CMS Physics Analysis Summary

Contact: cms-pog-conveners-btag@cern.ch 2013/08/15

Performance of b tagging at $\sqrt{s} = 8$ TeV in multijet, $t\bar{t}$ and boosted topology events

The CMS Collaboration

Abstract

The identification of jets originating from b quarks is an important component in analyses searching for new physics and for measurements of standard model processes. Particularly challenging are those involving top-quark or Higgs-boson production with decays in the boosted regime. Studies of b-quark jet algorithms and their performances are presented, using multijet events and $t\bar{t}$ events recorded in proton-proton collision data at $\sqrt{s} = 8$ TeV with the CMS detector in the year 2012. The b-tagging efficiency and misidentification probability are determined in data and in simulation over a wide range of jet transverse momenta from 20 GeV/c to more than 800 GeV/c. For the first time in CMS, dedicated studies of b-quark identification in boosted topologies are presented.

1 Introduction

Identification of b-quark jets (b tagging) is used in many physics analyses, from precision standard model and Higgs boson measurements to new physics searches. As the excluded mass regions of new physics continue to increase, the b quarks end up in boosted topologies with fully merged jets from top-quark or Higgs boson decays. These events are usually reconstructed with jet substructure methods such as the "top-tagging" [1] and "W/Z-tagging" [2] algorithms. Adding b-quark jet identification significantly improves the sensitivity of these methods, but it is challenging to measure the b-tagging efficiency in such environments.

The CMS detector is presented in [3]. A description of the b-tagging algorithms and of the methods to measure their performance in pp collisions at $\sqrt{s} = 7\,\text{TeV}$ taken in 2011 is detailed in [4]. The present document completes and extends these studies to the pp collision data recorded in 2012 at $\sqrt{s} = 8\,\text{TeV}$. It is organised as follows.

The data and simulation samples are described in Section 2. The reconstructed physics objects and the definition of the b-tagging algorithms are summarised in Section 3, where the distributions in data and simulation are compared for a number of b-tagging variables. For the analyses at $\sqrt{s} = 8\,\text{TeV}$, only the *Track Counting High Purity* (TCHP), *Jet Probability* (JP) and *Combined Secondary Vertex* (CSV) algorithms are used, as they are the most efficient (CSV, JP) or matched to a trigger condition with similar algorithms (Combined Secondary Vertex, Track Counting) and loose selection criteria.

Section 4 presents the measurements of the misidentification probability of light-parton (u, d, s, g) jets and of the b-tagging efficiency. We employ similar techniques as those presented in [4] to measure the data/simulation ratio of b-tagging efficiency measurements as a function of the b-quark jet transverse momentum in multijet and $\mathrm{t\bar{t}}$ samples. In addition, we look at a sample enriched in b jets by selecting events with a reconstructed J/ψ decaying into a pair of muons.

For the first time at CMS, these validation and performance studies have been also performed in events with boosted topology, as detailed in Section 5.

2 Data and simulation

Samples of inclusive multijet events for the measurement of efficiencies and misidentification probabilities were collected using single jet triggers with p_{T} thresholds between 40 to 320 GeV/c. For efficiency measurements, dedicated triggers were also used to enrich the data sample with jets from semimuonic b-hadron decays. These triggers required the presence of at least two jets with p_{T} thresholds ranging from 20 to 110 GeV/c, or at least one jet with p_{T} larger than 300 GeV/c. One of these jets was required to include a muon with $p_{\mathrm{T}} > 5\,\text{GeV}/c$ within a cone of $\Delta R = 0.4$ around the jet axis, where ΔR is defined as $\sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$. Here ϕ is the azimuthal angle in the (x, y) plane perpendicular to the beam axis and the pseudorapidity is defined as $\eta \equiv -\ln[\tan(\theta/2)]$ where θ is the polar angle relative to the (z) beam axis. All triggers were prescaled except those with the highest-p_{T} thresholds. Data for the analysis of $\mathrm{t\bar{t}}$ events were collected with unprescaled single- (e or μ) and double-lepton (ee or eμ or $\mu\mu$) triggers. For unprescaled triggers, both the multijet and $\mathrm{t\bar{t}}$ analyses used datasets with integrated luminosities of $19.8\,\text{fb}^{-1}$.

Monte Carlo (MC) simulated samples of multijet events were generated with PYTHIA 6.426 [5] using the Z2 tune [6]. In Section 4, the identification efficiencies in data will be compared to the simulation based on this PYTHIA generator and tuning. For b-jet tagging efficiency studies, dedicated multijet samples have been produced with the explicit requirement of a muon (with

$p_T > 5\,\text{GeV}/c$) in the final state.

In the simulation, a reconstructed jet is matched with a generated parton if the direction of the parton lies within a cone of radius $\Delta R = 0.3$ around the jet axis. The jet is then assigned the flavour of the parton. Should more than one parton be matched to a given jet, the flavour assigned is that of the heaviest parton. The b flavour is given priority over the c flavour, which in turn is given priority over light partons. According to this definition jets originating from gluon splitting to $b\bar{b}$ are classified as b jets.

Events involving $t\bar{t}$ production were simulated using MADGRAPH 5.1.3.30 [7], which was interfaced to PYTHIA for parton showering. A top quark mass of $m_t = 172.5\,\text{GeV}/c^2$ was assumed. Events resulting from single top-quark production were generated with POWHEG 3 [8]. All simulated event samples were normalised to the theoretical cross section of each process, as calculated with FEWZ [9] for W and Z production, with MCFM [10] for $t\bar{t}$ production, and taken from [11] for singly produced top quarks.

All generated events were passed through the GEANT4-based detector simulation [12]. The simulation also includes the effect of pileup, which refers to additional pp collisions occurring during the same bunch crossing or during the bunch crossing preceding or following the primary crossing. The simulated events were reweighted to match the pileup distribution observed in data.

For the validation of the agreement between data and simulation in Section 3, two different event samples have been considered:

- an inclusive multijet sample selected by requiring at least one jet with p_T of 60-500 GeV/c. The content of this sample is dominated by light-flavour quark and gluon jets;
- a sample selected by requiring an electron, a muon and at least two jets with $p_T > 30\,\text{GeV}/c$ ($e\mu+ \geq 2$ jets), this sample is dominated by top pair production, and so highly enriched in b-quark jets.

As these samples differ in the flavour composition and in the origin of heavy and light-parton jets they provide a good validation of the agreement between data and simulation.

3 b-tagging algorithms and data/MC comparisons

Jets are clustered from objects reconstructed by the particle-flow method [13, 14]. This algorithm combines information from all subdetectors to create a consistent set of reconstructed particles for each event. The particles are then clustered into jets using the anti-k_T clustering algorithm [15] with a distance parameter of 0.5. The raw jet energies are corrected to obtain a uniform response in η and an absolute calibration in p_T [16]. These jets are selected in the range $p_T > 20\,\text{GeV}/c$ and $|\eta| < 2.4$.

Tracks are reconstructed using an iterative tracking procedure [17]. The primary vertex (PV) 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. This is done by searching for clusters of tracks with similar z coordinates at their point of closest approach to the beam-line using a deterministic annealing method [18]. The one with the highest $\sum(p_T^{\text{track}})^2$ is selected to define the primary interaction vertex, whose position is determined from an adaptive vertex fit [19].

Only tracks fulfilling the following criteria are used for b-tagging:

- angular distance between track and jet axis $\Delta R < 0.3$ (except for the TCHP b-tagging algorithm, which uses $\Delta R < 0.5$)
- number of pixel hits ≥ 2 and number of tracker hits (including pixel) ≥ 8;
- distance smaller than 0.2 cm (17 cm) in the transverse plane (along the beam axis) between the track and the primary vertex at the point of closest approach of the trajectory to the PV in the transverse plane;
- transverse momentum $p_{\mathrm{T}} > 1\,\mathrm{GeV}/c$;
- normalised $\chi^2 < 5$;
- distance to jet axis < 0.07 cm, defined as the spatial distance between the trajectory and the jet axis at their point of closest approach, where the jet axis is reconstructed with respect to the primary vertex;
- decay length < 5 cm, defined as the spatial distance between the PV and the point of closest approach between the track trajectory and the jet axis.

A subset of these quantities are shown in Fig. 1 for the inclusive multijet sample. Throughout the whole Section, distributions from simulated events are normalized to data after applying the same event reconstruction and selection. The stacked distributions show the contributions from jets associated to b, c and light partons. For the b-quark distribution, the contribution from gluon splitting is displayed separately. The relative contributions from the different flavours are assumed to be correctly predicted by the simulation. The uncertainties shown in this and all following figures are statistical unless otherwise stated. The distributions show satisfactory agreement with the expectations from simulation. The distribution of the multiplicity of selected charged particles is shown in Fig. 2 for the two different event samples, introduced in Section 2. The track multiplicity, which is smaller in data than in simulation, and the lower part of the momentum spectrum (in Fig. 1) are particularly sensitive to the modelling of the particle multiplicity and kinematics by the Monte Carlo generator, as are other distributions such as the number of hits in the innermost pixel layers. Detector effects that are not modelled by the simulation, such as the dynamic readout inefficiency in the pixel system, can also contribute to the remaining discrepancies.

A very important quantity for b-tagging is the impact parameter (IP), which is defined as the distance from the PV to the track at their point of closest approach in space. The IP is given the same sign as the scalar product of the jet axis direction with the vector pointing from the primary vertex to this point of closest approach. The distributions of IP values and their uncertainties are shown in Fig. 3 for the inclusive multijet and $\mathrm{t\bar{t}}$ enriched samples. Good agreement with simulation is observed.

Secondary vertices (SV) within jets are reconstructed using the adaptive vertex fitter [19]. The resulting list of vertices is then subject to a cleaning procedure, rejecting SV candidates that share 65% or more of their tracks with the PV. The distance in space from the primary vertex to the secondary vertex, the flight distance, is used to identify long-lived particles such as b hadrons. The flight direction of the candidate SV has to be within a cone $\Delta R < 0.5$ around the jet direction. SV candidates with a radial distance from the primary vertex greater than 2.5 cm whose associated tracks form an invariant mass compatible with the K^0 mass are rejected.

Typical properties of secondary vertices, such as the invariant mass and total transverse momentum of the associated tracks, and the significance of the flight distance (flight distance divided by its uncertainty) are displayed in Figs. 4 and 5 for two different event samples. It is visible that the vertex quantities are very well described by the simulation.

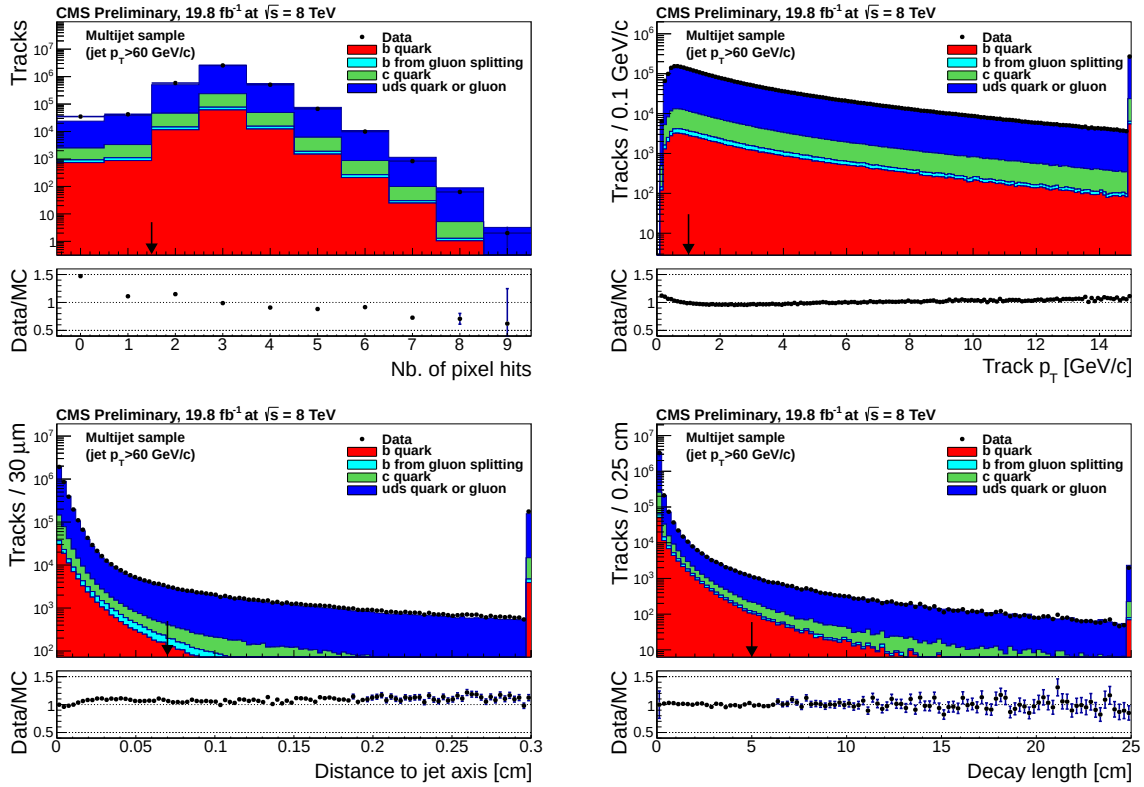


Figure 1: Track properties after basic selection (except for the variable plotted) in the inclusive multijet sample: (top left) number of hits in the pixel system, (top right) transverse momentum of tracks, (bottom left) distance to the jet axis, (bottom right) decay length. The arrow indicates the selection applied on the displayed variable. The distributions from simulation have been normalised to match the counts in data. The filled circles correspond to data. The stacked, coloured histograms indicate the contributions of different components from simulated multijet ("QCD") samples. Simulated events involving gluon splitting to b quarks ("b from gluon splitting") are indicated separately from the other b production processes ("b quark"). In each histogram, the last bin includes all events from the overflow.

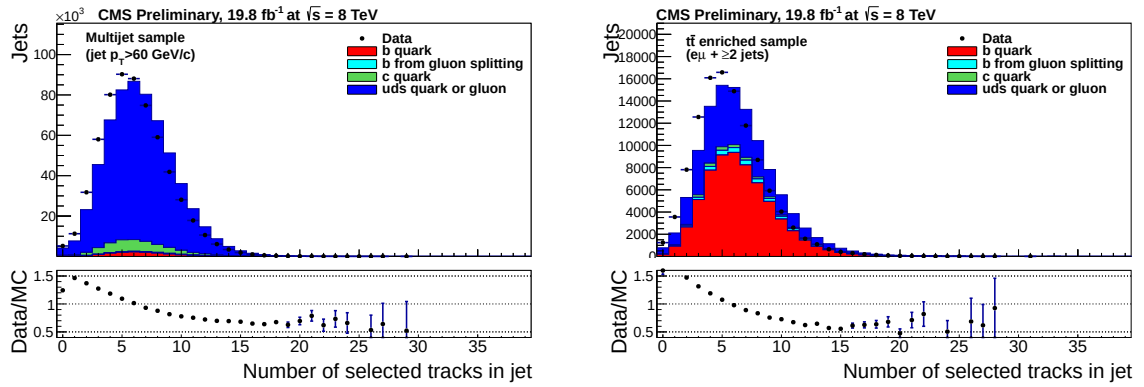


Figure 2: Multiplicity of selected tracks in jets for the (left) inclusive multijet and (right) $t\bar{t}$ enriched samples. Symbols are the same as in Fig. 1.

