



# 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

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

August 2021

# Abstract

The search for  $tWZ$  production using  $139 \text{ fb}^{-1}$  of  $pp$  collision data at a centre-of-mass energy of  $\sqrt{s} = 13 \text{ TeV}$ , recorded by the ATLAS experiment at CERN, is presented. Events containing exactly four electrons or muons (tetralepton) are selected with additional criteria based on the number of jets, the number of  $b$ -tagged jets, and the number of  $Z$  boson candidates are used to define signal and control regions. The large  $t\bar{t}Z$  and  $ZZ$  backgrounds are distinguished from signal by a BDT-based algorithm. Inputs to the BDT-based algorithm include, a kinematic reconstruction algorithm which reconstructs leptonically decaying top quarks and a BDT-based algorithm which aims to classify  $\ell b$  systems originating from top quarks. The expected signal strength is extracted via a blinded maximum-likelihood fit to multiple signal and control regions. The measured signal strength is  $\mu(tWZ) = 1.91^{+0.95}_{-0.82}$ , leading to an expected significance of  $1.44\sigma$ . An expected upper limit on the signal strength is set and is given by  $\mu_{up}^{exp} = 1.61^{+2.35}_{-1.16}$ . Furthermore, a combined blinded maximum-likelihood fit was performed across the tetralepton and trilepton (studied in an independent analysis) channels, to further increase the sensitivity of  $\sigma(tWZ)$ . The measured signal strength is  $\mu(tWZ) = 1.80^{+0.70}_{-0.65}$ , leading to an expected significance of  $1.61\sigma$ . An expected upper limit on the signal strength is set and is given by  $\mu_{up}^{exp} = 1.43^{+2.04}_{-1.03}$ .

# Declaration

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

# Acknowledgements

24

# Contents

25	<b>1 Introduction</b>	7
26	<b>2 Theory</b>	8
27	2.1 Standard Model of Particle Physics . . . . .	8
28	2.1.1 The Top Quark . . . . .	9
29	2.2 $tWZ$ . . . . .	11
30	2.2.1 Tetralepton Channel . . . . .	11
31	2.2.2 Comparison to Trilepton Channel . . . . .	12
32	<b>3 The ATLAS Experiment and Detector</b>	14
33	3.1 The ATLAS Experiment . . . . .	14
34	3.1.1 Large Hadron Collider (LHC) . . . . .	14
35	3.2 The ATLAS Detector . . . . .	16
36	3.2.1 Coordinate System and Kinematics . . . . .	16
37	3.2.2 Inner Detector . . . . .	17
38	3.2.3 Electromagnetic and Hadronic Calorimeters . . . . .	17
39	3.2.4 Muon Spectrometer . . . . .	17
40	3.2.5 Trigger and Data Acquisition System . . . . .	17
41	3.2.6 Particle Identification and Object Reconstruction . . . . .	17
42	<b>4 Analysis Setup and Strategy</b>	20
43	4.1 Data and Monte Carlo Simulation . . . . .	20
44	4.1.1 Data Samples . . . . .	20
45	4.1.2 Monte Carlo Samples . . . . .	20
46	4.2 Objects . . . . .	21
47	4.2.1 Leptons . . . . .	22
48	4.2.2 Jets . . . . .	22
49	4.2.3 $b$ -tagging . . . . .	23
50	4.2.4 Overlap Removal Procedure . . . . .	23
51	4.3 Kinematic cuts . . . . .	23
52	4.4 Regions and Event Selection . . . . .	23
53	4.4.1 Optimization studies for event selection . . . . .	25
54	4.5 Signal and Control Regions . . . . .	27
55	4.5.1 $tWZ$ OF SR . . . . .	27
56	4.5.2 $tWZ$ SF SR . . . . .	28
57	4.5.3 $t\bar{t}Z$ CR . . . . .	30
58	4.5.4 $Z Z b$ CR . . . . .	30
59	4.5.5 $(tWZ)_{\text{fake}}$ CR . . . . .	32
60	4.6 Fake Lepton Estimation . . . . .	32
61	4.7 Machine Learning Techniques . . . . .	36
62	4.7.1 Boosted Decision Trees . . . . .	36
63	4.7.2 Object-level BDT . . . . .	37
64	4.7.3 Two Neutrino Scanning Method ( $2\nu\text{SM}$ ) Algorithm . . . . .	43
65	4.7.4 Event-level BDT . . . . .	47
66	4.8 Systematic Uncertainties . . . . .	50

67	4.8.1	Experimental uncertainties . . . . .	52
68	4.8.2	Theoretical uncertainties . . . . .	53
69	4.8.3	Generic shape systematics . . . . .	54
70	4.9	Analysis Pipeline and TRexFitter . . . . .	54
71	4.9.1	Fitting Procedure . . . . .	55
72	4.10	Results . . . . .	56
73	4.10.1	Tetralepton Channel . . . . .	56
74	4.10.2	Trilepton and Tetralepton Channels . . . . .	60
75	<b>5</b>	<b>Summary and Conclusions</b>	<b>64</b>

76

# Chapter 1

77

## Introduction

78 The production of a single top quark in association with a  $W^\pm$  and  $Z$  boson ( $tWZ$ ) at the CERN LHC is sensitive  
 79 to both the neutral and charged electroweak couplings of the top quark as the process involves the simultaneous  
 80 production of a  $W$  boson and a  $Z$  boson in association with the top quark. Due to the very large coupling of the  
 81 top quark to the Higgs boson, the electroweak couplings of the top quark are a theoretically well-motivated area in  
 82 which to search for the first signs of new physics [80, 52] that could offer a resolution to the Hierarchy Problem [62,  
 83 31]. The recent lack of signs of new physics from LHC data tells us that new physics is either very heavy, or is very  
 84 weakly coupled to Standard Model particles, therefore signs of new physics might only be observed in anomalous  
 85 rates of well-chosen processes. A prime example of such a process is  $tWZ$ . This has an extremely low production  
 86 cross section ( $\approx 160$  fb for  $\sqrt{s} = 13$  TeV [26]), meaning that it is an extremely rare process to observe and subse-  
 87 quently, it has never been observed by any particle physics experiment.

88

89 The latest datasets recorded by the ATLAS experiment at the CERN LHC are sufficiently large to allow a potential  
 90 observation of  $tWZ$  production. In this analysis, the Full Run 2 dataset recorded by ATLAS is used to search for  
 91  $tWZ$  production in the tetralepton channel ( $tWZ$  with exactly four final state leptons). A kinematic reconstruction  
 92 technique is used which aims to discriminate between  $tWZ$  and our most prominent background process,  $t\bar{t}Z$ . In  
 93 addition to this, Machine Learning techniques are implemented to further isolate our  $tWZ$  signal. Backgrounds  
 94 from Standard Model (SM) processes including those in which one or more leptons originate from the semileptonic  
 95 decay of a heavy hadron or a photon conversion are estimated by fitting predictions from simulation to data in  
 96 dedicated control regions. As this work forms the basis of an official ATLAS analysis, only blinded results are shown.  
 97 A maximum likelihood fit is performed over our two  $tWZ$  signal regions and three control regions, to measure the  
 98 cross section of  $tWZ$  in the tetralepton channel. In this thesis only the tetralepton channel is explored, however,  
 99 the results of a combined fit that includes an entirely independent analysis of the trilepton channel is presented  
 100 with no overlap of events between the regions defined for these channels.

101

# Chapter 2

102

## Theory

### 103 2.1 Standard Model of Particle Physics

104 The SM is a model based on Quantum Field Theory (QFT) which classifies all known elementary particles and  
 105 describes their interactions. It is a well-tested model and has shown to be hugely successful in describing exper-  
 106 imental data to great precision [8, 7]. For example, in the top quark sector, the  $t\bar{t}$  cross section predictions have  
 107 been confirmed to 3.9% accuracy [40, 71]. It incorporates three of the four fundamental forces of nature: the  
 108 electromagnetic, the weak and the strong forces. In Figure 1, all known elementary particles described by the SM,  
 109 are shown.

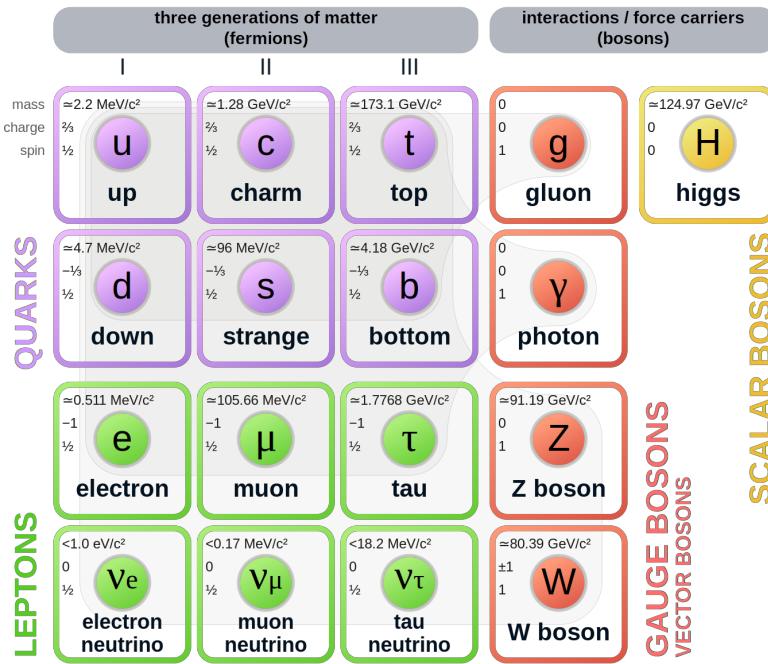


Figure 1: A summary of the elementary particles described by the SM [75] is shown. Fermions are shown on the left, with quarks shown in purple and leptons shown in green. Bosons are shown on the right, with gauge bosons shown in red and the Higgs boson shown in yellow. The mass, electric charge and spin of each particle is shown on the top left of each particle's block.

110 Particles in the SM are uniquely described by their quantum numbers: electric charge and spin. The SM particles  
 111 are split into main two classes, based off their spin quantum numbers. Particles which have half-integer spin are  
 112 called fermions, and those which have integer spin are called bosons. Fermions are further divided into three genera-  
 113 tions, each comprising of two quarks, one charged lepton and one neutrino. In a generation, the more massive quark  
 114 has an electric charge of  $+2/3$  (up-type) and the less massive quark has an electric charge of  $-1/3$  (down-type). All

115 charged leptons have an electric charge of  $-1$  and all neutrinos are electrically neutral. The masses of the particles  
 116 in a generation increase with increasing generation number, with generation 1 particles being the least massive and  
 117 generation 3 particles being the most massive. Quarks carry electric and colour charge, and can therefore interact  
 118 via the electromagnetic, weak and strong forces. Colour charge can take on three values: red, green and blue.  
 119 It is important to note that colour charge is completely unrelated to the everyday meaning of colour, and it just  
 120 represents the quantum state of the particle. Due to colour confinement [99], quarks cannot be isolated from one  
 121 another. They exist in colourless bound states, called hadrons, consisting of two or more quarks. Hadrons consisting  
 122 of an even number of quarks are known as mesons and those consisting of an odd number of quarks are known as  
 123 baryons. On the other hand, charged leptons (electron ( $e$ ), muon ( $\mu$ ) and tau ( $\tau$ )) only carry electric charge and  
 124 can therefore interact electromagnetically and weakly, but not through the strong interaction. The electric and  
 125 colour neutral fermions, neutrinos, can only interact via the weak force.

126  
 127 Particles are able to interact with one-another via the exchange of a gauge boson (boson with spin-1). Photons  
 128 are massless, spin-1 gauge bosons which mediate electromagnetic interactions between particles which carry electric  
 129 charge, such as quarks and charged leptons ( $e$ ,  $\mu$  and  $\tau$ ). The weak interaction is mediated by three massive gauge  
 130 bosons, the electrically charged  $W^+$  and  $W^-$  bosons and the electrically neutral  $Z$  boson. Gluons are massless,  
 131 spin-1 gauge bosons which mediate strong interactions between particles which carry colour charge, such as quarks.  
 132 Since gluons carry colour charge, they interact with themselves. The massive, spin-0, electrically neutral Higgs  
 133 boson mediates the Higgs field which gives mass to the  $W^\pm$  and  $Z$  bosons via the so-called Brout-Englert-Higgs  
 134 mechanism [46, 58, 57]. The Brout-Englert-Higgs mechanism induces spontaneous electroweak symmetry breaking  
 135 to provide mass terms for the  $W^\pm$  and  $Z$  bosons in the electroweak Lagrangian of the SM. All particles described in  
 136 the SM have their own antiparticle, with the same mass, but opposite charges. Some particles, such as the photon,  
 137 are their own antiparticle.

138  
 139 Although the SM has shown to be hugely successful, it is incomplete and fails to describe certain observed phenomena.  
 140 The most notable example being the absence of gravity from the SM. The gravitational force is  $\approx 10^{29}$  [91]  
 141 weaker than the weak force, therefore quantum gravitational effects are expected to only become significant at  
 142 energies much larger than that currently accessible by the LHC (known as the Planck scale  $\approx 10^9$  GeV) [62]. This  
 143 large difference in strength between the weak force and gravity is known as the Hierarchy Problem. Cosmological  
 144 observations infer that around 84% of the matter in the universe consists of gravitationally interacting matter known  
 145 as dark matter [61]. None of the particles described in the SM are good dark matter candidates, therefore the SM  
 146 only accounts for a small fraction of the total matter of the universe. The large discrepancy between the observed  
 147 amount of matter and antimatter in the universe, sometimes referred to as the matter-antimatter asymmetry, is not  
 148 fully explained by the SM. Neutrinos in the SM are assumed to be massless, however observations of neutrino oscillations  
 149 (neutrinos undergoing flavour change as they travel through space) imply that neutrinos do have mass [49].  
 150 Beyond the Standard Model (BSM) theories attempt to explain the phenomena which the SM cannot. For example,  
 151 a popular extension to the SM, Supersymmetry (SUSY) introduces new particles to the SM which are counterparts  
 152 to the existing SM particles with the same quantum numbers, except for their spins [64]. SUSY provides elegant  
 153 explanations to many shortcomings of the SM, however none of the supersymmetric particles described by SUSY  
 154 have been observed experimentally [34].

### 155 2.1.1 The Top Quark

156 The top quark is the heaviest particle in the SM, with a mass of  $172.76 \pm 0.30$  GeV [55]. According to the SM, since  
 157 the coupling to the Higgs boson is proportional to the the mass of the interacting particle, the top quark is strongly  
 158 coupled to the Higgs boson. Physics processes involving top quarks is therefore a theoretically well-motivated area  
 159 to search for new physics, since it is the most likely particle to couple to new physics theories at the TeV scale. Its  
 160 large mass also makes it highly unstable, with a mean lifetime of  $\approx 0.5 \times 10^{-24}$  s [55]. The top quark's lifetime  
 161 is shorter than that of the hadronisation process, and it therefore decays before hadronising. The top quark can  
 162 therefore be measured indirectly via its decay products. Top quarks almost always decay to a  $W$  boson and a  
 163  $b$ -quark ( $\frac{\Gamma(Wb)}{\Gamma(Wq(q=b,s,d))} = 0.957 \pm 0.034$  [55]). The  $b$ -quark is the second heaviest quark in the SM, however its  
 164 lifetime is still longer than the hadronisation time scale [55]. In hadron collider experiments,  $b$ -quarks travel a short  
 165 distance in the detector before hadronising to form jets. In Table 1, the dominant final state branching fractions of  
 166 the top quark are shown.

167 Hadronic final states are more than twice as likely than leptonic final states. Final state decays to different lepton  
 168 flavours are roughly equally probable.

Decay Mode	Branching Fraction ( $\frac{\Gamma_i}{\Gamma}$ )
$t \rightarrow Wb \rightarrow e\nu_e b$	$(11.10 \pm 0.30)\%$
$t \rightarrow Wb \rightarrow \mu\nu_\mu b$	$(11.40 \pm 0.20)\%$
$t \rightarrow Wb \rightarrow \tau\nu_\tau b$	$(10.70 \pm 0.50)\%$
$t \rightarrow Wb \rightarrow q\bar{q}b$	$(66.50 \pm 1.40)\%$

Table 1: The dominant final state branching fractions of the top quark [55] are shown.

169  
170 Top quark production can be placed into two main categories: pair production ( $t\bar{t}$ ) and single-top production ( $t$ ) [87].  
171 In the LHC, top quarks are mainly produced in pairs via strong interactions in gluon-gluon fusion ( $gg \rightarrow t\bar{t}$ ) or quark  
172 annihilation ( $q\bar{q} \rightarrow t\bar{t}$ ). Top quark production via gluon-gluon fusion is the dominating process [25]. The production  
173 cross section for  $t\bar{t}$  (leptonic final state) in  $pp$  collisions with  $\sqrt{s} = 13$  TeV was measured by ATLAS with a value  
174 of  $830 \pm 0.4(\text{stat}) \pm 36(\text{syst}) \pm 14(\text{lumi})$  pb [4], with good agreement between measurement and theoretical prediction.  
175  
176 Single top production occurs via the weak interaction. The most abundant production mechanisms leading to  
177 single top production are the  $s$ -,  $t$ - and  $Wt$ - channels [55]. In the  $s$ -channel, an initial quark annihilates with an  
178 anti-quark of different flavour, producing a virtual  $W$  boson which decays to a top quark and anti-bottom quark.  
179 In the  $t$ -channel, an initial  $b$  quark interacts with a different flavour quark via the exchange of a  $W$  boson. This  
180 interaction produces a top quark and another quark. In the  $Wt$ -channel, an initial gluon interacts with a  $b$  quark  
181 to produce a top quark and a  $W$  boson, either via the absorption of the gluon by the  $b$  quark or via the exchange  
182 of a top quark. In Table 2, single top production cross sections in  $pp$  collisions at  $\sqrt{s} = 13$  TeV for the  $s$ -,  $t$ - and  
183  $Wt$ -channels, are shown.

Channel	Process	Total Cross Section [pb]
$s$	$q\bar{q}' \rightarrow W \rightarrow \bar{b}t$	$10.32^{+0.40}_{-0.36}$
$t$	$bq' \rightarrow W \rightarrow tq$	$216.99^{+9.04}_{-7.71}$
$Wt$	$bg \rightarrow b/t \rightarrow Wt$	$71.7 \pm 3.85$

Table 2: Single top production cross sections in  $pp$  collisions at  $\sqrt{s} = 13$  TeV for the  $s$ -,  $t$ - and  $Wt$ -channels [35] are shown. The prime superscript on  $q'$  indicates that the quark has a different flavour to  $q$ .

184 Single top production is suppressed compared to pair produced top production, with  $t\bar{t}$  production (leptonic final  
185 state) being around three times as likely to occur than single top production across all decay channels.

### 186 2.1.1.1 Motivation for the search for $tWZ$ production in the tetralepton channel

187 The recent lack of signs of new physics from LHC data [90] tells us that new physics is either very heavy, or is  
188 very weakly coupled to SM particles. We therefore might only observe signs of new physics in anomalous rates of  
189 well-chosen processes. The  $tWZ$  process is a prime example of such a process. It has an extremely low production  
190 cross section (0.7 fb for  $\sqrt{s} = 13$  TeV [26]), and has subsequently never been observed by any particle physics  
191 experiment. Since  $tWZ$  involves a charged  $W$  boson and neutral  $Z$  boson, its cross section is sensitive to the  
192 charged and neutral couplings to the top quark. In turn, the top quark is strongly coupled to the Higgs boson, due  
193 to its large mass. Due to the top quark's large coupling to the Higgs boson, corrections to the Higgs boson mass  
194 diverge in the SM. The top quark's couplings are modified, in order to remove this divergence, in many scenarios of  
195 new physics that aim to resolve the Hierarchy Problem. Since the  $Z$  boson may be radiated from the initial-state  
196  $b$ -quark, the final-state top quark, or the final-state  $Z$  boson, the  $tWZ$  process embeds the  $b - Z$ ,  $t - Z$  and  $W - Z$   
197 electroweak couplings which are often modified in BSM physics. Therefore  $tWZ$  is an important process in the  
198 search for signs of new physics and BSM physics.

199 One such BSM theory which is sensitive to  $tWZ$  production [66, 73] is the Standard Model Effective Field  
200 Theory (SMEFT) [28]. The SMEFT attempts to describe physics at large energy scales which we have not yet been

able to probe experimentally. The SMEFT inherits the same QFT framework as the SM, and adds Lagrangian terms to the SM Lagrangian which describe the interactions of SM particles at higher energy scales. Analogous to the coupling constants found in the SM Lagrangian, which indicate the interaction strengths between different particles, SMEFT contains scalar coefficients which operate in the same way. These scalar coefficients are known as Wilson coefficients. It has been shown that the cross section of  $tWZ$  is sensitive to many Wilson coefficients. An experimental constraint on the cross section of  $tWZ$  is therefore expected to be impactful on a global fit on all the Wilson coefficients in SMEFT.

Prior to this analysis, only three experimental studies of  $tWZ$  in ATLAS have been performed. The first and third studies utilised the trilepton channel to search for  $tWZ$  production, whereas the second study utilised both the tri- and tetralepton channels. The first search utilised  $36 \text{ fb}^{-1}$  of ATLAS data and an upper limit on the cross section of  $tWZ$  was set at a value of  $\approx 6$  times the SM cross section [79]. The second study investigated the feasibility of a cross section measurement of  $tWZ$  production with CMS Run 3 data ( $300 \text{ fb}^{-1}$ ) [97]. The study showed that it is possible to exclude  $\mu(tWZ)$  at the  $7\sigma$  significance level using  $300 \text{ fb}^{-1}$  of data. This study needs to be further investigated, since its findings seem improbable given the results obtained in this thesis. The third search utilised  $139 \text{ fb}^{-1}$  of ATLAS data and an expected upper limit on the cross section of  $tWZ$  was set at a value of  $\approx 2.6$  times the SM cross section [98]. In Section 4.10.2, the latter analysis will be used in combination with this analysis, in order to further increase the sensitivity of the cross section of  $tWZ$ .

## 2.2 $tWZ$

### 2.2.1 Tetralepton Channel

In Figure 2, the Leading Order (LO) Feynman diagram for  $tWZ$  in the tetralepton channel, is shown.

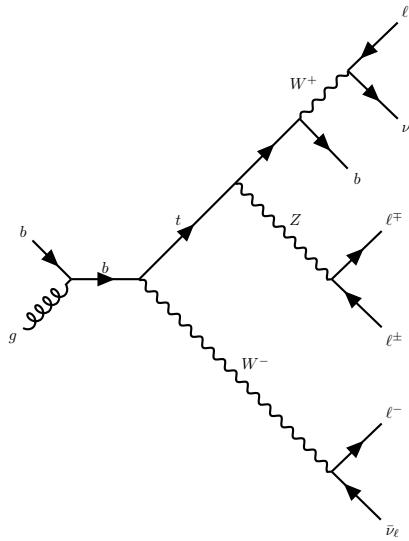


Figure 2: The LO Feynman diagram of  $tWZ$  production in the tetralepton channel is shown.

#### 2.2.1.1 Backgrounds

The main backgrounds for  $tWZ$  (tetralepton channel) are the production of a two tops, both in the  $\ell\nu b^1$  final state channel, together with a  $Z$  boson ( $t\bar{t}Z$ ) and diboson production with fully leptonic final states ( $ZZ$ ). In Figure 2.2.1.1, LO Feynman diagrams for  $t\bar{t}Z$  and  $ZZ$  in the tetralepton channel, are shown.

<sup>1</sup>In this thesis,  $\ell$  refers to an electron or muon,  $\nu$  refers to a neutrino or anti-neutrino and  $b$  refers to a bottom quark or anti-bottom quark

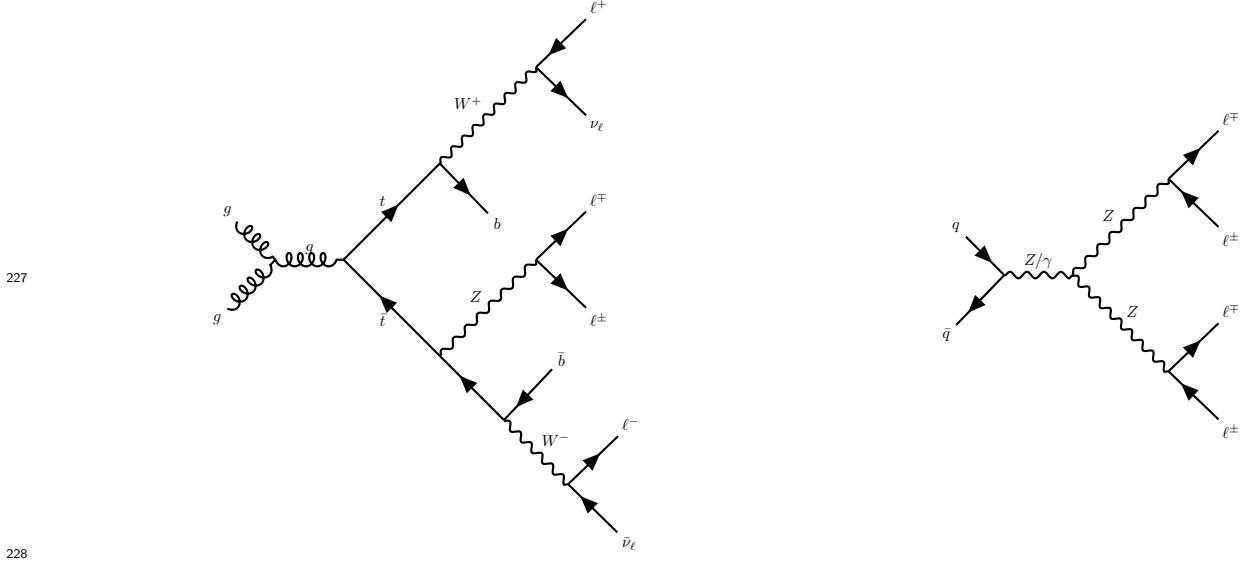


Figure 3: LO Feynman diagrams for  $t\bar{t}Z$  (left) and  $ZZ$  (right) in the tetralepton channel are shown.

229 The  $t\bar{t}Z$  process contains four leptons and two  $b$ -quarks in its final state (inclusive  $\sigma(t\bar{t}Z) = 0.95 \pm 0.08_{\text{stat}} \pm 0.10_{\text{syst}}$   
230 pb at  $\sqrt{s} = 13$  TeV [2]) and can easily mimic the  $tWZ$  signal process, for instance, by one of its  $b$ -jets getting  
231 missed during detection. The  $ZZ$  process contains four leptons and zero  $b$ -quarks in its final state (inclusive  
232  $\sigma(ZZ) = 14.6^{+1.9}_{-1.8}(\text{stat})^{+0.5}_{-0.3}(\text{syst}) \pm 0.2(\text{theo}) \pm 0.4(\text{lumi})$  pb at  $\sqrt{s} = 13$  TeV [63]). One way in which  $ZZ$  can mimic  
233 the  $tWZ$  signal process is by reconstruction of a non-prompt  $b$ -jet.

### 234 2.2.2 Comparison to Trilepton Channel

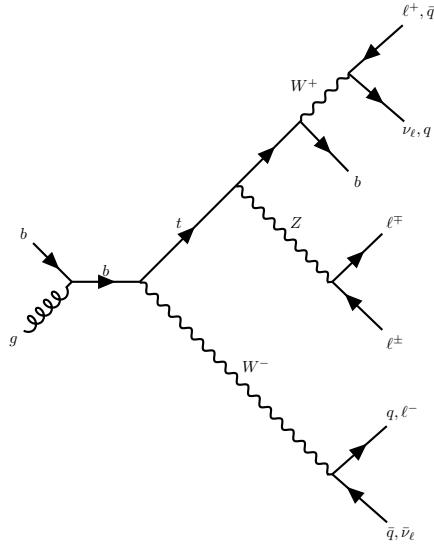


Figure 4: Example Feynman diagram of  $tWZ$  production in the tri-lepton channel.

235 The most apparent difference between the tri and tetralepton channels is the number of events present, with the  
236 tetralepton channel having far less events in its phase space than that of the tri-lepton channel. The lack of statistics  
237 in the tetralepton channel can be attributed to its low production cross section,  $\sigma_{(tW^\pm Z).Br(4\ell)}^{\text{NLO}} = 0.7$  fb [26].  
238 The tri-lepton channel has a production cross section ( $\sigma_{(tW^\pm Z).Br(3\ell)}^{\text{NLO}} = 3.9$  fb [26]) around a factor of 4 larger  
239 than that of the tetralepton channel. This difference between the production cross section of the two decay  
240 channels can be largely attributed to the difference in branching ratios ( $\frac{\Gamma_i}{\Gamma}$ ) between a hadronically decaying  $W$

<sup>241</sup> boson,  $\frac{\Gamma_{W \rightarrow had}}{\Gamma_W} = (67.41 \pm 0.27)\%$  [55], present in the tri-lepton channel and a leptonically decaying  $W$  boson,  
<sup>242</sup>  $\frac{\Gamma_{W \rightarrow \ell\nu}}{\Gamma_W} = (10.86 \pm 0.09)\%$  [55], present in the tetralepton channel.

<sup>243</sup>

<sup>244</sup> Despite the tetralepton channel's low statistics, it is not subject to the large  $WZ$  background present in the trilepton  
<sup>245</sup> channel [98]. The tetralepton channel has a substantial amount of  $ZZ$  background (not present in the trilepton  
<sup>246</sup> channel), fortunately this can be easily suppressed due to the full reconstructability of the two leptonically decaying  
<sup>247</sup>  $Z$ -bosons.

248

# Chapter 3

## The ATLAS Experiment and Detector

249

### 3.1 The ATLAS Experiment

251 The ATLAS (A Toroidal LHC ApparatuS) detector is one of two general purpose detectors at CERN (the European  
 252 Organization for Nuclear Research) near Geneva in Switzerland. These detectors collect data from the collisions  
 253 provided by the worlds highest energy particle accelerator [30], the Large Hadron Collider (LHC) situated at CERN.

254 In this section, information about the LHC and the ATLAS detector are given. This includes technical aspects  
 255 of the ATLAS detector and the processing of data into meaningful physics objects<sup>1</sup> to be used in analyses. The  
 256 following chapter consists of information from "The LHC Design Report" [30], "LHC Machine" [48] and "The  
 257 ATLAS Experiment at the CERN Large Hadron Collider" [41] unless otherwise stated.

#### 3.1.1 Large Hadron Collider (LHC)

258 The LHC is a circular 27 km particle accelerator located in an underground tunnel on the border between France  
 259 and Switzerland. The accelerator consists of supercooled, superconducting magnets which accelerate and collide  
 260 beams of protons at centre-of-mass energies up to  $\sqrt{s} = 13$  TeV at instantaneous luminosities of  $\mathcal{L} \sim 10^{34}$  cm $^{-2}$ s $^{-1}$ .  
 261 In the LHC,  $pp$  beams consist of bunches of protons which collide every 25 ns, corresponding to a frequency of 40  
 262 MHz. Several accelerator systems are used to accelerate protons and heavy ions to such high energies. Protons are  
 263 extracted from a tank of ionised hydrogen gas and are injected into the Linear Accelerator 2 (LINAC), where they  
 264 are linearly accelerated to momenta of 50 MeV. The proton bunches are then sequentially accelerated by a chain  
 265 of circular accelerators. The chain starts with the Booster which accelerates the protons to momenta of up to 1.4  
 266 GeV. The proton bunches are then fed through to the Proton Synchrotron (PS) and the Super Proton Synchrotron  
 267 (SPS) which accelerate the protons to momenta of up to 25 GeV and 450 GeV respectively. The protons are  
 268 then transferred to two beam pipes of the LHC where they travel in opposite directions. Both proton beams are  
 269 accelerated to their final momenta of 6.5 TeV, resulting in a centre-of-mass energy of 13 TeV. These proton beams  
 270 then collide at one of the four main interaction points (positions along the beam pipe where collisions occur)  
 271 situated along the LHC.

272 The four main experiments located at the interaction points are ATLAS, the Compact Muon Solenoid (CMS),  
 273 Large Hadron Collider Beauty (LHCb) Experiment and A Large Ion Collider Experiment (ALICE). The ATLAS  
 274 and CMS detectors are general-purpose detectors which investigate a wide range of physics processes. Since both  
 275 ATLAS and CMS can measure the same processes, they are able to cross-check and validate measurements taken  
 276 by one another. The LHCb detector is specifically designed to study decays of particles containing  $b$ -quarks. The  
 277 ALICE detector is designed to study the strongly interacting quark-gluon plasma which is formed at extremely  
 278 high energy densities. At the interaction points, the two proton beams which consist of protons in closely packed  
 279 bunches, travel in opposite directions to one another and collide. Many hard  $pp$  collisions (events) can occur per  
 280 bunch crossing, however it is the most energetic collision in the bunch crossing that is interesting for discovery  
 281 potential. The most energetic collision is therefore chosen to be studied and any additional collisions are aimed  
 282 to be rejected. These additional collisions are referred to as *pile-up*. Pileup complicates the reconstruction of the  
 283 particles originating from the hard collision of interest.

<sup>1</sup>meaningful physical systems which can be reconstructed from detector information (e.g leptons and jets)

287 **3.1.1.1 Luminosity**

288 This section consists of information from "Modern Particle Physics" [91], unless otherwise stated.

289

290 The event production rate at the LHC,  $R(t)$ , for a certain process of interest is given by,

$$R(t) = \mathcal{L}(t)\sigma \quad (3.1)$$

291 where  $\mathcal{L}(t)$  is the instantaneous luminosity and  $\sigma$  and is the cross section of the process of interest. The instantaneous  
292 luminosity,  $\mathcal{L}(t)$ , is independent on the process of interest, and depends on various collider and beam parameters.  
293  $\mathcal{L}(t)$  can be written in terms of these parameters as,

$$\mathcal{L}(t) = f \frac{N n_1 n_2}{4\pi \sigma_x \sigma_y} \quad (3.2)$$

294 where  $f$  is the beam revolution frequency,  $N$  is the number of proton bunches colliding per second,  $n_1$  and  $n_2$  are the  
295 number of protons in the colliding bunches,  $\sigma_x$  and  $\sigma_y$  are the beam spread in the  $x$  and  $y$  directions respectively.  
296 The average number of visible (particles that are measurable by the detector)  $pp$  interactions per bunch crossing,  
297  $\langle \mu \rangle$ , can be written as,

$$\langle \mu \rangle = \frac{\sigma_{inel} \mathcal{L}}{N f} \quad (3.3)$$

298 where  $\sigma_{inel}$  is the inelastic  $pp$  interaction cross section. The total integrated luminosity,  $L$ , across some time interval,  
299 is given by,

$$L = \int \mathcal{L} dt. \quad (3.4)$$

300 The units of  $L$  are inverse area, and are given by  $\text{fb}^{-1}$  at the LHC and the ATLAS detector.

301

302 In Figure 5, the total integrated luminosity delivered to ATLAS, recorded by ATLAS, and certified to be good  
303 enough for physics analyses (the data passes certain quality control criteria) for  $\sqrt{s} = 13$  TeV  $pp$  collisions at the  
304 LHC is shown [93].

