



## ATLAS Note

ANA-TOPQ-2019-29-INT1

13th June 2021



Draft version 0.1

# Search for $tWZ$ production in the tri- and tetra-lepton channels using the full Run II ATLAS dataset

The ATLAS Collaboration<sup>1</sup>, J. Keaveney<sup>a</sup>, B. Warren<sup>b</sup>, J. Reich<sup>a</sup>, S. Yacoob<sup>a</sup>, James Keaveney<sup>a</sup>, Benjamin Warren<sup>a</sup>, Jake Reich<sup>a</sup>

<sup>a</sup>*University of Cape Town*

A search for the Standard Model  $tWZ$  process using the data collected by the ATLAS experiment in 2016 to 2018 at a centre-of-mass energy of  $\sqrt{s} = 13$  TeV is presented. Events containing either exactly 3 or 4 electrons or muons are selected with additional criteria based on the number of jets, the number of b-tagged jets, and the number of Z boson candidates used to define signal, control, and validation regions. In the tri-lepton channel, the large background from  $WZ$  production is suppressed by reconstructing hadronically-decaying W bosons with a BDT-based algorithm. The signal strength is extracted via a maximum-likelihood fit to multiple signal and control regions. The measured signal strength is  $\mu_{tWZ} = XX_{-ZZ}^{+YY}$  (stat.)  $\pm AA$  (syst.) fb =  $XX_{-ZZ}^{+YY}$  fb. This result is converted to a cross section for  $pp \rightarrow t\ell^+\ell^-q$  in the fiducial phase space region  $m_{\ell\ell} > 30$  GeV using a MC-based correction factor. The cross section result is  $XX_{-ZZ}^{+YY}$  (stat.)  $\pm AA$  (syst.) fb =  $XX_{-ZZ}^{+YY}$  fb, and is to be compared with the Standard Model expectation of  $XX_{-ZZ}^{+YY}$  (stat.)  $\pm AA$  (syst.) fb =  $XX_{-ZZ}^{+YY}$  fb calculated at NLO in QCD and a five-flavour PDF scheme.

Not reviewed, for internal circulation only

## 23 **Contents**

24	<b>1 Introduction</b>	4
25	<b>2 Theoretical context</b>	5
26	<b>3 Experimental strategy</b>	8
27	3.1 General considerations	8
28	3.2 Data and simulation	8
29	3.3 Data samples	9
30	3.4 Monte Carlo samples	9
31	<b>4 Object Definitions</b>	13
32	4.1 Electrons	13
33	4.2 Muons	14
34	4.3 Jets	14
35	4.4 Missing transverse energy	15
36	4.5 Overlap removal	15
37	<b>5 The Trilepton Channel</b>	16
38	5.1 Signal Topology & General Considerations	16
39	5.2 Baseline selections	16
40	5.3 Signal and Control Regions	17
41	5.4 Fake lepton estimation	26
42	5.5 Hadronically-decaying W Boson identification	27
43	5.6 Event-level GBDT	36
44	<b>6 The Tetralepton Channel</b>	43
45	6.1 Signal Topology & General Considerations	43
46	6.2 Baseline selections	45
47	6.3 Signal and Control Regions	49
48	6.4 Fake Lepton Estimation (MC Template Method)	65
49	6.5 Analysis Strategy	67
50	<b>7 Systematics</b>	78
51	7.1 Experimental uncertainties	78
52	7.2 Theoretical uncertainties	79
53	7.3 Generic shape systematics	80
54	<b>8 Results</b>	82
55	8.1 Fitting Procedure	82
56	8.2 Tetralepton Channel	83
57	8.3 Trilepton Channel	88
58	8.4 Combined Tri- and Tetralepton Channels	92
59	<b>9 Summary and Conclusions</b>	95

---

60	<b>Appendices</b>	<b>101</b>
61	<b>A Tetralepton Control Plots</b>	<b>101</b>
62	A.1 $tWZ$ OF SR	101
63	A.2 $tWZ$ SF SR	107
64	A.3 $t\bar{t}Z$ CR	114
65	A.4 $ZZb$ CR	121
66	A.5 $(tWZ)_{\text{fake}}$ CR	129
67	<b>B TRexFitter Configuration</b>	<b>136</b>
68	B.1 Trilepton Channel	136
69	B.2 Tetralepton Channel	136
70	B.3 Combined	140

Not reviewed, for internal circulation only

## 71    1 Introduction

72    The production of a single top quark in association with a  $Z$  boson and  $W$  boson ( $tWZ$ ) from proton  
73    collisions is a rare process that is predicted by the Standard Model (SM). Before the work presented in this  
74    note was undertaken, the only experimental effort to investigate this process was the masters thesis [1].

## 75 2 Theoretical context

76 The production of a single top quark in association with a  $Z$  boson and  $W$  boson ( $tWZ$ ) in a proton collision  
 77 is a rare process predicted by the Standard Model (SM). The modelling of this process at NLO in QCD  
 78 has been explored in [2] and the sensitivity of the inclusive  $tWZ$  cross section to the coefficients of the  
 79 Standard Model Effective Field theory (SMEFT) has been quantified in [3]. Before the work presented in  
 80 this note was undertaken, the only experimental effort to investigate this process was a masters thesis [1].

81 The eight Feynman diagrams that contribute to  $tWZ$  production at leading order (LO) in QCD, are shown  
 82 in figure 1. As the  $Z$  boson may be radiated from the initial-state bottom quark, the final-state top quark, or  
 83 the final-state  $W$  boson the process already exhibits complex phenomenology at LO by simultaneously  
 84 embedding the  $b$ - $Z$ ,  $t$ - $Z$  and  $W$ - $Z$  electroweak couplings. At next-to-leading order (NLO) in QCD, an  
 85 overlap develops between the diagrams of  $tWZ$  process and those that would be normally attributed to the  
 86  $t\bar{t}Z$  process.

87 In order to avoid double-counting of amplitudes when combining simulations of the  $t\bar{t}Z$  and  $tWZ$  processes,  
 88 several theoretical techniques exist to suppress the component of the  $tWZ$  calculation that overlaps with  
 89 that of  $t\bar{t}Z$  [2]. The methods are known as Diagram Removal and Diagram Subtraction and are described  
 90 in the context of the overlap between  $tW$  and  $t\bar{t}$  production in [4]. The Diagram Removal method, of which  
 91 there are two variants, were extended to  $tWZ$  in [2] and are utilised in the Monte Carlo (MC) samples used  
 92 in this analysis.

93 In the Standard Model Effective Field Theory (SMEFT), the number of diagrams contributing to  $tWZ$  at  
 94 LO in QCD increases to 56 due to the 13 different dimension-6 operators that may be inserted in the  $tWZ$   
 95 production diagrams. The diagrams containing a dimension-6 operator can be categorised in five groups:

- 96    1. diagrams embedding the  $bW \rightarrow tZ$  vertex
- 97    2. diagrams embedding the heavy quarks' neutral EW coupling vertices  $t \rightarrow tZ$ ,  $b \rightarrow bZ$
- 98    3. diagrams embedding the heavy quarks' charged EW coupling vertices  $tWb$
- 99    4. diagrams embedding the heavy quark gluon gluon vertices  $g_{tt}$ ,  $g_{bb}$
- 100    5. diagrams embedding the triple-gauge-coupling vertex  $WWZ$

101 As a result, the cross section of  $tWZ$  production is sensitive to numerous Wilson coefficients that parametrise  
 102 the effects of the corresponding operators in the SMEFT. This sensitivity is quantified in [3] for 7 of the  
 103 13 operators and is comparable to the sensitivity of the  $tZq$  process which has been well explored in the  
 104 phenomenological literature [5]. This sensitivity to numerous SMEFT coefficients yields the potential  
 105 for an experimental constraint on the  $tWZ$  cross section to be an impactful input to future *global* fits of  
 106 SMEFT coefficients to top quark, Higgs boson and electroweak data.

Not reviewed, for internal circulation only

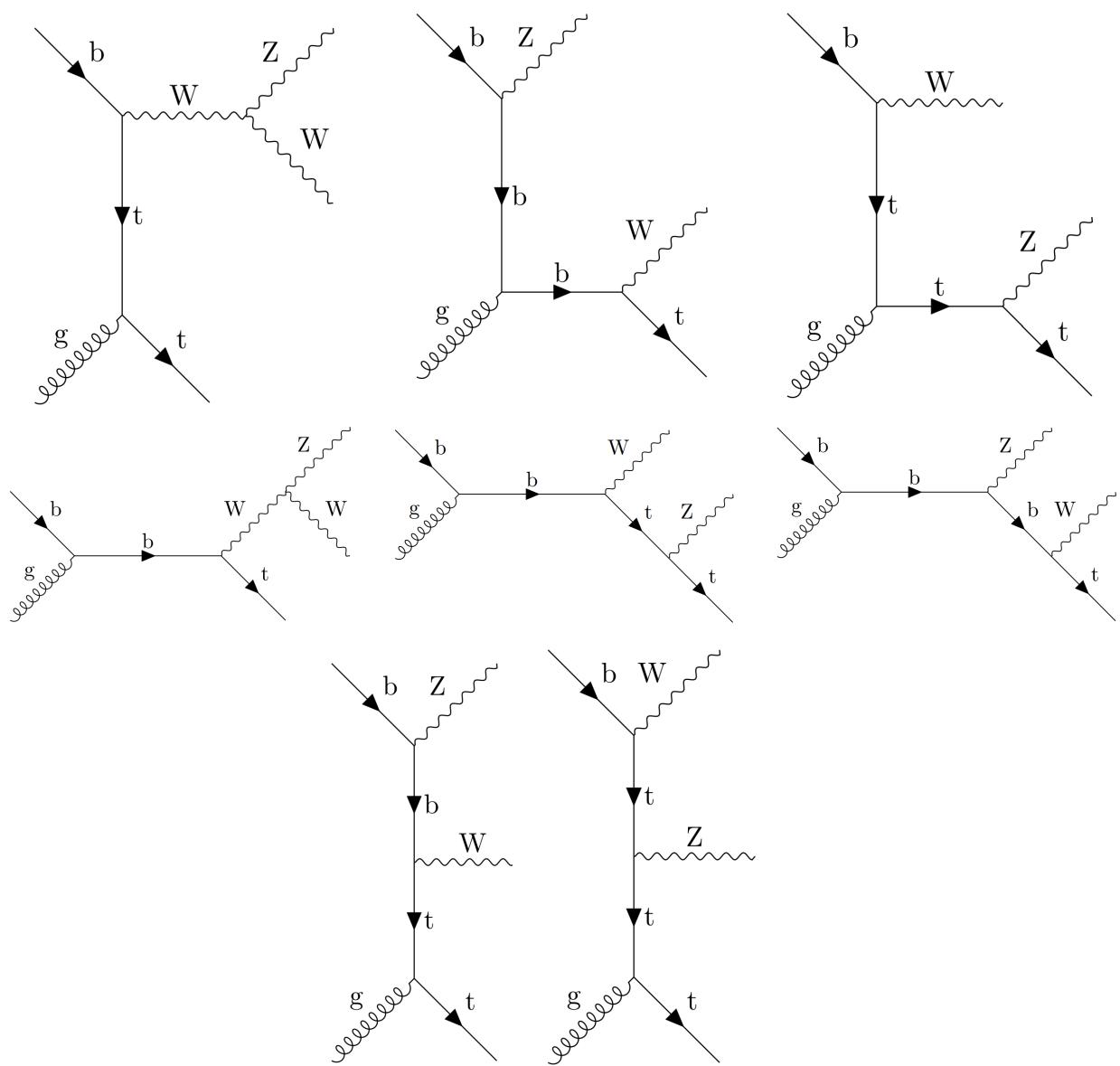


Figure 1: Feynman graphs to calculate the lowest order amplitudes.

Not reviewed, for internal circulation only

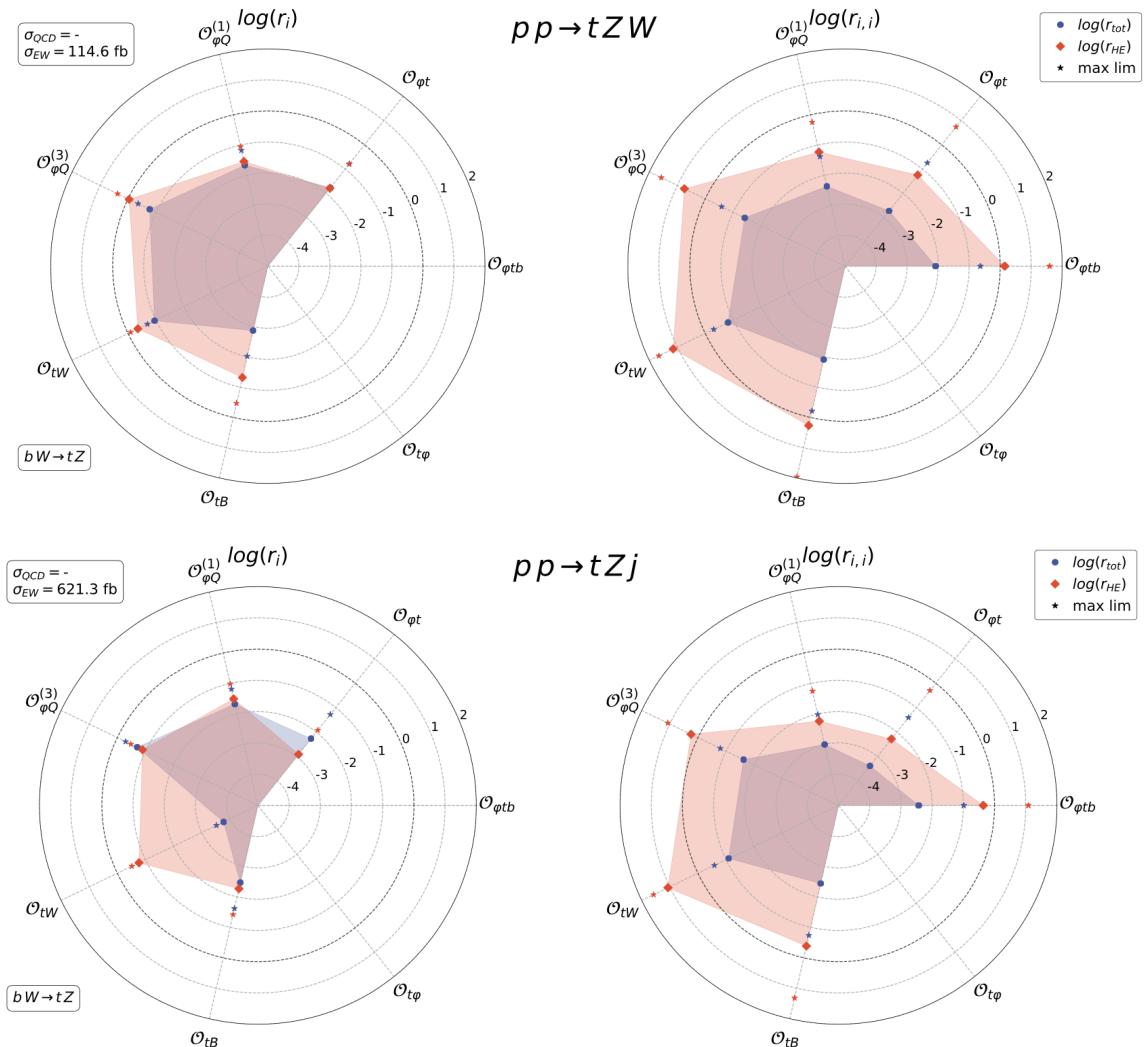


Figure 2

## 107 3 Experimental strategy

### 108 3.1 General considerations

109 Due to the small cross section and relatively large backgrounds of the SM  $tWZ$  process, only two final-state  
 110 topologies are experimentally feasible for a cross section measurement.

111 **Trilepton** In this channel the  $Z$  boson decays into a same-flavour charged lepton pair ( $e$  or  $\mu$ ), while either  
 112 the additional  $W$  boson or top quark decay produces a third charged lepton ( $e$  or  $\mu$ ). In either case this  
 113 topology contains a hadronically-deacying  $W$  boson. The backgrounds in this channel are numerous  
 114 with the  $WZ + b$  process providing the dominant contribution.

115 **Tetralepton** In this channel the  $Z$  boson decays into a same-flavour charged lepton pair ( $e$  or  $\mu$ ), while  
 116 both the additional  $W$  boson and top quark decay produce a charged lepton ( $e$  or  $\mu$ ). The smaller  
 117 branching ratio of this channel is compensated by a large reduction in backgrounds of which the  $t\bar{t}Z$   
 118 provides by far the dominant contribution.

### 119 3.2 Data and simulation

120 The first step in the prociessign chain of this analysis is the production of a dedicated set of ntuples common  
 121 to both the trilepton and tetralepton channels. These ntuples were prodced using AnalysisBase version  
 122 21.2.162 using a set of configuration files that can be found [6]. In the linked repository at [6], separate  
 123 configuration files for each era (MC16a, MC16d and MC16e) that differ only by the pile-up reweighting  
 124 procedure are provided, each with the tag *epping* to identify the latest ntuple version.

125 In Table ?? a table showing all MC simulation samples for the nominal analyses in both channels is  
 126 shown.

Table 1: Overview of the commons events selections applied for both the trilepton and tetralepton channels at the ntuple level

INITIAL, GRL GOODCALO, PRIVTX, RECO_LEVEL
2016: $296939 \leq \text{RUN\_NUMBER} \leq 311481$
2017: $324320 \leq \text{RUN\_NUMBER} \leq 341649$
2018: $\text{RUN\_NUMBER} \geq 348197$
$N_e \text{ or } \mu \geq 3$

Table 2: Overview of the common trigger selections applied for both the trilepton and tetralepton channels at the ntuple level for 2015 (labelled with \*) and 2016 / MC16a, 2017 / MC16d, 2018 / MC16e

Single-muon triggers	HLT_mu20_loose_L1MU15*, HLT_mu26_ivarmedium, HLT_mu50
Single-electron triggers	HLT_e24_lhmedium_L1EN20VH*, HLT_e26_lhtight_nod0_ivarloo, HLT_e60_lhmedium_nod0, HLT_e120_lhloose*, HLT_e140_lhloose_nod0
Di-muon triggers	HLT_mu18_mu8noL1*, HLT_mu22_mu8noL1
Di-electron triggers	HLT_2e12_lhloose_L12EM10VH1*, HLT_2e17_lhvloose_nod0
Electron-muon triggers	HLT_e17_lhloose_nod0_mu14

This section describes the data and simulated MC samples used by the analysis and the framework employed to process these samples. For both data, as well as simulated samples, TOPQ1\_DxAOD derivations are used. Most of the studies documented in this note are performed using flat ROOT Ntuples produced on the grid with a framework based on the EventSaverFlatNtuple framework. They contain the relevant object information (see Section ??) and are produced for signal and all relevant background samples.

### 3.3 Data samples

For this analysis, the full Run 2  $\sqrt{s} = 13$  TeV  $pp$  collision data, collected during the years 2015–2018 is used, corresponding to an integrated luminosity of  $139\text{ fb}^{-1}$ . Only events from lumi-blocks<sup>1</sup> in which LHC beams were stable and all ATLAS detectors were operational are selected. The list of suitable runs and lumi-blocks is summarised in the official Good Runs Lists for 2015–2018 data, as specified in Ref. [twiki-GRL]. Table 3 shows the integrated luminosities corresponding to the individual datasets per year.

Year	Data-taking periods	Luminosity [ $\text{pb}^{-1}$ ]
2015	D3–J	3219.6
2016	A–L	32988.1
2017	B–F, H, I, K	44307.4
2018	B–Q	58450.1

Table 3: The integrated luminosities for the years 2015–2018.

The uncertainty in the combined 2015–2018 luminosity is 1.7% ( $\pm 2.4\text{ fb}^{-1}$ ) and is derived from the calibration of the luminosity scale using  $x$ – $y$  beam-separation scans, following a methodology similar to that detailed in Ref. [7], and using the LUCID-2 detector [8] for the baseline luminosity measurements. The corresponding tag of the luminosity calibration software is `OfLLumi-13TeV-010` for 2015–2018 data [twiki-Luminosity]. The average number of interactions per bunch crossing (“pile-up”) ranges from 0.5 to  $\approx 80$  and has mean of around 34.2 [luminosityRunII].

### 3.4 Monte Carlo samples

Simulated Monte Carlo samples are used to predict the SM background. All MC samples used to model the contributions from the various SM processes, their cross section precisions and their settings (parton shower, PDF and MC tune) are summarised in Table 4. The samples are generated using a 25 ns bunch spacing configuration. A common PMG file [twiki-PMG] is used to obtain the cross section for the normalisation of each process, the generator filter efficiency and the  $k$ -factors. Different parton distribution functions (PDF) sets are used for the MC samples [twiki-PDF].

The effect of pile-up was modelled by overlaying the hard-scattering event with simulated minimum-bias events generated with PYTHIA8.186 [Sjostrand:2007gs] using the NNPDF2.31o [Ball:2012cx] set of PDFs and the A3 set of tuned MC parameters [14]. Separate MC production campaigns are used to model the different pile-up distributions observed in data for the years 2015/16, 2017 and 2018. The simulated

<sup>1</sup> Lumi-blocks denote specific time-intervals where the luminosity is assumed to be constant.

Process	Generator	Cross section calc.	Parton shower	PDF	MC
$t\bar{t}$	POWHEG [9]	NLO	PYTHIA8.230 [10]	NNPDF3.0nlo [11]	A14
Z + jets	SHERPA 2.2.1 [Bothmann:2019yzt]	NLO	SHERPA	NNPDF3.0nnlo	SHERPA
$t\bar{t}W$	MADGRAPH5_AMC@NLO 2.3.3 [13]	NLO	PYTHIA8.210	NNPDF3.0nlo	A14
	SHERPA 2.2.1	NLO	SHERPA	NNPDF3.0nnlo	SHERPA
	SHERPA 2.2.8	NLO	SHERPA	NNPDF3.0nnlo	SHERPA
$t\bar{t}Z$	MADGRAPH5_AMC@NLO 2.3.3	NLO	PYTHIA8.210	NNPDF3.0nlo	A14
$WZ/ZZ+$ jets	SHERPA 2.2.2	NLO	SHERPA	NNPDF3.0nnlo	SHERPA
$t\bar{t}H$	MADGRAPH5_AMC@NLO 2.6.0	NLO	PYTHIA8.230	NNPDF3.0nlo	A14
$tZq$	MADGRAPH5_AMC@NLO 2.3.3	NLO	PYTHIA8.230	NNPDF3.0nnlo	A14
$tWZ$	MADGRAPH5_AMC@NLO 2.3.3	NLO	PYTHIA8.235	NNPDF3.0nnlo	A14
$t\bar{t}WW$	MADGRAPH5_AMC@NLO 2.2.2	LO	PYTHIA8.186 [Sjostrand:2007gs]	NNPDF2.3lo [Ball:2012ex]	A14
$H + W/Z$	PYTHIA8.186	LO	PYTHIA8	NNPDF2.3lo	A14
$VVV (V = W/Z)$	SHERPA 2.2.2	NLO	SHERPA	NNPDF3.0nnlo	SHERPA
$t\bar{t}t, \bar{t}\bar{t}\bar{t}$	MADGRAPH5_AMC@NLO 2.2.2	LO	PYTHIA8.186	NNPDF2.3lo	A14

Table 4: List of Monte Carlo event generators and their settings used to predict the contributions from Standard Model processes.

Not reviewed, for internal circulation only

samples are reweighted to reproduce the observed distribution of the average number of collisions per bunch crossing using the officially recommended PileupReweightingTool package [twiki-PRW]:

- mc16a config files with 2015 + 2016 lumicalc files,
- mc16d config files with 2017 lumicalc files,
- mc16e config files with 2018 lumicalc files.

For mc16d and mc16e so-called “actual mu” files are used in addition to MC pile-up reweighting config files and the “lumicalc” files. Each distribution is normalised with the separate luminosity. Following the ATLAS pile-up reweighting recommendations, the  $\langle \mu \rangle$  value is scaled by 1/1.03, based on studies of the number of vertices as a function of  $\langle \mu \rangle$  as well as the results from inelastic cross-section measurement.

The simulation of detector effects is performed with either full ATLAS detector simulation [15] based on the GEANT [Geant] framework or a fast simulation (ATLFAST II) using a parameterisation of the performance of the electromagnetic and hadronic calorimeters and GEANT for the other detector components [16, 17].

The data and Monte Carlo samples are skimmed into TOPQ1\_DxAOD derivations with the following ATLAS software caches. The derivation software used corresponds always to a p-tag  $\geq 4346$ .

- Data: ATLAS derivation cache 21.2.111.0 (p-tag 4345)
- MC: Derivation cache 21.2.111.0 (p-tag 4344/4346)<sup>2</sup>

The production of  $t\bar{t}$  events is simulated with the POWHEG [9] generator providing matrix element (ME) calculations at NLO in the strong coupling constant ( $\alpha_S$ ) with the NNPDF3.0nlo [11] PDF set. The mass of the top quark is set to 172.5 GeV and the  $h_{\text{damp}}$  parameter, which controls the matching in POWHEG and regulates the high- $p_T$  radiation against which the  $t\bar{t}$  system recoils, is set to 1.5 times the nominal top quark mass. The events are interfaced with PYTHIA8.230 [10] for the parton shower, underlying event and

<sup>2</sup> Different derivation caches may be used for MC, since we decided to use the always the latest p-tag for each sample (and not samples used by this analysis were produced within the same derivation campaign)

hadronisation, using the A14 [12] set of tuned parameters and the NNPDF2.3lo PDF set. The  $t\bar{t}$  cross section is normalised to NNLO in QCD, including the resummation of next-to-next-to-leading logarithmic (NNLL) soft-gluon terms (NNLO+NNLL) [Melnikov:2011ta, Czakon:2012pz, LHCTopWGtbarXsec]. The decays of  $b$ - and  $c$ -hadrons are simulated with the EvtGen 1.6.0 [Lange:2001uf] program.

$Z + \text{jets}$  production is simulated with the SHERPA 2.2.1 generator [Bothmann:2019yzt]. In this setup, multiple MEs are matched and merged with the SHERPA parton shower based on the Catani-Seymour dipole factorisation scheme [18, 19] using the MEPS@NLO prescription [20–23]. The virtual QCD correction for MEs at NLO accuracy are provided by the OPENLOOPS library [24, 25]. Samples are generated using the NNPDF3.0nnlo PDF set. Up to two additional partons are included in the NLO ME, and three or four additional partons are included at LO in QCD. The sample cross sections are normalised to NNLO predictions. To have a sufficient number of simulated events in different kinematic regimes, the  $Z + \text{jets}$  samples are split according to the  $p_T$  of the  $Z$  boson and the  $H_T$  of the event, introducing a cut at generation level and producing samples for different slices in  $\max(p_T, H_T)$ . Furthermore, the samples are generated by applying different filters for  $b$ - and  $c$ -hadrons to select the flavour composition of the jets produced in association with the  $Z$  boson [twiki-Zfilter].

The  $t\bar{t}W$  and  $t\bar{t}Z$  processes are modelled using the MADGRAPH5\_AMC@NLO 2.3.3 [13] generator which provides a ME calculations at NLO in  $\alpha_S$  with the NNPDF3.0nlo PDF set. The functional form of the renormalisation and factorisation scales ( $\mu_R$  and  $\mu_F$ ) are set to  $\mu_{R,F} = 0.5 \times \sum_i \sqrt{m_i^2 + p_{T,i}^2}$ , where  $i$  runs over all final-state particles generated from the ME calculation. For the  $t\bar{t}Z$  process, the  $t\bar{t}\gamma^*$  contribution and the  $Z/\gamma^*$  interference is included, such that they feature events with dilepton invariant masses ( $m_{\ell\ell}$ ) down to 5 GeV. Top-quark decays are simulated at LO using the MADSPIN program [26, 27] to preserve all spin correlations. The events are interfaced with PYTHIA8.210 for the parton shower and hadronisation, using the A14 tune and the NNPDF3.0nlo PDF set. The decays of  $b$ - and  $c$ -hadrons are simulated using EvtGen 1.2.0.

Alternative  $t\bar{t}W$  samples are simulated with SHERPA 2.2.1 and SHERPA 2.2.8 generators at NLO accuracy. In both cases, dynamic  $\mu_R$  and  $\mu_F$  scales defined similarly to the nominal samples are used. The default SHERPA parton shower setup is used along with the NNPDF3.0nnlo PDF set. The samples are simulated using the MEPS@NLO prescription with up to one extra parton at NLO and with a merging scale of 30 GeV.

Diboson processes featuring the production of three charged leptons and one neutrino or four charged leptons (denoted as  $WZ$  or  $ZZ + \text{jets}$ , respectively) are simulated using the SHERPA 2.2.2 generator, with a similar setup as described for  $Z + \text{jets}$ . Events with up to one extra parton are simulated at NLO and two or three partons at LO precision.

The production of a single top quark (or antiquark) in association with a  $Z$  boson and one extra parton ( $tZq$ ) is generated using the MADGRAPH5\_AMC@NLO 2.3.3 generator at NLO with the NNPDF3.0nnlo PDF set. The events are interfaced with PYTHIA8.230 using the A14 tune and the NNPDF2.3lo PDF set. The  $tZq$  sample also includes off-shell  $Z$  decays to dilepton pairs with invariant masses in the range  $m_{\ell\ell} > 30$  GeV. Single top quark (antiquark) production in association with both a  $W$  and a  $Z$  boson ( $tWZ$ ) is simulated at NLO with MADGRAPH5\_AMC@NLO 2.2.2 and the NNPDF3.0nnlo PDF set, using PYTHIA8.235 for the parton-shower simulation. The interference between  $t\bar{t}Z$  and  $tWZ$  is removed following a diagram removal approach referred to as DR1 scheme [Demartin:2016axk].

Events featuring the associated production of a  $t\bar{t}$  pair with Higgs boson ( $t\bar{t}H$ ) are generated using NLO MEs in MADGRAPH5\_AMC@NLO 2.6.0 with the NNPDF3.0nlo PDF set. The samples are showered with

220 PYTHIA8.230 using the A14 tune. MC samples featuring Higgs production in association with a  $W$  or  $Z$   
221 boson ( $H+W/Z$ ) are generated at LO with PYTHIA8.186 using the A14 tune and the NNPDF2.3lo PDF  
222 set.

223 The production of three or four top quarks ( $t\bar{t}t$  and  $t\bar{t}t\bar{t}$ ) and the production of a  $t\bar{t}$  pair with two  $W$   
224 bosons ( $t\bar{t}WW$ ) are simulated at LO using MADGRAPH5\_AMC@NLO 2.2.2 interfaced to PYTHIA8.186  
225 with the A14 tune and the NNPDF2.3lo PDF set. Fully leptonic triboson processes ( $WWW$ ,  $WWZ$ ,  $WZZ$   
226 and  $ZZZ$ ) with up to six leptons in the final states are simulated with SHERPA 2.2.2 and the NNPDF3.0nlo  
227 PDF set. Final states with no additional partons are calculated at NLO, whereas final states with one, two  
228 and three additional partons are calculated at LO.

Not reviewed, for internal circulation only

## 229 4 Object Definitions

230 This section presents the definitions of the physics objects used in the analysis, namely the leptons (electrons  
 231 or muons), jets and missing transverse energy. Unless otherwise stated, the object recommendations are  
 232 used, as they are implemented in the `AnalysisBase` framework. The `AnalysisBase` release used for the  
 233 production of the analysis Ntuples is 21.2.162 [28].

234 **4.1 Electrons**

235 The requirements for pre-selected electrons (for overlap removal, see Section 4.5) and electrons used in  
 236 the analysis are summarised in Table 5 (left). Only electrons satisfying  $p_T > 7 \text{ GeV}$  and  $|\eta| < 2.47$  are  
 237 considered. A likelihood-based electron identification (ID) is used, as recommended by the egamma  
 238 CP group [29, 30]. Different working points (WPs) are defined and supported by the egamma CP  
 239 group. For pre-selected electrons, the `LooseAndBLayerLH` ID is used, whereas for electrons used in  
 240 the analysis selection a `TightLH` ID is used [31]. Furthermore, a requirement on the electron isolation,  
 241 corresponding to the `PLVTight` isolation WP [32] is applied. This isolation WP is defined by a multivariate  
 242 likelihood discriminant combining shower shape and track information to distinguish prompt electrons from  
 243 electron candidates from hadronic jets, photon conversions and heavy-flavor (HF) hadron decays (fake/non-  
 244 prompt).

245 The reconstructed track associated with the electron must pass the requirements  $|z_0 \sin(\theta)| < 0.5 \text{ mm}$  and  
 246  $|d_0|/\sigma(d_0) < 5$ , where  $z_0$  describes the longitudinal impact parameter (IP) relative to the reconstructed  
 247 primary vertex,<sup>3</sup>  $d_0$  is the transverse impact parameter relative to the beam axis, and  $\sigma(d_0)$  is the uncertainty  
 248 on  $d_0$ . Electron candidates are excluded if their calorimeter clusters lie within the transition region  
 249 between the barrel and the endcap of the electromagnetic calorimeter,  $1.37 < |\eta| < 1.52$ . Associated scale  
 250 factors (SFs) for electron reconstruction, identification, and isolation are applied in MC, to correct for the  
 251 efficiency differences between data and simulation if applying these requirements [33].

	<b>Pre-selected electron</b>	<b>Pre-selected muon</b>
Acceptance	$p_T > 7 \text{ GeV}$ , $ \eta^{\text{clust}}  < 2.47$ except $1.37 <  \eta^{\text{clust}}  < 1.52$	$p_T > 7 \text{ GeV}$ , $ \eta  < 2.5$
Identification	<code>LooseAndBLayerLH</code>	<code>Medium</code>
Overlap removal	see Section 4.5	
Impact parameter	$ d_0 /\sigma(d_0) < 5$ $ z_0 \sin(\theta)  < 0.5 \text{ mm}$	$ d_0 /\sigma(d_0) < 3$ $ z_0 \sin(\theta)  < 0.5 \text{ mm}$
	<b>Electron</b>	<b>Muon</b>
Identification	<code>TightLH</code>	<code>Medium</code>
Isolation	<code>PLVTight</code>	<code>PLVTight</code>

Table 5: Summary of the electron and muon object definitions: the selection requirements for actual electrons/muons are applied in addition to the pre-selected objects used for the overlap removal.

<sup>3</sup> The primary vertex is defined as the vertex with the highest scalar sum of the squared transverse momenta of associated tracks with  $p_T > 400 \text{ MeV}$ .

## 252 4.2 Muons

253 The muon selection criteria are summarised in Table 5 (right). Pre-selected muons used for the overlap  
 254 removal have to pass  $p_T > 7 \text{ GeV}$ ,  $|\eta| < 2.5$  and the Medium ID selection, as defined by the muon CP  
 255 group [34]. This WP sets requirements on the number of hits in the different Inner Detector and Muon  
 256 Spectrometer subsystems and on the significance of the charge-to-momentum ratio  $q/p$  [35, 36]. As for the  
 257 electrons, on top of the criteria for pre-selected muons, the PLVTight isolation is applied for muons used  
 258 in the analysis selection [32].

259 The recommended cuts for the longitudinal and transverse IPs are applied: the reconstructed track  
 260 associated with the muon candidate must have  $|z_0 \sin(\theta)| < 0.5 \text{ mm}$  and  $|d_0|/\sigma(d_0) < 3$ . Associated SFs  
 261 for identification and isolation [37] are applied as multiplicative factors to the MC event weight, to correct  
 262 for the efficiency differences between data and MC [35].

## 263 4.3 Jets

264 The jet definition is summarised in Table 6. Jets are reconstructed using the anti- $k_t$  jet algorithm [Cacciari:2008gp]  
 265 on particle-flow objects [38] with a distance parameter set to  $R = 0.4$ , as implemented in the FASTJET  
 266 package [Fastjet] (denoted as AntiKt4EMPFlowJets jet collection). The jet calibration is performed  
 267 using the standard jet calibration procedure recommended by the jet/ $E_T^{\text{miss}}$  CP group, which corrects the  
 268 jet energy to match on average the true jet energy at particle level and applies an in situ correction for  
 269 data [39]. A “Jet Vertex Tagger” (JVT) [40, 41] is used to select jets, as recommended by the jet/ $E_T^{\text{miss}}$   
 270 CP group [42]. The jets are only kept if they have  $p_T > 20 \text{ GeV}$  and are inside a pseudorapidity range of  
 271  $|\eta| < 4.5$ . In addition,  $JVT > 0.2$  is required for jets  $p_T < 60 \text{ GeV}$  and  $|\eta| < 2.4$ . For so-called “forward  
 272 jets” with  $2.5 < |\eta| < 4.5$  and  $p_T < 120 \text{ GeV}$ , a special JVT WP (fJVT) is applied:  $fJVT < 0.5$  in addition  
 273 to a timing-requirement on the jet [43].

274 The selection of jets containing  $b$ -hadrons (“ $b$ -jets”) is done with a multivariant deep-learning algorithm  
 275 denoted as DL1r [44], as recommended by the flavour-tagging CP group [45]. An approach denoted  
 276 as pseudo-continuous  $b$ -tagging (PCBT) is applied, where exclusive bins in the  $b$ -tagging discriminant  
 277 corresponding to different  $b$ -jet identification efficiencies are used. This allows for the use of different  
 278 calibrated  $b$ -tagging WPs for the pt-leading and sub-leading jets in the events.

Pre-selected jet	
Collection	AntiKt4EMPFlowJets
Acceptance	$p_T > 20 \text{ GeV},  \eta  < 4.5$
Overlap removal	see Section 4.5
Jet Vertex Tagger	$JVT > 0.2$ if $ \eta  < 2.4$ and $p_T < 60 \text{ GeV}$ $fJVT < 0.5$ if $2.5 <  \eta  < 4.5$ and $p_T < 120 \text{ GeV}$
$b$ -tagged jet	
Acceptance	$p_T > 20 \text{ GeV},  \eta  < 2.5$
$b$ -tagging	DL1r algorithm (pseudo-continuous)

Table 6: Summary of the jet selection criteria and  $b$ -tagging.

279 **4.4 Missing transverse energy**

280 The missing transverse energy ( $E_T^{\text{miss}}$ ) is computed using the xAOD container MET\_RefFinal as input and  
 281 using all identified and calibrated physics objects (except from tau-leptons), according to the aforementioned  
 282 definitions.

$$E_{x,y}^{\text{miss}} = \sum_{\mu} E_{x,y}^{\text{miss},\mu} + \sum_e E_{x,y}^{\text{miss},e} + \sum_{\gamma} E_{x,y}^{\text{miss},\gamma} + \sum_j E_{x,y}^{\text{miss,jet}} + E_{x,y}^{\text{miss,soft}}. \quad (1)$$

283 The soft-term ( $E_T^{\text{soft}}$ ) is obtained from tracks not associated to any reconstructed object. More details about  
 284 the  $E_T^{\text{miss}}$  computation can be found in Ref. [46].

285 **4.5 Overlap removal**

286 After the identification, overlaps between objects are resolved in order to avoid double-counting of physics  
 287 objects. Pre-selected leptons and jets are used for this procedure. This is the sequence of operations  
 288 that are performed to resolve these ambiguities, as implemented as the “harmonized” option in the  
 289 `AssociationUtils` package [47]. The overlap removal (OR) procedure is performed in the following  
 290 steps:

- 291 • Pre-selected electrons overlapping with pre-selected muons (as defined in Table 5) within a distance  
   292 of  $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.01$  (or sharing a track with them) are removed.
- 293 • If the distance between a jet (Table 6) and a pre-selected electron is  $\Delta R < 0.2$ , the jet candidate is  
   294 rejected.
- 295 • Any remaining electron or muon closer than  $\Delta R = 0.4$  to a jet is discarded.
- 296 • If the distance between a jet and a pre-selected muon is  $\Delta R < 0.4$  and the jet has more than two  
   297 associated tracks, then the muon is removed, otherwise the jet is removed.

## 298 5 The Trilepton Channel

### 299 5.1 Signal Topology & General Considerations

300 In order to reduce the number of background events in the search for  $tWZ$  production, a trilepton decay  
 301 channel was chosen. This was chosen because in the case of less than three leptons, hadronic background  
 302 processes dominate to such an extent as to make it impossible to make a measurement of  $tWZ$  production  
 303 in these channels. The trilepton and four lepton channels were considered for further study and the trilepton  
 304 channel was chosen because it offered manageable background production processes and offered reasonable  
 305 statistical power.

306 In Figure 3, a Feynman diagram of  $tWZ$  production in the trilepton channel is shown, where from an initial  
 307 state of a bottom quark and a gluon, a top quark,  $W$  boson, and  $Z$  boson are produced. The  $Z$  boson is  
 308 required to decay into a pair of opposite sign same flavour (OSSF) leptons, while the  $W$  bosons (both the  
 309  $W$  boson from the top quark decay and the one produced in association with the top quark) could decay in  
 310 the leptonic or hadronic channels.

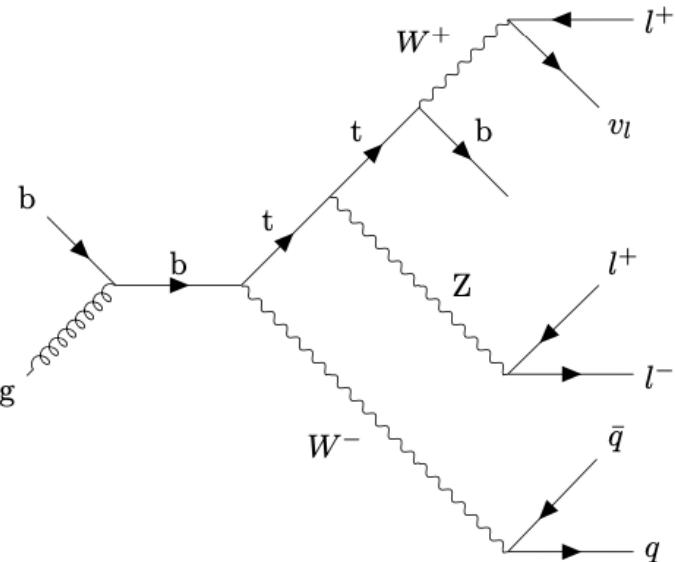


Figure 3: An example Feynman Diagram is shown for  $tWZ$  production in the trilepton channel.

### 311 5.2 Baseline selections

312 The baseline object selections described below are summarized in Table 7. Selections were applied to the  
 313  $p_T$  of the leading, next-to-leading (NL), and next-to-next-to-leading (NNL) leptons. The leading lepton  
 314 refers to the lepton with the highest  $p_T$  value, the NL lepton refers to the lepton with the second highest  $p_T$   
 315 value, and the NNL lepton refers to the lepton with the third highest  $p_T$  value. The leading, NL, and NNL  
 316 leptons were required to have  $p_T$  values greater than 30, 20, and 10 GeV respectively. Selections were also  
 317 applied to the  $\eta$  values of the leptons. Each muon ( $\mu$ ) was required to have an  $|\eta_\mu| \leq 2.5$ , and each electron

318 (e) was required to have an  $|\eta_e| \leq 2.47$  – excluding the region between  $1.37 \leq |\eta_e| \leq 1.52$  since this is a  
 319 transition region between the barrel and the end-caps with a relatively large amount of inactive material.

320 All jets were required to have a  $p_T > 25$  GeV, an  $|\eta_{jet}| < 2.5$ , and a jet vertex tagger (jvt) score greater than  
 321 0.5. The jvt above 0.5 was required to reject jets originating from pileup interactions. The jvt requirement  
 322 rejects jets originating from pileup interactions by using information about the jet shapes and topological  
 323 jet correlations. An operating point for the *DL1r* b-tagging algorithm with an efficiency of 77% was  
 324 chosen.

<b>Baseline object selections</b>
<b>Lepton selections</b>
Leading lepton $p_T > 30$ GeV
NL lepton $p_T > 20$ GeV
NNL lepton $p_T > 10$ GeV
$ \eta_\mu  \leq 2.5$
$ \eta_e  \leq 2.47$ excluding $1.37 \leq  \eta_e  \leq 1.52$
<b>Jet selections</b>
all jet $p_T > 25$ GeV
$ \eta_{jet}  < 2.5$
$jvt > 0.5$
DL1r eff = 77%

Table 7: This table is a summary of all object selections, with the lepton selections shown in the top half of the table, and the jet selections shown in the bottom half of the table.

325 The baseline event selections described below are summarized in Table 8. Exactly three leptons were  
 326 required in every event, where the three leptons could be any combination of electrons and muons. To  
 327 suppress events containing  $\Upsilon$  and  $J/\Psi$  particles, which are not included in the simulation, every OSSF  
 328 dilepton pair in an event was required to have an invariant mass greater than 10 GeV – this being greater than  
 329 the mass of these two particles. Exactly one  $Z$  candidate per event was required, where the  $Z$  candidates  
 330 are defined as OSSF dilepton pairs that were within a 10 GeV range of the  $Z$  boson mass. Furthermore, at  
 331 least two jets per event, and at least one b-tagged jet per event were also required.

### 332 5.3 Signal and Control Regions

333 The Signal Region (SR) and Control Region (CR) selections described below are summarized at the bottom  
 334 of Table 8. The regions were defined according to selections on the total number of jets and the number of  
 335 b-tagged jets contained within the total number of jets. The combination of the total number of jets and the  
 336 number of b-tagged jets contained within the total number of jets is referred to as the jet multiplicity.  
 337 The  $tWZ$  SR was chosen as the region that gave the highest significance score,  $(s/\sqrt{N})$ , where  $s$  is the  
 338 number of expected signal events in simulation that satisfy the SR selections, and  $N$  is the total number of  
 339 expected events for simulated signal and background production processes satisfying the SR selections.  
 340 This metric for significance was used to select a SR that was the most sensitive to the observation of  $tWZ$   
 341 production.

---

**Baseline event selections**


---

	Exactly 3 leptons ≥ 2 jets ≥ 1 b-tagged jet all $m_{OSSF_{ll}} > 10 \text{ GeV}$ 1 Z Candidate ( $ m_{OSSF_{ll}} - m_Z  < 10 \text{ GeV}$ )	
		<b>Regions</b>
$tWZ$ SR	WZ CR	$t\bar{t}Z$ CR
	Exactly 3 Tight leptons	
	≥ 3 jets 1 b-tagged jet	2 jets 1 b-tagged jet
		≥ 2 jets ≥ 2 b-tagged jets

Table 8: This table serves as a summary of all event selections and the region selections. The event selections are shown in the top half of the table, and the region selections are shown in the bottom half of the table.

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

351 In Figure 4, two-dimensional plots of the jet multiplicities for the SR and CRs are shown with the  
 352 significance scores shown by a colour axis on the right of each plot. The jet multiplicity chosen for the SR  
 353 was one with greater than or equal to three jets, and exactly one b-tagged jet ( $\geq 3j 1b$ ). A jet multiplicity  
 354 of exactly two jets, and exactly one b-tagged jet ( $2j 1b$ ) was chosen as the CR for  $WZ$  production, and a jet  
 355 multiplicity of greater than or equal to two jets and greater than or equal to two b-tagged jets ( $\geq 2j \geq 2b$ )  
 356 was chosen as the CR for  $t\bar{t}Z$  production.

357 In Table 9, the simulated number of events for each sample is shown for each region. The total number of  
 358 events for simulation is shown for each region. In the CRs, the number of events in data is also shown.  
 359 The work presented here forms the basis of an official ATLAS experiment analysis. Therefore, the SR  
 360 remains blinded to data to avoid potential biases in the future development of the official ATLAS analysis.  
 361 According to the total number of events for simulation and data, the simulation in the CRs seems to model  
 362 the data well within uncertainty. Furthermore, the production processes that these regions are chosen for  
 363 are in fact produced in their greatest quantities in these regions. Only samples that contributed to at least  
 364 one the the regions are shown in the table.

365  $WZ$  production is broken up into  $WZ + b$ ,  $WZ + c$ , and  $WZ + l$ . If an event has at least one true b-jet it is  
 366 included in  $WZ + b$ . If it has no true b-jet and at least one true c-jet it is included in  $WZ + c$ . Lastly, if it  
 367 has no true b-jet and no true c-jet it is included in  $WZ + l$ . However, no  $WZ + l$  jets pass the baseline event  
 368 selections in these regions.

Not reviewed, for internal circulation only

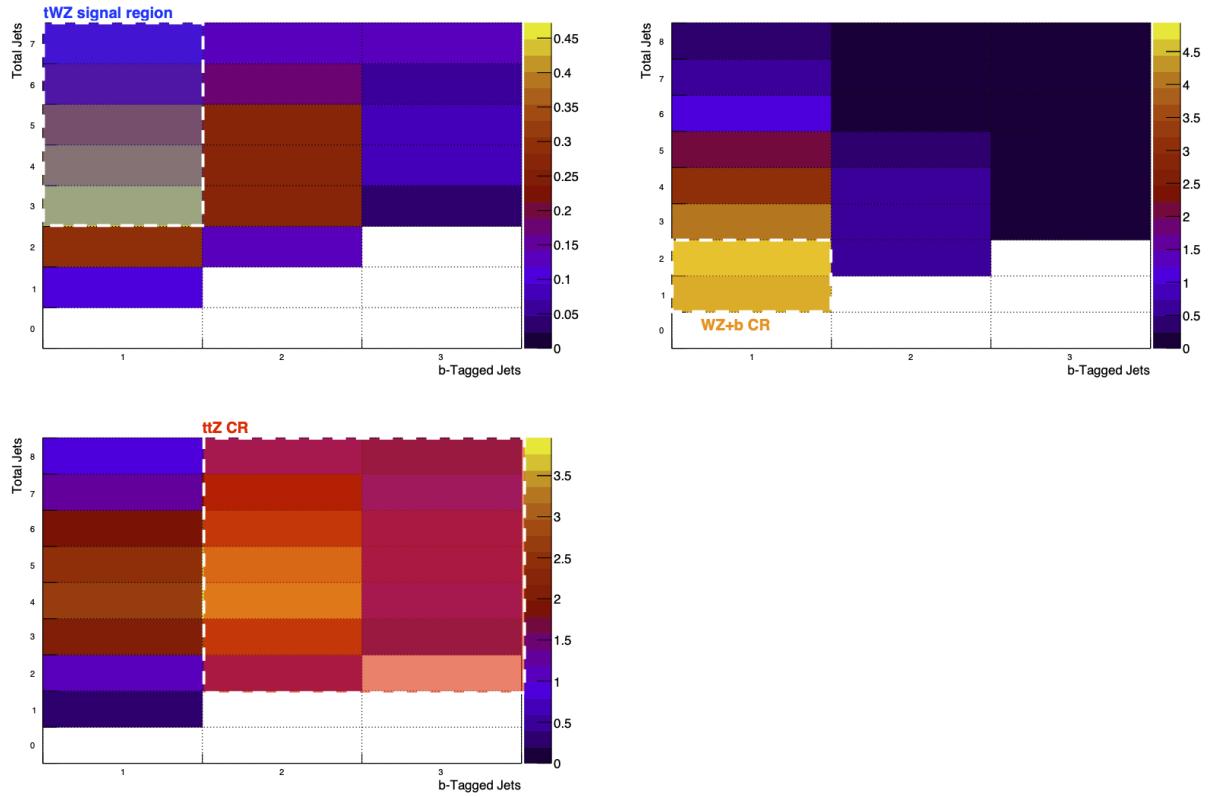


Figure 4: Two-dimensional plots of the jet multiplicities for the SR and CRs are shown with the significance scores shown by a colour axis on the right of each plot.

