DEPARTMENT OF PHYSICS
UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA • UNIVERSITEIT VAN KAAPSTAD

Search for tWZ production in the Full Run 2 ATLAS dataset using events with four leptons

Jake Reich
Student Number: RCHJAK001
Supervisor: Dr. James Keaveney
Co-Supervisor: Dr. Sahal Yacoob

October 2021

Abstract

Declaration

I certify that this assignment/report is my own work, based on my personal study and/or research and that I have acknowledged all material and sources used in its preparation, whether they be books, articles, reports, lecture notes, and any other kind of document, electronic or personal communication. I also certify that this assignment/report has not previously been submitted for assessment in any other unit, except where specific permission has been granted from all unit coordinators involved, or at any other time in this unit, and that I have not copied in part or whole or otherwise plagiarised the work of other students and/or persons.

Acknowledgements

Contents

Chapter 1

Introduction

Write similar to what is in nrf application.

Talk about previous paper's (tWZ 3-lep) findings - http://cds.cern.ch/record/2625170

Explain that SM aims to describe fundamental physics, but fails in certain cases (DM, gravity etc.)

Possibly talk about EFT? Finding tWZ cross section -¿ global fit. FOR REFERENCE:
The production of a single top quark in association with a W± and Z boson (tW±Z) is sensitive to both the neutral and charged electroweak couplings of the top quark as the process involves the simultaneous production of a W boson and a Z boson in association with the top quark. Due to the very large coupling of the top quark to the Higgs boson, the electroweak couplings of the top quark are a theoretically well-motivated area to expect the first signs of new physics. The recent lack of signs of new physics from LHC data tells us that new physics is either very heavy, or is very weakly coupled to Standard Model particles, therefore we might only observe signs of new physics in anomalous rates of well-chosen processes. A prime example of such a process is tWZ. This has an extremely low production cross section (0.7 fb), meaning that it is an extremely rare process to observe and subsequently, it has never been observed by any particle physics experiment. However, the latest datasets recorded by ATLAS are sufficiently large to allow a potential observation of this rare, novel process.

We aim to use the Full Run 2 dataset recorded by the ATLAS experiment at CERN to search for the production of a top quark together with a W± and Z boson for the channel with four leptons (two originating from the decay of the Z boson, one from the associated W boson and one from the W boson which decays from the top quark (together with a b quark)). The Standard Model of particle physics has been confirmed to an extraordinary degree of precision, however we know there are stark deficiencies therein. These include its incompatibility with the theory of gravity and an explanation of the matter-antimatter asymmetry in the universe. Especially relevant is the Standard Model's lack of an explanation for the vast differences in the strengths of the fundamental forces (The Hierarchy Problem), constraining the electroweak couplings of the top quark squarely addresses this fundamental scientific question.

Chapter 2

Theory

2.1 Standard Model of Particle Physics

What is SM (renormalisable qft), come from symmetry, brief description of group structure? Explain structure/properties (fermions, bosons, etc.), coupling constants. 'Particles carry colour/electric charge, some particles can interact with others/themselves, some can't' Where does it how? How well does it work? Where doesn't it work?

2.1.1 Electroweak Theory

Properties of W and Z. Decay channels. Z as the standard candle (distinct OSSF lepton signal).

2.1.2 Top Quark

Properties. History (when/how was it discovered/theorised). Why interesting (large mass). Hierarchy problem. Decay channels. Extremely short lifetime (makes b's so important for top ID).

2.2 tWZ

2.2.1 Tetra-lepton Channel

Provide Cross section.

The leading order Feynman diagram for tWZ in the tetra-lepton channel is shown below.

Backgrounds

The main backgrounds for tWZ (tetra-lepton channel) are the production of a two tops, both in the $\ell\nu b$ final state channel, together with a Z boson ($t\bar{t}Z$) and diboson production with fully leptonic final states (ZZ).

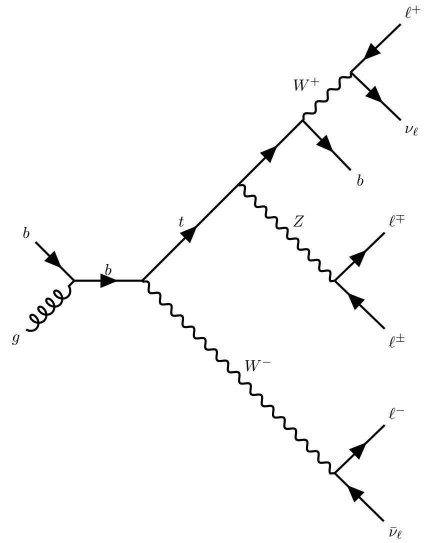

Figure 2.1: Example Feynman diagram of tWZ production in the tetra-lepton channel.

