Calibration of the ATLAS $b$-tagging algorithm in $t\bar{t}$ semileptonic events

The ATLAS Collaboration

The calibration of the ATLAS $b$-tagging algorithm in $t\bar{t}$ semileptonic events is presented. The calibration uses reconstructed $t\bar{t}$ candidate events collected by the ATLAS detector in 36.1 fb$^{-1}$ proton-proton collisions provided by the LHC at 13 TeV during 2015 and 2016, with a final state containing one charged lepton, missing transverse momentum and at least four jets. The $b$-tagging efficiencies are measured as a function of the transverse momentum and the pseudorapidity of the jets. The results for the data-to-simulation scale factors are presented in an extended transverse momentum region compared to $b$-tagging calibrations based on dileptonic $t\bar{t}$ events.
1 Introduction

The identification of jets containing $b$-hadrons, referred to as $b$-tagging, plays a crucial role in the physics programme of the ATLAS experiment [1]. This note presents a measurement of the data-to-simulation $b$-tagging efficiency scale factors $\kappa = \epsilon_{b}^{\text{data}} / \epsilon_{b}^{\text{sim}}$, where $\epsilon_{b}^{\text{data}}$ is the efficiency measured in data, while $\epsilon_{b}^{\text{sim}}$ represents the efficiency predicted by the simulation, of the ATLAS $b$-tagging algorithm MV2c10 [2]. The results are presented as a function of the transverse momentum ($p_T$) and the pseudorapidity ($\eta$) of the jets. The $b$-tagging efficiency is measured for several working points that are defined by sets of selection criteria applied on the output of the $b$-tagging algorithm in order to provide a specific $b$-jet tagging efficiency. These scale factors can be used per jet to correct the predicted yield after the application of $b$-tagging requirements, therefore physics analyses benefit from a precise measurement of the $b$-tagging efficiency. The correction of the $b$-tagging efficiency in the simulation via a scale factor is referred to as a calibration of the $b$-tagging efficiency.

The $b$-tagging calibrations in ATLAS are performed using top quark pair events. The decay of a $t\bar{t}$ pair provides a large sample of jets originating from the hadronisation of $b$-quarks, i.e. $b$-jets, given the fact that the $t \rightarrow Wb$ branching fraction is close to 1 [3]. In particular, most $b$-tagging calibrations are based on dilepton $t\bar{t}$ events, since they provide a pure $b$-jet sample with low background contamination. The choice of this decay channel therefore allows the systematic uncertainties related to the measurement to be reduced. The methods used on dilepton $t\bar{t}$ events to calibrate the ATLAS $b$-tagging algorithms are the combinatorial likelihood method (dilepton $t\bar{t}$ PDF) and the tag-and-probe method (dilepton $t\bar{t}$ T&P) [4], both of which have been proven to provide precise calibrations.

During Run 1, a $t\bar{t}$ based $b$-tagging calibration was performed on lepton+jets ($\ell$+jets) events, using a tag-and-probe method based on single lepton $t\bar{t}$ events ($t\bar{t}$ SL T&P) [5]. The use of semileptonic $t\bar{t}$ decays gives a larger sample of jets than from the dilepton sample, allowing data-to-simulation scale factors to be measured with higher statistical precision and the measurement to be extended to higher jet $p_T$. The measurement presented in this note is performed using 2015 and 2016 Run 2 data samples applying the same tag-and-probe method described in Ref. [5]. The probe jets are selected in $t\bar{t}$ candidate events with a final state containing exactly one charged lepton and at least four jets.

This note is organized as follows. Section 2 briefly describes the ATLAS detector, while Section 3 gives details of the Monte Carlo (MC) samples and the collected data used in this calibration. Section 4 presents the objects and all the selection requirements as well as the reconstruction procedure applied on the selected events. Section 5 describes the method used to measure the $b$-tagging efficiencies in data and in Section 6 the systematic uncertainties are described. The results of the calibration of MV2c10 are shown in Section 7 and conclusions are presented in Section 8.

---

1 ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the centre of the detector and the $z$-axis along the beam pipe. The $x$-axis points from the IP to the centre of the LHC ring, and the $y$-axis points upwards. Cylindrical coordinates ($r, \phi$) are used in the transverse plane, $\phi$ being the azimuthal angle around the $z$-axis. The pseudorapidity is defined in terms of the polar angle $\theta$ as $\eta = -\ln \tan(\theta/2)$. Angular distance is measured in units of $\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2}$. 

---
2 ATLAS detector

The ATLAS detector [1] at the LHC is a multi-purpose particle detector with a forward-backward symmetric cylindrical geometry and a near $4\pi$ coverage in solid angle. It consists of an inner tracking detector surrounded by a thin superconducting solenoid providing a 2 T axial magnetic field, electromagnetic and hadron calorimeters, and a muon spectrometer.

The inner tracking detector (ID) covers the pseudorapidity range $|\eta| < 2.5$. It consists of silicon pixel, silicon micro-strip, and transition radiation tracking detectors. Among them, the pixel detectors are crucial for b-jet identification. A new inner pixel layer, the Insertable B-Layer [6] (IBL), was added before the start of Run 2, at a mean sensor radius of 3.2 cm from the beam-line.

Outside the ID, the lead/liquid-argon (LAr) sampling calorimeters provide electromagnetic (EM) energy measurements with high granularity. A hadron (iron/scintillator-tile) calorimeter covers the central pseudorapidity range ($|\eta| < 1.7$). The end-cap and forward regions are instrumented with LAr calorimeters for both EM and hadronic energy measurements up to $|\eta| = 4.9$. The muon spectrometer surrounds the calorimeters and is based on three large air-core toroid superconducting magnets with eight coils each. Its bending power is in the range from 2.0 to 7.5 Tm. It includes a system of precision tracking chambers and fast detectors for triggering.

A two-level trigger system, using custom hardware followed by a software-based level, is used to reduce the event storage rate to a maximum of around 1 kHz.

3 Data and Monte Carlo samples

Events were selected from the data samples collected in 2015 and 2016 with the ATLAS detector at the LHC using $pp$ collisions at $\sqrt{s} = 13$ TeV. The corresponding total integrated luminosity is $36.1 \text{ fb}^{-1}$ ($3.2 \text{ fb}^{-1}$ in 2015 and $32.9 \text{ fb}^{-1}$ in 2016).

The $t\bar{t}$ process was modelled with a MC sample generated by Powheg+Pythia 8 [7, 8], using the A14 set of tunable parameters [9] for the showering and the NNPDF 3.0 [10] PDF set at next-to-leading order (NLO). The sample was normalized to the next-to-next-to-leading order (NNLO) cross-section computed with with top++ 2.0 [11] that includes resummation of next-to-next-to-leading logarithmic (NNLL) soft gluon terms. The sample did not include full hadronic decay, but only dilepton and semileptonic events. Single top-quark production in the $s$- and $t$-channel or in association with a $W$ boson was simulated by Powheg+Pythia 6 [12], using the Perugia2012 tune [13] for the showering and the CT10 [14] PDF set.

