Measurement of the cross-section for producing a $W$ boson in association with a single top quark in $pp$ collisions at $\sqrt{s} = 13$ TeV with ATLAS

The ATLAS collaboration

$E$-mail: atlas.publications@cern.ch

ABSTRACT: The inclusive cross-section for the associated production of a $W$ boson and top quark is measured using data from proton-proton collisions at $\sqrt{s} = 13$ TeV. The dataset corresponds to an integrated luminosity of $3.2\,\text{fb}^{-1}$, and was collected in 2015 by the ATLAS detector at the Large Hadron Collider at CERN. Events are selected requiring two opposite sign isolated leptons and at least one jet; they are separated into signal and control regions based on their jet multiplicity and the number of jets that are identified as containing $b$ hadrons. The $Wt$ signal is then separated from the $t\bar{t}$ background using boosted decision tree discriminants in two regions. The cross-section is extracted by fitting templates to the data distributions, and is measured to be $\sigma_{Wt} = 94\pm10\,\text{(stat.)}^{+28}_{-22}\,\text{(syst.)}\pm2\,\text{(lumi.)}\,\text{pb}$. The measured value is in good agreement with the SM prediction of $\sigma_{\text{theory}} = 71.7\pm1.8\,\text{(scale)}\pm3.4\,\text{(PDF)}\,\text{pb}$ [1].

KEYWORDS: Hadron-Hadron scattering (experiments)

ArXiv ePrint: 1612.07231
1 Introduction

Top quarks can be produced singly via electroweak interactions involving a $Wtb$ vertex. In the Standard Model (SM), single top quark production proceeds via three channels at leading order (LO), represented in figures 1 and 2: production in association with a $W$ boson ($Wt$), the $t$-channel and the $s$-channel. At the Large Hadron Collider (LHC), the $Wt$ channel is the mode with the second largest production cross-section, behind the dominant $t$-channel mode. The $Wt$ channel represents approximately 24% of the total single-top-quark production rate at $\sqrt{s} = 13$ TeV, making it experimentally accessible for detailed measurements.

The cross-section for each of the three single-top-quark production channels is sensitive to the coupling between the $W$ boson and the top quark. This coupling is parameterised by the relevant Cabibbo-Kobayashi-Maskawa (CKM) matrix element $V_{tb}$ and form factor $f_{V}^{L} [2-4]$ such that the proportionality is given by $|f_{V}^{L}V_{tb}|^{2} [5, 6]$, assuming a left-handed vector interaction as given in the SM. Single top quark production therefore presents an opportunity for testing the structure of the SM, as well as probing classes of new-physics models that can affect the $Wtb$ vertex. In contrast to the $t$- and $s$-channels, which are sensitive to both the existence of four-fermion operators and corrections to the $Wtb$ vertex,
the $Wt$ channel only depends on the latter; it is therefore important to study this channel separately to provide a comparison with the other channels [7, 8].

The $Wt$ channel was not accessible at the Tevatron due to its small cross-section in $pp$ collisions at $\sqrt{s} = 1.96\,\text{TeV}$. At the LHC, however, evidence of this process with 7 TeV collision data was presented by the ATLAS Collaboration [9] and by the CMS Collaboration [10]. With 8 TeV collision data, observations were made by the CMS Collaboration [11] and the ATLAS Collaboration [12] with cross-section measurements in good agreement with theoretical predictions.

The $Wt$ NLO cross-section at a $\sqrt{s} = 13\,\text{TeV}$ with next-to-next-to-leading logarithmic (NNLL) soft-gluon corrections is calculated as $\sigma_{\text{theory}} = 71.7 \pm 1.8\,\text{(scale)} \pm 3.4\,\text{(PDF)}\,\text{pb}$ [1], assuming a top quark mass ($m_{\text{top}}$) of 172.5 GeV. The first uncertainty accounts for the renormalisation and factorisation scale variations (from $m_{\text{top}}/2$ to $2\,m_{\text{top}}$), while the second uncertainty originates from uncertainties in the MSTW2008 NLO parton distribution function (PDF) sets [13].

This paper describes a measurement of the cross-section of the $Wt$ process using $\sqrt{s} = 13\,\text{TeV}$ proton-proton ($pp$) collisions with an integrated luminosity of $3.2\,\text{fb}^{-1}$. The data were recorded with the ATLAS detector in 2015. The measurement is made using events containing at least one $b$ jet (according to the definition in section 4) and exactly two oppositely charged leptons in the final state, where a lepton ($\ell$) is defined to be either an electron ($e$) or a muon ($\mu$), whether produced directly from the decay of a $W$ boson or from the decay of an intermediate $\tau$ lepton. The $Wt$ signal enters this final state when the top
quark decays into a $W$ boson and a quark (which is assumed to be a $b$-quark), with both $W$ bosons subsequently decaying into a neutrino and a lepton, as depicted in figure 1. A minimal selection is applied to reduce background contributions from $Z/\gamma^*+\text{jets}$ (hereafter called $Z+\text{jets}$) events, diboson events, and events containing leptons that are misidentified or arise from the decay of hadrons. A boosted decision tree (BDT) analysis is performed to construct discriminants capable of separating the $Wt$ signal from the dominant top quark pair ($t\bar{t}$) background, and these discriminants are used in a profile-likelihood fit to extract the $Wt$ cross-section. The top pair production background is described by simulation, which has been validated in previous ATLAS measurements \cite{14}.

The measurement technique is similar to that employed in the corresponding 8 TeV ATLAS measurement \cite{12}. The most significant changes include modifications to the BDT training and the binning of the distribution used in the likelihood fit (discussed in section 6 and section 8 respectively), and an optimisation of kinematic requirements to more effectively reject $Z+\text{jets}$ and other small backgrounds (discussed in section 5).

2 The ATLAS detector

The ATLAS detector \cite{15} at the LHC covers nearly the entire solid angle\footnote{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 upward. 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 $\eta = -\ln \tan(\theta/2)$, while the rapidity is defined in terms of particle energies and the z-component of particle momenta as $y = (1/2) \ln [(E + p_z)/(E - p_z)]$.} around the collision point, and consists of an inner tracking detector (ID) surrounded by a thin superconducting solenoid magnet producing a 2 T axial magnetic field, electromagnetic (EM) and hadronic calorimeters, and an external muon spectrometer (MS). The ID consists of a high-granularity silicon pixel detector and a silicon microstrip tracker, together providing precision tracking in the pseudorapidity range $|\eta| < 2.5$, complemented by a transition radiation tracker providing tracking and electron identification information for $|\eta| < 2.0$. The innermost pixel layer, the insertable B-layer, was added between Run 1 and Run 2 of the LHC, at an innermost radius of 33 mm around a new, thinner, beam pipe \cite{16}. A lead liquid-argon (LAr) electromagnetic calorimeter covers the region $|\eta| < 3.2$, and hadronic calorimetry is provided by steel/scintillator tile calorimeters within $|\eta| < 1.7$ and copper/LAr hadronic endcap calorimeters in the range $1.5 < |\eta| < 3.2$. A LAr forward calorimeter with copper and tungsten absorbers covers the range $3.1 < |\eta| < 4.9$. The MS consists of precision tracking chambers covering the region $|\eta| < 2.7$, and separate trigger chambers covering $|\eta| < 2.4$. A two-level trigger system, using a custom hardware level followed by a software-based level, selects from the 40 MHz of collisions a maximum of around 1 kHz of interesting events for offline storage.
3 Data and simulation

The data events analysed in this paper correspond to an integrated luminosity of $3.2 \, \text{fb}^{-1}$ collected from the operation of the LHC in 2015 at $\sqrt{s} = 13$ TeV with a bunch spacing of $25 \, \text{ns}$ and an average number of collisions per bunch crossing $\langle \mu \rangle$ of around 14. They are required to be recorded in periods where all detector systems are flagged as operating normally. Additionally, individual events identified as containing corrupted data are rejected.

