Measurement of the very forward inclusive jet cross section in pp collisions at $\sqrt{s} = 13$ TeV with CMS

The CMS Collaboration

Abstract

The measurement of very forward jets at 13 TeV in proton-proton collisions with CMS is reported. Such jets, in particular when correlated to other objects, are the most sensitive probes of the small-x parton dynamics in the proton. In this paper we demonstrate the fundamental capabilities of jet reconstruction with CMS using the CASTOR detector at a pseudorapidity of $-6.6 < \eta < -5.2$. The data is corrected for all detector effects and compared to model calculations. Models based on Gribov-Regge calculations tend to predict a slightly softer spectrum in contrast to PYTHIA tunes. The data indicates a spectrum more consistent with the shape predicted by PYTHIA, however, the data are compatible with all models within the experimental uncertainties.
1 Introduction

The most abundant processes at a hadron collider are QCD-mediated parton-parton scatterings. They represent a huge background for all processes and searches studied at LHC and must be modeled very precisely. In order to constrain the QCD background dedicated measurements of the underlying event and multi-parton interactions as well as general “minimum-bias” particle production are needed. Here we present a new measurement of the differential inclusive jet production cross section in the very forward direction, which is a powerful benchmark for QCD model predictions.

The most successful Monte Carlo event generators used to describe proton-proton collisions are based on the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) parton evolution equations and collinear factorization [1–4]. These models are extremely successful, but it is expected that the underlying assumptions break down when considering very small values of proton momentum fraction x. Extended models of parton dynamics are the Balitsky-Fadin-Kuraev-Lipatov (BFKL) [5–7] or the Ciafaloni-Catani-Fiorani-Marchesini (CCFM) [8–11] equations. It is an open quest to identify the phase space and exact characteristics of the breakdown of DGLAP. At LHC the most sensitive probe for this physics are gluons at small x and low $p_T$. Differential jet production cross sections are measured with the CASTOR calorimeter in CMS at $-6.6 < \eta < -5.2$ and are fully corrected for detector effects and all sources of systematic uncertainties are considered. This inclusive measurement turns out to be also very sensitive to multiple partonic interactions (MPI) in QCD collisions. The full potential to study the small-x parton dynamics will only be reached when very-forward jets are also correlated to jets at central rapidities, which is beyond the scope of this paper.

In this paper the very-forward differential jet production cross section measurement is compared to several model predictions. One set of models is based on Gribov-Regge field theory [12] and was developed to analyse ultra-high energy cosmic ray data. This is EPOS-LHC, which is a phenomenological model based on a parton-level Gribov-Regge implementation [13], and QGSJetII.4 [14], which is a more theory-motivated model implemented using Pomeron-tree and -loop diagrams calculated at all orders. Both models are based on DGLAP parton evolution and were re-tuned including LHC data up to 8 TeV. Furthermore, also various tunes of PYTHIA are also used. The PYTHIA6 [15] tune ZZ* is a tune derived from Z1 [16] and uses the CTEQ6L [17] parton distribution set. Also the PYTHIA8 [18] tune Monash was obtained by using minimum bias (MB) and underlying-event (UE) data from the LHC at $\sqrt{s} = 900$ GeV and 7 TeV to constrain the initial-state radiation and MPI parameters, and MB and UE data from the SPS at $\sqrt{s} = 200$ GeV and Tevatron at $\sqrt{s} = 0.3$ TeV, 0.9 TeV and 1.96 TeV to constrain the energy scaling. The tune CUETP8M1 [19, 20] is based on the tune Monash and uses the NNPDF2.3LO [21, 22] parton density function (PDF). The tunes were obtained by using UE data from CMS at $\sqrt{s} = 7$ TeV and from CDF at lower centre-of-mass energies. The PYTHIA8 event generator implementing the MBR [19, 23] model, also referred to as PYTHIA8-MBR, is used with the tunes 4C [24]. The tune 4C was obtained by using CDF UE and dijet data at $\sqrt{s} = 1.8$ TeV and retuned to early LHC data at $\sqrt{s} = 0.9$ TeV, 2.36 TeV, and 7 TeV. All these PYTHIA tunes describe parton showers using the DGLAP evolution equations. An overview of PYTHIA tuning can be found in [25].

2 The experimental setup

The central feature of the Compact Muon Solenoid (CMS) apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 T, which however was
The experimental setup turned off for the period when the data presented here were collected. Within the field volume are a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass/scintillator hadron calorimeter (HCAL). Muons are measured in gas-ionization detectors embedded in the steel return yoke. Extensive forward calorimetry complements the coverage provided by the barrel and endcap detectors.

