ATLAS Note

GROUP-2017-XX

17th February 2021

Draft version 0.1

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

The ATLAS Collaboration

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

Contents

1 Changes and outstanding items

1.1 Changelog

This is version 5

1.1.1 Changes relative to v4

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

1.1.2 Changes relative to v3

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

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

1.1.3 Changes relative to v2

- Added alternate VBS samples to include missing b-jet diagrams
- Included a section on tZ interference effects, **??**.
- 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 seperate 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

1.1.4 Changes relative to v1

- Added GRL list
- Fixed latex issue in line 92, typo in line 172
- Added tables **??** and **??**, summarizing the event and object selection
- Added table **??**, which includes the DSID of samples used
- Included reference to WZ inclusive paper in introduction

1.2 Outstanding Items

- Complete interference studies, apply any interference effects observed as a systematic
- Update results section with additional studies, possibly including:
  - Truth jet migration studies
  - Simultaneous fit over 1j and 2j
  - Impact of allowing tZ to float
- Unblind, update plots and fits to include data
- 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 rejcted. 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 fb^{-1} of data from pp collisions at $\sqrt{s} = 13$ 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 non-prompt Validation Regions.

Section **??** details the data and Monte Carlo (MC) samples used in the analysis. The reconstruction of various physics objects is described in section **??**. Section **??** 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 **??**. Section **??** describes the various sources of systematic uncertainties considered in the fit. Finally, the results of the analysis are summarized in section **??**, followed by a brief conclusion in section **??**.

The current state of thee analysis shows blinded results for thee full Run 2 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 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 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 **??**. A list of the specific samples used by data set ID is shown in table **??**.

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 functin (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
	(Sherpa 2.1.1)	(LO multileg)	(Sherpa)	(NNPDF 3.0 NLO)
$t\bar{t}(Z/\gamma^* \to ll)$	MG5_aMC	NLO	Pythia 8	NNPDF 3.0 NLO
$t\bar{t}H$	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 [**powhegtt**]	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				
$Z \to l^+l^-$	Sherpa 2.2.1	MEPS NLO	Sherpa	NNPDF 3.0 NLO

Sample	DSID
WZ	364253, 364739-42
VV	364250, 364254, 364255, 363355-60, 364890
$t\bar{t}W$	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.

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 **??**.

4.1 Trigger

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

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

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 $|\eta_{\text{cluster}}| < 2.47$. 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 $|\eta| < 2.5$. 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 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 ($|d_0|/\sigma_{d_0} < 5$) and the longitudinal impact parameter ($|\Delta z_0 \sin\theta_\ell| < 0.5$ mm).

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 $|\eta| < 2.5$. The longitutinal impact parameter is the same for both electrons and muons, while muons are required to pass a slightly tighter transverse imapact parameter, $|d_0|/\sigma_{d_0} < 3$.

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

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

4.4 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.

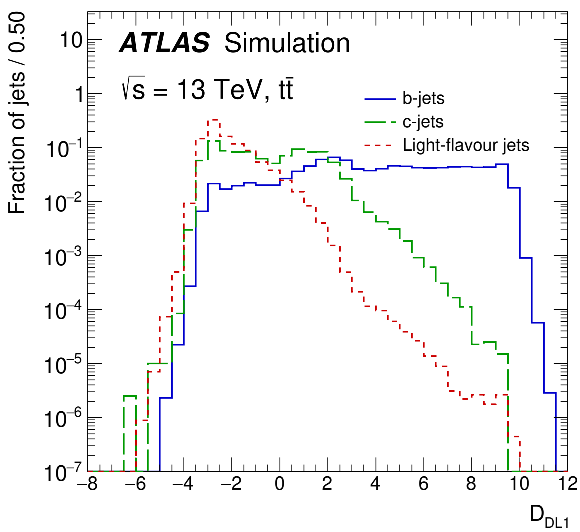


Figure 1: Output distribution of the DL1r algorithm for b-jets, charm jets, and light jets

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 **??**.

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

Table 4: 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 pseudo-continuous b-tag spectrum is used in this case to separate out WZ + b, WZ + charm, and WZ + light.

4.5 Missing transverse energy

Missing transverse momentum (E_T^{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^{miss} reconstruction are required to have p_T>10 GeV, while jets are required to have p_T>20 GeV.

5 Event Selection and Signal Region Definitions

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

5.1 Event Preselection

Events are required to include exactly three reconstructed light leptons passing the requirement described in **??**, which have a total charge of ± 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 tightly isolated, 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 ΔR, 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^{miss} > 20 GeV, and the transverse mass of the W candidate, $m(E_T^{miss} + l_{other}) > 30$ GeV, where E_T^{miss} is the missing transverse energy, and l_{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 **??**. 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 ±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, $m_T(E_T^{miss} + lep_{other}) > 30$ GeV
Missing transverse energy, $E_T^{miss} > 20$ GeV
One or two jets with $p_T > 25$ GeV

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

The event yields in the preselection region for both data and Monte Carlo are summarized in table **??**, which shows good agreement between data and Monte Carlo, and demonstrates that this region consists primarily of WZ events. The WZ events are split into WZ + b, WZ + c, and WZ + l based on the truth flavor of the associated jet in the event. Specifically, this determination is made based on the `HadronConeExclTruthLabelID` of the jet, as recommended by the b-tagging working group [**BtagWG**]. In this ordering b-jet supersede charm, which supersides light. That is, WZ + l events contain no charm and no b jets at truth level, WZ + c contain at 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
$t\bar{t}W$	16.8 ± 2.3
$t\bar{t}Z$	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
$t\bar{t}WW$	0.23 ± 0.05
$Z+$jets	600 ± 260
$V+\gamma$	37 ± 54
tZ	190 ± 70
tW	5.5 ± 1.2
WtZ	25.8 ± 1.1
VVV	26.2 ± 0.9
VH	94 ± 7
$t\bar{t}$	108.68 ± 8
$t\bar{t}H$	4.3 ± 0.5
Total	10600 ± 530
Data	10574

Table 6: Event yields in the preselection region at 139.0 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 figure **??**. Here, bins with 5% or more WZ+b are blinded.

WZ Fit Region - Inclusive

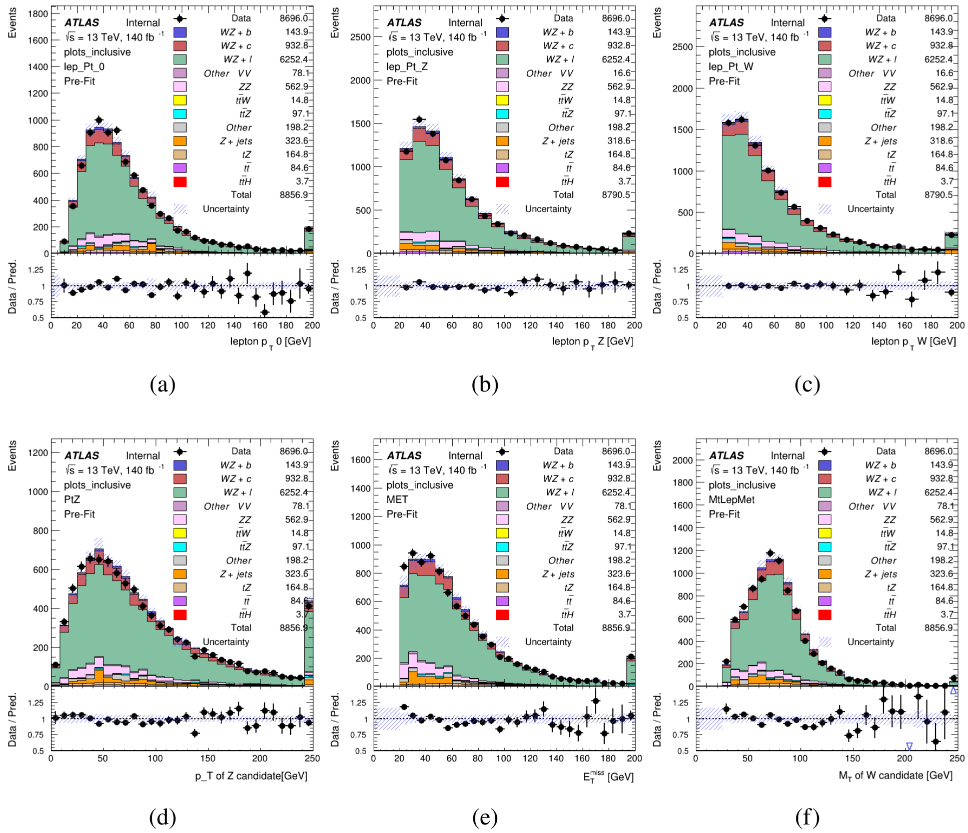


Figure 2: Comparisons between the data and MC distributions in the preselection region for (a) the p_T of the opposite sign lepton, (b) the p_T of the other lepton from the Z candidate, (c) the p_T of the lepton from the W candidate, (d) the p_T of the Z candidate, (e) the E_T^{miss}, and (f) the m_T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

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

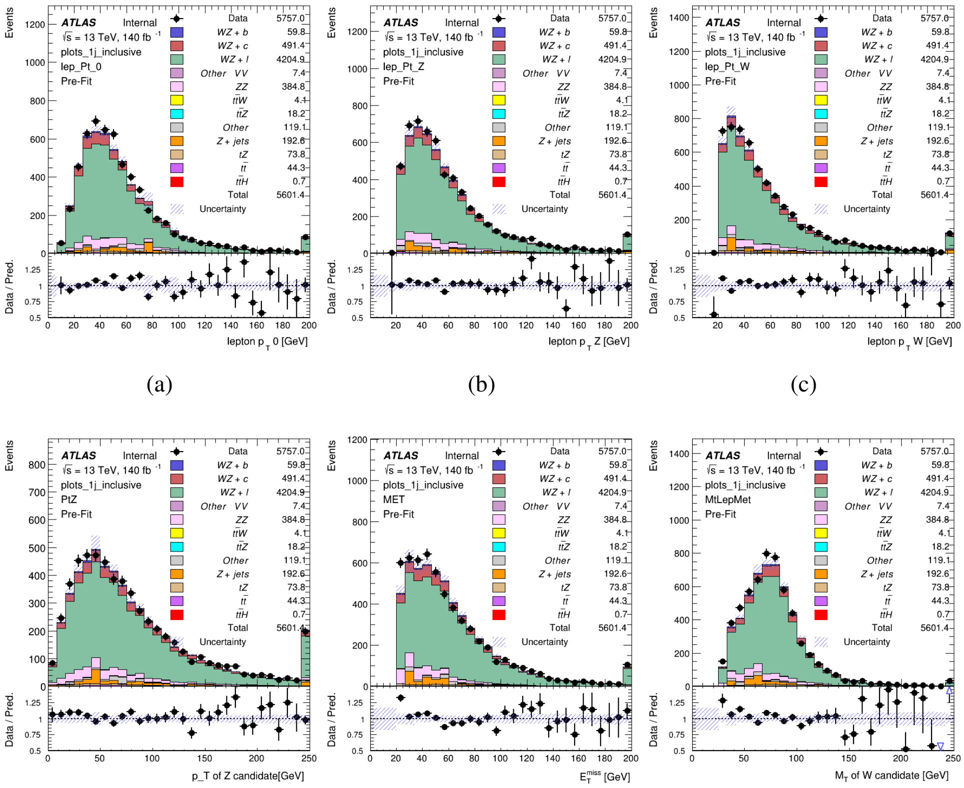
Region	Selection
1j, <85%	N_{jets} = 1, nJets_DL1r_85 = 0
1j, 85%-77%	N_{jets} = 1, nJets_DL1r_85 = 1, nJets_DL1r_77=0
1j, 77%-70%	N_{jets} = 1, nJets_DL1r_77 = 1, nJets_DL1r_70=0
1j, 70%-60%	N_{jets} = 1, nJets_DL1r_70 = 1, nJets_DL1r_60=0
1j, >60%	N_{jets} = 1, nJets_DL1r_60 = 1, tZ BDT > 0.725
1j tZ CR	N_{jets} = 1, nJets_DL1r_60 = 1, tZ BDT < 0.725
2j, <85%	N_{jets} = 2, nJets_DL1r_85 = 0
2j, 85%-77%	N_{jets} = 2, nJets_DL1r_85 >= 1, nJets_DL1r_77=0
2j, 77%-70%	N_{jets} = 2, nJets_DL1r_77 >= 1, nJets_DL1r_70=0
2j, 70%-60%	N_{jets} = 2, nJets_DL1r_70 >= 1, nJets_DL1r_60=0
2j, >60%	N_{jets} = 2, nJets_DL1r_60 >= 1, tZ BDT > 0.725
2j tZ CR	N_{jets} = 2, nJets_DL1r_60 >= 1, tZ BDT < 0.725

The working points discussed in section **??** 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 **??**. 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 **??**-**??**.

WZ Fit Region - 1j Inclusive

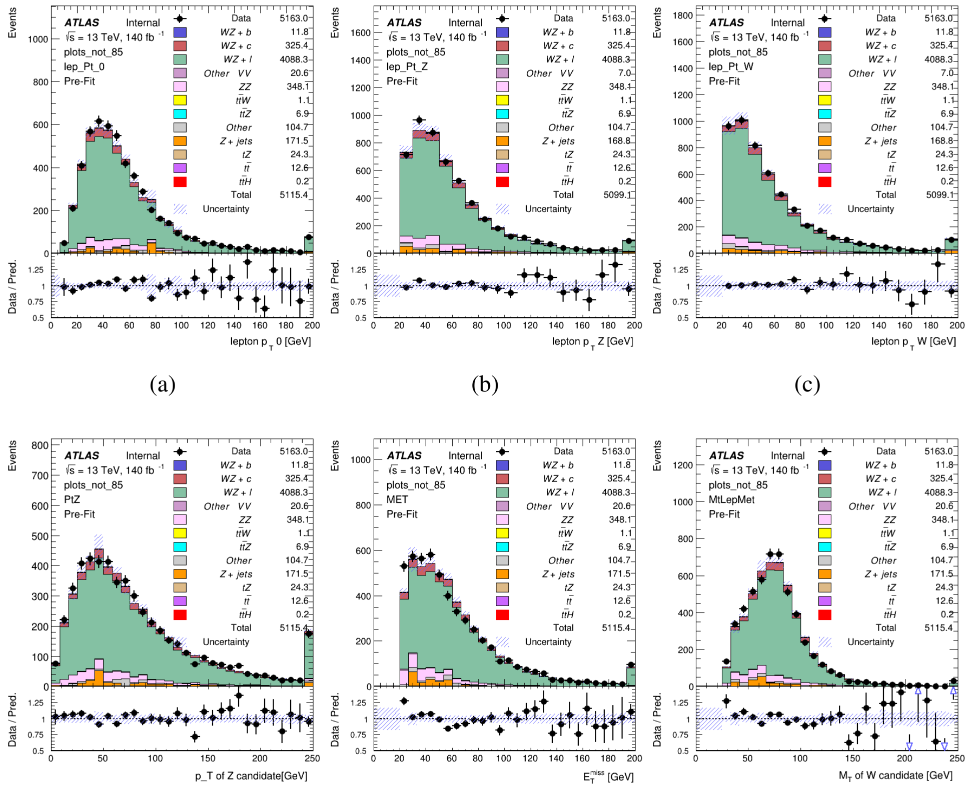


(a) (b) (c)

(d) (e) (f)