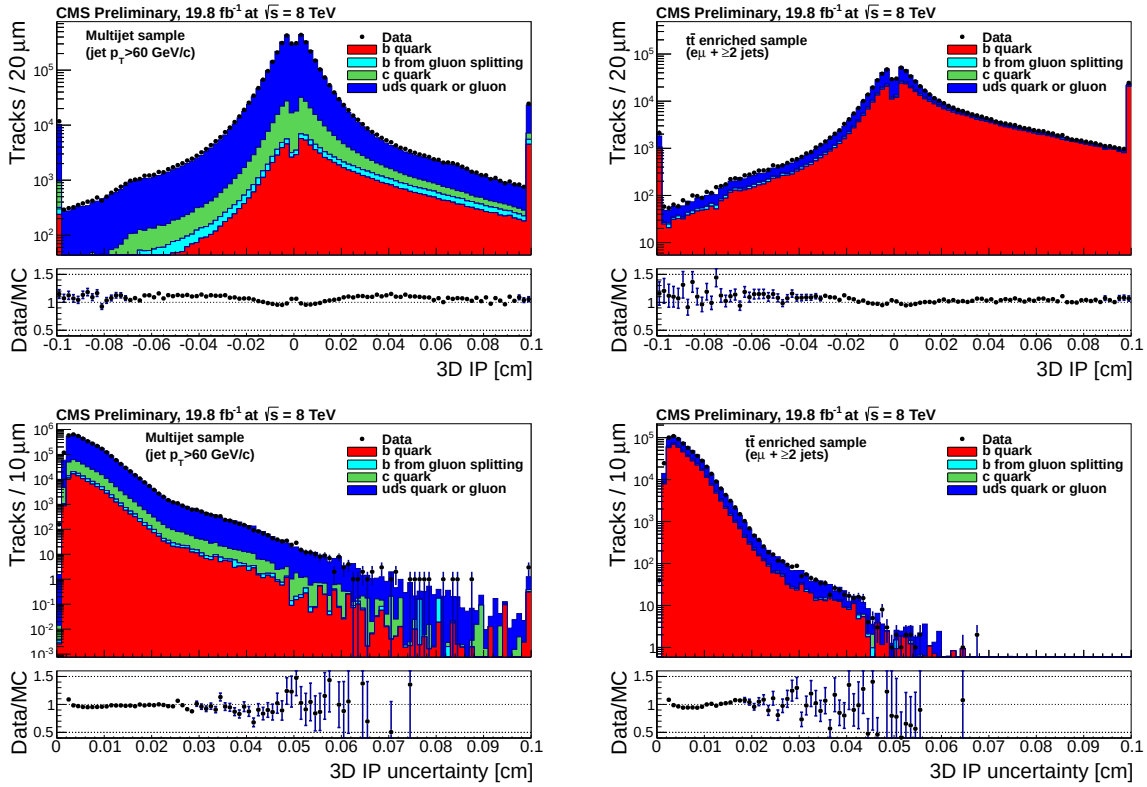


Figure 3: Distributions of (top) the 3D impact parameter and (bottom) the uncertainty on the 3D impact parameter for all selected tracks. The inclusive multijet, and $\mathrm{t\bar{t}}$ enriched samples are shown in the left and right panels, respectively. Symbols are the same as in Fig. 1. Underflow and overflow are added to the first and last bins, respectively.

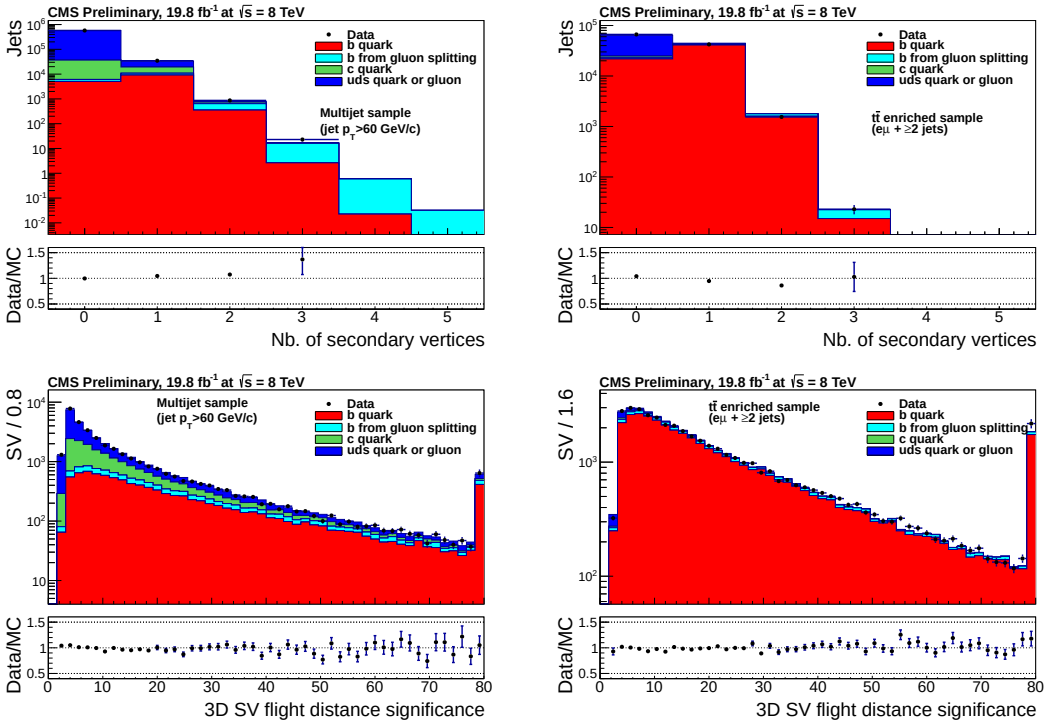


Figure 4: Properties of reconstructed decay vertices: (top) the multiplicity of reconstructed vertices within jets, (bottom) the significance of the 3D flight distance. The inclusive multijet and $\mathrm{t\bar{t}}$ enriched samples are shown in the left and right panels, respectively. Symbols and overflow are defined as in Fig. 1.

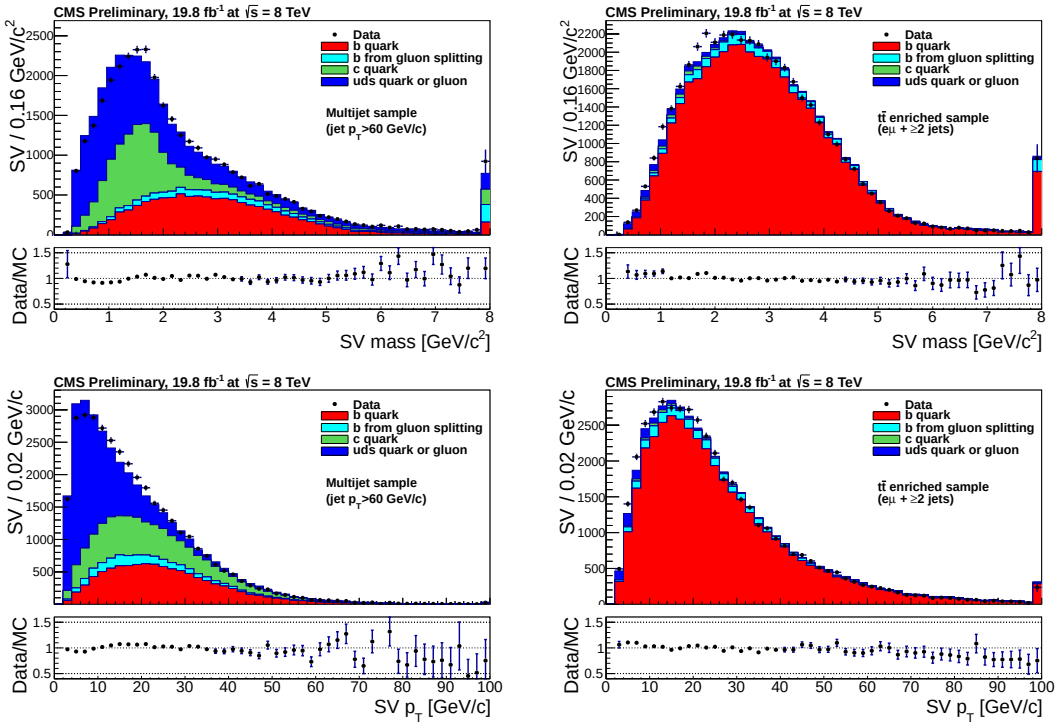


Figure 5: Same as Fig. 4 for (top) the "SV mass" defined as the invariant mass of tracks associated to the secondary vertex, (bottom) the "SV p_{T}" defined as the sum of transverse momenta of tracks associated to the secondary vertex.

Each b-tagging algorithm yields a single discriminator value for each jet. The following algorithms [4] were used for the 2012 data analyses:

- *Track Counting High Purity* (TCHP): the discriminator value is defined as the the impact parameter significance (IP/σ_{IP}, where σ_{IP} is the measurement uncertainty on the impact parameter) of the track with the third highest impact parameter significance;
- *Jet Probability* (JP): the jet is assigned a likelihood that all associated tracks come from the PV; the probability distribution of individual tracks is computed using the impact parameter significance of tracks with negative IP values, which are mostly produced in light-parton jets; this calibration is performed independently in data and simulation;
- *Combined Secondary Vertex* (CSV): secondary vertices and track-based lifetime information are used to build a likelihood-based discriminator to distinguish between jets from b-quarks and those from charm or light quarks and gluons.

The minimum thresholds on these discriminators define loose ("L"), medium ("M"), and tight ("T") operating points with a misidentification probability for light-parton jets of close to 10%, 1%, and 0.1%, respectively, with an average jet p_T of about 80 GeV/c. Throughout this paper, the tagging criteria will be labeled with the letter characterizing the operating point appended to the acronym of the algorithm in use. For instance, JPL will indicate the loose operating point for the JP tagger. The distributions of the TCHP, JP and CSV discriminator values are shown in Fig. 6 for two different event samples.

The quantities relevant for b-tagging exhibit a good agreement in data and simulations, with deviation generally within 20%, while larger discrepancies are observed only in the tails of the distributions. In order to take into account these differences in the analyses, the performance of the b-tagging algorithms are calibrated in data, as discussed in the next Section.

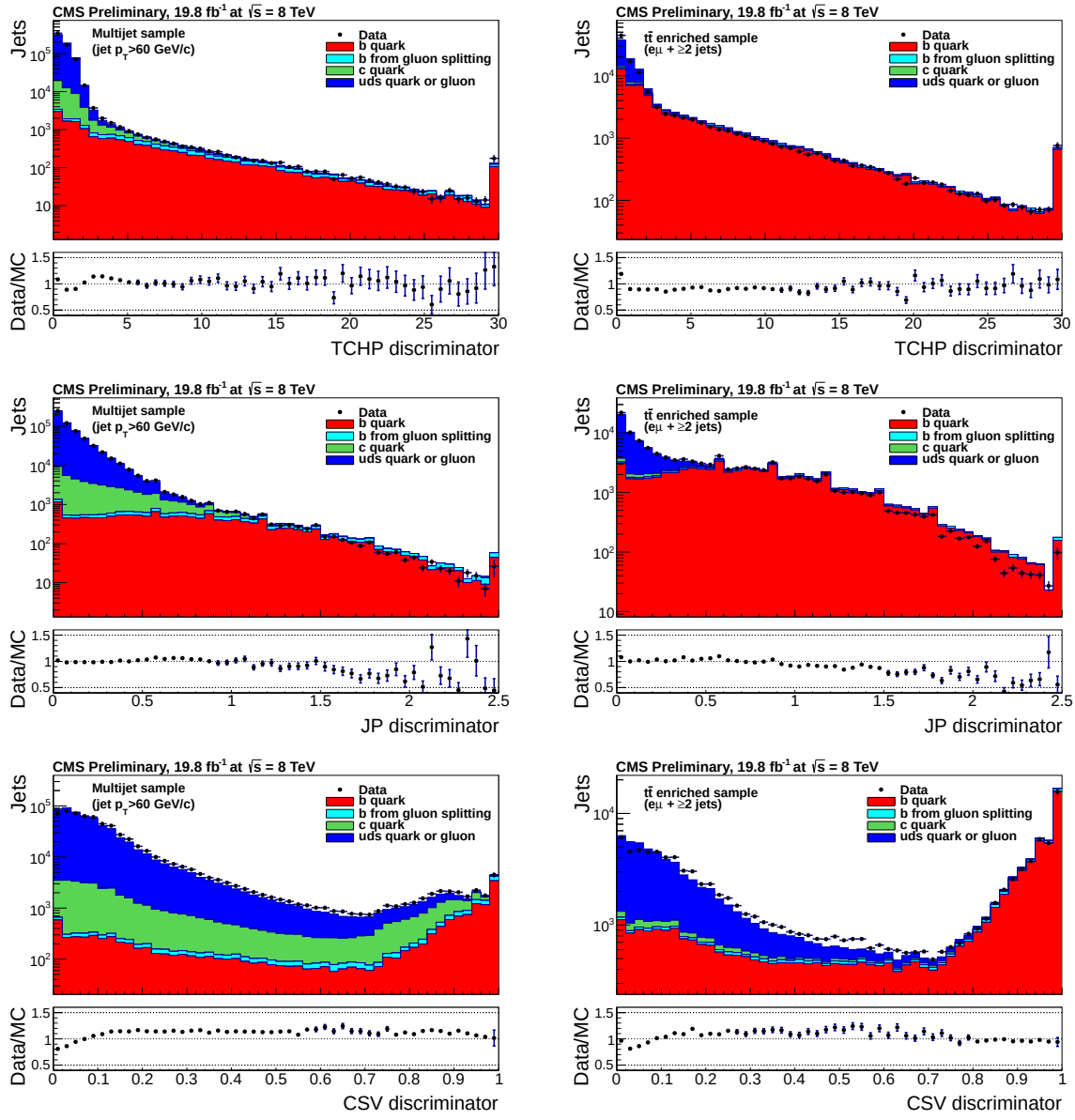


Figure 6: Discriminator values for the (top) TCHP, (middle) JP and (bottom) CSV algorithms. The inclusive multijet, and $t\bar{t}$ enriched samples are shown in the left and right panels, respectively. The small discontinuities in the JP distributions are due to the single track probabilities, which are required to be greater than 0.5%. Symbols are the same as in Fig. 1. Overflows are added to the last bins.

4 Performance measurements in multijet and $t\bar{t}$ events

Several methods have been developed in CMS to measure the b-jet tagging efficiency, ε_b, and the misidentification probability of light-parton jets [4]. Here only the improvements or changes introduced to study the 8 TeV data are detailed. Three sets of event samples are used in order to evaluate the b-tagging performance: inclusive jet samples triggered by the presence of a high-p_T jet; muon-jet samples triggered by the presence of two high-p_T jets, one of them containing a muon; $t\bar{t}$ samples triggered by the presence of one or two high-p_T leptons (e or μ) and enriched for $t\bar{t}$ through further selection criteria.

4.1 Misidentification probability

The probability of light-flavour quark and gluon jets being misidentified as b jets, $\varepsilon^{\text{misid}}$, is evaluated with negative tagging algorithms. The negative taggers are identical to the default algorithms, except that they are used only on tracks with negative IP values or on secondary vertices with negative decay lengths. To first order, the discriminator values for negative and positive taggers are expected to be symmetric for light-parton jets, since they are due only to resolution effects in track reconstruction. We can therefore derive the misidentification probability from the rate, ε^-, of negative-tagged jets in inclusive jet data. A correction factor, R_{light}, is evaluated from the simulation in order to correct for second-order asymmetries in the negative and positive tag rates of light-flavour quark and gluon jets, and for the heavy flavour contribution to the negative tags: $\varepsilon^{\text{misid}}_{\text{data}} = \varepsilon^-_{\text{data}} \cdot R_{\text{light}}$ where $R_{\text{light}} = \varepsilon^{\text{misid}}_{\text{MC}} / \varepsilon^-_{\text{MC}}$. The b-tagging discriminants are presented in Fig. 7 for the inclusive jet sample triggered with a p_T threshold larger than 40 GeV/c. For convenience, the discriminant values of the negative taggers are shown with a negative sign. The shapes from the MC simulation are normalised to the total number of jets in data. The data agree with the simulation within 20%. The difference can be due to physics-related effects (such as the fixed amount of b- and c-quark flavours in the simulation) or detector effects (as residual misalignments of the pixel detector).