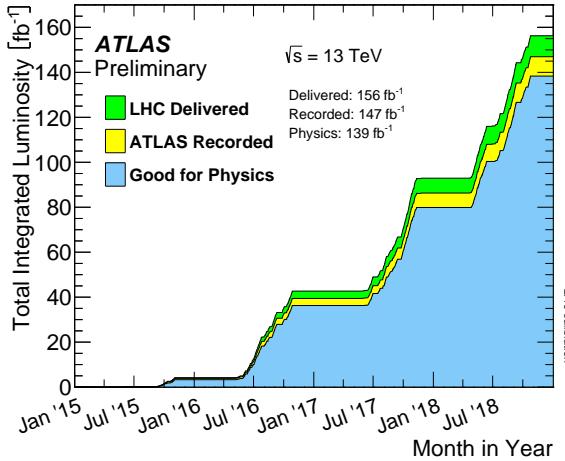
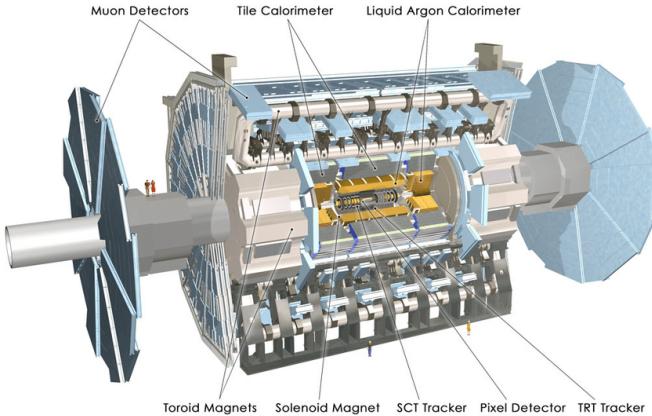


Figure 5: The total integrated luminosity delivered to ATLAS, recorded by ATLAS, and certified to be good enough for physics analyses (the data passes certain quality control criteria) for  $\sqrt{s} = 13$  TeV  $pp$  collisions at the LHC is shown [93]. The total integrated luminosity delivered by the LHC, recorded by ATLAS and certified to be good quality data are shown by the green, yellow and blue histograms respectively. The month and year of data taking is shown on the x-axis and the total integrated luminosity (in  $\text{fb}^{-1}$ ) is shown on the y-axis.

305 A total integrated luminosity of  $139 \text{ fb}^{-1}$  of data certified as good for physics was recorded by ATLAS between  
306 2015 and 2018. This data taking period is referred to as Run 2, since it proceeds the Run 1 data taking period  
307 (2011 and 2012) and the Long Shutdown 1 LHC upgrade period (2013 and 2014). In this analysis, we use the Full  
308 Run 2 dataset.

## 309 3.2 The ATLAS Detector

310 In Figure 6, the schematic of the ATLAS detector, is shown.



311 Figure 6: Schematic of the ATLAS detector [41]

312 The detector is cylindrically shaped which covers close to  $4\pi$  in solid angle. It has a length of 44 m, a diameter  
 313 of 25 m and a mass of 7000 tons. The ATLAS detector consists of four main sub-detectors arranged in concentric  
 314 cylindrical layers around the beam pipe. These include the inner detector, the electromagnetic calorimeter,  
 315 the hadronic calorimeters and the muon spectrometer. The sub-detectors record the momenta, energies and trajectories  
 316 of different particles produced in the collider, allowing for the reconstruction and identification of these particles to  
 be used in physics analyses.

### 317 3.2.1 Coordinate System and Kinematics

318 The ATLAS detector adopts a right-handed coordinate system. The origin is at the nominal interaction point  
 319 with the  $z$ -axis defined to be counter-clockwise along the beam line. The  $x - y$  plane (or transverse plane) is  
 320 perpendicular to the beam line, with the  $x$ -axis pointing towards the centre of the LHC ring and the  $y$ -axis  
 321 pointing upwards towards the Earth's surface. The azimuthal angle,  $\phi \in [-\pi, \pi]$ , is measured in the transverse  
 322 plane with respect to the positive  $x$ -axis. The polar angle,  $\theta \in [0, \pi]$ , is measured in the  $z - y$  plane with respect  
 323 to the positive  $y$ -axis. A quantity called the pseudorapidity,  $\eta \in [0, \infty]$  is defined as,

$$\eta = -\ln \tan\left(\frac{\theta}{2}\right) \quad (3.5)$$

324  $\eta$  is often used as a measure of the polar angle, instead of  $\theta$ , since the difference in  $\eta$  between two particles,  $\Delta\eta$ , is  
 325 invariant under a Lorentz boost in the  $z$ -direction [94]. The angular distance between two physics objects,  $\Delta R$ ,  
 326 can be written as,

$$\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} \quad (3.6)$$

327 where  $\Delta\phi$  is the difference in  $\phi$  between the two physics objects of interest. Quantities defined in the transverse  
 328 plane are often used to describe the kinematics of physics objects in hadron collider experiments. The transverse  
 329 momentum,  $p_T$ , is defined as,

$$p_T = \sqrt{(p_x)^2 + (p_y)^2} \quad (3.7)$$

330 where  $p_x$  and  $p_y$  are the  $x$  and  $y$  components of the physics object's momenta, respectively. The transverse energy,  
 331  $E_T$ , is defined as,

$$E_T = \sqrt{m^2 + p_T^2} \quad (3.8)$$

332 where  $m$  is the invariant mass of the physics object.

### 333 3.2.2 Inner Detector

334 The inner detector is the first layer of concentric cylindrical sub-detector layers in the ATLAS detector. It is used  
 335 to identify charged particles and reconstruct the trajectories of charged particles produced in the collisions via  
 336 energy deposition in semiconductor material (hits) and the ionisation of gas. It consists of three complementary  
 337 sub-detectors (in order from nearest to farthest from the beam pipe): the Pixel Detector, the Semiconductor Tracker  
 338 (SCT) and the Transition Radiation Detector (TRT). The Pixel Detector and SCT are based on semiconductor  
 339 technology and have the highest granularity of any sub-detector in ATLAS, in order to cope with the high frequency  
 340 of collisions near the interaction point. The TRT consists of drift tubes (straws) containing a gas mixture, which  
 341 allows measurement of the energy deposited by charged particles through the ionisation of the gas. Solenoid magnets  
 342 surround the inner detector and bend the trajectories of charged particles. The charges and transverse momenta  
 343 of particles can be inferred from their bent trajectories, which are reconstructed by the hits produced via energy  
 344 deposition in the Inner Detector.

### 345 3.2.3 Electromagnetic and Hadronic Calorimeters

346 The Electromagnetic Calorimeter (ECAL) and Hadronic Calorimeter (HCAL) surround the Inner Detector, with  
 347 the ECAL nearer to the beam line. The ECAL and HCAL provide accurate measurements of the energy of  
 348 particles which interact electromagnetically (e.g. photons and electrons) and hadronically (e.g. jets), respectively.  
 349 Particles entering the calorimeters interact with the detector material and create either a electromagnetic shower  
 350 (in the ECAL) or a hadronic shower (in the HCAL), depositing all their energy in the calorimeter cells. The  
 351 primary mechanism of energy deposition in the ECAL is through bremsstrahlung (for electrons) and pair production  
 352 (photons). Hadrons usually deposit a small amount of their energy in the ECAL, and interact via inelastic scattering  
 353 with the nuclei of the detector material. The hadronic showers (jets) produced in these nuclear interactions travel  
 354 much further than an electromagnetic shower, and for that reason, the volume of the HCAL is designed to occupy  
 355 a much larger space than that of the ECAL.

### 356 3.2.4 Muon Spectrometer

357 The Muon Spectrometer (MS) is the outermost sub-detector of ATLAS and surrounds the HCAL. Muons traverse  
 358 through the inner detector and calorimeters, with minimal energy loss, before reaching the MS. The MS consists  
 359 of trigger and high-precision tracking systems. Large superconducting toroid shaped magnets deflect the incoming  
 360 muons to measure their trajectories and subsequently their momenta via the curvature of the trajectories. The MS  
 361 measures muon trajectories as they ionize gas (filled with Ar and CO<sub>2</sub> gas) in the MS drift chambers.

### 362 3.2.5 Trigger and Data Acquisition System

363 The Trigger and Data Acquisition System (TDAQ) manages and handles the large amount of data produced within  
 364 the ATLAS detector. In Run 2,  $pp$  bunch crossings occur every 25 ns, corresponding to an event rate of 40 MHz.  
 365 The TDAQ system performs a fast preliminary reconstruction to select events with signatures which are interesting  
 366 for physics analyses. The information collected from these events are permanently stored for offline reconstruction  
 367 and analysis, and the rest (the vast majority of events) are discarded. The trigger system reduces the 40 MHz data  
 368 rate to around 1 kHz.

### 369 3.2.6 Particle Identification and Object Reconstruction

370 Particles originating from  $pp$  collisions, or from their subsequent decays, traverse through the ATLAS detector and  
 371 interact with its different sub-detectors, producing characteristic electronic signals. These signals are then processed  
 372 by various algorithms to reconstruct and identify the physics objects (e.g. electrons, muons, jets) in the event. This  
 373 section outlines the procedures used to define these physics objects.

#### 374 3.2.6.1 Tracks and primary vertices

375 The trajectories of charged particles, or tracks, are reconstructed in the ID. First, energy is deposited by charged  
 376 particles (hits) in pixels or strips, in the Pixel and SCT detectors respectively. Adjacent pixels or strips are grouped  
 377 together in *energy clusters*. Energy clusters define 3D space-points indicating the location where the charged  
 378 particle traversed. Track seeds are then defined as sets of three space-points, in either the Pixel or SCT detectors.  
 379 A Kalman filter [11] is then used to build track candidates from the track seeds. Often, multiple track candidates

are built per track seed, therefore an ambiguity solver [38] is needed for finding the track which best represents the traversal of the charged particle. The ambiguity solver ranks each track from a given seed based on, the number of associated hits, the number of holes (expected hits which are absent), track momenta and the  $\chi^2$  of the track fit. Low ranked tracks are then discarded. High ranked tracks are refitted, introducing information from the TRT.

The primary vertex is the location of the  $pp$  collision of interest (i.e. from the hard scatter). The primary vertex from the hard scatter needs to be identified, to isolate the event of interest from unwanted pile-up events. In the event reconstruction procedure [72], the primary vertex is defined as the vertex of the event with the largest sum of  $(p_T)^2$  (corresponding to the measured  $(p_T)^2$  of the particle from its reconstructed track) of its associated tracks. Furthermore, the primary vertex is required to have at least two associated tracks. To reduce contamination from fake tracks used in primary vertex reconstruction, only tracks which pass certain tight selection criteria are used in the reconstruction procedure. An iterative fitting procedure is then used to reconstruct the primary vertex by finding a set of reconstructed tracks which have a common vertex.

### 3.2.6.2 Electrons

Since electrons are charged particles, they give rise to tracks in the Inner Detector. They also deposit energy in the ECAL via electromagnetic showering. Electrons are therefore reconstructed and identified from signals in the Inner Detector and ECAL. Electrons are reconstructed using a dynamic clustering algorithm [1] which matches electron candidate tracks in the Inner Detector to energy clusters in the ECAL. The dynamic clustering algorithm matches tracks to energy clusters which have local maxima, to form electron candidates. A likelihood discriminant is used to identify electrons. Quantities measured in the Inner Detector and ECAL are used as input, such that they discriminate well between prompt isolated electrons and other physics objects (e.g. jets, electron from a photon conversion, electron from a semi-leptonically decaying hadron). Important input variables include the shape of the electromagnetic shower, track quality in the Inner Detector and information from the TRT.

### 3.2.6.3 Muons

Muons leave tracks in the Inner Detector and the MS. They traverse the ECAL and HCAL with no significant energy loss. Muons are therefore reconstructed and identified from information in the Inner Detector and MS. Tracks are reconstructed [5] in the Inner Detector and MS independently. Both tracks are combined, using a global  $\chi^2$  fit, resulting in reconstructed muon candidates.

Similar to electron identification, muons use a likelihood discriminant to identify prompt muons and suppress background contamination (mainly from pion and kaon decays).

### 3.2.6.4 Jets and $b$ -tagging

Coloured particles emerging from the interaction point result in collimated streams of colourless particles, known as jets. Jets can deposit energy in the Inner Detector and in the HCAL. Jets in ATLAS are reconstructed from topological clusters using the anti- $k_t$  algorithm [32]. Topological clusters are groups of adjacent calorimeter cells which contain energy deposition above the average amount of noise expected in the cell. Adjacent cells are grouped together under certain criteria to form topological clusters which form jets.

Different tagging algorithms are used to identify the quark flavour which initiated a jet.  $b$ -quark tagging is used extensively in top physics, due to the  $b$ -quark present in the top quark's dominant decay channel (See Table 1). Hadrons arising from  $b$ -quark hadronisation have mean lifetimes  $\sim 1.5$  ps and travel (on average) a few millimetres before decaying. This creates a secondary vertex within the jet (See Figure 7). This characteristic decay signature, along with several other unique features of  $b$ -jets, are exploited in  $b$ -tagging algorithms to distinguish  $b$ -jets from  $c$ - or light flavour jets. In Figure 7, an illustration of the production of a  $b$ -jet, is shown.

In this analysis, we use the recommended DL1r (Deep-Learning Flavour Tagger) tagging algorithm [19]. 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.

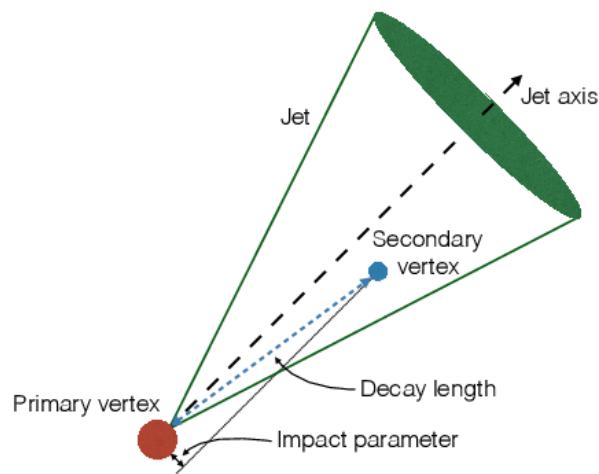


Figure 7: An illustration [42] of the production of a  $b$ -jet is shown. This illustrates the presence of a secondary vertex within a  $b$ -jet.

428

# Chapter 4

429

# Analysis Setup and Strategy

430 In this chapter, the setup of the analysis and the analysis strategy is presented for  $tWZ$  production in the tetralepton  
 431 channel based on an integrated luminosity of  $139 \text{ fb}^{-1}$  of data recorded by ATLAS. This includes a description of the  
 432 data and simulated Monte Carlo (MC) samples, definitions of physics objects, event selection and the subsequent  
 433 definition of signal and control regions used in this analysis. Furthermore, estimation of the fake lepton component  
 434 using the MC template method is described. Various Machine Learning techniques and a kinematic reconstruction  
 435 algorithm, used to discriminate between signal and background events, are described. An outline of the systematic  
 436 uncertainties affecting the measurement are presented. Finally, the results of the analysis are presented.

## 4.1 Data and Monte Carlo Simulation

### 4.1.1 Data Samples

437 The  $pp$  collision data used in this analysis was collected with the ATLAS detector at the LHC from 2015 to 2018. This  
 438 data period of data taking is referred to as Run 2. During this period,  $pp$  collisions at  $\sqrt{s} = 13 \text{ TeV}$ , corresponding  
 439 to an integrated luminosity ( $\mathcal{L}$ ) of  $156 \text{ fb}^{-1}$ , were delivered by the LHC. The ATLAS detector managed to record  
 440  $147 \text{ fb}^{-1}$  of this total delivered data.  $139 \text{ fb}^{-1}$  of the data recorded by ATLAS is considered to be good enough  
 441 for physics analyses (the data passes certain quality control criteria) and placed into the *Good Runs List* [54].  
 442 Specific time-intervals where the luminosity is assumed to be constant are known as *lumi-blocks*. Only events from  
 443 lumi-blocks in which LHC beams were stable and all ATLAS detectors were operational are selected. The list of  
 444 suitable run and lumi-blocks is summarised in the official Good Runs Lists for 2015-2018 data, as specified in  
 445 Ref. [53]. The integrated luminosities corresponding to the individual datasets for years 2015, 2016, 2017 and 2018  
 446 are  $3.2 \text{ fb}^{-1}$ ,  $33.0 \text{ fb}^{-1}$ ,  $44.3 \text{ fb}^{-1}$  and  $58.5 \text{ fb}^{-1}$  respectively.  
 447

### 4.1.2 Monte Carlo Samples

448 Simulated MC samples were generated and used to model the SM  $tWZ$  signal and its backgrounds.  
 449

450 The following background processes are considered:

- 451 •  **$t\bar{t}Z$ :**  $t\bar{t}$  with an associated  $Z$ -boson, in the tetralepton final state. Therefore, both top quarks decay leptonically  
   (e.g.  $t \rightarrow W^+ b \rightarrow \ell^+ \nu b$ ) and of these top quarks emits a  $Z$ -boson which decays leptonically ( $Z \rightarrow \ell^\pm \ell^\mp$  (OSSF  
   lepton pair)). This results in a final state with 4 leptons and 2 b-quarks.
- 452 •  **$ZZ$ :** Diboson production with a tetralepton final state, therefore both  $Z$ -bosons decay leptonically ( $Z \rightarrow \ell^\pm \ell^\mp$   
   (OSSF lepton pair)).
- 453 • **other:** Processes with a relatively minimal, but non-negligible background contribution
  - 454 -  $VVV$  ( $V = W/Z$ )
  - 455 -  $t\bar{t}$
  - 456 -  $t\bar{t}W$
  - 457 -  $t\bar{t}WW$

- 463      -  $t\bar{t}H$   
 464      -  $WZ$   
 465      -  $t\bar{t}t$   
 466      -  $t\bar{t}\bar{t}\bar{t}$   
 467      -  $tZq$

468 The MC simulations are achieved via the use of event generators and parton shower generators. Event generators  
 469 simulate the  $pp$  collisions (hard events) by sampling the proton's Parton Distribution Functions (PDFs) at the  
 470 desired energy scale. The parton shower generators simulate any incoming or outgoing particles from the hard  
 471 process, which carry QCD color charge and can therefore lead to parton showers.

472 The production of  $tWZ$  events is simulated with the **MADGRAPH5\_AMC@NLO 2.3.3** [9] generator providing matrix  
 473 element (ME) calculations at NLO (next-to-leading order) in QCD. The events are interfaced with **PYTHIA 8.235**  
 474 for the parton shower. The production of  $t\bar{t}Z$  and  $t\bar{t}W$  events are simulated with the **MADGRAPH5\_AMC@NLO 2.3.3**  
 475 generator providing ME calculations at NLO in QCD. The events are interfaced with **PYTHIA 8.210** [89] for the  
 476 parton shower. Event generation of  $tWZ$  and  $t\bar{t}Z$  results in diagrams which overlap with one another, that is,  
 477 these diagrams contain the same initial and final state particles. Several methods exist in order to separate between  
 478 the two processes, by removing the overlap, therefore avoiding double counting. There are two different diagram  
 479 removal procedures, diagram removal procedure 1 (DR1) [44] and diagram removal procedure 2 (DR2). The DR1  
 480 scheme is used to remove the overlap (interference) between  $tWZ$  and  $t\bar{t}Z$ . A comparison to the DR2 scheme is  
 481 used to estimate part of the theoretical systematic on the modelling of the  $tWZ$  signal (See Section 4.8.2). Diboson  
 482 processes which feature the three charged leptons and one neutrino or four charged lepton in their final states,  
 483 such as  $WZ$  and  $ZZ$ , are simulated using **SHERPA 2.2.2** [27] at NLO in QCD precision. The events are interfaced  
 484 with **SHERPA** for the parton shower. Triboson processes such as  $WWW$ ,  $WWZ$ ,  $WZZ$ , and  $ZZZ$  containing up  
 485 to six leptons in their final states are simulated using **SHERPA 2.2.2** at NLO in QCD precision. The events are  
 486 interfaced with **SHERPA** for the parton shower. The production of  $t\bar{t}$  events are simulated with the **POWHEG** [51]  
 487 generator providing ME calculations at NLO in QCD. The events are interfaced with **PYTHIA 8.210** for the parton  
 488 shower. The production of  $t\bar{t}t$ ,  $t\bar{t}\bar{t}\bar{t}$  and  $t\bar{t}WW$  are simulated using the **MADGRAPH5\_AMC@NLO 2.2.2** generator at LO  
 489 in QCD precision. The events are interfaced with **PYTHIA 8.186** [88] for the parton shower. The production of  
 490  $t\bar{t}$  with an associated Higgs boson,  $t\bar{t}H$ , are generated using the **MADGRAPH5\_AMC@NLO 2.6.0** generator at NLO in  
 491 QCD precision. The events are showered using **PYTHIA 8.230** [89]. The production of  $t\bar{t}$  events are simulated with  
 492 the **POWHEG** generator providing ME calculations at NLO in QCD. The events are showered using **PYTHIA 8.230**.  
 493 The production of a single top quark in association with a  $Z$ -boson and an extra parton,  $tZq$ , is simulated using  
 494 **MADGRAPH5\_AMC@NLO 2.3.3** at NLO in QCD precision. The events are interfaced with **PYTHIA 8.230** for the parton  
 495 shower. In Table 3, the event generator and parton shower used for each process's sample are shown.

Process	Event Generator	Cross section calculation	Parton Shower
$tWZ$	<b>MADGRAPH5_AMC@NLO 2.3.3</b>	NLO	<b>PYTHIA 8.235</b>
$t\bar{t}Z$	<b>MADGRAPH5_AMC@NLO 2.3.3</b>	NLO	<b>PYTHIA 8.210</b>
$ZZ, WZ$	<b>SHERPA 2.2.2</b>	NLO	<b>SHERPA</b>
$VVV (V = W/Z)$	<b>SHERPA 2.2.2</b>	NLO	<b>SHERPA</b>
$t\bar{t}$	<b>POWHEG</b>	NLO	<b>PYTHIA 8.230</b>
$t\bar{t}W$	<b>MADGRAPH5_AMC@NLO 2.3.3</b>	NLO	<b>PYTHIA 8.210</b>
$t\bar{t}WW$	<b>MADGRAPH5_AMC@NLO 2.2.2</b>	LO	<b>PYTHIA 8.186</b>
$t\bar{t}H$	<b>MADGRAPH5_AMC@NLO 2.6.0</b>	NLO	<b>PYTHIA 8.230</b>
$t\bar{t}t, t\bar{t}\bar{t}\bar{t}$	<b>MADGRAPH5_AMC@NLO 2.2.2</b>	LO	<b>PYTHIA 8.186</b>
$tZq$	<b>MADGRAPH5_AMC@NLO 2.3.3</b>	NLO	<b>PYTHIA 8.230</b>

Table 3: The event generator and parton shower used for the signal and background process's MC samples is shown.

## 497 4.2 Objects

498 In this section the physics objects (leptons, jets and  $b$ -tagged jets) used in this analysis are outlined.

### 499 4.2.1 Leptons

500 In this analysis only  $e$  and  $\mu$  leptons are considered, since  $\tau$  leptons are difficult to detect in the ATLAS detector.  $\tau$   
 501 leptons are challenging to detect since they have an extremely short lifetime ( $290.3 \pm 0.5\text{fs}$  [55]) which causes them  
 502 to decay before reaching any detector components and therefore can only be reconstructed via their decay products.  
 503 In addition to our selection criteria of exactly four leptons, it is required that the Leading (L), Next-to-Leading  
 504 (NL), Next-to-Next-to-Leading (NNL) and Next-to-Next-to-Next-to-Leading (NNNL) leptons have  $p_T$  greater than  
 505 28, 18, 10 and 10 GeV respectively. Relatively loose object-level cuts are chosen in an attempt to maximize our  
 506 signal statistics, since the analysis is heavily statistically limited. Reconstructed electrons are required to be within  
 507  $|\eta| < 2.47$  and excluding the transition region between the barrel and end-cap calorimeters at  $1.37 < |\eta| < 1.52$ .  
 508 Reconstructed muons are required to be within  $|\eta| < 2.5$ . The transverse impact parameter,  $d_0$ , is defined as the  
 509 minimal spatial distance between the object's (referring to leptons) trajectory and the primary vertex (the vertex  
 510 associated with the  $p$ - $p$  hard scatter). The longitudinal impact parameter,  $z_0$ , is defined as the value of  $z$  of the  
 511 point on the object's trajectory which determines  $d_0$ . To ensure consistency between the lepton and the primary  
 512 vertex, it is required that  $|\frac{d_0}{\sigma(d_0)}| < 5$ ,  $|z_0 \sin \theta| < 0.5\text{ mm}$  for electrons and  $|\frac{d_0}{\sigma(d_0)}| < 3$ ,  $|z_0 \sin \theta| < 0.5\text{ mm}$  for muons,  
 513 following the current recommendations [92]. To avoid instances where one detector signal can result in multiple  
 514 different reconstructed objects, an overlap removal procedure is applied which ignores all but one of these objects  
 515 (See Section 4.2.4). Electrons are selected using a likelihood based discriminant [1] which takes measurements  
 516 from the tracking system, calorimeter system and quantities derived from both the tracking and calorimeter system  
 517 as input. Muons are selected using the Muon Selection Tool [78]. Loose electrons are defined with the criteria  
 518 above, using the LooseAndBLayerLH ( $\sim 91\%$  selection efficiency for electrons with  $E_T > 30\text{ GeV}$  [45]) identification  
 519 algorithm (which has a certain cut applied). Similarly, tight electrons are defined with the criteria above, using  
 520 the TightLH ( $\sim 80\%$  selection efficiency for electrons with  $E_T > 30\text{ GeV}$  [45]) algorithm (which has a certain cut  
 521 applied). Both loose and tight muons use the Medium ( $\sim 95\%$  selection efficiency [5]) algorithm (which has a certain  
 522 cut applied). Tight leptons additionally require that they are sufficiently isolated from other particles produced  
 523 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  
 524 summing the  $p_T$  of all the reconstructed particles surrounding the particle of interest, situated within the cone.  
 525 A quantity,  $I_{rel}$ , is then defined as,  $I_{rel} = \frac{\sum p_T(\text{surrounding candidate})}{p_T(\text{candidate})}$ , the ratio of this sum to the  $p_T$  of the lepton  
 526 candidate. If this value is large, it is likely that the particle of interest originated from a jet (together with many  
 527 other particles), whereas a prompt decay product resulting from the hard scatter will have little to no energy  
 528 surrounding it ( $I_{rel} \ll 1$ ). The IsolationSelectionTool with the PLVTight ( $\sim 70\%$  efficiency [6]) and PLVTight  
 529 ( $\sim 70\%$  efficiency at  $p_T = 30\text{ GeV}$  [37]) algorithm are used for tight electrons and tight muons respectively (following  
 530 the current recommendations [84]). In Table 4, a summary of the selection criteria for leptons is shown.

	Electrons		Muons	
	Tight	Loose	Tight	Loose
$p_T$ cuts		$p_T(\ell_1, \ell_2, \ell_3, \ell_4) > (28, 18, 10, 10) \text{ GeV}$		
Overlap Removal		Described in Section 4.2.4		
$\eta$ cuts	$ \eta(\ell_e)  < 2.47$ excluding $1.37 <  \eta(\ell_e)  < 1.52$		$ \eta(\ell_\mu)  < 2.5$	
Impact Parameters	$ \frac{d_0}{\sigma(d_0)}  < 5$ , $ z_0 \sin \theta  < 0.5 \text{ mm}$		$ \frac{d_0}{\sigma(d_0)}  < 3$ , $ z_0 \sin \theta  < 0.5 \text{ mm}$	
Identification WP	TightLH	LooseAndBLayerLH	Medium	Medium
Isolation WP	PLVTight	Not Used	PLVTight	Not Used

Table 4: A summary of the requirements applied for selecting tight and loose leptons ( $e, \mu$ ) is shown.

### 531 4.2.2 Jets

532 Jets are reconstructed using the anti- $k_t$  algorithm (See Section 3.2.6.4). The AntiKt4EMPFflowjets ( $\sim 97\%$  average  
 533 efficiency with JVT (outlined in the subsequent paragraph)  $> 0.2$  [92]) algorithm (which has a certain cut applied)  
 534 is used, following the current recommendations [92]. The jet-vertex-tagger (JVT) and the forward jet-vertex-tagger  
 535 (fJVT) [50] are likelihood discriminant which aim to suppress pile-up jets. The Medium algorithm (which has a  
 536 certain cut applied) is used for the JVT and the fJVT (following the current recommendations [83]). Additionally,  
 537 a requirement that jets have a JVT value greater than 0.5 is applied. In the same way as with leptons, ambiguities  
 538 are removed where one detector signal can result in multiple different reconstructed objects, via overlap removal

(See Section 4.2.4). Jets are required to be within  $p_T(\text{jet}) > 20 \text{ GeV}$ . These relatively loose  $p_T$  cuts are applied in an attempt to increase our limited signal statistics. A forward jet is a signature of single top quark production, jets are therefore required to have  $|\eta| < 4.5$  in order to include these forward jets.

### 4.2.3 $b$ -tagging

The DL1r  $b$ -tagger [76] was used to identify jets as  $b$ -jets (See Section 3.2.6.4). Different DL1r working points are used 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%. The efficiency of the DL1r  $b$ -tagger is measured using control samples in data and in simulation. From these measurements, correction factors are derived to correct the tagging rates in the simulation. In the case of  $b$ -tagged jets, the correction factors and their uncertainties are estimated from data using dileptonic  $t\bar{t}$  events [17, 12]. Sources of uncertainty affecting the  $b$ -tagging efficiencies are evaluated as a function of jet  $p_T$ , including bin-to-bin correlations. Since this analysis is heavily statistically limited, the amount of statistics in our regions are aimed to be maximized. In an attempt to achieve this goal in the  $t\bar{t}Z$  CR,  $b$ -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 expected significance of  $\sigma(tWZ)$  (See Section 4.4.1).

### 4.2.4 Overlap Removal Procedure

An overlap removal procedure is performed to avoid instances where one detector signal can result in multiple different reconstructed objects. The overlap removal procedure is used on pre-selected leptons and jets. It is performed sequentially, in the following steps:

1. If the separation between a pre-selected electron and pre-selected muon is within  $\Delta R < 0.01$ , or they share a track, the pre-selected electron is discarded.
2. If the separation between a jet and a pre-selected electron is within  $\Delta R < 0.2$ , the jet is discarded.
3. Any remaining electron or muon closer than  $\Delta R = 0.4$  to a jet, is discarded.
4. If the distance between a jet and a pre-selected muon is  $\Delta R < 0.4$  and the jet has more than two associated tracks, then the muon is discarded, otherwise the jet is discarded.

## 4.3 Kinematic cuts

In order to suppress potential fakes and quarkonia (low mass resonances such as  $J/\psi$  and upsilon) a requirement that all OSSF lepton pairs have an invariant mass,  $m_{\text{OSSF}}$ , greater than 10 GeV is applied. The final state lepton charges must sum to zero. Therefore a requirement of  $\sum_{i=1}^4 \text{charge}(\ell_i) = 0$  is applied. 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  $e, \mu$  reconstruction and identification in the ATLAS detector has a high efficiency [67], these OSSF leptons are used to reconstruct  $Z$  bosons with relatively high confidence. A  $Z$  candidate is defined in this analysis as an OSSF lepton pair with an invariant mass,  $m_{\text{OSSF}}$ , satisfying the condition,  $|m_{\text{OSSF}} - m_Z| < 30 \text{ GeV}$ , where  $m(Z)$  is the nominal  $Z$  boson mass (91.1876 GeV [55]). This wider mass window is used in order to cover the full range of the  $m(Z)$  distribution, in an attempt to increase the number of events which pass our baseline selections. 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.

## 4.4 Regions and Event Selection

Two  $tWZ$  SRs are defined in an attempt to suppress and constrain the  $ZZ$  background. Both  $tWZ$  SRs are required to have exactly four tight leptons, exactly one  $Z$ -boson candidate, exactly one tight  $b$ -tagged jet (from the decay of the top quark) and greater than or equal to one jet. The two  $tWZ$  SR's differ by the flavours of their

leptons which don't originate from the decay of a  $Z$ -boson (non- $Z$  leptons). The  $ZZ$  background has two  $Z$ -bosons which decay into a pair of OSSF lepton pairs, in order to mimic the  $tWZ$  signal. This is taken advantage of, to define a  $tWZ$  region enriched in  $ZZ$  background and one with a minimal  $ZZ$  background component. This is done by requiring that one of the  $tWZ$  SRs has its two non- $Z$  leptons to have opposite flavour and the other  $tWZ$  SR is required to have its non- $Z$  leptons to have the same flavour. These two disjoint  $tWZ$  SRs are named  $tWZ$  OF SR and  $tWZ$  SF SR respectively. It is therefore expected that the  $tWZ$  SF SR contains the majority of the  $ZZ$  background events across both  $tWZ$  SRs.

In order to check the modelling of the most dominant background components in our signal region,  $t\bar{t}Z$  and  $ZZb$  control regions are defined. 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 it is required that there are 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). A  $ZZb$  region is defined, 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 [3, 39], 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, resulting in an implicit requirement on the number of jets ( $N_{jet} \geq 1$ ).

Fake leptons are objects reconstructed as leptons, but do not correspond to the leptons that are of interest in our analysis. Fake leptons can be split up into two main categories, irreducible (prompt) fakes and reducible (non-prompt) fakes. Irreducible fakes are true leptons which do not come from the process of interest. Reducible fakes are objects which are mis-identified or incorrectly reconstructed as leptons. In the ATLAS detector, the probability for a fake to occur is very low. In order to constrain the fake lepton component contained within the  $t\bar{t}Z$  sample, a  $(tWZ)_{fake}$  CR is defined which is as similar as possible to the  $tWZ$  SRs but is enhanced in fakes. This is achieved by defining the  $(tWZ)_{fake}$  CR to inherit the same selection criteria as the  $tWZ$  SRs however, in this case, a requirement of exactly 3 tight leptons and exactly 1 loose (and NOT tight) lepton (since looser leptons are more likely to be fakes, compared to tighter leptons) is applied. A  $(t\bar{t}Z)_{fake}$  CR (requiring exactly 3 tight leptons and exactly 1 loose (and NOT tight) lepton) was tried as an alternative to the  $(tWZ)_{fake}$  CR, however a much larger suppression of fakes were observed in this region (compared to the  $(tWZ)_{fake}$  CR). This suppression of fakes can be explained by the extra  $b$ -tagged jet requirement (exactly two  $b$ -tagged jets are required in the  $t\bar{t}Z$  region, compared to exactly one in the  $tWZ$  SRs (See Table 5)) which causes suppression of fakes via the overlap removal procedure (See Section 4.2.4).

In Table 5, a summary of the final selection criteria and region definitions is shown.

Baseline selections				
$N_\ell = 4$				
$p_T(\ell_1, \ell_2, \ell_3, \ell_4) > (28, 10, 10, 10)$ GeV				
$p_T(\text{jet}) > 20$ GeV, $ \eta(\text{jet})  < 4.5$ , $\text{jvt} > 0.5$				
$ \eta(\ell_e)  < 2.47$ excluding $1.37 <  \eta(\ell_e)  < 1.52$				
$ \eta(\ell_\mu)  < 2.5$				
$\sum_{i=1}^4 \text{charge}(\ell_i) = 0$				
All OSSF lepton pairs require $m_{\text{OSSF}} > 10$ GeV				
Regions				
$tWZ$ OF SR	$tWZ$ SF SR	$t\bar{t}Z$ CR	$ZZb$ CR	$(tWZ)_{fake}$ CR
$N_\ell(\text{tight}) = 4$	$N_\ell(\text{tight}) = 4$	$N_\ell(\text{tight}) = 4$	$N_\ell(\text{tight}) = 4$	$N_\ell(\text{tight}) = 3$ $N_\ell(\text{loose and NOT tight}) = 1$
$N_Z \text{ candidate} = 1$	$N_Z \text{ candidate} = 1$	$N_Z \text{ candidate} = 1$	$N_Z \text{ candidate} = 2$	$N_Z \text{ candidate} = 1$
$N_{jet} \geq 1$	$N_{jet} \geq 1$	$N_{jet} \geq 2$	$N_{jet} \geq 1$	$N_{jet} \geq 1$
$N_{b\text{-jet}}(\text{tight}) = 1$	$N_{b\text{-jet}}(\text{tight}) = 1$	$N_{b\text{-jet}}(\text{tight}) \geq 1$ $N_{b\text{-jet}}(\text{loose}) \geq 0$ $N_{b\text{-jet}}(\text{tight}) + N_{b\text{-jet}}(\text{loose}) = 2$	$N_{b\text{-jet}}(\text{tight}) = 1$	$N_{b\text{-jet}}(\text{tight}) = 1$
Opp. Flavour non- $Z$ leptons	Same Flavour non- $Z$ leptons	-	-	-

Table 5: A summary of the requirements applied for selecting events in the signal and control regions is shown.

