ELSEVIER

MSTW PDFs and impact of PDFs on cross sections at Tevatron and LHC

Graeme Watt

Theory Group, Physics Department, CERN, 1211 Geneva 23, Switzerland

Abstract

We briefly summarise the "MSTW 2008" determination of parton distribution functions (PDFs), and subsequent follow-up studies, before reviewing some topical issues concerning the PDF dependence of cross sections at the Tevatron and LHC. We update a recently published study of benchmark Standard Model total cross sections (W, Z, $gg \to H$ and $t\bar{t}$ production) at the 7 TeV LHC, where we account for all publicly available PDF sets and we compare to LHC data for W, Z, and $t\bar{t}$ production. We show the sensitivity of the Higgs cross sections to the gluon distribution, then we demonstrate the ability of the Tevatron jet data, and also the LHC $t\bar{t}$ data, to discriminate between PDF sets with different high-x gluon distributions. We discuss the related problem of attempts to extract the strong coupling α_S from only deep-inelastic scattering data, and we conclude that a direct data constraint on the high-x gluon distribution is required to obtain a meaningful result. We therefore discourage the use of PDF sets obtained from "non-global" fits where the high-x gluon distribution is not directly constrained by data.
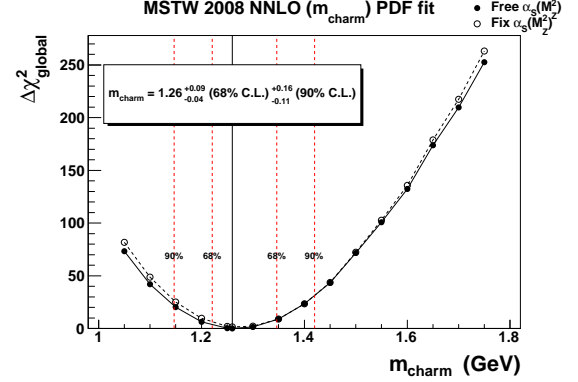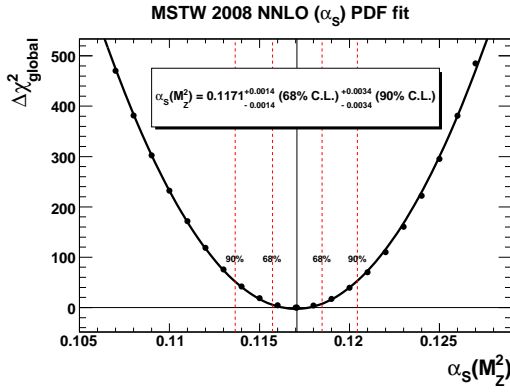
Keywords:

1. Introduction

The parton distribution functions (PDFs) of the proton are a non-negotiable input to almost all theory predictions at hadron colliders. The proton PDFs are determined by several groups from (global) analysis of a wide range of deep-inelastic scattering (DIS) and related hard-scattering data. The DIS data from HERA are perhaps the single most important input to global PDF fits. It is even possible to extract PDFs based *only* on HERA DIS data, albeit at the expense of leaving some kinematic regions and flavour combinations unconstrained, or alternatively by imposing a severe parameterisation constraint. The HERA data are therefore generally supplemented with DIS and Drell–Yan data from fixed-target experiments.

Of course, we need to know PDFs to predict Tevatron and LHC cross sections, but this argument can also work the other way around. If a cross section at a hadron collider is predicted with relatively small theoretical uncertainty, and it is sensitive to PDFs in a kinematic region poorly constrained by HERA and fixed-target data, then

precise measurements of hadron collider cross sections can give important information on PDFs. To give an example, inclusive jet production at the Tevatron is currently essential to directly constrain the high-x gluon.

The contents of this contribution to the proceedings are as follows. First in Sec. 2 we briefly review the status of the "MSTW" determination of PDFs. However, this write-up will mostly be based on two recent papers [1, 2], with some updates to account for new PDF sets and LHC data released subsequent to their publication. To avoid a deluge of plots we will only present a limited selection in this write-up, and a more extensive collection can be found at a public webpage [3]. In Sec. 3 we describe a recent benchmark exercise and give the status of the most recent PDF sets from all fitting groups (as of December 2011). In Sec. 4 we discuss W and Z production at the LHC, and the sensitivity to the quark distributions. In Sec. 5 we discuss Higgs, top-pair and jet production, and the sensitivity to the gluon distribution. In Sec. 6 we discuss the values of α_S obtained from DIS data. Finally, we summarise in Sec. 7.
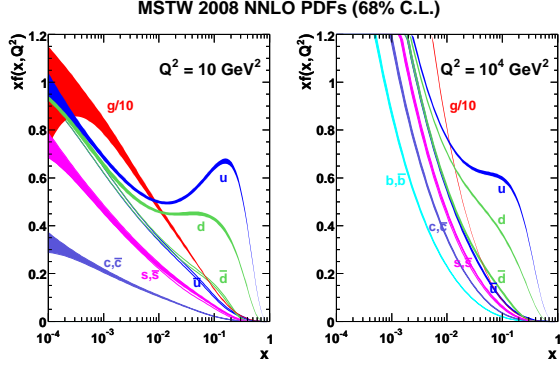
Figure 1: MSTW 2008 NNLO PDFs at two different Q^2 values [4].

Figure 3: $\Delta\chi^2_{\text{global}}$ as a function of (pole-mass) m_c [6].

LHC, $\sqrt{s}$ = 7 TeV	σ_W	σ_Z	σ_H
PDF-only uncertainty	+1.7% −1.6%	+1.7% −1.5%	+1.1% −1.6%
PDF+α_S uncertainty	+2.5% −1.9%	+2.5% −1.9%	+3.7% −2.9%
PDF+α_S+$m_{c,b}$ uncertainty	+2.7% −2.2%	+2.9% −2.4%	+3.7% −2.9%

Table 1: Impact of α_S and $m_{c,b}$ variation on LHC cross sections [6].

Figure 2: $\Delta\chi^2_{\text{global}}$ as a function of $\alpha_S(M_Z^2)$ [5].

2. Status of MSTW PDF analysis

The "MSTW 2008" determination of PDFs [4] at leading-order (LO), next-to-leading order (NLO) and next-to-next-to-leading order (NNLO) superseded the previously available "MRST" PDFs. New data sets fitted included neutrino structure functions (F_2 and xF_3) from NuTeV and CHORUS, neutrino dimuon cross sections from CCFR and NuTeV, HERA data on F_2^{charm} and on inclusive jet production in DIS, and Tevatron Run II data on inclusive jet production, the lepton charge asymmetry from W decays and the Z rapidity distribution. The CCFR/NuTeV dimuon cross sections allowed the strange-quark and -antiquark distributions to be fit directly for the first time. The Tevatron Run II jet data were found to prefer a softer gluon distribution at high x than the previous Run I data used in the MRST 2001–2006 fits. Uncertainties on the PDFs, shown in Fig. 1, were propagated from the experimental errors on the fit-

ted data points using a new dynamic procedure for each eigenvector of the covariance matrix. Subsequent studies used the same procedure to determine the experimental error on the best-fit $\alpha_S(M_Z^2)$, shown in Fig. 2 [5], and on the best-fit (pole-mass) m_c, shown in Fig. 3 [6].

Uncertainties in both $\alpha_S(M_Z^2)$ and the heavy-quark masses $m_{c,b}$ induce an additional uncertainty in cross-section predictions compared to the uncertainty arising only from PDFs. This increase in uncertainty is shown in Table 1 for the W, Z and $gg \rightarrow H$ (M_H = 120 GeV) NNLO total cross sections at the LHC with $\sqrt{s}$ = 7 TeV. Here, the PDF uncertainties are at 68% confidence-level (C.L.), and the parameters are varied in the ranges $\alpha_S(M_Z^2)$ = 0.1171 ± 0.0014, m_c = 1.40 ± 0.15 GeV and m_b = 4.75±0.25 GeV. Varying m_c leads to a change in $\sigma_{W,Z}$ of just over 1%, while varying m_b leads to a negligible change (0.1%) in $\sigma_{W,Z}$, and σ_H is insensitive to $m_{c,b}$ variation. Adding the $m_{c,b}$ uncertainty in quadrature to the "PDF+α_S" uncertainty does not result in a significant enhancement to the PDF+α_S uncertainty and it will therefore not be included in the remainder of this write-up. Further preliminary studies [7, 8] investigated problems in describing the precise $W \rightarrow \ell\nu$ charge asymmetry data at the Tevatron, and examined the impact of including the combined HERA I data [9] in the MSTW global fit, where the changes were not large enough to warrant an immediate update [8].

	MSTW08	**CT10**	**NNPDF2.1**	HERAPDF1.5	ABKM09	GJR08/JR09
HERA DIS	✔	✔	✔	✔	✔	✔
Fixed-target DIS	✔	✔	✔	✗	✔	✔
Fixed-target DY	✔	✔	✔	✗	✔	✔
Tevatron W,Z	✔	✔	✔	✗	✗	✗
Tevatron jets	✔	✔	✔	✗	✗	✔/✗
GM-VFNS	✔	✔	✔	✔	✗	✗
NNLO	✔	✗	✔	✔	✔	✔

Table 2: Comparison of major PDF sets considered, and their gross features distinguished by the main classes of data included (upper part of table) and important aspects of the theoretical treatment (lower part of table), specifically regarding the treatment of heavy quarks in DIS and the provision of NNLO PDFs. More refined differences between PDF sets are described in the text.

3. Benchmark exercise

Various fitting groups currently produce PDF sets: MSTW08 [4], CTEQ6.6/CT10 [10, 11], NNPDF2.1 [12, 13], HERAPDF1.0/1.5 [9, 14, 15], ABKM09 [16], GJR08/JR09 [17, 18, 19, 20]. Past experience has shown that results obtained with the different PDF sets often do not agree within the quoted uncertainties. Quantifying, understanding, then hopefully resolving differences in PDFs *between* groups is therefore as important, if not more important, as continued improvements in PDFs *within* groups. Some recent work has been initiated by the activities of the *LHC Higgs Cross Section Working Group* and the *PDF4LHC Working Group*. In particular, an exercise was proposed to use the most recent public NLO PDFs from all fitting groups to calculate some LHC benchmark processes at $\sqrt{s} = 7$ TeV, specifically total cross sections for production of $W^\pm$, Z^0, $t\bar{t}$ and $gg \rightarrow H$ for $M_H = \{120, 180, 240\}$ GeV. The aims were to establish the degree of compatibility and identify outliers amongst PDF sets, and to compare cross sections at the *same* α_S values, thereby showing to what extent differences in predictions are due to the different α_S values adopted by each group, rather than differences in the PDFs themselves. The results at NLO, initially presented by G.W. in a PDF4LHC meeting at CERN on 26th March 2010, formed the basis for the subsequent PDF4LHC *Interim Report* [21] and PDF4LHC *Interim Recommendations* [22] used in the *Handbook of LHC Higgs Cross Sections* [23]. An update of the NLO comparisons to include the recent CT10 [11] and NNPDF2.1 [12] analyses, and an extension of all comparisons to NNLO, was subsequently made in a later publication [1]. In this write-up we will make a further update to the comparison plots to include the even more recent NNPDF2.1 NNLO [13] and HERAPDF1.5 NNLO [15] analyses, and we will compare the predictions to the latest LHC data on the total cross sections for W, Z and $t\bar{t}$ production.

We will consider only *public* sets, defined to be those available for use with the latest LHAPDF V5.8.6 [24]. The broad distinctions between data sets fitted and aspects of the theoretical treatment are summarised in Table 2. Only three groups (MSTW, CT and NNPDF) make fully *global* fits to HERA and fixed-target DIS data, fixed-target Drell–Yan production, and Tevatron data on W, Z and jet production, although GJR08 includes all these processes other than Tevatron W and Z production. The HERAPDF1.0 fit includes *only* the combined HERA I inclusive data [9], while the HERAPDF1.5 fit additionally includes the preliminary combined HERA II inclusive data. The CT10 and NNPDF2.1 global fits include the combined HERA I inclusive data, while the other fits (MSTW08, CTEQ6.6, ABKM09, GJR08/JR09) include the older separate data from H1 and ZEUS. The MSTW08, CT10, NNPDF2.1 and GJR08 fits include Tevatron Run II data, while CTEQ6.6 uses only Tevatron Run I data. The original NNPDF2.1 fit has been reweighted to include Tevatron and LHC data on the $W \rightarrow \ell\nu$ charge asymmetry, denoted NNPDF2.2 [25], but this reweighted PDF set will not be considered here. Most groups now treat the heavy-quark contribution to DIS structure functions using a general-mass variable flavour number scheme (GM-VFNS), other than ABKM09 and GJR08/JR09 who use a fixed flavour number scheme (FFNS). The change from the inadequate zero-mass variable flavour number scheme (ZM-VFNS) to the GM-VFNS was the major improvement between NNPDF2.0 [26] and NNPDF2.1 [12], now allowing a meaningful comparison to other NLO global fits. The NNPDF fits parameterise the starting distributions at $Q_0^2 = 2$ GeV^2 as neural networks and use Monte Carlo methods for experimental error propagation. The other groups all use the more traditional approach of parameterising the input PDFs as some functional form in x, each with a handful of free parameters, and use the Hessian method for experimental error propagation
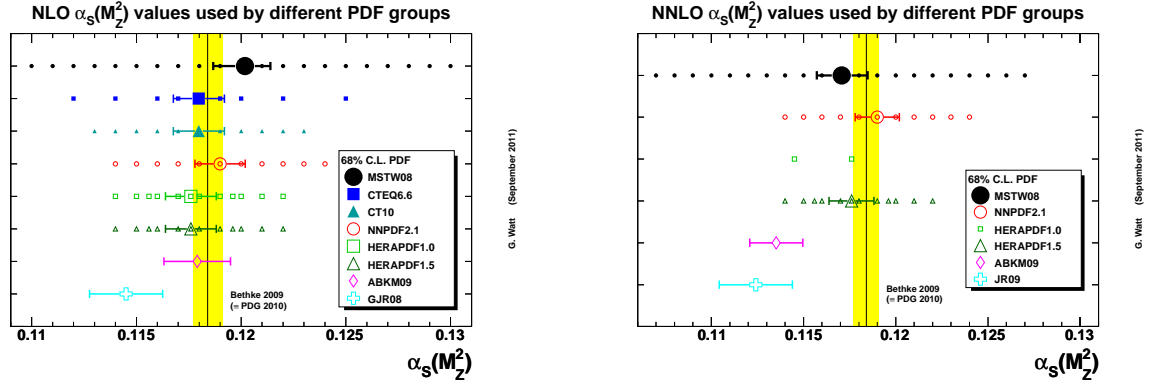
Figure 4: Values of $\alpha_S(M_Z^2)$, and their 1-σ uncertainties, used by different PDF fitting groups at NLO and NNLO. The smaller symbols indicate the PDF sets with alternative values of $\alpha_S(M_Z^2)$ provided by each fitting group. The shaded band indicates the Bethke 2009 world average $\alpha_S(M_Z^2)$.

with differing values of the tolerance $\Delta\chi^2$, that is, the change in the goodness-of-fit measure relative to the best-fit value. Contrary to the "standard" input parameterisation at $Q_0^2 \geq 1$ GeV^2, the GJR08/JR09 sets use a "dynamical" input parameterisation of valence-like input distributions at an optimally chosen $Q_0^2 < 1$ GeV^2, which gives a slightly worse fit quality and lower α_S values than the corresponding "standard" parameterisation, but is nevertheless favoured by the GJR08/JR09 authors. Public NNLO fits are available from MSTW08, NNPDF2.1, HERAPDF1.0/1.5, ABKM09 and JR09. (The first NNLO fits from the CT group should be available soon.) The Tevatron jet cross sections are excluded from the JR09 fit, where complete NNLO corrections are unavailable, whereas they are included in the MSTW08 and NNPDF2.1 NNLO fits by making the approximation of using the NLO partonic cross section supplemented by 2-loop threshold corrections [27].

