



# ATLAS Note

GROUP-2017-XX

18th January 2021



Draft version 0.1

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  $\text{fb}^{-1}$  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,  $WZ \rightarrow l\nu ll$ . The  
8 cross-section of WZ + b-jets is measured to be  $X \pm X \pm X$ , while the cross-section of WZ +  
9 charm is measured as X, with a correlation of X between the two processes.

11	<b>Contents</b>	
12	<b>1 Changes and outstanding items</b>	<b>3</b>
13	1.1 Changelog	3
14	1.1.1 Changes relative to v2	3
15	1.1.2 Changes relative to v1	3
16	1.2 Outstanding Items	3
17	<b>2 Executive Summary</b>	<b>4</b>
18	<b>3 Introduction</b>	<b>4</b>
19	<b>4 Data and Monte Carlo Samples</b>	<b>5</b>
20	4.1 Data Samples	6
21	4.2 Monte Carlo Samples	6
22	<b>5 Object Reconstruction</b>	<b>7</b>
23	5.1 Trigger	7
24	5.2 Light leptons	7
25	5.3 Jets	8
26	5.4 B-tagged Jets	10
27	5.5 Missing transverse energy	11
28	<b>6 Event Selection and Signal Region Definitions</b>	<b>11</b>
29	6.1 Event Preselection	11
30	6.2 Fit Regions	15
31	6.3 Non-Prompt Lepton Estimation	27
32	6.3.1 $t\bar{t}$ Validation	27
33	6.3.2 Z+jets Validation	30
34	<b>7 <math>tZ</math> Interference Studies and Separation Multivariate Analysis</b>	<b>32</b>
35	7.1 Interference Studies	32
36	7.2 Top Mass Reconstruction	33
37	7.3 $tZ$ BDT	34
38	<b>8 Systematic Uncertainties</b>	<b>37</b>
39	<b>9 Results</b>	<b>41</b>
40	9.1 1-jet Fit Results	42
41	9.2 2-jet Fit Results	48
42	<b>10 Conclusion</b>	<b>52</b>

## 1 Changes and outstanding items

### 1.1 Changelog

This is version 3

#### 1.1.1 Changes relative to v2

- Added alternate VBS samples to include missing b-jet diagrams
- Included a section on tZ interference effects, [7.1](#).
- Updated to reflect changes for 2018, including the move to PFlow jets, DL1r, updated trigger, and updated AnalysisBase version (now 21.2.127)
- Revised fit regions, using separate 1-jet and 2-jet fits, with all 2-j regions included
- updated plots for tZ BDT, added details about the model
- Included truth jet information

#### 1.1.2 Changes relative to v1

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

### 1.2 Outstanding Items

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

## 2 Executive Summary

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

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

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

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

Section 4 details the data and Monte Carlo (MC) samples used in the analysis. The reconstruction of various physics objects is described in section 5. Section 6 describes the event selection applied to these samples, along the definitions of the various regions used in the fit. The multivariate analysis techniques used to separate the tZ background from WZ + heavy flavor are described in section 7. Section 8 describes the various sources of systematic uncertainties considered in the fit. Finally, the results of the analysis are summarized in section 9, followed by a brief conclusion in section 10.

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

## 3 Introduction

The production of WZ in association with a heavy flavor jet represents an important background for many major analyses. This includes any process with leptons and b-jets in the final state,

such as  $t\bar{t}H$ ,  $t\bar{t}W$ , and  $t\bar{t}Z$ . While precise measurements have been made of  $WZ$  production [1],  $WZ$  + heavy flavor remains poorly understood. This is largely because the QCD processes involved in the production of the  $b$ -jet make it difficult to simulate accurately. This introduces a large uncertainty for analyses that include this process as a background.

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

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

Section 4 details the data and Monte Carlo (MC) samples used in the analysis. The reconstruction of various physics objects is described in section 5. Section 6 describes the event selection applied to these samples, along the definitions of the various regions used in the fit. The multivariate analysis techniques used to separate the  $tZ$  background from  $WZ$  + heavy flavor are described in section 7. Section 8 describes the various sources of systematic uncertainties considered in the fit. Finally, the results of the analysis are summarized in section 9, followed by a brief conclusion in section 10.

**The current state of the analysis shows blinded results for the full 2018 dataset. Regions containing >5%  $WZ+b$  events are blinded, and results are from Asimov, MC only fits.**

## 4 Data and Monte Carlo Samples

Both data and Monte Carlo samples used in this analysis were prepared in the  $xAOD$  format, which was used to produce a  $DxAOD$  sample in the HIGG8D1 derivation framework. The HIGG8D1 framework is designed for the  $t\bar{t}H$  multi-lepton analysis, which targets events with multiple 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  $p_T > 15 \text{ GeV}$  and subleading lepton  $p_T > 5 \text{ GeV}$
- at least one light lepton with  $p_T > 15 \text{ GeV}$  within a range  $|\eta| < 2.6$ , and at least two hadronic taus with  $p_T > 15 \text{ GeV}$ .

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

## 4.1 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  $139 \text{ fb}^{-1}$ . 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.

## 4.2 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
t $\bar{t}$ W	MG5_AMC	NLO	PYTHIA 8	NNPDF 3.0 NLO
t $\bar{t}$ (Z/ $\gamma^* \rightarrow \ell\ell$ )	(SHERPA 2.1.1)	(LO multileg)	(SHERPA)	(NNPDF 3.0 NLO)
t $\bar{t}$ H	MG5_AMC	NLO	PYTHIA 8	NNPDF 3.0 NLO
	MG5_AMC	NLO	PYTHIA 8	NNPDF 3.0 NLO [Ball:2014uwa]
	(MG5_AMC)	(NLO)	(HERWIG++)	(CT10 [ct10])
tHqb	MG5_AMC	LO	PYTHIA 8	CT10
tHW	MG5_AMC	NLO	HERWIG++	CT10
	(SHERPA 2.1.1)	(LO multileg)	(SHERPA)	(NNPDF 3.0 NLO)
tWZ	MG5_AMC	NLO	PYTHIA 8	NNPDF 2.3 LO
t $\bar{t}$ t, t $\bar{t}$ t $\bar{t}$	MG5_AMC	LO	PYTHIA 8	NNPDF 2.3 LO
t $\bar{t}$ W $^+W^-$	MG5_AMC	LO	PYTHIA 8	NNPDF 2.3 LO
t $\bar{t}$	POWHEG-BOX v2 [powheggt]	NLO	PYTHIA 8	NNPDF 3.0 NLO
t $\bar{t}\gamma$	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 $\rightarrow \ell^+\ell^-$	SHERPA 2.2.1	MEPS NLO	SHERPA	NNPDF 3.0 NLO

## 5 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 6.

### 5.1 Trigger

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

### 5.2 Light leptons

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

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

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

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

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

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

### 5.3 Jets

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



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

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

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

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

183 and only jets satisfying  $p_T > 25$  GeV and  $|\eta| < 2.5$  are used in this analysis. For jets with  
 184  $p_T < 60$  GeV and  $|\eta| < 2.4$ , a jet-track association algorithm is used to confirm that the jet  
 185 originates from the selected primary vertex, in order to reject jets arising from pileup collisions  
 186 [8].

## 5.4 B-tagged Jets

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

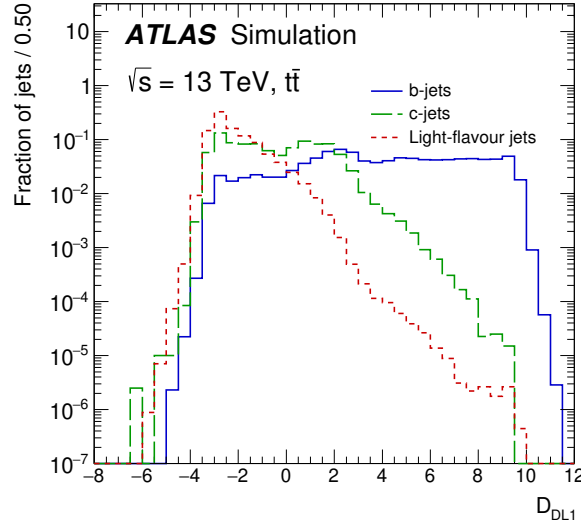


Figure 1: Output distribution of the DL1r algorithm for b-jets, charm jets, and light jets

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

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

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

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

## 5.5 Missing transverse energy

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

## 6 Event Selection and Signal Region Definitions

Event are required to pass a preselection described in section 6.1 and summarized in table 6. Those that pass this preselection are divided into various fit regions described in section 6.2, based on the number of jets in the event, and the b-tag score of those jets.

### 6.1 Event Preselection

Events are required to include exactly three reconstructed light leptons passing the requirement described in 5.2, which have a total charge of  $\pm 1$ . As the opposite sign lepton is found to be prompt the vast majority of the time [5], it is required to be loose and isolated, as defined though the standard `isolationFixedCutLoose` working point supported by combined performance groups. The same sign leptons are required to be very tight, as per the recommended `isolationFixedCutTight`.

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

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

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

Events are required to have one or two reconstructed jets passing the selection described in section 5.3. Events with more than two jets are rejected in order to reduce the contribution of backgrounds such as  $t\bar{t}Z$  and  $t\bar{t}W$ , which tend to have higher jet multiplicity.

---



---

Event Selection

---

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

---



---

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

235 The event yields in the preselection region for both data and Monte Carlo are summarized in  
236 table 6.1, which shows good agreement between data and Monte Carlo, and demonstrates that  
237 this region consists primarily of WZ events. The WZ events are split into WZ + b, WZ + c, and  
238 WZ + l based on the truth flavor of the heaviest associated jet in the event. Specifically, this  
239 determination is made based on the HadronConeExclTruthLabelID of the jet. That is, WZ +  
240 l events contain no charm and b jets at truth level, WZ + c contain at least one truth charm and  
241 no b-jets, and WZ + b contains at least one truth b-jet.

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

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

Here Other VV represents diboson processes other than WZ, and consists predominantly of  $ZZ \rightarrow llll$  events where one of the leptons is not reconstructed.

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

## WZ Fit Region - Inclusive

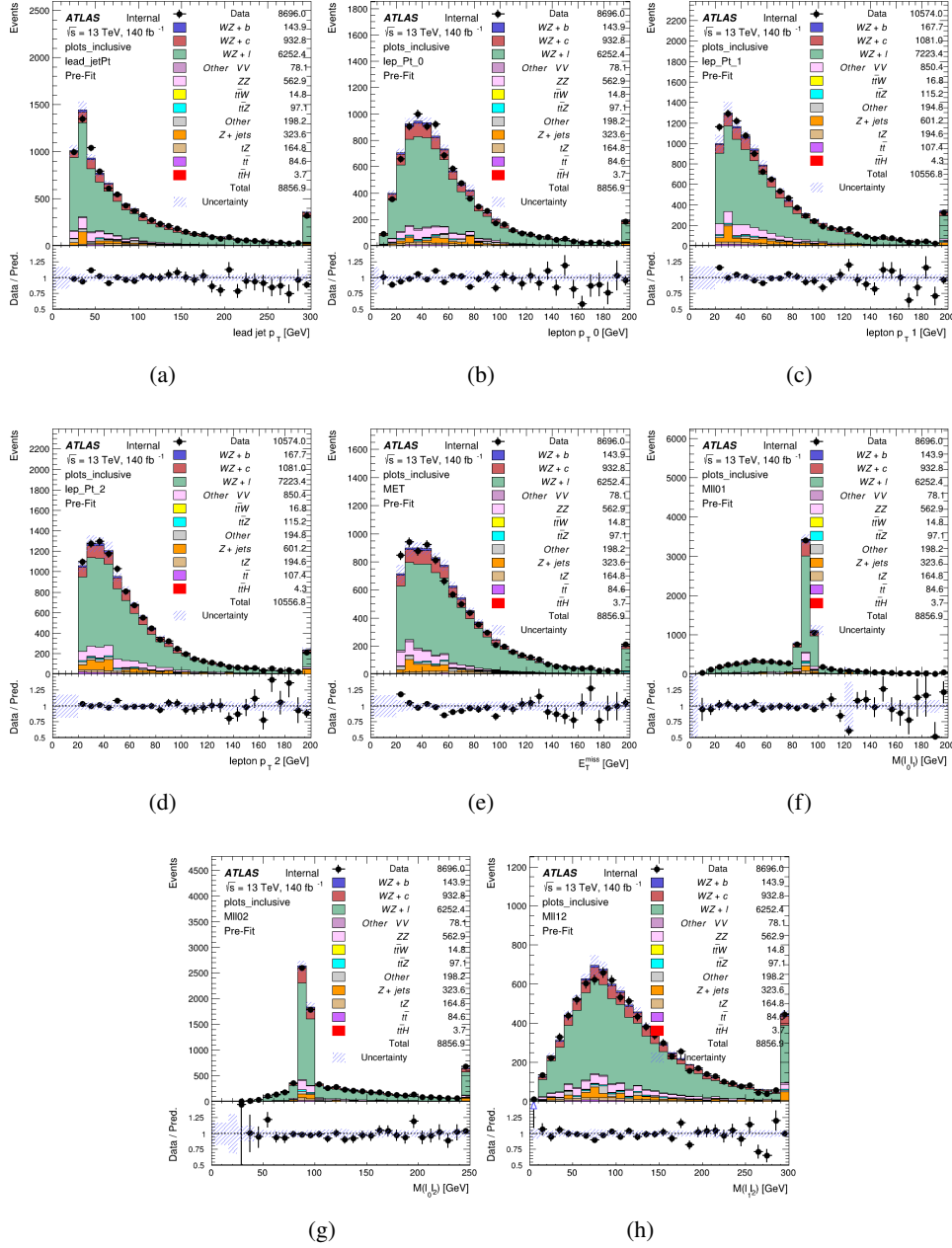


Figure 2: Comparisons between data and MC distributions in the preselection region for the  $p_T$  of (a) the leading jet, (b) lepton 0, (c) lepton 1, (d) lepton 2, (e) the missing transverse energy, and (f) the invariant mass of leptons 0 and 1, (g) the invariant mass of leptons 0 and 2, and (h) the invariant mass of leptons 1 and 2.