The central detectors of CMS are complemented by calorimeters in the forward direction, which are all relying on the production of Cherenkov photons by charged particles in quartz. The “hadron forward” (HF) calorimeters cover the pseudorapidity interval $3 < |\eta| < 5.2$ and are using quartz fibers embedded in a steel absorber. In the very-forward direction there is the CASTOR calorimeter located at a distance of 14.2 m from the interaction point at a radial distance from the LHC beam of about 4 to 15 cm. This corresponds to a pseudorapidity coverage of $-6.6 < \eta < -5.2$. CASTOR is a sampling calorimeter using layers of fused silica quartz plates and tungsten absorbers. CASTOR is segmented in 14 longitudinal and 16 azimuthal channels. The first two front channels correspond to a combined depth of 20 $\lambda_0$ and are used as the electromagnetic section, while the full depth of the calorimeter amounts to 10 $\lambda_I$. The data of CASTOR are reconstructed in 16 towers, each summing up the 14 longitudinal channels at the same azimuthal location. Towers are zero suppressed with a threshold of 650 MeV $\sqrt{N_{\text{channel}}}$ by considering the noise level in each of the $N_{\text{channel}}$ channels used to construct the tower.

In order to allow the installation around the beam-pipe, the CASTOR calorimeter consists of two half-cylinders split around the vertical plane. The position of CASTOR is measured after final installation by a survey team with laser targets to a precision of about 1 mm. Movements of CASTOR during the magnetic field ramp-up and ramp-down are observed with positioning sensors. The absolute position of CASTOR is therefore determined with the positioning sensors in combination with the laser measurements. On the side facing the interaction point, two infrared sensors are mounted on each half-cylinder of CASTOR to measure the distance to the beam pipe, which itself is fixed at multiple locations. Two potentiometer sensors measure the opening between the two halves. All sensor measurements are added to a global $\chi^2$ minimization for which the x- and y-positions of each half are free parameters. The position of CASTOR is determined with an uncertainty in x and y of better than 2 mm. The precision of this determination is not sufficient to measure a potential rotation of the calorimeter around the beam line, however, due to the large distance from the interaction point this does not further affect the acceptance of the detector in a meaningful way. Similarly, any possible tilt of the CASTOR detector around the axes perpendicular to the beam are neglected in this procedure.

The absolute energy scale of CASTOR is estimated from CMS collision data with a cross-calibration to HF at 7 TeV [26]. For the cross-calibration the hadron-level corrected energy measured in $3 < |\eta| < 5$ (with an energy scale uncertainty of 10%) is extrapolated with a large set of Monte Carlo generators to the acceptance of CASTOR (extrapolation uncertainty of 10%) and is correlated to the hadron level corrected response of CASTOR (uncertainty from non-compensation of 5%). The resulting energy scale is consistent with the value derived from test-beam measurements (corrected for various in-situ and aging effects). Summing all uncertainties in quadrature yields 15% energy scale uncertainty. Furthermore, all channels of CASTOR are intercalibrated to the observed response of halo-muon events in the range of low energy deposits. The per-channel muon response is also used to transport the absolute energy scale from 7 TeV to the 13 TeV collision data.

A more detailed description of the CMS detector can be found in Ref. [27], where also the coordinate system is defined.

The response of all detectors is simulated using the GEANT4 [28] framework in order to deter-
mine event-selection efficiencies and the correction to the stable particle level. For this purpose
Minimum Bias collisions are simulated with PYTHIA8 CUETP8M1, MBR as well as EPOS-LHC.

Special attention is paid to reproduce the conditions during data taking in these simulations.
Thus, no magnetic field is chosen as well as a realistic beamspot position. Also the position of
CASTOR as it is measured after the installation is correctly accounted for in the simulations,
and malfunctioning channels are excluded in simulations as well as in data.

Jets in CASTOR are reconstructed with FASTJET [29, 30] using the anti-$k_t$ (use as given in [31])
jet algorithm with a radius parameter of $R = \sqrt{\Delta\phi^2 + \Delta\eta^2} = 0.5$. Jets are assembled from
CASTOR towers. CASTOR towers are assigned a nominal pseudorapidity value of $\eta = -5.9$
corresponding to the center of the calorimeter acceptance.

The clustered jet energy, $E_{\text{jet}}$, is converted to $p_T$ by using the factor $\cosh(5.9)$ and neglecting
the mass of particles. Since the calculation has no information on the true jet pseudorapidity
within the acceptance of CASTOR, this calculation can be maximally inaccurate by a factor of
$\approx 2$ corresponding to the edges of CASTOR at $\eta = -5.2$ and $-6.6$. It is one of the tasks of the
analysis to correct for the acceptance and $p_T$ resolution by means of an unfolding procedure.
The resolution of reconstructing $p_T$ is studied with simulations and found to be $\approx 1.25$ GeV at
$p_T = 3$ GeV and $\approx 4$ GeV at $p_T = 13$ GeV.

