Measurement of differential cross-sections of a single top quark produced in association with a W boson at $\sqrt{s}=13$ TeV with ATLAS

ATLAS Collaboration

DOI
10.1140/epjc/s10052-018-5649-8

Publication date
2018

Document Version
Final published version

Published in
European Physical Journal C

License
CC BY

Citation for published version (APA):

Download date: 11 Dec 2022
Measurement of differential cross-sections of a single top quark produced in association with a $W$ boson at $\sqrt{s} = 13$ TeV with ATLAS

ATLAS Collaboration*

CERN, 1211 Geneva 23, Switzerland

Abstract  The differential cross-section for the production of a $W$ boson in association with a top quark is measured for several particle-level observables. The measurements are performed using 36.1 fb$^{-1}$ of $pp$ collision data collected with the ATLAS detector at the LHC in 2015 and 2016. Differential cross-sections are measured in a fiducial phase space defined by the presence of two charged leptons and exactly one jet matched to a $b$-hadron, and are normalised with the fiducial cross-section. Results are found to be in good agreement with predictions from several Monte Carlo event generators.

Contents  
1 Introduction ........................................ 1 
2 ATLAS detector ...................................... 2 
3 Data and Monte Carlo samples ...................... 3 
4 Object reconstruction ............................... 4 
5 Event selection ....................................... 5 
6 Separation of $tW$ signal from $t\bar{t}$ background ... 6 
7 Unfolding and cross-section determination ....... 8 
8 Systematic uncertainties ............................ 9 
  8.1 Sources of systematic uncertainty .............. 9 
  8.2 Procedure for estimation of uncertainty ...... 12 
9 Results .............................................. 13 
10 Conclusion ........................................ 14 
References .......................................... 15 

1 Introduction

Single-top-quark production proceeds via three channels through electroweak interactions involving a $Wtb$ vertex at leading order (LO) in the Standard Model (SM): the $t$-channel, the $s$-channel, and production in association with a $W$ boson ($tW$). The cross-section for each of these channels depends on the relevant Cabibbo–Kobayashi–Maskawa (CKM) matrix element $V_{tb}$ and form factor $f_V^t$ [1–3] such that the cross-section is proportional to $|f_V^t V_{tb}|^2$ [4,5], i.e. depends on the coupling between the $W$ boson, top and $b$ quarks. The $tW$ channel, represented in Fig. 1, has a $pp$ production cross-section at $\sqrt{s} = 13$ TeV of $\sigma_{\text{theory}} = 71.7 \pm 1.8$ (scale) $\pm 3.4$ (PDF) pb [6], and contributes approximately 24% of the total single-top-quark production rate at 13 TeV. At the LHC, evidence for this process with 7 TeV collision data was presented by the ATLAS Collaboration [7] (with a significance of 3.6$\sigma$), and by the CMS Collaboration [8] (with a significance of 4.0$\sigma$). With 8 TeV collision data, CMS observed the $tW$ channel with a significance of 6.1$\sigma$ [9] while ATLAS observed it with a significance of 7.7$\sigma$ [10]. This analysis extends an ATLAS analysis [11] which measured the production cross-section with 13 TeV data collected in 2015.

Accurate estimates of rates and kinematic distributions of the $tW$ process are difficult at higher orders in $\alpha_S$ since the process is not well-defined due to quantum interference with the $t\bar{t}$ production process. A fully consistent theoretical picture can be reached by considering $tW$ and $t\bar{t}$ to be components of the complete $Wb\bar{W}b$ final state in the four flavour scheme [12]. In the $t\bar{t}$ process the two $Wb$ systems are produced on the top quark mass shell, and so a proper treatment of this doubly resonant component is important in the study of $tW$ beyond leading order. Two commonly used approaches are diagram removal (DR) and diagram subtraction (DS) [13]. In the DR approach, all next-to-leading order (NLO) diagrams that overlap with the doubly resonant $t\bar{t}$ contributions are removed from the calculation of the $tW$ amplitude, violating gauge invariance. In the DS approach, a subtraction term is built into the amplitude to cancel out the $t\bar{t}$ component close to the top quark resonance while respecting gauge invariance.

This paper describes differential cross-section measurements in the $tW$ dilepton final state, where events contain two
oppositely charged leptons (henceforth “lepton” refers to an electron or muon) and two neutrinos. This channel is chosen because it has a better ratio of signal and background production over other background processes than the single lepton+jets channel, where large $W+$jets backgrounds are relatively difficult to separate from top quark events. Distributions are unfolded to observables based on stable particles produced in Monte Carlo (MC) simulation. Measurements are performed in a fiducial phase space, defined by the presence of two charged leptons as well as the presence of exactly one central jet containing $b$-hadrons ($b$-jet) and no other jets. This requirement on the jet multiplicity is expected to suppress the contribution from $t\bar{t}$ production, where a pair of $b$-jets is more commonly produced, as well as reducing the importance of $t\bar{t}$-$tW$ interference effects [12]. After applying the reconstruction-level selection of fiducial events (described in Sect. 5) backgrounds from $t\bar{t}$ and other sources are subtracted according to their predicted distributions from MC simulation. The definition of the fiducial event selection is chosen to match the lepton and jet requirements at reconstruction level. Exactly two leptons with $p_T > 20$ GeV and $|\eta| < 2.5$ are required, and at least one of the leptons must satisfy $p_T > 27$ GeV. Exactly one $b$-tagged jet satisfying $p_T > 25$ GeV and $|\eta| < 2.5$ must be present. No requirement is placed on $E_T^{\text{miss}}$ or $m_{T\ell}$. A boosted decision tree (BDT) is used to separate the $tW$ signal from the large $t\bar{t}$ background by placing a fixed requirement on the BDT response.

Although the top quark and the two $W$ bosons cannot be directly reconstructed due to insufficient kinematic constraints, one can select a ratio of observables that are correlated with kinematic properties of $tW$ production and are sensitive to differences in theoretical modelling. Particle energies and masses are also preferred to projections onto the transverse plane in order to be sensitive to polar angular information while keeping the list of observables as short as possible. Unfolded distributions are measured for:

- the energy of the $b$-jet, $E(b)$;
- the mass of the leading lepton and $b$-jet, $m(\ell_1b)$;
- the mass of the sub-leading lepton and the $b$-jet, $m(\ell_2b)$;
- the energy of the system of the two leptons and $b$-jet, $E(\ell\ell b)$;
- the transverse mass of the leptons, $b$-jet and neutrinos, $m_{T}(\ell\ell\nu\nu b)$; and
- the mass of the two leptons and the $b$-jet, $m(\ell\ell b)$.

