PROCEEDINGS OF SCIENCE

Inclusive jet measurements in CMS

Patrick L.S. Connor [a,b,*]

[a] *Institut für Experimentalphysik, Universität Hamburg*
Luruper Chaussee 149 (Geb. 68), D-22761 Hamburg

[b] *Center for Data Computing in natural Sciences, Universität Hamburg*
Notkestraße 11, D-22607 Hamburg

E-mail: patrick.connor@desy.de

Measurements of jet production in proton-proton collisions at the LHC are crucial for precise tests of QCD, improving the understanding of the proton structure and are important tools for searches for physics beyond the standard model. We describe the most recent set of inclusive jet measurements performed using CMS data, mentioning methodological specificities with respect to other analyses. Finally, we compare them to various theoretical predictions.

41th ICHEP
6-13 July 2022
Bologna, Italy

*on behalf of the CMS Collaboration

PoS(ICHEP2022)774

Inclusive jet cross section measurements allow for tests of state-of-the-art predictions of the standard model (SM) and of the SM supplemented with effective field theories (EFT), as well as for constraints on parton distribution functions (PDFs) at high momentum fraction x and on the strong coupling α_S. Indeed, the hadronic cross section may be factorised as follows [1]:

$$\sigma_{\mathrm{pp}\to\mathrm{jet}+X} = \sum_{ij\in gq\bar{q}} \underbrace{f_i(x_i,\mu_F^2)}_{\mathrm{PDF}} \otimes \underbrace{f_j(x_j,\mu_F^2)}_{\mathrm{PDF}} \otimes \underbrace{\hat{\sigma}_{ij\to\mathrm{jet}+X}\left(x_i, x_j, Q^2/\mu_F^2, Q^2/\mu_R^2, \alpha_S(\mu_R^2)\right)}_{\mathrm{SM(EFT)}} \quad (1)$$

where the partonic cross section $\hat{\sigma}$ can now be calculated at next-to-next-to-leading order (NNLO) [2–4] and next-to-leading order supplemented with next-to-leading logarithm (NLO+NLL) [5], and where, for each term of the sum on the flavours, the PDF f appears once per proton.

In these proceedings, we describe the recently published measurements at 13 TeV by the CMS Collaboration [6, 7] for two values of the R parameter of the anti-k_T clustering algorithm [8, 9]. We focus on the analysis of the experimental data and emphasise differences with respect to a previous measurement at 8 TeV [10], but leave the QCD interpretation to the dedicated proceedings of the same conference.

The measurements at $R = 0.4$ (0.7) are performed for jets clustered at particle level with high-pileup data recorded in 2016, corresponding to an integrated luminosity of 36.3 fb^{-1} (33.5 fb^{-1}):

$$\frac{\mathrm{d}^2\sigma_{\mathrm{pp}\to\mathrm{jet}+X}}{\mathrm{d}p_\mathrm{T}\,\mathrm{d}y} = \frac{1}{\mathcal{L}}\frac{N_\mathrm{jets}^\mathrm{eff}}{\Delta p_\mathrm{T}\,\Delta y} \quad (2)$$

with transverse momentum $p_\mathrm{T} > 97$ GeV and absolute rapidity $|y| < 2.0$; the measurement at 8 TeV was only provided for $R = 0.7$ with $\mathcal{L} = 19.7$ fb^{-1}, and whereas that measurement also included $p_\mathrm{T} > 21$ GeV and $|y| < 4.7$ (obtained with a low-pileup data set of lower luminosity), only data corresponding to $p_\mathrm{T} > 74$ GeV and $|y| < 3.0$ were included in the QCD interpretation). The $N_\mathrm{jets}^\mathrm{eff}$ corresponds to the effective number of jets after the various corrections for the detector effects, which will be described later on. The bin width Δp_T is taken from former measurements from the CMS Collaboration, with bins merged by two; the bin width Δy is not changed.

