



## ATLAS Note

GROUP-2017-XX

1st February 2021



Draft version 0.1

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>3</b>
13	1.1 Changelog	3
14	1.1.1 Changes relative to v3	3
15	1.1.2 Changes relative to v2	4
16	1.1.3 Changes relative to v1	4
17	1.2 Outstanding Items	4
18	<b>2 Executive Summary</b>	<b>5</b>
19	<b>3 Data and Monte Carlo Samples</b>	<b>5</b>
20	3.1 Data Samples	6
21	3.2 Monte Carlo Samples	6
22	<b>4 Object Reconstruction</b>	<b>7</b>
23	4.1 Trigger	7
24	4.2 Light leptons	7
25	4.3 Jets	8
26	4.4 B-tagged Jets	10
27	4.5 Missing transverse energy	11
28	<b>5 Event Selection and Signal Region Definitions</b>	<b>11</b>
29	5.1 Event Preselection	11
30	5.2 Fit Regions	14
31	5.3 Non-Prompt Lepton Estimation	29
32	5.3.1 $t\bar{t}$ Validation	29
33	5.3.2 Z+jets Validation	33
34	<b>6 tZ Interference Studies and Separation Multivariate Analysis</b>	<b>37</b>
35	6.1 Interference Studies	38
36	6.2 Top Mass Reconstruction	39
37	6.3 tZ BDT	40
38	<b>7 Systematic Uncertainties</b>	<b>43</b>
39	<b>8 Results</b>	<b>46</b>
40	8.1 1-jet Fit Results	47
41	8.2 2-jet Fit Results	53
42	<b>9 Conclusion</b>	<b>57</b>

---

43 **1 Changes and outstanding items**

44 **1.1 Changelog**

45 This is version 4

46 **1.1.1 Changes relative to v3**

- 47 • Merged introduction into executive summary, including unblinding details and list of  
48 SRs/CRs used
- 49 • listed ptag used (p4133), and release (AB 21.2.127)
- 50 • Included table reftab:xsecUnc listing x-sec uncertainties used
- 51 • Removed selection criteria listed in table 4 (QMisID, AmbiguityType) that were removed  
52 from the analysis
- 53 • specified variable used to make truth jet flavor determination (HadronConeExclTruthLa-  
54 belIID)
- 55 • fixed bug in MtLepMet calculation, updated selection/fits to account for this
- 56 • Included plots of MtLepMet and PtZ, swapped lep 1 and 2  $p_T$  plots for lep W and lep Z  
57 plots
- 58 • updated tZ BDT training to reduce overfitting, updated plots to include error bars, feature  
59 importance
- 60 • updated table 8 to clarify selection, fix the tZ\_BDT cut used
- 61 • replace a few broken ntuples which included large weight events
- 62 • include DL1r distribution for Z+jets and t $\bar{t}$  VRs
- 63 • Expanded section on fakes, included information on derived scale factors from VRs.
- 64 • Changed the kinematic plots to include  $p_T(Z)$  and  $m_T(W)$ , list lepton  $p_T$  based on W and  
65 Z candidates.

66 **1.1.2 Changes relative to v2**

- 67     • Added alternate VBS samples to include missing b-jet diagrams  
68     • Included a section on tZ interference effects, [6.1](#).  
69     • Updated to reflect changes for 2018, including the move to PFlow jets, DL1r, updated  
70       trigger, and updated AnalysisBase version (now 21.2.127)  
71     • Revised fit regions, using separate 1-jet and 2-jet fits, with all 2-j regions included  
72     • updated plots for tZ BDT, added details about the model  
73     • Included truth jet information

74 **1.1.3 Changes relative to v1**

- 75     • Added GRL list  
76     • Fixed latex issue in line 92, typo in line 172  
77     • Added tables [6](#) and [4](#), summarizing the event and object selection  
78     • Added table [2](#), which includes the DSID of samples used  
79     • Included reference to WZ inclusive paper in introduction

80 **1.2 Outstanding Items**

- 81     • Unblind, update plots and fits to include data  
82     • Include truth jet studies  
83     • Add cross-section, significance once unblinded

## 84 2 Executive Summary

85 The production of  $WZ$  in association with a heavy flavor jet represents an important background  
 86 for many major analyses. This includes any process with leptons and b-jets in the final state,  
 87 such as  $t\bar{t}H$ ,  $t\bar{t}W$ , and  $t\bar{t}Z$ . While precise measurements have been made of  $WZ$  production  
 88 [1],  $WZ +$  heavy flavor remains poorly understood. This is largely because the QCD processes  
 89 involved in the production of the b-jet make it difficult to simulate accurately. This introduces a  
 90 large uncertainty for analyses that include this process as a background.

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

96 Events that meet this selection criteria are sorted into pseudo-continuous b-tagging regions based  
 97 on the DL1r b-tag score of their associated jets. This is done to separate  $WZ +$  b-jet events  
 98 from  $WZ +$  charm and  $WZ +$  light jets. These orthogonal categories, in addition to a tZ Control  
 99 Region formed using an MVA, are the Signal Regions for the analysis. These regions are fit to  
 100 data in order make a more accurate estimate of the contribution of  $WZ +$  heavy-flavor, where  
 101 heavy-flavor jets include b-jets and charm jets. Separate fits are performed for 1-jet and 2-jet  
 102 events. The full Run-2 dataset collected by the ATLAS detector, representing  $139 \text{ fb}^{-1}$  of data  
 103 from pp collisions at  $\sqrt{s} = 13 \text{ TeV}$ , is used for this study.

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

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

113 The current state of the analysis shows blinded results for the full 2018 dataset. Regions containing  
 114  $>5\%$   $WZ+b$  events are blinded, and results are from Asimov, MC only fits. In addition to adding  
 115 some additional information to this note, remaining tasks include performing  $WZ/tZ$  interference  
 116 studies, finalizing the presentation of results, and unblinding.

## 117 3 Data and Monte Carlo Samples

118 Both data and Monte Carlo samples used in this analysis were prepared in the xAOD format,  
 119 which was used to produce a Dx AOD sample in the HIGG8D1 derivation framework. The HIGG8D1

120 framework is designed for the  $t\bar{t}H$  multi-lepton analysis, which targets events with multiple  
 121 leptons as well as tau hadrons. This framework skims the dataset to remove unneeded variables  
 122 as well as entire events. Events are removed from the derivations that do not meet the following  
 123 selection:

- 124 • at least two light leptons within a range  $|\eta| < 2.6$ , with leading lepton  $p_T > 15$  GeV and  
 125 subleading lepton  $p_T > 5$  GeV
- 126 • at least one light lepton with  $p_T > 15$  GeV within a range  $|\eta| < 2.6$ , and at least two hadronic  
 127 taus with  $p_T > 15$  GeV.

128 Samples were then generated from these HIGG8D1 derivations with p-tag of p4134 using a  
 129 modified version of AnalysisBase version 21.2.127.

### 130 3.1 Data Samples

131 The study uses a sample of proton-proton collision data collected by the ATLAS detector from  
 132 2015 through 2018 at an energy of  $\sqrt{s} = 13$  TeV, which represents an integrated luminosity of  
 133  $139 \text{ fb}^{-1}$ . This data set was collected with a bunch crossing rate of 25 ns. All data used in this  
 134 analysis was verified by data quality checks, having been included in the following Good Run  
 135 Lists:

- 136 • data15\_13TeV.periodAllYear\_DetStatus-v79-repro20-02\_DQDefects-00-02-02  
 137 \_PHYS\_StandardGRL\_All\_Good\_25ns.xml
- 138 • data16\_13TeV.periodAllYear\_DetStatus-v88-pro20-21\_DQDefects-00-02-04  
 139 \_PHYS\_StandardGRL\_All\_Good\_25ns.xml
- 140 • data17\_13TeV.periodAllYear\_DetStatus-v97-pro21-13\_Unknown\_PHYS\_StandardGRL  
 141 \_All\_Good\_25ns\_Triggerno17e33prim.xml
- 142 • data18\_13TeV.periodAllYear\_DetStatus-v102-pro22-04\_Unknown\_PHYS\_StandardGRL  
 143 \_All\_Good\_25ns\_Triggerno17e33prim.xml

144 Runs included from the AllYear period containers are included.

### 145 3.2 Monte Carlo Samples

146 Several different generators were used to produce Monte Carlo simulations of the signal and  
 147 background processes. For all samples, the response of the ATLAS detector is simulated using  
 148 Geant4. The WZ signal samples are simulated using Sherpa 2.2.2 [2]. Specific information  
 149 about the Monte Carlo samples being used can be found in table 1. A list of the specific samples  
 150 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)
t̄(Z/γ* → ll)	MG5_AMC	NLO	PYTHIA 8	NNPDF 3.0 NLO
t̄H	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
t̄W+W-	MG5_AMC	LO	PYTHIA 8	NNPDF 2.3 LO
t̄t	Powheg-BOX v2 [ <b>powheggtt</b> ]	NLO	PYTHIA 8	NNPDF 3.0 NLO
t̄t̄γ	MG5_AMC	LO	PYTHIA 8	NNPDF 2.3 LO
s-, t-channel, Wt single top	Powheg-BOX v1 [ <b>powhegstp</b> ]	NLO	PYTHIA 6	CT10
qqVV, VVV Z → l+l-	SHERPA 2.2.1	MEPS NLO	SHERPA	NNPDF 3.0 NLO

## 4 Object Reconstruction

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

### 4.1 Trigger

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

### 4.2 Light leptons

Electron candidates are reconstructed from energy clusters in the electromagnetic calorimeter that are associated with charged particle tracks reconstructed in the inner detector [3]. Electron candidates are required to have  $p_T > 10$  GeV and  $|\eta_{\text{cluster}}| < 2.47$ . Candidates in the transition region between different electromagnetic calorimeter components,  $1.37 < |\eta_{\text{cluster}}| < 1.52$ , are rejected. A multivariate likelihood discriminant combining shower shape and track information is used to distinguish real electrons from hadronic showers (fake electrons). To further reduce

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.

164 the non-prompt electron contribution, the track is required to be consistent with originating from  
 165 the primary vertex; requirements are imposed on the transverse impact parameter significance  
 166 ( $|d_0|/\sigma_{d_0}$ ) and the longitudinal impact parameter ( $|\Delta z_0 \sin \theta_\ell|$ ), as shown in table 4.

167 Muon candidates are reconstructed by combining inner detector tracks with track segments or  
 168 full tracks in the muon spectrometer [4]. Muon candidates are required to have  $p_T > 10$  GeV  
 169 and  $|\eta| < 2.5$ .

170 All leptons are required to be isolated, and pass a non-prompt BDT selection described in detail  
 171 in [5].

### 172 4.3 Jets

173 Jets are reconstructed from calibrated topological clusters built from energy deposits in the  
 174 calorimeters [6], using the anti- $k_t$  algorithm with a radius parameter  $R = 0.4$ . Jets with energy  
 175 contributions likely arising from noise or detector effects are removed from consideration [7],

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.

	$e$			$\mu$						
	L	$L^*$	T	L	$L^*$	T				
FixedCutLoose	No	Yes		No	Yes					
Non-prompt lepton BDT	No	Yes		No	Yes					
Identification	Loose	Tight		Loose	Medium					
Transverse impact parameter significance $ d_0 /\sigma_{d_0}$	$< 5$			$< 3$						
Longitudinal impact parameter $ z_0 \sin \theta $	$< 0.5$ mm									

Table 4: Loose (L), loose and minimally-isolated ( $L^*$ ), and tight (T) light lepton definitions.

and only jets satisfying  $p_T > 25$  GeV and  $|\eta| < 2.5$  are used in this analysis. For jets with  $p_T < 60$  GeV and  $|\eta| < 2.4$ , a jet-track association algorithm is used to confirm that the jet originates from the selected primary vertex, in order to reject jets arising from pileup collisions [8].

180 **4.4 B-tagged Jets**

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

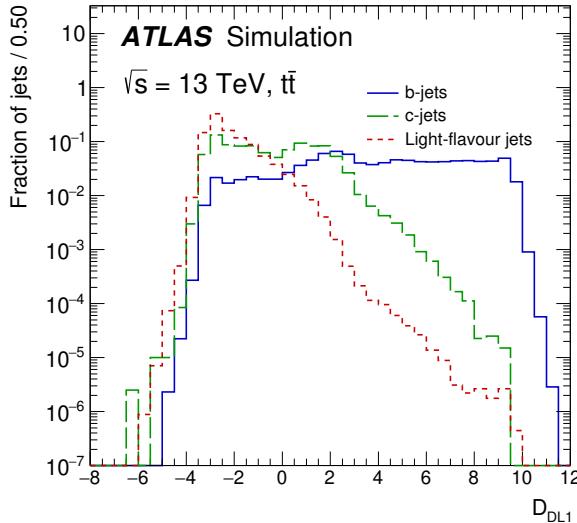


Figure 1: Output distribution of the DL1r algorithm for b-jets, charm jets, and light jets

186 From the output of the BDT, working points (WPs) are developed based on the efficiency of truth  
 187 b-jets at particular values of the DL1r algorithm. The working points used in this analysis are  
 188 summarized in table 5.

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

Table 5: B-tagging Working Points by tightness and b-jet efficiency

189 A tighter WP will accept fewer b-jets, but reject a higher fraction of charm and light jets.  
 190 Generally, analyses that include b-jets will use a fixed working point, for example, requiring that  
 191 a jet pass the 70% threshold. By instead treating these working point as bins, e.g. events with  
 192 jets that fall between the 85% and 77% WPs fall into one bin, while events with jets passing the  
 193 60% WP fall into another, and looking at the full psuedo-continuous DL1r spectrum of the jets,  
 194 additional information can be gained. The psuedo-continuous b-tag spectrum is used in this case  
 195 to separate out  $WZ + b$ ,  $WZ + \text{charm}$ , and  $WZ + \text{light}$ .

---

<sup>196</sup> **4.5 Missing transverse energy**

<sup>197</sup> Missing transverse momentum ( $E_T^{\text{miss}}$ ) is used as part of the event selection. The missing  
<sup>198</sup> transverse momentum vector is defined as the inverse of the sum of the transverse momenta of  
<sup>199</sup> all reconstructed physics objects as well as remaining unclustered energy, the latter of which is  
<sup>200</sup> estimated from low- $p_T$  tracks associated with the primary vertex but not assigned to a hard object,  
<sup>201</sup> with object definitions taken from [9]. Light leptons considered in the  $E_T^{\text{miss}}$  reconstruction are  
<sup>202</sup> required to have  $p_T > 10$  GeV, while jets are required to have  $p_T > 20$  GeV.

<sup>203</sup> **5 Event Selection and Signal Region Definitions**

<sup>204</sup> Events are required to pass a preselection described in section 5.1 and summarized in table 6.  
<sup>205</sup> Those that pass this preselection are divided into various fit regions described in section 5.2,  
<sup>206</sup> based on the number of jets in the event, and the b-tag score of those jets.

<sup>207</sup> **5.1 Event Preselection**

<sup>208</sup> Events are required to include exactly three reconstructed light leptons passing the requirement  
<sup>209</sup> described in 4.2, which have a total charge of  $\pm 1$ . As the opposite sign lepton is found to  
<sup>210</sup> be prompt the vast majority of the time [5], it is required to be loose and isolated, as defined  
<sup>211</sup> through the standard `isolationFixedCutLoose` working point supported by combined per-  
<sup>212</sup> formance groups. The same sign leptons are required to be very tight, as per the recommended  
<sup>213</sup> `isolationFixedCutTight`.

<sup>214</sup> The leptons are ordered in the analysis code as 0, 1, and 2. Lepton 0 is the lepton whose charge  
<sup>215</sup> is opposite the other two. Lepton 1 is the lepton closest to the opposite charge lepton, i.e. the  
<sup>216</sup> smallest  $\Delta R$ , leaving lepton 2 as the lepton further from the opposite charge lepton. Lepton 0  
<sup>217</sup> is required to have  $p_T > 10$  GeV, while the same sign leptons, 1 and 2, are required to have  
<sup>218</sup>  $p_T > 20$  GeV to reduce the contribution of non-prompt leptons.

<sup>219</sup> The invariant mass of at least one pair of opposite sign, same flavor leptons is required to fall  
<sup>220</sup> within 10 GeV of the mass of the Z boson, 91.2 GeV. Events where one of the opposite sign pairs  
<sup>221</sup> have an invariant mass less than 12 GeV are rejected in order to suppress low mass resonances.

<sup>222</sup> An additional requirement is placed on the missing transverse energy,  $E_T^{\text{miss}} > 20$  GeV, and the  
<sup>223</sup> transverse mass of the W candidate,  $m(E_T^{\text{miss}} + l_{\text{other}}) > 30$  GeV, where  $E_T^{\text{miss}}$  is the missing  
<sup>224</sup> transverse energy, and  $l_{\text{other}}$  is the lepton not included in the Z-candidate.