In Fig. 4 we show the default values of $\alpha_S(M_Z^2)$ used by different fitting groups at NLO and NNLO, and we compare to the world average value obtained by S. Bethke in 2009 [28] (and taken over by the Particle Data Group in 2010 [29]). The values for MSTW08, ABKM09 and GJR08/JR09 are obtained from a simultaneous fit with the PDF parameters, while $\alpha_S(M_Z^2)$ for other groups is applied as an external constraint, generally chosen to be close to the world average [28], and for those groups we assume a 1-σ uncertainty of ±0.0012 [21]. The smaller symbols indicate the PDF sets with alternative values of $\alpha_S(M_Z^2)$ values provided by each fitting group. The fitted NLO $\alpha_S(M_Z^2)$ values are always larger than the fitted NNLO $\alpha_S(M_Z^2)$ values in an attempt by the fit to mimic the missing higher-order corrections, which are generally positive.
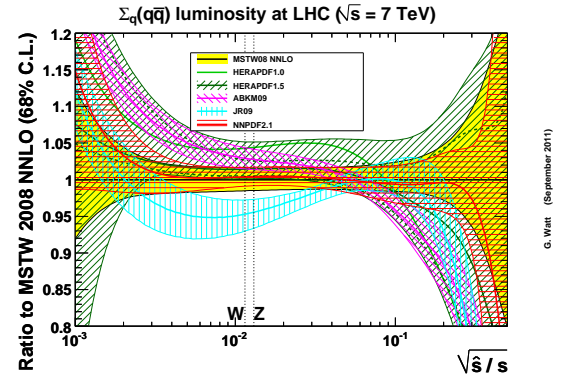

Figure 5: NNLO $q\bar{q}$ luminosities as the ratio to MSTW 2008.

This trend is repeated in the recent NNPDF determinations of $\alpha_S(M_Z^2) = 0.1191 \pm 0.0006$ at NLO [30] and $\alpha_S(M_Z^2) = 0.1173 \pm 0.0007$ at NNLO [31], where the experimental uncertainties are obtained using $\Delta\chi^2 = 1$, but these are not used as the default $\alpha_S(M_Z^2)$ values for the provided NNPDF2.1 PDFs.

4. *W* and *Z* production at the LHC

To understand properties of hadronic cross sections, such as PDF uncertainties or the dependence on collider energy, it is useful to consider the relevant parton–parton luminosities. We define a $q\bar{q}$ luminosity, relevant for the (rapidity-integrated) total cross sections for *W*
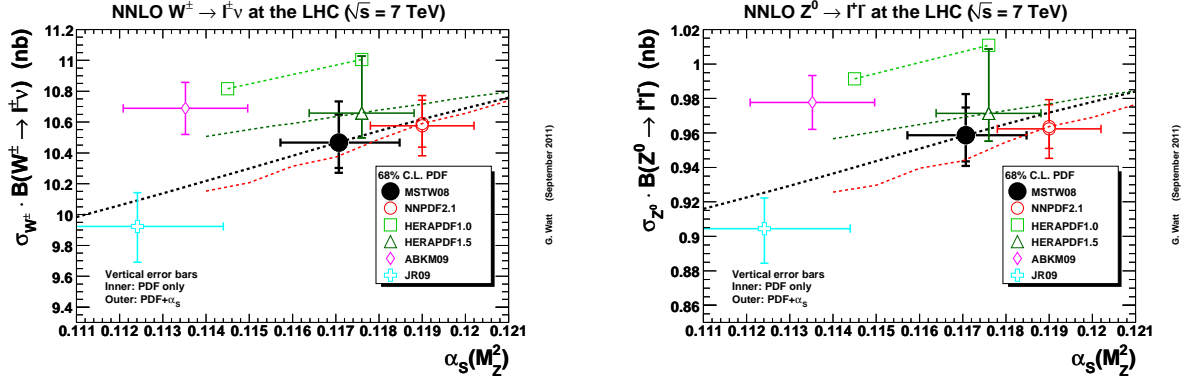
Figure 6: NNLO $W^\pm$ (= $W^+ + W^-$) and Z^0 total cross sections at the LHC, plotted as a function of $\alpha_S(M_Z^2)$.

and Z production at the LHC, as

$$\frac{\partial \mathcal{L}_{\Sigma_q(q\bar{q})}}{\partial \hat{s}} = \frac{1}{s} \int_\tau^1 \frac{dx}{x} \sum_{q=d,u,s,c,b} \left[q(x, \hat{s}) \bar{q}(\tau/x, \hat{s}) + (q \leftrightarrow \bar{q}) \right],$$

where $\tau \equiv \hat{s}/s$ and $\sqrt{\hat{s}}$ is the partonic centre-of-mass energy. Note that this generic $q\bar{q}$ luminosity does not specifically include the correct flavour combinations for $W^\pm$ or Z^0 production. More detailed studies would, for example, include the correct couplings of the vector bosons to quarks and antiquarks, or consider the specific $u\bar{d}$ (for W^+) or $d\bar{u}$ (for W^-) partonic luminosities.

In Fig. 5 we show the NNLO $q\bar{q}$ luminosities as the ratio with respect to the MSTW 2008 NNLO luminosities, for the LHC at $\sqrt{s} = 7$ TeV. We use the default α_S values for each set, shown in Fig. 4. The HER-APDF1.0 NNLO curve (without uncertainties) is for $\alpha_S(M_Z^2) = 0.1176$. The inner uncertainty bands (dashed lines) for HERAPDF1.5 correspond to the (asymmetric) experimental errors, while the outer uncertainty bands (shaded regions) also include the model and parameterisation errors, including uncertainties on heavy-quark masses but not on α_S. It is not possible to separate the "PDF only" uncertainty for ABKM09 and JR09, therefore the uncertainty bands for those sets also include the α_S uncertainty, and the uncertainty bands for ABKM09 also include uncertainties on heavy-quark masses. This is undesirable but unavoidable given that these groups do not provide PDF sets for fixed α_S (and fixed $m_{c,b}$ for ABKM09). The relevant values of $\sqrt{\hat{s}} = M_{W,Z}$ are indicated, and there is good agreement for the two global fits (MSTW08 and NNPDF2.1), but more variation for the other sets. The NNLO trend between groups is similar to at NLO [1, 3], with the exception of HERAPDF at large $\hat{s}$ values, where the HERAPDF NLO sets have a

much larger $q\bar{q}$ luminosity than other NLO PDF groups.

In Fig. 6 we show the $W^\pm$ (= $W^+ + W^-$) and Z^0 total cross section multiplied by the appropriate leptonic branching ratios, $B(W^\pm \to \ell^\pm \nu)$ or $B(Z^0 \to \ell^+ \ell^-)$, calculated at NNLO [32] with a scale choice $\mu_R = \mu_F = M_{W,Z}$, plotted as a function of $\alpha_S(M_Z^2)$. The markers are centred on the default $\alpha_S(M_Z^2)$ value and the corresponding predicted cross section of each PDF fitting group. The horizontal error bars span the $\alpha_S(M_Z^2)$ uncertainty, the inner vertical error bars span the "PDF only" uncertainty where possible (i.e. not for ABKM09 or JR09), and the outer vertical error bars span the "PDF+α_S" uncertainty. The effect of the additional α_S uncertainty is small for W and Z production. The dashed lines interpolate the cross-section predictions calculated with the alternative PDF sets with different $\alpha_S(M_Z^2)$ values provided by each group, represented by the smaller symbols in Fig. 4. The two global fits (MSTW08 and NNPDF2.1) are in good agreement, as was apparent from the $q\bar{q}$ luminosity plot in Fig. 5. The central HER-APDF1.5 prediction is close to the global fits at NNLO, contrary to the predictions using HERAPDF1.0 NNLO or HERAPDF1.0/1.5 NLO [1, 3].

The total cross-section *ratios*, $W^\pm/Z^0$ and W^+/W^-, are insensitive to NNLO corrections and the value of $\alpha_S(M_Z^2)$. The $W^\pm$ (= $W^+ + W^-$) and Z^0 total cross sections are highly correlated, which can be understood by considering the dominant partonic contributions arising from u and d quarks, i.e.

$$\frac{\sigma_{W^+} + \sigma_{W^-}}{\sigma_{Z^0}} \sim \frac{u(x_1) + d(x_1)}{0.29 \, u(\tilde{x}_1) + 0.37 \, d(\tilde{x}_1)}, \quad (1)$$

where we have neglected the contributions with $q \leftrightarrow \bar{q}$, assuming that $q(x_1)\bar{q}(x_2)$ dominates over $\bar{q}(x_1)q(x_2)$, and the numerical values in the denominator are the
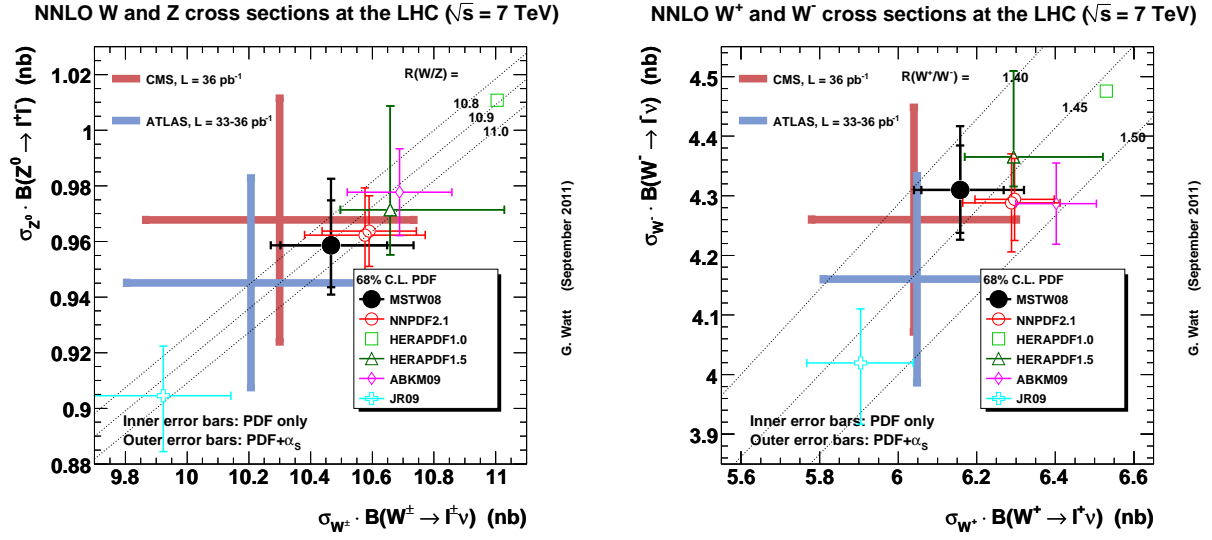
Figure 7: $W^{\pm}$ vs. Z^0 and W^+ vs. W^- total cross sections at NNLO, compared to data from CMS [33] and ATLAS [34].

appropriate sums of the squares of the vector and axial-vector couplings. We have also assumed that $\bar{u}(x_2) \approx \bar{d}(x_2)$. Here, the momentum fractions are $x_{1,2} = (M_W/\sqrt{s})\exp(\pm y_W)$ and $\tilde{x}_{1,2} = (M_Z/\sqrt{s})\exp(\pm y_Z)$, and y_W or y_Z should be interpreted as some "average" rapidity appropriate for the (rapidity-integrated) total cross section. The combination of u- and d-quark contributions is very similar (in numerical prefactors, in x values, and in $Q^2 = M_{W,Z}^2$ values) in both the numerator and denominator of Eq. (1), therefore the PDF dependence almost cancels in the $W^{\pm}/Z^0$ ratio. The W^+ and W^- total cross sections are much less correlated, since

$$\frac{\sigma_{W^+}}{\sigma_{W^-}} \sim \frac{u(x_1)\,\bar{d}(x_2)}{d(x_1)\,\bar{u}(x_2)} \sim \frac{u(x_1)}{d(x_1)},$$

and therefore the W^+/W^- cross-section ratio is a sensitive probe of the u/d ratio and there is more variation between different PDF sets.

