Measurement of the b-tagging efficiency using $t\bar{t}$ events

The CMS Collaboration

Abstract

Jet flavour tagging is one of the key ingredients of the diverse physics program of the CMS experiment. Various flavour taggers based on the lifetime and semileptonic decay of the heavy quarks have been implemented. They use jets, tracks, leptons and/or vertices. We exploit the $t\bar{t}$ samples collected at $\sqrt{s} = 7$ TeV running of the LHC during the first half of 2011, to extract the efficiencies for tagging the b-quark and c-quark jets. Event samples identified as semileptonic and fully-leptonic decays of the $t\bar{t}$ are used. Different techniques based on extraction of high purity b-flavoured jets, or on requiring the consistency between the observed and expected number of tags in the events are used to extract the performance of the heavy flavour algorithms directly from data and measure the differences between data and simulation.
1 Introduction

The ability to identify the jets arising from heavy flavour quarks plays a central role in many physics analyses such as precision measurements of top quark properties, searches for Higgs bosons, and searches for signals of physics beyond the Standard Model (SM). Flavour tagging algorithms are thus used to identify these jets and reduce the otherwise overwhelming background from processes involving jets from gluon and light quarks (u, d, s) fragmentation (light flavour, or \(\ell\)) and from charm-quark hadronisation. Analyses also need an accurate estimate of the performance (efficiency and mis-identification rate) of the chosen algorithm. In general, these quantities are functions of the transverse momentum and pseudorapidity of a jet [1]. The performance of the algorithms also depends on parameters like the material of the tracker, the hit resolution, the efficiency of the track reconstruction, the resolution of the reconstructed track parameters, and the track density in a jet. While the Compact Muon Solenoid (CMS) physics and detector simulation reproduces data with a remarkable degree of precision, it is difficult to control all the parameters relevant for b tagging. It is therefore essential to measure the performance of the algorithms directly from data.

The large number of top quark pairs produced at the Large Hadron Collider (LHC) can be used to measure the performance of the heavy flavour tagging algorithms directly from collider data, and further to calibrate the tagging algorithms. Indeed, with a \(t\bar{t}\) production cross section of \(165.8 \pm 13.3\) pb [2] as measured by CMS, a large sample of b jets originating from the decay of the top quarks is collected. This allows a wide range of transverse momenta \(p_T\) to be probed, with an average \(p_T\) of the jets of about 78 GeV/c. This range is relevant for many processes both within the SM and for many models beyond the SM.

In the framework of the SM, the top quark is expected to decay to a W boson and a b quark about 99.8% of the time [3]. With three quark generations, the magnitude of the Cabibbo-Kobayashi-Maskawa (CKM) matrix element \(|V_{tb}|\) is expected to be close to unity as a consequence of unitarity and the small size of \(|V_{td}|\) and \(|V_{ts}|\). Experimentally, the measurement of the heavy flavour content of \(t\bar{t}\) events can provide either a direct measurement of the branching fraction of the decay of the top quark to a W boson and a b quark, \(B(t \rightarrow Wb)\), or, assuming \(B(t \rightarrow Wb) = 1\), the b tagging performance.

In this note, we present several methods to study the heavy flavour content of \(t\bar{t}\) events. These methods allow the efficiency of tagging b jets to be measured in both the data and the Monte Carlo simulation over the average \(p_T\) and \(\eta\) range of jets in the top quark events. The differences in efficiencies observed between the data and Monte Carlo are provided as a data-to-Monte Carlo scale factor \(\text{SF}_b\). These methods use \(t\bar{t}\) events in the lepton+jets channel, in which one W boson decays into quarks and the other into a charged lepton and a neutrino, and the dilepton channel in which both W bosons decay into leptons. The measurements are performed on samples collected in 2011 by the CMS experiment at the LHC at a centre-of-mass energy of \(\sqrt{s} = 7\) TeV. These samples correspond to an integrated luminosity of \(2.18 \pm 0.11\) fb\(^{-1}\).

2 The CMS detector

The CMS detector is described in detail in Reference [4]. The central feature of the CMS apparatus is a superconducting solenoid that provides an axial magnetic field of 3.8 T. Charged particle trajectories are measured by the tracker, covering \(0 < \phi < 2\pi\) in azimuth and \(|\eta| < 2.5\), where the pseudorapidity \(\eta\) is defined as \(\eta = -\ln[\tan(\theta/2)]\), with \(\theta\) being the polar angle of the trajectory of the particle with respect to the counterclockwise beam direction. The calorimetry provides high-resolution energy and direction measurements of electrons and hadronic jets.
Muons are measured in gas-ionisation detectors embedded in the steel return yoke outside the solenoid. The detector is nearly hermetic, allowing for energy balance measurements in the plane transverse to the beam directions. A two-level trigger system selects the most interesting collision events for use in physics analysis.

The tracker [5, 6] is thus composed of a Pixel detector, providing two to three measurement points, followed by a Silicon Strip Tracker (SST) providing 10 to 14 measurement points per track. Due to the fine granularity, and hence the low occupancy obtained, pattern recognition problems are solved after the first few layers, and track parameter resolutions reach an asymptotic value after using only the first five to six hits. With the strong magnetic field, charged particles are reconstructed with an impact parameter resolution around 15 $\mu$m and a transverse momentum ($p_T$) resolution of $\sim 1.5\%$ for transverse momenta up to 100 GeV/$c$.

The Pixel detector is composed of three cylindrical barrel layers and two pairs of disks in the end-caps, such that three points are measured per tracks for $|\eta| < 2.2$. With a pixel size of $100 \times 150\mu m^2$, the hit resolution is approximately of 10 $\mu$m in $r - \phi$, using charge sharing induced by the Lorentz angle of 23$^\circ$, and 20 $\mu$m in $r - z$.

The SST is divided in four parts. The Inner Barrel is constituted of four cylindrical layers, enclosed by three pairs of disks (Inner Disks – TID). It is then followed by the six cylindrical layers of the Outer Barrel, and the End-Caps (TEC) are made of nine pairs of disks. The disks of the TID are composed of three rings of modules and the TEC disks of 7 rings. With this variety of detectors, there are 14 different sensor geometries, with pitches ranging from 80 to 205 $\mu$m.

### 3 The heavy flavour tagging algorithms

Several flavour tagging algorithms have been implemented within the reconstruction framework of the CMS experiment. They range from comparatively simple and robust approaches based on the presence of leptons to complex multi-variate techniques extracting lifetime and kinematic information from displaced vertices.

The main feature used to identify b jets from the light-flavour jets is the long lifetime of b hadrons. Simpler algorithms rely on the presence of tracks with large impact parameter with respect to the primary vertex, while others will attempt to identify the b-decay vertex and use the decay length as a rejection criterion against the mostly short-lived background. Further properties used are the large mass of the quark, its decay to final states with high charged particle multiplicities, with an average of five charged particles, and a hard fragmentation function. In addition, since the branching fraction of b quarks to an electron or a muon is of the order of 20%, the presence of either of these leptons can be used as well, although the efficiency is then limited by the branching fraction. Since only the performance of the lifetime taggers is presented in this note, these will be discussed in more detail in the following section.

