WW$) and associated production with a Higgs boson ($t\bar{t}H$).
- Single Top: The associated production of top quark with a Z boson, tZ is used. The tW sample has not been considered due to a very small branching fraction in the trileptonic channel, when requiring a Z boson.
- Triboson: The Triboson sample used consists of the WWW, WWZ, WZZ and ZZZ processes.

The two most important backgrounds with the highest contribution in this analysis are the WZ and $t\bar{t}Z$ processes. The WZ process has a large production cross section and already possess two particles that are also present in the signal process. Through ISR or FSR jets can form a final state composition similar to the signal. An example of this is when ISR causes a gluon split into a $b\bar{b}$ pair and another one to light jets ending up in a top-quark reconstruction. Due to the low probability of b-quark production and the low b-tagging misidentification rate, the WZ importance diminishes at high b-tagged jet multiplicities.

The $t\bar{t}Z$ process has a lower cross-section than WZ. However, it is more able to mimic the signal sample for its final states can exactly overlap with tWZ. Since $t\bar{t}Z$ has two b-jets, it is expected to be more dominant in high b-tagged jet multiplicity regions in contrast to WZ. Still, it can end up in the lower b-tagged jet regions, for instance through a missed b-jet.

Other backgrounds can also end up in the signal regions even when they don't normally possess three leptons. Non-prompt leptons refer to any leptons not resulting from the decay of a W or Z boson. This can be a mis-reconstructed jet mimicking the signature of a lepton, or a lepton emitted during hadronization or from photon conversion. The B-hadron being able to decay to leptons, and as this analysis focus on processes rich in b-quark production, the dominant source for non-prompt leptons is B-hadrons.

Process	Event Generator	Parton Shower
tWZ	aMC@NLO	Herwig
WZ	Sherpa	Sherpa
ZZ	Sherpa	Sherpa
Z + jets	Sherpa	Sherpa
$Z+\gamma$	Sherpa	Sherpa
tZ	MadGraph	PYTHIA8
$t\bar{t}$	Powheg	PYTHIA8
Triboson	Sherpa	Sherpa
$t\bar{t}H$	aMC@NLO	PYTHIA8
$t\bar{t}Z$	aMC@NLO	PYTHIA8
$t\bar{t}W$	aMC@NLO	PYTHIA8
$t\bar{t}WW$	aMC@NLO	PYTHIA8

Table 4.1: Processes considered in this study and their respective event and parton shower generators.

4.1.2 Dataset

The period of data taking at LHC during the years 2015–2018 is called Run 2. Data events used in this analysis are collected with the ATLAS detector during the first two years (2015–2016) of Run 2. Before data can be used in analyses it needs to be centrally processed and calibrated. Since this takes time, analyses of data from certain period commonly follows with some latency.

The LHC provided proton-proton collisions at a centre-of-mass energy of 13 TeV. Not all data events detected are used in physics analyses. A set of conditions are required for the quality of data to be ensured. These include successful operation of the machines, healthy performance of detector units and beam. The final set of accepted data is separated and named as *Good Run List*. A total integrated luminosity of 36 fb^{-1} has been reached in the 2015 - 2016 Good Run List. Total integrated luminosity delivered by the LHC, collected by the ATLAS detector and the amount good for physics during the Run 2 are shown in Figure 4.1.

The datasets are further categorized in so-called *period*s each labeled with a capital letter. This categorization is used to refer to different settings on operating conditions used for beams, machines and triggers at each run. These periods are listed in Table 4.2.

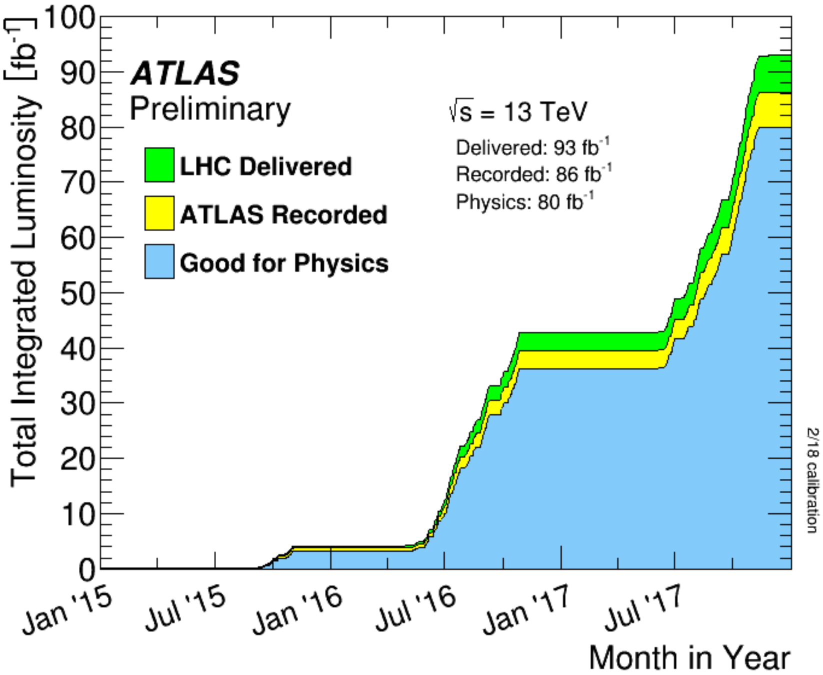


Figure 4.1: The total integrated luminosity collected by the ATLAS detector during th Run 2. In this thesis data collected during 2015-2016 period amounting to a total of 36 fb^{-1} are used [25].

Year	Periods
2015	D, E, F, G, H, J
2016	A, B, C, D, E, F, G, I, K, L

Table 4.2: Years and periods corresponding to collection of data used in this thesis.

4.2 Inputs

4.2.1 Objects

Leptons

In the ATLAS experiment there are three levels of lepton identification: *Loose*, *Medium* and *Tight*. These refer to how strict the identification requirements are. The loose selection gives the highest number of particles identified as leptons but has the lowest purity, since there will also be mis-identified particles, because less strict criteria are applied. On the contrary, the tight selection gives the smallest number of leptons, but they will have the highest purity thanks to strict selection requirements that needs to be passed. In this thesis the tight selection is used for electrons and medium selection is used for muons.

A multivariate discriminant exploiting the track, distance and energy information to reject non-prompt leptons has been developed within the ATLAS collaboration [26]. The studies done with the samples used in this thesis shows that this discriminant is especially effective in removing $t\bar{t}$ background. It is also observed that non-prompt rejection removes signal events in a proportion such that neither a gain nor a loss in the final sensitivity of analysis occurs. However, due to poor modeling of non-prompt objects it is preferred to have the same sensitivity with less non-prompt objects. Therefore the non-prompt rejection is kept. 75% reduction in $t\bar{t}$ and Z+jets, and 20% loss in signal yields are recorded relative to the case without non-prompt rejection discriminant.

Jets

Jets are reconstructed using a jet algorithm on the energy deposit clusters of the calorimeters. The jet algorithm used is the anti-k_{t} algorithm [27] with a radius parameter $R = 0.4$. Jet candidates are required to have $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$. To separate jets coming from the event of interest from the pile-up background jets, the *Jet Vertex Tagger* (JVT) discriminant is applied [28]. As recommended by ATLAS, JVT ≤ 0.59 is required when a jet candidate has $p_{\mathrm{T}} < 60$ GeV and $|\eta| < 2.5$.

b-Tagging

The 77% efficiency working point for b-jets to be tagged has been used, based on the MV2c10 multivariate classifier [23], corresponding to a requirement on the weight to be larger than 0.645.

Missing Transverse Momentum

Missing energy is calculated based on the MetRefFinal algorithm [29]. When a track is detected but not associated with any of the reconstructed events it is called a soft term and considered in the algorithm for the $E_{\mathrm{T}}^{\mathrm{miss}}$ detection.

4.2.2 Event Cleaning

- Trigger matching: The event has to have fired a single-lepton trigger. Additionally, events where none of the fired lepton trigger are corresponding to the selected leptons are excluded.
- Overlap removal between Z+jets and $Z+\gamma$: When the Z+jets and $Z+\gamma$ processes are used together, the case for double counting has to be addressed. The Z+jets process contains partially the $Z+\gamma$ process. To avoid this Z+jets events with a photon in it have been removed from the event selection [30].
- Overlap removal between tWZ and $t\bar{t}Z$: There is an overlap between $t\bar{t}Z$ and tWZ production. In order to avoid this, the signal sample produced with the so-called Diagram Removal 1 (DR1) is used. In DR1, the Feynman diagrams of $t\bar{t}Z$ production ending up in the tWZ final state are removed from the Monte Carlo sample [31].

4.3 Kinematic Pre-Selection Cuts

Z Boson Veto

The tWZ process contains one on-shell Z boson. The Z boson is among the most efficiently detected and reconstructed particles, often referred to as a SM candle. Therefore it is common to use properties of Z bosons to define otherwise strict cuts with more confidence. For a particle to be considered as an on-shell Z boson candidate two conditions are required to be simultaneously fullfilled:

- An opposite-sign and same-flavor (OSSF) lepton pair to exist: e^+e^- or $\mu^+\mu^-$;
- The invariant mass of this lepton pair to be within a mass window of 10 GeV centered at the Z boson pole mass of 91.2 GeV: $|m_{\ell^+\ell^-} - 91.2\ \mathrm{GeV}| < 10$ GeV

In the trilepton selection two OSSF pairs are possible. In this case a metric needs to be defined to choose which pair to assign. Two common ways are to use distance and invariant mass information. For the distance method, the lepton pair having the minimum distance ΔR is chosen. In the invariant mass method the lepton pair giving the invariant mass closest to the pole mass of the Z boson is chosen. A check has been made on the signal sample by comparing both methods to MC truth information. The mass method is found to be a better reconstruction metric with 96% of reconstructed particles within ΔR<0.4 of truth particles, compared to 76% in the distance method.

Low Mass Resonance Veto

In order to suppress quarkonium resonances occurring at low invariant mass a veto is applied on any OSSF lepton pairs with an invariant mass below 15 GeV, i.e. $m_{\ell^+\ell^-} > 15$ GeV for any such pair.

Cut on Non-Z Lepton p_T

The remaining lepton, after the assignment of two leptons to the Z boson will be called the non-Z lepton. The transverse momentum distribution of this lepton is shown in Figure 4.2. An accumulation of backgrounds other than WZ and $t\bar{t}Z$, categorized under the name "Others", are observed at the low values. This can be explained by non-prompt leptons emerging from the high cross-section Z+jets process or by prompt leptons from the four-lepton ZZ process when one of the lepton is not reconstructed. Therefore a cut at 20 GeV for the non-Z lepton p_T is applied to remove a substantial amount of background without losing more than 10% of signal events.

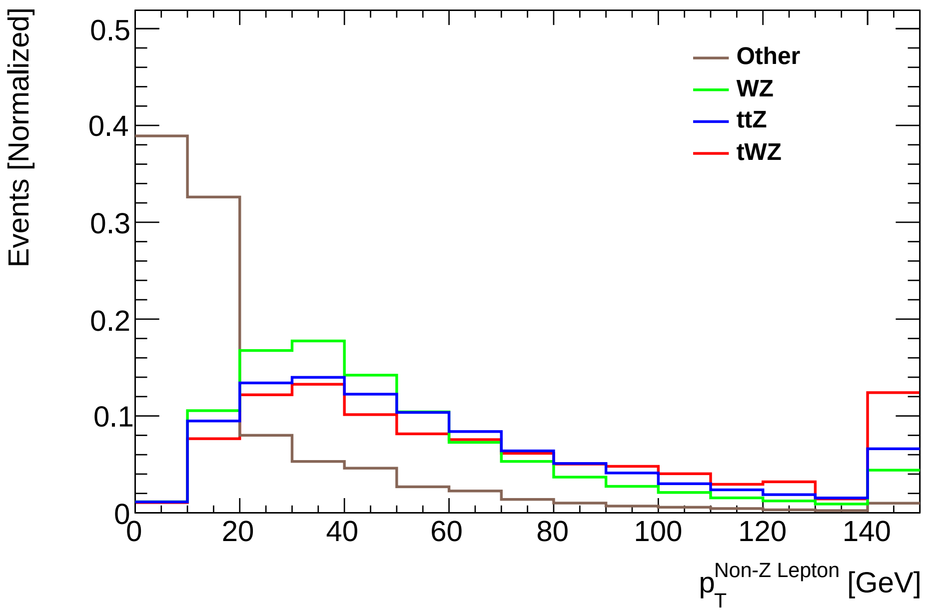


Figure 4.2: p_T of the non-Z lepton in the pre-selection region. Last bin contains the overflow.

4.4 Strategy

In particle physics analysis, the expected statistical significance of a measurement is commonly approximated [7] by $\sigma = \frac{S}{\sqrt{S+B}}$ where S stands for the number of signal events and B is the number of total background events.

To claim an evidence the particle physics community requires 3σ and for discovery a 5σ is required. When an analysis is not sensitive enough for either, than an upper limit can be set.

To increase the signal sensitivity, it is necessary to make selections that define a region which is high in signal fraction. Such regions are called *Signal Regions*. A critical aspect during the inspection of signal regions is to avoid unconscious human bias. One might be tempted to tune the selections, variables and other parameters so as to bring data and prediction to an agreement with previously expected outcomes. To avoid this, an analysis is done *blinded*. In a blinded analysis, the comparison between data and expectation in the signal region is not seen until the final stage of analysis, *unblinding*, after which results cannot be changed.

4.4.1 Regions and Yields

Since unblinding is done at a late step of an analysis, any mis-modeling between data and simulation cannot be fixed a posteriori. Thus, one needs handles on the modeling of the backgrounds processes, variables and regions defined by cuts. Such checks can be made on specific signal-depleted regions, where the signal process makes a very low fraction of the total (typically up to 5%), thus not biasing the experimenter. These regions are called *validation regions*. The closer the selection of validation region to the signal region, the better it is for the modeling will be more similar. Clearly, by definition there are differences between these two regions for one maximizes and the other minimizes the fraction of signal process.

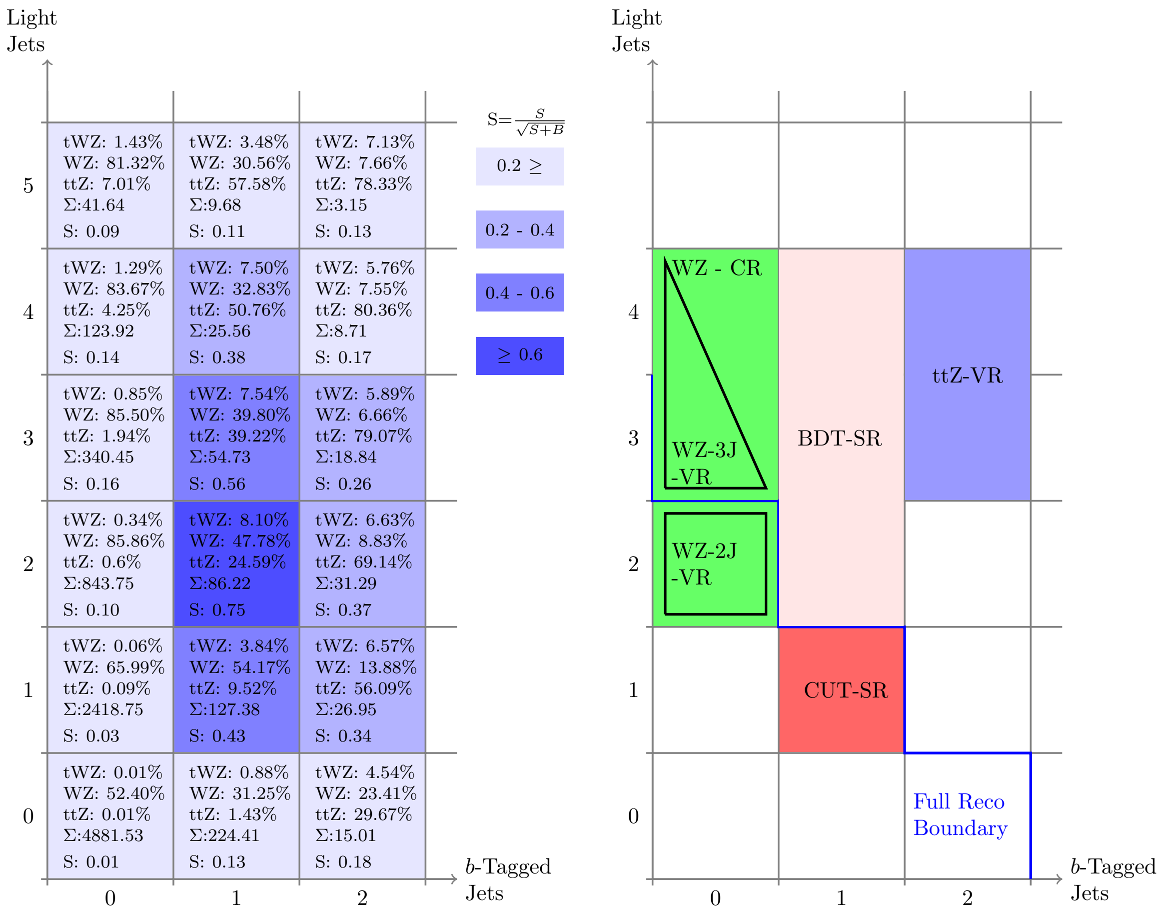
Another type of validation region is the *control region*, where one aims to create a region ideally dominated by a single background enabling one to investigate the match between data and simulation for that specific background. If the signal region is dominated by a certain background, that background's modeling will account for a large fraction of modeling. Control regions can also be used to constrain the background targeted, in the fit; if they have a large enough number of events and purity.

Due to their determining role in the event selections and topologies, jet multiplicities are commonly used in the definition of regions. In the following, any jet that is not b-tagged is referred as a light jet and denoted by the symbol "j". b-tagged jets are represented by the symbol "b". With these labels any orthogonal region based on jet multiplicity selections can be referred with their light jet and b-tagged jet numbers. For example, the 2j1b region refers to a selection of 2 light jets and 1 b-tagged jet.

In figure 4.3(a) the yields for signal and the two leading backgrounds ($t\bar{t}Z$ and WZ) as well as the total yield (Σ) and expected significance (S) for each orthogonal region defined by b-tagged (x-axis) and light jet-multiplicities (y-axis) are given. The regions with more than 5 light jets or 2 b-jets are not considered due to low yields. As expected, for the 3ℓ channel the most sensitive region is the 2j1b region. There are also significant fraction of signal events in the 1j1b region where a jet is missed by the detector.