#### 4.4.1 Optimization studies for event selection

In order to find the selection criteria for jets and leptons which maximized sensitivity to the  $tWZ$  signal, studies were performed by plotting the expected significance ( $Z_\mu^{exp}$ ) and expected upper limit ( $\mu_{up}^{exp}$ ) for different selection criteria. The fitting procedure as described in Section 4.9.1 was used to calculate the expected upper limits and expected significances in this study (these metrics are fully described later in Section 4.9.1). The same selection criteria and regions defined in Table 5 was used (unless otherwise specified), except for the selection(s) which were being optimised in each case.

In Figure 8 the expected significance ( $Z_\mu^{exp}$ ) and expected upper limits ( $\mu_{up}^{exp}$ ) for different  $\eta(jet)$  cuts are shown.

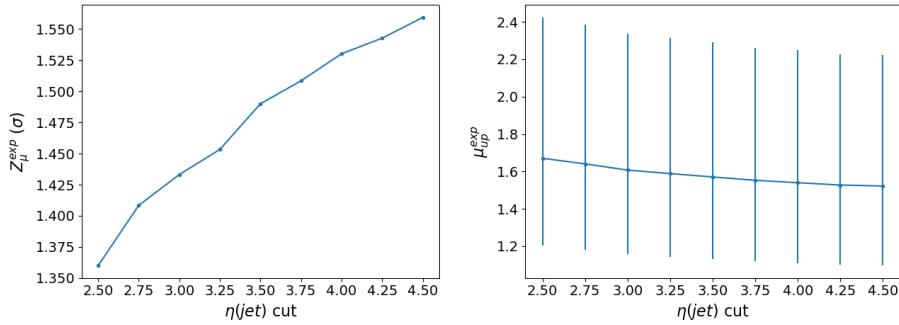


Figure 8: **Left:** Expected significance ( $Z_\mu^{exp}$ ) for different  $\eta(jet)$  cuts is shown. The cuts applied on the  $\eta(jet)$  are shown on the x-axis and corresponding expected significance from the likelihood fit is shown on the y-axis. **Right:** Expected upper limit ( $\mu_{up}^{exp}$ ) for different  $\eta(jet)$  cuts is shown. The cuts applied on the  $\eta(jet)$  are shown on the x-axis and corresponding expected upper limits are shown on the y-axis. Error bars representing the total uncertainty on the expected upper limits are shown as vertical lines.

From Figure 8, it can be seen that the  $\eta(jet)$  cut which maximizes the sensitivity of  $tWZ$  in the tetralepton channel is requiring that  $\eta(jet) < 4.5$ . This selection criteria was set for the  $\eta(jet)$  across all regions. In Figure 9 the expected significance ( $Z_\mu^{exp}$ ) and expected upper limits ( $\mu_{up}^{exp}$ ) for different  $p_T(jet)$  cuts are shown.

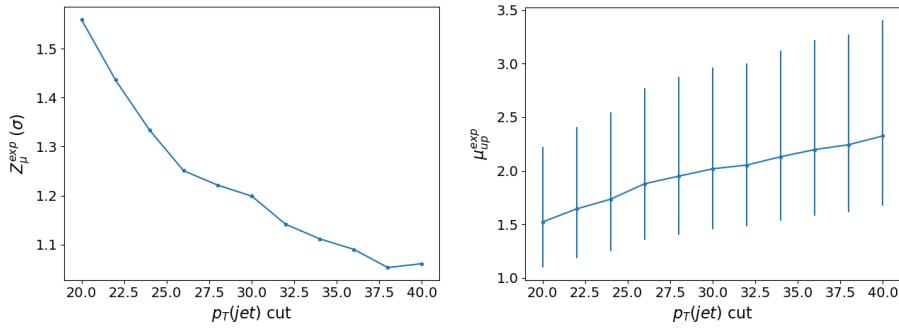


Figure 9: **Left:** Expected significance ( $Z_\mu^{exp}$ ) for different  $p_T(jet)$  cuts is shown. The cuts applied on the  $p_T(jet)$  are shown on the x-axis and corresponding expected significance from the likelihood fit is shown on the y-axis. **Right:** Expected upper limit ( $\mu_{up}^{exp}$ ) for different  $p_T(jet)$  cuts is shown. The cuts applied on the  $p_T(jet)$  are shown on the x-axis and corresponding expected upper limits are shown on the y-axis. Error bars representing the total uncertainty on the expected upper limits are shown as vertical lines.

From Figure 9, it can be seen that the  $p_T(jet)$  cut which maximizes the sensitivity of  $tWZ$  is requiring that  $p_T(jet) > 20$  GeV. This selection criteria was set for the  $p_T(jet)$  across all regions. In Figure 10 the expected significance ( $Z_\mu^{exp}$ ) and expected upper limits ( $\mu_{up}^{exp}$ ) for a range of different configurations of DL1r  $b$ -tagged jet working points across different regions.

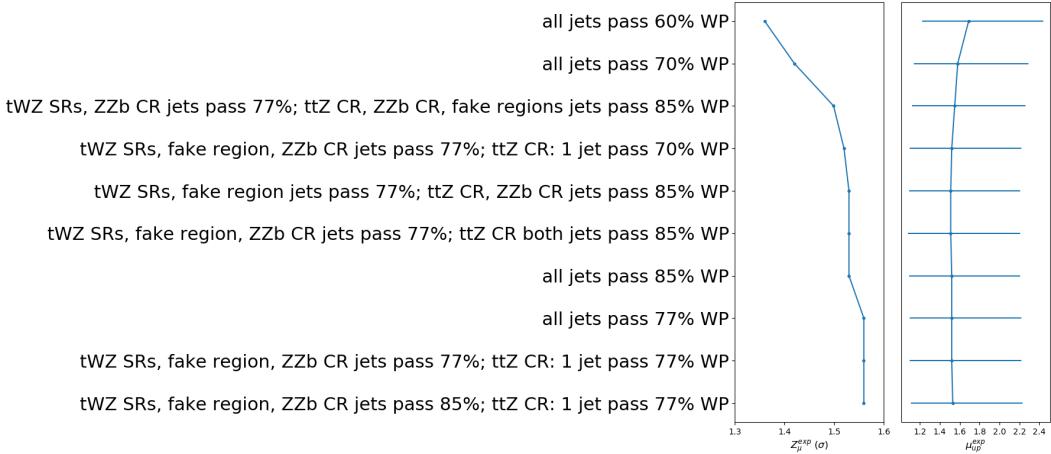


Figure 10: Expected significance ( $Z_\mu^{exp}$ ) and expected upper limit ( $\mu_{up}^{exp}$ ) for different configurations of DL1r  $b$ -tagged jet working points is shown. The common y-axis shows the different configurations of DL1r  $b$ -tagged jet working points. On the left panel, the expected significance from the likelihood fit is shown on the x-axis. On the right panel, the expected upper limit from the likelihood fit is shown on the x-axis (with the corresponding total uncertainty represented by horizontal lines).

From Figure 10, it can be seen that requiring that  $b$ -tagged jets pass the 77% DL1r WP in the  $tWZ$  SR, ( $tWZ$ )<sub>fake</sub> CR and the  $ZZb$  CR and that at least one  $b$ -tagged jet in the  $ttZ$  SR passes the 77% DL1r WP (the other jet is just required to pass the 85% DL1r WP) maximises the sensitivity overall (compared to the other investigated configurations). This configuration was chosen  $b$ -tagged jets. The  $p_T$ (L Lepton) is constrained by the single lepton triggers (Table ??). A cut was chosen to be applied on the  $p_T$ (NL Lepton) slightly tighter than the tightest single lepton  $p_T$  cut in the trigger. The  $p_T$ (NL Lepton) cut can be optimized by comparing the expected significance and limit for a range of  $p_T$ (NL Lepton) cuts to determine the cut which maximizes sensitivity. In Figure 11 the expected significance ( $Z_\mu^{exp}$ ) and expected upper limits ( $\mu_{up}^{exp}$ ) for different  $p_T$ (NL Lepton) cuts is shown.

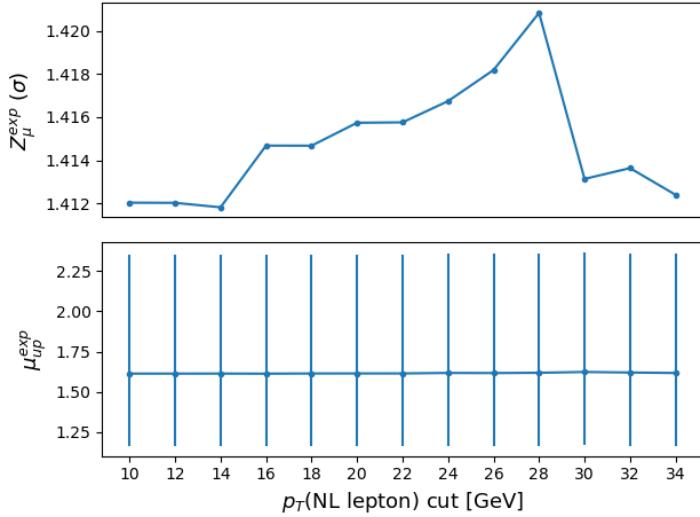


Figure 11: Expected significance ( $Z_\mu^{exp}$ ) and expected upper limit ( $\mu_{up}^{exp}$ ) for different  $p_T$ (NL Lepton) cuts is shown. The common x-axis shows cut applied to the  $p_T$  of the next-to-leading lepton. On the top panel, the expected significance from the likelihood fit is shown on the y-axis. On the bottom panel, the expected upper limit from the likelihood fit is shown on the y-axis (with the corresponding total uncertainty represented by vertical lines).

Since there is a very small change between the different  $p_T$ (NL Lepton) cuts on the sensitivity of  $tWZ$ , a  $p_T$ (NL Lepton) cut is applied at 18 GeV (avoiding a  $p_T$  cut near the sharp drop in expected significance after

642 28 GeV), therefore applying a cut above the tightest, looser dilepton trigger  $p_T$  cut (17 GeV) to suppress any  
 643 systematic from the modelling of the trigger efficiency.

## 644 4.5 Signal and Control Regions

645 In this section, expected number of events of variables in each region are shown. For each figure in this section,  
 646 the data is given by the black points and the MC predictions for each process are given by the filled histograms.  
 647 The vertical lines on the data points represent the statistical uncertainty in the data and the diagonally-lined  
 648 bands represent the total (statistical and systematic added in quadrature) uncertainty. The lower panel in each  
 649 plot shows the ratios of the data to the theoretical predictions. In order to suppress a bias towards large signal  
 650 observations in the development of the analysis, data has not been analysed in the SRs and in bins (in the CRs)  
 651 where the expected  $\frac{\text{signal}}{\text{background}}$  exceeds 0.1. Blinded bins in the CRs are shaded with black diagonal lines and their  
 652 data points are omitted. All data points are omitted in the SRs.

653 In Table 6, the expected number of events for each sample in each region are shown. The finite number of events

	$tWZ$ OF SR	$tWZ$ SF SR	$t\bar{t}Z$ CR	$ZZb$ CR	$(tWZ)$ fake CR
$t\bar{t}Z$	$13.9325 \pm 1.84643$	$10.1343 \pm 1.36039$	$31.7149 \pm 4.46776$	$5.26303 \pm 0.696828$	$19.1224 \pm 2.50011$
$t\bar{t}Z$ fakes	$0.0687541 \pm 0.0482172$	$0.032827 \pm 0.026286$	$0.0709734 \pm 0.043509$	$0.0474576 \pm 0.0301512$	$4.94775 \pm 2.48939$
$tWZ$	$3.81359 \pm 0.392241$	$2.57584 \pm 0.326401$	$2.61991 \pm 0.861557$	$1.4023 \pm 0.156686$	$4.93485 \pm 0.692143$
$ZZ$	$0.546045 \pm 0.18975$	$8.76232 \pm 2.66871$	$1.22357 \pm 0.376889$	$46.0616 \pm 13.9203$	$7.76724 \pm 2.36894$
$t\bar{t}$	$6e-06 \pm 3.04506e-06$	$0.250783 \pm 0.44226$	$0.269883 \pm 0.223373$	$6e-06 \pm 3.04506e-06$	$2.36284 \pm 0.927828$
$tZq$	$0.0827265 \pm 0.0399222$	$0.0757694 \pm 0.0355101$	$0.0637132 \pm 0.0293762$	$0.0590199 \pm 0.0244576$	$4.91371 \pm 0.754695$
$t\bar{t}W$	$0.00674747 \pm 0.00793546$	$0.00279491 \pm 0.00287747$	$6e-06 \pm 3.04506e-06$	$0.00221727 \pm 0.00562041$	$0.944039 \pm 0.296854$
$WZ$	$0.0439316 \pm 0.0241635$	$0.0397876 \pm 0.0154764$	$0.0134837 \pm 0.0128327$	$0.0474188 \pm 0.0330635$	$1.84471 \pm 0.397076$
other					
$t\bar{t}t$	$0.000987429 \pm 0.000768187$	$0.00249801 \pm 0.00138007$	$0.0141085 \pm 0.00486102$	$6e-06 \pm 3.04506e-06$	$0.0100745 \pm 0.00367677$
$t\bar{t}\bar{t}$	$0.00934516 \pm 0.0080725$	$0.0107503 \pm 0.00852049$	$0.0570846 \pm 0.0206271$	$6e-06 \pm 3.04506e-06$	$0.0216609 \pm 0.00999533$
$t\bar{t}WW$	$0.0293456 \pm 0.0263573$	$0.0296011 \pm 0.0196075$	$0.26412 \pm 0.0936908$	$0.013096 \pm 0.0323943$	$0.151267 \pm 0.0593376$
$VVV(V = W/Z)$	$0.280384 \pm 0.0866421$	$0.191257 \pm 0.0595588$	$0.0696624 \pm 0.0228108$	$0.171171 \pm 0.0526519$	$0.265957 \pm 0.0821857$
$t\bar{t}H$	$0.854064 \pm 0.177974$	$0.674566 \pm 0.141771$	$1.98187 \pm 0.406211$	$0.151447 \pm 0.0357703$	$2.22981 \pm 0.45726$
Total	$19.6684 \pm 1.95158$	$22.7832 \pm 3.10338$	$38.3633 \pm 4.6342$	$53.2187 \pm 13.9618$	$49.5163 \pm 4.77745$
data	-	-	36	49	57

Table 6: The expected number of events for each sample in each region is shown.

654 expected to be observed in data (MC simulation) carries an associated statistical uncertainty. To first order, this  
 655 uncertainty can be written as the square root of the expected number of events to be observed in data. In contrast  
 656 to this, predictions based on MC simulation carry uncertainties due to the finite number of simulated events utilised.  
 657 This uncertainty can be quantified by the Number of Equivalent Events [47],  $N_{equiv}$ , which relates the sample of  
 658  $N$  events (weighted by MC event weights) to  $N_{equiv}$  events with all MC event weights equal to 1, that would have  
 659 the same relative statistical fluctuation. The Number of Equivalent Events,  $N_{equiv}$ , can be written as,

$$N_{equiv} = \frac{(\sum_i^N w_i)^2}{\sum_i^N w_i^2} \quad (4.1)$$

660 where  $w_i$  is the MC event weight for event  $i$ . The standard uncertainty of  $N_{equiv}$  is given by  $u(N_{equiv}) = \sqrt{N_{equiv}}$ .  
 661 The Number of Equivalent Events for each sample in each region can be studied in order to ensure that the number  
 662 of events simulated for a given process is large in comparison to the number of events expected for that process in  
 663 data, thereby ensuring that uncertainties from MC statistics will be small (or sub-leading). In Table 7, the number  
 664 of equivalent events,  $N_{equiv}$  and its percentage uncertainty ( $\frac{u(N_{equiv})}{N_{equiv}} \times 100$ ), is shown for each sample in each region.

665  
 666  
 667  $N_{equiv}$  is much larger compared to the number of expected events (See Table 6) for the signal and background  
 668 processes in all regions. This tells us that there is a large number of simulated events for these samples. Therefore  
 669 ensuring that uncertainties resulting from MC statistics will be small (or sub-leading).

### 670 4.5.1 $tWZ$ OF SR

671 In this section, expected number of events of variables in the  $tWZ$  OF SR are shown.

672

	$tWZ$ OF SR	$tWZ$ SF SR	$t\bar{t}Z$ CR	$ZZb$ CR	$(tWZ)_{fake}$ CR			
	$N_{equiv}$	Uncertainty [%]	$N_{equiv}$	Uncertainty [%]	$N_{equiv}$	Uncertainty [%]	$N_{equiv}$	Uncertainty [%]
$tWZ$	$6463.29 \pm 80.39$	1.24	$4153.0 \pm 64.44$	1.55	$4800.67 \pm 69.29$	1.44	$2497.07 \pm 49.97$	2.0
$t\bar{t}Z$	$1363.87 \pm 36.93$	2.71	$1031.04 \pm 32.11$	3.11	$3237.01 \pm 56.89$	1.76	$561.41 \pm 23.69$	4.22
$ZZ$	$50.89 \pm 7.13$	14.02	$975.61 \pm 31.23$	3.2	$267.85 \pm 16.37$	6.11	$7023.35 \pm 83.81$	1.19
other	$748.0 \pm 27.35$	3.66	$2.47 \pm 1.57$	63.58	$4.23 \pm 2.06$	48.6	$255.32 \pm 15.98$	6.26
$t\bar{t}Z$ fakes	$6.72 \pm 2.59$	38.56	$1.31 \pm 1.14$	87.34	$16.05 \pm 4.01$	24.96	$7.15 \pm 2.67$	37.41
Total	$8632.77 \pm 92.91$	1.08	$6163.44 \pm 78.51$	1.27	$8325.81 \pm 91.25$	1.1	$10344.3 \pm 101.71$	0.98
							$12044.27 \pm 109.75$	0.91

Table 7: The number of equivalent events,  $N_{equiv}$  and its percentage uncertainty ( $\frac{u(N_{equiv})}{N_{equiv}} \times 100$ ), is shown for each sample in each region.

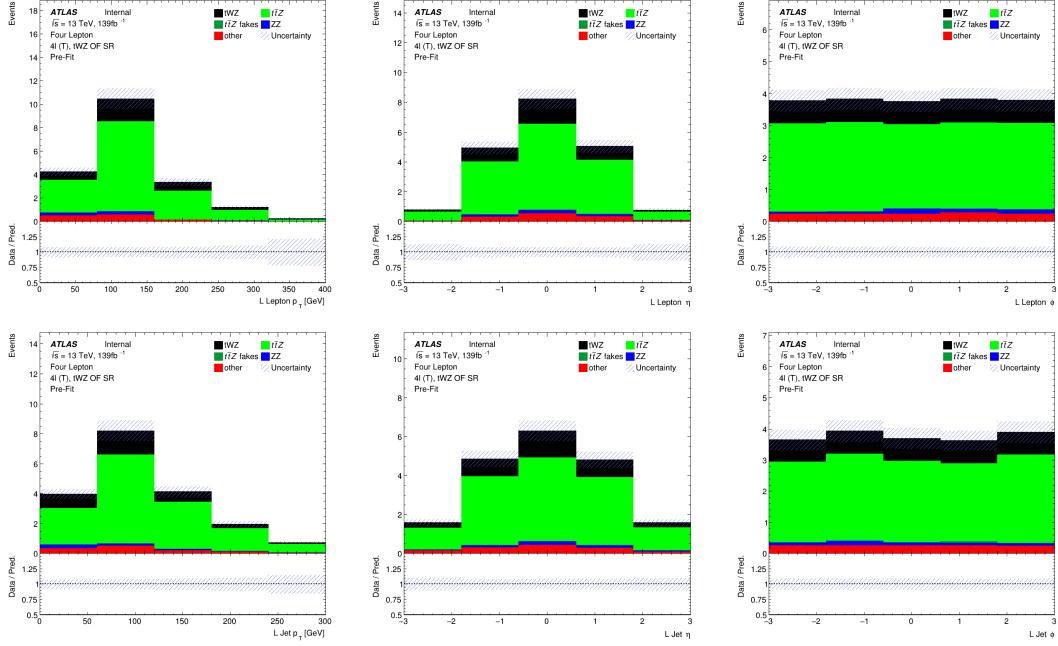


Figure 12: Comparisons of simulation and data of  $p_T$ ,  $\eta$  and  $\phi$  for leading (L) leptons (top row) and leading (NL) jets (bottom row) in the  $tWZ$  OF SR are shown.

In Figure 12, comparisons of simulation and data of  $p_T$ ,  $\eta$  and  $\phi$  for leading (L) leptons and leading (NL) jets in the  $tWZ$  OF SR are shown.

All bins for each plot in Figure 12 have  $\frac{signal}{background}$  exceeding 0.1 and are therefore blinded. This region is therefore enriched in  $tWZ$  signal events.

677

In Figure 13, comparisons of simulation and data of  $H_T$  (scalar sum of Jet  $p_T$ ), the Number of jets, the scalar sum of  $b$ -tagged jet  $p_T$  and the number of  $b$ -tagged jets in the  $tWZ$  OF SR are shown.

All bins for each plot in Figure 13 have  $\frac{signal}{background}$  exceeding 0.1 and are therefore blinded. This region is therefore enriched in  $tWZ$  signal events.

#### 4.5.2 $tWZ$ SF SR

In this section, expected number of events of variables in the  $tWZ$  SF SR are shown.

684

In Figure 14, comparisons of simulation and data of  $p_T$ ,  $\eta$  and  $\phi$  for leading (L) leptons and leading (NL) jets in the  $tWZ$  SF SR are shown.

All bins for each plot in Figure 14 have  $\frac{signal}{background}$  exceeding 0.1 and are therefore blinded. This region is therefore enriched in  $tWZ$  signal events.

689

In Figure 15, comparisons of simulation and data of  $H_T$  (scalar sum of Jet  $p_T$ ), the Number of jets, the scalar sum of  $b$ -tagged jet  $p_T$  and the number of  $b$ -tagged jets in the  $tWZ$  SF SR are shown.

The vast majority of bins for each plot in Figure 15 have  $\frac{signal}{background}$  exceeding 0.1 and are therefore blinded. This

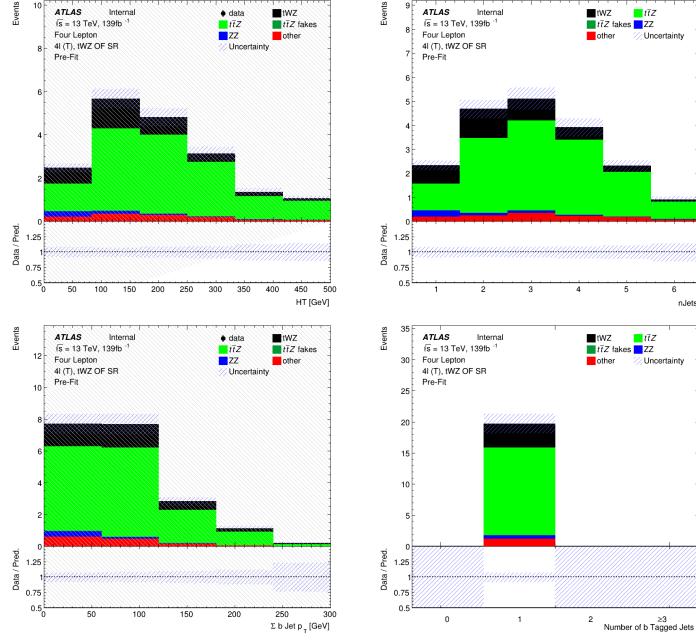


Figure 13: Comparisons of simulation and data of  $H_T$  (scalar sum of Jet  $p_T$ ), the Number of jets, the scalar sum of  $b$ -tagged jet  $p_T$  and the number of  $b$ -tagged jets (top left to bottom right) in the  $tWZ$  OF SR are shown.

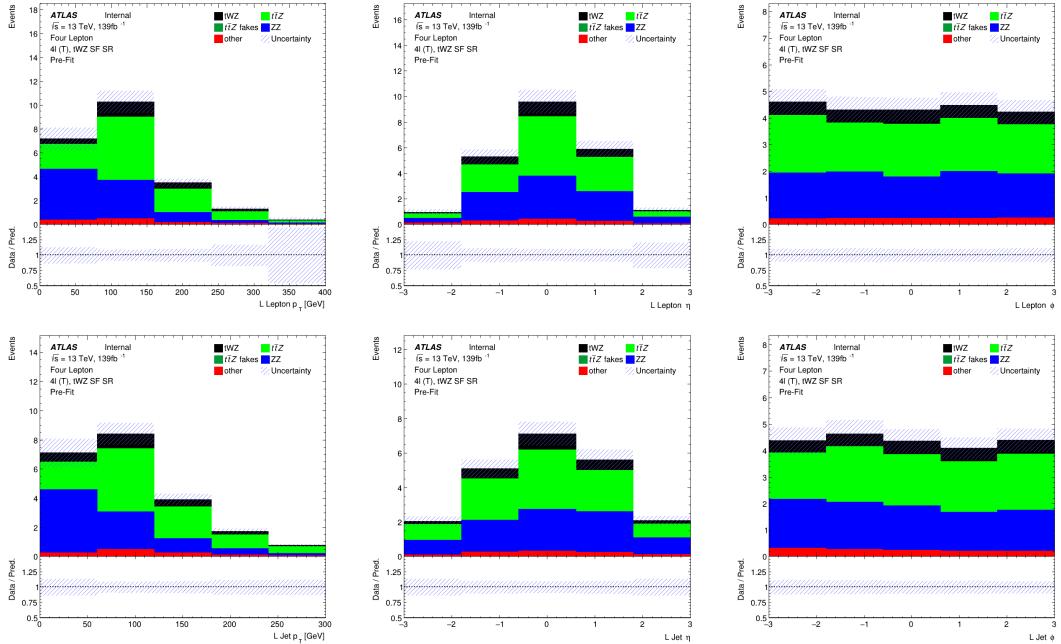


Figure 14: Comparisons of simulation and data of  $p_T$ ,  $\eta$  and  $\phi$  for leading (L) leptons (top row) and leading (NL) jets (bottom row) in the  $tWZ$  SF SR are shown.

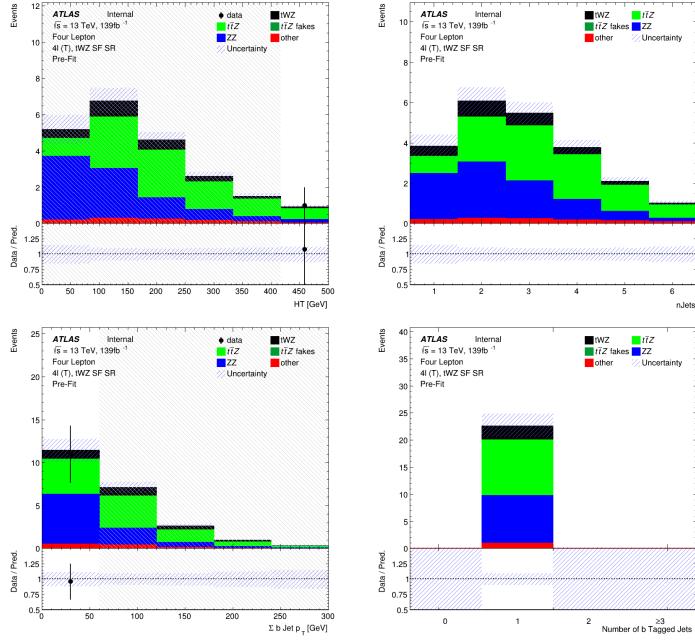


Figure 15: Comparisons of simulation and data of  $H_T$  (scalar sum of Jet  $p_T$ ), the Number of jets, the scalar sum of  $b$ -tagged jet  $p_T$  and the number of  $b$ -tagged jets (top left to bottom right) in the  $tWZ$  SF SR are shown.

region is therefore enriched in  $tWZ$  signal events. The deviations in data and simulation in the two bins (in the  $HT$  and  $\sigma b$  jet  $p_T$  distributions) which are not blinded, are within the expected uncertainties.

### 4.5.3 $t\bar{t}Z$ CR

In this section, expected number of events of variables in the  $t\bar{t}Z$  CR are shown.

In Figure 16, comparisons of simulation and data of  $p_T$ ,  $\eta$  and  $\phi$  for leading (L) leptons and leading (NL) jets in the  $t\bar{t}Z$  CR are shown.

The majority of the deviations in data and simulation for each plot in Figure 16 are within the expected uncertainties. The few plots which have bins where there is a disagreement between data and simulation are either within  $2\sigma$  (L Jet  $\phi$ ) or  $3\sigma$  (L Jet  $\eta$ ) standard uncertainties from one another, or are show more than a  $3\sigma$  (L Lepton  $p_T$ ) disagreement. The disagreement in the L Lepton  $p_T$  distribution could be due to statistical fluctuations in data or simulation, since there are so few events in these bins.

In Figure 17, comparisons of simulation and data of  $H_T$  (scalar sum of Jet  $p_T$ ), the Number of jets, the scalar sum of  $b$ -tagged jet  $p_T$  and the number of  $b$ -tagged jets in the  $t\bar{t}Z$  CR are shown.

Almost all of the deviations in data and simulation for each plot in Figure 17 are within the expected uncertainties. There is a  $2\sigma$  disagreement in one of the bins in the  $n\text{Jets}$  distribution and a large disagreement ( $> 5\sigma$ ) in one of the bins in the  $HT$  distribution. The large disagreement between data and simulation in the  $HT$  distribution is surprising since all other bins in the distribution agree within  $1\sigma$  uncertainties, and it is therefore not fully understood.

### 4.5.4 $ZZb$ CR

In this section, expected number of events of variables in the  $ZZb$  CR are shown.

In Figure 18, comparisons of simulation and data of  $p_T$ ,  $\eta$  and  $\phi$  for leading (L) leptons and leading (NL) jets in the  $ZZb$  CR are shown.

Most of the deviations in data and simulation for each plot in Figure 18 are within the expected uncertainties. There are a few bins with  $2\sigma$  and  $> 2\sigma$  disagreements between data and simulation in the L Lepton  $p_T$ , L Lepton  $\eta$  and L Jet  $p_T$  distributions, with the disagreement being much more noticeable in the L Lepton distributions.

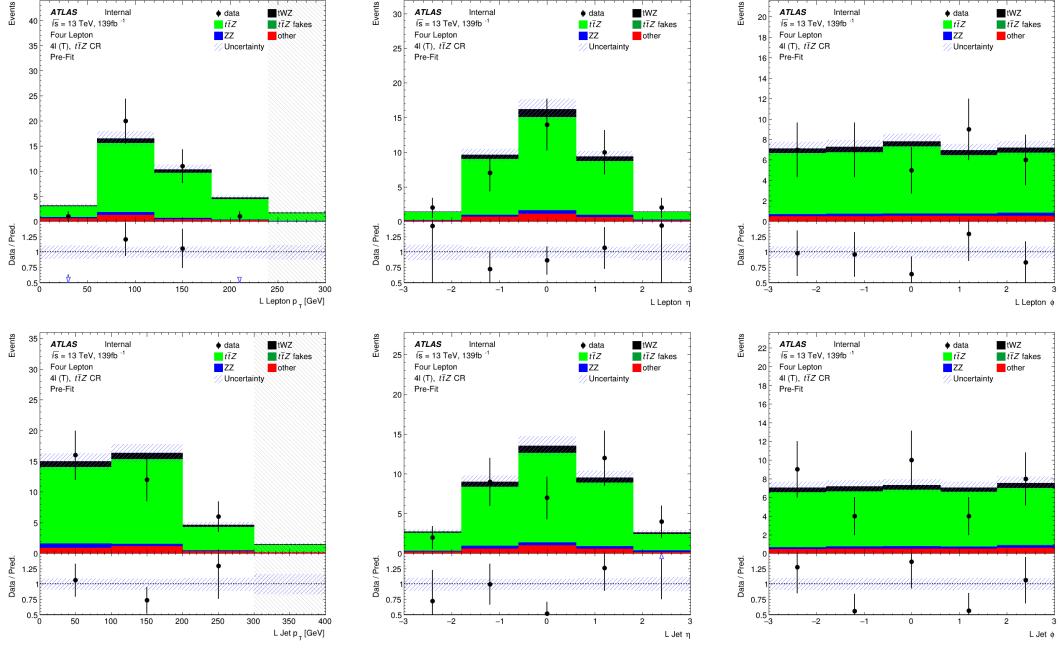


Figure 16: Comparisons of simulation and data of  $p_T$ ,  $\eta$  and  $\phi$  for leading (L) leptons (top row) and leading (NL) jets (bottom row) in the  $t\bar{t}Z$  CR are shown.

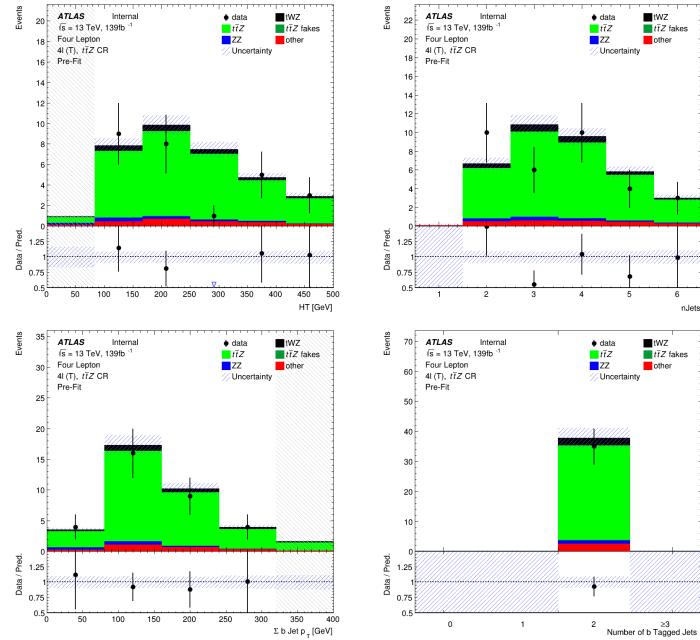


Figure 17: Comparisons of simulation and data of  $H_T$  (scalar sum of Jet  $p_T$ ), the Number of jets, the scalar sum of  $b$ -tagged jet  $p_T$  and the number of  $b$ -tagged jets (top left to bottom right) in the  $t\bar{t}Z$  CR are shown.

