Measurement of the Inclusive Jet Cross Section in pp Collisions at 7 TeV

The CMS Collaboration

Abstract

We report on a measurement of the inclusive jet cross section from pp collisions at a center-of-mass energy of $\sqrt{s} = 7$ TeV using data collected by the CMS experiment with an integrated luminosity of about $60 \text{nb}^{-1}$ and in the $p_T$ range of 18-700 GeV. Several different jet reconstruction methods are investigated using an anti-$k_T$ clustering algorithm. Studies of the systematic uncertainties in transverse momentum for each reconstruction method are presented, and the measured cross sections are found to be in agreement with next-to-leading order perturbative QCD calculations, within the experimental and theoretical uncertainties.
1 Introduction

The inclusive jet production cross section represents one of the basic measurements performed at hadron colliders, due to its very high rate. The most recent results from the CDF and DØ experiments at the Tevatron demonstrate agreement with next-to-leading-order theoretical predictions from perturbative QCD for jet $p_T$’s in the range of about 50-650 GeV, using about 1 fb$^{-1}$ of data at $\sqrt{s} = 1.96$ TeV [1, 2]. Using about 60 nb$^{-1}$ of data collected from proton–proton collisions delivered by the Large Hadron Collider at a center-of-mass energy of $\sqrt{s} = 7$ TeV and recorded by CMS, the jet transverse momentum spectrum is measured in the $p_T$ range of 18-700 GeV and for rapidities $|y| < 3.0$. Low $p_T$ jets are recorded with a prescaled Minimum Bias trigger, and the measurement is extended to high $p_T$ using single-jet triggers. Jets are reconstructed using the anti-$k_T$ algorithm with distance parameter of 0.5 [3] and studies of three different types of detector-level input to the jet clustering algorithm are presented, together with comparisons to next-to-leading order perturbative QCD theoretical predictions for the inclusive jet cross-section.

The relevant CMS detectors for this analysis consist of pixel and silicon-strip trackers up to $|\eta| < 2.4$ which, together with a 3.8 Tesla magnetic solenoid, provides track reconstruction down to about 100 MeV and a track momentum resolution of about 1% at 100 GeV; a granular electromagnetic crystal-calorimeter (ECAL) of $\Delta \eta \times \Delta \phi = 0.0174$ (barrel) extending up to $|\eta| < 3.0$ with an electromagnetic transverse energy resolution of about 3%/\sqrt{E_T}; a hermetic hadronic calorimeter (HCAL) extending up to $|\eta| < 5.0$ with a tower granularity of $\Delta \eta \times \Delta \phi = 0.087$ (barrel) and a transverse hadronic energy resolution of about 100%/\sqrt{E_T}; and an efficient muon system capable of reconstructing and identifying muons up to $|\eta| < 2.4$. Further details of the CMS Detector may be found in [4].

2 Jet Reconstruction

The anti-$k_T$ jet clustering algorithm with a distance parameter of 0.5 is used to reconstruct all jets. This jet clustering algorithm is both infrared and collinear safe. This summary presents studies of three different types of jets, each of which make use of the same anti-$k_T$ clustering algorithm, but which depend on the use of different detector information: calorimeter information only [5], calorimeter information combined with track information after jet clustering [6], and reconstructed particles from the CMS particle-flow algorithm [7, 8].

Calorimeter jets: Projective calorimeter towers, in the barrel region, are formed from single HCAL-cells and corresponding sets of $5 \times 5$ ECAL-crystals. Four-vectors are then formed from the energy of each calorimeter tower, taken to be the direct sum of the energy deposits in the associated HCAL-cell and ECAL-crystals. In the endcaps, calorimeter towers are also formed, but the combined HCAL-ECAL geometry is not projective. Calorimeter jets are formed by clustering the four-vectors of calorimeter towers. Additional information on calorimeter jets may be found in Ref. [5].