Monte Carlo (MC) simulation samples are used to estimate the efficiency to select signal and background events, train and test BDTs, estimate systematic uncertainties, and validate the analysis tools. All simulation samples are normalised to theoretical cross-section predictions. The nominal samples (used for estimating the central values for efficiencies and background templates) were simulated with a full ATLAS detector simulation [17] implemented in Geant4 [18]. Many of the samples used in the estimation of systematic uncertainties were instead produced using Atlfast2 [19], which differs from the full simulation in that the ATLAS calorimeters and their responses are simulated using a faster approximation. Pile-up (additional $pp$ collisions in the same or a nearby bunch crossing) was included in the simulation by overlaying collisions with the soft QCD processes of Pythia 8.186 [20] using a set of tuned parameters called the A2 tune [21] and the MSTW2008LO PDF set. Events were generated with a predefined distribution of the expected number of interactions per bunch crossing, then reweighted to match the actual observed data conditions. In all samples used for this analysis $m_{\text{top}}$ was set to 172.5 GeV and the $W \rightarrow \ell \nu$ branching ratio was set to 0.1080 per lepton flavour.

For the generation of $Wt$ and $t\bar{t}$ event samples [22], the Powheg-Box v1 (v2 for $t\bar{t}$) [23–27] generator with the CT10 PDF set [28] in the matrix element calculations is used. For these processes, top quark spin correlations are preserved. The parton shower, fragmentation, and underlying event were simulated using Pythia 6.428 [29] with the CTEQ6L1 PDF set [30] and the corresponding Perugia 2012 (P2012) tune [31]. The EvtGen v1.2.0 program [32] was used to simulate properties of the bottom and charmed hadron decays. The renormalisation and factorisation scales are set to $m_{\text{top}}$ for the $Wt$ process and $\sqrt{m_{\text{top}}^2 + p_T(t)^2}$ for the $t\bar{t}$ process. The diagram removal (DR) scheme [33], in which all next-to-leading order (NLO) diagrams that overlap with the $t\bar{t}$ definition are removed from the calculation of the $Wt$ amplitude, was employed to handle interference between $Wt$ and $t\bar{t}$ diagrams, and was applied to the $Wt$ sample.

The $t\bar{t}$ cross-section is set to $\sigma_{t\bar{t}} = 252.9^{+6.4}_{-8.6}$ (scale) $\pm 11.7$ (PDF + $\alpha_S$) pb as calculated with the Top++2.0 program to NNLO, including soft-gluon resummation to NNLL [34]. The first uncertainty comes from the independent variation of the factorisation and renormalisation scales, $\mu_F$ and $\mu_R$, while the second one is associated with variations in the PDF and $\alpha_S$, following the PDF4LHC prescription with the MSTW2008 68% CL NNLO, CT10 NNLO and NNPDF2.3 5f FFN PDF set [35–38]. Both calculations assume $m_{\text{top}} = 172.5$ GeV.

Additional $Wt$ samples were generated to estimate major systematic uncertainties. An alternative $Wt$ sample was generated using the diagram subtraction (DS) scheme instead of DR, where a gauge-invariant subtraction term modifies the NLO $Wt$ cross-
section to locally cancel the double-resonant $t\bar{t}$ contribution [33]. Another sample generated with \textsc{MadGraph5.aMC@NLO} [39] (instead of the \textsc{Powheg-Box}) interfaced with \textsc{Herwig++ 2.7.1} [40] using \textsc{Atlfast2} fast simulation is used to estimate uncertainties associated with the modelling of the NLO matrix element generator. A sample generated with \textsc{Powheg-Box} interfaced with \textsc{Herwig++} (instead of \textsc{Pythia 6}) is used to estimate uncertainties associated with the parton shower, hadronisation, and underlying-event models. In both cases the UE-EE-5 tune of ref. [41] was used for the underlying event, and \textsc{EvtGen v1.2.0} was used to simulate properties of the bottom and charmed hadron decays. Finally, in order to estimate uncertainties arising from additional QCD radiation in the $Wt$ events, a pair of samples were generated with \textsc{Powheg-Box} interfaced with \textsc{Pythia 6} using \textsc{Atlfast2} and the P2012 tune with higher and lower radiation relative to the nominal set, together with varied renormalisation and factorisation scales. In these samples the resummation damping factor was doubled in the case of higher radiation. In order to avoid comparing two different detector response models when estimating systematic uncertainties, another version of the nominal \textsc{Powheg-Box} with \textsc{Pythia 6} sample was also produced with fast simulation.

Additional $t\bar{t}$ samples were also generated to estimate major systematic uncertainties. As with the additional $Wt$ samples, these are used to estimate the uncertainties associated with the matrix element generator (a sample produced using \textsc{Atlfast2} fast simulation with \textsc{MadGraph5.aMC@NLO} interfaced with \textsc{Herwig++ 2.7.1}), parton shower and hadronisation models (a sample produced using \textsc{Atlfast2} with \textsc{Powheg-Box} interfaced with \textsc{Herwig++ 2.7.1}) and additional QCD radiation (a pair of samples produced using full simulation with the P2012 higher and lower radiation-varied sets of parameters, as well as with varied renormalisation and factorisation scales).

Samples used to model the $Z+\text{jets}$ background [42] were simulated with \textsc{Sherpa 2.1.1} [43]. Matrix elements were calculated for up to two partons at NLO and four partons at LO using the \textsc{Comix} [44] and \textsc{OpenLoops} [45] matrix element generators and merged with the \textsc{Sherpa} parton shower [46] using the ME+PS@NLO prescription [47]. The CT10 PDF set was used in conjunction with \textsc{Sherpa} parton shower tuning, with a generator-level cutoff on the dilepton invariant mass of $m_{\ell\ell}>40 \text{ GeV}$ applied. The $Z+\text{jets}$ events are normalised to NNLO cross-sections.

