Measurement of the production cross-section of a single top quark in association with a Z boson in proton–proton collisions at 13 TeV with the ATLAS detector

The ATLAS Collaboration

1. Introduction

At hadron colliders, the top quark is typically produced in \( t\bar{t} \) pairs through the strong interaction or as a single top or antitop quark through the electroweak interaction. The top quark was first observed via \( t\bar{t} \) production at the Tevatron [1,2]. This was followed by the observation of single-top-quark production [3–5] in the \( t \)- and \( s \)-channels, also at the Tevatron. The associated \( tW \) production was first observed in 8 TeV proton–proton collisions at the Large Hadron Collider (LHC) [6,7]. These single-top-quark channels allow a direct determination of the dominant \( tWb \) vertex and of the magnitude of the CKM matrix element \( |V_{tb}| \) [8] using their measured cross-sections.

With increasing energy and integrated luminosity, the ability to study rare Standard Model (SM) phenomena becomes possible. In the case of single top-quark production, examples include \( pp \rightarrow tZq \) [9] and \( pp \rightarrow tH \) [10]. The \( pp \rightarrow tZq \) process involves \( WWZ \) and \( tZ \) couplings and has not been observed so far [11].

Fig. 1 shows typical lowest-order Feynman diagrams for the process. This channel probes two SM couplings in a single process, whereas the similar final state \( t\bar{t}Z \) only probes the \( tZ \) coupling. The \( t\bar{t}Z \) process has been measured by the ATLAS [12,13] and CMS [14] collaborations. In addition, the production of \( pp \rightarrow tZq \) is a SM background to the \( tH \) final state [10].

This Letter presents evidence of the production of a single top quark in association with a Z boson in the \( t \)-channel process \( pp \rightarrow tZq \), where the Z boson decays into electrons or muons and the \( W \) boson from the top quark decays leptonically.

2. ATLAS detector

The ATLAS experiment [15] at the LHC is a multi-purpose particle detector with a forward–backward symmetric cylindrical geometry and a near 4π coverage in solid angle. It consists of an inner detector (ID) surrounded by a thin superconducting solenoid providing a 2T axial magnetic field, electromagnetic and hadron calorimeters, and a muon spectrometer. The inner detector covers the pseudorapidity range \(|\eta| < 2.5\). It consists of silicon pixel, silicon micro-strip and transition radiation tracking detectors. The innermost pixel layer, the insertable B-layer, was added between Run 1 and Run 2 of the LHC, at a radius of 33 mm around a new, thinner, beam pipe [16]. Lead/liquid-argon (LAr) sampling calorimeters provide electromagnetic (EM) energy measurements with high granularity. A hadron (steel/scintillator-tile) calorimeter covers the central pseudorapidity range \(|\eta| < 1.7\). The end-cap and forward regions are instrumented with LAr calorimeters for both the EM and hadronic energy measurements up to \(|\eta| = 4.9\).

\( 1 \) ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point [IP] in the centre of the detector and the \( z \)-axis along the beam pipe. The \( x \)-axis points from the IP to the centre of the LHC ring, and the \( y \)-axis points upwards. Cylindrical coordinates \((r, \phi)\) are used in the transverse plane, \( \phi \) being the azimuthal angle around the \( z \)-axis. The pseudorapidity is defined in terms of the polar angle \( \theta \) as \( \eta = -\ln \tan (\theta / 2) \). Distances in the \( \eta-\phi \) plane are measured in units of \( \Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \).
The muon spectrometer surrounds the calorimeters and is based on three large air-core toroidal superconducting magnets with eight coils each. The field integral of the toroids ranges between 2.0 and 6.0 T m across most of the detector. The muon spectrometer includes a system of precision tracking chambers and fast detectors for triggering. A two-level trigger system is used to select events. The first-level trigger is implemented in hardware and uses a subset of the detector information to reduce the accepted rate to at most 100 kHz. This is followed by a software-based trigger level that reduces the accepted event rate to 1 kHz on average.

