



ATLAS Note

GROUP-2017-XX

17th February 2021



1

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

4

The ATLAS Collaboration

5

A measurement of WZ produced with an associated heavy flavor jet is performed using 140
6 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ 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	1 Changes and outstanding items	3
13	1.1 Changelog	3
14	1.1.1 Changes relative to v4	3
15	1.1.2 Changes relative to v3	3
16	1.1.3 Changes relative to v2	4
17	1.1.4 Changes relative to v1	4
18	1.2 Outstanding Items	4
19	2 Executive Summary	5
20	3 Data and Monte Carlo Samples	5
21	3.1 Data Samples	6
22	3.2 Monte Carlo Samples	6
23	4 Object Reconstruction	8
24	4.1 Trigger	8
25	4.2 Light leptons	9
26	4.3 Jets	10
27	4.4 B-tagged Jets	10
28	4.5 Missing transverse energy	11
29	5 Event Selection and Signal Region Definitions	11
30	5.1 Event Preselection	11
31	5.2 Fit Regions	14
32	5.3 Non-Prompt Lepton Estimation	29
33	5.3.1 $t\bar{t}$ Validation	29
34	5.3.2 Z+jets Validation	33
35	6 tZ Interference Studies and Separation Multivariate Analysis	37
36	6.1 Top Mass Reconstruction	38
37	6.2 tZ BDT	38
38	7 Systematic Uncertainties	42
39	8 Results	48
40	8.1 1-jet Fit Results	48
41	8.2 2-jet Fit Results	56
42	8.3 tZ Inclusive Results	62
43	9 Conclusion	63
44	Appendices	64

45 **1 Changes and outstanding items**

46 **1.1 Changelog**

47 This is version 5

48 **1.1.1 Changes relative to v4**

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

54 **1.1.2 Changes relative to v3**

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

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

74 **1.1.3 Changes relative to v2**

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

82 **1.1.4 Changes relative to v1**

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

88 **1.2 Outstanding Items**

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

96 2 Executive Summary

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

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

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

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

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

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

127 3 Data and Monte Carlo Samples

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

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

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

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

3.1 Data Samples

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

- data15_13TeV.periodAllYear_DetStatus-v79-repro20-02_DQDefects-00-02-02
_PHYS_StandardGRL_All_Good_25ns.xml
- data16_13TeV.periodAllYear_DetStatus-v88-pro20-21_DQDefects-00-02-04
_PHYS_StandardGRL_All_Good_25ns.xml
- data17_13TeV.periodAllYear_DetStatus-v97-pro21-13_Unknown_PHYS_StandardGRL
_All_Good_25ns_Triggerno17e33prim.xml
- data18_13TeV.periodAllYear_DetStatus-v102-pro22-04_Unknown_PHYS_StandardGRL
_All_Good_25ns_Triggerno17e33prim.xml

Runs included from the AllYear period containers are included.

3.2 Monte Carlo Samples

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

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

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

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.

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.

4.2 Light leptons

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

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

188 4.3 Jets

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

196 4.4 B-tagged Jets

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

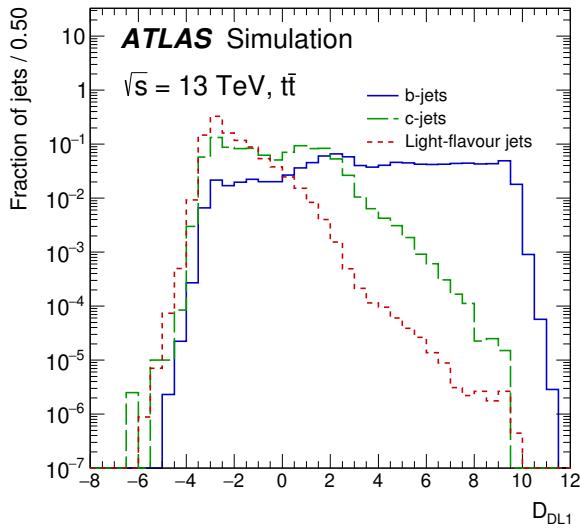


Figure 1: Output distribution of the DL1r algorithm for b-jets, charm jets, and light jets

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

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

212 4.5 Missing transverse energy

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

219 5 Event Selection and Signal Region Definitions

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

223 5.1 Event Preselection

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

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

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

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

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

241 Events are required to have one or two reconstructed jets passing the selection described in
 242 section 4.3. Events with more than two jets are rejected in order to reduce the contribution of
 243 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

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

252 Here Other VV represents diboson processes other than WZ, and consists predominantly of
253 ZZ → ll̄ll events where one of the leptons is not reconstructed.

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

WZ Fit Region - Inclusive

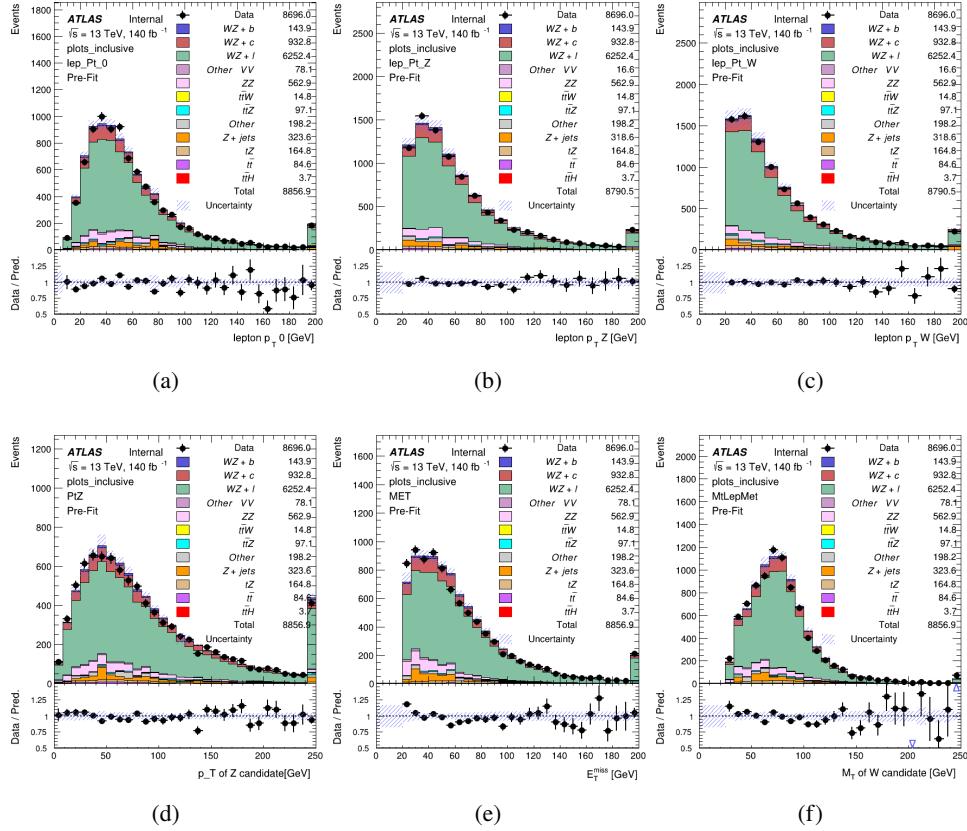


Figure 2: Comparisons between the data and MC distributions in the preselection region for (a) the p_T of the opposite sign lepton, (b) the p_T of the other lepton from the Z candidate, (c) the p_T of the lepton from the W candidate, (d) the p_T of the Z candidate, (e) the E_T^{miss} , and (f) the M_T 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%	$N_{\text{jets}} = 1, n\text{Jets_DL1r_85} = 0$
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	$N_{\text{jets}} = 2, n\text{Jets_DL1r_60} >= 1, tZ \text{ BDT} < 0.725$

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

WZ Fit Region - 1j Inclusive

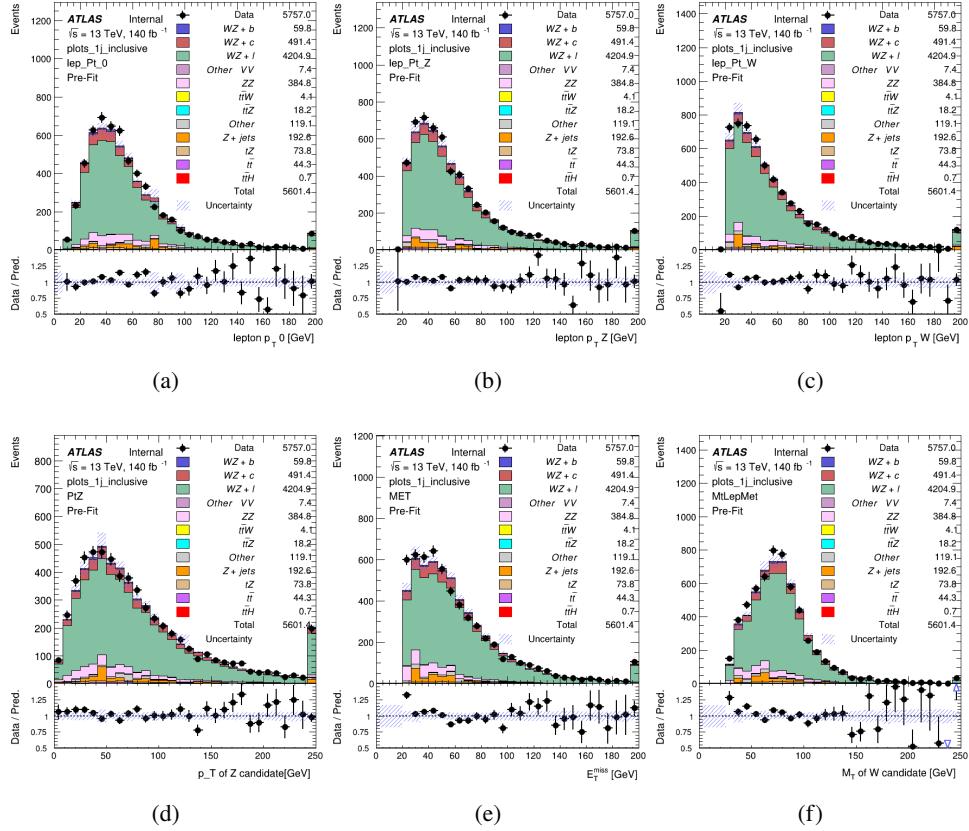


Figure 3: Comparisons between the data and MC distributions in the preselection region for (a) the p_T of the opposite sign lepton, (b) the p_T of the other lepton from the Z candidate, (c) the p_T of the lepton from the W candidate, (d) the p_T of the Z candidate, (e) the E_T^{miss} , and (f) the M_T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

WZ Fit Region - 1j < 85% WP

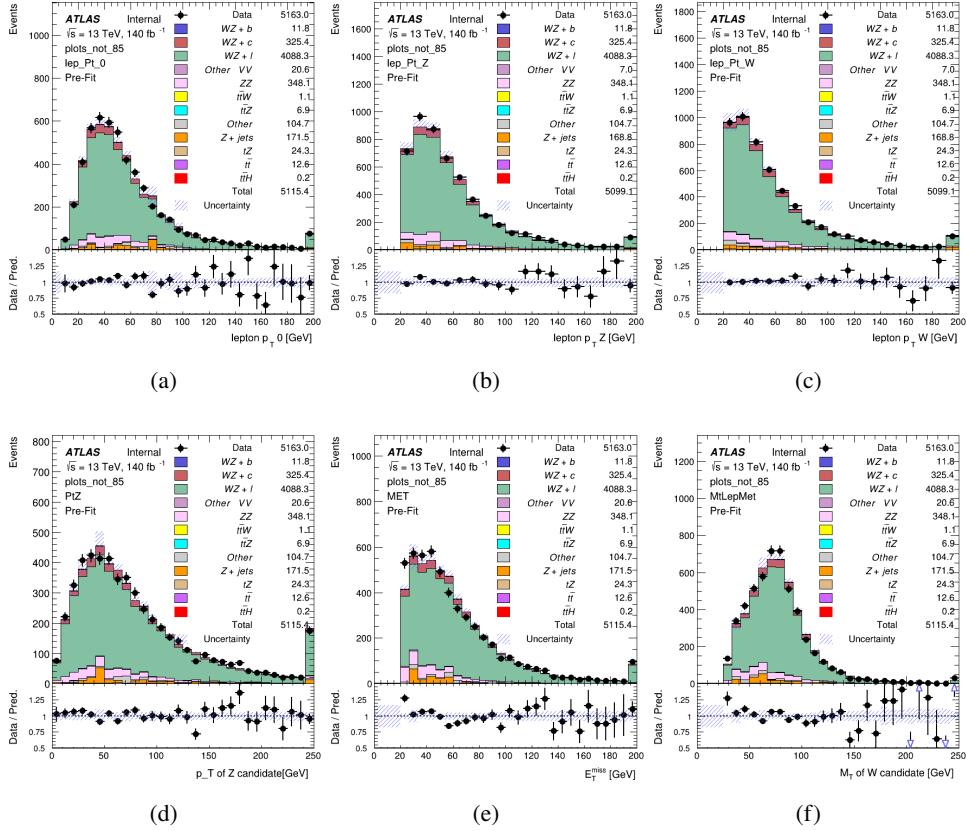


Figure 4: Comparisons between the data and MC distributions in the preselection region for (a) the p_T of the opposite sign lepton, (b) the p_T of the other lepton from the Z candidate, (c) the p_T of the lepton from the W candidate, (d) the p_T of the Z candidate, (e) the E_T^{miss} , and (f) the m_T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

WZ Fit Region - 1j 77-85% WP

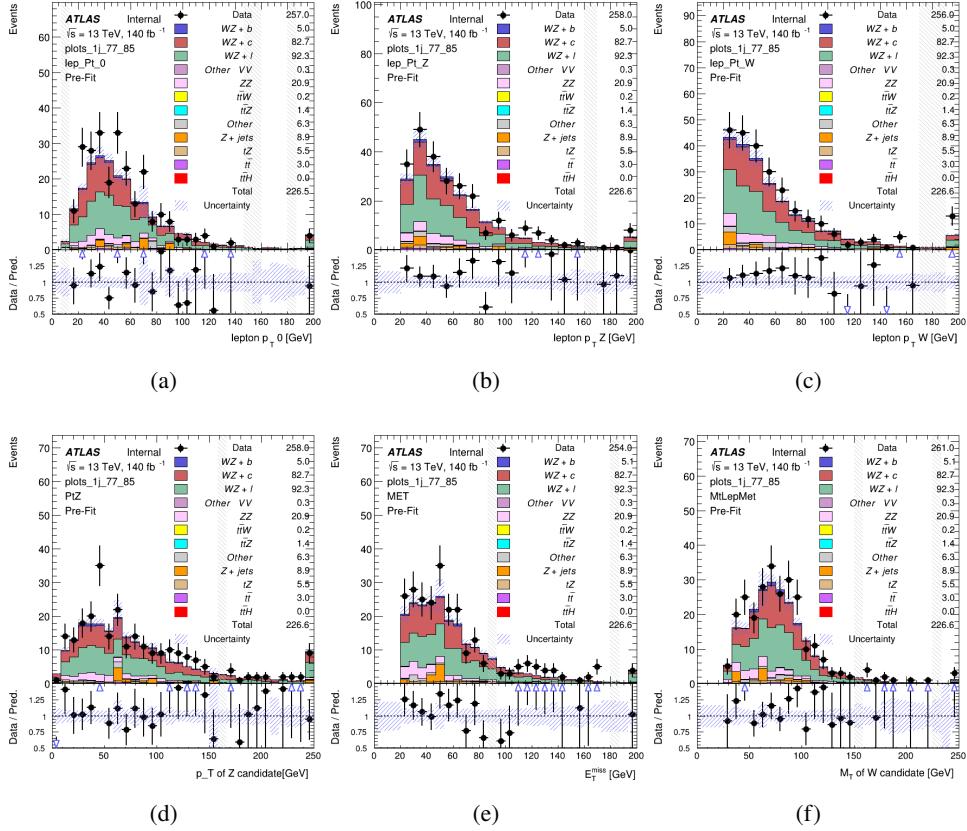


Figure 5: Comparisons between the data and MC distributions in the preselection region for (a) the p_T of the opposite sign lepton, (b) the p_T of the other lepton from the Z candidate, (c) the p_T of the lepton from the W candidate, (d) the p_T of the Z candidate, (e) the E_T^{miss} , and (f) the m_T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

