



ATLAS Note

ATL-COM-PHYS-2019-962

15th March 2021



Draft version 0.1

1

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

4 The ATLAS Collaboration¹, Aaron Webb^a, Peter Onyisi^a, Heather
5 Russell^b, Philip Sommer^c

6 ^a*Univ. of Texas at Austin*

7 ^b*McGill University*

8 ^c*Univ. of Sheffield*

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

© 2021 CERN for the benefit of the ATLAS Collaboration.
14 Reproduction of this article or parts of it is allowed as specified in the CC-BY-4.0 license.

15 **Contents**

16	1 Changes and outstanding items	5
17	1.1 Changelog	5
18	1.1.1 Changes relative to v5	5
19	1.1.2 Changes relative to v4	5
20	1.1.3 Changes relative to v3	6
21	1.1.4 Changes relative to v2	6
22	1.1.5 Changes relative to v1	7
23	1.2 Outstanding Items	7
24	2 Executive Summary	8
25	3 Data and Monte Carlo Samples	9
26	3.1 Data Samples	9
27	3.2 Monte Carlo Samples	10
28	4 Object Reconstruction	11
29	4.1 Trigger	11
30	4.2 Light leptons	12
31	4.3 Jets	13
32	4.4 B-tagged Jets	13
33	4.5 Missing transverse energy	14
34	5 Event Selection and Signal Region Definitions	14
35	5.1 Event Preselection	14
36	5.2 Fit Regions	17
37	5.3 Non-Prompt Lepton Estimation	33
38	5.3.1 $t\bar{t}$ Validation	33
39	5.3.2 Z+jets Validation	37
40	6 tZ Separation Multivariate Analysis	41
41	6.1 Top Mass Reconstruction	42
42	6.2 tZ BDT	42
43	7 Systematic Uncertainties	46
44	8 Results	53
45	9 Conclusion	65
46	Appendices	68
47	.1 tZ Interference Studies	68
48	.2 Alternate tZ Inclusive Fit	72

50 **List of contributions**

Aaron Webb	Main analyser. Responsible for ntuple production, fits, and note writing.
Peter Onyisi	Advisor to Aaron Webb. Analysis design and implementation strategy.
Heather Russell	EWK group convener. Note editing, developing analysis strategy.
Philip Sommer	EWK group convener. Note editing, developing analysis strategy.

53 **1 Changes and outstanding items**

54 **1.1 Changelog**

55 This is version 6

56 **1.1.1 Changes relative to v5**

- 57 • added list of DSIDs to an appendix
58 • included systematics on jet migrations
59 • Updated results section to include simultaneous fit over 1-jet and 2-jet bins, describe
60 unfolding procedure
61 • Updated other sections to account for this change
62 • Included info about migrations in Section 5.2

63 **1.1.2 Changes relative to v4**

- 64 • Fixed various typos, clarified wording
65 • Expanded info about JER uncertainties, electron systematics, theory uncertainties
66 • removed a table on lepton selection, included information in the text instead
67 • Plotted lepton Pt Z and W for Zjet CR, rather than lep 1 and 2
68 • fixed binning in kinematic plots
69 • Included prefit and postfit yield tables
70 • added signal modelling systematics
71 • included alternate fit studies with tZ included in signal

72 1.1.3 Changes relative to v3

- 73 • Merged introduction into executive summary, including unblinding details and list of
74 SRs/CRs used
- 75 • listed ptag used (p4133), and release (AB 21.2.127)
- 76 • Included table reftab:xsecUnc listing x-sec uncertainties used
- 77 • Removed selection criteria listed in table ?? (QMisID, AmbiguityType) that were removed
78 from the analysis
- 79 • specified variable used to make truth jet flavor determination (HadronConeExclTruthLa-
80 belID)
- 81 • fixed bug in MtLepMet calculation, updated selection/fits to account for this
- 82 • Included plots of MtLepMet and PtZ, swapped lep 1 and 2 p_T plots for lep W and lep Z
83 plots
- 84 • updated tZ BDT training to reduce overfitting, updated plots to include error bars, feature
85 importance
- 86 • updated table 7 to clarify selection, fix the tZ_BDT cut used
- 87 • replace a few broken ntuples which included large weight events
- 88 • include DL1r distribution for Z+jets and tt} VRs
- 89 • Expanded section on fakes, included information on derived scale factors from VRs.
- 90 • Changed the kinematic plots to include $p_T(Z)$ and $m_T(W)$, list lepton p_T based on W and
91 Z candidates.

92 1.1.4 Changes relative to v2

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

100 **1.1.5 Changes relative to v1**

- 101 • Added GRL list
102 • Fixed latex issue in line 92, typo in line 172
103 • Added tables 5 and ??, summarizing the event and object selection
104 • Added table 2, which includes the DSID of samples used
105 • Included reference to WZ inclusive paper in introduction

106 **1.2 Outstanding Items**

- 107 • Complete interference studies, apply any interference effects observed as a systematic
108 • Update results section with additional studies, possibly including:
109 – Truth jet migration studies
110 – Simultaneous fit over 1j and 2j
111 – Impact of allowing tZ to float
112 • Unblind, update plots and fits to include data
113 • Add cross-section, significance once unblinded

114 2 Executive Summary

115 The production of WZ in association with a heavy flavor jet represents an important background
 116 for many major analyses. This includes any process with leptons and b-jets in the final state,
 117 such as $t\bar{t}H$, $t\bar{t}W$, and $t\bar{t}Z$. While precise measurements have been made of WZ production
 118 [1], $WZ +$ heavy flavor remains poorly understood. This is largely because the QCD processes
 119 involved in the production of the b-jet make it difficult to simulate accurately. This introduces a
 120 large uncertainty for analyses that include this process as a background.

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

126 Events that meet this selection criteria are sorted into pseudo-continuous b-tagging regions based
 127 on the DL1r b-tag score of their associated jets. This is done to separate $WZ +$ b-jet events from
 128 $WZ +$ charm and $WZ +$ light jets. These regions are fit to data in order make a more accurate
 129 estimate of the contribution of $WZ +$ heavy-flavor, where heavy-flavor jets include b-jets and
 130 charm jets. The full Run-2 dataset collected by the ATLAS detector, representing 139 fb^{-1} of
 131 data from pp collisions at $\sqrt{s} = 13 \text{ TeV}$, is used for this study.

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

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

141 The analysis aims to report a cross-section measurement of $WZ+b$ and $WZ+charm$, along with
 142 their correlations, for both 1-jet and 2-jet exclusive regions. The proposed fiducial region for
 143 these measurements includes events with three leptons, where the invariant mass of at least one
 144 opposite charge, same flavor lepton pair falls within 10 GeV of 91.2 GeV , with 1 or 2 associated
 145 jets. An alternate version of the measurement is included in the appendix, which considers tZ as
 146 part of the $WZ+b$ signal.

147 The current state of the analysis shows blinded results for the full Run 2 dataset. Regions
 148 containing $>5\%$ $WZ+b$ events are blinded, and results are from Asimov, MC only fits. Expected
 149 significance and cross-section numbers are reported.

150 **3 Data and Monte Carlo Samples**

151 Both data and Monte Carlo samples used in this analysis were prepared in the `xAOD` format,
 152 which was used to produce a `DxAOD` sample in the `HIGG8D1` derivation framework. The `HIGG8D1`
 153 framework is designed for the $t\bar{t}H$ multi-lepton analysis, which targets events with multiple
 154 leptons as well as tau hadrons. This framework skims the dataset to remove unneeded variables
 155 as well as entire events. Events are removed from the derivations that do not meet the following
 156 selection:

- 157 • at least two light leptons within a range $|\eta| < 2.6$, with leading lepton $p_T > 15$ GeV and
 158 subleading lepton $p_T > 5$ GeV
- 159 • at least one light lepton with $p_T > 15$ GeV within a range $|\eta| < 2.6$, and at least two hadronic
 160 taus with $p_T > 15$ GeV.

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

163 **3.1 Data Samples**

164 The study uses a sample of proton-proton collision data collected by the ATLAS detector from
 165 2015 through 2018 at an energy of $\sqrt{s} = 13$ TeV, which represents an integrated luminosity of
 166 139 fb^{-1} . This data set was collected with a bunch crossing rate of 25 ns. All data used in this
 167 analysis was verified by data quality checks, having been included in the following Good Run
 168 Lists:

- 169 • `data15_13TeV.periodAllYear_DetStatus-v79-repro20-02_DQDefects-00-02-02`
 170 `_PHYS_StandardGRL_All_Good_25ns.xml`
- 171 • `data16_13TeV.periodAllYear_DetStatus-v88-pro20-21_DQDefects-00-02-04`
 172 `_PHYS_StandardGRL_All_Good_25ns.xml`
- 173 • `data17_13TeV.periodAllYear_DetStatus-v97-pro21-13_Unknown_PHYS_StandardGRL`
 174 `_All_Good_25ns_Triggerno17e33prim.xml`
- 175 • `data18_13TeV.periodAllYear_DetStatus-v102-pro22-04_Unknown_PHYS_StandardGRL`
 176 `_All_Good_25ns_Triggerno17e33prim.xml`

177 Runs included from the AllYear period containers are included.

178 **3.2 Monte Carlo Samples**

179 Several different generators were used to produce Monte Carlo simulations of the signal and
180 background processes. For all samples, the response of the ATLAS detector is simulated using
181 Geant4. The WZ signal samples are simulated using Sherpa 2.2.2 [2]. Specific information
182 about the Monte Carlo samples being used can be found in Table 1. A list of the specific samples
183 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
t̄tW	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 [3] (CT10 [4])
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
t̄t̄t, t̄tt̄t̄t̄	MG5_AMC	LO	PYTHIA 8	NNPDF 2.3 LO
t̄tW+W-	MG5_AMC	LO	PYTHIA 8	NNPDF 2.3 LO
t̄t	POWHEG-BOX v2 [5]	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 [6]	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.

184 4 Object Reconstruction

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

188 4.1 Trigger

189 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.190 **4.2 Light leptons**

191 Electron candidates are reconstructed from energy clusters in the electromagnetic calorimeter
 192 that are associated with charged particle tracks reconstructed in the inner detector [7]. Electron
 193 candidates are required to have $p_T > 10$ GeV and $|\eta_{\text{cluster}}| < 2.47$. Muon candidates are
 194 reconstructed by combining inner detector tracks with track segments or full tracks in the muon
 195 spectrometer [8]. Muon candidates are required to have $p_T > 10$ GeV and $|\eta| < 2.5$. Candidates
 196 in the transition region between different electromagnetic calorimeter components, $1.37 <$
 197 $|\eta_{\text{cluster}}| < 1.52$, are rejected. A multivariate likelihood discriminant combining shower shape
 198 and track information is used to distinguish real electrons from hadronic showers (fake electrons).
 199 To further reduce the non-prompt electron contribution, the track is required to be consistent
 200 with originating from the primary vertex; requirements are imposed on the transverse impact
 201 parameter significance ($|d_0|/\sigma_{d_0} < 5$) and the longitudinal impact parameter ($|\Delta z_0 \sin \theta_\ell| < 0.5$
 202 mm).

203 Muon candidates are reconstructed by combining inner detector tracks with track segments or
 204 full tracks in the muon spectrometer [8]. Muon candidates are required to have $p_T > 10$ GeV
 205 and $|\eta| < 2.5$. The longitudinal impact parameter is the same for both electrons and muons, while
 206 muons are required to pass a slightly tighter transverse impact parameter, $|d_0|/\sigma_{d_0} < 3$.

207 All leptons are required to be isolated, as defined through the standard `isolationFixedCutLoose`
 208 working point supported by combined performance groups. Leptons are additionally required to
 209 pass a non-prompt BDT selection described in detail in [9]. Optimized working points and scale
 210 factors for this BDT are taken from that analysis.

211 4.3 Jets

212 Jets are reconstructed from calibrated topological clusters built from energy deposits in the
 213 calorimeters [10], using the anti- k_t algorithm with a radius parameter $R = 0.4$. Jets with energy
 214 contributions likely arising from noise or detector effects are removed from consideration [11],
 215 and only jets satisfying $p_T > 25$ GeV and $|\eta| < 2.5$ are used in this analysis. For jets with
 216 $p_T < 60$ GeV and $|\eta| < 2.4$, a jet-track association algorithm is used to confirm that the jet
 217 originates from the selected primary vertex, in order to reject jets arising from pileup collisions
 218 [12].

219 4.4 B-tagged Jets

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

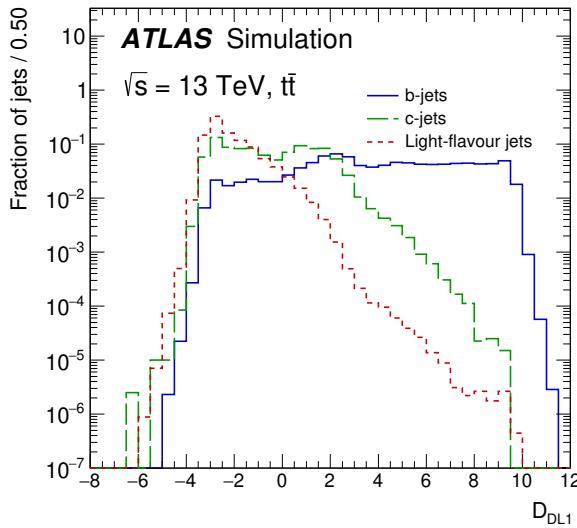


Figure 1: Output distribution of the DL1r algorithm for b-jets, charm jets, and light jets

225 From the output of the BDT, working points (WPs) are developed based on the efficiency of truth
 226 b-jets at particular values of the DL1r algorithm. The working points used in this analysis are
 227 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

228 A tighter WP will accept fewer b-jets, but reject a higher fraction of charm and light jets.
 229 Generally, analyses that include b-jets will use a fixed working point, for example, requiring that
 230 a jet pass the 70% threshold. By instead treating these working point as bins, e.g. events with
 231 jets that fall between the 85% and 77% WPs fall into one bin, while events with jets passing the
 232 60% WP fall into another, and looking at the full psuedo-continuous DL1r spectrum of the jets,
 233 additional information can be gained. The psuedo-continuous b-tag spectrum is used in this case
 234 to separate out WZ + b, WZ + charm, and WZ + light.

235 **4.5 Missing transverse energy**

236 Missing transverse momentum (E_T^{miss}) is used as part of the event selection. The missing
 237 transverse momentum vector is defined as the inverse of the sum of the transverse momenta of
 238 all reconstructed physics objects as well as remaining unclustered energy, the latter of which is
 239 estimated from low- p_T tracks associated with the primary vertex but not assigned to a hard object,
 240 with object definitions taken from [13]. Light leptons considered in the E_T^{miss} reconstruction are
 241 required to have $p_T > 10 \text{ GeV}$, while jets are required to have $p_T > 20 \text{ GeV}$.

242 **5 Event Selection and Signal Region Definitions**

243 Event are required to pass a preselection described in Section 5.1 and summarized in Table 5.
 244 Those that pass this preselection are divided into various fit regions described in Section 5.2,
 245 based on the number of jets in the event, and the b-tag score of those jets.

246 **5.1 Event Preselection**

247 Events are required to include exactly three reconstructed light leptons passing the requirement
 248 described in 4.2, which have a total charge of ± 1 . As the opposite sign lepton is found to be
 249 prompt the vast majority of the time [9], it is required to be loose and isolated, as defined though
 250 the standard `isolationFixedCutLoose` working point supported by combined performance

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

253 The leptons are ordered in the analysis code as 0, 1, and 2. Lepton 0 is the lepton whose charge
 254 is opposite the other two. Lepton 1 is the lepton closest to the opposite charge lepton, i.e. the
 255 smallest ΔR , leaving lepton 2 as the lepton further from the opposite charge lepton. Lepton 0
 256 is required to have $p_T > 10$ GeV, while the same sign leptons, 1 and 2, are required to have
 257 $p_T > 20$ GeV to reduce the contribution of non-prompt leptons.

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

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

264 Events are required to have one or two reconstructed jets passing the selection described in
 265 Section 4.3. Events with more than two jets are rejected in order to reduce the contribution of
 266 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 ± 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

267 The event yields in the preselection region for both data and Monte Carlo are summarized in
 268 Table 5.1, which shows good agreement between data and Monte Carlo, and demonstrates that
 269 this region consists primarily of WZ events. The WZ events are split into WZ + b, WZ + c, and
 270 WZ + l based on the truth flavor of the associated jet in the event. Specifically, this determination
 271 is made based on the `HadronConeExclTruthLabelID` of the jet, as recommended by the b-
 272 tagging working group [14]. In this ordering b-jet supersedes charm, which supersedes light. That
 273 is, WZ + l events contain no charm and no b jets at truth level, WZ + c contain at least one truth
 274 charm and no b-jets, and WZ + b contains at least one truth b-jet.

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

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

²⁷⁵ Here Other VV represents diboson processes other than WZ, and consists predominantly of
²⁷⁶ $ZZ \rightarrow ll\bar{l}\bar{l}$ events where one of the leptons is not reconstructed.

²⁷⁷ Simulations are further validated by comparing the kinematic distributions of the Monte Carlo
²⁷⁸ with data, which are shown in Figure 2. Here, bins with 5% or more WZ+b are blinded.

WZ Fit Region - Inclusive

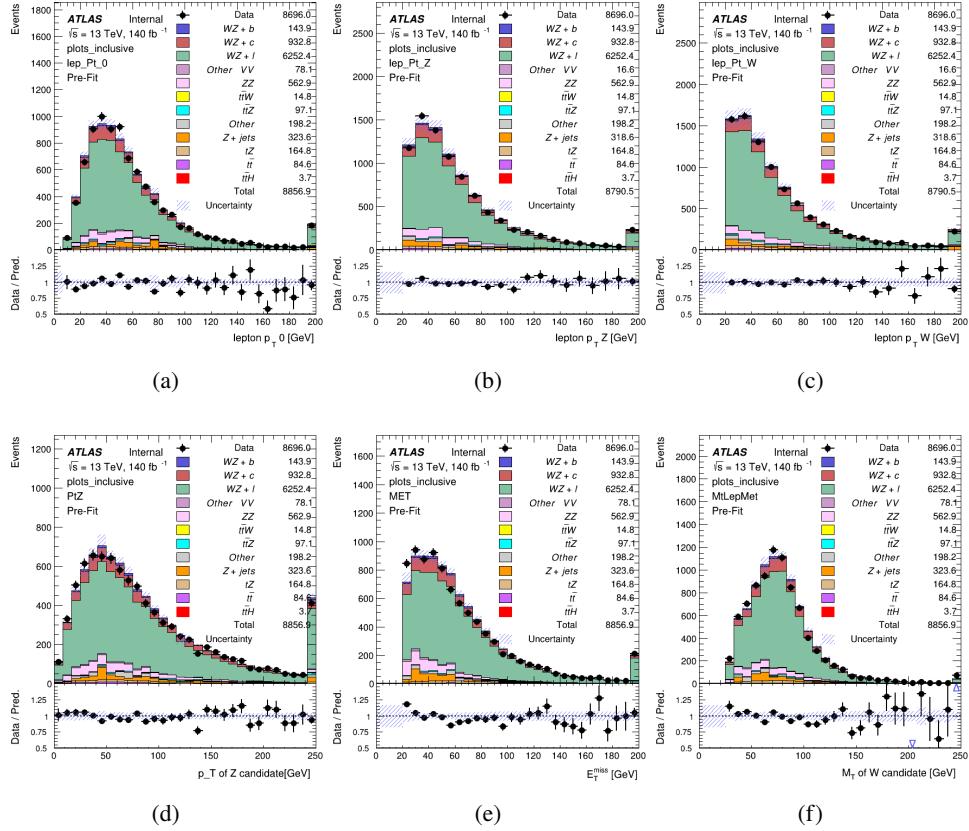


Figure 2: Comparisons between the data and MC distributions in the preselection region for (a) the p_T of the opposite sign lepton, (b) the p_T of the other lepton from the Z candidate, (c) the p_T of the lepton from the W candidate, (d) the p_T of the Z candidate, (e) the E_T^{miss} , and (f) the M_T 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%	$N_{\text{jets}} = 1, n\text{Jets}_{\text{DL1r}_85} = 0$
1j, 85%-77%	$N_{\text{jets}} = 1, n\text{Jets}_{\text{DL1r}_85} = 1, n\text{Jets}_{\text{DL1r}_77} = 0$
1j, 77%-70%	$N_{\text{jets}} = 1, n\text{Jets}_{\text{DL1r}_77} = 1, n\text{Jets}_{\text{DL1r}_70} = 0$
1j, 70%-60%	$N_{\text{jets}} = 1, n\text{Jets}_{\text{DL1r}_70} = 1, n\text{Jets}_{\text{DL1r}_60} = 0$
1j, >60%	$N_{\text{jets}} = 1, n\text{Jets}_{\text{DL1r}_60} = 1, tZ \text{ BDT} > 0.725$
1j tZ CR	$N_{\text{jets}} = 1, n\text{Jets}_{\text{DL1r}_60} = 1, tZ \text{ BDT} < 0.725$
2j, <85%	$N_{\text{jets}} = 2, n\text{Jets}_{\text{DL1r}_85} = 0$
2j, 85%-77%	$N_{\text{jets}} = 2, n\text{Jets}_{\text{DL1r}_85} \geq 1, n\text{Jets}_{\text{DL1r}_77} = 0$
2j, 77%-70%	$N_{\text{jets}} = 2, n\text{Jets}_{\text{DL1r}_77} \geq 1, n\text{Jets}_{\text{DL1r}_70} = 0$
2j, 70%-60%	$N_{\text{jets}} = 2, n\text{Jets}_{\text{DL1r}_70} \geq 1, n\text{Jets}_{\text{DL1r}_60} = 0$
2j, >60%	$N_{\text{jets}} = 2, n\text{Jets}_{\text{DL1r}_60} \geq 1, tZ \text{ BDT} > 0.725$
2j tZ CR	$N_{\text{jets}} = 2, n\text{Jets}_{\text{DL1r}_60} \geq 1, tZ \text{ BDT} < 0.725$

282 The working points discussed in Section 4.4 are used to separate events into fit regions based on
 283 the highest working point reached by a jet in each event. Because the background composition
 284 differs significantly based on the number of b-jets, events are further subdivided into 1-jet and
 285 2-jet regions in order to minimize the impact of background uncertainties.

