ATLAS NOTE

ATLAS-CONF-2012-043

March 28, 2012

Measurement of the b-tag Efficiency in a Sample of Jets Containing Muons with 5 fb^{-1} of Data from the ATLAS Detector

The ATLAS Collaboration

Abstract

The efficiency of several b-tagging algorithms has been measured with two methods using data from the ATLAS detector. The measurements are based on a sample of jets containing muons and are based on 5 fb^{-1} of data collected in 2011.

The measurements of the b-tag efficiency are provided in the form of jet-p_{T} dependent scale factors that correct the b-tagging performance in simulation to that observed in data. Good consistency is observed between the results of the two methods. The b-tag efficiency scale factors are found to be about 10% below unity for most tagging algorithms and operating points, but become consistent with unity for the loosest operating points, corresponding to the highest b-tagging efficiencies. The precision of the measurement depends on the calibration method, tagging algorithm, operating point and jet-p_{T} bin. For the highest performance MV1 tagging algorithm at an operating point corresponding to a 70% b-tag efficiency in a $t\bar{t}$ sample, the total uncertainty ranges from 5% to 19%, with the largest uncertainties corresponding to the highest p_{T} bins.

1 Introduction

The identification of jets originating from *b*-quarks is an important part of the LHC physics program. In precision measurements in the top quark sector as well as in the search for the Higgs boson and new phenomena, the suppression of background processes that contain predominantly light-flavour jets using *b*-tagging is of great use. It might also become critical to achieve an understanding of the flavour structure of any new physics (e.g. Supersymmetry) revealed at the LHC.

To use *b*-tagging in physics analyses, the efficiency ε_b with which a jet originating from a *b*-quark is tagged by a *b*-tagging algorithm needs to be measured. Other necessary pieces of information, not discussed in this note, are the probability of mistakenly tagging a jet originating from a *c*-quark or a light-flavour parton (*u*-, *d*-, *s*-quark or gluon *g*) as a *b*-jet, referred to as the *c*-tag efficiency and mistag rate respectively [1].

The *b*-tagging algorithms calibrated in this note are *SV0*, *IP3D+SV1*, *JetFitterCombNN*, *JetFitterCombNNc* and *MV1*. More details about SV0 can be found in [1] while the IP3D+SV1 and JetFitterCombNN algorithms are described in [2]. The JetFitterCombNNc algorithm is identical to JetFitterCombNN with the exception that the neural network is trained to reject *c*-jets rather than light-flavour jets. The MV1 algorithm is a neural network-based algorithm that uses the output weights of IP3D, SV1 and JetFitterCombNN as inputs.

For each *b*-tagging algorithm a set of operating points is defined, based on the inclusive *b*-tag efficiency in a simulated sample of $t\bar{t}$ events. The operating points for which calibration results are presented in this note are listed in Table 1. Figure 1 shows the expected performance of the various *b*-tagging algorithms in a simulated $t\bar{t}$ sample for jets with $p_{\mathrm{T}} > 15$ GeV and $|\eta| < 2.5$.

Tagging algorithm	ε_b(%) operating point
MV1	60, 70, 75, 85
JetFitterCombNN	57, 60, 70, 80
JetFitterCombNNc	50, 55
IP3D+SV1	60, 70, 80
SV0	50

Table 1: The tagging algorithms and operating points for which calibration results are presented in this note.

Several methods have been developed to measure the *b*-tag efficiency in data. This note describes two of them, both based on an inclusive sample of jets with muons inside. The two methods, referred to as $p_{\mathrm{T}}^{\mathrm{rel}}$ and *system8* are briefly described below. More details can be found in [1, 3].

The calibration results are presented as scale factors defined as the ratio of the *b*-tag efficiency in data to that in simulation:

$$\kappa_{\varepsilon_b}^{\mathrm{data/sim}} = \frac{\varepsilon_b^{\mathrm{data}}}{\varepsilon_b^{\mathrm{sim}}}, \tag{1}$$

where $\varepsilon_b^{\mathrm{sim}}$ is the fraction of *b*-jets which are tagged in simulated events, with the jet flavour defined by matching to generator level partons and $\varepsilon_b^{\mathrm{data}}$ is the *b*-tag efficiency measured by the $p_{\mathrm{T}}^{\mathrm{rel}}$ and system8 methods as described in Sections 3 and 4. With the methods described in this note, the *b*-tag efficiency can only be derived for semileptonic *b*-jets. The factor $\kappa_{\varepsilon_b}^{\mathrm{data/sim}}$ derived in this sample is assumed to be valid for all types of *b*-jets. Systematic uncertainties associated with this assumption are discussed in Section 5.

As the *b*-tagging performance depends strongly on the jet momentum and rapidity, the scale factors are derived in bins of jet p_{T} and jet η. The p_{T} bins used are 20 GeV $\leq p_{\mathrm{T}} <$ 30 GeV, 30 GeV $\leq p_{\mathrm{T}} <$

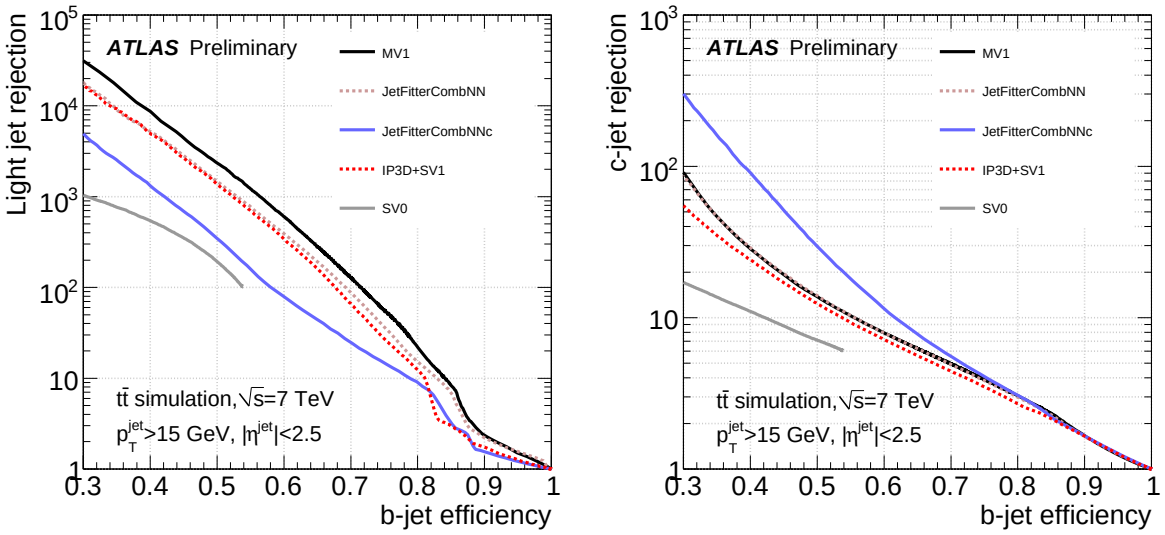

Figure 1: Light-jet rejection (left) and c-jet rejection (right) as a function of the b-tag efficiency for the b-tagging algorithms calibrated in this note, based on simulated $t\bar{t}$ events.

40 GeV, 40 GeV $\leq p_{\mathrm{T}} <$ 50 GeV, 50 GeV $\leq p_{\mathrm{T}} <$ 60 GeV, 60 GeV $\leq p_{\mathrm{T}} <$ 75 GeV, 75 GeV $\leq p_{\mathrm{T}} <$ 90 GeV, 90 GeV $\leq p_{\mathrm{T}} <$ 110 GeV, 110 GeV $\leq p_{\mathrm{T}} <$ 140 GeV and 140 GeV $\leq p_{\mathrm{T}} <$ 200 GeV, while the η bins are $0 \leq |\eta| < 0.6$, $0.6 \leq |\eta| < 1.2$, $1.2 \leq |\eta| < 1.8$ and $1.8 \leq |\eta| < 2.5$. The data-to-simulation scale factors do not show a strong dependence in either jet p_{T} or $|\eta|$, and the final results only include the subdivision in jet p_{T}.

2 Data and Simulation Samples, Object Selection

The data sample used in the analyses corresponds to approximately 5 fb^{-1} of 7 TeV proton-proton collision data collected by the ATLAS experiment during 2011. Events were collected with triggers that require a muon reconstructed from hits in the muon spectrometer that is spatially matched to a calorimeter jet. In each jet p_{T} bin of the analyses, the muon-jet trigger with the lowest jet threshold that has reached the efficiency plateau is used. In the lower p_{T} region (up to 60 GeV in the $p_{\mathrm{T}}^{\mathrm{rel}}$ analysis and up to 75 GeV in the system8 analysis) events with at least one jet with $E_{\mathrm{T}} > 10$ GeV at the last trigger level are used. Starting from 60 GeV (75 GeV) the $p_{\mathrm{T}}^{\mathrm{rel}}$ (system8) analysis uses events with at least one jet with $E_{\mathrm{T}} > 10$ GeV at the first trigger level. In the region between 110 and 200 GeV, the system8 analysis uses events with at least one jet $E_{\mathrm{T}} > 20$ or 30 GeV at the first trigger level. Each of the muon-jet triggers is collecting data at a fixed rate slightly below 1 Hz, meaning that the low jet threshold triggers are heavily prescaled.

The key objects for b-tagging are the reconstructed primary vertex, the calorimeter jets and tracks reconstructed in the inner detector. The tracks are associated with the calorimeter jets with a spatial matching in $\Delta R(\mathrm{jet}, \mathrm{track})$ [4]. The track-selection criteria depend on the b-tagging algorithm, and are detailed in [2, 5]. Jets are reconstructed from topological clusters [6] of energy in the calorimeter using the anti-k_t algorithm with a distance parameter of 0.4 [7–9]. The jet reconstruction is done at the electromagnetic scale and then a scale factor is applied in order to obtain the jet energy at the hadronic scale. The jet energy is further corrected for the energy of the muon and the average energy of the corresponding neutrino in simulated events, to arrive at the jet energy scale of an inclusive b-jet sample. The

measurement of the jet energy, the current status of the jet energy scale determination and the specific cuts used to reject jets of bad quality are described in [10]. The jets are required to have $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$. Since a well-reconstructed primary vertex is important in *b*-tagging analyses, the number of tracks associated to the primary vertex is required to be at least two. To measure the *b*-tag efficiency, the $p_{\mathrm{T}}^{\mathrm{rel}}$ and system8 methods make use of soft muons ($p_{\mathrm{T}} > 4$ GeV) associated to jets, using a spatial matching of $\Delta R(\mathrm{jet}, \mu) < 0.4$[1].

For quantities related to *b*- and *c*-jets, the analyses make use of a simulated muon-filtered QCD jet sample, referred to as the QCD μ-jet sample, where the events are required to have a muon with $p_{\mathrm{T}} > 3$ GeV at generator level. The sample is generated with PYTHIA 6 [11], utilising the ATLAS AUET2B LO** PYTHIA tune [12]. A total of 25.5 million events are simulated in four slices of $\hat{p}_{\perp}$, the momentum of the hard scatter process perpendicular to the beam line [11], starting from $\hat{p}_{\perp} = 17$ GeV. For estimates of inclusive flavour fractions, as well as quantities related to light-flavour jets, the analyses make use of an inclusive QCD jet sample for which the simulation has been carried out in six slices of $\hat{p}_{\perp}$. About 2.8 million events have been simulated per $\hat{p}_{\perp}$ slice. To simulate the detector response, the generated events are processed through a GEANT4 [13] simulation of the ATLAS detector, and then reconstructed and analysed in the same way as the data. The simulated geometry corresponds to a perfectly aligned detector and the majority of the disabled pixel modules and front-end chips seen in data are masked in the simulation. The ATLAS simulation infrastructure is described in more detail in Ref. [14].

The labeling of the flavour of a jet in simulation is done by spatially matching the jet with generator level partons [4]: if a *b*-quark is found within $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.3$ of the jet direction, the jet is labeled as a *b*-jet. If no *b*-quark is found the procedure is repeated for *c*-quarks and τ-leptons. A jet for which no such association could be made is labeled as a light-flavour jet.

3 The $p_{\mathrm{T}}^{\mathrm{rel}}$ Method

The number of *b*-jets before and after tagging can be obtained for a subset of all *b*-jets, namely those containing a reconstructed muon, using the variable $p_{\mathrm{T}}^{\mathrm{rel}}$ which is defined as the momentum of the muon transverse to the combined muon plus jet axis. Muons originating from *b*-hadron decays have a harder $p_{\mathrm{T}}^{\mathrm{rel}}$ spectrum than muons in *c*- and light-flavour jets. Templates of $p_{\mathrm{T}}^{\mathrm{rel}}$ are constructed for *b*-, *c*- and light-flavour jets separately, and these are fit to the $p_{\mathrm{T}}^{\mathrm{rel}}$ spectrum of muons in jets in data to obtain the fraction of *b*-jets before and after requiring a *b*-tag. To reduce the dependence on the modelling of $p_{\mathrm{T}}^{\mathrm{rel}}$ for muons in light-flavour jets, the heavy-flavour content in the $p_{\mathrm{T}}^{\mathrm{rel}}$ sample is increased by requiring that there is at least one jet in each event, other than those used in the $p_{\mathrm{T}}^{\mathrm{rel}}$ measurement, with a reconstructed secondary vertex with a signed decay length significance $L/\sigma(L) > 1$. The number of jets per jet p_{T} bin in the $p_{\mathrm{T}}^{\mathrm{rel}}$ sample after this additional selection is given in Table 2. This flavour-enhancement requirement is not applied in the sample used to derive the $p_{\mathrm{T}}^{\mathrm{rel}}$ template for light-flavour jets. As the templates from *c*- and light-flavour jets have a rather similar shape, the fit can only reliably separate the *b*-jets from non-*b*-jets. Therefore, the ratio of the *c*- and light-flavour fractions is constrained in the fit to the value observed in simulated events, which in the pre-tagged sample ranges from 2 at low p_{T} to 0.7 at high p_{T}. This ratio is then varied as a systematic uncertainty, as described in Section 5.