The top quark production is probed most directly by $E(b)$, the only final-state object that can unambiguously be matched to the decay products of the top quark. The top-quark decay is probed by $m(\ell_1 b)$ and $m(\ell_2 b)$, which are sensitive to angular correlations of decay products due to production spin correlations. The combined $tW$-system is probed by $E(\ell\ell b)$, $m_{T}(\ell\ell\nu\nu b)$, and $m(\ell\ell b)$. At reconstruction level, the transverse momenta of the neutrinos in $m_{T}(\ell\ell\nu\nu b)$ are represented by the measured $E_T^{\text{miss}}$ (reconstructed as described in Sect. 4).

At particle level the vector summed transverse momenta of simulated neutrinos (selected as defined in Sect. 4) are used in $m_{T}(\ell\ell\nu\nu b)$. All other quantities for leptons and jets are taken simply from the relevant reconstructed or particle-level objects. These observables are selected to minimise the bias introduced by the BDT requirement, as certain observables are highly correlated with the BDT discriminant. These cannot be effectively unfolded due to shaping effects that the BDT requirement imposes on the overall acceptance, and thus are not considered in this measurement. The background-subtracted data are unfolded using an iterative procedure [14] to correct for resolution and acceptance effects, biases, and particles outside the fiducial phase space of the measurement. The differential cross-sections are normalised with the fiducial cross-section, which cancels out many of the largest uncertainties.

2 ATLAS detector

The ATLAS detector [15] at the LHC covers nearly the entire solid angle\(^1\) around the collision point, and consists of an inner tracking detector (ID) surrounded by a thin superconducting solenoid 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 pseu-

\(^1\) 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 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 $y = (1/2) \ln \left[ (E + p_z)/(E - p_z) \right]$. Springer
dorapidity 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 [16], was added between Run 1 and Run 2 of the LHC, at a radius of 33 mm around a new, thinner, beam pipe. 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 [17], 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 events for offline storage.

3 Data and Monte Carlo samples

The data events analysed in this paper correspond to an integrated luminosity of 36.1 fb$^{-1}$ collected from the operation of the LHC in 2015 and 2016 at $\sqrt{s} = 13$ TeV with a bunch spacing of 25 ns and an average number of collisions per bunch crossing $\langle \mu \rangle$ of around 23. They are required to be recorded in periods where all detector systems are flagged as operating normally.

Monte Carlo simulated samples are used to estimate the efficiency to select signal and background events, train and test the BDT, estimate the migration of observables from parameterised detector simulation, estimate systematic uncertainties, and validate the analysis tools. The nominal samples, used for estimating the central values for efficiencies and background processes from PYTHIA 8.186 [21] using a set of tuned parameters called the A2 tune [22] and the MSTW2008LO parton distribution function (PDF) set [23]. 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 MC samples and fixed-order calculations used for this analysis the top quark mass $m_t$ is set to 172.5 GeV and the $W \rightarrow \ell\nu$ branching ratio is set to 0.108 per lepton flavour. The EVGEN v1.2.0 program [24] was used to simulate properties of the bottom and charmed hadron decays except for samples generated with SHERPA, which uses internal modules.

The nominal $tW$ event samples [25] were produced using the POWHEG-Box v1 [26–30] event generator with the CT10 PDF set [31] in the matrix-element calculations. The parton shower, hadronisation, and underlying event were simulated using PYTHIA 6.428 [32] with the CTEQ6L1 PDF set [33] and the corresponding Perugia 2012 (P2012) tune [34]. The DR scheme [13] was employed to handle the interference between $tW$ and $t\bar{t}$, and was applied to the $tW$ sample. For comparing MC predictions to data, the predicted $tW$ cross-section at $\sqrt{s} = 13$ TeV is scaled by a $K$-factor and set to the NLO value with next-to-next-to-leading logarithmic (NNLL) soft-gluon corrections: $\sigma_{\text{theory}} = 71.7 \pm 1.8$ (scale) $\pm 3.4$ (PDF) pb [6]. The first uncertainty accounts for the renormalisation and factorisation scale variations (from 0.5 to 2 times $m_t$), while the second uncertainty originates from uncertainties in the MSTW2008 NLO PDF sets.

Additional $tW$ samples were generated to estimate systematic uncertainties in the modelling of the signal process. An alternative $tW$ sample was generated using the DS scheme instead of DR. A $tW$ sample generated with MadGraph5 aMC@NLO v2.2.2 [35] (instead of the POWHEG-Box) interfaced with Herwig++ 2.7.1 [36] and processed through the ATLAST2 fast simulation was used to estimate uncertainties associated with the modelling of the NLO matrix-element event generator. A sample generated with POWHEG-Box interfaced with Herwig++ (instead of PYTHIA 6) is used to estimate uncertainties associated with the parton shower, hadronisation, and underlying-event models. This sample is also compared with the previously mentioned MadGraph5 aMC@NLO sample to estimate a matrix-element event generator uncertainty with a consistent parton shower event generator. In both cases, the UE-EE-5 tune of Ref. [37] was used for the underlying event. Finally, in order to estimate uncertainties arising from additional QCD radiation in the $tW$ events, a pair of samples were generated with POWHEG-Box interfaced with PYTHIA 6 using ATLAST2 and the P2012 tune with higher and lower radiation relative to the nominal set, together with varied renormalisation and factorisation scales. In order to avoid comparing two different detector response models when estimating systematic uncertainties, another version of the nominal POWHEG-Box with PYTHIA 6 sample was also produced with ATLAST2.

The nominal $t\bar{t}$ event sample [25] was produced using the POWHEG-Box v2 [26–30] event generator with the CT10 PDF set [31] in the matrix-element calculations. The parton shower, hadronisation, and underlying event were simulated using PYTHIA 6.428 [32] with the CTEQ6L1 PDF set [33] and the corresponding Perugia 2012 (P2012) tune [34]. The renormalisation and factorisation scales are set to $m_t$ for the
$tW$ process and to $\sqrt{m_\ell^2 + p_T^2}$ for the $t\bar{t}$ process, and the $h_{\text{damp}}$ resummation damping factor is set to equal the mass of the top quark.

Additional $t\bar{t}$ samples were generated to estimate systematic uncertainties. Like the additional $tW$ samples, these are used to estimate the uncertainties associated with the matrix-element event generator (a sample produced using ATLAST2 fast simulation with MadGraph5_aMC@NLO v2.2.2 interfaced with Herwig++ 2.7.1), parton shower and hadronisation models (a sample produced using ATLAST2 with POWHEG-BOX interfaced with Herwig++ 2.7.1) and additional QCD radiation. To estimate uncertainties on additional QCD radiation in $t\bar{t}$, a pair of samples is produced using full simulation with the varied sets of P2012 parameters for higher and lower radiation, as well as with varied renormalisation and factorisation scales. In these samples the resummation damping factor $h_{\text{damp}}$ is doubled in the case of higher radiation. The $t\bar{t}$ cross-section is set to $\sigma_{t\bar{t}} = 831.8^{+19.8}_{-29.2}$ (scale) $\pm 35.1$ (PDF + αS) pb as calculated with the Top++ 2.0 program to NNLO, including soft-gluon resummation to NNLL [38]. 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 αS, following the PDF4LHC prescription with the MSTW2008 68% CL NNLO, C10 NNLO and NNPDF2.3 5f FFN PDF sets [39–42].