The b-tagging algorithms rely thus on the capabilities and performance of the tracking detectors, particularly the pixel detector, and the associated reconstruction algorithms. Most of the b-hadron properties used for b tagging are exploited using charged particle tracks because only tracking detectors offer the spatial resolution needed to detect, for example, the significant decay length of b hadrons. Efficient track reconstruction, and in particular precise spatial reconstruction close to the interaction point, is thus the key ingredient for almost all b-tagging algorithms.

The jets to be tagged are reconstructed with the particle flow algorithm anti-$k_T$ clustering method with cone radius $\Delta R = \sqrt{\Delta \eta^2 + \Delta \phi^2} = 0.5$ [7, 8]. The energy scale of the jets is
corrected and jet identification criteria are applied to select jets of good quality [9]. The tagging algorithms are applied to jets with transverse momenta above 30 GeV/c and pseudorapidity within $|\eta| < 2.5$ to ensure that the particles in the jet are within the acceptance of the tracker. The jet direction is typically taken to approximate the original flight path of the $b$-hadron.

Tracks are reconstructed using an iterative tracking procedure based on the Kalman filter [10]. These are then associated to the jets if their distance to the jet axis is within $\Delta R < 0.5$. To suppress fake and badly reconstructed tracks, basic track quality requirements are imposed by the tagging algorithms, requiring for example that at least two pixel hits are used.

The primary vertex [11] is reconstructed from all tracks in the event that are compatible with the beam spot, the location of the LHC beam in the $(x,y)$ plane. Vertices are first identified using a deterministic annealing (DA) method and fitted using the Adaptive Vertex Fitter [12] algorithm. The vertex with the highest squared sum of the transverse momenta of the tracks is selected as the primary vertex in which the hard scattering occurred.

For each algorithm, standard operating points are defined according to the expected misidentification rate of light flavour jets. Three different operating points (“loose” (L), “medium” (M), and “tight” (T)) are thus defined to yield a mistag rate estimated from QCD Monte-Carlo simulation of 10%, 1%, or 0.1%, respectively, for jets with $p_T$ of about 80 GeV/c. In the following, acronyms will be used to quote the algorithms with operating points, where the first letters refer to the name of the algorithm, as described in this section, followed by the one letter abbreviation of the operating point. For instance, TCHPM refers to the “track-counting high purity” algorithm (TCHP) used at the “medium” (M) operating point.

### 3.1 Track impact parameter taggers

The simplest algorithms are based on the impact parameters of the tracks in the jets, as this is the most powerful single-track observable. The impact parameter (IP) of a track is defined as the distance between the track and the primary vertex or beamspot at the point of closest approach. It is calculated in three dimensions, taking advantage of the high resolution of the pixel detector along the $z$ axis. It is signed such that it is positive for tracks reconstructed to originate downstream from the primary vertex in the direction of the jet and negative otherwise. In the tagging algorithms, the impact parameter significance (the ratio of the track impact parameter to its uncertainty – $\text{IP}/\sigma_{\text{IP}}$) is used instead of the IP since the uncertainty on the IP can be of the same order as the IP itself.

The simplest algorithm is the track counting, which identifies a jet as a $b$-jet if it contains at least $N$ tracks with an IP significance above a given threshold. Specifically, tracks are ordered in decreasing impact parameter significance and the impact parameter significance of the $N$th track thus serves as the discriminator for this algorithm. Two variants of the algorithms are implemented, using the significance of either the second or the third track as discriminator. By using a second track a higher $b$-jet identification efficiency is obtained while a higher tagging purity is obtained by using the third track. The two algorithms are hence called track counting high efficiency (TCHE) and high purity (TCHP), respectively.

The jet probability algorithms are a natural extension of the track counting algorithms, where the information of all selected tracks is used instead of using only the first few tracks. For each track, the probability of coming from the primary vertex is computed and these probabilities are combined together to provide the jet probability. Two discriminators are defined. For the jet probability (JP) algorithm, the discriminator is defined as probability that all tracks with positive impact parameter come from the primary vertex. In the second algorithm, called jet B...
probability (JBP), the discriminator reflects the probability that the four most displaced tracks are compatible with the primary vertex.

### 3.2 Secondary vertex taggers

Due to the long life time of b hadrons, secondary vertices can be used to select jets from b hadrons with high purity. Secondary vertices are reconstructed using the “Adaptive Vertex Finder” [12], which performs a fully inclusive vertex search among the tracks associated to the jet. The approach is to fit a vertex from all available tracks and iteratively repeat the fit with the tracks that were not compatible with the vertices obtained in previous iterations. This procedure is repeated until the list of tracks is exhausted or the vertex fit fails. The vertices are fitted with the “Adaptive Vertex Fitter” [13], an iterative re-weighted fitter which down-weights tracks according to their standardized distance to the vertex. Outlying tracks (badly reconstructed tracks or tracks not compatible with the vertex) will thus have a low weight in the fit and can be re-used in subsequent iterations of the finder. The first vertex reconstructed will be the primary vertex, and this will not be considered subsequently. If more than one secondary vertex is reconstructed, the vertex with the smallest error on the flight distance is chosen as the preferred secondary vertex. This procedure is obviously limited by the branching fraction of b hadrons to final states with at least two charged particles.

A simple version, called simple secondary vertex [14] tagging algorithm relies on the presence of at least one reconstructed secondary vertex, with a discriminator based on the flight distance significance. No discriminator value is returned if no secondary vertex is found. This algorithm is found to be less sensitive to the alignment precision of the tracking system. Two variants of the algorithm are implemented. In the simple secondary vertex high efficiency (SSVHE), all vertices are considered, while in the simple secondary vertex high purity (SSVHP), only vertices which are reconstructed with at least three tracks are used.

A more complex approach involves the use of secondary vertices together with different topological and kinematic variables such as the $IP$ significance or the decay length. By using these additional variables, the combined secondary vertex (CSV) algorithm provides discrimination even when no secondary vertices are found, ensuring that the b-tagging efficiency is not limited by the secondary vertex reconstruction efficiency. In many cases, tracks with an $IP$ significance above 2 can be used together to compute a subset of the quantities derived from a secondary vertex even without an actual vertex fit. When even this is not possible, a “no vertex” category reverts simply to track based variables similar to the “jet probability” algorithm. All variables are used as input to a Likelihood Ratio, used twice to discriminate between b- and c-jets and between b- and light jets, and then combined additively with factors of 0.75 and 0.25, respectively.

### 4 Signal and Background Modeling

The efficiency for selecting lepton+jets signal events is determined using a simulated t$t$ event sample assuming a top quark mass of $m_t = 172.5\text{ GeV}/c^2$. The simulation of t$t$ events is performed using the MADGRAPH [15] Monte Carlo generator, where the top quark pairs are generated accompanied by up to three additional hard jets. The parton configurations generated by MADGRAPH are processed with PYTHIA [16] to provide showering of the generated particles. The showers are matched using the $k_T$-MLM prescription [17]. The generated events are then passed through the full CMS detector simulation based on GEANT4 [18].

The electro-weak production of single top quarks is considered a background process, and
is simulated using POWHEG. The production of W/Z + jets events where the vector boson decays leptonically has a signature similar to tt and constitutes the main background. These are simulated using MADGRAPH, with up to four jets. The W/Z + jets events are inclusive with respect to jet flavour. The bottom and charm components are separated from the light flavour (uds and gluon) components in the analysis by matching reconstructed jets to partons in the simulation. In addition, QCD multi-jets are also considered as a background and samples are produced using PYTHIA.