WZ Fit Region - 1j 70-77% WP

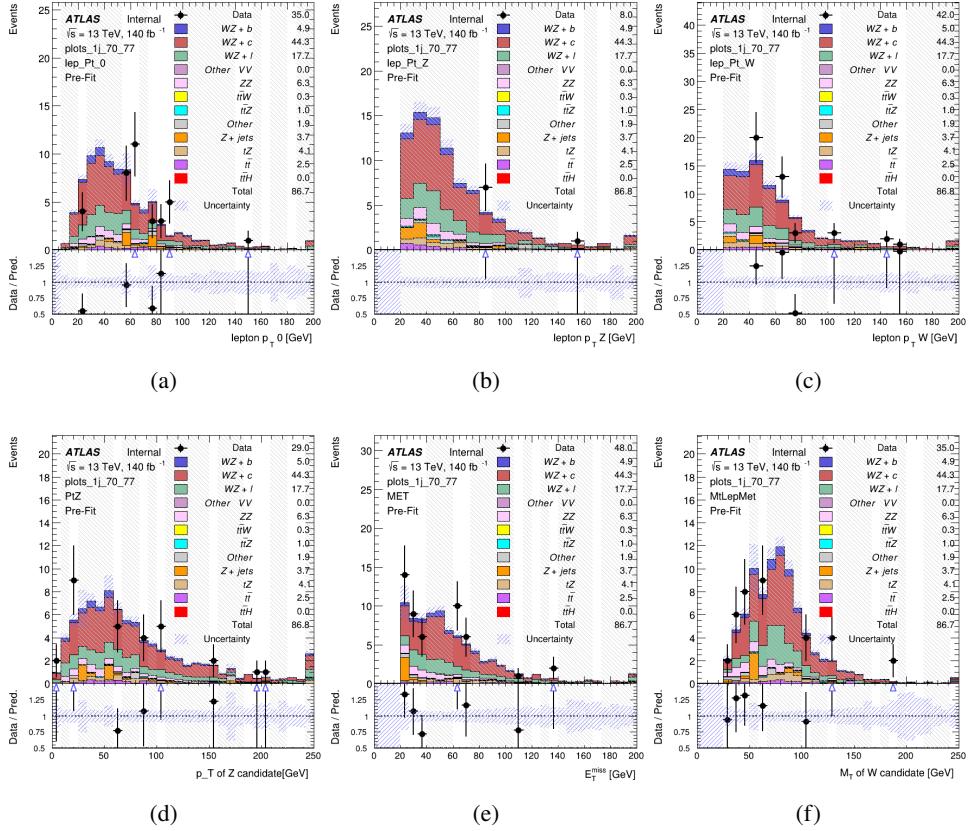


Figure 6: Comparisons between the data and MC distributions in the preselection region for (a) the p_T of the opposite sign lepton, (b) the p_T of the other lepton from the Z candidate, (c) the p_T of the lepton from the W candidate, (d) the p_T of the Z candidate, (e) the E_T^{miss} , and (f) the m_T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

WZ Fit Region - 1j 60-70% WP

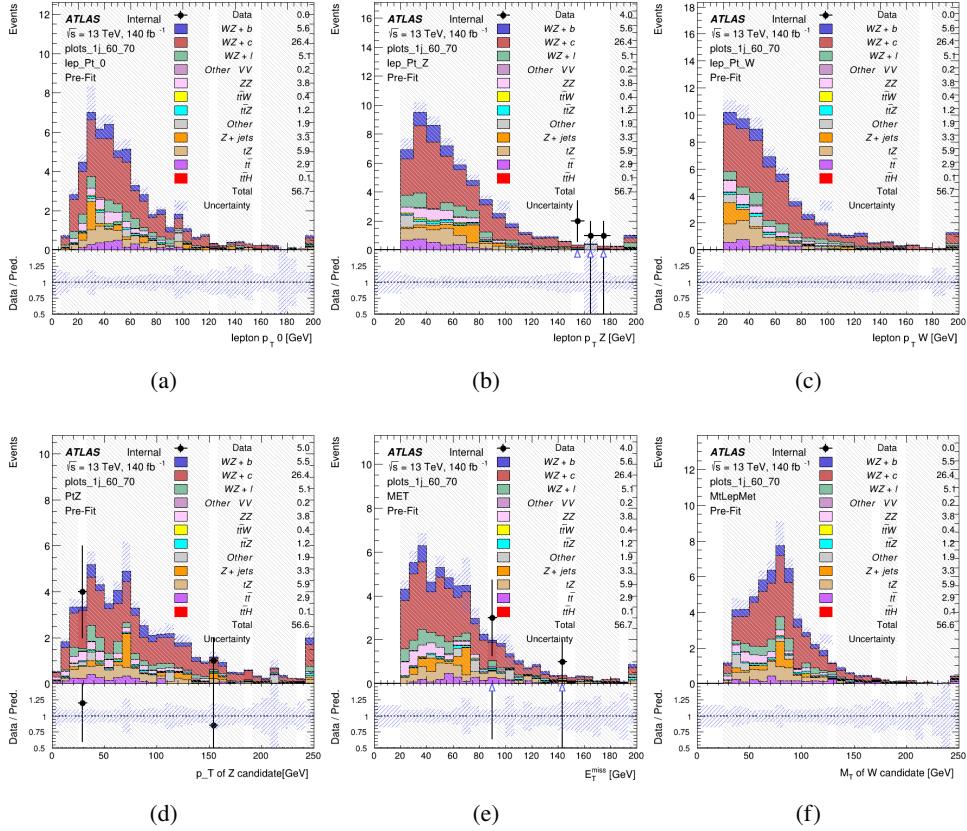


Figure 7: Comparisons between the data and MC distributions in the preselection region for (a) the p_T of the opposite sign lepton, (b) the p_T of the other lepton from the Z candidate, (c) the p_T of the lepton from the W candidate, (d) the p_T of the Z candidate, (e) the E_T^{miss} , and (f) the m_T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

WZ Fit Region - 1j 60% WP

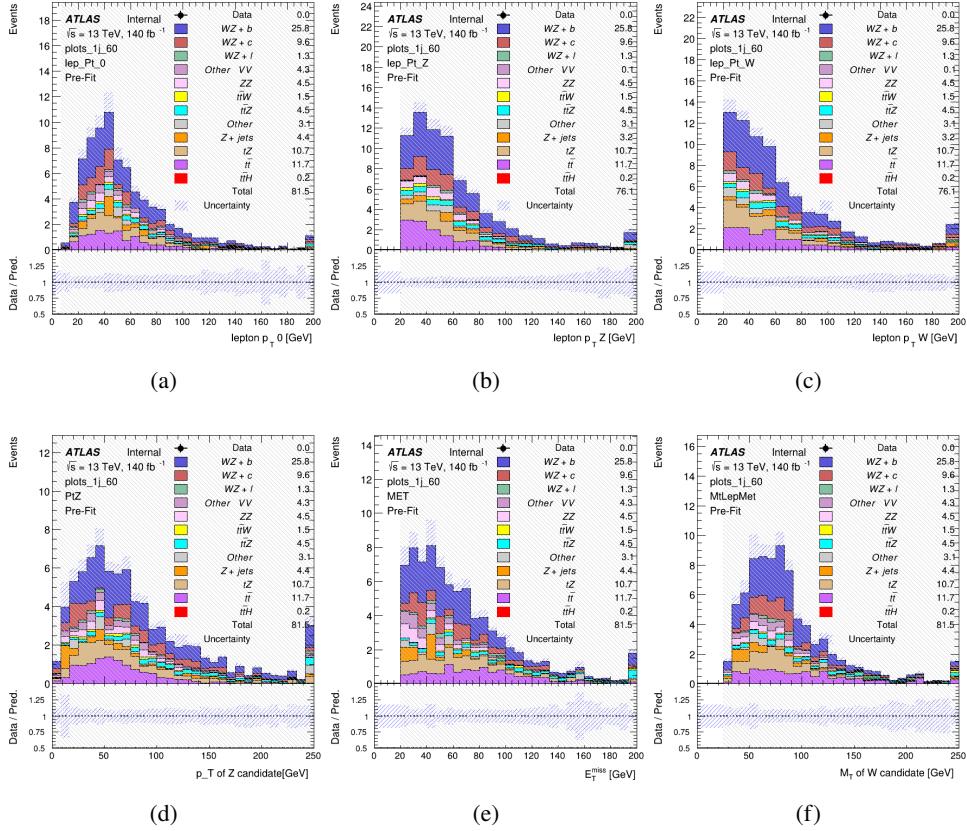


Figure 8: Comparisons between the data and MC distributions in the preselection region for (a) the p_T of the opposite sign lepton, (b) the p_T of the other lepton from the Z candidate, (c) the p_T of the lepton from the W candidate, (d) the p_T of the Z candidate, (e) the E_T^{miss} , and (f) the m_T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

WZ Fit Region - tZ-CR

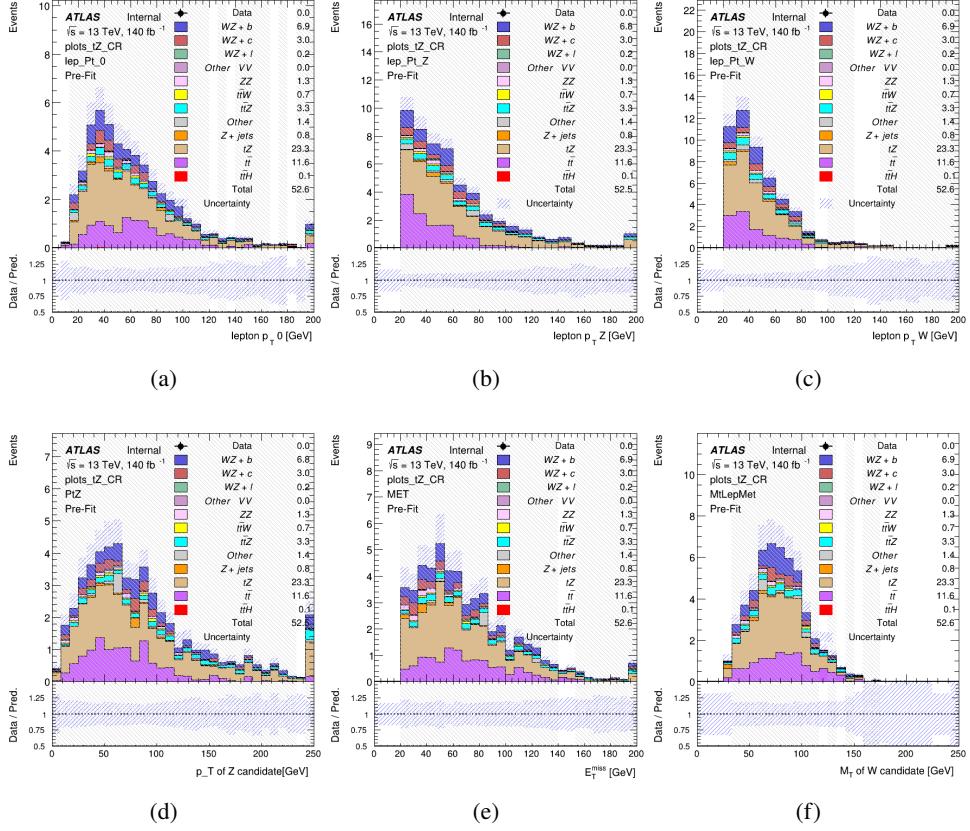


Figure 9: Comparisons between the data and MC distributions in the preselection region for (a) the p_T of the opposite sign lepton, (b) the p_T of the other lepton from the Z candidate, (c) the p_T of the lepton from the W candidate, (d) the p_T of the Z candidate, (e) the E_T^{miss} , and (f) the m_T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

WZ Fit Region - 2j Inclusive

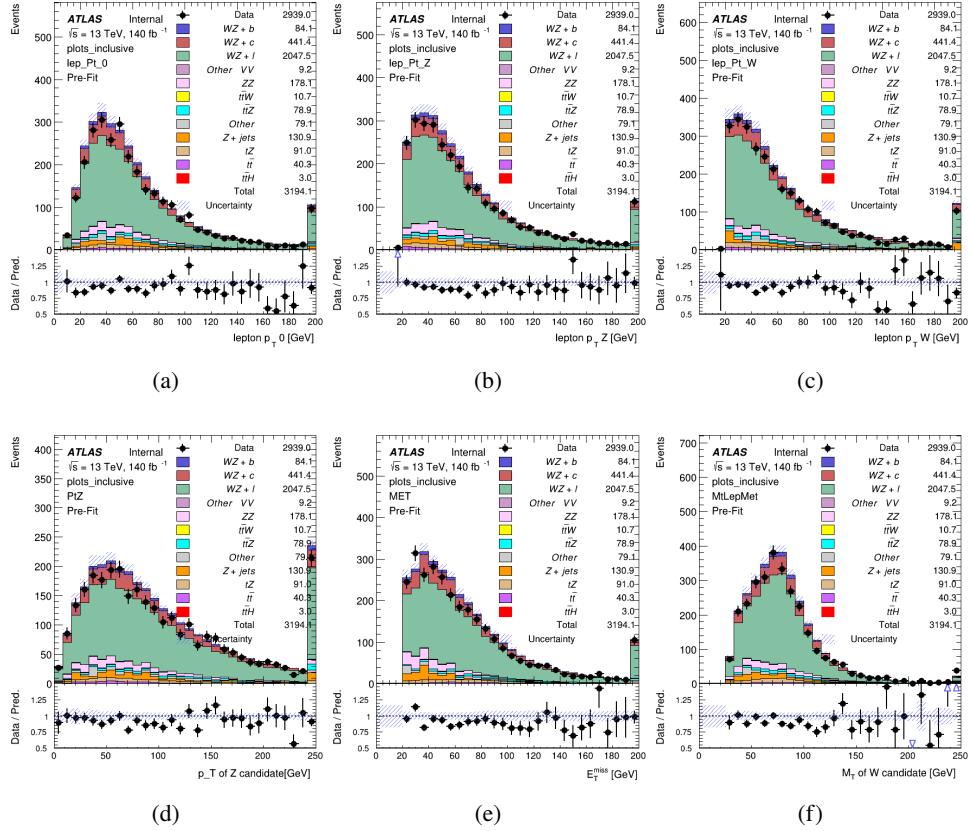


Figure 10: Comparisons between the data and MC distributions in the preselection region for (a) the p_T of the opposite sign lepton, (b) the p_T of the other lepton from the Z candidate, (c) the p_T of the lepton from the W candidate, (d) the p_T of the Z candidate, (e) the E_T^{miss} , and (f) the m_T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