3. Data and simulation samples

The pp collision data sample used in this measurement was collected with the ATLAS detector at the LHC during the 2015 and 2016 data-taking periods, corresponding to integrated luminosities of 3.3 fb$^{-1}$ and 32.8 fb$^{-1}$, respectively, for a total of 36.1 fb$^{-1}$, after requiring that the detector is fully operational. Events are considered if they were accepted by at least one of the single-muon or single-electron triggers [17,18]. The electron triggers select a cluster in the calorimeter matched to a track. Electrons must then satisfy identification criteria based on a multivariate technique using a likelihood discriminant. In 2015, electrons had to satisfy a 'medium' identification requirement and have a transverse energy of $E_T > 24$ GeV. In 2016, electrons had to satisfy a 'tight' identification together with an isolation criterion and have $E_T > 26$ GeV. To avoid efficiency loss due to isolation at high $E_T$, an additional trigger was used, selecting 'medium' electrons with $E_T > 60$ GeV. Muons are triggered on by matching tracks reconstructed in the muon spectrometer and in the inner detector. In 2015, muons had to satisfy a 'loose' isolation requirement and have a transverse momentum of $p_T > 20$ GeV. In 2016, the isolation criteria were tightened and the threshold increased to $p_T = 26$ GeV. In both years, another muon trigger without any isolation requirement was used, selecting muons with $p_T > 50$ GeV.

In order to evaluate the effects of the detector resolution and acceptance on signal and background and to estimate the SM background, a full GIAnt4-based detector simulation was used [19,20]. Event generators were used to estimate the expected signal and background contributions and their uncertainties. The top-quark mass in the event generators described below was set to 172.5 GeV. Multiple inelastic pp collisions (referred to as pile-up) are simulated with PYTHIA 8.186 [21], and are overlaid on each Monte Carlo (MC) event. Weights are assigned to the simulated events such that the distribution of the number of pile-up interactions in the simulation matches the corresponding distribution in the data. All simulation samples are processed through the same reconstruction algorithms as the data.

Monte Carlo $t\bar{t}Zq$ signal samples were generated at leading order (LO) in QCD using MG5_aMC@NLO 2.2.1 [22] in the four-flavour scheme, treating the $b$-quark as massive, with the CTEQ6L1 [23] LO parton distribution functions (PDFs). The $Z$ boson was simulated to be on-shell and off-shell $Z
on\text{\textit{p}}^* \text{\textit{p}}$ contributions and their interference are not taken into account. Following the discussion in Ref. [24], the renormalisation and factorisation scales ($\mu_T$ and $\mu_F$) used in MG5_aMC@NLO are set to $\mu_T = \mu_F = 4\sqrt{m_T^2 + p_T^2}$, where the $b$-quark is the external one produced from gluon splitting in the event. This choice is motivated by the total scale dependence being dominated by this external $b$-quark, shown in Fig. 1. The parton shower and the hadronisation of signal events were simulated with PYTHIA 6 [25] using the Perugia2012 set of tuned parameters [26]. The $t\bar{t}Zq$ total cross-section, calculated at next-to-leading order (NLO) using MG5_aMC@NLO 2.3.3 with the NNPDF3.0_nlo_as_0118 [27] PDF, is 800 fb, with an uncertainty of $\pm 6.1$%. The uncertainty is computed by varying the renormalisation and factorisation scales by a factor of two and by a factor of 0.5.

A comparison of the event kinematics before parton showering between the LO MG5_aMC@NLO 2.2.1 sample and a sample generated using NLO MG5_aMC@NLO 2.3.3 showed agreement within 10%, justifying the use of a LO sample for the detector simulation.

Monte Carlo simulated events are used to estimate the SM background that can produce three leptons and at least two jets in the final state. In $t\bar{t}$ production, if both $W$ bosons decay into leptons (referred to as 'prompt') and either a $b$- or $c$-hadron decays into a lepton (referred to as 'non-prompt') that is isolated, the final state can mimic the $t\bar{t}Zq$ final state. The nominal $t\bar{t}$ simulated sample was generated at NLO with the PowHEG-Box [28-30] event generator using the CT10 PDFs [31]. The cut-off parameter, $h_{\text{damp}}$, for the first emission of gluons was set to the top-quark mass. The events were then processed using PYTHIA 6 to perform the fragmentation and hadronisation, and to generate the underlying event.

