Measurement of single top-quark production in association with a W boson in the single-lepton channel at $\sqrt{s} = 8$ TeV with the ATLAS detector

ATLAS Collaboration

DOI
10.1140/epjc/s10052-021-09371-7

Publication date
2021

Document Version
Final published version

Published in
European Physical Journal C

License
CC BY

Citation for published version (APA):
Measurement of single top-quark production in association with a W boson in the single-lepton channel at \( \sqrt{s} = 8 \) TeV with the ATLAS detector

ATLAS Collaboration

CERN, 1211 Geneva 23, Switzerland

Abstract  The production cross-section of a top quark in association with a W boson is measured using proton–proton collisions at \( \sqrt{s} = 8 \) TeV. The dataset corresponds to an integrated luminosity of 20.2 fb\(^{-1}\), and was collected in 2012 by the ATLAS detector at the Large Hadron Collider at CERN. The analysis is performed in the single-lepton channel. Events are selected by requiring one isolated lepton (electron or muon) and at least three jets. A neural network is trained to separate the \( tW \) signal from the dominant \( t\bar{t} \) background. The cross-section is extracted from a binned profile maximum-likelihood fit to a two-dimensional discriminant built from the neural-network output and the invariant mass of the hadronically decaying W boson. The measured cross-section is \( \sigma_{tW} = 26 \pm 7 \) pb, in good agreement with the Standard Model expectation.

Contents

1 Introduction ........................................ 1
2 ATLAS detector ..................................... 2
3 Data and simulated event samples .................. 2
4 Object definitions ................................... 4
5 Event selection ...................................... 4
6 Separation of signal from background ............ 5
   6.1 Invariant mass of the hadronically decaying W boson .... 5
   6.2 Neural network ................................ 5
   6.3 Two-dimensional discriminant ................. 6
7 Systematic uncertainties ........................... 8
8 Statistical analysis .................................. 9
9 Cross-section measurement ......................... 9
10 Conclusion ....................................... 11
References ........................................... 13

© CERN for the benefit of the ATLAS collaboration 2021

1 Introduction

Single top quarks are produced in proton–proton collisions via the weak interaction. At leading order (LO) three different channels, which depend on the virtuality of the W boson involved, are defined: \( t\)-channel, \( s\)-channel or top-quark production in association with a W boson, called \( tW \) production. These processes, for which example Feynman diagrams are shown in Fig. 1, involve a \( Wtb \) vertex at LO in the Standard Model (SM). Calculations involving \( tW \) production beyond LO have to include quantum interference with \( t\bar{t} \) production. Measurements of single-top-quark cross-sections are used to study the properties of this vertex, as they are directly sensitive to the Cabibbo–Kobayashi–Maskawa (CKM) matrix element \( |V_{tb}| \). Deviations from the cross-sections predicted by the SM can originate from single top quarks produced with similar kinematics in the decays of unknown heavy particles predicted by physics beyond the Standard Model. If the masses of these particles are beyond the reach of direct searches, they might be revealed through their effects on the effective \( Wtb \) coupling [1]. Using measurements in all three channels of single top-quark production, physics beyond the SM can be probed systematically in the context of Effective Field Theory [2]. As each of the single-top-quark processes can be sensitive to different sources of new physics, it is also important to study each channel separately. In addition, the SM production of \( tW \) is an important background in direct searches for particles beyond the SM [3,4].

At the Large Hadron Collider (LHC), evidence for the \( tW \) production process was found by the ATLAS [5] and CMS Collaborations [6] at \( \sqrt{s} = 7 \) TeV and the process was observed by both experiments [7,8] at \( \sqrt{s} = 8 \) TeV. The \( tW \) cross-section has been also measured with 13 TeV collision data inclusively by the CMS Collaboration [9] as well as inclusively and differentially by the ATLAS Collaboration [10–12]. These measurements were performed in final states with two leptons, and the measured cross-sections agree with the theoretical expectations.
This paper presents evidence for \( tW \) production in final states with a single lepton using proton–proton (\( pp \)) collisions at \( \sqrt{s} = 8 \) TeV. This topology features a \( W \) boson in addition to a top quark, which decays mainly into another \( W \) boson and \( b \)-quark, leading to a \( W^+W^-b \) state. In the single-lepton channel, one of the \( W \) bosons decays leptonically (\( W_L \)) while the other one decays hadronically (\( W_H \)). Therefore, the experimental signature of event candidates is characterised by one isolated charged lepton (electron or muon), large missing transverse momentum (\( E_T^{\text{miss}} \)), and three jets with high transverse momentum (\( p_T \)), one of which contains a \( b \)-hadron and is labelled as a \( b \)-tagged jet, \( j_B \). In contrast to the dilepton analyses, the event signature contains only one neutrino, which originates from the leptonic \( W \)-boson decay. Hence, both the \( W \)-boson and the top-quark kinematics can be reconstructed and used to separate the signal from background. The main backgrounds are \( W + \text{jets} \) and \( t\bar{t} \) events, where the latter background poses a major challenge in this measurement because of its similar kinematics and a ten times larger cross-section compared to the \( tW \) signal. An artificial neural network is trained to separate the signal from the \( t\bar{t} \) background. The cross-section is extracted using a binned profile maximum-likelihood fit to a two-dimensional discriminant. This measurement, performed with \( tW \) single-lepton events, constitutes a cross-check of the previous results published in the dilepton channel.

