



ATLAS Paper
ATL-COM-PHYS-2019-962
24th May 2021



1

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

4 The ATLAS Collaboration¹, Aaron Webb^a, Peter Onyisi^a

5 ^a*Univ. of Texas at Austin*

6 A measurement of the cross-section for production of WZ with an associated heavy flavor jet
7 is performed using 140 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV from the ATLAS
8 experiment at the LHC. A measurement of the fully leptonic decay mode, $WZ \rightarrow \ell\nu\ell\ell$, is
9 performed in the fiducial phase space corresponding to $Z \rightarrow \ell^+\ell^-$ candidates with masses
10 between 81 GeV and 101 GeV. The cross-section of WZ + b-jets is measured to be $X \pm X \pm X$,
11 while the cross-section of WZ + charm is measured as X .

13	Contents	
14	1 Introduction	4
15	2 The ATLAS Detector	4
16	3 Data and Monte Carlo Samples	5
17	3.1 Data Samples	5
18	3.2 Monte Carlo Samples	5
19	4 Object Reconstruction	6
20	5 tZ Separation Classifier	8
21	6 tZ Interference Studies	9
22	7 Event Selection and Signal Region Definitions	10
23	7.1 Event Preselection	10
24	7.2 Fit Regions	11
25	7.3 Non-Prompt Lepton Estimation	11
26	8 Systematic Uncertainties	13
27	9 Results	15
28	9.1 Fit Procedure	15
29	9.2 Fiducial Region Definition	16
30	9.3 Results of the Simultaneous Fit	17
31	10 Conclusion	22
32	A Appendices	24
33	Appendices	24
34	A.1 Inclusive 1+2 Jet Cross Check	24
35	A.2 Alternate tZ Inclusive Fit	27
36	A.2.1 tZ Inclusive Fit	27
37	A.2.2 Floating tZ	27

³⁸ **List of contributions**

³⁹	Aaron Webb	Main analyser. Responsible for ntuple production, fits, and note writing.
	Peter Onyisi	Advisor to Aaron Webb. Analysis design and implementation strategy.

⁴⁰

1 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 multiple 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 inclusive 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.

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 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 $\sqrt{s} = 13 \text{ TeV}$, is used for this study.

Section 3 details the data and Monte Carlo (MC) samples used in the analysis. The reconstruction of various physics objects is described in Section 4. Section 7 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 5. 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 analysis reports a cross-section measurement of WZ+b and WZ+charm, along with their correlations, for both 1-jet and 2-jet exclusive regions. The proposed fiducial region for these measurements includes events with three leptons, where the invariant mass of at least one opposite charge, same flavor lepton pair falls within 10 GeV of 91.2 GeV, with 1 or 2 associated jets with $p_T > 25 \text{ GeV}$ and radius $R < 0.4$.

2 The ATLAS Detector

The ATLAS detector [2] at the LHC is a general purpose detector that covers nearly the entire solid angle around the collision point. It consists of several concentric subdetectors: The inner tracking detector, electromagnetic and hadronic calorimeters, and a muon spectrometer. The inner detector is made up of a high-granularity silicon pixel detector, designed to reconstruct the tracks of charged particles in a range of $|\eta| < 2.5$, and a transition radiation tracker which provides additional tracking and electron identification information for $|\eta| < 2.0$ [3]. A 2 T

axial magnetic field is produced in the inner detector, in order to bend the path of charged particles. The calorimeter system covers a pseudorapidity range of $|\eta| < 4.9$, with a lead/liquid-argon (LAr) electromagnetic calorimeter covering $|\eta| < 3.2$, and a steel/scintillator-tile hadronic calorimeter [4]. The rest of the solid-angle coverage of the calorimeter system comes from forward copper/LAr and tungsten/LAr modules. The muon spectrometer measures muons with $|\eta| < 2.7$ using several layered tracking chambers placed within a magnetic field of approximately 0.5 T. A two-level trigger system [5] is used to reduce the event rate from 40 MHz to around 1 kHz, using a hardware based Level-1 trigger, followed by a second software based High-Level Trigger (HLT).

3 Data and Monte Carlo Samples

3.1 Data Samples

This 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} [6]. 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 [7].

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 [8]. The WZ signal samples are simulated using Sherpa 2.2.2 [9]. Specific information about the Monte Carlo samples being used can be found in Table 1.

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, ZZ, WW	SHERPA 2.2.2	MEPS NLO	SHERPA	CT10 [10]
tZ	MG5_AMC [Alwall:2014hca]	NLO	PYTHIA 8	CTEQ6L1
t \bar{t} W	MG5_AMC (SHERPA 2.1.1)	NLO (LO multileg)	PYTHIA 8 (SHERPA)	NNPDF 3.0 NLO (NNPDF 3.0 NLO)
t \bar{t} (Z/ $\gamma^* \rightarrow \ell\ell$)	MG5_AMC	NLO	PYTHIA 8	NNPDF 3.0 NLO
t \bar{t} H	MG5_AMC (MG5_AMC)	NLO (NLO)	PYTHIA 8 (HERWIG++) [Bahr:2008pv]	NNPDF 3.0 NLO [11] (CT10 [10])
tHqb	MG5_AMC	LO	PYTHIA 8	CT10
tHW	MG5_AMC (SHERPA 2.1.1)	NLO (LO multileg)	HERWIG++ (SHERPA)	CT10 (NNPDF 3.0 NLO)
tWZ	MG5_AMC	NLO	PYTHIA 8	NNPDF 2.3 LO
t \bar{t} t, t \bar{t} t \bar{t}	MG5_AMC	LO	PYTHIA 8	NNPDF 2.3 LO [Ball:2012cx]
t \bar{t} W $^+W^-$	MG5_AMC	LO	PYTHIA 8	NNPDF 2.3 LO
t \bar{t}	POWHEG-BOX v2 [12]	NLO	PYTHIA 8	NNPDF 3.0 NLO
t \bar{t} γ	MG5_AMC	LO	PYTHIA 8	NNPDF 2.3 LO
s-, t-channel, Wt single top	POWHEG-BOX v1 [13]	NLO	PYTHIA 6	CT10
qqVV, VVV				
Z $\rightarrow \ell^+\ell^-$	SHERPA 2.2.1	MEPS NLO	SHERPA	NNPDF 3.0 NLO

4 Object Reconstruction

All regions defined in this analysis share a common lepton, jet, and overall event preselection. The selection applied to each physics object is detailed here; the event preselection, and the selection used to define the various fit regions, is described in Section 7.