In Fig. 7 we show $W^{\pm}$ (= $W^+ + W^-$) versus Z^0 and W^+ versus W^- total cross sections, and we draw dotted lines where the ratio of cross sections, $W^{\pm}/Z^0$ or W^+/W^-, is constant. We also compare to the experimental measurements using the 2010 LHC data from CMS [33] and ATLAS [34]. Here, the measured Z^0 cross sections have been corrected [1] for the small γ^* contribution and the finite invariant-mass range of the lepton pair (different for ATLAS and CMS) using a theory calculation at NNLO [35]. The spread in predictions using the different PDF sets is comparable to the (dominant) luminosity uncertainty of 4% (CMS) or 3.4% (ATLAS), and

perhaps the two extreme predictions (HERAPDF1.0 and JR09) are somewhat disfavoured. In Fig. 8 we show the same plots with ellipses drawn to account for the correlations between the two cross sections, both for the experimental measurements and for the theoretical predictions. Here, the ellipses are defined such that the projection onto either axes gives the 1-σ uncertainty for the individual cross sections, meaning that the area of the two-dimensional ellipse corresponds to a confidence-level somewhat smaller than the conventional 68%. The largest uncertainty in the ATLAS and CMS total cross-section ratios comes from extrapolating the measured (fiducial) cross section over the whole phase space. Indeed, improvements in the acceptance calculation led to the central value of the ATLAS W^+/W^- total cross-section ratio shifting from the preliminary result of 1.51 in March 2011 [36] to 1.45 in the final publication [34] following observations made in Ref. [1]. Therefore, data-to-theory comparisons are best made at the level of the fiducial cross section (i.e. within the acceptance), which is now possible using the public FEWZ [37, 38] and DYNNLO [39] codes, and indeed was done in the ATLAS publication [34].

Of course, as a constraint on PDFs, differential distributions are more useful than total or fiducial cross sections. The $W^{\pm}$ charge asymmetry is sensitive to the difference between up- and down-valence quark distributions, i.e.

$$A_W(y_W) \quad = \quad \frac{\mathrm{d}\sigma(W^+)/\mathrm{d}y_W - \mathrm{d}\sigma(W^-)/\mathrm{d}y_W}{\mathrm{d}\sigma(W^+)/\mathrm{d}y_W + \mathrm{d}\sigma(W^-)/\mathrm{d}y_W}$$
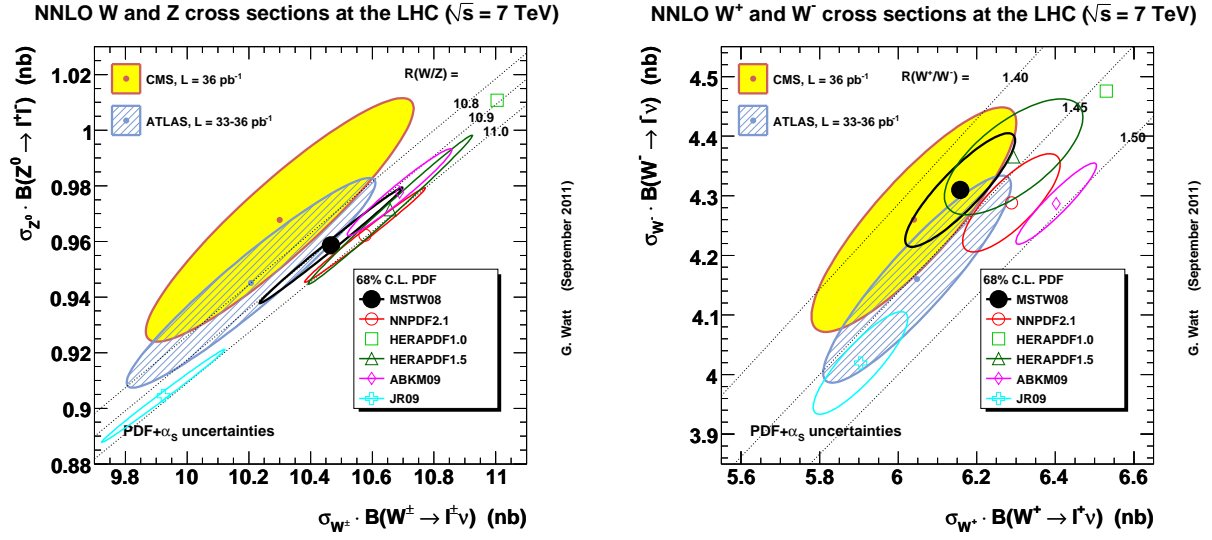
Figure 8: Same as Fig. 7, but with ellipses accounting for the correlations between the two cross sections.

$$\approx \quad \frac{u_v(x_1) - d_v(x_1)}{u(x_1) + d(x_1)}.$$

Experimentally, it is not possible to directly reconstruct the W rapidity since the longitudinal momentum of the decay neutrino is unknown, so instead the $W^\pm \to \ell^\pm \nu$ charge asymmetry is measured as a function of the charged-lepton pseudorapidity η_ℓ, i.e.

$$
\begin{aligned}
A_\ell(\eta_\ell) &= \frac{\mathrm{d}\sigma(\ell^+)/\mathrm{d}\eta_\ell - \mathrm{d}\sigma(\ell^-)/\mathrm{d}\eta_\ell}{\mathrm{d}\sigma(\ell^+)/\mathrm{d}\eta_\ell + \mathrm{d}\sigma(\ell^-)/\mathrm{d}\eta_\ell} \\
&\equiv A_W(y_W) \otimes (W^\pm \to \ell^\pm \nu),
\end{aligned}
$$

where the last line is symbolic and indicates the underlying $W^\pm$ charge asymmetry convoluted with the $W^\pm \to \ell^\pm \nu$ decay. Measurements of $A_\ell(\eta_\ell)$ have provided the first useful PDF constraint from LHC data, and indicate that $A_\ell(0)$ is underestimated by the MSTW08 fit and is better described by some other PDF sets, implying that $u_v - d_v$ is too small at $x \sim M_W/\sqrt{s} \sim 0.01$. This behaviour may be traced to the independent small-x powers, $xu_v \propto x^{0.29 \pm 0.02}$ and $xd_v \propto x^{0.97 \pm 0.11}$, at the input scale $Q_0^2 = 1$ GeV^2 for the MSTW08 NLO fit, compared to some other groups (e.g. CTEQ6.6/CT10) where the small-x powers of u_v and d_v are assumed to be equal. On the other hand, this implies some tension between the LHC $W \to \ell\nu$ charge asymmetry data and the data already included in the MSTW08 fit (e.g. the Tevatron $W \to \ell\nu$ asymmetry, the NMC F_2^d/F_2^p ratio, and the E866/NuSea Drell–Yan σ^{pd}/σ^{pp} ratio). Other tensions have been observed with the precise Tevatron

data on the $W \to \ell\nu$ charge asymmetry, and partially resolved by more flexible nuclear corrections for deuteron structure functions [8]. Further attempts to resolve these tensions will be necessary for any future update of the MSTW08 fit. The ATLAS Collaboration provide differential cross sections for $W^+ \to \ell^+ \nu$ and $W^- \to \ell^- \bar\nu$ with the complete information on correlated systematic uncertainties, which is potentially more useful for PDF fits than simply $A_\ell(\eta_\ell)$, and the individual $W^+ \to \ell^+ \nu$ and $W^- \to \ell^- \bar\nu$ cross sections seem to be better described by the MSTW08 fit than many other PDF sets [34].

Currently, perhaps the least well-known parton distributions are the strange-quark and -antiquark distributions, where the only experimental information comes from the CCFR/NuTeV dimuon cross sections in the region $0.01 \lesssim x \lesssim 0.2$. Although these data are included in the MSTW08, CT10 and NNPDF2.1 analysis, the three groups obtain quite different results for the s and $\bar s$ distributions in the data region. These differences are not well understood, but may be due to issues such as the acceptance calculation, different treatment of nuclear corrections, or differences in the treatment of charm production in charged-current DIS. In particular, the NNPDF2.1 fit includes the contribution initiated by the charm-quark PDF, leading to a smaller strange-quark PDF compared to the MSTW08 and CT10 fits, which do not include this contribution. In the x region outside of the CCFR/NuTeV data, the s and $\bar s$ distributions are largely determined by the assumed parameterisation (with the possible exception of NNPDF). A com-

plementary, and perhaps theoretically cleaner, process to probe strangeness at the LHC is W production with an associated charm-tagged jet. The dominant partonic subprocesses are $\bar{s}\,g \to W^+\,\bar{c}$ and $s\,g \to W^-\,c$, with a small Cabibbo-suppressed contribution from $\bar{d}\,g \to W^+\,\bar{c}$ (5%) and $d\,g \to W^-\,c$ (15%). A first preliminary measurement has been made by CMS [40] of the cross-section ratios $R_c \equiv \sigma(W+c)/\sigma(W+\text{jets})$, probing the strange content of the proton relative to other light-quark flavours, and $R_c^\pm \equiv \sigma(W^+ + \bar{c})/\sigma(W^- + c)$, potentially probing the strange asymmetry. The current CMS data [40] do not strongly discriminate between current PDF sets, although the measurement of R_c is in slightly better agreement with MSTW08 and CT10 and is more than 1-σ above the NNPDF2.1 prediction. With more precise measurements to come, including differential distributions, the W+charm process should enable powerful constraints to be made on the s and $\bar{s}$ distributions.

5. Higgs, top-pair and jet production

Whereas the cross sections for production of W and Z bosons are sensitive to the *quark* distributions, we now turn to processes that are sensitive to the *gluon* distribution. Of course, one of the main goals of the initial LHC physics programme is to either discover or exclude the Standard Model (SM) Higgs boson (H). This requires precise knowledge of the theoretical cross section, where the dominant production mechanism at both the Tevatron and LHC is gluon–gluon fusion ($gg \to H$) through a top-quark loop, with the gluon distribution being sampled at $x \sim M_H/\sqrt{s}$. The exclusion limits at 95% C.L. from December 2011 [41, 42, 43] are given in Table 3, together with the relevant x values probed.

We will calculate the total cross section (σ_H) for SM Higgs boson production (without decay) from gluon–gluon fusion via a top-quark loop, with a fixed scale choice of $\mu_R = \mu_F = M_H$. The m_t dependence is retained only at LO, with $m_t = 171.3$ GeV (PDG 2009 best value), and the higher-order corrections are calcu-

	M_H (GeV)	$x \sim M_H/\sqrt{s}$
Tevatron	156 – 177	0.08 – 0.09
ATLAS	112.7 – 115.5	0.02 – 0.02
	131 – 237	0.02 – 0.03
	251 – 453	0.04 – 0.06
CMS	127 – 600	0.02 – 0.09

Table 3: Exclusion limits at 95% C.L. for the SM Higgs boson (as of December 2011) [41, 42, 43] and the approximate x values probed.

