The PDF4LHC report on PDFs and LHC data: results from Run I and preparation for Run II

2015 J. Phys. G: Nucl. Part. Phys. 42 103103

(http://iopscience.iop.org/0954-3899/42/10/103103)

IOP Publishing Journal of Physics G: Nuclear and Particle Physics

J. Phys. G: Nucl. Part. Phys. **42** (2015) 103103 (51pp) doi:10.1088/0954-3899/42/10/103103

Topical Review

The PDF4LHC report on PDFs and LHC data: results from Run I and preparation for Run II

Juan Rojo[1], Alberto Accardi[2,3], Richard D Ball[4,5], Amanda Cooper-Sarkar[6], Albert de Roeck[5,7], Stephen Farry[8], James Ferrando[9], Stefano Forte[10], Jun Gao[11], Lucian Harland-Lang[12], Joey Huston[13], Alexander Glazov[14], Maxime Gouzevitch[15], Claire Gwenlan[6], Katerina Lipka[14], Mykhailo Lisovyi[16], Michelangelo Mangano[5], Pavel Nadolsky[17], Luca Perrozzi[18], Ringaile Plačakytė[14], Voica Radescu[16], Gavin P Salam[5,19] and Robert Thorne[12]

[1] Rudolf Peierls Centre for Theoretical Physics, University of Oxford 1 Keble Road, Oxford OX1 3NP, UK
[2] Hampton University, Hampton, VA 23668, USA
[3] Jefferson Lab, Newport News, VA 23606, USA
[4] The Higgs Centre for Theoretical Physics, University of Edinburgh, JCMB, KB, Mayfield Rd, Edinburgh EH9 3JZ, UK
[5] PH Department, CERN, CH-1211 Geneva 23, Switzerland
[6] Particle Physics, Department of Physics, University of Oxford, 1 Keble Road, Oxford OX1 3NP, UK
[7] Antwerp University, B2610 Wilrijk, Belgium
[8] Department of Physics, University of Liverpool, L69 7ZE Liverpool, UK
[9] SUPA, School of Physics and Astronomy, University of Glasgow, Glasgow, G12 8QQ, UK
[10] TIF Lab, Dipartimento di Fisica, Università di Milano and INFN, Sezione di Milano, Via Celoria 16, I-20133 Milano, Italy
[11] High Energy Physics Division, Argonne National Laboratory, Argonne, IL 60439, USA
[12] Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK
[13] Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, USA
[14] Deutsches Elektronen-Synchrotron (DESY), Notkestrasse 85, D-22607 Hamburg, Germany
[15] Université de Lyon, Université Claude Bernard Lyon 1, CNRS-IN2P3, Institut de Physique Nucléaire de Lyon, Villeurbanne, France
[16] Physikalisches Institut, Universität Heidelberg, Heidelberg, Germany
[17] Department of Physics, Southern Methodist University, Dallas, TX 75275–0181, USA
[18] Eidgenoessische Technische Hochschule (ETH), Hönggerberg 8093 Zürich, Switzerland

E-mail: juan.rojo@physics.ox.ac.uk

Received 3 July 2015
Accepted for publication 20 July 2015
Published 16 September 2015

Communicated by Professor Alan Martin.

Abstract

The accurate determination of the parton distribution functions (PDFs) of the proton is an essential ingredient of the Large Hadron Collider (LHC) program. PDF uncertainties impact a wide range of processes, from Higgs boson characterization and precision Standard Model measurements to new physics searches. A major recent development in modern PDF analyses has been to exploit the wealth of new information contained in precision measurements from the LHC Run I, as well as progress in tools and methods to include these data in PDF fits. In this report we summarize the information that PDF-sensitive measurements at the LHC have provided so far, and review the prospects for further constraining PDFs with data from the recently started Run II. This document aims to provide useful input to the LHC collaborations to prioritize their PDF-sensitive measurements at Run II, as well as a comprehensive reference for the PDF-fitting collaborations.

Keywords: structure functions, LHC, PDF4LHC

(Some figures may appear in colour only in the online journal)

1. Introduction and motivation

The initial state of hadronic collisions is the domain of the parton distribution functions (PDFs) of the proton, see [1–7] for recent reviews. Accurate PDFs are an essential ingredient for LHC phenomenology: PDF uncertainties limit the ultimate accuracy of the Higgs boson couplings extracted from LHC measurements [5, 8], degrade the reach of searches for massive new BSM particles at the TeV scale [9, 10] and are the dominant systematic uncertainties in the determination of fundamental parameters such as the W boson mass or $\sin^2\theta_{\rm eff}$, key ingredients for global stress-tests of the standard model [11–15]. Because they are non-perturbative objects, although their scale dependence is determined by the perturbative DGLAP evolution equations, they need to be extracted from global fits to hard-scattering data. Various PDF fitting collaborations provide regular updates of their QCD analysis: some of the latest PDF releases include ABM12 [16], CT14 [17], CJ12 [18], GR14 [19], HERAPDF2.0 [20], MMHT14 [21] and NNPDF3.0 [22].

A major recent development in PDF fits has been the inclusion of a wide variety of LHC data. Some of the LHC processes that are now used were already part of global PDF fits, mostly measured at the Tevatron, but the LHC data open a wider new kinematical range, as in the case of jet production [23–25] and inclusive electroweak boson production [26, 27]. On the other hand, other types of processes have only become available for PDF fits after their

[19] On leave from CNRS, UMR 7589, LPTHE, F-75005, Paris, France.

measurement at the LHC, like isolated direct photon production [28], W production in association with charm quarks [29–31], top quark pair production [32, 33], open charm and bottom production in proton–proton collisions [34, 35], low- and high-mass Drell–Yan (DY) production [36, 37] and W and Z production in association with jets [38, 39], among others. Remarkably, some of these processes open completely new avenues for PDF fits: data on $W+c$ production provides a clean handle on the strange PDF [29–31, 40] complementary to that of the low energy neutrino data [41], top quark-pair production allows for improved constraints on the large-x gluon complementary to jets [32, 33, 42] and forward heavy-flavour production probes the gluon distribution at small-x [34, 43].

The fact that the LHC provides measurements at different center-of-mass energies also allows one to construct novel observables with useful PDF sensitivity, the ratios and double ratios of cross-sections at different values of $\sqrt{s}$ where several theoretical and experimental uncertainties cancel [44]. This concept has already been validated by measurements by ATLAS of the ratio of jet cross-sections between 7 and 2.76 TeV [24], and by CMS of the ratio of DY cross-sections between 8 and 7 TeV [45]. Other similar ratios, this time between 13 and 8 TeV, will become possible with the availability of Run II data. Another important example of the relevance of LHC data for parton distributions is given by the close interplay between PDFs and the tunes of soft and semi-hard QCD models in the context of LO and next-to-leading order (NLO) event generators, where LHC measurements have been shown to provide invaluable constraints [46, 47]. Finally, PDFs are an important source of theoretical systematic uncertainty to the extraction of standard model parameters at the LHC, for instance the recent direct measurements of the strong coupling constant in the TeV region [48, 49].

Another important development for PDF studies in the recent years has been the availability of PDF fits being carried out also by the ATLAS and CMS collaborations themselves. Thanks to the know-how acquired from the HERA data analyses, and the availability of a public tool for PDF fits, HERAFITTER [50], both collaborations have developed an extensive program of PDF determinations from their own measurements [24, 26, 27, 48]. The aim of these studies is not to provide an alternative to global fits, but rather to study the constraining power of their new measurements on PDFs, ensure that all information in correlated systematics is suitably provided, and to perform checks of the data prior to publication using the QCD analysis as a diagnostic toolbox. In addition, the HERAFITTER developers are also performing a number of independent PDF studies [51, 52], providing useful input to the PDF community.

Given the importance of the LHC data in modern global PDF fits, and the recent restart of the LHC at 13 TeV, it is now timely to summarize what have we learned for PDFs from Run I data, and to set the stage for the corresponding measurements at Run II. One of the aims of this document is thus to review the constraints on PDFs that measurements at the large hadron collider (LHC) during the Run I data-taking have provided. With this motivation, we summarize the relevant measurements from ATLAS, CMS and LHCb, and discuss the available phenomenological studies that quantify the information on PDFs provided by these datasets. Then we move on to discuss the prospects for PDF-sensitive analysis at Run II. We explore how the increase in center-of-mass energy and luminosity can provide new opportunities for PDF studies beyond those available at Run I. We also quantify some of the constraints on PDFs that Run II can provide by means of a profiling analysis using W, Z and $t\bar{t}$ simulated pseudo-data as input, as well as study the impact of Run II inclusive jet data in the CT framework.

The outline of this document is the following. In section 2 we summarize the status of some of the latest PDF releases, with the emphasis on the role of LHC data. In section 3 we review recent studies that have quantified the PDF sensitivity of LHC measurements. In

section 4 a more detailed overview of the relevant measurements and their corresponding constraints on PDFs provided by the ATLAS, the CMS and the LHCb collaborations during Run I is given. In section 5, we discuss the prospects for PDF-sensitive measurements in Run II, including a profiling analysis using W, Z and $t\bar{t}$ simulated pseudo-data, and a study of inclusive jet production in the CT global analysis framework. In section 6 we present practical recommendations for the presentation and delivery of LHC measurements to be used in global PDF analysis. We conclude in section 7 with an outlook on the program of constraining PDFs with LHC data for the coming years.

This report summarizes the discussions that have taken place at various forums, in particular at the regular PDF4LHC meetings, during the last months. It is also indebted to the productive discussions that took place at the *Parton Distributions for the LHC* workshop that took place between the 15th and the 22nd of February 2015 at the Benasque Center for Science Pedro Pascual[20].

We would like to mention that, in parallel with the studies summarized in this report, an update of the benchmark comparisons between different PDF sets, as well as between different methods to combine them [53–55], is also being performed, with the aim of updating the current PDF4LHC recommendations [56, 57] for PDF usage at the LHC Run II. The results of these benchmark comparisons will be presented in a separate report.

2. PDF analysis at the dawn of the LHC Run II

We begin this document with a succinct review of the status of PDF fits at the dawn of the Run II of the LHC. This is specially timely since most PDF groups have provided major updates of their fits in time to be used along the Run II data in both theory predictions and in Monte Carlo simulations used in the data analysis. In this section we summarize the recent developments of these various groups, and emphasize the role that LHC data plays on each of these analyses. The reader is encouraged to consult the original publications for additional information about the updated PDF fits. All the PDF sets discussed below are available from the `LHAPDF6` [58] interface. Further, in this section we also review the development of new tools for PDF analysis.

While an extensive comparison between these updated PDF sets will be presented in the companion PDF4LHC recommendations paper, here for completeness we also show for illustrative purposes some comparisons between recent PDF sets and the corresponding parton luminosities.

2.1. CT14

CT14 provides PDFs at LO, NLO and NNLO [17]. These global PDF fits include LHC data for the first time, from ATLAS, CMS and LHCb, to go along with the data sets used for CT10 [59]. One important, recently released data set from the Tevatron has also been added, that of the D0 W electron asymmetry measurement using the full Run 2 data sample [60]. This data set provides important constraints on u and d quarks at high x. From the LHC experiments, the chosen data sets are vector boson (W, Z) production cross sections and asymmetries, for which NNLO predictions are available; and inclusive jet cross sections, for which the complete NNLO calculation is not yet available, but the estimated impact of the NNLO contributions is small compared to the current experimental uncertainties. The 7 TeV LHC W and Z data allow us to perform better separation of u and d (anti-)quark PDFs at $x \sim 0.02$, and

[20] http://benasque.org/2015lhc/.

also provide an independent constraint on the strangeness PDF, $s(x, Q)$. The LHC jet cross sections have a potential to probe the gluon PDF in a much wider x than the Tevatron ones.

There are a total of 2947 data points included in the NNLO fit, with data from 33 experiments. FastNLO and Applgrid interfaces [61, 62] have been used for quick calculations of NLO matrix elements in the global fits, supplemented by NNLO K-factors (for the NNLO fit). ResBos [63–66] has been used for the calculation of the NNLO K-factors for W/Z and W asymmetry data.

The PDF parametrization is more flexible than that of CT10. The PDFs are expressed as a linear combination of Bernstein polynomials, having an advantage that each basis polynomial peaks within a single x region. This serves for reduction of correlations among the parameters. PDF error sets with a total of 28 eigenvectors are provided at both NLO and NNLO. Correlated systematic errors from the experiments are included, when available, and have impact on some properties, such as the gluon PDF at $x > 0.1$. A central value of $\alpha_S(m_Z^2)$ of 0.118 has been assumed in the global fits at NLO and NNLO, and the PDF sets at alternative values of $\alpha_S(m_Z^2)$ in an expanded range are also provided. Similar to the CTEQ6 analysis [67], two versions of the LO PDFs are supplied, one with 1-loop evolution of α_S, with an input value of $\alpha_S(m_Z^2) = 0.130$; and the other with 2-loop evolution and $\alpha_S(m_Z^2) = 0.118$.

In general, the CT14 PDFs are similar to those from CT10, albeit with a somewhat smaller strange-quark distribution and a softer gluon at high x. Furthermore, CT14 and CT10 differ in the u and d quark distributions at moderate to large x, due to the inclusion of new data, both from the LHC and from the Tevatron, and new parametrization forms. Particular attention is paid to the behavior of the d/u and $\bar{u}/\bar{d}$ ratios in the limit as x approaches 1. The CT14 parameterizations are more likely to predict finite constant values for d/u and $\bar{u}/\bar{d}$ as $x \to 1$, besides the limits of zero or infinity that were preferred with the previous parametrization choices. (This change affects only extrapolations to very large x values that are not covered by the data.) At large x, $u_v(x)$ and $d_v(x)$ both vary as $(1-x)^{a_2}$, with the same value of a_2 for both (but allowing for different normalizations). This is consistent with the expectations of spectator counting rules.

2.2. CTEQ-JLab (CJ12)

The CTEQ-Jefferson (CJ) Laboratory global PDF fits are based on the world data on charged lepton DIS on proton and deuterium targets (including recent Jefferson Lab data), lepton pair production with a proton beam on proton and deuterium targets, W asymmetries in $\overline{p}p$ collisions, and jet production data from the Tevatron. For DIS data, the fits include subleading $\mathcal{O}(1/Q^2)$ corrections, such as target mass and higher twist effects, and nuclear corrections for the deuterium target data. The CJ fits incorporate data down to an invariant final state mass of $W^2 = 3\,\mathrm{GeV}^2$, weaker than the typical $\sim 12\,\mathrm{GeV}^2$ cut considered in the CT14, MMHT14, and NNPDF3.0 analysis, and therefore including much more DIS data from SLAC and Jefferson Lab extending to higher x values. The resulting fits [18, 68, 69] have culminated in the release of the CJ12 PDF sets, valid in the range $10^{-5} \lesssim x \lesssim 0.9$ and available on the CJ collaboration web page[21] as well as through the `LHAPDF6` interface. The fits were performed at NLO in the zero mass variable flavor number scheme, with α_S fixed to the world average value. Full heavy quark treatment and fits of the strong coupling constant will be included in the upcoming CJ15 PDF release, and fits of the relevant LHC data will be considered in a subsequent analysis.

[21] http://www.jlab.org/cj.

The CJ PDFs have been shown to be stable with the weaker cuts on W and Q^2, and the increased DIS data sample (of about 1000 additional points) has led to significantly reduced uncertainties, up to ∼50% on the d quark PDF at large $x \gtrsim 0.6$, where precise data have otherwise been until recently scarce [68]. Since a precise d quark flavor separation at high x depends largely on DIS on deuterium targets, corrections for nuclear Fermi motion and binding effects are included by convoluting the nucleon structure functions with a smearing function computed from the deuteron wave function. As the u quark is well constrained by data on proton targets, the d quark becomes directly sensitive to the nuclear corrections. The effect is a large suppression at high x, and a mild but non-negligible increase at intermediate x [68], still inside the 'safe' region defined by the larger W cut discussed above. These findings have subsequently been confirmed by Ball *et al* [70] and Martin *et al* [21].

The uncertainties on the d quark PDF from theoretical modeling of nuclear corrections (which we refer to as 'nuclear uncertainties') have been quantified in [18, 69]. These range from mild, corresponding to the hardest of the deuteron wave functions (WJC-1) coupled to a 0.3% nucleon off-shell correction, to strong, corresponding to the softest wave function (CD-Bonn) and a large, 2.1% nucleon off-shell correction; the central value corresponds to the AV18 deuteron wave function with a 1.2% off-shell correction [18]. The resulting PDFs are labeled 'CJ12min', 'CJ12max', and 'CJ12mid', respectively. This analysis demonstrates the usefulness of the deuterium data, even in the presence of the nuclear uncertainties that its use introduces.

A further source of theoretical uncertainty was investigated in [18, 69], where a more flexible parametrization was used for the valence d_v quark at large-x, with an admixture of the valence u_v PDF,

$$d_v(x) \to d'_v(x) = a_0^d \Big[d_v(x)/a_0^d + b\, x^c u_v(x) \Big], \tag{1}$$

where a_0^d is the d quark normalization, and b and c are two additional parameters. The result is that the d/u ratio at $x \to 1$ can now span the range $[0, \infty]$ rather than being limited to either 0 or ∞ as in all previous PDF fits. A finite, nonzero value of this ratio is in fact expected from several non-perturbative models of nucleon structure [7, 71, 72]. It is also required from a purely practical point of view because it avoids potentially large parametrization biases on the fitted d quark PDF, as explained in more detail in [73]. An analogous extended d-quark parametrization has been more recently considered also in the CT14 fits [17].

The ratios of the d to u PDFs for the three CJ12 sets are constrained up to $x \approx 0.8$ by the enlarged data set, and when extrapolated to $x = 1$ give the limiting value

$$d/u \underset{x\to 1}{\longrightarrow} 0.22 \pm 0.20(\text{PDF}) \pm 0.10\ (\text{nucl}), \tag{2}$$

where the first error is from the PDF fits and the second is from the nuclear correction models. These values encompass the full range ≈0–0.5 of available theoretical predictions [71, 72]. The impact of the very recent high-precision data on $W \to e + \nu_e$ and reconstructed W asymmetry from the D0 collaboration is being investigated in the context of the CJ15 fits, where the off-shell corrections are fitted to data (instead of being a-priori selected in a theoretically reasonable range) resulting in a further substantial reduction of the nuclear and statistical uncertainty.

The impact and relevance of precise large-x quark PDFs on forward rapidity observables at Tevatron and LHC, as well as on production of large mass particles, has been studied by Brady *et al* [74]. For example, the nuclear uncertainty from the CJ12 analysis becomes relevant for W production at rapidity greater than 2 at the Tevatron, and greater than 3 at the LHC. For particles of heavier mass, such as the putative W' and Z' bosons predicted in

scenarios beyond the standard model, the production cross section becomes sensitive to higher-x PDFs and the nuclear uncertainty may become larger than 20% above the lower mass limit of ≈2.5 TeV set by recent LHC data. This illustrates how nuclear and other large-x theoretical uncertainties may significantly affect the interpretation of signals of new particles and the determination of their properties, which requires a precise calculation of background QCD processes, and motivates further dedicated efforts (such as the upcoming CJ15 analysis) to reduce these uncertainties as much as possible. Conversely, measurements of large rapidity observables at LHCb can have an impact on the determination of large-x PDFs, as discussed in sections 3 and 4.3 in this review.

2.3. HERAPDF2.0

The HERAPDF fits are based only on data collected at the HERA ep collider. During HERA-I and HERA-II running approximately 1 fb^{-1} of data were collected divided roughly equally between e^+p and e^-p scattering. All of the published measurements on inclusive neutral-current (NC) and charged-current (CC) scattering have now been combined into a single coherent data set taking into account correlated systematic uncertainties [20]. This combination includes data taken at different proton beam energies, $E_p = 920, 820, 575, 460$ GeV. The data cover the ranges $6 \times 10^{-7} < x < 0.65$, $0.045 < Q^2 < 50\,000$ GeV^2. The combination led to significantly reduced uncertainties, below 1.5% over the kinematic range $3 < Q^2 < 500$ GeV^2. This combination supersedes the previous combination of only HERA-I data [75].

The availability of precision NC and CC data over such a large kinematic range allows the extraction of PDFs using only ep data without the need for heavy target corrections. The difference between the NC e^+p and e^-p cross sections at high Q^2 constrains the valence PDFs. The high-x CC data separate valence quark flavours. The CC e^-p data allow the extraction of the d_v PDF without assuming strong isospin symmetry. The lower-Q^2 NC data constrain the sea PDF directly and through their scaling violations they constrain the gluon PDF. A further constraint on the gluon comes from the data at different beam energies, which probe the longitudinal structure function F_L. The HERAPDF2.0 is based on the new data combination [20] and supersedes HERAPDF1.0 [75] and 1.5 which were based on previous partial combinations. The HERAPDF2.0 is available at LO, NLO and NNLO on `LHAPDF6`. The experimental uncertainties are presented as 14 pairs of Hessian eigenvectors evaluated by the standard criterion of $\Delta\chi^2 = 1$. For the NLO and NNLO PDFs, 13 further variations are supplied to cover uncertainties due to model assumptions and assumptions on the form of the parametrization. For the NLO and NNLO PDFs the standard value of $\alpha_S(m_Z^2) = 0.118$ but the PDFs are also supplied for values of $0.110 < \alpha_S(m_Z^2) < 0.130$ in steps of 0.001. Needless to say, this final HERA inclusive combination will be the backbone of all future PDF analyses, similarly as the HERA-I combination is the backbone of all available modern PDF sets.

Several further variations of the HERAPDF2.0 are also supplied: HERAPDF2.0HiQ2 for which only data with $Q^2 > 10$ GeV^2 are used to avoid possible bias from low-x, low-Q^2 effects; HERAPDF2.0AG for which the gluon takes form which is imposed to be positive definite for all x for which $Q^2 > 3.5$ GeV^2; HERAPDF2.0FF3A and FF3B, which use two different versions of the fixed flavour number scheme for heavy quarks; and finally HERAPDF2.0Jets which uses additional HERA data on jet production as well as the HERA combined charm data. The charm data mostly serve to constrain the uncertainty on the charm-quark mass parameter and this information is already used in the main HERAPDF2.0 PDFs,

whereas the jet data put further constraints on the gluon PDF, such that a simultaneous fit for $\alpha_S(m_Z^2)$ and the PDFs can be performed, resulting in a competitive determination of $\alpha_S(m_Z^2)$.

Let us briefly discuss also the prospects of future PDF-sensitive measurements from HERA. To begin with, the legacy HERA data on charm and beauty structure functions will be combined to provide further constraints in the heavy flavour sector and on the small-x gluon PDF. In addition, the final data on prompt photon production will provide information about QED contributions to PDFs, in particular thanks to photon-initiated processes [76, 77]. Also, the legacy HERA data on vector meson production and diffractive di-jet production will elucidate physics at low-x and thus address the question as to whether it is appropriate to include such data in PDF fits using the current conventional DGLAP formalism, or if instead a BFKL type of approach is necessary.