The production of $W$+jets, $Z$+jets and diboson ($WW$, $WZ$ and $ZZ$) were simulated using MC samples generated by Sherpa 2.2.0 [15], using NNPDF 3.0 PDF set at NNLO. For the specific cases of $W$+jets and $Z$+jets production, filters on vector boson $p_T$ and heavy flavour content have been used to guarantee enough statistics in the full phase space.

All simulated event samples were passed through the full ATLAS detector simulation [16, 17]. The simulated events were overlaid with additional inelastic $pp$ interactions that were simulated with Pythia 8 [7] in order to match the pile-up conditions observed in ATLAS data.
4 Selection of $t\bar{t}$ candidates

4.1 Object definition

Primary vertices are formed from reconstructed tracks, that satisfy the criteria described in Ref. [18], which are spatially compatible with the interaction region. The primary vertex is chosen to be the one with highest $\sum s_{\text{tracks}} p_{\text{T}}^2$ and at least two tracks with $p_{\text{T}} > 0.4 \text{ GeV}$ associated to the vertex.

Electron candidates are reconstructed by matching tracks in the ID to energy deposits in the EM calorimeter and they must satisfy a “tight” likelihood-based identification criterion [19]. The electron candidates are required to have a transverse momentum $p_{\text{T}} > 25 \text{ GeV}$ and a pseudorapidity $|\eta| < 2.47$, excluding the transition region between the barrel and the end-cap calorimeters (1.37 < |$\eta$| < 1.52). The associated track must have a longitudinal impact parameter $|z_0 \sin \theta| < 0.5 \text{ mm}$ ($z_0$ is the distance of closest approach to the primary vertex in the longitudinal plane and $\theta$ is the polar angle) and a transverse impact parameter $(d_0)$ significance $|d_0/\sigma(d_0)| < 5$, where $d_0$ is measured with respect to the beam line and $\sigma(d_0)$ is its resolution. Isolation requirements based on calorimeter and tracking quantities are used to reduce the background from non-prompt and fake electrons [20]. The isolation criteria are $p_{\text{T}}$- and $\eta$-dependent and ensure an efficiency of 90% for electrons with $p_{\text{T}}$ of 25 GeV and 99% efficiency for electrons with $p_{\text{T}}$ of 60 GeV. These efficiencies are measured using electrons from Z boson decays [21]. The track isolation cone size is given by the minimum of $\Delta R = 10 \text{ GeV}/p_{\text{T}}^{\text{el}}$ and $\Delta R = 0.2$, where $p_{\text{T}}^{\text{el}}$ is the electron $p_{\text{T}}$. Thus, the cone radius increases with decreasing $p_{\text{T}}$ up to a maximum of 0.2.

Muon candidates are identified by matching tracks in the muon spectrometer with tracks in the ID [22]. The track $p_{\text{T}}$ is determined through a global fit of the hits which takes into account the energy loss in the calorimeters. Muons are required to have $p_{\text{T}} > 25 \text{ GeV}$ and $|\eta| < 2.5$. The associated track must satisfy $|z_0 \sin \theta| < 0.5 \text{ mm}$ and $|d_0/\sigma(d_0)| < 3$. For this measurement only muon candidates that fulfill the medium ID [22] criteria are considered. Isolation requirements similar to those used for electrons are applied, except that the maximum $\Delta R$ in this case is 0.3. The resulting isolation efficiencies are the same as for electrons.

The jets are reconstructed by applying the anti-$k_t$ algorithm, as implemented in the FASTJet [23] package, to topological clusters made from adjoining calorimeter energy deposits and a distance parameter of $R = 0.4$. The jets are required to be matched to the primary vertex [24] and to have a transverse momentum of at least 20 GeV and an absolute value of the pseudorapidity below 2.5.

In order to avoid assigning single detector responses to more than one physics object, a sequential overlap removal procedure is applied. If the angular separation, $\Delta R(\text{electron}, \text{jet})$, between any selected jet and an electron candidate is smaller than 0.2, the jet is rejected. After that, if an electron is found to be close to a jet such that $\Delta R(\text{electron}, \text{jet}) < 0.4$, the electron is rejected. Selected muons that satisfy $\Delta R(\mu\text{on}, \text{jet}) < 0.04 + 10 \text{ GeV}/p_{\text{T}}^{\mu}$ are rejected, if the jet has at least 3 tracks originating from the primary vertex. Jets with less than 3 tracks that overlap with a muon are rejected.

The $E_{\text{T}}^{\text{miss}}$ vector is computed from the sum of the transverse momenta of the reconstructed calibrated objects (electrons, jets and muons) together with the transverse energy deposited in the calorimeter cells, calibrated using tracking information, not associated with these objects [25]. The contribution from muons is added using their momenta. To avoid double-counting of energy, the muon energy loss in the calorimeters is subtracted in the $E_{\text{T}}^{\text{miss}}$ calculation.
4.2 Jet truth labelling

The procedure used to assign a flavour label to jets in simulated events is based on an angular matching between the generator level particles and the reconstructed jets. If a jet has a $b$-hadron with a $p_T > 5$ GeV inside a cone of radius $\Delta R = 0.3$ around its axis, the jet is labelled as a $b$-jet. If two jets include the same $b$-hadron, only the closest jet will be labelled as $b$-jet. The same procedure is applied for $c$-hadrons and then for $\tau$ if jet with no association to a $b$-hadron is possible. Finally, a jet is labelled as light-flavoured if no association to one of the previous particles was possible.

4.3 $b$-tagging algorithms

The $b$-tagging algorithms used in ATLAS exploit the long lifetime of hadrons containing $b$-quarks and they are based either on the track impact parameters (IP2D or IP3D [2]) or on the properties of displaced vertices reconstructed inside a jet. For secondary vertex reconstruction, the used algorithms are the iterative vertex finder SV1 [26] and the JetFitter algorithm [27] (the latter goes beyond the secondary vertex and it also searches for tertiary vertices).

The MV2c10 algorithm [2] is a Boosted Decisions Tree (BDT) algorithm that uses the ROOT Toolkit for Multivariate Data Analysis (TMVA) [28] and combines the input of the IP3D, SV1 and JetFitter algorithms. The input variables can be found in Ref. [29]. The jet $p_T$ and $\eta$ are included in the training variables in order to exploit the correlations with other variables. The $b$-jet $p_T$ and $\eta$ distributions are reweighted to match the combined $c$-jet and light-flavour jet spectrum to avoid the possibility to use any discrepancy between the kinematic spectra of $b$-jet and background jets as a discriminant variable. The MV2c10 algorithm was trained on a subset of events from a simulated $t\bar{t}$ sample produced with Powheg + Pythia 6. This sample includes all the $t\bar{t}$ decay channels. The fraction of jets originating from $c$-quarks, the $c$-jets, used in the training phase of the MV2c10 algorithm is 7% such that the training is performed assigning $b$-jets as signal and a mixture of 93% jets originating from light quarks and 7% $c$-jets as background.