2 ATLAS detector

The ATLAS experiment [13] at the LHC is a multipurpose particle detector with a forward–backward symmetric cylindrical geometry and a near 4\( \pi \) coverage in solid angle.\(^1\) It consists of an inner tracking detector (ID) surrounded by a thin superconducting solenoid providing a 2T axial magnetic field, electromagnetic and hadron calorimeters, and a muon spectrometer (MS). The ID provides charged-particle tracking in the pseudorapidity range \( |\eta| < 2.5 \). It consists of silicon pixel, silicon microstrip, and transition-radiation tracking detectors. Lead/liquid-argon (LAr) sampling calorimeters provide electromagnetic (EM) energy measurements with high granularity. An iron/scintillator-tile hadron calorimeter covers the central pseudorapidity range (\( |\eta| < 1.7 \)). The endcap (\( 1.7 < |\eta| < 3.2 \)) and forward (\( 3.1 < |\eta| < 4.9 \)) regions are instrumented with LAr calorimeters for measurements of both EM and hadronic energy. The MS surrounds the calorimeters and includes a system of precision tracking chambers (\( |\eta| < 2.7 \)) and fast detectors for triggering (\( |\eta| < 2.4 \)). The magnet system for the MS consists of three large air-core toroidal magnets with eight superconducting coils. The field integral of the toroids ranges between 2.0 and 6.0 T m across most of the detector. Collisions producing interesting events are selected for storage with the trigger system [14]. For the data taken at \( \sqrt{s} = 8 \) TeV, a three-level trigger system was used to select events. The first-level trigger is implemented in hardware and uses a subset of the detector information. It reduced the accepted rate to at most 75 kHz. This was followed by two software-based trigger levels that together reduced the accepted event rate to 400 Hz on average, depending on the data-taking conditions.

3 Data and simulated event samples

The data considered in this analysis are from \( pp \) collisions at \( \sqrt{s} = 8 \) TeV and were taken with stable LHC beams and

\(^1\) 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 \( (r, \phi, z) \) 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 \( \eta = -\ln \tan(\theta/2) \). Angular distance is measured in units of \( \Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \).

Footnote 1 continued
the ATLAS detector fully operational, corresponding to an integrated luminosity of 20.2 fb\(^{-1}\).

Monte Carlo (MC) samples were produced using the full ATLAS detector simulation [15] implemented in GEANT 4 [16]. In addition, alternative MC samples, used to train the neural network and evaluate systematic uncertainties, were produced using ATLFAST2 [15], which provides a faster calorimeter simulation making use of parameterised showers to compute the energy deposited by the particles. Pile-up (additional pp interactions in the same or nearby bunch crossing) was modelled by overlaying simulated minimum-bias events generated with PYTHIA 8 [17]. Weights were assigned to the simulated events, such that the distribution of the number of pp interactions per bunch crossing in the simulation matches the corresponding distribution in the data, which has an average of 21 [18].

The \( t\bar{t} W \) signal events were simulated using the next-to-leading order (NLO) POWHEG method [19–21] implemented in the POWHEG BOX (v.1.0) generator (revision 2192) [22] with the CT10 parton distribution function (PDF) set [23] in the matrix-element calculation. The mass and width of the top-quark were set to \( m_t = 172.5 \text{ GeV} \) and \( \Gamma_t = 1.32 \text{ GeV} \), respectively. The top quark was assumed to decay exclusively into \( Wb \). The parton shower, hadronisation and underlying event were simulated using PYTHIA 6 (v6.426) [24] with the LO CTEQ6L1 PDF set [25] and a corresponding set of tuned parameters called the Perugia 2011 (P2011C) tune [26]. The factorisation scale, \( \mu_F \), and renormalisation scale, \( \mu_R \), were set to \( m_t \). Calculations involving \( t\bar{t} W \) production beyond LO included quantum interference with \( t\bar{t} \) production. Double-counting of the contributions was avoided by using either the diagram-removal (DR) or the diagram-subtraction (DS) scheme [27,28]. In the DR scheme, diagrams with a second on-shell top-quark propagator are removed from the amplitude, while in the DS scheme, a subtraction term cancels out the \( t\bar{t} \) contribution to the cross-section when the top-quark propagator becomes on shell. Nominal MC samples were generated using the DR scheme. For the evaluation of systematic uncertainties, alternative samples were generated using the DS scheme, or using POWHEG BOX or MC@NLO (v4.06) [29], each interfaced with HERWIG (v6.520) [30]. For the HERWIG samples, the AUET2 tune [31] with the CT10 PDF was used and the underlying event was generated with JIMMY (v4.31) [32]. In addition, PYTHIA 6 (v6.427) samples with variations of \( \mu_F \) and \( \mu_R \) and the radiation tunes were used. The SM \( t\bar{t} W \) cross-section prediction at NLO including next-to-next-to-leading-log (NNLL) soft gluon corrections [33,34] was calculated as \( \sigma_{t\bar{t} W}^{\text{NNLL}}(8 \text{ TeV}) = 22.4 \pm 0.6 \text{ (scale)} \pm 1.4 \text{ (PDF)} \) pb assuming a top-quark mass, \( m_t \), of 172.5 GeV. The first uncertainty accounts for renormalisation and factorisation scale variations (from \( m_t/2 \) to 2\( m_t \)) and the second term covers the uncertainty in the parton distribution functions, evaluated using the MSTW2008 PDF set [35] at next-to-next-to-leading order (NNLO).

The \( t\bar{t} \) sample was generated with POWHEG BOX (v1.1) interfaced with PYTHIA 6 (v6.427) [36]. In the POWHEG BOX event generator, the CT10 PDFs were used, while the CTEQ6L1 PDFs were used for PYTHIA. The \( h_{\text{damp}} \) parameter, which effectively regulates the high-\( p_T \) gluon radiation, was set to \( m_t \). The predicted \( t\bar{t} \) production cross-section, \( \sigma_{t\bar{t}}(8 \text{ TeV}) = 252.9_{-8.6}^{+8.4} \text{ (scale)} \pm 11.7 \text{ (PDF + } \alpha_s \text{)} \) pb, was calculated with the Top++2.0 program to NNLO in perturbative QCD, including soft-gluon resummation to NNLL [37]. The first uncertainty comes from the sum in quadrature of the effects of independently varying \( \mu_F \) and \( \mu_R \). The uncertainty associated with variations in the PDFs and strong coupling constant, \( \alpha_s \), was evaluated following the PDF4LHC NLO prescription [38,39], which defines the central value as the midpoint of the uncertainty envelope of three PDF sets: MSTW2008 NNLO [35], CT10 NNLO [40] and NNPDF2.3 5f FFN [41]. The same procedures as for the \( t\bar{t} W \) samples were employed to determine the uncertainties due to the NLO matching method and the parton shower and hadronisation. Samples to evaluate the scale uncertainties were produced in a similar way, varying \( \mu_F \) and \( \mu_R \) together with the Perugia tune, but also adding variations in the \( h_{\text{damp}} \) parameter (for the up-variation, \( h_{\text{damp}} \) was changed to 2\( m_t \), while for the down variation it was kept at \( m_t \)).