Calorimeter jets corrected with tracks: As in the case of calorimeter jets, calorimeter towers are first clustered into jets. A posteriori, each jet energy is corrected using the momentum of associated tracks: for each track inside the jet cone, an average single-particle response is removed from the calorimeter energy and replaced with the track momentum. The momenta of associated tracks outside the jet cone are simply added to the jet energy and explicit corrections are applied due to calorimeter readout threshold effects. Further details of the track corrections applied to calorimeter jets (referred to hereafter as “Jets Plus Tracks” or JPT) may be found in Ref. [6].
Particle-flow jets: The CMS particle-flow algorithm reconstructs individual particles (leptons, photons, charged and neutral hadrons) by linking tracks, ECAL clusters, and HCAL clusters. Each particle is reconstructed with the optimal momentum or energy, by considering information from all sub-detectors. Broadly speaking, charged hadrons are reconstructed from tracks; photons and neutral hadrons are reconstructed from energy clusters in the electromagnetic and hadron calorimeters. A detailed description of the particle-flow algorithm may be found in Ref. [7, 8]. Particle-flow jets are formed by clustering the four-vectors of reconstructed particles.

Differences across detector boundaries: Beyond $|\eta| > 2.0$, JPT jets do not include tracking information, whereas particle-flow jets exploit tracking information up to the full tracker acceptance of $|\eta| < 2.4$. Even for $2.4 < |\eta| < 2.9$, particle-flow jets may still possess reconstructed charged hadrons due to the 0.5 distance parameter in the $k_T$ clustering algorithm. Finally, above $|\eta| > 2.0$ some differences between calorimeter jets and JPT jets remain, as JPT jets apply explicit corrections accounting for calorimeter read-out thresholds.

2.1 Jet Energy Scale and Resolution Determination

The observed jet energy (in the case of all three jet types described above) is corrected on average, from detector level to hadron level [9, 10], and the corrections consist of two steps applied in sequence: a relative correction removes the pseudo-rapidity dependence of the jet energy response and an absolute correction restores the response to unity. Currently the energy scale corrections are derived from simulation, but will be determined from data in the future [11–14].

After applying the relative jet energy scale corrections to data from simulation, a residual discrepancy in the jet energy response, relative to the central region, is observed in the data when compared with that expected from simulation. The discrepancy increases linearly with $\eta$ and is up to 5% at the highest pseudo-rapidities. Using an exclusive sample of dijet events, pseudo-rapidity dependent residual-calibrations to the jet energy scale (relative to the central region) are determined and consequently applied to correct for this effect [15].

Jet energy resolutions are currently derived in the simulation and validated with data to within about 10%. Efforts to derive jet energy resolutions from data are underway using the jet energy asymmetry method in back-to-back dijet events [15, 16].

3 Cross-Section Measurement

The differential inclusive jet cross-section is measured in bins of jet transverse momentum $p_T$ and rapidity $y$ defined as:

$$\frac{d^2\sigma}{dp_T dy} = \frac{C_{\text{res}} \cdot N_{\text{jets}}}{L \cdot \epsilon \cdot \Delta p_T \cdot \Delta y}$$

(1)

where $N_{\text{jets}}$ is the measured number of jets per bin, $C_{\text{res}}$ is a correction factor for bin-to-bin migrations due to resolution effects, $\Delta p_T$ and $\Delta y$ are the bin widths in $p_T$ and $y$, $L$ is the total integrated luminosity and $\epsilon$ is the product of event and jet efficiencies, as calculated per jet. The measurements are reported in units of (pb/GeV).

3.1 Trigger Requirements and Spectrum Construction

The efficiency-corrected, luminosity-normalized $p_T$ spectra from Minimum Bias and single-jet triggers (uncorrected jet $p_T > 6$ GeV, 15 GeV, 30 GeV) are combined into a smooth, continuous
3.2 Event Selection

$p_T$ spectrum. To keep calculations simple, each trigger contributes to a discrete $p_T$ range, and no two triggers contribute to the same $p_T$ range.

The efficiency of the single-jet triggers is determined as a ratio to the Minimum Bias trigger, or to the lower $p_T$ trigger. Figures 1, 2, and 3 show the resulting trigger yield for the energy-scale-corrected jet $p_T$ spectra and the turn-on curves of the single-jet triggers for all jet types. Calorimeter jets use only the (uncorrected) 30 GeV single-jet trigger and require that each considered jet have a (corrected) $p_T$ above 50 GeV, well above the point at which the trigger becomes efficient. In the case of JPT jets and particle-flow jets, the best trigger and corresponding (corrected) jet $p_T$ threshold are selected by requiring close to 100% efficiency and highest available statistics. These requirements ensure that $\epsilon$ in Eq. 1 is close to unity.

