



# 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 . . . . .	15
36	3.2.1 Coordinate System and Kinematics . . . . .	16
37	3.2.2 Inner Detector . . . . .	16
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 . . . . .	42
65	4.7.4 Event-level BDT . . . . .	49
66	4.8 Systematics . . . . .	53

67	4.8.1	Experimental uncertainties . . . . .	53
68	4.8.2	Theoretical uncertainties . . . . .	56
69	4.8.3	Generic shape systematics . . . . .	56
70	4.9	Analysis Pipeline and TRexFitter . . . . .	57
71	4.9.1	Fitting Procedure . . . . .	58
72	4.10	Results . . . . .	59
73	4.10.1	Tetralepton Channel . . . . .	59
74	4.10.2	Trilepton and Tetralepton Channels . . . . .	62
75	<b>5</b>	<b>Conclusion</b>	<b>66</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 top  
 81 quark to the Higgs boson, the electroweak couplings of the top quark are a theoretically well-motivated area in which  
 82 to search for the first signs of new physics [**TopQuarkCouplings-searchNewPhysics, top-a-tool-for-disc**] that  
 83 could offer a resolution to the Hierarchy Problem [**Burdman 2015**, 54]. The recent lack of signs of new physics  
 84 from LHC data tells us that new physics is either very heavy, or is very weakly coupled to Standard Model particles,  
 85 therefore signs of new physics might only be observed in anomalous rates of well-chosen processes. A prime example  
 86 of such a process is  $tWZ$ . This has an extremely low production cross section ( $\approx 160$  fb for  $\sqrt{s} = 13$  TeV [25]),  
 87 meaning that it is an extremely rare process to observe and subsequently, it has never been observed by any particle  
 88 physics experiment.

89

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

102

# Chapter 2

103

## Theory

### 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 experi-  
 106 mental data to great precision [ALTARELLI 1998, 8]. For example, in the top quark sector, the  $t\bar{t}$  cross section  
 107 predictions have been confirmed to 3.9% accuracy [cms-ttbar, ATLAS-CONF-2019-041]. It incorporates three  
 108 of the four fundamental forces of nature: the electromagnetic, the weak and the strong forces. In Figure 1, all  
 109 known elementary particles described by the SM, are shown.  
 110

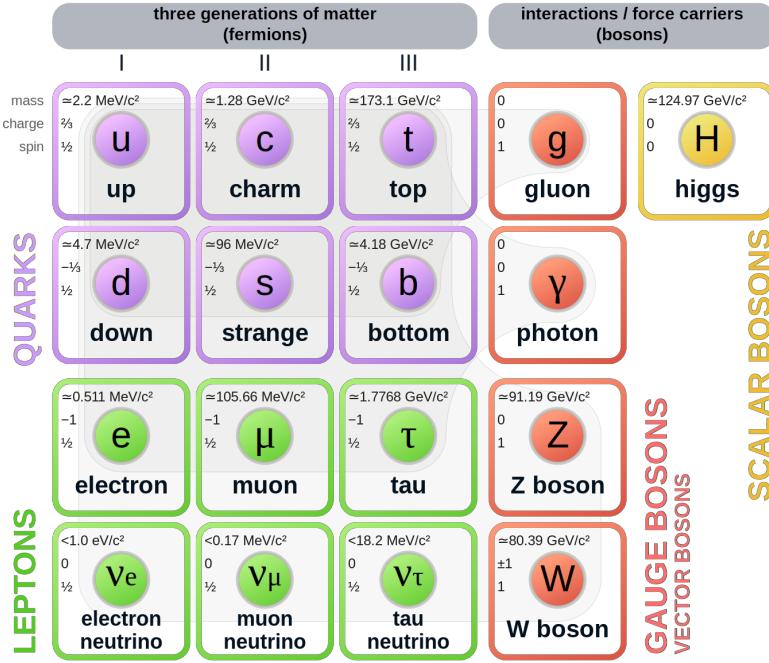


Figure 1: A summary of the elementary particles described by the SM [64] is shown. Fermions are shown on the left, with quarks shown in purple and leptons shown in green. Bosons are shown on the right, with gauge bosons shown in red and the Higgs boson shown in yellow. The mass, electric charge and spin of each particle is shown on the top left of each particle's block.

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

116 charged leptons have an electric charge of  $-1$  and all neutrinos are electrically neutral. The masses of the particles  
 117 in a generation increase with increasing generation number, with generation 1 particles being the least massive and  
 118 generation 3 particles being the most massive. Quarks carry electric and colour charge, and can therefore interact  
 119 via the electromagnetic, weak and strong forces. Colour charge can take on three values: red, green and blue.  
 120 It is important to note that colour charge is completely unrelated to the everyday meaning of colour, and it just  
 121 represents the quantum state of the particle. Due to colour confinement [82], quarks cannot be isolated from one  
 122 another. They exist in colourless bound states, called hadrons, consisting of two or more quarks. Hadrons consisting  
 123 of an even number of quarks are known as mesons and those consisting of an odd number of quarks are known as  
 124 baryons. On the other hand, charged leptons (electron ( $e$ ), muon ( $\mu$ ) and tau ( $\tau$ )) only carry electric charge and  
 125 can therefore interact electromagnetically and weakly, but not through the strong interaction. The electric and  
 126 colour neutral fermions, neutrinos, can only interact via the weak force.

127  
 128 Particles are able to interact with one-another via the exchange of a gauge boson (boson with spin-1). Photons  
 129 are massless, spin-1 gauge bosons which mediate electromagnetic interactions between particles which carry electric  
 130 charge, such as quarks and charged leptons ( $e$ ,  $\mu$  and  $\tau$ ). The weak interaction is mediated by three massive gauge  
 131 bosons, the electrically charged  $W^+$  and  $W^-$  bosons and the electrically neutral  $Z$  boson. Gluons are massless,  
 132 spin-1 gauge bosons which mediate strong interactions between particles which carry colour charge, such as quarks.  
 133 Since gluons carry colour charge, they interact with themselves. The massive, spin-0, electrically neutral Higgs  
 134 boson mediates the Higgs field which gives mass to the  $W^\pm$  and  $Z$  bosons via the so-called Brout-Englert-Higgs  
 135 mechanism [42, 50, 49]. The Brout-Englert-Higgs mechanism induces spontaneous electroweak symmetry breaking  
 136 to provide mass terms for the  $W^\pm$  and  $Z$  bosons in the electroweak Lagrangian of the SM. All particles described in  
 137 the SM have their own antiparticle, with the same mass, but opposite charges. Some particles, such as the photon,  
 138 are their own antiparticle.

139  
 140 Although the SM has shown to be hugely successful, it is incomplete and fails to describe certain observed phenomena.  
 141 The most notable example being the absence of gravity from the SM. The gravitational force is  $\approx 10^{29}$  [75]  
 142 weaker than the weak force, therefore quantum gravitational effects are expected to only become significant at  
 143 energies much larger than that currently accessible by the LHC (known as the Planck scale  $\approx 10^9$  GeV) [54]. This  
 144 large difference in strength between the weak force and gravity is known as the Hierarchy Problem. Cosmological  
 145 observations infer that around 84% of the matter in the universe consists of gravitationally interacting matter known  
 146 as dark matter [53]. None of the particles described in the SM are good dark matter candidates, therefore the SM  
 147 only accounts for a small fraction of the total matter of the universe. The large discrepancy between the observed  
 148 amount of matter and antimatter in the universe, sometimes referred to as the matter-antimatter asymmetry, is not  
 149 fully explained by the SM. Neutrinos in the SM are assumed to be massless, however observations of neutrino oscillations  
 150 (neutrinos undergoing flavour change as they travel through space) imply that neutrinos do have mass [45].  
 151 Beyond the Standard Model (BSM) theories attempt to explain the phenomena which the SM cannot. For example,  
 152 a popular extension to the SM, Supersymmetry (SUSY) introduces new particles to the SM which are counterparts  
 153 to the existing SM particles with the same quantum numbers, except for their spins [56]. SUSY provides elegant  
 154 explanations to many shortcomings of the SM, however none of the supersymmetric particles described by SUSY  
 155 have been observed experimentally [30].

### 156 2.1.1 The Top Quark

157 The top quark is the heaviest particle in the SM, with a mass of  $172.76 \pm 0.30$  GeV [48]. According to the SM, since  
 158 the coupling to the Higgs boson is proportional to the the mass of the interacting particle, the top quark is strongly  
 159 coupled to the Higgs boson. Physics processes involving top quarks is therefore a theoretically well-motivated area  
 160 to search for new physics, since it is the most likely particle to couple to new physics theories at the TeV scale. Its  
 161 large mass also makes it highly unstable, with a mean lifetime of  $\approx 0.5 \times 10^{-24}$  s [48]. The top quark's lifetime  
 162 is shorter than that of the hadronisation process, and it therefore decays before hadronising. The top quark can  
 163 therefore be measured indirectly via its decay products. Top quarks almost always decay to a  $W$  boson and a  
 164  $b$ -quark ( $\frac{\Gamma(Wb)}{\Gamma(Wq(q=b,s,d))} = 0.957 \pm 0.034$  [48]). The  $b$ -quark is the second heaviest quark in the SM, however its  
 165 lifetime is still longer than the hadronisation time scale [48]. In hadron collider experiments,  $b$ -quarks travel a short  
 166 distance in the detector before hadronising to form jets. In Table 1, the dominant final state branching fractions of  
 167 the top quark are shown.

168 Hadronic final states are more than twice as likely than leptonic final states. Final state decays to different lepton  
 169 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 [48] are shown.

170  
 Top quark production can be placed into two main categories: pair production ( $t\bar{t}$ ) and single-top production  
 171 ( $t$ ) [73]. In the LHC, top quarks are mainly produced in pairs via strong interactions in gluon-gluon fusion ( $gg \rightarrow t\bar{t}$ )  
 172 or quark annihilation ( $q\bar{q} \rightarrow t\bar{t}$ ). Top quark production via gluon-gluon fusion is the dominating process [24]. The  
 173 production cross section for  $t\bar{t}$  (leptonic final state) in  $pp$  collisions with  $\sqrt{s} = 13$  TeV was measured by ATLAS with  
 174 a value of  $830 \pm 0.4(\text{stat}) \pm 36(\text{syst}) \pm 14(\text{lumi})$  pb [4], with good agreement between measurement and theoretical  
 175 prediction.  
 176

177 Single top production occurs via the weak interaction. The most abundant production mechanisms leading to  
 178 single top production are the  $s$ -,  $t$ - and  $Wt$ - channels [48]. In the  $s$ -channel, an initial quark annihilates with an  
 179 anti-quark of different flavour, producing a virtual  $W$  boson which decays to a top quark and anti-bottom quark.  
 180 In the  $t$ -channel, an initial  $b$  quark interacts with a different flavour quark via the exchange of a  $W$  boson. This  
 181 interaction produces a top quark and another quark. In the  $Wt$ -channel, an initial gluon interacts with a  $b$  quark  
 182 to produce a top quark and a  $W$  boson, either via the absorption of the gluon by the  $b$  quark or via the exchange  
 183 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  
 184  $Wt$ -channels, are shown.  
 185

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 [31]  
 are shown. The prime superscript on  $q'$  indicates that the quark has a different flavour to  $q$ .

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

### 188 2.1.1.1 Motivation for the search for $tWZ$ production in the tetralepton channel

189 The recent lack of signs of new physics from LHC data [74] tells us that new physics is either very heavy, or is  
 190 very weakly coupled to SM particles. We therefore might only observe signs of new physics in anomalous rates of  
 191 well-chosen processes.  $tWZ$  is a prime example of such a process. It has an extremely low production cross section  
 192 (0.7 fb for  $\sqrt{s} = 13$  TeV [25]), and has subsequently never been observed by any particle physics experiment. Since  
 193  $tWZ$  involves a charged  $W$  boson and neutral  $Z$  boson, its cross section is sensitive to the charged and neutral  
 194 couplings to the top quark. In turn, the top quark is strongly coupled to the Higgs boson, due to its large mass. Due  
 195 to the top quark's large coupling to the Higgs boson, corrections to the Higgs boson mass diverge in the SM. The  
 196 top quark's couplings are modified, in order to remove this divergence, in many scenarios of new physics that aim  
 197 to resolve the Hierarchy Problem. Since the  $Z$  boson may be radiated from the initial-state  $b$ -quark, the final-state  
 198 top quark, or the final-state  $Z$  boson, the  $tWZ$  process embeds the  $b - Z$ ,  $t - Z$  and  $W - Z$  electroweak couplings  
 199 which are often modified in BSM physics. Therefore  $tWZ$  is an important process in the search for signs of new  
 200 physics and BSM physics.

One such BSM theory which is sensitive to  $tWZ$  production [57, 63] is the Standard Model Effective Field Theory (SMEFT) [26]. 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 [67]. The second study investigated the feasibility of a cross section measurement of  $tWZ$  production with CMS Run 3 data ( $300 \text{ fb}^{-1}$ ) [Tschida:2020ftz]. 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 [81]. 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.

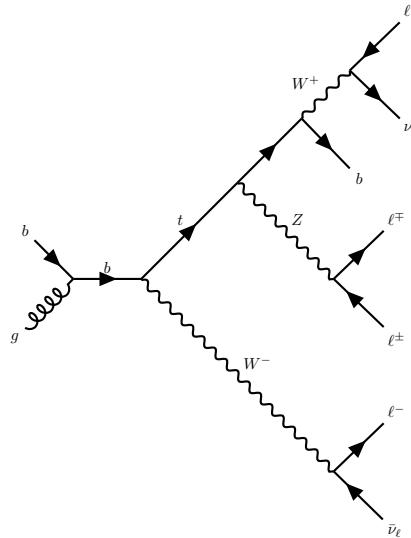


Figure 2: The LO Feynman diagram of  $tWZ$  production in the tetralepton channel is shown.

#### 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

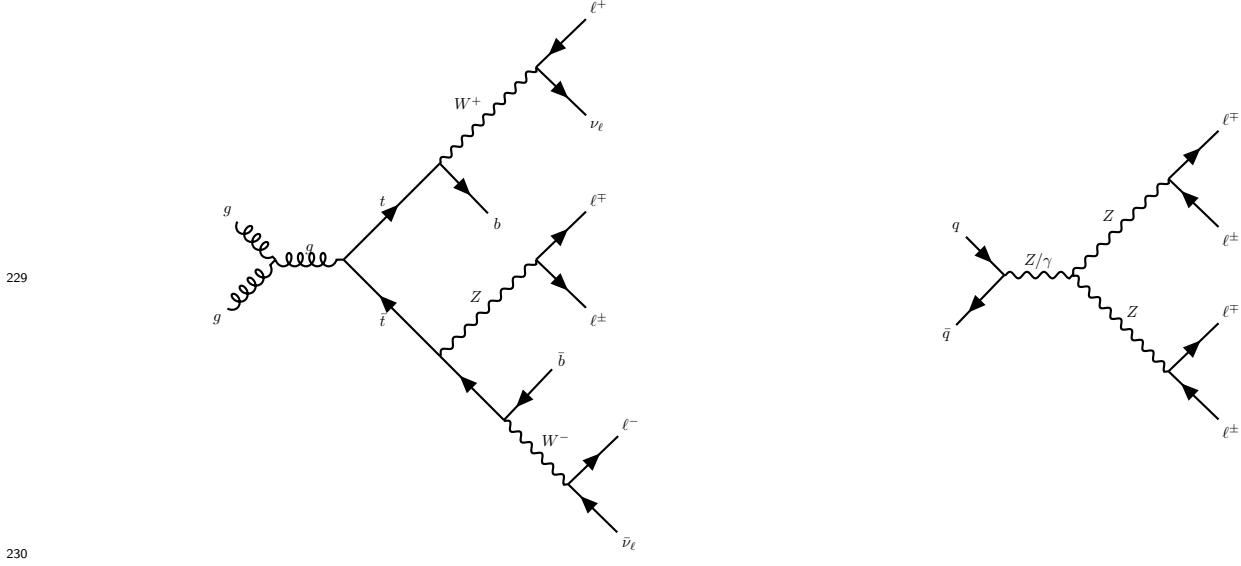


Figure 3: LO Feynman diagrams for  $t\bar{t}Z$  (left) and  $ZZ$  (right) in the tetralepton channel are shown.

231 The  $t\bar{t}Z$  process contains four leptons and two  $b$ -quarks in its final state (inclusive  $\sigma(t\bar{t}Z) = 0.95 \pm 0.08_{\text{stat}} \pm 0.10_{\text{syst}}$   
232 pb at  $\sqrt{s} = 13$  TeV [2]) and can easily mimic the  $tWZ$  signal process, for instance, by one of its  $b$ -jets getting  
233 missed during detection. The  $ZZ$  process contains four leptons and zero  $b$ -quarks in its final state (inclusive  
234  $\sigma(ZZ) = 14.6^{+1.9}_{-1.8}(\text{stat})^{+0.5}_{-0.3}(\text{syst}) \pm 0.2(\text{theo}) \pm 0.4(\text{lumi})$  pb at  $\sqrt{s} = 13$  TeV [55]). One way in which  $ZZ$  can mimic  
235 the  $tWZ$  signal process is by reconstruction of a non-prompt  $b$ -jet.

### 236 2.2.2 Comparison to Trilepton Channel

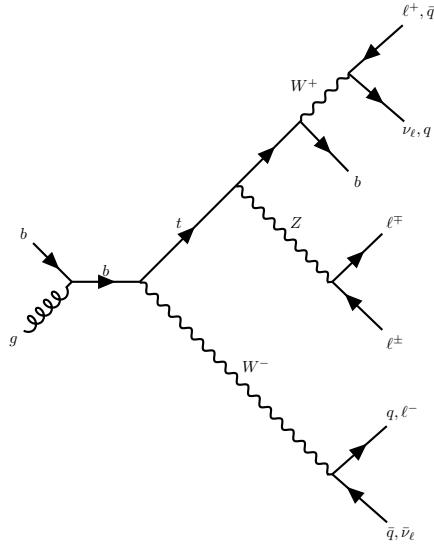


Figure 4: Example Feynman diagram of  $tWZ$  production in the tri-lepton channel.

237 The most apparent difference between the tri and tetralepton channels is the number of events present, with the  
238 tetralepton channel having far less events in its phase space than that of the tri-lepton channel. The lack of statistics  
239 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 [25].  
240 The tri-lepton channel has a production cross section ( $\sigma_{(tW^\pm Z).Br(3\ell)}^{\text{NLO}} = 3.9$  fb [25]) around a factor of 4 larger  
241 than that of the tetralepton channel. This difference between the production cross section of the two decay  
242 channels can be largely attributed to the difference in branching ratios ( $\frac{\Gamma_i}{\Gamma}$ ) between a hadronically decaying  $W$

243 boson,  $\frac{\Gamma_{W \rightarrow had}}{\Gamma_W} = (67.41 \pm 0.27)\%$  [48], present in the tri-lepton channel and a leptonically decaying  $W$  boson,  
244  $\frac{\Gamma_{W \rightarrow \ell\nu}}{\Gamma_W} = (10.86 \pm 0.09)\%$  [48], present in the tetralepton channel.

245  
246 Despite the tetralepton channel's low statistics, it is not subject to the large  $WZ$  background present in the trilepton  
247 channel [81]. The tetralepton channel has a substantial amount of  $ZZ$  background (not present in the trilepton  
248 channel), fortunately this can be easily suppressed due to the full reconstructability of the two leptonically decaying  
249  $Z$ -bosons.

250

# Chapter 3

## The ATLAS Experiment and Detector

251

### 3.1 The ATLAS Experiment

253 The ATLAS (A Toroidal LHC ApparatuS) detector is one of two general purpose detectors at CERN (the European  
 254 Organization for Nuclear Research) near Geneva in Switzerland. These detectors collect data from the collisions  
 255 provided by the worlds highest energy particle accelerator [27], the Large Hadron Collider (LHC) situated at CERN.

256 In this section, information about the LHC and the ATLAS detector are given. This includes technical aspects  
 257 of the ATLAS detector and the processing of data into meaningful physics objects<sup>1</sup> to be used in analyses. The  
 258 following chapter consists of information from "The LHC Design Report" [27], "LHC Machine" [44] and "The  
 259 ATLAS Experiment at the CERN Large Hadron Collider" [37] unless otherwise stated.

#### 3.1.1 Large Hadron Collider (LHC)

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

275 The four main experiments located at the interaction points are ATLAS, the Compact Muon Solenoid (CMS),  
 276 Large Hadron Collider Beauty (LHCb) Experiment and A Large Ion Collider Experiment (ALICE). The ATLAS  
 277 and CMS detectors are general-purpose detectors which investigate a wide range of physics processes. Since both  
 278 ATLAS and CMS can measure the same processes, they are able to cross-check and validate measurements taken  
 279 by one another. The LHCb detector is specifically designed to study decays of particles containing  $b$ -quarks. The  
 280 ALICE detector is designed to study the strongly interacting quark-gluon plasma which is formed at extremely  
 281 high energy densities. At the interaction points, the two proton beams which consist of protons in closely packed  
 282 bunches, travel in opposite directions to one another and collide. Many hard  $pp$  collisions (events) can occur per  
 283 bunch crossing, however it is the most energetic collision in the bunch crossing that is interesting for discovery  
 284 potential. The most energetic collision is therefore chosen to be studied and any additional collisions are aimed  
 285 to be rejected. These additional collisions are referred to as *pile-up*. Pileup complicates the reconstruction of the  
 286 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)

289 **3.1.1.1 Luminosity**

290 This section consists of information from "Modern Particle Physics" [75], unless otherwise stated.

291

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

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

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

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

296 where  $f$  is the beam revolution frequency,  $N$  is the number of proton bunches colliding per second,  $n_1$  and  $n_2$  are the  
297 number of protons in the colliding bunches,  $\sigma_x$  and  $\sigma_y$  are the beam spread in the  $x$  and  $y$  directions respectively.  
298 The total integrated luminosity,  $L$ , across some time interval, is given by,

$$L = \int \mathcal{L} dt. \quad (3.3)$$

299 The units of  $L$  are inverse area, and are given by  $\text{fb}^{-1}$  at the LHC and the ATLAS detector. In Figure 5, the  
300 total integrated luminosity delivered to ATLAS, recorded by ATLAS, and certified to be good enough for physics  
301 analyses (the data passes certain quality control criteria) for  $\sqrt{s} = 13$  TeV  $pp$  collisions at the LHC is shown [77].

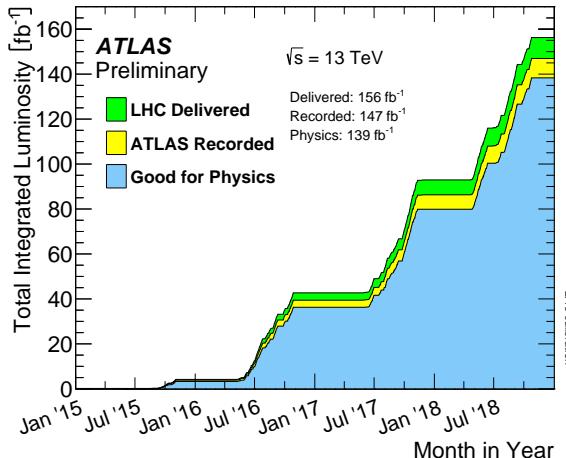


Figure 5: The total integrated luminosity delivered to ATLAS, recorded by ATLAS, and certified to be good enough for physics analyses (the data passes certain quality control criteria) for  $\sqrt{s} = 13$  TeV  $pp$  collisions at the LHC is shown [77]. The total integrated luminosity delivered by the 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.

302 A total integrated luminosity of  $139 \text{ fb}^{-1}$  of data certified as good for physics was recorded by ATLAS between  
303 2015 and 2018. This data taking period is referred to as Run 2, since it proceeds the Run 1 data taking period  
304 (2011 and 2012) and the Long Shutdown 1 LHC upgrade period (2013 and 2014). In this analysis, we use the Full  
305 Run 2 dataset.

306 **3.2 The ATLAS Detector**

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

308 The detector is cylindrically shaped which covers close to  $4\pi$  in solid angle. It has a length of 44 m, a diameter  
309 of 25 m and a mass of 7000 tons. The ATLAS detector consists of four main sub-detectors arranged in concentric  
310 cylindrical layers around the beam pipe. These include the inner detector, the electromagnetic calorimeter, the

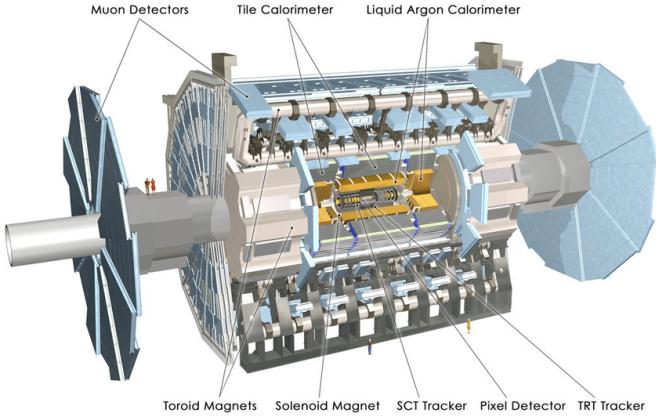


Figure 6: Schematic of the ATLAS detector [37]

hadronic calorimeters and the muon spectrometer. The sub-detectors record the momenta, energies and trajectories of different particles produced in the collider, allowing for the reconstruction and identification of these particles to be used in physics analyses.