Events from the associated production of a $t\bar{t}$ pair and a boson ($W/Z/H$) provide additional modes for the production of leptons in the final state. For $t\bar{t} + W$ the MC simulated events were generated using MG5_aMC@NLO 2.2.2 [22], while the $t\bar{t} + H$ and $t\bar{t} + Z$ MC simulated events were generated using MG5_aMC@NLO 2.2.3. The generated events were then processed with PYTHIA 8 [21] to perform the fragmentation and hadronisation, and to generate the underlying event, using the NNPDF2.3LO PDF set and the A14 tune [32].

Processes that include the production of $WW$, $WZ$ and $ZZ$ events were simulated using SHERRA 2.1.1 at LO with up to three additional partons and the CT10 PDF set. In the trilepton topology, the diboson background consists mainly of $WZ$ events, while the contribution to the background from $WW$ final states, corresponding to the case where a jet is misidentified as a lepton, is negligible. The $ZZ$ background gives a small contribution of 9% of all diboson events. The gluon-induced diboson production,
which amounts to about 10% of the quark-induced diboson production, is therefore negligible in the $t\bar{t}Zq$ signal region, and is not included in the diboson samples. In order to estimate the systematic uncertainty, additional diboson samples were simulated using the PowHEG-Box generator in combination with PYTHIA 8 and the CTEQ6L1 PDF sets.

Of the aforementioned single-top-quark production channels, only the $tW$ channel contributes to the trilepton final state. This sample was produced using the NLO POWHEG-Box event generator with the CT10 PDF set. The events were then processed with PYTHIA 6 to perform the fragmentation and hadronisation, and produce the underlying event. A sample of $tWZ$ events was produced using the MG5_aMC@NLO 2.2.3 generator and showered with PYTHIA 8, using the NNPDF3.0_NLO PDF set and the A14 tune.

4. Object reconstruction

The reconstruction of the basic physics objects used in this analysis is described in the following. The primary vertex is chosen as the proton–proton vertex candidate with the highest sum of the squared transverse momenta of all associated tracks with $p_T > 400$ MeV.

Electron candidates are reconstructed from energy deposits in the electromagnetic calorimeter that match a reconstructed track [33–36]. The clusters are required to be within $|\eta| < 2.47$ excluding the transition region between the barrel and end-cap calorimeters at 1.37 < $|\eta|$ < 1.52. Electron candidates must also satisfy a transverse energy requirement of $E_T > 15$ GeV. A likelihood-based discriminant is constructed from a set of variables that enhance the electron selection, while rejecting photon conversions and hadrons misidentified as electrons [34]. An $|\eta|$- and $p_T$-dependent selection on the likelihood discriminant is applied, such that it has an 80% efficiency when used to identify electrons from the Z-boson decay. This working point corresponds to an approximate rejection factor against jets of 700 at a $p_T$ of 40 GeV. Electrons are further required to be isolated using criteria based on ID tracks and topological clusters in the calorimeter, with an isolation efficiency of 90 (99)% for $p_T = 25$ (60) GeV. Correction factors are applied to simulated electrons to take into account the small differences in reconstruction, identification and isolation efficiencies between data and MC simulation.

Muon candidates are required to have $|\eta| < 2.5$ and $p_T > 15$ GeV, and are reconstructed by combining a reconstructed track from the inner detector with one from the muon spectrometer [37]. To reject misidentified muon candidates, primarily from pion and kaon decays, several quality requirements are imposed on the muon candidate. An isolation requirement based on ID tracks and topological clusters in the calorimeter is imposed, and results in an isolation efficiency of 90 (99)% for $p_T = 25$ (60) GeV. The overall efficiency obtained for muons from W-boson decays in simulated $pp \rightarrow t\bar{t}$ events is 96% and the rejection factor for non-prompt muons with $p_T > 20$ GeV is approximately 600. As for electrons, correction factors are applied to muons to account for the small differences between data and simulation.

Jets are reconstructed from topological clusters using the anti-$k_t$ algorithm [38,39] with the radius parameter set to $R = 0.4$. They are reconstructed for $p_T > 30$ GeV in the region with $|\eta| < 4.5$. To account for inhomogeneities and the non-compensating response of the calorimeter, the reconstructed jet energies are corrected using $p_T$- and $\eta$-dependent factors that are derived in MC simulation and validated in data. Any remaining differences in the jet energy scale are corrected using in situ techniques, where a well-defined reference object is momentum-balanced with a jet [40]. To suppress pile-up, a discriminant called the jet-vertex-trigger (JVT) is constructed using a two-dimensional likelihood method [41]. For jets with $p_T < 60$ GeV and $|\eta| < 2.4$ a JVT requirement corresponding to a 92% efficiency, while rejecting 98% of jets from pile-up and noise, is imposed.