286 An unfolding procedure is performed to account for differences in the number of reconstructed
 287 jets compared to the number of truth jets in each event. In order to account for migration of
 288 WZ+1-jet and WZ+2-jet events between the 1-jet and 2-jet bins at reco level, the signal samples
 289 are separated based on the number of truth jets. Events with 0 jets or more than 3 jets at truth
 290 level, yet fall within one of the categories listed in Table 7, are categorized as WZ + other, and
 291 treated as a background. The migration matrix in the number of jets at truth level versus reco
 292 level is shown in Figure 3. The composition of the number of truth jets in each reco jet bin is
 293 taken from MC, with uncertainties in these estimates described in detail in Section 7.

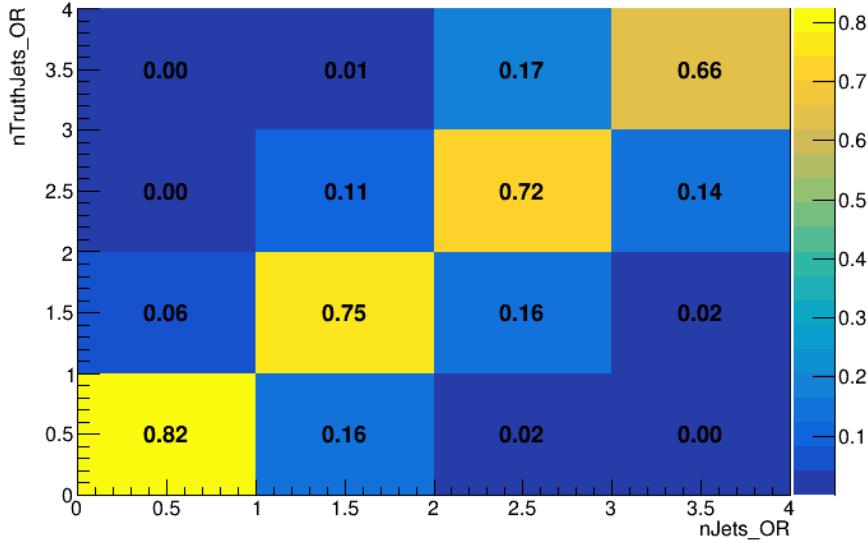


Figure 3: Number of truth jets compared to the number of reconstructed jets in each WZ event falling in the preselection region. Each row is normalized to unity.

294 An additional tZ control region is created based on the BDT described in Section 6. The region
 295 with 1-jet passing the 60% working point is split in two - a signal enriched region of events with
 296 a BDT score greater than 0.03, and a tZ control region including events with less than 0.03. This
 297 cutoff is arrived at by performing an Asimov fit with a variety of cutoffs, and selecting the value
 298 that produces the highest significance for the measurement of WZ + b.

299 The modeling in each region is validated by comparing data and MC predictions for various
 300 kinematic distributions. These plot are shown in Figures 4-17.

WZ Fit Region - 1j Inclusive

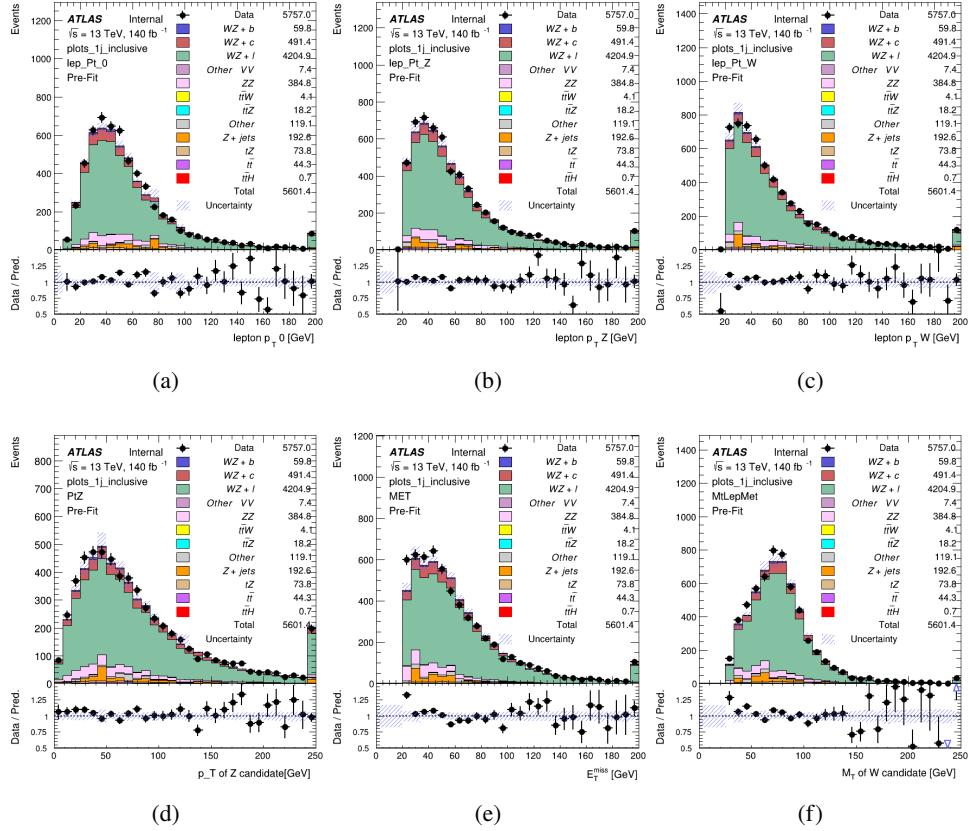


Figure 4: Comparisons between the data and MC distributions in the preselection region for (a) the p_T of the opposite sign lepton, (b) the p_T of the other lepton from the Z candidate, (c) the p_T of the lepton from the W candidate, (d) the p_T of the Z candidate, (e) the E_T^{miss} , and (f) the M_T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

WZ Fit Region - 1j < 85% WP

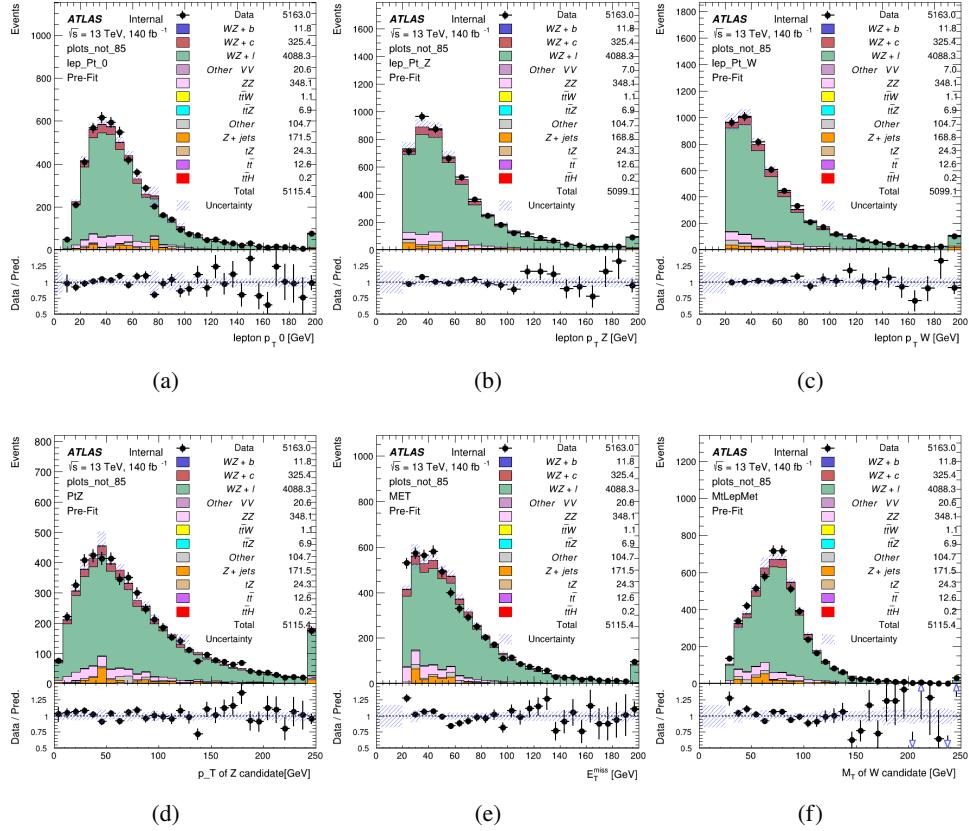


Figure 5: Comparisons between the data and MC distributions in the preselection region for (a) the p_T of the opposite sign lepton, (b) the p_T of the other lepton from the Z candidate, (c) the p_T of the lepton from the W candidate, (d) the p_T of the Z candidate, (e) the E_T^{miss} , and (f) the m_T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

WZ Fit Region - 1j 77-85% WP

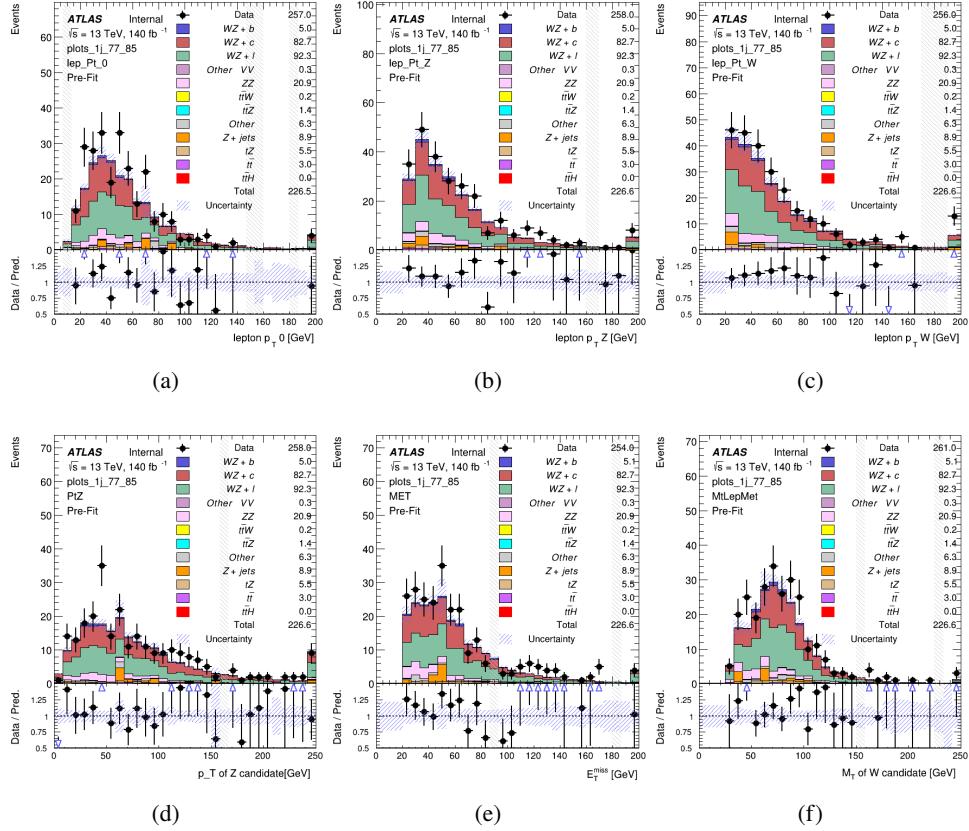


Figure 6: Comparisons between the data and MC distributions in the preselection region for (a) the p_T of the opposite sign lepton, (b) the p_T of the other lepton from the Z candidate, (c) the p_T of the lepton from the W candidate, (d) the p_T of the Z candidate, (e) the E_T^{miss} , and (f) the m_T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

WZ Fit Region - 1j 70-77% WP

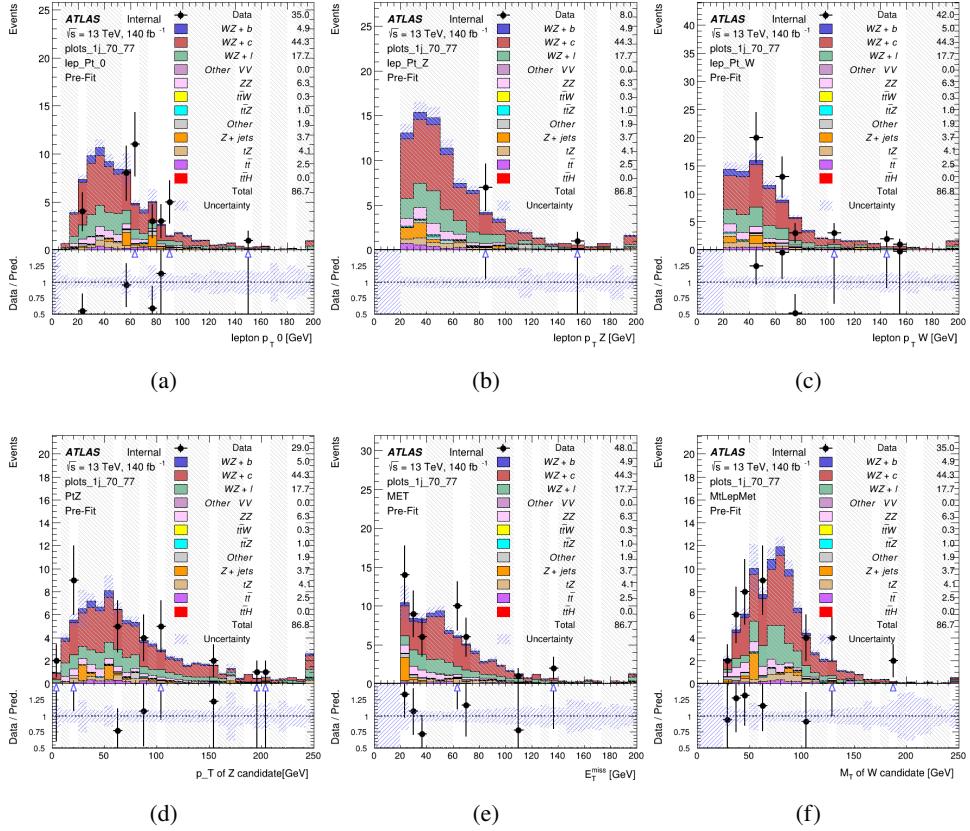


Figure 7: Comparisons between the data and MC distributions in the preselection region for (a) the p_T of the opposite sign lepton, (b) the p_T of the other lepton from the Z candidate, (c) the p_T of the lepton from the W candidate, (d) the p_T of the Z candidate, (e) the E_T^{miss} , and (f) the m_T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

WZ Fit Region - 1j 60-70% WP

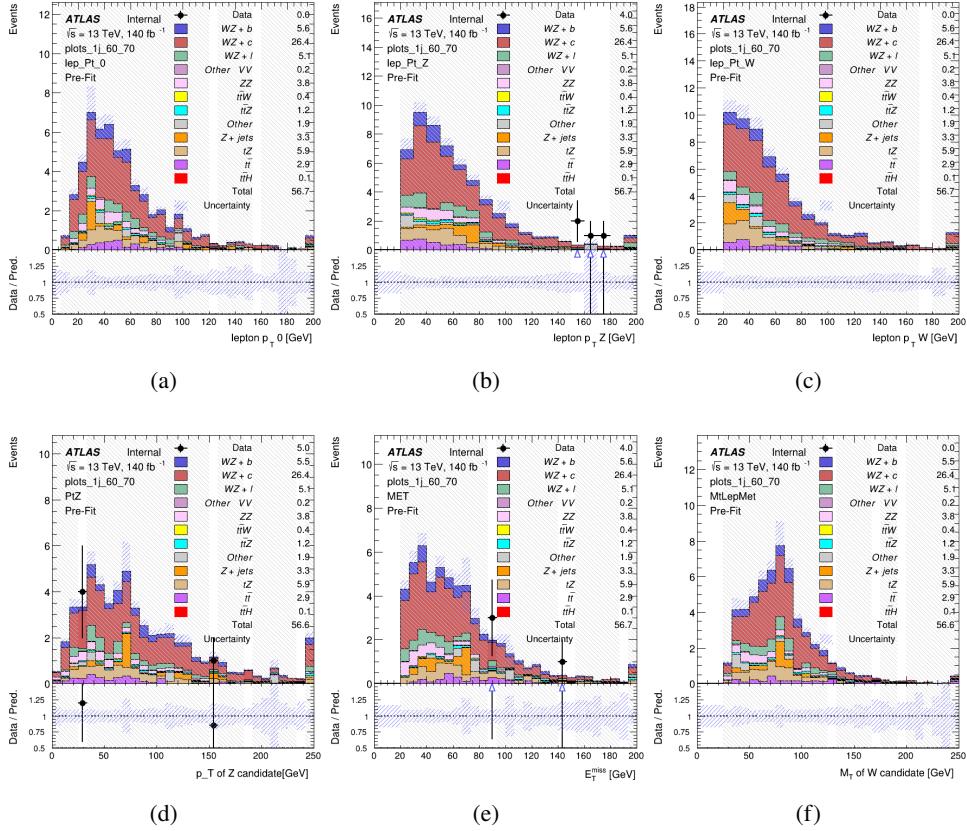


Figure 8: Comparisons between the data and MC distributions in the preselection region for (a) the p_T of the opposite sign lepton, (b) the p_T of the other lepton from the Z candidate, (c) the p_T of the lepton from the W candidate, (d) the p_T of the Z candidate, (e) the E_T^{miss} , and (f) the m_T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