### 3.2.1 Coordinate System and Kinematics

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

$$\eta = -\ln \tan\left(\frac{\theta}{2}\right) \quad (3.4)$$

$\eta$  is often used as a measure of the polar angle, instead of  $\theta$ , since the difference in  $\eta$  between two particles,  $\Delta\eta$ , is invariant under a Lorentz boost in the  $z$ -direction [78]. The angular distance between two physics objects,  $\Delta R$ , can be written as,

$$\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} \quad (3.5)$$

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

$$p_T = \sqrt{(p_x)^2 + (p_y)^2} \quad (3.6)$$

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

$$E_T = \sqrt{m^2 + p_T^2} \quad (3.7)$$

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

### 3.2.2 Inner Detector

The inner detector is the first layer of concentric cylindrical sub-detector layers in the ATLAS detector. It is used to identify charged particles and reconstruct the trajectories of charged particles produced in the collisions via energy deposition in semiconductor material (hits) and the ionisation of gas. It consists of three complementary sub-detectors (in order from nearest to farthest from the beam pipe): the Pixel Detector, the Semiconductor Tracker (SCT) and the Transition Radiation Detector (TRT). The Pixel Detector and SCT are based on semiconductor technology and have the highest granularity of any sub-detector in ATLAS, in order to cope with the high frequency

337 of collisions near the interaction point. The TRT consists of drift tubes (straws) containing a gas mixture, which  
 338 allows measurement of the energy deposited by charged particles through the ionisation of the gas. Solenoid magnets  
 339 surround the inner detector and bend the trajectories of charged particles. The charges and transverse momenta  
 340 of particles can be inferred from their bent trajectories, which are reconstructed by the hits produced via energy  
 341 deposition in the Inner Detector.

### 342 3.2.3 Electromagnetic and Hadronic Calorimeters

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

### 353 3.2.4 Muon Spectrometer

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

### 359 3.2.5 Trigger and Data Acquisition System

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

### 366 3.2.6 Particle Identification and Object Reconstruction

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

#### 371 3.2.6.1 Tracks and primary vertices

372 The trajectories of charged particles, or tracks, are reconstructed in the ID. First, energy is deposited by charged  
 373 particles (hits) in pixels or strips, in the Pixel and SCT detectors respectively. Adjacent pixels or strips are grouped  
 374 together in *energy clusters*. Energy clusters define 3D space-points indicating the location where the charged  
 375 particle traversed. Track seeds are then defined as sets of three space-points, in either the Pixel or SCT detectors.  
 376 A Kalman filter [10] is then used to build track candidates from the track seeds. Often, multiple track candidates  
 377 are built per track seed, therefore an ambiguity solver [35] is needed for finding the track which best represents the  
 378 traversal of the charged particle. The ambiguity solver ranks each track from a given seed based on, the number of  
 379 associated hits, the number of holes (expected hits which are absent), track momenta and the  $\chi^2$  of the track fit.  
 380 Low ranked tracks are then discarded. High ranked tracks are refitted, introducing information from the TRT.

381  
 382 The primary vertex is the location of the  $pp$  collision of interest (i.e. from the hard scatter). The primary vertex  
 383 from the hard scatter needs to be identified, to isolate the event of interest from unwanted pile-up events. In the

384 event reconstruction procedure [62], the primary vertex is defined as the vertex of the event with the largest sum  
 385 of  $(p_T)^2$  (corresponding to the measured  $(p_T)^2$  of the particle from its reconstructed track) of its associated tracks.  
 386 Furthermore, the primary vertex is required to have at least two associated tracks. To reduce contamination from  
 387 fake tracks used in primary vertex reconstruction, only tracks which pass certain tight selection criteria are used  
 388 in the reconstruction procedure. An iterative fitting procedure is then used to reconstruct the primary vertex by  
 389 finding a set of reconstructed tracks which have a common vertex.

### 390 3.2.6.2 Electrons

391 Since electrons are charged particles, they give rise to tracks in the Inner Detector. They also deposit energy in the  
 392 ECAL via electromagnetic showering. Electrons are therefore reconstructed and identified from signals in the Inner  
 393 Detector and ECAL. Electrons are reconstructed using a dynamic clustering algorithm [1] which matches electron  
 394 candidate tracks in the Inner Detector to energy clusters in the ECAL. The dynamic clustering algorithm matches  
 395 tracks to energy clusters which have local maxima, to form electron candidates.

396 A likelihood discriminant is used to identify electrons. Quantities measured in the Inner Detector and ECAL  
 397 are used as input, such that they discriminate well between prompt isolated electrons and other physics objects  
 398 (e.g. jets, electron from a photon conversion, electron from a semi-leptonically decaying hadron). Important input  
 399 variables include the shape of the electromagnetic shower, track quality in the Inner Detector and information from  
 400 the TRT.

### 401 3.2.6.3 Muons

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

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

### 409 3.2.6.4 Jets and $b$ -tagging

410 Coloured particles emerging from the interaction point result in collimated streams of colourless particles, known  
 411 as jets. Jets can deposit energy in the Inner Detector and in the HCAL. Jets in ATLAS are reconstructed from  
 412 topological clusters using the anti- $k_t$  algorithm [28]. Topological clusters are groups of adjacent calorimeter cells  
 413 which contain energy deposition above the average amount of noise expected in the cell. Adjacent cells are grouped  
 414 together under certain criteria to form topological clusters which form jets.

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

422 In this analysis, we use the recommended DL1r (Deep-Learning Flavour Tagger) tagging algorithm [18]. The DL1r  
 423 algorithm combines outputs from several low-level tagging algorithms using a Deep Neural Network and outputs  
 424 the probability that a given input jet is identified as a  $b$ ,  $c$  or light flavoured jet.

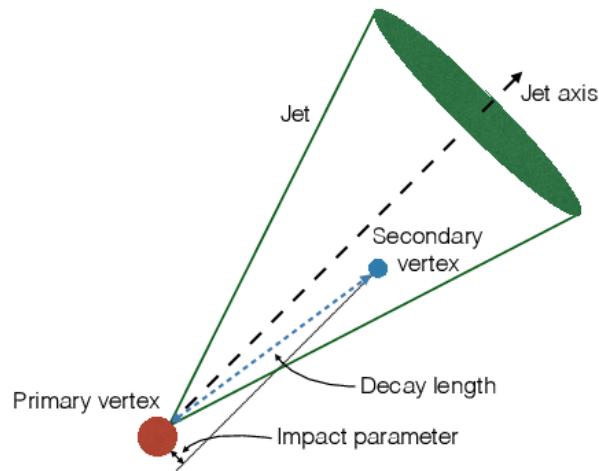


Figure 7: An illustration [38] of the production of a  $b$ -jet is shown. This illustrates the presence of a secondary vertex within a  $b$ -jet.

425

# Chapter 4

426

# Analysis Setup and Strategy

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

434 The  $pp$  collision data used in this analysis was collected with the ATLAS detector at the LHC from 2015 to 2018. This  
 435 data period of data taking is referred to as Run 2. During this period,  $pp$  collisions at  $\sqrt{s} = 13 \text{ TeV}$ , corresponding  
 436 to an integrated luminosity ( $\mathcal{L}$ ) of  $156 \text{ fb}^{-1}$ , were delivered by the LHC. The ATLAS detector managed to record  
 437  $147 \text{ fb}^{-1}$  of this total delivered data.  $139 \text{ fb}^{-1}$  of the data recorded by ATLAS is considered to be good enough  
 438 for physics analyses (the data passes certain quality control criteria) and placed into the *Good Runs List* [47].  
 439 Specific time-intervals where the luminosity is assumed to be constant are known as *lumi-blocks*. Only events from  
 440 lumi-blocks in which LHC beams were stable and all ATLAS detectors were operational are selected. The list of  
 441 suitable run and lumi-blocks is summarised in the official Good Runs Lists for 2015-2018 data, as specified in  
 442 Ref. [goodrunlist]. The integrated luminosities corresponding to the individual datasets for years 2015, 2016,  
 443 2017 and 2018 are  $3.2 \text{ fb}^{-1}$ ,  $33.0 \text{ fb}^{-1}$ ,  $44.3 \text{ fb}^{-1}$  and  $58.5 \text{ fb}^{-1}$  respectively.  
 444

### 4.1.2 Monte Carlo Samples

445 Simulated MC samples were generated and used to model the SM  $tWZ$  signal and its backgrounds.  
 446

447 The following background processes are considered:

- 448 •  **$t\bar{t}Z$** :  $t\bar{t}$  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.
- 449 •  **$ZZ$** : Diboson production with a tetralepton final state, therefore both  $Z$ -bosons decay leptonically ( $Z \rightarrow \ell^\pm \ell^\mp$   
   (OSSF lepton pair)).
- 450 • **other**: Processes with a relatively minimal, but non-negligible background contribution
  - 451 -  $VVV$  ( $V = W/Z$ )
  - 452 -  $t\bar{t}$
  - 453 -  $t\bar{t}W$
  - 454 -  $t\bar{t}WW$

- 460        -  $t\bar{t}H$   
 461        -  $WZ$   
 462        -  $t\bar{t}t$   
 463        -  $t\bar{t}\bar{t}\bar{t}$   
 464        -  $tZq$

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

469  
 470 The production of  $tWZ$  events is simulated with the **MADGRAPH5\_AMC@NLO 2.3.3** generator providing matrix element  
 471 (ME) calculations at NLO (next-to-leading order) in QCD. The events are interfaced with **PYTHIA 8.235** for the  
 472 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** generator  
 473 providing ME calculations at NLO in QCD. The events are interfaced with **PYTHIA 8.210** for the parton shower.  
 474 Event generation of  $tWZ$  and  $t\bar{t}Z$  results in diagrams which overlap with one another, that is, these diagrams contain  
 475 the same initial and final state particles. Several methods exist in order to separate between the two processes,  
 476 by removing the overlap, therefore avoiding double counting. There are two different diagram removal procedures,  
 477 diagram removal procedure 1 (DR1) [40] and diagram removal procedure 2 (DR2). The DR1 scheme is used to  
 478 remove the overlap (interference) between  $tWZ$  and  $t\bar{t}Z$ . A comparison to the DR2 scheme is used to estimate  
 479 part of the theoretical systematic on the modelling of the  $tWZ$  signal (See Section 4.8.2). Diboson processes which  
 480 feature the three charged leptons and one neutrino or four charged lepton in their final states, such as  $WZ$  and  
 481  $ZZ$ , are simulated using **SHERPA 2.2.2** at NLO in QCD precision. The events are interfaced with **SHERPA** for the  
 482 parton shower. Triboson processes such as  $WWW$ ,  $WWZ$ ,  $WZZ$ , and  $ZZZ$  containing up to six leptons in their  
 483 final states are simulated using **SHERPA 2.2.2** at NLO in QCD precision. The events are interfaced with **SHERPA** for  
 484 the parton shower. The production of  $t\bar{t}$  events are simulated with the **POWHEG** generator providing ME calculations  
 485 at NLO in QCD. The events are interfaced with **PYTHIA 8.210** for the parton shower. The production of  $t\bar{t}t$ ,  $t\bar{t}\bar{t}$   
 486 and  $t\bar{t}WW$  are simulated using the **MADGRAPH5\_AMC@NLO 2.2.2** generator at LO in QCD precision. The events are  
 487 interfaced with **PYTHIA 8.186** for the parton shower. The production of  $t\bar{t}$  with an associated Higgs boson,  $t\bar{t}H$ , are  
 488 generated using the **MADGRAPH5\_AMC@NLO 2.6.0** generator at NLO in QCD precision. The events are showered using  
 489 **PYTHIA 8.230**. The production of  $t\bar{t}$  events are simulated with the **POWHEG** generator providing ME calculations at  
 490 NLO in QCD. The events are showered using **PYTHIA 8.230**. The production of a single top quark in association  
 491 with a  $Z$ -boson and an extra parton,  $tZq$ , is simulated using **MADGRAPH5\_AMC@NLO 2.3.3** at NLO in QCD precision.  
 492 The events are interfaced with **PYTHIA 8.230** for the parton shower. In Table 3, the event generator and parton  
 493 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>
$t\bar{t}t, t\bar{t}\bar{t}$	<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.

## 4.2 Objects

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

### 4.2.1 Leptons

In this analysis only  $e$  and  $\mu$  leptons are considered, since  $\tau$  leptons are difficult to detect in the ATLAS detector.  $\tau$  leptons are challenging to detect since they have an extremely short lifetime ( $290.3 \pm 0.5$  fs [48]) which causes them to decay before reaching any detector components and therefore can only be reconstructed via their decay products. In addition to our selection criteria of exactly four leptons, it is required that the Leading (L), Next-to-Leading (NL), Next-to-Next-to-Leading (NNL) and Next-to-Next-to-Next-to-Leading (NNNL) leptons have  $p_T$  greater than 28, 18, 10 and 10 GeV respectively. Relatively loose object-level cuts are chosen in an attempt to maximize our signal statistics, since the analysis is heavily statistically limited. Reconstructed electrons are required to be within  $|\eta| < 2.47$  and excluding the transition region between the barrel and end-cap calorimeters at  $1.37 < |\eta| < 1.52$ . Reconstructed muons are required to be within  $|\eta| < 2.5$ . The transverse impact parameter,  $d_0$ , is defined as the minimal spatial distance between the object's (referring to leptons) trajectory and the primary vertex (the vertex associated with the  $p$ - $p$  hard scatter). The longitudinal impact parameter,  $z_0$ , is defined as the value of  $z$  of the point on the object's trajectory which determines  $d_0$ . To ensure consistency between the lepton and the primary vertex, it is required that  $|\frac{d_0}{\sigma(d_0)}| < 5$ ,  $|z_0 \sin \theta| < 0.5$  mm for electrons and  $|\frac{d_0}{\sigma(d_0)}| < 3$ ,  $|z_0 \sin \theta| < 0.5$  mm for muons, following the current recommendations [76]. To avoid instances where one detector signal can result in multiple different reconstructed objects, an overlap removal is applied which ignores all but one of these objects (See Section 4.2.4). Electrons are selected using a likelihood based discriminant [1] which takes measurements from the tracking system, calorimeter system and quantities derived from both the tracking and calorimeter system as input. Muons are selected using the Muon Selection Tool [66]. Loose electrons are defined with the criteria above, using the `LooseAndBLayerLH` ( $\sim 91\%$  selection efficiency for electrons with  $E_T > 30$  GeV [41]) identification algorithm (which has a certain cut applied). Similarly, tight electrons are defined with the criteria above, using the `TightLH` ( $\sim 80\%$  selection efficiency for electrons with  $E_T > 30$  GeV [41]) algorithm (which has a certain cut applied). Both loose and tight muons use the `Medium` ( $\sim 95\%$  selection efficiency [5]) algorithm (which has a certain cut applied). Tight leptons additionally require that they are sufficiently isolated from other particles produced in the collision. This is done by defining a cone of radius  $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$  around the particle of interest and summing the  $p_T$  of all the reconstructed particles surrounding the particle of interest, situated within the cone. 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 candidate. If this value is large, it is likely that the particle of interest originated from a jet (together with many other particles), whereas a prompt decay product resulting from the hard scatter will have little to no energy surrounding it ( $I_{rel} \ll 1$ ). The `IsolationSelectionTool` with the `PLVTight` ( $\sim 70\%$  efficiency [7]) and `PLVTight` ( $\sim 70\%$  efficiency at  $p_T = 30$  GeV [34]) algorithm are used for tight electrons and tight muons respectively (following the current recommendations [70]). In Table 4, a summary of the selection criteria for leptons is shown.

	Electrons		Muons	
	Tight	Loose	Tight	Loose
$p_T$ cuts		$p_T(\ell_1, \ell_2, \ell_3, \ell_4) > (28, 18, 10, 10)$ 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$ mm		$ \frac{d_0}{\sigma(d_0)}  < 3$ , $ z_0 \sin \theta  < 0.5$ 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.

### 4.2.2 Jets

Jets are reconstructed using the anti- $k_t$  algorithm (See Section 3.2.6.4). The `AntiKt4EMPFflowjets` ( $\sim 97\%$  average efficiency with JVT (outlined in the subsequent paragraph)  $> 0.2$  [76]) algorithm (which has a certain cut applied) is used, following the current recommendations [76]. The jet-vertex-tagger (JVT) and the forward jet-vertex-tagger (fJVT) are likelihood discriminant which aim to suppress pile-up jets. The `Medium` algorithm (which has a certain cut applied) is used for the JVT and the fJVT (following the current recommendations [69]). Additionally, a requirement that jets have a JVT value greater than 0.5 is applied. In the same way as with leptons, ambiguities are removed where one detector signal can result in multiple different reconstructed objects, via overlap removal

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

### 539 4.2.3 *b*-tagging

540 The DL1r *b*-tagger [65] was used to identify jets as *b*-jets (See Section 3.2.6.4). Different DL1r working points are  
 541 used to identify *b*-jets in our event selection (See Section 4.4). The working points are defined based off a cut on  
 542 the DL1r score corresponding to a *b*-jet tagging efficiency of 60%, 70%, 77% and 85%.

543 Since this analysis is heavily statistically limited, the amount of statistics in our regions are aimed to be maximized.  
 544 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  
 545 tight *b*-tagged jet is defined as a jet which passes the 77%, 70%, 65% or 60% DL1r *b*-tagger working point. A loose  
 546 *b*-tagged jet is defined as a jet which passes 85% DL1r *b*-tagger working point, but not the 77%, 70%, 65% or 60%  
 547 DL1r *b*-tagger working points. Different numbers (and definitions) of tight and loose *b*-tagged jets were tried in  
 548 each region, with the final selection criteria being chosen which maximised the expected significance of  $\sigma(tWZ)$   
 549 (See Section 4.4.1).

### 551 4.2.4 Overlap Removal Procedure

552 The overlap removal procedure is used on pre-selected leptons and jets. It is performed sequentially, in the following  
 553 steps:

- 554 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.
- 555 2. If the separation between a jet and a pre-selected electron is within  $\Delta R < 0.2$ , the jet is discarded.
- 556 3. Any remaining electron or muon closer than  $\Delta R = 0.4$  to a jet, is discarded.
- 557 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.

## 560 4.3 Kinematic cuts

561 In order to suppress potential fakes and quarkonia (low mass resonances such as  $J/\psi$  and upsilon) a requirement  
 562 that all OSSF lepton pairs have an invariant mass,  $m_{\text{OSSF}}$ , greater than 10 GeV is applied. The final state lepton  
 563 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  
 564 OSSF lepton pair coming from the  $Z$  boson must equal the invariant mass of the  $Z$  boson, and noting that  $e,\mu$   
 565 reconstruction and identification in the ATLAS detector has a high efficiency [58], these OSSF leptons are used  
 566 to reconstruct  $Z$  bosons with relatively high confidence. A  $Z$  candidate is defined in this analysis as an OSSF  
 567 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  
 568 nominal  $Z$  boson mass (91.1876 GeV [48]). This wider mass window is used in order to cover the full range of the  
 569  $m(Z)$  distribution, in an attempt to increase the number of events which pass our baseline selections. Multiple  $Z$   
 570 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  
 571 has an invariant mass closest to the nominal  $Z$  boson mass is chosen.

## 572 4.4 Regions and Event Selection

573 Two  $tWZ$  SRs are defined in an attempt to suppress and constrain the  $ZZ$  background. Both  $tWZ$  SRs are  
 574 required to have exactly four tight leptons, exactly one  $Z$ -boson candidate, exactly one tight *b*-tagged jet (from  
 575 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  
 576 leptons which don't originate from the decay of a  $Z$ -boson (non- $Z$  leptons). The  $ZZ$  background has two  $Z$ -bosons  
 577 which decay into a pair of OSSF lepton pairs, in order to mimic the  $tWZ$  signal. This is taken advantage of, to  
 578 define a  $tWZ$  region enrich in  $ZZ$  background and one with a minimal  $ZZ$  background component. This is done  
 579 by requiring that one of the  $tWZ$  SRs has its two non- $Z$  leptons to have opposite flavour and the other  $tWZ$  SR

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	$(tWZ)_{\text{fake}}$ CR
$N_\ell(\text{tight}) = 4$	$N_\ell(\text{tight}) = 4$	$N_\ell(\text{tight}) = 4$	$N_\ell(\text{tight}) = 4$	$N_\ell(\text{tight}) = 3$ $N_\ell(\text{loose and NOT tight}) = 1$
$N_Z$ candidate = 1	$N_Z$ candidate = 1	$N_Z$ candidate = 1	$N_Z$ candidate = 2	$N_Z$ candidate = 1
$N_{\text{jet}} \geq 1$	$N_{\text{jet}} \geq 1$	$N_{\text{jet}} \geq 2$	$N_{\text{jet}} \geq 1$	$N_{\text{jet}} \geq 1$
$N_{\text{b-jet}}(\text{tight}) = 1$	$N_{\text{b-jet}}(\text{tight}) = 1$	$N_{\text{b-jet}}(\text{tight}) \geq 1$ $N_{\text{b-jet}}(\text{loose}) \geq 0$ $N_{\text{b-jet}}(\text{tight}) + N_{\text{b-jet}}(\text{loose}) = 2$	$N_{\text{b-jet}}(\text{tight}) = 1$	$N_{\text{b-jet}}(\text{tight}) = 1$
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.

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, 36], 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 ( $N_{\text{jet}} \geq 1$ ).

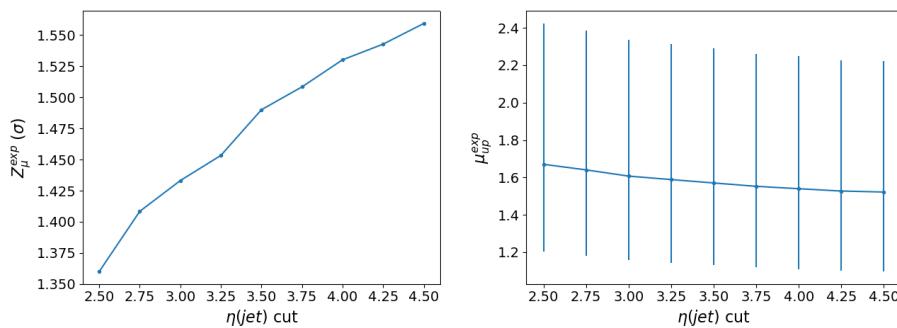
Fake leptons are objects reconstructed as leptons, but do not correspond to the leptons that are of interest in our analysis. Fake leptons can be split up into two main categories, irreducible (prompt) fakes and reducible (non-prompt) fakes. Irreducible fakes are true leptons which do not come from the 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  $(tWZ)_{\text{fake}}$  CR is defined which is as similar as possible to the  $tWZ$  CR but is enhanced in fakes. This is achieved by defining the  $(tWZ)_{\text{fake}}$  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  $(t\bar{t}Z)_{\text{fake}}$  CR (requiring exactly 3 tight leptons and exactly 1 loose (and NOT tight) lepton) was tried as an alternative to the  $(tWZ)_{\text{fake}}$  CR, however a much larger suppression of fakes were observed in this region (compared to the  $(tWZ)_{\text{fake}}$  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.

#### 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.

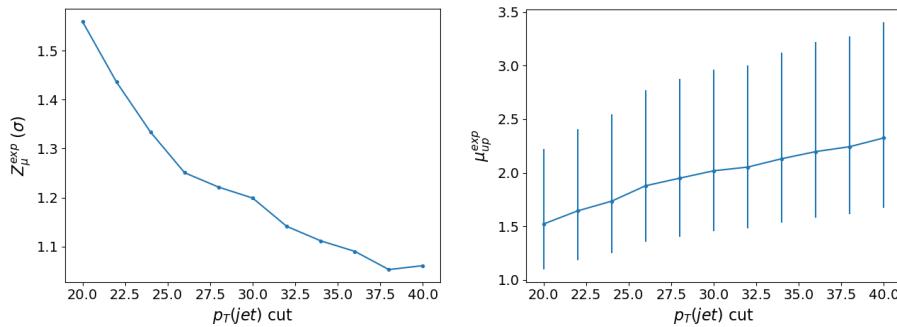
In Figure 8 the expected significance ( $Z_\mu^{exp}$ ) and expected upper limits ( $\mu_{up}^{exp}$ ) for different  $\eta(jet)$  cuts are shown.