The background processes are normalized to next-to leading order (NLO) and next-to next-leading order (NNLO) cross section calculations, as listed below, with the exception of the QCD background, which is normalized using the fit to the missing transverse energy ($E_T$) distribution in data.

The NLO tt production cross section has been calculated to be $\sigma_{tt} = 157^{+23}_{-24}$ pb, using MCFM [19]. The uncertainty in this cross section includes the scale uncertainties, estimated by varying the factorization and renormalization scales by a factor 2 around the nominal scale of $(2m_t)^2 + (\sum_{j} p_{T,j})^2$, with $m_t = 172.5$ GeV/$c^2$. The uncertainties from the parton distribution functions (PDF) and the value of $\alpha_S$ are estimated following the procedures from the MSTW2008 [20], CTEQ6.6 [21], and NNPDF2.0 [22] sets. The uncertainties are then combined according to the PDF4LHC prescriptions [23].

Similarly, the t-channel single top NLO cross-section has been determined to be $\sigma_t = 64.6^{+3.4}_{-3.2}$ pb using MCFM [19, 24–26]. The uncertainty is evaluated similarly as for top-quark pair production. The single top-quark associated production (tW) cross section has been set to $\sigma_{tW} = 15.7 \pm 1.2$ pb [27]. The s-channel single top next-to next-to leading log (NNLL) cross-section has been determined as $\sigma_t = 4.6 \pm 0.06$ pb [28].

The inclusive NNLO cross section of the production of W bosons multiplied by its branching fraction to leptons has been determined to be $\sigma_{W\rightarrow l\nu} = 31.3 \pm 1.6$ nb using FEWZ [29], setting renormalization and factorization scales (the so-called "Q$^2$" scale) to $(m_W)^2 + (\sum_{j} p_{T,j})^2$ with $m_W = 80.398$ GeV/$c^2$. The uncertainty was determined in a similar way as for top-quark pair production. The normalizations of the W+b jets and W+c jets components have been determined in a measurement of the top pair production cross section in the lepton+jet channel [30] where a simultaneous fit of the tt cross section and the normalization of the main backgrounds was performed.

Finally, the Drell-Yan (DY) production cross section at NNLO has been calculated using FEWZ as $\sigma_{Z/\gamma^*\rightarrow ll}(m_{ll} > 50$ GeV/$c^2) = 3048 \pm 132$ pb, where $m_{ll}$ is the invariant mass of the two leptons and the scales were set using the Z boson mass $m_Z = 91.1876$ GeV/$c^2$ [3]. In the analyses using the dilepton decay channel, where DY events are the dominant background, the number of these events remaining after all selection requirements is measured on data. Two different methods are used, and the two estimates are compatible. In the analysis (c.f. Section 7), the ratio of DY events outside and inside the dilepton invariant mass window $R_{out/in}$, which is estimated from the simulation, is used to estimate the DY background using the number of data events inside the dilepton invariant mass window [31]. As contamination from non-DY backgrounds can still be present in the Z mass window, this contribution is subtracted using the ee channel scaled according to the event yields in ee and $\mu\mu$ channels. In the Flavour Tag Matching method (c.f. Section 9), the number of DY events is estimated from the shape of the distribution of the angle between the momentum of the two leptons.
Table 1: Number of observed and expected events in the lepton+jets sample after applying all selection requirements. All MC samples have been scaled to a luminosity of 2.18 fb$^{-1}$. The uncertainties include the uncertainties on the luminosity and the cross sections. The TCHEL operating point has been used for the b-tagging requirement.

<table>
<thead>
<tr>
<th></th>
<th>no b-tagging</th>
<th>$\geq$ 1 b-tagged jets</th>
<th>$\geq$ 2 b-tagged jets</th>
</tr>
</thead>
<tbody>
<tr>
<td>tt</td>
<td>7741 ± 1161</td>
<td>6514 ± 977</td>
<td>2939 ± 441</td>
</tr>
<tr>
<td>Single top</td>
<td>437±105</td>
<td>350±83</td>
<td>108±23</td>
</tr>
<tr>
<td>W+Jets</td>
<td>5673 ± 1701</td>
<td>1247 ± 374</td>
<td>153 ± 46</td>
</tr>
<tr>
<td>Z+Jets</td>
<td>422 ± 127</td>
<td>72 ± 22</td>
<td>12 ± 4</td>
</tr>
<tr>
<td>QCD</td>
<td>22 ± 7</td>
<td>8 ± 2</td>
<td>0</td>
</tr>
<tr>
<td>Sum MC</td>
<td>14295 ± 2066</td>
<td>8191 ± 1049</td>
<td>3210 ± 444</td>
</tr>
<tr>
<td>Data</td>
<td>13935</td>
<td>8205</td>
<td>3246</td>
</tr>
</tbody>
</table>

5 The event selection

The event reconstruction used herein follows closely the event selection performed for the tt production cross section measurements [2], with the obvious exception of the b-tagging requirements.

In the lepton+jets channel, the final state is composed of four jets, one energetic isolated muon and missing transverse energy. Events are required to pass the single-muon trigger. After offline reconstruction, events are selected requiring exactly one isolated muon with $p_T > 30$ GeV/c and $|\eta| < 2.1$ and at least four jets with $p_T > 30$ GeV/c and $|\eta| < 2.4$. The Flavour Tag Consistency method (c.f. Section 8) further requires that the two leading jets have transverse momenta greater than 70 GeV/c and 50 GeV/c respectively, and that the transverse momenta of the muon is greater than 35 GeV/c. The jets and the missing transverse energy ($E_T$) are reconstructed with the particle flow algorithm.

In the dilepton channel, the final state is composed of two jets, two energetic isolated leptons (electron or muon) and missing transverse energy. Events are required to pass dilepton triggers in which two muon, two electron, or one electron and one muon were required to be present.

After offline reconstruction, events are selected with two isolated, oppositely-charged leptons (electrons or muons) with $p_T > 20$ GeV/c and $|\eta| < 2.5$ (2.4) for electrons (muons), at least two jets with $p_T > 30$ GeV/c and $|\eta| < 2.4$, and $E_T > 30$ GeV for ee/$\mu\mu$ events. The selected leptons and jets are required to originate from the same primary interaction vertex. Events with same-flavour lepton pairs in the dilepton mass region $76 < m_{ll} < 106$ GeV/c$^2$ are removed to suppress the dominant Z+jet background. Dilepton pairs from heavy flavour resonances and low-mass Drell-Yan are also removed by requiring a minimum invariant mass of 12 GeV/c$^2$.

The numbers of observed and expected events in the lepton+jets channel and the dilepton channel are given in Table 1 and Table 2 respectively. The Drell-Yan background in the ee and $\mu\mu$ channels is estimated by the above-mentioned data-driven method. The uncertainties include the uncertainties on the luminosity and the cross sections. For all MC predictions, events are reweighted to take into account differences in trigger and lepton selection efficiency between data and simulation. The lepton selection efficiency scale factors are estimated with a tag and probe method in Z events. The trigger efficiencies are estimated on a sample of data events using a trigger that is weakly correlated to the dilepton triggers. The dilepton trigger selection efficiency is then estimated on events which contain two leptons that fulfill the complete dilepton event selection.
Table 2: Number of observed and expected events in the Dilepton sample after applying all selection requirements. The number of DY events is estimated from data with the $R_{\text{out/in}}$ method. All other MC samples have been scaled to a luminosity of 2.18 fb$^{-1}$. The uncertainties include the uncertainties on the luminosity and the cross sections. The TCHEL operating point has been used for the b-tagging requirement.