The steeply falling cross section has to be determined with subpercent level statistical precision. At 8 TeV, a 1% bin-to-bin uncorrelated systematic uncertainty was included in the measurement to cover residual effects. One of the challenges of the present measurements was to reduce the size of the bin-to-bin variations to provide even more stringent constraints in the QCD interpretation.

Jets are reconstructed offline with the particle flow (PF) algorithm (like at 8 TeV), now also supplemented with charge hadron subtraction (CHS). The CMS standard quality criteria are applied, such as event filters, or selection on the primary vertex and on jet constituents.

The high production rate prevents from recording all events, therefore, like in the 8 TeV analysis, events are recorded using various single-jet triggers. In the real data, events are here considered if the leading jet within $|y| < 2.5$ with an online reconstruction [11] has fired one of the single-jet triggers (at 8 TeV, each jet was triggered on independently) in the region of >99.5% of efficiency (99%) as a function of the offline reconstructed p_T. The contribution of each event is then multiplied by a time- and energy-dependent prescale factor in the count of jets; in contrast, at 8 TeV, the prescale factor was assumed constant over the whole year. In addition, we now correct the residual, relative trigger inefficiency; the differences between various methods to determine the efficiency are addressed with a bin-to-bin uncorrelated uncertainty of 0.2%, which is to be compared with the statistical precision, reaching 0.3-0.4% at medium p_T. In the analysis of the 8 TeV data, the residual trigger inefficiency was mitigated with the aforementioned 1% bin-to-bin uncorrelated uncertainties. The count of the jets, as opposed to the analysis of the 8 TeV, also had to be corrected for losses of events due to time jitter in the trigger by applying an extra efficiency correction as a function of the p_T and pseudorapidity η of the jets present in $2.0 < |\eta| < 3.0$; this uncertainty is minor at high p_T and at most a few percents at medium p_T.

Large samples of simulated data are generated with LO generators, assuming a pure QCD jet signal, and undergo the same reconstruction procedure as the real data. In the following, we first describe the jet energy and pileup corrections applied to both real and simulated data at detector level, then how the simulated data are used to unfold the real data.

Similarly to the 8 TeV analysis, the detector level jet p_T is corrected to match in average the particle level jet p_T and to ensure a realistic response in the real and simulated samples [12, 13]. However, the treatment of the resolution was refined to better account for the non-Gaussian tails of the jet response, as it was found to impact the final cross section beyond the percent level; in particular, we improved the identification of the tails by taking the numerical derivative of the logarithm of the jet response, whose linear and non-linear regimes respectively correspond to the Gaussian and non-Gaussian regimes of the response itself. These jet energy corrections come with various, standard sources of uncertainty with various shapes and at orders of magnitude from subpercent to percent levels.

The particles from the pileup may be clustered together with the jets from the signal or clustered as additional jets. The former cause an offset in the jet energy scale, and the latter provide additional jets in the total count of jets. Furthermore, the larger particle multiplicity reduces the jet resolution. The residual offset after CHS is now corrected within the jet scale corrections. The number of additional jets does not play a significant role in the phase space of this measurement. However the impact on jet resolution is significant, therefore, as in the analysis of the 8 TeV data, the pileup profile of the simulated data is corrected to match that of the real data; an additional uncertainty of 4.6% from the minimum-bias cross section is propagated to the measurement.

The simulated data are used to build response matrices (RMs) directly, unlike the 8 TeV measurement where the RMs were constructed from the convolution of fixed-order predictions and of a Gaussian response estimated from the simulated data sets. This change of strategy is also motivated by the deviation from the Gaussian-response hypothesis; the statistical uncertainties from the limited statistics of the simulated samples, smaller than that of the real data, are also propagated to the final cross section. To unfold, we use a least-square minimisation without Tikhonov regularisation, as implemented in the TUnfold package [14, 15]; the full phase space with unfolded in a single fit to account properly for all correlations among y bins. This contrasts from the 8 TeV analysis where the unfolding was performed in each y bin separately with the approach proposed by D'Agostini and implemented in the RooUnfold package [16, 17]. Unmatched jets in the construction of the RMs are now accounted for at both detector and particle levels in the form of background subtraction and efficiency corrections, whereas in the 8 TeV analysis, these were also addressed with the aforementioned 1% bin-to-bin uncorrelated uncertainties.