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

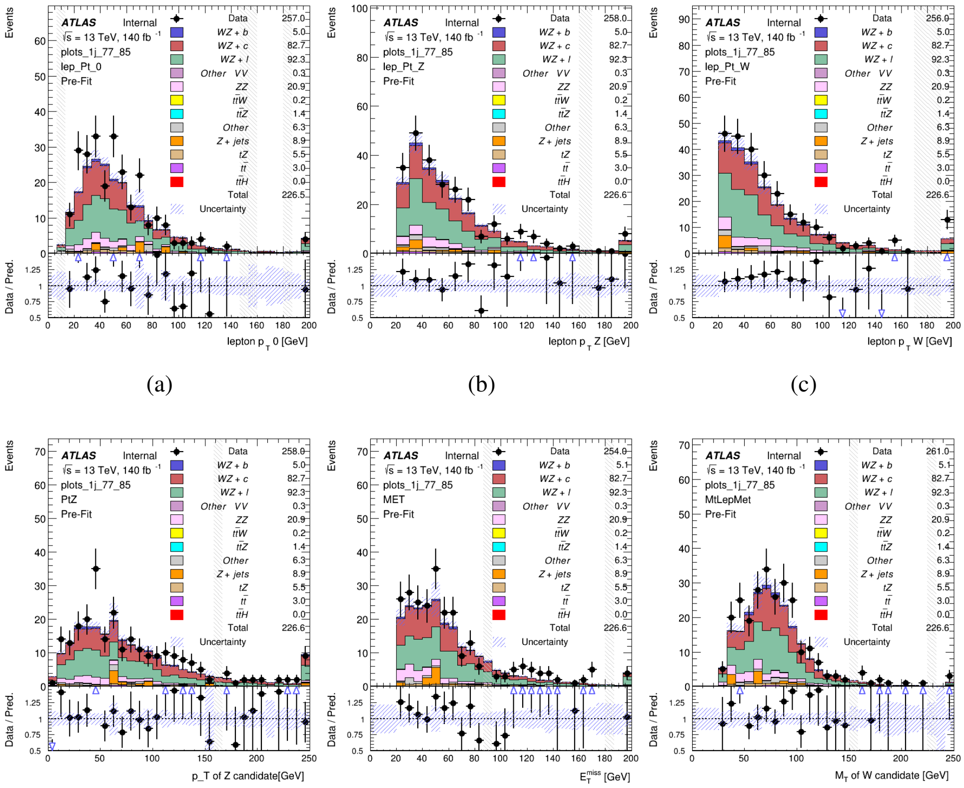


(a) (b) (c)

(d) (e) (f)

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

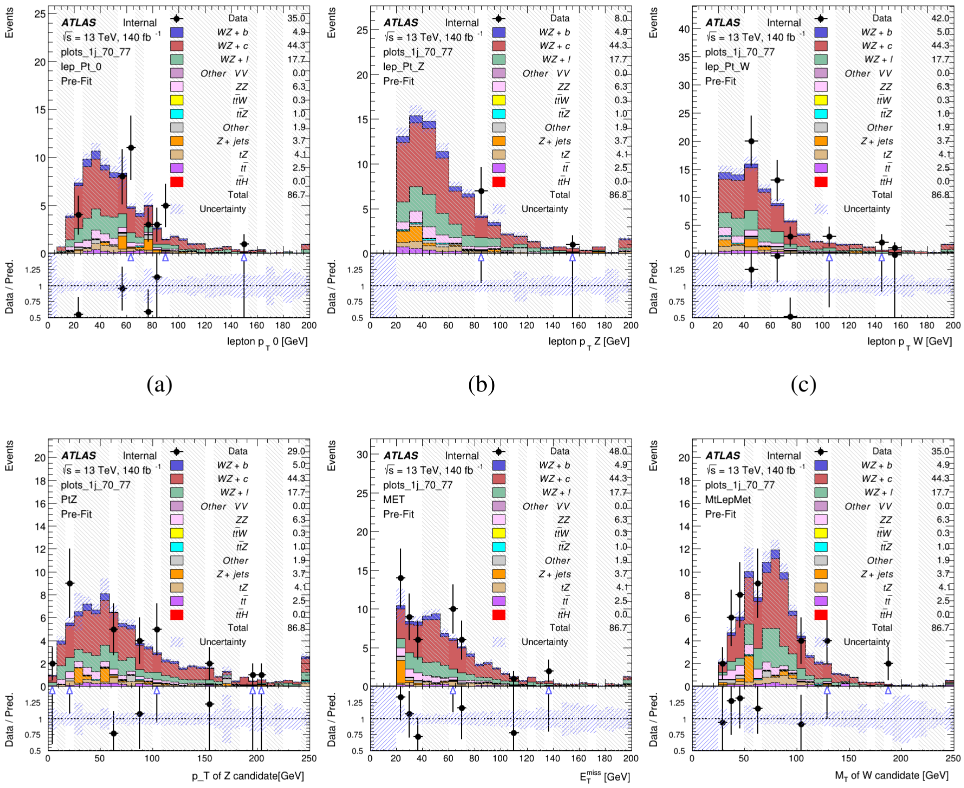


(a) (b) (c)

(d) (e) (f)

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

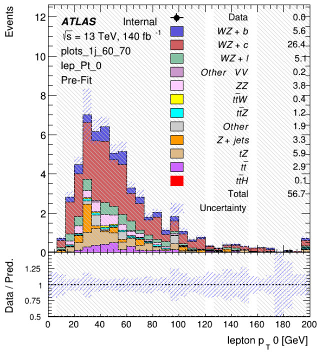


(a) (b) (c)

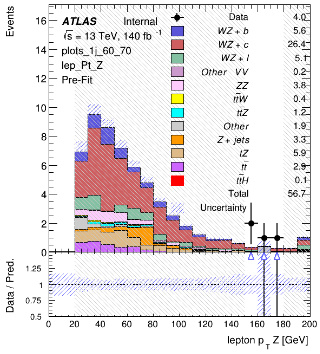
(d) (e) (f)

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.

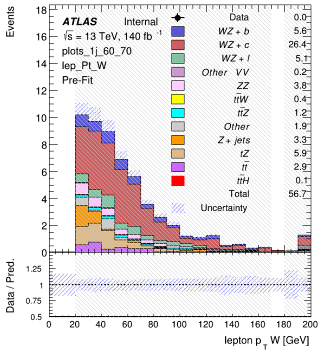
WZ Fit Region - 1j 60-70% WP



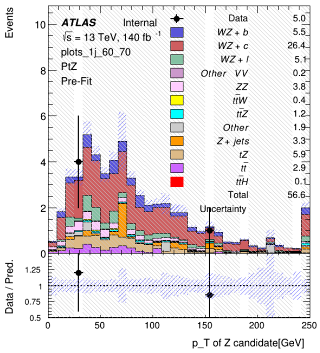
(a)



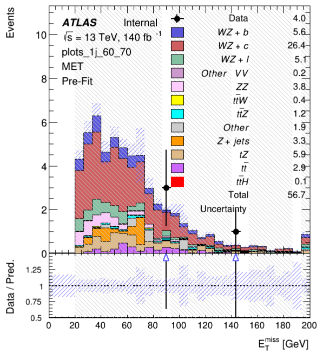
(b)



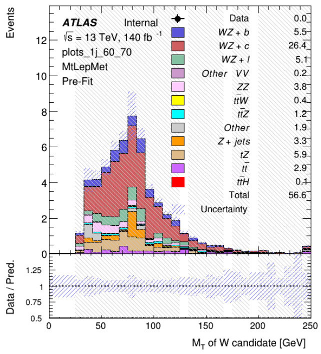
(c)



(d)



(e)



(f)

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

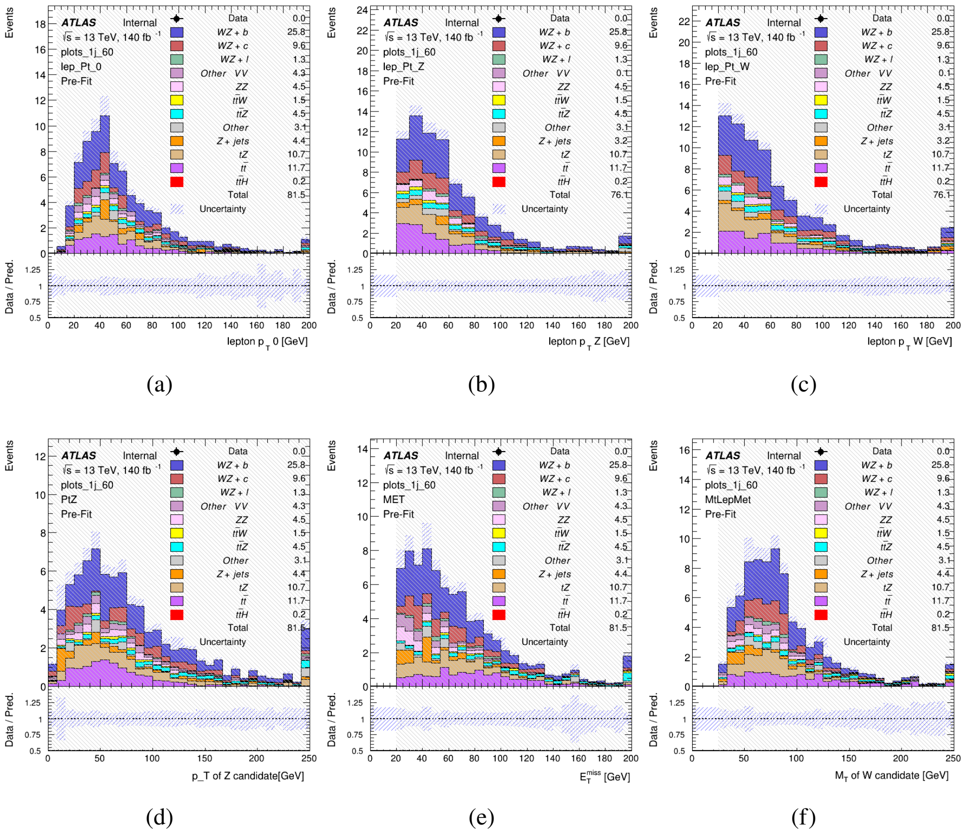


Figure 8: Comparisons between the data and MC distributions in the preselection region for (a) the p_T of the opposite sign lepton, (b) the p_T of the other lepton from the Z candidate, (c) the p_T of the lepton from the W candidate, (d) the p_T of the Z candidate, (e) the E_T^{miss}, and (f) the m_T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

WZ Fit Region - tZ-CR

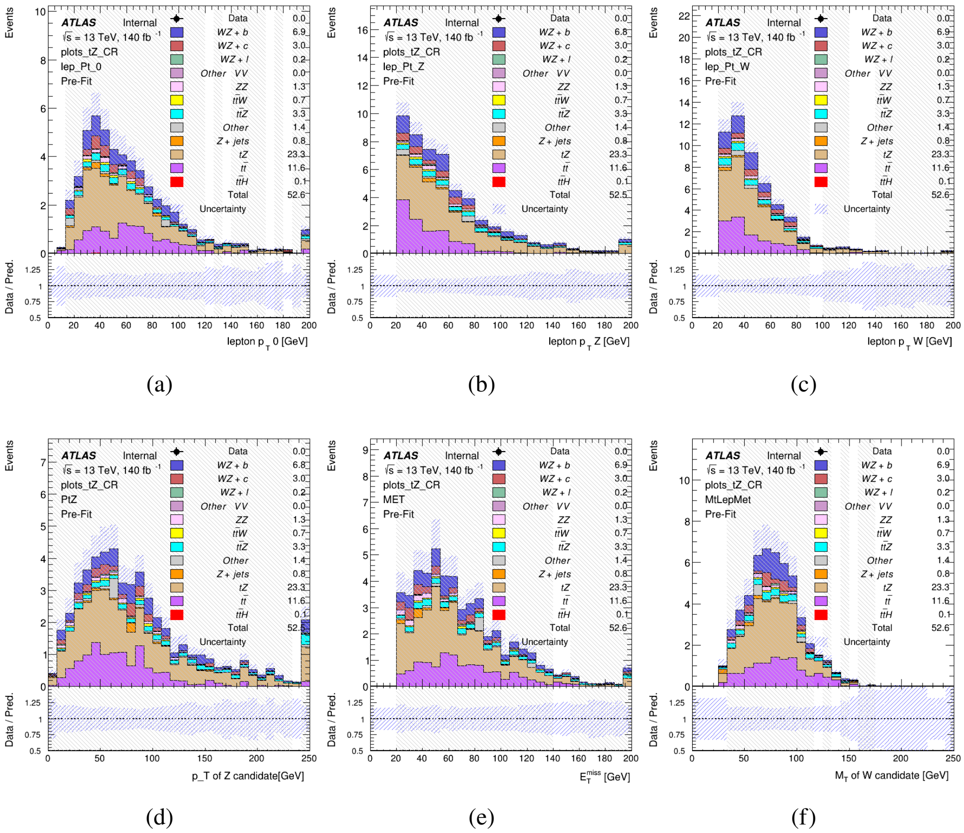
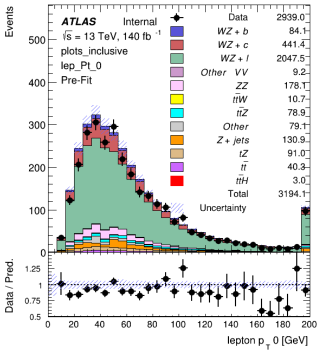
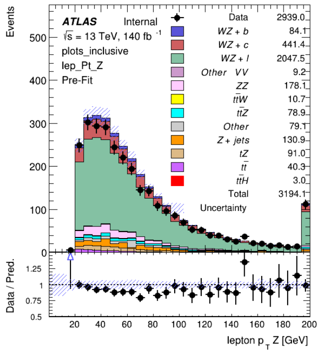


Figure 9: Comparisons between the data and MC distributions in the preselection region for (a) the p_T of the opposite sign lepton, (b) the p_T of the other lepton from the Z candidate, (c) the p_T of the lepton from the W candidate, (d) the p_T of the Z candidate, (e) the E_T^{miss}, and (f) the m_T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

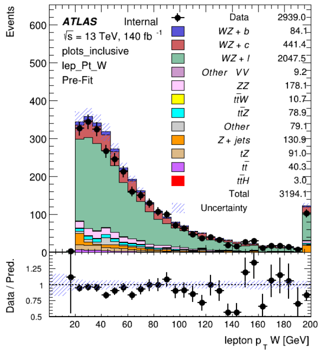
WZ Fit Region - 2j Inclusive



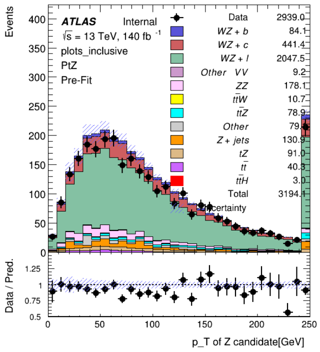
(a)



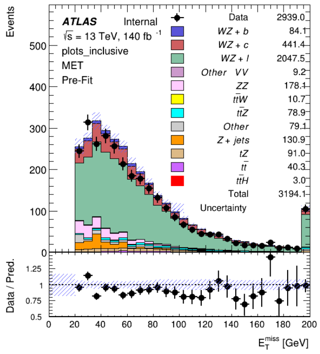
(b)



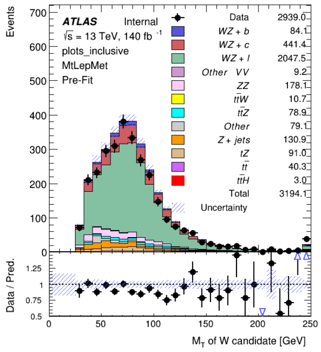
(c)



(d)



(e)



(f)