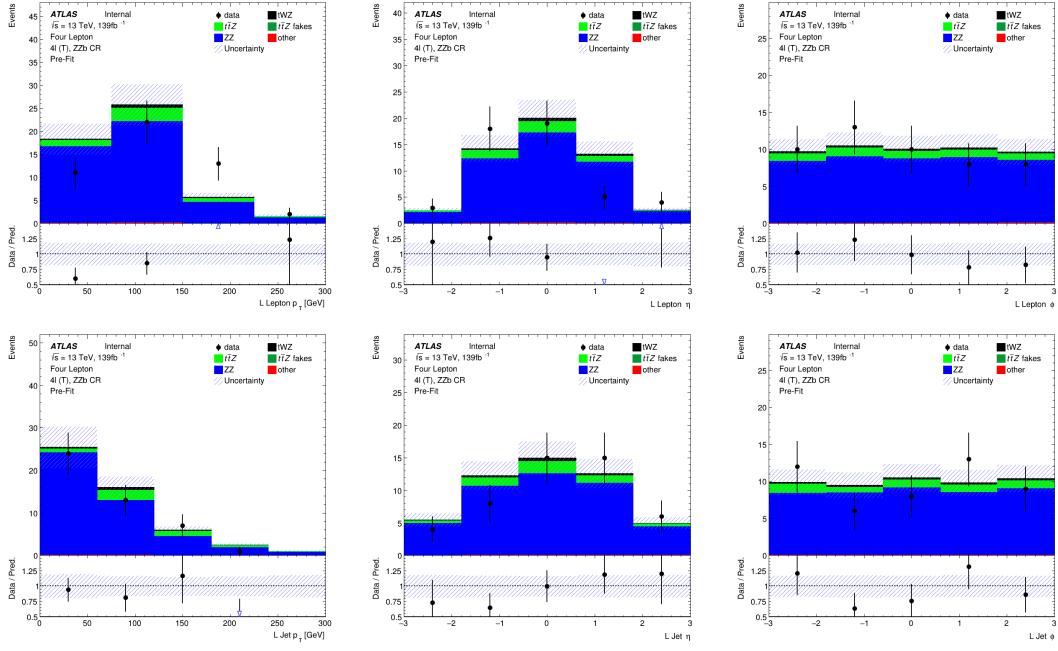


Figure 18: Comparisons of simulation and data of  $p_T$ ,  $\eta$  and  $\phi$  for leading (L) leptons (top row) and leading (NL) jets (bottom row) in the  $ZZb$  CR are shown.

This could suggest some mis-modelling for L Leptons in this region.

In Figure 19, comparisons of simulation and data of  $H_T$  (scalar sum of Jet  $p_T$ ), the Number of jets, the scalar sum of  $b$ -tagged jet  $p_T$  and the number of  $b$ -tagged jets in the  $ZZb$  CR are shown.

Most of the deviations in data and simulation for each plot in Figure 18 are within the expected uncertainties.

#### 4.5.5 $(tWZ)_{\text{fake}}$ CR

In this section, expected number of events of variables in the  $(tWZ)_{\text{fake}}$  CR are shown.

In Figure 20, comparisons of simulation and data of  $p_T$ ,  $\eta$  and  $\phi$  for leading (L) leptons and leading (NL) jets in the  $(tWZ)_{\text{fake}}$  CR are shown.

The vast majority of bins for each plot in Figure 20 have  $\frac{\text{signal}}{\text{background}}$  exceeding 0.1 and are therefore blinded. This region is therefore enriched in  $tWZ$  signal events. Most deviations in data and simulation in the bins which are not blinded, are within the expected uncertainties. Only two out of seven unblinded bins are not within expected uncertainties and are within a  $2\sigma$  uncertainty.

In Figure 21, comparisons of simulation and data of  $H_T$  (scalar sum of Jet  $p_T$ ), the Number of jets, the scalar sum of  $b$ -tagged jet  $p_T$  and the number of  $b$ -tagged jets in the  $(tWZ)_{\text{fake}}$  CR are shown.

The majority of bins for each plot in Figure 21 have  $\frac{\text{signal}}{\text{background}}$  exceeding 0.1 and are therefore blinded. This region is therefore enriched in  $tWZ$  signal events. Most deviations in data and simulation in the bins which are not blinded, are within the expected uncertainties. Only two out of seven unblinded bins are not within expected uncertainties and are within a  $2\sigma$  uncertainty.

## 4.6 Fake Lepton Estimation

Fake leptons are objects reconstructed as leptons, but do not correspond to the leptons which originate from the process of interest. The sources of fake leptons include those originating from heavy hadron decays, light hadron decays or via the conversion of a photon to a lepton. In the ATLAS detector, the probability for a fake to occur is very low. In this section, the method used to estimate the fake lepton contribution is described.

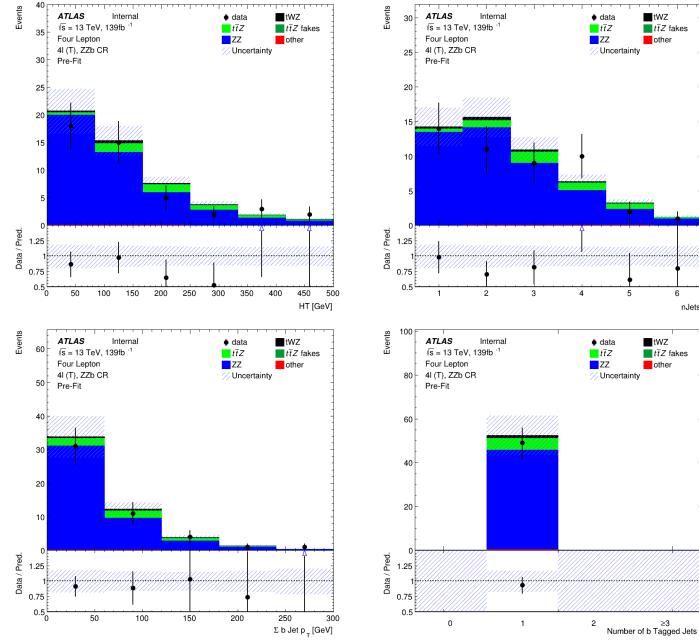


Figure 19: Comparisons of simulation and data of  $H_T$  (scalar sum of Jet  $p_T$ ), the Number of jets, the scalar sum of  $b$ -tagged jet  $p_T$  and the number of  $b$ -tagged jets (top left to bottom right) in the  $ZZb$  CR are shown.

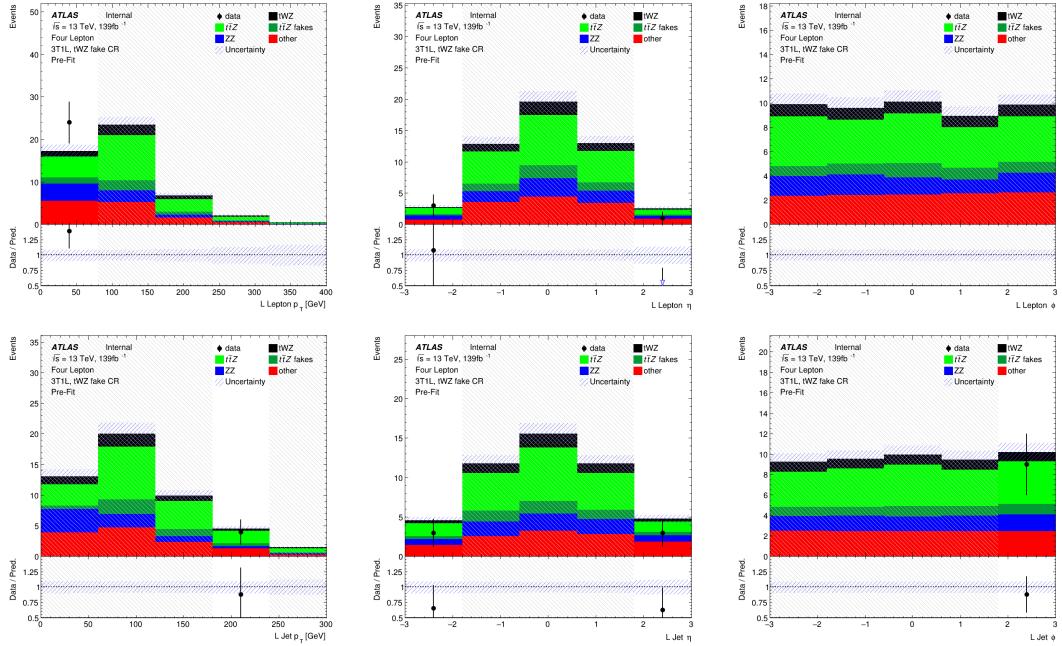


Figure 20: Comparisons of simulation and data of  $p_T$ ,  $\eta$  and  $\phi$  for leading (L) leptons (top row) and leading (NL) jets (bottom row) in the  $(tWZ)_{\text{fake}}$  CR are shown.

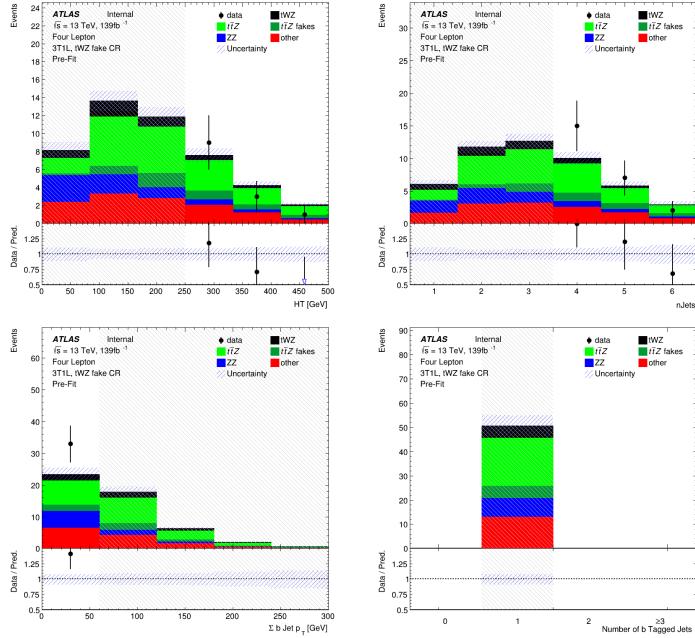


Figure 21: Comparisons of simulation and data of  $H_T$  (scalar sum of Jet  $p_T$ ), the Number of jets, the scalar sum of  $b$ -tagged jet  $p_T$  and the number of  $b$ -tagged jets (top left to bottom right) in the  $(tWZ)_{\text{fake}}$  CR are shown.

As  $t\bar{Z}$  is the dominant background process ( $\sim 75\%$  of the total background contribution), it is assumed that  $t\bar{Z}$  will also dominate the events containing fake leptons. The fake lepton efficiency,  $\epsilon$ , can be written as  $\epsilon = \frac{N_{\text{fake}}^{\text{tight}}}{N_{\text{fake}}^{\text{loose}}}$ , where  $N_{\text{fake}}^{\text{tight}}$  is the number of fake leptons which pass the tight lepton selection (See Section 4.2.1) and  $N_{\text{fake}}^{\text{loose}}$  is the number of fake leptons which pass the loose lepton selection (See Section 4.2.1). The probability of one fake lepton to occur,  $P(\text{one fake } \ell)$ , is proportional to  $\epsilon_1 \ll 1$  [65, 82] and the probability for two fakes to occur is,  $P(\text{two fakes } \ell)$ , is proportional to  $\epsilon_2 < \epsilon_1 \ll 1$ . In this analysis, an estimation of the fake lepton component to the highest order is investigated and therefore the case where at least one fake lepton occurs in a  $t\bar{Z}$  event is considered.

Firstly, the dominant  $t\bar{Z}$  background is split up into  $t\bar{Z}$  and  $(t\bar{Z})_{\text{fake}}$  components. Secondly, a  $(tWZ)_{\text{fake}}$  CR (See Section 4.4) is defined which is enhanced in fakes and aims to constrain the  $(t\bar{Z})_{\text{fake}}$  background in the SR. All events which contribute to the  $(t\bar{Z})_{\text{fake}}$  background are determined by the **IFF Truth Classifier** [59]. The **IFF Truth Classifier** is a tool which aims to classify leptons based off their truth information. It uses the more general **MCTruthClassifier** [70] tool's output as input and returns one of the following lepton categories: **Unknown**, **KnownUnknown** (leptons which can (in principle) be classified, but the **MCTruthClassifier** fails to classify the lepton's truth type or origin), **IsoElectron**, **ChargeFlipIsoElectron**, **PromptMuon**, **PromptPhotonConversion**, **ElectronFromMuon**, **TauDecay**, **BHadronDecay**, **CHadronDecay** or **LightFlavorDecay** (More details [60]). Given these categories, leptons are considered as fake if they are classified as **PromptPhotonConversion**, **BHadronDecay**, **CHadronDecay** or **LightFlavorDecay** (i.e. a lepton originating from the decay of a  $b$ -Hadron,  $c$ -Hadron or light-flavour jet). Events which contribute to the  $(t\bar{Z})_{\text{fake}}$  background are those where at least one lepton from the  $t\bar{Z}$  sample are classified by the **IFF Truth Classifier** with one of the four aforementioned categories.

The  $(tWZ)_{\text{fake}}$  CR aims to be as similar as possible to the  $tWZ$  SRs, but enhanced in fakes. This CR can then be used to constrain the normalisation of the  $(t\bar{Z})_{\text{fake}}$  template. To ensure that this region is enhanced in fakes, it is required that it contains 3 tight leptons and 1 loose lepton, since loose leptons are more likely to be fakes. Leptons from heavy decays are produced in jets and are typically surrounded by other energetic particles. Since the loose lepton definition relaxes the isolation requirement, leptons satisfying the loose criteria are more enhanced in these fake leptons. By using the  $p_T$  of the loose lepton ( $p_T(\text{Loose Lepton})$ ) in this region as the variable used in the fit, the shape (and normalisation) of the  $(t\bar{Z})_{\text{fake}}$  template can be constrained.

In Figure 22, the number of leptons classified as fake, split up by their IFF Truth classification, in each region is shown.

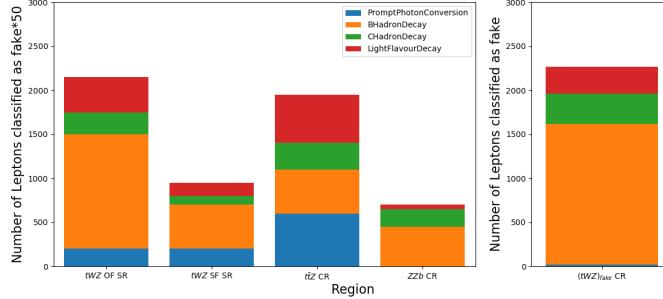


Figure 22: The number of leptons classified as fake, split up by their IFF Truth classification, in each region is shown. The left panel shows the number of leptons classified as fakes, scaled by a factor of 50, on the y-axis. The right panel shows the number of leptons classified as fakes (unscaled), on the y-axis. The different signal and control regions are shown on the x-axes of the left and right panels. The IFF truth classification of the leptons are shown in the legend and correspond to the different coloured stacked histograms.

779 Around 50 times more fake leptons pass our selection criteria for the  $(tWZ)_{\text{fake}}$  CR, compared to remaining four  
 780 regions. This relative abundance of fake leptons present in the  $(tWZ)_{\text{fake}}$  CR further justifies our use of this region  
 781 to constrain the fake lepton component. In Figure 23, the relative dominance of the different classifications for fake  
 782 leptons (classified by the IFF truth classified) in each region, is shown.

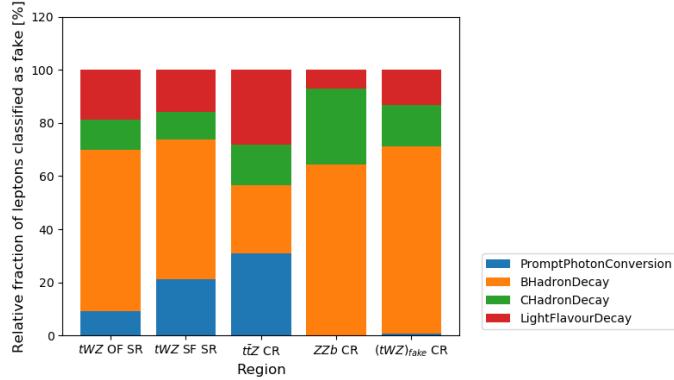


Figure 23: The relative dominance of the different classifications for fake leptons (classified by the IFF truth classified) in each region, is shown. The relative dominance of leptons classified as fakes, as a fraction of the total number of fake leptons (in each region), is shown on the y-axis. The different signal and control regions are shown on the x-axis. The IFF truth classification of the leptons are shown in the legend and correspond to the different coloured stacked histograms.

783 The majority of fake leptons which pass our selection criteria originate from the decay of  $b$ -hadrons, in all regions  
 784 but the  $t\bar{t}Z$  CR. The smaller proportion of fake leptons originating from  $b$ -hadron decays in the  $t\bar{t}Z$  CR could  
 785 possibly be due to statistical fluctuations resulting from the low number of fake leptons which pass our selection  
 786 criteria in this region ( $\sim 40$  fake leptons). In Figure 24, the amount of fake and real  $t\bar{t}Z$  events which pass our  
 787 selection criteria, in each region, is shown.

788 Around 20% of all  $t\bar{t}Z$  events are classified as fake events (having one or more of its leptons being classified as fake)  
 789 in the  $(tWZ)_{\text{fake}}$  CR. The  $tWZ$  OF SR,  $tWZ$  SF SR,  $t\bar{t}Z$  CR and  $ZZb$  CR have less than 1% of their total  $t\bar{t}Z$   
 790 events being fake. The non-negligible amount of fake  $t\bar{t}Z$  events present in the  $(tWZ)_{\text{fake}}$  CR, allows the  $t\bar{t}Z$  fake  
 791 background to be constrained by the  $(tWZ)_{\text{fake}}$  CR.

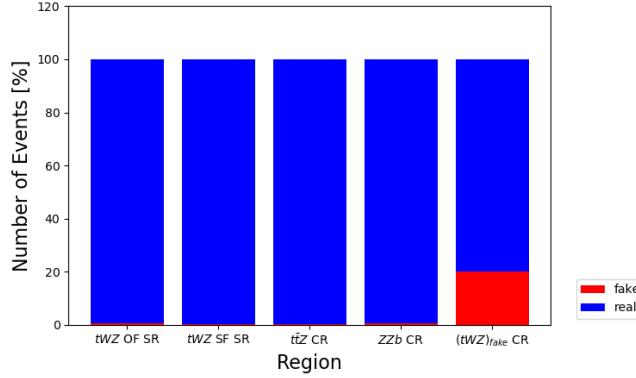


Figure 24: The percentage of fake and real  $t\bar{t}Z$  events which pass our selection criteria, in each region, is shown. The relative number of fake and real events (in % of the total number of events in the nominal and fake  $t\bar{t}Z$  background samples) is shown on the y-axis. The different signal and control regions are shown on the x-axis. The blue and red histograms represent the percentage of real and fake events (out of the total number of events in the nominal and fake  $t\bar{t}Z$  background samples), respectively.

## 792 4.7 Machine Learning Techniques

793 The presence of different numbers of top quarks is a key discriminator between signal and the dominant background  
 794 process,  $t\bar{t}Z$ . This information is aimed to be exploited by reconstructing  $\ell b$  systems as a proxy for top quarks  
 795 (since,  $t \rightarrow W(\rightarrow \ell\nu)b$ ). This is done in two ways, firstly by implementing a Boosted Decision Tree (BDT)  
 796 which is used to classify  $\ell b$  systems originating from top quarks, and secondly, by implementation of a kinematic  
 797 reconstruction algorithm (Two Neutrino Scanning Method) which aims to determine the likelihood of an event  
 798 containing two top quarks. In this thesis, this BDT is referred to as an *object-level* BDT.

799 Certain variables constructed from event information show discrimination between signal and background events.  
 800 This information can be exploited to discriminate between signal and background events by constructing an algo-  
 801 rithm which takes advantage of these discriminating variables. A BDT is implemented and is used to discriminate  
 802 between  $tWZ$  events and its major backgrounds,  $t\bar{t}Z$  and  $ZZ$ . In this thesis, this BDT is referred to as an *event-level*  
 803 BDT. The discriminator output from the object-level BDT can be converted to a variable which can then be used  
 804 as an input to the event-level BDT.

### 806 4.7.1 Boosted Decision Trees

807 Machine Learning techniques can be used to build multivariate algorithms that exploit information from many  
 808 weak discriminators (a model that is used to classify data in a dataset) to form one strong discriminator. A BDT  
 809 is a Machine Learning technique which classifies data in a dataset into different categories by iteratively applying  
 810 binary cuts on features of the data (variables, in the context of this analysis) to segregate the data [100]. The  
 811 method in which a BDT combines weak discriminators to build a single strong discriminator is called *boosting*. In  
 812 boosting, weak discriminators are sequentially combined, where each model iteration is fitted to the residuals from  
 813 the previous step, such that the model performance improves [56]. A few concepts related to Machine Learning  
 814 and BDTs that are used in this analysis, are described briefly in the proceeding text.

815 Performance metrics can be used to evaluate how well a model performs in a classification problem [74]. A perfor-  
 816 mance metric used extensively in this analysis is the *accuracy* of a model. The accuracy is defined as the percentage  
 817 of correct predictions for the test dataset ( $\text{accuracy} = \frac{\text{correct number of predictions}}{\text{total number of predictions}}$ ). Machine Learning models can be  
 818 susceptible to learning a training dataset too well, in such a way as to negatively affect its performance on unseen  
 819 data. This is known as *over-training*. Over-training occurs when noise or random fluctuations in the training  
 820 dataset are learnt by the model [29]. Cross Validation [77] is a procedure used to evaluate a Machine Learning  
 821 algorithm. Cross validation gives an estimate on how the model is expected to perform on unseen data and it  
 822 can be useful tool to protect against over-training. In this analysis we use a type of cross validation called, *k-fold*  
 823 cross validation. In k-fold cross validation, the training dataset is randomly split up into k subsets, or folds, of

825 approximately equal size. A fold is defined as a test dataset and the remaining k-1 folds are used to train the model.  
 826 The model is then evaluated on the test set and a performance metric (or multiple) is evaluated. This procedure is  
 827 performed once on each unique fold. Hyper-parameters are user-defined parameters of a model that are govern the  
 828 entire training process. Typical examples of hyper-parameters include the learning rate, the number of discriminators  
 829 and the type of loss function to be minimised. The learning rate determines the step size at each iteration in  
 830 determining the minimum of the loss function. Hyper-parameter optimisation is a process which aims to determine  
 831 the best hyper-parameters for a model, based off some performance metric. In this analysis hyper-parameter  
 832 optimisation is performed using a *grid search*. In a grid search, a user-defined list of hyper-parameter values  
 833 are chosen for each hyper-parameter that one aims to optimise. The model is then trained using each permutation  
 834 of hyper-parameters and determines the set of hyper-parameters in which the performance metric is maximised.

835  
 836 BDTs are chosen to use in this analysis, as opposed to another ML algorithm, since they are not very sensitive to  
 837 over-training and perform well with minimal optimisation or tweaking of the hyper-parameters. A multi-layered  
 838 sequential neutral network was tried, however, it was out-performed by a BDT. More specifically, Scikit-Learn’s  
 839 `GradientBoostingClassifier` [85] was used.

#### 840 4.7.2 Object-level BDT

841 The object-level BDT was trained on an alternative  $t\bar{t}$  sample (simulated using the same generator, parton shower  
 842 and to the same order of QCD as the  $t\bar{t}$  sample described in Section 4.1.2) with a baseline selection of exactly 1 tight  
 843 lepton with  $p_T > 28$  GeV. Additionally, jets in this sample are required to have  $p_T > 20$  GeV. Jets are identified  
 844 as  $b$ -tagged jets by the 77% DL1r working point. These baseline selections were chosen to mimic those used in the  
 845 event selection of the analysis (outlined in Table 5). The leptons and  $b$ -jets used in training the object-level BDT  
 846 are required to pass the aforementioned baseline selections. This alternative  $t\bar{t}$  sample was utilised in training the  
 847 BDT to avoid using a subset of events from the MC samples used in the rest of the analysis, therefore maximizing  
 848 the amount of generated events available to use in other parts of the analysis.

849  
 850 The signal class is defined to consist of reconstructed  $\ell b$  systems (defined as the sum of the 4-vectors of a lepton  
 851 and a  $b$ -tagged jet) originating from top quarks which are well matched to their truth counterparts. All possible  
 852 combinations of  $\ell$  and  $b$ -tagged jets are selected from the events. In particular, it is required that  $\Delta R$  between  
 853 the reconstructed and truth  $\ell b$  system is less than 0.05. An additional requirement is implemented such that the  
 854 reconstructed lepton and the truth top quark have charges with the same sign (since  $t \rightarrow b\ell^+\bar{\nu}_\ell$  and  $\bar{t} \rightarrow \bar{b}\ell^-\nu_\ell$ ).  
 855 The background class is defined to consist of all reconstructed  $\ell b$  systems which fail to pass the criteria for  $\ell b$   
 856 systems which are labelled as signal. These definitions for the signal and background classes ensure that the signal  
 857 class consists of mostly  $\ell b$  systems originating from top quarks and the background class consists of mostly  $\ell b$   
 858 systems which do not originate from top quarks.

859  
 860 The variables chosen to be used as inputs to the object-level BDT are relate to measurable quantities of  $\ell b$   
 861 systems. The optimum values for the hyper-parameters used were determined via the use of a grid-search (See  
 862 Section 4.7.1) that determined the set of hyper-parameters which maximized the mean accuracy (based off 5 fold  
 863 kfold cross-validation). After hyper-parameter optimisation, the mean accuracy of each fold increased from 0.76  
 864 to 0.77 ( $\sim 1\%$  increase). Variables can be assigned a score called *variable importance*, based on their usefulness  
 865 on predicting a target variable (in this case, a signal or background event). The variable importance for any  
 866 given variable was obtained by computing the mean accuracy of the model, removing the variable from training,  
 867 retraining the model and computing the mean accuracy of this new model. The difference between mean accuracies  
 868 of the unaltered model and the retrained model (after removal of the variable) gives us the variable importance of  
 869 the variable of interest. This method returns positive values for variables which increase the mean accuracy of the  
 870 model and negative values for variables which decrease the mean accuracy of the model. Variables with negative  
 871 variable importances were completely removed from training.

872  
 873 In Table 8, the variables used in training the object-level BDT are shown.  
 874 In Figure 25, normalised distributions of the variables used in the object-level BDT, for the signal and background  
 875 classes are shown.  
 876 The variables used in the object-level BDT show a clear distinction between signal and background  $\ell b$  systems.  
 877  
 878 The modelling of the variables used in the object-level BDT can be checked by studying the agreement between  
 879 data and simulation in the  $t\bar{t}Z$  CR. In Figure 26, MC predictions for the variables used in the object-level BDT in

Variable	Description	Variable Importance
$m(\ell b)$	Invariant mass of the $\ell b$ system	0.0025
$p_T(\ell b)$	$p_T$ of the $\ell b$ system	0.0005
$\Delta\eta(\ell, b)$	$\Delta\eta$ between the $\ell$ and $b$ -tagged jet	0.0003
$\Delta\phi(\ell, b)$	$\Delta\phi$ between the $\ell$ and $b$ -tagged jet	0.0003
$\Delta R(\ell, b)$	$\Delta R$ between the $\ell$ and $b$ -tagged jet	0.0001

Table 8: A list of the variables used in the object-level BDT, ordered by variable importance (descending, top to bottom) is shown.

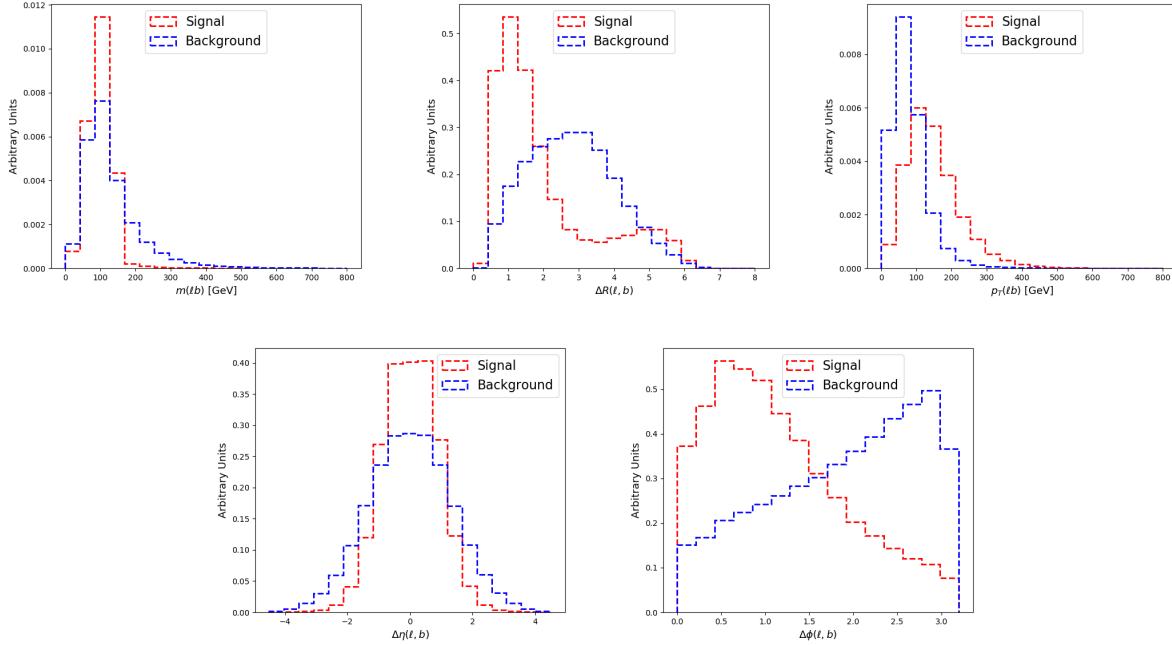


Figure 25: Normalised distributions of the variables used in the object-level BDT (ordered from top left to bottom right via decreasing variable importance), for the signal and background classes are shown. **From top left to bottom right:** Invariant mass of the  $\ell b$  system.  $\Delta R$  between the  $\ell$  and  $b$ -tagged jet. The  $p_T$  of the  $\ell b$  system.  $\Delta\eta$  between the  $\ell$  and  $b$ -tagged jet.  $\Delta\phi$  between the  $\ell$  and  $b$ -tagged jet. The red and blue dotted lined histograms represent the signal and background classes events (from the training set), respectively. These histograms are normalised to an area of 1. The variable used in training is shown on the x-axis. The y-axis shows the relative number of events for the signal and background classes (in arbitrary units).

880 the  $t\bar{t}Z$  CR are shown.

881 Overall, there is good agreement between data and simulation for the variables used in the object-level BDT, in  
882 the  $t\bar{t}Z$  CR. This suggests that the variables used in the object-level BDT are well-modelled and are reasonable to  
883 include as inputs to the object-level BDT.

884 A final check can be done to study the similarity of the  $\ell b$  systems present in the alternative  $t\bar{t}$  sample which are  
885 used for training the object-level BDT, and the  $\ell b$  systems which are aimed to be identified using the object-level  
886 BDT. More specifically, the study is done to ensure that the modelling of the  $\ell b$  systems in the alternative  $t\bar{t}$  sample  
887 are sufficiently similar to those in the  $tWZ$  and  $t\bar{t}Z$  samples (see Table 3). This is done to understand how well  
888 the BDT (trained on  $\ell b$  systems in the alternative  $t\bar{t}$  sample) generalises to classifying  $\ell b$  systems in the analysis  
889 ( $tWZ$  and  $t\bar{t}Z$  samples). In Figure 27, normalised distributions of the variables used in the object-level BDT for  
890 the alternative  $t\bar{t}$ ,  $tWZ$  and  $t\bar{t}Z$  samples, are shown.

891 The  $m(\ell b)$ ,  $p_T(\ell b)$  and  $\Delta\eta(\ell, b)$  distributions show little discrimination between the samples. However, the  $\Delta\phi(\ell, b)$   
892 and  $\Delta R(\ell, b)$  distributions show a large amount of discrimination between the three samples. This suggests that the  
893 use of the alternative  $t\bar{t}$  sample in training the object-level BDT may be sub-optimal. However, it still represents

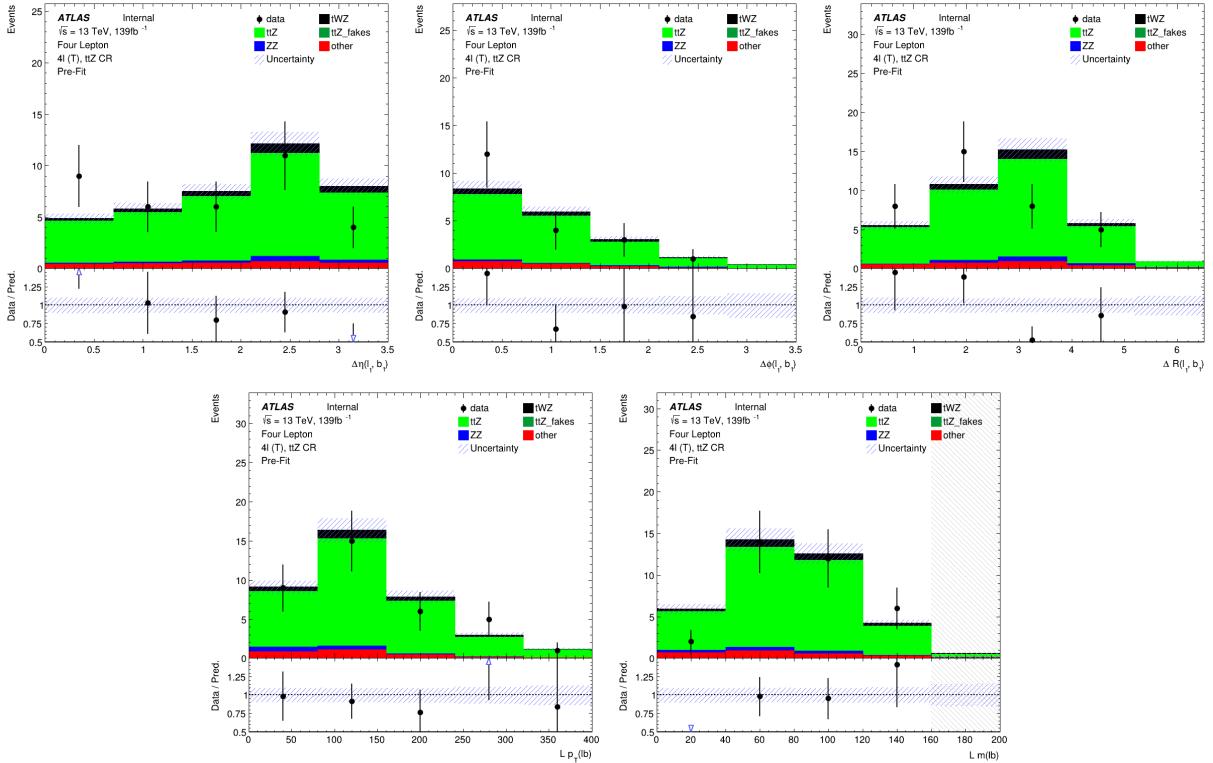


Figure 26: The expected number of events of variables used as input to the object-level BDT (ordered from top left to bottom right via decreasing variable importance), in the  $t\bar{t}Z$  CR, are shown. **From top left to bottom right:**  $\Delta\eta$  between the lepton and  $b$ -jet of the leading  $\ell b$  system.  $\Delta\phi$  between the lepton and  $b$ -jet of the leading  $\ell b$  system.  $\Delta R$  between the lepton and  $b$ -jet of the leading  $\ell b$  system.  $p_T$  of the leading  $\ell b$  system. Mass of the leading  $\ell b$  system. The data is given by the black points and the MC predictions for each process are given by the filled histograms. The vertical lines on the data points represent the statistical uncertainty in the data and the blue diagonally-lined bands represent the total (statistical and systematic added in quadrature) uncertainty. The lower panel in each plot shows the ratios of the data to the theoretical predictions. Bins with  $\frac{signal}{background} > 0.1$  are kept blinded. Blinded bins are shaded with black diagonal lines and their data points are omitted.

