

Parton Distribution Functions

R. Plačakytė (on behalf of the H1 and ZEUS Collaborations)

*Institut für Exp. Physik,
Universität Hamburg,
Germany*

The parton distribution functions (PDFs) of the proton, a necessary input to almost all theory predictions for hadron colliders, are reviewed in this document. An introduction to the PDF determination by global analyses of the main PDF fitting groups with an emphasis on HERA PDFs is presented. Finally, theory predictions based on different PDFs are compared to some recent relevant LHC and TEVATRON measurements.

1. INTRODUCTION

A precise knowledge of the Parton Distribution Functions (PDFs) of the proton is essential in order to make predictions for the Standard Model and beyond the Standard Model processes at hadron colliders. The parton density function $f_i(x, Q^2)$ gives the probability of finding in the proton a parton of flavour i (quarks or gluon) carrying a fraction x of the proton momentum with Q being the energy scale of the hard interaction. Cross sections are calculated by convoluting the parton level cross section with the PDFs. Since QCD does not predict the parton content of the proton, the shapes of the PDFs are determined by a fit to data from experimental observables in various processes, using the DGLAP evolution equation [1]. The knowledge of proton PDFs mainly comes from the Deep Inelastic Scattering (DIS) HERA, fixed target and TEVATRON data. The recent precise TEVATRON and LHC data has a potential to improve constraints of the PDFs further. For example, through the W lepton asymmetry the different quark contributions can be accessed and with inclusive jet productions it is possible to constrain the gluon PDF.

2. PARTON DISTRIBUTION FUNCTIONS

2.1. Proton Structure and DIS

Deep inelastic lepton nucleon scattering probes the structure of matter at small distance scales. At HERA (Hadron Elektron Ring Anlage), DESY, electrons (or positrons) were collided with protons at centre-of-mass energies $\sqrt{s} = 225 - 318$ GeV. Two different deep inelastic ep scattering processes are measured at HERA: neutral current (NC), $ep \rightarrow eX$, and charged current (CC), $ep \rightarrow \nu X$. In neutral current reactions the interaction proceeds via the exchange of a photon or a Z boson while in charged current scattering a W^\pm boson is exchanged (figure 1). The NC (and similarly CC) cross section can be ex-

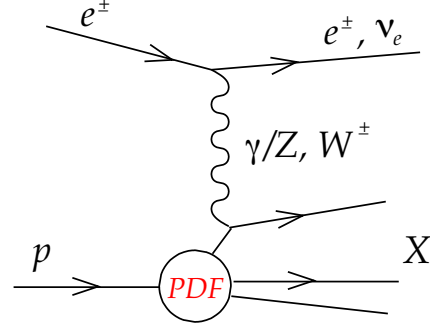


Figure 1: Diagrams of neutral NC and charged CC current deep inelastic scattering processes. The symbols denote the particles, the label "X" denotes the hadronic final state.

pressed in terms of structure functions:

$$\frac{d^2\sigma_{NC}^{e^\pm p}}{dx dQ^2} = \frac{2\pi\alpha^2}{xQ^4} [Y_+ \tilde{F}_2^\pm \mp Y_- x\tilde{F}_3^\pm - y^2 \tilde{F}_L^\pm],$$

where $Y_\pm = 1 \pm (1-y)^2$ with y being the inelasticity. The structure function \tilde{F}_2 is the dominant contribution to the cross section, $x\tilde{F}_3$ is important at high Q^2 and \tilde{F}_L is sizable only at high y . In the framework of perturbative QCD the structure functions are directly related to the parton distribution functions, i.e. in leading order (LO) F_2 is the momentum sum of quark and anti-quark distributions, $F_2 \approx x \sum e_q^2 (q + \bar{q})$, and xF_3 is related to their difference, $xF_3 \approx x \sum 2e_q a_q (q - \bar{q})$. At higher orders, terms related to the gluon density distribution ($\alpha_s g$) appear.

In analogy to neutral currents, the inclusive CC ep cross section can be expressed in terms of structure functions and in LO the e^+p and e^-p cross sections are sensitive to different quark densities:

$$\begin{aligned} e^+ : \quad \tilde{\sigma}_{CC}^{e^+ p} &= x[\bar{u} + \bar{c}] + (1-y)^2 x[d + s] \\ e^- : \quad \tilde{\sigma}_{CC}^{e^- p} &= x[u + c] + (1-y)^2 x[\bar{d} + \bar{s}]. \end{aligned}$$

The measurements of the NC and CC cross sections from HERA extend the kinematic regime in Q^2 by more than two orders of magnitude with respect to the fixed target experiments.