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

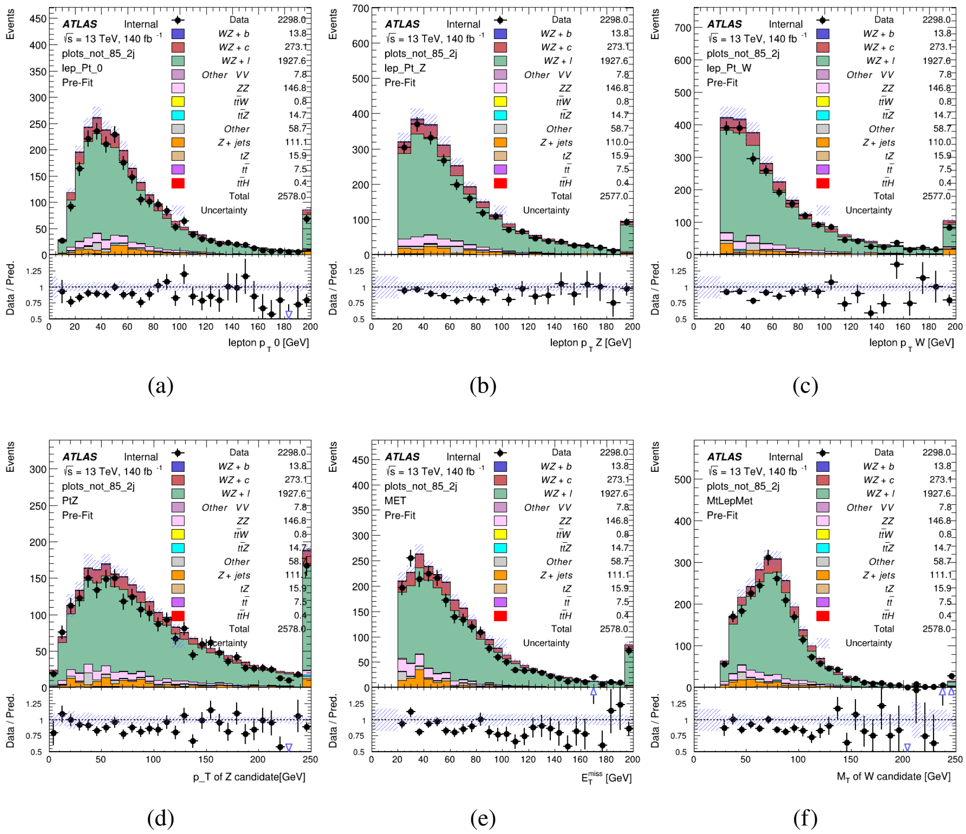
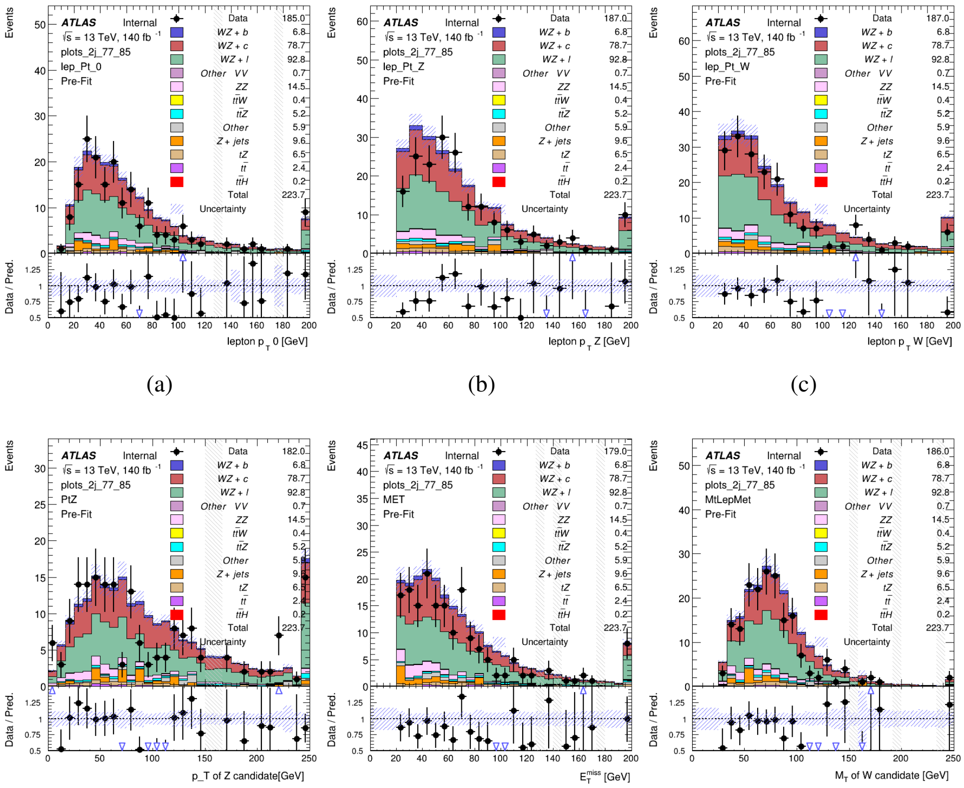


Figure 11: Comparisons between the data and MC distributions in the preselection region for (a) the p_T of the opposite sign lepton, (b) the p_T of the other lepton from the Z candidate, (c) the p_T of the lepton from the W candidate, (d) the p_T of the Z candidate, (e) the E_T^{miss}, and (f) the m_T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

WZ Fit Region - 2j 77-85% WP



(a) (b) (c)

(d) (e) (f)

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

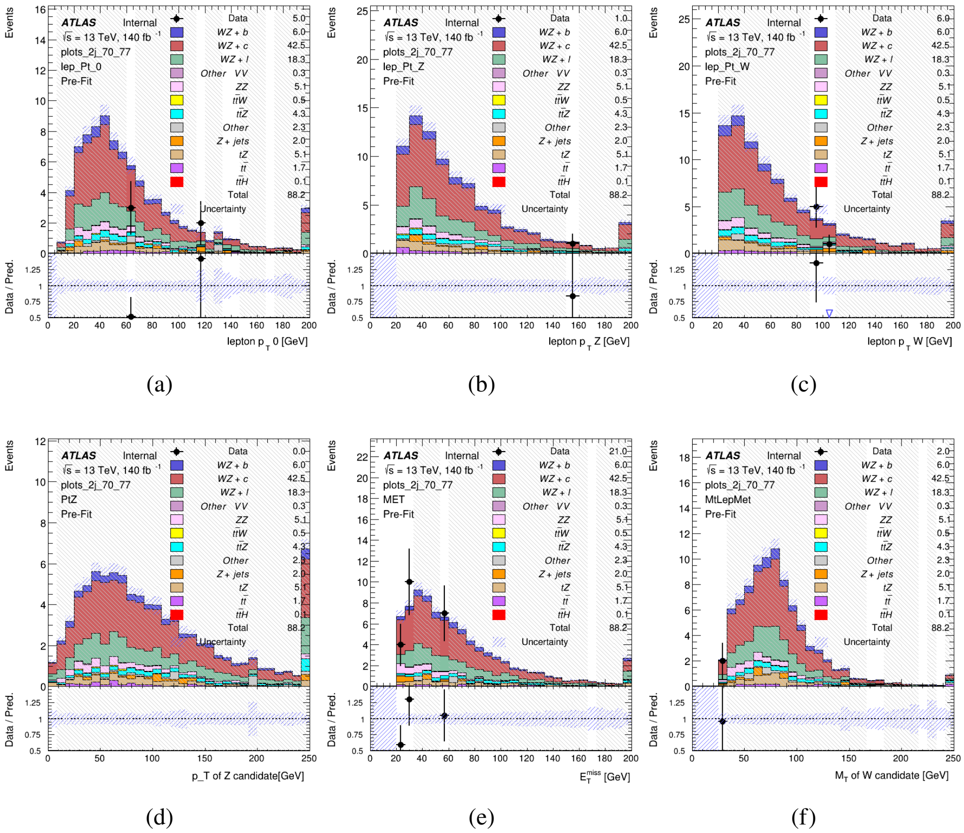


Figure 13: Comparisons between the data and MC distributions in the preselection region for (a) the p_T of the opposite sign lepton, (b) the p_T of the other lepton from the Z candidate, (c) the p_T of the lepton from the W candidate, (d) the p_T of the Z candidate, (e) the E_T^{miss}, and (f) the m_T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

WZ Fit Region - 2j 60-70% WP

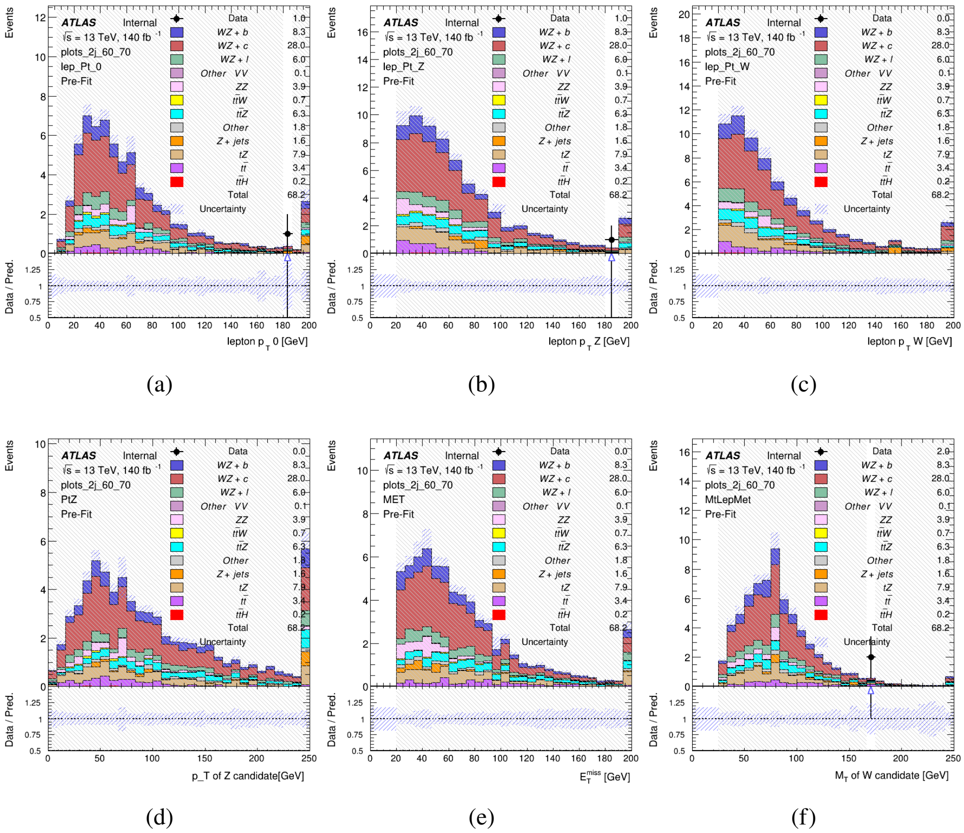
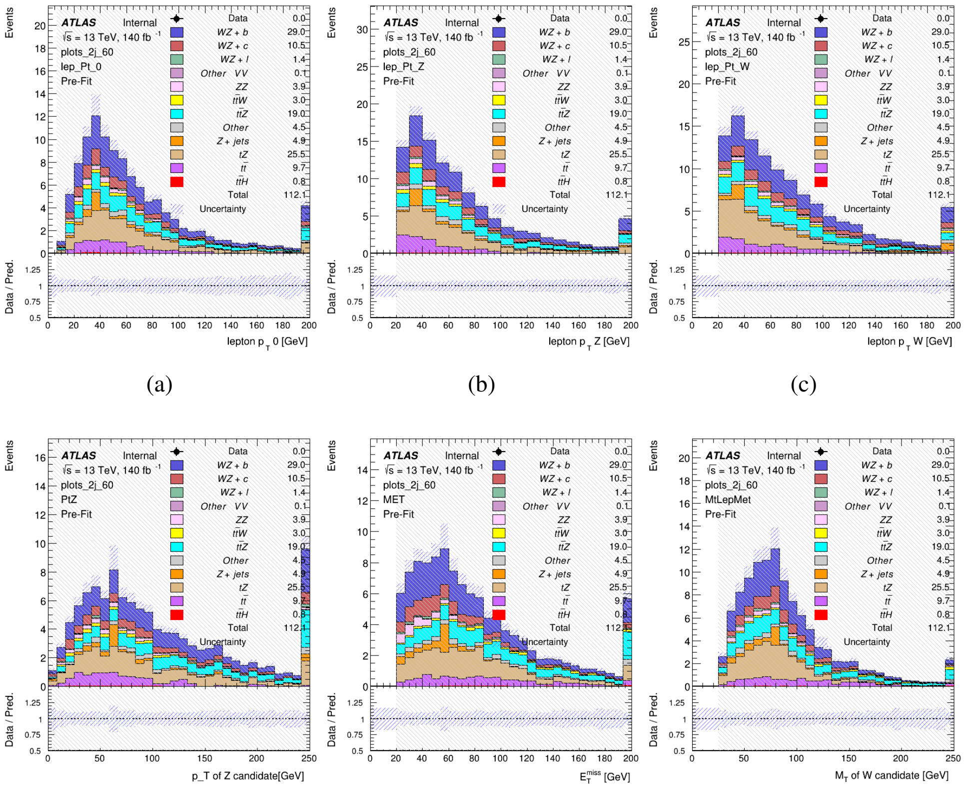


Figure 14: Comparisons between the data and MC distributions in the preselection region for (a) the p_T of the opposite sign lepton, (b) the p_T of the other lepton from the Z candidate, (c) the p_T of the lepton from the W candidate, (d) the p_T of the Z candidate, (e) the E_T^{miss}, and (f) the m_T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

WZ Fit Region - 2j 60% WP

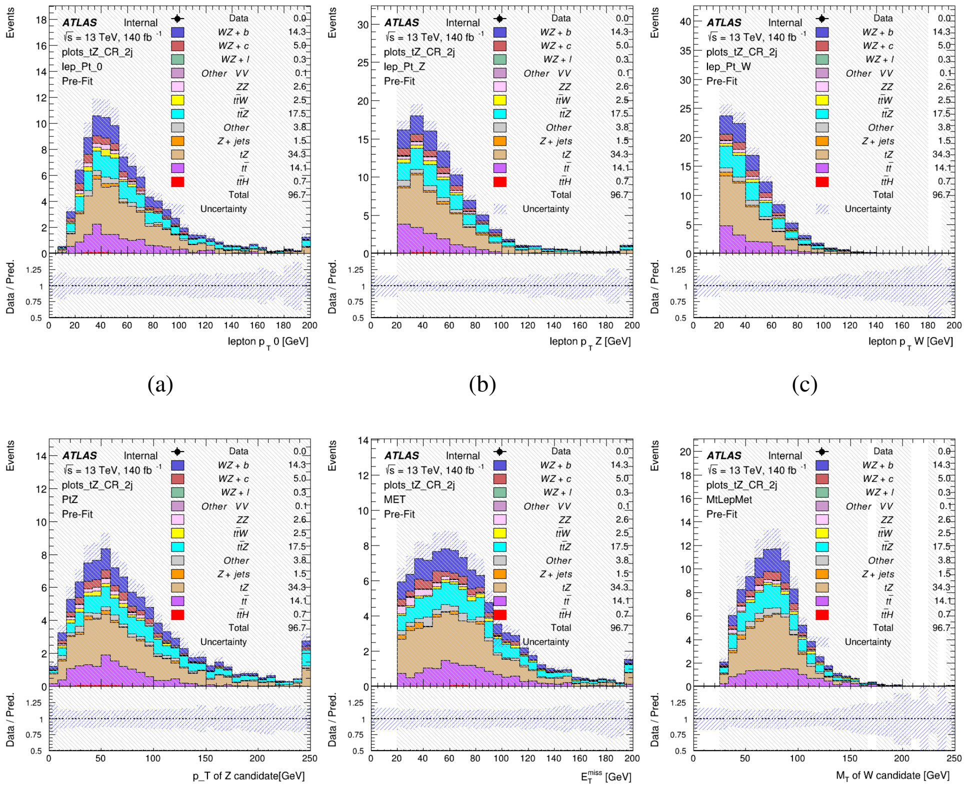


(a) (b) (c)

(d) (e) (f)

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



(a) (b) (c)

(d) (e) (f)

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 3l region when an additional non-prompt lepton appears in the event. The contribution of these processes is estimated with Monte Carlo simulations, which are validated using enriched validation regions.