Samples used to model the $Z +$ jets background [43] were simulated with SHERPA 2.2.1 [44]. In these, the matrix element is calculated for up to two partons at NLO and four partons at LO using COMIX [45] and OPENLOOPS [46], and merged with the SHERPA parton shower [47] using the ME+PS@NLO prescription [48]. The NNPDF3.0 NNLO PDF set [49] was used in conjunction with SHERPA parton shower tuning, with a generator-level cut-off on the dilepton invariant mass of $m_{\ell\ell} > 40$ GeV applied. The $Z +$ jets events are normalised using NNLO cross-sections computed with FEWZ [50].

Diboson processes with four charged leptons, three charged leptons and one neutrino, or two charged leptons and two neutrinos [51] were simulated using the SHERPA 2.1.1 event generator. The matrix elements contain all diagrams with four electroweak vertices. NLO calculations were 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 COMIX and OPENLOOPS were used. Their outputs were combined with the SHERPA parton shower using the ME+PS@NLO prescription [48]. The CT10 PDF set with dedicated parton shower tuning was used. The cross-sections provided by the event generator (which are already at NLO) were used for diboson processes.

4 Object reconstruction

Electron candidates are reconstructed from energy deposits in the EM calorimeter associated with ID tracks [17]. 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 a transverse momentum of $p_T > 20$ GeV. Further requirements on the electromagnetic shower shape, ratio of calorimeter energy to tracker momentum, and other variables are combined into a likelihood-based discriminant [52], with signal electron efficiencies measured to be at least 85%, increasing for higher $p_T$. Candidate electrons also must satisfy requirements on the distance from the ID track to the beamline or 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 with respect to the beamline, $d_0$, must satisfy $|d_0|/\sigma_{d_0} < 5$, where $\sigma_{d_0}$ is the uncertainty in $d_0$. The longitudinal impact parameter, $z_0$, must satisfy $|\Delta z_0 \sin \theta| < 0.5$ mm, where $\Delta z_0$ is the longitudinal distance from the primary vertex along the beamline and $\theta$ is the angle of the track to the beamline. Furthermore, electrons must satisfy isolation requirements based on ID tracks and topological clusters in the calorimeter [53], designed to achieve an isolation efficiency of 90% (99%) for $p_T = 25(60)$ GeV.

Muon candidates are identified by matching MS tracks with ID tracks [54]. The candidates must satisfy requirements on hits in the MS and on the compatibility of 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 from the combined ID and MS track to the beamline or 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, where $d_0$ and $z_0$ are defined as above for electrons. An isolation requirement based on ID tracks and topological clusters in the calorimeter is imposed, which targets an isolation efficiency of 90% (99%) for $p_T = 25(60)$ GeV.

Jets are reconstructed from topological clusters of energy deposited in the calorimeter [53] using the anti-$k_t$ algorithm [55] with a radius parameter of 0.4 implemented in the FastJet package [56]. Their energies are corrected to account for pile-up and calibrated using a $p_T$- and $\eta$-dependent correction derived from Run 2 data [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 is constructed using a two-dimensional likelihood method [58]. For jets with $p_T < 60$ GeV and $|\eta| < 2.4$, a jet-vertex-tagger requirement corresponding to a 92% efficiency while rejecting 98% of jets from pile-up and noise is imposed.
The tagging of $b$-jets uses a multivariate discriminant which exploits the long lifetime of $b$-hadrons and large invariant mass of their decay products relative to $c$-hadrons and unstable light hadrons [59,60]. The discriminant is calibrated to achieve a 77% $b$-tagging efficiency and a rejection factor of about 4.5 against jets containing charm quarks ($c$-jets) and 140 against light-quark and gluon jets in a sample of simulated $t\bar{t}$ events. The jet 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, $E_{T}^{\text{miss}}$, is a measure of the transverse momentum imbalance, primarily due to neutrinos that escape detection. In addition to the identified jets, electrons and muons, a track-based soft term is included in the $E_{T}^{\text{miss}}$ calculation by considering tracks associated with the hard-scattering vertex in the event which are not also associated with an identified jet, electron, 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. First, identified muons that deposit energy in the calorimeter and share a track with an electron are removed, followed by the removal of any remaining electrons sharing a track with a muon. This step is designed to avoid cases where a muon mimics an electron through radiation of a hard photon. Next, the jet closest to each electron within a $y$–$\phi$ cone of size $\Delta R_{y,\phi} \equiv \sqrt{(\Delta y)^2 + (\Delta \phi)^2} = 0.2$ is removed to reduce the proportion of electrons being reconstructed as jets. Next, electrons with a distance $\Delta R_{y,\phi} < 0.4$ from any of the remaining jets are removed to reduce backgrounds from non-prompt, non-isolated electrons originating from heavy-flavour hadron decays. Jets with fewer than three tracks and distance $\Delta R_{y,\phi} < 0.2$ from a muon are then removed to reduce the number of jet fakes from muons depositing energy in the calorimeters. Finally, muons with a distance $\Delta R_{y,\phi} < 0.4$ from any of the surviving jets are removed to avoid contamination due to non-prompt muons from heavy-flavour hadron decays.

Definitions of particle-level objects in MC simulation are based on stable ($c\tau > 10\text{ mm}$) outgoing particles [64]. Particle-level prompt charged leptons and neutrinos that arise from decays of $W$ bosons or $Z$ bosons are accepted. The charged leptons are then dressed with nearby photons, considering all photons that satisfy $\Delta R_{y,\phi}(\ell, \gamma) < 0.1$ and do not originate from hadrons, adding the four-momenta of all selected photons to the bare lepton to obtain the dressed lepton four-momentum. Particle-level jets are built from all remaining stable particles in the event after excluding leptons and the photons used to dress the leptons, clustering them using the anti-$k_t$ algorithm with $R = 0.4$. Particle-level jet $b$-tagging is performed by checking the jets for any associated $b$-hadron with $p_T > 5\text{ GeV}$. This association is achieved by reclustering jets with $b$-hadrons included in the input list of particles, but with their $p_T$ scaled down to negligibly small values. Jets containing $b$-hadrons after this reclustering are considered to be associated to a $b$-hadron.

5 Event selection

Events passing the reconstruction-level selection are required to have at least one interaction vertex, to pass a single-electron or single-muon trigger, and to contain at least one jet with $p_T > 25\text{ GeV}$. Single-lepton triggers used in this analysis are designed to select events containing a well-identified charged lepton with high transverse momentum [17]. They require a $p_T$ of at least 20 GeV (26 GeV) for muons and 24 GeV (26 GeV) for electrons for the 2015 (2016) data set, 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 $p_T$.