Using different jet-trigger p_T thresholds, the misidentification probability can be computed in a wide jet p_T range from 20 to 1000 GeV/c, as illustrated in Figs. 8 and 9 for the JPL and CSVM tagging algorithms. The data/MC scale factor of the misidentification probabilities, $SF_{\text{light}} = \varepsilon^{\text{misid}}_{\text{data}} \,/\, \varepsilon^{\text{misid}}_{\text{MC}}$, is inferred. The misidentification probabilities measured with data and the SF_{light} values are given in Table 1 for jets with p_T between 80 and 120 GeV/c. Various sources of systematic uncertainties are considered, as detailed in [4]. The main systematic uncertainty is due to the the event kinematics, as the misidentification probability of a light-parton jet depends on whether there are other jets with higher p_T in the event. This uncertainty is evaluated by comparing different choices of inclusive jet trigger thresholds and ranges from 3 to 20%, depending on the operating point. Other relevant sources of uncertainties are due to: secondary interactions of charged particles in the detector material and the decay of long-lived K^0_S and Λ, which are corrected to the measured production rate as explained in [4]; the sensitivity of the negative tagging efficiency to the fraction of tracks from b and c jets, which is varied by a relative $\pm$20% with respect to its nominal value [20]. Uncertainties of 3 to 13% is estimated for each of these sources. Smaller contributions to the misidentification probability uncertainty come from the fraction of gluon jets in the simulation, the reconstructed tracks not associated to a charged particle, the change of sign of the impact parameter due to mismeasurement of the jet or tracks directions, and pileup effects.

The scale factors for the misidentification probability have also been measured as a function of the jet p_T for jets in several pseudorapidity intervals: $|\eta| < 0.5$, $0.5 \leq |\eta| < 1.0$, $1.0 \leq |\eta| < 1.5$ and $1.5 \leq |\eta| < 2.4$ for the loose operating points and $|\eta| < 0.8$, $0.8 \leq |\eta| < 1.6$ and $1.6 \leq$

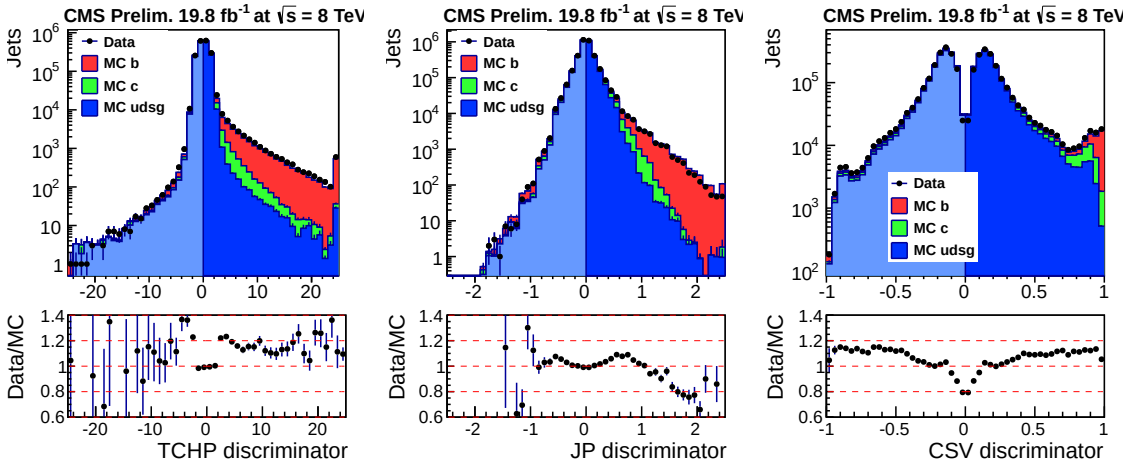


Figure 7: Comparison of the discriminator distributions for negative (negative sides of the distributions) and positive taggers (positive sides) in data (dots) and simulation for light-parton jets (blue histogram, with a lighter colour for the discriminator values from negative taggers), c jets (green histogram), and b jets (red histogram) for the (left) TCHP, (middle) JP, (right) CSV algorithms. A jet-trigger p_{T} threshold of 40 GeV/c is required for both data and MC. The simulation is normalised to the number of entries in the data. Underflow and overflow entries are added to the first and last bins, respectively. Typical values for the b-tagging requirement on these discriminators in the analyses range from 1 to 3.4 for TCHP and from 0.2 to 0.9 for JP and CSV.

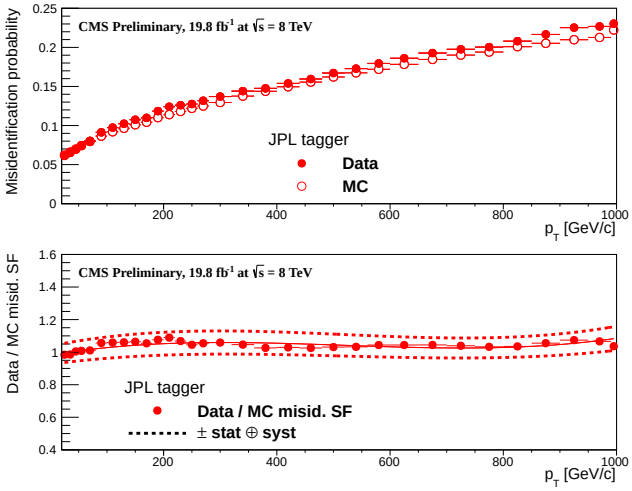


Figure 8: For the JPL tagging criterion: (top) misidentification probability in data (filled circles) and simulation (open circles); (bottom) scale factor for the misidentification probability. The last p_{T} bin in each plot includes all jets with $p_{\mathrm{T}} > 1000$ GeV/c. The solid curve is the result of a polynomial fit to the data points. The dashed curves represent the overall statistical and systematic uncertainties on the measurements.

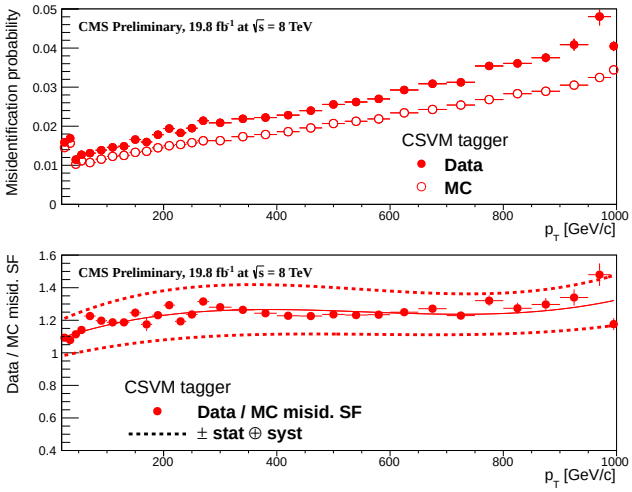


Figure 9: Same as Fig. 8 but for the CSVM tagging criterion.

Table 1: Misidentification probabilities in data and the data/MC scale factors SF_{light} for different algorithms and operating points for jet p_{T} in the range 80–120 GeV/c. The statistical uncertainties are quoted for the misidentification probabilities, while both the statistical and the systematic uncertainties are given for the scale factors.

b tagger	misidentification probability	SF_{light}
JPL	0.0944 ± 0.0004	$1.03 \pm 0.01 \pm 0.07$
CSVL	0.0990 ± 0.0004	$1.10 \pm 0.01 \pm 0.05$
JPM	0.0105 ± 0.0002	$1.10 \pm 0.02 \pm 0.20$
CSVM	0.0142 ± 0.0002	$1.17 \pm 0.02 \pm 0.15$
TCHPT	0.0026 ± 0.0001	$1.27 \pm 0.06 \pm 0.27$
JPT	0.0013 ± 0.0001	$1.11 \pm 0.07 \pm 0.31$
CSVT	0.0016 ± 0.0001	$1.26 \pm 0.07 \pm 0.28$

$|\eta| < 2.4$ for the medium operating points. For each b-tagging algorithm, the scale factors are compatible with one within about 10-30%. From the loose to tight operating points, the SF_{light} values are measured with a precision of about 5-10 to 30%, respectively. The deviations of SF_{light} from unity increase as the cut on the discriminant is tightened, as tails of the distributions are more difficult to describe in the simulations. Also, we notice that the values of SF_{light} for JP are closer to one, as a consequence of the fact that this algorithm uses a calibration mostly based on tracks from light-parton jets, both in data and simulation.

4.2 b-tagging efficiency

4.2.1 b-tagging efficiency from multijet events

The b-tagging efficiency is measured in data using several methods applied to multijet events [4]. Comparing the measured value with the identification efficiency for b jets in the simulation, a data/MC scale factor of the b-tagging efficiency is inferred: $SF_{\text{b}} = \varepsilon_{\text{b}}^{\text{data}} / \varepsilon_{\text{b}}^{\text{MC}}$.

The PtRel, IP3D and LT methods are based on a sample of jets enriched in heavy flavour content by requiring a soft muon within a cone $\Delta R < 0.4$ around the jet axis (*muon-jet*). The fraction of jets from b-quark hadronisation in the selected sample is estimated by fitting the data distribution of a discriminant variable, providing a good separation between b and light-parton

jets, to the expected templates for the different jet flavours. The b-tagging efficiency in data is measured by the ratio of the number of b jets estimated by the fit in the muon-jet sample and the one estimated by repeating the fit on the subsample of muon jets that pass the tagging requirement. In addition, the LT method is implemented using a sample tagged by a reconstructed $J/\psi \to \mu^+\mu^-$ decay. The System8 method is also based on the muon-jet sample, but it is devised to make a minimal use of MC information. Further details on the different methods are given in the following:

- PtRel: the transverse momentum of the muon relative to the jet axis, p_T^{rel}, is used as the discriminant variable. In order to reduce the light-flavour background, another tagged jet ("away tag") is required in the event. The p_T^{rel} shape from b jets is taken from the simulation. The sum of the shapes from c and light-parton jets in MC is corrected from the ratio of shapes observed for inclusive charged particles in data and MC. An example of fits to p_T^{rel} is given in Fig. 10.
- IP3D: the 3D impact parameter of the muon track is used as a discriminant variable. An away-tagged jet is required. Shapes from b, c and udsg jets are taken from the simulation. An example of IP3D fits is given in Fig. 11.
- LT: the discriminant variable is chosen to be the b-tag discriminator of the JP algorithm for the measurement of the b-tagging efficiency of the CSV and TCHP algorithms. Shapes from b, c and udsg jets are taken from the simulation.

  The b-tagging efficiency of the Jet Probability algorithm is obtained by fitting the CSV distribution. Since the CSV tagger is not calibrated on data, a correction is computed by comparing the b-tagging efficiencies measured for the Track Counting algorithm when either JP fits or CSV fits are performed.
- LTJ/ψ: instead of considering muon jets, events are selected where a J/ψ decaying to a muon pair is reconstructed inside a jet. Applying a tight selection on the J/ψ decay length, the J/ψ is mainly produced from a b-hadron decay and another b-quark jet is expected in the event. Fits are thus performed with the LT method for a jet opposite to the J/ψ.
- System8: in muon-jet events, instead of relying on simulated templates of some discriminant variable, this method applies three weakly correlated b-tagging requirements to the data. The reference lifetime tagger and a supplementary p_T^{rel} based selection are tested on the muon jet; the third weakly correlated requirement is a separate lifetime tagger selection on the leading away jet. The numbers of events satisfying each of the eight possible tagging conditions can be written as a function of the unknown numbers of b and non-b jets in the sample, and of the unknown tagging efficiencies, which can be derived by resolving the system of eight equations.

Each method has discrimination power between b and light-parton jets in different ranges of jet transverse momentum. The PtRel and the System8 methods provide precise measurements for the lower part of the spectrum ($20 < p_T < 120\,\text{GeV}/c$). The J/$\psi$ method is suited for $p_T < 160\,\text{GeV}/c$. The IP3D method has been designed for high jet $p_T > 120\,\text{GeV}/c$. The LT method can be applied on the full p_T range from 20 up to $800\,\text{GeV}/c$.

Various systematic uncertainties are considered, some common to all methods and some more specific [4]. Among the common systematic uncertainties are the pileup description, the rate of gluon splitting into b quark pairs, the muon p_T selection, the b-fragmentation modelling, the requirement of an away-jet tag (for PtRel, IP3D and System8), the ΔR requirement between the muon direction and the jet axis (for PtRel and System8).

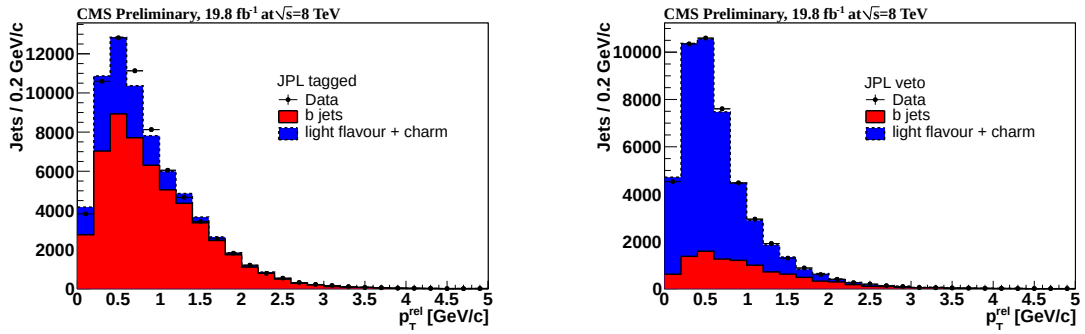


Figure 10: Fits of the summed b and non-b templates, for simulated muon jets, to the muon $p_{\mathrm{T}}^{\mathrm{rel}}$ distributions from data. (left) and (right) show the results for muon jets that pass (tagged) or fail (veto) the JPL b-tagging criterion, respectively. The muon-jet p_{T} is between 80 and 120 GeV/c.

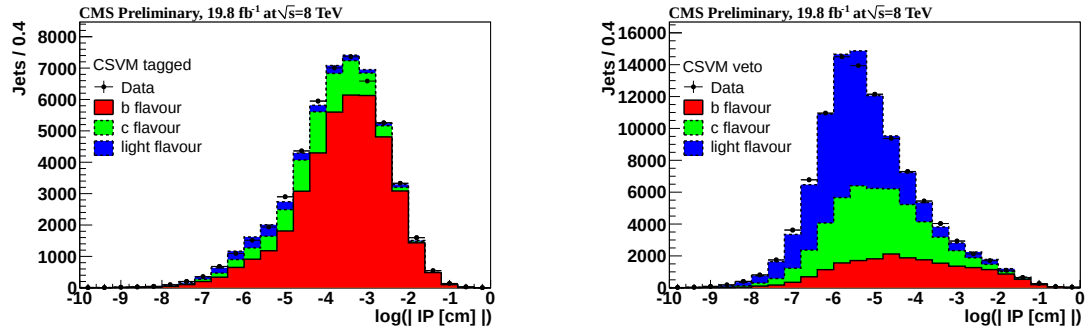


Figure 11: Same as Fig. 10 using the $\ln(|\mathrm{IP}|)$ distributions of the muon with the CSVM b-tagging criterion. The muon-jet p_{T} is between 160 and 320 GeV/c.

Specific to PtRel is the relative fraction of c and udsg jets. Specific to System8 is the $p_{\mathrm{T}}^{\mathrm{rel}}$ selection. Specific to LT and LTJ/ψ are the fraction of b jets with JP information (0.8% and 1.2% for CSVM at $80<p_{\mathrm{T}}<120$ GeV/c, respectively) and the correction when a fit to the CSV discriminant is performed (1-2% for JPL at $80<p_{\mathrm{T}}<120$ GeV/c). Specific to LTJ/ψ are the J/ψ selection criteria (0.9% for CSVM at $80<p_{\mathrm{T}}<120$ GeV/c).