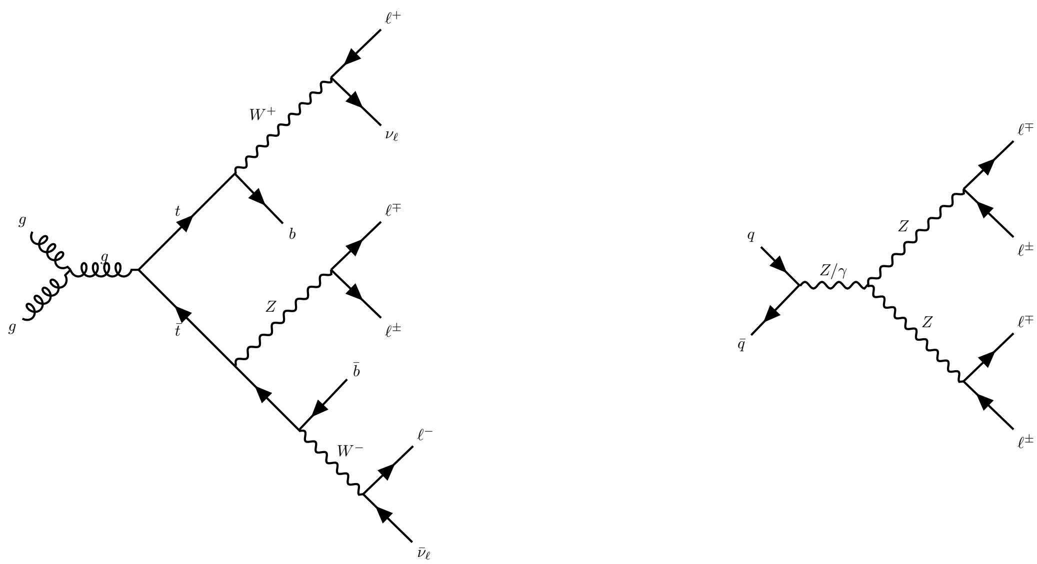

Figure 2.2: Example Feynman diagrams for $t\bar{t}Z$ (left) and ZZ (right) in the tetra-lepton channel.

2.2.2 Comparison to Tri-lepton Channel

Less backgrounds to deal with (in tetralepton). However lower stats (in tetra). Give cross sections (and feynman diagram). Maybe talk a bit about analysis related challenges (trilepton has a hadronically decaying W, does this make the analysis easier or more difficult?).

The most apparent difference between the tri and tetra-lepton channels is the amount of statistics present, with the tetra-lepton channel having far less events in its phase space than that of the tri-lepton channel. The lack of statistics in the tetra-lepton channel can be attributed to its low production cross section, $\sigma^{\mathrm{NLO}}_{(tW^{\pm}Z).Br(4\ell)} = 0.7\,\mathrm{fb}$[5]. The tri-lepton channel has a production cross section ($\sigma^{\mathrm{NLO}}_{(tW^{\pm}Z).Br(3\ell)} = 3.9\,\mathrm{fb}$[5]) around a factor of 4 larger than that of the tetra-lepton channel. This difference between the production cross section of the two decay channels can be largely attributed to the difference in branching ratios ($\frac{\Gamma_i}{\Gamma}$) between a hadronically decaying W boson, $\frac{\Gamma_{W\rightarrow had}}{\Gamma_W} = (67.41 \pm 0.27)\%$[3], present in the tri-lepton channel and a leptonically decaying W boson, $\frac{\Gamma_{W\rightarrow \ell\nu}}{\Gamma_W} = (10.86 \pm 0.09)\%$[3], present in the tetra-lepton channel.