WZ Fit Region - 2j < 85% WP

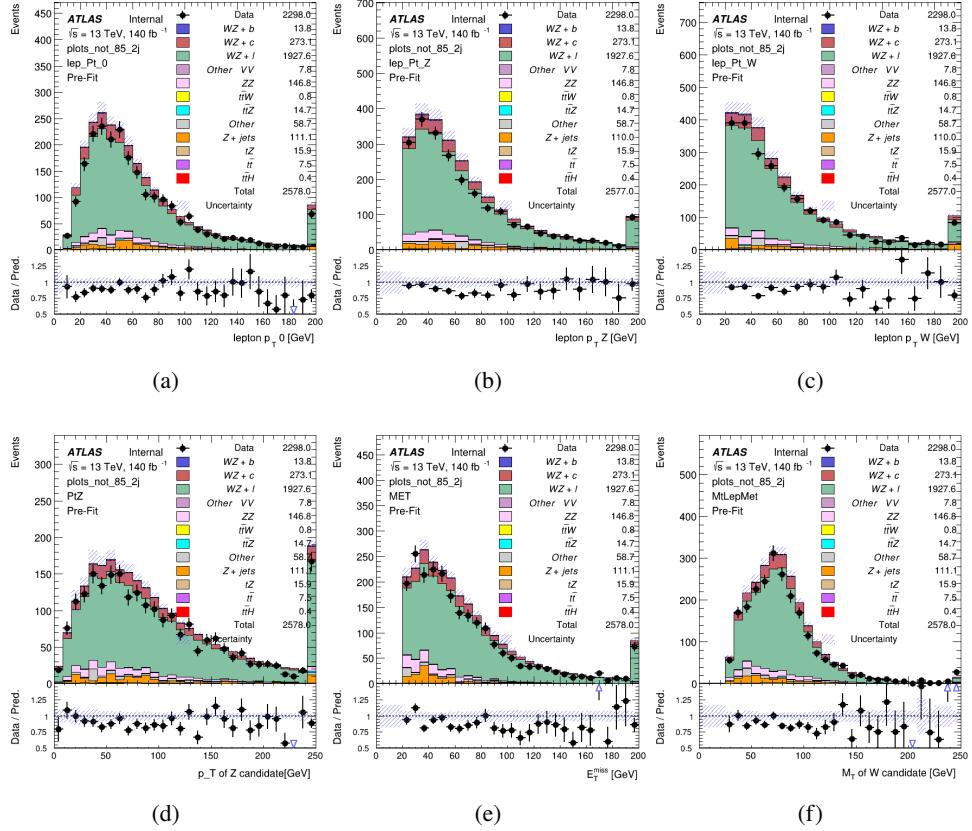


Figure 11: Comparisons between the data and MC distributions in the preselection region for (a) the p_T of the opposite sign lepton, (b) the p_T of the other lepton from the Z candidate, (c) the p_T of the lepton from the W candidate, (d) the p_T of the Z candidate, (e) the E_T^{miss} , and (f) the m_T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

WZ Fit Region - 2j 77-85% WP

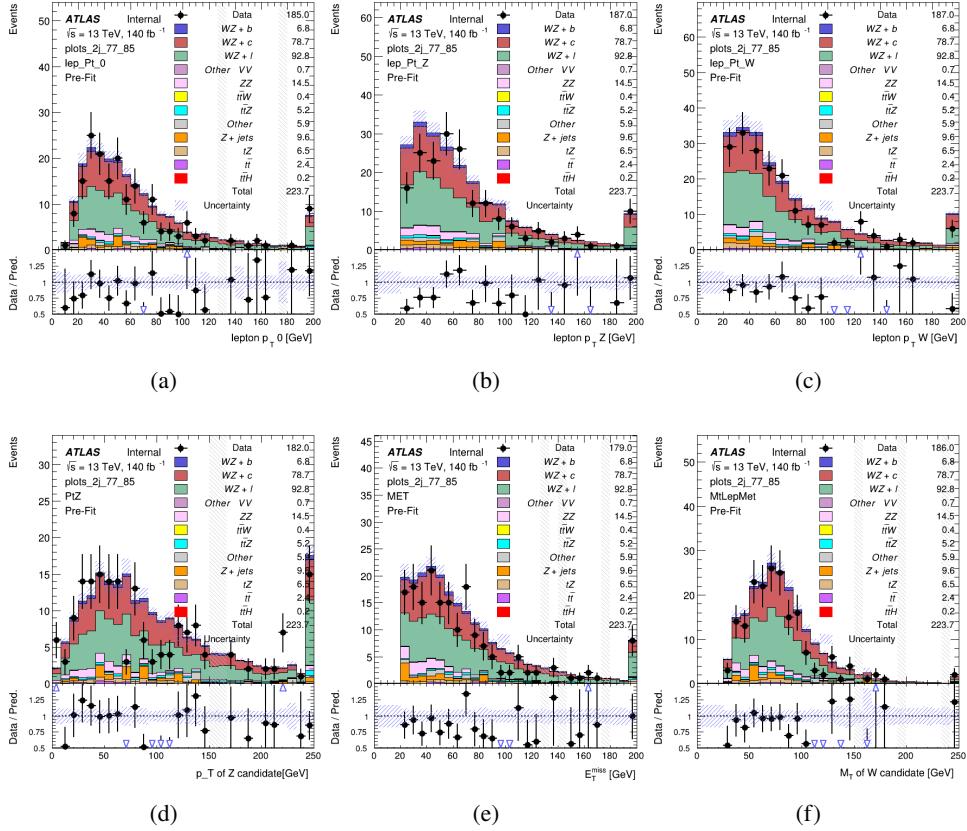


Figure 12: Comparisons between the data and MC distributions in the preselection region for (a) the p_T of the opposite sign lepton, (b) the p_T of the other lepton from the Z candidate, (c) the p_T of the lepton from the W candidate, (d) the p_T of the Z candidate, (e) the E_T^{miss} , and (f) the m_T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

WZ Fit Region - 2j 70-77% WP

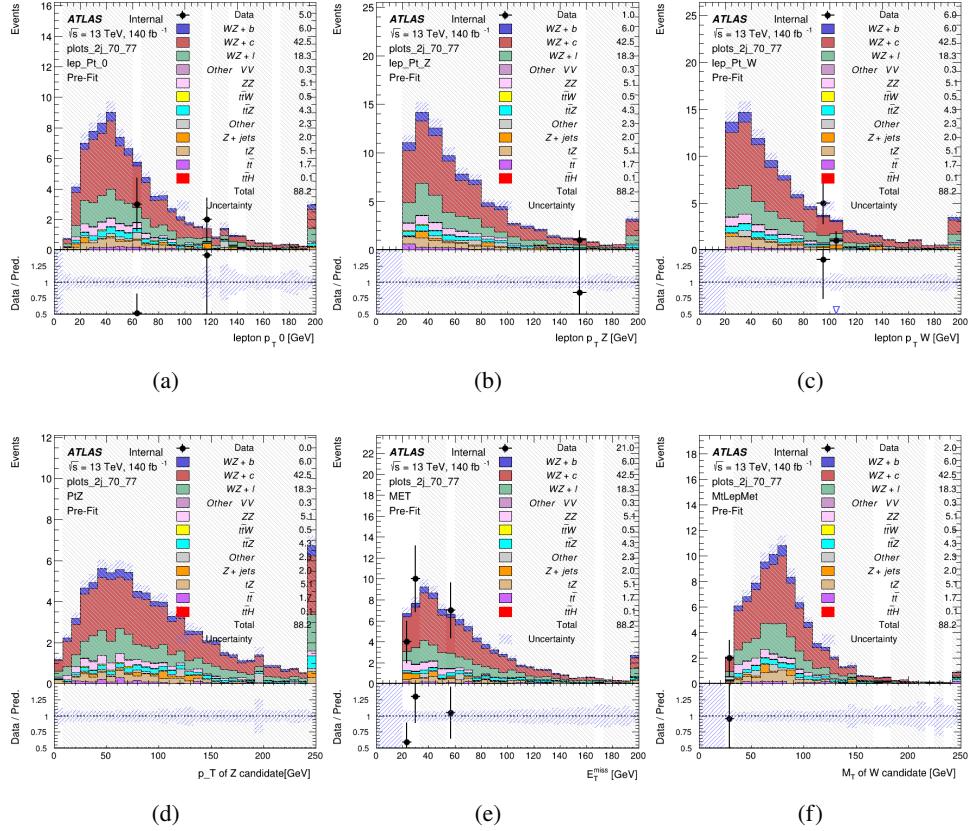


Figure 13: Comparisons between the data and MC distributions in the preselection region for (a) the p_T of the opposite sign lepton, (b) the p_T of the other lepton from the Z candidate, (c) the p_T of the lepton from the W candidate, (d) the p_T of the Z candidate, (e) the E_T^{miss} , and (f) the m_T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

WZ Fit Region - 2j 60-70% WP

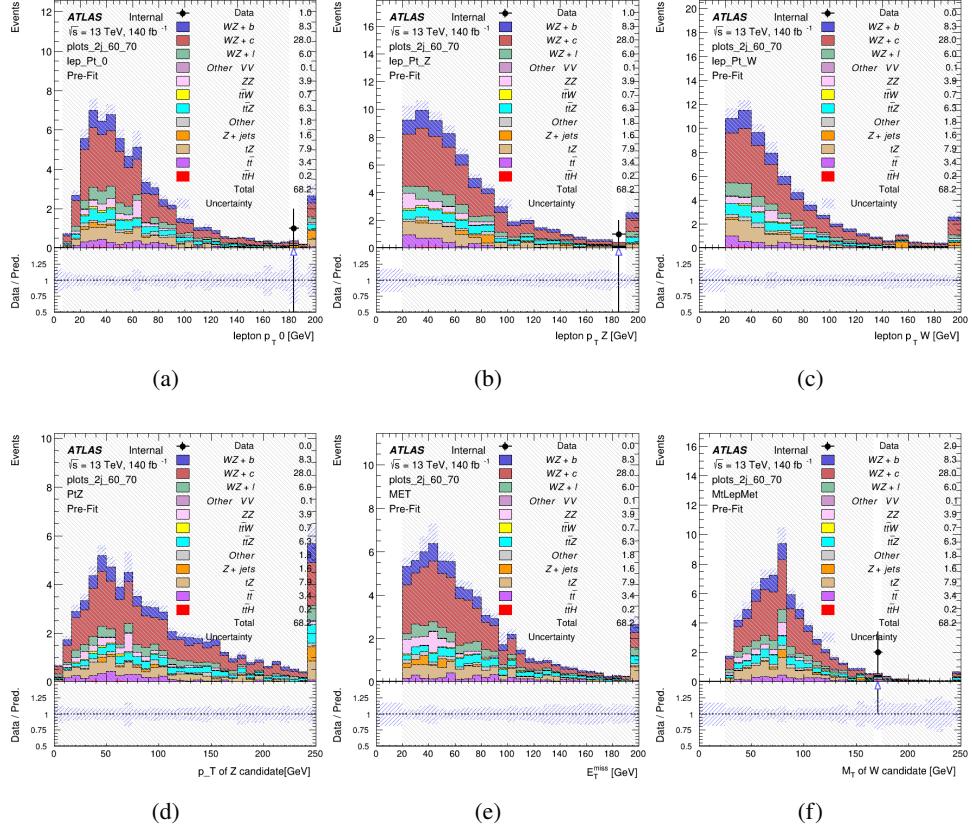


Figure 14: Comparisons between the data and MC distributions in the preselection region for (a) the p_T of the opposite sign lepton, (b) the p_T of the other lepton from the Z candidate, (c) the p_T of the lepton from the W candidate, (d) the p_T of the Z candidate, (e) the E_T^{miss} , and (f) the m_T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

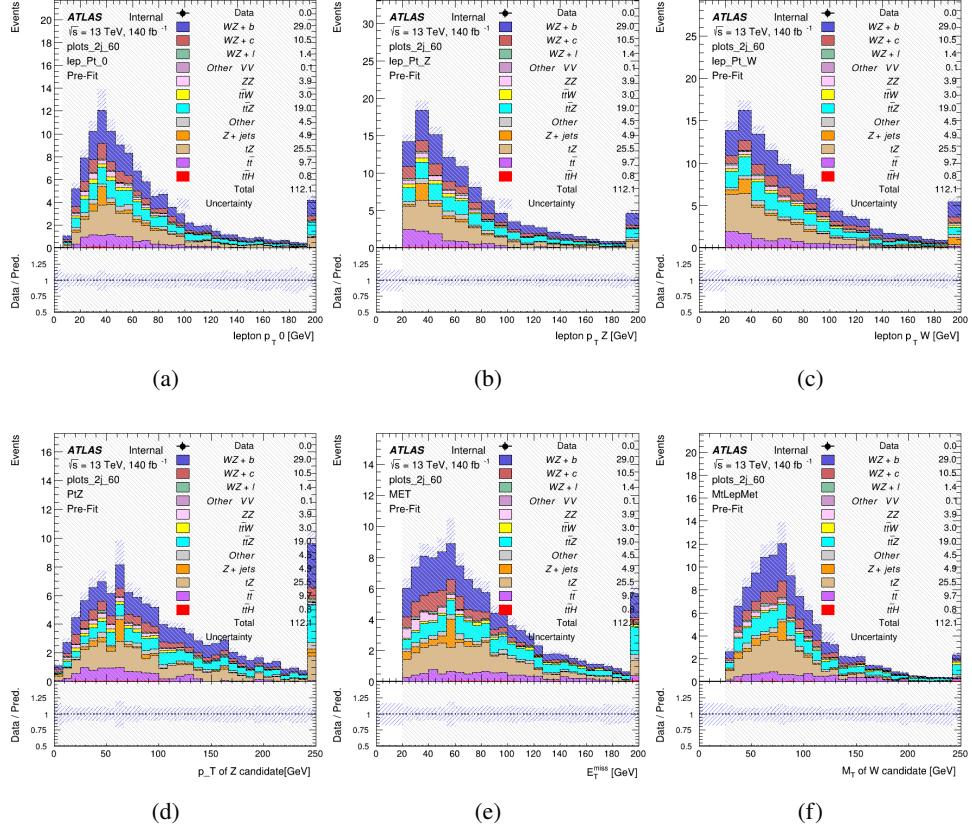
WZ Fit Region - 2j 60% WP

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

WZ Fit Region - tZ-CR-2j

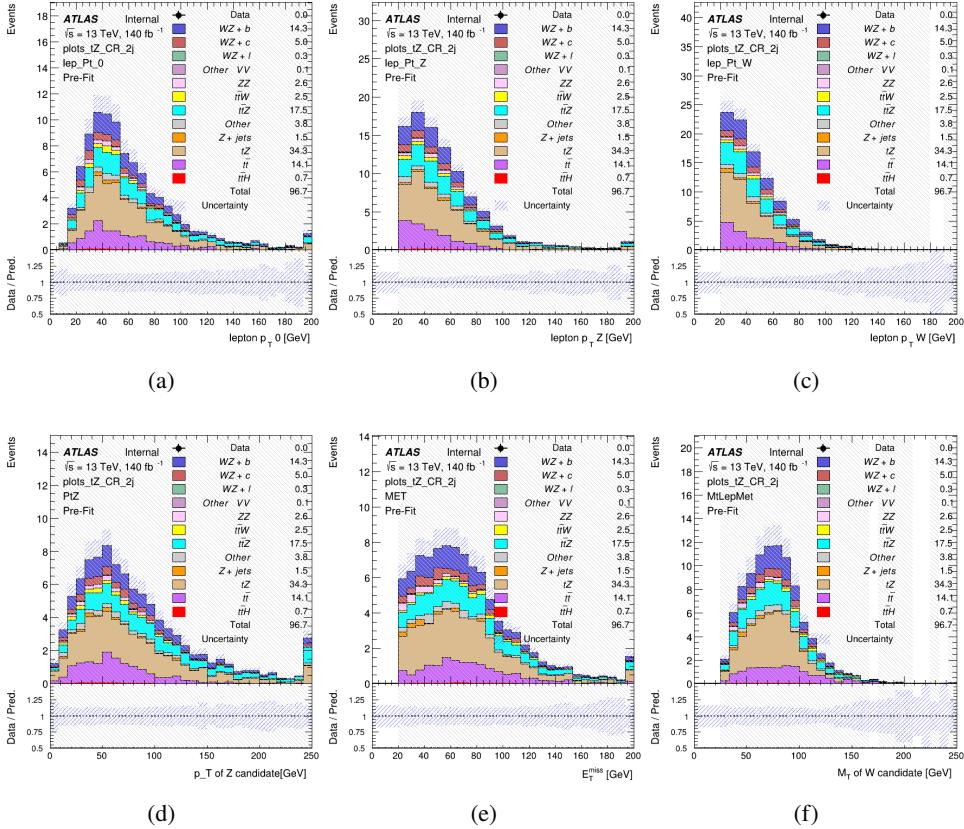


Figure 16: Comparisons between the data and MC distributions in the preselection region for (a) the p_T of the opposite sign lepton, (b) the p_T of the other lepton from the Z candidate, (c) the p_T of the lepton from the W candidate, (d) the p_T of the Z candidate, (e) the E_T^{miss} , and (f) the m_T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