4.4 Event preselection

Top quark pair candidate events in the semileptonic decay channel are selected by using single-lepton triggers and requiring the candidate lepton to be trigger matched [30]. Exactly one electron or one muon that passes the full object definition requirements has to be identified within the acceptance of the detector and its $p_T$ has to exceed 25 GeV for 2015 data and 27 GeV for 2016 data.

In order to reduce the background due to non-prompt and fake leptons, which will be described in Sec. 4.6, the value of the missing transverse momentum $E_T^{\text{miss}}$ is required to be at least 20 GeV. An additional requirement is applied on the transverse mass of the $W$ boson candidate

$$m_T^W = \sqrt{2p_T^l E_T^{\text{miss}} (1 - \cos \phi_{l\nu})},$$

where $\phi_{l\nu}$ corresponds to the azimuthal angle between the lepton candidate and the $E_T^{\text{miss}}$ vector. For both the electron and the muon channel, the sum of $m_T^W$ and $E_T^{\text{miss}}$ has to be larger than 60 GeV.
Candidate events are also required to contain at least four jets. Additional quality requirements are applied removing events that contain noise bursts in the LAr calorimeter or any jet with $p_T > 20$ GeV matched to noise in the calorimeter or out-of-time activity with respect to the $pp$ collision.

### 4.5 Event reconstruction

The $b$-tagging efficiencies and the corresponding data-to-simulation scale factors are measured in a sample of $b$-jets, exploiting $tt$ events. The reconstruction of top quark pair decays is performed by using a kinematic fit procedure on events containing a single charged lepton and at least four jets. The reconstruction of the neutrino four-momentum, needed in the leptonic side of the top-quark decay reconstruction, is obtained from applying a $W$ boson mass constraint on the invariant mass of the charged-lepton–neutrino system for the $z$ component, while the $x$ and $y$ components of the neutrino momentum are set to the corresponding components of the missing transverse momentum. If the resulting quadratic equation does not have a real solution, the missing momentum vector is rotated including a parameter in the expression of the lepton-neutrino system that is minimized in order to find a real solution. In case of two real solutions of the quadratic equation, both solutions are tested in the reconstruction and the one that minimizes the $\chi^2$ value is chosen.

The $\chi^2$ minimisation procedure is based on constraints on the top quark and $W$ boson masses and on the kinematics of a $tt$ event. All the reconstructed objects that pass the selection are taken into account and the permutation which leads to the smallest $\chi^2$ value is chosen, where the $\chi^2$ is defined as:

$$
\chi^2 = \left[ \frac{m_{jj} - m_W}{\sigma_W} \right]^2 + \left[ \frac{m_{jjj} - m_{t_W} - m_{W}}{\sigma_{t_W}} \right]^2 + \left[ \frac{m_{b\ell\nu} - m_{t_\ell}}{\sigma_{t_\ell}} \right]^2 + \left[ \frac{(p_T^{j Bj} - p_T^{b\ell\nu}) - (p_T^{t_\ell} - p_T^{t_\ell})}{\sigma_{diff_{pT}}} \right]^2 .
$$

In the $\chi^2$ definition $t_W$ and $t_\ell$ refer to the hadronically and semileptonically decaying top quarks, respectively. The $m_W$, $\sigma_W$, $m_{t_W}$, $\sigma_{t_W}$, $m_{t_\ell}$, $\sigma_{t_\ell}$, $p_T^{t_\ell}$, $p_T^{t_\ell}$ and $\sigma_{diff_{pT}}$ parameters contained in Eq. 2 are kept constant during the minimisation procedure. Their values are obtained from the simulation following the procedure described in Ref. [31]. The first term is the constraint from the hadronically decaying $W$ boson and $m_W$ and $\sigma_W$ represent, respectively, the mean and the standard deviation of the invariant mass distribution related to the hadronically decaying $W$ boson. The second term corresponds to the hadronically decaying top quark. Since the mass of the hadronically decaying $W$ boson, $m_{jj}$, and the mass of the hadronically decaying top quark, $m_{jjj}$, are heavily correlated, the hadronically decaying $W$-boson was subtracted from $m_{t_W}$ to decouple this term from the previous one. The third term represents the semileptonically decaying top quark and it includes the information of the $b$-jet-lepton-neutrino system $m_{b\ell\nu}$ and the mass and the standard deviation ($m_{t_\ell}$ and $\sigma_{t_\ell}$) of the leptonic top quark. In the last term, the leptonic and hadronic top quark transverse momenta, $p_T^{b\ell\nu}$ and $p_T^{t_\ell}$, and the standard deviation of the difference between them, $\sigma_{diff_{pT}}$, are used to constraint the top quark transverse momenta to be balanced.

In order to reduce the background contamination, the candidate events are required to satisfy $\log_{10}(\chi^2) < -0.9$. In addition, the jet assigned to the leptonic quark decay is required to be $b$-tagged, while the two jets assigned to the hadronic $W$ boson decay are required not to be $b$-tagged (this requirement will be called anti-tag in the following). The MV2c10 algorithm is used for the $b$-tagging assessment at a working point that corresponds to an overall efficiency of 70% in the same simulated $tt$ sample used for the training of the $b$-tagging algorithm, described in 4.3. All of the jets, except the jet assigned to the $b$-jet candidate on the hadronic side of the events, are required to have a transverse momentum $p_T > 20$ GeV in order to improve
the purity of the selection. The $b$-tagging efficiencies are measured in data on the $b$-jet candidates on the hadronic side of the $t\bar{t}$ events and this jet sample will be considered the probe $b$-jet sample. The hadronic $b$-jet candidate is chosen in order to probe also the hadronic side of the $t\bar{t}$ decay, which have larger jet multiplicity with respect to the leptonic side. A comparison between the results on the hadronic $b$-jet candidate and the dileptonic $t\bar{t}$ calibrations has been performed to test the consistency. In addition, the leptonic $b$-jet has a better resolution, so the inclusion in the $\chi^2$ minimization leads to better reconstruction of the $t\bar{t}$ system.