To infer the propagation of the systematic uncertainties from detector level to particle level, the unfolding procedure is repeated for each systematic variation. The model dependence of the RMs is mitigated with a systematic variation of the model where the simulated data are reweighted to resemble the real data; at 8 TeV, no model dependence was included.

Equation 1 suggests that the inclusive jet spectrum be smooth, i.e. that the deviations of the experimental points around a smooth function be described by the bin-to-bin uncertainties and partial correlations. However, steps or outliers may be introduced in the course of the data reduction, e.g. in case the contributions from the successive triggers are not normalised consistently. To test the smoothness of the experimental spectrum, a polynomial fit of the data is performed in each y bin, following the least-square minimisation described in Ref. [18]:

$$\chi_n^2 = \min_{b_i} \left[\left(\mathbf{x} - \mathbf{y}_{b_i}\right)^\top \mathbf{V}_{\mathbf{x}}^{-1} \left(\mathbf{x} - \mathbf{y}_{b_i}\right) \right] \quad \text{with} \quad y_{b_i}^j = \frac{1}{\Delta p_{\mathrm{T}}^j} \int_{p_{\mathrm{T}}^j} \exp\left(\sum_{i=0}^{n} b_i \, T_i \left(\log p_{\mathrm{T}}\right) \right) \mathrm{d}p_{\mathrm{T}} \tag{3}$$

where $\mathbf{x}$ corresponds to the experimental data, as given in Eq. 2. This test is applied at the successive stages of the analysis to ensure that the each procedure preserves the data quality. A polynomial of order 6 is sufficient in all y bins after the unfolding. Furthermore, as fluctuations may propagate differently in the various systematic variations, hence introduce tensions in a QCD interpretation of the data, the systematic variations are smoothed following the same approach.

The measurements and their comparisons to theoretical predictions may be found to Ref. [6]. In Fig. 1, we show a subsample of the comparisons with predictions at NNLO and NLO+NLL precision for $R = 0.4$.

- The NNLO predictions are provided with two scale choices; the statistical precision of the calculation (not shown in the figure) is similar to that of the experimental data. The H_{T} scale seems to provide a slightly better description of the data than the jet p_{T} scale.

- The NLO+NLL predictions [5] are shown for various choices of global PDF [19–23] sets. In comparison with the NNLO predictions, the NLO+NLL predictions seem slightly further away from the data for $|y| < 1.0$, although the difference is still within uncertainties. Among the different PDFs, two should be noted: HERAPDF2.1 shows a significant difference, which illustrates the expected impact of inclusive jet production on the PDF determination; ABMP16 also disagrees with the data beyond uncertainties.

A direct comparison of the measurements at 8 and 13 TeV is not possible, as the predictions are performed at different orders, and as all global PDF sets have now been updated to newer versions, with the exception of MMHT14; nonetheless, a similar dependence on y at high p_{T} can still be observed.

In the original publication, both NNLO and NLO+NLL results were obtained with the k factor technique. An addendum to the publication was recently accepted, showing comparisons with NNLO predictions obtained from NNLO interpolation grids [24].

To summarise, the CMS Collaboration has produced two measurements of inclusive jet production in pp collisions at 13 TeV. The experimental analysis includes corrections to the jet count, the jet energy, and the pileup; all effects are corrected via the procedure of unfolding. Tests of smoothness have been applied to the data at all steps of the analysis; furthermore, the bin-to-bin fully correlated systematic variations have been smoothed to remove artificial tensions in the QCD interpretation. Data have been compared to fixed-order predictions at NLO+NLL and NNLO. Systematic differences with respect to the analogous analysis with 8 TeV data have been described.