## 6.2 Fit Regions

Once preselection has been applied, the remaining events are categorized into one of twelve orthogonal regions. The regions used in the fit are summarized in table 8.

Table 8: A list of the regions used in the fit and the selection used for each.

Region	Selection
1j, <85%	$N_{\text{jets}} = 1, n\text{Jets\_DL1r\_85} = 0$
1j, 85%-77%	$N_{\text{jets}} = 1, n\text{Jets\_DL1r\_85} = 1, n\text{Jets\_DL1r\_77}=0$
1j, 77%-70%	$N_{\text{jets}} = 1, n\text{Jets\_DL1r\_77} = 1, n\text{Jets\_DL1r\_70}=0$
1j, 70%-60%	$N_{\text{jets}} = 1, n\text{Jets\_DL1r\_70} = 1, n\text{Jets\_DL1r\_60}=0$
1j, >60%	$N_{\text{jets}} = 1, n\text{Jets\_DL1r\_60} = 1, tZ \text{ BDT} > 0.525$
1j tZ CR	$N_{\text{jets}} = 1, n\text{Jets\_DL1r\_60} = 1, tZ \text{ BDT} < 0.525$
2j, <85%	$N_{\text{jets}} = 2, n\text{Jets\_DL1r\_85} = 0$
2j, 85%-77%	$N_{\text{jets}} = 2, n\text{Jets\_DL1r\_85} \geq 1, n\text{Jets\_DL1r\_77}=0$
2j, 77%-70%	$N_{\text{jets}} = 2, n\text{Jets\_DL1r\_77} \geq 1, n\text{Jets\_DL1r\_70}=0$
2j, 70%-60%	$N_{\text{jets}} = 2, n\text{Jets\_DL1r\_70} \geq 1, n\text{Jets\_DL1r\_60}=0$
2j, >60%	$N_{\text{jets}} = 2, n\text{Jets\_DL1r\_60} \geq 1, tZ \text{ BDT} > 0.525$
2j tZ CR	$N_{\text{jets}} = 2, n\text{Jets\_DL1r\_60} \geq 1, tZ \text{ BDT} < 0.525$

The working points discussed in section 5.4 are used to separate events into fit regions based on the highest working point reached by a jet in each event. Because the background composition differs significantly based on the number of b-jets, events are further subdivided into 1 jet and 2 jet regions in order to minimize the impact of background uncertainties.

An additional tZ control region is created based on the BDT described in section 7. The region with 1-jet passing the 60% working point is split in two - a signal enriched region of events with a BDT score greater than 0.03, and a tZ control region including events with less than 0.03. This cutoff is arrived at by performing an Asimov fit with a variety of cutoffs, and selecting the value that produces the highest significance for the measurement of  $WZ + b$ .

The modeling in each region is validated by comparing data and MC predictions for various kinematic distributions. These plot are shown in figures 3-14.

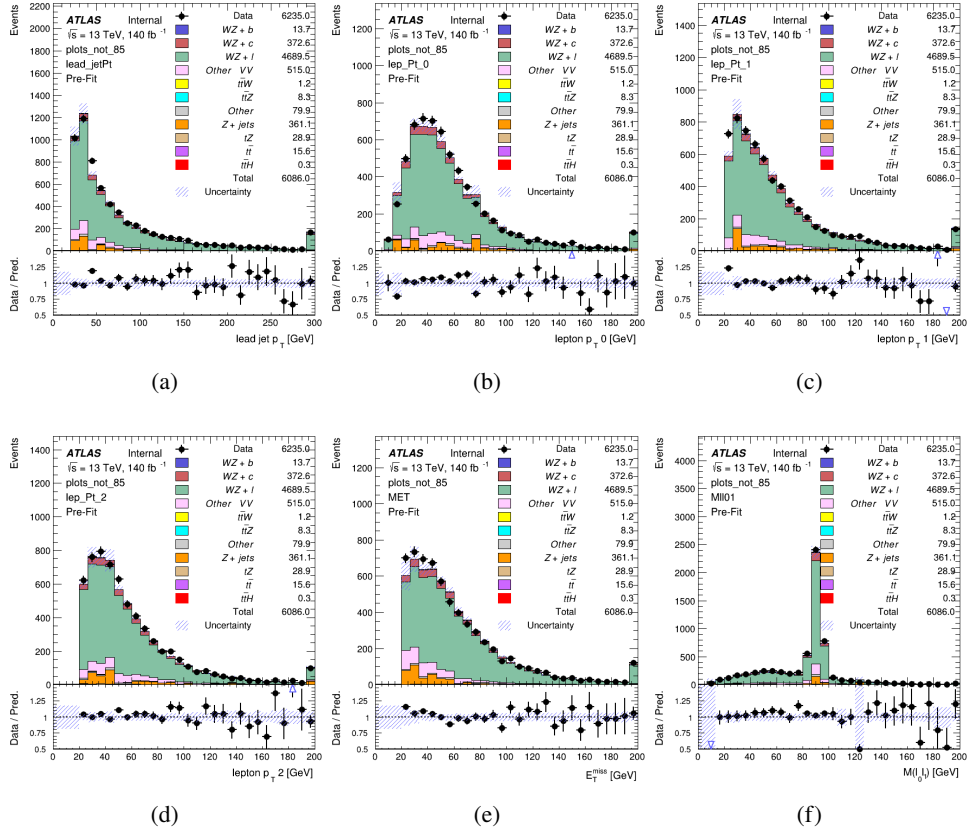
WZ Fit Region -  $1j < 85\%$  WP

Figure 3: Comparisons between the data and MC distributions in the preselection region for the  $p_T$  of (a) the leading jet, (b) lepton 0, (c) lepton 1, (d) lepton 2, (e) the missing transverse energy, and (f) the invariant mass of lepton 0 and 1.



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

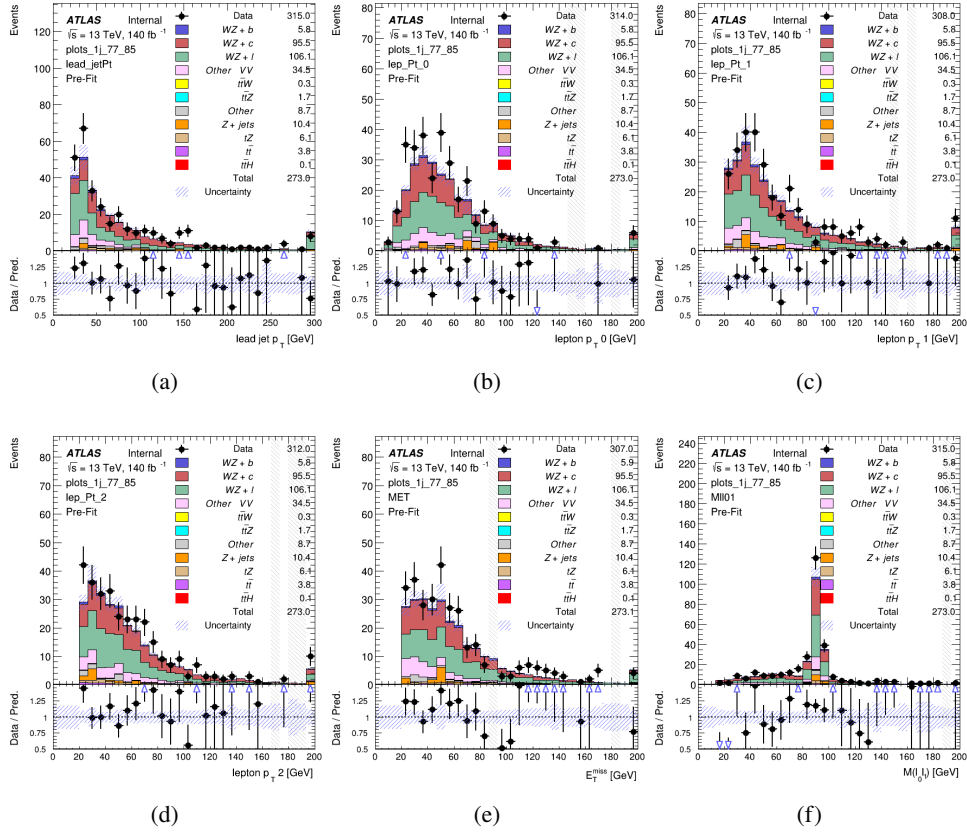


Figure 4: Comparisons between the data and MC distributions in the preselection region for the  $p_T$  of (a) the leading jet, (b) lepton 0, (c) lepton 1, (d) lepton 2, (e) the missing transverse energy, and (f) the invariant mass of lepton 0 and 1.

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

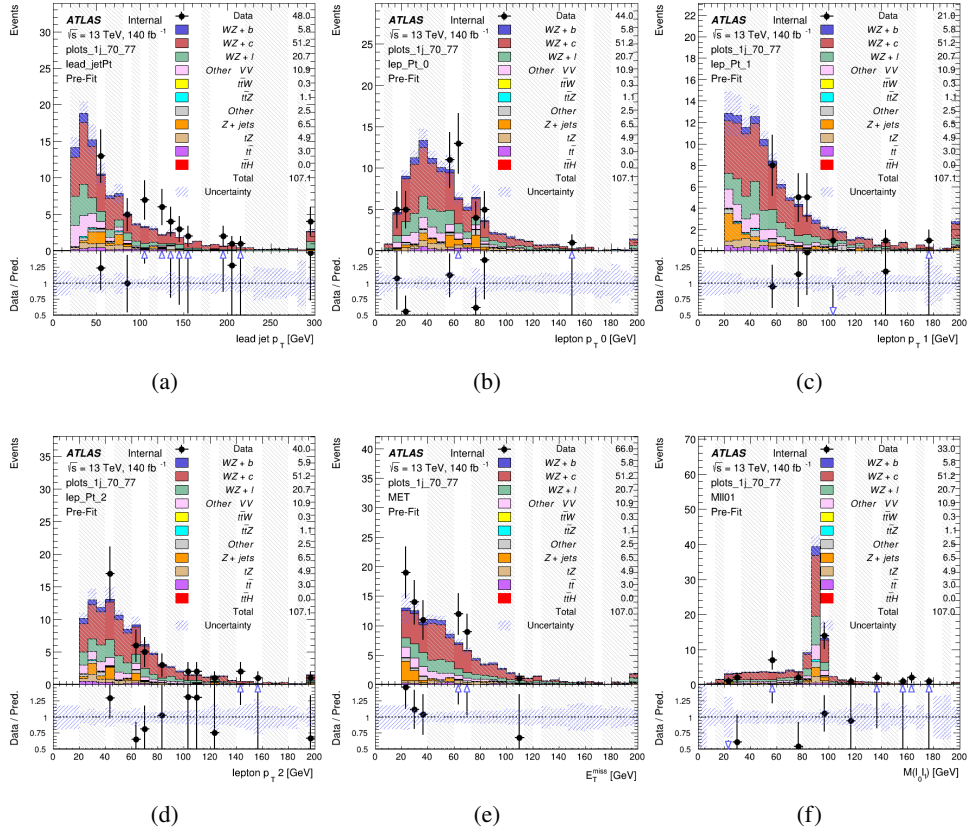


Figure 5: Comparisons between the data and MC distributions in the preselection region for the  $p_T$  of (a) the leading jet, (b) lepton 0, (c) lepton 1, (d) lepton 2, (e) the missing transverse energy, and (f) the invariant mass of lepton 0 and 1.

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

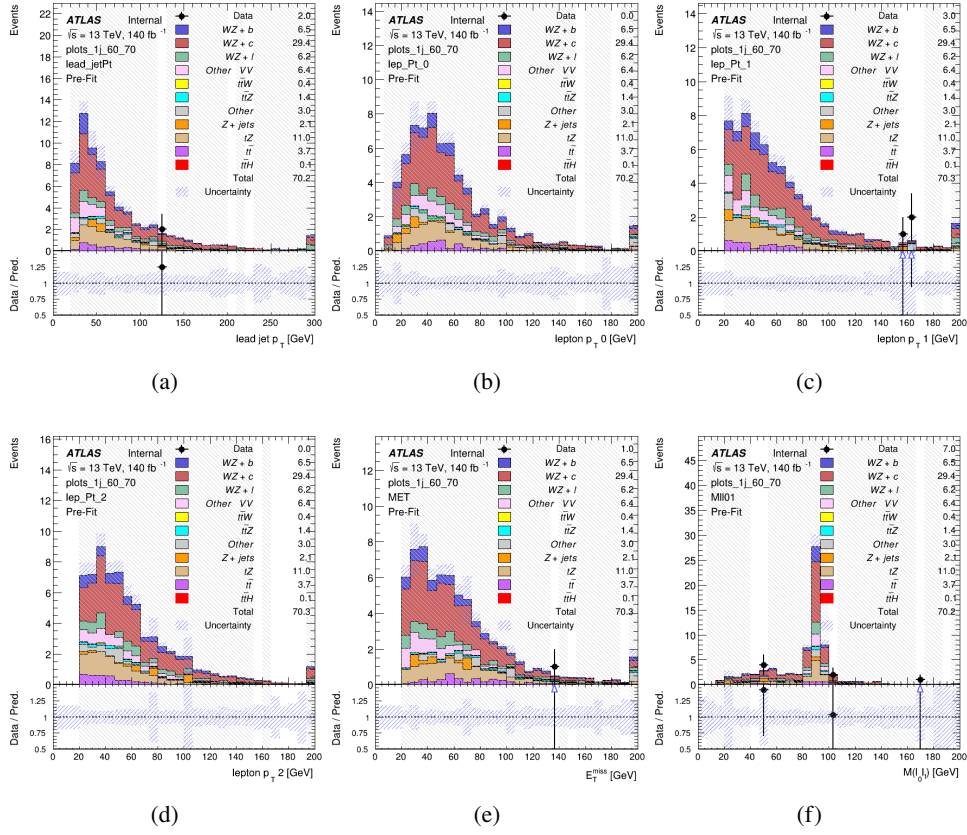


Figure 6: Comparisons between the data and MC distributions in the preselection region for the  $p_T$  of (a) the leading jet, (b) lepton 0, (c) lepton 1, (d) lepton 2, (e) the missing transverse energy, and (f) the invariant mass of lepton 0 and 1.