Figure 2 shows examples of template fits to the $p_{\mathrm{T}}^{\mathrm{rel}}$ distribution in data before (left) and after (right) *b*-tagging. Having obtained the flavour composition of jets containing muons from the $p_{\mathrm{T}}^{\mathrm{rel}}$ fits, the *b*-tag efficiency is defined as

$$\varepsilon_b^{\mathrm{data}} = \frac{f_b^{\mathrm{tag}} \cdot N^{\mathrm{tag}}}{f_b \cdot N} \cdot C, \tag{2}$$

[1] The impact of varying the muon association requirement has been studied, and the effect found to be negligible.

p_{T}(GeV)	pre-tagged $p_{\mathrm{T}}^{\mathrm{rel}}$ sample
20-30	44104
30-40	28988
40-50	16397
50-60	9076
60-75	36077
75-90	19257
90-110	11776
110-140	7402
140-200	3903

Table 2: The number of jets per jet p_{T} bin in data in the pre-tagged $p_{\mathrm{T}}^{\mathrm{rel}}$ sample.

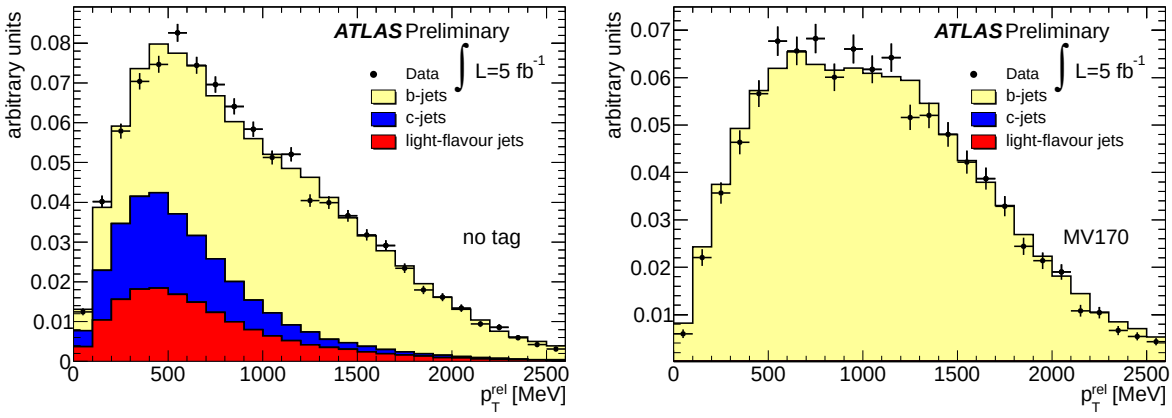

Figure 2: Examples of template fits to the $p_{\mathrm{T}}^{\mathrm{rel}}$ distribution in data before (left) and after b-tagging by applying the MV1 tagging algorithm at 70% efficiency (right) in the 40-50 GeV bin. After applying the MV1 tagging criteria the fitted b-fraction is 100%.

where f_b and f_b^{tag} are the fractions of b-jets in the pre-tagged and tagged samples of jets containing muons, and N and N^{tag} are the total number of jets in those two samples. The factor C corrects the efficiency for the biases introduced through differences between data and simulation in the modelling of the b-hadron direction and through heavy flavour contamination of the $p_{\mathrm{T}}^{\mathrm{rel}}$ template for light-flavour jets, as described below. The efficiency measured for b-jets with a semileptonically decaying b-hadron in data is compared to the efficiency for the same kind of jets in simulated events to compute the data-to-simulation scale factor defined in Eq. 1.

Both the pre-tagged and the tagged samples are fit using templates derived from all jets passing the jet selection criteria defined in Section 2, without requiring any b-tagging criteria. The $p_{\mathrm{T}}^{\mathrm{rel}}$ templates for b- and c-jets are derived from the simulated QCD μ-jet sample, using muons associated with b- and c-jets. It has been verified that the pre-tagged and tagged template shapes agree within statistical uncertainties. The template for light-quark jets is derived from muons in jets in a light-flavour dominated data sample. The sample is constructed by requiring that no jet in the event is b-tagged by the IP3D+SV1 tagging algorithm [2], using an operating point that yields a b-tag efficiency of approximately 80% in simulated $t\bar{t}$ events. This requirement rejects most events containing b-jets and yields a sample dominated by c- and light-flavour jets. The b-contamination in this sample varies between 2 and 6% depending on the

p_{T} bin. The small bias introduced in the measurement from the *b*-contamination in the light-template is corrected for in the final result.

As the $p_{\mathrm{T}}^{\mathrm{rel}}$ method is directly affected by how well the *b*-hadron direction and the calorimeter jet axis are modeled in the simulation, a difference in the jet direction resolution between data and simulation, or e.g. an improper modelling of the angle between the *b*-quark and the *b*-hadron in simulation would cause the $p_{\mathrm{T}}^{\mathrm{rel}}$ spectra in simulation and data to disagree, introducing a bias in the measurement. To study this effect, an independent jet axis was formed by vectorially adding the momenta of all tracks in the jet. The difference between this track-based and the standard calorimeter-based jet axis in the azimuth angle ϕ and the pseudorapidity η, $\Delta\phi(\mathrm{calo},\mathrm{track})$ and $\Delta\eta(\mathrm{calo},\mathrm{track})$, was derived in both data and simulation. The difference between the track-based jet axis direction and the calorimeter-based jet axis direction is observed to be a bit larger in data than in simulation, and the ϕ and η of the calorimeter-based jet axis in simulation were therefore smeared such that the $\Delta\phi(\mathrm{calo},\mathrm{track})$ and $\Delta\eta(\mathrm{calo},\mathrm{track})$ distributions agreed better with those from data. Smearing based on a Gaussian distribution with a width of 0.004 in ϕ and 0.008 in η was found to give good agreement between data and simulation in all bins of jet p_{T} [15]. The $p_{\mathrm{T}}^{\mathrm{rel}}$ templates for *b*- and *c*-jets were rederived from this smeared sample, and the $p_{\mathrm{T}}^{\mathrm{rel}}$ distribution in data was fit using these altered templates. The efficiency measured in data is corrected for half of the difference in the efficiency observed in the unsmeared and the smeared scenarios, and the full size of the correction is taken as a systematic uncertainty.

4 The System8 Method

The *system8* method uses three uncorrelated selection criteria to construct a system of eight equations based on the number of events surviving any given subset of these criteria. The system, which is fully constrained, is used to solve for eight unknowns: the efficiencies for *b* and non-*b* jets to pass each of the three selection criteria, and the number of *b* and non-*b* jets originally present in the sample. As there are not sufficient degrees of freedom to make a complete separation into (*c*, *s*, *d*, *u*, *g*) jet flavours, these are combined into one category and denoted *cl*. In simulated events, the flavour composition of the sample is relatively independent of jet p_{T} in the range studied, while the efficiency to pass each of the selection criteria has a strong p_{T} dependence. To avoid unphysical variations in the fitted flavour composition of the samples, the fraction of *b*-jets in the sample prior to applying the three selection criteria is fixed to the average of the *b*-fractions from the unconstrained fit in the p_{T} bin under study and the two neighbouring p_{T} bins. In the first and last p_{T} bins the *b*-fraction is left unchanged. The *b*-tag efficiency obtained without the constraint is taken as a systematic uncertainty in the system8 analysis, as described in Section 5.

The three selection criteria chosen are:

- The lifetime tagging criterion under study.
- A requirement that the muon $p_{\mathrm{T}}^{\mathrm{rel}}$ is greater than 700 MeV.
- A requirement that the event contains an *opposite-jet*, with $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.5$, satisfying $\pi - |\Delta\phi_{jj}| < 1$ where $\Delta\phi_{jj}$ is the difference in azimuth between the opposite-side jet and the jet under study. The opposite-jet is required to be *b*-tagged by the presence of a reconstructed secondary vertex with a decay length significance $L/\sigma(L) > 1$.

The primary motivation for these criteria is that inherently, little correlation between them is expected. In addition, the requirement of a soft muon facilitates efficient triggering and selection of calibration events. However, even if tagger correlations are small in practice, they must be accounted for. As it is impossible to isolate independent corresponding samples in data, these correlations are inferred from

p_{T}(GeV)	system8 n sample	system8 p sample
20-30	983148	13251
30-40	656723	11887
40-50	257678	5915
50-60	112386	2810
60-75	71282	1915
75-90	210529	6282
90-110	125733	3644
110-140	153312	4362
140-200	157013	4674

Table 3: The number of jets per jet p_{T} bin in data in the system8 p and n samples.

simulated samples. The resulting system of equations can be written as follows:

$$\begin{aligned}
n &= n_b + n_{cl} \\
p &= p_b + p_{cl} \\
n^{LT} &= \varepsilon_b^{LT} n_b + \varepsilon_{cl}^{LT} n_{cl} \\
p^{LT} &= \alpha_6 \varepsilon_b^{LT} p_b + \alpha_4 \varepsilon_{cl}^{LT} p_{cl} \\
n^{MT} &= \varepsilon_b^{MT} n_b + \varepsilon_{cl}^{MT} n_{cl} \\
p^{MT} &= \alpha_5 \varepsilon_b^{MT} p_b + \alpha_3 \varepsilon_{cl}^{MT} p_{cl} \\
n^{LT,MT} &= \alpha_1 \varepsilon_b^{LT} \varepsilon_b^{MT} n_b + \alpha_2 \varepsilon_{cl}^{LT} \varepsilon_{cl}^{MT} n_{cl} \\
p^{LT,MT} &= \alpha_7 \alpha_6 \alpha_5 \varepsilon_b^{LT} \varepsilon_b^{MT} p_b + \alpha_8 \alpha_4 \alpha_3 \varepsilon_{cl}^{LT} \varepsilon_{cl}^{MT} p_{cl}
\end{aligned} \tag{3}$$

In these equations, the superscripts LT and MT denote the lifetime tagging criterion and soft muon tagging criterion, respectively. The n and p numbers denote the size of the samples without (n) and with (p) the application of the opposite-jet criterion; these samples are referred to as the "n" sample and the "p" sample, respectively. The events in both samples are required to have at least one jet containing a muon, and it is to these jets that the lifetime tagging, soft muon tagging and opposite-jet criteria are applied. The number of jets per jet p_{T} bin in the p and n samples is given in Table 3. The size of the system8 samples differ from the $p_{\mathrm{T}}^{\mathrm{rel}}$ pre-tagged sample due to a slight difference in the triggers used and the $\Delta\phi_{jj}$ requirement which is not made in the $p_{\mathrm{T}}^{\mathrm{rel}}$ analysis.

Finally, the *correlation factors* $\alpha_i, i = 1,\ldots,8$, are defined as:

$$\begin{aligned}
\alpha_1 &= \varepsilon_b^{LT,MT,n} / (\varepsilon_b^{LT,n} \varepsilon_b^{MT,n}) & \alpha_2 &= \varepsilon_{cl}^{LT,MT,n} / (\varepsilon_{cl}^{LT,n} \varepsilon_{cl}^{MT,n}) \\
\alpha_5 &= \varepsilon_b^{MT,p} / \varepsilon_b^{MT,n} & \alpha_3 &= \varepsilon_{cl}^{MT,p} / \varepsilon_{cl}^{MT,n} \\
\alpha_6 &= \varepsilon_b^{LT,p} / \varepsilon_b^{LT,n} & \alpha_4 &= \varepsilon_{cl}^{LT,p} / \varepsilon_{cl}^{LT,n} \\
\alpha_7 &= \varepsilon_b^{LT,MT,p} / (\varepsilon_b^{LT,p} \varepsilon_b^{MT,p}) & \alpha_8 &= \varepsilon_{cl}^{LT,MT,p} / (\varepsilon_{cl}^{LT,p} \varepsilon_{cl}^{MT,p})
\end{aligned} \tag{4}$$

A lack of correlation between two criteria thus implies that the related correlation factors are equal to unity. The correlation factors for b- and c-jets are derived from the simulated QCD μ-jets sample, while the correlation factors associated to light-flavour jets are derived from the simulated inclusive QCD jet sample. As light-flavour jets only rarely have reconstructed muons associated with them, the statistical uncertainty on the correction factors would be unacceptably large if they were derived from muons matched to light-flavour jets in simulation. Instead, a charged particle track, fulfilling the requirements

made for the inner detector track matched to reconstructed muons, is chosen at random and treated subsequently as if it was a muon. To ensure that inner detector tracks model the kinematic properties of reconstructed muons in light-flavour jets, the tracks are weighted to account for the p_{T}- and η-dependent probability that a muon reconstructed as a track in the inner detector also gets reconstructed in the muon system, as well as the sculpting of the muon kinematics by the muon trigger turn-on curve. An additional correction factor is applied to account for the probability that an in-flight muon gets associated to the jet.

As system8 only includes correlation factors for *b* and non-*b*-jets, the *c*- and light-flavour samples have to be combined to obtain the *cl* correlation factors. The relative normalisation of the charm and light-flavour samples is inferred from the simulated inclusive QCD sample leading to a charm-to-light ratio in the *n*- and *p*-samples which ranges from 0.6 to 1.5 depending on sample and jet p_{T} bin. The variation of the charm fraction in the combined sample is treated as a systematic uncertainty, as discussed in Section 5. The actual values of the correlation factors depend on the tagging algorithm, operating point and jet p_{T} bin. For the MV1 tagging algorithm at 70% efficiency, 95% of the correlation factors are within 15% from unity and no correlation factor is differing from unity by more than 30%.