<sup>225</sup> Events are required to have one or two reconstructed jets passing the selection described in  
<sup>226</sup> section 4.3. Events with more than two jets are rejected in order to reduce the contribution of  
<sup>227</sup> 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 same-charge leptons with $p_T > 20 \text{ GeV}$
One opposite charge lepton with $p_T > 10 \text{ GeV}$
$m(l^+l^-)$ within 10 GeV of 91.2 GeV
Transverse mass of W-candidate, $m_T(E_T^{\text{miss}} + \text{lep}_{\text{other}}) > 30 \text{ GeV}$
Missing transverse energy, $E_T^{\text{miss}} > 20 \text{ GeV}$
One or two jets with $p_T > 25 \text{ GeV}$

Table 6: Summary of the selection applied to events for inclusion in the fit

228 The event yields in the preselection region for both data and Monte Carlo are summarized in  
 229 table 5.1, which shows good agreement between data and Monte Carlo, and demonstrates that  
 230 this region consists primarily of WZ events. The WZ events are split into WZ + b, WZ + c, and  
 231 WZ + 1 based on the truth flavor of the heaviest associated jet in the event. Specifically, this  
 232 determination is made based on the HadronConeExclTruthLabelID of the jet. That is, WZ +  
 233 1 events contain no charm and b jets at truth level, WZ + c contain at least one truth charm and  
 234 no b-jets, and WZ + b contains at least one truth b-jet.

Process	Events
$WZ + b$	$167.64 \pm 6.45$
$WZ + c$	$1080.91 \pm 39.28$
$WZ + l$	$7223.37 \pm 312.53$
Other VV	$849.79 \pm 142.13$
$t\bar{t}W$	$16.81 \pm 2.31$
$t\bar{t}Z$	$114.68 \pm 17.40$
rare Top	$2.20 \pm 0.14$
Single top	$0.10 \pm 0.45$
Three top	$0.01 \pm 0.01$
Four top	$0.02 \pm 0.01$
$t\bar{t}WW$	$0.23 \pm 0.05$
$Z + \text{jets}$	$601.16 \pm 260.13$
$V + \gamma$	$36.51 \pm 54.34$
$tZ$	$194.64 \pm 65.74$
$tW$	$5.49 \pm 1.24$
$WtZ$	$25.80 \pm 1.07$
$VVV$	$26.21 \pm 0.87$
$VH$	$94.34 \pm 7.35$
$t\bar{t}$	$107.68 \pm 8.14$
$t\bar{t}H$	$4.28 \pm 0.46$
Total	$10556.8 \pm 533.4$
Data	10574

Table 7: Events yields in the preselection region at  $138.9 \text{ fb}^{-1}$ 

<sup>235</sup> Here Other VV represents diboson processes other than WZ, and consists predominantly of  
<sup>236</sup>  $ZZ \rightarrow ll\bar{l}\bar{l}$  events where one of the leptons is not reconstructed.

<sup>237</sup> Simulations are further validated by comparing the kinematic distributions of the Monte Carlo  
<sup>238</sup> with data, which are shown in figures 2. Here, bins with 5% or more WZ+b are blinded.

### WZ Fit Region - Inclusive

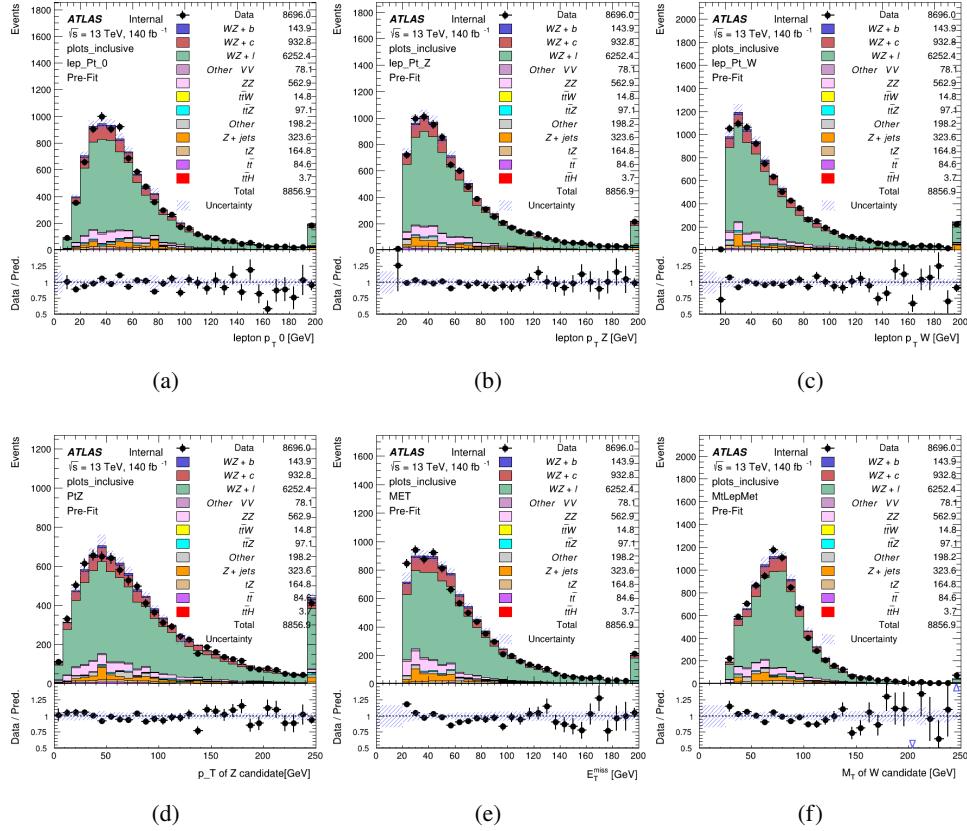


Figure 2: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the Z candidate, (c) the  $p_T$  of the lepton from the W candidate, (d) the  $p_T$  of the Z candidate, (e) the  $E_T^{\text{miss}}$ , 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 8.

The working points discussed in section 4.4 are used to separate events into fit regions based on the highest working point reached by a jet in each event. Because the background composition differs significantly based on the number of b-jets, events are further subdivided into 1 jet and 2 jet regions in order to minimize the impact of background uncertainties.

An additional tZ control region is created based on the BDT described in section 6. The region with 1-jet passing the 60% working point is split in two - a signal enriched region of events with a BDT score greater than 0.03, and a tZ control region including events with less than 0.03. This

Table 8: 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	$N_{\text{jets}} = 2, n\text{Jets\_DL1r\_60} >= 1, tZ \text{ BDT} < 0.725$

249 cutoff is arrived at by performing an Asimov fit with a variety of cutoffs, and selecting the value  
 250 that produces the highest significance for the measurement of  $WZ + b$ .

251 The modeling in each region is validated by comparing data and MC predictions for various  
 252 kinematic distributions. These plot are shown in figures 3-16.

## WZ Fit Region - 1j Inclusive

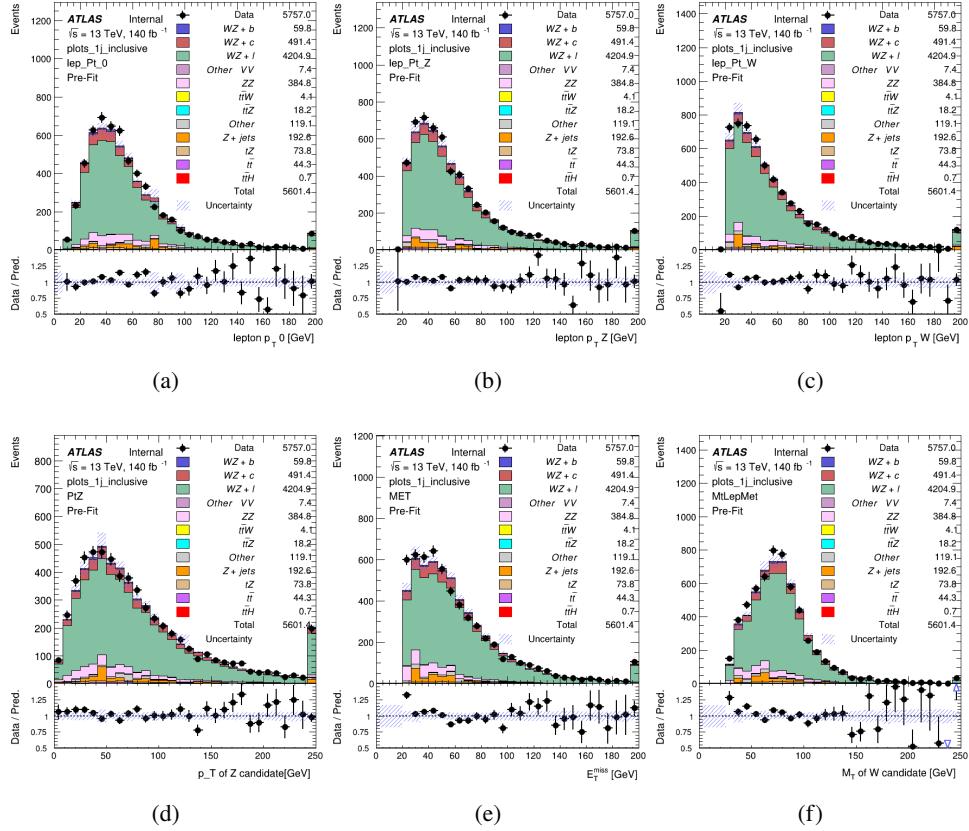


Figure 3: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the  $Z$  candidate, (c) the  $p_T$  of the lepton from the  $W$  candidate, (d) the  $p_T$  of the  $Z$  candidate, (e) the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

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

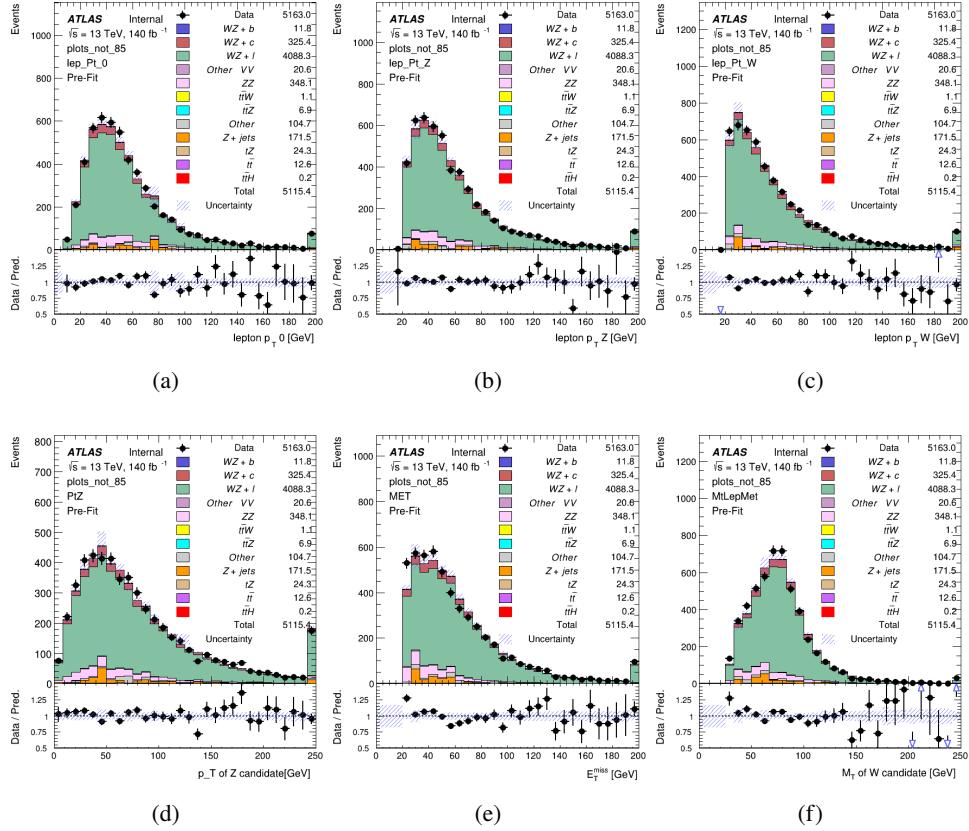


Figure 4: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the  $Z$  candidate, (c) the  $p_T$  of the lepton from the  $W$  candidate, (d) the  $p_T$  of the  $Z$  candidate, (e) the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

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

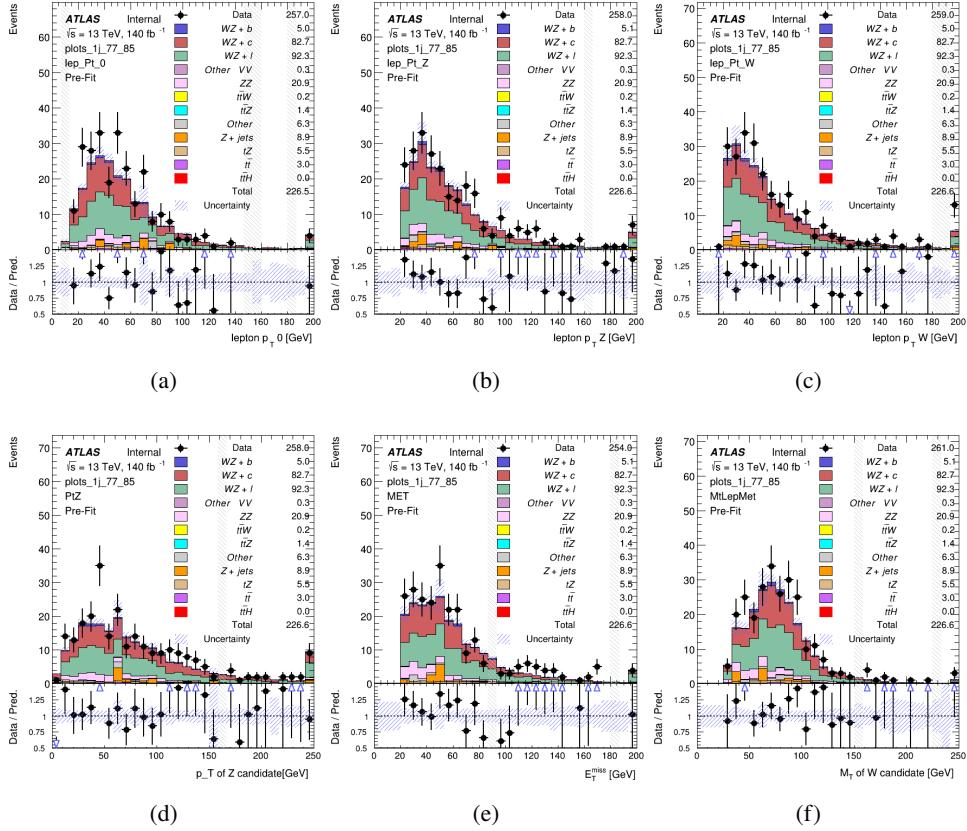


Figure 5: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the Z candidate, (c) the  $p_T$  of the lepton from the W candidate, (d) the  $p_T$  of the Z candidate, (e) the  $E_T^{\text{miss}}$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

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

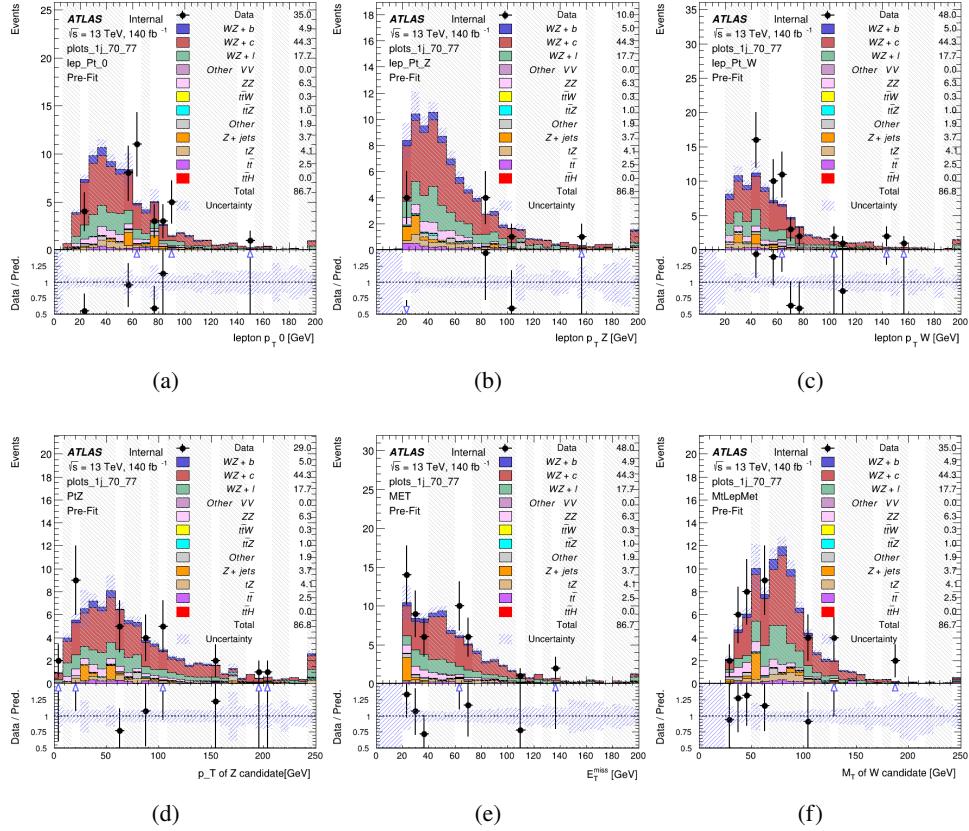


Figure 6: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the  $Z$  candidate, (c) the  $p_T$  of the lepton from the  $W$  candidate, (d) the  $p_T$  of the  $Z$  candidate, (e) the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