WZ Fit Region - 1j 60% WP

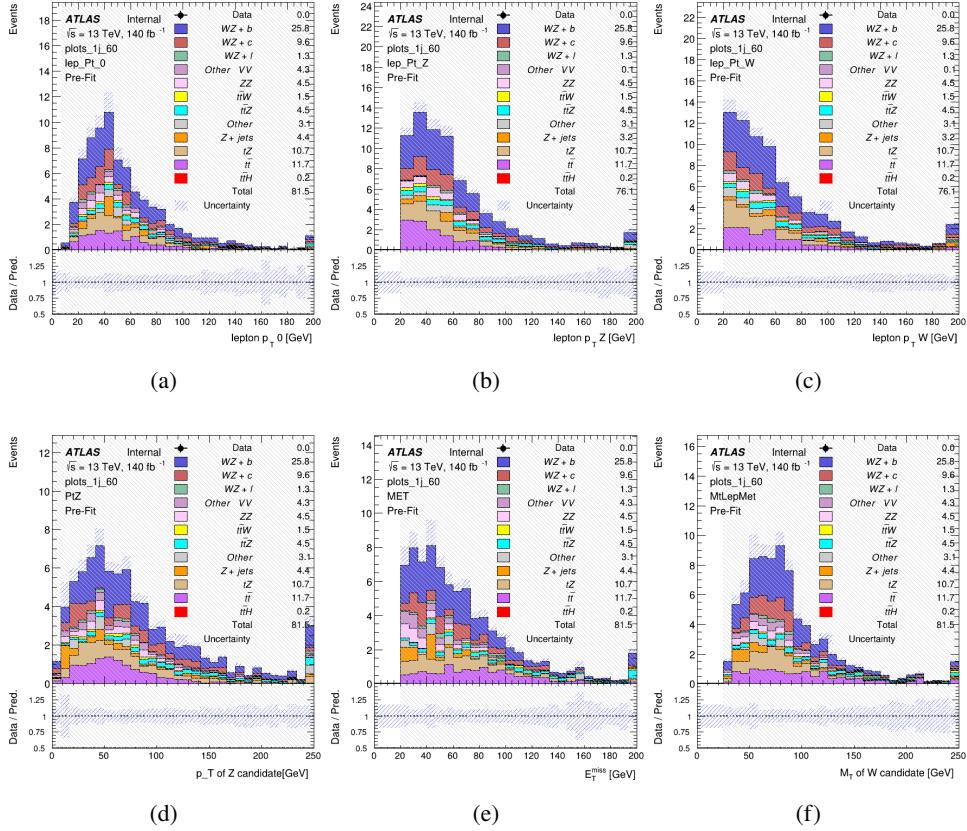


Figure 9: Comparisons between the data and MC distributions in the preselection region for (a) the p_T of the opposite sign lepton, (b) the p_T of the other lepton from the Z candidate, (c) the p_T of the lepton from the W candidate, (d) the p_T of the Z candidate, (e) the E_T^{miss} , and (f) the m_T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

WZ Fit Region - tZ-CR

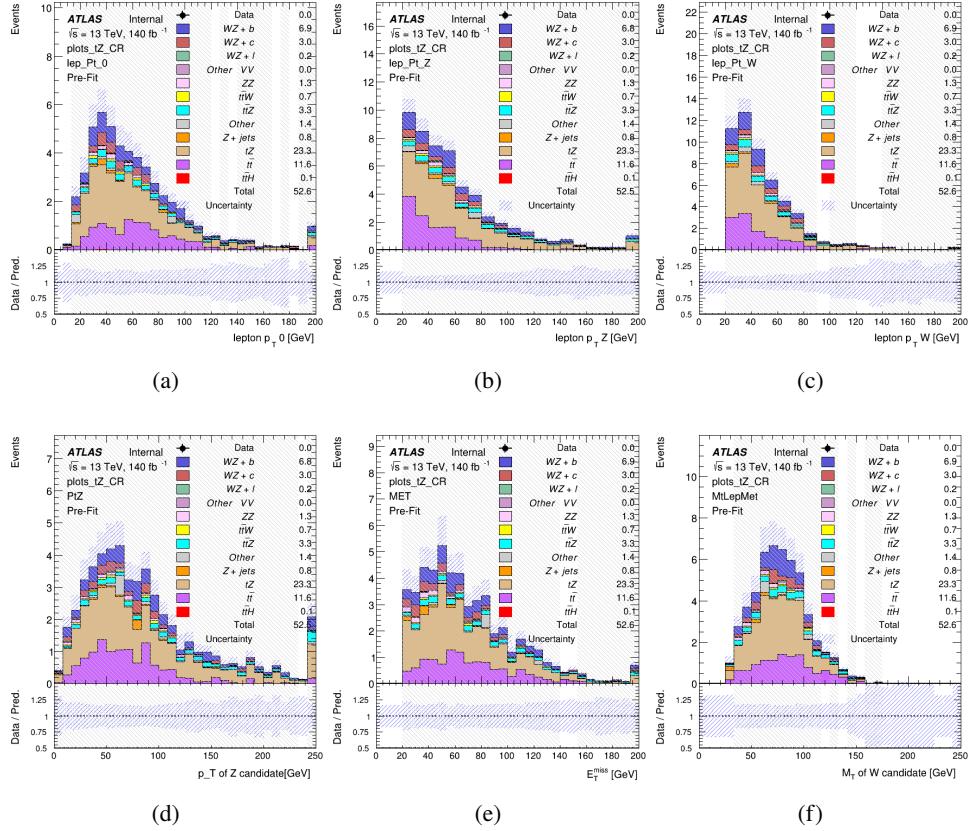


Figure 10: Comparisons between the data and MC distributions in the preselection region for (a) the p_T of the opposite sign lepton, (b) the p_T of the other lepton from the Z candidate, (c) the p_T of the lepton from the W candidate, (d) the p_T of the Z candidate, (e) the E_T^{miss} , and (f) the M_T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

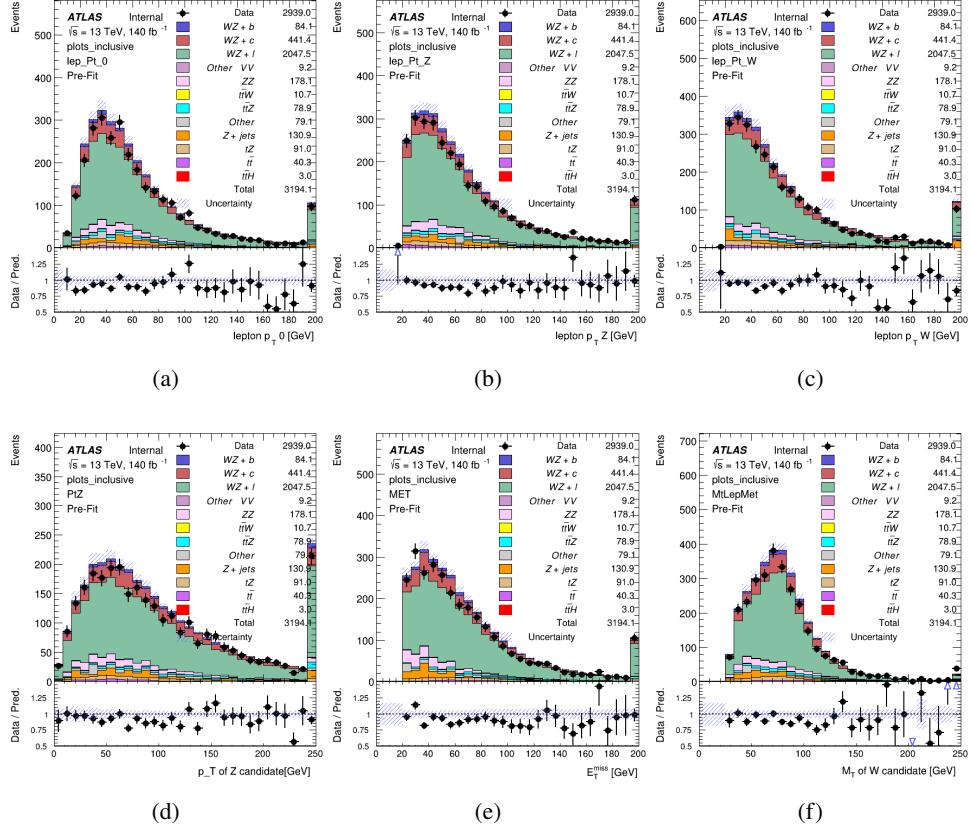
WZ Fit Region - 2j Inclusive

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^{miss} , and (f) the m_T of the W candidate.

WZ Fit Region - 2j < 85% WP

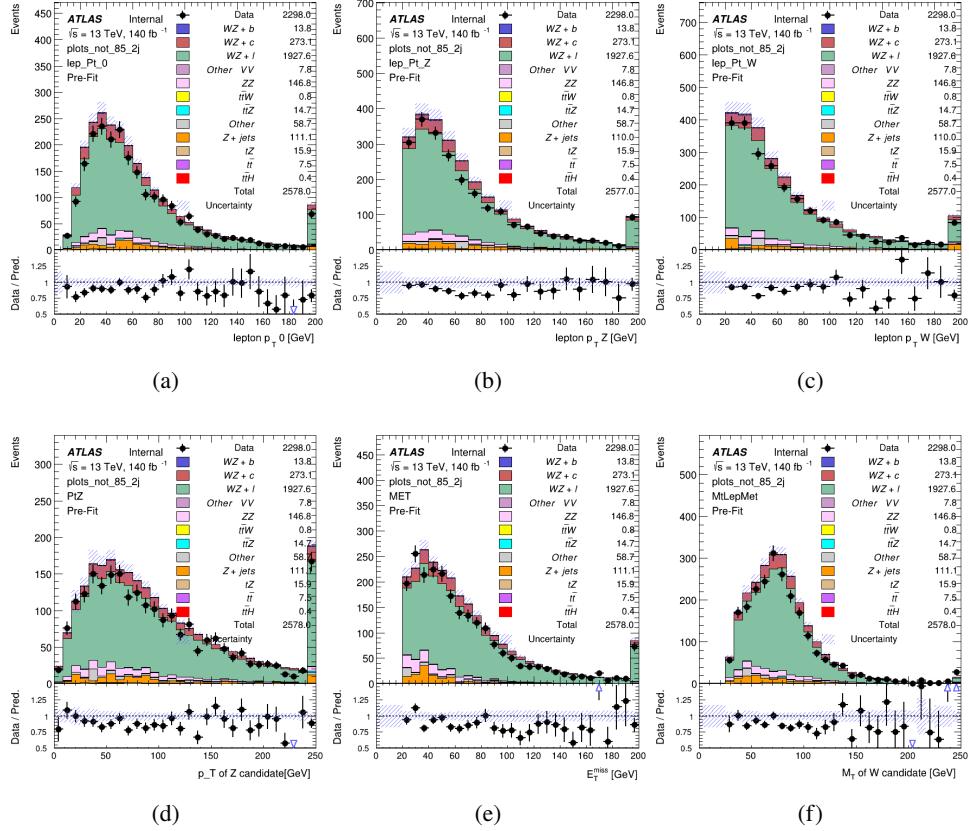


Figure 12: Comparisons between the data and MC distributions in the preselection region for (a) the p_T of the opposite sign lepton, (b) the p_T of the other lepton from the Z candidate, (c) the p_T of the lepton from the W candidate, (d) the p_T of the Z candidate, (e) the E_T^{miss} , and (f) the M_T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

WZ Fit Region - 2j 77-85% WP

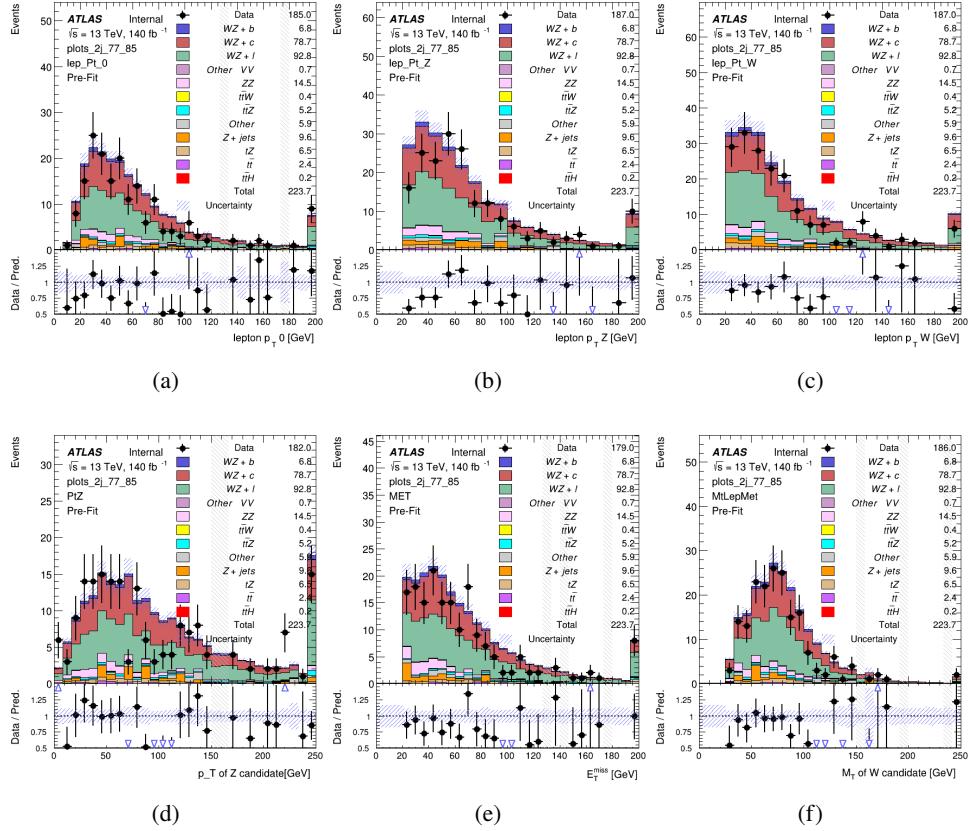


Figure 13: Comparisons between the data and MC distributions in the preselection region for (a) the p_T of the opposite sign lepton, (b) the p_T of the other lepton from the Z candidate, (c) the p_T of the lepton from the W candidate, (d) the p_T of the Z candidate, (e) the E_T^{miss} , and (f) the m_T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

WZ Fit Region - 2j 70-77% WP

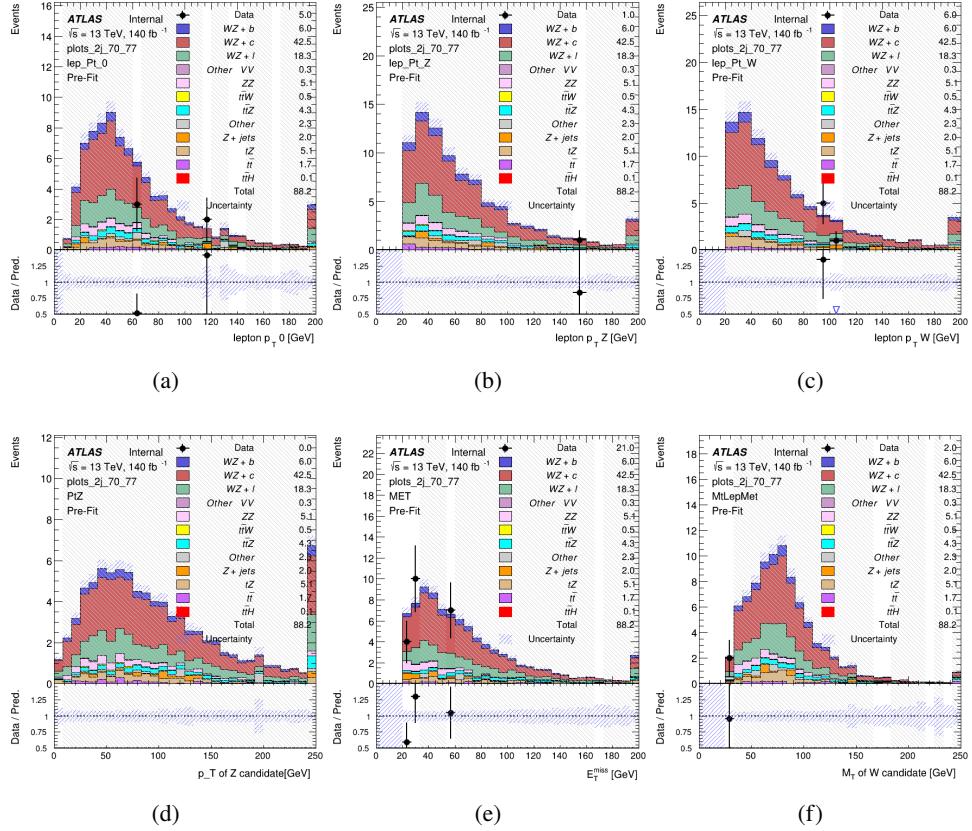


Figure 14: Comparisons between the data and MC distributions in the preselection region for (a) the p_T of the opposite sign lepton, (b) the p_T of the other lepton from the Z candidate, (c) the p_T of the lepton from the W candidate, (d) the p_T of the Z candidate, (e) the E_T^{miss} , and (f) the m_T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

WZ Fit Region - 2j 60-70% WP

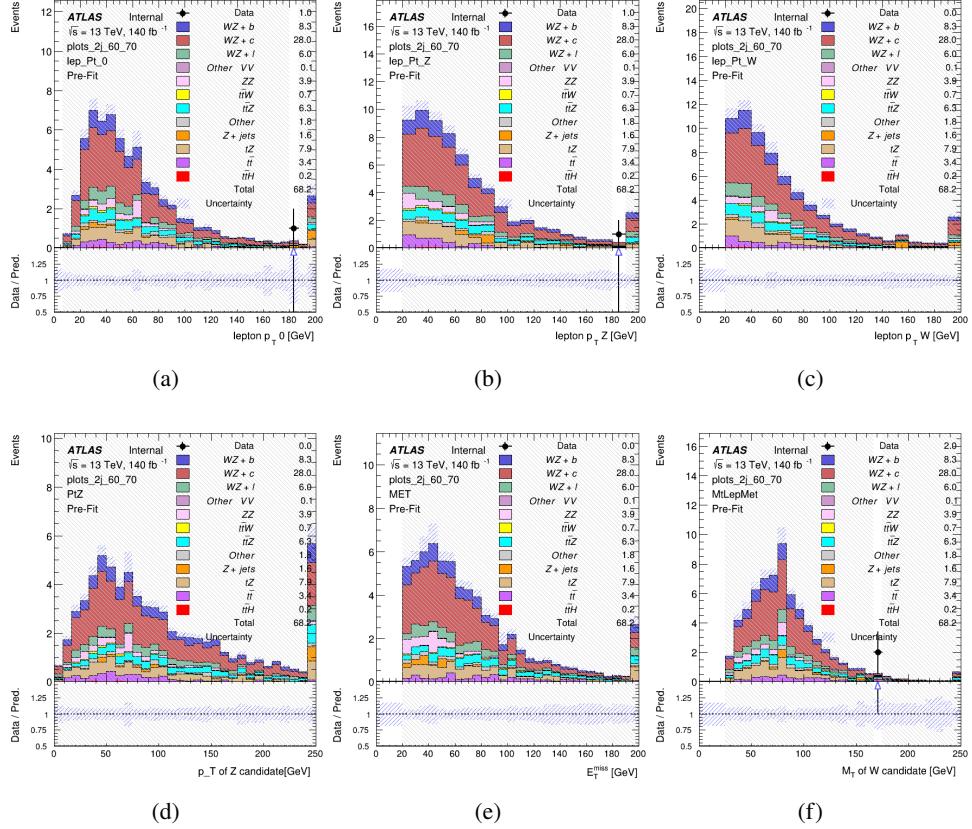


Figure 15: Comparisons between the data and MC distributions in the preselection region for (a) the p_T of the opposite sign lepton, (b) the p_T of the other lepton from the Z candidate, (c) the p_T of the lepton from the W candidate, (d) the p_T of the Z candidate, (e) the E_T^{miss} , and (f) the m_T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

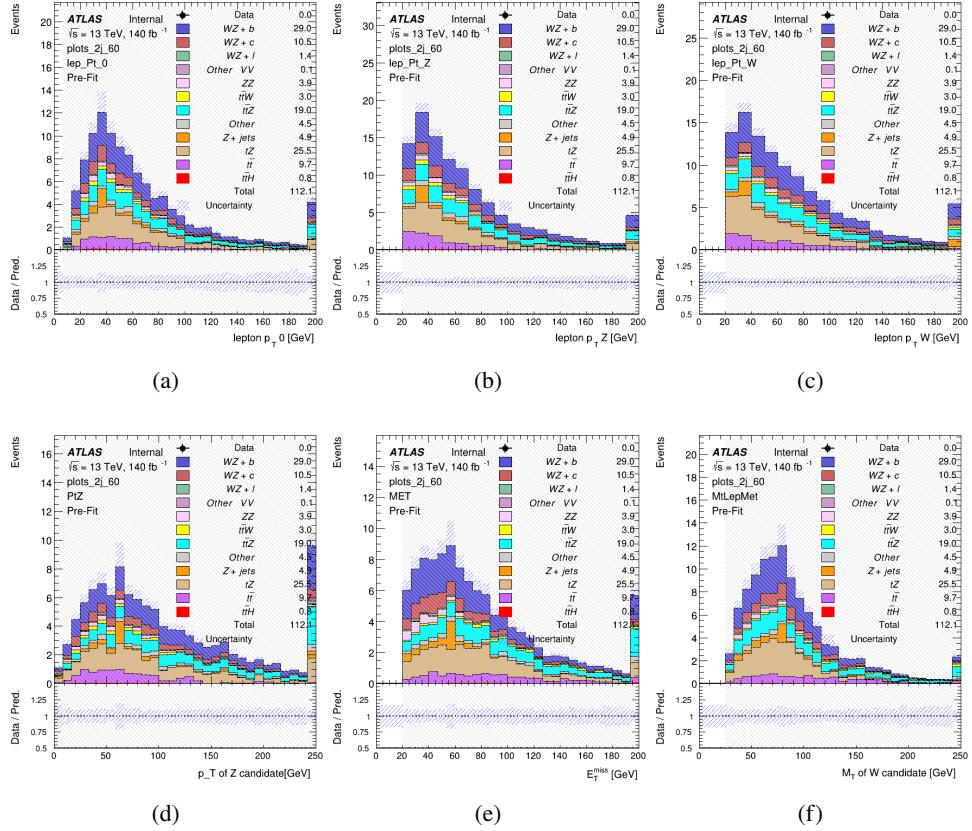
WZ Fit Region - 2j 60% WP

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^{miss} , and (f) the m_T of the W candidate.