Despite the tetra-lepton channel's low statistics, it is not subject to the large WZ background present in

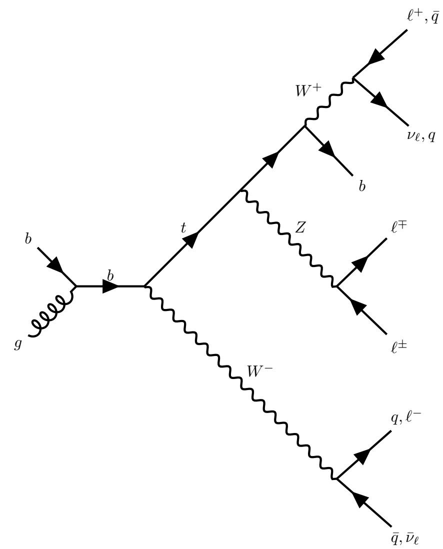

Figure 2.3: Example Feynman diagram of tWZ production in the tri-lepton channel.

the tri-lepton channel.

2.3 Effective Field Theory (EFT)

Brief overview of EFT. What is EFT? Why important to pp as a whole? Why important in twz (high sensitivity to wilson coefficients, expected to have a large impact on global fit)? Similar to what james says in INT note.

2.4 Machine Learning in the Context of Particle Physics Analyses

Brief overview of ML as a whole; History; increase in popularity in recent years (why increased in popularity $\rightarrow$ novel techniques developed, increase in computing power for your buck). Where does it fit into pp (event selection, object reconstruction and ID (jet reco, b-tagging)). Explain concepts (vocabulary), overtraining, training, testing, classifier, classification. Why use x train/test ratio in pp (use some for analysis/use some for training)? Popular tools which are used today (scikit learn, TMVA, xgboost what is theano, what is keras).

Maybe a subsection on bdt (if end up using it) on the specific algorithm, and minimizing cost function (general). Explanation on ROC curve and why we can use it as a proxy to determine how well our bdt/nn is doing (and where it fails/can fail $\rightarrow$ things to be aware of cautious of when straight up using ROC integral (e.g. overtraining)).

2.5 Statistical Techniques

Brief overview: frequentist and bayes approach in general in pp, why we use frequentist in this analysis

2.5.1 Maximum Likelihood Fitting

Go through theory. Varying histograms.

2.5.2 p-value and χ^2

Go through theory. Go through story of getting p-value, what it means, getting chi-squared, what it means, where the chi squared distribution comes in, degrees of freedom, what is a 'good' chi squared value and what is 'bad' and what do different values mean/infer/suggest.

2.5.3 Significance

Go through theory. What is means/interpreted as. 3 sigma observation, 5 sigma discovery.

2.5.4 Limit Setting

Go through theory.

Chapter 3

The ATLAS Experiment and Detector

3.1 The ATLAS Experiment

Brief overview of history of ATLAS and CERN in general. State accomplishments (higgs discovery - completed list of SM particles).

3.1.1 Large Hadron Collider (LHC)

Topology, parts of it.

3.2 The ATLAS Detector

3.2.1 Coordinate System and Kinematics

Physical dimensions, properties (mass). Coordinate system. Definition of ΔR, η.

In the next subsections:
Where are different particles detected? How are they detected at each part of the detector? How well do the different parts detect the different particles and how do they take advantage of different particle properties each part wishes to detect in an engineering/physics perspective (e.g. why we WANT some particles to get detected at some parts of the detector and how we do this (by use of correct materials which interact specifically with those particles which we wish to measure and NOT with other particles))

3.2.2 Tracking Detectors

3.2.3 Calorimeter System

Electromagnetic Calorimeter

Hadronic Calorimeter

3.2.4 Muon Spectrometer

3.2.5 Trigger and Data Acquisition System

3.2.6 Particle Identification

Chapter 4

Analysis Setup and Strategy

4.1 Data and Monte Carlo Simulation

4.1.1 Data Samples

Brief overview of Full Run 2 and where the data comes from and what time period.

Luminosity of full run 2.

4.1.2 Monte Carlo Samples

Signal and Background samples. Choice of cuts at ntuple level. Specifically why each background is chosen (and why others excluded) → e.g. talk about branching fractions, cross sections, topology. How are these backgrounds passing our event selection (e.g. ttz → b can be lost/untagged/mis-id'ed) → provide an explanation for each background.