369 In Figure 5, the number of events for each production process and the total number of events for simulation  
 370 and data are shown by a histogram with each bin corresponding to a different region. Simulation is shown  
 371 with different production processes stacked on top of one another, with the non-dominant background  
 372 production processes grouped together into a class labelled as "other".  $tWZ$  production is shown in black.  
 373  $WZ + b$  is shown in light yellow,  $WZ + c$  is shown in dark yellow, and  $WZ + l$  is shown in orange. However,  
 374 no  $WZ + l$  events contributed.  $t\bar{t}Z$  production is shown in green,  $tZq$  production is shown in pink, and the  
 375 other class of background production processes is shown in red. Data is shown by black points, and at the  
 376 bottom of the histogram the ratio of the total number of data events divided by the total number of events in  
 377 simulation is shown. This is not shown in the SR because the SR is blind to data. In this figure, the  $t\bar{t}Z$  CR  
 378 is shown with an abundance of  $t\bar{t}Z$  events with minimal  $WZ$ ,  $tWZ$ , and other background events, while  
 379 an abundance of  $tZq$  events is observed. In the  $WZ$  CR, an abundance of  $WZ$  events is observed with  
 380 minimal  $t\bar{t}Z$ ,  $tWZ$ , and other background events, while an abundance of  $tZq$  events is observed. From this  
 381 histogram, one can see that  $tZq$  production is abundant in both regions but is particularly prevalent in the  
 382  $WZ$  CR. This histogram also shows that simulation and data agree in the  $t\bar{t}Z$  and  $WZ$  CRs – as can be seen  
 383 by the ratio of Data/Pred. around 1. This shows that the simulation successfully models the data.  
 384 The CRs are also used to check the modeling of the jet and lepton kinematics –  $\eta$ ,  $\phi$ , and  $p_T$  – and certain  
 385 event level variables – the number of jets, and the number of b-tagged jets. These variables are used  
 386 to construct higher level variables that are then used in the search for  $tWZ$  production. Therefore, it is  
 387 essential to check that the data is well modelled by simulation before using the kinematics to construct

Not reviewed, for internal circulation only

Process	Sample(s)	$tWZ$ SR	$t\bar{t}Z$ CR	WZ CR
$tWZ$	$tWZ$	$17.10 \pm 0.61$	$8.30 \pm 0.30$	$5.18 \pm 0.20$
$WZ$	$WZ + b$	$31.56 \pm 15.83$	$13.79 \pm 6.92$	$30.07 \pm 15.10$
	$WZ + c$	$39.87 \pm 16.02$	$2.50 \pm 1.04$	$46.62 \pm 18.76$
$t\bar{t}Z$	$t\bar{t}Zee$	$34.04 \pm 1.42$	$40.77 \pm 1.60$	$4.27 \pm 0.24$
	$t\bar{t}Z\mu\mu$	$47.91 \pm 1.81$	$61.86 \pm 2.33$	$6.33 \pm 0.33$
	$t\bar{t}Z\tau\tau$	$0.24 \pm 0.07$	$0.45 \pm 0.09$	$0.09 \pm 0.04$
$tZq$	$tZq$	$69.99 \pm 10.23$	$57.33 \pm 8.39$	$89.45 \pm 13.07$
Diboson	$VBF_{EWK}$	$1.15 \pm 0.07$	$0.09 \pm 0.02$	$0.75 \pm 0.06$
	$ZZgg$	$0.42 \pm 0.09$	$0.07 \pm 0.02$	$0.54 \pm 0.11$
	$ZZqq$	$9.41 \pm 1.92$	$2.20 \pm 0.45$	$13.39 \pm 2.77$
	$zz\_lowMllPtComplement$	$1.91 \pm 1.50$	$0.91 \pm 2.22$	
	$vvllljjEWK$	$3.33 \pm 0.14$	$0.32 \pm 0.03$	$2.80 \pm 0.13$
$t\bar{t}$	$t\bar{t}$	$37.05 \pm 7.74$	$16.77 \pm 3.62$	$67.37 \pm 13.88$
$t\bar{t} + X$	$t\bar{t}H$ dilep	$1.24 \pm 0.26$	$1.93 \pm 0.40$	$0.35 \pm 0.08$
	$t\bar{t}H$ semilep	$0.90 \pm 0.22$	$1.35 \pm 0.31$	$0.02 \pm 0.03$
	$t\bar{t}W$	$2.07 \pm 0.20$	$5.79 \pm 0.36$	$2.77 \pm 0.23$
	$t\bar{t}tt$	$0.01 \pm 0.01$	$0.23 \pm 0.03$	$0.00 \pm 0.00$
	$t\bar{t}WW$	$0.19 \pm 0.06$	$0.16 \pm 0.06$	$0.00 \pm 0.00$
$ZH$	$ZH$	$0.57 \pm 1.41$	$0.00 \pm 0.00$	$1.66 \pm 1.72$
$Z + jets$				
$Z\mu\mu\_MAXHTPTV$				
	$0\_70\_BFilter$	$2.27 \pm 5.67$	$0.00 \pm 0.00$	$6.28 \pm 12.11$
	$70\_140\_CFilterBVeto$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$2.66 \pm 6.64$
	$70\_140\_BFilter$	$0.81 \pm 2.91$	$1.82 \pm 1.68$	$7.20 \pm 4.41$
	$140\_280\_CFilterBVeto$	$1.14 \pm 1.31$	$0.00 \pm 0.00$	$0.00 \pm 0.00$
	$140\_280\_BFilter$	$4.47 \pm 2.33$	$1.58 \pm 0.87$	$6.49 \pm 3.33$
	$280\_500\_CFilterBVeto$	$1.03 \pm 0.90$	$0.00 \pm 0.00$	$0.26 \pm 0.64$
	$280\_500\_BFilter$	$1.35 \pm 0.71$	$0.49 \pm 0.28$	$0.75 \pm 0.42$
	$500\_1000$	$0.81 \pm 0.46$	$0.00 \pm$	$0.13 \pm 0.12$
	$1000\_E\_CMS$	$0.05 \pm 0.04$	$0.03 \pm 0.05$	$0.01 \pm 0.02$
	$Zee\_MAXHTPTV$			
	$70\_140\_BFilter$	$2.37 \pm 1.62$	$0.00 \pm 1.12$	$8.01 \pm 5.28$
	$140\_280\_CFilterBVeto$	$0.23 \pm 0.58$	$0.00 \pm 0.00$	$0.61 \pm 0.90$
	$140\_280\_BFilter$	$2.67 \pm 1.44$	$0.70 \pm 0.42$	$3.65 \pm 1.90$
	$280\_500\_BFilter$	$0.95 \pm 0.51$	$0.24 \pm 0.17$	$0.51 \pm 0.29$
	$500\_1000$	$0.50 \pm 0.30$	$0.00 \pm 0.00$	$0.26 \pm 0.19$
	$1000\_E\_CMS$	$0.07 \pm 0.05$	$0.01 \pm 0.01$	$0.03 \pm 0.03$
	Total	$292.36 \pm 29.82$	$207.83 \pm 13.89$	$258.95 \pm 38.39$
Data		—	199	249

Table 9: The number of expected events for each simulated sample is shown in the SR and CRs. The total number of predicted events in simulation is shown in each region at the bottom of the table alongside the number of events in data in the CRs.

Not reviewed, for internal circulation only

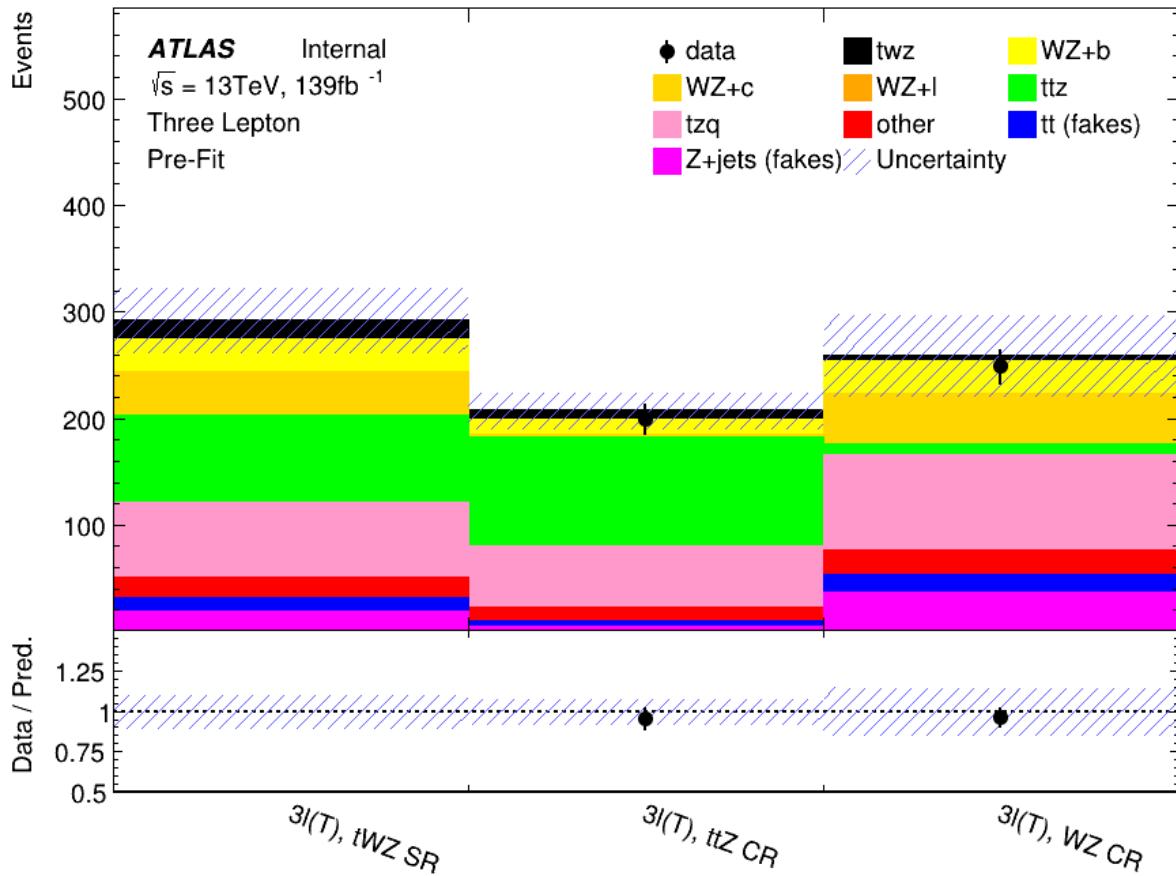


Figure 5: The number of events in the SR and CRs are shown for simulation and data in the upper panel. Simulation is shown with different production processes stacked on top of one another in different colours, and data in the CRs is shown by the black points. The agreement between simulation and data is shown by the Data/Pred. ratio in the lower panel.

388 these high level variables.

389 In Figure 6, the kinematics of the leading, NL, and NNL leptons are shown for simulation and data in  
 390 the  $t\bar{t}Z$  CR. In Figure 7, the same is shown in the  $WZ$  CR. There is generally good agreement between  
 391 simulation and data for all the object kinematics as shown by the Data/Pred. ratio in each plot.

392 In Figure 8, the kinematics of the leading and NL jets are shown for simulation and data in the  $t\bar{t}Z$  CR.  
 393 The number of jets, and the number of b-tagged jets are also shown for simulation and data in the  $t\bar{t}Z$   
 394 CR. In Figure 9 the same variables are shown in the  $WZ$  CR. Generally good agreement is seen between  
 395 simulation and data in both CRs for these variables.

Not reviewed, for internal circulation only

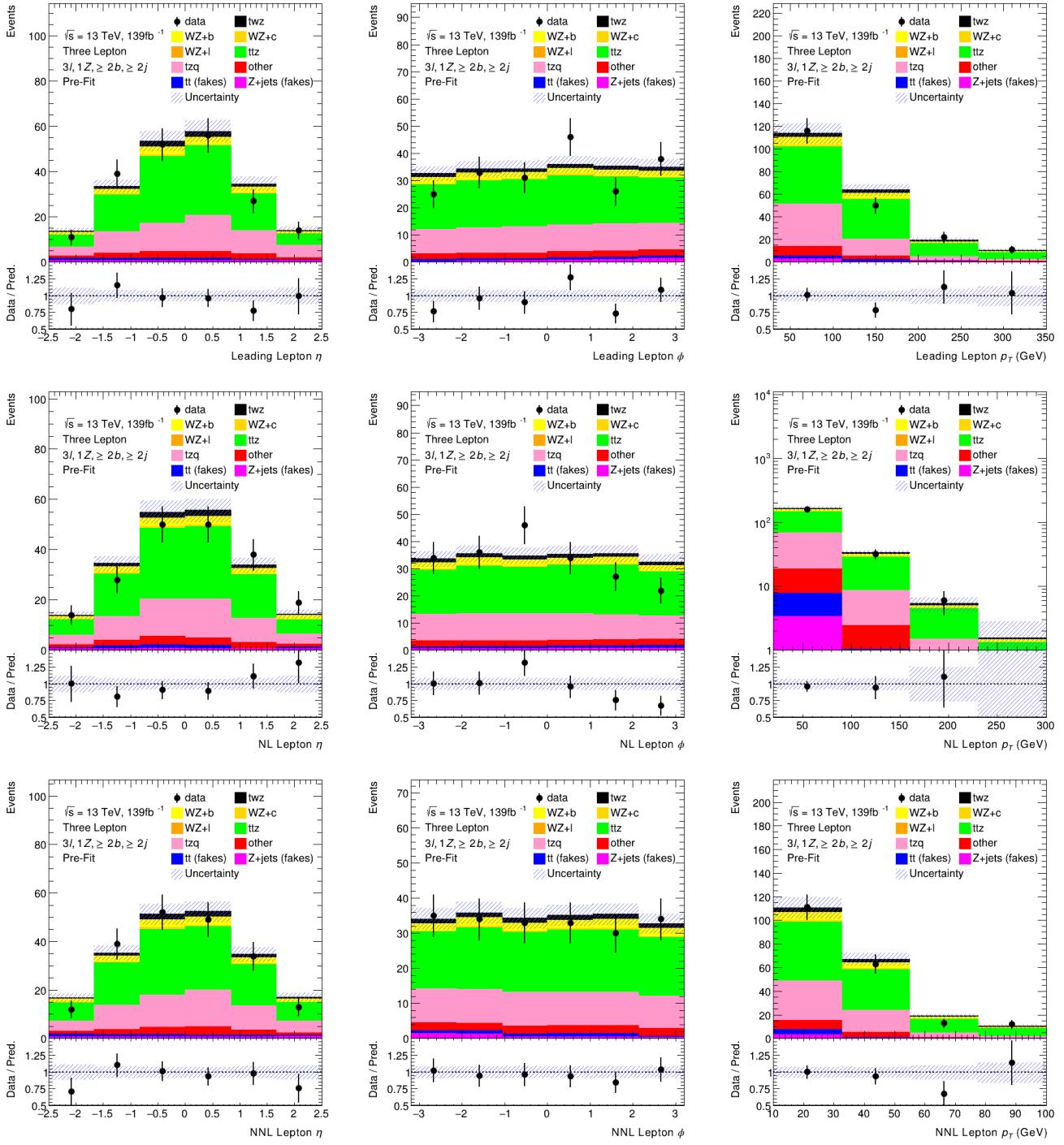


Figure 6: A comparison of simulation and data for the object kinematics –  $\eta$ ,  $\phi$ , and  $p_T$  – is shown for the leading, NL, and NNL leptons in the  $t\bar{t}Z$  CR

Not reviewed, for internal circulation only

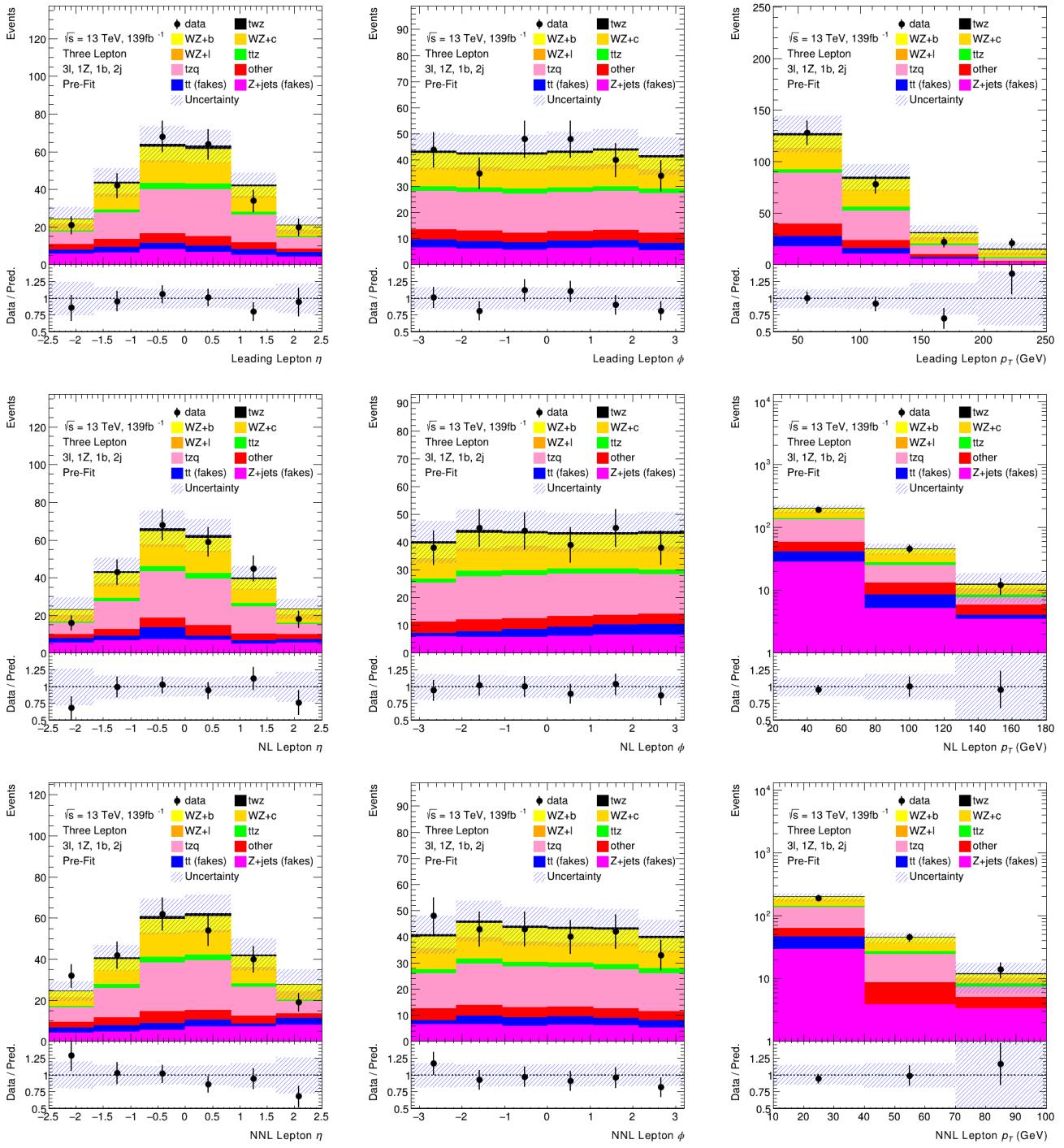


Figure 7: A comparison of simulation and data for the object kinematics –  $\eta$ ,  $\phi$ , and  $p_T$  – is shown for the leading, NL, and NNL leptons in the  $WZ$  CR

Not reviewed, for internal circulation only

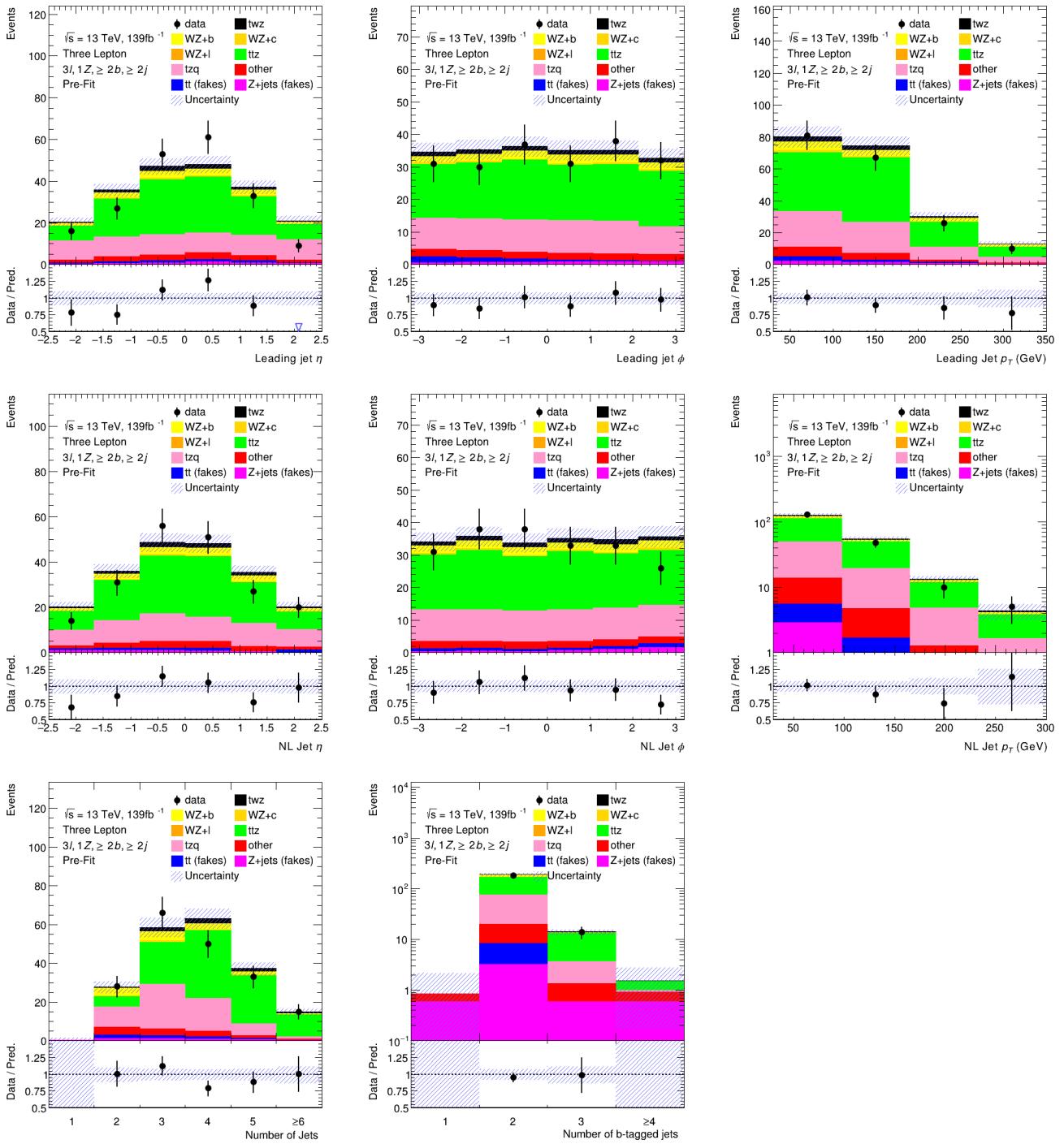


Figure 8: A comparison of simulation and data for the object kinematics –  $\eta$ ,  $\phi$ , and  $p_T$  – is shown for the leading, and NL jets in the  $t\bar{t}Z$  CR

Not reviewed, for internal circulation only

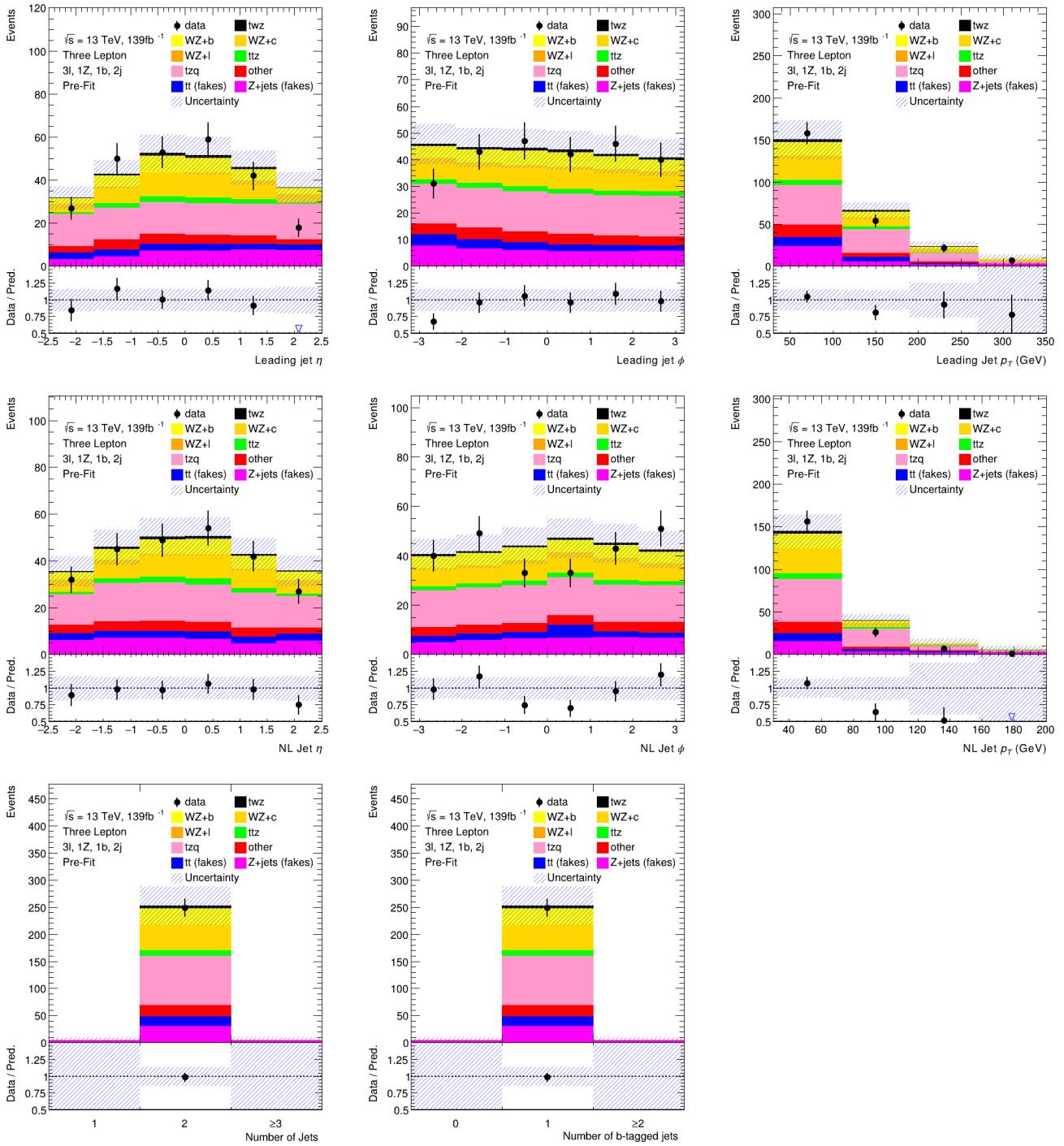


Figure 9: A comparison of simulation and data for the object kinematics –  $\eta$ ,  $\phi$ , and  $p_T$  – is shown for the leading, and NL jets in the  $WZ$  CR

### 396 5.4 Fake lepton estimation

397 The trilepton channel requires three leptons which come from a W or Z boson. However, it is possible that  
398 certain events pass the trilepton selection even if one of the leptons is not a prompt lepton. Such non-prompt  
399 (or fake) leptons can originate from decays of bottom or charm hadrons, a jet that is misidentified as an  
400 electron, leptons from kaon or pion decays or even electrons coming from photon conversions.

401 The  $t\bar{t}$  and  $Z + jets$  processes are expected to be the dominant source of fake leptons in the trilepton  
402 channel because they should not contribute at all in this channel without fake leptons. This is because only  
403 two leptons should be produced in each case. These being from the leptonic decay of the top quarks for  $t\bar{t}$ ,  
404 and from the leptonic decay of the Z Boson for  $Z + jets$ . However, due to the large cross sections of these  
405 processes, some events do pass the event selections due to these fake leptons.

406 The contributions from these two processes were estimated using a MC-based method called the MC  
407 template method. This method was implemented by defining three orthogonal CRs that were enhanced in  
408 fake leptons. With these regions included, the normalization uncertainties on the  $t\bar{t}$  and  $Z + jets$  cross  
409 sections were then allowed to vary in the fit in order to estimate the contribution from these two processes.

410 In Figure 10, the different CRs chosen to estimate the contribution from events with fake leptons are shown.  
411 The number of events for each production process and the total number of events for simulation and data  
412 are shown by a histogram with each bin corresponding to a different CR. The regions chosen to model the  
413 contribution from fake leptons had the same jet multiplicity selections as the SR and CRs. However, these  
414 fake CRs differed in the isolation requirement on one of the leptons. Instead of three tight leptons being  
415 required, two tight leptons and one loose (and not tight) lepton were required. This was done because a  
416 loose (and not tight) lepton is more likely to be fake. These regions were labelled as  $tWZ$  fk CR,  $WZ$  fk  
417 CR, and  $t\bar{t}Z$  fk CR to reflect the jet multiplicity selections chosen for these regions.

Not reviewed, for internal circulation only

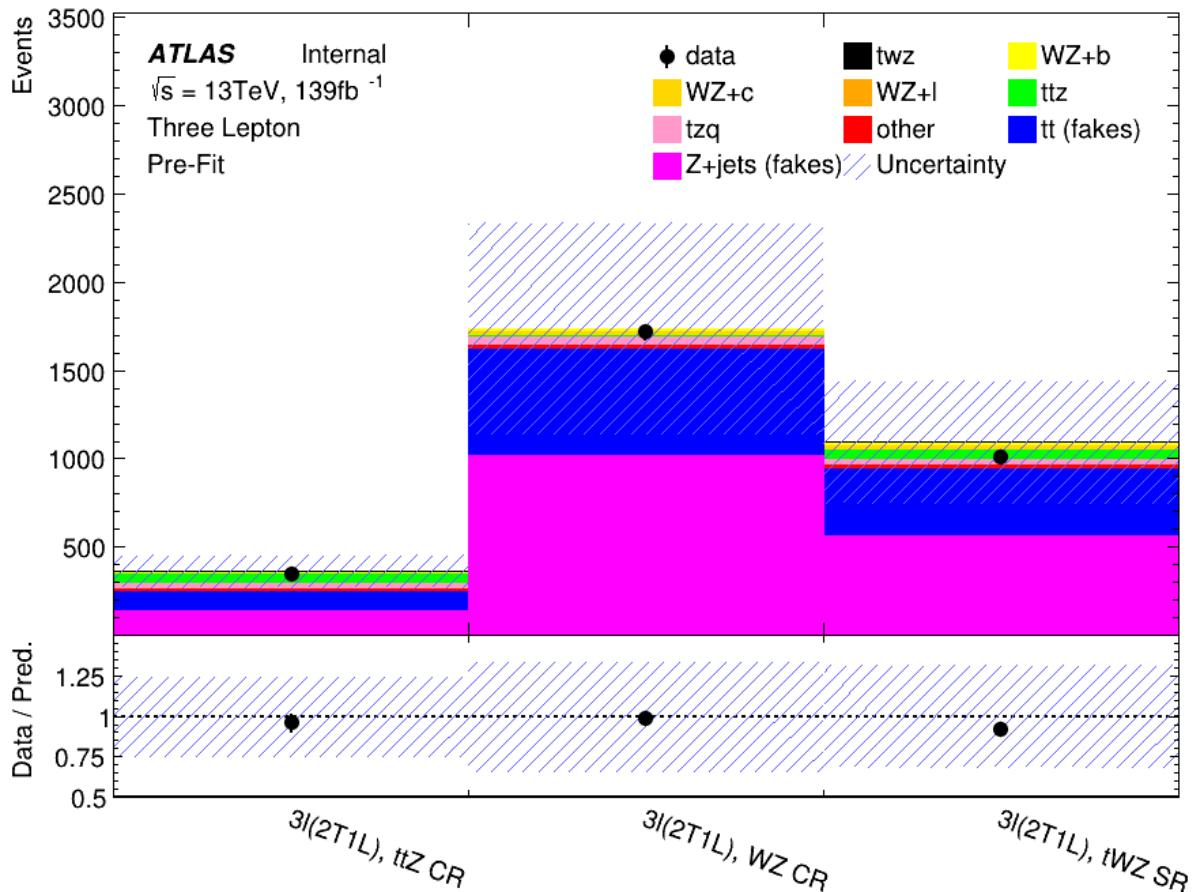


Figure 10: The number of events in the fake lepton CRs are shown for simulation and data in the upper panel. Simulation is shown with different production processes stacked on top of one another in different colours, and data is shown by the black points. The agreement between simulation and data is shown by the Data/Pred. ratio in the lower panel.

## 418 5.5 Hadronically-decaying W Boson identification

419 In a previous search for  $tWZ$  production in the trilepton channel using  $36.1\text{ fb}^{-1}$  of  $\sqrt{s} = 13\text{ TeV}$  pp  
420 collision data from the ATLAS detector [1], the leading source of uncertainty was due to the uncertainty  
421 on the size of the background from  $WZ$  production. This analysis aimed to minimize the uncertainty  
422 due to this background by using the presence of a hadronically-decaying W Boson ( $W_{had}$ ) to distinguish  
423 between  $tWZ$  and  $WZ$  production. The presence of a  $W_{had}$  was used because in events with three leptons  
424 originating from  $tWZ$  production, there will always be a  $W_{had}$  (either coming from the  $W$  produced in  
425 association with the top quark or from the  $W$  coming from the decay of the top quark). However, there will  
426 never be a  $W_{had}$  in events with three leptons originating from  $WZ$  production. In Figures 11(a) and 11(b),  
427 two Feynman diagrams show the two possible decay channels of  $tWZ$  to the three leptons, where a  $W_{had}$   
428 must be present. However, in Figure 11(c),  $WZ$  production in the trilepton channel can be seen with no  
429  $W_{had}$  present.

430 The identification of  $W_{had}$  in an event was achieved using a Gradient Boosted Decision Tree (GBDT)  
431 trained on dijets from a simulated  $t\bar{t}$  sample with a single lepton selection. To train the GBDT, labels and

Not reviewed, for internal circulation only

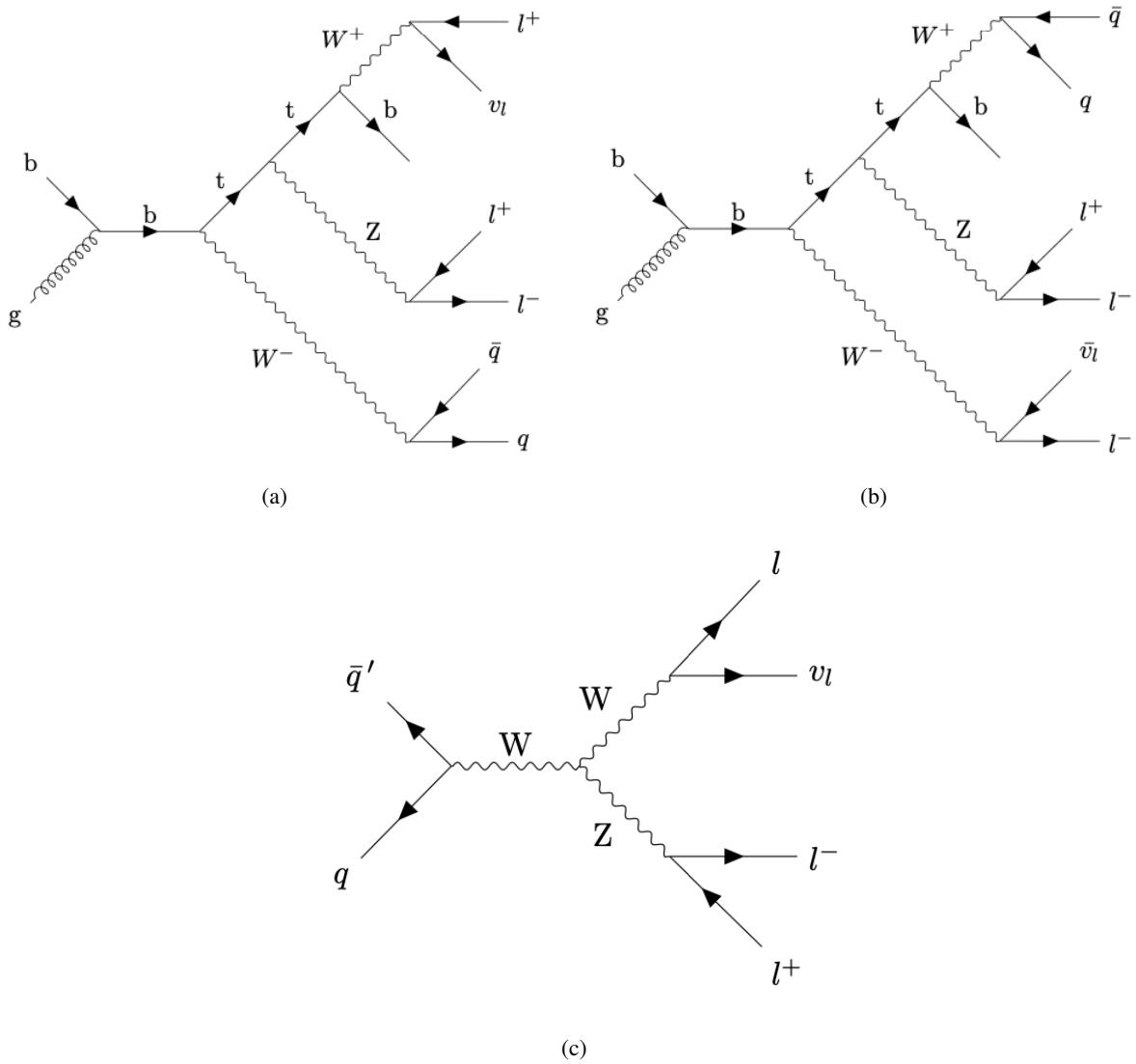


Figure 11: The top two Feynman diagrams show the two possible decays of  $tWZ$  production with a  $W_{had}$  present. The bottom Feynman diagram shows the decay of  $WZ$  production with no  $W_{had}$  present.

432 features for signal and background dijets were needed, where a dijet is the combination of two jets in an  
 433 event, a signal dijet is a dijet found to originate a  $W_{had}$  in an event, and a background dijet is any other dijet  
 434 combination in an event.

### 435 **Labelling dijets**

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

In Figure 12, a flow chart depicting the procedure for labelling signal and background dijets is shown. First all the dijet combinations in an event are constructed. Then to associate a dijet with a  $W$  boson, the spatial separation between a  $W$  boson, found in the truth information, and a dijet, formed using reco information, was used as a metric to quantify how likely the dijet is to originate from a true  $W_{had}$ . The spatial separation is given by:

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

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

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

461 In Figure 13, two examples are shown of Gaussian fits (in red) to the  $m_{jj}$  distributions of the signal dijets  
 462 (in blue) for cuts of  $\Delta R_{Truth,Reco} < 0.05$  and  $\Delta R_{Truth,Reco} < 0.45$ . The signal dijets are taken from a  
 463  $t\bar{t}$  sample with a single lepton selection. In Figure 14, the increase in  $\sigma_{Truth,Reco}$  – shown in Figures  
 464 13(a) and 13(b) as the sigma values in the top right block – as a function of the cut applied to define the  
 465 signal dijet class can be seen. Figure 14 shows that as the cut on  $\Delta R_{Truth,Reco}$  increases,  $\sigma_{Truth,Reco}$   
 466 increases dramatically up until a cut of around 0.5, and then increases slowly for larger values of the cut

Not reviewed, for internal circulation only

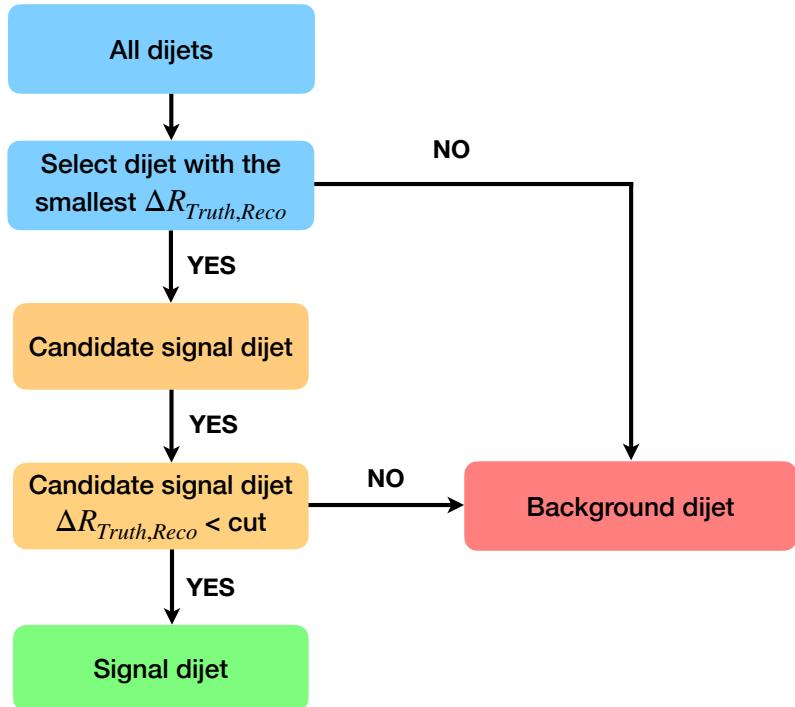


Figure 12: This flow chart details the procedure for taking all dijets and labelling them as either signal or background dijets.

467 value. This highlights the fact that as the cut increases, the signal dijet class becomes more contaminated  
 468 with erroneously added background dijets. Based on Figure 14, a cut of  $\Delta R_{Truth,Reco} \leq 0.05$  was chosen  
 469 because it is as tight a cut as possible that still provides a sufficient number of signal dijets for training the  
 470 GBDT, has a mean value close to the mass of the  $W$  boson, and has a small  $\sigma_{Truth,Reco}$ .

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

Not reviewed, for internal circulation only

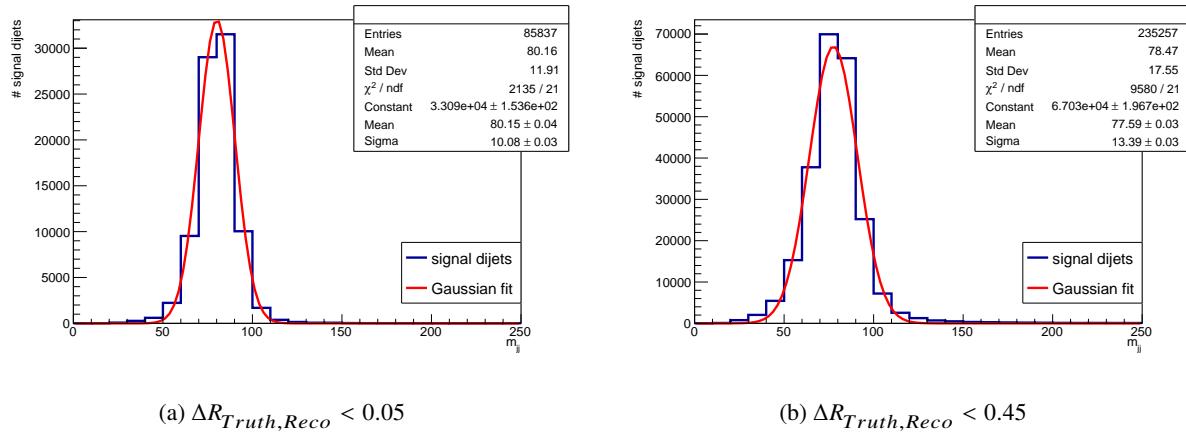


Figure 13: Two examples of the  $m_{jj}$  distributions for signal dijets (in blue) with a gaussian fit (in red) to the distributions are shown for all signal dijets. The signal dijets are taken from a  $t\bar{t}$  sample with a single lepton selection. In Figure 13(a) the  $m_{jj}$  distribution is shown for signal dijets satisfying a cut of  $\Delta R_{Truth, Reco} < 0.05$ , and in Figure 13(b) the  $m_{jj}$  distribution is shown for signal dijets satisfying a cut of  $\Delta R_{Truth, Reco} < 0.45$ .

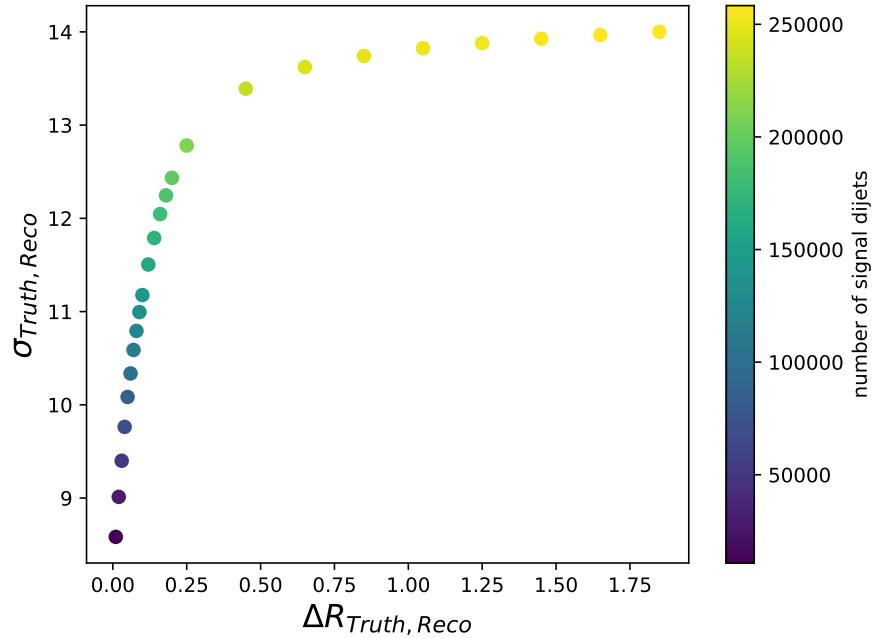


Figure 14: The  $\sigma_{Truth, Reco}$  of the fits to the  $m_{jj}$  distributions, as a function of the cut on  $\Delta R_{Truth, Reco}$ , is shown with the number of signal dijets indicated by the colour axis

## 478 Features

479 For the training of the GBDT, only three features were found to have a significant effect when identifying  
 480  $W_{had}$ 's. This is because there are only a few degrees of freedom given by the two four-vectors and

481 identification criteria of the dijets. In Figure 15, normalized distributions of the chosen features for signal  
 482 and background dijets are shown. These plots show a clear distinction between signal and background  
 483 dijets. The three features used for the training were:

484 • **The separation of the jets –  $\Delta R_{jj}$**

485 The separation of the jets is given by  $\Delta R_{jj} = \sqrt{(\eta_{j1} - \eta_{j2})^2 + (\phi_{j1} - \phi_{j2})^2}$  where  $\eta_{j1} - \eta_{j2}$  is the  
 486 difference in  $\eta$  of the jets, and  $\phi_{j1} - \phi_{j2}$  is the difference in  $\phi$  of the jets. Signal dijets should be  
 487 closer together than background dijet combinations because the signal dijets are decaying from  
 488 the same  $W$  boson. In Figure 15(a), a normalized distribution of this feature is shown for signal  
 489 and background dijets, where the signal dijet distribution of this feature peaks close to 1, and the  
 490 background dijet distribution is more uniformly distributed with a lower peak that is close to 3.

491 • **The invariant mass of the dijet system –  $m_{jj}$**

492 The invariant mass of the dijet system is given by  $m_{jj} = \sqrt{E_{jj}^2 - p_{jj}^2}$ , where  $E_{jj}$  is the energy of  
 493 the dijet system, and  $p_{jj}$  is the momentum of the dijet system. Signal dijets should have an  $m_{jj}$   
 494 distribution around the  $W$  boson mass, whereas background dijets should not be sharply peaked. In  
 495 Figure 15(b), a normalized distribution of this feature is shown for signal and background dijets,  
 496 where the signal dijet distribution peaks around the  $W$  boson mass, and the background dijets are  
 497 more uniformly distributed.

498 • **Dijet b-tagging information –  $Oneb$**

499 This feature is constructed to show if at least one of the jets in a dijet system is b-tagged. It is 1 if at  
 500 least one of the jets in the dijet system is b-tagged and 0 otherwise. Signal dijets should not have a  
 501 b-tagged jet, whereas a background dijet could, but does not have to, have a b-tagged jet. In Figure  
 502 15(c), a normalized distribution of this feature is shown for signal and background dijets, where the  
 503 signal dijet distribution has most dijets with no b-tagged jets, and the background dijet distribution  
 504 has close to an equal number of dijets with and without at least one b-tagged jet.

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

506 Any simulated events that a GBDT is trained on have to be discarded because the GBDT is biased towards  
 507 these events. Furthermore, an event-level GBDT is used in section 5.6, which means that any events used  
 508 for the the training of the event-level GBDT must also be discarded. Therefore, the training of these two  
 509 GBDTs would require a large percentage of the simulated events to be discarded for the fit. Specifically  
 510 this would mean that a large percentage of  $tWZ$  events would have to be discarded, which is problematic  
 511 because there are only a limited number of simulated  $tWZ$  events. To combat this problem, dijets from a  $t\bar{t}$   
 512 sample with a single lepton selection were used to train the  $W_{had}$  GBDT.

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

Not reviewed, for internal circulation only

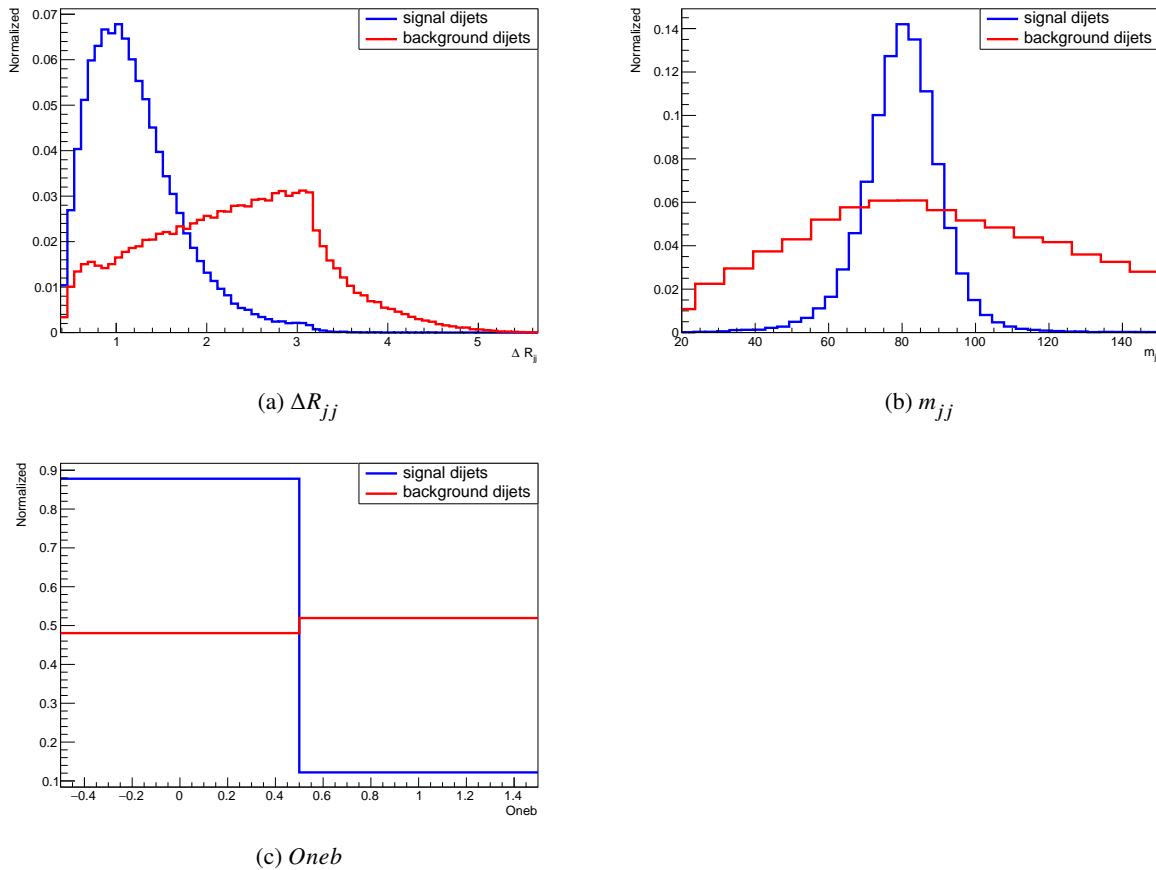


Figure 15: Normalized distributions of the GBDT features for signal dijets (in blue) and background dijets (in red) are shown. The dijets used are from a  $t\bar{t}$  sample with a single lepton selection. Figure 15(a) shows the  $\Delta R_{jj}$  distribution, Figure 15(b) shows the  $m_{jj}$  distribution, and Figure 15(c) shows the  $Oneb$  distribution

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

528 subsectionMax  $W_{had}$  GBDT score

529 In Figure 16, a histogram of the  $W_{had}$  GBDT score – where the score is the probability of a dijet being a  
 530 signal dijet – is shown for a test dataset comprising an equal number of signal and background dijets. The  
 531 histogram shows a clear distinction between the signal and background dijets with the former peaking at 1  
 532 and the latter at 0.

533 The  $W_{had}$  GBDT score is used to construct an event-level variable that can be used to identify  $W_{had}$ 's in  
 534 an event, so that  $tWZ$  events can be distinguished from  $WZ$  events. This variable is called the Max  $W_{had}$

Not reviewed, for internal circulation only

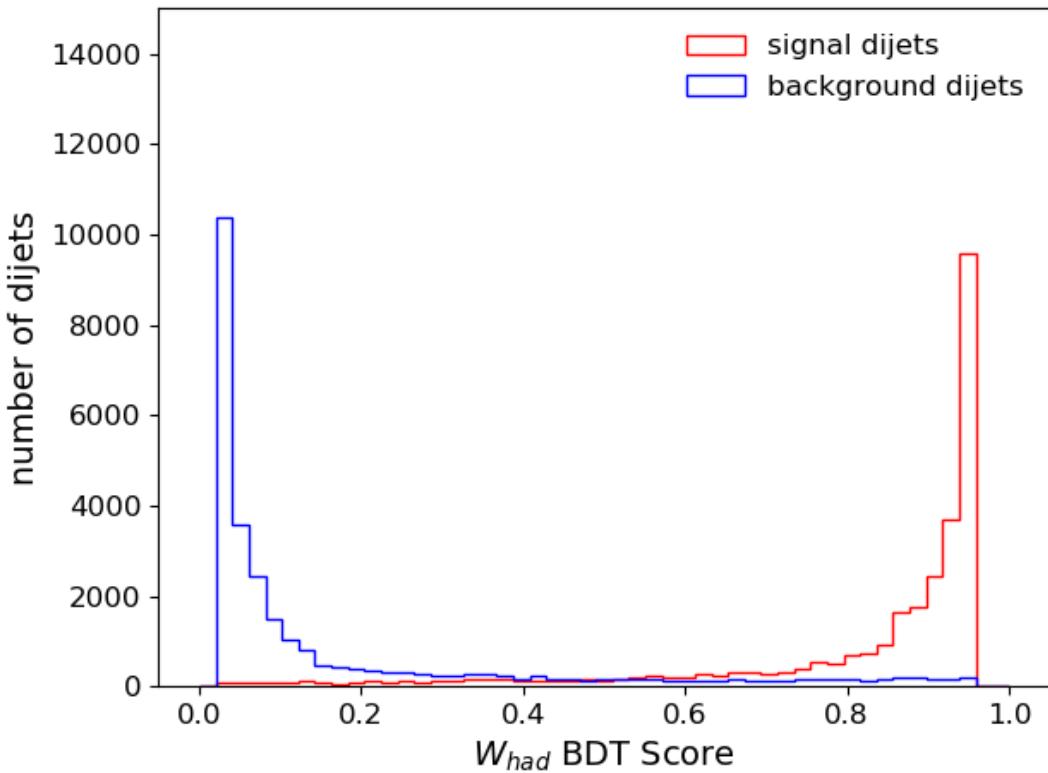


Figure 16:  $W_{had}$  GBDT scores are shown for an equal number of signal dijets (in blue) and background dijets (in red).

535 GBDT score and is constructed by taking the maximum scoring dijet per event. Events that have a  $W_{had}$   
 536 can be identified because the Max  $W_{had}$  GBDT score should be high for an event with a  $W_{had}$  present and  
 537 low for an event without a  $W_{had}$  present. In Figure 17, the Max  $W_{had}$  GBDT score is shown in the SR and  
 538 CRs for simulation and data. In the SR,  $tWZ$  events, normalized to the total number of background events,  
 539 are shown by a dotted line. The normalized  $tWZ$  events show that this variable discriminates well between  
 540 signal and background, since the  $tWZ$  events are placed predominantly to the right of the histogram since  
 541 they should have a  $W_{had}$  present. In the CRs, there is good agreement between data and simulation, except  
 542 in the last bin of the  $WZ$  CR. This disagreement in the  $WZ$  CR is under investigation, but because there are  
 543 few events in this bin, it is not expected to have a significant impact on the analysis. This could also be  
 544 evidence that the estimation of backgrounds with fake leptons entirely from simulation is an inadequate  
 545 approach.