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

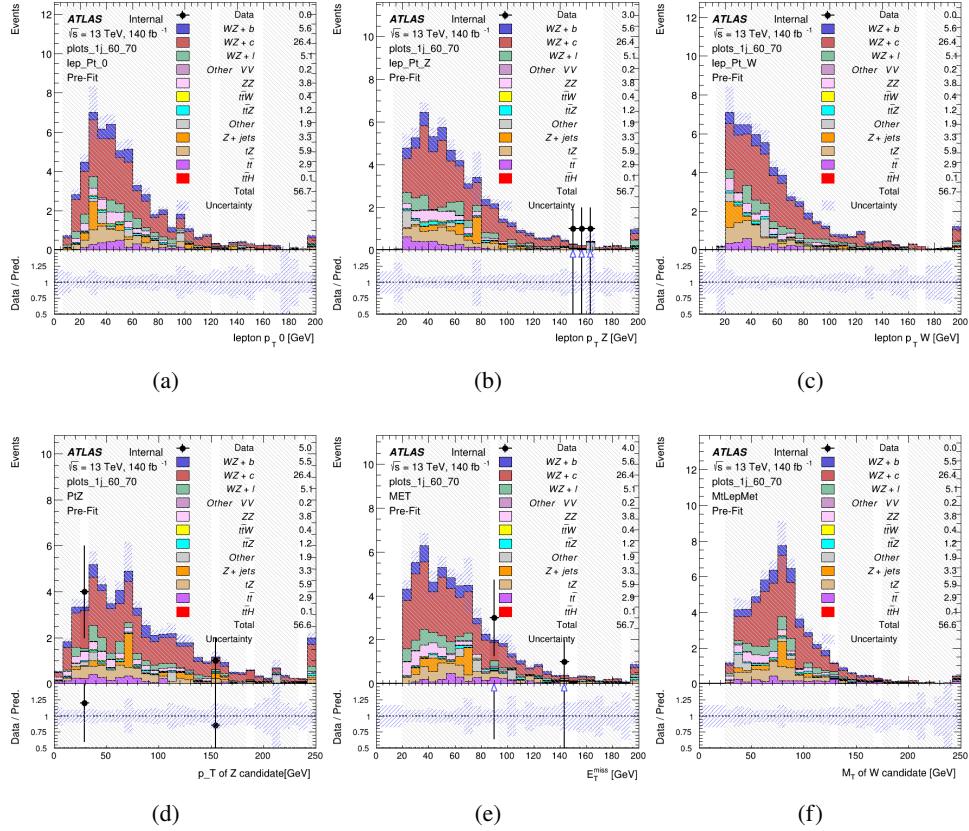


Figure 7: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the  $Z$  candidate, (c) the  $p_T$  of the lepton from the  $W$  candidate, (d) the  $p_T$  of the  $Z$  candidate, (e) the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

## WZ Fit Region - 1j 60% WP

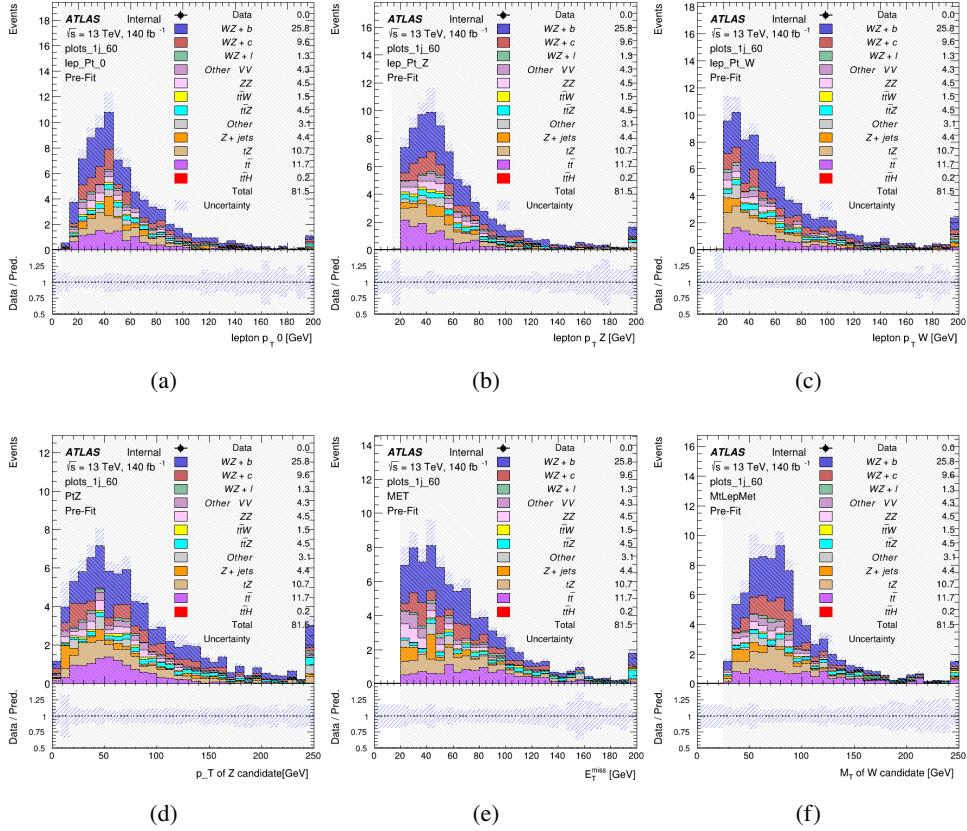


Figure 8: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the Z candidate, (c) the  $p_T$  of the lepton from the W candidate, (d) the  $p_T$  of the Z candidate, (e) the  $E_T^{\text{miss}}$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

## WZ Fit Region - tZ-CR

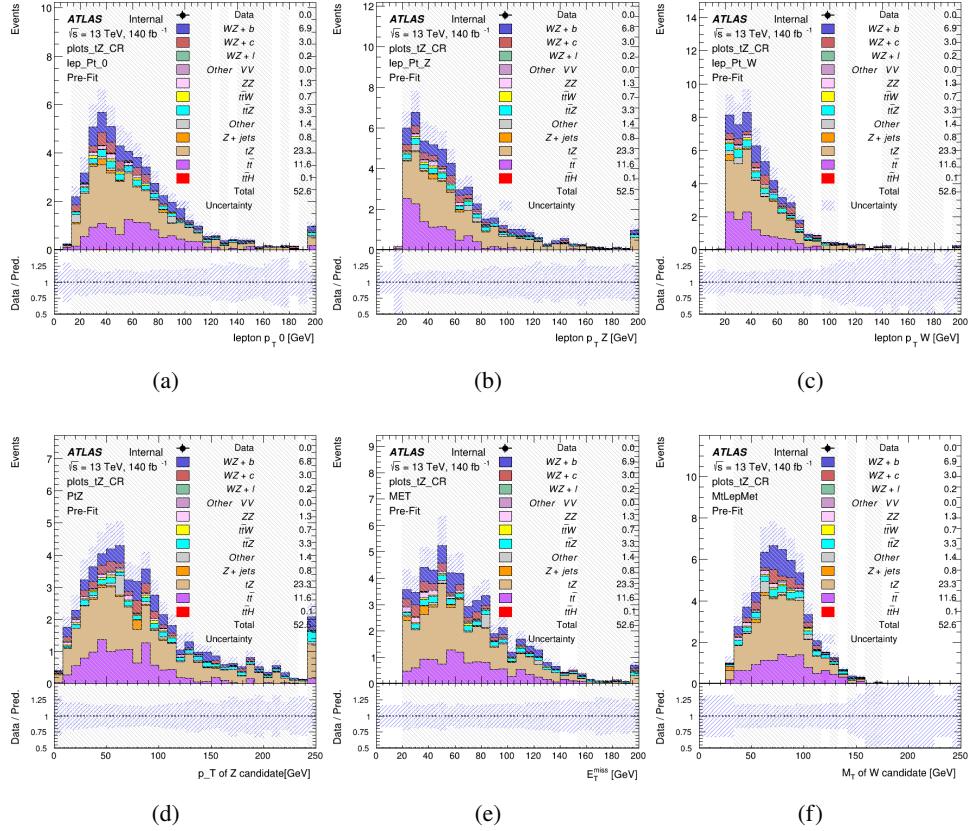


Figure 9: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the Z candidate, (c) the  $p_T$  of the lepton from the W candidate, (d) the  $p_T$  of the Z candidate, (e) the  $E_T^{\text{miss}}$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

## WZ Fit Region - 2j Inclusive

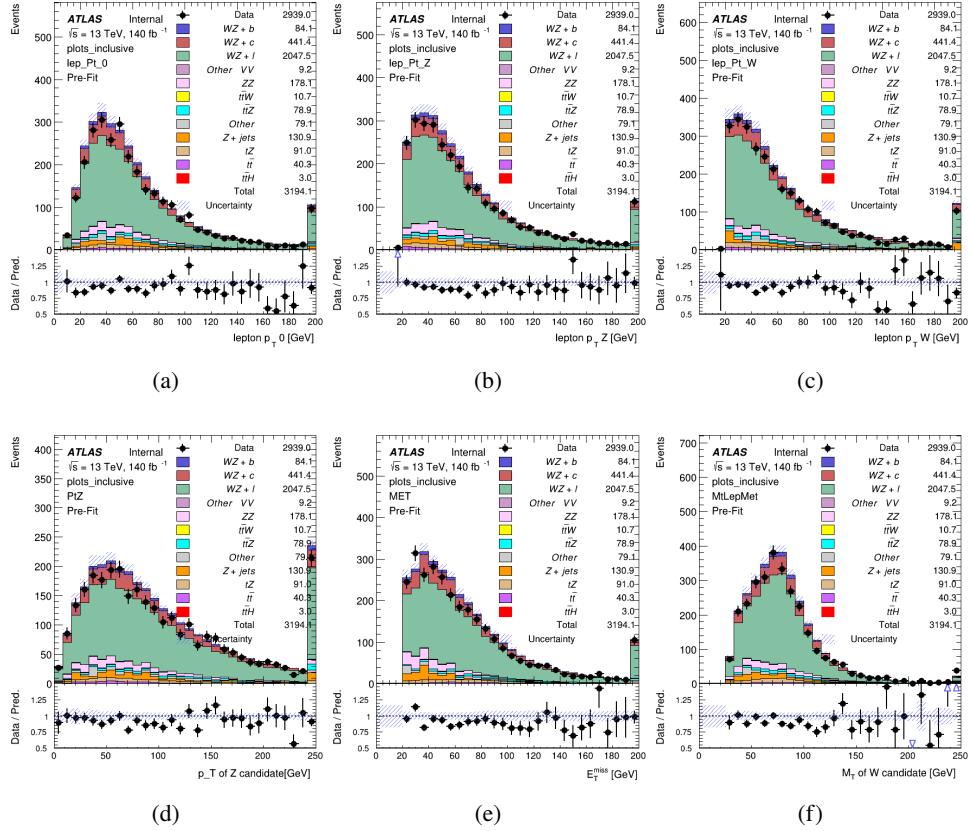


Figure 10: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the Z candidate, (c) the  $p_T$  of the lepton from the W candidate, (d) the  $p_T$  of the Z candidate, (e) the  $E_T^{\text{miss}}$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

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

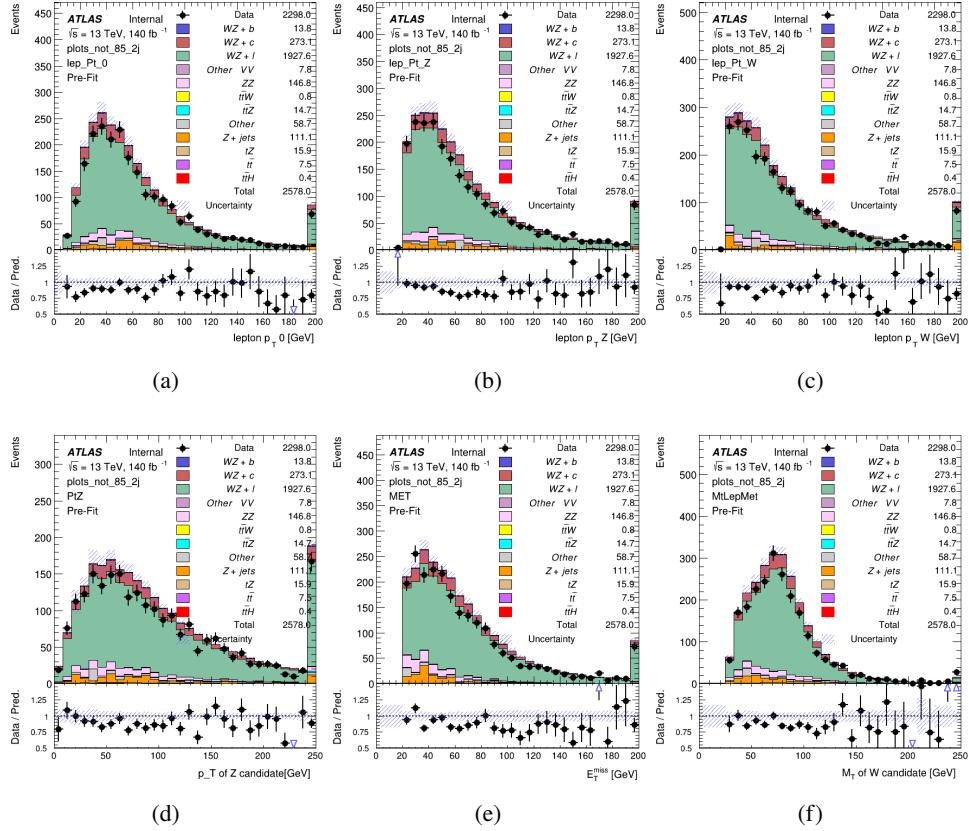


Figure 11: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the Z candidate, (c) the  $p_T$  of the lepton from the W candidate, (d) the  $p_T$  of the Z candidate, (e) the  $E_T^{\text{miss}}$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

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

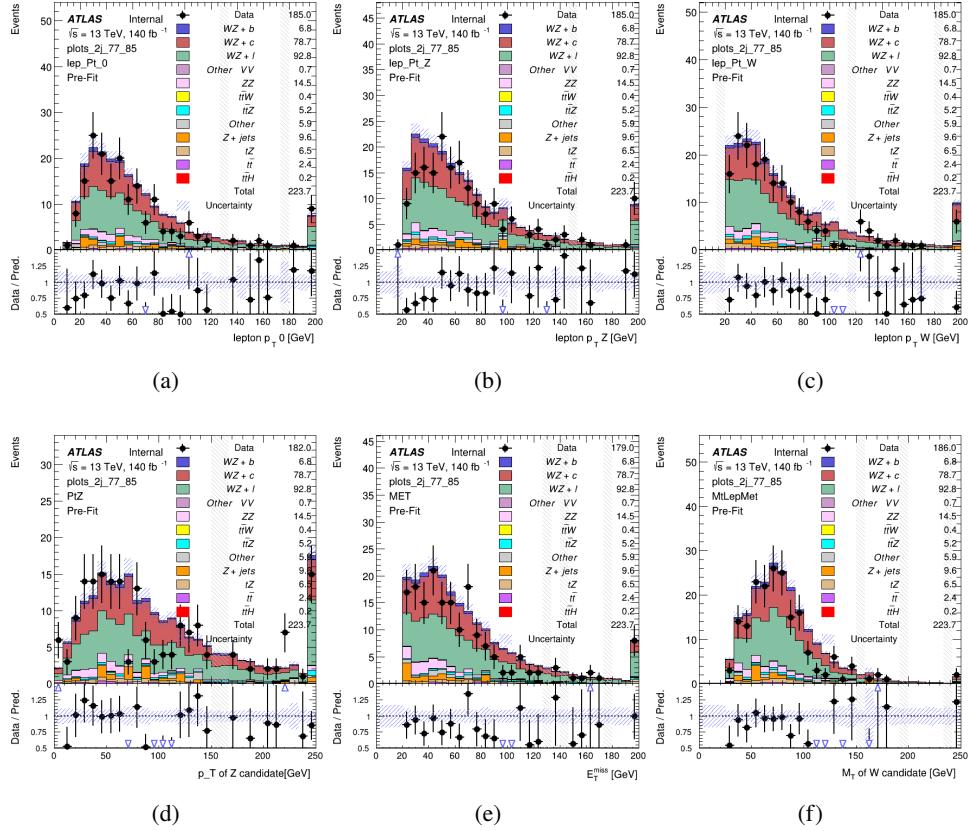


Figure 12: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the Z candidate, (c) the  $p_T$  of the lepton from the W candidate, (d) the  $p_T$  of the Z candidate, (e) the  $E_T^{\text{miss}}$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