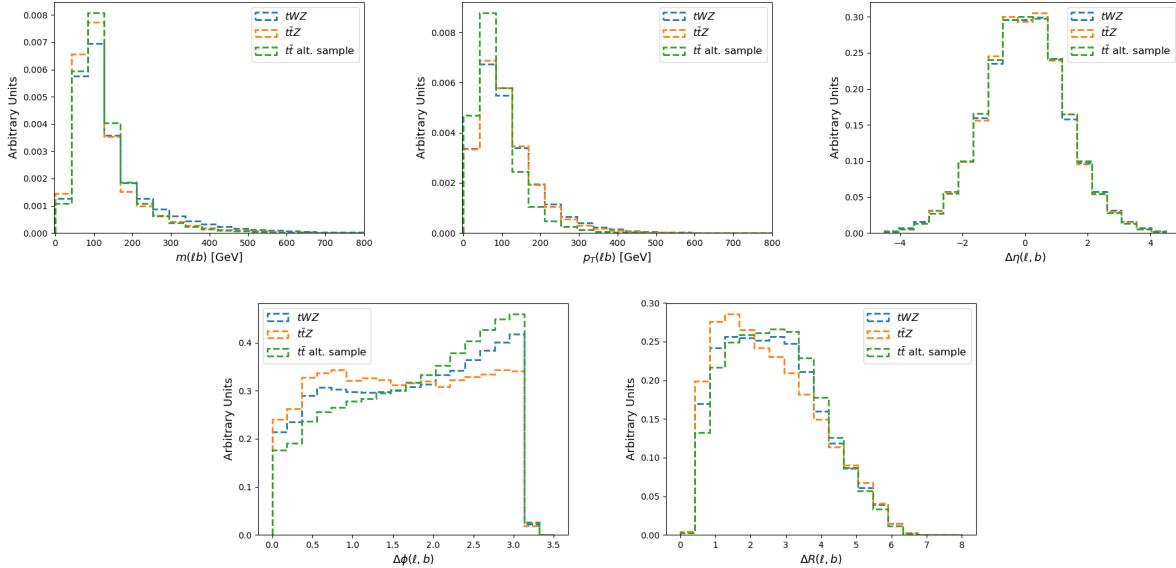


Figure 27: Normalised distributions of the variables (ordered from top left to bottom right via decreasing variable importance) used in the object-level BDT for the alternative  $t\bar{t}$ ,  $tWZ$  and  $t\bar{t}Z$  samples, are shown. **From top left to bottom right:** Invariant mass of the  $\ell\bar{\ell}$  system.  $\Delta R$  between the  $\ell$  and  $b$ -tagged jet. The  $p_T$  of the  $\ell\bar{\ell}$  system.  $\Delta\eta$  between the  $\ell$  and  $b$ -tagged jet.  $\Delta\phi$  between the  $\ell$  and  $b$ -tagged jet. The green, blue and orange dotted histograms represent events from the alternative  $t\bar{t}$ ,  $tWZ$  and  $t\bar{t}Z$  samples, respectively. These histograms are normalised to an area of 1. The variable used in training is shown on the x-axis. The y-axis shows the relative number of events (in arbitrary units).

the best option available, since our other options involve utilising of a subset of generated events used in the other parts of the analysis. This would result in a smaller number of generated events used in the background prediction, leading to larger statistical uncertainties, in an already statistically limited analysis.

898

In Table 9, the hyper-parameters used in the object-level BDT is shown.

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	200	The number of boosting stages to perform
learning_rate	0.1	The step size at each iteration during optimisation
max_depth	6	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

Table 9: A list of the hyper-parameters used in the object-level BDT is shown. Hyper-parameters not listed in this table use the default values as stated in the Scikit-learn Documentation[86].

899

The number of events used in training for the signal and background classes were 49871 and 384152 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 BDT performing sub-optimally. In order to correct this dataset imbalance, it is ensured 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. In order to avoid over-training, the BDT outputs to the training set and a test set can be studied. If over-training occurs, the BDT will fit the data in the training set too closely, resulting in the BDT outputs of the training and test sets to differ. In Figure 28 the normalised histograms of the training and test sets (extracted from fold 5 from a 5 fold kfold cross validation) for signal and background is shown. The shapes of the training and test sets for both signal and background agree within uncertainties in the vast majority of bins. This is a good indicator that no over-training occurred, since it indicates that statistical fluctuations (or noise) present in the training set was not learnt during training. Another over-training check is

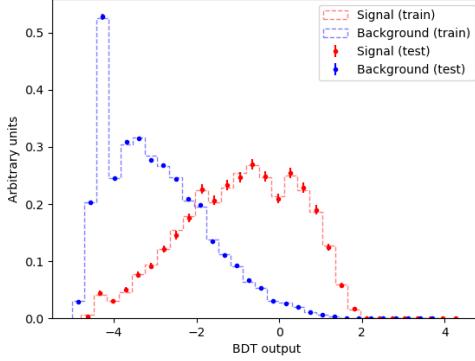


Figure 28: Normalised histograms of the object-level BDT discriminator output from the signal and background classes for the training and test sets from the 5th fold in a 5 fold kfold cross validation is shown. The output of the object-level BDT is shown on the x-axis and the relative number of events (in arbitrary units) is shown on the y-axis. The training set for the signal class is shown by the red dotted histogram. The test set for the signal class is shown by the red points, with the total uncertainty represented by the vertical error bars. The training set for the background class is shown by the blue dotted histogram. The test set for the background class is shown by the blue points, with the total uncertainty represented by the vertical error bars.

performed using 5 fold kfold cross validation. To ensure that the BDT is not over-training, it is ensured that the variance of the mean accuracy of each folds' test set in cross validation is substantially small. This tells us that the BDT does not perform better on one subset of a dataset over another and it is therefore not prone to learning statistical fluctuations of a subset, which would result in a BDT which does not generalise well to unseen datasets. For the object-level BDT, a variance of  $3.24 \times 10^{-7}$  was calculated for the mean accuracies of each folds' test set in cross validation. This small variance therefore provides further evidence that no over-training occurred.

The output of the object-level BDT is converted to an event-level variable to be used in the event-level BDT. This variable,  $\text{BDTScore}(\frac{\text{Best}}{\text{2nd Best}})$ , takes the ratio of the scores of the top scoring  $\ell b$  system to the 2nd best scoring  $\ell b$  system. The 2nd best scoring  $\ell b$  system in a  $tWZ$  event is expected to be low, since there is only one  $\ell b$  system originating from a top quark. Thus  $\text{BDTScore}(\frac{\text{Best}}{\text{2nd Best}})$  is expected to be large for  $tWZ$  events and closer to one for  $t\bar{t}Z$  events, therefore providing discrimination between  $tWZ$  and  $t\bar{t}Z$ . In Figure 29, normalised distributions of the signal and total background of the  $\text{BDTScore}(\frac{\text{Best}}{\text{2nd Best}})$  variable in the  $tWZ$  OF SR,  $tWZ$  SF SR and  $t\bar{t}Z$  CR are shown.

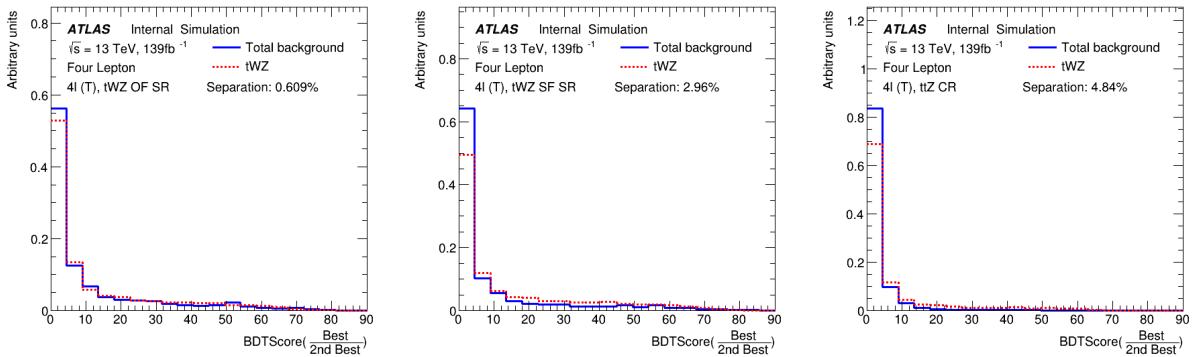


Figure 29: Normalised distributions of the signal and total background of the  $\text{BDTScore}(\frac{\text{Best}}{\text{2nd Best}})$  variable in the  $tWZ$  OF SR,  $tWZ$  SF SR and  $t\bar{t}Z$  CR are shown (left to right). The dotted red and solid blue lines represent the distributions of the signal and total background events respectively. These histograms are normalised to an area of 1. The x-axis shows the  $\text{BDTScore}(\frac{\text{Best}}{\text{2nd Best}})$  and the y-axis show the relative number of events (in arbitrary units).

The amount of discrimination can be quantified by the separation metric, which gives the percentage of the total

area of the distributions which do not overlap. A value of 1 indicates that the distributions are fully separated (no overlap) and a value of 0 indicates that the distributions have no separation (fully overlapped). The separation between signal and background for  $\text{BDTScore}(\frac{\text{Best}}{\text{2nd Best}})$  in the  $tWZ$  OF SR,  $tWZ$  SF SR and  $t\bar{t}Z$  CR are 0.609%, 2.96% and 4.84% respectively. The larger separation in the  $t\bar{t}Z$  CR, compared to the  $tWZ$  SRs, can be explained since there is a larger proportion of  $t\bar{t}Z$  events (events with two  $\ell b$  systems) in this region, due to the baseline selection requirement of exactly two  $b$ -tagged jets. In a similar way, the smaller separation in the two  $tWZ$  SRs can be explained by the tighter selection on the number of  $b$ -tagged jets (exactly one) leading to regions which are enriched in only one  $\ell b$  system which originates from a top quark. Using the  $\text{BDTScore}(\frac{\text{Best}}{\text{2nd Best}})$  variable in training in the event-level BDT (see Section 4.7.4) improves the mean accuracy of the BDT. This tells us that the event-level BDT is taking advantage of the discrimination between signal and background present in the  $\text{BDTScore}(\frac{\text{Best}}{\text{2nd Best}})$  variable.

In an attempt to optimise the performance of the object-level BDT, signal events which are pure in  $\ell b$  systems originating from top quarks are targeted for training the BDT. Similarly, background events which are pure in  $\ell b$  systems which do not originate from top quarks are targeted for training the BDT. This is done by studying the distribution of  $\Delta R$  between the reconstructed  $\ell b$  system and the truth  $\ell b$  system ( $\Delta R((lb)_{reco}, (lb)_{truth})$ ), and excluding  $\ell b$  systems from training which are moderately matched in  $\Delta R$  to their truth counterparts, leaving well matched  $\ell b$  systems being labelled as signal and badly matched  $\ell b$  systems labelled as background. The  $\Delta R$  range where  $\ell b$  systems are excluded from training is referred to as the exclusion region. In Figure 30, the distribution of  $\Delta R$  between the reconstructed  $\ell b$  system and the truth  $\ell b$  system ( $\Delta R((lb)_{reco}, (lb)_{truth})$ ) in the alternative  $t\bar{t}$  sample, along with the exclusion region, is shown.

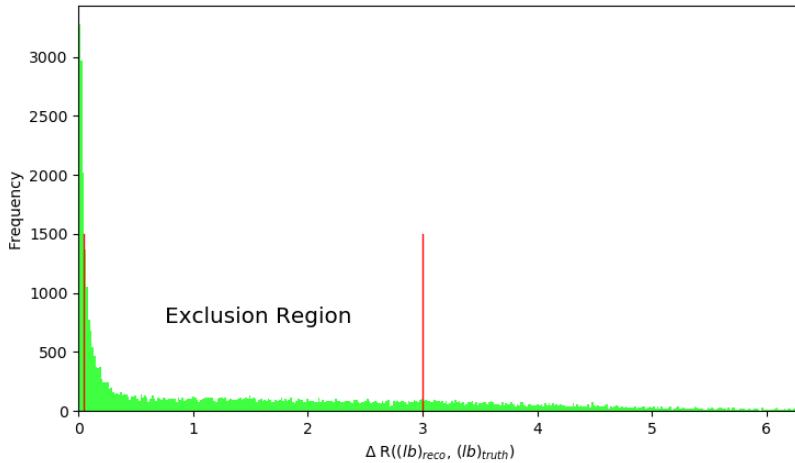


Figure 30: The distribution of  $\Delta R$  between the reconstructed  $\ell b$  system and the truth  $\ell b$  system ( $\Delta R((lb)_{reco}, (lb)_{truth})$ ) in the alternative  $t\bar{t}$  sample, along with the exclusion region, is shown. The  $\Delta R$  distribution is shown in green.  $\Delta R$  between the reconstructed  $\ell b$  system and the truth  $\ell b$  system ( $\Delta R((lb)_{reco}, (lb)_{truth})$ ) is shown on the x-axis. The bin frequency is shown on the y-axis. The exclusion region is shown between the vertical red lines situated at  $\Delta R((lb)_{reco}, (lb)_{truth}) = 0.05$  and  $\Delta R((lb)_{reco}, (lb)_{truth}) = 3.0$ . Reconstructed  $\ell b$  systems with  $\Delta R((lb)_{reco}, (lb)_{truth}) \leq 0.05$  are labelled as signal and reconstructed  $\ell b$  systems with  $\Delta R((lb)_{reco}, (lb)_{truth}) \geq 3.0$  are labelled as background. All reconstructed  $\ell b$  systems with  $0.05 < \Delta R((lb)_{reco}, (lb)_{truth}) < 3.0$  are excluded from training.

A large number of reconstructed  $\ell b$  systems have  $\Delta R((lb)_{reco}, (lb)_{truth})$  at values near 0. These are matched (in  $\Delta R$ ) extremely well to truth  $\ell b$  systems originating from top quarks. Therefore our exclusion region is defined to be between  $0.05 < \Delta R((lb)_{reco}, (lb)_{truth}) < 3.0$ , such that all reconstructed  $\ell b$  systems with  $\Delta R((lb)_{reco}, (lb)_{truth}) \leq 0.05$  are labelled as signal and reconstructed  $\ell b$  systems with  $\Delta R((lb)_{reco}, (lb)_{truth}) \geq 3.0$  are labelled as background. All reconstructed  $\ell b$  systems with  $0.05 < \Delta R((lb)_{reco}, (lb)_{truth}) < 3.0$  are excluded from training.

The performance of the object-level BDT with and without the exclusion region can be compared by studying the discrimination between signal and background events in the  $\text{BDTScore}(\frac{\text{Best}}{\text{2nd Best}})$  variable (object-level output converted to an event-level variable to be used in the event-level BDT) for both object-level BDTs. In Figure 31, normalised distributions of  $\text{BDTScore}(\frac{\text{Best}}{\text{2nd Best}})$  using the object-level BDT without the exclusion region (see Figure

956 30) for the  $tWZ$  OF SR,  $tWZ$  SF SR and  $t\bar{t}Z$  CR are shown.

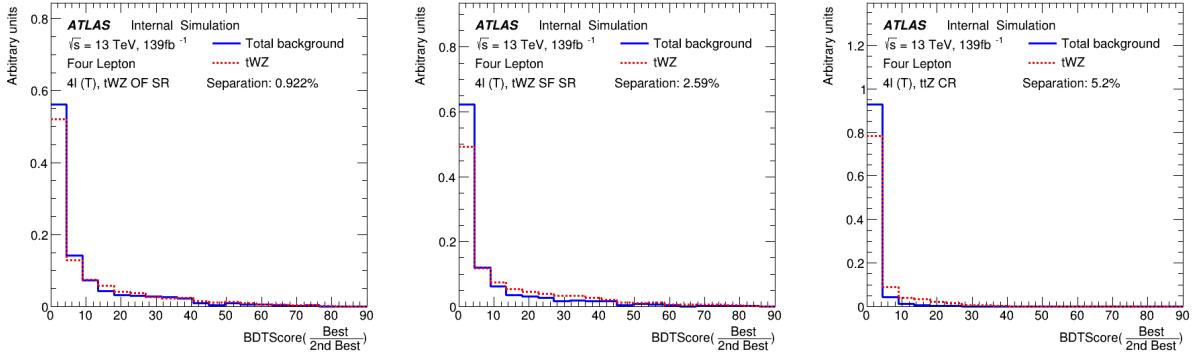


Figure 31: Normalised distributions of  $\text{BDTScore}(\frac{\text{Best}}{\text{2nd Best}})$  using the object-level BDT without the exclusion region (see Figure 30) for the  $tWZ$  OF SR,  $tWZ$  SF SR and  $t\bar{t}Z$  CR are shown (left to right). The dotted red and solid blue lines represent the distributions of the signal and total background events respectively. These histograms are normalised to an area of 1. The x-axis shows the  $\text{BDTScore}(\frac{\text{Best}}{\text{2nd Best}})$  and the y-axis show the relative number of events (in arbitrary units).

957 The separation metrics can be compared between the  $\text{BDTScore}(\frac{\text{Best}}{\text{2nd Best}})$  variable in the  $tWZ$  OF SR,  $tWZ$  SF  
 958 SR and  $t\bar{t}Z$  CR for the object-level BDT with (Figure 29) and without (Figure 31) the exclusion region by taking  
 959 the absolute difference between the two values in each region. The differences are 0.31%, 0.37% and 0.36% for the  
 960  $tWZ$  OF SR,  $tWZ$  SF SR and  $t\bar{t}Z$  CR, respectively. These differences are minimal and the object-level BDT with  
 961 the exclusion region outperforms the object-level BDT without the exclusion region in the  $tWZ$  SF SR. Due to the  
 962 small differences in performance between the two BDTs, the BDT with the exclusion region was chosen to be kept.

### 963 4.7.3 Two Neutrino Scanning Method ( $2\nu\text{SM}$ ) Algorithm

964 The difference in the number of resonant top quarks in the  $tWZ$  signal and the dominant background,  $t\bar{t}Z$ , is a key  
 965 feature which can be exploited in order to discriminate between these two processes. In Section 4.7.2, a BDT was  
 966 implemented which exploits this information by aiming to identify  $\ell b$  systems originating from top quarks. In this  
 967 section, a kinematic reconstruction algorithm (Two Neutrino Scanning Method) is implemented which exploits the  
 968 same feature.

969 The Two Neutrino Scanning Method ( $2\nu\text{SM}$ ) algorithm<sup>1</sup> [69, 68] aims to reconstruct  $t\bar{t}$  systems in the  $2\ell$ ,  $3\ell$  and  
 970  $4\ell$  final states (e.g.  $2\ell$  case:  $t\bar{t} \rightarrow \ell^+ \nu_\ell b\bar{b} - \bar{\nu}_\ell \bar{b}$ ). The  $2\nu\text{SM}$  algorithm aims to reconstruct a  $t\bar{t}$  system by finding  
 971 two neutrinos ( $\nu_1$  and  $\nu_2$ ) which are most likely to correspond to the neutrinos that originate from the decay of  
 972 a  $t\bar{t}$  system. This algorithm can be used in our analysis to discriminate between  $tWZ$  and  $t\bar{t}Z$ , since the OSSF  
 973 leptons which decay from the  $Z$  boson can be easily reconstructed and removed before inputting the event into the  
 974 algorithm. It would then be expected that the  $2\nu\text{SM}$  algorithm returns a higher score from a  $t\bar{t}Z$  event ( $\sim 1$ , i.e. it  
 975 looks like a  $t\bar{t}$  event after removal of the  $Z$  boson) and a lower score from a  $tWZ$  event ( $\sim 0$ , i.e. it does not look  
 976 like a  $t\bar{t}$  event after removal of the  $Z$  boson).

977 The first step in the  $2\nu\text{SM}$  algorithm involves stating four equations which correspond to the invariant masses of  
 978 the top quark ( $m(t)$ ) and  $W$  boson ( $m(W)$ ) for the two top quark decays (i.e.  $t \rightarrow W^+ b \rightarrow \ell^+ \nu_\ell$ ) in a dileptonic  $t\bar{t}$   
 979 event. These can be written as,

$$(\ell_1 + \nu_1)^2 = m(W)^2 = (80.385 \text{ GeV})^2 \quad (4.2)$$

$$(\ell_1 + \nu_1 + b_{1,2})^2 = m(t)^2 = (172.5 \text{ GeV})^2 \quad (4.3)$$

$$(\ell_2 + \nu_2)^2 = m(W)^2 = (80.385 \text{ GeV})^2 \quad (4.4)$$

$$(\ell_2 + \nu_2 + b_{2,1})^2 = m(t)^2 = (172.5 \text{ GeV})^2 \quad (4.5)$$

981 where the subscripts indicate that these particles originate from the decay of two different top quarks in a  $t\bar{t}$  system.  
 982 An assumption is made such that the mass of the neutrinos ( $\nu_1$  and  $\nu_2$ ) are exactly zero, which leaves us with 6

<sup>1</sup>software tool and weights provided by Thomas McCarthy ( $t\bar{t}Z$  analysis group - Max Planck Institute)

983 unknowns,  $p_{T,\nu_1}$ ,  $\phi_{\nu_1}$ ,  $\eta_{\nu_1}$ ,  $p_{T,\nu_2}$ ,  $\phi_{\nu_2}$  and  $\eta_{\nu_2}$  (components of the two neutrino's 4-vectors). The 4-vectors of the two  
 984 reconstructed leptons (not from the  $Z$  boson) and the two jets with the highest DL1r  $b$ -tagger score are used as  
 985 input to the algorithm. For each neutrino ( $\nu_1$  and  $\nu_2$ ), a scan over a range of possible  $\eta$  and  $\phi$  values is performed.  
 986 These values were chosen to be  $\phi_{\nu_1}$ ,  $\phi_{\nu_2} \in [-\pi, \pi]$  with a step size of  $\approx 0.25$  and  $\eta_{\nu_1}$ ,  $\eta_{\nu_2} \in [-5, 5]$  with a step  
 987 size of  $\approx 0.31$ . These ranges were chosen to maximize accuracy and minimize computation time. For each of these  
 988 possible  $\eta$  and  $\phi$  values, the corresponding  $p_T$  for each neutrino is calculated ( $p_{T,\nu_1}$  and  $p_{T,\nu_2}$ ) via,

$$p_{T,\nu} = \frac{\frac{1}{2}(m(W)^2 - m(\ell)^2)}{E_\ell \cosh \eta_\nu - p_{\ell,z} \sinh \eta_\nu - p_{\ell,x} \cos \phi_\nu - p_{\ell,y} \sin \phi_\nu} \quad (4.6)$$

989 where  $E_\ell$  is the energy of the lepton and  $p_{\ell,z}$ ,  $p_{\ell,x}$ ,  $p_{\ell,y}$  are the  $z$ ,  $x$  and  $y$  components of lepton's momentum.  
 990 After computing  $p_{T,\nu}$  for both neutrinos for each possible  $\eta$  and  $\phi$  combination in the defined ranges, a collection of  
 991 possible 4-vectors for  $\nu_1$  and  $\nu_2$  can be reconstructed. Using  $\nu_1$  and  $\nu_2$ , two possible  $t\bar{t}$  systems are reconstructed,

$$t_1 = \ell_1 + b_1 + \nu_1 \text{ and } t_2 = \ell_2 + b_2 + \nu_2 \quad (4.7)$$

**OR**

$$t_1 = \ell_1 + b_2 + \nu_1 \text{ and } t_2 = \ell_2 + b_1 + \nu_2 \quad (4.8)$$

992 The  $2\nu$ SM algorithm is extremely computationally intensive. The computation time depends on the number step  
 993 size of the  $\phi$  and  $\eta$  ranges which are scanned over to reconstruct the neutrinos. For example, consider the step sizes  
 994 chosen in this analysis,  $\Delta\eta \approx 0.31$  and  $\Delta\phi \approx 0.25$  which corresponds to 32 values for  $\eta$  and 25 values for  $\phi$ . There  
 995 will be  $(32)(32)(25)(25) = 640\,000$  possible pairs of neutrinos ( $\nu_1$  and  $\nu_2$ ) to consider per event. Since two possible  
 996  $t\bar{t}$  systems (See Equations 4.7 and 4.8) are considered, this number effectively increases to  $(2)(640\,000) = 128\,000$   
 997 0 iterations per event. In order to reduce the number of  $t\bar{t}$  systems needed to be considered, therefore decreasing  
 998 computation time, distributions of variables from  $t\bar{t}$  events are studied to apply a veto to a possible reconstructed  
 999  $t\bar{t}$  system if the variable in question is improbable or unlikely to be observed in a  $t\bar{t}$  event. To achieve this, an  
 1000 allowed range is defined for these variables (See Figure ?? and Figure ??), and if the possible reconstructed  $t\bar{t}$   
 1001 system's corresponding value for this variable lies outside this range, it is vetoed and the algorithm continues with  
 1002 the next iteration.

1003 The first variable which is considered, is the difference between average mass of the two possible  $\ell b$  system combi-  
 1004 nations,  $\Delta\langle m(\ell b) \rangle$ . The two possible  $\ell b$  system combinations are,

$$(\ell b)_1 = \ell_1 + b_1 \text{ and } (\ell b)_2 = \ell_2 + b_2 \quad (4.9)$$

**OR**

$$(\ell b)_1 = \ell_1 + b_2 \text{ and } (\ell b)_2 = \ell_2 + b_1 \quad (4.10)$$

$$(4.11)$$

1006  $\Delta\langle m(\ell b) \rangle$  is therefore defined as,

$$\Delta\langle m(\ell b) \rangle = \frac{1}{2} |[(m(\ell_1 b_1) + m(\ell_2 b_2)) - (m(\ell_1 b_2) + m(\ell_2 b_1))]| \quad (4.12)$$

1007 The idea here is that, in events where the average masses of the two possible  $\ell b$  system combinations differ greatly,  
 1008 the correct combination is usually given by the combination with the smaller average mass. Furthermore, it  
 1009 was shown in Ref. [69] that reconstructed top quarks in a  $t\bar{t}$  system that contain  $b$ -tagged jets in opposite hemi-  
 1010 spheres<sup>2</sup> ( $\eta(b_1) \times \eta(b_2) < 0$ ) of the ATLAS detector are easier to determine the correct  $\ell b$  system combination  
 1011 than reconstructed  $t\bar{t}$  systems that contain  $b$ -tagged jets in the same hemispheres ( $\eta(b_1) \times \eta(b_2)$ ). To illustrate this,  
 1012 the distributions (constructed from  $t\bar{t}$  events) of the probability of choosing the correct  $\ell b$  system combination,  
 1013 given that the one with the minimum  $\langle m(\ell b) \rangle$  chosen ( $P(\text{Correct combination of } \ell b \text{ systems} | \text{minimum} \langle m(\ell b) \rangle)$ )  
 1014 vs  $\Delta\langle m(\ell b) \rangle$  for  $b$ -tagged jets in the same and opposite hemispheres are investigated. In Figure 32, the  
 1015  $P(\text{Correct combination of } \ell b \text{ systems} | \text{minimum} \langle m(\ell b) \rangle)$  vs  $\Delta\langle m(\ell b) \rangle$ , for  $b$ -tagged jets in the same and opposite  
 1016 hemispheres, constructed from  $t\bar{t}$  events is shown.

1017 From Figure 32, for both cases where the  $b$ -tagged jets are in the same and opposite hemispheres, the probability  
 1018 for a correct  $\ell b$  system being chosen, given that the  $\ell b$  system with the minimum average mass is under con-  
 1019 sideration, is an increasing function which plateaus to 1 at  $\sim 90$  GeV. One of these two distributions are used

<sup>2</sup>The ATLAS detector can be split into two regions or *hemispheres*, defined where  $z > 0$  and  $z < 0$

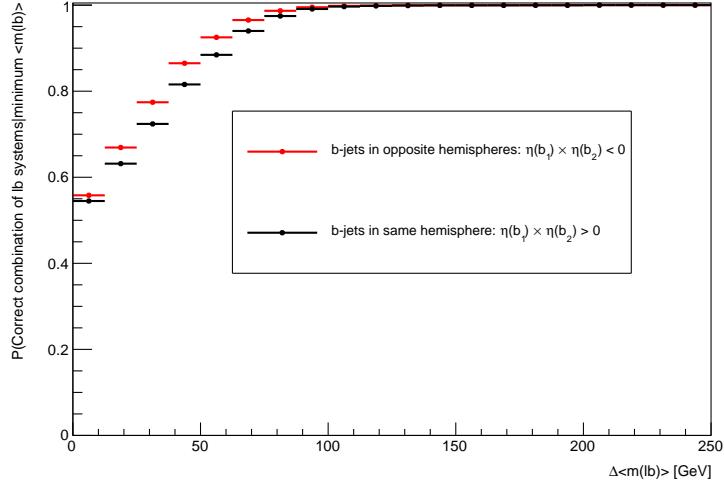


Figure 32:  $P(\text{Correct combination of } lb \text{ systems}|\text{minimum}\langle m(lb) \rangle)$  vs  $\Delta\langle m(lb) \rangle$ , for  $b$ -tagged jets in the same and opposite hemispheres, constructed from  $t\bar{t}$  events is shown. The horizontal red lines represent the distribution in the case when the two  $b$ -jets are in opposite hemispheres. The dot in the middle of the line represents the midpoint of the line. The horizontal black lines represent the distribution in the case when the two  $b$ -jets are in the same hemispheres. The dot in the middle of the line represents the midpoint of the line. The average  $m(lb)$  is shown on the x-axis. The  $P(\text{Correct combination of } lb \text{ systems}|\text{minimum}\langle m(lb) \rangle)$  is shown on the y-axis.

(depending on whether or not the two  $b$ -tagged jets are in the same or opposite hemispheres) to interpolate the  $P(\text{Correct combination of } lb \text{ systems}|\text{minimum}\langle m(lb) \rangle)$  from  $\Delta\langle m(lb) \rangle$ . A veto is applied to the  $lb$  combination with the maximum  $\Delta\langle m(lb) \rangle$  if  $P(\text{Correct combination of } lb \text{ systems}|\text{minimum}\langle m(lb) \rangle) > 0.8$ , indicating that there is at least an 80% certainty that the  $lb$  combination with the minimum  $\langle m(lb) \rangle$  is the correct combination. If  $P(\text{Correct combination of } lb \text{ systems}|\text{minimum}\langle m(lb) \rangle) < 0.8$ , both possible  $lb$  system combinations need to be considered.

The  $\eta$  of the  $b\bar{b}\ell\ell$  system,  $\eta(b\bar{b}\ell\ell)$ , to veto improbable  $\eta(\nu_1)$  and  $\eta(\nu_2)$  values is then considered. In the same way as for  $\Delta\langle m(lb) \rangle$ , a distribution is generated to determine values  $\eta(\nu)$  which are improbable for a  $t\bar{t}$  event. In this case, a 2D histogram from simulated  $t\bar{t}$  events (dileptonic final state) at generator-level of  $\eta(\nu)$  vs  $\eta(b\bar{b}\ell\ell)$  is generated. Using this histogram, a veto region (where a  $t\bar{t}$  event is extremely unlikely to occur) is defined which contains 95% of events. A veto is applied if either possible neutrino lies within this region. In Figure 33, a heatmap of occupancy for  $\eta(\nu)$  vs  $\eta(b\bar{b}\ell\ell)$  (produced from simulated  $t\bar{t}$  events) and its corresponding veto region are shown. The final kinematic constraint which is considered is the scalar sum of lepton  $p_T$ ,  $L_T = p_T(\ell_1) + p_T(\ell_2)$ , which is used to veto certain possible neutrinos,  $\nu_1$  and  $\nu_2$ . Again, a distribution is generated to determine (and veto) improbable possible neutrinos in simulated  $t\bar{t}$  events (dilepton final state). Using the same method as described for Figure 33, a veto region is defined where a veto is applied if either possible neutrino lies within this region. In Figure 34, a heatmap of occupancy for  $\Delta R(\ell, \nu)$  vs  $L_T$  (produced from simulated  $t\bar{t}$  events) and its corresponding veto region are shown.

In order to choose the solution which best represents the two top quarks in a  $t\bar{t}$  system, the likelihood of each solution is evaluated in the SM  $t\bar{t}$  hypothesis. This is performed using the product of probabilities derived from certain distributions of variables from simulated  $t\bar{t}$  events. The events in these distributions are obtained from an ATLAS simulation of generated  $t\bar{t}$  events in the dileptonic final state.

A normalised distribution of the mass of reconstructed top quarks,  $m_{b\ell\nu}$ , from a  $t\bar{t}$  sample is generated to determine the probabilities  $P_{m_{t_1}}$  and  $P_{m_{t_2}}$  which correspond to the likelihood of the reconstructed top quarks under the SM  $t\bar{t}$  hypothesis. The distribution is generated from reco-level leptons, generator-level neutrinos and reco-level jets matched in  $\Delta R$  to generator-level  $b$ -quarks, therefore only filling the distribution with correct detector-level objects. For both reconstructed top quarks,  $m(b\ell\nu)$  is calculated and interpolated<sup>3</sup> (i.e. estimate the value of the distribution for some value of the independent variable) against the  $m_{b\ell\nu}$  distribution which returns a weight

<sup>3</sup>performed via linear interpolation based on the two nearest bin centres [ROOTTH1C87:online]

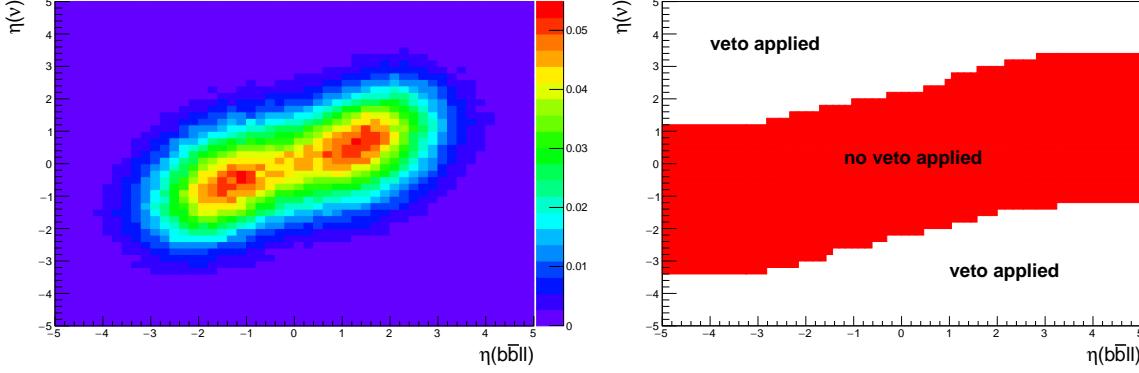


Figure 33: **Right:** Heatmap of occupancy for  $\eta(\nu)$  vs  $\eta(b\bar{b}\ell\ell)$  produced from simulated  $t\bar{t}$  events (dileptonic final state) at generator-level is shown.  $\eta$  of the  $b\bar{b}\ell\ell$  system is shown on the x-axis.  $\eta$  of the neutrino is shown on the y-axis. The colorbar on the right represents the fraction of events in the phase space. **Left:** The regions where vetoes are applied for the  $\eta(b_1 b_2 \ell_1 \ell_2)$  constraint is shown.  $\eta$  of the  $b\bar{b}\ell\ell$  system is shown on the x-axis.  $\eta$  of the neutrino is shown on the y-axis. The red band shows the region where the neutrino would not be vetoed. The white areas (above and below the red band) are regions where the neutrino is vetoed.