<table>
<thead>
<tr>
<th>Processes</th>
<th>Channel ee</th>
<th>Channel $\mu\mu$</th>
<th>Channel $e\mu$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without b-tagging requirement</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t\bar{t}$ signal</td>
<td>932 ± 141</td>
<td>1223 ± 184</td>
<td>3313 ± 500</td>
</tr>
<tr>
<td>$t\bar{t}$ background</td>
<td>11.0 ± 5.5</td>
<td>3.2 ± 1.6</td>
<td>22.7 ± 11.3</td>
</tr>
<tr>
<td>Single top</td>
<td>46.7 ± 14.0</td>
<td>60.2 ± 18.1</td>
<td>157 ± 47</td>
</tr>
<tr>
<td>Di-bosons</td>
<td>21.4 ± 6.4</td>
<td>28.0 ± 8.4</td>
<td>47.4 ± 14.2</td>
</tr>
<tr>
<td>Z+jets</td>
<td>409 ± 204</td>
<td>545 ± 273</td>
<td>200 ± 100</td>
</tr>
<tr>
<td>W+jets</td>
<td>12.0 ± 6.0</td>
<td>0.0</td>
<td>11.4 ± 5.7</td>
</tr>
<tr>
<td>Sum MC</td>
<td>1432 ± 249</td>
<td>1860 ± 329</td>
<td>3751 ± 512</td>
</tr>
<tr>
<td>Data</td>
<td>1442</td>
<td>1773</td>
<td>3898</td>
</tr>
<tr>
<td>With $\geq 1$ b-tagged jets</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sum MC</td>
<td>1056 ± 165</td>
<td>1388 ± 212</td>
<td>3259 ± 456</td>
</tr>
<tr>
<td>Data</td>
<td>1080</td>
<td>1364</td>
<td>3375</td>
</tr>
<tr>
<td>With $\geq 2$ b-tagged jets</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sum MC</td>
<td>514 ± 71</td>
<td>677 ± 94</td>
<td>1757 ± 253</td>
</tr>
<tr>
<td>Data</td>
<td>554</td>
<td>686</td>
<td>1854</td>
</tr>
</tbody>
</table>

In the selected dilepton events, the average $p_T$ of the two leading jets is 83.1 GeV/c, while a similar selection in the $t\bar{t}$ simulation yields an average of 85.6 GeV/c. After requiring that at least one jet is tagged using the TCHEL criterion, the average $p_T$ of the b-tagged jets is 83.8 GeV/c for the data, and 85.7 GeV/c in simulation.

6 Systematic uncertainties

Most of the sources of systematic uncertainties are common to all methods, and several methods have specific additional contributions. A description of the common systematic uncertainties is given in this section. The description of the procedure to estimate the systematic uncertainties in each analysis and the influence of the different sources will be given separately for each analysis in its relevant section.

There are different sources of uncertainties originating from detector knowledge or related to the theory and the Monte Carlo simulation. These uncertainties can affect the description of the distributions in two different ways. They can play the role of a simple normalisation factor or they can distort the distributions.

The dominant sources of uncertainty are due to the Monte Carlo simulation. The uncertainty due to modeling of the underlying event is estimated by comparing results between the main sample generated with the Z2 tune and that with the D6T tune [32]. The effect due to the scale used to match clustered jets to partons (i.e., jet-parton matching) is estimated with dedicated samples generated by varying the nominal matching $p_T$ thresholds by factors of 2 and 1/2. Effects due to the definition of the renormalisation and factorisation scales used in the simulation of the signal are studied with dedicated MC samples with the scales varied by a factor of two.
The uncertainties related to the PDFs used to model the hard scattering of the proton-proton collisions is estimated by varying the PDFs by its uncertainties according to the PDF4LHC prescription [21, 23]. Variations in the relative composition of the simulated samples are studied by varying the contributions of each background with respect to the signal and each other.

Several sources pertain to the modeling of the CMS detector in the Monte Carlo simulations. Important sources of uncertainty are the energy scales of the jets and, to a lesser extent, of the leptons, as they shift the momenta of the reconstructed objects. Similarly, the uncertainty in jet energy resolution has also been considered. The effects of the jet energy scale are taken into account by varying the energy scale of the jets according to its uncertainty [9]. A further source comes from the uncertainties associated to the measurement of the trigger and lepton selection efficiencies.

Since the Monte Carlo samples have been generated with a different pile-up distribution than observed in data, simulated events are reweighted to match the observed pile-up distribution. The uncertainty on this reweighing procedure is taken into account by varying the mean value of the measured distribution by ±0.6 events.

7 Profile Likelihood Ratio method using dilepton events

The technique uses dilepton events to estimate a scale factor for the b-tagging efficiency with a profile likelihood ratio (PLR) method using the 2-dimensional distribution of the jet multiplicity versus the b-tagged jet multiplicity. The uncertainties in the event yield and the shape of the distribution are considered as nuisance parameters in the likelihood function and are then fitted during the minimization procedure. This leads to combined statistical and systematic uncertainties associated with the measurement of the scale factor.

The method is applied to a binned 2D distribution of jet multiplicity versus the b-jet multiplicity. The likelihood function for a given dilepton channel \( j \) (ee, e\( \mu \) or \( \mu \mu \)) and a given bin \( i \) of this distribution (corresponding to \( n \) jets and \( m \) b-tagged jets) is written as [33]:

\[
L_{i,j}(S_{F_b}, N_{i,j}^{\text{obs}}, \{U_k\}) = \text{Poisson}(N_{i,j}^{\text{obs}}, \mu_{i,j}(S_{F_b}, \{U_k\})) \times \prod_k \text{Gauss}(U_k, 0, 1). \tag{1}
\]

where \( N_{i,j}^{\text{obs}} \) is the number of observed events, \( \mu_{i,j} \) the number of expected events, and \( U_k \) the nuisance parameters. The likelihood function for a given channel \( j \) is then the product of the likelihood functions over all the bins of the distribution:

\[
L_j(S_{F_b}, N_{i,j}^{\text{obs}}, \{U_k\}) = \prod_i L_{i,j}(S_{F_b}, N_{i,j}^{\text{obs}}, \{U_k\}). \tag{2}
\]

Since the decay channels are statistically independent, the overall likelihood function is then simply the product of the individual channel likelihoods:

\[
L(S_{F_b}, N_{i,j}^{\text{obs}}, \{U_k\}) = \prod_j L_j. \tag{3}
\]

The expression of the profile likelihood ratio (PLR) is then

\[
\text{PLR}(S_{F_b}) = \frac{L(S_{F_b}, \hat{U}_i)}{L(S_{F_b}, \hat{U}_i)}, \tag{4}
\]

where \( \hat{U}_i \) represents the conditional Maximum Likelihood estimates of \( U_i \) obtained with the scale factor \( S_{F_b} \) fixed while \( \hat{S_{F_b}} \) and \( \hat{U}_i \) are the estimates obtained with \( S_{F_b} \) free.
The distribution of $-2 \log(PLR(SF_b))$ is asymptotically distributed as a $\chi^2$ distribution with one degree of freedom (Wilk's theorem) which allows to establish 68% confidence intervals. A PLR curve is obtained by scanning the values of $SF_b$ in a given range. The curve is then fitted with a third degree polynomial function and used to derive the interval error at $\pm 1\sigma$ by changing the minimum negative log-likelihood by 0.5, in order to get the Gaussian equivalence at 68% C.L. These errors are the combination of the statistical uncertainty and the systematic uncertainties considered as nuisance parameters. All the nuisance parameters are common to the three channels except the data-driven estimation of the backgrounds for W+jets and Z+jets. The Z+jets background is estimated from data as described in Section 4. The small W+jets background is estimated on data using the matrix method [34].

