



## ATLAS Note

GROUP-2017-XX

17th February 2021



1

# 2 WZ + Heavy Flavor Production in pp collisions 3 at $\sqrt{s} = 13 \text{ TeV}$

4

The ATLAS Collaboration

5

A measurement of WZ produced with an associated heavy flavor jet is performed using 140  
6  $\text{fb}^{-1}$  of proton-proton collision data at  $\sqrt{s} = 13 \text{ TeV}$  from the ATLAS experiment at the  
7 LHC. The measurement is performed in the fully leptonic decay mode,  $\text{WZ} \rightarrow \ell\bar{\nu}\ell\bar{\nu}$ . The  
8 cross-section of  $\text{WZ} + b\text{-jets}$  is measured to be  $X \pm X \pm X$ , while the cross-section of  $\text{WZ} +$   
9 charm is measured as  $X$ , with a correlation of  $X$  between the two processes.

10

© 2021 CERN for the benefit of the ATLAS Collaboration.

Reproduction of this article or parts of it is allowed as specified in the CC-BY-4.0 license.

# 11 Contents

12	<b>1 Changes and outstanding items</b>	<b>4</b>
13	1.1 Changelog	4
14	1.1.1 Changes relative to v4	4
15	1.1.2 Changes relative to v3	4
16	1.1.3 Changes relative to v2	5
17	1.1.4 Changes relative to v1	5
18	1.2 Outstanding Items	5
19	<b>2 Executive Summary</b>	<b>6</b>
20	<b>3 Data and Monte Carlo Samples</b>	<b>6</b>
21	3.1 Data Samples	7
22	3.2 Monte Carlo Samples	7
23	<b>4 Object Reconstruction</b>	<b>9</b>
24	4.1 Trigger	9
25	4.2 Light leptons	10
26	4.3 Jets	11
27	4.4 B-tagged Jets	11
28	4.5 Missing transverse energy	12
29	<b>5 Event Selection and Signal Region Definitions</b>	<b>12</b>
30	5.1 Event Preselection	12
31	5.2 Fit Regions	15
32	5.3 Non-Prompt Lepton Estimation	30
33	5.3.1 $t\bar{t}$ Validation	30
34	5.3.2 Z+jets Validation	34
35	<b>6 tZ Interference Studies and Separation Multivariate Analysis</b>	<b>38</b>
36	6.1 Top Mass Reconstruction	39
37	6.2 tZ BDT	39
38	<b>7 Systematic Uncertainties</b>	<b>43</b>
39	<b>8 Results</b>	<b>49</b>
40	8.1 1-jet Fit Results	49
41	8.2 2-jet Fit Results	57
42	8.3 tZ Inclusive Results	63
43	<b>9 Conclusion</b>	<b>64</b>

44	<b>Appendices</b>	<b>65</b>
45	.1 tZ Interference Studies	65

---

<sup>46</sup> **1 Changes and outstanding items**

<sup>47</sup> **1.1 Changelog**

<sup>48</sup> This is version 5

<sup>49</sup> **1.1.1 Changes relative to v4**

- <sup>50</sup>     • Fixed various typos, clarified wording  
<sup>51</sup>     • Expanded info about JER uncertainties, theory uncertainties  
<sup>52</sup>     • removed a table on lepton selection, included information in the text instead  
<sup>53</sup>     • Plotted lepton Pt Z and W for Zjet CR, rather than lep 1 and 2  
<sup>54</sup>     • fixed binning in kinematic plots

<sup>55</sup> **1.1.2 Changes relative to v3**

- <sup>56</sup>     • Merged introduction into executive summary, including unblinding details and list of  
<sup>57</sup>     SRs/CRs used  
<sup>58</sup>     • listed ptag used (p4133), and release (AB 21.2.127)  
<sup>59</sup>     • Included table reftab:xsecUnc listing x-sec uncertainties used  
<sup>60</sup>     • Removed selection criteria listed in table ?? (QMisID, AmbiguityType) that were removed  
<sup>61</sup>     from the analysis  
<sup>62</sup>     • specified variable used to make truth jet flavor determination (HadronConeExclTruthLa-  
<sup>63</sup>     belID)  
<sup>64</sup>     • fixed bug in MtLepMet calculation, updated selection/fits to account for this  
<sup>65</sup>     • Included plots of MtLepMet and PtZ, swapped lep 1 and 2  $p_T$  plots for lep W and lep Z  
<sup>66</sup>     plots  
<sup>67</sup>     • updated tZ BDT training to reduce overfitting, updated plots to include error bars, feature  
<sup>68</sup>     importance  
<sup>69</sup>     • updated table 7 to clarify selection, fix the tZ\_BDT cut used  
<sup>70</sup>     • replace a few broken ntuples which included large weight events  
<sup>71</sup>     • include DL1r distribution for Z+jets and  $t\bar{t}$  VRs  
<sup>72</sup>     • Expanded section on fakes, included information on derived scale factors from VRs.

- 73     • Changed the kinematic plots to include  $p_T(Z)$  and  $m_T(W)$ , list lepton  $p_T$  based on W and  
74     Z candidates.

75     **1.1.3 Changes relative to v2**

- 76     • Added alternate VBS samples to include missing b-jet diagrams  
77     • Included a section on tZ interference effects, ??.  
78     • Updated to reflect changes for 2018, including the move to PFlow jets, DL1r, updated  
79     trigger, and updated AnalysisBase version (now 21.2.127)  
80     • Revised fit regions, using separate 1-jet and 2-jet fits, with all 2-j regions included  
81     • updated plots for tZ BDT, added details about the model  
82     • Included truth jet information

83     **1.1.4 Changes relative to v1**

- 84     • Added GRL list  
85     • Fixed latex issue in line 92, typo in line 172  
86     • Added tables 5 and ??, summarizing the event and object selection  
87     • Added table 2, which includes the DSID of samples used  
88     • Included reference to WZ inclusive paper in introduction

89     **1.2 Outstanding Items**

- 90     • Complete interference studies, apply any interference effects observed as a systematic  
91     • Update results section with additional studies, possibly including:  
92        – Truth jet migration studies  
93        – Simultaneous fit over 1j and 2j  
94        – Impact of allowing tZ to float  
95     • Unblind, update plots and fits to include data  
96     • Add cross-section, significance once unblinded

## 97 2 Executive Summary

98 The production of  $WZ$  in association with a heavy flavor jet represents an important background  
 99 for many major analyses. This includes any process with leptons and b-jets in the final state,  
 100 such as  $t\bar{t}H$ ,  $t\bar{t}W$ , and  $t\bar{t}Z$ . While precise measurements have been made of  $WZ$  production  
 101 [1],  $WZ +$  heavy flavor remains poorly understood. This is largely because the QCD processes  
 102 involved in the production of the b-jet make it difficult to simulate accurately. This introduces a  
 103 large uncertainty for analyses that include this process as a background.

104 Motivated by its relevance to the  $t\bar{t}H$  multilepton analysis, we perform a study of the fully  
 105 leptonic decay mode of this channel; that is, events where both the W and Z decay leptonically.  
 106 Because  $WZ$  has no associated jets at leading order, while the major backgrounds for this channel  
 107 tend to have high jet multiplicity, events with more than two jets are rejected. This gives a final  
 108 state signature of three leptons and one or two jets.

109 Events that meet this selection criteria are sorted into pseudo-continuous b-tagging regions based  
 110 on the DL1r b-tag score of their associated jets. This is done to separate  $WZ +$  b-jet events from  
 111  $WZ +$  charm and  $WZ +$  light jets. These regions are fit to data in order make a more accurate  
 112 estimate of the contribution of  $WZ +$  heavy-flavor, where heavy-flavor jets include b-jets and  
 113 charm jets. The full Run-2 dataset collected by the ATLAS detector, representing  $139 \text{ fb}^{-1}$  of  
 114 data from pp collisions at  $\sqrt{s} = 13 \text{ TeV}$ , is used for this study.

115 All backgrounds are accounted for using Monte Carlo, with the simulation of non-prompt lepton  
 116 backgrounds -  $Z+jets$  and  $t\bar{t}$  - validated using non-prompt Validation Regions.

117 Section 3 details the data and Monte Carlo (MC) samples used in the analysis. The reconstruction  
 118 of various physics objects is described in section 4. Section 5 describes the event selection applied  
 119 to these samples, along the definitions of the various regions used in the fit. The multivariate  
 120 analysis techniques used to separate the  $tZ$  background from  $WZ +$  heavy flavor are described in  
 121 section 6. Section 7 describes the various sources of systematic uncertainties considered in the  
 122 fit. Finally, the results of the analysis are summarized in section 8, followed by a brief conclusion  
 123 in section 9.

124 The current state of thee analysis shows blinded results for thee full Run 2 dataset. Regions  
 125 containing  $>5\%$   $WZ+b$  events are blinded, and results are from Asimov, MC only fits. In  
 126 addition to adding some additional information to this note, remaining tasks include performing  
 127  $WZ/tZ$  interference studies, finalizing the presentation of results, and unblinding.

## 128 3 Data and Monte Carlo Samples

129 Both data and Monte Carlo samples used in this analysis were prepared in the `xAOD` format,  
 130 which was used to produce a `DxAOD` sample in the `HIGG8D1` derivation framework. The `HIGG8D1`  
 131 framework is designed for the  $t\bar{t}H$  multi-lepton analysis, which targets events with multiple  
 132 leptons as well as tau hadrons. This framework skims the dataset to remove unneeded variables

133 as well as entire events. Events are removed from the derivations that do not meet the following  
 134 selection:

- 135 • at least two light leptons within a range  $|\eta| < 2.6$ , with leading lepton  $p_T > 15$  GeV and  
 136 subleading lepton  $p_T > 5$  GeV
- 137 • at least one light lepton with  $p_T > 15$  GeV within a range  $|\eta| < 2.6$ , and at least two hadronic  
 138 taus with  $p_T > 15$  GeV.

139 Samples were then generated from these HIGG8D1 derivations with p-tag of p4134 using Ana-  
 140 lysisBase version 21.2.127 modified to include custom variables..

### 141 **3.1 Data Samples**

142 The study uses a sample of proton-proton collision data collected by the ATLAS detector from  
 143 2015 through 2018 at an energy of  $\sqrt{s} = 13$  TeV, which represents an integrated luminosity of  
 144  $139 \text{ fb}^{-1}$ . This data set was collected with a bunch crossing rate of 25 ns. All data used in this  
 145 analysis was verified by data quality checks, having been included in the following Good Run  
 146 Lists:

- 147 • data15\_13TeV.periodAllYear\_DetStatus-v79-repro20-02\_DQDefects-00-02-02  
 148 \_PHYS\_StandardGRL\_All\_Good\_25ns.xml
- 149 • data16\_13TeV.periodAllYear\_DetStatus-v88-pro20-21\_DQDefects-00-02-04  
 150 \_PHYS\_StandardGRL\_All\_Good\_25ns.xml
- 151 • data17\_13TeV.periodAllYear\_DetStatus-v97-pro21-13\_Unknown\_PHYS\_StandardGRL  
 152 \_All\_Good\_25ns\_Triggerno17e33prim.xml
- 153 • data18\_13TeV.periodAllYear\_DetStatus-v102-pro22-04\_Unknown\_PHYS\_StandardGRL  
 154 \_All\_Good\_25ns\_Triggerno17e33prim.xml

155 Runs included from the AllYear period containers are included.

### 156 **3.2 Monte Carlo Samples**

157 Several different generators were used to produce Monte Carlo simulations of the signal and  
 158 background processes. For all samples, the response of the ATLAS detector is simulated using  
 159 Geant4. The WZ signal samples are simulated using Sherpa 2.2.2 [2]. Specific information  
 160 about the Monte Carlo samples being used can be found in table 1. A list of the specific samples  
 161 used by data set ID is shown in table 2.

Table 1: The configurations used for event generation of signal and background processes, including the event generator, matrix element (ME) order, parton shower algorithm, and parton distribution function (PDF).

Process	Event generator	ME order	Parton Shower	PDF
WZ, VV	SHERPA 2.2.2	MEPS NLO	SHERPA	CT10
tZ	MG5_AMC	LO	PYTHIA 6	CTEQ6L1
t̄W	MG5_AMC (SHERPA 2.1.1)	NLO (LO multileg)	PYTHIA 8 (SHERPA)	NNPDF 3.0 NLO (NNPDF 3.0 NLO)
t̄t(Z/γ* → ll)	MG5_AMC	NLO	PYTHIA 8	NNPDF 3.0 NLO
t̄tH	MG5_AMC (MG5_AMC)	NLO (NLO)	PYTHIA 8 (HERWIG++)	NNPDF 3.0 NLO [Ball:2014uwa] (CT10 [ <b>ct10</b> ])
tHqb	MG5_AMC	LO	PYTHIA 8	CT10
tHW	MG5_AMC (SHERPA 2.1.1)	NLO (LO multileg)	HERWIG++ (SHERPA)	CT10 (NNPDF 3.0 NLO)
tWZ	MG5_AMC	NLO	PYTHIA 8	NNPDF 2.3 LO
t̄t, t̄t̄t̄	MG5_AMC	LO	PYTHIA 8	NNPDF 2.3 LO
t̄tW+W-	MG5_AMC	LO	PYTHIA 8	NNPDF 2.3 LO
t̄t	Powheg-BOX v2 [ <b>powheg</b> tt]	NLO	PYTHIA 8	NNPDF 3.0 NLO
t̄tγ	MG5_AMC	LO	PYTHIA 8	NNPDF 2.3 LO
s-, t-channel, Wt single top	Powheg-BOX v1 [ <b>powheg</b> st <sub>p</sub> ]	NLO	PYTHIA 6	CT10
qqVV, VVV Z → l+l-	SHERPA 2.2.1	MEPS NLO	SHERPA	NNPDF 3.0 NLO

Sample	DSID
WZ	364253, 364739-42
VV	364250, 364254, 364255, 363355-60, 364890
t̄W	410155
t̄Z	410156, 410157, 410218-20
low mass t̄Z	410276-8
Rare Top	410397, 410398, 410399
single Top	410658-9, 410644-5
three Top	304014
four Top	410080
t̄WW	410081
Z + jets	364100-41
low mass Z + jets	364198-215
W + jets	364156-97
Vγ	364500-35
tZ	410560
tW	410013-4
WtZ	410408
VVV	364242-9
VH	342284-5
WtH	341998
t̄tγ	410389
t̄t	410470
t̄tH	345873-5, 346343-5

Table 2: List of Monte Carlo samples by data set ID used in the analysis.

## 162 4 Object Reconstruction

163 All regions defined in this analysis share a common lepton, jet, and overall event preselection.  
 164 The selection applied to each physics object is detailed here; the event preselection, and the  
 165 selection used to define the various fit regions, is described in section 5.

### 166 4.1 Trigger

167 Events are required to be selected by dilepton triggers, as summarized in table 3.

Dilepton triggers (2015)	
$\mu\mu$ (asymm.)	HLT_mu18_mu8noL1
$ee$ (symm.)	HLT_2e12_lhloose_L12EM10VH
$e\mu, \mu e$ ( $\sim$ symm.)	HLT_e17_lhloose_mu14
Dilepton triggers (2016)	
$\mu\mu$ (asymm.)	HLT_mu22_mu8noL1
$ee$ (symm.)	HLT_2e17_lhvloose_nod0
$e\mu, \mu e$ ( $\sim$ symm.)	HLT_e17_lhloose_nod0_mu14
Dilepton triggers (2017)	
$\mu\mu$ (asymm.)	HLT_mu22_mu8noL1
$ee$ (symm.)	HLT_2e24_lhvloose_nod0
$e\mu, \mu e$ ( $\sim$ symm.)	HLT_e17_lhloose_nod0_mu14
Dilepton triggers (2018)	
$\mu\mu$ (asymm.)	HLT_mu22_mu8noL1
$ee$ (symm.)	HLT_2e24_lhvloose_nod0
$e\mu, \mu e$ ( $\sim$ symm.)	HLT_e17_lhloose_nod0_mu14

Table 3: List of lowest  $p_T$ -threshold, un-prescaled dilepton triggers used for 2015-2018 data taking.

## 168 4.2 Light leptons

- 169 Electron candidates are reconstructed from energy clusters in the electromagnetic calorimeter  
 170 that are associated with charged particle tracks reconstructed in the inner detector [3]. Electron  
 171 candidates are required to have  $p_T > 10$  GeV and  $|\eta_{\text{cluster}}| < 2.47$ . Muon candidates are  
 172 reconstructed by combining inner detector tracks with track segments or full tracks in the muon  
 173 spectrometer [4]. Muon candidates are required to have  $p_T > 10$  GeV and  $|\eta| < 2.5$ . Candidates  
 174 in the transition region between different electromagnetic calorimeter components,  $1.37 <$   
 175  $|\eta_{\text{cluster}}| < 1.52$ , are rejected. A multivariate likelihood discriminant combining shower shape  
 176 and track information is used to distinguish real electrons from hadronic showers (fake electrons).  
 177 To further reduce the non-prompt electron contribution, the track is required to be consistent  
 178 with originating from the primary vertex; requirements are imposed on the transverse impact  
 179 parameter significance ( $|d_0|/\sigma_{d_0} < 5$ ) and the longitudinal impact parameter ( $|\Delta z_0 \sin \theta_\ell| < 0.5$   
 180 mm).
- 181 Muon candidates are reconstructed by combining inner detector tracks with track segments or  
 182 full tracks in the muon spectrometer [4]. Muon candidates are required to have  $p_T > 10$  GeV  
 183 and  $|\eta| < 2.5$ . The longitudinal impact parameter is the same for both electrons and muons, while  
 184 muons are required to pass a slightly tighter transverse impact parameter,  $|d_0|/\sigma_{d_0} < 3$ .

185 All leptons are required to be isolated, as defined through the standard `isolationFixedCutLoose`  
 186 working point supported by combined performance groups. Leptons are additionally required to  
 187 pass a non-prompt BDT selection described in detail in [5]. Optimized working points and scale  
 188 factors for this BDT are taken from that analysis.

### 189 4.3 Jets

190 Jets are reconstructed from calibrated topological clusters built from energy deposits in the  
 191 calorimeters [6], using the anti- $k_t$  algorithm with a radius parameter  $R = 0.4$ . Jets with energy  
 192 contributions likely arising from noise or detector effects are removed from consideration [7],  
 193 and only jets satisfying  $p_T > 25$  GeV and  $|\eta| < 2.5$  are used in this analysis. For jets with  
 194  $p_T < 60$  GeV and  $|\eta| < 2.4$ , a jet-track association algorithm is used to confirm that the jet  
 195 originates from the selected primary vertex, in order to reject jets arising from pileup collisions  
 196 [8].

### 197 4.4 B-tagged Jets

198 In order to make a measurement of  $WZ +$  heavy flavor it is necessary to distinguish these events  
 199 from  $WZ +$  light jets. For this purpose, the DL1r b-tagging algorithm is used to distinguish  
 200 heavy flavor jets from lighter ones. The DL1r algorithm uses jet kinematics, particularly jet  
 201 vertex information, as input for a neural network which assigns each jet a score designed to  
 202 reflect how likely that jet is to have originated from a b-quark.

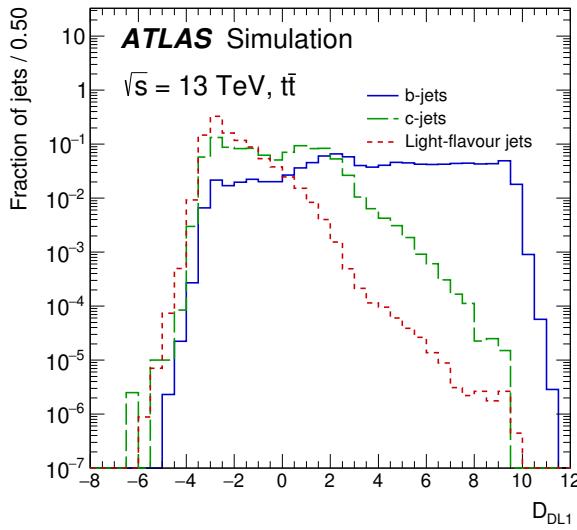


Figure 1: Output distribution of the DL1r algorithm for b-jets, charm jets, and light jets

203 From the output of the BDT, working points (WPs) are developed based on the efficiency of truth  
 204 b-jets at particular values of the DL1r algorithm. The working points used in this analysis are  
 205 summarized in table 4.

WP	none	loose	medium	tight	tightest
b eff.	-	85%	77%	70%	60%

Table 4: B-tagging Working Points by tightness and b-jet efficiency

206 A tighter WP will accept fewer b-jets, but reject a higher fraction of charm and light jets.  
 207 Generally, analyses that include b-jets will use a fixed working point, for example, requiring that  
 208 a jet pass the 70% threshold. By instead treating these working point as bins, e.g. events with  
 209 jets that fall between the 85% and 77% WPs fall into one bin, while events with jets passing the  
 210 60% WP fall into another, and looking at the full psuedo-continuous DL1r spectrum of the jets,  
 211 additional information can be gained. The psuedo-continuous b-tag spectrum is used in this case  
 212 to separate out WZ + b, WZ + charm, and WZ + light.

#### 213 4.5 Missing transverse energy

214 Missing transverse momentum ( $E_T^{\text{miss}}$ ) is used as part of the event selection. The missing  
 215 transverse momentum vector is defined as the inverse of the sum of the transverse momenta of  
 216 all reconstructed physics objects as well as remaining unclustered energy, the latter of which is  
 217 estimated from low- $p_T$  tracks associated with the primary vertex but not assigned to a hard object,  
 218 with object definitions taken from [9]. Light leptons considered in the  $E_T^{\text{miss}}$  reconstruction are  
 219 required to have  $p_T > 10 \text{ GeV}$ , while jets are required to have  $p_T > 20 \text{ GeV}$ .

