Measurement of the top quark pair cross section with ATLAS in pp collisions at $\sqrt{s} = 7$ TeV using final states with an electron or a muon and a hadronically decaying $\tau$ lepton

ATLAS Collaboration

1. Introduction

Measuring the top quark pair ($t\bar{t}$) production cross section ($\sigma_{t\bar{t}}$) in different decay channels is of interest because it can indicate physics beyond the Standard Model (SM). In the SM, the top quark decays with a branching ratio close to 100% into a $W$ boson and a $b$ quark, and $t\bar{t}$ pairs are identified by either the hadronic or leptonic decays of the $W$ bosons and the presence of additional jets. The ATLAS Collaboration has previously used the single-lepton channel [1] from a 35 pb$^{-1}$ data sample, and the dilepton channels including only electrons and muons [2] from a 0.7 fb$^{-1}$ data sample, to perform cross-section measurements at 7 TeV proton–proton centre-of-mass energy. Similar measurements have been performed by the CMS Collaboration [3–5]. All these measurements are systematics limited.

The large data samples for $t\bar{t}$ production at the Large Hadron Collider (LHC) provide an opportunity to measure $\sigma_{t\bar{t}}$ using final states with an electron or a muon and a $\tau$ lepton with high precision. The $\sigma_{t\bar{t}}$ in this channel has been measured at the Tevatron in $pp$ collisions at 1.96 TeV with 25% precision [6] and recently by the CMS Collaboration at the LHC with 18% precision [7]. A deviation from the value of $\sigma_{t\bar{t}}$ measured in other final states would be an indication of non-Standard Model decays of the top quark, such as a decay to a charged Higgs ($H^{\pm}$) and a $b$ quark with $H^{\pm}$ decaying to a $\tau$ lepton and a $\tau$ neutrino, or of contributions from other non-Standard Model processes [8–10]. Upper limits on the branching ratio of top quark decays to $H^{\pm}$ bosons decaying to a $\tau$ lepton and a neutrino have been published by Tevatron and LHC experiments [11–13].

2. Analysis overview

This analysis uses 2.05 fb$^{-1}$ of data collected by ATLAS from $pp$ collisions in the LHC at a centre-of-mass energy of 7 TeV between March and August 2011. The $t\bar{t}$ events are selected with kinematic criteria that make use of the fact that they feature two $W$ bosons and two $b$ quarks. The selections favour events with one $W$ decay to a charged $\ell$ (with $\ell$ denoting an electron or a muon; either prompt or from a $\tau$ lepton decay to $\ell$) and a neutrino and the other $W$ decays to a $\tau$ lepton and a neutrino with the $\tau$ lepton in turn decaying hadronically. In addition at least one jet is tagged ($b$-tag) as originating from a $b$ quark ($b$-jet) by means of an algorithm that can identify $b$-jets with high efficiency while maintaining a high rejection of light-quark jets. Isolated electrons and muons are well identified, but because of the large cross section for multi-jet production the background from jets misidentified as isolated electrons or muons is not negligible. This background is reduced by requiring significant missing transverse momentum signalling the presence of energetic neutrinos. Hadronic $\tau$ lepton decays are more difficult to identify and require elaborate techniques to reject jets and electrons misidentified as a $\tau$ lepton.

Section 5 describes how the objects used in the event selection are defined. After all selections given in Section 5.2, the dominant background to the $t\bar{t} \rightarrow \ell + \tau + X$ channel is the $t\bar{t} \rightarrow t\bar{t}$ channel in which the $\tau$ candidate is from jets misidentified as hadronic $\tau$ lepton decays. Therefore, $\tau$ lepton identification ($\tau$ ID) is critical.
for separating signal and background. The τ ID methodology employed in this analysis exploits a multivariate technique to build a discriminant [14]. A boosted decision tree (BDT) algorithm is used [15,16]. The number of τ leptons in the selected samples is extracted by fitting the distributions of BDT outputs to background and signal templates. Section 6 describes how the background templates are constructed using control data samples. They exploit the fact that events with ℓ and τ candidates of opposite sign charge (OS) contain real τ leptons while those with same sign charge (SS) are pure background. Events with τ leptons are not all from ℓ+τ; the contribution from processes other than ℓ+ℓ + τ + X is estimated from Monte Carlo simulation. Section 7 describes the fitting procedure and the results of the fit. The fit results are also checked using an alternative method, referred to as the "matrix method", based on a cut on the BDT output (Section 7.1). The measured cross section is given in Section 8 and the conclusions are in Section 9.

3. ATLAS detector

The ATLAS detector [17] at the LHC covers nearly the entire solid angle around the collision point. It consists of an inner tracking detector surrounded by a thin superconducting solenoid, electromagnetic (EM) and hadronic calorimeters, and an external muon spectrometer incorporating three large superconducting toroid magnet assemblies. The inner tracking detector provides tracking information in a pseudorapidity range |η| < 2.5. The liquid-argon (LAr) EM sampling calorimeters cover a range of |η| < 3.2 with fine granularity. An iron-scintillator tile calorimeter provides hadronic energy measurements in the central rapidity range |η| < 1.7. The endcap and forward regions are instrumented with LAr calorimeters for both EM and hadronic energy measurements covering |η| < 4.9. The muon spectrometer provides precise tracking information in a range of |η| < 2.7.

ATLAS uses a three-level trigger system to select events. The level-1 trigger is implemented in hardware using a subset of detector information to reduce the event rate to below 75 kHz. This is followed by two software-based trigger levels, level-2 and the event filter, which together reduce the event rate to about 300 Hz recorded for analysis.