All events are required to be selected by dilepton triggers. The p_T thresholds of the dilepton trigger on two electrons were 12 GeV in 2015, 17 GeV in 2016, and 24 GeV in 2017 and 2018, while for the dimuon triggers the p_T thresholds on the leading (sub-leading) muon were 18 GeV (8 GeV) in 2015, and 22 GeV (8 GeV) in 2016-2018. For the electron+muon triggers, the p_T thresholds on the electron (muon) were 17 GeV (14 GeV) for all datasets.

Electron candidates are reconstructed from energy clusters in the electromagnetic calorimeter that are associated with charged particle tracks reconstructed in the inner detector [14]. 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 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). Electron candidates are required to pass TightLH identification.

Muon candidates are reconstructed by combining inner detector tracks with track segments or full tracks in the muon spectrometer [15]. Muon candidates are required to have $p_T > 10$ GeV and $|\eta| < 2.5$. The longitudinal impact parameter is the same for both electrons and muons, while muons are required to pass a slightly tighter transverse impact parameter selection, $|d_0|/\sigma_{d_0} < 3$. Muons are also required to pass Medium ID requirements.

Leptons are additionally required to pass a non-prompt BDT selection. This BDT and the WPs used are summarized in Appendix ??, and described in detail in [16]. Optimized working points and scale factors for this BDT are taken from that analysis.

Jets are reconstructed from calibrated topological clusters built from energy deposits in the calorimeters [17], using the anti- k_t algorithm with a radius parameter $R = 0.4$. Particle Flow, or PFlow, jets are used in the analysis, which are hadronic objects reconstructed using information from both the tracker and the calorimeter. Jets with energy contributions likely arising from noise or detector effects are removed from consideration [18], 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 [19].

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 [20]. 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.

From the output of the BDT, calibrated 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 2.

WP	Rejection	
	b-jet eff.	light jet
85%	2.6	29
77%	4.9	130
70%	9.4	390
60%	27	1300

Table 2: c-jet and light-flavor jet rejections corresponding to each b-tagging Working Point by b-jet efficiency, evaluated on $t\bar{t}$ events.

As shown in table 2, 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 pseudo-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.

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 [21]. 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.

To avoid double counting objects and remove leptons originating from decays of hadrons, overlap removal is performed in the following order: any electron candidate within $\Delta R = 0.1$ of another electron candidate with higher p_T is removed; any electron candidate within $\Delta R = 0.1$ of a muon candidate is removed; any jet within $\Delta R = 0.3$ of an electron candidate is removed; if a muon candidate and a jet lie within $\Delta R = \min(0.4, 0.04 + 10[\text{GeV}]/p_T(\text{muon}))$ of each other, the jet is kept and the muon is removed. This algorithm is applied to the preselected objects.

5 tZ Separation Classifier

An important process to consider in this analysis is tZ: the top almost always decays into a W boson and b-quark, and when both the W and Z decay leptonically, this gives three leptons and a heavy flavor jet in the final state. Because tZ can produce a final state identical to the 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 XGBoost [22] to separate WZ + heavy flavor from tZ using kinematic quantities. 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.

The kinematic variables used as inputs to train this BDT include the invariant mass of the reconstructed top candidate, the p_T of each of the leptons and associated jets, the invariant mass of each combination of lepton pairs, E_T^{miss} , the distance between each combination of leptons, $\Delta R(\text{ll})$, and the distance between each lepton and the jet, $\Delta R(\text{lj})$.

Here the top candidate is reconstructed based on the procedure described in section 6.1 of [23]. Broadly, 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 training samples included only events meeting the requirements of the 1-jet, >60% region, i.e. passing all the selection described in section 7 and having exactly one jet which passes the tightest (60%) DL1r working point. 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 1.

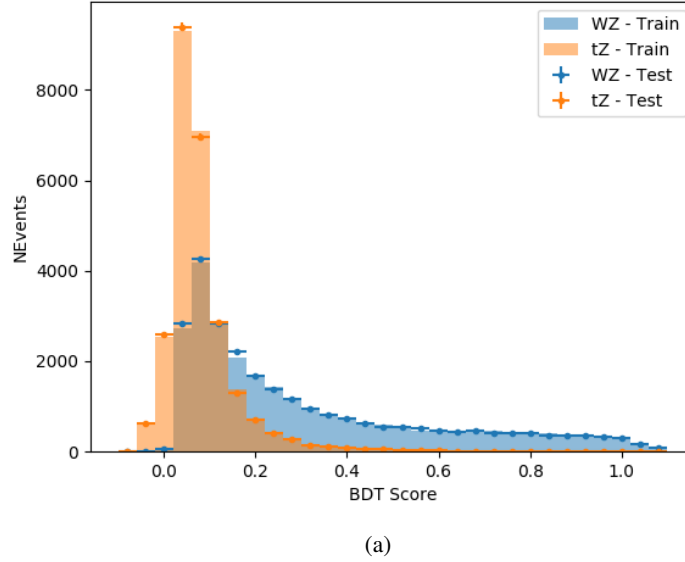


Figure 1: Distribution of the BDT response for WZ+b (blue) and tZ (orange) events, for both training and testing samples.

A BDT score of 0.12 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 this cutoff in stat-only Asimov fits, and selecting the value that minimizes the statistical uncertainty on WZ + b.

6 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 $t\bar{t}Z$ 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 overall cross-section with and without interference effects 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.

7 Event Selection and Signal Region Definitions

Event are required to pass a preselection described in Section 7.1. Those that pass this preselection are divided into various fit regions described in Section 7.2, based on the number of jets in the event, and the b-tag score of those jets.

7.1 Event Preselection

Events are required to include exactly three reconstructed light leptons passing the requirement described in 4, which have a total charge of ± 1 . As the opposite sign lepton is found to be prompt the vast majority of the time [24], it is required to have $p_T > 10$ GeV, while the same sign leptons 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, $m_T(E_T^{\text{miss}} + l_{\text{other}}) > 30$ 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 exactly one or two reconstructed. 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.

The WZ events are split into $WZ + b$, $WZ + c$, and $WZ + \text{light}$ based on the truth flavor of the associated jet in the event. In this ordering b-jet supersede charm, which supersides light. That is, $WZ + \text{light}$ 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.

7.2 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 3.

Table 3: 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%	$N_{\text{jets}} = 1, n\text{Jets_DL1r_60} = 1, tZ \text{ BDT} > 0.12$
1j tZ CR	$N_{\text{jets}} = 1, n\text{Jets_DL1r_60} = 1, tZ \text{ BDT} < 0.12$
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%	$N_{\text{jets}} = 2, n\text{Jets_DL1r_60} \geq 1, tZ \text{ BDT} > 0.12$
2j tZ CR	$N_{\text{jets}} = 2, n\text{Jets_DL1r_60} \geq 1, tZ \text{ BDT} < 0.12$

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 unfolding procedure is performed to account for differences in the number of reconstructed jets compared to the number of truth jets in each event. In order to account for migration of WZ+1-jet and WZ+2-jet events between the 1-jet and 2-jet bins at reco level, the signal samples are separated based on the number of truth jets. Events with 0 jets or more than 3 jets at truth level, yet fall within one of the categories listed in Table 3, are categorized as WZ + other, and treated as background. The composition of the number of truth jets in each reco jet bin is taken from MC, with uncertainties in these estimates described in detail in Section 8.