The other single-top-quark production processes, \( s \)-channel and \( t \)-channel, were also generated with POWHEG BOX (v1.1) coupled to PYTHIA 6 (v6.426), using the same PDF sets as described for the other top-quark processes above. The predicted cross-sections at \( \sqrt{s} = 8 \text{ TeV} \) were calculated at NLO plus NNLL as 5.6 \pm 0.2 pb for the \( s \)-channel [42,43], and 87.8 \pm 3.4 pb for the \( t \)-channel [44,45] process.

The multi-leg LO generator SHERPA (v1.4.1) [46–48], together with the CT10 PDF sets, was used to simulate vector-boson production in association with jets. SHERPA was used to generate the hard process as well as the parton shower and the modelling of the underlying event. Double-counting between the inclusive \( V + n \) parton samples (with \( V = W \) or \( Z \)) and samples with associated heavy-quark pair production was avoided consistently by using massive \( c- \) and \( b- \)quarks in the shower. The predicted NNLO \( W + \)jets cross-section with \( W \) decaying leptonically was calculated as \( \sigma(p p \rightarrow \ell^\pm \nu_\ell X) = 36.3 \pm 1.9 \text{ nb} \) [49].\(^2\) For \( Z + \) jets the cross-section was calculated at NNLO in QCD for leptonic \( Z \) decays as \( \sigma(p p \rightarrow \ell^+ \ell^- X) = 3.72 \pm 0.19 \text{ nb} \) [49]. The ATLFAST2 simulation was used to generate these samples with sufficient statistics. For cross-checks of the \( W + \)jets modelling, an alternative sample generated with ALPGEN (v2.14) [50] with up to five additional partons, PYTHIA 6 (v6.426) and the CTEQ6L1 PDFs were used. Dibo-

\(^2\) The quoted cross-section is the sum of decays into \( e, \mu \) and \( \tau \).
son samples (WW/WZ/ZZ+jets) were generated with Herwig (v6.520) at LO QCD using the CTEQ6L1 PDF. The theoretical NLO cross-section for events with one lepton is 29.4 ± 1.5 pb [51].

Multijet events are selected in the analysis when they contain jets or photons misidentified as leptons or contain non-prompt leptons from hadron decays (both referred to as a ‘fake’ lepton). This background was estimated directly from data using the matrix method [52], which exploits differences in lepton identification and isolation properties between prompt and fake leptons. The data were processed with a second, ‘loose’ set of lepton selection criteria. The resulting sample was then corrected for efficiency differences between the two sets of cuts, and the contamination from non-prompt leptons from hadron decays (both referred to as the ‘fake’ lepton). This background was estimated directly from data using the matrix method [52], as input to the jet finding. The jet energy is further corrected by subtracting the contribution from pile-up events and applying an MC-based and a data-based calibration. The jet vertex fraction (JVF) [60] variable is used to identify the primary vertex from which the jet originated. The JVF criterion suppresses pile-up jets with $p_T < 50$ GeV and $|\eta| < 2.4$. To avoid possible overlap between jets and electrons, jets that are closer than $\Delta R = 0.2$ to an electron are removed. Afterwards, remaining electron candidates overlapping with jets within a distance of $\Delta R = 0.4$ are rejected. Finally, muons overlapping with jets within $\Delta R = 0.4$ are removed.

The identification of jets originating from the hadronisation of a $b$-quark (b-tagging) is based on various algorithms exploiting the long lifetime, high mass and high decay multiplicity of $b$-hadrons as well as the properties of the $b$-quark fragmentation. The outputs of these algorithms are combined in a neural network classifier to maximise the $b$-tagging performance [61]. The choice of $b$-tagging working point represents a trade-off between the efficiency for identifying $b$-jets and rejection of other jets. The chosen working point for this analysis corresponds to a $b$-tagging efficiency of 70%. The corresponding $c$-quark-jet rejection factor is about 5 and the light-quark-jet rejection factor is about 120. These efficiencies and rejection factors were obtained using simulated $t\bar{t}$ events. The tagging efficiencies in the simulation are corrected to match the efficiencies measured in data [61].

The $p_T^{\text{miss}}$ of the event, defined as the momentum imbalance in the plane transverse to the beam axis, is primarily due to neutrinos that escape detection. It is calculated as the negative vector sum of the transverse momenta of the reconstructed electrons, muons, jets and the clusters that are not associated with any of the previous objects (the ‘soft term’) [62]. Its magnitude is denoted $E_T^{\text{miss}}$.

5 Event selection

Events are required to have a hard-collision primary vertex. They also have to pass a single-lepton trigger requirement [14,63] and contain at least one electron or muon candidate with $p_T > 30$ GeV matched to the lepton that fired the trigger. The electron trigger requires an electron candidate, formed by an EM calorimeter cluster matched with a track, either with $E_T > 60$ GeV or with $E_T > 24$ GeV and additional isolation requirements. The muon trigger requires a muon candidate, defined as a reconstructed track in the muon spectrometer, either with $p_T > 36$ GeV or with $p_T > 24$ GeV.
and isolation requirements. If there is another lepton candidate with a transverse momentum above 25 GeV, the event is rejected. This lepton veto guarantees orthogonality with respect to the dilepton analysis. The contribution from leptonically decaying \( \tau \)-leptons is included. In the following, the electron or muon candidate is referred to as the lepton.

Events identified as containing jets from cosmic rays or beam-induced backgrounds or due to noise hot spots in the calorimeter are removed. Only jets with \( p_T > 30 \text{ GeV} \) and \( |\eta| < 2.4 \) are considered in the analysis. Additionally, a requirement of \( E_T^{\text{miss}} > 30 \text{ GeV} \) is applied, and the transverse mass\(^3\) of the leptonically decaying \( W \) boson must satisfy \( m_T(W_L) > 50 \text{ GeV} \).