The expected number of b-tagged jets in events with $n$ jets of a given dilepton final state, $\mu_{ij}$ in Eq.(1), is derived from pre-tagged simulated events with $n$ jets by applying per-jet b tagging efficiencies, considering all jet tagging combinations. These efficiencies are derived as a function of $p_T$ and $\eta$, using simulated $t\bar{t}$ events for b-jets and data samples dominated by light flavour jets for the light flavour jets. A constant scale factor $SF_b$ is applied to the b- and c-jet efficiencies to model the b-tag efficiency in data. The value of $SF_b$ is then extracted by minimizing the PLR as described above. A closure test is performed on simulated signal events to check that for a unit scale factor, the b-tagged jet multiplicity distribution obtained with the reweighting procedure is the same as the one obtained directly from Monte Carlo using a requirement on the b-tagging discriminant.

## 7.1 Systematic uncertainties

Several uncertainties are considered as nuisance parameters in the likelihood function and are then fitted during the minimization procedure. These are the uncertainties on the energy scale of the jets and the leptons, the expected number of events of the different contributions, and the uncertainty on the light jet scale factor.

Further contributions to the systematic uncertainties are estimated outside the PLR procedure. The expected input distributions to the PLR method are re-derived using MC samples with varied parameters and the b-tagging scale factors are re-measured. The relative differences of $SF_b$ with respect to the nominal values are taken as systematic uncertainties, and added in quadrature to the total uncertainty from the fit. These uncertainties include the uncertainties on the jet-parton matching scale, the parton-shower/Matrix-Element threshold and the top mass. The factorization scale is the dominant systematic, as it affects the jet multiplicity distribution, with an uncertainty of approximately 0.017 for the CSVL operating point.

The second largest contribution is from the uncertainty on the $t\bar{t}$ event yield, which is estimated to be 20%. It includes the uncertainties on the $t\bar{t}$ cross section, the trigger and lepton selection efficiencies and the branching ratio of the decays of the W bosons. This results in an uncertainty of 0.0136 for the CSVL operating point. Further, the statistical uncertainty on the b-tagging efficiency in the simulation was found to range between 0.4% and 1.6% depending on the operating point considered. A 1.6% systematic uncertainty was therefore chosen for all the operating points.

Finally, to account for a possible uncertainty coming from the fitting algorithm itself, an additional systematic is estimate by comparing the Minuit [35] framework with the Theta framework [36]. This is taken as a 1% relative uncertainty.
8 The Flavor Tag Consistency Method

The Flavor Tag Consistency method (PtCM) requires consistency between the observed and expected number of tags in the events to study the performance of the heavy flavor algorithms.

In a sample of t#bar{t} pair candidates in the lepton+jets channel, the expected number of events with \( n \) b-tagged jets \( \langle N_n \rangle \) can be written as

\[
\langle N_n \rangle = L \cdot \sigma_{t#bar{t}} \cdot \epsilon \cdot \sum_{i,j,k} \sum_{i',j',k'=n}^{i+j+k \leq k} [C_i^b \epsilon_b^j (1 - \epsilon_b^j)(i-j') C_j^c \epsilon_c^j (1 - \epsilon_c^j)(j' - j) C_k^l \epsilon_l^j (1 - \epsilon_l^j)(k-k')] ,
\]

where \( L \) is the integrated luminosity, \( \sigma_{t#bar{t}} \) is the t#bar{t} cross section, \( \epsilon \) is the pre-tagging selection efficiency, \( C^b \) is the binomial coefficient and \( \epsilon_b, \epsilon_c \) and \( \epsilon_l \) are the b-, c- and light-jet tagging efficiencies. The factors \( F_{ijk} \) are the fraction of events with \( i \) b jets, \( j \) c jets and \( k \) light jets. They are derived from the t#bar{t} simulation in which the true flavour of the jets is known.

As an example, the \( F_{112} \) term contributes to the expected number of events with 1 b-tagged jet \( \langle N_1 \rangle \) in the following way,

\[
\langle N_1 \rangle \propto F_{112} \times \left( \frac{1 \cdot \epsilon_b (1 - \epsilon_c) (1 - \epsilon_l)}{\text{the b jet}} + \frac{1 \cdot (1 - \epsilon_b) \epsilon_c (1 - \epsilon_l) \epsilon_l}{\text{the c jet}} + \frac{2 \cdot (1 - \epsilon_b) (1 - \epsilon_c) \epsilon_l (1 - \epsilon_l)}{\text{the light jet}} \right) .
\]

To account for the non negligible amount of background, equation 5 is modified to include each background sample:

\[
\langle N_n \rangle = \langle N_n^{t#bar{t}} \rangle + \langle N_n^{\text{background}} \rangle \\
= L \cdot \sigma_{t#bar{t}} \cdot \epsilon \cdot \sum_{i,j,k} \sum_{i',j',k'=n}^{i+j+k \leq k} \left( \cdots \right) \\
+ \frac{\sigma_{\text{background}}}{\sigma_{t#bar{t}}} \cdot \frac{\epsilon_{\text{background}}}{\epsilon_{t#bar{t}}} \cdot \sum_{i,j,k} \sum_{i',j',k'=n}^{i+j+k \leq k} \left( \cdots \right) ,
\]

where \( \left( \cdots \right) \) stands for the expression in square brackets from equation 5.

The tagging efficiencies and the t#bar{t} production cross section are then measured from the b-tag multiplicity distribution by minimizing the log-likelihood function:

\[
\mathcal{L} = -2 \log \prod_n \text{Poisson}(N_n, \langle N_n \rangle) ,
\]

where \( N_n \) is the number of observed events with \( n \) b-tagged jets. The distribution of the number of b-tagged jets observed in data and predicted in the simulation for t#bar{t} and background events is shown in Fig. 1.

In the current implementation the likelihood only uses the b-tagged jet multiplicity in t#bar{t} lepton+jets events with between four to seven reconstructed jets, as it emphasizes the measurement of the heavy flavor b tagging efficiency. The b-tagging efficiencies and t#bar{t} cross section are thus treated as free parameters in the fit, while the c- and light-jet tagging efficiencies are taken from the simulation corrected for the data-to-Monte Carlo scale factors [1].
8.1 Systematic uncertainties

The systematic uncertainties are determined from ensembles of pseudo-experiments. In each of these pseudo-experiments, the number of signal and background events are generated using Poisson statistics, using as mean values the number of expected events in each channel. Events are then randomly chosen in the simulated samples and the b-tag multiplicity distributions are populated according to the simulated jet multiplicity in each event. The measurement is then performed as described above using the factors $F_{ijk}$ from the nominal simulation. The average b-tagging efficiency is compared to the average b-tagging efficiency measured in ensemble tests with the nominal samples and the difference is taken as a systematic uncertainty.