5 Systematic Uncertainties

The systematic uncertainties affecting the $p_{\mathrm{T}}^{\mathrm{rel}}$ and system8 methods are common to a large extent. One important class of common systematic uncertainties are those addressing how well the simulation models heavy flavour production, decays and fragmentation. Other common systematic uncertainties are those arising from the imperfect knowledge of the jet energy scale and resolution as well as the modeling of the additional pileup interactions. The systematic uncertainty coming from limited simulation statistics is almost fully uncorrelated between the two analyses despite the fact that they are using the same simulated samples. A systematic uncertainty which applies only to the $p_{\mathrm{T}}^{\mathrm{rel}}$ analysis is the heavy-flavour contamination in the $p_{\mathrm{T}}^{\mathrm{rel}}$ light-flavour data control sample while a systematic uncertainty which only applies to the system8 analysis arises from varying the muon $p_{\mathrm{T}}^{\mathrm{rel}}$ cut which is used as the soft muon tagging criteria.

The systematic uncertainties on the data-to-simulation scale factor $\kappa_{\varepsilon_b}^{\mathrm{data/sim}}$ of the MV1 tagging algorithm at 70% efficiency is shown in Table 4 and 5 for the $p_{\mathrm{T}}^{\mathrm{rel}}$ and system8 methods respectively. The systematic uncertainties have a positive (negative) sign if the difference between the shift in the scale factor when applying a positive and a negative variation of the underlying parameter is positive (negative), i.e. if $\kappa_{\varepsilon_b}^{\mathrm{data/sim}}(\mathrm{up}) - \kappa_{\varepsilon_b}^{\mathrm{data/sim}}(\mathrm{down}) > 0$ (< 0). The estimates of the systematic uncertainties, especially in the system8 analysis, suffer from the limited simulation statistics which leads to unphysical bin-to-bin variations in some cases. However, when the two methods are combined these irregularities get smoothed out.

Simulation Statistics

The limited simulation statistics results in statistical fluctuations on the $p_{\mathrm{T}}^{\mathrm{rel}}$ templates in the case of the $p_{\mathrm{T}}^{\mathrm{rel}}$ analysis and in statistical uncertainties on the system8 correlation factors in the system8 analysis.

The effect from limited template statistics in the $p_{\mathrm{T}}^{\mathrm{rel}}$ analysis is assessed through pseudo-experiments as described in Ref. [1]. In the system8 analysis, the limited statistics available for the samples used to estimate the correlation factors is accounted for using an extra contribution to the fit χ^2, as described in Ref. [3]. To estimate the contribution to the uncertainty from this limited statistics, the uncertainty for the fit without this addition is subtracted quadratically from the uncertainty for the fit including it.

In addition, the limited simulation statistics propagate to an uncertainty on the denominator in the scale factor expression, denoted *simulation tagging efficiency* in Tables 4 and 5.

	Jet p_{T} [GeV]								
Source	20 to 30	30 to 40	40 to 50	50 to 60	60 to 75	75 to 90	90 to 110	110 to 140	140 to 200
simulation statistics	2.2	2.0	1.4	1.8	2.7	3.4	6.0	6.0	4.7
simulation tagging efficiency	1.0	1.0	0.4	0.6	1.6	0.9	0.6	0.5	0.1
modelling of b-production	0.5	0.2	< 0.1	0.4	0.1	0.4	0.7	0.8	1.0
modelling of c-production	0.1	0.3	0.1	0.2	0.2	0.4	0.2	0.6	2.5
b-hadron direction modelling	0.1	0.2	0.2	0.4	0.4	1.1	0.7	2.5	1.4
b-fragmentation fraction	0.1	0.5	0.1	0.1	< 0.1	0.1	0.4	0.6	0.7
b-fragmentation function	< 0.1	0.1	< 0.1	< 0.1	< 0.1	0.1	0.3	0.8	0.9
b-decay branching fractions	< 0.1	0.1	0.1	0.1	0.1	0.1	< 0.1	0.1	0.3
b-decay p^* spectrum	-0.6	-0.7	-0.8	-0.9	-1.0	-0.7	-0.6	-0.4	-1.0
charm-light ratio	3.0	2.4	2.1	2.0	1.2	0.4	0.9	1.5	16.0
muon p_{T} spectrum	0.1	< 0.1	0.1	< 0.1	< 0.1	< 0.1	0.2	< 0.1	1.4
fake muons in b-jets	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	0.1	-0.1	-0.2
$p_{\mathrm{T}}^{\mathrm{rel}}$ light template contamination	-0.1	-0.2	-0.4	-0.4	-0.4	-0.8	-0.7	-0.7	-1.0
jet energy resolution	-0.1	0.5	< 0.1	0.2	< 0.1	0.5	-0.8	1.7	0.4
jet energy scale	< 0.1	0.2	0.4	0.3	0.4	0.5	< 0.1	0.5	0.4
semileptonic correction	-0.5	-0.1	< 0.1	-0.2	-0.2	-0.1	0.1	< 0.1	-0.3
jet vertex fraction	2.9	0.6	0.3	0.4	0.3	0.3	0.4	0.1	0.2
pileup μ reweighting	0.1	< 0.1	< 0.1	0.1	< 0.1	0.3	0.1	0.5	1.9
scale factor for inclusive b-jets	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
total systematic	6.4	5.3	4.9	5.1	5.4	5.7	7.5	8.1	17.7
statistical	1.8	1.9	2.5	3.7	2.1	3.5	5.2	7.4	14.8

Table 4: Relative statistical and systematic uncertainties, in %, on the data-to-simulation scale factor $\kappa_{\varepsilon_b}^{\mathrm{data/sim}}$ from the $p_{\mathrm{T}}^{\mathrm{rel}}$ method for the MV1 tagging algorithm at 70% efficiency.

	Jet p_{T} [GeV]								
	20 to	30 to	40 to	50 to	60 to	75 to	90 to	110 to	140 to
Source	30	40	50	60	75	90	110	140	200
simulation statistics	5.1	2.8	2.4	5.5	2.2	4.7	5.4	26.5	18.8
simulation tagging efficiency	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	0.1	0.1	0.2	0.3
modelling of b-production	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	-0.3	< 0.1	-0.8	-3.0
modelling of c-production	< 0.1	-0.1	-0.1	-0.1	-0.1	-0.3	-0.4	-0.3	0.4
b-hadron direction modelling	-0.6	< 0.1	-0.3	0.9	0.1	0.4	-1.6	-9.1	1.6
b-fragmentation fraction	-0.3	1.0	2.1	3.9	2.3	3.4	2.4	-1.2	-8.0
b-fragmentation function	0.3	-0.1	-0.2	0.6	< 0.1	-0.2	-0.1	-1.3	-5.7
b-decay branching fractions	-0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	0.1	< 0.1	-7.8
b-decay p^* spectrum	-0.3	0.3	0.1	0.1	0.3	0.1	-0.8	0.7	-0.8
charm-light ratio	-0.2	-0.1	< 0.1	-0.7	< 0.1	-0.6	< 0.1	-4.0	-0.7
$p_{\mathrm{T}}^{\mathrm{rel}}$ cut variation	0.4	0.4	0.4	0.4	0.4	-0.2	-0.2	-0.2	-0.2
b-fraction constraint	< 0.1	0.1	6.2	11.0	14.3	1.2	2.5	24.4	< 0.1
jet energy resolution	2.1	-1.7	-0.5	-4.0	-1.4	2.5	0.1	-8.5	-5.7
jet energy scale	-1.7	-1.1	1.7	1.5	0.5	-1.2	< 0.1	0.5	1.4
jet vertex fraction	-2.2	0.9	-0.2	0.1	-0.2	-0.6	0.3	-3.0	-0.5
pileup μ reweighting	-0.4	0.3	0.2	-0.1	0.1	0.5	0.3	9.1	5.2
scale factor for inclusive b-jets	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
total systematic	7.4	5.5	8.3	14.2	15.3	7.7	7.8	39.8	24.6
statistical	2.6	1.8	2.8	5.7	5.2	5.4	6.5	16.3	13.2

Table 5: Relative statistical and systematic uncertainties, in %, on the data-to-simulation scale factor $\kappa_{\varepsilon_b}^{\mathrm{data/sim}}$ from the system8 method for the MV1 tagging algorithm at 70% efficiency.

The systematic uncertainty arising from limited simulation statistics is treated as fully uncorrelated between p_{T} bins but fully correlated between the $p_{\mathrm{T}}^{\mathrm{rel}}$ and system8 analyses for a given p_{T} bin.

Modelling of *b*- and *c*-Production

As the properties of jets with two *b*- or *c*-quarks inside (originating e.g. from gluon splitting) are different from those containing only a single *b*- or *c*-quark, a possible mismodelling of the fraction of double-*b* or double-*c* jets in simulation has to be taken into account. Jets which have two associated *b*-quarks or *c*-quarks are either given a weight of zero or a weight of two (effectively removing or doubling the double-*b* or double-*c* contribution) when constructing the $p_{\mathrm{T}}^{\mathrm{rel}}$ templates and system8 correlation factors.

b-Hadron Direction Modelling

Both the $p_{\mathrm{T}}^{\mathrm{rel}}$ and the system8 analyses make use of the momentum of the associated muon transverse to the combined muon plus jet axis, where the muon plus jet axis is a measure of the *b*-hadron direction. A different jet direction resolution in data and simulation would therefore affect both analyses. This is accounted for by smearing the calorimeter jet direction by 0.004 in ϕ and 0.008 in η [15], as discussed in Section 3. The $p_{\mathrm{T}}^{\mathrm{rel}}$ analysis corrects the measured efficiency by half of the difference between the efficiency obtained with the smeared and unsmeared samples, and assigns the full size of the correction as a systematic uncertainty while the system8 analysis uses the result in the unsmeared sample as the central value and treats the full difference to the smeared sample as a systematic uncertainty.

b-Quark Fragmentation

An incorrect modelling of the *b*-quark fragmentation in simulation can affect the momentum spectrum of the muons from *b*-decays and thus alter which muons pass the selection criteria. To investigate the impact of fragmentation on the data-to-simulation scale factor, the $p_{\mathrm{T}}^{\mathrm{rel}}$ templates and system8 correlation factors have been rederived on a simulated sample where the *b* fragmentation function was reweighed so that the average fraction of the *b*-quark energy given to the *b*-hadron was changed by 5%.

The production fractions of the various *b*-flavoured hadrons have been measured both at LEP and the Tevatron [16, 17], and the results for *b*-baryon production are only compatible at the 2 σ level. The production fractions in the simulated samples used in this note are in reasonable agreement with the fractions as measured by LEP. A systematic uncertainty is evaluated by considering the difference in the result obtained by reweighting all of the events so that the distribution of hadron species matches the measured Tevatron numbers.

b-Decay

The muon momentum spectrum in the *b*-hadron restframe, denoted as p^*, directly affects the shape of the $p_{\mathrm{T}}^{\mathrm{rel}}$ distribution for *b*-jets. Uncertainties in the modelling of the p^* spectrum have to be taken into account and propagated through the analysis. The p^* spectrum has two components, direct $b \to \mu + X$ decays and cascade $b \to c/\overline{c} \to \mu + X$ decays. Their branching fractions are $BF(b \to \ell X) = (10.69 \pm 0.22)\%$ and $BF(b \to c/\overline{c} \to \ell X) = (9.62 \pm 0.53)\%$, respectively [16], giving the ratio $BF(b \to \ell X)/BF(b \to c/\overline{c} \to \ell X) = 1.11 \pm 0.07$, where ℓ denotes either a muon or an electron. This ratio of branching fractions has been varied within the quoted uncertainty. To investigate the effect of variations of the p^* spectra, a weighting function has been applied to the p^* spectrum of muons from the direct $b \to \mu + X$ decay. This weighting function has been derived by comparing the direct p^* spectrum of $b \to \ell + X$ decays in PYTHIA as used in the analysis with the corresponding spectrum measured in [18].

Charm-to-Light Ratio

Both the $p_{\mathrm{T}}^{\mathrm{rel}}$ and system8 methods are sensitive to the relative fractions of *c*- and light-flavour jets in the simulation. As the $p_{\mathrm{T}}^{\mathrm{rel}}$ templates for *c*- and light-flavour jets have a very similar shape, the $p_{\mathrm{T}}^{\mathrm{rel}}$ fits can become unstable if both components are allowed to vary freely in the fit. Therefore the fits to the $p_{\mathrm{T}}^{\mathrm{rel}}$ templates are performed with the ratio of the charm and light fractions fixed to the simulated value. In the system8 method, the ratio of *c*- and light-flavour fractions in the *n* and *p* samples are also fixed to their simulated values. The relative fractions of *c*- and light-flavour jets in these samples will affect the correction factors related to non-*b*-jets, namely α_2, α_3, α_4 and α_8. In both analyses the impact of the constrained charm-to-light ratio has been addressed by varying the ratio up and down by a factor of two. The charm-to-light ratio variation is one of the dominant systematic uncertainties in the $p_{\mathrm{T}}^{\mathrm{rel}}$ analysis, especially in the highest p_{T} bin where the *b*- and *c*-templates start to look very similar. The system8 analysis, which is only relying on the $p_{\mathrm{T}}^{\mathrm{rel}}$ cut to increase the fraction of *b*-jets in the sample, is less affected.