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

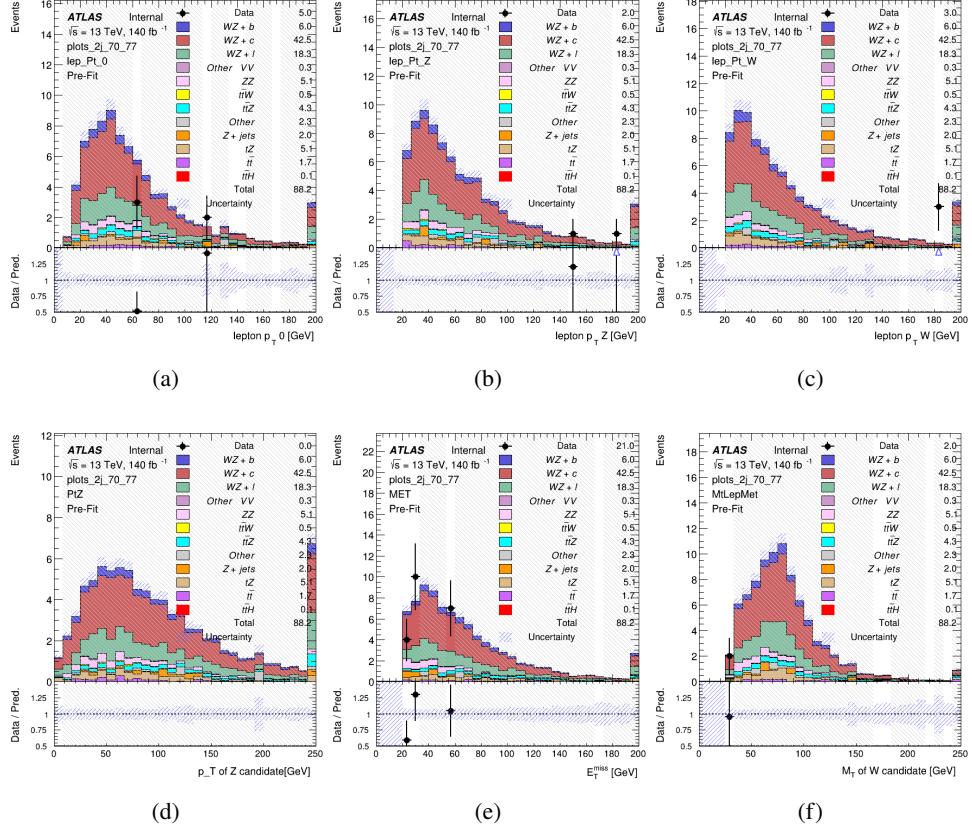


Figure 13: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the Z candidate, (c) the  $p_T$  of the lepton from the W candidate, (d) the  $p_T$  of the Z candidate, (e) the  $E_T^{\text{miss}}$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

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

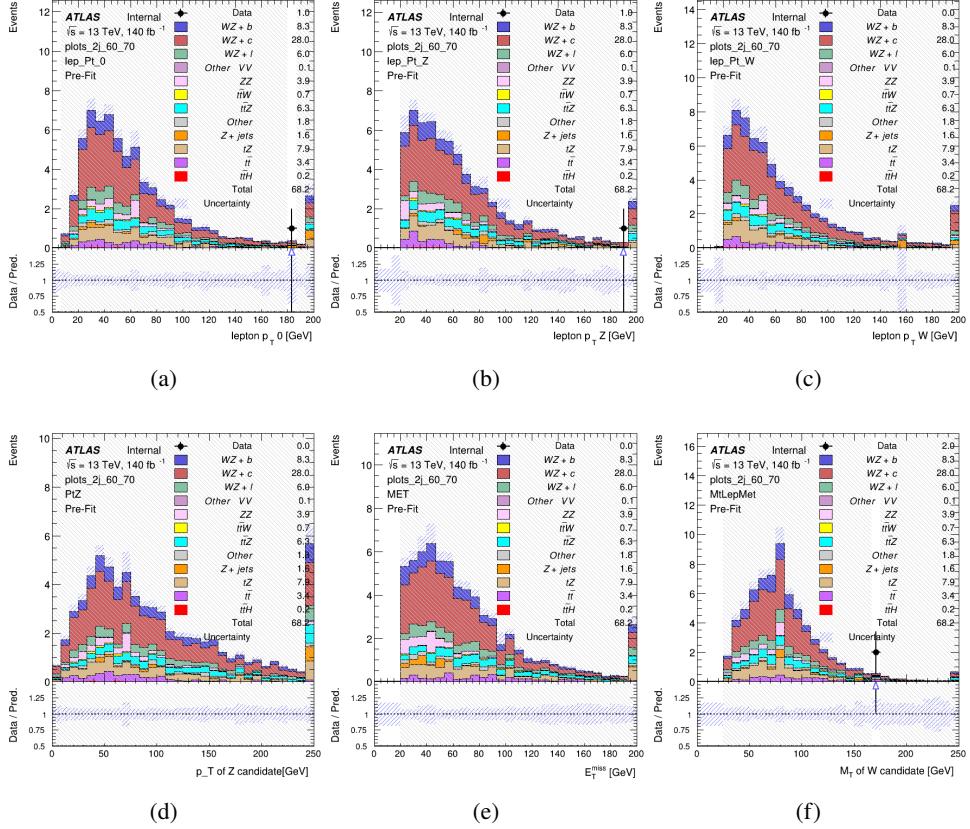


Figure 14: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the Z candidate, (c) the  $p_T$  of the lepton from the W candidate, (d) the  $p_T$  of the Z candidate, (e) the  $E_T^{\text{miss}}$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

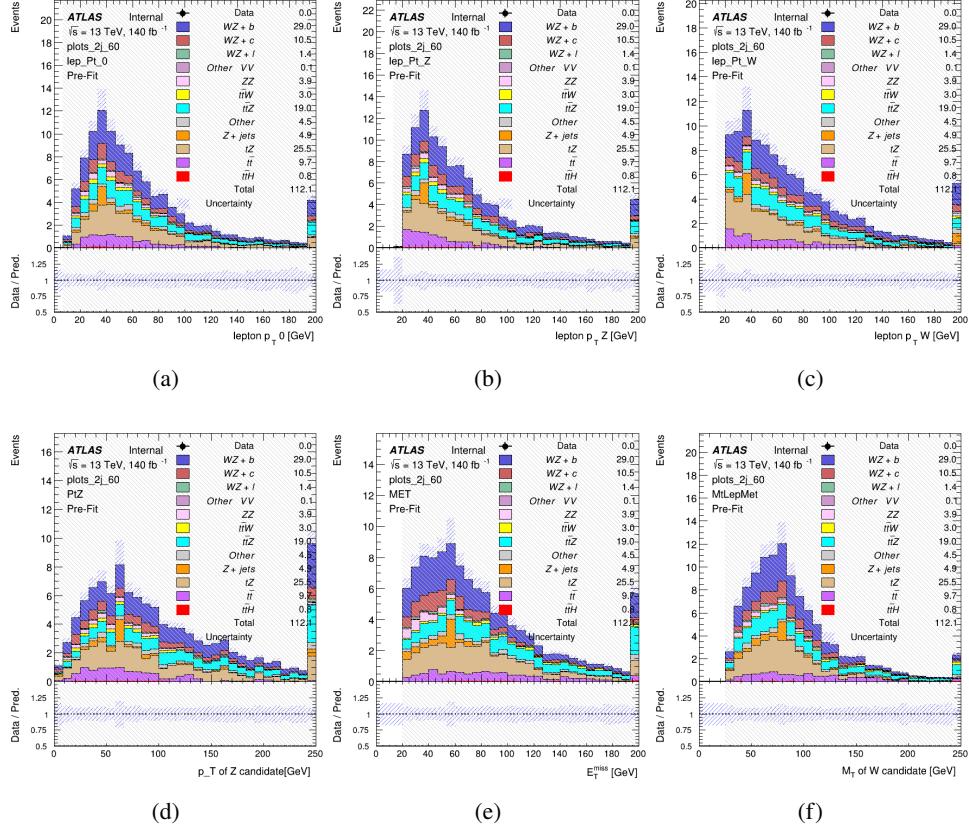
**WZ Fit Region - 2j 60% WP**

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

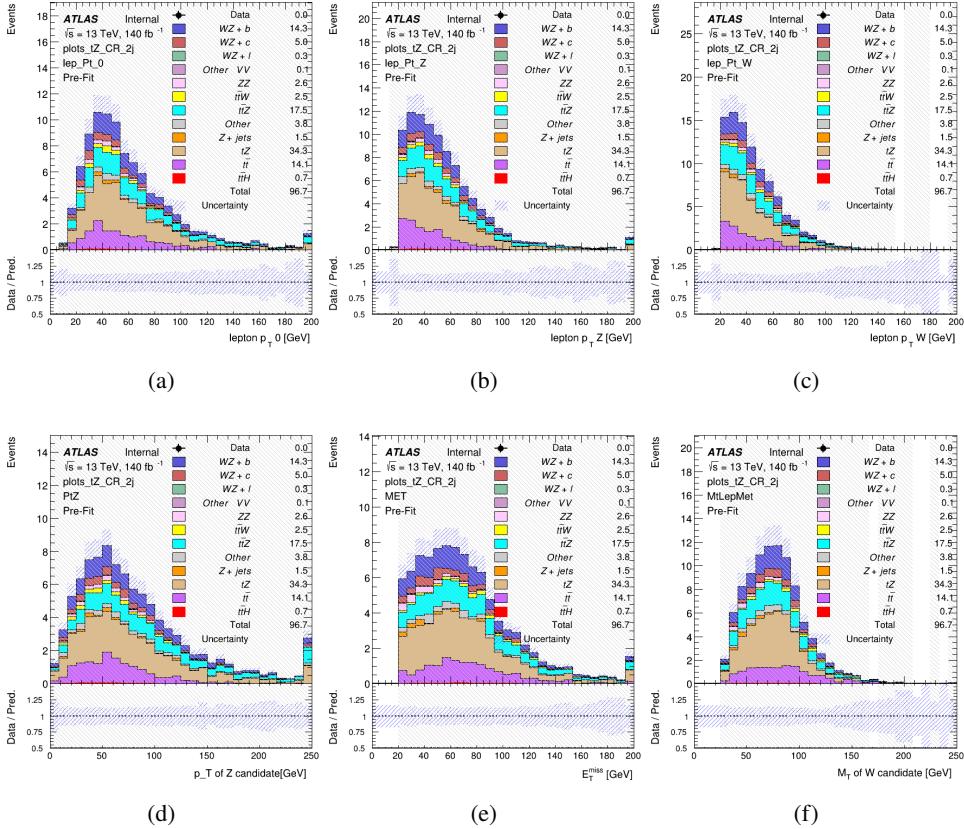


Figure 16: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the  $Z$  candidate, (c) the  $p_T$  of the lepton from the  $W$  candidate, (d) the  $p_T$  of the  $Z$  candidate, (e) the  $E_T^{\text{miss}}$ , 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 3l 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

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

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

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

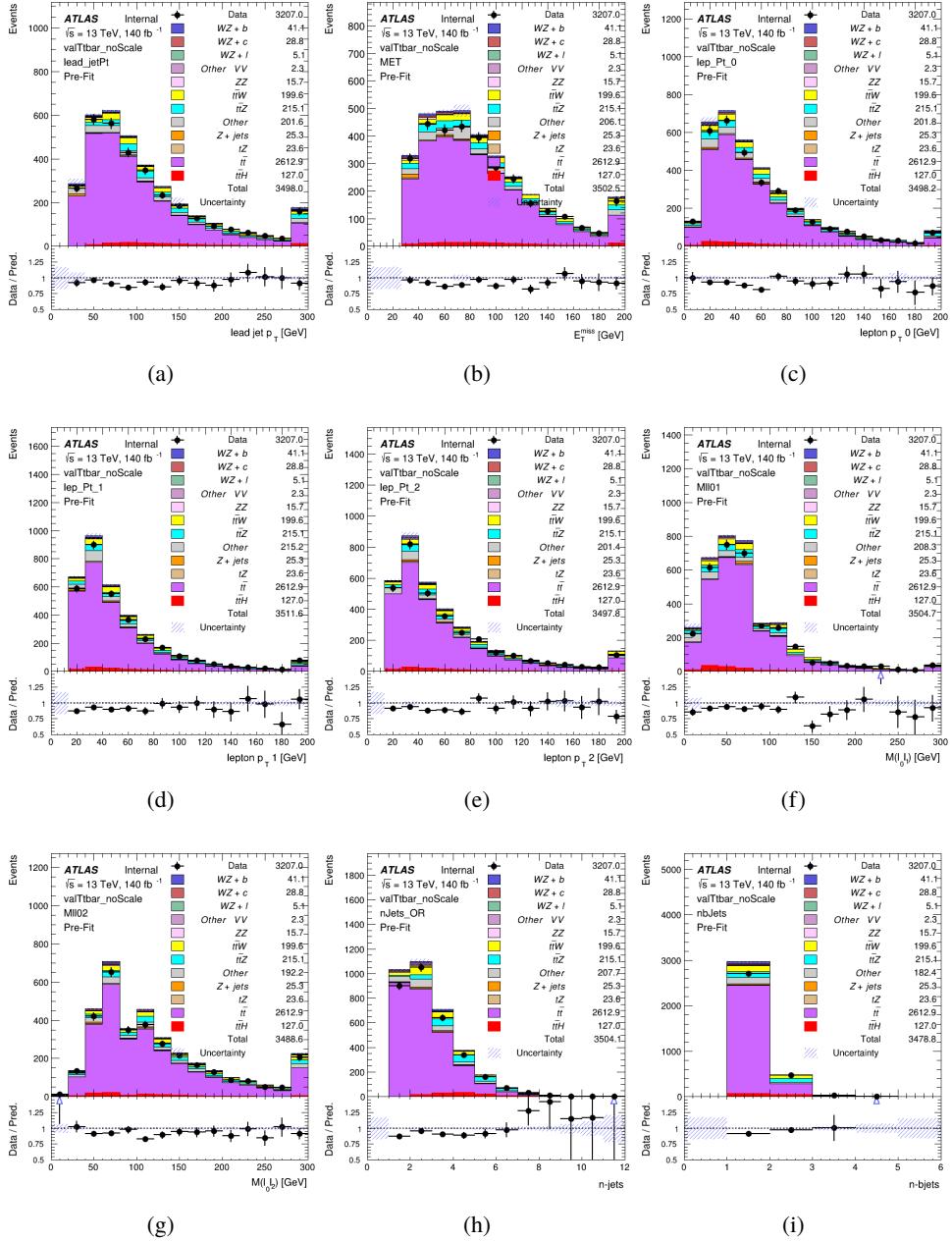


Figure 17: Comparisons between the data and MC distributions in the  $t\bar{t}$  validation region for (a) the  $p_T$  of the leading jet, (b) the missing transverse energy, (c) the  $p_T$  of lepton 0, (d)  $p_T$  of lepton 1, (e)  $p_T$  of lepton 2, (f) the invariant mass of leptons 0 and 1, (g) the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