To identify jets containing a b-hadron (b-tagging), a multivariate algorithm is employed [42]. This algorithm uses the impact parameter and reconstructed secondary vertex information of the tracks contained in the jet as input for a neural network. Due to its use of the inner detectors, the reconstruction of b-jets is done in the region with $|\eta| < 2.5$. Jets initiated by b-quarks are selected by setting the algorithm’s output threshold such that a 77% b-jet selection efficiency is achieved in simulated $t\bar{t}$ events. With this setting, the misidentification rate for jets initiated by light-flavour quarks or gluons is 1%, while it is 17% for jets initiated by c-quarks [43]. Correction factors are derived and applied to correct for the small differences in b-quark selection efficiency between data and MC simulation [42].

The missing transverse momentum, with magnitude $E_T^{miss}$, is calculated as the negative of the vector sum of the transverse momenta of all reconstructed objects, $p_T^{miss}$. In addition to the identified jets, electrons and muons, a track-based ‘soft’ term is included in the $p_T^{miss}$ calculation, by considering tracks associated with the primary vertex in the event but not with an identified jet, electron, or muon [44,45].

To avoid cases where the detector response to a single physical object is reconstructed as two separate final-state objects, several steps are followed to remove such overlaps, following Ref. [46].

5. Signal, control and validation regions

The reconstructed $t\bar{t}Zq$ final state consists of three charged leptons (electron and/or muon), a b-tagged jet, an additional jet and $E_T^{miss}$. Reconstructing the Z boson and the top quark is important in order to identify specific features that help to separate the signal from the background. For example, the Z-boson mass distributions can contribute to the reduction of top-quark backgrounds, as these do not include a Z boson in the final state, while the untagged-jet pseudorapidity distribution differs in shape between $t\bar{t}$Z signal events and diboson and $t\bar{t}Z$ events, which constitute some of the largest backgrounds.

The signal region (SR) definition reflects the $t\bar{t}Zq$ final state by selecting only events that have exactly three charged leptons, one b-tagged jet and one additional jet, referred to as the untagged jet as no b-tagging requirement is applied. In order to better separate the $t\bar{t}Zq$ signal from background, additional requirements are imposed on the properties of the selected objects. The three leptons are sorted by their $p_T$, irrespective of flavour, and required to have transverse momenta of at least 28, 25 and 15 GeV, respectively. Both jets are required to have $p_T > 30$ GeV.

An opposite-sign, same-flavour (OSSF) lepton pair is required in order to reconstruct the Z boson. In the $ee\mu\mu$ and $e\mu\mu\mu$ channels, the pair is uniquely identified. For the $ee\mu\mu\mu$ events, both possible combinations are considered and the pair that has the invariant mass closest to the Z-boson mass is chosen. The W boson is reconstructed from the remaining lepton and the missing transverse momentum, using as constraint the W-boson mass to evaluate the z component of the neutrino momentum. The top quark is reconstructed from the reconstructed W boson and the b-tagged jet.

To suppress background sources that do not contain a Z boson, the invariant mass of the leptons is required to be between 81 and 101 GeV. Because a W boson is expected in the final state,

\footnotesize
\begin{enumerate}
\item In case of an imaginary solution, the $p_T^{miss}$ value is varied until one real solution is found.
\end{enumerate}
the reconstructed transverse mass of the W-boson candidate is required to be $\sqrt{T^2 + m_{T^2}^2} > 20\text{GeV}$.

The selection criteria that define the SR are summarised in Table 1. In total, 141 events are selected using these criteria. The criteria are modified to define validation regions, which are used to check the modelling of the main background contributions. Two validation regions (VR) are defined as follows: the diboson VR uses the same event selection as the SR, except that only one jet is required in the event and no $b$-tagging requirement is applied. The $t\bar{t}$ VR also uses the same selection as the SR, except that the invariant mass of the $t\bar{t}$ pair must be outside the $Z$-mass window $(m_{tt} < 81\text{ GeV} \text{ or } m_{tt} > 101\text{ GeV})$. In addition, two control regions (CR) are defined, from which the normalisations of the diboson and the $t\bar{t}$ background sources are computed, as explained in Section 6. The diboson CR is defined in the same way as the diboson VR, except with a tighter requirement on $m_T(W, \nu)$. The $t\bar{t}$ CR instead has the same selection as the SR but it requires an opposite-sign, different-flavour (OSDF) lepton pair and rejects events with an OSDF pair.