**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 maximises 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 maximises 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. From Figure 10, it can be seen that requiring that  $b$ -tagged jets pass the 77% DL1r WP in the  $tWZ$  SR,  $(tWZ)_{fake}$  CR and the  $ZZb$  CR and that at least one  $b$ -tagged jet in the  $t\bar{t}Z$  SR passes the 77% DL1r WP (the other jet is

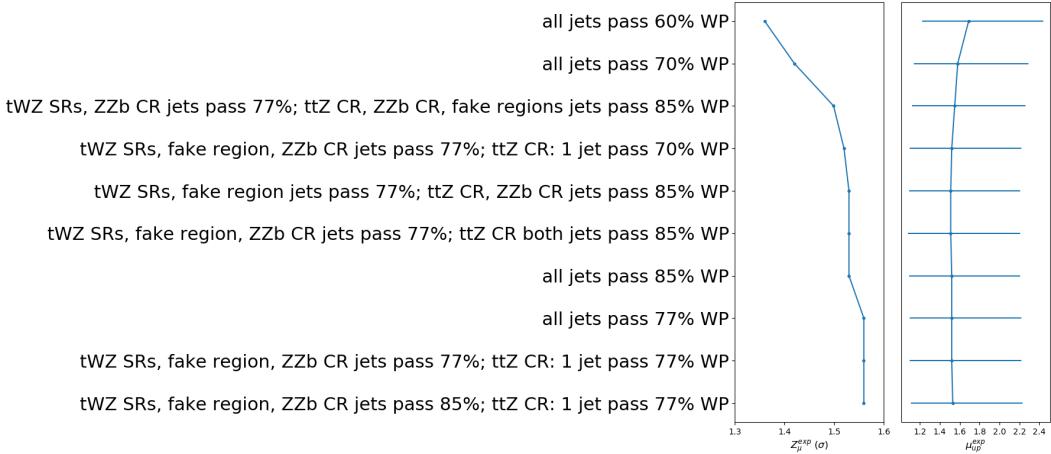


Figure 10: Expected significance ( $Z_\mu^{exp}$ ) and expected upper limit ( $\mu_{up}^{exp}$ ) for different configurations of DL1r  $b$ -tagged jet working points is shown. The common y-axis shows the different configurations of DL1r  $b$ -tagged jet working points. On the left panel, the expected 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).

630 just required to pass the 85% DL1r WP) maximises the sensitivity overall (compared to the other investigated  
631 configurations). This configuration was chosen  $b$ -tagged jets.

632 The  $p_T$ (L Lepton) is constrained by the single lepton triggers (Table ??). A cut was chosen to be applied on the  
633  $p_T$ (NL Lepton) slightly tighter than the tightest single lepton  $p_T$  cut in the trigger. The  $p_T$ (NL Lepton) cut can  
634 be optimized by comparing the expected significance and limit for a range of  $p_T$ (NL Lepton) cuts to determine the  
635 cut which maximizes sensitivity.

636 In Figure 11 the expected significance ( $Z_\mu^{exp}$ ) and expected upper limits ( $\mu_{up}^{exp}$ ) for different  $p_T$ (NL Lepton) cuts is  
637 shown.

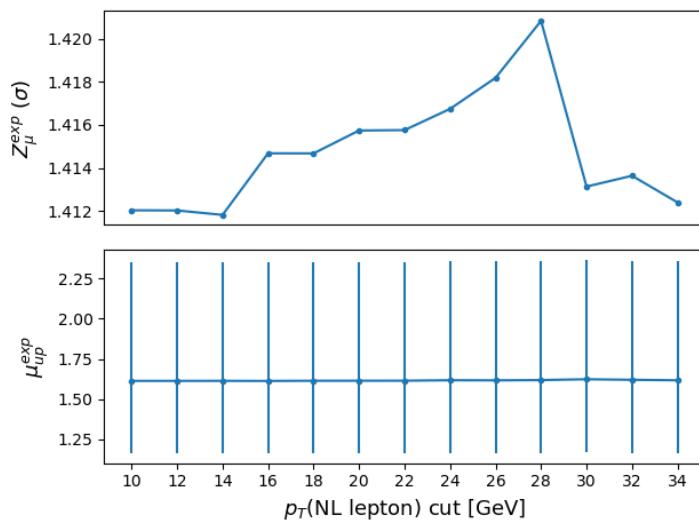


Figure 11: Expected significance ( $Z_\mu^{exp}$ ) and expected upper limit ( $\mu_{up}^{exp}$ ) for different  $p_T$ (NL Lepton) cuts is shown. The common x-axis shows cut applied to the  $p_T$  of the next-to-leading lepton. On the top panel, the expected significance from the likelihood fit is shown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

640 Since there is a very small change between the different  $p_T$ (NL Lepton) cuts on the sensitivity of  $tWZ$ , a  
 641  $p_T$ (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  
 646 section, the data is given by the black points and the MC predictions for each process are given by the filled  
 647 histograms. The vertical lines on the data points represent the statistical uncertainty in the data and the  
 648 diagonally-lined bands represent the total (statistical and systematic added in quadrature) uncertainty. The  
 649 lower panel in each plot shows the ratios of the data to the theoretical predictions. In order to suppress a bias  
 650 towards large signal observations in the development of the analysis, data has not been analysed in bins where the  
 651 expected  $\frac{\text{signal}}{\text{background}}$  exceeds 0.1. Blinded bins are shaded with black diagonal lines and their data points are omitted.  
 652

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	$(tWZ)_{\text{fake}}$ 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	$0.0687541 \pm 0.0482172$	$0.032827 \pm 0.026286$	$0.0709734 \pm 0.043509$	$0.0474576 \pm 0.0301512$	$4.94775 \pm 2.48939$
$tWZ$	$3.81359 \pm 0.392241$	$2.57584 \pm 0.326401$	$2.61991 \pm 0.861557$	$1.4023 \pm 0.156686$	$4.93485 \pm 0.692143$
$ZZ$	$0.546045 \pm 0.18975$	$8.76232 \pm 2.66871$	$1.22357 \pm 0.376889$	$46.0616 \pm 13.9203$	$7.76724 \pm 2.36894$
other	$t\bar{t}$	$6e-06 \pm 3.04506e-06$	$0.250783 \pm 0.44226$	$0.269883 \pm 0.223373$	$6e-06 \pm 3.04506e-06$
	$tZq$	$0.0827265 \pm 0.0399222$	$0.0757694 \pm 0.0355101$	$0.0637132 \pm 0.0293762$	$0.0590199 \pm 0.0244576$
	$t\bar{t}W$	$0.00674747 \pm 0.00793546$	$0.00279491 \pm 0.00287747$	$6e-06 \pm 3.04506e-06$	$0.00221727 \pm 0.00562041$
	$WZ$	$0.0439316 \pm 0.0241635$	$0.0397876 \pm 0.0154764$	$0.0134837 \pm 0.0128327$	$0.0474188 \pm 0.0330635$
	$t\bar{t}t$	$0.000987429 \pm 0.000768187$	$0.00249801 \pm 0.00138007$	$0.0141085 \pm 0.00486102$	$6e-06 \pm 3.04506e-06$
	$t\bar{t}\bar{t}$	$0.00934516 \pm 0.0080725$	$0.0107503 \pm 0.00852049$	$0.0570846 \pm 0.0206271$	$6e-06 \pm 3.04506e-06$
	$t\bar{t}WW$	$0.0293456 \pm 0.0263573$	$0.0296011 \pm 0.0196075$	$0.26412 \pm 0.0936908$	$0.013096 \pm 0.0323943$
	$VVV (V = W/Z)$	$0.280384 \pm 0.0866421$	$0.191257 \pm 0.0595588$	$0.0696624 \pm 0.0228108$	$0.171171 \pm 0.0526519$
	$t\bar{t}H$	$0.854064 \pm 0.177974$	$0.674566 \pm 0.141771$	$1.98187 \pm 0.406211$	$0.151447 \pm 0.0357703$
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.

653 expected to be observed in data (MC simulation) carries an associated statistical uncertainty. To first order, this  
 654 uncertainty can be written as the square root of the expected number of events to be observed in data. In contrast  
 655 to this, predictions based on MC simulation carry uncertainties due to the finite number of simulated events utilised.  
 656 This uncertainty can be quantified by the Number of Equivalent Events [43],  $N_{\text{equiv}}$ , which relates the sample of  
 657  $N$  events (weighted by MC event weights) to  $N_{\text{equiv}}$  events with all MC event weights equal to 1, that would have  
 658 the same relative statistical fluctuation. The Number of Equivalent Events,  $N_{\text{equiv}}$ , can be written as,  
 659

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

660 where  $w_i$  is the MC event weight for event  $i$ . The standard uncertainty of  $N_{\text{equiv}}$  is given by  $u(N_{\text{equiv}}) = \sqrt{N_{\text{equiv}}}$ .  
 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,  $N_{\text{equiv}}$  and its percentage uncertainty ( $\frac{u(N_{\text{equiv}})}{N_{\text{equiv}}} \times 100$ ), is shown for each sample in each region.  
 665

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

### 669 4.5.1 $tWZ$ OF SR

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

671

	$tWZ$ 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.

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

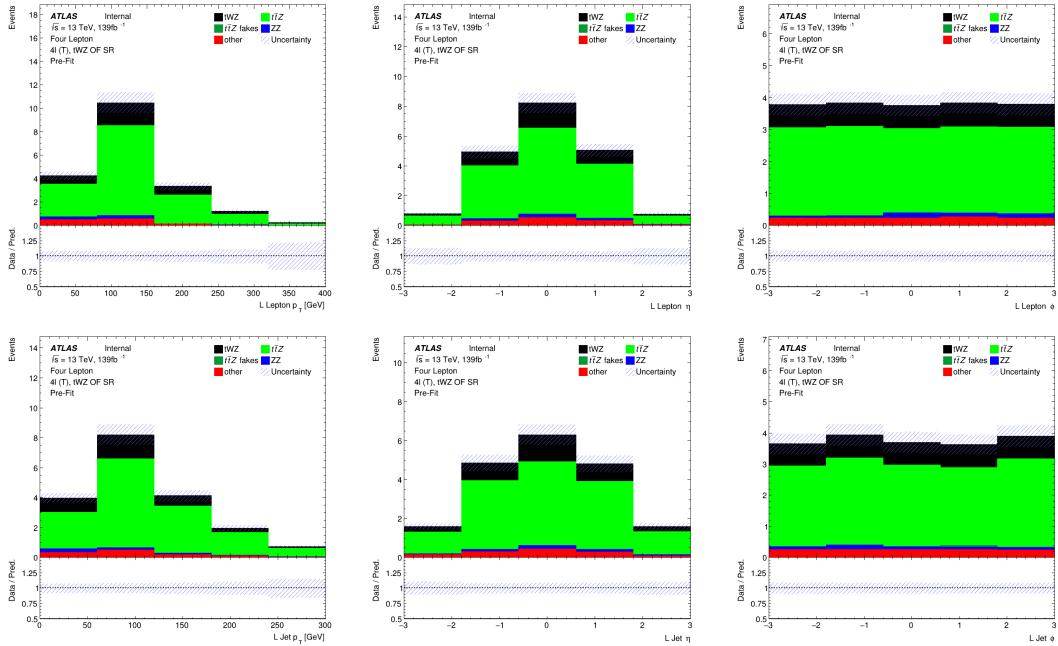


Figure 12: Comparisons of simulation and data of  $p_T$ ,  $\eta$  and  $\phi$  for leading (L) leptons (top row) and leading (NL) jets (bottom row) in the  $tWZ$  OF SR are shown.

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

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

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

#### 681 4.5.2 $tWZ$ SF SR

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

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

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

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

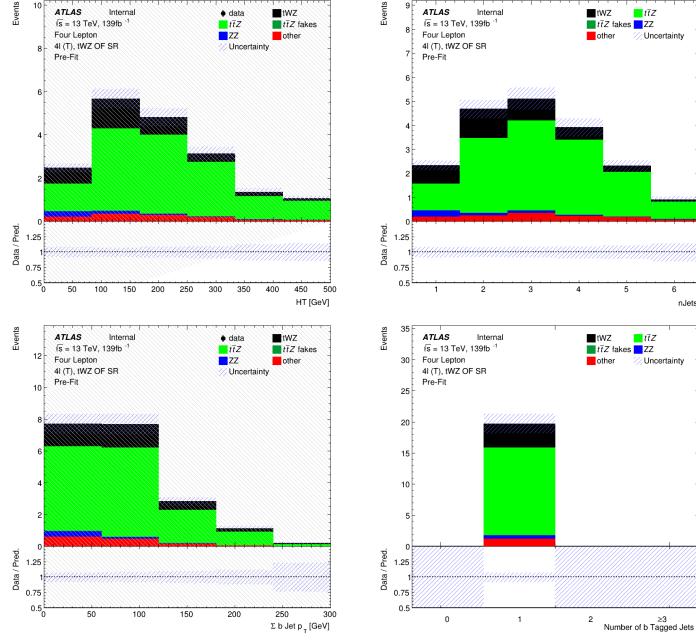


Figure 13: Comparisons of simulation and data of  $H_T$  (scalar sum of Jet  $p_T$ ), the Number of jets, the scalar sum of  $b$ -tagged jet  $p_T$  and the number of  $b$ -tagged jets (top left to bottom right) in the  $tWZ$  OF SR are shown.

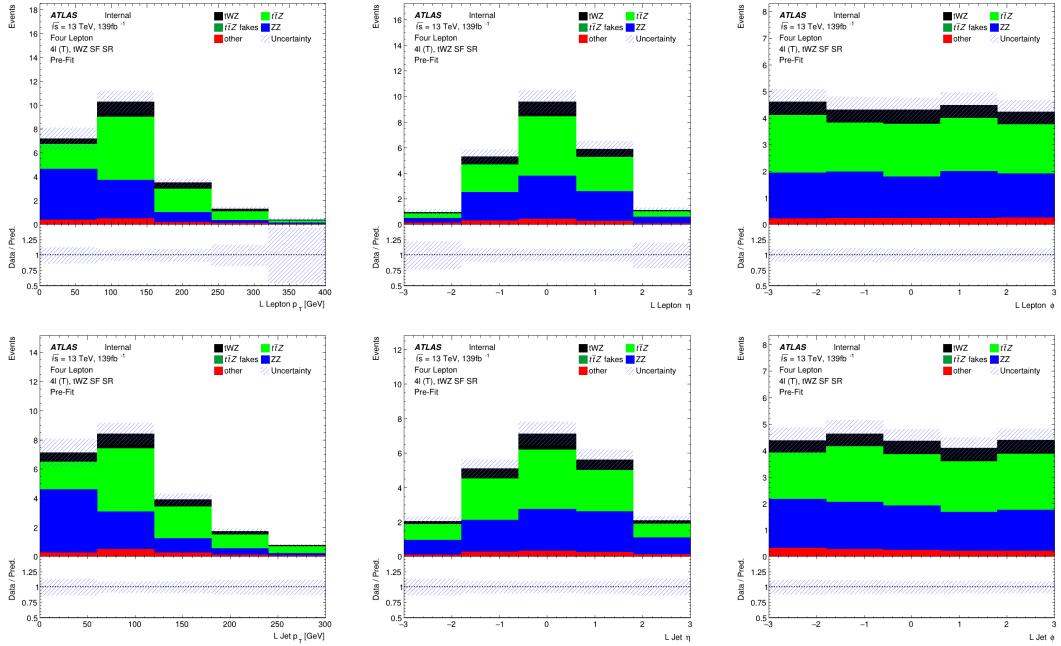


Figure 14: Comparisons of simulation and data of  $p_T$ ,  $\eta$  and  $\phi$  for leading (L) leptons (top row) and leading (NL) jets (bottom row) in the  $tWZ$  SF SR are shown.

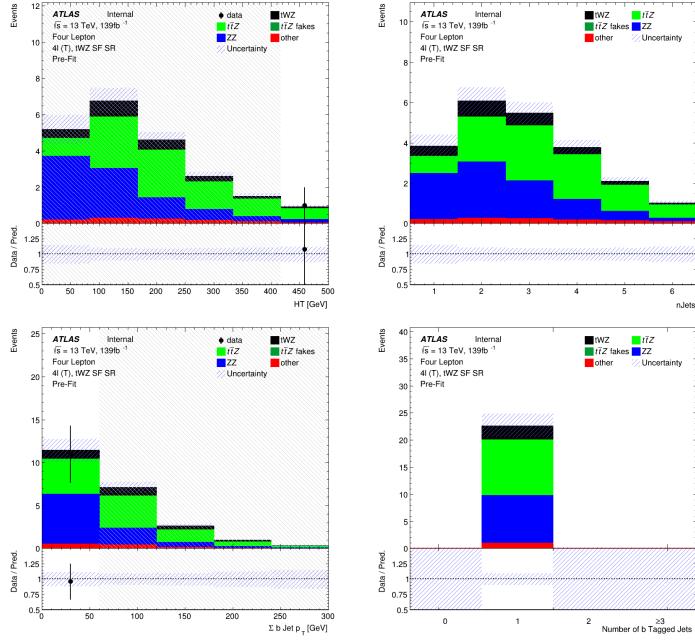


Figure 15: Comparisons of simulation and data of  $H_T$  (scalar sum of Jet  $p_T$ ), the Number of jets, the scalar sum of  $b$ -tagged jet  $p_T$  and the number of  $b$ -tagged jets (top left to bottom right) in the  $tWZ$  SF SR are shown.

691 The vast majority of bins for each plot in Figure 15 have  $\frac{\text{signal}}{\text{background}}$  exceeding 0.1 and are therefore blinded. This  
 692 region is therefore enriched in  $tWZ$  signal events. The deviations in data and simulation in the two bins (in the  
 693  $HT$  and  $\sigma b$  jet  $p_T$  distributions) which are not blinded, are within the expected uncertainties.

#### 694 4.5.3 $t\bar{t}Z$ CR

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

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

698 The majority of the deviations in data and simulation for each plot in Figure 16 are within the expected  
 699 uncertainties. The few plots which have bins where there is a disagreement between data and simulation are either  
 700 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  
 701  $p_T$ ) disagreement. The disagreement in the L Lepton  $p_T$  distribution could be due to statistical fluctuations in  
 702 data or simulation, since there are so few events in these bins.

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

706 Almost all of the deviations in data and simulation for each plot in Figure 17 are within the expected uncertainties.  
 707 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  
 708 of the bins in the  $HT$  distribution. The large disagreement between data and simulation in the  $HT$  distribution  
 709 is surprising since all other bins in the distribution agree within  $1\sigma$  uncertainties, and it is therefore not fully  
 710 understood.

#### 712 4.5.4 $ZZb$ CR

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

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

716 Most of the deviations in data and simulation for each plot in Figure 18 are within the expected uncertainties.  
 717 There are a few bins with  $2\sigma$  and  $> 2\sigma$  disagreements between data and simulation in the L Lepton  $p_T$ , L Lepton

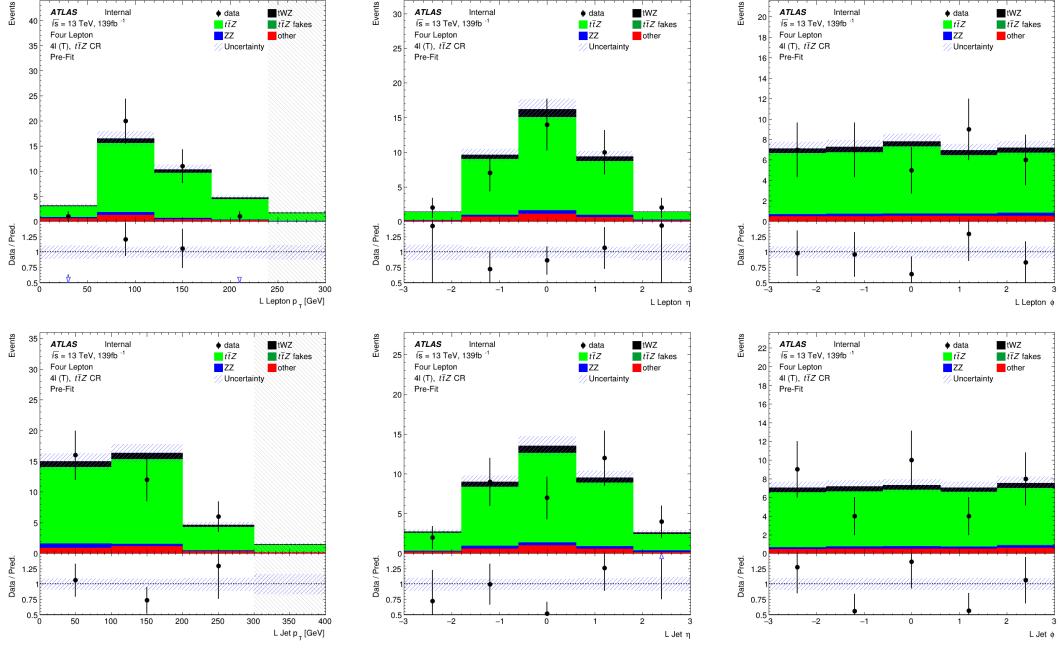


Figure 16: Comparisons of simulation and data of  $p_T$ ,  $\eta$  and  $\phi$  for leading (L) leptons (top row) and leading (NL) jets (bottom row) in the  $t\bar{t}Z$  CR are shown.

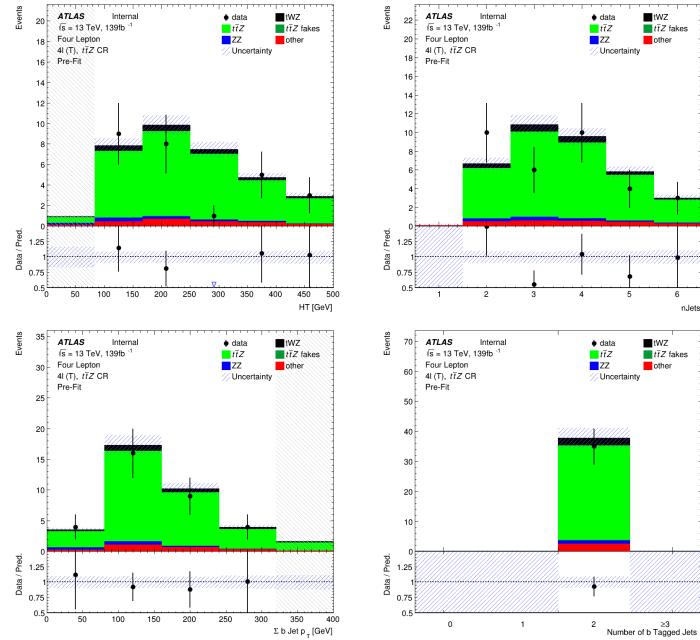


Figure 17: Comparisons of simulation and data of  $H_T$  (scalar sum of Jet  $p_T$ ), the Number of jets, the scalar sum of  $b$ -tagged jet  $p_T$  and the number of  $b$ -tagged jets (top left to bottom right) in the  $t\bar{t}Z$  CR are shown.

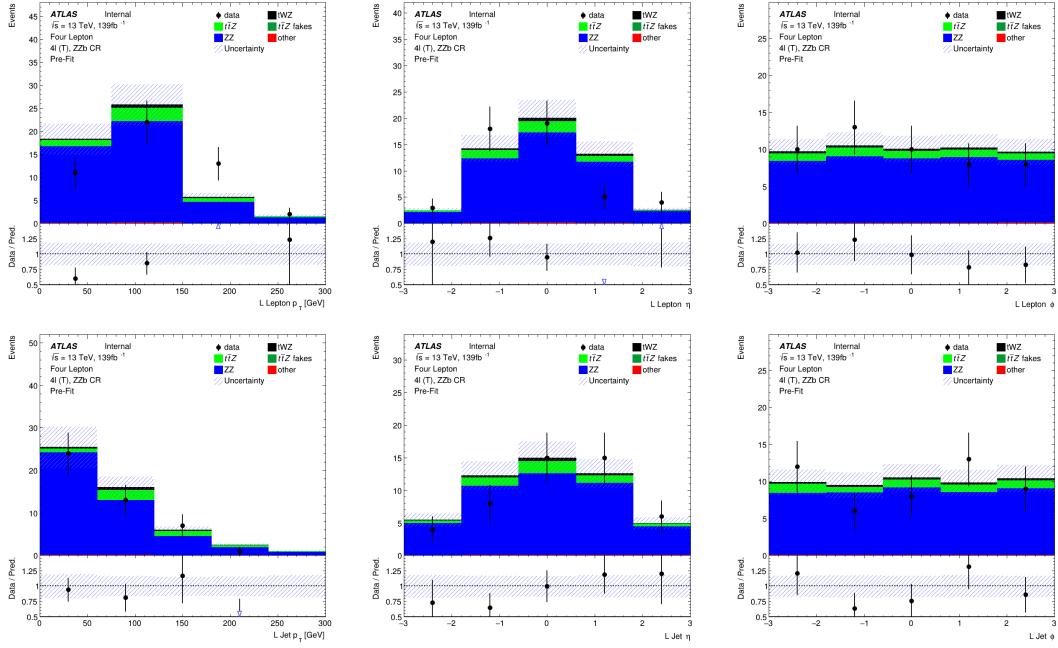


Figure 18: Comparisons of simulation and data of  $p_T$ ,  $\eta$  and  $\phi$  for leading (L) leptons (top row) and leading (NL) jets (bottom row) in the  $ZZb$  CR are shown.

719  $\eta$  and L Jet  $p_T$  distributions, with the disagreement being much more noticeable in the L Lepton distributions.  
720 This could suggest some mis-modelling for L Leptons in this region.

721 In Figure 19, comparisons of simulation and data of  $H_T$  (scalar sum of Jet  $p_T$ ), the Number of jets, the scalar sum  
722 of  $b$ -tagged jet  $p_T$  and the number of  $b$ -tagged jets in the  $ZZb$  CR are shown.  
723

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

#### 725 4.5.5 $(tWZ)_{\text{fake}}$ CR

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

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

729 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  
730 region is therefore enriched in  $tWZ$  signal events. Most deviations in data and simulation in the bins which are  
731 not blinded, are within the expected uncertainties. Only two out of seven unblinded bins are not within expected  
732 uncertainties and are within a  $2\sigma$  uncertainty.