546 In Figure 18, the features of the Max  $W_{had}$  GBDT score are shown for simulation and data in the  $t\bar{t}Z$   
 547 CR (on the top row), and the  $WZ$  CR (on the bottom row). These are used to check the modeling of the  
 548 simulation for these features. In both CRs, there is good agreement between data and simulation as shown  
 549 by the Data/Pred. at the bottom the histograms. However, there is disagreement in the second last bin of the  
 550 Max Scoring  $W_{had} m_{jj}$  variable. This disagreement in the  $WZ$  CR is under investigation, but because there  
 551 are few events in this bin, it is not expected to have a significant impact on the analysis. This could also be  
 552 evidence that the estimation of backgrounds with fake leptons entirely from simulation is an inadequate  
 553 approach.

Not reviewed, for internal circulation only

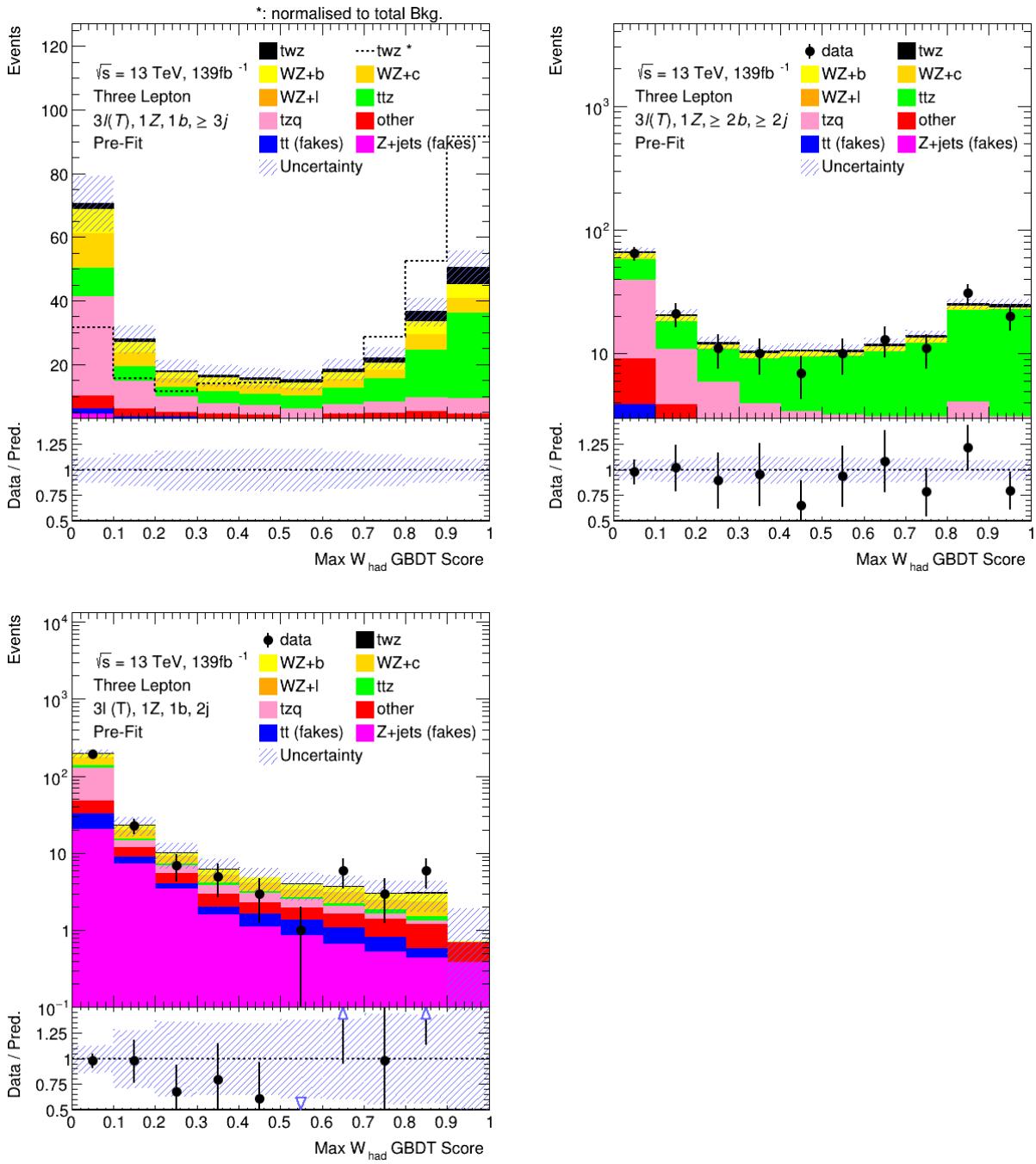


Figure 17: The  $\text{Max } W_{\text{had}}$  GBDT score is shown for simulation in the SR and CRs, and for data in the CRs. A distribution of  $tWZ$  events, normalized to the total number of background events, is shown in the SR by a dotted line.

Not reviewed, for internal circulation only

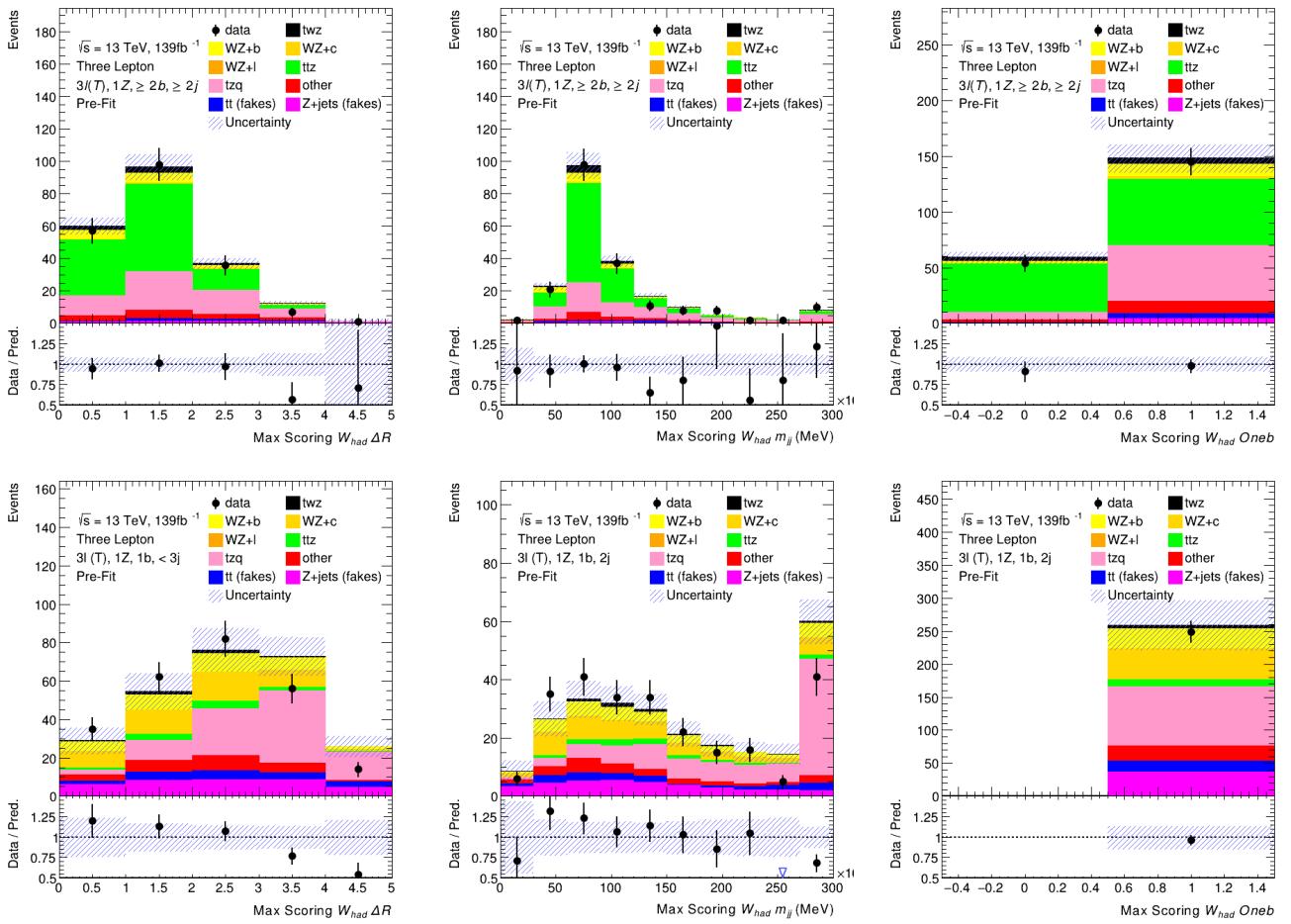


Figure 18: A comparison of data and simulation for  $W_{had}$  GBDT input features –  $\Delta R_{jj}$ ,  $m_{jj}$ , and  $Oneb$  – is shown for the maximum scoring dijet in the  $t\bar{t}Z$  CR – shown on the top row – and the  $WZ$  CR – shown on the bottom row.

## 554 5.6 Event-level GBDT

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

## 560 Labels and features

561 The events from  $tWZ$  production were labelled as the signal class and all background events were labelled  
 562 as the background class. To account for the expected number of signal and background events found in  
 563 the detector, each event was assigned a weight in the training that was equal to its simulated event weight.  
 564 Additional re-weighting of the signal class was applied to account for the imbalance in the signal and  
 565 background classes due to the rarity of  $tWZ$  events, this is dealt with in section 5.6.

566 In Figure 19, normalized distributions of the features used for training are shown for the half of all simulated  
 567 events used for training the event-level GBDT, with all the object and event selections applied.  $tWZ$   
 568 events are shown in red,  $t\bar{t}Z$  events in dark blue,  $WZ$  events in dark green,  $tZq$  events in purple, and the  
 569 combination of all the other smaller backgrounds shown in yellow.

570 • **Max  $W_{had}$  GBDT score**

571 This feature is the event-level discriminator outlined in the section 5.5. It is used to distinguish  
 572 between  $tWZ$  and  $WZ$  production. The normalized distributions for the different production  
 573 processes are shown in Figure 19(a). This feature does not help distinguish between  $tWZ$  and  $t\bar{t}Z$ ,  
 574 because  $t\bar{t}Z$  also has a  $W_{had}$  in the trilepton channel, but it does help distinguish between the other  
 575 backgrounds since a  $W_{had}$  is not expected in  $tZq$  or any of the other smaller background events.

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

577 In the event selections, exactly one  $Z$  Boson is required in all events. This leaves one lepton that does  
 578 not come from the  $Z$  Boson. This lepton is called the Non- $Z$  lepton. This feature is the minimum  
 579 invariant mass of the vectorial sum of the Non- $Z$  lepton and a b-tagged jet per event. This feature is  
 580 used to distinguish between  $tWZ$  and  $t\bar{t}Z$ , since one of the top quarks in  $t\bar{t}Z$  has to decay leptonically  
 581 in the three lepton channel. Therefore, the minimum vectorial sum of a lepton and b-tagged jet  
 582 originating from  $t\bar{t}Z$  should be more sharply peaked since these two objects decay from the same  
 583 top quark. In Figure 19(b), a normalized histogram of this variable is plotted for all the different  
 584 production processes. The difference between  $tWZ$  and these backgrounds can be seen since  $tWZ$   
 585 has a right shifted peak and a longer tail compared to  $t\bar{t}Z$  and  $tZq$ , whereas  $WZ$  and the other  
 586 production processes have a distribution that is similar to  $tWZ$  events.

587 • **HT**

588 This feature is the scalar sum of jet  $p_T$ 's. It does not help distinguish between  $tWZ$  and  $t\bar{t}Z$ , but does  
 589 help distinguish between  $tWZ$  and several other backgrounds. In Figure 19(c), normalized histogram  
 590 of this variable is plotted for all the different production processes. The difference between  $tWZ$  and  
 591 its backgrounds can be seen where the distributions of  $tWZ$  and  $t\bar{t}Z$  are shifted to the right compared  
 592 to the other backgrounds which are more prominent at low values of  $HT$ .

593 • **LT**

594 This feature is the scalar sum of lepton  $p_T$ 's. It is used to distinguish between  $tWZ$ ,  $tZq$ , and  
 595 combination of smaller background processes. In Figure 19(d), a normalized histogram of this  
 596 variable is plotted for all the different production processes, where the  $tZq$ , and other backgrounds  
 597 are more sharply peaked at lower values of  $LT$  compared to  $tWZ$ ,  $t\bar{t}Z$ , and  $WZ$ .

598 In Figure 20, the event-level GBDT features are shown for simulation and data in the  $t\bar{t}Z$  CR. In Figure  
 599 21 the same is shown in the  $WZ$  CR. In both cases there is generally good agreement, except in the Max  
 600  $W_{had}$  GBDT score in the  $WZ$  CR, as discussed previously, and in the  $LT$  feature in the  $WZ$  CR at high  $LT$ .  
 601 This is under investigation, but due to the small number of entries in the bins, these discrepancies are not  
 602 expected to have a significant effect. This could also be evidence that the estimation of backgrounds with  
 603 fake leptons entirely from simulation is an inadequate approach.

Not reviewed, for internal circulation only

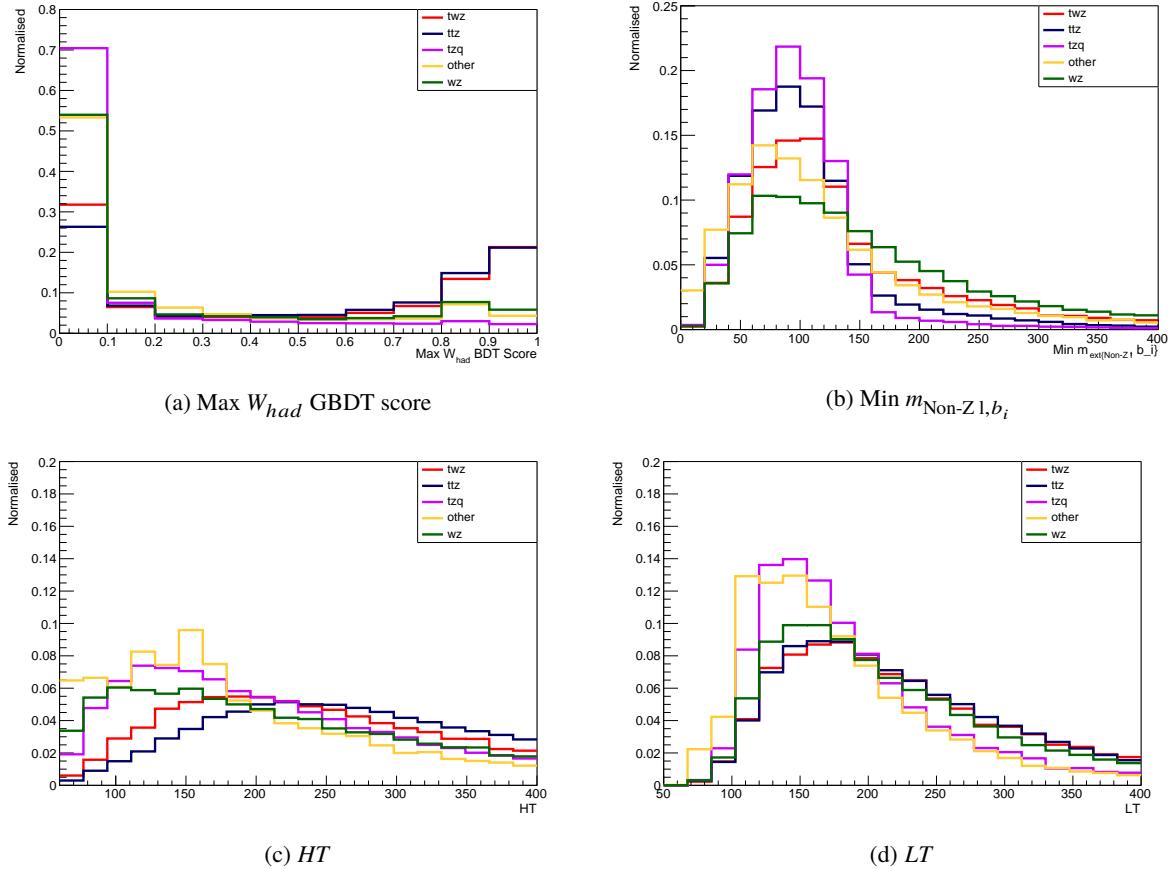


Figure 19: Normalized distribution of event GBDT features for  $tWZ$  and background events are shown. (a) shows the Max  $W_{had}$  GBDT score feature, (b) shows the Min  $m_{\text{Non-Z1}, b_i}$  feature, (c) shows the  $HT$  feature, and (d) shows  $LT$  feature

#### 604 Re-weighting of signal event weights

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

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

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

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

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

607 In addition to the re-weighting of signal event weights, all events with negative MC weights had to be  
608 removed before training and before the signal re-weighting was applied. This was done because of a

Not reviewed, for internal circulation only

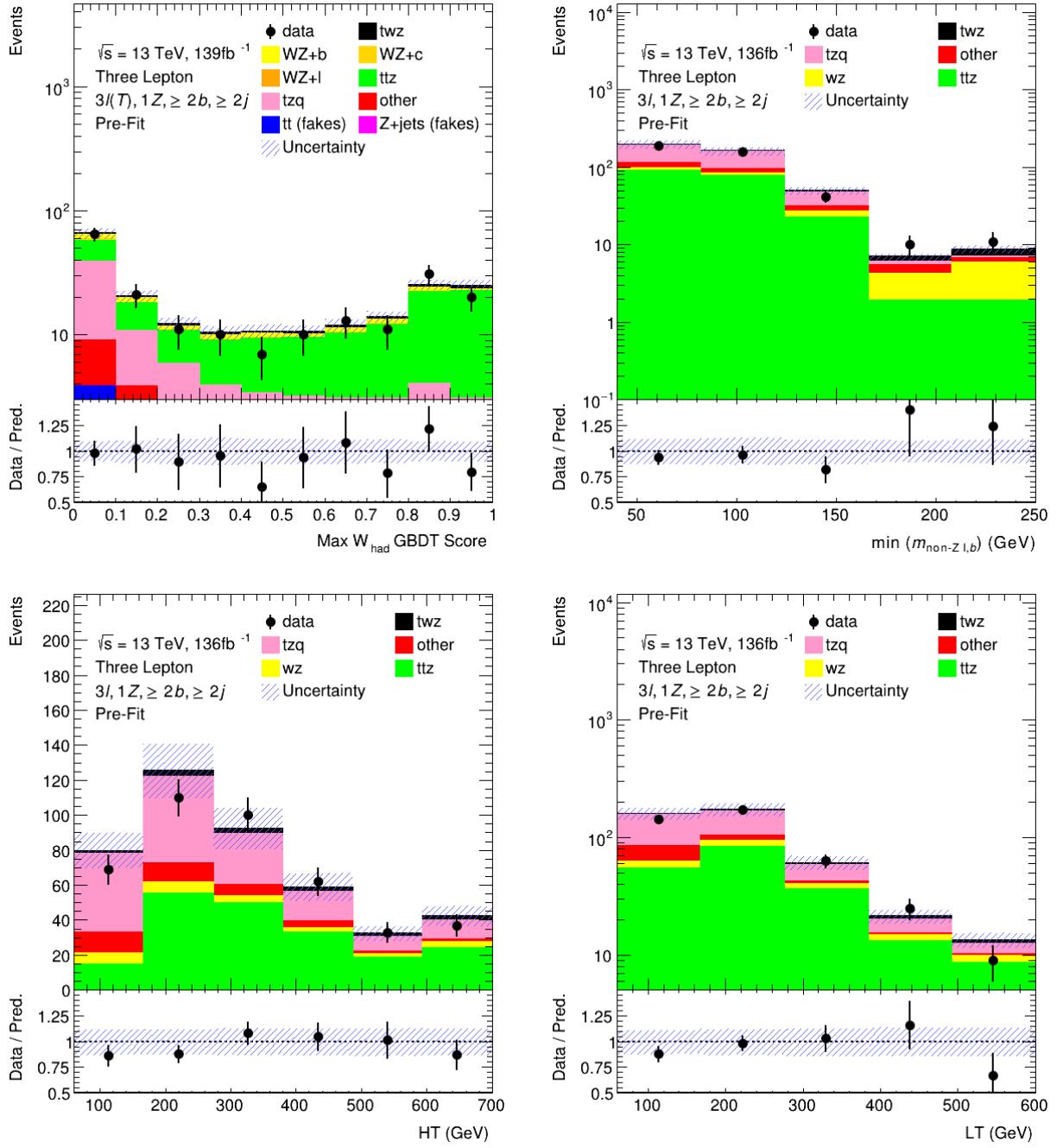


Figure 20: A comparison of data and simulation for the event-level GBDT input features –  $\text{Max } W_{\text{had}}$  GBDT score,  $\text{Min } m_{\text{Non-Z1},b}$ ,  $\text{HT}$ , and  $\text{LT}$  – is shown in the  $t\bar{t}Z$  CR

Not reviewed, for internal circulation only

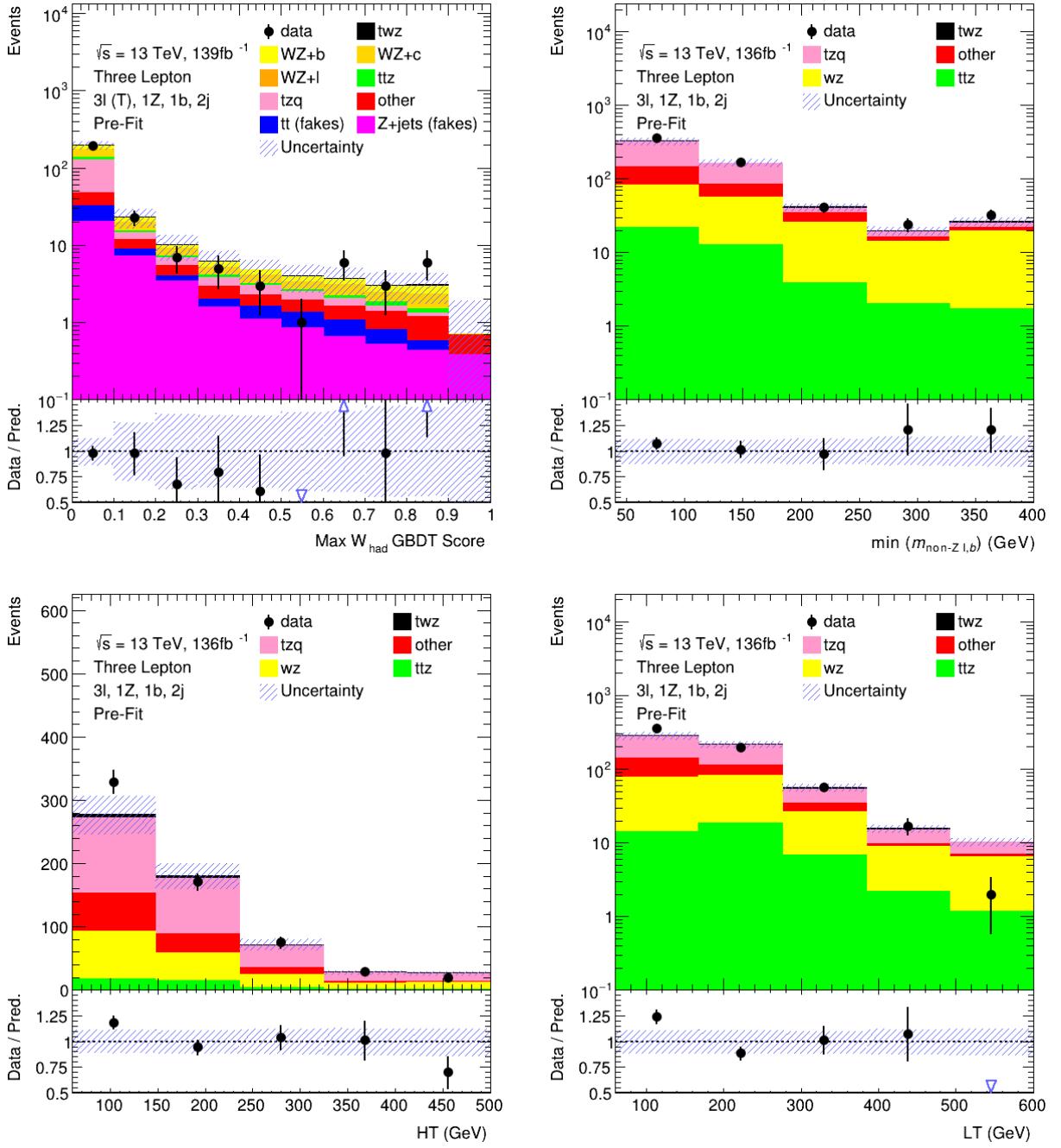


Figure 21: A comparison of data and simulation for the event GBDT input features – Max  $W_{had}$  GBDT score, Min  $m_{\text{Non-Z } l,b_i}$ ,  $HT$ , and  $LT$  – is shown in the  $WZ$  CR

609 technical limitation in the scikit-learn library. This means that the events used for training are not entirely  
610 representative of what is expected in data from the ATLAS detector. This is an unavoidable issue when  
611 using the sci-kit learn library and the effect is not expected to be significant due to the relatively small  
612 number of negatively weighted events compared to the total number of events.

613 **5.6.1 Event GBDT Score**

614 In Figure 22, the Event GBDT Score is shown for simulation and data in the SR and CRs, where the  
615 histograms have been auto-binned to minimize uncertainty due to statistical fluctuations. The autobinning  
616 algorithm starts from the bin with largest output and merges bins until a certain fraction of signal and  
617 background events is obtained [48]. In Figure 22(a), the Event GBDT Score is shown in the SR and is  
618 blind to data. A distribution of  $tWZ$  events, normalized to the total number of background events, is shown  
619 by a dotted line. This distribution shows that the Event GBDT Score discriminates well between signal and  
620 backgrounds. In Figures 22(b) and 22(c), the Event GBDT Score is shown in the  $t\bar{t}Z$  and  $WZ$  CRs. There  
621 is good agreement in the  $WZ$  CR, however, in the  $t\bar{t}Z$  CR there is disagreement in the bin from 0.3 to 0.4.  
622 This could be evidence that the estimation of backgrounds with fake leptons entirely from simulation is an  
623 inadequate approach.

Not reviewed, for internal circulation only

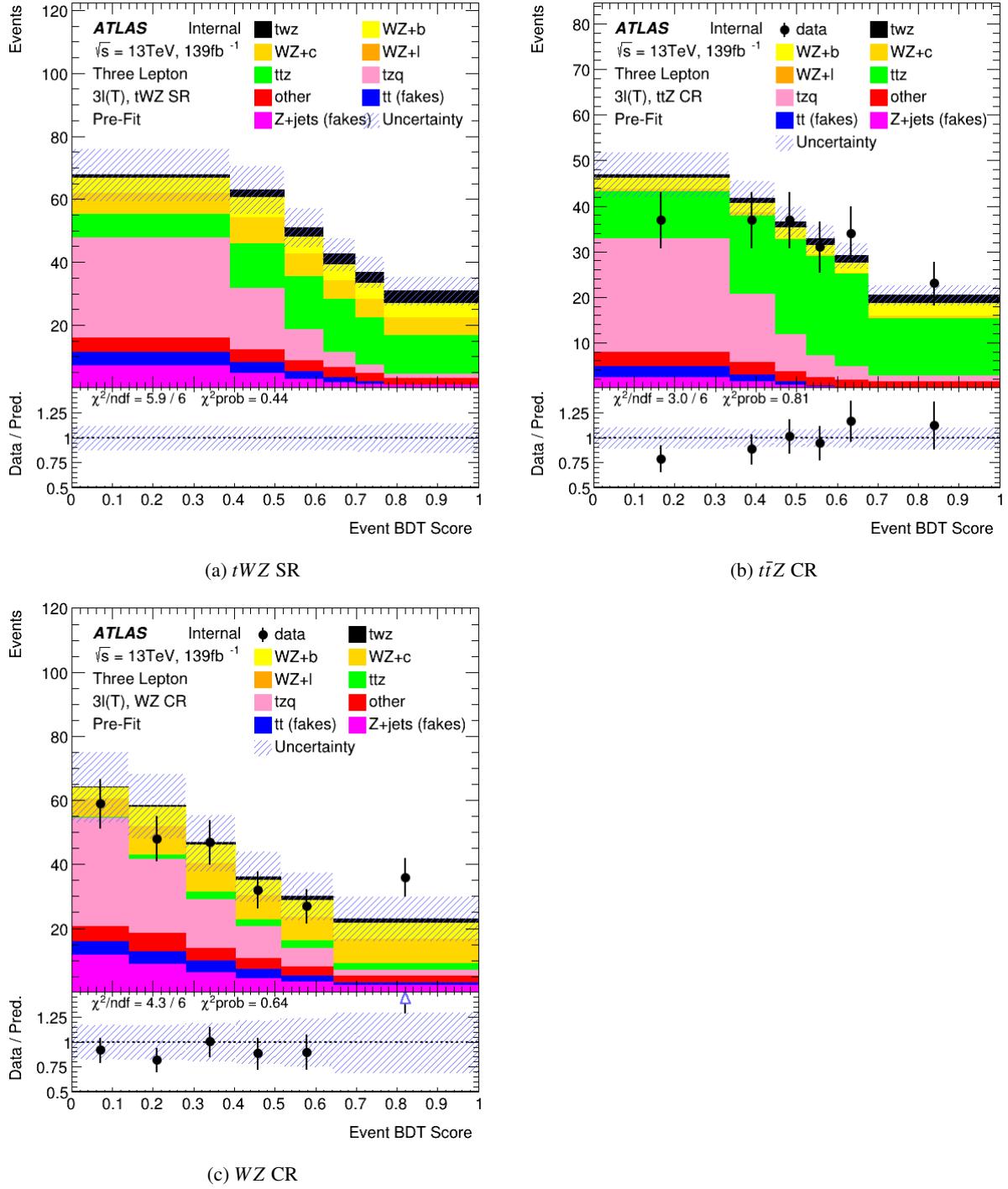


Figure 22: The Event GBDT Score is shown for simulation and data in the  $t\bar{t}Z$  CR (b) and the  $WZ$  CR (c). The  $tWZ$  SR is blind to data and a distribution of  $tWZ$  events, normalized to the total number of background events, is shown by a dotted line. In all regions the histograms have been auto-binned.

## 624 6 The Tetralepton Channel

### 625 6.1 Signal Topology & General Considerations

626 The  $tWZ \rightarrow 4\ell$  channel consists of a leptonically decaying  $W$  boson and a top quark which emits a  
 627 leptonically decaying  $Z$  boson before it decays into another leptonically decaying  $W$  boson and a  $b$ -quark.  
 628 The two leptons which decay from the  $Z$  boson must be of Opposite Sign and Same Flavour (OSSF).  
 629 Our final state consists of four leptons and a  $b$ -quark which motivates us to require a baseline selection  
 630 consisting of exactly four leptons and exactly one  $b$ -tagged jet. We can also note that, due to conservation  
 631 of charge, the sum of all the final state lepton charges must equal zero.

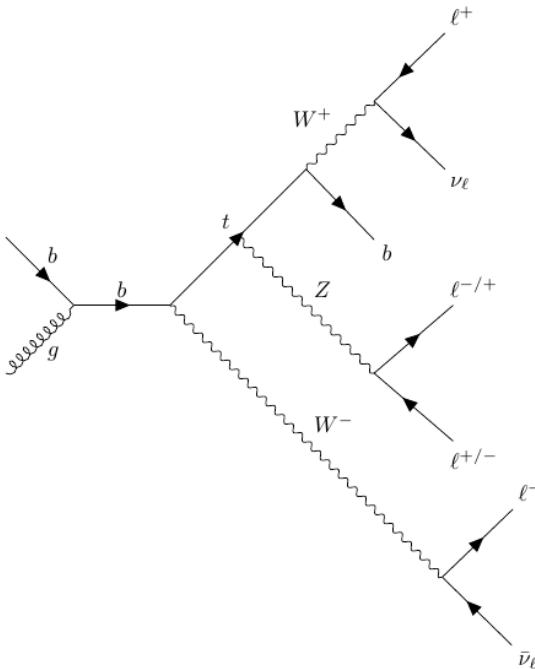


Figure 23: NLO Feynman diagram of  $tWZ$  in the tetralepton channel

#### 632 6.1.1 Comparison to the trilepton channel

633 The most apparent difference between the tri and tetralepton channels is the amount of statistics present,  
 634 with the tetralepton channel having far less events in its phase space than that of the trilepton channel.  
 635 The lack of statistics in the tetralepton channel can be attributed to its low production cross section,  
 636  $\sigma_{(tW^\pm Z).Br(4\ell)}^{\text{NLO}} = 0.7 \text{ fb}$ [1]. The trilepton channel has a production cross section ( $\sigma_{(tW^\pm Z).Br(3\ell)}^{\text{NLO}} = 3.9 \text{ fb}$ [1])  
 637 around a factor of 4 larger than that of the tetralepton channel. This difference between the production  
 638 cross section of the two decay channels can be largely attributed to the difference in branching ratios ( $\frac{\Gamma_i}{\Gamma_W}$ )  
 639 between a hadronically decaying  $W$  boson,  $\frac{\Gamma_{W \rightarrow \text{had}}}{\Gamma_W} = (67.41 \pm 0.27)\%$ [49], present in the trilepton channel  
 640 and a leptonically decaying  $W$  boson,  $\frac{\Gamma_{W \rightarrow \ell\nu}}{\Gamma_W} = (10.86 \pm 0.09)\%$ [49], present in the tetralepton channel.

641 Despite the tetralepton channel's low statistics, it is not subject to the large  $WZ$  background present in  
642 the trilepton channel. The tetralepton channel has a relatively large  $ZZ$  background (not present in the  
643 trilepton channel), fortunately this can be easily suppressed due to the full reconstructability of the two  
644 leptonically decaying  $Z$ -bosons.

[Not reviewed, for internal circulation only]

## 645 6.2 Baseline selections

646 In addition to our selection criteria of exactly four tight leptons, we require that the leading, next-to-leading,  
 647 next-to-next-to-leading and next-to-next-to-next-to-leading leptons have  $p_T$  greater than 28, 18, 10 and  
 648 10 GeV respectively. Here we have chosen to apply looser object-level cuts than the tri-lepton channel in  
 649 an attempt to maximize the amount of events in our signal regions, as the tetralepton channel is heavily  
 650 statistically limited. For the same reasoning, we have required  $p_T(\text{jet}) > 20 \text{ GeV}$ .

651 The final state lepton charges must sum to zero.

652 We therefore require  $\sum_{i=1}^4 \text{charge}(\ell_i) = 0$ .

653 Reconstructed electrons are required to be within  $|\eta| < 2.47$  and excluding the transition region between  
 654 the barrel and end-cap calorimeters at  $1.37 < |\eta| < 1.52$ . Reconstructed muons and jets are required to be  
 655 within  $|\eta| < 2.5$ . The jet-vertex-tagger (jvt) on jets are required to have a value greater than 0.5, in an  
 656 attempt to reject effects caused by pile-up interactions.

657 The invariant mass of the OSSF lepton pair coming from the  $Z$  boson must equal the invariant mass of  
 658 the  $Z$  boson, and noting that lepton reconstruction and identification in the ATLAS detector has a high  
 659 accuracy [], we can use these OSSF leptons to reconstruct the  $Z$  boson with relatively high confidence.  
 660 We therefore define a  $Z$  candidate as an OSSF lepton pair with an invariant mass,  $m_{\text{OSSF}}$ , satisfying the  
 661 condition,  $|m_{\text{OSSF}} - m_Z| < 30 \text{ GeV}$ , where  $m_Z$  is the nominal  $Z$  boson mass (91.1876 GeV [49]). Multiple  
 662  $Z$  candidates can be present in certain decay channels (e.g.  $eeee$ ,  $\mu\mu ee$ ,  $\mu\mu\mu\mu$ ). In these cases, the  $Z$   
 663 candidate which has an invariant mass closest to the nominal  $Z$  boson mass is chosen.

664 In order to suppress quarkonia (low mass resonances such as  $J/\psi$  and  $\Upsilon$ ) we require that all OSSF  
 665 lepton pairs have an invariant mass,  $m_{\text{OSSF}}$ , greater than 10 GeV.

666 We define two  $tWZ$  SRs, as opposed to one  $tWZ$  SR, in an attempt to suppress and constrain the  $ZZ$   
 667 background. We require that both  $tWZ$  SRs have exactly four tight leptons, exactly one  $Z$ -boson candidate,  
 668 exactly one tight  $b$ -tagged jet (from the decay of the top quark) and greater than or equal to one jet. The  
 669 two  $tWZ$  SR's differ by the flavours of their leptons which don't originate from the decay of a  $Z$ -boson  
 670 (Non- $Z$  leptons). The  $ZZ$  background has two  $Z$ -bosons which decay into a pair of OSSF lepton pairs, in  
 671 order to mimic the  $tWZ$  signal. We take advantage of this to define a  $tWZ$  region rich in  $ZZ$  background  
 672 and one with a minimal  $ZZ$  background component. This is done by requiring that one of the  $tWZ$  SRs has  
 673 its two non- $Z$  leptons to have opposite flavour and the other  $tWZ$  SR is required to have its non- $Z$  leptons  
 674 to have the same flavour. These two disjoint  $tWZ$  SRs are named  $tWZ$  OF SR and  $tWZ$  SF SR respectively.  
 675 We therefore expect that the  $tWZ$  SF SR contains the majority of  $tWZ$ 's  $ZZ$  background events.

676 The selection criteria which define the SR and the CRs are summarised in Table 10. In order to check  
 677 the modelling of the most dominant background components in our signal region, we have modified our  
 678 selection criteria to define  $t\bar{t}Z$  and  $ZZb$  control regions. The  $t\bar{t}Z$  control region has the same requirement  
 679 on the number of reconstructed  $Z$  boson candidates in the signal region (due to a commonality on the  
 680 number of  $Z$  bosons present in both processes), however we require at least two jets and that exactly two of  
 681 these jets are  $b$ -tagged (corresponding to the  $b$ -quark jets originating from the two top-quark decays). We  
 682 choose to define a  $ZZb$  region, as opposed to a  $ZZ$  region, since the  $ZZ$  background present in the  $tWZ$   
 683 signal region contains exactly one  $b$ -tagged jet. Therefore defining a region with  $ZZ$  plus exactly one  $b$ -jet  
 684 more closely resembles the  $ZZ$  background present in the signal region. In addition to this, mis-modelling  
 685 of  $ZZ$  has been seen in other analyses [50, 51], further motivating the use of a  $ZZb$  control region over a

Baseline selections				
$N_\ell = 4$				
$p_T(\ell_1, \ell_2, \ell_3, \ell_4) > (28, 18, 10, 10) \text{ GeV}$				
$p_T(\text{jet}) > 20 \text{ GeV},  \eta(\text{jet})  < 4.5, \text{jvt} > 0.5$				
$ \eta(\ell_e)  < 2.47 \text{ 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 $m_{\text{OSSF}} > 10 \text{ GeV}$				
Regions				
$tWZ$ OF SR	$tWZ$ SF SR	$t\bar{t}Z$ CR	$ZZb$ CR	$(tWZ)_{\text{fake}}$ CR
$N_\ell(\text{tight}) = 4$	$N_\ell(\text{tight}) = 4$	$N_\ell(\text{tight}) = 4$	$N_\ell(\text{tight}) = 4$	$N_\ell(\text{tight}) = 3$ $N_\ell(\text{loose and NOT tight}) = 1$
$N_Z \text{ candidate} = 1$	$N_Z \text{ candidate} = 1$	$N_Z \text{ candidate} = 1$	$N_Z \text{ candidate} = 2$	$N_Z \text{ candidate} = 1$
$N_{\text{jet}} \geq 1$	$N_{\text{jet}} \geq 1$	$N_{\text{jet}} \geq 2$	$N_{\text{jet}} \geq 1$	$N_{\text{jet}} \geq 1$
$N_{\text{b-jet}}(\text{tight}) = 1$	$N_{\text{b-jet}}(\text{tight}) = 1$	$N_{\text{b-jet}}(\text{tight}) \geq 1$ $N_{\text{b-jet}}(\text{loose}) \geq 0$ $N_{\text{b-jet}}(\text{tight}) + N_{\text{b-jet}}(\text{loose}) = 2$	$N_{\text{b-jet}}(\text{tight}) = 1$	$N_{\text{b-jet}}(\text{tight}) = 1$
Opp. Flavour Non-Z leptons	Same Flavour Non-Z leptons	-	-	-

Table 10: Summary of the requirements applied for selecting events in the signal and control regions.

686  $ZZ$  CR. The  $ZZb$  CR requires exactly two  $Z$  boson candidates and exactly one b-tagged jet, resulting in an  
 687 implicit requirement on the number of jets ( $N_{\text{jet}} \geq 1$ ).

688 Since we are heavily statistically limited, we aim to increase the amount of events in our regions. In an  
 689 attempt to achieve this goal in the  $t\bar{t}Z$  CR,  $b$ -tagged jets were placed under *tight* and *loose* definitions. A  
 690 tight  $b$ -tagged jet is defined as a jet which passes the 77%, 70%, 65% or 60% DL1r  $b$ -tagger working point.  
 691 A loose  $b$ -tagged jet is defined as a jet which passes 85% DL1r  $b$ -tagger working point, but not the 77%,  
 692 70%, 65% or 60% DL1r  $b$ -tagger working points. Different numbers (and definitions) of tight and loose  
 693  $b$ -tagged jets were tried in each region, with the final selection criteria being chosen which maximised the  
 694 expect significance of  $tWZ$ .

695 In order to constrain the fake lepton component contained within the  $t\bar{t}Z$  sample, we define a  $(tWZ)_{\text{fake}}$  CR  
 696 which is as similar as possible to the  $tWZ$  CR but is enhanced in fakes. This is achieved by defining the  
 697  $(tWZ)_{\text{fake}}$  CR to inherit the same selection criteria as the  $t\bar{t}Z$  CR however, in this case, we require exactly 3  
 698 tight leptons and exactly 1 loose (and NOT tight) lepton (since looser leptons are more likely to be fakes,  
 699 compared to tighter leptons).

700 The final selection criteria and region definitions are summarised in Table 10.

### 701 6.2.1 Optimization studies for event selection

702 In order to find the selection criteria for jets and leptons which maximized sensitivity to the  $tWZ$  signal,  
 703 studies were performed by plotting the expected significance ( $Z_\mu^{\text{exp}}$ ) and expected upper limit ( $\mu_{\text{up}}^{\text{exp}}$ ) for  
 704 different selection criteria.

705 The expected significance and limits were computed for a range of  $\eta_{\text{jet}}$  cuts.

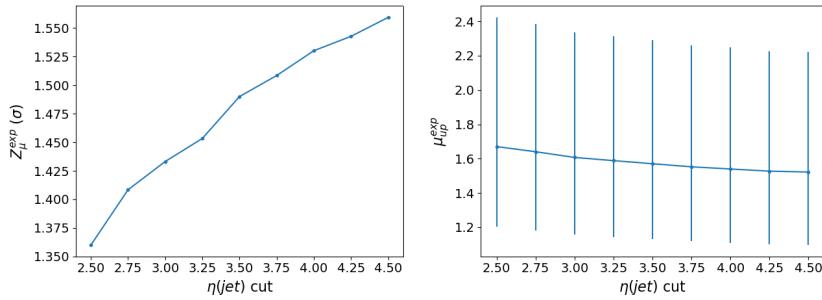


Figure 24: Left: Expected significance ( $Z_\mu^{\text{exp}}$ ) for different  $\eta(\text{jet})$  cuts. Right: Expected upper limit ( $\mu_{\text{up}}^{\text{exp}}$ ) for different  $\eta(\text{jet})$  cuts.

From Figure 24, we can see that the  $\eta(\text{jet})$  cut which maximises the sensitivity of  $tWZ$  in the tetralepton channel is requiring that  $\eta(\text{jet}) < 4.5$ . This selection criteria was set for the  $\eta(\text{jet})$  across all regions.

The expected significance and limits were computed for a range of  $p_T(\text{jet})$  cuts.

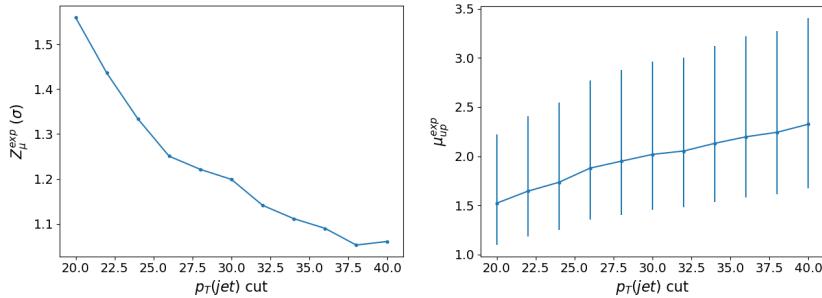


Figure 25: Left: Expected significance ( $Z_\mu^{\text{exp}}$ ) for different  $p_T(\text{jet})$  cuts. Right: Expected upper limit ( $\mu_{\text{up}}^{\text{exp}}$ ) for different  $p_T(\text{jet})$  cuts

From Figure 25, we can see that the  $p_T(\text{jet})$  cut which maximises the sensitivity of  $tWZ$  in the tetralepton channel is requiring that  $p_T(\text{jet}) > 20\text{GeV}$ . This selection criteria was set for the  $p_T(\text{jet})$  across all regions.

The expected significance and limits were computed for a range of different configurations of DL1r  $b$ -tagged jet working points across different regions.

From Figure 26, we can see that requiring that  $b$ -tagged jets pass the 77% DL1r WP in the  $tWZ$  SR,  $(tWZ)_{\text{fake}}$  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 just required to pass the 85% DL1r WP) maximises the sensitivity overall (compared to the other investigated configurations). This configuration was chosen  $b$ -tagged jets.

The  $p_T(\text{L Lepton})$  is constrained by the single lepton triggers (Table 2). We choose to apply a cut on the  $p_T(\text{NL Lepton})$  slightly tighter than the tightest single lepton  $p_T$  cut in the trigger. We can however, try optimising the  $p_T(\text{NL Lepton})$  cut by comparing the expected significance and limit for a range of  $p_T(\text{NL Lepton})$  cuts to determine the cut which maximizes sensitivity.

Not reviewed, for internal circulation only

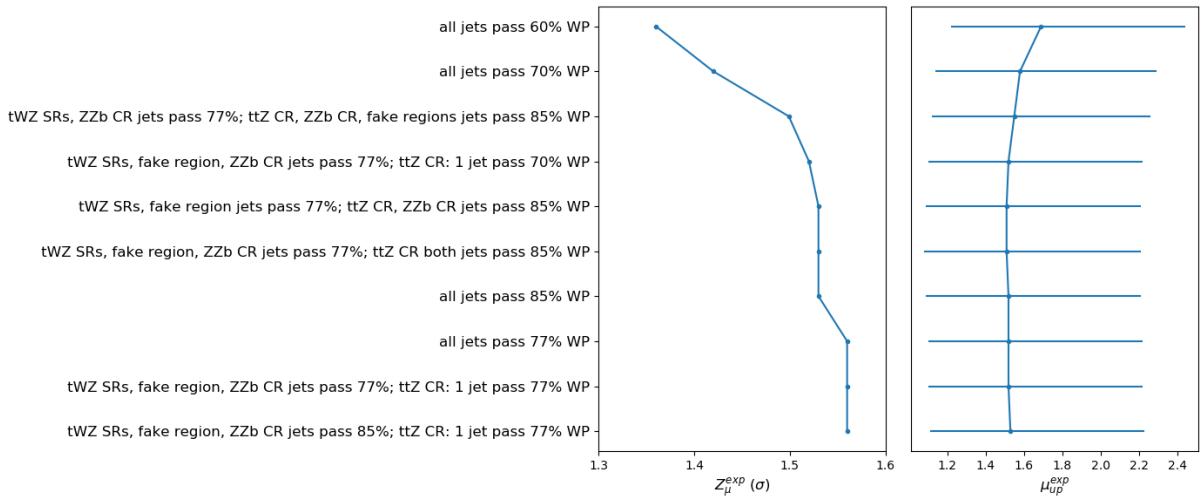


Figure 26: Expected significance ( $Z_\mu^{exp}$ ) and expected upper limit ( $\mu_{up}^{exp}$ ) for different configurations of DL1r  $b$ -tagged jet working points

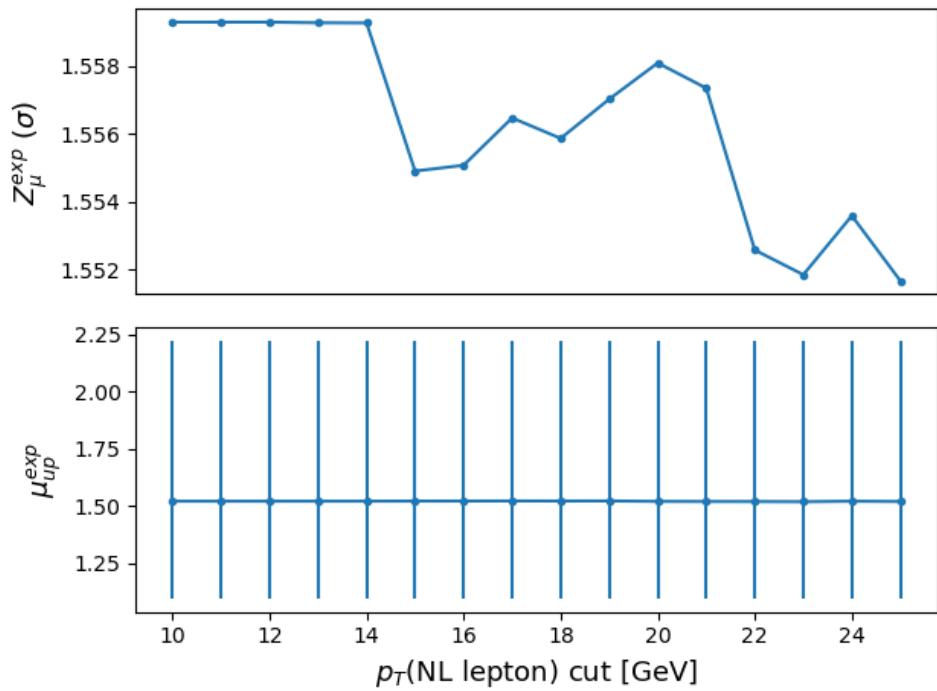


Figure 27: Expected significance ( $Z_\mu^{exp}$ ) and expected upper limit ( $\mu_{up}^{exp}$ ) for different  $p_T(\text{NL Lepton})$  cuts

722 Since there is a very small change between the different  $p_T(\text{NL Lepton})$  cuts on the sensitivity of  $tWZ$  in  
 723 the tetralepton channel, we choose to apply a  $p_T(\text{NL Lepton})$  cut at 18 GeV, therefore applying a cut above  
 724 the trigger  $p_T$  cut to suppress any systematic from the modelling of the trigger efficiency.

<sup>725</sup> **6.3 Signal and Control Regions**

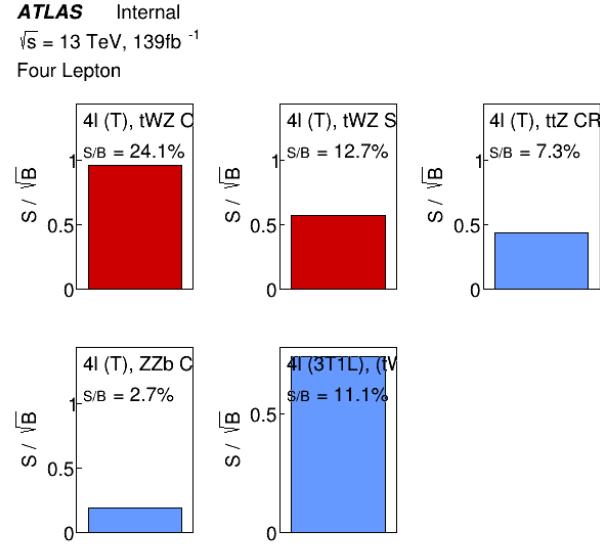


Figure 28: Plots of  $\frac{S}{\sqrt{B}}$  ( $\frac{\text{Signal}}{\sqrt{\text{Background}}}$ ) for the signal and control regions

	tWZ OF SR	tWZ SF SR	t̄tZ CR	ZZb CR	(tWZ) <sub>fake</sub> CR
t̄tZ	$13.9325 \pm 1.84643$	$10.1343 \pm 1.36039$	$31.7149 \pm 4.46776$	$5.26303 \pm 0.696828$	$19.1224 \pm 2.50011$
t̄tZ 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	$6e-06 \pm 3.04506e-06$	$0.250783 \pm 0.44226$	$0.269883 \pm 0.223373$	$6e-06 \pm 3.04506e-06$	$2.36284 \pm 0.927828$
	$0.0827265 \pm 0.0399222$	$0.0757694 \pm 0.0355101$	$0.0637132 \pm 0.0293762$	$0.0590199 \pm 0.0244576$	$4.91371 \pm 0.754695$
	$0.00674747 \pm 0.00793546$	$0.00279491 \pm 0.00287747$	$6e-06 \pm 3.04506e-06$	$0.00221727 \pm 0.00562041$	$0.944039 \pm 0.296854$
	$0.0439316 \pm 0.0241635$	$0.0397876 \pm 0.0154764$	$0.0134837 \pm 0.0128327$	$0.0474188 \pm 0.0330635$	$1.84471 \pm 0.397076$
	$0.000987429 \pm 0.000768187$	$0.00249801 \pm 0.00138007$	$0.0141085 \pm 0.00486102$	$6e-06 \pm 3.04506e-06$	$0.0100745 \pm 0.00367677$
	$0.00934516 \pm 0.0080725$	$0.0107503 \pm 0.00852049$	$0.0570846 \pm 0.0206271$	$6e-06 \pm 3.04506e-06$	$0.0216609 \pm 0.00999533$
	$0.0293456 \pm 0.0263573$	$0.0296011 \pm 0.0196075$	$0.26412 \pm 0.0936908$	$0.013096 \pm 0.0323943$	$0.151267 \pm 0.0593376$
	$0.280384 \pm 0.0866421$	$0.191257 \pm 0.0595588$	$0.0696624 \pm 0.0228108$	$0.171171 \pm 0.0526519$	$0.265957 \pm 0.0821857$
	$0.854064 \pm 0.177974$	$0.674566 \pm 0.141771$	$1.98187 \pm 0.406211$	$0.151447 \pm 0.0357703$	$2.22981 \pm 0.45726$
	$0.0439316 \pm 0.0260839$	$0.0397876 \pm 0.0178514$	$0.0134837 \pm 0.0131821$	$0.0474188 \pm 0.0347221$	$1.84471 \pm 0.572552$
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	24	26	36	49	57

Table 11: Pre-fit yields for each sample in each region

726 The statistical uncertainty associated with the number of raw MC events (weighted by MC event weights)  
 727 can be quantified by the Number of Equivalent Events [52],  $N_{equiv}$ , which relates the sample of  $N$  weighted  
 728 events to  $N_{equiv}$  events with all MC event weights equal to 1, that would have the same relative statistical  
 729 fluctuation.  $N_{equiv}$  can be written as,

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

730 where  $w_i$  is the MC event weight for event  $i$ . The standard uncertainty of  $N_{equiv}$  is given by  $u(N_{equiv}) =$   
 731  $\sqrt{N_{equiv}}$ .