In order to perform the measurement and validate the result, selected events are divided into different categories based on the jet and \( b \)-tagged jet multiplicities. The region with three jets of which one is \( b \)-tagged (3j1b) is called the signal region and is used to extract the \( tW \) cross-section. The region with four jets, two of them \( b \)-tagged (4j2b), contains a very pure sample of \( t\bar{t} \) events and is used as the \( t\bar{t} \) validation region to check the modelling of this background. Table 1 shows the expected and the observed numbers of events in the signal region after the event selection. All backgrounds except fake leptons, which is estimated using data-driven methods, are normalised to their expected cross-sections. The \( tW \) events constitute about 5\% of the total number of events. The major backgrounds are \( t\bar{t} \) production with about 58\%, and \( W \)+jets production with about 28\% of the total number of events. The \( W \)+jets events are subdivided into heavy flavour (HF), where a \( W \) boson is produced in association with \( b \)- or \( c \)-jets, and light flavour (LF). The total numbers of expected events agree within a few percent with the observed numbers of events.

6 Separation of signal from background

Differences between signal and background event kinematics are exploited to better separate them. The \( t\bar{t} \) background is inherently difficult to distinguish from the signal, motivating the use of an artificial neural network (NN) implemented in the NeuroBayes framework [64, 65]. Detailed information about how the NN is used in single-top-quark analyses can be found in Ref. [66]. The NN input variables are selected such that they contribute significantly to the statistical separation power between signal and background, while avoiding variables that would lead to an increase of the expected uncertainty in the signal cross-section. The observable \( m_T(W_3) \) (Fig. 2) provides very good separation of the signal from the background, but is strongly affected by uncertainties in the reconstructed jet energies as well as uncertainties in the \( b \)-tagging in \( t\bar{t} \) events. For this reason, \( m(W_3) \) is not used in the NN; instead a two-dimensional discriminant is constructed from \( m(W_3) \) and the response of the NN. The two-dimensional discriminant, explained in the following sections, allows the nuisance parameters affecting the variable \( m(W_3) \) to be partially constrained.

### 6.1 Invariant mass of the hadronically decaying \( W \) boson

The variable \( m(W_3) \) is computed from the four-momenta of the two selected untagged jets. For the signal and the \( t\bar{t} \) background, the distribution of \( m(W_3) \) exhibits a peak near the mass of the \( W \) boson, shown in Fig. 2a. The peak results from events where the two untagged jets are correctly matched to the hadronically decaying \( W \) boson. This is less likely to happen for \( t\bar{t} \) events than for \( tW \) events due to the higher \( b \)-jet multiplicity and the limited \( b \)-tagging efficiency. On the other hand, the \( W \)+jets background does not feature such a peak since the \( W \) boson must decay leptonically for the events to pass the selection. Figure 2b shows the pre-fit distribution of \( m(W_3) \), and also demonstrates good pre-fit modelling of the data.

### 6.2 Neural network

The NN is trained using simulated events with the two reconstructed untagged jets matched within \( \Delta R < 0.35 \) to the generator-level jets originating from a \( W \)-boson decay in the MC simulation and having a reconstructed mass of \( 65 \text{ GeV} < m(W_3) < 92.5 \text{ GeV} \). As events are required to

\[ \text{Table 1 Expected signal and background and observed number of events in the signal 3j1b region. The cross-section for } tW \text{ production is taken to be the theory prediction. The uncertainties include statistical and systematic uncertainties.} \]

<table>
<thead>
<tr>
<th>Process</th>
<th>Signal region (3j1b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( tW ) ((\sigma_{tW} = 22.4 \text{ pb}))</td>
<td>6300 ± 600</td>
</tr>
<tr>
<td>( t\bar{t} )</td>
<td>77000 ± 6000</td>
</tr>
<tr>
<td>( t\bar{t} ), ( t )-channel</td>
<td>4180 ± 290</td>
</tr>
<tr>
<td>( t\bar{t} ), ( s )-channel</td>
<td>307 ± 19</td>
</tr>
<tr>
<td>( W )+jets, HF</td>
<td>31000 ± 14000</td>
</tr>
<tr>
<td>( W )+jets, LF</td>
<td>6000 ± 3000</td>
</tr>
<tr>
<td>( Z )+ jets</td>
<td>3900 ± 1700</td>
</tr>
<tr>
<td>( W/WZW/ZZ+jets )</td>
<td>650 ± 280</td>
</tr>
<tr>
<td>Fake leptons</td>
<td>4300 ± 1900</td>
</tr>
<tr>
<td>Total background</td>
<td>128000 ± 18000</td>
</tr>
<tr>
<td>Total signal + background</td>
<td>134000 ± 18000</td>
</tr>
<tr>
<td>Observed</td>
<td>13463</td>
</tr>
</tbody>
</table>

\(^3\) The transverse mass is calculated using the momentum of the lepton associated with the \( W \) boson, \( p_T^{\text{miss}} \), and the azimuthal angle between the two: \( m_T(W_L) = m_T(\ell\nu) = \sqrt{2p_T(\ell) \cdot E_T^{\text{miss}}[1 - \cos(\Delta\phi(\ell, p_T^{\text{miss}}))]} \).
contain a lepton, only $tW$, $t\bar{t}$ and diboson events can have a pair of jets matched to the hadronic $W$-boson decay. Given that the contribution from diboson production is very small, the background sample used for the training consists entirely of $t\bar{t}$ events. Following the training procedure mentioned before, the following four variables (ordered by significance) are selected as input for the NN:

- the transverse momentum of the $tW$ system, $p_T(W_H W_L j_B)$, divided by the sum of the objects’ transverse momenta,

$$\rho_T(W_H, W_L, j_B) = \frac{p_T(W_H W_L j_B)}{p_T(W_H) + p_T(W_L) + p_T(j_B)},$$

where the four-momentum of $W_L$ is the sum of the four-momenta of the electron or muon and the neutrino, and the four-momentum of the neutrino is determined using $E_T^{\text{miss}}$ from the solution of a quadratic equation. The use of $\rho_T(W_H, W_L, j_B)$, instead of the transverse momentum of the $tW$ system, decreases the background contribution in the signal-like region of the NN response and results in a gain of sensitivity;

- the invariant mass of the reconstructed $tW$ system, $m(W_H W_L j_B)$;

- the absolute value of the difference between the pseudorapidities of the lepton and the leading untagged jet in $p_T, |\Delta\eta(\ell, j_1)|$;

- the absolute value of the pseudorapidity of the lepton, $|\eta(\ell)|$.