Diboson processes with four charged leptons, three charged leptons and one neutrino, or two charged leptons and two neutrinos [48] were simulated using the \textsc{Sherpa 2.1.1} generator. The matrix elements contain all diagrams with four electroweak vertices. The NLO corrections are used for the purely leptonic final states as well as for final states with two or four charged leptons plus one additional parton. For other final states with up to three additional partons, the LO calculations of the \textsc{Comix} and \textsc{OpenLoops} generators are used. Their outputs are combined with the \textsc{Sherpa} parton shower using the ME+PS@NLO prescription [47]. The PDF set used was CT10 with dedicated parton shower tuning. The generator-calculated cross-sections are used for diboson processes (already at NLO).

Finally, the very small $W+\text{jets}$ contribution was simulated using \textsc{Powheg-Box v2} interfaced to the \textsc{Pythia 8.186} [20] parton shower model. The CT10 PDF set was used in the matrix element. The AZNLO [49] tune was used, with PDF set CTEQ6L1, for
the modelling of non-perturbative effects, and the EvtGen v1.2.0 program was used to simulate properties of the bottom and charmed hadron decays.

4 Object selection

Electron candidates are reconstructed from energy deposits in the EM calorimeter associated with ID tracks [50]. The deposits are required to be in the $|\eta| < 2.47$ region, with the transition region between the barrel and endcap EM calorimeters, $1.37 < |\eta| < 1.52$, excluded. The candidate electrons are required to have transverse energy $p_T > 20$ GeV. Further requirements on the electromagnetic shower shape, calorimeter energy to tracker momentum ratio, and other discriminating variables are combined into a likelihood-based object quality selection [50], optimised for strong background rejection. Candidate electrons also must satisfy requirements on the distance of the ID track to the reconstructed primary vertex in the event, which is identified as the vertex with the largest summed $p_T^2$ of associated tracks. The transverse impact parameter significance must satisfy $|d_0|/\sigma_{d_0} < 5$, and the longitudinal impact parameter must satisfy $|\Delta z_0 \sin \theta| < 0.5$ mm. Electrons are further required to be isolated based on ID tracks and topological clusters in the calorimeter [51], with an isolation efficiency of 90(99) % for $p_T = 25(60)$ GeV.

Muon candidates are identified by matching MS segments with ID tracks [52]. The candidates must satisfy requirements on hits in the MS and on the compatibility between ID and MS momentum measurements to remove fake muon signatures. Furthermore, they must have $p_T > 20$ GeV as well as $|\eta| < 2.5$ to ensure they are within coverage of the ID. Candidate muons must satisfy the following requirements on the distance of the combined ID and MS track to the primary vertex: the transverse impact parameter significance must satisfy $|d_0|/\sigma_{d_0} < 3$, and the longitudinal impact parameter must satisfy $|\Delta z_0 \sin \theta| < 0.5$ mm. An isolation requirement is imposed based on ID tracks and topological clusters in the calorimeter, and results in an isolation efficiency of 90(99) % for $p_T = 25(60)$ GeV.

Single-lepton triggers used in this analysis are designed to select events containing a high-$p_T$, well-identified charged lepton [53]. They require a $p_T$ of at least 20 GeV for muons and 24 GeV for electrons, and also have requirements on the lepton quality and isolation. These are complemented by triggers with higher $p_T$ thresholds and relaxed isolation and identification requirements to ensure maximum efficiency at higher lepton $p_T$.

Jets are reconstructed from topological clusters in the calorimeter [54] using the anti-$k_t$ algorithm [55, 56] with a radius parameter of 0.4. They are energy-corrected to account for pile-up and calibrated using a $p_T$- and $\eta$-dependent correction derived from simulation [57]. They are required to have $p_T > 25$ GeV and $|\eta| < 2.5$. To suppress pile-up, a discriminant called the jet-vertex-tagger (JVT) is constructed using a two-dimensional likelihood method [58]. 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 pileup and noise is imposed.

Jets containing $b$-hadrons ($b$-jets) are tagged using a multivariate discriminant which exploits the long lifetime and large invariant mass of $b$-hadron decay products relative to $c$-hadrons and unstable light hadrons [59]. The discriminant is calibrated to achieve a 77% $b$-tagging efficiency and rejection factor of about 4.5 against jets containing charm quarks.
(c-jet) and 140 against light-quark and gluon jets in a sample of simulated $t\bar{t}$ events [60]. The $b$-tagging efficiency in simulation is corrected to the efficiency in data [61].

The missing transverse momentum vector is calculated as the negative vectorial sum of the transverse momenta of particles in the event. Its magnitude $E^{\text{miss}}_T$ is a measure of the transverse momentum imbalance, primarily due to neutrinos that escape detection. Energy deposits in the calorimeters are uniquely assigned in order of priority to electrons, jets, and muons found in the event, thus avoiding double counting of signals. This approach also obviates the need for further overlap removal in the $E^{\text{miss}}_T$ calculation, since a single energy deposit cannot be re-assigned to two nearby reconstructed signals. In addition to the identified electrons, jets and muons, a track-based soft term is included in the $E^{\text{miss}}_T$ calculation by considering tracks associated with the hard-scattering vertex in the event but not with an identified electron, jet, or muon [62, 63].

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. Bremsstrahlung radiation by a muon can result in ID tracks and a calorimeter energy deposit that are also reconstructed as an electron candidate. Therefore in cases where an electron and muon candidate share an ID track, the object is considered to be a muon, and the electron candidate is rejected.

The overlap of objects is measured using the Lorentz-invariant distance $\Delta R_{y,\phi} = \sqrt{(\Delta y)^2 + (\Delta \phi)^2}$. Due to the isolation requirements placed on electron candidates, any jets that closely overlap an electron candidate within a cone $\Delta R_{y,\phi} < 0.2$ are likely to be reconstructions of the electron and so are rejected. When jets and electrons are found within the larger hollow cone $0.2 < \Delta R_{y,\phi} < 0.4$, it is more likely that a real hadronic jet is present and that the electron is a non-prompt constituent of the jet arising from the decay of heavy-flavour hadrons. Hence and electron candidates found within a cone $\Delta R_{y,\phi} < 0.4$ of any remaining jet is rejected.

Muons can be accompanied by a hard photon due to bremsstrahlung or collinear final state radiation, and the muon-photon system can then be reconstructed as both a jet and muon candidate. Non-prompt muons can arise from the decay of hadronic jets, however these muons are associated with a higher ID track multiplicity than those accompanied by hard photons. In order to resolve these ambiguities between nearby jet and muon candidates, first any jets having fewer than three ID tracks and within a cone $\Delta R_{y,\phi} < 0.4$ of any muon candidate are rejected, then any muon candidates within a cone $\Delta R_{y,\phi} < 0.4$ of any remaining jet is rejected.