5.3.1 $t\bar{t}$ Validation

$t\bar{t}$ events can produce two prompt leptons from the decay of each of the tops. These top decays produce two b-quarks, the decay of which can produce additional non-prompt leptons, which occasionally pass the event preselection. In order to validate that the Monte Carlo accurately simulates this process accurately, the MC prediction in a non-prompt $t\bar{t}$ enriched validation region is compared to data.

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

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

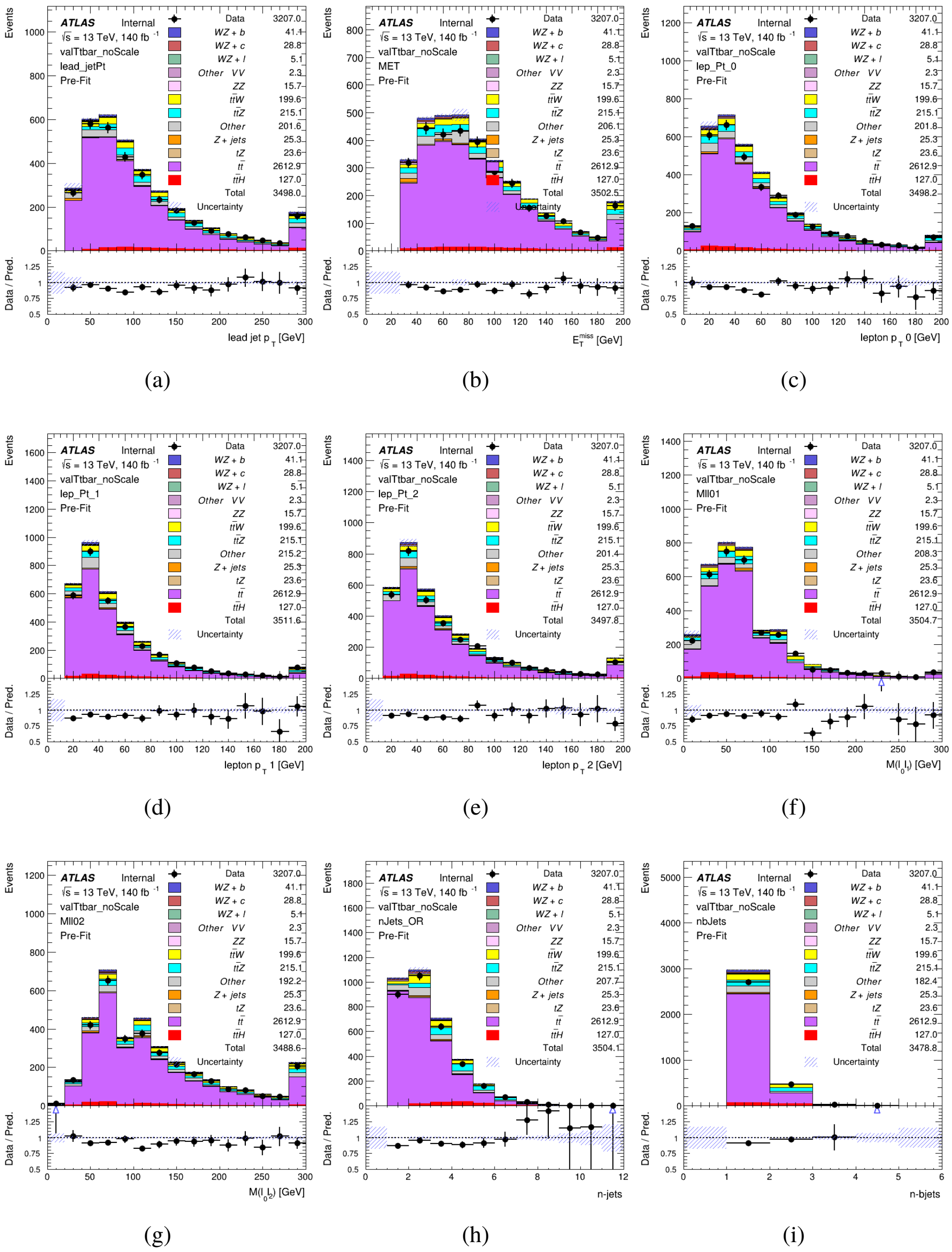


Figure 17: Comparisons between the data and MC distributions in the $t\bar{t}$ validation region for (a) the p_T of the leading jet, (b) the missing transverse energy, (c) the p_T of lepton 0, (d) p_T of lepton 1, (e) p_T of lepton 2, (f) the invariant mass of leptons 0 and 1, (g) the invariant mass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

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

259 MC prediction. Plots showing the kinematics of the $t\bar{t}$ VR after this correction factor has been
260 applied are shown in figure **??**.

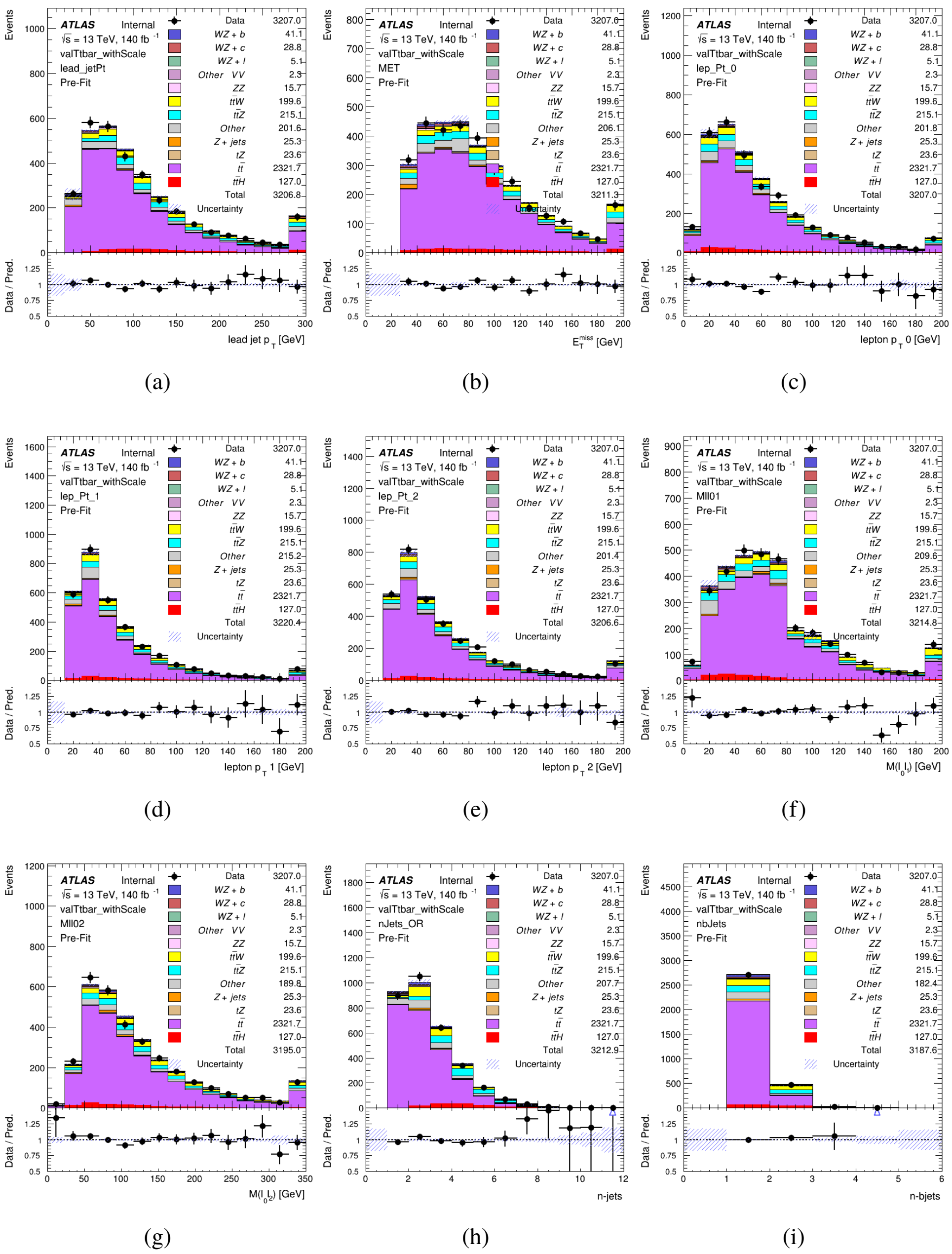


Figure 18: Comparisons between the data and MC distributions in the $t\bar{t}$ validation region after the correction factor has been applied for (a) the p_T of the leading jet, (b) the missing transverse energy, (c) the p_T of lepton 0, (d) p_T of lepton 1, (e) p_T of lepton 2, (f) the invariant mass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

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

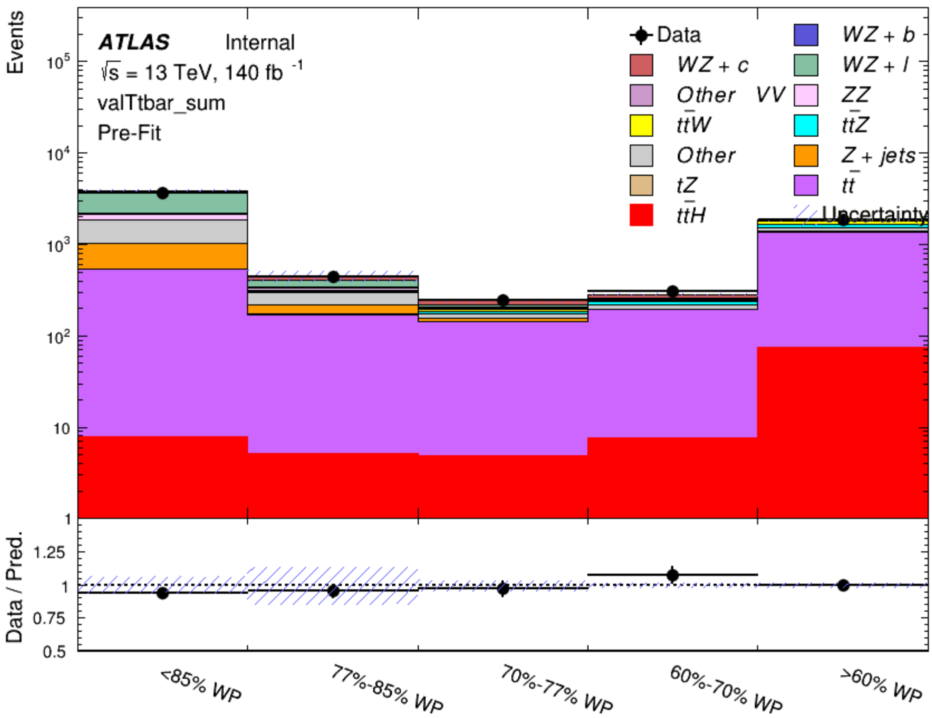


Figure 19: Data and MC comparisons for each DL1r WP for both 1-jet and 2-jet regions, after the $t\bar{t}$ VR selection and correction factor have been applied

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

5.3.2 Z+jets Validation

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

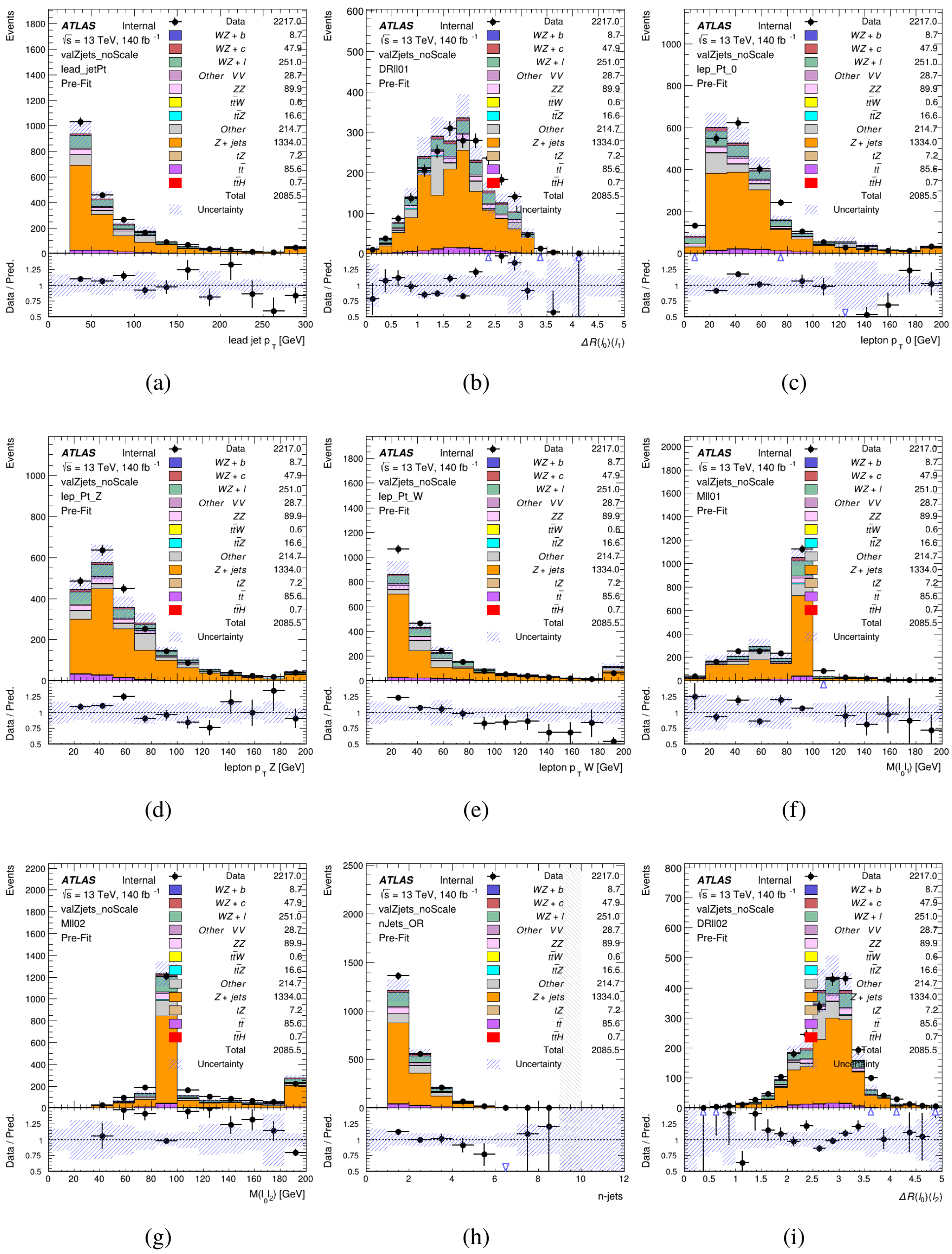


Figure 20: Comparisons between the data and MC distributions in the Z+jets validation region for (a) the p_T of the leading jet, (b) ΔR between leptons 0 and 1, (c) the p_T of lepton 0, (d) p_T of SS lepton from the Z candidate, (e) p_T of the SS lepton from the W candidate, (f) the invariant mass of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

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

of the p_T spectrum of the lepton from the W is found to differ. To account for this discrepency, a variable correction factor is applied to Z+jets. χ^2 minimization of the lepton 2 p_T spectrum is performed to derive a correction factor of $1.53 - 6.6 * 10^{-6}(\mathtt{lep\_Pt\_}W)$. Kinematic plots of the Z + jets validation region after this correction factor has been aplied are shown in figure **??**.

