Not reviewed, for internal circulation only

ATLAS Paper Draft

ANA-TOPQ-2016-06
Version 1.0
Target journal: Eur. Phys. J. C

Comments are due by: 23rd June 2019

Supporting internal notes

tW cross-section measurement in the lepton+jets channel at $\sqrt{s} = 8$ TeV:
https://cds.cern.ch/record/2133344

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

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 requiring one isolated lepton 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. The CKM matrix element multiplied by a form factor is also extracted: $f_{VL} \cdot |V_{tb}| = 1.08 \pm 0.15$.

Document created on 13th June 2019 using ATLAS LaTeX Version 07-02-00.

Analysis Team

[*email:* atlas-ANA-TOPQ-2016-06-editors@cern.ch]

Ian Brock, Sebastian Mergelmeyer, Regina Moles-Valls

Not reviewed, for internal circulation only

ATLAS Paper

ANA-TOPQ-2016-06

13th June 2019

Journal: Eur. Phys. J. C

Draft version 1.0

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

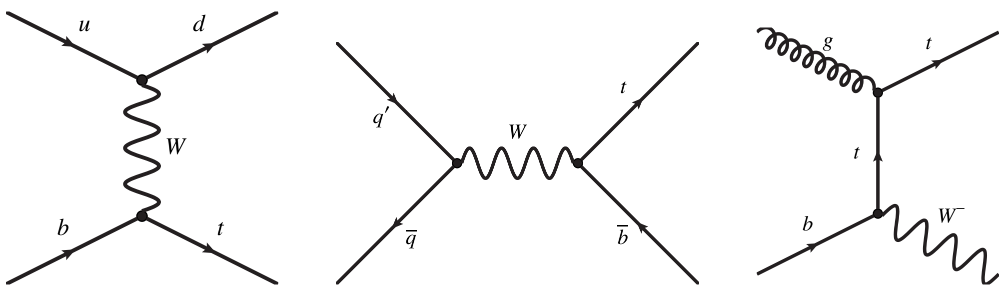
The ATLAS Collaboration

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 requiring one isolated lepton 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. The CKM matrix element multiplied by a form factor is also extracted: $f_{VL} \cdot |V_{tb}| = 1.08 \pm 0.15$.

Contents

1 Introduction

Single top quarks are produced via weak interactions through three differerent channels which are defined by the virtuality of the W boson involved: t-channel, s-channel or W-boson production in association with a top quark, called tW production. These processes, shown in Figure 1, involve a Wtb vertex at leading order (LO) in the Standard Model (SM). 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-section predicted by the SM can originate from single top quarks produced in the decays of unknown heavy particles predicted by new physics models, such as vector-like or excited quarks [1, 2] or superpartners of the top quark as predicted by supersymmetry [3]. 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 [4]. 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 [5]. As each of the single-top-quark processes is sensitive to different sources of new physics, is it important to study each channel separately. In addition, the SM production of tW is an important background to direct searches for particles beyond the SM [6, 7].



(a) t-channel (b) s-channel (c) tW production

Figure 1: LO Feynman diagrams of single-top-quark production: (a) t-channel, (b) s-channel and (c) tW production.

At the Large Hadron Collider (LHC), evidence for the tW production process was found by the ATLAS [8] and CMS Collaborations [9] at $\sqrt{s} = 7$ TeV and the process was observed by both experiments [10, 11] at $\sqrt{s} = 8$ TeV. The tW cross-section has been also measured with 13 TeV collision data inclusively by the CMS Collaboration [12] and inclusively and differentially by the ATLAS Collaboration [13, 14]. These measurements were done 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 contains a W boson in addition to a top quark, which decays predominantly 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 ($E_{\mathrm{T}}^{\mathrm{miss}}$), two light jets with high transverse momentum (p_{T}), and one jet identified as containing a b-hadron (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+ jets and $t\bar{t}$ events, where the latter poses a major challenge to this measurement because of its similar kinematics, and a 10 times larger cross-section. An artificial neural network (NN) is trained to separate the signal from the $t\bar{t}$ background. The cross-section is extracted using a binned profile maximum-likelihood fit of a two-dimensional discriminant. The discriminant combines the neural-network response with the reconstructed invariant mass of the hadronic W-boson decay, $m(W_{\mathrm{H}})$, allowing the fit to disentangle better the signal from the large background and constrain systematic effects. This measurement, performed for the first time with tW single-lepton events, constitutes a cross-check of the previous results published in the dilepton channel as well as a starting point for future differential cross-section measurements as a function of the reconstructed top-quark kinematics.

2 ATLAS detector

The ATLAS experiment [15] at the LHC is a multi-purpose particle detector with a forward-backward symmetric cylindrical geometry and a near 4π coverage in solid angle[1]. It consists of an inner tracking detector (ID) surrounded by a thin superconducting solenoid providing a 2 T axial magnetic field, electromagnetic and hadron calorimeters, and a muon spectrometer (MS). The ID provides a 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. A hadronic (iron/scintillator-tile) calorimeter covers the central pseudorapidity range ($|\eta| < 1.7$). The end-cap ($1.5 < |\eta| < 3.2$) and forward ($3.1 < |\eta| < 4.9$) regions are instrumented with LAr calorimeters for both EM and hadronic energy measurements. The MS surrounds the calorimeters and includes a system of precision tracking chambers and fast detectors for triggering. The magnet system for the MS consists of three large air-core toroid magnets with eight superconducting coils. The field integral of the toroids ranges between 2.0 and 7.5 T m across most of the detector. Collisions with a potential physics interest are captured with the trigger system. For the data taken at $\sqrt{s} = 8$ TeV, a three-level trigger system is used to select events.The first-level trigger is implemented in hardware and uses a subset of the detector information to reduce the accepted rate to at most 75 kHz. This is 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 Monte Carlo samples

Only pp data periods at $\sqrt{s} = 8$ TeV taken with stable LHC beams and the ATLAS detector fully operational are considered, corresponding to an integrated luminosity of 20.2 fb^{-1}.

Monte Carlo samples are produced using the full ATLAS detector simulation [16] implemented in Geant 4 [17]. In addition, alternative MC samples, used to train the neural network and evaluate systematic uncertainties, are produced using AtlFast2 [18], that provides a faster calorimeter simulation making use of parameterized showers to compute the energy deposited by the particles. Pile-up (additional pp interactions in the same or nearby bunch crossing) are modelled by overlaying simulated minimum-bias events generated with Pythia 8 [19]. Weights are assigned to the simulated events, such that the distribution of the number of pile-up interactions in the simulation matches the corresponding distribution in the data, which has an average of 21 [20].

[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, ϕ) are used in the transverse plane, ϕ being the azimuthal angle around the z-axis. The pseudorapidity is defined in terms of the polar angle θ as $\eta = -\ln\tan(\theta/2)$.

The tW signal events are simulated using the NLO Powheg method [21, 22] implemented in the Powheg-Box (v.1.0) generator (revision 2192) [23] with the CT10 Parton Distribution Function (PDF) set [24] in the matrix-element calculation. The mass and width of the top-quark are set to $m_t = 172.5$ GeV and $\Gamma = 1.32$ GeV, respectively. The top quark is assumed to decay exclusively into Wb. The parton shower, hadronisation and underlying event are simulated using Pythia 6 (v6.426) [25] with the LO CTEQ6L1 PDF set [26] and the corresponding Perugia 2011 (P2011C) set of tuned parameters [27]. The factorisation scale, μ_F, and renormalisation scale, μ_R, are set to m_t. Calculations involving tW production beyond LO include quantum interference with $t\bar{t}$ production. Double-counting of the contributions is avoided by using the diagram-removal (DR) scheme [28, 29], in which diagrams with a second on-shell top-quark propagator are removed from the amplitude. The SM tW cross-section prediction at next-to-leading order (NLO) including next-to-next-to-leading-log (NNLL) soft gluon corrections is calculated as $\sigma^{\text{th.}}_{tW}(8\,\text{TeV}) = 22.4 \pm 0.6\,(\text{scale}) \pm 1.4\,(\text{PDF})$ 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 $2m_t$) and the second term covers the uncertainty in the parton distribution function (PDF), evaluated using the MSTW2008 PDF set [30] at next-to-next-to-leading order (NNLO).

The diagram-subtraction (DS) scheme is used as to evaluate the systematic uncertainty associated with the tW–$t\bar{t}$ overlap. In the DS scheme, a subtraction term cancels the $t\bar{t}$ contribution to the cross-section when the top-quark propagator becomes on-shell. The uncertainty associated with the NLO matrix-element generator is estimated by comparing Powheg-Box with MC@NLO (v4.06) [31] both interfaced with Herwig (v6.520) [32]. Parton showering and hadronisation model uncertainties are assessed by comparing Powheg-Box interfaced to Herwig (instead of Pythia 6). For the Herwig samples, the AUET2 tune [33] with the CT10 PDF is used and the underlying event is generated with Jimmy (v4.31) [34]. Uncertainties associated with different μ_R and μ_F scales are evaluated using Powheg-Box interfaced with Pythia 6 (v6.6427) samples, by varying the scales simultaneously in the matrix element and in the parton shower. In these samples the variation of both, μ_R and μ_F, by a factor of 0.5 is combined with a Perugia 2012radHi tune, while the variation of the scale parameters by a factor of 2.0 is combined with the Perugia 2012radLo tune.

The $t\bar{t}$ sample is generated with Powheg-Box (v1.1) interfaced with Pythia 6 (v.6427). In the Powheg-Box event generator, the CT10 PDFs are used, while the CTEQ6L1 PDFs are used for Pythia. The h_{damp} parameter, which effectively regulates the high-p_T gluon radiation, is set to m_t. The predicted $t\bar{t}$ production cross-section is $\sigma_{t\bar{t}}(8\,\text{TeV}) = 252.9^{+6.4}_{-8.6}\,(\text{scale}) \pm 11.7\,(\text{PDF}+\alpha_S)$ pb, calculated with the Top++2.0 program to NNLO in perturbative QCD, including soft-gluon resummation to NNLL [35]. The first uncertainty comes from the quadratic sum of the independent variation of μ_R and μ_F. The uncertainty associated with variations in the PDFs and strong coupling constant, α_S, is evaluated following the PDF4LHC NLO prescription [36, 37], which defines the central value as the midpoint of the uncertainty envelope of three PDF sets: MSTW2008 NNLO [30], CT10 NNLO [38] and NNPDF2.3 5f FFN [39]. The same procedures as for the tW samples are employed to determine the uncertainties due to the NLO matching method and the parton shower and hadronisation. Samples to evaluate the scale uncertainties are produced in a similar way, varying the μ_R and μ_F together with the Perugia tune, but adding also variations in the h_{damp} parameter (for the up-variation, h_{damp} is changed to $2m_t$, while for the down variation it is kept at m_t).

The other single-top-quark production processes, s-channel and t-channel, are also generated with Powheg-Box (v1.1) coupled to Pythia 6 (v.6426), using the same PDF sets as described for the other top-quark processes previously. The predicted cross-sections at $\sqrt{s}$ =8 TeV calculated at NLO plus NNLL are 5.6 ± 0.2 pb for the s-channel [40], and $87.8^{+3.4}_{-1.9}$ pb for the t-channel [41].

Vector-boson production in association with jets is simulated using the multi-leg LO generator Sherpa (v1.4.1) [42–44] and the CT10 PDF sets. Sherpa is 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 and samples with associated heavy-quark pair production is avoided consistently by using massive c- and b-quarks in the shower. The predicted NNLO W+jets cross-section with W decaying leptonically is $\sigma(pp \to \ell^{\pm}\nu_{\ell}X) = 36.3 \pm 1.9$ nb [45]. For Z+jets the cross-section calculated at NNLO in QCD for leptonic Z decays is: $\sigma(pp \to \ell^{+}\ell^{-}X) = 3.72 \pm 0.19$ nb [45]. The AtlFast2 simulation is used to generate these samples with sufficient statistics. Diboson samples are generated with Herwig at LO QCD and CTEQ6L1 PDF. The theoretical NLO cross-section for events with one lepton is 29.4 pb [45].

