Constraints on parton distribution functions and extraction of the strong coupling constant from the inclusive jet cross section in pp collisions at s = 7 TeV


The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.
Constraints on parton distribution functions and extraction of the strong coupling constant from the inclusive jet cross section in pp collisions at \( \sqrt{s} = 7 \) TeV

CMS Collaboration

CERN, 1211 Geneva 23, Switzerland

Received: 24 October 2014 / Accepted: 2 June 2015 / Published online: 26 June 2015
© CERN for the benefit of the CMS collaboration 2015. This article is published with open access at Springerlink.com

Abstract The inclusive jet cross section for proton–proton collisions at a centre-of-mass energy of 7 TeV was measured by the CMS Collaboration at the LHC with data corresponding to an integrated luminosity of \( 5.0 \) fb\(^{-1}\). The measurement covers a phase space up to 2 TeV in jet transverse momentum and 2.5 in absolute jet rapidity. The statistical precision of these data leads to stringent constraints on the parton distribution functions of the proton. The data provide important input for the gluon density at high fractions of the proton momentum and for the strong coupling constant at large energy scales. Using predictions from perturbative quantum chromodynamics at next-to-leading order, complemented with electroweak corrections, the constraining power of these data is investigated and the strong coupling constant at the Z boson mass \( \alpha_S(M_Z) \) is determined to be \( \alpha_S(M_Z) = 0.1185 \pm 0.0019 \) (exp)\(^{+0.0060}_{-0.0037}\) (theo), which is in agreement with the world average.

1 Introduction

Collimated streams of particles, conventionally called jets, are abundantly produced in highly energetic proton–proton collisions at the LHC. At high transverse momenta \( p_T \) these collisions are described by quantum chromodynamics (QCD) using perturbative techniques (pQCD). Indispensable ingredients for QCD predictions of cross sections in pp collisions are the proton structure, expressed in terms of parton distribution functions (PDFs), and the strong coupling constant \( \alpha_S \), which is a fundamental parameter of QCD. The PDFs and \( \alpha_S \) both depend on the relevant energy scale \( Q \) of the scattering process, which is identified with the jet \( p_T \) for the reactions considered in this report. In addition, the PDFs, defined for each type of parton, depend on the fractional momentum \( x \) of the proton carried by the parton.

The large cross section for jet production at the LHC and the unprecedented experimental precision of the jet measurements allow stringent tests of QCD. In this study, the theory is confronted with data in previously inaccessible phase space regions of \( Q \) and \( x \). When jet production cross sections are combined with inclusive data from deep-inelastic scattering (DIS), the gluon PDF for \( x \gtrsim 0.01 \) can be constrained and \( \alpha_S(M_Z) \) can be determined. In the present analysis, this is demonstrated by means of the CMS measurement of inclusive jet production [1]. The data, collected in 2011 and corresponding to an integrated luminosity of \( 5.0 \) fb\(^{-1}\), extend the accessible phase space in jet \( p_T \) up to 2 TeV, and range up to \( |y| = 2.5 \) in absolute jet rapidity. A PDF study using inclusive jet measurements by the ATLAS Collaboration is described in Ref. [2].

This paper is divided into six parts. Sect. 2 presents an overview of the CMS detector and of the measurement, published in Ref. [1], and proposes a modified treatment of correlations in the experimental uncertainties. Theoretical ingredients are introduced in Sect. 3. Section 4 is dedicated to the determination of \( \alpha_S \) at the scale of the Z-boson mass \( M_Z \), and in Sect. 5 the influence of the jet data on the PDFs is discussed. A summary is presented in Sect. 6.

2 The inclusive jet cross section

2.1 Overview of the CMS detector and of the measurement

The central feature of the CMS detector is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 T. Within the superconducting solenoid volume are a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass/scintillator hadron calorimeter, each composed of a barrel and two endcap sections. Muons are measured in gas-ionisation detectors embedded in the steel flux-return yoke outside the solenoid.
Extensive forward calorimetry (HF) complements the coverage provided by the barrel and endcap detectors. A more detailed description of the CMS detector, together with a definition of the coordinate system used and the relevant kinematic variables, can be found in Ref. [3].

Jets are reconstructed with a size parameter of $R = 0.7$ using the collinear- and infrared-safe anti-$k_T$ clustering algorithm [4] as implemented in the FASTJET package [5]. The published measurements of the cross sections were corrected for detector effects, and include statistical and systematic experimental uncertainties as well as bin-to-bin correlations for each type of uncertainty. A complete description of the measurement can be found in Ref. [1].

The double-differential inclusive jet cross section investigated in the following is derived from observed inclusive jet yields via

$$\frac{d^2\sigma}{dp_T dy} = \frac{1}{\epsilon \cdot \mathcal{L}_{\text{int}} \cdot \Delta p_T} \frac{N_{\text{jets}}}{(2 \cdot |y|)},$$

where $N_{\text{jets}}$ is the number of jets in the specific kinematic range (bin), $\mathcal{L}_{\text{int}}$ is the integrated luminosity, $\epsilon$ is the product of trigger and event selection efficiencies, and $\Delta p_T$ and $\Delta |y|$ are the bin widths in $p_T$ and $|y|$. The factor of two reflects the folding of the distributions around $y = 0$.

2.2 Experimental uncertainties

The inclusive jet cross section is measured in five equally sized bins of $\Delta |y| = 0.5$ up to an absolute rapidity of $|y| = 2.5$. The inner three regions roughly correspond to the barrel part of the detector, the outer two to the endcaps. Tracker coverage extends up to $|y| = 2.4$. The minimum $p_T$ imposed on any jet is 114 GeV. The binning in jet $p_T$ follows the jet $p_T$ resolution of the central detector and changes with $p_T$. The upper reach in $p_T$ is given by the available data and decreases with $|y|$.

Four categories [1] of experimental uncertainties are defined: the jet energy scale (JES), the luminosity, the corrections for detector response and resolution, and all remaining uncorrelated effects.

The JES is the dominant source of systematic uncertainty, because a small shift in the measured $p_T$ translates into a large uncertainty in the steeply falling jet $p_T$ spectrum and hence in the cross section for any given value of $p_T$. The JES uncertainty is parameterized in terms of jet $p_T$ and pseudorapidity $\eta \equiv -\ln \tan(\theta/2)$ and amounts to 1–2% [6], which translates into a 5–25% uncertainty in the cross section. Because of its particular importance for this analysis, more details are given in Sect. 2.3.

The uncertainty in the integrated luminosity is 2.2% [7] and translates into a normalisation uncertainty that is fully correlated across $|y|$ and $p_T$.

The effect of the jet energy resolution (JER) is corrected for using the D’Agostini method [8] as implemented in the ROOUnfold package [9]. The uncertainty due to the unfolding comprises the effects of an imprecise knowledge of the JER, of residual differences between data and the Monte Carlo (MC) modelling of detector response, and of the unfolding technique applied. The total unfolding uncertainty, which is fully correlated across $\eta$ and $p_T$, is 3–4%.

Additionally, the statistical uncertainties are propagated through the unfolding procedure, thereby providing the correlations between the statistical uncertainties of the unfolded measurement. A statistical covariance matrix must be used to take this into account.

Remaining effects are collected into an uncorrelated uncertainty of $\approx 1\%$.

2.3 Uncertainties in JES

The procedure to calibrate jet energies in CMS and ways to estimate JES uncertainties are described in Ref. [10]. To use CMS data in fits of PDFs or $\alpha_S(M_Z)$, it is essential to account for the correlations in these uncertainties among different regions of the detector. The treatment of correlations uses 16 mutually uncorrelated sources as in Ref. [1]. Within each source, the uncertainties are fully correlated in $p_T$ and $\eta$. Any change in the jet energy calibration (JEC) is described through a linear combination of sources, where each source is assumed to have a Gaussian probability density with a zero mean and a root-mean-square of unity. In this way, the uncertainty correlations are encoded in a fashion similar to that provided for PDF uncertainties using the Hessian method [11]. The total uncertainty is defined through the quadratic sum of all uncertainties. The full list of sources together with their brief descriptions can be found in Appendix A.

The JES uncertainties can be classified into four broad categories: absolute energy scale as a function of $p_T$, jet flavour dependent differences, relative calibration of JES as a function of $\eta$, and the effects of multiple proton interactions in the same or adjacent beam crossings (pileup). The absolute scale is a single fixed number such that the corresponding uncertainty is fully correlated across $p_T$ and $\eta$. Using photon + jet and $Z +$ jet data, the JES can be constrained directly in the jet $p_T$ range 30–600 GeV. The response at larger and smaller $p_T$ is extrapolated through MC simulation. Extra uncertainties are assigned to this extrapolation based on the differences between MC event generators and the single-particle response of the detector. The absolute calibration is the most relevant uncertainty in jet analyses at large $p_T$.

The categories involving jet flavour dependence and pileup effects are important mainly at small $p_T$ and have relatively little impact for the phase space considered in this report.
The third category parameterizes $\eta$-dependent changes in relative JES. The measurement uncertainties within different detector regions are strongly correlated, and thus the $\eta$-dependent sources are only provided for wide regions: barrel, endcap with upstream tracking, endcap without upstream tracking, and the HF calorimeter. In principle, the $\eta$-dependent effects can also have a $p_T$ dependence. Based on systematic studies on data and simulated events, which indicate that the $p_T$ and $\eta$ dependence of the uncertainties factorise to a good approximation, this is omitted from the initial calibration procedure. However, experiences with the calibration of data collected in 2012 and with fits of $\alpha_S(M_Z)$ reported in Sect. 4 show that this is too strong an assumption. Applying the uncertainties and correlations in a fit of $\alpha_S(M_Z)$ to the inclusive jet data separately for each bin in $|y|$ leads to results with values of $\alpha_S(M_Z)$ that scatter around a central value. Performing the same fit taking all $|y|$ bins together and assuming 100% correlation in $|y|$ within the JES uncertainty sources results in a bad fit quality (high $\chi^2$ per number of degrees of freedom $n_{dof}$) and a value of $\alpha_S(M_Z)$ that is significantly higher than any value observed for an individual bin in $|y|$. Changing the correlation in the JES uncertainty from 0 to 100% produces a steep rise in $\chi^2/n_{dof}$, and influences the fitted value of $\alpha_S(M_Z)$ for correlations near 90%, indicating an assumption on the correlations in $|y|$ that is too strong. The technique of nuisance parameters, as described in Sect. 5.2.2, helped in the analysis of this issue.

To implement the additional $\eta$-decorrelation induced by the $p_T$-dependence in the $\eta$-dependent JEC introduced for the calibration of 2012 data, the source from the single-particle response JEC2, which accounts for extrapolation uncertainties at large $p_T$ as discussed in Appendix A, is decorrelated versus $\eta$ as follows:

1. in the barrel region ($|y| < 1.5$), the correlation of the single-particle response source among the three bins in $|y|$ is set to 50%,
2. in the endcap region ($1.5 \leq |y| < 2.5$), the correlation of the single-particle response source between the two bins in $|y|$ is kept at 100%,
3. there is no correlation of the single-particle response source between the two detector regions of $|y| < 1.5$ and $1.5 \leq |y| < 2.5$.

The additional freedom of $p_T$-dependent corrections versus $\eta$ hence leads to a modification of the previously assumed full correlation between all $\eta$ regions to a reduced estimate of 50% correlation of JEC2 within the barrel region, which always contains the tag jet of the dijet balance method [10]. In addition, the JEC2 corrections are estimated to be uncorrelated between the barrel and endcap regions of the detector because of respective separate $p_T$-dependences of these corrections.

Technically, this can be achieved by splitting the single-particle response source into five parts (JEC2a–e), as shown in Table 8. Each of these sources is a duplicate of the original single-particle response source that is set to zero outside the respective ranges of $|y| < 1.5$, $1.5 \leq |y| < 2.5$, $|y| < 0.5$, $0.5 \leq |y| < 1.0$, and $1.0 \leq |y| < 1.5$, such that the original full correlation of

$$\text{corr}_{\text{JEC2,old}} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

is replaced by the partially uncorrelated version of

$$\text{corr}_{\text{JEC2,new}} = \begin{pmatrix} 1 & 0.5 & 0.5 & 0 & 0 \\ 0.5 & 1 & 0.5 & 0 & 0 \\ 0.5 & 0.5 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

which is more accurate as justified by studies based on 2012 data. For the proper normalisation of the five new correlated sources, normalisation factors of 1, 0.5, 0.5, and 1 (JEC2a–JEC2d) and 1 (JEC2b) must be applied. With these factors, the sum of the five sources reproduces the original uncertainty for each $|y|$, while the additional freedom gives the estimated level of correlation among the $|y|$ regions.