	$tWZ$ SR		$ttZ \bar{t}tZ$ CR		$ZZb$ $ZZb$ CR	
	$N_{equiv}$	$\frac{u(N_{equiv})}{N_{equiv}} \times 100$ [%]	$N_{equiv}$	$\frac{u(N_{equiv})}{N_{equiv}} \times 100$ [%]	$N_{equiv}$	$\frac{u(N_{equiv})}{N_{equiv}} \times 100$ [%]
$tWZ$	$2131.35 \pm 46.17$	2.17	$629.67 \pm 25.09$	3.99	$410.79 \pm 20.27$	4.93
$t\bar{t}Z$	$7984.99 \pm 89.36$	1.12	$6230.91 \pm 78.94$	1.27	$1678.56 \pm 40.97$	2.44
$ZZ$	$2846.93 \pm 53.36$	1.87	$340.73 \pm 18.46$	5.42	$15387.97 \pm 124.05$	0.81
other	$14.06 \pm 3.75$	26.67	$10.83 \pm 3.29$	30.39	$361.12 \pm 19.0$	5.26
Total	$7212.15 \pm 84.92$	1.18	$12977.33 \pm 113.92$	0.88	$17838.44 \pm 133.56$	0.75

Table 12:  $N_{equiv}$  and  $\frac{u(N_{equiv})}{N_{equiv}} \times 100$  for each sample in each region ( $tWZ$  SR,  $t\bar{t}Z$  CR,  $ZZb$  CR)

Not reviewed, for internal circulation only

### 732 6.3.1 $tWZ$ OF SR

733 In this section, pre-fit distributions of variables in the  $tWZ$  OF signal region are shown. More pre-fit  
 734 distributions for the  $tWZ$  OF SR are shown in the appendix (Section A.1).

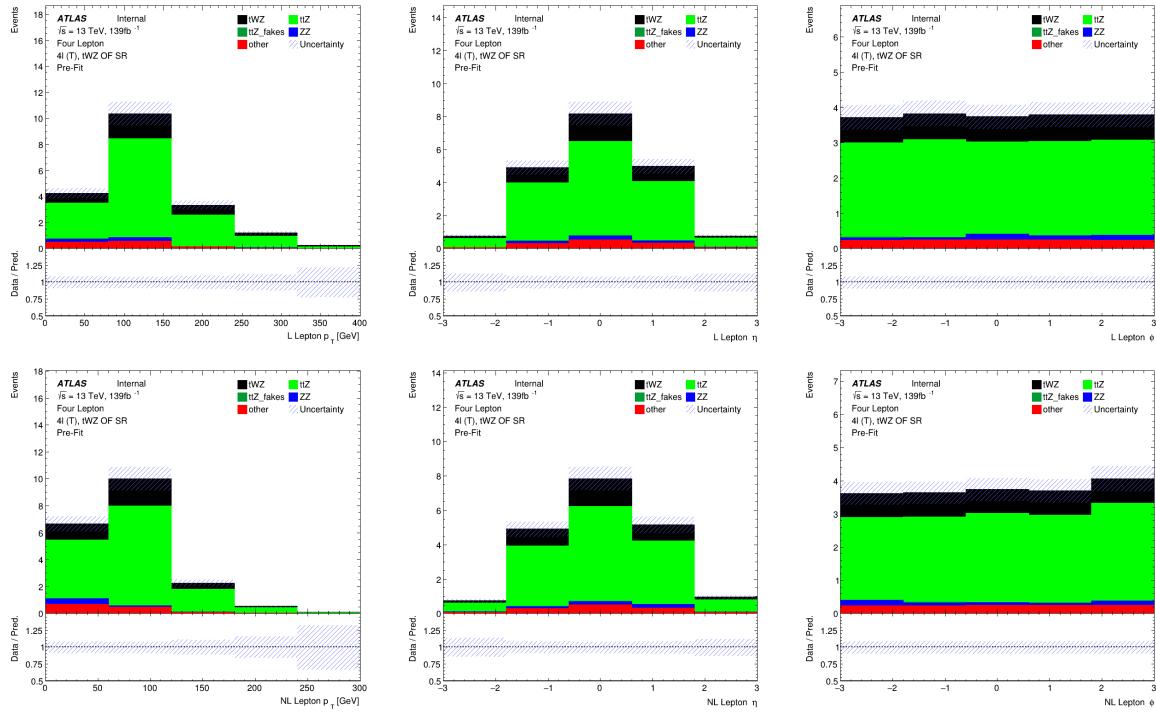


Figure 29: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  for leading (L) leptons (top row) and next-to-leading (NL) leptons (bottom row) in the  $tWZ$  SR region (*blinded*)

### 735 6.3.2 $tWZ$ SF SR

736 In this section, pre-fit distributions of variables in the  $tWZ$  SF signal region are shown. More pre-fit  
 737 distributions for the  $tWZ$  SF SR are shown in the appendix (Section A.2).

Not reviewed, for internal circulation only

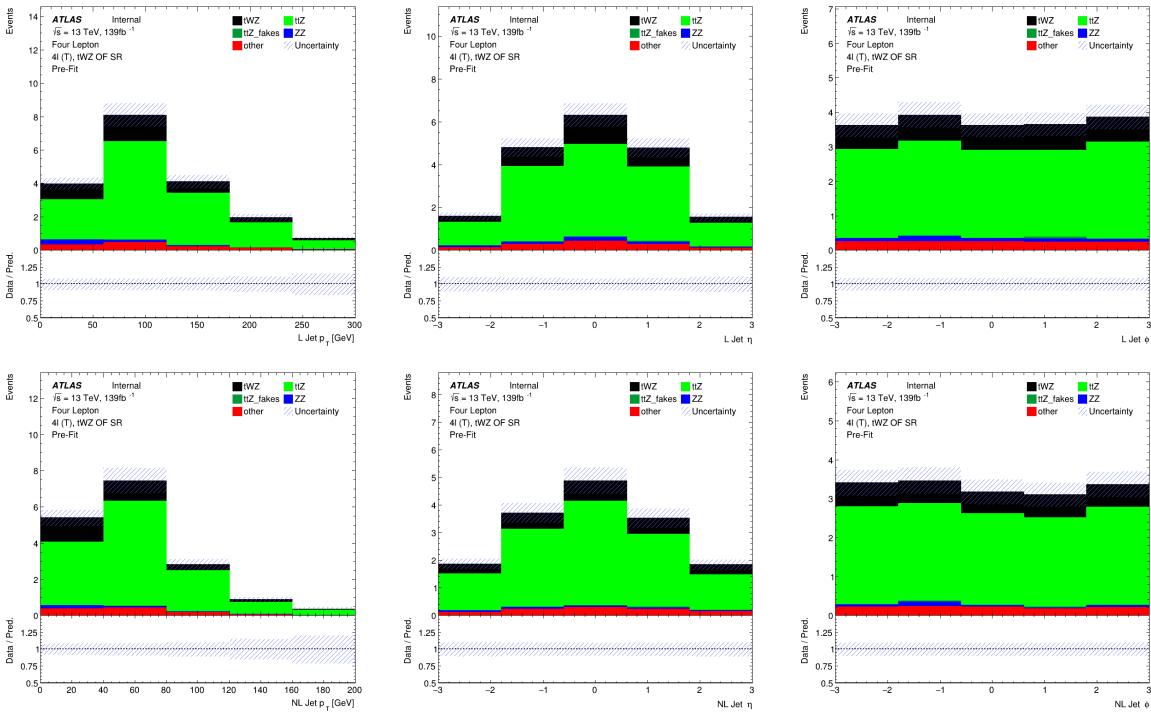


Figure 30: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  leading (L) jets (top row) and next-to-leading (NL) jets (bottom row) in the  $tWZ$  SR region (*blinded*)

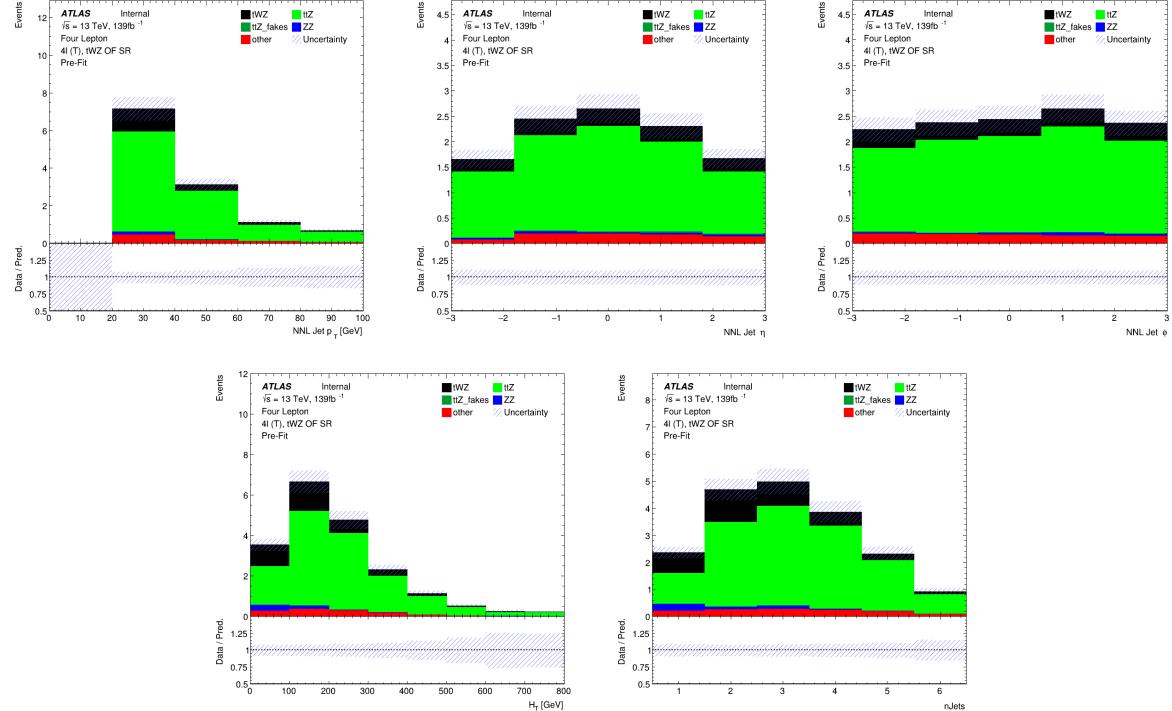


Figure 31: Top row: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  for next-to-next-to-leading (NNL) jets in the  $tWZ$  SR region (*blinded*). Bottom row: MC predictions for  $H_T$  (scalar sum of Jet  $p_T$ ) (left) and the Number of jets (right) in the  $tWZ$  SR region (*blinded*)

Not reviewed, for internal circulation only

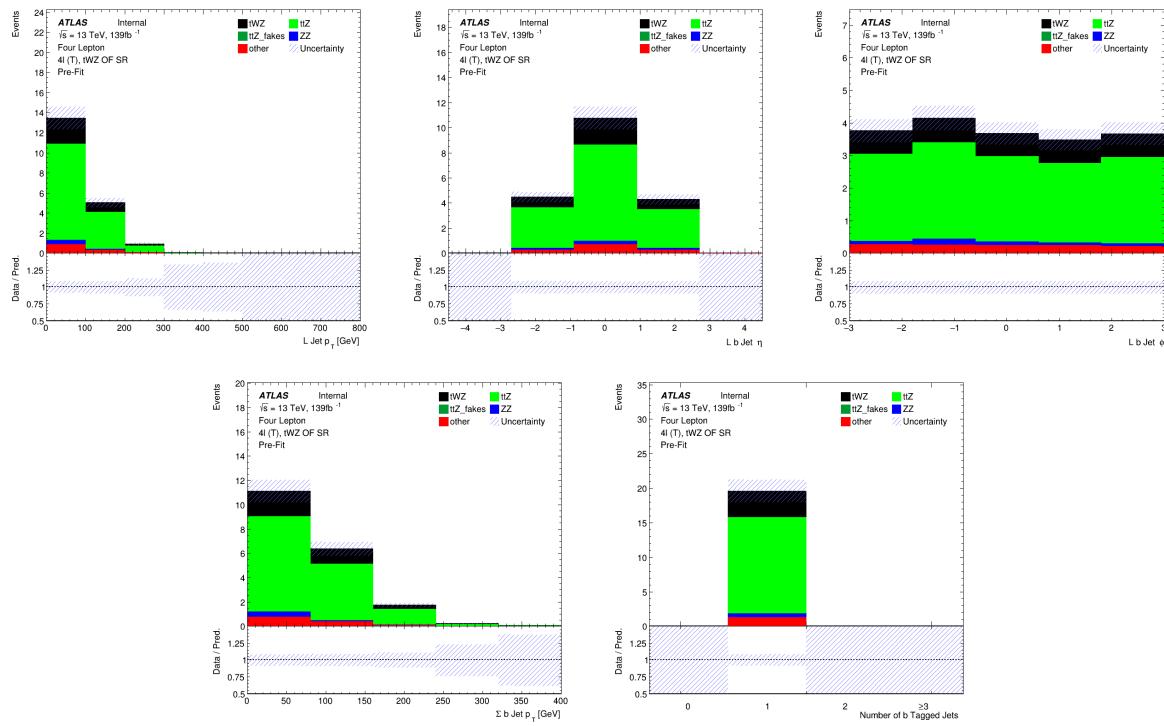


Figure 32: Top row: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  for leading b-tagged jets in the  $t\bar{W}Z$  SR region (*blinded*). Bottom row: MC predictions for the scalar sum of b-tagged jet  $p_T$  (left) and the Number of b-tagged jets (right) in the  $t\bar{W}Z$  SR region (*blinded*)

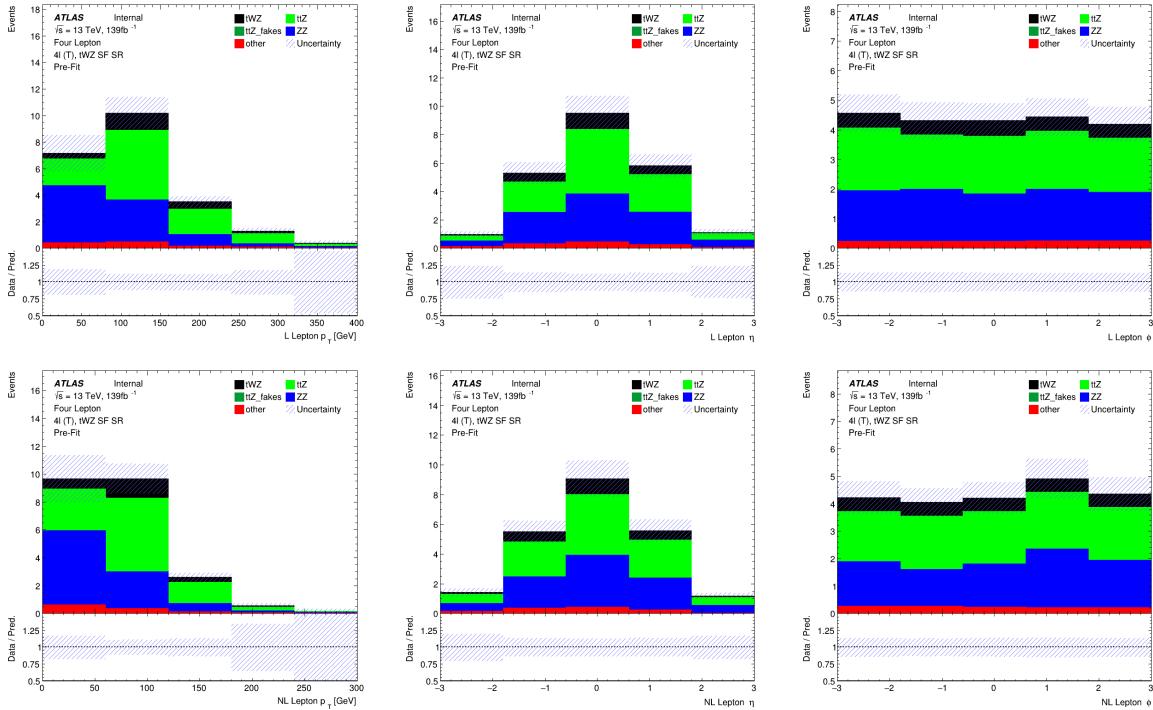


Figure 33: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  for leading (L) leptons (top row) and next-to-leading (NL) leptons (bottom row) in the  $t\bar{W}Z$  SR region (*blinded*)

Not reviewed, for internal circulation only

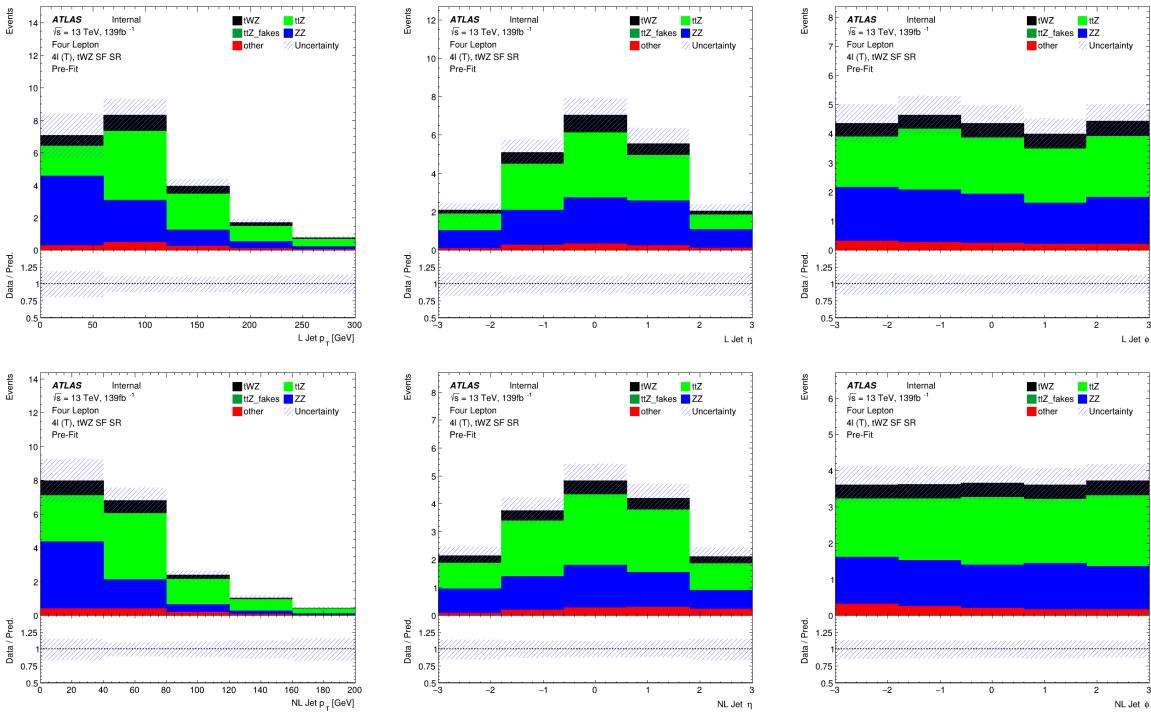


Figure 34: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  leading (L) jets (top row) and next-to-leading (NL) jets (bottom row) in the  $tWZ$  SR region (*blinded*)

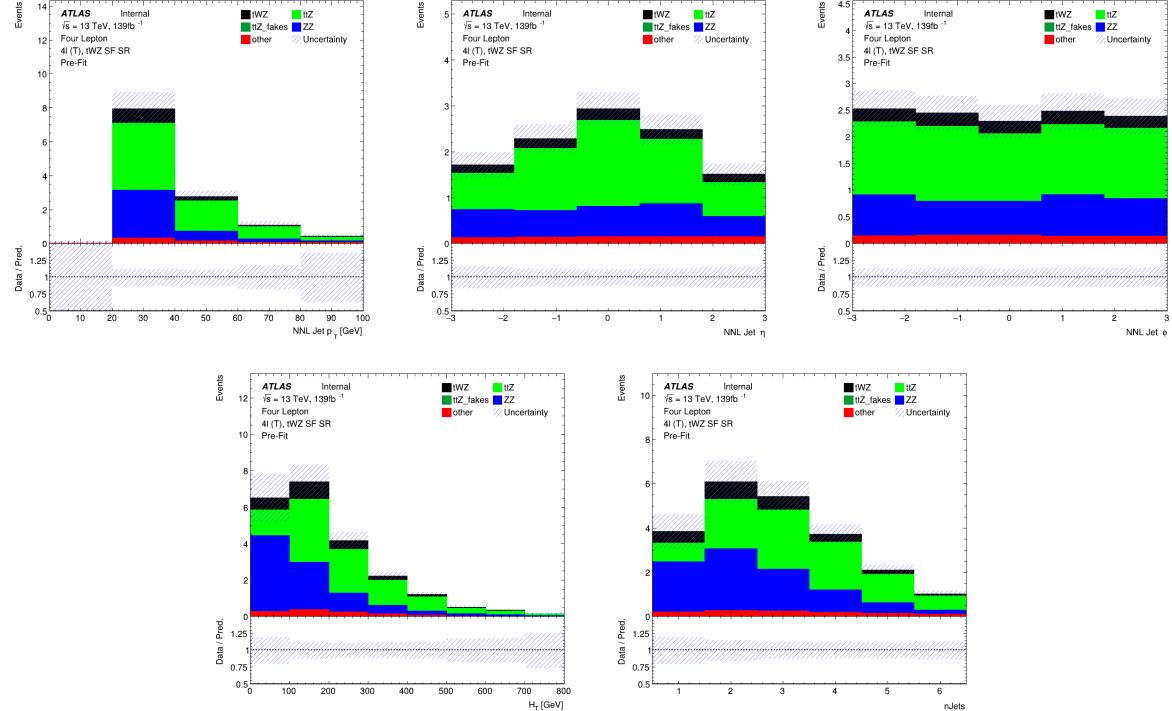


Figure 35: Top row: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  for next-to-next-to-leading (NNL) jets in the  $tWZ$  SR region (*blinded*). Bottom row: MC predictions for  $H_T$  (scalar sum of Jet  $p_T$ ) (left) and the Number of jets (right) in the  $tWZ$  SR region (*blinded*)

Not reviewed, for internal circulation only

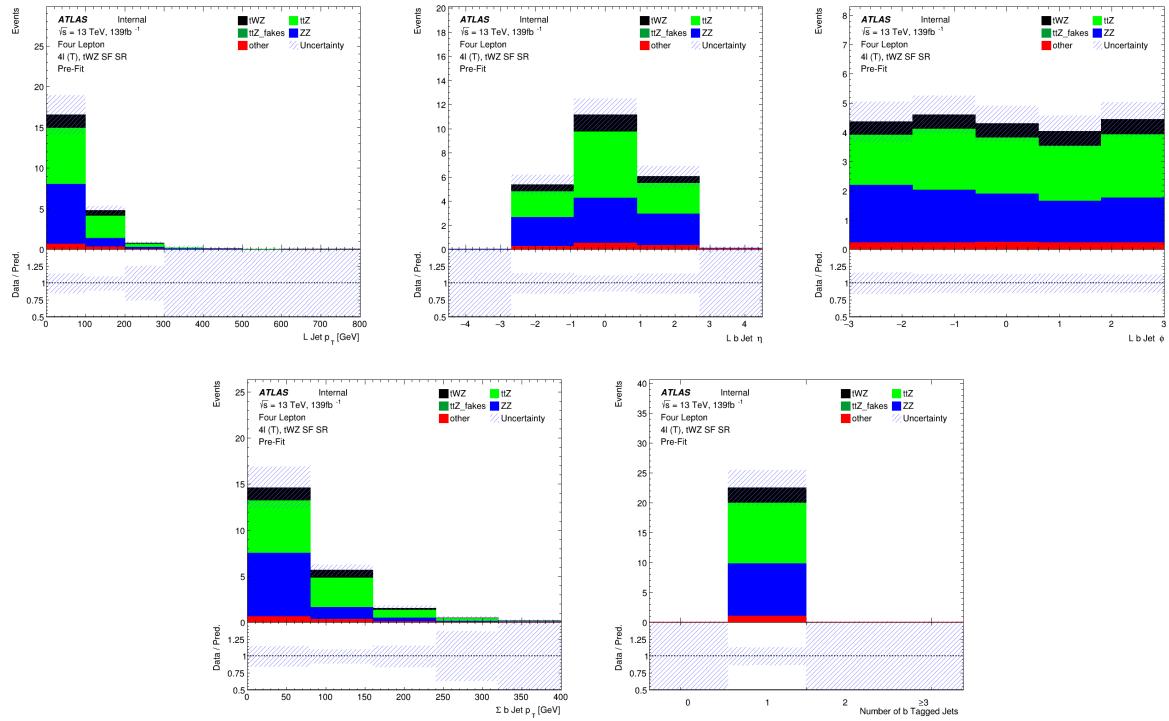


Figure 36: Top row: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  for leading b-tagged jets in the  $tWZ$  SR region (*blinded*). Bottom row: MC predictions for the scalar sum of b-tagged jet  $p_T$  (left) and the Number of b-tagged jets (right) in the  $tWZ$  SR region (*blinded*)

738 **6.3.3  $t\bar{t}Z$  CR**

739 In this section, the pre-fit distributions of variables in the  $t\bar{t}Z$  CR region are shown. More pre-fit distributions  
 740 for the  $t\bar{t}Z$  CR are shown in the appendix (Section A.3).

Not reviewed, for internal circulation only

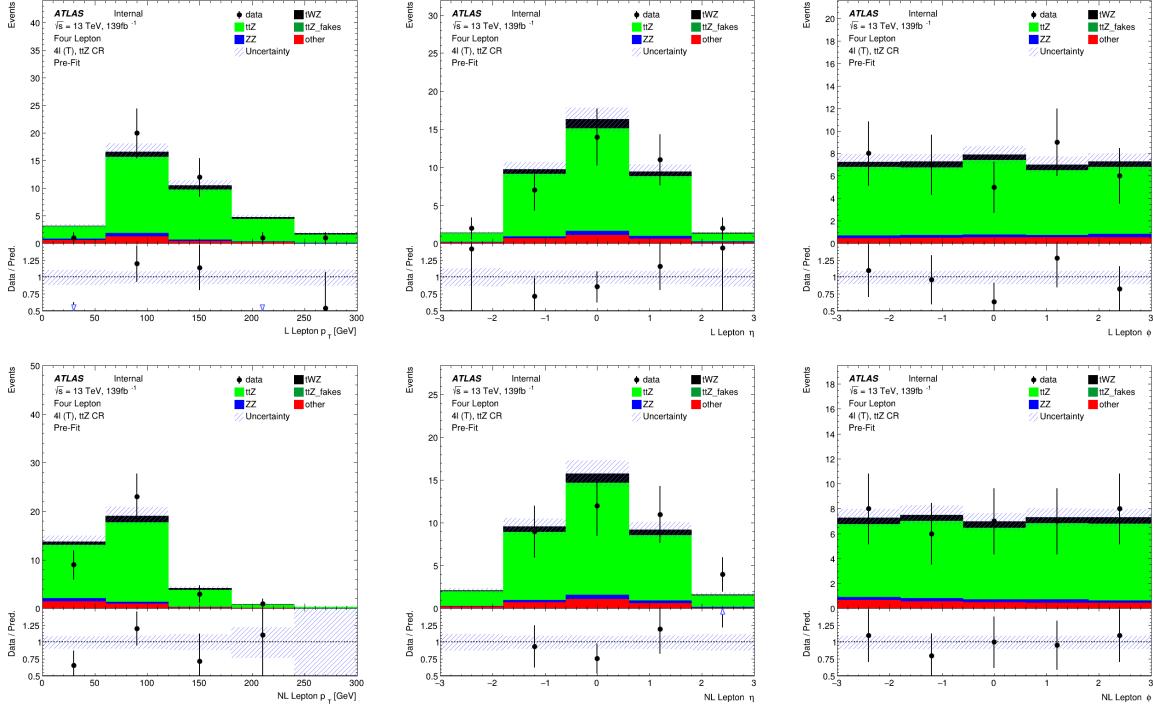


Figure 37: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  for leading (L) leptons (top row) and next-to-leading (NL) leptons (bottom row) in the  $t\bar{t}Z$  CR region

Not reviewed, for internal circulation only

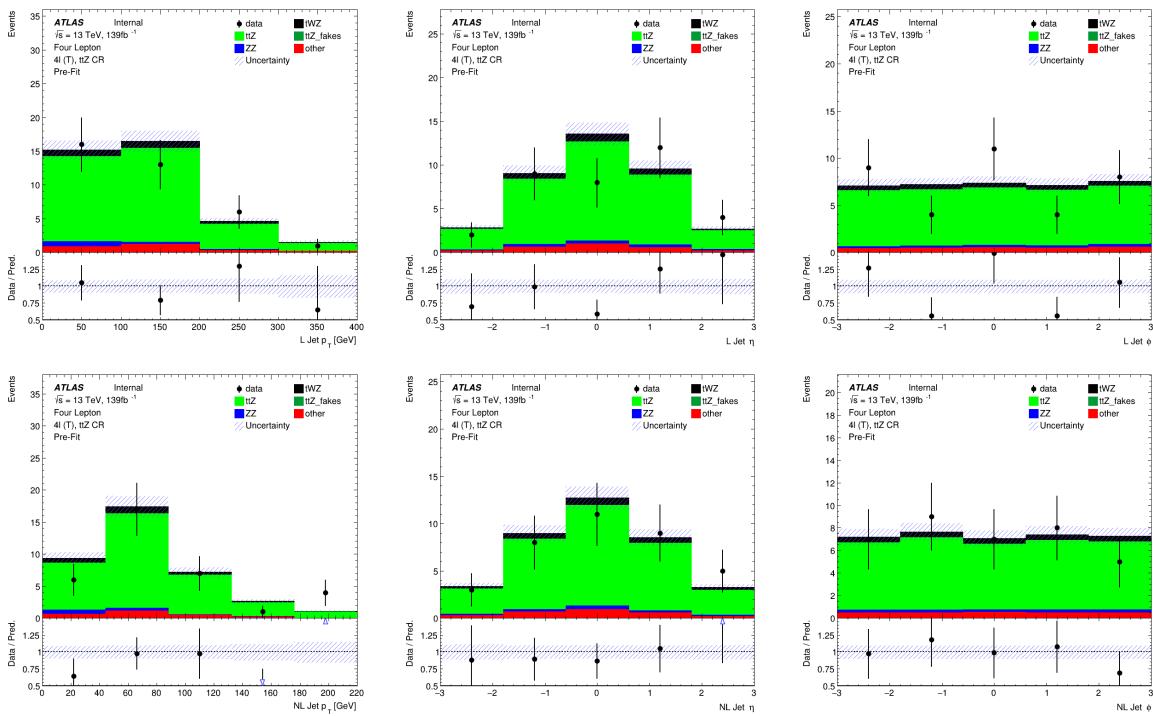


Figure 38: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  for leading (L) jets (top row) and next-to-leading (NL) jets (bottom row) in the  $t\bar{t}Z$  CR region

Not reviewed, for internal circulation only

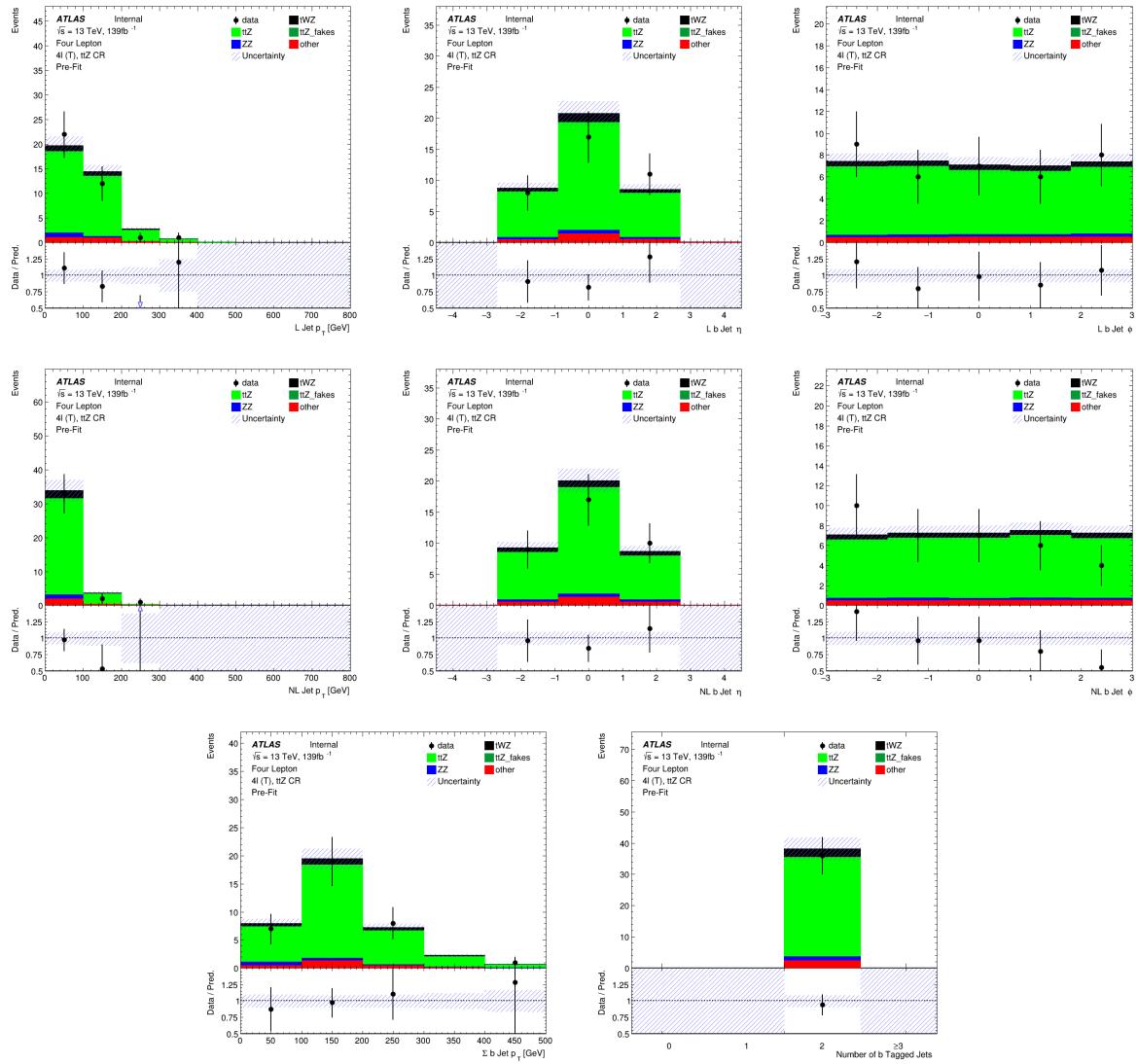


Figure 39: Top row: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  for leading b-tagged jets in the  $t\bar{t}Z$  CR region. Middle row: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  for next-to-leading b-tagged jets in the  $t\bar{t}Z$  CR region. Bottom row: MC predictions for the scalar sum of b-tagged jet  $p_T$  (left) and the Number of b-tagged jets (right) in the  $t\bar{t}Z$  CR region

741 **6.3.4 ZZ<sub>b</sub> CR**

742 In this section, the pre-fit distributions of variables in the ZZ<sub>b</sub> CR region are shown. More pre-fit  
 743 distributions for the ZZ<sub>b</sub> CR are shown in the appendix (Section A.4).

Not reviewed, for internal circulation only

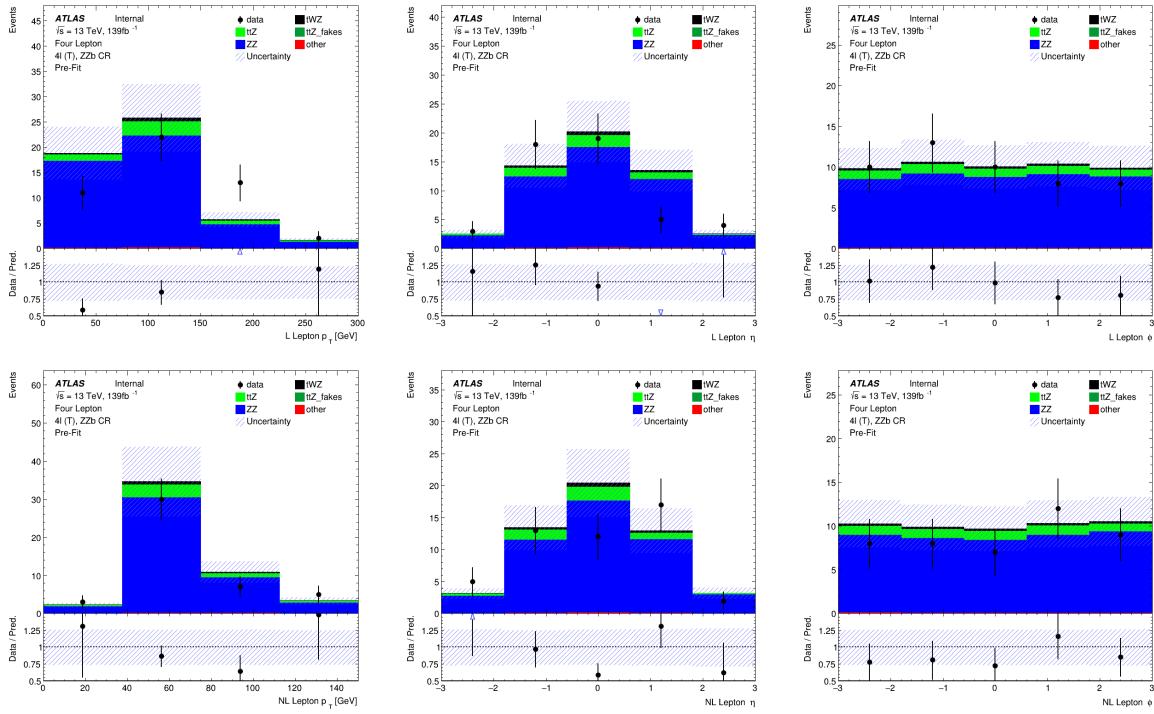


Figure 40: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  for leading (L) leptons (top row) and next-to-leading (NL) leptons (bottom row) in the ZZ<sub>b</sub> CR region

Not reviewed, for internal circulation only

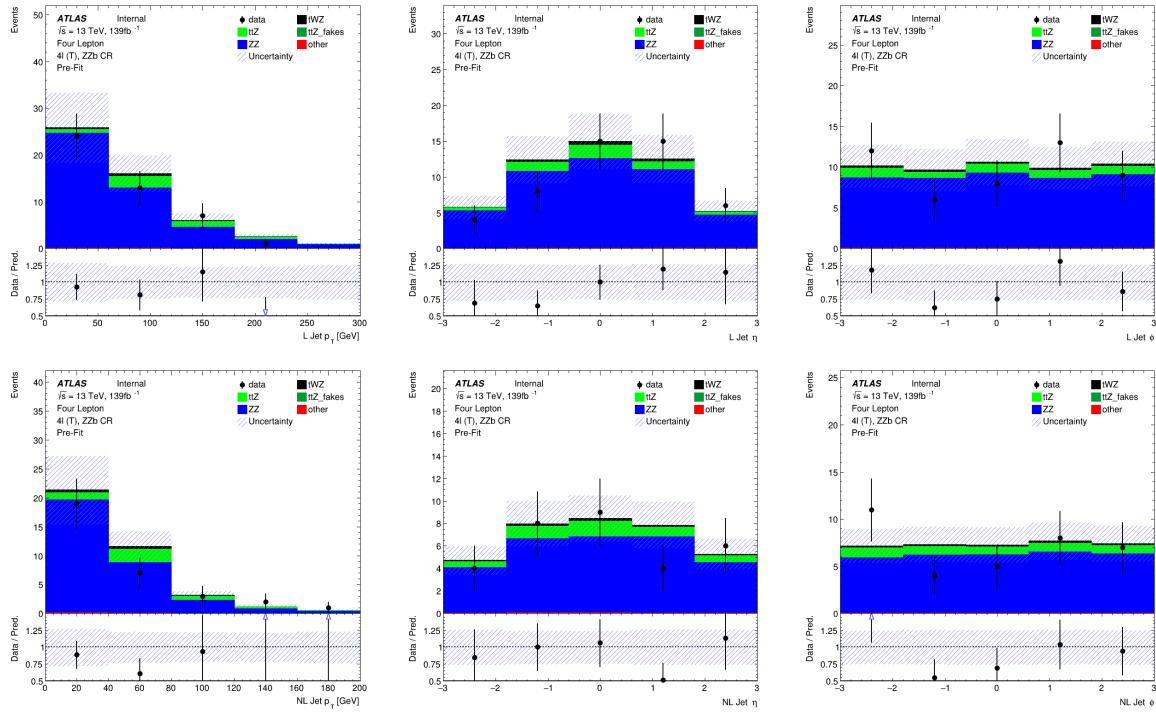


Figure 41: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  for leading (L) jets (top row) and next-to-leading (NL) jets (bottom row) in the  $ZZb$  CR region

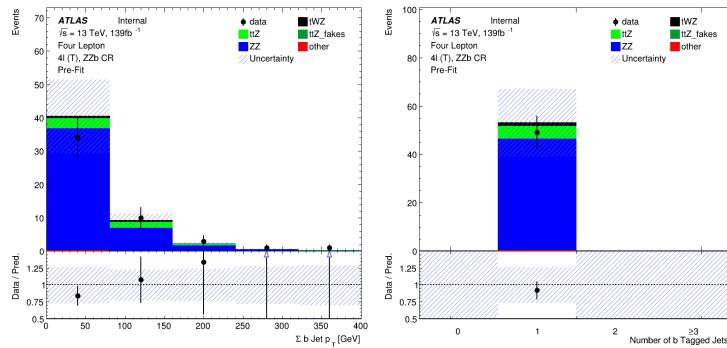


Figure 42: Top row: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  for b-tagged jets in the  $ZZb$  CR region . Bottom row: MC predictions for the scalar sum of b-tagged jet  $p_T$  (left) and the Number of b-tagged jets (right) in the  $ZZb$  CR region

744 **6.3.5 ( $tWZ$ )<sub>fake</sub> CR**

745 In this section, the pre-fit distributions of variables in the ( $tWZ$ )<sub>fake</sub> CR region are shown. More pre-fit  
 746 distributions for the ( $tWZ$ )<sub>fake</sub> CR are shown in the appendix (Section A.5).

Not reviewed, for internal circulation only

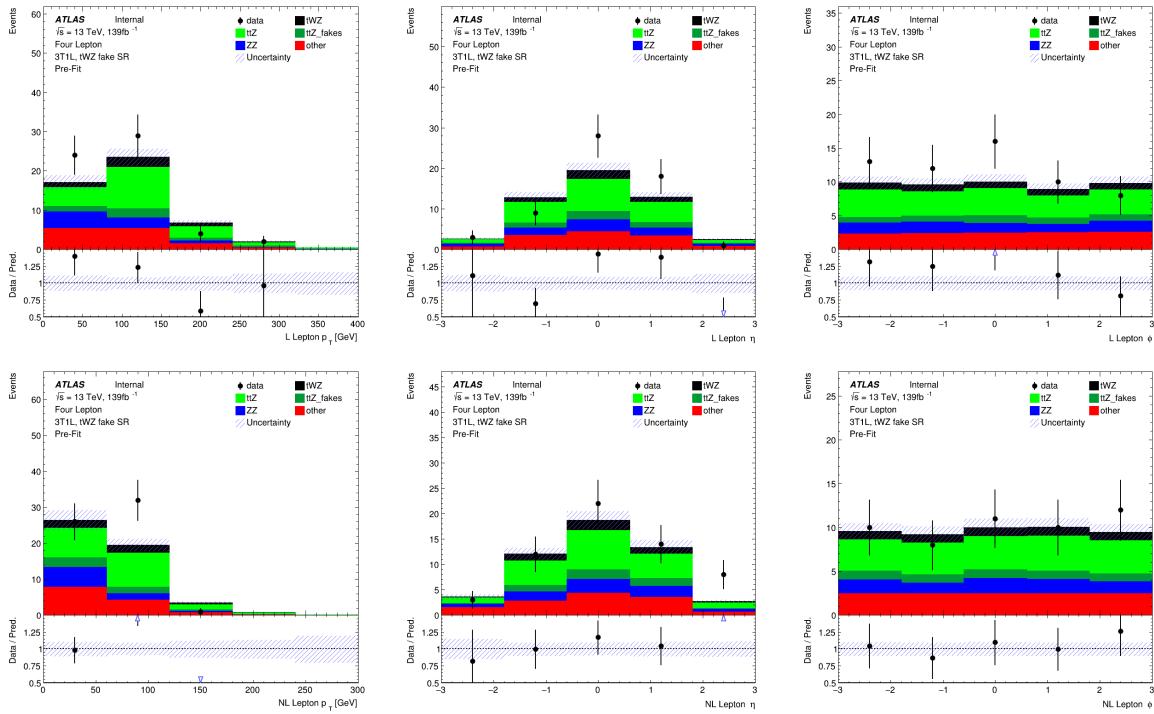


Figure 43: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  for leading (L) leptons (top row) and next-to-leading (NL) leptons (bottom row) in the ( $tWZ$ )<sub>fake</sub> CR region

Not reviewed, for internal circulation only

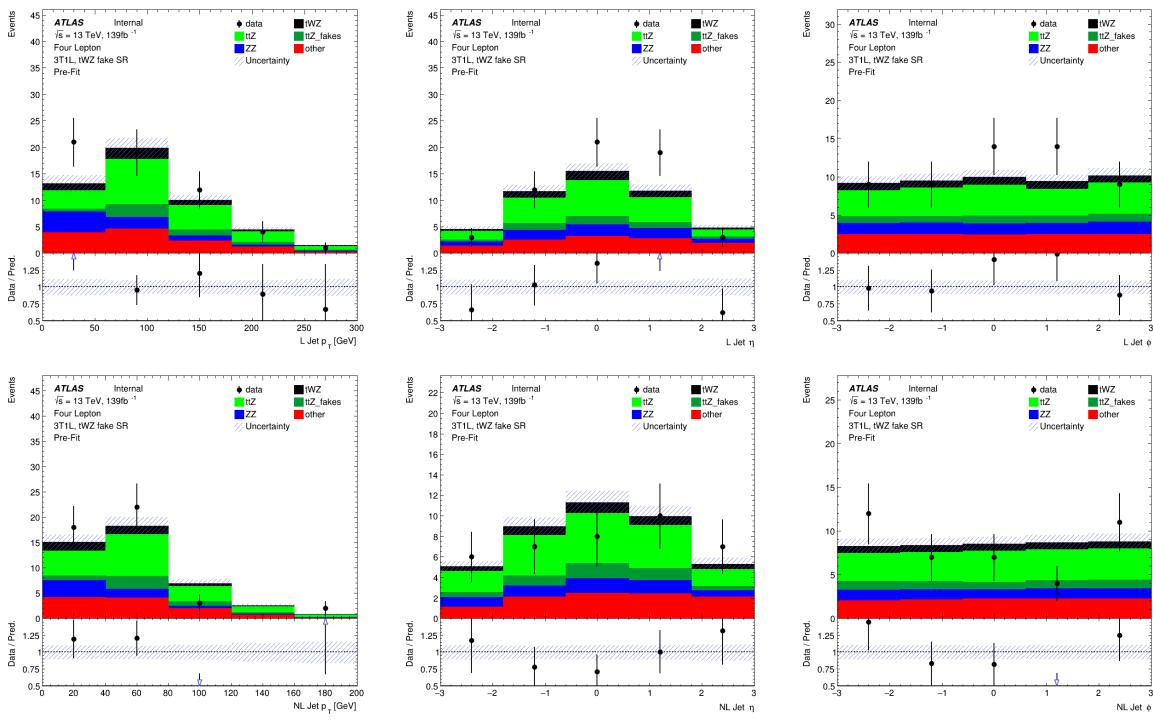


Figure 44: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  leading (L) jets (top row) and next-to-leading (NL) jets (bottom row) in the  $(tWZ)_{\text{fake}}$  CR region

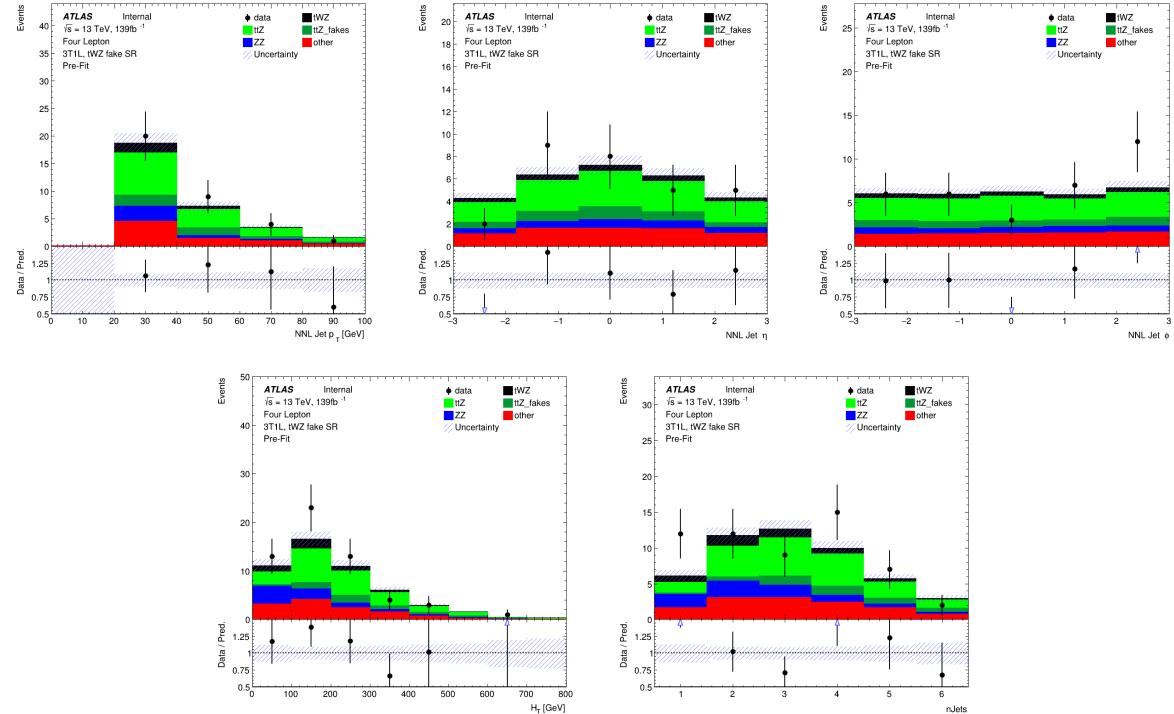


Figure 45: Top row: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  for next-to-next-to-leading (NNL) jets in the  $(tWZ)_{\text{fake}}$  CR region . Bottom row: MC predictions for  $H_T$  (scalar sum of Jet  $p_T$ ) (left) and the Number of jets (right) in the  $(tWZ)_{\text{fake}}$  CR region

Not reviewed, for internal circulation only

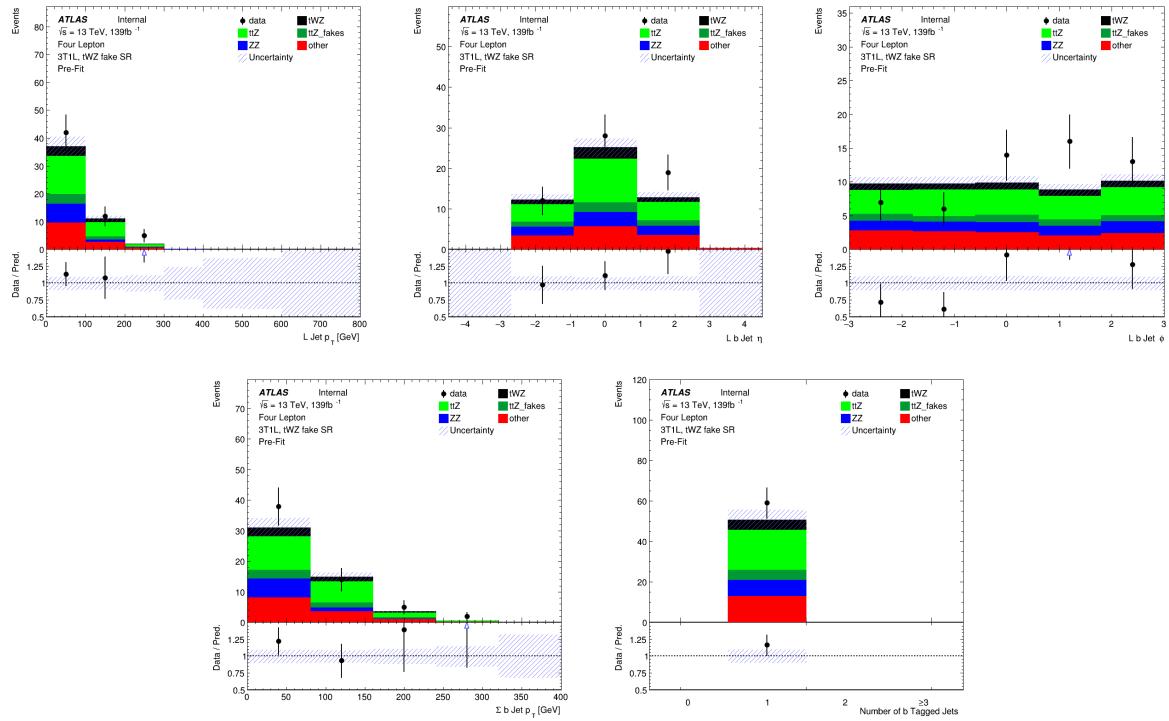


Figure 46: Top row: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  for leading b-tagged jets in the  $(tWZ)_{\text{fake}}$  CR region . Bottom row: MC predictions for the scalar sum of b-tagged jet  $p_T$  (left) and the Number of b-tagged jets (right) in the  $(tWZ)_{\text{fake}}$  CR region

747 **6.3.6 Fit Variables**

748 The variables used in the ML fit, in each region, are outlined in this section.

Not reviewed, for internal circulation only

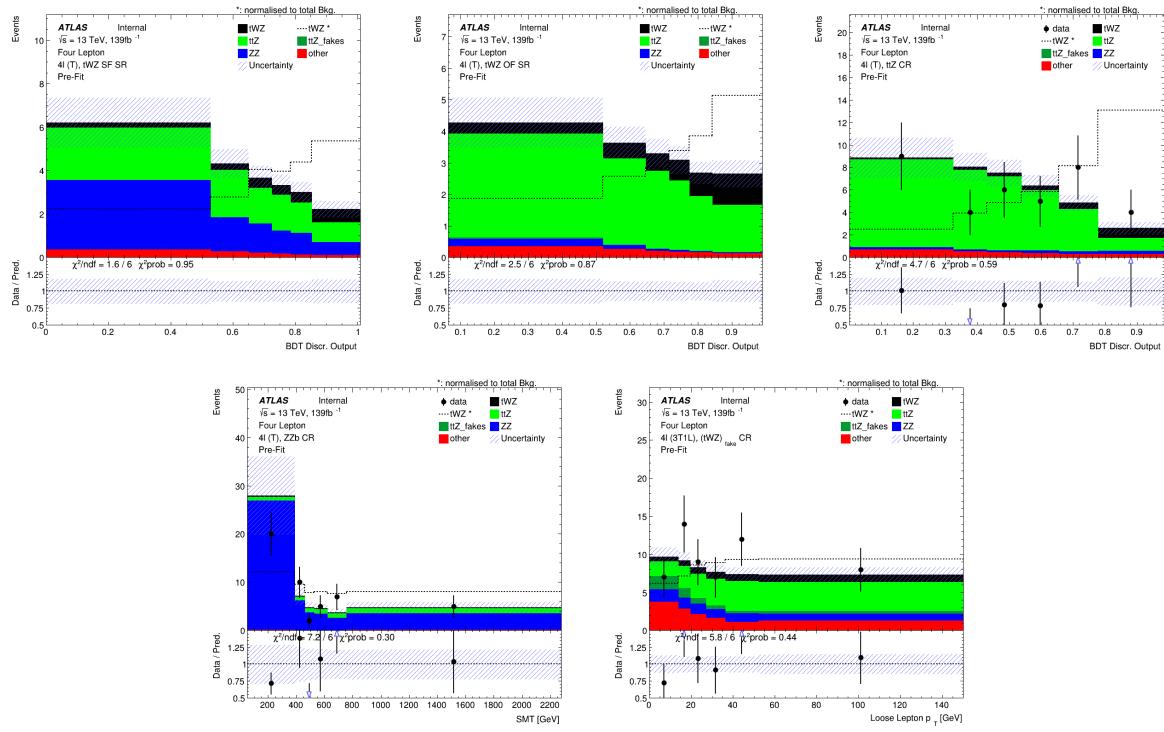


Figure 47: Pre-fit Plots of fit variables

Figure 48: Pre-fit Plots (blinded) of variables used in the fit. Top left: event-level BDT Disc. Output in the  $tWZ$  SF (4T) SR region. Middle Top: event-level BDT Disc. Output in the  $tWZ$  SF (4T) SR region. Top right: event-level BDT Disc. Output in the  $t\bar{t}Z$  CR region. Bottom left:  $SMT = \sum p_T(\ell) + \sum p_T(jet) + E_T^{\text{miss}}$  in the  $ZZb$  CR. Bottom right:  $p_T$ (loose lepton) in the  $(tWZ)_\text{fake}$  (3T1L) CR

749 Given the limited statistics which we are presented with in the tetralepton channel, we nevertheless observe  
 750 relatively good agreement overall between data and MC.