Muon p_{T} Spectrum

The muon p_{T} spectrum is softer in data than in simulation. To estimate the effect of this mismodelling on the $p_{\mathrm{T}}^{\mathrm{rel}}$ distribution, the $p_{\mathrm{T}}^{\mathrm{rel}}$ analysis is repeated after reweighting the muon p_{T} spectrum in simulation to agree with that in data. The difference between the two measurements is assigned as a systematic uncertainty.

Fake Muons in *b*-Jets

The $p_{\mathrm{T}}^{\mathrm{rel}}$ templates are obtained from the simulated QCD μ-jet sample where a muon with $p_{\mathrm{T}} > 3$ GeV is required at generator level. This filter suppresses *b*-jets containing a fake muon over those where the muon originated from a *b*-decay. The fraction of fake muons in this sample is therefore likely to be lower than in data which could potentially impact the $p_{\mathrm{T}}^{\mathrm{rel}}$ *b*-template shapes. To address this, the $p_{\mathrm{T}}^{\mathrm{rel}}$ measurement has been repeated with the fake muon fraction in the *b*-template increased by a factor of three, which was found to have a negligible impact on the final result.

$p_{\mathrm{T}}^{\mathrm{rel}}$ Light-Template Contamination

In the $p_{\mathrm{T}}^{\mathrm{rel}}$ method, the templates for light-flavour jets are obtained from a light-flavour enriched data sample. A measurement bias can arise from *b*-jet contamination in the light-flavour template. This *b*-jet contamination in the light-flavour template is estimated from simulation to be between 4% and 6% depending on jet p_{T} bin. The bias introduced by this is corrected for in the final result, and the result of a variation by 25% of the *b*-jet contamination is taken as systematic uncertainty.

$p_{\mathrm{T}}^{\mathrm{rel}}$ Cut Variation

The system8 analysis uses a cut on the $p_{\mathrm{T}}^{\mathrm{rel}}$ of the muon associated to the jet to arrive at a sample with enhanced heavy flavour content. The $p_{\mathrm{T}}^{\mathrm{rel}}$ cut, which is nominally placed at 700 MeV, was varied between 600 and 800 MeV, and the difference to the nominal result was taken as a systematic uncertainty.

b-Fraction Constraint

The fraction of *b*-jets in the *n* sample is one of the unknowns in system8, as defined in Eq. 3. To avoid unphysical variations in the fitted flavour composition of the samples, the fraction of *b*-jets in the sample prior to applying the three selection criteria is fixed to the average of the *b*-fractions from the

unconstrained fit in the p_{T} bin under study and the two neighbouring p_{T} bins. In the first and last p_{T} bins the *b*-fraction is left unchanged. The *b*-tag efficiency obtained from the fit without the *b*-fraction constraint is taken as a systematic uncertainty in the system8 analysis. As the difference between the *b*-fraction obtained in the floating fit and the average used in the constrained fit is large for some bins while smaller for others, this systematic uncertainty varies significantly between p_{T} bins.

Jet Energy Resolution

The jet energy resolution in simulation is smeared to agree with that in data [19]. As a systematic uncertainty an additional smearing, accounting for the uncertainties in the nominal smearing parameters, has been applied.

Jet Energy Scale

A jet energy scale in simulation that is different from that in data would bias the p_{T} spectrum of the simulated events used to extract the $p_{\mathrm{T}}^{\mathrm{rel}}$ templates and the system8 correlation factors. The systematic uncertainty originating from the jet energy scale is obtained by scaling the p_{T} of each jet in the simulation up and down by one standard deviation, according to the uncertainty of the jet energy scale [10].

Semileptonic Correction

Both the $p_{\mathrm{T}}^{\mathrm{rel}}$ and system8 analyses measure the *b*-tag efficiency of jets containing semileptonically decaying *b*-hadrons. As part of the jet energy is taken by the muon and the neutrino, the jet energy scale for such jets are different from a sample of inclusive *b*-jets. Both analyses are therefore making an extra jet energy scale correction, to correct the jet energy to the inclusive scale. The uncertainty on this correction, which amounts to about 2%, reflects how sensitive the correction is to the modelling of *b*-jets in the simulation, the correlation between the correction and the *b*-tag output weights and how well the correction agrees in data and simulation. The uncertainty on the semileptonic correction is propagated through the $p_{\mathrm{T}}^{\mathrm{rel}}$ analysis as a systematic uncertainty. Systematic sources which affect both the semileptonic correction and the $p_{\mathrm{T}}^{\mathrm{rel}}$ templates are varied in a correlated manner. The system8 analysis does not account for this uncertainty, but given its small size this will have a negligible impact on the final result.

Jet Vertex Fraction

Jets originating from pileup interactions can be rejected by requiring that the tracks associated to the jet are compatible with originating from the selected primary vertex. The fraction of compatible tracks is referred to as the Jet Vertex Fraction or JVF. Many physics analyses in ATLAS consider only jets with a large JVF (typically above 0.75). The data-to-simulation scale factors derived only from jets with a JVF above 0.75 is compared to those obtained without a cut on the JVF, and the difference is treated as a systematic uncertainty.

Pileup μ Reweighting

During 2011, the maximum average number of interactions per bunch crossing, referred to as μ, increased from 4 to 17. Simulation studies show that the impact on the *b*-tagging performance from the change in pileup conditions during 2011 is relatively small compared to the precision of the $p_{\mathrm{T}}^{\mathrm{rel}}$ and system8 analyses. The change in light-jet rejection at fixed *b*-tag efficiency exceeds 5% only for the tightest operating points. However, to ensure that a pileup dependence does not bias the scale-factor

measurement, the simulation has been reweighted to reproduce the μ-distribution in data. As a systematic uncertainty, the μ values in simulation have been scaled up and down by 9% prior to reweighting, based on the level of agreement in other variables sensitive to the amount of pileup.

Scale Factor for Inclusive *b*-Jets

The $p_{\mathrm{T}}^{\mathrm{rel}}$ and system8 methods can only measure the *b*-tag efficiency in data for *b*-jets with a semileptonic *b*-hadron decay. As these jets always contain a high-momentum and typically well-measured muon track, whereas the hadronic *b*-jets do not, the *b*-tag efficiency is different for these two types of *b*-jets. However, the calibration results in this note are to first order insensitive to this effect as they are given in the form of data-to-simulation scale factors. As long as the simulation adequately models the relative differences in *b*-tag efficiencies between semileptonic and hadronic *b*-jets the same data-to-simulation scale factor is valid for both types of jets. Therefore, the data-to-simulation scale factor derived as the ratio between the semileptonic *b*-tag efficiency in data and simulation, is assumed to be identical for hadronic *b*-jets.

To investigate the validity of this assumption, the data-to-simulation scale factor was measured separately for jets with and without muons using a high purity sample of *b*-jets in $t\bar{t}$ dilepton events. The ratio of the data-to-simulation scale factors in jets with and without muons was found to be consistent with unity for all tagging algorithms and operating points. The uncertainty of the measurement, which is approximately 4%, is assigned as a systematic uncertainty on the data-to-simulation scale factors obtained with the muon-based methods.

6 Results

The *b*-tag efficiencies measured in data using the $p_{\mathrm{T}}^{\mathrm{rel}}$ and system8 methods, the corresponding values from simulation and the resulting data-to-simulation scale factors for the MV1 tagging algorithm at 70% efficiency are shown in Fig. 3 for the $p_{\mathrm{T}}^{\mathrm{rel}}$ and system8 methods. Results for other tagging algorithms and operating points can be found in Appendix A. The *b*-tag efficiencies from simulation are determined by directly applying the *b*-tagging algorithms to the sample of jets labelled as *b*-jets (defined in Section 2). As the sample of jets selected for the $p_{\mathrm{T}}^{\mathrm{rel}}$ and system8 measurements are different, the fraction of *b*-tagged jets are not necessarily equal.

The agreement in the data-to-simulation scale factors between the two methods is very good. For this particular tagging algorithm and operating point the efficiency in data tends to be about 5 to 10% lower than that in simulation, leading to a data-to-simulation scale factors ranging between 0.9 and 0.95. In general, the data-to-simulation scale factors can be as low as 0.85 for the tightest operating points while they are consistent with unity for the loosest operating points.

7 Combination

To improve the precision of the *b*-tag efficiency measurement, a combination of the $p_{\mathrm{T}}^{\mathrm{rel}}$ and system8 results is performed. In each jet p_{T} bin, the best estimate of the true data-to-simulation scale factor $\hat{\kappa}$ is extracted by maximising the likelihood that each measurement κ_i, associated with a statistical uncertainty $\delta\kappa_i^{\mathrm{stat}}$ and a set of systematic uncertainties $\delta\kappa_i^{\mathrm{syst}_j}$, originates from a Gaussian probability density function $\mathscr{P}_i$. As an example, a combination of two measurements κ_1 and κ_2, with statistical uncertainties $\delta\kappa_1^{\mathrm{stat}}$ and $\delta\kappa_2^{\mathrm{stat}}$ and two sources of systematic uncertainty syst_1 and syst_2 affecting the two measurements by the amounts $\delta\kappa_1^{\mathrm{syst}_1}$, $\delta\kappa_2^{\mathrm{syst}_1}$, $\delta\kappa_1^{\mathrm{syst}_2}$ and $\delta\kappa_2^{\mathrm{syst}_2}$, is performed by maximising the following

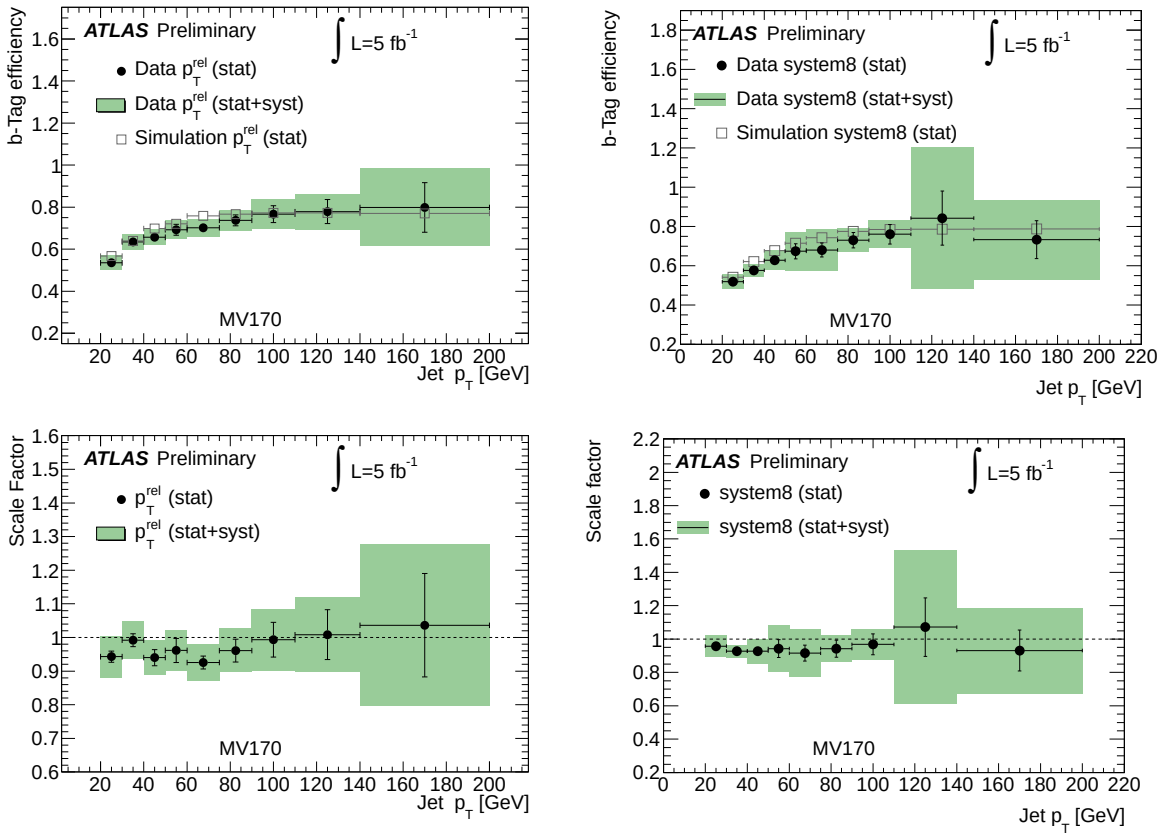

Figure 3: The *b*-tag efficiency in data and simulation (top) and the data-to-simulation scale factor (bottom) for the MV1 tagging algorithm at 70% efficiency obtained with the $p_{\mathrm{T}}^{\mathrm{rel}}$ method (left) and the system8 method (right).

likelihood expression:

$$\begin{aligned}\mathscr{L} = \prod_{i=1}^{2} \mathscr{P}_i \quad = \quad & G(\kappa_1|\hat{\kappa}(1+\delta\kappa_1^{\text{syst}_1}\lambda^{\text{syst}_1}+\delta\kappa_1^{\text{syst}_2}\lambda^{\text{syst}_2}),\delta\kappa_1^{\text{stat}})\times \\ & G(\kappa_2|\hat{\kappa}(1+\delta\kappa_2^{\text{syst}_1}\lambda^{\text{syst}_1}+\delta\kappa_2^{\text{syst}_2}\lambda^{\text{syst}_2}),\delta\kappa_2^{\text{stat}})\times \\ & G(\lambda^{\text{syst}_1}|0,1)\times G(\lambda^{\text{syst}_2}|0,1)\end{aligned} \tag{5}$$

Here, $G(\kappa_i|\hat{\kappa}(1+\delta\kappa_i^{\text{syst}_1}\lambda^{\text{syst}_1}+\delta\kappa_i^{\text{syst}_2}\lambda^{\text{syst}_2}),\delta\kappa_i^{\text{stat}})$ with $i=1,2$, are Gaussian distributions centered at $1+\delta\kappa_i^{\text{syst}_1}\lambda^{\text{syst}_1}+\delta\kappa_i^{\text{syst}_2}\lambda^{\text{syst}_2}$ with a width of $\delta\kappa_i^{\text{stat}}$, evaluated at point κ_i. The parameters λ^{syst_j} with $j=1,2$, which control by how many standard deviations the systematic uncertainty syst_j is shifting the central value of the Gaussian distribution, are controlled by Gaussian probability terms $G(\lambda^{\text{syst}_j}|0,1)$ constraining the parameters λ^{syst_j} to originate from Gaussian distributions which are centered at zero and have unit width.