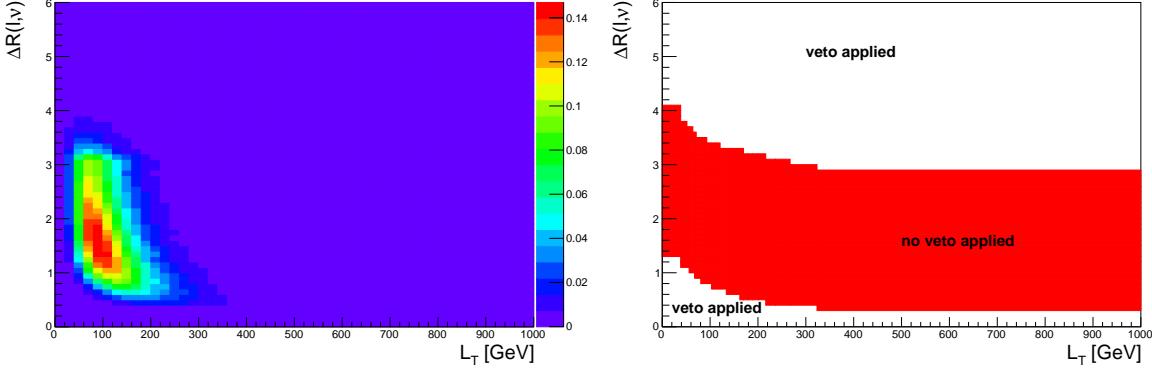


Figure 34: **Left:** A heatmap of occupancy for  $\Delta R(\ell, \nu)$  vs  $L_T$  produced from simulated  $t\bar{t}$  events (dileptonic final state) at generator-level is shown.  $\Delta R$  between leptons and neutrinos is shown on the x-axis.  $L_T$  (scalar sum of lepton  $p_T$ ) is shown on the y-axis. The colourbar on the right represents the fraction of events in the phase space. **Right:** The regions where vetoes are applied for the  $L_T$  constraint is shown.  $\Delta R$  between leptons and neutrinos is shown on the x-axis.  $L_T$  (scalar sum of lepton  $p_T$ ) is shown on the y-axis. The red band shows the region where the neutrino would not be vetoed. The white areas (above and below the red band) are regions where the neutrino is vetoed.

value from 0 to 1, with higher values corresponding to a reconstructed top quark which has a mass close to that of a top quark from a  $t\bar{t}$  system. This interpolation is done for both reconstructed top quarks,  $t_1$  and  $t_2$ , corresponding to probabilities  $P_{m_{t_1}}$  and  $P_{m_{t_2}}$ . A similar method is used to determine  $P_{\Delta E_x}$  and  $P_{\Delta E_y}$ , which corresponds to the likelihood of the reconstructed top quarks under the SM  $t\bar{t}$  hypothesis. In this case, a weight distribution of  $\Delta E_x = (p_{T,\nu_1})_x + (p_{T,\nu_2})_x - (E_T^{\text{miss}})_x$  based off simulated  $t\bar{t}$  events is generated. In particular, this distribution is generated using reco-level  $E_T^{\text{miss}}$  and generator-level neutrinos. The use of this distribution lies under the assumption that neutrinos are the dominant source of  $E_T^{\text{miss}}$ , and therefore,  $(E_T^{\text{miss}})_x \approx (p_{T,\nu_1})_x + (p_{T,\nu_2})_x$  and  $(E_T^{\text{miss}})_y \approx (p_{T,\nu_1})_y + (p_{T,\nu_2})_y$ . This distribution is then used to interpolate the value of  $\Delta E_x$  and  $\Delta E_y$  from our reconstructed neutrinos. This returns a weight value from 0 to 1, with higher values corresponding to  $\Delta E_x$  and  $\Delta E_y$  (and in turn our reconstructed neutrino's  $p_T$ ) closer to those observed in a  $t\bar{t}$  event. It is expected that the  $\Delta E_x$  and  $\Delta E_y$  distributions have the same shapes, therefore only one is needed to be generated. In this case the the  $\Delta E_x$  distribution was chosen. In Figure 35, the  $m_{b\ell\nu}$  and  $\Delta E_x$  distributions (generated from simulated  $t\bar{t}$  events) are shown.

A final weight,  $w_{2\nu SM} \in [0, 1]$ , is then calculated by combining the four probabilities ( $P_{m_{t_1}}$ ,  $P_{m_{t_2}}$ ,  $P_{\Delta E_x}$  and  $P_{\Delta E_y}$ )

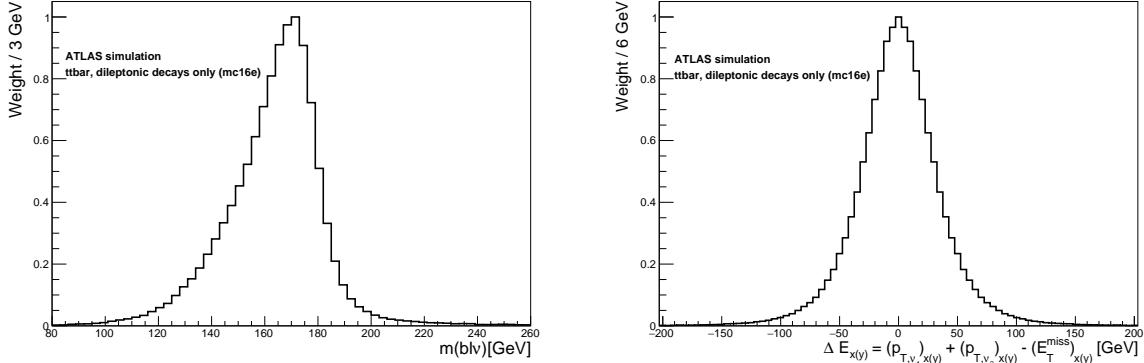


Figure 35: **Left:**  $m_{b\ell\nu}$  distribution generated from simulated  $t\bar{t}$  events, used to calculate  $P_{m_{t_1}}$  and  $P_{m_{t_2}}$  is shown. The  $m_{b\ell\nu}$  distribution is shown by the black lined histogram. The mass of the  $b\ell\nu$  system is shown on the x-axis. The corresponding weight of the  $m_{b\ell\nu}$  distribution is shown on the y-axis. **Right:**  $\Delta E_x$  distribution generated from simulated  $t\bar{t}$  events, used to calculate  $P_{\Delta E_x}$  and  $P_{\Delta E_y}$  is shown. The  $\Delta E_x$  distribution is shown by the black lined histogram.  $\Delta E_x$  is shown on the x-axis. The corresponding weight of  $\Delta E_x$  distribution is shown on the y-axis.

described above. This final weight represents a total probability of the reconstructed top quarks under the SM  $t\bar{t}$  hypothesis, and can be written as,

$$w_{2\nu SM} = P_{m_{t_1}} \times P_{m_{t_2}} \times P_{\Delta E_x} \times P_{\Delta E_y} \quad (4.13)$$

The  $w_{2\nu SM}$  is calculated for each pair of reconstructed neutrinos (or reconstructed  $t\bar{t}$  systems), with the maximum value being chosen as the final value for the event.

#### 4.7.4 Event-level BDT

The event-level BDT is used to distinguish between signal and its major background events,  $t\bar{t}Z$  and  $ZZ$ . The key difference between the object-level BDT and the event-level BDT is that while the former exploits information associated with  $\ell b$  systems and thus distinguishes between  $\ell b$  systems, the event-level BDT exploits information based on the entirety of the event and thus distinguishes between events.

The event-level BDT was trained on 50% of the  $tWZ$  MC sample's events for the signal class and similarly, 50% of the  $t\bar{t}Z$  and  $ZZ$  MC sample's events were used for the background class. The variables used to train the BDT are chosen on the basis that they are somewhat uncorrelated from one another and show some discrimination between  $tWZ$  and  $t\bar{t}Z$ . Similarly to the object-level BDT, the optimum values for the hyper-parameters used were determined via a grid-search (See Section 4.7.1) that determined the set of hyper-parameters which maximized the mean accuracy (based off 5 fold kfold cross-validation). After hyper-parameter optimisation, the mean accuracy of each fold (determined from 5 fold kfold cross validation) increased from 0.72 to 0.74 ( $\sim 3\%$  increase). The variable importance of each variable was computed in the same way as described for the object-level BDT (See Section 4.7.2). In Table 10, the variables used in training the event-level BDT are shown.

Variable	Description	Variable Importance
$2\nu SM$	Maximum weight from the $2\nu SM$ algorithm (See Section 4.7.3)	0.029
$HT$	Scalar sum of jet $p_T$	0.016
$LT$	Scalar sum of lepton $p_T$	0.011
$\sum p_T(b-jet)$	Scalar sum of $b$ -tagged jet $p_T$	0.006
$BDTScore(\frac{\text{Best}}{\text{2nd Best}})$	Ratio of the top scoring $\ell b$ system to the 2nd best scoring $\ell b$ system from the output of the object-level BDT (object-level BDT)	0.006
$\Delta\eta(\ell_{1,\text{non-Z}}, \ell_{2,\text{non-Z}})$	$\Delta\eta$ between the two leptons, not coming from a $Z$ candidate	0.005

Table 10: A list of the variables used in the event-level BDT, ordered by variable importance (descending, top to bottom) is shown.

In Figure 36, normalised distributions of the variables used in the event-level BDT, for the signal and background classes are shown. The variables used in the event-level BDT show a clear distinction between signal and background events. In particular the output weight from the  $2\nu SM$  algorithm shows the most discrimination. When determining which

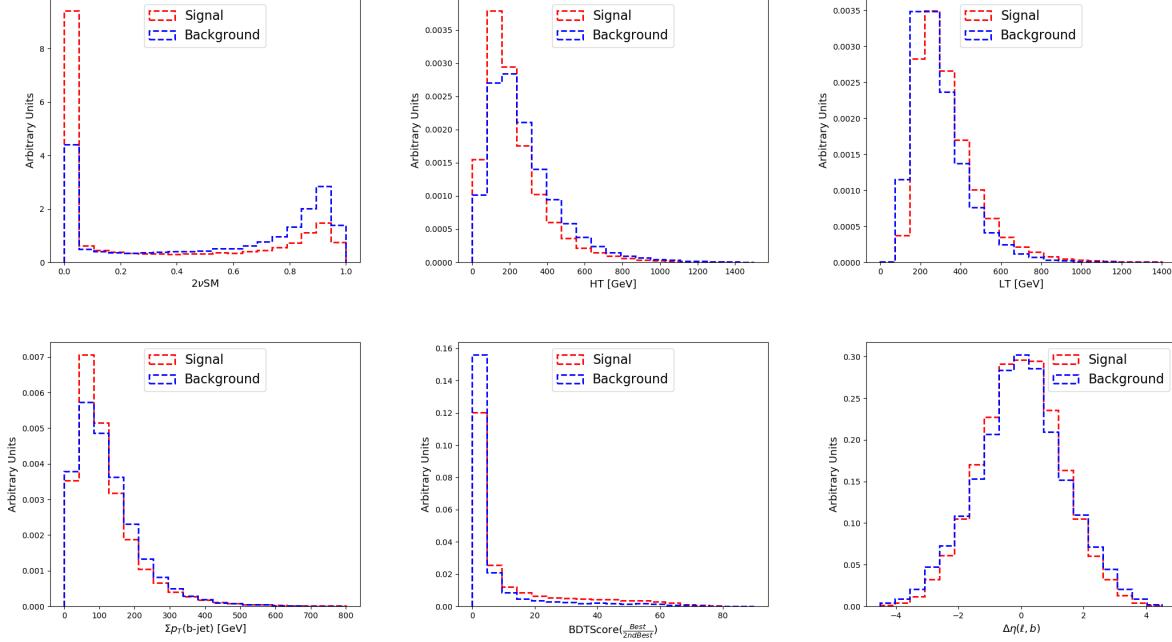


Figure 36: Normalised distributions of the variables used in the event-level BDT (ordered from top left to bottom right via decreasing variable importance), for the signal and background classes are shown. **From top left to bottom right:** Output weight from the  $2\nu\text{SM}$  algorithm (See Section 4.7.3). Scalar sum of jet  $p_T$ . Scalar sum of lepton  $p_T$ . Sum of  $b$ -tagged jet  $p_T$ . Ratio of the top scoring  $\ell b$  system to the 2nd best scoring  $\ell b$  system from the output of the object-level BDT (See Section 4.7.2).  $\Delta\eta$  between the two leptons, not coming from a  $Z$  candidate. The red and blue dotted lined histograms represent the signal and background classes events, respectively. These histograms are normalised to an area of 1. The variable used in training is shown on the x-axis. The y-axis shows the relative number of events for the signal and background classes (in arbitrary units).

variables to use in training the event-level BDT, the output weight from  $2\nu\text{SM}$  was shown to provide the most sizeable boost in performance of the BDT. Surprisingly, the least important variable,  $\Delta\phi$  between the non- $Z$  lepton system (leptons not originating from a  $Z$ -candidate) and the leading  $b$ -tagged jet, seem to discriminate well between signal and background. A possible explanation for its low ranking variable importance is due to it being relatively highly correlated with many of the other variables used in the BDT.

The modelling of the variables used in the event-level BDT can be checked by referring to the expected number of events of data and simulation in control regions where they are defined. Note that certain variables which are ill-defined in certain regions (e.g.  $\Delta\eta(\ell_{1,\text{non-}Z}, \ell_{2,\text{non-}Z})$ ) in the  $ZZb$  CR, as all leptons originate from a  $b$ -jet in this region) will not be shown. In Figure 37, MC predictions for the variables used in the event-level BDT in the  $t\bar{t}Z$  CR are shown. The deviations between data and simulation, across all variables used in the  $t\bar{t}Z$  CR, in all but three bins are within expected uncertainties. In Figure 38, MC predictions for the variables used in the event-level BDT in the  $ZZb$  CR are shown. The deviations between data and simulation, across all variables used in the  $ZZb$  CR, are within expected uncertainties. In Figure 39, MC predictions for the variables used in the event-level BDT in the  $(tWZ)_{\text{fake}}$  CR are shown.

The deviations between data and simulation, across all variables used in the  $(tWZ)_{\text{fake}}$  CR, in all but one bin are within expected uncertainties. Overall, the vast majority of predictions between data and simulation in the bins of the event-level BDT distributions in the  $t\bar{t}Z$  CR,  $ZZb$  CR and  $(tWZ)_{\text{fake}}$  CR, are within the expected uncertainties. Therefore, these variables are well-modelled and reasonable to include as inputs to the event-level BDT.

In Table 11, the hyper-parameters used in the event-level BDT are shown. Since the training is performed on  $t\bar{t}Z$  and  $ZZ$  events for the background class, it is ensured that the relative weighting of these events are such that they mimic the amount of  $t\bar{t}Z$  and  $ZZ$  expected to be present in the regions

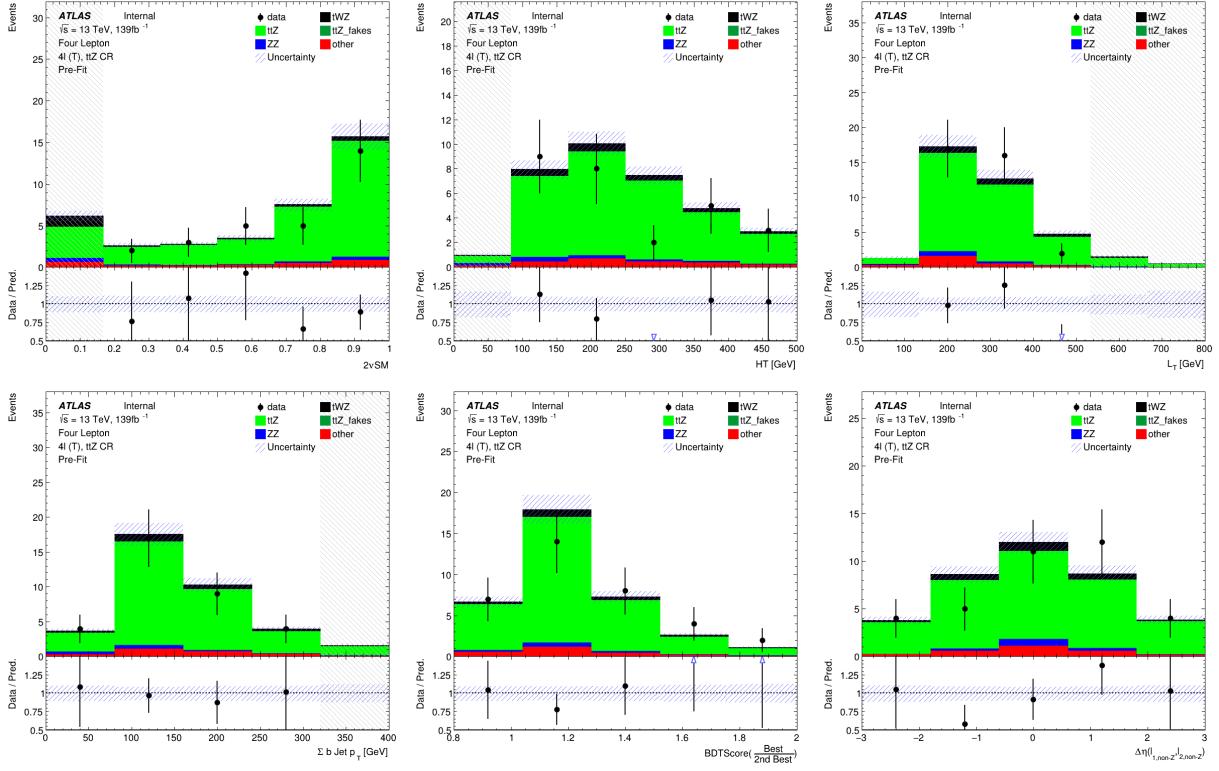


Figure 37: The expected number of events of variables used as input to the event-level BDT (ordered from top left to bottom right via decreasing variable importance), in the  $t\bar{t}Z$  CR, are shown. **From top left to bottom right:** Output weight from the  $2\nu\text{SM}$  algorithm (See Section 4.7.3). Scalar sum of jet  $p_T$ . Scalar sum of lepton  $p_T$ . Sum of  $b$ -tagged jet  $p_T$ . Ratio of the top scoring  $\ell b$  system to the 2nd best scoring  $\ell b$  system from the output of the object-level BDT (See Section 4.7.2).  $\Delta\eta$  between the two leptons, not coming from a  $Z$  candidate. The data is given by the black points and the MC predictions for each process are given by the filled histograms. The vertical lines on the data points represent the statistical uncertainty in the data and the blue diagonally-lined bands represent the total (statistical and systematic added in quadrature) uncertainty. The lower panel in each plot shows the ratios of the data to the theoretical predictions. Bins with  $\frac{\text{signal}}{\text{background}} > 0.1$  are kept blinded. Blinded bins are shaded with black diagonal lines and their data points are omitted.

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	200	The number of boosting stages to perform
learning_rate	0.1	The step size at each iteration during optimisation
max_depth	6	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

Table 11: A list of the hyper-parameters used in the event-level BDT is shown. Hyper-parameters not listed in this table use the default values as stated in the Scikit-learn Documentation[86].

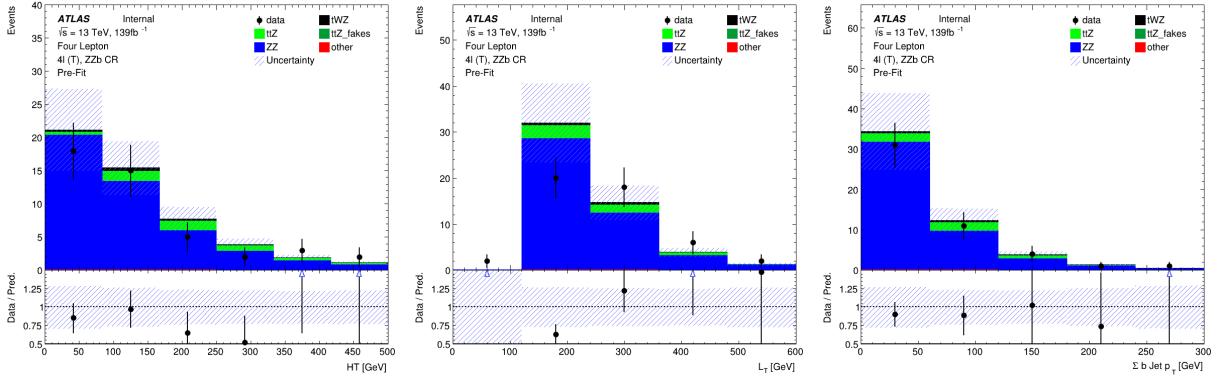


Figure 38: The expected number of events of variables used as input to the event-level BDT (ordered from top left to bottom right via decreasing variable importance), in the  $Z Z b$  CR, are shown. **From left to right:** Scalar sum of jet  $p_T$ . Scalar sum of lepton  $p_T$ . Sum of  $b$ -tagged jet  $p_T$ . The data is given by the black points and the MC predictions for each process are given by the filled histograms. The vertical lines on the data points represent the statistical uncertainty in the data and the blue diagonally-lined bands represent the total (statistical and systematic added in quadrature) uncertainty. The lower panel in each plot shows the ratios of the data to the theoretical predictions. Bins with  $\frac{\text{signal}}{\text{background}}$  greater than 0.1 are kept blinded. Blinded bins are shaded with black diagonal lines and their data points are omitted.

1111 where the BDT discriminator ( $tWZ$  SRs and  $t\bar{Z}$  CR) is used. This is done by applying normalization weights to  
 1112 each event, defined as,

$$W = \frac{\sigma \mathcal{L} \text{weight(MC)}}{\text{totalWeight(MC)}} \quad (4.14)$$

1113 where  $\sigma$  is the cross section of the process,  $\mathcal{L}$  is the integrated luminosity, weight(MC) is the weight assigned to  
 1114 the event by the MC generator and totalWeight(MC) is the sum of those weights for all the generated events. The  
 1115 number of events used in training for the signal and background classes were 41066 and 22608 respectively. Similarly  
 1116 to the object-level BDT, there is a dataset imbalance. This imbalance is corrected for (in the same way as before  
 1117 with the object-level BDT) by ensuring that the relative weighting of each event is such that the sum of the signal  
 1118 weights is equal to the sum of the background weights. In order to avoid over-training, in the same way that was  
 1119 performed for the object-level BDT (See Section 4.7.2), the BDT outputs to the training set and a test set can be  
 1120 studied. In Figure 40 the normalised histograms of the training and test sets (extracted from fold 5 from a 5 fold  
 1121 kfold validation) for signal and background is shown.

1122 The shapes of the training and test sets for both signal and background agree within uncertainties in the vast  
 1123 majority of bins. This is a good indicator that no over-training occurred, since it indicates that statistical fluctuations  
 1124 (or noise) present in the training set was not learnt during training. As with the object-level BDT, another over-  
 1125 training check is performed, by ensuring that the variance of the mean accuracy of each folds' test set in a 5 fold  
 1126 kfold cross validation is sufficiently small. This indicates that fluctuations in features from different training sets  
 1127 are not learnt by the BDT. For the event-level BDT, a variance of 0.00026 was calculated for the mean accuracies  
 1128 of each folds' test set in cross validation, providing further evidence that no over-training occurred. In Figure 41,  
 1129 normalised distributions of the signal and total background of the event-level BDT discriminator output in the  
 1130  $tWZ$  OF SR,  $tWZ$  SF SR and  $t\bar{Z}$  CR, are shown.

1131 The event-level BDT discriminates well between signal and background events in the  $tWZ$  OF SR,  $tWZ$  SF SR  
 1132 and  $t\bar{Z}$  CR, with separations of 8.98%, 10.6% and 20.6%, respectively.

## 1133 4.8 Systematic Uncertainties

1134 The final results are obtained with a maximum likelihood fit (described in Section 4.9.1) in which normalisations and  
 1135 shapes of distributions are allowed to vary to account for uncertainties of both experimental and theoretical origin.  
 1136 Systematic uncertainties are allowed to vary within a pre-determined envelope (consisting of two MC templates)  
 1137 which defines the bounds within which the systematic can vary. In this section, the methods used to determine the  
 1138 variations that form the inputs to the fit are described.

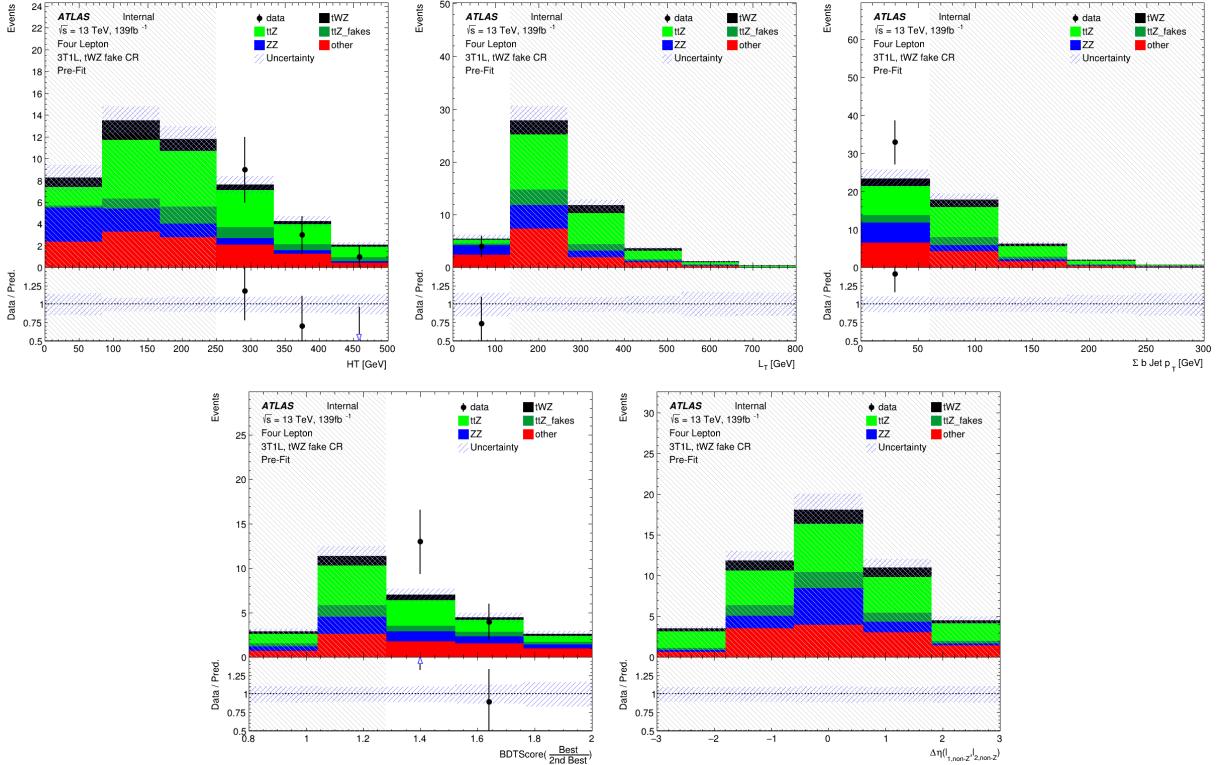


Figure 39: The expected number of events of variables used as input to the event-level BDT (ordered from top left to bottom right via decreasing variable importance), in the ( $tWZ$ )<sub>fake</sub> CR, are shown. **From top left to bottom right:** Output weight from the  $t\nu$ SM algorithm (See Section 4.7.3). Scalar sum of jet  $p_T$ . Scalar sum of lepton  $p_T$ . Sum of  $b$ -tagged jet  $p_T$ . Ratio of the top scoring  $\ell b$  system to the 2nd best scoring  $\ell b$  system from the output of the object-level BDT (See Section 4.7.2).  $\Delta\eta$  between the two leptons, not coming from a  $Z$  candidate. The data is given by the black points and the MC predictions for each process are given by the filled histograms. The vertical lines on the data points represent the statistical uncertainty in the data and the blue diagonally-lined bands represent the total (statistical and systematic added in quadrature) uncertainty. The lower panel in each plot shows the ratios of the data to the theoretical predictions. Bins with  $\frac{\text{signal}}{\text{background}}$  greater than 0.1 are kept blinded. Blinded bins are shaded with black diagonal lines and their data points are omitted.

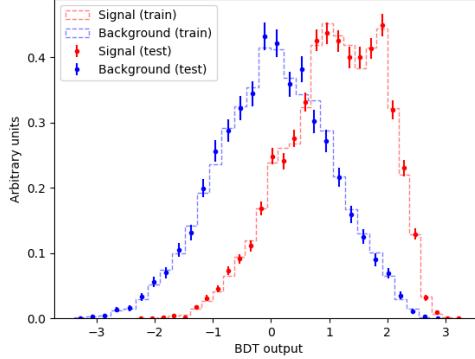


Figure 40: Normalised histograms of the event-level BDT discriminator output from the signal and background classes for the training and test sets from the 5th fold in a 5 fold kfold cross validation are shown. The output of the event-level BDT is shown on the x-axis and the relative number of events (normalised to have an area of 1, in arbitrary units) is shown on the y-axis. The training set for the signal class is shown by the red dotted histogram. The test set for the signal class is shown by the red points, with the total uncertainty represented by the vertical error bars. The training set for the background class is shown by the blue dotted histogram. The test set for the background class is shown by the blue points, with the total uncertainty represented by the vertical error bars.

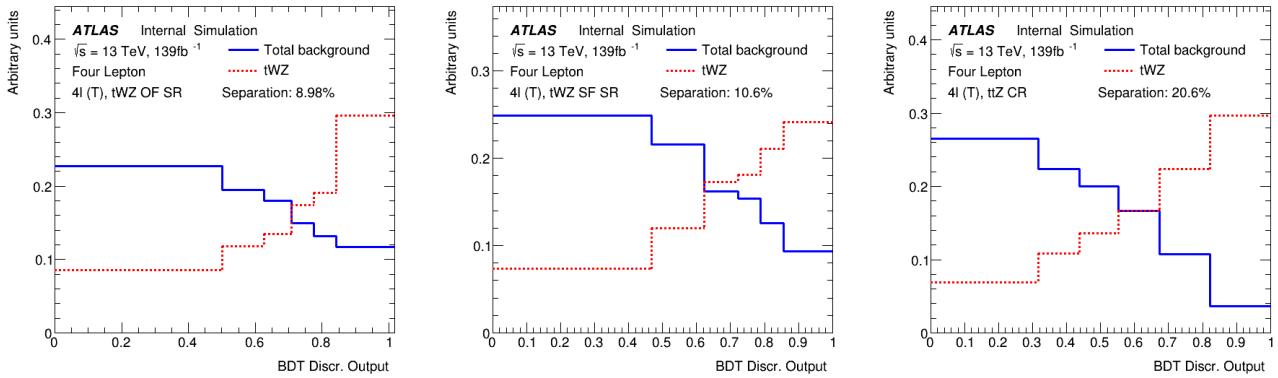


Figure 41: Normalised distributions of the signal and total background of the event-level BDT discriminator output in the  $tWZ$  OF SR,  $tWZ$  SF SR and  $t\bar{t}Z$  CR are shown (left to right). The dotted red and solid blue lines represent the distributions of the signal and total background events respectively. These histograms are normalised to an area of 1. The x-axis shows the event-level BDT discriminator output and the y-axis shows the relative number of events (in arbitrary units).

### 1139 4.8.1 Experimental uncertainties

1140 In this section, the experimental systematics are outlined.

- 1141 • **Luminosity:**

1142 The 2015-2018 luminosity estimate of  $139\text{fb}^{-1}$  has a relative uncertainty of 3%. The uncertainty associated  
1143 with the luminosity is obtained using the LUCID-2 detector [24], which is the primary luminosity monitor for  
1144 ATLAS. This systematic uncertainty affects all processes modelled using MC simulations.

- 1145 • **Pile-up modelling:**

1146 An uncertainty related to the SFs used for MC to account for differences in pile-up distributions between  
1147 MC and data is applied. This uncertainty is obtained by re-scaling the  $\langle\mu\rangle$  value in data by 1.00 and 1/1.18  
1148 corrections are only applied to MC.

- 1149 • **Jet vertex tagger:**

1150 Uncertainties associated to the *JVT* are applied via the `JetJvtEfficiency` package [22] which account for  
 1151 the residual contamination from pile-up jets after pile-up suppression and the MC generator choice [15].

- 1152 • **Heavy- and light-flavor tagging:**

1153 An additional uncertainty is assigned to account for the extrapolation of the *b*-tagging efficiency measurement  
 1154 from the  $p_T$  region used to determine the correction factors to regions with higher  $p_T$ .

- 1155 • **Electron efficiency:**

1156 Uncertainties associated with the electron efficiency SFs arise from the reconstruction, ID, isolation and trigger  
 1157 efficiencies [20]. They correct for the efficiency difference between data and MC [14] and are measured with  
 1158 a "tag-and-probe" method in  $Z \rightarrow e^+e^-$  and  $J/\psi \rightarrow e^+e^-$  events. The information on the correlation of  
 1159 the different components of the systematic uncertainties are provided for all efficiency measurements. The  
 1160 default correlation model for the uncertainties is used, which provides one up/down variation for each of the  
 1161 SF components separately [20, 21].

- 1162 • **Muon efficiency:**

1163 As for electrons, SFs obtained from  $Z \rightarrow \mu^+\mu^-$  and  $J/\psi \rightarrow \mu^+\mu^-$  events are applied to correct for the  
 1164 differences between data and MC in the muon ID, isolation and trigger efficiencies [18]. Uncertainties on these  
 1165 SFs are provided by the muon CP group [23] and applied as up/down variations of the nominal SFs for each  
 1166 component.

## 1167 4.8.2 Theoretical uncertainties

1168 In this section, the theoretical systematics are outlined.

- 1169 •  **$t\bar{t}Z$  background:**

1170 An overall normalization uncertainty of 10% is considered for the  $t\bar{t}Z$  background. Two generic shape sys-  
 1171 tematics are considered for the  $t\bar{t}Z$  background. They are constructed (see Section 4.8.3) by either applying a  
 1172 linear or triangular interpolation to up and down variations which are defined to be  $\pm 20\%$  from the nominal  
 1173  $t\bar{t}Z$  background. These linear and triangular variations are applied to distributions used in the maximum  
 1174 likelihood fit (described in Section 4.9.1).

- 1175 •  **$ZZ$  background:**

1176 An overall normalization uncertainty of 30% is considered for the  $ZZ$  background.

- 1177 •  **$t\bar{t}H$  background:**

1178 An overall normalization uncertainty of 20% is considered for the  $t\bar{t}H$  background.

- 1179 •  **$tZq$  background:**

1180 An overall normalization uncertainty of 14% is considered for the  $tZq$  background.

- 1181 •  **$t\bar{t}Z$  fake background:**

1182 An overall normalization uncertainty of 50% is considered for the  $t\bar{t}Z$  fake background.

- 1183 • **other background processes:**

1184 The 'other' background consists of many processes which have minimal but non-negligible contribution in  
 1185 the signal regions (See Table 6). An overall normalization uncertainty of 30% is considered for the 'other'  
 1186 background processes.

- 1187 •  **$tWZ$ :**

1188 A modelling uncertainty on  $tWZ$  is considered by comparing the nominal sample (using the DR1 scheme)  
 1189 and a DR2 sample.

1190 Two generic shape systematics are considered for the  $tWZ$  background. They are constructed (see Section  
 1191 4.8.3) by either applying a linear or triangular interpolation to up and down variations which are defined to  
 1192 be  $\pm 20\%$  from the nominal  $tWZ$  background.

### 4.8.3 Generic shape systematics

The quantity which is measured from the maximum likelihood fit, the expected signal strength of  $tWZ$  ( $\mu(tWZ)$ ), is defined as the ratio of the expected cross section of  $tWZ$ ,  $\sigma(tWZ)$ , to the SM cross section of  $tWZ$ ,  $\sigma_{SM}(tWZ)$  (described in more detail in Section 4.9.1). It is evident that the tetralepton channel is statistically limited, as seen by the small number of expected events in the control plots shown previously in Section 4.5. Therefore it is expected that the uncertainty on  $\mu(tWZ)$  is dominated by statistical uncertainty and that the impact of shape systematics will be negligible in comparison. In this analysis, all the sources of systematic uncertainties have not been precisely estimated, as this involves time consuming analysis of extra event samples, especially those related to the modelling of background processes. However, as described above, these missing systematic uncertainties are likely to have negligible impact. To cross check this assumption, we generate generic shape systematics as a proxy for the missing shape systematics for the modelling of processes. If these generic shape systematics have negligible impact on the measurement of the expected  $\mu(tWZ)$ , it would give us confidence that the missing systematics will not greatly alter our results when they are eventually added (this is outside the scope of this thesis).