## 751 6.4 Fake Lepton Estimation (MC Template Method)

752 Fake leptons are objects reconstructed as leptons, but do not correspond to the leptons which we are  
 753 interested in our analysis. Fake leptons can be split up into two main categories, irreducible (prompt) fakes  
 754 and reducible (non-prompt) fakes. Irreducible fakes are true leptons which do not come from the process  
 755 of interest. Reducible fakes are objects which are mis-identified or incorrectly reconstructed as leptons. In  
 756 the ATLAS detector, the probability for a fake to occur is very low.

757

758

759 We aim to estimate the fake lepton contribution in this analysis. We start off by noticing that  $t\bar{t}Z$  is our most  
 760 dominant background ( $\sim 75\%$  of the total background contribution) and will therefore have the largest  
 761 fake component compared to all other samples considered in the analysis. The fake lepton efficiency,

762  $\epsilon$ , 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  
 763 selection (See Section ??) and  $N_{\text{fake}}^{\text{loose}}$  is the number of fake leptons which pass the loose lepton selection  
 764 (See Section ??). The probability of one fake lepton to occur,  $P(\text{one fake } \ell)$ , is proportional to  $\epsilon$  and  
 765 the probability for two fakes to occur is simply,  $P(\text{two fake } \ell) = (P(\text{one fake } \ell))^2 \propto \epsilon^2$ . Since  $\epsilon < 1$ , we  
 766 have  $P(\text{one fake } \ell) \ll P(\text{two fake } \ell)$ . For the purposes of this analysis we shall investigate the fake lepton  
 767 component to first order and therefore will only consider the case where one fake lepton occurs in a  $t\bar{t}Z$   
 768 event.

769

770

771 Firstly, we split up the dominant  $t\bar{t}Z$  background into  $t\bar{t}Z$  and  $(t\bar{t}Z)_{\text{fake}}$  components. Secondly, we define a  
 772  $(tWZ)_{\text{fake}}$  CR (See Section ??) which is enhanced in fakes and aims to constrain the  $(t\bar{t}Z)_{\text{fake}}$  background  
 773 in the SR.

774 All events which contribute to the  $(t\bar{t}Z)_{\text{fake}}$  background are determined by the IFF Truth Classifier [53].  
 775 The IFF Truth Classifier is a tool which aims to classify leptons based off their truth information.  
 776 The function returns one of the following lepton categories: Unknown, KnownUnknown, IsoElectron,  
 777 ChargeFlipIsoElectron, PromptMuon, PromptPhotonConversion, ElectronFromMuon, TauDecay,  
 778 BHadronDecay, CHadronDecay or LightFlavorDecay (More details [54]). Given these categories,  
 779 we consider leptons classified as BHadronDecay, CHadronDecay or LightFlavorDecay (i.e. a lepton  
 780 originating from the decay of a  $b$ -Hadron,  $c$ -Hadron or light-flavour jet) to be fakes. Since we only  
 781 consider events with one fake lepton to contribute to our  $(t\bar{t}Z)_{\text{fake}}$  background, we require that events which  
 782 contribute to this background are those where exactly one lepton from the  $t\bar{t}Z$  sample are classified by the  
 783 IFF Truth Classifier with one of the three aforementioned categories.

784 The  $(tWZ)_{\text{fake}}$  CR aims to be as similar as possible to the  $tWZ$  SRs, but enhanced in fakes. This CR can  
 785 then be used to constrain the normalisation of the  $(t\bar{t}Z)_{\text{fake}}$  template from data, since we use the *mixed data*  
 786 and *MC fit* (See Section ??) which uses the data in the CRs to construct a modified ASIMOV dataset from  
 787 the fitted nuisance parameters from a background only fit to the CRs. To ensure that this region is enhanced  
 788 in fakes, we require that it contains 3 tight leptons and 1 loose lepton, since loose leptons are more likely to  
 789 be fakes. By using the  $p_T$  of the loose lepton ( $p_T(\text{Loose Lepton})$ ) in this region as the variable used in the  
 790 fit, the shape (and normalisation) of the  $(t\bar{t}Z)_{\text{fake}}$  template can be constrained.

791 The proportion of fake events in a region can be calculated by taking the ratio of the  $(t\bar{t}Z)_{\text{fake}}$  MC yield in  
 792 the region and the total MC yield in the region. These are  $0.0035 \pm 0.0025$ ,  $0.0014 \pm 0.0012$  and  $0.10 \pm$

793 0.051 for the  $tWZ$  OF SR,  $tWZ$  SF SR and  $(tWZ)_{\text{fake}}$  CR respectively. This shows that there is significantly  
794 more  $(t\bar{t}Z)_{\text{fake}}$  events in the  $(tWZ)_{\text{fake}}$  CR compared to the  $tWZ$  SRs, therefore justifying our inclusion of  
795 the  $(tWZ)_{\text{fake}}$  CR in the fit (to constrain the  $(t\bar{t}Z)_{\text{fake}}$  background).

[Not reviewed, for internal circulation only]

## 796 6.5 Analysis Strategy

797 The dominant backgrounds in the  $tWZ$  SRs are  $t\bar{t}Z$  ( $\sim 75\%$  of the total background contribution) and  $ZZ$   
 798 ( $\sim 20\%$  of the total background contribution). We therefore aim to find variables which discriminate well  
 799 between  $tWZ$  and the dominant  $t\bar{t}Z$  and  $ZZ$  backgrounds.

800 We use two MVAs (an event-level Boosted Decision Tree (BDT) and an object-level BDT) and a kinematic  
 801 reconstruction algorithm which targets the discrimination between  $tWZ$  and  $t\bar{t}Z$  events.

### 802 6.5.1 Machine Learning Techniques

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

809 Two different BDTs were used, one which aims to discriminate between  $tWZ$  and  $t\bar{t}Z$  events and one which  
 810 aims to discriminate between  $\ell b$  systems which originate from the decay of a top quark ( $t \rightarrow W(\rightarrow \ell\nu)b$ )  
 811 and those which do not. We refer to these two BDTs as an **event-level** and an **object-level** classifier  
 812 respectively. The discriminator output from the object-level BDT can be converted to a variable which can  
 813 then be used as input to the event-level BDT.

815 **Object-level BDT** The object-level BDT was trained on a  $t\bar{t}$  sample, using the same jet and lepton  
 816 baseline selections as outlined in Table 10. Minor differences occur between this  $t\bar{t}$  sample and the MC  
 817 samples used for the rest of the analysis, such as the use of the older MV2c10  $b$ -tagger. This is such  
 818 since the  $t\bar{t}$  sample was used in a separate analysis and was re-purposed for this analysis. Despite these  
 819 differences, the trained object-level BDT is sufficient in this case, since we expect that  $\ell b$  systems present  
 820 in the  $t\bar{t}$  sample are similar enough to those present in the MC samples used in the rest of the analysis.  
 821 Additionally, we opted to use this disjoint  $t\bar{t}$  sample as to avoid resorting to use our MC samples used in the  
 822 rest of the analysis which is heavily limited on statistics, therefore maximizing the amount of MC statistics  
 823 used in the fitting procedure and the training of the event-level BDT.

824

825

826 The signal class is defined to consist of reconstructed  $\ell b$  systems (defined as the sum of the 4-vectors of the  
 827 lepton and  $b$ -jet) coming from top quarks which are well matched to their truth counterparts. In particular,  
 828 we require that  $\Delta R$  between the reconstructed and truth  $\ell b$  system is less than 0.05. We additionally require  
 829 that the reconstructed lepton and the truth top have charges with the same sign (since  $t \rightarrow b\ell^+\bar{\nu}_\ell$  and  
 830  $\bar{t} \rightarrow \bar{b}\ell^-\nu_\ell$ ). Conversely, the background class is defined to consist of reconstructed  $\ell b$  systems which are  
 831 not well matched to their truth counterparts. In particular, we require that  $\Delta R$  between the reconstructed  
 832 and truth  $\ell b$  system is greater than 0.05. We additionally require that the reconstructed lepton and the truth  
 833 top have charges with opposite sign. These definitions for the signal and background classes ensure that  
 834 the signal class consists of mostly  $\ell b$  systems originating from tops and the background class consists of  
 835 mostly  $\ell b$  systems which do not originate from a top decay.

836  
837

838 Different observables corresponding to an  $\ell b$  system were used as input to training. The optimum values  
 839 for the hyper-parameters used were determined by training the BDT with a range of different values for the  
 840 hyper-parameters and choosing the set of values which maximized the mean accuracy (based off 5 fold  
 841 kfold cross-validation). This method is more commonly referred to as hyper-parameter optimisation or  
 842 tuning. After hyperparameter optimisation, the mean accuracy of each fold increased from 0.76 to 0.77  
 843 ( $\sim 1\%$  increase). The observables and hyper-parameters used in training are summarised in Table 13 and  
 844 Table 14 respectively.

Observable	Description
$m(\ell b)$	Invariant mass of the $\ell b$ system
$p_T(\ell b)$	$p_T$ of the $\ell b$ system
$\Delta\eta(\ell, b)$	$\Delta\eta$ between the $\ell$ and $b$ -tagged jet
$\Delta\phi(\ell, b)$	$\Delta\phi$ between the $\ell$ and $b$ -tagged jet
$\Delta R(\ell, b)$	$\Delta R$ between the $\ell$ and $b$ -tagged jet

Table 13: A list of the observables used in the object-level BDT, ordered by importance (descending, top to bottom).

Hyper-parameter	Value	Description
loss	deviance	The loss function to be optimised
criterion	friedman_mse	The function used to measure the quality of a split
n_estimators	100	The number of boosting stages to perform
learning_rate	0.1	The step size at each iteration during optimisation
max_depth	4	The maximum depth of the individual regression estimators
min_samples_split	2	The minimum number of samples (events) required to split an internal node
min_samples_leaf	1	The minimum number of samples (events) required to be at a leaf node
validation_fraction	0.1	The proportion of training data to set aside as validation set for early stopping
n_iter_no_change	20	Training terminates when the validation score (determined by the validation set) does not improve in all of the previous n_iter_no_change number of iterations

Table 14: A list of the hyper-parameters used in the object-level BDT. Hyperparameters not listed in this table use the default values as stated in the Scikit-learn Documentation[55].

845 The number of events used in training for the signal and background classes were 347952 and 266636  
 846 respectively. Imbalanced datasets can cause ML classifiers to ignore small classes while concentrating on  
 847 classifying large classes more accurately, which may result in the trained classifier performing sub-optimally.  
 848 In order to correct this dataset imbalance, we ensure that the relative weighting of each event is such that the  
 849 sum of the signal weights is equal to the sum of the background weights. The resulting BDT discriminator  
 850 plots are shown in Figure 49.

851 We can see that the shapes of the training and test sets for both signal and background are very similar.  
 852 This is a good indicator that no over-training occurred. Another over-training check is performed using 5  
 853 fold kfold cross validation. We ensure that the variance of the mean accuracy of each folds' test set in cross  
 854 validation is substantially small. This indicates that fluctuations in features from different training sets are  
 855 not learnt by the classifier. For the object-level classifier, a variance of  $3.24 \times 10^{-7}$  was calculated for the  
 856 mean accuracies of each folds' test set in cross validation, providing further evidence that no overtraining  
 857 occurred.

858  
859

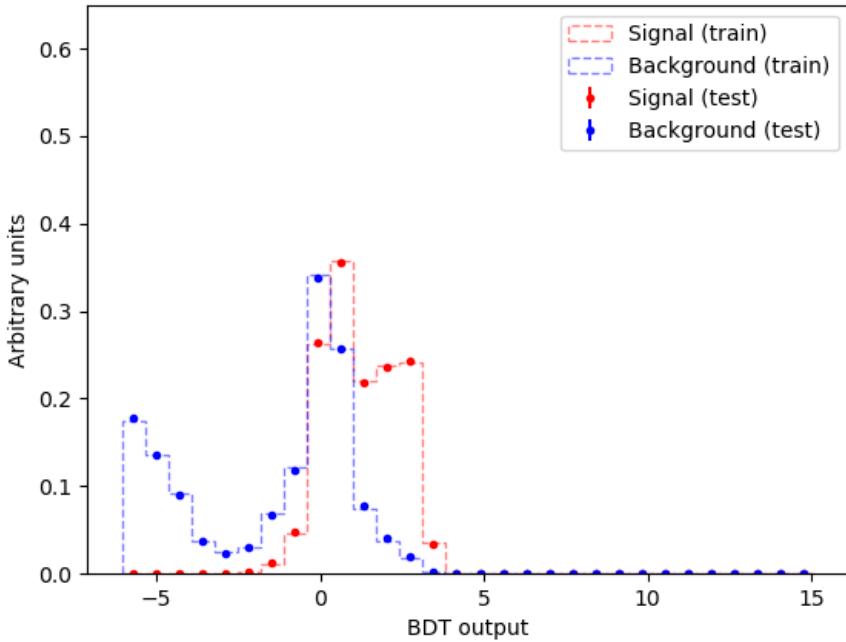


Figure 49: Normalised histograms of the object-level BDT discriminator output from the signal and background classes for the training and test sets from the 5th fold in a 5 fold kfold cross validation.

The output from the object-level BDT was used to construct a variable to be used as input to the event-level BDT. The event-level BDT aims to discriminate between  $tWZ$  and our most prominent background,  $t\bar{t}Z$ . We therefore aim to construct a variable from the output of the object-level BDT which discriminates well between  $tWZ$  and  $t\bar{t}Z$ . Since  $tWZ$  events contain one top quark and  $t\bar{t}Z$  events contain two top quarks, we expect that  $tWZ$  events have one  $\ell b$  combination which scores well and we expect that  $t\bar{t}Z$  events have two  $\ell b$  combinations which score well. We construct a variable,  $BDTScore(\frac{\text{Best}}{\text{2nd Best}})$ , which takes the ratio of the scores of the top scoring  $\ell b$  system to the 2nd best scoring  $\ell b$  system. We expect this variable to be large for  $tWZ$  events and closer to one for  $t\bar{t}Z$  events, therefore providing discrimination between  $tWZ$  and  $t\bar{t}Z$ .

**Event-level BDT** The event-level BDT was trained on 50% of the  $tWZ$  MC sample's events for the signal class and similarly, 50% of the  $t\bar{t}Z$  and  $ZZ$  MC sample's events were used for the background class. The samples we train on are individual events, with the features being carefully chosen observables. These observables are chosen on the basis that they are somewhat uncorrelated from one another and show a relatively large amount of separation power between  $tWZ$  and  $t\bar{t}Z$ . Similarly to the object-level BDT, the optimum values for the hyper-parameters used were determined via hyper-parameter optimisation. After hyperparameter optimisation, the mean accuracy of each fold (determined from 5 fold kfold cross validation) increased from 0.72 to 0.74 ( $\sim 3\%$  increase). The observables and hyper-parameters used in training are summarised in Table 15 and Table 16 respectively.

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

Observable	Description
$2\nu\text{SM}$	Maximum weight from the $2\nu\text{SM}$ algorithm
$HT$	Scalar sum of jet $p_T$
$LT$	Scalar sum of lepton $p_T$
$\sum p_T(b - \text{jet})$	Scalar sum of $b$ -tagged jet $p_T$
$\text{BDTScore}(\frac{\text{Best}}{\text{2nd Best}})$	Ratio of the top scoring $\ell b$ system to the 2nd best scoring $\ell b$ system from the output of the object-level BDT ( $\ell b$ classifier)
$\Delta\eta(\ell_{1,\text{Non-Z}}, \ell_{2,\text{Non-Z}})$	$\Delta\eta$ between the two leptons, not coming from a $Z$ candidate
$\min(m(\ell b))$	Mass of the $\ell b$ system with the smallest mass
$\Delta\phi(\ell_{1,\text{Non-Z}}, \ell_{2,\text{Non-Z}}, b_1)$	$\Delta\phi$ between the non- $Z$ lepton system (leptons not originating from a $Z$ -candidate) and the leading $b$ -tagged jet

Table 15: A list of the observables used in the event-level BDT, ordered by importance (descending, top to bottom).

Hyper-parameter	Value	Description
loss	deviance	The loss function to be optimised
criterion	friedman_mse	The function used to measure the quality of a split
n_estimators	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 n_iter_no_change number of iterations

Table 16: A list of the hyper-parameters used in the event-level BDT. Hyperparameters not listed in this table use the default values as stated in the Scikit-learn Documentation[55].

880 where we aim to use the BDT discriminator ( $tWZ$  SRs and  $t\bar{t}Z$  CR). This is done by applying normalization  
 881 weights to each event, defined as,

$$W = \frac{\sigma \mathcal{L} \text{weight(MC)}}{\text{totalWeight(MC)}} \quad (3)$$

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

885 The number of events used in training for the signal and background classes were 31584 and 14010  
 886 respectively. Similarly to the object-level BDT, there is a dataset imbalance. We correct this imbalance (in  
 887 the same way as before with the object-level BDT) by ensuring that the relative weighting of each event is  
 888 such that the sum of the signal weights is equal to the sum of the background weights.

889 The resulting BDT discriminator plots are shown in Figure 50.

890 We can see that the shapes of the training and test sets for both signal and background are very similar.  
 891 This is a good indicator that no over-training occurred. As with the object-level bdt, we perform another  
 892 overtraining check, by ensuring that the variance of the mean accuracy of each folds' test set in a 5 fold kfold  
 893 cross validation is substantially small. This indicates that fluctuations in features from different training sets  
 894 are not learnt by the classifier. For the event-level classifier, a variance of 0.00026 was calculated for the  
 895 mean accuracies of each folds' test set in cross validation, providing further evidence that no overtraining  
 896 occurred.

### 897 6.5.2 Two Neutrino Scanning Method ( $2\nu\text{SM}$ ) Algorithm

898 The Two Neutrino Scanning Method ( $2\nu\text{SM}$ ) algorithm<sup>4</sup> [56, 57] aims to reconstruct  $t\bar{t}$  systems in the  
 899  $2\ell$ ,  $3\ell$  and  $4\ell$  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  $t\bar{t}$

<sup>4</sup> software tool and weights provided by Thomas McCarthy ( $t\bar{t}Z$  analysis group - Max Planck Institute)

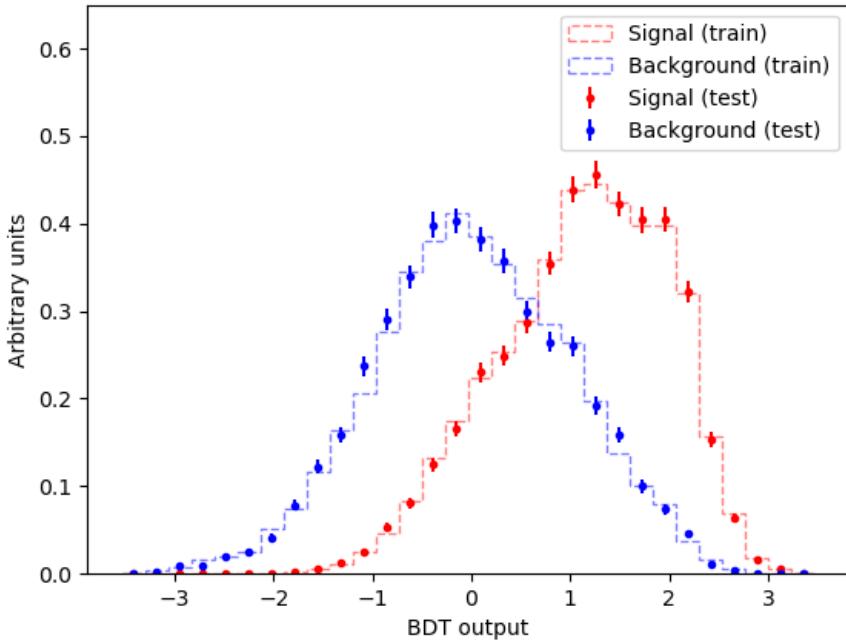


Figure 50: Normalised histograms of the event-level BDT discriminator output from the signal and background classes for the training and test sets from the 5th fold in a 5 fold kfold cross validation.

background in the  $t\bar{t}Z$  analysis. We can re-purpose this algorithm to distinguish between  $tWZ$  and  $t\bar{t}Z$  by removing the easily-identifiable  $Z$  boson.

The  $2\nu$ SM algorithm reconstructs a  $t\bar{t}$  system by scanning through the components of two possible neutrino 4-vectors ( $\nu_1$  and  $\nu_2$ ). It then aims to determine which  $\nu_1$  and  $\nu_2$  correspond to the two neutrinos which originate 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  $t\bar{t}$  dilipoton final state hypothesis. We are able to use this algorithm in our analysis to discriminate between  $tWZ$  and  $t\bar{t}Z$  since we can easily reconstruct the OSSF leptons which decay from the  $Z$  boson and remove it before inputting the event into the algorithm. We would then expect that the  $2\nu$ SM algorithm returns a higher score from a  $t\bar{t}Z$  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. it does not look like a  $t\bar{t}$  event after removal of the  $Z$  boson).

## 912 The algorithm

The  $2\nu$ SM algorithm starts off by writing down four equations which correspond to the invariant masses of the top quark ( $m(t)$ ) and  $W$  boson ( $m(W)$ ) for the two top decays (i.e.  $t \rightarrow W^+ b \rightarrow \ell^+ \nu_\ell$ ) in a dileptonic  $t\bar{t}$  event. These can be written as,

$$(\ell_1 + \nu_1)^2 = m(W)^2 = (80.385 \text{ GeV})^2 \quad (4)$$

$$(\ell_1 + \nu_1 + b_{1,2})^2 = m(t)^2 = (172.5 \text{ GeV})^2 \quad (5)$$

$$(\ell_2 + \nu_2)^2 = m(W)^2 = (80.385 \text{ GeV})^2 \quad (6)$$

$$(\ell_2 + \nu_2 + b_{2,1})^2 = m(t)^2 = (172.5 \text{ GeV})^2 \quad (7)$$

where the subscripts indicate that these particles originate from the decay of two different top quarks in a  $t\bar{t}$  system. We assume 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).

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

$$p_{T_\nu} = \frac{\frac{1}{2}(m(W)^2 - m(\ell)^2)}{E_\ell \cosh \eta_\nu - p_{\ell,z} \sinh \eta_\nu - p_{\ell,x} \cos \phi_\nu - p_{\ell,y} \sin \phi_\nu} \quad (8)$$

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. At this stage, we have possible 4-vectors for  $\nu_1$  and  $\nu_2$ . Using these possible neutrino 4-vectors, we reconstruct the two possible  $t\bar{t}$  systems,

$$t_1 = \ell_1 + b_1 + \nu_1 \text{ and } t_2 = \ell_2 + b_2 + \nu_2 \quad (9)$$

**OR**

$$t_1 = \ell_1 + b_2 + \nu_1 \text{ and } t_2 = \ell_2 + b_1 + \nu_2 \quad (10)$$

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 from 0 to 1) is defined as a product of four probabilities (described below) and can be written as,

$$w_{2\nu SM} = P_{m_{t_1}} \times P_{m_{t_2}} \times P_{\Delta E_x} \times P_{\Delta E_y} \quad (11)$$

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

### Calculating $w_{2\nu SM}$

We use distributions of well modelled observables ( $m_{b\ell\nu}$  and  $\Delta E_x$ ) from simulated  $t\bar{t}$  events in order to determine how well our reconstructed neutrinos (and in turn top quarks) resemble neutrinos (and top quarks) present in a  $t\bar{t}$  event.

Not reviewed, for internal circulation only

**938  $P_{m_{t_1}}$  and  $P_{m_{t_2}}$**  A normalised distribution of the mass of reconstructed top quarks ( $m_{b\ell\nu}$ ) from a  $t\bar{t}$   
**939** sample is generated to determine the probabilities  $P_{m_{t_1}}$  and  $P_{m_{t_2}}$ . The distribution is generated from  
**940** reco-level leptons, generator-level neutrinos and reoc-level jets matched in  $\Delta R$  to generator-level  $b$ -quarks,  
**941** therefore only filling the distribution with correct detector-level objects. We then use the distribution to  
**942** interpolate our two reconstructed top quarks, which returns a weight value from 0 to 1, with higher values  
**943** corresponding to a reconstructed top quark which has a mass close to that of a top quark from a  $t\bar{t}$  system.  
**944** This interpolation is done for both reconstructed tops,  $t_1$  and  $t_2$ , corresponding to probabilities  $P_{m_{t_1}}$  and  
**945**  $P_{m_{t_2}}$ . The distribution used is shown in Figure 51.

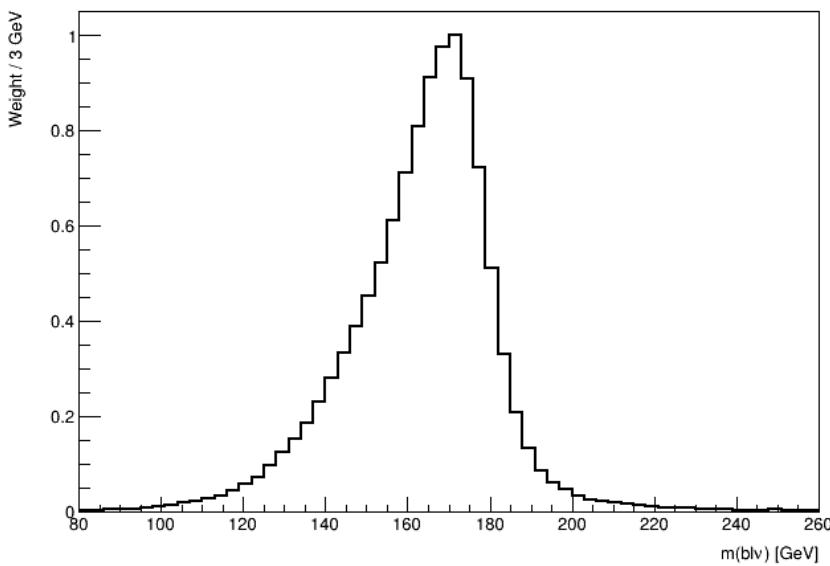


Figure 51:  $m_{b\ell\nu}$  distribution based on simulated  $t\bar{t}$  events, used to calculate  $P_{m_{t_1}}$  and  $P_{m_{t_2}}$

**946  $P_{\Delta E_x}$  and  $P_{\Delta E_y}$**  A similar method is used to determine  $P_{\Delta E_x}$  and  $P_{\Delta E_y}$ . In this case we generate a  
**947** weight distribution of  $\Delta E_x = (p_{T,\nu_1})_x + (p_{T,\nu_2})_x - (E_T^{\text{miss}})_x$  based off simulated  $t\bar{t}$  events. In particular,  
**948** this distribution is generated using reco-level  $E_T^{\text{miss}}$  and generator-level neutrinos. The use of this  
**949** distribution lies under the assumption that neutrinos are the dominant source of  $E_T^{\text{miss}}$ , and therefore,  
**950**  $(E_T^{\text{miss}})_x \approx (p_{T,\nu_1})_x + (p_{T,\nu_2})_x$  and  $(E_T^{\text{miss}})_y \approx (p_{T,\nu_1})_y + (p_{T,\nu_2})_y$ . We then use the distribution to interpolate  
**951** the value of  $\Delta E_x$  and  $\Delta E_y$  from our reconstructed neutrinos. This returns a weight value from 0 to 1, with  
**952** higher values corresponding to  $\Delta E_x$  and  $\Delta E_y$  (and in turn our reconstructed neutrino's  $p_T$ ) closer to those  
**953** observed in a  $t\bar{t}$  event. We expect the  $\Delta E_x$  and  $\Delta E_y$  distributions to have the same shapes, therefore we  
**954** only need to generate one (we have chosen  $\Delta E_x$ ). The distribution used is shown Figure 52.

## 955 Kinematic Veto

**956** The  $2\nu\text{SM}$  algorithm is extremely computationally intensive. The computation time depends on the number  
**957** step size of the  $\phi$  and  $\eta$  ranges which we scan over to reconstruct the neutrinos. For example, consider the  
**958** step 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  
**959** values for  $\phi$ . There will be  $(32)(32)(25)(25) = 640\,000$  possible pairs of neutrinos ( $\nu_1$  and  $\nu_2$ ) to consider

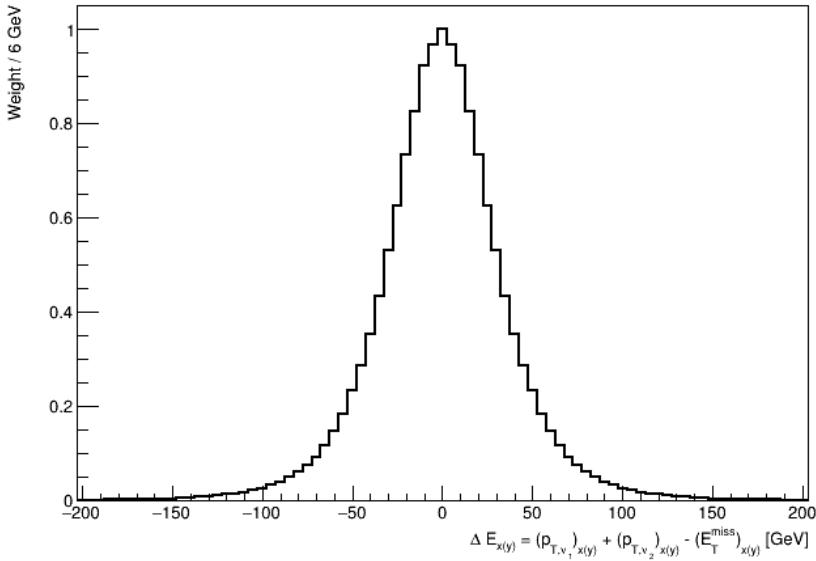


Figure 52:  $\Delta E_x$  distribution based on simulated  $t\bar{t}$  events, used to calculate  $P_{\Delta E_x}$  and  $P_{\Delta E_y}$

**per event.** Since we have to consider two possible  $t\bar{t}$  systems (See Equations 9 and 10), this number effectively increases to  $(2)(640000) = 128\ 000\ 0$  iterations **per event**. In order to reduce the number of  $t\bar{t}$  systems we need to consider, therefore decreasing computation time, we look at distributions of well modelled observables from  $t\bar{t}$  events and veto (discard) a possible reconstructed  $t\bar{t}$  system if the observable in question is improbable or unlikely to be observed in a  $t\bar{t}$  event. To achieve this, we define a threshold range for these observables, and if the possible reconstructed  $t\bar{t}$  system's corresponding value for this observable lies outside this range, it is vetoed and the algorithm continues with the next iteration.

**$\Delta\langle m(\ell b)\rangle$**  The first observable which we consider 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,

$$(\ell b)_1 = \ell_1 + b_1 \text{ and } (\ell b)_2 = \ell_2 + b_2 \quad (12)$$

**OR**

$$(\ell b)_1 = \ell_1 + b_2 \text{ and } (\ell b)_2 = \ell_2 + b_1 \quad (13)$$

$$(14)$$

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

$$\Delta\langle m(\ell b)\rangle = \frac{1}{2} |[(m(\ell_1 b_1) + m(\ell_1 b_1)) - (m(\ell_1 b_2) + m(\ell_2 b_1))]| \quad (15)$$

The idea here is that, if  $\Delta\langle m(\ell b)\rangle$  is large, it's more likely that we can simply select the  $\ell b$  combination with the smaller (minimum) average mass. To illustrate this, we look at the distribution (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 ( $\eta(b_1) \times \eta(b_2) < 0$ ).

Not reviewed, for internal circulation only

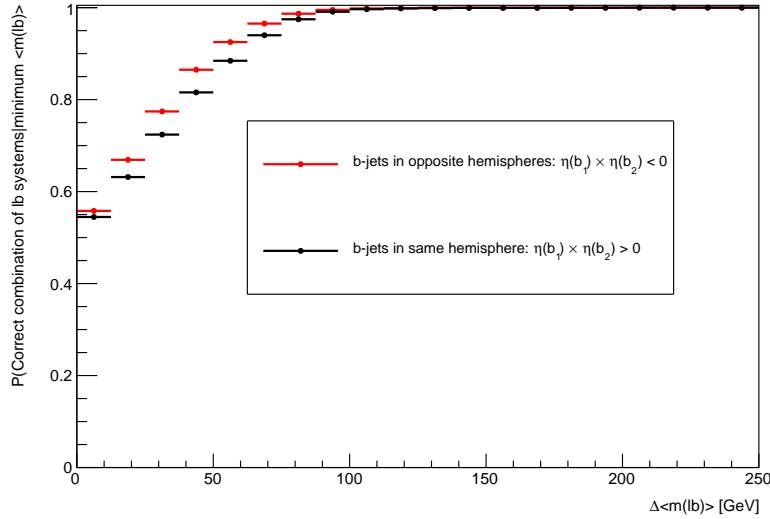


Figure 53:  $P(\text{Correct combination of } \ell b \text{ systems} | \text{minimum}\langle m(\ell b) \rangle)$  vs  $\Delta\langle m(\ell b) \rangle$ , for  $b$ -tagged jets in the same and opposite hemispheres, constructed from  $t\bar{t}$  events

From the Figure above, 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 we are considering the  $\ell b$  system with minimum average mass is an increasing function which plateaus to 1 at  $\sim 90$  GeV. We use these two distributions to interpolate the  $P(\text{Correct combination of } \ell b \text{ systems} | \text{minimum}\langle m(\ell b) \rangle)$  from  $\Delta\langle m(\ell b) \rangle$ . We require 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.

**982  $\eta(b\bar{b}\ell\ell)$**  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.

983

984

985 In the same way as for  $\Delta\langle m(\ell b) \rangle$ , we generate a distribution to determine values  $\eta(\nu)$  which are improbable  
986 for a  $t\bar{t}$  event. In this case, we generate a 2D histogram from simulated  $t\bar{t}$  events (dileptonic final state)  
987 at generator-level of  $\eta(\nu)$  vs  $\eta(b\bar{b}\ell\ell)$ . A heatmap showing the number of entries for these observables is  
988 shown in Figure 54 .

989 Using the above heatmap, we define a veto region (where a  $t\bar{t}$  event is extremely unlikely to occur) based  
990 off double-sided 95% limits (\*\*something here on confidence limit??\*\*). We apply a veto if either possible  
991 neutrino lies within this region. The veto region is shown in Figure 55 .

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

994

995 Again, we generate a distribution to determine (and veto) improbable possible neutrinos in simulated

Not reviewed, for internal circulation only

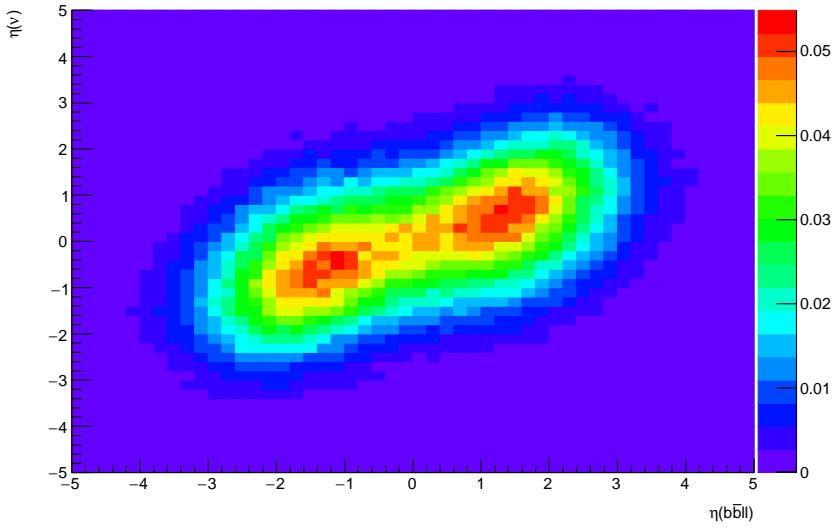


Figure 54: Heatmap of occupancy for  $\eta(\nu)$  vs  $\eta(b\bar{b}\ell\ell)$  produced from simulated  $t\bar{t}$  events (dileptonic final state) at generator-level

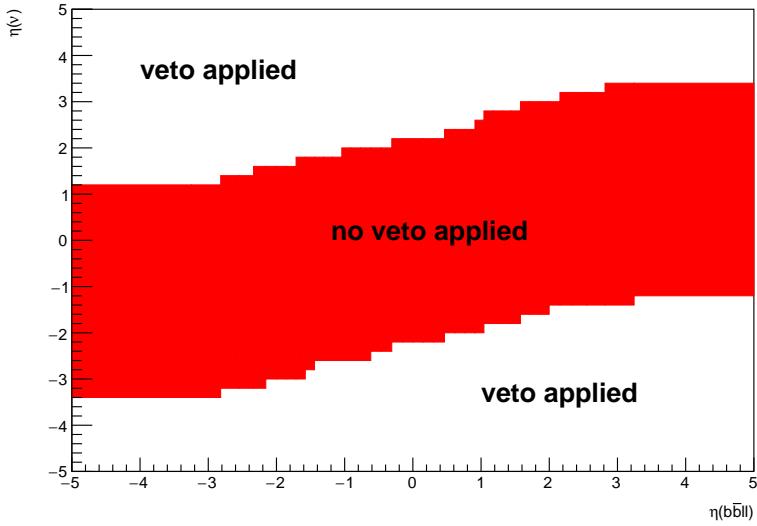


Figure 55: Plot showing the regions where vetoes are applied for the  $\eta(b_1 b_2 \ell_1 \ell_2)$  constraint

996  $t\bar{t}$  events (dilepton final state). A 2D histogram of  $\Delta R(\ell, \nu)$  vs  $L_T$  is generated from these events (at  
997 generator-level). A heatmap showing the number of entries for these observables is shown in Figure 56.

998 Using the same method as described in Section 6.5.2, we define a veto region where a veto is applied if  
999 either possible neutrino lies within this region. The veto region is shown in Figure 57.

Not reviewed, for internal circulation only

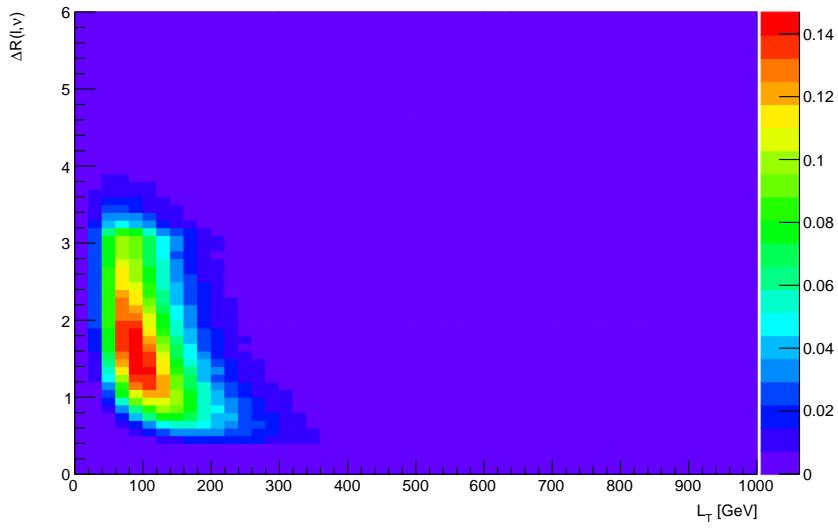


Figure 56: Heatmap of occupancy for  $\Delta R(\ell, \nu)$  vs  $L_T$  produced from simulated  $t\bar{t}$  events (dileptonic final state) at generator-level

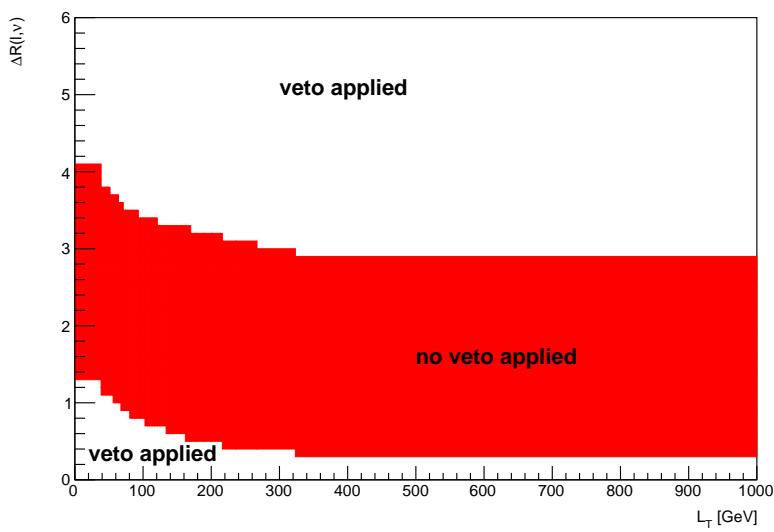


Figure 57: Plot showing the regions where vetoes are applied for the  $L_T$  constraint

## 1000 7 Systematics

1001 The systematic uncertainties can be separated into experimental (detector) systematics, which are related to  
 1002 the reconstruction of physics objects in the detector and theoretical uncertainties related to the modelling  
 1003 of the different processes background.

### 1004 7.1 Experimental uncertainties

1005 The experimental systematics (listed below) are used in the trilepton and tetralepton channels.

- 1006 • **Luminosity:**

1007 The 2015–2018 luminosity estimate of  $139 \text{ fb}^{-1}$  has a relative uncertainty of 3%. This uncertainty is  
 1008 obtained using the LUCID-2 detector [8] for the primary luminosity measurements. This systematic  
 1009 uncertainty affects all processes modelled using MC simulations.

- 1010 • **Pile-up reweighting:**

1011 An uncertainty related to the SFs used for MC to account for differences in pile-up distributions  
 1012 between MC and data is applied. This uncertainty is obtained by re-scaling the  $\langle \mu \rangle$  value in data by  
 1013 1.00 and 1/1.18 corrections are only applied to MC.

- 1014 • **Jet vertex tagger:**

1015 Uncertainties associated to the JVT are applied via the `JetJvtEfficiency` package [42] which  
 1016 account for the residual contamination from pile-up jets after pile-up suppression and the MC  
 1017 generator choice [58].

- 1018 • **Heavy- and light-flavor tagging:**

1019 The efficiency of the flavour-tagging algorithm is measured for each jet flavour using control samples  
 1020 in data and in simulation. From these measurements, correction factors are derived to correct  
 1021 the tagging rates in the simulation. In the case of  $b$ -tagged jets, the correction factors and their  
 1022 uncertainties are estimated from data using dileptonic  $t\bar{t}$  events [44, 59]. In the case of  $c$ -jets, they  
 1023 are derived from jets arising from  $W$  boson decays in  $t\bar{t}$  events [60]. In the case of light-flavour jets,  
 1024 the correction factors are derived using dijet events [61]. Sources of uncertainty affecting the  $b$ -  
 1025 and  $c$ -tagging efficiencies are evaluated as a function of jet  $p_T$ , including bin-to-bin correlations.  
 1026 The uncertainties in the efficiency for tagging light-flavour jets depend on the jet  $p_T$  and on  $\eta$ .  
 1027 An additional uncertainty is assigned to account for the extrapolation of the  $b$ -tagging efficiency  
 1028 measurement from the  $p_T$  region used to determine the correction factors to regions with higher  $p_T$ .

- 1029 • **Electron efficiency:**

1030 Uncertainties associated with the electron efficiency SFs are provided by the egamma CP group [33]  
 1031 and arise from the reconstruction, ID, isolation and trigger efficiencies. They correct for the efficiency  
 1032 difference between data and MC [30] and are measured with a “tag-and-probe” method in  $Z \rightarrow e^+e^-$   
 1033 and  $J/\psi \rightarrow e^+e^-$  events. The information on the correlation of the different components of the  
 1034 systematic uncertainties are provided for all efficiency measurements. The default correlation model  
 1035 for the uncertainties is used, which provides one up/down variation for each of the SF components  
 1036 separately [33, 62].

1037 • **Muon efficiency:**

1038 As for electrons, SFs obtained from  $Z \rightarrow \mu^+ \mu^-$  and  $J/\psi \rightarrow \mu^+ \mu^-$  events are applied to correct for the  
 1039 differences between data and MC in the muon ID, isolation and trigger efficiencies [36]. Uncertainties  
 1040 on these SFs are provided by the muon CP group [37] and applied as up/down variations of the  
 1041 nominal SFs for each component.

1042 **7.2 Theoretical uncertainties**

1043 The theoretical systematics which are used for both the tri- and tetralepton channels are listed below.  
 1044 Theoretical systematics which aren't common across both channels are listed for the tri- and tetralepton  
 1045 channels in Sections 7.2.1 and 7.2.2 respectively.

1046 •  **$t\bar{t}Z$  background:**

1047 An overall normalization uncertainty of 10% is considered for the  $t\bar{t}Z$  background. Two generic  
 1048 shape systematics are considered for the  $t\bar{t}Z$  background. They are constructed (see Section 7.3) by  
 1049 either applying a linear or triangular interpolation to up and down variations which are defined to be  
 1050  $\pm 20\%$  from the nominal  $t\bar{t}Z$  background.

1051 •  **$ZZ$  background:**

1052 An overall normalization uncertainty of 20% and 30% is considered for the  $t\bar{t}Z$  background in the  
 1053 trilepton and tetralepton channels respectively.

1054 •  **$t\bar{t}H$  background:**

1055 An overall normalization uncertainty of 20% is considered for the  $t\bar{t}H$  background.

1056 •  **$tZq$  background:**

1057 An overall normalization uncertainty of 14% is considered for the  $tZq$  background.

1058 **7.2.1 Trilepton Channel**

1059 –  **$WZ$  background:**

1060 The overall normalization uncertainties on  $WZ + b$ ,  $WZ + c$ , and  $WZ + l$  were set to 50%,  
 1061 40%, and 30% respectively. Additionally, generic triangular and linear shape systematics  
 1062 (see Section 7.3) are considered for the  $WZ + b$  and  $WZ + c$  backgrounds. Generic shape  
 1063 systematics are not considered for the  $WZ + l$  background, since it contributes very few events  
 1064 in the SR and CEs

1065 –  **$t\bar{t}$  background:** The overall normalization uncertainty on  $t\bar{t}$  was set to 50%.

1066 –  **$Z + jets$  background:** The overall normalization uncertainty on  $Z + jets$  was set to 50%.

1067    **7.2.2 Tetralepton Channel**

1068    –  **$t\bar{t}Z$  fake background:**

1069    An overall normalization uncertainty of 50% is considered for the  $t\bar{t}Z$  fake background.

1070    – **other background processes:**

1071    The 'other' background consists of many processes which have minimal but non-negligible  
 1072    contribution in the signal regions (See Section ??). An overall normalization uncertainty of  
 1073    30% is considered for the 'other' background processes in the tetralepton channel.

1074    –  **$tWZ$  background:**

1075    A modelling uncertainty on  $tWZ$  is considered by comparing the nominal sample (using the  
 1076    DR1 scheme) and a minimal DR2 sample (see Section ??).

1077    Two generic shape systematics are considered for the  $tWZ$  background. They are constructed  
 1078    (see Section 7.3) by either applying a linear or triangular interpolation to up and down variations  
 1079    which are defined to be  $\pm 20\%$  from the nominal  $tWZ$  background.

1080    **7.3 Generic shape systematics**

1081    It is evident that the tri- and tetralepton channels are statistically limited. We therefore expect that the  
 1082    uncertainty on  $u_{tWZ}$  is dominated by statistical uncertainty and that the impact of shape systematics will  
 1083    be negligible in comparison.

1084    In order to include shape uncertainties related to the modelling of our samples, we construct generic shape  
 1085    systematics for any given sample process. Given that we choose a sufficiently large set of values for which  
 1086    the systematics can take in the fit, the constructed systematics could represent many shape systematics  
 1087    which we have not yet considered to include in the fitting procedure.

1088    We start by constructing an envelope (error bars) consisting of two MC templates. One with the nominal MC  
 1089    template increased by 20% on its normalisation and the other with the nominal MC template decreased by  
 1090    20% on its normalisation. The templates are then modified from their original shape either by doing linear  
 1091    interpolation (from the leftmost-up variation to the rightmost-down variation) or triangular interpolation  
 1092    (shape is set to zero at the rightmost and leftmost parts and reaches the envelope in the middle). The linear  
 1093    and triangular interpolation is done using TRF's ForceShape option [63]. This envelope now represents  
 1094    the bounds which the systematic can vary in the fit.

1095    The envelope before and after the shape change, for both the linear and triangular interpolations, are shown  
 1096    in the Figure ?? (for the  $t\bar{t}Z$  background in the  $t\bar{t}Z$  CR in the tetralepton channel).

1097    We consider two shape systematics (linear and triangular shapes) for  $tWZ$  and the most dominant background  
 1098    processes across both channels. In particular, we consider these shape systematics for  $t\bar{t}Z$  for both the tri-  
 1099    and tetralepton channels. Additionally, we consider these shape systematics for the  $WZ + b$  and  $WZ + c$   
 1100    backgrounds in the trilepton channel.