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

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

## 741 4.6 Fake Lepton Estimation

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

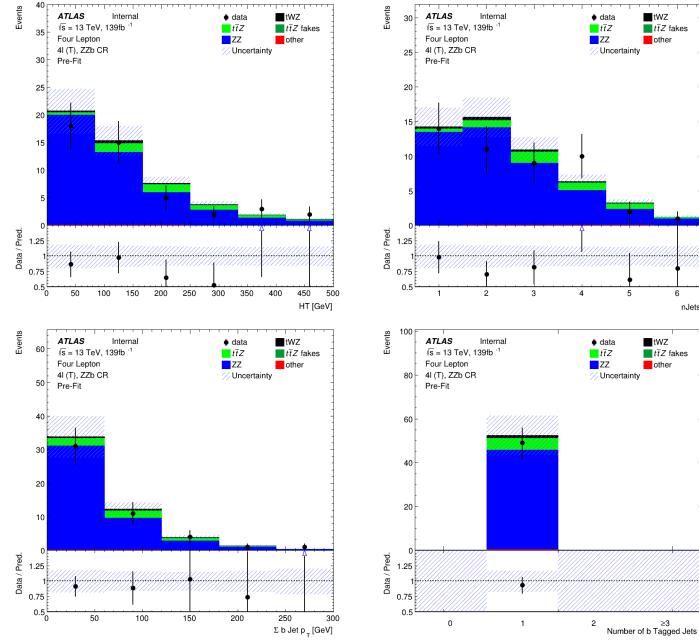


Figure 19: Comparisons of simulation and data of  $H_T$  (scalar sum of Jet  $p_T$ ), the Number of jets, the scalar sum of  $b$ -tagged jet  $p_T$  and the number of  $b$ -tagged jets (top left to bottom right) in the  $ZZb$  CR are shown.

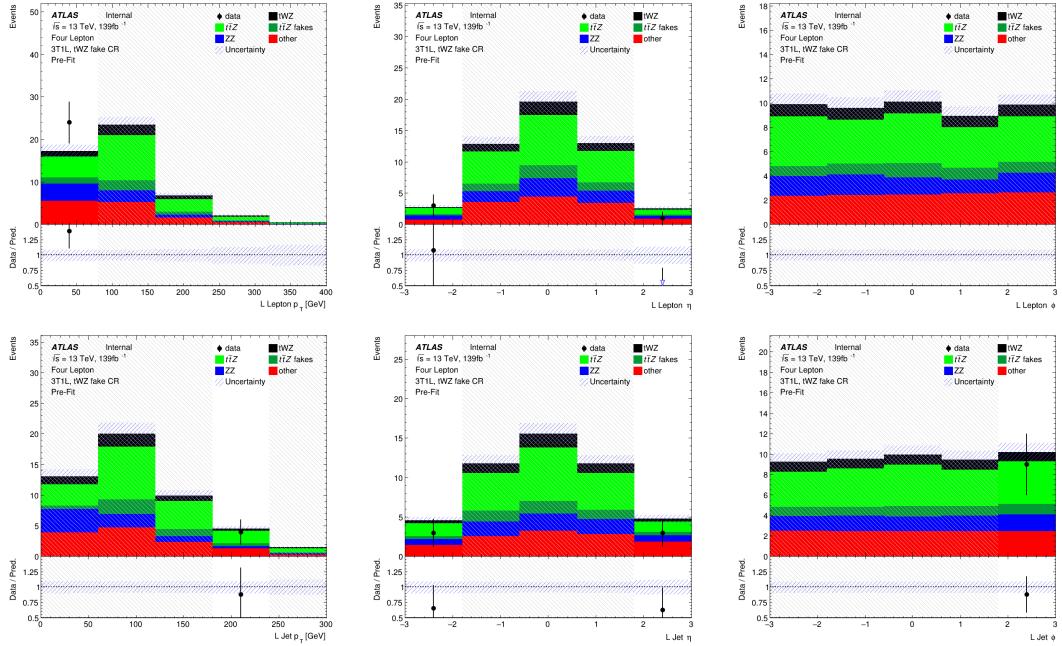


Figure 20: Comparisons of simulation and data of  $p_T$ ,  $\eta$  and  $\phi$  for leading (L) leptons (top row) and leading (NL) jets (bottom row) in the  $(tWZ)_{\text{fake}}$  CR are shown.

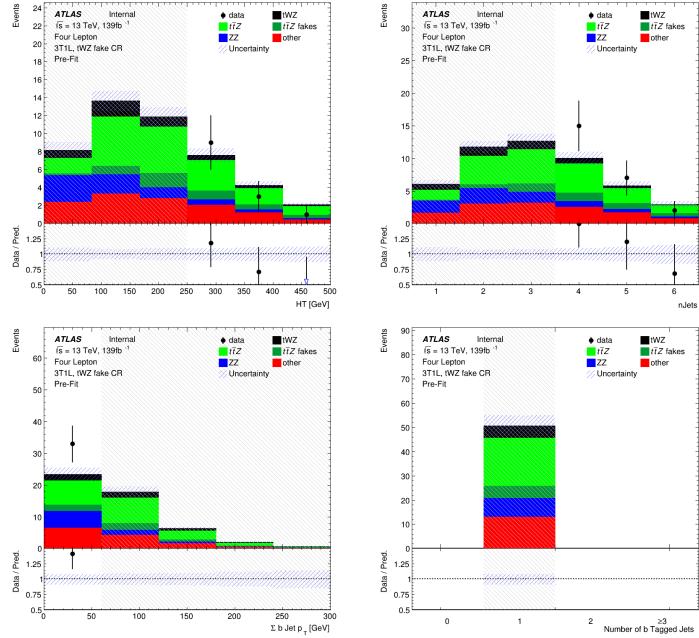


Figure 21: Comparisons of simulation and data of  $H_T$  (scalar sum of Jet  $p_T$ ), the Number of jets, the scalar sum of  $b$ -tagged jet  $p_T$  and the number of  $b$ -tagged jets (top left to bottom right) in the  $(tWZ)_{\text{fake}}$  CR are shown.

As  $t\bar{t}Z$  is the dominant background process ( $\sim 75\%$  of the total background contribution), it is assumed that  $t\bar{t}Z$  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  $\epsilon_1 \ll 1$  [lesage2017lepton, ATLAS-CONF-2012-048] 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  $t\bar{t}Z$  event is considered.

Firstly, the dominant  $t\bar{t}Z$  background is split up into  $t\bar{t}Z$  and  $(t\bar{t}Z)_{\text{fake}}$  components. Secondly, a  $(tWZ)_{\text{fake}}$  CR (See Section 4.4) is defined which is enhanced in fakes and aims to constrain the  $(t\bar{t}Z)_{\text{fake}}$  background in the SR. All events which contribute to the  $(t\bar{t}Z)_{\text{fake}}$  background are determined by the IFF Truth Classifier [51]. The IFF Truth Classifier is a tool which aims to classify leptons based off their truth information. It uses the more general MCTruthClassifier [61] 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 [52]). 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  $(t\bar{t}Z)_{\text{fake}}$  background are those where at least one lepton from the  $t\bar{t}Z$  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  $(t\bar{t}Z)_{\text{fake}}$  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  $(t\bar{t}Z)_{\text{fake}}$  template can be constrained.

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

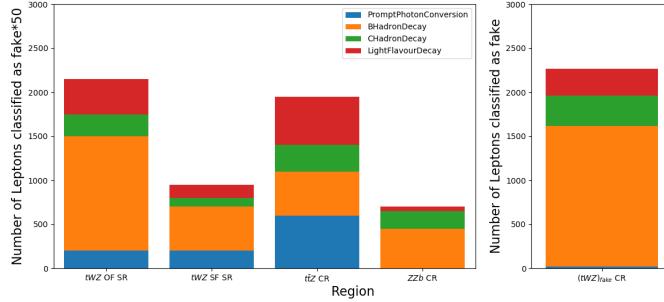


Figure 22: The number of leptons classified as fake, split up by their IFF Truth classification, in each region is shown. The left panel shows the number of leptons classified as fakes, scaled by a factor of 50, on the y-axis. The right panel shows the number of leptons classified as fakes (unscaled), 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.

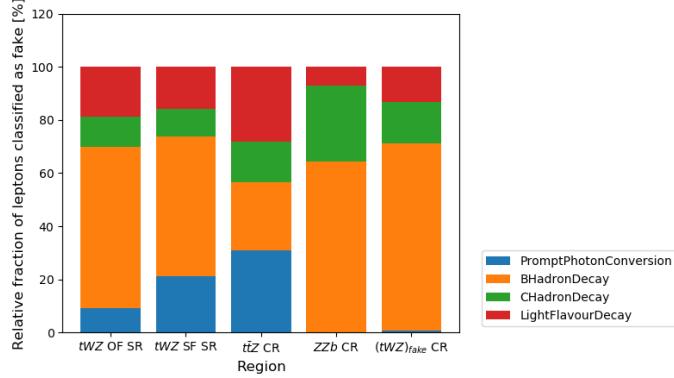


Figure 23: The relative dominance of the different classifications for fake leptons (classified by the IFF truth classified) in each region, is shown. The relative dominance of leptons classified as fakes, as a fraction of the total number of fake leptons (in each region), is shown on the y-axis. The 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  $ttZ$  CR. The smaller proportion of fake leptons originating from  $b$ -hadron decays in the  $ttZ$  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,  $ttZ$  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.

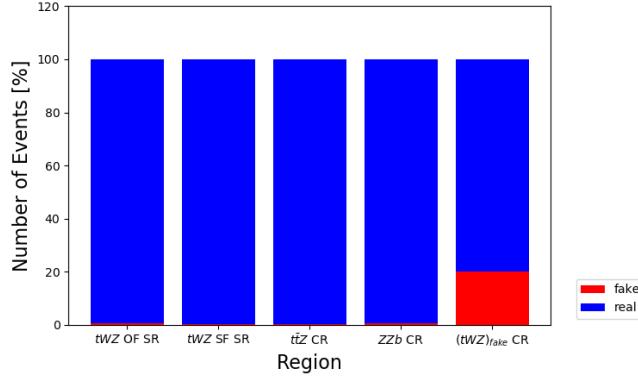


Figure 24: The percentage of fake and real  $t\bar{Z}$  events which pass our selection criteria, in each region, is shown. The relative number of fake and real events (in % of the total number of events in the nominal and fake  $t\bar{Z}$  background samples) is shown on the y-axis. The different signal 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  $t\bar{Z}$  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,  $t\bar{Z}$ . This information is aimed to be exploited by reconstructing  $\ell b$  systems as a proxy for top quarks (since,  
 795  $t \rightarrow W(\rightarrow \ell\nu)b$ ). This is done in two ways, firstly by implementing a Boosted Decision Tree (BDT) which is used  
 796 to classify  $\ell b$  systems originating from top quarks, and secondly, by implementation of a kinematic reconstruction  
 797 algorithm (Two Neutrino Scanning Method) which aims to determine the likelihood of an event containing two top  
 798 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  
 801 algorithm which takes advantage of these discriminating variables. A BDT is implemented and is used to  
 802 discriminate between  $tWZ$  events and its major backgrounds,  $t\bar{Z}$  and  $ZZ$ . In this thesis, this BDT is referred to  
 803 as an *event-level* BDT. The discriminator output from the object-level BDT can be converted to a variable which  
 804 can then be used as an input to the event-level BDT.

805 BDTs are chosen, as opposed to another ML algorithm, since they are not very sensitive to overtraining  
 806 and perform well with minimal/no optimisation or tweaking of the hyper-parameters. A multi-layered se-  
 807 quential neutral network was tried, however, it was out-performed by a BDT. More specifically, Scikit-Learn's  
 808 `GradientBoostingClassifier` [71] was used.

### 810 4.7.1 Boosted Decision Trees

811 Machine Learning techniques can be used to build multivariate algorithms that exploit information from many  
 812 weak discriminators (a model that is used to classify data in a dataset) to form one strong discriminator. A BDT  
 813 is a Machine Learning technique which classifies data in a dataset into different categories by iteratively applying  
 814 binary cuts on features of the data (variables, in the context of this analysis) to segregate the data. The method in  
 815 which a BDT combines weak discriminators to build a single strong discriminator is called *boosting*. In boosting,  
 816 weak discriminators are sequentially combined, where each model iteration is fitted to the residuals from the  
 817 previous step, such that the model performance improves.

818 There are a few concepts related to Machine Learning and BDTs that are used in this analysis. These concepts are  
 819 described briefly in the proceeding text.

#### 821 Hyper-parameter optimisation

822 Hyper-parameters are user-defined parameters of a model that are govern the entire training process. Typical

examples of hyper-parameters include the learning rate, the number of discriminators and the type of loss function used to be optimised. The learning rate determines the step size at each iteration in determining the minimum of the loss function. Hyper-parameter optimisation is a process which aims to determine the best hyper-parameters for a model, based off some performance metric. In this analysis hyper-parameter optimisation is performed using a *grid search*. In a grid search, a user-defined list of hyper-parameter values are chosen for each hyper-parameter that one aims to optimise. The model is then trained using each permutation of hyper-parameters and determines the set of hyper-parameters in which the performance metric is maximised.

### 831 Cross Validation

832 as

#### 834 4.7.2 Object-level BDT

835 The object-level BDT was trained on an alternative  $t\bar{t}$  sample (simulated using the same generator, parton shower  
836 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  
837 lepton with  $p_T > 28$  GeV. Additionally, jets in this sample are required to have  $p_T > 20$  GeV. Jets are identified  
838 as  $b$ -tagged jets by the 77% DL1r working point. These baseline selections were chosen to mimic those used in the  
839 event selection of the analysis (outlined in Table 5). The leptons and  $b$ -jets used in training the object-level BDT  
840 are required to pass the aforementioned baseline selections. This alternative  $t\bar{t}$  sample was utilised in training the  
841 BDT to avoid using a subset of events from the MC samples used in the rest of the analysis, therefore maximizing  
842 the amount of generated events available to use in other parts of the analysis.

843  
844 The signal class is defined to consist of reconstructed  $\ell b$  systems (defined as the sum of the 4-vectors of a lepton  
845 and a  $b$ -tagged jet) originating from top quarks which are well matched to their truth counterparts. All possible  
846 combinations of  $\ell$  and  $b$ -tagged jets are selected from the events. In particular, it is required that  $\Delta R$  between  
847 the reconstructed and truth  $\ell b$  system is less than 0.05. An additional requirement is implemented such that the  
848 reconstructed lepton and the truth top quark have charges with the same sign (since  $t \rightarrow b\ell^+\bar{\nu}_\ell$  and  $\bar{t} \rightarrow \bar{b}\ell^-\nu_\ell$ ).  
849 The background class is defined to consist of all reconstructed  $\ell b$  systems which fail to pass the criteria for  $\ell b$   
850 systems which are labelled as signal. These definitions for the signal and background classes ensure that the signal  
851 class consists of mostly  $\ell b$  systems originating from top quarks and the background class consists of mostly  $\ell b$   
852 systems which do not originate from top quarks.

853  
854 Different variables corresponding to an  $\ell b$  system were used as input to training. The optimum values for the hyper-  
855 parameters used were determined by training the BDT with a range of different values for the hyper-parameters  
856 and choosing the set of values which maximized the mean accuracy (based off 5 fold kfold cross-validation).  
857 This method is more commonly referred to as hyper-parameter optimisation or tuning. After hyper-parameter  
858 optimisation, the mean accuracy of each fold increased from 0.76 to 0.77 ( $\sim 1\%$  increase). Variables can be assigned  
859 a score called *variable importance*, based on their usefulness on predicting a target variable (in this case, a signal or  
860 background event). The variable importance for any given variable was obtained by computing the mean accuracy  
861 of the model, removing the variable from training, retraining the model and computing the mean accuracy of this  
862 new model. The difference between mean accuracies of the unaltered model and the retrained model (after removal  
863 of the variable) gives us the variable importance of the variable of interest. This method returns positive values  
864 for variables which increase the mean accuracy of the model and negative values for variables which decrease the  
865 mean accuracy of the model. Variables with negative variable importances were completely removed from training.

866  
867 In Table 8, the variables used in training the object-level BDT are shown.

868 In Figure 25, normalised distributions of the variables used in the object-level BDT, for the signal and background  
869 classes are shown.

870 Overall the BDT input variables show a large amount of discrimination.

871  
872 The modelling of the variables used in the object-level BDT can be checked by studying the agreement between  
873 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  
874 the  $t\bar{t}Z$  CR are shown.

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

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.

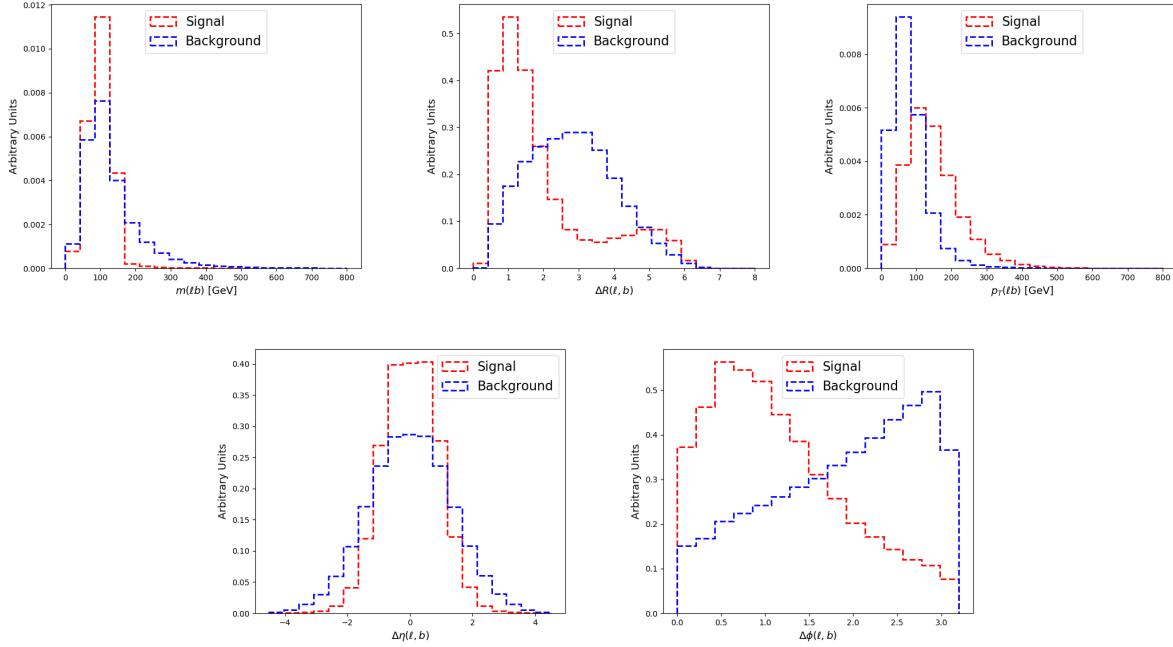


Figure 25: Normalised distributions of the variables used in the object-level BDT (ordered from top left to bottom right via decreasing variable importance), for the signal and background classes are shown. **From top left to bottom right:** Invariant mass of the  $\ell b$  system.  $\Delta R$  between 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).

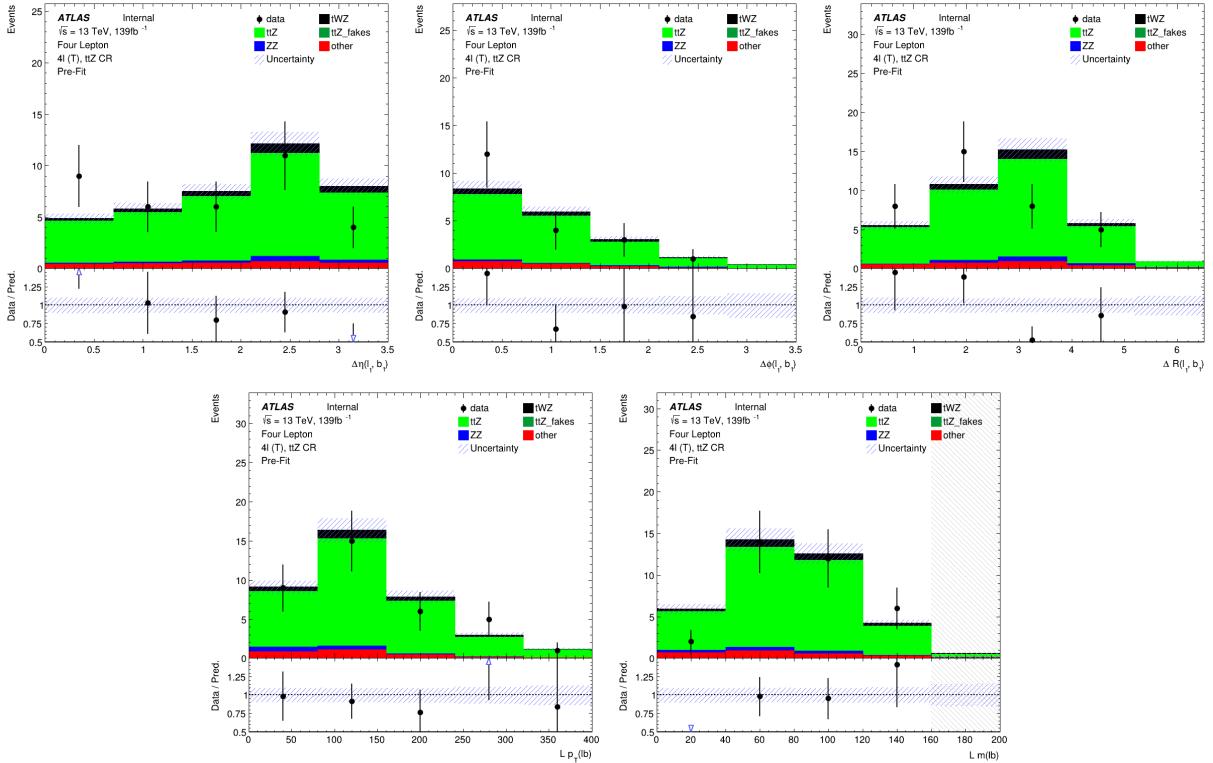


Figure 26: The expected number of events of variables used as input to the object-level BDT (ordered from top left to bottom right via decreasing variable importance), in the  $t\bar{t}Z$  CR, are shown. **From top left to bottom right:**  $\Delta\eta$  between the lepton and  $b$ -jet of the leading  $\ell b$  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.

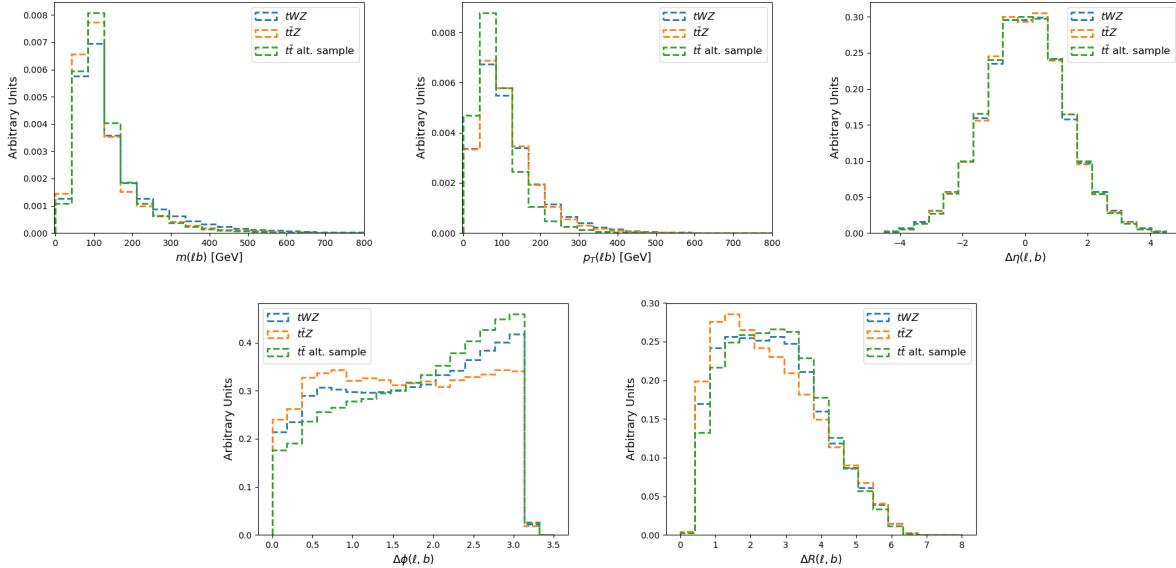


Figure 27: Normalised distributions of the variables (ordered from top left to bottom right via decreasing variable importance) used in the object-level BDT for the alternative  $t\bar{t}$ ,  $t\bar{W}Z$  and  $t\bar{t}Z$  samples, are shown. **From top left to bottom right:** Invariant mass of the  $\ell b$  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 green, blue and orange dotted histograms represent events from the alternative  $t\bar{t}$ ,  $t\bar{W}Z$  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).

877 include as inputs to the object-level BDT.

878 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  
 879 used for training the object-level BDT, and the  $\ell b$  systems which are aimed to be identified using the object-level  
 880 BDT. More specifically, the study is done to ensure that the modelling of the  $\ell b$  systems in the alternative  $t\bar{t}$  sample  
 881 are sufficiently similar to those in the  $t\bar{W}Z$  and  $t\bar{t}Z$  samples (see Table 3). This is done to understand how well  
 882 the BDT (trained on  $\ell b$  systems in the alternative  $t\bar{t}$  sample) generalises to classifying  $\ell b$  systems in the analysis  
 883 ( $t\bar{W}Z$  and  $t\bar{t}Z$  samples). In Figure 27, normalised distributions of the variables used in the object-level BDT for  
 884 the alternative  $t\bar{t}$ ,  $t\bar{W}Z$  and  $t\bar{t}Z$  samples, are shown.