Details of each sample (event generator, parton shower).
The following background processes are considered:

- **t$\bar{\mathbf{t}}$Z**: $t\bar{t}$ with an associated Z-boson, in the tetralepton final state. Therefore, both top-quarks decay leptonically (e.g. $t \to W^+b \to \ell^+\nu b$) and of these top-quarks emits a Z-boson which decays leptonically ($Z \to \ell^\pm\ell^\mp$ (OSSF lepton pair)). This results in a final state with 4 leptons and 2 b-quarks.
- **ZZ**: Diboson production with a tetralepton final state, therefore both Z-bosons decay leptonically ($Z \to \ell^\pm\ell^\mp$ (OSSF lepton pair)).
- **other**: Processes with a relatively minimal, but non-negligible background contribution in the tWZ signal region
  - Triboson:

    $WZZ \to \ell\ell\ell\ell\ell\nu$
    $ZZZ \to \ell\ell\ell\ell\ell\ell$
    $ZZZ \to \ell\ell\ell\ell\nu\nu$
    $WWZ \to \ell\ell\ell\ell\nu\nu$
  - tZq: Top quark in association with a Z-boson and another quark.

4.2 Objects

In the subsections below:

Explain why we applied each cut/selection.

4.2.1 Leptons

In this analysis we only consider e and μ leptons, since τ leptons are difficult to detect in the ATLAS detector. τ leptons are challenging to detect since they have an extremely short lifetime (290.3 ± 0.5fs [3]) which causes them to decay before reaching any detector components and therefore can only be reconstructed via their decay products.

In addition to our selection criteria of exactly four leptons, we require that the Leading (L), Next-to-Leading (NL), Next-to-Next-to-Leading (NNL) and Next-to-Next-to-Next-to-Leading (NNNL) leptons have p_T greater than 28, 10, 10 and 10 GeV respectively. Here we have chosen to apply relativity loose object-level cuts in an attempt to maximize our signal statistics, since the analysis is heavily statistically limited.

Reconstructed electrons are required to be within $|\eta| < 2.47$ and excluding the transition region between the barrel and end-cap calorimeters at $1.37 < |\eta| < 1.52$. Reconstructed muons are required to be within $|\eta| < 2.5$.

The transverse impact parameter, d_0, is defined as the minimal spacial distance between the object's (here we are referring to leptons) trajectory and the primary vertex (the vertex associated with the p-p hard scatter). The longitudinal impact parameter, z_0, is defined as the value of z of the point on the object's trajectory which determines d_0. To ensure consistency between the lepton and the primary vertex, we require that $|\frac{d_0}{\sigma(d_0)}| < 5$, $|z_0 \sin\theta| < 0.5\,\mathrm{mm}$ for electrons and $|\frac{d_0}{\sigma(d_0)}| < 3$, $|z_0 \sin\theta| < 0.5\,\mathrm{mm}$ for muons, following the current recommendations [**TopRecoObjTwikiModel**].

To avoid instances where one detector signal can result in multiple different reconstructed objects, an overlap removal is applied which ignores all but one of these objects. We use the current recommended configuration [**Overlap-removal-slides**].

Electrons are selected using a likelihood based discriminant [**electronRecoAndID:paper**] which takes measurements from the tracking system, calorimeter system and quantities derived from both the tracking and calorimeter system as input. Muons are selected using AnalysisBase's `Muon Selection Tool` [**muon-selection-tool**].

Loose electrons are defined with the criteria above, using the `LooseAndBLayerLH` identification working point. Similarly, tight electrons are defined with the criteria above, using the `TightLH` identification working point. Both loose and tight muons use the `Medium` identification working point.
Tight leptons additionally require that they are sufficiently isolated from other particles produced in the collision. This is done by defining a cone of radius $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ around the particle of interest and summing the p_T of all the reconstructed particles surrounding the particle of interest, situated within the cone. I_{rel} is defined as ,$I_{rel} = \frac{\Sigma p_T(\text{surrounding candidate})}{p_T(\text{candidate})}$, the ratio of this sum to the p_T of the lepton candidate. If this value is large, is it likely that the particle of interest originated from a jet (together with many other particles), whereas a prompt decay product resulting from the hard scatter will have little to no energy surrounding it ($I_{rel} \ll 1$). We use Analysis Base's `IsolationSelectionTool` with the `PLVTight` and `PflowTight_FixedRad` for tight electrons and tight muons respectively (following the current recommendations [**recommendedIsolationWPs**]).

The selection criteria for leptons are summarised in the Table **??** below.

4.2.2 Jets

Jets are required to be within $|\eta| < 2.5$ and $p_T(\text{jet}) > 25\,\mathrm{GeV}$. We apply these looser p_T cuts in an attempt to increase our limited signal statistics.

We use the `Medium` working point for the jet-vertex-tagger (jvt) and the forward jet-vertex-tagger (forward jvt). Jets are required to have a jvt value greater than 0.5, in an attempt to reject effects caused by pile-up interactions.