WZ Fit Region - tZ-CR-2j

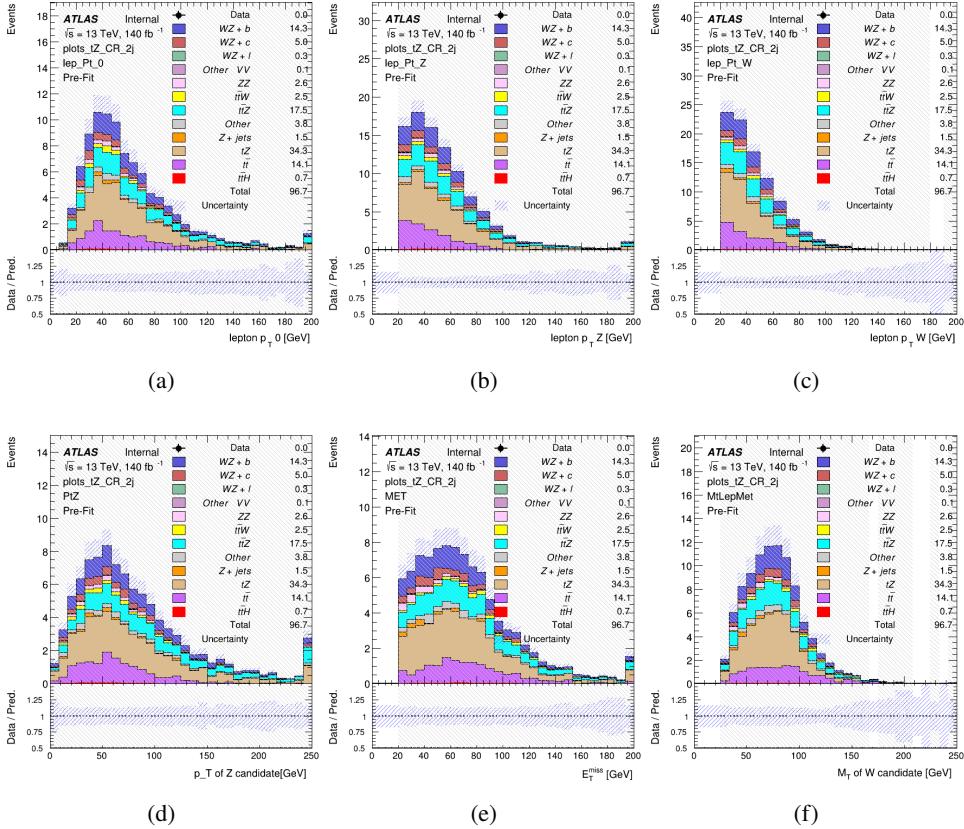


Figure 17: Comparisons between the data and MC distributions in the preselection region for (a) the p_T of the opposite sign lepton, (b) the p_T of the other lepton from the Z candidate, (c) the p_T of the lepton from the W candidate, (d) the p_T of the Z candidate, (e) the E_T^{miss} , and (f) the m_T 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

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

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

317 Further, because the jet multiplicity of $t\bar{t}$ events tends to be higher than WZ, the number of jets
318 in each event is required to be greater than 1. As b-jets are almost invariably produced from top
319 decays, at least one b-tagged jet passing the 70% DL1r WP in each event is required. Various
320 kinematic plots of this region are shown in Figure 18.

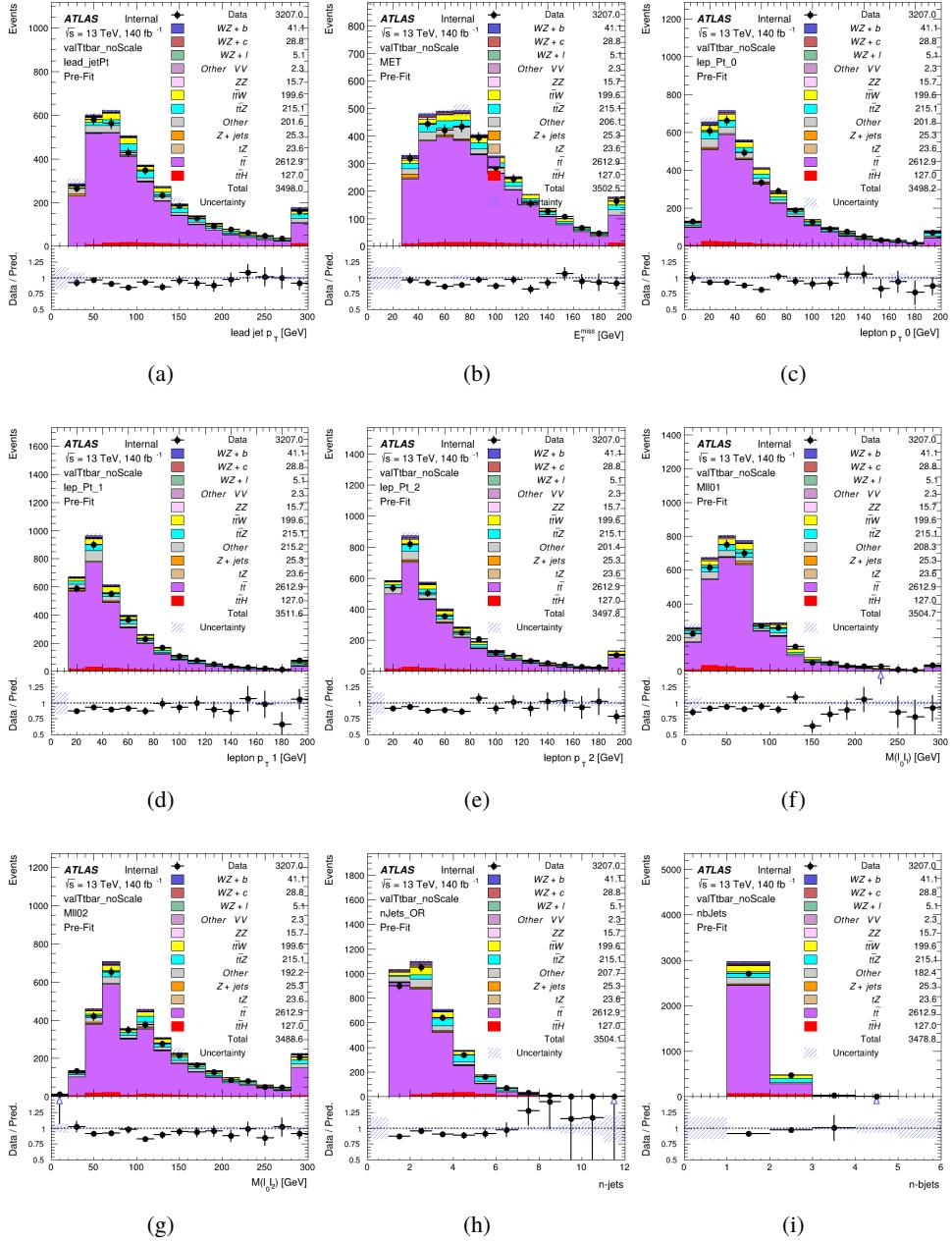


Figure 18: Comparisons between the data and MC distributions in the $t\bar{t}$ validation region for (a) the p_T of the leading jet, (b) the missing transverse energy, (c) the p_T of lepton 0, (d) p_T of lepton 1, (e) p_T of lepton 2, (f) the invariant mass of leptons 0 and 1, (g) the invariant mass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

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

323 MC prediction. Plots showing the kinematics of the $t\bar{t}$ VR after this correction factor has been
 324 applied are shown in Figure 19.

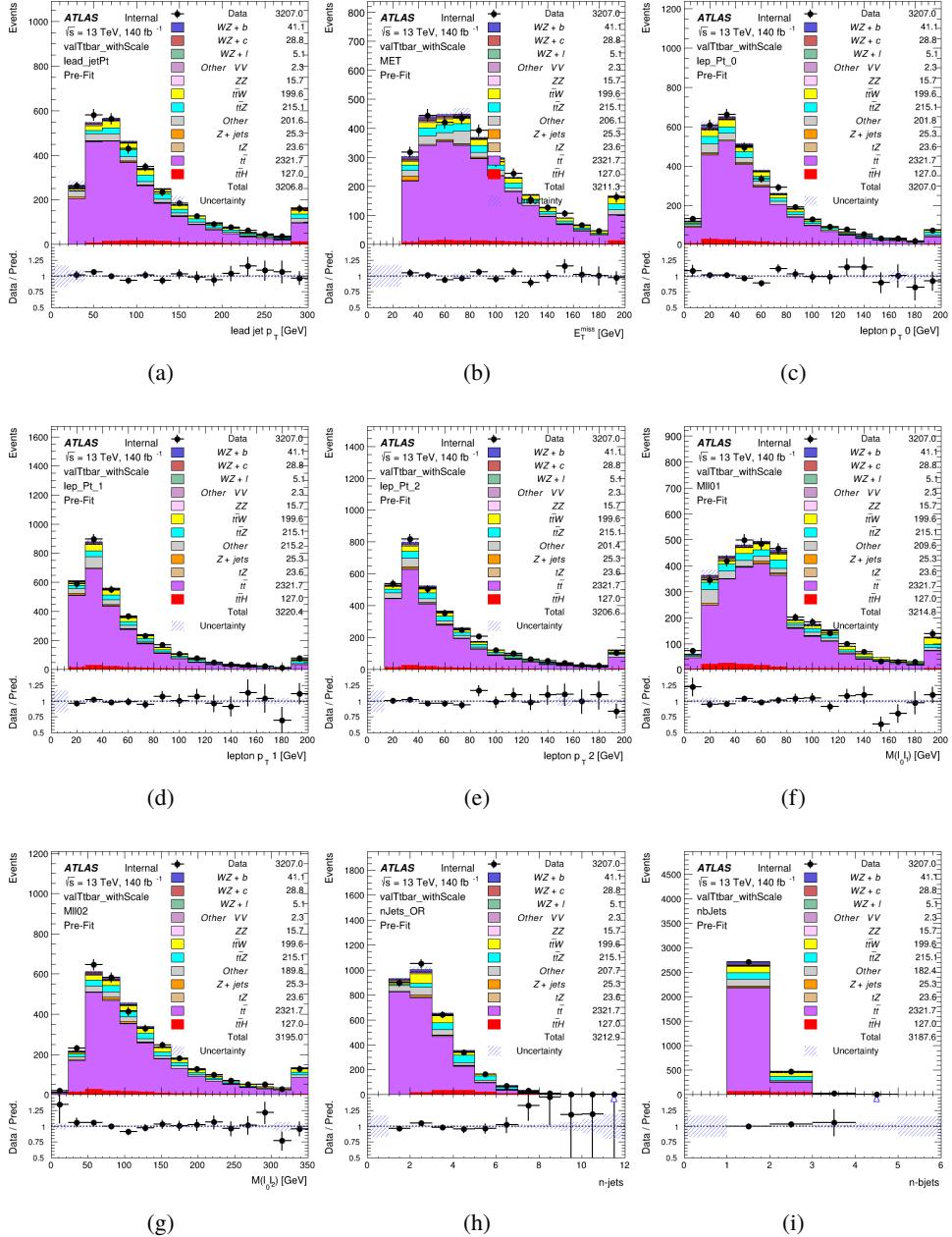


Figure 19: Comparisons between the data and MC distributions in the $t\bar{t}$ validation region after the correction factor has been applied for (a) the p_T of the leading jet, (b) the missing transverse energy, (c) the p_T of lepton 0, (d) p_T of lepton 1, (e) p_T of lepton 2, (f) the invariant mass 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.

325 The modeling is further validated by looking at the yield in the $t\bar{t}$ VR for each DL1r WP, giving
 326 a clearer correspondence to the signal regions used in the fit. Each region shown in Figure 23
 327 requires one or more jets pass the listed WP, with no jets passing the next highest WP.

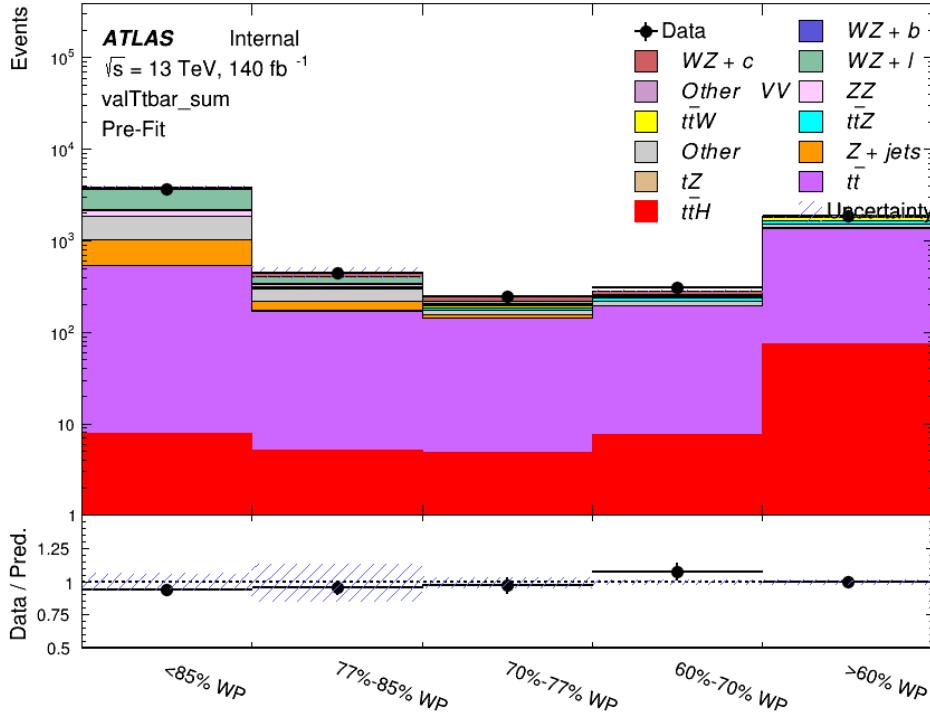


Figure 20: Data and MC comparisons for each DL1r WP for both 1-jet and 2-jet regions, after the $t\bar{t}$ VR selection and correction factor have been applied

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

330 5.3.2 Z+jets Validation

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

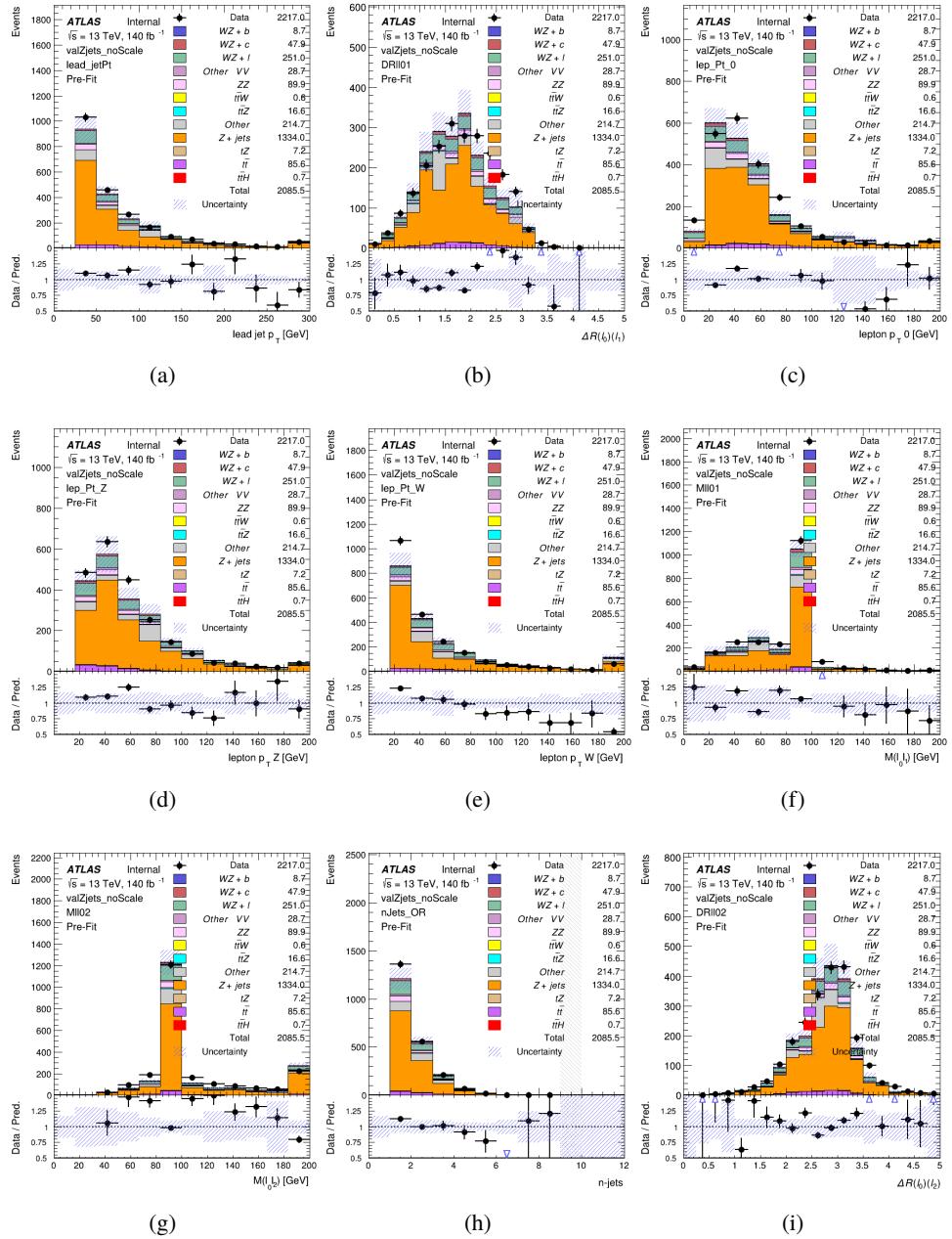


Figure 21: Comparisons between the data and MC distributions in the $Z+jets$ validation region for (a) the p_T of the leading jet, (b) ΔR between leptons 0 and 1, (c) the p_T of lepton 0, (d) p_T of SS lepton from the Z candidate, (e) p_T of the SS lepton from the W candidate, (f) the invariant mass of leptons 0 and 1, (g) the invariant mass of leptons 0 and 2, (h) the number of jets, (i) ΔR between leptons 0 and 2. Includes only statistical uncertainties

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

³³⁹ of the p_T spectrum of the lepton from the W is found to differ. To account for this discrepancy,
³⁴⁰ a variable correction factor is applied to $Z+jets$. χ^2 minimization of the lepton 2 p_T spectrum is
³⁴¹ performed to derive a correction factor of $1.53 - 6.6 * 10^{-6}(lep_Pt_W)$. Kinematic plots of the
³⁴² $Z + jets$ validation region after this correction factor has been applied are shown in Figure 22.