### 220 5 Event Selection and Signal Region Definitions

221 Event are required to pass a preselection described in section 5.1 and summarized in table 5.  
 222 Those that pass this preselection are divided into various fit regions described in section 5.2,  
 223 based on the number of jets in the event, and the b-tag score of those jets.

#### 224 5.1 Event Preselection

225 Events are required to include exactly three reconstructed light leptons passing the requirement  
 226 described in 4.2, which have a total charge of  $\pm 1$ . As the opposite sign lepton is found to be  
 227 prompt the vast majority of the time [5], it is required to be loose and isolated, as defined though  
 228 the standard `isolationFixedCutLoose` working point supported by combined performance

229 groups. The same sign leptons are required to be very tightly isolated, as per the recommended  
 230 `isolationFixedCutTight`.

231 The leptons are ordered in the analysis code as 0, 1, and 2. Lepton 0 is the lepton whose charge  
 232 is opposite the other two. Lepton 1 is the lepton closest to the opposite charge lepton, i.e. the  
 233 smallest  $\Delta R$ , leaving lepton 2 as the lepton further from the opposite charge lepton. Lepton 0  
 234 is required to have  $p_T > 10$  GeV, while the same sign leptons, 1 and 2, are required to have  
 235  $p_T > 20$  GeV to reduce the contribution of non-prompt leptons.

236 The invariant mass of at least one pair of opposite sign, same flavor leptons is required to fall  
 237 within 10 GeV of the mass of the Z boson, 91.2 GeV. Events where one of the opposite sign pairs  
 238 have an invariant mass less than 12 GeV are rejected in order to suppress low mass resonances.

239 An additional requirement is placed on the missing transverse energy,  $E_T^{\text{miss}} > 20$  GeV, and the  
 240 transverse mass of the W candidate,  $m(E_T^{\text{miss}} + l_{\text{other}}) > 30$  GeV, where  $E_T^{\text{miss}}$  is the missing  
 241 transverse energy, and  $l_{\text{other}}$  is the lepton not included in the Z-candidate.

242 Events are required to have one or two reconstructed jets passing the selection described in  
 243 section 4.3. Events with more than two jets are rejected in order to reduce the contribution of  
 244 backgrounds such as  $t\bar{t}Z$  and  $t\bar{t}W$ , which tend to have higher jet multiplicity.

---

#### Event Selection

---

- Exactly three leptons with charge  $\pm 1$
  - Two tight Iso, tight ID same-charge leptons with  $p_T > 20$  GeV
  - One loose Iso, medium ID opposite charge lepton with  $p_T > 10$  GeV
  - $m(l^+l^-)$  within 10 GeV of 91.2 GeV
  - Transverse mass of W-candidate,  $m_T(E_T^{\text{miss}} + \text{lep}_{\text{other}}) > 30$  GeV
  - Missing transverse energy,  $E_T^{\text{miss}} > 20$  GeV
  - One or two jets with  $p_T > 25$  GeV
- 

Table 5: Summary of the selection applied to events for inclusion in the fit

245 The event yields in the preselection region for both data and Monte Carlo are summarized in  
 246 table 5.1, which shows good agreement between data and Monte Carlo, and demonstrates that  
 247 this region consists primarily of WZ events. The WZ events are split into WZ + b, WZ + c, and  
 248 WZ + l based on the truth flavor of the associated jet in the event. Specifically, this determination  
 249 is made based on the `HadronConeExclTruthLabelID` of the jet, as recommended by the b-  
 250 tagging working group [BtagWG]. In this ordering b-jet supersedes charm, which supersedes  
 251 light. That is, WZ + l events contain no charm and no b jets at truth level, WZ + c contain at  
 252 least one truth charm and no b-jets, and WZ + b contains at least one truth b-jet.

Process	Events
WZ + b	$167.6 \pm 6.5$
WZ + c	$1080 \pm 40$
WZ + l	$7220 \pm 310$
Other VV	$850 \pm 140$
t̄tW	$16.8 \pm 2.3$
t̄tZ	$115 \pm 17$
rare Top	$2.2 \pm 0.1$
Single top	$0.10 \pm 0.45$
Three top	$0.01 \pm 0.01$
Four top	$0.02 \pm 0.01$
t̄tWW	$0.23 \pm 0.05$
Z + jets	$600 \pm 260$
V + $\gamma$	$37 \pm 54$
tZ	$190 \pm 70$
tW	$5.5 \pm 1.2$
WtZ	$25.8 \pm 1.1$
VVV	$26.2 \pm 0.9$
VH	$94 \pm 7$
t̄t	$108.68 \pm 8$
t̄tH	$4.3 \pm 0.5$
Total	$10600 \pm 530$
Data	10574

Table 6: Event yields in the preselection region at  $139.0 \text{ fb}^{-1}$ 

<sup>253</sup> Here Other VV represents diboson processes other than WZ, and consists predominantly of  
<sup>254</sup>  $ZZ \rightarrow ll\bar{l}\bar{l}$  events where one of the leptons is not reconstructed.

<sup>255</sup> Simulations are further validated by comparing the kinematic distributions of the Monte Carlo  
<sup>256</sup> with data, which are shown in figure 2. Here, bins with 5% or more WZ+b are blinded.

### WZ Fit Region - Inclusive

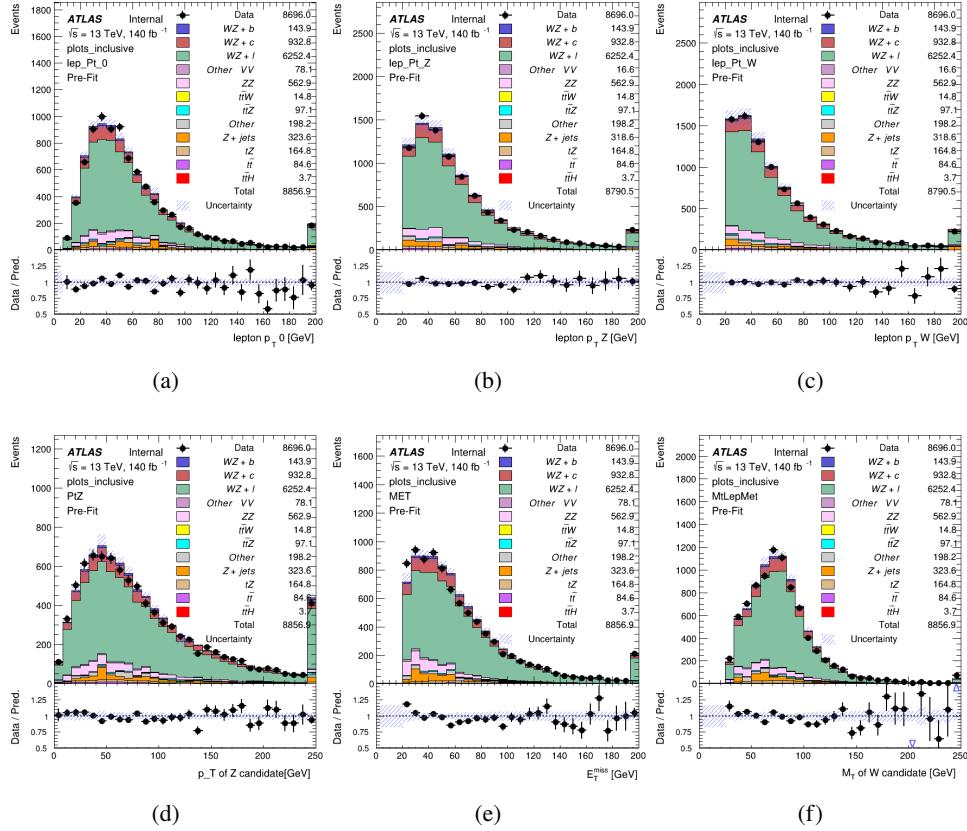


Figure 2: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the Z candidate, (c) the  $p_T$  of the lepton from the W candidate, (d) the  $p_T$  of the Z candidate, (e) the  $E_T^{\text{miss}}$ , and (f) the  $m_T$  of the W candidate.

## 5.2 Fit Regions

Once preselection has been applied, the remaining events are categorized into one of twelve orthogonal regions. The regions used in the fit are summarized in table 7.

Table 7: A list of the regions used in the fit and the selection used for each.

Region	Selection
1j, <85%	$N_{\text{jets}} = 1, n\text{Jets\_DL1r\_85} = 0$
1j, 85%-77%	$N_{\text{jets}} = 1, n\text{Jets\_DL1r\_85} = 1, n\text{Jets\_DL1r\_77}=0$
1j, 77%-70%	$N_{\text{jets}} = 1, n\text{Jets\_DL1r\_77} = 1, n\text{Jets\_DL1r\_70}=0$
1j, 70%-60%	$N_{\text{jets}} = 1, n\text{Jets\_DL1r\_70} = 1, n\text{Jets\_DL1r\_60}=0$
1j, >60%	$N_{\text{jets}} = 1, n\text{Jets\_DL1r\_60} = 1, tZ \text{ BDT} > 0.725$
1j tZ CR	$N_{\text{jets}} = 1, n\text{Jets\_DL1r\_60} = 1, tZ \text{ BDT} < 0.725$
2j, <85%	$N_{\text{jets}} = 2, n\text{Jets\_DL1r\_85} = 0$
2j, 85%-77%	$N_{\text{jets}} = 2, n\text{Jets\_DL1r\_85} >= 1, n\text{Jets\_DL1r\_77}=0$
2j, 77%-70%	$N_{\text{jets}} = 2, n\text{Jets\_DL1r\_77} >= 1, n\text{Jets\_DL1r\_70}=0$
2j, 70%-60%	$N_{\text{jets}} = 2, n\text{Jets\_DL1r\_70} >= 1, n\text{Jets\_DL1r\_60}=0$
2j, >60%	$N_{\text{jets}} = 2, n\text{Jets\_DL1r\_60} >= 1, tZ \text{ BDT} > 0.725$
2j tZ CR	$N_{\text{jets}} = 2, n\text{Jets\_DL1r\_60} >= 1, tZ \text{ BDT} < 0.725$

- 260 The working points discussed in section 4.4 are used to separate events into fit regions based on  
 261 the highest working point reached by a jet in each event. Because the background composition  
 262 differs significantly based on the number of b-jets, events are further subdivided into 1 jet and 2  
 263 jet regions in order to minimize the impact of background uncertainties.
- 264 An additional tZ control region is created based on the BDT described in section 6. The region  
 265 with 1-jet passing the 60% working point is split in two - a signal enriched region of events with  
 266 a BDT score greater than 0.03, and a tZ control region including events with less than 0.03. This  
 267 cutoff is arrived at by performing an Asimov fit with a variety of cutoffs, and selecting the value  
 268 that produces the highest significance for the measurement of WZ + b.
- 269 The modeling in each region is validated by comparing data and MC predictions for various  
 270 kinematic distributions. These plot are shown in figures 3-16.

## WZ Fit Region - 1j Inclusive

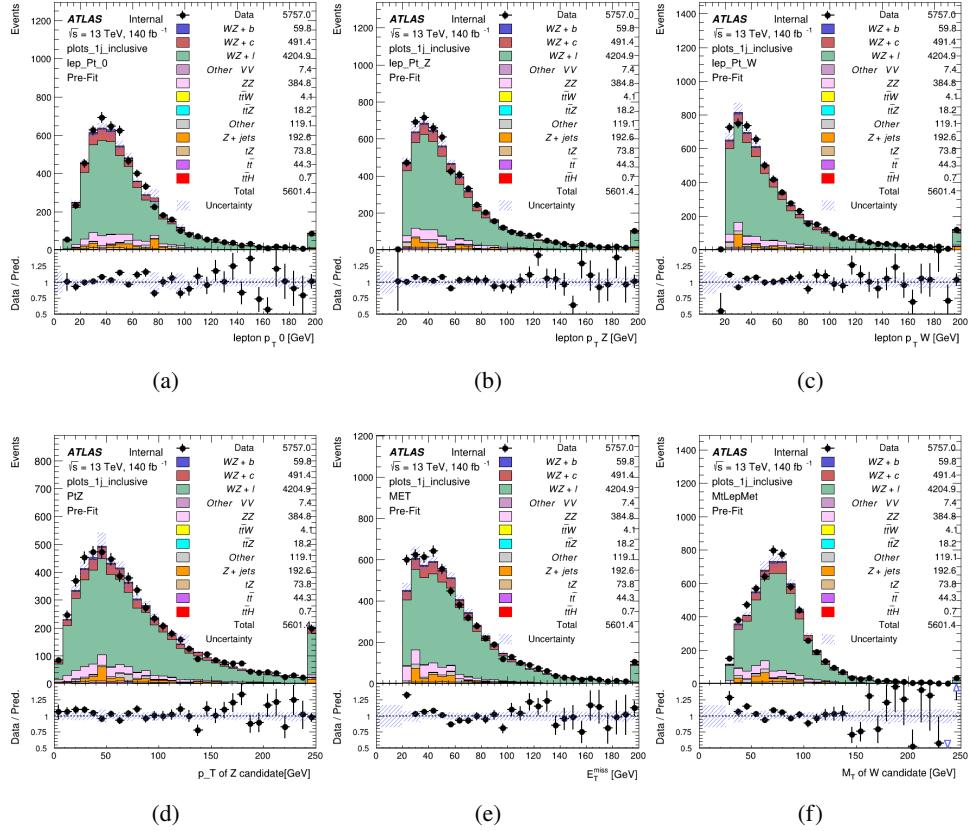


Figure 3: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the Z candidate, (c) the  $p_T$  of the lepton from the W candidate, (d) the  $p_T$  of the Z candidate, (e) the  $E_T^{\text{miss}}$ , and (f) the  $m_T$  of the W candidate.

## WZ Fit Region - 1j &lt; 85% WP

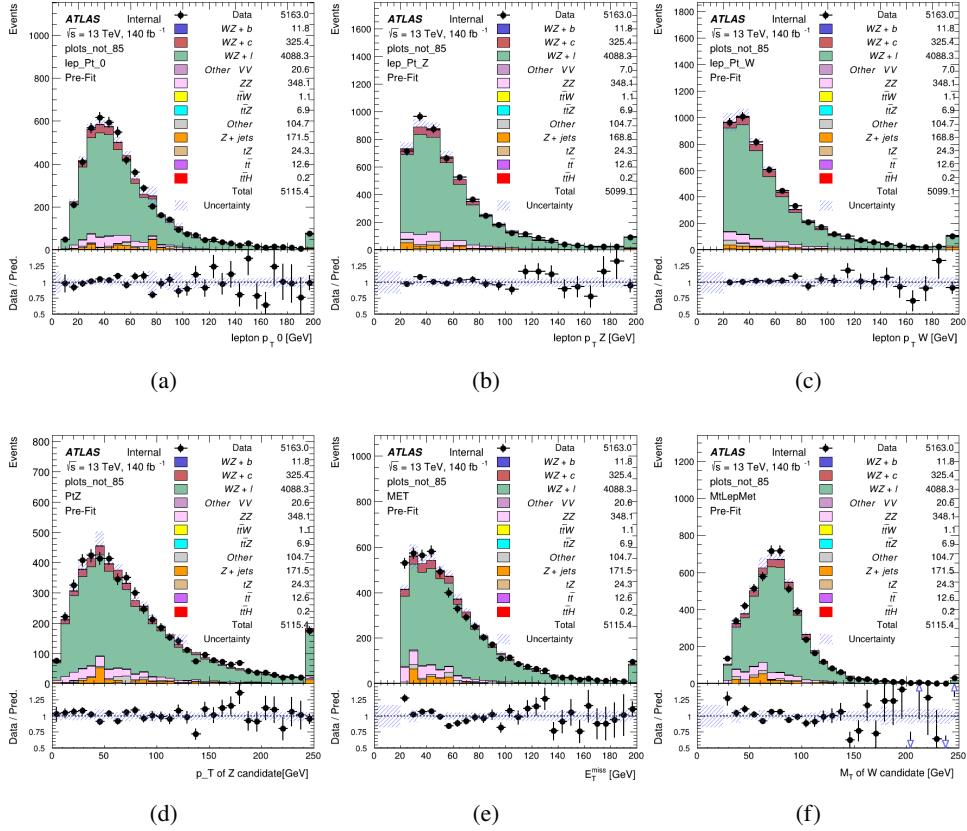


Figure 4: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the  $Z$  candidate, (c) the  $p_T$  of the lepton from the  $W$  candidate, (d) the  $p_T$  of the  $Z$  candidate, (e) the  $E_T^{\text{miss}}$ , and (f) the  $m_T$  of the  $W$  candidate.

## WZ Fit Region - 1j 77-85% WP

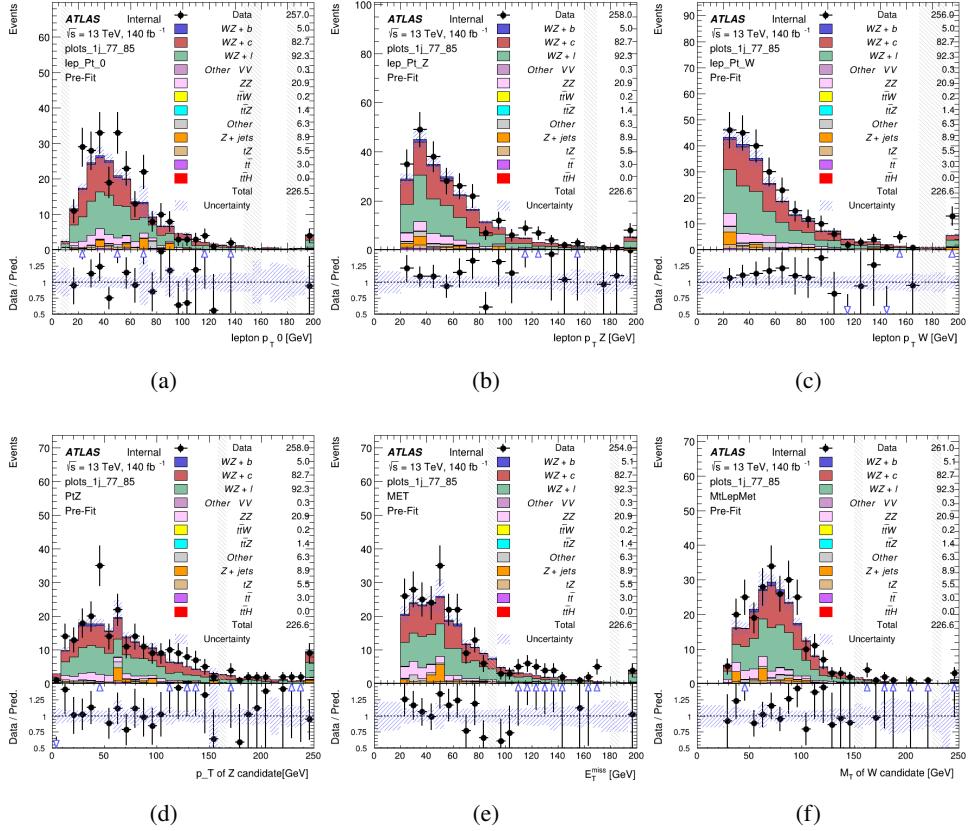


Figure 5: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the  $Z$  candidate, (c) the  $p_T$  of the lepton from the  $W$  candidate, (d) the  $p_T$  of the  $Z$  candidate, (e) the  $E_T^{\text{miss}}$ , and (f) the  $m_T$  of the  $W$  candidate.

## WZ Fit Region - 1j 70-77% WP

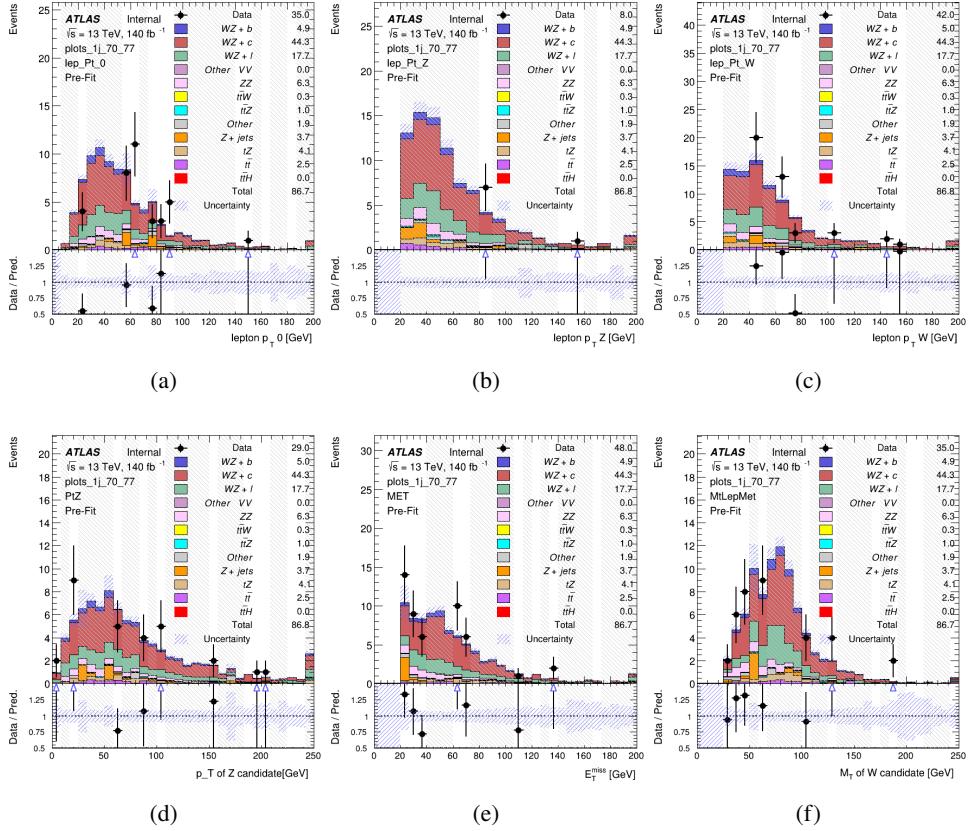


Figure 6: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the  $Z$  candidate, (c) the  $p_T$  of the lepton from the  $W$  candidate, (d) the  $p_T$  of the  $Z$  candidate, (e) the  $E_T^{\text{miss}}$ , and (f) the  $m_T$  of the  $W$  candidate.