Events are required to contain exactly two oppositely charged leptons with $p_T > 20\text{ GeV}$; events with a third charged lepton with $p_T > 20\text{ GeV}$ are rejected. At least one lepton must have $p_T > 27\text{ GeV}$, and at least one of the selected electrons (muons) must be matched within a $\Delta R_{y,\phi}$ 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 heavy-flavour 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, regardless of whether they are $tW$ fake events or fake events from other sources, they 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 + $ jets, diboson and fake-lepton backgrounds. The samples consist almost entirely of $tW$ signal and $t\bar{t}$ background, which are subsequently separated by the BDT discriminant. Events in which the two leptons have the same flavour and an invariant mass consistent with a $Z$ boson ($81 < m_{\ell\ell} < 101\text{ GeV}$) are vetoed, as well as those with an invariant mass $m_{\ell\ell} < 40\text{ GeV}$. Further requirements placed on $E_{T}^{\text{miss}}$ and $m_{\ell\ell}$ depend on the flavour of the selected leptons. Events with different-flavour leptons contain backgrounds from $Z \rightarrow \tau\tau$, and are required to
have $E_T^{\text{miss}} > 20\text{ GeV}$, with the requirement raised to
$E_T^{\text{miss}} > 50\text{ GeV}$ when the dilepton invariant mass satisfies
$m_{\ell\ell} < 80\text{ GeV}$. All events with same-flavour leptons,
which contain backgrounds from $Z \to \ell\ell$ and $Z \to \mu\mu$,
must satisfy $E_T^{\text{miss}} > 40\text{ GeV}$. For same-flavour leptons,
the $Z + \text{jets}$ background is concentrated in a region of the
$m_{\ell\ell} - E_T^{\text{miss}}$ plane corresponding to values of $m_{\ell\ell}$ near the $Z$
mass, and towards low values of $E_T^{\text{miss}}$. Therefore, a selection
in $E_T^{\text{miss}}$ and $m_{\ell\ell}$ is used to remove these backgrounds:
events with $40\text{ GeV} < m_{\ell\ell} < 81\text{ GeV}$ are required to satisfy
$E_T^{\text{miss}} > 1.25 \times m_{\ell\ell}$ while events with $m_{\ell\ell} > 101\text{ GeV}$ are
required to satisfy $E_T^{\text{miss}} > 300\text{ GeV} - 2 \times m_{\ell\ell}$.

Finally, events are required to have exactly one jet which is
$b$-tagged. For validation of the signal and background
models, additional regions are also defined according to the
number of jets and the number of $b$-tagged jets, but are not
used in the differential cross-section measurement, primarily
due to the lower signal purity in these regions. These
regions are labelled by the number $n$ of selected jets and
the number $m$ of selected $b$-tagged jets as $njbmb$ (for example
the 2j1b region consists of events with 2 selected jets
of which 1 is $b$-tagged), and show good agreement between
data and predictions. The event yields for signal and back-
grounds with their total systematic uncertainties, as well
as the number of observed events in the data in the signal
and validation regions are shown in Fig. 2, and the yields
in the signal region are shown in Table 1. Distributions of
the events passing these requirements are shown in Fig. 3
at reconstruction level. Most of the predictions agree well
with data within the systematic errors, which are highly cor-
related bin-to-bin due to the dominance of a small number
of sources of large normalisation uncertainties. The distrib-
ution of $m_T(\ell\ell\nu\nu)$, which shows a slope in the ratio of
data to prediction, has a $p$ value of 2–4% for the predictions
to describe the observed distribution after taking bin-to-bin
correlations into account.

### Table 1 Predicted and observed yields in the 1j1b signal region before
and after the application of the BDT requirement

<table>
<thead>
<tr>
<th>Process</th>
<th>Events</th>
<th>Events BDT response &gt; 0.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$tW$</td>
<td>8300 ± 1400</td>
<td>1970 ± 560</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>38,400 ± 6600</td>
<td>3400 ± 1300</td>
</tr>
<tr>
<td>$Z + \text{jets}$</td>
<td>620 ± 310</td>
<td>159 ± 80</td>
</tr>
<tr>
<td>Diboson</td>
<td>230 ± 58</td>
<td>81 ± 20</td>
</tr>
<tr>
<td>Fakes</td>
<td>220 ± 220</td>
<td>19 ± 19</td>
</tr>
<tr>
<td>Predicted</td>
<td>47,800 ± 7300</td>
<td>5600 ± 1700</td>
</tr>
<tr>
<td>Observed</td>
<td>45,273</td>
<td>5043</td>
</tr>
</tbody>
</table>

### 6 Separation of $tW$ signal from $t\bar{t}$ background

To separate $tW$ signal events from background $t\bar{t}$ events, a
BDT technique [65] is used to combine several observables
into a single discriminant. In this analysis, the BDT imple-
mentation is provided by the TMVA package [66], using
the GradientBoost algorithm. The approach is based on the
BDT developed for the inclusive cross-section measurement
in Ref. [11].

The BDT is optimised by using the sum of the nominal
tW MC sample, the alternative tW MC sample with the diagram
subtraction scheme and the nominal $t\bar{t}$ MC sample; for
each sample, half of the events are used for training while
the other half is reserved for testing. A large list of variables
is prepared to serve as inputs to the BDT. An optimisation
procedure is then carried out to select a subset of input vari-
ableS and a set of BDT parameters (such as the number of
trees in the ensemble and the maximum depth of the indi-
vidual decision trees). The optimisation is designed to provide
the best separation between the $tW$ signal and the $t\bar{t}$ back-
ground 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 Sect. 4 for each event. For a set of objects $o_1 \cdots o_n$:
$pT_\ell(o_1 \cdots o_n)$ is the transverse momentum of vector sums
of various subsets; $\sum E_T$ is the scalar sum of the transverse
momenta of all objects which contribute to the $E_T^{\text{miss}}$ calu-
Fig. 3 Distributions of the observables chosen to be unfolded after selection at the reconstruction level but before applying the BDT selection. The signal and backgrounds are normalised to their theoretical predictions, and the error bands represent the total systematic uncertainties in the MC predictions. The last bin of each distribution contains overflow events. The panels give the yields in number of events, and the ratios of the numbers of observed events to the total prediction in each bin.

\[ \eta(o_1 \cdots o_n) \] is the pseudorapidity of vector sums of various subsets; \( m(o_1 \cdots o_n) \) is the invariant mass of various subsets. For vector sums of two systems of objects \( s_1 \) and \( s_2 \): \( \Delta p_T(s_1, s_2) \) is the \( p_T \) difference; and \( C(s_1, s_2) \) is the ratio of the scalar sum of \( p_T \) to the sum of energy, called the centrality.
Table 2 The variables used in the signal region BDT and their separation power (denoted $S$). The variables are derived from the four-momenta of the leading lepton ($\ell_1$), sub-leading lepton ($\ell_2$), the $b$-jet ($b$) and $E_T^{miss}$. The last row gives the separation power of the BDT discriminant response