The dominant contribution is the uncertainty on the jet energy scale, with an uncertainty of 0.023 on the scale factor of the CSVL operating point. The second largest uncertainty arises from the uncertainty on the production cross section of the $W+$heavy flavour jets, with an uncertainty of 0.01. The uncertainties due to the factorisation scale and the jet-parton matching are 0.0042 and 0.0036 respectively for the CSVL operating point.

9 The flavour Tag Matching Method

The flavour Tag Matching method (FrMM) requires consistency between the observed and expected number of tags in dilepton events to study the performance of the heavy flavour algorithms.

The expected number of events with $n$ b-tagged jets $\langle N_n \rangle$ is written as

$$\langle N_n \rangle = \sum_{k \text{ jets}=2}^{\text{all jets}} n_k \cdot P_{n,k},$$

where $n_k$ is the observed number of events with $k$ jets, and $P_{n,k}$ is the probability to count $n$ b tags in a $k$-jet event. These probability functions are written in terms of the tagging efficiencies and the expected jet composition.

In order to illustrate explicitly the construction of the probability functions, the exclusive two-
The flavour Tag Matching Method

The jet multiplicity bin is used and the following expression is obtained:

\[
P_{n,2} = \sum_{i \text{ jets}=0}^{2} \alpha_i \cdot P_{n,2,i},
\]  

(10)

where \( P_{n,2,i} \) is the probability that \( k \) b tags are observed in an event with two jets of which \( i \) jets come from \( t\bar{t} \) decays.

The mis-assignment probabilities \( \alpha_i \) denote the probability in the sample that \( i \) jets from the decay of the \( t\bar{t} \) pair have been reconstructed and selected. These are normalised such that \( \sum_i \alpha_i = 1 \). For example, \( \alpha_2 \) is the probability that both b jets from the \( t\bar{t} \) decay have been selected. They take into account both the contribution from the background, which is small in the dilepton channel, and jet mis-assignment. Either or both of the jets from the decays of the two top quarks may not be selected, and jets from initial state radiation (ISR) or final state radiation (FSR), or jets from the proton recoil may enter the selection, further diluting the sample.

As an example, for the case where two tagged jets are found in a two-jet event, the probabilities can be explicitly written as:

\[
P_{2,2,0} = \varepsilon_q^2 \quad \text{if no jets are from } t\bar{t} \text{ decays,}
\]

\[
P_{2,2,1} = 2\varepsilon_b \varepsilon_q \quad \text{if 1 jet is from } t\bar{t} \text{ decays,}
\]

\[
P_{2,2,2} = \varepsilon_b^2 \quad \text{if 2 jets are from } t\bar{t} \text{ decays.}
\]

(11)

The mistag rate \( \varepsilon_q \) is an effective measurement of the probability of tagging light quark, gluon and charmed jets in the dilepton sample. Similar expressions can easily be derived for the other jet multiplicity bins.

The mis-assignment probabilities are determined from the data, and used in the subsequent likelihood of the b-tag multiplicity distribution. In order to estimate the actual fraction of b jets from top quark decays in the selected sample, kinematic properties of the top decay topology are used. The invariant mass of the lepton-jet pairs from a \( t \rightarrow Wb \) decay have a kinematic end-point at \( M_{t,b}^{\text{max}} \equiv \sqrt{m_t^2 - m_W^2} \approx 156 \text{ GeV}/c^2 \). The invariant mass of mis-assigned lepton-jet pairs exhibits a longer tail towards high mass values. The shape of the mis-assigned pairs can be modeled to a good approximation using combinatorics, i.e. mixing lepton-jet pairs from different events or randomly rotating the lepton momentum direction. The fraction of jets from \( t \rightarrow Wb \) decays can thus be measured normalizing the spectrum obtained from the combinatorial model to the number of pairs observed in the tail (i.e. \( M_{t,j} > 180 \text{ GeV}/c^2 \)). This is estimated independently for each dilepton channel and for each jet-multiplicity bin. The procedure is checked and found to be unbiased from MC based pseudo-experiments. Taking into account the expected contribution of \( t\bar{t} \) and single top events to the final sample, the sample composition in terms of events with 2, 1 or 0 correctly reconstructed and selected b jets is estimated.

The b tagging efficiency \( \varepsilon_b \) can then be measured from the b-tag multiplicity distributions by maximizing the likelihood function:

\[
\mathcal{L}(\varepsilon_b, \varepsilon_q, \alpha_i) = \prod_{n=0}^{\text{all jets}} \text{Poisson}(N_n, \langle N_n \rangle),
\]

(12)

where \( N_n \) is the observed number of events with with \( n \) b-tagged jets.
The likelihood only uses the b-tagged jet multiplicity in $t\bar{t}$ dilepton events with two and three reconstructed jets. Gaussian constraints are added for the effective c- and light-jet tagging efficiency $\varepsilon_q$ and the mis-assignment probabilities:

$$L = \prod_{n=0}^{\text{all jets}} \text{Poisson}(N_n, \langle N_n \rangle) \cdot \prod_i \text{Gauss}(\alpha_i, \hat{\alpha}_i, \sigma_{\alpha_i}) \cdot \text{Gauss}(\varepsilon_q, \hat{\varepsilon}_q, \sigma_{\varepsilon_q}).$$  \hspace{1cm} (13)$$

The central value and width of $\varepsilon_q$ are determined from the simulation. For the mis-assignment probabilities $\alpha_i$, the central values are taken from the measurement described above, and the width derived from the uncertainty of the expected contribution of $t\bar{t}$ and single top events to the final sample.

### 9.1 Systematic uncertainties

The systematic uncertainties affect the measurement of the b-tagging probability through their effect on the parameters of the fit, namely the measured mis-assignment probabilities and the mistag rate.

The effect on the measured mis-assignment probabilities is determined from ensembles of pseudo-experiments, where, for each source of uncertainty, the bias on the probabilities is determined. Most sources of uncertainties such as jet energy scale and resolution and pileup have little effect as the method used to derive the mis-assignment probabilities is based on data-driven templates for the lepton-jet invariant mass. Other sources which might affect the contribution from top decays and from radiation jets to the final sample are evaluated using samples where the $Q^2$ and the jet-parton matching scales are varied. In the pseudo-experiments the standard $t\bar{t}$ sample is substituted by each one of these samples and the process is repeated.

This bias computed as described above is then used to shift the measured mis-assignment probabilities and the likelihood fit of the data is repeated with the modified values. The difference with respect to the nominal result is taken as systematic uncertainty. The same procedure is applied to evaluate the mistag rate uncertainty.

The final uncertainty is dominated by factors which tend to increase the contamination of background or alter the jet environment. The main uncertainties are thus the factorisation scale and to a smaller extent the jet-parton matching with uncertainties of 0.023 and 0.014 respectively for the CSVL operating point. The second largest uncertainty arises from the uncertainty on the light jet tagging efficiency, with an uncertainty of 0.015.

### 10 Efficiency measurement from a b-enriched jet sample

In this method, the b-jet efficiency is measured from a b-jet-enriched sample (bSample). The contamination of this sample due to light jets is estimated from data and subtracted.