## WZ Fit Region - 1j 60-70% WP

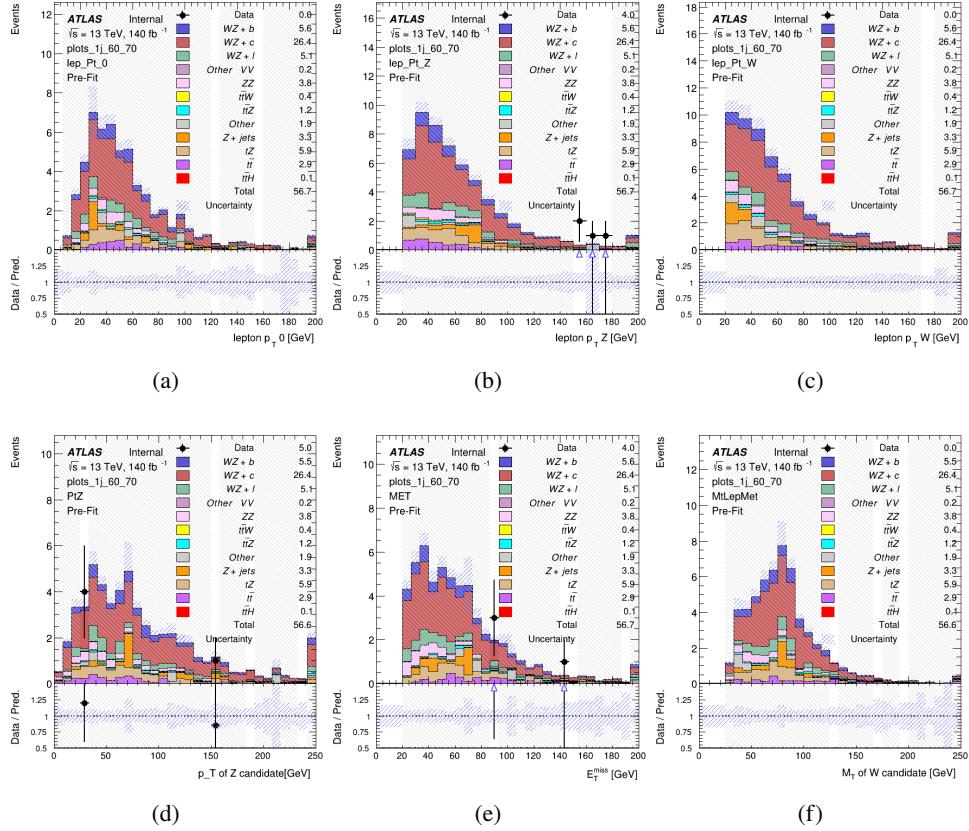


Figure 7: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the  $Z$  candidate, (c) the  $p_T$  of the lepton from the  $W$  candidate, (d) the  $p_T$  of the  $Z$  candidate, (e) the  $E_T^{\text{miss}}$ , and (f) the  $m_T$  of the  $W$  candidate.

## WZ Fit Region - 1j 60% WP

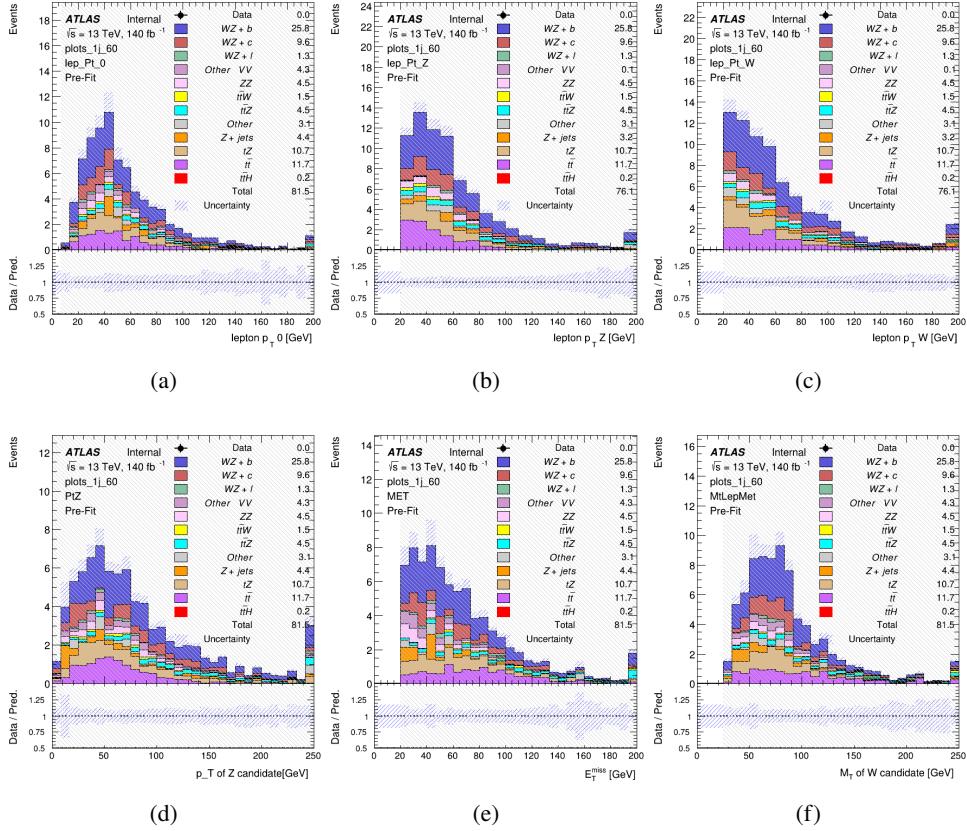


Figure 8: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the  $Z$  candidate, (c) the  $p_T$  of the lepton from the  $W$  candidate, (d) the  $p_T$  of the  $Z$  candidate, (e) the  $E_T^{\text{miss}}$ , and (f) the  $m_T$  of the  $W$  candidate.

## WZ Fit Region - tZ-CR

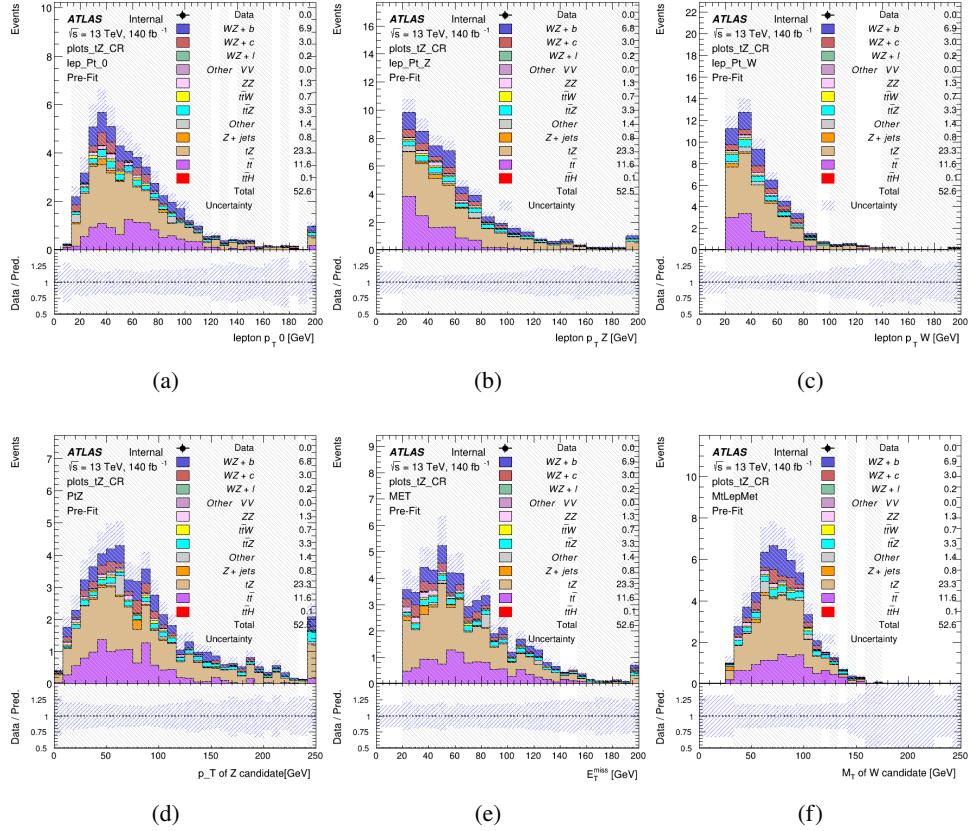


Figure 9: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the  $Z$  candidate, (c) the  $p_T$  of the lepton from the  $W$  candidate, (d) the  $p_T$  of the  $Z$  candidate, (e) the  $E_T^{\text{miss}}$ , and (f) the  $m_T$  of the  $W$  candidate.

## WZ Fit Region - 2j Inclusive

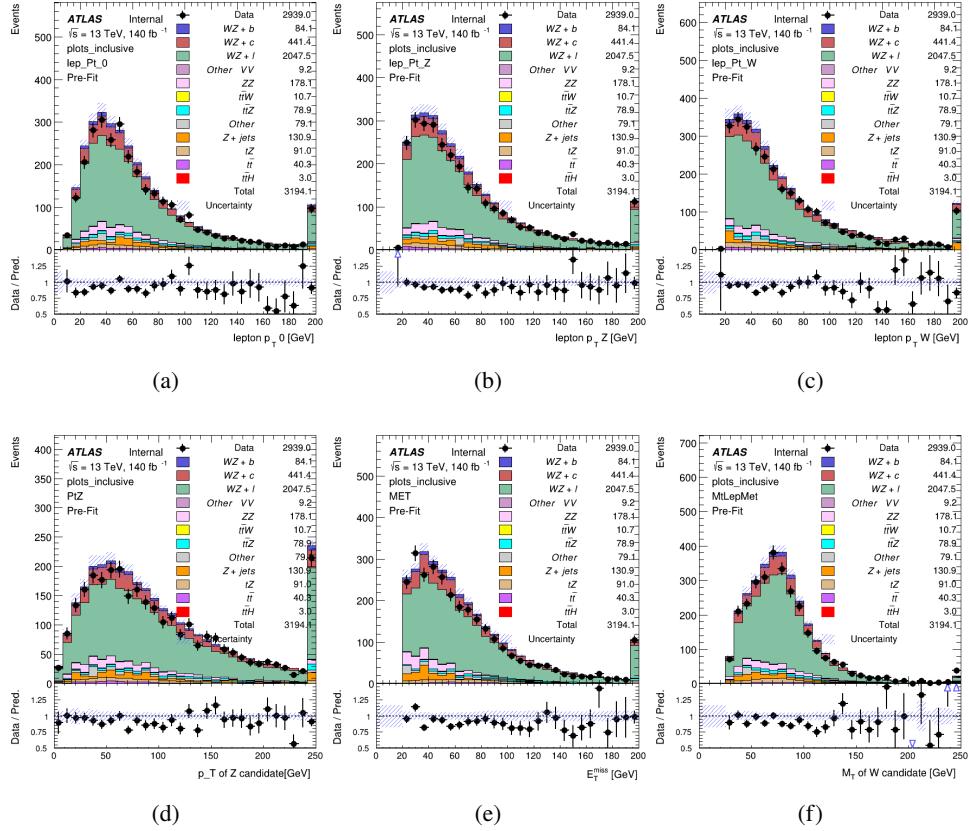


Figure 10: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the Z candidate, (c) the  $p_T$  of the lepton from the W candidate, (d) the  $p_T$  of the Z candidate, (e) the  $E_T^{\text{miss}}$ , and (f) the  $m_T$  of the W candidate.

## WZ Fit Region - 2j &lt; 85% WP

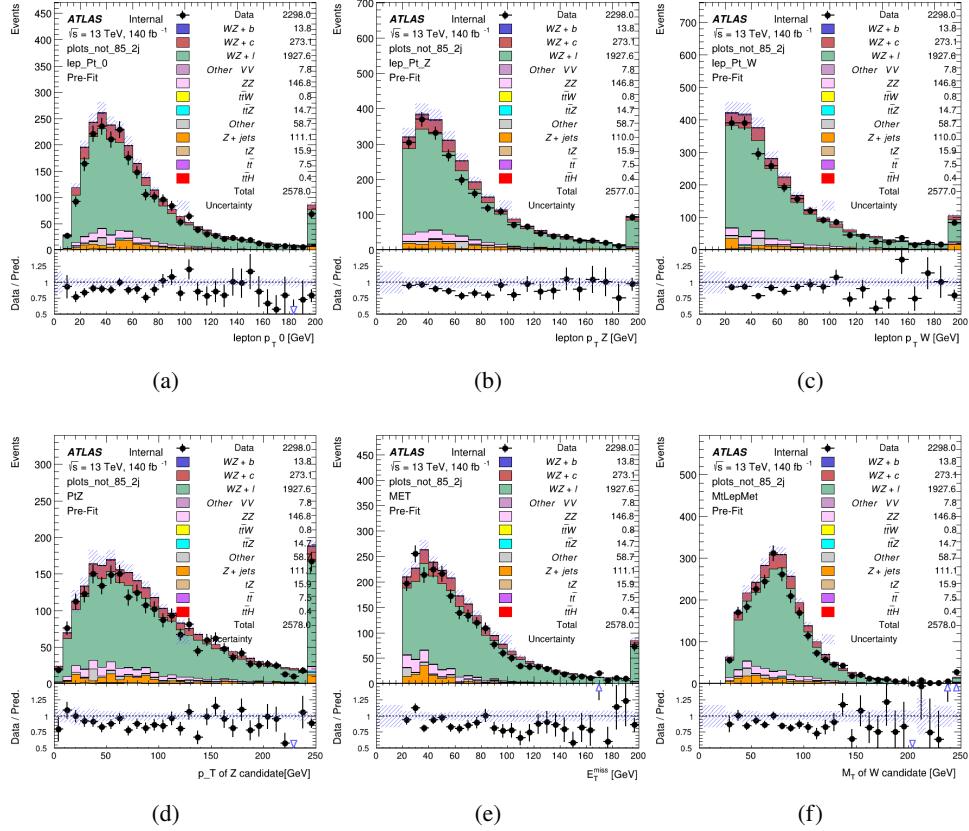


Figure 11: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the Z candidate, (c) the  $p_T$  of the lepton from the W candidate, (d) the  $p_T$  of the Z candidate, (e) the  $E_T^{\text{miss}}$ , and (f) the  $m_T$  of the W candidate.

## WZ Fit Region - 2j 77-85% WP

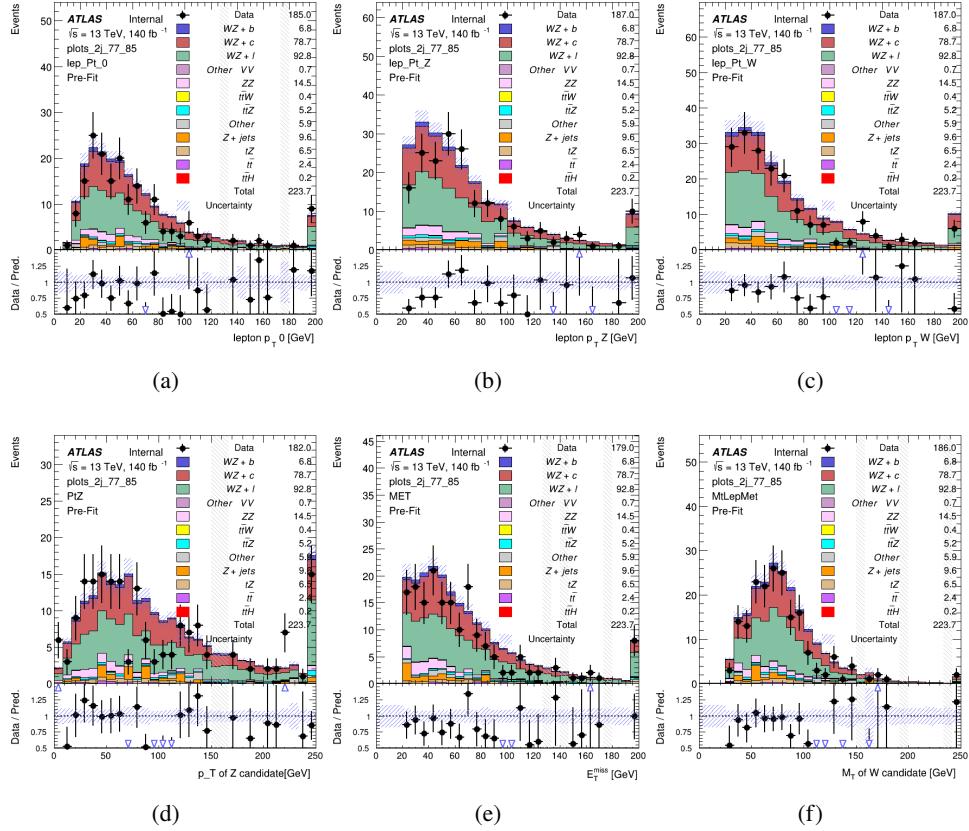


Figure 12: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the Z candidate, (c) the  $p_T$  of the lepton from the W candidate, (d) the  $p_T$  of the Z candidate, (e) the  $E_T^{\text{miss}}$ , and (f) the  $m_T$  of the W candidate.

## WZ Fit Region - 2j 70-77% WP

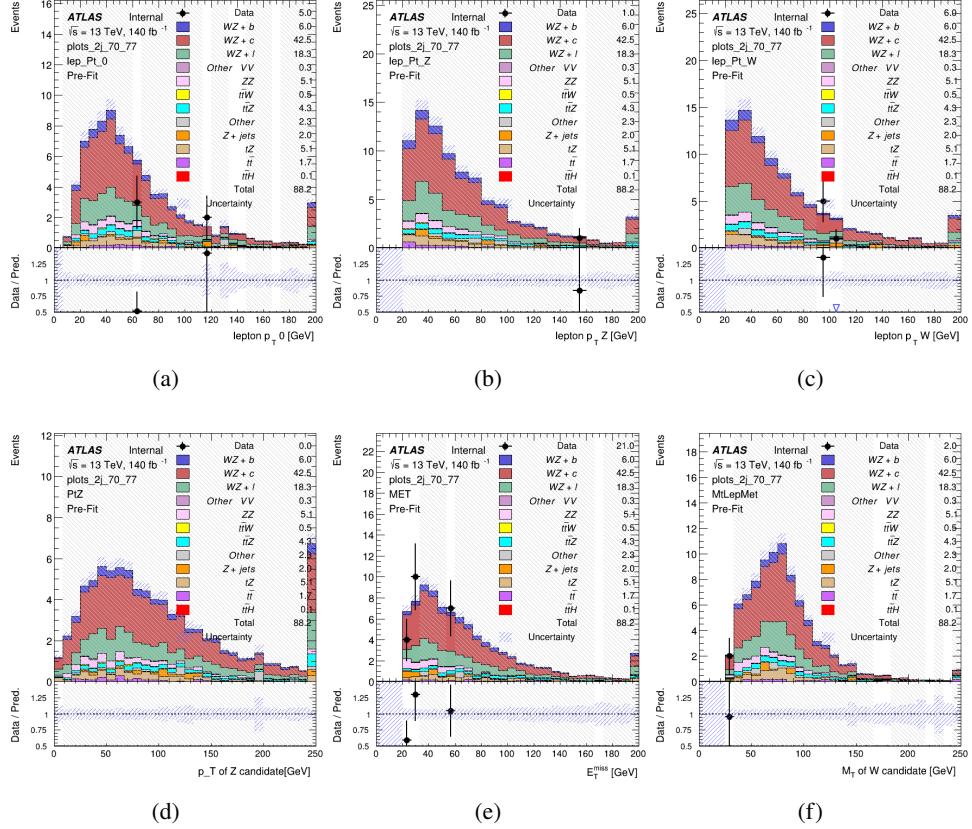


Figure 13: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the Z candidate, (c) the  $p_T$  of the lepton from the W candidate, (d) the  $p_T$  of the Z candidate, (e) the  $E_T^{\text{miss}}$ , and (f) the  $m_T$  of the W candidate.

## WZ Fit Region - 2j 60-70% WP

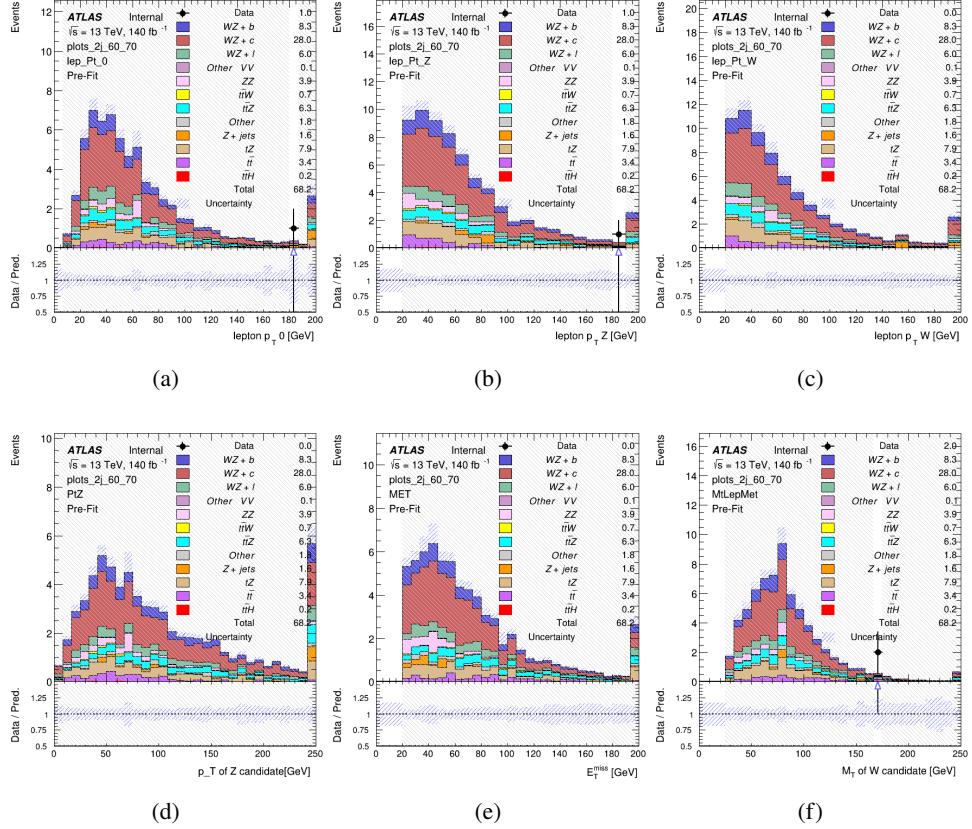


Figure 14: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the Z candidate, (c) the  $p_T$  of the lepton from the W candidate, (d) the  $p_T$  of the Z candidate, (e) the  $E_T^{\text{miss}}$ , and (f) the  $m_T$  of the W candidate.