An envelope consisting of two MC templates is constructed that represents the bounds in which the systematic can vary in the fit. One with the nominal MC template increased by 20% on its normalisation and the other with the nominal MC template decreased by 20% on its normalisation. The templates are then modified from their original shape either by doing linear interpolation (from the leftmost-up variation to the rightmost-down variation) or triangular interpolation (shape is set to zero at the higher and lower parts and reaches the envelope in the middle). The linear and triangular interpolation is done using TRF's ForceShape option [81], which alters the original templates (as described above). In Figure 42 the envelope before and after the shape change, for both the linear and triangular interpolations, for the  $t\bar{t}Z$  background in the  $t\bar{t}Z$  CR is shown.

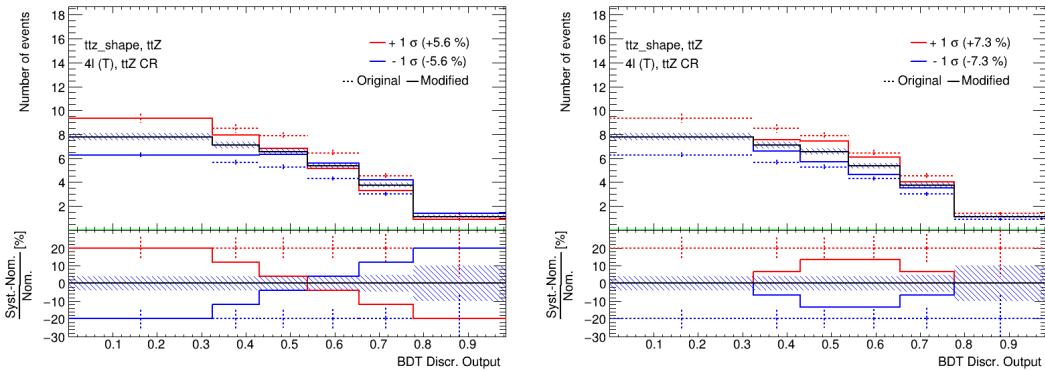


Figure 42:  $t\bar{t}Z$  generic shape systematic before (original) and after (modified) linear (left) and triangular (right) interpolation in the  $t\bar{t}Z$  CR is shown. The output from the event-level BDT shown on the x-axis. In the upper panel, the number of events is shown on the y-axis. In the lower panel, the difference between the systematic variation (the envelope's templates) and the nominal template, divided by the nominal template, is shown on the y-axis. The nominal  $t\bar{t}Z$  template is shown by the solid black lined histogram, with the diagonal lined bands representing its total uncertainty. The templates of the upper and lower envelopes (before modification) is given by the dotted red and blue lined histograms respectively. The templates of the upper and lower envelopes (after modification) is given by the solid red and blue lined histograms respectively, with the vertical dotted lines representing its total uncertainty.

Two shape systematics (linear and triangular shapes) are considered for  $tWZ$  and the most dominant background processes across the tri- and tetralepton channels. In particular, we consider these shape systematics for  $t\bar{t}Z$  for both the tri- and tetralepton channels. Additionally, we consider these shape systematics for the  $WZ + b$  and  $WZ + c$  backgrounds in the trilepton channel.

## 4.9 Analysis Pipeline and TRExFitter

We make use of industry standard **ROOT**<sup>4</sup> 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

<sup>4</sup>CERN's high energy physics data analysis framework (written in C++)

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[95]. In this analysis, we used TRexFitter (tag: TRexFitter-00-04-13) to produce all plots (including fit statistics, e.g. limit, significance,  $\mu_{tWZ}$ ).

The analysis pipeline starts with sample derivations (data from ATLAS with certain selection criteria applied in order to reduce its file size such that it is more manageable to work with) 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, variables, 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 scripts are used to define the different regions and apply the final cuts and selections outlined in Table 5. In addition to this purpose, they are used to train the two BDTs and to produce the output from these trained BDTs. As each event is looped over, the cuts and selection criteria are checked for the given event and is either discarded (if the event does not pass the selection criteria), or gets written to a ROOT file (if the event passes the selection criteria) corresponding to the MC sample and Run 2 data-set which it belongs to. These ROOT files contain all variables, weights and scale factors (corresponding to an event) which we wish to use in TRexFitter. TRexFitter then takes these files as input, runs a maximum likelihood fit and produces relevant plots (e.g. expected number of events, pull plots) and statistical parameters (e.g. limit, significance,  $\mu_{tWZ}$ ).

#### 4.9.1 Fitting Procedure

Maximum likelihood fits can be used to calculate the probability of data being consistent with a given hypothesis. In this analysis, the hypothesis is represented by the signal strength of  $tWZ$  production,  $\mu(tWZ)$ , defined as,

$$\mu(tWZ) = \frac{\sigma(tWZ)}{\sigma_{SM}(tWZ)} \quad (4.15)$$

where  $\sigma(tWZ)$  expected cross section of  $tWZ$  and  $\sigma_{SM}(tWZ)$  is the SM cross section of  $tWZ$ . A value of  $\mu(tWZ) = 0$  represents the background-only hypothesis, where no  $tWZ$  is present. A value of  $\mu(tWZ) = 1$  represents the signal+background hypothesis, where the expected  $tWZ$  signal is exactly the SM prediction. The maximum likelihood fit is performed by finding the set of values for the parameters which maximize a likelihood function or equivalently by minimizing the negative logarithm of the likelihood function. The data in a given bin of a distribution is expected to follow a Poisson probability distribution around its expected number of events. The likelihood function,  $L(\mu, \theta)$ , used in analyses involving binned histograms of signal and background events is therefore constructed as a product of Poisson probability terms for all bins ( $N_{bins}$ ) and is given by [BARLOW1993219],

$$L(\mu, \theta) = \prod_{N_{bins}}^i \frac{(\mu s_i(\vec{\theta}) + b_i(\vec{\theta}))^{n_i}}{n_i!} e^{-(\mu s_i(\vec{\theta}) + b_i(\vec{\theta}))} \prod_{\theta \in \vec{\theta}} \rho(\theta) \quad (4.16)$$

where  $\mu$  is the signal strength of the process of interest,  $s_i$  and  $b_i$  represent the expected number of signal and background events in bin  $i$  of the distribution respectively,  $\theta$  represents a nuisance parameter contained in the set of all nuisance parameters ( $\vec{\theta}$ ) and  $\rho(\theta)$  represents the probability density functions which define the nuisance parameter's distribution. In analyses with multiple regions, the likelihood function contains all bins across the regions. A fit across all regions in the tetralepton channel is performed to determine the sensitivity of  $tWZ$  in this channel. Plots shown prior to the fit are referred to as *pre-fit* and those shown after the fit are referred to as *post-fit*. In this analysis, the *mixed data and MC* fit setup [96] is used. This is done to obtain the most accurate prediction of the expected results while keeping the signal regions blinded. For this setup, first a background-only fit to the control regions using real data is performed to determine estimates of the nuisance parameters. Then these estimates are used to construct a modified Asimov dataset in the signal regions. Finally, the fit is performed using real data in the control regions and the aforementioned modified Asimov data-set in the signal regions. In these fits, the POI is  $\mu(tWZ)$ . The POI is ultimately the quantity to be measured and it is set as a *free parameter* in the fit. This means that during the fitting procedure,  $\mu(tWZ)$  is unconstrained and can take any value in the fit. The nuisance parameters are assigned to the systematic uncertainties outlined in Section 4.8. Furthermore, a gamma ( $\gamma$ ) nuisance parameter for a bin is added to the likelihood function if the statistical uncertainty in the bin exceeds 0.1% of its nominal value. Pruning is done per sample and per region on the shape and normalisation uncertainties for samples. A sample's shape and normalisation nuisance parameter is pruned (removed from the limit/fit) if the impact of the systematic uncertainty on the POI is less than 0.01. An auto-binning algorithm,

1269 1270 1271 1272 1273 1274 1275 1276 1277 1278 1279 1280 1281 1282 1283 1284 1285 1286 1287 1288 1289 1290 1291 1292 1293 1294 1295 1296 1297 1298 1299 1300 1301 1302 1303 1304 1305 1306 1307 1308 1309

TransfoD [Calvet:2296985, 33], was used to define the binning. This aims to maximise  $\frac{\text{signal}}{\text{background}}$  in each bin. Furthermore, it aims to avoid defining bins with a low number of events. The auto-binning algorithm works by scanning through the bins of the given distribution with user-defined bins of equal width, and merges bins until a certain fraction of  $\frac{\text{signal}}{\text{background}}$  events is obtained. The merging threshold,  $MT$ , is defined as,

$$MT = z_s \frac{n_s}{N_s} + z_b \frac{n_b}{N_b} \quad (4.17)$$

where  $n_s$  and  $n_b$  are the number of signal and background events in the merging bin, respectively. The total number of signal and background events is given by  $N_s$  and  $N_b$ , respectively. User-defined parameters,  $z_s$  and  $z_b$ , control the maximum fraction of signal and background events in each bin with the condition,  $z_s + z_b = \text{Number of Bins}$ . A bin is formed when  $MT \geq 1$ .

To characterise the sensitivity and associated uncertainty of our measurement of  $\mu(tWZ)$ , we compute two metrics: the expected significance ( $Z_\mu^{\exp}$ ) and the expected upper limit ( $\mu_{\text{up}}^{\exp}$ ). In this context, the expected significance can be interpreted as, the probability that the expected signal is due to a background fluctuation. The expected significance,  $Z_\mu^{\exp}$ , is related to the  $p$ -value. The  $p$ -value can be defined as the probability, under the assumption of a given hypothesis (in the context of this analysis, this would refer to the background-only hypothesis), of finding data of equal or greater incompatibility with the predictions of the hypothesis [43]. The expected significance can be written in terms of the  $p$ -value ( $p$ ) by,

$$Z_\mu^{\exp} = \Phi^{-1}(1 - p) \quad (4.18)$$

where  $\Phi^{-1}$  is the inverse of the cumulative function of the standard Gaussian distribution. This is defined such that a Gaussian distributed variable found  $Z_\mu^{\exp}$  standard deviations above its mean has an area under its rightmost tail equal to  $p$ . Larger values indicate lower probabilities and smaller values indicate higher probabilities. Particle physicists have adopted a standard to define the sensitivity necessary for evidence and discovery of a particular particle or phenomena. A  $3\sigma$  (corresponding to a background fluctuation probability of  $\approx 10^{-3}$ ) significance is considered to be evidence for observation and a  $5\sigma$  (corresponding to a background fluctuation probability of  $\approx 10^{-7}$ ) is considered to be a discovery. The expected upper limit is a single-sided interval test statistic, associated with the parameter of interest (POI) in the maximum-likelihood fit ( $\mu(tWZ)$ , in our case). In this context, the expected upper limit can be understood in the following way: consider running an ensemble of MC toy experiments, each with their own confidence interval (a range of possible values for  $\mu(tWZ)$ ). An expected upper limit, at some fixed percentage  $x\%$  (or *confidence level*), can be determined from this ensemble. The expected upper limit tells us that,  $x\%$  of the toy MC experiment's confidence intervals will contain the true value of  $\mu(tWZ)$ . A commonly used percentage in particle physics is 95%, which we adopt for this analysis. In particle physics, this is referred to as the *CLs Method* [43]. The CLs test statistic can be calculated 'brute force' by running these MC toy experiments, however this is very CPU intensive. Asymptotic formulae are able to describe the underlying CLs test statistic distributions under certain approximations [10]. Instead of running toy MC experiments, we use asymptotic formulae to perform the CLs method, which reduces computation time from several millions of CPU hours to several CPU minutes. The significance and upper limits which are calculated in this analysis are given a prefix of 'expected' to indicate that these are results from a blinded analysis.

## 4.10 Results

In the section, an expected upper limit and an expected significance are set on the signal strength of  $tWZ$ . This is performed for the current analysis in the tetralepton channel as well as for a combined analysis across the trilepton and tetralepton channels. The trilepton analysis was performed as an independent study by Benjamin Warren (UCT) [98]. Note that throughout this section, all signal regions remain blinded.

### 4.10.1 Tetralepton Channel

In Figure 43 expected number of events for the variables used in the likelihood fit in each region are shown. In Figure 44 expected number of events for the variables used in the likelihood fit in each region are shown. In Table 12, the expected number of events (after the fit) for each sample in each region is shown. The expected upper limit of  $tWZ$  in the tetralepton channel is measured as,  $\mu_{\text{up}}^{\exp} = 1.61^{+2.35}_{-1.16}$ . The expected significance of  $tWZ$  in the tetralepton channel is measured as,  $Z_\mu^{\exp} = 1.44\sigma$ . The best-fit value of the signal strength,  $\mu(tWZ) = \frac{\sigma(tWZ)}{\sigma(tWZ)_{SM}}$ , from the likelihood fit is measured as,  $\mu(tWZ) = 1.91^{+0.95}_{-0.82}$ . The best fit value

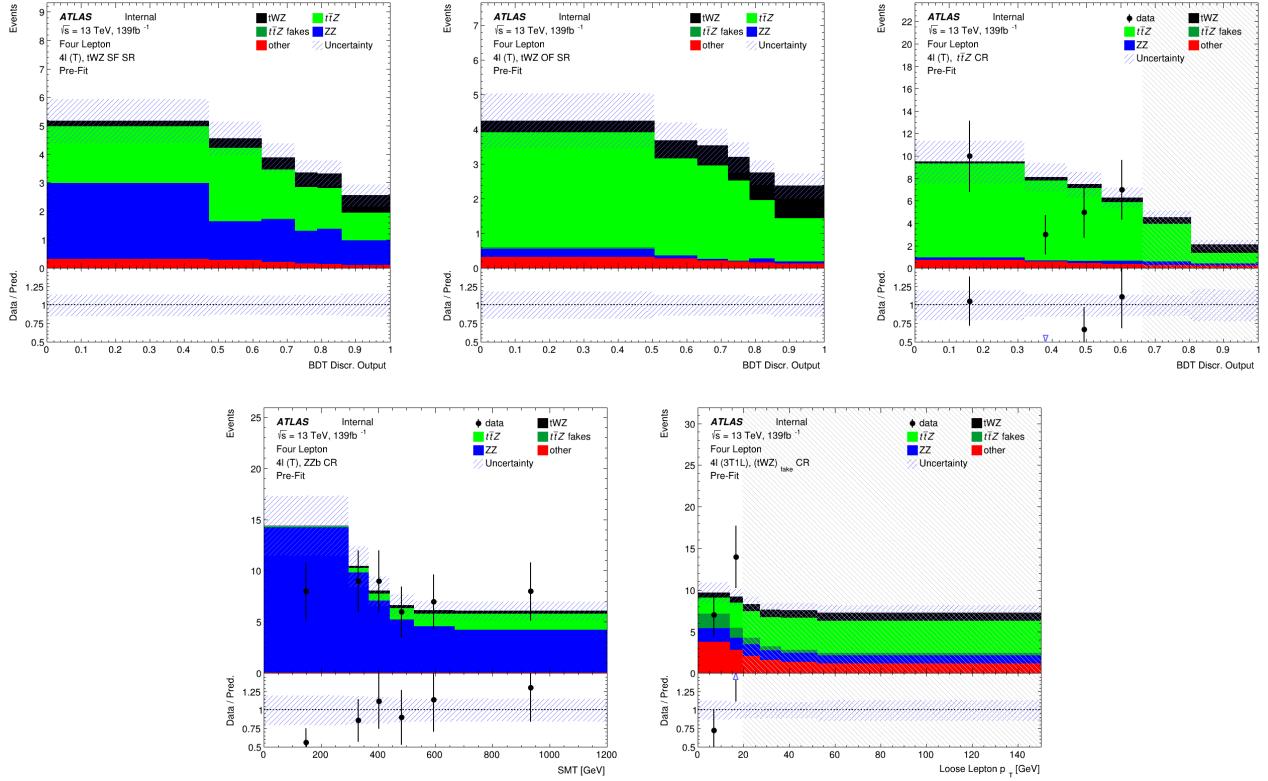


Figure 43: Pre-fit distributions (blinded) of variables used in the fit are shown. **From top left to bottom right:** The event-level BDT Disc. Output in the  $tWZ$  SF SR. The event-level BDT Disc. Output in the  $tWZ$  OF SR. The event-level BDT Disc. Output in the  $t\bar{t}Z$  CR.  $SMT = \sum p_T(\ell) + \sum p_T(jet) + E_T^{\text{miss}}$  in the  $ZZb$  CR.  $p_T$ (loose lepton) in the  $(tWZ)_\text{fake}$  (3T1L) CR. The data is given by the black points and the MC predictions for each process are given by the filled histograms. The vertical lines on the data points represent the statistical uncertainty in the data and the diagonally-lined bands represent the total (statistical and systematic added in quadrature) uncertainty. The lower panel in each plot shows the ratios of the data to the theoretical predictions. Bins in the CRs with  $\frac{\text{signal}}{\text{background}}$  greater than 0.1 are kept blinded. Blinded bins in the CRs are shaded with black diagonal lines and their data points are omitted. The plots in the  $tWZ$  OF SR and  $tWZ$  SF SR are kept blinded by omitting the data points.

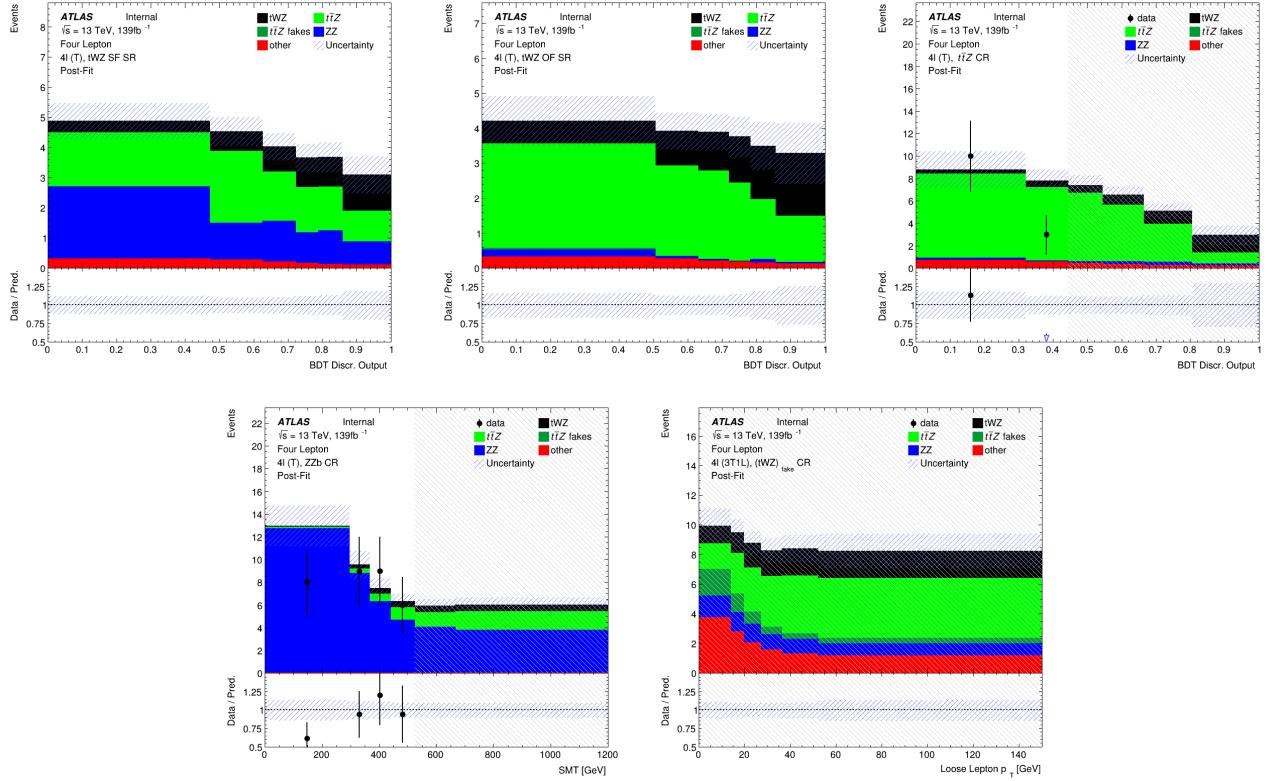


Figure 44: Post-fit distributions (blinded) of variables used in the fit are shown. **From top left to bottom right:** The event-level BDT Disc. Output in the  $tWZ$  SF SR. The event-level BDT Disc. Output in the  $tWZ$  OF SR. The event-level BDT Disc. Output in the  $t\bar{Z}$  CR.  $SMT = \sum p_T(\ell) + \sum p_T(jet) + E_T^{\text{miss}}$  in the  $ZZb$  CR.  $p_T$  (loose lepton) in the  $(tWZ)_\text{fake}$  (3T1L) CR. The data is given by the black points and the MC predictions for each process are given by the filled histograms. The vertical lines on the data points represent the statistical uncertainty in the data and the diagonally-lined bands represent the total (statistical and systematic added in quadrature) uncertainty. The lower panel in each plot shows the ratios of the data to the theoretical predictions. Bins in the CRs with  $\frac{\text{signal}}{\text{background}}$  greater than 0.1 are kept blinded. Blinded bins in the CRs are shaded with black diagonal lines and their data points are omitted. The plots in the  $tWZ$  OF SR and  $tWZ$  SF SR are kept blinded by omitting the data points.

	$tWZ$ OF SR	$tWZ$ SF SR	$t\bar{t}Z$ CR	$ZZb$ CR	$(tWZ)_{\text{fake}}$ CR
$t\bar{t}Z$	$13.2379 \pm 1.52295$	$9.62061 \pm 1.12291$	$29.9054 \pm 3.60908$	$5.08899 \pm 0.620121$	$18.5393 \pm 2.23036$
$t\bar{t}Z$ fakes	$0.0702522 \pm 0.0468691$	$0.0334067 \pm 0.0256903$	$0.0723509 \pm 0.0418526$	$0.0485273 \pm 0.029133$	$5.04378 \pm 2.34732$
$tWZ$	$7.83414 \pm 3.31679$	$5.33547 \pm 2.24801$	$5.69373 \pm 2.58041$	$2.89889 \pm 1.23837$	$10.278 \pm 4.3345$
$ZZ$	$0.481776 \pm 0.119774$	$7.72372 \pm 1.2351$	$1.07955 \pm 0.182461$	$40.6067 \pm 6.26078$	$6.86097 \pm 1.11443$
other	$tt$	$6.00553e-06 \pm 0.02819e-06$	$0.252557 \pm 0.442116$	$0.273507 \pm 0.223201$	$6.00553e-06 \pm 0.02819e-06$
	$tZq$	$0.0827905 \pm 0.0398773$	$0.0756107 \pm 0.0354584$	$0.063585 \pm 0.0293325$	$0.05884 \pm 0.0244084$
	$t\bar{t}W$	$0.00668643 \pm 0.00792217$	$0.00279748 \pm 0.00287361$	$6.00553e-06 \pm 0.02819e-06$	$0.002306 \pm 0.00564349$
	$WZ$	$0.0442934 \pm 0.024156$	$0.0396511 \pm 0.0154282$	$0.0133471 \pm 0.0128199$	$0.0472562 \pm 0.0330315$
	$t\bar{t}t$	$0.000987164 \pm 0.000766266$	$0.00247481 \pm 0.00136945$	$0.0140869 \pm 0.00479496$	$6.00553e-06 \pm 0.02819e-06$
	$t\bar{t}\bar{t}$	$0.00934035 \pm 0.0080554$	$0.0107458 \pm 0.00849984$	$0.0571373 \pm 0.0204011$	$6.00553e-06 \pm 0.02819e-06$
	$t\bar{t}WW$	$0.0294618 \pm 0.0263174$	$0.029771 \pm 0.0195582$	$0.264364 \pm 0.0926252$	$0.0129431 \pm 0.0323803$
	$VVV(V = W/Z)$	$0.280643 \pm 0.0853411$	$0.191433 \pm 0.0586778$	$0.0697266 \pm 0.0225059$	$0.171142 \pm 0.0518102$
	$t\bar{t}H$	$0.846054 \pm 0.175495$	$0.669375 \pm 0.140107$	$1.96662 \pm 0.401199$	$0.150025 \pm 0.0353826$
Total	$22.9243 \pm 2.96284$	$23.9876 \pm 2.11249$	$39.4734 \pm 3.44937$	$49.0856 \pm 6.04162$	$53.4282 \pm 4.31683$
data	-	-	36	49	57

Table 12: The expected number of events (after the fit) for each sample in each region is shown.

of the signal strength is inconsistent with 1 by a small margin. This indicates a disagreement between data and the post-fit model within post-fit model uncertainties. This may suggest that there is some mis-modelling in the backgrounds of the post-fit model. Since there is a small number of events in the regions of tetralepton channel, the disagreement may be due to statistical fluctuations in data or simulation.

1320

In Figure 45, a ranking plot showing the impact of the systematic uncertainties on the POI,  $\mu(tWZ)$  is shown.

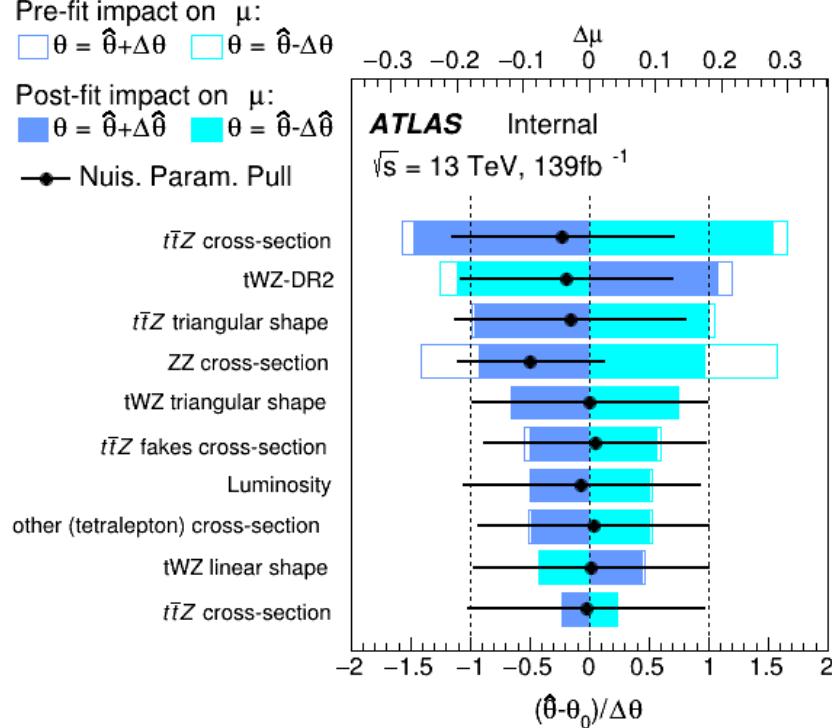


Figure 45: A ranking plot showing the impact (ordered from top to bottom via decreasing impact) of the systematic uncertainties (top 10) on the POI,  $\mu(tWZ)$ , in the tetralepton channel is shown.  $\hat{\theta}$  is the best-fit value of the nuisance parameter.  $\Delta\hat{\theta}$  and  $\Delta\theta$  are the post-fit and pre-fit uncertainties respectively. The post-fit and pre-fit impact of each nuisance parameter on  $\mu(tWZ)$  are shown with the solid and lined rectangles respectively. The empty and solid blue rectangles correspond to the pre-fit and post-fit impacts on  $\mu(tWZ)$  respectively. These impacts are shown on the upper axis ( $\Delta\mu$ ). On the lower axis, the nuisance parameter pull,  $\frac{\hat{\theta} - \theta_0}{\Delta\theta}$ , is shown ( $\theta_0$  is the nominal pre-fit value of the nuisance parameter). The nuisance parameter pull is indicated by the black points, with their relative post-fit errors ( $\frac{\Delta\hat{\theta}}{\Delta\theta}$ ) shown by the black horizontal error bars.

The most important systematics are the cross sections of  $t\bar{t}Z$  and  $ZZ$ , and shape modelling on  $t\bar{t}Z$  ( $t\bar{t}Z$  triangular shape) and  $tWZ$  ( $tWZ$ -DR2 and  $tWZ$  triangular shape). The cross section of  $ZZ$  is significantly shifted down

in the fit from its nominal value. The  $t\bar{t}Z$  cross section,  $tWZ$ -DR2 and  $t\bar{t}Z$  triangular shape nuisance parameters are similarly shifted down in the fit from their nominal values, but to a much lesser degree than the cross section of  $ZZ$ . These pulls are all within  $1\sigma$  uncertainty and are thus relatively small. It is expected that the modelling uncertainties (shape and normalisations) of the most dominant backgrounds (e.g.  $t\bar{t}Z$ ,  $ZZ$ ) have relatively large impacts on  $\mu(tWZ)$ , since the uncertainty of the analysis is dominated by statistical uncertainty.

#### 4.10.2 Trilepton and Tetralepton Channels

In the section, an expected upper limit and an expected significance are set on the signal strength of  $tWZ$  ( $\mu(tWZ)$ ) from the combined fit across all regions of  $tWZ$  in the tetralepton and trilepton channels.

The trilepton channel is an entirely independent analysis [98] that uses separate data and simulated events compared to the tetralepton channel. Although the trilepton channel uses separate data and simulated events, the treatment of the statistical and systematic uncertainties (and therefore the statistical model) in the combined fit is entirely coherent with the tetralepton channel. Therefore the inclusion of the trilepton channel in the combined fit is only a matter of adding the regions in the trilepton channel to the likelihood definition. The trilepton analysis follows a similar analysis strategy to that of the tetralepton analysis. It includes an event-level BDT which aims to discriminate between the  $tWZ$  and all background as well as an object-level BDT which aims to identify hadronically decaying  $W$  bosons to discriminate between  $tWZ$  and the large  $WZ$  background. One  $tWZ$  SR is defined and five CRs are defined.  $WZ$  and  $t\bar{t}Z$  CRs are defined to constrain the dominant  $WZ$  and  $t\bar{t}Z$  backgrounds. Three CRs which require that one of the three selected leptons are loose, are defined for  $WZ$ ,  $t\bar{t}Z$  and  $tWZ$  in order to constraint the fake lepton component (using the MC template method - similar to the method used in Section 4.6 to estimate the fake lepton component).

The expected upper limit of  $tWZ$  in the trilepton channel is measured as,  $\mu_{up}^{exp} = 2.65^{+3.67}_{-1.91}$ . The expected significance of  $tWZ$  in the trilepton channel is measured as,  $Z_\mu^{exp} = 0.75\sigma$ . The best-fit value of the signal strength,  $\mu(tWZ) = \frac{\sigma(tWZ)}{\sigma(tWZ)_{SM}}$ , from the likelihood fit is measured as,  $\mu(tWZ) = 1.16^{+1.33}_{-1.30}$ . The best fit value of the signal strength is consistent with 1. This indicates an agreement between data and the post-fit model within post-fit model uncertainties. Therefore, this demonstrates that the constructed model able to sufficiently model the background processes.

In Table 13, the nuisance parameters used in the fit, including which channel's regions are affected by each, are shown.

The expected upper limit of  $tWZ$  across both channels is measured as,  $\mu_{up}^{exp} = 1.43^{+2.04}_{-1.03}$ . The expected significance of  $tWZ$  across both channels is measured as,  $Z_\mu^{exp} = 1.61\sigma$ . The best-fit value of the signal strength,  $\mu(tWZ) = \frac{\sigma(tWZ)}{\sigma(tWZ)_{SM}}$ , from the likelihood fit is measured as,  $\mu(tWZ) = 1.80^{+0.70}_{-0.65}$ . In Figure 46, the expected upper limits of the trilepton channel, tetralepton channel and both channels combined are shown.

It can be seen that the sensitivity of  $tWZ$  is mostly driven by the tetralepton analysis, with the trilepton analysis attributing a small decrease in the expected upper limit of the combined analysis, and its associated uncertainty. In Figure 47, the best-fit values of  $\mu(tWZ)$  from the fit for the trilepton channel, tetralepton channel and both channels combined are shown.

It can be seen from Figures 46 and 47 that the best fit value for the signal strength on  $tWZ$ ,  $\mu(tWZ)$ , and the expected limits for the tri- and tetralepton channels are consistent with one-another (their uncertainties overlap). Therefore it is appropriate to combine these two analyses.

The effect of each systematic uncertainty on the POI,  $\mu(tWZ)$ , in the fit can be studied from a *ranking plot* [stats-at-atlas-cms-slides]. A ranking plot lists the nuisance parameters used in the fit in order of their *impact* on the POI. The impact of a nuisance parameter describes to what degree the nuisance parameter effects the extracted value of the POI from the fit. The impact of a nuisance parameter (denoted by  $\theta$ ) is calculated by first fixing the pre-fit and post-fit uncertainties (denoted by  $\Delta\theta$  and  $\Delta\hat{\theta}$  respectively) of all the nuisance parameters to  $\pm 1$ , performing the fit, and extracting the nominal value of the POI. The impact of a given nuisance parameter is then defined by the difference in this reference POI value to a POI value extracted from a fit performed by removing the given nuisance parameter. In Figure 48, a ranking plot showing the impact of the systematic uncertainties on the POI,  $\mu(tWZ)$ , in the combined fit across both the tri- and tetralepton channels is shown.

Some nuisance parameters are pulled down from their nominal pre-fit values, however these are all within  $1\sigma$  uncertainty and are thus relatively small. The most important systematics are the cross sections of  $t\bar{t}Z$ ,  $ZZ$

Nuisance Parameter	Channel Affected	
	trilepton	tetralepton
$\sigma(t\bar{t}H)$	✓	✓
$\sigma(t\bar{t}Z)$	✓	✓
$\sigma(WZ)$	✓	✓
$\sigma(tZq)$	✓	✓
$\sigma(ZZ)$	✓	✓
$\sigma(\text{other(trilepton)})$	✓	✗
$\sigma(\text{other(tetralepton)})$	✗	✓
$\sigma(t\bar{t}Z)_{\text{fakes}}$	✗	✓
$\sigma(t\bar{t})_{\text{fakes}}$	✓	✗
$\sigma(Z + \text{jets})_{\text{fakes}}$	✓	✗
Luminosity	✓	✓
$jvt$	✓	✓
pileup	✓	✓
DL1r SF (b jets)	✓	✓
DL1r SF (light jets)	✓	✓
$\sigma(tWZ - DR2)$	✗	✓
lepton SF	✓	✓
$t\bar{t}Z$ triangular shape	✓	✓
$t\bar{t}Z$ linear shape	✓	✓
$WZ + b$ triangular shape	✓	✗
$WZ + b$ linear shape	✓	✗
$WZ + c$ triangular shape	✓	✗
$WZ + c$ linear shape	✓	✗

Table 13: A summary of the nuisance parameters used in the combined fit is shown. The channels which are affected by each nuisance parameter are indicated with a ✓ (is affected) or a ✗ (is not affected).