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

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

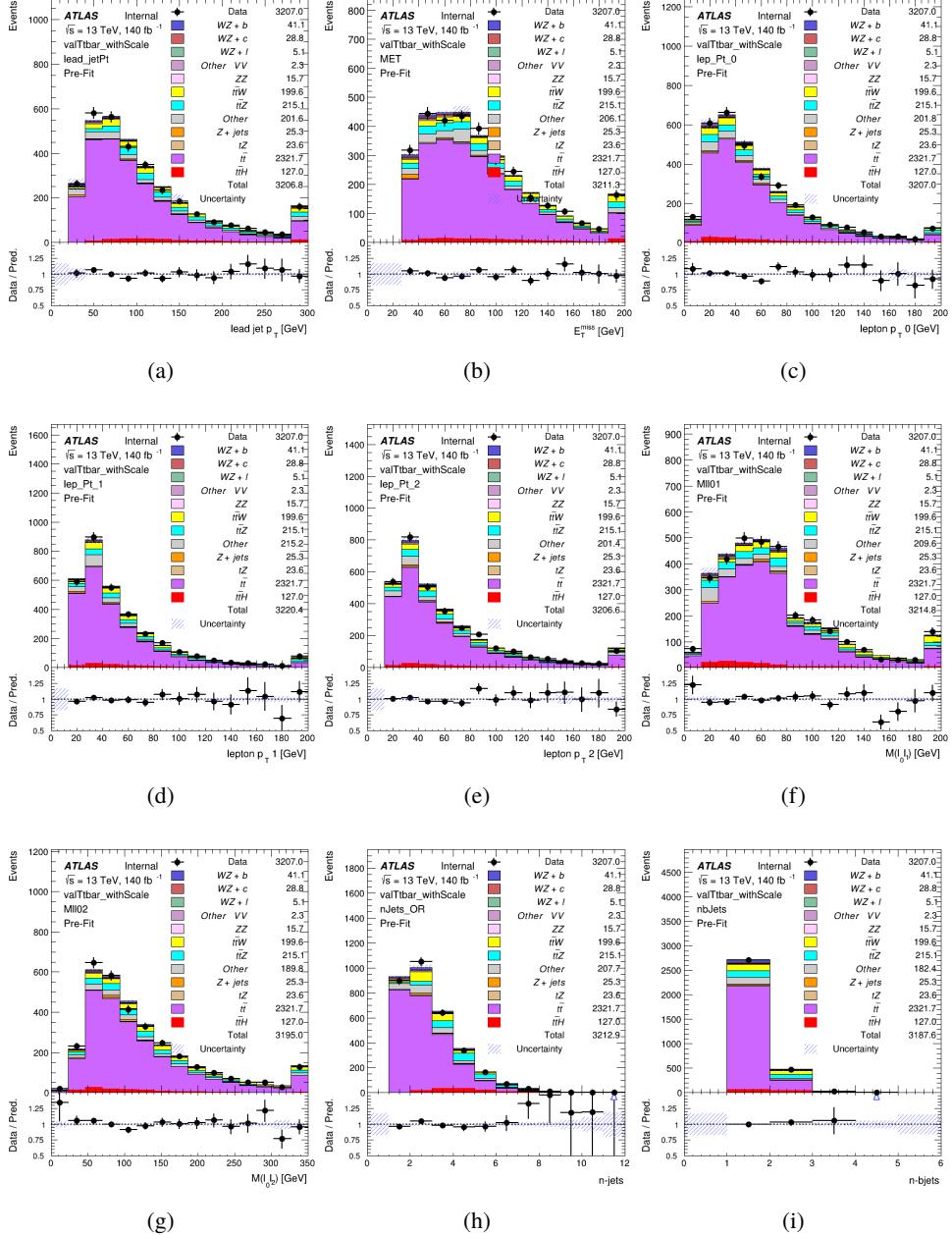


Figure 18: Comparisons between the data and MC distributions in the  $t\bar{t}$  validation region after the correction factor has been applied for (a) the  $p_T$  of the leading jet, (b) the missing transverse energy, (c) the  $p_T$  of lepton 0, (d)  $p_T$  of lepton 1, (e)  $p_T$  of lepton 2, (f)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

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

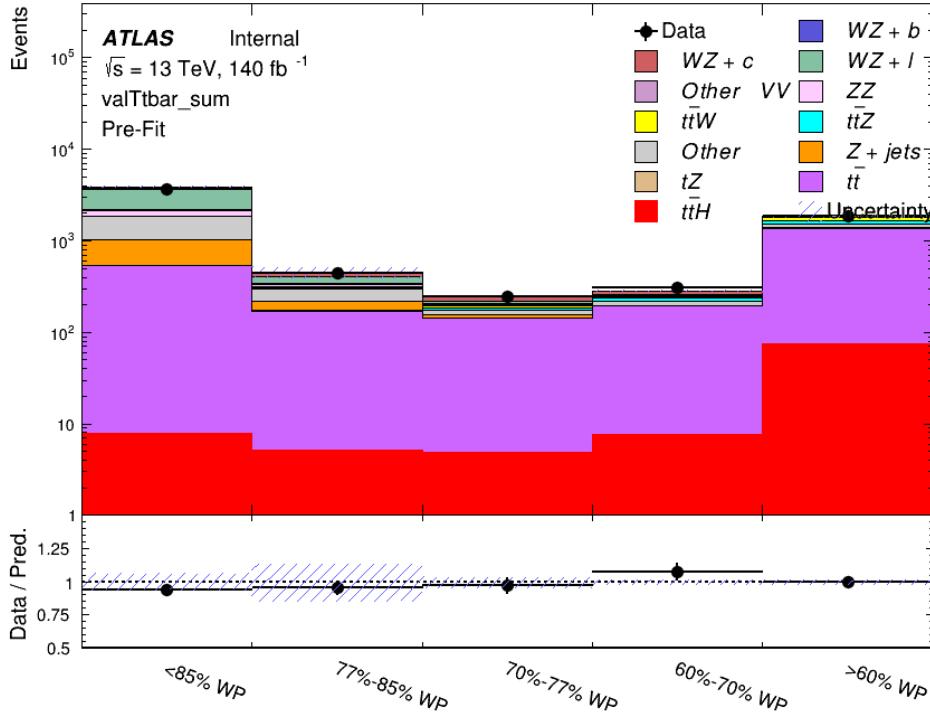


Figure 19: Data and MC comparisons for each DL1r WP for both 1-jet and 2-jet regions, after the  $t\bar{t}$  VR selection and correction factor have been applied

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

### 282 5.3.2 Z+jets Validation

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

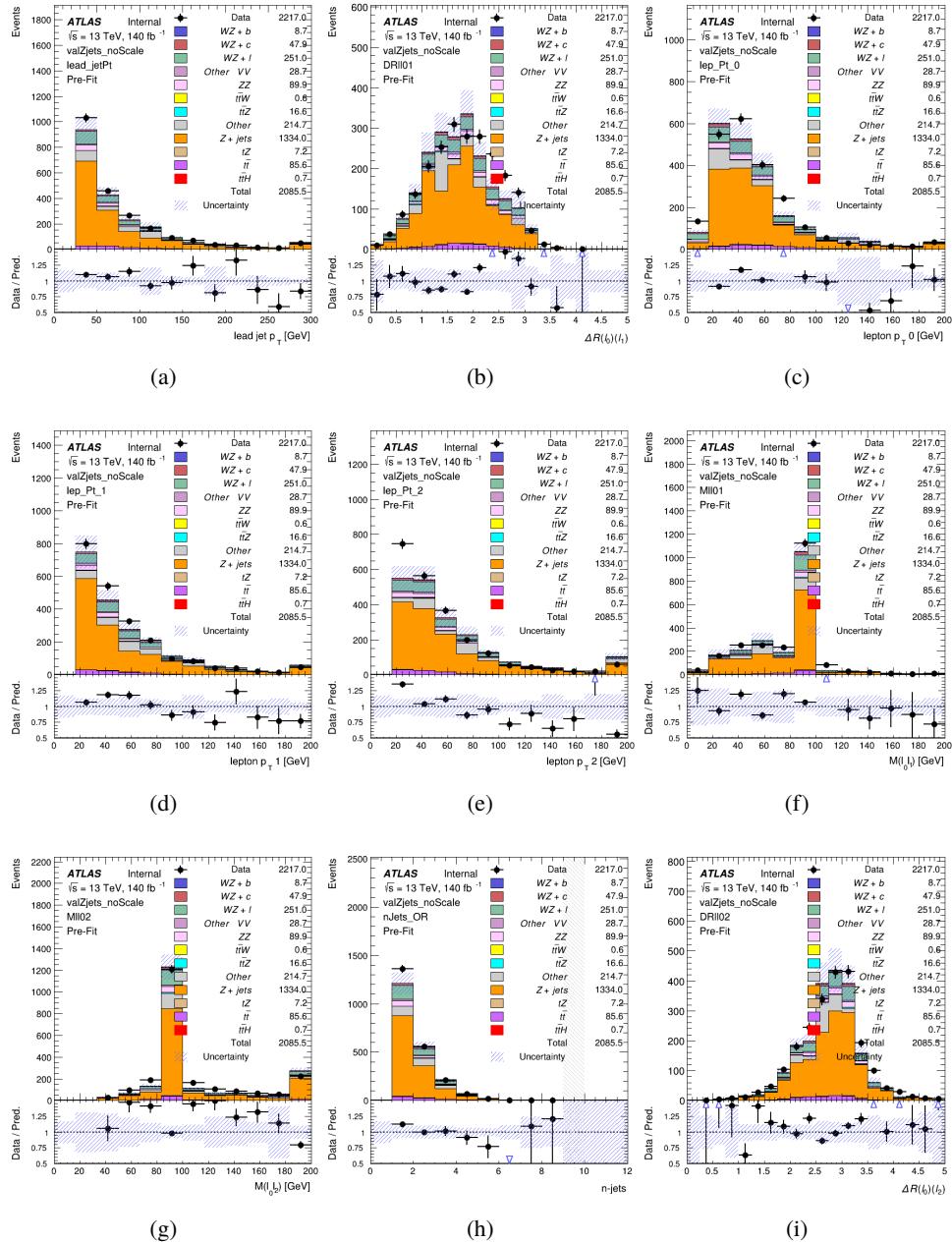


Figure 20: Comparisons between the data and MC distributions in the Z+jets validation region for (a) the  $p_T$  of the leading jet, (b)  $\Delta R$  between leptons 0 and 1, (c) the  $p_T$  of lepton 0, (d)  $p_T$  of lepton 1, (e)  $p_T$  of lepton 2, (f) the invariant mass of leptons 0 and 1, (g)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

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

<sup>291</sup> of the  $p_T$  spectrum of lepton 2 is found to differ. To account for this discrepancy, a variable  
<sup>292</sup> correction factor is applied to Z+jets.  $\chi^2$  minimization of the lepton 2  $p_T$  spectrum is performed  
<sup>293</sup> to derive a correction factor of  $1.53 - 6.6 * 10^{-6}(\text{lep\_Pt\_2})$ . Kinematic plots of the Z + jets  
<sup>294</sup> validation region after this correction factor has been aplied are shown in figure 21.

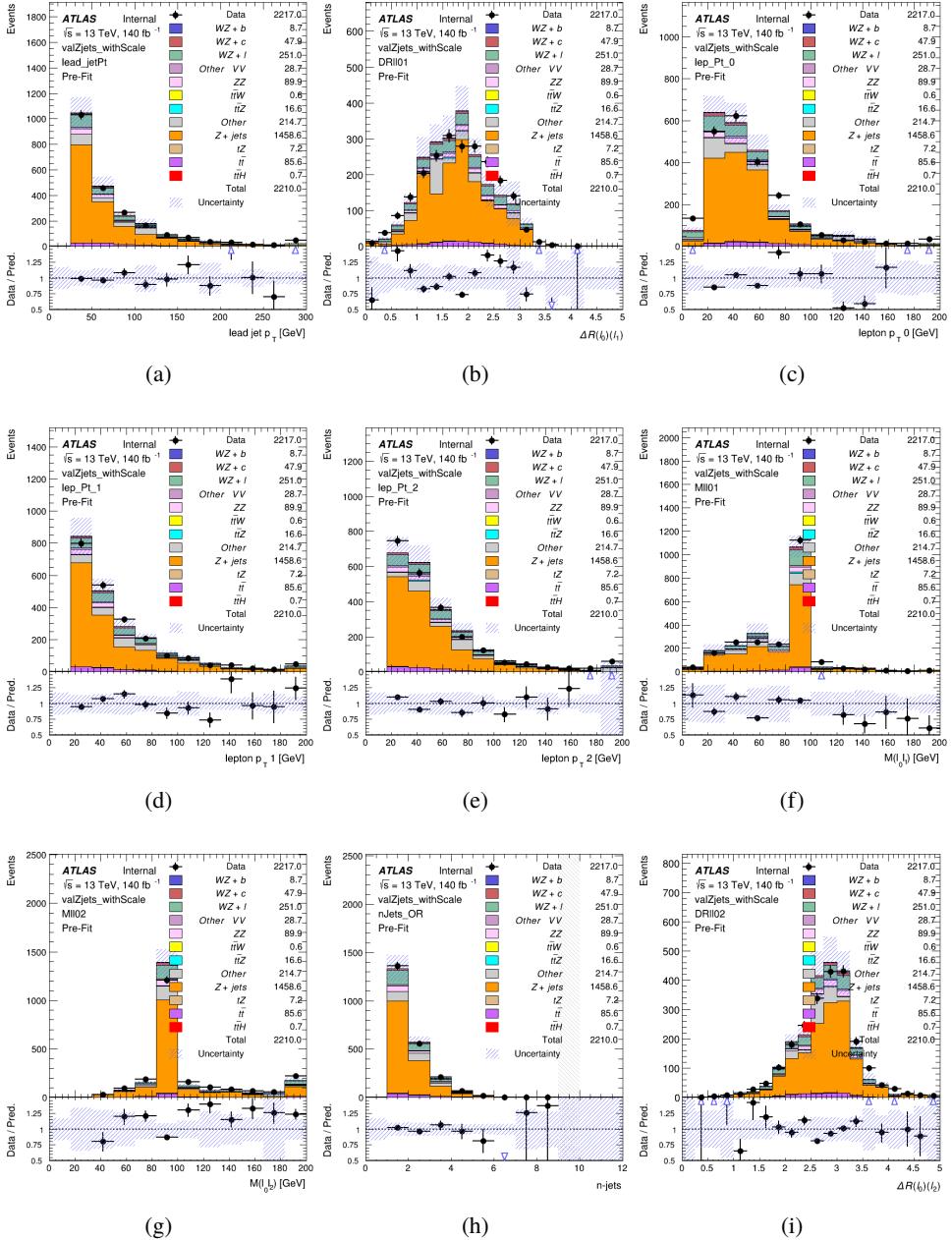


Figure 21: Comparisons between the data and MC distributions in the Z+jets validation region after the correction factor has been applied for (a) the  $p_T$  of the leading jet, (b)  $\Delta R$  between leptons 0 and 1, (c) the  $p_T$  of lepton 0, (d)  $p_T$  of lepton 1, (e)  $p_T$  of lepton 2, (f)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

295 The modeling is further validated by looking at the yield in the Z+jets VR for each DL1r WP,

296 giving a clearer correspondence to the signal regions used in the fit. Each region shown in figure  
 297 22 requires one or more jets pass the listed WP, with no jets passing the next highest WP.

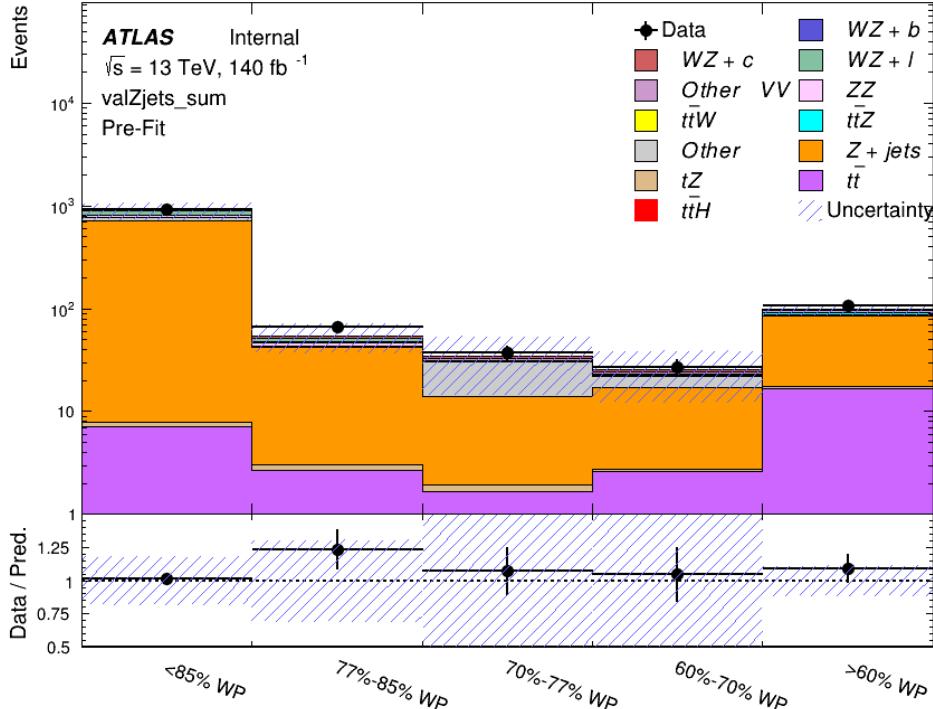


Figure 22: Data and MC comparisons for each DL1r WP for both 1-jet and 2-jet regions, after the Z+jets VR selection and correction factor have been applied

298 For each of the working points considered, the data falls within 20% of the MC prediction once  
 299 this correction factor has been applied. Therefore, a 20% systematic uncertainty is applied to Z  
 300 + jets in the analysis.

## 301 6 tZ Interference Studies and Separation Multivariate Analysis

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

307 Because tZ produces a final state identical to signal, it represents a predominant background in  
 308 the most signal enriched regions. That is, the region with one jet passing the 60% DL1r WP.