All results presented in this paper are based on this improved treatment of the correlation of JES uncertainties. While some decorrelation of these uncertainties versus $\eta$ is important for the fits of $\alpha_S(M_Z)$ described in Sect. 4, the exact size of the estimated decorrelation is not. Varying the assumptions according to Eq. (3) from 50% to 20 or 80% in the barrel region, from 100 to 80% in the endcap region, or from 0 to 20% between the barrel and endcap region leads to changes in the fitted value of $\alpha_S(M_Z)$ that are negligible with respect to other experimental uncertainties.

### 3 Theoretical ingredients

The theoretical predictions for the inclusive jet cross section comprise a next-to-leading order (NLO) $p$QCD calculation with electroweak corrections (EW) [12, 13]. They are complemented by a nonperturbative (NP) factor that corrects for multiple-parton interactions (MPI) and hadronization (HAD) effects. Parton shower (PS) corrections, derived from NLO predictions with matched parton showers, are tested in an additional study in Sect. 4.3, but are not applied to the main result.
Table 1 The PDF sets used in comparisons to the data together with the evolution order (Evol.), the corresponding number of active flavours $N_f$, the assumed masses $M_t$ and $M_Z$ of the top quark and the Z boson, respectively, the default values of $\alpha_S(M_Z)$, and the range in $\alpha_S(M_Z)$ variation available for fits. For CT10 the updated versions of 2012 are taken.

<table>
<thead>
<tr>
<th>Base set</th>
<th>Refs.</th>
<th>Evol.</th>
<th>$N_f$</th>
<th>$M_t$ (GeV)</th>
<th>$M_Z$ (GeV)</th>
<th>$\alpha_S(M_Z)$</th>
<th>$\alpha_S(M_Z)$ range</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABM11</td>
<td>[17]</td>
<td>NLO</td>
<td>5</td>
<td>180</td>
<td>91.174</td>
<td>0.1180</td>
<td>0.110–0.130</td>
</tr>
<tr>
<td>ABM11</td>
<td>[17]</td>
<td>NNLO</td>
<td>5</td>
<td>180</td>
<td>91.174</td>
<td>0.1134</td>
<td>0.104–0.120</td>
</tr>
<tr>
<td>CT10</td>
<td>[18]</td>
<td>NLO</td>
<td>$\leq 5$</td>
<td>172</td>
<td>91.188</td>
<td>0.1180</td>
<td>0.112–0.127</td>
</tr>
<tr>
<td>CT10</td>
<td>[18]</td>
<td>NNLO</td>
<td>$\leq 5$</td>
<td>172</td>
<td>91.188</td>
<td>0.1180</td>
<td>0.110–0.130</td>
</tr>
<tr>
<td>HERAPDF1.5</td>
<td>[19]</td>
<td>NLO</td>
<td>$\leq 5$</td>
<td>180</td>
<td>91.187</td>
<td>0.1176</td>
<td>0.114–0.122</td>
</tr>
<tr>
<td>HERAPDF1.5</td>
<td>[19]</td>
<td>NNLO</td>
<td>$\leq 5$</td>
<td>180</td>
<td>91.187</td>
<td>0.1176</td>
<td>0.114–0.122</td>
</tr>
<tr>
<td>MSTW2008</td>
<td>[20,21]</td>
<td>NLO</td>
<td>$\leq 5$</td>
<td>$10^{10}$</td>
<td>91.1876</td>
<td>0.1202</td>
<td>0.110–0.130</td>
</tr>
<tr>
<td>MSTW2008</td>
<td>[20,21]</td>
<td>NNLO</td>
<td>$\leq 5$</td>
<td>$10^{10}$</td>
<td>91.1876</td>
<td>0.1171</td>
<td>0.107–0.127</td>
</tr>
<tr>
<td>NNPDF2.1</td>
<td>[22]</td>
<td>NLO</td>
<td>$\leq 6$</td>
<td>175</td>
<td>91.2</td>
<td>0.1190</td>
<td>0.114–0.124</td>
</tr>
<tr>
<td>NNPDF2.1</td>
<td>[22]</td>
<td>NNLO</td>
<td>$\leq 6$</td>
<td>175</td>
<td>91.2</td>
<td>0.1190</td>
<td>0.114–0.124</td>
</tr>
</tbody>
</table>

3.1 Fixed-order prediction in perturbative QCD

The same NLO prediction as in Ref. [1] is used, i.e. the calculations are based on the parton-level program NLOJET++ version 4.1.3 [14,15] and are performed within the FASTNLO framework version 2.1 [16]. The renormalization and factorisation scales, $\mu_r$ and $\mu_f$ respectively, are identified with the individual jet $p_T$. The number of active (massless) flavours $N_f$ in NLOJET++ has been set to five.

Five sets of PDFs are available for a series of values of $\alpha_S(M_Z)$, which is a requisite for a determination of $\alpha_S(M_Z)$ from data. For an overview, these PDF sets are listed in Table 1 together with the respective references. The ABM11 PDF set employs a fixed-flavour number scheme with five active flavours, while the other PDF sets use a variable-flavour number scheme with a maximum of five flavours, $N_f, \text{max} = 5$, except for NNPDF2.1 which has $N_f, \text{max} = 6$.

All sets exist at next-to-leading and next-to-next-to-leading evolution order. The PDF uncertainties are provided at 68.3% confidence level (CL) except for CT10, which provides uncertainties at 90% CL. For a uniform treatment of all PDFs, the CT10 uncertainties are down-scaled by a factor of $\sqrt{2}\text{erf}^{-1}(0.9) \approx 1.645$.

The electroweak corrections to the hard-scattering cross section have been computed with the CT10-NLO PDF set for a fixed number of five flavours and with the $p_T$ of the leading jet, $p_T, \text{max}$, as scale choice for $\mu_r$ and $\mu_f$ instead of the $p_T$ of each jet. At high jet $p_T$ and central rapidity, where the electroweak effects become sizeable, NLO calculations with either of the two scale settings differ by less than one percent. Given the small impact of the electroweak corrections on the final results in Sects. 4 and 5, no uncertainty on their size has been assigned.

3.2 Theoretical prediction from MC simulations including parton showers and nonperturbative effects

The most precise theoretical predictions for jet measurements are usually achieved in fixed-order pQCD, but are available at parton level only. Data that have been corrected for detector effects, however, refer to measurable particles, i.e. to colour-neutral particles with mean decay lengths such that $c\tau > 10 \text{ mm}$. Two complications arise when comparing fixed-order perturbation theory to these measurements: emissions of additional partons close in phase space, which are not sufficiently accounted for in low-order approximations, and effects that cannot be treated by perturbative methods. The first problem is addressed by the parton shower concept [23–25] within pQCD, where multiple parton radiation close in phase space is taken into account through an all-orders approximation of the dominant terms including coherence effects. Avoiding double counting, these parton showers are combined with leading-order (LO) calculations in MC event generators, such as PYTHIA [26] and HERWIG++ [27].

The second issue concerns NP corrections, which comprise supplementary parton-parton scatters within the same colliding protons, i.e. MPI, and the hadronization process including particle decays. The MPI [28,29] model for additional soft-particle production, which is detected as part of the underlying event, is implemented in PYTHIA as well as HERWIG++. Hadronization describes the transition phase from coloured partons to colour-neutral particles, where perturbative methods are no longer applicable. Two models for hadronization are in common use, the Lund string fragmentation [30–32] that is used in PYTHIA, and the cluster fragmentation [33] that has been adopted by HERWIG++.
Beyond LO combining fixed-order predictions with parton showers, MPI, and hadronization models is much more complicated. Potential double counting of terms in the perturbative expansion and the PS has to be avoided. In recent years programs have become available for dijet production at NLO that can be matched to PS MC event generators. In the following, one such program, the POWHEG package [34,35] will be used for comparisons with dijet events [36] to the LO MC event generators.

3.3 NP corrections from PYTHIA6 and HERWIG++

For the comparison of theoretical predictions to the measurement reported in Ref. [1], the NP correction was derived as usual [37] from the average prediction of two LO MC event generators and more specifically from PYTHIA version 6.4.22 tune Z2 and HERWIG++ version 2.4.2 with the default tune of version 2.3. Tune Z2 is identical to tune Z1 described in [38] except that Z2 employs the CTEQ6L1 [39] PDF set, while Z1 uses the CTEQ5L [40] PDF set. The NP correction factor can be defined for each bin in \( p_T \) and \( |y| \) as

\[
C_{\text{NP}} = \frac{\sigma_{\text{LO+PS+HAD+MPI}}}{\sigma_{\text{LO+PS}}} \quad (4)
\]

where \( \sigma \) represents the inclusive jet cross section and the subscripts “LO+PS+HAD+MPI” and “LO+PS” indicate which steps of a general MC event generation procedure have been run, see also Refs. [37,41]. The central value is calculated by taking the average of the two predictions from PYTHIA6 and HERWIG++.

In applying these factors as corrections for NP effects to NLO theory predictions, it is assumed that the NP corrections are universal, i.e. they are similar for LO and NLO.

3.4 NP and PS corrections from POWHEG + PYTHIA6

Alternative corrections are derived, which use the POWHEG BOX revision 197 with the CT10-NLO PDF set for the hard subprocess at NLO plus the leading emission [42] complemented with the matched showering, MPI, and hadronization from PYTHIA6 version 6.4.26. The NLO event generation within the POWHEG framework, and the showering and hadronization process performed by PYTHIA6 are done in independent steps.

For illustration, Fig. 1 shows the comparison of the inclusive jet data with the POWHEG + PYTHIA6 tune Z2* particle-level prediction complemented with electroweak corrections. The tune Z2* is derived from the earlier tune Z2, where the PYTHIA6 parameters PARP(82) and PARP(90) that control the energy dependence of the MPI are retuned, yielding 1.921 and 0.227, respectively. The error boxes indicate statistical uncertainties. Ratio plots of this comparison for each separate region in \( |y| \) can be found in Appendix B.

The corrections to NLO parton-level calculations that are derived this way consist of truly nonperturbative contributions, which are optionally complemented with parton shower effects. They are investigated separately in the following two sections. A previous investigation can be found in Ref. [43].

3.4.1 NP corrections from POWHEG + PYTHIA6

The NP corrections using a NLO prediction with a matched PS event generator can be defined analogously as in Eq. (4):

\[
C_{\text{NP}} = \frac{\sigma_{\text{NLO+PS+HAD+MPI}}}{\sigma_{\text{NLO+PS}}} \quad (5)
\]

i.e. the numerator of this NP correction is defined by the inclusive cross section, where parton showers, hadronization, and multiparton interactions are turned on, while the inclusive cross section in the denominator does not include hadronization and multiparton interactions. A NLO calculation can then be corrected for NP effects as

\[
\frac{d^2\sigma}{dp_T dy} = \frac{d^2\sigma_{\text{NLO}}}{dp_T dy} \cdot C_{\text{NP}} \quad (6)
\]

In contrast to the LO MC event generation with PYTHIA6, the parameters of the NP and PS models, however, have not been retuned to data for the use with NLO + PS predictions by POWHEG. Therefore two different underlying event tunes of PYTHIA6 for LO + PS predictions, P11 [44] and Z2*, are used. In both cases a parameterization using a functional
form of $a_0 + a_1/p_T^a$ is employed to smoothen statistical fluctuations. For $p_T > 100$ GeV the difference in the NP correction factor between the two tunes is very small such that their average is taken as $C_{\text{NP}}^{\text{NLO}}$.

Since procedures to estimate uncertainties inherent to the NLO + PS matching procedure are not yet well established and proper tunes to data for POWHEG + PYTHIA6 are lacking, the centre of the envelope given by the three curves from PYTHIA6, HERWIG++, and the POWHEG + PYTHIA6 average of tunes Z2* and P11 is adopted as the final NP correction for the central results in Sects. 4 and 5. Half the spread among these three predictions defines the uncertainty.