Multijet events are selected in the analysis when they contain jets or photons misidentified as leptons or non-prompt leptons from hadron decays (both referred to as a ‘fake’ lepton). This background is estimated directly from data using the matrix method [46], which exploits differences in lepton identification and isolation properties between prompt and non-prompt leptons. The shape and normalisation of the multijet background are determined in both electron and muon channels.

4 Object definitions

In the interaction region, primary vertex (PV) candidates are reconstructed from at least five tracks that satisfy $p_{\mathrm{T}} > 400$ MeV. The candidate with the highest sum of p_{T}^2 over all associated tracks is chosen as the hard-collision PV [47].

Muon candidates are reconstructed by matching segments or tracks in the MS with tracks found in the ID [48]. The candidates must have $p_{\mathrm{T}} > 25$ GeV and be in a pseudorapidity range $|\eta| < 2.5$. The longitudinal impact parameter of the track with respect to the hard-collision PV, $|z_{\mathrm{vtx}}|$, is required to be smaller than 2 mm. In order to reject non-prompt muons, an isolation criterion is applied. The isolation variable, defined as the scalar sum of the transverse momenta of all tracks with $p_{\mathrm{T}} > 1$ GeV (excluding the muon track) within a cone of size $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 10\,\mathrm{GeV}/p_{\mathrm{T}}(\mu)$, is required to be lower than 0.05. The selection efficiency after this requirement is measured to be about 97 % in $Z \to \mu^{+}\mu^{-}$ events.

Electron candidates are selected from energy deposits (clusters) in the EM calorimeter, which match a well-reconstructed track in the ID [49]. Requirements on the transverse and longitudinal impact parameter of $|d_{\mathrm{vtx}}| < 1$ mm and $|z_{\mathrm{vtx}}| < 2$ mm, respectively, are applied. Electron candidates must have $E_{\mathrm{T}} > 25$ GeV and $\left|\eta_{\mathrm{cluster}}\right| < 2.47$, where η_{cluster} denotes the pseudorapidity of the cluster. Clusters in the calorimeter barrel–endcap transition region, $1.37 < |\eta| < 1.52$, are excluded. An isolation requirement based on the deposited transverse energy in a cone of size $\Delta R < 0.2$ around the direction of the electron and the p_{T} sum of the tracks in a cone with $\Delta R < 0.3$ around the same direction is applied. This requirement is chosen to give a nearly uniform selection efficiency of 85 % in p_{T} and η, as measured in $Z \to e^{+}e^{-}$ events. Electron candidates that share the ID track with a reconstructed muon candidate are vetoed.

Jets are reconstructed using the anti-k_t algorithm [50] with a radius parameter of $R = 0.4$ using topological clusters [51], calibrated with the Local Cluster Weighting 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) [53] variable is used to identify the primary vertex from which the jet originated. The JVF criterion applied supress pile-up jets with $p_{\mathrm{T}} < 50$ GeV and $|\eta| < 2.4$. To avoid possible overlap between jets and electrons, jets that are close to an electron within a cone of size $\Delta R < 0.2$ are removed. Afterwards, remaining electron candidates overlapping with jets within a distance of $\Delta R < 0.4$ are rejected. Also an overlap removal between muons and jets is applied; in this case the 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 inside b-jets as well as the properties of the b-quark fragmentation. The output of these algorithms are combined in a neural network classifier to maximize the b-tagging performance [54] . 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 of about 5 and the light-quark-jet rejection factor is of about 120. These efficiencies and rejection factors were obtained using $t\bar{t}$ events.

The $E_{\mathrm{T}}^{\mathrm{miss}}$ of the event is defined as the momentum imbalance in the plane transverse to the beam axis, primarily due to neutrinos that escape detection. It is calculated as the negative vector sum of the momenta of the reconstructed electrons, muons and jets as well as any clusters that are not associated with any of the previous objects [55].

5 Event selection

Events are required to have a hard-collision PV. They also have to pass a single-lepton trigger [56, 57] and to contain at least one electron or muon candidate with $p_{\mathrm{T}} > 30$ GeV matched to the lepton selected by the trigger. The electron trigger requires an electron candidate, formed by an EM calorimeter cluster matched with a track, with $E_{\mathrm{T}} > 60$ GeV or $E_{\mathrm{T}} > 24$ GeV and additional isolation requirements. The muon trigger requires a muon candidate, defined as a reconstructed track in the muon spectrometer, with $p_{\mathrm{T}} > 36$ GeV or $p_{\mathrm{T}} > 24$ 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 will be referred to as the lepton.

Events identified as containing jets from cosmic rays, beam-induced backgrounds or due to noise hot spots in the calorimeter are removed. Only jets with $p_{\mathrm{T}} > 30$ GeV and $|\eta| < 2.4$ are considered in the analysis. Additionally, requirements on the $E_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV and the transverse mass[2] of the leptonically decaying W boson, $m_{\mathrm{T}}(W_L) > 50$ GeV, are applied.

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 and exactly one b-tagged jet (`3j1b`) with the best signal-to-background ratio is denoted the signal region and is used to extract the tW cross-section. Table 1 shows the expected number 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 and the major backgrounds are $t\bar{t}$ production with about 58 %, and W+jets production with about 30 % of the total number of events. The W+jets contribution is mostly composed of events in which a W-boson is produced in association with heavy flavour (HF) jets (b- and c-jets). The total numbers of expected events agree within a few percent with the observed number of events shown in Table 3. Nearby region with four jets,

[2] The transverse mass is calculated using the momentum of the lepton associated with the W boson, $E_{\mathrm{T}}^{\mathrm{miss}}$ and the azimuthal angular difference between the two: $m_{\mathrm{T}}(W_L) = m_{\mathrm{T}}(\ell\nu) = \sqrt{2p_{\mathrm{T}}(\ell) \cdot E_{\mathrm{T}}^{\mathrm{miss}}\left[1 - \cos(\Delta\phi(\ell, E_{\mathrm{T}}^{\mathrm{miss}}))\right]}$.

two of them b-tagged (`4j2b`), contains a very pure sample of $t\bar{t}$ events and is used to check the modelling of this background (Section 9).

Process	Signal region (`3j1b`)
tW ($\sigma_{tW}^{\text{th.}} = 22.4$ pb)	6300 ± 600
$t\bar{t}$	77000 ± 6000
t, t-channel	4180 ± 290
t, s-channel	307 ± 19
W+jets, HF	31000 ± 14000
W+jets, other	6000 ± 3000
Z+jets	3900 ± 1700
$WW/WZ/ZZ$+jets	650 ± 280
Fake leptons	4300 ± 1900
Total background	128000 ± 18000
Total signal + background	134000 ± 18000
Observed	134633

Table 1: Expected signal and background and observed number of events in the signal (`3j1b`) region. The cross-section for tW production has been fixed to the theory prediction. The uncertainties include statistical and systematic uncertainties.

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 NN (implemented in the NeuroBayes framework [58, 59]). More detailed information about how the NN is used in single-top-quark analyses can be found in Ref. [60]. The NN variables are selected such that they significantly contribute to the statistical separation power between signal and background, while avoiding variables that would lead to an increase of the expected systematic uncertainty. In order to study the impact on the systematic uncertainty the complete analysis chain is rerun for any list of NN input variables. The observable $m(W_{\text{H}})$ provides a 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, the $m(W_{\text{H}})$ is not used in the NN but a two-dimensional discriminant is constructed using the response of a NN and $m(W_{\text{H}})$. This procedure, explained in more detail in the following subsections, allows the uncertainties affecting the variable $m(W_{\text{H}})$ to be (partially) absorbed into nuisance parameters.

6.1 Invariant mass of the hadronically decaying W boson

The variable $m(W_{\text{H}})$ is computed from the four-vectors of the two selected untagged jets. For the signal and the $t\bar{t}$ background, the distribution of $m(W_{\text{H}})$ exhibits a peak near the mass of the W boson, shown in Figure 2(a). The peak results from events where the two untagged jets are correctly associated 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. For the shape comparison plots in Figure 2(a) and Figure 6(a), only the $W+$jets HF contribution is shown, which corresponds approximately to 80% of the total $W+$jets background in the signal region. Figure 2(b) shows the pre-fit distribution of $m(W_\mathrm{H})$, and also demonstrates good pre-fit modelling of the data.

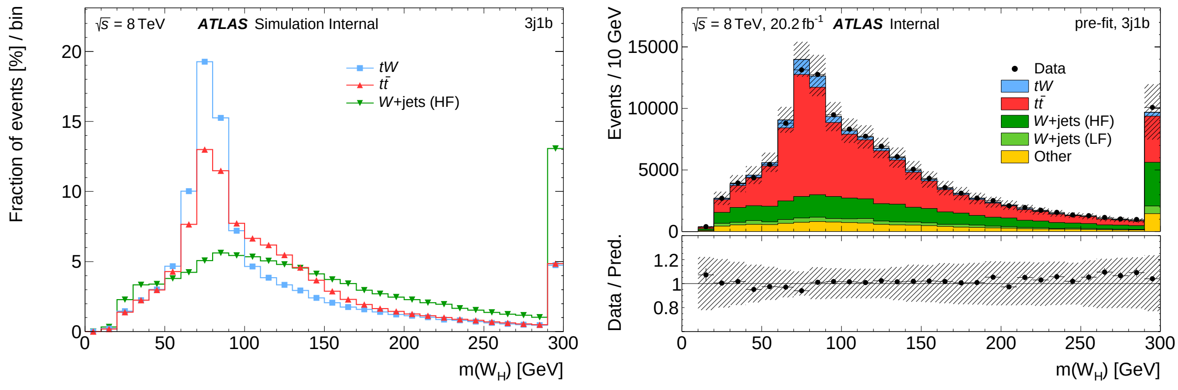


Figure 2: (a) Shape of the reconstructed $m(W_\mathrm{H})$ distribution for signal and most important backgrounds in the signal (`3j1b`) region. The distribution for each process normalised to unity is shown.(b) Pre-fit $m(W_\mathrm{H})$ distribution in the `3j1b` region. Small backgrounds are subsumed under 'Other'. The simulated distributions are normalised to their theoretical cross-sections. The dashed uncertainty band includes statistical and systematic uncertainties. The last bin includes the overflow events. The lower panel shows the ratio of the observed and the expected number of events in each bin.

6.2 Neural network

The NN is trained using simulated events with a well-reconstructed hadronic W-boson decay. That means the two reconstructed untagged jets are matched within $\Delta R < 0.35$ to the truth jets coming from a W-boson decay in the MC simulation. Such events are expected to have a reconstructed W-boson mass close to the PDG mass, so a requirement of $65\,\mathrm{GeV} < m(W_\mathrm{H}) < 92.5\,\mathrm{GeV}$ is applied. Since the $W+$jets events and other background events cannot have a well-reconstructed hadronic W-boson decay, the background sample used for the training consists entirely of $t\bar{t}$ events, where a tiny contribution from diboson production has been neglected. Therefore, the network is trained only against $t\bar{t}$ production. 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, $p_\mathrm{T}(W_\mathrm{H}W_\mathrm{L}j_\mathrm{B})$, divided by the sum of the object transverse momenta
$$\rho_\mathrm{T}(W_H, W_L, j_\mathrm{B})^{3} = \frac{p_\mathrm{T}(W_\mathrm{H}W_\mathrm{L}j_\mathrm{B})}{p_\mathrm{T}(W_\mathrm{H}) + p_\mathrm{T}(W_\mathrm{L}) + p_\mathrm{T}(j_\mathrm{B})};$$
- the invariant mass of the reconstructed tW system, $m(W_\mathrm{L}W_\mathrm{H}j_\mathrm{B})$;

[3] The use of $\rho_\mathrm{T}(W_H, W_L, j_\mathrm{B})$, 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 absolute value of the difference between the pseudorapidities of the lepton and the leading untagged jet, $|\Delta\eta(\ell, j_{L1})|$;
- the absolute value of the pseudorapidity of the lepton, $|\eta(\ell)|$.