2.4. MMHT2014

The MMHT2014 PDF sets were released in December 2014 [21]. They are the first major update based on this framework since the MSTW2008 PDFs [78]. However, the updates incorporate the improvements to the parametrization and deuteron corrections already presented in [79]. This study showed that the new parameterizations, which use Chebyshev polynomials in $(1 - 2\sqrt{x})$ rather than simple powers of $\sqrt{x}$ and up to seven free parameters for a particular PDF, can reproduce functions obtained from a much greater number of parameters up to a small fraction of percent over a wide range in x. The more flexible deuteron corrections improve the fit quality and result in a shape similar to the models in e.g. [18]. The new PDF sets also use the optimal variable flavour number scheme of [80], updated heavy nucleus corrections [81], a modified central value and uncertainty for the branching ratio $B_\mu = B(D \to \mu)$ used in the determination of the strange quark from dimuon data, and use the multiplicative rather than additive definition for correlated systematic uncertainties [82].

The data used in the fits have been very significantly updated from the MSTW2008 analysis, with relevant data sets published before the beginning of 2014 included, as summarized in section 4. In particular, the combined HERA total cross section data [75] and combined charm data [83] are now used, along with some updates on Tevatron W production. Moreover, a variety of LHC data, including W, Z and $\gamma^\star$ data from ATLAS, CMS and LHCb, inclusive jet data from ATLAS and CMS and total top quark-pair production from ATLAS and CMS (and the Tevatron combined result [84]) have now been included. Although they have not been used to determine the PDFs, data on $W + c$ production and differential top quark-pair production have been checked against the QCD predictions using these PDFs and give good agreement. NLO calculations are produced for LHC data using [61, 62] and K-factors employed at NNLO. The LHC inclusive jet data are currently not used at NNLO, leading to a very slight increase in the uncertainty for the high- x gluon at NNLO compared to NLO.

The resulting PDFs are made available together with 25 eigenvector pairs of uncertainties given at 68% confidence level at LO, NLO and NNLO and correspond to $\alpha_S(m_Z^2))$ values of 0.130 at LO, 0.118 and 0.120 at NLO and 0.118 at NNLO. The increase in the number of eigenvectors from 20 in the MSTW2008 sets is related to the increased flexibility of the PDFs, partially made possible by extra constraints coming from new LHC processes. The value of $\alpha_S(m_Z^2)$ is left free in fits in the first instance, resulting in best fits near 0.135 at LO, 0.120 at NLO and 0.118 at NNLO. Therefore, the choice of the values for the eigenvector sets, though 0.118 is available as well as 0.120 at NLO since this is close to the world average for $\alpha_S(m_Z^2))$, and a set with this value may be required by users. Each eigenvector set is

accompanied by a central set with $\alpha_S(m_Z^2))$ values with ± 0.001 in order to enable uncertainties due to variations of $\alpha_S(m_Z^2))$ to be calculated.

A dedicated study about the uncertainties in $\alpha_S(m_Z^2))$ in the MMHT14 analysis has been presented in [85], and the corresponding sets with a wide variety of $\alpha_S(m_Z^2))$ values have also been released. PDF sets with a variety of values of charm and bottom mass values will also follow soon, as well as PDF sets in the three and four flavour schemes. The MMHT2014 PDFs generally give similar results for the LHC observables as the MSTW2008 PDFs, and have comparable uncertainties. The main change in the MMHT2014 PDFs is in the small-x valence quarks, related to the improved parameterization and deuteron corrections, and an increase in the uncertainty (and to a lesser extent the central value) of the strange quark. The latter is due to the quite generous uncertainty allowed on $B_\mu = B(D \to \mu)$ and an extra free parameter for the strange quark contributing to the eigenvectors. This combination of extra freedom in the strange PDF is then given an extra constraint by the LHC W- and Z-boson production data.

2.5. NNPDF3.0

The NNPDF3.0 sets were released in October 2014 [22]. As compared with previous NNPDF global analysis [86–91], NNPDF3.0 is the result of a extensive redevelopment of the NNPDF code, including constraints from new experimental data, theoretical calculations of new processes, and a major code re-organization. Furthermore, NNPDF3.0 is the first set of global PDFs with a fitting methodology validated through a closure test.

Regarding experimental data, NNPDF3.0 includes the fixed target, HERA, Tevatron and LHC data already included in NNPDF2.3, and in addition all the published HERA-II data from H1 and ZEUS, and a wide range of more recent ATLAS, CMS and LHCb data on jet production, weak boson production and asymmetries, DY, W+charm and top quark pair production. A total of 4276 data points are fitted at NLO and 4078 at NNLO. The complete list of LHC measurements that have been included in NNPDF3.0 is summarized in section 4.

Concerning theory calculations, all collider processes have been computed using fast NLO interfaces [61, 62, 92], supplemented by NNLO and electroweak K-factors when required. Inclusive jets are treated at NNLO using the approximate threshold calculation [93, 94], validated on the exact calculation in the gg channel [95]. Heavy quark mass effects are computed in the FONLL General-Mass variable-flavor number scheme [96], with the main difference being that at NLO it is the FONLL-B scheme that is used, rather than FONLL-A as previously, since this provides a better description of the low-Q^2 charm production data.

All the fitting code has been rewritten from FORTRAN to C++ and PYTHON, making it robust and modular, so that the modification of the theoretical calculations, the addition of new datasets, or the generation of entirely new sets of pseudo-data for use in closure testing, can be done easily and quickly with no need to modify the rest of the code. Improved positivity constraints and dynamical preprocessing exponents have also been implemented.

As both data and theory improve, it becomes even more necessary to ensure that the fitting methodology is consistent and unbiased. To this end the NNPDF methodology has now been subjected to a closure test [22]. This is performed by generating pseudodata based on an assumed prior PDF (for example MSTW08) and a particular theory (for example NLO perturbative QCD with given α_S, heavy quark scheme, etc). To make the test as realistic as possible, the pseudodata are generated using the experimental uncertainties of the current global dataset. In the context of the closure test, the pseudodata are 'perfect': in particular they are fully consistent both with each other and with the assumed theory. A full fit to the

pseudodata is thus a rigorous test of the fitting methodology: fitted PDFs should have manifestly unbiased central values, and statistically meaningful uncertainties and correlations (so that for example the fitted PDF agrees with the assumed prior at one sigma 68% of the time).

The success of the NNPDF3.0 closure test proves that the NNPDF3.0 PDF sets fitted to real data are unbiased and have statistically meaningful uncertainties and correlations. This in turn confirms that most modern datasets are consistent (both internally and in the context of the global dataset), in the sense that their systematic errors have been sensibly estimated, and furthermore that NNLO QCD is sufficient to describe the global dataset within a common universal framework. In particular the NNPDF3.0 fits to real data show no sign of tension between deep inelastic and hadronic data.

The NNPDF3.0 PDFs are available on `LHAPDF6` as sets of 100 replicas. Baseline fits are available for $\alpha_S(m_Z) = 0.118$ at LO, NLO and NNLO, and also at $\alpha_S(m_Z) = 0.115$, 0.117, 0.119, 0.121 at NLO and NNLO. A LO fit with $\alpha_S(m_Z) = 0.130$ is also provided. The baseline fits are provided with 5 active flavours: alternative fits with $N_f = 3$, 4, 6 active flavours are also available. NNPDF also provide fits to reduced datasets, for studies by the LHC experimental collaborations: HERA-only, HERA+ATLAS, HERA+CMS, no-LHC, and no-jets. All these are available at NLO and NNLO, for $\alpha_S(m_Z) = 0.117$, 0.118, 0.119. The fits with $\alpha_S(m_Z) = 0.118$ are also provided with 1000 replicas, for use in reweighting studies. In addition, improved delivery tools, such as reducing the number of replicas [55], or by provision of Hessian eigenvectors [54], have also become available recently.

2.6. PDF analysis tools

When performing a QCD analysis to determine PDFs there are various assumptions and choices to be made concerning, for example, the functional form of the input parametrization, the treatment of heavy quarks and their mass values, alternative theoretical calculations or representations of the fit quality estimator, χ^2, and for different ways of treating correlated systematic uncertainties. It is useful to discriminate or quantify the effect of a chosen anstaz within a common framework and the `HERAFitter`, an open source QCD fit analysis project [50, 97], is optimally designed for such tests.

`HERAfitter` incorporates results from a wide range of experimental measurements in lepton–proton deep inelastic scattering, proton–proton and proton–antiproton collisions. These are complemented with a variety of theoretical options for calculating PDF-dependent cross section predictions corresponding to the measurements. The framework covers a large number of the existing methods (e.g. `fastNLO` and `APPLgrid`, described later in this section) and schemes used for PDF determination. The data and theoretical predictions are confronted by means of numerous methodological options for performing PDF fits and plotting tools to help visualize the results. For example, recently the `HERAFitter` framework has been used to study the consistency of the legacy measurements of the *W*-boson charge asymmetry and of the *Z*-boson production cross sections from Tevatron with the NLO QCD theoretical predictions, which are found in good agreement [98] and illustrate the importance of the Tevatron data to constrain the *d*-quark and the valence PDFs. In summary, with sufficient options to reproduce the majority of the different theoretical choices made in global PDF fits, `HERAFitter` is a valuable tool for benchmarking and understanding differences in the phenomenology of PDF fits by different groups and it can be used to study the impact of new precision measurements at hadron colliders.

Precise measurements require accurate theoretical predictions in order to maximize their impact in PDF fits. Perturbative calculations become more complex and time-consuming at higher orders due to the increasing number of relevant Feynman diagrams. The direct inclusion of computationally demanding higher-order calculations into iterative fits is thus not possible currently. However, a full repetition of the perturbative calculation for small changes in input parameters is not necessary at each step of the iteration. Two methods have been developed which take advantage of this to solve the problem: the *K*-factor technique and the *fast grid* technique.

In the *K*-factor method, the ratio of the prediction of a higher-order pQCD calculation, usually time-consuming, to a lower-order calculation using the same PDF, are estimated once for a given PDF, stored into a table of *K*-factors, and applied multiplicatively to the theory prediction derived from the fast lower-order calculation throughout the iterative process in minimizing the χ^2. Hence, this technique avoids iteration of the higher-order calculation at each step. This procedure, however, neglects the fact that the *K*-factors are PDF dependent, and as a consequence, they have to be re-evaluated for the newly determined PDF at the end of the fit until input and output *K*-factors have converged (typically 2–3 iterations are needed). This method has been used for the NNLO QCD fits to the DY measurements.

In the *fast grid* method, a generic PDF can be approximated by a set of interpolating functions with a sufficient number of support points. The accuracy of this approximation is checked and optimized such that the approximation bias is negligible compared to the experimental and theoretical accuracy. Hence, this method can be used to perform the time consuming higher-order calculations only once for the set of interpolating functions. Further iterations of the calculation for a particular PDF set are fast, involving only sums over the set of interpolators multiplied by factors depending on the PDF. This approach can be used to calculate the cross sections of processes involving one or two hadrons in the initial state and to assess their renormalization and factorization scale variation.

There are three projects most commonly used to exploit the described techniques: `FastNLO` [99, 100], `APPLgrid` [61, 101] and `aMCfast` [92]. The packages differ in their interpolation and optimization strategies, but they all construct tables with grids for each bin of an observable in two steps: in the first step, the accessible phase space in the parton momentum fractions x and the renormalization and factorization scales μ_r and μ_f is explored in order to optimize the table size. In the second step the grid is filled for the requested observables. Higher-order cross sections can then be obtained very efficiently from the pre-computed grids while varying externally provided PDF sets, μ_r and μ_f, or $\alpha_S(\mu_R)$. This approach can be extended to arbitrary processes. This requires an interface between the higher-order theory programs and the fast interpolation frameworks.

The open-source project `fastNLO`[22], has been interfaced to the `NLOjet++` program [102] for the calculation of jet production in DIS [103] as well as 2- and 3-jet production in hadron–hadron collisions at NLO [104, 105]. Threshold corrections at 2-loop order, which approximate NNLO for the inclusive jet cross section for pp and $p\bar{p}$, have also been included into the framework [62] following [106]. The latest version of the `fastNLO` convolution program [107] allows for the creation of tables in which renormalization and factorization scales can be varied as a function of two pre-defined observables. More recently, the differential calculation of top-pair production in hadron collisions at approximate NNLO [33] has been interfaced to `fastNLO` [108].

In the `APPLgrid` package [109], in addition to jet cross sections for $pp\,(p\bar{p})$ and DIS processes, calculations of DY production and other processes are also implemented using an

[22] http://fastnlo.hepforge.org.

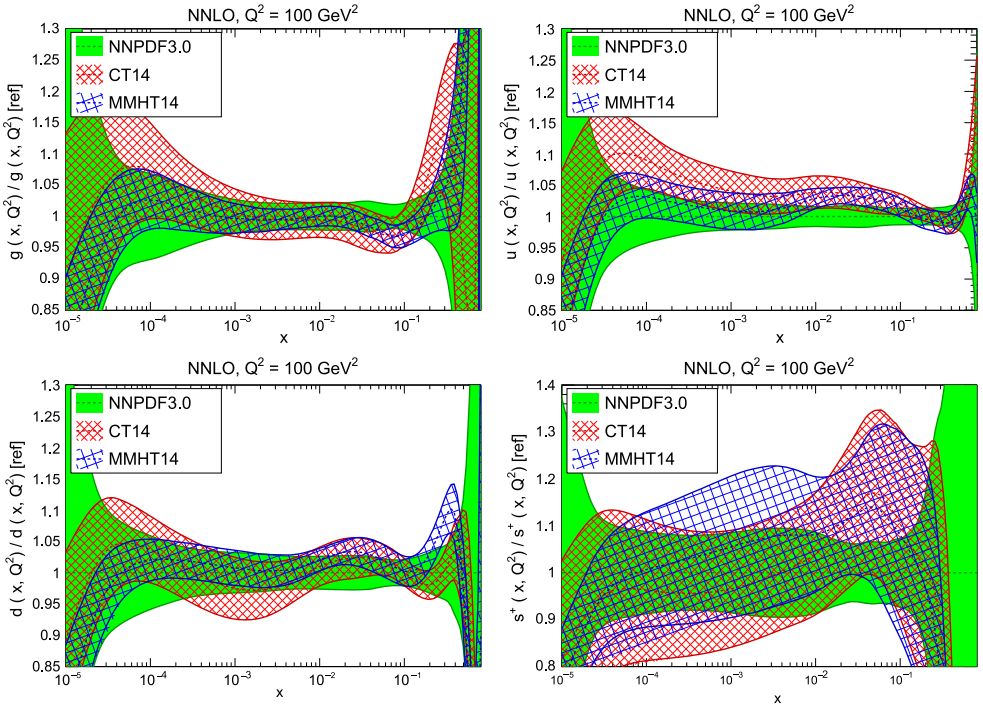

Figure 1. Comparison of PDFs at $Q^2 = 10^2$ GeV^2 between the NNPDF3.0, CT14 and MMHT14 sets, all of them at NNLO, with $\alpha_S(m_Z^2)) = 0.118$. From top to bottom, and from left to right, we show the gluon, the up quark, the down quark, and the total strangeness PDFs. Results are shown normalized to the central value of NNPDF3.0.

interface to the `MCFM` parton level generator [110–112]. Variation of the renormalization and factorization scales is possible *a posteriori* when calculating theory predictions with the `APPLgrid` tables, using the `HOPPET` program [113], and independent variation of α_S is also allowed. The `aMCfast` project is based on using the same `APPLgrid` interpolation methods within the fully automated `MadGraph5_aMC@NLO` [114] framework to achieve the automation of fast NLO QCD calculations for PDF fits, for arbitrary processes. Work in progress in the `aMCfast` code is directed towards achieving the same automation for NLO calculations matched to parton showers and to the inclusion in PDF fits of generic NLO electroweak corrections.

2.7. Comparison of PDFs and parton luminosities

To conclude this section, we compare some of the recent releases from the various PDF groups in terms of PDFs and parton luminosities. For this purpose, the `APFEL-Web` online PDF plotting interface [115, 116] has been used. First, we compare the NNPDF3.0, CT14 and MMHT14 NNLO sets in figure 1, at a scale of $Q^2 = 100$ GeV^2. From top to bottom, and from left to right, we show the gluon, the up quark, the down quark, and the total strangeness PDFs. Results are shown normalized to the central value of NNPDF3.0.

The comparisons in figure 1 indicate that there is reasonable agreement at the level of the one standard deviation of the PDF uncertainties between the three groups. In some cases the agreement is only marginal, for instance for the d PDF at large-x. In general, it is at small and

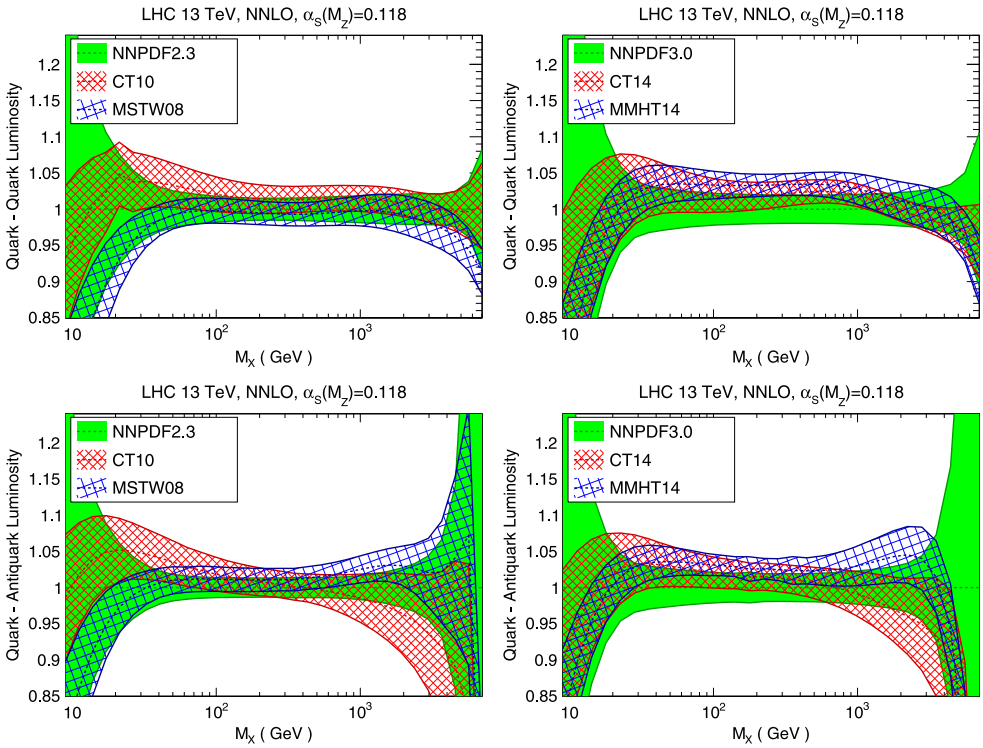

Figure 2. Comparison of the quark–quark (upper plots) and quark–antiquark (lower plots) PDF luminosities between NNPDF2.3, MSTW08 and CT10 (left column) and the more recent NNPDF3.0, MMHT14 and CT14 (right column) PDF sets. The comparison has been performed at NNLO and $\alpha_S(m_Z^2)) = 0.118$ for the LHC with a center of mass energy of 13 TeV, as a function of the invariant mass of the final state system M_X. Results are shown normalized to the central value of the NNPDF sets.

large values of x, in regions with limited kinematical coverage, that the differences between the three fits are more marked. For some PDF combinations, the size of the PDF uncertaintiy can also show differences between the three groups, such in the total strangeness s^+ PDF at intermediate values of x.

Next, we turn to a comparison of PDF luminosities for the LHC at a center of mass energy of 13 TeV. To illustrate the differences between the previous releases from NNPDF, CT and MSTW/MMHT, in figure 2 we compare the quark–quark and quark–antiquark luminosities in NNPDF2.3, CT10 and MSTW08 with the same results from NNPDF3.0, CT14 and MMHT14. The corresponding comparison for the gluon–gluon and quark–gluon luminosities is shown in figure 3. The comparison has been performed at NNLO and $\alpha_S(m_Z^2) = 0.118$, as a function of the invariant mass of the final state system M_X. Results are shown normalized to the central value of the NNPDF sets.

Comparing the newer and the older PDF sets, we notice that in general there has been improved agreement between the three sets in a number of phenomenologically important regions, like the gg luminosity at intermediate values of the final-state invariant mass M_X. For the four luminosities that are compared here, the three PDF sets agree at the one-sigma level or better in all the relevant range of M_X values. The differences are larger at large invariant masses, a key region for massive New Physics searches at the LHC, where also the intrinsic

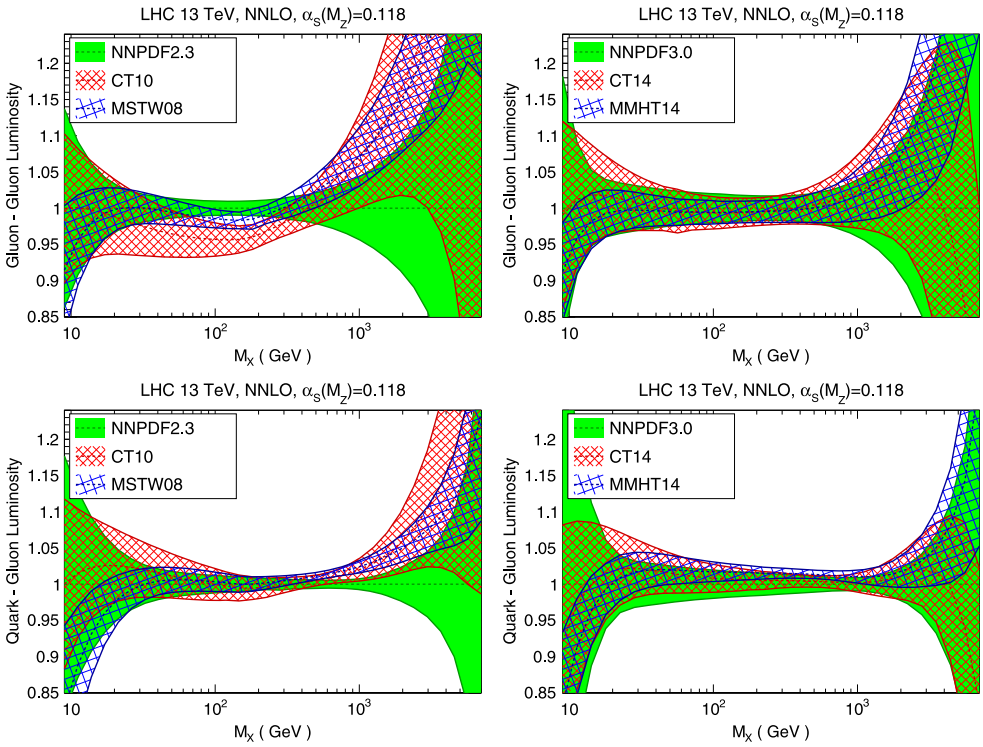

Figure 3. Same as figure 2 now for the gluon–gluon (upper plots) and quark–gluon (lower plots) PDF luminosities

PDF uncertainties for each group are substantial due to the lack of experimental constraints. We also find that in some cases, like the quark–quark luminosity, the agreement is only marginal, driven by the differences at the level of u and d PDFs observed in figure 1.