## WZ Fit Region - 1j 60% WP

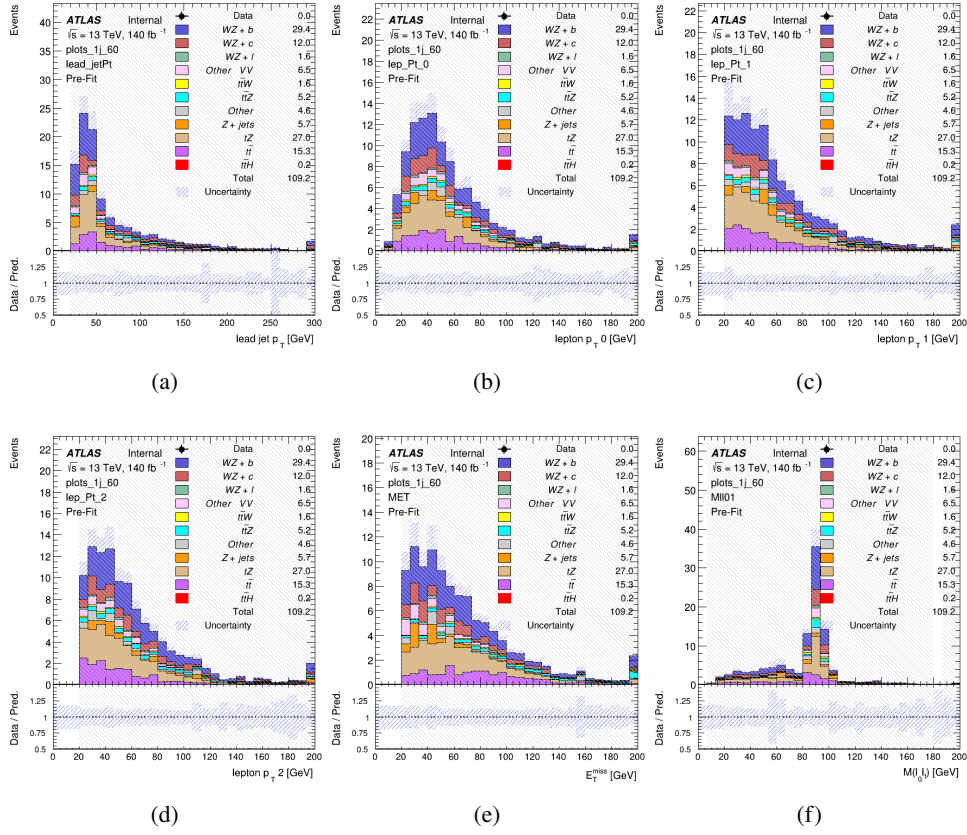


Figure 7: Comparisons between the data and MC distributions in the preselection region for the  $p_T$  of (a) the leading jet, (b) lepton 0, (c) lepton 1, (d) lepton 2, (e) the missing transverse energy, and (f) the invariant mass of lepton 0 and 1.

## WZ Fit Region - tZ-CR

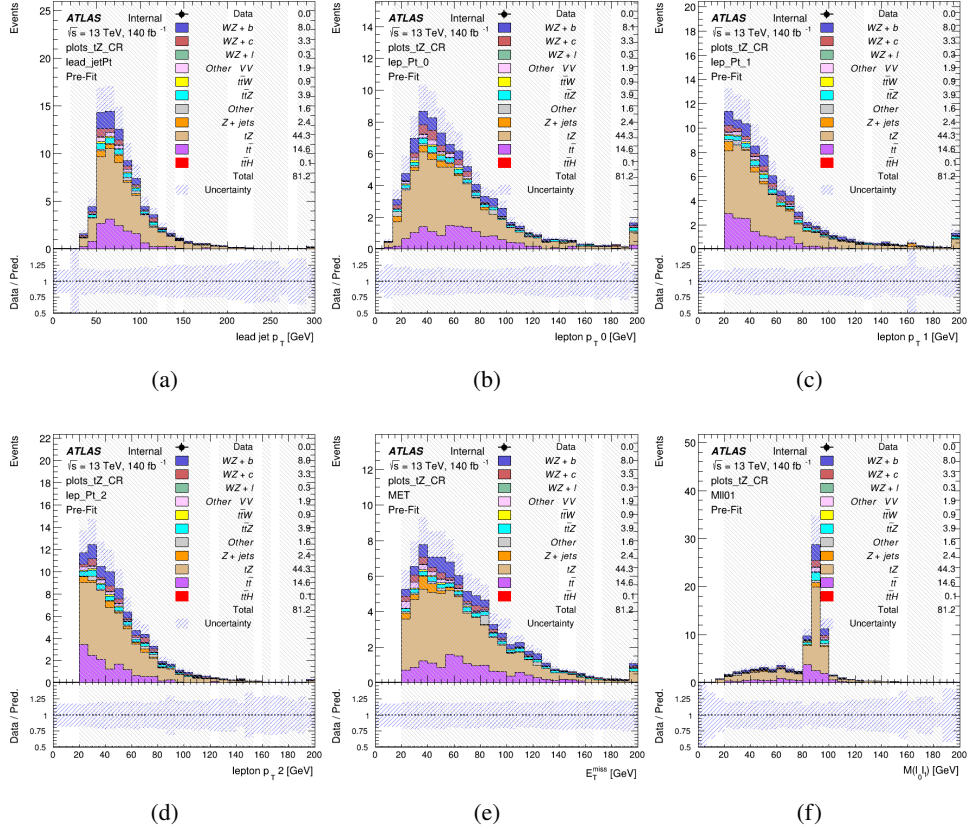


Figure 8: Comparisons between the data and MC distributions in the preselection region for the  $p_T$  of (a) the leading jet, (b) lepton 0, (c) lepton 1, (d) lepton 2, (e) the missing transverse energy, and (f) the invariant mass of lepton 0 and 1.

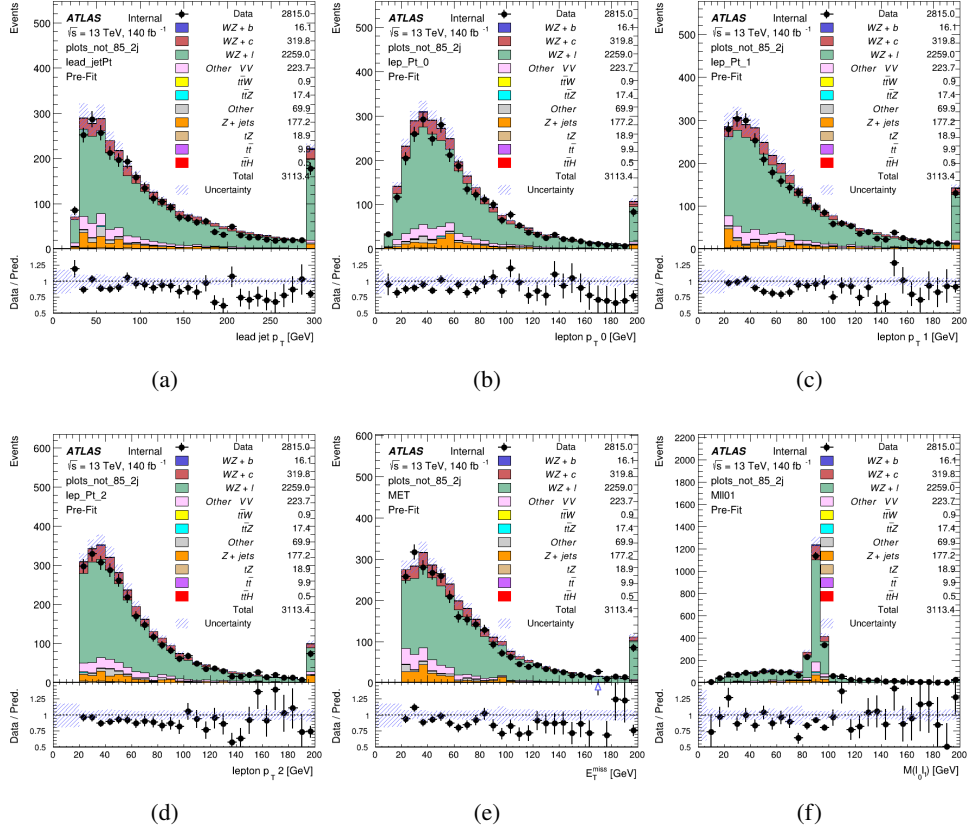
WZ Fit Region -  $2j < 85\%$  WP

Figure 9: Comparisons between the data and MC distributions in the preselection region for the  $p_T$  of (a) the leading jet, (b) lepton 0, (c) lepton 1, (d) lepton 2, (e) the missing transverse energy, and (f) the invariant mass of lepton 0 and 1.

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

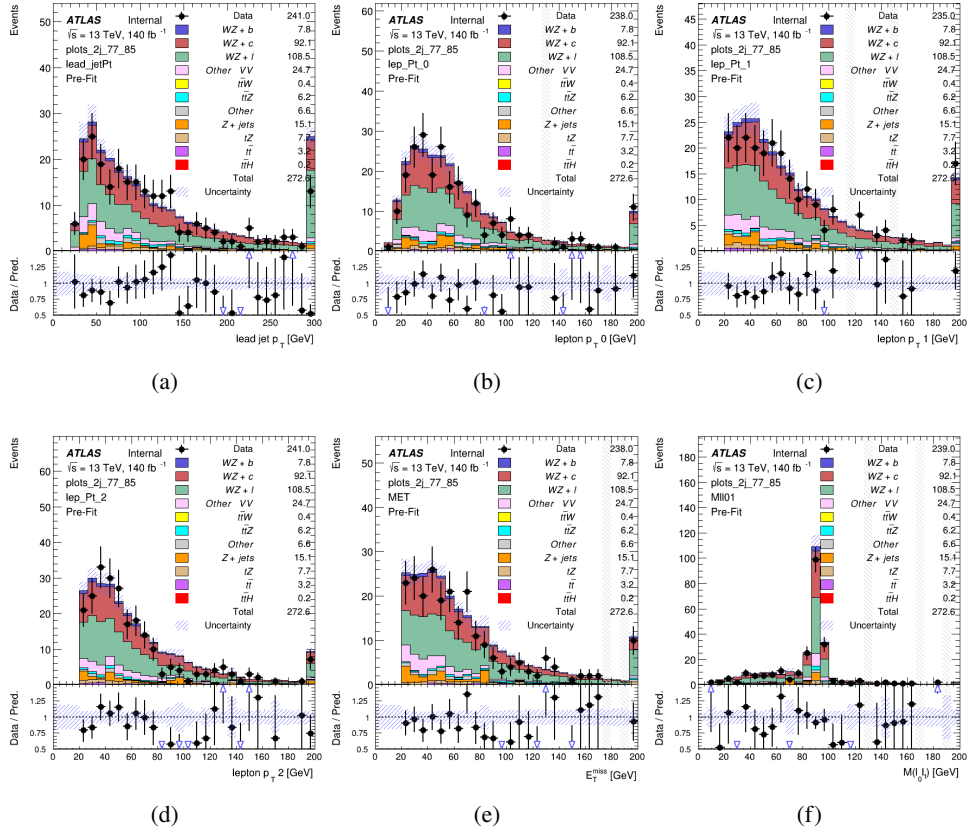


Figure 10: Comparisons between the data and MC distributions in the preselection region for the  $p_T$  of (a) the leading jet, (b) lepton 0, (c) lepton 1, (d) lepton 2, (e) the missing transverse energy, and (f) the invariant mass of lepton 0 and 1.



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

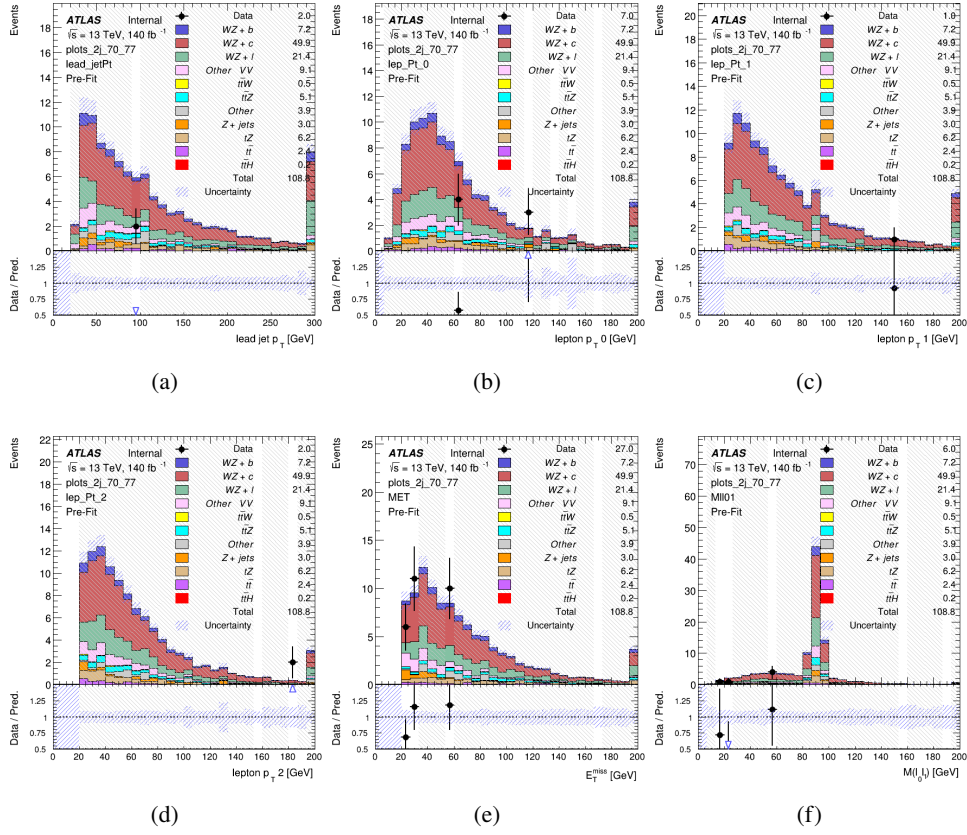


Figure 11: Comparisons between the data and MC distributions in the preselection region for the  $p_T$  of (a) the leading jet, (b) lepton 0, (c) lepton 1, (d) lepton 2, (e) the missing transverse energy, and (f) the invariant mass of lepton 0 and 1.