Figure 3 compares the data to the prediction for the NN input variables. Good modelling of the variables was also confirmed in the $t\bar{t}$ validation region.

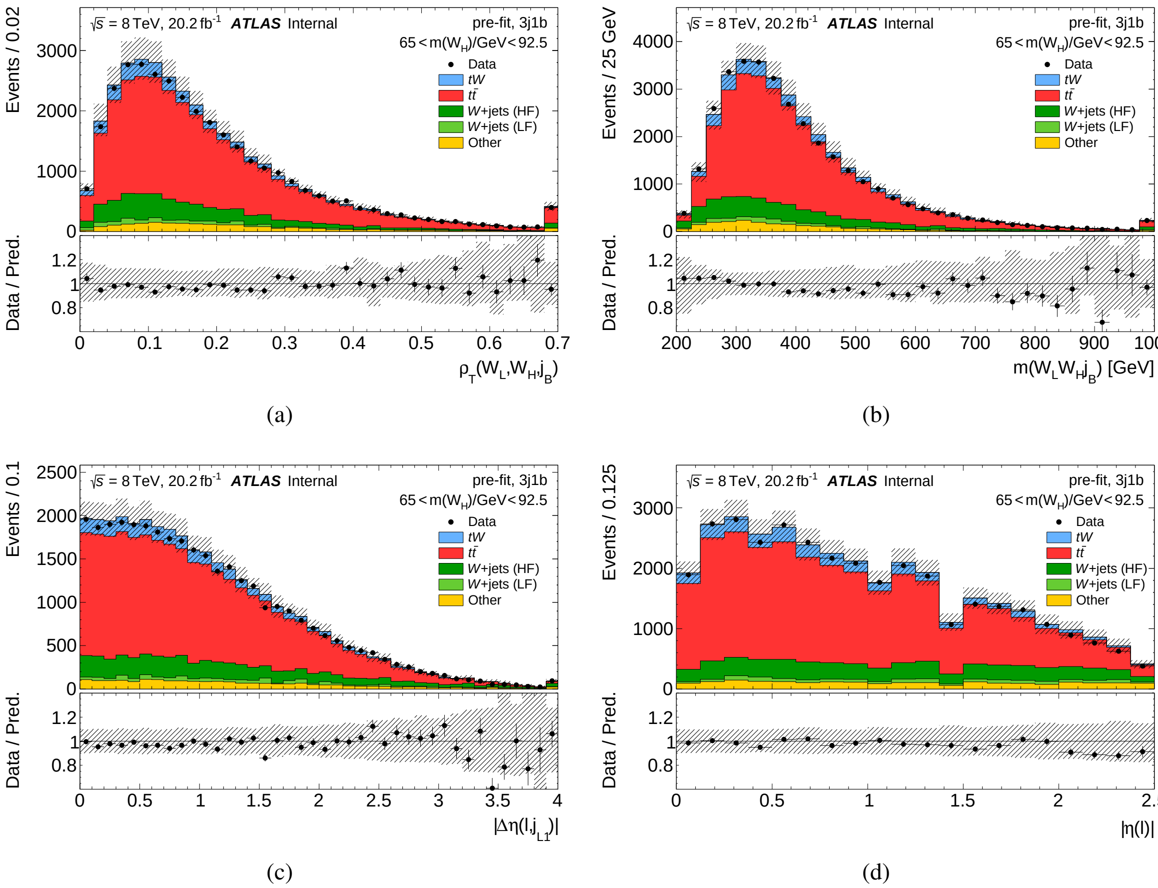
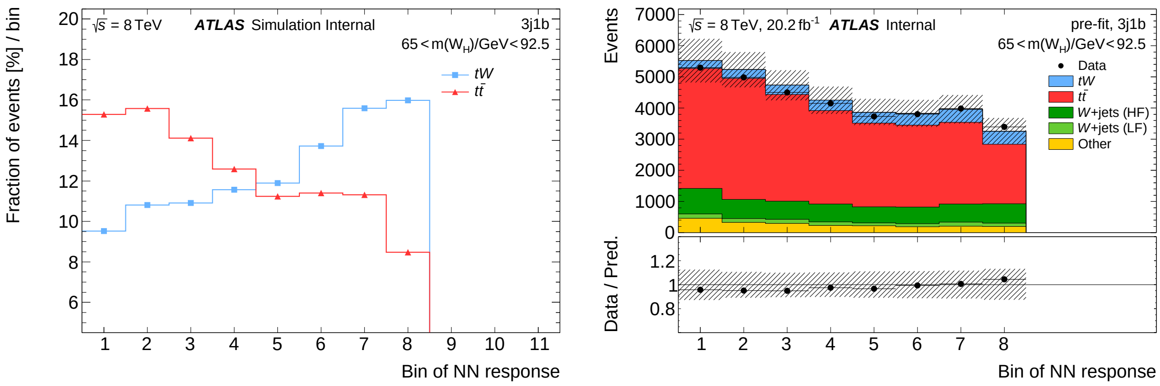


Figure 3: Pre-fit distributions of the NN input variables in the tW signal (3j1b) region with 65 GeV $\leq m(W_{\mathrm{H}}) \leq$ 92.5 GeV. Small backgrounds are subsumed under 'Other'. The simulated distributions are normalised to their theoretical cross-sections. The dashed uncertainty band includes statistical and systematic uncertainties. The last bin includes the overflow events. The lower panels show the ratio of the observed and the expected number of events in each bin.

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 4(a) shows the shape of the NN response for the tW and $t\bar{t}$ processes and Figure 4(b) presents the comparison between data and Monte Carlo.



(a) (b)

Figure 4: (a) Shape of the NN response in the signal (3j1b) region. The distribution contains those events with 65 GeV $\leq m(W_{\mathrm{H}}) \leq$ 92.5 GeV. The distribution for the tW process and the $t\bar{t}$ process normalised to unity is 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 last bin includes the overflow events. The lower panel shows the ratio of the observed and the expected number of events in each bin.

6.3 Two-dimensional discriminant

For the two-dimensional discriminant, $m(W_{\mathrm{H}})$ is used on the first and the NN response on the second axis of the two-dimensional discriminant. Outside of the aforementioned $m(W_{\mathrm{H}})$ range from 65 GeV 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 Figure 5.

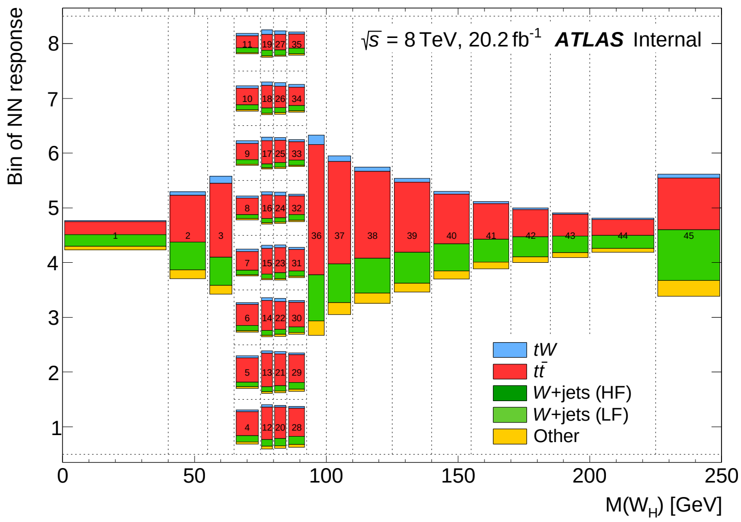


Figure 5: Predicted distribution of the two-dimensional discriminant in the signal (3j1b) region. The proportions of the coloured areas reflect the expected composition in terms of tW, $t\bar{t}$, $W+$jets and other processes. The numbers correspond to the bin order when projecting the discriminant on one axis as in Figure 6.

The bins are then rearranged on a one-dimensional axis in row-major order. The resulting one-dimensional distribution is presented in Figure 6, together with a comparison of the shapes. The first three bins and the last ten bins correspond directly to the bins of $m(W_{\text{H}})$ 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 $m(W_{\text{H}})$. 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 [52] is evaluated by varying the reconstructed energies of the jets in the simulated samples. It is split in multiple components, taking into account the uncertainty in the calorimeter response, the detector simulation, the choice of the 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 η regions of the detector, the difference between data and MC simulation, as well as the noise contribution in the forward detector region [61]. The uncertainty in jet reconstruction efficiency is estimated by randomly dropping simulated jets from the events according to the jet reconstruction inefficiency measured with di-jet events [52]. The JVF uncertainty is evaluated by varying the JVF criterion [53].

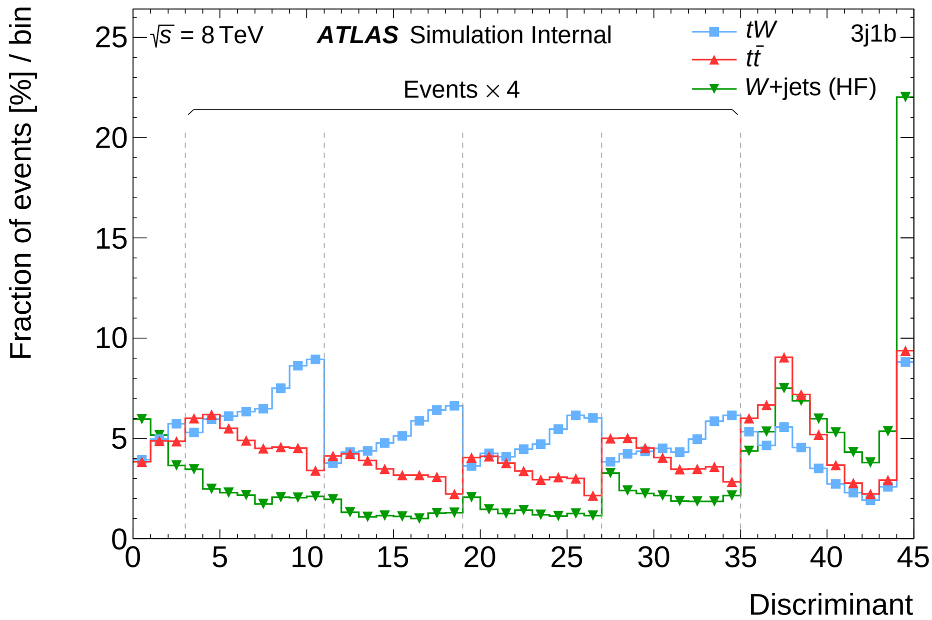
Uncertainties in the scale factors to correct the b-tagging efficiency in simulation to the efficiency in data are varied separately for b-jet, c-jet and light-flavour jets. Several methods are developed to measure the b-tagging efficiency, c-tagging efficiency and mistag rate using 8 TeV data [54, 62, 63]. Independent sources of uncertainty affecting the b-jet tagging efficiency and c-jet mis-tagging efficiency are considered depending on the jet kinematics, i.e. 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 [48, 49].