Figure 3 compares the data with the prediction for the NN input variables. For all variables, the simulation provides a good description of the data.

The distribution of the NN response is subdivided into eight bins, with the edges placed approximately at the 12.5% quantiles of a 50:50 mixture of $tW$ and $t\bar{t}$ events. Figure 4a shows the shape of the NN response for the $tW$ and $t\bar{t}$ processes and Fig. 4b presents the comparison between data and Monte Carlo simulation.

6.3 Two-dimensional discriminant

For the two-dimensional discriminant, $m(W_H)$ is used on the abscissa and the NN response on the ordinate of the two-dimensional discriminant. Outside of the aforementioned $m(W_H)$ range from 65 to 92.5 GeV, the bins corresponding to different values of the NN response are merged, i.e. the NN response is ignored. The two-dimensional distribution is presented in Fig. 5.

The bins are then rearranged on a one-dimensional axis in column-major order. The resulting one-dimensional distribution is presented in Fig. 6, together with a comparison of the shapes. The first three bins and the last ten bins correspond directly to the bins of $m(W_H)$ below 65 GeV and above 92.5 GeV respectively. In between are four blocks of eight bins, corresponding to the NN output in slices of $m(W_H)$. 
Fig. 3  Pre-fit distributions of the NN input variables in the $tW$ signal (3j1b) region with $65 \, \text{GeV} \leq m(W_H) \leq 92.5 \, \text{GeV}$. Small backgrounds are subsumed under ‘Other’. The simulated distributions are normalised to their theoretical cross-sections. The dashed uncertainty band includes statistical and systematic uncertainties. The last bin includes the overflow events. The lower panels show the ratio of the observed and the predicted number of events in each bin.

Fig. 4  
(a) Shape of the NN response in the signal (3j1b) region. The distribution contains those events with $65 \, \text{GeV} \leq m(W_H) \leq 92.5 \, \text{GeV}$. The distributions for the $tW$ process and the $t\bar{t}$ process normalised to unity are shown.  
(b) Pre-fit NN output distribution in the 3j1b region. Small backgrounds are subsumed under ‘Other’. The simulated distributions are normalised to their theoretical cross-sections. The dashed uncertainty band includes statistical and systematic uncertainties. The lower panels show the ratio of the observed and the predicted number of events in each bin.
Fig. 5 Predicted distribution of the two-dimensional discriminant in the signal (3j1b) region. The proportions of the coloured areas reflect the expected composition in terms of $tW$, $t\bar{t}$, $W+$-jets and other processes. The numbers correspond to the bin order when projecting the discriminant onto one axis as in Fig. 6. The last bin on the horizontal axis includes the overflow events.

Inside each of the blocks, the $tW$-to-$t\bar{t}$ ratio increases significantly from left to right.

7 Systematic uncertainties

Uncertainties in the jet reconstruction arise from the jet energy scale (JES), jet energy resolution (JER), JVF requirement and jet reconstruction efficiency. The effect of the uncertainty in the JES [59] is evaluated by varying the reconstructed energies of the jets in the simulated samples. It is split into multiple components, taking into account the uncertainty in the calorimeter response, the detector simulation, the choice of MC event generator, the subtraction of pile-up, and differences in the detector response for jets initiated by a gluon, a light-flavour quark, or a $b$-quark. In a similar way, the JER uncertainty is represented using several components, which account for the uncertainty in different $p_T$ and $\eta$ regions of the detector, the difference between data and MC simulation, as well as the noise contribution in the forward detector region [59]. The uncertainty in jet reconstruction efficiency is estimated by randomly removing simulated jets from the events according to the jet reconstruction inefficiency measured with dijet events [67]. The JVF uncertainty is evaluated by varying the JVF criterion [60].

The scale factors used to correct the $b$-tagging efficiency in simulation compared to the efficiency in data are varied separately for $b$-jet, $c$-jet and light-flavour jets. Independent sources of uncertainty affecting the $b$-jet tagging efficiency and $c$-jet mis-tagging efficiency are considered depending on the jet kinematics, e.g. the variation of the $b$-quark jets is subdivided into 6 components. Uncertainties associated with the lepton selection arise from the trigger, reconstruction, identification, isolation and lepton momentum scale and resolution [54,68,69].

All systematic uncertainties in the reconstruction of jets and leptons are propagated to the uncertainty in $E_T^{miss}$. In addition, dedicated uncertainties are assigned to the soft term of the $E_T^{miss}$, which accounts for energy deposits in the calorimeter which are not matched to high-$p_T$ physics objects [62].

The uncertainty in the integrated luminosity for the data set used in this analysis is 1.9%. It is derived following the methodology detailed in Ref. [18]. This systematic uncertainty is applied to all contributions determined from the MC simulation.