4. Simulated event samples

Monte Carlo (MC) simulation samples are used to optimise selection procedures, to calculate the signal acceptance and to evaluate the background contributions from single top quark, W + W, W + Z and ZZ production and Z → τ + τ decay. The event generation, all samples are processed with the GEANT4 [18] simulation of the ATLAS detector, the trigger simulation and are then subject to the same reconstruction algorithms as the data [19].

For the ℓ+ℓ− single top quark final states, the next-to-leading-order (NLO) generator MC@NLO [20–22] is used with a top quark mass of 172.5 GeV and with the NLO parton distribution function (PDF) set CTEQ6.6 [23]. The MC@NLO program uses HERWIG [24] to simulate the parton shower and hadronise the partons. The "diagram removal scheme" is used to remove overlaps between the single top quark and the ℓ+ℓ− final states. The ℓ+ℓ− cross section is normalised to the prediction of HATHOR (164 ± 16 pb) [25], which employs an approximate next-to-next-to-leading-order (NNLO) per-turbative Quantum Chromodynamics (QCD) calculation. The diboson samples are generated with HERWIG. W + jets events and Z/γ∗ + jets events (with dilepton invariant mass m_{ℓ+ℓ−} > 40 GeV) are generated by ALPGEN generator [26] with up to five outgoing partons from the hard scattering process, in addition to the lepton pairs. The MLM matching scheme [27] of the ALPGEN generator is used to remove overlaps between matrix-element and parton-shower products. Parton evolution and hadronisation is handled by HERWIG, as is the generation of diboson events. The next-order PDF set CTEQ6L is used for all backgrounds described above.

All samples that use HERWIG for parton shower evolution and hadronisation rely on JIMMY [28] for the underlying event model. The ℓ+ℓ− decays are handled by TAUOLA [29]. The effect of multiple pp interactions per bunch crossing ("pile-up") is modelled by overlaying simulated minimum bias events over the original hard-scattering event [30]. MC events are then reweighted so that the distribution of interactions per crossing in the MC simulation matches that observed in data. The average number of pile-up interactions in the sample is 63.4.

5. Object identification and event selection

The event selection uses nearly the same object definition as in the ℓ+ℓ− cross-sectional measurement in the dilepton channel [2] with the exception of a τ candidate instead of a second electron or muon candidate. The electrons must be isolated and have E_T > 25 GeV and |η_cluster| < 2.47, excluding the barrel-endcap transition region (1.37 < |η_cluster| < 1.52), where E_T is the transverse energy and η_cluster is the pseudorapidity of the calorimeter energy cluster associated with the candidate. The electron is defined as isolated if the E_T deposited in the calorimeter and not associated with the electron in a cone in η−φ space of radius ΔR = 0.2 is less than 4 GeV. The muons must also be isolated and have transverse momentum p_T > 20 GeV and |η| < 2.5. For isolated muons, both the corresponding E_T and the analogous track isolation transverse momentum must be less than 4 GeV in a cone of ΔR = 0.3.

The track isolation p_T is calculated from the sum of the track transverse momenta for tracks with p_T > 1 GeV around the muon. Jets are reconstructed with the anti-kT algorithm [31] with a radius parameter R = 0.4, starting from energy deposits (clusters) in the calorimeter reconstructed using the scale established using Z → e+e− events for electromagnetic objects. These jets are then calibrated to the hadronic energy scale using p_T- and η-dependent correction factors obtained from simulation [32]. The jet candidates are required to have p_T > 25 GeV and |η| < 2.5. A jet is tagged as a b-jet by a vertex tagging algorithm that constructs a likelihood ratio of b- and light-quark jet hypothesis using the following discriminating variables: the signed impact parameter significance of well measured tracks associated with a given jet, the decay length significance associated with a reconstructed secondary vertex, the invariant mass of all tracks associated to the secondary vertex, the ratio of the sum of the energies of the tracks associated with the secondary vertex to the sum of the energies of all tracks in the jet assuming a pion hypothesis, and the number of two-track vertices that can be formed at the secondary vertex. The cut on the combined likelihood ratio has been chosen to give an average efficiency of 70% for b-quark jets from ℓ+ℓ− events and a 1% efficiency for light-quark and gluon jets [33].

1. Atlas uses a right-handed coordinate system with its origin at the nominal interaction point in the centre of the detector and the z-axis along the beam pipe. The x-axis points to the centre of the LHC ring, and the y-axis points upwards. The azimuthal angle φ is measured around the beam axis and the polar angle θ is the angle from the beam axis. The pseudorapidity is defined as η = − ln(tan(θ/2)). The distance ΔR in η−φ space is defined as ΔR = √((Δη)² + (Δφ)²).

2. The fraction of events with m_{ℓ+ℓ−} < 40 GeV is estimated to be less than 0.2% of the total after all selections. The estimate is based on ALPGEN samples for Drell–Yan simulation and confirmed by a good agreement with data.
The missing transverse momentum is constructed from the vector sum of all calorimeter cells with $|η| < 4.5$, projected onto the transverse plane. Its magnitude is denoted $E_T^{miss}$. The hadronic energy scale is used for the energies of cells associated with jets; τ candidates are treated as jets. Contributions from cells associated with electrons employ the electromagnetic energy calibration. Contributions from the $p_T$ of muon tracks are included, removing the contributions of any calorimeter cells associated with the muon.

5.1. τ reconstruction and identification

The reconstruction and identification of hadronically decaying τ leptons proceed as follows:

1. the τ candidate reconstruction starts by considering each jet as a τ candidate;
2. energy clusters in the calorimeter associated with the τ candidate are used to calculate kinematic quantities (such as $E_T$) and the associated tracks are found;
3. identification variables are calculated from the tracking and calorimeter information;
4. these variables are combined into multivariate discriminants and the outputs of the discriminants are used to separate jets and electrons misidentified as τ leptons decaying hadronically from τ leptons.