885 The  $m(\ell b)$ ,  $p_T(\ell b)$  and  $\Delta\eta(\ell, b)$  distributions show little discrimination between the samples. However, the  $\Delta\phi(\ell, b)$   
 886 and  $\Delta R(\ell, b)$  distributions show a large amount of discrimination between the three samples. This suggests that the  
 887 use of the alternative  $t\bar{t}$  sample in training the object-level BDT may be sub-optimal. However, it still represents  
 888 the best option available, since our other options involve utilising of a subset of generated events used in the other  
 889 parts of the analysis. This would result in a smaller number of generated events used in the background prediction,  
 890 leading to larger statistical uncertainties, in an already statistically limited analysis.

892

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[72].

893 The number of events used in training for the signal and background classes were 49871 and 384152 respectively.  
 894 Imbalanced datasets can cause ML classifiers to ignore small classes while concentrating on classifying large classes  
 895 more accurately, which may result in the trained BDT performing sub-optimally. In order to correct this dataset  
 896 imbalance, it is ensured that the relative weighting of each event is such that the sum of the signal weights is  
 897 equal to the sum of the background weights. In Figure 28 the normalised histograms of the training and test sets  
 898 (extracted from fold 5 from a 5 fold kfold cross validation) for signal and background is shown.

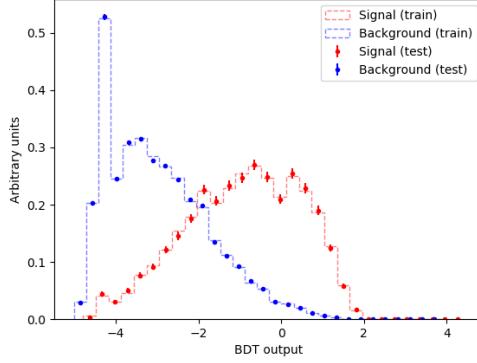


Figure 28: Normalised histograms of the object-level BDT discriminator output from the signal and background classes for the training and test sets from the 5th fold in a 5 fold kfold cross validation is shown. The output of the object-level BDT is shown on the x-axis and the relative number of events (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.

899 The shapes of the training and test sets for both signal and background are very similar. This is a good indicator  
 900 that no over-training occurred. Another over-training check is performed using 5 fold kfold cross validation. It is  
 901 ensured that the variance of the mean accuracy of each folds' test set in cross validation is substantially small. This  
 902 indicates that fluctuations in features from different training sets are not learnt by the BDT. For the object-level  
 903 BDT, a variance of  $3.24 \times 10^{-7}$  was calculated for the mean accuracies of each folds' test set in cross validation.  
 904 This tells us that the BDT does not perform better on one subset of a dataset over another and it is therefore  
 905 not prone to learning statistical fluctuations of a subset, which would result in a BDT which does not gener-  
 906 alise well to unseen datasets. This small variance therefore provides further evidence that no over-training occurred.  
 907

908 The output of the object-level BDT is converted to an event-level variable to be used in the event-level BDT. This  
 909 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  $\ell b$   
 910 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  
 911 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  
 912 for  $t\bar{t}Z$  events, therefore providing discrimination between  $tWZ$  and  $t\bar{t}Z$ . In Figure 29, normalised distributions of  
 913 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  
 914 are shown.

915 The amount of discrimination can be quantified by the separation metric, which gives the percentage of the total  
 916 area of the distributions which do not overlap. A value of 1 indicates that the distributions are fully separated (no  
 917 overlap) and a value of 0 indicates that the distributions have no separation (fully overlapped). The separation  
 918 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%,  
 919 2.96% and 4.84% respectively. The larger separation in the  $t\bar{t}Z$  CR, compared to the  $tWZ$  SRs, can be explained  
 920 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 selec-  
 921 tion requirement of exactly two  $b$ -tagged jets. In a similar way, the smaller separation in the two  $tWZ$  SRs can be  
 922 explained by the tighter selection on the number of  $b$ -tagged jets (exactly one) leading to regions which are enriched  
 923 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  
 924 event-level BDT (see Section 4.7.4) improves the mean accuracy of the BDT. This tells us that the event-level BDT  
 925 is taking advantage of the discrimination between signal and background present in the  $\text{BDTScore}(\frac{\text{Best}}{\text{2nd Best}})$  variable.

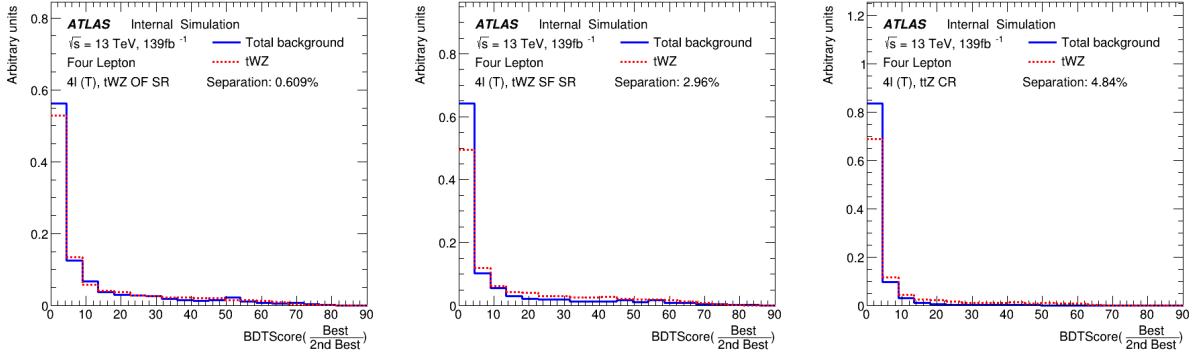


Figure 29: Normalised distributions of the signal and total background of the  $\text{BDTScore}(\frac{\text{Best}}{\text{2nd Best}})$  variable in the  $tWZ$  OF SR,  $tWZ$  SF SR and  $t\bar{t}Z$  CR are shown (left to right). The dotted red and solid blue lines represent the distributions of the signal 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).

926

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.

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.

941

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

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

### 953 4.7.3 Two Neutrino Scanning Method (2 $\nu$ SM) Algorithm

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

959

The Two Neutrino Scanning Method (2 $\nu$ SM) algorithm<sup>1</sup> [60, 59] aims to reconstruct  $t\bar{t}$  systems in the  $2\ell$ ,  $3\ell$  and

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

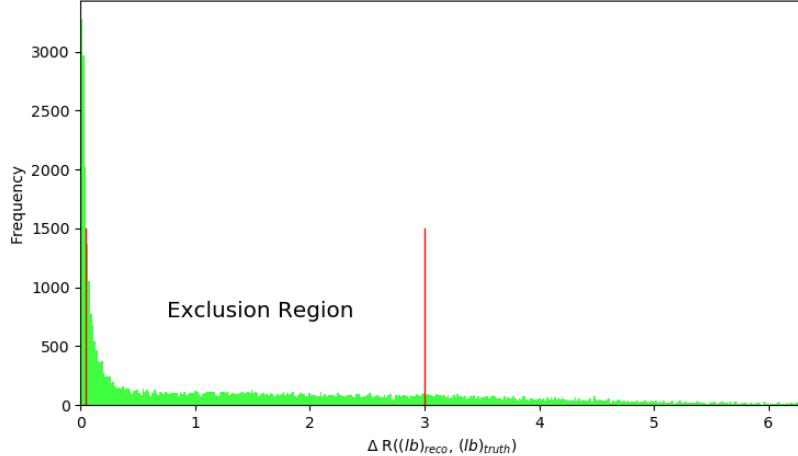


Figure 30: The distribution of  $\Delta R$  between the reconstructed  $\ell b$  system and the truth  $\ell b$  system ( $\Delta R((lb)_{reco}, (lb)_{truth})$ ) in the alternative  $t\bar{t}$  sample, along with the exclusion region, is shown. The  $\Delta R$  distribution is shown in green.  $\Delta R$  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.

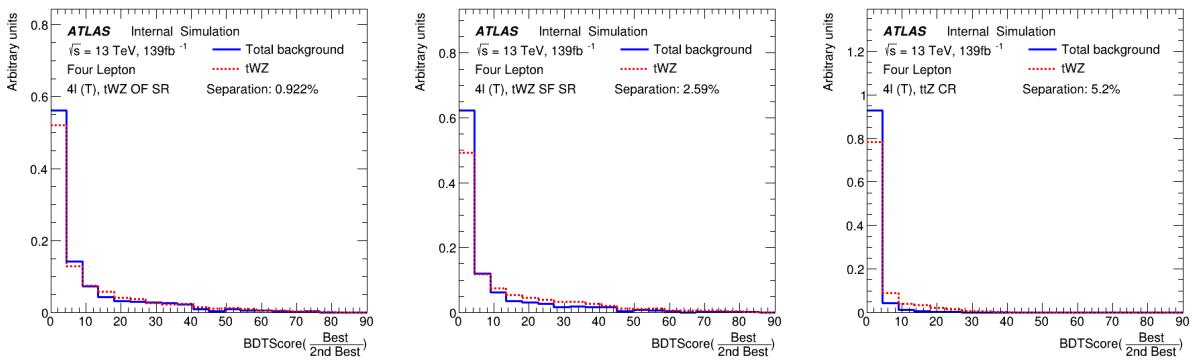


Figure 31: Normalised distributions of  $BDTScore(\frac{\text{Best}}{\text{2nd Best}})$  using the object-level BDT without the exclusion region (see Figure 30) for the  $tWZ$  OF SR,  $tWZ$  SF SR and  $ttZ$  CR are shown (left to right). The dotted red and solid blue lines represent the distributions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

961 4 $\ell$  final states (e.g. 2 $\ell$  case:  $t\bar{t} \rightarrow \ell^+\nu_\ell b\ell^-\bar{\nu}_\ell \bar{b}$ ). This was initially designed to suppress the  $t\bar{t}$  background in the  $t\bar{t}Z$   
 962 analysis. This algorithm can be re-purposed for our analysis, to distinguish between  $tWZ$  and  $t\bar{t}Z$ , by removing  
 963 the easily-identifiable  $Z$  boson.

964  
 965 The 2 $\nu$ SM algorithm reconstructs a  $t\bar{t}$  system by scanning through the components of two possible neutrino 4-  
 966 vectors ( $\nu_1$  and  $\nu_2$ ). It then aims to determine which  $\nu_1$  and  $\nu_2$  correspond to the two neutrinos which originate  
 967 from the decay of a  $t\bar{t}$  system the best (quantified by an output weight,  $w_{2\nu SM}$ ).  $w_{2\nu SM}$  is the likelihood under the  
 968  $t\bar{t}$  dilipeton final state hypothesis. This algorithm can be used in our analysis to discriminate between  $tWZ$  and  $t\bar{t}Z$ ,  
 969 since we can easily reconstruct the OSSF leptons which decay from the  $Z$  boson and remove it before inputting the  
 970 event into the algorithm. It would then be expected that the 2 $\nu$ SM algorithm returns a higher score from a  $t\bar{t}Z$   
 971 event ( $\sim 1$ , i.e. it 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.  
 972 it does not look like a  $t\bar{t}$  event after removal of the  $Z$  boson).

#### 973 4.7.3.1 The algorithm

974 The first step in the 2 $\nu$ SM algorithm involves stating four equations which correspond to the invariant masses of  
 975 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  $t\bar{t}$   
 976 event. These can be written as,

$$(977) (\ell_1 + \nu_1)^2 = m(W)^2 = (80.385 \text{ GeV})^2 \quad (4.2)$$

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

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

$$(978) (\ell_2 + \nu_2 + b_{2,1})^2 = m(t)^2 = (172.5 \text{ GeV})^2 \quad (4.5)$$

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

980  
 981 The 2 $\nu$ SM algorithm takes the 4-vectors of the two reconstructed leptons (not from the  $Z$  boson) and the two jets  
 982 with the highest DL1r  $b$ -tagger score as input. For each neutrino ( $\nu_1$  and  $\nu_2$ ), a scan over a range of possible  $\eta$   
 983 and  $\phi$  values is performed. These values were chosen to be  $\phi_{\nu_1}, \phi_{\nu_2} \in [-\pi, \pi]$  with a step size of  $\approx 0.25$  and  $\eta_{\nu_1},$   
 984  $\eta_{\nu_2} \in [-5, 5]$  with a step size of  $\approx 0.31$ . These ranges were chosen to maximize accuracy and minimize computation  
 985 time. For each of these possible  $\eta$  and  $\phi$  values, the corresponding  $p_T$  for each neutrino is calculated. The transverse  
 986 momentum of a neutrino,  $p_{T\nu}$ , can be calculated via (\*\*\*\*\*referecne somewhere here\*\*\*\*\*),

$$(987) p_{T\nu} = \frac{\frac{1}{2}(m(W)^2 - m(\ell)^2)}{E_\ell \cosh \eta_\nu - p_{\ell,z} \sinh \eta_\nu - p_{\ell,x} \cos \phi_\nu - p_{\ell,y} \sin \phi_\nu} \quad (4.6)$$

988 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.  
 989 At this stage, possible 4-vectors for  $\nu_1$  and  $\nu_2$  are defined. Using these possible neutrino 4-vectors, 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

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

990 These reconstructed  $t\bar{t}$  systems are then used to calculate a weight,  $w_{2\nu SM}$ . The  $w_{2\nu SM}$  weight (a value ranging  
 991 from 0 to 1) is defined as a product of four probabilities (described below) and can be written as,

$$w_{2\nu SM} = P_{m_{t_1}} \times P_{m_{t_2}} \times P_{\Delta E_x} \times P_{\Delta E_y} \quad (4.9)$$

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

994 **4.7.3.2 Calculating  $w_{2\nu SM}$** 

995 Distributions of well modelled variables ( $m_{b\ell\nu}$  and  $\Delta E_x$ ) from simulated  $t\bar{t}$  events are used in order to determine  
 996 how well our reconstructed neutrinos (and in turn top quarks) resemble neutrinos (and top quarks) present in a  $t\bar{t}$   
 997 event.

998  **$P_{m_{t_1}}$  and  $P_{m_{t_2}}$** 

999  
 1000 A normalised distribution of the mass of reconstructed top quarks ( $m_{b\ell\nu}$ ) from a  $t\bar{t}$  sample is generated to determine  
 1001 the probabilities  $P_{m_{t_1}}$  and  $P_{m_{t_2}}$ . The distribution is generated from reco-level leptons, generator-level neutrinos  
 1002 and reoc-level jets matched in  $\Delta R$  to generator-level  $b$ -quarks, therefore only filling the distribution with correct  
 1003 detector-level objects. The distribution is then used to interpolate our two reconstructed top quarks, which returns  
 1004 a weight value from 0 to 1, with higher values corresponding to a reconstructed top quark which has a mass close  
 1005 to that of a top quark from a  $t\bar{t}$  system. This interpolation is done for both reconstructed top quarks,  $t_1$  and  $t_2$ ,  
 1006 corresponding to probabilities  $P_{m_{t_1}}$  and  $P_{m_{t_2}}$ . The distribution used is shown in Figure 32.

1007  
 1008 In Figure 32, the  $m_{b\ell\nu}$  distribution (generated from simulated  $t\bar{t}$  events), used to calculate  $P_{m_{t_1}}$  and  $P_{m_{t_2}}$  is shown.

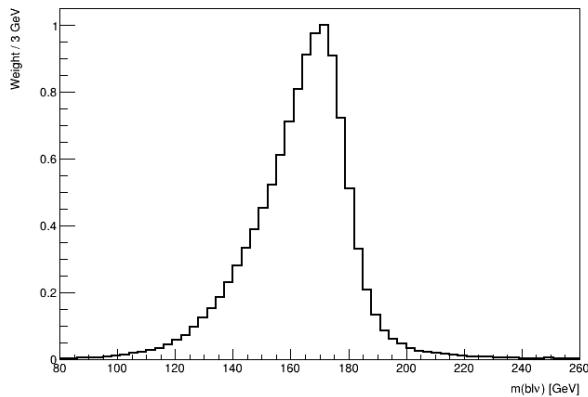


Figure 32:  $m_{b\ell\nu}$  distribution generated from simulated  $t\bar{t}$  events, used to calculate  $P_{m_{t_1}}$  and  $P_{m_{t_2}}$  is shown. The  $m_{b\ell\nu}$  distribution is shown by the black lined histogram. The mass of the  $b\ell\nu$  system is shown on the x-axis. The corresponding weight of the  $m_{b\ell\nu}$  distribution is shown on the y-axis.

1009  **$P_{\Delta E_x}$  and  $P_{\Delta E_y}$** 

1010  
 1011 A similar method is used to determine  $P_{\Delta E_x}$  and  $P_{\Delta E_y}$ . In this case, a weight distribution of  $\Delta E_x = (p_{T,\nu_1})_x +$   
 1012  $(p_{T,\nu_2})_x - (E_T^{\text{miss}})_x$  based off simulated  $t\bar{t}$  events is generated. In particular, this distribution is generated using  
 1013 reco-level  $E_T^{\text{miss}}$  and generator-level neutrinos. The use of this distribution lies under the assumption that neutrinos  
 1014 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  $(E_T^{\text{miss}})_y \approx (p_{T,\nu_1})_y + (p_{T,\nu_2})_y$ .  
 1015 This distribution is then used to interpolate the value of  $\Delta E_x$  and  $\Delta E_y$  from our reconstructed neutrinos. This  
 1016 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  
 1017 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  
 1018 same shapes, therefore we only need to generate one (the  $\Delta E_x$  distribution was chosen). In Figure 33, the  $m_{b\ell\nu}$   
 1019 distribution (generated from simulated  $t\bar{t}$  events), used to calculate  $P_{m_{t_1}}$  and  $P_{m_{t_2}}$  is shown.

1020 **4.7.3.3 Kinematic Veto**

1021 The  $2\nu SM$  algorithm is extremely computationally intensive. The computation time depends on the number step  
 1022 size of the  $\phi$  and  $\eta$  ranges which are scanned over to reconstruct the neutrinos. For example, consider the step  
 1023 sizes 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  $\phi$ .  
 1024 There will be  $(32)(32)(25) = 640\,000$  possible pairs of neutrinos ( $\nu_1$  and  $\nu_2$ ) to consider **per event**. Since two  
 1025 possible  $t\bar{t}$  systems (See Equations 4.7 and 4.8) are considered, this number effectively increases to  $(2)(640\,000) =$   
 1026  $128\,000$  iterations **per event**. In order to reduce the number of  $t\bar{t}$  systems needed to be considered, therefore  
 1027 decreasing computation time, distributions of well modelled variables from  $t\bar{t}$  events are studied to apply a veto

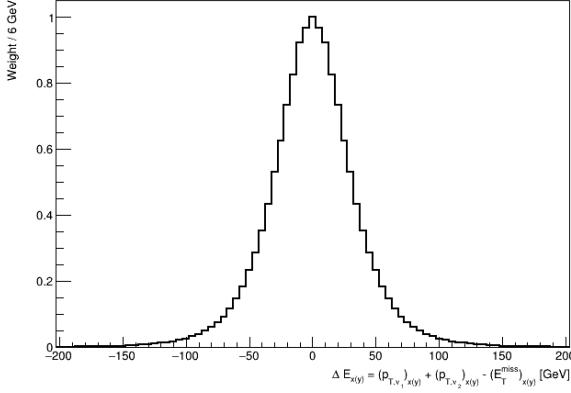


Figure 33:  $\Delta E_x$  distribution generated from simulated  $t\bar{t}$  events, used to calculate  $P_{\Delta E_x}$  and  $P_{\Delta E_y}$  is shown. The  $\Delta E_x$  distribution is shown by the black lined histogram.  $\Delta E_x$  is shown on the x-axis. The corresponding weight of  $\Delta E_x$  distribution is shown on the y-axis.

(discard) a possible reconstructed  $t\bar{t}$  system if the variable in question is improbable or unlikely to be observed in a  $t\bar{t}$  event. To achieve this, a threshold range is defined for these variables (See Figure 36 and Figure 38), and if the possible reconstructed  $t\bar{t}$  system's corresponding value for this variable lies outside this range, it is vetoed and the algorithm continues with the next iteration.

$$\Delta \langle m(\ell b) \rangle$$

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

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

**OR**

$$(\ell b)_1 = \ell_1 + b_2 \text{ and } (\ell b)_2 = \ell_2 + b_1 \quad (4.11)$$

$$(4.12)$$

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

$$\Delta \langle m(\ell b) \rangle = \frac{1}{2} |[(m(\ell_1 b_1) + m(\ell_1 b_1)) - (m(\ell_1 b_2) + m(\ell_2 b_1))| \quad (4.13)$$

The idea here is that, if  $\Delta \langle m(\ell b) \rangle$  is large, it's more likely that one can simply select the  $\ell b$  combination with the smaller (minimum) average mass. To illustrate this, the distributions (constructed from  $t\bar{t}$  events) of  $P(\text{Correct combination of } \ell b \text{ systems} | \text{minimum} \langle m(\ell b) \rangle)$  vs  $\Delta \langle m(\ell b) \rangle$  for  $b$ -tagged jets in the same ( $\eta(b_1) \times \eta(b_2) \geq 0$ ) and opposite hemispheres ( $\eta(b_1) \times \eta(b_2) < 0$ ) are investigated. In Figure 34 the  $P(\text{Correct combination of } \ell b \text{ systems} | \text{minimum} \langle m(\ell b) \rangle)$  vs  $\Delta \langle m(\ell b) \rangle$ , for  $b$ -tagged jets in the same and opposite hemispheres, constructed from  $t\bar{t}$  events is shown.

From Figure 34, for both cases where the  $b$ -tagged jets are in the same and opposite hemispheres, the probability for a correct  $\ell b$  system being chosen, given that the  $\ell b$  system with the minimum average mass is under consideration, is an increasing function which plateaus to 1 at  $\sim 90$  GeV. These two distributions are used to interpolate the  $P(\text{Correct combination of } \ell b \text{ systems} | \text{minimum} \langle m(\ell b) \rangle)$  from  $\Delta \langle m(\ell b) \rangle$ . It is required that  $P(\text{Correct combination of } \ell b \text{ systems} | \text{minimum} \langle m(\ell b) \rangle) > 0.8$ , before vetoing any  $\ell b$  combination, such that we have are at least 80% certain that we know the correct  $\ell b$  combination. In this case, the  $\ell b$  combination with the maximum  $\Delta \langle m(\ell b) \rangle$  is vetoed. If  $P(\text{Correct combination of } \ell b \text{ systems} | \text{minimum} \langle m(\ell b) \rangle) < 0.8$  we need to consider both possible  $\ell b$  system combinations.

$$\eta(b\bar{b}\ell\ell)$$

We consider  $\eta$  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.

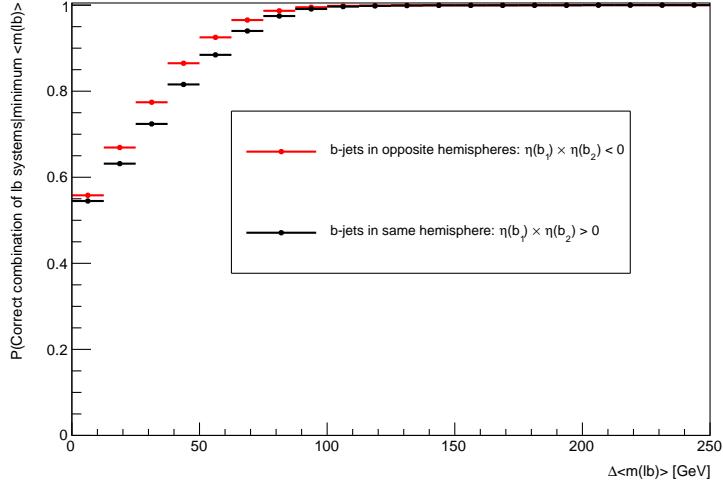


Figure 34:  $P(\text{Correct combination of } \ell b \text{ systems} | \text{minimum} \langle m(\ell b) \rangle)$  vs  $\Delta \langle m(\ell b) \rangle$ , for  $b$ -tagged jets in the same and opposite hemispheres, constructed from  $t\bar{t}$  events is shown. The horizontal red lines represent the 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  $m(\ell b)$  is shown on the x-axis. The  $P(\text{Correct combination of } \ell b \text{ systems} | \text{minimum} \langle m(\ell b) \rangle)$  is shown on the y-axis.

1055 In the same way as for  $\Delta \langle m(\ell b) \rangle$ , we generate a distribution to determine values  $\eta(\nu)$  which are improbable for a  $t\bar{t}$   
 1056 event. In this case, we generate a 2D histogram from simulated  $t\bar{t}$  events (dileptonic final state) at generator-level  
 1057 of  $\eta(\nu)$  vs  $\eta(b\bar{b}\ell\ell)$ .

1058  
 1059 In Figure 35, a heatmap of occupancy for  $\eta(\nu)$  vs  $\eta(b\bar{b}\ell\ell)$  (produced from simulated  $t\bar{t}$  events) is shown.