PoS(ICHEP2022)774

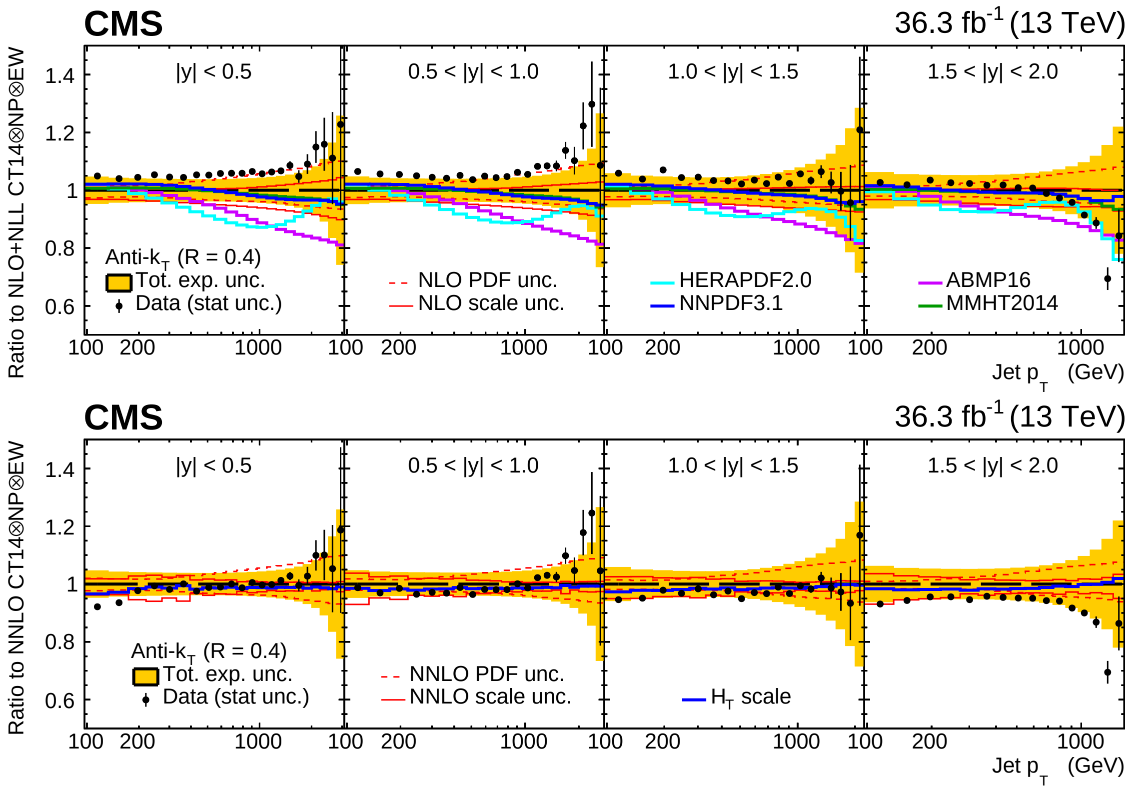

Figure 1: Double-differential cross section of inclusive jet production as a function of p_T and $|y|$, presented as ratios to the QCD predictions at NLO+NLL (above) and NNLO (below). The data points are shown by the black markers, with total bin-to-bin uncorrelated uncertainties shown by vertical error bars, while the total experimental uncertainty is shown with the orange band around unity. Alternative predictions are shown with thick, coloured lines; theoretical uncertainties are shown with thin red lines.

References

[1] J.C. Collins, D.E. Soper and G.F. Sterman, *Factorization of Hard Processes in QCD*, Adv. Ser. Direct. High Energy Phys. **5** (1989) 1 [hep-ph/0409313].

[2] J. Currie, E.W.N. Glover and J. Pires, *Next-to-next-to leading order QCD predictions for single jet inclusive production at the LHC*, Phys. Rev. Lett. **118** (2017) 072002 [1611.01460].