### Table 1
Overview of the requirements applied for selecting events in the signal, validation and control regions.

<table>
<thead>
<tr>
<th>Common Selections</th>
<th>Diboson VR / CR</th>
<th>$t\bar{t}$ VR</th>
<th>$t\bar{t}$ CR</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\geq$ 1 OSDF pair</td>
<td>$\geq$ 1 OSDF pair</td>
<td>$\geq$ 1 OSDF pair</td>
<td>$\geq$ 1 OSDF pair</td>
</tr>
<tr>
<td>$m_{T^2} &lt; 10\text{ GeV}$</td>
<td>$m_{T^2} &lt; 10\text{ GeV}$</td>
<td>$m_{T^2} &lt; 10\text{ GeV}$</td>
<td>$m_{T^2} &lt; 10\text{ GeV}$</td>
</tr>
<tr>
<td>2 jets, $</td>
<td>\eta</td>
<td>&lt; 4.5$</td>
<td>1 jet, $</td>
</tr>
<tr>
<td>1 b-jet, $</td>
<td>\eta</td>
<td>&lt; 2.5$</td>
<td>$\geq$ 1 OSSF pair</td>
</tr>
<tr>
<td>$m_T(W, \nu) &gt; 20/60\text{GeV}$</td>
<td>$\geq$ 1 OSSDF pair</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

6. Background estimation

Different SM processes are considered as background sources for this analysis. These are either processes such as diboson or $t\bar{t}V + t\bar{t}H$ production, in which three or more prompt leptons are produced, or processes with only two prompt leptons in the final state (such as $Z +$ jets and $t\bar{t}$ production) and one additional non-prompt or fake lepton that meets the selection criteria. Such non-prompt or fake leptons can originate from decays of bottom or charm hadrons, a jet that is misidentified as an electron, leptons from kaon or pion decays, or electrons from photon conversions.

The dominant source of background originates from diboson production. This consists mainly of $WZ$ events with a small fraction of $ZZ$ events in which the fourth lepton is missed (roughly 9% of the total number of diboson events). Studies in the diboson VR indicated that the number of events predicted by the Sherpa MC samples is lower than the number observed. The kinematic distributions are otherwise well described. Hence, the total number of diboson events predicted by the Sherpa samples is scaled by a factor of 1.47, leading to an expected number of diboson events in the SR of 53. The scale factor is derived from the diboson CR, defined in Section 5, by comparing the data-to-MC ratio for events that satisfy the condition $m_T(W) > 60\text{GeV}$. This selection is applied in order to reduce the $Z +$ jets contamination and ensure a diboson-dominated region. The uncertainty in the scale factor is estimated by varying the requirement for the $m_T(W)$ selection. An additional uncertainty in the diboson estimate is assigned by evaluating the difference in the number of events in the signal region when using the default Sherpa samples and a set of POWHEG samples. This results in an estimated uncertainty of 30%, also taking into account the extrapolation of the scale factor from the CR to the SR.

The main sources of non-prompt or fake-lepton background for this analysis are $t\bar{t}$ and $Z +$ jets events. These two contributions are evaluated separately. This choice is motivated by MC generator-level studies showing that although very similar in origin, the source of the non-prompt or fake lepton is usually different for processes involving top quarks compared to $Z +$ jets events. For $t\bar{t}$ events, in most cases, it is the softer of the two leptons assigned to the reconstructed $Z$ boson, while for $Z +$ jets events it is the lepton not assigned to the $Z$ boson.