Not reviewed, for internal circulation only

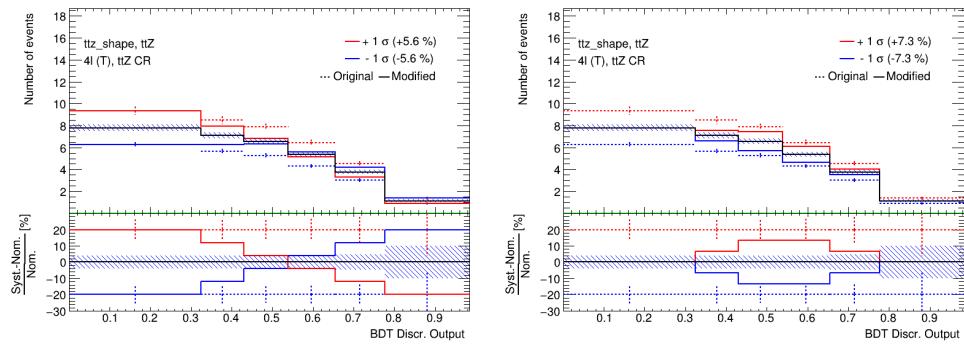


Figure 58:  $t\bar{t}Z$  generic shape systematic before (original) and after (modified) linear (left) and triangular (right) interpolation in the  $t\bar{t}Z$  CR (tetralepton channel)

---

## 1101 8 Results

### 1102 8.1 Fitting Procedure

1103 Using the TRExFitter framework, binned profile-likelihood fits are performed to determine the signal  
 1104 strength  $\mu_{tWZ} = \sigma_{obs}(tWZ)/\sigma_{SM}(tWZ)$  of  $tWZ$  production. First, separate fits are done in the trilepton  
 1105 and tetralpepton channels to determine the sensitivity of these channels. Then, a combined fit is performed  
 1106 for a more accurate measurement of  $\mu_{tWZ}$ .

1107 In the separate and combined fits, the *mixed data and MC* fit setup [64] is used. This is done to obtain  
 1108 a more accurate prediction of expected results while keeping the signal regions blinded. For this setup,  
 1109 first a background-only fit to the control regions using real data is done to determine estimates of the  
 1110 nuisance parameters. Then these estimates are used to construct a modified ASIMOV dataset in the signal  
 1111 regions. Finally, the fit is performed using real data in the control regions and the aforementioned modified  
 1112 ASIMOV data-set in the signal regions.

1113 In these fits, the parameter of interest (POI) is  $\mu_{tWZ}$ , and nuisance parameters are assigned to the systematic  
 1114 uncertainties outlined in 7 Furthermore, a gamma ( $\gamma$ ) nuisance parameter for a bin is added to the likelihood  
 1115 function if the statistical uncertainty in the bin exceeds 0.1% of its nominal value.

1116 Pruning is done per sample and per region on the shape and normalisation uncertainties for samples. A  
 1117 sample's shape and normalisation nuisance parameter is pruned (removed from the limit/fit) if the fraction  
 1118 of signal yield to the total yield (signal + background) is less than 0.01.

1119 Certain MC samples with minimal background contribution to the  $tWZ$  signal are not grouped with the  
 1120 'other' sample when defining their cross section pre-fit uncertainties (e.g.  $t\bar{t}h$ ). This is done for samples  
 1121 which are included in both the trilepton and tetralepton channels, such that in the combined fit, there is one  
 1122 nuisance parameter in the likelihood function which can vary for the cross section of that process. The  
 1123 events of these samples are still present in the data/MC plots in the 'other' category.

Not reviewed, for internal circulation only

## 1124 8.2 Tetralepton Channel

1125 The results shown in this section are blinded. We use **TRExFitter**'s *mixed data and MC fit* [64], which  
 1126 aims to obtain the most accurate representation for the expected results, while keeping the SR completely  
 1127 blind to data. It does this by performing a background only fit, using real data, to the CRs. This fit provides  
 1128 a set of fitted values for the nuisance parameters. These fitted nuisance parameters are then used to construct  
 1129 a modified ASIMOV dataset in the SR. The fitting procedure is then run using the modified ASIMOV  
 1130 dataset in the SR and real data in the CRs.

Not reviewed, for internal circulation only

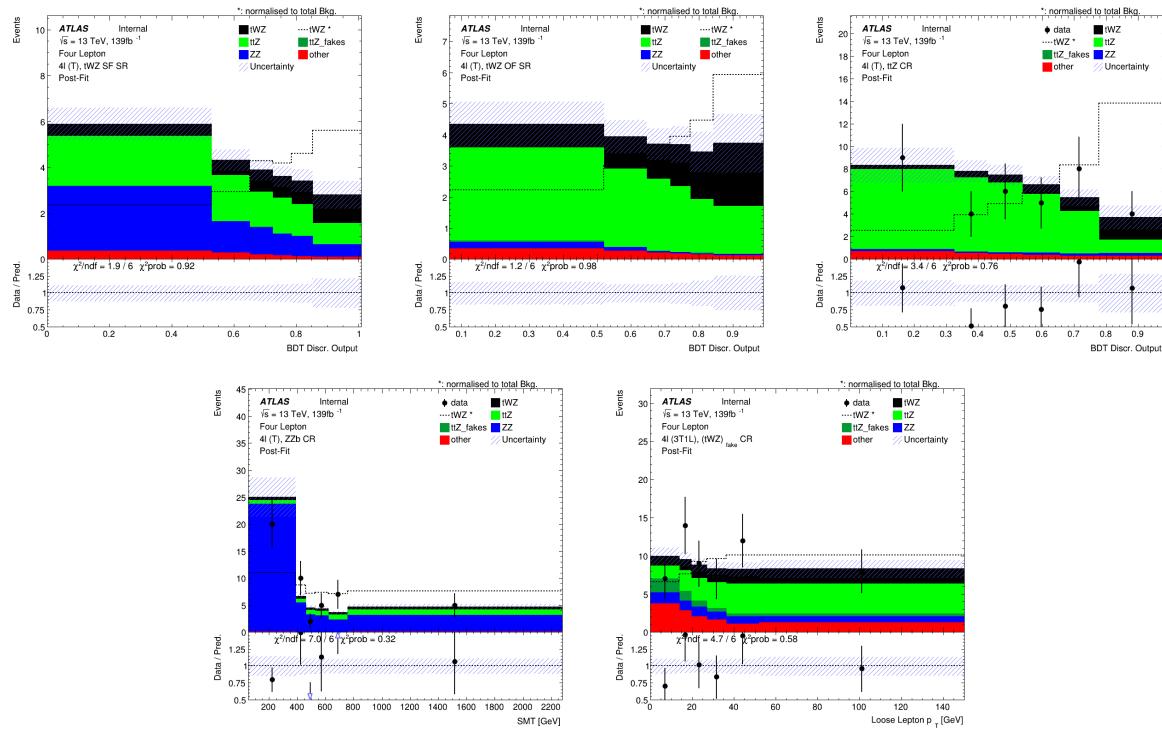
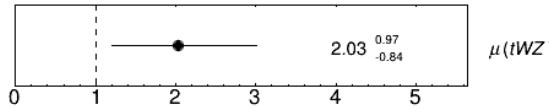


Figure 59: Post-fit Plots of fit variables

Figure 60: Post-fit Plots (blinded) of variables used in the fit. Top left: event-level BDT Disc. Output in the  $tWZ$  SF (4T) SR region. Middle Top: event-level BDT Disc. Output in the  $tWZ$  SF (4T) SR region. Top right: event-level BDT Disc. Output in the  $t\bar{Z}$  CR region. Bottom left:  $SMT = \sum p_T(\ell) + \sum p_T(jet) + E_T^{\text{miss}}$  in the  $ZZb$  CR. Bottom right:  $p_T$ (loose lepton) in the  $(tWZ)$ fake (3T1L) CR

	$tWZ$ 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	$t\bar{t}$ $tZq$ $t\bar{t}tW$ $WZ$ $t\bar{t}t$ $t\bar{t}\bar{t}$ $t\bar{t}WW$ $VVV(V = W/Z)$ $t\bar{t}H$ $WZ$	$6.00553e-06 \pm 3.02819e-06$ $0.0827905 \pm 0.0398773$ $0.00668643 \pm 0.00792217$ $0.0442934 \pm 0.024156$ $0.000987164 \pm 0.000766266$ $0.00934035 \pm 0.0080554$ $0.0294618 \pm 0.0263174$ $0.280643 \pm 0.0853411$ $0.846054 \pm 0.175495$ $0.0444253 \pm 0.0259642$	$0.252557 \pm 0.442116$ $0.0756107 \pm 0.0354584$ $0.00279748 \pm 0.00287361$ $0.0396511 \pm 0.0154282$ $0.00247481 \pm 0.00136945$ $0.0107458 \pm 0.00849984$ $0.029771 \pm 0.0195582$ $0.191433 \pm 0.0586778$ $0.669375 \pm 0.140107$ $0.0397691 \pm 0.0176416$	$0.273507 \pm 0.223201$ $0.063585 \pm 0.0293325$ $6.00553e-06 \pm 3.02819e-06$ $0.0133471 \pm 0.0128199$ $0.0140869 \pm 0.00479496$ $0.0571373 \pm 0.0204011$ $0.264364 \pm 0.0926252$ $0.0697266 \pm 0.0225059$ $1.96662 \pm 0.401199$ $0.0133868 \pm 0.0131418$	$6.00553e-06 \pm 3.02819e-06$ $0.05884 \pm 0.0244084$ $6.00553e-06 \pm 0.00564349$ $0.0472562 \pm 0.0330315$ $6.00553e-06 \pm 3.02819e-06$ $0.00553e-06 \pm 3.02819e-06$ $0.0129431 \pm 0.0323803$ $0.171142 \pm 0.0518102$ $0.150025 \pm 0.0353826$ $0.0473967 \pm 0.0345694$	$2.35427 \pm 0.917112$ $4.90963 \pm 0.745354$ $0.943182 \pm 0.292148$ $1.83567 \pm 0.392459$ $0.0100558 \pm 0.00363157$ $0.0216809 \pm 0.00992819$ $0.152448 \pm 0.058965$ $0.266502 \pm 0.0810194$ $2.21264 \pm 0.451623$ $1.84113 \pm 0.557788$
Total	24	26	36	49	57	
data						

Table 17: Post-fit yields for each sample in each region

1131  
Expected limit,  $\mu$ :  $\mu_{\text{expected}} = 1.62^{+2.37}_{-1.17}$ 1132  
Expected Significance =  $1.44\sigma$ Figure 61: Signal strength,  $\mu$ , of the fit

Not reviewed, for internal circulation only

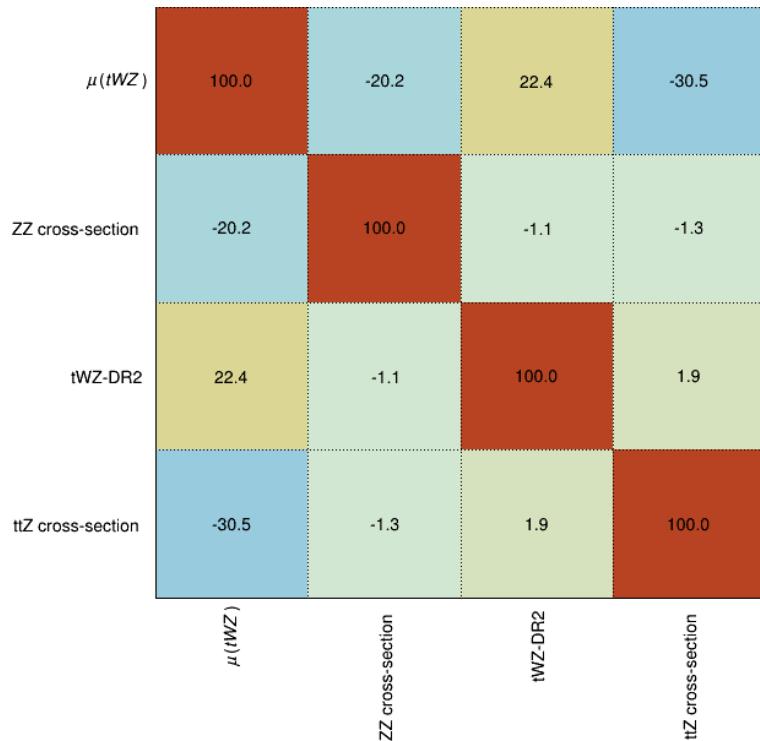


Figure 62: Correlation matrix of the fit

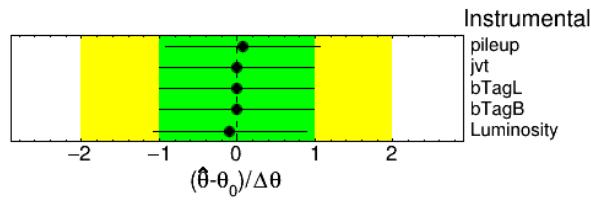


Figure 63: Pulls and constraints of the nuisance parameters (Instrumental)

Not reviewed, for internal circulation only

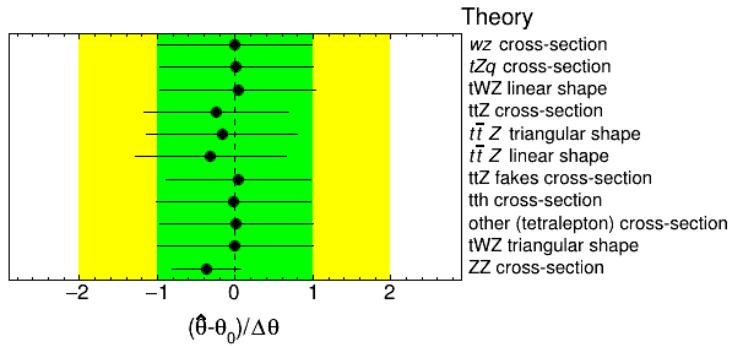


Figure 64: Pulls and constraints of the nuisance parameters (Theory)

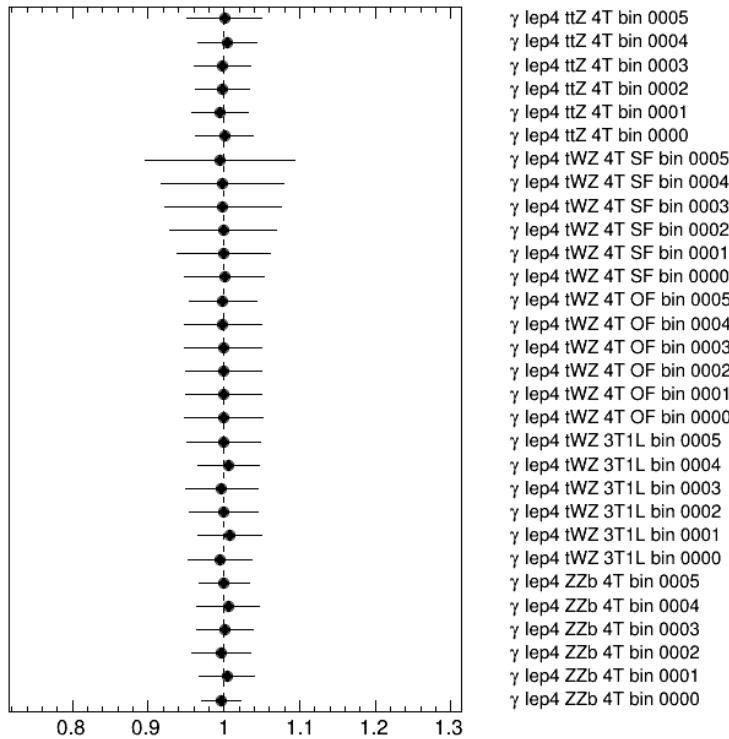


Figure 65: Gamma parameters of the unblinded fit

Not reviewed, for internal circulation only

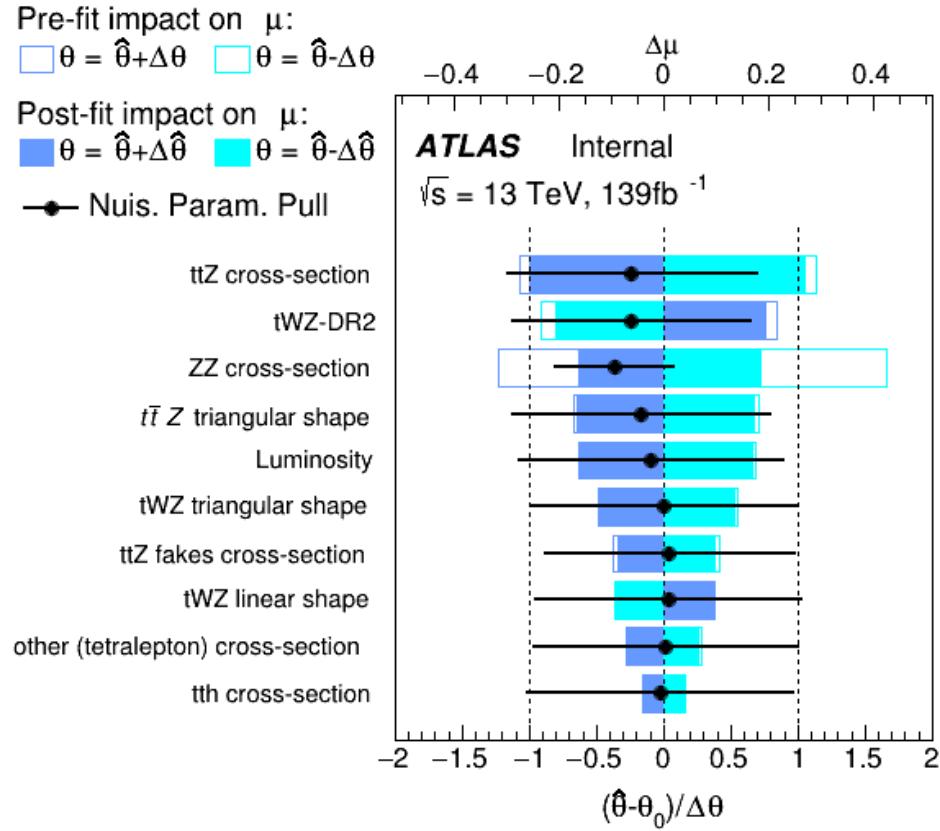


Figure 66: Ranking of the nuisance parameters of the fit

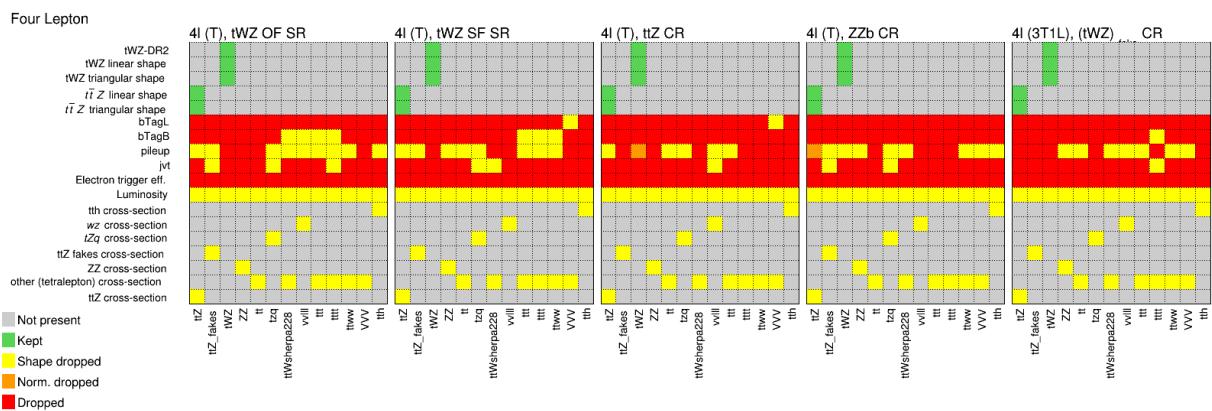


Figure 67: Pruning of the nuisance parameters of the fit

### 1133 8.3 Trilepton Channel

1134 The fit in the trilepton channel is performed using the histograms of the event GBDT score in the SR and  
 1135 CRs, with the autobinning setting turned on to minimize statistical fluctuations.

1136 The expected signal strength resulting from the fit to the modified Asimov dataset is:

$$\hat{\mu} = 1.20^{+1.37}_{-1.29}$$

1137 The best fit value is consistent with  $\mu = 1$ , as expected from a fit to a modified Asimov dataset. In Figure  
 1138 68, the post-fit histograms of the event GBDT score are shown for simulation and data, where the estimates  
 1139 of nuisance parameters extracted from the background-only fit to the CRs are used. Comparing these to the  
 1140 pre-fit plots in Figure 22, one sees that the uncertainty bands have narrowed due to the use of the nuisance  
 1141 parameters estimated from the background-only fit in the CRs. In Figure 68(c), there is good agreement  
 1142 between simulation and data in the  $WZ$  CR even given the narrowing of uncertainty bands.

1143 The expected significance of  $\mu$  was found to be

$$Z_\mu^{\text{exp}} = 0.77 \sigma$$

1144 Therefore, given the current analysis setup, one does not expect to be able to measure  $tWZ$  production in  
 1145 the trilepton channel at the level necessary to constitute evidence for this process.

1146 When one does not expect to observe a process, the result of the fit is reported by giving an upper limit on  
 1147 the parameter of interest. Since this is the case here, this is what is done. In TRExFitter, the upper limit is  
 1148 calculated using the  $CL_s$  procedure [65]. By convention, a confidence level of 95% ( $\beta = 0.05$ ) is chosen  
 1149 such that the interval  $[0, \mu_{\text{up}}]$  will cover  $\mu$  with a probability of at least 95%. The expected upper limit  
 1150 determined from the fit to the modified Asimov dataset was found to be

$$\mu_{\text{up}}^{\text{exp}} = 2.64^{+2.34}_{-1.22}$$

1151 According to Figure 69, the leading source of uncertainty is found to be the uncertainty due to the  $WZ + b$   
 1152 production cross section. Most nuisance parameter pulls are consistent with 0, as expected from a fit to a  
 1153 modified Asimov dataset.

#### 1154 8.3.1 Studying the expected significance and expected upper limit at higher luminosity

1155 In Figure 70, plots of  $Z_\mu^{\text{exp}}$  and  $\mu_{\text{up}}^{\text{exp}}$  are shown for fits to Asimov datasets where the simulation has been  
 1156 scaled to a higher integrated luminosity,  $\mathcal{L}$ , for each fit. A modified Asimov dataset, like that used for the  
 1157 extraction of the signal strength, was not used since there is  $139 \text{ fb}^{-1}$  of data that could be used. Fits to  
 1158 Asimov datasets with simulation scaled to higher luminosity were done in order to determine if this  
 1159 analysis is statistically limited. In Figure 70(a),  $Z_\mu^{\text{exp}}$  increases with increased luminosity, showing that a  
 1160 more accurate measurement could be achieved with more data. In Figure 70(b),  $\mu_{\text{up}}^{\text{exp}}$  and the uncertainties  
 1161 on  $\mu_{\text{up}}^{\text{exp}}$  decrease with increased luminosity. This shows that the current analysis in this channel is strongly  
 1162 statistically limited.

Not reviewed, for internal circulation only

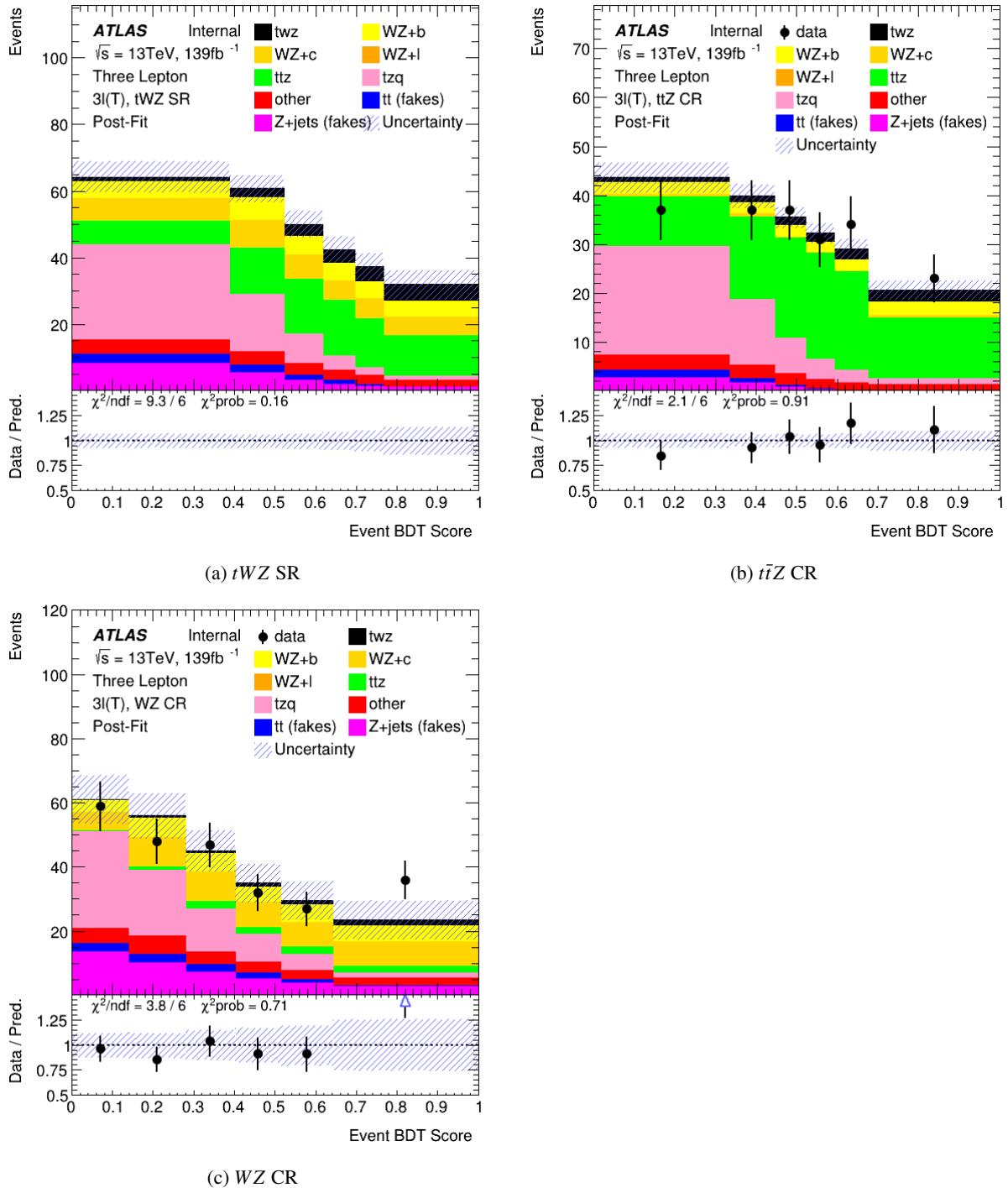


Figure 68: The post-fit event GBDT score histograms are shown with the post-fit model resulting in narrowed uncertainty bands.

Not reviewed, for internal circulation only

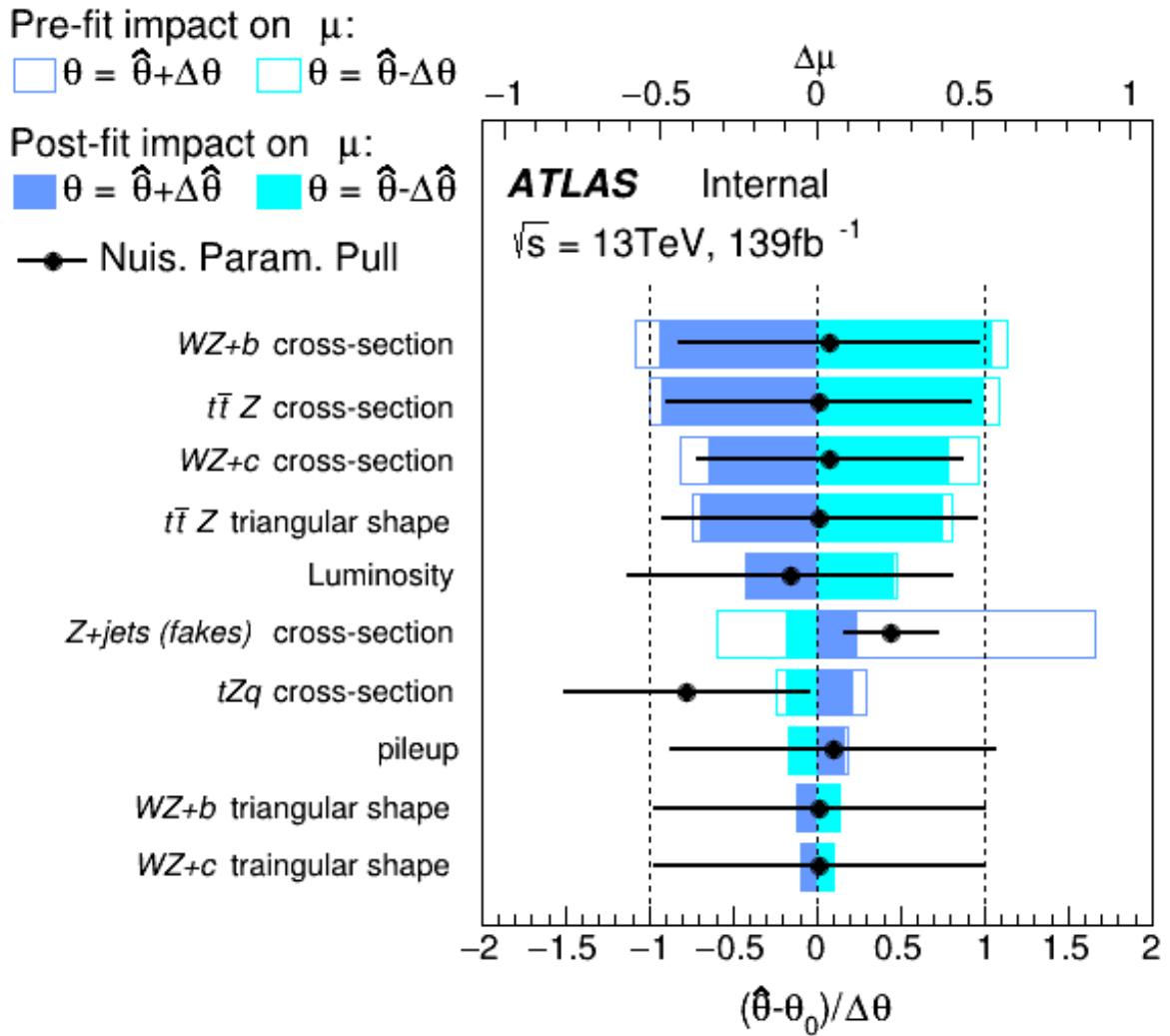
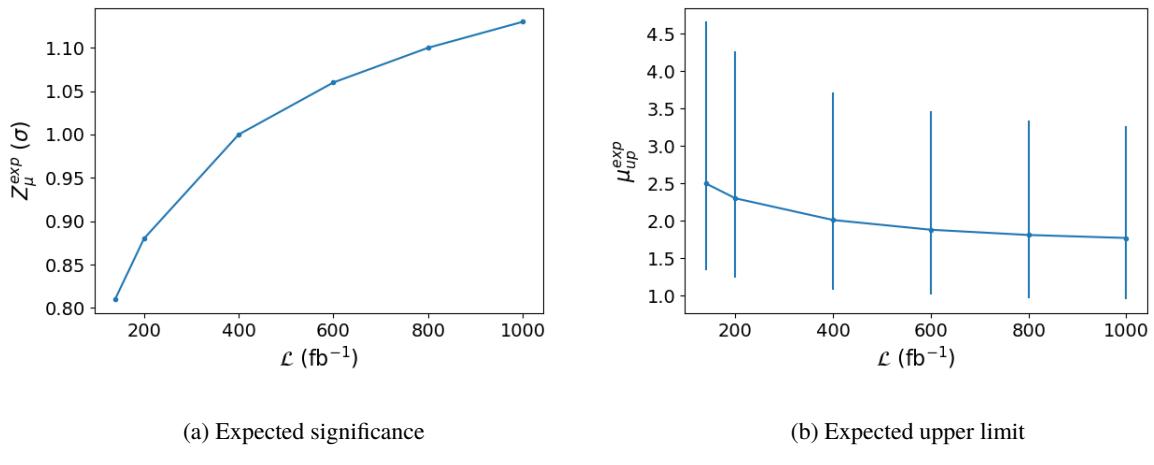


Figure 69: A ranking plot of the impact of the nuisance parameters is shown. The pre-fit and post-fit impacts are shown by empty and coloured rectangles, and the nuisance parameter pull values are shown by the black points. On the upper axis the impact,  $\Delta\mu$ , is shown and on the lower axis the nuisance parameter pull is shown.

Not reviewed, for internal circulation only



(a) Expected significance

(b) Expected upper limit

Figure 70: Plots of  $Z_\mu^{\text{exp}}$  (a) and  $\mu_{\text{up}}^{\text{exp}}$  (b) are shown for increasing  $\mathcal{L}$

Not reviewed, for internal circulation only

Nuisance Parameter	Channel Affected	
	trilepton	tetralepton
$\sigma(t\bar{t}H)$	✓	✓
$\sigma(t\bar{t}Z)$	✓	✓
$\sigma(WZ)$	✓	✓
$\sigma(tZq)$	✓	✓
$\sigma(ZZ)$	✓	✓
$\sigma(\text{other(trilepton)})$	✓	✗
$\sigma(\text{other(tetralepton)})$	✗	✓
$\sigma(t\bar{t}Z)_{fakes}$	✗	✓
$\sigma(t\bar{t})_{fakes}$	✓	✗
$\sigma(Z + \text{jets})_{fakes}$	✓	✗
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 18: Summary of the nuisance parameters used in the combined fit. The channels which are affected by each nuisance parameter are indicated with a ✓ (is affected) or a ✗ (is not affected).

## 1163 8.4 Combined Tri- and Tetralepton Channels

1164 In this section, fit results and plots of the combined tri and tetralepton channels are shown.

1165 The likelihood fit is performed over all regions defined for the tri and tetralepton channels (Figure ??  
 1166 (trilepton analysis) and Figure ?? (tetralepton analysis)), where the signal strength  $\mu tWZ$  of  $tWZ$  is  
 1167 considered as a free parameter in the fit (parameter of interest, POI). The results shown in this section<sup>5</sup> are  
 1168 blinded, where the signal regions contain a modified ASIMOV dataset (using TRexFitter's *mixed data*  
 1169 and *MC fit*). The  $WZ$  and  $tZq$  (trilepton),  $t\bar{t}Z$  (trilepton),  $t\bar{t}Z$  (tetralepton),  $ZZb$  (tetralepton) and  $(tWZ)_{fake}$   
 1170 control regions use real data.

1171 The nuisance parameters used in the fit, including which channel's regions are affected by each, are  
 1172 summarised in Table ?? .

1173 Expected Limit:  $1.5^{+2.1}_{-1.1} \times \sigma_{SM}^{tWZ} \text{ fb}$

1174 Expected significance:  $1.6\sigma$

<sup>5</sup> unless otherwise specified

Not reviewed, for internal circulation only

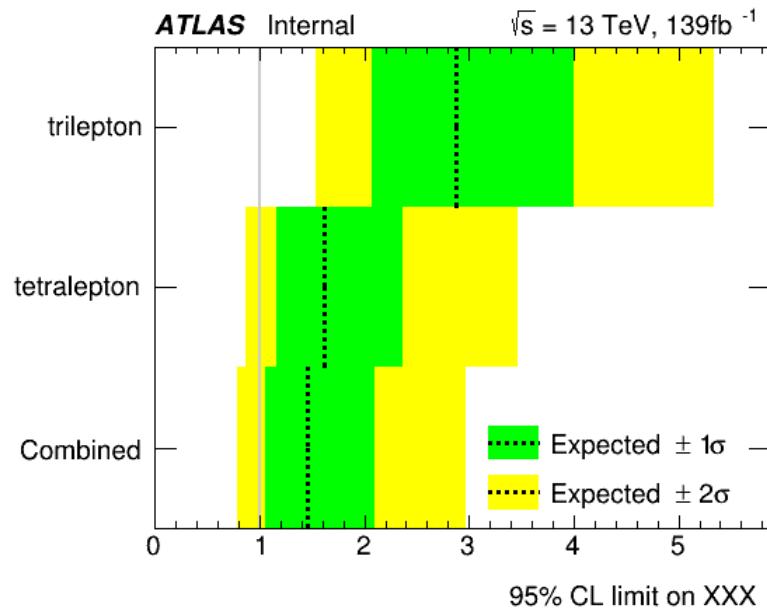


Figure 71: Trilepton, Tetralepton and Combined Expected Limits

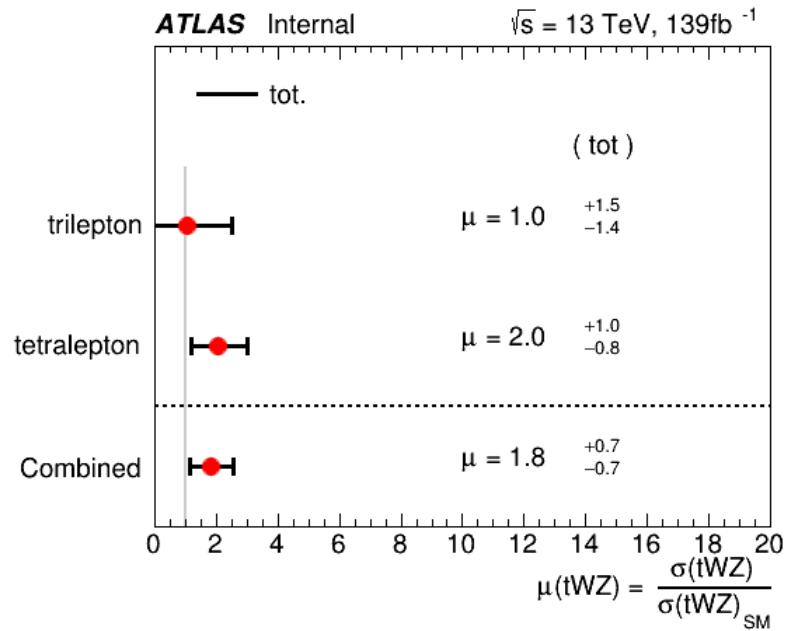


Figure 72: Trilepton, Tetralepton and Combined  $\mu_{\text{best-fit}}$

1175 It can be seen from Figures 71 and 72 that the best fit value for the signal strength on  $tWZ$ ,  $\mu(tWZ)$ , and  
 1176 the expected limits for the tri- and tetralepton channels are consistent with one another (their uncertainties  
 1177 overlap). Therefore it is appropriate to combine these two analyses.

Not reviewed, for internal circulation only

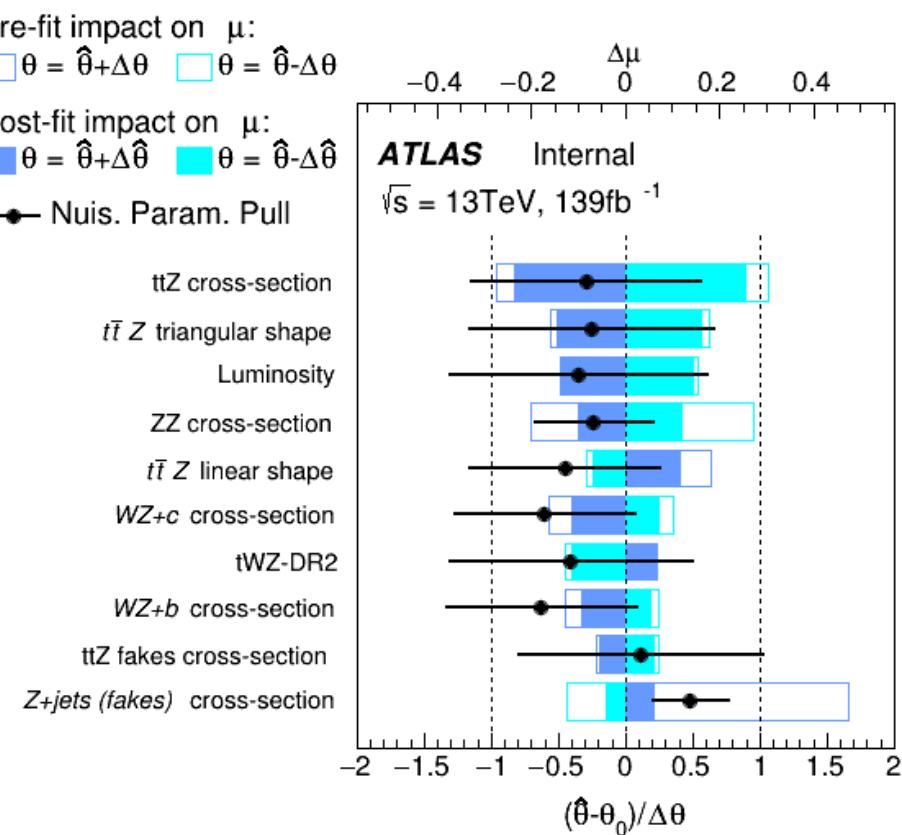


Figure 73: Ranking of the nuisance parameters of the fit

<sup>1178</sup> **9 Summary and Conclusions**

[Not reviewed, for internal circulation only]

1179 You can find some text snippets that can be used in papers in `template/atlas-snippets.tex`. Some of  
 1180 the snippets need the `jetetmiss` option passed to `atlasphysics`.

Not reviewed, for internal circulation only

## 1181 References

- 1182 [1] O. Oncel, *Search for Single Top Quark Production in Association with a W and a Z Boson in the 3  
 1183 Lepton Final State with the ATLAS Experiment at 13 TeV*, URL: [https://cds.cern.ch/record/  
 1184 2625170](https://cds.cern.ch/record/2625170) (cit. on pp. 4, 5, 27, 43).
- 1185 [2] O. Bessidskaia Bylund, ‘Top quark and heavy vector boson associated production at the ATLAS  
 1186 experiment : Modelling, measurements and effective field theory’, PhD thesis: Stockholm U., 2017  
 1187 (cit. on p. 5).
- 1188 [3] F. Maltoni, L. Mantani and K. Mimasu, *Top-quark electroweak interactions at high energy*, **JHEP**  
 1189 **10** (2019) 004, arXiv: [1904.05637 \[hep-ph\]](https://arxiv.org/abs/1904.05637) (cit. on p. 5).
- 1190 [4] S. Frixione, E. Laenen, P. Motylinski, B. R. Webber and C. D. White, *Single-top hadroproduction in  
 1191 association with a W boson*, **JHEP** **07** (2008) 029, arXiv: [0805.3067 \[hep-ph\]](https://arxiv.org/abs/0805.3067) (cit. on p. 5).
- 1192 [5] C. Degrande, F. Maltoni, K. Mimasu, E. Vryonidou and C. Zhang, *Single-top associated production  
 1193 with a Z or H boson at the LHC: the SMEFT interpretation*, **Journal of High Energy Physics** **2018**  
 1194 (2018) 5, ISSN: 1029-8479, URL: [https://doi.org/10.1007/JHEP10\(2018\)005](https://doi.org/10.1007/JHEP10(2018)005) (cit. on p. 5).
- 1195 [6] *AnalysisBase configuration files for tWZ ntuple production*, URL: [https://gitlab.cern.ch/  
 1196 atlas-sa-uct/uctanalysistop/-/tree/master/ATNtupleProduction](https://gitlab.cern.ch/atlas-sa-uct/uctanalysistop/-/tree/master/ATNtupleProduction) (cit. on p. 8).
- 1197 [7] ATLAS Collaboration, *Luminosity determination in pp collisions at  $\sqrt{s} = 8$  TeV using the ATLAS  
 1198 detector at the LHC*, **Eur. Phys. J. C** **76** (2016) 653, arXiv: [1608.03953 \[hep-ex\]](https://arxiv.org/abs/1608.03953) (cit. on p. 9).
- 1199 [8] G. Avoni et al., *The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*,  
 1200 **JINST** **13** (2018) P07017 (cit. on pp. 9, 78).
- 1201 [9] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with Parton Shower  
 1202 simulations: the POWHEG method*, **JHEP** **11** (2007) 070, arXiv: [0709.2092 \[hep-ph\]](https://arxiv.org/abs/0709.2092) (cit. on  
 1203 p. 10).
- 1204 [10] T. Sjöstrand et al., *An Introduction to PYTHIA 8.2*, **Comput. Phys. Commun.** **191** (2015) 159, arXiv:  
 1205 [1410.3012 \[hep-ph\]](https://arxiv.org/abs/1410.3012) (cit. on p. 10).
- 1206 [11] NNPDF Collaboration, R.D. Ball et al., *Parton distributions for the LHC Run II*, **JHEP** **04** (2015) 040,  
 1207 arXiv: [1410.8849 \[hep-ph\]](https://arxiv.org/abs/1410.8849) (cit. on p. 10).
- 1208 [12] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021, 2014, URL:  
 1209 <https://cds.cern.ch/record/1966419> (cit. on pp. 10, 11).
- 1210 [13] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential  
 1211 cross sections, and their matching to parton shower simulations*, **JHEP** **1407** (2014) 079, arXiv:  
 1212 [1405.0301 \[hep-ph\]](https://arxiv.org/abs/1405.0301) (cit. on pp. 10, 11).
- 1213 [14] ATLAS Collaboration, *The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic  
 1214 measurements incorporating the Donnachie–Landshoff diffractive model*, ATL-PHYS-PUB-2016-  
 1215 017, 2016, URL: <https://cds.cern.ch/record/2206965> (cit. on p. 9).
- 1216 [15] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, **Eur. Phys. J. C** **70** (2010) 823, arXiv:  
 1217 [1005.4568 \[physics.ins-det\]](https://arxiv.org/abs/1005.4568) (cit. on p. 10).

- [16] ATLAS Collaboration, *The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim*, ATL-PHYS-PUB-2010-013, 2010, URL: <https://cds.cern.ch/record/1300517> (cit. on p. 10).
- [17] ATLAS Collaboration, *The new Fast Calorimeter Simulation in ATLAS*, ATL-SOFT-PUB-2018-002, 2018, URL: <https://cds.cern.ch/record/2630434> (cit. on p. 10).
- [18] T. Gleisberg and S. Höche, *Comix, a new matrix element generator*, JHEP **12** (2008) 039, arXiv: [0808.3674 \[hep-ph\]](https://arxiv.org/abs/0808.3674) (cit. on p. 11).
- [19] S. Schumann and F. Krauss, *A Parton shower algorithm based on Catani-Seymour dipole factorisation*, JHEP **03** (2008) 038, arXiv: [0709.1027 \[hep-ph\]](https://arxiv.org/abs/0709.1027) (cit. on p. 11).
- [20] S. Höche, F. Krauss, M. Schonherr and F. Siegert, *A critical appraisal of NLO+PS matching methods*, JHEP **09** (2012) 049, arXiv: [1111.1220 \[hep-ph\]](https://arxiv.org/abs/1111.1220) (cit. on p. 11).
- [21] S. Höche, F. Krauss, M. Schonherr and F. Siegert, *QCD matrix elements + parton showers: The NLO case*, JHEP **04** (2013) 027, arXiv: [1207.5030 \[hep-ph\]](https://arxiv.org/abs/1207.5030) (cit. on p. 11).
- [22] S. Catani, F. Krauss, R. Kuhn and B. R. Webber, *QCD matrix elements + parton showers*, JHEP **11** (2001) 063, arXiv: [hep-ph/0109231](https://arxiv.org/abs/hep-ph/0109231) (cit. on p. 11).
- [23] S. Höche, F. Krauss, S. Schumann and F. Siegert, *QCD matrix elements and truncated showers*, JHEP **05** (2009) 053, arXiv: [0903.1219 \[hep-ph\]](https://arxiv.org/abs/0903.1219) (cit. on p. 11).
- [24] F. Caccioli, P. Maierhofer and S. Pozzorini, *Scattering Amplitudes with Open Loops*, Phys. Rev. Lett. **108** (2012) 111601, arXiv: [1111.5206 \[hep-ph\]](https://arxiv.org/abs/1111.5206) (cit. on p. 11).
- [25] A. Denner, S. Dittmaier and L. Hofer, *Collier: a fortran-based Complex One-Loop Library in Extended Regularizations*, Comput. Phys. Commun. **212** (2017) 220, arXiv: [1604.06792 \[hep-ph\]](https://arxiv.org/abs/1604.06792) (cit. on p. 11).
- [26] S. Frixione, E. Laenen, P. Motylinski and B. R. Webber, *Angular correlations of lepton pairs from vector boson and top quark decays in Monte Carlo simulations*, JHEP **0704** (2007) 081, arXiv: [hep-ph/0702198](https://arxiv.org/abs/hep-ph/0702198) (cit. on p. 11).
- [27] P. Artoisenet, R. Frederix, O. Mattelaer and R. Rietkerk, *Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations*, JHEP **1303** (2013) 015, arXiv: [1212.3460 \[hep-ph\]](https://arxiv.org/abs/1212.3460) (cit. on p. 11).
- [28] ATLAS Internal, *ASG Analysis Release 21.2.X series. Release Notes*, 2021, URL: [https://twiki.cern.ch/twiki/bin/view/AtlasProtected/AnalysisBaseReleaseNotes21\\_2](https://twiki.cern.ch/twiki/bin/view/AtlasProtected/AnalysisBaseReleaseNotes21_2) (cit. on p. 13).
- [29] ATLAS Collaboration, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at  $\sqrt{s} = 13$  TeV*, Eur. Phys. J. C **79** (2019) 639, arXiv: [1902.04655 \[hep-ex\]](https://arxiv.org/abs/1902.04655) (cit. on p. 13).
- [30] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data*, (2019), arXiv: [1908.00005 \[hep-ex\]](https://arxiv.org/abs/1908.00005) (cit. on pp. 13, 78).
- [31] ATLAS Internal, *Electron and Photon Selection and Identification for Run 2*, 2021, URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/EGammaIdentificationRun2> (cit. on p. 13).
- [32] ATLAS Internal, *Lepton Isolation Working points*, 2021, URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/RecommendedIsolationWPs> (cit. on pp. 13, 14).

- [33] ATLAS Internal, *Electron Efficiencies for Analyses*, 2021, URL: <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/ElectronEfficienciesForAnalysis> (cit. on pp. 13, 78).
- [34] ATLAS Internal, *MCP Analysis Guidelines*, 2021, URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/MCPAnalysisGuidelinesMC16> (cit. on p. 14).
- [35] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at  $\sqrt{s} = 13 \text{ TeV}$* , *Eur. Phys. J. C* **76** (2016) 292, arXiv: [1603.05598 \[hep-ex\]](https://arxiv.org/abs/1603.05598) (cit. on p. 14).
- [36] ATLAS Collaboration, *Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at  $\sqrt{s} = 13 \text{ TeV}$* , (2020), arXiv: [2012.00578 \[hep-ex\]](https://arxiv.org/abs/2012.00578) (cit. on pp. 14, 79).
- [37] ATLAS Internal, *Muon Efficiencies for Analyses*, 2021, URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/MuonEfficienciesForAnalysis> (cit. on pp. 14, 79).
- [38] ATLAS Collaboration, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, *Eur. Phys. J. C* **77** (2017) 466, arXiv: [1703.10485 \[hep-ex\]](https://arxiv.org/abs/1703.10485) (cit. on p. 14).
- [39] ATLAS Collaboration, *Jet energy scale and resolution measured in proton–proton collisions at  $\sqrt{s} = 13 \text{ TeV}$  with the ATLAS detector*, (2020), arXiv: [2007.02645 \[hep-ex\]](https://arxiv.org/abs/2007.02645) (cit. on p. 14).
- [40] ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, ATLAS-CONF-2014-018, 2014, URL: <https://cds.cern.ch/record/1700870> (cit. on p. 14).
- [41] ATLAS Collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at  $\sqrt{s} = 8 \text{ TeV}$  using the ATLAS detector*, *Eur. Phys. J. C* **76** (2016) 581, arXiv: [1510.03823 \[hep-ex\]](https://arxiv.org/abs/1510.03823) (cit. on p. 14).
- [42] ATLAS Internal, *Jet Vertex Tagger for Run 2 in reco and analysis*, 2021, URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/JetVertexTaggerTool> (cit. on pp. 14, 78).
- [43] ATLAS Internal, *Pileup jet recommendations*, 2021, URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/PileupJetRecommendations> (cit. on p. 14).
- [44] ATLAS Collaboration, *ATLAS b-jet identification performance and efficiency measurement with  $t\bar{t}$  events in pp collisions at  $\sqrt{s} = 13 \text{ TeV}$* , (2019), arXiv: [1907.05120 \[hep-ex\]](https://arxiv.org/abs/1907.05120) (cit. on pp. 14, 78).
- [45] ATLAS Internal, *Flavor Tagging Recommendations in Release 21*, 2021, URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/BTagTaggerRecommendationsRelease21> (cit. on p. 14).
- [46] ATLAS Internal, *EtMiss recommendations for Rel. 21*, 2021, URL: <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/EtmissRecommendationsFullRun2> (cit. on p. 15).
- [47] ATLAS Internal, *AnalysisTop Overlap Removal*, 2021, URL: [https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TopRecoObjTwikiModel#Overlap\\_Removal](https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TopRecoObjTwikiModel#Overlap_Removal) (cit. on p. 15).
- [48] T. P. Calvet, ‘Search for the production of a Higgs boson in association with top quarks and decaying into a b-quark pair and b-jet identification with the ATLAS experiment at LHC’, Presented 08 Nov 2017, 2017, URL: <https://cds.cern.ch/record/2296985> (cit. on p. 41).
- [49] P. D. Group et al., *Review of Particle Physics*, *Progress of Theoretical and Experimental Physics* **2020** (2020), 083C01, ISSN: 2050-3911, eprint: <https://academic.oup.com/ptep/article-pdf/2020/8/083C01/33653179/ptaa104.pdf>, URL: <https://doi.org/10.1093/ptep/ptaa104> (cit. on pp. 43, 45).