4.6 Background estimation

The background contributions for the production of single-top quarks, $Z$+jets and diboson are estimated by using MC simulation. When the precision of the MC simulation is not adequate, data-driven techniques are used, as in the case of $W$ boson production in association with jets and the multijet background originating from jets mimicking the signature of charged leptons. Multijet production processes have a large cross section and have a signature similar to the one of $\ell$+jets, due to jets misidentified as prompt leptons (fake leptons) or due to semileptonic decays of heavy-flavour hadrons (non-prompt real leptons). The background contamination due to the non-prompt and fake lepton background, which will be indicated as fake lepton background, is estimated using a matrix method. The fake-lepton efficiency is extracted from control regions dominated by the multijet background, subtracting the real-lepton contribution using MC simulation. Thereal-leptonefficiencyisextractedusingatag-and-probetaotechniqueonleptonscomingfromZbosondecays. The corresponding formalism and the details on the extraction of the efficiencies for real and fake leptons, as well as the definitions of the control regions, are extensively explained in Ref. [32].

The estimation of the $W$+jets background is performed using a combination of MC simulations and data-driven techniques. The MC generator is used to estimate the contribution from the $W$+jets process. The normalisation and the heavy-flavour fractions in this process, which are affected by large theoretical uncertainties, are determined from data. The overall $W$+jets normalisation is obtained by exploiting the underlying charge asymmetry of the $W$+jets production. This asymmetry is evaluated using the MC simulation with the assumption that other processes are charge symmetric. A small contamination coming from single-top quarks, $t\bar{t}V$ and $WZ$ events is subtracted using the MC information. The total number of $W$+jets events in data, $N_{W^+} + N_{W^-}$, is given by:

$$N_{W^+} + N_{W^-} = \frac{r_{\text{MC}} + 1}{r_{\text{MC}} - 1} (D_{\text{corr}+} - D_{\text{corr}-})$$  

(3)

where $r_{\text{MC}}$ is the ratio given by MC simulation of the number of $W$+jets events with a positively charged lepton to the events with a negatively charged lepton and $D_{\text{corr}+(-)}$ is the number of events with a positively (negatively) charged lepton. The relative fractions of $W$-bosons associated to light-flavour and heavy-flavour jets are estimated in a dedicated control region enriched in $W$+jets events. A system of three equations is solved to obtain correction factors for MC based fractions. The correction factors are then extrapolated from the dedicated control region enriched in $W$+jets events to the region in which the measurement is performed. The scale factors and further information on this method (e.g. definition of the control regions) can be found in Ref. [31].
4.7 Event yields and validation plots

The event yields after the full selection and reconstruction chain are displayed in Table 1 for data, simulated signal and backgrounds. The number of events observed are 331406 (329792) in the electron (muon) channel, while the estimate of the contribution for signal and background processes predicts approximately 330000 (325000) events. The dominant background contributions originate from the non-prompt and fake lepton background, single top quarks and the production of $W$ boson in association with jets. Smaller contributions are due to $Z$+jets, diboson and the production of $t\bar{t}$ in association with a vector boson.

In general, the predictions from the simulation and the observation in data are compatible within the total uncertainties for both signal and background processes.

<table>
<thead>
<tr>
<th>Source</th>
<th>$N_{e+\text{jets}}$</th>
<th>$N_{\mu+\text{jets}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$tt$</td>
<td>283000 ± 50000</td>
<td>280000 ± 40000</td>
</tr>
<tr>
<td>$t\bar{t}+V$</td>
<td>730 ± 50</td>
<td>700 ± 40</td>
</tr>
<tr>
<td>Single Top</td>
<td>10000 ± 1000</td>
<td>10000 ± 1000</td>
</tr>
<tr>
<td>$W$+jets</td>
<td>15000 ± 2400</td>
<td>14000 ± 3000</td>
</tr>
<tr>
<td>$Z$+jets</td>
<td>4500 ± 1700</td>
<td>2300 ± 500</td>
</tr>
<tr>
<td>Diboson</td>
<td>470 ± 70</td>
<td>430 ± 70</td>
</tr>
<tr>
<td>Fake leptons</td>
<td>15000 ± 8000</td>
<td>18000 ± 10000</td>
</tr>
<tr>
<td>Total prediction</td>
<td>330000 ± 50000</td>
<td>325000 ± 50000</td>
</tr>
<tr>
<td>Observed</td>
<td>331406</td>
<td>329792</td>
</tr>
</tbody>
</table>

Table 1: Predicted yields after the full selection and reconstruction chain for signal and background contributions together with the observed yields from the combined 2015 and 2016 data sample. The uncertainties include the combined statistical and systematic uncertainties.

Data-to-prediction comparisons for the relevant properties of the selected probe jets, which are the $b$-jet candidate on the hadronic side of the reconstructed event, are displayed in Fig. 1. The transverse momenta and the absolute value of the pseudorapidities of these jets are shown. These distributions are obtained after applying the full selection and reconstruction requirements to the candidate events, including the tag and anti-tag requirements and the cut on the corresponding $\log_{10}(\chi^2)$ value, shown in Fig. 2. The validation plots show good agreement between the observations in data and the predictions for almost all of the observables. The difference, between data and prediction, observed in the slope of the $p_T$ spectrum of the probe jet is due to a mismodelling of the $t\bar{t}$ system that is covered by the total systematic uncertainty. The $b$-tagging calibration results are presented in the following section based on the combination of both electron and muon channels and years.
Figure 1: Distribution of the transverse momentum (a), (b) and the absolute pseudorapidity (c), (d) of the selected probe jets displayed separately for the $e$+jets (left column) and the $\mu$+jets (right column) channels. The simulated MC samples are normalised according to their predicted cross-sections for an integrated luminosity of 3.2 fb$^{-1}$ for 2015 data and of 32.9 fb$^{-1}$ for 2016 data. Data-to-prediction ratios are shown at the bottom of each plot. The shaded area indicates the total statistical and systematic uncertainties.
Figure 2: Distribution of the minimum $\chi^2$ values of the $t\bar{t}$ kinematic reconstruction applied to events passing all the selection requirements displayed separately for the $e$+jets (a) and the $\mu$+jets (b) channels. The simulated MC samples are normalised according to their predicted cross-sections for an integrated luminosity of 3.2 fb$^{-1}$ for 2015 data and of 32.9 fb$^{-1}$ for 2016 data. Data-to-prediction ratios are shown at the bottom of each plot. The shaded area indicates the total statistical and systematic uncertainties.
5 Measurement of the b-tagging efficiency in data