Details, including the variable definitions used in the multivariate discriminants, are given in Ref. [9]. In this analysis the outputs of BDT discriminants are used. Reconstructed τ candidates are required to have $20 \text{ GeV} < E_T < 100 \text{ GeV}$. They must also have $|η| < 2.3$, and one, two or three associated tracks. A track is associated with the τ candidate if it has $p_T > 1 \text{ GeV}$ and is inside a cone of $ΔR < 0.4$ around the jet axis. The associated track with highest $p_T$ must have $p_T > 4 \text{ GeV}$. The charge is given by the sum of the charges of the associated tracks, and is required to be non-zero. The probability of misidentifying the τ lepton charge sign is about 1%. The charge misidentification rate for muons and electrons is negligible.

If the τ candidate overlaps with a muon (with $p_T > 4 \text{ GeV}$ and without an isolation requirement) or an electron candidate within $ΔR(ℓ, τ) < 0.4$, the τ candidate is removed. To remove electrons misidentified as τ leptons, an additional criterion is used that relies on a BDT trained to separate τ leptons and electrons (BDTτ) using seven variables shown to be well modelled by comparing $Z → e^+e^−$ and $Z → τ^+τ^−$ events in data and in MC simulation. The variables were chosen after ranking a large set by their effectiveness.\(^3\) The most effective variables for BDTτ are $E/p$, the EM fraction (the ratio of the τ candidate energy measured in the EM calorimeter to the total τ candidate energy measured in the calorimeter), and the cluster-based shower width. The BDT output tends to be near 1 (0) if the τ candidate is a τ lepton (electron). BDTτ was trained using $Z → e^+e^−$ and $Z → τ^+τ^−$ Monte Carlo samples. The τ candidate is required to satisfy $BDT_{τ} > 0.51$; 85% of reconstructed τ leptons decaying hadronically satisfy this requirement, as measured in $Z → τ^+τ^−$ events. The additional rejection for electrons is a factor of 60.

The majority of objects reconstructed as τ candidates in a multi-jet environment are jets misidentified as τ leptons. A jet or an electron misidentified as a τ lepton will be referred to as a τ. Another BDT (BDTj) based on eight variables is used to separate τ leptons in τ candidates with one track (denoted $τ_1$) from such jets. For candidates with more than one track (denoted $τ_2$) BDTj includes ten variables. The BDTj was trained using multi-jet events as background and $Z → τ^+τ^−$ Monte Carlo as signal. The most effective variables for BDTj are calorimeter and track isolation, cluster-based jet mass, and the fraction of energy within $ΔR = 0.1$ of the jet axis. The BDTj distributions are fit with templates for background and signal to extract the number of τ leptons in the sample. Details are given in Section 7. The fake τ background in the $τ_3$ sample is significantly higher than in the $τ_1$ sample, leading to very different BDTj distributions. Hence independent measurements are carried out for $τ_1$ and $τ_3$ candidate events and the results are combined at the end. If there is a $τ_1$ and a $τ_3$ candidate in the event, the $τ_3$ candidate is kept as the probability that the $τ_1$ is a τ lepton is much higher. If there are two $τ_1$ or $τ_3$ candidates, both are kept.

5.2. Event selection

For this analysis, events are selected using a single-muon trigger with a $p_T$ threshold of 18 GeV or a single-electron trigger with a $p_T$ threshold of 20 GeV, rising to 22 GeV during periods of high instantaneous luminosity. The offline requirements are based on data quality criteria and optimised using Monte Carlo simulation:

- a primary vertex with at least five tracks, each with $p_T > 400 \text{ MeV}$, associated with it;
- one and only one isolated high-$p_T$ muon and no identified electrons for the $μ + τ$ channel, or one and only one isolated electron and no isolated muons for the $e + τ$ channel;
- at least one τ candidate (as defined in Section 5.1);
- at least two jets not overlapping with a τ candidate, i.e. $ΔR(τ, jet) > 0.4$;
- $E_T^{miss} > 30 \text{ GeV}$ to reduce the multi-jet background, and the scalar sum of the $p_T$ of the leptons (including τ), jets, and $E_T^{miss}$ must be greater than 200 GeV, to reduce the $W + jets$ background.

The $e + τ$ samples are divided into events with no jets identified as a $b$-quark jet (0 $b$-tag control sample) and those with at least one such jet ($≥1$ $b$-tag $ℓ$ sample). The 0 $b$-tag sample is used to estimate the background in the $≥1$ $b$-tag $ℓ$ sample. Each sample is split into two, one with the τ candidate and $ℓ$ having the opposite sign charge (OS), and the other one with τ and $ℓ$ having the same sign charge (SS). While the τ candidates in the SS samples are almost all fake τ leptons, the OS samples have a mixture of τ leptons and fake τ leptons. The numbers of observed and expected events in the above samples are shown in Table 1. The $e + jets$ is the contribution from all processes with a $ℓ$ and a τ candidate that is a jet misidentified as a τ lepton other than from $ℓt$ ($→ ℓ + jets$). The $τ$ entries require the reconstructed τ candidate be matched to a generated τ lepton. The matching criterion is $ΔR < 0.1$ between the τ candidate and the observable component of the generated τ lepton.