In order to take into account a possible difference between data and MC simulation for $t\bar{t}$ events, the number of events containing a non-prompt or fake lepton in the MC simulation is scaled by a data/MC factor that is derived in the $t\bar{t}$ CR defined in Section 5. This $t\bar{t}$ control region and the signal region have very similar non-prompt lepton compositions. Requiring a pair of opposite-sign, different-flavour leptons, and rejecting events with an OSDF pair, ensures that there is no contamination from $Z +$ jets events and from the SR. Different electron–muon invariant mass windows around the $Z$ mass, with widths ranging from 20 GeV to 60 GeV, were investigated and the average of the obtained factors is used for scaling the $t\bar{t}$ background in the signal region. The total uncertainty in the scaling factor is calculated taking into account this variation and the statistical uncertainty of the sample. This leads to a data/MC scale factor of 1.21 ± 0.51. Deriving separate factors depending on the fake lepton’s flavour or on the lepton $p_T$ was also investigated. All approaches are consistent with each other within the assigned uncertainties. The expected number of $t\bar{t}$ events in the SR is $18 \pm 9$. According to the MC prediction, the $tW$ contribution is found to be less than one event.

A data-driven technique called the fake-factor method is used to estimate the $Z +$ jets background contribution. A region defined by selecting events with $m_T(W) < 20\text{GeV}$ is used for deriving the fake factors. Since it is observed that the number of non-prompt or fake electrons and muons can be very different, the estimation is done separately for the electron and muon channel. Fake factors are defined as the ratio of data events that have three isolated leptons to events in which one of the leptons fails the isolation requirement. They are derived in bins of the $p_T$ of the lepton not associated with the $Z$ boson. According to MC simulation, this lepton is in over 95% of the cases the non-prompt or fake lepton. These factors are then applied to events passing the signal region selection (including a $m_T(W) > 20\text{GeV}$ cut) that have one of the three leptons failing the isolation requirement. Contamination from other background sources, which is about 50% and mainly coming from $t\bar{t}$, is taken into account and subtracted before making the

---

3 The transverse mass is calculated using the momentum of the lepton associated with the $W$ boson, $p_T^{\text{miss}}$, and the azimuthal angular difference between the two:

$$m_T(\ell, p_T^{\text{miss}}) = \sqrt{2p_T(\ell)p_T^{\text{miss}}[1 - \cos(\Delta \phi(\ell, p_T^{\text{miss}}))]}. $$
The expected number of Z + jets events in the SR is 37. Different sources of uncertainty are investigated, including consistency checks of the fake-factor method using MC Z + jets samples, the effect of changing the diboson scale factor and the statistical uncertainties in the estimated and observed number of events. All these amount to a total uncertainty of 40%.

The expected t\bar{t}V, t\bar{t}H and tWZ contributions are evaluated from the MC samples normalised to their predicted NLO cross-sections [22]. The t\bar{t}V + t\bar{t}H contribution is approximately 10% of the total background estimate, while tWZ events amount to 3%. The expected number of t\bar{t}V + t\bar{t}H + tWZ events is 20 ± 3. The uncertainty in the predictions is taken to be 13% [22].

### 7. Multivariate analysis

A multivariate analysis is used to separate the signal from the large number of background events. The neural-network package NeuroBayes [47,48] is used, which combines a three-layer feed-forward neural network with a complex robust preprocessing. Several variables are combined into one discriminant, then mapped onto the interval [0, 1], such that background-like events have an output value, O_{NN}, closer to 0 and signal-like events have an output closer to 1. All background processes are considered in the training except t\bar{t} production, due to the very small number of available MC events that meet the selection criteria. Only variables that provide separation power and are well modelled are taken into account in the final neural network (NN). For the NN training, the ten variables with the highest separation power are used. These variables are explained in the order of their importance in Table 2. They include simple variables, such as the \pt and \eta of jets and of the lepton not associated with the Z boson. Information about the reconstructed W boson, Z boson and top quark, such as their \pt as well as their masses, is also used. In addition, the \Delta R between the untagged jet and the Z boson is employed as an input.

The modelling of the input variables is checked both in the validation regions defined in Table 1 and in the signal region. The distributions of some input variables in the signal region are shown in Fig. 2, normalised to the expected number of events, including the scale factors determined in Section 6. Good agreement between data and the prediction is observed.

The output of the NN is checked in the validation regions, shown in Fig. 3. Good agreement between the expected and ob-
served numbers of events and in the shape of the NN output distribution are seen, demonstrating reliable background modelling. NeuroBayes includes extensive protection against overtraining and several further checks confirm that it functions well.