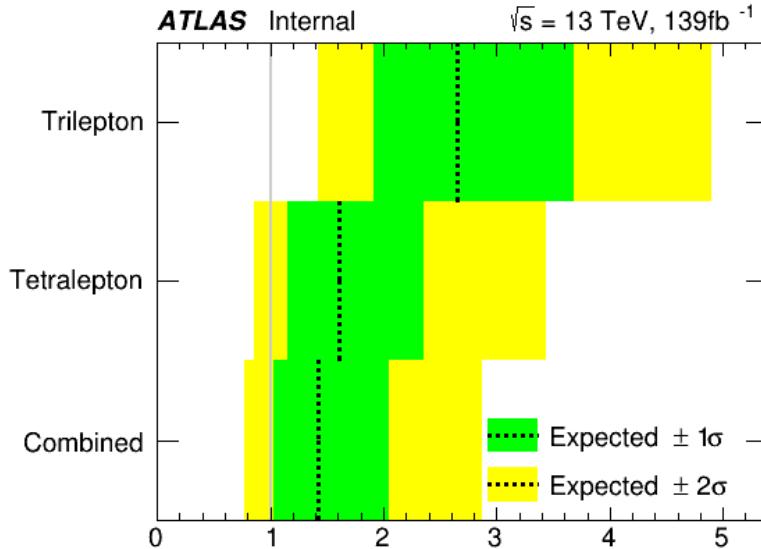


Figure 46: The expected upper limits of the trilepton channel, tetralepton channel and both channels combine are shown. The y-axis shows the channels in which the fitting procedure was performed. The expected limits are represented by the vertical dotted line. One- and two-  $\sigma$  uncertainty bands are shown in green and yellow respectively. The vertical grey line indicates when  $\mu(tWZ) = 1$ .

and  $WZ + b$ , and shape modelling on  $t\bar{t}Z$  ( $t\bar{t}Z$  triangular shape). The aforementioned nuisance parameters are significantly shifted down in the fit from its nominal value. The less important systematics are pulled in the fit from their nominal values, but to a much lesser degree than the cross section of those mentioned above. It is expected that the modelling uncertainties (shape and normalisations) of the most dominant backgrounds (e.g.  $t\bar{t}Z$ ,  $ZZ$  and  $WZ + b$ ) have relatively large impacts on  $\mu(tWZ)$ , since the uncertainty of the analysis is dominated by statistical uncertainty.

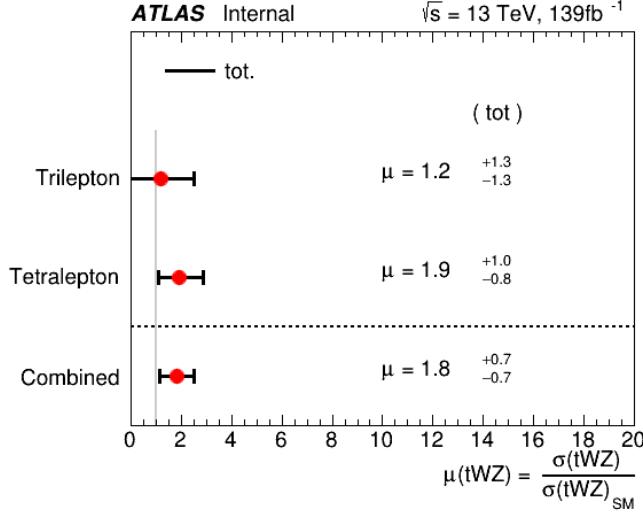


Figure 47: The best-fit values of  $\mu(tWZ)$  from the fit for the trilepton channel, tetralepton channel and both channels combined are shown. The y-axis shows the channels in which the fitting procedure was performed. The signal strength  $\mu(tWZ)$  is shown on the x-axis. The nominal signal strengths are represented by the red dots. The total uncertainty associated with the best-fit  $\mu(tWZ)$  value is shown by the black error bars.

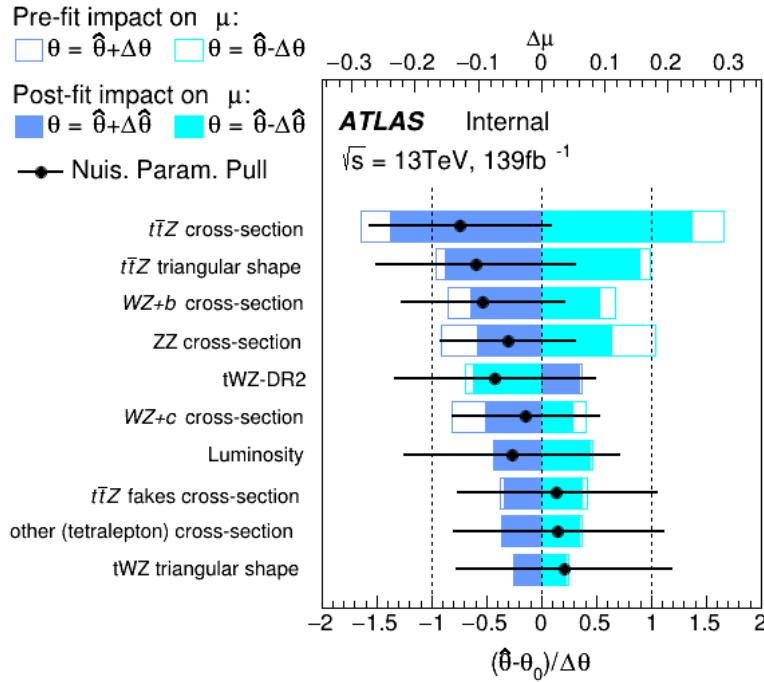


Figure 48: A ranking plot showing the impact of systematic uncertainties (top 10) on the POI,  $\mu(tWZ)$ , in the combined fit across both the tri- and tetralepton channels is shown.  $\hat{\theta}$  is the best-fit value of the nuisance parameter.  $\Delta\hat{\theta}$  and  $\Delta\theta$  are the post-fit and pre-fit uncertainties respectively. The post-fit and pre-fit impact of each nuisance parameter on  $\mu(tWZ)$  are shown with the solid and lined rectangles respectively. The empty and solid blue rectangles correspond to the pre-fit and post-fit impacts on  $\mu(tWZ)$  respectively. These impacts are shown on the upper axis ( $\Delta\mu$ ). On the lower axis, the nuisance parameter pull,  $\frac{\hat{\theta} - \theta_0}{\Delta\theta}$ , is shown ( $\theta_0$  is the nominal pre-fit value of the nuisance parameter). The nuisance parameter pull is indicated by the black points, with their relative post-fit errors ( $\frac{\Delta\hat{\theta}}{\Delta\theta}$ ) shown by the black horizontal error bars.

#### 1384 4.10.2.1 Projection to Higher Luminosity

1385 The expected significance of  $tWZ$  across both channels was measured as  $1.61\sigma$  and it is likely that the analysis is  
 1386 statistically limited. Therefore, the question naturally arises whether or not it is possible to observe  $tWZ$  with a  $3\sigma$

1387 or  $5\sigma$  significance and how much data one would need to achieve this. In this section we apply a fully blinded fit to  
 1388 the Asimov dataset for integrated luminosities larger than the  $139\text{fb}^{-1}$  currently available from the ATLAS Full Run  
 1389 2 dataset. This study gives us insight into the sensitivity of  $tWZ$  which we could expect if we were to replicate the  
 1390 current analysis, given more data. Given the upgrades planned for the LHC and the ATLAS detector, we will soon  
 1391 expect a large increase in available  $pp$  collision data (reaching  $\mathcal{L} = 3000\text{fb}^{-1}$  [36]). This study therefore gives us an  
 1392 idea of the sensitivity of  $tWZ$  which we can expect in the future. It would also provide a clear indication whether  
 1393 or not the uncertainties in this analysis is dominated by statistical uncertainties. In Figure 49, the expected upper  
 1394 limit and significance for combined fully blinded fits to the Asimov dataset across both the tri- and tetralepton  
 1395 channels for a range integrated luminosities are shown.

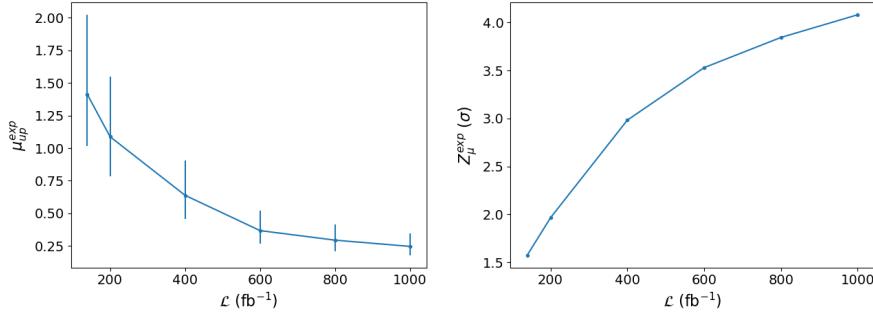


Figure 49: The expected upper limit (left) and significance (right) for combined fully blinded fits to the Asimov dataset across both the tri- and tetralepton channels for a range integrated luminosities are shown. **Left:** The integrated luminosity,  $\mathcal{L}$ , is shown on the x-axis. The expected upper limit,  $\mu_{up}^{exp}$ , is shown on the y-axis. The vertical lines represent the total uncertainty ( $\pm 1 \sigma$ ) on the expected upper limit. **Right:** The integrated luminosity,  $\mathcal{L}$ , is shown on the x-axis. The expected significance,  $Z_\mu^{exp}$ , is shown on the y-axis. The vertical lines represent the total uncertainty ( $\pm 1 \sigma$ ) on the expected significance.

1396 As we increase the total integrated luminosity, the sensitivity of  $tWZ$  increases. The increase in sensitivity is shown  
 1397 by the rapid decrease of the expected upper limit and the rapid increase of the expected significance with increasing  
 1398 integrated luminosity. It can be noted that a  $3\sigma$  expected significance is achieved at  $\mathcal{L} \sim 400 \text{ fb}^{-1}$ , indicating  
 1399 evidence for observation of  $tWZ$ . The  $5\sigma$  significance ‘discovery’ standard is above our  $1000\text{fb}^{-1}$  luminosity range  
 1400 of study. Furthermore, these plots show that the sensitivity of  $tWZ$  is hindered by the low amount of events we  
 1401 observe, rather than the systematic effects. This is evident since increasing the luminosity, therefore increasing the  
 1402 number of events we see in the detector, directly causes a steady increase in sensitivity.

1403

# Chapter 5

1404

## Summary and Conclusions

The  $tWZ$  process is an important process in the search for new physics since its cross section is sensitive to the charged and neutral couplings to the top quark, which is strongly coupled to the Higgs boson. The top quark's couplings are often modified in many scenarios of new physics that aim to resolve the Hierarchy Problem, therefore a constraint placed on the cross section of  $tWZ$  is expected to be impactful in constraining such BSM models. A search for  $tWZ$  production using  $139 \text{ fb}^{-1}$  of  $pp$  collision data at a centre-of-mass energy of  $\sqrt{s} = 13 \text{ TeV}$ , recorded by the ATLAS experiment at CERN, has been presented. This thesis targeted the tetralepton final state channel. Two SRs and three CRs were defined. Two SRs, instead of one, were defined in order to suppress and constrain the  $ZZ$  background. The dominant background processes,  $t\bar{t}Z$  and  $ZZ$  were constrained by the definition of  $t\bar{t}Z$  and  $ZZ$  CRs, respectively. The dominant source of fake leptons, originating from the  $t\bar{t}Z$  background, was constrained by the  $(tWZ)_{\text{fake}}$  CR, using the MC template method. Two BDTs were implemented: an object-level BDT which aims to classify between  $\ell b$  systems coming from top quarks and an event-level BDT which aims to discriminate between  $tWZ$  and our major backgrounds,  $t\bar{t}Z$  and  $ZZ$ . The output from the object-level BDT was converted to an event-level variable to be used as input to the event-level BDT. A kinematic reconstruction algorithm,  $2\nu\text{SM}$ , was used to reconstruct top quarks in order to discriminate between  $tWZ$  and  $t\bar{t}Z$ . The output from this algorithm was used as an input variable to the event-level BDT. The trained BDT was shown to discriminate well between signal and background events. Using a modified Asimov dataset in the SRs and real data in the CRs, a blinded maximum-likelihood fit was performed across all regions in the tetralepton channel. The best-fit value of the signal strength in the tetralepton channel was,

$$\mu(tWZ) = 1.91^{+0.95}_{-0.82} \quad (5.1)$$

with an expected significance of  $1.44\sigma$ . The expected upper limit on the signal strength of  $tWZ$  in the tetralepton channel was,

$$\mu_{up}^{\text{exp}} = 1.61^{+2.35}_{-1.16} \quad (5.2)$$

To further increase the sensitivity of  $tWZ$ , a blinded maximum-likelihood fit was performed across all regions across the trilepton (studied in an independent analysis by Benjamin Warren (UCT) [98]) and tetralepton channels. The best-fit value of the signal strength across both the trilepton and tetralepton channels were,

$$\mu(tWZ) = 1.80^{+0.70}_{-0.65} \quad (5.3)$$

with an expected significance of  $1.61\sigma$ . The expected upper limit on the signal strength of  $tWZ$  across both the trilepton and tetralepton channels were,

$$\mu_{up}^{\text{exp}} = 1.43^{+2.04}_{-1.03} \quad (5.4)$$

Although this result does not satisfy the  $3\sigma$  evidence nor the  $5\sigma$  discovery standards, an unblinding of this analysis in the future may still provide the tightest ever constraint on the  $tWZ$  process. The results in this analysis are heavily statistically limited, it is therefore expected that future analyses of this process, using larger datasets (such as that from the HL-LHC), would significantly improve the results.

1434

# Bibliography

- [1] M. Aaboud et al. “Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at  $\sqrt{s} = 13$  TeV”. In: *The European Physical Journal C* 79.8 (Aug. 2019). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-019-7140-6. URL: <http://dx.doi.org/10.1140/epjc/s10052-019-7140-6>.
- [2] M. Aaboud et al. “Measurement of the  $t\bar{t}Z$  and  $t\bar{t}W$  cross sections in proton-proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector”. In: *Phys. Rev. D* 99 (7 Apr. 2019), p. 072009. DOI: 10.1103/PhysRevD.99.072009. URL: <https://link.aps.org/doi/10.1103/PhysRevD.99.072009>.
- [3] 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: %7B[http://dx.doi.org/10.1007/JHEP10\(2019\)127](http://dx.doi.org/10.1007/JHEP10(2019)127)%7D.
- [4] G. Aad et al. “Measurement of the  $t\bar{t}$  production cross-section in the lepton+jets channel at  $s=13$  TeV with the ATLAS experiment”. In: *Physics Letters B* 810 (2020), p. 135797. ISSN: 0370-2693. DOI: <https://doi.org/10.1016/j.physletb.2020.135797>. URL: <https://www.sciencedirect.com/science/article/pii/S0370269320306006>.
- [5] G. Aad et al. “Muon reconstruction performance of the ATLAS detector in proton–proton collision data at  $\sqrt{s} = 13$  TeV”. In: *The European Physical Journal C* 76.5 (May 2016). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-016-4120-y. URL: <http://dx.doi.org/10.1140/epjc/s10052-016-4120-y>.
- [6] Shunichi Akatsuka and Shion Chen. *Isolation WPs summary: PLV + LowPtPLV*. Oct. 2019. URL: <https://indico.cern.ch/event/854783/contributions/3595486/attachments/1929380/3195230/PLV-Summary.pdf>.
- [7] G. ALTARELLI, R. BARBIERI, and F. CARAVAGLIOS. “ELECTROWEAK PRECISION TEST: A CONSISE REVIEW”. In: *International Journal of Modern Physics A* 13.07 (Mar. 1998), pp. 1031–1058. ISSN: 1793-656X. DOI: 10.1142/s0217751x98000469. URL: <http://dx.doi.org/10.1142/S0217751X98000469>.
- [8] Guido Altarelli. *The Higgs and the Excessive Success of the Standard Model*. 2014. arXiv: 1407.2122 [hep-ph].
- [9] J. Alwall et al. “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”. In: *JHEP* 07 (2014), p. 079. DOI: 10.1007/JHEP07(2014)079. arXiv: 1405.0301 [hep-ph].
- [10] Aaron Armbruster. *Asymptotic Formulae*. Feb. 2013. URL: [https://indico.cern.ch/event/233551/contributions/493678/attachments/389871/542293/asymptotics\\_armbruster.pdf](https://indico.cern.ch/event/233551/contributions/493678/attachments/389871/542293/asymptotics_armbruster.pdf).
- [11] Pierre Astier et al. “Kalman filter track fits and track breakpoint analysis”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 450.1 (2000), pp. 138–154. ISSN: 0168-9002. DOI: [https://doi.org/10.1016/S0168-9002\(00\)00154-6](https://doi.org/10.1016/S0168-9002(00)00154-6). URL: <https://www.sciencedirect.com/science/article/pii/S0168900200001546>.
- [12] ATLAS Collaboration. “ATLAS  $b$ -jet identification performance and efficiency measurement with  $t\bar{t}$  events in  $pp$  collisions at  $\text{sqrt}s = 13$  TeV”. In: *Eur. Phys. J. C* 79 (2019), p. 970. DOI: 10.1140/epjc/s10052-019-7450-8. arXiv: 1907.05120 [hep-ex].
- [13] ATLAS Collaboration. *Calibration of light-flavour  $b$ -jet mistagging rates using ATLAS proton–proton collision data at  $\text{sqrt}s = 13$  TeV*. ATLAS-CONF-2018-006. 2018. URL: <https://cds.cern.ch/record/2314418>.

- [14] ATLAS Collaboration. “Electron and photon performance measurements with the ATLAS detector using the 2015-2017 LHC proton-proton collision data”. In: *JINST* 14 (2019), P12006. doi: 10.1088/1748-0221/14/12/P12006. arXiv: 1908.00005 [hep-ex].
- [15] ATLAS Collaboration. “Identification and rejection of pile-up jets at high pseudorapidity with the ATLAS detector”. In: *Eur. Phys. J. C* 77 (2017), p. 580. doi: 10.1140/epjc/s10052-017-5081-5. arXiv: 1705.02211 [hep-ex].
- [16] ATLAS Collaboration. *Measurement of b-tagging efficiency of c-jets in  $t\bar{t}$  events using a likelihood approach with the ATLAS detector*. ATLAS-CONF-2018-001. 2018. URL: <https://cds.cern.ch/record/2306649>.
- [17] ATLAS Collaboration. “Measurements of  $b$ -jet tagging efficiency with the ATLAS detector using  $t\bar{t}$  events at  $\sqrt{s} = 13\text{TeV}$ ”. In: *JHEP* 08 (2018), p. 089. doi: 10.1007/JHEP08(2018)089. arXiv: 1805.01845 [hep-ex].
- [18] ATLAS Collaboration. “Muon reconstruction and identification efficiency in ATLAS using the full Run 2  $pp$  collision data set at  $\sqrt{s} = 13\text{TeV}$ ”. In: (2020). arXiv: 2012.00578 [hep-ex].
- [19] Manuel Guth on behalf of the ATLAS collaboration. *Deep-Neural-Network-based b-Tagging as Basis for Improvements in Top Analyses*. URL: <https://cds.cern.ch/record/2693088/files/ATL-PHYS-SLIDE-2019-751.pdf>.
- [20] ATLAS Internal. *Electron Efficiencies for Analyses*. 2021. URL: <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/ElectronEfficienciesForAnalysis>.
- [21] ATLAS Internal. *Electron Efficiency Correlation Model*. 2021. URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/ElectronEfficiencyCorrelationModel>.
- [22] ATLAS Internal. *Jet Vertex Tagger for Run 2 in reco and analysis*. 2021. URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/JetVertexTaggerTool>.
- [23] ATLAS Internal. *Muon Efficiencies for Analyses*. 2021. URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/MuonEfficienciesForAnalysis>.
- [24] G. Avoni et al. “The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS”. In: *JINST* 13.07 (2018), P07017. doi: 10.1088/1748-0221/13/07/P07017.
- [25] Richard D. Ball et al. “Parton distributions for the LHC run II”. In: *Journal of High Energy Physics* 2015.4 (Apr. 2015). ISSN: 1029-8479. doi: 10.1007/jhep04(2015)040. URL: [http://dx.doi.org/10.1007/JHEP04\(2015\)040](http://dx.doi.org/10.1007/JHEP04(2015)040).
- [26] Olga Bessidskaia Bylund. *Measurement of  $t\bar{t}Z$  and  $t\bar{t}W$  production at ATLAS in 13 TeV data, using trilepton and same charge dimuon final states*. Tech. rep. Geneva: CERN, Aug. 2016. doi: 10.22323/1.276.0237. URL: <http://cds.cern.ch/record/2211022>.
- [27] Enrico Bothmann et al. “Event generation with Sherpa 2.2”. In: *SciPost Physics* 7.3 (Sept. 2019). ISSN: 2542-4653. doi: 10.21468/scipostphys.7.3.034. URL: <http://dx.doi.org/10.21468/SciPostPhys.7.3.034>.
- [28] Ilaria Brivio and Michael Trott. “The standard model as an effective field theory”. In: *Physics Reports* 793 (Feb. 2019), pp. 1–98. ISSN: 0370-1573. doi: 10.1016/j.physrep.2018.11.002. URL: <http://dx.doi.org/10.1016/j.physrep.2018.11.002>.
- [29] Jason Brownlee. *Overfitting and Underfitting With Machine Learning Algorithms*. URL: <https://machinelearningmastery.com/overfitting-and-underfitting-with-machine-learning-algorithms/>.
- [30] Oliver Sim Brüning et al. *LHC Design Report*. CERN Yellow Reports: Monographs. Geneva: CERN, 2004. doi: 10.5170/CERN-2004-003-V-1. URL: <https://cds.cern.ch/record/782076>.
- [31] Gustavo Burdman et al. “Colorless top partners, a 125 GeV Higgs boson, and the limits on naturalness”. In: *Physical Review D* 91.5 (Mar. 2015). ISSN: 1550-2368. doi: 10.1103/physrevd.91.055007. URL: <http://dx.doi.org/10.1103/PhysRevD.91.055007>.
- [32] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. “The anti- $k_t$  jet clustering algorithm”. In: *JHEP* 04 (2008), p. 063. doi: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ph].
- [33] Thomas Calvet. *Automatic binning implementation in TTHFitter - Htop(bb)*. URL: [https://indico.cern.ch/event/455289/contributions/1953694/attachments/1209081/1762963/Calvet\\_binning\\_Htop-160108.pdf](https://indico.cern.ch/event/455289/contributions/1953694/attachments/1209081/1762963/Calvet_binning_Htop-160108.pdf).

- [34] Anadi Canepa. "Searches for supersymmetry at the Large Hadron Collider". In: *Reviews in Physics* 4 (2019), p. 100033. ISSN: 2405-4283. DOI: <https://doi.org/10.1016/j.revip.2019.100033>. URL: <https://www.sciencedirect.com/science/article/pii/S2405428318300091>.
- [35] CERN Twiki - TOP WG Summary Plots. URL: <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCTopWGSummaryPlots>.
- [36] CERN Yellow Reports: Monographs. *CERN Yellow Reports: Monographs, Vol. 10 (2020): High-Luminosity Large Hadron Collider (HL-LHC): Technical design report.* en. 2020. DOI: 10.23731/CYRM-2020-0010. URL: <https://e-publishing.cern.ch/index.php/CYRM/issue/view/127>.
- [37] Shion Chen. *Track isolation variable for the PFlow WPs.* Oct. 2019. URL: [https://indico.cern.ch/event/854783/contributions/3595529/attachments/1926980/3190772/IFF\\_20191003\\_PflowWPs.pdf](https://indico.cern.ch/event/854783/contributions/3595529/attachments/1926980/3190772/IFF_20191003_PflowWPs.pdf).
- [38] KyungEon Choi. "Tracking and Vertexing with the ATLAS Inner Detector in the LHC Run-2". In: *Springer Proc. Phys.* 213 (2018). Ed. by Zhen-An Liu, pp. 400–403. DOI: 10.1007/978-981-13-1316-5\_75.
- [39] CMS Collaboration. *Measurements of  $pp \rightarrow ZZ$  production cross sections and constraints on anomalous triple gauge couplings at  $\sqrt{s} = 13$  TeV.* 2020. arXiv: {2009.01186} (hep-ex).
- [40] CMS Collaboration. "Measurement of the ttbar production cross section using events with one lepton and at least one jet in pp collisions at  $\sqrt{s}=13$  TeV". In: (Jan. 2017).
- [41] The ATLAS Collaboration et al. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08003–S08003. DOI: 10.1088/1748-0221/3/08/s08003. URL: <https://doi.org/10.1088/1748-0221/3/08/s08003>.
- [42] I. Connelly. "Performance and calibration of b-tagging with the ATLAS experiment at LHC Run-2". In: 2017.
- [43] Glen Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *The European Physical Journal C* 71.2 (Feb. 2011). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-011-1554-0. URL: <http://dx.doi.org/10.1140/epjc/s10052-011-1554-0>.
- [44] Federico Demartin et al. "tWH associated production at the LHC". In: *Eur. Phys. J. C* 77.1 (2017), p. 34. DOI: 10.1140/epjc/s10052-017-4601-7. arXiv: 1607.05862 [hep-ph].
- [45] *Electron identification efficiency in data for electrons with  $E_T > 30\text{GeV}$ .* URL: [https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/EGAM-2018-01/fig\\_16.png](https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/EGAM-2018-01/fig_16.png).
- [46] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". In: *Phys. Rev. Lett.* 13 (9 Aug. 1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.13.321>.
- [47] *Errors in weighted histograms.* URL: <https://www.zeuthen.desy.de/~wischnew/amanda/discussion/wgterror/working.html>.
- [48] Lyndon Evans and Philip Bryant. "LHC Machine". In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08001–S08001. DOI: 10.1088/1748-0221/3/08/s08001. URL: <https://doi.org/10.1088/1748-0221/3/08/s08001>.
- [49] Guido Fantini et al. *The formalism of neutrino oscillations: an introduction.* 2020. arXiv: 1802.05781 [hep-ph].
- [50] *Forward Jet Vertex Tagging: A new technique for the identification and rejection of forward pileup jets.* Tech. rep. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2015-034>. Geneva: CERN, Aug. 2015. URL: <https://cds.cern.ch/record/2042098>.
- [51] Stefano Frixione, Paolo Nason, and Carlo Oleari. "Matching NLO QCD computations with parton shower simulations: the POWHEG method". In: *JHEP* 11 (2007), p. 070. DOI: 10.1088/1126-6708/2007/11/070. arXiv: 0709.2092 [hep-ph].
- [52] Michele Gallinaro. "Top quark physics: A tool for discoveries". In: *Journal of Physics Conference Series* 447 (July 2013), pp. 2012–. DOI: 10.1088/1742-6596/447/1/012012.
- [53] *Good Run Lists for Analysis Run 2 - ATLAS Twiki.* URL: <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/GoodRunListsForAnalysisRun2>.
- [54] *GoodRunListsForAnalysisRun2.* URL: <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/GoodRunListsForAnalysisRun2>.

- [55] 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>.
- [56] Trevor Hastie. *The elements of statistical learning : data mining, inference, and prediction*. New York: Springer, 2009. ISBN: 978-0-387-84858-7.
- [57] Peter W. Higgs. “Broken Symmetries and the Masses of Gauge Bosons”. In: *Phys. Rev. Lett.* 13 (1964). Ed. by J. C. Taylor, pp. 508–509. DOI: 10.1103/PhysRevLett.13.508.
- [58] Peter W. Higgs. “Spontaneous Symmetry Breakdown without Massless Bosons”. In: *Phys. Rev.* 145 (1966), pp. 1156–1163. DOI: 10.1103/PhysRev.145.1156.
- [59] *IFFTruthClassifier GitLab Repository*. URL: <https://gitlab.cern.ch/ATLAS-IFF/IFFTruthClassifier/-/tree/master>.
- [60] *IFFTruthClassifier Lepton Categories*. URL: <https://gitlab.cern.ch/ATLAS-IFF/IFFTruthClassifier/-/tree/master#3-details-about-the-lepton-categories>.
- [61] N. Jarosik et al. “SEVEN-YEAR WILKINSON MICROWAVE ANISOTROPY PROBE ( WMAP ) OBSERVATIONS: SKY MAPS, SYSTEMATIC ERRORS, AND BASIC RESULTS”. In: *The Astrophysical Journal Supplement Series* 192.2 (Jan. 2011), p. 14. ISSN: 1538-4365. DOI: 10.1088/0067-0049/192/2/14. URL: <http://dx.doi.org/10.1088/0067-0049/192/2/14>.
- [62] F. Jegerlehner. “The hierarchy problem of the electroweak Standard Model revisited”. In: (May 2013).
- [63] V. Khachatryan et al. “Measurement of the ZZ production cross section and  $Z \rightarrow \ell^+\ell^-\ell^+\ell^-$  branching fraction in pp collisions at  $\sqrt{s} = 13\text{TeV}$ ”. In: *Physics Letters B* 763 (2016), pp. 280–303. ISSN: 0370-2693. DOI: <https://doi.org/10.1016/j.physletb.2016.10.054>. URL: <https://www.sciencedirect.com/science/article/pii/S0370269316306256>.
- [64] H. J. W. Kirsten. *Introduction to supersymmetry*. Singapore Hackensack, NJ: World Scientific, 2010. ISBN: 978-9814293426.
- [65] Arthur Alexis Jules Lesage. *Lepton and photon performance at ATLAS and CMS*. 2017. arXiv: 1709.02598 [hep-ex].
- [66] Fabio Maltoni, Luca Mantani, and Ken Mimasu. “Top-quark electroweak interactions at high energy”. In: *Journal of High Energy Physics* 2019.10 (Oct. 2019). ISSN: 1029-8479. DOI: 10.1007/jhep10(2019)004. URL: [http://dx.doi.org/10.1007/JHEP10\(2019\)004](http://dx.doi.org/10.1007/JHEP10(2019)004).
- [67] Luigi Marchese. *Lepton and photon reconstruction and identification performance in ATLAS and CMS*. Tech. rep. Geneva: CERN, Sept. 2019. DOI: 10.22323/1.350.0237. URL: <https://cds.cern.ch/record/2688452>.
- [68] Thomas McCarthy. *Macro developed to compare t/W/Z reconstruction performance (2 $\ell$ , 3 $\ell$ , 4 $\ell$ )*. URL: [https://indico.cern.ch/event/986357/contributions/4172907/attachments/2169451/3666801/reco\\_performance\\_macro\\_20210112.pdf](https://indico.cern.ch/event/986357/contributions/4172907/attachments/2169451/3666801/reco_performance_macro_20210112.pdf).
- [69] Thomas McCarthy and Florian Fischer. *Exploiting full/partial  $t\bar{t}$  reconstruction for background suppression in 2 $\ell$* . URL: [https://indico.cern.ch/event/955360/contributions/4016465/attachments/2102418/3534816/top\\_reco\\_bkgd\\_suppression\\_2L\\_20200915.pdf](https://indico.cern.ch/event/955360/contributions/4016465/attachments/2102418/3534816/top_reco_bkgd_suppression_2L_20200915.pdf).
- [70] *MCTruthClassifier - ATLAS Twiki*. URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/MCTruthClassifier>.
- [71] *Measurement of the  $t\bar{t}$  production cross-section and lepton differential distributions in  $e\mu$  dilepton events from pp collisions at  $\sqrt{s} = 13\text{ TeV}$  with the ATLAS detector*. Tech. rep. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2019-041>. Geneva: CERN, Aug. 2019. URL: <https://cds.cern.ch/record/2686255>.
- [72] F. Meloni. “Primary vertex reconstruction with the ATLAS detector”. In: *Journal of Instrumentation* 11.12 (Dec. 2016), pp. C12060–C12060. DOI: 10.1088/1748-0221/11/12/c12060. URL: <https://doi.org/10.1088/1748-0221/11/12/c12060>.
- [73] Ken Mimasu. *Top quark electroweak interactions*. 2021. arXiv: 2105.10261 [hep-ph].
- [74] Aditya Mishra. *Metrics to Evaluate your Machine Learning Algorithm*. URL: <https://towardsdatascience.com/metrics-to-evaluate-your-machine-learning-algorithm-f10ba6e38234>.

- [75] Users MissMJ and Cush. *Standard model of elementary particles - Wikimedia Commons*. URL: [https://en.wikipedia.org/wiki/File:Standard\\_Model\\_of\\_Elementary\\_Particles.svg](https://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg).
- [76] *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>.
- [77] Hussain Mujtaba. *What is Cross Validation in Machine learning? Types of Cross Validation*. URL: <https://www.mygreatlearning.com/blog/cross-validation/>.
- [78] *MuonSelectionTool, ATLAS TWiki*. URL: <https://twiki.cern.ch/twiki/bin/view/Atlas/MuonSelectionTool>.
- [79] 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>.
- [80] António Onofre. “Top Quark Couplings and Search for New Physics at the LHC”. In: *Journal of Physics Conference Series* 447 (July 2013), pp. 2030–. DOI: 10.1088/1742-6596/447/1/012030.
- [81] *Option to force a shape withing an error band by hand - TRF documentation*. URL: [https://trexfitter-docs.web.cern.ch/trexfitter-docs/model\\_building/shape/](https://trexfitter-docs.web.cern.ch/trexfitter-docs/model_building/shape/).
- [82] *Performance of the ATLAS Electron and Photon Trigger in p-p Collisions at sqrt(s) = 7 TeV in 2011*. Tech. rep. All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2012-048>. Geneva: CERN, May 2012. URL: <https://cds.cern.ch/record/1450089>.
- [83] *Pileup jet recommendations*. URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/PileupJetRecommendations>.
- [84] *Recommended isolation working points*. URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/RecommendedIsolationWPs>.
- [85] *scikit-learn Documentation - Gradient Boosting Classifier*. URL: <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.GradientBoostingClassifier.html>.
- [86] *Scikit-Learn GradientBoostingClassifier Documentation*. URL: <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.GradientBoostingClassifier.html>.
- [87] Pedro Ferreira da Silva. *Top quark production at the LHC*. 2016. arXiv: 1605.05343 [hep-ex].
- [88] T. Sjöstrand, S. Mrenna, and P. Skands. “A brief introduction to PYTHIA 8.1”. In: *Comput. Phys. Commun.* 178 (2008), pp. 852–867. DOI: 10.1016/j.cpc.2008.01.036. arXiv: 0710.3820 [hep-ph].
- [89] Torbjörn Sjöstrand et al. “An introduction to PYTHIA 8.2”. In: *Comput. Phys. Commun.* 191 (2015), p. 159. DOI: 10.1016/j.cpc.2015.01.024. arXiv: 1410.3012 [hep-ph].
- [90] Jory Sonneveld. *Searches for physics beyond the standard model at the LHC*. 2019. arXiv: 1905.06239 [hep-ex].
- [91] Mark Thomson. *Modern particle physics*. Cambridge, United Kingdom New York: Cambridge University Press, 2013. ISBN: 9781107034266.
- [92] *TopRecoObjTwikiModel*. URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TopRecoObjTwikiModel>.
- [93] *Total Integrated Luminosity and Data Quality in 2015-2018 - LuminosityPublicResultsRun2*. URL: <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2>.
- [94] Daniel R. Tovey. “Transformation properties of the transverse mass under transverse Lorentz boosts at hadron colliders”. In: *The European Physical Journal C* 79.4 (Apr. 2019). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-019-6813-5. URL: <http://dx.doi.org/10.1140/epjc/s10052-019-6813-5>.
- [95] *TRExFitter*. URL: <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/TtHFitter>.
- [96] *TRExFitter: Mixed data and MC fit*. URL: <https://trexfitter-docs.web.cern.ch/trexfitter-docs/AdvancedTutorial2020/Mixed/>.
- [97] Tommy Tschida. “Cross section measurements of processes with a single top quark and two vector bosons with the CMS experiment”. MA thesis. Vienna, Tech. U., 2020. DOI: 10.34726/hss.2020.79420.
- [98] Benjamin Warren. “A search for tWZ production in the trilepton channel using Run 2 data from the ATLAS experiment.” In: (2021).

- 1676 [99] Kenneth G. Wilson. “Confinement of quarks”. In: *Phys. Rev. D* 10 (8 Oct. 1974), pp. 2445–2459. DOI:  
1677 10.1103/PhysRevD.10.2445. URL: <https://link.aps.org/doi/10.1103/PhysRevD.10.2445>.
- 1678 [100] Katherine Woodruff. *Introduction to boosted decision trees*. URL: <https://indico.fnal.gov/event/15356/contributions/31377/attachments/19671/24560/DecisionTrees.pdf>.  
1679