Measurement of the b-jet to inclusive jet ratio in PbPb and pp collisions at $\sqrt{s_{NN}} = 2.76$ TeV with the CMS detector

The CMS Collaboration

Abstract

Modification to jets in high-energy heavy-ion collisions is expected to depend on the flavor of the fragmenting parton. To disentangle this flavor dependence, jets from heavy quark fragmentation are identified for the first time in heavy ion collisions. Jets are first tagged by their secondary vertices and the contribution from bottom quarks is extracted using template fits to their secondary vertex mass distributions. The bottom quark jet to inclusive jet ratio is measured with the CMS detector from PbPb and pp collisions at a center-of-mass energy of 2.76 TeV per nucleon. This b-jet fraction is measured in the range of $80 < \text{jet } p_T < 200$ GeV/c in PbPb collisions and found to lie in the range of $2.9–3.5\% \pm 0.6–1.1\%$ depending on jet $p_T$. The measured values are comparable to those predicted by PYTHIA simulation. The PbPb b-jet fraction is also compatible with the pp b-jet fraction, within sizeable uncertainties. The measurement is sufficiently precise to demonstrate that b-jets are subject to jet quenching, although a precise comparison of light and b-jet quenching would require a reduction of the statistical and systematic uncertainties.
1 Introduction

By colliding heavy nuclei at the LHC, one expects to reach sufficiently large energy densities to form a quark-gluon plasma (QGP) which is characterized by the effective deconfinement of the color degrees of freedom. Hard-scattered partons suffer energy loss as they transverse the QGP via elastic and inelastic interactions [1]. This phenomenon, known as “jet quenching”, is one of the experimental signatures of the QGP formation. The energy loss measurement is expected to reveal fundamental thermodynamical and transport properties of this phase of matter (see [2, 3] for recent reviews).

The quenching of jets in heavy-ion collisions is expected to depend on the flavor of the fragmenting parton. For example, under the assumption that radiative energy loss is the dominant mechanism, gluon jets should be quenched more strongly than light quark jets, due to the larger color factor for gluon emission from gluons than from quarks. On the other hand, jets initiated by heavy quarks, particularly bottom, are expected to radiate less than light ones. This is due to the so-called dead cone effect whereby radiation is suppressed in the direction of propagation as a result of coherence effects [4].

Identification of reconstructed b-jets has so far not been performed in heavy-ion collisions. However, recent CMS data demonstrates that non-prompt J/ψ, i.e., those coming from decay of B mesons, are indeed suppressed in PbPb collisions with respect to the pp expectation [5]. This indirect measurement, which samples typical transverse momentum of J/ψ meson values on the order of 10 GeV/c, provides a strong motivation to perform a more direct measurement using fully reconstructed jets. Such a measurement would enable a direct comparison of b quark energy loss to that of inclusive jets at much larger values of jet $p_T$, where the flavor dependence of parton energy loss can be probed in detail.

Jets formed from heavy flavor quark fragmentation can be tagged by the presence of displaced vertices, either by direct reconstruction of these vertices or by the impact parameter (i.e., the distance of closest approach to the primary vertex) of tracks originating from these vertices [6]. Information from these tracks and vertices are typically combined into a quantity which optimizes their discrimination between heavy and light flavor jets. In this analysis, we use a discriminator to tag b-jets which is based on the flight distance of the reconstructed secondary vertex (SV) with respect to the primary vertex of the interaction. The efficiency of this SV tagging is then evaluated both directly from simulation, and with a data-driven method using a weakly correlated tagger based on the impact parameter of tracks associated to the jets.

2 Event & Jet Selection

This analysis is based on 150 $\mu$b$^{-1}$ of PbPb collisions at $\sqrt{s_{NN}} = 2.76$ TeV recorded with the CMS detector in 2011. A detailed description of the CMS experiment can be found in [7]. Control data from 231 nb$^{-1}$ of pp collisions at $\sqrt{s} = 2.76$ TeV are also analyzed to benchmark possible nuclear effects. The event selection for this analysis is the same as in [8]. Events are selected using high-level jet triggers. Offline selections are applied to remove backgrounds from beam-gas interactions and anomalous detector artifacts. The collision centrality is related to the degree of overlap of the incoming nuclei. Centrality classes are defined as percentiles of the total inelastic cross section using the energy deposited in the forward hadron calorimeters. Simulated data are generated with PYTHIA Tune Z2 [9] and embedded into PbPb events generated by HYDJET [10] version 1.8.