## WZ Fit Region - 2j 60% WP

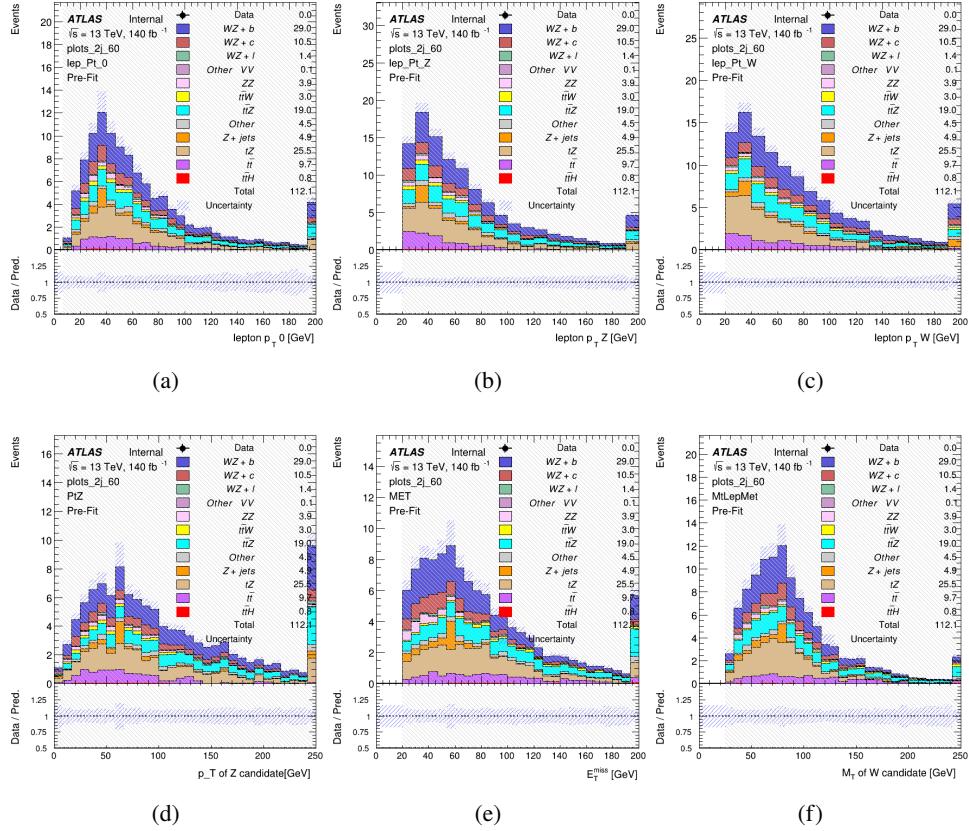


Figure 15: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the Z candidate, (c) the  $p_T$  of the lepton from the W candidate, (d) the  $p_T$  of the Z candidate, (e) the  $E_T^{\text{miss}}$ , and (f) the  $m_T$  of the W candidate.

## WZ Fit Region - tZ-CR-2j

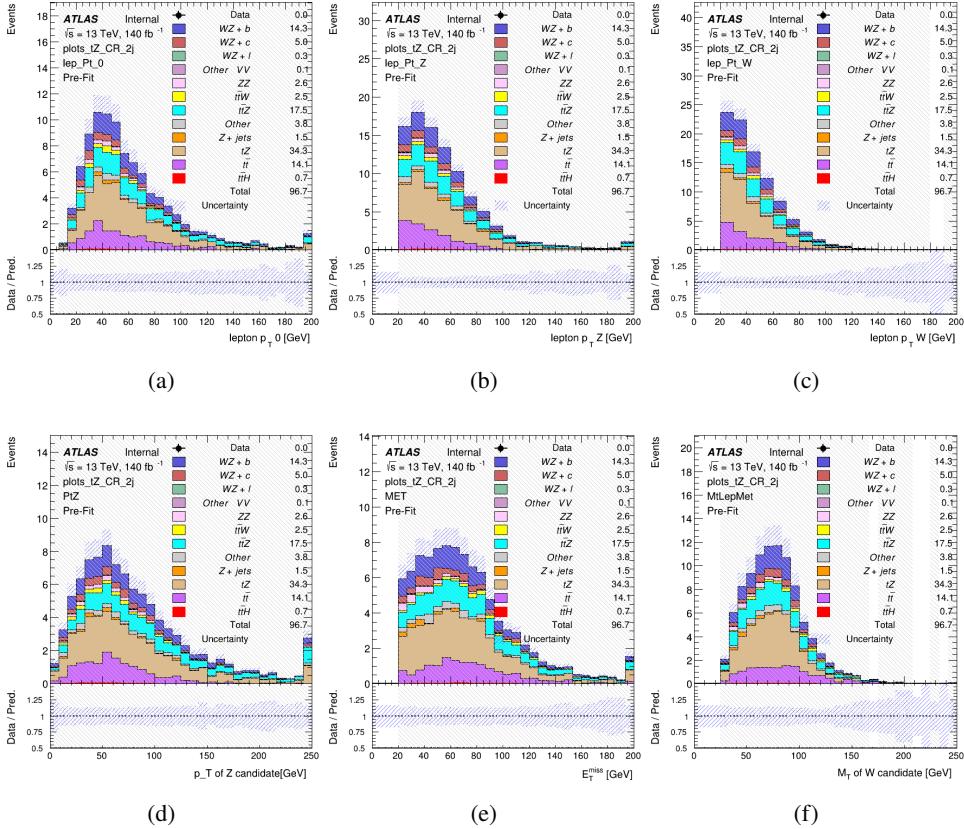


Figure 16: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the  $Z$  candidate, (c) the  $p_T$  of the lepton from the  $W$  candidate, (d) the  $p_T$  of the  $Z$  candidate, (e) the  $E_T^{\text{miss}}$ , and (f) the  $m_T$  of the  $W$  candidate.

### 5.3 Non-Prompt Lepton Estimation

Two processes act as sources of non-prompt leptons appear in the analysis:  $t\bar{t}$  and  $Z+jet$  production both produce two prompt leptons, and each contribute to the 31 region when an additional non-prompt lepton appears in the event. The contribution of these processes is estimated with Monte Carlo simulations, which are validated using enriched validation regions.

#### 5.3.1 $t\bar{t}$ Validation

$t\bar{t}$  events can produce two prompt leptons from the decay of each of the tops. These top decays produce two b-quarks, the decay of which can produce additional non-prompt leptons, which occasionally pass the event preselection. In order to validate that the Monte Carlo accurately

280 simulates this process accurately, the MC prediction in a non-prompt  $t\bar{t}$  enriched validation  
281 region is compared to data.

282 The  $t\bar{t}$  validation region is similar to the preselection region - three leptons meeting the criteria  
283 described in section 5 are required, and the requirements on  $E_T^{\text{miss}}$  remain the same. However,  
284 the selection requiring a lepton pair form a Z-candidate are reversed. Events where the invariant  
285 mass of any two opposite sign, same flavor leptons falls within 10 GeV of 91.2 GeV are rejected.  
286 This ensures the  $t\bar{t}$  validation region is orthogonal to the preselection region.

287 Further, because the jet multiplicity of  $t\bar{t}$  events tends to be higher than WZ, the number of jets  
288 in each event is required to be greater than 1. As b-jets are almost invariably produced from top  
289 decays, at least one b-tagged jet passing the 70% DL1r WP in each event is required. Various  
290 kinematic plots of this region are shown in figure 17.

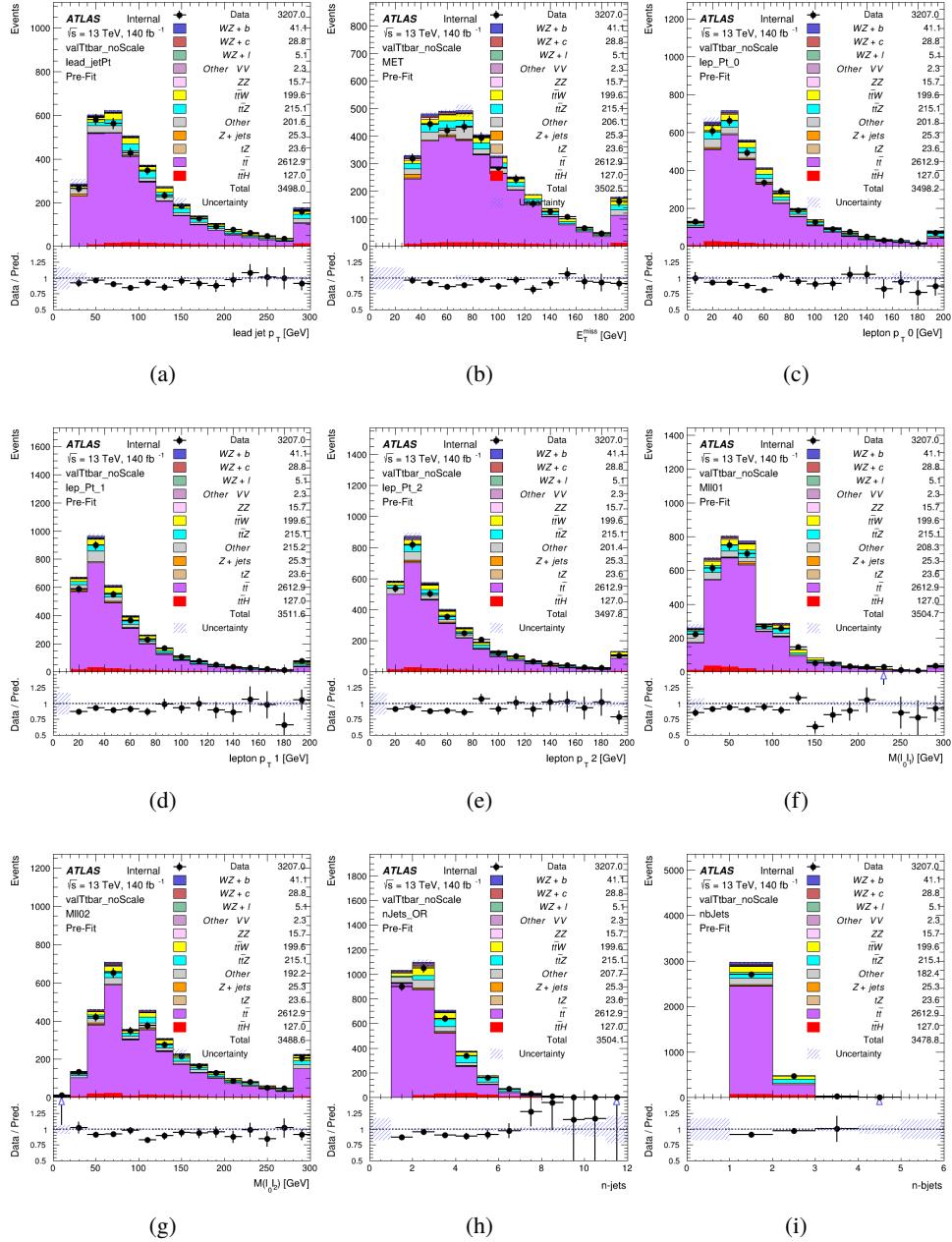


Figure 17: Comparisons between the data and MC distributions in the  $t\bar{t}$  validation region for (a) the  $p_T$  of the leading jet, (b) the missing transverse energy, (c) the  $p_T$  of lepton 0, (d)  $p_T$  of lepton 1, (e)  $p_T$  of lepton 2, (f) the invariant mass of leptons 0 and 1, (g) the invariant mass of leptons 0 and 2, (h) the number of jets, (i) the number of b-tagged jets.

291 The shape of each distribution agrees quite well between data and MC, with a constant offset  
 292 between the two. This is accounted for by applying a constant correction factor of 0.883 to the  $t\bar{t}$

293 MC prediction. Plots showing the kinematics of the  $t\bar{t}$  VR after this correction factor has been  
 294 applied are shown in figure 18.

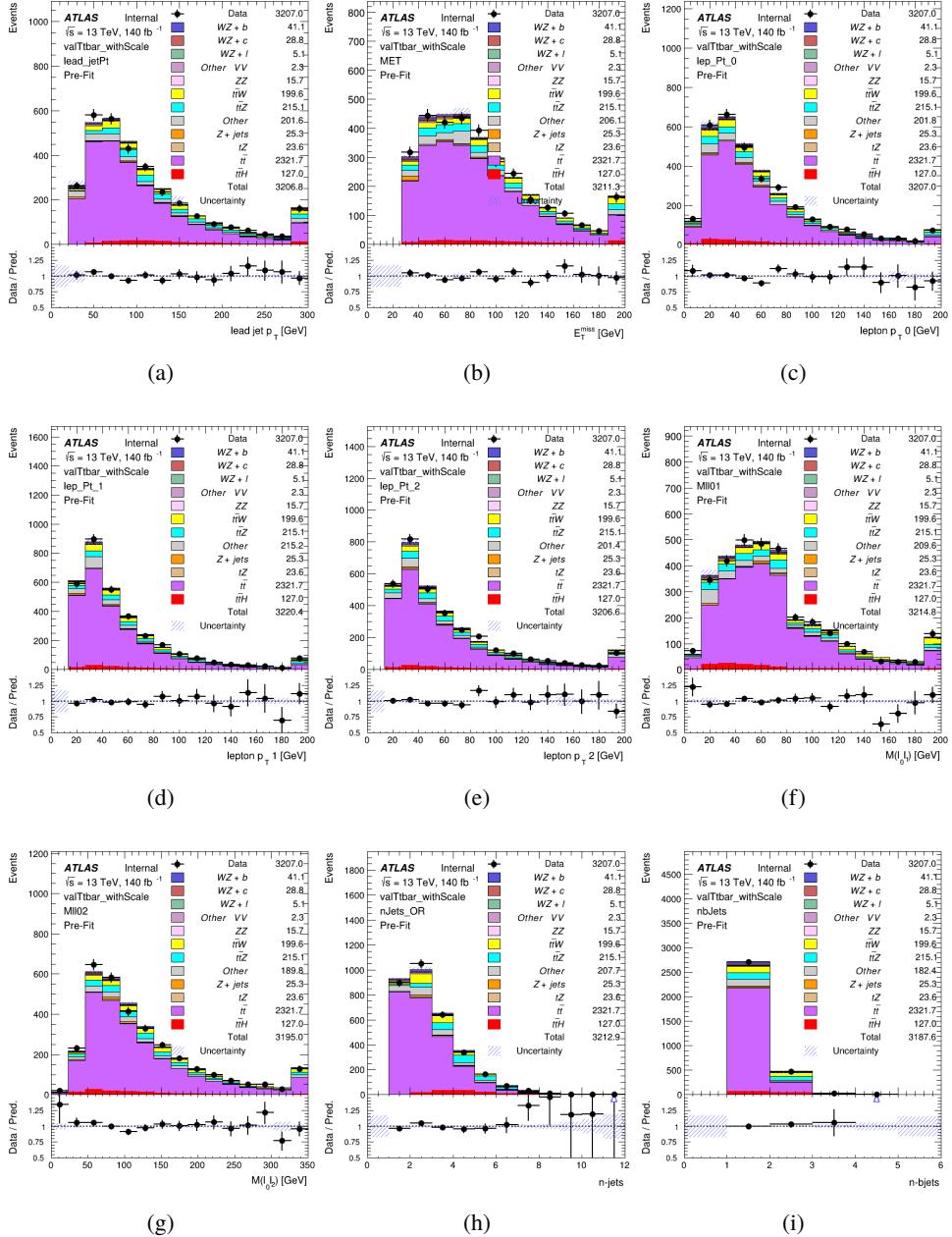


Figure 18: Comparisons between the data and MC distributions in the  $t\bar{t}$  validation region after the correction factor has been applied for (a) the  $p_T$  of the leading jet, (b) the missing transverse energy, (c) the  $p_T$  of lepton 0, (d)  $p_T$  of lepton 1, (e)  $p_T$  of lepton 2, (f) the invariant mass of leptons 0 and 1, (g) the invariant mass of leptons 0 and 2, (h) the number of jets, (i) the number of b-tagged jets.

295 The modeling is further validated by looking at the yield in the  $t\bar{t}$  VR for each DL1r WP, giving  
 296 a clearer correspondence to the signal regions used in the fit. Each region shown in figure 22  
 297 requires one or more jets pass the listed WP, with no jets passing the next highest WP.

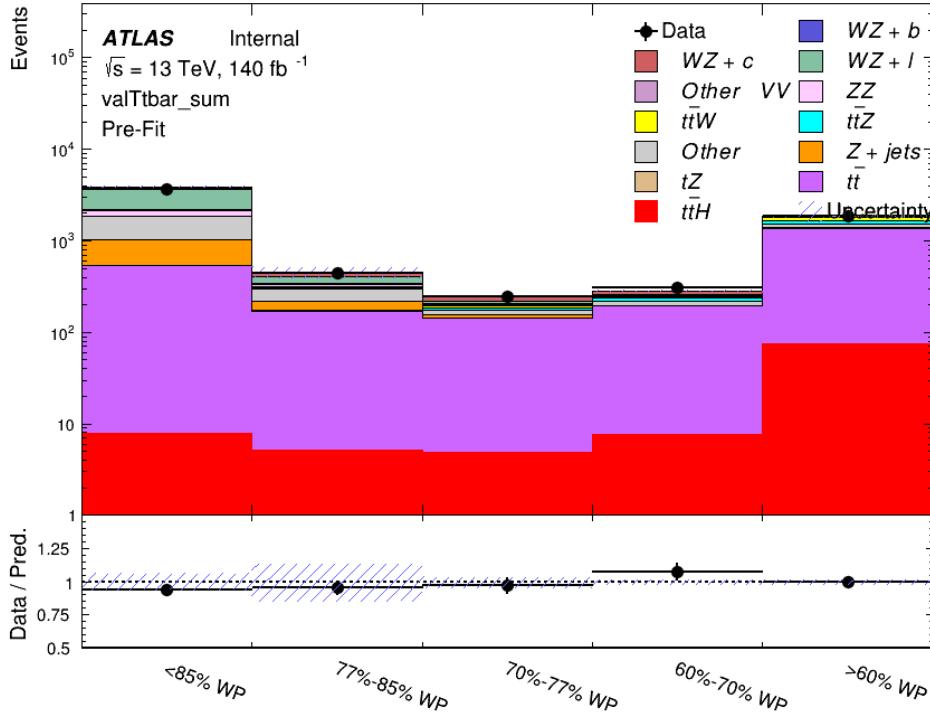


Figure 19: Data and MC comparisons for each DL1r WP for both 1-jet and 2-jet regions, after the  $t\bar{t}$  VR selection and correction factor have been applied

298 As data and MC are found to agree within 10% for each of these working points, a 10% systematic  
 299 uncertainty on the  $t\bar{t}$  prediction is included for the analysis.

### 300 5.3.2 Z+jets Validation

301 Similar to  $t\bar{t}$ , a non-prompt Z+jets validation region is produced in order to validate the MC  
 302 predictions. The lepton requirements remain the same as the preselection region. Because no  
 303 neutrinos are present for this process, the  $E_T^{\text{miss}}$  cut is reversed, requiring  $E_T^{\text{miss}} < 30 \text{ GeV}$ . This  
 304 also ensures this validation region is orthogonal to the preselection region. Further, the number  
 305 of jets in each event is required to be greater than or equal to one. Various kinematic plots of this  
 306 region are shown below. The general agreement between data and MC in each of these suggests  
 307 that the non-prompt contribution of Z+jets is well modeled by Monte Carlo.