To estimate the multi-jet background from data, an event selection identical to the $μ + τ$ ($e + τ$) event selection except for an inverted muon (electron) isolation cut is used to obtain a multi-jet template for the shape of the transverse mass, $m_T = \sqrt{(p_T^\mu + E_T^{miss})^2 - (p_T^\mu + E_T^{miss})^2 - (p_T^e + E_T^{miss})^2}$. The normalisation of each selected data sample is obtained by fitting the $m_T$ distribution of the selected data samples with the multi-jet template and the sum of non-multi-jet processes predicted by MC, allowing the amount of both to float. The uncertainty on the multi-jet

---

\(^3\) The effectiveness is quantified by quadratically summing over the change in the purity between the mother and daughter leaves for every node in which the given variable is used in a decision tree.
Table 1
Number of $\ell + \tau$ candidates for Monte Carlo samples and data. $t\ell(\ell + e)$ are $t\ell$ events with one identified lepton and an electron reconstructed as a $\tau$ candidate. $t\ell(\ell + jets)$ are $t\ell$ events with one identified lepton and a jet reconstructed as a $\tau$ candidate. $t\ell(\ell + jets)$ are events with one identified lepton and a jet reconstructed as a $\tau$ candidate from sources other than $t\ell(\ell + jets)$ and multi-jets. Sources contributing to jet fakes are $W + jets$, $Z + jets$, single top quark, diboson events and $t\ell$. $W(\ell + \tau)$ is $W + \tau$ production with one $W$ decaying to $\ell$ and another to $\tau$. The uncertainties are statistical only, except for the multi-jet estimates. MC samples are normalised to the data integrated luminosity.

<table>
<thead>
<tr>
<th>$\mu + \tau$</th>
<th>$\mu(\mu + \tau)$</th>
<th>OS</th>
<th>SS</th>
<th>OS</th>
<th>SS</th>
<th>OS</th>
<th>SS</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t\ell(\mu + \tau)$</td>
<td>60 ± 2</td>
<td>&lt;1</td>
<td>390 ± 4</td>
<td>2 ± 1</td>
<td>17 ± 1</td>
<td>1 ± 1</td>
<td>118 ± 3</td>
</tr>
<tr>
<td>$t\ell(\mu + jets)$</td>
<td>310 ± 4</td>
<td>163 ± 3</td>
<td>1528 ± 9</td>
<td>660 ± 6</td>
<td>685 ± 6</td>
<td>443 ± 5</td>
<td>3484 ± 13</td>
</tr>
<tr>
<td>$\mu + jets$</td>
<td>3010 ± 70</td>
<td>3020 ± 60</td>
<td>496 ± 17</td>
<td>297 ± 17</td>
<td>12230 ± 120</td>
<td>8670 ± 90</td>
<td>1293 ± 28</td>
</tr>
<tr>
<td>Multi-jets</td>
<td>470 ± 140</td>
<td>540 ± 160</td>
<td>117 ± 35</td>
<td>150 ± 40</td>
<td>990 ± 300</td>
<td>1120 ± 340</td>
<td>460 ± 140</td>
</tr>
<tr>
<td>$W(\mu + \tau)$</td>
<td>7 ± 1</td>
<td>&lt;1</td>
<td>18 ± 1</td>
<td>1 ± 1</td>
<td>2 ± 1</td>
<td>&lt;1</td>
<td>5 ± 1</td>
</tr>
<tr>
<td>$Z \rightarrow \tau \tau$</td>
<td>301 ± 13</td>
<td>2 ± 1</td>
<td>16 ± 3</td>
<td>&lt;1</td>
<td>75 ± 7</td>
<td>1 ± 1</td>
<td>3 ± 2</td>
</tr>
<tr>
<td>Total</td>
<td>6160 ± 160</td>
<td>3730 ± 170</td>
<td>2580 ± 40</td>
<td>1110 ± 40</td>
<td>14000 ± 320</td>
<td>10230 ± 350</td>
<td>5370 ± 140</td>
</tr>
<tr>
<td>Data</td>
<td>5450</td>
<td>3700</td>
<td>2472</td>
<td>1332</td>
<td>13322</td>
<td>10193</td>
<td>5703</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$e + \tau$</th>
<th>$e(\ell + e)$</th>
<th>OS</th>
<th>SS</th>
<th>OS</th>
<th>SS</th>
<th>OS</th>
<th>SS</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t\ell(e + \tau)$</td>
<td>54 ± 7</td>
<td>1 ± 1</td>
<td>342 ± 19</td>
<td>3 ± 2</td>
<td>15 ± 4</td>
<td>&lt;1</td>
<td>103 ± 10</td>
</tr>
<tr>
<td>$t\ell(e + jets)$</td>
<td>2 ± 1</td>
<td>&lt;1</td>
<td>11 ± 3</td>
<td>1 ± 1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>2 ± 1</td>
</tr>
<tr>
<td>$t\ell(e + jets)$</td>
<td>273 ± 17</td>
<td>146 ± 12</td>
<td>1340 ± 40</td>
<td>599 ± 25</td>
<td>633 ± 25</td>
<td>399 ± 20</td>
<td>3000 ± 60</td>
</tr>
<tr>
<td>$e + jets$</td>
<td>3950 ± 60</td>
<td>2590 ± 50</td>
<td>380 ± 20</td>
<td>256 ± 16</td>
<td>10140 ± 100</td>
<td>7530 ± 90</td>
<td>1120 ± 33</td>
</tr>
<tr>
<td>Multi-jets</td>
<td>600 ± 180</td>
<td>620 ± 190</td>
<td>170 ± 50</td>
<td>140 ± 40</td>
<td>2000 ± 600</td>
<td>2000 ± 600</td>
<td>690 ± 210</td>
</tr>
<tr>
<td>$Z \rightarrow \ell \ell$</td>
<td>92 ± 10</td>
<td>3 ± 2</td>
<td>9 ± 3</td>
<td>&lt;1</td>
<td>11 ± 3</td>
<td>2 ± 1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>$W(\ell + e)$</td>
<td>7 ± 1</td>
<td>&lt;1</td>
<td>17 ± 4</td>
<td>&lt;1</td>
<td>1 ± 1</td>
<td>&lt;1</td>
<td>5 ± 2</td>
</tr>
<tr>
<td>$Z \rightarrow \tau \tau$</td>
<td>217 ± 15</td>
<td>2 ± 1</td>
<td>15 ± 4</td>
<td>&lt;1</td>
<td>65 ± 7</td>
<td>1 ± 1</td>
<td>3 ± 2</td>
</tr>
<tr>
<td>Total</td>
<td>5190 ± 190</td>
<td>3360 ± 200</td>
<td>2280 ± 70</td>
<td>990 ± 50</td>
<td>12900 ± 600</td>
<td>9900 ± 600</td>
<td>5020 ± 220</td>
</tr>
<tr>
<td>Data</td>
<td>5111</td>
<td>3462</td>
<td>2277</td>
<td>1107</td>
<td>12102</td>
<td>9635</td>
<td>5033</td>
</tr>
</tbody>
</table>