3. Overview of PDF-sensitive measurements at the LHC

In this section we review LHC processes relevant for PDF constraints, summarized in table 1. Emphasis is put on information not previously accounted for in PDF fits. For each process, the sensitivity to specific PDF flavours is briefly described and the probed ranges of x and Q^2 are listed. The corresponding measurements performed by the ATLAS, CMS and LHCb collaborations, when already available, are presented in the next section.

3.1. Jet production

Jet production allows one to constrain quarks and gluons at medium and large-x, for $x \gtrsim 0.005$ [17, 117], a region where the constraints from deep-inelastic scattering data are only indirect. Inclusive jet production has been used in PDF fits since the first measurements at the Tevatron. Nowadays, a number of precise LHC measurements of inclusive jet, dijet and trijet production are available. In addition, jet production provides a unique possibility for direct determinations of the strong coupling $\alpha_S(Q)$ in the TeV range way above any other existing measurements, providing information on BSM physics.

Table 1. Summary of LHC processes sensitive to PDFs. For each process, we quote the corresponding measured distribution, the PDFs that are probed, and the approximate ranges of x and Q^2 that can be accessible using available Run I data. These ranges have been obtained assuming the Born kinematics.

Reaction	Observable	PDFs	x	Q
$pp \to W^{\pm} + X$	$\mathrm{d}\sigma(W^{\pm})/\mathrm{d}y_l$	q, $\bar{q}$	$10^{-3} \lesssim x \lesssim 0.7$	$\sim M_W$
$pp \to \gamma^*/Z + X$	$\mathrm{d}^2\sigma(\gamma^*/Z)/\mathrm{d}y_{ll}\mathrm{d}M_{ll}$	q, $\bar{q}$	$10^{-3} \lesssim x \lesssim 0.7$	5 GeV $\lesssim Q \lesssim$ 2 TeV
$pp \to \gamma^*/Z + \mathrm{jet} + X$	$\mathrm{d}\sigma(\gamma^*/Z)/\mathrm{d}p_T^{ll}$	q, g	$10^{-2} \lesssim x \lesssim 0.7$	200 GeV $\lesssim Q \lesssim$ 1 TeV
$pp \to \mathrm{jet}+X$	$\mathrm{d}\sigma(\mathrm{jet})/\mathrm{d}p_T\,\mathrm{d}y$	q, g	$10^{-2} \lesssim x \lesssim 0.8$	20 GeV $\lesssim Q \lesssim$ 3 TeV
$pp \to \mathrm{jet} + \mathrm{jet}+X$	$\mathrm{d}\sigma(\mathrm{jet})/\mathrm{d}M_{jj}\,\mathrm{d}y_{jj}$	q, g	$10^{-2} \lesssim x \lesssim 0.8$	500 GeV $\lesssim Q \lesssim$ 5 TeV
$pp \to t\bar{t}+X$	$\sigma(t\bar{t})$, $\mathrm{d}\sigma(t\bar{t})/\mathrm{d}M_{t\bar{t}}$, ...	g	$0.1 \lesssim x \lesssim 0.7$	350 GeV $\lesssim Q \lesssim$ 1 TeV
$pp \to c\bar{c}+X$	$\mathrm{d}\sigma(c\bar{c})/\mathrm{d}p_{T,c}\mathrm{d}y_c$	g	$10^{-5} \lesssim x \lesssim 10^{-3}$	1 GeV $\lesssim Q \lesssim$ 10 GeV
$pp \to b\bar{b}+X$	$\mathrm{d}\sigma(b\bar{b})/\mathrm{d}p_{T,c}\mathrm{d}y_c$	g	$10^{-4} \lesssim x \lesssim 10^{-2}$	5 GeV $\lesssim Q \lesssim$ 30 GeV
$pp \to W+c$	$\mathrm{d}\sigma(W+c)/\mathrm{d}\eta_l$	s, $\bar{s}$	$0.01 \lesssim x \lesssim 0.5$	$\sim M_W$

Jet production can be presented in a number of complementary ways, the most traditional are the measurements of the inclusive jet and dijet cross-sections, but measurements of three-jet and multi-jet cross-sections have also became available recently. The impact of ATLAS and CMS jet data on PDFs has been quantified in a number of studies, both from global PDF fitting groups [17, 22, 23] and from the LHC collaborations themselves [24, 48]. In addition to the gluon, also information on the large-x quarks can be obtained, since the quark–quark scattering mechanism dominates at the highest values of the jet transverse momentum, due to the steeper fallout of the gluon PDF at large-x.

While the NLO calculation for inclusive and dijet production has been available for more than 20 years, only recently the first partial results on the NNLO calculation have become available [95, 118]. While the full calculation is not yet available, it has been proposed that a subset of jet data can still be consistently included in NNLO fits by using the approximate NNLO threshold calculation. This strategy, presented in [94], has been used to include LHC jet data in the NNPDF3.0 fit.

In figure 4 we illustrate the impact of the CMS 2011 inclusive jet data on the large-x gluon PDF, from [48]. In the same figure we also show the constraints that the ATLAS measurement on the ratio of inclusive jet cross-sections between 7 TeV and 2.76 TeV imposes on the gluon PDF, from [24]. In both cases, the PDF fits have been performed using the HERAFitter framework.

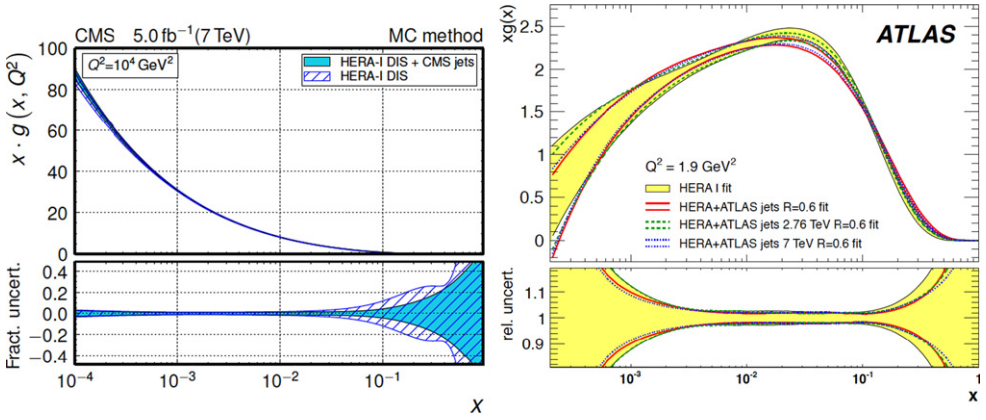

Figure 4. Left plot: impact of the CMS 2011 inclusive jet data on the gluon PDF, from [48]. Right plot: impact of the ATLAS measurement on the ratio of inclusive jet cross-sections between 7 TeV and 2.76 TeV on the gluon PDF, from [24].

3.2. Prompt photon production

Direct photon production measurements from fixed-target experiments were recognized as a useful probe of the gluon PDFs a long time ago [119], but their use was very limited, because of inconsistencies between the various experiments, especially after the competitive Tevatron jet production data were published. More recently, the use of LHC isolated photon data in the PDF fits was advocated in [28], where a reanalysis of all available fixed target and collider isolated direct-photon production data was performed, finding good consistency with NLO QCD calculations. In [28] it was also shown that direct photon data can potentially constrain the gluon PDF in an intermediate range of x, around $\sim$0.01, which is the region relevant for Higgs-boson production via gluon fusion. The main obstacle for the full inclusion of direct photon data into PDF fits is the large scale uncertainties that affect the NLO QCD calculation. The possibility to use isolated photon production in association with additional jets has also been explored [120], however a substantial reduction of the experimental uncertainties, with respect to that of available measurements, would be needed before this data could be used effectively in the PDF fits. While LHC photon data has still not been directly included in global PDF fits, a systematic comparison between different PDF sets and direct photon data was presented by ATLAS [121].

3.3. Inclusive W and Z production and asymmetries

Inclusive production of W and Z bosons, presented in the form of total cross sections, differential distributions in leptonic rapidities, and corresponding asymmetries, has been important in the global PDF fits since the first such measurements were made at the Tevatron. As compared to inclusive DIS, where only flavor symmetric components $q + \bar{q}$ can be constrained, inclusive W and Z production provides a clean handle on quark flavour separation. At the LHC, the kinematical range in terms of the underlying x has substantially increased as compared to the Tevatron, reaching both smaller and larger values of x. To pin down the PDF quark flavor separation, a number of measurements have been presented by ATLAS, CMS and LHCb, as will be discussed in more detail in section 4. In addition, as shown by ATLAS, once the rapidity distributions of W and Z bosons are measured

simultaneously accounting for the correlated systematics between the various distributions [122], an additional handle on the strangeness content of the nucleon can be provided [27].

3.4. High and low mass DY production

Data from fixed-target DY experiments, such as E605 [123] and E866 [124], have been included in global PDF fits since many years. However, these data are affected by some drawbacks since they miss information on the systematic correlations, and because $\sqrt{s}$ is small, thus leading to potentially large perturbative and non-perturbative corrections to fixed-order calculations. This has provided the motivation to perform, for the first time, the measurements of the off-peak DY processes at a hadron collider. At low mass, DY provides interesting constraints on the low-x quarks and gluons (occurring in gluon radiation from the quarks), as well as tests of perturbative QCD, like the possible breakdown of DGLAP evolution [125]. The high-mass region, instead, provides information from the high-x quarks and anti-quarks, which are affected by substantial uncertainties (the latter in particular). To maximize the impact of the data in PDF fits, it is essential to use the most updated NNLO QCD and NLO EW theory calculations [126, 127].

In addition, both the low and high-mass DY production are sensitive to the photon PDF $\gamma(x, Q^2)$: indeed, since $\gamma\gamma \to l^+l^-$ production is t-channel, as opposed to the s-channel quark-induced diagrams $q\bar{q} \to l^+l^-$, photon-initiated contributions become comparable to quark-initiated at low and high invariant di-lepton masses. Therefore, off-peak DY production provides important constraints on the photon PDF [128] for PDF fits which account for QED corrections [76, 129].

Finally, the high-mass region provides a crucial validation of theory calculations in a region which is instrumental for new physics searches.

3.5. The transverse momentum of W and Z bosons

The transverse momentum, p_T, of the W and Z bosons is a key observable for hadron collider phenomenology. At low p_T, it is used to validate Monte Carlo predictions, analytically resummed calculations, and is important for many precision measurements like the W mass. At large values of the transverse momentum, we would expect that fixed-order theory provides a reasonable description of the data. For large p_T, the transverse momentum distribution of W and Z bosons uniquely depends on the combination $\alpha_S \times q \times g$, where the fraction of gluon-initiated contributions increases with p_T. Therefore, one might want to use the high-p_T spectrum of W and Z bosons as a direct probe of the gluon PDF. This option seems particularly robust for the case of the Z p_T, where a high precision measurement can be performed in terms of leptonic variables only [130].

One possible issue for the inclusion of these measurements in PDF fits is that available data on the Z boson transverse momentum from ATLAS [130] and CMS [131] exhibit a $\mathcal{O}(10\%)$ discrepancy with pure NLO calculations in the region around 100 GeV, where the accuracy of the experimental measurement is around $\mathcal{O}(1\%)$. In this respect, having the full NNLO results for the Z p_T will shed light on the origin of this discrepancy, though available results for the NNLO W+jets calculation [132] suggest that higher-order corrections on top of NLO might not be enough to explain the differences observed between theory and data .

The measurements of the ratios of W and Z cross sections as a function of boson p_T would provide additional information on PDFs. As motivated in [38], various ratios of W and Z cross sections at high p_T provide a handle on the proton's flavour decomposition, while cancelling various theoretical uncertainties like higher order QCD and EW effects. In figure 6

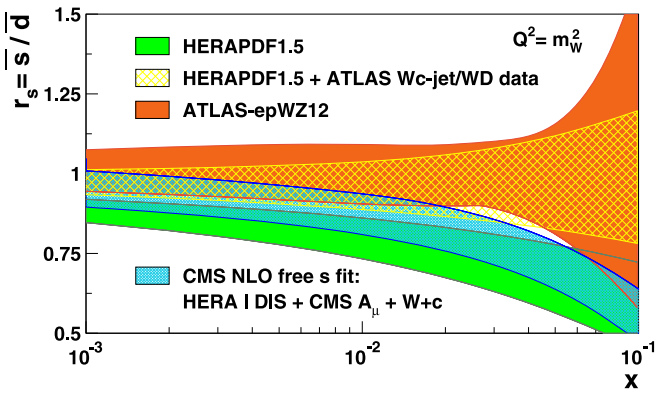

Figure 5. The ratio of the s and d PDFs, as a function of x, compared in different analysis. We show the ATLAS results based on Z and W -boson production measurements [27] and on the associated production of W with charmed hadrons [30] as well as the CMS result [26], based on the $W + c$ production measurement of [29]. For comparison, the HERAPDF1.5 result is also shown, where the constraints on the strange quark distribution are obtained from the neutrino-scattering experiments.

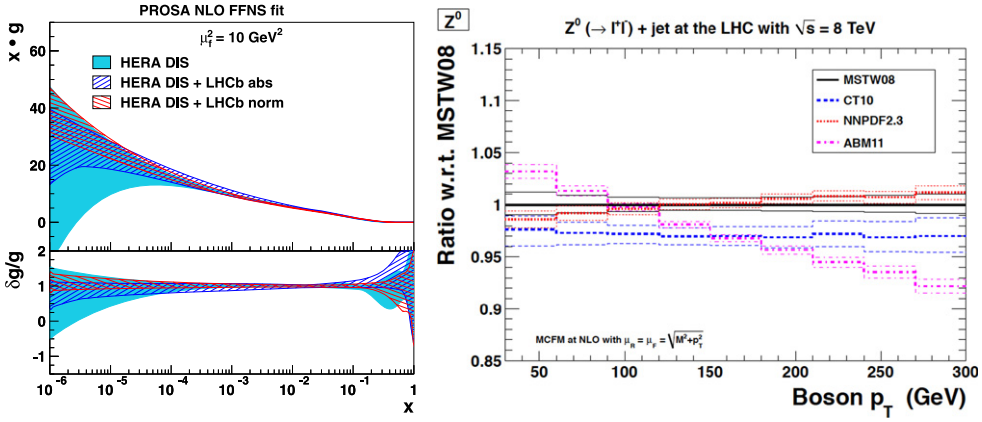

Figure 6. Left plot: gluon distribution (top) and its relative uncertainty (bottom), shown as a function of x at the factorization scale of 10 GeV^2, comparing the results of the NLO QCD analysis [34] of HERA-I DIS data only (filled band) and that including either absolute or normalized cross sections of heavy-flavor production at LHCb (dashed bands). Right plot: the transverse momentum distribution of Z bosons at the LHC 8 TeV computed at NLO using various sets of PDFs, shown as ratio to the MSTW08 prediction, from [38]

we show the transverse momentum distribution of Z bosons at the LHC 8 TeV computed at NLO using various sets of PDFs, shown as ratio to the MSTW08 prediction [38].

In addition, a number of measurements of W and Z boson production in association with jets has been performed at the LHC. The main motivation for these measurements is to validate Monte Carlo event generators, but given the fact that the underlying dynamics are the same as those that generate the vector-boson p_T, it is conceivable that these can also be used for PDF fits. However, these measurements will be affected by larger theoretical uncertainties (due to the higher final state multiplicity) and experimental uncertainties (due to the presence

of jets) than the inclusive W and Z p_T measurement, and thus might not be competitive with the latter.

3.6. W production in association with charm quarks

Production of W bosons in association with charm quarks has been proposed for a long time as a direct probe of the strangeness content of the proton [133]. Indeed, before the LHC start-up, constraints on strangeness from global fits were provided mostly by low-energy neutrino data, in particular by the measurements of charm production through di-muon final states [87, 134]. At the LHC, independent constraints on the strangeness can now be provided by the measurement of the $W + c$ process cross section [40], and also the differences between s^+ and s^- content can be potentially be investigated with cross section ratios such as $(W^+ + c)/(W^- + c)$. As we will discuss in the next section, this measurement has been recently published by both ATLAS and CMS, and is already part of several global PDF fits.

A topic that has attracted sizable attention recently is whether the LHC $W + c$ data suggest a symmetric strange sea, opposite to neutrino charm data which clearly indicates a strangeness suppression. In figure 5 we show the strangeness fraction in the quark sea, obtained by ATLAS and CMS by using inclusive W and Z measurements and the $W + c$ data. In addition we show the HERAPDF1.5 result, where the constraints on the strange quark distribution are obtained from the neutrino-scattering experiments. While CMS data prefer a suppressed strangeness, and the ATLAS measurements indicate a symmetric light quark sea, both results are consistent within uncertainties. Moreover, recent global analyses combining both fixed-target and collider data sensitive to the strangeness [22, 31] demonstrated the general consistency of the LHC data among each other and with the measurements of neutrino experiments in the x-range accessible by the LHC measurements. Future, higher precision data from Run II will shed more light on this issue.

3.7. Top quark pair production

Top quarks are abundantly produced at the LHC, which can be considered a real 'top factory' due to the high center of mass energy and luminosity. As opposed to the Tevatron, where top quark pairs are produced predominantly via quark–anti-quark annihilation, at the LHC they are produced mostly in the gluon-gluon channel. Therefore, they provide potentially useful information on the gluons for $x \geqslant 0.1$, a region which is only covered by jet production in PDF global fits. In addition, for differential distributions sensitive to large-x PDFs, such as the $t\bar{t}$ invariant mass distribution or the tail of the p_T^t distribution, there is also sensitivity to quarks and anti-quarks.

While NLO calculations are affected by large scale uncertainties, the completion of the full NNLO calculation for total production cross-sections [135] and for differential distributions [136, 137] will allow for consistent use of the top quark-pair data in the fits at NNLO. Furthermore, their availability allows for more precise extractions of fundamental QCD parameters, like top-quark mass and α_S [138]. Since the exact differential NNLO calculation is not yet available in a form suitable for QCD analyses, its approximate version [33], featuring the methods of threshold resummation, might be used.

Up to now, a number of studies has quantified the sensitivity of top quark pair production data to the gluon PDFs using the total top-quark pair production cross-sections, showing that available data from ATLAS and CMS already provide powerful constrains on the large-x gluon [32, 42]. Among other collaborations that include top data in their fits, ABM has explored their impact showing that it can lead to a shift in the gluon PDF up to one-sigma [16]

in units of the PDF uncertainties. The impact of total cross-sections in PDF fits is only moderate, but the full constraining power of top quark data will be assessed using the differential distributions. A first study on this respect, based on the approximate NNLO from threshold-resummed calculation, has been presented in [33].

3.8. Charm and bottom pair production

Production of heavy quark pairs in hadron collisions is a powerful test of perturbative QCD. While top pair production at the LHC is nowadays included in PDF fits, this is not the case for charm and bottom quarks. On the other hand, their differential p_T and rapidity distributions ($\mathrm{d}^2\sigma/\mathrm{d}p_T\,\mathrm{d}y$) are directly sensitive to the small-x gluon PDF at low scale. The downside is a large scale dependence of theoretical predictions (currently available only at NLO), which, however, may be mitigated by analyzing ratios of differential rates in various experimental bins, instead of their absolute values [34].

The constraints are expected to be particularly powerful in measurements in the LHCb acceptance. The LHCb detector is suitable to precisely tag and measure the properties of heavy quark mesons. Its forward coverage allows one to access the small-x region that cannot be accessed by ATLAS and CMS. A first study in this direction has been performed by the PROSA Collaboration [34], which evaluated the impact of recent measurements of heavy-flavour production at LHCb using the HERAFitter framework in a QCD analysis at NLO in a fixed-flavour number scheme. Significant reduction of the gluon and the sea-quark distribution uncertainties is found down to $x \sim 5 \times 10^{-6}$, as illustrated in figure 6. A related analysis of the constraints of the LHCb 7 TeV charm production data using the NNPDF reweighting method has been presented in [43]. Further developments of the underlying theory and the generalization of the NNLO calculations for top quark-pair production [33, 136] for the case of charm and beauty production is highly desirable to exploit the full constraining power of these data.

3.9. Central exclusive production of heavy-flavours

As mentioned above, the production of charm and bottom quarks at the LHC provides useful constraints on the low-x gluon. Related processes, which can potentially give similar information on the gluon, are the central exclusive production of J/ψ and related mesons, such as the Υ or $\psi(2S)$. These processes have the advantage of providing very clean experimentally final states. Calculations for these processes have been presented in [139, 140] within the diffractive photo-production formalism and have been compared to the LHCb data [141, 142]. While this theoretical approach is distinct from the standard collinear factorization picture, the predictions may be extended so as to be related to the collinear gluon PDF, allowing such data to be incorporated consistently into PDF fits. The full NLO contribution has been recently obtained and the scale dependence can be reduced by using an appropriate choice of the factorization scale [143]. Exclusive J/ψ photo production in ultra peripheral p–$\mathrm{P_b}$ collisions has also been measured by ALICE at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV [144].

3.10. Ratios of cross-sections for different center-of-mass energies

The availability of LHC data at different center of mass energies ($\sqrt{s}$), allows us to construct novel types of observables, namely the ratios and double ratios of cross-sections measured at different $\sqrt{s}$ [44]. The main advantage of this approach is that several experimental and theoretical systematics cancel to some approximation. Predictions reach higher accuracy because they are less sensitive to higher orders, and measurements become more precise due

to the reduction of systematic uncertainties such as jet energy scale and luminosities. On the other hand, the interest in these ratios relies on the fact that the PDF dependence does not factorize out, because data at different $\sqrt{s}$ will probe different range in x and Q^2, as well as slightly different flavor content, and therefore might be used to constrain PDFs. Their constraints can be complementary to those of the absolute cross-sections in many cases: in top quark production, for example, the dependence on the exact value of m_t will completely cancel in the ratio.

From the experimental point of view, the crucial point is to quantify the degree of correlation of systematic uncertainties between measurements at different values of $\sqrt{s}$. Up to now, this idea has been implemented in two cases: the ratio of inclusive jet cross-sections between 7 TeV and 2.76 TeV from ATLAS [24], and the ratio of DY cross-sections between 7 and 8 TeV at CMS [45]. While these two measurements are important as proof-of-concept, their impact on PDF fits will be moderate due to the limited statistics of the 2.76 TeV sample (for the ATLAS analysis) and the small lever arm between 7 and 8 TeV (in the CMS measurement). However, during Run II, several ratios between 13 and 8 TeV measurements should be performed, and could become a valuable source of PDF constraints in the global analysis.