309 Therefore, a boosted decision tree (BDT) algorithm is trained using TMVA [10] to separate WZ  
310 + heavy flavor from tZ.

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

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

316 **6.1 Interference Studies**

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

322 MadGraph simulations of 10,000 tZ and 10,000 WZ + b-jet events are produced, along with  
323 20,000 events where both are present, in the fiducial region where three leptons and at least one  
324 jet are produced.

325 The kinematics of these samples are shown below:

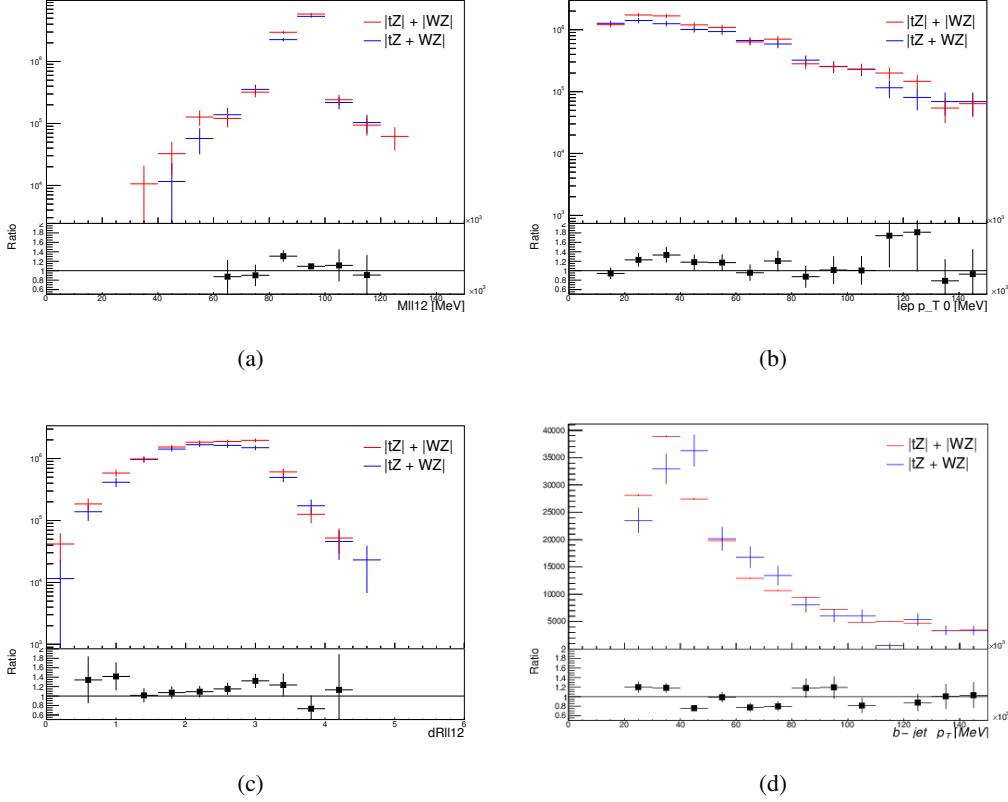


Figure 23: Comparisons between (a) the invariant mass of the Z-candidate, (b) the  $p_T$  of the leading lepton, (c)  $\Delta(R)$  of the two leptons that form the Z-candidate, and (d) the  $p_T$  of the b-jet, for WZ and tZ events generated with interference effects (blue) and without interference effects (red).

## 326 6.2 Top Mass Reconstruction

327 The reconstruction of the top mass follows the procedure described in detail in section 6.1 of  
 328 [11]. The mass of the top quark candidate is reconstructed from the jet, the lepton not included  
 329 in the Z-candidate, and a reconstructed neutrino. In the case that there is one jet in the event,  
 330 there is only possible b-jet candidate. For events with two jets, the jet with the highest DL1r  
 331 score is used.

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

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

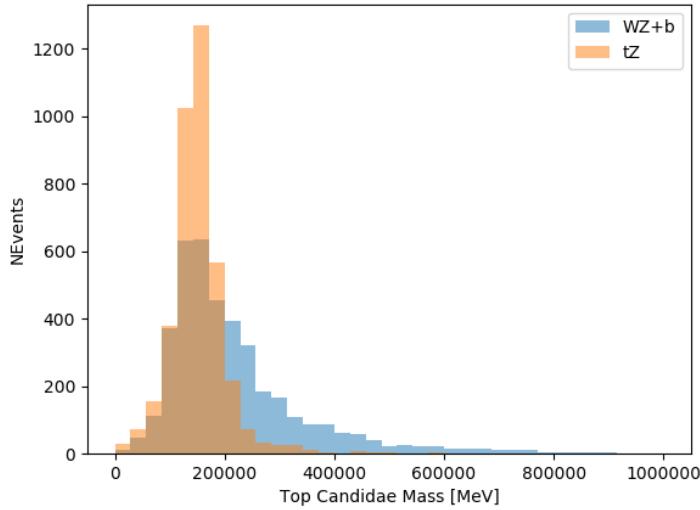


Figure 24: Reconstructed top mass distributions for tZ and WZ + b, measured in MeV.

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

<sup>337</sup> Expanding this out into components, this equation gives:

$$\sqrt{p_{T\nu}^2 + p_{z\nu}^2} E_l = \frac{m_W^2 - m_l^2}{2} + p_{T\nu}(p_{lx}\cos\phi_\nu + p_{ly}\sin\phi_\nu) + p_{lz}p_{\nu z}$$

<sup>340</sup> This equation gives two solutions for  $p_{\nu z}$ . For cases where only one of these solutions is real,  
<sup>341</sup> that is taken as the value of  $p_{\nu z}$ . For instances with two real solutions, the one which is shown  
<sup>342</sup> to be correct in the largest fraction of simulations is taken. For cases when no real solution is  
<sup>343</sup> found, often because of detector effects, the value of  $E_T^{\text{miss}}$  is varied in decreasing increments of  
<sup>344</sup> 100 MeV until a real solution is found.

<sup>345</sup> The reconstructed top mass distribution for tZ and WZ + b can be seen in figure 24.

### <sup>346</sup> 6.3 tZ BDT

<sup>347</sup> A Boosted Decision Tree (BDT), specifically XGBoost [[xgboost\\_cite](#)], is used to provide separation  
<sup>348</sup> between tZ and WZ+b. The following kinematic variables are used as inputs:

- <sup>349</sup> • The invariant mass of the reconstructed top candidate
- <sup>350</sup> •  $p_T$  of each of the leptons, jet
- <sup>351</sup> • The invariant mass of each combination of lepton pairs,  $M(l\bar{l})$

- 352     •  $E_T^{\text{miss}}$   
 353     • Distance between each combination of leptons,  $\Delta R(l_1 l_2)$   
 354     • Distance between each lepton and the jet,  $\Delta R(l_j)$

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

358 The distributions of a few of these features for both signal and background is shown in figure  
 359 25.

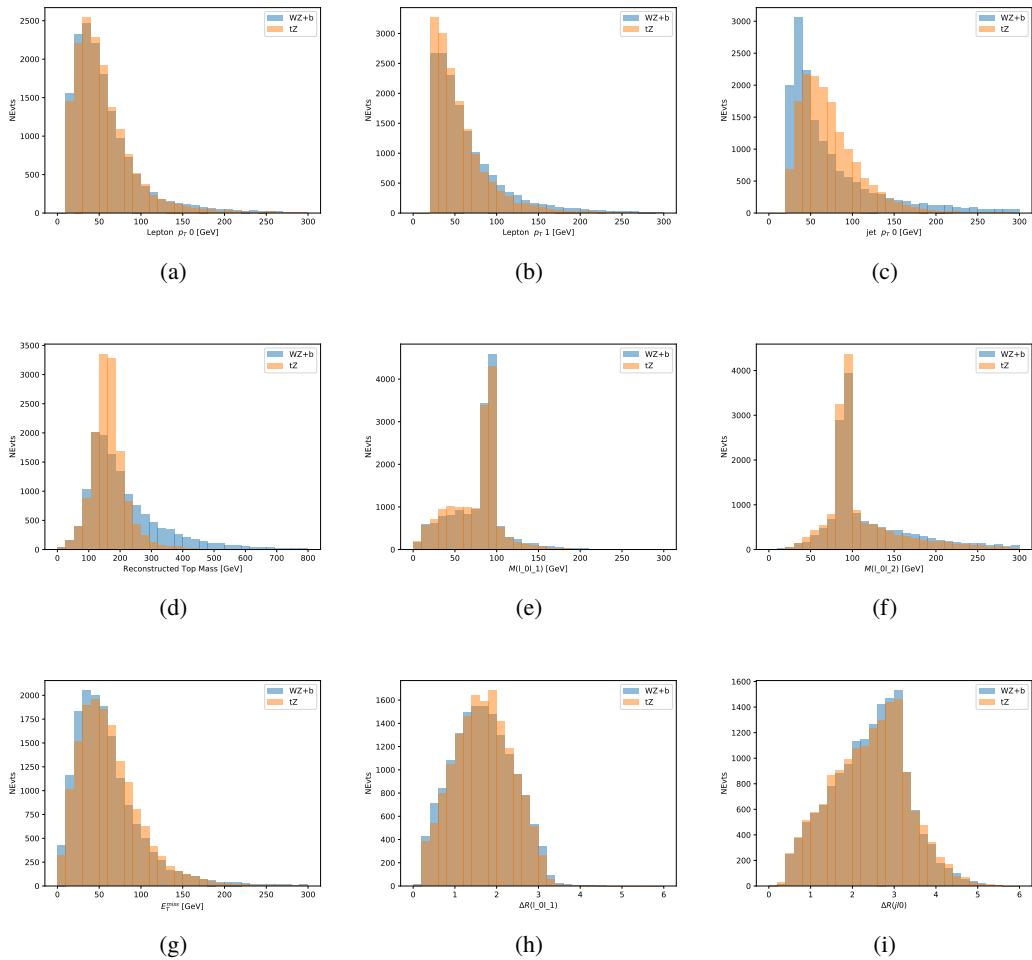


Figure 25: Distribution of input features of the BDT for signal (WZ) and background (tZ). Both are scaled to an equal number of events. (a), (b) and (c) show the  $p_T$  of lepton 0, lepton 1, and the jet, (d) show the reconstructed top mass, (e) and (f) show the invariant mass of leptons 0 and 1, leptons 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.

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

365 The results of the BDT training are shown in figure 26. The output scores for both signal and  
 366 background events is shown on the left. The right shows the receiving operating characteristic  
 367 (ROC) curve that results from the MVA. The ROC curve represents the background rejection  
 368 as a function of signal efficiency, where each point on the curve represents a different response  
 369 score. The ROC curve of the BDT is compared to the performance of using an optimal set of flat  
 370 selections on the same set of input variables.

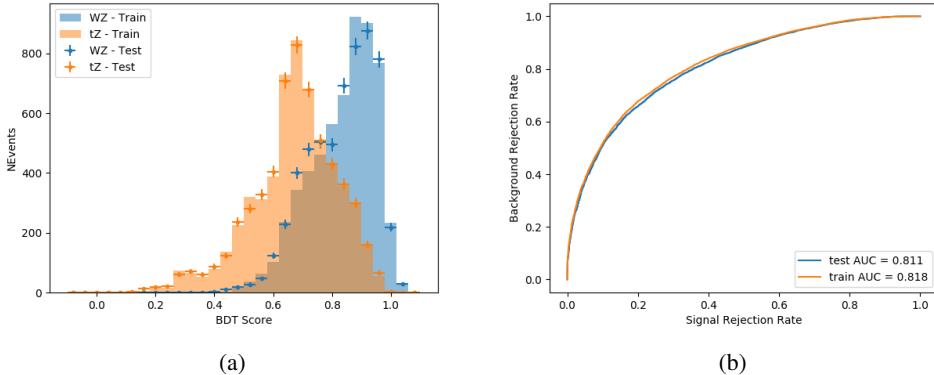


Figure 26: Distribution of the BDT response for signal and background events on the left, the ROC curve for the BDT on the right.

371 The relative important of each input feature in the model, measured by how often they appeared  
 372 in the decision trees, is shown in figure 27.

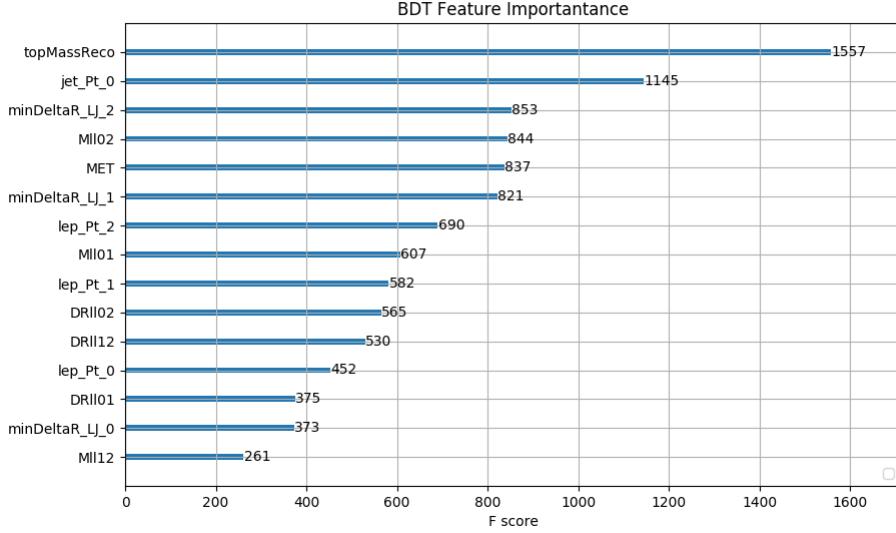


Figure 27: Relative importance of each input feature in the model.

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

## 375 7 Systematic Uncertainties

376 The systematic uncertainties that are considered are summarized in table 9. These are imple-  
 377 mented in the fit either as a normalization factors or as a shape variation or both in the signal  
 378 and background estimations. The numerical impact of each of these uncertainties is outlined in  
 379 section 8.

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

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

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

Table 9: 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
<b>Total (Overall)</b>	<b>226</b>

<sup>390</sup> The systematic uncertainties associated with the signal and background processes are accounted  
<sup>391</sup> for by varying the cross-section of each process within its uncertainty.

<sup>392</sup> The full list of systematic uncertainties considered in the analysis is summarized in tables 10, 11  
<sup>393</sup> and 12.

<sup>394</sup>

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 10: 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
1st February 2021 – Run 44 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

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 12: 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.

395 Theoretical uncertainties applied to backgrounds, including cross section, PDF, and scale un-  
 396 certainties are taken from theory calculations, with the exception of non-prompt and diboson  
 397 backgrounds. Derivation of the non-prompt background uncertainties, Z+jets and  $t\bar{t}$ , are ex-  
 398 plained in detail in section 5.3. Because the other VV + heavy flavor processes are also poorly  
 399 understood, a conservative 50% uncertainty is applied to those samples. The cross-section  
 400 uncertainties applied to the predominate background estimates are summarized in table 13.

## 8 Results

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

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

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	$\pm 20$ placeholder
Z + jets	$\pm 25$ placeholder

Table 13: Summary of theoretical uncertainties for MC prediction of backgrounds in the analysis.

414 **8.1 1-jet Fit Results**

415 **The results of the fit are currently blinded.** The post-fit yields in each region are summarized  
416 in figure 28.