The NP correction, as defined for POWHEG + PYTHIA6, is shown in Fig. 2 together with the original factors from PYTHIA6 and HERWIG++, as a function of the jet $p_T$ for five ranges in absolute rapidity $|y|$ of size 0.5 up to $|y| = 2.5$. The factors derived from both, LO + PS and NLO + PS MC event generators, are observed to decrease with increasing

**Fig. 2** NP corrections for the five regions in $|y|$ as derived in Ref. [1], using PYTHIA6 tune Z2 and HERWIG++ with the default tune of version 2.3, in comparison to corrections obtained from POWHEG using PYTHIA6 for showering with the two underlying event tunes P11 and Z2*.
3.4.2 PS corrections from POWHEG + PYTHIA6

Similarly to the NP correction of Eq. (5), a PS correction factor can be defined as the ratio of the differential cross section including PS effects divided by the NLO prediction, as given by POWHEG, i.e. including the leading emission:

$$C_{\text{PS}}^{\text{NLO}} = \frac{\sigma_{\text{NLO+PS}}}{\sigma_{\text{NLO}}}.$$  \hfill (7)

The combined correction for NP and PS effects can then be written as

$$\frac{d^2\sigma_{\text{theo}}}{d\pt dy} = \frac{d^2\sigma_{\text{NLO}}}{d\pt dy} \cdot C_{\text{NP}}^{\text{NLO}} \cdot C_{\text{PS}}^{\text{NLO}}.$$  \hfill (8)

Fig. 3 PS corrections for the five regions in $|y|$ obtained from POWHEG using PYTHIA6 for showering for different upper scale limits of the parton shower evolution in PYTHIA6 tune Z2*. The curves parameterize the correction factors as a function of the jet $p_T$. Within modelling uncertainties, the assumption of universal NP corrections that are similar for LO + PS and NLO + PS MC event generation holds approximately above a jet $p_T$ of a few hundred GeV.
The PS corrections derived with POWHEG + PYTHIA6 are presented in Fig. 3. They are significant at large $p_T$, particularly at high rapidity, where the factors approach $-20\%$. However, the combination of POWHEG + PYTHIA6 has never been tuned to data and the $Z^{\pm}$ tune strictly is only valid for a LO + PS tune with PYTHIA6, but not with showers matched to POWHEG. Moreover, POWHEG employs the CT10-NLO PDF, while the $Z^{\pm}$ tune requires the CTEQ6L1-LO PDF to be used for the showering part. Therefore, such PS corrections can be considered as only an illustrative test, as reported in Sect. 4.3.

The maximum parton virtuality allowed in the parton shower evolution, $\mu^2_{PS}$, is varied by factors of 0.5 and 1.5 by changing the corresponding parameter PARP(67) in PYTHIA6 from its default value of 4–2 and 6, respectively. The resulting changes in the PS factors are shown in Fig. 3. The POWHEG + PYTHIA6 PS factors employed in an illustrative test later are determined as the average of the predictions from the two extreme scale limits. Again, a parameterization using a functional form of $a_0 + a_1/\mu^2_{PS}$ is employed to smoothen statistical fluctuations.

Finally, Fig. 4 presents an overview of the NP, PS, and combined corrections for all five ranges in $|y|$.

4 Determination of the strong coupling constant

The measurement of the inclusive jet cross section [1], as described in Sect. 2, can be used to determine $\alpha_S(M_Z)$, where the proton structure in the form of PDFs is taken as a prerequisite. The necessary theoretical ingredients are specified in Sect. 3. The choice of PDF sets is restricted to global sets that fit data from different experiments, so that only the most precisely known gluon distributions are employed. Combined fits of $\alpha_S(M_Z)$ and the gluon content of the proton are investigated in Sect. 5.5.

In the following, the sensitivity of the inclusive jet cross section to $\alpha_S(M_Z)$ is demonstrated. Subsequently, the fitting procedure is given in detail before presenting the outcome of the various fits of $\alpha_S(M_Z)$.

4.1 Sensitivity of the inclusive jet cross section to $\alpha_S(M_Z)$

Figures 5, 6, 7 and 8 present the ratio of data to the theoretical predictions for all variations in $\alpha_S(M_Z)$ available for the PDF sets ABM11, CT10, MSTW2008, and NNPDF2.1 at next-to-leading evolution order, as specified in Table 1. Except for the ABM11 PDF set, which leads to QCD predictions significantly different in shape to the measurement, all PDF sets give satisfactory theoretical descriptions of the data and a strong sensitivity to $\alpha_S(M_Z)$ is demonstrated. Because of the discrepancies, ABM11 is excluded from further investigations. The CT10-NLO PDF set is chosen for the main result on $\alpha_S(M_Z)$, because the value of $\alpha_S(M_Z)$ preferred by the CMS jet data is rather close to the default value of this PDF set. As crosschecks fits are performed with the NNPDF2.1-NLO and MSTW2008-NLO sets. The CT10-NNLO, NNPDF2.1-NNLO, and MSTW2008-NNLO PDF sets are employed for comparison.
Fig. 5 Ratio of the inclusive jet cross section to theoretical predictions using the ABM11-NLO PDF set for the five rapidity bins, where the $\alpha_S(M_Z)$ value is varied in the range 0.110–0.130 in steps of 0.001. The error bars correspond to the total uncertainty.

4.2 The fitting procedure

The value of $\alpha_S(M_Z)$ is determined by minimising the $\chi^2$ between the $N$ measurements $D_i$ and the theoretical predictions $T_i$. The $\chi^2$ is defined as

$$\chi^2 = \sum_{ij} (D_i - T_i)C_{ij}^{-1}(D_j - T_j),$$

where the covariance matrix $C_{ij}$ is composed of the following terms:

$$C = \text{cov}_{\text{stat}} + \text{cov}_{\text{uncor}} + \left( \sum_{\text{sources}} \text{cov}_{\text{JES}} \right) + \text{cov}_{\text{unfolding}} + \text{cov}_{\text{lumi}} + \text{cov}_{\text{PDF}},$$

and the terms in the sum represent
Fig. 6 Ratio of the inclusive jet cross section to theoretical predictions using the CT10-NLO PDF set for the five rapidity bins, where the $\alpha_s(M_Z)$ value is varied in the range 0.112–0.126 in steps of 0.001. The error bars correspond to the total uncertainty:

1. cov\text{stat}: statistical uncertainty including correlations induced through unfolding;
2. cov\text{uncor}: uncorrelated systematic uncertainty summing up small residual effects such as trigger and identification inefficiencies, time dependence of the jet $p_T$ resolution, or the uncertainty on the trigger prescale factor;
3. cov\text{JES sources}: systematic uncertainty for each JES uncertainty source;
4. cov\text{unfolding}: systematic uncertainty of the unfolding;
5. cov\text{lumi}: luminosity uncertainty; and
6. cov\text{PDF}: PDF uncertainty.

All JES, unfolding, and luminosity uncertainties are treated as 100% correlated across the $p_T$ and $|y|$ bins, with the exception of the single-particle response JES source as described in Sect. 2.3. The JES, unfolding, and luminosity uncertainties are treated as multiplicative to avoid the sta-
Fig. 7 Ratio of the inclusive jet cross section to theoretical predictions using the MSTW2008-NLO PDF set for the five rapidity bins, where the $\alpha_s(M_Z)$ value is varied in the range 0.110–0.130 in steps of 0.001. The error bars correspond to the total uncertainty.

The statistical bias that arises when estimating uncertainties from data [45–47].

The derivation of PDF uncertainties follows prescriptions for each individual PDF set. The CT10 and MSTW PDF sets both employ the eigenvector method with upward and downward variations for each eigenvector. As required by the use of covariance matrices, symmetric PDF uncertainties are computed following Ref. [39]. The NNPDF2.1 PDF set uses the MC pseudo-experiments instead of the eigenvector method in order to provide PDF uncertainties. A hundred so-called replicas, whose averaged predictions give the central result, are evaluated following the prescription in Ref. [48] to derive the PDF uncertainty for NNPDF.

As described in Sect. 3.4.1, the NP correction is defined as the centre of the envelope given by PYTHIA6, HERWIG++, and the POWHEG + PYTHIA6 average of tunes Z2* and P11. Half the spread among these three numbers is taken as the uncertainty. This is the default NP correction used in this
Fig. 8 Ratio of the inclusive jet cross section to theoretical predictions using the NNPDF2.1-NLO PDF set for the five rapidity bins, where the $\alpha_S(M_Z)$ value is varied in the range 0.116–0.122 in steps of 0.001. The error bars correspond to the total uncertainty analysis. Alternatively, the PS correction factor, defined in Sect. 3.4.2, is applied in addition as an illustrative test to complement the main results.

The uncertainty in $\alpha_S(M_Z)$ due to the NP uncertainties is evaluated by looking for maximal offsets from a default fit. The theoretical prediction $T$ is varied by the NP uncertainty $\Delta$NP as $T \cdot NP \rightarrow T \cdot (NP \pm \Delta$NP). The fitting procedure is repeated for these variations, and the deviation from the central $\alpha_S(M_Z)$ values is considered as the uncertainty in $\alpha_S(M_Z)$.

Finally the uncertainty due to the renormalization and factorisation scales is evaluated by applying the same method as for the NP corrections: $\mu_r$ and $\mu_f$ are varied from the default choice of $\mu_r = \mu_f = p_T$ between $p_T/2$ and $2p_T$ in the following six combinations: $(\mu_r/p_T, \mu_f/p_T) = (1/2, 1/2), (1/2, 1), (1, 1/2), (1, 2), (2, 1),$ and $(2, 2)$. The $\chi^2$ minimisation with respect to $\alpha_S(M_Z)$ is repeated in each case. The contribution from the $\mu_r$ and $\mu_f$ scale variations to the uncertainty is evaluated by considering the maximal upwards and downwards deviation of $\alpha_S(M_Z)$ from the central result.
4.3 The results on $\alpha_S(M_Z)$

The values of $\alpha_S(M_Z)$ obtained with the CT10-NLO PDF set are listed in Table 2 together with the experimental, PDF, NP, and scale uncertainties for each bin in rapidity and for a simultaneous fit of all rapidity bins. To disentangle the uncertainties of experimental origin from those of the PDFs, additional fits without the latter uncertainty source are performed. An example for the evaluation of the uncertainties in a $\chi^2$ fit is shown in Fig. 9. The NP and scale uncertainties are determined via separate fits, as explained above.

For the two outer rapidity bins ($1.5 < |y| < 2.0$ and $2.0 < |y| < 2.5$) the series in values of $\alpha_S(M_Z)$ of the CT10-NLO PDF set does not reach to sufficiently low values of $\alpha_S(M_Z)$. As a consequence the shape of the $\chi^2$ curve at minimum up to $\chi^2 + 1$ cannot be determined completely. To avoid extrapolations based on a polynomial fit to the available points, the alternative $\alpha_S$ evolution code of the HOPPET package [49] is employed. This is the same evolution code as chosen for the creation of the CT10 PDF set. Replacing the original $\alpha_S$ evolution in CT10 by HOPPET, $\alpha_S(M_Z)$ can be set freely and in particular different from the default value used in a PDF set, but at the expense of losing the correlation between the value of $\alpha_S(M_Z)$ and the fitted PDFs. Downwards or upwards deviations from the lowest and highest values of $\alpha_S(M_Z)$, respectively, provided in a PDF series are accepted for uncertainty evaluations up to a limit of $|\Delta \alpha_S(M_Z)| = 0.003$. Applying this method for comparisons, within the available range of $\alpha_S(M_Z)$ values, an additional uncertainty is estimated to be negligible.

For comparison the CT10-NNLO PDF set is used for the determination of $\alpha_S(M_Z)$. These results are presented in Table 3.