4. Constraining PDFs with LHC data at Run I

In this section we present an overview of LHC Run I measurements with potential sensitivity to PDFs, along the lines discussed in section 3. We report here the specific analyses that have been performed and, whenever applicable, the studies used to document their PDF constraining power. This section is split by experiment: we begin with the summary of the results from ATLAS, then we move to CMS and finally to LHCb. A summary table, including only measurements that are published or submitted to a journal, is provided for each experiment. Measurements which are still in preliminary form, or have been superseded, are not included in this report.

4.1. Constraints from ATLAS

The measurement from the ATLAS collaboration are summarized in table 2.

The available ATLAS jet production measurements, relevant for PDF studies, are the inclusive and dijet differential cross sections from the 2010 dataset [157], the ratio of 2.76 to 7 TeV inclusive jet cross sections [24] and, more recently, the inclusive [158], dijet [159] and trijet [160] differential cross sections from the 2011 dataset. The jets are reconstructed using the anti-k_T clustering algorithm [174], and the cross section measurements are available separately for two radius parameters: $R = 0.4$ and $R = 0.6$. The different radii have different sensitivity to final state radiation and underlying event effects. Non-perturbative corrections are supplied in the relevant publications, as well as electroweak corrections for the 2011 inclusive and dijet measurements. The cross sections for either $R = 0.4$ or $R = 0.6$ jets can be included in a PDF fit, though not both at the same time since the measurements are correlated.

A PDF fit to demonstrate the sensitivity to the large-x gluon has been performed, adding the ratio of 2.76 TeV over 7 TeV inclusive jet data to the HERA I data, and comparing to a baseline fit determined using only the HERA data [24]. While the measurement has limited statistics, it provides a powerful proof-of-principle of the enhanced PDF sensitivity of such ratio measurements, due to the partial cancellation of the dominant theoretical and experimental systematic uncertainties, as discussed in the previous section and in [44]. The data

Table 2. Overview of published PDF-sensitive measurements from the LHC Run I from the ATLAS experiment, where we provide the center-of-mass energy, year of data, and the integrated luminosity, its motivation in terms of PDF sensitivity, the publication reference and the references where these measurements have been used to quantify PDF constraints.

ATLAS				
Measurement	$\sqrt{s}$, year of data, $\mathcal{L}_{\text{int}}$	Motivation	Reference	PDF fits
W, Z rapidity	7 TeV, 2010, 36 pb^{-1}	Section 3.3	[123]	[16, 21, 22, 27, 91]
High mass Drell–Yan	7 TeV, 2011, 4.9 fb^{-1}	Section 3.4	[37]	[21, 22, 130]
Low mass Drell–Yan	7 TeV, 2011+2010, 1.6 fb^{-1}+35 pb^{-1}	Section 3.4	[145]	—
ZA_{FB}	7 TeV, 2011, 4.8 fb^{-1}	Section 3.4	[14]	—
W+charm production	7 TeV, 2011, 4.6 fb^{-1}	Section 3.6	[30]	[30]
W + beauty production	7 TeV, 2010, 35 pb^{-1}	Section 3.6	[146]	—
W + beauty production	7 TeV, 2011, 4.6 fb^{-1}	Section 3.6	[147]	—
Z + beauty production	7 TeV, 2010, 36 pb^{-1}	Section 3.6	[148]	—
Z + beauty production	7 TeV, 2011, 4.6 fb^{-1}	Section 3.6	[149]	—
Zp_T	7 TeV, 2010, 40 pb^{-1}	Section 3.5	[150]	—
Zp_T	7 TeV, 2011, 4.7 fb^{-1}	Section 3.5	[131]	—
Wp_T	7 TeV, 2010, 31 pb^{-1}	Section 3.5	[151]	[22]
Z + jets	7 TeV, 2010, 36 pb^{-1}	Section 3.5	[152]	—
Z + jets	7 TeV, 2011, 4.6 fb^{-1}	Section 3.5	[153]	—
W + jets	7 TeV, 2010, 36 pb^{-1}	Section 3.5	[154]	—
W + jets	7 TeV, 2011, 4.6 fb^{-1}	Section 3.5	[155]	—
R_{jets}(W +jets/Z +jets)	7 TeV, 2011, 4.6 fb^{-1}	Section 3.5	[156]	—
Inclusive jets	7 TeV, 2010, 37 pb^{-1}	Section 3.1	[157]	[21, 22, 91]
Inclusive jets	7 TeV, 2011, 4.5 fb^{-1}	Section 3.1	[158]	—
Inclusive jets (+7 TeV ratio)	2.76 TeV, 2010, 0.2 pb^{-1}	Section 3.1, 3.10	[24]	[21, 22, 24]
Dijets	7 TeV, 2010, 37 pb^{-1}	Section 3.1	[157]	—
Dijets	7 TeV, 2011, 4.6 fb^{-1}	Section 3.1	[159]	—
Trijets	7 TeV, 2011, 4.5 fb^{-1}	Section 3.1	[160]	—
γ inclusive production	7 TeV, 2010, 35 pb^{-1}	Section 3.2	[161]	—
γ inclusive production	7 TeV, 2011, 4.6 fb^{-1}	Section 3.2	[162]	[122]

Table 2. (Continued.)

		ATLAS		
γ + jets	7 TeV, 2010, 37 pb^{-1}	Section 3.2	[163]	—
$t\bar{t}$ incl (single lepton, dilepton)	7 TeV, 2010, 2.9 pb^{-1}	Section 3.7	[164]	[21]
$t\bar{t}$ incl (dilepton)	7 TeV, 2010, 35 pb^{-1}	Section 3.7	[165]	[21]
$t\bar{t}$ incl (single lepton)	7 TeV, 2010, 35 pb^{-1}	Section 3.7	[166]	[21]
$t\bar{t}$ incl (dilepton)	7 TeV, 2011, 0.70 fb^{-1}	Section 3.7	[167]	[21, 22]
$t\bar{t}$ incl (e/μ + τ)	7 TeV, 2011, 2.05 fb^{-1}	Section 3.7	[168]	[21]
$t\bar{t}$ incl (tau+jets)	7 TeV, 2011, 1.67 fb^{-1}	Section 3.7	[169]	[21]
$t\bar{t}$ incl (e μ b-tag jets)	7 + 8 TeV, 2012, 24.9 fb^{-1}	Section 3.7	[170]	[22]
$t\bar{t}$ differential	7 TeV, 2011, 2.05 fb^{-1}	Section 3.7	[171]	—
$t\bar{t}$ differential	7 TeV, 2011, 4.6 fb^{-1}	Section 3.7	[172]	—
WW, $Z \to \tau\tau$, $t\bar{t}$ xsec	7 TeV, 2011, 4.6 fb^{-1}	Section 3.3	[173]	—

were shown to be sensitive to the large-x gluon, both reducing the uncertainty and favoring a larger gluon at high x compared to the fit including only HERA I data.

The jet measurements performed using 2011 data have substantially higher precision compared to previous measurements, and span a wider kinematic range, extending up to 2 TeV in inclusive jet p_T, and up to 5 TeV in dijet or trijet invariant mass. They are particularly interesting since, for the first time, the full statistical and systematic correlations between various measurements are provided, allowing their simultaneous inclusion in a PDF fit, thus enhancing the constraining information. Jet measurements using 8 TeV data are also ongoing.

Measurements of top quark pair production provide complementary information on the gluon PDF at high x, and at the same time are sensitive to the strong coupling and the top mass. The ATLAS collaboration has released many measurements of the top quark pair production cross section [164–170] with increasing precision and using a variety of final states. Several of these measurements have been used in recent global fits [21, 22]. Differential cross sections, using the 2011 [171, 172] dataset, have also been measured. In the most recent of these measurements, using the full 2011 dataset, the normalized differential cross sections are measured as a function of the top quark transverse momentum, and of the top quark pair mass, transverse momentum and rapidity. Normalizing the cross sections reduces the dependence on higher order QCD corrections, though it also slightly degrades the sensitivity to the gluon PDF overall normalization. When compared to a range of PDFs, the measured data distributions tend to be softer than the predictions, indicating either their power to constrain the gluon, and/or the importance of higher order QCD and/or electroweak corrections. Further measurements at 8 TeV are also ongoing. These, as well as future precision measurements at higher energy, could place more significant constraints on the gluon.

Additional constraints on the gluon PDF at medium and large-x could be provided by isolated prompt photon data. Compared to jets, prompt photons feature a cleaner experimental environment, though current measurements are of limited precision. ATLAS has measured isolated prompt photon production cross sections in the 2010 [161] and 2011 [162] datasets. The latter provide differential cross sections as a function of photon transverse energy and pseudo-rapidity, in central and forward pseudo-rapidity regions. The ATLAS collaboration has studied the sensitivity of these data to PDFs [121], providing a quantitative χ^2 comparison of the agreement with NLO QCD predictions for a range of PDFs. The results show some tension between data and theory for several current PDFs, indicating potential to constrain the shape and uncertainty of the gluon, though theoretical scale uncertainties are also large. Measurements of prompt photons in association with jets [163] are also available.

Inclusive electroweak boson production data provide constraints on quarks and quark flavor separation. ATLAS has published W and Z rapidity distributions using the 2010 dataset, including full experimental correlations [122]. This has become a standard dataset, now widely used in the global PDF fits. In a separate ATLAS analysis, these data were used in a full PDF fit using the `HERAfitter` framework, in order to quantify the sensitivity to the strange quark content of the proton [122]. The result favors a non-suppressed strangeness. The data also provide constraints on the valence quarks. The ATLAS inclusive W and Z analysis of 2011 data is also close to publication. The larger dataset will allow for more differential measurements, such as the double differential Z/γ^* cross sections binned in dilepton mass and rapidity, and $W^{\pm}$ cross sections in bins of lepton pseudo-rapidity, as well as double differentially in lepton pseudo-rapidity and lepton transverse momentum. Importantly, these data have the potential to provide constraints on valence quark PDFs, and will shed further light on the strangeness content of the proton.

ATLAS has also published low [145] and high [37] mass DY measurements, providing additional and complementary information to the measurements around the Z mass peak. The ATLAS low mass DY cross sections are measured as a function of dilepton invariant mass with coverage $12 < M_{ll} < 26$ GeV for the 2010 dataset, and $26 < M_{ll} < 66$ GeV on a subset of the 2011 dataset. High-mass DY measurements can provide important constraints on large x quarks and anti-quarks. The ATLAS analysis, based on the 2011 dataset, reaches dilepton invariant masses up to 1.5 TeV and has been included in recent PDF fits [21, 22]. In addition, in the presence of QED corrections, this measurement can also be used to constrain the photon content of the proton, and has been included in the NNPDF2.3QED fit [129]. A higher precision measurement at 8 TeV is in preparation. ATLAS has also studied the forward-backward asymmetry in DY events [14], using 2011 data, which may provide sensitivity to PDFs. In the same paper the effective weak mixing angle was extracted.

ATLAS has explored the production of vector bosons in association with heavy flavours. The production of W in association with charm quarks has been measured differentially in charged lepton pseudo-rapidity, using the 2011 dataset [30]. These data give a direct handle on the strange content of the proton, and the ratio of $W^+ + \bar{c}$ to $W^- + c$ is also sensitive to the strange asymmetry, $s - \bar{s}$. In the same analysis, the data are analysed and found to be consistent with a range of PDFs, and indicate a preference for PDFs with an $SU(3)$ symmetric light quark sea, consistent with the result found with the W and Z rapidity distributions. Work is underway to include such data fully and consistently in a PDF fit for the first time. Measurements of vector boson in association with beauty are also available [146–149]. These are especially useful to test pQCD calculational schemes [175].

Vector boson production in association with jets provides further sensitivity to the gluon and sea quark PDFs. ATLAS has released a number of measurements of $W+$ jets and $Z+$ jets using 2010 and 2011 data [152–155]. In addition, ATLAS has measured the $W+$ jets to $Z+$ jets ratio [156], which is complementary to the individual measurements, and is especially interesting due to the large cancellation of experimental systematic uncertainties and non-perturbative QCD effects. Vector boson transverse momentum distributions are also sensitive to the gluon and quark PDFs over a wide range of x. ATLAS has measured the W p_T [151] distribution with 2010 data, and the Zp_T distribution with both 2010 [150] and 2011 [130] data.

4.2. Constraints from CMS

The results from the CMS collaboration sensitive to PDFs are summarized in table 3.

High-precision measurements of the cross-sections of multi-jet production in proton–proton collisions have been performed by the CMS collaboration and the systematic correlations have been investigated. Also, the potential of several jet measurements to constrain the PDFs and determine the strong coupling has been demonstrated.

Jets are reconstructed with the same anti-k_T clustering algorithm used by ATLAS. A different value of radius parameter, $R = 0.7$, is chosen for jet analyses performed with only jets in the final state. This is motivated by the fact that a smaller cone is more sensitive to the final state radiation effects, which are not well described by the NLO predictions in pQCD. However, in the case of the associated production of jets with vector bosons, the value of the jet radius $R = 0.5$ is preferred.

The measurement of inclusive jet production cross-sections in pp collisions at $\sqrt{s} = 7$ TeV based on the data collected in 2011, has been published in [25] as a function of jet kinematics. Furthermore, the correlations of the systematic uncertainties have been reanalyzed and the recommendations for usage of the measurement in the PDF fits published [48].

Table 3. Same as table 2, for the CMS experiment. In the last column, we also indicate which of these measurements have been used as input for either a determination of PDFs or of the strong coupling α_s.

CMS				
Measurement	$\sqrt{s}$, $\mathcal{L}_{\rm int}$	Motivation	Reference	Used in PDF or α_S fits
High and low mass Drell–Yan	7 TeV, 5 fb^{-1}	Section 3.4	[36]	[21, 118]
High and low mass Drell–Yan	8 TeV, 20 fb^{-1}	Section 3.4	[45]	—
Drell–Yan AFB	7 TeV, 5 fb^{-1}	Section 3.4	[176]	—
W asymmetry	7 TeV, 36 pb^{-1}	Section 3.3	[177]	—
We asymmetry	7 TeV, 880 pb^{-1}	Section 3.3	[178]	—
$W\mu$ asymmetry	7 TeV, 4.7 fb^{-1}	Section 3.3	[26]	[26, 118]
W, Z production and rapidity	7 TeV, 3 pb^{-1}	Section 3.3	[179]	—
W, Z inclusive production	7 TeV, 36 pb^{-1}	Section 3.3	[180]	—
W, Z inclusive production	8 TeV, 19 pb^{-1}	Section 3.3	[181]	—
Zp_T and rapidity	7 TeV, 36 pb^{-1}	Section 3.5, 3.3	[182]	—
Zp_T and rapidity	8 TeV, 19.7 fb^{-1}	Section 3.5, 3.3	[132]	—
Inclusive jets	7 TeV, 5 fb^{-1}	Section 3.1	[25, 183]	[21, 48, 91]
Dijets	7 TeV, 5 fb^{-1}	Section 3.1	[25]	—
Three-jets	7 TeV, 5 fb^{-1}	Section 3.1	[184]	[184]
Three-jets/Di-jets ratio	7 TeV, 5 fb^{-1}	Section 3.1	[49]	[49]
W+charm	7 TeV, 5 fb^{-1}	Section 3.6	[29]	[26, 31, 91]
Z+beauty	7 TeV, 5 fb^{-1}	Section 3.6	[185]	—
γ inclusive production	7 TeV, 36 pb^{-1}	Section 3.2	[186]	[28]
γ+jets	7 TeV, 2.1 fb^{-1}	Section 3.2	[187]	—
$t\bar{t}$ inclusive	7 TeV, 2.3 fb^{-1}	Section 3.7	[188]	[32, 33, 138]
$t\bar{t}$ differential	7 TeV, 5.0 fb^{-1}	Section 3.7	[189]	[33]
$t\bar{t}$ inclusive	8 TeV, 1.14 fb^{-1}	Section 3.7	[190]	[32]
$t\bar{t}$ inclusive	8 TeV, 2.8 fb^{-1}	Section 3.7	[191]	[32]
$t\bar{t}$ inclusive	8 TeV, 2.4 fb^{-1}	Section 3.7	[192]	[33]
$t\bar{t}$ differential	8 TeV, 19.7 fb^{-1}	Section 3.7	[193]	—

Another analysis [183], designed to test the performance and result of different jet radii, has measured the inclusive jets cross section ratio using the same data with two different radii parameters: 0.5 and 0.7. In this latter paper, an inclusive jet cross section with $R = 0.5$ is also presented, as well as the cross section with $R = 0.7$ extrapolated towards lower p_T.

A comprehensive QCD analysis [48] of the inclusive jet cross-section measurement at 7 TeV has been performed by the CMS collaboration to demonstrate the impact of these data on the PDFs and to determine the strong coupling constant. The impact of the inclusive jet measurement on the PDFs of the proton is investigated in detail using the `HERAFitter` tool [50], using both the Hessian [194] and the Monte Carlo methods [88]. When the CMS inclusive jet data are used together with the HERA-I DIS measurements, the uncertainty in the gluon distribution is reduced in particular at large x, and a significant reduction of the parametric uncertainty is observed. At the same time, a modest reduction of the uncertainties on u and d valence quark distributions is observed, consistent with the dominance of qq scattering of jet production at high p_T. The inclusion of the CMS inclusive jet data also allows for a combined fit of $\alpha_S(m_Z)$ and of the PDFs, which is not possible with the HERA data alone. As summarized in table 3, these inclusive jets results are already used by several PDF

collaborations. Further inclusive jets measurements at CMS are still ongoing, and are expected to extend and complete the Run I legacy picture.

Two measurements of the three-jet cross section have been performed and optimized for the extraction of α_S running. The first one [49] has used the ratio between the three jets cross section and the dijet cross section, that is proportional to α_S at leading order. This observable has a reduced dependence on the proton PDF and is used to partially decouple the measurement of α_S from the gluon density. The second analysis [184] makes usage of the three jet mass spectrum, which is proportional to α_S^3 at leading order. In principle this observable is more sensitive than the previous one, but is also more dependent on the choice of the proton PDF and suffers from larger systematic uncertainties. The three measurements together provide for the first time a stringent test of the strong coupling running in the region between 100 GeV and 2 TeV. In particular, these are the first direct measurements of the strong coupling constant at the TeV scale, that can be used to provide constraints on BSM scenarios [195].

The dijet cross-sections [25] have been measured using the CMS data collected in 2011 at $\sqrt{s} = 7$ TeV. This measurements exhibit a significant statistical correlation with the inclusive jets case. Since no statistical correlation matrix has been provided between the various CMS jet measurements, it is not possible to use them at the same time in a PDF analysis.

A significant effort within the CMS collaboration has been devoted to the precise measurements of the inclusive vector boson production. Three sets of measurements can be identified: the neutral and charged DY production with a particular attention dedicated to the Z peak; the charged lepton radial asymmetry in the W production (hereafter referred as W asymmetry); the high p_T bosons production. While the first two measurements are expected to be mainly sensitive to quark density, the third one should provide additional constrain on the gluon density.

The inclusive measurements in electron and muon channels of the on-peak neutral and charged DY cross section have been performed at 7 and 8 TeV using the first low luminosity data in order to reduce the contamination from the pile-up [179–181]. Subsequently a precise measurement of the double-differential cross sections as a function of lepton-pair mass and rapidity has been produced, normalized to the peak cross section [36, 45]. A full correlation matrix between bins of the normalized measurement as well as the peak cross section has been provided. Moreover the 8 TeV analysis has been designed to simplify the measurement of the cross section ratio between 8 and 7 TeV. This extremely precise result, with typical uncertainties at a percent level in the bulk of the cross section data, is ready to be used by the PDF extraction groups and its sensitivity to the parton densities still needs to be assessed.

The tensor properties of the DY events have been studied using the forward–backward asymmetry in DY events [176] and subsequently used to extract the effective Weinberg angle [196]. The former measurement may also provide a certain sensitivity to the PDFs, which has not been studied yet.

The lepton-charge asymmetry measurements in W-boson production has been performed separately with muons and electrons that are sensitive to different experimental systematic effects. The most precise measurement available today [26] remains the muon charge asymmetry measurement at CMS, performed with the full 7 TeV data set, while the electron charge asymmetry is limited by the available p_T single lepton trigger to 880 pb^{-1} data sample [178]. The sensitivity of the muon charge asymmetry to the valence quark density has been studied in a QCD analysis at NLO in [26]. A significant reduction of the uncertainty on the d-valence and u -valence distributions is observed with respect to a PDF fit in which only HERA-I inclusive DIS data are used. The lepton-charge asymmetry is now a standard

component of the PDF extraction by many global fit groups. Even more precise measurements of muon charge asymmetry at $\sqrt{s} = 8$ TeV by the CMS collaboration is ongoing.

The on-peak Z-boson production cross section has been measured double differentially in p_T and y in the muon channel with the experimental precision at a percent level [131]. This measurement should provide an additional constraint on the gluon density. The available predictions for the boson production with $p_T(Z) \approx M_Z$ are available only at NLO, while an NNLO prediction would to be necessary to explore the full advantage of the experimental precision.

CMS has also explored the production of vector-boson associated with heavy quarks. The cross sections of the associated production of the W boson together with charm quark has been measured differentially as a function of charged lepton rapidity at 7 TeV [29]. This measurement provides a direct probe of the strange-quark content of the proton sea, as demonstrated by the CMS collaboration in a QCD analysis [26] at NLO, in which HERA-I DIS data, measurements of muon charge asymmetry and the cross sections of $W+$ charm production are used. The strange-quark content, as determined by the analysis, is demonstrated to be consistent with results of the neutrino-scattering experiments. The $Z + b$ production at 7 TeV [185] is measured single-differentially due to a lack of statistics, but a differential measurement is expected to be provided at 8 TeV. Besides to their sensitivity to the PDFs, the measurements of gauge boson production in association with heavy quarks provide useful information about applicability of different heavy-quark schemes in the probed energy regime.

Measurements of the top-pair production at the LHC probe the gluon distribution at high x and at the same time provide constraints on the top-quark mass and the strong coupling constant. For the first time, the value of α_s has been determined [138] at NNLO using the inclusive $t\bar{t}$ production cross section measured by the CMS collaboration [188]. The impact of the inclusive cross section of $t\bar{t}$ production on the gluon distribution is studied [32], where the CMS measurements at $\sqrt{s} = 7$ TeV [188] and $\sqrt{s} = 8$ TeV [190, 191] are included. In the QCD analysis [33], the inclusive and differential cross sections of the top-quark pair production are included and a moderate reduction of the uncertainty on the gluon distribution at high x is demonstrated. In this analysis, the CMS measurements of total [188, 192] and differential [189] top-pair production cross sections are used. More significant improvement of the precision of the gluon distribution is expected with more precise data of the LHC at higher energies. It is important to notice that, for future PDF fits using the top-pair production measurements, the parton-level cross sections provided in the full phase space should be supplemented by the information about correlations of the statistic and systematic uncertainties, also between the data sets of different energies and between inclusive and (normalized) differential cross section measurements.