5 Event selection and background estimation

Events are required to have at least one well-reconstructed interaction vertex, to pass a single-electron or single-muon trigger, and to contain at least one jet with $p_T > 25$ GeV. Events are required to contain exactly two charged leptons of opposite charge with $p_T > 20$ GeV; events with a third lepton with $p_T > 20$ GeV are rejected. At least one lepton must have $p_T > 25$ GeV, and at least one of the selected electrons (muons) must be matched
within a $\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2}$ cone of size 0.07 (0.1) to the electron (muon) selected online by the corresponding trigger.

In simulated events, information recorded by the event generator is used to identify events in which any selected lepton does not originate promptly from the hard-scatter process. These non-prompt or fake leptons arise from processes such as the decay of a $b$-hadron, photon conversion or hadron misidentification, and are identified when the electron or muon does not originate from the decay of a $W$ or $Z$ boson (or a $\tau$ lepton itself originating from a $W$ or $Z$). Events with a selected lepton which is non-prompt or fake are themselves labelled as fake and are treated as a contribution to the background.

After this selection has been made, a further set of requirements is imposed with the aim of reducing the contribution from the $Z + \text{jets}$, diboson and fake/non-prompt lepton backgrounds. The resultant sample is intended to consist almost entirely of $Wt$ signal and $t\bar{t}$ background (a breakdown of the expected signal contributions and background compositions in all regions can be seen in figure 3), which are subsequently separated by the BDT analysis. Events in which the two leptons have the same flavour and an invariant mass consistent with a $Z$ boson ($81 < m_{ll} < 101$ GeV) are vetoed, as well as those with an invariant mass $m_{ll} < 40$ GeV. Further requirements on $E_T^{\text{miss}}$ and $m_{ll}$ are chosen based on the flavour of the selected leptons (as shown in table 1). Events with different-flavour leptons are required to have $E_T^{\text{miss}} > 20$ GeV, with the requirement raised to $E_T^{\text{miss}} > 50$ GeV when the dilepton invariant mass satisfies $m_{ll} < 80$ GeV. All events with same-flavour leptons must satisfy $E_T^{\text{miss}} > 40$ GeV. For same-flavour events, the $Z + \text{jets}$ background is concentrated in a region of the $m_{ll} - E_T^{\text{miss}}$ plane corresponding to values of $m_{ll}$ near the $Z$ mass, and towards low values of $E_T^{\text{miss}}$. Therefore, a selection in $E_T^{\text{miss}}$ and $m_{ll}$ is used to remove these backgrounds: events with $40$ GeV $< m_{ll} < 81$ GeV are required to satisfy $4 \times E_T^{\text{miss}} > 5 \times m_{ll}$ while events with $m_{ll} > 101$ GeV are required to satisfy $2 \times m_{ll} + E_T^{\text{miss}} > 300$ GeV. The requirements for the same- and different-flavour events are chosen separately due to the kinematically different processes contributing to the $Z + \text{jets}$ background, namely $Z \rightarrow ee/\mu\mu$ in same-flavour events and $Z \rightarrow \tau\tau$ in different-flavour events. These requirements reduce the $Z + \text{jets}$ contributions in the signal regions to 12% according to simulation. The partitioning of events into different selections based on lepton flavour, $E_T^{\text{miss}}$, and $m_{ll}$ is described well by the simulation, motivating the choice to merge these selection regions into the signal regions described below.

The sample of selected events is divided into regions based on the number of jets and $b$-tagged jets. At LO, the signal process results in a final state with one $b$-jet arising from the top quark decay, and no additional jets, while the $t\bar{t}$ process results in two $b$-jets from the top quark decays. Events with additional jets are also studied since the underlying event, higher order QCD and other effects may produce additional jets in signal events.

Corresponding to these expected final states, two signal regions are defined by the presence of exactly one $b$-tagged jet and either zero (denoted $1j1b$) or one (denoted $2j1b$) additional jet. A $t\bar{t}$-enriched control region is defined by the presence of exactly two jets, which are both $b$-tagged (denoted $2j2b$). This control region is used to constrain the $t\bar{t}$ background normalisation, and is expected to contain only a small ($< 1\%$) proportion of signal events. These three regions — $1j1b$, $2j1b$ and $2j2b$ — are called the fit regions,
At least one jet with $p_T > 25$ GeV, $|\eta| < 2.5$

Exactly two leptons of opposite charge with $p_T > 20$ GeV, $|\eta| < 2.5$ for muons and $|\eta| < 2.47$ excluding $1.37 < |\eta| < 1.52$ for electrons

At least one lepton with $p_T > 25$ GeV, veto if third lepton with $p_T > 20$ GeV

At least one lepton matched to the trigger object

<table>
<thead>
<tr>
<th>Different flavour</th>
<th>$E_T^{\text{miss}} &gt; 50$ GeV, if $m_{\ell\ell} &lt; 80$ GeV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$E_T^{\text{miss}} &gt; 20$ GeV, if $m_{\ell\ell} &gt; 80$ GeV</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Same flavour</th>
<th>$E_T^{\text{miss}} &gt; 40$ GeV, always</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>veto, if $m_{\ell\ell} &lt; 40$ GeV</td>
</tr>
<tr>
<td></td>
<td>$4E_T^{\text{miss}} &gt; 5m_{\ell\ell}$, if $40$ GeV $&lt; m_{\ell\ell} &lt; 81$ GeV</td>
</tr>
<tr>
<td></td>
<td>veto, if $81$ GeV $&lt; m_{\ell\ell} &lt; 101$ GeV</td>
</tr>
<tr>
<td></td>
<td>$2m_{\ell\ell} + E_T^{\text{miss}} &gt; 300$ GeV, if $m_{\ell\ell} &gt; 101$ GeV</td>
</tr>
</tbody>
</table>

Table 1. Summary of event selection criteria used in the analysis.

as they are used in the simultaneous fit described in section 8. The total efficiency in simulation to accept a dilepton $Wt$ signal event into one of the signal or control regions is about 12%, while the efficiency to accept a dilepton $t\bar{t}$ background event to the same regions is about 5% estimated in simulation. Event yields for each fit region are presented in section 9. Two additional regions, in which events are required to contain one (denoted 1j0b) or two (denoted 2j0b) jets but no $b$-tagged jets are used to validate the description of the data by the simulation. A schematic view of the regions definition is shown in figure 4.

### 6 Separation of signal from background

After the event selection is performed, the data sample consists primarily of $t\bar{t}$ events with a significant number of $Wt$ signal events (see for example figure 3). As there is no single observable that clearly discriminates between the $Wt$ signal and the $t\bar{t}$ background, several observables are combined into a single discriminator using a BDT technique [64]. A collection of decision trees is created that weakly separates events into signal and background based on a number of binary decisions considering a single observable at a time. A boosting algorithm is then used to assign weights to each tree such that the ensemble of weak classifiers performs as a strong classifier [65]. In this analysis, the BDT implementation is provided by the TMVA package [66], using the GradientBoost algorithm.