An additional tZ control region is created based on the BDT described in Section 5. 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 0.12, and a tZ control region including events with less than 0.12. This cutoff is optimized for significance of WZ + b.

7.3 Non-Prompt Lepton Estimation

Two processes that act as sources of non-prompt leptons appear in the analysis: $t\bar{t}$ and Z+jet production both produce two prompt leptons, but can meet the selection of this analysis 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 non-prompt enriched regions. These validation regions are used to derive correction factors and uncertainties for the non-prompt contribution.

$t\bar{t}$ 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 7 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.

Data is compared to MC predictions in the region for a variety of kinematic variable, as well as various b-tag WPs. A constant normalization discrepancy between data and MC predictions of approximately 10% is found, which is accounted for by applying a constant correction factor of 0.9 to the $t\bar{t}$ MC prediction. As data and MC are found to agree within 20% for each of the b-tag WPs considered, a 20% systematic uncertainty on the $t\bar{t}$ prediction is included for the analysis.

Similar to $t\bar{t}$, a 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 $E_T^{\text{miss}} < 30$ 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.

While there is general agreement between data and MC, the shape of the p_T spectrum of the lepton from the W is found to differ. To account for this discrepancy, a variable correction factor is applied to Z+jets. χ^2 minimization of the W lepton p_T spectrum is performed to derive a correction factor.

The systematic uncertainty in the Z + jets prediction is evaluated by comparing data to MC for each of the continuous b-tag WPs. For each of the regions considered, the data falls within 25% of the MC prediction once this correction factor has been applied. Therefore, a 25% systematic uncertainty is applied to Z + jets in the analysis.

8 Systematic Uncertainties

The systematic uncertainties that are considered are summarized in Table 4. 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.

Table 4: Sources of systematic uncertainty considered in the analysis.

Systematic uncertainty	Components
Luminosity	1
Pileup reweighting	1
Physics Objects	
Electron	6
Muon	15
Jet energy scale	28
Jet energy resolution	8
Jet vertex fraction	1
Jet flavor tagging	131
E_T^{miss}	3
Total (Experimental)	194
Signal Modeling	
Shape modelling	3
Renormalization and factorization scales	5
nJet Migration	5
Background Modeling	
Cross section	15
Renormalization and factorization scales	12
Total (Signal and background modeling)	35
Total (Overall)	230

The uncertainty in the combined integrated luminosity is derived from a calibration of the luminosity scale performed for 13 TeV proton-proton collisions [6], [25].

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) [26] are decomposed into uncorrelated components and treated as independent sources of uncertainty in the analysis. A similar approach is used for the jet energy resolution (JER) uncertainty.

The uncertainties in the b-tagging efficiencies measured in dedicated calibration analyses [27] 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 for each WP bin..

Theoretical 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 $t\bar{t}Z$ is taken from [28]. Derivation of the non-prompt background uncertainties, Z +jets and $t\bar{t}$, are explained in Section 7.3. These normalization uncertainties are chosen so as to account for the complete uncertainty in the non-prompt contribution, and therefore no additional modelling uncertainties are considered for Z +jets and $t\bar{t}$.

The other VV + heavy flavor processes (namely $VV+b$ and VV +charm, which primarily consist of ZZ events) are also poorly understood, because these processes involve the same physics as WZ + heavy flavor, and have also not been measured. Therefore, a conservative 50% uncertainty is applied to those samples. While this uncertainty is large, it is found to have little impact on the significance of the final result.

The theory uncertainties applied to the predominant background estimates are summarized in Table 5.

Process	X-section [%]
$t\bar{t}Z$	X-sec: ± 15.2 QCD Scale: $+5.2$ PDF($+\alpha_S$): -1.3 PDF($+\alpha_S$): ± 1.2
$t\bar{t}H$ (aMC@NLO+Pythia8)	QCD Scale: $+5.8$ PDF($+\alpha_S$): -9.2 PDF($+\alpha_S$): ± 3.6
$t\bar{t}Z$ (aMC@NLO+Pythia8)	QCD Scale: $+9.6$ PDF($+\alpha_S$): -11.3 PDF($+\alpha_S$): ± 4
$t\bar{t}W$ (aMC@NLO+Pythia8)	QCD Scale: $+12.9$ PDF($+\alpha_S$): -11.5 PDF($+\alpha_S$): ± 3.4
$VV + b/\text{charm}$ (Sherpa 2.2.1)	± 50
$VV + \text{light}$ (Sherpa 2.2.1)	± 6
$t\bar{t}$	± 20
$Z + \text{jets}$	± 25
Others	± 50

Table 5: Summary of theoretical uncertainties for normalization of 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. Separate systematics are included in the fit for $WZ + b$, $WZ + \text{charm}$ and $WZ + \text{light}$, where the distribution among each of the fit regions is varied based on the prediction of the Powheg sample.

A similar approach is taken to account for uncertainties in migrations between the number of reco and truth jets. The fraction of events with 1 truth jet which fall into the 1 jet bin versus the 2 jet bin at reco level is compared for Sherpa and Powheg. The same is done for events with 2 truth jets. A systematic is included where events are shifted between the 1-jet and 2-jet regions

based on the differences between these two shapes. This is done independently for each of the WZ + b, WZ + charm, and WZ + light templates.

Additional systematics are included to account for the uncertainty in the contamination of 0 jet and 3 or more jet events (as defined at truth level) in the 1 and 2 reco jet bins. Because these events fall outside the scope of this measurement, these events are included as a background. As such, a normalization, rather than a shape, uncertainty is applied for this background. The number of WZ events with 0-jets and ≥ 3 -jets in the reconstructed 1-jet and 2-jet regions are compared for Sherpa and Powheg, and these differences are taken as separate normalization systematics on the yield of WZ+0-jet and WZ+ ≥ 3 -jet events.

9 Results

9.1 Fit Procedure

A maximum-likelihood fit is performed over the total yields in the various fit regions described in Section 7 in order to extract the best-fit value of the WZ + b-jet and WZ + charm jet contributions for events with both 1 and 2 associated jets.