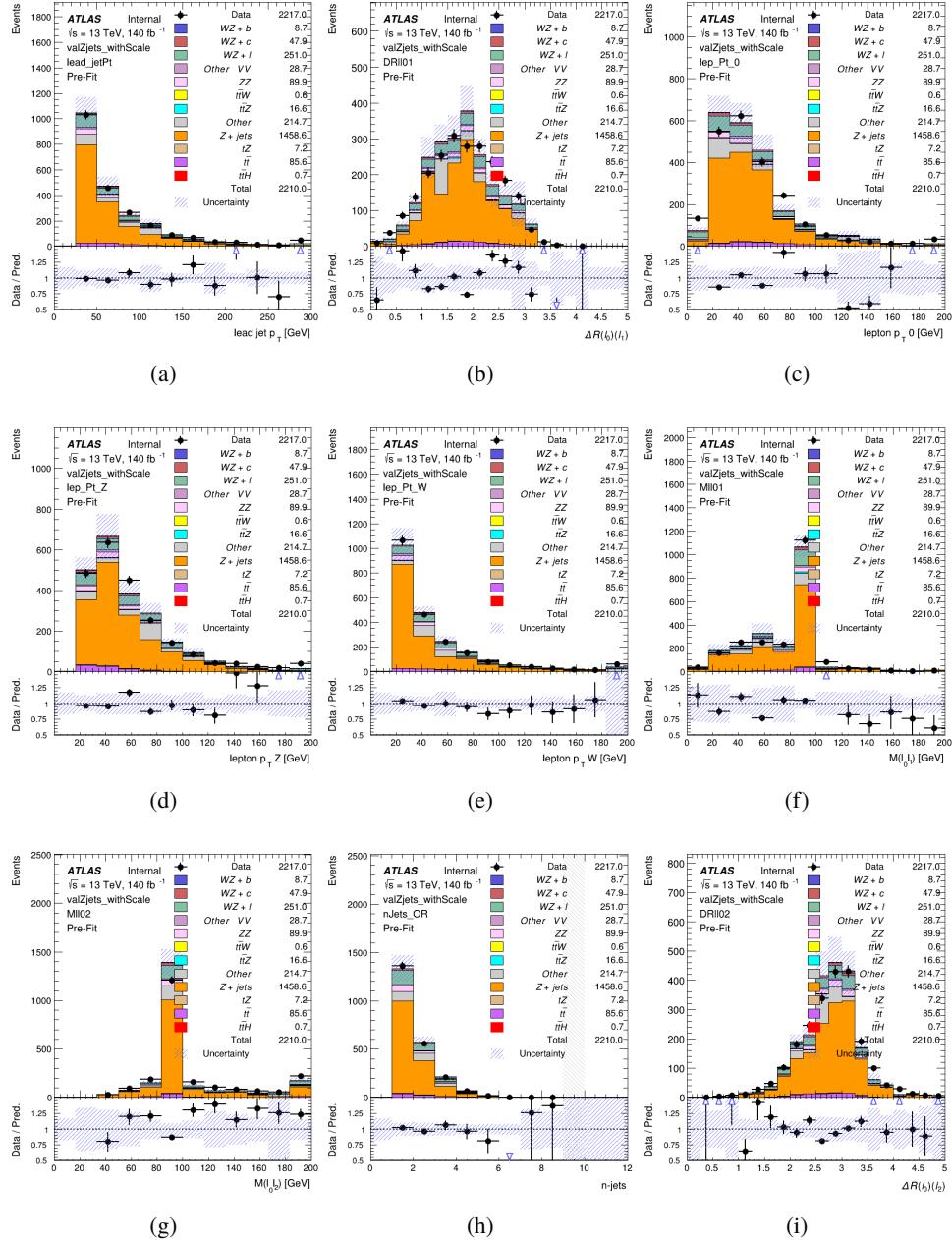


Figure 22: Comparisons between the data and MC distributions in the Z+jets validation region after the correction factor has been applied for (a) the p_T of the leading jet, (b) ΔR between leptons 0 and 1, (c) the p_T of lepton 0, (d) p_T of SS lepton from the Z candidate, (e) p_T of the SS lepton from 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) ΔR between leptons 0 and 2

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

344 giving a clearer correspondence to the signal regions used in the fit. Each region shown in Figure
 345 23 requires one or more jets pass the listed WP, with no jets passing the next highest WP.

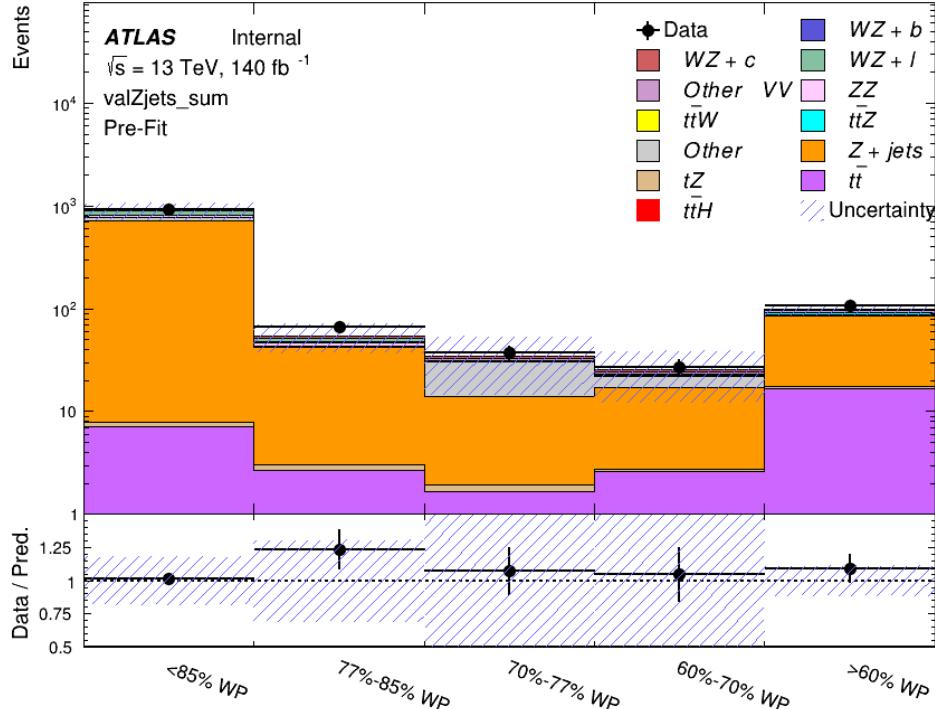


Figure 23: Data and MC comparisons for each DL1r WP for both 1-jet and 2-jet regions, after the Z+jets VR selection and correction factor have been applied

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

349 6 tZ Separation Multivariate Analysis

350 Because tZ produces a final state identical to signal, it represents a predominant background in
 351 the most signal enriched regions. That is, the region with one jet passing the 60% DL1r WP.
 352 Therefore, a boosted decision tree (BDT) algorithm is trained using TMVA [15] to separate WZ
 353 + heavy flavor from tZ.

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

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

359 **6.1 Top Mass Reconstruction**

360 The reconstruction of the top mass follows the procedure described in detail in section 6.1 of
 361 [16]. The mass of the top quark candidate is reconstructed from the jet, the lepton not included
 362 in the Z-candidate, and a reconstructed neutrino. In the case that there is one jet in the event,
 363 there is only possible b-jet candidate. For events with two jets, the jet with the highest DL1r
 364 score is used.

365 The neutrino from the W decay is expected to be the only source of E_T^{miss} . Therefore, the E_T
 366 and ϕ of the neutrino are taken from the E_T^{miss} measurement. This leaves the z-component of
 367 the neutrino momentum, $p_{\nu z}$ as the only unknown.

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

$$370 \quad (p_l + p_\nu)^2 = m_W^2$$

371 Expanding this out into components, this equation gives:

$$372 \quad \sqrt{p_{T\nu}^2 + p_{z\nu}^2} E_l = \frac{m_w^2 - m_l^2}{2} + p_{T\nu}(p_{lx}\cos\phi_\nu + p_{ly}\sin\phi_\nu) + p_{lz}p_{\nu z}$$

373 This equation gives two solutions for $p_{\nu z}$. For cases where only one of these solutions is real,
 374 that is taken as the value of $p_{\nu z}$. For instances with two real solutions, the one which is shown
 375 to be correct in the largest fraction of simulations is taken. For cases when no real solution is
 376 found, often because of detector effects, the value of E_T^{miss} is varied in decreasing increments of
 377 100 MeV until a real solution is found.

378 The reconstructed top mass distribution for tZ and WZ + b can be seen in figure 24.

379 **6.2 tZ BDT**

380 A Boosted Decision Tree (BDT), specifically XGBoost [17], is used to provide separation between
 381 tZ and WZ+b. The following kinematic variables are used as inputs:

- 382 • The invariant mass of the reconstructed top candidate
- 383 • p_T of each of the leptons, jet
- 384 • The invariant mass of each combination of lepton pairs, $M(l\bar{l})$
- 385 • E_T^{miss}

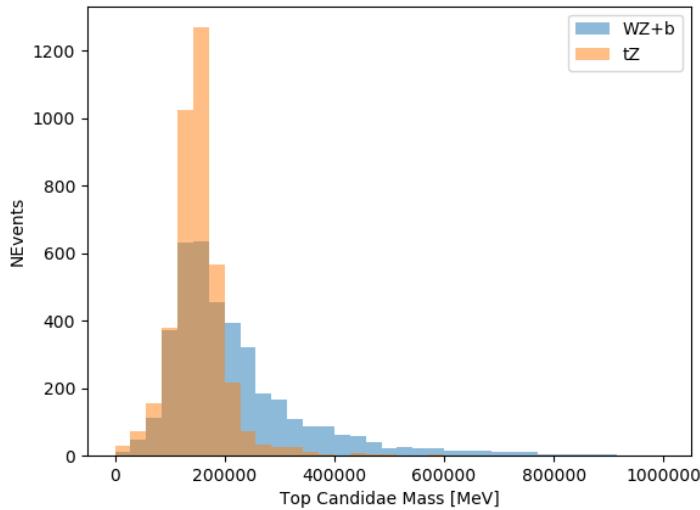


Figure 24: Reconstructed top mass distributions for tZ and WZ + b, measured in MeV.

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

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

391 The distributions of a few of these features for both signal and background is shown in figure
 392 25.

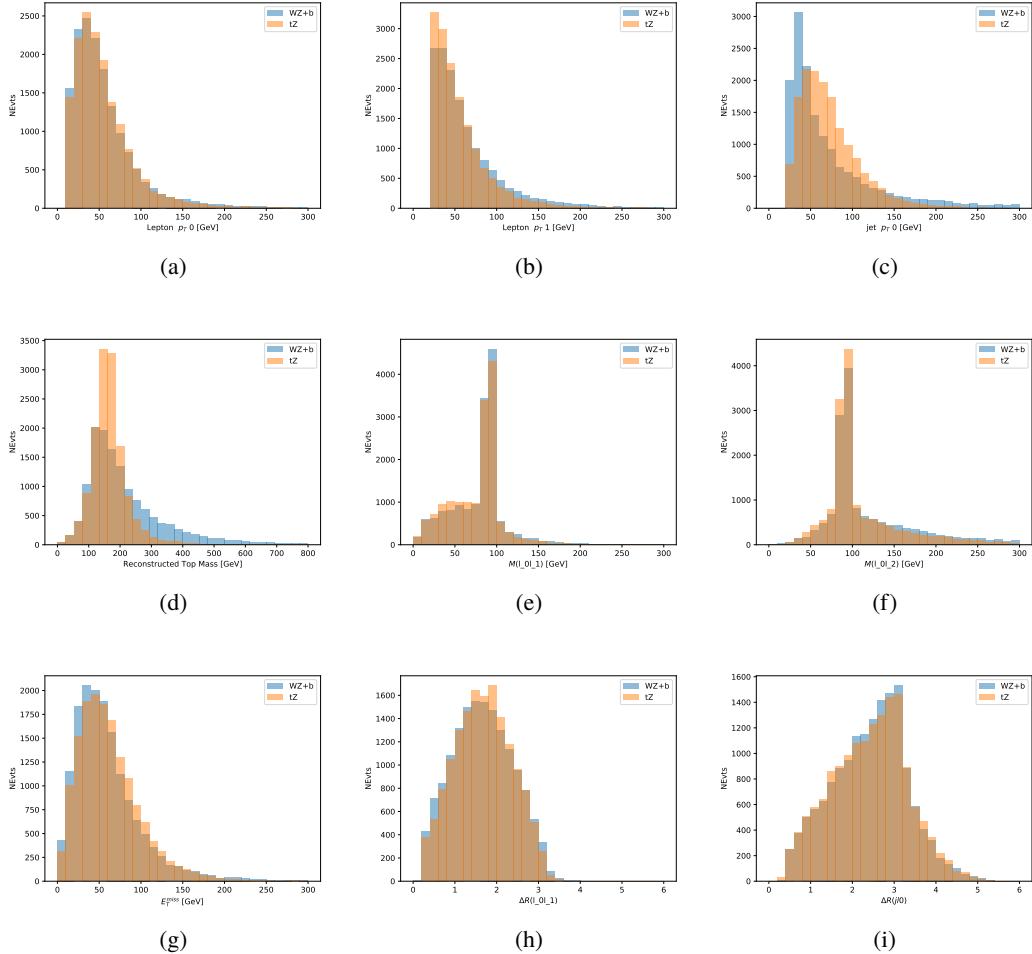


Figure 25: Distribution of input features of the BDT for signal (WZ) and background (tZ). Both are scaled to an equal number of events. (a), (b) and (c) show the p_T of lepton 0, lepton 1, and the jet, (d) show the reconstructed top mass, (e) and (f) show the invariant mass of leptons 0 and 1, leptons 0 and 2. (g) shows the E_T^{miss} of each event. (h) and (i) show the ΔR between lepton 0 and lepton 1, and the jet.

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

398 The results of the BDT training are shown in figure 26. The output scores for both signal and
 399 background events is shown on the left. The right shows the receiving operating characteristic
 400 (ROC) curve that results from the MVA. The ROC curve represents the background rejection
 401 as a function of signal efficiency, where each point on the curve represents a different response

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

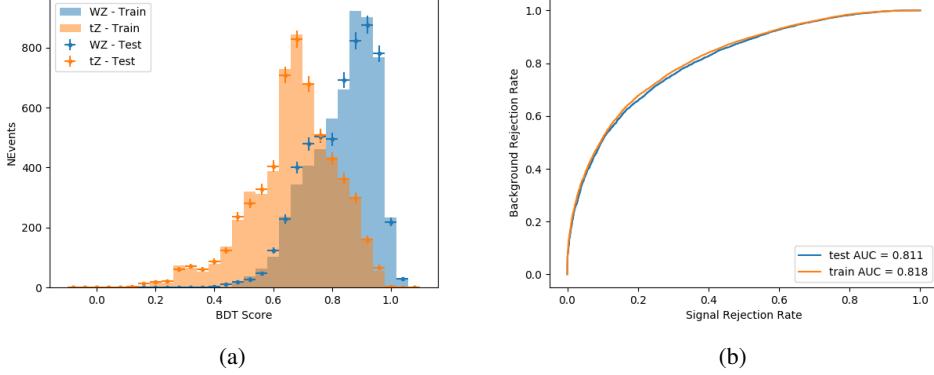


Figure 26: Distribution of the BDT response for signal and background events on the left, the ROC curve for the BDT on the right.

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

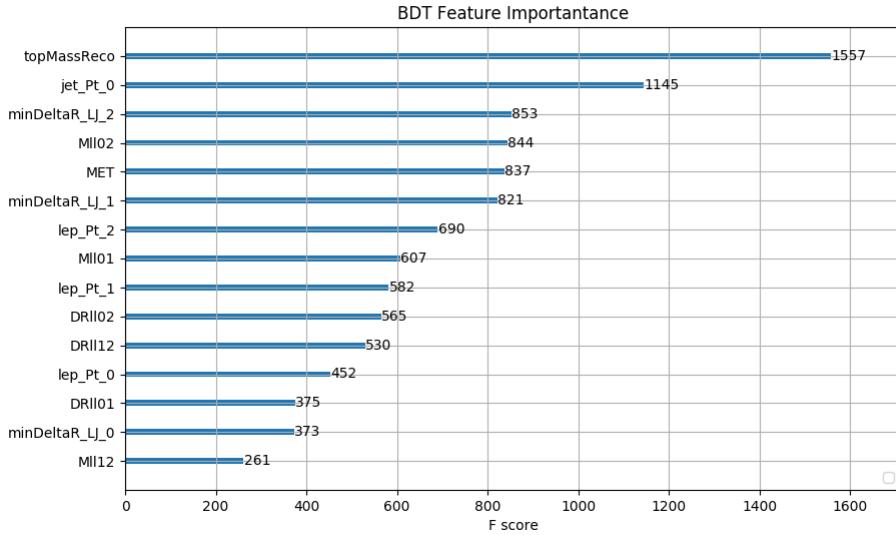


Figure 27: Relative importance of each input feature in the model.

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

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

412 7 Systematic Uncertainties

413 The systematic uncertainties that are considered are summarized in Table 8. These are imple-
 414 mented in the fit either as a normalization factors or as a shape variation or both in the signal
 415 and background estimations. The numerical impact of each of these uncertainties is outlined in
 416 Section 8.

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

Systematic uncertainty	Components
Luminosity	1
Pileup reweighting	1
Physics Objects	
Electron	6
Muon	15
Jet energy scale	28
Jet energy resolution	8
Jet vertex fraction	1
Jet flavor tagging	131
E_T^{miss}	3
Total (Experimental)	194
Signal Modeling	
Shape modelling	3
Renormalization and factorization scales	5
nJet Migration	4
Background Modeling	
Cross section	15
Renormalization and factorization scales	12
Total (Signal and background modeling)	35
Total (Overall)	229

417 The uncertainty in the combined integrated luminosity is derived from a calibration of the
 418 luminosity scale performed for 13 TeV proton-proton collisions [18], [19].

419 The experimental uncertainties are related to the reconstruction and identification of light leptons
 420 and b-tagging of jets, and to the reconstruction of E_T^{miss} . The TOTAL electron ID correlation

421 model is used, corresponding to 1 electron ID systematic. Electron ID is found to be a subleading
422 systematic that is unconstrained by the fit, making it an appropriate choice for this analysis.

423 The sources which contribute to the uncertainty in the jet energy scale (JES) [20] are decom-
424 posed into uncorrelated components and treated as independent sources in the analysis. The
425 CategoryReduction model is used to account for JES uncertainties, which decomposes the uncer-
426 tainties into 30 nuisance parameters included in the fit. The SimpleJER model is used to account
427 for jet energy resolution (JER) uncertainties, and 8 JER uncertainty components unclued as
428 NPs in the fit.

429 The uncertainties in the b-tagging efficiencies measured in dedicated calibration analyses [21] are
430 also decomposed into uncorrelated components. The large number of components for b-tagging
431 is due to the calibration of the distribution of the MVA discriminant.

432 The full list of systematic uncertainties considered in the analysis is summarized in Tables 9, 10
433 and 11.

434

Experimental Systematics on Leptons and E_T^{miss}			
Type	Description	Systematics Name	Application
Trigger			
Scale Factors	Trigger Efficiency	lepSFTrigTight_MU(EL)_SF_Trigger_STAT(SYST)	Event Weight
Muons			
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
Electrons			
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
E_T^{miss}			
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^{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
	η 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, η 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 η	DL1r_Continuous_EventWeight_B0-29
	DL1r b-tagger efficiency on c originated jets in bins of η	DL1r_Continuous_EventWeight_C0-19
	DL1r b-tagger efficiency on light flavoured originated jets in bins of η 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.

- 435 Theoretical uncertainties applied to MC predictions, including cross section, PDF, and scale
 436 uncertainties are taken from theory calculations, with the exception of non-prompt and diboson
 437 backgrounds. The cross-section uncertainty on tZ is taken from [22]. Derivation of the non-
 438 prompt background uncertainties, Z+jets and tt}, are explained in detail in Section 5.3.
 439 The other VV + heavy flavor processes (namely VV+b and VV+charm, which primarily consist
 440 of ZZ events) are also poorly understood, because these processes involve the same physics as
 441 WZ + heavy flavor, and have also not been measured. Therefore, a conservative 50% uncertainty
 442 is applied to those samples. While this uncertainty is large, it is found to have little impact on
 443 the significance of the final result.
 444 The theory uncertainties applied to the predominate background estimates are summarized in
 445 Table 12.