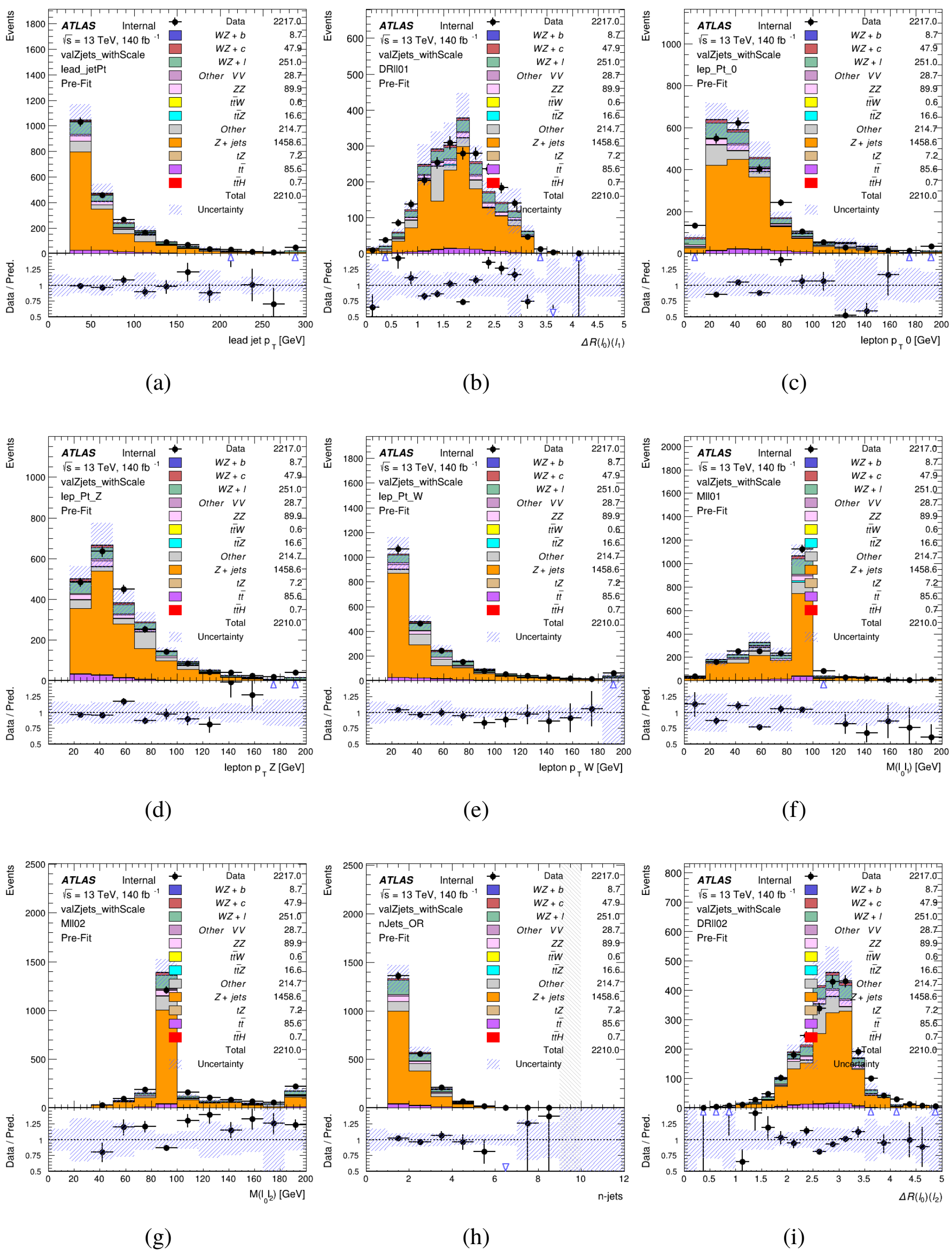


Figure 21: Comparisons between the data and MC distributions in the Z+jets validation region after the correction factor has been applied for (a) the p_T of the leading jet, (b) ΔR between leptons 0 and 1, (c) the p_T of lepton 0, (d) p_T of SS lepton from the Z candidate, (e) p_T of the SS lepton from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

The modeling is further validated by looking at the yield in the Z+jets VR for each DL1r WP, giving a clearer correspondence to the signal regions used in the fit. Each region shown in figure **??** requires one or more jets pass the listed WP, with no jets passing the next highest WP.

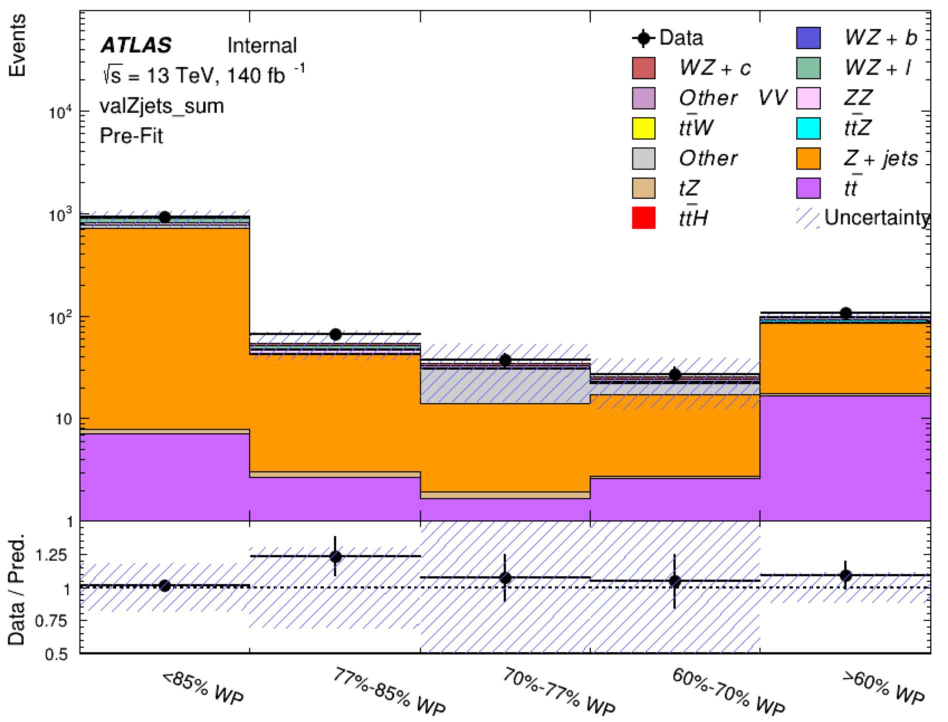


Figure 22: Data and MC comparisons for each DL1r WP for both 1-jet and 2-jet regions, after the Z+jets VR selection and correction factor have been applied

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

6 tZ Interference Studies and Separation Multivariate Analysis

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* + heavy flavor from tZ.

Separation between tZ and *WZ* + 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 + 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 + heavy flavor.

6.1 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^{miss}. Therefore, the E_T and ϕ of the neutrino are taken from the E_T^{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:

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

Expanding this out into components, this equation gives:

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

This equation gives two solutions for $p_{\nu z}$. For cases where only one of these solutions is real, that is taken as the value of $p_{\nu z}$. 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^{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 **??**.

6.2 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, $M(ll)$
- E_T^{miss}

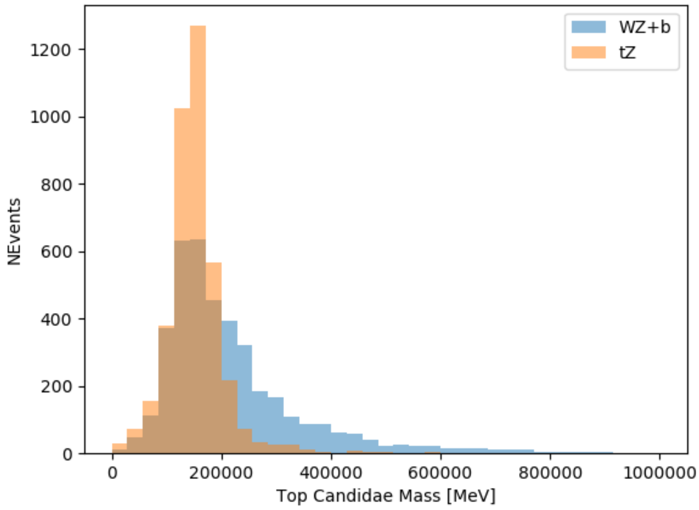


Figure 23: Reconstructed top mass distributions for tZ and WZ + b, measured in MeV.

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

The training samples included only events meeting the requirements of the 1-jet, >60% region, i.e. passing all the selection described in section **??** 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 **??**.

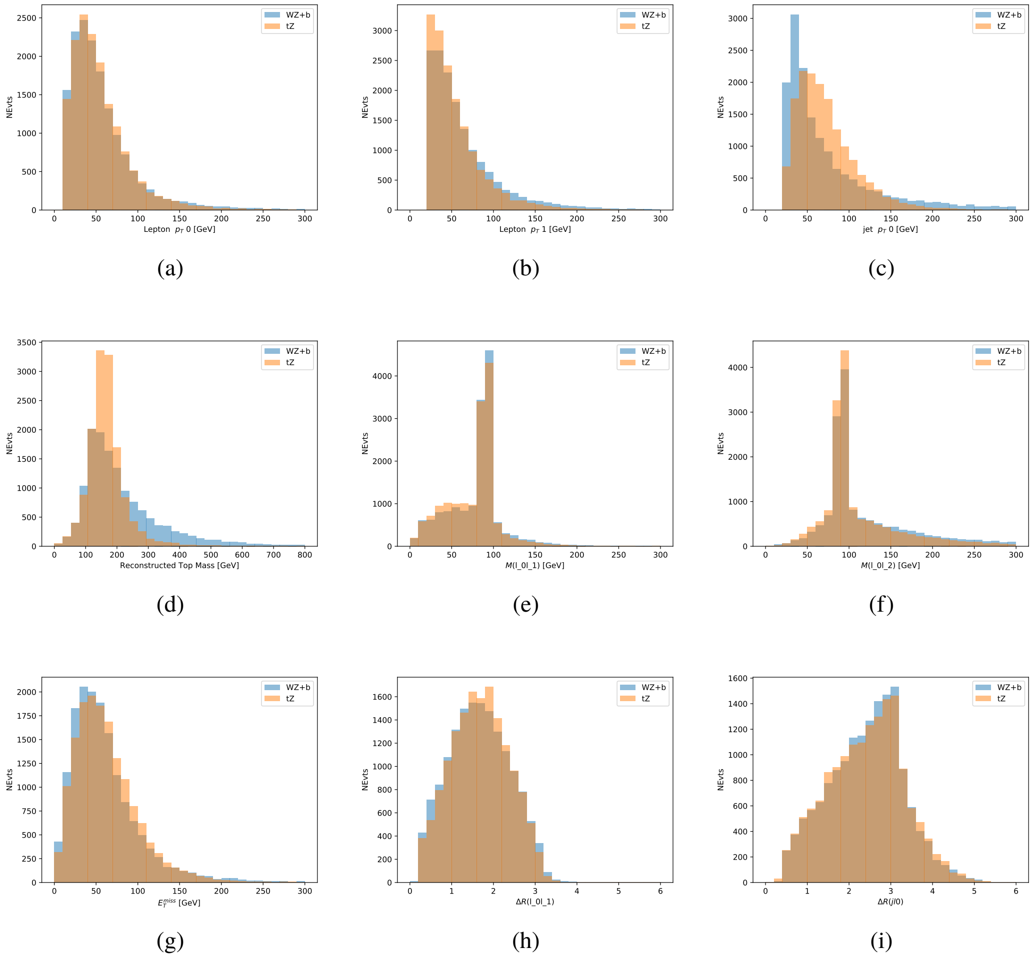
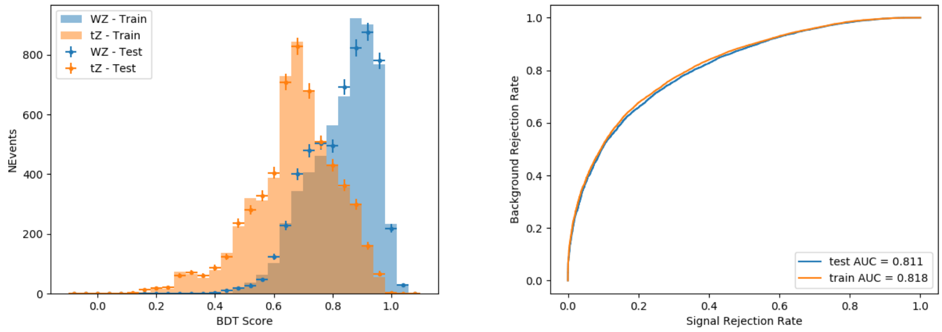


Figure 24: Distribution of input features of the BDT for signal (WZ) and background (tZ). Both are scaled to an equal number of events. (a), (b) and (c) show the p_T of lepton 0, lepton 1, and the jet, (d) show the reconstructed top mass, (e) and (f) show the invariant mass of leptons 0 and 1, leptons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

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

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



(a) (b)

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

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

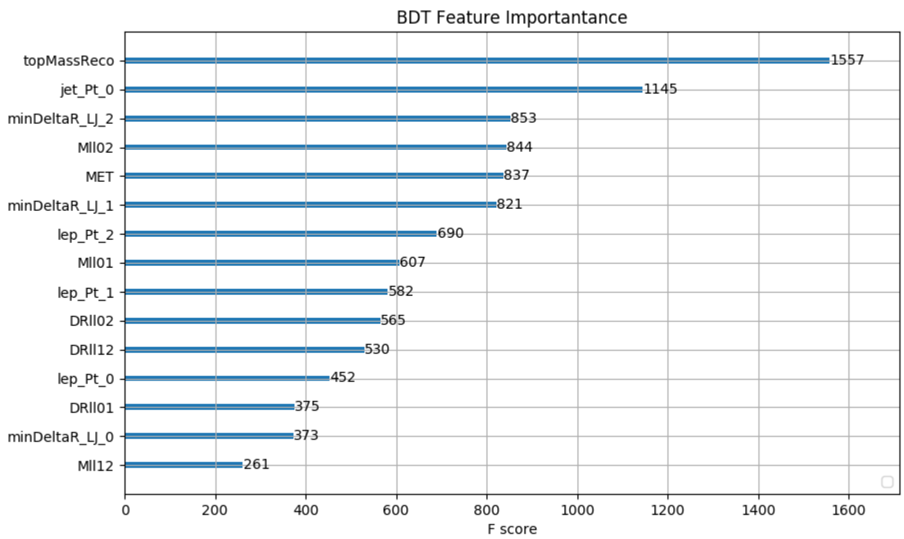


Figure 26: Relative importance of each input feature in the model.

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

7 Systematic Uncertainties

The systematic uncertainties that are considered are summarized in table **??**. 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 **??**.

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

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^{miss}. The sources which contribute to the uncertainty in the jet energy scale (JES) [13] are decomposed into uncorrelated components and treated as independent sources in the analysis. The CategoryReduction model is used to account for JES uncertainties, which decomposes the uncertainties into 30 nuiscance parameters included in the fit. The SimpleJER model is used to account for jet energy resolution (JER) uncertainties, and 8 JER uncertainty components uncluded as NPs in the fit.

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.

The full list of systematic uncertainties considered in the analysis is summarized in tables **??**, **??** and **??**.

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_	p_T Correction
		JET_EffectiveNP_1-19	p_T Correction
	η 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 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.

Theoritical uncertainties applied to MC predictions, including cross section, PDF, and scale uncertainties are taken from theory calculations, with the exception of non-prompt and diboson backgrounds. The cross-section uncertainty on tZ is taken from [**tZ_paper**]. Derivation of the non-prompt background uncertainties, Z+jets and $t\bar{t}$, are explained in detail in section **??**.