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

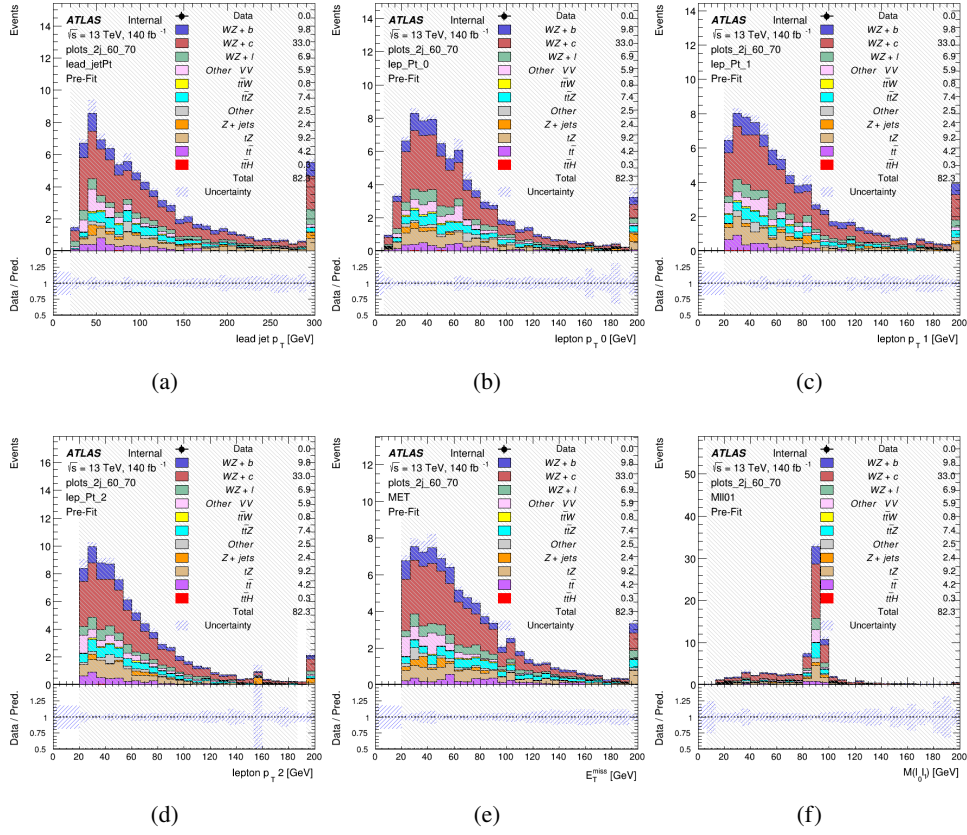


Figure 12: Comparisons between the data and MC distributions in the preselection region for the  $p_T$  of (a) the leading jet, (b) lepton 0, (c) lepton 1, (d) lepton 2, (e) the missing transverse energy, and (f) the invariant mass of lepton 0 and 1.

## WZ Fit Region - 2j 60% WP

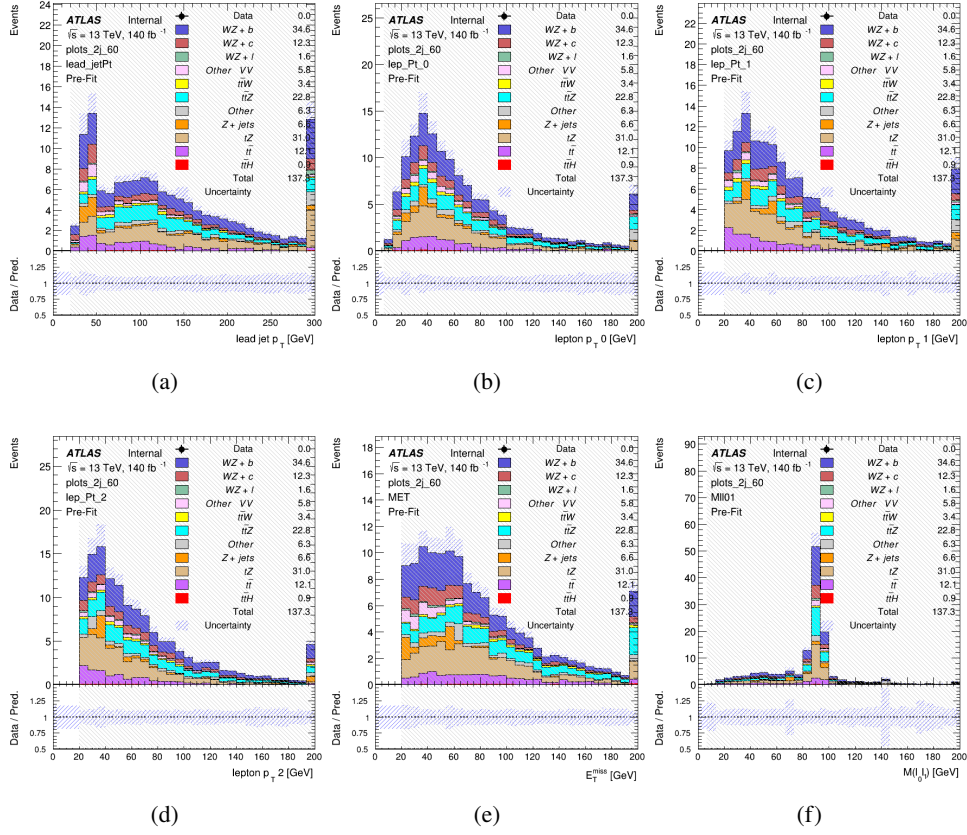


Figure 13: Comparisons between the data and MC distributions in the preselection region for the  $p_T$  of (a) the leading jet, (b) lepton 0, (c) lepton 1, (d) lepton 2, (e) the missing transverse energy, and (f) the invariant mass of lepton 0 and 1.

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

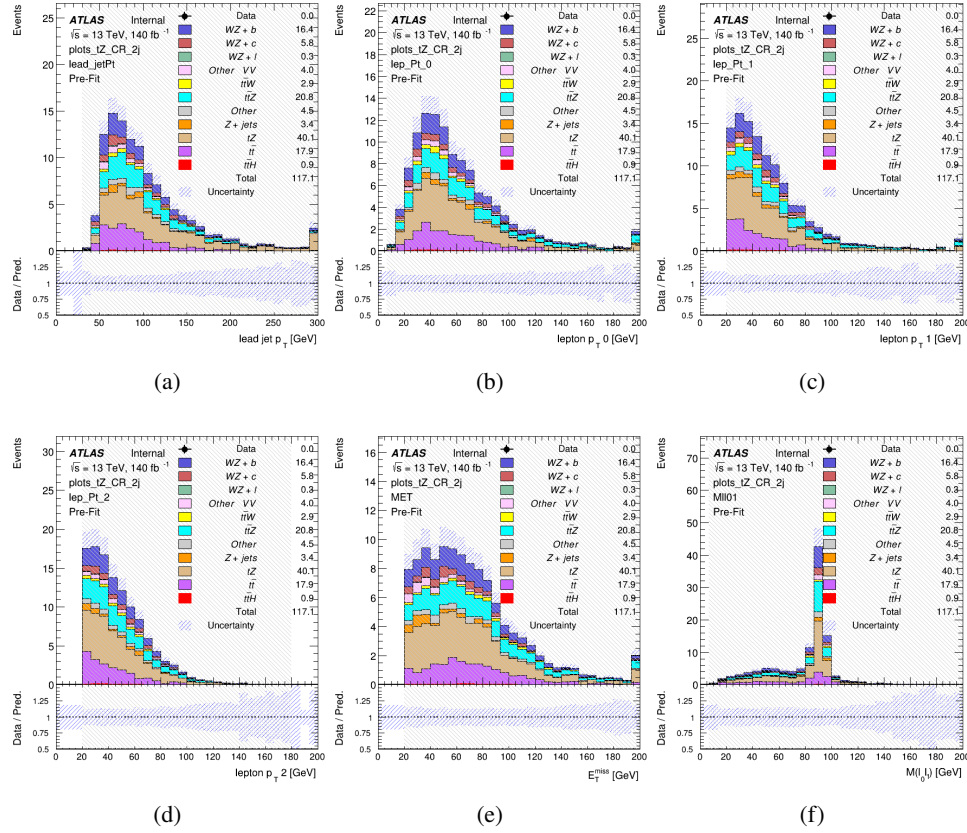


Figure 14: Comparisons between the data and MC distributions in the preselection region for the  $p_T$  of (a) the leading jet, (b) lepton 0, (c) lepton 1, (d) lepton 2, (e) the missing transverse energy, and (f) the invariant mass of lepton 0 and 1.

### 6.3 Non-Prompt Lepton Estimation

Two processes act as sources of non-prompt leptons appear in the analysis:  $t\bar{t}$  and  $Z$ +jet production both produce two prompt leptons, and each contribute to the 3l region when an additional non-prompt lepton appears in the event. The contribution of these processes is estimated with Monte Carlo simulations, which are validated using enriched validation regions.

#### 6.3.1 $t\bar{t}$ 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

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

271 The  $t\bar{t}$  validation region is similar to the preselection region - three leptons meeting the criteria  
272 described in section 6 are required, and the requirements on  $E_T^{\text{miss}}$  remain the same. However,  
273 the selection requiring a lepton pair form a Z-candidate are reversed. Events where the invariant  
274 mass of any two opposite sign, same flavor leptons falls within 10 GeV of 91.2 GeV are rejected.  
275 This ensures the  $t\bar{t}$  validation region is orthogonal to the preselection region.

276 Further, because the jet multiplicity of  $t\bar{t}$  events tends to be higher than WZ, the number of jets  
277 in each event is required to be greater than 1. As b-jets are almost invariably produced from top  
278 decays, at least one b-tagged jet passing the 70% DL1r WP in each event is required.

279 Various kinematic plots of this region are shown below. The general agreement between data and  
280 MC in each of these suggests that the non-prompt contribution of  $t\bar{t}$  is well modeled by Monte  
281 Carlo.

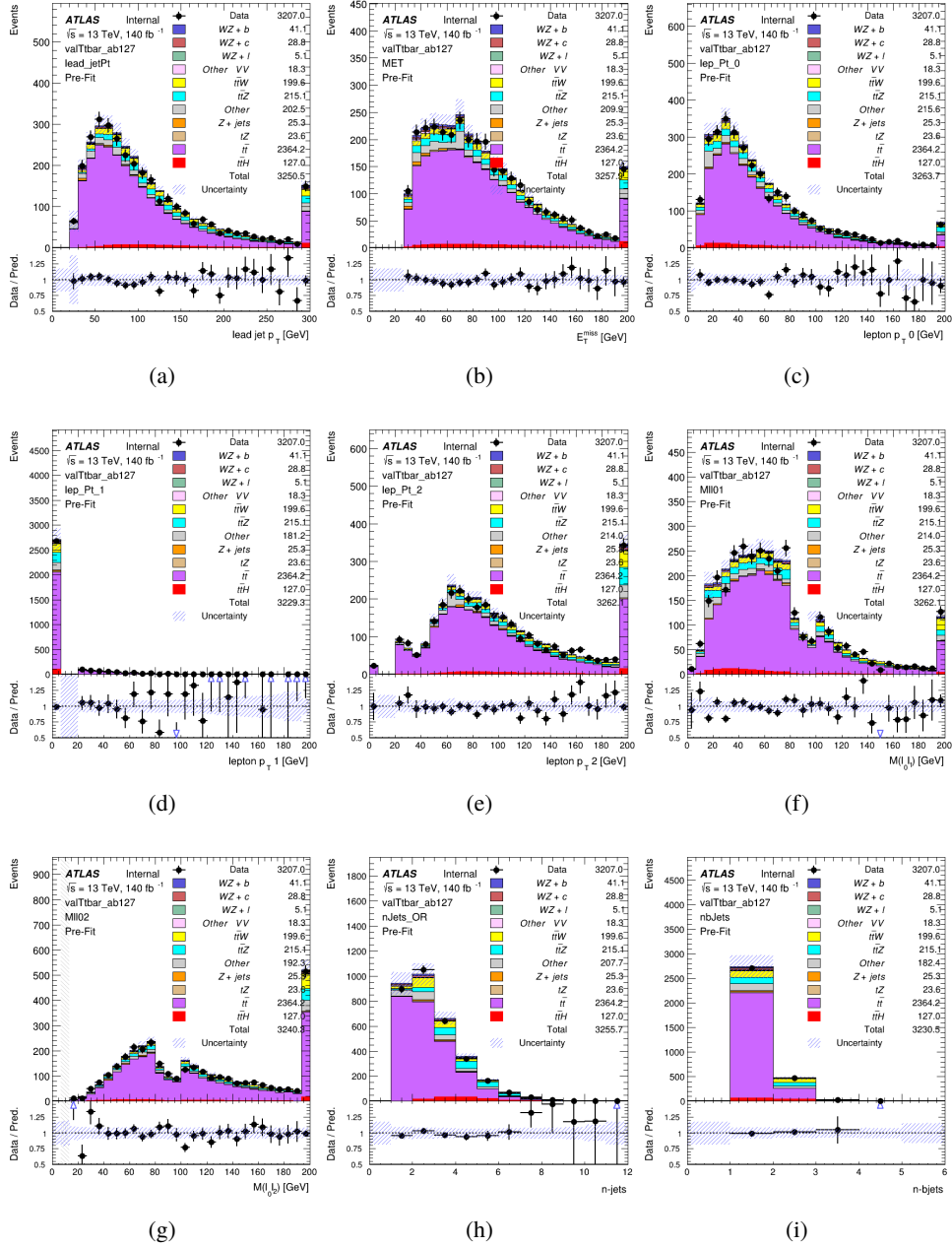


Figure 15: Comparisons between the data and MC distributions in the  $t\bar{t}$  validation region for (a) the  $p_T$  of the leading jet, (b) the missing transverse energy, (c) the  $p_T$  of lepton 0, (d)  $p_T$  of lepton 1, (e)  $p_T$  of lepton 2, (f) the invariant mass of leptons 0 and 1, (g) the invariant mass of leptons 0 and 2, (h) the number of jets, (i) the number of b-tagged jets.