Separate BDTs are prepared for the analysis regions 1j1b and 2j1b. Due to the low efficiency to accept a $Wt$ event in the 2j2b region, the computing cost to simulate events in this region is especially large. Since the expected gain in signal precision from subdividing the 2j2b region is minimal, no BDT is constructed here and a single bin is used. The BDTs are optimised to distinguish between $Wt$ and $t\bar{t}$ by using the nominal
Figure 3. Expected event yields for signal and backgrounds with their total systematic uncertainty (discussed in section 8) and the number of observed events in the data are shown in the three fit regions (1j1b, 2j1b, and 2j2b) and the two additional regions (1j0b and 2j0b). The signal and backgrounds are normalised to their theoretical predictions, and the error bands represent the total systematic uncertainties which are used in this analysis. The upper panel gives the yields in number of events per bin, while the lower panel gives the ratios of the numbers of observed events to the total prediction in each bin.

Figure 4. A schematic view of signal, control and validation regions. Signal and control regions are used in fits.

Wt MC sample, the alternative Wt MC sample with diagram subtraction scheme and the nominal tt MC sample; for each sample, half of the events are used for training while the other half is reserved for testing. For each region, a large list of variables is prepared for the BDT. An optimisation procedure is then carried out in each region to select a subset of input variables and a set of BDT parameters (such as the number of trees in the ensemble and the maximum depth of the individual decision trees). The optimisation is designed to provide the best separation between the Wt signal and tt background while avoiding sensitivity to statistical fluctuations in the training sample.
The variables considered are derived from the kinematic properties of subsets of the selected physics objects defined in section 4 for each event. For a set of objects \( o_1 \ldots o_n \): 

\[ p_T^{\text{sys}}(o_1 \ldots o_n) \] 

is the transverse momenta of various subsets; 

\[ H_T(o_1 \ldots o_n) \] 

is the scalar sum of transverse momenta; 

\[ E_T^{\text{miss}}(o_1 \ldots o_n) \] 

is the scalar sum of transverse momenta of all objects which contribute to the \( E_T^{\text{miss}} \) calculation; 

\[ \sigma(p_T^{\text{sys}}) \] 

is the ratio of \( p_T^{\text{sys}} \) to \( (H_T + \sum E_T) \); 

\[ m(o_1 \ldots o_n) \] 

is the invariant mass of various subsets; 

\[ m_T(o_1 \ldots o_n) \] 

is the transverse mass (i.e. the sum of the invariant masses of \( o_1 \ldots o_n \) each projected onto the transverse plane); and 

\[ E/m(o_1 \ldots o_n) \] 

is the ratio of energy to invariant mass. Two-dimensional vectors such as \( E_T^{\text{miss}} \) are assigned four-momenta by assuming zero mass and \( z \)-component. For two systems of objects \( s_1 \) and \( s_2 \): 

\[ R(s_1; s_2) \] 

is the separation in \( \phi - \eta \) space; 

\[ \Delta p_T(s_1, s_2) \] 

is the \( p_T \) difference; 

\[ \Delta \phi(s_1, s_2) \] 

is the \( \phi \) difference; and 

\[ C(s_1, s_2) \] 

is the centrality, is the ratio of the scalar sum of \( p_T \) to the sum of energy.

The final set of input variables used in each BDT is listed in table 2 along with the separating power of each variable. In order to check that the variables and their correlations in \( Wt \) signal and the background events are well modelled by simulation, the distributions of these variables and the BDTs are compared between the MC prediction and the observed data, using a Kolmogorov-Smirnov (KS) statistical test [67] to check agreement. The distributions of the two most powerful variables in each fit region are shown in figure 5. The MC predictions describe the data well, within the total systematic uncertainties.

7 Systematic uncertainties

Systematic uncertainties are divided into experimental and theoretical sources. Each uncertainty is assigned a Gaussian-constrained nuisance parameter, which allows the uncertainty to be constrained by data.

The experimental sources of uncertainty include the measurement of the luminosity, lepton efficiency scale factors used to correct simulation to data, lepton energy scale and resolution, \( E_T^{\text{miss}} \) soft-term calculation, jet energy scale and resolution, and the \( b \)-tagging efficiency. Among these, the dominant sources of uncertainty are due to the determination of the jet energy scale (JES) and jet energy resolution. Table 3 gives a breakdown of uncertainties in the final fitted cross-section.

The JES uncertainty [57] is divided into a total of 18 components, which are derived using \( \sqrt{s} = 13 \) TeV data. The uncertainties from \textit{in situ} analyses including studies of \( Z/\gamma + \text{jet} \) and dijet events are represented with six orthogonal components (JES Eff1–6). The full description of jet uncertainties and correlations is reduced to obtain this set of uncertainty components that can be used as nuisance parameters in a likelihood fit. This is done by diagonalising the covariance matrix describing the jet uncertainties to obtain a set of reduced

\[ \langle S^2 \rangle = \frac{1}{2} \int \frac{(Y_\ell(y) - Y_b(y))^2}{(Y_\ell(y) + Y_b(y))^2} dy \]

where \( Y_\ell(y) \) and \( Y_b(y) \) are the signal and background probability distribution functions of each variable \( y \), respectively.
Table 2. The variables used in each BDT and their separating powers (a measure of the difference between probability distributions of signal and background in the variable, denoted $S$). The variables are derived from the four-momenta of the leading (sub-leading) lepton $\ell_1$ ($\ell_2$), the leading (sub-leading) jet $j_1$ ($j_2$) and $E_T^{\text{miss}}$. The last row gives the separation power of the BDT discriminant output.