Uncertainties stemming from theoretical models are evaluated using alternative MC samples for $tW$ and $t\bar{t}$ processes.
The renormalisation and factorisation scales are varied in the matrix element and in the parton shower together with the amount of QCD radiation. Both scales are varied simultaneously in the matrix element and in the parton shower. The variation of both \( \mu_r \) and \( \mu_f \) by a factor of 0.5 is combined with the Perugia 2012radHi tune, while the variation of the scale parameters by a factor of 2.0 is combined with the Perugia 2012radLo tune [26]. This (radiation) uncertainty is considered uncorrelated between the \( tW \) and \( tt \) processes. The NLO matrix element generator uncertainty is estimated by comparing two NLO matching methods: POWHEG BOX and MC@NLO, both interfaced with HERWIG. The parton shower, hadronisation and underlying-event systematic uncertainties are computed by comparing POWHEG BOX with either PYTHIA or HERWIG. These are treated as fully correlated between the \( tW \) and \( tt \) processes. The uncertainty due to the treatment of the interference effects of the \( tW \) and \( tt \) processes is evaluated by using the \( tW \) DS scheme instead of the DR scheme, both generated using POWHEG BOX with PYTHIA. The effect of the PDF uncertainties on the acceptance is taken into account for both the \( tW \) signal and the \( tt \) background and treated as uncorrelated between the processes, following the studies in Ref. [70].

The uncertainties in the theoretical cross-section calculations are process dependent and vary from 4% for the \( t \)-channel to 6% for \( tt \) (see Sect. 3). In addition, there are large uncertainties in the \( Z/W + \text{jets} \) production cross-sections. For every jet an additional uncertainty of 24% is assumed [71]. The uncertainty in the normalisation of \( W/Z \)-boson production in association with three jets is 42%, and the rate of \( \ell \)-boson events with heavy-flavour jets is allowed to vary by an additional 20%.

The modelling of the \( W + \text{jets} \) background was crosschecked using ALPGEN with PYTHIA. The shape of the \( W + \text{jets} \) background was found to be consistent with the nominal SHERPA prediction. Hence no dedicated systematic uncertainty is assigned to the choice of generator, in order to avoid double-counting of the statistical uncertainty of the prediction (model statistics).

Uncertainties related to the modelling of the fake-lepton background take into account the choice of control region for the determination of the fake- and real-lepton efficiencies, the choice of parameterisation, and the normalisation of the prompt-lepton backgrounds in the determination of the efficiencies [52].

The uncertainty due to the limited size of the simulated samples and the fake-lepton background (model statistics) is estimated through the procedure detailed in Refs. [72,73]: for every bin of the discriminant, an independent parameter is assigned which describes the variation of the predicted event rate constrained by its statistical uncertainty.

8 Statistical analysis

A binned profile maximum-likelihood fit to the discriminant in the signal region is used to determine the \( tW \) cross-section. The likelihood function is defined as a product of Poisson probability terms over all the bins of the discriminant in the signal region and Gaussian penalty terms,

\[
L(\mu, \vec{\theta}; \vec{n}) = \prod_{i} \text{Pois}(n_i; \nu_i(\mu, \vec{\theta})) \prod_{k} G(\theta_k; 0, 1),
\]

where the \( n_i(\nu_i) \) is the observed (expected) number of events in each bin \( i \) of the discriminant. The expected number of events depends on the signal-strength parameter, \( \mu \), which is a multiplicative factor on the predicted signal cross-section. Nuisance parameters (NPs), \( \theta_k \), are used to encode the effects of the systematic uncertainties in the expected number of events. The Gaussian penalty terms model the external constraints on these parameters. The estimated parameters, denoted by \( \hat{\mu} \) and \( \hat{\vec{\theta}} \), are obtained by maximising \( L(\mu, \vec{\theta}; \vec{n}) \).

The likelihood function is composed and evaluated with the HISTFACTORY program [74], part of the ROOSTATS framework [75]. The minimisation is performed with the MINUIT package [76], using MINOS to compute the error estimates.

The statistical significance, \( Z \), of the result is estimated by comparing the likelihood values of two hypotheses. The background-only hypothesis is that there is no signal in the data (or equivalently, \( \mu = 0 \)). The signal-plus-background hypothesis is that the signal exists with the signal strength obtained from the fit to data. With the asymptotic approximation [77], the significance is calculated using a test statistic based on the profile likelihood ratio,

\[
Z^2 = -2 \log \frac{L(\mu = 0, \vec{\theta} = \hat{\vec{\theta}}_{\mu=0})}{L(\mu = \hat{\mu}, \vec{\theta} = \hat{\vec{\theta}})},
\]

where \( \hat{\theta}_{\mu=0} \) denotes the estimates of the nuisance parameters that maximise the likelihood function under the background-only hypothesis. The expected significance is calculated by replacing \( \hat{n} \) in the likelihood function with the Asimov dataset for the nominal signal-plus-background hypothesis \( (\mu = 1, \theta = \hat{\theta}) \).

9 Cross-section measurement

The \( tW \) cross-section is extracted from the fit to data in the signal region. Given the Standard Model prediction, the extracted signal strength is expected to be \( \hat{\mu} = 1.00 \pm 0.35 \). The measured value is \( \hat{\mu} = 1.16 \pm 0.31 \), corresponding to an observed cross-section of \( \sigma_{tW}^{\text{obs}} = 26 \pm 7 \text{ pb} \), which is con-
The (post-fit) impact of each systematic uncertainty on the measured signal strength is estimated by means of conditional fits, i.e. the fit is repeated while keeping the corresponding nuisance parameter fixed at the ±1 standard deviation value of the post-fit error interval. The resulting change in the estimate of the signal strength quantifies the impact of the uncertainty. For each nuisance parameter, the +1 and −1 sigma variations are found to be symmetric about the best-fit value to a very good approximation. Table 2 shows the impacts of the systematic uncertainties on the observed fit result, where the impacts of uncertainties with similar sources have been added in quadrature. The dominant uncertainties are due to the amount of QCD radiation in signal and t¯t background. Also, the nuisance parameter for the NLO matching for tW and t¯t is constrained: the choice of MC@NLO is not supported by the data, reducing the impact of the choice from 9% pre-fit to 3% post-fit.

A few nuisance parameters are pulled away from the pre-fit expectation. For the parameter associated with the choice of parton-shower generator, a blend of PYTHIA and HERWIG gives the best description of the data, while the nominal PYTHIA prediction is disfavoured at the two-sigma level. The b-tagging parameter with the largest effect on the overall b-tagging efficiency is pulled by about one sigma, corresponding to a decrease of about 1% to 2% in the b-tagging efficiency compared to the pre-fit expectation. Given that the b-tagging calibration partially relies on dijet events [61], which correspond to a different environment regarding the production mechanism of the b-jets, the pull is reasonable.