Specific to LT is a comparison of SF_{b} results between muon jets and inclusive jets (where an away-tagged jet is required) in order to assess if the muon requirement changes the performance of the lifetime taggers (1.8% for CSVM at $80<p_{\mathrm{T}}<120$ GeV/c). Fits to the JP distribution are sensitive to the assumed shapes for b, c and light-parton jets in the simulation. The track-probability distributions are calibrated independently in data and simulation to account for impact parameter resolution effects. As mostly tracks from light-parton jets are used in the calibration, this improves the agreement between data and MC for udsg jets. The inaccurate simulation of the tracker resolution can distort the JP distribution for c and b jets, but as shown in Fig. 3 the level of agreement between data and simulation is consistent for different flavour compositions. A relative variation of $\pm 3\%$ of the track multiplicity from b-hadron decays is also considered [21], the inferred SF_{b} value is found compatible within the quoted systematic uncertainty. The uncertainty on the tracking efficiency [22] would lead to a similar variation of the track multiplicity in b jets.

As an illustration, the breakdown of the systematic uncertainties is given in Table 2 for the CSVM tagging criterion in the jet p_{T} range 80–120 GeV/c.

Table 2: Relative systematic uncertainties on SF_b for the CSVM tagging criterion in the muon-jet p_T range 80–120 GeV/c.

method	pileup	g $\to$ b$\bar{b}$	b frag.	p_T^{μ}	away jet	$\Delta R(\mu$ jet)	specific	total
PtRel	0.2%	0.0%	0.4%	2.8%	3.3%	1.9%	1.0%	4.9%
System8	0.8%	0.0%	0.6%	1.0%	2.1%	1.1%	1.0%	2.9%
LT	0.1%	0.9%	0.4%	0.2%	0.0%	0.0%	2.0%	2.2%
J/ψ	0.3%	0.0%	0.4%	0.0%	0.0%	0.0%	1.5%	1.6%

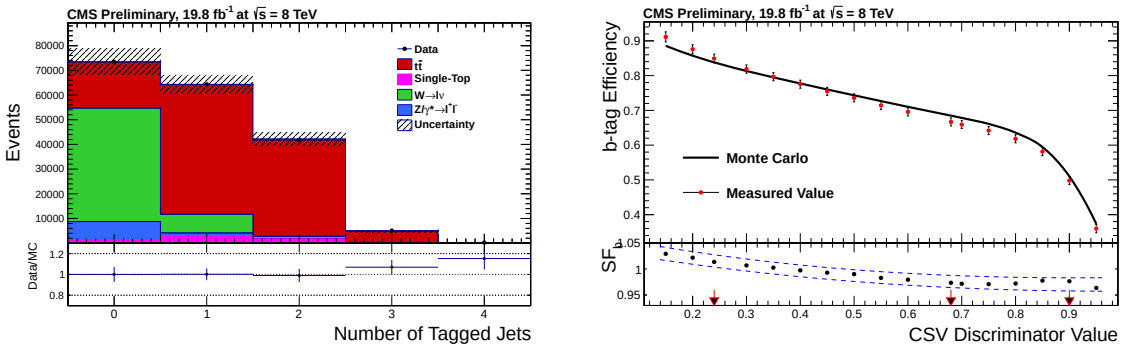


Figure 12: With the FTC method in the lepton+jet channel: (left) Number of tagged jets per event with the CSVM operating point in data (filled circles) and simulation (solid histograms, normalised to the integrated luminosity) after the FTC fit. The hatched area corresponds to the combined statistical and systematic uncertainty. (right) b-jet tagging efficiency as a function of the discriminator threshold for the CSV algorithm. The efficiency measured in data and predicted from simulation are shown in the upper panel. The scale factor SF_b is shown in the lower panel, where the blue dashed lines represent the combined statistical and systematic uncertainty. The arrows indicate the standard operating points.

4.2.2 b-tagging efficiency from $t\bar{t}$ events

Several methods are applied to measure the b-tagging performance using $t\bar{t}$ events in both the decays into lepton+jets and dilepton final states. In the lepton+jets channel, the flavour tag consistency (FTC) method and the bSample method are used, as detailed in [4]. In the dilepton channel, the flavour tag matching (FTM) method [4] is used as well as the LT method (denoted LTtop), which can be applied on the same events. For the FTC, bSample and FTM methods, only an average b-tagging efficiency is measured, which corresponds to b jets with an average $p_T \simeq 83\,\text{GeV}/c$ (for $p_T > 30\,\text{GeV}/c$) in $t\bar{t}$ events. The LTtop method also allows a p_T dependent measurement of the data/MC scale factor value of the b-jet tagging efficiency.

The FTC method (FTM method) requires consistency between the observed and expected number of tags in the lepton+jets (dilepton) events. A log-likelihood fit is performed with the b-jet tagging efficiency as free variable. For the FTC method, the distributions of the number of b-tagged jets observed in data and predicted in the simulation for $t\bar{t}$ and background events after the fit are shown in Fig. 12(left). These $t\bar{t}$ methods can be used to measure the b-tagging efficiency and scale factor as a function of the operating point value as illustrated in Fig. 12(right).

In the bSample method, the b-jet tagging efficiency is measured from a sample enriched with b jets through the selection of $t\bar{t}$ events in lepton+jets final states. A kinematic fit is used to

associate the four leading jets to the quarks from top decay. The candidate b jet from the leptonically decaying top quark can be used to define a b-jet sample. This sample is divided into two subsamples, one enriched in b jets and the other depleted, by looking at the low mass region of the distribution of the invariant mass of the charged lepton + candidate b-jet system (see Fig. 13). In the b-enriched sample, the b-tagging discriminator distribution for b jets can be reconstructed by subtracting the discriminator distribution in the b-depleted subsample. In this subtraction, the b-depleted distribution is scaled by a factor, which accounts for the different fraction of non-b jets between the two samples. This scale factor is determined by looking at the same regions of the invariant mass between the charged lepton and the candidate jets from W boson decay, which provide a highly pure sample of light-parton jets.

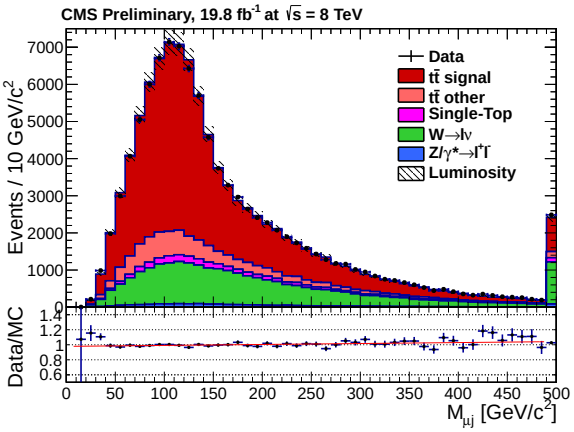


Figure 13: Distribution of the jet-muon mass in the data (filled circles) and the simulation (solid histograms). The simulated distribution is normalised to the integrated luminosity. The component "$\mathrm{t\bar{t}}$ *signal*" stands for the lepton+jet events. The component "$\mathrm{t\bar{t}}$ *other*" contains the events in all other decay channels. The hatched area corresponds to the combined statistical and systematic uncertainty. Overflow entries are added to the last bin.

In the LTtop method with dileptons, fits to the JP discriminant are shown before and after applying the CSVM tagging criterion in Fig. 14. The data/MC scale factor of the CSV b-tagging efficiency can be measured as a function of the jet p_{T}, $|\eta|$ or number of primary vertices. As shown in Fig. 15, although the efficiencies display some dependencies, the SF_{b} distributions are rather flat.

Most of the sources of systematic uncertainties are common to all methods in $\mathrm{t\bar{t}}$ events [4]: modelling of the underlying events, scale of the jet-parton matching, definition of the renormalisation and factorisation scales, choice of the parton distribution function, pileup description, jet energy scale and jet energy resolution in the CMS detector, uncertainty on the production cross section for the physical backgrounds and light-parton tagging efficiency (for FTC and FTM), top-quark mass and definition of the b-depleted region (for bSample), limited number of events in the MC templates (for LTtop).

4.2.3 Combination of b-tagging efficiency measurements

The PtRel, IP3D, LT and LTJ/ψ measurements of the data/MC scale factor SF_{b} are combined in each jet p_{T} bin. As an alternative, the SF_{b} measurement from the LTtop method can also be included in the combination in order to provide results for analyses which are not measuring top quark properties in the same final states with two charged leptons as the ones used by the LTtop method.

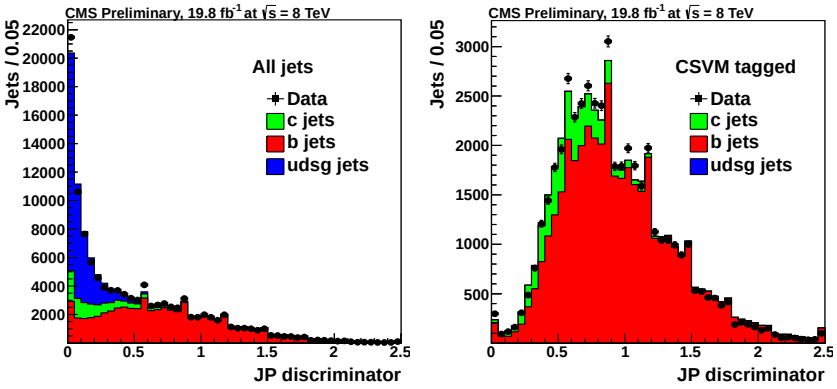


Figure 14: Fits of the summed b, c and light-parton templates, for simulated jets, to the JP-discriminant distributions from $t\bar{t}$ into dilepton (ee, eμ, $\mu\mu$) candidate events in data. (left) and (right) show the results for jets before and after identification with the CSVM tagging criterion, respectively. The small discontinuities in the JP distributions are due to the single track probabilities, which are required to be greater than 0.5%. The black line is the sum of the contributions from the templates. Overflows are displayed in the rightmost bins.

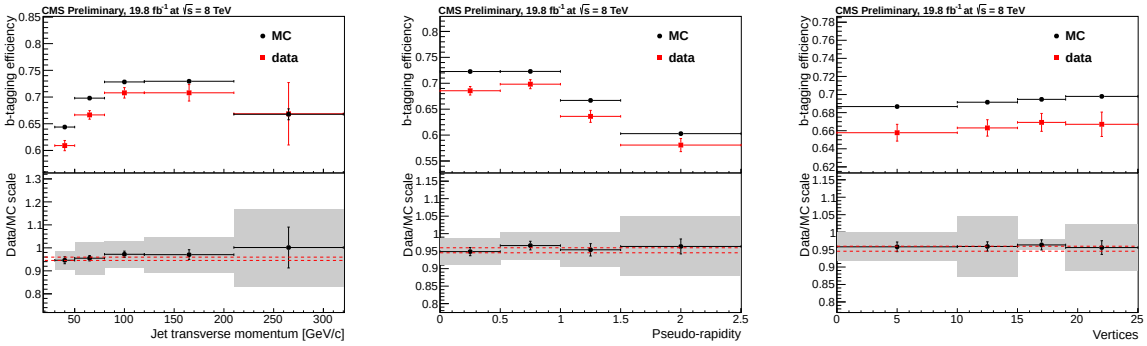


Figure 15: In $t\bar{t}$ into dilepton events with the LTtop method for the CSVM criterion: (upper panels) b-tagging efficiencies and (lower panels) data/MC scale factor SF_b as a function of (left) jet p_T, (middle) jet $|\eta|$ and (right) number of primary vertices. In the lower panels, the grey filled areas represent the total statistical and systematic uncertainties, whereas the dotted lines are the average SF_b values within statistical uncertainties.

The combination is based on a weighted mean of the scale factors [23], taking into account correlated and uncorrelated uncertainties. The shared fraction of events between the muon-jet measurements has been evaluated. Typical values are 6–24% between the LT and PtRel/IP3D methods, 46–58% between the PtRel and System8 methods, and 10–34% between the System8 and LT methods. This overlap has been taken into account in the combination, whereas the $t\bar{t}$ into dilepton sample is independent from the muon-jet sample. Correlation between systematic uncertainties has been discussed at the end of sections 4.2.1 and 4.2.2. If the χ^2 from the fit exceeds the number of degrees of freedom, the uncertainty is scaled by the square root of the normalised χ^2, when combining the results.

Summaries for the individual and combined scale factor measurements for the JPL and the CSVM tagging algorithms are shown in Figs. 16 and 17. Also shown are the parameterisations of the combined scale factor of the form $SF_b(p_T) = \alpha(1+\beta p_T)/(1+\gamma p_T)$.

The four data/MC scale factor values measured in $t\bar{t}$ events in section 4.2.2 can be also com-

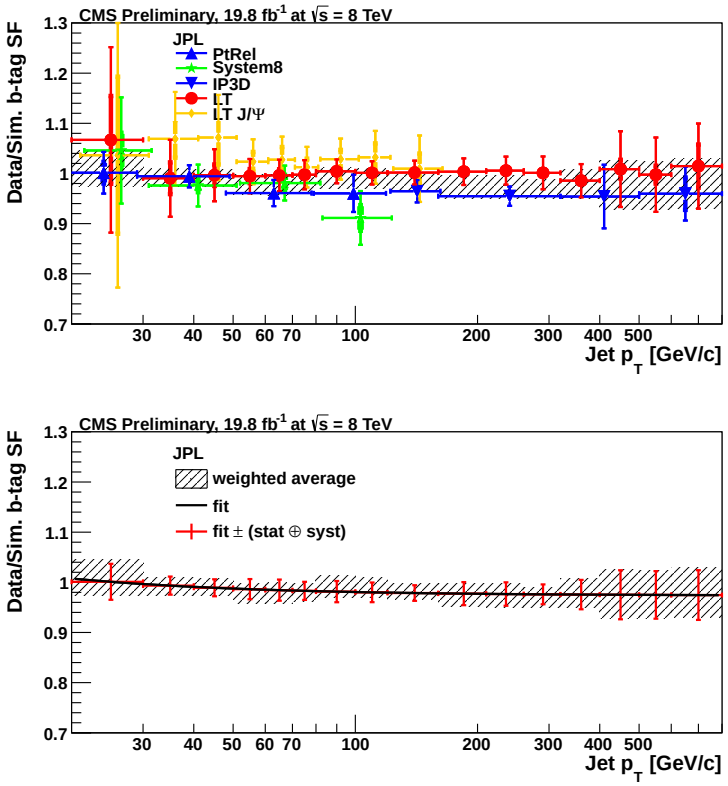


Figure 16: Individual and combined measurements of the data/MC scale factor SF_b of the b-tagging efficiencies for the JPL tagging criterion. (top) individual measurements from the muon p_T^{rel} ("PtRel"), System8, muon IP ("IP3D"), lifetime tagger ("LT") and J/ψ methods. The inner and outer error bars indicate the statistical and the overall statistical and systematic uncertainties, respectively. The hatched areas represent the combined measurements from muon-jet data. (bottom) the combined measurement has been parameterised as $SF_b(p_T) = \alpha(1+\beta p_T)/(1+\gamma p_T)$. The error bars attached to the function have the same size as the uncertainties from the combined measurement in each bin.

pared. The FTC and FTM methods are used to calculate a combined scale factor, by taking the weighted mean of the scale factors from each method [23]. 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 provided uncertainty on the combined result is the RMS of the four results, see Table 3.

In order to compare muon-jets, J/ψ and $t\bar{t}$ results, the p_T-dependent scale factors measured in multijet events have been reweighted based on the observed jet p_T spectrum in $t\bar{t}$ events. For SF_b from muon-jet events, the combined result of the different methods is used. The results are shown in Table 4 and in Fig. 18: a good agreement is obtained between the measurements. These measurements also show that the scale factors for the muon jets and inclusive jets are compatible.