3.2 Event Selection

All triggered events are required to have a good primary vertex (PV) consistent with the measured transverse position of the beam (referred to as the beam spot, or BS) using the following selection: at least one vertex must be reconstructed in the event, the z-coordinate of the PV must lie within the luminous region $|z_{PV}| < 15$ cm, the fit for the PV must be sufficiently constrained with at least five associated tracks, and the radial distance of the PV must be less than 0.15 cm, consistent with the BS. This selection, while powerfully rejecting non-collision and beam related backgrounds, is highly efficient (~100%) for this analysis.

In the cases of calorimeter jets and JPT jets, loose quality criteria [17] are applied requiring that each jet (1) has an electromagnetic energy fraction greater than 1% if that jet is well contained within the detector acceptance with $|\eta| < 2.6$, (2) has at least 90% of its energy shared amongst at least two reconstructed calorimeter cells or crystals, and (3) has less than 98% of its energy contained in the hottest HCAL electronic readout channel. This selection is essentially 100% efficient for jets having corrected $p_T > 50$ GeV.

In the case of particle-flow jets, loose quality criteria [18] are applied requiring that each jet have (1) at least two constituent particles of which at least one must be charged, (2) at least some fraction of energy from charged hadrons, (3) an energy fraction which can not solely be attributed to electrons, and (4) at most 90% of energy from neutral particles, be they photons or neutral hadrons. The selection is more than 98% efficient for jets at $p_T > 10$ GeV and more than 99.5% efficient at $p_T > 30$ GeV.

3.3 Corrections due to finite jet resolution

The measured jet $p_T$ spectra are corrected for resolution effects using an ansatz: the true jet $p_T$ spectrum is assumed to be modeled by some parameterization as a function of jet transverse momentum, $f(p_T)$, which is then smeared using known jet resolutions and the parameters of the model are determined by fitting the smeared transverse momentum spectrum $F(p_T)$ to data.

$$f(p_T) = N p_T^{-\alpha} \cdot \left(1 - \cosh(y_{min}) \frac{2p_T}{\sqrt{s}} \right)^\beta \exp(-\gamma/p_T)$$

$$F(p_T) = \int_0^\infty f(p'_T) R(p'_T - p_T; \sigma) dp'_T$$

The function $R(p'_T - p_T; \sigma)$ is a smearing function, assumed to be Gaussian, and the ansatz $f(p_T)$ is based on early phenomenological fits, partly motivated by the parton model [19, 20], partly motivated by studies at the Tevatron [21, 22], as well as a desire to simultaneously fit the spectra from heavy flavour jets [23]. The $\alpha$ power term is associated with hard (power law)
production. A typical exponent for single particle production is $\alpha = 4 - 6$. The $\beta$ power term represents the suppression effect at the edges of the parton distribution function phase space. This threshold term does not typically contribute to the overall ansatz until the spectrum has reached roughly half of the kinematic range, or about 1750 GeV at 7 TeV, and is currently fixed to an effective value of $\beta = 8$ because of the low kinematic range accessible for the current integrated luminosity.

The unsmearing corrections $C_{\text{res}} = f(p_T)/F(p_T)$, plotted in Fig. 4, are obtained as the ratio between the original ansatz and the smeared ansatz, at the center of each bin in measured jet $p_T$. The most important input to the convolution are the parameterized jet $p_T$ resolutions, reported in [16]. The $\sigma$ in the Gaussian resolution function, $R(p_T' - p_T; \sigma)$, is taken from the following parameterization of the resolution as a function of the true jet $p_T'$:

$$
\left( \frac{\sigma}{p_T'} \right)^2 = \text{sgn}(N) \left( \frac{N}{p_T'} \right)^2 + \left( \frac{S}{\sqrt{p_T'}} (p_T')^m \right)^2 + C^2
$$

(4)

where $N$ is a noise term, $S$ is a stochastic term and $C$ is a constant term. The traditional fit for calorimeter jets is obtained by constraining $N \geq 0$ and $m = 0$. For particle-flow jets, which include tracking information in addition to calorimeter information, the best results are obtained by freeing the sign on $N$ and fixing $C = 0$. In this case the $m \neq 0$ term fits the appropriate effective behavior between the traditional calorimeter-based stochastic $S$ and constant $C$ terms. The values for $N$, $S$, $C$, and $m$ are currently taken from simulation. Generally speaking, the noise terms for all three jet types are typically less than about 5 GeV and the constant terms are zero. The stochastic terms are currently estimated to be about 54% for calorimeter jets, about 36% for JPT jets, and about 30% for particle-flow jets.