Table 3 shows the post-fit event yields of each process. The uncertainties in the yields are computed taking the correlations between nuisance parameters and processes into account. The post-fit estimates are well within the uncertainties of the pre-fit expectation (Table 1), while most of their uncertainties are reduced. The normalisation uncertainty for W + HF jets changes from almost 50% to about 10%.

Figure 7 shows the post-fit distributions for the NN input variables, the NN output response and the m(Wt) in the signal region. The post-fit plots use the parameter estimates obtained in the fit of the discriminant, including their uncertainties, and demonstrate a good description of the data.

Figure 8a shows that the data are well described by the model in the signal region. Figure 8b shows the strongest support for the validity of the fit result by comparing the

---

Table 2: List of systematic uncertainties considered in the analysis and their relative impact on the observed signal strength, evaluated as described in the text. The ‘model statistics’ uncertainty is dominated by the W + jets background:

<table>
<thead>
<tr>
<th>Source</th>
<th>Uncertainty (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jet energy scale</td>
<td>10</td>
</tr>
<tr>
<td>b-tagging</td>
<td>8</td>
</tr>
<tr>
<td>Jet energy resolution</td>
<td>7</td>
</tr>
<tr>
<td>E_T^miss reconstruction</td>
<td>7</td>
</tr>
<tr>
<td>Lepton reconstruction</td>
<td>4</td>
</tr>
<tr>
<td>Luminosity</td>
<td>3</td>
</tr>
<tr>
<td>Jet vertex fraction</td>
<td>3</td>
</tr>
<tr>
<td>t¯t radiation</td>
<td>10</td>
</tr>
<tr>
<td>tW radiation</td>
<td>9</td>
</tr>
<tr>
<td>tW→t¯t interference</td>
<td>7</td>
</tr>
<tr>
<td>t¯t cross-section normalisation</td>
<td>6</td>
</tr>
<tr>
<td>Other background cross-section normalisations</td>
<td>5</td>
</tr>
<tr>
<td>tW and t¯t parton shower</td>
<td>4</td>
</tr>
<tr>
<td>tW and t¯t NLO matching</td>
<td>3</td>
</tr>
<tr>
<td>PDF</td>
<td>1</td>
</tr>
<tr>
<td>Model statistics</td>
<td>11</td>
</tr>
<tr>
<td>Data statistics</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>27</td>
</tr>
</tbody>
</table>

QCD radiation in signal and t¯t background. Also, the nuisance parameter for the NLO matching for tW and t¯t is constrained: the choice of MC@NLO is not supported by the data, reducing the impact of the choice from 9% pre-fit to 3% post-fit.
expected distributions and observed distributions in the $t\bar{t}$ validation region. It shows that the uncertainty due to the extrapolation from the signal region is small, and therefore provides a stringent test that the main background is well modelled.

**10 Conclusion**

The inclusive cross-section for the production of a single top quark in association with a $W$ boson in the single-lepton channel is measured using an integrated luminosity...
of 20.2 fb$^{-1}$ of $\sqrt{s} = 8$ TeV proton–proton collision data collected by the ATLAS detector at the LHC in 2012. A neural network is used to separate the signal from the $t\bar{t}$ background. A two-dimensional discriminant, built from the neural-network response and the mass of the hadronically decaying $W$ boson, is used to extract the cross-section. Evidence for $tW$ production in the single-lepton channel is obtained with an observed (expected) significance of 4.5 (3.9$\sigma$) standard deviations. The measured cross-section is:

$$\sigma_{tW}^{\text{meas}} = 26 \pm 7 \text{ pb},$$

which is consistent with the SM expectation of $\sigma_{tW}^{\text{SM}} = 22.4 \pm 1.5 \text{ pb}$.

Acknowledgements We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently. We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWFW and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; ANID, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DSRF and DSSRC, Denmark; IN2P3-CNRS and CEA-DRF/IRFU, France; SRNSFG, Georgia; BMHEF and MPG, Germany; GSRT, Greece; RGC and Hong Kong SAR, China; ISF and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; NWO, Netherlands; RCN, Norway; MNiSW and NCN, Poland; FCT, Portugal; MINE/IFA, Romania; JINR; MES of Russia and NRC KI, Russian Federation; MESTD, Serbia; MSSR, Slovakia; ARRS and MIZŠ, Slovenia; DST/NRF, South Africa; MICINN, Spain; SRC and Wallace Foundation, Sweden; SERI, SNSF and Cantons of Bern and Geneva, Switzerland; MOST, Taiwan; TAEK, Turkey; STFC, United Kingdom; DOE and NSF, USA. In addition, individual groups and members have received support from BCKDF, CANARIE, Compute Canada, CRC and IVADO, Canada; Beijing Municipal Science & Technology Commission, China; COST, ERC, ERDF, Horizon 2020 and Marie Skłodowska-Curie Actions, European Union; Investissements d’Avenir Labex, Investissements d’Avenir Idex and ANR, France; DFG and AvH Foundation, Germany; Herakleitos, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSRF, Greece; BSF-NSF and GIF, Israel; La Caixa Banking Foundation, CERCA Programme Generalitat de Catalunya and PROMETEO and GenT Programmes Generalitat Valenciana, Spain; Göran Gustafssons Stiftelse, Sweden; 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. [78].

Data Availability Statement This manuscript has no associated data or the data will not be deposited. [Authors’ comment: “All ATLAS scientific output is published in journals, and preliminary results are made available in Conference Notes. All are openly available, without restriction on use by external parties beyond copyright law and the standard conditions agreed by CERN. Data associated with journal publications are also made available: tables and data from plots (e.g. cross section values, likelihood profiles, selection efficiencies, cross section limits, ...) are stored in appropriate repositories such as HEPDATA (http://hepdata.cedar.ac.uk/).” ATLAS also strives to make additional material related to the paper available that allows a reinterpretation of the data in the context of new theoretical models. For example, an extended encapsulation of the analysis is often provided for measurements in the framework of RIVET (http://rivet.hepforge.org/).” This information is taken from the ATLAS Data Access Policy, which is a public document that can be downloaded from http://opendata.cern.ch/record/413 [opendata.cern.ch].]