8. Systematic uncertainties

Systematic uncertainties in the normalisation of the individual backgrounds and in the signal acceptance, as well as uncertainties in the shape of the NN distributions, are taken into account when determining the $t\bar{t}q$ cross-section. For uncertainties where variations as a function of the NN distribution are consistent with being due to statistical fluctuations, only the normalisation difference is taken into account. The uncertainties are split into the following categories:

Reconstruction efficiency and calibration uncertainties Systematic uncertainties affecting the reconstruction and energy calibration of jets, electrons and muons are propagated through the analysis. The dominant sources of uncertainty for this measurement are the jet energy scale (JES) calibration, including the modelling of pile-up, and the b-jet tagging efficiencies.

The uncertainties due to lepton reconstruction, identification, isolation requirements and trigger efficiencies are estimated using tag-and-probe methods in $Z \rightarrow \ell\ell$ events. Correction factors are derived to match the simulation to observed distributions in collision data and associated uncertainties are estimated. Uncertainties in the lepton momentum scale and resolution are also assessed using $Z \rightarrow \ell\ell$ events.

Several components of the JES uncertainty are considered [40, 50]. Uncertainties derived from different dijet-$p_T$-balance measurements as well as uncertainties associated with other in situ calibration techniques are considered. Furthermore, the presence of nearby jets and the modelling of pile-up affect the jet calibration. The uncertainty in the flavour composition covers effects due to the difference in quark–gluon composition between the jets used in the calibration and the jets used in the analysis. Also an uncertainty due to the different calorimeter responses to light-quark and gluon jets is taken into account. Finally, the JES uncertainty is estimated for $b$-jets by varying the modelling of $b$-quark fragmentation. The uncertainty in the jet energy resolution (JER) and the one associated with the JVT requirement are also considered [51]. The jet-related uncertainties with the highest impact on the final result are the JER and the flavour composition.

The impact of a possible miscalibration on the soft-track component of $E_T^{\text{miss}}$ is derived from data–MC comparisons of the $p_T$ balance between the hard and soft $E_T^{\text{miss}}$ components [45]. The uncertainty associated with the leptons and jets is propagated from the corresponding uncertainties in the energy/momentum scales and resolutions, and it is classified together with the uncertainty associated with the corresponding objects.

Since the analysis makes use of $b$-tagging, the uncertainties in the b-tagging efficiency and the mistag rate are taken into account. These uncertainties were determined using $\sqrt{s} = 8$ TeV data as described in Ref. [52] for $b$-jets and Ref. [53] for light jets, with additional uncertainties to account for the presence of the newly added inner layer of the pixel detector and the extrapolation to $\sqrt{s} = 13$ TeV.

Signal PDF and radiation The systematic effects due to uncertainties in the parton distribution functions are taken into account for the signal. As it was generated at LO, the uncertainty is evaluated using the 30 eigenvectors of the NNPDF3.0_lo_as_0118 [27] PDF set, in the four-flavour scheme. The events are reweighted according to each of the PDF uncertainty eigenvectors. As a cross-check, the PDF uncertainty is also evaluated following the updated PDF4LHC recommendation [54] by using the PDF4LHC15 NLO PDF set. This has a smaller effect; hence the uncertainty from the LO PDF set is used.

Variations of the amount of additional radiation are studied by changing the hard-scatter scales and the scales in the parton shower simultaneously in the $t\bar{t}q$ sample. A variation of the factorisation and renormalisation scale by a factor of two is combined with the Perugia2012 set of tuned parameters with lower radiation (P2012radlo) than the nominal set; while a variation of both scales by a factor of 0.5 is combined with the Perugia2012 set of tuned parameters with higher radiation (P2012radHi).

Luminosity The uncertainty in the combined 2015+2016 integrated luminosity is 2.1%. It is derived, following a methodology similar to that detailed in Ref. [55], from a calibration of the luminosity scale using $x$–$y$ beam-separation scans performed in August 2015 and May 2016.

The effects of the above uncertainties on the number of signal events are summarised in Table 3. This does not include the impact of the background uncertainties.