The Gaussian response function assumed in the smeared ansatz does not always provide a good description of the jet resolution, especially in cases where the jet energy is badly reconstructed [15]. In simulations, a small fraction of jets are found to lie in non-Gaussian tails, the vast majority of which reside below the calibrated, average jet energy response; such jets are found not to play any systematic role due the rapidly falling $p_T$ spectrum. The fraction of jets lying in the non-gaussian tails above the calibrated, average jet energy response is tiny and the effect that they have on the inclusive jet $p_T$ spectrum is currently estimated, from simulation, to be negligible.

### 3.4 Experimental Systematic Uncertainties

The primary sources of systematic uncertainties in the measurement of the jet cross-section are the jet energy scale and the jet energy resolution, and are summarized in Fig. 5.

#### 3.4.1 Jet Energy Scale

Due to the small integrated luminosity currently available, precise quantitative estimates of the jet-energy scale (and uncertainty) using, e.g. photon-jet balancing from data, are not yet available. Therefore, corrections to the jet energy are derived from simulation. Nevertheless, an understanding of the absolute jet energy scale in the barrel region from data is assisted by several facts: (1) the agreement between data and simulation for the $\pi^0$ mass peak already sets the ECAL energy scale to better than 1% [24]; (2) using isolated charged hadrons, the single pion response in the hadron calorimeter for high $p_T$ tracks agrees to better than a few percent between data and simulation [24, 25]. Considering these two facts in the absence of a direct absolute energy scale measurement in data, a very conservative 10% absolute uncertainty on the energy scale for calorimeter jets is assumed.
Further, (3) the magnetic field map has been measured in the tracker volume to within 0.1% [26, 27]; (4) the particle-flow algorithm has been commissioned [24] with the 7 TeV collision data and the particle composition of jets (relative energy fraction of charged hadrons, photons, and neutral hadrons) agree to better than a few percent between data and simulation; (5) comparisons of jet energy corrections derived using Pythia [28] and Herwig [29] limit the impact of fragmentation to less than about 2-3% for particle-flow jets; finally, (6) direct comparisons of calorimeter jets and particle-flow jet energies above $p_T > 40$ GeV show agreement to better than 2% in the barrel at $|y| < 1.4$. Given all of these constraints, but in the absence of a direct quantitative scale estimate from data, a conservative 5% absolute uncertainty on the energy scale for JPT jets, and particle-flow jets is assumed.

Following the residual-calibration, applied for the observed variations of the relative jet-energy scale across detector boundaries in data (as briefly summarized in Section 2.1), a 2% $\times |\eta|$ systematic uncertainty that linearly increases with pseudo-rapidity on the relative energy scale, is assumed for calorimeter jets, JPT jets and particle-flow jets.

### 3.4.2 Jet Energy Resolution

Quantitative estimates of the jet-energy resolution (and uncertainty), e.g. via asymmetry measurements in dijet balancing from data, are also not yet available. Nevertheless, evidence from data suggests that the jet energy resolution (as derived from simulation) is under control: the agreement between data and simulation for the $\pi^0$ width as measured in the ECAL agrees to better than 3% [24]; numerous meson and baryon resonances have been observed in the tracker and their measured widths agree with simulation to very high precision (better than 1% in many cases) [30]. The hadron energy resolution will be measured with collision data; currently, early test-beam data agrees with the simulation at the level of 10%. Hence, in the absence of a direct quantitative resolution estimate from data, a 10% uncertainty on the energy resolution for calorimeter jets, JPT jets, and particle-flow jets is assumed.

### 3.4.3 Luminosity

The luminosity of proton–proton collisions occurring within the interaction region of the CMS detector is monitored by two forward hadron calorimeters (HF), with a coverage of $3 < |\eta| < 5$. Vertexing multiple tracks within the interaction region provides an independent luminosity measurement cross-check and is found to be compatible with that from the HF. Knowledge of the LHC beam currents and a precise study of the transverse size of the beams provides an 11% accuracy on the absolute calibration of the luminosity measurement [31], and directly translates into an 11% normalization uncertainty on the inclusive jet cross section.