5.1 Tag-and-probe method

The measurement of the b-tagging efficiency in data is performed on the hadronic top candidate of the reconstructed \( t\bar{t} \) decay, testing the b-jet candidate sample. The probe jet sample consists of all jets in data and MC passing the requirements discussed in Sec. 4 and assigned to the hadronic b-jet by the kinematic fitter described in Sec. 4.5. The contamination of the jet sample due to the fraction of c- and light-flavour jets has to be taken into account in order to properly measure the b-tagging efficiency in data. The equation used to calculate the b-tagging efficiency in data is

\[
\varepsilon_b = \frac{1}{f_{b\text{-jets}}} \cdot (f_{\text{tag}} - e_c f_{c\text{-jets}} - e_l f_{l\text{-jets}} - e_{\text{fake}} f_{\text{fake}})
\]

where \( f_{b\text{-jets}}, f_{c\text{-jets}} \) and \( f_{l\text{-jets}} \) represent the fractions of b-, c- and light flavour jets in the sample of probe jets, while \( f_{\text{fake}} \) denotes the fraction of jets coming from the non-prompt lepton background estimated from data. The fraction of jets that are b-tagged at a specific working point is represented by \( f_{\text{tag}} \), while the flavour fractions \( f_{b\text{-jets}}, f_{c\text{-jets}} \) and \( f_{l\text{-jets}} \) are taken from the simulation. The mistag efficiencies \( e_c \) and \( e_l \) for c- and light-flavour jets are taken from the simulation and corrected using the most recent results obtained by the measurements in the calibrations using c-jets in \( t\bar{t} \) events [33] and the negative tag method [34]. The mistag efficiency due to the jets coming from non-prompt lepton background \( e_{\text{fake}} \) is extracted from a control region in data described in Sec. 5.4.

5.2 Estimated b-tagging efficiencies

Fig. 3 shows the estimated tagging efficiencies (estimated from simulation for b-, c- and light-flavour jets, and data for \( e_{\text{fake}} \)) in the probe jet sample for a 85% working point of the MV2c10 b-tagging algorithm. For this particular working point the efficiencies are approximately flat as a function of the transverse momentum and the absolute value of the pseudorapidity of the probe jet, as shown in Fig. 3(a) and 3(b). Fig. 4 shows the measured b-tagging efficiencies for the 60% working points. A trend in the tagging efficiencies becomes evident as the working point decreases. In particular, the c- and light-flavour tagging efficiency values are significantly smaller than the b-tagging and fake tagging efficiencies.

5.3 Flavour composition of the selected jet sample

The flavour composition of the probe jet sample is crucial for the measurement of the b-tagging efficiency performed using Eq. (4). The accuracy of the estimate of the flavour fractions depends on the quality of the kinematic reconstruction of \( t\bar{t} \) candidate events. The flavour fractions and, as a consequence, the measured b-tagging efficiencies are affected by systematic uncertainties due to mis-reconstruction effects.

The estimated fraction of b-, c- and light-flavour jets within the probe jet sample obtained from MC simulation together with the fraction of jets coming from the non-prompt lepton background derived in data are shown in Fig. 5(a) as a function of the probe jet \( p_T \). For jets with a transverse momentum between 20 GeV and 30 GeV the fraction of b-jets is 30% and increases to 70% for jets with transverse momentum above 200 GeV. The second contribution in terms of flavour fraction is due to light-flavour jets, which is 60% in the region between 20 GeV and 30 GeV and decreases as a function of the jet \( p_T \) to 20%.
Figure 3: $b$-tagging efficiencies for $b$-, $c$- and light-flavour jets (estimated from the simulation) and for jets coming from the non-prompt and fake lepton background (obtained from a control region in data) of the MV2c10 algorithm at a working point providing 85% efficiency. The efficiencies of the various jet flavours are shown as a function of the jet $p_T$ (a) and $|\eta|$ (b). The shaded areas represent the total statistical and systematic uncertainties.

The contamination due to jets coming from the non-prompt lepton background is below 2%, while the fraction of $c$-jets varies between 5% and 7%. The contamination of the probe jet sample due to $c$- and the light-flavour jets is caused in large part by the assignment of the wrong permutation of the jets to the decay products of the top quark.

5.4 Measurement of the $b$-tagging efficiency for jets coming from the non-prompt lepton background

The estimate of the $b$-tagging efficiencies due to jets arising from the non-prompt lepton background $\epsilon_{\text{fake}}$ is performed in data. A control region is defined by inverting the selection requirements on the $E_T^{\text{miss}}$, $m_W^{\text{miss}}$ and $\log_{10}(\chi^2_{\text{total}})$ obtained by the kinematic fit. The corresponding cut values are set to $E_T^{\text{miss}} < 20\text{ GeV}$, $E_T^{\text{miss}} + m_W^{\text{miss}} < 60\text{ GeV}$ and $\log_{10}(\chi^2_{\text{total}}) > 0.9$. In order to enrich this control region in non-prompt leptons, the reconstructed lepton candidate has to pass the loose requirement and fail the tight lepton requirement. This inversion of the selection requirements provides a control region completely dominated by the fake lepton contribution and in which the $b$-tagging efficiency $\epsilon_{\text{fake}}$ is determined directly (i.e. the fraction of $b$-tagged jets). In order to check the possible dependence of the efficiency of fake leptons on the definition of the control region definition, the extraction of the $\epsilon_{\text{fake}}$ is repeated in different control regions in which one of the cuts on the $E_T^{\text{miss}}$, $E_T^{\text{miss}} + m_W^{\text{miss}}$ and $\log_{10}(\chi^2_{\text{total}})$ is not inverted. Then, the largest variation is taken as uncertainty on the $\epsilon_{\text{fake}}$ extraction and propagated to the measurement of the $b$-tagging efficiency in data. The impact of this source of uncertainty on the final measurement is negligible.

6 Systematic uncertainties

The evaluation of the systematic uncertainties related to the measurements of the $b$-tagging efficiencies and the corresponding scale factors, is performed by varying the properties of the different objects in the simulation, obtaining a modified jet sample on which all the selection and measurement chain is applied.
Figure 4: $b$-tagging efficiencies for $b$-, $c$- and light-flavour jets (estimated from the simulation) and for jets coming from the non-prompt and fake lepton background (obtained from a control region in data) of the MV2c10 algorithm at a working point providing 60% efficiency. The efficiencies of the various jet flavours are shown as a function of the jet $p_T$ (a) and $|\eta|$ (b). The shaded areas represent the total statistical and systematic uncertainties.

Figure 5: Relative flavour fractions of the selected probe jet sample as a function of the jet $p_T$ (a) and $|\eta|$ (b). The shaded bands represent the total statistical and systematic uncertainties on the expected flavour fractions.
Then, all of the individual uncertainties are summed in quadrature in order to obtain the total uncertainty. During the calculations of the various systematic uncertainties corresponding to the measurement of the \( b \)-tagging scale factors \( \kappa = \varepsilon^\text{data}_{b}/\varepsilon^\text{sim.}_{b} \), the denominator is kept constant at the values obtained from the nominal jet sample predicted by the simulation. In this way, the systematic uncertainty is evaluated taking into account only the effect of the different variations on the measurement of \( b \)-tagging efficiencies in data but not on the \( b \)-tagging efficiencies in the simulation. The largest uncertainty contributions are expected to arise from systematics that lead to a significant change in the flavour composition of the selected jet sample.