Jets are reconstructed from particle flow objects using the anti-$k_T$ algorithm with a radius pa-
rameter of $R = 0.3$ [11, 12]. The background from the underlying heavy ion event is subtracted differentially in $\eta$, event-by-event, according to the method described in [13]. Jets are required to have $p_T > 80$ GeV/$c$ as well as $|\eta| < 2$, which insures that they fragment mostly into the tracker acceptance ($|\eta| < 2.4$).

3 Data Analysis

The performance of lifetime-based tagging relies on the high efficiency and low fake rate of reconstruction of charge particle tracks from displaced vertices. The standard heavy-ion tracking algorithm [14] in CMS is largely restricted to the reconstruction of charged particles from the primary vertex. To enhance the efficiency of tracks from secondary vertices, additional track reconstruction is performed. Secondary track reconstruction in central heavy-ion events is extremely resource intensive, due to the large number of possible hit combinations when incorporating hits over a sufficiently large region to find such displaced tracks. To mitigate the number of possible track candidates, the additional track reconstruction uses reconstructed jets as seeds and limits the search window for tracker hits to a region defined around the jet axis.

Identification of b-jets is achieved using a discriminating variable, denoted Simple Secondary Vertex High Efficiency (SSVHE) [6], which is based on the flight distance significance (flight distance divided by its uncertainty) of reconstructed secondary vertices. Detailed studies of quantities related to SV reconstruction show that it is well modeled by PbPb simulations, lending confidence to the use of this discriminator to identify b-jets. Another discriminating variable, called Jet Probability (JP), is also used in this analysis. The JP algorithm orders jet-associated tracks based on their impact parameter significance (impact parameter divided by its uncertainty) and calculates the likelihood that they come from the same primary vertex [6]. The JP algorithm provides a means of discrimination for nearly all b-jets, even in the case when no SV is reconstructed. This property of the JP tagger is exploited to obtain a data-driven estimate of the SSVHE tagging efficiency [6]. This is done by evaluating the fraction of jets from bottom fragmentation with and without the SSVHE tagging requirement, fitting the JP discriminator distribution with the same template fitting method described below.

The performance of the discriminators is typically benchmarked by plotting the b-jet tagging efficiency vs. the light and charm jet mis-tag efficiencies, such that the resulting curves are independent of the underlying b-jet fraction. Fig. 1 shows the b-jet tagging efficiency vs. the light jet and charm jet mis-tag efficiencies from PYTHIA and PYTHIA+HYDJET. No centrality selection is applied to the PYTHIA+HYDJET sample, but its centrality distribution is weighted to match the distribution in data after the jet selection. The two taggers used in this analysis, SSVHE and JP, are shown. The performance of the b-tagging degrades somewhat with the increased multiplicity of PbPb collisions, giving roughly a factor of 3 poorer rejection of light jets for a b-jet efficiency of 50%, relative to PYTHIA alone. The effect on the charm rejection is somewhat more modest, particularly for the SSVHE tagger. Despite the reduced performance, one is still able to achieve roughly a factor of 100 (10) rejection of light (charm) jets for a b-jet efficiency around 50%.

The b-jet fraction is extracted from the data using the SV mass. The SV mass, which is simply the invariant mass of the charged tracks used in the SV reconstruction, provides good separation between the charm and bottom contributions. Unbinned maximum likelihood fits are then performed where the b-jet and non-b-jet shapes are fixed by the simulation, but their normalizations are allowed to float. The four panels of Fig. 2 show the SV mass distributions in data and corresponding simulation templates for the jet $p_T$ bins used in this analysis. The ratio of the charm to light normalizations are fixed in the template fits according to their relative
Figure 1: The b-jet tagging efficiency vs. the light jet (top) and charm jet (bottom) mis-tag efficiencies for simulated pp events from PYTHIA (left) and simulated PbPb events from PYTHIA embedded in HYDJET (right) for the SSVHE and JP discriminators. The red cross marks the working point of the SSVHE discriminator used in this analysis.

The b-jet purity is defined to be the fraction of jets from bottom quarks in the SSVHE-tagged sample. The left panel of Fig. 3 shows the b-jet purity of the tagged sample as a function of the jet $p_T$ extracted from the template fits in Fig. 2. These purity values are close to those from the input simulation. The SSVHE tagging efficiency is shown in the right panel of Fig. 3. The result obtained using JP as a reference tagger is compared to the result obtained directly from simulation. The difference, which is found to be within about 5%, is included in the systematic uncertainties, as discussed in the next section.

4 Corrections and Uncertainties

There are several sources of systematic uncertainty which affect b-jet to inclusive jet ratio. The uncertainty assigned to the efficiency of the SSVHE tagger is taken from the difference between the simulation and data-driven efficiency determinations. The central value is taken from the data-driven efficiency.
Figure 2: Template fits to the SV mass distributions in PbPb collisions, after tagging with the SSVHE discriminator. Several $p_T$ ranges are shown as indicated on the figures. The colored lines represent the statistical uncertainties on the MC templates.