lated in the limit of an infinite top-quark mass, with NNLO corrections from Ref. [44]. We do not include the small bottom-quark loop contributions to the $gg \to H$ cross section. The size of the higher-order corrections to the $gg \to H$ total cross sections is substantial. Taking the appropriate MSTW08 PDFs and α_S values consistently at each perturbative order for σ_H with $M_H = 160$ GeV, then the NLO/LO ratio is 2.1 (Tevatron) or 1.9 (LHC), the NNLO/LO ratio is 2.7 (Tevatron) or 2.4 (LHC), and so the NNLO/NLO ratio is 1.3 (Tevatron and LHC). The perturbative series is therefore slowly convergent, mandating the use of (at least) NNLO calculations together with the corresponding NNLO PDFs and α_S values. The convergence can be improved by using a scale choice $\mu_R = \mu_F = M_H/2$, which mimics the effect of soft-gluon resummation. However, we aim to study only the PDF and α_S dependence of the $gg \to H$ cross sections, and we do *not* aim to come up with a single "best" prediction together with a complete evaluation of all sources of theoretical uncertainty.
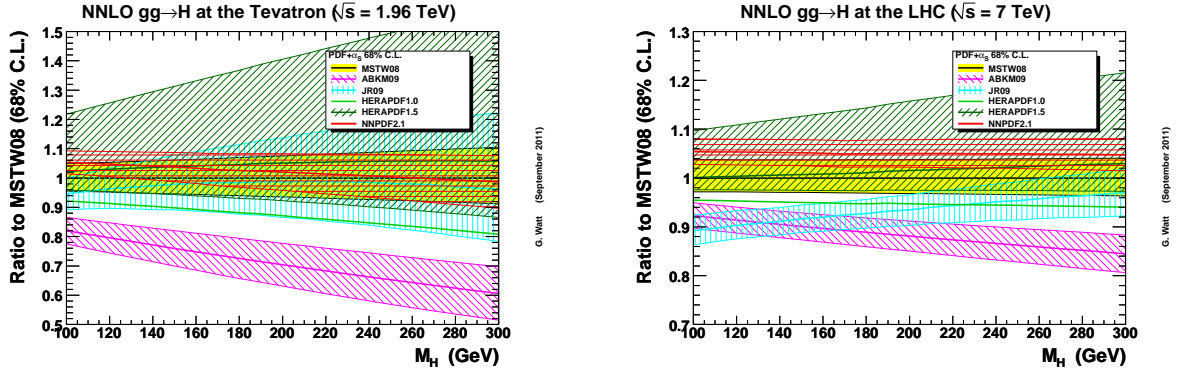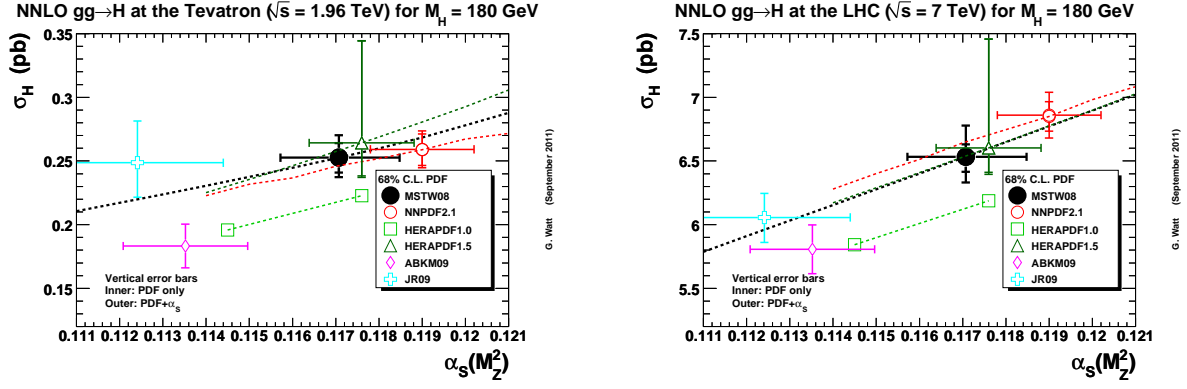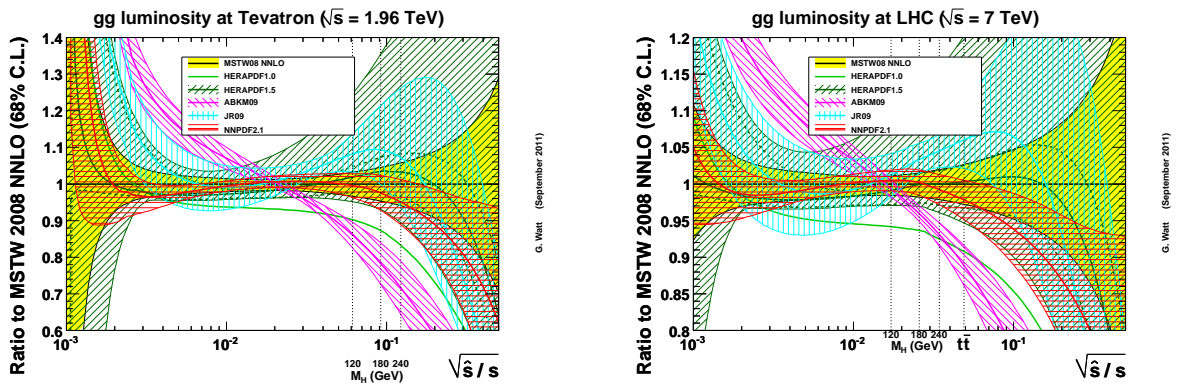
The ratios of the NNLO $gg \to H$ cross sections with respect to the MSTW08 predictions, plotted against the SM Higgs mass M_H, are shown for the Tevatron and LHC in Fig. 9, where PDF+α_S uncertainty bands at 68% C.L. are plotted. It can be seen that there is good agreement for the two global fits (MSTW08 and NNPDF2.1) at NNLO. The central value of HERA-PDF1.5 is also in good agreement, but it has a very large uncertainty in the upwards direction, and we will return to this feature later. The HERAPDF1.0 prediction with $\alpha_S(M_Z^2) = 0.1176$ lies somewhat below MSTW08, NNPDF2.1 and HERAPDF1.5. However, the ABKM09 prediction lies even further below MSTW08 at the LHC, and especially at the Tevatron, even allowing for the PDF+α_S uncertainties, by an amount much larger than the scale uncertainty (~10%). The Tevatron and LHC exclusion bounds are based on predictions using the MSTW08 PDFs, and the decision not to use *all* available NNLO PDFs has drawn some misguided criticism from certain quarters [45, 46], particularly before the LHC exclusion results became available due to the more significant discrepancies between PDF sets for Tevatron predictions.

The $gg \to H$ cross sections at the Tevatron and LHC start at $O(\alpha_S^2)$ at LO, with anomalously large higher-order corrections, therefore they are directly sensitive to the value of $\alpha_S(M_Z^2)$. Moreover, there is a known correlation between the value of α_S and the gluon distribution, which additionally affects the $gg \to H$ cross sections. In Fig. 10 we show this sensitivity by plotting the Higgs cross sections versus $\alpha_S(M_Z^2)$ at the Tevatron and LHC for a Higgs mass $M_H = 180$ GeV. The for-

Figure 9: Ratio to MSTW 2008 NNLO $gg \rightarrow H$ total cross section, plotted as a function of M_H, with PDF+α_S uncertainty bands at 68% C.L.

Figure 10: NNLO $gg \rightarrow H$ total cross sections, plotted as a function of $\alpha_S(M_Z^2)$, for $M_H = 180$ GeV.

Figure 11: NNLO gluon–gluon luminosities as the ratio with respect to MSTW 2008, plotted as a function of $\sqrt{\hat{s}/s}$.
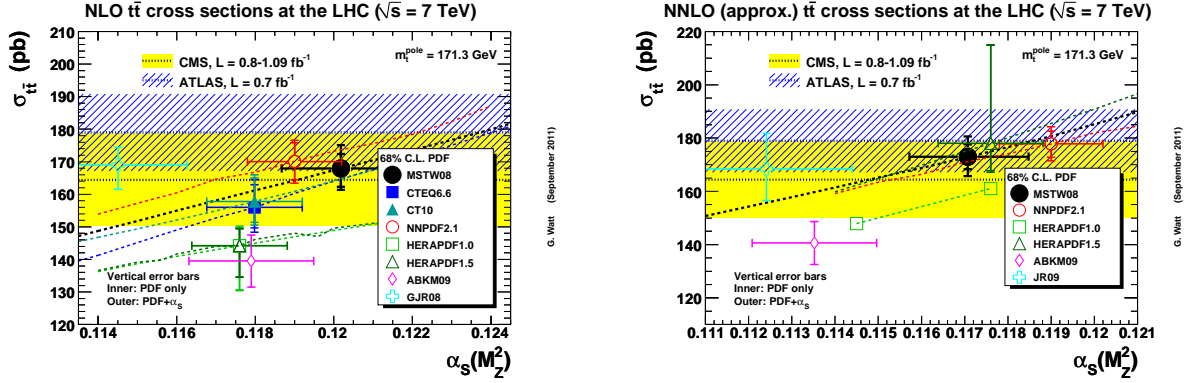
Figure 12: NLO and NNLO (approx.) $t\bar{t}$ total cross sections at the LHC, plotted as a function of $\alpha_S(M_Z^2)$, for $m_t = 171.3$ GeV, and compared to the single most precise current LHC measurements from CMS [47] and ATLAS [48].

mat of the plots is the same as in Fig. 6, but in this case the effect of the additional α_S uncertainty is more sizeable. It is apparent from the plots that at least part of the MSTW08/ABKM09 discrepancy for Higgs cross sections is due to using quite different values of $\alpha_S(M_Z^2)$ at NNLO, specifically $\alpha_S(M_Z^2) = 0.1135 \pm 0.0014$ for ABKM09 [16] compared to $\alpha_S(M_Z^2) = 0.1171 \pm 0.0014$ for MSTW08 [4, 5]. Comparing cross-section predictions at the same value of $\alpha_S(M_Z^2)$ would reduce the MSTW08/ABKM09 discrepancy at the LHC, but there would still be a significant discrepancy at the Tevatron (see also the later Table 7).

At LO, the PDF dependence of the $gg \rightarrow H$ total cross section is simply given by the gluon–gluon luminosity evaluated at a partonic centre-of-mass energy $\sqrt{\hat{s}} = M_H$, i.e.

$$\frac{\partial \mathcal{L}_{gg}}{\partial \hat{s}} = \frac{1}{s} \int_\tau^1 \frac{\mathrm{d}x}{x} g(x, \hat{s}) g(\tau/x, \hat{s}),$$

where $g(x, \mu^2 = \hat{s})$ is the gluon distribution and $\tau \equiv \hat{s}/s$. In Fig. 11 we show the gluon–gluon luminosities calculated using different PDF sets and taken as the ratio with respect to the MSTW 2008 NNLO value, at centre-of-mass energies corresponding to the Tevatron and LHC. The relevant values of $\sqrt{\hat{s}} = M_H = \{120, 180, 240\}$ GeV are indicated, along with the threshold for $t\bar{t}$ production at the LHC, $\sqrt{\hat{s}} = 2m_t$ with $m_t = 171.3$ GeV, where this process is predominantly gg-initiated at the LHC. Indeed, $t\bar{t}$ production at the LHC is strongly correlated with $gg \rightarrow H$ production at the Tevatron, with both processes probing the gluon distribution at similar x values, as seen from Fig. 11. There is reasonable agreement for the global fits (MSTW08 and NNPDF2.1), but more variation for the other sets, particularly at large $\hat{s}$, where

the HERAPDF1.0 set with $\alpha_S(M_Z^2) = 0.1176$, and especially the ABKM09 set, has a much softer high-x gluon distribution, and this feature has a direct impact on the $gg \rightarrow H$ cross sections, particularly at the Tevatron (see Fig. 9). Again, we note that the central value of HERAPDF1.5 is in good agreement with the global fits, but it has a very large uncertainty in the upwards direction, and we will return to this feature later.

More than 80% of the NLO $t\bar{t}$ cross section comes from the gg channel for the LHC with $\sqrt{s} = 7$ TeV, rising to almost 90% at $\sqrt{s} = 14$ TeV, compared to less than 15% at the Tevatron ($\sqrt{s} = 1.96$ TeV). The significant difference in the initial parton composition for $t\bar{t}$ production is due partly to the lower Tevatron energy (pp collisions at $\sqrt{s} = 1.96$ TeV would give around 50% of the $t\bar{t}$ cross section from the gg channel), but mainly due to the valence–valence nature of the $q\bar{q} \rightarrow t\bar{t}$ channel in $p\bar{p}$ collisions. The partonic subprocess is $O(\alpha_S^2)$ at LO. There is therefore a strong dependence on both the gluon distribution (at $x \sim 2m_t/\sqrt{s} = 0.05$) and α_S. We calculate $t\bar{t}$ production (without decay) for a top-quark pole mass $m_t = 171.3$ GeV (PDG 2009 best value), with a fixed scale choice of $\mu_R = \mu_F = m_t$. We show the $t\bar{t}$ total cross sections at the LHC ($\sqrt{s} = 7$ TeV), plotted as a function of $\alpha_S(M_Z^2)$, in Fig. 12, with 68% C.L. PDF+α_S uncertainties. The NNLO calculation of the total cross section for $t\bar{t}$ production is still in progress, although various approximations based on threshold resummation are available. We use the HATHOR [49] public code with the default settings for an approximate "NNLO" calculation, although we make no attempt to quantify the theoretical uncertainty (other than from PDFs and α_S). A more complete study of the theoretical uncertainties in the approx-
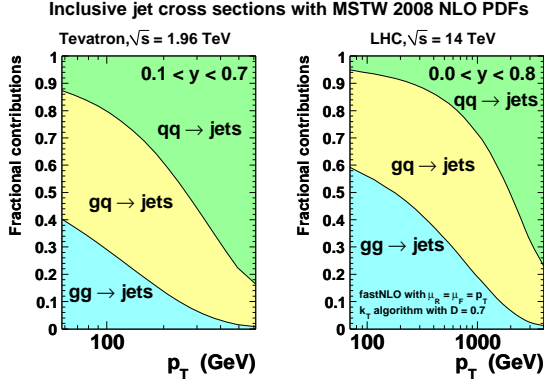
Figure 13: Fractional contributions of the *gg*-, *gq*- and *qq*-initiated processes to inclusive jet production as a function of p_T [5].
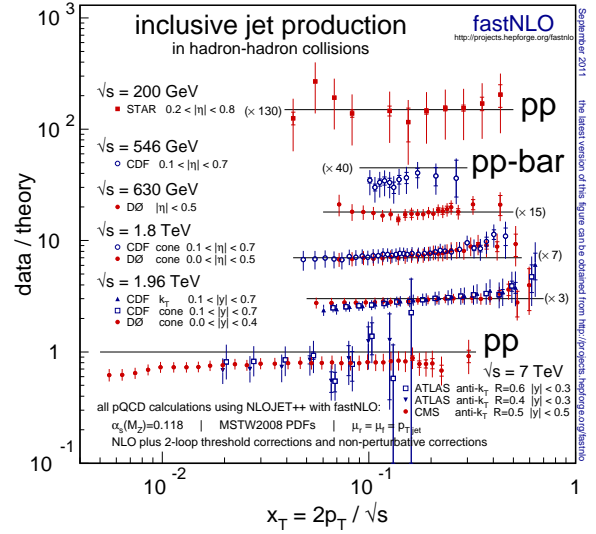

Figure 14: Ratios of data over theory (using MSTW 2008 PDFs) for inclusive jet cross sections at hadron colliders as a function of the scaling variable $x_T = 2p_T / \sqrt{s}$. Plot taken from Ref. [54].

imate NNLO calculation is clearly important, but it is beyond the scope of this work (see Refs. [50, 51, 52] for some recent studies). As a rough indication, the scale uncertainties are estimated to be 13% at NLO and significantly smaller in the approximate NNLO calculations (for example, 8% at NLO+NNLL [52]). The predicted $t\bar{t}$ cross section has a fairly strong dependence on the assumed top-quark mass m_t, such that comparison of the measured cross section with theory predictions even allows an extraction of m_t. As some indication of the m_t dependence, increasing m_t by 2 GeV to give a value close to the current Tevatron average of $m_t = 173.2 \pm 0.9$ GeV [53] decreases the $t\bar{t}$ cross section at the 7 TeV LHC by about 10 pb (or 6%) at both NLO and NNLO with MSTW08 PDFs. Bearing these caveats in mind, we compare to the single most precise current CMS measurement (e/μ+jets+b-tag) of [47]

$$\sigma_{t\bar{t}} = 164.4 \pm 2.8(\text{stat.}) \pm 11.9(\text{syst.}) \pm 7.4(\text{lumi.}) \text{ pb},$$

which is close to a recent CMS combination of $t\bar{t}$ cross-section measurements giving [55]

$$\sigma_{t\bar{t}} = 165.8 \pm 2.2(\text{stat.}) \pm 10.6(\text{syst.}) \pm 7.8(\text{lumi.}) \text{ pb},$$

and to the single most precise current ATLAS measurement (using kinematic information of lepton+jets events) of [48]

$$\sigma_{t\bar{t}} = 179.0 \pm 9.8(\text{stat. + syst.}) \pm 6.6(\text{lumi.}) \text{ pb}.$$

The approximate NNLO prediction using MSTW08 PDFs shown in Fig. 12 is consistent with both the ATLAS and CMS measurements, while the central value using ABKM09 is almost 2-σ below CMS and more than 3-σ below ATLAS. The discrepancy for ABKM09

would increase further if using the more up-to-date value of $m_t = 173.2 \pm 0.9$ GeV [53] rather than $m_t = 171.3$ GeV used in the plots, which would reduce all theory predictions by around 6%. The HERAPDF1.0/1.5 sets at NLO are also disfavoured by the LHC data, as is the HERAPDF1.0 NNLO set with $\alpha_S(M_Z^2) = 0.1145$.