**Fake lepton background estimation** Uncertainties related to the multijet background normalization procedure, described in Sec. 4.6 and extensively explained in Ref. [32], are estimated choosing a different control region for the fake efficiency extraction, a different method to extract the real efficiency for the matrix method, a variation in all MC samples subtracted from the fake control region and finally a different parametrization. All of the above variations are compared to the nominal fake lepton background and are taken into account to estimate the dependency of the measurement with respect to this procedure. This uncertainty varies from 1% to 3% as a function of the jet \( p_T \) and \(|\eta|\).

**Systematic uncertainty on the integrated luminosity** The uncertainty in the combined 2015+2016 integrated luminosity is 2.1%. It is derived, following a methodology similar to that detailed in Ref. [35], from a preliminary calibration of the luminosity scale using x-y beam-separation scans performed in August 2015 and May 2016. The uncertainty related to the integrated luminosity is smaller than 1% on the measured scale factors.

**Electron and muon** The lepton reconstruction is connected to many sources of systematic uncertainties. The trigger, the reconstruction and the identification efficiencies as well as the resolution and the scale of the momentum have been taken into account in the estimation of the systematic uncertainties [19, 22]. These systematic uncertainties are one of the smallest contribution to the total systematic uncertainty and are smaller than 0.1% for a large part of the jet \( p_T \) and \(|\eta|\) bins.

**Jet energy scale and resolution** The systematic uncertainties related to the jet energy scale measurement are taken into account and in particular the flavour-specific calorimeter response (i.e. different responses for gluon and light quark initiated jets), the modelling of \( b \)-hadron decays and the fact that a substantial fraction of heavy-flavour hadron decays contains neutrinos, which escape undetected. Additional contributions from jet-flavour composition, \( \eta \)-intercalibration, hadrons passing through the calorimeter without interacting (punch-through), single-particle response, calorimeter response to different jet flavours, and pile-up are taken into account. This systematic uncertainty is one of the most significant for the calibration and the scale factors vary from below 2% to 7% as a function of the jet \( p_T \) and \(|\eta|\). The uncertainty in the jet energy resolution is obtained with an in situ measurement of the jet response in dijet events [36]. The uncertainty in the jet energy resolution is estimated by an additional smearing of the jet energies in simulated events [37], included as one of the nuisance parameters and the only variation taken into account is the one which worsens the resolution. This systematic uncertainty has an impact on the total systematic uncertainty of about 1% as a function of the jet \( p_T \) and \(|\eta|\).
**Top quark modelling**  The following sources of systematic uncertainty related to the $t\bar{t}$ modelling are considered:

- **Hard scattering**: this uncertainty is evaluated as the difference between the Powheg+Pythia 8 and MC@NLO+Pythia 8 [38] predictions and then is symmetrized.

- **Parton showering**: the difference between Powheg+Pythia 8 and Powheg+Herwig 7 [39, 40] predictions is used to evaluate the uncertainty and then is symmetrized.

- **Initial and final state radiation**: the differences between the nominal Powheg+Pythia 8 and two different Powheg+Pythia 8 samples [41], in which the factorization and renormalization scales as well as the shower radiation and NLO radiation (hdamp = 3$m_{\text{top}}$ and 1.5$m_{\text{top}}$) have been modified, are used to estimate this uncertainty.

The total $t\bar{t}$ modelling uncertainty is evaluated by summing in quadrature the above listed contributions. This is the dominant source of systematic uncertainty and varies from 4% to 15% depending on the jet kinematics.

**7 Results**

The $b$-tagging efficiencies for the MV2c10 algorithm measured in data collected by the ATLAS detector during the 2015 and 2016 runs of the LHC by using the tag-and-probe ($t\bar{t}$ SL T&P) method, described in section 5, are presented in this section. The results are compared with the efficiencies predicted by the simulation and the corresponding data-to-simulation scale factors are extracted at different working points. These measurements have been performed for each working point as a function of the probe jet $p_T$ and the probe jet $|\eta|$. The values of $\varepsilon_b$ in the simulation, shown in the following plots, are extracted directly from the MC. A comparison between the values of $\varepsilon_b$ by applying Eq. 4 to MC and the true values derived directly from the MC simulation has been performed and the results are compatible within the statistical uncertainty.

Fig. 6 shows the $b$-tagging efficiency measurements for the 85% working point. Tables 2 and 3 list the values of the data-to-simulation scale factors with their statistical, systematic and total uncertainties for the probe jet $p_T$ and $\eta$. The results for the 77%, 70%, 60% working points are shown in Fig. 7, 8 and 9 and the corresponding scale factors are listed in Tables 4-5, 6-7 and 8-9, respectively. The measurements as a function of the probe jet $p_T$ and $\eta$ provide scale factors that are compatible with unity within their combined statistical and systematic uncertainties. The dominant contributions to the total systematic uncertainties are due to the choice of the MC generator, the choice of the fragmentation and hadronisation model, the jet energy scale and resolution, the fake lepton background normalization. Systematic uncertainties such as the jet energy resolution and scale, as well as the hadronisation model show a strong $p_T$ or $\eta$ dependence. Their contribution to the total uncertainties decreases for increasing $p_T$ values, while their contribution increases for increasing $|\eta|$ values. In contrast, the systematic uncertainties related to the choice of the $t\bar{t}$ MC generator become more relevant at higher jet $p_T$ values. The combined statistical and systematic uncertainty is below 7% in the range between 60 GeV and 300 GeV and approximately 10% for jets with a $p_T$ above 300 GeV, dominated by the statistical uncertainty. In the measurements as a function of the jet $|\eta|$, the total uncertainty range is from 7% for low $|\eta|$ values to about 13%.