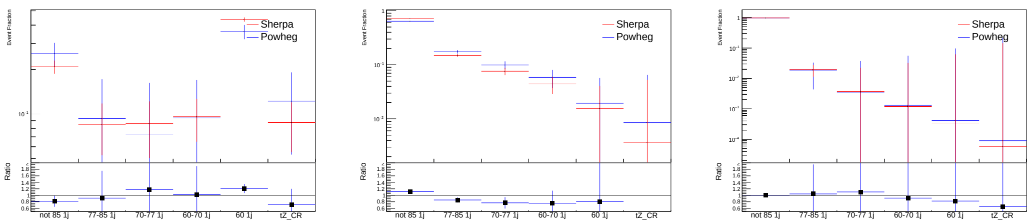
The other VV + heavy flavor processes (namely VV+b and VV+charm, which primarily comprise of ZZ events) are also poorly understood, because these processes involve much of the same physics as WZ + heavy flavor, and have also not been measured. Therefore, a conservative 50% uncertainty is applied to those samples.

The theory uncertainties applied to the predominate background estimates are summarized in table **??**.

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
$t\bar{t}$ 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)	± 50
VV + light (Sherpa 2.2.1)	± 6
$t\bar{t}$	± 10
Z + jets	± 20

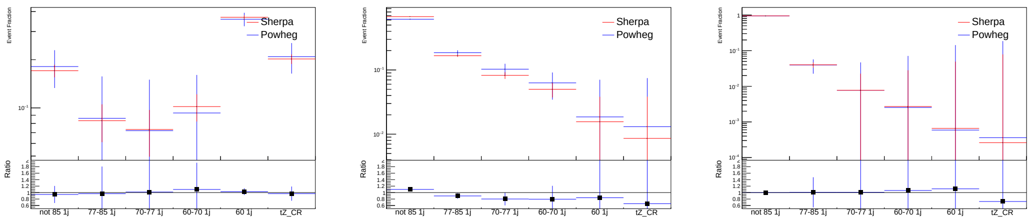
Table 12: Summary of theoretical uncertainties for MC predictions in the analysis.

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



(a) (b) (c)

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.



(a) (b) (c)

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.

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

8 Results

A separate maximum-likelihood fit is performed over the 1-jet and 2-jet fit regions in order to 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 **??**. The parameters $\mu_{WZ+\mathrm{b}}$, $\mu_{WZ+\mathrm{charm}}$, $\mu_{WZ+\mathrm{light}}$, where $\mu = \sigma_{\mathrm{observed}}/\sigma_{SM}$, are extracted from the fit.

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

8.1 1-jet Fit Results

The results of the fit are currently blinded.

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

	1j, <85% WP	1j, 77%-85% WP	1j, 70%-77% WP
$WZ + b$	11.10±1.5	4.68±0.372802	4.61±0.417424
$WZ + b$	0.62±0.0201084	0.35±0.0122749	0.28±0.0101779
$WZ + c$	318.25±22.0592	80.74±6.26045	43.26±3.60719
$WZ + c$	7.19±0.232501	1.98±0.0639804	1.04±0.0332902
$WZ + l$	4 017.60±252.306	90.52±13.0391	17.30±3.08617
$WZ + l$	70.71±2.33476	1.82±0.1109	0.39±0.0443128
Other VV	0.09±0.0880758	$1.00 \cdot 10^{-06}$±0.220715	$1.00 \cdot 10^{-06}$±0.014
Other VV	0.67±0.377189	0.08±0.119764	0.03±0.0287662
Other VV	6.20±0.642919	0.19±0.0699289	0.02±0.00800664
ZZ	1.33±nan	0.69±nan	0.49±nan
ZZ	10.27±nan	2.40±nan	1.45±nan
ZZ	336.48±nan	17.81±nan	4.33±nan
$t\bar{t}W$	1.09±0.206622	0.23±0.0687169	0.28±0.0786439
$t\bar{t}Z$	6.83±1.18797	1.41±0.253789	0.99±0.189632
$t\bar{t}lllowmass$	0.04±0.024398	$1.00 \cdot 10^{-06}$±0.00554059	0.01±0.0118307
rare Top	0.14±0.0438576	0.04±0.0230025	0.04±0.0233746
Single top	$1.00 \cdot 10^{-06}$±1.00051e-06	$1.00 \cdot 10^{-06}$±1.00051e-06	$1.00 \cdot 10^{-06}$±1.000
Single top t	$1.00 \cdot 10^{-06}$±1.00051e-06	$1.00 \cdot 10^{-06}$±1.00051e-06	$1.00 \cdot 10^{-06}$±1.000
Single top s	$1.00 \cdot 10^{-06}$±1.00051e-06	$1.00 \cdot 10^{-06}$±1.00051e-06	$1.00 \cdot 10^{-06}$±1.000
Three top	0.00±0.000642123	$1.00 \cdot 10^{-06}$±1.11849e-06	$1.00 \cdot 10^{-06}$±1.118
Four top	$1.00 \cdot 10^{-06}$±1.00051e-06	$1.00 \cdot 10^{-06}$±1.00051e-06	$1.00 \cdot 10^{-06}$±1.000
$t\bar{t}WW$	0.04±0.0274542	0.01±0.0195713	$1.00 \cdot 10^{-06}$±1.000
$Z +$ jets	168.84±37.8442	8.87±1.93588	3.73±0.780807
low mass $Z +$ jets	2.56±2.11056	$1.00 \cdot 10^{-06}$±1.02153e-06	0.20±0.286559
$W +$ jets	$1.00 \cdot 10^{-06}$±1.00051e-06	$1.00 \cdot 10^{-06}$±1.00051e-06	$1.00 \cdot 10^{-06}$±1.000
$V + \gamma$	45.67±58.1897	1.98±2.40971	0.13±0.121788
tZ	24.25±4.33518	5.49±1.08562	4.12±0.833384
tW	1.37±0.82139	0.18±0.264118	$1.00 \cdot 10^{-06}$±0.152
WtZ	2.33±1.20478	0.55±0.308675	0.27±0.165516
VVV	12.38±0.451037	0.93±0.0576788	0.35±0.0334005
VH	40.19±6.2286	2.56±1.41069	0.90±0.767229
$t\bar{t}$	0.45±0.294906	0.07±0.0820946	$1.00 \cdot 10^{-06}$±1.122
$t\bar{t}$	12.14±1.59161	2.88±0.555655	2.46±0.503466
$t\bar{t}H$	0.24±0.0319849	0.05±0.00829275	0.04±0.00800996
Total	5 099.12±nan	226.54±nan	86.71±nan

Table 13: Yields of the analysis

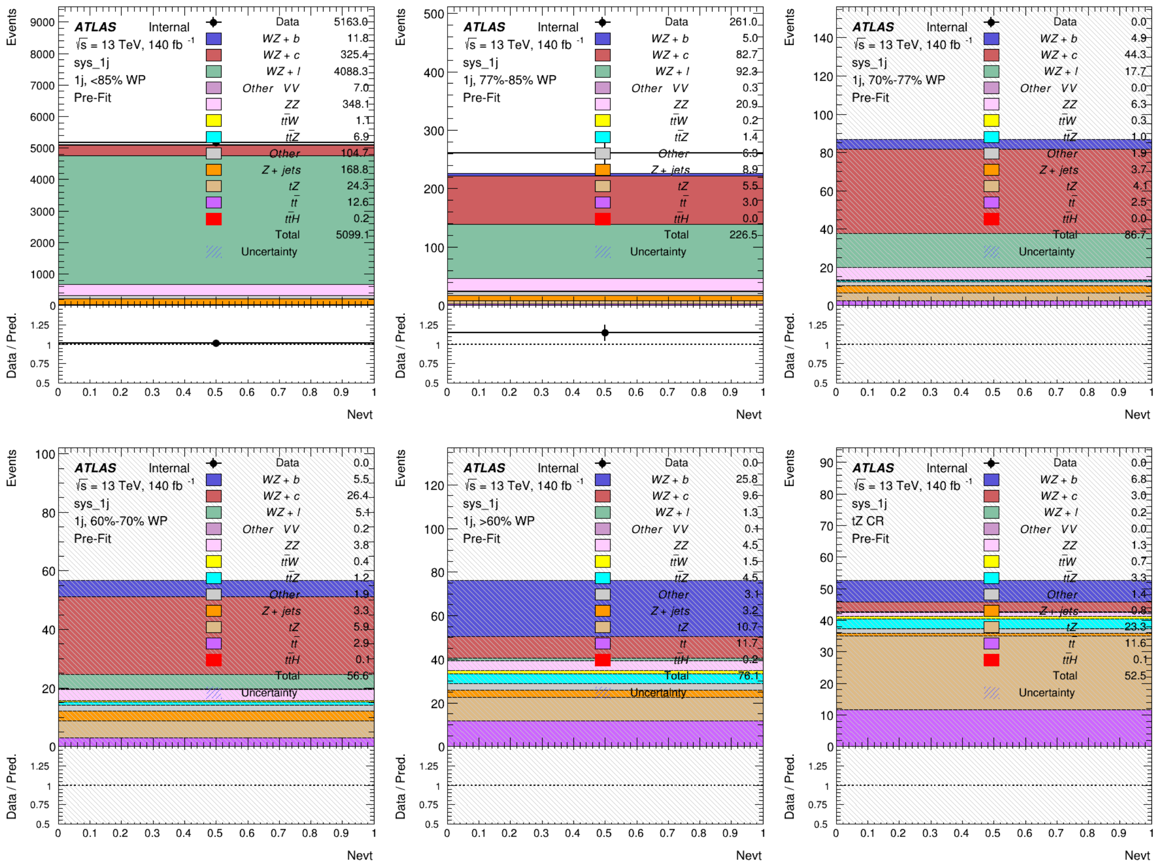


Figure 29: Data/MC yields in each of the 1-jet regions before the fit has been performed.

406 The post-fit yields in each region are summarized in figure **??**.

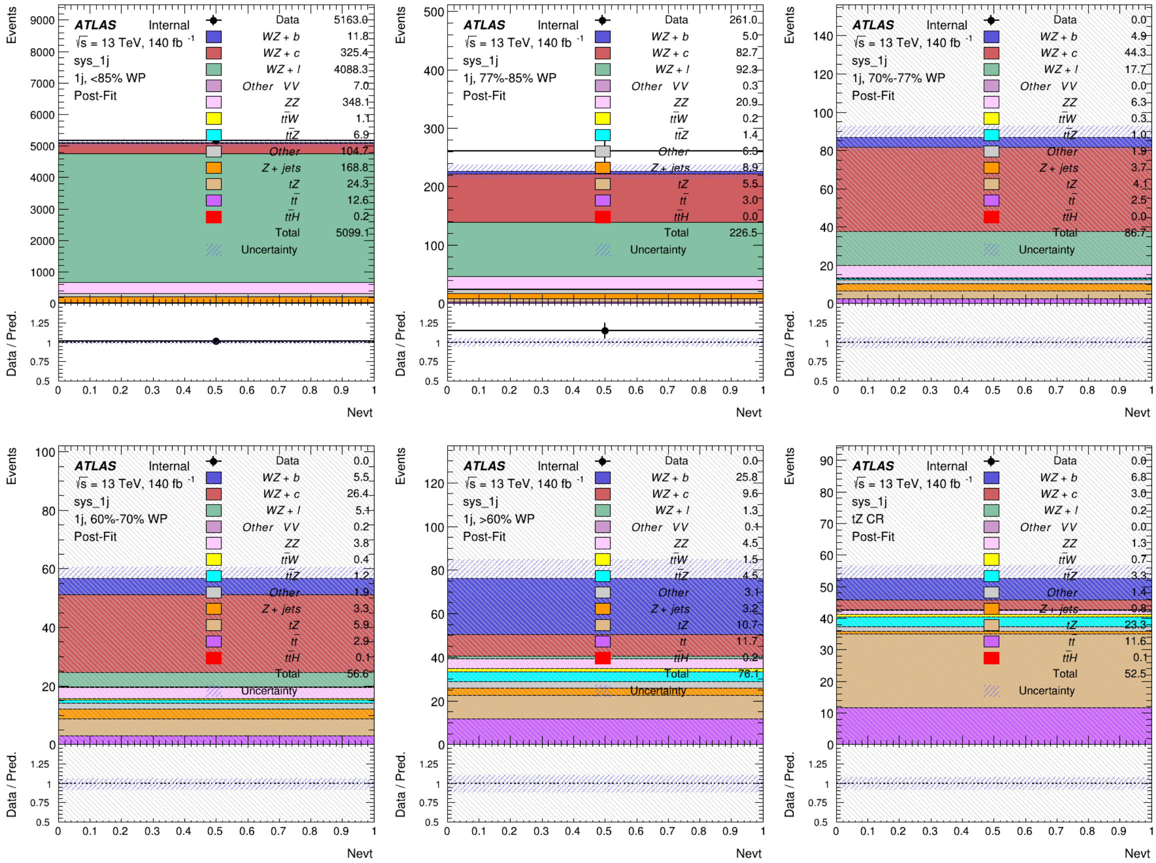


Figure 30: Data/MC results in each of the 1-jet regions after the fit has been performed.

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

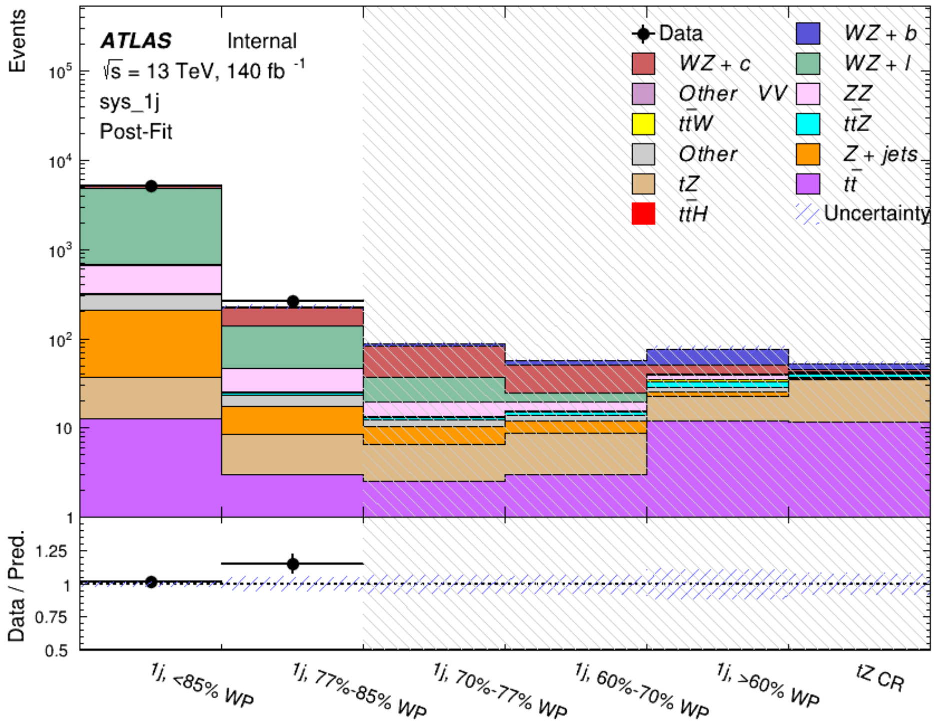


Figure 31: Post-fit summary of the 1-jet fit regions.

As described in section **??**, 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 NP 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 sources of systematic uncertainties is summarized in table **??**.

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-sction	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 $WZ + \mathrm{b}$ with exactly one associated jet.