It would be a worrying situation if the $t\bar{t}$ cross section at the LHC was needed to discriminate between PDF sets. The measured $t\bar{t}$ cross section is commonly used to constrain new physics contributions, therefore it is questionable whether it should be used directly to constrain PDFs. Rather, we would hope that the gluon distribution (and α_S) would be sufficiently constrained by other data sets that the $t\bar{t}$ cross section is a *prediction* rather than a direct PDF constraint. If *only* the ABKM09 PDF set was available, one can only imagine what new physics scenarios could be conjured up to explain the "excess" of data over theory seen in Fig. 12.

Measurements of the scaling violations of DIS structure functions can be used to constrain the small-x gluon, although there is no direct constraint on the large-x gluon from inclusive DIS. To constrain the high-x gluon distribution we need to look for processes where the gluon appears in the initial state at LO, and the best example is inclusive jet production at hadron colliders. In Fig. 13 we show the composition of the initial state as a function of the jet transverse momentum. There is
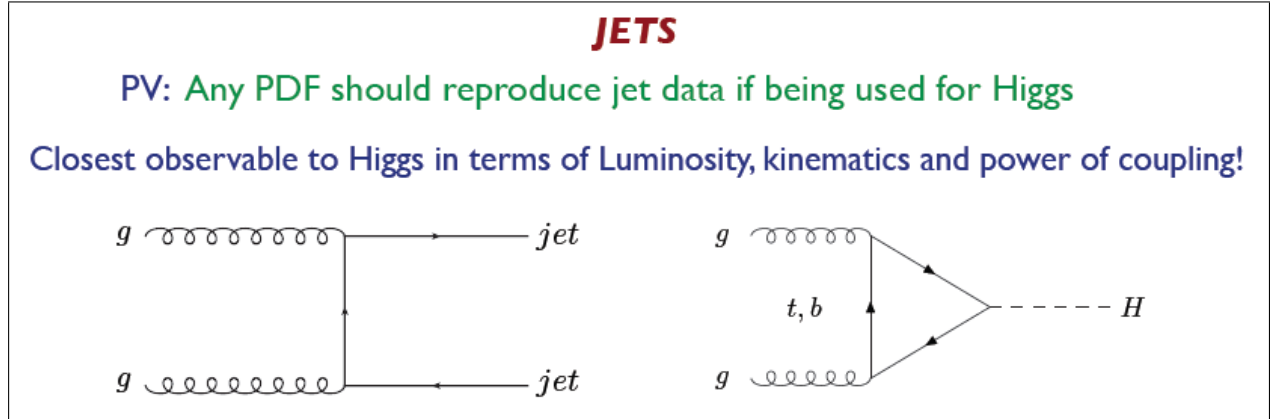
Figure 15: Jets as a discriminator of the high-x gluon distribution (extract from talk by D. de Florian [56]).

a transition from gluon–gluon and gluon–quark at low p_T values to quark–quark at high p_T values, but even at high p_T values the gluon–quark contribution is still significant. (Here, we do not distinguish between quarks and antiquarks.) Since the quark distributions can be constrained by other data sets, it means that jet production mainly constrains the gluon distribution. The plot in Fig. 14, taken from Ref. [54], shows data plotted as a function of the scaling variable $x_T = 2p_T/\sqrt{s}$, which is related to the x values probed at central rapidity. The current LHC data are generally at lower x_T than at the Tevatron, so presently the best constraint on the high-x gluon comes from the Tevatron jet data. Indeed, other than the $t\bar{t}$ cross section, jets are perhaps the closest observable to Higgs in terms of the partonic luminosity, kinematics and power of coupling; see Fig. 15 taken from Ref. [56]. The crucial question we must address, therefore, is how well do the "non-global" fits describe the Tevatron jet data?

We will not consider the less reliable Tevatron Run I data, which prefer a much harder high-x gluon distribution [4], and are obtained using less sophisticated jet algorithms. The three data sets on inclusive jet production from the Tevatron Run II [57, 58, 59] were all found to be compatible [4]. The MSTW 2008 analysis [4] included the CDF Run II inclusive jet data using the k_T jet algorithm [57] and the DØ Run II inclusive jet data using a cone jet algorithm [59]. Consistency was checked with the CDF Run II inclusive jet data using the cone-based Midpoint jet algorithm [58], but this data set was not included in the final MSTW08 fit, since it is essentially the same measurement (using 1.13 fb^{-1}) as Ref. [57] (using 1.0 fb^{-1}), differing mainly by the choice of jet algorithm. The k_T jet algorithm is theoretically preferred due to its property of infrared safety.

We therefore focus here on the CDF Run II inclusive jet data using the k_T jet algorithm [57], and the description using NNLO PDFs, but the other Tevatron Run II jet data sets and the description using NLO PDFs are considered in detail in Ref. [2].

One obvious problem is that the complete NNLO partonic cross section ($\hat{\sigma}$) for inclusive jet production is currently unknown, and needs to be approximated with the NLO $\hat{\sigma}$ supplemented by 2-loop threshold corrections [27]. We calculate jet cross sections using FASTNLO [60] (based on NLOJET++ [61, 62]), which includes these 2-loop threshold corrections. Following the usual way of estimating theoretical uncertainties due to unknown higher-order corrections, we take different scale choices $\mu_R = \mu_F = \mu = \{p_T/2, p_T, 2p_T\}$ as some indication of the theoretical uncertainty. Smaller scale choices raise the partonic cross section, so favour softer high-x gluon distributions [4], and the central $\mu = p_T$ was chosen for the final MSTW08 fit [4]. More comments on the scale dependence are given in Ref. [2].

It is important to account for *correlated* systematic uncertainties of the experimental data points. The full correlated error information is accounted for by using a goodness-of-fit (χ^2) definition given by [63, 64]

$$\chi^2 = \sum_{i=1}^{N_{\text{pts.}}} \left(\frac{\hat{D}_i - T_i}{\sigma_i^{\text{uncorr.}}}\right)^2 + \sum_{k=1}^{N_{\text{corr.}}} r_k^2, \qquad (2)$$

where T_i are the theory predictions and $\hat{D}_i \equiv D_i - \sum_{k=1}^{N_{\text{corr.}}} r_k\,\sigma_{k,i}^{\text{corr.}}$ are the data points allowed to shift by the systematic errors in order to give the best fit. Here, $i = 1, \ldots, N_{\text{pts.}}$ labels the individual data points and $k = 1, \ldots, N_{\text{corr.}}$ labels the individual correlated systematic errors. The data points D_i have uncorrelated (statistical and systematic) errors $\sigma_i^{\text{uncorr.}}$ and correlated systematic errors $\sigma_{k,i}^{\text{corr.}}$. The optimal shifts of the data

NNLO PDF	$\alpha_S(M_Z^2)$	$\mu = p_T/2$	$\mu = p_T$	$\mu = 2p_T$
MSTW08	0.1171	1.39 (+0.35)	0.69 (−0.45)	0.97 (−*1.30*)
NNPDF2.1	0.1190	0.68 (−0.77)	0.71 (−*2.02*)	0.71 (−**3.46**)
HERAPDF1.0	0.1145	2.37 (−**2.65**)	1.48 (−**3.64**)	1.29 (−**4.12**)
HERAPDF1.0	0.1176	2.24 (−0.48)	1.13 (−*1.60*)	1.09 (−*2.23*)
HERAPDF1.5	0.1176	1.61 (+*1.22*)	0.77 (+0.30)	1.06 (−0.39)
ABKM09	0.1135	1.53 (−**4.27**)	1.23 (−**5.05**)	1.44 (−**5.65**)
JR09	0.1124	0.75 (+0.13)	1.26 (−0.61)	2.20 (−*1.22*)

Table 4: Values of $\chi^2/N_{\text{pts.}}$ for the CDF Run II inclusive jet data using the k_T jet algorithm [57] with $N_{\text{pts.}} = 76$ and $N_{\text{corr.}} = 17$, for different NNLO PDF sets and different scale choices $\mu_R = \mu_F = \mu = \{p_T/2, p_T, 2p_T\}$. No restriction is imposed on the shift in normalisation and the optimal value of "$-r_{\text{lumi.}}$" is shown in brackets, where the data points are shifted as $D_i \to D_i(1 - 0.058\,r_{\text{lumi.}})$. Values of $|r_{\text{lumi.}}| \in [1,3]$ are shown in *italics* and values $|r_{\text{lumi.}}| > 3$ are shown in **bold**.

points by the systematic errors, r_k, are solved for analytically by minimising the χ^2 in Eq. (2). There is a clear trade-off between the systematic shifts r_k and the parameters of the gluon distribution. Deficiencies in the theory calculation can be masked to some extent by large systematic shifts, therefore it is important to check that the optimal r_k values are not unreasonable. This is straightforward when using a χ^2 definition like Eq. (2), but is more difficult using an equivalent form

$$\chi^2 = \sum_{i=1}^{N_{\text{pts.}}} \sum_{i'=1}^{N_{\text{pts.}}} (D_i - T_i)\left(V^{-1}\right)_{ii'} (D_{i'} - T_{i'}), \qquad (3)$$

written in terms of the inverse of the experimental covariance matrix,

$$V_{ii'} = \delta_{ii'}\left(\sigma_i^{\text{uncorr.}}\right)^2 + \sum_{k=1}^{N_{\text{corr.}}} \sigma_{k,i}^{\text{corr.}}\,\sigma_{k,i'}^{\text{corr.}},$$

as used by the ABKM and NNPDF fitting groups. More precisely, NNPDF use a refinement to treat normalisation errors as multiplicative [65], while Alekhin (ABKM) treats all correlated systematic errors as multiplicative [66, 67].

In Table 4 we give the χ^2 per data point, calculated using Eq. (2), for the CDF Run II data on inclusive jet production using the k_T jet algorithm [57], for different NNLO PDF sets and different scale choices $\mu_R = \mu_F = \mu = \{p_T/2, p_T, 2p_T\}$, where p_T is the jet transverse momentum. We treat the luminosity uncertainty as any other correlated systematic. However, we find that the relevant systematic shift $r_{\text{lumi.}} \sim 3$–5 for some PDF sets with soft high-x gluon distributions (e.g. ABKM09 and HERAPDF1.0), which is clearly completely unreasonable, as it means that the data points are normalised downwards by 3–5 times the nominal luminosity uncertainty (5.8% for CDF). The penalty term $r_{\text{lumi.}}^2$ will contribute only 9–25 units to the total χ^2 given by Eq. (2),

which can therefore still lead to reasonably low overall χ^2 values.

It is the usual situation at collider experiments that the luminosity determination is common to all cross sections measured from a given data set, so the requirement of a single common luminosity is mandatory when fitting multiple measurements taken during a single running period. All NNLO PDFs are in good agreement with the Tevatron W and Z cross sections. If the Tevatron jet data were normalised downwards by 20–30% (i.e. 3–5 times the luminosity uncertainty), the Tevatron W and Z total cross sections would need to normalised downwards by the same amount, resulting in complete disagreement with all theory predictions. This example illustrates the utility of simultaneously fitting W and Z cross sections together with jet cross sections at the Tevatron (and LHC). The luminosity shifts, common to both data sets, are effectively determined by the more precise W and Z cross sections. The luminosity uncertainty is then effectively removed from the jet cross sections, thereby allowing the jet data to provide a tighter constraint on the gluon distribution (and α_S).