Background The uncertainties in the normalisation of the various background processes use the uncertainty estimated in Section 6. For the $t\bar{t}$ sample, the systematic effects due to uncertainties in the scale and the amount of radiation are included.
9. Results

Using the 141 selected events, a maximum-likelihood fit is performed to extract the $t\bar{Z}q$ signal strength, $\mu$, defined as the ratio of the measured signal yield to the NLO Standard Model prediction. The statistical analysis of the data employs a binned likelihood function $L(\mu, \bar{\theta})$, constructed as the product of Poisson probability terms, to estimate $\mu$ [56]. The likelihood is maximised on the NN output distribution in the signal region. The background normalisations are allowed to vary within the uncertainties given in Section 6.

The impact of systematic uncertainties on the expected numbers of signal and background events is described by nuisance parameters, $\bar{\theta}$, which are each parameterised by a Gaussian or log-normal constraint for each bin of the NN output distribution. If the variation of the uncertainty in each bin is consistent with being due to statistical fluctuations, only the overall change in normalisation is included as a nuisance parameter. The uncertainties are set to be symmetric in the fit, using the average of the variations up and down. The expected numbers of signal and background events in each bin are functions of $\bar{\theta}$. The test statistic, $q_{\mu}$, is constructed according to the profile likelihood ratio:

$$q_{\mu} = -2\ln(L(\hat{\mu}, \hat{\theta})/L(\bar{\mu}, \bar{\theta})),$$

where $\hat{\mu}$ and $\hat{\theta}$ are the parameters that maximise the likelihood, and $\bar{\mu}$ and $\bar{\theta}$ are the nuisance parameter values that maximise the likelihood for a given $\mu$. This test statistic is used to determine a probability for accepting the background-only hypothesis for the observed data.

Fig. 4 shows the NN discriminant in the signal region with background normalisations, signal normalisation and nuisance parameters adjusted by the profile likelihood fit.

The results for the numbers of fitted signal and background events are summarised in Table 4. The table also shows the result of a fit to the Asimov dataset [56]. The total uncertainty in the number of fitted events includes the effect of correlations, which are large among the background sources, as the $O_{\text{NN}}$ distributions have a similar shape. The strongest correlation is found to be between the diboson and the $Z + \text{jets}$ contributions and it is about $-0.5$ for both the Asimov dataset and the data.

After performing the binned maximum-likelihood fit and estimating the total uncertainty, the fitted value for $\mu$ is $0.75 \pm 0.21$ (stat.) $\pm 0.17$ (syst.) $\pm 0.05$ (th.). The quoted theory (th.) uncertainty in $\mu$ includes the $t\bar{Z}q$ NLO cross-section uncertainty given in Section 3. This is not taken into account when evaluating the cross-section. The statistical uncertainty in the cross-section is determined by performing a fit to the data, including only the statistical uncertainties. The total systematic uncertainty is determined by subtracting this value in quadrature from the total uncertainty. The cross-section for $t\bar{Z}q$ production is measured to be $600 \pm 170$ (stat.) $\pm 140$ (syst.) fb, assuming a top-quark mass of $m_t = 172.5$ GeV.

The probability $p_0$ of obtaining a result at least as signal-like as observed in the data if no signal were present is calculated using the test statistic $q_{\mu=0}$ in the asymptotic approximation [56]. The observed $p_0$ value is $3.2 \times 10^{-5}$. The resulting significance is $4.2\sigma$, to be compared with the expected significance of $5.4\sigma$.

10. Conclusion

The cross-section for $t\bar{Z}q$ production has been measured using $36.1\, \text{fb}^{-1}$ of proton-proton collision data collected by the ATLAS experiment at the LHC in 2015 and 2016 at a centre-of-mass energy of $\sqrt{s} = 13\, \text{TeV}$. Evidence for the signal is obtained with a measured (expected) significance of $4.2\sigma$ ($5.4\sigma$). The measured cross-section is $600 \pm 170$ (stat.) $\pm 140$ (syst.) fb. This result is in agreement with the predicted SM $t\bar{Z}q$ cross-section, calculated at NLO to be 800 fb with a scale uncertainty of $\pm 7.1\%$.

Acknowledgements

We thank CERN for the 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; BMWFW and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI,
Canada; CERN; CONICYT, Chile; CAS, MOST, and NSFC, China; CONICET, Argentina; CNRS/IN2P3, CEA/DSM/IRFU, France; BMBF, HGF, and MPG, Germany; GSRT, Greece; RGC, Hong Kong SAR, China; ISP, 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; ARES 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. [57].

References

The ATLAS Collaboration