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

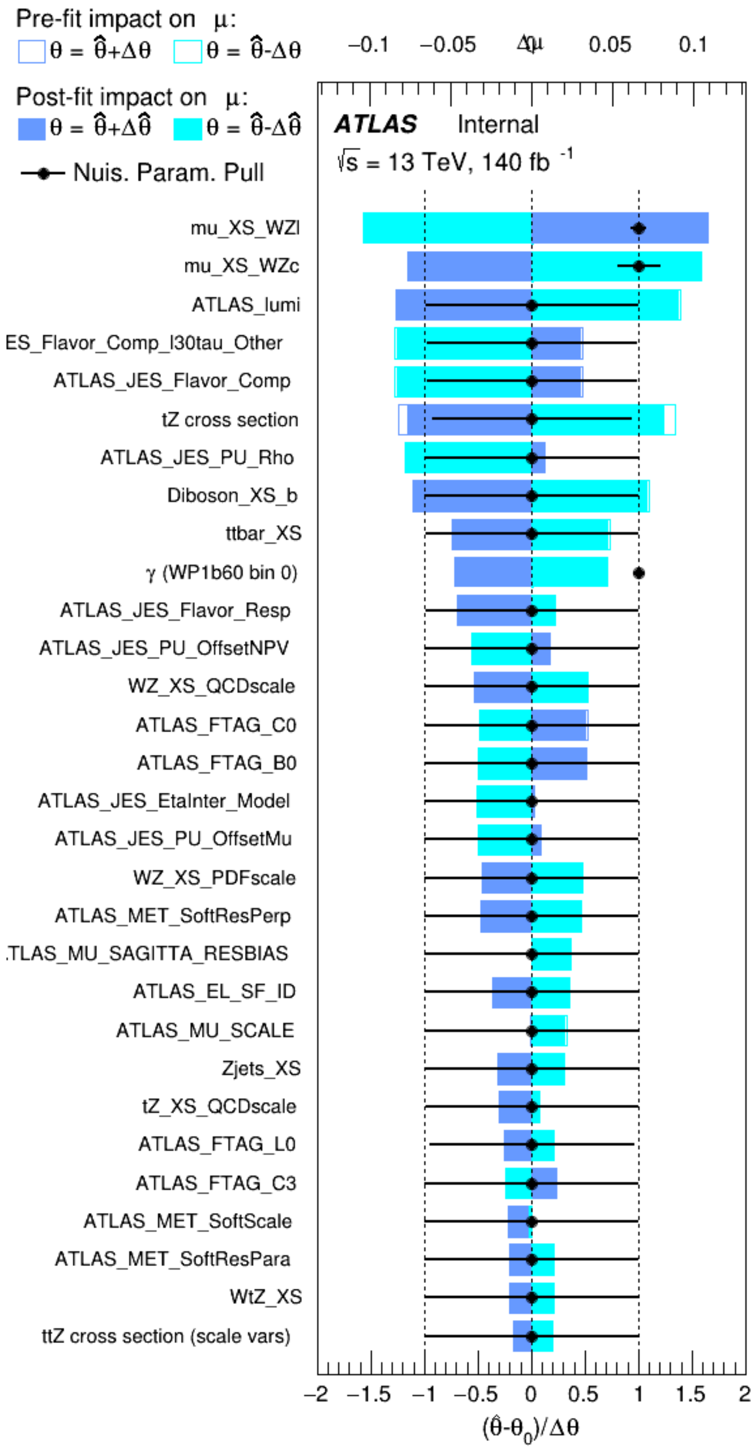


Figure 32: Impact of systematic uncertainties on the signal-strength of WZ + b for events with exactly one jet

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

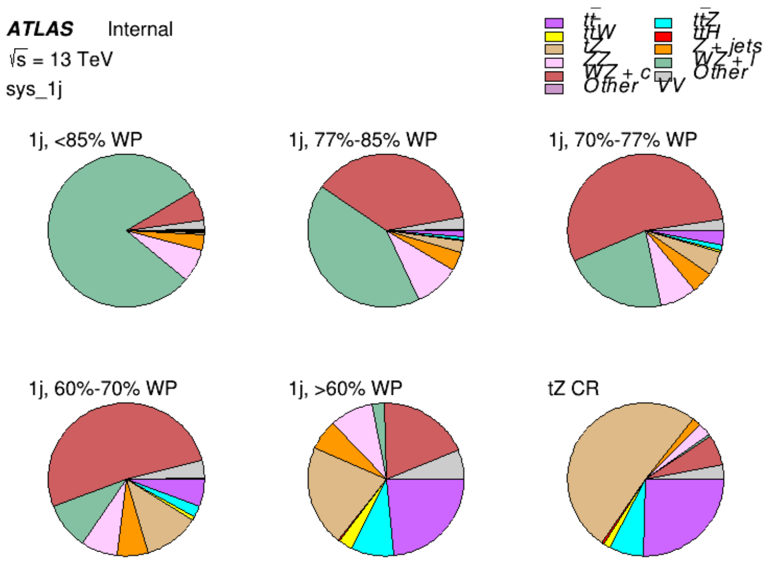


Figure 33: Post-fit background composition of the fit regions.

The correlations between these nuisance parameters are summarized in figure **??**.

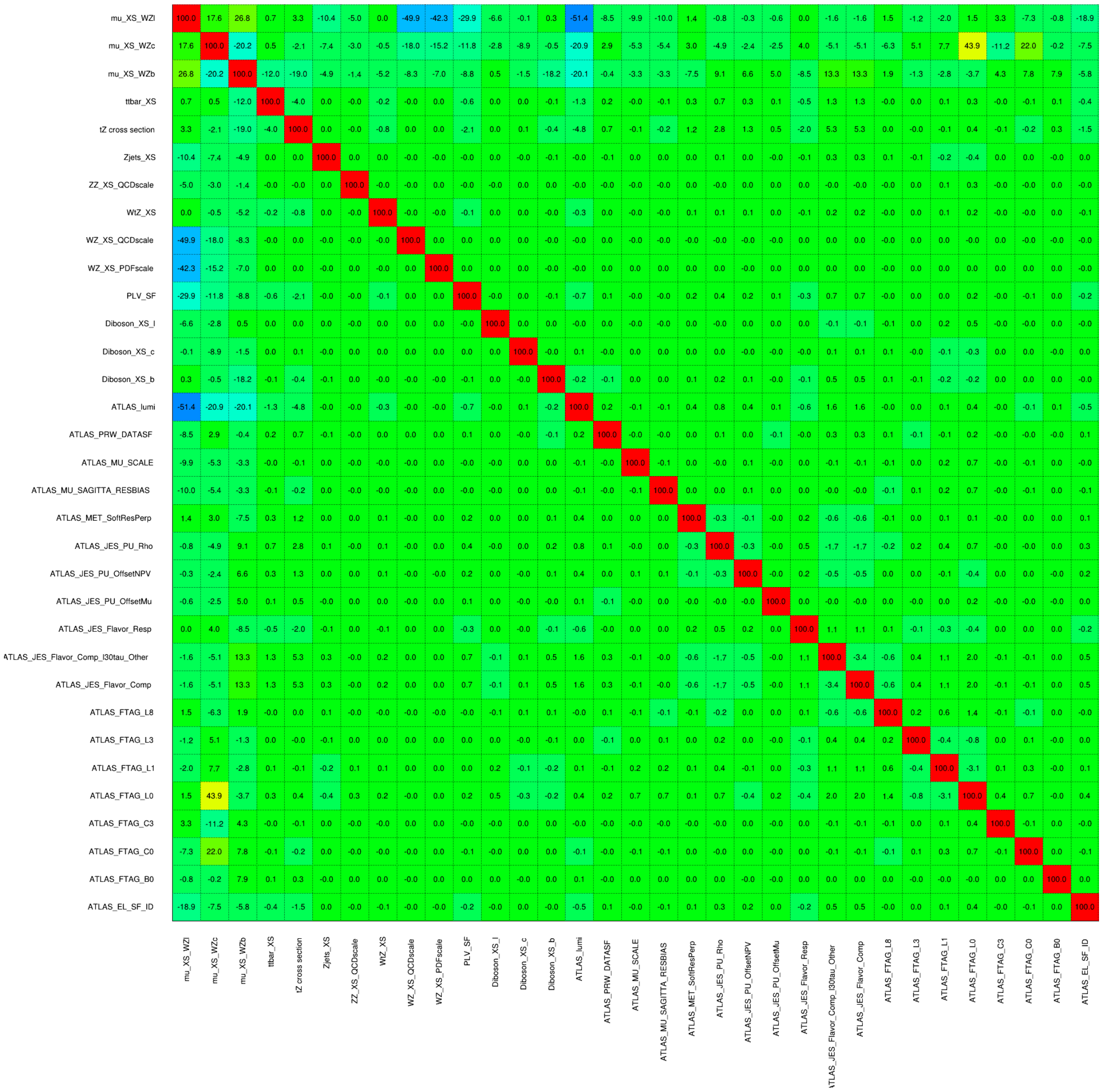


Figure 34: Correlations between nuisance parameters

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

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

The high correlation between the luminosity and $\mu_{WZ+\text{light}}$ arises from the fact that the uncertainty on $\mu_{WZ+\text{light}}$ 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.

8.2 2-jet Fit Results

The results of the fit are currently blinded.

Pre-fit yields in each of the 2-jet fit regions are shown in figure **??**.

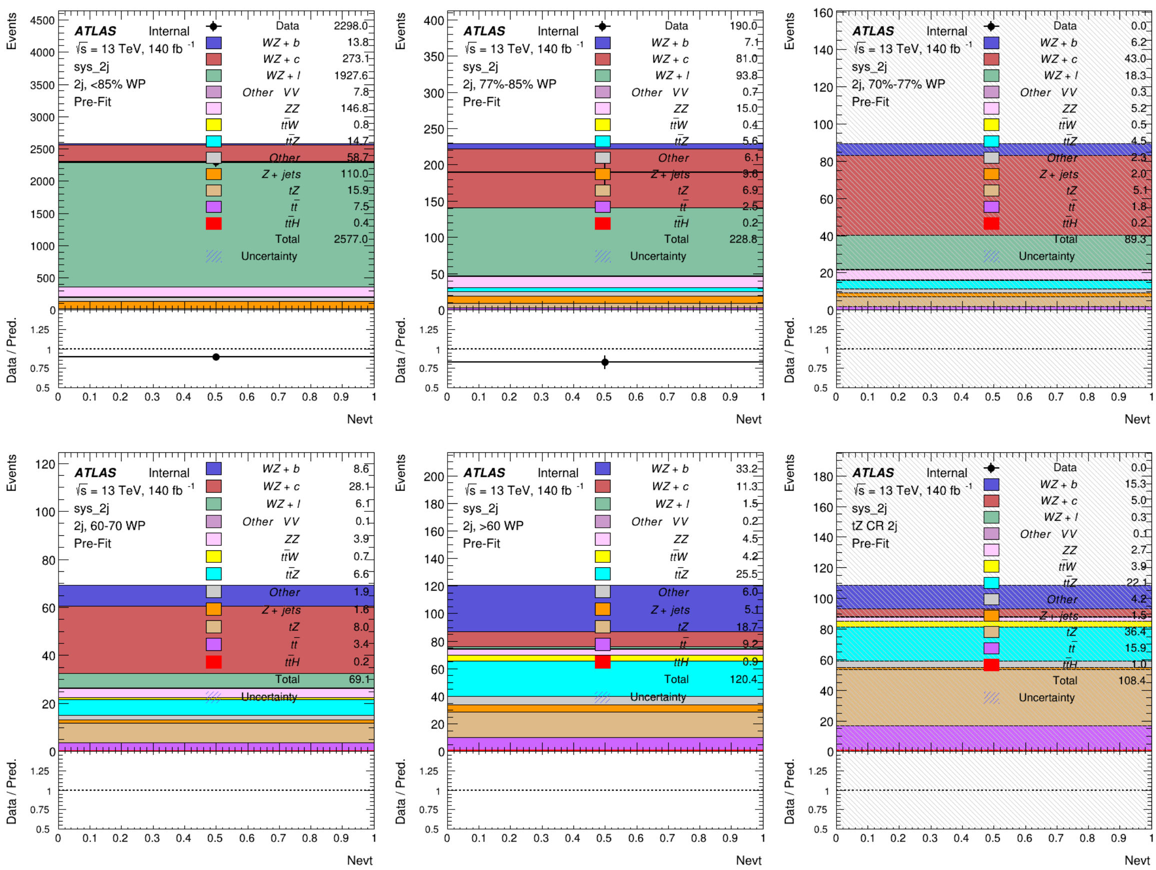


Figure 35: Data/MC yields in each of the regions in the 2-jet fit before the fit has been performed.

The post-fit yields in each region are summarized in figure **??**.

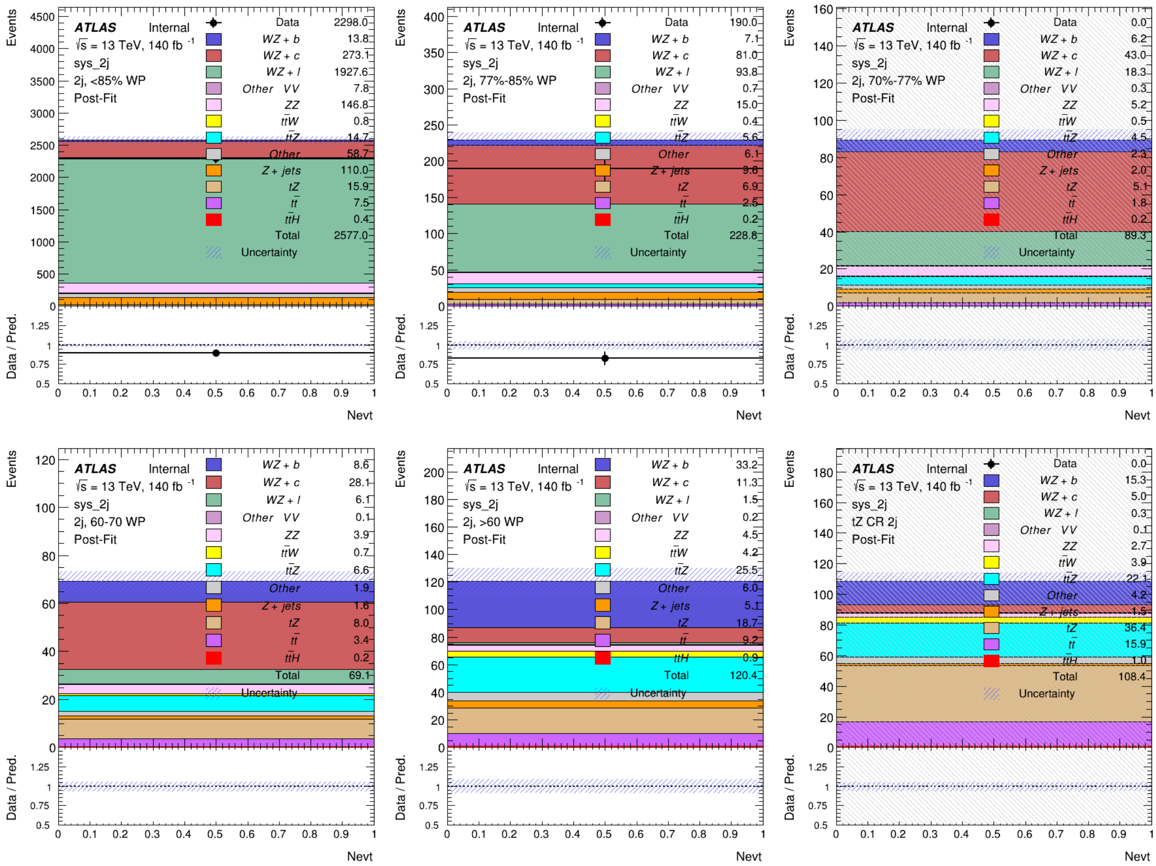


Figure 36: Data/MC results in each of the regions in the 2-jet fit after the fit has been performed.