To avoid these completely unrealistic luminosity shifts, $r_{\text{lumi.}} \sim 3$–5, without going into the complication of simultaneously including W and Z cross sections in the χ^2 computation, we will calculate the χ^2 values for the Tevatron jet data using Eq. (2), but with the simple restriction that the relevant systematic shift $|r_{\text{lumi.}}| \leq 1$. More practically, this means that if $|r_{\text{lumi.}}| > 1$ for any particular PDF set, we fix $r_{\text{lumi.}}$ at ± 1 and reevaluate Eq. (2) with the luminosity removed from the list of correlated systematics. The results are given in Table 5. We highlight in bold the χ^2 values lying inside the 90% C.L. region defined as

$$\chi^2 < \left(\frac{\chi_0^2}{\xi_{50}}\right)\xi_{90}, \qquad (4)$$

NNLO PDF	$\alpha_S(M_Z^2)$	$\mu = p_T/2$	$\mu = p_T$	$\mu = 2p_T$
MSTW08	0.1171	1.39 (+0.35)	**0.69** (−0.45)	0.97 (−1.00)
NNPDF2.1	0.1190	**0.68** (−0.77)	**0.81** (−1.00)	1.29 (−1.00)
HERAPDF1.0	0.1145	2.64 (−1.00)	2.15 (−1.00)	2.20 (−1.00)
HERAPDF1.0	0.1176	2.24 (−0.48)	1.17 (−1.00)	1.23 (−1.00)
HERAPDF1.5	0.1176	1.61 (+1.00)	**0.77** (+0.30)	1.06 (−0.39)
ABKM09	0.1135	2.55 (−1.00)	2.76 (−1.00)	3.41 (−1.00)
JR09	0.1124	**0.75** (+0.13)	1.26 (−0.61)	2.21 (−1.00)

Table 5: Same as Table 4, but at most a 1-σ shift in normalisation is allowed. The optimal value of "$-r_{\text{lumi.}}$" is shown in brackets, subject to this restriction, where the data points are shifted as $D_i \rightarrow D_i(1 - 0.058\, r_{\text{lumi.}})$. We highlight in bold those values lying inside the 90% C.L. region, defined by Eq. (4), which gives $\chi^2/N_{\text{pts.}} < 0.83$.

where ξ_{50} and ξ_{90} are the 50th and 90th percentiles of the χ^2-distribution with $N_{\text{pts.}} = 76$ degrees of freedom. (These quantities are defined in detail in Sect. 6.2 of Ref. [4].) Here, χ_0^2 is defined as the lowest χ^2 value of all theory predictions, i.e. assumed to be close to the best possible fit, so that the rescaling factor χ_0^2/ξ_{50} in Eq. (4) empirically accounts for any unusual fluctuations preventing the best possible fit having $\chi^2 \simeq \xi_{50} \simeq N_{\text{pts.}}$ [63]. The 90% C.L. region given in this way is used to determine the PDF uncertainties according to the "dynamical tolerance" prescription introduced in Ref. [4], so PDF sets with χ^2 values far outside this region cannot be considered to give an acceptable description of the data. We see from Table 5 that MSTW08, NNPDF2.1 and HERAPDF1.5 give an acceptable description for $\mu = p_T$, while HERAPDF1.0 (with the lower α_S value) and ABKM09 give $\chi^2/N_{\text{pts.}} \sim 2$–3. The JR09 set and the HERAPDF1.0 set with $\alpha_S(M_Z^2) = 0.1176$ give a better description, and give predictions for $gg \rightarrow H$ cross sections at the Tevatron which are closer to the MSTW08 predictions than those from ABKM09 and the HERA-PDF1.0 NNLO set with $\alpha_S(M_Z^2) = 0.1145$. The same trend is apparent, but to a somewhat lesser extent, for the CDF Run II inclusive jet data using the cone-based Midpoint jet algorithm [58] and the DØ Run II inclusive jet data using a cone jet algorithm [59]; see Ref. [2] for the details and more discussion.

To summarise, comparison with Tevatron jet data is subtle because of the large correlated systematic uncertainties. The systematic shifts can compensate for inadequacies in the theory calculation. The traditional χ^2 definition in terms of the experimental covariance matrix, Eq. (3), can hide such systematic shifts. In particular, we find that the Tevatron jet data need to be normalised downwards by typically between 3-σ and 5-σ (see Table 4) to achieve the best agreement with some PDF sets, particularly the ABKM09 predictions. Even if the luminosity shift is artificially constrained,

the other systematic shifts move by large amounts for the inclusive jet data, incompatible with the Gaussian expectation. No such problems are observed for the MSTW08 predictions. It can also be seen from the plots in Ref. [68] that the unshifted Tevatron jet data lie significantly above the theory predictions even after including these data in variants of the ABKM09 fit. Constraining the Tevatron luminosity shifts, for example, so that the predicted W and Z cross sections agreed with Tevatron data, would increase the constraining power of the Tevatron jet data and thereby very likely give a larger α_S and high-x gluon distribution than the current ABM studies [68]. Even with the existing treatment, the NNLO Tevatron $gg \rightarrow H$ cross section for $M_H = 165$ GeV goes up by 15% when including the CDF Run II (k_T jet algorithm) [57] data set in a variant of the ABKM09 fit [68].

6. Strong coupling α_S from DIS

A recent claim has been made [46] that the bulk of the MSTW08/ABKM09 difference in both the extracted $\alpha_S(M_Z^2)$ value and the $gg \rightarrow H$ predictions is explained by the treatment of NMC data [69]. The differential cross section for DIS of charged leptons off nucleons, $\ell N \rightarrow \ell X$, neglecting the nucleon and lepton masses, and assuming single-photon exchange, is

$$\frac{d^2\sigma}{dx\,dQ^2} \simeq \frac{4\pi\alpha^2}{x\,Q^4}\left[1 - y + \frac{y^2/2}{1 + R(x, Q^2)}\right] F_2(x, Q^2), \quad (5)$$

where $R = \sigma_L/\sigma_T \simeq F_L/(F_2 - F_L)$ is the ratio of the $\gamma^* N$ cross sections for longitudinally and transversely polarised photons, Q^2 is the photon virtuality, x is the Bjorken variable and $y \simeq Q^2/(x\,s)$ is the inelasticity (with $\sqrt{s}$ the ℓN centre-of-mass energy). The ABKM09 [16] analysis fitted the NMC differential cross sections directly, calculating F_L to $O(\alpha_S^2)$ and including empirical higher-twist corrections. The

ABKM09	MSTW08
Fit NMC cross section	Fit NMC F_2
Fit $Q^2 \geq 2.5 \text{ GeV}^2$	Fit $Q^2 \geq 2 \text{ GeV}^2$
Fit $W^2 \geq 3.24 \text{ GeV}^2$	Fit $W^2 \geq 15 \text{ GeV}^2$
Fit higher-twist	Neglect higher-twist
Separated energies E_μ	Averaged energies E_μ
Correlated systematics	Neglect correlations
3 gluon parameters	7 gluon parameters
No jet data	**Tevatron jet data**

Table 6: Main differences in the treatment of NMC data by the ABKM09 [16] and MSTW08 [4] fits.

MSTW08 [4] analysis instead fitted the NMC F_2 values corrected for R, where $R(x, Q^2) = R_{\text{NMC}}(x)$ if $x < 0.12$ or $R(x, Q^2) = R_{1990}(x, Q^2)$ if $x > 0.12$ [69]. Here, $R_{\text{NMC}}(x)$ was a (Q^2-independent) value extracted from NMC data, while $R_{1990}(x, Q^2)$ was a Q^2-dependent empirical parameterisation of SLAC data dating from 1990 [70]. By replacing the NMC differential cross-section data by NMC F_2 data, ABM [46] found that their best-fit $\alpha_S(M_Z^2)$ moved from 0.1135 to 0.1170 and their $gg \rightarrow H$ cross sections at the Tevatron and LHC moved closer to the MSTW08 values. ABM [46] therefore concluded that the use of NMC F_2 data in the MSTW08 fit rather than the differential cross section is the main reason for the higher $\alpha_S(M_Z^2)$ and Higgs cross sections obtained with MSTW08.

It is certainly more consistent to fit directly to the NMC differential cross-section data, and so this rather dramatic assertion made by ABM [46] would obviously be very worrying if correct. However, there are many other differences between the two analyses besides the treatment of F_L for the NMC data, with some relevant differences given in Table 6 and where the last row (inclusion of jet data) is highlighted as being the most important. Nevertheless, we carried out a detailed investigation of the sensitivity to NMC data in Ref. [2]. Rather than repeat the MSTW08 analysis by fitting the NMC differential cross sections, we noted that the original NMC paper [69] made an alternative extraction of F_2 values using the SLAC R_{1990} parameterisation [70]. We observed that the MSTW08 NNLO prediction, with F_L calculated to $O(\alpha_S^3)$ and without any higher-twist corrections, gives a good description of the SLAC R_{1990} parameterisation, demonstrating that fitting the alternative NMC F_2 data extracted using the SLAC R_{1990} parameterisation will give very similar results to fitting the NMC differential cross sections. In Table 7 we show the effect of repeating the MSTW08 NNLO fit with the NMC F_2 data extracted using R_{1990} on $\alpha_S(M_Z^2)$ and the

Higgs cross sections (for $M_H = 165$ GeV) at the Tevatron and LHC, and in Fig. 16 we show the change in the gluon distribution at the corresponding scale. We make other fits either cutting the NMC F_2 data for $x < 0.1$, above which the R correction in Eq. (5) is very small indeed, or completely removing all NMC F_2 data. In all cases there is very little change in $\alpha_S(M_Z^2)$, the gluon distribution, and the Higgs cross section. We conclude that the treatment of NMC data cannot explain the difference between the MSTW08 and ABKM09 results. Similar stability has been found by the NNPDF group [71], but in a somewhat less relevant study with fixed α_S.

The cuts on DIS data are not explicitly given in the ABKM09 paper [16], but the previous AMP06 paper [72] mentions that DIS data are removed with $Q^2 < 2.5$ GeV^2 and $W^2 < (1.8 \text{ GeV})^2 = 3.24 \text{ GeV}^2$, compared to the MSTW08 fit which removes DIS data with $Q^2 < 2$ GeV^2 and $W^2 < 15$ GeV^2. The much weaker cut on the hadronic invariant mass (squared), $W^2 \simeq Q^2(1/x - 1)$, clearly explains why higher-twist corrections are more important in the ABKM09 analysis. To investigate the possible effect of neglected higher-twist corrections on the MSTW08 NNLO fit we raised the cuts to remove DIS data with $W^2 < 20$ GeV^2 and either $Q^2 < 5$ GeV^2 or $Q^2 < 10$ GeV^2. The results are shown in Table 7 and Fig. 16. The changes in α_S, the gluon distribution and the Higgs cross sections are generally small and within uncertainties, although with the strongest Q^2 cut there is no data constraint below $x = 10^{-4}$ and little just above, so the PDFs differ but have large uncertainties at low x values.

In Table 7 and Fig. 16 we show the results of the MSTW08 NNLO fit with a fixed $\alpha_S(M_Z^2) = 0.113$ [5] (slightly below the ABKM09 value), and even in this case the gluon distribution and Higgs cross sections move only part of the way towards the ABKM09 result, as already seen in Fig. 10. The MSTW08 NNLO input gluon parameterisation [4] has 7 free parameters compared to only 3 for ABKM09, only 2 for JR09 and HERAPDF1.0 (although the value of Q_0^2 is optimised in the case of JR09), and only 4 for HERAPDF1.5 NNLO. In the lack of any direct data constraint on the high-x gluon distribution, the other fits are therefore constrained by the form of the input parameterisation, avoiding potential pathological behaviour such as a negative high-x gluon distribution. In an attempt to mimic the ABKM09 fit we performed a variant of the MSTW08 NNLO fit without jet data and with the input gluon distribution forced to be positive. The results of this fit are shown in Table 7 and Fig. 16 and it goes some way towards reproducing the high-x gluon of the ABKM09 fit and the

NNLO PDF	$\alpha_S(M_Z^2)$	σ_H at Tevatron	σ_H at 7 TeV LHC
MSTW08	**0.1171**	**0.342 pb**	**7.91 pb**
Use R_{1990} for NMC F_2	0.1167	−0.7%	−0.9%
Cut NMC F_2 ($x < 0.1$)	0.1162	−1.2%	−2.1%
Cut all NMC F_2 data	0.1158	−0.7%	−2.1%
Cut $Q^2 < 5$ GeV^2, $W^2 < 20$ GeV^2	0.1171	−1.2%	+0.4%
Cut $Q^2 < 10$ GeV^2, $W^2 < 20$ GeV^2	0.1164	−3.0%	−1.7%
Fix $\alpha_S(M_Z^2)$	0.1130	−11%	−7.6%
Input $xg > 0$, no jets	0.1139	−17%	−4.9%
ABKM09	0.1135	−26%	−11%

Table 7: Effect of NMC treatment on $\alpha_S(M_Z^2)$ and Higgs cross sections ($M_H = 165$ GeV). We also show the effect of raising the cuts imposed on the DIS data compared to the default of removing data with $Q^2 < 2$ GeV^2 and $W^2 < 15$ GeV^2. Finally, we show the effect of simply fixing $\alpha_S(M_Z^2)$ to be close to the ABKM09 value, or performing a fit with a positive-definite input gluon distribution and no jet data, and we compare directly to ABKM09. Table taken from Ref. [2].
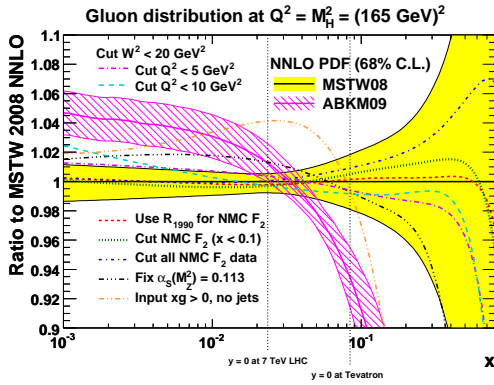

Figure 16: Effect of NMC treatment on the gluon distribution at a scale $Q^2 = (165 \text{ GeV})^2$. The values of $x = M_H/\sqrt{s}$ relevant for central production (assuming $p_T^H = 0$) of a SM Higgs boson of mass $M_H = 165$ GeV at the Tevatron and LHC are indicated. We also show the effect of raising the cuts imposed on the DIS data compared to the default of removing data with $Q^2 < 2$ GeV^2 and $W^2 < 15$ GeV^2. Finally, we show the effect of simply fixing $\alpha_S(M_Z^2)$ to be close to the ABKM09 value, or performing a fit with a positive-definite input gluon distribution and no jet data, and we compare directly to ABKM09. Plot taken from Ref. [2].

corresponding Tevatron $gg \to H$ prediction, certainly closer than we come with other modifications.

Other differences between the two analyses are that ABKM09 used the NMC data for separate muon beam energies, whereas MSTW08 used the NMC data averaged over beam energies, which reduces the maximum effect of the change in R for a particular data point, i.e. at a given x and Q^2, a data point at high y, and so very sensitive to R at a low beam energy, is at lower y for a higher beam energy. In the case of the averaged NMC data, correlated systematic uncertainties are unavailable, so the MSTW08 fit simply added errors (other than nor-

malisation) in quadrature. As with the Tevatron jet data, deficiencies in the theory calculation may be hidden, without much trace, by large systematic shifts implicit in the χ^2 definition, Eq. (3), similar to that used in the ABKM09 analysis. We conclude that the greater sensitivity to the treatment of NMC data found by ABM [46] is due to a variety of reasons, but perhaps most significantly, the inclusion of higher-twist corrections due to the weaker cuts on DIS data, and, as we have repeatedly emphasised, the lack of additional constraints provided by the Tevatron jet data to pin down the high-x gluon distribution.