<table>
<thead>
<tr>
<th>Variable</th>
<th>$S$ [$10^{-2}$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_T^{\text{sys}}(\ell_1/\ell_2, E_T^{\text{miss}} j_1)$</td>
<td>5.3</td>
</tr>
<tr>
<td>$\Delta p_T(\ell_1/\ell_2, E_T^{\text{miss}} j_1)$</td>
<td>2.9</td>
</tr>
<tr>
<td>$\sum E_T$</td>
<td>2.7</td>
</tr>
<tr>
<td>$\Delta p_T(\ell_1/\ell_2, E_T^{\text{miss}})$</td>
<td>1.2</td>
</tr>
<tr>
<td>$p_T^{\text{sys}}(\ell_1 E_T^{\text{miss}} j_1)$</td>
<td>0.9</td>
</tr>
<tr>
<td>$C(\ell_1/\ell_2)$</td>
<td>0.9</td>
</tr>
<tr>
<td>$\Delta p_T(\ell_1, E_T^{\text{miss}})$</td>
<td>0.8</td>
</tr>
<tr>
<td>BDT discriminant</td>
<td>8.6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>$S$ [$10^{-2}$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_T(\ell_1/\ell_2)$</td>
<td>1.7</td>
</tr>
<tr>
<td>$\Delta R(\ell_1/\ell_2, E_T^{\text{miss}} j_1/j_2)$</td>
<td>1.7</td>
</tr>
<tr>
<td>$\Delta R(\ell_1/\ell_2, j_1/j_2)$</td>
<td>1.5</td>
</tr>
<tr>
<td>$m(\ell_1/j_2)$</td>
<td>1.4</td>
</tr>
<tr>
<td>$\Delta p_T(\ell_1/\ell_2, E_T^{\text{miss}})$</td>
<td>1.4</td>
</tr>
<tr>
<td>$\Delta p_T(\ell_1, j_1)$</td>
<td>1.4</td>
</tr>
<tr>
<td>$m(\ell_1/j_1)$</td>
<td>1.3</td>
</tr>
<tr>
<td>$p_T(\ell_1)$</td>
<td>1.3</td>
</tr>
<tr>
<td>$\sigma(p_T^{\text{sys}})(\ell_1/\ell_2 E_T^{\text{miss}} j_1)$</td>
<td>1.2</td>
</tr>
<tr>
<td>$\Delta R(\ell_1/j_1)$</td>
<td>1.2</td>
</tr>
<tr>
<td>$p_T(j_2)$</td>
<td>0.9</td>
</tr>
<tr>
<td>$\sigma(p_T^{\text{sys}})(\ell_1/\ell_2 E_T^{\text{miss}} j_1/j_2)$</td>
<td>0.9</td>
</tr>
<tr>
<td>$m(\ell_1/j_1 j_2)$</td>
<td>0.3</td>
</tr>
<tr>
<td>$m(\ell_2/j_1)$</td>
<td>0.3</td>
</tr>
<tr>
<td>$m(\ell_2/j_2)$</td>
<td>0.1</td>
</tr>
<tr>
<td>BDT discriminant</td>
<td>10.9</td>
</tr>
</tbody>
</table>

uncertainties corresponding to the eigenvector-eigenvalue pairs as demonstrated in ref. [68]. Other components are model uncertainties (such as flavour composition, $\eta$ intercalibration model), and other systematics in the JES determination (such as pile-up jet area $\rho$). The most significant JES uncertainty components for this analysis are the in situ calibration and the flavour composition uncertainty, which is the dependence of the jet calibration on the fraction of quark or gluon jets in data. The jet energy resolution uncertainty estimate [57] is based on comparisons of simulation and data using in situ studies with Run-1 data. These studies are then cross-calibrated and checked to confirm good agreement with Run-2 data.

As discussed in section 4, the $E_T^{\text{miss}}$ calculation includes contributions from hard sources, including leptons and jets, in addition to soft terms which arise primarily from low-$p_T$ pile-up jets and underlying-event activity. The uncertainty associated with the hard terms is propagated from the corresponding uncertainties in the energy/momentum scales and resolutions for jets and leptons, and is classified together with the uncertainty asso-
Figure 5. Distributions of the two most powerful BDT input variables in each fit region: in the $1\ell 1b$ region (a) $p_T^{\text{sys}}(\ell_1, \ell_2, E_T^{\text{miss}} j_1)$ and (b) $\Delta p_T(\ell_1, \ell_2, E_T^{\text{miss}} j_1)$; in the $2\ell 1b$ region (c) $p_T^{\text{sys}}(\ell_1, \ell_2)$ and (d) $\Delta R(\ell_1, \ell_2, E_T^{\text{miss}} j_1, j_2)$. The signal and backgrounds are normalised to their theoretical predictions, and the error bands represent the total systematic uncertainties in the Monte Carlo predictions. The first and last bins of each distribution contain overflow events. The upper panels give the yields in number of events per bin, while the lower panels give the ratios of the numbers of observed events to the total prediction in each bin.

associated with the hard objects. The uncertainty associated with the soft term is estimated by comparing the simulated scale and resolution to that in data, including differences in uncertainties due to model dependence.

Uncertainties in the scale factors to correct the $b$-tagging efficiency in simulation to the efficiency in data are assessed using independent eigenvectors for the efficiency of $b$-jets, $c$-jets, light-parton jets, and two extrapolation uncertainty factors. These $b$-tagging uncertainties are determined with $\sqrt{s} = 13$ TeV data for $b$-jets, while for $c$-jets and light-parton jets they are determined in $\sqrt{s} = 8$ TeV data, then extrapolated to and checked with
\( \sqrt{s} = 13 \text{ TeV} \) data. The extrapolation is performed by adding additional uncertainties to cover changes made in the inner detector and tracking algorithms between Run-1 and Run-2 data, accounting for fake tracks, tracking efficiency and tracking resolution. These \( c \)-jet and light-parton jet scale factors were later checked against similar scale factors derived on \( \sqrt{s} = 13 \text{ TeV} \) data, and the scale factors with added uncertainties were found to agree well with the full run-2 scale factors.

Systematic uncertainties in lepton momentum resolution and scale, trigger efficiency, isolation efficiency, and identification efficiency are also considered. These uncertainties arise from corrections to simulation based on studies of \( Z \rightarrow ee \) and \( Z \rightarrow \mu\mu \) data. In this analysis the effects of the uncertainties in these corrections are relatively small.

A 2.1\% uncertainty is assigned to the integrated luminosity determination for 2015 data. It is derived, following a methodology similar to that detailed in ref. [69], from a calibration of the luminosity scale using \( x-y \) beam-separation scans performed in August 2015.

Uncertainties stemming from theoretical models are estimated by comparing a set of predicted distributions produced with different assumptions and applying the difference observed as a weight to the nominal \( Wt \) or \( \ttbar \) distribution. The main uncertainties are due to the NLO matrix element (ME) generator, parton shower and hadronisation generator, initial- and final-state radiation (I/FSR) tuning and the PDF. The NLO matrix element uncertainty is estimated by comparing two NLO matching methods: the predictions of POWHEG-BOX and MADGRAPH5_aMC@NLO, both interfaced with Herwig++. The parton shower, hadronisation, and underlying-event model uncertainty is estimated by comparing POWHEG-BOX interfaced with either PYTHIA 6 or HERWIG++. The uncertainty from the matrix element generator is treated as uncorrelated between the \( Wt \) and \( \ttbar \) processes, while the uncertainty from the parton shower generator is treated as correlated. The I/FSR tuning uncertainty is estimated by taking half of the difference between samples with POWHEG-BOX interfaced with PYTHIA 6 tuned with either more or less radiation, and is uncorrelated between the \( Wt \) and \( \ttbar \) processes. The choice of scheme to account for the interference between the \( Wt \) and \( \ttbar \) processes constitutes another source of systematic uncertainty for the signal modelling, and it is estimated by comparing samples using either the diagram removal scheme or the diagram subtraction scheme, both generated with POWHEG-BOX+PYTHIA 6. The uncertainty due to the choice of PDF is estimated using the PDF4LHC15 combined PDF set [70]. The difference between the central CT10 [28] prediction and the central PDF4LHC15 prediction (PDF central value) is taken and symmetrised together with the internal uncertainty set provided with PDF4LHC15. For \( \ttbar \) and \( Wt \) modelling, the NLO matrix element model, parton shower model, and PDF uncertainties are estimated using fast-simulated samples; for \( Wt \), fast simulation is also used for I/FSR. In each case where results from two samples must be compared, fast simulated samples are only compared to other fast simulated samples.