### 4 Comparison to Theory

Perturbative QCD (pQCD) theoretical predictions for the inclusive jet cross-section at next-to-leading order (NLO) accuracy are calculated using NLOJET++ [32] within the fastNLO [33, 34] framework with CTEQ-6.6 parton distribution function (PDF) sets [35]. A phenomenological correction to the NLO parton-level prediction is required to account for non-perturbative effects such as hadronization and multiple parton interactions. Such corrections are estimated by comparing a leading-order (LO) parton-level prediction for inclusive jet production with the corresponding prediction for inclusive jet production after hadronization, etc.

Two different models, Pythia [28] and Herwig++ [29], are used to predict these non-perturbative (NP) corrections, and the mean of the two predictions taken as the correction. Fig. 6a displays
the derived corrections as a function of jet $p_T$ and for different bins in rapidity $y$.

The theoretical uncertainty associated with the NP corrections is taken to be half of the difference between the Pythia and Herwig++ predictions; the two predictions can differ by as much as 100% at low $p_T$. Additional theoretical uncertainties arise from the PDFs, the variations of which are estimated by repeated evaluations of the NLO predicted inclusive jet cross section for the different PDFs in the CTEQ-6.6 set. Finally, uncertainties from any residual dependence on the choice of renormalization and factorization scales are estimated by varying the scale between $\mu_R = \mu_F = p_T/2$ and $\mu_R = \mu_F = 2p_T$, where the default choice is $\mu_R = \mu_F = p_T$.

Figure 6b shows that the leading sources of theoretical uncertainty are the NP corrections for $p_T < 50$ GeV; in the central region, PDF uncertainties dominate at higher $p_T$; in the forward regions, uncertainties in the non-perturbative corrections dominate at higher $p_T$.

5 Results

The unfolded, measured jet $p_T$ spectra are summarized in Figs. 7, 8, and 9, which also includes the NLO predicted curve. The integrated luminosity for the three different cross-section measurements correspond to about 60 nb$^{-1}$. The systematic uncertainties due to experimental effects are shown as yellow bands.

Each of the three different jet reconstruction methods are compared in Fig. 10, which shows the ratio of the unfolded inclusive jet cross-section in data to the NLO theoretical prediction, for different bins in rapidity $y$. The results from these three methods agree with the theoretical prediction and with each other to within 20% over most of the measured $p_T$ and rapidity ranges, well within the uncertainties associated with particle-flow jet energy scale, which is the most accurate of the three measurements. The theoretical systematic uncertainty band is plotted about the NLO prediction and the experimental systematic uncertainty band is plotted about the ansatz fit to the data points for particle-flow jets.