[3] J. Currie, E.W.N. Glover, T. Gehrmann, A. Gehrmann-De Ridder, A. Huss and J. Pires, *Single jet inclusive production for the individual jet p_t scale choice at the LHC*, Acta Phys. Polon. B **48** (2017) 955 [1704.00923].

[4] T. Gehrmann et al., *Jet cross sections and transverse momentum distributions with NNLOJET*, PoS **RADCOR2017** (2018) 074 [1801.06415].

[5] X. Liu, S.-O. Moch and F. Ringer, *Phenomenology of single-inclusive jet production with jet radius and threshold resummation*, Phys. Rev. D **97** (2018) 056026 [1801.07284].

[6] CMS Collaboration, *Measurement and QCD analysis of double-differential inclusive jet cross sections in proton-proton collisions at $\sqrt{s}$ = 13 TeV*, JHEP **02** (2022) 142 [2111.10431].

[7] CMS Collaboration, *The CMS experiment at the CERN LHC*, *JINST* **3** (2008) S08004.

[8] M. Cacciari, G.P. Salam and G. Soyez, *The anti-k_t jet clustering algorithm*, *JHEP* **04** (2008) 063 [0802.1189].

[9] M. Cacciari, G.P. Salam and G. Soyez, *Fastjet user manual*, *Eur. Phys. J. C* **72** (2012) 1896 [1111.6097].

[10] CMS Collaboration, *Measurement and qcd analysis of double-differential inclusive jet cross sections in pp collisions at $\sqrt{s}$ = 8 TeV and cross section ratios to 2.76 and 7 TeV*, *JHEP* **03** (2017) 156 [1609.05331].

[11] CMS Collaboration, *The CMS trigger system*, *JINST* **12** (2017) P01020 [1609.02366].

[12] CMS Collaboration, *Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV*, *JINST* **12** (2017) P02014 [1607.03663].

[13] CMS Collaboration, *Jet algorithms performance in 13 TeV data*, CMS Physics Analysis Summary CMS-PAS-JME-16-003 (2017).

[14] S. Schmitt, *TUnfold: an algorithm for correcting migration effects in high energy physics*, *JINST* **7** (2012) T10003 [1205.6201].

[15] S. Schmitt, *Data Unfolding Methods in High Energy Physics*, *EPJ Web Conf.* **137** (2017) 11008 [1611.01927].

[16] G. D'Agostini, *A Multidimensional unfolding method based on Bayes' theorem*, *Nucl. Instrum. Meth.* **A362** (1995) 487.

[17] G. D'Agostini, *Improved iterative Bayesian unfolding*, *ArXiv e-prints* (2010) [1010.0632].

[18] P.L.S. Connor and R. Žlebčík, *Step: a tool to perform tests of smoothness on differential distributions based on expansion of polynomials*, 2111.09968.

[19] H1 and ZEUS Collaborations, *Combination of measurements of inclusive deep inelastic $e^{\pm}p$ scattering cross sections and QCD analysis of HERA data*, *Eur. Phys. J. C* **75** (2015) 580 [1506.06042].

[20] S. Dulat, T.-J. Hou, J. Gao, M. Guzzi, J. Huston, P. Nadolsky et al., *New parton distribution functions from a global analysis of quantum chromodynamics*, *Phys. Rev. D* **93** (2016) 033006 [1506.07443].

[21] NNPDF Collaboration, *Parton distributions from high-precision collider data*, *Eur. Phys. J. C* **77** (2017) 663 [1706.00428].

[22] L.A. Harland-Lang, A.D. Martin, P. Motylinski and R.S. Thorne, *Parton distributions in the LHC era: MMHT 2014 PDFs*, *Eur. Phys. J. C* **75** (2015) 204 [1412.3989].

[23] S. Alekhin, J. Blümlein, S. Moch and R. Placakyte, *Parton distribution functions, α_s, and heavy-quark masses for LHC run II*, *Phys. Rev. D* **96** (2017) 014011 [1701.05838].

[24] D. Britzger et al., *NNLO interpolation grids for jet production at the LHC*, 2207.13735.

PoS(ICHEP2022)774