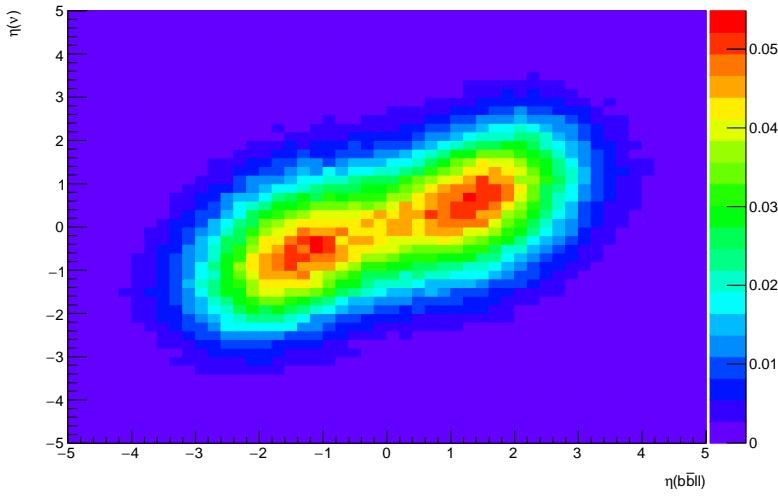


Figure 35: Heatmap of occupancy for  $\eta(\nu)$  vs  $\eta(b\bar{b}\ell\ell)$  produced from simulated  $t\bar{t}$  events (dileptonic final state) at generator-level is shown.  $\eta$  of the  $b\bar{b}\ell\ell$  system is shown on the x-axis.  $\eta$  of the neutrino is shown on the y-axis. The colorbar on the right represents the occupancy (normalised) in the phase space.

1060 Using the above heatmap, we define a veto region (where a  $t\bar{t}$  event is extremely unlikely to occur) based off  
 1061 double-sided 95% limits (\*\*something here on confidence limit??\*\*). We apply a veto if either possible neutrino  
 1062 lies within this region. The veto region is shown in Figure 36.

1063

1064 In Figure 36, the veto region (extracted from Figure 35) for vetoing improbable neutrinos is shown.

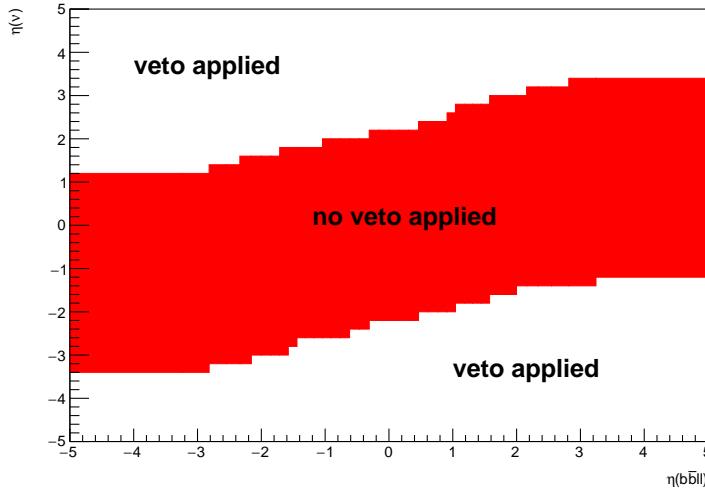


Figure 36: The regions where vetoes are applied for the  $\eta(b_1 b_2 \ell_1 \ell_2)$  constraint is shown.  $\eta$  of the  $b\bar{b}\ell\ell$  system is shown on the x-axis.  $\eta$  of the neutrino is shown on the y-axis. The red band shows the region where the neutrino would not be vetoed. The white areas (above and below the red band) are regions where the neutrino is vetoed.

1065  $L_T$

1066

1067 The final kinematic constraint which we consider is the scalar sum of lepton  $p_T$ ,  $L_T = p_T(\ell_1) + p_T(\ell_2)$  which we  
1068 use to veto certain possible neutrinos,  $\nu_1$  and  $\nu_2$ .

1069

1070 Again, we generate a distribution to determine (and veto) improbable possible neutrinos in simulated  $t\bar{t}$  events  
1071 (dilepton final state).

1072 In Figure 37, a heatmap of occupancy for  $\Delta R(\ell, \nu)$  vs  $L_T$  (produced from simulated  $t\bar{t}$  events) is shown.

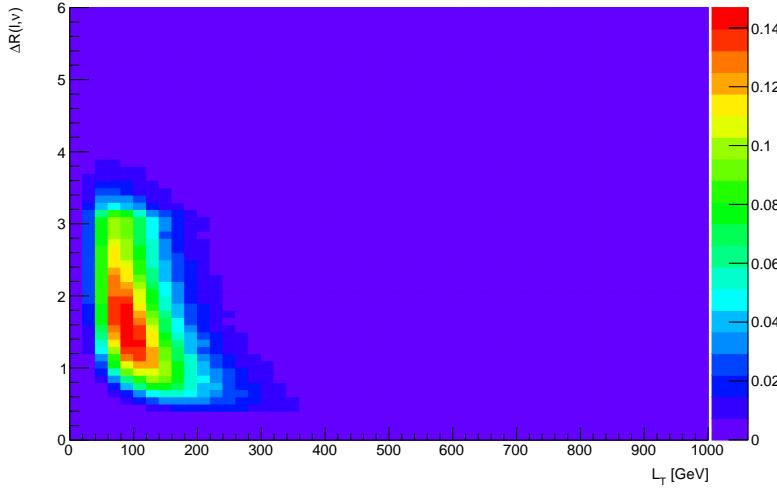


Figure 37: A heatmap of occupancy for  $\Delta R(\ell, \nu)$  vs  $L_T$  produced from simulated  $t\bar{t}$  events (dileptonic final state) at generator-level is shown.  $\Delta R$  between leptons and neutrinos is shown on the x-axis.  $L_T$  (scalar sum of lepton  $p_T$ ) is shown on the y-axis. The colorbar on the right represents the occupancy (normalised) in the phase space.

1073 Using the same method as described in Section 4.7.3.3, we define a veto region where a veto is applied if either

<sup>1074</sup> possible neutrino lies within this region. In Figure 36, the veto region (extracted from Figure 37) for vetoing improbable neutrinos values is shown.

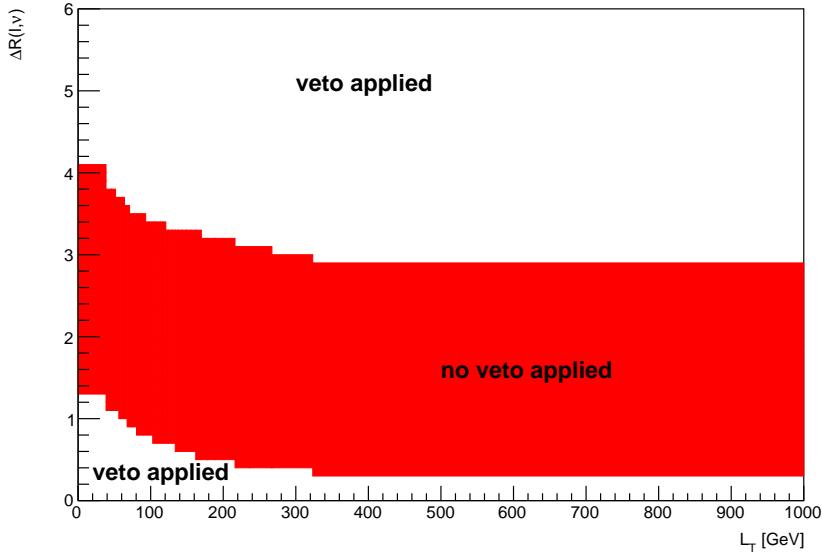


Figure 38: The regions where vetoes are applied for the  $L_T$  constraint is shown.  $\Delta R$  between leptons and neutrinos is shown on the x-axis.  $L_T$  (scalar sum of lepton  $p_T$ ) is shown on the y-axis. The red band shows the region where the neutrino would not be vetoed. The white areas (above and below the red band) are regions where the neutrino is vetoed.

<sup>1075</sup>

#### <sup>1076</sup> 4.7.4 Event-level BDT

<sup>1077</sup> The event-level BDT is used to distinguish between signal and its major background events,  $t\bar{t}Z$  and  $ZZ$ . The key  
<sup>1078</sup> difference between the object-level BDT and the event-level BDT is that while the former exploits information  
<sup>1079</sup> associated with  $\ell b$  systems and thus distinguishes between  $\ell b$  systems, the event-level BDT exploits information  
<sup>1080</sup> based on the entirety of the event and thus distinguishes between events.

<sup>1081</sup>

<sup>1082</sup> The event-level BDT was trained on 50% of the  $tWZ$  MC sample's events for the signal class and similarly, 50%  
<sup>1083</sup> 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  
<sup>1084</sup> are chosen on the basis that they are somewhat uncorrelated from one another and show some discrimination  
<sup>1085</sup> between  $tWZ$  and  $t\bar{t}Z$ . Similarly to the object-level BDT, the optimum values for the hyper-parameters used were  
<sup>1086</sup> determined via hyper-parameter optimisation. After hyper-parameter optimisation, the mean accuracy of each  
<sup>1087</sup> fold (determined from 5 fold kfold cross validation) increased from 0.72 to 0.74 ( $\sim 3\%$  increase). The variable  
<sup>1088</sup> importance of each variable was computed in the same way as described for the object-level BDT (See Section 4.7.2).

<sup>1089</sup>

In Table 10, the variables used in training the event-level BDT are shown.

Variable	Description	Variable Importance
$2\nu\text{SM}$	Maximum weight from the $2\nu\text{SM}$ algorithm (See Section 4.7.3)	0.029
$HT$	Scalar sum of jet $p_T$	0.016
$L_T$	Scalar sum of lepton $p_T$	0.011
$\Delta\eta(\ell_{1,\text{non-}Z}, \ell_{2,\text{non-}Z})$	$\Delta\eta$ between the two leptons, not coming from a $Z$ candidate	0.006
$\text{BDTScore}\left(\frac{\text{Best}}{\text{2nd Best}}\right)$	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
$\sum p_T(b - \text{jet})$	Scalar sum of b-tagged jet $p_T$	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.

<sup>1090</sup>

<sup>1091</sup> In Figure 39, normalised distributions of the variables used in the event-level BDT, for the signal and background  
<sup>1092</sup> classes are shown.

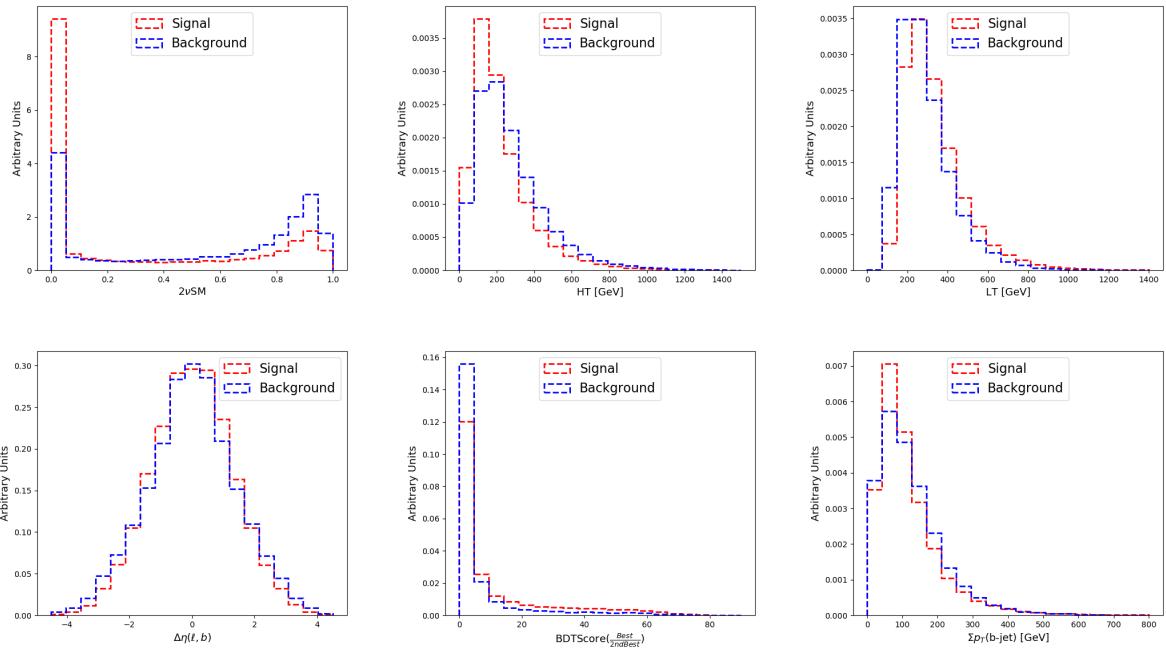


Figure 39: Normalised distributions of the variables used in the event-level BDT (ordered from top left to bottom right via decreasing variable importance), for the signal and background classes are shown. **From top left to bottom right:** Output weight from the  $2\nu\text{SM}$  algorithm (See Section 4.7.3). Scalar sum of jet  $p_T$ . Scalar sum of lepton  $p_T$ .  $\Delta\eta$  between the two leptons, not coming from a  $Z$  candidate. 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). Sum of  $b$ -tagged jet  $p_T$ . 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).

Overall the variables used in the BDT show a reasonable amount of discrimination. In particular the output weight from the  $2\nu\text{SM}$  algorithm shows the most discrimination. When determining which 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 40, 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 41, 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 42, 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.

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[72].

Since we are training on  $t\bar{t}Z$  and  $ZZ$  events for the background class, we ensure 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 where we aim to use the BDT discriminator ( $tWZ$  SRs and  $t\bar{t}Z$  CR). This is done by applying normalization weights to each event, defined as,

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

where  $\sigma$  is the cross section of the process,  $\mathcal{L}$  is the integrated luminosity,  $\text{weight(MC)}$  is the weight assigned to the event by the MC generator and  $\text{totalWeight(MC)}$  is the sum of those weights for all the generated events.

The number of events used in training for the signal and background classes were 41066 and 22608 respectively. Similarly to the object-level BDT, there is a dataset imbalance. We correct this imbalance (in the same way as before with the object-level BDT) by ensuring 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 Figure 43 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.

We can see that the shapes of the training and test sets for both signal and background are very similar. This is a good indicator that no over-training occurred. As with the object-level BDT, we perform another over-training

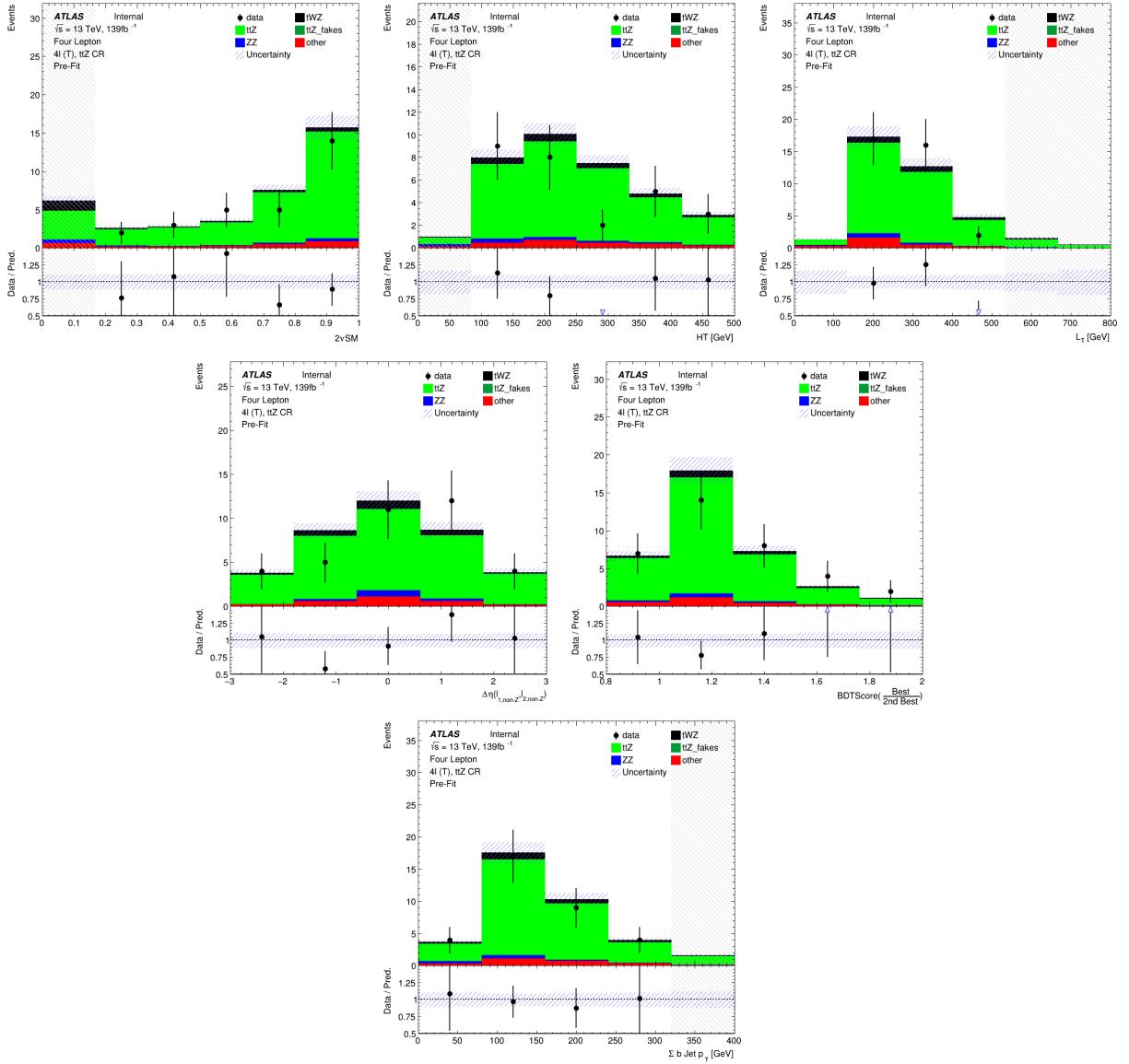


Figure 40: The expected number of events of variables used as input to the event-level BDT (ordered from top left to bottom right via decreasing variable importance), in the  $t\bar{t}Z$  CR, are shown. **From top left to bottom right:** Output weight from the  $2\nu$ SM algorithm (See Section 4.7.3). Scalar sum of jet  $p_T$ . Scalar sum of lepton  $p_T$ .  $\Delta\eta$  between the two leptons, not coming from a  $Z$  candidate. 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). 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.

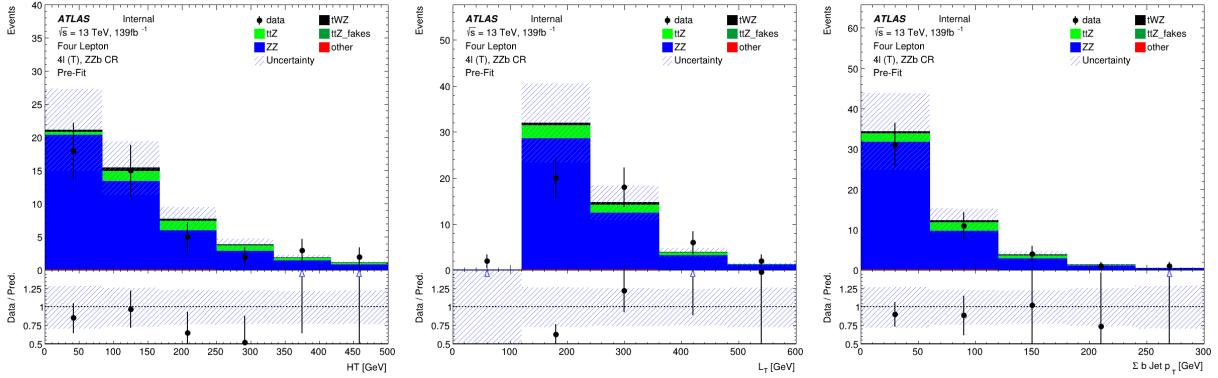


Figure 41: The expected number of events of variables used as input to the event-level BDT (ordered from top left to bottom right via decreasing variable importance), in the  $Z Z b$  CR, are shown. **From left to right:** Scalar sum of jet  $p_T$ . Scalar sum of lepton  $p_T$ . Sum of  $b$ -tagged jet  $p_T$ . 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.

check, by ensuring that the variance of the mean accuracy of each folds' test set in a 5 fold kfold cross validation is sufficiently small. This indicates that fluctuations in features from different training sets are not learnt by the BDT. For the event-level BDT, a variance of 0.00026 was calculated for the mean accuracies of each folds' test set in cross validation, providing further evidence that no over-training occurred. In Figure 44, 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.  
The event-level BDT discriminates well between signal and background events in the  $tWZ$  OF SR,  $tWZ$  SF SR and  $t\bar{t}Z$  CR, with separations of 8.98%, 10.6% and 20.6%, respectively.

## 4.8 Systematics

The systematic uncertainties can be separated into experimental (detector) systematics, which are related to the reconstruction of physics objects in the detector and theoretical uncertainties related to the modelling of the different processes background.

### 4.8.1 Experimental uncertainties

In this section, the experimental systematics are outlined.

- **Luminosity:**

The 2015–2018 luminosity estimate of  $139\text{fb}^{-1}$  has a relative uncertainty of 3%. This uncertainty is obtained using the LUCID-2 detector [23] for the primary luminosity measurements. This systematic uncertainty affects all processes modelled using MC simulations.

- **Pile-up reweighting:**

An uncertainty related to the SFs used for MC to account for differences in pile-up distributions between 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 corrections are only applied to MC.

- **Jet vertex tagger:**

Uncertainties associated to the  $JVT$  are applied via the `JetJvtEfficiency` package [21] which account for the residual contamination from pile-up jets after pile-up suppression and the MC generator choice [14].

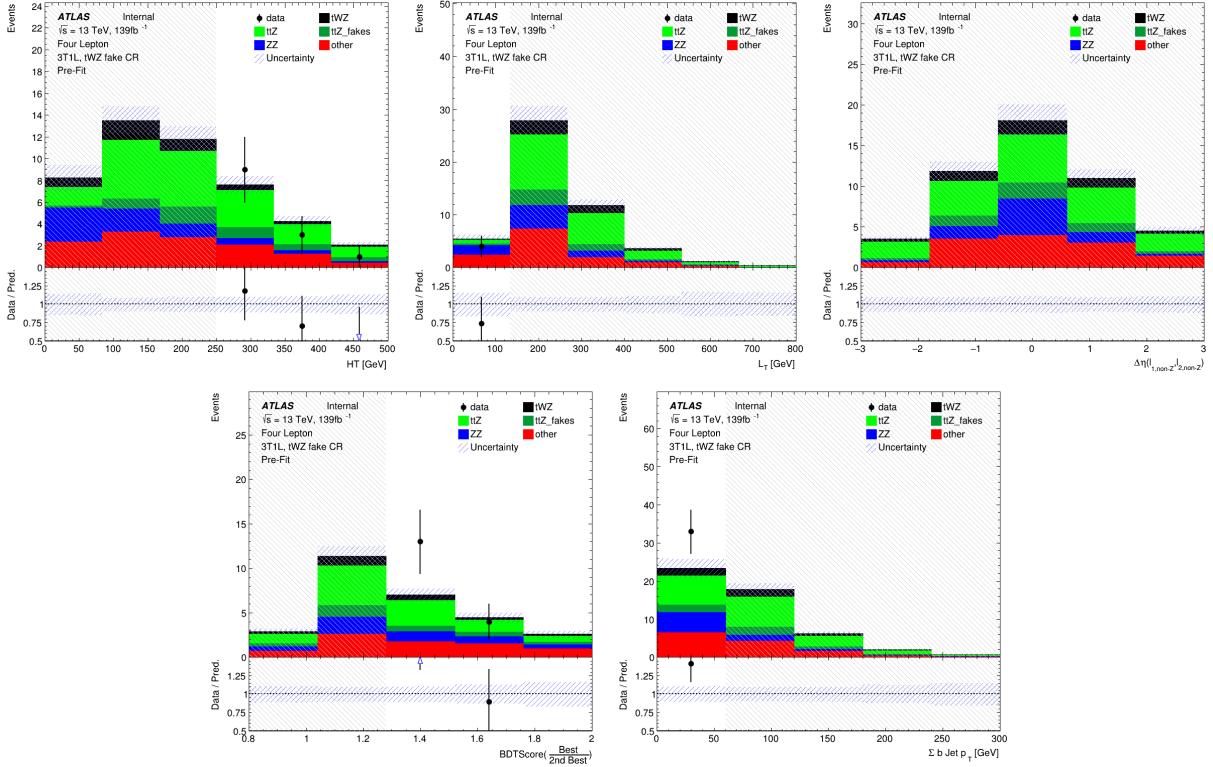


Figure 42: The expected number of events of variables used as input to the event-level BDT (ordered from top left to bottom right via decreasing variable importance), in the ( $tWZ$ )<sub>fake</sub> CR, are shown. **From top left to bottom right:** Output weight from the  $2\nu$ SM algorithm (See Section 4.7.3). Scalar sum of jet  $p_T$ . Scalar sum of lepton  $p_T$ .  $\Delta\eta$  between the two leptons, not coming from a  $Z$  candidate. 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). 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.