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

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

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

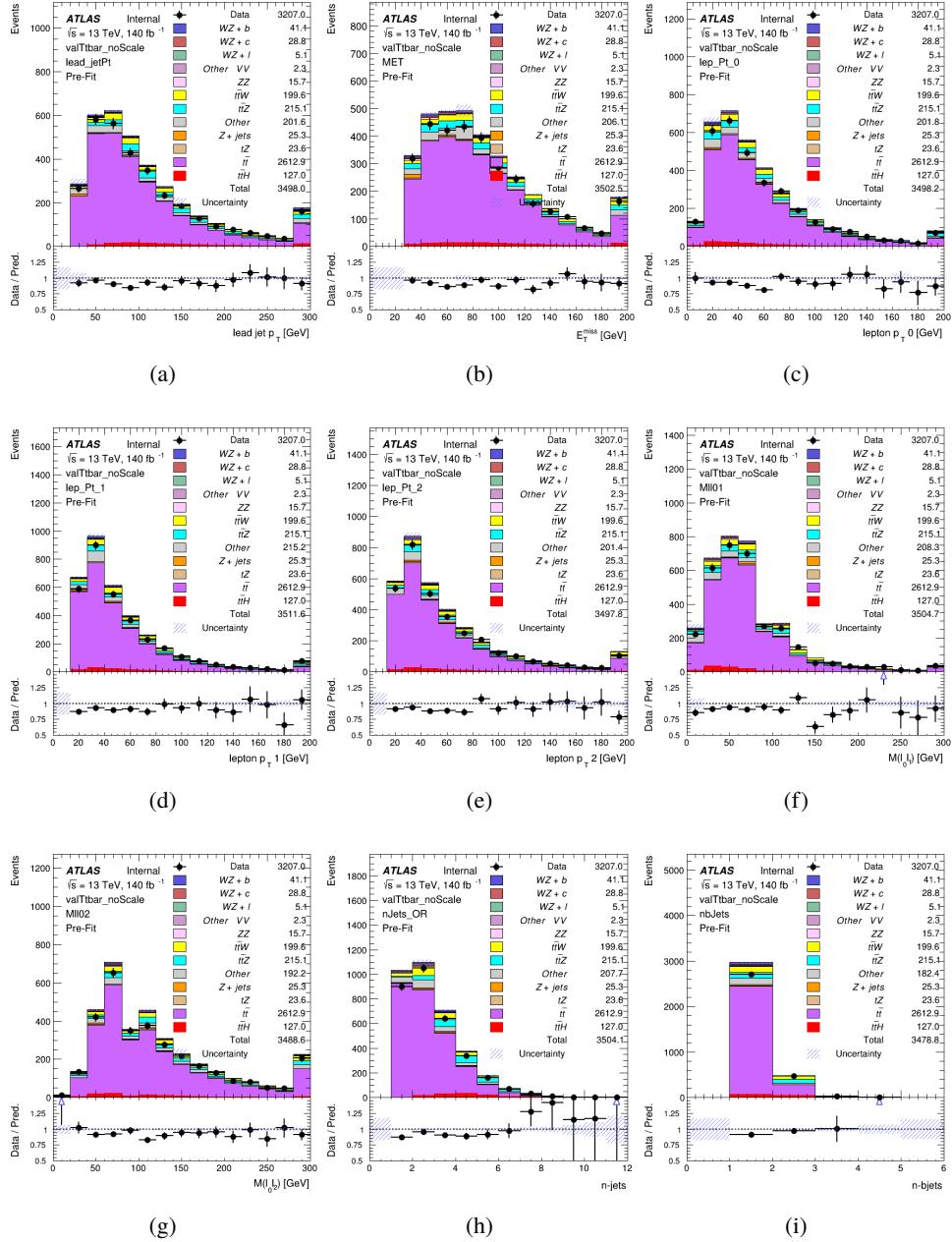


Figure 17: Comparisons between the data and MC distributions in the $t\bar{t}$ validation region for (a) the p_T of the leading jet, (b) the missing transverse energy, (c) the p_T of lepton 0, (d) p_T of lepton 1, (e) p_T of lepton 2, (f) the invariant mass of leptons 0 and 1, (g) the invariant mass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

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

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

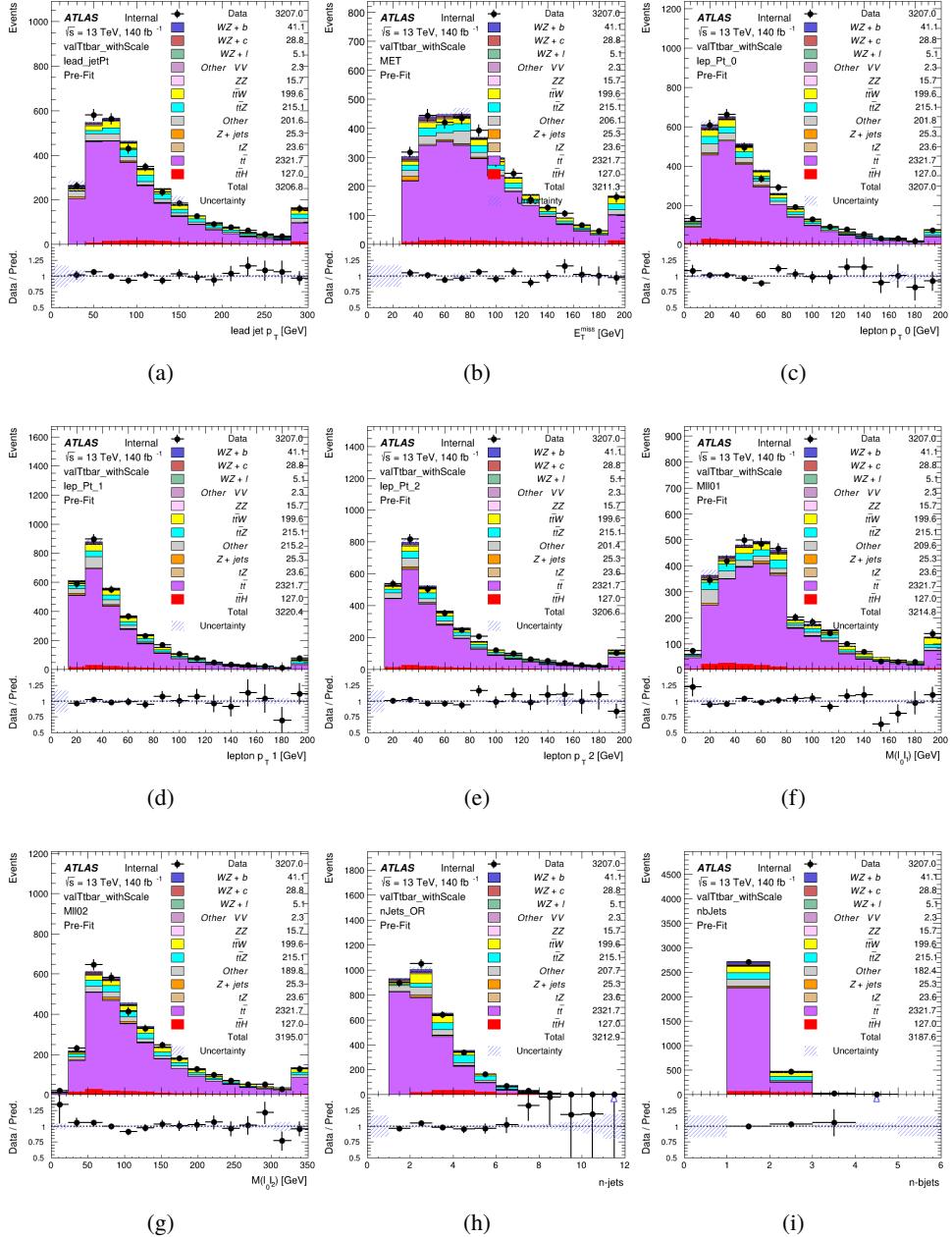


Figure 18: Comparisons between the data and MC distributions in the $t\bar{t}$ validation region after the correction factor has been applied for (a) the p_T of the leading jet, (b) the missing transverse energy, (c) the p_T of lepton 0, (d) p_T of lepton 1, (e) p_T of lepton 2, (f) the invariant mass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

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

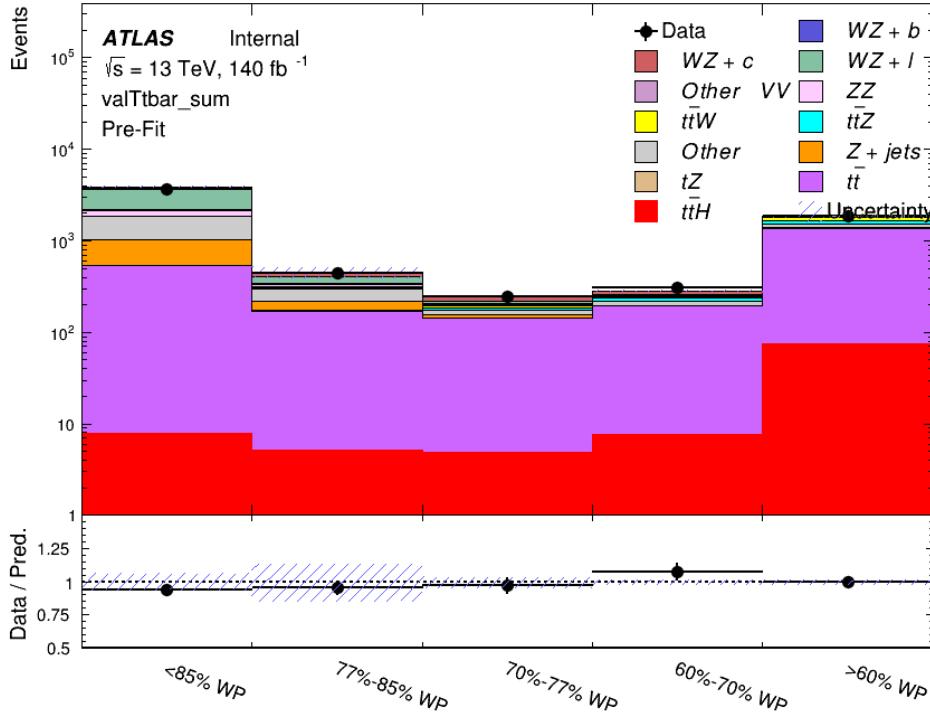


Figure 19: Data and MC comparisons for each DL1r WP for both 1-jet and 2-jet regions, after the $t\bar{t}$ VR selection and correction factor have been applied

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

299 5.3.2 Z+jets Validation

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

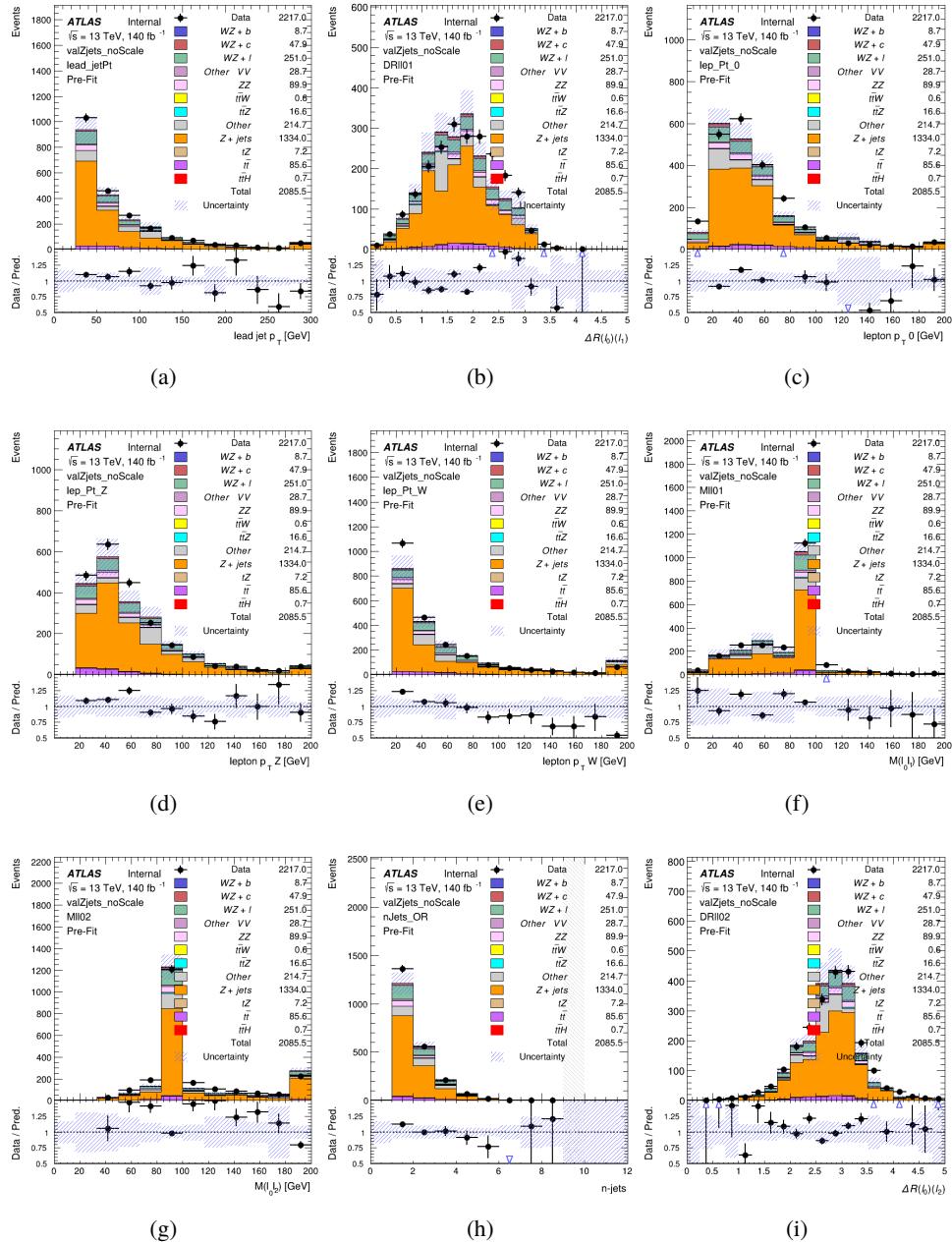


Figure 20: Comparisons between the data and MC distributions in the $Z+jets$ validation region for (a) the p_T of the leading jet, (b) ΔR between leptons 0 and 1, (c) the p_T of lepton 0, (d) p_T of SS lepton from the Z candidate, (e) p_T of the SS lepton from the W candidate, (f) the invariant mass 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

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

³⁰⁸ of the p_T spectrum of the lepton from the W is found to differ. To account for this discrepancy,
³⁰⁹ a variable correction factor is applied to Z+jets. χ^2 minimization of the lepton 2 p_T spectrum is
³¹⁰ performed to derive a correction factor of $1.53 - 6.6 \cdot 10^{-6}(\text{lep}_\text{Pt}_W)$. Kinematic plots of the
³¹¹ Z + jets validation region after this correction factor has been applied are shown in figure 21.