For higher light jet multiplicities, in the 1 b-tagged jet region, again, an increased fraction of signal events are expected, inversely proportional to jet multiplicity.

In the zero b-tagged jet region a domination of WZ process is seen in all jet multiplicities. In the two b-tagged jet region $t\bar{t}Z$ process is dominant as expected, with its two b-quarks coming from top-quark pair production. There is also a relatively high percentage of signal events which can be understood via the Feynman diagrams at Figure 2.3(c). It is strongest in the 2j2b region. However, it should be noted



(a) Yields (in percent) for signal and background, total number of events Σ and significance S as a measure for discovery potential, are presented for each orthogonal region based on their b-tagged and light jet multiplicities.

(b) Four regions and the full reconstruction boundary are presented for each orthogonal region based on their b-tagged and light jet multiplicities. CR stands for control region, SR for signal region and VR for validation region.

Figure 4.3: Yields and definition of regions as described in Section 4.4.1 are visualized in the above graphics.

that event yields in these regions are low. This short survey indicates that it is not possible to isolate a high fraction of signal.

Variables other than jet multiplicities can also be used to reduce the backgrounds and gain further sensitivity, as in the case of the non-Z lepton cut in the pre-selection. A full event reconstruction in which all particles in the event are reconstructed; provides many variables such as angles, distances and kinematic relations between the particles. The $tWZ \rightarrow 3\ell$ channel event reconstruction assumes the presence of one top quark, one Z boson and one W boson. The Z boson reconstruction is always possible in any given jet-multiplicity and already discussed in Section 4.3 . The reconstruction of the top quark and the W boson requires at least 3 jets, one of which is b-tagged. If the calibrated b-tag working point is not used, one can reconstruct the event only based on ≥3 jets requirement and with identifying the jet with the highest MV2c10 score as the ‘b-jet’. This enables one to extrapolate event reconstruction to the zero b-tagged jet region whenever at least 3 jets are present. In a similar manner, when there are two b-tagged jets, the one with the higher MV2c10 score can be used as the b-tagged jet and the second one can be treated as a non-b-tagged jet. The border of full event reconstruction availability is illustrated in Figure 4.3(b) with a blue line labeled as “full reco boundary”.

Given the above examination, a WZ control region with 0 b-tagged jets and 2-4 light jets is defined. Two signal regions are determined in the 1 b-tagged jet region, separated by the full reconstruction boundary so that the full event reconstructed case can be separately validated. The first signal region where the full event reconstruction can be done is defined for 2-4 light jets spanning the highly sensitive region. This selection will be further optimized through a Boosted Decision Tree (BDT) training, which is described in the next section and implemented in the next chapter. Thus, this region is named BDT-SR. The second signal region is defined with 1 light jet and 1 b-tagged jet and named CUT-SR. Two validation regions are defined in the 0 b-tagged region to measure the modeling of the WZ background, one below and one above the full reconstruction boundary for the respective signal regions. The VR used for the modeling checks of CUT-SR is defined with 2 light jets and named WZ-2J-VR. The second one used for BDT-SR is defined with 3-4 light jets and named WZ-3J-VR. To check the behaviour of the $t\bar{t}Z$ background for the BDT-SR, another validation region is defined in the 2 b-tagged jet region with 3-4 light jets, named ttZ-VR. However, this region has more than 5% signal contamination, i.e above the previously defined limit. Because of that, the $t\bar{t}Z$ validation region will also be subject to further selection in the next chapter. In conclusion, all regions, except BDT-SR and ttZ-VR, are finalized at this point. The regions are also illustrated in Figure 4.3(b). The yields and other quantitites of interest are given in Table 4.3.

Process	WZ-CR	WZ-2J-VR	WZ-3J-VR	BDT-SR	CUT-SR	ttZ-VR
tWZ	7.33 ± 2.89	2.84 ± 1.15	4.48 ± 1.78	13.02 ± 4.93	4.89 ± 1.77	1.61 ± 0.75
WZ	1120 ± 288	724 ± 182	394 ± 107	71.36 ± 39.58	69.04 ± 39.33	1.91 ± 1.08
$t\bar{t}Z$	17.05 ± 3.25	5.10 ± 0.84	11.85 ± 2.53	55.63 ± 11.52	12.12 ± 1.88	21.90 ± 5.86
tZ	8.89 ± 1.34	6.52 ± 0.95	2.36 ± 0.43	9.69 ± 1.03	20.6 ± 2.30	0.93 ± 0.15
Others	156.2 ± 19.4	104.8 ± 14.3	50.90 ± 6.20	16.78 ± 2.01	20.74 ± 4.63	1.19 ± 0.23
Total	1310 ± 290	843 ± 183	464.4 ± 108.4	166.5 ± 41.8	127.4 ± 40.1	27.54 ± 6.10
Data	1275	842	433	Blinded	Blinded	33
$\frac{S}{\sqrt{S+B}}$	<0.20	<0.20	<0.20	1.00	0.43	0.30
Data/MC	0.97	0.99	0.93	Blinded	Blinded	1.19
Signal %	<1%	<1%	<1%	8.0%	6.0%	5.8%
Leading Background %	85%	85%	85%	42%	85%	79%

Table 4.3: The expected number of events in the different regions defined for signal and background. Uncertainties include systematic and statistical uncertainties. Numbers are rounded.

4.4.2 Boosted Decision Trees

Machine learning tools are used to build multivariate classifiers that can create a single strong discriminator from many weaker ones, in order to separate signal from the background in a classification problem. In this thesis the TMVA (Toolkit for Multivariate Analysis)[32] computational package has been used. This section largely follows from Reference [32]. A decision tree is built through sequential binary cut splits. Each split point is called a node and at each node a decision is made based on a single cut on a single variable, with the goal of separating the background events from signal events. A schematic illustration of a decision tree is shown in Figure 4.4.

This variable and cut is chosen according to the maximum separation among all variable and cut combinations possible at that node. the maximum separation is defined as the separation gain between

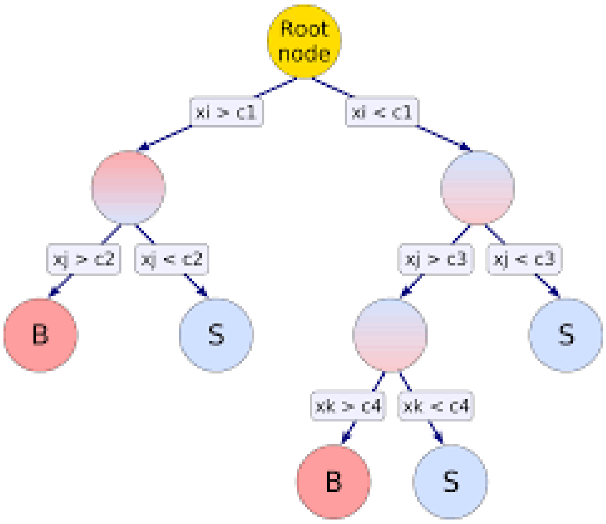


Figure 4.4: Schematic description of a decision tree. Starting from the root node, a series of cuts is applied classifying events as Signal or Background at the final node, known as leaf. Illustration taken from [32].

the parent node and the sum of children nodes, where gain is defined as $G = p \cdot (1 - p)$ with p being the purity.

Boosting is the process of learning for the classifier. In this study the *Adaptive Boosting* algorithm is used. Boosting requires multiple decision trees to be trained. The first tree is trained with the original event weights. The mis-classified events are re-weighted and given higher weights so that the training will focus more on those. The re-weighting factor, α is common and derived from the mis-classification rate, *err* of the previous tree.

$$\alpha = \frac{1 - \mathrm{err}}{\mathrm{err}} \tag{4.1}$$

The weights of the entire event sample are then renormalized such that the sum of weights remains constant.

The final output of an individual classifier can be defined as $h(x)$ with x being the collection of input variables. h will be equal to +1 for signal and -1 for the background. The boosted event classification $y(x)$ is then given by

$$y(x) = \frac{1}{\mathrm{N_{collection}}} \Sigma_i^{\mathrm{N_{collection}}} \ln(\alpha_i) \cdot h_i(x) \tag{4.2}$$

where the sum is over all classifiers. If the classifier is close to -1 an event is classified more background like. If it is closer to +1 it classified more signal-like.

Training Parameters

The TMVA package provides many configuration options. Among them, the most commonly used five parameters have been chosen for optimization. The abbreviations specified in the parenthesis next to each parameter will be used later.

- Number of trees (Trees): This parameter determines how many trees will be trained, that is the number of iterations.

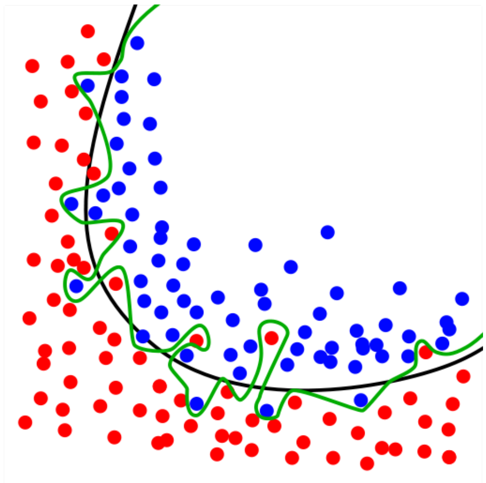


Figure 4.5: An example of overtraining. The classifier attempts at a separation of red and blue events. The black curve is an optimal cut. The green drawing shows an overtrained classifier, having learned individual fluctuations of individual events. Illustration taken from [33].

- Maximum depth (Max Depth): This option controls the growth length of the tree. It provides the maximum number of allowed node splits a tree can have.
- Minimum percentage of events (nMin%): This parameter puts a lower bound on the minimum percentage of training events that a leaf node has to have. If the next split makes the event content of the leaf drop below the specified value, then it is terminated.
- Learning rate (LearnR): This parameter determines the rate of adjustment of weights, which means how steep the learning is happening.
- Number of cuts (nCuts): This parameter determines the scanning granularity of the distribution of the variables used in the training.

Receiver-Operator-Characteristic (ROC) Curve

ROC curves are commonly used graphics for evaluating the classification performance of BDTs. A ROC curve is a plot whose x-axis gives the signal efficiency rate and y-axis giving the background rejection rate, both normalized to 1. The point (0,1) describes a situation where all backgrounds are rejected successfully but with no correct signal classification. The other extreme point (1,0) is the case where all signal events are classified correctly but no backgrounds are rejected. The optimum point that maximizes this trade-off between effectively selecting the signal and rejecting the background gives the best performance for the BDT training. This is quantified through the Area Under the Curve (AUC), which in this thesis is referred to as ROC Integral. A ROC integral of 0.5 corresponding to a diagonal means no separation power beyond a random outcome. On the other hand a ROC integral of 1.0 means a perfect classification.

Overtraining

When working with BDTs, one has to check against *overtraining*. Overtraining happens when a BDT training starts to learn individual fluctuations in a sample. An illustration of overtraining is shown in

Figure 4.5. This makes BDT's performance dependent on specific events in a sample and reducing its capacity for generalization to a statistically independent sample, without loss of classification power. Because of that, overtraining might give a false sense of increased classification capability. A test against overtraining can be done by splitting the samples into two and training on one half and testing the trained classifier output by applying it on the other half. Since both subsets belong to the same distribution, the BDT should in principle, perform similarly on both. In this thesis this method has been implemented by dividing samples into half, based on their event numbers being even or odd. A training on odd numbered events are tested on the even numbered events and a training done with even numbered events are tested on the odd numbered events. Ideally, a training should not perform better than a testing output. This information is used to quantify the overtraining as relative performance of training to testing sample.

$$\text{Overtraining} = \frac{(\text{ROC}_{\text{Even}}^{\text{Training}} + \text{ROC}_{\text{Odd}}^{\text{Training}}) - (\text{ROC}_{\text{Even}}^{\text{Test}} + \text{ROC}_{\text{Odd}}^{\text{Test}})}{\text{ROC}_{\text{Even}}^{\text{Training}} + \text{ROC}_{\text{Odd}}^{\text{Training}}} \tag{4.3}$$

Using this equation, Overtraining ≤ 0 corresponds to no overtraining. There is no established rule on how much overtraining is unacceptable. Generally the final decisions are made by inspecting the distributions of the training and testing response of the BDTs. From the studies made during the preparation of this work, it has been decided to set Overtraining $= 0.003$ as an upper limit corresponding to the point beyond which the distributions start to look characteristically different for training and testing samples. It should however be noted that this estimate is rather liberal and still allows for some degree of overtraining. Motivations behind BDT parametrization choices are discussed on Chapter 7.

Performance

As described in the previous subsection, the classification efficiency will almost always be higher during the training, relative to testing. Testing results are therefore expected to be more representative of the actual classification capability. Because of that, testing ROC integrals are used to quantify the relative performance of BDTs. The Performance is defined as the arithmetic average of testing ROC integrals on even and odd tests.

$$\text{Performance} = \frac{\text{ROC}_{\text{Even}}^{\text{Test}} + \text{ROC}_{\text{Odd}}^{\text{Test}}}{2} \tag{4.4}$$

CHAPTER 5

Search for tWZ Production

This chapter elaborates on the analysis approach described in the previous chapter. The first part describes the full event reconstruction for the tWZ process in the trilepton channel. The second part describes modified BDT-SR and ttZ-VR regions, including results of optimization and selection studies done with the BDTs for the former.

5.1 $tWZ \to 3\ell$ Event Reconstruction

The $tWZ \to 3\ell$ channel event reconstruction assumes the presence of one top quark decaying into Wb, one Z boson decaying into dileptons and one other W boson. The number of leptons dictate that only one of the two W bosons can decay leptonically. The other W boson will then decay via the hadronic channel. This complementary nature of two W bosons is illustrated in Figure 5.1, showing a Feynman diagram for the $tWZ \to 3\ell$ final states.

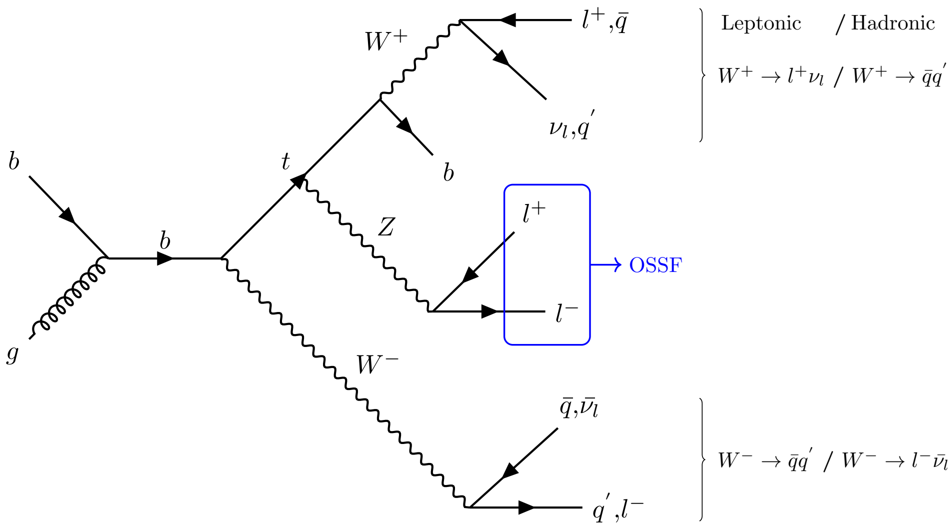


Figure 5.1: Feynman diagram of tWZ decay modes described in this chapter. OSSF stands for Opposite-Sign Same-Flavor, describing the criteria used in the reconstruction of the dileptonically decaying Z boson candidate, which is the first particle to be determined in the full event reconstruction.

At the beginning of the full event reconstruction, two groups of particles are distinguished, based on their dependence on the number of jets. Two particles can always be reconstructed because they don't decay to jets. The dileptonically decaying Z boson and leptonically decaying W boson are independent of the jet multiplicity. However, it is not possible to directly construct the four-vector of the neutrino, which needs to be combined with the non-Z lepton four-vector, because $E_{\mathrm{T}}^{\mathrm{miss}}$ does not have a z-component. In the following a method to reconstruct the z-component of the neutrino is implemented based on Reference [34].

Neutrino Reconstruction

The z-axis component of the neutrino can be estimated by constraining the mass of its parent particle, the W boson, to its on-shell mass of $M_W^2 = (80.4\ \mathrm{GeV})^2$. This is not the current best mass but a convention used to make different analyses comparable. Similarly for top quark 172.5 GeV is used. Having fixed one degree of freedom, an expression for the z-component of $E_{\mathrm{T}}^{\mathrm{miss}}$ can be found [34].

$$p_{z,\nu}^{\pm} = \frac{\mu p_{z,\ell}}{p_{T,\ell}} \pm \sqrt{\frac{\mu^2 p_{z,\ell}^2}{p_{T,\ell}^4} - \frac{E_\ell^2 p_{T,\nu}^2 - \mu^2}{p_{T,\ell}^2}} \tag{5.1}$$

$$\text{with } \mu = \frac{M_W^2}{2} + p_{T,\ell} p_{T,\nu} \cos(\Delta\phi) \tag{5.2}$$

where $\Delta\phi$ is the angle between $p_{T,\ell}$ and $p_{T,\nu}$. Equation 5.1 can yield to two real solutions when radicand is positive or imaginary solutions when the radicand is negative. The two cases are treated differently.