417 A post-fit summary plot of the 1-jet fitted regions is shown in figure 29:

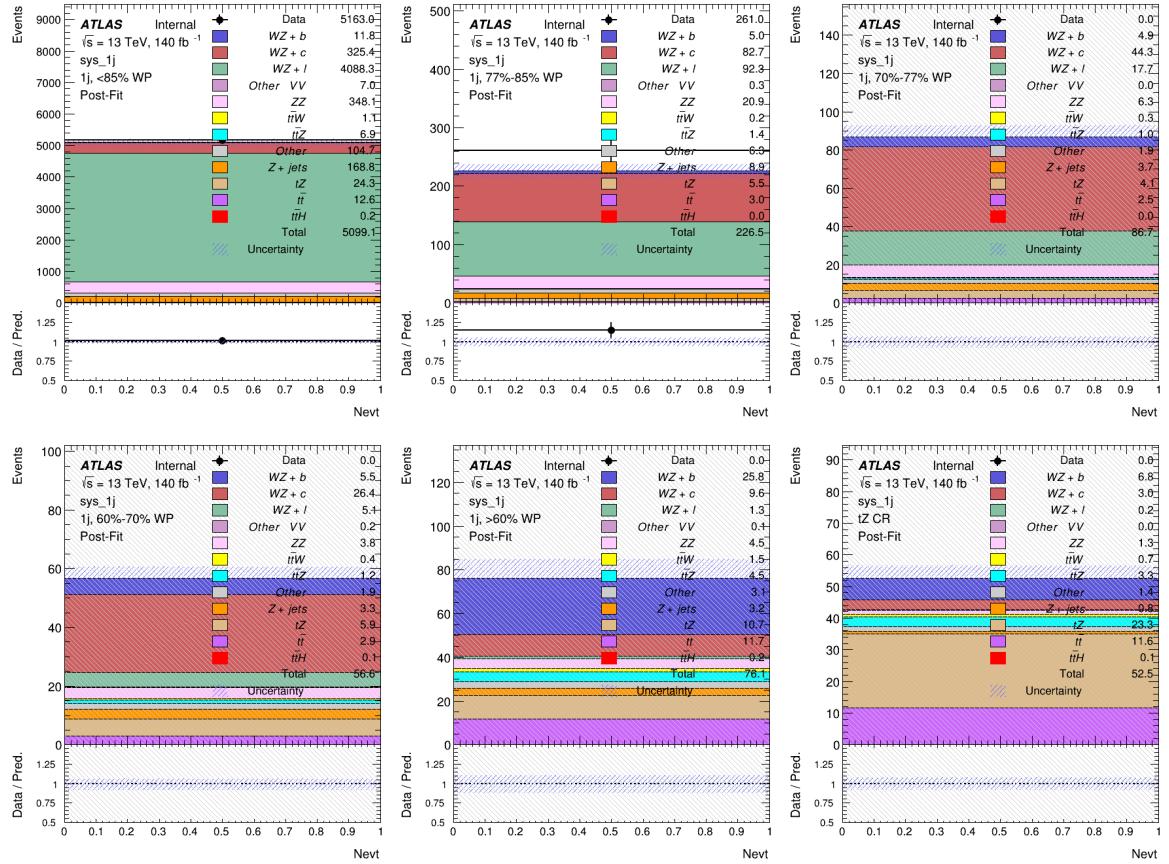


Figure 28: Data/MC results in each of the 1-jet regions after the fit has been performed.

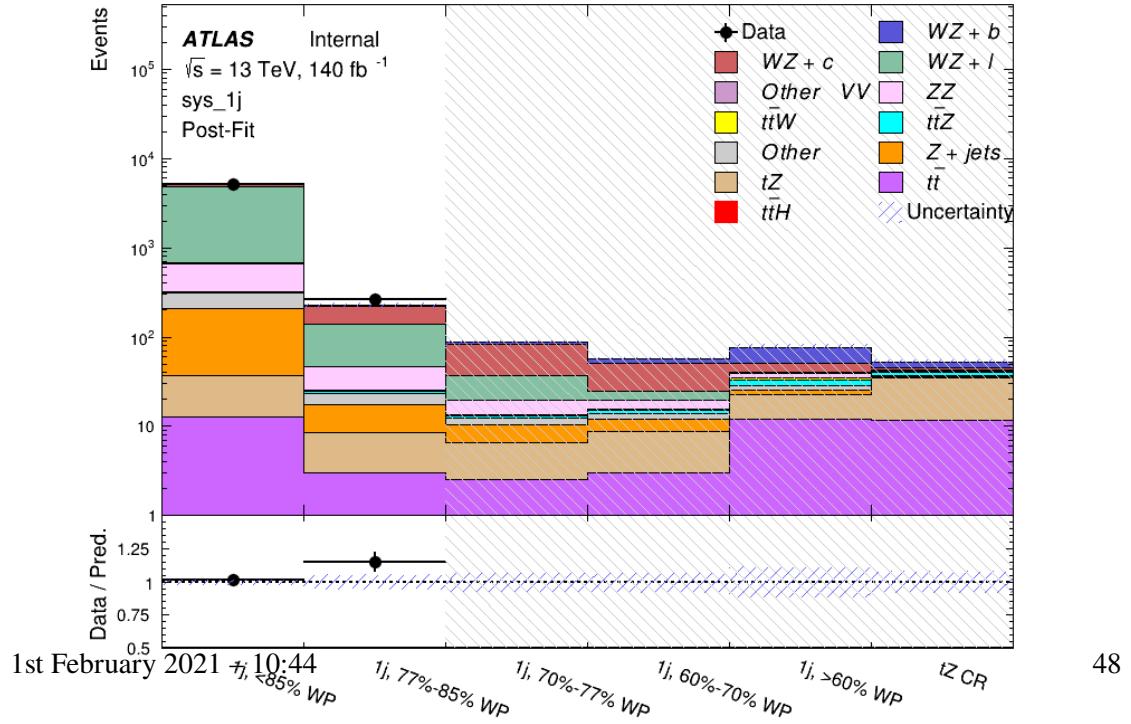


Figure 29: Post-fit summary of the 1-jet fit regions.

418 As described in section 7, there are 226 systematic uncertainties that are considered as NPs in  
 419 the fit. These NPs are constrained by Gaussian or log-normal probability density functions. The  
 420 latter are used for normalisation factors to ensure that they are always positive. The expected  
 421 numbers of signal and background events are functions of the likelihood. The prior for each  
 422 NP is added as a penalty term, decreasing the likelihood as it is shifted away from its nominal  
 423 value.

424 The impact of each NP is calculated by performing the fit with the parameter of interest held  
 425 fixed, varied from its fitted value by its uncertainty, and calculating  $\Delta\mu$  relative to the baseline  
 426 fit. The impact of the most significant sources of systematic uncertainties is summarized in table  
 427 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 with exactly one associated jet.

428 The ranking and impact of those nuisance parameters with the largest contribution to the overall  
 429 uncertainty is shown in figure 30.

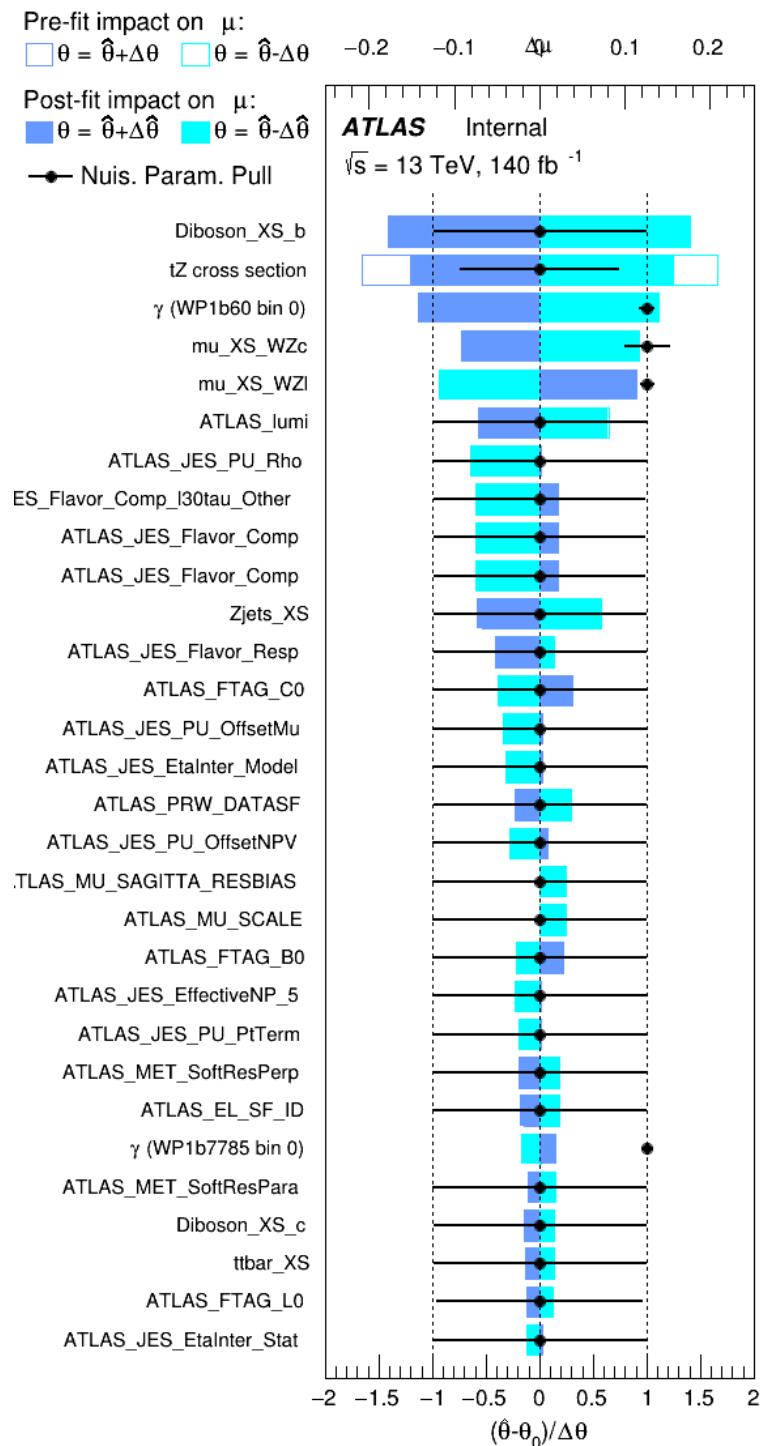


Figure 30: Impact of systematic uncertainties on the signal-strength of  $WZ + b$  for events with exactly one jet

430 The large impact of the Jet Energy Scale and Jet Flavor Tagging is unsurprising, as the shape  
 431 of the fit regions depends heavily on the modeling of the jets. The other major sources of  
 432 uncertainty come from background modelling and cross-section uncertainty. The pie charts in  
 433 figure 31 show that for the modelling uncertainties that contribute most correspond to the most  
 434 significant backgrounds.

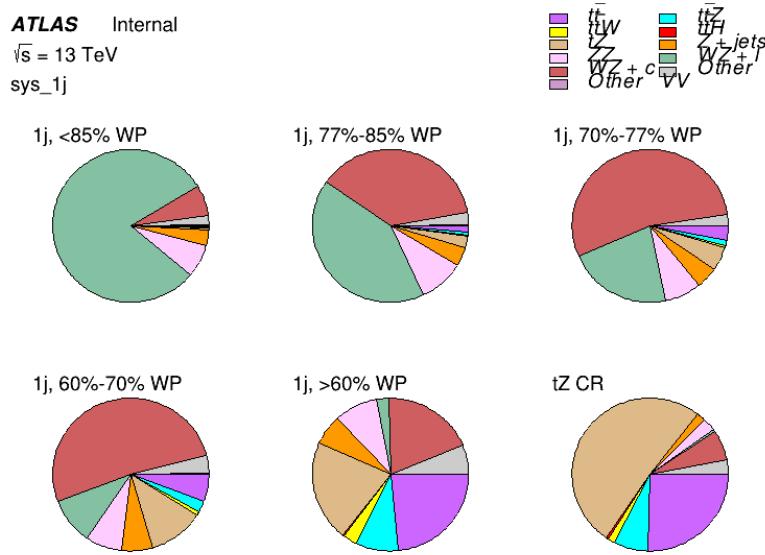


Figure 31: Post-fit background composition of the fit regions.

435 The correlations between these nuisance parameters are summarized in figure 32.

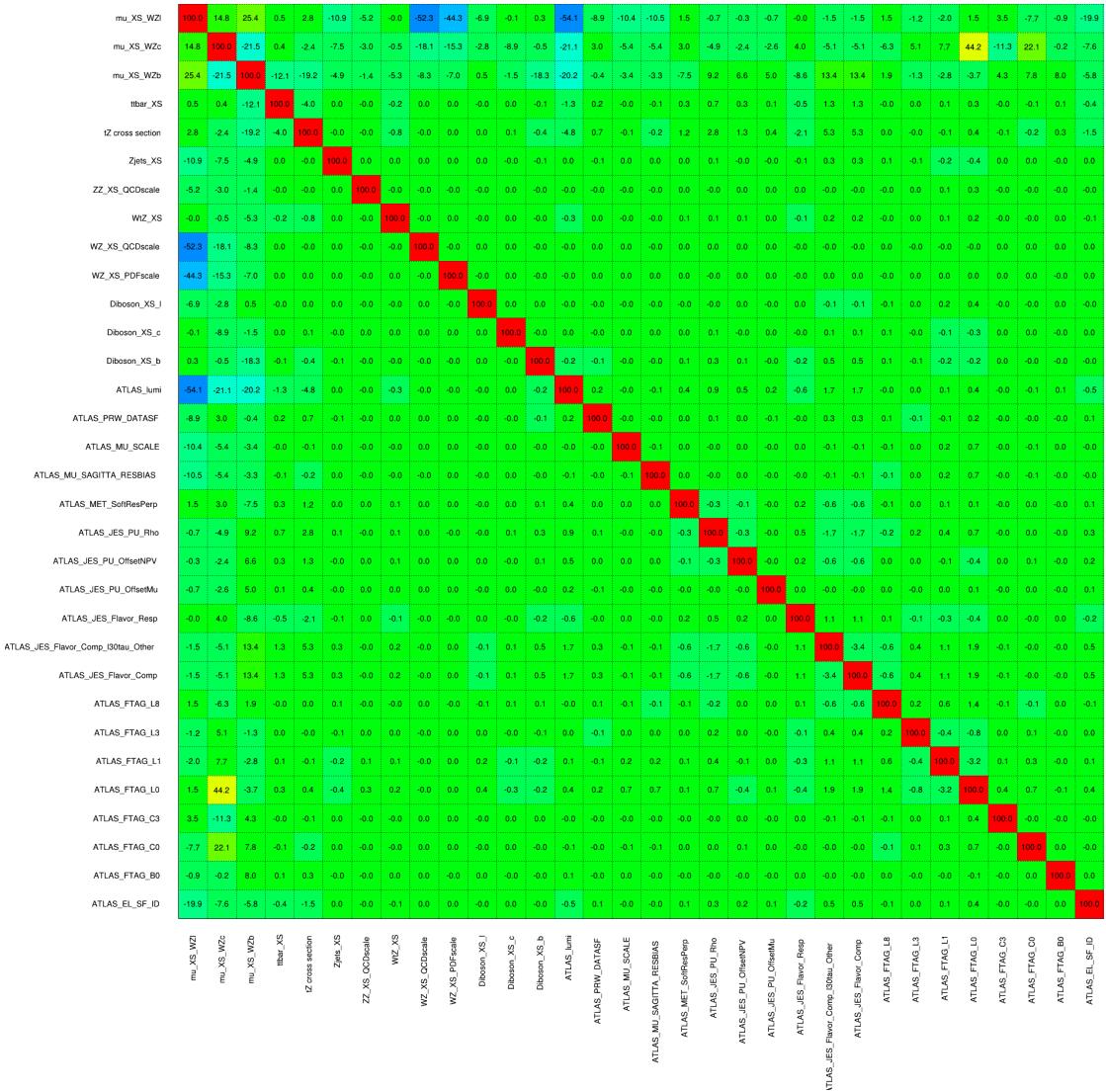


Figure 32: Correlations between nuisance parameters

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

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

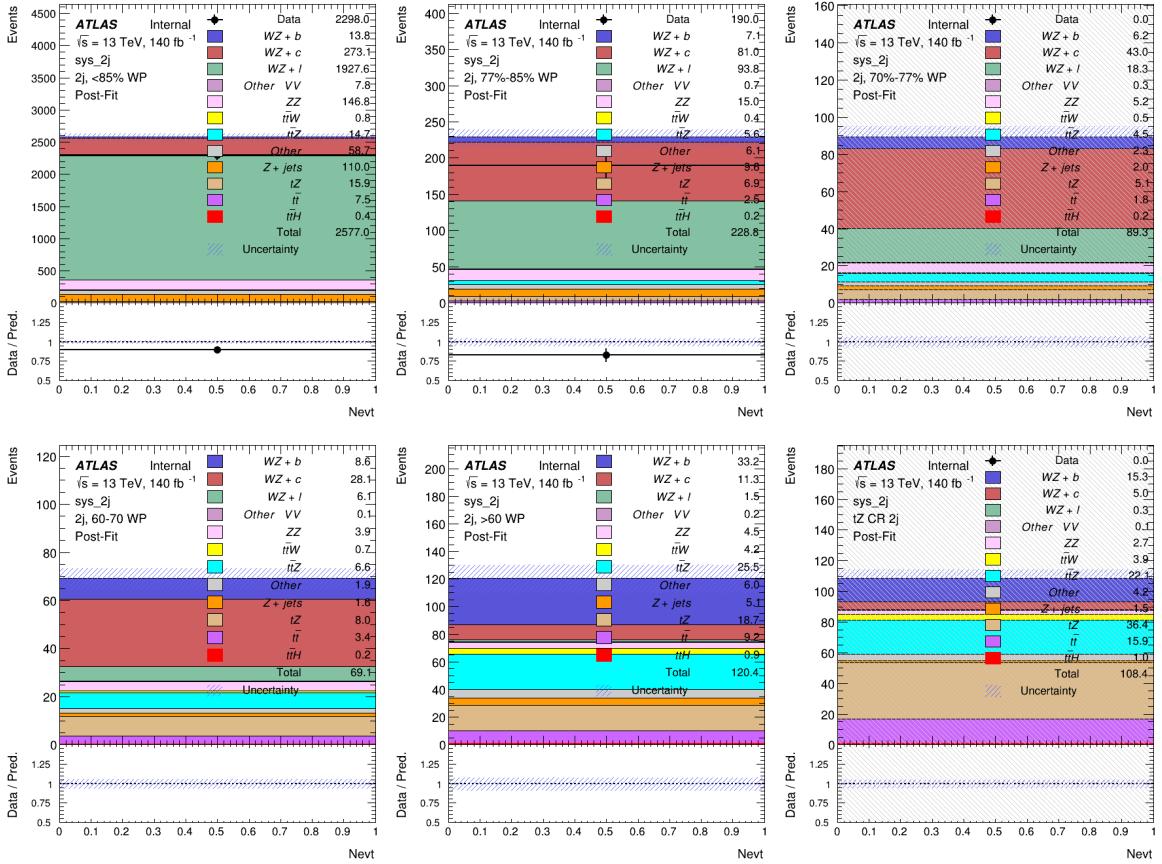


Figure 33: Data/MC results in each of the regions in the 2-jet fit after the fit has been performed.