- [50] 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*, Journal of High Energy Physics **2019** (2019), ISSN: 1029-8479, URL: [%7Bhttp://dx.doi.org/10.1007/JHEP10\(2019\)127%7D](http://dx.doi.org/10.1007/JHEP10(2019)127) (cit. on p. 45).
- [51] 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](https://arxiv.org/abs/2009.01186) (hep-ex) (cit. on p. 45).
- [52] *Errors in weighted histograms*, URL: <https://www.zeuthen.desy.de/~wischnew/amanda/discussion/wgterror/working.html> (cit. on p. 50).
- [53] *IFFTruthClassifier GitLab Repository*, URL: <https://gitlab.cern.ch/ATLAS-IFF/IFFTruthClassifier/-/tree/master> (cit. on p. 65).
- [54] *IFFTruthClassifier Lepton Categories*, URL: <https://gitlab.cern.ch/ATLAS-IFF/IFFTruthClassifier/-/tree/master#3-details-about-the-lepton-categories> (cit. on p. 65).
- [55] *Scikit-Learn GradientBoostingClassifier Documentation*, URL: <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.GradientBoostingClassifier.html> (cit. on pp. 68, 70).
- [56] T. McCarthy and F. Fischer, *Exploiting full/partial  $t\bar{t}$  reconstruction for background suppression in  $2\ell$* , URL: [https://indico.cern.ch/event/955360/contributions/4016465/attachments/2102418/3534816/top\\_reco\\_bkgd\\_suppression\\_2L\\_20200915.pdf](https://indico.cern.ch/event/955360/contributions/4016465/attachments/2102418/3534816/top_reco_bkgd_suppression_2L_20200915.pdf) (cit. on p. 70).
- [57] T. McCarthy, *Macro developed to compare  $tW/Z$  reconstruction performance ( $2\ell, 3\ell, 4\ell$ )*, URL: [https://indico.cern.ch/event/986357/contributions/4172907/attachments/2169451/3666801/reco\\_performance\\_macro\\_20210112.pdf](https://indico.cern.ch/event/986357/contributions/4172907/attachments/2169451/3666801/reco_performance_macro_20210112.pdf) (cit. on p. 70).
- [58] ATLAS Collaboration, *Identification and rejection of pile-up jets at high pseudorapidity with the ATLAS detector*, Eur. Phys. J. C **77** (2017) 580, arXiv: [1705.02211](https://arxiv.org/abs/1705.02211) [hep-ex] (cit. on p. 78), Erratum: Eur. Phys. J. C **77** (2017) 712.
- [59] ATLAS Collaboration, *Measurements of  $b$ -jet tagging efficiency with the ATLAS detector using  $t\bar{t}$  events at  $\sqrt{s} = 13$  TeV*, JHEP **08** (2018) 089, arXiv: [1805.01845](https://arxiv.org/abs/1805.01845) [hep-ex] (cit. on p. 78).
- [60] 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> (cit. on p. 78).
- [61] 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> (cit. on p. 78).
- [62] ATLAS Internal, *Electron Efficiency Correlation Model*, 2021, URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/ElectronEfficiencyCorrelationModel> (cit. on p. 78).
- [63] *TRExFitter: Force shape by hand*, URL: <https://trexfitter-docs.web.cern.ch/trexfitter-docs/BasicTutorial2020/Shape/> (cit. on p. 80).
- [64] *TRExFitter: Mixed data and MC fit*, URL: <https://trexfitter-docs.web.cern.ch/trexfitter-docs/AdvancedTutorial2020/Mixed/> (cit. on pp. 82, 83).

- [65] T. Junk, *Confidence level computation for combining searches with small statistics*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **434** (1999) 435, ISSN: 0168-9002, URL: [http://dx.doi.org/10.1016/S0168-9002\(99\)00498-2](http://dx.doi.org/10.1016/S0168-9002(99)00498-2) (cit. on p. 88).
- [66] *TRExFitter config files for tWZ in the tetralepton channel*, URL: [https://gitlab.cern.ch/atlas-sa-uct/uctanalysistop/-/tree/master/analyses/twz\\_4lep/config\\_files](https://gitlab.cern.ch/atlas-sa-uct/uctanalysistop/-/tree/master/analyses/twz_4lep/config_files) (cit. on p. 136).
- [67] Cowan, Glen and Cranmer, Kyle and Gross, Eilam and Vitells, Ofer, *Asymptotic formulae for likelihood-based tests of new physics*, The European Physical Journal C **71** (2011), ISSN: 1434-6052, URL: <http://dx.doi.org/10.1140/epjc/s10052-011-1554-0> (cit. on p. 136).

Not reviewed, for internal circulation only

1351 The supporting notes for the analysis should also contain a list of contributors. This information should  
1352 usually be included in `mydocument-metadata.tex`. The list should be printed either here or before the  
1353 Table of Contents.

Not reviewed, for internal circulation only

1354 **List of contributions**

J. Keaveney	Overall project supervision, hadronic W reconstruction, lepton cut optimisation, framework development, ntuple production, statistical fitting
B. Warren	Tri-lepton channel - hadronic W reconstruction, lepton cut optimisation, statistical fitting
J. Reich	Tetra-lepton channel - statistical fitting, kinematic reconstruction, cut optimisation, statistical combination
S. Yacoob	Project supervision

# 1357 Appendices

## 1358 A Tetralepton Control Plots

### 1359 A.1 $tWZ$ OF SR

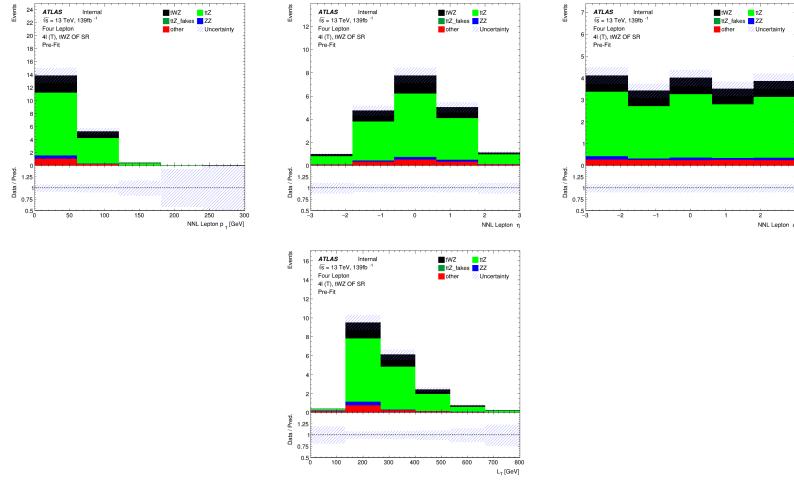


Figure 74: Top row: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  for next-to-next-to-leading (NNL) leptons in the  $tWZ$  OF SR region (*blinded*). Bottom row: MC predictions for  $L_T$  (scalar sum of lepton  $p_T$ ) in the  $tWZ$  OF SR region (*blinded*)

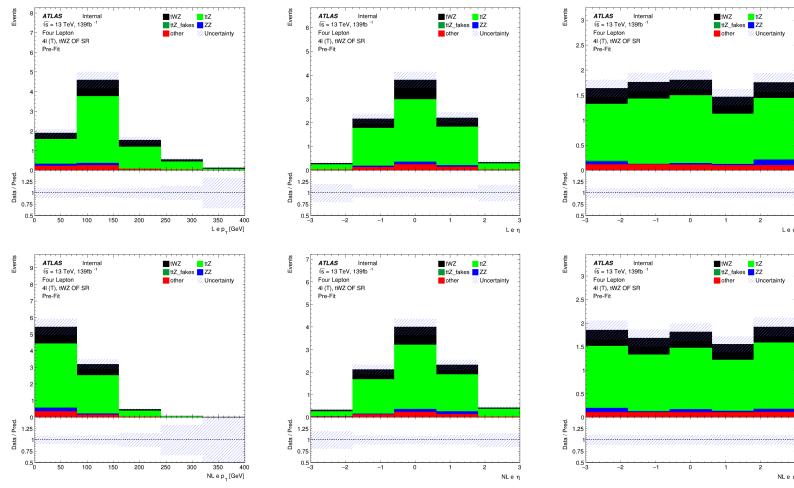


Figure 75: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  for leading (L) electrons (top row) and next-to-leading (NL) electrons (bottom row) in the  $tWZ$  OF SR region (*blinded*)

Not reviewed, for internal circulation only

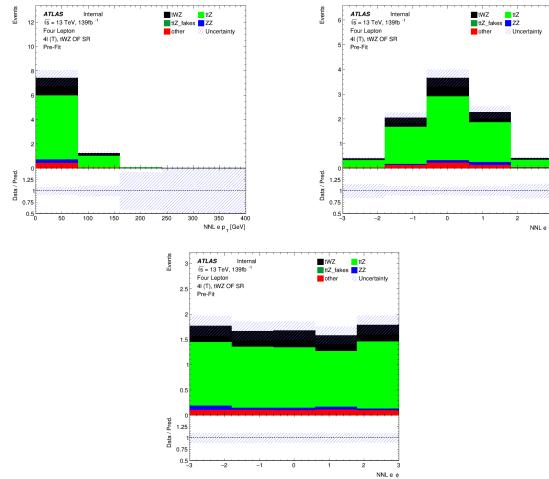


Figure 76: MC predictions for  $p_T$ ,  $\eta$  (top row) and  $\phi$  (bottom row) for next-to-next-to-leading (NNL) electrons in the  $t\text{WZ OF SR}$  region

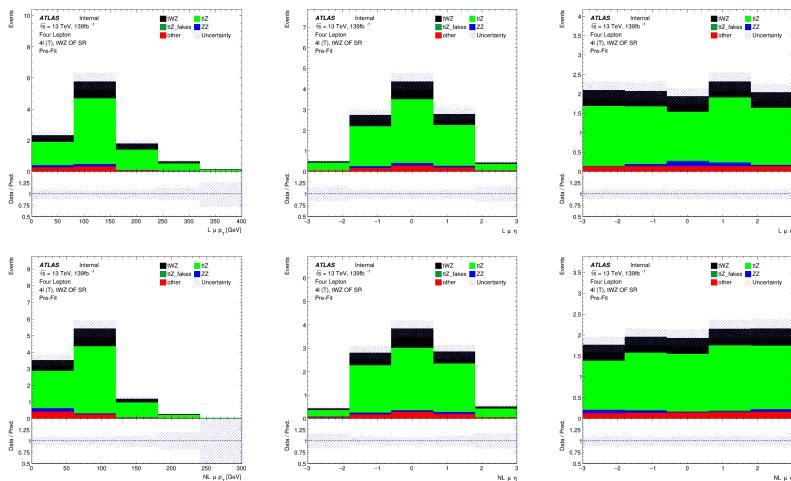


Figure 77: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  for leading (L) muons (top row) and next-to-leading (NL) muons (bottom row) in the  $t\text{WZ OF SR}$  region (*blinded*)

Not reviewed, for internal circulation only

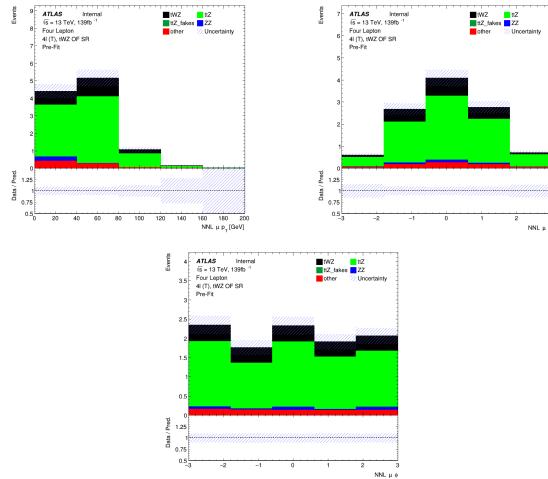


Figure 78: MC predictions for  $p_T$ ,  $\eta$  (top row) and  $\phi$  (bottom row) for next-to-next-to-leading (NNL) muons in the  $tWZ$  OF SR region (*blinded*)

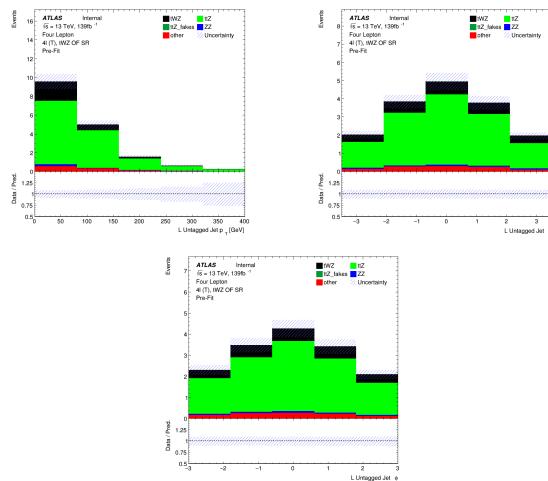


Figure 79: MC predictions for  $p_T$ ,  $\eta$  (top row) and  $\phi$  (bottom row) for untagged jets in the  $tWZ$  OF SR region (*blinded*)

Not reviewed, for internal circulation only

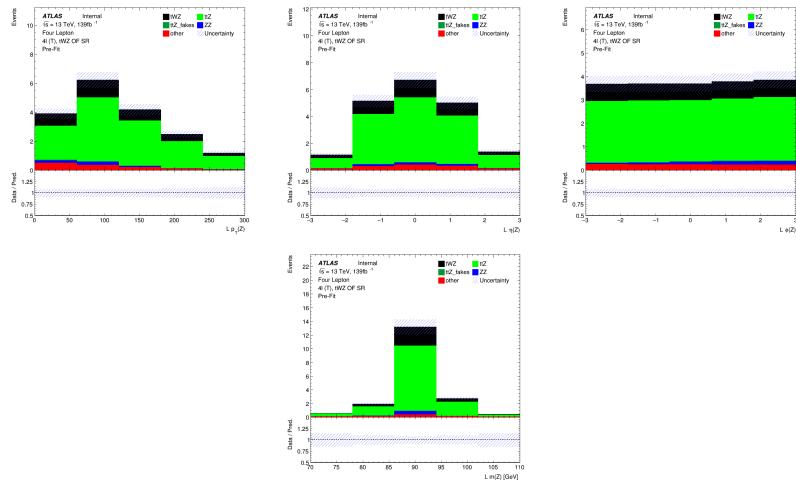


Figure 80: MC predictions for  $p_T$ ,  $\eta$ ,  $\phi$  (top row) and mass ( $m_Z$ ) (bottom row) of the leading reconstructed  $Z$  candidate (OSSF lepton pair with  $|m_{\text{OSSF}} - m_Z| < 30 \text{ GeV}$ ) in the  $t\text{WZ}$  OF SR region (*blinded*)

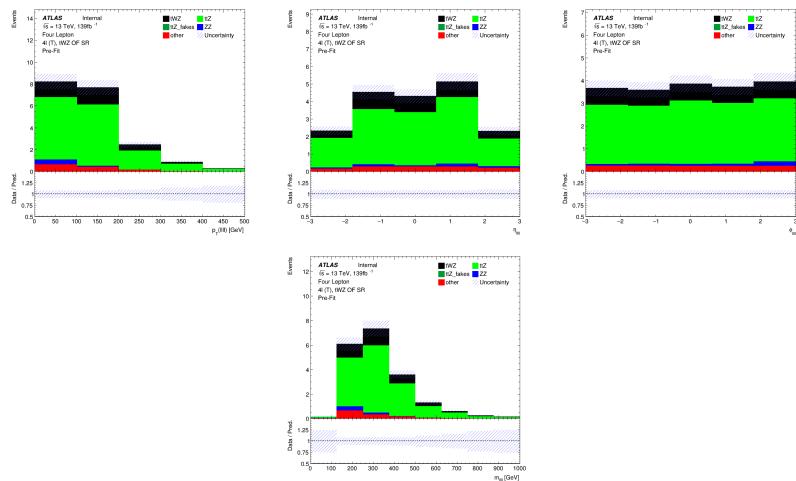


Figure 81: MC predictions for  $p_T$ ,  $\eta$ ,  $\phi$  (top row) and mass (bottom row) of the lepton system ( $\ell\ell\ell\ell$ ) in the  $t\text{WZ}$  OF SR region (*blinded*)

Not reviewed, for internal circulation only

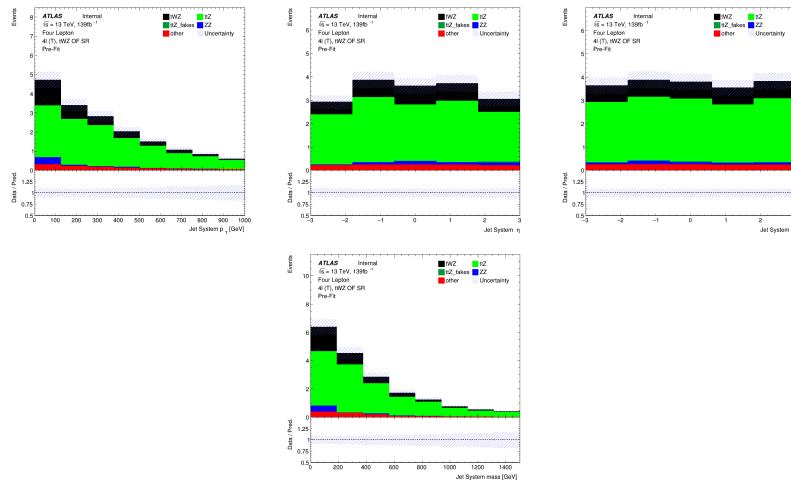


Figure 82: MC predictions for  $p_T$ ,  $\eta$ ,  $\phi$  (top row) and mass (bottom row) of the jet systems in the  $tWZ$  OF SR region (*blinded*)

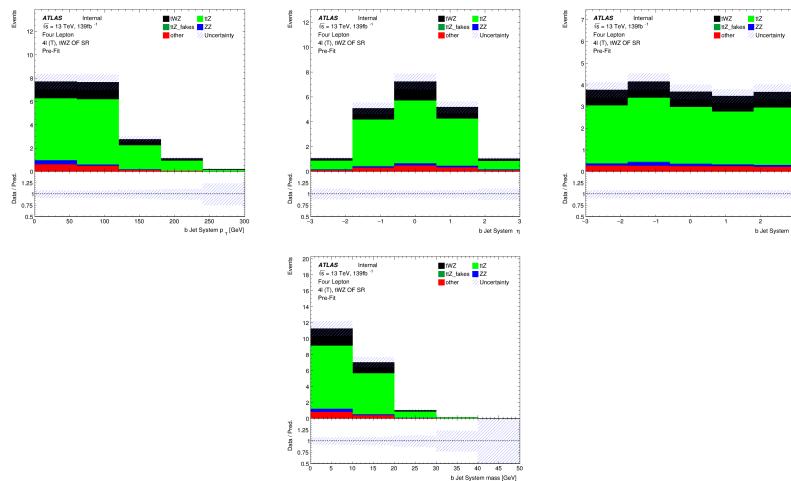


Figure 83: MC predictions for  $p_T$ ,  $\eta$ ,  $\phi$  (top row) and mass (bottom row) of the b-tagged jet systems in the  $tWZ$  SR region (*blinded*)

Not reviewed, for internal circulation only

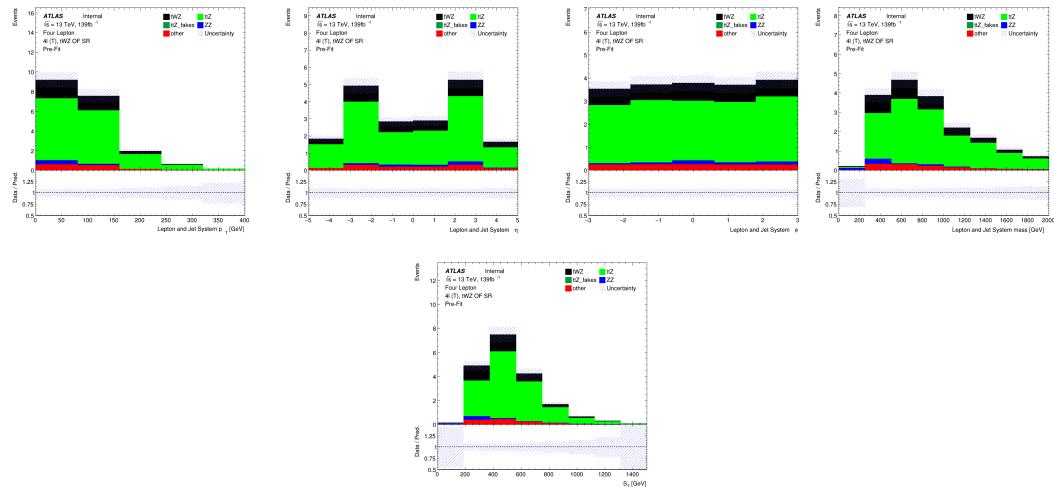


Figure 84: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  (top row) and mass (bottom left) for the lepton + jet systems ( $\ell\ell\ell\ell + \text{jets}$ ) in the  $t\text{WZ OF SR}$  region (*blinded*). Bottom right: MC predictions for  $S_T$  (scalar sum of jet  $p_T$  and lepton  $p_T$ ) in the  $t\text{WZ OF SR}$  region (*blinded*)

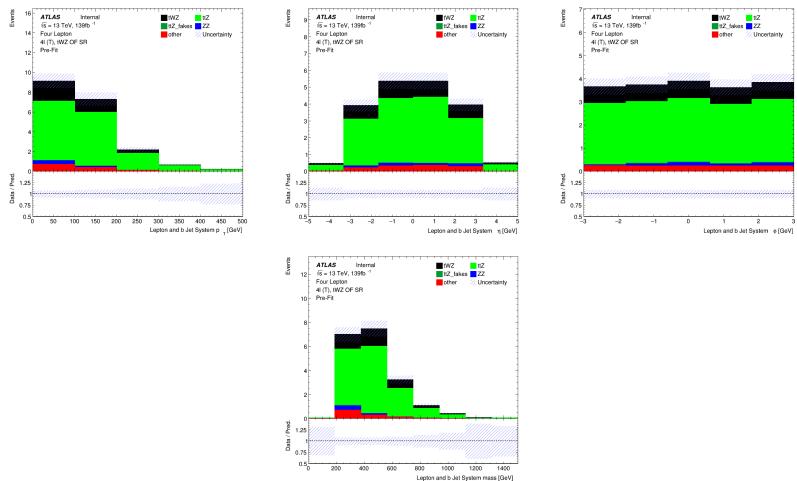


Figure 85: MC predictions for  $p_T$ ,  $\eta$ ,  $\phi$  (top row) and mass (bottom row) of the lepton + b-tagged jet systems ( $\ell\ell\ell\ell + \text{b-tagged jets}$ ) in the  $t\text{WZ OF SR}$  region (*blinded*)

Not reviewed, for internal circulation only

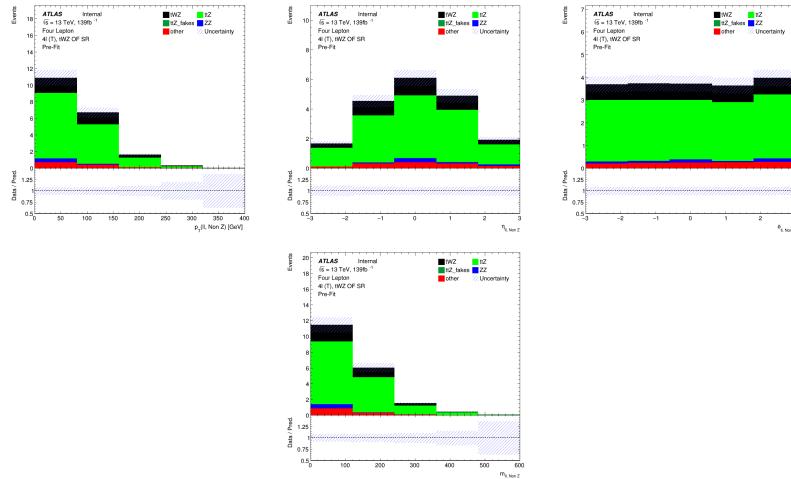


Figure 86: MC predictions for  $p_T$ ,  $\eta$ ,  $\phi$  (top row) and mass (bottom row) of reconstructed Non Z leptons (lepton pairs which don't originate from a Z candidate) in the  $tWZ$  OF SR region (*blinded*)

## 1360 A.2 $tWZ$ SF SR

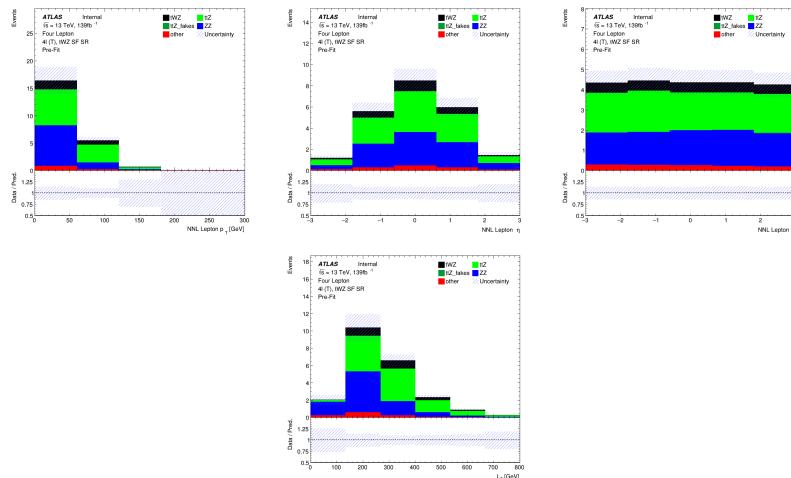


Figure 87: Top row: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  for next-to-next-to-leading (NNL) leptons in the  $tWZ$  SF SR region (*blinded*). Bottom row: MC predictions for  $L_T$  (scalar sum of lepton  $p_T$ ) in the  $tWZ$  SF SR region (*blinded*)

Not reviewed, for internal circulation only

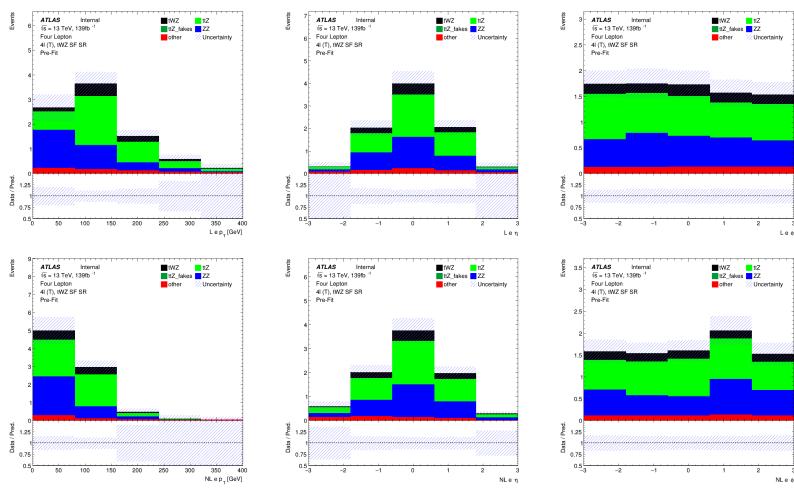


Figure 88: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  for leading (L) electrons (top row) and next-to-leading (NL) electrons (bottom row) in the  $tWZ$  SF SR region (*blinded*)

Not reviewed, for internal circulation only

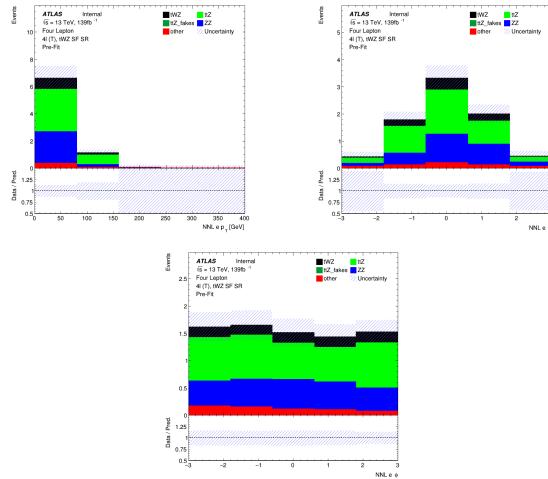


Figure 89: MC predictions for  $p_T$ ,  $\eta$  (top row) and  $\phi$  (bottom row) for next-to-next-to-leading (NNL) electrons in the  $t\text{WZ}$  SF SR region

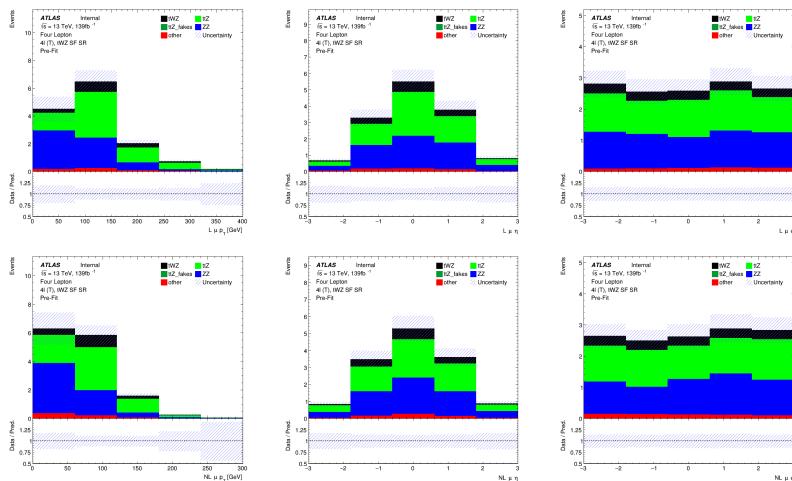


Figure 90: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  for leading (L) muons (top row) and next-to-leading (NL) muons (bottom row) in the  $t\text{WZ}$  SF SR region (*blinded*)

Not reviewed, for internal circulation only

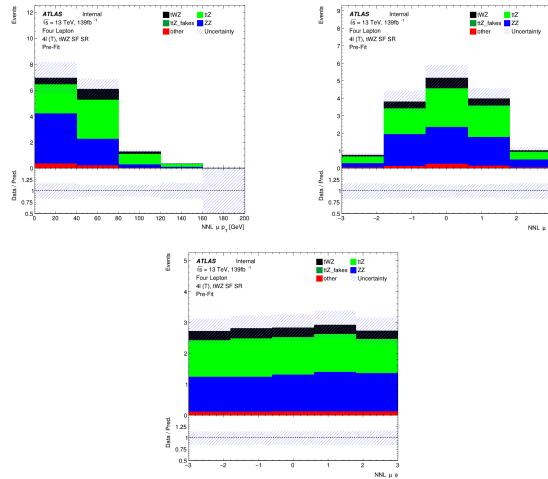


Figure 91: MC predictions for  $p_T$ ,  $\eta$  (top row) and  $\phi$  (bottom row) for next-to-next-to-leading (NNL) muons in the  $tWZ$  SF SR region (*blinded*)

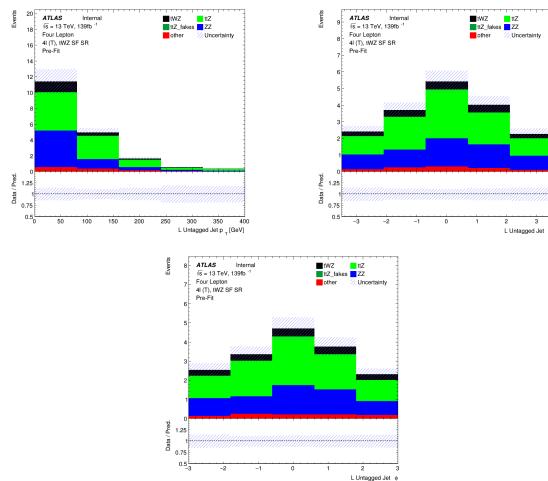


Figure 92: MC predictions for  $p_T$ ,  $\eta$  (top row) and  $\phi$  (bottom row) for untagged jets in the  $tWZ$  SF SR region (*blinded*)

Not reviewed, for internal circulation only

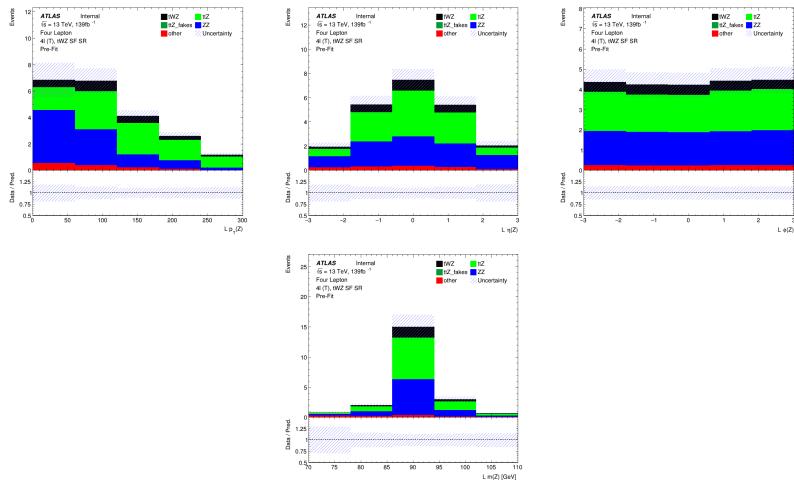


Figure 93: MC predictions for  $p_T$ ,  $\eta$ ,  $\phi$  (top row) and mass ( $m_Z$ ) (bottom row) of the leading reconstructed  $Z$  candidate (OSSF lepton pair with  $|m_{\text{OSSF}} - m_Z| < 30 \text{ GeV}$ ) in the  $tWZ$  SF SR region (*blinded*)

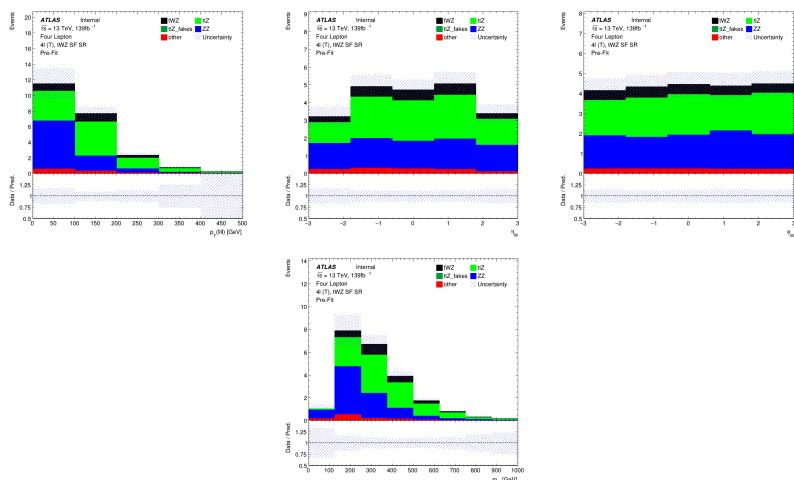


Figure 94: MC predictions for  $p_T$ ,  $\eta$ ,  $\phi$  (top row) and mass (bottom row) of the lepton system ( $\ell\ell\ell\ell$ ) in the  $tWZ$  SF SR region (*blinded*)

Not reviewed, for internal circulation only

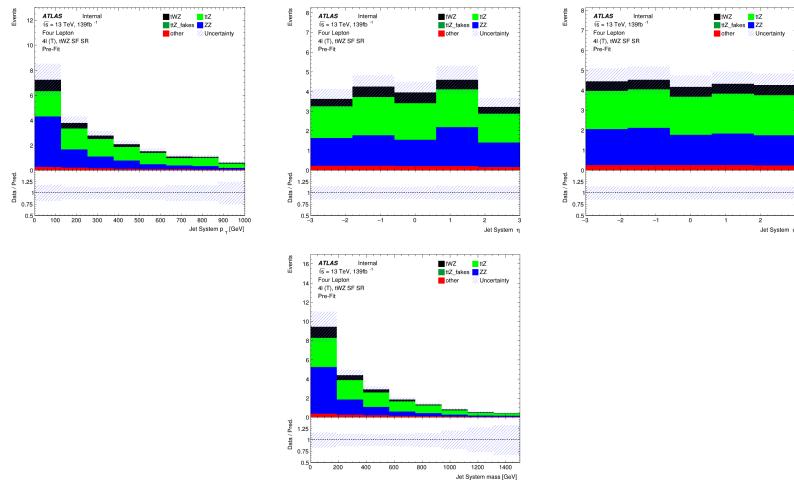


Figure 95: MC predictions for  $p_T$ ,  $\eta$ ,  $\phi$  (top row) and mass (bottom row) of the jet systems in the  $tWZ$  SF SR region (*blinded*)

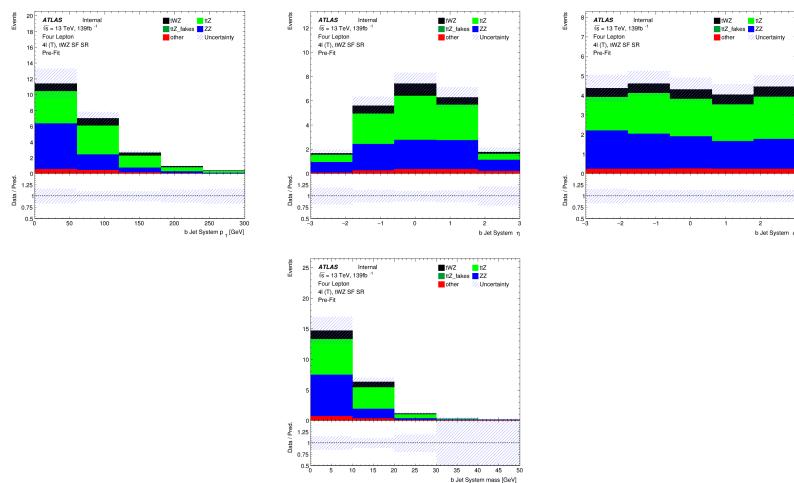


Figure 96: MC predictions for  $p_T$ ,  $\eta$ ,  $\phi$  (top row) and mass (bottom row) of the b-tagged jet systems in the  $tWZ$  SF SR region (*blinded*)

Not reviewed, for internal circulation only

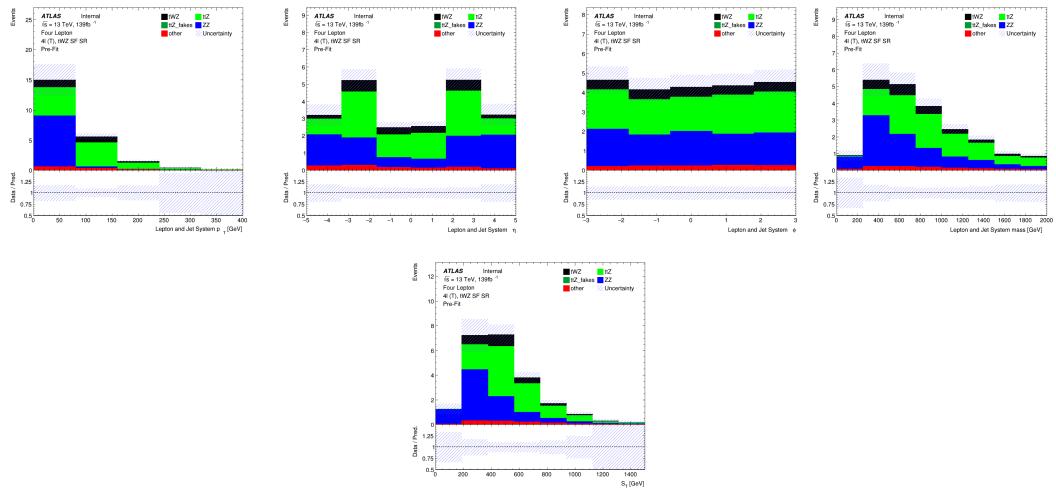


Figure 97: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  (top row) and mass (bottom left) for the lepton + jet systems ( $\ell\ell\ell\ell + \text{jets}$ ) in the  $t\text{WZ}$  SF SR region (*blinded*). Bottom right: MC predictions for  $S_T$  (scalar sum of jet  $p_T$  and lepton  $p_T$ ) in the  $t\text{WZ}$  SF SR region (*blinded*)

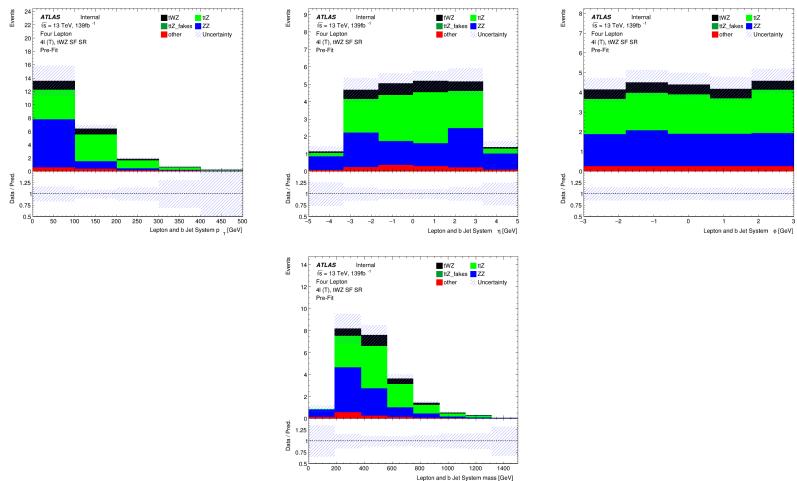


Figure 98: MC predictions for  $p_T$ ,  $\eta$ ,  $\phi$  (top row) and mass (bottom row) of the lepton + b-tagged jet systems ( $\ell\ell\ell\ell + \text{b-tagged jets}$ ) in the  $t\text{WZ}$  SF SR region (*blinded*)

Not reviewed, for internal circulation only

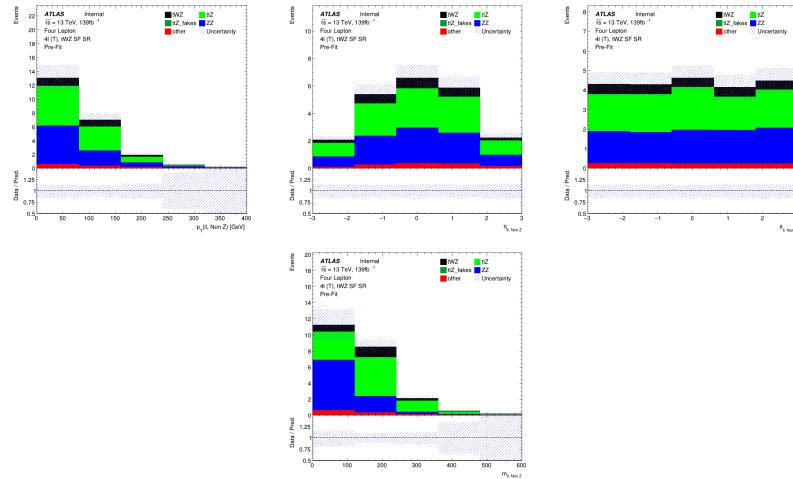


Figure 99: MC predictions for  $p_T$ ,  $\eta$ ,  $\phi$  (top row) and mass (bottom row) of reconstructed Non Z leptons (lepton pairs which don't originate from a Z candidate) in the  $tWZ$  SF SR region (*blinded*)

### 1361 A.3 $t\bar{t}Z$ CR

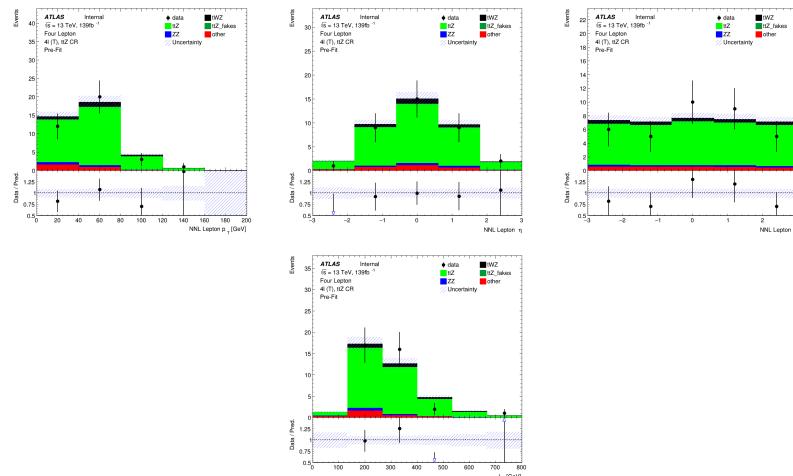


Figure 100: Top row: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  for next-to-next-to-leading (NNL) leptons in the  $t\bar{t}Z$  CR region . Bottom row: MC predictions for  $L_T$  (scalar sum of lepton  $p_T$ ) in the  $t\bar{t}Z$  CR region

Not reviewed, for internal circulation only

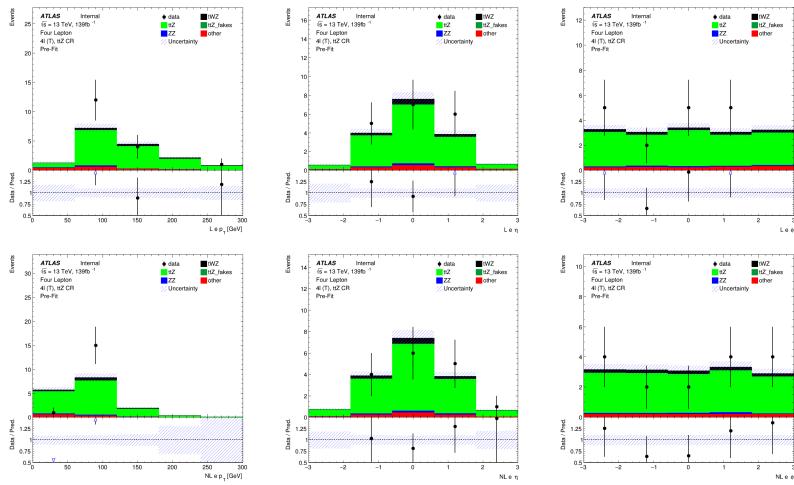


Figure 101: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  for leading (L) electrons (top row) and next-to-leading (NL) electrons (bottom row) in the  $t\bar{t}Z$  CR region

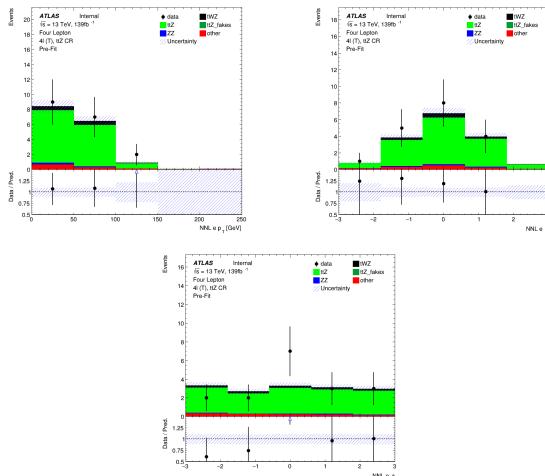


Figure 102: MC predictions for  $p_T$ ,  $\eta$  (top row) and  $\phi$  (bottom row) for next-to-next-to-leading (NNL) electrons in the  $t\bar{t}Z$  CR region

Not reviewed, for internal circulation only

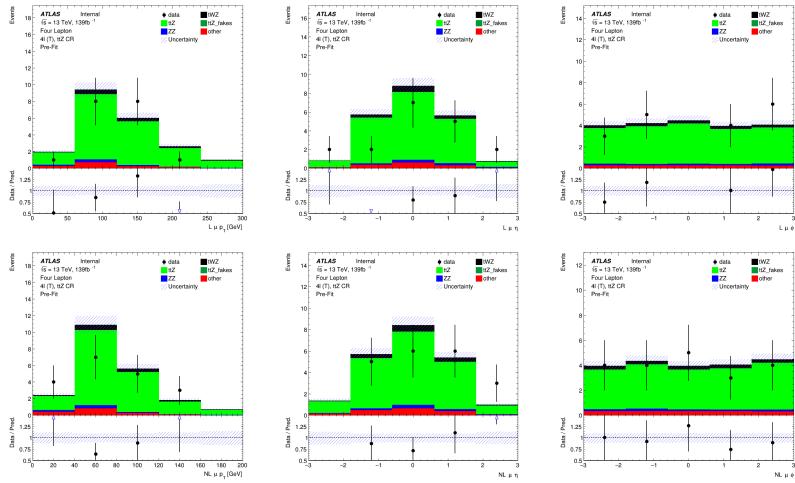


Figure 103: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  for leading (L) muons (top row) and next-to-leading (NL) muons (bottom row) in the  $t\bar{t}Z$  CR region

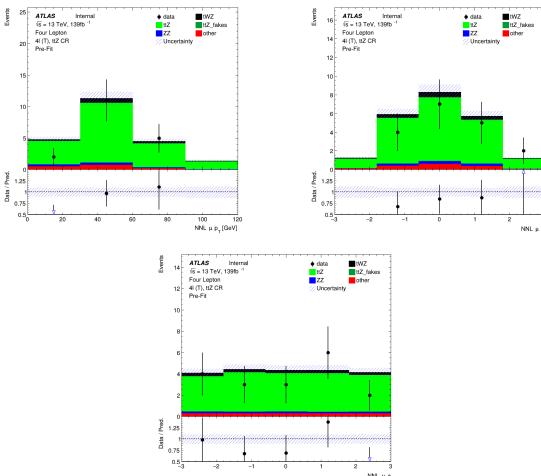


Figure 104: MC predictions for  $p_T$ ,  $\eta$  (top row) and  $\phi$  (bottom row) for next-to-next-to-leading (NNL) muons in the  $t\bar{t}Z$  CR region

Not reviewed, for internal circulation only

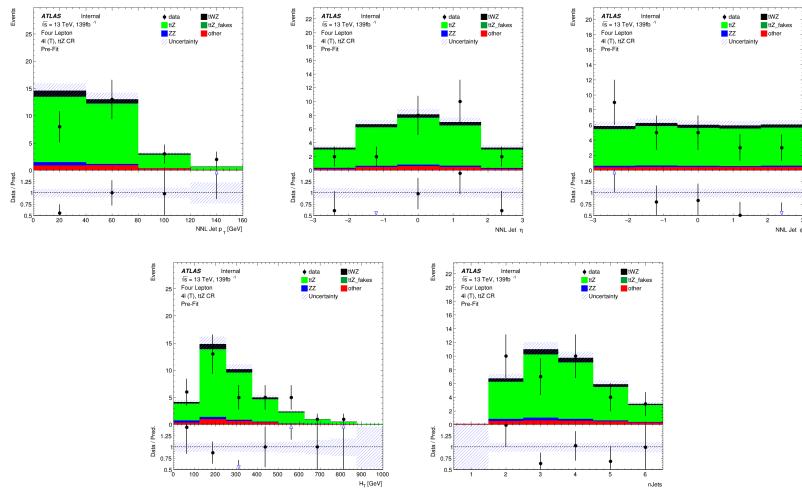


Figure 105: Top row: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  for next-to-next-to-leading (NNL) jets in the  $t\bar{t}Z$  CR region .  
Bottom row: MC predictions for  $H_T$  (scalar sum of Jet  $p_T$ ) (left) and the Number of jets (right) in the  $t\bar{t}Z$  CR region

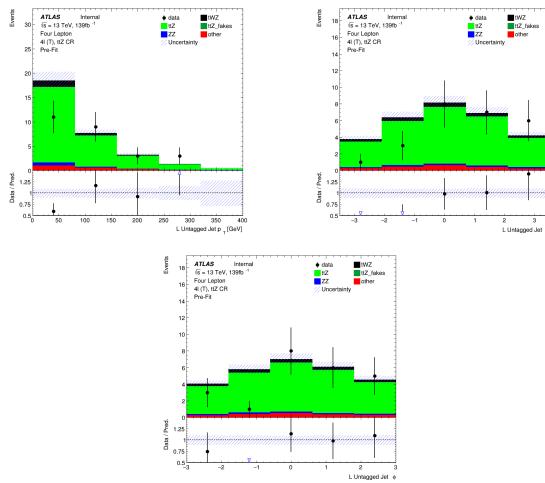


Figure 106: MC predictions for  $p_T$ ,  $\eta$  (top row) and  $\phi$  (bottom row) for untagged jets in the  $t\bar{t}Z$  CR region

Not reviewed, for internal circulation only

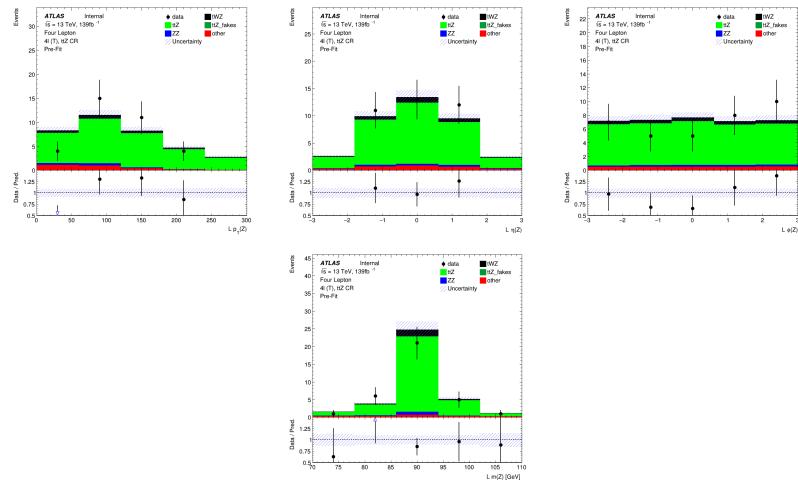


Figure 107: MC predictions for  $p_T$ ,  $\eta$ ,  $\phi$  (top row) and mass ( $m_Z$ ) (bottom row) of reconstructed leading  $Z$  candidates (OSSF lepton pair with  $|m_{\text{OSSF}} - m_Z| < 30 \text{ GeV}$ ) in the  $t\bar{t}Z$  CR region

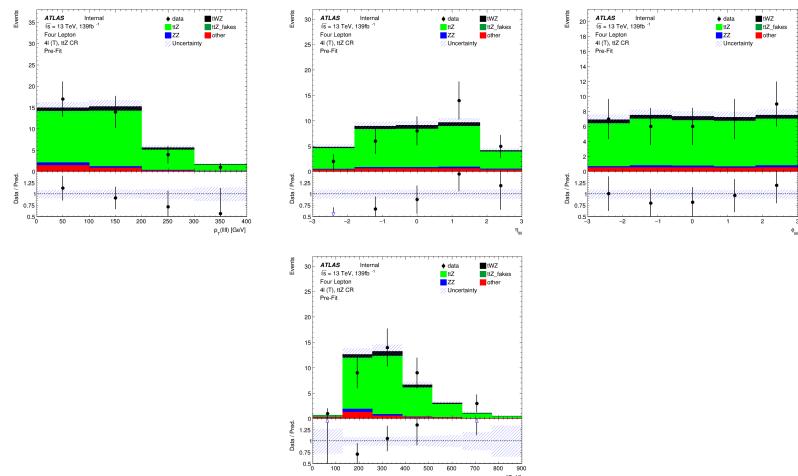


Figure 108: MC predictions for  $p_T$ ,  $\eta$ ,  $\phi$  (top row) and mass (bottom row) of the lepton system ( $\ell\ell\ell\ell$ ) in the  $t\bar{t}Z$  CR region