- Imaginary solutions: Imaginary solutions can be understood through a re-formulation of the radicand in Equation 5.1.

  $$\frac{\mu^2 p_{z,\ell}^2}{p_{T,\ell}^4} - \frac{E_\ell^2 p_{T,\nu}^2 - \mu^2}{p_{T,\ell}^2} < 0 \tag{5.3}$$

  Multiplying the equation by $p_{T,\ell}^4$ gives

  $$\mu^2 (p_{z,\ell}^2 + p_{T,\ell}^2) - E_\ell^2 p_{T,\ell}^2 p_{T,\nu}^2 < 0 \tag{5.4}$$

  Substituting $p_{z,\ell}^2 + p_{T,\ell}^2$ with E_ℓ^2 and simplifying reduces the inequality to

  $$\mu^2 < p_{T,\ell}^2 p_{T,\nu}^2. \tag{5.5}$$

  Replacing μ with 5.2 back and taking the square root gives the following relation that can be physically interpreted:

  $$M_W < \sqrt{2 p_{T,\ell} p_{T,\nu} (1 - \cos(\Delta\phi))}. \tag{5.6}$$

  The right-hand side of Equation 5.6 is the transverse mass of the W boson. The imaginary solutions for the z-component of the neutrino thus occur when the transverse mass of the W boson exceeds its pole mass. In this treatment it is assumed that reconstructed masses exceeding the pole mass are due to the limited resolution of $E_{\mathrm{T}}^{\mathrm{miss}}$, since lepton reconstruction has a better resolution. Thus, the x and y components of $E_{\mathrm{T}}^{\mathrm{miss}}$ will be modified minimally so as to make the square root term zero. The smallest modification possible to ensure a real z-component is when $m_W^T = M_W = 80.4$ GeV. Substituting this value into Equation 5.6 results in a quadratic relation between $p_{x,\nu}$ and $p_{y,\nu}$.

$$p^{\pm}_{y,\nu}(p_{x,\nu}) = \frac{M_W^2 p_{y,\ell} + 2p_{x,\ell}p_{y,\ell}p_{x,\nu}}{2p_{x,\ell}^2} \pm \frac{M_W p_{T,\ell}}{2p_{x,\ell}^2}\sqrt{M_W^2 + 4p_{x,\ell}p_{x,\nu}} \tag{5.7}$$

To search for minimum possible deviation from the originally measured components of $E_{\mathrm{T}}^{\mathrm{miss}}$, the distance between modified and measured values are defined. Since $p_{y,\nu}$ is parametrized by $p_{x,\nu}$ the two distances $d^{\pm}$ depend only on the $p_{x,\nu}$.

$$d^{\pm}(p_{x,\nu}) = \sqrt{(p_{x,\nu} - E_{\mathrm{T},x}^{\mathrm{miss}})^2 + (p^{\pm}_{y,\nu}(p_{x,\nu}) - E_{\mathrm{T},y}^{\mathrm{miss}})^2} \tag{5.8}$$

An iterative minimization process is implemented using the BrentMinimizer1D minimization method in the Root package to find the $p^{\pm}_{y,\nu}$ corresponding to the minimum distance $d(p_{x,\nu})$. The distance with the smallest absolute value is then chosen.

- Real solutions: In contrast to the imaginary case there are two real solutions for $p^{\pm}_{z,\nu}$. A check with MC truth information of the signal sample has been done in order to find out which choice leads to a better match with the leptonically decaying W boson. The solution with smaller absolute value is found to be more accurate as illustrated in Figure 5.2. This is qualitatively in agreement with findings based on $t\bar{t}$ in the study done in reference [34].

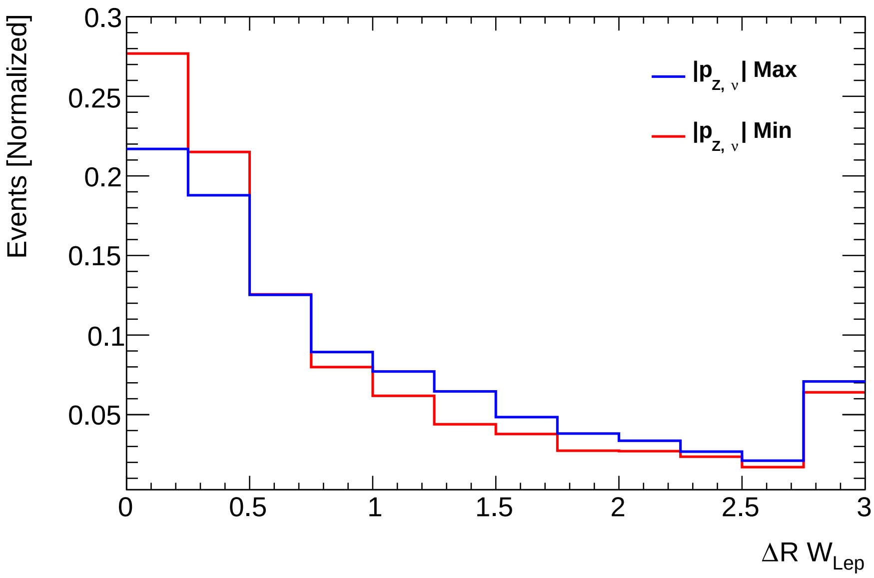


Figure 5.2: Distance between generator level and reconstructed level particle for the two cases of $p_{z,\nu}$ selection. The last bin contains overflow.

The second type of particles subject to reconstruction are the top quark and the hadronically decaying W boson. At this point no b-tagging information is used. First, all possible top-quark reconstruction combinations are considered to find the one giving the invariant mass closest to the on-shell top mass of 172.5 GeV. One semi-leptonically decaying top candidate is possible for each jet combining with W_{Lep}. For hadronically decaying top quarks, all possible different combinations of three jets are used, two for W_{Had} and one for the b-jet.

Among all these top candidates the one with the invariant mass closest to 172.5 GeV is chosen. In order to categorize two W bosons, the W boson decaying from the top quark is referred to as W_{Best} and the W boson not decaying from the top quark is labeled as W_{Other}.

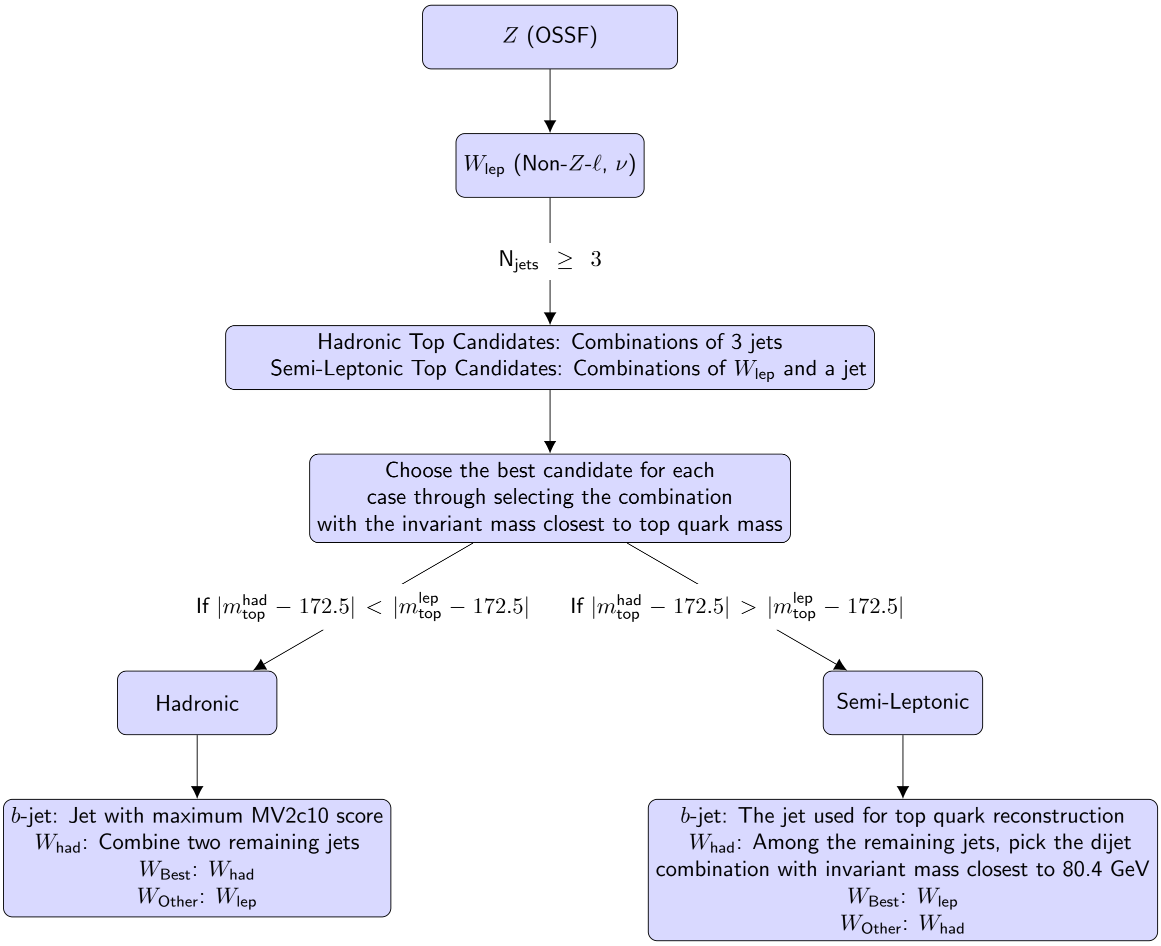


Figure 5.3: Event reconstruction flowchart.

If the top quark decays semi-leptonically, W_{Lep} becomes W_{Best} and the jet combined with W_{Lep} to make the top-quark candidate with the mass closest to top quark mass is chosen as the 'b-jet'. The remaining jets are combined again and the combination that gives an invariant mass closest to the W pole mass of 80.4 GeV is used to reconstruct the hadronically decaying W boson which becomes W_{Other}.

If the top quark decays hadronically, among the three jets used to reconstruct the top quark, the one with the highest MV2c10 score is chosen as 'b-jet' and the two other jets are used to reconstruct the hadronically decaying W boson which becomes W_{Best}. In this case W_{Lep} is identified as W_{Other}. The flowchart for event reconstruction is shown on Figure 5.3.

5.2 Analysis Regions

In this section the selections for ttZ-VR and BDT-SR are finalized. Both, BDT-SR and CUT-SR are validated in previously described validation regions.

5.2.1 $t\bar{t}Z$ Validation Region

The $t\bar{t}Z$ validation region selection given in the previous chapter has a signal purity of 5.8%, i.e. slightly above 5%. An improvement has been made to further decrease the signal fraction using the mass of the reconstructed top quark, m_{top}. A restriction on the mass range between 150–200 GeV is found to be giving the best improvement for the signal fraction. So a new requirement is added,

$$150 \text{ GeV} < m_{\text{top}} < 200 \text{ GeV} \tag{5.9}$$

With this cut, the signal fraction falls to 4.6% and $t\bar{t}Z$ fraction rises to 83% from 79%. This improvement can be understood through the fact that the two processes have a different number of top quarks. The restriction around the on-shell top-quark mass window is easier for $t\bar{t}Z$ process to fulfill because it is more likely to contain one on-shell top quark.

ttZ-VR	Old	New
tWZ	1.61 ± 0.75	0.79 ± 0.45
WZ	1.91± 1.08	1.02 ± 1.61
$t\bar{t}Z$	21.95 ± 5.86	14.40 ± 3.82
tZ	0.93 ± 0.15	0.44 ± 0.09
Others	1.19 ± 0.23	0.58 ± 0.10
Total	27.54 ± 6.10	17.25 ± 3.95
Data	33	21
Signal%	5.8%	4.6%
$t\bar{t}Z$ %	79%	83%

Table 5.1: The change in yields and the signal fraction for signal and $t\bar{t}Z$ are shown in two different ttZ-CR selections, as explained in the text. Both statistical and systematic uncertainties are considered.

5.2.2 BDT Signal Region

To maximize the signal-to-background separation in the BDT-SR a BDT is used. Among the set of additional variables made possible through full event reconstruction, a set of 62 composed from combinations of particles and kinematic relations such as invariant masses, centralities ($\text{C} = \sum_{\text{i}} \frac{p_{\text{T,i}}}{E_{\text{i}}}$), angles and combinations of momenta and energy are given to a BDT as training input. The optimization of the BDT is done through the following steps, which are detailed in the next parts of this section.

- 62 initial variables are ranked by the TMVA package based on their separation power. Starting from the top of the list each variable is checked for modeling and correlation. 20 best variables that are non-correlated (<70%) and well-modeled in validation regions are chosen for parameter optimization. The variables are listed and described on Table 5.2.
- A survey on the effects of 5 BDT parameters on the overtraining and performance are studied by varying one parameter at a time and training a BDT for a range of values.
- Upon finding the optimal intersection for low overtraining and high performance: The effect of individual variables on the performance are tested by removing one variable at a time and checking the ROC integral resulting from the training. If the removal of a variable does not worsen the Overtraining/Performance balance then it is dropped and the procedure is repeated. This approach is called Iterative Removal [35] .

Top 20 Variables	Description	IR Rank	Final Selection
$W_{\text{Had}}^{\text{Mass}}$	Mass of the hadronically decaying W boson	1	✓
$m(b_1$,Non-Z-$\ell)$	Invariant mass of the selected b-jet and non-Z lepton	2	✓
$\Delta R(W_{\text{Had}}, j_1, j_2)$	Distance between W_{Had}, leading and sub-leading jets	3	✓
nJets	Number of jets	4	✓
$p_{\text{T}}^{\text{Non-Z Lepton}}$	p_{T} of the non-Z-lepton	5	✓
$m_{\text{top}}^{\text{mass}}$	Mass of the top quark	6	✓
$H_{\text{T}}(\ell_1, \ell_2, \ell_3, E_{\text{T}}^{\text{miss}})$	Scalar sum of the p_{T} of all leptons and $E_{\text{T}}^{\text{miss}}$	7	✓
$p_{\text{T}}^{\text{Jet 2}}$	p_{T} of the sub-leading jet	8	✓
$m_{W_{\text{Best}}}$	Mass of W_{Best}	9	✓
$m(j_1, \ell_3)$	Invariant mass of selected b-jet and lowest p_{T} lepton	10	✓
ΔR(Sum)	Sum of distances of all reconstructed objects	11	✓
$\Delta p_{\text{T}}(\ell_1, j_2)$	Δp_{T} of leading lepton and sub-leading jet	12	✓
$E_{\text{T}}^{\text{miss}}$	Missing transverse energy	13	✓
C(Top)	Centrality of the top quark	14	✓
$m(b_1, \ell_1)$	Invariant mass of the selected b-jet and leading lepton	15	✓
$\eta^{\text{Non-Z Lepton}}$	Pseudo-rapidity of the non-Z lepton	16	✓
$\text{C}(\ell_1, \ell_3)$	Centrality of leading and lowest p_{T} leptons	17	✓
$p_{\text{T}}^{W_{\text{Other}}}$	p_{T} of W_{Other}	18	✓
$\text{C}(\ell_1, \ell_2)$	Centrality of leading and sub-leading leptons	19	✗
$\text{C}(W_{\text{Had}}, j_1, j_2)$	Centrality of W_{Had}, leading and sub-leading jets	20	✗

Table 5.2: The top 20 variables fulling the criteria described in the text, according to the BDT output. Their descriptions are given on the corresponding row, followed by their rank according to the IR method. Variables that are selected are marked with ✓ on the last column. Variables that are dropped are marked by ✗.

BDT Parameter Optimization

An initial stable configuration below the chosen overtraining limit of 0.003 has been used as the reference setting. The reference is as follows: 100 Trees, 2 Max Depth, nMin% 10, LearnR 0.1. It has an overtraining of 0.0025 with a performance of 0.645.

With this variables and settings fixed, one parameter is varied at a time for a range of values to examine how overtraining and performance evolves. The result is presented in Figure 5.4 for 5 parameters.

Changes in performance is found to be more stable compared to changes in overtraining, where larger fluctuations are observed. The Tree and Max Depth parameters behave as expected. 400 trees are chosen for at this value overtraining has a dip and at higher ranges not much performance gain occurs. Max Depth is found to be strongly correlated with overtraining, thus the initial setting of 2 is kept. For LearnR, extremes are not observed at the highest and lowest values but somewhere in the middle. 0.1 is chosen as final LearnR where a dip in overtraining is again contrasted by almost peak performance.

BDT Settings and Output	Trees	Max Depth	nMin%	nCuts	LearnR	Overtraining	Performance
20 Variable Initial	100	2	10	20	0.1	0.0025	0.645
20 Variables Optimized	400	2	15	30	0.1	0.0025	0.660
18 Final Variables	400	2	15	30	0.1	0.0030	0.665

Table 5.3: Summary table for the three different setups used during the BDT optimization studies and their corresponding overtraining and performance outputs.

Overtraining for nCuts and nMin% shows a different behaviour at the highest values against the

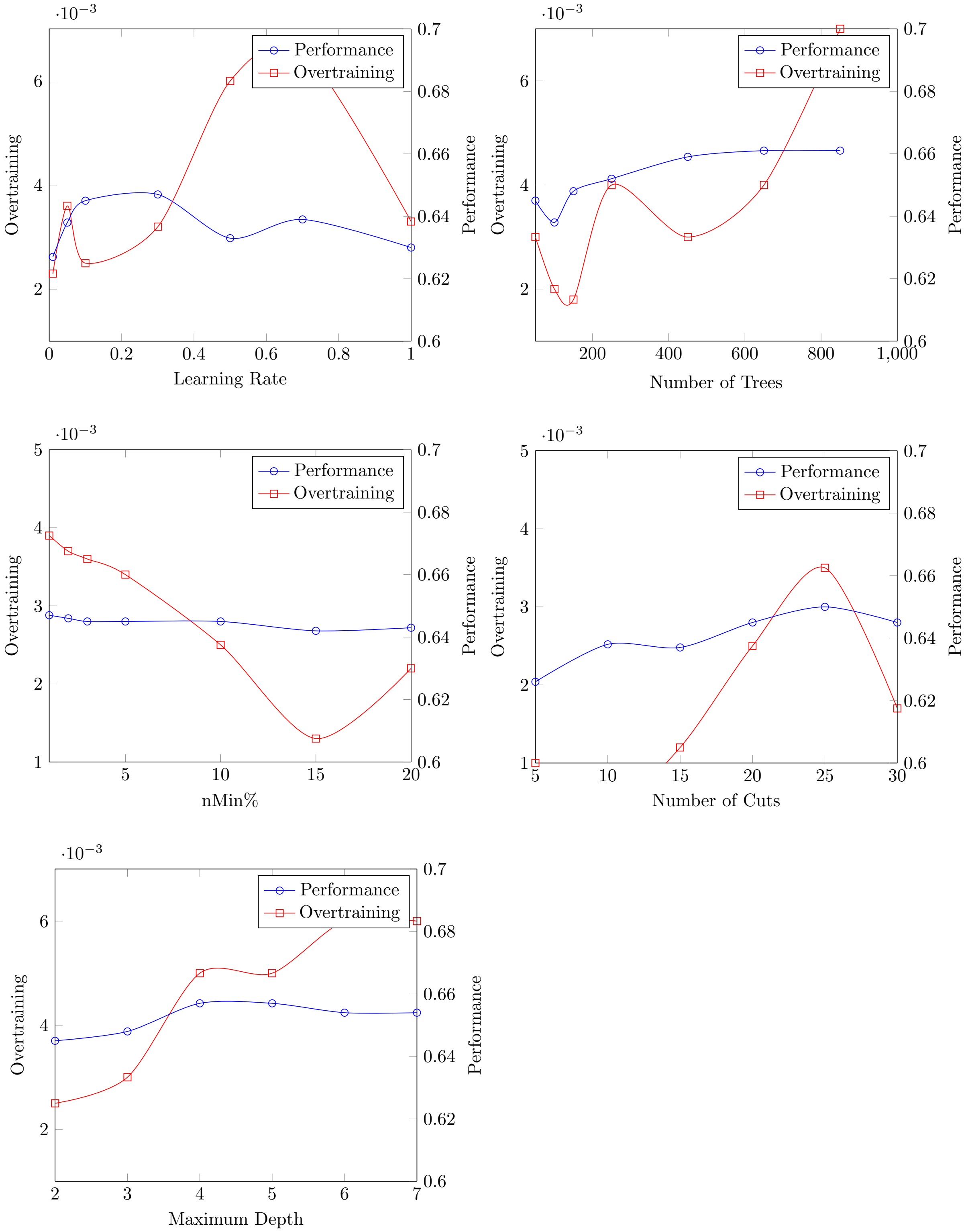


Figure 5.4: Overtraining and performance plots for the five BDT parameters considered.

expected trend. For nMin% a clear global overtraining minimum occurs at 15. Since the performance is almost flat on the whole range, this value is chosen. For nCuts, 30 is chosen where a slightly increased performance can be reached with a local overtraining minimum. After the training with the new settings the same overtraining value as with reference setting is found with an increased performance of 0.660. The results and settings are shown in Table 5.3. The corresponding ROC integrals and overtraining check plots provided by TMVA package are shown in Figure 5.6.