2.2. Determination of PDFs

The general procedure to determine PDFs is as follows. Starting from a parameterisation of the non-perturbative PDFs at a low scale, either by making ad-hoc assumptions on their analytical form or by using the neural-net technology, fits to various sets of experimental data (mainly to DIS data) are performed within the DGLAP evolution scheme. The resulting PDFs depend on the choice of the input data, the order in which the perturbative QCD calculation is performed, the assumptions about the PDFs, the treatment of heavy quarks, the correlation between α_s and the PDFs and the treatment of the uncertainties. Presently, the determination of PDFs is carried out by several groups, namely MSTW [2], CTEQ [3], NNPDF [4], HERAPDF [5], AB(K)M [6] and GJR [7]. In the following, a short description of each group with the relevant data sets used in PDF fits is given.

The MSTW (Martin-Stirling-Thorne-Watt) PDFs are determined from a global analysis of hard-scattering data within the framework of leading-twist fixed-order collinear factorisation in the $\overline{\text{MS}}$ scheme [2]. PDF sets in LO, NLO and NNLO are available. This determination is based on the HERA DIS data (with the exception of the latest combined HERA I data) as the main input for constraining quarks and gluon PDFs at low x , fixed target DIS data constraining quarks and gluons in the high- x region, fixed target Drell-Yan (DY) data which help to constrain high- x sea quarks, TEVATRON jet data contributing to high- x gluon PDF and W , Z data which provide an access to the different quark contributions.

The CTEQ PDFs [3] are obtained by a global analysis of hard scattering data in the framework of general-mass perturbative QCD. This analysis is based on the same data mentioned above (CT10 also includes the recent combined HERA I data and more TEVATRON data). PDF sets in NLO and NNLO QCD are available.

The Neural Net PDF (NNPDF) group [4] parametrise non-perturbative PDFs by training a neural network (i.e. minimising the χ^2) on MC replicas of the experimental data. With such a method NNPDFs are free of assumptions made by other groups (however there is some ambiguity from the procedure itself). It also includes almost all above mentioned experimental data in the fit and provides LO, NLO and NNLO PDFs.

The HERAPDFs [5] are based only on HERA DIS data which allows the usage of the conventional χ^2 tolerance of $\Delta\chi^2 = 1$. Since this analysis is solely based on ep data, the PDFs do not depend on the approach for nuclear corrections needed for fixed target data. The HERAPDF has been extracted in NLO and NNLO. More details about HERA PDFs are given in the next section.

The ABKM [6] group provides NLO and NNLO

PDFs based on the inclusive NC DIS world data from HERA and from fixed-target experiments. They also provide the strong coupling constant α_s in NNLO.

The GJR [7] group uses a more restrictive dynamical approach in the PDFs determination, i.e. a dynamical model is used in the evolution of the input distributions starting at very low Q^2 . Fits to DIS, DY and TEVATRON jet data provide LO, NLO and NNLO PDF sets.

The large number of PDF parameters and their treatment in the fitting procedure within the different groups results in differences of the PDFs provided. In order to study these differences, a benchmarking exercise is being carried out by the PDF4LHC working group formed by the members of the PDF fitting groups mentioned above. As an example of this exercise, the NLO prediction of the Higgs cross section ($M_H = 120$ GeV) for the LHC is shown in figure 2 for different PDF sets as a function of α_s . It can be seen that the different PDF groups are using different default values of α_s and the cross section predictions differ by a few percent when using the default values of α_s as well as when using the same value for α_s . Furthermore, the different predictions have different uncertainties (e.g. HERA PDF has asymmetric uncertainties due to considered parametrisation variation as explained in the next section).

Further comparisons and an overview of the PDF4LHC working group activities can be found in [8].

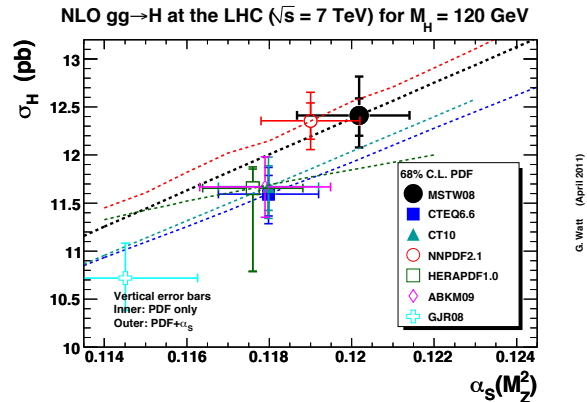


Figure 2: *NLO Higgs cross section predictions* ($M_H = 120$ GeV) using different PDFs at the LHC with $\sqrt{s} = 7$ TeV.