Because the fit regions are defined by the number of associated jets at reco-level, the signal is split into separate samples based on the number of truth jets in order to account for differences in the number of truth jets compared to the number of reco-jets. The WZ + b, WZ + charm and WZ + light contributions are separated into independent samples based on the number of truth jets in each event. WZ + 1 truth-jet and WZ + 2 truth-jets are treated as signal samples, while WZ + 0 truth-jets and WZ + ≥ 3 truth-jets are treated as an additional background.

A maximum likelihood fit to data is performed simultaneously in the regions described in Section 7, summarized in figure 2. The six signal templates, which include WZ+b 1-jet, WZ+c 1-jet, WZ+l 1-jet, WZ+b 2-jets, WZ+c 2-jets, WZ+l 2-jets, are allowed to float, while the remaining background contributions are held fixed. The parameters $\mu_{\text{WZ+b-1-jet}}$, $\mu_{\text{WZ+charm1-jet}}$, $\mu_{\text{WZ+light-1-jet}}$, $\mu_{\text{WZ+b-2-jet}}$, $\mu_{\text{WZ+charm2-jet}}$, $\mu_{\text{WZ+light-2-jet}}$, where $\mu = \sigma_{\text{observed}}/\sigma_{\text{SM}}$, are extracted from the fit. A simultaneous fit is performed over all 1-jet and 2-jet regions.

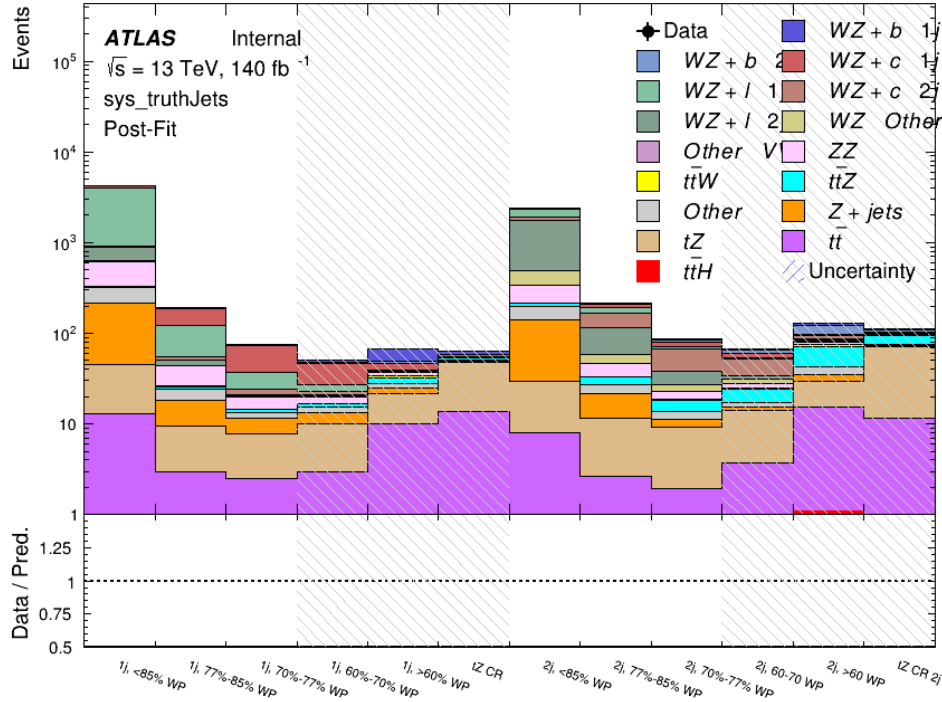


Figure 2: Post-fit summary of the fit regions.

As described in Section 8, there are 230 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 number 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.

Several alternate fit strategies are documented in Appendices A.1-A.2.1. These include a measurement of $WZ + 1$ or 2 jets inclusively, a fit where tZ is allowed to float, and a case where tZ is included as part of the signal.

9.2 Fiducial Region Definition

The fiducial volume at particle level is defined based on the number of stable leptons and jets in each event. Three light leptons with total charge ± 1 and one or two associated jets are required. This is separated into four observations based on the number and flavor of associated jets.

Only leptons which do not originate from hadron or τ decays are considered. The phase space definitions use dressed kinematics of the final state particles. Leptons are dressed by summing

the momentum of photons within a cone of $\Delta R < 0.1$ of the lepton to correct the leptons energy. Particle level jets are reconstructed using the anti- k_t algorithm with a radius of $R = 0.4$.

The kinematic selection used at particle level closely follows the selection used at reconstructed level. Three light leptons with total charge ± 1 and one or two associated jets Leptons and jets are required to have $|\eta| < 2.5$, with the transition region included. The OS leptons is required to have $p_T > 10$ GeV, while the SS leptons are required to have $p_T > 20$ GeV. Jets are required to have $p_T > 25$ GeV. The base fiducial region definition is summarized below:

- Three light leptons with total charge ± 1 , $|\eta| < 2.5$
- OS lepton with $p_T > 10$ GeV, SS leptons with $p_T > 20$ GeV
- One OSSF lepton pair with $|M(\ell\ell) - 91.2 \text{ GeV}| < 10 \text{ GeV}$
- One or two associated truth jets with $p_T > 25 \text{ GeV}$, $|\eta| < 2.5$

The result of the fit is used to extract the cross-section in this fiducial region for WZ + b and WZ + charm with one associated jet, and WZ + b and WZ + charm with two associated jets, where the number and flavor of the jets is determined at particle level. Events with both charm and b-jets are counted as WZ + b.

9.3 Results of the Simultaneous Fit

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

Process	μ	σ
WZ + b - 1-jet	$1.00^{+0.47}_{-0.43}(\text{stat})^{+0.30}_{-0.27}(\text{sys})$	
WZ + c - 1-jet	$1.00 \pm 0.17(\text{stat}) \pm 0.17(\text{sys})$	
WZ + b - 2-jet	$1.00^{+0.53}_{-0.51}(\text{stat})^{+0.39}_{-0.34}(\text{sys})$	
WZ + c - 2-jet	$1.00 \pm 0.25(\text{stat}) \pm 0.21(\text{sys})$	

Table 6: Correlations between the various components of WZ

The expected cross-section of WZ+b with 1-jet is $1.74^{+0.82}_{-0.75}(\text{stat})^{+0.53}_{-0.48}(\text{sys})$ fb, and $14.6 \pm 2.5(\text{stat}) \pm 2.3(\text{sys})$ fb for WZ + charm, with a correlation of -0.15 between them. An expected significance of 2.0 is observed for WZ + b in this region.

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

The expected WZ + b cross-section in the 2-jet region is $2.5^{+1.3}_{-1.3}(\text{stat})^{+0.95}_{-0.83}(\text{sys})$ fb with an expected significance of 1.7σ . The 2-jet expected cross-section of WZ + charm is $12.7 \pm 3.2(\text{stat}) \pm 2.7(\text{sys})$ fb, and the correlation between WZ + charm and WZ + b is -0.22.