4.2.3 b-tagging

The DL1r b-tagger [4] was used to identify jets as b-jets. The DL1r algorithm combines outputs from several low-level tagging algorithms using a Deep Neural Network and outputs the probability that a given input jet is identified as a b, c or light flavoured jet. We use different DL1r working points to identify b-jets in our event selection (See Section 4.4). The working points are defined based off a cut on the DL1r score corresponding to a b-jet tagging efficiency of 60%, 70%, 77% and 85%.

Since we are heavily statistically limited, we aim to increase the amount of statistics in our four regions. In an attempt to achieve this goal, b-tagged jets were placed under *tight* and *loose* definitions. A tight b-tagged jet is defined as a jet which passes the 77%, 70%, 65% or 60% DL1r b-tagger working point. A loose b-tagged jet is defined as a jet which passes 85% DL1r b-tagger working point, but not the 77%, 70%, 65% or 60% DL1r b-tagger working points. Different numbers (and definitions) of tight and loose b-tagged jets were tried in each region, with the final selection criteria being chosen which maximised the expect significance of tWZ.

4.3 Kinematic Pre-selection cuts

The invariant mass of the OSSF lepton pair coming from the Z boson must equal the invariant mass of the Z boson, and noting that e,μ reconstruction and identification in the ATLAS detector has a high accuracy [], we can use these OSSF leptons to reconstruct Z bosons with relatively high confidence. We therefore define a Z candidate as an OSSF lepton pair with an invariant mass, m_{OSSF}, satisfying the condition, $|m_{\text{OSSF}} - m_Z| < 30\,\text{GeV}$, where m_Z is the nominal Z boson mass (91.1876 GeV [3]). Multiple Z candidates can be present in certain decay channels (e.g. $eeee$, $\mu\mu ee$, $\mu\mu\mu\mu$). In these cases, the Z candidate which has an invariant mass closest to the nominal Z boson mass is chosen.

In order to suppress potential fakes and quarkonia (low mass resonances such as J/ψ and upsilon) we require that all OSSF lepton pairs have an invariant mass, m_{OSSF}, greater than 10 GeV.

Due to conservation of charge, the final state lepton charges must sum to zero.
We therefore require $\displaystyle\sum_{i=1}^{4} \text{charge}(\ell_i) = 0$.

4.4 Regions and Event Selection

The selection criteria which define the SR and the CRs are summarised in Table 4.1. In order to check the modelling of the most dominant background components in our signal region, we have modified our selection criteria to define $t\bar{t}Z$ and ZZb control regions. The $t\bar{t}Z$ control region has the same requirement on the number of reconstructed Z boson candidates in the signal region (due to a commonality on the number of Z bosons present in both processes), however we require at least two jets and that exactly two of these jets are b-tagged (corresponding to the b-quark jets originating from the two top-quark decays). We choose to define a ZZb region, as opposed to a ZZ region, since the ZZ background present in the tWZ signal region contains exactly one b-tagged jet. Therefore defining a region with ZZ plus exactly one b-jet more closely resembles the ZZ background present in the signal region. In addition to this, mis-modelling of ZZ has been seen in other analyses [1, 2], further motivating the use of a ZZb control region over a ZZ CR. The ZZb CR requires exactly two Z boson candidates and exactly one b-tagged jet, with no requirement on the number of jets. In order to constrain the fake lepton component contained within the $t\bar{t}Z$ sample, we define a $t\bar{t}Z$ fake CR which is as similar as possible to the $t\bar{t}Z$ CR but is enhanced in fakes. This is achieved by defining the $t\bar{t}Z$ fake CR to inherit the same selection criteria as the $t\bar{t}Z$ CR however, in this case, we require exactly 3 tight leptons and exactly 1 loose lepton (since loose leptons are more likely to be fakes, compared to tight leptons).

The final selection criteria and region definitions are summarised in Table 4.1 below.