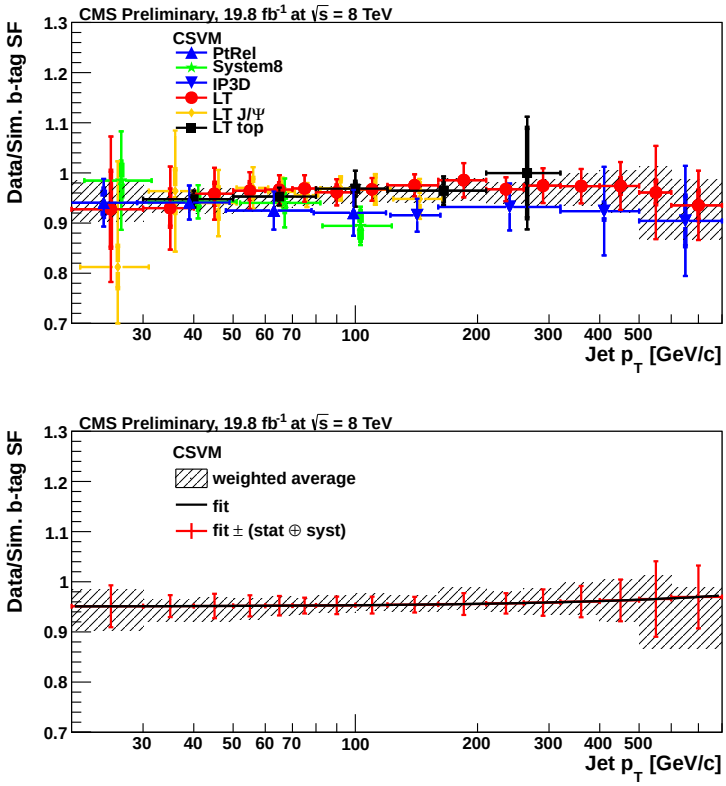


Figure 17: Same as Fig. 16 for the CSVM tagging criterion. Results from $t\bar{t}$ into dilepton events ("LTtop") are also included.

Table 3: b-tagging efficiency scale factors derived from $t\bar{t}$ events. The overall uncertainties are given.

Method	bSample	FTC	LTtop	FTM	Combined
JPL	0.992 ± 0.018	0.965 ± 0.012		0.967 ± 0.013	0.966 ± 0.015
CSVL	1.007 ± 0.015	1.014 ± 0.014	0.98 ± 0.03	0.977 ± 0.009	0.987 ± 0.018
JPM	0.974 ± 0.026	0.966 ± 0.015		0.960 ± 0.012	0.961 ± 0.012
CSVM	0.967 ± 0.017	0.973 ± 0.013	0.95 ± 0.04	0.952 ± 0.009	0.953 ± 0.012
TCHPT	0.930 ± 0.024	0.922 ± 0.017	0.92 ± 0.06	0.921 ± 0.009	0.921 ± 0.010
JPT	0.926 ± 0.032	0.923 ± 0.017		0.927 ± 0.022	0.922 ± 0.017
CSVT	0.977 ± 0.021	0.976 ± 0.015	0.93 ± 0.05	0.928 ± 0.010	0.926 ± 0.036

Table 4: Scale factors SF_b obtained in muon-jet data, J/ψ events and $t\bar{t}$ data for b jets in the p_T range of $t\bar{t}$ events. The overall uncertainties are given.

b tagger	SF_b in muon jets	SF_b in LTJ/ψ events	SF_b in $t\bar{t}$ events
JPL	0.982 ± 0.020	1.003 ± 0.056	0.966 ± 0.015
CSVL	0.983 ± 0.017	0.985 ± 0.070	0.987 ± 0.018
JPM	0.947 ± 0.034	0.977 ± 0.066	0.961 ± 0.012
CSVM	0.951 ± 0.024	0.964 ± 0.071	0.953 ± 0.012
TCHPT	0.896 ± 0.035	0.926 ± 0.084	0.921 ± 0.010
JPT	0.866 ± 0.036	0.901 ± 0.080	0.922 ± 0.017
CSVT	0.916 ± 0.032	0.920 ± 0.104	0.926 ± 0.036

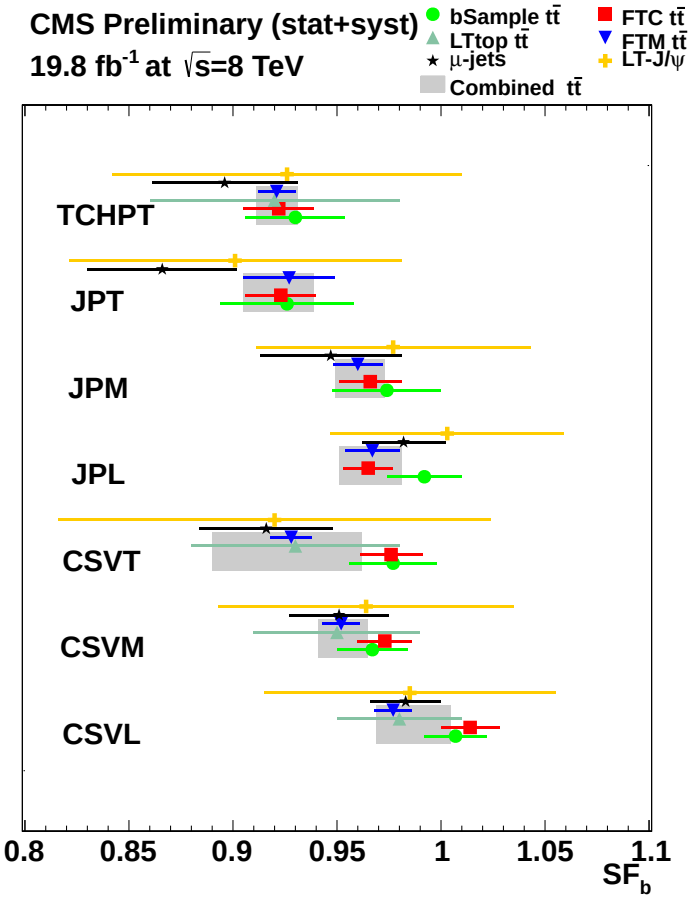


Figure 18: Summary of the average SF_b measurements in the b-jet p_T range of $t\bar{t}$ events. The combined $t\bar{t}$ values are displayed as grey filled areas.

5 b tagging in boosted topologies

High-mass resonances, which are predicted in various models of physics beyond the standard model, such as heavy resonances decaying to $t\bar{t}$ or heavy quarks decaying to standard model quarks and Higgs bosons, may decay to final states containing several b quarks. As the excluded mass regions continue to increase, searches for new physics increasingly focus on event topologies with boosted top quarks and Higgs bosons. The key feature of such event topologies is that top quarks and Higgs bosons are imparted large enough momentum that their motion becomes highly relativistic and their decay products highly collimated. Consequently, the decay products can end up clustered in a single fat jet, depending on the momentum of the decaying particle and the jet size, as can be inferred from Fig. 19. An approximate spread of a hadronic top-quark decay at the partonic level can be determined from the ΔR distances between the three main decay products. The four-momenta of the two quarks with the smallest distance, $\Delta R(q_1, q_2)$, are vectorially summed and the ΔR distance between the sum vector and the third quark, $\Delta R(q_{1+2}, q_3)$, is evaluated. The maximum distance between $\Delta R(q_1, q_2)$ and $\Delta R(q_{1+2}, q_3)$ indicates the approximate size ΔR_{bij} of a Cambridge/Aachen (CA) jet [24, 25] containing the entire top-quark decay. The spread of a Higgs boson decay into a pair of b quarks is quantified using the ΔR distance between the two quarks $\Delta R(b, \bar{b})$.

Boosted topologies are usually reconstructed and interpreted using the jet substructure methods such as top- [1] and W/Z-tagging [2] algorithms. However, if one or more of the decay products is a b quark, adding b-jet identification in the jet substructure can significantly improve the sensitivity of these methods. In the following we present methods of applying b tagging in boosted topologies along with a validation and measurement of the performance. The event topologies with boosted top quarks are studied in samples of simulated $T'\overline{T'} \to tH\bar{t}H$ events with a T' mass of 1 TeV/c^2 and a Higgs boson mass of 120 GeV/c^2, while the event topologies with boosted Higgs bosons are studied in samples of simulated $b'\overline{b'} \to bH\bar{b}H$ events with a b' mass of 1 TeV/c^2 and 1.5 TeV/c^2 and a Higgs boson mass of 120 GeV/c^2. For topologies with boosted top quarks, the comparison of the simulation with the data is performed using a $t\bar{t}$-enriched sample. However, for topologies with boosted Higgs bosons, direct measurements of the b-tagging performance are currently not possible due to a limited sample size. Instead, a data sample enriched in gluon splitting to $b\bar{b}$ is used.

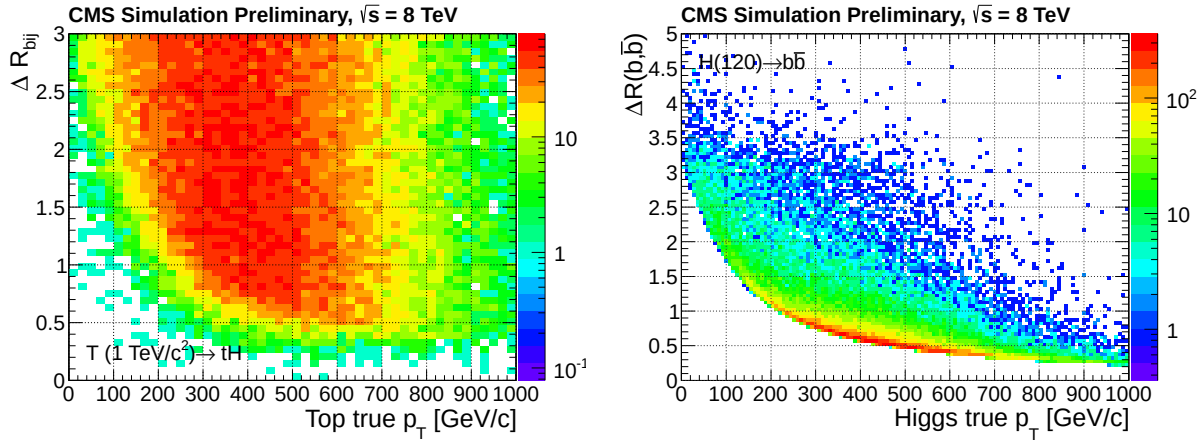


Figure 19: (left) Approximate ΔR distance between the three top-quark decay products versus p_T of the top quark from a simulated T' pair production sample, with a T' mass of 1 TeV. (right) $\Delta R(b, \bar{b})$ between the two b quarks from the Higgs boson decays versus p_T of the Higgs boson obtained from a simulated $b'\overline{b'} \to bH\bar{b}H$ sample with a b' mass of 1 TeV.

5.1 b tagging in jet substructure

A variety of algorithms for the identification of fully merged top jets, so called top-tagging algorithms, have been suggested in the last couple of years [1]. One important criterion for such top-tagging algorithms is the size of the jet. Larger jets are more likely to include all top-quark decay products. The optimal jet size depends on the momentum of the top quark. As the data-driven validation and performance measurement of such top jets strongly depend on the available statistics of boosted top-quark samples, we use a top-tagging algorithm, which works sufficiently well at low transverse momenta. We therefore use the `HEPTopTagger` algorithm [1], which is based on Cambridge/Aachen jets of size $R = 1.5$ (CA15). It allows moderately boosted top quarks to be reconstructed, down to transverse momenta of the order of 200 GeV/c [1]. For each fat jet j, the substructure of the jet is identified undoing the Cambridge/Aachen clustering of the jet, with an additional mass drop criterion. Fat jets with less than three hard subjets are discarded. Otherwise, for all possible pairings of three subjets a filtering is applied with variable resolution $R_{\text{filt}} = min(0.3, \Delta R_{ij}/2)$, where i and j are the closest subjets in ΔR. For each pairing, only the five hardest filtered constituents are kept and the mass of the fat jet is calculated from their vectorial sum. In order for the fat jet to be considered top-tagged, the masses of the three subjets are required to satisfy specific criteria, which are based on the pairing of two subjets to the mass of the W boson. Additionally, the reconstructed mass of the fat jet is required to be within 115 GeV/c^2 and 250 GeV/c^2.

Two baseline b-tagging approaches have been investigated:

- application of b tagging to fat jets,
- application of b tagging to subjets, which are reconstructed within fat jets.

The CMS b-tagging chain starts with an association of tracks to jets, based on the angular distance between the tracks and the jet axis. Most b-tagging algorithms are using a selection of $\Delta R < 0.3$. However, when applying this to a fat jet of size $R = 1.5$, the criterion may not be optimal anymore. For the study of applying b tagging to fat jets, this angular distance has therefore been enlarged to $\Delta R < 1.5$. For the application of b tagging to subjets, the angular distance remains at the default value of $\Delta R < 0.3$.

The simulated performance of the CSV algorithm on subjets and fat jets is shown in Fig. 20 for two different p_{T} intervals of the fat jet. At least one of the three subjets identified by the `HEPTopTagger` algorithm is required to be b-tagged in this case. It is clearly visible that the performance on subjets is better than on fat jets.

The overall performance when combining the `HEPTopTagger` algorithm with a subjet b-tag is shown in Fig.21. Two different operating points of the CSV algorithm have been studied, loose (CSVL) and medium (CSVM). A strong reduction of the misidentification probability of a factor of ten is observed, while the tagging efficiency is only moderately decreased.

A similar investigation is made for boosted Higgs bosons decaying to pairs of b quarks. In this case, however, smaller-size CA8 jets ($R = 0.8$) are used. As in the case of boosted top quarks, the two options of b-tagging fat jets and subjets are considered. Since the tracks used for b tagging are collected within a cone around the jet axis, in order to get the best possible performance, it is important to define the subjets in such a way that their momenta strongly correlate with the direction of the two b quarks from the Higgs boson decay. Pruned subjets, obtained by undoing the last clustering iteration of the pruned jet clustering [26], were found to have the best performance.

For a mass of the Higgs boson of 120 GeV/c^2 and a jet size $R = 0.8$, the boosted regime starts at

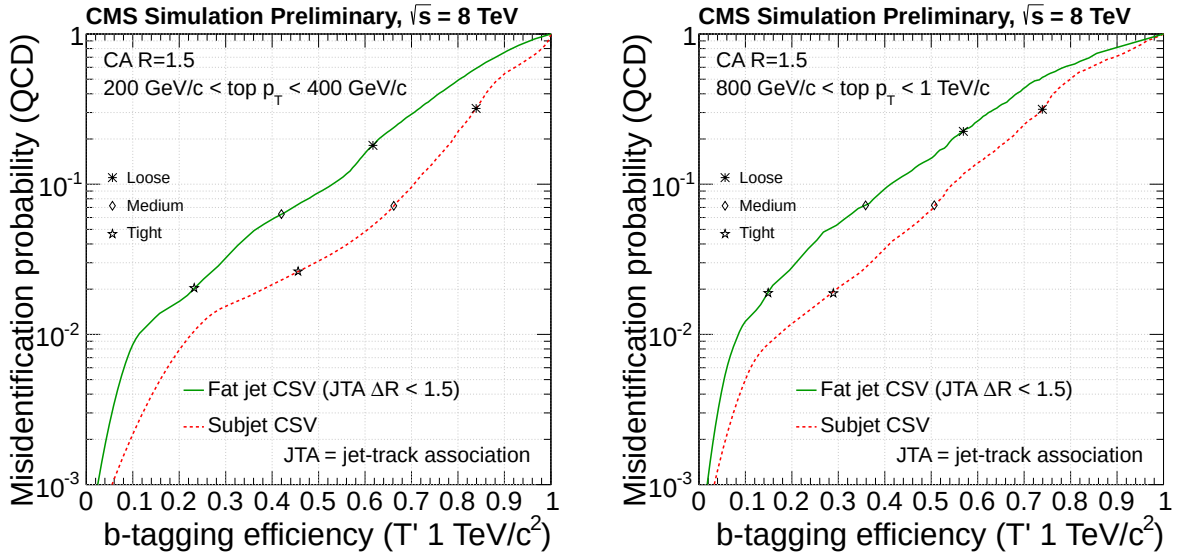


Figure 20: Performance of the CSV algorithm on CA fat jets of size $R = 1.5$ and on subjets of the same jets. The misidentification probability on inclusive QCD jets is shown versus the b-tagging efficiency on a sample of boosted hadronically decaying top quarks from $T' \to tH$ decays. The results are shown in two intervals of the p_T of the fat jet: (left) for $200 < p_T < 400$ GeV/c and (right) $800 < p_T < 1000$ GeV/c.