A summary of the correlation between the various WZ components is summarized in Table 7.

	WZ + b - 1-jet	WZ + c - 1-jet	WZ + l - 1-jet	WZ + b - 2-jet	WZ + c - 2-jet	WZ + l - 2-jet
WZ + b - 1-jet	1.00	-0.15	0.28	-0.13	-0.22	0.17
WZ + c - 1-jet	-	1.00	0.36	0.13	-0.14	-0.16
WZ + l - 1-jet	-	-	1.00	0.10	-0.20	-0.39
WZ + b - 2-jet	-	-	-	1.00	-0.22	0.17
WZ + c - 2-jet	-	-	-	-	1.00	0.23
WZ + l - 2-jet	-	-	-	-	-	1.00

Table 7: Correlations between the various components of WZ

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 on WZ + b with one associated jet is summarized in Table 8.

Uncertainty Source	$\Delta\sigma/\sigma_{\text{nominal}}$	
WZ + 1-jet light cross-section	0.13	-0.15
WZ + 1-jet charm cross-section	-0.10	0.12
Jet Energy Scale	0.1	-0.13
Other Diboson + b cross-section	-0.09	0.09
tZ cross-section	-0.08	0.08
WZ 1-jet/2-jet Migration	0.08	-0.07
Jet Energy Resolution	-0.07	0.08
Luminosity	-0.06	0.07
Flavor tagging	0.05	0.05
t \bar{t} cross-section	-0.05	0.05
Total Systematic Uncertainty	0.28	0.33

Table 8: Summary of the most significant sources of systematic uncertainty on the measurement of WZ + b with exactly one associated jet.

The ranking and impact of those nuisance parameters with the largest contribution to the overall uncertainty is shown in Figure 3.

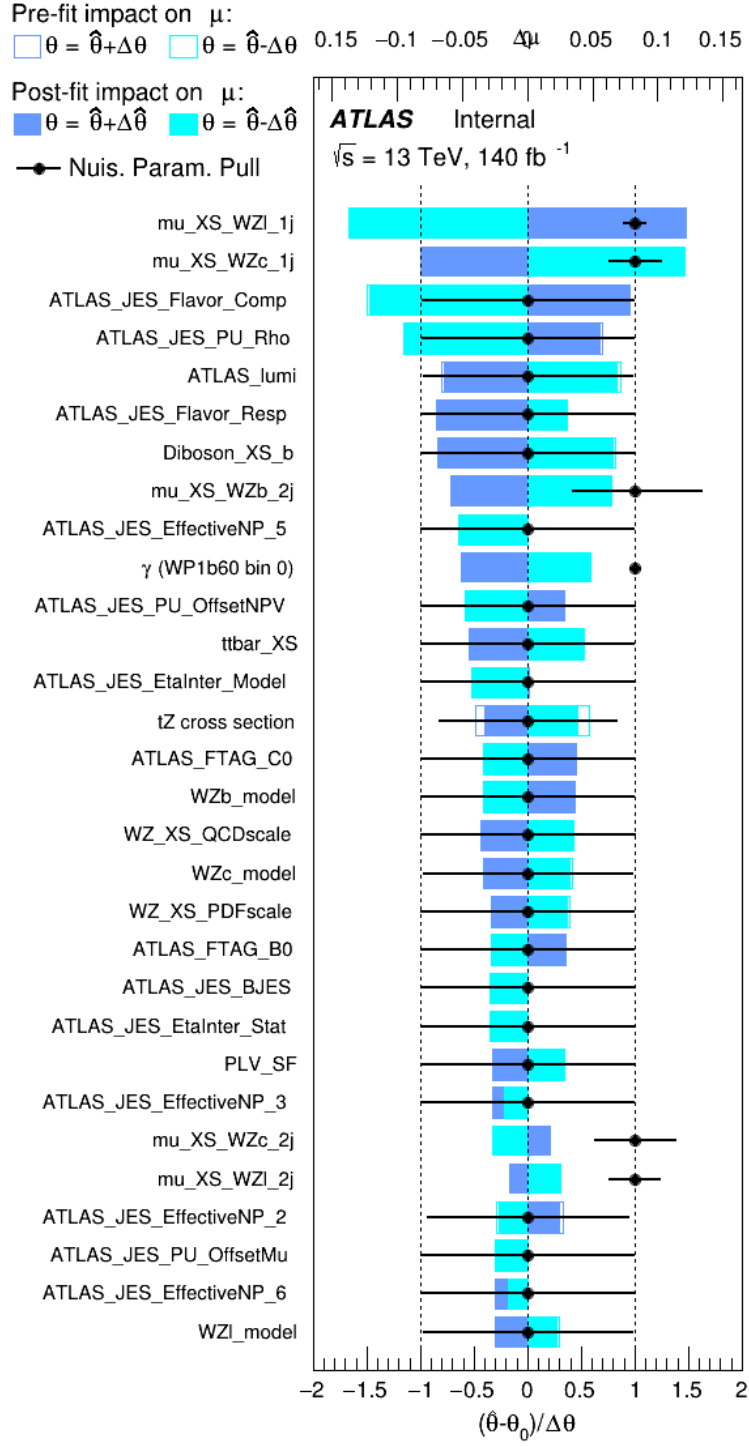


Figure 3: Impact of systematic uncertainties on the signal-strength of WZ + b for events with exactly one jet

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 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 9.

Uncertainty Source	$\Delta\sigma/\sigma_{\text{nominal}}$	
WZ + c 2-jet cross-section	-0.13	0.16
WZ + l 2-jet cross-section	0.12	-0.09
ttZ cross-section - QCD scale	-0.10	0.13
WZ + b 1-jet cross-section	-0.11	0.10
Jet Energy Scale	-0.11	0.11
Luminosity	-0.11	0.12
tZ cross-section	-0.11	0.11
WtZ cross-section	-0.07	0.07
Flavor tagging	0.05	0.05
Other VV + b cross-section	-0.05	0.05
Total	0.35	0.37

Table 9: 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 4.