Selection of Final BDT Variables

With optimized parameters, a study is done using the Iterative Removal technique to rank the variables based on their individual contribution to Performance. First, 20 trainings are made each of which only differ by the removal of a single variable. The resulting performances are used to rank the variables. If the removal of lowest ranking variable does not worsen the performance, it is dropped. The procedure is repeated for the set of remaining variables until 10 variables remain. After this study, it had been possible to exclude 2 variables out of 20 without passing the predetermined overtraining threshold. A slight increase in performance to 0.665 is observed as shown at the bottom row in Table 5.3. The variables are listed and described on Table 5.2. The modeling of final set of variables in WZ-3J-VR and ttZ-VR are shown in Figures 5.7, 5.8, 5.9 and 5.10. No mis-modeling is observed.

Selection of the BDT Cut

The final BDT distribution has a range between -0.6 and +0.6. A scan of the BDT range is done with 0.01 increments steps and statistical significant is checked at each cut. A maximum stastistical significance of 1.021 is found for the cut at BDT≥-0.33. The optimization resulted in less than 2% improvement over the case without a BDT in which a statistical significance of 1.009 had been reached.

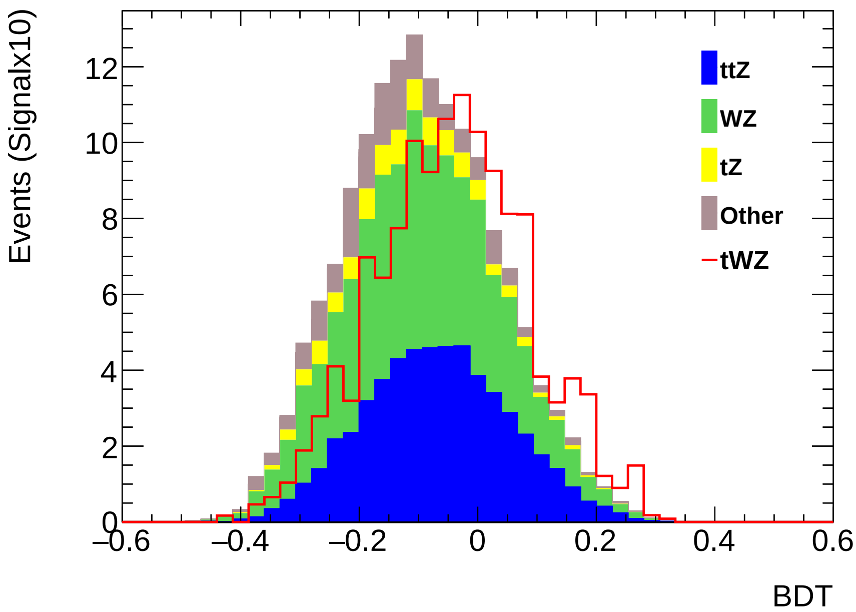


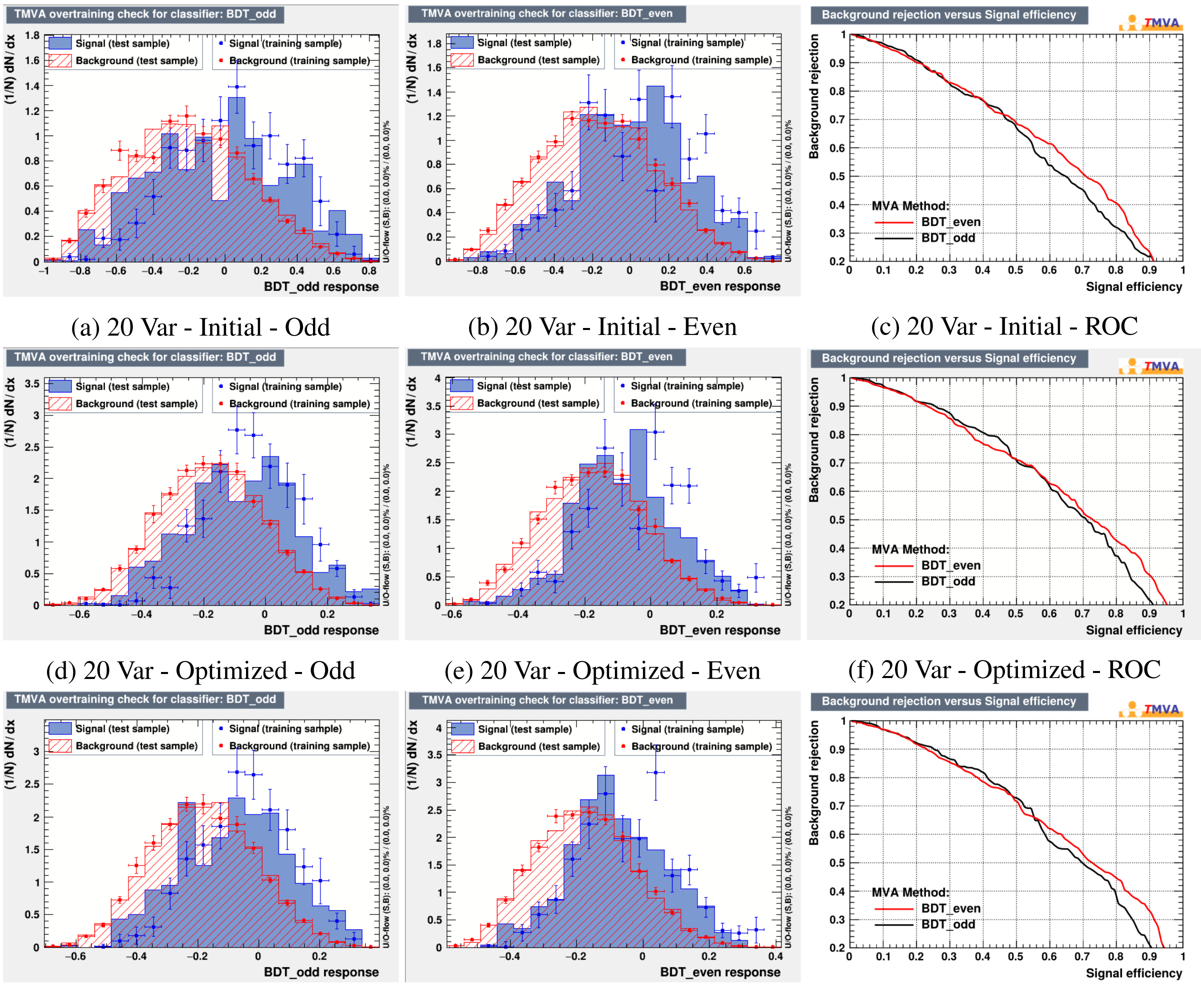
Figure 5.5: The BDT response of final optimized setting in BDT-SR. Backgrounds are presented as a stack plot. Signal is overlayed on the stack as a red line and multiplied by a factor of 10 so as to make its shape visible.

Combining the statistical significance of two signal regions a total statistical significance of 1.1 is reached. As discussed in the previous section, this value is too low for claiming an evidence or discovery. Thus, it has been decided to set an upper limit on the *tWZ* production cross section. The BDT is re-optimized in the next chapter so as to give the best upper limit.

BDT-SR	No Cut	BDT > -0.33
tWZ	13.02 ± 4.93	12.88 ± 4.88
WZ	71.36 ± 39.58	67.49 ± 37.23
ttZ	55.63 ± 11.52	54.25 ± 10.97
tZ	9.69 ± 1.04	15.40 ± 1.88
Others	16.78 ± 2.10	9.03 ± 0.96
Total	166.5 ± 41.9	159.1 ± 39.5
$\frac{S}{\sqrt{S+B}}$	1.009	1.021

Table 5.4: Pre and post BDT cut yield and significance comparison for the BDT-SR. Both systematic and statistical uncertainties are included.

For the CUT-SR signal region, modeling of seven variables are checked in the WZ-2J-VR. These are the p_{T} of leptons, of the two leading jets and $E_{\mathrm{T}}^{\mathrm{miss}}$. Validation process showed good modeling. Plots are shown in Figure 5.11.



(a) 20 Var - Initial - Odd (b) 20 Var - Initial - Even (c) 20 Var - Initial - ROC

(d) 20 Var - Optimized - Odd (e) 20 Var - Optimized - Even (f) 20 Var - Optimized - ROC

(g) 18 Var - Optimized - Odd (h) 18 Var - Optimized - Even (i) 18 Var - Optimized - ROC

Figure 5.6: Overtraining plots for even and odd tests, and ROC integrals for the three cases studied, are shown. The upper row plots belong to initial training with 20 variables. Middle row plots show optimized training with 20 variables and bottom row shows training with 18 variables in the final setting.

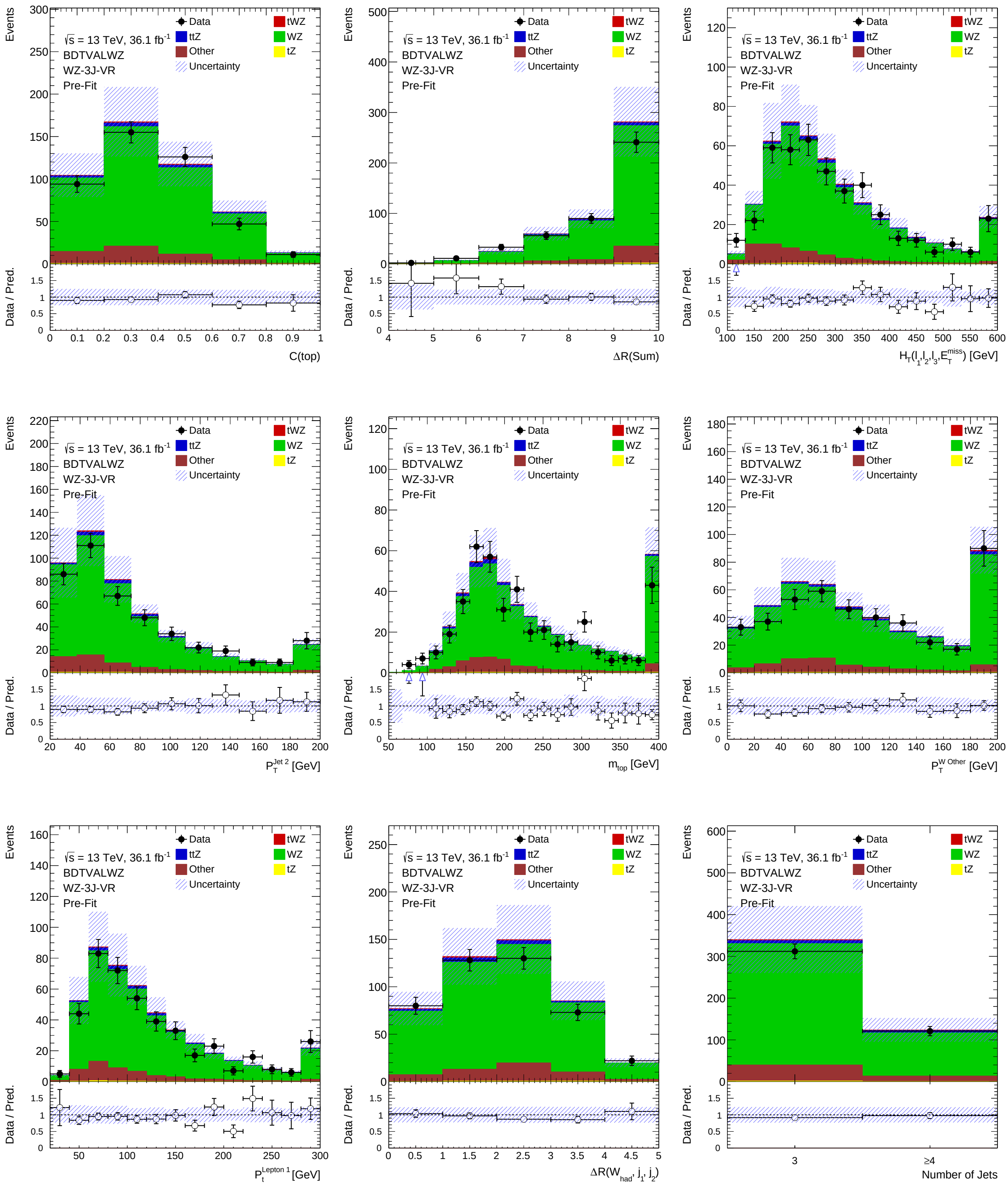


Figure 5.7: Comparison of data and expectation for BDT variables in the WZ-3J-VR validation region. Both statistical and systematical uncertainties are included. The overflow is included in the last bin of each histogram.

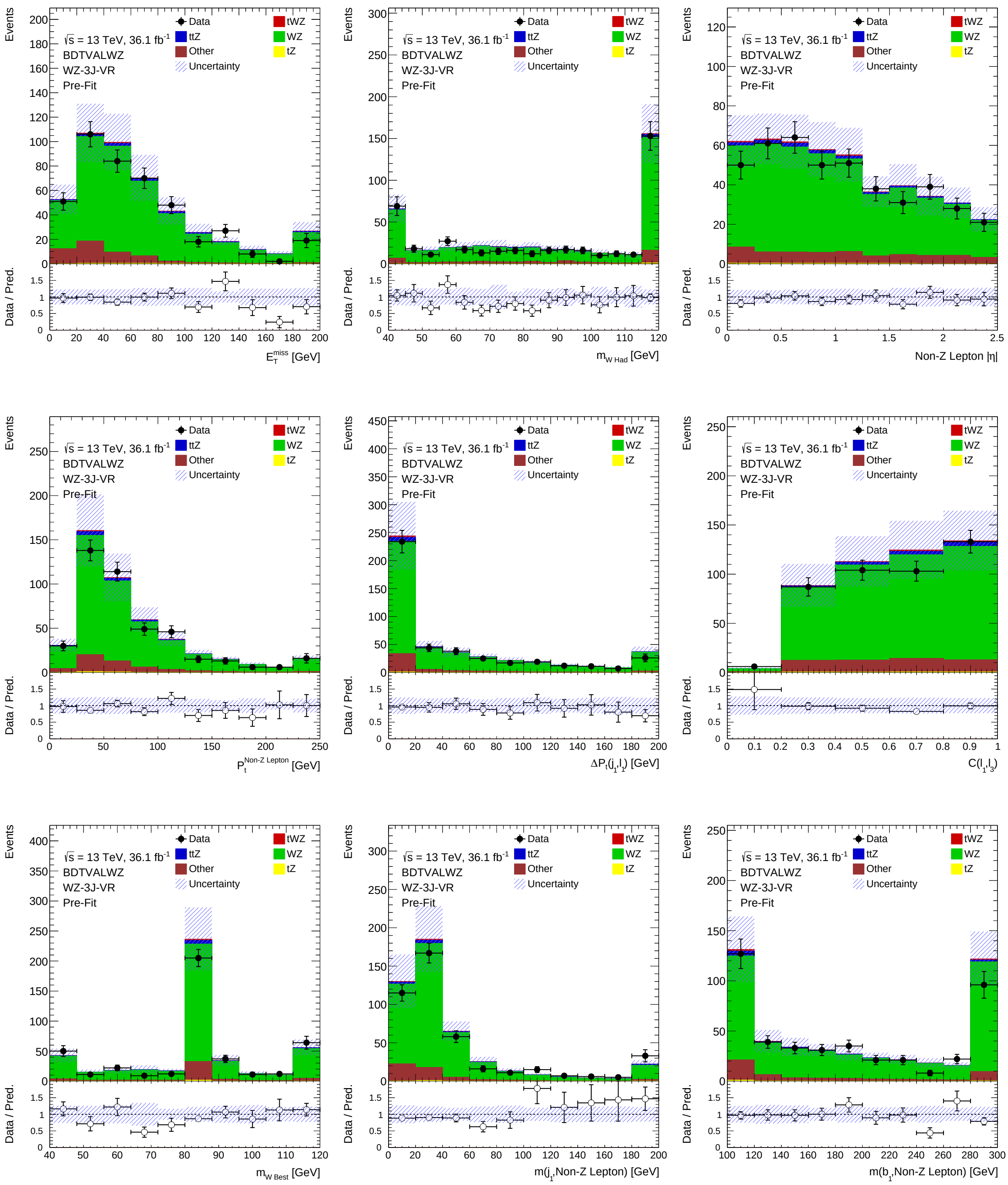


Figure 5.8: Comparison of data and expectation for BDT variables in the WZ-3J-VR validation region. Both statistical and systematical uncertainties are included. The overflow is included in the last bin of each histogram.