4.3. Constraints from LHCb

The LHCb experiment, thanks to its unique forward coverage, extends the kinematical range covered by ATLAS and CMS and in particular allows to explore in better detail the small-x region [197]. Therefore, even for the same underlying physical process, LHCb measurements are fully complementary to those of ATLAS and CMS. The corresponding overview of LHCb results are summarized in table 4.

Measurements of W and Z production using muon final states have been performed with 37 pb $^{-1}$ of data collected in 2010 [198]. These measurements, along with those of Z production in the di-electron channel at 7 TeV [199], have been incorporated by the CT, MMHT and NNPDF collaborations into their latest PDF fits [21, 22]. Updated measurements of the W

Table 4. Same as table 2, for the LHCb experiment.

LHCb				
Measurement	$\sqrt{s}$, $\mathcal{L}_{\rm int}$	Motivation	Reference	Used in PDF fits
W, Z muon rap dist	7 TeV, 1.0 fb $^{-1}$	Section 3.3	[201]	[21, 22]
$Z\rightarrow$ ee rap dist	7 TeV, 0.94 fb $^{-1}$	Section 3.3	[199]	[21, 22]
$Z\rightarrow$ ee rap dist	8 TeV, 2.0 fb $^{-1}$	Section 3.3	[203]	—
$W + b/c$	7,8 TeV, 3.0 fb $^{-1}$	Section 3.6	[207]	—
$c\bar{c}$ production	7 TeV, 15 nb $^{-1}$	Section 3.8	[35]	[34, 43]
$b\bar{b}$ production	7 TeV, 0.36 fb $^{-1}$	Section 3.8	[210]	[34]
Exclusive J/ψ production	7 TeV, 1.0 fb $^{-1}$	Section 3.9	[142]	—
Exclusive Υ production	7, 8 TeV, 3.0 fb $^{-1}$	Section 3.9	[211]	—

and Z production cross-sections and their ratio have since been performed with the full 2011 dataset [200, 201]. Among these, [201] contains the most up-to-date and precise measurement of both the W and Z cross-sections. The precision is significantly improved due to the larger data sample, a better understanding of the detector effects, and an improved luminosity determination [202]. As regards the dataset collected in 2012 at a centre-of-mass energy of 8 TeV, Z production has been measured in the di-electron channel [203], with W and Z measurements in the more precise muon channels expected to follow in 2015.

Low-mass DY measurements at LHCb are sensitive to x values as low as 8×10^{-6} at $Q^2 = 25$ GeV^2. A preliminary measurement has been performed by the collaboration at 7 TeV [204] and work is ongoing to finalize the result with the Run-I dataset. Measurements of the associated production of Z bosons with b-quarks and D mesons have been performed in [205, 206] while more recent measurements of W production in association with beauty and charm jets are also presented in [207]. In the latter measurement, the jets are identified using the algorithm outlined in [208] achieving a 65% (25%) efficiency for identifying beauty (charm) jets with a corresponding light-jet mis-tag rate of 0.3%. The first observation of top quark production in the forward region, relevant for constraining the large-x gluon PDF, has been also presented in [209].

Measurements of inclusive beauty and charm quark production have been performed [35, 210] using data collected in 2010 and 2011 at 7 TeV. The measurements exploit LHCb's particle identification and vertexing capabilities to fully reconstruct B and D mesons using hadronic decay modes. As discussed in section 3.8, heavy flavor production can be used to constrain the gluon distribution at low-x and the impact of these results on the PDFs is under study by a number of groups [34].,

As discussed in section 3.9, precise measurements of J/ψ and Υ photo-production can also lead to strong constraints on the low-x gluon distribution [139]. As these processes are characterized by events containing just two muon tracks and a large rapidity gap, LHCb is well suited to their detection due to its relatively low pile-up running conditions and partial backward coverage. Measurements have been made of central exclusive J/ψ production at a centre-of-mass energy of 7 TeV [142] with Υ production in collisions at 7 and 8 TeV [211].

5. Prospects for LHC Run II measurements

In this section we present a general overview of the plans for the ATLAS, CMS and LHCb collaborations concerning PDF-sensitive measurements for the LHC Run II, including a

possible time-line. In addition, we present the results of a profiling analysis which provides an estimate of the impact on PDFs on a number of Run II measurements for W and Z bosons and $t\bar{t}$ production, as well of an estimate of the impact of Run II inclusive jet measurements performed in the frameworks of the CT global analysis. It is worth reminding that, in the near future, complementary measurements relevant for PDF fits will be provided also by other experiments, including HERA and JLab and among others, but their characteristics will not be discussed here.

5.1. Prospects for the LHC experiments

The LHC Run II will produce proton–proton collisions at 13 TeV center-of-mass energy, with integrated luminosity up to 300 fb^{-1}. Compared to Run I, the higher center-of-mass energy implies larger cross sections and extended kinematic reach for many processes of interest like jets, DY, prompt photons, $t\bar{t}$ and vector bosons in association with heavy quarks. For example, an increase by a factor 2 for the electroweak vector bosons, and a factor 4 for the $t\bar{t}$, are expected for the inclusive production cross sections at 13 TeV compared to 8 TeV. Therefore, Run II data will provide complementary PDF sensitivity with respect to the measurements performed during LHC Run I at 7 and 8 TeV. Furthermore, the increased integrated luminosity will lead to a significant reduction of the statistical and systematic uncertainties.

In addition to the total and fiducial cross sections, a special role will be reserved for the measurements of cross section ratios, also involving different center-of-mass energies, which provide more stringent PDF constraints thanks to substantial cancellation of systematic uncertainties, provided a careful treatment of the correlations.

The time-line behind the program for measurements sensitive to PDF during the LHC Run II will be most likely based on the optimal usage of data collected with different running conditions, which are expected to change substantially over time. The very first set of data delivered by the LHC is expected to contain limited pile-up (PU), typically below five interactions per bunch crossing. Depending on the integrated luminosity, which could sum up to 30 pb $^{-1}$ or more, these data could be used for a quick but precise determination of the benchmark cross sections for the inclusive Z and W-boson production. In particular, a limited amount of PU significantly simplifies the extraction of the W cross section, affected by the performance of the missing transverse energy, and the use of low trigger threshold for the electron channel, otherwise affected by large fake rates. If the statistical uncertainty of the sample allows, measurements of differential distributions as functions of the boson p_T and y could be performed on the same dataset.

A subsequent period of data taking with bunch spacing of 50 ns and integrated luminosity up to 1 fb^{-1} is foreseen. These data will be provided with pileup conditions very similar to those occurring at the end of Run I (PU = 20) and will represent a perfect candidate to measure the cross section ratios at different center-of-mass energies. The rest of the data, corresponding to the largest part of the integrated luminosity, will be collected with a bunch spacing of 25 ns, with pileup rapidly increasing from 20 to 40, and probably more. These data will be used for the long term program of Run II, where measurements with increases statistical accuracy and wider phase space coverage will be delivered.

5.1.1. ATLAS and CMS. Run II measurements of DY production as a function of the dilepton invariant mass distribution will potentially improve their experimental precision, providing information down to x $\sim 10^{-4}$ in the low mass region where PDF uncertainties are large. High mass measurements will also benefit significantly from the new conditions,

substantially improving their statistical precision and allowing extended coverage up to 3 TeV, thus providing direct constraints on the poorly known quark and antiquark PDFs at large x and provide constraints on the photon PDF.

Further measurements of vector boson p_T distributions, and of vector bosons in association with jets (including their ratios) are planned, where both the kinematic reach in p_T and the experimental uncertainties can be improved as compared to the corresponding 8 TeV measurements.

Measurements of vector boson in association with heavy flavor production are also of significant interest for Run II. In fact, measurements like $W + c/W + D^*$ performed by ATLAS, are statistically limited and the new data can substantially reduce the statistical uncertainty. A factor of 2 in this respect might be achieved already with 2015 data, potentially allowing a widening of phase space (with the extended coverage at low track p_T, provided by the newly inserted Insertable B-layer, IBL in ATLAS). Given that the ATLAS inclusive W and Z in Run I have suggested an enhanced strangeness content of the proton, supported by the current ATLAS Run I $W + c$ data, it will be important to measure this process at Run II with the highest precision possible, to shed further light. The same emphasis will be put by CMS on detailed study of heavy flavour production. Both the collaborations will able to make higher precision measurements of vector bosons in association with bottom quarks, providing a means to explore different heavy flavor schemes, among other things. While $Z + b$ is known to be a channel more sensitive to the flavour scheme used in PDF evolution than the PDF content itself, $W + c$ was demonstrated to provide an impact on strangeness content of the proton. Finally the Z or $\gamma + c$ and $W + b$ channels are expected to provide for the first time constraints for the intrinsic charm content of the proton [212].

Jet measurements at Run II will allow an extended kinematic reach up to inclusive jet transverse momenta of around 3.5 TeV. Again, ratio measurements at different center-of-mass energy, which will require careful consideration of correlated systematics between Run I and Run II data, can give a better control of dominant systematic uncertainties, as already demonstrated by the previous ATLAS measurement of the ratio of the 2.76 to 7 TeV inclusive jet cross sections (see section 4.1). As for Run I, measurements of dijet, trijet and multi-jet cross sections will also be possible, extending to higher scales and potentially providing further constraints on PDFs and α_S. The CMS plans also include maintaining the effort and expertize to extend the tests of α_S running in the multi-TeV range. In particular the dijet production is expected to be measured triple-differentially in m_{jj}, y_{j1} and y_{j2}. This setup was proposed by the authors of [119] to take the best advantage of the NNLO calculations once these results become public and can be used for the α_S and PDF extraction.

Prompt photon production will also benefit from Run II, providing improved precision on the measurements, which is required for these data to have significant PDF-constraining power. At 13 TeV, the top quark pair production cross section is increased by a factor of 4.7 (3.3) compared to 7 (8) TeV. ATLAS will be able to perform higher precision measurements of total and differential (normalized and absolute) cross sections, as well as ratio measurements at different centre of mass energies, which can help constrain and disentangle the high x gluon PDF, and α_S.

While the potential of $t\bar{t}$ differential production to constrain the gluon PDF was demonstrated with the Run I data, a statistically larger sample is required to make a sizeable impact on PDFs. A number of differential distributions will be measured, in particular allowing to extend the coverage of the gluon PDFs towards larger values of x.

Finally, both the ATLAS and CMS experiments foresee the measurement of cross section ratios different center-of-mass energies, as well as double ratios of different processes (e.g. $t\bar{t}$, Z, W^+, W^-).

5.1.2. LHCb. The increased centre-of-mass energy extends the kinematic range of the experiment to lower x values for W, Z and low-mass DY production. As shown in [39], LHCb measurements of the differential charged lepton asymmetry in $W+$ jet events in Run II have the potential to provide important PDF constraints. The forward acceptance of the LHCb detector, in addition to a significant p_{T} requirement on the jet, extends the sensitivity of the measurement to x values of greater than 0.5, where reductions of up to 35% on the d-quark PDF uncertainty are achievable. Larger cross-sections are also expected in Run II for the production of W and Z bosons in association with heavy quarks, and more precise measurements can be expected. In particular, measurements of W production in association with charm jets or D mesons will provide information on the strange content of the proton complementary to that from ATLAS and CMS.

The greater centre-of-mass energy in Run II will also result in a dramatic increase in the $t\bar{t}$ production cross-section in the LHCb fiducial region. Consequently measurements of $t\bar{t}$ production can be made with a much improved statistical precision. Such a measurement, originally proposed in the context of the forward-backward asymmetry [213], will provide important information on the large-x gluon PDF [214].

In addition to extending coverage to an even lower x-region, measurements of $b\bar{b}$ and $c\bar{c}$ production in Run II will allow a determination of the production ratio of heavy quarks at different centre of mass energies. The relatively large theoretical uncertainties present in the predictions for these processes make the ratios particularly attractive as a partial cancellation is expected. As such, the ratios may provide more stringent constraints on the PDFs than the individual measurements.

The installation of a dedicated forward shower counter system (HERSCHEL) on the LHCb detector ahead of Run II has the potential to improve the precision of measurements of central exclusive production by extending its coverage into the very forward region. Current LHCb measurements of exclusive J/ψ and $\psi(2S)$ production [142] contain large backgrounds arising from inelastic production where the dissociation of one or both protons is not detected. HERSCHEL allows such events to be rejected by identifying forward showers through the interaction of high rapidity particles with the beam pipe. Consequently, a higher purity and precision can be expected also for these Run II measurements.

5.2. Constraining PDFs with Run II data: a profiling analysis

The upcoming Run II data will provide rich information on PDFs. Compared to the Run I data, higher center-of-mass energy extends the probed kinematic range while larger data samples should lead to reduced uncertainties. In the following a possible impact of the LHC data is estimated using the Hessian PDF profiling method which is implemented in the `HERAFitter` program. For this purpose, benchmark measurements, such as inclusive W, Z and $t\bar{t}$ production are considered. An estimate of the data uncertainties is based on the existing Run I measurements which were published by the ATLAS and CMS collaborations. The inclusive measurements are typically dominated by the systematic uncertainties already for the Run I based results, however several components of the systematic uncertainty may be reduced with increased data statistics. Thus a simplified procedure is used to estimate the uncertainties of the Run II measurements. Three possible scenarios are considered: baseline, when the data uncertainties are taken to be similar to those of the Run I measurements; conservative, when the data uncertainties are scaled up by factor of two; and aggressive, when the data uncertainties are reduced by factor of two.

The study is an indication of the LHC Run II data sensitivity however it is not meant to be an exhaustive investigation. For example, other measurements such as off-peak NC DY

production, $W^{\pm}c^{\mp}$ charge asymmetry, and vector boson production in the forward region, which can be measured at the LHCb, are not considered.

5.2.1. PDF profiling and theoretical predictions. The impact of a pseudo-data set on a Hessian PDF set can be quantitatively estimated with a profiling procedure [51, 215] [23]. The profiling can be performed by minimizing a χ^2 function comparing data and theory predictions which includes both the experimental uncertainties and the theoretical uncertainties arising from PDF variations:

$$\chi^2\left(\boldsymbol{\beta}_{exp}, \boldsymbol{\beta}_{th}\right) = \sum_{i=1}^{N_{\rm data}} \frac{\left(\sigma_i^{\rm exp} + \sum_j \Gamma_{ij}^{\rm exp}\beta_{j,\rm exp} - \sigma_i^{\rm th} - \sum_k \Gamma_{ik}^{\rm th}\beta_{k,\rm th}\right)^2}{\Delta_i^2} + \sum_j \beta_{j,\rm exp}^2 + \sum_k \beta_{k,\rm th}^2. \quad (3)$$

The correlated experimental and theoretical uncertainties are included using the nuisance parameter vectors $\boldsymbol{\beta}_{\rm exp}$ and $\boldsymbol{\beta}_{\rm th}$, respectively. Their influence on the data and theory predictions is described by the $\Gamma_{ij}^{\rm exp}$ and $\Gamma_{ik}^{\rm th}$ matrices. The index i runs over all $N_{\rm data}$ data points, whereas the index $j(k)$ corresponds to the experimental (theoretical) uncertainty nuisance parameters. The measurements and the uncorrelated experimental uncertainties are given by $\sigma_i^{\rm exp}$ and Δ_i, respectively, and the theory predictions are $\sigma_i^{\rm th}$. Following [51], the profiling procedure is generalized to account for asymmetric uncertainties:

$$\Gamma_{ik}^{\rm th} \rightarrow \Gamma_{ik}^{\rm th} + \Omega_{ik}^{\rm th}\beta_{k,\rm th}, \quad (4)$$

where $\Gamma_{ik}^{\rm th} = 0.5(\Gamma_{ik}^{\rm th+} - \Gamma_{ik}^{\rm th-})$ and $\Omega_{ik}^{\rm th} = 0.5(\Gamma_{ik}^{\rm th+} + \Gamma_{ik}^{\rm th-})$ are determined from the shifts of predictions corresponding to up ($\Gamma_{ik}^{\rm th+}$) and down ($\Gamma_{ik}^{\rm th-}$) PDF uncertainty eigenvectors.

The values at the minimum of the nuisance parameters $\beta_{k,\rm th}^{\rm min}$ can be interpreted as optimization ('profiling') of PDFs to describe the data. When profiling is performed using pseudo-data, for which the data central values coincide with the prediction, the shifts of the PDF nuisance parameters vanish. However after the profiling the nuisance parameters have reduced uncertainties which directly affects the uncertainty bands of the PDFs.

The predictions for DY production are obtained using the FEWZ program [218]. The predictions of $t\bar{t}$ production are calculated using the TOP++ program [219]. All the calculations are performed at NNLO accuracy. The PDF sets used for the profiling are CT10nnlo [59], MMHT14 [21] and a Hessian version of NNPDF3.0 [22, 54] [24]. When needed, the PDF uncertainties are re-scaled to the 68 confidence level.

The central values of the pseudo-data are taken to be equal to the central values of the predictions. The profiling uses $\Delta\chi^2 = 1$ criterion for the uncertainty estimate, thus the impact of the data on the PDF uncertainties may differ compared to inclusion in a full PDF fit, especially if there is a tension among different data sets. Recall that Hessian global PDF fits use alternative methods that produce PDF uncertainties that are larger than those for the $\Delta\chi^2 = 1$ condition[25].

[23] For Monte Carlo sets instead one should use the Bayesian reweighting method [216, 217].

[24] Using the `mc2hessian` algorithm developed in [54], any Monte Carlo PDF set can be converted into a Hessian representation and thus the profiling method can be applied. Usage of the profiling method on a hessian version of a Monte Carlo PDF set was checked on a MMHT2014_hessian set, that was extracted from the hessian → MC → hessian transformation. The size of the observed constraints was found to be similar to those on the original MMHT2014 PDF set.

[25] In particular, CT10 uses a two-tier method for the computation of the PDF uncertainty that is not equivalent to the $\Delta\chi^2 = 100$ tolerance [220].

5.2.2. Generation of pseudo-data. The pseudo-measurements selected for the study satisfy the following criteria:

1. There are NNLO predictions available. This requirement ensures that the theoretical uncertainties are smaller than the PDF uncertainties and comparable to the ultimate data uncertainties. In the following, other theoretical uncertainties such as scale variations are neglected.
2. The data have $\sim$1% accuracy and can be described by a simple correlation model. This criterion excludes final states with jets, such as inclusive jet and vector boson plus jet production. With recent developments of NNLO calculations, these data may have the power to place strong constraints on the PDFs. However the impact of the data depends strongly on measurement-specific correlation model, investigation of which is beyond this study.
3. The measurement can be expressed in a simplified phase-space region with well-defined particle to parton-level corrections. This excludes observables such as $W+$ charm production.
4. Only data from the central detectors ATLAS and CMS are considered.

The observables are also selected such that the correlations among them are reduced. This leads to a preference for ratio measurements rather than absolute cross-section determinations. Measurements of absolute cross sections with full correlation information may lead to better PDF constraints, however they depend on detector-specific correlation model, which is difficult to follow in this simplified investigation.

Taking into account these requirements, the four pseudo-measurements used in the present study of the PDF sensitivity of the LHC Run II at $\sqrt{s} = 13$ TeV data are the following:

- Ratio of inclusive cross sections of W-boson to Z-boson production, $R_{W/Z}$. The reference measurements for this observable are the ATLAS measurement performed at $\sqrt{s} = 7$ TeV [122] and the CMS measurement at $\sqrt{s} = 8$ TeV [181]. The ratio is considered for the fiducial region defined by the lepton transverse momentum and pseudorapidity cuts, $p_t > 25$ GeV and $|\eta| < 2.5$. The baseline uncertainty is taken to be 1%.
- Ratio of inclusive cross sections of $t\bar{t}$ to Z-boson production, $R_{\mathrm{t\bar{t}}/Z}$. The $t\bar{t}$ pseudo-data are based on the ATLAS 7 and 8 TeV total cross-section measurement in $e\mu$ channel with b-tagged jets [170]. This measurement reached 2% accuracy, excluding the luminosity uncertainty. The luminosity uncertainty cancels for the $t\bar{t}$ to Z cross-section ratio. If the Z cross-section measurement is obtained using both $Z \to \mathrm{e}^+\mathrm{e}^-$ and $Z \to \mu^+\mu^-$ channels, a significant additional cancellation of uncertainties may be also achieved for the reconstruction of leptons. Thus 2% uncertainty on $R_{\mathrm{t\bar{t}}/Z}$ is considered as a baseline. The fiducial definition for the $Z \to \ell\ell$ cross-section measurement is taken to be the same as for $R_{W/Z}$.
- Lepton charge asymmetry for W decays, A_ℓ. The pseudo-data are based on the CMS measurement of the muon charge asymmetry [26]. The data are considered in fiducial region $p_T > 25$ GeV and $|\eta_\ell| < 2.5$. The data are binned in 10 bins with bin width $\Delta\,|\eta_\ell| = 0.25$. The baseline statistical uncertainty is taken to be 0.0005 per bin, which roughly corresponds to integrated luminosity of $10\,\mathrm{fb}^{-1}$ of $\sqrt{s} = 13$ TeV data. The baseline systematic uncertainty varies from 0.0020 to 0.0036 for the data from the most central to the most forward bin. The bin-to-bin correlation model for the systematic

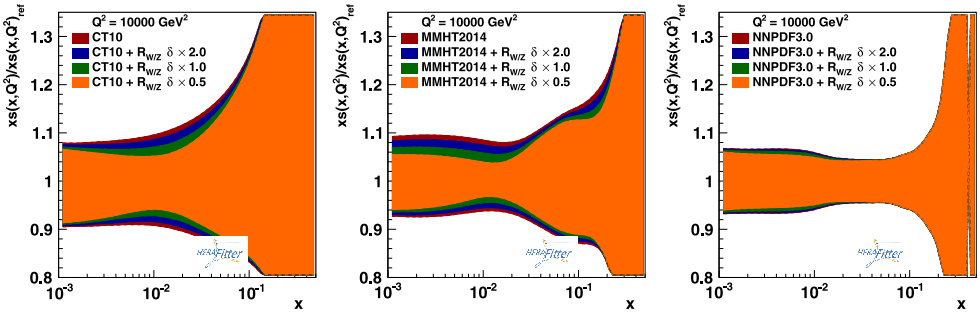

Figure 7. Relative uncertainty of the strange-quark distribution as a function of x for $Q^2 = 10^4\,\mathrm{GeV}^2$ estimated based on CT10nnlo (left), MMHT14 (middle) and NNPDF3.0 (right) PDF sets, respectively. The outer uncertainty band corresponds to the original PDF uncertainty. The embedded bands represent results of the PDF profiling using $R_{W/Z}$ pseudo-data at 13 TeV corresponding to (from outermost to innermost band) conservative, baseline, aggressive model of the data uncertainties.

uncertainties is taken similar to the CMS analysis, as implemented in the `HERAfitter` package, with the correlation coefficient between 0.2 and 0.3.