The final result using all rapidity bins and the CT10-NLO PDF set is (last row of Table 2)

| $|y|$ range | No. of data points | $\alpha_S(M_Z)$ | $\chi^2/n_{dof}$ |
|------------|-------------------|----------------|-----------------|
| $|y| < 0.5$ | 33                | $0.1189 \pm 0.0024 \text{(exp)} \pm 0.0030 \text{(PDF)}$ | $16.2/32$ |
|           |                   | $\pm 0.0008 \text{(NP)}^{+0.0045}_{-0.0027} \text{(scale)}$ |                 |
| $0.5 \leq |y| < 1.0$ | 30                | $0.1182 \pm 0.0024 \text{(exp)} \pm 0.0029 \text{(PDF)}$ | $25.4/29$ |
|           |                   | $\pm 0.0008 \text{(NP)}^{+0.0050}_{-0.0035} \text{(scale)}$ |                 |
| $1.0 \leq |y| < 1.5$ | 27                | $0.1165 \pm 0.0027 \text{(exp)} \pm 0.0024 \text{(PDF)}$ | $9.5/26$ |
|           |                   | $\pm 0.0008 \text{(NP)}^{+0.0043}_{-0.0020} \text{(scale)}$ |                 |
| $1.5 \leq |y| < 2.0$ | 24                | $0.1146 \pm 0.0035 \text{(exp)} \pm 0.0031 \text{(PDF)}$ | $20.2/23$ |
|           |                   | $\pm 0.0013 \text{(NP)}^{+0.0037}_{-0.0020} \text{(scale)}$ |                 |
| $2.0 \leq |y| < 2.5$ | 19                | $0.1161 \pm 0.0045 \text{(exp)} \pm 0.0054 \text{(PDF)}$ | $12.6/18$ |
|           |                   | $\pm 0.0015 \text{(NP)}^{+0.0034}_{-0.0032} \text{(scale)}$ |                 |
| $|y| < 2.5$ | 133               | $0.1185 \pm 0.0019 \text{(exp)} \pm 0.0028 \text{(PDF)}$ | $104.1/132$ |
|           |                   | $\pm 0.0004 \text{(NP)}^{+0.0053}_{-0.0024} \text{(scale)}$ |                 |

Fig. 9 The $\chi^2$ minimisation with respect to $\alpha_S(M_Z)$ using the CT10-NLO PDF set and data from all rapidity bins. The experimental uncertainty is obtained from the $\alpha_S(M_Z)$ values for which $\chi^2$ is increased by one with respect to the minimum value, indicated by the dashed line. The curve corresponds to a second-degree polynomial fit through the available $\chi^2$ points.

$$\alpha_S(M_Z) = 0.1185 \pm 0.0019 \text{(exp)}$$
$$\pm 0.0028 \text{(PDF)} \pm 0.0004 \text{(NP)}^{+0.0053}_{-0.0024} \text{(scale)}$$
$$= 0.1185 \pm 0.0034 \text{(all except scale)}^{+0.0053}_{-0.0024} \text{(scale)}$$
$$= 0.1185^{+0.0063}_{-0.0042}. \quad (11)$$

where experimental, PDF, NP, and scale uncertainties have been added quadratically to give the total uncertainty. The result is in agreement with the world average value of $\alpha_S(M_Z) = 0.1185 \pm 0.0006$ [50], with the Tevatron results [51–53], and recent results obtained with LHC data [54–56]. The determination of $\alpha_S(M_Z)$, which is based on the CT10-NLO PDF set, is also in agreement with the result obtained using the NNPDF2.1-NLO and MSTW2008-NLO sets, as shown in Table 4. For comparison this table also shows the results using the CT10, MSTW2008, and NNPDF2.1 PDF.
The results are obtained by a simultaneous fit in all rapidity bins and the CT10-NLO PDF set. The last row presents the result of a simultaneous fit in all rapidity bins. The α values are determined with the H1 [57–59], ZEUS [60], and D0 [52,53] collaborations. These extract the α values at lower scales determined by the H1 [57–59], ZEUS [60], and D0 [52,53] collaborations are shown for comparison. Recent CMS measurements [55,56], which are in agreement with the α(MZ) determination of this study, are displayed as well. The results on α(MZ) reported here are consistent with the energy dependence predicted by the RGE.

<table>
<thead>
<tr>
<th>[y] range</th>
<th>No. of data points</th>
<th>α(MZ)</th>
<th>χ²/ndof</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[y] &lt; 0.5</td>
<td>33</td>
<td>0.1180 ± 0.0017 (exp) ± 0.0027 (PDF)</td>
<td>15.4/32</td>
</tr>
<tr>
<td>0.5 ≤</td>
<td>0.0006 (NP) +0.0031</td>
<td>23.9/29</td>
<td></td>
</tr>
<tr>
<td>1.0 ≤</td>
<td>0.0006 (NP) +0.0033</td>
<td>10.5/26</td>
<td></td>
</tr>
<tr>
<td>1.5 ≤</td>
<td>0.0006 (NP) +0.0039</td>
<td>22.3/23</td>
<td></td>
</tr>
<tr>
<td>2.0 ≤</td>
<td>0.0010 (NP) +0.0044</td>
<td>13.8/18</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.0004 (NP) +0.0044</td>
<td>105.7/132</td>
</tr>
</tbody>
</table>

### Table 4

<table>
<thead>
<tr>
<th>PDF set</th>
<th>α(MZ)</th>
<th>χ²/ndof</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CT10-NLO</td>
<td>0.1185 ± 0.0019 (exp) ± 0.0028 (PDF)</td>
<td>104.1/132</td>
</tr>
<tr>
<td></td>
<td>± 0.0004 (NP) +0.0051</td>
<td>103.5/132</td>
</tr>
<tr>
<td></td>
<td>± 0.0003 (NP) +0.0051</td>
<td>107.9/132</td>
</tr>
<tr>
<td></td>
<td>± 0.0001 (NP) +0.0044</td>
<td>105.7/132</td>
</tr>
<tr>
<td></td>
<td>± 0.0004 (NP) +0.0044</td>
<td>103.0/132</td>
</tr>
<tr>
<td></td>
<td>± 0.0001 (NP) +0.0044</td>
<td>108.8/132</td>
</tr>
</tbody>
</table>

The α(MZ) values are in agreement among the different PDF sets within the uncertainties.

Applying the PS correction factor to the NLO theory prediction in addition to the NP correction as discussed in Sect. 3.4.2, the fit using all rapidity bins and the CT10-NLO PDF set yields α(MZ) = 0.1204 ± 0.0018 (exp). This value is in agreement with our main result of Eq. (11), which is obtained using only the NP correction factor.

To investigate the running of the strong coupling, the fitted region is split into six bins of pt and the fitting procedure is repeated in each of these bins. The six extractions of α(MZ) are reported in Table 5. The α(MZ) values are evolved to the corresponding energy scale \( Q \) using the two-loop solution to the renormalization group equation (RGE) within HOPPET. The value of \( Q \) is calculated as a cross section weighted average in each fit region. These average scale values \( Q \), derived again with the FASTNLO framework, are identical within about 1 GeV for different PDFs. To emphasise that theoretical uncertainties limit the achievable precision, Tables 6 and 7 present for the six bins in pt the total uncertainty as well as the experimental, PDF, and scale components, where the six experimental uncertainties are all correlated.

Figure 10 presents the running of the strong coupling \( \alpha_S(Q) \) and its total uncertainty as determined in this analysis. The extractions of \( \alpha_S(Q) \) in six separate ranges of \( Q \), as presented in Table 5, are also shown. In the same figure the values of \( \alpha_S \) at lower scales determined by the H1 [57–59], ZEUS [60], and D0 [52,53] collaborations are shown for comparison. Recent CMS measurements [55,56], which are in agreement with the \( \alpha_S(M_Z) \) determination of this study, are displayed as well. The results on \( \alpha_S \) reported here are consistent with the energy dependence predicted by the RGE.
measurements are displayed as well. The uncertainties represent error bars set for different processes and phenomenological frameworks designed among other things to fit PDFs to data. Theory is used at the determination of the strong coupling constant $\alpha_s$ as presented in Table 5 are shown in Fig. 10. The strong coupling $\alpha_s(Q)$ (full line) and its total uncertainty (band) as determined in this analysis using a two-loop solution to the RGE as a function of the momentum transfer $Q = p_T$. The extractions of $\alpha_s(Q)$ in six separate ranges of $Q$ as shown in Table 5 are shown together with results from the H1 [58, 59], ZEUS [60], and D0 [52, 53] experiments at the HERA and Tevatron colliders. Other recent CMS measurements [55, 56] are displayed as well. The uncertainties represented by error bars are subject to correlations.

### Table 5 Determination of $\alpha_s$ in separate bins of jet $p_T$ using the CT10-NLO PDF set

<table>
<thead>
<tr>
<th>$p_T$ range (GeV)</th>
<th>$Q$ (GeV)</th>
<th>$\alpha_s(M_Z)$</th>
<th>$\alpha_s(Q)$</th>
<th>No. of data points</th>
<th>$\chi^2/\text{dof}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>114–196</td>
<td>136</td>
<td>0.1172$\pm$0.0058</td>
<td>0.1106$\pm$0.0052</td>
<td>20</td>
<td>6.2/19</td>
</tr>
<tr>
<td>196–300</td>
<td>226</td>
<td>0.1180$\pm$0.0063</td>
<td>0.1038$\pm$0.0048</td>
<td>20</td>
<td>7.6/19</td>
</tr>
<tr>
<td>300–468</td>
<td>345</td>
<td>0.1194$\pm$0.0064</td>
<td>0.0993$\pm$0.0044</td>
<td>25</td>
<td>8.1/24</td>
</tr>
<tr>
<td>468–638</td>
<td>521</td>
<td>0.1187$\pm$0.0055</td>
<td>0.0940$\pm$0.0032</td>
<td>20</td>
<td>10.6/19</td>
</tr>
<tr>
<td>638–905</td>
<td>711</td>
<td>0.1192$\pm$0.0074</td>
<td>0.0909$\pm$0.0042</td>
<td>22</td>
<td>11.2/21</td>
</tr>
<tr>
<td>905–2116</td>
<td>1007</td>
<td>0.1176$\pm$0.0111</td>
<td>0.0866$\pm$0.0036</td>
<td>26</td>
<td>33.6/25</td>
</tr>
</tbody>
</table>

### Table 6 Uncertainty composition for $\alpha_s(M_Z)$ from the determination of $\alpha_s(Q)$ in bins of $p_T$ using the CT10-NLO PDF set

<table>
<thead>
<tr>
<th>$p_T$ range (GeV)</th>
<th>$Q$ (GeV)</th>
<th>$\alpha_s(M_Z)$</th>
<th>Exp.</th>
<th>PDF</th>
<th>NP</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>114–196</td>
<td>136</td>
<td>0.1172</td>
<td>±0.0031</td>
<td>±0.0018</td>
<td>±0.0007</td>
<td>+0.0045</td>
</tr>
<tr>
<td>196–300</td>
<td>226</td>
<td>0.1180</td>
<td>±0.0034</td>
<td>±0.0019</td>
<td>±0.0011</td>
<td>-0.0022</td>
</tr>
<tr>
<td>300–468</td>
<td>345</td>
<td>0.1194</td>
<td>±0.0032</td>
<td>±0.0023</td>
<td>±0.0010</td>
<td>+0.0049</td>
</tr>
<tr>
<td>468–638</td>
<td>521</td>
<td>0.1187</td>
<td>±0.0029</td>
<td>±0.0031</td>
<td>±0.0006</td>
<td>+0.0052</td>
</tr>
<tr>
<td>638–905</td>
<td>711</td>
<td>0.1192</td>
<td>±0.0034</td>
<td>±0.0032</td>
<td>±0.0005</td>
<td>+0.0057</td>
</tr>
<tr>
<td>905–2116</td>
<td>1007</td>
<td>0.1176</td>
<td>±0.0047</td>
<td>±0.0040</td>
<td>±0.0002</td>
<td>+0.0092</td>
</tr>
</tbody>
</table>

### Table 7 Uncertainty composition for $\alpha_s(Q)$ in bins of $p_T$ using the CT10-NLO PDF set

<table>
<thead>
<tr>
<th>$p_T$ range (GeV)</th>
<th>$Q$ (GeV)</th>
<th>$\alpha_s(Q)$</th>
<th>Exp.</th>
<th>PDF</th>
<th>NP</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>114–196</td>
<td>136</td>
<td>0.1106</td>
<td>±0.0028</td>
<td>±0.0016</td>
<td>±0.0006</td>
<td>+0.0040</td>
</tr>
<tr>
<td>196–300</td>
<td>226</td>
<td>0.1038</td>
<td>±0.0026</td>
<td>±0.0015</td>
<td>±0.0008</td>
<td>+0.0037</td>
</tr>
<tr>
<td>300–468</td>
<td>345</td>
<td>0.0993</td>
<td>±0.0022</td>
<td>±0.0016</td>
<td>±0.0007</td>
<td>+0.0033</td>
</tr>
<tr>
<td>468–638</td>
<td>521</td>
<td>0.0940</td>
<td>±0.0018</td>
<td>±0.0019</td>
<td>±0.0004</td>
<td>+0.0032</td>
</tr>
<tr>
<td>638–905</td>
<td>711</td>
<td>0.0909</td>
<td>±0.0019</td>
<td>±0.0018</td>
<td>±0.0003</td>
<td>+0.0032</td>
</tr>
<tr>
<td>905–2116</td>
<td>1007</td>
<td>0.0866</td>
<td>±0.0025</td>
<td>±0.0021</td>
<td>±0.0001</td>
<td>+0.0048</td>
</tr>
</tbody>
</table>

5 Study of PDF constraints with HERAFITTER

The PDFs of the proton are an essential ingredient for precision studies in hadron-induced reactions. They are derived from experimental data involving collider and fixed-target experiments. The DIS data from the HERA-I ep collider cover most of the kinematic phase space needed for a reliable PDF extraction. The pp inclusive jet cross section contains additional information that can constrain the PDFs, in particular the gluon, in the region of high fractions $x$ of the proton momentum.