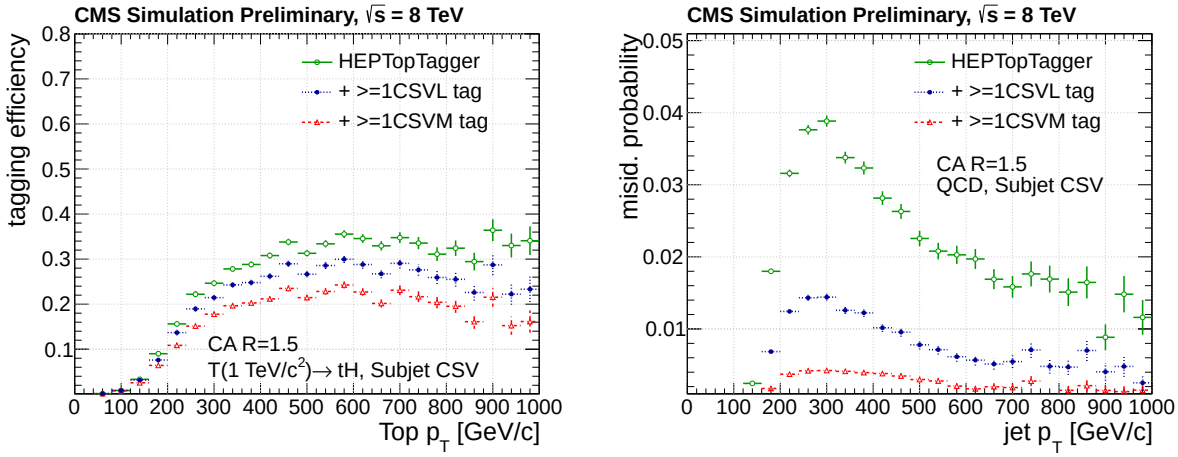


Figure 21: The performance of the `HEPTopTagger` alone compared to a combination with a subjet b-tagging requirement. (left) Tagging efficiency in boosted top quarks from simulated $T' \to tH$ events. The efficiency is shown versus the true p_T of the top quark. (right) Inclusive misidentification probability from QCD jets versus the jet p_T.

$p_T \sim 300$ GeV/c. A requirement on the pruned jet mass [26], $75 < m_{jet} < 135$ GeV/c^2, is used, which rejects a large fraction of QCD background. In addition to this jet mass requirement, the b tagging is applied either on the fat jet or on the two subjets, where both subjets are required to pass the same selection on the CSV discriminator. The (double) b-tagging efficiency for this topology is shown in Fig. 22. As in the case of boosted top quarks, the subjet tagging outperforms the fat jet tagging in the largest part of the phase space.

The resulting (double) b-tagging and overall Higgs-tagging performance at the CSVL operating point for the $H \to b\bar{b}$ signal and several different backgrounds, where the Higgs tagging is defined as a combination of a requirement on the pruned jet mass ($75 < m_{jet} < 135$ GeV/c^2)

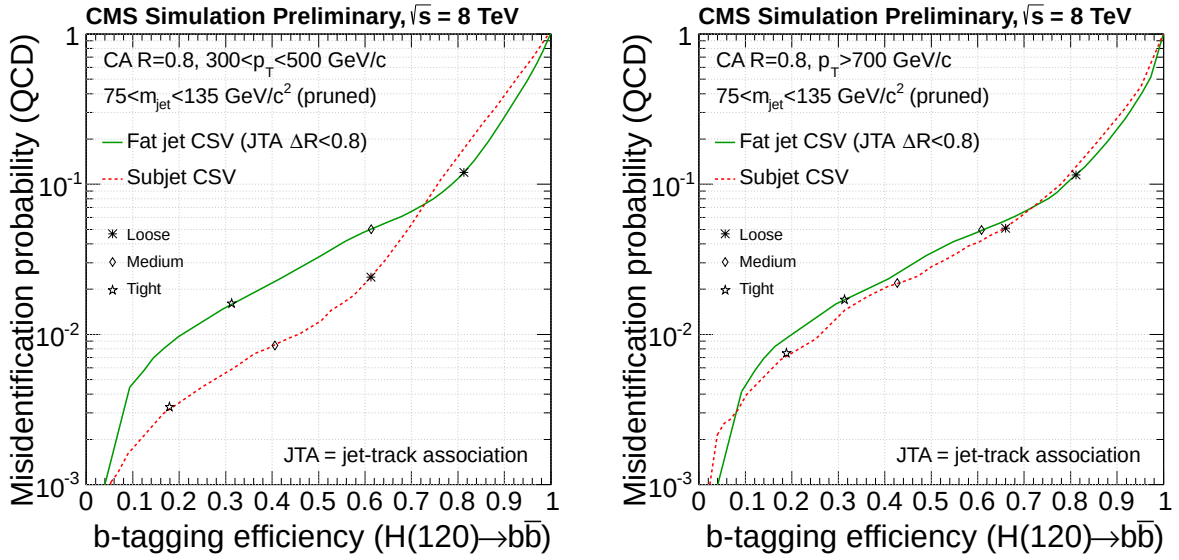


Figure 22: Misidentification probability as a function of b-tagging efficiency for boosted $H \to b\bar{b}$ jets and inclusive QCD jets for the CSV algorithm applied to fat jets and pruned subjets for fat jets with (left) $300 < p_T < 500$ GeV/c and (right) $p_T > 700$ GeV/c. For b tagging on fat jets, a relaxed jet-track association cone of $\Delta R < 0.8$ is used. For b tagging on subjets, both subjets are required to pass the same selection on the CSV discriminator. Loose, medium, and tight operating points are indicated.

and subjet b tagging, is displayed in Fig. 23. While the Higgs-tagging efficiency for the boosted $H \to b\bar{b}$ jets is between 40% and 50%, the QCD background is reduced to about 0.4%. The dominant non-QCD background comes from boosted Z bosons and top quarks.

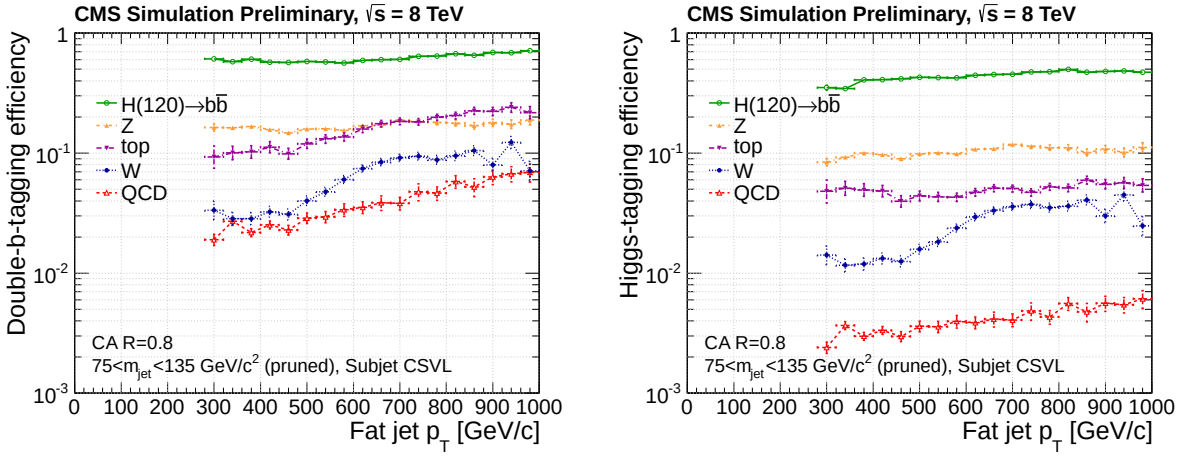


Figure 23: (left) Double-b-tagging efficiency and (right) Higgs-tagging efficiency as a function of the fat jet p_T for the boosted $H \to b\bar{b}$ signal and several different backgrounds. The b-tagging efficiency shown is for fat jets passing a requirement on the pruned jet mass ($75 < m_{jet} < 135$ GeV/c^2). The Higgs tagging is defined as a combination of this requirement on the pruned jet mass and subjet b tagging (CSVL).

It should be pointed out that the use of a fixed-size jet-track association cone inevitably leads to track sharing between the subjets of fat jets once their angular separation becomes comparable or smaller than the size of the association cone. For boosted top jets, the average fraction of shared tracks, defined as the ratio of the number of tracks within $\Delta R < 0.3$ from more than

one subjet and the number of all tracks within $\Delta R < 0.3$ from any of the subjets, ranges from a few percent at a fat jet p_{T} of 400 GeV/c and increases to ∼40% at a fat jet p_{T} of 1 TeV/c. For boosted Higgs jets the same fraction ranges from a few percent at a fat jet p_{T} of 400 GeV/c and increases to ∼40% at a fat jet p_{T} of 700 GeV/c and ∼ 80% at a fat jet p_{T} of 1 TeV/c. Because of track sharing, the b-tagging probabilities for individual subjets become increasingly correlated as the fat jet transverse momentum increases. For boosted Higgs jets, where both subjets are required to be b-tagged, this finally results in the subjet b-tagging performance approaching the fat jet b-tagging performance at large fat jet transverse momenta, as can be seen in Fig. 22.

5.2 Validation of b-tagging variables

The validation of b tagging in boosted top jets is done by selecting a sample of semi-leptonic $\mathrm{t\bar{t}}$ events, in which one top quark decays hadronically and the other decays leptonically. The events are selected by requiring an isolated muon with $p_{\mathrm{T}} > 45$ GeV/c and $|\eta| < 2.1$. Additional muons or electrons in the event are vetoed. The isolated muon divides the event into two hemispheres. The leptonically decaying top quark is selected by requiring at least one CSVM-tagged AK5 jet with $p_{\mathrm{T}} > 30$ GeV/c and $|\eta| < 2.5$ in the leptonic hemisphere. In the hadronic hemisphere exactly one CA15 fat jet is required, with $p_{\mathrm{T}} > 150$ GeV/c and $|\eta| < 2.5$ and fulfilling the top-tagging requirements of the `HEPTopTagger` algorithm. The three subjets of this fat jet clustered according to the `HEPTopTagger` requirements are used to measure the scale factors for the subjet b tagging.

Furthermore, the validation of b tagging in boosted Higgs jets requires the presence of a jet with two b quarks within it. One such standard model process consists of jets from gluon splitting to $\mathrm{b\bar{b}}$ in which the two b quarks hadronise inside one jet. Event samples containing high-p_{T} jets reconstructed using the CA8 algorithm are used to select such gluon splitting to $\mathrm{b\bar{b}}$ jets. The CA8 jets are required to have a transverse momentum $p_{\mathrm{T}} > 400$ GeV/c. The jet-pruning algorithm is used to decompose the fat jets into their subjet components. To remove infrared unsafe pruned subjet configurations, the following requirement is imposed [27]: $\Delta R(\text{subjets}) > R_{\text{cut}} \times 2m/p_{\mathrm{T}}$. Here $\Delta R(\text{subjets})$ is the angular separation between the two subjets of a pruned CA8 jet, R_{cut} $(= 0.5)$ is the jet-pruning parameter, m the mass of the unpruned CA8 jet and p_{T} is its transverse momentum. This cut removes fat jets containing one very soft subjet with the other subjet nearly collinear with the fat jet axis. The selection described so far defines what is referred to as the inclusive sample of fat jets. In order to enrich this sample in heavy-flavour jets, at least one of the two subjets can be required to be muon-tagged where a muon candidate with $p_{\mathrm{T}} > 5$ GeV/c needs to be matched to a subjet within $\Delta R < 0.4$. Finally, to additionally enrich a sample of fat jets in gluon splitting to $\mathrm{b\bar{b}}$ component ($\mathrm{g \to b\bar{b}}$), used as an analogue of boosted $\mathrm{H \to b\bar{b}}$ jets, the fat jets can be required to be double-muon-tagged with both subjets matched to distinct muon candidates.

In order to compare different flavour compositions and different boosted event topologies, four different jet samples are used for the validation purposes:

- a sample of subjets of the `HEPTop`-tagged CA15 fat jets in a sample of semi-leptonic $\mathrm{t\bar{t}}$ events, as described above,
- a sample of inclusive CA8 fat jets,
- a sample of muon-tagged subjets of CA8 fat jets, and
- a sample of double-muon-tagged CA8 fat jets, as described above.

For these topologies, a set of quantities, which are relevant for b tagging, is displayed in Figs. 24 and 25. CA8 fat jets are labeled as b jets from gluon splitting ($\mathrm{g \to b\bar{b}}$) if at least two generator-

level b hadrons are found within $\Delta R < 0.8$ from the jet axis. However, for subjets of fat jets only the standard three flavour categories (udsg, c, and b) are used. A comparison with Section 3 shows that the agreement between data and simulation is at a similar level.

5.3 Performance measurements

5.3.1 Boosted top jets

The measurement of the b-tagging efficiency in boosted topologies is a particular challenge. The standard methods of efficiency measurement use non-boosted objects and their results are therefore not necessarily applicable to boosted topologies. A modified implementation of the "Flavor-tag Consistency Method" (FTC), described in Section 4.2.2, has been developed to measure the b-tagging efficiency in boosted top-quark events. In the following the FTC concept is applied to subjets of fat jets of the `HEPTopTagger` algorithm. This modified FTC requires consistency between the number of tags in the data and in the simulation for subjets in boosted top events. A log-likelihood fit is performed in which the b-tagging efficiency scale factor SF_b and the $t\bar{t}$ cross section are free parameters. The light-flavour misidentification scale factor SF_{light} is fixed to the standard value for regular anti-k_T jets with a size parameter of 0.5. A simultaneous fit of SF_{light}, SF_b and the $t\bar{t}$ cross section has also been performed. This method relies on Monte Carlo simulation for the flavor composition of the sample of subjets.

The stability and consistency of the FTC method in boosted top jets is tested in 200 toy experiments in which known values for SF_{light} and SF_b are injected. When fitting SF_b only, with SF_{light} fixed to the nominal value, no difference between the fitted and injected values are observed. The statistical uncertainty is about 1%. When fitting both SF_b and SF_{light}, the result for SF_b is at the same level of stability as when fitting SF_b alone. The result for SF_{light} is compatible with the injected values as well, but the statistical uncertainties are 5% and 30% for the loose and medium operating points of the CSV algorithm, respectively.

The systematic uncertainties for this measurement are:

- No-flavour subjets: to predict the number of b-tags in the simulation the true flavour of the jet needs to be known. The jet flavour identification fails in a small fraction of subjets. The impact on the final scale factor is less than 2%.
- Background normalisation: the relative normalisation of the background processes with respect to the signal can affect the measurement of SF_b. A conservative systematic uncertainty of 50% is assigned to the background normalisation. The impact on the final scale factor is less than 1%.
- Scale factor uncertainties: the scale factors for those flavours that are not left as free parameters of the fit are fixed to the values measured on regular anti-k_T jets in the non-boosted regime. The FTC fit with SF_{light} fixed has been repeated, changing SF_c and SF_{light} by $\pm 1\sigma$. The impact on the final scale factor is of about 1%, 1% and 3% for the loose, medium and tight working points of the CSV algorithm, respectively.

Figure 26 shows the result of the FTC fit when leaving both SF_b and SF_{light} as free parameters in the fit. A comparison of the measured SF_b with the result from the non-boosted analysis is given in Table 5.