In the combination of the $p_{\text{T}}^{\text{rel}}$ and system8 results, κ_1 and κ_2 correspond to the scale factors measured by the two analyses. The number of systematic uncertainties amounts to 21, where all but seven are in common between the two analyses. All systematic uncertainties are treated as fully correlated across p_{T} bins. The only exceptions are the systematic uncertainties arising from limited simulation statistics and the systematic uncertainty from the b-fraction constraint in the system8 analysis which are treated as fully uncorrelated across p_{T} bins. The common systematic uncertainties are also treated as fully correlated between the $p_{\text{T}}^{\text{rel}}$ and system8 measurements.

As the $p_{\text{T}}^{\text{rel}}$ and system8 analyses are using partly overlapping samples, the statistical uncertainty is actually partially but not fully correlated. The correlation coefficients have been derived using toy experiments in which somewhat simplified versions of the $p_{\text{T}}^{\text{rel}}$ and system8 analyses were performed. The statistical correlation of the two analyses was found to be below 66% for all tagging algorithms, operating points and p_{T} bins. The smallest correlations are observed in the p_{T} bins that suffer from large statistical uncertainties, while the largest correlations are found for bins in the lower p_{T} range where the statistical uncertainties are smaller. There are also p_{T} bins in which the two analyses use different, but highly prescaled, triggers, leading to a negligible correlation. The correlation of the statistical uncertainty is accounted for in the combination by dividing it into two components, one which is treated as fully correlated and the one which is treated as fully uncorrelated. The correlation of the systematic uncertainty arising from limited simulation statistics was found to be negligible.

The result of the combination for the MV1 tagging algorithm at the 70% operating point is shown in Fig. 4. Combined results for other tagging algorithms and operating points can be found in Appendix A. Note that the combined value of the scale factor in the first and last jet p_{T} bins is slightly below both input values. This is a consequence of the fact that the likelihood function assumes that the systematic uncertainties of a given origin are correlated in all p_{T} bins. A systematic uncertainty which is constrained in the low jet p_{T} bins can thus cause the combined scale factor in the higher p_{T} bins to move below or above both inputs values and vice versa. For the tagging algorithms and operating points where this happens, the difference between the combined scale factor and the input values is always small with respect to the uncertainty of the measurement.

8 Conclusions

Two methods, based on a sample of jets containing muons, have been used to measure the b-tag efficiency of several algorithms with 5 fb^{-1} of data from the ATLAS detector. The results are expressed in terms of scale factors, correcting the efficiencies in simulated events to those measured in data. The scale factors measured with the two methods are consistent with each other within uncertainties. For operating

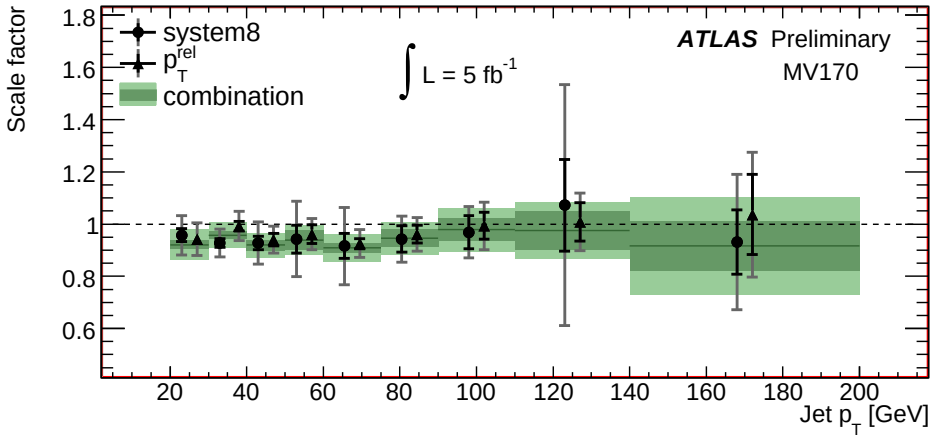

Figure 4: The data-to-simulation scale factor for the MV1 tagging algorithm at 70% efficiency, obtained by combining the p_T^{rel} and system8 results. The dark green band represents the statistical uncertainty of the combined scale factor while the light green band shows the total uncertainty. The data points showing the p_T^{rel} and system8 measurements have been separated a little along the x-axis to make the plot more readable.

points corresponding to a high b-tag efficiency, the scale factors are consistent with unity, while the b-tag efficiency in data is measured to be up to 15% lower than in simulated events for operating points corresponding to a b-tag efficiency of 50 or 60%. The uncertainties on the scale factors depend on tagging algorithm and jet p_T. For the MV1 tagging algorithm at 70% efficiency they range from 5% in the intermediate p_T range to as much as 19% in the high p_T region.

References

[1] ATLAS Collaboration, *Calibrating the b-Tag Efficiency and Mistag Rate in 35 pb^{-1} of Data with the ATLAS Detector*, ATLAS-CONF-2011-089, June, 2011.

[2] ATLAS Collaboration, *Commissioning of high performance b-tagging algorithms with the ATLAS detector*, ATLAS-CONF-2011-102, July, 2011.

[3] ATLAS Collaboration, *b-tagging Efficiency Calibration in 35 pb^{-1} of Data with the ATLAS Detector using the System8 Method*, ATLAS-CONF-2011-143, October, 2011.

[4] ATLAS Collaboration, *Expected performance of the ATLAS experiment: Detector, Trigger and Physics, Volume 1*, CERN-OPEN-2008-020, December, 2008.

[5] ATLAS Collaboration, *Performance of the ATLAS Secondary Vertex b-tagging Algorithm in 7 TeV Collision Data*, ATLAS-CONF-2010-042, May, 2010.

[6] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[7] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-k(t) jet clustering algorithm*, JHEP **04** (2008) 063.

[8] M. Cacciari and G. P. Salam, *Dispelling the N**3 myth for the k(t) jet-finder*, Phys. Lett. **B641** (2006) 057.

[9] M. Cacciari, G. P. Salam, and G. Soyez. http://fastjet.fr/.

[10] ATLAS Collaboration, *Jet energy measurement with the ATLAS detector in proton- proton collisions at sqrt(s) = 7 TeV*, `arXiv:1112.6426 [hep-ex]`.

[11] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **05** (2006) 026.

[12] ATLAS Collaboration, *Further tunes of PYTHIA6 and Pythia 8*, ATL-PHYS-PUB-2011-014, November, 2011.

[13] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250–303.

[14] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. **C70** (2010) 823–874, `arXiv:1005.4568 [physics.ins-det]`.

[15] ATLAS Collaboration, *Measurement of the b-jet production cross section using muons in jets with ATLAS in pp Collisions at* $\sqrt{s} = 7$ *TeV*, ATLAS-CONF-2011-057, April, 2011.

[16] Particle Data Group Collaboration, K. Nakamura et al., *Review of particle physics*, J. Phys. **G37** (2010) 075021.

[17] CDF Collaboration, T. Aaltonen et al., *Measurement of Ratios of Fragmentation Fractions for Bottom Hadrons in p anti-p Collisions at* $\sqrt{s} = 1.96$ *TeV*, Phys. Rev. **D77** (2008) 072003, `arXiv:0801.4375 [hep-ex]`.

[18] DELPHI Collaboration, J. Abdallah et al., *Determination of heavy quark non-perturbative parameters from spectral moments in semileptonic B decays*, Eur. Phys. J. **C45** (2006) 35–59.

[19] ATLAS Collaboration, *Jet energy resolution and reconstruction efficiencies from in-situ techniques with the ATLAS Detector Using Proton-Proton Collisions at a Center of Mass Energy* $\sqrt{s} = 7$ *TeV*, ATLAS-CONF-2010-054, July, 2010.

A Results for Additional Tagging Algorithms and Operating Points

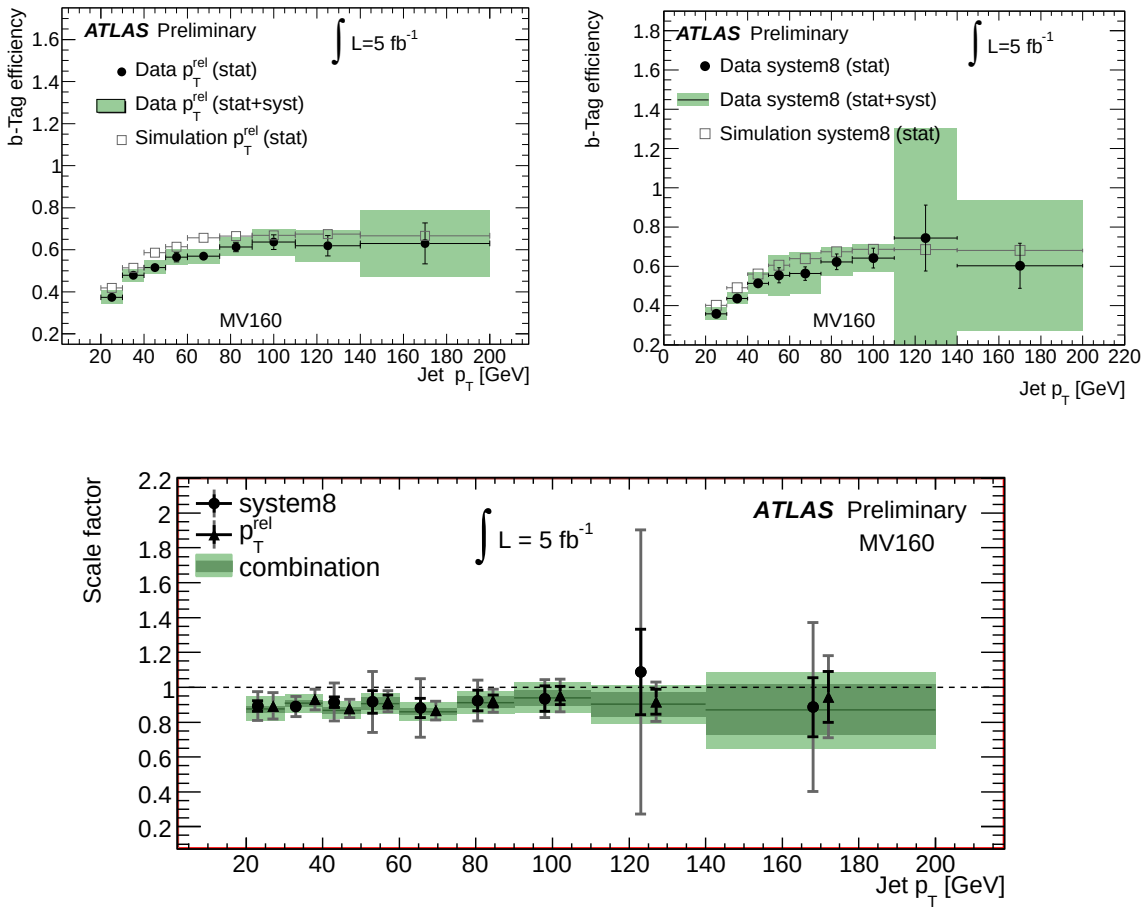

Figure 5: The *b*-tag efficiency in data and simulation for the $p_{\mathrm{T}}^{\mathrm{rel}}$ method (top left) and the system8 method (top right) as well as the individual and combined data-to-simulation scale factors (bottom) for the MV1 tagging algorithm at 60% efficiency. The dark green band represents the statistical uncertainty of the combined scale factor while the light green band shows the total uncertainty. The data points showing the $p_{\mathrm{T}}^{\mathrm{rel}}$ and system8 measurements have been separated a little along the *x*-axis to make the plot more readable.