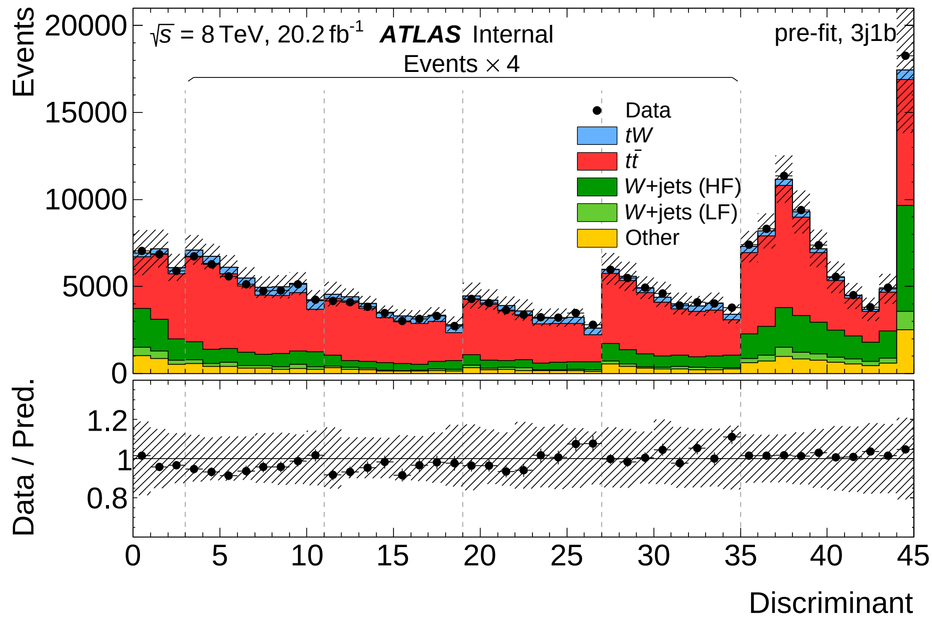
All systematic uncertainties in the reconstruction of jets and leptons are propagated to the $E_{\text{T}}^{\text{miss}}$. In addition, uncertainties in the soft terms of the $E_{\text{T}}^{\text{miss}}$, which account for energy deposits in the calorimeter which are not matched to high-p_{T} physics objects [55].

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. [20]. 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. This uncertainty is considered uncorrelated between the tW and the $t\bar{t}$ 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 systematics are computed by comparing Powheg-Box with either Pythia or Herwig. These ones are treated as fully correlated between the tW and the $t\bar{t}$ processes. The uncertainty due to the treatment of the interference effects of the tW and the $t\bar{t}$ processes is evaluated by using the tW DS instead of the DR scheme, both generated with Powheg-Box with Pythia. The effect of the PDF uncertainties on the acceptance is taken into account for both, the tW signal and the $t\bar{t}$ background and treated as uncorrelated between the processes, following the studies in Ref. [64].



(a)



(b)

Figure 6: (a) Shape distribution of the reconstructed discriminant in the tW signal (3j1b) region rearranged onto a one-dimensional distribution. The distribution for each process normalised to unity is shown.(b) Pre-fit distributions of the discriminant in the tW signal (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 last bin includes the overflow events. The lower panel shows the ratio of the observed and the expected number of events in each bin.The first three bins and the last ten bins correspond directly to (non-uniform) bins of $m(W_{\mathrm{H}})$. In between are four blocks of eight bins, corresponding to the NN output in slices of $m(W_{\mathrm{H}})$. Inside each of the blocks, the numbers of events are scaled by a factor of four for better visibility.

The uncertainties in the theoretical cross-sections are process dependent and vary from 4 % for the t-channel to 6 % for $t\bar{t}$, see Section 3. In addition, there are large uncertainties on the Z/W + jets production cross-sections. For every jet an additional uncertainty of 24 % is assumed [65]. The uncertainty in the normalisation of W/Z-boson production in association with three jets is 42 %, and in addition, the rate of W-boson events with heavy-flavour jets is allowed to vary by 20 %.

No dedicated modelling uncertainties are used for the W+ jets background, since its large uncertainty in the normalisation and flavour fraction dominate over the uncertainties due to the choice of generator and scale variations.

The uncertainty due to the limited size of the simulated samples is estimated by applying the Barlow–Beeston light treatment [66, 67]: for every bin of the discriminant, an independent parameter is assigned which describes the variation of the predicted event rate by its statistical uncertainty.

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 efficiency, the choice of the parametrisation and the normalisation of the prompt-lepton backgrounds in the determination of the efficiencies [46].

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:

$$L(\mu, \theta; \vec{n}) = \prod_{i}^{\text{bins}} \text{Pois}(n_i; \nu_i(\mu, \theta))\, G(\theta; 0, 1)\,,$$

where the n_i (ν_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, μ, which is a multiplicative factor on the predicted signal cross-section. Nuisance parameters, θ, 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 $\hat{\mu}$, $\hat{\theta}$, are obtained by minimising $-2\log L(\mu, \theta; \vec{n})$.

The likelihood function is composed and evaluated with the HistFactory program [68], part of the RooStats framework [69]. The minimisation is performed with the Minuit package [70], using Minos to compute the error estimates.

The statistical significance, Z, of the result is estimated by comparing two hypotheses: the background-only hypothesis, which states that the signal does not exist (or equivalently, $\mu = 0$) and the observed data can be explained using only the background processes; and the signal-plus-background hypothesis, using the fitted signal strength. With the so-called asymptotic approximation [71], the significance is calculated using a test statistic based on the profile likelihood ratio:

$$Z^2 = -2 \log \frac{L(\mu = 0, \theta = \hat{\theta}_0)}{L(\mu = \hat{\mu}, \theta = \hat{\theta})}\,, \tag{1}$$

where $\hat{\theta}_0$ denotes the estimates of the nuisance parameters that maximise the likelihood function under the background-only hypothesis. The expected significance is calculated by replacing $\vec{n}$ in the likelihood function with the Asimov dataset for the nominal signal-plus-background hypothesis ($\mu = 1$, $\theta = 0$).

9 Cross-section measurement

The tW cross-section is extracted from the fit to data in the signal region. The measured signal strength is $\hat{\mu} = 1.16 \pm 0.31$, consistent with the expected value $\mu = 1.0^{+0.36}_{-0.33}$. This signal strength corresponds to an observed cross-section of $\sigma_{tW}^{\text{obs}} = 26 \pm 7$ pb and an observed (expected) significance of 4.5σ (4.1σ).

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 sigma 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 very good approximation. Table 2 shows the impacts of the systematic uncertainties on the observed fit result, where the impacts of similar uncertainties have been added in quadrature. The dominant uncertainties are due to the amount of QCD radiation in signal and $t\bar{t}$ background, the JES and the model statistics that includes the uncertainty due to the limited MC statistics and the fake-lepton background determination.

Source	Uncertainty [%]
Jet Energy Scale	10
b-tagging	8
Jet Energy Resolution	7
$E_{\text{T}}^{\text{miss}}$ reconstruction	7
Lepton reconstruction	4
Luminosity	3
Jet Vertex Fraction	3
$t\bar{t}$ radiation	10
tW radiation	9
$tW - t\bar{t}$ interference	7
$t\bar{t}$ cross-section normalisation	6
Other background cross-section normalisations	5
tW and $t\bar{t}$ Parton shower	4
tW and $t\bar{t}$ NLO matching	3
PDF	1
Model statistics	11
Data statistics	4
Total	27

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.

Some nuisance parameters are constrained by the data. For example, the normalisation uncertainty for $W+$jets events is reduced, because the assigned initial uncertainty is large and this background can be separated well from tW and $t\bar{t}$ events. By design of the discriminant, combinations of nuisance parameters that shift the peak in the $m(W_{\text{H}})$ distribution are constrained, primarily the JES and choice of renormalisation scale together with the amount of QCD radiation in signal and $t\bar{t}$ background. Also, the NP for the NLO matching for tW and $t\bar{t}$ 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 to 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. One of the parameters corresponding to the b-tagging efficiency, 'B5', 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 relies partially on the $p_{\mathrm{T}}^{\mathrm{rel}}$ method, which operates in a different environment regarding the production mechanism of the b-jets, the pull is reasonable.

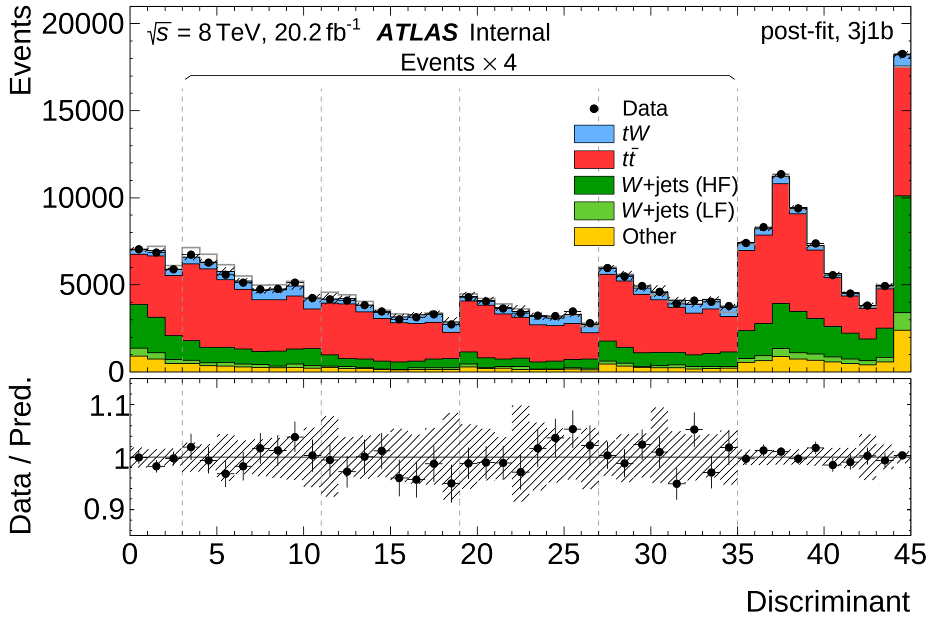
Table 3 shows the post-fit event yields of each process. The post-fit estimates are well within the uncertainties of the pre-fit expectation (Table 1), while most of their uncertainties are reduced. The normalisation uncertainty for W +HF jets changes from almost 50 % to about 10 %. The agreement between the observed number of data events and the prediction calculated using the post-fit values of nuisance parameters in the $t\bar{t}$ validation region indicates a correct estimation of the model parameters determined by the fit in the signal region.

Figures 8 shows the post-fit distributions for the NN input variables, the NN output response and the $m(W_{\mathrm{H}})$ in the signal region. The post-fit plots use the parameter estimates obtained in the fit of the discriminant, including their uncertainties. The distributions demonstrate that the simulation with the updated parameters give a good description of the data.