In order to select the correct jets originating from the decay of the top quarks, a $\chi^2$ is calculated for each jet-parton combination based on the masses of the reconstructed W boson $m_{qq}$ and the hadronically-decaying top quark $m_{bqq}$:

$$\chi^2 = \left(\frac{m_{bqq} - m_t}{\sigma_t}\right)^2 + \left(\frac{m_{qq} - m_W}{\sigma_W}\right)^2,$$

where the mean masses and width are obtained from the $t\bar{t}$ simulation using a Gaussian fit to the mass distributions of the combination with the correct jet to quark assignment. The mean
and width of the reconstructed top quark are $172.5\text{ GeV}/c^2$ and $16.3\text{ GeV}/c^2$ respectively, and the mean and width of the W boson are $82.9\text{ GeV}/c^2$ and $9.5\text{ GeV}/c^2$ respectively. Using the four leading jets with a transverse momentum above $30\text{ GeV}/c$, there are 12 combinations to pair the four reconstructed jets with the quarks from $t\bar{t}$ decay, and the combination with the lowest $\chi^2$ is selected to represent the event topology. Furthermore, the event is rejected if the lowest $\chi^2$ is above 90.

A generic $b$-candidate sample is thus constructed by taking the jet assigned to the leptonic $b$ quark. This sample is further sub-divided into a $b$-enriched and a $b$-depleted subsample by using the invariant mass of the leptonic $b$ jet and the reconstructed muon (called the jet-muon mass, $m_{\mu j}$). The distribution of this variable is shown in Fig. 2. For the $b$-enriched subsample, the jet-muon mass is required to be in the range $80\text{ GeV}/c^2 < m_{\mu j} < 150\text{ GeV}/c^2$, and for the $b$-depleted subsample in the region $150\text{ GeV}/c^2 < m_{\mu j} < 250\text{ GeV}/c^2$. Based on the simulation, the purity of the two subsamples is 44.7% and 15.9% respectively.

![Figure 2: Distribution of the jet-muon mass for all relevant processes. The data is compared to simulation. The simulated distribution is normalized to an integrated luminosity of 2.18 fb$^{-1}$.](image)

The distribution of the discriminators of the taggers for true $b$ jets, $\hat{\Delta}_b^{enr}$, is then obtained by subtracting the discriminator distribution of the $b$-depleted subsample, $\Delta_b^{depl}$ from the discriminator distribution of the $b$-enriched subsample, $\Delta_b^{enr}$:

$$\hat{\Delta}_b^{enr} = \Delta_b^{enr} - F \times \Delta_b^{depl}. \quad (15)$$

The factor $F$ represents the ratio of the number of non-$b$ jets in the $b$-enriched and $b$-depleted subsamples. It is measured from a background dominated sample composed mainly of light quark jets. This sample is obtained by using the jets attributed to the decay of the W boson and ensuring that they both fail the $b$-tagging requirements of the TCHEM operating point. Both jets are then used to construct a jet-muon mass distribution, and the same sub-samples are defined as in the signal sample. The purity of light quark jets is 92.0% in the region $80\text{ GeV}/c^2 < m_{\mu j} < 150\text{ GeV}/c^2$ and 95.0% in the region $150\text{ GeV}/c^2 < m_{\mu j} < 250\text{ GeV}/c^2$. To match the shape
of the jet-muon mass distribution of this background sample to the jet-muon mass distribution of the signal sample, the jets in the background sample are reweighted according to the \((\eta, p_T)\) of the signal sample. After this reweighting, the shape of the jet-muon mass distribution agrees well between the two samples, and the factor \(F\) is taken as the ratio of the number of events in the \(80\,\text{GeV}/c^2 < m_{\mu j} < 150\,\text{GeV}/c^2\) and the \(150\,\text{GeV}/c^2 < m_{\mu j} < 250\,\text{GeV}/c^2\) regions in the background sample, and is found to be \(1.164 \pm 0.018\).

Finally, a small correlation between the jet-muon mass and the discriminators has to be corrected for. This correlation is attributed to the correlation between the transverse momentum of the jet and the jet-muon mass and distorts the distribution of the discriminants of the b-tagging algorithms in the b-depleted subsample with respect to the distribution of the non-b jets in the b-enhanced subsample. This effect is corrected by reweighting the jets in the b-depleted subsample according to the transverse momentum distribution of the jets in the b-enhanced subsample.

### 10.1 Systematic uncertainties

The systematic uncertainties are determined as the absolute difference on the results for the b-tagging efficiency and the scale factors between our nominal simulation sample and the sample with adapted parameters. Since this analysis is purely data-driven, the uncertainties due to the Monte Carlo simulation have no effect on the final result. The main source of systematic uncertainty is due to the bias on the method, which is determined with a relatively large uncertainty. For the CSVL operating point, the uncertainty is 0.031. This uncertainty is thus taken as a systematic. The uncertainty on the method bias is driven by the statistics of the smallest simulated sample in the set, describing the W+jets process. The bias itself is inherent to the method and does not depend on the details of the simulation.

The jet energy scale and resolution have a small contribution by the change in the mean masses and widths used for the \(\chi^2\). For the CSVL, the uncertainties on the scale factor are 0.014 and 0.022 respectively. Finally, a small uncertainty of 0.005 is due to the choice of the boundaries of the b-depleted region. The high tail of the jet-muon mass distribution is composed mainly of background events and wrongly combined jets hence not reflecting the kinematics of the signal events. The effect of the upper limit of the region is assessed by varying the boundary between 200 and 300\,\text{GeV}/c^2.

### 11 Results

The statistical properties of the four methods have been studied using ensembles of pseudo-experiments based on the expected numbers of signal and background events. The characteristic distributions of the estimated values and their errors show that the methods are unbiased, with pull distributions with mean values close to zero and standard deviations close to one.

The scale factors \(S_{F_b} = \epsilon_{eb}^{\text{meas}}/\epsilon_{eb}^{\text{MC}}\) as measured for the different algorithms using 2.18\,fb\(^{-1}\) of data are shown in Table 3. The measurements are compared to the scale factors measured from dijet events enriched with jets from semimuonic b-hadron decays [1]. The measured efficiencies and scale factors for the CSV algorithm are shown in Figures 3 to 5.

A combined scale factor to be used in analyses is derived as a weighted mean of the scale factors measured by the PLR and FtCM methods. The two methods are chosen because each has the smallest uncertainty among the analyses in its respective decay channel. By choosing one analysis in the dilepton channel and one in the lepton+jets channel, there is no statistical correlation between the two measurements as the samples are mutually exclusive. The uncertainty
Table 3: Summary of the $S_{b}$ as measured by the Profile Likelihood Ratio method (PLR), Flavour Tag Consistency method (FtCM), Flavour Tag Matching method (FtMM), $b$Sample, and the weighted mean (WM) compared to measurements using dijet events [1].