Note from Figs. 9, 10, 11 and 12 that the HERA-PDF1.5 NNLO prediction has a very large uncertainty in the upwards direction. For example, the uncertainties on the $t\bar{t}$ cross section can be broken down as

$$\sigma_{t\bar{t}} = 178^{+4}_{-6}(\text{exp.})^{+37}_{-6}(\text{model})^{+0}_{-6}(\text{param.}) \pm 6(\alpha_S) \text{ pb.}$$

The dominant model uncertainty comes from varying the minimum Q^2 cut from the default value of $Q^2_{\text{min}} = 3.5$ GeV^2. Lowering the cut to $Q^2_{\text{min}} = 2.5$ GeV^2 increases $\sigma_{t\bar{t}}$ by 9 pb, while raising the cut to $Q^2_{\text{min}} = 5$ GeV^2 increases $\sigma_{t\bar{t}}$ by 35 pb, almost all of the total uncertainty on $\sigma_{t\bar{t}}$ in the upwards direction. It is maybe worrying that the NLO version does not exhibit the same sensitivity, perhaps because of the more restrictive parameterisation at NLO (only 2 gluon parameters) than at NNLO (4 gluon parameters). As shown in Fig. 16, the high-x gluon in the MSTW08 fit is relatively insensitive to raising $Q^2_{\text{min}} = 2 \to \{5, 10\}$ GeV^2, primarily because the Tevatron jet data stabilise the fit and so lessen sensitivity to the fine details of the treatment of the DIS data.

There is a common lore (see, for example, Ref. [73]) that DIS-only fits prefer low $\alpha_S(M_Z^2)$ values, but Ref. [5]

showed that not all DIS data sets prefer low $\alpha_S(M_Z^2)$ values. In particular, this was found to be true only for BCDMS data, and for E665 and SLAC *ep* data, while NMC, SLAC *ed* and HERA data preferred high $\alpha_S(M_Z^2)$ values within the context of the global fit [5]. See also the recent NNPDF studies using an "unbiased" PDF parameterisation [30, 31]. It is well known that α_S is highly *anticorrelated* with the low-x gluon distribution through scaling violations of HERA data: $\partial F_2/\partial \ln(Q^2) \sim \alpha_S \, g$. Then α_S is *correlated* with the high-x gluon distribution through the momentum sum rule; see, for example, Fig. 14(b) of Ref. [5]. Restrictive gluon parameterisations, without the negative small-x term allowed by MSTW [4], can therefore bias the extracted α_S value. For example, the default MSTW08 NNLO fit obtained $\alpha_S(M_Z^2) = 0.1171 \pm 0.0014$, while imposing the restriction of a positive input gluon at $Q_0^2 = 1$ GeV^2 gave a best-fit $\alpha_S(M_Z^2) = 0.1157$, but with a χ^2 worse by 63 units for the global fit to 2615 data points.

What is α_S from only DIS data in the MSTW08 NNLO fit? Recall that the global fit gave $\alpha_S(M_Z^2) = 0.1171 \pm 0.0014$ [5]. To expand on the studies made in Ref. [5], we performed a new NNLO DIS-only fit, which gave a best-fit $\alpha_S(M_Z^2) = 0.1104$, but with an input gluon distribution which went negative for $x > 0.4$ due to lack of any data constraint. This implies a negative charm structure function, F_2^{charm}, and a terrible description ($\chi^2/N_{\mathrm{pts.}} \sim 10$ including correlated systematic errors) of Tevatron jet data using the obtained PDFs. A DIS-only fit fixing the high-x gluon parameters to prevent such bad behaviour gave $\alpha_S(M_Z^2) = 0.1172$, i.e. very similar to the global fit. However, a NNLO fit which imposed the condition of the positive low-x gluon, which stopped the gluon from going negative at high x values, and which also omitted the Tevatron jet data, gave $\alpha_S(M_Z^2) = 0.1139$, rather closer to the ABKM09 value. The very low value of $\alpha_S(M_Z^2) = 0.1104$ found in the DIS-only fit is due to the dominance of BCDMS data. We can show this explicitly by removing the BCDMS data from the DIS-only fit, then the best-fit $\alpha_S(M_Z^2)$ moves from 0.1104 to 0.1193. Repeating the *global* fit with BCDMS data removed gives $\alpha_S(M_Z^2) = 0.1181$, i.e. a change by less than the quoted experimental uncertainty of ± 0.0014. The conclusion is that the Tevatron jet data are vital to pin down the high-x gluon, giving a smaller low-x gluon and therefore a larger α_S in the global fit compared to a DIS-only fit, at the expense of some deterioration in the fit quality of the BCDMS data. The benefits of including the Tevatron jet data to obtain sensible results in a simultaneous fit of PDFs and α_S therefore greatly outweighs any dis-

advantage such as lack of complete NNLO corrections.

The only input DIS value to the current world average $\alpha_S(M_Z^2)$ [28, 29] is the BBG06 value [74], which is from a non-singlet analysis and therefore *in principle* free of assumptions made about the gluon distribution. A value of

$$\alpha_S(M_Z^2) = \left\{ 0.1148^{+0.0019}_{-0.0019}, 0.1134^{+0.0019}_{-0.0021}, 0.1141^{+0.0020}_{-0.0022} \right\}$$

was obtained at {NLO, NNLO, N^{3}LO}, by fitting proton and deuteron structure functions, F_2^p and F_2^d, for $x \geq 0.3$ (assuming only valence quarks, neglecting the singlet contribution), and the less precise $F_2^{\mathrm{NS}} = 2(F_2^p - F_2^d)$ for $x < 0.3$. However, using the MSTW08 NNLO central fit, contributions other than valence quarks are found to make up about 10% (2%) of F_2^p at $x = 0.3$ ($x = 0.5$). As an exercise we performed the MSTW08 NNLO DIS-only fit just to F_2^p and F_2^d for $x > 0.3$ (comprising 282 data points, 160 of these from BCDMS), which gave $\alpha_S(M_Z^2) = 0.1103$ (0.1130) without (with) the singlet contribution included. This is even lower than the BBG06 value presumably due to lack of the $y > 0.3$ cut on BCDMS data applied in the BBG06 analysis. The low value of $\alpha_S(M_Z^2)$ found by BBG06 [74] is therefore due to both dominance of BCDMS data and by what we conclude is the unjustified neglect of the singlet contribution to F_2^p and F_2^d for $x \geq 0.3$. Given that it was argued above that the Tevatron jet data are needed to pin down the high-x gluon, we conclude that an extraction of $\alpha_S(M_Z^2)$ only from inclusive DIS data is not meaningful, and the closest possible to a reliable extraction is the MSTW08 NNLO combined analysis of DIS, Drell–Yan and jet data [4, 5],

$$\alpha_S(M_Z^2) = 0.1171 \pm 0.0014 \text{ (exp.)} \pm 0.002 \text{ (th.)},$$

or, alternatively, the more recent NNLO determination by the NNPDF Collaboration [31],

$$\alpha_S(M_Z^2) = 0.1173 \pm 0.0007 \text{ (exp.)} \pm 0.0009 \text{ (th.)}.$$

These values are the only NNLO determinations, from a simultaneous fit with PDFs, which are in agreement with the current world average $\alpha_S(M_Z^2) = 0.1184 \pm 0.0007$ [28, 29]; see Fig. 4.

With all these problems in α_S determinations from non-global PDF fits, it is therefore disconcerting that the 2011 update of the world average α_S by S. Bethke [75], intended for the PDG 2012 review, has chosen to treat the MSTW08, ABKM09, JR09 and BBG06 determinations on an equal footing in forming a "DIS" value of $\alpha_S(M_Z^2) = 0.1148 \pm 0.0024$, which lies significantly below other classes of measurements and the overall preliminary world average of $\alpha_S(M_Z^2) = 0.1185 \pm 0.0008$ [75].

7. Summary

The "MSTW 2008" determination of parton distribution functions (PDFs) [4, 5, 6] is still fairly current with no immediate update planned. Another update may be appropriate after the final HERA II combined data (including F_2^{charm}) are published, together with making the first use of input data from the LHC, such as differential Z/γ^*, W, jet and W+charm distributions. However, provided that the new data are reasonably consistent with the old data, we would not expect to see substantial changes to the PDFs. On the other hand, perhaps a more pressing concern is the compatibility of existing PDFs from different groups with each other. We have updated a recently-published study of benchmark cross sections [1, 2] to include results obtained using the NNPDF2.1 NNLO [13] and HERAPDF1.5 NNLO [15] PDF sets, and to compare to LHC data on W, Z and $t\bar{t}$ production. Supplementary plots can be obtained from a public webpage [3]. There is now reasonably good agreement between the NLO *global* fits from MSTW08, CT10 and NNPDF2.1, all using variants of a GM-VFNS to treat DIS structure functions, together with the NNLO *global* fits from MSTW08 and NNPDF2.1. More variation is seen with other PDF sets using more limited data sets and/or restrictive input PDF parameterisations. The latest HERAPDF1.5 NNLO set is surprisingly close to MSTW08, at least for the central value, unlike the analogous HERAPDF1.5 NLO set, but it has a large uncertainty in the high-x gluon distribution due to variation of the minimum Q^2 cut, not seen at NLO perhaps due to the more restricted parameterisation. The Tevatron jet data are important to pin down the high-x gluon, with an indirect effect on the value of $\alpha_S(M_Z^2)$ extracted, reducing sensitivity to the fine details of the treatment of DIS data. The LHC measurements of top-pair production tend to favour PDF sets where the high-x gluon is determined using Tevatron jet data. Conversely, both the LHC $t\bar{t}$ cross sections and the Tevatron jet data strongly disfavour PDF sets with soft high-x gluon distributions and low α_S values (specifically, ABKM09 [16]), which give anomalously low Higgs cross sections at both the Tevatron and LHC. We therefore caution against the use of PDF sets obtained from "non-global" fits where the high-x gluon distribution is generally constrained by assumptions made on the form of parameterisation in the absence of a direct data constraint.

References

[1] G. Watt, Parton distribution function dependence of benchmark Standard Model total cross sections at the 7 TeV LHC, JHEP 1109 (2011) 069. arXiv:1106.5788, doi:10.1007/JHEP09(2011)069.

[2] R. S. Thorne, G. Watt, PDF dependence of Higgs cross sections at the Tevatron and LHC: Response to recent criticism, JHEP 1108 (2011) 100. arXiv:1106.5789, doi:10.1007/JHEP08(2011)100.

[3] http://projects.hepforge.org/mstwpdf/pdf4lhc/.

[4] A. D. Martin, W. J. Stirling, R. S. Thorne, G. Watt, Parton distributions for the LHC, Eur.Phys.J. C63 (2009) 189–285. arXiv:0901.0002, doi:10.1140/epjc/s10052-009-1072-5.

[5] A. D. Martin, W. J. Stirling, R. S. Thorne, G. Watt, Uncertainties on α_S in global PDF analyses and implications for predicted hadronic cross sections, Eur.Phys.J. C64 (2009) 653–680. arXiv:0905.3531, doi:10.1140/epjc/s10052-009-1164-2.

[6] A. D. Martin, W. J. Stirling, R. S. Thorne, G. Watt, Heavy-quark mass dependence in global PDF analyses and 3- and 4-flavour parton distributions, Eur.Phys.J. C70 (2010) 51–72. arXiv:1007.2624, doi:10.1140/epjc/s10052-010-1462-8.

[7] G. Watt, Parton Distributions: HERA-Tevatron-LHC, PoS HCP2009 (2009) 014. arXiv:1001.3954.

[8] R. S. Thorne, A. D. Martin, W. J. Stirling, G. Watt, The effects of combined HERA and recent Tevatron $W \rightarrow \ell\nu$ charge asymmetry data on the MSTW PDFs, PoS DIS2010 (2010) 052. arXiv:1006.2753.

[9] F. D. Aaron, et al., Combined Measurement and QCD Analysis of the Inclusive $e^\pm p$ Scattering Cross Sections at HERA, JHEP 1001 (2010) 109. arXiv:0911.0884, doi:10.1007/JHEP01(2010)109.

[10] P. M. Nadolsky, H.-L. Lai, Q.-H. Cao, J. Huston, J. Pumplin, et al., Implications of CTEQ global analysis for collider observables, Phys.Rev. D78 (2008) 013004. arXiv:0802.0007, doi:10.1103/PhysRevD.78.013004.

[11] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, et al., New parton distributions for collider physics, Phys.Rev. D82 (2010) 074024. arXiv:1007.2241, doi:10.1103/PhysRevD.82.074024.

[12] R. D. Ball, V. Bertone, F. Cerutti, L. Del Debbio, S. Forte, et al., Impact of Heavy Quark Masses on Parton Distributions and LHC Phenomenology, Nucl.Phys. B849 (2011) 296–363. arXiv:1101.1300, doi:10.1016/j.nuclphysb.2011.03.021.

[13] R. D. Ball, et al., Unbiased global determination of parton distributions and their uncertainties at NNLO and at LO, Nucl.Phys. B855 (2012) 153–221. arXiv:1107.2652, doi:10.1016/j.nuclphysb.2011.09.024.

[14] F. D. Aaron, et al., PDF fits including HERA-II high Q^2 data, H1prelim-10-142, ZEUS-prel-10-018.

[15] F. D. Aaron, et al., HERAPDF1.5 NNLO, H1prelim-11-042, ZEUS-prel-11-002.

[16] S. Alekhin, J. Blümlein, S. Klein, S. Moch, The 3, 4, and 5-flavor NNLO Parton from Deep-Inelastic-Scattering Data and at Hadron Colliders, Phys.Rev. D81 (2010) 014032. arXiv:0908.2766, doi:10.1103/PhysRevD.81.014032.