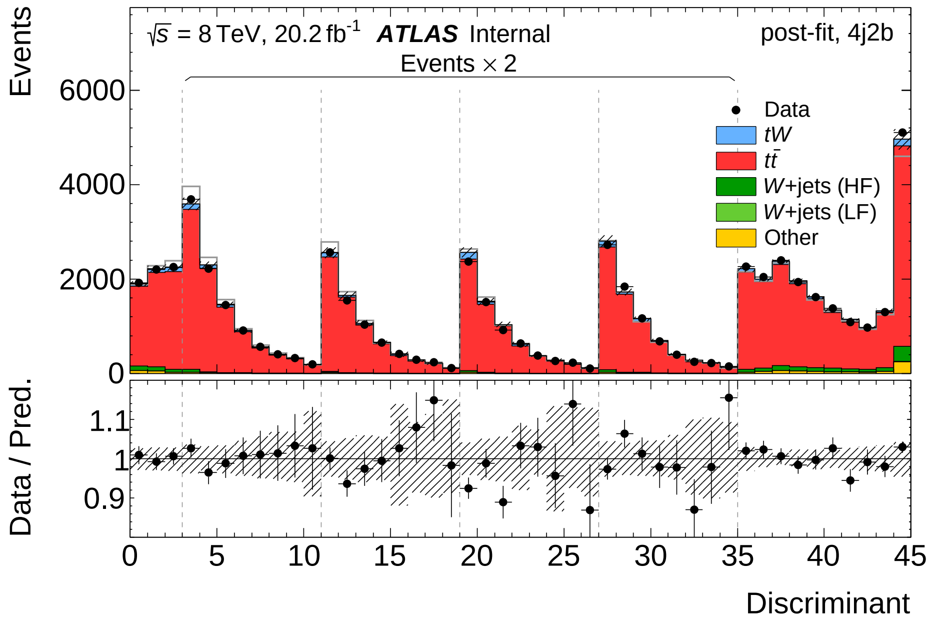
Figure 7 shows the post-fit distributions of the discriminant in the signal and validation region. While 7(a) shows that the data are well described by the model in the signal region, the strongest support for the validity of the fit result comes from the comparisons of the expected and the observed distributions in the $t\bar{t}$ validation, region 7(b), where the uncertainty due to the extrapolation from the signal region is small, and therefore provides a stringent test that the main background is understood very well.

Process	Signal region (3j1b)	$t\bar{t}$ region (4j2b)
tW	7800 ± 1800	1300 ± 400
$t\bar{t}$	74500 ± 2100	36700 ± 2300
t, t-channel	4250 ± 200	590 ± 40
t, s-channel	315 ± 15	63 ± 4
W+ jets, HF	34700 ± 3300	1400 ± 500
W+ jets, other	5700 ± 1800	27 ± 19
Z + jets	3800 ± 1500	180 ± 90
$WW/WZ/ZZ$+jets	640 ± 270	23 ± 13
Fake leptons	3000 ± 1600	5 ± 22
Total background	126900 ± 1900	38900 ± 2400
Total signal + background	134700 ± 500	40200 ± 2300
Observed	134633	41738

Table 3: Post-fit signal and background and observed number of events in the signal and the $t\bar{t}$ validation region. The uncertainties include statistical plus all systematic uncertainties (cf. Section 7).



(a)



(b)

Figure 7: Post-fit distributions of the discriminant in the signal (7(a)) and validation region (7(b)). Small backgrounds are subsumed under 'Other'. The dashed uncertainty band includes statistical and systematic uncertainties. The lower panels show the ratio of the observed and the expected number of events in each bin.The first three bins and the last ten bins correspond directly to (non-uniform) bins of $m(W_{\mathrm{H}})$. In between are four blocks of eight bins, corresponding to the NN output in slices of $m(W_{\mathrm{H}})$. Inside each of the blocks, the numbers of events are scaled by a factor of four (factor of two in `4j2b`) for better visibility.

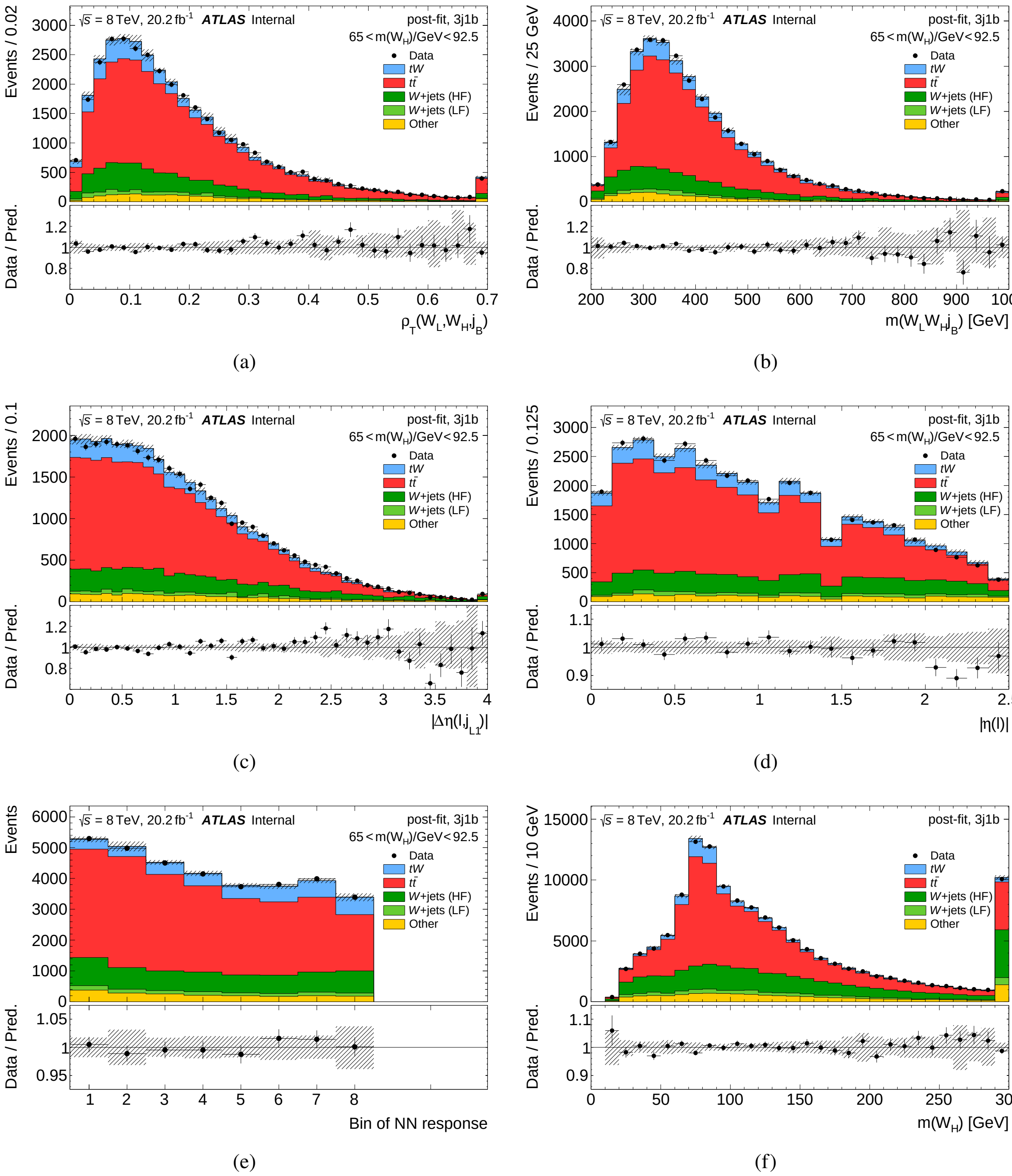


Figure 8: Post-fit distributions of the NN input variables (8(a), 8(b), 8(c), 8(d)), NN discriminant (8(e)) and $m(W_{\mathrm{H}})$ (8(f)) in the signal region. Small backgrounds are subsumed under ‘Other’. The dashed uncertainty band includes statistical and systematic uncertainties. The last bin includes the overflow events. The lower panels show the ratio of the observed and the expected number of events in each bin.

10 Extraction of the CKM matrix element $|V_{tb}|$

The production rate of single-top-quark processes is proportional to the square of the left-handed vector coupling at the Wtb vertex. In the SM, the coupling is given by the CKM matrix element V_{tb}. A direct estimate of this coupling can be extracted from the ratio of the measured single-top-quark cross-section to the theoretical prediction: $|f_{\mathrm{LV}} \cdot V_{tb}|^2 = \sigma_{\mathrm{meas.}}/\sigma_{\mathrm{th.}}$, with f_{LV} being a left-handed form factor. The SM predicts a V_{tb} value close to one and f_{LV} exactly one, but new physics could alter the value of the form factor significantly. Combination of single-top-quark cross-sections and $|f_{\mathrm{LV}} \cdot V_{tb}|$ have been perform by ATLAS and CMS Collaboration using Run 1 data [72].

The measured cross-section can be interpreted in terms of V_{tb} under the following assumptions:

- $|V_{tb}| \gg |V_{ts}|, |V_{td}|$ so the cross-section is proportional to $|V_{tb}|^2$, and no extra hypothesis is needed on the unitarity of the CKM matrix.
- decays of the top quark into particles not described by the SM can be neglected.

Two additional sources of uncertainties enter into the $|V_{tb}|$ calculation: the theoretical uncertainty in the tW cross-section calculated to be 6.8 % [73] at $m_t = 172.5$ GeV, including the variation of the renormalisation and factorisation scales as well as the dependence on the PDFs; and an uncertainty in the theoretical cross-section of 3.4 % due to a variation of the top-quark mass by 1.0 GeV. The uncertainties are added in quadrature to obtain the total uncertainty on the measured cross-section.

The result obtained from the cross-section measured in the present analysis is:

$$|f_{\mathrm{LV}} \cdot V_{tb}| = \sqrt{\frac{\sigma_{tW}^{\mathrm{meas.}}}{\sigma_{tW}^{\mathrm{th.}}}} = 1.08 \pm 0.15\,,$$

in agreement with the SM prediction. Assuming f_{LV}=1, a lower limit on $|V_{tb}|$ is extracted: $|V_{tb}| > 0.84$ at 95 % confidence level.

11 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 data collected by the ATLAS detector at $\sqrt{s} = 8$ TeV in 2012. An NN is used to separate the signal from the $t\bar{t}$ background and a two-dimensional discriminant, built from the NN response and the mass of the hadronically decaying W boson, is used to extract the cross-section. Evidence for the tW production in the single-lepton channel is obtained with an observed (expected) significance of 4.5 (4.1) standard deviations. The measured cross-section is:

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

which is consistent with the SM expectation of $\sigma_{tW}^{\mathrm{th.}} = 22.4 \pm 1.5$ pb. The value of the CKM matrix element $f_{\mathrm{LV}}|V_{tb}|$ is extracted from the measured cross-section: 1.08 ± 0.15.

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; 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 and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; NWO, Netherlands; RCN, Norway; MNiSW and NCN, Poland; FCT, Portugal; MNE/IFA, Romania; MES of Russia and NRC KI, Russian Federation; JINR; MESTD, Serbia; MSSR, Slovakia; ARRS and MIZŠ, Slovenia; DST/NRF, South Africa; MINECO, Spain; SRC and Wallenberg Foundation, Sweden; SERI, SNSF and Cantons of Bern and Geneva, Switzerland; MOST, Taiwan; TAEK, Turkey; STFC, United Kingdom; DOE and NSF, United States of America. In addition, individual groups and members have received support from BCKDF, CANARIE, CRC and Compute Canada, Canada; COST, ERC, ERDF, Horizon 2020, and Marie Skłodowska-Curie Actions, European Union; Investissements d' Avenir Labex and Idex, 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; CERCA Programme Generalitat de Catalunya, 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. [74].

References

[1] J. Nutter, R. Schwienhorst, D. G. E. Walker and J.-H. Yu, *Single Top Production as a Probe of B-prime Quarks*, Phys. Rev. **D86** (2012) 094006, arXiv: 1207.5179 [hep-ph] (cit. on p. 3).

[2] ATLAS Collaboration, *Search for the production of single vector-like and excited quarks in the Wt final state in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, JHEP **02** (2016) 110, arXiv: 1510.02664 [hep-ex] (cit. on p. 3).

[3] ATLAS Collaboration, *Search for top squarks in final states with one isolated lepton, jets, and missing transverse momentum in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector*, Phys. Rev. D **94** (2016) 052009, arXiv: 1606.03903 [hep-ex] (cit. on p. 3).