Additionally, normalisation uncertainties of 100\% are assumed for the fake/non-prompt lepton backgrounds. The \( Z + \text{jets} \) backgrounds with one \( b \)-tagged jet are assigned a 50\% uncertainty, while a 100\% uncertainty is assumed for \( Z + \text{jets} \) events with two \( b \)-tagged jets. These uncertainties are chosen to be consistent with previous ATLAS studies of these processes in dedicated validation regions. Diboson backgrounds are assigned an uncertainty
of 25% to cover the difference between the predictions of the Sherpa and Powheg-Box generators. These uncertainties are treated as uncorrelated across the various regions of jet and $b$-tagged jet multiplicity.

8 Extraction of signal cross-section

The $Wt$ cross-section is extracted from the data using a profile-likelihood fit that combines inputs from each signal and control region to constrain backgrounds and systematic uncertainties. The fit uses the HistFitter [71] software framework, which is in turn built on the HistFactory, RooStats, and RooFit [72] frameworks.

The fit uses the binned BDT response for MC events in two of the three fit regions ($1j1b$ and $2j1b$) and a single bin in the $2j2b$ region to construct templates for the $Wt$ signal and each modelled background ($tt$, $Z + \text{jets}$, diboson, fake or non-prompt leptons). For each signal and background template, an additional template is constructed for each of the MC sample variations (see section 7) accounting for a systematic uncertainty. Systematic uncertainties are considered by allowing Gaussian-constrained nuisance parameters to deform fit templates while simultaneously varying the normalisation of the templates. The normalisation of the $tt$ background, $\mu_{tt}$, is also determined in the fit by assigning an unconstrained parameter to the $tt$ normalisation. Other backgrounds are constrained within their systematic uncertainties by Gaussian-constrained nuisance parameters, and all templates are affected by the overall luminosity uncertainty.

A global likelihood function is constructed to describe the level of agreement between data and prediction as a function of the parameter of interest, namely the $Wt$ signal strength $\sigma_{Wt}$, and a list of nuisance parameters each describing the influence of a different source of systematic uncertainty. The $Wt$ cross-section and its uncertainty are extracted from the fitted value of $\sigma_{Wt}$, with a value of unity corresponding to the predicted NLO+NNLL theory value.

9 Results

The expected and fitted yields from data are measured in the three fit regions. The majority of signal events fall in the $1j1b$ and $2j1b$ regions, with the former giving the better signal to background ratio as well as the larger yield of signal events. Table 4 shows the fitted yields of each process. From the fitted $Wt$ yield, a cross-section is then extracted. The result is a measured cross-section of $\sigma_{Wt} = 94 \pm 10 \text{(stat.)}^{+28}_{-22} \text{(syst.)} \pm 2 \text{(lumi.)} \text{pb}$, corresponding to an observed (expected) significance of 4.5$\sigma$ (3.9$\sigma$). Most pairs of parameters in the fit show small correlations, generally at the 25% level or less. The most correlated pairs of nuisance parameters are the modelling uncertainties due to matrix element and parton shower (57%), and the parameters related to JES flavour composition and parton shower uncertainties (45%).

Figure 6 shows the fit parameters ($\theta$) with the highest post-fit impact on the signal strength, and also gives the pre-fit impacts as well as post-fit parameter values. The post-fit parameters, $\theta$, are shifted and re-scaled by $(\hat{\theta} - \theta_0)/\Delta\theta$, where $\theta_0$ and $\hat{\theta}$ are the pre- and
<table>
<thead>
<tr>
<th>Source</th>
<th>( \Delta \sigma_{Wt}/\sigma_{Wt} [%] )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jet energy scale</td>
<td>21</td>
</tr>
<tr>
<td>Jet energy resolution</td>
<td>8.6</td>
</tr>
<tr>
<td>( E_T^{\text{miss}} ) soft terms</td>
<td>5.3</td>
</tr>
<tr>
<td>( b )-tagging</td>
<td>4.3</td>
</tr>
<tr>
<td>Luminosity</td>
<td>2.3</td>
</tr>
<tr>
<td>Lepton efficiency, energy scale and resolution</td>
<td>1.3</td>
</tr>
<tr>
<td>NLO matrix element generator</td>
<td>18</td>
</tr>
<tr>
<td>Parton shower and hadronisation</td>
<td>7.1</td>
</tr>
<tr>
<td>Initial-/final-state radiation</td>
<td>6.4</td>
</tr>
<tr>
<td>Diagram removal/subtraction</td>
<td>5.3</td>
</tr>
<tr>
<td>Parton distribution function</td>
<td>2.7</td>
</tr>
<tr>
<td>Non-( t\bar{t} ) background normalisation</td>
<td>3.7</td>
</tr>
<tr>
<td>Total systematic uncertainty</td>
<td>30</td>
</tr>
<tr>
<td>Data statistics</td>
<td>10</td>
</tr>
<tr>
<td>Total uncertainty</td>
<td>31</td>
</tr>
</tbody>
</table>

**Table 3.** Relative uncertainties in the \( Wt \) cross-section. These are estimated by fixing each uncertainty parameter to its post-fit \( \pm 1\sigma \) uncertainties, re-fitting, and assessing the change in the signal strength. Due to correlations between parameters, the individual uncertainty categories are not expected to add up to the total systematic uncertainty. The statistical uncertainty is evaluated by fitting without any nuisance parameters corresponding to systematic uncertainties in the fit, and the total systematic uncertainty is evaluated by subtracting the statistical uncertainty from the total uncertainty in quadrature.

post-fit values of \( \theta \), while \( \Delta \theta \) is the pre-fit uncertainty on \( \theta \). Here the impact \( (\Delta \mu) \) of a parameter is defined as the change in signal strength observed when fixing this parameter to its \( \pm 1\sigma \) values, fixing all other parameters to their nominal values, and fitting the signal strength. The change is taken with respect to the nominal pre-fit value for pre-fit impact and with respect to the nominal post-fit value for the post-fit impact. The pre-fit and post-fit impact are differentiated based on whether pre-fit or post-fit values of \( \pm 1\sigma \) variations are assumed for the parameter under consideration. The parameters with the highest post-fit impact are jet energy scale uncertainties and the modelling uncertainties due to parton shower and \( t\bar{t} \) initial- and final-state radiation. Some parameters fit to values which are significantly different from unity; \( t\bar{t} \) initial- and final-state radiation and the component JES \( \text{Eff1} \) each exhibit this effect. This behaviour is expected as a few parameters could be pulled outside of the \( \pm 1\sigma \) band when there are a large number of parameters being fitted, while the majority should fall within the \( \pm 1\sigma \) range. Certain parameters are assigned post-fit uncertainties significantly smaller than the nominal pre-fit uncertainty values and are thus profiled or constrained by the observed data. For example, the uncertainty due
Table 4. Fit results for an integrated luminosity of 3.2 fb$^{-1}$. The errors shown are the final fitted uncertainties in the yields, including uncertainties in the fitted signal strength, systematic uncertainties, and statistical uncertainties, taking into account correlations and constraints induced by the fit.

to parton shower generator would be among the most dominant uncertainties without the constraints from profiling, with pre-fit impacts exceeding 60% of the signal strength. However, information from the 2j2b region about the $t\bar{t}$ normalisation and the relative yields in the signal regions significantly constrain these uncertainties. Another feature observed in this plot is how the sum in quadrature of the individual impacts is substantially smaller than the final uncertainty shown in table 3. This is due to the correlations between the uncertainties, and in particular to the constraint provided by the $t\bar{t}$ normalisation. Some of the larger uncertainties such as the parton shower generator uncertainty have an asymmetric impact on the signal strength. These asymmetries have been traced to originate from the large normalisation uncertainty on the $t\bar{t}$ background.