<table>
<thead>
<tr>
<th>Variable</th>
<th>$S \times 10^{-2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_T(\ell_1 \ell_2 E_T^{miss}b)$</td>
<td>4.1</td>
</tr>
<tr>
<td>$\Delta p_T(\ell_1 \ell_2 b, E_T^{miss})$</td>
<td>2.5</td>
</tr>
<tr>
<td>$\sum E_T$</td>
<td>2.3</td>
</tr>
<tr>
<td>$\eta(\ell_1 \ell_2 E_T^{miss}b)$</td>
<td>1.3</td>
</tr>
<tr>
<td>$\Delta p_T(\ell_1 \ell_2, E_T^{miss})$</td>
<td>1.1</td>
</tr>
<tr>
<td>$p_T(\ell_1 \ell_2 b)$</td>
<td>1.0</td>
</tr>
<tr>
<td>$C(\ell_1 \ell_2)$</td>
<td>0.9</td>
</tr>
<tr>
<td>$m(\ell_2, b)$</td>
<td>0.2</td>
</tr>
<tr>
<td>$m(\ell_1, b)$</td>
<td>0.1</td>
</tr>
<tr>
<td>BDT response</td>
<td>8.1</td>
</tr>
</tbody>
</table>

The final set of input variables used in the BDT is listed in Table 2 along with the separation power of each variable.\(^2\) The distributions of these variables are compared between the MC predictions and observed data, and found to be well modelled. The BDT discriminant distributions from MC predictions and data are compared and shown in Fig. 4.

To select a signal-enriched portion of events in the signal region, the BDT response is required to be larger than 0.3. The effect of this requirement on event yields is shown in Table 1. The BDT requirement lowers systematic uncertainties by reducing contributions from the $t\bar{t}$ background, which is subject to large modeling uncertainties. For example, the total systematic uncertainty in the fiducial cross-section is reduced by 16% of the total when applying the BDT response requirement, compared to having no requirement. The exact value of the requirement is optimised to reduce the total uncertainty of the measurement over all bins, considering both statistical and systematic uncertainties.

7 Unfolding and cross-section determination

The iterative Bayesian unfolding technique in Ref. [14], as implemented in the RooUnfold software package [67], is used to correct for detector acceptance and resolution effects and the efficiency to pass the event selection. The unfolding procedure includes bin-by-bin correction for out-of-fiducial ($C_{\text{off}}$) events which are reconstructed but fall outside the fiducial acceptance at particle level:

$$C_{\text{off}} = \frac{N_{\text{fid}}}{N_{\text{reco}}}$$

followed by the iterative matrix unfolding procedure. The matrix $M$ is the migration matrix, and $M^{-1}$ represents the application of the iterative unfolding procedure with migration information from $M$. The iterative unfolding is followed by another bin-by-bin correction to the efficiency to reconstruct a fiducial event ($C_{\text{eff}}$):

$$\frac{1}{C_{\text{eff}}} = \frac{N_{\text{fid}}}{N_{\text{fid}} N_{\text{reco}}}$$

In both expressions, “fid” refers to events passing the fiducial selection, “reco” refers to events passing reconstruction-level requirements, and “fid&reco” refers to events passing both. This full unfolding procedure is then described by the expression for the number of unfolded events in bin $i$ ($N_{\text{ufd}}$) of the particle-level distribution:

$$N_{\text{ufd}} = \frac{1}{C_{\text{eff}}} \sum_j M_{ij}^{-1} C_{\text{off}} \left( N_{\text{data}} - B_j \right),$$

where $i$ ($j$) indicates the bin at particle (reconstruction) level, $N_{\text{data}}$ is the number of events in data and $B_j$ is the sum of all background contributions. Table 3 gives the number of iterations used for each observable in this unfolding step. The bias is defined as the difference between the unfolded and true values. The number of iterations is chosen to minimise
the growth of the statistical uncertainty propagated through the unfolding procedure while operating in a regime where the bias is sufficiently independent of the number of iterations. The optimal number of iterations is small for most observables, but a larger number is picked for \(E(b)\), where larger off-diagonal elements of the migration matrix cause slower convergence of the method.

The list of observables chosen was also checked for shaping induced by the requirement on the BDT response, since strong shaping can make the unfolding unstable. These shaping effects were found to be consistently well-described by the various MC models considered. Any residual differences in the predictions of different MC event generators would increase MC modelling uncertainties, thus ensuring shaping effects of the BDT are covered by the total uncertainties.

Unfolded event yields \(N_i^{\text{ufd}}\) are converted to cross-section values as a function of an observable \(X\) using the expression:

\[
\frac{d\sigma_i}{dX} = \frac{N_i^{\text{ufd}}}{L \Delta_i},
\]

where \(L\) is the integrated luminosity of the data sample and \(\Delta_i\) is the width of bin \(i\) of the particle-level distribution. Differential cross-sections are divided by the fiducial cross-section to create a normalised distribution. The fiducial cross-section is simply the sum of the cross-sections in each bin multiplied by the corresponding bin widths:

\[
\sigma_i^{\text{fid}} = \sum_i \left( \frac{d\sigma_i}{dX} \cdot \Delta_i \right) = \sum_i \frac{N_i^{\text{ufd}}}{L}.
\]

### 8 Systematic uncertainties

#### 8.1 Sources of systematic uncertainty

The experimental sources of uncertainty include the uncertainty in the lepton efficiency scale factors used to correct simulation to data, the lepton energy scale and resolution, the \(E_T^{\text{miss}}\) soft-term calculation, the jet energy scale and resolution, the \(b\)-tagging efficiency, and the luminosity.

The JES uncertainty [57] is divided into 18 components, which are derived using \(\sqrt{s} = 13\) TeV data. The uncertainties from data-driven calibration studies of \(Z/\gamma +\text{jet}\) and dijet events are represented with six orthogonal components using the eigenvector decomposition procedure, as demonstrated in Ref. [68]. Other components include model uncertainties (such as flavour composition, \(\eta\) intercalibration model). The most significant JES uncertainty components for this measurement are the data-driven 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 studies of Run-1 data. These studies are then cross-calibrated and checked to confirm good agreement with Run-2 data.

As discussed in Sect. 4, the \(E_T^{\text{miss}}\) calculation includes contributions from leptons and jets in addition to soft terms which arise primarily from low-\(p_T\) pile-up jets and underlying-event activity [62,63]. The uncertainty associated with the leptons and jets is propagated from the corresponding uncertainties in the energy/momentum scales and resolutions, and it is classified together with the uncertainty associated with the corresponding objects. The uncertainty associated with the soft term is estimated by comparing the simulated soft-jet energy scale and resolution to that in data.