6. Background models

The jet origin can strongly influence the $\tau$-lepton fake probability. Due to their narrow shower width and lower track multiplicity, light-quark jets have a higher probability of faking a $\tau$ lepton than other jet types. Thus the BDT$_j$ distributions have a strong dependence on the jet type. It is therefore crucial to build a background model which properly reflects the jet composition in order to correctly estimate the fake $\tau$ contamination in the signal region. Deriving this background model from control regions in data rather than MC simulation is preferable in order to avoid systematic effects related to jet composition in the MC models.

The gluon component of the fake $\tau$ leptons is charge symmetric; therefore it is expected to have the same shape in SS events as in OS events and should contribute the same number of fake $\tau$ leptons in each sample. The contribution of fake $\tau$ leptons from gluons can be removed by subtracting the distribution of any quantity for SS events from the corresponding distribution for OS events. The multi-jet background also cancels, as can be seen in Table 1. The resulting distributions are labelled OS-SS. Similarly, since each sample is expected to have an almost equal contribution from $b$-jets and $b$-jets, the small $b$-jet component should also be removed by OS-SS (the asymmetry in $b$ production, mainly from single top quark final states, is negligible). The only jet types remaining in the OS-SS distributions are light-quark jets. MC studies indicate that the BDT$_j$ distributions of $c$-quark jets misidentified as $\tau$ leptons are not noticeably different from those of light-quark jets.

One can construct a background BDT$_j$ distribution from the 0 $b$-tag data sample by subtracting the expected amount of $\tau$-lepton signal. The signal is mainly from $Z \rightarrow \tau^+ \tau^-$ and can be reliably predicted from MC. A control sample dominated by $W + jets$ events is considered as a check. The latter sample is selected by requiring events with a muon and a $\tau$ candidate, no additional jets, $E_T^{miss} > 30$ GeV and 40 GeV $< m_T < 100$ GeV. According to MC simulation, in $W + jets$ events where exactly one jet is required, 90% of the fake $\tau$ leptons are from light-quark jets and 10% from gluons. This sample is labelled $W + 1$ jet.

The BDT$_j$ background shapes for the OS-SS 0 $b$-tag and $\geq 1$ $b$-tag data samples are not identical to the $W + 1$ jet distributions for two reasons: (1) the shape depends on the jet multiplicity, (2) different OS/SS ratios are observed in the samples. The dependence on the OS/SS ratio comes from the differences in jet fragmentation producing a leading particle with the opposite charge and the same charge as the initial quark; thus the OS BDT$_j$ shape from light-quark jets differs from the equivalent SS BDT$_j$ shape. The ratio of OS-SS BDT$_j$ background distributions derived from $W + 1$ jet and $\geq 1$ $b$-tag simulated events show that significant corrections are needed (30% for BDT$_j > 0.8$, a region dominated by the true $\tau$ signal). For the 0 $b$-tag sample the corresponding corrections are much smaller (5% in the same region). Both the 0 $b$-tag and the $W + 1$ jet data samples are used to obtain statistically independent estimates of the background in the $\geq 1$ $b$-tag sample.
Two different approaches are used for constructing backgrounds in the \( \geq 1 \) b-tag data sample. One, used by the fit method (Section 7), is to reweight the BDT\(_j\) distribution of the background bin-by-bin using the MC-based ratio of the \( \geq 1 \) b-tag background to the background model. In this case the \( \geq 1 \) b-tag sample is preferred as it requires smaller corrections derived from MC simulation; the \( W + 1 \) jet is used as a cross check. The other approach is to split the background into bins of some variable within which the shapes of BDT\(_j\) distributions of the background model are close to those from the \( \geq 1 \) b-tag background. This approach, used in the matrix method cross check (Section 7.1), avoids using MC corrections, but assumes the data and MC simulation behave similarly as function of the binning variable.

7. Fits to BDT\(_j\) distributions

The contribution from \( t\bar{t} \rightarrow \ell + \tau + X \) signal is derived from the \( \geq 1 \) b-tag data sample by a \( \chi^2 \) fit to the OS-SS BDT\(_j\) distribution with a background template and a signal template. The parameters of the fit are the amount of background and the amount of signal. The shapes of the templates are fixed.

Two statistically independent background templates corrected by MC, as discussed in Section 6, are used: one derived from 0 b-tag data, the other (purely as a check) derived from the \( W + 1 \) jet data sample. The signal BDT\(_j\) templates for 0 b-tag and \( \geq 1 \) b-tag are derived from \( \tau \) leptons in \( t\bar{t} \) and \( Z \rightarrow \tau^+\tau^- \) MC simulation. Contributions to the BDT\(_j\) distributions from electrons passing the BDT\(_e\) cut cannot be distinguished from \( \tau \) leptons so they are treated as part of the signal.