The results obtained by using the $t\bar{t}$ SL T&P method are compared in Fig. 10(a)-10(c) to the scale factors provided by the dilepton $t\bar{t}$ PDF and the dilepton $t\bar{t}$ T&P calibrations [4]. In general, all three methods
are compatible within their combined statistical and systematic uncertainties. The comparison between
the values obtained by the different methods is shown for the working points of 85%, 77% and 70%,
respectively. The dominant uncertainty in the high-$p_T$ region for the dilepton $t\bar{t}$ PDF and the dilepton
$t\bar{t}$ T&P measurements is the statistical uncertainty. The tag-and-probe method on semileptonic $t\bar{t}$ events
provides a scale factor for $b$-jets with a $p_T$ between 300 GeV and 500 GeV. Thus the method described in
this note allows to extend the $p_T$ region where data-to-simulation $b$-tagging scale factors can be derived
of 200 GeV compared to the measurements performed using dilepton $t\bar{t}$ events.
7.1 Calibration for the MV2c10 85% working point

Figure 6: The $b$-tagging efficiencies for the MV2c10 algorithm at the 85% working point as a function of the transverse momentum (a) and the absolute pseudorapidity (c) of the probe jet obtained by selecting $t\bar{t}$ single lepton (SL) events. The $b$-tagging efficiencies for the predictions extracted from the simulation are shown as a red line while the efficiencies measured in data by using the tag-and-probe (T&P) method are shown as black dots. The vertical error bars represent the statistical uncertainty on the measurement. The green band indicates the total statistical and systematic uncertainties on the measurement. The corresponding data-to-simulation scale factors are presented as a function of the jet $p_T$ (b) and $|\eta|$ (d).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale factor</td>
<td>1.061</td>
<td>1.043</td>
<td>1.036</td>
<td>1.021</td>
<td>1.000</td>
<td>1.022</td>
<td>1.010</td>
</tr>
<tr>
<td>Total uncertainty [%]</td>
<td>±15.9</td>
<td>±6.4</td>
<td>±5.8</td>
<td>±6.0</td>
<td>±5.1</td>
<td>±5.1</td>
<td>±5.5</td>
</tr>
<tr>
<td>Statistical uncertainty [%]</td>
<td>±1.0</td>
<td>±0.4</td>
<td>±0.4</td>
<td>±0.4</td>
<td>±0.7</td>
<td>±1.3</td>
<td>±3.8</td>
</tr>
<tr>
<td>Systematic uncertainty [%]</td>
<td>±15.8</td>
<td>±6.4</td>
<td>±5.8</td>
<td>±6.0</td>
<td>±5.1</td>
<td>±4.9</td>
<td>±4.0</td>
</tr>
</tbody>
</table>

Table 2: Measured data-to-simulation scale factors in the various jet $p_T$ regions with statistical and systematic uncertainties (in %). The results are presented for a representative working point that corresponds to an overall $b$-tagging efficiency of 85%.
<table>
<thead>
<tr>
<th>Jet</th>
<th>0–0.50</th>
<th>0.50–1</th>
<th>1–1.50</th>
<th>1.50–2</th>
<th>2–2.50</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale factor (%)</td>
<td>1.005</td>
<td>1.017</td>
<td>1.024</td>
<td>1.048</td>
<td>1.073</td>
</tr>
<tr>
<td>Total uncertainty (%)</td>
<td>±7.9</td>
<td>±7.3</td>
<td>±7.2</td>
<td>±7.5</td>
<td>±8.8</td>
</tr>
<tr>
<td>Statistical uncertainty (%)</td>
<td>±0.4</td>
<td>±0.4</td>
<td>±0.5</td>
<td>±0.7</td>
<td>±1.0</td>
</tr>
<tr>
<td>Systematic uncertainty (%)</td>
<td>±7.9</td>
<td>±7.3</td>
<td>±7.2</td>
<td>±7.5</td>
<td>±8.8</td>
</tr>
</tbody>
</table>

Table 3: Measured data-to-simulation scale factors in the various jet $|\eta|$ regions with statistical and systematic uncertainties (in %). The results are presented for a representative working point that corresponds to an overall $b$-tagging efficiency of 85%.
7.2 Calibration for the MV2c10 77% working point

Figure 7: The $b$-tagging efficiencies for the MV2c10 algorithm at the 77% working point as a function of the transverse momentum (a) and the absolute pseudorapidity (c) of the probe jet obtained by selecting $t\bar{t}$ single lepton (SL) events. The $b$-tagging efficiencies for the predictions extracted from the simulation are shown as a red line while the efficiencies measured in data by using the tag-and-probe (T&P) method are shown as black dots. The vertical error bars represent the statistical uncertainty on the measurement. The green band indicates the total statistical and systematic uncertainties on the measurement. The corresponding data-to-simulation scale factors are presented as a function of the jet $p_T$ (b) and $|\eta|$ (d).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale factor</td>
<td>1.014</td>
<td>1.032</td>
<td>1.029</td>
<td>1.016</td>
<td>0.998</td>
<td>1.014</td>
<td>1.034</td>
</tr>
<tr>
<td>Total uncertainty [%]</td>
<td>±14.3</td>
<td>±7.4</td>
<td>±6.9</td>
<td>±6.7</td>
<td>±5.1</td>
<td>±5.5</td>
<td>±8.1</td>
</tr>
<tr>
<td>Statistical uncertainty [%]</td>
<td>±1.4</td>
<td>±0.6</td>
<td>±0.6</td>
<td>±0.6</td>
<td>±1.1</td>
<td>±2.1</td>
<td>±6.2</td>
</tr>
<tr>
<td>Systematic uncertainty [%]</td>
<td>±14.3</td>
<td>±7.4</td>
<td>±6.9</td>
<td>±6.7</td>
<td>±4.9</td>
<td>±5.1</td>
<td>±5.2</td>
</tr>
</tbody>
</table>

Table 4: Measured data-to-simulation scale factors in the various jet $p_T$ regions with statistical and systematic uncertainties (in %). The results are presented for a representative working point that corresponds to an overall $b$-tagging efficiency of 77%.
Table 5: Measured data-to-simulation scale factors in the various jet $|\eta|$ regions with statistical and systematic uncertainties (in %). The results are presented for a representative working point that corresponds to an overall $b$-tagging efficiency of 77%.
7.3 Calibration for the MV2c10 70% working point

Figure 8: The $b$-tagging efficiencies for the MV2c10 algorithm at the 70% working point as a function of the transverse momentum (a) and the absolute pseudorapidity (c) of the probe jet obtained by selecting $t\bar{t}$ single lepton (SL) events. The $b$-tagging efficiencies for the predictions extracted from the simulation are shown as a red line while the efficiencies measured in data by using the tag-and-probe (T&P) method are shown as black dots. The vertical error bars represent the statistical uncertainty on the measurement. The green band indicates the total statistical and systematic uncertainties on the measurement. The corresponding data-to-simulation scale factors are presented as a function of the jet $p_T$ (b) and $|\eta|$ (d).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale factor</td>
<td>0.998</td>
<td>1.028</td>
<td>1.027</td>
<td>1.016</td>
<td>0.993</td>
<td>1.013</td>
<td>1.021</td>
</tr>
<tr>
<td>Total uncertainty [%]</td>
<td>±15.4</td>
<td>±6.3</td>
<td>±6.4</td>
<td>±6.2</td>
<td>±5.3</td>
<td>±5.6</td>
<td>±9.1</td>
</tr>
<tr>
<td>Statistical uncertainty [%]</td>
<td>±1.0</td>
<td>±0.4</td>
<td>±0.4</td>
<td>±0.4</td>
<td>±0.8</td>
<td>±1.6</td>
<td>±4.6</td>
</tr>
<tr>
<td>Systematic uncertainty [%]</td>
<td>±15.3</td>
<td>±6.3</td>
<td>±6.4</td>
<td>±6.2</td>
<td>±5.2</td>
<td>±5.4</td>
<td>±7.9</td>
</tr>
</tbody>
</table>