The HERAFITTER project [61, 62] is an open-source framework designed among other things to fit PDFs to data. It has a modular structure, encompassing a variety of theoretical predictions for different processes and phenomenological approaches for determining the parameters of the PDFs. In this study, the recently updated HERAFITTER version 1.1.1 is employed to estimate the impact of the CMS inclusive jet data on the PDFs and their uncertainties. Theory is used at...
NLO for both processes, i.e. up to order $\alpha_s^3$ for DIS and up to order $\alpha_s^2$ for inclusive jet production in pp collisions.

5.1 Correlation between inclusive jet production and the PDFs

The potential impact of the CMS inclusive jet data can be illustrated by the correlation between the inclusive jet cross section $\sigma_{\text{jet}}(Q)$ and the PDF $x f(x, Q^2)$ for any parton flavour $f$. The NNPDF Collaboration [63] provides PDF sets in the form of an ensemble of replicas $i$, which sample variations in the PDF parameter space within allowed uncertainties. The correlation coefficient $\rho_f(x, Q)$ between the inclusive jet cross section and the PDF for flavour $f$ at a point $(x, Q)$ can be computed by evaluating means and standard deviations from an ensemble of $N$ replicas as

$$\rho_f(x, Q) = \frac{\langle \sigma_{\text{jet}}(Q)_i \cdot x f(x, Q^2)_i \rangle - \langle \sigma_{\text{jet}}(Q)_i \rangle \cdot \langle x f(x, Q^2)_i \rangle}{\Delta_{\sigma_{\text{jet}}(Q)} \Delta_{x f(x, Q^2)}}.$$  

(12)

Here, the angular brackets denote the averaging over the replica index $i$, and $\Delta$ represents the evaluation of the corresponding standard deviation for either the jet cross section, $\Delta_{\sigma_{\text{jet}}(Q)}$, or a PDF, $\Delta_{x f(x, Q^2)}$. Figure 11 presents the correlation coefficient between the inclusive jet cross section and the gluon, u valence quark, and d valence quark PDFs in the proton.

The correlation between the gluon PDF and the inclusive jet cross section is largest at central rapidity for most jet $p_T$. In contrast, the correlation between the valence quark distributions and the jet cross section is rather small except for very high $p_T$ such that some impact can be expected at high $x$ from including these jet data in PDF fits. In the forward region the correlation between the valence quark distributions and the jet cross sections is more pronounced at high $x$ and smaller jet $p_T$. Therefore, a significant reduction of the PDF uncertainties is expected by including the CMS inclusive jet cross section into fits of the proton structure.

5.2 The fitting framework

5.2.1 The HERAFITTER setup

The impact of the CMS inclusive jet data on proton PDFs is investigated by including the jet cross section measurement in a combined fit at NLO with the HERA-I inclusive DIS cross sections [19], which were the basis for the determination of the HERAPDF1.0 PDF set. The analysis is performed within the HERAFITTER framework using the Dokshitzer–Gribov–Lipatov–Altarelli–Parisi [64–66] evolution scheme at NLO as implemented in the QCDNUM pack-
age [67] and the generalised-mass variable-flavour number Thorne–Roberts scheme [68,69].

In contrast to the original HERAPDF fit, the results presented here require the DIS data to fulfill $Q^2 > Q^2_{\text{min}} = 7.5 \text{GeV}^2$ instead of $3.5 \text{GeV}^2$. The amount of DIS data left out by the increased $Q^2_{\text{min}}$ threshold is rather small and concerns a phase space where a perturbative description is less reliable. A similar, higher cutoff has been applied by the ATLAS Collaboration [70,71]. As a crosscheck all fits have been performed for a cutoff of $Q^2 > Q^2_{\text{min}} = 3.5 \text{GeV}^2$, and the results are consistent with the ones obtained using the more stringent cutoff. Differences beyond the expected reduction of uncertainties at low $x$ have not been observed.

The following PDFs are independent in the fit procedure: $xu_i(x)$, $xd_i(x)$, $xg(x)$, and $x\bar{U}(x)$, $x\bar{D}(x)$, where $x\bar{U}(x) = x\bar{u}(x)$, and $x\bar{D}(x) = x\bar{d}(x) + x\bar{s}(x)$. Similar to Ref. [72], a parameterization with 13 free parameters is used. At the starting scale $Q_0$ of the QCD evolution, chosen to be $Q_0^2 = 1.9 \text{GeV}^2$, the PDFs are parameterized as follows:

$$xg(x) = A_g x^{B_g} (1 - x)^{C_g} - A_g' x^{B_g'} (1 - x)^{C_g'},$$

$$xu_i(x) = A_{u_i} x^{B_{u_i}} (1 - x)^{C_{u_i}} (1 + E_{u_i} x^2),$$

$$xd_i(x) = A_{d_i} x^{B_{d_i}} (1 - x)^{C_{d_i}},$$

$$x\bar{U}(x) = A_{\bar{T}u} x^{B_{\bar{T}u}} (1 - x)^{C_{\bar{T}u}},$$

$$x\bar{D}(x) = A_{\bar{T}d} x^{B_{\bar{T}d}} (1 - x)^{C_{\bar{T}d}}.$$  

(13)

The normalisation parameters $A_g$, $A_{u_i}$, and $A_{d_i}$ are constrained by QCD sum rules. Additional constraints $B_{\bar{T}u} = B_{\bar{T}d}$ and $A_{\bar{T}u} = A_{\bar{T}d} (1 - f_s)$ are applied to ensure the same normalisation for the $\bar{u}$ and $\bar{d}$ densities for $x \to 0$. The strangeness fraction is set to $f_s = 0.31$, as obtained from neutrino-induced dimuon production [73]. The parameter $C'_g$ is fixed to 25 [20,69] and the strong coupling constant $\alpha_s(M_Z) = 0.1176$.

5.2.2 Definition of the goodness-of-fit estimator

The agreement between the $N$ data points $D_i$ and the theoretical predictions $T_i$ is quantified via a least-squares method, where

$$\chi^2 = \sum_{ij}^N (D_i - T_i - \sum_k r_k \beta_{ik})$$

$$\times C_{ij}^{-1} \left( D_j - T_j - \sum_k r_k \beta_{jk} \right) + \sum_k r_k^2.$$  

(14)

For fully correlated sources of uncertainty following a Gaussian distribution with a zero mean and a root-mean-square of unity as assumed here, this definition is equivalent to Eq. (9) [74]. As a bonus, the systematic shift of the nuisance param-
Fig. 11 The correlation coefficient between the inclusive jet cross section and the gluon (top row), the $u$ valence quark (middle row), and the $d$ valence quark PDFs (bottom row), as a function of the momentum fraction $x$ of the proton and the energy scale $Q$ of the hard process. The correlation is shown for the central rapidity region $|y| < 0.5$ (left) and for $2.0 < |y| < 2.5$ (right).
eter $r_k$ for each source in a fit is determined. Numerous large shifts in either direction indicate a problem as for example observed while fitting $\alpha_s(M_Z)$ with this technique and the old uncertainty correlation prescription.

In the following, the covariance matrix is defined as $C = \text{cov}_{\text{stat}} + \text{cov}_{\text{uncor}}$, while the JES, unfolding, and luminosity determination are treated as fully correlated systematic uncertainties $\beta_{ik}$ with nuisance parameters $r_k$. Including also the NP uncertainties, treated via the offset method in Sect. 4, in the form of one nuisance parameter in total $K$ such sources are defined. Of course, PDF uncertainties emerge as results of the fits performed here, in contrast to serving as inputs, as they do in the fits of $\alpha_s(M_Z)$ presented in Sect. 4.

All the fully correlated sources are assumed to be multiplicative to avoid the statistical bias that arises from uncertainty estimations taken from data [45–47]. As a consequence, the covariance matrix of the remaining sources has to be re-evaluated in each iteration step. To inhibit the compensation of large systematic shifts by increasing simultaneously the theoretical prediction and the statistical uncertainties, the systematic shifts of the theory are taken into account before the rescaling of the statistical uncertainty. Otherwise alternative minima in $\chi^2$ can appear that are associated with large theoretical predictions and correspondingly large shifts in the nuisance parameters. These alternative minima are clearly undesirable [62].

5.2.3 Treatment of CMS data uncertainties

The JES is the dominant source of experimental systematic uncertainty in jet cross sections. As described in Sect. 2.3, the $p_T$- and $\eta$-dependent JES uncertainties are split into 16 uncorrelated sources that are fully correlated in $p_T$ and $\eta$. Following the modified recommendation for the correlations versus rapidity of the single-particle response source as given in Sect. 2.3, it is necessary to split this source into five parts for the purpose of using the uncertainties published in Ref. [1] within the $\chi^2$ fits. The complete set of uncertainty sources is shown in Table 8.

By employing the technique of nuisance parameters, the impact of each systematic source of uncertainty on the fit result can be examined separately. For an adequate estimation of the size and the correlations of all uncertainties, the

<table>
<thead>
<tr>
<th>Systematic source</th>
<th>Shift in standard deviations</th>
</tr>
</thead>
<tbody>
<tr>
<td>JEC0 absolute jet energy scale</td>
<td>0.09</td>
</tr>
<tr>
<td>JEC1 MC extrapolation</td>
<td>0.00</td>
</tr>
<tr>
<td>JEC2a single-particle response barrel</td>
<td>1.31</td>
</tr>
<tr>
<td>JEC2b single-particle response endcap</td>
<td>-1.46</td>
</tr>
<tr>
<td>JEC2c single-particle decorrelation $</td>
<td>y</td>
</tr>
<tr>
<td>JEC2d single-particle decorrelation $0.5 \leq</td>
<td>y</td>
</tr>
<tr>
<td>JEC2e single-particle decorrelation $1.0 \leq</td>
<td>y</td>
</tr>
<tr>
<td>JEC3 jet flavor correction</td>
<td>0.04</td>
</tr>
<tr>
<td>JEC4 time-dependent detector effects</td>
<td>-0.15</td>
</tr>
<tr>
<td>JEC5 jet $p_T$ resolution in endcap 1</td>
<td>0.76</td>
</tr>
<tr>
<td>JEC6 jet $p_T$ resolution in endcap 2</td>
<td>-0.42</td>
</tr>
<tr>
<td>JEC7 jet $p_T$ resolution in HF</td>
<td>0.01</td>
</tr>
<tr>
<td>JEC8 correction for final-state radiation</td>
<td>0.03</td>
</tr>
<tr>
<td>JEC9 statistical uncertainty of $\eta$-dependent correction for endcap</td>
<td>-0.42</td>
</tr>
<tr>
<td>JEC10 statistical uncertainty of $\eta$-dependent correction for HF</td>
<td>0.00</td>
</tr>
<tr>
<td>JEC11 data-MC difference in $\eta$-dependent pileup correction</td>
<td>0.91</td>
</tr>
<tr>
<td>JEC12 residual out-of-time pileup correction for prescaled triggers</td>
<td>-0.17</td>
</tr>
<tr>
<td>JEC13 offset dependence in pileup correction</td>
<td>-0.03</td>
</tr>
<tr>
<td>JEC14 MC pileup bias correction</td>
<td>0.39</td>
</tr>
<tr>
<td>JEC15 jet rate dependent pileup correction</td>
<td>0.29</td>
</tr>
<tr>
<td>Unfolding</td>
<td>-0.26</td>
</tr>
<tr>
<td>Luminosity</td>
<td>-0.07</td>
</tr>
<tr>
<td>NP correction</td>
<td>0.60</td>
</tr>
</tbody>
</table>

Table 8 The 19 independent sources of systematic uncertainty considered in the CMS inclusive jet measurement. Out of these, 16 are related to the JES and are listed first. In order to implement the improved correlation treatment as described in Sect. 2.3, the single-particle response source JEC2, see also Appendix A, has been split up into five sources: JEC2a–JEC2e. The shift from the default value in each source of systematic uncertainty is determined by nuisance parameters in the fit and is presented in units of standard deviations.
Table 9: Partial \( \chi^2 \) values, \( \chi^2_p \), for each data set in the HERA-I DIS (middle section) or in the combined fit including CMS inclusive jet data (right section). Here, \( n_{\text{data}} \) is the number of data points available for the determination of the 13 parameters. The bottom two lines show the total \( \chi^2 \) and \( \chi^2/n_{\text{dof}} \). The difference between the sum of all \( \chi^2_p \) and the total \( \chi^2 \) for the combined fit is attributed to the nuisance parameters.