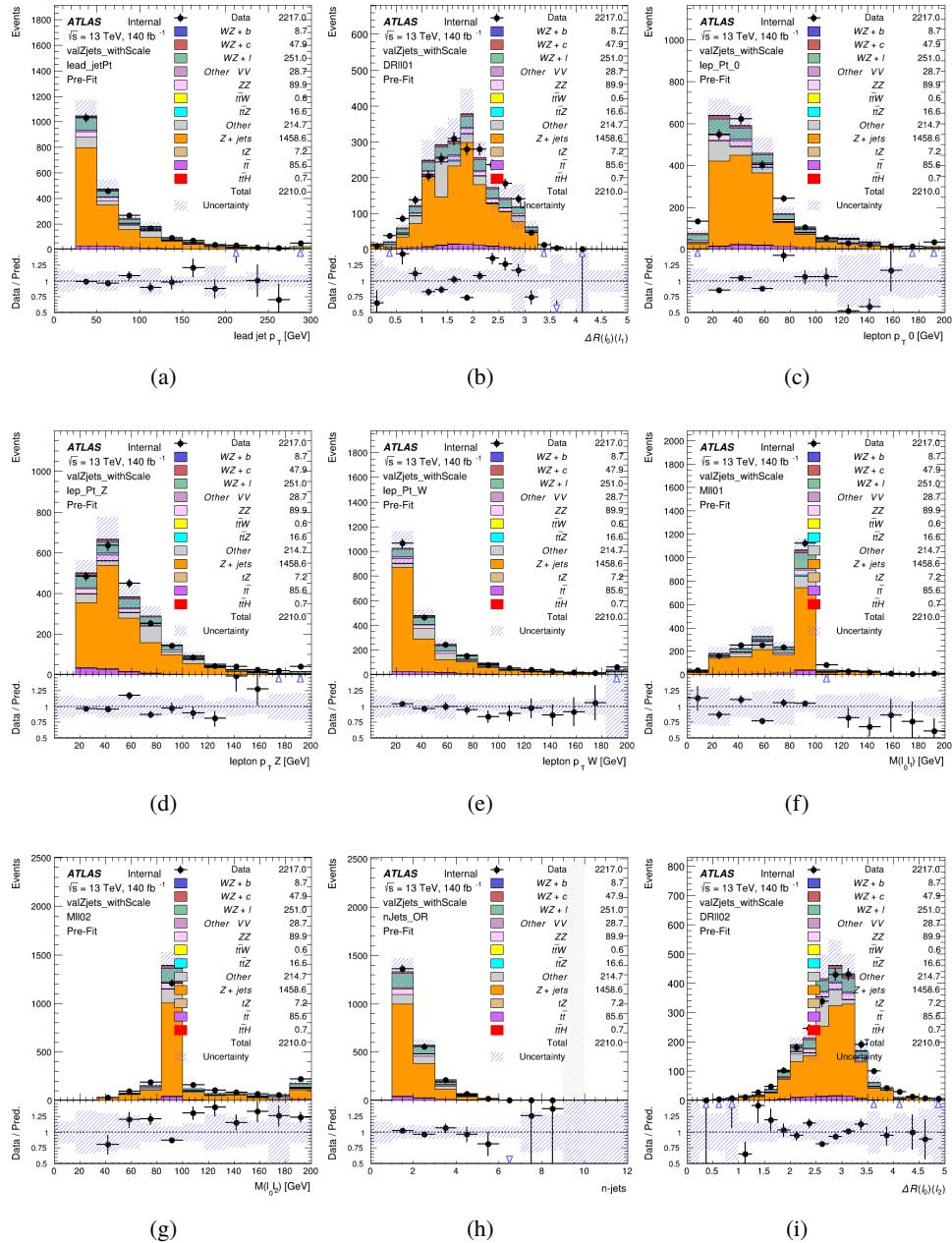


Figure 21: Comparisons between the data and MC distributions in the Z+jets validation region after the correction factor has been applied for (a) the p_T of the leading jet, (b) ΔR between leptons 0 and 1, (c) the p_T of lepton 0, (d) p_T of SS lepton from the Z candidate, (e) p_T of the SS lepton from 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

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

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

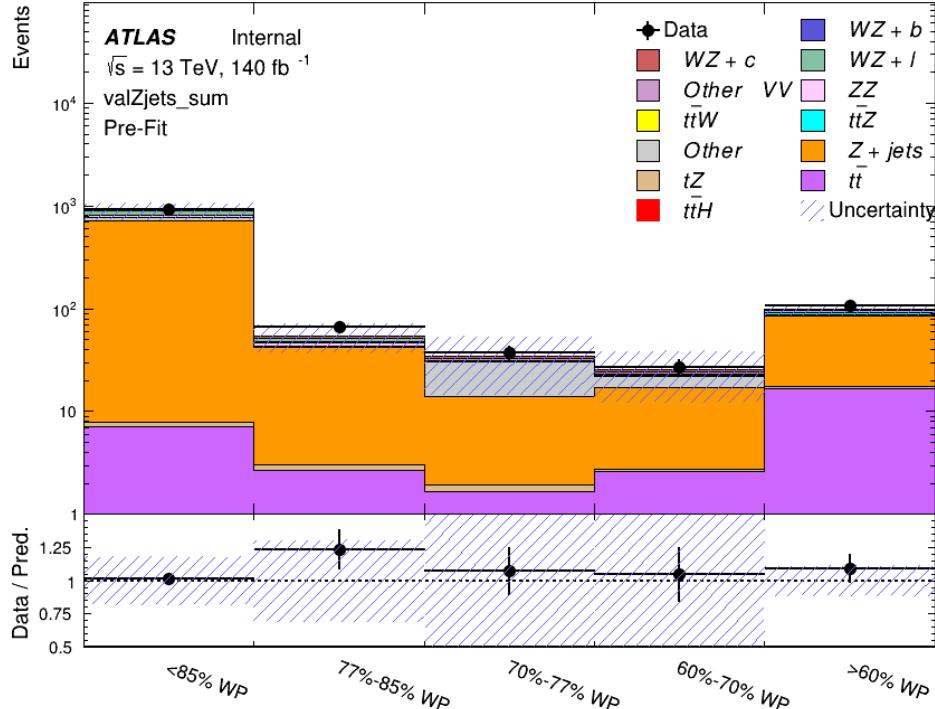


Figure 22: Data and MC comparisons for each DL1r WP for both 1-jet and 2-jet regions, after the Z+jets VR selection and correction factor have been applied

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

318 6 tZ Interference Studies and Separation Multivariate Analysis

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

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

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

328 **6.1 Top Mass Reconstruction**

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

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

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

$$339 \quad (p_l + p_\nu)^2 = m_W^2$$

340 Expanding this out into components, this equation gives:

$$341 \quad \sqrt{p_{T\nu}^2 + p_{z\nu}^2} E_l = \frac{m_w^2 - m_l^2}{2} + p_{T\nu}(p_{lx}\cos\phi_\nu + p_{ly}\sin\phi_\nu) + p_{lz}p_{\nu z}$$

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

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

348 **6.2 tZ BDT**

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

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

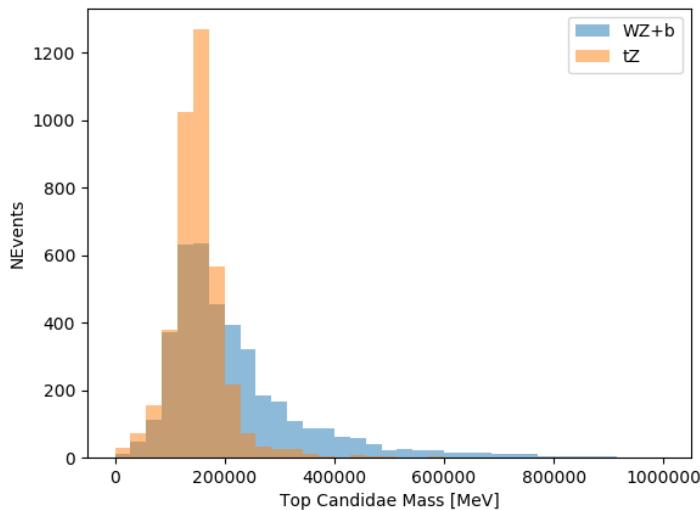


Figure 23: Reconstructed top mass distributions for tZ and WZ + b, measured in MeV.

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

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

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

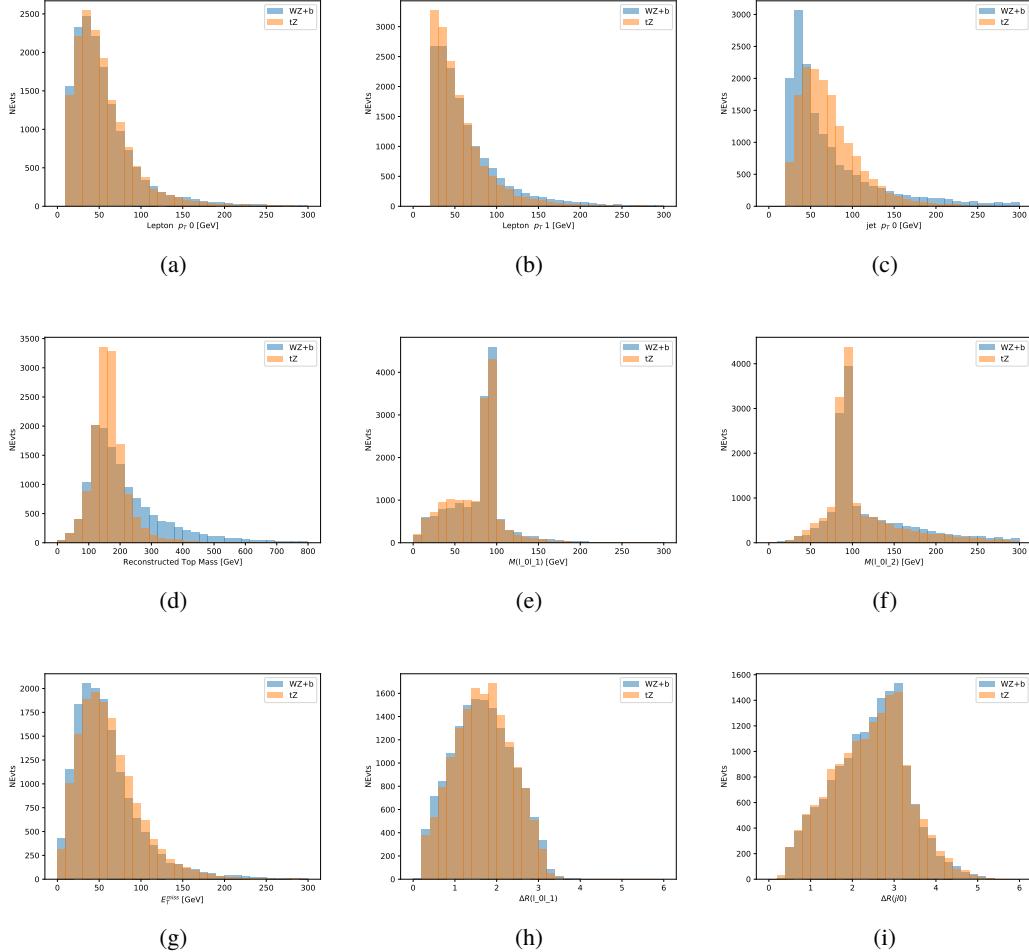


Figure 24: Distribution of input features of the BDT for signal (WZ) and background (tZ). Both are scaled to an equal number of events. (a), (b) and (c) show the p_T of lepton 0, lepton 1, and the jet, (d) show the reconstructed top mass, (e) and (f) show the invariant mass of leptons 0 and 1, leptons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

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

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

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

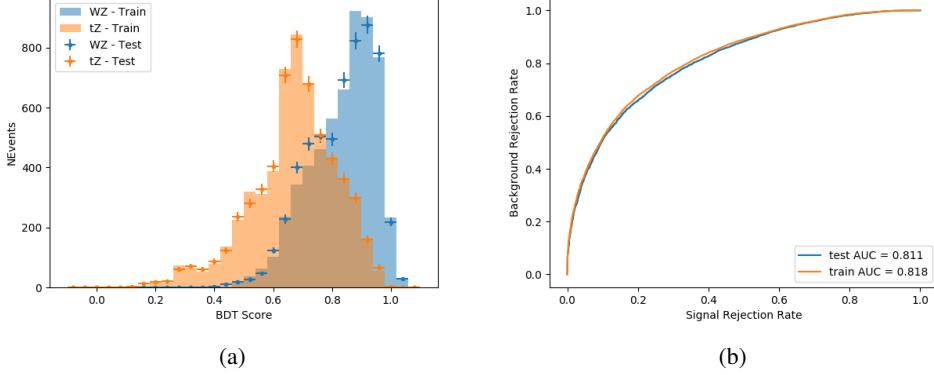


Figure 25: Distribution of the BDT response for signal and background events on the left, the ROC curve for the BDT on the right.

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

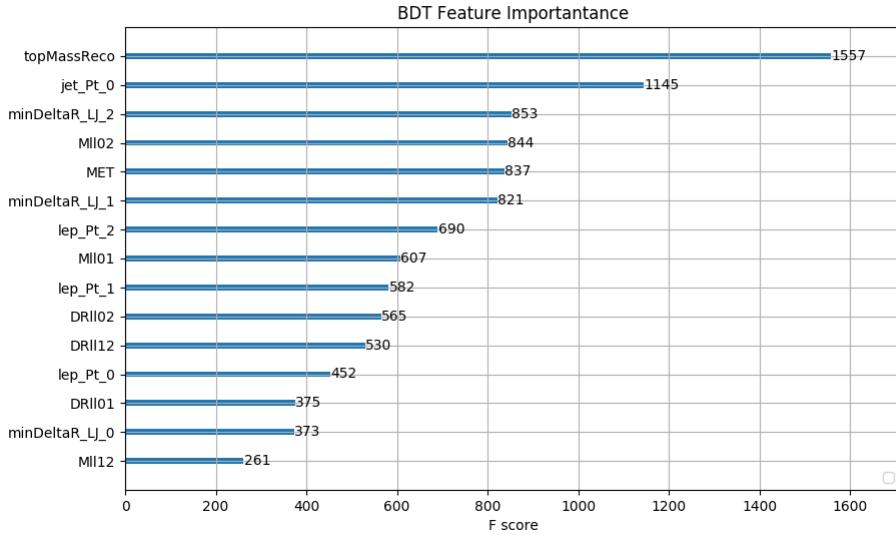


Figure 26: Relative importance of each input feature in the model.

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

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

381 7 Systematic Uncertainties

382 The systematic uncertainties that are considered are summarized in table 8. These are imple-
 383 mented in the fit either as a normalization factors or as a shape variation or both in the signal
 384 and background estimations. The numerical impact of each of these uncertainties is outlined in
 385 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 and resolution	28
Jet vertex fraction	1
Jet flavor tagging	131
E_T^{miss}	3
Total (Experimental)	186
Background Modeling	
Cross section	24
Renormalization and factorization scales	10
Parton shower and hadronization model	2
Shower tune	4
Total (Signal and background modeling)	40
Total (Overall)	226

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

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

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

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

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

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

402

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.