Uncertainties in the scale factors used 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 the extrapolation uncertainty for high-\(p_T\) jets [59,60].

Systematic uncertainties in lepton momentum resolution and scale, trigger efficiency, isolation efficiency, and identification efficiency are also considered [52–54]. These uncertainties arise from corrections to simulation based on studies of \(Z \rightarrow ee\) and \(Z \rightarrow \mu\mu\) data. In this measurement, the effects of the uncertainties in these corrections are relatively small.

A 2.1\% uncertainty is assigned to the integrated luminosity. 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.

Uncertainties stemming from theoretical models are estimated by comparing a set of predicted distributions produced with different assumptions. The main uncertainties are due to the NLO matrix-element (ME) event generator, paron shower and hadronisation event generator, radiation tuning and scale choice 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 paron 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 event generator is treated as uncor-

---

**Table 3** Number of iterations chosen in the unfolding procedure for each of the observables used in the measurement

<table>
<thead>
<tr>
<th>Observable</th>
<th>Number of iterations</th>
</tr>
</thead>
<tbody>
<tr>
<td>(E(b))</td>
<td>15</td>
</tr>
<tr>
<td>(m(\ell\ell\nu\nu))</td>
<td>7</td>
</tr>
<tr>
<td>(m(\ell\ell\nu\nu))</td>
<td>5</td>
</tr>
<tr>
<td>(E(\ell\ell\nu\nu))</td>
<td>5</td>
</tr>
<tr>
<td>(m(\ell\ell\nu\nu))</td>
<td>7</td>
</tr>
<tr>
<td>(m(\ell\ell\nu\nu))</td>
<td>5</td>
</tr>
</tbody>
</table>
Table 4 Summary of the measured normalised differential cross-sections, with uncertainties shown as percentages. The uncertainties are divided into statistical and systematic contributions

<table>
<thead>
<tr>
<th>$E(b)$</th>
<th>bin [GeV]</th>
<th>(1/$\sigma$) $d\sigma/dx$ [GeV$^{-1}$]</th>
<th>Stat. uncertainty</th>
<th>Total syst. uncertainty</th>
<th>Total uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[25, 60]</td>
<td>0.00438</td>
<td>25</td>
<td>33</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>[60, 100]</td>
<td>0.00613</td>
<td>20</td>
<td>28</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>[100, 135]</td>
<td>0.00474</td>
<td>12</td>
<td>22</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>[135, 175]</td>
<td>0.00252</td>
<td>28</td>
<td>28</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>[175, 500]</td>
<td>0.00103</td>
<td>37</td>
<td>37</td>
<td>74</td>
</tr>
<tr>
<td>$m(\ell_1b)$</td>
<td>bin [GeV]</td>
<td>(1/$\sigma$) $d\sigma/dx$ [GeV$^{-1}$]</td>
<td>Stat. uncertainty</td>
<td>Total syst. uncertainty</td>
<td>Total uncertainty</td>
</tr>
<tr>
<td></td>
<td>[0, 60]</td>
<td>0.000191</td>
<td>130</td>
<td>39</td>
<td>169</td>
</tr>
<tr>
<td></td>
<td>[60, 100]</td>
<td>0.00428</td>
<td>22</td>
<td>22</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>[100, 150]</td>
<td>0.00806</td>
<td>12</td>
<td>24</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>[150, 200]</td>
<td>0.00333</td>
<td>14</td>
<td>33</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td>[200, 250]</td>
<td>0.00114</td>
<td>22</td>
<td>22</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>[250, 400]</td>
<td>0.000879</td>
<td>32</td>
<td>29</td>
<td>61</td>
</tr>
<tr>
<td>$m(\ell_2b)$</td>
<td>bin [GeV]</td>
<td>(1/$\sigma$) $d\sigma/dx$ [GeV$^{-1}$]</td>
<td>Stat. uncertainty</td>
<td>Total syst. uncertainty</td>
<td>Total uncertainty</td>
</tr>
<tr>
<td></td>
<td>[0, 50]</td>
<td>0.00184</td>
<td>30</td>
<td>30</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>[50, 100]</td>
<td>0.00845</td>
<td>11</td>
<td>20</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>[100, 150]</td>
<td>0.00531</td>
<td>11</td>
<td>24</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>[150, 200]</td>
<td>0.00114</td>
<td>14</td>
<td>33</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td>[200, 250]</td>
<td>0.000879</td>
<td>14</td>
<td>29</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>[250, 400]</td>
<td>0.000879</td>
<td>14</td>
<td>33</td>
<td>47</td>
</tr>
<tr>
<td>$E(\ell\ell b)$</td>
<td>bin [GeV]</td>
<td>(1/$\sigma$) $d\sigma/dx$ [GeV$^{-1}$]</td>
<td>Stat. uncertainty</td>
<td>Total syst. uncertainty</td>
<td>Total uncertainty</td>
</tr>
<tr>
<td></td>
<td>[50, 175]</td>
<td>0.000597</td>
<td>30</td>
<td>30</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>[175, 275]</td>
<td>0.00322</td>
<td>12</td>
<td>22</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>[275, 375]</td>
<td>0.00185</td>
<td>12</td>
<td>24</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>[375, 500]</td>
<td>0.00135</td>
<td>14</td>
<td>33</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td>[500, 700]</td>
<td>0.000832</td>
<td>14</td>
<td>29</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>[700, 1200]</td>
<td>0.000167</td>
<td>14</td>
<td>29</td>
<td>43</td>
</tr>
<tr>
<td>$m_T(\ell\ell\nu\nu b)$</td>
<td>bin [GeV]</td>
<td>(1/$\sigma$) $d\sigma/dx$ [GeV$^{-1}$]</td>
<td>Stat. uncertainty</td>
<td>Total syst. uncertainty</td>
<td>Total uncertainty</td>
</tr>
<tr>
<td></td>
<td>[50, 275]</td>
<td>0.0033</td>
<td>7.1</td>
<td>29</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>[275, 375]</td>
<td>0.00123</td>
<td>7.8</td>
<td>38</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>[375, 500]</td>
<td>0.000856</td>
<td>7.8</td>
<td>38</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>[500, 1000]</td>
<td>5.51 x 10^{-5}</td>
<td>7.8</td>
<td>38</td>
<td>45</td>
</tr>
<tr>
<td>$m(\ell\ell b)$</td>
<td>bin [GeV]</td>
<td>(1/$\sigma$) $d\sigma/dx$ [GeV$^{-1}$]</td>
<td>Stat. uncertainty</td>
<td>Total syst. uncertainty</td>
<td>Total uncertainty</td>
</tr>
<tr>
<td></td>
<td>[0, 125]</td>
<td>0.00051</td>
<td>35</td>
<td>35</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>[125, 175]</td>
<td>0.00533</td>
<td>15</td>
<td>25</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>[175, 225]</td>
<td>0.00538</td>
<td>15</td>
<td>25</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>[225, 300]</td>
<td>0.00242</td>
<td>15</td>
<td>25</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>[300, 400]</td>
<td>0.000949</td>
<td>15</td>
<td>25</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>[400, 1000]</td>
<td>0.000208</td>
<td>15</td>
<td>25</td>
<td>40</td>
</tr>
</tbody>
</table>