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

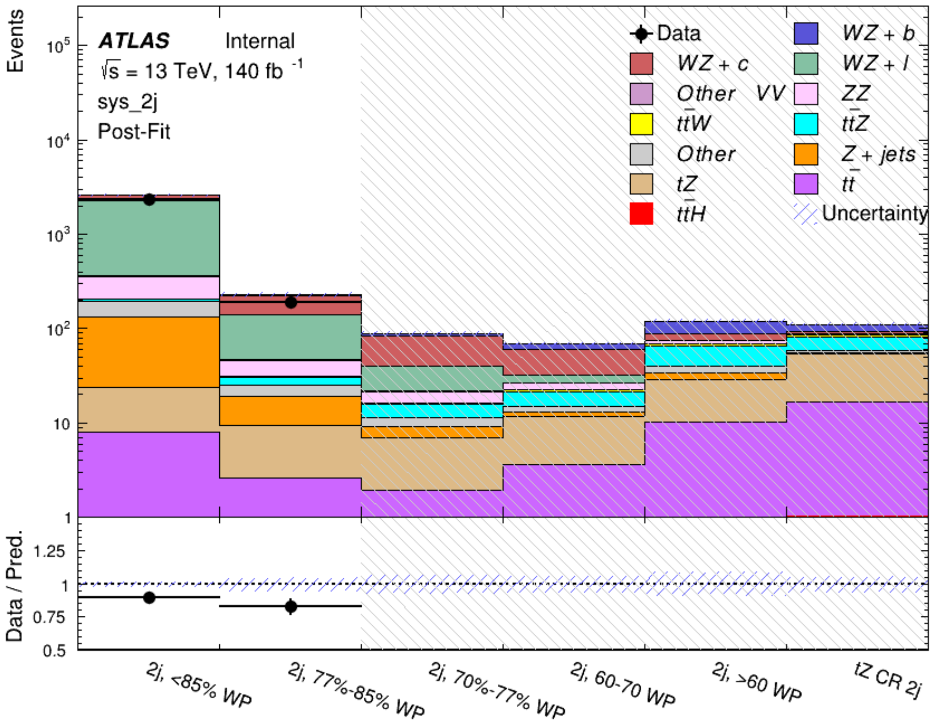


Figure 37: Post-fit summary of the fit over 2-jet regions.

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

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-sction	-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-sction	0.0207	-0.0202
Muon Trigger Scale Factor	0.019	0.0209
Total Systematic Uncertainty	0.3511	0.3679

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

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

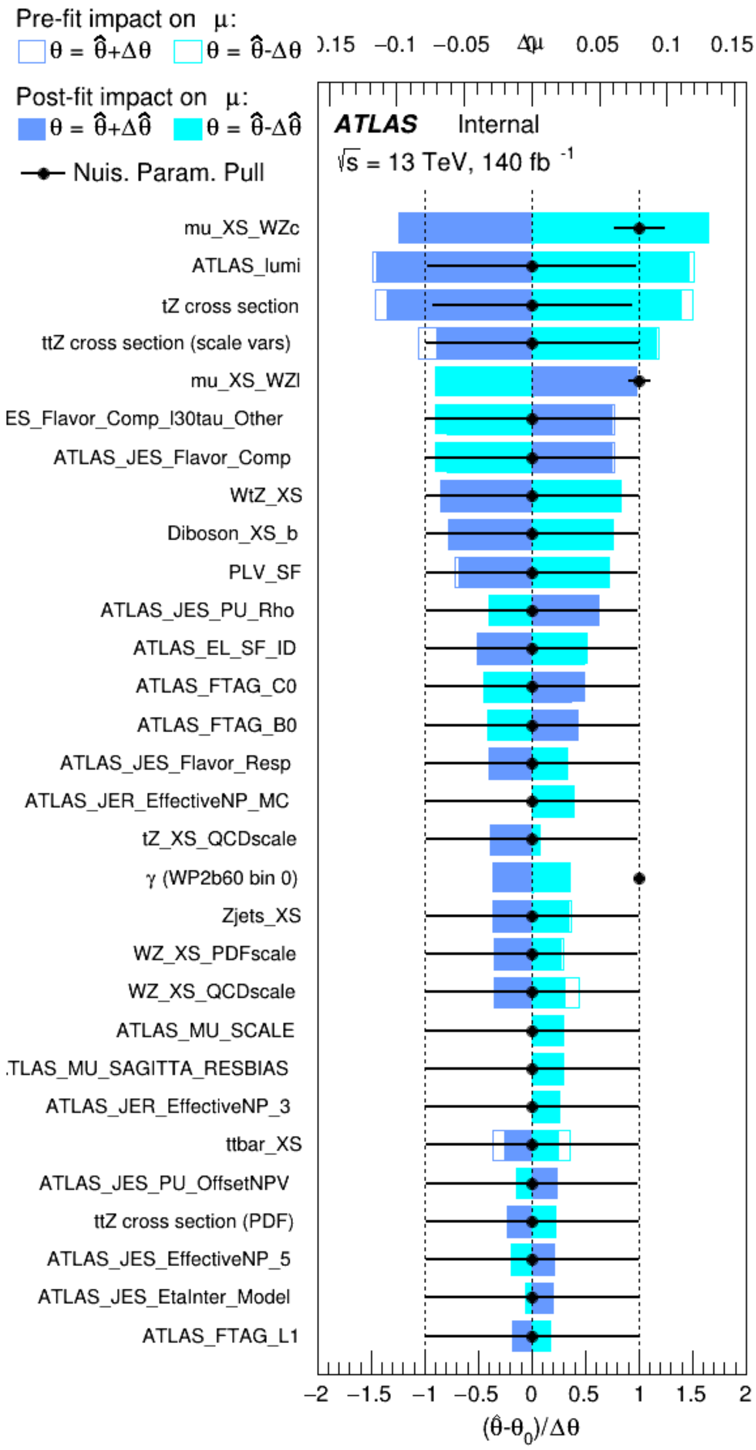


Figure 38: Impact of systematic uncertainties on the signal-strength of WZ + b in 2-jet events.

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

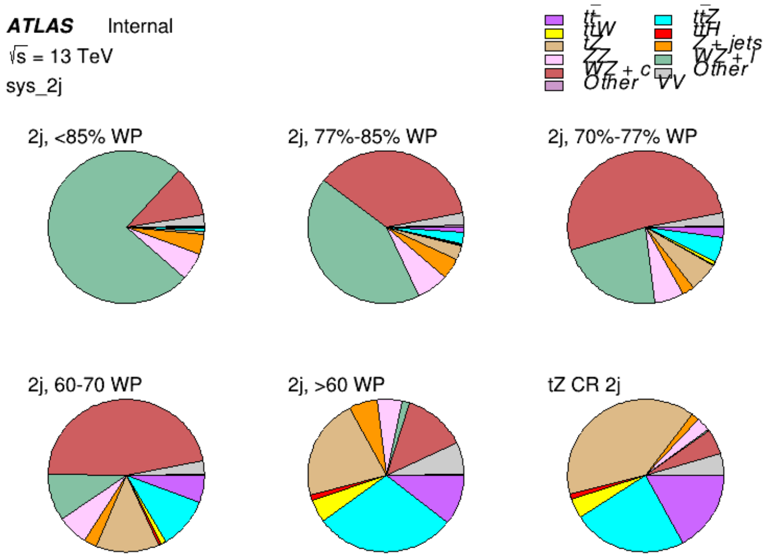


Figure 39: Post-fit background composition of the 2-jet fit regions.

The correlations between these nuisance parameters are summarized in figure **??**.

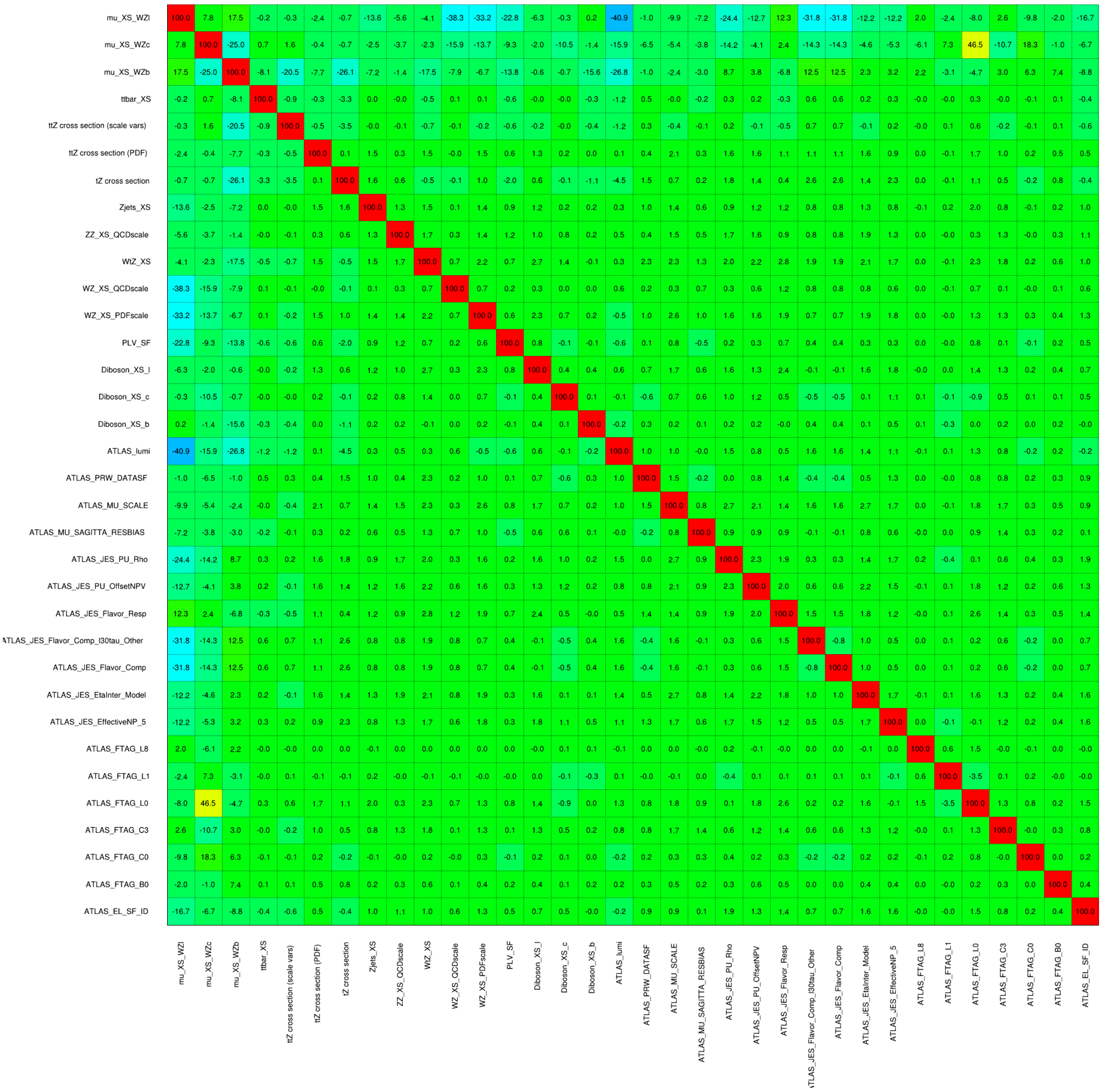


Figure 40: Correlations between nuisance parameters in the 2-jet fit

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

8.3 tZ Inclusive Results

While tZ is often considered as a distinct process form WZ + b,

9 Conclusion

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

Appendices

tZ Interference Studies

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.

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.

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

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

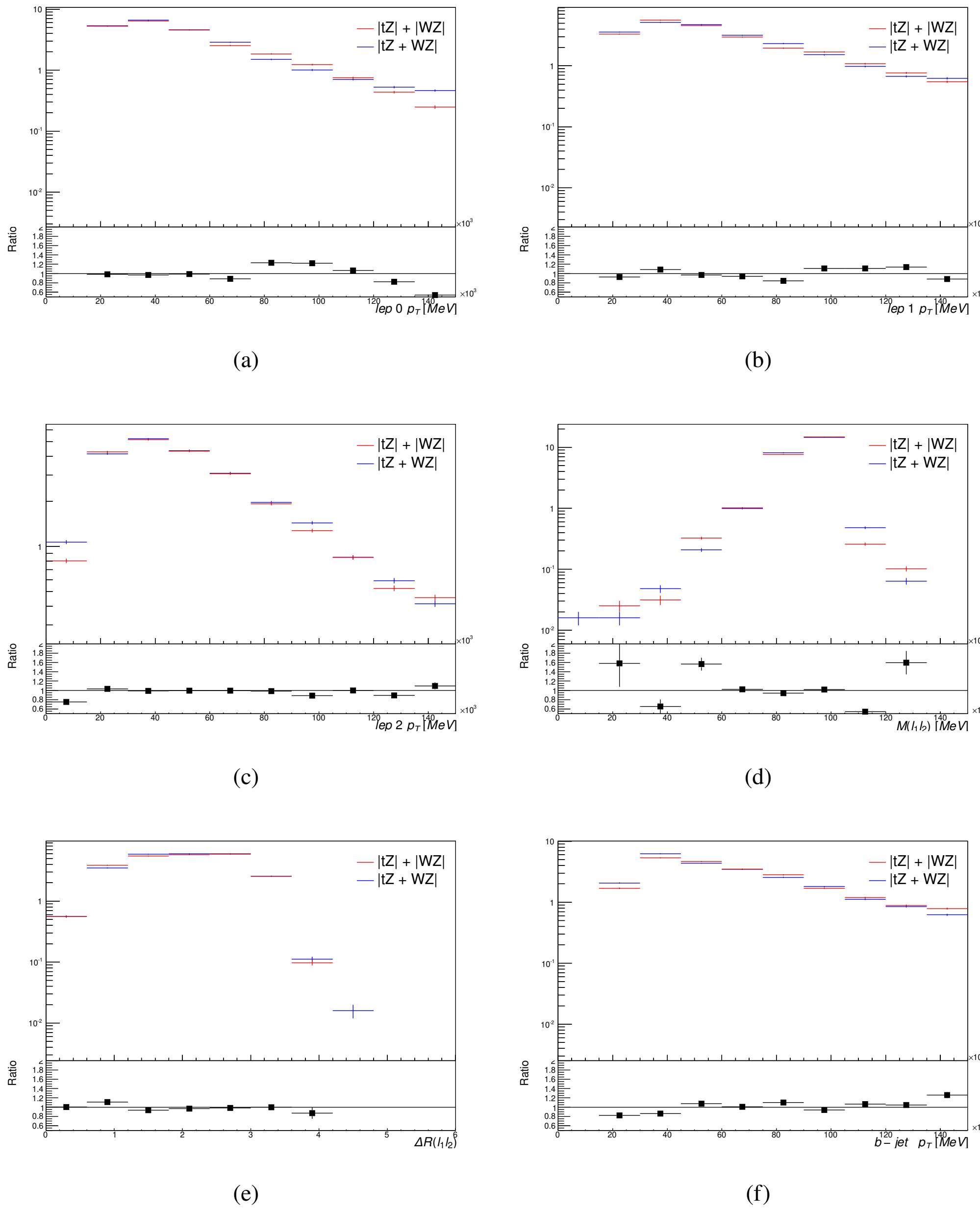


Figure 41: Comparisons between (a) the p_T of the lepton from the W, (b) and (c) show the p_T of the lepton from the Z, (d) the invariant mass of the Z-candidate, (e) ΔR of the leptons from the Z, and (f) the p_T of the b-jet, for WZ and tZ events generated with interference effects (blue) and without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

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

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.* B793 (2019), pp. 469–492. DOI: 10.1016/j.physletb.2019.05.012. arXiv: 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. arXiv: 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. arXiv: 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. arXiv: 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.

[12] ATLAS Collaboration. 'Luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC'. In: *Eur. Phys. J. C* 71 (2011), p. 1630. DOI: 10.1140/epjc/s10052-011-1630-5. arXiv: 1101.2185 [hep-ex].

[13] G. Aad et al. 'Jet energy resolution in proton-proton collisions at $\sqrt{s} = 7$ TeV recorded in 2010 with the ATLAS detector'. In: *The European Physical Journal C* 73.3 (Mar. 2013), p. 2306. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-013-2306-0. URL: https://doi.org/10.1140/epjc/s10052-013-2306-0.

[14] A. Collaboration. 'Performance of b -jet identification in the ATLAS experiment'. In: *Journal of Instrumentation* 11.04 (2016), P04008. URL: http://stacks.iop.org/1748-0221/11/i=04/a=P04008.