A second category of uncertainty is composed of effects which may impact the estimate of the b-tagging purity. One such uncertainty is the relative contribution from charm and light jets, which changes the shape of the non-b-jet template. The relative normalization is fixed in our template fits, and then varied to obtain a systematic uncertainty. For PbPb collisions we perform a second set of secondary vertex mass fits in which the charm and light normalizations are allowed to float independently, and take the difference between the floating and fixed normalizations as a contribution to the uncertainties. For pp collisions we instead vary the charm contribution by 20%, which is consistent with the procedure employed in pp collisions at 7 TeV [15].

To determine the uncertainty on the modeling of the non-b-jet template shapes by the simulation, we compared to results obtained with a data-driven template. This template was formed by requiring an anti-b-tag of $|P_\tau| < 0.5$ in data. An additional source of uncertainty on the template shapes was estimated by varying the working point of the SSVHE tagger. This was done with both the simulation and data-derived non-b-jet templates and the maximum deviation was taken as the error.

The difference between the b-jet and inclusive jet energy scale (JES) is corrected for based on
simulated. The response of b-jets tends to be lower than that of inclusive jets for two reasons. First, b-jets are more likely to contain neutrinos, which are included in our generator-level jet definition. Second, jets are reconstructed prior to the jet-seeded track reconstruction, which has the consequence that secondary tracks are not included in the jet reconstruction. Instead, the contribution of the particles that give rise to these tracks is based solely on their calorimeter deposits, which tend to have a lower response than the combined track and calorimeter information in the particle flow reconstruction. While an overall shift in the JES will cancel in the b-jet to inclusive jet ratio, a possible residual difference between the b-jet and inclusive jet JES is considered as a source of uncertainty. For the pp data, we conservatively quote the entire JES uncertainty of 1.5% for this difference as was done for the 7 TeV analysis [15]. For the PbPb data, the overall variation of the JES with centrality of 2% is added in quadrature. These uncertainties propagate to a relative uncertainty of about 7% and 14%, on the pp and PbPb b-jet fractions, respectively.

Several production mechanisms contribute to heavy quark production. The splitting of a gluon into a $q-ar{q}$ pair tends to give rise to jets with small opening angle and may also give rise a single reconstructed jet. The b-tagging performance may be slightly different for such jets and the relative contribution of this next-to-leading-order process may not be perfectly modeled in PYTHIA. To test this we varied the contribution of gluon splitting by $\pm 50\%$. The effect was found to be negligible compared to other sources of systematic uncertainty.

The systematic uncertainties, which are in the range of roughly 20-40%, are compiled in Table 1. It should be noted that the statistical uncertainty plays a large role in the size of the systematic errors, which are derived bin-by-bin. The stability of the fits also varies from one bin to the next, which affects the size of the uncertainties.

The jet triggers used in this analysis trigger on jets above a $p_T$ threshold of 80 and 40 GeV/c for PbPb and pp, respectively, where the jets are reconstructed from calorimeter towers with a cone algorithm with a radius $R = 0.5$. The efficiencies of both triggers are close to unity for the jets selected in this analysis. The trigger efficiency have a non-negligible effect on the results only for the jets of $80 < p_T < 100$ GeV/c in the PbPb sample. For 0-100% centrality, the trigger efficiency in this $p_T$ bin is found to be 91.9% for inclusive jets and 95.8% for b-jets in simulation.
Table 1: Relative systematic uncertainties on the b-jet to inclusive jet ratio

<table>
<thead>
<tr>
<th>$p_T$ (GeV/c)</th>
<th>Tagging eff.</th>
<th>udsg;c norm.</th>
<th>Non-b shape</th>
<th>Working point</th>
<th>JES</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>PbPb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80-100</td>
<td>5%</td>
<td>3%</td>
<td>2%</td>
<td>11%</td>
<td>14%</td>
<td>19%</td>
</tr>
<tr>
<td>100-120</td>
<td>9%</td>
<td>17%</td>
<td>8%</td>
<td>25%</td>
<td>14%</td>
<td>35%</td>
</tr>
<tr>
<td>120-150</td>
<td>8%</td>
<td>11%</td>
<td>12%</td>
<td>15%</td>
<td>14%</td>
<td>27%</td>
</tr>
<tr>
<td>150-200</td>
<td>4%</td>
<td>15%</td>
<td>6%</td>
<td>20%</td>
<td>14%</td>
<td>30%</td>
</tr>
<tr>
<td>pp</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80-100</td>
<td>12%</td>
<td>7%</td>
<td>36%</td>
<td>18%</td>
<td>7%</td>
<td>43%</td>
</tr>
<tr>
<td>100-120</td>
<td>16%</td>
<td>6%</td>
<td>11%</td>
<td>18%</td>
<td>7%</td>
<td>28%</td>
</tr>
<tr>
<td>120-200</td>
<td>25%</td>
<td>5%</td>
<td>6%</td>
<td>22%</td>
<td>7%</td>
<td>35%</td>
</tr>
</tbody>
</table>