6 Conclusions

Using about 60 nb$^{-1}$ of data collected from proton–proton collisions delivered by the Large Hadron Collider at a center-of-mass energy of $\sqrt{s} = 7$ TeV and recorded by the CMS detector, the jet transverse momentum spectrum is measured in the $p_T$ range of 18-700 GeV and for different rapidities $|y| < 3.0$. Three different jet reconstruction methods were compared as a function of $p_T$ and for different bins in rapidity $y$. The largest systematic uncertainties arise from a current imprecise knowledge of the absolute jet energy scale and the jet energy scale relative to the central region of the detector. Within the current experimental and theoretical uncertainties, the theoretical calculation predicts the inclusive jet cross section observed in the data well, both in transverse momentum and in rapidity.
Figure 1: (Left) Jet $p_T$ spectra of the trigger stream used for calorimeter jets. (Right) Turn-on curve in the central rapidity region versus jet $p_T$ for the applied trigger stream for calorimeter jets. The markers give the estimated turn-on point, after which the trigger is more than 99% efficient. The trigger turn-on for calorimeter jets is sharp due to their close similarity with the calorimeter jets used by the trigger.
Figure 2: (Left) Jet $p_T$ spectra for each of the applied trigger streams, for JPT jets. (Right) Turn-on curves in the central rapidity region versus jet $p_T$ for each of the applied trigger streams, normalized to the preceding one, for JPT jets. The markers give the estimated turn-on point, after which the trigger is more than 99% efficient. The trigger turn-on for JPT jets is slower than for calorimeter jets due to the use of tracking information, which smears the JPT jet energy with respect to the calorimeter jets used by the trigger.
Figure 3: (Left) Jet $p_T$ spectra for each of the applied trigger streams, for particle-flow jets. (Right) Turn-on curves in the central rapidity region versus jet $p_T$ for each of the applied trigger streams, normalized to the preceding one, for particle-flow jets. The markers give the estimated turn-on point, after which the trigger is more than 99% efficient. The trigger turn-on for particle-flow jets is slower than for calorimeter jets, due to the use of tracking information, which smears the particle-flow jet energy with respect to the calorimeter jets used by the trigger (the turn-on is also slightly slower than that for JPT jets, as JPT jets are more closely based on calorimeter jets).
Figure 4: Unsmearing correction factors for different rapidity ranges for (a) calorimeter jets, (b) JPT jets, and (c) particle-flow jets.
Figure 5: Fractional experimental systematic uncertainties for (a) calorimeter jets, (b) JPT jets, and (c) particle-flow jets. The total systematic uncertainty band includes separately the 11% uncertainty from the integrated luminosity measurement.
Figure 6: (a) Non-perturbative corrections to NLO QCD calculations for the anti-$k_T$ algorithm, as determined using Pythia [28] and separately by Herwig++ [29]; the mean of the corrections determined using Pythia and Herwig++. (b) Fractional theoretical uncertainties; the uncertainties associated with the non-perturbative corrections are taken to be half of the difference between the Pythia and Herwig++ predictions.
Figure 7: Comparison between the unfolded measured spectra and the theory predictions for calorimeter jets. For better visibility the spectra are multiplied by arbitrary factors (indicated in the legend).
Figure 8: Comparison between the unfolded measured spectra and the theory predictions for JPT jets. For better visibility the spectra are multiplied by arbitrary factors (indicated in the legend).
Figure 9: Comparison between the unfolded measured spectra and the theory predictions for particle-flow jets. For better visibility the spectra are multiplied by arbitrary factors (indicated in the legend).
Figure 10: The unfolded measured spectra in data plotted as the ratio of data to theory prediction for (triangles) calorimeter jets, (squares) JPT jets, (circles) particle-flow jets for (a) $|y| < 0.5$, (b) $0.5 < |y| < 1.0$, (c) $1.0 < |y| < 1.5$, (d) $1.5 < |y| < 2.0$, (e) $2.0 < |y| < 2.5$, and (f) $2.5 < |y| < 3.0$. The experimental uncertainty band corresponds to that of particle-flow jets and is plotted about the ansatz fit to the data from particle-flow jets.
Figure 11: The unfolded measured spectra in data plotted as the ratio of data to theory prediction for calorimeter jets for (a) $|y| < 0.5$, (b) $0.5 < |y| < 1.0$, (c) $1.0 < |y| < 1.5$, (d) $1.5 < |y| < 2.0$, (e) $2.0 < |y| < 2.5$, and (f) $2.5 < |y| < 3.0$. The experimental uncertainty band corresponds to that of calorimeter jets and is plotted about the ansatz fit to the data from calorimeter jets.
Figure 12: The unfolded measured spectra in data plotted as the ratio of data to theory prediction for JPT jets for (a) $|y| < 0.5$, (b) $0.5 < |y| < 1.0$, (c) $1.0 < |y| < 1.5$, (d) $1.5 < |y| < 2.0$, (e) $2.0 < |y| < 2.5$, and (f) $2.5 < |y| < 3.0$. The experimental uncertainty band corresponds to that of JPT jets and is plotted about the ansatz fit to the data from JPT jets.
Figure 13: The unfolded measured spectra in data plotted as the ratio of data to theory prediction for particle-flow jets for (a) $|y| < 0.5$, (b) $0.5 < |y| < 1.0$, (c) $1.0 < |y| < 1.5$, (d) $1.5 < |y| < 2.0$, (e) $2.0 < |y| < 2.5$, and (f) $2.5 < |y| < 3.0$. The experimental uncertainty band corresponds to that of particle-flow jets and is plotted about the *ansatz* fit to the data from particle-flow jets.
References