Common selections			
$N_\ell = 4$ $p_T(\ell_1, \ell_2, \ell_3, \ell_4) > (28, 10, 10, 10)$ GeV $p_T(\text{jet}) > 25\,\text{GeV}$, $\lvert\eta(\text{jet})\rvert < 2.5$, jvt > 0.5 $\lvert\eta(\ell_e)\rvert < 2.47$ excluding $1.37 < \lvert\eta(\ell_e)\rvert < 1.52$ $\lvert\eta(\ell_\mu)\rvert < 2.5$ $\sum_{i=1}^{4} \text{charge}(\ell_i) = 0$ All OSSF lepton pairs require $m_{\text{OSSF}} > 10\,\text{GeV}$			
tWZ SR	$t\bar{t}Z$ CR	ZZb CR	$t\bar{t}Z$ fake CR
$N_\ell(\text{tight}) = 4$	$N_\ell(\text{tight}) = 4$	$N_\ell(\text{tight}) = 4$	$N_\ell(\text{tight}) = 3$ $N_\ell(\text{loose}) = 1$
$N_{\text{Z candidate}} = 1$	$N_{\text{Z candidate}} = 1$	$N_{\text{Z candidate}} = 2$	$N_{\text{Z candidate}} = 1$
$N_{\text{jet}} \geq 1$	$N_{\text{jet}} \geq 2$	no requirement	$N_{\text{jet}} \geq 2$
$N_{\text{b-jet}}(\text{tight}) = 1$	$N_{\text{b-jet}}(\text{tight}) \geq 1$ $N_{\text{b-jet}}(\text{loose}) \geq 0$ $N_{\text{b-jet}}(\text{tight}) + N_{\text{b-jet}}(\text{loose}) = 2$	$N_{\text{b-jet}}(\text{tight}) = 1$	$N_{\text{b-jet}}(\text{tight}) \geq 1$ $N_{\text{b-jet}}(\text{loose}) \geq 0$ $N_{\text{b-jet}}(\text{tight}) + N_{\text{b-jet}}(\text{loose}) = 2$

Table 4.1: Summary of the requirements applied for selecting events in the signal and control regions.

4.5 Machine Learning Techniques

What tool did we use, how did we use it, parameters of bdt/nn, input variables/importance, conversion of event level bdt output (bdtscore) to variable for fitting. Used ROC curve integral as a proxy for how good bdt was doing.

Now that we have our baseline selections applied and our regions defined, we implement a Boosted Decision Tree (BDT) in order to discriminate between tWZ and our most prominent background process, $t\bar{t}Z$. We chose to use a BDT, as opposed to another ML algorithm, since they are very stable and perform well with minimal/no optimisation or tweaking of the hyper parameters. A multi-layered sequential neutral network was tried, however, it was out-performed by a BDT. More specifically, Scikit-Learn's `GradientBoostingClassifier` was used.

The BDT was trained on 50% of the tWZ MC sample's events for the signal class and similarly, 50% of the $t\bar{t}Z$ MC sample's events were used for the background class. The samples we train on are individual events, with the features being carefully chosen observables. These observables are chosen on the basis that they are somewhat uncorrelated from one another and show a relatively large amount of separation power between tWZ and $t\bar{t}Z$. Since we train on observables from individual events, we refer to this BDT as an *event-level* BDT. The optimum values for the hyper-parameters used were determined by training the BDT with a range of different values for the hyper-parameters and choosing the set of values which maximized the accuracy (based off cross-validation). This method is more commonly referred to as hyper-parameter optimisation or tuning. The observables and hyper-parameters used in training are summarised in Table 4.2 and Table 4.3 respectively.

Observable	Description
2νSM	Maximum weight from the 2νSM algorithm
$\Delta\eta(\ell\ell_{\text{non-Z}}, b_2)$	$\Delta\eta$ between the dilepton system not from a Z candidate and the Next-to-Leading b-tagged jet
HT	Scalar sum of jet p_T
$\Delta\eta(\ell_{1,\text{non-Z}}, \ell_{2,\text{non-Z}})$	$\Delta\eta$ between the Leading lepton not from a Z candidate and Next-to-Leading lepton not from a Z candidate
$\Delta R(b_1, Z_1)$	ΔR between the Leading b-tagged jet and the Leading Z candidate
$\Sigma p_T(b)$	Scalar sum of b-tagged jet p_T
$\Delta\eta(\ell\ell_{\text{non-Z}}, b_1)$	$\Delta\eta$ between the dilepton system not from a Z candidate and the Leading b-tagged jet

Table 4.2: A list of the observables used in the event-level BDT, ordered by importance (descending, top to bottom).