The FTC method measures the average value of the scale factors for the considered sample. The dependency of the scale factors on the transverse momenta of the subjets is not provided by this technique. In order to exclude significant changes of the scale factors when moving to the boosted regime, the FTC method has been repeated for three different p_T regions of the top

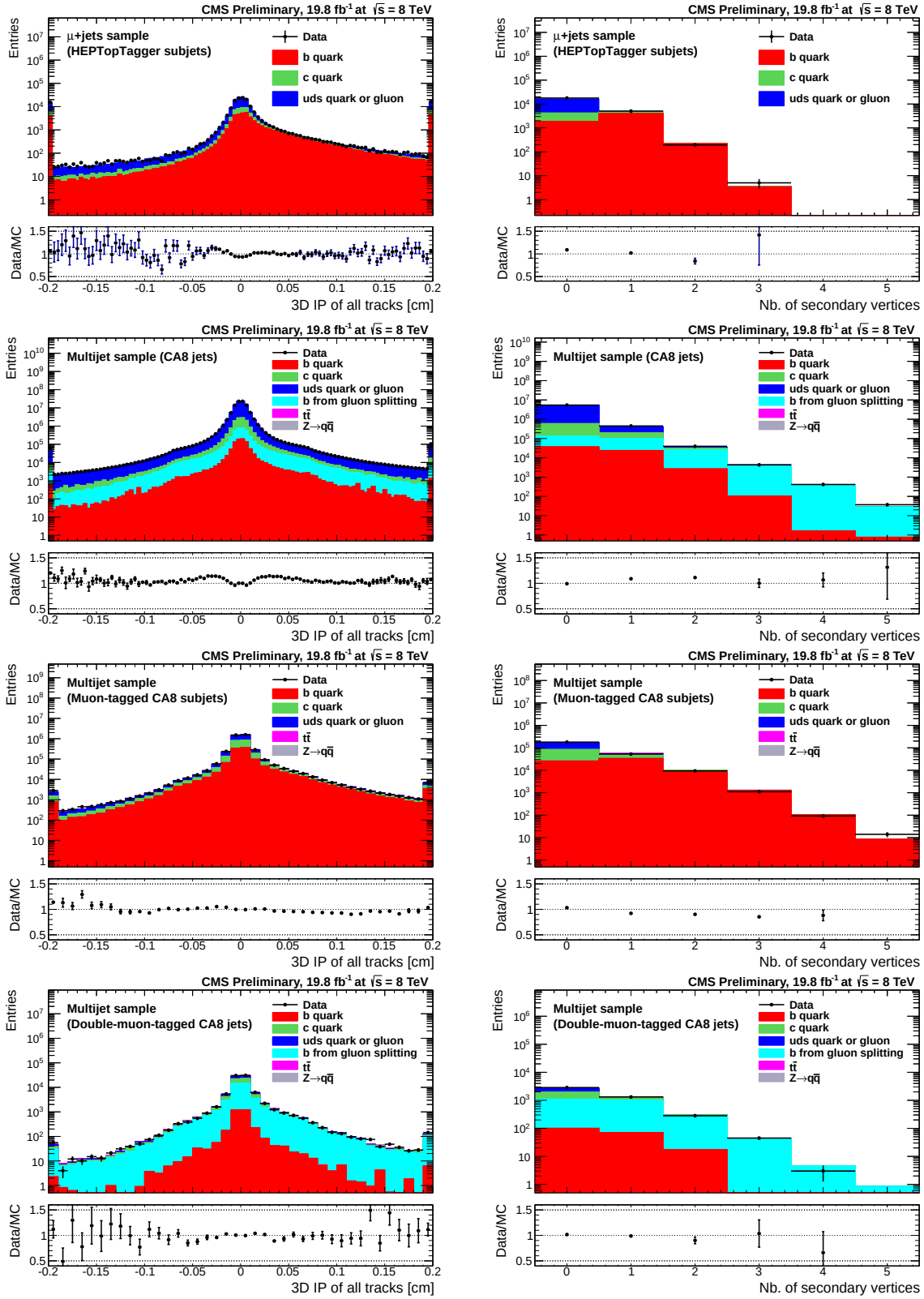


Figure 24: Left: 3D track impact parameter, Right: multiplicity of secondary vertices. First row: subjets of `HEPTopTagger` CA15 fat jets, second row: inclusive CA8 fat jets, third row: muon-tagged subjets of CA8 fat jets, fourth row: double-muon-tagged CA8 fat jets.

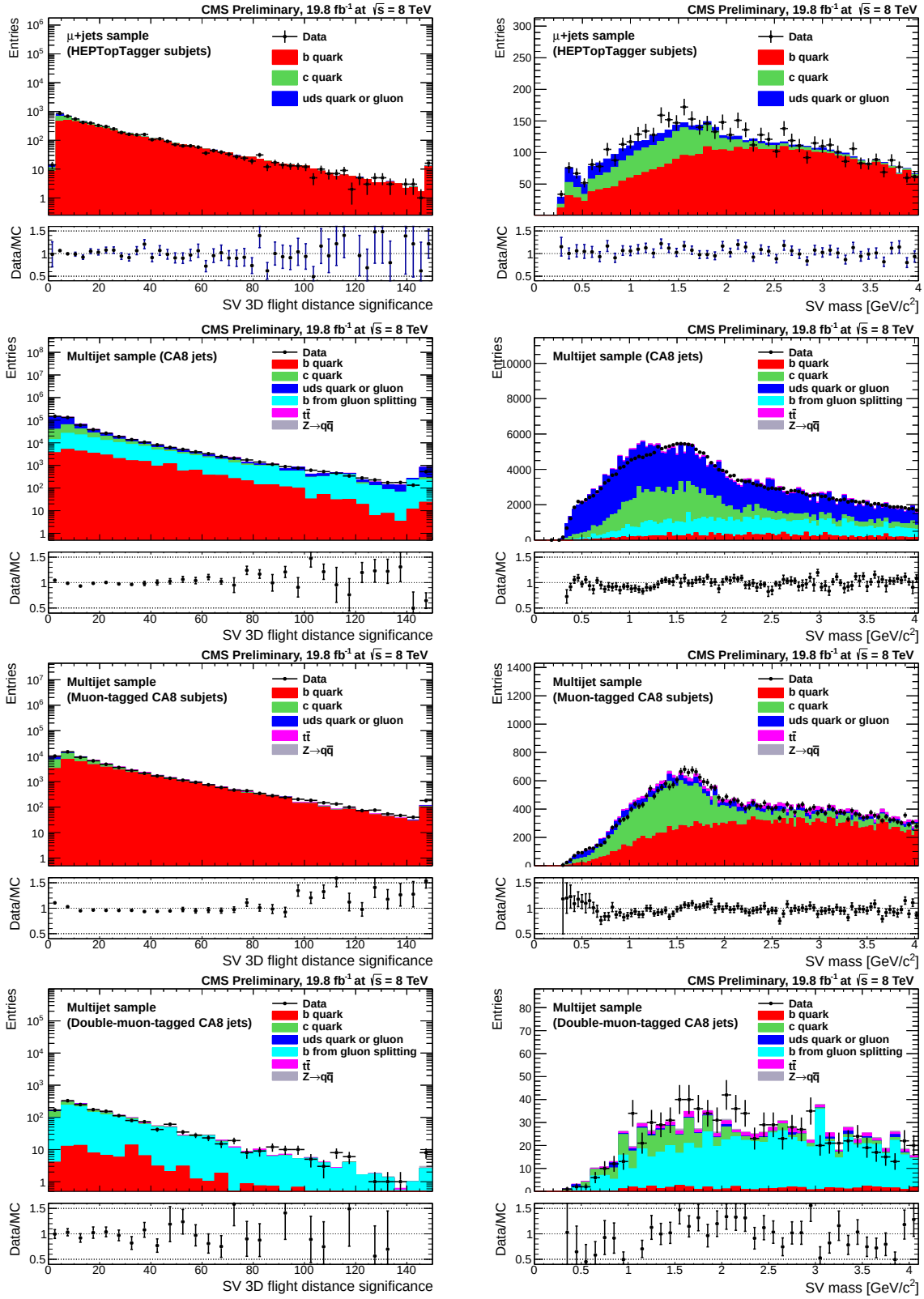
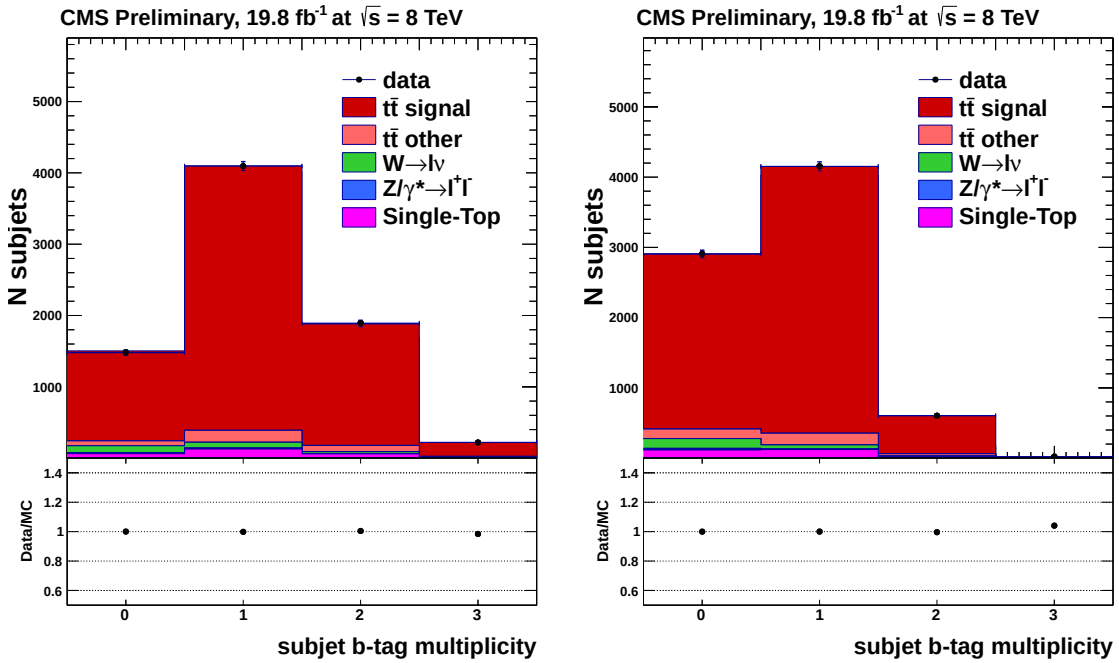


Figure 25: Left: secondary vertex flight distance significance, Right: secondary vertex mass. First row: subjets of `HEPTopTagger` CA15 fat jets, second row: inclusive CA8 fat jets, third row: muon-tagged subjets of CA8 fat jets, fourth row: double-muon-tagged CA8 fat jets.



(a) (b)

Figure 26: Distributions of the subjet b-tag multiplicities after the FTC fit. SF_b, SF_light and $\sigma_{\mathrm{t\bar{t}}}$ are free parameters in the fit; (left) results for CSVL, (right) results for CSVM.

Table 5: Comparison of the measured scale factors obtained for regular anti-k_T jets and for subjets of boosted top jets, which are reconstructed with the `HEPTopTagger`. The combined statistical and systematic uncertainties are quoted for both SF_b and SF_light. The measured values of SF_b for CSVM are also shown for different p_T regions of the top fat jet.

	CSVL	CSVM	CSVT
SF_b for non-boosted jets	1.010±0.013	0.970±0.013	0.950±0.015
SF_b for `HEPTopTagger` subjets	1.003±0.026	0.979±0.023	0.960±0.036
$150 \leq p_\mathrm{T} < 350$ GeV/c	—	$0.978^{+0.023}_{-0.023}$	—
$p_\mathrm{T} \geq 350$ GeV/c	—	$0.993^{+0.034}_{-0.034}$	—
$p_\mathrm{T} \geq 450$ GeV/c	—	$0.997^{+0.067}_{-0.067}$	—
SF_light for non-boosted jets	$1.080^{+0.063}_{-0.072}$	$1.136^{+0.090}_{-0.110}$	$1.088^{+0.039}_{-0.086}$
SF_light for `HEPTopTagger` subjets	1.185±0.080	1.58±0.47	—

fat jet: $150 \leq p_\mathrm{T} < 350$ GeV/c, $p_\mathrm{T} \geq 350$ GeV/c and $p_\mathrm{T} \geq 450$ GeV/c.

The measured values of SF_b for the three samples are shown in Table. 5. The values are statistically consistent and a significant deviation of the scale factors for an increased boost of the top quark is excluded. The selected semi-leptonic top-quark sample is not suitable for probing extremely boosted regimes, as the statistics of selected top fat jets with $p_\mathrm{T} \geq 500$ GeV/c is insufficient and does not allow the convergence of the FTC likelihood fit.

We conclude that the simulation reproduces the b-tagging efficiencies in data equally well in boosted and in non-boosted top-quark events. For the first time, the misidentification probability has also been measured in top-quark events, albeit with larger uncertainties than with the standard methods. Also the misidentification probability is in good agreement between boosted and non-boosted topologies.

5.3.2 Boosted Higgs jets

Since boosted topologies involve high-p_{T} jets and a wide range of subjet transverse momenta, a method suitable for measurements of the b-tagging efficiency in such topologies is the LT method (see Section 4). In the following the LT method is applied to the pruned subjets of CA8 fat jets. The subjets are required to be muon-tagged and only those with $p_{\mathrm{T}} > 40$ GeV/c are considered. The resulting fits to the JP discriminator distributions for subjets with $160 < p_{\mathrm{T}} < 210$ GeV/c, before and after btagging with the CSVL tagger, are shown in Fig. 27. In addition, the LT method is applied to a subset of muon-tagged subjets where the companion subjet is b-tagged with the CSVL tagger. This type of subjet configuration is enriched in b jets from gluon splitting and therefore closer to the configurations expected for boosted $\mathrm{H} \to \mathrm{b\bar{b}}$ jets. The measured b-tagging efficiencies and data/MC scale factors in different subjet p_{T} bins are shown in Fig. 28 and compared with the corresponding measurements for the standard AK5 jets.

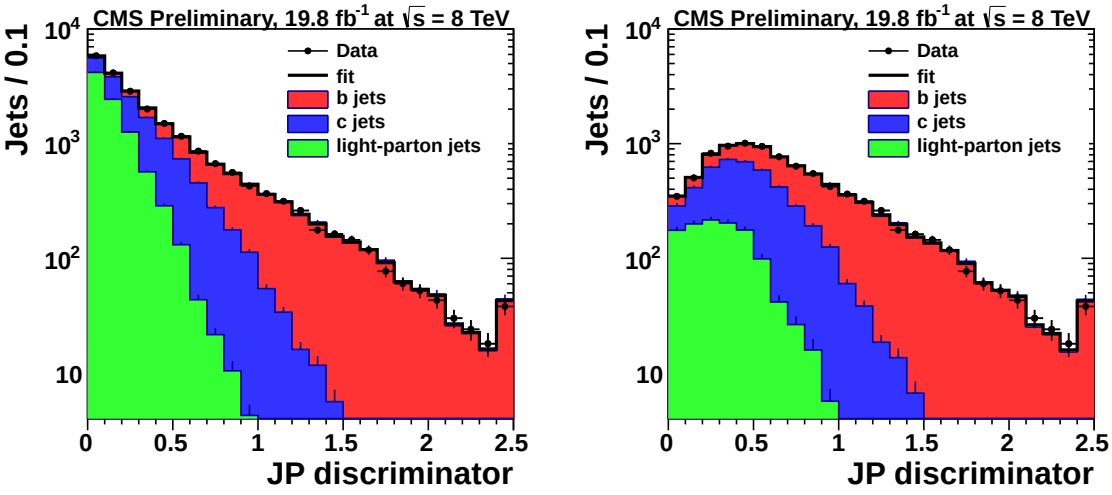


Figure 27: Fits to the JP discriminator distributions for muon-tagged subjets of CA8 fat jets with subjet $160 < p_{\mathrm{T}} < 210$ GeV/c (left) before and (right) after btagging with the CSVL tagger. The overflow entries are shown in the rightmost bin.

Similarly to the LT method, a misidentification probability measurement based on the negative tags can be adapted to boosted topologies and applied to the subjets of fat jets. This allows a direct measurement of the misidentification probability and corresponding scale factors for subjets. The resulting misidentification probability and corresponding scale factors for the CSVL tagger measured on subjets of fat jets using the negative tags are shown in Fig. 29.

Overall, the scale factors measured in boosted topologies are found to be in good agreement with those measured in the standard, non-boosted topologies. This confirms observations from studies with boosted top jets, which indicate that the simulation reproduces the b-tagging performance in boosted and non-boosted event topologies equally well. It should be noted, however, that a detailed study of systematic uncertainties performed in non-boosted topologies has not been repeated for boosted topologies.