<table>
<thead>
<tr>
<th>Data set</th>
<th>( n_{\text{data}} )</th>
<th>( \chi^2_p )</th>
<th>( \chi^2_p/n_{\text{data}} )</th>
<th>( \chi^2 )</th>
<th>( \chi^2/n_{\text{dof}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>NC HERA-I H1-ZEUS combined e(^-)p</td>
<td>145</td>
<td>109</td>
<td>0.75</td>
<td>109</td>
<td>0.75</td>
</tr>
<tr>
<td>NC HERA-I H1-ZEUS combined e(^+)p</td>
<td>337</td>
<td>309</td>
<td>0.91</td>
<td>311</td>
<td>0.92</td>
</tr>
<tr>
<td>CC HERA-I H1-ZEUS combined e(^-)p</td>
<td>34</td>
<td>20</td>
<td>0.59</td>
<td>22</td>
<td>0.65</td>
</tr>
<tr>
<td>CC HERA-I H1-ZEUS combined e(^+)p</td>
<td>34</td>
<td>29</td>
<td>0.85</td>
<td>35</td>
<td>1.03</td>
</tr>
<tr>
<td>CMS inclusive jets</td>
<td>133</td>
<td>–</td>
<td>–</td>
<td>102</td>
<td>0.77</td>
</tr>
<tr>
<td>Data set(s)</td>
<td>( n_{\text{dof}} )</td>
<td>( \chi^2 )</td>
<td>( \chi^2/n_{\text{dof}} )</td>
<td>( \chi^2 )</td>
<td>( \chi^2/n_{\text{dof}} )</td>
</tr>
<tr>
<td>HERA-I data</td>
<td>537</td>
<td>468</td>
<td>0.87</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>HERA-I and CMS data</td>
<td>670</td>
<td>–</td>
<td>–</td>
<td>591</td>
<td>0.88</td>
</tr>
</tbody>
</table>

Fig. 12: The gluon (top) and sea quark (bottom) PDFs as a function of \( x \) as derived from HERA-I inclusive DIS data alone (left) and in combination with CMS inclusive jet data (right). The PDFs are shown at the starting scale \( Q^2 = 1.9 \text{ GeV}^2 \). The experimental (inner band), model (middle band), and parameterization uncertainties (outer band) are successively added quadratically to give the total uncertainty.

In contrast, with the original assumption of full correlation within the 16 JES systematic sources across all \(|y|\) bins, shifts beyond two standard deviations were apparent and led to a re-examination of this issue and the improved correlation.
5.3 Determination of PDF uncertainties according to the HERAPDF prescription

The uncertainty in the PDFs is subdivided into experimental, model, and parameterization uncertainties that are studied separately. In the default setup of the HERAFIT framework, experimental uncertainties are evaluated following a Hessian method [74], and result from the propagated statistical and systematic uncertainties of the input data.

For the model uncertainties, the offset method [75] is applied considering the following variations of model assumptions:

1. The strangeness fraction $f_s$, by default equal to 0.31, is varied between 0.23 and 0.38.
2. The $b$-quark mass is varied by $\pm 0.25$ GeV around the central value of 4.75 GeV.
3. The $c$-quark mass, with the central value of 1.4 GeV, is varied to 1.35 and 1.65 GeV. For the downwards variation the charm production threshold is avoided by changing the starting scale to $Q_0^2 = 1.8$ GeV$^2$ in this case.
4. The minimum $Q^2$ value for data used in the fit, $Q_{\text{min}}^2$, is varied from 7.5 to 5.0 and 10 GeV$^2$.

The PDF parameterization uncertainty is estimated as described in Ref. [19]. By employing the more general form of parameterizations

$$x g(x) = A_g x^B g (1-x)^C g (1 + D_g x + E_g x^2)$$

$$ - A'_g x^{B' g} (1-x)^{C'_ g}$$

$$x f(x) = A_f x^B f (1-x)^C f (1 + D_f x + E_f x^2)$$

for gluons and the nongluon flavours, respectively, it is tested whether the successive inclusion of additional fit parameters leads to a variation in the shape of the fitted results. Furthermore, the starting scale $Q_0$ is changed to $Q_0^2 = 1.5$ and 2.5 GeV$^2$. The maximal deviations of the resulting PDFs

---

Fig. 13 The $u$ valence quark (top) and $d$ valence quark (bottom) PDFs as a function of $x$ as derived from HERA-I inclusive DIS data alone (left) and in combination with CMS inclusive jet data (right). The PDFs are shown at the starting scale $Q^2 = 1.9$ GeV$^2$. The experimental (inner band), model (middle band), and parameterization uncertainties (outer band) are successively added quadratically to give the total uncertainty.

---

by employing the more general form of parameterizations

$$x g(x) = A_g x^B g (1-x)^C g (1 + D_g x + E_g x^2)$$

$$ - A'_g x^{B' g} (1-x)^{C'_ g}$$

$$x f(x) = A_f x^B f (1-x)^C f (1 + D_f x + E_f x^2)$$

for gluons and the nongluon flavours, respectively, it is tested whether the successive inclusion of additional fit parameters leads to a variation in the shape of the fitted results. Furthermore, the starting scale $Q_0$ is changed to $Q_0^2 = 1.5$ and 2.5 GeV$^2$. The maximal deviations of the resulting PDFs

---

by employing the more general form of parameterizations

$$x g(x) = A_g x^B g (1-x)^C g (1 + D_g x + E_g x^2)$$

$$ - A'_g x^{B' g} (1-x)^{C'_ g}$$

$$x f(x) = A_f x^B f (1-x)^C f (1 + D_f x + E_f x^2)$$

for gluons and the nongluon flavours, respectively, it is tested whether the successive inclusion of additional fit parameters leads to a variation in the shape of the fitted results. Furthermore, the starting scale $Q_0$ is changed to $Q_0^2 = 1.5$ and 2.5 GeV$^2$. The maximal deviations of the resulting PDFs

---

by employing the more general form of parameterizations

$$x g(x) = A_g x^B g (1-x)^C g (1 + D_g x + E_g x^2)$$

$$ - A'_g x^{B' g} (1-x)^{C'_ g}$$

$$x f(x) = A_f x^B f (1-x)^C f (1 + D_f x + E_f x^2)$$

for gluons and the nongluon flavours, respectively, it is tested whether the successive inclusion of additional fit parameters leads to a variation in the shape of the fitted results. Furthermore, the starting scale $Q_0$ is changed to $Q_0^2 = 1.5$ and 2.5 GeV$^2$. The maximal deviations of the resulting PDFs

---

by employing the more general form of parameterizations

$$x g(x) = A_g x^B g (1-x)^C g (1 + D_g x + E_g x^2)$$

$$ - A'_g x^{B' g} (1-x)^{C'_ g}$$

$$x f(x) = A_f x^B f (1-x)^C f (1 + D_f x + E_f x^2)$$

for gluons and the nongluon flavours, respectively, it is tested whether the successive inclusion of additional fit parameters leads to a variation in the shape of the fitted results. Furthermore, the starting scale $Q_0$ is changed to $Q_0^2 = 1.5$ and 2.5 GeV$^2$. The maximal deviations of the resulting PDFs

---

by employing the more general form of parameterizations

$$x g(x) = A_g x^B g (1-x)^C g (1 + D_g x + E_g x^2)$$

$$ - A'_g x^{B' g} (1-x)^{C'_ g}$$

$$x f(x) = A_f x^B f (1-x)^C f (1 + D_f x + E_f x^2)$$

for gluons and the nongluon flavours, respectively, it is tested whether the successive inclusion of additional fit parameters leads to a variation in the shape of the fitted results. Furthermore, the starting scale $Q_0$ is changed to $Q_0^2 = 1.5$ and 2.5 GeV$^2$. The maximal deviations of the resulting PDFs

---

by employing the more general form of parameterizations

$$x g(x) = A_g x^B g (1-x)^C g (1 + D_g x + E_g x^2)$$

$$ - A'_g x^{B' g} (1-x)^{C'_ g}$$

$$x f(x) = A_f x^B f (1-x)^C f (1 + D_f x + E_f x^2)$$

for gluons and the nongluon flavours, respectively, it is tested whether the successive inclusion of additional fit parameters leads to a variation in the shape of the fitted results. Furthermore, the starting scale $Q_0$ is changed to $Q_0^2 = 1.5$ and 2.5 GeV$^2$. The maximal deviations of the resulting PDFs

---

by employing the more general form of parameterizations

$$x g(x) = A_g x^B g (1-x)^C g (1 + D_g x + E_g x^2)$$

$$ - A'_g x^{B' g} (1-x)^{C'_ g}$$

$$x f(x) = A_f x^B f (1-x)^C f (1 + D_f x + E_f x^2)$$

for gluons and the nongluon flavours, respectively, it is tested whether the successive inclusion of additional fit parameters leads to a variation in the shape of the fitted results. Furthermore, the starting scale $Q_0$ is changed to $Q_0^2 = 1.5$ and 2.5 GeV$^2$. The maximal deviations of the resulting PDFs

---

by employing the more general form of parameterizations

$$x g(x) = A_g x^B g (1-x)^C g (1 + D_g x + E_g x^2)$$

$$ - A'_g x^{B' g} (1-x)^{C'_ g}$$

$$x f(x) = A_f x^B f (1-x)^C f (1 + D_f x + E_f x^2)$$

for gluons and the nongluon flavours, respectively, it is tested whether the successive inclusion of additional fit parameters leads to a variation in the shape of the fitted results. Furthermore, the starting scale $Q_0$ is changed to $Q_0^2 = 1.5$ and 2.5 GeV$^2$. The maximal deviations of the resulting PDFs

---

by employing the more general form of parameterizations

$$x g(x) = A_g x^B g (1-x)^C g (1 + D_g x + E_g x^2)$$

$$ - A'_g x^{B' g} (1-x)^{C'_ g}$$

$$x f(x) = A_f x^B f (1-x)^C f (1 + D_f x + E_f x^2)$$

for gluons and the nongluon flavours, respectively, it is tested whether the successive inclusion of additional fit parameters leads to a variation in the shape of the fitted results. Furthermore, the starting scale $Q_0$ is changed to $Q_0^2 = 1.5$ and 2.5 GeV$^2$. The maximal deviations of the resulting PDFs
Fig. 14 The gluon (top left), sea quark (top right), u valence quark (bottom left), and d valence quark (bottom right) PDFs as a function of $x$ as derived from HERA-I inclusive DIS data alone (dashed line) and in combination with CMS inclusive jet data (full line).

The PDFs are determined employing the HERAPDF method with a $Q^2_{\text{min}} = 7.5 \text{ GeV}^2$ selection criterion. The PDFs are shown at the starting scale $Q^2 = 1.9 \text{ GeV}^2$. Only the total uncertainty in the PDFs is shown (hatched and solid bands).

From those obtained in the central fit define the parameterization uncertainty. The experimental, model, and parameterization uncertainties are added in quadrature to give the final PDF uncertainty according to the HERAPDF prescription [19].