Hyper-parameter	Value	Description
`loss`	`deviance`	The loss function to be optimised
`criterion`	`friedman_mse`	The function used to measure the quality of a split
`n_estimators`	`100`	The number of boosting stages to perform
`learning_rate`	`0.1`	The step size at each iteration during optimisation
`max_depth`	`3`	The maximum depth of the individual regression estimators
`min_samples_split`	`2`	The minimum number of samples (events) required to split an internal node
`min_samples_leaf`	`1`	The minimum number of samples (events) required to be at a leaf node
`validation_fraction`	`0.1`	The proportion of training data to set aside as validation set for early stopping
`n_iter_no_change`	`20`	Training terminates when the validation score (determined by the validation set) does not improve in all of the previous `n_iter_no_change` number of iterations

Table 4.3: A list of the hyper-parameters used in the event-level BDT. Hyperparameters not listed in this table use the default values as stated in the Scikit-learn Documentation[6].

The number of events used in training for the signal and background classes were 31584 and 14010 respectively. Imbalanced datasets can cause ML classifiers to ignore small classes while concentrating on classifying large classes more accurately, which may result in the trained classifier performing sub-optimally. In order to correct this dataset imbalance, we ensure that the relative weighting of each event is such that the sum of the signal weights is equal to the sum of the background weights. This is done by using the Scikit-Learn's `utils.class_weight.compute_sample_weight` function.

The resulting BDT discriminator plots are shown below in Figure 4.1.

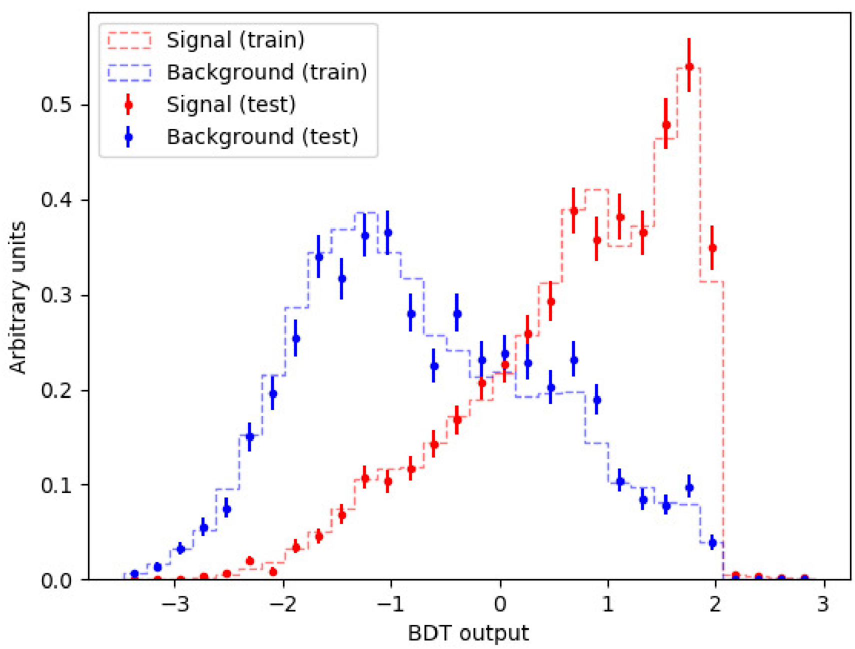

Figure 4.1: Normalised histograms of the BDT discriminator output from the signal and background classes for the training and test sets.

We can see that the shapes of the training and test sets for both signal and background are very similar. This is a good indicator that no overtraining occurred.

4.6 Fake Lepton Estimation

Expected to be a small effect, why? Brief, general explanation/idea of methods used (full explanation/description of what we did and the results/plots/etc. later)

4.7 Analysis Pipeline/Workflow and TRExFitter

What is TRExFitter? What can it do? At which stage(s) in the analysis did we use it? Which version did we use? Binning method. Explain calculation of error bars in TRF.

Include general flow chart of analysis (not sure where)
We make use of industry standard `ROOT`[1] wrappers in this analysis, namely, `PyROOT` and `TRExFitter`.

`Python` is used extensively in many fields of science (not limited to physics and data science) due to its simplicity and ongoing support by the communities which utilize it. `PyROOT` allows users to access the full `ROOT` functionality within `Python`. More specifically, `PyROOT` provides `Python` bindings for `ROOT`.

`TRExFitter` is a framework for binned template profile likelihood fits[7]. In this analysis, we used TRExFitter (tag: `TRExFitter-00-04-11`) to produce all pre-fit and post-fit plots (including fit statistics, e.g. limit, significance, $\mu_{best-fit}$).