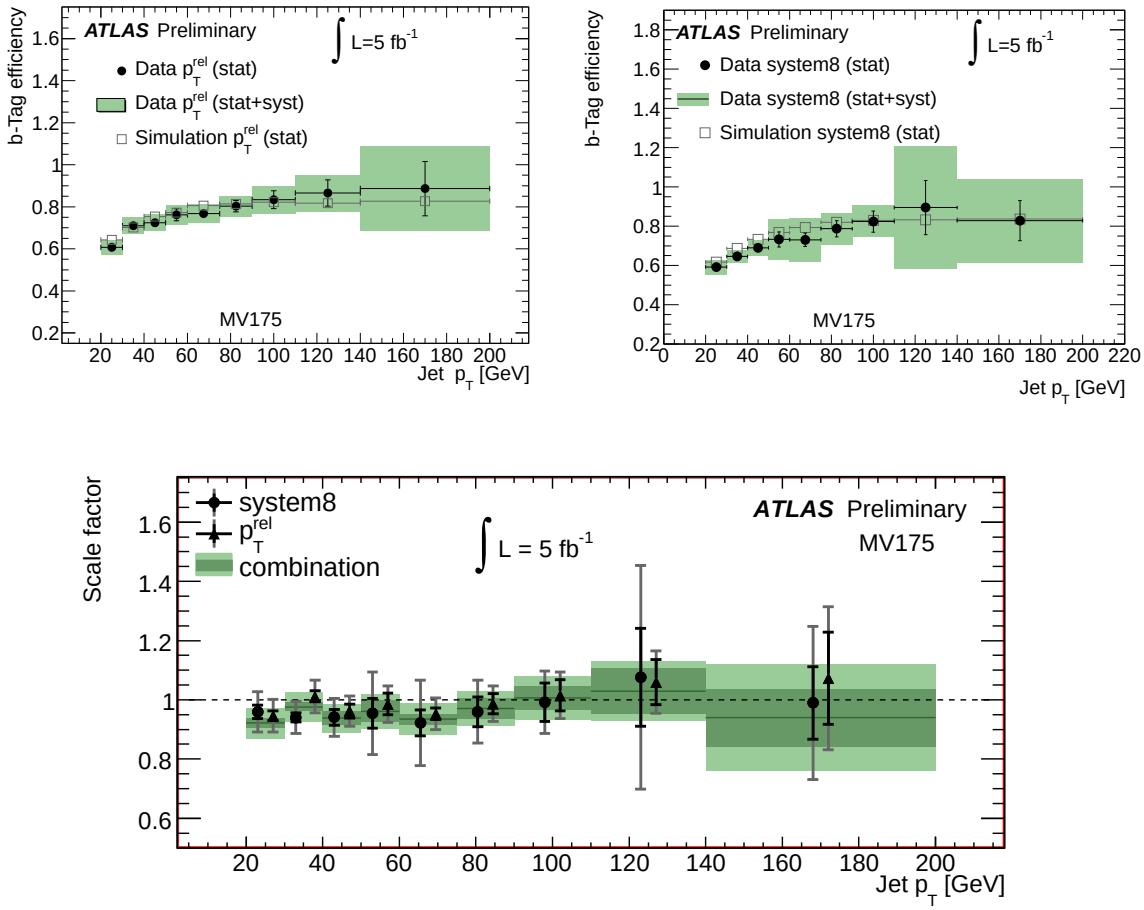

Figure 6: The *b*-tag efficiency in data and simulation for the $p_{\mathrm{T}}^{\mathrm{rel}}$ method (top left) and the system8 method (top right) as well as the individual and combined data-to-simulation scale factors (bottom) for the MV1 tagging algorithm at 75% efficiency. The dark green band represents the statistical uncertainty of the combined scale factor while the light green band shows the total uncertainty. The data points showing the $p_{\mathrm{T}}^{\mathrm{rel}}$ and system8 measurements have been separated a little along the *x*-axis to make the plot more readable.

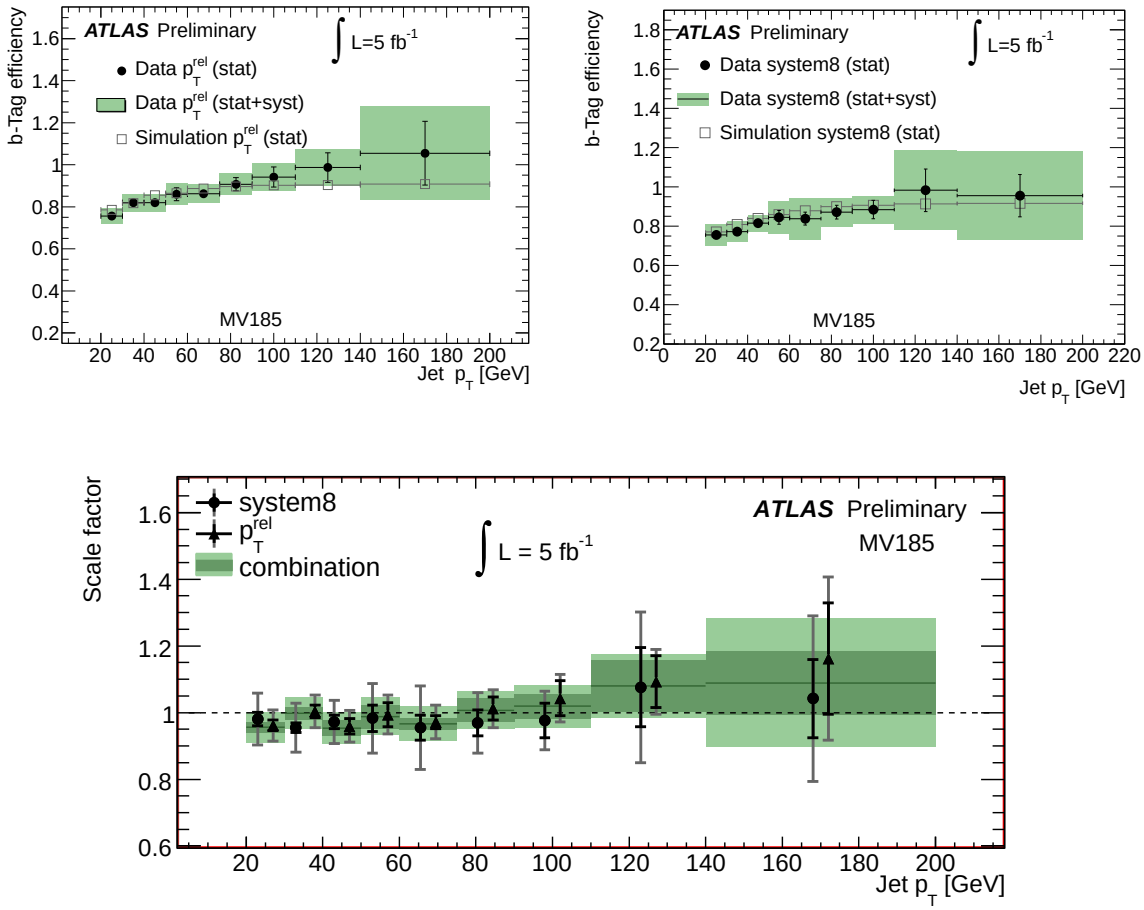

Figure 7: The *b*-tag efficiency in data and simulation for the $p_{\mathrm{T}}^{\mathrm{rel}}$ method (top left) and the system8 method (top right) as well as the individual and combined data-to-simulation scale factors (bottom) for the MV1 tagging algorithm at 85% efficiency. The dark green band represents the statistical uncertainty of the combined scale factor while the light green band shows the total uncertainty. The data points showing the $p_{\mathrm{T}}^{\mathrm{rel}}$ and system8 measurements have been separated a little along the *x*-axis to make the plot more readable. For this high-efficiency operating point, the efficiency measured in some p_{T} bins exceeds 100%, but is still compatible with 100% within uncertainties.

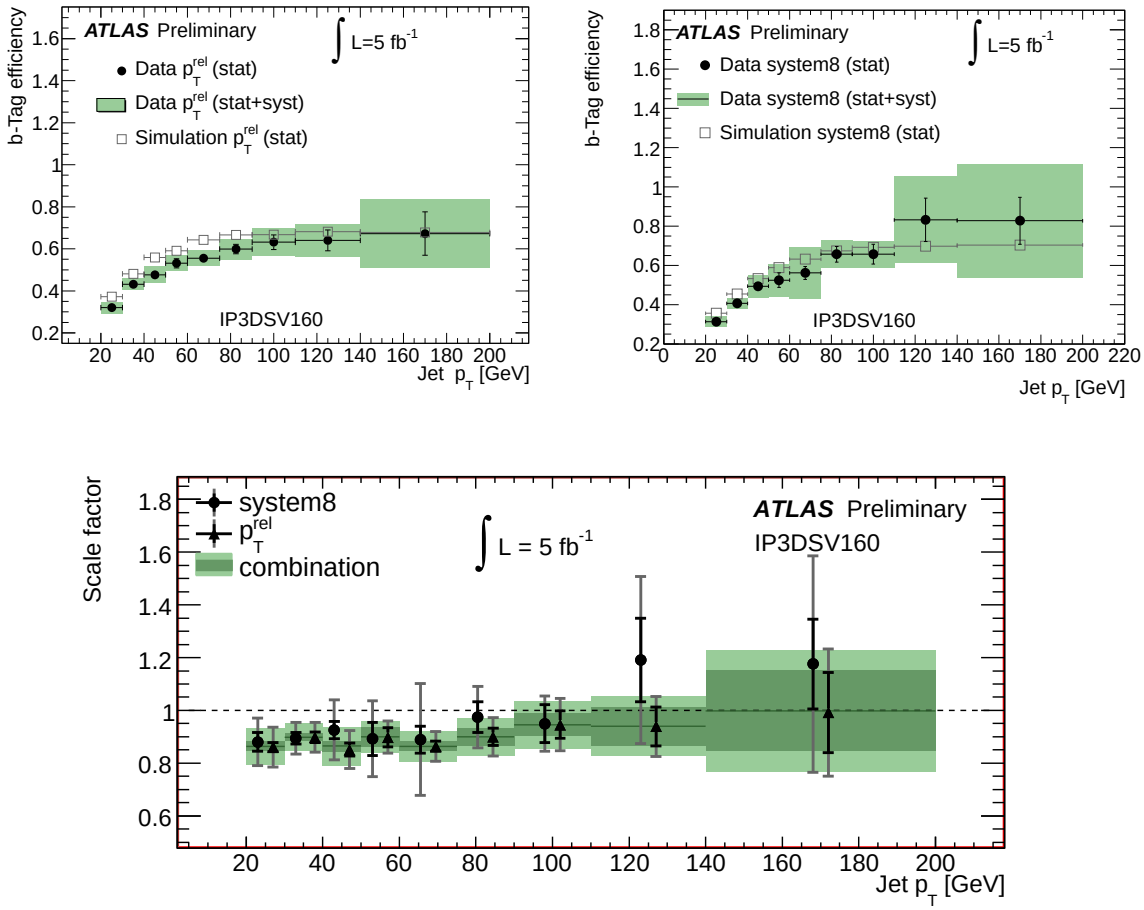

Figure 8: The *b*-tag efficiency in data and simulation for the $p_{\mathrm{T}}^{\mathrm{rel}}$ method (top left) and the system8 method (top right) as well as the individual and combined data-to-simulation scale factors (bottom) for the IP3D+SV1 tagging algorithm at 60% efficiency. The dark green band represents the statistical uncertainty of the combined scale factor while the light green band shows the total uncertainty. The data points showing the $p_{\mathrm{T}}^{\mathrm{rel}}$ and system8 measurements have been separated a little along the *x*-axis to make the plot more readable.

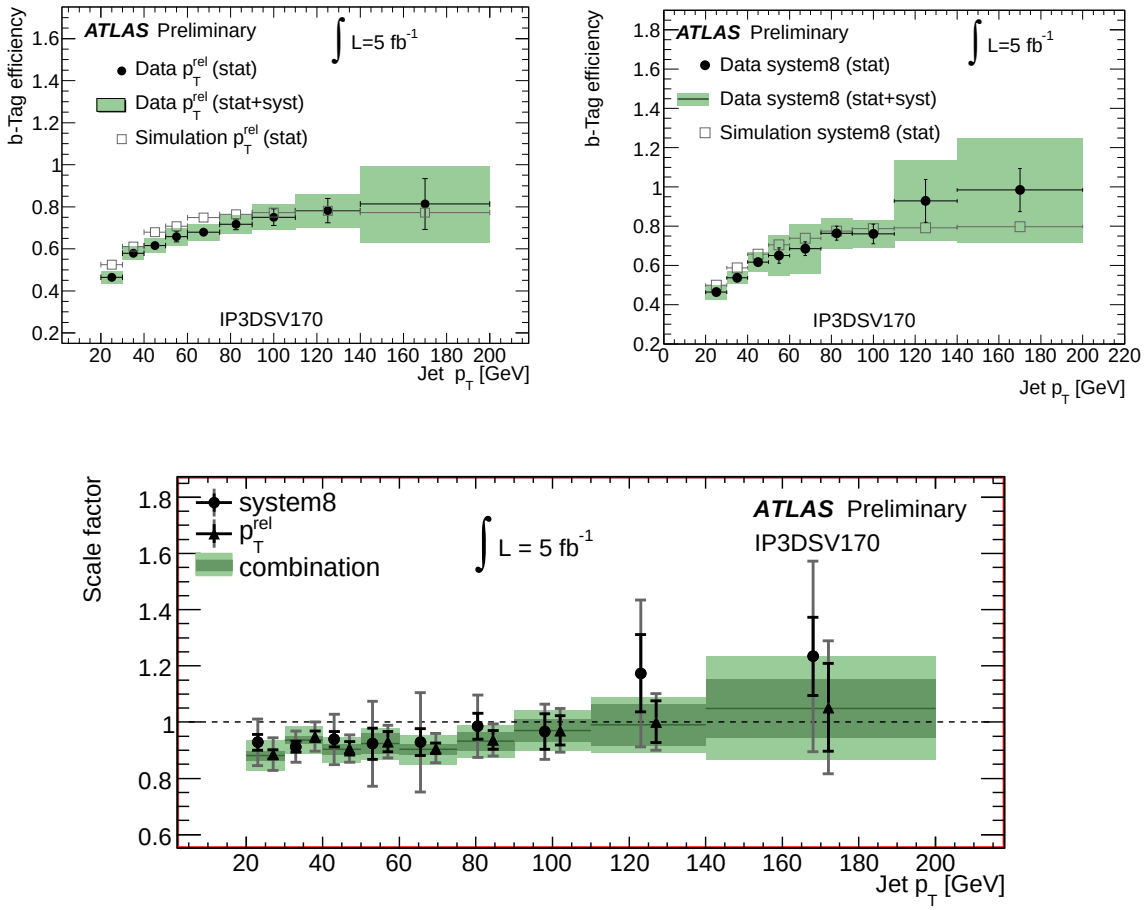

Figure 9: The *b*-tag efficiency in data and simulation for the $p_{\mathrm{T}}^{\mathrm{rel}}$ method (top left) and the system8 method (top right) as well as the individual and combined data-to-simulation scale factors (bottom) for the IP3D+SV1 tagging algorithm at 70% efficiency. The dark green band represents the statistical uncertainty of the combined scale factor while the light green band shows the total uncertainty. The data points showing the $p_{\mathrm{T}}^{\mathrm{rel}}$ and system8 measurements have been separated a little along the *x*-axis to make the plot more readable.