### 282 6.3.2 Z+jets Validation

283 Similar to  $t\bar{t}$ , a non-prompt Z+jets validation region is produced in order to validate the MC  
284 predictions. The lepton requirements remain the same as the preselection region. Because no  
285 neutrinos are present for this process, the  $E_T^{\text{miss}}$  cut is reversed, requiring  $E_T^{\text{miss}} < 30$  GeV. This  
286 also ensures this validation region is orthogonal to the preselection region. Further, the number  
287 of jets in each event is required to be greater than or equal to one.

288 Various kinematic plots of this region are shown below. The general agreement between data  
289 and MC in each of these suggests that the non-prompt contribution of Z+jets is well modeled by  
290 Monte Carlo.

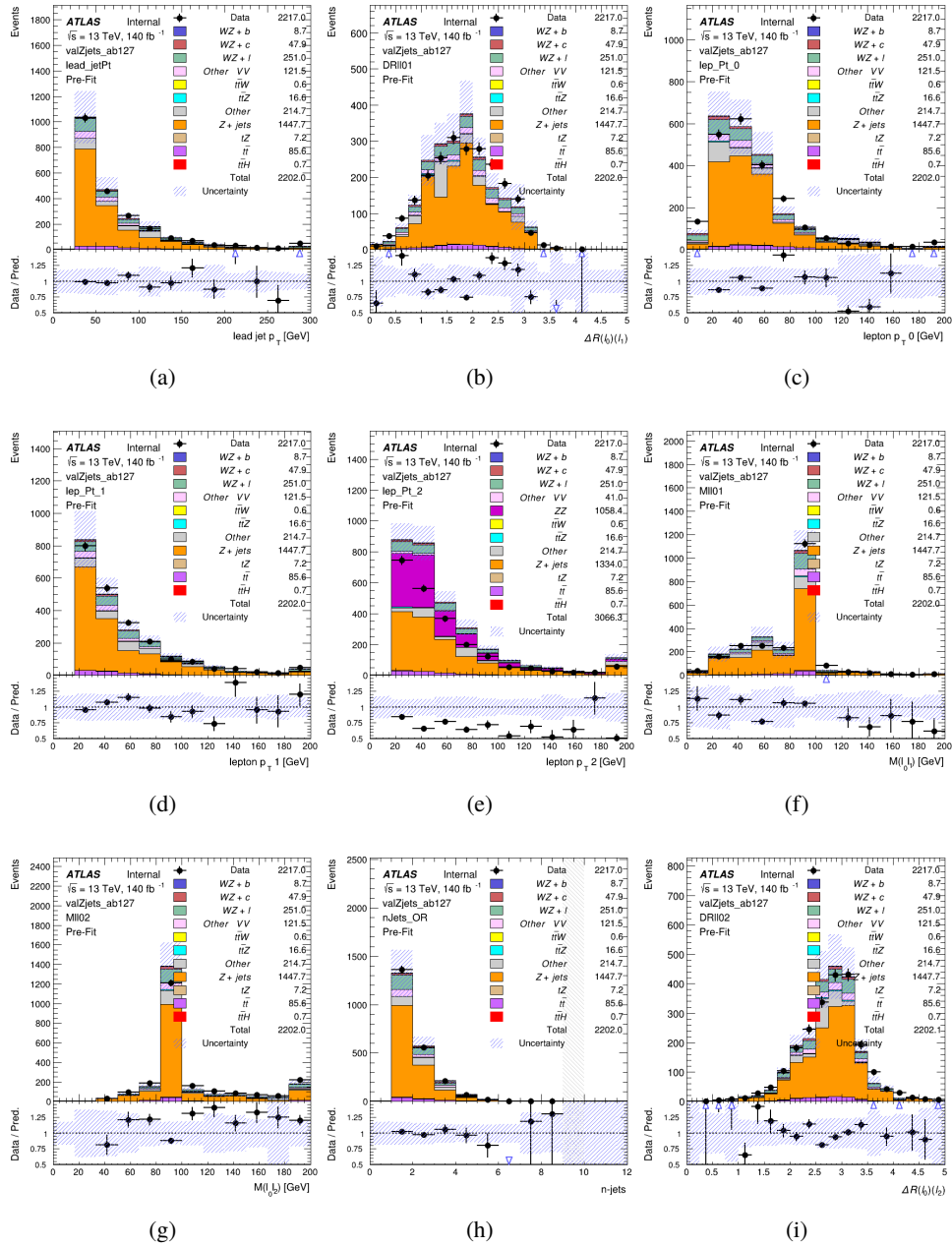


Figure 16: Comparisons between the data and MC distributions in the Z+jets validation region for (a) the  $p_T$  of the leading jet, (b)  $\Delta R$  between leptons 0 and 1, (c) the  $p_T$  of lepton 0, (d)  $p_T$  of lepton 1, (e)  $p_T$  of lepton 2, (f) the invariant mass of leptons 0 and 1, (g) the invariant mass of leptons 0 and 2, (h) the number of jets, (i)  $\Delta R$  between leptons 0 and 2

## 7 $tZ$ Interference Studies and Separation Multivariate Analysis

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

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

Separation between  $tZ$  and  $WZ + \text{heavy flavor}$  is achieved in part by reconstructing the invariant mass of the top candidate, which clusters more closely to the top mass for  $tZ$  than  $WZ + \text{heavy flavor}$ .

The result of this BDT is used to create a  $tZ$  enriched region in the fit, reducing its impact on the measurement of  $WZ + \text{heavy flavor}$ .

### 7.1 Interference Studies

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

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

The kinematics of these samples are shown below:



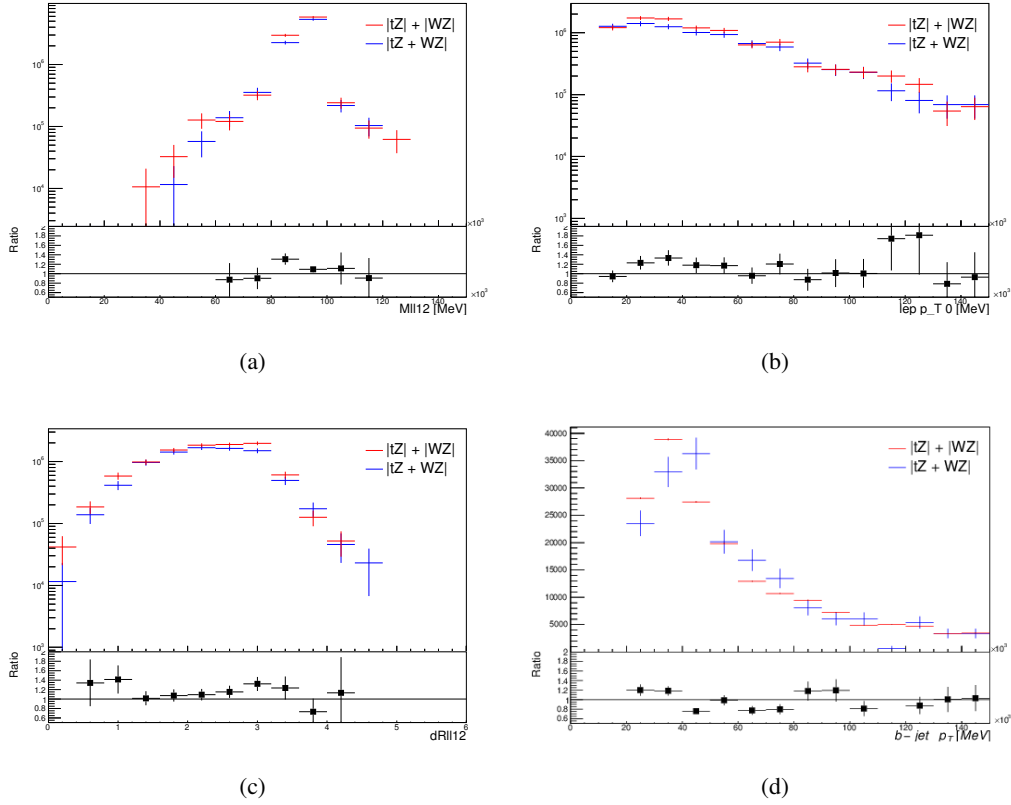


Figure 17: Comparisons between (a) the invariant mass of the Z-candidate, (b) the  $p_T$  of the leading lepton, (c)  $\Delta(R)$  of the two leptons that form the Z-candidate, and (d) the  $p_T$  of the b-jet, for WZ and tZ events generated with interference effects (blue) and without interference effects (red).

## 7.2 Top Mass Reconstruction

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

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

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

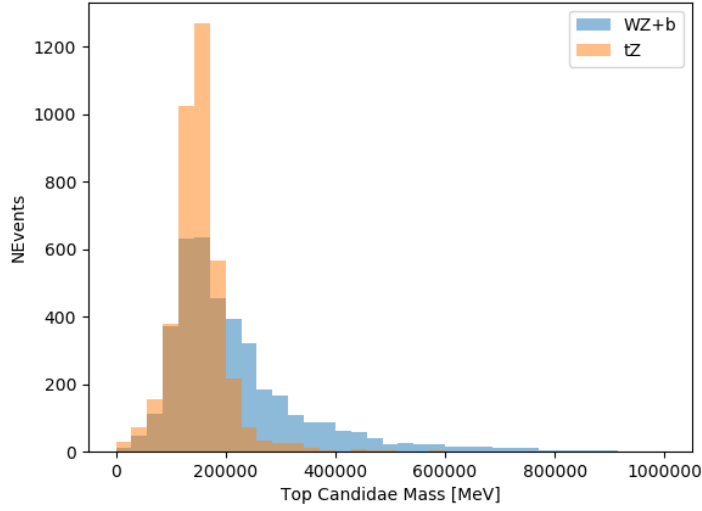


Figure 18: Reconstructed top mass distributions for tZ and WZ + b, measured in MeV.

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

Expanding this out into components, this equation gives:

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

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

The reconstructed top mass distribution for tZ and WZ + b can be seen in figure 18.

### 7.3 tZ BDT

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

- The invariant mass of the reconstructed top candidate
- $p_T$  of each of the leptons, jet
- The invariant mass of each combination of lepton pairs,  $M(l\bar{l})$

- $E_T^{\text{miss}}$
- Distance between each combination of leptons,  $\Delta R(l_l)$
- Distance between each lepton and the jet,  $\Delta R(l_j)$

The training samples included only events meeting the requirements of the 1-jet, >60% region, i.e. passing all the selection described in section 6 and having exactly one jet which passes the tightest (60%) DL1r working point.

The distributions of a few of these features for both signal and background is shown in figure 19.

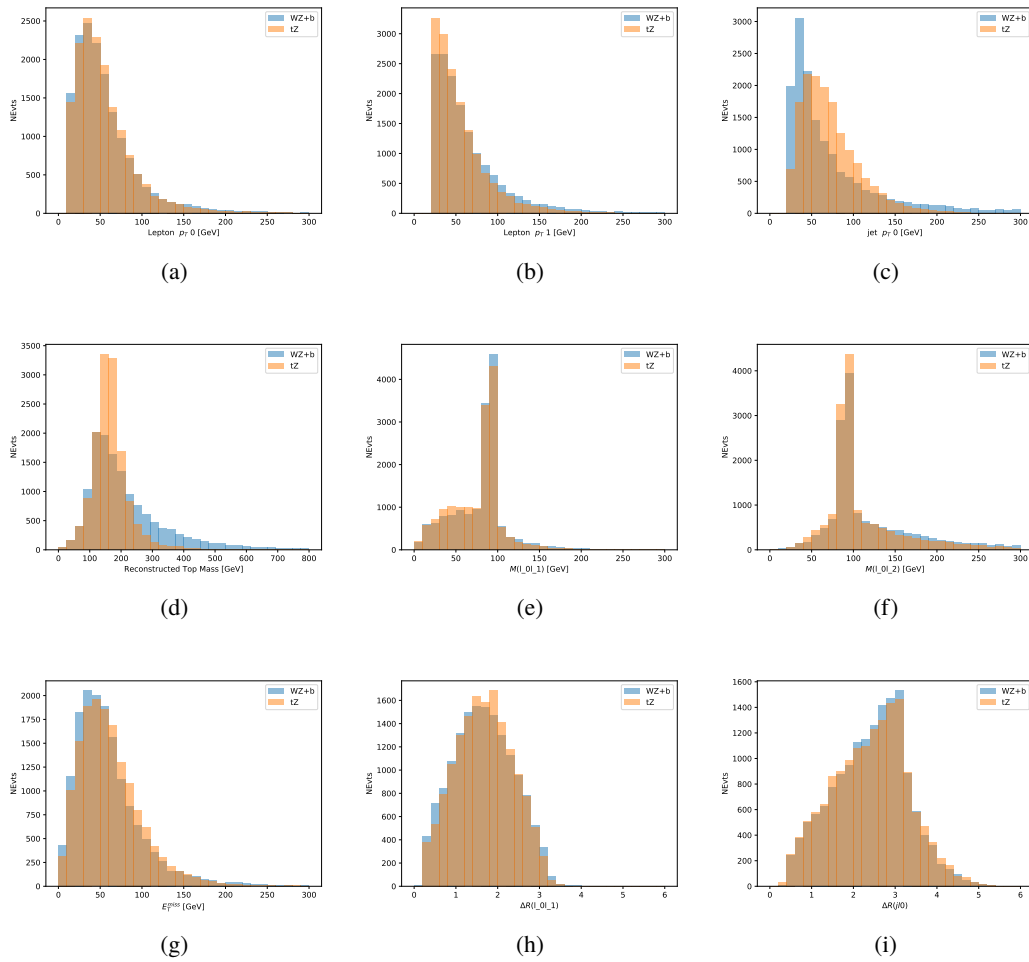


Figure 19: Distribution of input features of the BDT for signal (WZ) and background (tZ). Both are scaled to an equal number of events. (a), (b) and (c) show the  $p_T$  of lepton 0, lepton 1, and the jet, (d) show the reconstructed top mass, (e) and (f) show the invariant mass of leptons 0 and 1, leptons 0 and 2. (g) shows the  $E_T^{\text{miss}}$  of each event. (h) and (i) show the  $\Delta R$  between lepton 0 and lepton 1, and the jet.