Using this fitting setup, the partial $\chi^2$ values per number of data points, $n_{\text{data}}$, are reported in Table 9 for each of the neutral current (NC) and charged current (CC) data sets in the HERA-I DIS fit and for the combined fit including the CMS inclusive jet data. The achieved fit qualities demonstrate the compatibility of all data within the presented PDF fitting framework. The resulting PDFs with breakdown of the uncertainties for the gluon, the sea, u valence, and d valence quarks with and without CMS inclusive jet data are arranged next to each other in Figs. 12 and 13. Figure 14 provides direct comparisons of the two fit results with total uncertainties. The parameterization and model uncertainties of the gluon distribution are significantly reduced for almost the whole $x$ range from $10^{-4}$ up to 0.5. When DIS data below $Q^2_{\text{min}} = 7.5 \text{ GeV}^2$ are included in the fit, the effect is much reduced for the low $x$ region $x < 0.01$, but remains important for medium to high $x$. Also, for the u valence, d valence, and sea quark distributions some reduction in their uncertainty is visible at high $x$ ($x \gtrsim 0.1$).

At the same time, some structure can be seen, particularly in the parameterization uncertainties that might point to a still insufficient flexibility in the parameterizations. Therefore, a comparison is presented in the next Sect. 5.4, using the MC method with the regularisation based on data, which is also implemented within the HERAFITTER framework.

5.4 Determination of PDF uncertainties using the MC method with regularisation

To study more flexible PDF parameterizations, a MC method based on varying the input data within their correlated uncertainties is employed in combination with a data-based regularisation technique. This method was first used by the NNPDF Collaboration and uses a more flexible parameterization to describe the $x$ dependence of the PDFs [63]. To
avoid the fitting of statistical fluctuations present in the input data (over-fitting) a data-based stopping criterion is introduced. The data set is split randomly into a “fit” and a “control” sample. The $\chi^2$ minimisation is performed with the “fit” sample while simultaneously the $\chi^2$ of the “control” sample is calculated using the current PDF parameters. It is observed that the $\chi^2$ of the “control” sample at first decreases and then starts to increase again because of over-fitting. At this point, the fit is stopped. This regularisation technique is used in combination with a MC method to estimate the central value and the uncertainties of the fitted PDFs. Before a fit, several hundred replica sets are created by allowing the central values of the measured cross section to fluctuate within their statistical and systematic uncertainties while taking into account all correlations. For each replica, a fit to NLO QCD is performed, which yields an optimum value and uncertainty for each parameter. The collection of all replica fits can then provide an ensemble average and root-mean-square. Moreover, the variations to derive the model dependence of the HERAPDF prescription do not lead to any further increase of the uncertainty.

Similarly to Fig. 14 for the HERAPDF method, a direct comparison of the two fit results with total uncertainties is shown in Fig. 15 for the MC method. The total uncertainty derived with the MC method is almost always larger than with the HERAPDF technique, and in the case of the gluon at low $x$, it is much larger. In both cases a significant reduction of the uncertainty in the gluon PDF is observed, notably in the $x$ range from $10^{-2}$ up to 0.5. Both methods also lead to a decrease in the gluon PDF between $10^{-2}$ and $10^{-1}$ and an increase for larger $x$. Although this change is more pronounced when applying the MC method, within the respective uncertainties both results are compatible. For the sea quark only small differences in shape are observed, but, in contrast to the HERAPDF method that exhibits reduced uncertainties for $x > 0.2$, this is not visible when using the MC method. Both methods agree on a very modest reduction in uncertainty at high $x > 0.05$ in the u valence quark.
PDF and a somewhat larger improvement for the d valence quark PDF, which is expected from the correlations, studied in Fig. 11, where the quark distributions are constrained via the \( q\bar{q} \) contribution to jet production at high \( |y| \) and \( p_T \). Changes in shape of the d valence quark PDF go into opposite directions for the two methods, but are compatible within uncertainties.

All preceding figures presented the PDFs at the starting scale of the evolution of \( Q^2 = 1.9 \text{ GeV}^2 \). For illustration, Fig. 16 displays the PDFs derived with the regularised MC method after evolution to a scale of \( Q^2 = 10^4 \text{ GeV}^2 \). Finally, Fig. 17 shows an overview of the gluon, sea, u valence, and d valence distributions at the starting scale of \( Q^2 = 1.9 \text{ GeV}^2 \) for both techniques, the HERAPDF and the regularised MC method.

5.5 Combined fit of PDFs and the strong coupling constant

Inclusive DIS data alone are not sufficient to disentangle effects on cross section predictions from changes in the gluon distribution or \( \alpha_S(M_Z) \) simultaneously. Therefore \( \alpha_S(M_Z) \) was always fixed to 0.1176 in the original HERAPDF1.0 derivation. When the CMS inclusive jet data are added, this constraint can be dropped and \( \alpha_S(M_Z) \) and its uncertainty (without \( Q \) scale variations) is determined to \( \alpha_S(M_Z) = 0.1192^{+0.0023}_{-0.0015} \) (all except scale). Repeating the fit with the regularised MC method gives \( \alpha_S(M_Z) = 0.1188 \pm 0.0041 \) (all except scale).

Since a direct correspondence among the different components of the uncertainty can not easily be established, only the quadratic sum of experimental, PDF, and NP uncertainties are presented, which is equivalent to the total uncertainty without scale uncertainty. For example, the HERA-I DIS data contribute to the experimental uncertainty in the combined fits, but contribute only to the PDF uncertainty in separate \( \alpha_S(M_Z) \) fits. The HERAPDF prescription for PDF fits tends to small uncertainties, while the uncertainties of the MC method with data-derived regularisation are twice as large. For comparison, the corresponding uncertainty in \( \alpha_S(M_Z) \) using more precisely determined PDFs...
are derived for a series of fixed values of $\alpha_s(M_Z)$ as done for the global PDF sets. Using this series of PDFs with varying values of $\alpha_s(M_Z)$, the combination of PDF and $\alpha_s(M_Z)$ that best fits the HERA-I DIS and CMS inclusive jet data is found. The $\alpha_s(M_Z)$ values determined both ways are consistent with each other. The fits are now repeated for the same scale factor variations, and the maximal observed upward and downward changes of $\alpha_s(M_Z)$ with respect to the default scale factors are taken as scale uncertainty: $\Delta \alpha_s(M_Z) = +0.0024$ (scale).

In contrast to the scale uncertainty of the first procedure, there is less freedom for compensating effects between different gluon distributions and $\alpha_s(M_Z)$ values in the second procedure, and the latter procedure leads to a larger scale uncertainty as expected. In overall size the uncertainty is similar to the final results on $\alpha_s(M_Z)$ reported in the last section: $\Delta \alpha_s(M_Z) = +0.0053$ (scale).

6 Summary

An extensive QCD study has been performed based on the CMS inclusive jet data in Ref. [1]. Fits dedicated to determine $\alpha_s(M_Z)$ have been performed involving QCD predictions at NLO complemented with electroweak and NP corrections. Employing global PDFs, where the gluon is constrained through data from various experiments, the strong coupling constant has been determined to be

$$\alpha_s(M_Z) = 0.1185 \pm 0.0019 \text{ (exp)} \pm 0.0028 \text{ (PDF)}$$

$$\pm 0.0004 \text{ (NP)} = +0.0053$$

which is consistent with previous results.

It was found that the published correlations of the experimental uncertainties adequately reflect the detector characteristics and reliable fits of standard model parameters could be performed within each rapidity region. However, when combining several rapidity regions, it was discovered that the assumption of full correlation in rapidity $y$ had to be revised for one source of uncertainty in the JES, which suggested a modified correlation treatment that is described and applied in this work.

To check the running of the strong coupling, all fits have also been carried out separately for six bins in inclusive jet $p_T$, where the scale $Q$ of $\alpha_s(Q)$ is identified with $p_T$. The observed behaviour of $\alpha_s(Q)$ is consistent with the energy scale dependence predicted by the renormalization group equation of QCD, and extends the H1, ZEUS, and D0 results to the TeV region.

The impact of the inclusive jet measurement on the PDFs of the proton is investigated in detail using the HERAFITTER tool. When the CMS inclusive jet data are used together with the HERA-I DIS measurements, the uncertainty in the
gluon distribution is significantly reduced for fractional parton momenta $x \gtrsim 0.01$. Also, a modest improvement in uncertainty in the $u$ and $d$ valence quark distributions is observed.

The inclusion of the CMS inclusive jet data also allows a combined fit of $\alpha_S(M_Z)$ and of the PDFs, which is not possible with the HERA-I inclusive DIS data alone. The result is consistent with the reported values of $\alpha_S(M_Z)$ obtained from fits employing global PDFs.

Acknowledgments We congratulate our colleagues in the CERN accelerator departments for the excellent performance of the LHC and thank the technical and administrative staffs at CERN and at other CMS institutes for their contributions to the success of the CMS effort. In addition, we gratefully acknowledge the computing centres and personnel of the Worldwide LHC Computing Grid for delivering so effectively the computing infrastructure essential to our analyses. Finally, we acknowledge the enduring support for the construction and operation of the LHC and the CMS detector provided by the following funding agencies: the Austrian Federal Ministry of Science, Research and Economy and the Austrian Science Fund; the Belgian Fonds de la Recherche Scientifique, and Fonds voor Wetenschappelijk Onderzoek; the Brazilian Funding Agencies (CNPq, CAPES, FAPERJ, and FAPESP); the Bulgarian Ministry of Education and Science; CERN; the Chinese Academy of Sciences, Ministry of Science and Technology, and National Natural Science Foundation of China; the Colombian Funding Agency (COLCIENCIAS); the Croatian Ministry of Science, Education and Sport, and the Croatian Science Foundation; the Research Promotion Foundation, Cyprus; the Ministry of Education and Research, Estonian Research Council via IUT23-4 and IUT23-6 and European Regional Development Fund, Estonia; the Academy of Finland, Finnish Ministry of Education and Culture, and Helsinki Institute of Physics; the Institut National de Physique Nucléaire et de Physique des Particules/CNRS, and Commissariat à l’Énergie Atomique et aux Énergies Alternatives/CEA, France; the Bundesministerium für Bildung und Forschung, Deutsche Forschungsgemeinschaft, and Helmholtz-Gemeinschaft Deutscher Forschungszentren, Germany; the General Secretariat for Research and Technology, Greece; the National Scientific Research Foundation, and National Innovation Office, Hungary; the Department of Atomic Energy and the Department of Science and Technology, India; the Institute for Studies in Theoretical Physics and Mathematics, Iran; the Science Foundation, Ireland; the Istituto Nazionale di Fisica Nucleare, Italy; the Korean Ministry of Education, Science and Technology and the World Class University program of KU; the Czech Ministry of Education, Youth and Sports (MEYS) of the Czech Republic; the Council of Science and Industrial Research, India; the HOMING PLUS programme of Foundation for Polish Science, cofinanced from European Union, Regional Development Fund; the Compagnia di San Paolo (Torino); the Consorzio per la Fisica (Trieste); MIUR Project 2010ST4XTM (Italy); the Thalis and Aristeia programmes cofinanced by EU-ESF and the Greek NSRF; and the National Priorities Research Program by Qatar National Research Fund.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Funded by SCOAP3.

Appendix A: Sources of uncertainty in the calibration of jet energies in CMS

In the following, the full list of uncertainty sources of the jet energy calibration procedure that were originally considered by CMS and that were used in Ref. [1] is presented including a short description. It is recommended to apply the procedure with updated correlations for the JEC2 source, as described in Sect. 2.3. A general description of the jet energy calibration procedure of CMS is given in Ref. [10].

When simulations were employed, the following event generators have been used: PYTHIA version 6.4.22 [26] tune Z2 and HERWIG++ version 2.4.2 [27] with the default tune of version 2.3.

JEC0 Absolute uncertainty. Using data with photon + jet and $Z +$ jet events an absolute calibration of jet energies is performed in the jet $p_T$ range of 30–600 GeV. Uncertainties in the determination of electromagnetic energies in the ECAL, of the muon momenta from $Z \rightarrow \mu \mu$ decays, and of the corrections for initial- and final-state (ISR and FSR) radiation are propagated together with the statistical uncertainty to give the absolute JES uncertainty.

JEC1 High- and low-$p_T$ extrapolation uncertainty.
Where an absolute calibration with data is not possible, events are produced with the event generators \textsc{pythia6} and \textsc{herwig++} and are subsequently processed through the CMS detector simulation based on \textsc{geant4} [79]. Differences in particular in modelling the fragmentation process and the underlying event lead to an extrapolation uncertainty relative to the directly calibrated jet $p_T$ range of 30–600 GeV.