The MC predictions and data yields for the BDT response after setting all fit parameters to their final best-fit values are shown in figure 7, with error bands representing the total uncertainties in the fitted results. The NLO+NNLL cross-section prediction agrees well with the measured value, and $\mu_{t\bar{t}}$ is fitted to $0.98 \pm 0.05$. 

<table>
<thead>
<tr>
<th></th>
<th>1j1b</th>
<th>2j1b</th>
<th>2j2b</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed events</td>
<td>4254</td>
<td>6138</td>
<td>4912</td>
</tr>
<tr>
<td>Fitted events</td>
<td>4257</td>
<td>6139</td>
<td>4908</td>
</tr>
<tr>
<td>Fitted $Wt$ events</td>
<td>910 $\pm$ 210</td>
<td>640 $\pm$ 160</td>
<td>210 $\pm$ 82</td>
</tr>
<tr>
<td>Fitted $t\bar{t}$ events</td>
<td>3230 $\pm$ 210</td>
<td>5340 $\pm$ 160</td>
<td>4670 $\pm$ 110</td>
</tr>
<tr>
<td>Fitted $Z + \text{jets}$ events</td>
<td>69 $\pm$ 35</td>
<td>87 $\pm$ 46</td>
<td>7.6 $\pm$ 7.5</td>
</tr>
<tr>
<td>Fitted fake events</td>
<td>30 $\pm$ 26</td>
<td>40 $\pm$ 38</td>
<td>15 $\pm$ 14</td>
</tr>
<tr>
<td>Fitted diboson events</td>
<td>23.5 $\pm$ 6.0</td>
<td>24.8 $\pm$ 6.2</td>
<td>0.91 $\pm$ 0.23</td>
</tr>
</tbody>
</table>
Figure 6. List of fit parameters ranked by post-fit impact on the signal strength. The fit parameters ($\theta$) here correspond to the nuisance parameters from section 8. Impact ($\Delta \mu$) is calculated by fixing the parameter to its $\pm 1\sigma$ values, fixing all other parameters to their nominal values, re-fitting the signal strength, and evaluating the change in signal strength with respect to the nominal fit. Green bands indicate the impacts computed with $\sigma$ corresponding to the pre-fit uncertainty, and hatched purple bands indicate the impacts computed with $\sigma$ corresponding to the post-fit uncertainty. The black points represent $(\bar{\theta} - \theta_0)/\Delta \theta$, the shifted and scaled post-fit parameter values, while the error bars are the post-fit errors of the fit parameter. The meanings of the labels and abbreviations are detailed in section 7.
10 Conclusion

The inclusive cross-section for the associated production of a $W$ boson and top quark is measured using 3.2 fb$^{-1}$ of $pp$ collision data collected at $\sqrt{s} = 13$ TeV by the ATLAS detector at the LHC. The analysis uses dilepton events with at least one $b$-tagged jet. Events are separated into signal and control regions based on the number of jets and $b$-tagged jets, and the $Wt$ signal is separated from the $t\bar{t}$ background using a BDT discriminant. The cross-section is extracted by fitting templates to the BDT output distribution, and is measured to be $\sigma_{Wt} = 94 \pm 10 \text{(stat.)} \pm 28 \text{(syst.)} \pm 2 \text{(lumi.)}$ pb. The measured value is in good agreement with the SM prediction of $\sigma_{\text{theory}} = 71.7 \pm 1.8 \text{(scale)} \pm 3.4 \text{(PDF)}$ pb [1].

Acknowledgments

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; 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; COLCIENCIAS, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DNRF and DNSRC, Denmark; IN2P3-CNRS, CEA-DSM/IRFU, France; SRNSF, 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 (U.K.) and BNL (U.S.A.), the Tier-2 facilities worldwide and large non-WLCG resource providers. Major contributors of computing resources are listed in ref. [73].

Open Access. This article is distributed under the terms of the Creative Commons Attribution License (CC-BY 4.0), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

References


The ATLAS collaboration

aa Also at Manhattan College, New York NY, United States of America
ab Also at Academia Sinica Grid Computing, Institute of Physics, Academia Sinica, Taipei, Taiwan
ac Also at School of Physics, Shandong University, Shandong, China
ad Also at Department of Physics, California State University, Sacramento CA, United States of America
ae Also at Moscow Institute of Physics and Technology State University, Dolgoprudny, Russia
af Also at Departement de Physique Nucleaire et Corpusculaire, Université de Genève, Geneva, Switzerland
ag Also at Eotvos Lorand University, Budapest, Hungary
ah Also at Departments of Physics & Astronomy and Chemistry, Stony Brook University, Stony Brook NY, United States of America
ai Also at International School for Advanced Studies (SISSA), Trieste, Italy
aj Also at Department of Physics and Astronomy, University of South Carolina, Columbia SC, United States of America
ak Also at Institut de Física d’Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Barcelona, Spain
al Also at School of Physics and Engineering, Sun Yat-sen University, Guangzhou, China
am Also at Institute for Nuclear Research and Nuclear Energy (INRNE) of the Bulgarian Academy of Sciences, Sofia, Bulgaria
an Also at Faculty of Physics, M.V.Lomonosov Moscow State University, Moscow, Russia
ao Also at Institute of Physics, Academia Sinica, Taipei, Taiwan
ap Also at National Research Nuclear University MEPhI, Moscow, Russia
aq Also at Department of Physics, Stanford University, Stanford CA, United States of America
ar Also at Institute for Particle and Nuclear Physics, Wigner Research Centre for Physics, Budapest, Hungary
as Also at Giresun University, Faculty of Engineering, Turkey
at Also at Flensburg University of Applied Sciences, Flensburg, Germany
au Also at University of Malaya, Department of Physics, Kuala Lumpur, Malaysia
av Also at CPPM, Aix-Marseille Université and CNRS/IN2P3, Marseille, France
aw Also at LAL, Univ. Paris-Sud, CNRS/IN2P3, Université Paris-Saclay, Orsay, France

* Deceased