Process	X-section [%]
tZ	X-sec: ± 15.2 QCD Scale: $^{+5.2}_{-1.3}$ PDF($+\alpha_S$): ± 1.2
t <bar>t} H (aMC@NLO+Pythia8)</bar>	QCD Scale: $^{+5.8}_{-9.2}$ PDF($+\alpha_S$): ± 3.6
t <bar>t} Z (aMC@NLO+Pythia8)</bar>	QCD Scale: $^{+9.6}_{-11.3}$ PDF($+\alpha_S$): ± 4
t <bar>t} W (aMC@NLO+Pythia8)</bar>	QCD Scale: $^{+12.9}_{-11.5}$ PDF($+\alpha_S$): ± 3.4
VV + b/charm (Sherpa 2.2.1)	± 50
VV + light (Sherpa 2.2.1)	± 6
t <bar>t}</bar>	± 10
Z + jets	± 20
Others	± 50

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

446 Additional signal uncertainties are estimated by comparing estimates from the nominal Sherpa
 447 WZ samples with alternate WZ samples generated with Powheg+PYTHIA8 (DSID 361601).
 448 The shape of the templates used in the fit are compared between these two samples for WZ + b,
 449 WZ + charm and WZ + light, as shown in Figures 28 and 29. Each of these plots are normalized
 450 to unity in order to capture differences in shape.

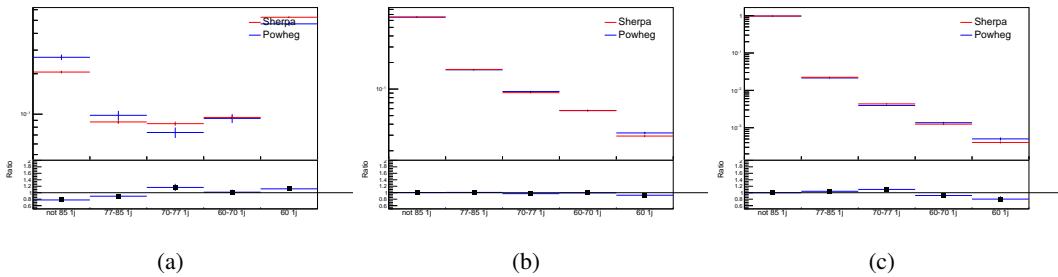


Figure 28: Comparison between Sherpa and Powheg predictions of the distribution of (a) WZ + b, (b) WZ + charm, and (c) WZ + light among the various b-tag WPs used in the 1-jet fit.

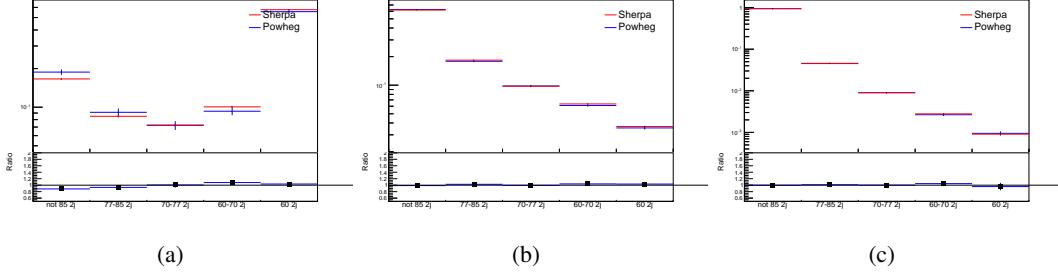


Figure 29: Comparison between Sherpa and Powheg predictions of the distribution of (a) WZ + b, (b) WZ + charm, and (c) WZ + light among the various b-tag WPs used in the 2-jet fit.

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

454 A similar approach is taken to account for uncertainties in migrations between the number of
 455 reco and truth jets. The fraction of events with 1 truth jet which fall into the 1 jet bin versus the
 456 2 jet bin at reco level is compared for Sherpa and Powheg. The same is done for events with 2
 457 truth jets. This comparison is shown in figure 30.

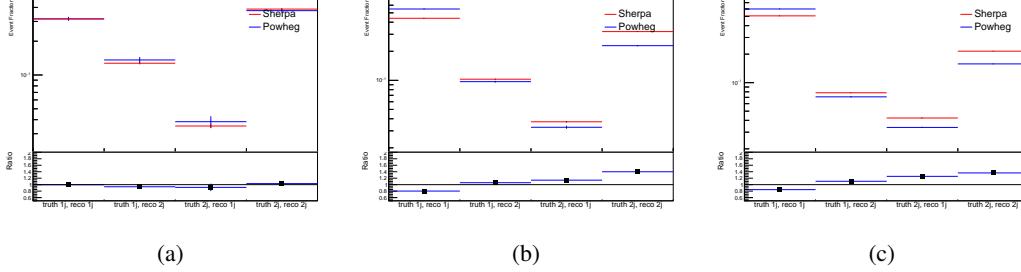


Figure 30: Comparison between Sherpa and Powheg predictions for truth jet migrations between the 1 and 2 jet reco bins for (a) WZ + b, (b) WZ + charm, and (c) WZ + light

458 A systematic is included where events are shifted between the 1-jet and 2-jet regions based on
 459 the differences between these two shapes. This is done independently for each of the WZ + b,
 460 WZ + charm, and WZ + light templates.

461 Additional systematics are included to account for the uncertainty in the contamination of 0 jet
 462 and 3 or more jet events (at truth level) in the 1 and 2 reco jet bins. Because these events fall
 463 outside the scope of this measurement, these events are included as a background. As such, a
 464 normalization, rather than a shape, uncertainty is applied for this background.

465 The number of WZ events with 0-jets and ≥ 3 -jets in the reconstructed 1-jet and 2-jet regions
 466 are compared for Sherpa and Powheg, as seen in figure 31. These differences are taken as separate
 467 normalization systematics on the yield of WZ+0-jet and WZ+ ≥ 3 -jet events.

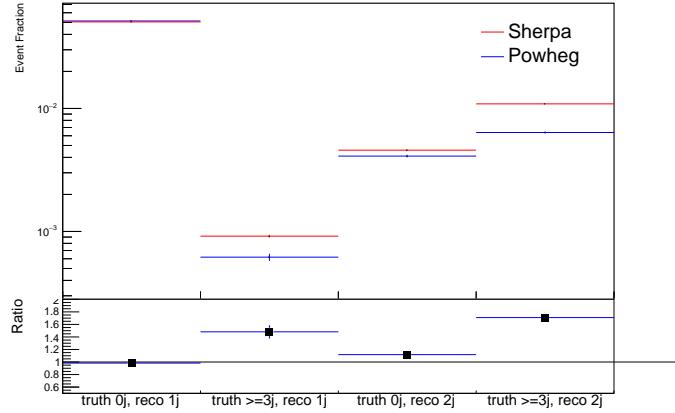


Figure 31: Comparison between Sherpa and Powheg predictions for 0 and ≥ 3 truth jet contributions in the 1 and 2 jet reco bins

468 8 Results

469 A separate maximum-likelihood fit is performed over the various fit regions in order to extract
 470 the best-fit value of the WZ + b-jet and WZ + charm jet contributions for events with both 1 and 2
 471 associated jets. The WZ + b, WZ + charm and WZ + light contributions, are separated into 1-jet
 472 and 2-jet samples at truth level. The signal templates are allowed to float, with the remaining
 473 background contributions are held fixed. The result of the fit is used to extract the cross-section
 474 of WZ + heavy-flavor production.

475 A maximum likelihood fit to data is performed simultaneously in the regions described in Section
 476 5, summarized in figure 32. The parameters $\mu_{WZ+b-1-jet}$, $\mu_{WZ+charm1-jet}$, $\mu_{WZ+light-1-jet}$,
 477 $\mu_{WZ+b-2-jet}$, $\mu_{WZ+charm2-jet}$, $\mu_{WZ+light-2-jet}$, where $\mu = \sigma_{\text{observed}}/\sigma_{\text{SM}}$, are extracted from
 478 the fit. A simultaneous fit is performed over all 1-jet and 2-jet regions.

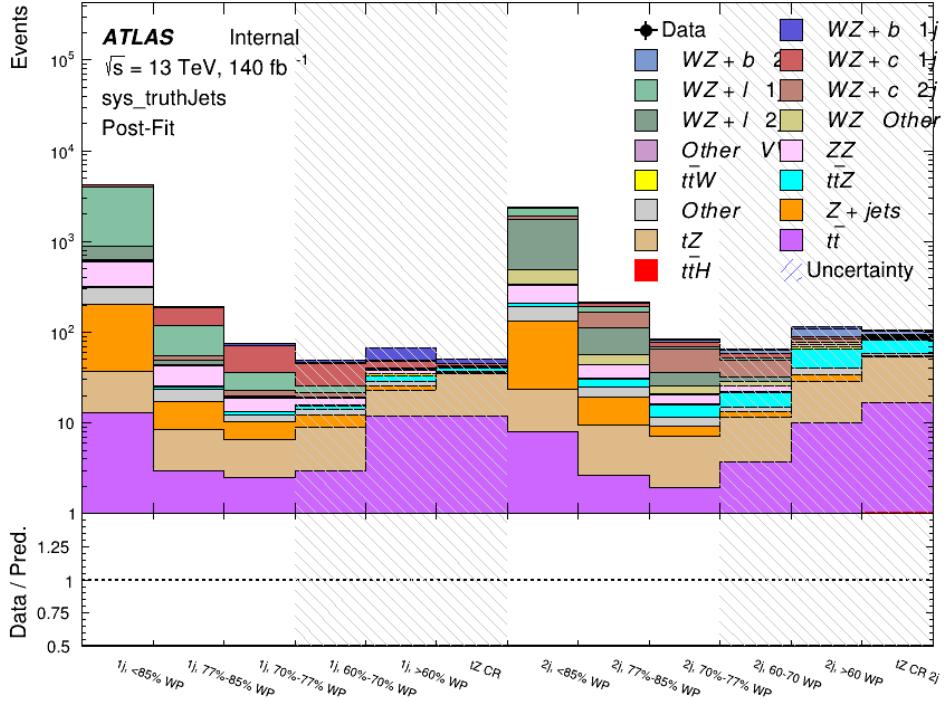


Figure 32: Post-fit summary of the fit regions.

479 As described in Section 7, there are 229 systematic uncertainties that are considered as NPs in
 480 the fit. These NPs are constrained by Gaussian or log-normal probability density functions. The
 481 latter are used for normalisation factors to ensure that they are always positive. The expected
 482 number of signal and background events are functions of the likelihood. The prior for each NP is
 483 added as a penalty term, decreasing the likelihood as it is shifted away from its nominal value.

484 The Asimov fit for 1-jet events gives an expected μ value of $1.00^{+0.47}_{-0.43}(\text{stat})^{+0.30}_{-0.27}(\text{sys})$ for WZ
 485 + b. The fitted cross-section modifiers for $WZ + \text{charm}$ and $WZ + \text{light}$ are $1.00 \pm 0.17 \pm 0.17$
 486 and $1.00 \pm 0.06 \pm 0.14$, respectively.

487 The expected cross-section of $WZ+b$ with 1-jet is $1.74^{+0.82}_{-0.75}(\text{stat})^{+0.53}_{-0.48}(\text{sys})$ fb, and $14.6 \pm$
 488 $2.5(\text{stat}) \pm 2.3(\text{sys})$ fb for $WZ + \text{charm}$, with a correlation of -0.22 between them. An expected
 489 significance of 2.0 is observed for $WZ + b$ in this region.

490 For 2-jet events, the fit gives an expected μ value of $1.00^{+0.53}_{-0.51}(\text{stat})^{+0.39}_{-0.34}(\text{sys})$ for $WZ + b$.
 491 The fitted cross-section modifiers for $WZ + \text{charm}$ and $WZ + \text{light}$ are $1.00 \pm 0.25 \pm 0.21$ and
 492 $1.00 \pm 0.06 \pm 0.16$, respectively.

493 The expected $WZ + b$ cross-section in the 2-jet region is $2.5^{+1.3}_{-1.3}(\text{stat})^{+0.95}_{-0.83}(\text{sys})$ fb with an
 494 expected significance of 1.7σ . The 2-jet expected cross-section of $WZ + \text{charm}$ is $12.7 \pm$
 495 $3.2(\text{stat}) \pm 2.7(\text{sys})$ fb, and the correlation between $WZ + \text{charm}$ and $WZ + b$ is -0.26.

⁴⁹⁶ The pre-fit yields in each of the 1-jet regions used in the fit are shown in Table 8.

⁴⁹⁷

Sample	1j, <85% WP	1j, 77%-85% WP	1j, 70%-77% WP	1j, 60%-70% WP	1j, >60% WP	tZ CR
WZ + b - 1j	8.1 ± 1.6	4.7 ± 0.5	4.6 ± 0.4	5.1 ± 0.4	18.2 ± 2.4	4.8 ± 0.6
WZ + c - 1j	260 ± 22	81 ± 6	43.1 ± 3.6	25.8 ± 2.6	9.4 ± 1.8	2.9 ± 0.6
WZ + l - 1j	3090 ± 250	91 ± 13	17 ± 3	4.9 ± 1.6	1.3 ± 0.4	0.2 ± 0.1
WZ + b - 2j	1.10 ± 0.37	0.44 ± 0.11	0.39 ± 0.06	0.62 ± 0.14	2.1 ± 0.5	0.59 ± 0.14
WZ + c - 2j	21 ± 5	5.6 ± 1.2	3.0 ± 0.7	2.0 ± 0.5	0.70 ± 0.20	0.30 ± 0.08
WZ + l - 2j	250 ± 60	5.7 ± 1.6	0.73 ± 0.53	0.31 ± 0.15	0.07 ± 0.06	0.01 ± 0.01
WZ - Other	13 ± 5	1.4 ± 0.4	0.42 ± 0.08	0.2 ± 0.01	0.30 ± 0.05	0.67 ± 0.15
Other VV	6.2 ± 0.6	0.2 ± 0.4	0.2 ± 0.04	0.07 ± 0.1	0.1 ± 0.1	0.1 ± 0.2
ZZ	336 ± 26	17.8 ± 2.1	4.3 ± 0.6	1.7 ± 0.5	0.36 ± 0.08	0.10 ± 0.03
t̄W	1.1 ± 0.2	0.2 ± 0.1	0.3 ± 0.1	0.4 ± 0.1	1.5 ± 0.3	0.7 ± 0.2
t̄Z	6.8 ± 1.2	1.4 ± 0.3	1.0 ± 0.2	1.2 ± 0.2	4.4 ± 0.8	3.2 ± 0.6
Z + jets	169 ± 38	8.9 ± 1.9	3.7 ± 0.8	3.3 ± 0.7	3.2 ± 0.7	0.8 ± 0.17
V + γ	45 ± 28	1.9 ± 2.4	0.1 ± 0.1	0.02 ± 0.01	1.0 ± 0.9	0.02 ± 0.03
tZ	24.3 ± 4.3	5.5 ± 1.1	4.1 ± 0.8	5.9 ± 1.1	10.7 ± 2.0	23 ± 4
tW	1.4 ± 0.8	0.2 ± 0.5	0.0 ± 0.2	0.7 ± 0.6	0.26 ± 0.42	0.39 ± 0.41
WtZ	2.3 ± 1.2	0.6 ± 0.3	0.3 ± 0.21	0.27 ± 0.2	1.1 ± 0.7	0.6 ± 0.5
VVV	12.4 ± 0.5	0.93 ± 0.06	0.35 ± 0.03	0.13 ± 0.02	0.14 ± 0.03	0.02 ± 0.01
VH	40 ± 6	2.6 ± 1.4	0.9 ± 0.8	0.7 ± 0.8	0.5 ± 0.6	0.0 ± 0.0
t̄t	12.1 ± 1.6	2.9 ± 0.6	2.5 ± 0.5	2.8 ± 0.5	11.2 ± 1.4	10.9 ± 1.5
t̄tH	0.24 ± 0.03	0.05 ± 0.01	0.04 ± 0.01	0.06 ± 0.01	0.20 ± 0.03	0.13 ± 0.02
Total	5010 ± 260	227 ± 24	88 ± 12	57 ± 8	76 ± 16	53 ± 8

Table 13: Pre-fit yields in each of the 1-jet regions.

⁴⁹⁸ The post-fit yields in each region are summarized in Figure 8.

⁴⁹⁹

	1j, <85% WP	1j, 77%-85% WP	1j, 70%-77% WP	1j, 60%-70% WP	1j, >60% WP	tZ CR
WZ + b	8.1 ± 4.9	4.7 ± 2.0	4.6 ± 2.0	5.1 ± 2.1	18 ± 10	5.0 ± 2.5
WZ + c	260 ± 60	80 ± 14	43 ± 7	26 ± 5	9.4 ± 2.3	2.9 ± 0.7
WZ + l	3090 ± 130	90 ± 11	17.3 ± 2.8	4.9 ± 1.6	1.3 ± 0.4	0.23 ± 0.13
WZ + b - 2j	1.10 ± 0.37	0.44 ± 0.11	0.39 ± 0.06	0.62 ± 0.14	2.1 ± 0.5	0.59 ± 0.14
WZ + c - 2j	21 ± 5	5.6 ± 1.2	3.0 ± 0.7	2.0 ± 0.5	0.70 ± 0.20	0.30 ± 0.08
WZ + l - 2j	250 ± 60	5.7 ± 1.6	0.73 ± 0.53	0.31 ± 0.15	0.07 ± 0.06	0.01 ± 0.01
WZ - Other	13 ± 5	1.4 ± 0.4	0.42 ± 0.08	0.2 ± 0.01	0.30 ± 0.05	0.67 ± 0.15
Other VV	6.2 ± 0.6	0.92 ± 0.07	0.02 ± 0.01	0.01 ± 0.01	0.01 ± 0.01	0.01 ± 0.01
ZZ	346 ± 57	19 ± 5	4.3 ± 0.8	2.7 ± 0.5	2.4 ± 0.1	2.1 ± 0.6
t̄W	1.09 ± 0.21	0.2 ± 0.1	0.1 ± 0.1	0.4 ± 0.1	1.5 ± 0.3	0.1 ± 0.2
t̄Z	6.8 ± 1.2	1.4 ± 0.3	1.0 ± 0.2	1.2 ± 0.2	4.4 ± 0.7	3.2 ± 0.5
rare Top	0.14 ± 0.04	0.04 ± 0.02	0.04 ± 0.0	0.1 ± 0.03	0.14 ± 0.04	0.15 ± 0.05
t̄WW	0.04 ± 0.03	0.01 ± 0.02	0.01 ± 0.01	0.01 ± 0.01	0.01 ± 0.02	0.01 ± 0.01
Z + jets	169 ± 37	8.9 ± 1.9	3.7 ± 0.8	3.3 ± 0.7	3.2 ± 0.7	0.8 ± 0.2
W + jets	0.01 ± 0.01					
V + γ	46 ± 28	1.9 ± 2.4	0.1 ± 0.1	0.0 ± 0.2	1.0 ± 0.9	0.0 ± 0.0
tZ	24 ± 4	5.5 ± 1.0	4.1 ± 0.8	5.9 ± 1.1	10.7 ± 1.8	23.3 ± 3.7
tW	1.37 ± 0.82	0.18 ± 0.26	0.01 ± 0.12	0.67 ± 0.64	0.26 ± 0.42	0.39 ± 0.41
WtZ	2.3 ± 1.2	0.6 ± 0.3	0.3 ± 0.2	0.3 ± 0.2	1.1 ± 0.6	0.6 ± 0.3
VVV	12.4 ± 0.4	0.9 ± 0.1	0.4 ± 0.1	0.13 ± 0.02	0.14 ± 0.03	0.02 ± 0.01
VH	40 ± 6	2.6 ± 1.4	0.9 ± 0.8	0.7 ± 0.8	0.4 ± 0.6	0.01 ± 0.01
t̄t	12.1 ± 1.6	2.9 ± 0.6	2.5 ± 0.5	2.8 ± 0.5	11.2 ± 1.5	10.9 ± 1.4
t̄tH	0.24 ± 0.03	0.05 ± 0.01	0.04 ± 0.01	0.06 ± 0.01	0.20 ± 0.03	0.13 ± 0.02
Total	5100 ± 110	227 ± 12	87 ± 6	56.7 ± 4.4	76 ± 9	52.5 ± 4.2

Table 14: Post-fit yields in each of the 1-jet regions.

500 The impact of each NP is calculated by performing the fit with the parameter of interest held
 501 fixed, varied from its fitted value by its uncertainty, and calculating $\Delta\mu$ relative to the baseline
 502 fit. The impact of the most significant sources of systematic uncertainties on WZ + b with one
 503 associated jet is summarized in Table 15.