JEC2 High-$p_T$ extrapolation uncertainty. This source accounts for a $\pm 3\%$ variation in the single-particle response that is propagated to jets using a parameterized fast simulation of the CMS detector [80].

JEC3 Jet flavour related uncertainty. Differences in detector response to light, charm, and bottom quark as well as gluon jets relative to the mixture predicted by QCD for the measured processes are evaluated on the basis of simulations with \textsc{pythia6} and \textsc{herwig++}.

JEC4 Uncertainty caused by time dependent detector effects. This source considers residual time-dependent variations in the detector conditions such as the endcap ECAL crystal transparency.

JEC5–JEC10 $\eta$-dependent uncertainties coming from the dijet balance method:

JEC5–JEC7 Caused by the jet energy resolution. These three sources are assumed to be fully correlated for the endcap with upstream tracking detectors (JEC5), the endcap without upstream tracking detectors (JEC6), and the HF calorimeter (JEC7).

JEC8 $\eta$-dependent uncertainty caused by corrections for final-state radiation. The uncertainty is correlated from one region to the other and increases towards HF.

JEC9–JEC10 Statistical uncertainty in the determination of $\eta$-dependent corrections. These are two separate sources for the endcap without upstream tracking detectors (JEC9), and the HF calorimeter (JEC10).

JEC11–JEC15 Uncertainties for the pileup corrections:

JEC11 Parameterizes differences between data and MC events versus $\eta$ in zero-bias data.

JEC12 Estimates residual out-of-time pileup for prescaled triggers, if MC events are reweighted to unprescaled data.

JEC13 Covers an offset dependence on jet $p_T$ (due to, e.g. zero-suppression effects), when the correction is calibrated for jets in the $p_T$ range of 20–30 GeV.

JEC14 Accounts for differences in measured offset from zero-bias MC events and from generator-level information in a QCD sample.

JEC15 Covers observed jet rate variations versus the average number of reconstructed primary vertices in the 2011 single-jet triggers after applying L1 corrections.

Appendix B: Comparison to theoretical predictions by \textsc{powheg} + \textsc{pythia6}

Figure 18 presents ratios of data over theory predictions at NLO using the CT10-NLO PDF set multiplied by electroweak and NP corrections including PDF uncertainties.
Fig. 18 Ratio of data to pQCD at NLO with the CT10-NLO PDF set multiplied by electroweak and NP corrections for the five bins in rapidity together with bands representing the CT10 PDF uncertainty (hatched), and the quadratically added scale and NP uncertainty (dashed lines). In addition, the ratio of the prediction by powheg + pythia6 tune Z2* at particle level is shown with boxes indicating the statistical uncertainty. The error bars and the grey boxes correspond to the statistical and systematic uncertainty in the data.
INFIN Sezione di Napoli, Università di Napoli ‘Federico II’, Università della Basilicata (Potenza),
Università G. Marconi (Roma), Naples, Italy

INFIN Sezione di Padova, Università di Padova, Università di Trento (Trento), Padua, Italy

INFIN Sezione di Pavia, Università di Pavia, Pavia, Italy
M. Gabusi, S. P. Ratti, V. Re, C. Riccardi, P. Salvini, P. Vitulo

INFIN Sezione di Perugia, Università di Perugia, Perugia, Italy
M. Biasini, G. M. Bilei, D. Ciangottini, L. Fanò, P. Lariccia, G. Mantovani, M. Menichelli, A. Saha, A. Santocchia, A. Spiezia

INFIN Sezione di Pisa, Università di Pisa, Scuola Normale Superiore di Pisa, Pisa, Italy

INFIN Sezione di Roma, Università di Roma, Rome, Italy

INFIN Sezione di Torino, Università di Torino, Università del Piemonte Orientale (Novara), Turin, Italy

INFIN Sezione di Trieste, Università di Trieste, Trieste, Italy
S. Belforte, V. Candelsi, M. Casarsa, F. Cossutti, G. Della Ricca, B. Gobbo, C. La Licata, M. Marone, A. Schizzi, T. Umer, A. Zanetti

Kangwon National University, Chunchon, Korea
S. Chang, T. A. Kropivnitskaya, S. K. Nam

Kyungpook National University, Taegu, Korea
D. H. Kim, G. N. Kim, M. S. Kim, S. D. Kong, S. Lee, Y. D. Oh, H. Park, A. Sakharov, D. C. Son

Chonbuk National University, Chonju, Korea
T. J. Kim

Chonnam National University, Institute for Universe and Elementary Particles, Kwangju, Korea
J. Y. Kim, S. Song

Korea University, Seoul, Korea
S. Choi, D. Gyun, B. Hong, M. Jo, H. Kim, Y. Kim, B. Lee, K. S. Lee, S. K. Park, Y. Roh

University of Seoul, Seoul, Korea
M. Choi, J. H. Kim, I. C. Park, G. Ryu, M. S. Ryu

Sungkyunkwan University, Suwon, Korea
Vilnius University, Vilnius, Lithuania
A. Juodagalvis

National Centre for Particle Physics, Universiti Malaya, Kuala Lumpur, Malaysia
J. R. Komaragiri, M. A. B. Md Ali

Centro de Investigacion y de Estudios Avanzados del IPN, Mexico City, Mexico

Universidad Iberoamericana, Mexico City, Mexico
S. Carrillo Moreno, F. Vazquez Valencia

Benemerita Universidad Autonoma de Puebla, Puebla, Mexico
I. Pedraza, H. A. Salazar Ibarguen

Universidad Autónoma de San Luis Potosí, San Luis Potosí, Mexico
A. Morelos Pineda

University of Auckland, Auckland, New Zealand
D. Krofcheck

University of Canterbury, Christchurch, New Zealand
P. H. Butler, S. Reucroft

National Centre for Physics, Quaid-I-Azam University, Islamabad, Pakistan
A. Ahmad, M. Ahmad, Q. Hassan, H. R. Hoorani, W. A. Khan, T. Khurshid, M. Shoaib

National Centre for Nuclear Research, Swierk, Poland

Institute of Experimental Physics, Faculty of Physics, University of Warsaw, Warsaw, Poland

Laboratório de Instrumentação e Física Experimental de Partículas, Lisbon, Portugal

Joint Institute for Nuclear Research, Dubna, Russia

Petersburg Nuclear Physics Institute, Gatchina, St. Petersburg, Russia

Institute for Nuclear Research, Moscow, Russia

Institute for Theoretical and Experimental Physics, Moscow, Russia
V. Epshteyn, V. Gavrilov, N. Lychkovskaya, V. Popov, I. Pozdnyakov, G. Safronov, S. Semenov, A. Spiridonov, V. Stolin, E. Vlasov, A. Zhokin

P. N. Lebedev Physical Institute, Moscow, Russia

Skobeltsyn Institute of Nuclear Physics, Lomonosov Moscow State University, Moscow, Russia
A. Belyaev, E. Boos, M. Dubinin, L. Dudko, A. Ershov, A. Gribushin, V. Klyukhin, O. Kodolova, I. Lokhtin, S. Obraztsov, S. Petrushanko, V. Savrin, A. Snigirev
State Research Center of Russian Federation, Institute for High Energy Physics, Protvino, Russia

Faculty of Physics and Vinca Institute of Nuclear Sciences, University of Belgrade, Belgrade, Serbia
P. Azdic, M. Ekmekzic, J. Milosevic, V. Rekovic

Centro de Investigaciones Energéticas Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain

Universidad Autónoma de Madrid, Madrid, Spain
C. Albajar, J. F. de Trocóniz, M. Missiroli, D. Moran

Universidad de Oviedo, Oviedo, Spain
H. Brun, J. Cuevas, J. Fernandez Menendez, S. Folgueras, I. Gonzalez Caballero

Instituto de Física de Cantabria (IFCA), CSIC-Universidad de Cantabria, Santander, Spain

CERN, European Organization for Nuclear Research, Geneva, Switzerland

Paul Scherrer Institut, Villigen, Switzerland

Institute for Particle Physics, ETH Zurich, Zurich, Switzerland

Universität Zürich, Zurich, Switzerland

National Central University, Chung-Li, Taiwan

National Taiwan University (NTU), Taipei, Taiwan

Springer
3: Also at Institut Pluridisciplinaire Hubert Curien, Université de Strasbourg, Université de Haute Alsace Mulhouse, CNRS/IN2P3, Strasbourg, France
4: Also at National Institute of Chemical Physics and Biophysics, Tallinn, Estonia
5: Also at Skobeltsyn Institute of Nuclear Physics, Lomonosov Moscow State University, Moscow, Russia
6: Also at Universidade Estadual de Campinas, Campinas, Brazil
7: Also at Laboratoire Leprince-Ringuet, Ecole Polytechnique, IN2P3-CNRS, Palaiseau, France
8: Also at Joint Institute for Nuclear Research, Dubna, Russia
9: Also at Suez University, Suez, Egypt
10: Also at Cairo University, Cairo, Egypt
11: Also at Fayoum University, El-Fayoum, Egypt
12: Also at British University in Egypt, Cairo, Egypt
13: Now at Sultan Qaboos University, Muscat, Oman
14: Also at Université de Haute Alsace, Mulhouse, France
15: Also at Brandenburg University, of Technology, Cottbus, Germany
16: Also at Institute of Nuclear Research ATOMKI, Debrecen, Hungary
17: Also at Eötvös Loránd University, Budapest, Hungary
18: Also at University of Debrecen, Debrecen, Hungary
19: Also at University of Visva-Bharati, Santiniketan, India
20: Now at King Abdulaziz University, Jeddah, Saudi Arabia
21: Also at University of Ruhuna, Matara, Sri Lanka
22: Also at Isfahan University of Technology, Isfahan, Iran
23: Also at University of Tehran, Department of Engineering Science, Tehran, Iran
24: Also at Plasma Physics Research Center, Science and Research Branch, Islamic Azad University, Tehran, Iran
25: Also at Laboratori Nazionali di Legnaro dell’INFN, Legnaro, Italy
26: Also at Università degli Studi di Siena, Siena, Italy
27: Also at Centre National de la Recherche Scientifique (CNRS)-IN2P3, Paris, France
28: Also at Purdue University, West Lafayette, USA
29: Also at Universidad Michoacana de San Nicolas de Hidalgo, Morelia, Mexico
30: Also at Institute for Nuclear Research, Moscow, Russia
31: Also at St. Petersburg State Polytechnical University, St. Petersburg, Russia
32: Also at California Institute of Technology, Pasadena, USA
33: Also at Faculty of Physics, University of Belgrade, Belgrade, Serbia
34: Also at Facoltà Ingegneria, Università di Roma, Rome, Italy
35: Also at Scuola Normale e Sezione dell’INFN, Pisa, Italy
36: Also at University of Athens, Athens, Greece
37: Also at Paul Scherrer Institut, Villigen, Switzerland
38: Also at Institute for Theoretical and Experimental Physics, Moscow, Russia
39: Also at Albert Einstein Center for Fundamental Physics, Bern, Switzerland
40: Also at Gaziosmanpasa University, Tokat, Turkey
41: Also at Adiyaman University, Adiyaman, Turkey
42: Also at Cag University, Mersin, Turkey
43: Also at Anadolu University, Eskisehir, Turkey
44: Also at Izmir Institute of Technology, Izmir, Turkey
45: Also at Necmettin Erbakan University, Konya, Turkey
46: Also at Mimar Sinan University, Istanbul, Istanbul, Turkey
47: Also at Ozyegin University, Istanbul, Turkey
48: Also at Marmara University, Istanbul, Turkey
49: Also at Kafkas University, Kars, Turkey
50: Also at Yildiz Technical University, Istanbul, Turkey
51: Also at Rutherford Appleton Laboratory, Didcot, UK
52: Also at School of Physics and Astronomy, University of Southampton, Southampton, UK
53: Also at University of Belgrade, Faculty of Physics and Vinca Institute of Nuclear Sciences, Belgrade, Serbia
54: Also at Argonne National Laboratory, Argonne, USA
55: Also at Erzincan University, Erzincan, Turkey
56: Also at Texas A&M University at Qatar, Doha, Qatar
57: Also at Kyungpook National University, Daegu, Korea