related between the $tW$ and $t\bar{t}$ processes, while the uncertainty from the parton shower event generator is treated as correlated. The radiation tuning and scale choice 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 $tW$ and $t\bar{t}$ processes. These choices of correlations are based on Ref. [11], and were checked to be no less conservative than the alternative options. The choice of scheme to account for the interference between the $tW$ and $t\bar{t}$ 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 [31] prediction and the central PDF4LHC15 prediction (PDF central value) is taken and symmetrised together with the internal uncertainty set provided with PDF4LHC15.

Additional normalisation uncertainties are applied to each background. A 100% uncertainty is applied to the normalisation of the background from non-prompt and fake leptons, an uncertainty of 50% is applied to the Z+jets background, and a 25% normalisation uncertainty is assigned to diboson backgrounds. These uncertainties are based on earlier ATLAS studies of background simulation in top.
Fig. 5 Normalised differential cross-sections unfolded from data, compared with selected MC models, with respect to $E(b)$, $m(ℓ_1b)$, $m(ℓ_2b)$, and $E(ℓℓb)$. Data points are placed at the horizontal centre of each bin, and the error bars on the data points show the statistical uncertainties. The total uncertainty in the first bin of the $m(ℓ_1b)$ distribution (not shown) is 140%. See Sect. 1 for a description of the observables plotted.

These normalisation uncertainties are not found to have a large impact on the final measurement due to the small contribution of these backgrounds in the signal region as well as their cancellation in the normalised cross-section measurement. An uncertainty of 5.5% is applied to the $t\bar{t}$ normalisation to account for the quark analyses [71].
scale, $\alpha_S$, and PDF uncertainties in the NNLO cross-section calculation.

Uncertainties due to the size of the MC samples are estimated using pseudoexperiments. An ensemble of pseudodata is created by fluctuating the MC samples within the statistical uncertainties. Each set of pseudodata is used to construct $M_{ij}$, $C_{ij}^{\text{eff}}$, and $C_{ij}^{\text{off}}$, and the nominal MC sample is unfolded. The width of the distribution of unfolded values from this ensemble is taken as the statistical uncertainty. Additional non-closure uncertainties are added in certain cases after stress-testing the unfolding procedure with injected Gaussian or linear functions. Each distribution is tested by reweighting the input MC sample according to the injected function, unfolding, and checking that the weights are recovered in the unfolded distribution. The extent to which the unfolded weighted data are biased with respect to the underlying weighted generator-level distribution is taken as the unfolding non-closure uncertainty.

8.2 Procedure for estimation of uncertainty

The propagation of uncertainties through the unfolding process proceeds by constructing the migration matrix and efficiency corrections with the baseline sample and unfolding with the varied sample as input. In most cases, the baseline sample is from POWHEG-BOX+PYTHIA 6 and produced with the full detector simulation, but in cases where the varied sample uses the ATLFAST2 fast simulation, the baseline sample is also changed to use ATLFAST2. For uncertainties modifying background processes, varied samples are prepared by taking into account the changes in the background induced by a particular systematic effect. Experimental uncertainties are treated as correlated between signal and background in this procedure. The varied samples are unfolded and compared to the corresponding particle-level distribution from the MC event generator; the relative difference in each bin is the estimated systematic uncertainty.

The covariance matrix $C$ for each differential cross-section measurement is computed following a procedure similar to that used in Ref. [72]. Two covariance matrices are summed to form the final covariance. The first one is computed using 10,000 pseudoexperiments and includes statistical uncertainties as well as systematic uncertainties from experimental sources. The statistical uncertainties are included by independently fluctuating each bin of the data distribution according to Poisson distributions for each pseudoexperiment. Each bin of the resulting pseudodata distribution is then fluctuated according to a Gaussian distribution for each experimental uncertainty, preserving bin-to-bin correlation information for each uncertainty. The other matrix includes the systematic uncertainties from event generator model uncertainties, PDF uncertainties, unfolding non-closure uncertainties, and MC statistical uncertainties. In this
second matrix, the bin-to-bin correlation value is set to zero for the non-closure and MC statistical uncertainties, and set to unity for the other uncertainties. The impact of setting the bin-to-bin correlation value to unity was compared for the non-closure uncertainty, and this choice was found to have negligible impact on the results. This covariance matrix is used to compute a $\chi^2$ and corresponding $p$ value to assess how well the measurements agree with the predictions. The $\chi^2$ values are computed using the expression:

$$\chi^2 = \mathbf{v}^T \mathbf{C}^{-1} \mathbf{v},$$

where $\mathbf{v}$ is the vector of differences between the measured cross-sections and predictions.

9 Results

Unfolded particle-level normalised differential cross-sections are given in Table 4. In Figs. 5 and 6, the results are shown compared to the predictions of various MC event generators, and in Fig. 7 the main systematic uncertainties for each distribution are summarised. The results show that the largest uncertainties come from the size of the data sample as well as $t\bar{t}$ and $tW$ MC modelling.

The comparison between the data and Monte Carlo predictions is summarised in Table 5, where $\chi^2$ values and corresponding $p$ values are listed. In general, most of the MC models show fair agreement with the measured cross-sections, with no particularly low $p$ values observed. Notably, for each
distribution there is a substantial negative slope in the ratio of predicted to observed cross-sections, indicating there are more events with high-momentum final-state objects than several of the MC models predict. This effect is most visible in the $E(\ell \ell \bar{b})$ distribution, where the lower $p$ values for all MC predictions reflect this. In most cases, differences between the MC predictions are smaller than the uncertainty on the data, but there are some signs that Powheg-Box+Herwig++ deviates more from the data and from the other predictions in certain bins of the $E(\ell \ell \bar{b})$, $m(\ell \ell \bar{b})$, and $m(\ell_1 \bar{b})$ distributions. The predictions of DS and DR samples likewise give very similar results for all observables as expected from the fiducial selection. The predictions of Powheg-Box+Pythia 6 with varied initial- and final-state radiation tuning were also examined but not found to give significantly different distributions in the fiducial phase space of this analysis.