[4] T. M. P. Tait and C. Yuan, *Single top quark production as a window to physics beyond the standard model*, Phys. Rev. D **63** (2000) 014018, arXiv: hep-ph/0007298 [hep-ph] (cit. on p. 3).

[5] Q.-H. Cao, J. Wudka and C. Yuan, *Search for new physics via single top production at the LHC*, Phys. Lett. B **658** (2007) 50, arXiv: 0704.2809 [hep-ph] (cit. on p. 3).

[6] ATLAS Collaboration, *Search for $B-L$ R-parity-violating top squarks in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS experiment*, Phys. Rev. D **97** (2018) 032003, arXiv: 1710.05544 [hep-ex] (cit. on p. 3).

[7] ATLAS Collaboration, *Search for top-squark pair production in final states with one lepton, jets, and missing transverse momentum using $36\,fb^{-1}$ of $\sqrt{s} = 13$ TeV pp collision data with the ATLAS detector*, JHEP **06** (2018) 108, arXiv: 1711.11520 [hep-ex] (cit. on p. 3).

[8] ATLAS Collaboration, *Evidence for the associated production of a W boson and a top quark in ATLAS at $\sqrt{s} = 7$ TeV*, Phys. Lett. B **716** (2012) 142, arXiv: 1205.5764 [hep-ex] (cit. on p. 3).

[9] CMS Collaboration, *Evidence for associated production of a single top quark and W boson in pp collisions at $\sqrt{s} = 7$ TeV*, Phys. Rev. Lett. **110** (2013) 022003, arXiv: 1209.3489 [hep-ex] (cit. on p. 3).

[10] ATLAS Collaboration, *Measurement of the production cross-section of a single top quark in association with a W boson at 8 TeV with the ATLAS experiment*, JHEP **01** (2016) 064, arXiv: 1510.03752 [hep-ex] (cit. on p. 3).

[11] CMS Collaboration, *Observation of the associated production of a single top quark and a W boson in pp collisions at $\sqrt{s} = 8$ TeV*, Phys. Rev. Lett. **112** (2014) 231802, arXiv: 1401.2942 [hep-ex] (cit. on p. 3).

[12] CMS Collaboration, *Measurement of the production cross section for single top quarks in association with W bosons in proton–proton collisions at $\sqrt{s} = 13$ TeV*, JHEP **10** (2018) 117, arXiv: 1805.07399 [hep-ex] (cit. on p. 3).

[13] ATLAS Collaboration, *Measurement of the cross-section for producing a W boson in association with a single top quark in pp collisions at $\sqrt{s} = 13$ TeV with ATLAS*, JHEP **01** (2018) 063, arXiv: 1612.07231 [hep-ex] (cit. on p. 3).

[14] ATLAS Collaboration, *Measurement of differential cross-sections of a single top quark produced in association with a W boson at $\sqrt{s} = 13$ TeV with ATLAS*, Eur. Phys. J. C **78** (2018) 186, arXiv: 1712.01602 [hep-ex] (cit. on p. 3).

[15] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003 (cit. on p. 4).

[16] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70** (2010) 823, arXiv: 1005.4568 [physics.ins-det] (cit. on p. 4).

[17] S. Agostinelli et al., *GEANT4: A simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250 (cit. on p. 4).

[18] T. Yamanaka, *The ATLAS calorimeter simulation FastCaloSim*, J. Phys. Conf. Ser. **331** (2011) 032053 (cit. on p. 4).

[19] T. Sjöstrand, S. Mrenna and P. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2007) 852, URL: https://cds.cern.ch/record/1064095 (cit. on p. 4).

[20] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC*, Eur. Phys. J. C **76** (2016) 653, arXiv: 1608.03953 [hep-ex] (cit. on pp. 4, 12).

[21] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **11** (2004) 040, arXiv: hep-ph/0409146 [hep-ph] (cit. on p. 5).

[22] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, JHEP **11** (2007) 070, arXiv: 0709.2092 [hep-ph] (cit. on p. 5).

[23] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **06** (2010) 1, ISSN: 1029-8479, arXiv: 1002.2581 [hep-ph] (cit. on p. 5).

[24] H.-L. Lai et al., *New parton distributions for collider physics*, Phys. Rev. D **82** (7 2010) 074024, arXiv: 1007.2241 [hep-ph] (cit. on p. 5).

[25] T. Sjöstrand, S. Mrenna and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026, arXiv: hep-ph/0603175 [hep-ph] (cit. on p. 5).

[26] P. M. Nadolsky et al., *Implications of CTEQ global analysis for collider observables*, Phys. Rev. **D78** (2008) 013004, arXiv: 0802.0007 [hep-ph] (cit. on p. 5).

[27] P. Z. Skands, *Tuning Monte Carlo Generators: The Perugia Tunes*, Phys. Rev. D **82** (2010) 074018, arXiv: 1005.3457 [hep-ph] (cit. on p. 5).

[28] S. Frixione, E. Laenen, P. Motylinski, B. R. Webber and C. D. White, *Single-top hadroproduction in association with a W boson*, JHEP **07** (2008) 029, arXiv: 0805.3067 [hep-ph] (cit. on p. 5).

[29] C. D. White, S. Frixione, E. Laenen and F. Maltoni, *Isolating tW production at the LHC*, JHEP **11** (2009) 074, arXiv: 0908.0631 [hep-ph] (cit. on p. 5).

[30] A. Martin, W. Stirling, R. Thorne and G. Watt, *Parton distributions for the LHC*, The European Physical Journal C **63** (2009) 189, ISSN: 1434-6052, arXiv: 0901.0002 [hep-ph] (cit. on p. 5).

[31] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, JHEP **06** (2002) 029, arXiv: hep-ph/0204244 [hep-ph] (cit. on p. 5).

[32] G. Corcella et al., *HERWIG 6: an event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, JHEP **01** (2001) 010, arXiv: hep-ph/0011363 [hep-ph] (cit. on p. 5).

[33] ATLAS Collaboration, *New ATLAS event generator tunes to 2010 data*, ATL-PHYS-PUB-2011-008, 2011, URL: https://cds.cern.ch/record/1345343 (cit. on p. 5).

[34] J. Butterworth, J. Forshaw and M. Seymour, *Multiparton interactions in photoproduction at HERA*, Zeitschrift für Physik C: Particles and Fields **72** (1996) 637, ISSN: 1431-5858, arXiv: hep-ph/9601371 [hep-ph] (cit. on p. 5).

[35] M. Czakon and A. Mitov, *Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders*, Comput. Phys. Commun. **185** (2014) 2930, arXiv: 1112.5675 [hep-ph] (cit. on p. 5).

[36] M. Botje et al., *The PDF4LHC Working Group Interim Recommendations*, 2011, URL: https://cds.cern.ch/record/1318947 (cit. on p. 5).

[37] F. Demartin, S. Forte, E. Mariani, J. Rojo and A. Vicini, *The impact of PDF and α_s uncertainties on Higgs Production in gluon fusion at hadron colliders*, Phys. Rev. D **82** (2010) 014002, arXiv: 1004.0962 [hep-ph] (cit. on p. 5).

[38] J. Gao et al., *CT10 next-to-next-to-leading order global analysis of QCD*, Phys. Rev. D **89** (2014) 033009, arXiv: 1302.6246 [hep-ph] (cit. on p. 5).

[39] R. D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. B **867** (2013) 244, arXiv: 1207.1303 [hep-ph] (cit. on p. 5).

Not reviewed, for internal circulation only

[40] N. Kidonakis, *NNLL resummation for s-channel single top quark production*, Phys. Rev. D **81** (2010) 054028, arXiv: 1001.5034 [hep-ph] (cit. on p. 5).

[41] N. Kidonakis, *Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production*, Phys. Rev. D **83** (2011) 091503, arXiv: 1103.2792 [hep-ph] (cit. on p. 5).

[42] T. Gleisberg, S. Höche, F. Krauss, M. Schönherr, S. Schumann et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, arXiv: 0811.4622 [hep-ph] (cit. on p. 6).

[43] S. Höche, F. Krauss, S. Schumann and F. Siegert, *QCD matrix elements and truncated showers*, JHEP **05** (2009) 053, arXiv: 0903.1219 [hep-ph] (cit. on p. 6).

[44] S. Schumann and F. Krauss, *A Parton shower algorithm based on Catani-Seymour dipole factorisation*, JHEP **03** (2008) 038, arXiv: 0709.1027 [hep-ph] (cit. on p. 6).

[45] J. Butterworth et al., *Single Boson and Diboson Production Cross Sections in pp Collisions at sqrts=7 TeV*, tech. rep. ATL-COM-PHYS-2010-695, CERN, 2010, URL: https://cds.cern.ch/record/1287902 (cit. on p. 6).

[46] ATLAS Collaboration, *Estimation of non-prompt and fake lepton backgrounds in final states with top quarks produced in proton–proton collisions at $\sqrt{s}$ = 8 TeV with the ATLAS Detector*, ATLAS-CONF-2014-058, 2014, URL: https://cds.cern.ch/record/1951336 (cit. on pp. 6, 14).

[47] ATLAS Collaboration, *Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton–proton collisions at the LHC*, Eur. Phys. J. C **77** (2017) 332, arXiv: 1611.10235 [hep-ex] (cit. on p. 6).

[48] ATLAS Collaboration, *Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton–proton collision data*, Eur. Phys. J. C **74** (2014) 3130, arXiv: 1407.3935 [hep-ex] (cit. on pp. 6, 12).

[49] ATLAS Collaboration, *Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton–proton collision data*, Eur. Phys. J. C **74** (2014) 2941, arXiv: 1404.2240 [hep-ex] (cit. on pp. 6, 12).

[50] M. Cacciari, G. P. Salam and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **04** (2008) 063, arXiv: 0802.1189 [hep-ph] (cit. on p. 6).

[51] W. Lampl et al., *Calorimeter Clustering Algorithms: Description and Performance*, ATL-LARG-PUB-2008-002, 2008, URL: https://cds.cern.ch/record/1099735 (cit. on p. 6).

[52] ATLAS Collaboration, *Jet energy measurement and its systematic uncertainty in proton–proton collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector*, Eur. Phys. J. C **75** (2015) 17, arXiv: 1406.0076 [hep-ex] (cit. on pp. 6, 12).

[53] ATLAS Collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s}$ = 8 TeV using the ATLAS detector*, Eur. Phys. J. C **76** (2016) 581, arXiv: 1510.03823 [hep-ex] (cit. on pp. 6, 12).

[54] ATLAS Collaboration, *Performance of b-jet identification in the ATLAS experiment*, JINST **11** (2016) P04008, arXiv: 1512.01094 [hep-ex] (cit. on pp. 7, 12).

[55] ATLAS Collaboration, *Performance of algorithms that reconstruct missing transverse momentum in $\sqrt{s}$ = 8 TeV proton–proton collisions in the ATLAS detector*, Eur. Phys. J. C **77** (2017) 241, arXiv: 1609.09324 [hep-ex] (cit. on pp. 7, 12).

[56] ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2010*, Eur. Phys. J. C **72** (2012) 1849, arXiv: 1110.1530 [hep-ex] (cit. on p. 7).

[57] ATLAS Collaboration, *Performance of the ATLAS muon trigger in pp collisions at $\sqrt{s} = 8$ TeV*, Eur. Phys. J. C **75** (2015) 120, arXiv: 1408.3179 [hep-ex] (cit. on p. 7).

[58] M. Feindt and U. Kerzel, *The NeuroBayes neural network package*, Nucl. Instrum. Meth. A **559** (2006) 190 (cit. on p. 8).

[59] M. Feindt, *A Neural Bayesian Estimator for Conditional Probability Densities*, (2004), arXiv: physics/0402093 [physics.data-an] (cit. on p. 8).