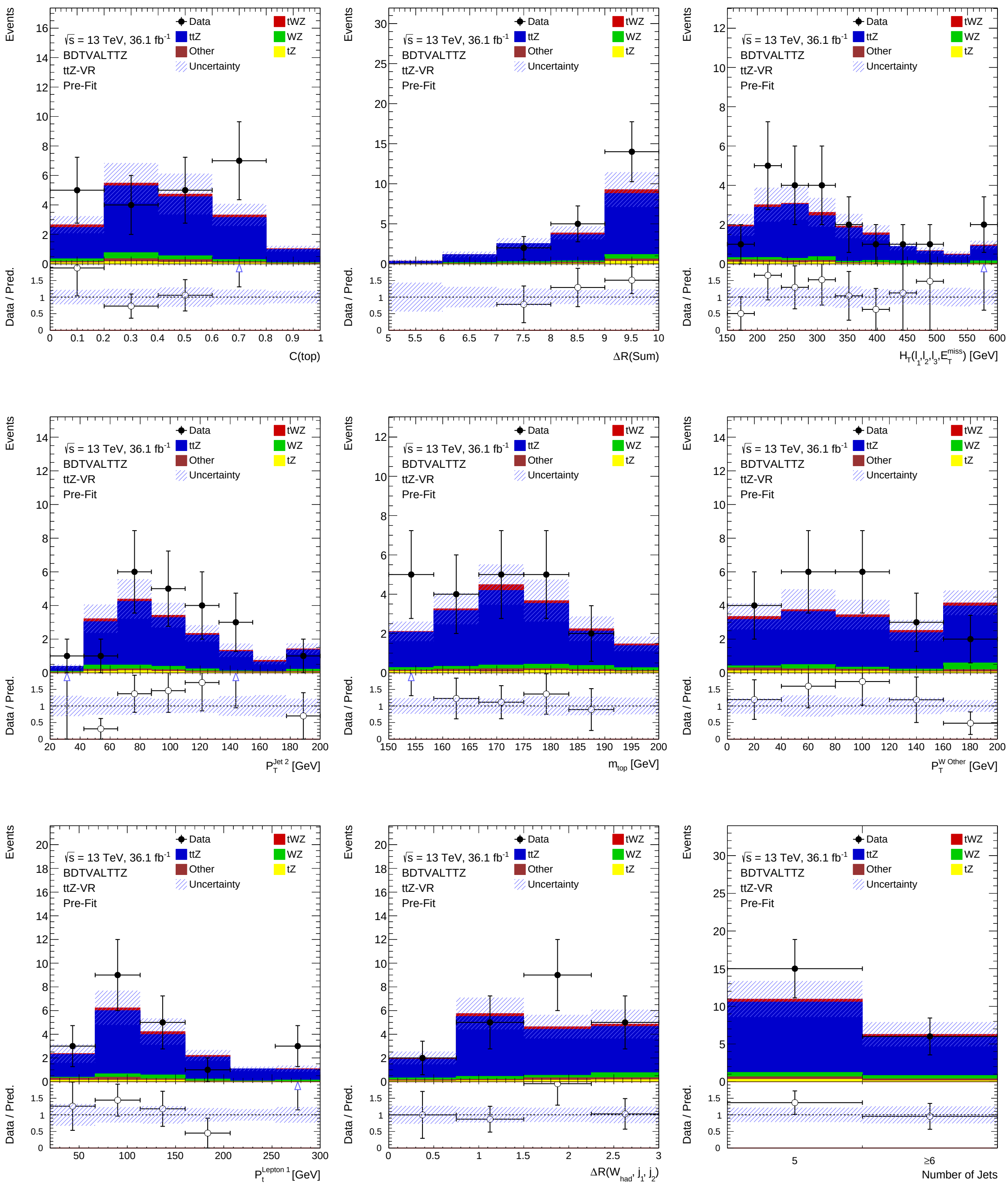


Figure 5.9: Comparison of data and expectation for BDT variables in the ttZ-VR validation region. Both statistical and systematical uncertainties are included. The overflow is included in the last bin of each histogram.

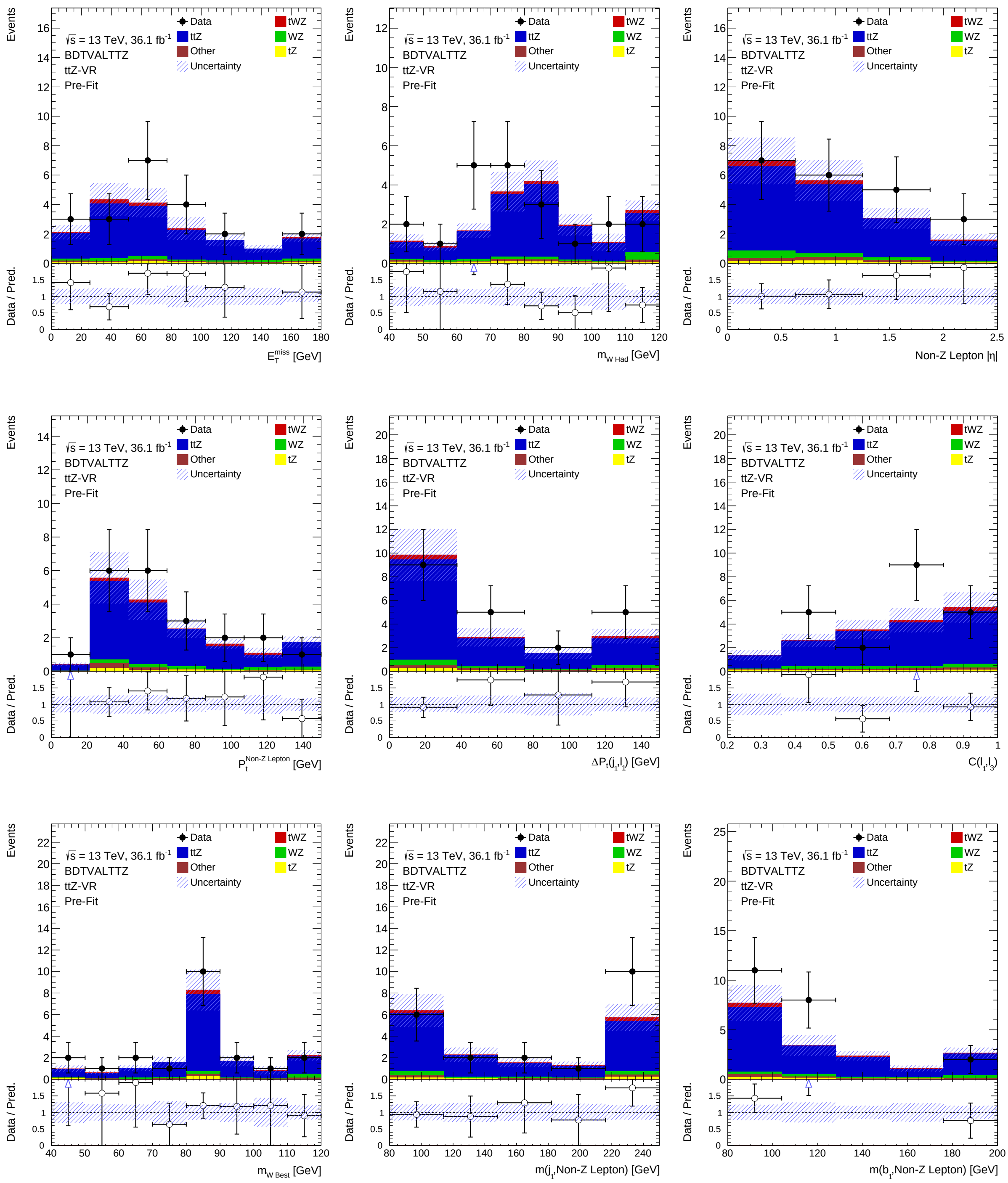


Figure 5.10: Comparison of data and expectation for BDT variables in the ttZ-VR validation region. Both statistical and systematical uncertainties are included. The overflow is included in the last bin of each histogram.

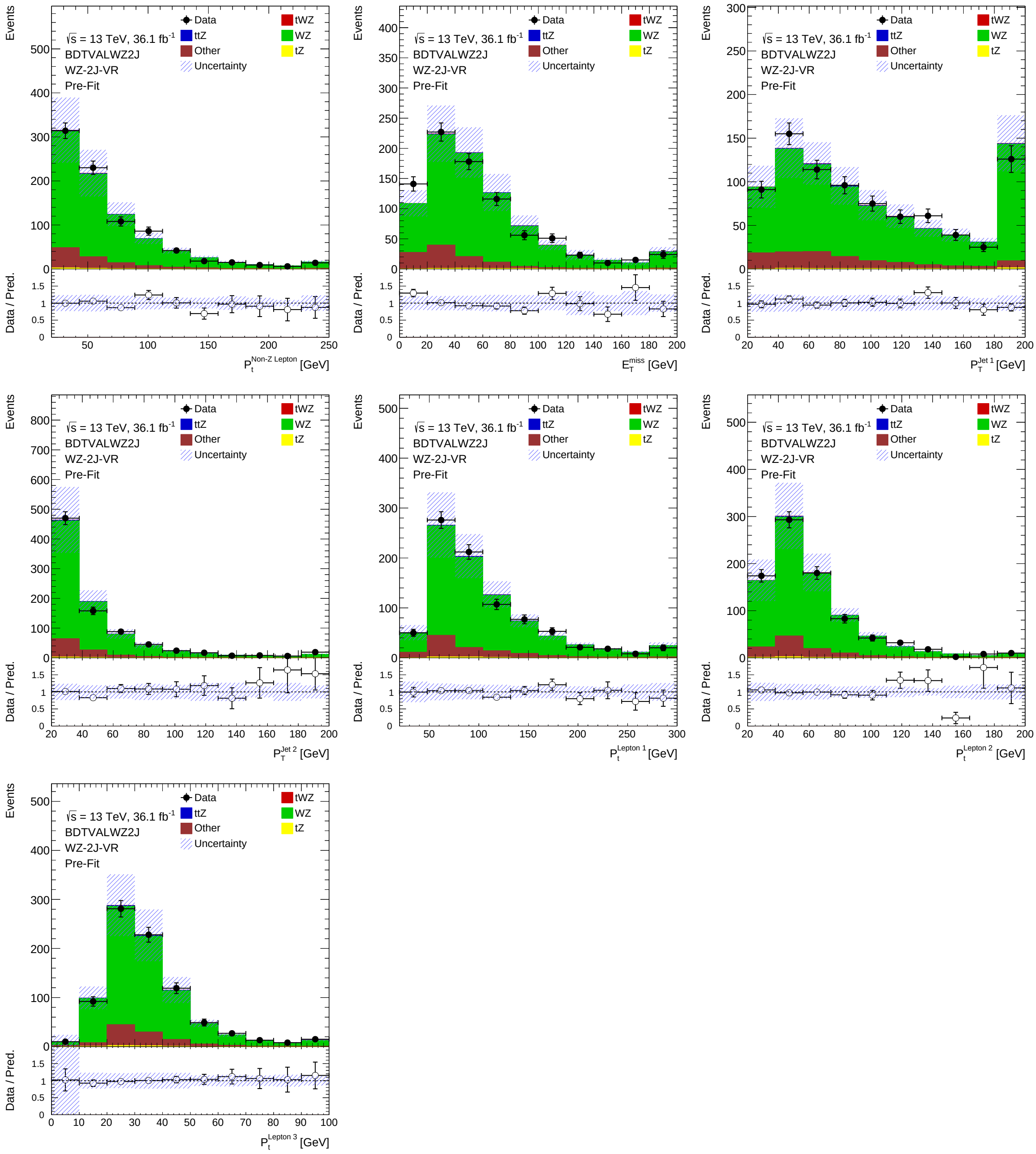


Figure 5.11: Comparison of data and expectation for WZ-2J-VR. Both statistical and systematical uncertainties are included. The overflow is included in the last bin of each histogram.

CHAPTER 6

Setting an Upper Limit on the Standard Model *tWZ* Production Cross Section

In this chapter an upper limit on the production cross-section is set. This is done by applying a fit, which is performed on a distribution containing 3 bins, corresponding to the two signal regions and the WZ control region. An upper limit is calculated using the $\mathrm{CL_s}$ method [36]. Both procedures have been executed using the ROOT data analysis framework and the TReXFitter computer package, which is using the HistFactory tool [37]. Before the fit could be done, systematic uncertainties needs to be considered. Thus, first a description of the systematic uncertainties considered in this analysis are given. Fitting is done in two steps. First, the BDT cut is optimized, with the simulation samples only, so as to give the best exclusion limit. At the end a fit to data is applied and the final result is reported.

6.1 Systematic Uncertainties

Estimating and reporting uncertainties are an integral part of any experimental effort. They provide information about the precision of the measurement and crucial in assessing the importance of the result. Reduction in uncertainties could promote the status of a measurement to a higher rank, such as evidence or discovery. In physics experiments uncertainties on measurements are categorized as statistical or systematical uncertainties. Statistical uncertainties reflect the finite sample size used in the measurement and the sensitivity of this selected subset to the fluctuations. Statistical uncertainties are denoted by $\pm 1\sigma$ standard deviation from the mean assuming a Gaussian distribution. Statistical uncertainties can be reduced by having more data.

Systematic uncertainties have diverse sources and generally they are defined as any type of uncertainty that is not statistical. In this study systematic uncertainties are divided into four categories. Signal, background, experimental and reconstruction uncertainties. Signal uncertainties apply only to signal sample. As it is the main object of interest, several possible uncertainties has to be consider for it. The background systematic uncertainties apply only to backgrounds. The experimental systematic uncertainties affect all samples and consist of luminosity and pile-up uncertainties. The reconstruction uncertainties are applied to signal and two leading background samples, $t\bar{t}Z$ and WZ only. They account for the efficiencies of object reconstructions. Due to their low importance, other processes have only been assigned theoretical modeling uncertainties, which is placed under the background uncertainties category.

The fixed percentages applied follow the choices made of other ATLAS analyses, like in Reference [15]. The detector uncertainties are assigned according to ATLAS recommendations.

Signal Uncertainties

- Parton Showering: Different parton showering models are used in different simulation methods. The tWZ signal used in the analysis is generated with Herwig. Another sample generated with Pythia is used to account for uncertainties stemming from this choice. After comparison of the two samples a 10% uncertainty is assigned.
- Diagram Removal: Two methods can be used for overlap removal between the signal sample and $t\bar{t}Z$ process. An uncertainty of 28% has been set to account for differences between the two methods and another 10% of uncertainty is included for the uncertainty of the used sample itself, based on studies done in Reference [38]
- Modeling Uncertainty: A 10% uncertainty is defined for uncertainty on the cross-section calculation.
- Renormalization and Factorization Scale Choice: The uncertainties due to approximations made in the modeling of parton fragmentation have been included by varying μ_F and μ_R scale factors with factors of two.

Background Uncertainties

- WZ: The normalization of the WZ process has been left as a free-floating parameter to be determined by the fit. Since the WZ control region is defined in the zero b-tag region, an uncertainty of 50% is considered to account for extrapolating its effect to the 1 b-tag signal regions.
- $t\bar{t}Z$: 10% theoretical cross-section modeling uncertainty is applied.
- Others: 10% theoretical cross-section modeling uncertainty is applied.

Experimental Uncertainties

- Luminosity: The uncertainty coming from luminosity is set to 2.1% based on the methods described in Reference [39].
- Pile-Up: Uncertainties coming from pile-up are another source of experimental uncertainty. The events are re-weighted such that they correspond to a reduced or augmented pile-up regime corresponding to the uncertainty on the observed pile-up distribution.

Reconstruction Uncertainties

- Jets: For the uncertainties on jets, three separate families of components are considered: Jet Energy Scale, Jet Energy Resolution (JER) and Jet Vertex Tagging efficiencies. The latter accounts for the uncertainty on identifying the jets originating from the correct event. The Jet Energy Scale uncertainty accounts for the difference between the energy measurement of simulation and actual data, based on calibration studies. The JER uncertainty accounts for the difference between the resolution of the detector for jet measurements. The uncertainties are related to the η and p_{T} values of the jets, thus these components are used.
- b-Tagging: Three different flavour tagging categories are considered for this uncertainty. The b-quark tagging efficiency, the c-quark tagging efficiency and the tagging of all other jets, corresponding to a misidentification.

- Leptons: For the leptons, the main components are reconstruction and identification efficiencies. The difference in efficiencies in data and simulation are used as scale factors. These factors depend on the energy and η of the given lepton. The uncertainty on these parameters are used for the estimation of systematic uncertainties.

- Missing Transverse Energy: The missing transverse energy is amongst the most difficult to reconstruct objects, because its definition is based on the combination of other objects. Therefore the uncertainties on these objects are propagated to the uncertainty on the missing transverse energy measurement. Another uncertainty that is used is the soft-term uncertainty. Soft term stands for the transverse momentum recorded in the detector that has not been associated with any reconstructed objects. Because of that, this contribution is not included in the propagated uncertainty from reconstructed objects and included separately [40].

6.2 Optimization of the Fit

For the fit and limit setting a test statistic is generated using the profile likelihood method in two signal regions and the WZ control region. The effect of systematic uncertainties are included in the likelihood function as *nuisance paramaters* (θ). Each systematic uncertainty is defined as a nuisance parameter and the uncertainties on the signal and backgrounds change according to their combined impact on the fit. As SM is assumed, the signal strength μ (ratio of observed cross section to the SM prediction) is set to 1. The test statistics is used by program to calculate the exclusion upper limit with the $\mathrm{CL_s}$ method as described in reference [41], at the 95% confidence level (CL). In order to decide for the best cut, a scan of the BDT response in units of 0.01 is checked for the strictest exclusion limit. The best cut is found to be at BDT>-0.10 giving an expected exclusion upper limit of $4.6^{+7.6}_{-3.3}$.

6.3 Results

After the unblinding, a good agreement between data and simulation is observed, especially in the CUT-SR region. The yields for pre-fit and post-fit cases are shown in Table 6.1.

Table 6.1: Expected number of events compared to data before (left) and after (right) the fit.