Uncertainty Source	$\Delta\mu$	
WZ + 1-jet light XS	0.13	-0.18
WZ + charm cross-section	-0.10	0.12
Jet Energy Scale	0.1	-0.13
tZ cross-section	-0.10	0.10
WZ 1-jet/2-jet Migration	0.08	-0.07
Jet Energy Resolution	-0.10	0.10
Luminosity	-0.08	0.09
Other Diboson + b cross-section	-0.07	0.07
Flavor tagging	0.05	0.05
t <bar>t} cross-section</bar>	-0.05	0.05
Total Systematic Uncertainty	0.29	0.34

Table 15: Summary of the most significant sources of systematic uncertainty on the measurement of WZ + b with exactly one associated jet.

504 The ranking and impact of those nuisance parameters with the largest contribution to the overall
 505 uncertainty is shown in Figure 33.

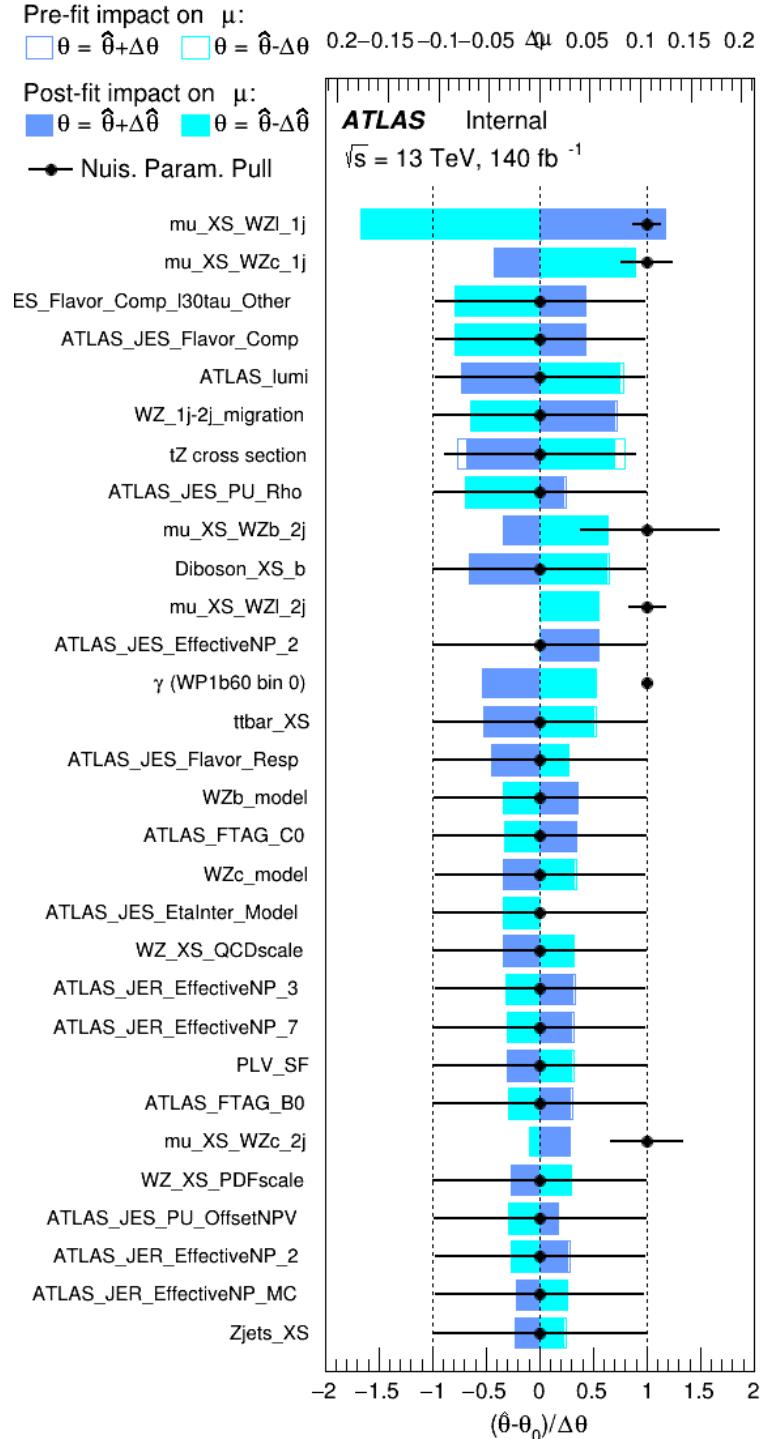


Figure 33: Impact of systematic uncertainties on the signal-strength of $WZ + b$ for events with exactly one jet

506 The large impact of the Jet Energy Scale and Jet Flavor Tagging is unsurprising, as the shape of
 507 the fit regions depends heavily on the modeling of the jets. The other major sources of uncertainty
 508 come from background modelling and cross-section uncertainty.

509 The correlations between these nuisance parameters are summarized in Figure 34.

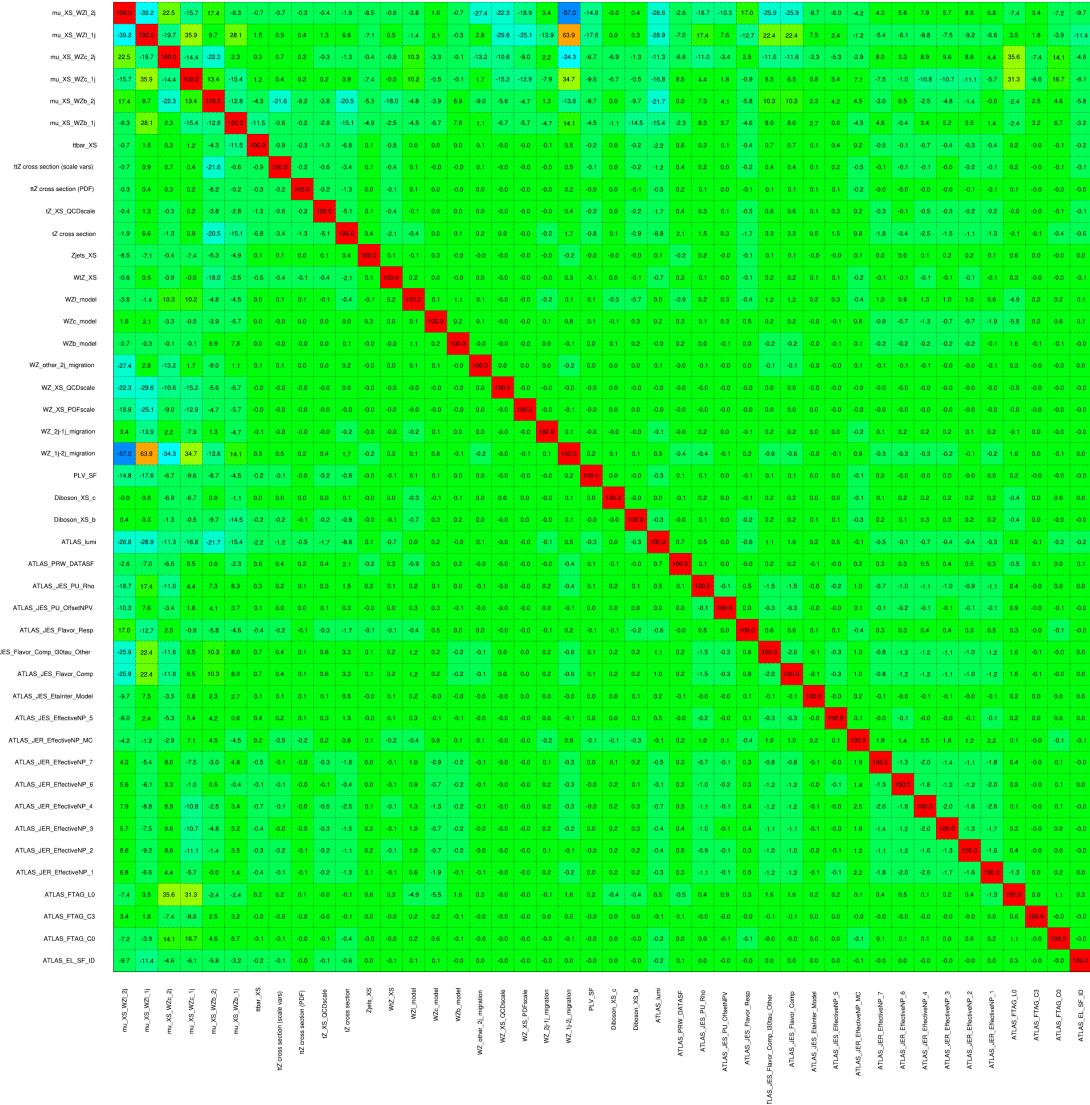


Figure 34: Correlations between nuisance parameters

510 The negative correlations between $\mu_{WZ+charm}$ and μ_{WZ+b} and $\mu_{WZ+light}$ are expected: WZ +
 511 charm is present in both the WZ + b and WZ + light enriched regions, therefore increasing the

	2j, <85% WP	2j, 77%-85% WP	2j, 70%-77% WP	2j, 60%-70% WP	2j, >60% WP	tZ CR 2j
WZ + b	13 ± 6	6.7 ± 2.9	5.8 ± 2.5	8.0 ± 3.5	31 ± 13	14 ± 5
WZ + c	260 ± 60	77 ± 15	41 ± 8	26 ± 5	10.9 ± 2.4	4.8 ± 1.1
WZ + l	1860 ± 90	90 ± 12	17.6 ± 2.8	5.8 ± 1.3	1.4 ± 0.4	0.3 ± 0.2
WZ + b - 1j	3.4 ± 0.6	1.52 ± 0.35	1.58 ± 0.23	1.95 ± 0.39	6.7 ± 1.1	1.9 ± 0.6
WZ + c - 1j	56 ± 14	17.6 ± 4.0	8.6 ± 2.2	6.3 ± 1.8	3.0 ± 0.9	0.7 ± 0.2
WZ + l - 1j	427 ± 120	24 ± 7	4.7 ± 2.3	1.6 ± 0.7	0.3 ± 0.2	0.01 ± 0.01
WZ - Other	129 ± 29	6.1 ± 4.6	1.2 ± 0.3	0.3 ± 0.2	2.9 ± 0.5	3.6 ± 0.6
Other VV	7.6 ± 0.6	0.3 ± 0.3	0.3 ± 0.1	0.1 ± 0.06	0.03 ± 0.02	0.1 ± 0.1
ZZ	145 ± 30	11.3 ± 4.4	2.7 ± 1.6	1.0 ± 0.3	4.0 ± 0.1	2.4 ± 0.1
t̄tW	0.8 ± 0.2	0.4 ± 0.1	0.54 ± 0.12	0.74 ± 0.15	4.3 ± 0.6	3.9 ± 0.6
t̄tZ	14.7 ± 2.2	5.6 ± 0.8	4.5 ± 0.7	6.5 ± 1.0	25.4 ± 3.9	21.9 ± 3.3
rare Top	0.14 ± 0.04	0.07 ± 0.03	0.03 ± 0.02	0.09 ± 0.03	0.4 ± 0.1	0.6 ± 0.1
t̄tWW	0.04 ± 0.03	0.02 ± 0.02	0.01 ± 0.01	0.01 ± 0.01	0.03 ± 0.03	0.01 ± 0.01
Z + jets	110 ± 23	9.6 ± 2.0	2.1 ± 0.5	1.6 ± 0.4	5.1 ± 1.1	1.5 ± 0.3
W + jets	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
V + γ	25 ± 19	0.5 ± 0.2	0.1 ± 0.1	0.13 ± 0.14	0.0 ± 0.02	0.05 ± 0.07
tZ	15.9 ± 2.7	6.9 ± 1.2	5.1 ± 0.9	8.0 ± 1.4	18.7 ± 3.0	36 ± 6
tW	0.1 ± 0.7	0.2 ± 0.3	0.0 ± 0.1	0.0 ± 0.0	0.8 ± 0.6	0.2 ± 0.2
WtZ	4.9 ± 2.5	1.5 ± 0.8	1.1 ± 0.6	1.3 ± 0.7	4.6 ± 2.3	3.3 ± 1.7
VVV	7.4 ± 0.3	1.0 ± 0.1	0.36 ± 0.03	0.19 ± 0.03	0.13 ± 0.03	0.04 ± 0.01
VH	19 ± 4	2.8 ± 1.6	0.7 ± 0.7	0.1 ± 0.2	0.0 ± 0.0	0.0 ± 0.0
t̄t	6.8 ± 1.0	2.4 ± 0.5	1.8 ± 0.4	3.3 ± 0.6	8.4 ± 1.2	13.6 ± 1.7
t̄tH	0.40 ± 0.05	0.19 ± 0.03	0.16 ± 0.02	0.23 ± 0.03	0.94 ± 0.11	1.03 ± 0.11
Total	2580 ± 60	229 ± 11	89 ± 6	69.1 ± 4.1	120 ± 10	108 ± 6

Table 17: Post-fit yields in each of the 2-jet regions.

523 The same set of systematic uncertainties consider for the 1-jet fit are included in the 2-jet fit as
 524 well. The impact of the most significant systematic uncertainties is summarized in Table 18.

Uncertainty Source	$\Delta\mu$	
WZ + c 2-jet cross-section	0.19	0.15
WZ 0-jet, $>=3$ -jet cross-section	0.14	-0.14
Luminosity	-0.11	0.12
tZ cross-section	-0.11	0.11
Jet Energy Scale	-0.11	0.11
ttZ cross-section - QCD scale	-0.08	0.09
WZ + l cross-section	0.08	-0.07
WtZ cross-section	-0.07	0.07
Flavor tagging	0.05	0.05
Other VV + b cross-section	-0.05	0.05
Total	0.32	0.36

Table 18: Summary of the most significant sources of systematic uncertainty on the measurement of WZ + b 2-jet events.

⁵²⁵ The ranking and impact of those nuisance parameters with the largest contribution to the overall
⁵²⁶ uncertainty is shown in Figure 35.

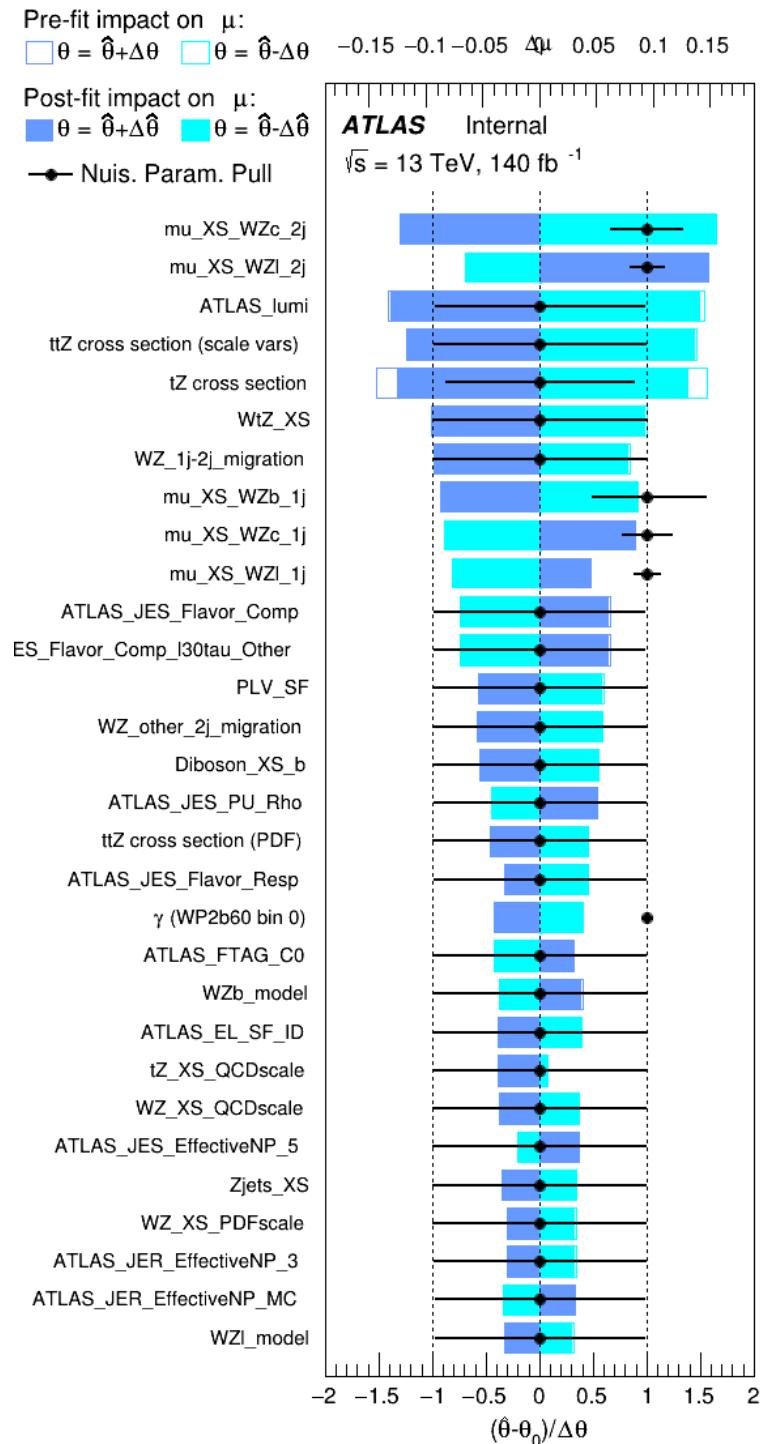


Figure 35: Impact of systematic uncertainties on the signal-strength of $WZ + b$ in 2-jet events.

527 The large impact of the Jet Energy Scale and Jet Flavor Tagging is unsurprising, as the shape of
 528 the fit regions depends heavily on the modeling of the jets. The other major sources of uncertainty
 529 come from background modelling and cross-section uncertainty.

530 The correlations between these nuisance parameters are summarized in Figure 36.

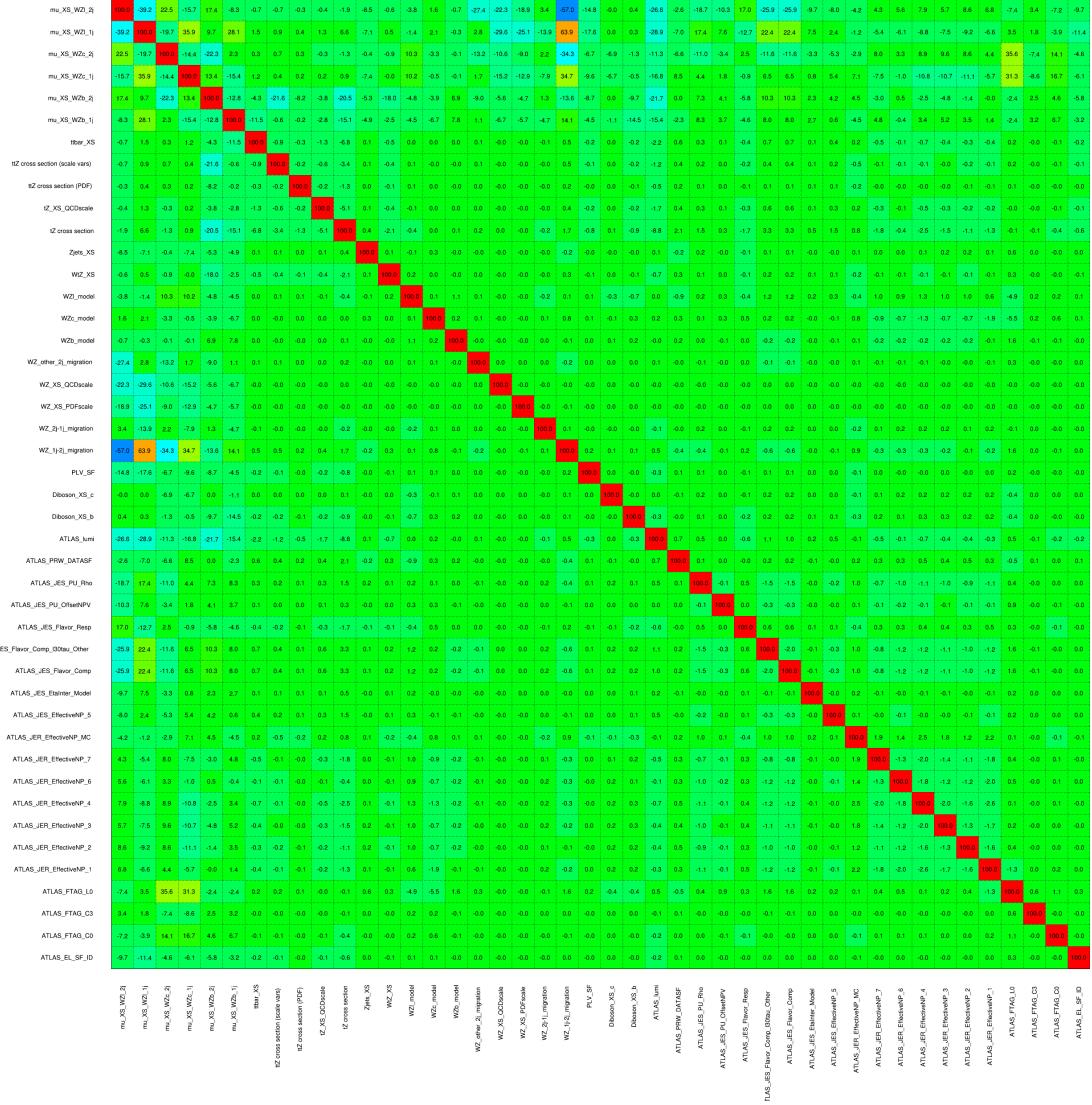


Figure 36: Correlations between nuisance parameters in the 2-jet fit

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

533 9 Conclusion

534 A measurement of $WZ +$ heavy flavor is performed using 140 fb^{-1} of $\sqrt{s} = 13 \text{ TeV}$ proton-
 535 proton collision data collected by the ATLAS detector at the LHC. The expected cross-section
 536 of $WZ+b$ with 1-jet is $1.74_{-0.75}^{+0.82}(\text{stat})_{-0.48}^{+0.53}(\text{sys}) \text{ fb}$, and $14.6 \pm 2.5(\text{stat}) \pm 2.3(\text{sys}) \text{ fb}$ for WZ
 537 + charm, with a correlation of -0.22 between them. An expected significance of 2.0 is observed
 538 for $WZ + b$ in this region.