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

Process	X-section [%]
WZ	QCD Scale: $^{+3.7}_{-3.4}$ PDF($+\alpha_S$): ± 3.1
tZ	X-sec: ± 15.2 QCD Scale: $^{+5.2}_{-1.3}$ PDF($+\alpha_S$): ± 1.2
t̄H (aMC@NLO+Pythia8)	QCD Scale: $^{+5.8}_{-9.2}$ PDF($+\alpha_S$): ± 3.6
t̄Z (aMC@NLO+Pythia8)	QCD Scale: $^{+9.6}_{-11.3}$ PDF($+\alpha_S$): ± 4
t̄W (aMC@NLO+Pythia8)	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
t̄t	± 10
Z + jets	± 20

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

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

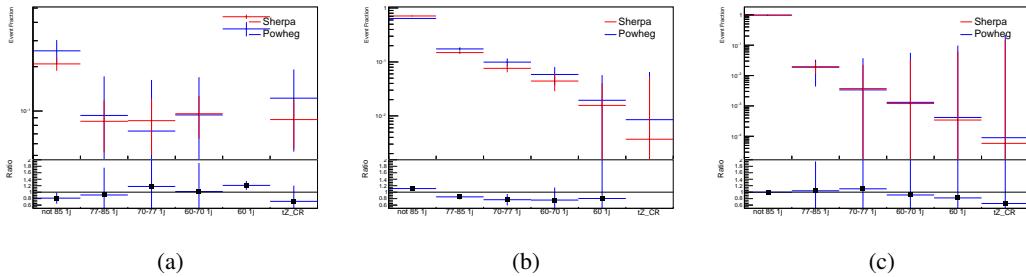


Figure 27: Comparison between Sherpa and Powheg predictions of the distribution of (a) WZ + b, (b) WZ + charm, and (c) WZ + light among the various b-tag WPs used in the 1-jet fit.

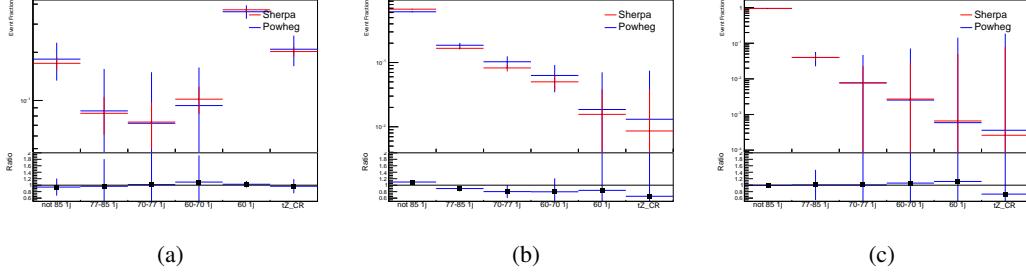


Figure 28: Comparison between Sherpa and Powheg predictions of the distribution of (a) WZ + b, (b) WZ + charm, and (c) WZ + light among the various b-tag WPs used in the 2-jet fit.

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

421 8 Results

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

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

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

436 8.1 1-jet Fit Results

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

438 The prefit yields in each of the regions used in the fit are shown in figure 29.

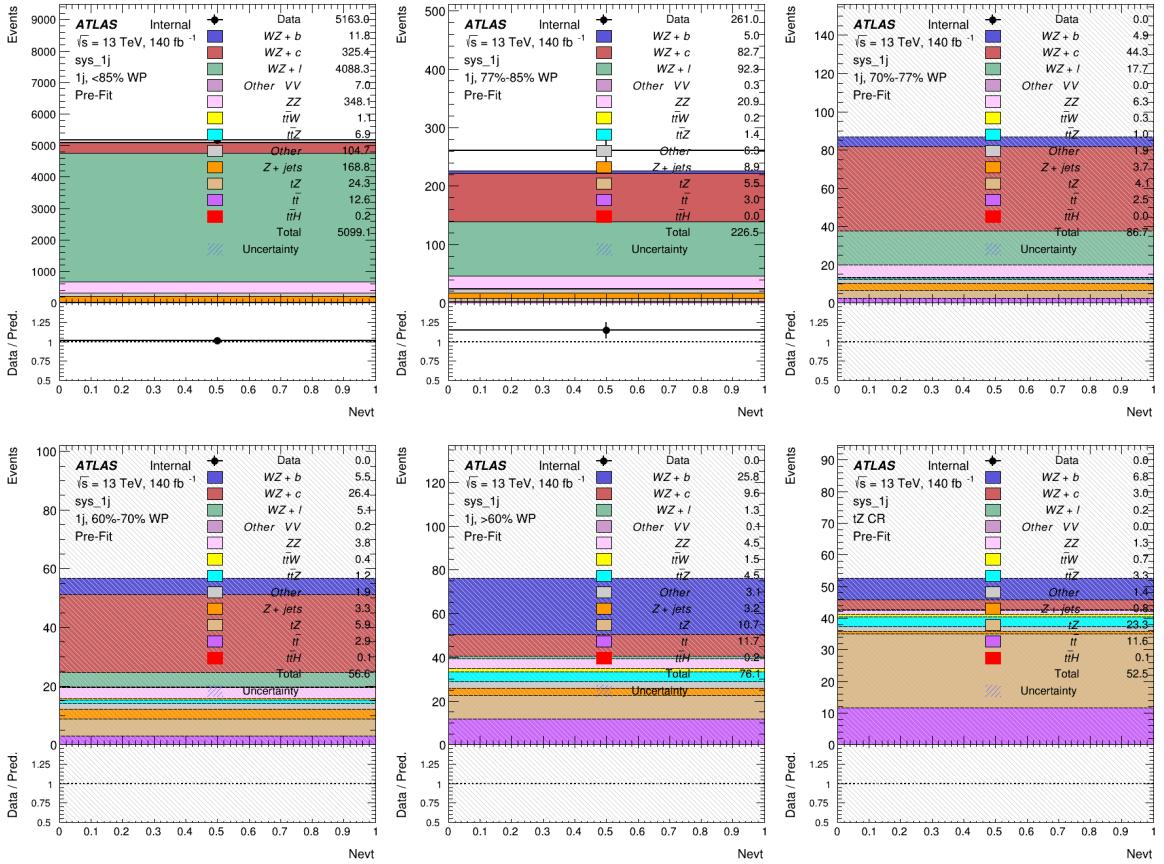


Figure 29: Data/MC yields in each of the 1-jet regions before the fit has been performed.

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

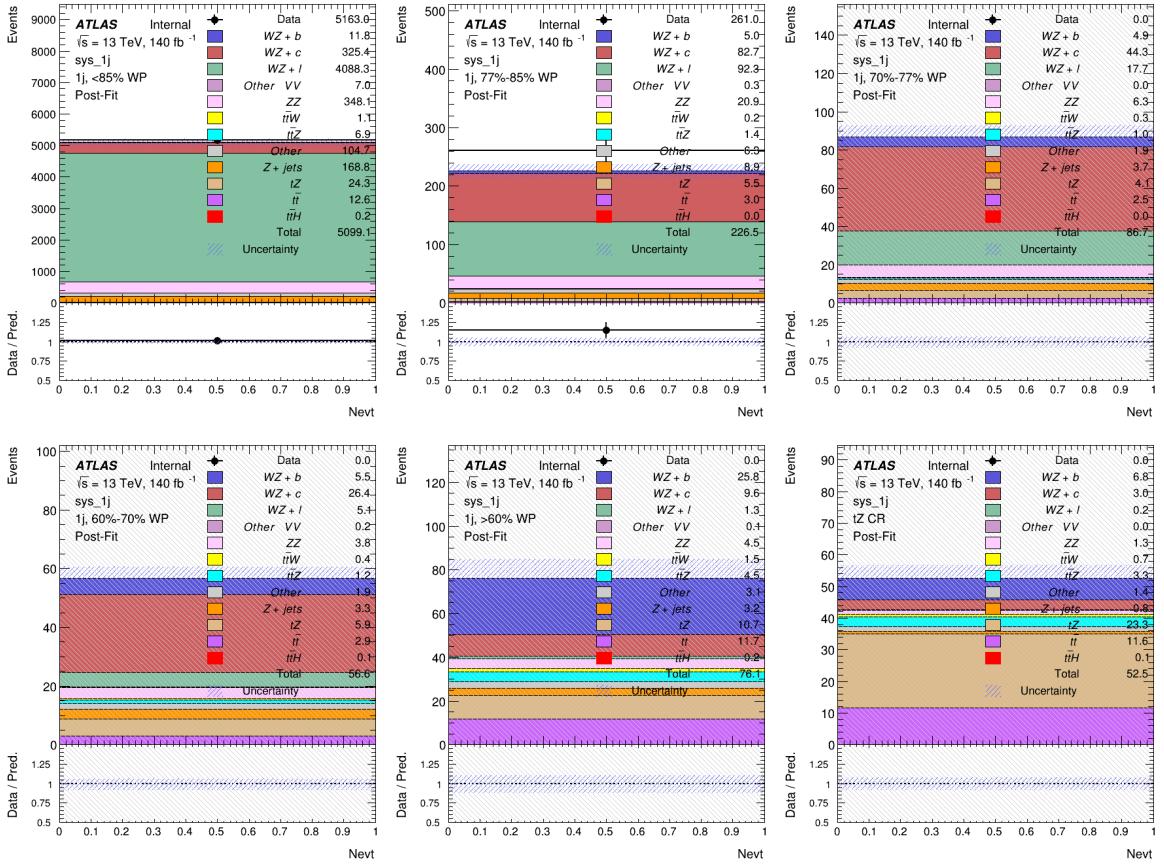


Figure 30: Data/MC results in each of the 1-jet regions after the fit has been performed.

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

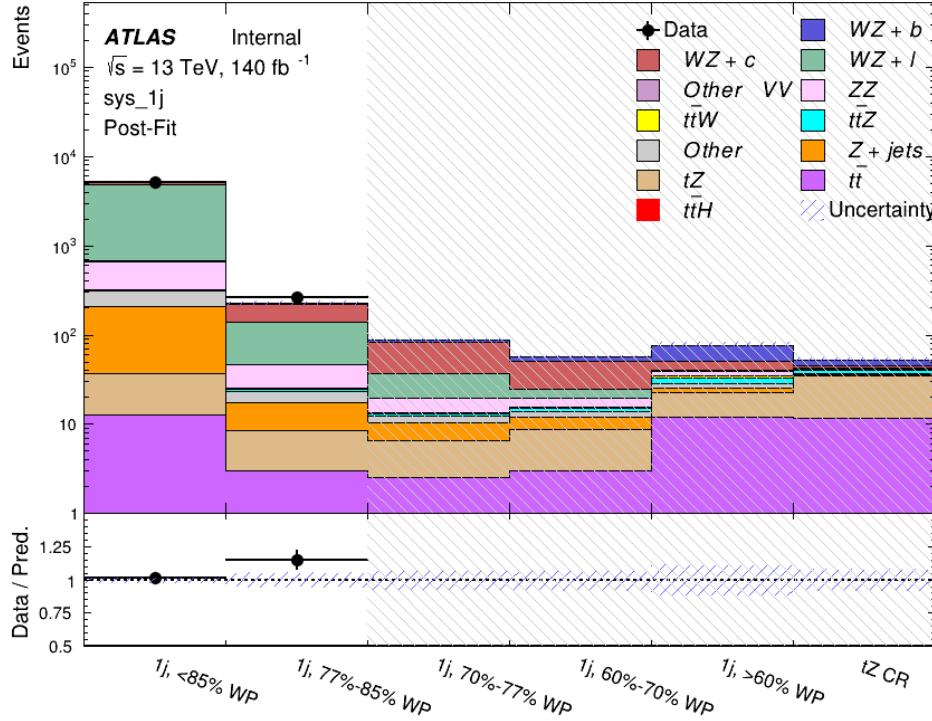


Figure 31: Post-fit summary of the 1-jet fit regions.

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

447 The impact of each NP is calculated by performing the fit with the parameter of interest held
 448 fixed, varied from its fitted value by its uncertainty, and calculating $\Delta\mu$ relative to the baseline
 449 fit. The impact of the most significant sources of systematic uncertainties is summarized in table
 450 13.

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

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

451 The ranking and impact of those nuisance parameters with the largest contribution to the overall
 452 uncertainty is shown in figure 32.

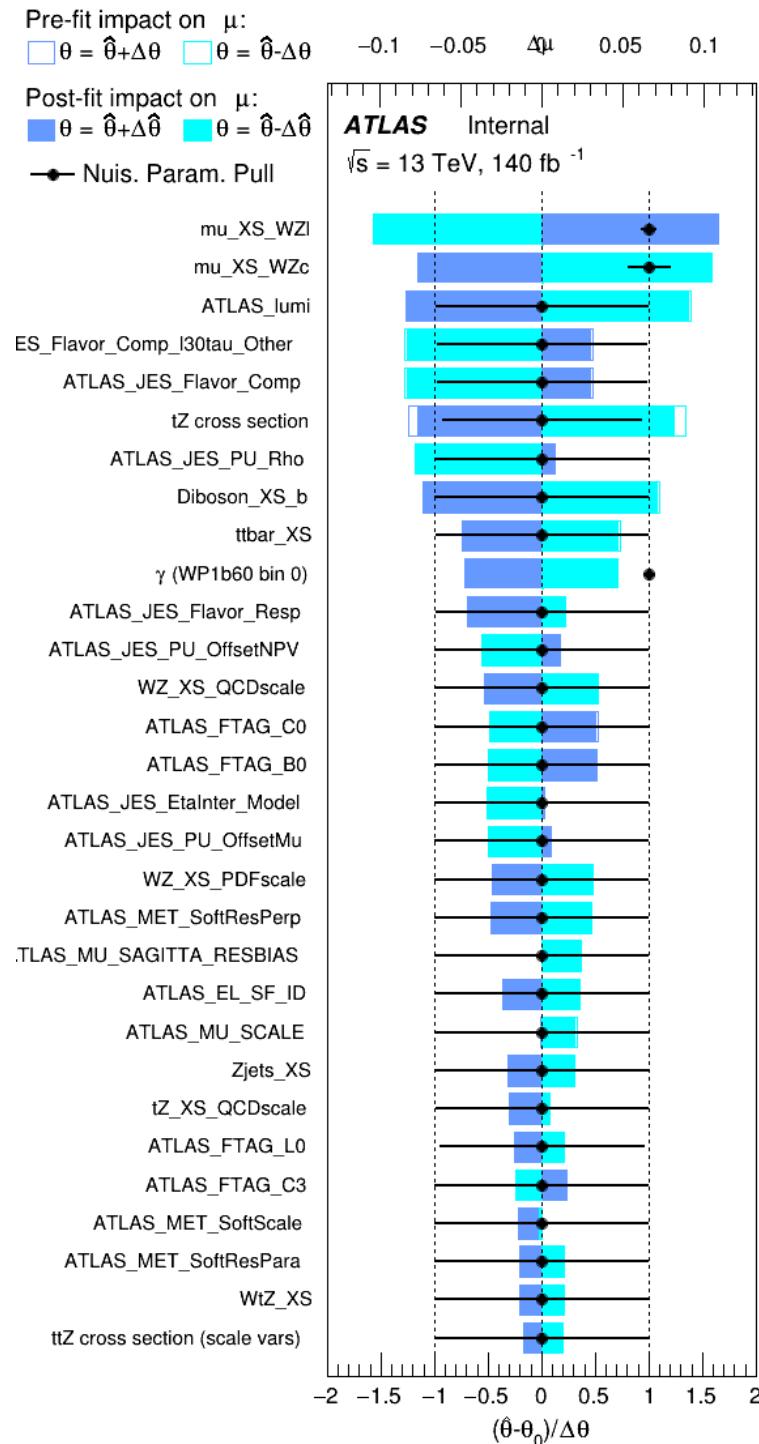


Figure 32: Impact of systematic uncertainties on the signal-strength of $WZ + b$ for events with exactly one jet

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

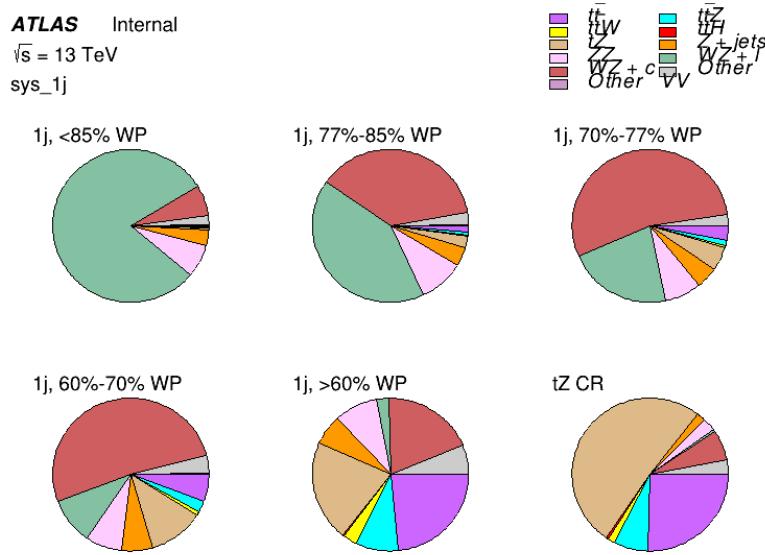


Figure 33: Post-fit background composition of the fit regions.