The choice of $p_T$ as the physical quantity to characterize jets is driven by the future plans to correlate
CASTOR jets with other objects measured by CMS in the more central acceptance. Since
CASTOR cannot measure the jet angle within its acceptance the conversion of reconstructed
energy to $p_T$ introduces increased migration effects. While this is corrected for in the unfolding
step, the measurement of a jet-energy instead of a jet-$p_T$ spectrum would have the advantage
to be less affected by migration effects.

For this analysis the data of the CASTOR calorimeter are analyzed. The performance of the

Figure 1: Left panel: Multiplicity distribution of reconstructed jets in CASTOR. Right panel:
Distribution of the number of ranges of CASTOR towers merged into a jet. Both plots are
for jets with $p_T > 3$ GeV. Simulations are done with CASTOR at measured position and the
magnetic field of CMS switched off.
CASTOR data as well as the implementation of CASTOR in the detector simulations are thoroughly evaluated. In Fig. 1 some characteristics of reconstructed jets in CASTOR are compared to the fully simulated detector response. Fig. 2 compares the reconstructed jet spectrum in CASTOR to simulations on detector level. The uncorrected jet-$p_T$ spectra of PYTHIA are a bit harder compared to the simulations. EPOS has the best agreement on detector level.

3 Data analysis

The analysis is based on data recorded in the very beginning of LHC Run2, during a period where the CMS solenoid was not switched on. The data has an interaction probability per bunch crossing of about 6% and the integrated luminosity corresponding to the data used is $0.212 \text{ nb}^{-1}$ with an uncertainty of 2.9% [32] on the luminosity scale, which is determined using the HF calorimeters. An unbiased trigger only requiring the presence of two colliding bunches is used to select events. Thus, there are no event selection steps involved besides the presence of a reconstructed jet in CASTOR well above the noise level of the calorimeter. It is verified, with data where only one LHC beam was present, that the background fraction per bin of the measurement is well below $10^{-5}$.

Jets in the CASTOR acceptance at generator level are also reconstructed with FASTJET and the anti-$k_t$ clustering algorithm with a radius of $R = 0.5$. Unlike at detector level the generator level jets have a continuous spectrum of $\eta$ and $\phi$ values.

Since jets in CASTOR are essentially clusters of measured energy in the calorimeter, this measurement must first be corrected for known calorimetric non-compensation effects. From test beam data [33] it is known that the response to hadrons (pions) is reduced with respect to that of electrons of the same energy by about a factor of two. A leading-order calibration relation based on reconstructed jet energies, $p_T^{\text{cal}} = f(p_T^{\text{rec}})$, is determined from Monte Carlo simula-
Figure 3: The jet $p_T$ calibration curve determined from Monte Carlo simulations with EPOS-LHC. Shown is the correlation between reconstructed jet-$p_T$ to generated jet-$p_T$ for a tight high-quality isolated jet selection. The red dots show the average value including the RMS of it, which is parameterized with a linear function in order to obtain the calibration relation.

Figure 4: Response matrix showing the distribution of reconstructed CASTOR jet-$p_T$ to their matched generator jet-$p_T$. Left panel: PYTHIA8 CUETP8M1-Tune. Middle panel: PYTHIA8 4C-Tune with MBR. Right panel: EPOS-LHC.
Table 1: Summary of systematic uncertainties in different $p_T$ bins when the result is normalized by luminosity.

<table>
<thead>
<tr>
<th>$p_T$-bin [ GeV]</th>
<th>lumi [%]</th>
<th>uncertainty w/o lumi &amp; CES [%]</th>
<th>energy scale [%]</th>
<th>total [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-4</td>
<td>±2.9</td>
<td>$-19/+18$</td>
<td>$-40/+55$</td>
<td>$-45/+58$</td>
</tr>
<tr>
<td>4-5</td>
<td>±2.9</td>
<td>$-36/+32$</td>
<td>$-46/+60$</td>
<td>$-58/+68$</td>
</tr>
<tr>
<td>5-6</td>
<td>±2.9</td>
<td>$-49/+45$</td>
<td>$-49/+72$</td>
<td>$-69/+85$</td>
</tr>
<tr>
<td>6-8</td>
<td>±2.9</td>
<td>$-56/+55$</td>
<td>$-58/+76$</td>
<td>$-80/+93$</td>
</tr>
<tr>
<td>8-10</td>
<td>±2.9</td>
<td>$-59/+55$</td>
<td>$-67/+81$</td>
<td>$-89/+98$</td>
</tr>
<tr>
<td>10-13</td>
<td>±2.9</td>
<td>$-49/+48$</td>
<td>$-78/+183$</td>
<td>$-92/+189$</td>
</tr>
</tbody>
</table>

4 Uncertainties

All possibly relevant sources of additional uncertainties have been investigated and are propagated to the final result if they are identified to be not negligible.