Pre-fit	BDT-SR	CUT-SR	WZ-CR	Post-fit	BDT-SR	CUT-SR	WZ-CR
tWZ	8.69 ± 3.23	4.89 ± 1.80	7.33 ± 2.91	tWZ	14.06 ± 16.83	7.91 ± 9.61	11.86 ± 14.29
$t\bar{t}Z$	30.86 ± 6.00	12.13 ± 1.89	16.96 ± 3.27	$t\bar{t}Z$	30.87 ± 5.95	12.13 ± 1.88	16.96 ± 3.24
WZ	32.90 ± 18.35	69.04 ± 39.33	1 120 ± 288	WZ	31.33 ± 8.26	65.71 ± 16.25	1 080 ± 43
Other	5.30 ± 0.80	20.71 ± 4.58	155.8 ± 21.2	Other	5.29 ± 0.80	20.70 ± 4.53	155.7 ± 21.1
tZ	3.38 ± 0.44	20.64 ± 2.35	8.89 ± 1.37	tZ	3.38 ± 0.44	20.64 ± 2.32	8.89 ± 1.36
Total	81.1 ± 19.7	127.4 ± 40.1	1 310 ± 290	Total	84.9 ± 10.3	127.1 ± 10.5	1 270 ± 38
Data	85	127	1275	Data	85	127	1275

No modeling problems are observed. The fit is repeated with data at the previously optimized BDT cut of -0.10. Pre-fit and post-fit plots for signal regions are provided in the Appendix A. An upper limit of 6.6 is set on the SM production cross section with 95% CL. The increase in the upper limit is in agreement with the extracted signal strength μ =1.62 i.e. 62% above the SM but still within uncertainties. Therefore no deviations from SM is observed. Results are summarized in Table 6.2.

Results	95% CL upper limit	$\mu = \frac{\sigma_{tWZ,\text{measured}}}{\sigma_{tWZ,\text{SM}}}$
Expected	$4.6^{+7.6}_{-3.3}$	$1.00^{+1.03}_{-0.97}$(Stat.) $^{+1.85}_{-1.50}$(Sys.)
Observed	6.6	$1.62^{+1.10}_{-0.90}$(Stat.) $^{+2.01}_{-1.46}$(Sys.)

Table 6.2: Expected and observed results for 95% CL upper limit, for the signal strength first uncertainties are statistical (Stat.) second are systematic (Sys.) uncertainties.

Systematic uncertainties with highest impact on the fit are shown in the Figure 6.1. Full list of nuisance parameters and correlation matrix is provided in the Appendix C. Most important uncertainty is the heavy flavor extrapolation uncertainty for the WZ process. The modeling uncertainty on the Diagram Removal method and reconstruction uncertainties of $t\bar{t}Z$ process are also found to be important. The WZ normalization is 0.97 ± 0.25. The fit had a constraining effect on the WZ uncertainty.

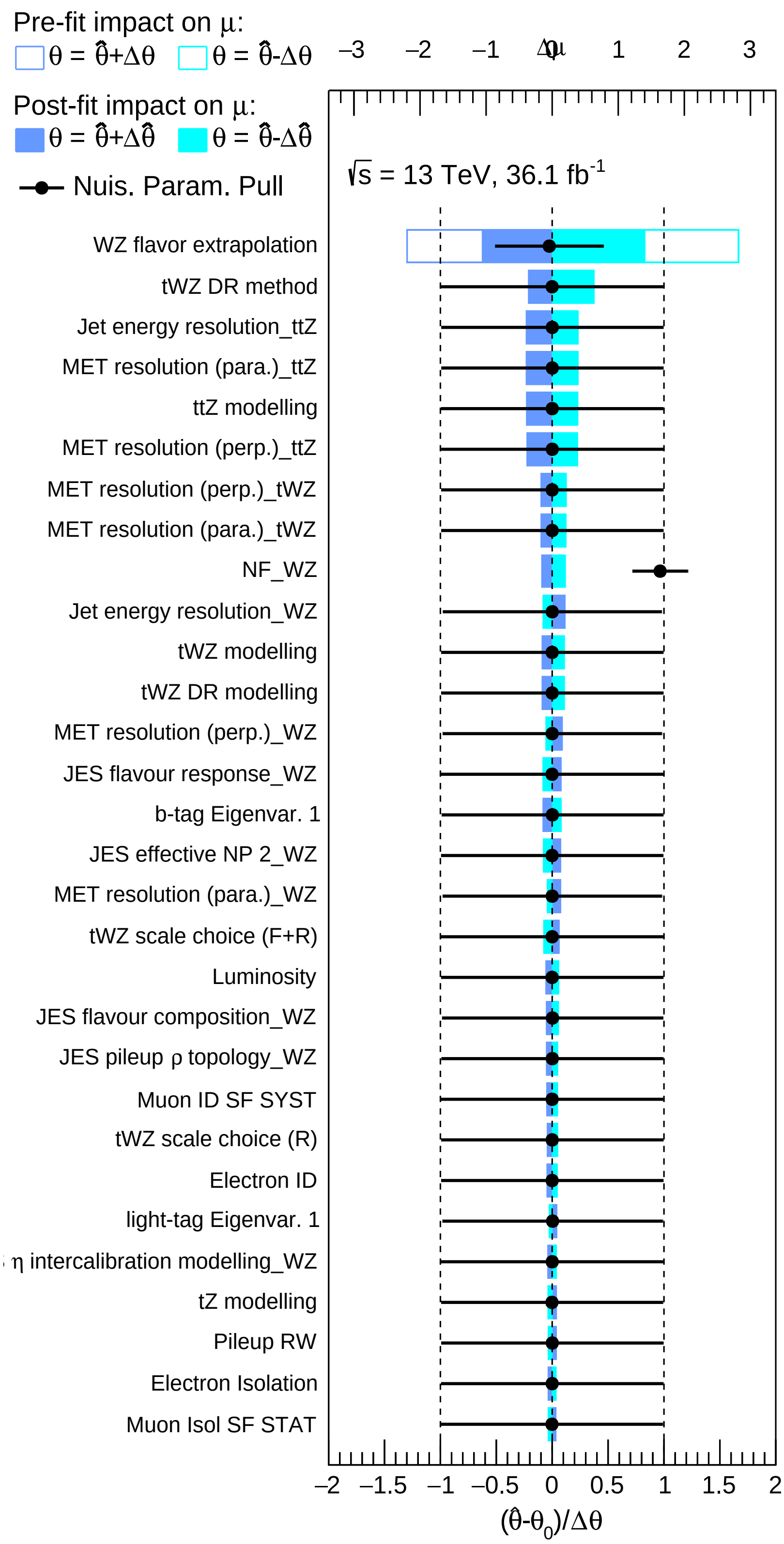


Figure 6.1: Ranking plot showing the systematic uncertainties with highest impact on the signal strength.

CHAPTER 7

Summary and Outlook

This chapter gives a summary of the study that has been done in this thesis and discusses possible ways to extend and improve the analysis for future exploration.

With its unique mass and decay properties, top-quark physics is an active and important field of research. The single top quark production presents a possible way of investigating these properties. In this thesis a measurement on the associated top quark production with a W and a Z boson has been made. This is the first measurement made on this process. The tWZ channel could be instrumental in the search for new physics, for example through investigation of weak coupling. Also, it is an important background in other top quark measurements such as tW and $t\bar{t}Z$. Thus, a better understanding of this process is of great use.

The analysis has been done using ATLAS detector proton–proton collision data collected at the LHC in the trilepton channel. Two signal regions are defined in the 1 b-jet region. In the 2j region a simple cut based selection is applied. For the higher light-jet multiplicity signal region a BDT is trained. The variables used are validated in separate regions targeting WZ and $t\bar{t}Z$ in the 0 and 2 b-tagged jet regions respectively. In addition, a WZ control region is defined in the 0 b-tagged jet region to constrain this process. A likelihood fit is applied to three bins composed of WZ control region and two signal regions. The fit is optimized so as to give the strictest upper limit on tWZ production cross section.

The result of this thesis have been reached through several steps in the analysis process. In each of these steps several decisions had to be made in order to proceed. A complete survey of deeper details were not possible due to limitations on time and resources. Naturally, the result obtained depends on these decisions and their cumulative effect.

A better understanding of overlap between tWZ and $t\bar{t}Z$ processes in the 2 b-jet region could make a signal region in this jet multiplicity possible. Although this overlap is similar to the overlap problem between the tW and $t\bar{t}$ process. Preliminary studies showed that the presence of Z boson does change the event for example making the second b-tagged jet emitting a Z boson less-forward in the detector.

Generator level studies were only done for the signal sample. A comparative study with the main backgrounds $t\bar{t}Z$ and WZ might reveal some new information that could render better filtering of these irreducible backgrounds possible.

Despite a dedicated effort, the BDT response was relatively poor with a ROC integral of 0.665, and only less than 2% improvement over the pre-BDT selection. This is also attributed to the limited sample size of the signal process 1200 raw training events. The low number of events had two effects on the BDT training. First, specific signal regions such as the 2j1b golden channel could not be targeted. The initial idea was to train two separate BDTs in this channel, one for each top-quark decay mode. In the end, a much more inclusive selection had to be chosen to preserve enough events for the training. Secondly, the training parameters could not be set to high performance, for overtraining prevails very quickly. Samples with a higher number of events could enable one to train more powerful BDTs in more specific regions.

In this analysis the priority had been given to the agreement between data and simulation. To achieve this in a limited amount of time, selections have been chosen in a conservative manner. For example a $E_{\mathrm{T}}^{\mathrm{miss}}$ cut could increase the WZ purity in the control region but it has not been preferred due to concerns in the modeling. Given enough scrutiny and understanding such cuts could be in principle implemented without concern.

Non-prompt lepton estimation is done through Monte-Carlo samples. Proper estimation of non-prompt leptons is another important point that could improve the analysis.

Preliminary studies done with a shape fit on the BDT distribution instead of the single bin with the optimal cut hints to a room for improvement. This method could not have been implemented in this study due to lack of time required to investigate the behaviour of systematic uncertainties.

In conclusion, a measurement has been done on the tWZ process for the first time and an upper limit of 6.6 has been set on the Standard Model prediction. The high fraction of irreducible backgrounds and low production rate are the main challenges in this channel, but the outlook is positive and many ideas are waiting to be picked.

Bibliography

[1] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. C680519 (1968) 367.

[2] S. L. Glashow, *Partial-symmetries of weak interactions*, Nuclear Physics **22** (1961) 579, ISSN: 0029-5582.

[3] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264.

[4] D. Griffiths, *Introduction to elementary particles*, 2008, ISBN: 9783527406012.

[5] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*, 1995, ISBN: 9780201503975.

[6] M. D. Schwartz, *Quantum Field Theory and the Standard Model*, Cambridge University Press, 2014, ISBN: 1107034736, 9781107034730.

[7] C. Patrignani et al. (Particle Data Group), *Review of Particle Physics*, Chin. Phys. **C40** (2016) 100001.

[8] MissJl, Nasfarley88 et al., *Standard Model of Elementary Particle Physics*, URL: `http://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles_modified_version.svg` (visited on 20/05/2018).

[9] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508.

[10] P. W. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons*, Phys. Rev. **145** (1966) 1156.

[11] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321.

[12] F. Abe et al. (CDF Collaboration), *Observation of Top Quark Production in $p\bar{p}$ Collisions with the Collider Detector at Fermilab*, Phys. Rev. Lett. **74** (1995) 2626, arXiv: `hep-ex/9503002`.

[13] S. Abachi et al. (DØ Collaboration), *Observation of the top quark*, Phys. Rev. Lett. **74** (1995) 2632, arXiv: `hep-ex/9503003`.

[14] Abazov, V. M. and others (DØ Collaboration), *Observation of Single Top Quark Production*, Phys. Rev. Lett. **103** (2009) 092001, arXiv: `0903.0850`.

[15] M. Aaboud et al. (ATLAS Collaboration), *Measurement of the $t\bar{t}Z$ and $t\bar{t}W$ production cross sections in multilepton final states using 3.2 fb^{-1} of pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, Eur. Phys. J. **C77** (2017) 40, arXiv: `1609.01599`.

[16] M. Aaboud et al. (ATLAS Collaboration), *Measurement of differential cross-sections of a single top quark produced in association with a W boson at $\sqrt{s}$ = 13 TeV with ATLAS*, Eur. Phys. J. **C78** (2018) 186, arXiv: `1712.01602`.

[17] M. Aaboud et al. (ATLAS Collaboration), *Measurement of the production cross-section of a single top quark in association with a Z boson in proton–proton collisions at 13 TeV with the ATLAS detector*, Phys. Lett. **B780** (2018) 557, arXiv: 1710.03659.

[18] M. Aaboud et al. (ATLAS Collaboration), *Standard Model Total Production Cross Section Measurements*, URL: http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/SM/ (visited on 20/05/2018).

[19] M. Aaboud et al. (ATLAS Collaboration), *Electron identification measurements in ATLAS using* $\sqrt{s} = 13$ *TeV data with* 50 *ns bunch spacing*, (2015), ATL-PHYS-PUB-2015-041, URL: http://cdsweb.cern.ch/record/2048202.

[20] G. Aad et al. (ATLAS Collaboration), *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1, arXiv: 1207.7214.

[21] A. Airapetian et al. (ATLAS Collaboration), *ATLAS: Detector and physics performance technical design report. Volume 1*, (1999).

[22] G. Aad et al. (ATLAS Collaboration), *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[23] G. Aad et al. (ATLAS Collaboration), *Expected performance of the ATLAS b-tagging algorithms in Run-2*, (2015), ATL-PHYS-PUB-2015-022, URL: http://cds.cern.ch/record/2037697.

[24] N. Bartosik, *Diagram showing the common principle of identification of jets initiated by b-hadron decays*, URL: http://en.wikipedia.org/wiki/B-tagging#/media/File:B-tagging_diagram.png (visited on 20/05/2018).

[25] M. Aaboud et al. (ATLAS Collaboration), *ATLAS Luminosity Public Results*, URL: http://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2 (visited on 20/05/2018).

[26] M. Aaboud et al. (ATLAS Collaboration), *Evidence for the associated production of the Higgs boson and a top quark pair with the ATLAS detector*, Phys. Rev. **D97** (2018) 072003, arXiv: 1712.08891.

[27] M. Cacciari, G. P. Salam and G. Soyez, *The Anti-*k_t *jet clustering algorithm*, JHEP **04** (2008) 063, arXiv: 0802.1189.

[28] G. Aad et al. (ATLAS Collaboration), *Tagging and suppression of pileup jets with the ATLAS detector*, (2014), ATLAS-CONF-2014-018, URL: https://cds.cern.ch/record/1700870.

[29] G. Aad et al. (ATLAS Collaboration), *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, (2009), arXiv: 0901.0512.

[30] J. Krause and F. Siegert, *NLO QCD predictions for* $Z + \gamma +$ *jets production with Sherpa*, Eur. Phys. J. **C78** (2018) 161, arXiv: 1708.06283.

[31] O. Bessidskaia Bylund,
Top quark and heavy vector boson associated production at the ATLAS experiment,
PhD thesis: Stockholm U., 2017,
URL: http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-147262.

[32] A. Höcker et al., *TMVA - Toolkit for Multivariate Data Analysis*, PoS ACAT (2007) 040,
arXiv: physics/0703039.

[33] Chabacano, *Overtraining Plot*,
URL: http://en.wikipedia.org/wiki/Overfitting#/media/File:Overfitting.svg
(visited on 20/05/2018).

[34] T. Chwalek, *Measurement of W -Boson Helicity-Fractions in Top-Quark Decays with the CDF II Experiment and Prospects for an Early $t\bar{t}$ Cross-Section Measurement with the CMS Experiment*,
PhD thesis: Karlsruhe U., 2010,
URL: http://inspirehep.net/record/1087965/files/CERN-THESIS-2010-255.pdf.

[35] S. An, *Optimising Variable Selection for Machine Learning Analysis in ATLAS $t\bar{t}H$ Search*, URL:
http://www.desy.de/f/students/2017/reports/SitongAn.pdf (visited on 20/05/2018).

[36] A. L. Read, *Presentation of search results: The CL(s) technique*, J. Phys. **G28** (2002) 2693.

[37] K. Cranmer et al.,
HistFactory: A tool for creating statistical models for use with RooFit and RooStats, (2012),
URL: http://cds.cern.ch/record/1456844.

[38] O. Bessidskaia Bylund, "Modelling Wt and tWZ production at NLO for ATLAS analyses",
Proceedings, 9th International Workshop on Top Quark Physics (TOP 2016): Olomouc, Czech Republic, September 19-23, 2016, arXiv: 1612.00440.

[39] M. Aaboud et al. (ATLAS Collaboration),
Luminosity determination in pp collisions at $\sqrt{s}$ = 8 TeV using the ATLAS detector at the LHC,
Eur. Phys. J. **C76** (2016) 653, arXiv: 1608.03953.

[40] G. Aad et al. (ATLAS Collaboration), *Expected performance of missing transverse momentum reconstruction for the ATLAS detector at $\sqrt{s}$ = 13 TeV*, (), ATL-PHYS-PUB-2015-023, 2015,
URL: http://cds.cern.ch/record/2037700.

[41] G. Cowan et al., *Asymptotic formulae for likelihood-based tests of new physics*,
Eur. Phys. J. **C71** (2011) 1554, arXiv: 1007.1727.