[60] ATLAS Collaboration, *Comprehensive measurements of t-channel single top-quark production cross sections at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Phys. Rev. D **90** (2014) 112006, arXiv: 1406.7844 [hep-ex] (cit. on p. 8).

[61] ATLAS Collaboration, *Jet energy resolution in proton–proton collisions at $\sqrt{s} = 7$ TeV recorded in 2010 with the ATLAS detector*, Eur. Phys. J. C **73** (2013) 2306, arXiv: 1210.6210 [hep-ex] (cit. on p. 12).

[62] ATLAS Collaboration, *Calibration of the performance of b-tagging for c and light-flavour jets in the 2012 ATLAS data*, ATLAS-CONF-2014-046, 2014, URL: https://cds.cern.ch/record/1741020 (cit. on p. 12).

[63] ATLAS Collaboration, *Calibration of b-tagging using dileptonic top pair events in a combinatorial likelihood approach with the ATLAS experiment*, ATLAS-CONF-2014-004, 2014, URL: https://cds.cern.ch/record/1664335 (cit. on p. 12).

[64] ATLAS Collaboration, *Study of correlation of PDF uncertainty in single top and top pair production at the LHC*, ATL-PHYS-PUB-2015-010, 2015, URL: https://cds.cern.ch/record/2020601 (cit. on p. 12).

[65] ATLAS Collaboration, *Measurements of the W production cross sections in association with jets with the ATLAS detector*, Eur. Phys. J. C **75** (2015) 82, arXiv: 1409.8639 [hep-ex] (cit. on p. 14).

[66] R. J. Barlow and C. Beeston, *Fitting using finite Monte Carlo samples*, Comput. Phys. Commun. **77** (1993) 219 (cit. on p. 14).

[67] J. Conway, *Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra*, (2011), arXiv: 1103.0354 [physics.data-an] (cit. on p. 14).

[68] K. Cranmer, G. Lewis, L. Moneta, A. Shibata and W. Verkerke, *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, (2012), URL: https://cds.cern.ch/record/1456844 (cit. on p. 14).

[69] L. Moneta, K. Belasco, K. S. Cranmer, S. Kreiss, A. Lazzaro et al., *The RooStats Project*, PoS **ACAT2010** (2010) 057, arXiv: 1009.1003 [physics.data-an] (cit. on p. 14).

[70] F. James and M. Roos, *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations*, Comput. Phys. Commun. **10** (1975) 343 (cit. on p. 14).

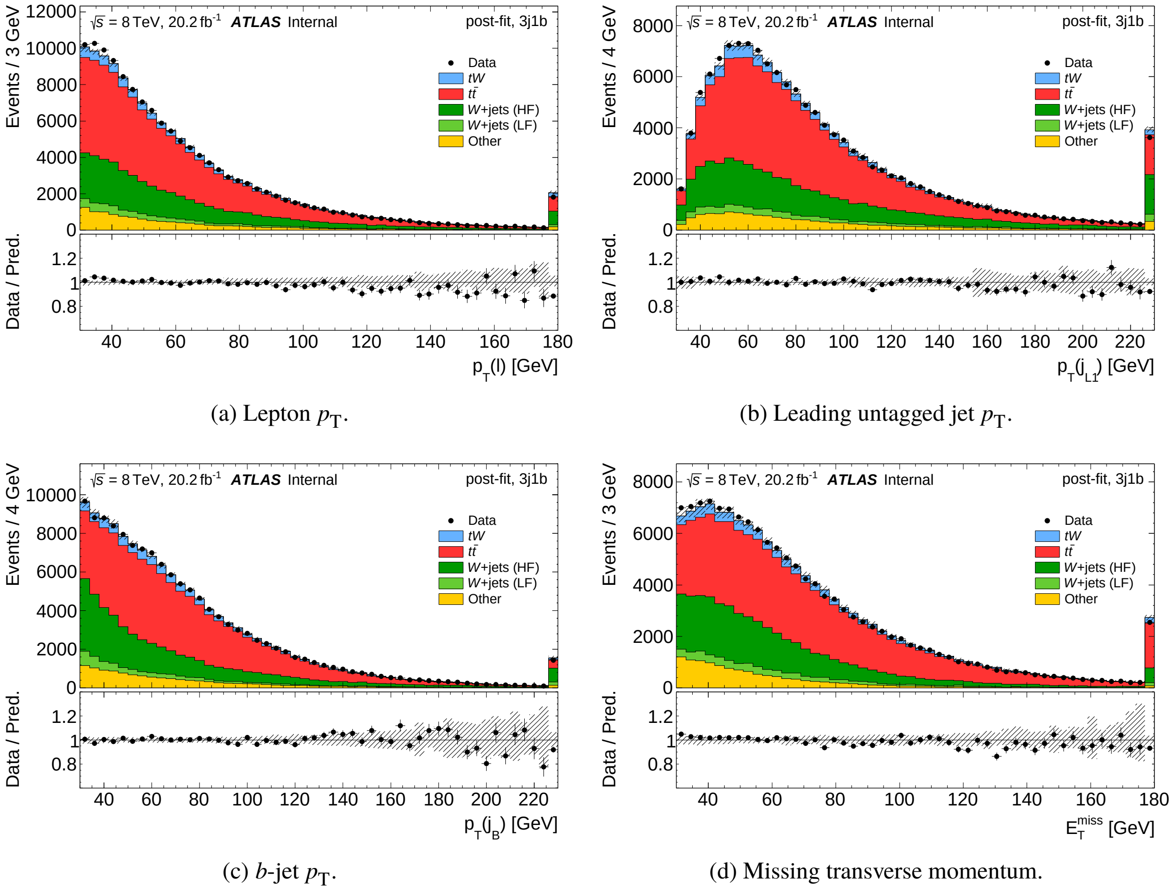
[71] G. Cowan, K. Cranmer, E. Gross and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. C **71** (2011) 1554, arXiv: 1007.1727 [physics.data-an] (cit. on p. 14), Erratum: Eur. Phys. J. C **73** (2013) 2501.

[72] ATLAS and CMS Collaborations, *Combinations of single-top-quark production cross-section measurements and $|f_{LV}V_{tb}|$ determinations at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS experiments*, (2019), arXiv: 1902.07158 [hep-ex] (cit. on p. 19).

[73] N. Kidonakis, *Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-*, Phys. Rev. D **82** (2010) 054018, arXiv: 1005.4451 [hep-ph] (cit. on p. 19).

[74] ATLAS Collaboration, *ATLAS Computing Acknowledgements*, ATL-GEN-PUB-2016-002, URL: https://cds.cern.ch/record/2202407 (cit. on p. 20).

Auxiliary material



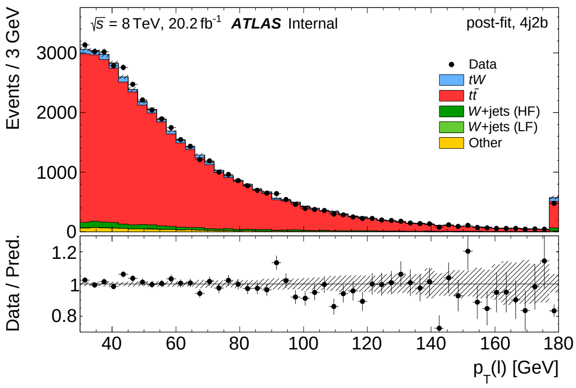
(a) Lepton p_{T}.

(b) Leading untagged jet p_{T}.

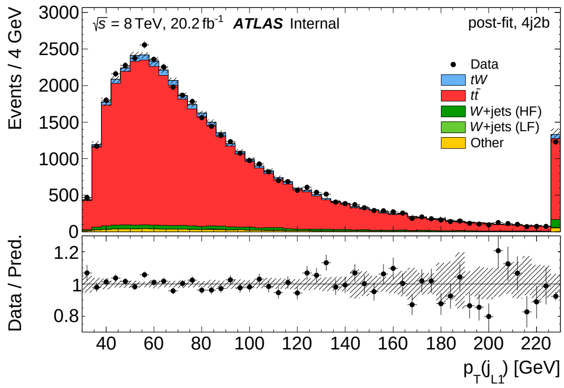
(c) b-jet p_{T}.

(d) Missing transverse momentum.

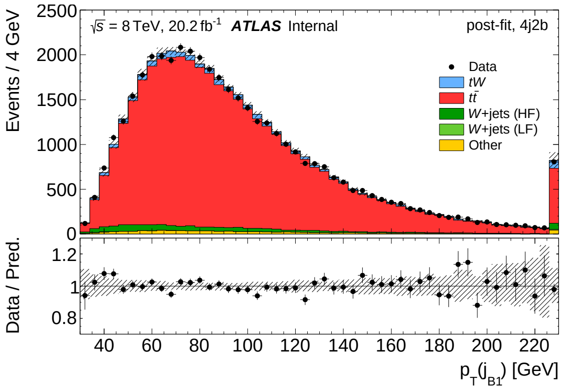
Figure 9: Post-fit transverse momentum distributions of (a) the selected lepton ($l = \mu, e$), (b) the leading untagged jet, (c) the leading b-tagged jet and (d) $E_{\mathrm{T}}^{\mathrm{miss}}$, in the tW signal region (3j1b). Small backgrounds are subsumed under ‘Other’. The dashed uncertainty band includes statistical and systematic uncertainties. The last bin includes the overflow events. The lower panels show the ratio of the observed and the expected number of events in each bin.



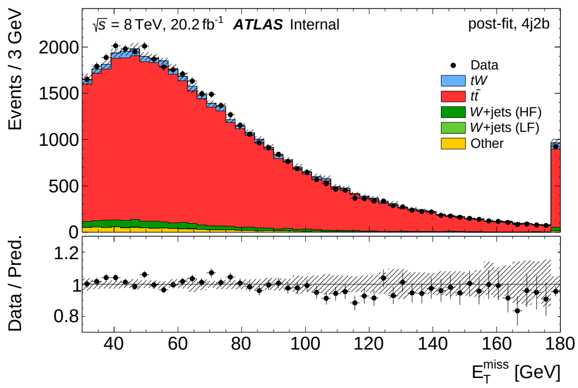
(a) Lepton p_{T}.



(b) Leading untagged jet p_{T}.

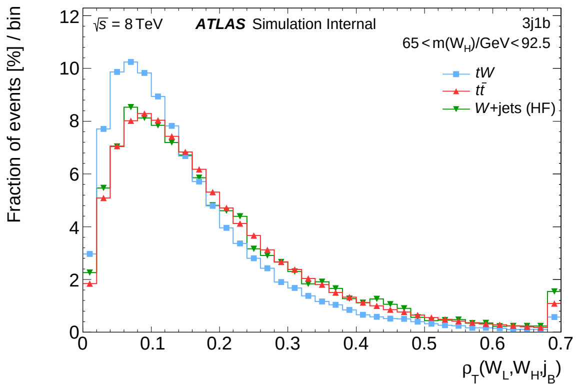


(c) Leading b-jet p_{T}.

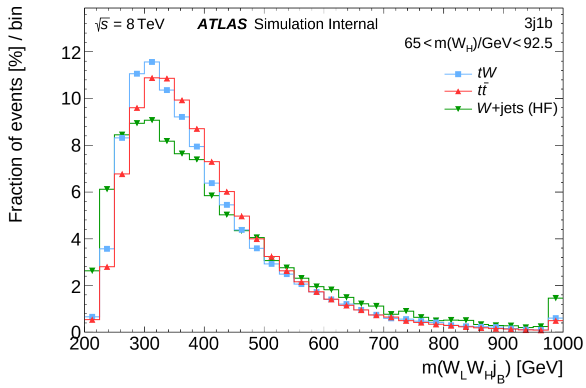


(d) Missing transverse momentum.