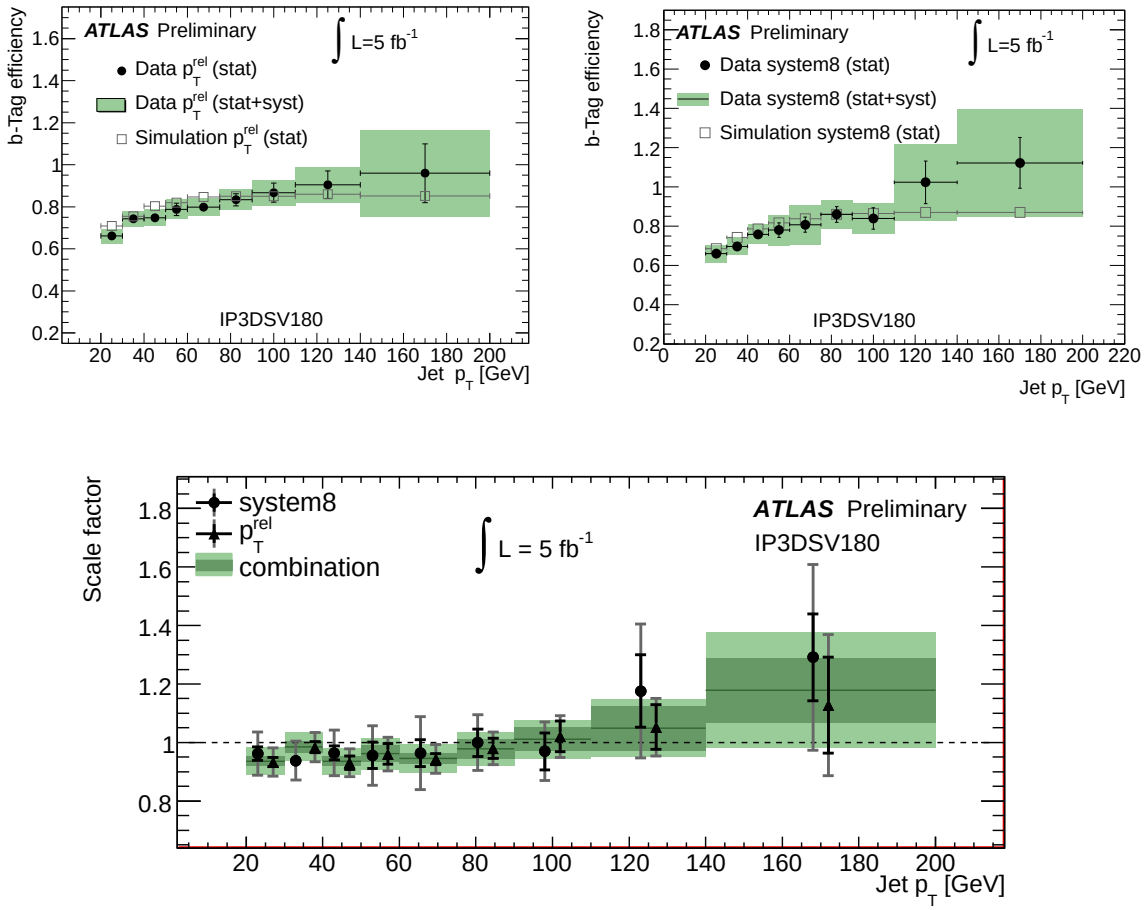

Figure 10: The *b*-tag efficiency in data and simulation for the $p_{\mathrm{T}}^{\mathrm{rel}}$ method (top left) and the system8 method (top right) as well as the individual and combined data-to-simulation scale factors (bottom) for the IP3D+SV1 tagging algorithm at 80% efficiency. The dark green band represents the statistical uncertainty of the combined scale factor while the light green band shows the total uncertainty. The data points showing the $p_{\mathrm{T}}^{\mathrm{rel}}$ and system8 measurements have been separated a little along the *x*-axis to make the plot more readable. For this high-efficiency operating point, the efficiency measured in some p_{T} bins exceeds 100%, but is still compatible with 100% within uncertainties.

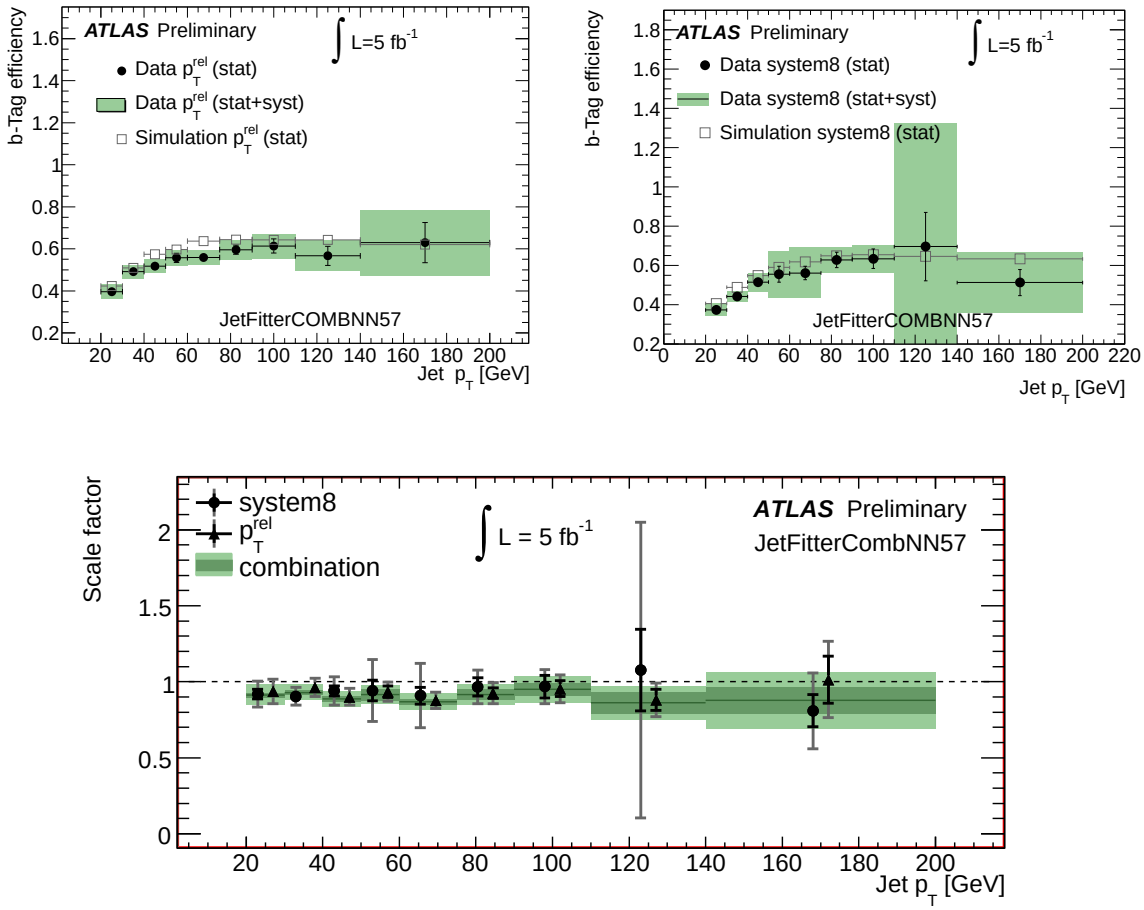

Figure 11: The *b*-tag efficiency in data and simulation for the $p_{\mathrm{T}}^{\mathrm{rel}}$ method (top left) and the system8 method (top right) as well as the individual and combined data-to-simulation scale factors (bottom) for the JetFitterCombNN tagging algorithm at 57% efficiency. The dark green band represents the statistical uncertainty of the combined scale factor while the light green band shows the total uncertainty. The data points showing the $p_{\mathrm{T}}^{\mathrm{rel}}$ and system8 measurements have been separated a little along the *x*-axis to make the plot more readable.

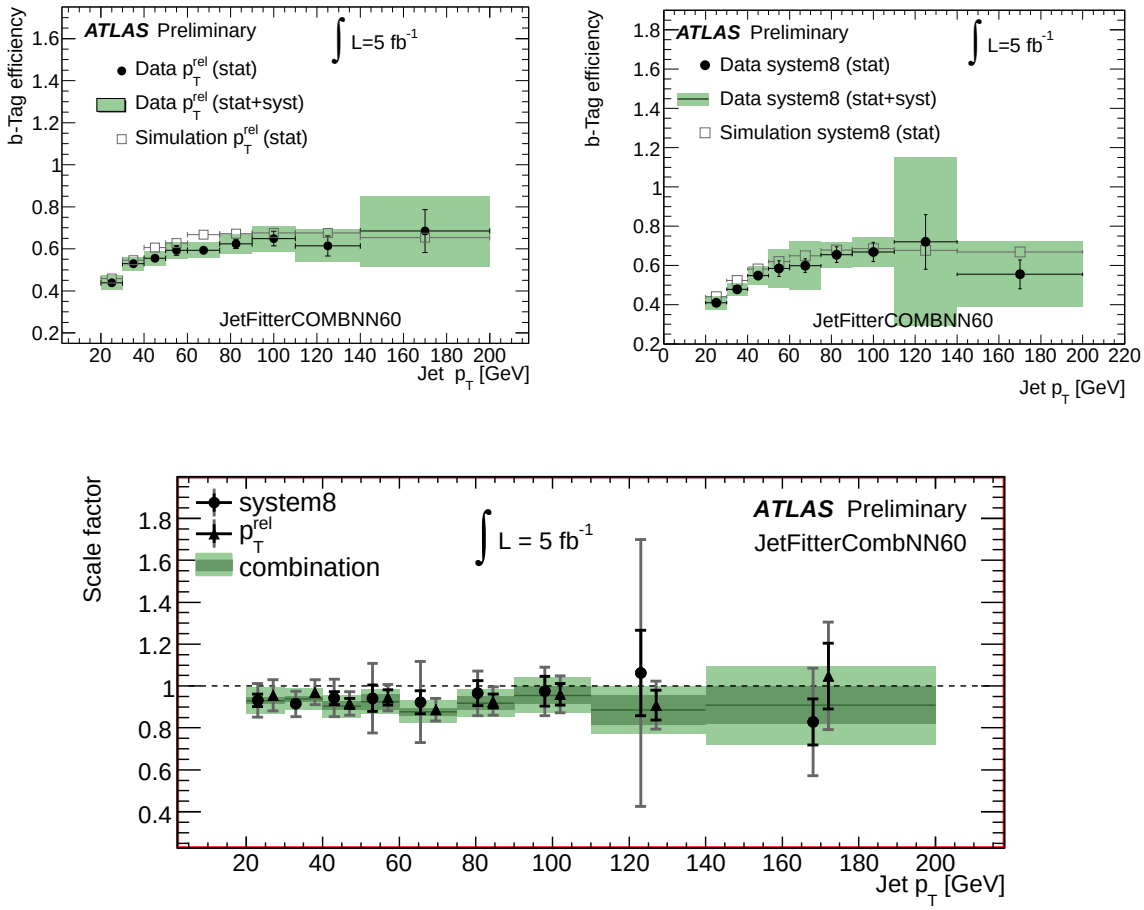

Figure 12: The *b*-tag efficiency in data and simulation for the $p_{\mathrm{T}}^{\mathrm{rel}}$ method (top left) and the system8 method (top right) as well as the individual and combined data-to-simulation scale factors (bottom) for the JetFitterCombNN tagging algorithm at 60% efficiency. The dark green band represents the statistical uncertainty of the combined scale factor while the light green band shows the total uncertainty. The data points showing the $p_{\mathrm{T}}^{\mathrm{rel}}$ and system8 measurements have been separated a little along the *x*-axis to make the plot more readable.