APPENDIX A

Validation Plots for Signal Regions

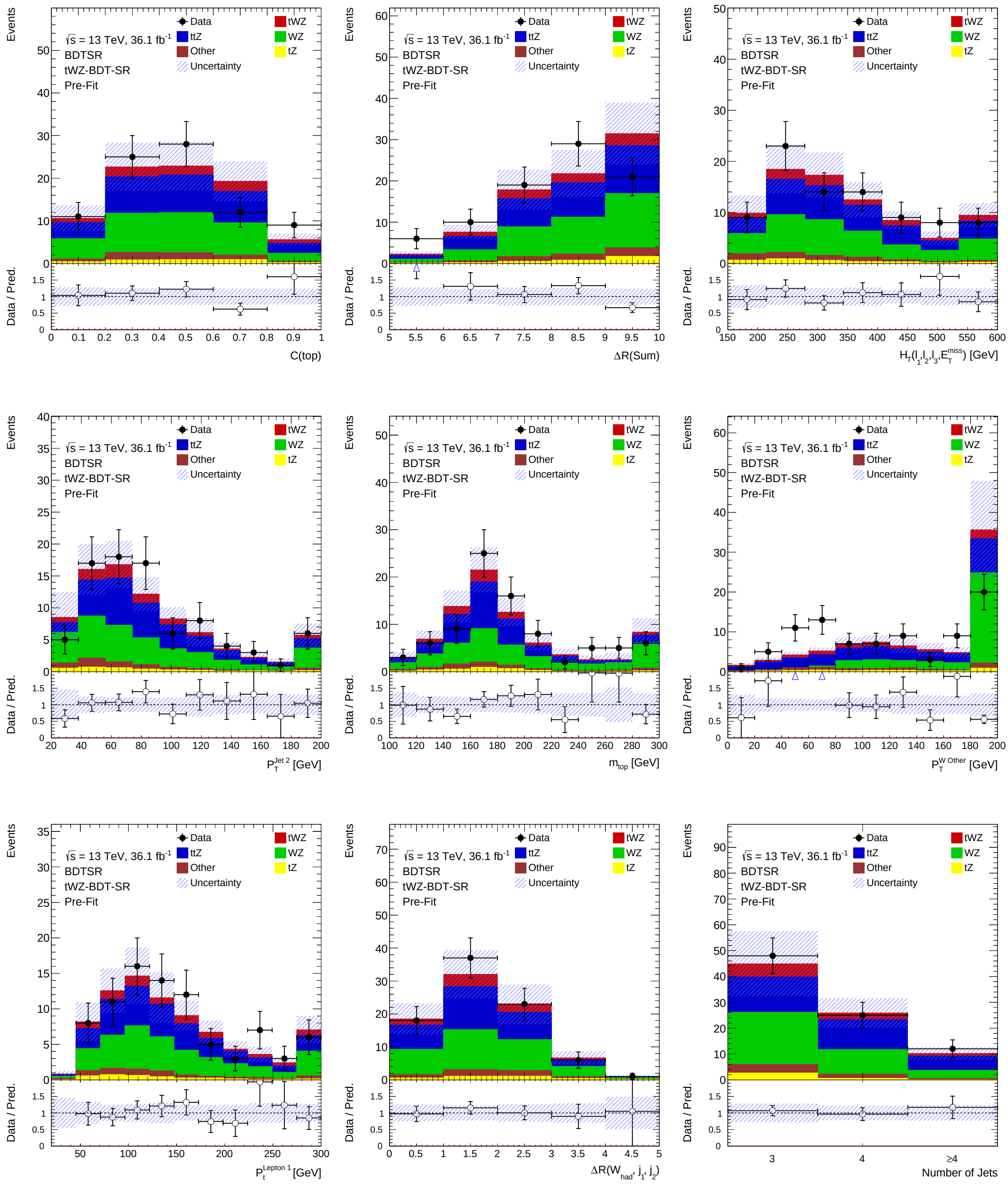


Figure A.1: Pre-Fit validation plots for BDT-SR. Both statistical and systematical uncertainties are included. The overflow is included in the last bin of each histogram.

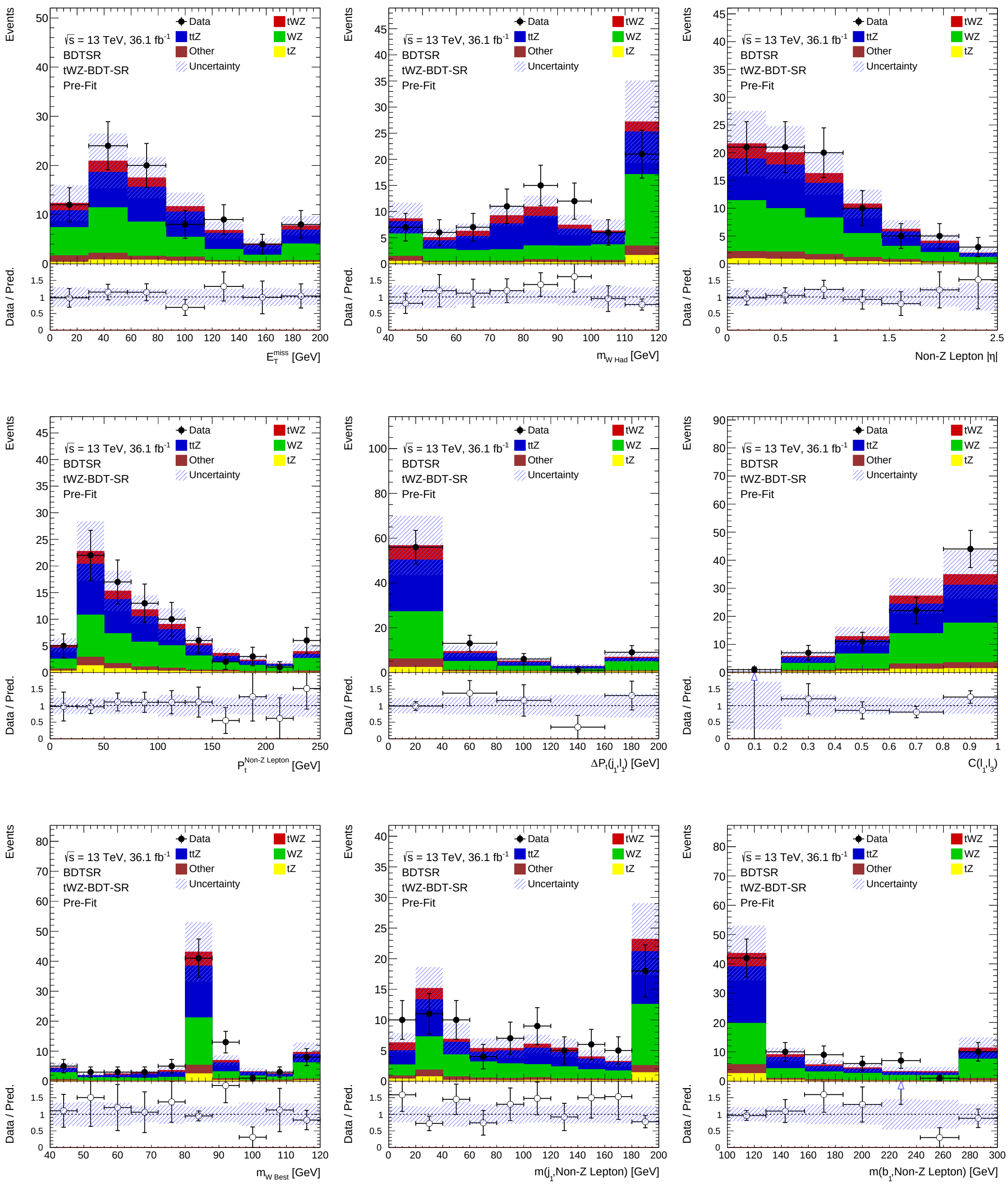


Figure A.2: Pre-Fit validation plots for BDT-SR. Both statistical and systematical uncertainties are included. The overflow is included in the last bin of each histogram.

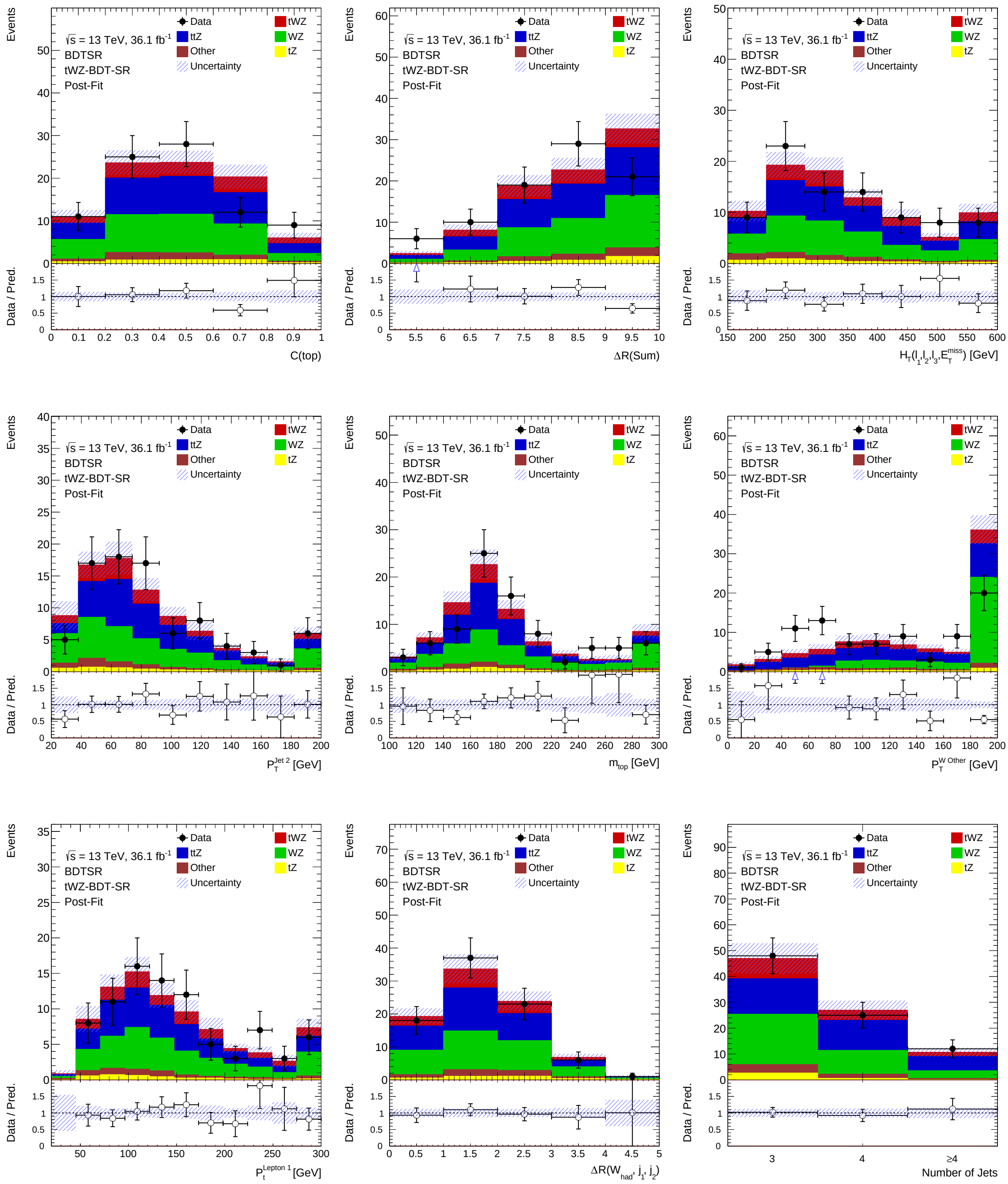


Figure A.3: Post-Fit validation plots for BDT-SR. Both statistical and systematical uncertainties are included. The overflow is included in the last bin of each histogram.

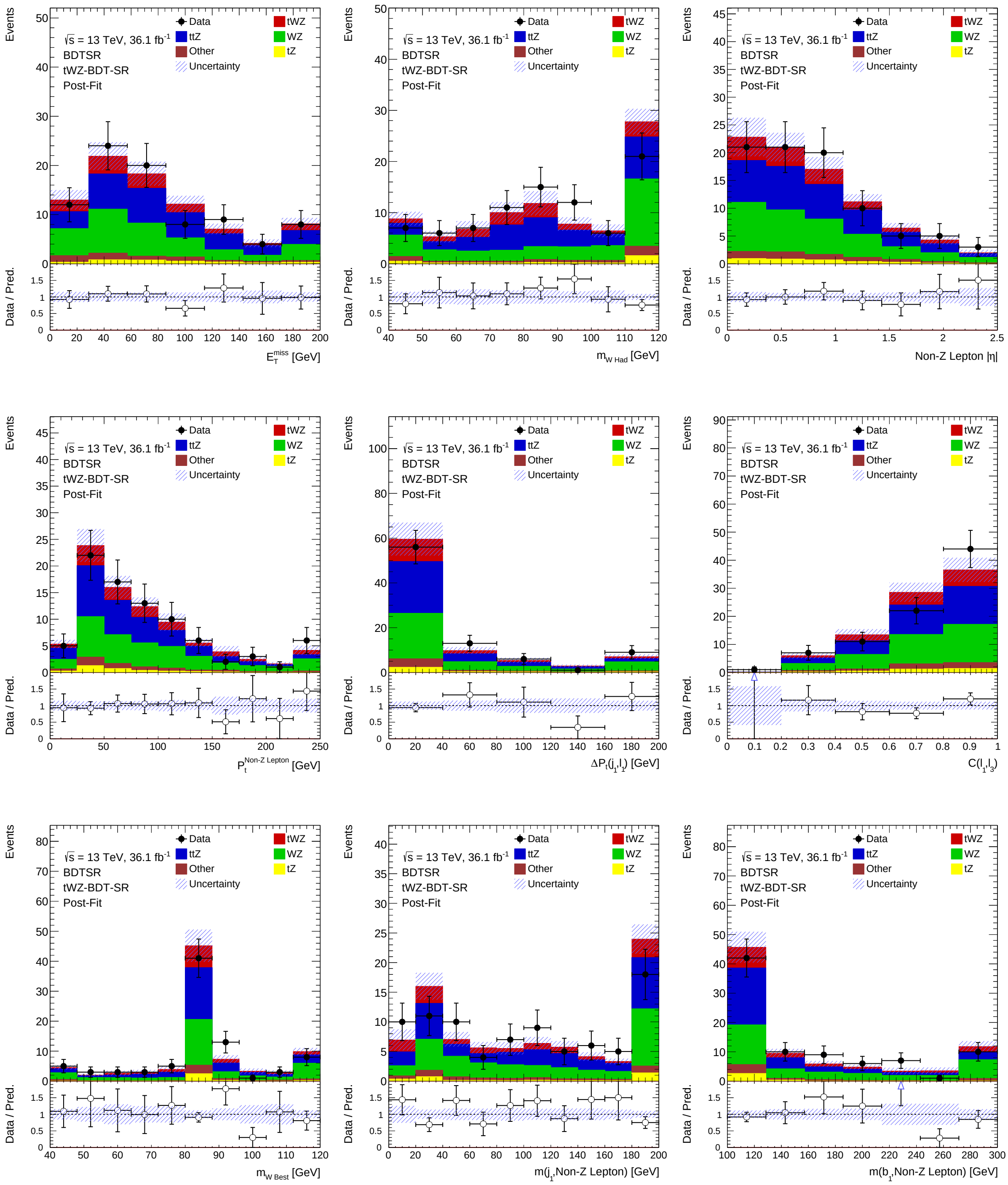


Figure A.4: Post-Fit validation plots for BDT-SR. Both statistical and systematical uncertainties are included. The overflow is included in the last bin of each histogram.

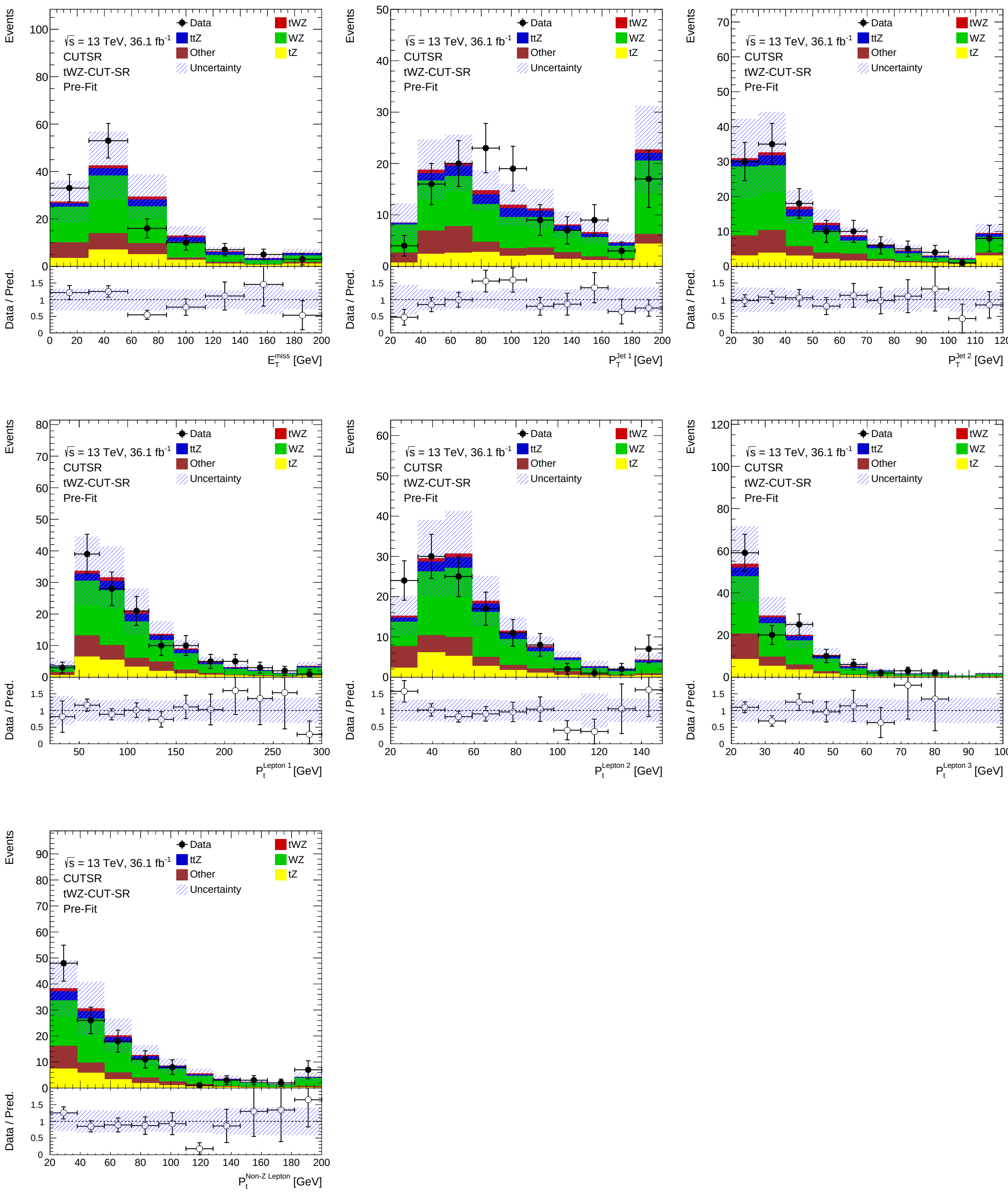


Figure A.5: Pre-Fit validation plots for CUT-SR. Both statistical and systematical uncertainties are included. The overflow is included in the last bin of each histogram.