The uncertainty on the background templates is determined by the numbers of data and MC simulated events. The signal template for the 0 b-tag control sample has a non-negligible statistical uncertainty (2% for \( \tau_1 \), 5% for \( \tau_2 \)) because of the low acceptance.

The fitting procedure was tested extensively with MC simulation before applying it to data. In the fits to the \( \geq 1 \) b-tag data, applying MC corrections to the 0 b-tag background template increases the statistical uncertainty because of the uncertainty due to the number of simulated events but raises the measured cross section by only 1%.

Fig. 1 shows the BDT\(_j\) (OS-SS) distributions of \( \ell + \tau \) in the \( \geq 1 \) b-tag sample: (a) \( \tau_1 \), (b) \( \tau_3 \). The normalisation of each template is derived from a fit to the data. The fitted contributions are shown as the light/red (signal), dashed/blue (background derived from 0 b-tag after applying MC corrections) and dark/black (total) lines. Shaded/blue bands are the statistical uncertainty of the background template.
fitted separately. The combined $\ell + t_1$ results are obtained by fitting the sum of the distributions. After adding $\ell + t_1$ and $\ell + t_3$ signals obtained from a $\chi^2$ fit to the combined $e + \tau$ and $\mu + \tau$ distributions and subtracting the small contributions to the signal from $Z\rightarrow\tau^+\tau^-$, $Z\rightarrow e^+e^-$ and $t\bar{t}\rightarrow e + \ell$ (given in Table 1), the result is $R_{\mu + \tau} = 0.453 \pm 0.032$ ($456 \pm 32$) and $R_{e + \tau} = 0.044 \pm 0.032$ ($430 \pm 50$) with the combined $R_{\tau} = 0.497 \pm 0.031$ ($456 \pm 50$). The uncertainty is from the fit only and does not include systematic uncertainties. The results are in good agreement with the values expected from MC simulation.

Table 2. Results of template fits to $\mu + \tau$, $e + \tau$ and the combined $BDT_j$ distributions. The combined results are obtained by fitting the sum of the $\mu + \tau$ and $e + \tau$ $BDT_j$ distributions. The first column gives the channel and the second the $\tau$ type. The third column shows the extracted signal (sum of $\tau$ leptons and electrons misidentified as $\tau$ leptons) with the background template derived from 0 $b$-tag data distributions. The fourth column shows the expected number of $t\bar{t}$ production cross section (164 pb).

<table>
<thead>
<tr>
<th></th>
<th>$N_{\text{signal}}$</th>
<th>$N_{\text{background}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu + \tau$</td>
<td>$t_1$ 490 ± 40</td>
<td>$t_2$ 135 ± 33</td>
</tr>
<tr>
<td></td>
<td>$t_3$ 166 ± 32</td>
<td>$r_2$ 116 ± 32</td>
</tr>
<tr>
<td>$e + \tau$</td>
<td>$t_1$ 440 ± 50</td>
<td>$t_2$ 116 ± 32</td>
</tr>
<tr>
<td></td>
<td>$t_3$ 200 ± 60</td>
<td>$r_2$ 166 ± 32</td>
</tr>
<tr>
<td>Combined</td>
<td>$t_1$ 930 ± 70</td>
<td>$t_2$ 260 ± 60</td>
</tr>
<tr>
<td></td>
<td>$t_3$ 260 ± 60</td>
<td>$r_2$ 260 ± 40</td>
</tr>
</tbody>
</table>

Fig. 3 shows the OS-SS distribution of the number of jets with at least one $b$-tag event adding all channels for two $BDT_j$ regions: $BDT_j < 0.7$, which is dominated by $t\bar{t}\rightarrow\ell +$ jets, and $BDT_j > 0.7$, in which the largest contribution is from $t\bar{t}\rightarrow\ell + \tau + X$. As expected, the multiplicity of jets peaks at four when $BDT_j < 0.7$ and three when $BDT_j > 0.7$ (the $\tau$ is counted as a jet). Fig. 4 shows the invariant mass of a selected jet with the $\tau$ candidate for $BDT_j < 0.7$ and $BDT_j > 0.7$ for events with a $\tau$ candidate and three or more jets. The selected jet is the highest $p_T$ untagged jet in events with more than one $b$-tag and the second highest $p_T$ untagged jet in events with one $b$-tag. The distribution shows clearly the presence of a $W$ decaying to two jets in the $BDT_j < 0.7$ region dominated by $t\bar{t}\rightarrow\ell +$ jets. The mass distribution in the $BDT_j > 0.7$ signal region is significantly broader as expected for $t\bar{t}\rightarrow\ell + \tau + X$. The signal and background shown in these figures are based on the fit using the 0 $b$-tag background template.