[17] M. Glück, P. Jimenez-Delgado, E. Reya, Dynamical parton distributions of the nucleon and very small-x physics, Eur.Phys.J. C53 (2008) 355–366. arXiv:0709.0614, doi:10.1140/epjc/s10052-007-0462-9.

[18] M. Glück, P. Jimenez-Delgado, E. Reya, C. Schuck, On the role of heavy flavor parton distributions at high energy colliders, Phys.Lett. B664 (2008) 133–138. arXiv:0801.3618, doi:10.1016/j.physletb.2008.04.063.

[19] P. Jimenez-Delgado, E. Reya, Dynamical NNLO parton distributions, Phys.Rev. D79 (2009) 074023. arXiv:0810.4274, doi:10.1103/PhysRevD.79.074023.

[20] P. Jimenez-Delgado, E. Reya, Variable Flavor Number Parton Distributions and Weak Gauge and Higgs Boson Production

at Hadron Colliders at NNLO of QCD, Phys.Rev. D80 (2009) 114011. arXiv:0909.1711, doi:10.1103/PhysRevD.80.114011.

[21] S. Alekhin, S. Alioli, R. D. Ball, V. Bertone, J. Blümlein, et al., The PDF4LHC Working Group Interim Report (2011). arXiv:1101.0536.

[22] M. Botje, J. Butterworth, A. Cooper-Sarkar, A. de Roeck, J. Feltesse, et al., The PDF4LHC Working Group Interim Recommendations (2011). arXiv:1101.0538.

[23] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino, R. Tanaka (Eds.), Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables (CERN, Geneva, 2011). arXiv:1101.0593.

[24] M. R. Whalley, D. Bourilkov, R. C. Group, The Les Houches accord PDFs (LHAPDF) and LHAGLUE (2005). arXiv:hep-ph/0508110.

[25] R. D. Ball, V. Bertone, F. Cerutti, L. Del Debbio, S. Forte, et al., Reweighting and Unweighting of Parton Distributions and the LHC W lepton asymmetry data, Nucl.Phys. B855 (2012) 608–638. arXiv:1108.1758, doi:10.1016/j.nuclphysb.2011.10.018.

[26] R. D. Ball, L. Del Debbio, S. Forte, A. Guffanti, J. I. Latorre, et al., A first unbiased global NLO determination of parton distributions and their uncertainties, Nucl.Phys. B838 (2010) 136–206. arXiv:1002.4407, doi:10.1016/j.nuclphysb.2010.05.008.

[27] N. Kidonakis, J. F. Owens, Effects of higher-order threshold corrections in high-E_T jet production, Phys.Rev. D63 (2001) 054019. arXiv:hep-ph/0007268, doi:10.1103/PhysRevD.63.054019.

[28] S. Bethke, The 2009 World Average of α_S, Eur.Phys.J. C64 (2009) 689–703. arXiv:0908.1135, doi:10.1140/epjc/s10052-009-1173-1.

[29] K. Nakamura, et al., Review of particle physics, J.Phys.G G37 (2010) 075021. doi:10.1088/0954-3899/37/7A/075021.

[30] S. Lionetti, R. D. Ball, V. Bertone, F. Cerutti, L. Del Debbio, et al., Precision determination of α_S using an unbiased global NLO parton set, Phys.Lett. B701 (2011) 346–352. arXiv:1103.2369, doi:10.1016/j.physletb.2011.05.071.

[31] R. D. Ball, V. Bertone, L. Del Debbio, S. Forte, A. Guffanti, et al., Precision NNLO determination of $\alpha_S(M_Z^2)$ using an unbiased global parton set (2011). arXiv:1110.2483.

[32] R. Hamberg, W. L. van Neerven, T. Matsuura, A complete calculation of the order α_S^2 correction to the Drell-Yan K-factor, Nucl.Phys. B359 (1991) 343–405. doi:10.1016/0550-3213(91)90064-5, 10.1016/0550-3213(91)90064-5.

[33] S. Chatrchyan, et al., Measurement of the Inclusive W and Z Production Cross Sections in pp Collisions at $\sqrt{s}$ = 7 TeV with the CMS experiment, JHEP 1110 (2011) 132. arXiv:1107.4789, doi:10.1007/JHEP10(2011)132.

[34] G. Aad, et al., Measurement of the inclusive $W^\pm$ and Z/γ^* cross sections in the electron and muon decay channels in pp collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector (2011). arXiv:1109.5141.

[35] C. Anastasiou, L. J. Dixon, K. Melnikov, F. Petriello, High precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at NNLO, Phys.Rev. D69 (2004) 094008. arXiv:hep-ph/0312266, doi:10.1103/PhysRevD.69.094008.

[36] A measurement of the total $W^\pm$ and Z/γ^* cross sections in the e and μ decay channels and of their ratios in pp collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector, ATLAS-CONF-2011-041 (18th March 2011).

[37] K. Melnikov, F. Petriello, Electroweak gauge boson production at hadron colliders through $O(\alpha_S^2)$, Phys.Rev. D74 (2006) 114017. arXiv:hep-ph/0609070, doi:10.1103/PhysRevD.74.114017.

[38] R. Gavin, Y. Li, F. Petriello, S. Quackenbush, FEWZ 2.0: A code for hadronic Z production at next-to-next-to-

leading order, Comput.Phys.Commun. 182 (2011) 2388–2403. arXiv:1011.3540, doi:10.1016/j.cpc.2011.06.008.

[39] S. Catani, L. Cieri, G. Ferrera, D. de Florian, M. Grazzini, Vector boson production at hadron colliders: A Fully exclusive QCD calculation at NNLO, Phys.Rev.Lett. 103 (2009) 082001. arXiv:0903.2120, doi:10.1103/PhysRevLett.103.082001.

[40] Measurement of associated charm production in W final states at $\sqrt{s}$ = 7 TeV, CMS PAS EWK-11-013 (22nd July 2011).

[41] Combined CDF and DØ Upper Limits on Standard Model Higgs Boson Production with up to 8.6 fb^{-1} of Data (2011). arXiv:1107.5518.

[42] Combination of Higgs Boson Searches with up to 4.9 fb^{-1} of pp Collision Data Taken at $\sqrt{s}$ = 7 TeV with the ATLAS Experiment at the LHC, ATLAS-CONF-2011-163 (13th December 2011).

[43] Combination of CMS searches for a Standard Model Higgs boson, CMS PAS HIG-11-032 (13th December 2011).

[44] R. V. Harlander, W. B. Kilgore, Next-to-next-to-leading order Higgs production at hadron colliders, Phys.Rev.Lett. 88 (2002) 201801. arXiv:hep-ph/0201206, doi:10.1103/PhysRevLett.88.201801.

[45] J. Baglio, A. Djouadi, S. Ferrag, R. M. Godbole, The Tevatron Higgs exclusion limits and theoretical uncertainties: A Critical appraisal, Phys.Lett. B699 (2011) 368–371. arXiv:1101.1832, doi:10.1016/j.physletb.2011.04.039, 10.1016/j.physletb.2011.06.074.

[46] S. Alekhin, J. Blümlein, S. Moch, Higher order constraints on the Higgs production rate from fixed-target DIS data, Eur.Phys.J. C71 (2011) 1723. arXiv:1101.5261, doi:10.1140/epjc/s10052-011-1723-1.

[47] Measurement of $t\bar{t}$ Pair Production Cross Section at $\sqrt{s}$ = 7 TeV using b-quark Jet Identification Techniques in Lepton + Jet Events, CMS PAS TOP-11-003 (3rd October 2011).

[48] Measurement of the $t\bar{t}$ production cross-section in pp collisions at $\sqrt{s}$ = 7 TeV using kinematic information of lepton+jets events, ATLAS-CONF-2011-121 (20th August 2011).

[49] M. Aliev, H. Lacker, U. Langenfeld, S. Moch, P. Uwer, et al., HATHOR: HAdronic Top and Heavy quarks crOss section calculatoR, Comput.Phys.Commun. 182 (2011) 1034–1046. arXiv:1007.1327, doi:10.1016/j.cpc.2010.12.040.

[50] N. Kidonakis, B. D. Pecjak, Top-quark production and QCD (2011). arXiv:1108.6063.

[51] M. Beneke, P. Falgari, S. Klein, C. Schwinn, Hadronic top-quark pair production with NNLL threshold resummation, Nucl.Phys. B855 (2012) 695–741. arXiv:1109.1536, doi:10.1016/j.nuclphysb.2011.10.021.

[52] M. Cacciari, M. Czakon, M. L. Mangano, A. Mitov, P. Nason, Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation (2011). arXiv:1111.5869.

[53] M. Lancaster, Combination of CDF and DØ results on the mass of the top quark using up to 5.8 fb^{-1} of data (2011). arXiv:1107.5255.

[54] M. Wobisch, D. Britzger, T. Kluge, K. Rabbertz, F. Stober, Theory-Data Comparisons for Jet Measurements in Hadron-Induced Processes (2011). arXiv:1109.1310.

[55] Combination of top quark pair production cross section measurements, CMS PAS TOP-11-024 (23rd November 2011).

[56] D. de Florian, Status of parton distributions and impact on the Higgs, talk at "Higgs Hunting 2011", Orsay, France (28th July 2011).

[57] A. Abulencia, et al., Measurement of the Inclusive Jet Cross Section using the k_T algorithm in $p\bar{p}$ Collisions at $\sqrt{s}$ = 1.96 TeV with the CDF II Detector, Phys.Rev. D75 (2007) 092006. arXiv:hep-

ex/0701051, doi:10.1103/PhysRevD.75.092006, 10.1103/PhysRevD.75.119901.

[58] T. Aaltonen, et al., Measurement of the Inclusive Jet Cross Section at the Fermilab Tevatron $p\bar{p}$ Collider Using a Cone-Based Jet Algorithm, Phys.Rev. D78 (2008) 052006. arXiv:0807.2204, doi:10.1103/PhysRevD.78.052006, 10.1103/PhysRevD.79.119902.

[59] V. Abazov, et al., Measurement of the inclusive jet cross-section in $p\bar{p}$ collisions at $\sqrt{s}$ = 1.96 TeV, Phys.Rev.Lett. 101 (2008) 062001. arXiv:0802.2400, doi:10.1103/PhysRevLett.101.062001.

[60] T. Kluge, K. Rabbertz, M. Wobisch, FastNLO: Fast pQCD calculations for PDF fits (2006). arXiv:hep-ph/0609285.

[61] Z. Nagy, Three jet cross-sections in hadron hadron collisions at next-to-leading order, Phys.Rev.Lett. 88 (2002) 122003. arXiv:hep-ph/0110315, doi:10.1103/PhysRevLett.88.122003.

[62] Z. Nagy, Next-to-leading order calculation of three-jet observables in hadron–hadron collision, Phys.Rev. D68 (2003) 094002. arXiv:hep-ph/0307268, doi:10.1103/PhysRevD.68.094002.

[63] D. Stump, J. Pumplin, R. Brock, D. Casey, J. Huston, et al., Uncertainties of predictions from parton distribution functions. 1. The Lagrange multiplier method, Phys.Rev. D65 (2001) 014012. arXiv:hep-ph/0101051, doi:10.1103/PhysRevD.65.014012.

[64] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, P. M. Nadolsky, et al., New generation of parton distributions with uncertainties from global QCD analysis, JHEP 0207 (2002) 012. arXiv:hep-ph/0201195.

[65] R. D. Ball, et al., Fitting Parton Distribution Data with Multiplicative Normalization Uncertainties, JHEP 1005 (2010) 075. arXiv:0912.2276, doi:10.1007/JHEP05(2010)075.

[66] S. Alekhin, Extraction of parton distributions and α_S from DIS data within the Bayesian treatment of systematic errors, Eur.Phys.J. C10 (1999) 395–403. arXiv:hep-ph/9611213, doi:10.1007/s100520050763.

[67] S. I. Alekhin, Global fit to the charged leptons DIS data: α_S, parton distributions, and high twists, Phys.Rev. D63 (2001) 094022. arXiv:hep-ph/0011002, doi:10.1103/PhysRevD.63.094022.

[68] S. Alekhin, J. Blümlein, S.-O. Moch, Parton distributions and Tevatron jet data (2011). arXiv:1105.5349.

[69] M. Arneodo, et al., Measurement of the proton and deuteron structure functions, F_2^p and F_2^d, and of the ratio σ_L/σ_T, Nucl.Phys. B483 (1997) 3–43. arXiv:hep-ph/9610231, doi:10.1016/S0550-3213(96)00538-X.

[70] L. W. Whitlow, S. Rock, A. Bodek, E. Riordan, S. Dasu, A Precise extraction of $R = \sigma_L/\sigma_T$ from a global analysis of the SLAC deep inelastic ep and ed scattering cross-sections, Phys.Lett. B250 (1990) 193–198. doi:10.1016/0370-2693(90)91176-C.

[71] R. D. Ball, et al., On the Impact of NMC Data on NLO and NNLO Parton Distributions and Higgs Production at the Tevatron and the LHC, Phys.Lett. B704 (2011) 36–42. arXiv:1102.3182, doi:10.1016/j.physletb.2011.08.055.

[72] S. Alekhin, K. Melnikov, F. Petriello, Fixed target Drell-Yan data and NNLO QCD fits of parton distribution functions, Phys.Rev. D74 (2006) 054033. arXiv:hep-ph/0606237, doi:10.1103/PhysRevD.74.054033.

[73] S. Alekhin, J. Blümlein, H. Böttcher, S.-O. Moch, $\alpha_S(M_Z^2)$ in NNLO Analyses of Deep-Inelastic World Data (2011). arXiv:1104.0469.

[74] J. Blümlein, H. Böttcher, A. Guffanti, Non-singlet QCD analysis of deep inelastic world data at $O(\alpha_S^3)$, Nucl.Phys. B774 (2007) 182–207. arXiv:hep-ph/0607200, doi:10.1016/j.nuclphysb.2007.03.035.

[75] S. Bethke, World summary of α_S, talk at this workshop (26th September 2011).