- Normalized inclusive Z-boson rapidity, y_Z. The pseudo-data are based on the CMS measurement of the NC DY production at 7 TeV [36]. The data are considered in fiducial region $p_T > 25$ GeV and $|\eta_\ell| < 2.5$. The pseudo-data are binned in 12 bins with bin width $\Delta\,|y_Z| = 0.2$. The statistical uncertainty is expected to be negligible compared to the systematics for the Run II dataset. The baseline total uncertainty varies between 0.00155 and 0.00050 for the central and the most forward regions, respectively. The bin-to-bin correlation model for the systematic uncertainties is taken similar to the CMS analysis with strong correlation for the neighboring bins, $\approx$0.7, and some anti-correlation between far-apart bins, up to -0.5.

Basic properties of the pseudo-data samples are listed in table 5. The correlation model was kept unchanged between the baseline, aggressive and conservative scenarios for data uncertainties.

5.2.3. Results. Firstly, the effect of PDF profiling is studied separately for each individual pseudo-data set. Profiling of different PDF sets show qualitatively similar behavior, however the size of the constraints differs depending on how strongly the published PDF set was constrained by the input data used in the original fit. For the main comparisons described below, the CT10nnlo set is used. It is however important to compare results for several sets. The MMHT14 and NNPDF3.0 sets are of particular interest since these two sets include the published Run I data from the LHC experiments.

The PDF uncertainties are reported at momentum transfers squared Q^2 roughly corresponding to the data scale, $Q^2 = 10^4\,\mathrm{GeV}^2$. For profiling of each pseudo-data sample the PDF or a combination of PDFs which are most affected by the measurement are reported. The results are given for the baseline, conservative and aggressive scenarios of the data uncertainties.

The profiling of the $R_{W/Z}$ pseudo-data has the largest impact on the strange-quark distribution which is shown in figure 7. For pseudo-data have maximal impact at $x \sim 0.01$ which agrees with the ATLAS observation [27]. The uncertainty reduction improves significantly as the data accuracy improves; with respect to the CT10 set the aggressive

Table 5. Features of the pseudo-measurements considered for the $\sqrt{s} = 13$ TeV profiling studies

	$R_{W/Z}$	$R_{t\bar{t}/Z}$	A_ℓ	y_Z
Kinematic range			$p_{t,\ell} > 25$ GeV, $\lvert\eta_\ell\rvert < 2.5$	
Number of bins	1	1	10	12
Baseline accuracy per bin	1%	2%	≈1.5%	≈1.5%

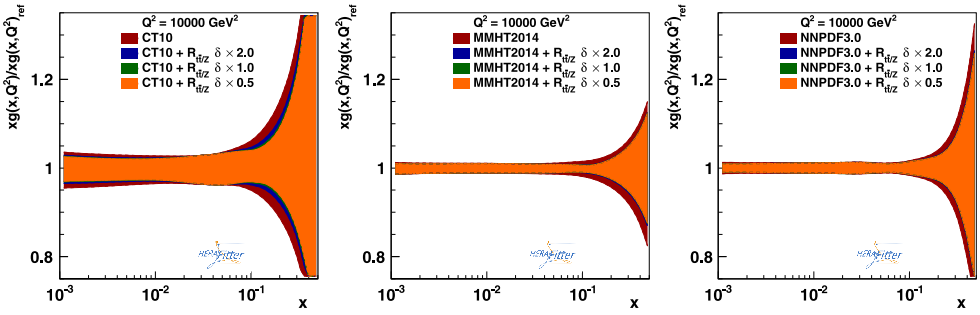

Figure 8. Same as figure 7 this time for the gluon PDF, using the measurement of the $t\bar{t}/Z$ ratio as input to the profiling.

scenario leads to close to factor of 2 reduction. The other PDFs are not affected significantly by the inclusion of $R_{W/Z}$ data apart from moderate reduction in uncertainty for $\bar{d}$ and $\bar{u}$ distributions.

The profiling of the $R_{t\bar{t}/Z}$ pseudo-data affects the gluon distribution the most, see figure 8. The uncertainties are reduced for $x \geqslant 0.1$ and $x < 0.01$ regions. Contrary to the other observables, the difference in pseudo-data accuracy does not affect the gluon density uncertainty significantly. The other PDF which has notable reduction in uncertainty at low x is the total light sea, $x\Sigma = x\bar{u} + x\bar{d} + x\bar{s}$. This reduction can be explained by high degree of correlation between the gluon and sea distributions at low x. Note also that constraints from measurements of the $t\bar{t}$ cross sections depend strongly on the values of the top mass and $\alpha_s(m_Z^2)$.

A study was performed to clarify the dependence of PDF uncertainty reduction as a function of the $R_{t\bar{t}/Z}$ pseudo-data uncertainty. Using the procedure described in [221] the PDFs eigenvectors were re-diagonalized to isolate a linear combination of them which affects the $R_{t\bar{t}/Z}$ observable the most. For a single measurement such as $R_{t\bar{t}/Z}$ this procedure returns a single re-diagonalized eigenvector which affects the measurement while others have no impact. This eigenvector has a significant contribution to the gluon density uncertainty at $x = 0.1$, however it does not saturate the uncertainty band. As a consequence, while the eigenvector is constrained progressively as the pseudo-data accuracy increases, the other irreducible uncertainty component prevents from further improvement in the total gluon density uncertainty.

The lepton-asymmetry measurement has the largest impact on the difference of the u- and d-valence distributions, $u_v - d_v$, which is shown in figure 9. There is a sizable reduction in the uncertainty for $x \sim 0.03$ and $x < 0.003$ kinematic regions which becomes more significant as the pseudo-data accuracy increases.

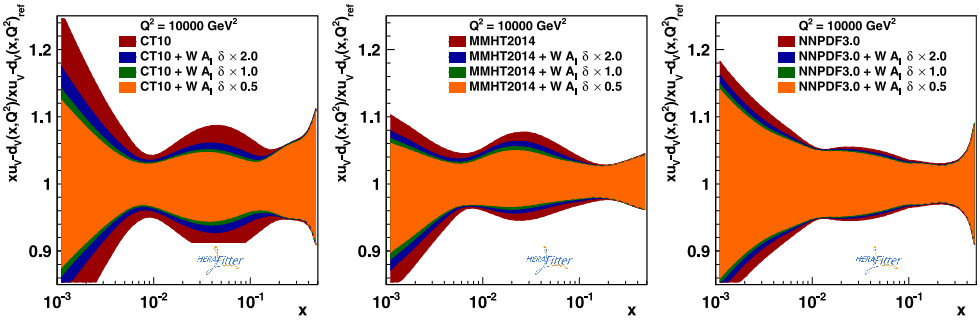

Figure 9. Same as figure 7 this time for the difference between u_V and d_V PDFs, using the measurement of the W lepton asymmetry as input to the profiling.

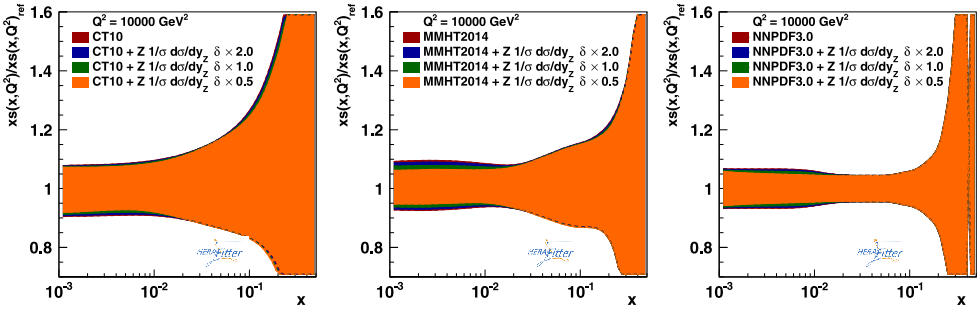

Figure 10. Same as figure 7 this time for the strange PDF, using the measurement of the normalized rapidity distributions of Z bosons as input for Run II.

The data on y_Z also has largest impact on the strange-quark distribution which is shown in figure 10. The effect is complementary to the impact of the W/Z cross-section ratio pseudo-data, compared which the reduction of the uncertainty is more concentrated in the small $x < 0.01$ region. Similarly to $R_{W/Z}$, the data also constrain the $\bar{u}$ and $\bar{d}$ light sea-quark distributions.

It is interesting to notice that the level of uncertainty reduction due to inclusion of the pseudo-data is rather similar for the CT10nnlo and MMHT14 sets while it is significantly smaller for the NNPDF3.0 set. This behavior can be most likely explained by the difference of input data used in the sets and different level of parameterization flexibility.

Finally, all the pseudo-data samples are profiled together in a simultaneous fit. Figure 11 shows result of this profiling for the CT10nnlo sample and for the most affected PDF distributions. The simultaneous fit yields to quantitatively similar reduction of PDF uncertainties compared to the fits to the individual observables. This is not unexpected since with exception of $R_{W/Z}$ and y_Z, the observables are sensitive to different PDF combinations and they are not correlated experimentally.

To summarize, the $\sqrt{s} = 13$ TeV LHC data will make a contribution for reduction of PDF uncertainties. Measurements of the cross-section ratios of the W- to Z-boson and $t\bar{t}$ to Z-boson production, W-boson lepton asymmetry and Z-boson rapidity distribution can be used to constrain strange-quark, gluon and valence-quark distributions. Additional constraints from 13 TeV LHC data will be provided from more differential distributions, provided the statistical and systematic experimental uncertainties can be kept under control. The results of

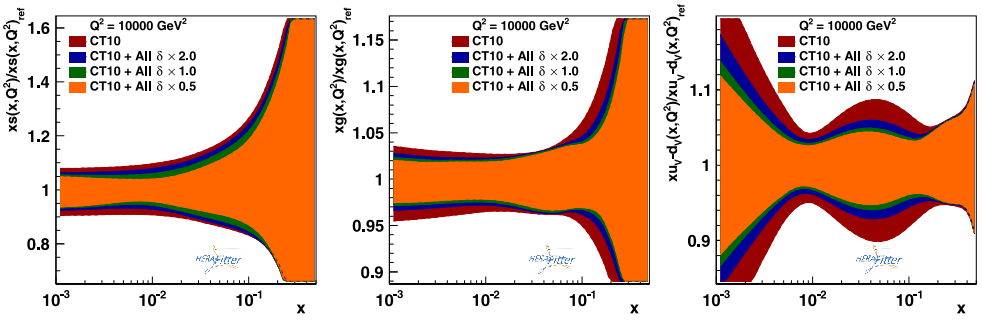

Figure 11. Relative uncertainty of the strange-quark (left), gluon (center) and $u_v - d_v$ (right) distributions as a function of x for $Q^2 = 10^4\,\mathrm{GeV}^2$ estimated based on CT10nnlo PDF set. The outer uncertainty band corresponds to the original PDF uncertainty. The embedded bands represent results of the PDF profiling using the complete set of observables considered in this exercise: $R_{W/Z}$, $R_{t\bar{t}/Z}$, A_ℓ and y_Z pseudo-data at 13 TeV. The various bands correspond to (from outermost to innermost band) conservative, baseline, aggressive model of the data uncertainties.

figure 11 also nicely illustrate the advantages of achieving a reduction of experimental uncertainties in terms of improved PDF constraints.

5.3. Projected impact of the LHC inclusive jet data

Single-inclusive jet production, a key benchmark process at hadron-hadron colliders, proceeds through multiple parton scattering channels. Under LHC conditions, much of the PDF uncertainty of inclusive jet cross sections arises from the gluon PDF; hence they can constrain $g(x, Q)$ in a wide range of x. Potential impact of future LHC jet cross sections has been recently examined in the context of the CTEQ-TEA global analysis. The CTEQ series of PDFs include single-inclusive jet cross sections from Tevatron D0 and CDF collaborations [222, 223], and, starting with CT14, from ATLAS [157] and CMS [25]. Section 2 of [17] shows that the PDF uncertainty of inclusive jet data is correlated with $g(x, Q)$ at $x \gtrsim 0.07$ at CDF and at $x \gtrsim 0.005$ at ATLAS; i.e. the reach in x of jet production is extended at least by an order of magnitude at the LHC.

Let us illustrate how the gluon PDF changes upon including various data sets on inclusive jet production, using the framework of the CT10 NNLO QCD analysis [59] as an example. We start by including all experiments used in CT10 NNLO, except for jet experiments, and assuming the world-average value of the QCD coupling constant, $\alpha_s(m_Z) = 0.118$. The 90% confidence level error PDFs are found by following the Hessian approach, as summarized in [220]. Single-inclusive jet cross sections are evaluated using fast interpolation interfaces [61, 62] to the theoretical calculation at NLO in QCD [224]. We set the factorization and renormalization scales equal to p_T of the jet in each experimental bin, which minimizes both the residual scale dependence at NLO [4, 225] and the NNLO/NLO correction [94] in the partial NNLO calculation in the gg sub-channel [95, 119]. Thus, the unknown NNLO corrections are believed to be inconsequential for the present study.

We include the full ATLAS data sample (7 TeV, 37 pb^{-1}, cone size $R = 0.6$). Similar outcomes are obtained with the ATLAS data set for $R = 0.4$. As an option, we also estimate the possible impact of the NLO scale dependence and missing NNLO contributions on

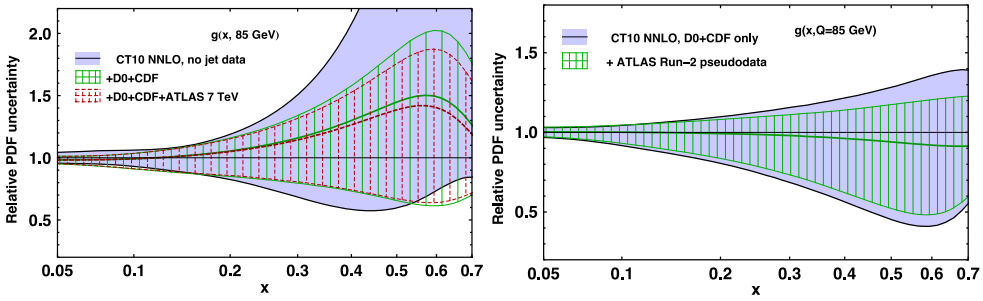

Figure 12. Left plot: Gluon PDF uncertainties at 90% C.L. for the CT10-like fits without any jet data included, compared with the fits with the Run II D0 and CDF jet data, and with the Run I ATLAS jet data included. Right plot: same comparison, now for the fits including the Run 2 D0 and CDF jet data, and in addition with the ATLAS Run II simulated pseudo-data.

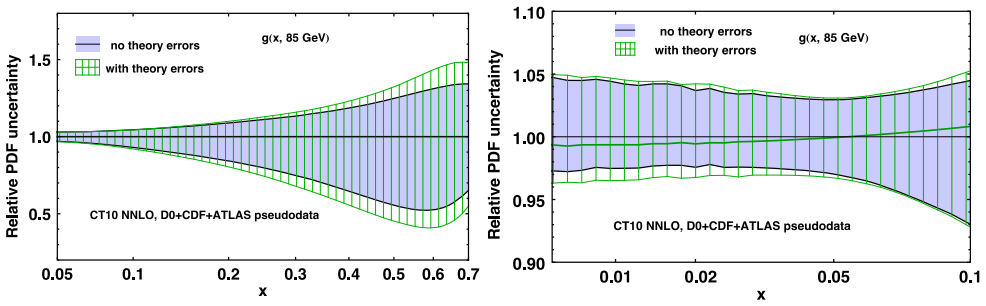

Figure 13. Gluon PDF uncertainties at 90% C.L. for the fits with and without theoretical errors. Left plot: large-x region. Right plot: intermediate-x region.

$g(x, Q)$ using a phenomenological approach that is similar to the ones proposed in [226, 227]. This is done by treating additional theoretical uncertainties as correlated systematic errors and including corresponding columns in the correlation matrix of each experiment, in addition to the usual experimental systematic errors. We notice, in particular, that the estimated theoretical uncertainties do not exceed the experimental uncertainties in all bins on the ATLAS data set.

If we compare the uncertainties on the gluon PDF in CT10-like fits without and with the Tevatron Run II and ATLAS jet data, and not accounting for theoretical uncertainties, the resulting 90% C.L. error bands on $g(x, Q = 85\ \text{GeV})$ in the range 0.05–0.7 are shown in the left sub-panel of figure 12. The jet data modify the central gluon PDF and, as seen in the figure, greatly reduces the PDF uncertainty. The early ATLAS jet data do not improve the constraints on the gluon PDF much, since they have large experimental errors.

To estimate effects of future LHC jet measurements, we introduce two new pseudo-data sets with the same kinematics as the ATLAS measurement [157], and $R = 0.6$ and 0.4. We assume the statistical errors in the measurements to be reduced by a factor of 20, and the jet energy scale errors, which dominate the experimental systematic errors, to be reduced by a factor of 3. The central values of the pseudo-data sets are generated randomly based on the theoretical predictions from one PDF set. The right sub-panel of figure 12 shows the effects that the pseudo-data sets have on the gluon PDF. The PDF uncertainty in this case is reduced

by about 20 percentage points in the large-x region. At moderate $x \sim 0.05$, relevant to central Higgs boson production, the reduction in the uncertainty is less pronounced.

Next, we add four correlated shifts in theoretical predictions that are allowed by NLO theoretical uncertainties according to our method. The impact is illustrated in figure 13, comparing $g(x, Q)$ in the fits with and without the theoretical errors included. With the additional correlated shifts due to theoretical errors, we obtain a slightly harder best-fit gluon in the large-x region. Most importantly, the PDF uncertainty increases by up to 15 percentage points at large x, and by up to 1 percentage point at moderate x (in the Higgs production region). Needless to say, these preliminary estimates of the theoretical uncertainty at NLO (dominated by QCD scale dependence) will likely be reduced once the full NNLO computation is completed.

To summarize, using pseudo-data sets of inclusive jet measurements at the LHC, we have estimated the potential for reduction of the uncertainty in the gluon PDF upon inclusion in the future. Although the ATLAS pseudo-data sets in this exercise correspond to $\sqrt{s} = 7\,\text{TeV}$, similar constraints are expected from the measurements at 8 TeV. Measurements at 13 TeV will probe the gluon PDF at even smaller x; their prospects will still be dependent on improvements in experimental systematic errors, as at the other two energies.

6. On the presentation of LHC data for PDF fits

We conclude this report with a number of practical suggestions concerning the presentation and delivery of LHC measurements to be used in global PDF analysis:

- It is important that measurements are provided with the complete information on correlated systematic uncertainties. While this has not always been the case in pre-LHC experiments, it is now common practice by ATLAS, CMS and LHCb, and thus it should be encouraged to continue.
  The preference is to make available the full breakdown of individual sources of experimental systematic uncertainties, in terms of nuisance parameters. If this is not possible, the full correlation matrix of the experimental data needs to be be provided.
- It is common practice that experimental data is made publicly available through `HepForge`, and it would be beneficial for all parties if this practice is continued. When available, also the corresponding `Rivet` analysis could be made public there.
- Whenever possible, the LHC experiments should try to provide information on cross-correlations between individual datasets, as well as between different data-taking years when available. These cross-correlations can be both of statistical and of systematic origin. This is necessary to consistently include at the same time in a common global analysis measurements like inclusive jets and dijets from the same dataset, which have both statistical and systematic correlations.
- In addition, it would be beneficial for PDF analysis if the LHC experiments could agree on which systematic uncertainties are correlated between ATLAS, CMS and LHCb, even partially, since this maximizes the constraints that can be extracted from the LHC measurements.
- The LHC experiments should give clear indications of which measurements are the most updated ones and should then be used on PDF analysis, and which ones are superseded and should thus not be used. This information can be make public for example in the analysis group webpages for each experiment.

- It would be advantageous if the LHC collaborations could agree on common settings for their PDF-sensitive measurements, for instance the jet radius R for jet measurements, since this streamlines the comparisons between different measurements and their impact on PDFs.
- With a similar motivation, it would be beneficial that the LHC experiments agree on a common set of observables and distributions to deliver in PDF-sensitive measurements. For instance, for immediate use in PDF fits, parton-level corrected data is needed, or, alternatively, experiments could supplement hadron-level data with the corresponding correction factors.

  This said, future fits should move towards using hadron level data, which is closer to the actual measurements and that cancels the hadron-to-parton modelling ambiguities. This is especially important for complicated processes like $W+$ charm, where assessing the theoretical uncertainties in the hadron-to-parton correction factors is challenging. Note that the technology to include hadron-level theory calculations in PDF fits is already available [92].
- Fully differential measurements in the fiducial regions are typically preferred over integrated cross-sections, due to the theoretical uncertainty induced by the extrapolation from the fiducial to the full phase space. This is now a common practice for both experiments, and should be encouraged to continue in future measurements.

  If presenting fiducial measurements is not possible, it would be important to provide the conversion factors used to extrapolate to the full phase space. In this case, as well as the conversion factors, additional useful information would be provided by the systematic uncertainties associated with the PDF dependence in this conversion, if it they are at all significant.
- When theory calculations are generated explicitly for PDF comparisons in a LHC analysis, using codes such as `APPLgrid` or `FastNLO`, it would ease their inclusion in global PDF fits if the corresponding fast grids could also be released together with the experimental data. This would also apply to other ingredients of the theoretical calculations, such as hadron-to-parton corrections or NNLO K-factors. This has already been done for a number of important analysis, and it will be important to ensure that releasing theory grids becomes a standard practice in the future.
- Ideally, it would be useful if experiments could agree on the usage of common theory tools for constructing fast grids, such that theory calculations corresponding to measurements of one specific type are delivered as fast grids with a common format. In the cases where different tools are used, for example for jet production measurements, it would be important to ensure that all theory settings, such as factorization and renormalization scales, or heavy-quark schemes, are spelled out in detail in the corresponding publications, allowing for *a posteriori* comparison between the various theory codes.
- Whenever possible, it would be useful to agree on a common treatment of theory corrections to be applied to the data: for example, all gauge-boson production data being presented with the same treatment of final-state QED radiation and electroweak corrections.

These suggestions should be helpful to streamline the process of adding new LHC measurements into PDF fits, maximizing the constraining power of the data and minimizing the theoretical and methodological uncertainties associated to PDF-sensitive measurements.

7. Summary and outlook

In this report we have summarized the constraints on PDFs that have been obtained from LHC measurements during Run I. The impressive wealth and quality of PDF-sensitive measurements at the LHC Run II have been summarized in tables 2 for the ATLAS, table 3 for the CMS and table 4 for the LHCb collaborations. Many of these measurements are already included in recent global PDF fits to date, where they provide constraints in a wide range of x and PDF flavours. It is especially remarkable that the high energy of the LHC also allows us to introduce new types of the processes in the PDF analyses for the first time, from $t\bar{t}$ production and $W + c$ production to high-mass DY data.