539 For the 2-jet regions, an expected significance of 1.7 is observed for $WZ + b$, with an ex-
 540 pected cross-section of $2.5_{-1.3}^{+1.3}(\text{stat})_{-0.83}^{+0.95}(\text{sys}) \text{ fb}$. For $WZ +$ charm, a cross-section of
 541 $12.7 \pm 3.2(\text{stat}) \pm 2.7(\text{sys}) \text{ fb}$ is expected for 2-jet events. A correlation of -0.26 is observed
 542 for $WZ+b$ and $WZ +$ charm.

543 **This section will be include final results once unblinded.**

544 **References**

- 545 [1] M. Aaboud et al. ‘Observation of electroweak $W^\pm Z$ boson pair production in association
 546 with two jets in pp collisions at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS detector’. In: *Phys.
 547 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:
 548 [1812.09740 \[hep-ex\]](https://arxiv.org/abs/1812.09740).
- 549 [2] T. Gleisberg et al. ‘Event generation with SHERPA 1.1’. In: *JHEP* 02 (2009), p. 007. doi:
 550 [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).
- 551 [3] R. D. Ball et al. ‘Parton distributions for the LHC Run II’. In: *JHEP* 04 (2015), p. 040.
 552 doi: [10.1007/JHEP04\(2015\)040](https://doi.org/10.1007/JHEP04(2015)040). arXiv: [1410.8849 \[hep-ph\]](https://arxiv.org/abs/1410.8849).
- 553 [4] H.-L. Lai et al. ‘New parton distributions for collider physics’. In: *Phys. Rev. D* 82 (2010),
 554 p. 074024. doi: [10.1103/PhysRevD.82.074024](https://doi.org/10.1103/PhysRevD.82.074024). arXiv: [1007.2241 \[hep-ph\]](https://arxiv.org/abs/1007.2241).
- 555 [5] S. Frixione, G. Ridolfi and P. Nason. ‘A positive-weight next-to-leading-order Monte Carlo
 556 for heavy flavour hadroproduction’. In: *JHEP* 09 (2007), p. 126. doi: [10.1088/1126-6708/2007/09/126](https://doi.org/10.1088/1126-

 557 6708/2007/09/126). arXiv: [0707.3088 \[hep-ph\]](https://arxiv.org/abs/0707.3088).
- 558 [6] E. Re. ‘Single-top Wt-channel production matched with parton showers using the POWHEG
 559 method’. In: *Eur. Phys. J. C* 71 (2011), p. 1547. doi: [10.1140/epjc/s10052-011-1547-z](https://doi.org/10.1140/epjc/s10052-011-

 560 1547-z). arXiv: [1009.2450 \[hep-ph\]](https://arxiv.org/abs/1009.2450).
- 561 [7] ATLAS Collaboration. *Electron efficiency measurements with the ATLAS detector using
 562 the 2015 LHC proton–proton collision data*. ATLAS-CONF-2016-024. 2016. URL: <https://cds.cern.ch/record/2157687>.
- 564 [8] ATLAS Collaboration. ‘Measurement of the muon reconstruction performance of the
 565 ATLAS detector using 2011 and 2012 LHC proton–proton collision data’. In: *Eur. Phys.
 566 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
 567 \[hep-ex\]](https://arxiv.org/abs/1407.3935).

- 568 [9] *Evidence for the associated production of the Higgs boson and a top quark pair with the*
569 *ATLAS detector*. Tech. rep. ATLAS-CONF-2017-077. Geneva: CERN, Nov. 2017. URL:
570 <https://cds.cern.ch/record/2291405>.
- 571 [10] ATLAS Collaboration. *Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s} = 13$ TeV*. ATL-PHYS-PUB-2015-015. 2015. URL:
572 <https://cds.cern.ch/record/2037613>.
- 574 [11] ATLAS Collaboration. *Selection of jets produced in 13 TeV proton–proton collisions with*
575 *the ATLAS detector*. ATLAS-CONF-2015-029. 2015. URL: <https://cds.cern.ch/record/2037702>.
- 577 [12] ATLAS Collaboration. ‘Performance of pile-up mitigation techniques for jets in pp col-
578 lisions at $\sqrt{s} = 8$ TeV using the ATLAS detector’. In: *Eur. Phys. J. C* 76 (2016), p. 581.
579 DOI: [10.1140/epjc/s10052-016-4395-z](https://doi.org/10.1140/epjc/s10052-016-4395-z). arXiv: [1510.03823](https://arxiv.org/abs/1510.03823) [hep-ex].
- 580 [13] ATLAS Collaboration. *Performance of missing transverse momentum reconstruction with*
581 *the ATLAS detector in the first proton–proton collisions at $\sqrt{s} = 13$ TeV*. ATL-PHYS-
582 PUB-2015-027. 2015. URL: <https://cds.cern.ch/record/2037904>.
- 583 [14] 2021. URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/BTagCalibrationRecommendationsRelease21#Tools_for_Flavor_Tagging_Calibra.
- 586 [15] P. S. A. Hoecker. ‘TMVA 4 Toolkit for Multivariate Data Analysis with ROOT’. In:
587 *arXiv:physics/0703039* (2013).
- 588 [16] F. Cardillo et al. ‘Measurement of the fiducial and differential cross-section of a top quark
589 pair in association with a Z boson at 13 TeV with the ATLAS detector’. In: ATL-COM-
590 PHYS-2019-334 (Apr. 2019). URL: <https://cds.cern.ch/record/2672207>.
- 591 [17] T. Chen and C. Guestrin. ‘XGBoost: A Scalable Tree Boosting System’. In: *Proceedings*
592 *of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data*
593 *Mining*. KDD ’16. San Francisco, California, USA: ACM, 2016, pp. 785–794. ISBN: 978-
594 1-4503-4232-2. doi: [10.1145/2939672.2939785](https://doi.acm.org/10.1145/2939672.2939785). URL: <http://doi.acm.org/10.1145/2939672.2939785>.
- 596 [18] ATLAS Collaboration. ‘Luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using
597 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](https://arxiv.org/abs/1101.2185) [hep-ex].
- 599 [19] G. Avoni et al. ‘The new LUCID-2 detector for luminosity measurement and monitoring in
600 ATLAS’. In: *JINST* 13.07 (2018), P07017. DOI: [10.1088/1748-0221/13/07/P07017](https://doi.org/10.1088/1748-0221/13/07/P07017).
- 601 [20] G. Aad et al. ‘Jet energy resolution in proton-proton collisions at $\sqrt{s} = 7$ TeV recorded
602 in 2010 with the ATLAS detector’. In: *The European Physical Journal C* 73.3 (Mar.
603 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:
604 <https://doi.org/10.1140/epjc/s10052-013-2306-0>.
- 605 [21] A. Collaboration. ‘Performance of b -jet identification in the ATLAS experiment’. In:
606 *Journal of Instrumentation* 11.04 (2016), P04008. URL: <http://stacks.iop.org/1748-0221/11/i=04/a=P04008>.

- 608 [22] ‘Observation of the associated production of a top quark and a Z boson at 13 TeV with
609 ATLAS’. In: (July 2020). URL: <https://cds.cern.ch/record/2722504>.

610 **Appendices**

611 **.1 tZ Interference Studies**

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

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

622 MadGraph simulations of 10,000 tZ and 10,000 WZ + b-jet events are produced, along with
623 20,000 events where both are present, in the fiducial region where three leptons and at least one
624 jet are produced.

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

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

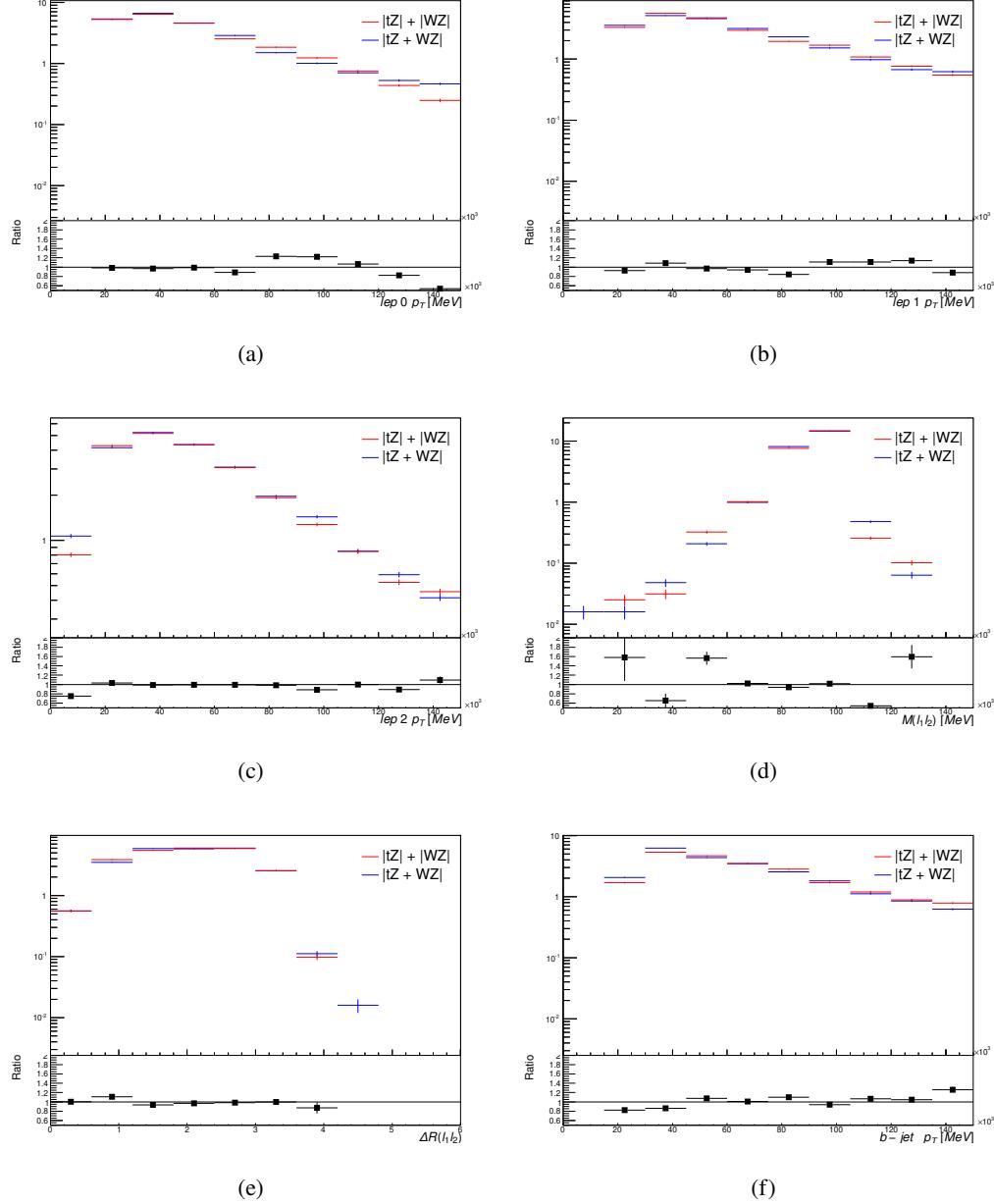


Figure 37: Comparisons between (a) the p_T of the lepton from the W, (b) and (c) show the p_T of the lepton from the Z, (d) the invariant mass of the Z-candidate, (e) ΔR of the leptons from the Z, and (f) the p_T of the b-jet, for WZ and tZ events generated with interference effects (blue) and without interference effects (red).

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

632 significantly impact the results.

633 .2 Alternate tZ Inclusive Fit

634 While tZ is often considered as a distinct process from WZ + b, this could also be considered part
 635 of the signal. Alternate studies are performed where, using the same framework as the nominal
 636 analysis, a measurement of WZ + b is performed that includes tZ as part of WZ+b.

637 Because of this change, the tZ CR is no longer necessary, and only the five pseudo-continuous
 638 b-tag regions are used in the fit. Further, systematics related to the tZ cross-section are removed
 639 from the fit, as they are now encompassed by the normalization measurement of WZ + b. All
 640 other systematic uncertainties are carried over from the nominal analysis.

641 A post-fit summary of the 1-jet regions used in the fit are shown in Figure 38.

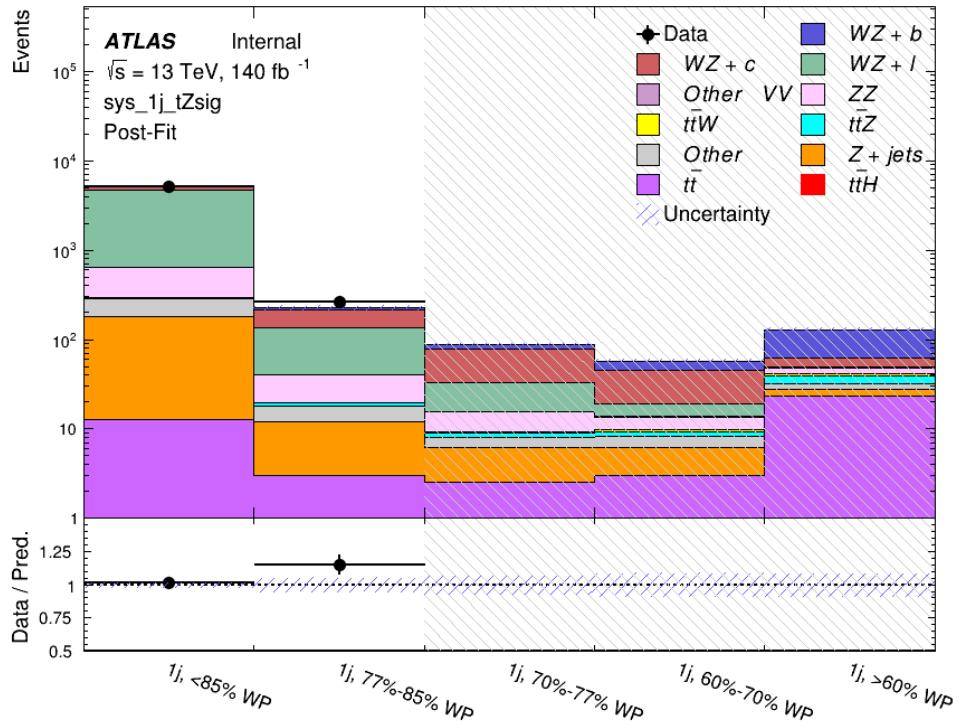


Figure 38: Post-fit summary of the 1-jet fit regions.

642 An expected WZ + b cross-section of $4.1^{+0.78}_{-0.74}(\text{stat})^{+0.53}_{-0.52}(\text{sys}) \text{ fb}$ is extracted from the fit, with
 643 an expected significance of 4.0σ .

644 The impact of the predominate systematics are summarized in Table 19.

Uncertainty Source	$\Delta\mu$	
WZ + light cross-section	0.08	-0.08
Jet Energy Scale	-0.06	0.08
Luminosity	-0.05	0.06
WZ + charm cross-section	-0.04	0.05
Other Diboson + b cross-section	-0.04	0.04
WZ cross-section - QCD scale	-0.04	0.03
t <bar>t</bar>	-0.03	0.03
Jet Energy Resolution	-0.03	0.03
Flavor tagging	-0.03	0.03
Z+jets cross section	-0.02	0.02
Total Systematic Uncertainty	-0.15	0.16

Table 19: Summary of the most significant sources of systematic uncertainty on the measurement of WZ + b with exactly one associated jet.

645 A post-fit summary of the 2-jet regions used in the fit are shown in Figure 39.

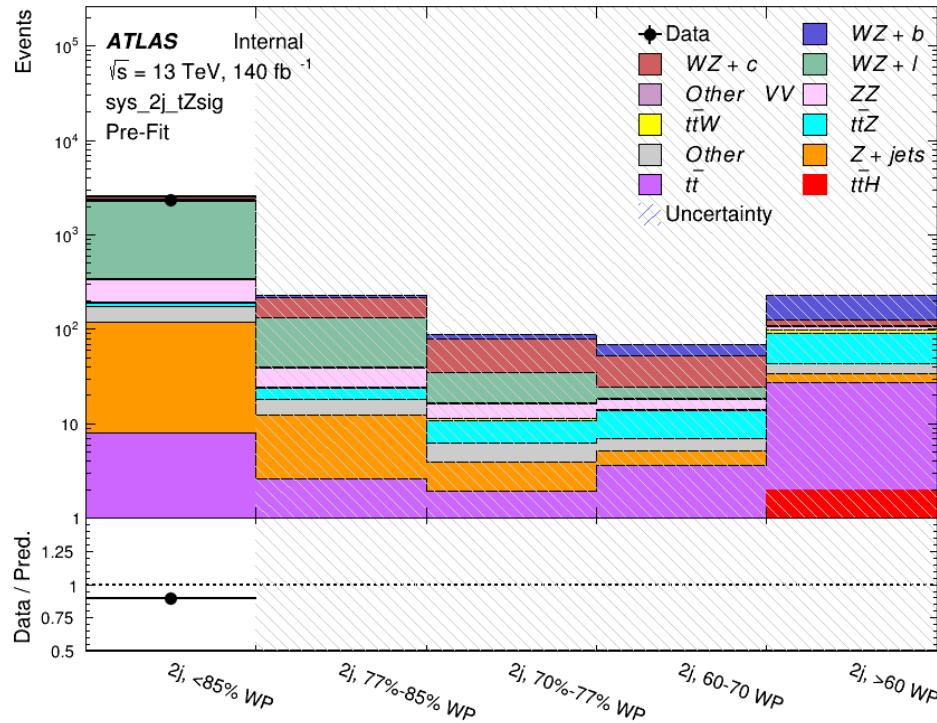


Figure 39: Post-fit summary of the 2-jet fit regions.

646 An expected WZ + b cross-section of $5.9^{+0.9}_{-0.9}(\text{stat})^{+0.9}_{-0.8}(\text{sys}) \text{ fb}$ is extracted from the fit, with an

⁶⁴⁷ expected significance of 5.3σ .

⁶⁴⁸ The impact of the predominate systematics are summarized in Table 20.

Uncertainty Source	$\Delta\mu$	
Luminosity	-0.07	0.07
Jet Energy Scale	0.06	0.05
ttZ cross-section - QCD scale	-0.05	0.05
WZ+l cross-section	0.05	-0.05
WZ+c cross-section	-0.03	0.05
WtZ cross-section	-0.03	0.03
WZ cross-section QCDscale	-0.03	0.03
Diboson cross-section b	-0.03	0.03
WZ cross-section - PDF	-0.03	0.03
Flavor Tagging	0.03	0.02
Total Systematic Uncertainty	-0.14	0.16

Table 20: Summary of the most significant sources of systematic uncertainty on the measurement of WZ + b with exactly one associated jet.

649 .3 DSID list