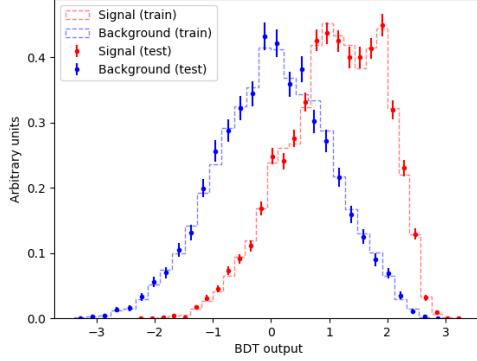


Figure 43: Normalised histograms of the event-level BDT discriminator output from the signal and background classes for the training and test sets from the 5th fold in a 5 fold kfold cross validation are shown. The output of the event-level BDT is shown on the x-axis and the relative number of events (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.

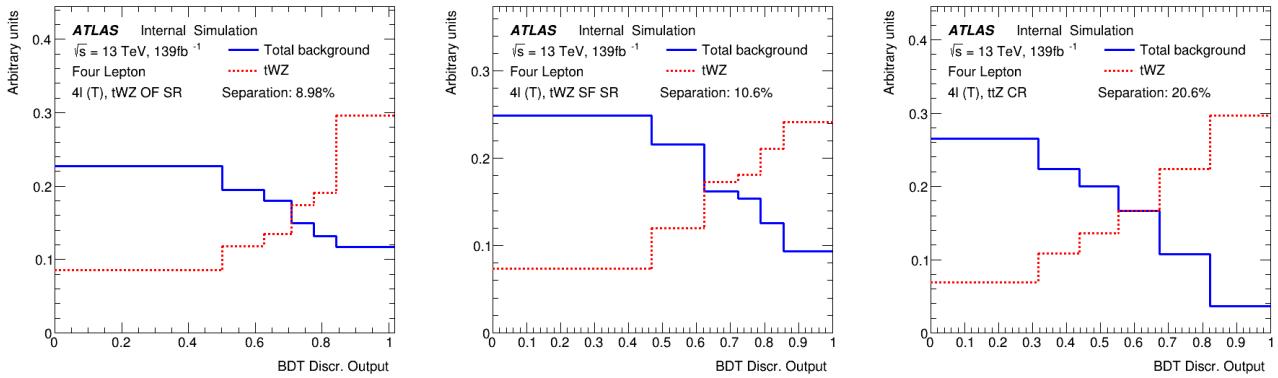


Figure 44: Normalised distributions of the signal and total background of the event-level BDT discriminator output in the  $tWZ$  OF SR,  $tWZ$  SF SR and  $t\bar{t}Z$  CR are shown (left to right). The dotted red and solid blue lines represent the distributions of the signal and total background events 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).

#### • Heavy- and light-flavor tagging:

The efficiency of the flavour-tagging algorithm is measured for each jet flavour 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 [16, 11]. In the case of  $c$ -jets, they are derived from jets arising from  $W$  boson decays in  $t\bar{t}$  events [15]. In the case of light-flavour jets, the correction factors are derived using dijet events [12]. Sources of uncertainty affecting the  $b$ - and  $c$ -tagging efficiencies are evaluated as a function of jet  $p_T$ , including bin-to-bin correlations. The uncertainties in the efficiency for tagging light-flavour jets depend on the jet  $p_T$  and on  $\eta$ . An additional uncertainty is assigned to account for the extrapolation of the  $b$ -tagging efficiency measurement from the  $p_T$  region used to determine the correction factors to regions with higher  $p_T$ .

#### • Electron efficiency:

Uncertainties associated with the electron efficiency SFs are provided by the egamma CP group [19] and arise from the reconstruction, ID, isolation and trigger efficiencies. They correct for the efficiency difference

1173 between data and MC [13] and are measured with a “tag-and-probe” method in  $Z \rightarrow e^+e^-$  and  $J/\psi \rightarrow e^+e^-$   
 1174 events. The information on the correlation of the different components of the systematic uncertainties are  
 1175 provided for all efficiency measurements. The default correlation model for the uncertainties is used, which  
 1176 provides one up/down variation for each of the SF components separately [19, 20].

- 1177 • **Muon efficiency:**

1178 As for electrons, SFs obtained from  $Z \rightarrow \mu^+\mu^-$  and  $J/\psi \rightarrow \mu^+\mu^-$  events are applied to correct for the  
 1179 differences between data and MC in the muon ID, isolation and trigger efficiencies [17]. Uncertainties on these  
 1180 SFs are provided by the muon CP group [22] and applied as up/down variations of the nominal SFs for each  
 1181 component.

#### 1182 4.8.2 Theoretical uncertainties

1183 In this section, the theoretical systematics are outlined.

- 1184 •  **$t\bar{t}Z$  background:**

1185 An overall normalization uncertainty of 10% is considered for the  $t\bar{t}Z$  background. Two generic shape sys-  
 1186 tematics are considered for the  $t\bar{t}Z$  background. They are constructed (see Section 4.8.3) by either applying a  
 1187 linear or triangular interpolation to up and down variations which are defined to be  $\pm 20\%$  from the nominal  
 1188  $t\bar{t}Z$  background.

- 1189 •  **$ZZ$  background:**

1190 An overall normalization uncertainty of 30% is considered for the  $ZZ$  background.

- 1191 •  **$t\bar{t}H$  background:**

1192 An overall normalization uncertainty of 20% is considered for the  $t\bar{t}H$  background.

- 1193 •  **$tZq$  background:**

1194 An overall normalization uncertainty of 14% is considered for the  $tZq$  background.

- 1195 •  **$t\bar{t}Z$  fake background:**

1196 An overall normalization uncertainty of 50% is considered for the  $t\bar{t}Z$  fake background.

- 1197 • **other background processes:**

1198 The ‘other’ background consists of many processes which have minimal but non-negligible contribution in  
 1199 the signal regions (See Table 6). An overall normalization uncertainty of 30% is considered for the ‘other’  
 1200 background processes.

- 1201 •  **$tWZ$ :** A modelling uncertainty on  $tWZ$  is considered by comparing the nominal sample (using the DR1  
 1202 scheme) and a minimal DR2 sample.

1203 Two generic shape systematics are considered for the  $tWZ$  background. They are constructed (see Section  
 1204 4.8.3) by either applying a linear or triangular interpolation to up and down variations which are defined to  
 1205 be  $\pm 20\%$  from the nominal  $tWZ$  background.

#### 1206 4.8.3 Generic shape systematics

1207 It is evident that the tetralepton channel is statistically limited. We therefore expect that the uncertainty on  $u_{tWZ}$   
 1208 is dominated by statistical uncertainty and that the impact of shape systematics will be negligible in comparison.

1209 In order to include shape uncertainties related to the modelling of our samples, we construct generic shape  
 1210 systematics for any given sample process. Given that we choose a sufficiently large set of values for which the  
 1211 systematics can take in the fit, the constructed systematics could represent many shape systematics which we have  
 1212 not yet considered to include in the fitting procedure.

1215 We start by constructing an envelope (error bars) consisting of two MC templates. One with the nominal MC  
 1216 template increased by 20% on its normalisation and the other with the nominal MC template decreased by 20% on  
 1217 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 rightmost and leftmost parts and reaches the envelope in the middle). The linear and triangular interpolation is done using TRF's `ForceShape` option [68], which alters the original templates (as described above). This envelope now represents the bounds which the systematic can vary in the fit.

In Figure 45 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.

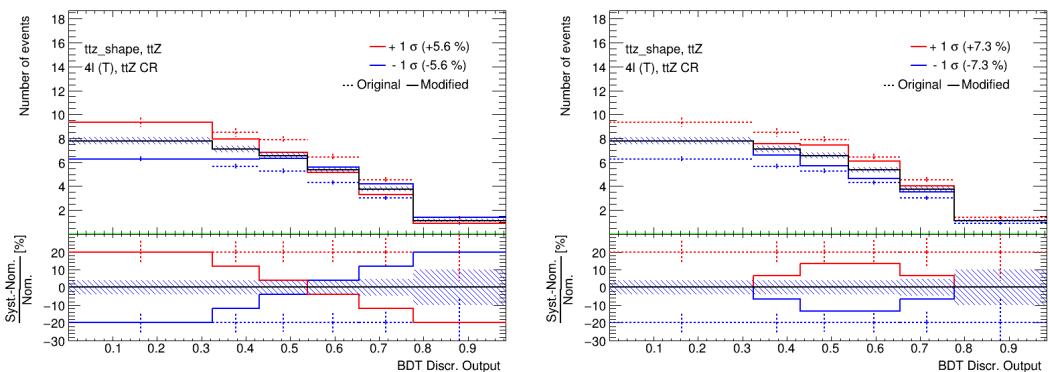


Figure 45:  $t\bar{t}Z$  generic shape systematic before (original) and after (modified) linear (left) and triangular (right) interpolation in the  $t\bar{t}Z$  CR is shown. The output from the event-level BDT shown on the x-axis. In the upper panel, the number of events is shown on the y-axis. In the 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.

We consider two shape systematics (linear and triangular shapes) for  $tWZ$  and the most dominant background processes across both 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>2</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 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[79]. In this analysis, we used `TRexFitter` (tag: `TRexFitter-00-04-13`) to produce all plots (including fit statistics, e.g. limit, significance,  $\mu_{best-fit}$ ).

The analysis pipeline starts with sample derivations (derived dataset) being submitted to the grid for ntuple production. This applies cuts and selections to the already reduced derivations and produces ntuples with trees containing variables (e.g. scale factors, 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 thrown away (if the event does not pass the selection criteria), or gets written to a `ROOT` file (if

<sup>2</sup>CERN's HEP data analysis framework (written in C++)

1247 the event passes the selection criteria) corresponding to the MC sample and Run 2 data-set which it belongs to.  
 1248 These ROOT files contain all variables, weights and scale factors (corresponding to an event) which we wish to use  
 1249 in **TRExFitter**. **TRExFitter** then takes these files as input, runs a maximum likelihood fit and produces relevant  
 1250 plots (e.g. expected number of events, pull plots) and statistical parameters (e.g. limit, significance,  $\mu_{best-fit}$ ).

### 1251 4.9.1 Fitting Procedure

1252 Using the **TRExFitter** framework, binned profile-likelihood fits are performed to determine the signal strength,  
 1253  $\mu_{tWZ} = \frac{\sigma_{obs}(tWZ)}{\sigma_{SM}(tWZ)}$ , of  $tWZ$  production. A fit across all regions in the tetralepton channel is performed to determine  
 1254 the sensitivity  $tWZ$  in this channel. In Section 4.10.2, a combined fit is performed across all regions in the trilepton  
 1255 and tetralepton channels to take advantage of the sensitivity of  $tWZ$  in both channels in order to further boost the  
 1256 sensitivity of  $tWZ$ . Plots shown prior to the fit are referred to as *pre-fit* and those shown after the fit are referred  
 1257 to as *post-fit*.

1258 To characterise the sensitivity and associated uncertainty of our measurement of  $\mu(tWZ)$ , we compute two metrics:  
 1259 the expected significance ( $Z_\mu^{exp}$ ) and the expected upper limit ( $\mu_{up}^{exp}$ ). In this context, the expected significance  
 1260 can be interpreted as, the probability that the measured signal is due to a background fluctuation. Larger values  
 1261 indicate lower probabilities and smaller values indicate higher probabilities. Particles physicists have adopted a  
 1262 standard to define the sensitivity necessary for evidence and discovery of a particular particle or phenomena. A  
 1263  $3\sigma$  (corresponding to a background fluctuation probability of  $\approx 10^{-3}$ ) significance is considered to be evidence  
 1264 for observation and a  $5\sigma$  (corresponding to a background fluctuation probability of  $\approx 10^{-7}$ ) is considered to be  
 1265 a discovery. The expected upper limit is a single-sided interval test statistic, associated with the POI in the  
 1266 maximum-likelihood fit ( $\mu(tWZ)$ , in our case). In this context, the expected upper limit can be understood in the  
 1267 following way: consider running an ensemble of MC toy experiments, each with their own confidence interval (a  
 1268 range of possible values for  $\mu(tWZ)$ ). An expected upper limit, at some fixed percentage  $x\%$  (or *confidence level*),  
 1269 can be determined from this ensemble. The expected upper limit tells us that,  $x\%$  of the toy MC experiment's  
 1270 confidence intervals will contain the true value of  $\mu(tWZ)$ . A commonly used percentage in particle physics is  
 1271 95%, which we adopt for this analysis. In particle physics, this is referred to as the *CLs Method* [39]. The CLs  
 1272 test statistic can be calculated 'brute force' by running these MC toy experiments, however this is very CPU  
 1273 intensive. Asymptotic formulae are able to describe the underlining CLs test statistic distributions under certain  
 1274 approximations [9]. Instead of running toy MC experiments, we use asymptotic formulae to perform the CLs  
 1275 method, which reduces computation time from several millions of CPU hours to several CPU minutes. The  
 1276 significance and upper limits which are calculated in this analysis are given a prefix of 'expected' to indicate that  
 1277 these are results from a blinded analysis.

1278 In the separate and combined fits, the *mixed data and MC* fit setup [80] is used. This is done to obtain the  
 1279 most accurate prediction of the expected results while keeping the signal regions blinded. For this setup, first a  
 1280 background-only fit to the control regions using real data is done to determine estimates of the nuisance parameters.  
 1281 Then these estimates are used to construct a modified ASIMOV dataset in the signal regions. Finally, the fit is  
 1282 performed using real data in the control regions and the aforementioned modified ASIMOV data-set in the signal  
 1283 regions.

1284 In these fits, the parameter of interest (POI) is  $\mu_{tWZ}$ . The POI is ultimately the quantity which we wish to  
 1285 measure and is set as a free parameter (unconstrained; can take any value in the fit). The nuisance parameters are  
 1286 assigned to the systematic uncertainties outlined in Section 4.8. Furthermore, a gamma ( $\gamma$ ) nuisance parameter  
 1287 for a bin is added to the likelihood function if the statistical uncertainty in the bin exceeds 0.1% of its nominal value.

1288 Pruning is done per sample and per region on the shape and normalisation uncertainties for samples. A  
 1289 sample's shape and normalisation nuisance parameter is pruned (removed from the limit/fit) if the fraction of  
 1290 the expected number of signal events to the expected number of total events (signal + background) is less than 0.01.

1291 An auto-binning algorithm, **TransfoD** [29], was used to define the binning. This aims to maximise  $\frac{signal}{background}$  in  
 1292 each bin. Furthermore, it aims to avoid defining bins with a low number of events.

## 1297 4.10 Results

1298 In the section, an expected upper limit and an expected significance are set on the cross section of  $tWZ$ . This is  
 1299 performed for the current analysis in the tetralepton channel as well as for a combined analysis across the trilepton  
 1300 and tetralepton channels. The trilepton analysis was performed as an independent study by Benjamin Warren  
 1301 (UCT) [81]. Note that throughout this section, all signal regions remain blinded.

### 1302 4.10.1 Tetralepton Channel

1303 In Figure 46 expected number of events for the variables used in the likelihood fit in each region are shown.

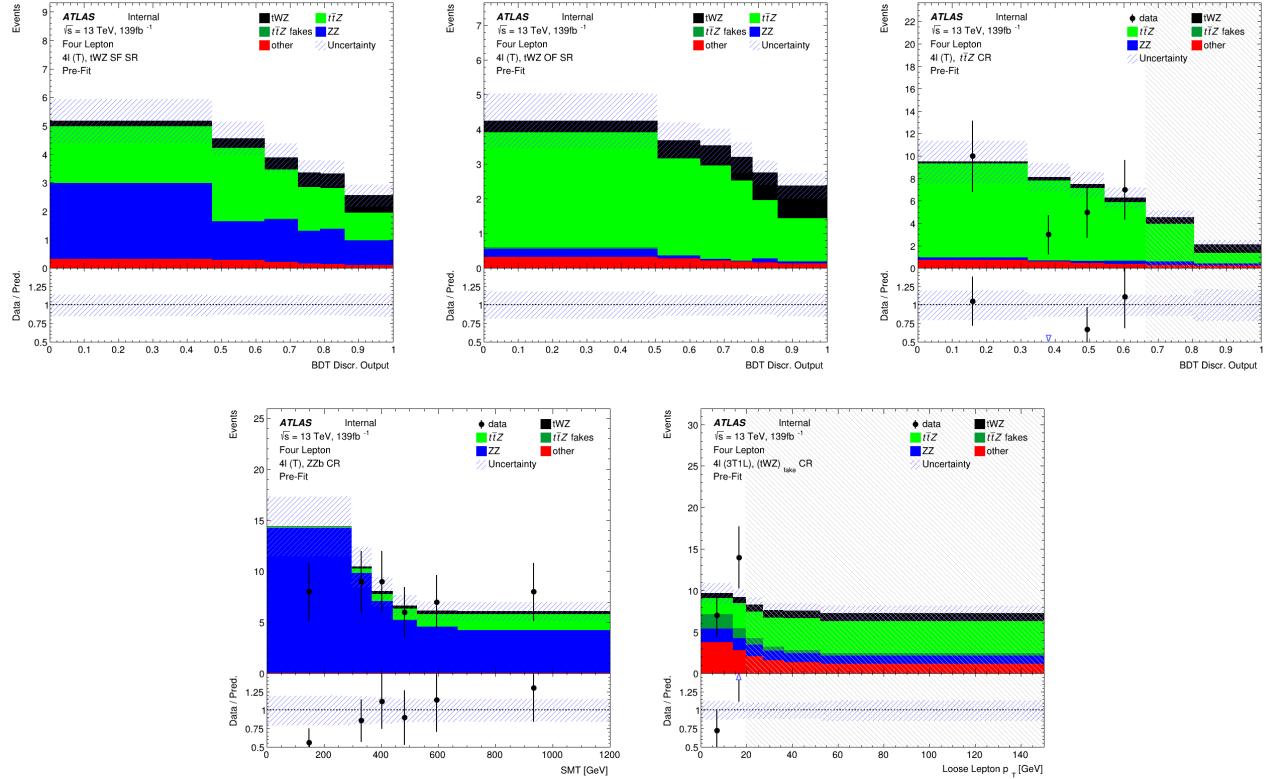


Figure 46: Pre-fit distributions (blinded) of variables used in the fit are shown. **Top left:** The event-level BDT Disc. Output in the  $tWZ$  SF (4T) SR region is shown. **Middle Top:** The event-level BDT Disc. Output in the  $tWZ$  SF (4T) SR region is shown. **Top right:** The event-level BDT Disc. Output in the  $t\bar{Z}$  CR region is shown. **Bottom left:**  $SMT = \sum p_T(\ell) + \sum p_T(jet) + E_T^{\text{miss}}$  in the  $ZZb$  CR is shown. **Bottom right:**  $p_T$ (loose lepton) in the  $(tWZ)_\text{fake}$  (3T1L) CR is shown. 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. The plots in the  $tWZ$  OF SR and  $tWZ$  SF SR are kept blinded by omitting the data points.

1304 In Figure 47 expected number of events for the variables used in the likelihood fit in each region are shown.

1305 In Table 12, the expected number of events (after the fit) for each sample in each region is shown.

1306 The expected upper limit of  $tWZ$  in the tetralepton channel is measured as,

$$\mu_{up}^{exp} = 1.61^{+2.35}_{-1.16} \quad (4.15)$$

1307 The expected significance of  $tWZ$  in the tetralepton channel is measured as,

$$Z_\mu^{exp} = 1.44\sigma \quad (4.16)$$

1308 The best-fit value of the signal strength,  $\mu(tWZ) = \frac{\sigma(tWZ)}{\sigma(tWZ)_{SM}}$ , from the likelihood fit is measured as,

$$\mu(tWZ) = 1.91^{+0.95}_{-0.82} \quad (4.17)$$

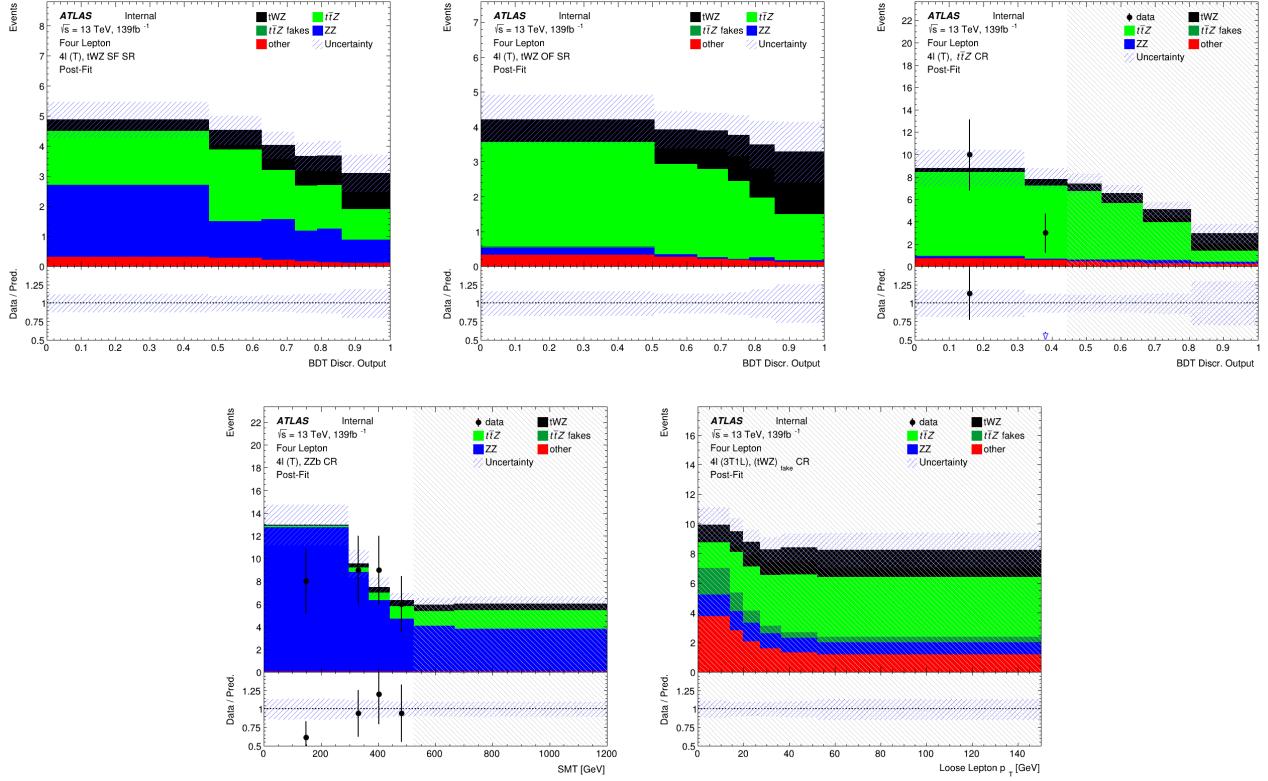


Figure 47: Post-fit distributions (blinded) of variables used in the fit are shown. **Bottom left:**  $SMT = \sum p_T(\ell) + \sum p_T(jet) + E_T^{\text{miss}}$  in the  $ZZb$  CR is shown. **Bottom right:**  $p_T$ (loose lepton) in the  $(tWZ)_{\text{fake}}$  (3T1L) CR is shown. The data is given by 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. The plots in the  $tWZ$  OF SR and  $tWZ$  SF SR are kept blinded by omitting the data points. **Top left:** The event-level BDT Disc. Output in the  $tWZ$  SF (4T) SR region is shown. **Middle Top:** The event-level BDT Disc. Output in the  $tWZ$  SF (4T) SR region is shown. **Top right:** The event-level BDT Disc. Output in the  $ttZ$  CR region is shown.

	$tWZ$ OF SR	$tWZ$ SF SR	$t\bar{t}Z$ CR	$ZZb$ CR	$(tWZ)_{fake}$ 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	$tt$	$6.00553e-06 \pm 0.02819e-06$	$0.252557 \pm 0.442116$	$0.273507 \pm 0.223201$	$6.00553e-06 \pm 0.02819e-06$
	$tZq$	$0.0827905 \pm 0.0398773$	$0.0756107 \pm 0.0354584$	$0.063585 \pm 0.0293325$	$0.05884 \pm 0.0244084$
	$t\bar{t}W$	$0.00668643 \pm 0.00792217$	$0.00279748 \pm 0.00287361$	$6.00553e-06 \pm 3.02819e-06$	$0.002306 \pm 0.00564349$
	$WZ$	$0.0442934 \pm 0.024156$	$0.0396511 \pm 0.0154282$	$0.0133471 \pm 0.0128199$	$0.0472562 \pm 0.0330315$
	$t\bar{t}t$	$0.000987164 \pm 0.000766266$	$0.00247481 \pm 0.00136945$	$0.0140869 \pm 0.00479496$	$6.00553e-06 \pm 3.02819e-06$
	$t\bar{t}\bar{t}$	$0.00934035 \pm 0.0080554$	$0.0107458 \pm 0.00849984$	$0.0571373 \pm 0.0204011$	$6.00553e-06 \pm 3.02819e-06$
	$t\bar{t}WW$	$0.0294618 \pm 0.0263174$	$0.029771 \pm 0.0195582$	$0.264364 \pm 0.0926252$	$0.0129431 \pm 0.0323803$
	$VVV(V = W/Z)$	$0.280643 \pm 0.0853411$	$0.191433 \pm 0.0586778$	$0.0697266 \pm 0.0225059$	$0.171142 \pm 0.0518102$
	$t\bar{t}H$	$0.846054 \pm 0.175495$	$0.669375 \pm 0.140107$	$1.96662 \pm 0.401199$	$0.150025 \pm 0.0353826$
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.