In addition to the above-described dedicated scale-factor measurements, a crosscheck is performed where the standard scale factors derived for AK5 jets are applied to a sample of double-muon-tagged jets in order to validate their suitability for topologies with boosted Higgs bosons. The validation procedure is performed as follows. A sample of double-muon-tagged fat jets in

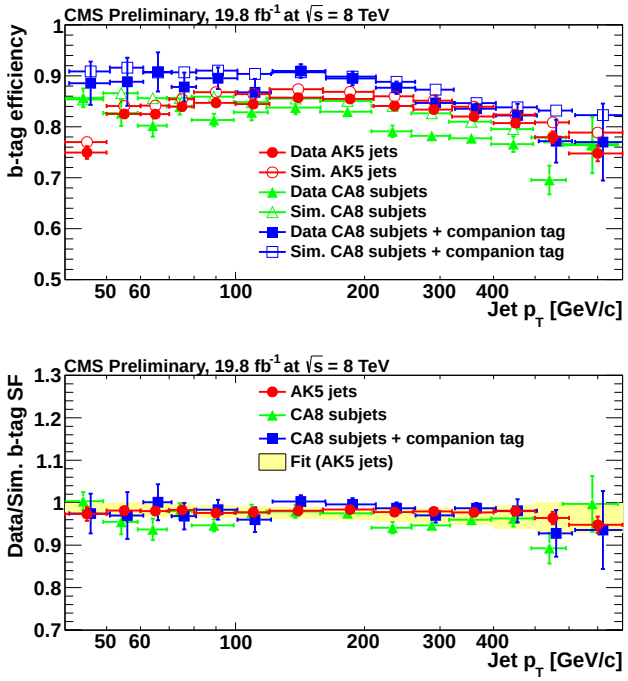


Figure 28: (top panel) b-tagging efficiency and (bottom panel) data/MC scale factors for muon-tagged subjets of CA8 fat jets, with and without the CSVL tag applied to the companion subjet, compared to the standard measurements for AK5 jets. In both cases the measurement is performed using the LT method. The error bars indicate the statistical uncertainties only. The yellow band shows the combined p_{T}-dependent data/MC scale factors for AK5 jets, along with corresponding combined statistical and systematic uncertainties.

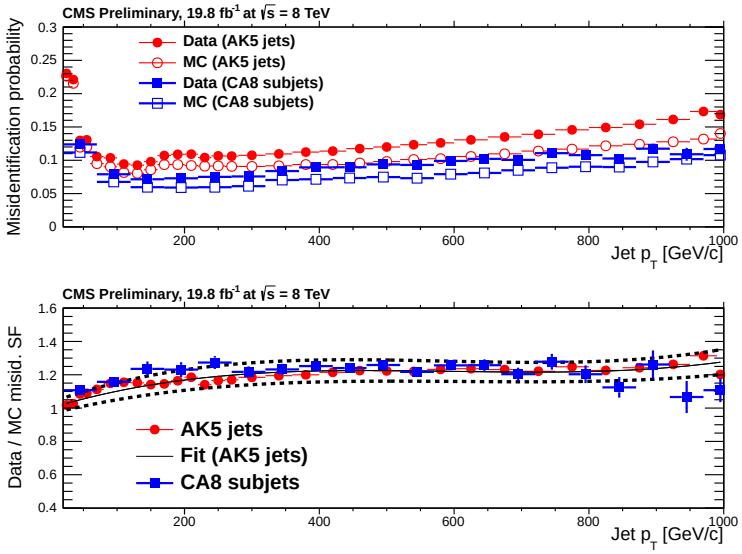


Figure 29: (top panel) Misidentification probability in data and MC and (bottom panel) corresponding scale factors for subjets of CA8 fat jets compared with the standard measurements for AK5 jets. In both cases the measurement is performed for the CSVL tagger using the negative tags. A fit to the scale factors for AK5 jets (solid curve) and the corresponding overall statistical and systematic uncertainties (represented by the dashed curves) are also shown.

simulation is first normalised to the data. This is performed by normalising the $t\bar{t}$ and $Z \to q\bar{q}$ components to 19.8 fb^{-1} of data using the next-to-next-to-leading-order $t\bar{t}$ [28] and Z [29, 30] cross sections and rescaling the remaining QCD component so that the total matches to the data. At this stage the size of the $t\bar{t}$ and $Z \to q\bar{q}$ contributions is ∼5% and ∼1%, respectively, with the overall data/MC ratio equal to unity by construction. At the next stage, fat jets are required to be double-b-tagged with both pruned subjets CSVL-tagged. The QCD normalisation is kept from the previous, double-muon-tagged stage. The size of the $t\bar{t}$ and $Z \to q\bar{q}$ contributions at this stage is ∼3% and ∼2%, respectively. The resulting data/MC ratio is equal to $0.94 \pm 0.03\,(\text{stat.})$, i.e., it is no longer consistent with unity. After applying the scale factors the overall data/MC ratio is equal to $0.98 \pm 0.03\,(\text{stat.})^{+0.04}_{-0.05}\,(\text{syst.})$, i.e., it is consistent with unity within both statistical and systematic uncertainties. A distribution of ΔR separation between the subjets of fat jets for double-muon- and double-CSVL-tagged fat jets, before and after applying the b-tagging scale factors, are shown in Fig. 30. The same exercise repeated with CSVM-tagged subjets results in an overall data/MC ratio of $0.91 \pm 0.04\,(\text{stat.})$ before and $0.98 \pm 0.04\,(\text{stat.})^{+0.06}_{-0.05}\,(\text{syst.})$ after applying the scale factors.

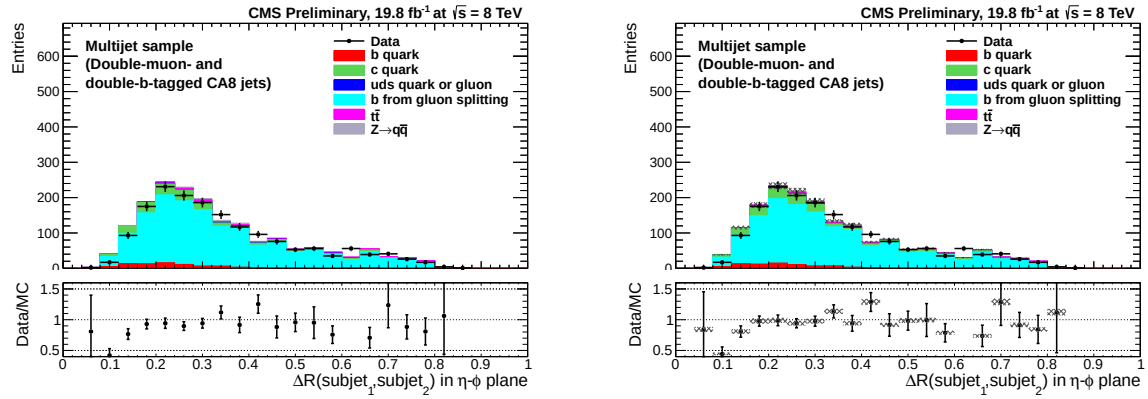


Figure 30: $\Delta R(\text{subjet}_1, \text{subjet}_2)$ distribution for double-muon- and double-CSVL-tagged fat jets (left) before and (right) after applying the b-tagging scale factors. The hatched areas represent the systematic uncertainty due to variations of the scale factors within 1σ of their uncertainties. Only the heavy-flavour scale factors SF_b are varied since the fat jet sample is dominated by heavy-flavour jets.

A potential concern for applying the scale factors are configurations in which subjets of fat jets are close to one another. Because of track sharing and possible correlation between the b-tagging probabilities for the two subjets, a straightforward application of the scale factors, which treats individual subjets as independent objects, may lead to an overcorrection of the simulated events. Since the scale factors for CSVL- and CSVM-tagged jets are less than one, the overcorrection would result in an overall data/MC ratio greater than one. The control sample of double-muon-tagged fat jets used for the scale-factors validation is dominated by configurations with nearby subjets, as can be seen in Fig. 30, and has an average fraction of shared tracks of ∼40%. Given that no overcorrection in the used control sample is observed, the level of correlation between the b-tagging probabilities for the two subjets with a fraction of shared tracks of ∼40% is still relatively moderate. Translated to the boosted Higgs jets, this implies that applying the scale factors should be safe up to a fat jet p_T of at least ∼700 GeV/c, possibly higher.

6 Conclusions

Several methods have been applied to measure the identification efficiency of b-quark jets. The results are used to determine data-to-simulation scale factors for the b-tagging efficiency, SF_b. Results based on jets with a semimuonic b-hadron decay, events with b decay to J/ψ, and $\mathrm{t\bar{t}}$ events have been compared using data collected in 2012 with an integrated luminosity of 19.8 fb^{-1}. The results obtained by these different methods are found to be consistent with each other. A combination is presented as a function of the b-quark jet transverse momentum ranging from 20 up to 800 GeV/c. The combined scale factors range from about 0.87 to 0.98. For jet p_T within 80-120 GeV, the achieved relative precision on SF_b is 2-4% and it is about 5-9% at the highest considered jet transverse momenta. Comparisons are performed between the methods relying on multijet events and on top-quark pair events in the p_T range of b jets from top quark decays: a fairly good agreement is obtained.

The scale factor for the misidentification probability from light-flavour quark (u, d, s) and gluon jets, SF_{light}, has been measured as a function of the jet p_T between 20 and 1000 GeV/c. For each b-tagging algorithm, the scale factors are compatible with unity within 30%. From the loose to tight operating points, the SF_{light} values are measured with a precision of about 5-10 to 30%, respectively.

Overall, the b-tagging performance measurements at $\sqrt{s} = 8$ TeV are in good agreement with those observed at 7 TeV [4]. The increased integrated luminosity and higher pp collision energy allow the measurements to be extended to both lower and higher transverse momenta.

For the first time at CMS, results of a study of b tagging in boosted topologies have been presented. The topologies studied are those of boosted top quarks and Higgs bosons which are imparted large enough momentum that their decay products end up clustered in a single fat jet. Studies of the performance of b-tagging algorithms applied to fat jets and their subjets are presented. Dedicated studies using suitably defined control samples have been performed to validate the agreement of the data with the simulation and to measure the performance of b tagging in boosted environments. A good agreement is found between measurements performed in boosted topologies with those performed in the standard, non-boosted topologies. These are the first systematic studies of b tagging in boosted topologies performed in the CMS experiment.

References

[1] T. Plehn and M. Spannowsky, "Top Tagging", *J. Phys. G* **39** (2012) 083001, doi:10.1088/0954-3899/39/8/083001, arXiv:1112.4441.

[2] A. Altheimer et al., "Jet Substructure at the Tevatron and LHC: New results, new tools, new benchmarks", *J. Phys. G* **39** (2012) 063001, doi:10.1088/0954-3899/39/6/063001, arXiv:1201.0008.

[3] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

[4] CMS Collaboration, "Identification of b-quark jets with the CMS experiment", *JINST* **8** (2013) P04013, doi:10.1088/1748-0221/8/04/P04013.

[5] T. Sjöstrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 physics and manual", *JHEP* **05** (2006) 026, doi:10.1088/1126-6708/2006/05/026, arXiv:hep-ph/0603175.

[6] R. Field, "Early LHC underlying event data - findings and surprises", (2010). arXiv:1010.3558.

[7] J. Alwall et al., "MadGraph 5: going beyond", *JHEP* **06** (2011) 128, doi:10.1007/JHEP06(2011)128, arXiv:1106.0522.

[8] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with parton shower simulations: the POWHEG method", *JHEP* **11** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.

[9] K. Melnikov and F. Petriello, "Electroweak gauge boson production at hadron colliders through $\mathcal{O}(\alpha_s^2)$", *Phys. Rev. D* **74** (2006) 114017, doi:10.1103/PhysRevD.74.114017, arXiv:hep-ph/0609070.

[10] J. M. Campbell and R. K. Ellis, "MCFM for the Tevatron and the LHC", *Nucl. Phys. Proc. Suppl.* **205-206** (2010) 10, doi:10.1016/j.nuclphysbps.2010.08.011, arXiv:1007.3492.

[11] N. Kidonakis, "NNLL threshold resummation for top-pair and single-top production", (2012). arXiv:1210.7813.

[12] GEANT4 Collaboration, "GEANT4—a simulation toolkit", *Nucl. Instrum. Meth. A* **506** (2003) 250, doi:10.1016/S0168-9002(03)01368-8.

[13] CMS Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and Missing E_T", CMS Physics Analysis Summary CMS-PAS-PFT-09-001, (2009).

[14] CMS Collaboration, "Commissioning of the Particle-Flow Reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV", CMS Physics Analysis Summary CMS-PAS-PFT-10-002, (2010).

[15] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm", *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.

[16] CMS Collaboration, "Determination of jet energy calibration and transverse momentum resolution in CMS", *JINST* **06** (2011) P11002, doi:10.1088/1748-0221/6/11/P11002, arXiv:1107.4277.

[17] CMS Collaboration, "CMS Tracking Performance Results from Early LHC Operation", *Eur. Phys. J. C* **70** (Jul, 2010) 1165. 29 p.

[18] K. Rose, "Deterministic Annealing for Clustering, Compression, Classification, Regression and related Optimisation Problems", *Proceedings of the IEEE* **Vol. 86, Issue 11** (1998).

[19] W. Waltenberger, R. Frühwirth, and P. Vanlaer, "Adaptive Vertex Fitting", *J. Phys. G: Nucl. Part. Phys.* **34** (2007), no. 12, N343, doi:10.1088/0954-3899/34/12/N01.

[20] CMS Collaboration, "Inclusive b-jet production in pp collisions at $\sqrt{s} = 7$ TeV", *JHEP* **04** (2012) 084, doi:10.1007/JHEP04(2012)084, arXiv:1202.4617.

[21] CLEO Collaboration, "Charged track multiplicity in B meson decay", *Phys. Rev. D* **61** (2000) 072002, doi:10.1103/PhysRevD.61.072002, arXiv:hep-ex/9907057.

[22] CMS Collaboration, "Measurement of Tracking Efficiency", CMS Physics Analysis Summary CMS-PAS-TRK-10-002, (2010).

[23] L. Lyons, D. Gibaut, and P. Clifford, "How to combine correlated estimates of a single physical quantity", *Nucl. Instrum. Meth. A* **270** (1988) 110, doi:10.1016/0168-9002(88)90018-6.

[24] M. Wobisch and T. Wengler, "Hadronization corrections to jet cross-sections in deep inelastic scattering", (1998). arXiv:hep-ph/9907280.

[25] Y. L. Dokshitzer, G. Leder, S. Moretti, and B. Webber, "Better jet clustering algorithms", *JHEP* **08** (1997) 001, doi:10.1088/1126-6708/1997/08/001, arXiv:hep-ph/9707323.

[26] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, "Techniques for improved heavy particle searches with jet substructure", *Phys. Rev. D* **80** (2009) 051501, doi:10.1103/PhysRevD.80.051501, arXiv:0903.5081.

[27] M. Dasgupta, A. Fregoso, S. Marzani, and G. P. Salam, "Towards an understanding of jet substructure", (2013). arXiv:1307.0007.

[28] M. Czakon, P. Fiedler, and A. Mitov, "The total top quark pair production cross-section at hadron colliders through $\mathcal{O}(\alpha_S^4)$", (2013). arXiv:1303.6254.

[29] K. Melnikov and F. Petriello, "Electroweak gauge boson production at hadron colliders through $\mathcal{O}(\alpha_s^2)$", *Phys. Rev. D* **74** (2006) 114017, doi:10.1103/PhysRevD.74.114017, arXiv:hep-ph/0609070.

[30] R. Gavin, Y. Li, F. Petriello, and S. Quackenbush, "FEWZ 2.0: A code for hadronic Z production at next-to-next-to-leading order", *Comput. Phys. Commun.* **182** (2011) 2388–2403, doi:10.1016/j.cpc.2011.06.008, arXiv:1011.3540.