- One of the dominant sources of uncertainty is the energy scale of CASTOR (CES). The impact on the jet cross section exceeds 50% over the full range in $p_T$.
- The luminosity scale in this very early data is determined based on the HF calorimeters. The assigned uncertainty is 2.9%.
Table 2: Summary of systematic uncertainties in different $p_T$ bins when the result is normalized by number of events.

<table>
<thead>
<tr>
<th>$p_T$-bin [GeV]</th>
<th>uncertainty w/o CES [%]</th>
<th>energy scale [%]</th>
<th>total [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-4</td>
<td>$-22^{+23}_{-23}$</td>
<td>$-5^{+10}_{-10}$</td>
<td>$-23^{+25}_{-25}$</td>
</tr>
<tr>
<td>4-5</td>
<td>$-11^{+10}_{-10}$</td>
<td>$-2^{+0}_{-2}$</td>
<td>$-11^{+10}_{-10}$</td>
</tr>
<tr>
<td>5-6</td>
<td>$-29^{+28}_{-28}$</td>
<td>$-5^{+5}_{-5}$</td>
<td>$-29^{+28}_{-28}$</td>
</tr>
<tr>
<td>6-8</td>
<td>$-38^{+39}_{-39}$</td>
<td>$-22^{+8}_{-22}$</td>
<td>$-44^{+39}_{-39}$</td>
</tr>
<tr>
<td>8-10</td>
<td>$-45^{+43}_{-43}$</td>
<td>$-40^{+11}_{-40}$</td>
<td>$-60^{+44}_{-44}$</td>
</tr>
<tr>
<td>10-13</td>
<td>$-44^{+44}_{-44}$</td>
<td>$-59^{+73}_{-59}$</td>
<td>$-73^{+85}_{-85}$</td>
</tr>
</tbody>
</table>

- Besides the energy scale, the model dependence is the other dominant source of uncertainty. The model dependence of the unfolding procedure is evaluated with two different tunes of PYTHIA8, which are CUETPM1 and 4C with MBR and an Monte Carlo sample with EPOS. The model dependence uncertainty is around 20-50%.
- The impact of pile-up on the jet reconstruction is evaluated by using two data samples with a different Poissonian interaction probability, $\lambda$, per bunch crossing. Comparison of data at $\lambda \approx 2\%$ and $\approx 6\%$ does not indicate any deviation within the limits of the statistical precision.
- The location of CASTOR and its true acceptance. During ramps of the main CMS magnet the whole region around CASTOR moves on the order of a cm. Furthermore, CASTOR was de-installed and installed several times over the course of time. The position of CASTOR during exposure to collision data varies by about 1 cm and is determined to a precision of better than 2 mm. It is understood that changes in the location by 1 cm will make a significant impact on the selection of observed jets on detector level. This effect is evaluated with different Monte Carlo samples, where the position of CASTOR is varied within its known precision. The full unfolding process is repeated for these samples. The impact on the jet cross sections is in the range of 10-30%.

The magnitude of all uncertainties is summarized in Tab. 1 where CES refers to the CASTOR energy scale.

5 Results and Summary

In Fig. 5 and Fig. 6 (left) the differential jet cross section in the pseudorapidity interval $-6.6 < \eta < -5.2$ in proton-proton collisions at 13 TeV fully corrected for all detector effects is shown in the range $3 < p_T < 13$ GeV. The covariance matrix of this data is illustrated in Fig. 7. In the right panel of Fig. 5 and Fig. 6 the jet yield normalized by the number of jets in the visible $p_T$-range is shown. Due to the significant experimental uncertainties all of the models are essentially consistent with the data. However, it is interesting to note that all the PYTHIA tunes tend to slightly overpredict the cross section, while the two Gribov-Regge models EPOS-LHC and QGSJETII.4 tend to underpredict the cross section. The data favours a cross section in between these two extremes. The shape of EPOS-LHC and QGSJETII.4 is a bit softer than the data indicate, while all PYTHIA versions tend to reproduce the shape very well. The presented differential spectra have only a moderate sensitivity to the underlying PDF set of the model, however, they are very sensitive to MPI.
Results and Summary

Figure 5: Final unfolded differential jet-\(p_T\) spectrum in CASTOR compared to different model predictions. Left panel: the cross section. Right panel: the jet yield normalized by number of visible jets.

Figure 6: Final unfolded differential jet-\(p_T\) spectrum in CASTOR compared to different model predictions. Left panel: the cross section. Right panel: the jet yield normalized by number of visible jets.
Figure 7: The final covariance matrix of the statistical errors of the unfolded jet-$p_T$ spectrum with EPOS. The entries marked with a black box indicate negative values.

References