<table>
<thead>
<tr>
<th>Point</th>
<th>Dijet</th>
<th>PLR</th>
<th>FtCM</th>
<th>FtMM</th>
<th>$b$Sample</th>
<th>WM</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCHEM</td>
<td>0.95±0.09</td>
<td>0.96±0.03</td>
<td>0.94±0.04</td>
<td>0.95±0.03</td>
<td>0.92±0.05</td>
<td>0.95±0.04</td>
</tr>
<tr>
<td>TCHEM</td>
<td>0.94±0.09</td>
<td>0.96±0.03</td>
<td>0.97±0.04</td>
<td>0.96±0.03</td>
<td>0.93±0.06</td>
<td>0.96±0.04</td>
</tr>
<tr>
<td>TCHET</td>
<td>-</td>
<td>0.93±0.03</td>
<td>0.99±0.05</td>
<td>1.01±0.04</td>
<td>0.95±0.08</td>
<td>0.95±0.04</td>
</tr>
<tr>
<td>TCHPL</td>
<td>-</td>
<td>0.95±0.03</td>
<td>0.90±0.04</td>
<td>0.95±0.03</td>
<td>0.92±0.06</td>
<td>0.93±0.04</td>
</tr>
<tr>
<td>TCHPM</td>
<td>0.91±0.09</td>
<td>0.94±0.03</td>
<td>0.93±0.04</td>
<td>0.95±0.04</td>
<td>0.92±0.06</td>
<td>0.93±0.04</td>
</tr>
<tr>
<td>TCHPT</td>
<td>0.88±0.09</td>
<td>0.93±0.03</td>
<td>0.92±0.05</td>
<td>0.94±0.04</td>
<td>0.91±0.07</td>
<td>0.93±0.04</td>
</tr>
<tr>
<td>SSVHEM</td>
<td>0.95±0.09</td>
<td>0.95±0.03</td>
<td>0.98±0.04</td>
<td>0.98±0.04</td>
<td>0.95±0.07</td>
<td>0.96±0.04</td>
</tr>
<tr>
<td>SSVHET</td>
<td>-</td>
<td>0.94±0.03</td>
<td>0.96±0.05</td>
<td>0.99±0.04</td>
<td>0.96±0.08</td>
<td>0.95±0.04</td>
</tr>
<tr>
<td>SSVHPT</td>
<td>0.90±0.09</td>
<td>0.95±0.03</td>
<td>0.95±0.04</td>
<td>0.96±0.04</td>
<td>0.89±0.07</td>
<td>0.95±0.04</td>
</tr>
<tr>
<td>CSVL</td>
<td>1.00±0.10</td>
<td>1.00±0.03</td>
<td>1.03±0.03</td>
<td>0.99±0.04</td>
<td>0.97±0.05</td>
<td>1.01±0.04</td>
</tr>
<tr>
<td>CSVM</td>
<td>0.96±0.10</td>
<td>0.97±0.03</td>
<td>0.98±0.04</td>
<td>0.97±0.04</td>
<td>0.95±0.06</td>
<td>0.97±0.04</td>
</tr>
<tr>
<td>CSVT</td>
<td>0.94±0.09</td>
<td>0.95±0.03</td>
<td>0.97±0.05</td>
<td>0.96±0.03</td>
<td>0.95±0.06</td>
<td>0.96±0.04</td>
</tr>
</tbody>
</table>

of the resulting scale factor is taken as 0.04 for all points. It is based on the typical systematic uncertainty of the measurements with the PLR method, 0.03, and the average standard deviation of the four measurements, 0.02. For each operating point, the standard deviation of the four measurements is between 0.01 and 0.03.

Analyses which use the $b$-tagging discriminators in multi-variate analysis methods need a continuous function for the scale factors. For this function, the linear function fitted on the distribution of the scale factors measured with the FtCM (Fig. 3) is offset vertically to match the weighted mean of the medium operating point. This is illustrated in Figure 6.

12 Conclusion

Four methods to measure the $b$-tagging efficiency and scale factors on $t\bar{t}$ events are presented. The measurements are done on a sample corresponding to an integrated luminosity of $2.18±0.11$ fb$^{-1}$ collected by the CMS experiment in 2011. Good agreement is found between data and simulation, and the scale factors are measured within 10% of unity, with uncertainties of a few percent. The different methods yield scale factors which agree within uncertainties, and, to provide a single scale factor for each working point, a weighted mean between the results with the smallest uncertainties is proposed.

Acknowledgements

We wish to congratulate our colleagues in the CERN accelerator departments for the excellent performance of the LHC machine. We thank the technical and administrative staff at CERN and other CMS institutes, and acknowledge support from: FMSR (Austria); FNRS and FWO (Belgium); CNPq, CAPES, FAPERJ, and FAPESP (Brazil); MES (Bulgaria); CERN; CAS, MoST, and NSFC (China); CONICYT grant (Chile); COLCIENCIAS (Colombia); MSES (Croatia); RPF (Cyprus); Academy of Sciences and NICPB (Estonia); Academy of Finland, MEC, and HIP (Finland); CEA and CNRS/IN2P3 (France); BMBF, DFG, and HGF (Germany); GSRT (Greece); OTKA and NKTH (Hungary); DAE and DST (India); IPM (Iran); SFI (Ireland); INFN (Italy); NRF and WCU (Korea); LAS (Lithuania); CINVESTAV, CONACYT, SEP, and UASLP-FAI (Mexico); MSI (New Zealand); PAEC (Pak-
Figure 3: Measured b-tagging efficiency as a function of the flavour discriminator threshold for the CSV algorithm, measured with the Flavour Tag Consistency method. The absolute b-tagging efficiency measured from data and predicted from simulation are shown in the upper histogram. The scale factors are shown in the lower histogram. The values of $SF_b$ are shown together with the combined statistical and systematic uncertainty. The lines show the functional form of $SF_b$ and the uncertainty band. The arrows indicate the standard operating points.

Figure 4: Measured b-tagging efficiency as a function of the discriminator threshold for the CSV algorithm, measured with the Flavour Tag Matching method. The absolute b-tagging efficiency measured from data and predicted from simulation are shown in the upper histogram. The scale factors are shown in the lower histogram. The total uncertainties are represented by a blue band for the scale factor. The arrows indicate the standard operating points.
Figure 5: Measured b-tagging efficiency as a function of the discriminator threshold for the CSV
algorithm, measured with the bSample method. The absolute b-tagging efficiency measured
from data and predicted from simulation are shown in the upper histogram. The scale factors
are shown in the lower histogram, where the blue band represents the total uncertainties. The
arrows indicate the standard operating points.
Figure 6: Scale factors measured with the PLR and Flavour Tag Consistency method and weighted mean as a function of the discriminator threshold for the CSV algorithm. The black function is derived from a fit to the values measured with the FtCM, and the red function labeled “Final function” corresponds to the same function fixed to the weighted mean of the medium operating point.
istan); SCSR (Poland); FCT (Portugal); JINR (Armenia, Belarus, Georgia, Ukraine, Uzbekistan); MST, MAE and RFBR (Russia); MSTD (Serbia); MICINN and CPAN (Spain); Swiss Funding Agencies (Switzerland); NSC (Taipei); TUBITAK and TAEK (Turkey); STFC (United Kingdom); DOE and NSF (USA).

Individuals have received support from the Marie-Curie programme and the European Research Council (European Union); the Leventis Foundation; the A. P. Sloan Foundation; the Alexander von Humboldt Foundation; the Belgian Federal Science Policy Office; the Fonds pour la Formation à la Recherche dans l’Industrie et dans l’Agriculture (FRIA-Belgium); the Agentschap voor Innovatie door Wetenschap en Technologie (IWT-Belgium); and the Council of Science and Industrial Research, India.

References