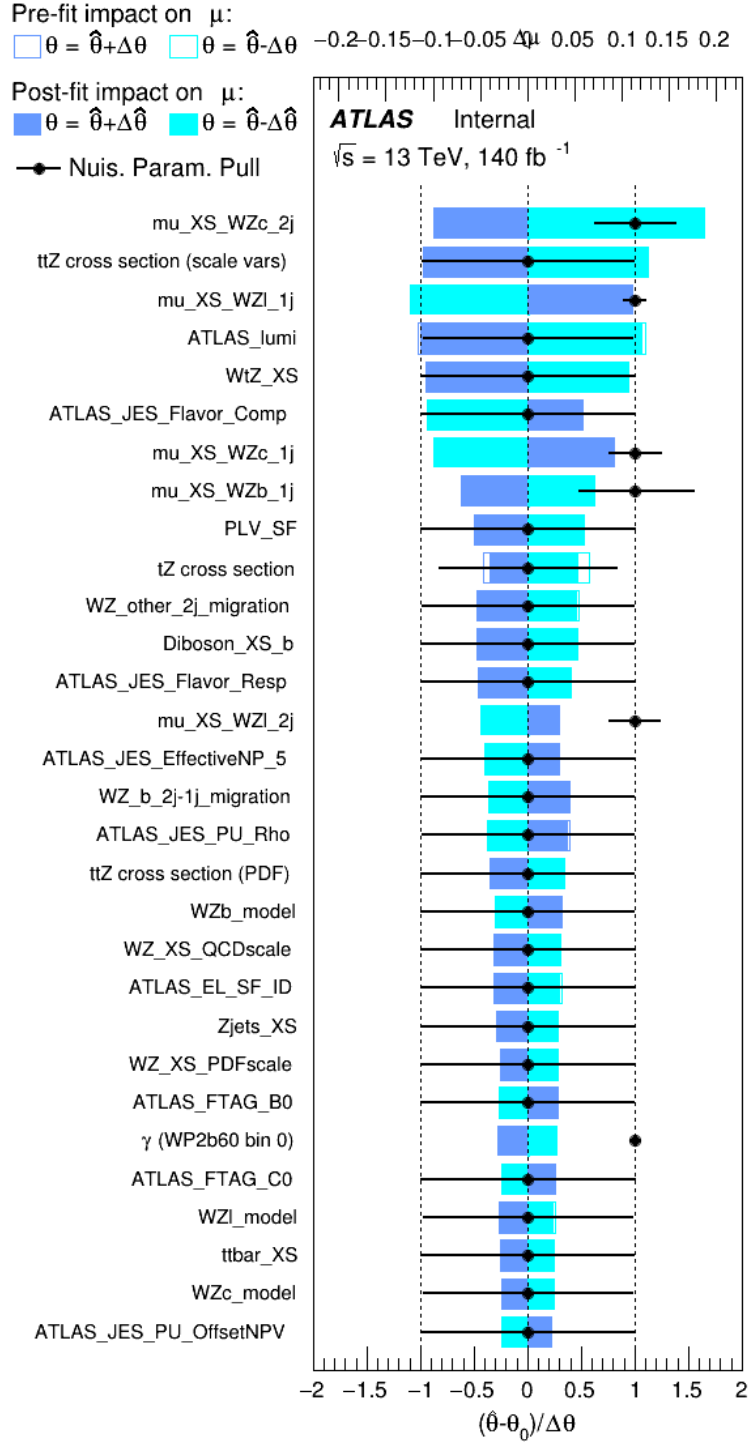


Figure 4: Impact of systematic uncertainties on the signal-strength of WZ + b in 2-jet events.

10 Conclusion

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

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

This section will be include final results once unblinded.

Acknowledgements

We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently.

We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BM-WFW and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DNRF and DNSRC, Denmark; IN2P3-CNRS, CEA-DRF/IRFU, France; SRNSFG, Georgia; BMBF, HGF, and MPG, Germany; GSRT, Greece; RGC, Hong Kong SAR, China; ISF, I-CORE and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; NWO, Netherlands; RCN, Norway; MNiSW and NCN, Poland; FCT, Portugal; MNE/IFA, Romania; MES of Russia and NRC KI, Russian Federation; JINR; MESTD, Serbia; MSSR, Slovakia; ARRS and MIZŠ, Slovenia; DST/NRF, South Africa; MINECO, Spain; SRC and Wallenberg Foundation, Sweden; SERI, SNSF and Cantons of Bern and Geneva, Switzerland; MOST, Taiwan; TAEK, Turkey; STFC, United Kingdom; DOE and NSF, United States of America. In addition, individual groups and members have received support from BCKDF, the Canada Council, CANARIE, CRC, Compute Canada, FQRNT, and the Ontario Innovation Trust, Canada; EPLANET, ERC, ERDF, FP7, Horizon 2020 and Marie Skłodowska-Curie Actions, European Union; Investissements d'Avenir Labex and Idex, ANR, Région Auvergne and Fondation Partager le Savoir, France; DFG and AvH Foundation, Germany; Herakleitos, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSRF; BSF, GIF and Minerva, Israel; BRF, Norway; CERCA Programme Generalitat de Catalunya, Generalitat Valenciana, Spain; the Royal Society and Leverhulme Trust, United Kingdom.

The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN, the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (UK) and BNL (USA), the Tier-2 facilities worldwide and large non-WLCG resource providers. Major contributors of computing resources are listed in Ref. [29].

A Appendices

A.1 Inclusive 1+2 Jet Cross Check

An alternate fit is performed which combines the WZ + 1-jet and WZ + 2-jet samples rather than fitting them independently. This is done primarily as a cross-check of the nominal analysis, to see if measuring 1-jet and 2-jet events separately and combining them gives drastically different results than measuring them together.

For this study, three signal templates, WZ + b, WZ + charm and WZ + light, are fit to data, and the systematics accounting for migrations between 1-jet and 2-jet bins are removed. All other background and nuisance parameters remain the same as the nominal fit.

The measured μ value for WZ + b is $\mu = 1.00^{+0.30}_{-0.29}(\text{stat})^{+0.25}_{-0.23}(\text{sys})$, with a significance of 2.8σ , and the uncertainty on WZ + charm is $\mu = 1.00 \pm 0.12(\text{stat}) \pm 0.13(\text{sys})$. This is compared to combined uncertainty of $\mu = 1.00^{+0.32}_{-0.30}(\text{stat})^{+0.24}_{-0.23}(\text{sys})$ for WZ + b when 1-jet and 2-jet events are measured separately and then combined.