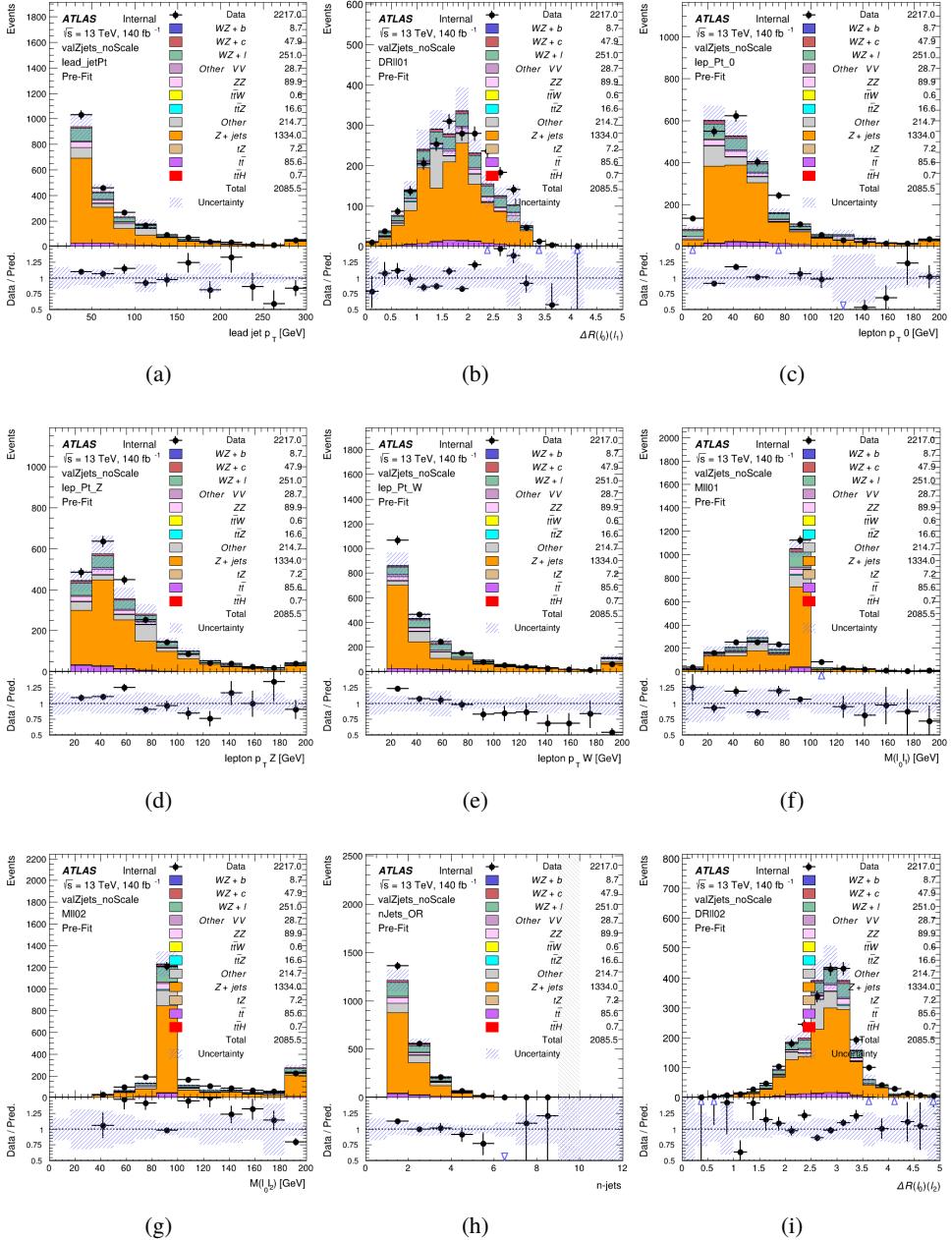


Figure 20: Comparisons between the data and MC distributions in the Z+jets validation region for (a) the  $p_T$  of the leading jet, (b)  $\Delta R$  between leptons 0 and 1, (c) the  $p_T$  of lepton 0, (d)  $p_T$  of SS lepton from the Z candidate, (e)  $p_T$  of the SS lepton from the W candidate, (f) the invariant mass of leptons 0 and 1, (g) the invariant mass of leptons 0 and 2, (h) the number of jets, (i)  $\Delta R$  between leptons 0 and 2. Includes only statistical uncertainties

308 While there is general agreement between data and MC within statistical uncertainty, the shape

<sup>309</sup> of the  $p_T$  spectrum of the lepton from the W is found to differ. To account for this discrepancy,  
<sup>310</sup> a variable correction factor is applied to Z+jets.  $\chi^2$  minimization of the lepton 2  $p_T$  spectrum is  
<sup>311</sup> performed to derive a correction factor of  $1.53 - 6.6 \cdot 10^{-6}(\text{lep\_Pt\_W})$ . Kinematic plots of the  
<sup>312</sup> Z + jets validation region after this correction factor has been applied are shown in figure 21.

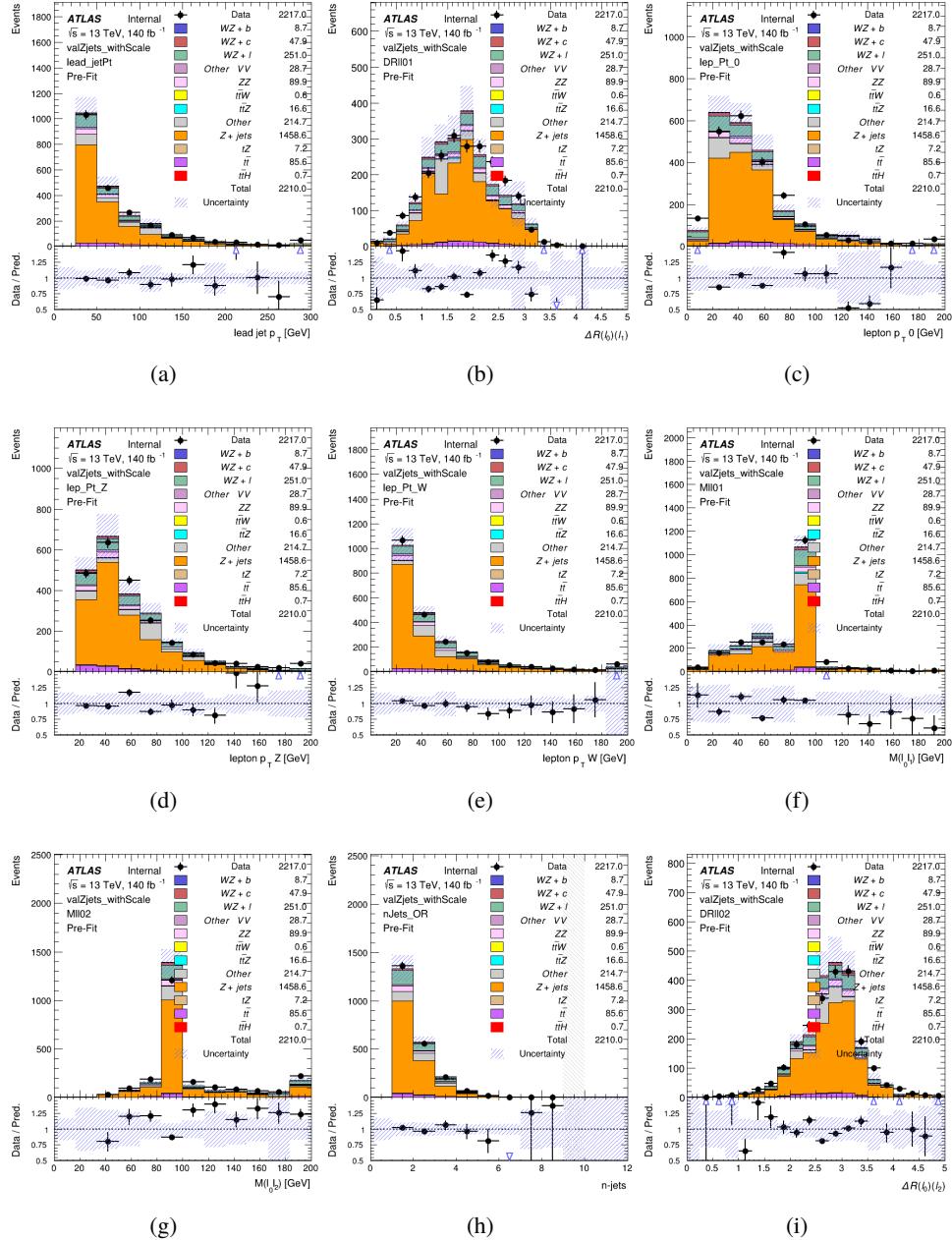


Figure 21: Comparisons between the data and MC distributions in the  $Z + \text{jets}$  validation region after the correction factor has been applied for (a) the  $p_\text{T}$  of the leading jet, (b)  $\Delta R$  between leptons 0 and 1, (c) the  $p_\text{T}$  of lepton 0, (d)  $p_\text{T}$  of SS lepton from the  $Z$  candidate, (e)  $p_\text{T}$  of the SS lepton from the  $W$  candidate, (f) the invariant mass of leptons 0 and 1, (g) the invariant mass of leptons 0 and 2, (h) the number of jets, (i)  $\Delta R$  between leptons 0 and 2

313 The modeling is further validated by looking at the yield in the  $Z + \text{jets}$  VR for each DL1r WP,

314 giving a clearer correspondence to the signal regions used in the fit. Each region shown in figure  
 315 22 requires one or more jets pass the listed WP, with no jets passing the next highest WP.

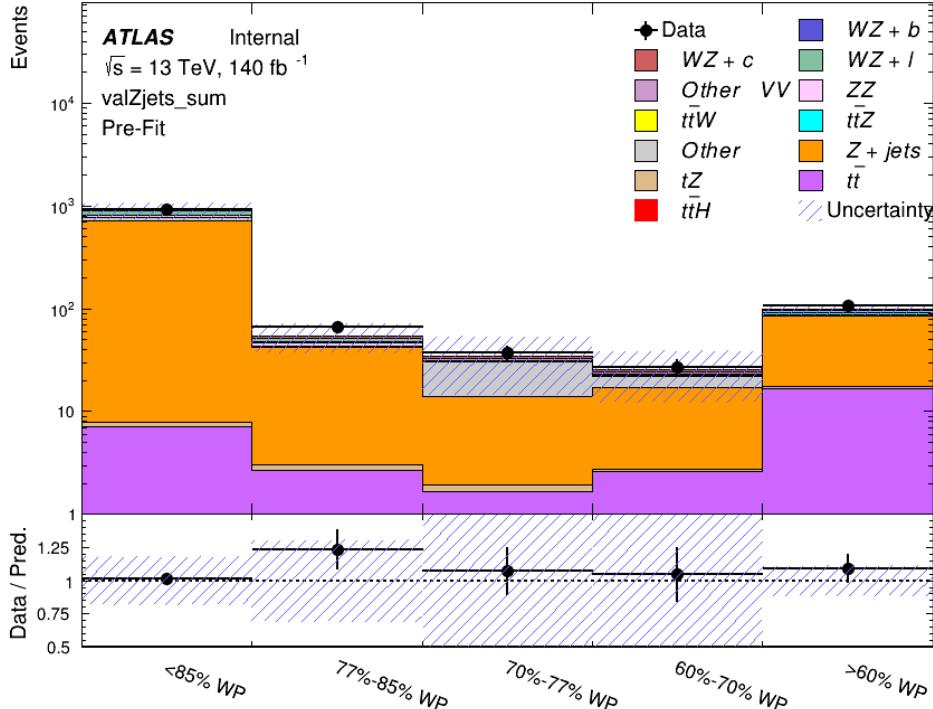


Figure 22: Data and MC comparisons for each DL1r WP for both 1-jet and 2-jet regions, after the Z+jets VR selection and correction factor have been applied

316 For each of the b-tagging working points considered, the data falls within 20% of the MC  
 317 prediction once this correction factor has been applied. Therefore, a 20% systematic uncertainty  
 318 is applied to Z + jets in the analysis.

## 319 6 tZ Interference Studies and Separation Multivariate Analysis

320 Because tZ produces a final state identical to signal, it represents a predominant background in  
 321 the most signal enriched regions. That is, the region with one jet passing the 60% DL1r WP.  
 322 Therefore, a boosted decision tree (BDT) algorithm is trained using TMVA [10] to separate WZ  
 323 + heavy flavor from tZ.

324 Separation between tZ and WZ + heavy flavor is achieved in part by reconstructing the invariant  
 325 mass of the top candidate, which clusters more closely to the top mass for tZ than WZ + heavy  
 326 flavor.

327 The result of this BDT is used to create a tZ enriched region in the fit, reducing its impact on the  
 328 measurement of WZ + heavy flavor.

329 **6.1 Top Mass Reconstruction**

330 The reconstruction of the top mass follows the procedure described in detail in section 6.1 of  
 331 [11]. The mass of the top quark candidate is reconstructed from the jet, the lepton not included  
 332 in the Z-candidate, and a reconstructed neutrino. In the case that there is one jet in the event,  
 333 there is only possible b-jet candidate. For events with two jets, the jet with the highest DL1r  
 334 score is used.

335 The neutrino from the W decay is expected to be the only source of  $E_T^{\text{miss}}$ . Therefore, the  $E_T$   
 336 and  $\phi$  of the neutrino are taken from the  $E_T^{\text{miss}}$  measurement. This leaves the z-component of  
 337 the neutrino momentum,  $p_{vz}$  as the only unknown.

338 This unknown is solved for by taking the combined invariant mass of the lepton and neutrino to  
 339 give the invariant mass of the W boson:

$$340 \quad (p_l + p_\nu)^2 = m_W^2$$

341 Expanding this out into components, this equation gives:

$$342 \quad \sqrt{p_{Tv}^2 + p_{zv}^2} E_l = \frac{m_w^2 - m_l^2}{2} + p_{Tv}(p_{lx} \cos\phi_\nu + p_{ly} \sin\phi_\nu) + p_{lz} p_{vz}$$

343 This equation gives two solutions for  $p_{vz}$ . For cases where only one of these solutions is real,  
 344 that is taken as the value of  $p_{vz}$ . For instances with two real solutions, the one which is shown  
 345 to be correct in the largest fraction of simulations is taken. For cases when no real solution is  
 346 found, often because of detector effects, the value of  $E_T^{\text{miss}}$  is varied in decreasing increments of  
 347 100 MeV until a real solution is found.

348 The reconstructed top mass distribution for tZ and WZ + b can be seen in figure 23.

349 **6.2 tZ BDT**

350 A Boosted Decision Tree (BDT), specifically XGBoost [xgboost\_cite], is used to provide separa-  
 351 tion between tZ and WZ+b. The following kinematic variables are used as inputs:

- 352 • The invariant mass of the reconstructed top candidate
- 353 •  $p_T$  of each of the leptons, jet
- 354 • The invariant mass of each combination of lepton pairs,  $M(l\bar{l})$
- 355 •  $E_T^{\text{miss}}$

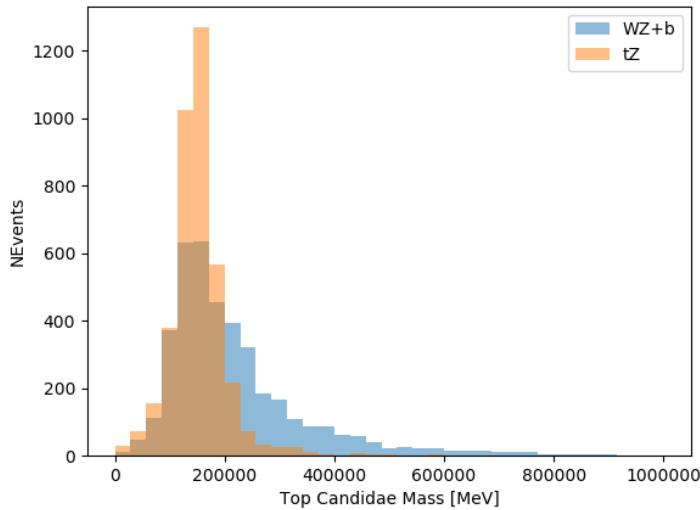


Figure 23: Reconstructed top mass distributions for tZ and WZ + b, measured in MeV.

- 356     • Distance between each combination of leptons,  $\Delta R(ll)$
- 357     • Distance between each lepton and the jet,  $\Delta R(lj)$

358     The training samples included only events meeting the requirements of the 1-jet,  $>60\%$  region,  
 359     i.e. passing all the selection described in section 5 and having exactly one jet which passes the  
 360     tightest (60%) DL1r working point.

361     The distributions of a few of these features for both signal and background is shown in figure  
 362     24.

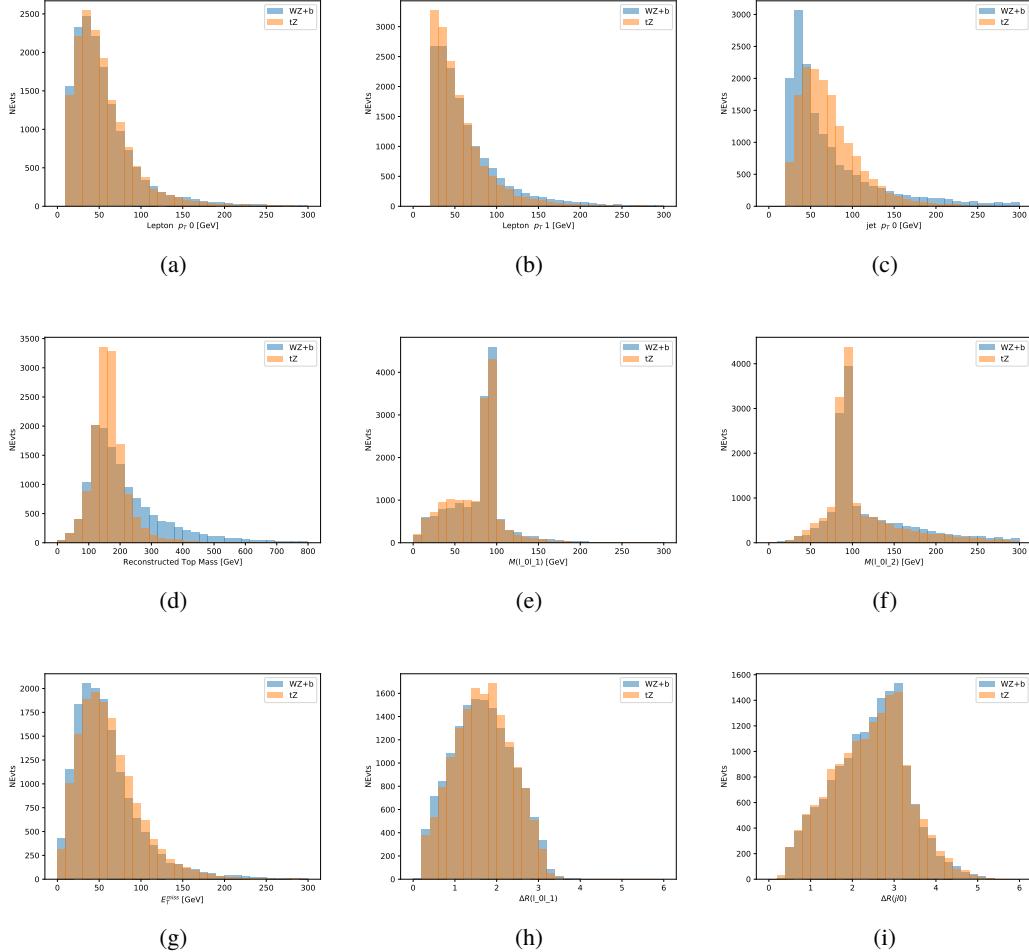


Figure 24: Distribution of input features of the BDT for signal (WZ) and background (tZ). Both are scaled to an equal number of events. (a), (b) and (c) show the  $p_T$  of lepton 0, lepton 1, and the jet, (d) show the reconstructed top mass, (e) and (f) show the invariant mass of leptons 0 and 1, leptons 0 and 2. (g) shows the  $E_T^{\text{miss}}$  of each event. (h) and (i) show the  $\Delta R$  between lepton 0 and lepton 1, and the jet.

363 A sample of 20,000 background (tZ) and signal (WZ+b) Monte Carlo events are used to train  
 364 the BDT. And additional 5,000 events are reserved for testing the model, in order to prevent  
 365 over-fitting. A total of 750 decision trees with a maximum depth of 6 branches are used to build  
 366 the model. These parameters are chosen empirically, by training several models with different  
 367 parameters and selecting the one that gave the best separation for the test sample.

368 The results of the BDT training are shown in figure 25. The output scores for both signal and  
 369 background events is shown on the left. The right shows the receiving operating characteristic  
 370 (ROC) curve that results from the MVA. The ROC curve represents the background rejection  
 371 as a function of signal efficiency, where each point on the curve represents a different response

372 score. The ROC curve of the BDT is compared to the performance of using an optimal set of flat  
 373 selections on the same set of input variables.

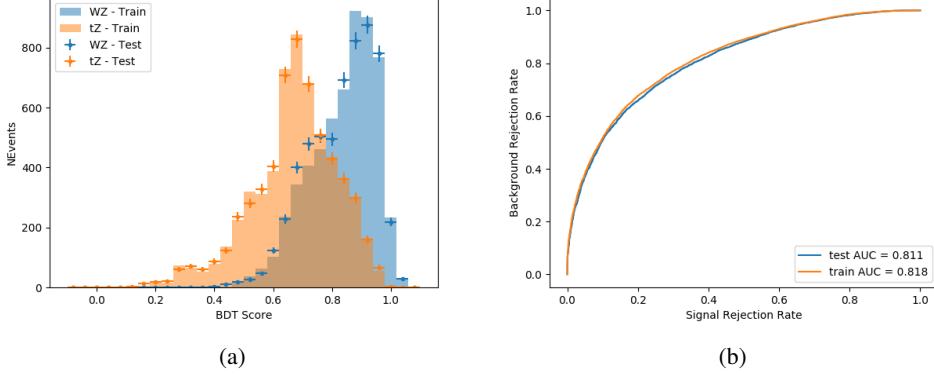


Figure 25: Distribution of the BDT response for signal and background events on the left, the ROC curve for the BDT on the right.

374 The relative important of each input feature in the model, measured by how often they appeared  
 375 in the decision trees, is shown in figure 26.