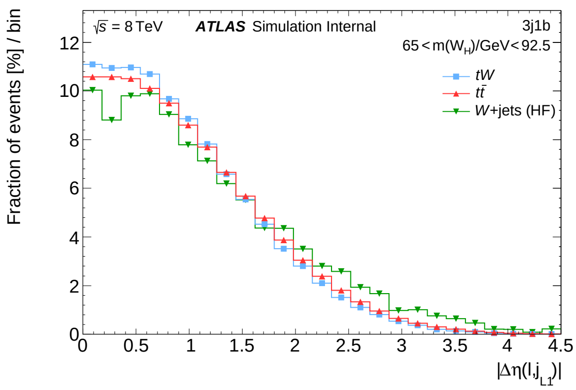
Figure 10: Post-fit transverse momentum distributions of (a) the selected lepton ($l = \mu, e$), (b) the leading untagged jet, (c) the leading b-tagged jet and (d) $E_{\mathrm{T}}^{\mathrm{miss}}$ in the $t\bar{t}$ validation region (4j2b). Small backgrounds are subsumed under 'Other'. The dashed uncertainty band includes statistical and systematic uncertainties. The last bin includes the overflow events. The lower panels show the ratio of the observed and the expected number of events in each bin.



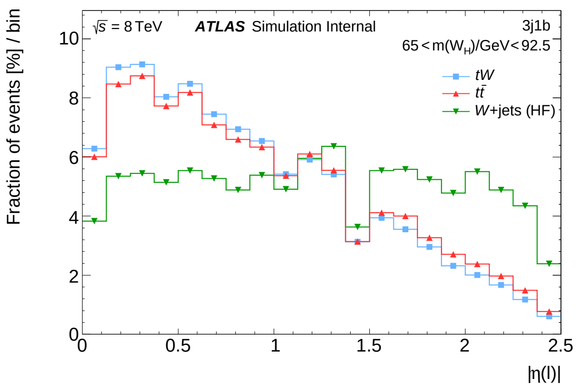
(a) p_{T} balance of the tW system.



(b) Invariant mass of the tW system.

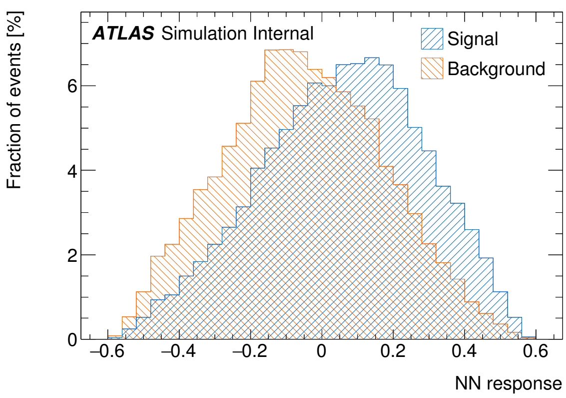


(c) Lepton and light jet pseudorapidity difference.

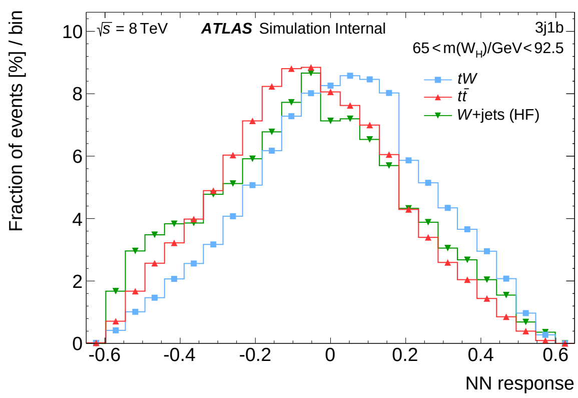


(d) Pseudorapidity of the lepton (l=μ,e).

Figure 11: Shape distributions of the input neural-network variables in the signal (3j1b) region for events with invariant mass in the range 65 GeV $\leq m(W_{\mathrm{H}}) \leq$ 92.5 GeV (11(a), 11(b), 11(c), 11(d)). The distribution for each process normalised to unity is shown.

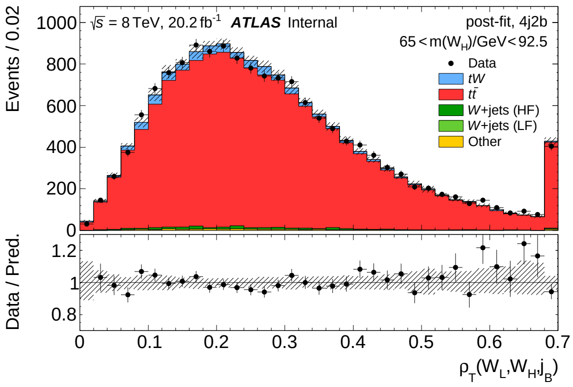


(a)

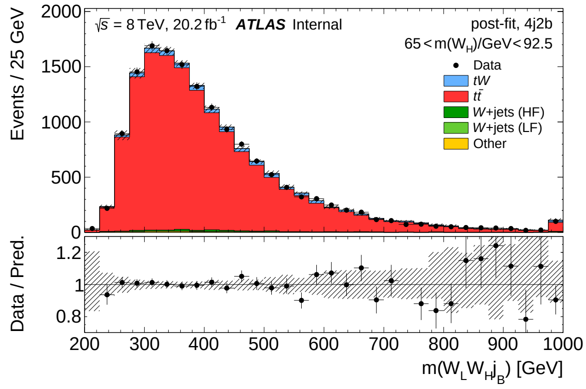


(b)

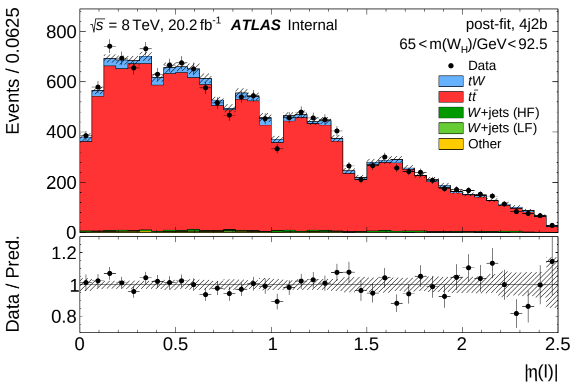
Figure 12: (a) Neural-network (NN) response for the signal (blue) and background (red) samples used for the NN training, which consist only of events with a well-reconstructed hadronic W-boson decay. (b) NN shape distributions in the signal (3j1b) region for events with invariant mass between 65 GeV $\leq m(W_{\mathrm{H}}) \leq$ 92.5 GeV. The distribution for each process normalised to unity is shown.



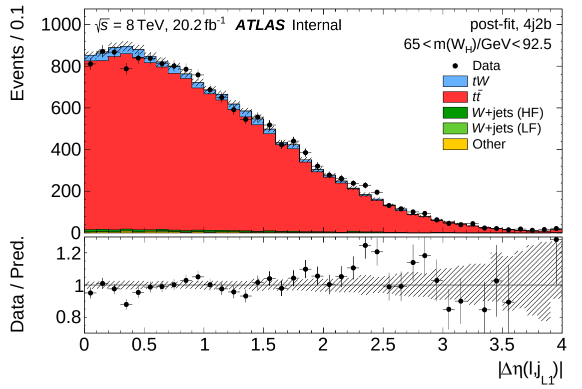
(a) p_{T} balance of the tW system.



(b) Invariant mass of the tW system.



(c) Lepton and light jet pseudorapidity difference.



(d) Pseudorapidity of the lepton ($l = \mu, e$).

Figure 13: Post-fit distributions of the input neural-network variables in the $t\bar{t}$ validation (4j2b) region with $65\,\mathrm{GeV} \leq m(W_{\mathrm{H}}) \leq 92.5\,\mathrm{GeV}$ (13(a), 13(b), 13(c), 13(d)). Small backgrounds are subsumed under 'Other'. 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 expected number of events in each bin.

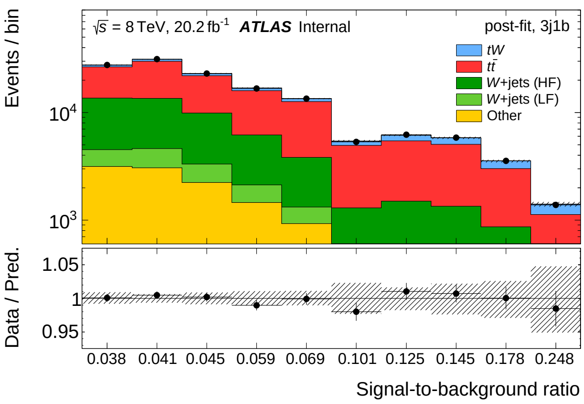


Figure 14: Rebinned post-fit distribution of the discriminant in the signal region. The bins in the distribution shown in 7(a) have been ordered by their signal-to-background ratio, and bins with similar ratios have been merged. Small backgrounds are subsumed under 'Other'. The dashed uncertainty band includes statistical and systematic uncertainties. The last bin includes the overflow events. The lower panel shows the ratio of the observed and the expected number of events in each bin.

Not reviewed, for internal circulation only

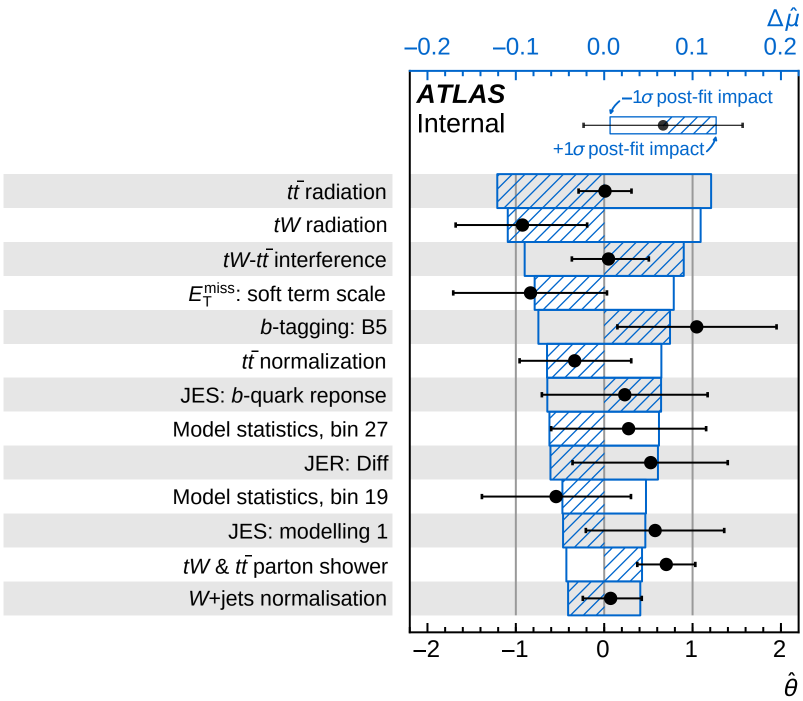


Figure 15: The measured values of the nuisance parameters after fitting the model to the observed data ($\hat{\theta}$). The black points represent $\hat{\theta}$ and the error bars are the post-fit errors of the fit parameter. The blue boxes shown are the post-fit impact ($\Delta\hat{\mu}$) of each nuisance parameter, see Section 8. The hatched part of the box indicates whether the measured signal strength has a positive or a negative correlation to the nuisance parameter. '*b*-tagging: B5' corresponds to the largest eigenvariation of the uncertainty in the *b*-tagging efficiency, 'JES: modelling 1' to the largest eigenvariation of the modelling uncertainties in the JES and the 'JER: Diff' the largest JER eigenvariation associated with data and MC differences.