7.1. Check with matrix method

From Figs. 3 and 4 one can see that a $BDT_j > 0.7$ requirement separates well a region dominated by $t\bar{t}\rightarrow\ell +$ jets from a region dominated by $t\bar{t}\rightarrow\ell + \tau + X$. One can extract the signal from the same OS-SS $\geq 1$ $b$-tag sample used by the fit method via a matrix method. All $\tau$ candidates are labelled “loose”, and $\tau$ candidates with $BDT_j > 0.7$ are labelled “tight”. The probability that the loose $\tau$ candidates are also tight $\tau$ candidates, for both $\tau$ leptons and fake $\tau$ leptons, is defined as

$$P_{\text{fake}} = \frac{N_{\text{fake}}}{N_{\text{loose}}}$$

where the “real” subscript denotes $\tau$ lepton, the “fake” subscript denotes fake $\tau$ and $N$ is the number of $\tau$ candidates. The number of “tight” $\tau$ leptons is then given by

$$N_{\text{real}} = N_{\text{data}} \cdot \frac{P_{\text{real}}}{P_{\text{fake}}}$$

The value of $P_{\text{fake}}$ is measured utilising the OS-SS $BDT_j$ distributions from the background control samples; $P_{\text{real}}$ is derived from MC (including all processes that contribute a $\tau$ lepton or an electron misidentified as a $\tau$ lepton) and was tested using $Z\rightarrow\ell^+\ell^-$ events. This method uses the binning approach described in Section 6 to estimate the background. Values of $P_{\text{fake}}$ and $P_{\text{real}}$ are measured separately for three EM-fraction bins. The EM-fraction, the ratio of the energy measured in the EM calorimeter to the total $\tau$ candidate energy measured in the calorimeter, is an effective variable for splitting the data into regions where the shapes of MC OS-SS $BDT_j$ distributions from the background control samples are similar. Table 3 shows the number of signal events obtained with the matrix method using the background derived from the 0 $b$-tag data sample and from the $W + 1$ jet data sample. The numbers in each pair are in good agreement and consistent with the numbers obtained by fitting the OS-SS $BDT_j$ distributions.

7.2. Systematic uncertainty

Several experimental and theoretical sources of systematic uncertainty are considered. Lepton trigger, reconstruction and selection efficiencies are assessed by comparing the $Z\rightarrow\ell^+\ell^-$ events
Fig. 4. OS-SS invariant mass of jet and τ candidate for events with at least one b-tag. The jet is the highest \( p_T \) untagged jet in events with more than one b-tag and the second highest \( p_T \) untagged jet in events with one b-tag. The \( \mu + \tau \) and \( e + \tau \) channels have been summed together. The solid circles indicate data and the histograms indicate the expected signal and backgrounds. The normalisation of the expected signal and the backgrounds are based on the fit result. The uncertainty includes statistical and systematic contributions. The fraction of each background is expected signal and the backgrounds are based on the fit result. The uncertainty histograms indicate the expected signal and backgrounds. The normalisation of the background template derived from 0 b-tag data distributions. The fourth column shows the extracted signal with the background template derived from \( W + 1 \) jet. In order to compare the matrix method results to the fit results the number of signal events shown is \( \sum N_{\text{tag}}/E_{\text{real}} \), where \( E_{\text{real}} \) is the \( E_{\text{real}} \) averaged over the three EM-fraction bins. The uncertainties are statistical only.

Table 3
Number of signal events obtained with the matrix method for \( \mu + \tau \), \( e + \tau \) and the combined channels. The first column gives the channel and the second the \( \tau \) type. The third column shows the extracted signal with the background template derived from 0 b-tag data distributions. The fourth column shows the extracted signal with the background template derived from \( W + 1 \) jet. In order to compare the matrix method results to the fit results the number of signal events shown is \( \sum N_{\text{tag}}/E_{\text{real}} \), where \( E_{\text{real}} \) is the \( E_{\text{real}} \) averaged over the three EM-fraction bins. The uncertainties are statistical only.

<table>
<thead>
<tr>
<th>Source</th>
<th>( \mu + \tau )</th>
<th>( e + \tau )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mu ) (ID/Trigger)</td>
<td>( -1.1/1.5 )</td>
<td>-</td>
</tr>
<tr>
<td>( e ) (ID/Trigger)</td>
<td>-</td>
<td>( +2.9 )</td>
</tr>
<tr>
<td>JES</td>
<td>( -2.0/2.2 )</td>
<td>( -1.9/2.8 )</td>
</tr>
<tr>
<td>JER</td>
<td>( +1.0 )</td>
<td>( +1.2 )</td>
</tr>
<tr>
<td>ISR/FSR</td>
<td>( +4.8 )</td>
<td>( +3.5 )</td>
</tr>
<tr>
<td>Generator</td>
<td>( +0.7 )</td>
<td>( +0.7 )</td>
</tr>
<tr>
<td>PDF</td>
<td>( +2.0 )</td>
<td>( +2.1 )</td>
</tr>
<tr>
<td>b-tag</td>
<td>( -7.7/9.0 )</td>
<td>( -7.5/8.9 )</td>
</tr>
<tr>
<td>( \tau ) ID</td>
<td>( -3.0/3.2 )</td>
<td>( -2.7/3.0 )</td>
</tr>
<tr>
<td>( \tau ) ID</td>
<td>( -3.1/3.4 )</td>
<td>( -2.9/3.2 )</td>
</tr>
</tbody>
</table>

The uncertainty in the kinematic distributions of the \( tt \) signal events gives rise to systematic uncertainties in the signal acceptance, with contributions from the choice of generator, the modelling of initial- and final-state radiation (ISR/FSR) and the choice of the PDF set. The generator uncertainty is evaluated by comparing the MC@NLO predictions with those of POWHEG [36–38] interfaced to either HERWIG or PYTHIA. The uncertainty due to ISR/FSR is evaluated using the AcerMC generator [39] which relies on the MADGRAPH package [40] interfaced to the PYTHIA shower model, and by varying the parameters controlling ISR and FSR in a range consistent with experimental data [19]. Finally, the PDF uncertainty is evaluated using a range of current PDF sets: NNPDF, CTEQ, and MSTW [41–43]. Each one comes with a set of error PDFs, the RMS of the variations was taken as the PDF uncertainty. The dominant uncertainty in this category of systematic uncertainties is the modelling of ISR/FSR.