The small difference between the trigger efficiencies is applied as a correction.

5 Results and Conclusions

The b-jet to inclusive jet ratio is calculated from $(N_{\text{tagged}}/N_{\text{jets}})(P/\epsilon)$, where $N_{\text{tagged}}$ and $N_{\text{jets}}$ are the number of tagged jets and the total number of jets counted from the data, respectively, and $P$ and $\epsilon$ correspond to the purity and efficiency plotted in the left and right panels of Fig. 3, respectively. The left panel of Fig. 4 shows this ratio as function of jet $p_T$ in 0-100% PbPb centrality collisions. The b-jet fraction is around 2.9–3.5% with no significant $p_T$ dependence, with an absolute uncertainty in the range of 0.6–1.1%, increasing with jet $p_T$. The values predicted by PYTHIA+HYDJET, which decrease from about 3% to 2.5% as a function of $p_T$, are consistent with the data. The b-jet fraction in pp collisions as a function of jet $p_T$ is shown in the right panel of Fig. 4. The values predicted by PYTHIA are consistent with the data.

The left panel in Fig. 5 shows the b-jet fraction as a function of PbPb collision centrality for $80 < \text{jet } p_T < 120 \text{ GeV/c}$. No dependence is observed within the uncertainties. The ratio of the b-jet fraction in 0-100% PbPb to pp collisions is shown in the right panel of Fig. 5. The

Figure 4: The b-jet to inclusive jet ratio in 0-100% PbPb collisions (left) and pp collisions (right) as a function of jet $p_T$ compared to PYTHIA embedded in HYDJET (PbPb) and PYTHIA (pp). Data and MC have not been corrected for bin migration effects from finite jet resolution.
Figure 5: The b-jet to inclusive jet ratio for \(80 < \text{jet } p_T < 100 \text{ GeV/c}\) as a function of PbPb collision centrality (left) and the ratio of the b-jet fraction in PbPb to the b-jet fraction in pp as a function of jet \(p_T\) (right). Data and MC have not been corrected for bin migration effects from finite jet resolution.

Systematic uncertainties in the two measurements were treated as uncorrelated, except for the 1.5% uncertainty on the jet energy scale which is common to the two measurements. The remaining uncertainty on the PbPb energy scale translates to a 10% relative error. An additional 1% uncertainty is attributed to the difference between the pp and PbPb jet resolutions, which has a very small effect due to weak dependence of the b-jet fraction on jet \(p_T\). The ratio of the b-jet fractions is about 1.6 in the lowest \(p_T\) bin, but with a very large uncertainty. Above 100 GeV/c the double ratio is consistent with unity and constrained to be less than about 1.5 at the 1 \(\sigma\) level.

For inclusive jets the nuclear modification factor, \(R_{AA}\), has been measured to be \(0.50 \pm 0.01\) (stat.) \(\pm 0.06\) (syst.), for \(100 < \text{jet } p_T < 120 \text{ GeV/c}\) [16], which indicates that the PbPb yield is suppressed by about a factor two compared to the pp expectation. For consistency with this analysis in which no unfolding is performed, the inclusive jet \(R_{AA}\) is quoted with no correction for the PbPb jet resolution. Instead, a smearing procedure is applied to the jets from pp data to match the resolution in PbPb. The b-jet \(R_{AA}\) is the product of the ratio of b-jet fraction in PbPb and pp, and the inclusive jet \(R_{AA}\) and has a value of \(0.48 \pm 0.09\) (stat.) \(\pm 0.18\) (syst.), for \(100 < \text{jet } p_T < 120 \text{ GeV/c}\).

Taken as a whole, the data disfavor an extreme scenario in which b-jets suffer no energy loss in PbPb collision, at least for jet \(p_T > 100 \text{ GeV/c}\). A more detailed understanding of the jet \(p_T\) dependence of the b-jet to inclusive jet ratio and, more generally, the parton mass and flavor dependence of energy loss, would require a reduction of the statistical and systematic uncertainties.
References