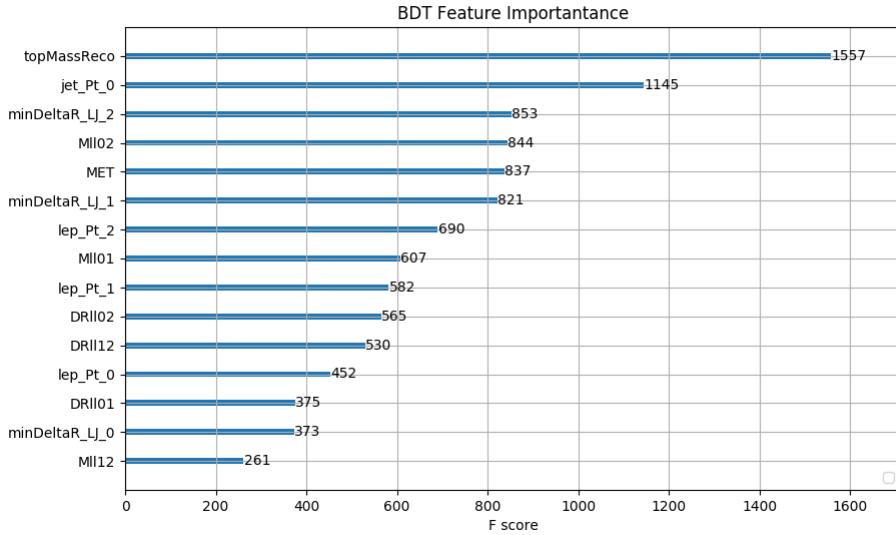


Figure 26: Relative importance of each input feature in the model.

376 These results suggest that some amount of separation can be achieved between these two pro-  
 377 cesses, with a high BDT score selecting a set of events that is pure in  $WZ + b$ . A BDT score

378 of 0.725 is selected as a cutoff, where events with scores higher than this form a signal enriched  
 379 region, and events with scores lower than this form a tZ control region. This cutoff is selected by  
 380 varying the value of this cutoff in stat-only Asimov fits, and selecting the value that minimizes  
 381 the statistical uncertainty on WZ + b.

## 382 7 Systematic Uncertainties

383 The systematic uncertainties that are considered are summarized in table 8. These are imple-  
 384 mented in the fit either as a normalization factors or as a shape variation or both in the signal  
 385 and background estimations. The numerical impact of each of these uncertainties is outlined in  
 386 section 8.

Table 8: Sources of systematic uncertainty considered in the analysis.

Systematic uncertainty	Components
Luminosity	1
Pileup reweighting	1
<b>Physics Objects</b>	
Electron	6
Muon	15
Jet energy scale and resolution	28
Jet vertex fraction	1
Jet flavor tagging	131
$E_T^{\text{miss}}$	3
Total (Experimental)	186
<b>Background Modeling</b>	
Cross section	24
Renormalization and factorization scales	10
Parton shower and hadronization model	2
Shower tune	4
Total (Signal and background modeling)	40
Total (Overall)	226

387 The uncertainty in the combined integrated luminosity is derived from a calibration of the  
 388 luminosity scale performed for 13 TeV proton-proton collisions [12], [**LUCID2**].

389 The experimental uncertainties are related to the reconstruction and identification of light leptons  
 390 and b-tagging of jets, and to the reconstruction of  $E_T^{\text{miss}}$ . The sources which contribute to the  
 391 uncertainty in the jet energy scale (JES) [13] are decomposed into uncorrelated components and  
 392 treated as independent sources in the analysis. The CategoryReduction model is used to account  
 393 for JES uncertainties, which decomposes the uncertainties into 30 nuisance parameters included

394 in the fit. The SimpleJER model is used to account for jet energy resolution (JER) uncertainties,  
395 and 8 JER uncertainty components unclded as NPs in the fit.

396 The uncertainties in the b-tagging efficiencies measured in dedicated calibration analyses [14] are  
397 also decomposed into uncorrelated components. The large number of components for b-tagging  
398 is due to the calibration of the distribution of the MVA discriminant.

399 The systematic uncertainties associated with the signal and background processes are accounted  
400 for by varying the cross-section of each process within its uncertainty.

401 The full list of systematic uncertainties considered in the analysis is summarized in tables 9, 10  
402 and 11.

403

Experimental Systematics on Leptons and $E_T^{\text{miss}}$			
Type	Description	Systematics Name	Application
<b>Trigger</b>			
Scale Factors	Trigger Efficiency	lepSFTrigTight_MU(EL)_SF_Trigger_STAT(SYST)	Event Weight
<b>Muons</b>			
Efficiencies	Reconstruction and Identification	lepSFObjTight_MU_SF_ID_STAT(SYST)	Event Weight
	Isolation	lepSFObjTight_MU_SF_Isol_STAT(SYST)	Event Weight
	Track To Vertex Association	lepSFObjTight_MU_SF_TTVA_STAT(SYST )	Event Weight
$p_T$ Scale	$p_T$ Scale	MUONS_SCALE	$p_T$ Correction
Resolution	Inner Detector Energy Resolution	MUONS_ID	$p_T$ Correction
	Muon Spectrometer Energy Resolution	MUONS_MS	$p_T$ Correction
<b>Electrons</b>			
Efficiencies	Reconstruction	lepSFObjTight_EL_SF_ID	Event Weight
	Identification	lepSFObjTight_EL_SF_Reco	Event Weight
	Isolation	lepSFObjTight_EL_SF_Isol	Event Weight
Scale Factor	Energy Scale	EG_SCALE_ALL	Energy Correction
Resolution	Energy Resolution	EG_RESOLUTION_ALL	Energy Correction
<b><math>E_T^{\text{miss}}</math></b>			
Soft Tracks Terms	Resolution	MET_SoftTrk_ResoPerp	$p_T$ Correction
	Resolution	MET_SoftTrk_ResoPara	$p_T$ Correction
	Scale	MET_SoftTrk_ScaleUp	$p_T$ Correction
	Scale	MET_SoftTrk_ScaleDown	$p_T$ Correction

Table 9: Summary of experimental systematics considered for leptons and  $E_T^{\text{miss}}$ . Includes type, description, name of systematic as used in the fit, and mode of application. The mode of application indicates the systematic evaluation, e.g. as an overall event re-weighting (Event Weight) or rescaling ( $p_T$  Correction).

Experimental Systematics on Jets			
Type	Origin	Systematics Name	Application
Jet Vertex Tagger		JVT	Event Weight
Energy Scale	Calibration Method	JET_21NP_ JET_EffectiveNP_1-19	$p_T$ Correction $p_T$ Correction
	$\eta$ inter-calibration	JET_EtaIntercalibration_Modelling JET_EtaIntercalibration_NonClosure JET_EtaIntercalibration_TotalStat	$p_T$ Correction $p_T$ Correction $p_T$ Correction
	High $p_T$ jets	JET_SingleParticle_HighPt	$p_T$ Correction
	Pile-Up	JET_Pileup_OffsetNPV JET_Pileup_OffsetMu JET_Pileup_PtTerm JET_Pileup_RhoTopology	$p_T$ Correction $p_T$ Correction $p_T$ Correction $p_T$ Correction
	Non Closure	JET_PunchThrough_MC15	$p_T$ Correction
	Flavour	JET_Flavor_Response JET_BJES_Response JET_Flavor_Composition	$p_T$ Correction $p_T$ Correction $p_T$ Correction
Resolution		JET_JER_SINGLE_NP	Event Weight

Table 10: Jet systematics take into account effects of jets calibration method,  $\eta$  inter-calibration, high  $p_T$  jets, pile-up, and flavor response. They are all diagonalised into effective parameters.

Experimental Systematics on b-tagging		
Type	Origin	Systematic Name
Scale Factors	DL1r b-tagger efficiency on b originated jets in bins of $\eta$	DL1r_Continuous_EventWeight_B0-29
	DL1r b-tagger efficiency on c originated jets in bins of $\eta$	DL1r_Continuous_EventWeight_C0-19
	DL1r b-tagger efficiency on light flavoured originated jets in bins of $\eta$ and $p_T$	DL1r_Continuous_EventWeight_Light0-79
	DL1r b-tagger extrapolation efficiency	DL1r_Continuous_EventWeight_extrapolation DL1r_Continuous_EventWeight_extrapolation_from_charm

Table 11: Summary of experimental systematics to be included for b-tagging of jets in the analysis, using the continuous DL1r tagging algorithm. All of the b-tagging related systematics are applied as event weights. From left: type, description, and the name of systematic used in the fit.

- 404 Theoretical uncertainties applied to MC predictions, including cross section, PDF, and scale  
 405 uncertainties are taken from theory calculations, with the exception of non-prompt and diboson  
 406 backgrounds. The cross-section uncertainty on tZ is taken from ?? . Derivation of the non-prompt  
 407 background uncertainties, Z+jets and t $\bar{t}$ , are explained in detail in section 5.3.  
 408 The other VV + heavy flavor processes (namely VV+b and VV+charm, which primarily comprise  
 409 of ZZ events) are also poorly understood, because these processes involve much of the same  
 410 physics as WZ + heavy flavor, and have also not been measured. Therefore, a conservative 50%  
 411 uncertainty is applied to those samples.  
 412 The theory uncertainties applied to the predominate background estimates are summarized in  
 413 table 12.

Process	X-section [%]
WZ	QCD Scale: $^{+3.7}_{-3.4}$ PDF( $+\alpha_S$ ): $\pm 3.1$
tZ	X-sec: $\pm 15.2$ QCD Scale: $^{+5.2}_{-1.3}$ PDF( $+\alpha_S$ ): $\pm 1.2$
t̄H (aMC@NLO+Pythia8)	QCD Scale: $^{+5.8}_{-9.2}$ PDF( $+\alpha_S$ ): $\pm 3.6$
t̄Z (aMC@NLO+Pythia8)	QCD Scale: $^{+9.6}_{-11.3}$ PDF( $+\alpha_S$ ): $\pm 4$
t̄W (aMC@NLO+Pythia8)	QCD Scale: $^{+12.9}_{-11.5}$ PDF( $+\alpha_S$ ): $\pm 3.4$
VV + b/charm (Sherpa 2.2.1)	$\pm 50$
VV + light (Sherpa 2.2.1)	$\pm 6$
t̄t	$\pm 10$
Z + jets	$\pm 20$

Table 12: Summary of theoretical uncertainties for MC predictions in the analysis.

414 Additional signal uncertainties are estimated by comparing estimates from the nominal Sherpa  
 415 WZ samples with alternate WZ samples generated with Powheg+PYTHIA8 (DSID 361601).  
 416 The shape of the templates used in the fit are compared between these two samples for WZ + b,  
 417 WZ + charm and WZ + light, as shown in figures 27 and 28. Each of these plots are normalized  
 418 to unity in order to capture differences in shape.

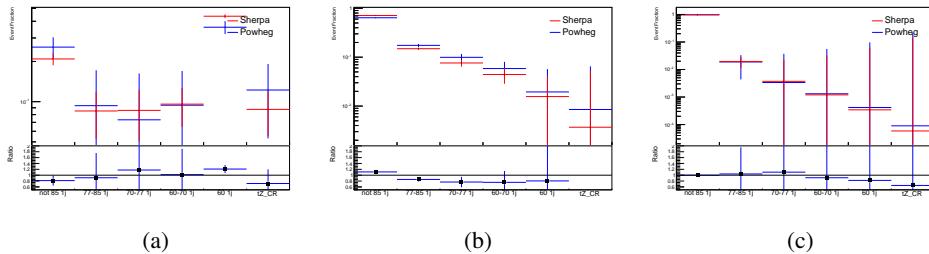


Figure 27: Comparison between Sherpa and Powheg predictions of the distribution of (a) WZ + b, (b) WZ + charm, and (c) WZ + light among the various b-tag WPs used in the 1-jet fit.

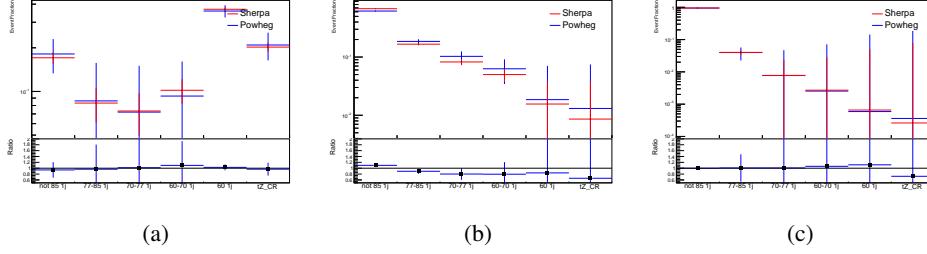


Figure 28: Comparison between Sherpa and Powheg predictions of the distribution of (a) WZ + b, (b) WZ + charm, and (c) WZ + light among the various b-tag WPs used in the 2-jet fit.

419 Separate systematics are included in the fit for WZ + b, WZ + charm and WZ + light, where  
 420 the distribution among each of the fit regions is varied based on the prediction of the Powheg  
 421 sample.

## 422 8 Results

423 A separate maximum-likelihood fit is performed over the 1-jet and 2-jet fit regions in order to  
 424 extract the best-fit value of the WZ + b-jet and WZ + charm jet contributions. The WZ + b,  
 425 WZ + charm and WZ + light contributions are allowed to float, with the remaining background  
 426 contributions are held fixed. **The current fit strategy treats the WZ + b-jet contribution as**  
 427 **the parameter of interest, with the normalization of the WZ + charm and the WZ + light**  
 428 **contributions taken as systematic uncertainties. This could however be adjusted, depending**  
 429 **on whether it is decided the goal of the analysis should be to measure WZ+b specifically or**  
 430 **WZ + heavy flavor overall.** The result of the fit is used to extract the cross-section of WZ +  
 431 heavy-flavor production.

432 A maximum likelihood fit to data is performed simultaneously in the regions described in section  
 433 5. The parameters  $\mu_{WZ+b}$ ,  $\mu_{WZ+charm}$ ,  $\mu_{WZ+light}$ , where  $\mu = \sigma_{\text{observed}}/\sigma_{\text{SM}}$ , are extracted  
 434 from the fit.

435 While in the nominal fit events with an intermediate top quark are treated as a separate process, an  
 436 alternative version of the measurement is performed which considers tZ as part of the signal.

### 437 8.1 1-jet Fit Results

438 **The results of the fit are currently blinded.**

439 The prefit yields in each of the regions used in the fit are shown in figure ??.

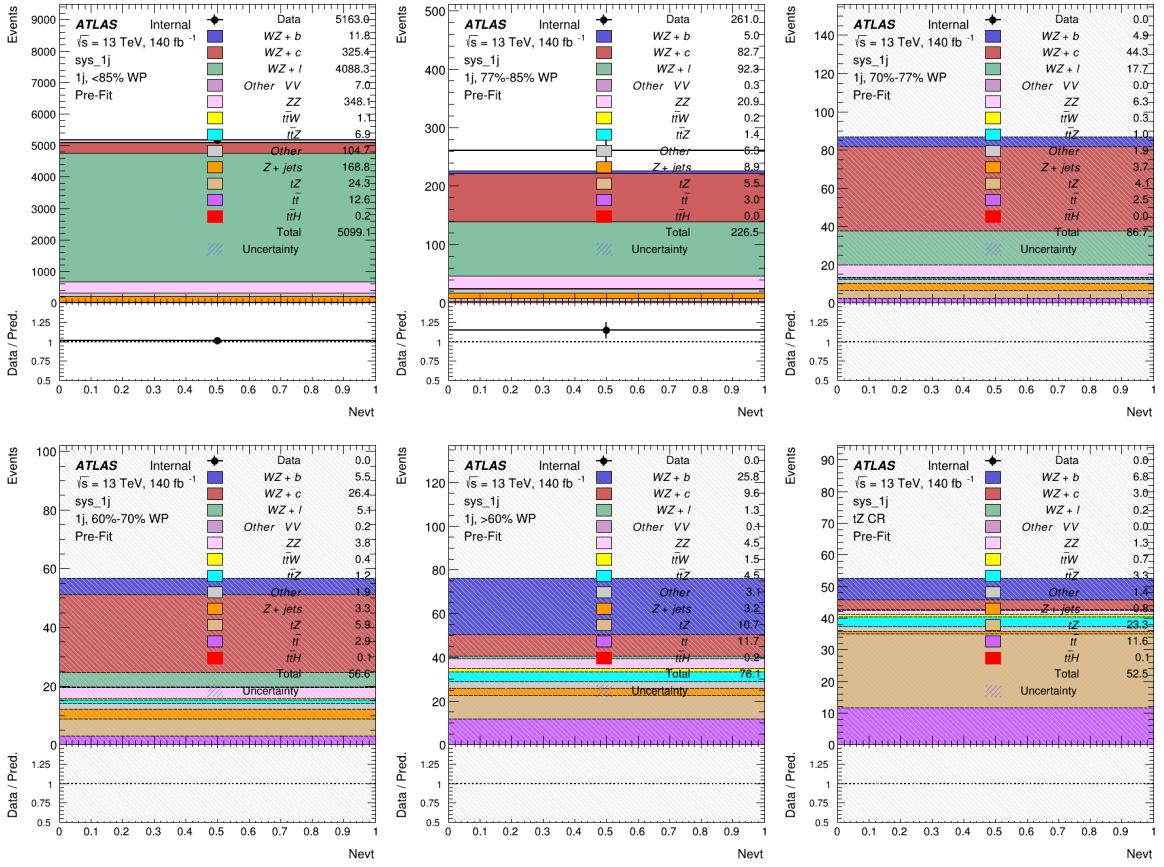


Figure 29: Data/MC yields in each of the 1-jet regions before the fit has been performed.

440 The post-fit yields in each region are summarized in figure 30.

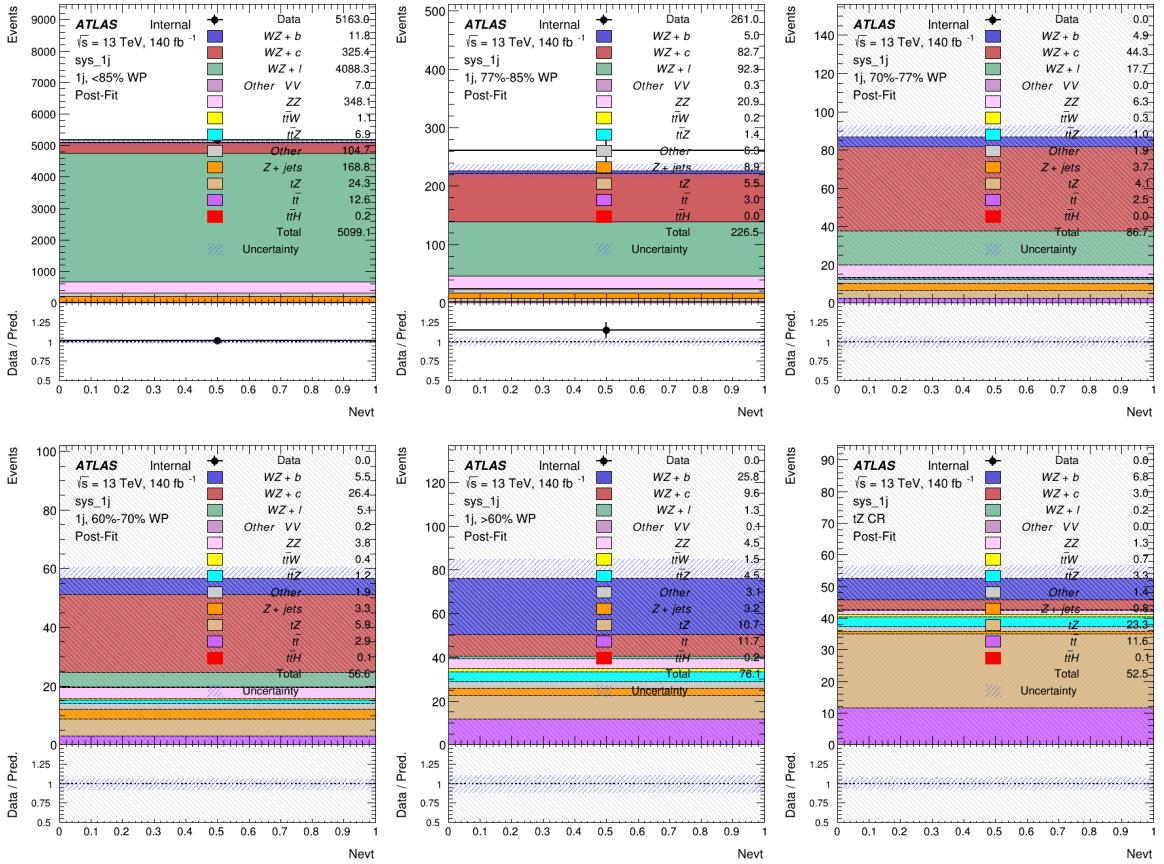


Figure 30: Data/MC results in each of the 1-jet regions after the fit has been performed.

441 A post-fit summary plot of the 1-jet fitted regions is shown in figure 31:

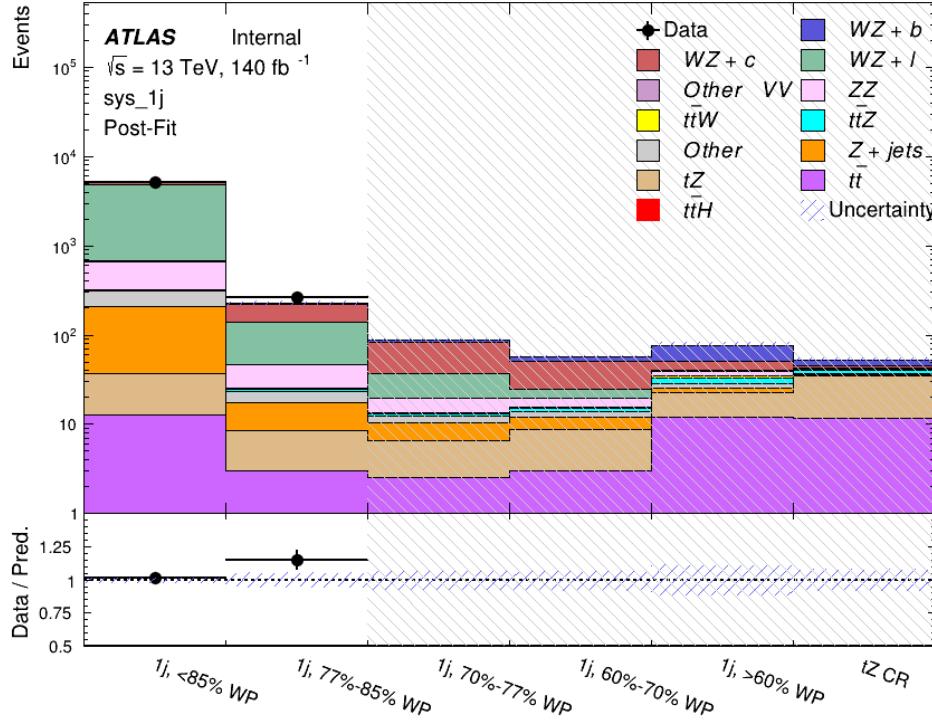


Figure 31: Post-fit summary of the 1-jet fit regions.