Table 6: Measured data-to-simulation scale factors in the various jet $p_T$ regions with statistical and systematic uncertainties (in %). The results are presented for a representative working point that corresponds to an overall $b$-tagging efficiency of 70%.
Table 7: Measured data-to-simulation scale factors in the various jet $|\eta|$ regions with statistical and systematic uncertainties (in %). The results are presented for a representative working point that corresponds to an overall $b$-tagging efficiency of 70%.

| Jet $|\eta|$ | 0–0.50 | 0.50–1 | 1–1.50 | 1.50–2 | 2–2.50 |
|-----------|--------|--------|--------|--------|--------|
| Scale factor | 0.995 | 0.997 | 1.009 | 1.032 | 1.032 |
| Total uncertainty [%] | $\pm 8.2$ | $\pm 7.2$ | $\pm 7.1$ | $\pm 7.6$ | $\pm 9.1$ |
| Statistical uncertainty [%] | $\pm 0.4$ | $\pm 0.4$ | $\pm 0.6$ | $\pm 0.7$ | $\pm 1.2$ |
| Systematic uncertainty [%] | $\pm 8.2$ | $\pm 7.2$ | $\pm 7.1$ | $\pm 7.6$ | $\pm 9.0$ |
7.4 Calibration for the MV2c10 60% working point

Figure 9: The $b$-tagging efficiencies for the MV2c10 algorithm at the 60% working point as a function of the transverse momentum (a) and the absolute pseudorapidity (c) of the probe jet obtained by selecting $t\bar{t}$ single lepton (SL) events. The $b$-tagging efficiencies for the predictions extracted from the simulation are shown as a red line while the efficiencies measured in data by using the tag-and-probe (T&P) method are shown as black dots. The vertical error bars represent the statistical uncertainty on the measurement. The green band indicates the total statistical and systematic uncertainties on the measurement. The corresponding data-to-simulation scale factors are presented as a function of the jet $p_T$ (b) and $|\eta|$ (d).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale factor</td>
<td>0.988</td>
<td>1.022</td>
<td>1.024</td>
<td>1.014</td>
<td>0.986</td>
<td>1.011</td>
<td>1.055</td>
</tr>
<tr>
<td>Total uncertainty [%]</td>
<td>±15.3</td>
<td>±6.3</td>
<td>±6.4</td>
<td>±6.2</td>
<td>±5.2</td>
<td>±5.6</td>
<td>±8.7</td>
</tr>
<tr>
<td>Statistical uncertainty [%]</td>
<td>±1.2</td>
<td>±0.5</td>
<td>±0.5</td>
<td>±0.5</td>
<td>±0.9</td>
<td>±1.8</td>
<td>±5.8</td>
</tr>
<tr>
<td>Systematic uncertainty [%]</td>
<td>±15.2</td>
<td>±6.3</td>
<td>±6.3</td>
<td>±6.2</td>
<td>±5.1</td>
<td>±5.3</td>
<td>±6.5</td>
</tr>
</tbody>
</table>

Table 8: Measured data-to-simulation scale factors in the various jet $p_T$ regions with statistical and systematic uncertainties (in %). The results are presented for a representative working point that corresponds to an overall $b$-tagging efficiency of 60%.
Table 9: Measured data-to-simulation scale factors in the various jet $|\eta|$ regions with statistical and systematic uncertainties (in %). The results are presented for a representative working point that corresponds to an overall $b$-tagging efficiency of 60%.

| Jet $|\eta|$ | 0–0.50 | 0.50–1 | 1–1.50 | 1.50–2 | 2–2.50 |
|------------|--------|--------|--------|--------|--------|
| Scale factor | 0.989 | 0.994 | 1.003 | 1.027 | 1.017 |
| Total uncertainty [%] | ±8.3 | ±7.0 | ±7.0 | ±7.5 | ±8.9 |
| Statistical uncertainty [%] | ±0.5 | ±0.5 | ±0.6 | ±0.8 | ±1.3 |
| Systematic uncertainty [%] | ±8.2 | ±6.9 | ±7.0 | ±7.4 | ±8.8 |

Figure 10: Comparison between the data-to-simulation scale factors for the MV2c10 algorithm at the 85% (a), the 77% (b) and the 70% (c) working points obtained by using the tag-and-probe method (T&P) applied to $t\bar{t}$ single lepton (SL) candidate events and those obtained by using $t\bar{t}$ dilepton events. The results for the combinatorial likelihood (PDF) method using $t\bar{t}$ dilepton events [4] are presented as red squares and the scale factors measured with the tag-and-probe method (T&P) using $t\bar{t}$ dilepton events [4] are presented as blue triangles.
8 Conclusion

A measurement of the $b$-tagging efficiencies and the corresponding calibration scale factors of the MV2c10 algorithm has been performed on $t\bar{t}$ single lepton candidate events, requiring a final state of exactly one charged lepton and at least four jets. The $b$-jet sample is obtained by the reconstruction of the hadronically decaying top quark. The $b$-tagging efficiency measurement is performed in data collected by the ATLAS detector in $36.1 \text{ fb}^{-1}$ proton-proton collisions provided by the LHC at $13 \text{ TeV}$ during 2015 and 2016. The measurement has been done by evaluating the number of $b$-tagged jets and correcting for the expected fractions of $b$-tagged $c$- and light-flavour jets extracted from the simulation.

The MV2c10 algorithm is calibrated using this tag-and-probe method. The $b$-tagging calibration results have been presented in the form of data-to-simulation scale factors measured as a function of the $p_T$ and $\eta$. The results of the $t\bar{t}$ calibration using $\ell$+jets events are compatible within the total statistical and systematic uncertainties with the results obtained by using dilepton $t\bar{t}$ events with PDF and T&P methods. This technique allows the calibration to be extracted for jets of up to 500 GeV, beyond the reach of the dilepton $t\bar{t}$ calibrations.

References


[12] T. Sjostrand, S. Mrenna and P. Z. Skands, 

[13] P. Skands, 

[14] J. Gao et al., 


[16] ATLAS Collaboration, 

[17] S. Agostinelli et al., 

[18] ATLAS Collaboration, 

[19] ATLAS Collaboration, 

[20] ATLAS Collaboration, 

[21] ATLAS Collaboration, 

[22] ATLAS Collaboration, 

[23] M. Cacciari, G. P. Salam and G. Soyez, 

[24] ATLAS Collaboration, 

[25] ATLAS Collaboration, 

[26] ATLAS Collaboration, 

[27] G. Piacquadio and C. Weiser, 