1309 The expected upper limit is in agreement with the extracted best-fit value on the signal strength, therefore no  
 1310 deviations from the SM cross section of  $tWZ$  is observed. Neither the  $3\sigma$  evidence nor the  $5\sigma$  discovery standards  
 1311 are reached for the expected significance. This is not surprising, given the low amount of events present in the  
 1312 tetralepton channel.

1313 In Figure 48, a ranking plot showing the impact of the systematic uncertainties on the POI,  $\mu(tWZ)$  is shown.  
 1314

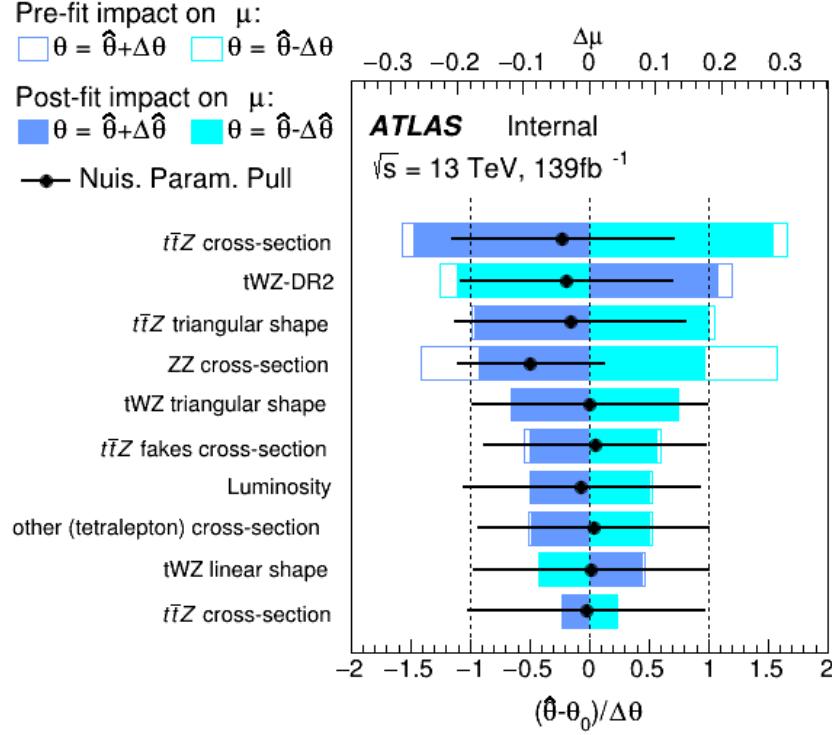


Figure 48: A ranking plot showing the impact (ordered from top to bottom via decreasing impact) of the systematic uncertainties (top 10) on the POI,  $\mu(tWZ)$ , in the tetralepton channel is shown.  $\hat{\theta}$  is the best-fit value of the nuisance parameter.  $\Delta\hat{\theta}$  and  $\Delta\theta$  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.

1315 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  
 1316 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 cross section of  $tWZ$  from the combined fit across all regions of  $tWZ$  in the tetralepton and trilepton channels.

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,

$$\mu_{up}^{exp} = 2.65^{+3.67}_{-1.91} \quad (4.18)$$

The expected significance of  $tWZ$  in the trilepton channel is measured as,

$$Z_\mu^{exp} = 0.75\sigma \quad (4.19)$$

The best-fit value of the signal strength,  $\mu(tWZ) = \frac{\sigma(tWZ)}{\sigma(tWZ)_{SM}}$ , from the likelihood fit is measured as,

$$\mu(tWZ) = 1.16^{+1.33}_{-1.30} \quad (4.20)$$

The expected upper limit is in agreement with the extracted best-fit value on the signal strength, therefore no deviations from the SM cross section of  $tWZ$  is observed.

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,

$$\mu_{up}^{exp} = 1.43^{+2.04}_{-1.03} \quad (4.21)$$

The expected significance of  $tWZ$  across both channels is measured as,

$$Z_\mu^{exp} = 1.61\sigma \quad (4.22)$$

The best-fit value of the signal strength,  $\mu(tWZ) = \frac{\sigma(tWZ)}{\sigma(tWZ)_{SM}}$ , from the likelihood fit is measured as,

$$\mu(tWZ) = 1.80^{+0.70}_{-0.65} \quad (4.23)$$

In Figure 49, 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 50, 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 49 and 50 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.

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).

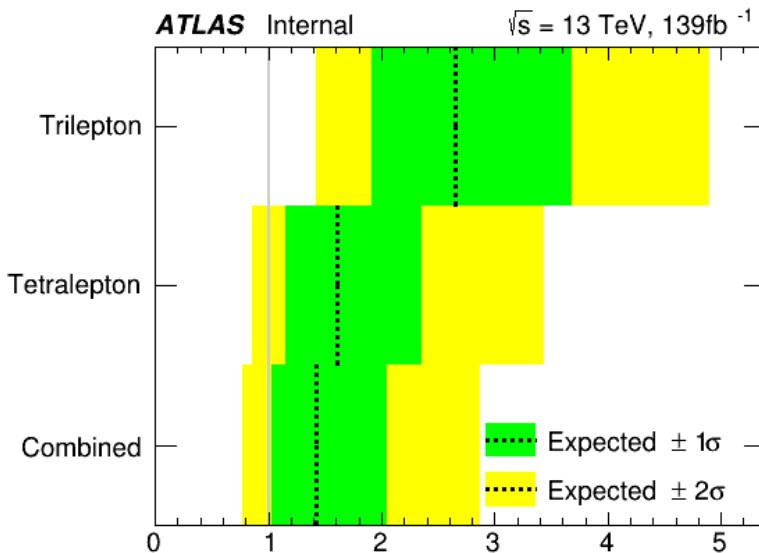


Figure 49: The expected upper limits of the trilepton channel, tetralepton channel and both channels combine are shown. The y-axis shows the channels in which the fitting procedure was performed. The expected limits are represented by the vertical dotted line. One- and two-  $\sigma$  uncertainty bands are shown in green and yellow respectively. The vertical grey line indicates when  $\mu(tWZ) = 1$ .

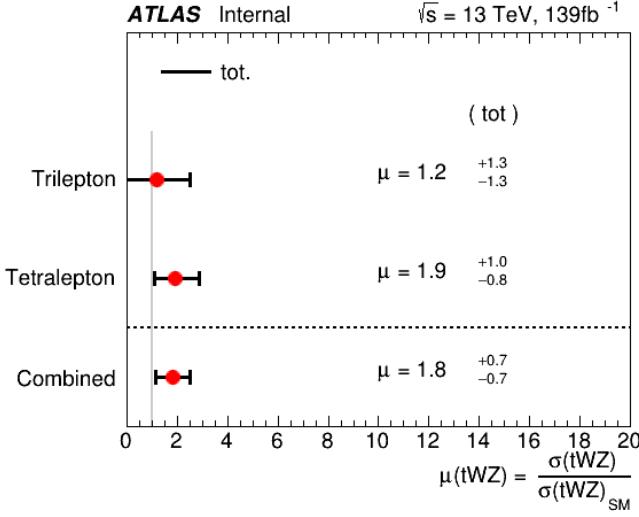


Figure 50: The best-fit values of  $\mu(tWZ)$  from the fit for the trilepton channel, tetralepton channel and both channels combined are shown. The y-axis shows the channels in which the fitting procedure was performed. The signal strength  $\mu(tWZ)$  is shown on the x-axis. The nominal signal strengths are represented by the red dots. The total uncertainty associated with the best-fit  $\mu(tWZ)$  value is shown by the black error bars.

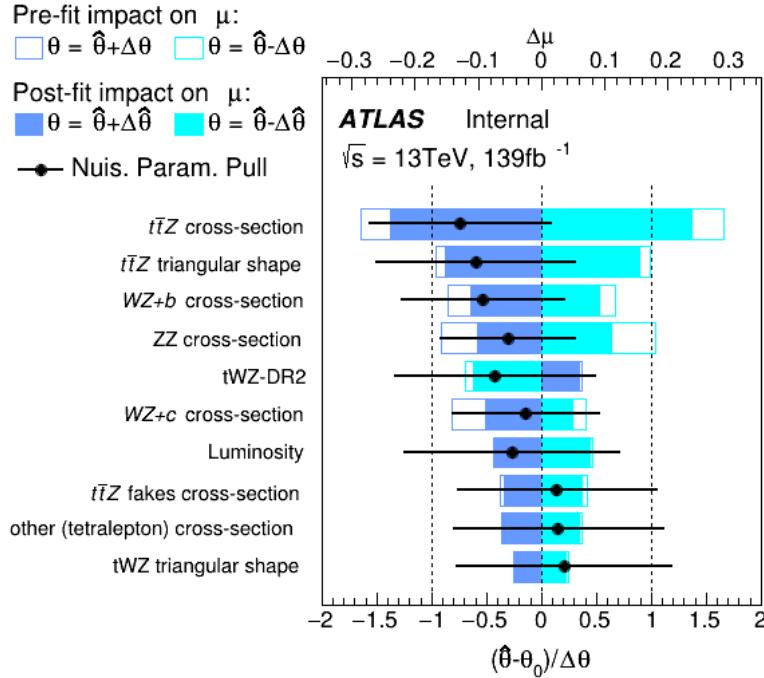


Figure 51: A ranking plot showing the impact of the systematic uncertainties (top 10) on the POI,  $\mu(tWZ)$ , in the combined fit across both the tri- and tetralepton channels is shown.  $\hat{\theta}$  is the best-fit value of the nuisance parameter.  $\Delta\hat{\theta}$  and  $\Delta\theta$  are 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.

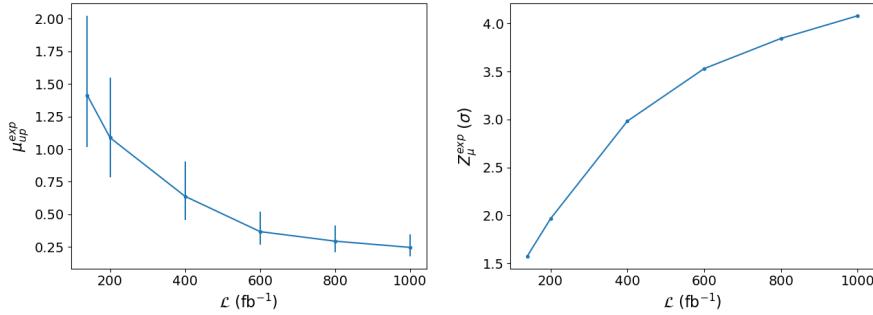


Figure 52: The expected upper limit (left) and significance (right) for combined fully blinded fits to the ASIMOV dataset across both the tri- and tetralepton channels for a range integrated luminosities are shown. **Left:** The integrated luminosity,  $\mathcal{L}$ , is shown on the x-axis. The expected 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.

1356 In Figure 51, a ranking plot showing the impact of the systematic uncertainties on the POI,  $\mu(tWZ)$ , in the  
 1357 combined fit across both the tri- and tetralepton channels is shown.

1358 Some nuisance parameters are pulled down from their nominal pre-fit values, however these are all within  $1\sigma$   
 1359 uncertainty and are thus relatively small. It is expected that the modelling uncertainties (shape and normalisa-  
 1360 tions) of the most dominant backgrounds (e.g.  $t\bar{t}Z$ ,  $ZZ$ ,  $WZ$ ) have relatively large impacts on  $\mu(tWZ)$ , since the  
 1361 uncertainty of the analysis is dominated by statistical uncertainty.

1362 The most important systematics are the cross sections of  $t\bar{t}Z$ ,  $ZZ$  and  $WZ + b$ , and shape modelling on  $t\bar{t}Z$   
 1363 ( $t\bar{t}Z$  triangular shape). The aforementioned nuisance parameters are significantly shifted down in the fit from its  
 1364 nominal value. The less important systematics are pulled in the fit from their nominal values, but to a much lesser  
 1365 degree than the cross section of those mentioned above. It is expected that the modelling uncertainties (shape and  
 1366 normalisations) of the most dominant backgrounds (e.g.  $t\bar{t}Z$ ,  $ZZ$  and  $WZ + b$ ) have relatively large impacts on  
 1367  $\mu(tWZ)$ , since the uncertainty of the analysis is dominated by statistical uncertainty.  
 1368

#### 1369 4.10.2.1 Projection to Higher Luminosity

1370 In this section we apply a fully blinded fit to the ASIMOV dataset for integrated luminosities larger than the  $139\text{fb}^{-1}$   
 1371 currently available from the ATLAS Full Run 2 dataset. This study gives us insight into the sensitivity of  $tWZ$   
 1372 which we could expect if we were to replicate the current analysis, given more data. Given the upgrades planned  
 1373 for the LHC and the ATLAS detector, we will soon expect a large increase in available  $pp$  collision data (reaching  
 1374  $\mathcal{L} = 3000\text{fb}^{-1}$  [32]). This study therefore gives us an idea of the sensitivity of  $tWZ$  which we can expect in the future.  
 1375

1376 In Figure 52, the expected upper limit and significance for combined fully blinded fits to the ASIMOV dataset  
 1377 across both the tri- and tetralepton channels for a range integrated luminosities are shown.

1378 As we increase the total integrated luminosity, the sensitivity of  $tWZ$  increases. The increase in sensitivity is shown  
 1379 by the decrease of the expected upper limit and the increase of the expected significance with increasing integrated  
 1380 luminosity. It can be noted that a  $3\sigma$  expected significance is achieved at  $\mathcal{L} \sim 400 \text{ fb}^{-1}$ , indicating evidence for  
 1381 observation of  $tWZ$ . The  $5\sigma$  significance 'discovery' standard is above our  $1000\text{fb}^{-1}$  luminosity range of study.  
 1382 Furthermore, these plots show that the sensitivity of  $tWZ$  is hindered by the low amount of events we observe,  
 1383 rather than the systematic effects. This is evident since increasing the luminosity, therefore increasing the number  
 1384 of events we see in the detector, directly causes a steady increase in sensitivity.

1385

# Chapter 5

1386

## Conclusion

1387 The search for  $tWZ$  production using  $139 \text{ fb}^{-1}$  of  $pp$  collision data at a centre-of-mass energy of  $\sqrt{s} = 13 \text{ TeV}$ ,  
 1388 recorded by the ATLAS experiment at CERN, has been presented. This thesis targeted the tetralepton final state  
 1389 channel. To further increase the sensitivity of  $tWZ$ , a combined analysis was done across the tetralepton and  
 1390 trilepton (studied in an independent analysis by Benjamin Warren (UCT) [81]) channels.

1391 Two SRs and Three CRs were defined. Two SRs, instead of one, were defined in order to suppress and constrain  
 1392 the  $ZZ$  background. The definition of the  $tWZ$  OF SR and the  $tWZ$  SF SR take advantage of the OSSF lepton  
 1393 pairs which originate from the decay of a  $Z$  boson, by requiring that non- $Z$  leptons in the event have the opposite  
 1394 and same flavours, for the  $tWZ$  OF SR and  $tWZ$  SF SR respectively. The resulting SRs successfully separate the  
 1395  $ZZ$  background, with the  $tWZ$  OF SR containing around 6% of the total  $ZZ$  background yield across both regions  
 1396 (implying that the remaining  $\sim 94\%$  is contained in the  $tWZ$  SF SR). The dominant background processes,  $t\bar{t}Z$   
 1397 and  $ZZ$  were constrained by the definition of  $t\bar{t}Z$  and  $ZZ$  CRs, respectively. The dominant source of fake leptons,  
 1398 originating from the  $t\bar{t}Z$  background, was constrained by the  $(tWZ)_{fake}$  CR, using the MC template method.  
 1399

1400 Two BDTs were implemented: an object-level BDT which aims to classify between  $\ell b$  systems coming from top  
 1401 quarks and an event-level BDT which aims to discriminate between  $tWZ$  and our major backgrounds,  $t\bar{t}Z$  and  
 1402  $ZZ$ . The output from the object-level BDT was converted to an event-level variable to be used as input to  
 1403 the event-level BDT. A kinematic reconstruction algorithm,  $2\nu\text{SM}$ , was used to reconstruct top quarks in order  
 1404 to discriminate between  $tWZ$  and  $t\bar{t}Z$ . The output from this algorithm was used as an input variable to the  
 1405 event-level BDT. The trained BDT was shown to discriminate well between signal and background events.

1406 Using a modified ASIMOV dataset in the SRs and real data in the CRs, a blinded maximum-likelihood fit was  
 1407 performed across all regions in the tetralepton channel. The best-fit value of the signal strength in the tetralepton  
 1408 channel was,

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

1409 with an expected significance of  $1.44\sigma$ . The expected upper limit on the signal strength of  $tWZ$  in the tetralepton  
 1410 channel was,

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

1411 Furthermore, a blinded maximum-likelihood fit was performed across all regions across the trilepton and tetralepton  
 1412 channels. The best-fit value of the signal strength across both the trilepton and tetralepton channels were,

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

1413 with an expected significance of  $1.61\sigma$ . The expected upper limit on the signal strength of  $tWZ$  across both the  
 1414 trilepton and tetralepton channels were,

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

1415 Although this result does not satisfy the  $3\sigma$  evidence nor the  $5\sigma$  discovery standards, this is the tightest ever  
 1416 constraint on the  $tWZ$  process. The results in this analysis are heavily statistically limited, it is therefore expected  
 1417 that future analyses of this process, using larger datasets (such as that from the HL-LHC), would significantly  
 1418 improve the results.

1421

# 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%7D](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] Georges Aad et al. “Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1. Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1”. In: *Eur. Phys. J. C* 77 (Mar. 2016), 490. 87 p. DOI: 10.1140/epjc/s10052-017-5004-5. arXiv: 1603.02934. URL: <https://cds.cern.ch/record/2138166>.
- [7] 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>.
- [8] Guido Altarelli. *The Higgs and the Excessive Success of the Standard Model*. 2014. arXiv: 1407.2122 [hep-ph].
- [9] 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).
- [10] 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>.
- [11] ATLAS Collaboration. “ATLAS  $b$ -jet identification performance and efficiency measurement with  $t\bar{t}$  events in  $pp$  collisions at  $\sqrt{s} = 13$  TeV”. In: *Eur. Phys. J. C* 79 (2019), p. 970. DOI: 10.1140/epjc/s10052-019-7450-8. arXiv: 1907.05120 [hep-ex].
- [12] ATLAS Collaboration. *Calibration of light-flavour  $b$ -jet mistagging rates using ATLAS proton–proton collision data at  $\sqrt{s} = 13$  TeV*. ATLAS-CONF-2018-006. 2018. URL: <https://cds.cern.ch/record/2314418>.
- [13] 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].

- [14] 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].
- [15] 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>.
- [16] 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].
- [17] 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].
- [18] 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>.
- [19] ATLAS Internal. *Electron Efficiencies for Analyses*. 2021. URL: <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/ElectronEfficienciesForAnalysis>.
- [20] ATLAS Internal. *Electron Efficiency Correlation Model*. 2021. URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/ElectronEfficiencyCorrelationModel>.
- [21] ATLAS Internal. *Jet Vertex Tagger for Run 2 in reco and analysis*. 2021. URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/JetVertexTaggerTool>.
- [22] ATLAS Internal. *Muon Efficiencies for Analyses*. 2021. URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/MuonEfficienciesForAnalysis>.
- [23] 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.
- [24] 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).
- [25] Olga Bessidskaia Bylund. *Measurement of ttZ and ttW 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: <https://cds.cern.ch/record/2211022>.
- [26] 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>.
- [27] 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>.
- [28] 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].
- [29] Thomas Calvet. *Automatic binning implementation in TTHTfitter - 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).
- [30] 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>.
- [31] CERN Twiki - TOP WG Summary Plots. URL: <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCTopWGSummaryPlots>.
- [32] 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>.
- [33] Masud Chaichian, Ricardo Gonzalez Felipe, and Katri Huitu. “On quadratic divergences and the Higgs mass”. In: *Physics Letters B* 363.1-2 (Nov. 1995), pp. 101–105. ISSN: 0370-2693. DOI: 10.1016/0370-2693(95)01191-r. URL: [http://dx.doi.org/10.1016/0370-2693\(95\)01191-R](http://dx.doi.org/10.1016/0370-2693(95)01191-R).

- [34] 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).
- [35] 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.
- [36] 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).
- [37] 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>.
- [38] I. Connolly. “Performance and calibration of b-tagging with the ATLAS experiment at LHC Run-2”. In: 2017.
- [39] 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>.
- [40] 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].
- [41] *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).
- [42] 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>.
- [43] *Errors in weighted histograms*. URL: <https://www.zeuthen.desy.de/~wischnew/amanda/discussion/wgterror/working.html>.
- [44] 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>.
- [45] Guido Fantini et al. *The formalism of neutrino oscillations: an introduction*. 2020. arXiv: 1802.05781 [hep-ph].
- [46] Steve Farrell. *Overlap Removal Tools, FTAG/Hbb Workshop*. URL: [https://indico.cern.ch/event/631313/contributions/2683959/attachments/1518878/2373377/Farrell\\_ORTools\\_ftaghbb.pdf](https://indico.cern.ch/event/631313/contributions/2683959/attachments/1518878/2373377/Farrell_ORTools_ftaghbb.pdf).
- [47] *GoodRunListsForAnalysisRun2*. URL: <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/GoodRunListsForAnalysisRun2>.
- [48] 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>.
- [49] 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.
- [50] Peter W. Higgs. “Spontaneous Symmetry Breakdown without Massless Bosons”. In: *Phys. Rev.* 145 (1966), pp. 1156–1163. DOI: 10.1103/PhysRev.145.1156.
- [51] *IFFTruthClassifier GitLab Repository*. URL: <https://gitlab.cern.ch/ATLAS-IFF/IFFTruthClassifier/-/tree/master>.
- [52] *IFFTruthClassifier Lepton Categories*. URL: <https://gitlab.cern.ch/ATLAS-IFF/IFFTruthClassifier/-/tree/master#3-details-about-the-lepton-categories>.
- [53] 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>.
- [54] F. Jegerlehner. “The hierarchy problem of the electroweak Standard Model revisited”. In: (May 2013).

- [55] 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>.
- [56] H. J. W. Kirsten. *Introduction to supersymmetry*. Singapore Hackensack, NJ: World Scientific, 2010. ISBN: 978-9814293426.
- [57] 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).
- [58] 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>.
- [59] Thomas McCarthy. *Macro developed to compare t/W/Z reconstruction performance (2ℓ, 3ℓ, 4ℓ)*. 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).
- [60] Thomas McCarthy and Florian Fischer. *Exploiting full/partial  $t\bar{t}$  reconstruction for background suppression in 2ℓ*. 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).
- [61] *MCTruthClassifier - ATLAS Twiki*. URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/MCTruthClassifier>.
- [62] 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>.
- [63] Ken Mimasu. *Top quark electroweak interactions*. 2021. arXiv: 2105.10261 [hep-ph].
- [64] 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).
- [65] *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>.
- [66] *MuonSelectionTool, ATLAS TWiki*. URL: <https://twiki.cern.ch/twiki/bin/view/Atlas/MuonSelectionTool>.
- [67] 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>.
- [68] *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/).
- [69] *Pileup jet recommendations*. URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/PileupJetRecommendations>.
- [70] *Recommended isolation working points*. URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/RecommendedIsolationWPs>.
- [71] *scikit-learn Documentation - Gradient Boosting Classifier*. URL: <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.GradientBoostingClassifier.html>.
- [72] *Scikit-Learn GradientBoostingClassifier Documentation*. URL: <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.GradientBoostingClassifier.html>.
- [73] Pedro Ferreira da Silva. *Top quark production at the LHC*. 2016. arXiv: 1605.05343 [hep-ex].
- [74] Jory Sonneveld. *Searches for physics beyond the standard model at the LHC*. 2019. arXiv: 1905.06239 [hep-ex].
- [75] Mark Thomson. *Modern particle physics*. Cambridge, United Kingdom New York: Cambridge University Press, 2013. ISBN: 9781107034266.
- [76] *TopRecoObjTwikiModel*. URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TopRecoObjTwikiModel>.
- [77] *Total Integrated Luminosity and Data Quality in 2015-2018 - LuminosityPublicResultsRun2*. URL: <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2>.

- 1611 [78] Daniel R. Tovey. “Transformation properties of the transverse mass under transverse Lorentz boosts at hadron  
1612 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>.
- 1614 [79] *TRExFitter*. URL: <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/TtHFitter>.
- 1615 [80] *TRExFitter: Mixed data and MC fit*. URL: <https://trexfitter-docs.web.cern.ch/trexfitter-docs/AdvancedTutorial2020/Mixed/>.
- 1617 [81] Benjamin Warren. “A search for tWZ production in the trilepton channel using Run 2 data from the ATLAS  
1618 experiment.” In: (2021).
- 1619 [82] Kenneth G. Wilson. “Confinement of quarks”. In: *Phys. Rev. D* 10 (8 Oct. 1974), pp. 2445–2459. DOI: 10.  
1620 1103/PhysRevD.10.2445. URL: <https://link.aps.org/doi/10.1103/PhysRevD.10.2445>.