- 442 As described in section 7, there are 226 systematic uncertainties that are considered as NPs in  
 443 the fit. These NPs are constrained by Gaussian or log-normal probability density functions. The  
 444 latter are used for normalisation factors to ensure that they are always positive. The expected  
 445 numbers of signal and background events are functions of the likelihood. The prior for each  
 446 NP is added as a penalty term, decreasing the likelihood as it is shifted away from its nominal  
 447 value.  
 448 The impact of each NP is calculated by performing the fit with the parameter of interest held  
 449 fixed, varied from its fitted value by its uncertainty, and calculating  $\Delta\mu$  relative to the baseline  
 450 fit. The impact of the most significant sources of systematic uncertainties is summarized in table  
 451 13.

Uncertainty Source	$\Delta\mu$	
WZ + charm cross-section	-0.1966	0.2171
tZ cross-section	-0.1521	0.1518
WZ + light cross-section	0.1485	-0.1411
Other VV + b cross-section	-0.1115	0.1163
Flavor Tagging	0.0955	0.0957
Jet Energy Scale	0.0613	0.081
t <bar>t</bar>	-0.0662	0.0654
Luminosity	-0.0609	0.0655
Z + jets cross-section	-0.0284	0.0284
Other VV + charm cross-section	0.0207	-0.0202
Muon Trigger Scale Factor	0.019	0.0209
Total Systematic Uncertainty	0.3511	0.3679

Table 13: Summary of the most significant sources of systematic uncertainty on the measurement of WZ + b with exactly one associated jet.

<sup>452</sup> The ranking and impact of those nuisance parameters with the largest contribution to the overall  
<sup>453</sup> uncertainty is shown in figure 32.

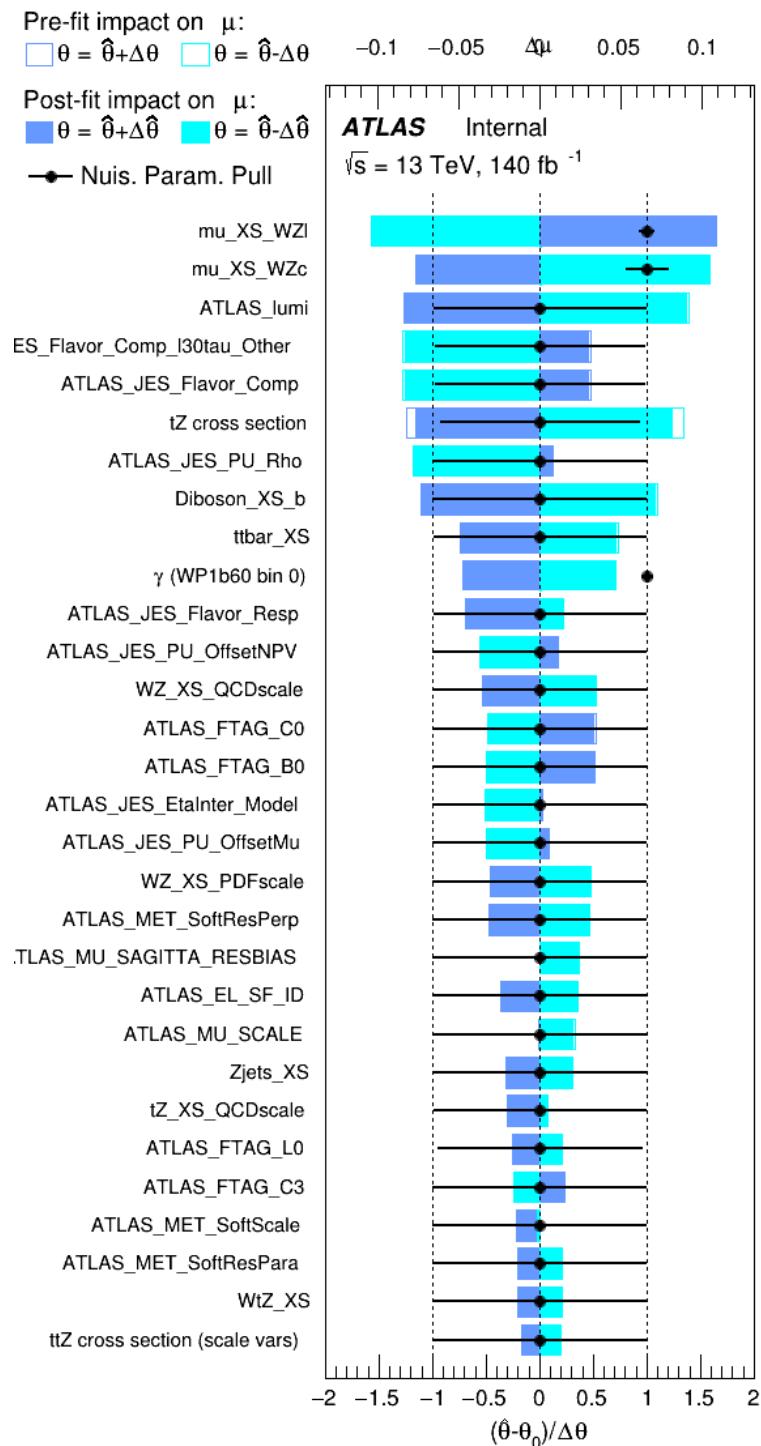


Figure 32: Impact of systematic uncertainties on the signal-strength of  $WZ + b$  for events with exactly one jet

454 The large impact of the Jet Energy Scale and Jet Flavor Tagging is unsurprising, as the shape  
 455 of the fit regions depends heavily on the modeling of the jets. The other major sources of  
 456 uncertainty come from background modelling and cross-section uncertainty. The pie charts in  
 457 figure 33 show that for the modelling uncertainties that contribute most correspond to the most  
 458 significant backgrounds.

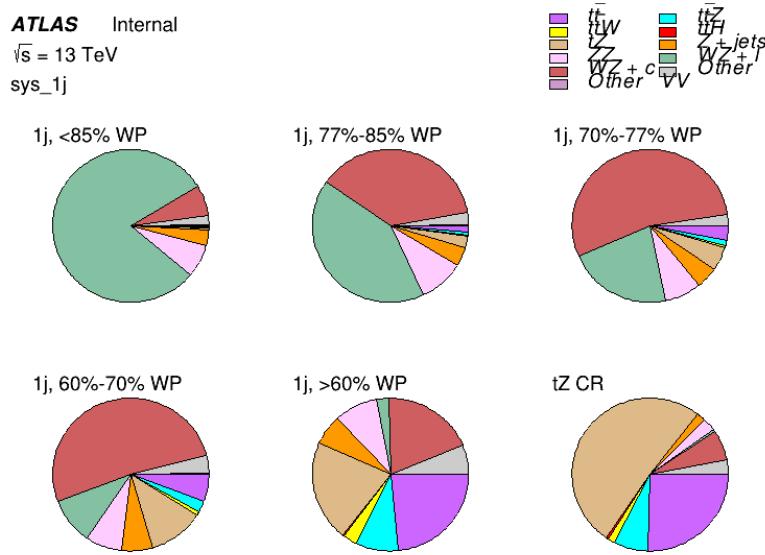


Figure 33: Post-fit background composition of the fit regions.

459 The correlations between these nuisance parameters are summarized in figure 34.

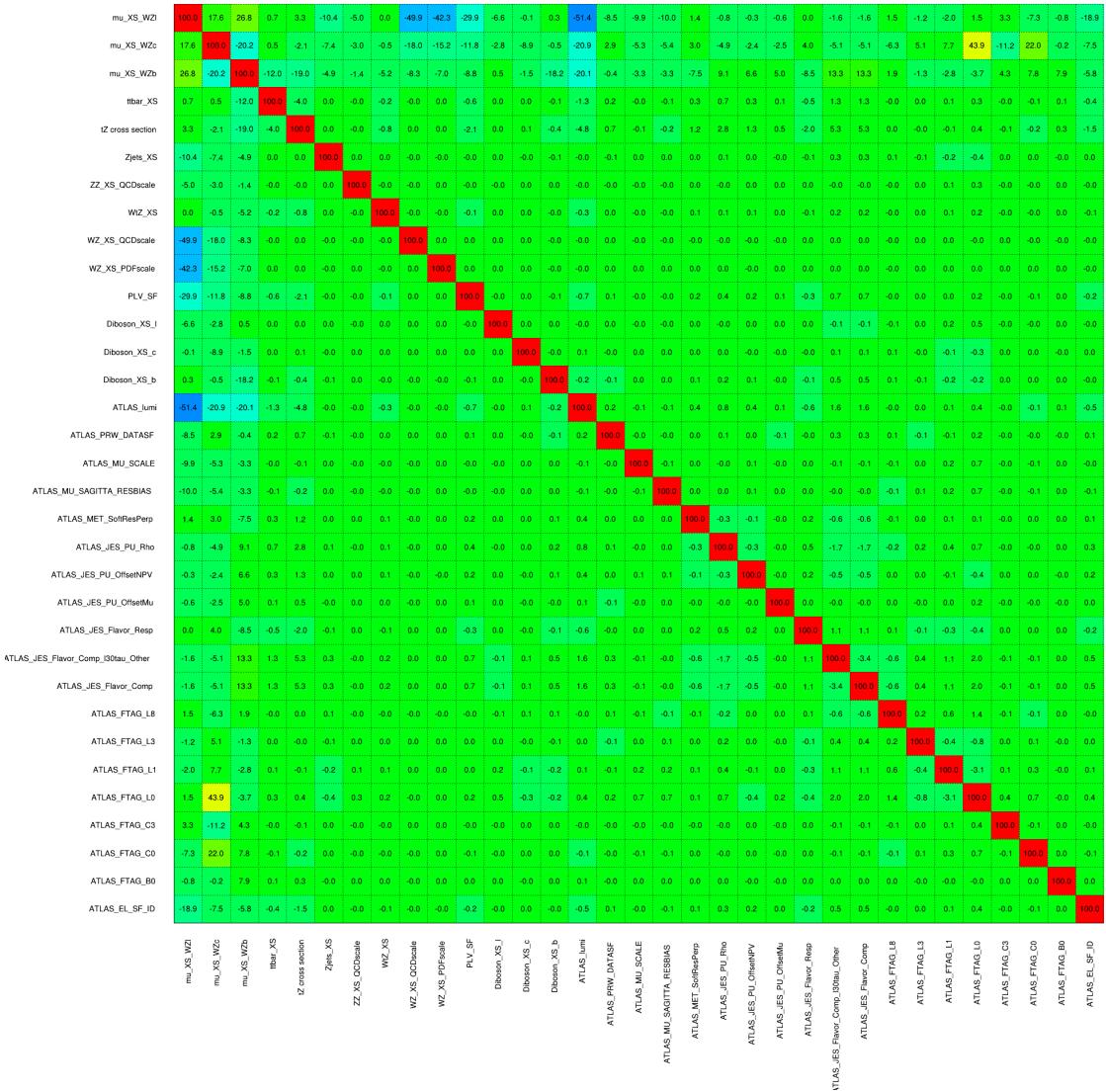


Figure 34: Correlations between nuisance parameters

The negative correlations between  $\mu_{WZ+\text{charm}}$  and  $\mu_{WZ+b}$  and  $\mu_{WZ+\text{light}}$  are expected: WZ + charm is present in both the WZ + b and WZ + light enriched regions, therefore increasing the fraction of charm requires increasing the fraction of WZ + b and WZ + light. This reasoning also explains the positive correlation between  $\mu_{WZ+b}$  and  $\mu_{WZ+\text{light}}$ .

Two of the major backgrounds in the region with the highest purity of WZ + b are tZ and Other VV + b, explaining the negative correlations between  $\mu_{WZ+b}$  and the tZ cross section, and the VV + b cross section.

467 The high correlation between the luminosity and  $\mu_{WZ+\text{light}}$  arises from the fact that the uncer-  
 468 tainty on  $\mu_{WZ+\text{light}}$  is very low (around 4%). Small changes in luminosity cause a change in  
 469 the yield of  $WZ + \text{light}$  that is large compared to its uncertainty, producing a large correlation  
 470 between these two parameters.

## 471 8.2 2-jet Fit Results

### 472 The results of the fit are currently blinded.

473 Pre-fit yields in each of the 2-jet fit regions are shown in figure 35.

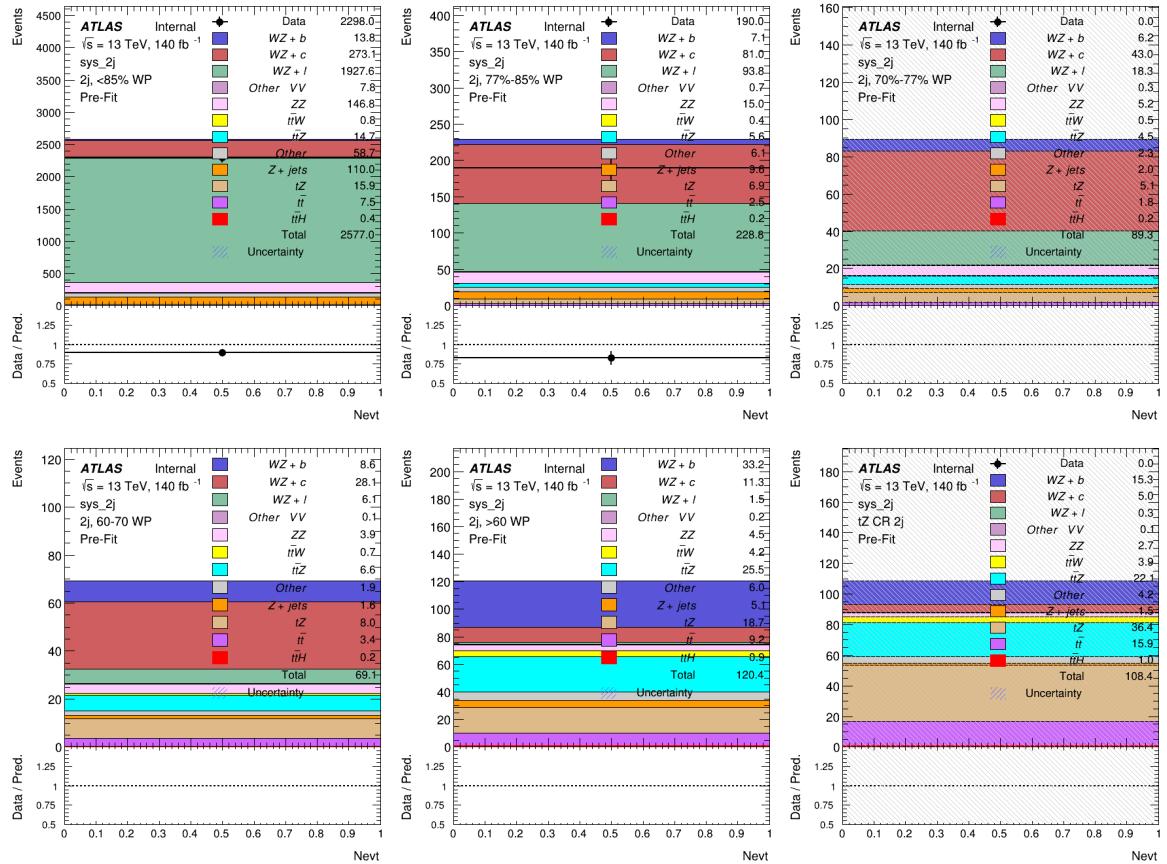


Figure 35: Data/MC yields in each of the regions in the 2-jet fit before the fit has been performed.

474 The post-fit yields in each region are summarized in figure 36.

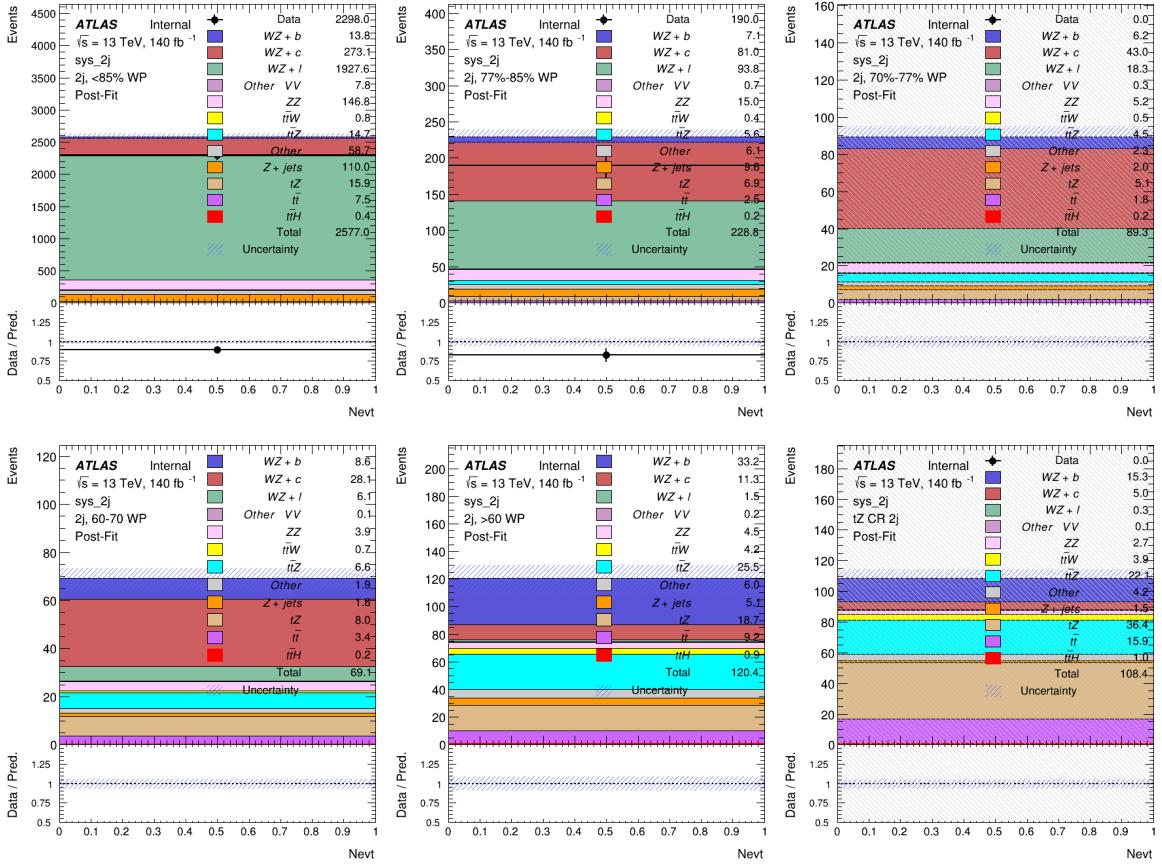


Figure 36: Data/MC results in each of the regions in the 2-jet fit after the fit has been performed.

475 A post-fit summary plot of the fitted regions is shown in figure 37:

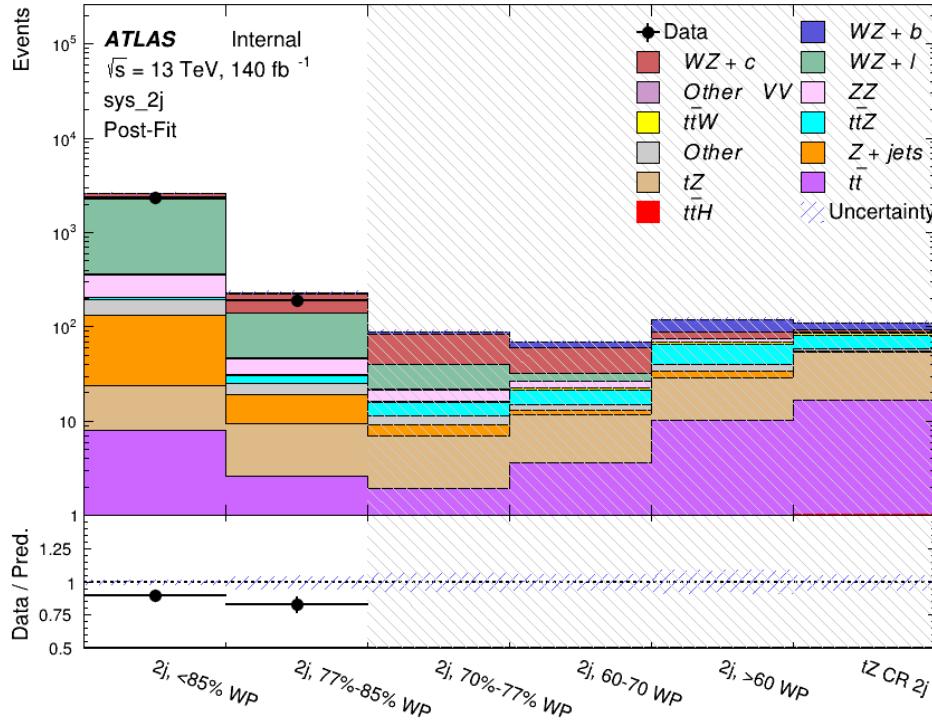


Figure 37: Post-fit summary of the fit over 2-jet regions.