We have also reviewed the prospects for PDF studies with the recently started 13 TeV Run II data-taking. A number of improvements are foreseen, thanks to the increase on center-of-mass energy and in the integrated luminosity, especially for those measurements that at Run I were limited by statistical or statistics-related systematic uncertainties. We have also performed a quantitative estimate of the impact on PDFs based on Run II pseudo-data for W, Z and $t\bar{t}$ production using the profiling method. Our results show that PDF uncertainties in a number of PDF flavors can be reduced with 13 TeV measurements of these processes, and emphasize the importance of reducing the systematic uncertainties in PDF-sensitive measurements.

This report addresses both the experimental collaborations, in order to identify their priorities for PDF-sensitive measurements at Run II, as well as the PDF fitting groups to have a clear perspective of which measurements are already available and which ones will also become available in the near future. Exploiting the full potential of Run II data for PDF constrains is essential for the LHC physics program for the next years, and in turn feeds into many other analysis like improved determination of the Higgs couplings or of precision SM measurements like the W mass.

In this report we have concentrated only on the experimental constraints provided by the LHC data. Equally important for PDF determinations is to use state-of-the-art theoretical calculations, especially exploiting the recent developments in NNLO calculations for LHC processes. The use of higher-order perturbative calculations is essential for reducing sources of theoretical uncertainties in PDF fits, which are presently not even estimated. In this respect, a careful benchmarking of NLO and NNLO codes for processes relevant for PDF studies would be certainly interesting. A collaborative effort is of particular importance for the progress of the benchmarking exercise, which was performed in the past [4, 225, 228, 229] and is necessary to understand and reduce the differences between the results of different PDF groups.

Acknowledgments

We are grateful to all the participants of the various PDF4LHC meetings that stimulated many enthusiastic discussions on the measurements and methodological developments reviewed in this report. We are grateful to José Ignacio Latorre and all the Pedro Pascual Benasque Center for Science team for their invaluable support during the workshop *Parton Distributions for the LHC* in February 2015, where many of the authors of this paper were present.

The work of A A was supported by the DOE contract No. DE-AC05-06OR23177, under which Jefferson Science Associates, LLC operates Jefferson Lab, and by the DOE contract DE-SC008791. The work of S Farry is supported by a research fellowship from the Royal Commission for the Exhibition of 1851. The work of S Forte is supported in part by an Italian

PRIN2010 grant and by a European Investment Bank EIBURS grant. The work of LH-L is supported by the Science and Technology Facilities Council (STFC) for via the grant award ST/L000377/1. The research of JG in the High Energy Physics Division at Argonne is supported by the US Department of Energy, High Energy Physics, Office of Science, under Contract No. DE-AC02-06CH11357. The work of PN is supported by the US Department of Energy under grants DE-SC0003870 and DE-SC0013681. The work of J R is supported by an STFC Rutherford Fellowship ST/K005227/1 and by an European Research Council Starting Grant 'PDF4BSM'. The work of R T is supported partly by the London Centre for Terauniverse Studies (LCTS), using funding from the European Research Council via the Advanced Investigator Grant 267352. R T would also like to thank the Science and Technology Facilities Council (STFC) for support via grant awards ST/J000515/1 and ST/L000377/1, and the IPPP, Durham, for the award of a Research Associateship.

References

[1] Forte S 2010 Parton distributions at the dawn of the LHC *Acta Phys. Pol.* B **41** 2859–920
[2] Forte S and Watt G 2013 Progress in the determination of the partonic structure of the proton *Annu. Rev. Nucl. Part. Sci.* **63** 291–328
[3] Perez E and Rizvi E 2013 The Quark and Gluon structure of the proton *Rep. Prog. Phys.* **76** 046201
[4] Ball R D, Carrazza S, del Debbio L, Forte S, Gao J *et al* 2013 Parton distribution benchmarking with LHC data *J. High Energy Phys.* JHEP04(2013)125
[5] Watt G 2011 Parton distribution function dependence of benchmark standard model total cross sections at the 7 TeVLHC *J. High Energy Phys.* JHEP09(2011)069
[6] de Roeck A and Thorne R 2011 Structure functions *Prog. Part. Nucl. Phys.* **66** 727–81
[7] Jimenez-Delgado P, Melnitchouk W and Owens J 2013 Parton momentum and helicity distributions in the nucleon *J. Phys. G: Nucl. Part. Phys.* **40** 093102
[8] Dittmaier S *et al* (LHC Higgs Cross section Working Group Collaboration) 2011 *Handbook of LHC Higgs cross sections: 1. Inclusive observables* (arXiv:1101.0593)
[9] Abelleira Fernandez J *et al* (LHeC Study Group Collaboration) 2012 A large hadron electron collider at CERN: report on the physics and design concepts for machine and detector *J. Phys. G: Nucl. Part. Phys.* **39** 075001
[10] Borschensky C *et al* 2014 Squark and gluino production cross sections in pp collisions at $\sqrt{s}$ = 13 14, 33 and 100 TeV (arXiv:1407.5066)
[11] Bozzi G, Rojo J and Vicini A 2011 The impact of PDF uncertainties on the measurement of the W boson mass at the Tevatron and the LHC *Phys. Rev.* D **83** 113008
[12] ATLAS Collaboration 2014 Studies of theoretical uncertainties on the measurement of the mass of the W boson at the LHC *ATLAS Public Note* ATL-PHYS-PUB-2014-015
[13] Bozzi G, Citelli L and Vicini A 2015 Parton density function uncertainties on the W boson mass measurement from the lepton transverse momentum distribution *Phys. Rev.* D **91** 113005
[14] ATLAS Collaboration 2015 Measurement of the forward–backward asymmetry of electron and muon pair-production in pp collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector (arXiv:1503.03709)
[15] Baak M *et al* (Gfitter Group Collaboration) 2014 The global electroweak fit at NNLO and prospects for the LHC and ILC *Eur. Phys. J.* C **74** 3046
[16] Alekhin S, Bluemlein J and Moch S 2014 The ABM parton distributions tuned to LHC data *Phys. Rev.* D **89** 054028
[17] Dulat S *et al* 2015 The CT14 global analysis of quantum chromodynamics arXiv:1506.07443
[18] Owens J, Accardi A and Melnitchouk W 2013 Global parton distributions with nuclear and finite-ιt Q^2 corrections *Phys. Rev.* D **87** 094012
[19] Jimenez-Delgado P and Reya E 2014 Delineating parton distributions and the strong coupling *Phys. Rev.* D **89** 074049
[20] Abramowicz H *et al* (H1 and ZEUS Collaboration) 2015 Combination of measurements of inclusive deep inelastic $e^{\pm}p$ scattering cross sections and QCD analysis of HERA data (arXiv:1506.06042)

[21] Harland-Lang L, Martin A, Motylinski P and Thorne R 2015 Parton distributions in the LHC era: MMHT 2014 PDFs *Eur. Phys. J.* C **75** 204
[22] Ball R D *et al* (NNPDF Collaboration) 2015 Parton distributions for the LHC Run II *J. High Energy Phys.* JHEP04(2015)040
[23] Watt B, Motylinski P and Thorne R 2014 The effect of LHC jet data on MSTW PDFs *Eur. Phys. J.* C **74** 2934
[24] Aad G *et al* (ATLAS Collaboration) 2013 Measurement of the inclusive jet cross section in pp collisions at sqrt(s) = 2.76 TeV and comparison to the inclusive jet cross section at sqrt (s) = 7 TeV using the ATLAS detector *Eur. Phys. J.* C **73** 2509
[25] Chatrchyan S *et al* (CMS Collaboration) 2013 Measurements of differential jet cross sections in proton–proton collisions at $\sqrt{s}$ = 7 TeV with the CMS detector *Phys. Rev.* D **87** 112002
[26] Chatrchyan S *et al* (CMS Collaboration) 2014 Measurement of the muon charge asymmetry in inclusive pp to WX production at sqrt(s) = 7 TeV and an improved determination of light parton distribution functions *Phys. Rev.* D **90** 032004
[27] Aad G *et al* (ATLAS Collaboration) 2012 Determination of the strange quark density of the proton from ATLAS measurements of the W, Z cross sections *Phys. Rev. Lett.* **109** 012001
[28] d'Enterria D and Rojo J 2012 Quantitative constraints on the gluon distribution function in the proton from collider isolated-photon data *Nucl. Phys.* B **860** 311–38
[29] Chatrchyan S *et al* (CMS Collaboration) 2014 Measurement of associated W+ charm production in pp collisions at $\sqrt{s}$ = 7 TeV *J. High Energy Phys.* JHEP02(2014)013
[30] Aad G *et al* (ATLAS Collaboration) 2014 Measurement of the production of a W boson in association with a charm quark in pp collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector *J. High Energy Phys.* JHEP05(2014)068
[31] Alekhin S *et al* 2015 Determination of strange sea quark distributions from fixed-target and collider data *Phys. Rev.* D **91** 094002
[32] Czakon M, Mangano M L, Mitov A and Rojo J 2013 Constraints on the gluon PDF from top quark pair production at hadron colliders *J. High Energy Phys.* JHEP07(2013)167
[33] Guzzi M, Lipka K and Moch S-O 2015 Top-quark pair production at hadron colliders: differential cross section and phenomenological applications with DiffTop *J. High Energy Phys.* JHEP01(2015)082
[34] Zenaiev O *et al* 2015 Impact of heavy-flavour production cross sections measured by the LHCb experiment on parton distribution functions at low x (arXiv:1503.04581)
[35] Aaij R *et al* (LHCb Collaboration) 2013 Prompt charm production in pp collisions at sqrt (s) = 7 TeV *Nucl. Phys.* B **871** 1–20
[36] Chatrchyan S *et al* (CMS Collaboration) 2013 Measurement of the differential and double-differential Drell–Yan cross sections in proton–proton collisions at $\sqrt{s}$ = 7 TeV *J. High Energy Phys.* JHEP12(2013)030
[37] Aad G *et al* (ATLAS Collaboration) 2013 Measurement of the high-mass Drell–Yan differential cross-section in pp collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector *Phys. Lett.* B **725** 223–42
[38] Malik S A and Watt G 2014 Ratios of W and Z cross sections at large boson p_T as a constraint on PDFs and background to new physics *J. High Energy Phys.* JHEP02(2014)025
[39] Farry S and Gauld R 2015 Prospects for measurements of $W^{\pm}$ boson production in association with jets at LHCb, and PDF constraints at large-x (arXiv:1505.01399)
[40] Stirling W and Vryonidou E 2012 Charm production in association with an electroweak gauge boson at the LHC *Phys. Rev. Lett.* **109** 082002
[41] Mason D *et al* 2007 Measurement of the nucleon strange–antistrange asymmetry at next-to-leading order in QCD from NuTeV Dimuon data *Phys. Rev. Lett.* **99** 192001
[42] Beneke M, Falgari P, Klein S, Piclum J, Schwinn C *et al* 2012 Inclusive top-pair production phenomenology with TOPIXS *J. High Energy Phys.* JHEP07(2012)194
[43] Gauld R, Rojo J, Rottoli L and Talbert J 2015 Charm production in the forward region: constraints on the small-x gluon and backgrounds for neutrino astronomy (arXiv:1506.08025)
[44] Mangano M L and Rojo J 2012 Cross section ratios between different CM energies at the LHC: opportunities for precision measurements and BSM sensitivity *J. High Energy Phys.* JHEP08 (2012)010
[45] Khachatryan V *et al* (CMS Collaboration) 2015 Measurements of differential and double-differential Drell–Yan cross sections in proton-proton collisions at 8 TeV *Eur. Phys. J.* C **75** 147

[46] Skands P, Carrazza S and Rojo J 2014 Tuning PYTHIA 8.1: the Monash 2013 Tune *Eur. Phys. J.* **74** 3024

[47] ATLAS Collaboration 2014 ATLAS Run 1 Pythia8 tunes *ATLAS Public Note* ATL-PHYS-PUB-2014-021

[48] Khachatryan V *et al* (CMS Collaboration) 2015 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 *Euro. Phys. J.* C **75** 288

[49] Chatrchyan S *et al* (CMS Collaboration) 2013 Measurement of the ratio of the inclusive 3-jet cross section to the inclusive 2-jet cross section in pp collisions at $\sqrt{s} = 7$ TeV and first determination of the strong coupling constant in the TeV range *Eur. Phys. J.* C **73** 2604

[50] Alekhin S *et al* 2015 HERAFitter *Euro. Phys. J.* C **75** 304

[51] Camarda S *et al* (HERAFitter developers' Team Collaboration) 2015 QCD analysis of W- and Z-boson production at Tevatron (arXiv:1503.05221)

[52] Belov P *et al* (HERAFitter developers' Team Collaboration) 2014 Parton distribution functions at LO, NLO and NNLO with correlated uncertainties between orders *Eur. Phys. J.* C **74** 3039

[53] Gao J and Nadolsky P 2014 A meta-analysis of parton distribution functions *J. High Energy Phys.* JHEP07(2014)035

[54] Carrazza S, Forte S, Kassabov Z, Latorre J I and Rojo J 2015 An unbiased Hessian representation for Monte Carlo PDFs *Euro. Phys. J.* C **75** 369

[55] Carrazza S, Latorre J I, Rojo J and Watt G 2015 A compression algorithm for the combination of PDF sets (arXiv:1504.06469)

[56] Alekhin S *et al* 2011 *PDF4LHC Working Group Interim Report* (arXiv:1101.0536)

[57] Botje M *et al* 2011 *The PDF4LHC Working Group Interim Recommendations* (arXiv:1101.0538)

[58] Buckley A *et al* 2015 LHAPDF6: parton density access in the LHC precision era *Eur. Phys. J.* C **75** 132

[59] Gao J *et al* 2014 CT10 next-to-next-to-leading order global analysis of QCD *Phys. Rev.* D **89** 033009

[60] Abazov V M *et al* (D0 Collaboration) 2015 Measurement of the electron charge asymmetry in $p\bar{p} \to W + X \to e\nu + X$ decays in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV *Phys. Rev.* D **1** 032007

[61] Carli T *et al* 2010 *A posteriori* inclusion of parton density functions in NLO QCD final-state calculations at hadron colliders: the APPLGRID Project *Eur. Phys. J.* C **66** 503–24

[62] Wobisch M, Britzger D, Kluge T, Rabbertz K, Stober F (fastNLO Collaboration) 2011 Theory-data comparisons for jet measurements in hadron-induced processes (arXiv:1109.1310)

[63] Balazs C, Qiu J-w and Yuan C 1995 Effects of QCD resummation on distributions of leptons from the decay of electroweak vector bosons *Phys. Lett.* B **355** 548–54

[64] Balazs C and Yuan C 1997 Soft gluon effects on lepton pairs at hadron colliders *Phys. Rev.* D **56** 5558–83

[65] Landry F, Brock R, Nadolsky P M and Yuan C 2003 Tevatron Run-1 Z boson data and Collins–Soper–Sterman resummation formalism *Phys. Rev.* D **67** 073016

[66] Guzzi M, Nadolsky P M and Wang B 2014 Nonperturbative contributions to a resummed leptonic angular distribution in inclusive neutral vector boson production *Phys. Rev.* D **90** 014030

[67] Pumplin J *et al* 2002 New generation of parton distributions with uncertainties from global QCD analysis *J. High Energy Phys.* JHEP07(2002)012

[68] Accardi A *et al* 2010 New parton distributions from large-x and low-Q^2 data *Phys. Rev.* D **81** 034016

[69] Accardi A *et al* 2011 Uncertainties in determining parton distributions at large x *Phys. Rev.* D **84** 014008

[70] Ball R D *et al* (The NNPDF Collaboration) 2013 Theoretical issues in PDF determination and associated uncertainties *Phys. Lett.* B **723** 330–9

[71] Melnitchouk W and Thomas A W 1996 Neutron/proton structure function ratio at large x *Phys. Lett.* B **377** 11–17

[72] Holt R J and Roberts C D 2010 Distribution functions of the nucleon and pion in the valence region *Rev. Mod. Phys.* **82** 2991–3044

[73] Accardi A 2013 Large-x connections of nuclear and high-energy physics *Mod. Phys. Lett.* A **28** 1330032

[74] Brady L, Accardi A, Melnitchouk W and Owens J 2012 Impact of PDF uncertainties at large x on heavy boson production *J. High Energy Phys.* JHEP06(2012)019

[75] Aaron F *et al* (H1 and ZEUS Collaboration) 2010 Combined measurement and QCD analysis of the inclusive e^+-p scattering cross sections at HERA *J. High Energy Phys.* **1001** JHEP01(2010)109
[76] Martin A D, Roberts R G, Stirling W J and Thorne R S 2005 Parton distributions incorporating QED contributions *Eur. Phys. J.* C **39** 155–61
[77] Martin A and Ryskin M 2014 The photon PDF of the proton *Eur. Phys. J.* C **74** 3040
[78] Martin A D, Stirling W J, Thorne R S and Watt G 2009 Parton distributions for the LHC *Eur. Phys. J.* C **63** 189–285
[79] Martin A *et al* 2013 Extended parameterisations for MSTW PDFs and their effect on lepton charge asymmetry from *W* decays *Eur. Phys. J.* C **73** 2318
[80] Thorne R 2012 The effect of changes of variable flavour number scheme on PDFs and predicted cross sections *Phys. Rev.* D **86** 074017
[81] de Florian D, Sassot R, Zurita P and Stratmann M 2012 Global analysis of nuclear parton distributions *Phys. Rev.* D **85** 074028
[82] Ball R D *et al* (The NNPDF Collaboration) 2010 Fitting parton distribution data with multiplicative normalization uncertainties *J. High Energy Phys.* JHEP05(2010)075
[83] Abramowicz H *et al* (H1, ZEUS Collaboration) 2013 Combination and QCD analysis of charm production cross section measurements in deep-inelastic ep scattering at HERA *Eur. Phys. J.* C **73** 2311
[84] Aaltonen T A *et al* (CDF, D0 Collaboration) 2014 Combination of measurements of the top-quark pair production cross section from the Tevatron collider *Phys. Rev.* D **89** 072001
[85] Harland-Lang L, Martin A, Motylinski P and Thorne R 2015 Uncertainties on α_S in the MMHT2014 global PDF analysis and implications for SM predictions (arXiv:1506.05682)
[86] Ball R D *et al* (The NNPDF Collaboration) 2009 A determination of parton distributions with faithful uncertainty estimation *Nucl. Phys.* B **809** 1–63
[87] Ball R D *et al* (The NNPDF Collaboration) 2009 Precision determination of electroweak parameters and the strange content of the proton from neutrino deep-inelastic scattering *Nucl. Phys.* B **823** 195–233
[88] Ball R D *et al* (The NNPDF Collaboration) 2010 A first unbiased global NLO determination of parton distributions and their uncertainties *Nucl. Phys.* B **838** 136–206
[89] Ball R D *et al* (The NNPDF Collaboration) 2011 Impact of heavy quark masses on parton distributions and LHC phenomenology *Nucl. Phys.* B **849** 296–363
[90] Ball R D *et al* (The NNPDF Collaboration) 2012 Unbiased global determination of parton distributions and their uncertainties at NNLO and at LO *Nucl. Phys.* B **855** 153–221
[91] Ball R D *et al* (NNPDF Collaboration) 2013 Parton distributions with LHC data *Nucl. Phys.* B **867** 244–89
[92] Bertone V, Frederix R, Frixione S, Rojo J and Sutton M 2014 aMCfast: automation of fast NLO computations for PDF fits *J. High Energy Phys.* JHEP08(2014)166
[93] de Florian D, Hinderer P, Mukherjee A, Ringer F and Vogelsang W 2014 Approximate next-to-next-to-leading order corrections to hadronic jet production *Phys. Rev. Lett.* **112** 082001
[94] Carrazza S and Pires J 2014 Perturbative QCD description of jet data from LHC Run-I and Tevatron Run-II *J. High Energy Phys.* JHEP10(2014)145
[95] Gehrmann-de Ridder A, Gehrmann T, Glover E and Pires J 2013 Second order QCD corrections to jet production at hadron colliders: the all-gluon contribution *Phys. Rev. Lett.* **110** 162003
[96] Forte S, Laenen E, Nason P and Rojo J 2010 Heavy quarks in deep-inelastic scattering *Nucl. Phys.* B **834** 116–62
[97] HERAFitter, An open source QCD fit framework (http://herafitter.org)
[98] Camarda S 2015 HERAFitter—An open source framework to determine PDFs (arXiv:1505.04973)
[99] Adloff C *et al* (H1 Collaboration) 2001 Measurement and QCD analysis of jet cross-sections in deep inelastic positron–proton collisions at s**(1/2) of 300 GeV *Eur. Phys. J.* C **19** 289–311
[100] Kluge T, Rabbertz K and Wobisch M 2006 Fast pQCD calculations for PDF fits (arXiv:hep-ph/0609285)
[101] Carli T, Salam G P and Siegert F 2005 *A posteriori* inclusion of PDFs in NLO QCD final-state calculations (arXiv:hep-ph/0510324)
[102] Nagy Z and Trocsanyi Z 1999 Next-to-leading order calculation of four jet observables in electron positron annihilation *Phys. Rev.* D **D59** 014020