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

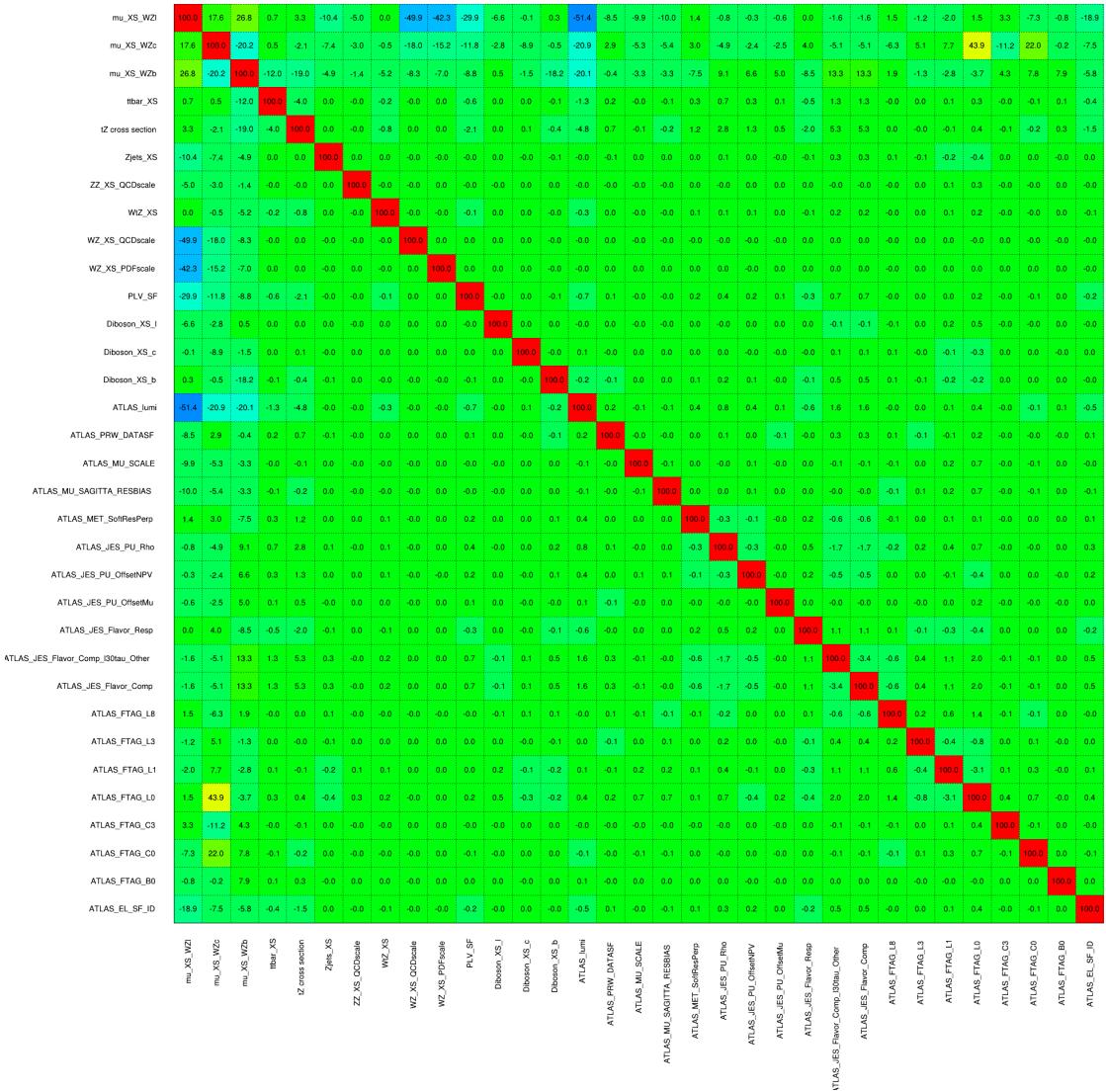


Figure 34: Correlations between nuisance parameters

459 The negative correlations between $\mu_{WZ+\text{charm}}$ and μ_{WZ+b} and $\mu_{WZ+\text{light}}$ are expected: WZ +

460 charm is present in both the WZ + b and WZ + light enriched regions, therefore increasing the

461 fraction of charm requires increasing the fraction of WZ + b and WZ + light. This reasoning

462 also explains the positive correlation between μ_{WZ+b} and $\mu_{WZ+\text{light}}$.

463 Two of the major backgrounds in the region with the highest purity of WZ + b are tZ and Other

464 VV + b, explaining the negative correlations between μ_{WZ+b} and the tZ cross section, and the

465 VV + b cross section.

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

470 8.2 2-jet Fit Results

471 The results of the fit are currently blinded.

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

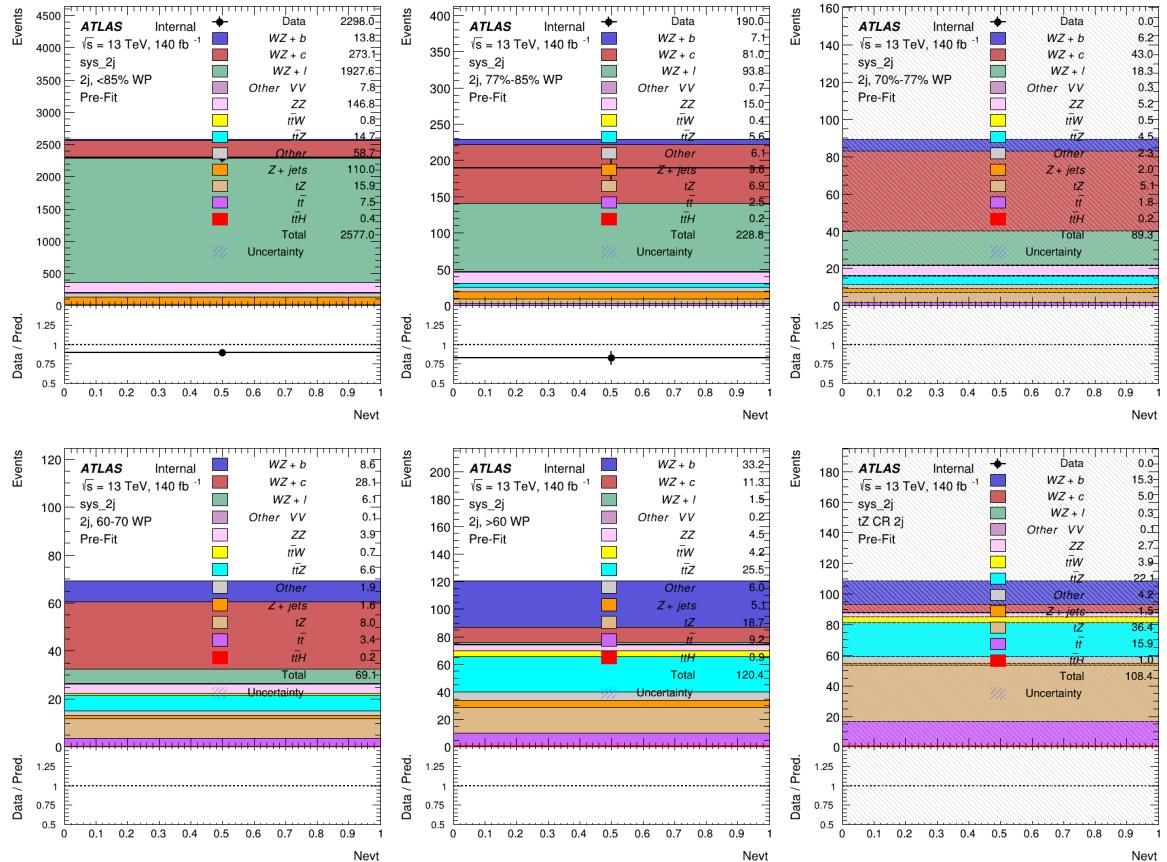


Figure 35: Data/MC yields in each of the regions in the 2-jet fit before the fit has been performed.

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

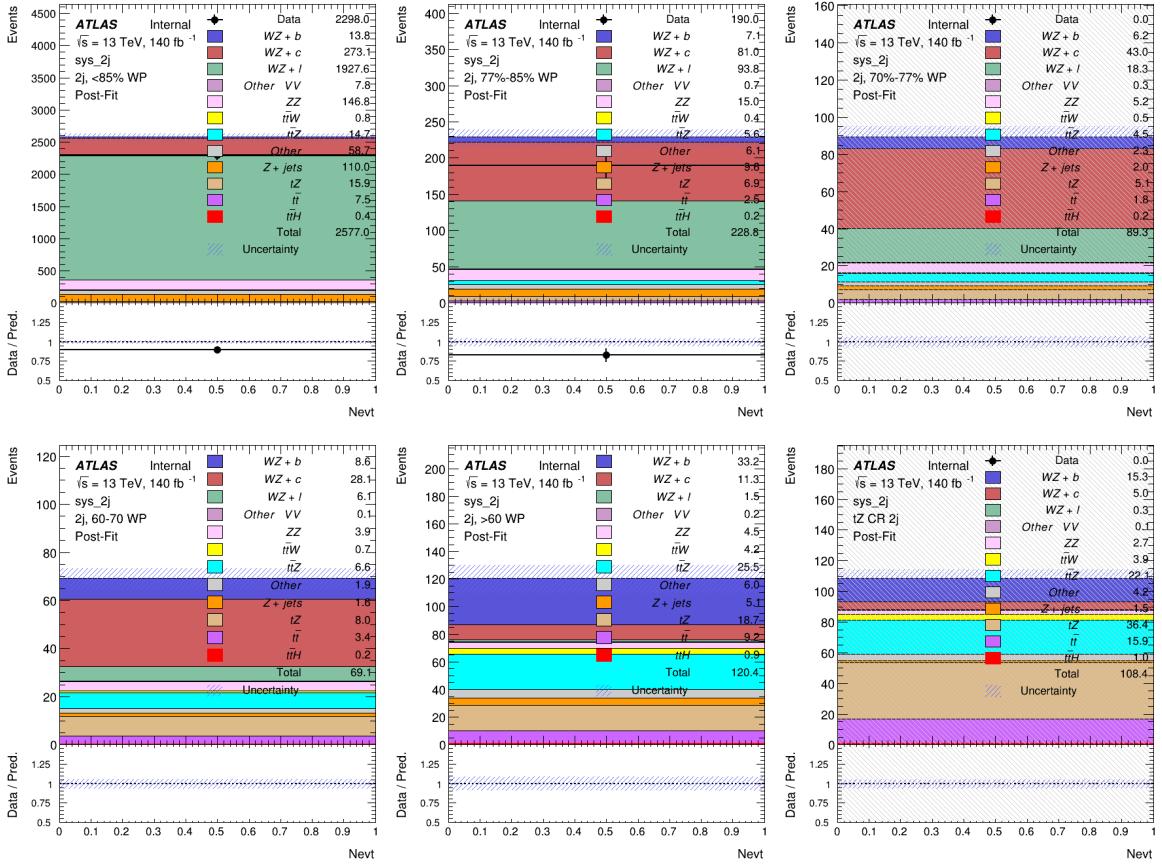


Figure 36: Data/MC results in each of the regions in the 2-jet fit after the fit has been performed.

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

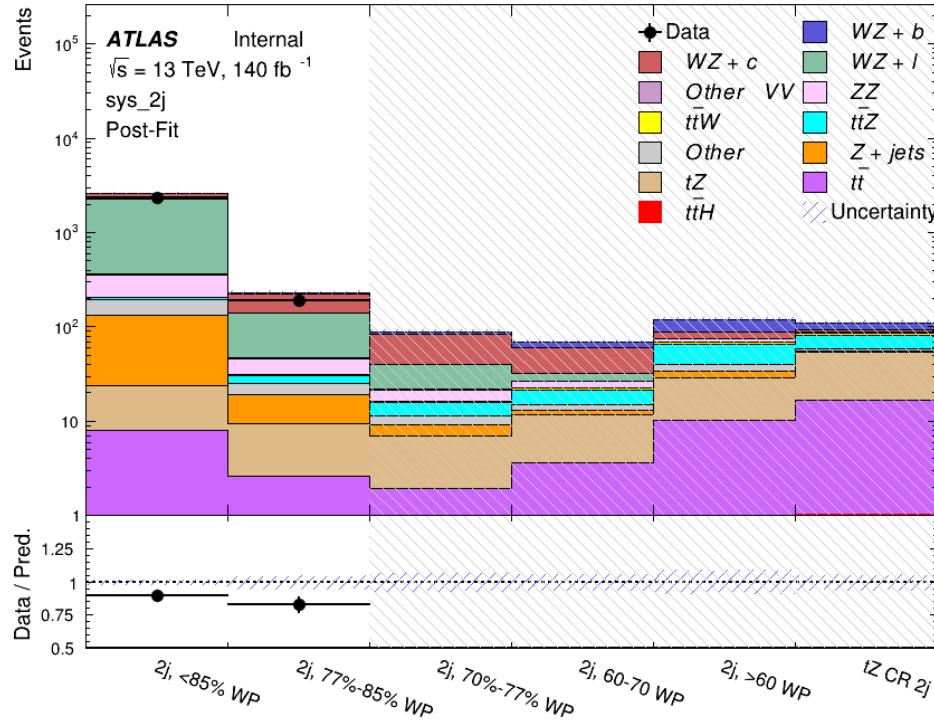


Figure 37: Post-fit summary of the fit over 2-jet regions.

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

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

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

⁴⁷⁷ The ranking and impact of those nuisance parameters with the largest contribution to the overall
⁴⁷⁸ uncertainty is shown in figure 38.