<sup>476</sup> The same set of systematic uncertainties consider for the 1-jet fit are included in the 2-jet fit as  
<sup>477</sup> well. The impact of the most significant systematic uncertainties is summarized in table 14.

Uncertainty Source	$\Delta\mu$	
WZ + charm cross-section	-0.1966	0.2171
tZ cross-section	-0.1521	0.1518
WZ + light cross-section	0.1485	-0.1411
Other VV + b cross-section	-0.1115	0.1163
Flavor Tagging	0.0955	0.0957
Jet Energy Scale	0.0613	0.081
t <bar>t</bar>	-0.0662	0.0654
Luminosity	-0.0609	0.0655
Z + jets cross-section	-0.0284	0.0284
Other VV + charm cross-section	0.0207	-0.0202
Muon Trigger Scale Factor	0.019	0.0209
Total Systematic Uncertainty	0.3511	0.3679

Table 14: Summary of the most significant sources of systematic uncertainty on the measurement of WZ + b 2-jet events.

<sup>478</sup> The ranking and impact of those nuisance parameters with the largest contribution to the overall  
<sup>479</sup> uncertainty is shown in figure 38.

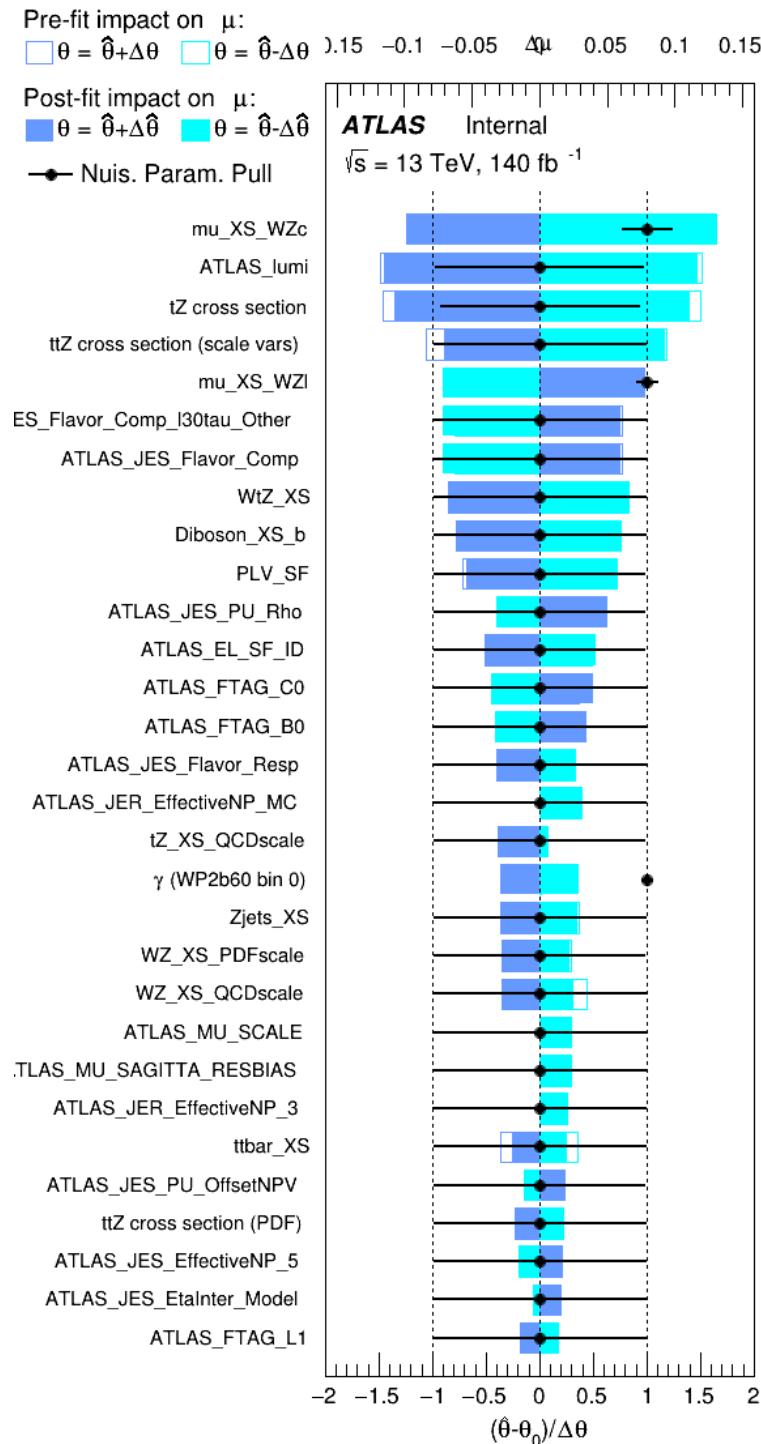


Figure 38: Impact of systematic uncertainties on the signal-strength of WZ + b in 2-jet events.

480 The large impact of the Jet Energy Scale and Jet Flavor Tagging is unsurprising, as the shape  
 481 of the fit regions depends heavily on the modeling of the jets. The other major sources of  
 482 uncertainty come from background modelling and cross-section uncertainty. The pie charts in  
 483 figure 39 show that for the modelling uncertainties that contribute most correspond to the most  
 484 significant backgrounds.

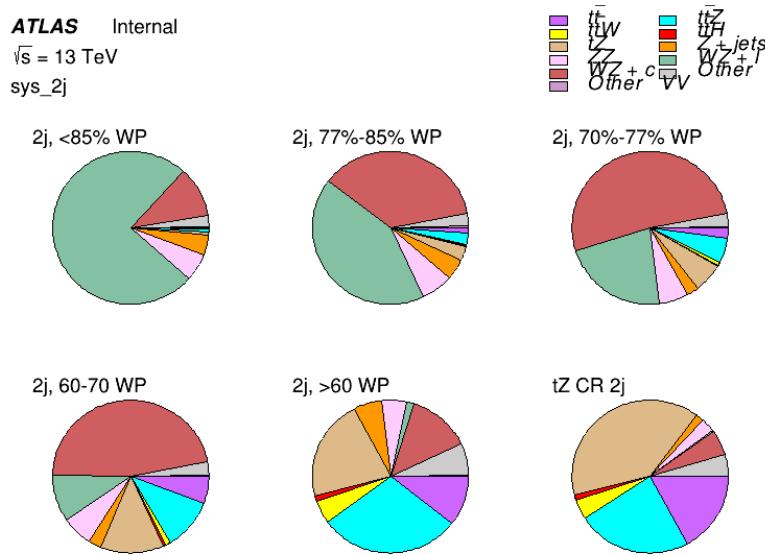


Figure 39: Post-fit background composition of the 2-jet fit regions.

485 The correlations between these nuisance parameters are summarized in figure 40.

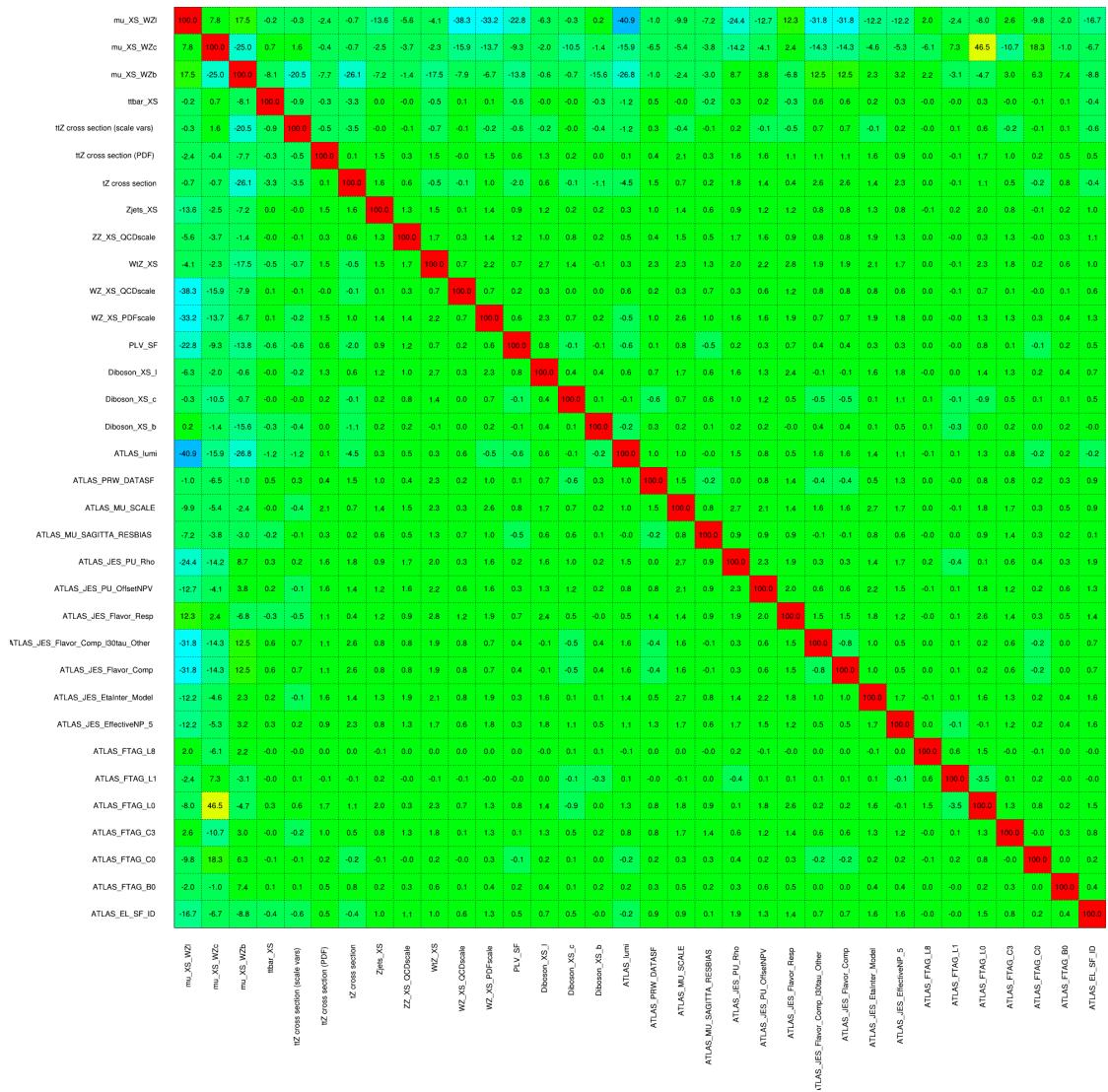


Figure 40: Correlations between nuisance parameters in the 2-jet fit

486 As in the 1-jet case, no significant, unexpected correlations are found between nuisance para-  
487 meters.

### 8.3 tZ Inclusive Results

489 While  $tZ$  is often considered as a distinct process from  $WZ + b$ ,

## 490 9 Conclusion

491 A measurement of  $WZ + \text{heavy flavor}$  is performed using  $140 \text{ fb}^{-1}$  of  $\sqrt{s} = 13 \text{ TeV}$  proton-  
492 proton collision data collected by the ATLAS detector at the LHC. **This section will be include**  
493 **final results once unblinded.**

494 **Appendices**

495 **.1 tZ Interference Studies**

496 Because it includes an on-shell Z boson as well as a b-jet and W from the top decay, tZ  
497 production represents an identical final state to WZ + b-jet. This implies the possibility of matrix  
498 level interference between these two processes not accounted for in the Monte Carlo simulations,  
499 which consider the two processes independently. Truth level studies are performed in order to  
500 estimate the impact of these interference effects.

501 In order to estimate the matrix level interference effects between tZ and WZ + b-jet, two different  
502 sets of simulations are produced using MadGraph 5 [**Madgraph**] - one which simulates these  
503 two processes independently, and another where they are produced simultaneously, such that  
504 interference effects are present. These two sets of samples are then compared, and the difference  
505 between them can be taken to represent any interference effects.

506 MadGraph simulations of 10,000 tZ and 10,000 WZ + b-jet events are produced, along with  
507 20,000 events where both are present, in the fiducial region where three leptons and at least one  
508 jet are produced.

509 A selection mimicking the preselection used in the main analysis is applied to the samples: The  
510 SS leptons are required to have  $p_T > 20$  GeV, and  $> 10$  GeV is required for the OS lepton. The  
511 associated b-jet is required to have  $p_T > 25$  GeV, and all physics objects are required to fall in a  
512 range of  $|\eta| < 2.5$ .

513 The kinematics of these samples after the selection has been applied are shown below:

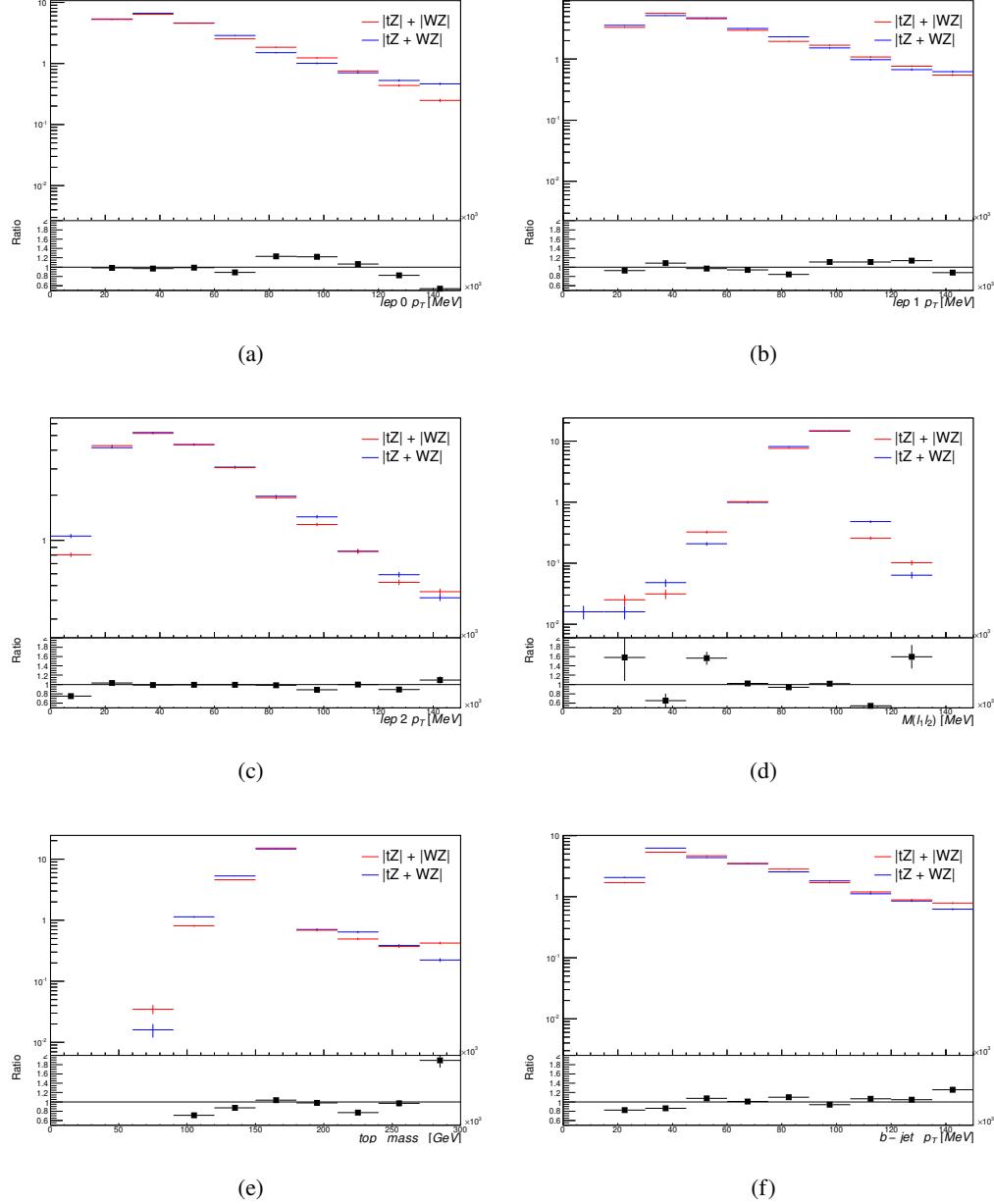


Figure 41: Comparisons between (a) the  $p_T$  of the lepton from the W, (b) and (c) show the  $p_T$  of the lepton from the Z, (d) the invariant mass of the Z-candidate, (e) the mass of the top candidate, and (f) the  $p_T$  of the b-jet, for WZ and tZ events generated with interference effects (blue) and without interference effects (red).

514 The overall cross-section of the two methods agree within error, and no significant differences  
515 in the kinematic distributions are seen. It is therefore concluded that interference effects do not

516 significantly impact the results.

517 **References**

- 518 [1] M. Aaboud et al. ‘Observation of electroweak  $W^\pm Z$  boson pair production in association  
519 with two jets in pp collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector’. In: *Phys.*  
520 *Lett.* B793 (2019), pp. 469–492. doi: [10.1016/j.physletb.2019.05.012](https://doi.org/10.1016/j.physletb.2019.05.012). arXiv:  
521 [1812.09740 \[hep-ex\]](https://arxiv.org/abs/1812.09740).
- 522 [2] T. Gleisberg et al. ‘Event generation with SHERPA 1.1’. In: *JHEP* 02 (2009), p. 007. doi:  
523 [10.1088/1126-6708/2009/02/007](https://doi.org/10.1088/1126-6708/2009/02/007). arXiv: [0811.4622 \[hep-ph\]](https://arxiv.org/abs/0811.4622).
- 524 [3] ATLAS Collaboration. *Electron efficiency measurements with the ATLAS detector using*  
525 *the 2015 LHC proton–proton collision data*. ATLAS-CONF-2016-024. 2016. URL: <https://cds.cern.ch/record/2157687>.
- 527 [4] ATLAS Collaboration. ‘Measurement of the muon reconstruction performance of the  
528 ATLAS detector using 2011 and 2012 LHC proton–proton collision data’. In: *Eur. Phys.*  
529 *J. C* 74 (2014), p. 3130. doi: [10.1140/epjc/s10052-014-3130-x](https://doi.org/10.1140/epjc/s10052-014-3130-x). arXiv: [1407.3935 \[hep-ex\]](https://arxiv.org/abs/1407.3935).
- 531 [5] *Evidence for the associated production of the Higgs boson and a top quark pair with the*  
532 *ATLAS detector*. Tech. rep. ATLAS-CONF-2017-077. Geneva: CERN, Nov. 2017. URL:  
533 <https://cds.cern.ch/record/2291405>.
- 534 [6] ATLAS Collaboration. *Jet Calibration and Systematic Uncertainties for Jets Reconstructed*  
535 *in the ATLAS Detector at  $\sqrt{s} = 13$  TeV*. ATL-PHYS-PUB-2015-015. 2015. URL:  
536 <https://cds.cern.ch/record/2037613>.
- 537 [7] ATLAS Collaboration. *Selection of jets produced in 13 TeV proton–proton collisions with*  
538 *the ATLAS detector*. ATLAS-CONF-2015-029. 2015. URL: <https://cds.cern.ch/record/2037702>.
- 540 [8] ATLAS Collaboration. ‘Performance of pile-up mitigation techniques for jets in pp col-  
541 lisions at  $\sqrt{s} = 8$  TeV using the ATLAS detector’. In: *Eur. Phys. J. C* 76 (2016), p. 581.  
542 doi: [10.1140/epjc/s10052-016-4395-z](https://doi.org/10.1140/epjc/s10052-016-4395-z). arXiv: [1510.03823 \[hep-ex\]](https://arxiv.org/abs/1510.03823).
- 543 [9] ATLAS Collaboration. *Performance of missing transverse momentum reconstruction with*  
544 *the ATLAS detector in the first proton–proton collisions at  $\sqrt{s} = 13$  TeV*. ATL-PHYS-  
545 PUB-2015-027. 2015. URL: <https://cds.cern.ch/record/2037904>.
- 546 [10] P. S. A. Hoecker. ‘TMVA 4 Toolkit for Multivariate Data Analysis with ROOT’. In:  
547 *arXiv:physics/0703039* (2013).
- 548 [11] F. Cardillo et al. ‘Measurement of the fiducial and differential cross-section of a top quark  
549 pair in association with a Z boson at 13 TeV with the ATLAS detector’. In: ATL-COM-  
550 PHYS-2019-334 (Apr. 2019). URL: <https://cds.cern.ch/record/2672207>.
- 551 [12] ATLAS Collaboration. ‘Luminosity determination in pp collisions at  $\sqrt{s} = 7$  TeV using  
552 the ATLAS detector at the LHC’. In: *Eur. Phys. J. C* 71 (2011), p. 1630. doi: [10.1140/epjc/s10052-011-1630-5](https://doi.org/10.1140/epjc/s10052-011-1630-5). arXiv: [1101.2185 \[hep-ex\]](https://arxiv.org/abs/1101.2185).

- 554 [13] G. Aad et al. ‘Jet energy resolution in proton-proton collisions at  $\sqrt{s} = 7$  TeV recorded  
555 in 2010 with the ATLAS detector’. In: *The European Physical Journal C* 73.3 (Mar.  
556 2013), p. 2306. ISSN: 1434-6052. DOI: [10.1140/epjc/s10052-013-2306-0](https://doi.org/10.1140/epjc/s10052-013-2306-0). URL:  
557 <https://doi.org/10.1140/epjc/s10052-013-2306-0>.
- 558 [14] A. Collaboration. ‘Performance of b -jet identification in the ATLAS experiment’. In:  
559 *Journal of Instrumentation* 11.04 (2016), P04008. URL: <http://stacks.iop.org/1748-0221/11/i=04/a=P04008>.  
560