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

355 The results of the BDT training are shown in figure 20. The output scores for both signal and  
 356 background events is shown on the left. The right shows the receiving operating characteristic  
 357 (ROC) curve that results from the MVA. The ROC curve represents the background rejection  
 358 as a function of signal efficiency, where each point on the curve represents a different response  
 359 score. The ROC curve of the BDT is compared to the performance of using an optimal set of flat  
 360 selections on the same set of input variables.

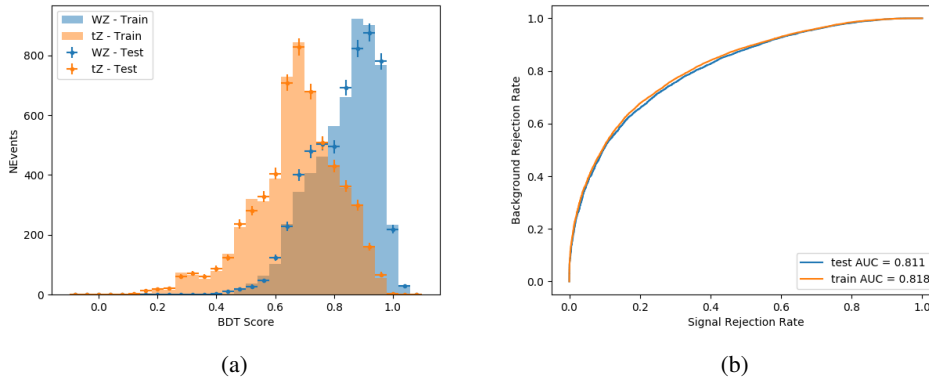


Figure 20: Distribution of the BDT response for signal and background events on the left, the ROC curve for the BDT on the right.

361 The relative important of each input feature in the model, measured by how often they appeared  
 362 in the decision trees, is shown in figure 21.

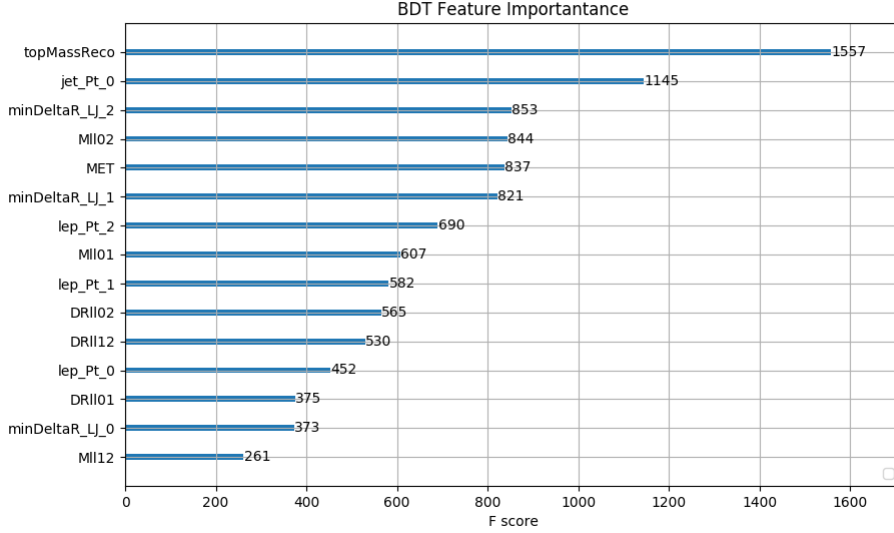


Figure 21: Relative importance of each input feature in the model.

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

## 8 Systematic Uncertainties

The systematic uncertainties that are considered are summarized in table 9. These are implemented in the fit either as a normalization factors or as a shape variation or both in the signal and background estimations. The numerical impact of each of these uncertainties is outlined in section 9.

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

The experimental uncertainties are related to the reconstruction and identification of light leptons and b-tagging of jets, and to the reconstruction of  $E_T^{\text{miss}}$ . The sources which contribute to the uncertainty in the jet energy scale [13] are decomposed into uncorrelated components and treated as independent sources in the analysis.

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

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

Table 9: Sources of systematic uncertainty considered in the analysis.

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

381 The full list of systematic uncertainties considered in the analysis is summarized in tables 10, 11  
382 and 12.

383

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

Table 10: Summary of experimental systematics considered for leptons and  $E_T^{\text{miss}}$ . Includes type, description, name of systematic as used in the fit, and mode of application. The mode of application indicates the systematic evaluation, e.g. as an overall event re-weighting (Event Weight) or rescaling ( $p_T$  Correction).

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

Table 11: Jet systematics take into account effects of jets calibration method,  $\eta$  inter-calibration, high  $p_T$  jets, pile-up, and flavor response. They are all diagonalised into effective parameters.



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

Table 12: Summary of experimental systematics to be included for b-tagging of jets in the analysis, using the continuous DL1r tagging algorithm. All of the b-tagging related systematics are applied as event weights. From left: type, description, and the name of systematic used in the fit.

384 The cross-section uncertainties applied to background estimates are summarized in table 13.

Process	X-section [%]
tZ	$^{+36}_{-31}$
t $\bar{t}$ H (aMC@NLO+Pythia8)	QCD Scale: $^{+5.8}_{-9.2}$ PDF( $+\alpha_S$ ): $\pm 3.6$
t $\bar{t}$ Z (aMC@NLO+Pythia8)	QCD Scale: $^{+9.6}_{-11.3}$ PDF( $+\alpha_S$ ): $\pm 4$
t $\bar{t}$ W (aMC@NLO+Pythia8)	QCD Scale: $^{+12.9}_{-11.5}$ PDF( $+\alpha_S$ ): $\pm 3.4$
VV + b/charm (Sherpa 2.2.1)	$\pm 50$
VV + light (Sherpa 2.2.1)	$\pm 6$
t $\bar{t}$	$\pm 20$ <b>placeholder</b>
Z + jets	$\pm 25$ <b>placeholder</b>
Others	$\pm 50$

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

## 385 9 Results

386 A separate maximum-likelihood fit is performed over the 1-jet and 2-jet fit regions in order to  
387 extract the best-fit value of the WZ + b-jet and WZ + charm jet contributions. The WZ + b,

WZ + charm and WZ + light contributions are allowed to float, with the remaining background contributions are held fixed. **The current fit strategy treats the WZ + b-jet contribution as the parameter of interest, with the normalization of the WZ + charm and the WZ + light contributions taken as systematic uncertainties. This could however be adjusted, depending on whether it is decided the goal of the analysis should be to measure WZ+b specifically or WZ + heavy flavor overall.** The result of the fit is used to extract the cross-section of WZ + heavy-flavor production.

A maximum likelihood fit to data is performed simultaneously in the regions described in section 6. The parameters  $\mu_{WZ+b}$ ,  $\mu_{WZ+charm}$ ,  $\mu_{WZ+light}$ , where  $\mu = \sigma_{\text{observed}}/\sigma_{\text{SM}}$ , are extracted from the fit.

## 9.1 1-jet Fit Results

**The results of the fit are currently blinded.** The post-fit yields in each region are summarized in figure 22.

A post-fit summary plot of the 1-jet fitted regions is shown in figure 23:

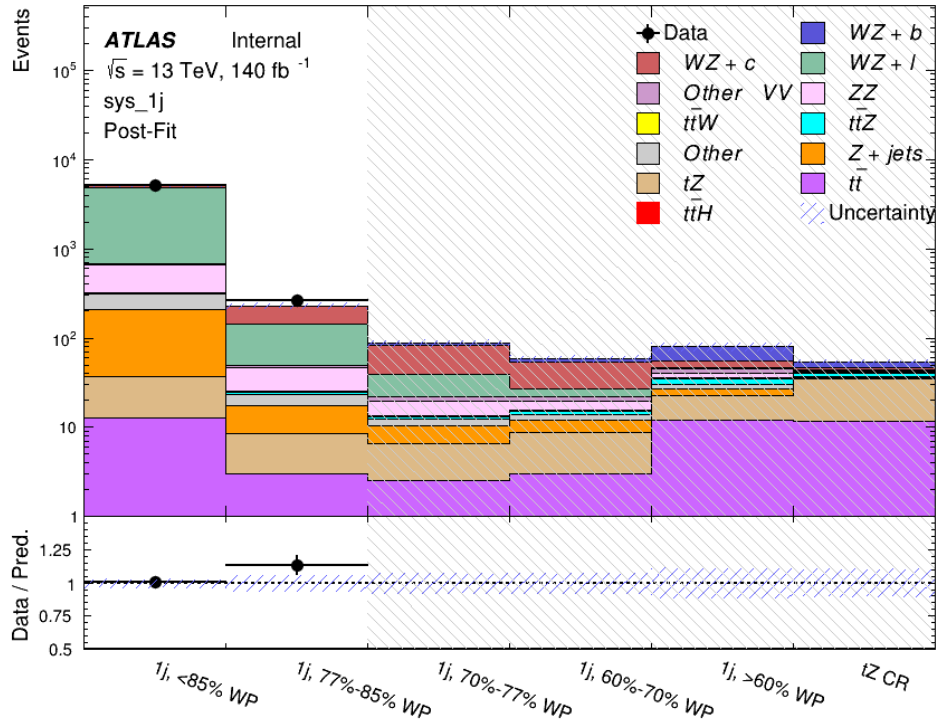


Figure 23: Post-fit summary of the 1-jet fit regions.

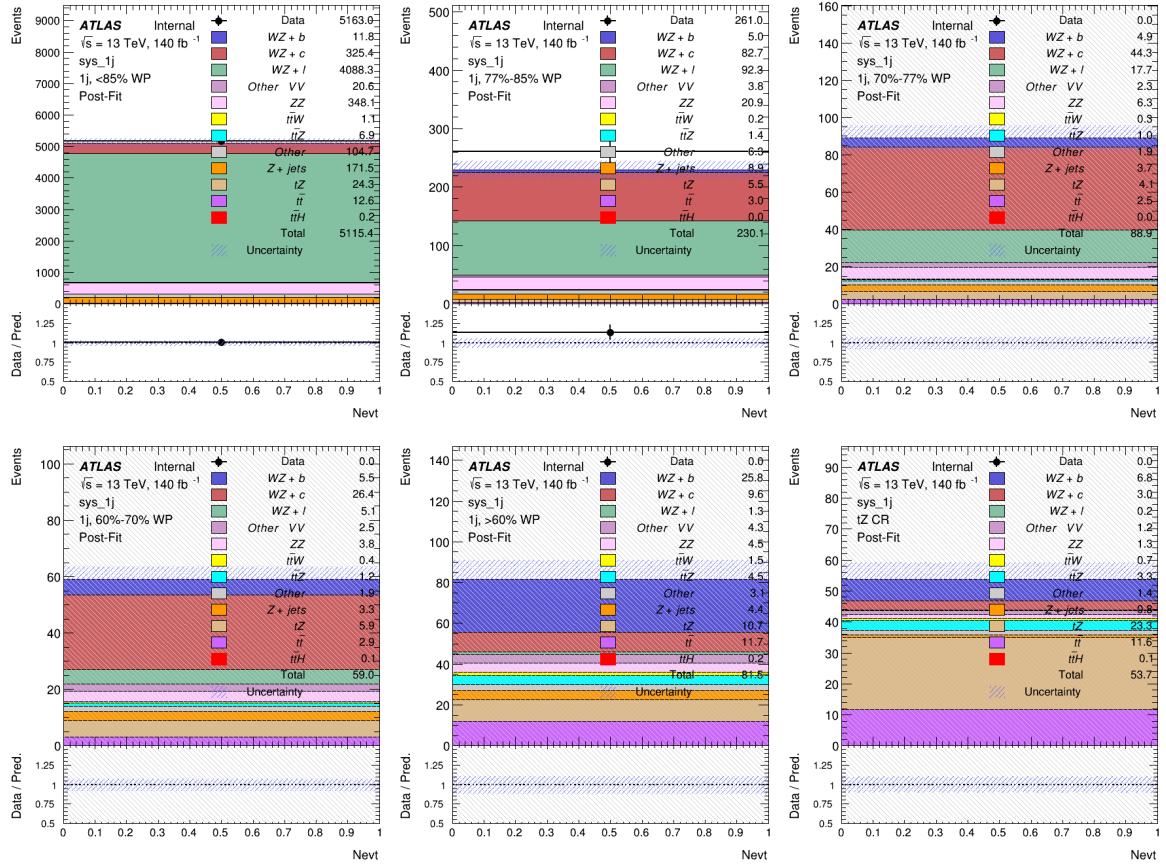


Figure 22: Data/MC results in each of the 1-jet regions after the fit has been performed.

As described in section 8, there are 226 systematic uncertainties that are considered as NPs in the fit. These NPs are constrained by Gaussian or log-normal probability density functions. The latter are used for normalisation factors to ensure that they are always positive. The expected numbers of signal and background events are functions of the likelihood. The prior for each NP is added as a penalty term, decreasing the likelihood as it is shifted away from its nominal value.

The impact of each systematic uncertainty is calculated by performing the fit with the parameter of interest held fixed, varied from its fitted value by its uncertainty, and calculating  $\delta\mu$  relative to the baseline fit. 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 $\bar{t}$ cross-section	-0.0662	0.0654
Luminosity	-0.0609	0.0655
Z + jets cross-section	-0.0284	0.0284
Other VV + charm cross-section	0.0207	-0.0202
Muon Trigger Scale Factor	0.019	0.0209
Total Systematic Uncertainty	0.3511	0.3679

Table 14: Summary of the most significant sources of systematic uncertainty on the measurement of WZ + b with exactly one associated jet.

412 The ranking and impact of those nuisance parameters with the largest contribution to the overall  
413 uncertainty is shown in figure 24.