[103] Nagy Z and Trocsanyi Z 2001 Multijet cross-sections in deep inelastic scattering at next-to-leading order *Phys. Rev. Lett.* **87** 082001
[104] Nagy Z 2003 Next-to-leading order calculation of three-jet observables in hadron hadron collision *Phys. Rev.* D **68** 094002
[105] Nagy Z 2002 Three jet cross-sections in hadron hadron collisions at next-to-leading order *Phys. Rev. Lett.* **88** 122003
[106] Kidonakis N and Owens J F 2001 Effects of higher-order threshold corrections in high-$E(T)$ jet production *Phys. Rev.* D **63** 054019
[107] Britzger D, Rabbertz K, Stober F and Wobisch M 2012 New features in version 2 of the fastNLO project (arXiv:1208.3641)
[108] Britzger D, Guzzi M, Rabbertz K, Sieber G, Stober F and Wobisch M 2014 Generalization of the fastNLO approach to NNLO calculations *Proc. of the 22nd Int. Workshop on Deep-Inelastic Scattering and Related Subjects* p 202
[109] http://applgrid.hepforge.org
[110] Campbell J M and Ellis R K 1999 An update on vector boson pair production at hadron colliders *Phys. Rev.* D **60** 113006
[111] Campbell J M and Ellis R K 2000 Radiative corrections to Zbbbar production *Phys. Rev.* D **62** 114012
[112] Campbell J M and Ellis R K 2010 MCFM for the Tevatron and the LHC *Nucl. Phys.* B **205–206** 10–15
[113] Salam G P and Rojo J 2009 A higher order perturbative parton evolution toolkit (HOPPET) *Comput. Phys. Commun.* **180** 120–56
[114] Alwall J *et al* 2014 Automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations *J. High Energy Phys.* JHEP07 (2014)079
[115] Bertone V, Carrazza S and Rojo J 2014 APFEL: a PDF evolution library with QED corrections *Comput. Phys. Commun.* **185** 1647–68
[116] Carrazza S, Ferrara A, Palazzo D and Rojo J 2015 APFEL Web: a web-based application for the graphical visualization of parton distribution functions *J. Phys. G: Nucl. Part. Phys.* **42** 057001
[117] Rojo J 2014 Constraints on parton distributions and the strong coupling from LHC jet data (arXiv:1410.7728)
[118] Currie J, Gehrmann-de Ridder A, Glover E and Pires J 2014 NNLO QCD corrections to jet production at hadron colliders from gluon scattering *J. High Energy Phys.* JHEP01 (2014)110
[119] Vogelsang W and Vogt A 1995 Constraints on the proton's gluon distribution from prompt photon production *Nucl. Phys.* B **453** 334–54
[120] Carminati L *et al* 2013 Sensitivity of the LHC isolated-gamma+jet data to the parton distribution functions of the proton *Europhys. Lett.* **101** 61002
[121] ATLAS Collaboration 2013 A study of the sensitivity to the proton parton distributions of the inclusive photon production cross section in pp collisions at 7 TeV measured by the ATLAS experiment at the LHC *ATLAS Public Note* ATL-PHYS-PUB-2013-018
[122] Aad G *et al* (ATLAS Collaboration) 2012 Measurement of the inclusive W+- and Z/gamma cross sections in the electron and muon decay channels in pp collisions at sqrt(s) = 7 TeV with the ATLAS detector *Phys. Rev.* D **85** 072004
[123] Moreno G *et al* 1991 Dimuon production in proton–copper collisions at $\sqrt{s} = 38.8$ GeV *Phys. Rev.* D **43** 2815–36
[124] Towell R S *et al* (FNAL E866/NuSea Collaboration) 2001 Improved measurement of the anti-d/anti-u asymmetry in the nucleon sea *Phys. Rev.* D **64** 052002
[125] Marzani S and Ball R D 2009 High energy resummation of Drell–Yan processes *Nucl. Phys.* B **814** 246–64
[126] Catani S, Cieri L, Ferrera G, de Florian D and Grazzini M 2009 Vector boson production at hadron colliders: a fully exclusive QCD calculation at NNLO *Phys. Rev. Lett.* **103** 082001
[127] Gavin R, Li Y, Petriello F and Quackenbush S 2013 W Physics at the LHC with FEWZ 2.1 *Comput. Phys. Commun.* **184** 208–14
[128] Boughezal R, Li Y and Petriello F 2014 Disentangling radiative corrections using high-mass Drell–Yan at the LHC *Phys. Rev.* D **89** 034030
[129] Ball R D *et al* (NNPDF Collaboration) 2013 Parton distributions with QED corrections *Nucl. Phys.* B **877** 290–320

[130] Aad G *et al* (ATLAS Collaboration) 2014 Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector *J. High Energy Phys.* JHEP09(2014)145
[131] Khachatryan V *et al* (CMS Collaboration) 2015 Measurement of the Z boson differential cross section in transverse momentum and rapidity in proton-proton collisions at 8 TeV (arXiv:1504.03511)
[132] Boughezal R, Focke C, Liu X and Petriello F 2015 *W*-boson production in association with a jet at next-to-next-to-leading order in perturbative QCD *Phys. Rev. Lett.* **115** 062022
[133] Baur U, Halzen F, Keller S, Mangano M L and Riesselmann K 1993 The Charm content of W + 1 jet events as a probe of the strange quark distribution function *Phys. Lett.* B **318** 544–8
[134] Goncharov M *et al* (NuTeV Collaboration) 2001 Precise measurement of dimuon production cross-sections in nu/mu Fe and anti-nu/mu Fe deep inelastic scattering at the Tevatron *Phys. Rev.* D **64** 112006
[135] Czakon M, Fiedler P and Mitov A 2013 The total top quark pair production cross-section at hadron colliders through $O(\alpha_S^4)$ *Phys. Rev. Lett.* **110** 252004
[136] Czakon M, Fiedler P and Mitov A 2015 Resolving the Tevatron top quark forward-backward asymmetry puzzle *Phys. Rev. Lett.* **115** 052001
[137] Abelof G, Ridder A G-D and Majer I 2015 Top quark pair production at NNLO in the quark-antiquark channel (arXiv:1506.04037)
[138] Chatrchyan S *et al* (CMS Collaboration) 2014 Determination of the top-quark pole mass and strong coupling constant from the t t-bar production cross section in pp collisions at $\sqrt{s} = 7$ TeV *Phys. Lett.* B **728** 496–517
[139] Jones S, Martin A, Ryskin M and Teubner T 2013 Probes of the small x gluon via exclusive J/ψ and Υ production at HERA and the LHC *J. High Energy Phys.* JHEP11(2013)085
[140] Jones S, Martin A, Ryskin M and Teubner T 2014 Predictions of exclusive (2S) production at the LHC *J. Phys. G: Nucl. Part. Phys.* **41** 055009
[141] Aaij R *et al* (LHCb Collaboration) 2013 Exclusive J/ψ and ψ(2S) production in pp collisions at $\sqrt{s} = 7$ TeV *J. Phys. G: Nucl. Part. Phys.* **40** 045001
[142] Aaij R *et al* (LHCb Collaboration) 2014 Updated measurements of exclusive J/ψ and ψ(2S) production cross-sections in pp collisions at $\sqrt{s} = 7$ TeV *J. Phys. G: Nucl. Part. Phys.* **41** 055002
[143] Jones S P, Martin A D, Ryskin M G and Teubner T Exclusive J/ψ and Υ photo production and the low x gluon (arXiv:1507.06942)
[144] Abelev B *et al* (ALICE Collaboration) 2014 Exclusive J/ψ photoproduction off protons in ultraperipheral p–P_b collisions at $\sqrt{s_{NN}} = 5.02$ TeV *Phys. Rev. Lett.* **113** 232504
[145] Aad G *et al* (ATLAS Collaboration) 2014 Measurement of the low-mass Drell–Yan differential cross section at $\sqrt{s} = 7$ TeV using the ATLAS detector *J. High Energy Phys.* JHEP06(2014)112
[146] Aad G *et al* (ATLAS Collaboration) 2012 Measurement of the cross section for the production of a W boson in association with b-jets in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector *Phys. Lett.* B **707** 418–37
[147] Aad G *et al* (ATLAS Collaboration) 2013 Measurement of the cross-section for W boson production in association with b-jets in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector *J. High Energy Phys.* JHEP06(2013)084
[148] Aad G *et al* (ATLAS Collaboration) 2012 Measurement of the cross-section for b jets produced in association with a Z boson at $\sqrt{s} = 7$ TeV with the ATLAS detector *Phys. Lett.* B **706** 295–313
[149] Aad G *et al* (ATLAS Collaboration) 2014 Measurement of differential production cross-sections for a Z boson in association with b-jets in 7 TeV proton–proton collisions with the ATLAS detector *J. High Energy Phys.* JHEP10(2014)141
[150] Aad G *et al* (ATLAS Collaboration) 2011 Measurement of the transverse momentum distribution of Z/gamma* bosons in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector *Phys. Lett.* B **705** 415–34
[151] Aad G *et al* (ATLAS Collaboration) 2012 Measurement of the transverse momentum distribution of W Bosons in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector *Phys. Rev.* D **85** 012005
[152] Aad G *et al* (ATLAS Collaboration) 2012 Measurement of the production cross section for Z/gamma* in association with jets in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector *Phys. Rev.* D **85** 032009

[153] Aad G *et al* (ATLAS Collaboration) 2013 Measurement of the production cross section of jets in association with a Z boson in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector *J. High Energy Phys.* JHEP07(2013)032
[154] Aad G *et al* (ATLAS Collaboration) 2012 Study of jets produced in association with a W boson in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector *Phys. Rev.* D **85** 092002
[155] Aad G *et al* (ATLAS Collaboration) 2015 Measurements of the W production cross sections in association with jets with the ATLAS detector *Eur. Phys. J.* C **75** 82
[156] Aad G *et al* (ATLAS Collaboration) 2014 A measurement of the ratio of the production cross sections for W and Z bosons in association with jets with the ATLAS detector *Eur. Phys. J.* C **74** 3168
[157] Aad G *et al* (ATLAS Collaboration) 2012 Measurement of inclusive jet and dijet production in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector *Phys. Rev.* D **86** 014022
[158] Aad G *et al* (ATLAS Collaboration) 2015 Measurement of the inclusive jet cross-section in proton–proton collisions at $\sqrt{s} = 7$ TeV using 4.5 fb^{-1} of data with the ATLAS detector *J. High Energy Phys.* JHEP02(2015)153
[159] Aad G *et al* (ATLAS Collaboration) 2014 Measurement of dijet cross sections in pp collisions at 7 TeV centre-of-mass energy using the ATLAS detector *J. High Energy Phys.* JHEP05 (2014)059
[160] Aad G *et al* (ATLAS Collaboration) 2015 Measurement of three-jet production cross-sections in pp collisions at 7 TeV centre-of-mass energy using the ATLAS detector *Eur. Phys. J.* C **75** 228
[161] Aad G *et al* (ATLAS Collaboration) 2011 Measurement of the inclusive isolated prompt photon cross-section in pp collisions at $\sqrt{s} = 7$ TeV using 35 pb-1 of ATLAS data *Phys. Lett.* B **706** 150–67
[162] Aad G *et al* (ATLAS Collaboration) 2014 Measurement of the inclusive isolated prompt photons cross section in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector using 4.6 fb^{-1} *Phys. Rev.* D **89** 052004
[163] Aad G *et al* (ATLAS Collaboration) 2012 Measurement of the production cross section of an isolated photon associated with jets in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector *Phys. Rev.* D **85** 092014
[164] Aad G *et al* (ATLAS Collaboration) 2011 Measurement of the top quark-pair production cross section with ATLAS in pp collisions at $\sqrt{s} = 7$ TeV *Eur. Phys. J.* C **71** 1577
[165] Aad G *et al* (ATLAS Collaboration) 2012 Measurement of the top quark pair production cross section in pp collisions at $\sqrt{s} = 7$ TeV in dilepton final states with ATLAS *Phys. Lett.* B **707** 459–77
[166] Aad G *et al* (ATLAS Collaboration) 2012 Measurement of the top quark pair production cross-section with ATLAS in the single lepton channel *Phys. Lett.* B **711** 244–63
[167] Aad G *et al* (ATLAS Collaboration) 2012 Measurement of the cross section for top-quark pair production in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector using final states with two high-pt leptons *J. High Energy Phys.* JHEP05(2012)059
[168] Aad G *et al* (ATLAS Collaboration) 2012 Measurement of the top quark pair cross section with ATLAS in pp collisions at sqrt(s) = 7 TeV using final states with an electron or a muon and a hadronically decaying τ lepton *Phys. Lett.* B **717** 89–108
[169] Aad G *et al* (ATLAS Collaboration) 2013 Measurement of the ttbar production cross section in the tau+jets channel using the ATLAS detector *Eur. Phys. J.* C **73** 2328
[170] Aad G *et al* (ATLAS Collaboration) 2014 Measurement of the $t\bar{t}$ production cross-section using $e\mu$ events with b -tagged jets in pp collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS detector *Eur. Phys. J.* C **74** 3109
[171] Aad G *et al* (ATLAS Collaboration) 2013 Measurements of top quark pair relative differential cross-sections with ATLAS in pp collisions at $\sqrt{s} = 7$ TeV *Eur. Phys. J.* C **73** 2261
[172] Aad G *et al* (ATLAS Collaboration) 2014 Measurements of normalized differential cross sections for $t\bar{t}$ production in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector *Phys. Rev.* D **90** 072004
[173] Aad G *et al* (ATLAS Collaboration) 2015 Simultaneous measurements of the $t\bar{t}$, W^+W^-, and $Z/\gamma^* \rightarrow \tau\tau$ production cross-sections in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector *Phys. Rev.* D **91** 052005
[174] Cacciari M, Salam G P and Soyez G 2008 The Anti-k(t) jet clustering algorithm *J. High Energy Phys.* JHEP04(2008)063

[175] Maltoni F, Ridolfi G and Ubiali M 2012 B-initiated processes at the LHC: a reappraisal *J. High Energy Phys.* JHEP07(2012)022
[176] Chatrchyan S *et al* (CMS Collaboration) 2013 Forward–backward asymmetry of Drell–Yan lepton pairs in pp collisions at $\sqrt{s} = 7$ TeV *Phys. Lett.* B **718** 752–72
[177] Chatrchyan S *et al* (CMS Collaboration) 2011 Measurement of the lepton charge asymmetry in inclusive W production in pp collisions at $\sqrt{s} = 7$ TeV *J. High Energy Phys.* JHEP04 (2011)050
[178] Chatrchyan S *et al* (CMS Collaboration) 2012 Measurement of the electron charge asymmetry in inclusive W production in pp collisions at sqrt(s) = 7 TeV *Phys. Rev. Lett.* **109** 111806
[179] Khachatryan V *et al* (CMS Collaboration) 2011 Measurements of Inclusive W and Z Cross sections in pp Collisions at $\sqrt{s} = 7$ TeV *J. High Energy Phys.* JHEP01(2011)080
[180] Chatrchyan S *et al* (CMS Collaboration) 2011 Measurement of the Inclusive W and Z Production Cross sections in pp Collisions at $\sqrt{s} = 7$ TeV *J. High Energy Phys.* JHEP10(2011)132
[181] Chatrchyan S *et al* (CMS Collaboration) 2014 Measurement of inclusive W and Z boson production cross sections in pp collisions at $\sqrt{s} = 8$ TeV *Phys. Rev. Lett.* **112** 191802
[182] Chatrchyan S *et al* (CMS Collaboration) 2012 Measurement of the rapidity and transverse momentum distributions of Z Bosons in pp collisions at $\sqrt{s} = 7$ TeV *Phys. Rev.* D **85** 032002
[183] Chatrchyan S *et al* (CMS Collaboration) 2014 Measurement of the ratio of inclusive jet cross sections using the anti- k_T algorithm with radius parameters $R = 0.5$ and 0.7 in pp collisions at $\sqrt{s} = 7$ TeV *Phys. Rev.* D **90** 072006
[184] Khachatryan V *et al* (CMS Collaboration) 2015 Measurement of the inclusive 3-jet production differential cross section in proton-proton collisions at 7 TeV and determination of the strong coupling constant in the TeV range *Eur. Phys. J.* C **75** 186
[185] Chatrchyan S *et al* (CMS Collaboration) 2014 Measurement of the production cross sections for a Z boson and one or more b jets in pp collisions at sqrt(s) = 7 TeV *J. High Energy Phys.* JHEP06(2014)120
[186] Chatrchyan S *et al* (CMS Collaboration) 2011 Measurement of the differential cross section for isolated prompt photon production in pp collisions at 7 TeV *Phys. Rev.* D **84** 052011
[187] Chatrchyan S *et al* (CMS Collaboration) 2014 Measurement of the triple-differential cross section for photon+jets production in proton-proton collisions at $\sqrt{s} = 7$ TeV *J. High Energy Phys.* JHEP06(2014)009
[188] Chatrchyan S *et al* (CMS Collaboration) 2012 Measurement of the $t\bar{t}$ production cross section in the dilepton channel in pp collisions at $\sqrt{s} = 7$ TeV *J. High Energy Phys.* JHEP11(2012)067
[189] Chatrchyan S *et al* (CMS Collaboration) 2013 Measurement of top-quark pair production cross sections in pp colisions at $\sqrt{s} = 7$ TeV *Eur. Phys. J.* C **73** 2339
[190] CMS Collaboration CMS 2012 Measurement of the $t\bar{t}$ production cross section in the dilepton channel in pp collisions at $\sqrt{s} = 7$ TeV *J. High Energy Phys.* JHEP11(2012)067
[191] CMS Collaboration CMS 2012 Top pair cross section in e/mu+jets at 8 TeV *CMS Phys. Anal. Summary* CMS PAS TOP-12-006
[192] CMS Collaboration CMS 2014 Measurement of the $t\bar{t}$ production cross section in the dilepton channel in pp collisions at $\sqrt{s} = 8$ TeV *J. High Energy Phys.* JHEP02(2014)024
[193] Khachatryan V *et al* (CMS Collaboration) 2015 Measurement of the differential cross section for top quark pair production in pp collisions at sqrt(s) = 8 TeV (arXiv:1505.04480)
[194] Martin A D, Roberts R G, Stirling W J and Thorne R S 2003 Uncertainties of predictions from parton distributions: I. Experimental errors. ((T)) *Eur. Phys.* J **C28** 455–73
[195] Becciolini D, Gillioz M, Nardecchia M, Sannino F and Spannowsky M 2015 Constraining new colored matter from the ratio of 3 to 2 jets cross sections at the LHC *Phys. Rev.* D **91** 015010
[196] Chatrchyan S *et al* (CMS Collaboration) 2011 Measurement of the weak mixing angle with the Drell–Yan process in proton-proton collisions at the LHC *Phys. Rev.* D **84** 112002
[197] Thorne R, Martin A, Stirling W and Watt G 2008 Parton distributions and QCD at LHCb arXiv:0808.1847
[198] Aaij R *et al* (LHCb Collaboration) 2012 Inclusive W and Z production in the forward region at $\sqrt{s} = 7$ TeV *J. High Energy Phys.* JHEP06(2012)058
[199] Aaij R *et al* (LHCb collaboration) 2013 Measurement of the cross-section for $Z \to e^+e^-$ production in pp collisions at $\sqrt{s} = 7$ TeV *J. High Energy Phys.* JHEP02(2013)106
[200] Aaij R *et al* (LHCb Collaboration) 2014 Measurement of the forward W boson cross-section in pp collisions at $\sqrt{s} = 7$ TeV *J. High Energy Phys.* JHEP12(2014)079

[201] Aaij R *et al* (LHCb Collaboration) 2015 Measurement of the forward Z boson production cross-section in pp collisions at $\sqrt{s} = 7$ TeV *J. High Energy Phys.* JHEP08(2015)039
[202] Aaij R *et al* (LHCb Collaboration) 2014 Precision luminosity measurements at LHCb *J. Instrum.* **9** P12005
[203] Aaij R *et al* (LHCb Collaboration) 2015 Measurement of forward $Z \to e^+e^-$ production at $\sqrt{s} = 8$ TeV *J. High Energy Phys.* JHEP05(2015)109
[204] LHCb Collaboration 2012 Inclusive low mass Drell–Yan production in the forward region at $\sqrt{s} = 7$ TeV *LHCb Note* LHCb-CONF-2012-013
[205] Aaij R *et al* (LHCb Collaboration) 2015 Measurement of the Z+ b-jet cross-section in pp collisions at $\sqrt{s} = 7$ TeV in the forward region *J. High Energy Phys.* JHEP01(2015)064
[206] Aaij R *et al* (LHCb Collaboration) 2014 Observation of associated production of a Z boson with a D meson in the forward region *J. High Energy Phys.* JHEP04(2014)091
[207] Aaij R *et al* (LHCb Collaboration) 2015 Study of W boson production in association with beauty and charm (arXiv:1505.04051)
[208] Aaij R *et al* (LHCb Collaboration) 2015 Identification of beauty and charm quark jets at LHCb (arXiv:1504.07670)
[209] Aaij R *et al* (LHCb Collaboration) 2015 First observation of top quark production in the forward region (arXiv:1506.00903)
[210] Aaij R *et al* (LHCb Collaboration) 2013 Measurement of B meson production cross-sections in proton-proton collisions at $\sqrt{s} = 7$ TeV *J. High Energy Phys.* JHEP08(2013)117
[211] Aaij R *et al* (LHCb Collaboration) 2015 Measurement of the exclusive Υ production cross-section in pp collisions at $\sqrt{s} = 7$ TeV and 8 TeV (arXiv:1505.08139)
[212] Bednyakov V, Demichev M, Lykasov G, Stavreva T and Stockton M 2013 Searching for intrinsic charm in the proton at the LHC *EPJ Web Conf.* **60** 20047
[213] Kagan A L, Kamenik J F, Perez G and Stone S 2011 Top LHCb Physics *Phys. Rev. Lett.* **107** 082003
[214] Gauld R 2014 Feasibility of top quark measurements at LHCb and constraints on the large-x gluon PDF *J. High Energy Phys.* JHEP02(2014)126
[215] Paukkunen H and Zurita P 2014 PDF reweighting in the Hessian matrix approach *J. High Energy Phys.* JHEP12(2014)100
[216] Ball R D *et al* 2012 Reweighting and unweighting of parton distributions and the LHC W lepton asymmetry data *Nucl. Phys.* B **855** 608–38
[217] Ball R D *et al* (The NNPDF Collaboration) 2011 Reweighting NNPDFs: the W lepton asymmetry *Nucl. Phys.* B **849** 112–43
[218] Anastasiou C, Dixon L J, Melnikov K and Petriello F 2004 High precision QCD at hadron colliders: electroweak gauge boson rapidity distributions at NNLO *Phys. Rev.* D **69** 094008
[219] Czakon M and Mitov A 2014 Top^{++}: a program for the calculation of the top-pair cross-section at hadron colliders *Comput. Phys. Commun.* **185** 2930–8
[220] Lai H-L *et al* 2010 New parton distributions for collider physics *Phys. Rev.* D **82** 074024
[221] Pumplin J, Huston J, Lai H L, Tung W-K and Yuan C P 2009 Collider inclusive jet data and the gluon distribution *Phys. Rev.* D **80** 014019
[222] Abazov V *et al* (D0 Collaboration) 2010 Measurement of the dijet invariant mass cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV *Phys. Lett.* B **693** 531–8
[223] Aaltonen T *et al* (CDF Collaboration) 2008 Measurement of the inclusive jet cross section at the Fermilab Tevatron p-pbar collider using a cone-based jet algorithm *Phys. Rev.* D **78** 052006
[224] Ellis S D, Kunszt Z and Soper D E 1992 Two jet production in hadron collisions at order alpha-s**3 in QCD *Phys. Rev. Lett.* **69** 1496–9
[225] Gao J *et al* 2013 MEKS: a program for computation of inclusive jet cross sections at hadron colliders *Comput. Phys. Commun.* **184** 1626–42
[226] Cacciari M and Houdeau N 2011 Meaningful characterisation of perturbative theoretical uncertainties *J. High Energy Phys.* JHEP09(2011)039
[227] Olness F I and Soper D E 2010 Correlated theoretical uncertainties for the one-jet inclusive cross section *Phys. Rev.* D **81** 035018
[228] Butterworth J *et al* 2014 Les Houches physics at TeV colliders: standard model working group report (arXiv:1405.1067)
[229] Rojo J *et al* 2010 The SM and NLO multileg working group: summary report chapter 22 (arXiv:1003.1241)