A post-fit summary plot of the fit regions is shown in Figure 5:

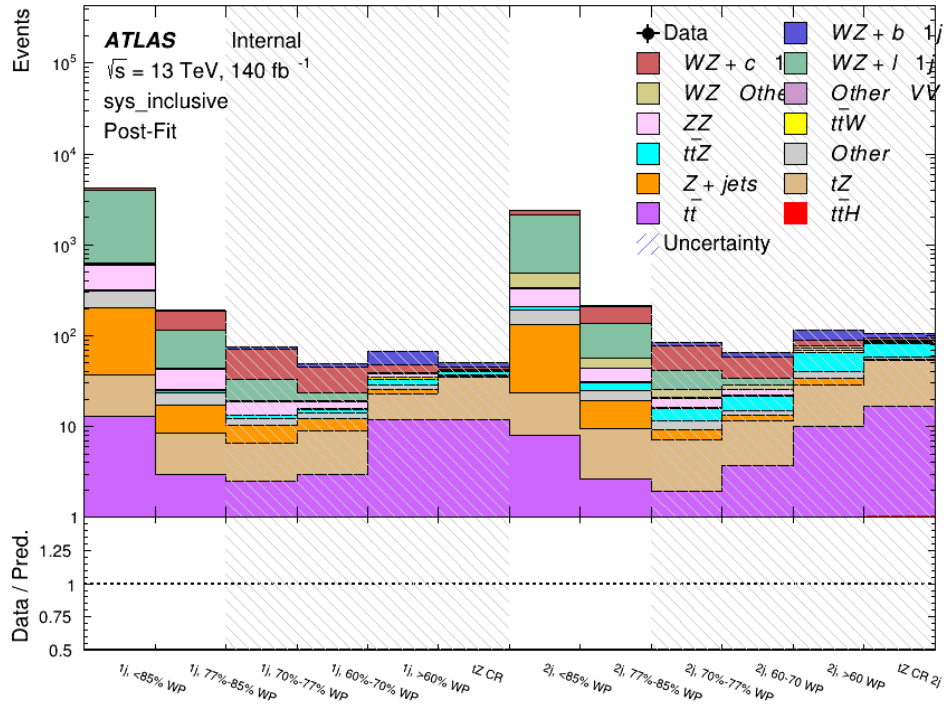


Figure 5: Post-fit summary of the 1-jet fit regions.

463 The impact of the most significant sources of systematic uncertainties on the measurement of
 464 WZ+b is summarized in Table 10.

Uncertainty Source	$\Delta\mu$	
WZ + light cross-section	0.13	-0.12
WZ + charm cross-section	-0.10	0.12
Jet Energy Scale	0.08	0.13
tZ cross-section	-0.10	0.10
Jet Energy Resolution	-0.10	0.10
Luminosity	-0.08	0.09
Other Diboson + b cross-section	-0.07	0.07
Flavor tagging	0.05	0.05
t \bar{t} cross-section	-0.05	0.05
WZ cross-section - QCD scale	-0.04	0.03
Total Systematic Uncertainty	0.28	0.32

Table 10: Summary of the most significant sources of systematic uncertainty on the measurement of WZ + b with one or two associated jets.

465 The ranking and impact of those nuisance parameters with the largest contribution to the overall
 466 uncertainty is shown in Figure 6.

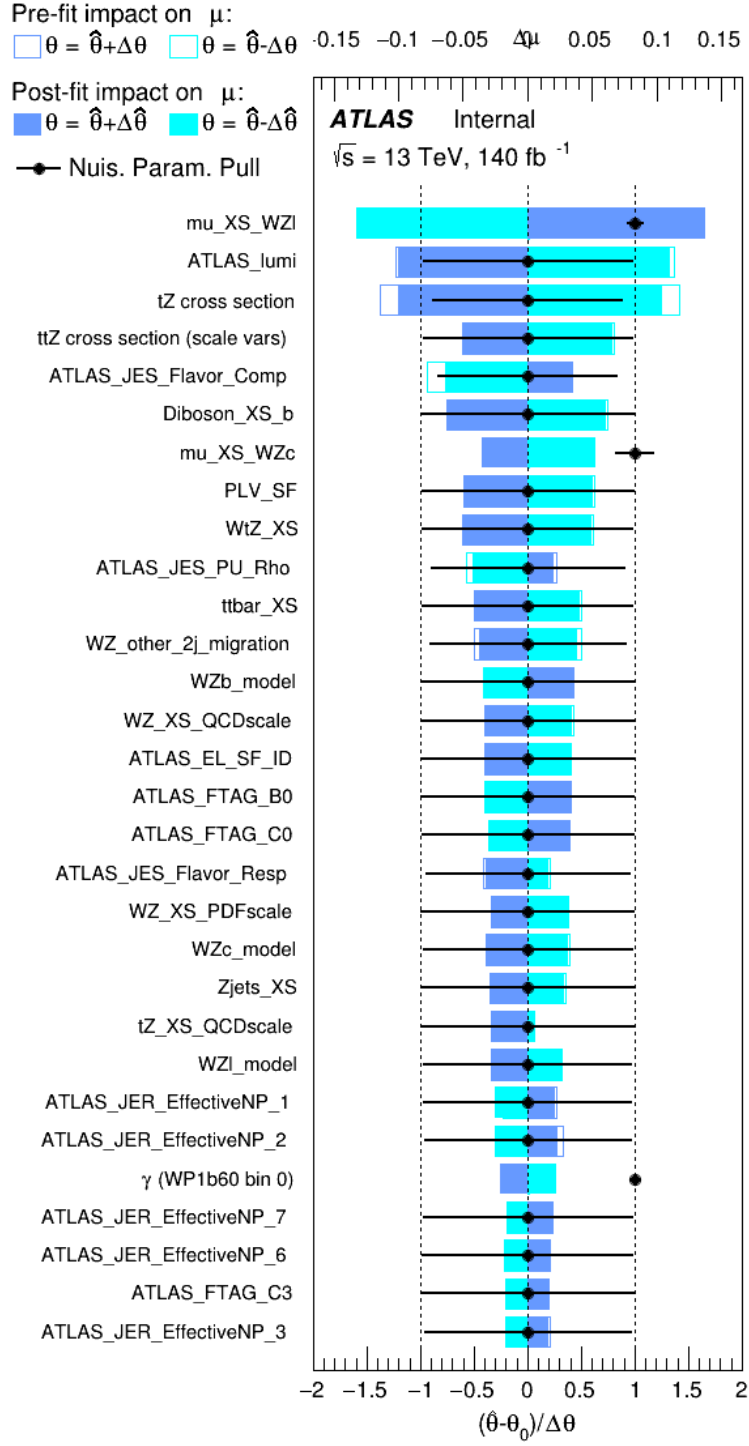


Figure 6: Impact of systematic uncertainties on the signal-strength of WZ + b for events with one or two jets

A.2 Alternate tZ Inclusive Fit

A.2.1 tZ Inclusive Fit

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

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

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

The impact of the predominate systematics are summarized in Table 11.

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

Table 11: Summary of the most significant sources of systematic uncertainty on the measurement of WZ + b with exactly one associated jet.

A.2.2 Floating tZ

In order to quantify the impact of the tZ uncertainty on the fit, an alternate fit strategy is used where the tZ normalization is allowed to float. This normalization factor replaces the cross-section uncertainty on tZ, and all other parameters of the fit remain the same.

An uncertainty of 17% on the normalization of tZ is extracted from the fit, compared to a theory uncertainty of 15% applied to the tZ cross-section. The measured uncertainties on WZ remain the same.