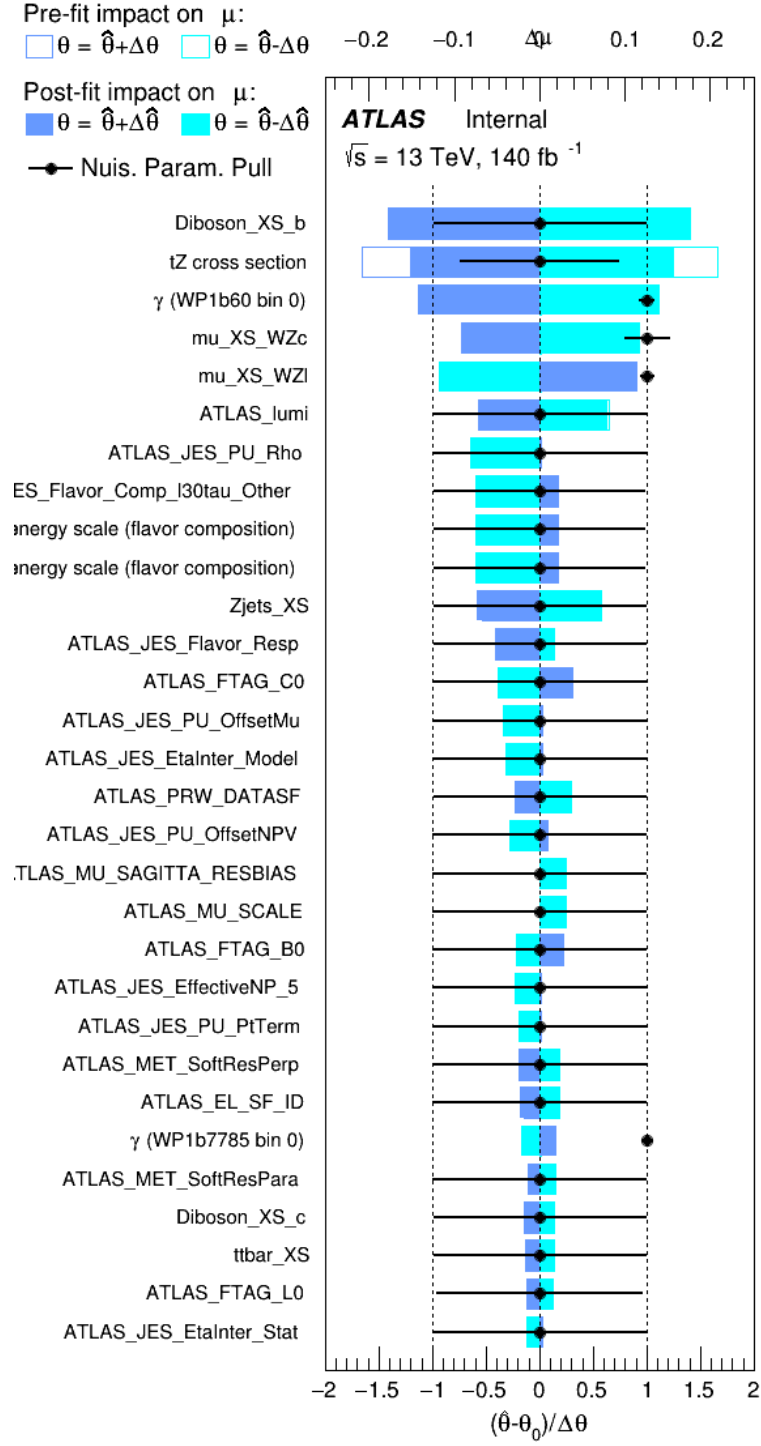


Figure 24: Impact of systematic uncertainties on the signal-strength of WZ + b for events with exactly one jet

414 The large impact of the Jet Energy Scale and Jet Flavor Tagging is unsurprising, as the shape  
 415 of the fit regions depends heavily on the modeling of the jets. The other major sources of  
 416 uncertainty come from background modelling and cross-section uncertainty. The pie charts in  
 417 figure 25 show that for the modelling uncertainties that contribute most correspond to the most  
 418 significant backgrounds.

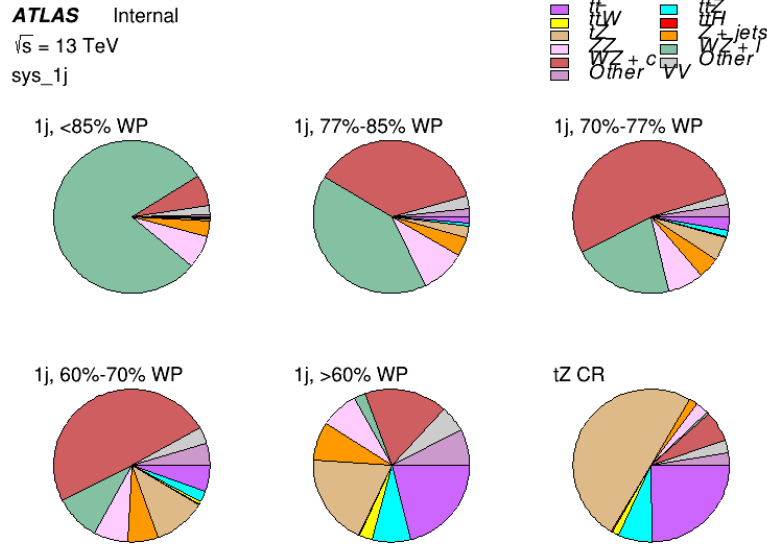


Figure 25: Background composition of the fit regions.

419 The correlations between these nuisance parameters are summarized in figure 26.

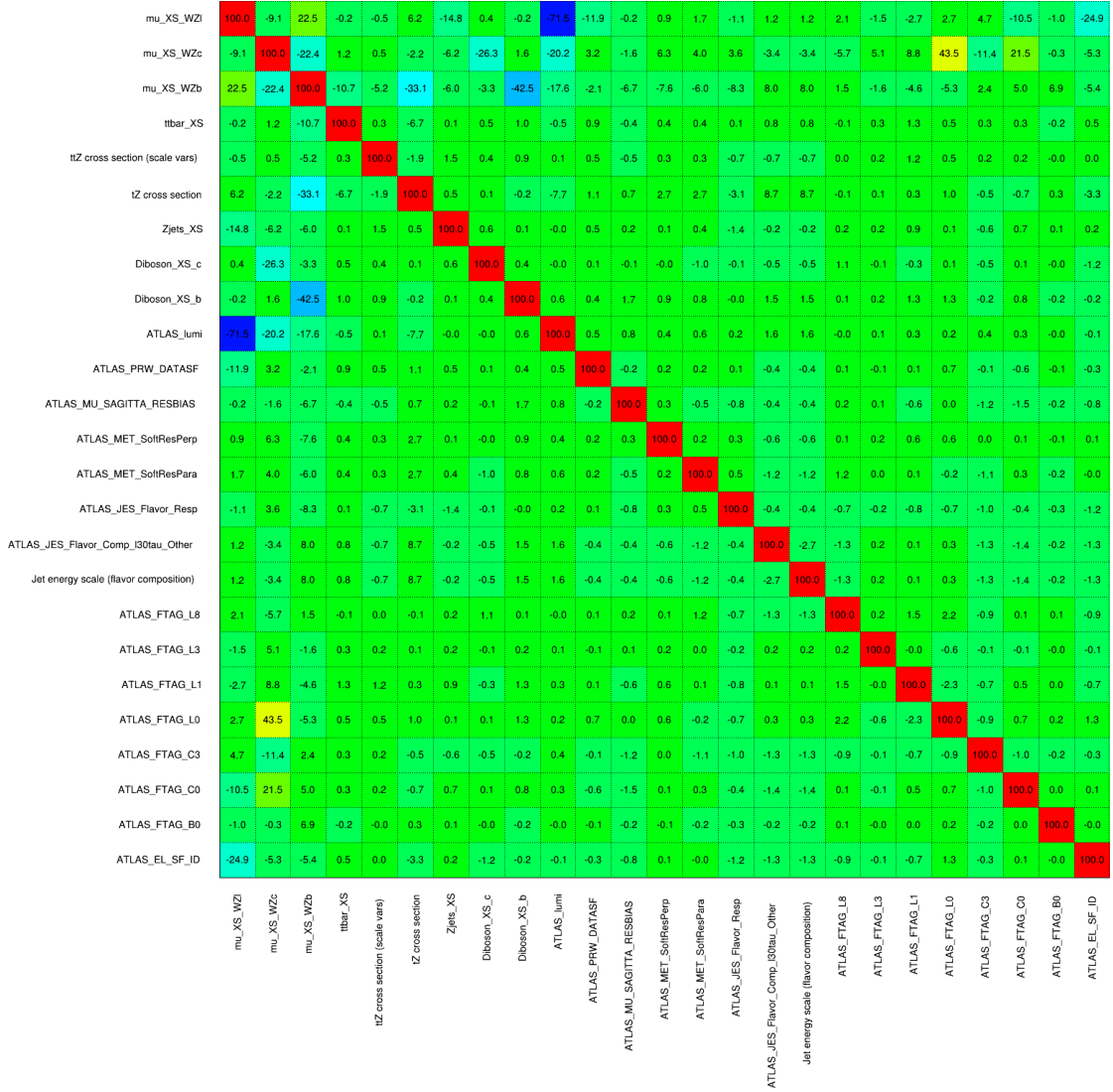


Figure 26: Correlations between nuisance parameters

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

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

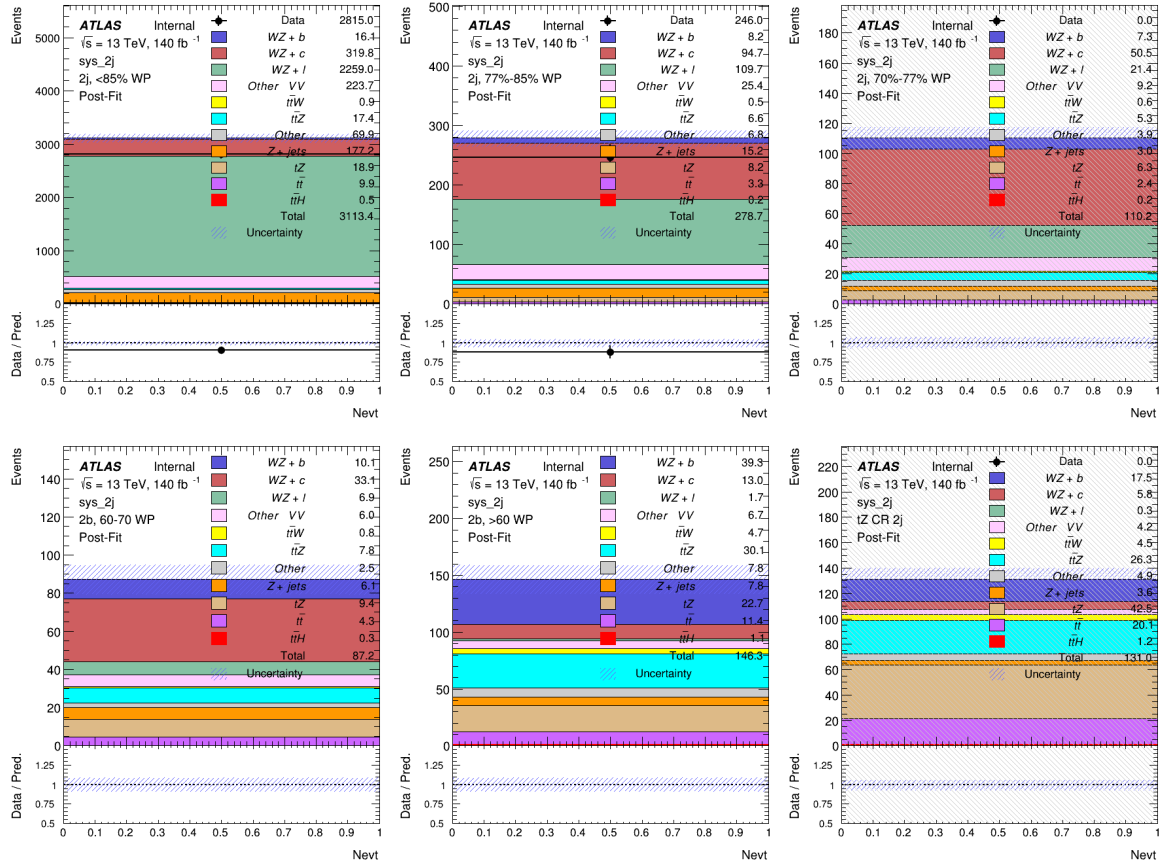


Figure 27: Data/MC results in each of the regions in the 2-jet fit after the fit has been performed.

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

## 9.2 2-jet Fit Results

**The results of the fit are currently blinded.** The post-fit yields in each region are summarized in figure 27.

A post-fit summary plot of the fitted regions is shown in figure 28:

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



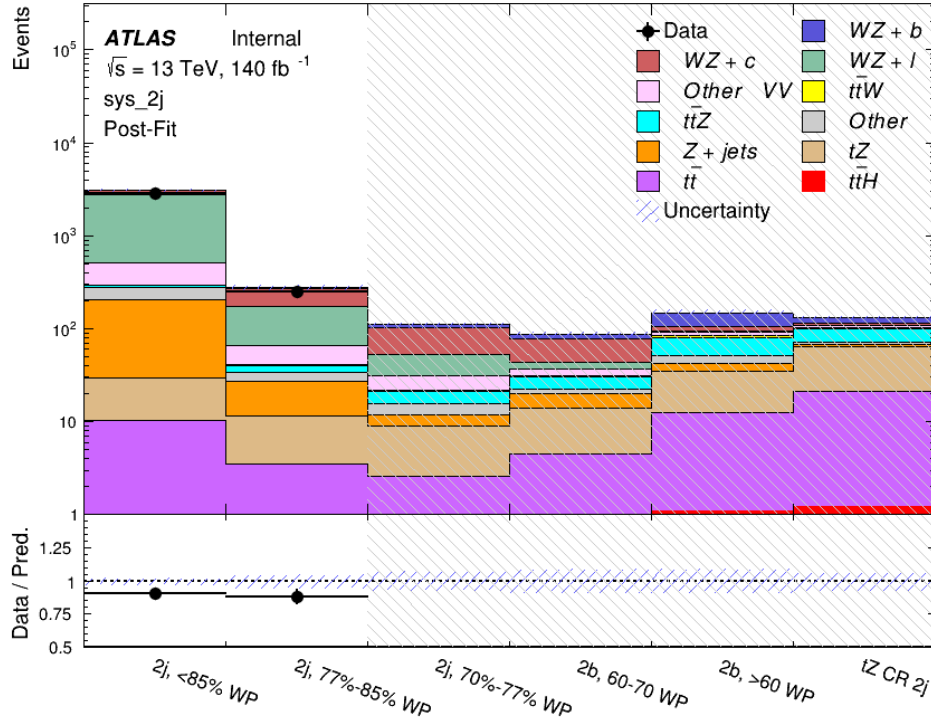


Figure 28: Post-fit summary of the fit over 2-jet regions.