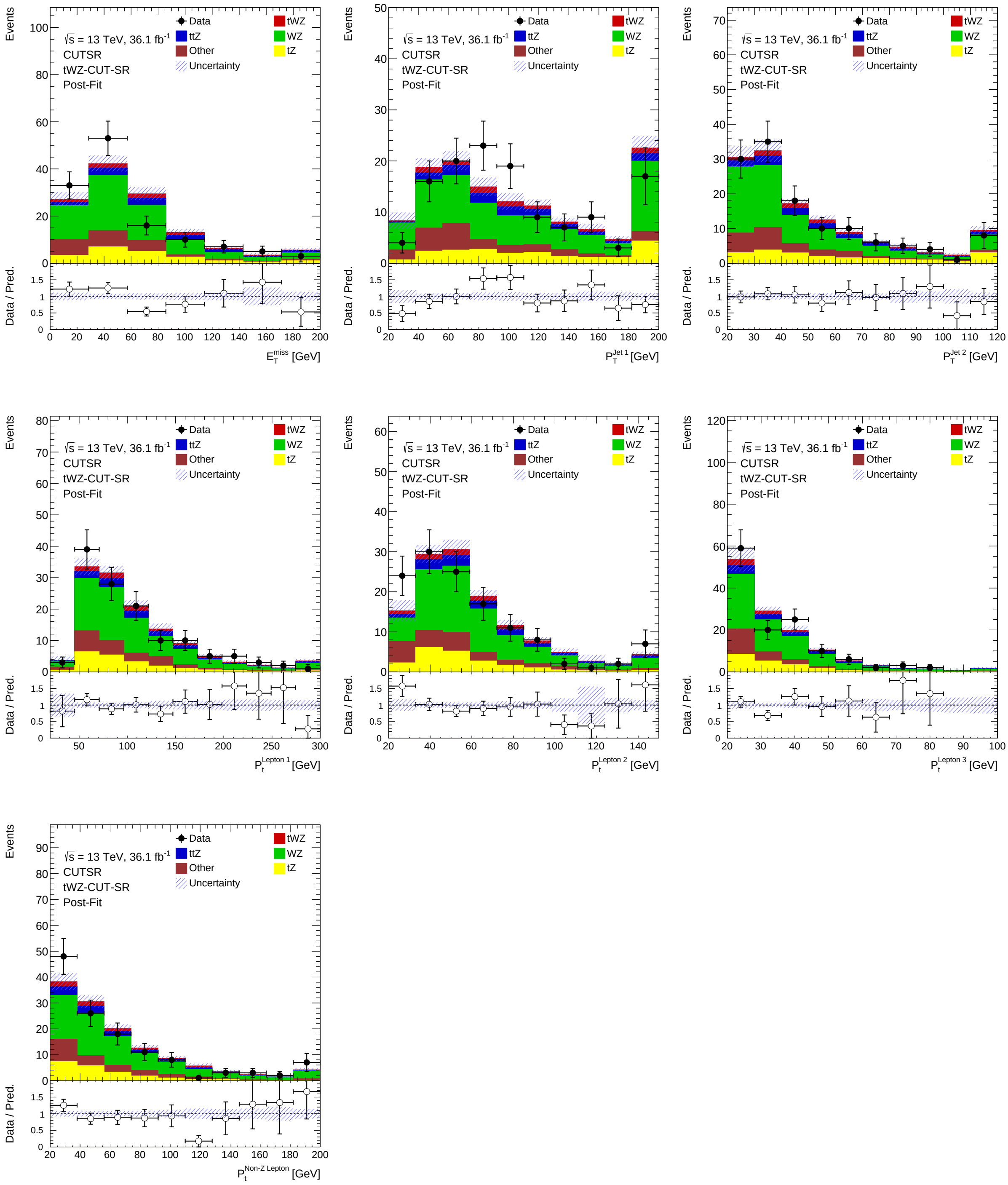


Figure A.6: Post-Fit validation plots for CUT-SR. Both statistical and systematical uncertainties are included. The overflow is included in the last bin of each histogram.

APPENDIX B

List of Monte-Carlo Simulation Samples

PROCESS	DATASET NUMBER	Notes
tWZ	410217	Nominal
	410215	Shower Systematics (Pythia)
$t\bar{t}Z$	410218	$Z \rightarrow$ee
	410219	$Z \rightarrow \mu\mu$
	410220	$Z \rightarrow \tau\tau$
WZ	361071	$\ell\ell\ell\nu$jj
	363491	$\ell\ell\ell\nu$
tZ	410050	
$t\bar{t}$	410503	
$t\bar{t}H$	343365	Dileptonic
	343366	Semileptonic
	343367	All hadronic
$t\bar{t}W$	410155	
ZZ	361072	$\ell\ell\ell\ell$jj
	361073	gg$\ell\ell\ell\ell$
	363490	$\ell\ell\ell\ell$
Z+jets	364100 to 364113	Z$\rightarrow \mu\mu$
	364114 to 364127	Z$\rightarrow$ee
	364128 to 364141	Z$\rightarrow \tau\tau$
Triboson	361620	WWW$\rightarrow$3ℓ3v
	361621	$WWZ \rightarrow 4\ell 2\nu$
	361622	$WWZ \rightarrow 2\ell 4\nu$
	361623	$WZZ \rightarrow 5\ell 1\nu$
	361624	$WZZ \rightarrow 3\ell 3\nu$
	361625	$ZZZ \rightarrow 6\ell 0\nu$
	361626	$ZZZ \rightarrow 4\ell 2\nu$
	361627	$ZZZ \rightarrow 2\ell 4\nu$

Table B.1: List of simulation samples used in this thesis.

Process	Dataset Number	Notes	
Z+γ	301535	Z+γ →eeγ	p_{T}: 10-35 GeV
	301536	Z+γ →$\mu\mu\gamma$	p_{T}: 10-35 GeV
	301887	Z+γ →e$\nu\gamma$	p_{T}: 10-35 GeV
	301888	Z+γ →$\mu\nu\gamma$	p_{T}: 10-35 GeV
	301889	Z+γ →$\tau\nu\gamma$	p_{T}: 10-35 GeV
	301890	Z+γ →e$\nu\gamma$	p_{T}: 35-70 GeV
	301891	Z+γ →e$\nu\gamma$	p_{T}: 70-140 GeV
	301892	Z+γ →e$\nu\gamma$	p_{T}: >140 GeV
	301893	Z+γ →$\mu\nu\gamma$	p_{T}: 35-70 GeV
	301894	Z+γ →$\mu\nu\gamma$	p_{T}: 70-140 GeV
	301895	Z+γ →$\mu\nu\gamma$	p_{T}: >140 GeV
	301896	Z+γ →$\tau\nu\gamma$	p_{T}: 35-70 GeV
	301897	Z+γ →$\tau\nu\gamma$	p_{T}: 70-140 GeV
	301898	Z+γ →$\tau\nu\gamma$	p_{T}: >140 GeV
	301899	Z+γ →eeγ	p_{T}: 35-70 GeV
	301900	Z+γ →eeγ	p_{T}: 70-140 GeV
	301901	Z+γ →eeγ	p_{T}: >140 GeV
	301902	Z+γ →$\mu\mu\gamma$	p_{T}: 35-70 GeV
	301903	Z+γ →$\mu\mu\gamma$	p_{T}: 70-140 GeV
	301904	Z+γ →$\mu\mu\gamma$	p_{T}: >140 GeV
	301905	Z+γ →$\tau\tau\gamma$	p_{T}: 35-70 GeV
	301906	Z+γ →$\tau\tau\gamma$	p_{T}: 70-140 GeV
	301907	Z+γ →$\tau\tau\gamma$	p_{T}: >140 GeV
	304776	Z+γ →$\tau\tau\gamma$	p_{T}: 10-35 GeV

Table B.2: List of simulation samples used in this thesis (Continued).

APPENDIX C

Nuisance Parameters

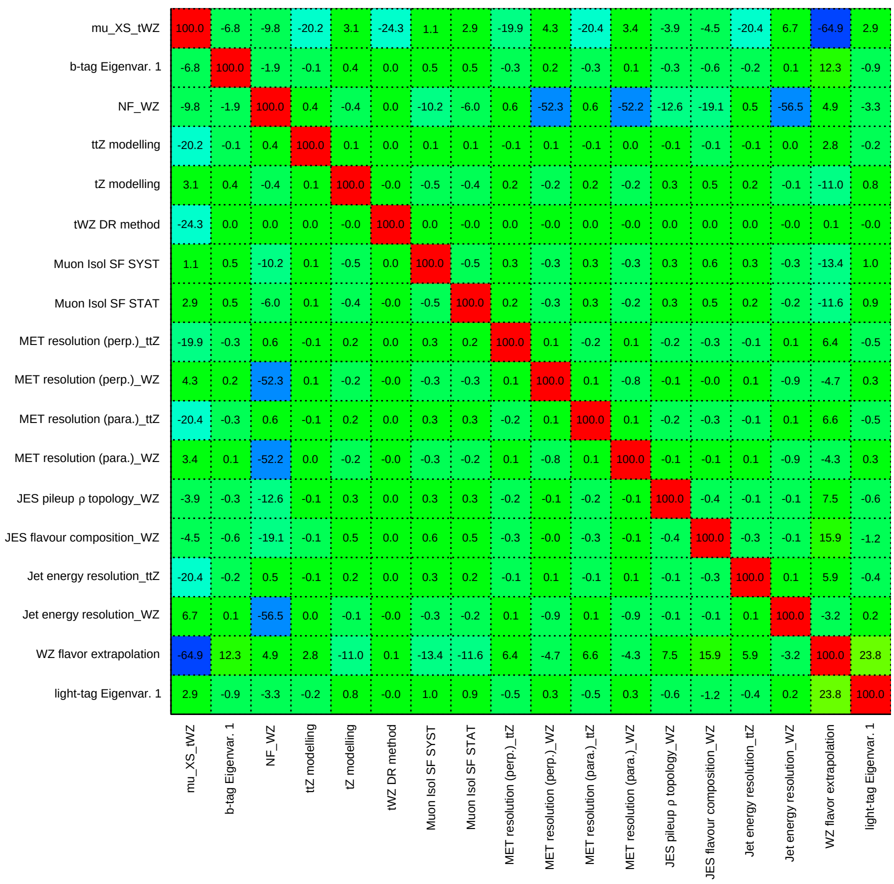


Figure C.1: Correlation matrix for systematic uncertainties.

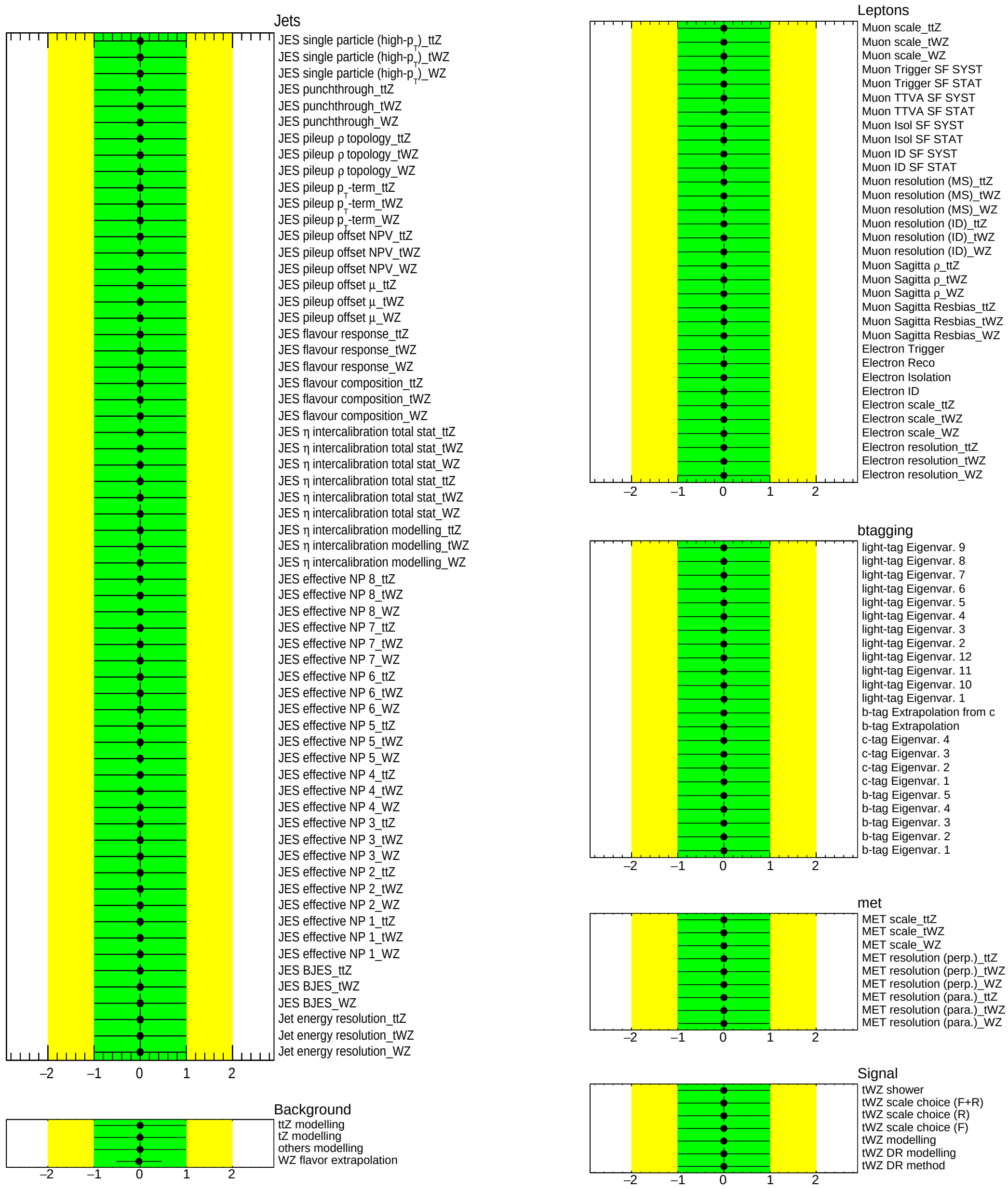


Figure C.2: Nuisance parameters for different categories of systematic uncertainties. Green and yellow belts correspond to $\pm 1\sigma$ and $\pm 2\sigma$ gaussian variations respectively.

Acknowledgements

First of all, I would like to express my gratitude for PD Dr. Markus Cristinziani for giving me this opportunity. Under his supervision, I have been able to learn a lot in a relatively short amount of time, explore my curiosities, and get a taste of scientific research. Aside of physics, I also learned much from his experience and insight on the social and institutional workings of academia. His cheerful, friendly, helpful character and fair treatment made this period as stress-free as it could be. I also thank him for sending me to the events in DESY, CERN, Dortmund and Würzburg, where I had the opportunity to present my work in a talk at DPG Spring Meeting.

I would like to thank Prof. Dr. Jan Jolie for overlooking my thesis despite the topic being out of his field of expertise. He was always quick to respond to my questions, helpful with mediation of official paperwork and thesis requirements between Bonn and Cologne. I would like to thank Prof. Dr. Norbert Wermes for hosting our group and his management of the Physics Institute in Bonn. I have benefited a lot from the positive atmosphere, collaboration and dialogue between the working groups, regular seminars and colloquia. I also would like to thank my academic mentor in Cologne, Prof. Dr. Achim Rosch for his help in dealing with requests of *Ausländeramt*, his valuable counsel on studies and for pointing me out about the possibility of working in particle physics in Bonn.

During my year in Bonn, I have had the good luck to be a part of team consistent of friendly and competent people. Their contributions were crucial to the completion of this work: I would like to thank Andrea Sciandra for his clear explanations on many topics, especially about the BDT and its computational implementation. He generously provided tools and shared shortcuts that helped me a lot. I also thank him for occasional chats about science, art, culture and history. I would like to thank my office mate Arshia Ruina for her help in debugging and cross checking codes as well as for nice talks, food and her guidance at CERN. I would like to thank Carlo A. Gottardo for his explanations of many concepts in analyses, providing macros and especially for his patience and efforts in teaching me C++ and ROOT during my first weeks. I also thank him for nice chats out of physics and contributing to my Italian vocabulary. I would like to thank Dr. Julien Caudron for always being available, willing to share his diverse, extensive knowledge and for his humble personality. Additionally, I am grateful for his time and efforts in reading my thesis, suggesting corrections and helping me with colloquium preparations. I would like to thank Sebastian Heer for always being available to help me at each step of the way, both in academic and in general contexts. I am grateful for assistance with paperwork, academic writing and presentations throughout. His patience, experience and optimism was invaluable to me. I would like to thank Dr. Vadim Kostyukhin for providing clear explanations to challenging and complex topics, and for his useful comments and suggestions which were very helpful in noticing the otherwise overlooked, important details. I would also like to thank Dr. Olga Bessidskaia Bylund for the useful discussions we had at CERN, David Hohn for sharing his expertise in TRExFitter program, IT teams in Bonn and at CERN for making our daily work proceed smoothly.

During my residence in Cologne, I have had many new friends. I am grateful to Anıl, Mehmet and Melek for all of their help and support. They became my second family here and were always generous and kind to me. Also, I would like to thank Chiara, Julia, Melanie and Sarah for their friendship that made my life in Germany much easier, colorful and pleasant.

I would like to thank all my friends for their great support and understanding during my studies, especially to Beyza, Ceylan, Duygu, Selcan and Kemal. I would like to thank Hakan Uzun for teaching me the value of patience and perseverance. I would like to thank all members of the Demirtaş family for their continuous support. Their presence have been reassuring and gave me peace of mind while away from home. I am also grateful to the Damschen family for their positive attitude and emotional support.

I would like to thank my brother for his support, useful advice and his genuine interest in my studies. I would like to thank my aunts and cousins for their love and support.

I am forever grateful to Amelie for her love, care and endless optimism. She provided all the support I needed during this period. She made sure that I don't get detached from life because of my studies, creating many moments and breaks that were refreshing and relieving. She managed to keep my feet on the ground and put a smile on my face, even when it seemed quiet impossible to me.

Finally, I would like to thank my mother and grandmother for valuing education above all, for their sacrifice, sincere and selfless efforts whilst raising and supporting me throughout my life. They have installed in me that I should only pursue an opportunity if I can actualise it in a fair, kind and honest way- a lesson I have aspired to carry with me throughout my studies here. Beyond all, I thank them for this most important lesson and gift in my life.

Erklärung

Ich versichere hiermit, dass die vorliegende Masterarbeit selbständig verfasst und keine weiteren als die angegebenen Hilfsmittel benutzt sowie die Stellen der Arbeit, die in anderen Werken dem Worlaut oder dem Sinn nach entnommen sind, durch Angaben der Quellen sichtbar gemacht wurden.