The \( \tau \) ID uncertainty is derived from a template fit to a \( Z \rightarrow \tau^+ \tau^- \) data sample selected with the same \( \mu \) and \( \tau \) candidate requirements as the sample for this analysis, but with fewer than two jets and \( m_\tau < 20 \) GeV to remove \( W + \) jets events. The fit relies on the \( W + 1 \) jet data sample for a background template and \( Z \rightarrow \tau^+ \tau^- \) MC events for a signal template. The uncertainty includes the statistical uncertainty of the data samples, the uncertainty in the \( Z/\gamma^* \) cross section measured by ATLAS [44] (excluding the luminosity uncertainty) and the jet energy scale uncertainty. It also includes the uncertainty on the number of misidentified electrons (<0.5%, determined from \( Z \rightarrow e^+e^- \) data).

The effect of these variations on the final result is evaluated by varying each source of systematic uncertainty by \( \pm 1\sigma \), applying the selection criteria to obtain new signal and background templates and recalculating the cross section.
The uncertainties obtained for the fit method using the 0 b-tag background template are shown in Table 4. The systematic uncertainties for the matrix method are very similar. The uncertainty on the measured integrated luminosity is 3.7% [45]. This translates into a 3.5% uncertainty on the cross section.

8. Measuring the $t\bar{t}$ cross section

The cross section is derived from the number of observed OS-SS signal events in the $1 b$-tag data sample assuming the only top quark decay mode is $t \to Wb$, and subtracting from that number the small contribution from $t\bar{t} \to e+\ell$ (from electrons faking $\tau$ leptons) and $\tau$ leptons from $Z \to \tau^+\tau^-$ (Table 1). The systematic uncertainties are estimated as the quadratic sum of all uncertainties given in Table 4, which includes the uncertainty from the subtraction.

The results are given separately for $\tau_1$ and $\tau_2$ and then combined (weighted by their statistical uncertainty and assuming all systematic uncertainties other than from $\tau$ ID are fully correlated). The results using the 0 b-tag background template are shown in Table 5.

The results for the $\mu + \tau$ and $e + \tau$ channels are combined taking into account the correlated uncertainties using the BLUE [Best Linear Unbiased Estimator] technique [46]. Combining them does not improve the systematic uncertainty as the systematic uncertainties are almost 100% correlated.

The results for each lepton type are:

$$\mu + \tau: \quad \sigma_{\ell \ell} = 186 \pm 15 \text{ (stat.)} \pm 20 \text{ (syst.)} \pm 7 \text{ (lumi.)} \text{ pb},$$

$$e + \tau: \quad \sigma_{\ell \ell} = 187 \pm 18 \text{ (stat.)} \pm 20 \text{ (syst.)} \pm 7 \text{ (lumi.)} \text{ pb}.$$

Combining both channels one obtains:

$$\sigma_{\ell \ell} = 186 \pm 13 \text{ (stat.)} \pm 20 \text{ (syst.)} \pm 7 \text{ (lumi.)} \text{ pb}.$$

To check the fit measurements, the cross sections can be calculated using the matrix method and the results obtained with the $W + 1$ jet background to minimise the correlation with the fit results. The combination of the matrix method and the fit results with the BLUE method shows they are compatible at the 45% and 10% confidence level for $\mu + \tau$ and $e + \tau$, respectively.

9. Conclusions

The cross section for $t\bar{t}$ production in $pp$ collisions at 7 TeV has been measured in the $\mu + \tau$ and the $e + \tau$ channels in which the $\tau$ decays hadronically. The number of $\tau$ leptons in these channels has been extracted using multivariate discriminators to separate $\tau$ leptons from electrons and jets misidentified as hadronically decaying $\tau$ leptons. These numbers were obtained by fitting the discriminator outputs and checked with a matrix method. Combining the measurements from $\mu + \tau$ and $e + \tau$ events, the cross section is measured to be

$$\sigma_{\ell \ell} = 186 \pm 13 \text{ (stat.)} \pm 20 \text{ (syst.)} \pm 7 \text{ (lumi.)} \text{ pb},$$

in good agreement with the cross section measured by ATLAS in other channels [1,2], with the cross-section measurements by the CMS Collaboration [3–5,7] and with the SM prediction, 164$^{+11}_{-10}$ pb [25].

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, Austria; ANAS, Azerbaijan; SSTC, Belarus; CINPq and FAPESP, Brazil; NSERC and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MZMT CR, MPO CR and VSC CR, Czech Republic; DEFA, DFG and GSI, Germany; INFN, Italy; JINR, Russia; Kحمد and KSU, Jordan; CNRST, Morocco; FOM and NWO, Netherlands; RCN, Norway; MNISW, Poland; GRICES and FCT, Portugal; MEN-RS, Tunisia; RCN, Norway; MNISW, Poland; GRICES and FCT, Portugal; MURST (MECTS), Romania; MES of Russia and ROSATOM, Russian Federation; JINR, Russia; SAR, Slovakia; ARRS and MIZT, Slovenia; DST/NRF, South Africa; MINECO, Spain; SRC and Wallenberg Foundation, Sweden; SER, SNSF and Cantons of Bern and Geneva, Switzerland; NSF, Taiwan; TAEK, Turkey; STFC, the Royal Society and Leverhulme Trust, United Kingdom; DOE and NSF, United States of America.

The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN and the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NLT1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (UK) and BNL (USA) and in the Tier-2 facilities worldwide.

Open access

This article is published Open Access at sciencedirect.com. It is distributed under the terms of the Creative Commons Attribution License 3.0, which permits unrestricted use, distribution, and reproduction in any medium, provided the original authors and source are credited.

References