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

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

437 The ranking and impact of those nuisance parameters with the largest contribution to the overall  
 438 uncertainty is shown in figure 29.

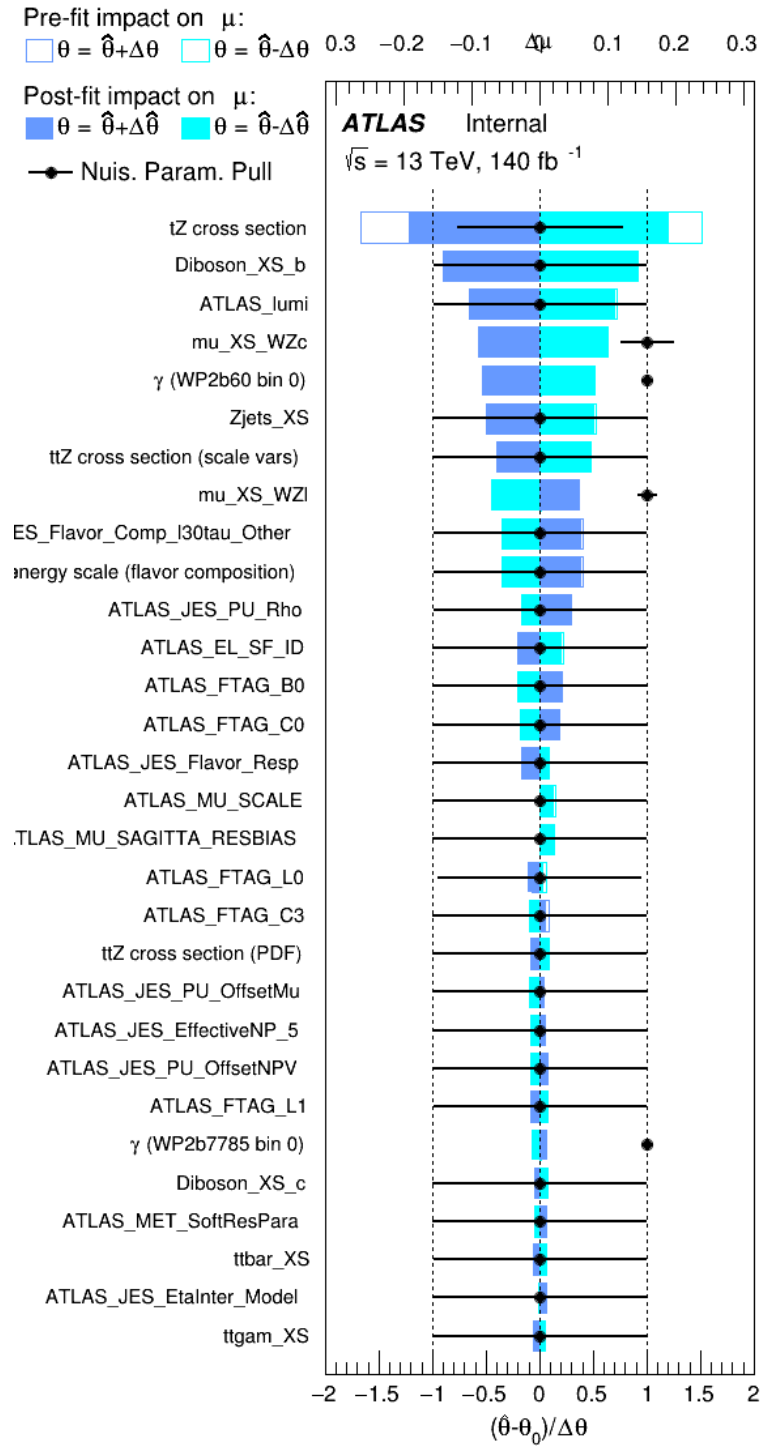


Figure 29: Impact of systematic uncertainties on the signal-strength of WZ + b in 2-jet events.

439 The large impact of the Jet Energy Scale and Jet Flavor Tagging is unsurprising, as the shape  
 440 of the fit regions depends heavily on the modeling of the jets. The other major sources of  
 441 uncertainty come from background modelling and cross-section uncertainty. The pie charts in  
 442 figure 30 show that for the modelling uncertainties that contribute most correspond to the most  
 443 significant backgrounds.

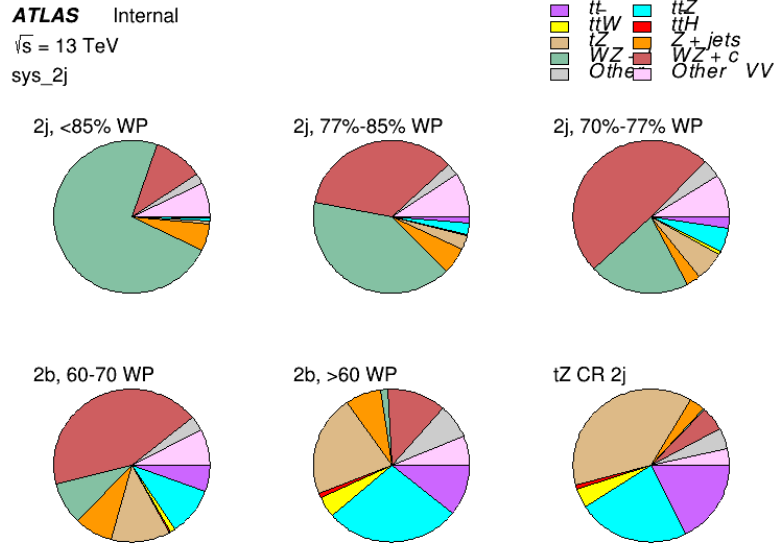


Figure 30: Background composition of the 2-jet fit regions.

444 The correlations between these nuisance parameters are summarized in figure 31.

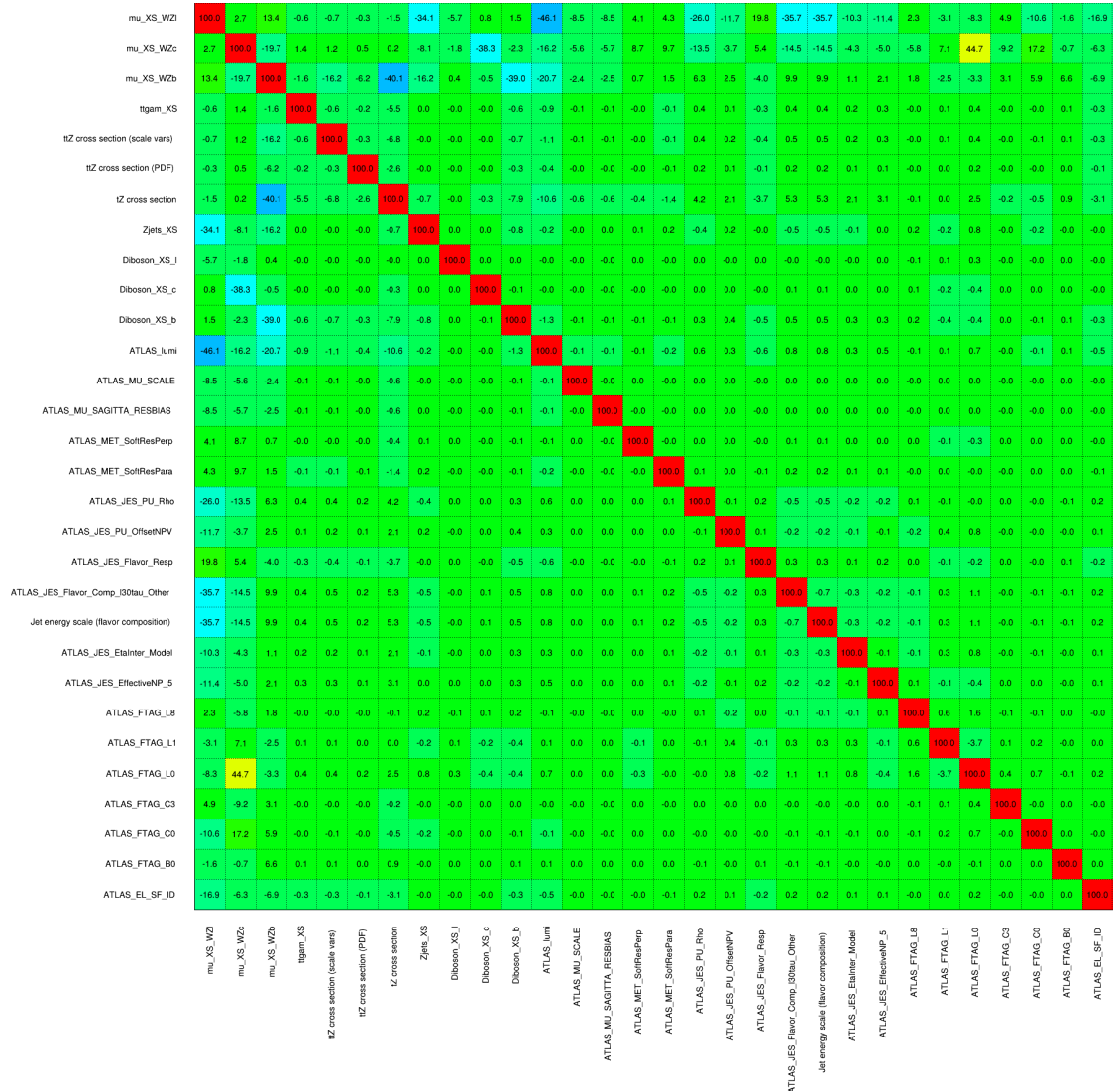


Figure 31: Correlations between nuisance parameters in the 2-jet fit

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

## 10 Conclusion

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

final results once unblinded.

## References

- [1] M. Aaboud et al. ‘Observation of electroweak  $W^\pm Z$  boson pair production in association with two jets in pp collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector’. In: *Phys. Lett. B* 793 (2019), pp. 469–492. DOI: [10.1016/j.physletb.2019.05.012](https://doi.org/10.1016/j.physletb.2019.05.012). arXiv: [1812.09740](https://arxiv.org/abs/1812.09740) [hep-ex].
- [2] T. Gleisberg et al. ‘Event generation with SHERPA 1.1’. In: *JHEP* 02 (2009), p. 007. DOI: [10.1088/1126-6708/2009/02/007](https://doi.org/10.1088/1126-6708/2009/02/007). arXiv: [0811.4622](https://arxiv.org/abs/0811.4622) [hep-ph].
- [3] ATLAS Collaboration. *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton–proton collision data*. ATLAS-CONF-2016-024. 2016. URL: <https://cds.cern.ch/record/2157687>.
- [4] ATLAS Collaboration. ‘Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton–proton collision data’. In: *Eur. Phys. 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](https://arxiv.org/abs/1407.3935) [hep-ex].
- [5] *Evidence for the associated production of the Higgs boson and a top quark pair with the ATLAS detector*. Tech. rep. ATLAS-CONF-2017-077. Geneva: CERN, Nov. 2017. URL: <https://cds.cern.ch/record/2291405>.
- [6] 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: <https://cds.cern.ch/record/2037613>.
- [7] ATLAS Collaboration. *Selection of jets produced in 13 TeV proton–proton collisions with the ATLAS detector*. ATLAS-CONF-2015-029. 2015. URL: <https://cds.cern.ch/record/2037702>.
- [8] ATLAS Collaboration. ‘Performance of pile-up mitigation techniques for jets in pp collisions at  $\sqrt{s} = 8$  TeV using the ATLAS detector’. In: *Eur. Phys. J. C* 76 (2016), p. 581. 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].
- [9] ATLAS Collaboration. *Performance of missing transverse momentum reconstruction with the ATLAS detector in the first proton–proton collisions at  $\sqrt{s} = 13$  TeV*. ATL-PHYS-PUB-2015-027. 2015. URL: <https://cds.cern.ch/record/2037904>.
- [10] P. S. A. Hoecker. ‘TMVA 4 Toolkit for Multivariate Data Analysis with ROOT’. In: *arXiv:physics/0703039* (2013).
- [11] F. Cardillo et al. ‘Measurement of the fiducial and differential cross-section of a top quark pair in association with a Z boson at 13 TeV with the ATLAS detector’. In: ATL-COM-PHYS-2019-334 (Apr. 2019). URL: <https://cds.cern.ch/record/2672207>.

- 485 [12] ATLAS Collaboration. ‘Luminosity determination in pp collisions at  $\sqrt{s} = 7$  TeV using  
486 the ATLAS detector at the LHC’. In: *Eur. Phys. J. C* 71 (2011), p. 1630. doi: [10.1140/  
487 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].
- 488 [13] G. Aad et al. ‘Jet energy resolution in proton-proton collisions at  $\sqrt{s} = 7$  TeV recorded  
489 in 2010 with the ATLAS detector’. In: *The European Physical Journal C* 73.3 (Mar.  
490 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:  
491 <https://doi.org/10.1140/epjc/s10052-013-2306-0>.
- 492 [14] A. Collaboration. ‘Performance of b -jet identification in the ATLAS experiment’. In:  
493 *Journal of Instrumentation* 11.04 (2016), P04008. URL: [http://stacks.iop.org/  
494 1748-0221/11/i=04/a=P04008](http://stacks.iop.org/1748-0221/11/i=04/a=P04008).