Not reviewed, for internal circulation only

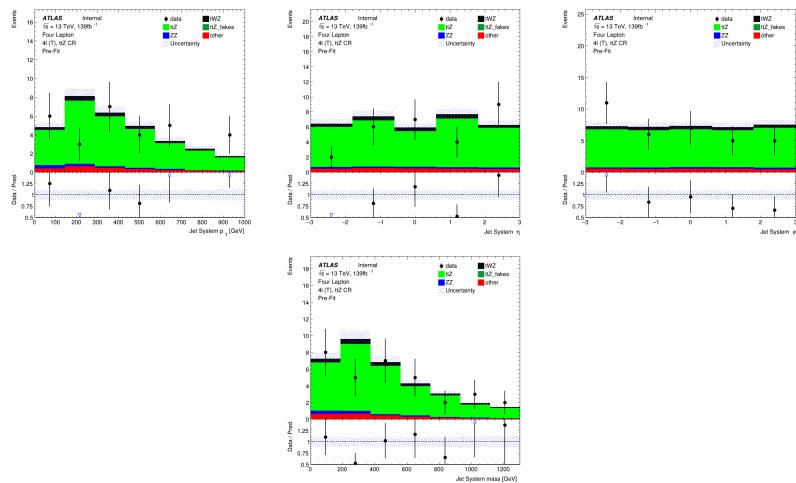


Figure 109: MC predictions for  $p_T$ ,  $\eta$ ,  $\phi$  (top row) and mass (bottom row) of the jet systems in the  $t\bar{t}Z$  CR region

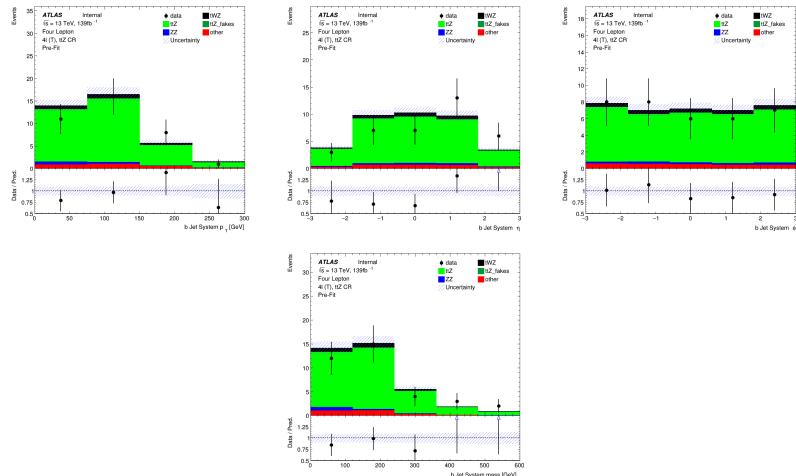


Figure 110: MC predictions for  $p_T$ ,  $\eta$ ,  $\phi$  (top row) and mass (bottom row) of the b-tagged jet systems in the  $t\bar{t}Z$  CR region

Not reviewed, for internal circulation only

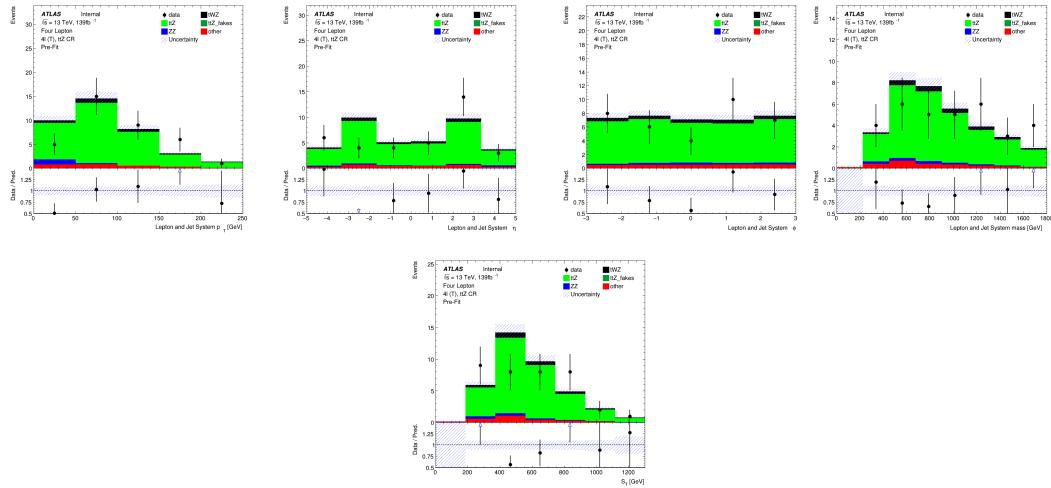


Figure 111: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  (top row) and mass (bottom left) for the lepton + jet systems ( $\ell\ell\ell\ell + \text{jets}$ ) in the  $t\bar{t}Z$  CR region . Bottom right: MC predictions for  $S_T$  (scalar sum of jet  $p_T$  and lepton  $p_T$ ) in the  $t\bar{t}Z$  CR region

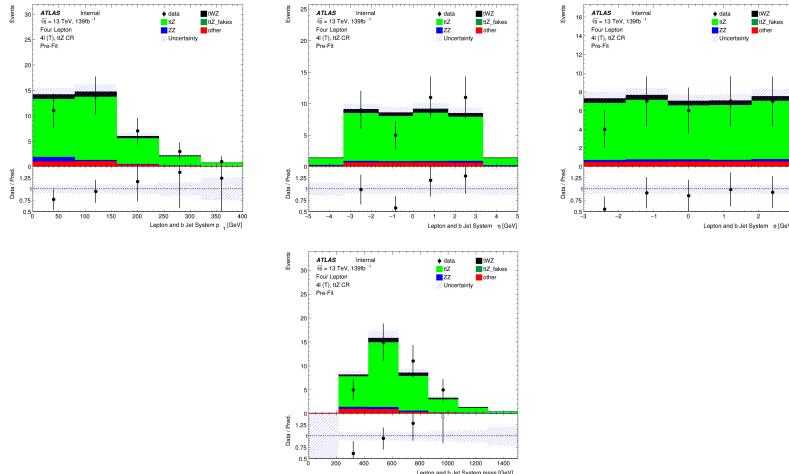


Figure 112: MC predictions for  $p_T$ ,  $\eta$ ,  $\phi$  (top row) and mass (bottom row) of the lepton + b-tagged jet systems ( $\ell\ell\ell\ell + \text{b-tagged jets}$ ) in the  $t\bar{t}Z$  CR region

Not reviewed, for internal circulation only

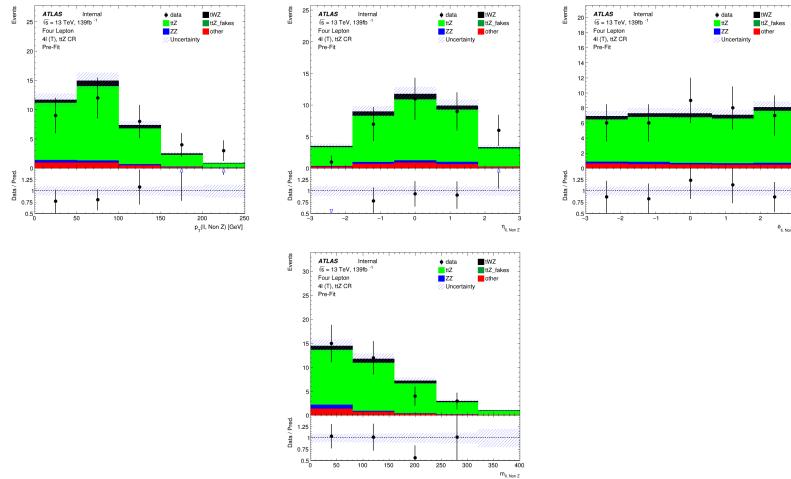


Figure 113: MC predictions for  $p_T$ ,  $\eta$ ,  $\phi$  (top row) and mass (bottom row) of reconstructed Non Z leptons (lepton pairs which don't originate from a  $Z$  candidate) in the  $t\bar{t}Z$  CR region

#### 1362 A.4 ZZb CR

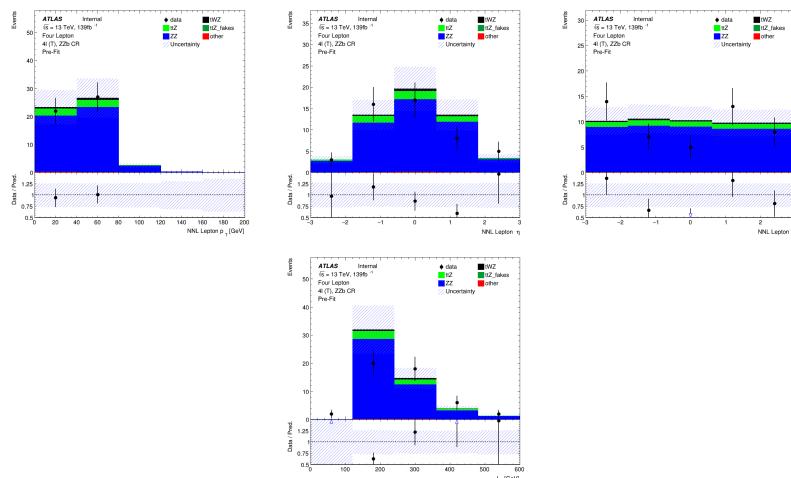


Figure 114: Top row: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  for next-to-next-to-leading (NNL) leptons in the  $ZZb$  CR region . Bottom row: MC predictions for  $L_T$  (scalar sum of lepton  $p_T$ ) in the  $ZZb$  CR region

Not reviewed, for internal circulation only

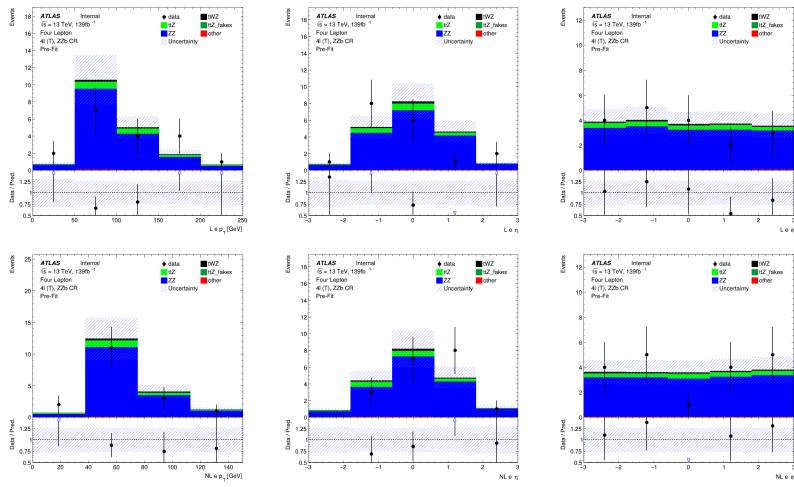


Figure 115: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  for leading (L) electrons (top row) and next-to-leading (NL) electrons (bottom row) in the  $ZZb$  CR region

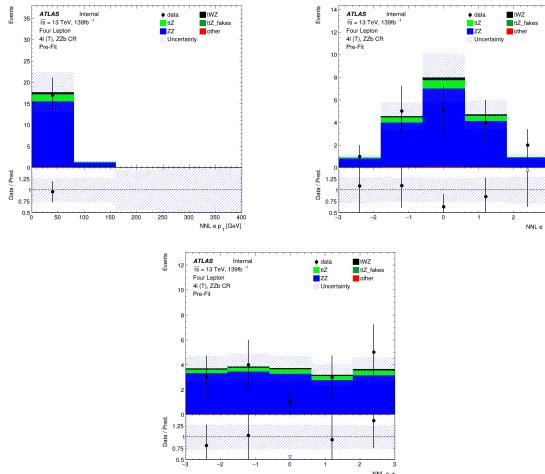


Figure 116: MC predictions for  $p_T$ ,  $\eta$  (top row) and  $\phi$  (bottom row) for next-to-next-to-leading (NNL) electrons in the  $ZZb$  CR region

Not reviewed, for internal circulation only

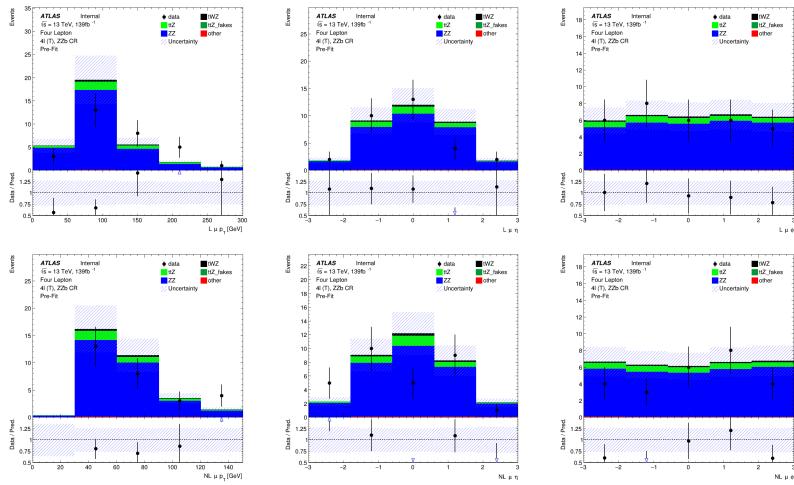


Figure 117: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  for leading (L) muons (top row) and next-to-leading (NL) muons (bottom row) in the  $ZZb$  CR region

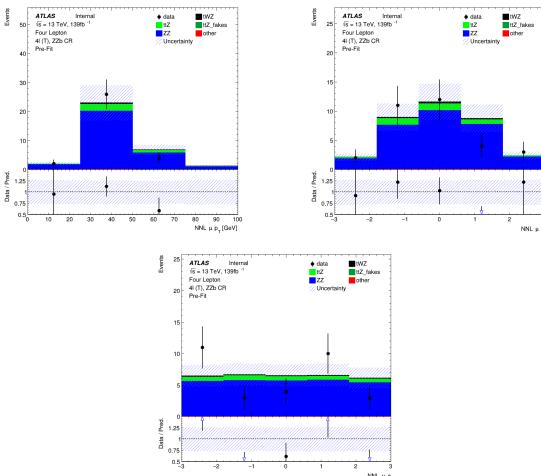


Figure 118: MC predictions for  $p_T$ ,  $\eta$  (top row) and  $\phi$  (bottom row) for next-to-next-to-leading (NNL) muons in the  $ZZb$  CR region

Not reviewed, for internal circulation only

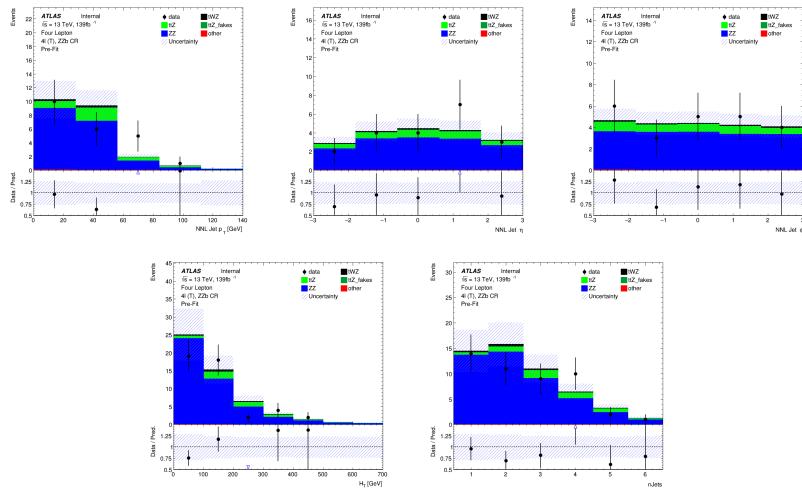


Figure 119: Top row: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  for next-to-next-to-leading (NNL) jets in the  $ZZb$  CR region . Bottom row: MC predictions for  $H_T$  (scalar sum of Jet  $p_T$ ) (left) and the Number of jets (right) in the  $ZZb$  CR region

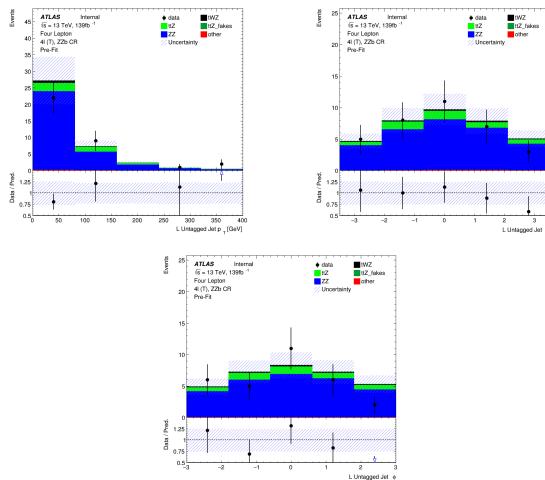


Figure 120: MC predictions for  $p_T$ ,  $\eta$  (top row) and  $\phi$  (bottom row) for untagged jets in the  $ZZb$  CR region

Not reviewed, for internal circulation only

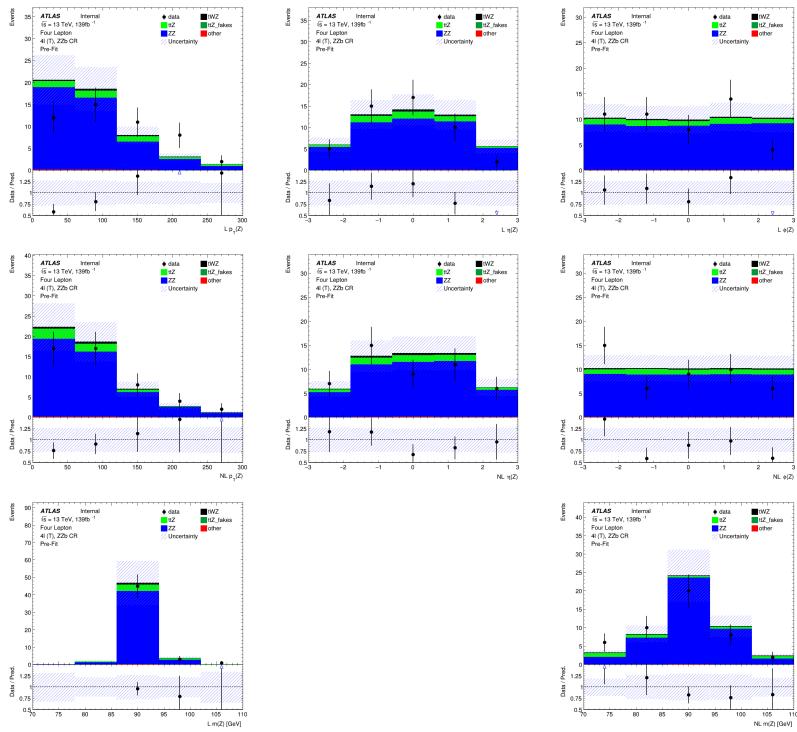


Figure 121: MC predictions for  $p_T$ ,  $\eta$ ,  $\phi$  of the leading (top row), next-to-leading (middle row) and mass ( $m_Z$ ) (bottom row) of reconstructed Z candidates (OSSF lepton pair with  $|m_{\text{OSSF}} - m_Z| < 30 \text{ GeV}$ ) in the  $ZZb$  CR region

Not reviewed, for internal circulation only

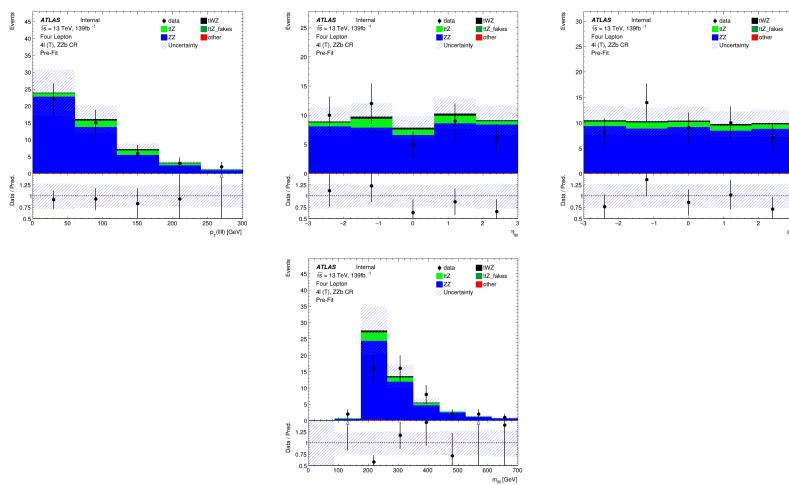


Figure 122: MC predictions for  $p_T$ ,  $\eta$ ,  $\phi$  (top row) and mass (bottom row) of the lepton system ( $\ell\ell\ell\ell$ ) in the  $ZZb$  CR region

Not reviewed, for internal circulation only

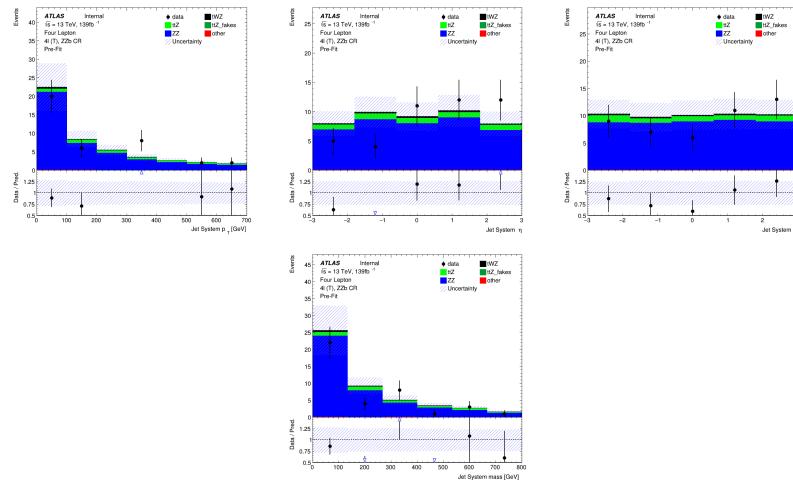


Figure 123: MC predictions for  $p_T$ ,  $\eta$ ,  $\phi$  (top row) and mass (bottom row) of the jet systems in the  $ZZb$  CR region

Not reviewed, for internal circulation only

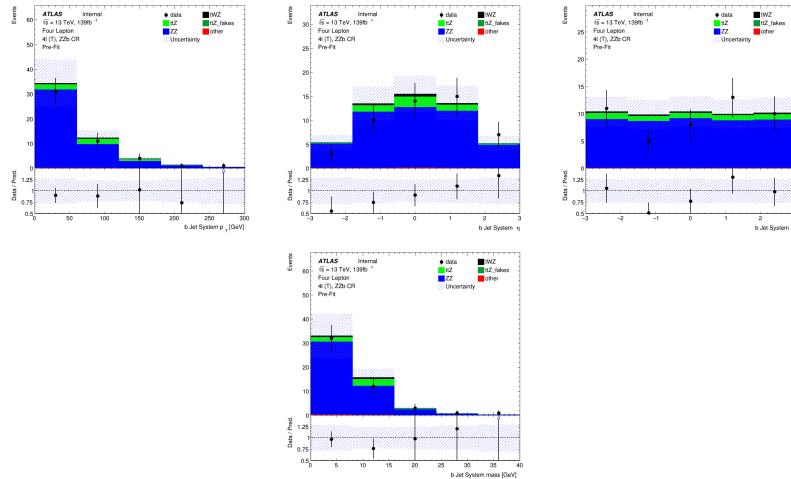


Figure 124: MC predictions for  $p_T$ ,  $\eta$ ,  $\phi$  (top row) and mass (bottom row) of the b-tagged jet systems in the  $ZZb$  CR region

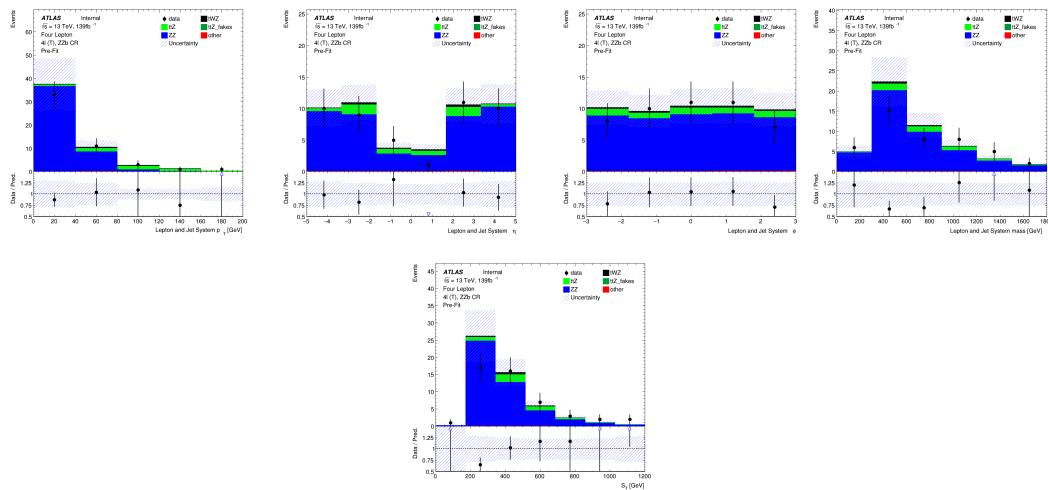


Figure 125: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  (top row) and mass (bottom left) for the lepton + jet systems ( $llll + jets$ ) in the  $ZZb$  CR region . Bottom right: MC predictions for  $S_T$  (scalar sum of jet  $p_T$  and lepton  $p_T$ ) in the  $ZZb$  CR region

Not reviewed, for internal circulation only

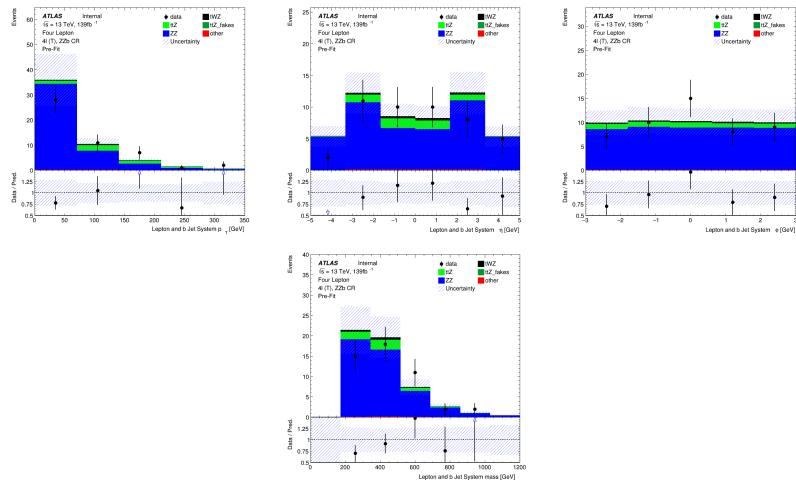


Figure 126: MC predictions for  $p_T$ ,  $\eta$ ,  $\phi$  (top row) and mass (bottom row) of the lepton + b-tagged jet systems ( $\ell\ell\ell\ell + b$ -tagged jets) in the  $ZZb$  CR region

### 1363 A.5 $(tWZ)_{\text{fake}}$ CR

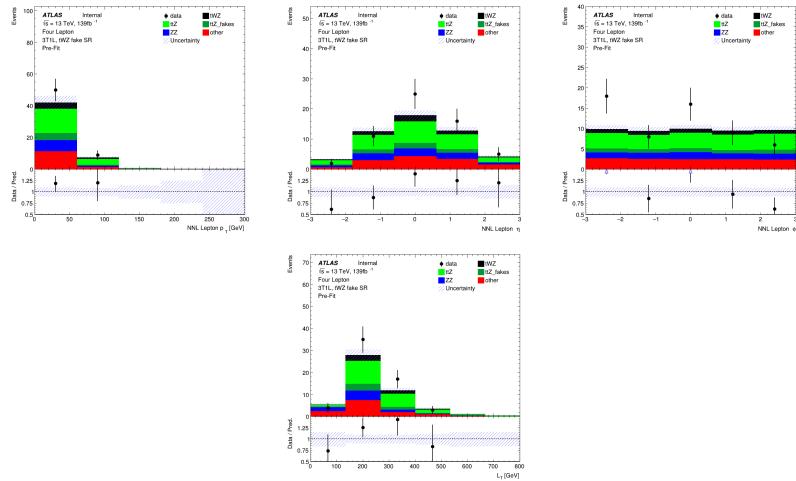


Figure 127: Top row: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  for next-to-next-to-leading (NNL) leptons in the  $(tWZ)_{\text{fake}}$  CR region . Bottom row: MC predictions for  $L_T$  (scalar sum of lepton  $p_T$ ) in the  $(tWZ)_{\text{fake}}$  CR region

Not reviewed, for internal circulation only

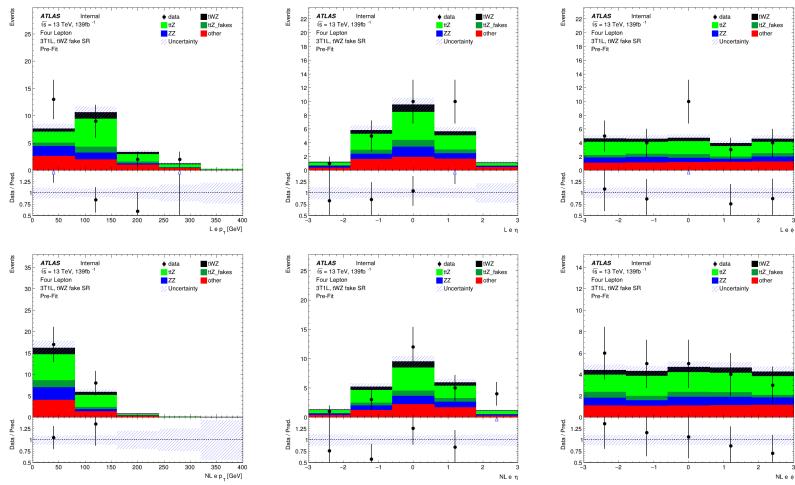


Figure 128: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  for leading (L) electrons (top row) and next-to-leading (NL) electrons (bottom row) in the  $(tWZ)_{\text{fake}}$  CR region

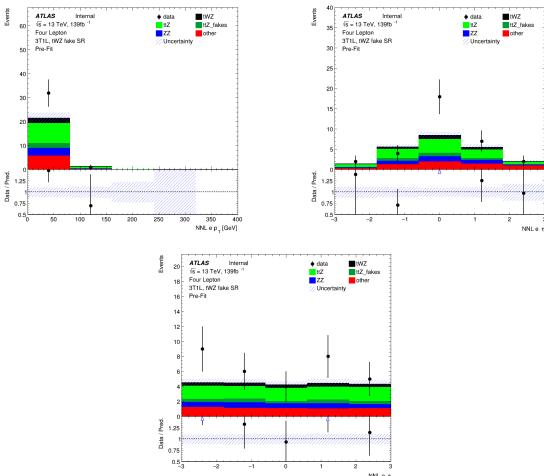


Figure 129: MC predictions for  $p_T$ ,  $\eta$  (top row) and  $\phi$  (bottom row) for next-to-next-to-leading (NNL) electrons in the  $(tWZ)_{\text{fake}}$  CR region

Not reviewed, for internal circulation only

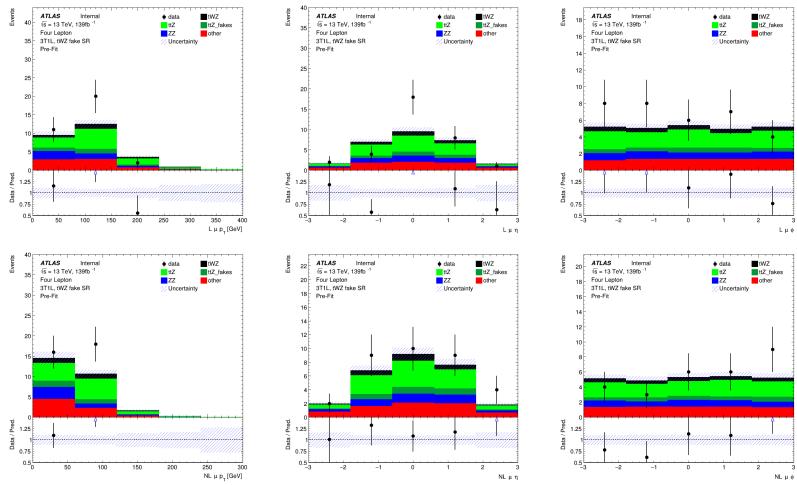


Figure 130: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  for leading (L) muons (top row) and next-to-leading (NL) muons (bottom row) in the  $(tWZ)_{\text{fake}}$  CR region

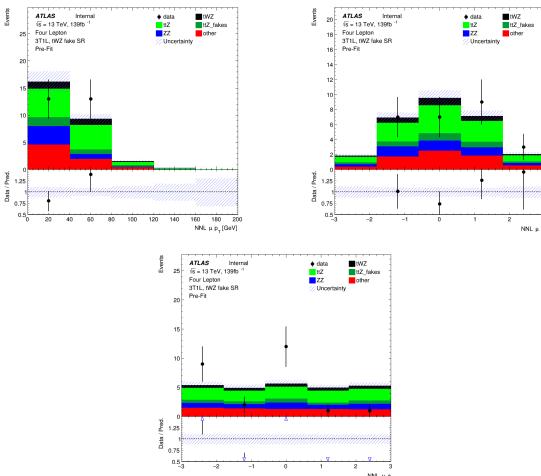


Figure 131: MC predictions for  $p_T$ ,  $\eta$  (top row) and  $\phi$  (bottom row) for next-to-next-to-leading (NNL) muons in the  $(tWZ)_{\text{fake}}$  CR region

Not reviewed, for internal circulation only

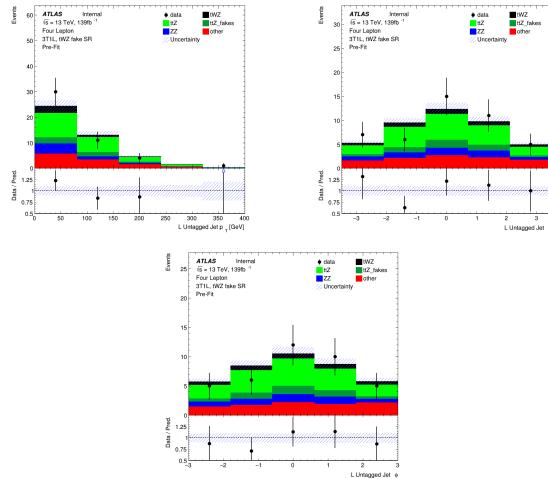


Figure 132: MC predictions for  $p_T$ ,  $\eta$  (top row) and  $\phi$  (bottom row) for untagged jets in the  $(tWZ)_{\text{fake}}$  CR region

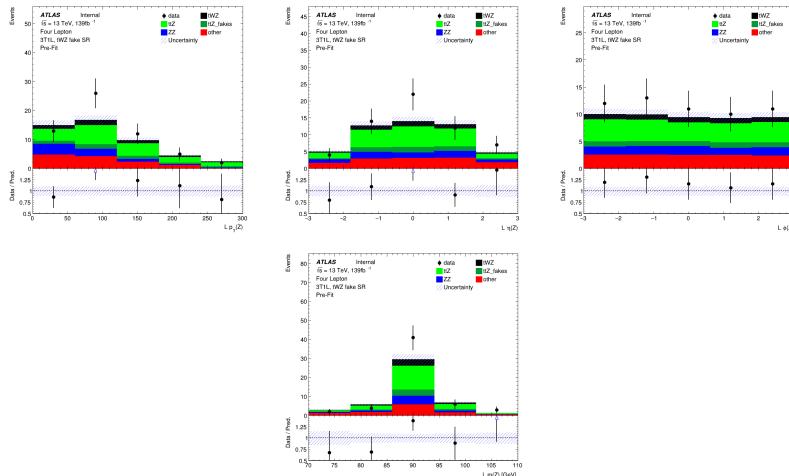


Figure 133: MC predictions for  $p_T$ ,  $\eta$ ,  $\phi$  (top row) and mass ( $m_Z$ ) (bottom row) of the leading reconstructed  $Z$  candidate (OSSF lepton pair with  $|m_{\text{OSSF}} - m_Z| < 30$  GeV) in the  $(tWZ)_{\text{fake}}$  CR region

Not reviewed, for internal circulation only

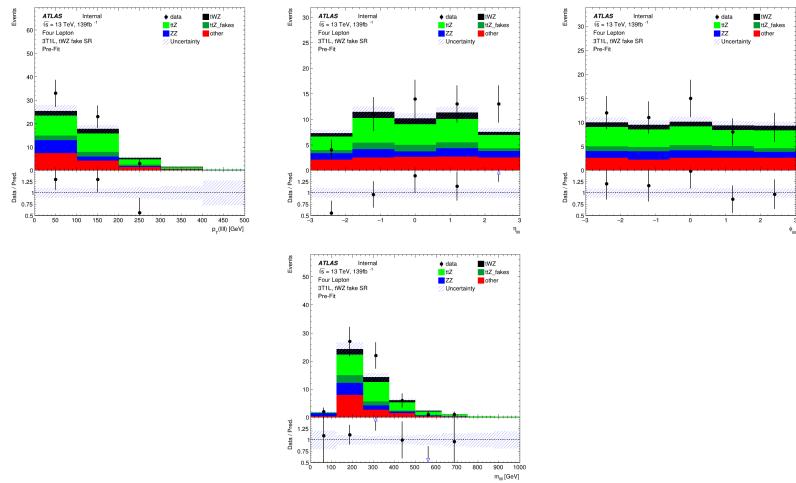


Figure 134: MC predictions for  $p_T$ ,  $\eta$ ,  $\phi$  (top row) and mass (bottom row) of the lepton system ( $\ell\ell\ell\ell$ ) in the  $(tWZ)_{\text{fake}}$  CR region

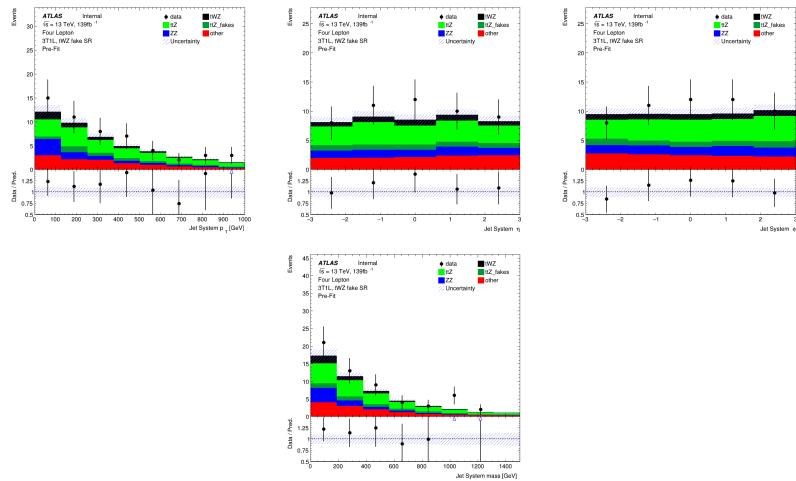


Figure 135: MC predictions for  $p_T$ ,  $\eta$ ,  $\phi$  (top row) and mass (bottom row) of the jet systems in the  $(tWZ)_{\text{fake}}$  CR region

Not reviewed, for internal circulation only

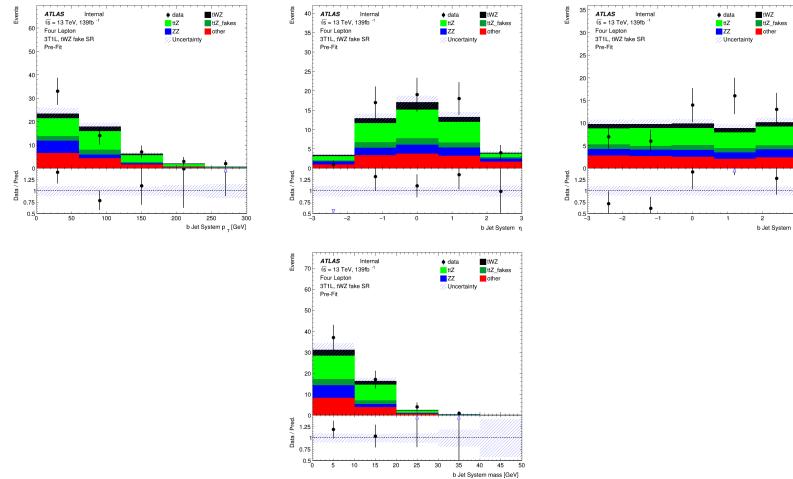


Figure 136: MC predictions for  $p_T$ ,  $\eta$ ,  $\phi$  (top row) and mass (bottom row) of the b-tagged jet systems in the  $(tWZ)_{\text{fake}}$  CR region

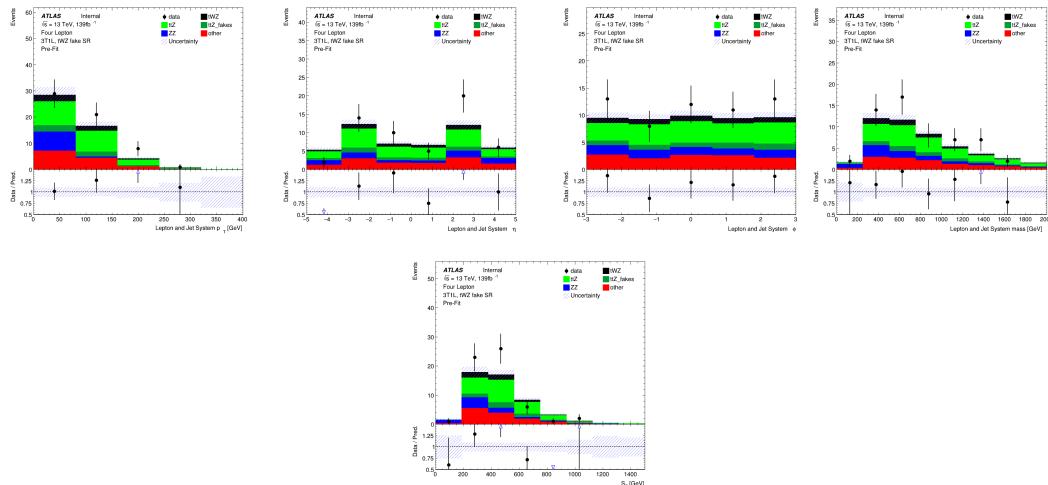


Figure 137: MC predictions for  $p_T$ ,  $\eta$  and  $\phi$  (top row) and mass (bottom left) for the lepton + jet systems ( $\ell\ell\ell\ell +$  jets) in the  $(tWZ)_{\text{fake}}$  CR region . Bottom right: MC predictions for  $S_T$  (scalar sum of jet  $p_T$  and lepton  $p_T$ ) in the  $(tWZ)_{\text{fake}}$  CR region

Not reviewed, for internal circulation only

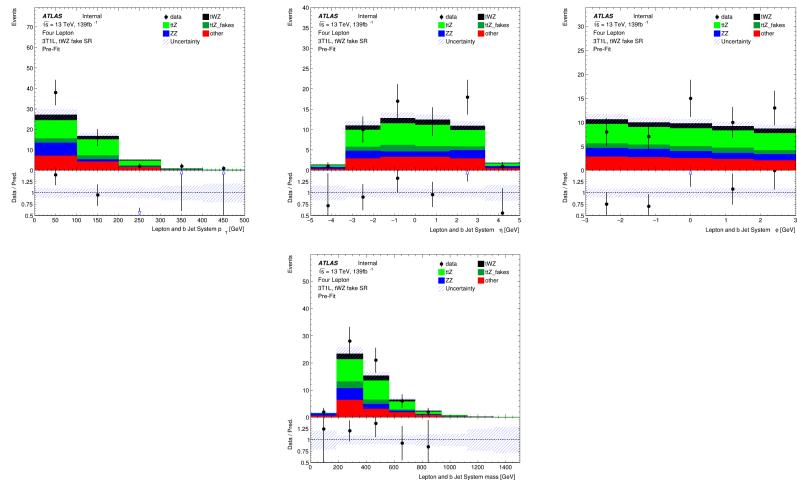


Figure 138: MC predictions for  $p_T$ ,  $\eta$ ,  $\phi$  (top row) and mass (bottom row) of the lepton + b-tagged jet systems ( $\ell\ell\ell\ell + b$ -tagged jets) in the  $(tWZ)_{\text{fake}}$  CR region

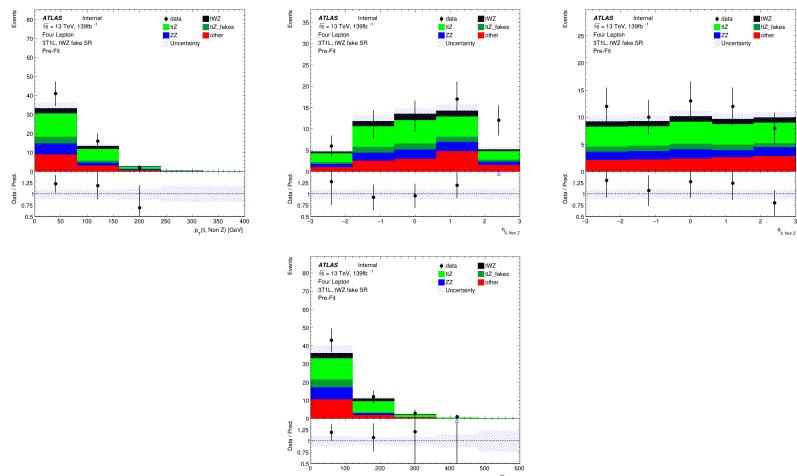


Figure 139: MC predictions for  $p_T$ ,  $\eta$ ,  $\phi$  (top row) and mass (bottom row) of reconstructed Non  $Z$  leptons (lepton pairs which don't originate from a  $Z$  candidate) in the  $(tWZ)_{\text{fake}}$  CR region

## 1364 B TRexFitter Configuration

1365 In this section, an outline of the various settings chosen in the TRexFitter config files [66] (used for  
 1366 plotting and fitting), in the trilepton channel, tetralepton channel and the combined analysis, are given.  
 1367 Settings which only affect cosmetics are not shown, for more information see the [TRexFitter settings](#)  
 1368 documentation.

### 1369 B.1 Trilepton Channel

### 1370 B.2 Tetralepton Channel

- 1371 • MC event weights

1372 The ingredients included in the MC event weights are applied to each MC event multiplicatively.

$$1373 - \text{normWeight} = \frac{\sigma \mathcal{L} w_{MC\text{Gen}}}{\sum w_{MC\text{Gen}}},$$

1374 where  $\sigma$  is the cross section of the process of interest,  $\mathcal{L}$  is the integrated luminosity,  $w_{MC\text{Gen}}$   
 1375 is the weight assigned to the the event by the MC event generator and  $\sum w_{MC\text{Gen}}$  is the sum of  
 1376 these event weights for all generated events

- 1377 – `weight_leptonSF` → electron and muon efficiency scale factors
- 1378 – `weight_pileup` → detector pile up
- 1379 – `weight_jvt` → jet vertex tagger for PFlow jets
- 1380 – `weight_bTagSF_DL1r_77` → DL1r b-tagger scale factors

- 1381 • `MCstatConstraint`: "Poisson", constraint used for MC stat uncertainties

- 1382 • `MCstatThreshold`: **0.001**, a nuisance parameter will be added for each bin with a MC statistical  
 1383 uncertainty greater than `MCstatThreshold`

- 1384 • `SystPruningShape`: **0.01**, lower threshold to remove a shape systematic from the fit/limit  
 1385 (supression is done per sample and per region)

- 1386 • `SystPruningNorm`: **0.01**, lower threshold to remove a normalisation systematic from the fit/limit  
 1387 (supression is done per sample and per region)

- 1388 • `FitType`: SPLUSB, fit under the signal background hypothesis

- 1389 • `POIAsimov`: 1, value of the parameter of interest ( $\mu$ ) in the AsimovDataset used in the fit (used in  
 1390 blinded fits)

- 1391 • `LimitType`: ASYMPTOTIC, asymptotic formulae [67] used in the limit calculation

1392 Normalisation Factors:

$$1393 - \text{"SigXsecOverSM"} \rightarrow \mu_{tWZ} = \frac{\sigma_{\text{measured}}(tWZ)}{\sigma_{SM}(tWZ)}$$

- 1394 – `Nominal`: 1, nominal value
- 1395 – `Min`: **-100**, minimum value
- 1396 – `Max`: **100**, maximum value



- 1428           – Samples: tzq, samples to apply the systematic
- 1429       • "tt cross section"
- 1430           – Type: OVERALL, HistFactory shape systematic
- 1431           – OverallUp:  $0.20$ , the relative "up" shift (+3%)
- 1432           – OverallDown:  $-0.20$ , the relative "down" shift (-3%)
- 1433           – Samples: tt, samples to apply the systematic
- 1434       • "tth cross section"
- 1435           – Type: OVERALL, HistFactory shape systematic
- 1436           – OverallUp:  $0.20$ , the relative "up" shift (+3%)
- 1437           – OverallDown:  $-0.20$ , the relative "down" shift (-3%)
- 1438           – Samples: tth, samples to apply the systematic
- 1439       • "vv cross section"
- 1440           – Type: OVERALL, HistFactory shape systematic
- 1441           – OverallUp:  $0.20$ , the relative "up" shift (+3%)
- 1442           – OverallDown:  $-0.20$ , the relative "down" shift (-3%)
- 1443           – Samples: vv111, samples to apply the systematic
- 1444       • "Lumi" → Luminosity
- 1445           – Type: OVERALL, HistFactory shape systematic
- 1446           – OverallUp:  $0.03$ , the relative "up" shift (+3%)
- 1447           – OverallDown:  $-0.03$ , the relative "down" shift (-3%)
- 1448           – Samples: all, samples to apply the systematic
- 1449       • "LepSF\_EL\_Trigger" → Electron trigger efficiency
- 1450           – Type: HISTO, HistFactory shape systematic
- 1451           – Symmetrisation: TWOSIDED,  $\frac{up-down}{2}$  variation is calculated bin-by-bin (this is used as up variation and then mirrored to down variation)
- 1453           – WeightSufUp: "weight\_leptonSF\_EL\_SF\_Trigger\_UP/weight\_leptonSF", additional weight for systematic up variation
- 1455           – WeightSufDown: "weight\_leptonSF\_EL\_SF\_Trigger\_DOWN/weight\_leptonSF", additional weight for systematic down variation
- 1457           – Samples: all, samples to apply the systematic
- 1458       • "jvt" → jet vertex tagger
- 1459           – Type: HISTO, HistFactory shape systematic

Not reviewed, for internal circulation only

- Not reviewed, for internal circulation only
- Symmetrisation: TWOSIDED,  $\frac{up-down}{2}$  variation is calculated bin-by-bin (this is used as up variation and then mirrored to down variation)
  - WeightSufUp: "weight\_jvt\_UP/weight\_jvt", additional weight for systematic up variation
  - WeightSufDown: "weight\_jvt\_DOWN/weight\_jvt", additional weight for systematic down variation
  - Samples: all, samples to apply the systematic
  - "pileup"
    - Type: HISTO, HistFactory shape systematic
    - Symmetrisation: TWOSIDED,  $\frac{up-down}{2}$  variation is calculated bin-by-bin (this is used as up variation and then mirrored to down variation)
    - WeightSufUp: "weight\_pileup\_UP/weight\_pileup", additional weight for systematic up variation
    - WeightSufDown: "weight\_pileup\_DOWN/weight\_pileup", additional weight for systematic down variation
    - Samples: all, samples to apply the systematic
  - "bTagB"
    - Type: HISTO, HistFactory shape systematic
    - Symmetrisation: TWOSIDED,  $\frac{up-down}{2}$  variation is calculated bin-by-bin (this is used as up variation and then mirrored to down variation)
    - WeightSufUp: "weight\_bTagSF\_DL1r\_77\_eigenvars\_B\_up/weight\_bTagSF\_DL1r\_77", additional weight for systematic up variation
    - WeightSufDown: "weight\_bTagSF\_DL1r\_77\_eigenvars\_B\_down/weight\_bTagSF\_DL1r\_77", additional weight for systematic down variation
    - Samples: all, samples to apply the systematic
  - "bTagL"
    - Type: HISTO, HistFactory shape systematic
    - Symmetrisation: TWOSIDED,  $\frac{up-down}{2}$  variation is calculated bin-by-bin (this is used as up variation and then mirrored to down variation)
    - WeightSufUp: "weight\_bTagSF\_DL1r\_77\_eigenvars\_Light\_up/weight\_bTagSF\_DL1r\_77", additional weight for systematic up variation
    - WeightSufDown: "weight\_bTagSF\_DL1r\_77\_eigenvars\_Light\_down/weight\_bTagSF\_DL1r\_77", additional weight for systematic down variation
    - Samples: all, samples to apply the systematic
  - "tWZ-DR2" → shape systematic for  $tWZ$  using diagram removal scheme 2 (DR2)

- 
- 1495        – Type: HISTO, HistFactory shape systematic  
1496        – Symmetrisation: ONESIDED, one-sided symmetrisation  
1497        – Samples: twz, samples to apply the systematic

1498        Certain MC samples with minimal background contribution to the  $tWZ$  signal are not grouped with the  
1499        'other' sample when defining their cross section pre-fit uncertainties (e.g.  $t\bar{t}$  and  $t\bar{t}h$ ). This is done for  
1500        samples which are included in both the tri and tetralepton channels, such that in the combined fit, there is  
1501        one nuisance parameter in the likelihood function which can vary for the cross section of that process.

1502        **B.3 Combined**