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

## 447 8.2 2-jet Fit Results

448 **The results of the fit are currently blinded.** The post-fit yields in each region are summarized  
 449 in figure 33.

450 A post-fit summary plot of the fitted regions is shown in figure 34:

451 The same set of systematic uncertainties consider for the 1-jet fit are included in the 2-jet fit as  
 452 well. The impact of the most significant systematic uncertainties is summarized in table 15.

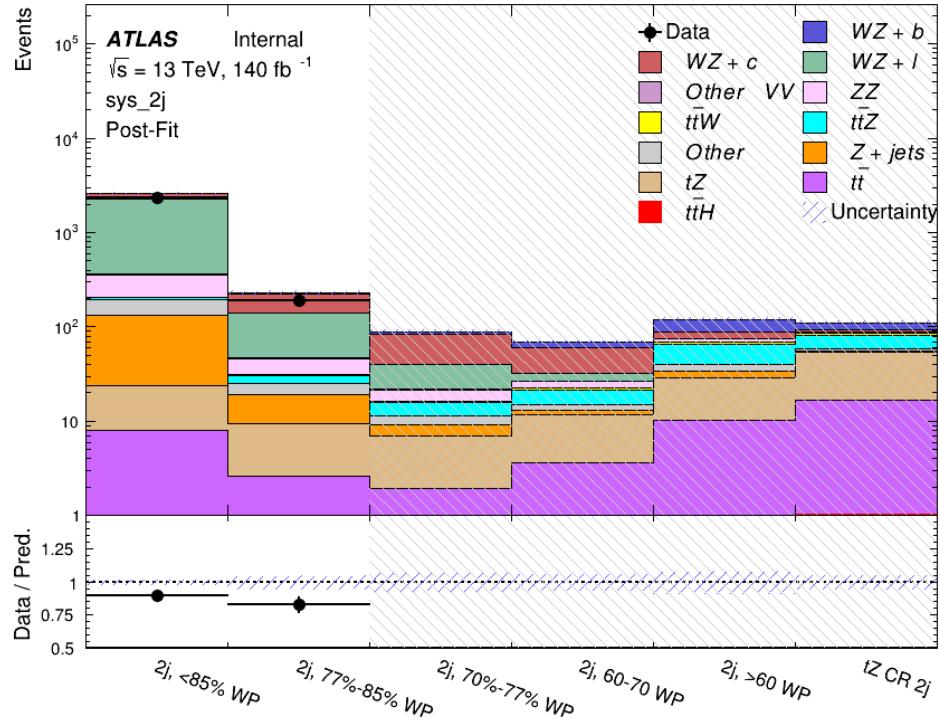


Figure 34: Post-fit summary of the fit over 2-jet regions.

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

The ranking and impact of those nuisance parameters with the largest contribution to the overall uncertainty is shown in figure 35.

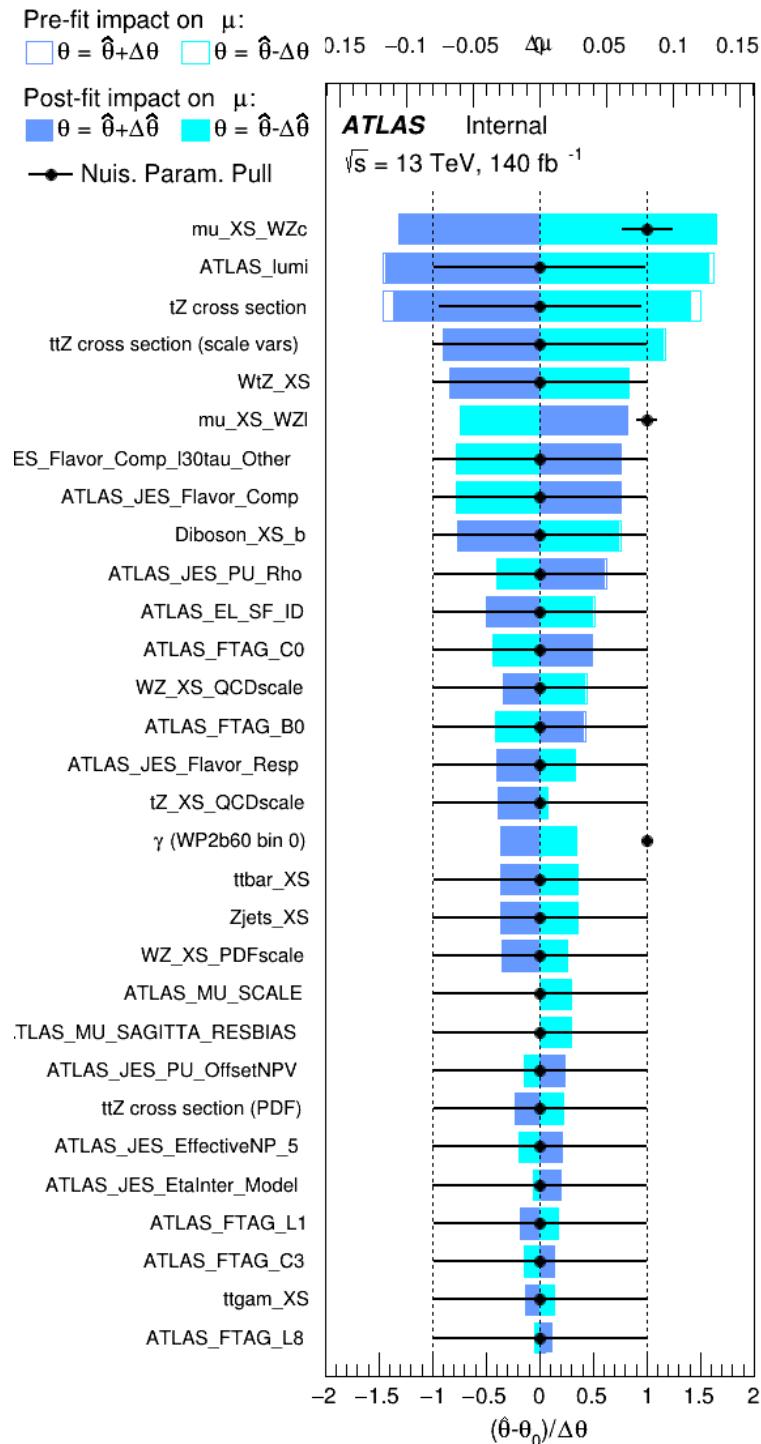


Figure 35: Impact of systematic uncertainties on the signal-strength of WZ + b in 2-jet events.

455 The large impact of the Jet Energy Scale and Jet Flavor Tagging is unsurprising, as the shape  
 456 of the fit regions depends heavily on the modeling of the jets. The other major sources of  
 457 uncertainty come from background modelling and cross-section uncertainty. The pie charts in  
 458 figure 36 show that for the modelling uncertainties that contribute most correspond to the most  
 459 significant backgrounds.

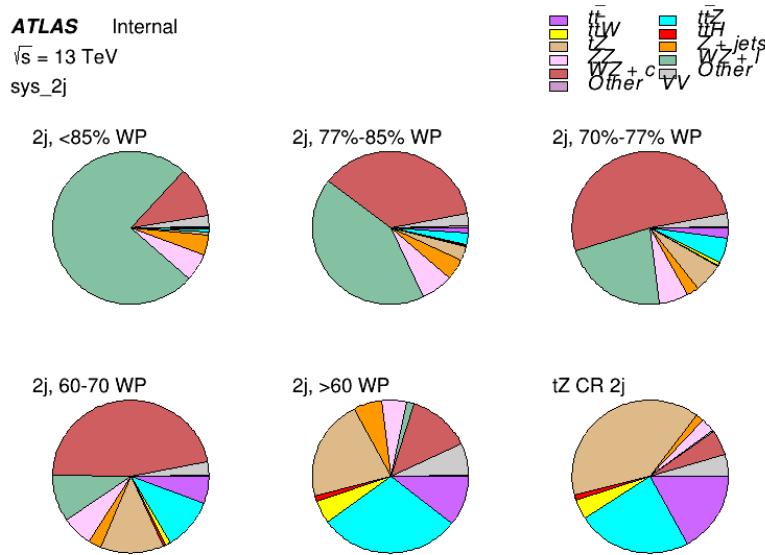


Figure 36: Post-fit background composition of the 2-jet fit regions.

460 The correlations between these nuisance parameters are summarized in figure 37.

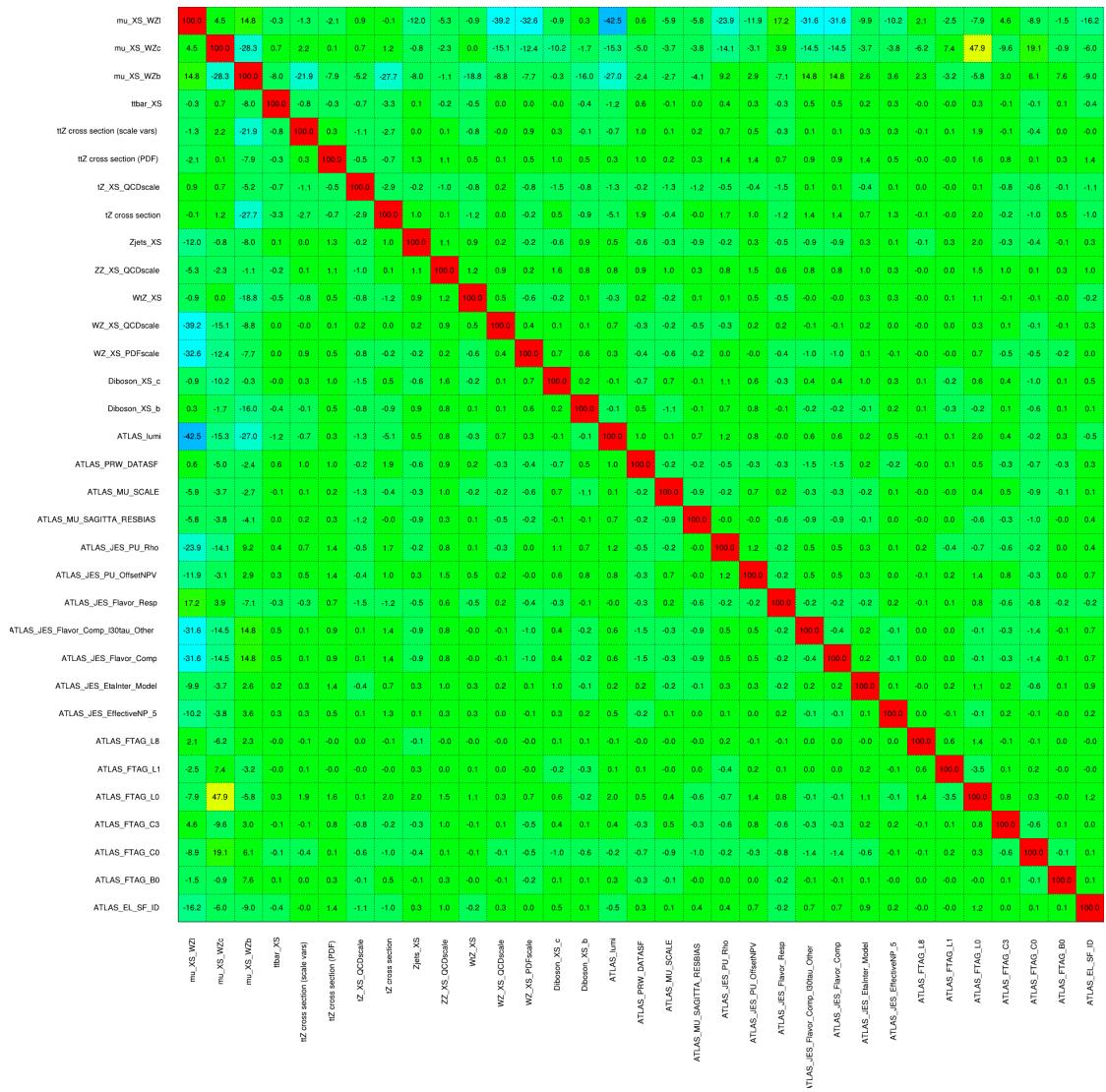


Figure 37: Correlations between nuisance parameters in the 2-jet fit

As in the 1-jet case, no significant, unexpected correlations are found between nuisance parameters.

## 9 Conclusion

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

466 final results once unblinded.

467 **References**

- 468 [1] M. Aaboud et al. ‘Observation of electroweak  $W^\pm Z$  boson pair production in association  
469 with two jets in pp collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector’. In: *Phys.  
470 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:  
471 [1812.09740 \[hep-ex\]](https://arxiv.org/abs/1812.09740).
- 472 [2] T. Gleisberg et al. ‘Event generation with SHERPA 1.1’. In: *JHEP* 02 (2009), p. 007. doi:  
473 [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).
- 474 [3] ATLAS Collaboration. *Electron efficiency measurements with the ATLAS detector using  
475 the 2015 LHC proton–proton collision data*. ATLAS-CONF-2016-024. 2016. URL: <https://cds.cern.ch/record/2157687>.
- 476 [4] ATLAS Collaboration. ‘Measurement of the muon reconstruction performance of the  
477 ATLAS detector using 2011 and 2012 LHC proton–proton collision data’. In: *Eur. Phys.  
478 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).
- 479 [5] *Evidence for the associated production of the Higgs boson and a top quark pair with the  
480 ATLAS detector*. Tech. rep. ATLAS-CONF-2017-077. Geneva: CERN, Nov. 2017. URL:  
481 <https://cds.cern.ch/record/2291405>.
- 482 [6] ATLAS Collaboration. *Jet Calibration and Systematic Uncertainties for Jets Reconstruc-  
483 ted in the ATLAS Detector at  $\sqrt{s} = 13$  TeV*. ATL-PHYS-PUB-2015-015. 2015. URL:  
484 <https://cds.cern.ch/record/2037613>.
- 485 [7] ATLAS Collaboration. *Selection of jets produced in 13 TeV proton–proton collisions with  
486 the ATLAS detector*. ATLAS-CONF-2015-029. 2015. URL: [https://cds.cern.ch/  
487 record/2037702](https://cds.cern.ch/record/2037702).
- 488 [8] ATLAS Collaboration. ‘Performance of pile-up mitigation techniques for jets in pp col-  
489 lisions at  $\sqrt{s} = 8$  TeV using the ATLAS detector’. In: *Eur. Phys. J. C* 76 (2016), p. 581.  
490 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).
- 491 [9] ATLAS Collaboration. *Performance of missing transverse momentum reconstruction with  
492 the ATLAS detector in the first proton–proton collisions at  $\sqrt{s} = 13$  TeV*. ATL-PHYS-  
493 PUB-2015-027. 2015. URL: <https://cds.cern.ch/record/2037904>.
- 494 [10] P. S. A. Hoecker. ‘TMVA 4 Toolkit for Multivariate Data Analysis with ROOT’. In:  
495 *arXiv:physics/0703039* (2013).
- 496 [11] F. Cardillo et al. ‘Measurement of the fiducial and differential cross-section of a top quark  
497 pair in association with a Z boson at 13 TeV with the ATLAS detector’. In: ATL-COM-  
498 PHYS-2019-334 (Apr. 2019). URL: <https://cds.cern.ch/record/2672207>.

- 501 [12] ATLAS Collaboration. ‘Luminosity determination in pp collisions at  $\sqrt{s} = 7$  TeV using  
502 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).
- 504 [13] G. Aad et al. ‘Jet energy resolution in proton-proton collisions at  $\sqrt{s} = 7$  TeV recorded  
505 in 2010 with the ATLAS detector’. In: *The European Physical Journal C* 73.3 (Mar.  
506 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:  
507 <https://doi.org/10.1140/epjc/s10052-013-2306-0>.
- 508 [14] A. Collaboration. ‘Performance of b -jet identification in the ATLAS experiment’. In:  
509 *Journal of Instrumentation* 11.04 (2016), P04008. URL: <http://stacks.iop.org/1748-0221/11/i=04/a=P04008>.