The analysis pipeline starts with sample derivations (derived dataset) being submitted to the grid for ntuple production. This applies cuts and selections to the already reduced derivations and produces ntuples with trees containing variables (e.g. scale factors, observables, MC truth flags) that will be used at future stages in the analysis. These ntuples are then read by `PyROOT` where the events are looped over, before being written to `ROOT` files as input to `TRExFitter`. The `Python` script's main purpose is to define the different regions and apply the final cuts and selections outlined in Table 4.1. As each event is looped over, these cut and selection criteria are checked for the given event and is either thrown away, or gets written to a `ROOT` file corresponding to the MC sample and Full Run 2 data-set (mc16a, mc16d, mc16e) which it belongs to. `TRExFitter` then takes these files as input, runs a maximum likelihood fit and produces relevant plots (e.g. pre-fit, post-fit, pull plots) and statistical parameters (e.g. limit, significance, $\mu_{best-fit}$).

Throughout this analysis, we ensured that the signal region was kept blinded. We did this by implementing `TRExFitter's` *'mixed data and MC'* [8] fit, which aims to obtain the most accurate prediction for the expected results (while keeping the signal region blinded). It does this by performing a background only fit to the control regions (using real data). The set of fitted values for all the nuisance parameters from the background only fit are then used to construct a modified ASIMOV data-set. Finally, the fit is performed using real data in the control regions and the aforementioned modified ASIMOV data-set in the signal region.

[1] CERN's HEP data analysis framework (written in C++)

Chapter 5

Search for tWZ Production

5.1 Backgrounds

Maybe remove subsections below...

Yields table, signal percentage, background percentage in each region (and each sample)

5.1.1 $t\bar{t}Z$

5.1.2 ZZ

5.1.3 other

5.2 Control Plots

Fit variables only. Why these variables were chosen (maybe provide separation plots or just the separation values of different variables). Comment on data/mc agreement.

5.3 Post-Fit Plots

Plots, yields.

Input into fitting procedure (systematics, normfactors, nuisance parameters , why these were chosen) $\rightarrow$ does this go in a previous section?

What did fit do? Explain results from push-pull (e.g. the fit constrained (had a constraining effect) x background uncertainty)

5.4 Results

Commentary on results, limit on σ(tWZ), within uncertainties of SM?

Chapter 6

Conclusion and Outlook

Summary of study. Possible ways to improve/extend analysis. Better understand ttz background? Is there anything we can say about Z, t, b, W, whatever that we can take from the study (e.g. x is difficult to detect/discriminate in such analyses due to a and b, however it can be improved by doing y). Maybe talk about what preliminary studies showed, but couldn't fully implement that idea due to whatever restriction/limitation.

Appendix A

Appendix

Bibliography

[1] Aaboud, M. and Aad, G. and Abbott, B. and Abbott, D. C. and Abdinov, O. and Abed Abud, A. and Abhayasinghe, D. K. and Abidi, S. H. and AbouZeid, O. S. and et al. “Measurement of ZZ production in the $\ell\ell\nu\nu$ final state with the ATLAS detector in pp collisions at s $\sqrt{s} = 13$ TeV”. In: *Journal of High Energy Physics* 2019.10 (Oct. 2019). ISSN: 1029-8479. DOI: {10.1007/jhep10(2019)127}. URL: %7Bhttp://dx.doi.org/10.1007/JHEP10(2019)127%7D.

[2] CMS Collaboration. *Measurements of $pp \to ZZ$ production cross sections and constraints on anomalous triple gauge couplings at $\sqrt{s} = 13$ TeV*. 2020. arXiv: {2009.01186} (hep-ex).

[3] Particle Data Group et al. “Review of Particle Physics”. In: *Progress of Theoretical and Experimental Physics* 2020.8 (Aug. 2020). 083C01. ISSN: 2050-3911. DOI: 10.1093/ptep/ptaa104. eprint: https://academic.oup.com/ptep/article-pdf/2020/8/083C01/33653179/ptaa104.pdf. URL: https://doi.org/10.1093/ptep/ptaa104.

[4] *Monte Carlo to Monte Carlo scale factors for flavour tagging efficiency calibration*. Tech. rep. ATL-PHYS-PUB-2020-009. Geneva: CERN, May 2020. URL: https://cds.cern.ch/record/2718610.

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

[6] *Scikit-Learn GradientBoostingClassifier Documentation*. URL: https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.GradientBoostingClassifier.html.

[7] *TRExFitter*. URL: https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/TtHFitter.

[8] *TRExFitter: Mixed data and MC fit*. URL: https://trexfitter-docs.web.cern.ch/trexfitter-docs/AdvancedTutorial2020/Mixed/.