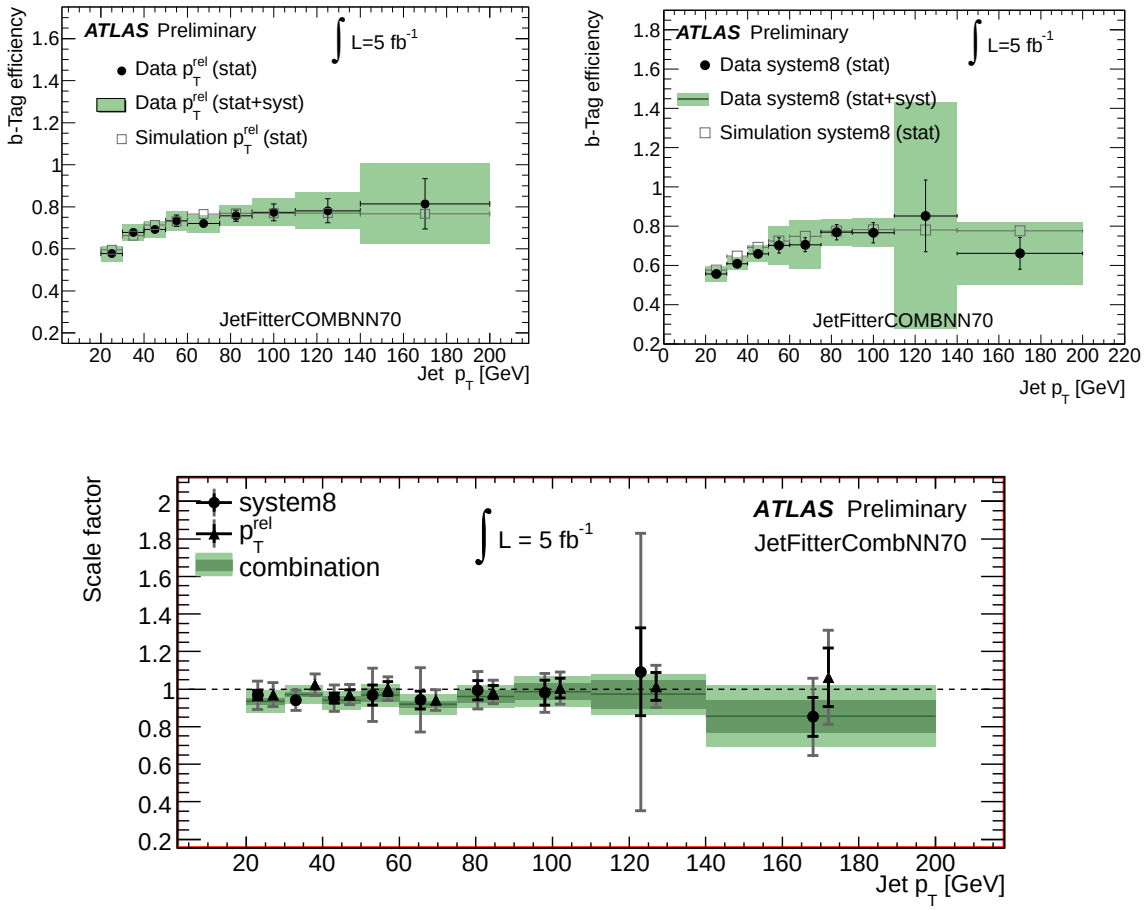

Figure 13: The *b*-tag efficiency in data and simulation for the $p_{\mathrm{T}}^{\mathrm{rel}}$ method (top left) and the system8 method (top right) as well as the individual and combined data-to-simulation scale factors (bottom) for the JetFitterCombNN tagging algorithm at 70% efficiency. The dark green band represents the statistical uncertainty of the combined scale factor while the light green band shows the total uncertainty. The data points showing the $p_{\mathrm{T}}^{\mathrm{rel}}$ and system8 measurements have been separated a little along the *x*-axis to make the plot more readable.

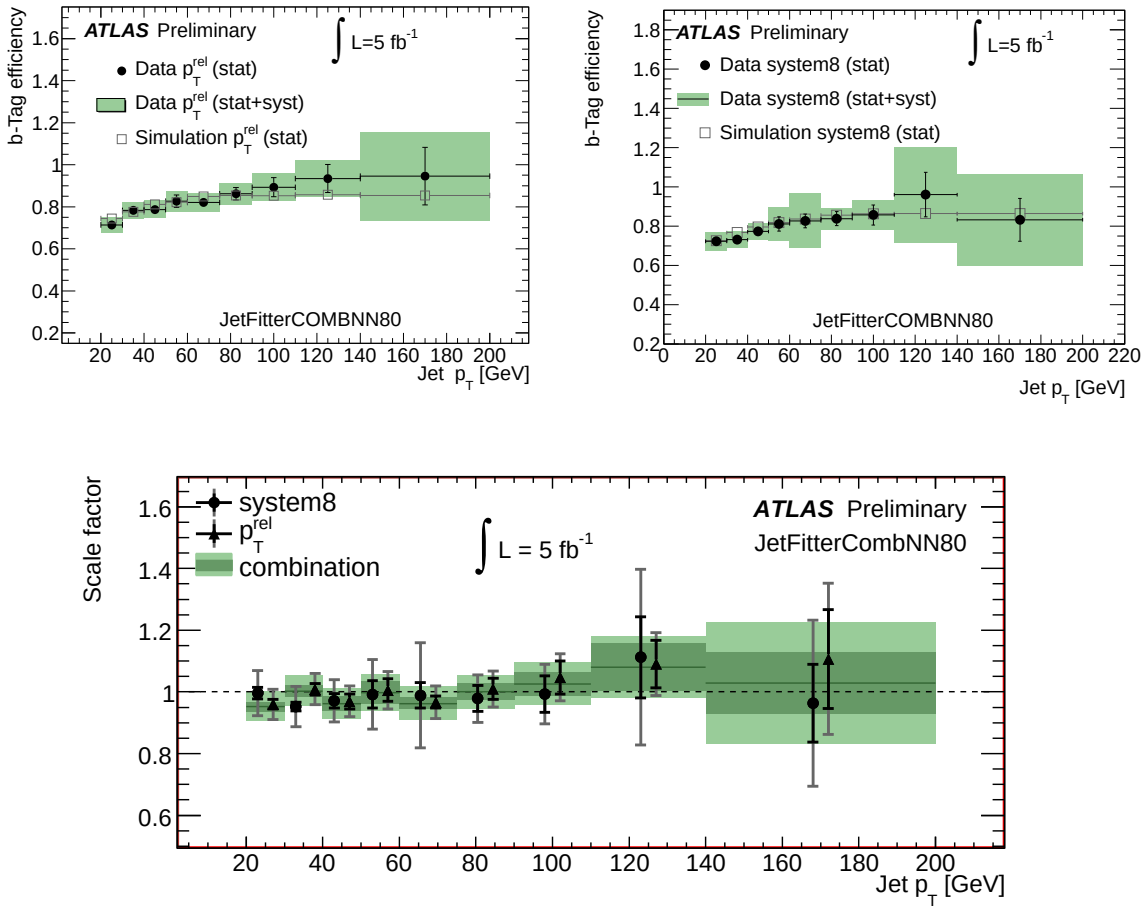

Figure 14: The *b*-tag efficiency in data and simulation for the $p_{\mathrm{T}}^{\mathrm{rel}}$ method (top left) and the system8 method (top right) as well as the individual and combined data-to-simulation scale factors (bottom) for the JetFitterCombNN tagging algorithm at 80% efficiency. The dark green band represents the statistical uncertainty of the combined scale factor while the light green band shows the total uncertainty. The data points showing the $p_{\mathrm{T}}^{\mathrm{rel}}$ and system8 measurements have been separated a little along the *x*-axis to make the plot more readable. For this high-efficiency operating point, the efficiency measured in some p_{T} bins exceeds 100%, but is still compatible with 100% within uncertainties.

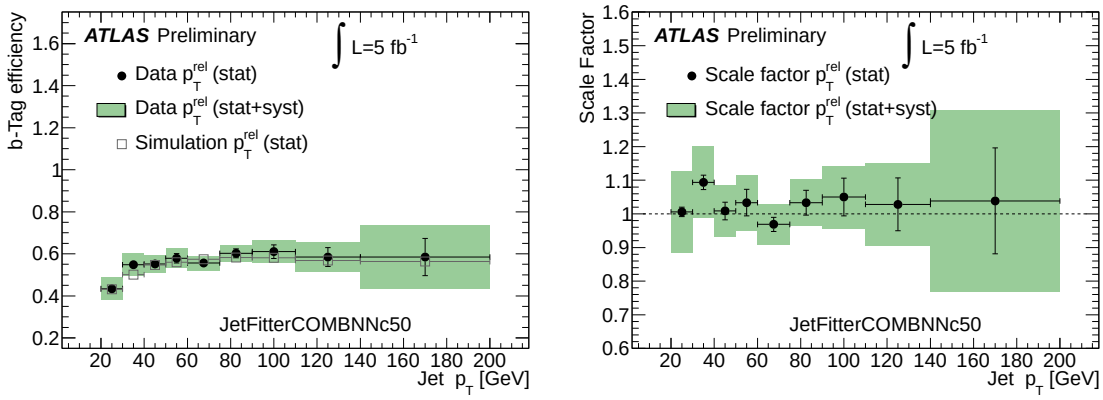

Figure 15: The *b*-tag efficiency in data and simulation (left) and the data-to-simulation scale factor (right) for the JetFitterCombNNc tagging algorithm at 50% efficiency obtained with the $p_{\mathrm{T}}^{\mathrm{rel}}$ method.

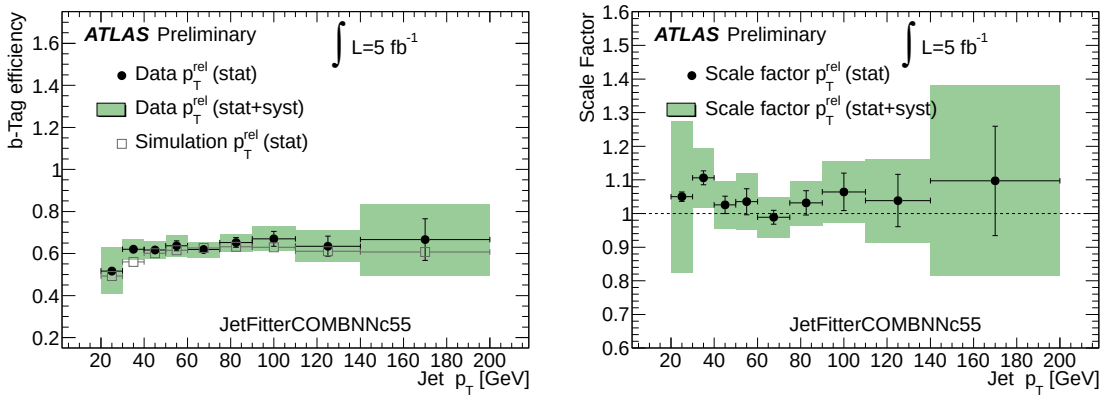

Figure 16: The *b*-tag efficiency in data and simulation (left) and the data-to-simulation scale factor (right) for the JetFitterCombNNc tagging algorithm at 55% efficiency obtained with the $p_{\mathrm{T}}^{\mathrm{rel}}$ method.

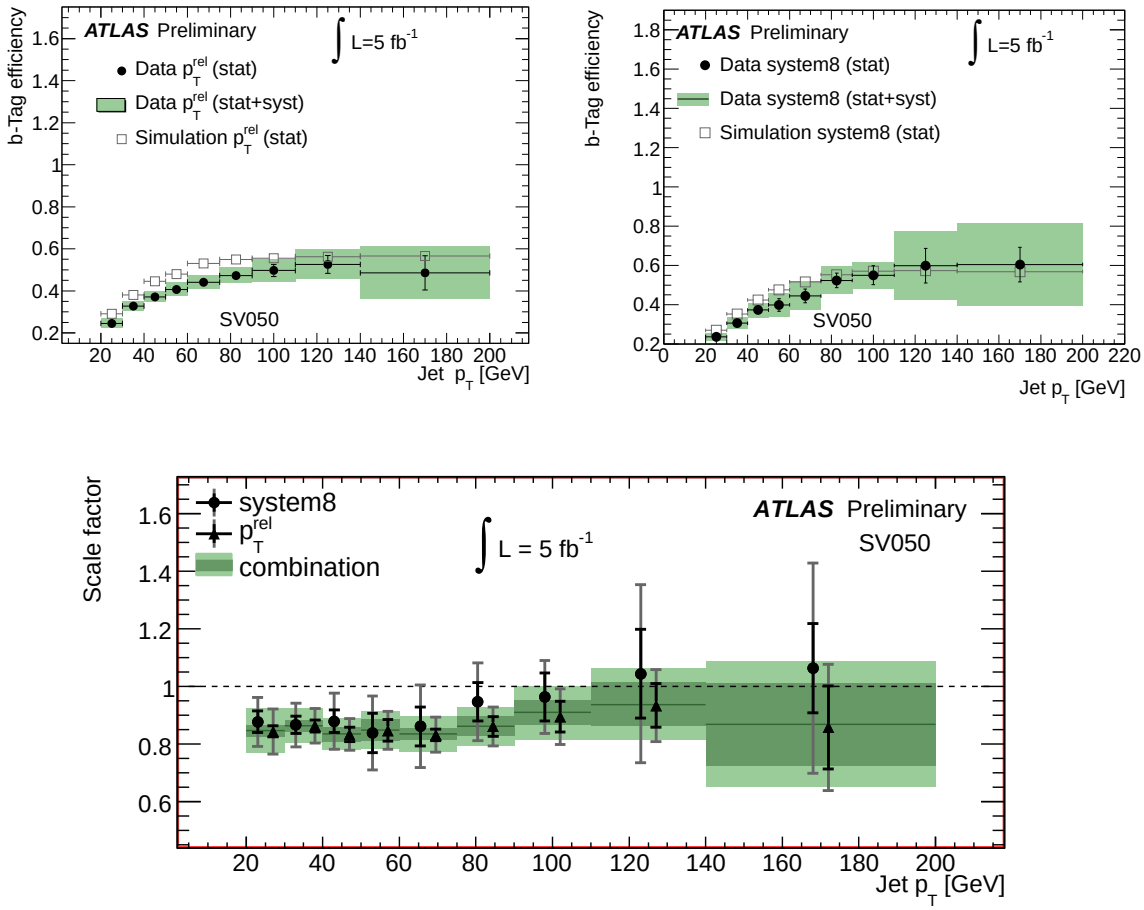

Figure 17: The *b*-tag efficiency in data and simulation for the $p_{\mathrm{T}}^{\mathrm{rel}}$ method (top left) and the system8 method (top right) as well as the individual and combined data-to-simulation scale factors (bottom) for the SV0 tagging algorithm at 50% efficiency. The dark green band represents the statistical uncertainty of the combined scale factor while the light green band shows the total uncertainty. The data points showing the $p_{\mathrm{T}}^{\mathrm{rel}}$ and system8 measurements have been separated a little along the *x*-axis to make the plot more readable.