Both the statistical and systematic uncertainties have a significant impact on the result. The exact composition varies bin-to-bin but there is no single source of uncertainty that dominates each normalised measurement. Some of the largest systematic uncertainties are those related to $t \bar{t}$ and $tW$ modelling. The cancellation in the normalised differential cross-sections is very effective at reducing a number of systematic uncertainties. The most notable cancellation is related to the $t \bar{t}$ parton shower model uncertainty, which is quite dominant prior to dividing by the fiducial cross-section.

### 10 Conclusion

The differential cross-section for the production of a $W$ boson in association with a top quark is measured for several particle-level observables. The measurements are performed using $36.1 \,fb^{-1}$ of $pp$ collision data with $\sqrt{s} = 13$ TeV collected in 2015 and 2016 by the ATLAS detector at the LHC. Cross-sections are measured in a fiducial phase space defined by the presence of two charged leptons and exactly one jet identified as containing $b$-hadrons. Six observables are chosen, constructed from the masses and energies of leptons and jets as well as the transverse momenta of neutrinos. Measurements are normalised with the fiducial cross-section, causing several of the main uncertainties to cancel out. Dominant uncertainties arise from limited data statistics, signal modelling, and $t \bar{t}$ background modelling. Results are found to be in good agreement with predictions from several MC event generators.

### 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; BMWF and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NERSC and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DNRF and DNSRC, Denmark; IN2P3-CNRS, CEA-DRF/IRFU, France; SRNSFG, Georgia; BMBF, HGF, and MPG, Germany; GSRT, Greece; RGC, Hong Kong SAR, China; ISF, I-CORE and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; NWO, Netherlands; RCN, Norway; MNiSW and NCN, Poland; FCT, Portugal; MNE/IFA, Romania; MES of Russia and NRC KI, Russian Federation; JINR; MESTD, Serbia; MSSR, Slovakia; ARRS and MIZŠ, Slovenia; STFC, Switzerland; MOST, Taiwan; TAEK, Turkey; STFC, United Kingdom; DOE and NSF, United States of America. In addition, individual groups and members have received support from BCKDF, the Canada Council, CANARIE, CRC, Compute Canada, FQRNT, and the Ontario Innovation Trust, Canada; EPLANET, ERC, ERDF, FP7, Horizon 2020 and Marie Skłodowska-Curie Actions, European Union; Investissements d’Avenir Labex and Idex, ANR, Région Auvergne and Fondation Partager le Savoir, France; DFG and AvH Foundation, Germany; Herakleitos, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSRF; BSF, GIF and Minerva, Israel; BRF, Norway; CERCA Programme Generalitat de Catalunya, Generalitat Valenciana, Spain; the Royal Society and Leverhulme Trust, United Kingdom. The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN, the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (UK) and BNL (USA), the Tier-2 facilities worldwide and large non-WLCG resource providers. Major contributors of computing resources are listed in Ref. [73].

### Open Access

This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution,
(a)Universidade Federal do Rio De Janeiro COPPE/EE/IF, Rio de Janeiro, Brazil; (b)Electrical Circuits Department, Federal University of Juiz de Fora (UFJF), Juiz de Fora, Brazil; (c)Federal University of Sao Joao del Rei (UFSJ), Sao Joao del Rei, Brazil; (d)Instituto de Fisica, Universidade de Sao Paulo, Sao Paulo, Brazil

Physica Department, Brookhaven National Laboratory, Upton, NY, USA

(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)Physics Department, National Institute for Research and Development of Isotopic and Molecular Technologies, Cluj Napoca, Romania; (e)University Politehnica Bucharest, Bucharest, Romania; (f)West University in Timisoara, Timisoara, Romania

Departamento de Física, Universidad de Buenos Aires, Buenos Aires, Argentina

Cavendish Laboratory, University of Cambridge, Cambridge, UK

Department of Physics, Carleton University, Ottawa, ON, Canada

CERN, Geneva, Switzerland

Enrico Fermi Institute, University of Chicago, Chicago, IL, USA

(a)Departamento de Física, Pontificia Universidad Católica de Chile, Santiago, Chile; (b)Departamento de Física, Universidad Técnica Federico Santa María, Valparaiso, Chile

(a)Institute of High Energy Physics, Chinese Academy of Sciences, Beijing, China; (b)Department of Physics, Nanjing University, Nanjing, Jiangsu, China; (c)Physics Department, Tsinghua University, Beijing 100084, China; (d)University of Chinese Academy of Science (UCAS), Beijing, China

(a)Department of Modern Physics and Key Laboratory of Particle Detection and Electronics, University of Science and Technology of China, Anhui, China; (b)School of Physics, Shandong University, Shandong, China; (c)Department of Physics and Astronomy, Key Laboratory for Particle Physics, Astrophysics and Cosmology, Ministry of Education, Shanghai Key Laboratory for Particle Physics and Cosmology, Shanghai Jiao Tong University, Tsung-Dao Lee Institute, Shanghai, China

Université Clermont Auvergne, CNRS/IN2P3, LPC, Clermont-Ferrand, France

Nevis Laboratory, Columbia University, Irvington, NY, USA

Niels Bohr Institute, University of Copenhagen, Copenhagen, Denmark

(a)INFN Gruppo Collegato di Cosenza, Laboratori Nazionali di Frascati, Frascati, Italy; (b)Dipartimento di Fisica, Università della Calabria, Rende, Italy

(a)Faculty of Physics and Applied Computer Science, AGH University of Science and Technology, Kraków, Poland; (b)Marian Smoluchowski Institute of Physics, Jagiellonian University, Kraków, Poland

Institute of Nuclear Physics, Polish Academy of Sciences, Kraków, Poland

Physics Department, Southern Methodist University, Dallas, TX, USA

Physics Department, University of Texas at Dallas, Richardson, TX, USA

DESY, Hamburg and Zeuthen, Germany

Lehrstuhl für Experimentelle Physik IV, Technische Universität Dortmund, Dortmund, Germany

Institut für Kern- und Teilchenphysik, Technische Universität Dresden, Dresden, Germany

Department of Physics, Duke University, Durham, NC, USA

SUPA-School of Physics and Astronomy, University of Edinburgh, Edinburgh, UK

INFN e Laboratori Nazionali di Frascati, Frascati, Italy

(a)E. Andronikashvili Institute of Physics, Iv. Javakhishvili Tbilisi State University, Tbilisi, Georgia; (b)High Energy Physics Institute, Tbilisi State University, Tbilisi, Georgia

II Physikalisches Institut, Justus-Liebig-Universität Giessen, Giessen, Germany

SUPA-School of Physics and Astronomy, University of Glasgow, Glasgow, UK

II Physikalisches Institut, Georg-August-Universität, Göttingen, Germany

Laboratoire de Physique Subatomique et de Cosmologie, Université Grenoble-Alpes, CNRS/IN2P3, Grenoble, France

Laboratory for Particle Physics and Cosmology, Harvard University, Cambridge, MA, USA