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] ATLAS Collaboration. ‘The ATLAS Experiment at the CERN Large Hadron Collider’. In: *JINST* 3 (2008), S08003. DOI: [10.1088/1748-0221/3/08/S08003](https://doi.org/10.1088/1748-0221/3/08/S08003).
- [3] ATLAS Collaboration. ‘The ATLAS Inner Detector commissioning and calibration’. In: *Eur. Phys. J. C* 70 (2010), p. 787. DOI: [10.1140/epjc/s10052-010-1366-7](https://doi.org/10.1140/epjc/s10052-010-1366-7). arXiv: [1004.5293](https://arxiv.org/abs/1004.5293) [hep-ex].
- [4] ATLAS Collaboration. ‘Readiness of the ATLAS liquid argon calorimeter for LHC collisions’. In: *Eur. Phys. J. C* 70 (2010), p. 723. DOI: [10.1140/epjc/s10052-010-1354-y](https://doi.org/10.1140/epjc/s10052-010-1354-y). arXiv: [0912.2642](https://arxiv.org/abs/0912.2642) [hep-ex].
- [5] ATLAS Collaboration. ‘Performance of the ATLAS Trigger System in 2010’. In: *Eur. Phys. J. C* 72 (2012), p. 1849. DOI: [10.1140/epjc/s10052-011-1849-1](https://doi.org/10.1140/epjc/s10052-011-1849-1). arXiv: [1110.1530](https://arxiv.org/abs/1110.1530) [hep-ex].
- [6] 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](https://doi.org/10.1140/epjc/s10052-011-1630-5). arXiv: [1101.2185](https://arxiv.org/abs/1101.2185) [hep-ex].
- [7] ATLAS Collaboration. ‘Performance of the ATLAS detector using first collision data’. In: *JHEP* 09 (2010), p. 056. DOI: [10.1007/JHEP09\(2010\)056](https://doi.org/10.1007/JHEP09(2010)056). arXiv: [1005.5254](https://arxiv.org/abs/1005.5254) [hep-ex].
- [8] S. Agostinelli et al. ‘GEANT4: A Simulation toolkit’. In: *Nucl. Instrum. Meth. A* 506 (2003), pp. 250–303. DOI: [10.1016/S0168-9002\(03\)01368-8](https://doi.org/10.1016/S0168-9002(03)01368-8).
- [9] 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].
- [10] H.-L. Lai et al. ‘New parton distributions for collider physics’. In: *Phys. Rev. D* 82 (2010), p. 074024. DOI: [10.1103/PhysRevD.82.074024](https://doi.org/10.1103/PhysRevD.82.074024). arXiv: [1007.2241](https://arxiv.org/abs/1007.2241) [hep-ph].
- [11] R. D. Ball et al. ‘Parton distributions for the LHC Run II’. In: *JHEP* 04 (2015), p. 040. DOI: [10.1007/JHEP04\(2015\)040](https://doi.org/10.1007/JHEP04(2015)040). arXiv: [1410.8849](https://arxiv.org/abs/1410.8849) [hep-ph].
- [12] S. Frixione, G. Ridolfi and P. Nason. ‘A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction’. In: *JHEP* 09 (2007), p. 126. DOI: [10.1088/1126-6708/2007/09/126](https://doi.org/10.1088/1126-6708/2007/09/126). arXiv: [0707.3088](https://arxiv.org/abs/0707.3088) [hep-ph].
- [13] E. Re. ‘Single-top Wt -channel production matched with parton showers using the POWHEG method’. In: *Eur. Phys. J. C* 71 (2011), p. 1547. DOI: [10.1140/epjc/s10052-011-1547-z](https://doi.org/10.1140/epjc/s10052-011-1547-z). arXiv: [1009.2450](https://arxiv.org/abs/1009.2450) [hep-ph].
- [14] 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>.

- [15] 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].
- [16] R. Narayan et al. *Measurement of the total and differential cross sections of a top-quark-antiquark pair in association with a W boson in proton-proton collisions at a centre-of-mass energy of 13 TeV with ATLAS detector at the Large Hadron Collider*. Tech. rep. ATL-COM-PHYS-2020-217. Geneva: CERN, Mar. 2020. URL: <https://cds.cern.ch/record/2712986>.
- [17] 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>.
- [18] 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>.
- [19] 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].
- [20] M. Aaboud et al. ‘Measurements of b-jet tagging efficiency with the ATLAS detector using $t\bar{t}$ events at $\sqrt{s} = 13$ TeV’. In: *Journal of High Energy Physics* 2018.8 (Aug. 2018). ISSN: 1029-8479. DOI: [10.1007/jhep08\(2018\)089](https://doi.org/10.1007/jhep08(2018)089). URL: [http://dx.doi.org/10.1007/JHEP08\(2018\)089](http://dx.doi.org/10.1007/JHEP08(2018)089).
- [21] 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>.
- [22] T. Chen and C. Guestrin. ‘XGBoost: A Scalable Tree Boosting System’. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD ’16. San Francisco, California, USA: ACM, 2016, pp. 785–794. ISBN: 978-1-4503-4232-2. DOI: [10.1145/2939672.2939785](https://doi.org/10.1145/2939672.2939785). URL: <http://doi.acm.org/10.1145/2939672.2939785>.
- [23] 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>.
- [24] *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>.
- [25] G. Avoni et al. ‘The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS’. In: *JINST* 13.07 (2018), P07017. DOI: [10.1088/1748-0221/13/07/P07017](https://doi.org/10.1088/1748-0221/13/07/P07017).

- 563 [26] G. Aad et al. ‘Jet energy resolution in proton-proton collisions at $\sqrt{s} = 7$ TeV recorded
564 in 2010 with the ATLAS detector’. In: *The European Physical Journal C* 73.3 (Mar.
565 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:
566 <https://doi.org/10.1140/epjc/s10052-013-2306-0>.
- 567 [27] A. Collaboration. ‘Performance of b -jet identification in the ATLAS experiment’. In:
568 *Journal of Instrumentation* 11.04 (2016), P04008. URL: [http://stacks.iop.org/](http://stacks.iop.org/1748-0221/11/i=04/a=P04008)
569 [1748-0221/11/i=04/a=P04008](http://stacks.iop.org/1748-0221/11/i=04/a=P04008).
- 570 [28] ‘Observation of the associated production of a top quark and a Z boson at 13 TeV with
571 ATLAS’. In: (July 2020). URL: <https://cds.cern.ch/record/2722504>.
- 572 [29] ATLAS Collaboration. *ATLAS Computing Acknowledgements*. ATL-GEN-PUB-2016-
573 002. URL: <https://cds.cern.ch/record/2202407>.