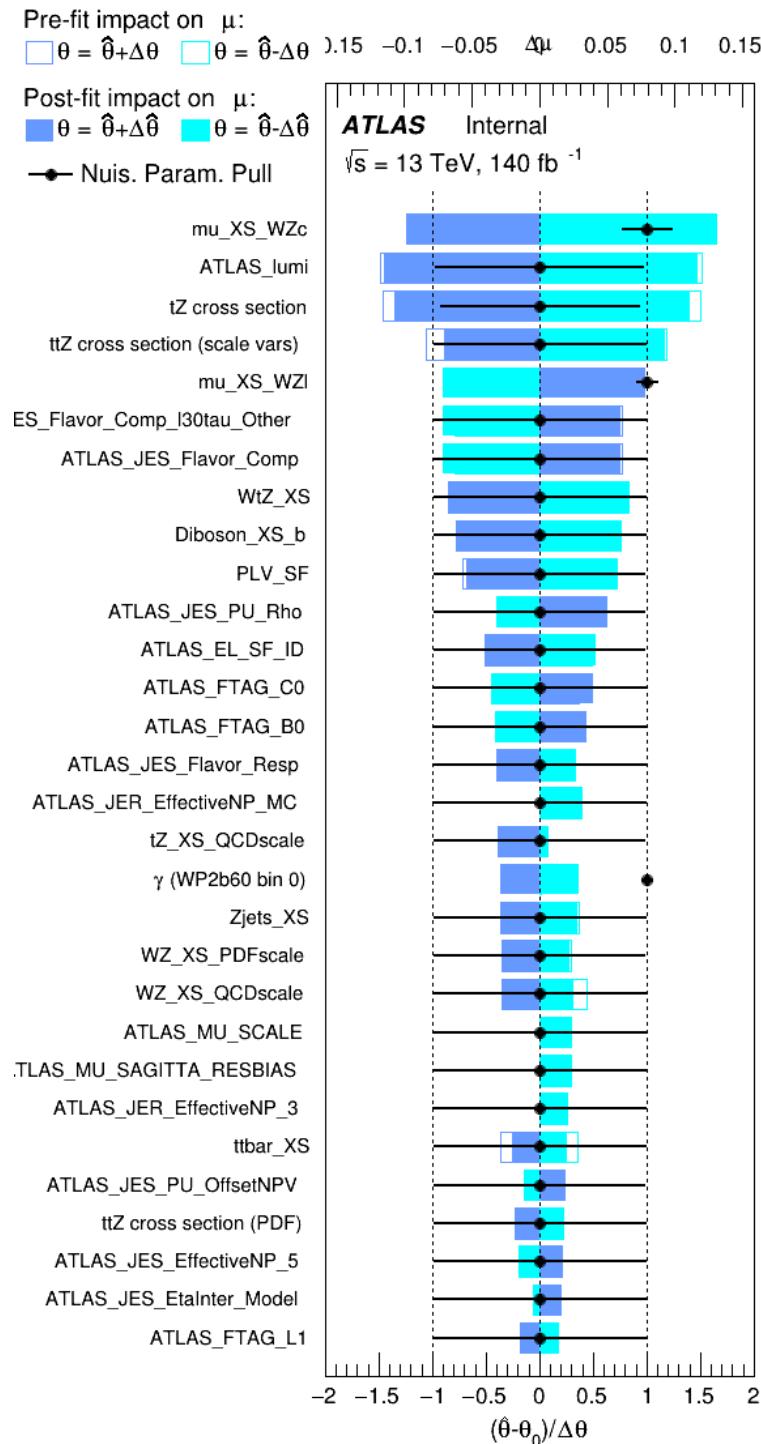


Figure 38: Impact of systematic uncertainties on the signal-strength of WZ + b in 2-jet events.

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

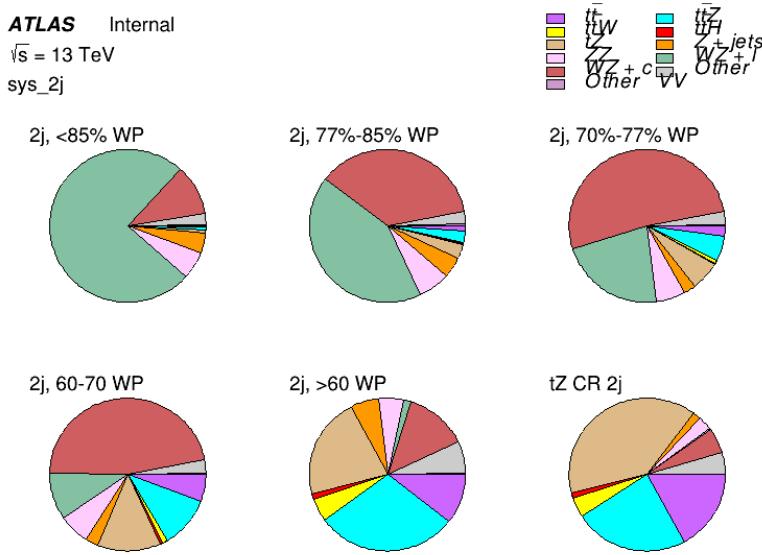


Figure 39: Post-fit background composition of the 2-jet fit regions.

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

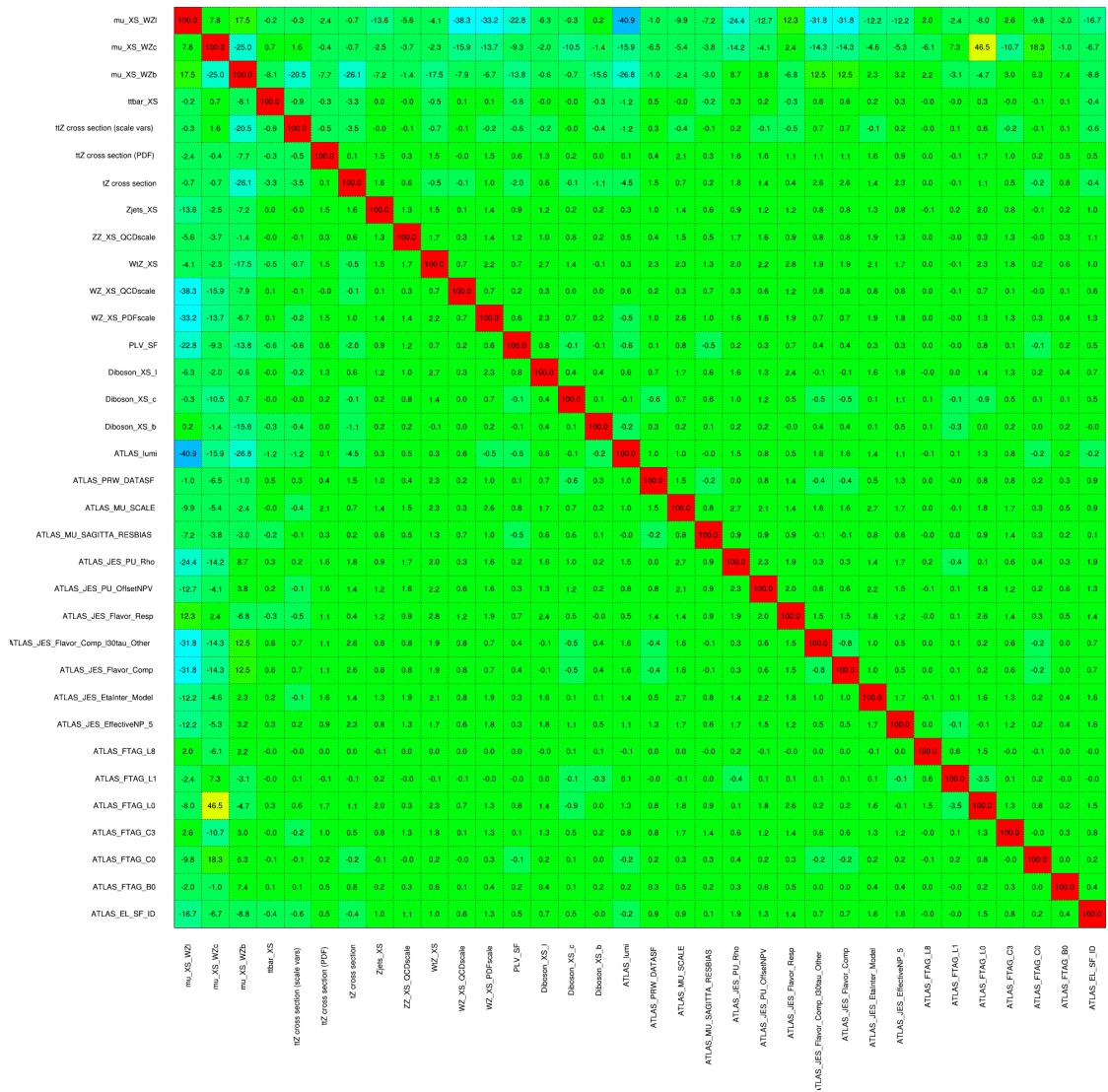


Figure 40: Correlations between nuisance parameters in the 2-jet fit

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

8.3 tZ Inclusive Results

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

489 9 Conclusion

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

493 **Appendices**

494 **tZ Interference Studies**

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

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

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

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

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

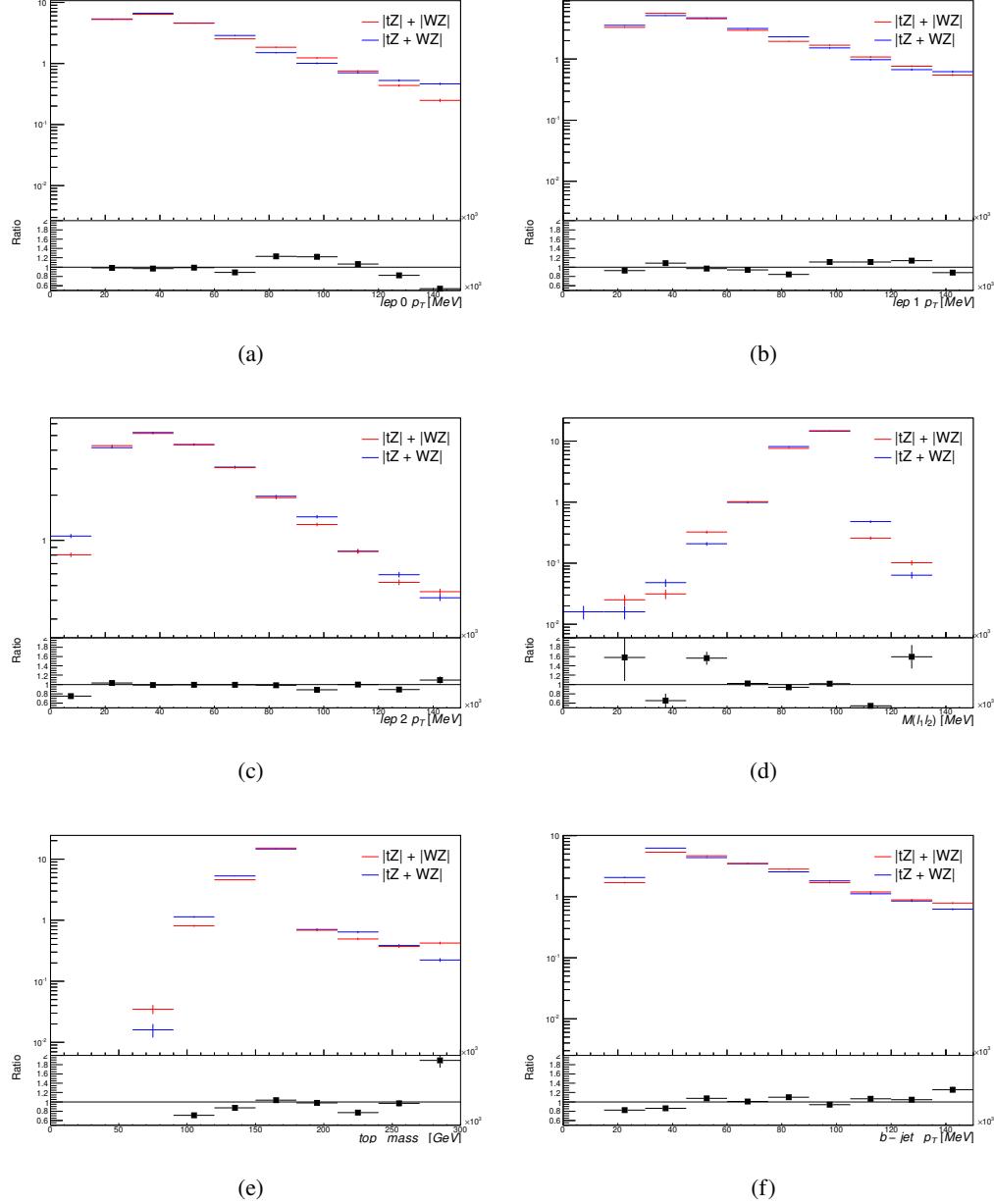


Figure 41: Comparisons between (a) the p_T of the lepton from the W, (b) and (c) show the p_T of the lepton from the Z, (d) the invariant mass of the Z-candidate, (e) the mass of the top candidate, and (f) the p_T of the b-jet, for WZ and tZ events generated with interference effects (blue) and without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

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

515 significantly impact the results.

516 **References**

- 517 [1] M. Aaboud et al. ‘Observation of electroweak $W^\pm Z$ boson pair production in association
 518 with two jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector’. In: *Phys.*
 519 *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:
 520 [1812.09740 \[hep-ex\]](https://arxiv.org/abs/1812.09740).
- 521 [2] T. Gleisberg et al. ‘Event generation with SHERPA 1.1’. In: *JHEP* 02 (2009), p. 007. doi:
 522 [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).
- 523 [3] ATLAS Collaboration. *Electron efficiency measurements with the ATLAS detector using*
 524 *the 2015 LHC proton–proton collision data*. ATLAS-CONF-2016-024. 2016. URL: <https://cds.cern.ch/record/2157687>.
- 526 [4] ATLAS Collaboration. ‘Measurement of the muon reconstruction performance of the
 527 ATLAS detector using 2011 and 2012 LHC proton–proton collision data’. In: *Eur. Phys.*
 528 *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).
- 530 [5] *Evidence for the associated production of the Higgs boson and a top quark pair with the*
 531 *ATLAS detector*. Tech. rep. ATLAS-CONF-2017-077. Geneva: CERN, Nov. 2017. URL:
 532 <https://cds.cern.ch/record/2291405>.
- 533 [6] ATLAS Collaboration. *Jet Calibration and Systematic Uncertainties for Jets Reconstructed*
 534 *in the ATLAS Detector at $\sqrt{s} = 13$ TeV*. ATL-PHYS-PUB-2015-015. 2015. URL:
 535 <https://cds.cern.ch/record/2037613>.
- 536 [7] ATLAS Collaboration. *Selection of jets produced in 13 TeV proton–proton collisions with*
 537 *the ATLAS detector*. ATLAS-CONF-2015-029. 2015. URL: <https://cds.cern.ch/record/2037702>.
- 539 [8] ATLAS Collaboration. ‘Performance of pile-up mitigation techniques for jets in pp col-
 540 lisions at $\sqrt{s} = 8$ TeV using the ATLAS detector’. In: *Eur. Phys. J. C* 76 (2016), p. 581.
 541 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).
- 542 [9] ATLAS Collaboration. *Performance of missing transverse momentum reconstruction with*
 543 *the ATLAS detector in the first proton–proton collisions at $\sqrt{s} = 13$ TeV*. ATL-PHYS-
 544 PUB-2015-027. 2015. URL: <https://cds.cern.ch/record/2037904>.
- 545 [10] P. S. A. Hoecker. ‘TMVA 4 Toolkit for Multivariate Data Analysis with ROOT’. In:
 546 *arXiv:physics/0703039* (2013).
- 547 [11] F. Cardillo et al. ‘Measurement of the fiducial and differential cross-section of a top quark
 548 pair in association with a Z boson at 13 TeV with the ATLAS detector’. In: ATL-COM-
 549 PHYS-2019-334 (Apr. 2019). URL: <https://cds.cern.ch/record/2672207>.
- 550 [12] ATLAS Collaboration. ‘Luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using
 551 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).

- 553 [13] G. Aad et al. ‘Jet energy resolution in proton-proton collisions at $\sqrt{s} = 7$ TeV recorded
554 in 2010 with the ATLAS detector’. In: *The European Physical Journal C* 73.3 (Mar.
555 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:
556 <https://doi.org/10.1140/epjc/s10052-013-2306-0>.
- 557 [14] A. Collaboration. ‘Performance of b -jet identification in the ATLAS experiment’. In:
558 *Journal of Instrumentation* 11.04 (2016), P04008. URL: <http://stacks.iop.org/1748-0221/11/i=04/a=P04008>.
559