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. Funded by SCOAP3.

Springer
12 (a) Bahcesehir University, Faculty of Engineering and Natural Sciences, Istanbul, Turkey; (b) Istanbul Bilgi University, Faculty of Engineering and Natural Sciences, Istanbul, Turkey; (c) Department of Physics, Bogazici University, Istanbul, Turkey; (d) Department of Physics Engineering, Gaziantep University, Gaziantep, Turkey
13 Institute of Physics, Azerbaijan Academy of Sciences, Baku, Azerbaijan
14 Institut de Física d’Altes Energies (IFAE), Barcelona Institute of Science and Technology, Barcelona, Spain
15 (a) Institute of High Energy Physics, Chinese Academy of Sciences, Beijing, China; (b) Physics Department, Tsinghua University, Beijing, China; (c) Department of Physics, Nanjing University, Nanjing, China; (d) University of Chinese Academy of Science (UCAS), Beijing, China
16 Institute of Physics, University of Belgrade, Belgrade, Serbia
17 Department for Physics and Technology, University of Bergen, Bergen, Norway
18 Physics Division, Lawrence Berkeley National Laboratory and University of California, Berkeley, CA, USA
19 Institut für Physik, Humboldt Universität zu Berlin, Berlin, Germany
20 Albert Einstein Center for Fundamental Physics and Laboratory for High Energy Physics, University of Bern, Bern, Switzerland
21 School of Physics and Astronomy, University of Birmingham, Birmingham, UK
22 (a) Facultad de Ciencias y Centro de Investigaciones, Universidad Antonio Nariño, Bogotá, Colombia; (b) Departamento de Física, Universidad Nacional de Colombia, Bogotá, Colombia, Colombia
23 (a) Dipartimento di Fisica, INFN Bologna and Universita’ di Bologna, Bologna, Italy; (b) INFN Sezione di Bologna, Bologna, Italy
24 Physikalisches Institut, Universität Bonn, Bonn, Germany
25 Department of Physics, Boston University, Boston, MA, USA
26 Department of Physics, Brandeis University, Waltham, MA, USA
27 (a) Transilvania University of Brasov, Brasov, Romania; (b) Horia Hulubei National Institute of Physics and Nuclear Engineering, Bucharest, Romania; (c) Department of Physics, Alexandru Ioan Cuza University of Iasi, Iasi, Romania; (d) National Institute for Research and Development of Isotopic and Molecular Technologies, Physics Department, Cluj-Napoc, Romania; (e) University Politehnica Bucharest, Bucharest, Romania; (f) West University in Timisoara, Timisoara, Romania
28 (a) Faculty of Mathematics, Physics and Informatics, Comenius University, Bratislava, Slovak Republic; (b) Department of Subnuclear Physics, Institute of Experimental Physics of the Slovak Academy of Sciences, Kosice, Slovak Republic
29 Physics Department, Brookhaven National Laboratory, Upton, NY, USA
30 Departamento de Física, Universidad de Buenos Aires, Buenos Aires, Argentina
31 California State University, Long Beach, CA, USA
32 Cavendish Laboratory, University of Cambridge, Cambridge, UK
33 (a) Department of Physics, University of Cape Town, Cape Town, South Africa; (b) iThemba Labs, Western Cape, South Africa; (c) Department of Mechanical Engineering Science, University of Johannesburg, Johannesburg, South Africa; (d) University of South Africa, Department of Physics, Pretoria, South Africa; (e) School of Physics, University of the Witwatersrand, Johannesburg, South Africa
34 Department of Physics, Carleton University, Ottawa, ON, Canada
35 (a) Faculté des Sciences Ain Chock, Réseau Universitaire de Physique des Hautes Energies - Université Hassan II, Casablanca, Morocco; (b) Faculté des Sciences, Université Ibn-Tofail, Kénitra, Morocco; (c) Faculté des Sciences Semlalia, Université Cadi Ayyad, L.P.H.E.A.-Marrakech, Marrakesh, Morocco; (d) Faculté des Sciences, Université Mohammed Premier and LPTPM, Oujda, Switzerland; (e) Faculté des sciences, Université Mohammed V, Rabat, Morocco
36 CERN, Geneva, Switzerland
37 Enrico Fermi Institute, University of Chicago, Chicago, IL, USA
38 LPC, Université Clermont Auvergne, CNRS/IN2P3, Clermont-Ferrand, France
39 Nevis Laboratory, Columbia University, Irvington, NY, USA
40 Niels Bohr Institute, University of Copenhagen, Copenhagen, Denmark
41 (a) Dipartimento di Fisica, Università della Calabria, Rende, Italy; (b) INFN Gruppo Collegato di Cosenza, Laboratori Nazionali di Frascati, Frascati, Italy
42 Physics Department, Southern Methodist University, Dallas, TX, USA
43 Physics Department, University of Texas at Dallas, Richardson, TX, USA
44 National Centre for Scientific Research “Demokritos”, Agia Paraskevi, Greece
45 (a) Department of Physics, Stockholm University, Stockholm, Sweden; (b) Oskar Klein Centre, Stockholm, Sweden

Spranger
aq Also at National Research Nuclear University MEPhI, Moscow, Russia
ar Also at Physics Department, An-Najah National University, Nablus, Palestine
as Also at Physics Dept, University of South Africa, Pretoria, South Africa
at Also at Physikalisches Institut, Albert-Ludwigs-Universität Freiburg, Freiburg, Germany
au Also at School of Physics, Sun Yat-sen University, Guangzhou, China
av Also at The City College of New York, New York, NY, USA
aw Also at The Collaborative Innovation Center of Quantum Matter (CICQM), Beijing, China
ax Also at TRIUMF, Vancouver, BC, Canada
ay Also at Universita di Napoli Parthenope, Naples, Italy
* Deceased