2.3. HERA PDFs

In the HERAPDF1.0 fit [5], the combined set on NC and CC ep inclusive cross-sections for the first running period of HERA (HERA I) is used. The

full statistics HERA inclusive CC and NC data are used for NLO and NNLO QCD fits resulting in HERAPDF1.5 [9]. The same formalism, model and parametrisation assumptions as in the HERAPDF1.0 are used in the HERAPDF1.5(NLO) fit.

The QCD predictions for the structure functions are obtained by solving the DGLAP evolution equations at NLO (or NNLO) in the \overline{MS} scheme with the renormalisation and factorisation scales chosen to be Q^2 . The DGLAP equations yield the PDFs at all values of Q^2 above the input scale Q_0^2 at which they are parametrised as a functions of x . The starting scale Q_0^2 is chosen to be 1.9 GeV^2 such that the starting scale is below the charm mass threshold.

The QCD predictions for the structure functions are obtained by convolution of the PDFs with the NLO coefficient functions calculated using the general mass variable favour number RT scheme [10].

For the parametrisation of PDFs at the input scale the generic form $xf(x) = Ax^B(1-x)^C(1+Ex^2)$ is used. The parametrised PDFs are the gluon distribution xg , the valence quark distributions xu_v , xd_v , and the u-type and d-type anti-quark distributions $x\bar{U}$, $x\bar{D}$. At the starting scale $Q_0^2 = 1.9 \text{ GeV}^2$ $x\bar{U} = x\bar{u}$ and $x\bar{D} = x\bar{d} + x\bar{s}$. The central fit parametrisation is:

$$\begin{aligned} xg(x) &= A_g x^{B_g} (1-x)^{C_g}, \\ xu_v(x) &= A_{u_v} x^{B_{u_v}} (1-x)^{C_{u_v}} (1 + E_{u_v} x^2), \\ xd_v(x) &= A_{d_v} x^{B_{d_v}} (1-x)^{C_{d_v}}, \\ x\bar{U}(x) &= A_{\bar{U}} x^{B_{\bar{U}}} (1-x)^{C_{\bar{U}}}, \\ x\bar{D}(x) &= A_{\bar{D}} x^{B_{\bar{D}}} (1-x)^{C_{\bar{D}}}. \end{aligned}$$

The normalisation parameters A are constrained by the quark number sum-rules and momentum sum-rule, extra constrains for small- x behaviour of d - and u -type quarks $B_{u_v} = B_{d_v}$, $B_{\bar{U}} = B_{\bar{D}}$ and $A_{\bar{U}} = A_{\bar{D}}(1 - f_s)$ (f_s is the strange quark distribution) which ensures that $x\bar{u} \rightarrow x\bar{d}$ as $x \rightarrow 0$.

The break-up of the HERA PDFs into different flavours is illustrated in figure 3. Model uncertainties (shown as yellow bands in the figure) of the central fit solution is evaluated by varying the input assumptions: Q_{min}^2 , f_s , mass of heavy quarks m_C and m_B . Parametrisation uncertainties (green band) is formed by an envelope of the maximal deviation from the central fit varying parametrisation assumptions in the fit and therefore has an asymmetric shape. The determination of parameterisation uncertainties are unique to HERAPDFs.

An example of the parton distribution functions from HERAPDF1.5 at NNLO is shown in figure 4. HERAPDF1.5NLO and NNLO sets are the recommended HERA PDFs to be used for the predictions of processes at LHC.

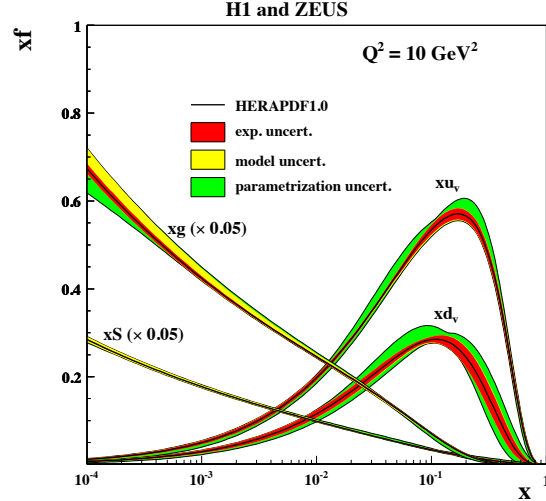


Figure 3: The parton distribution functions from HERAPDF1.0 at $Q^2 = 10 \text{ GeV}^2$. The gluon and sea distributions are scaled down by a factor of 20. The experimental, model and parametrisation uncertainties are shown separately.

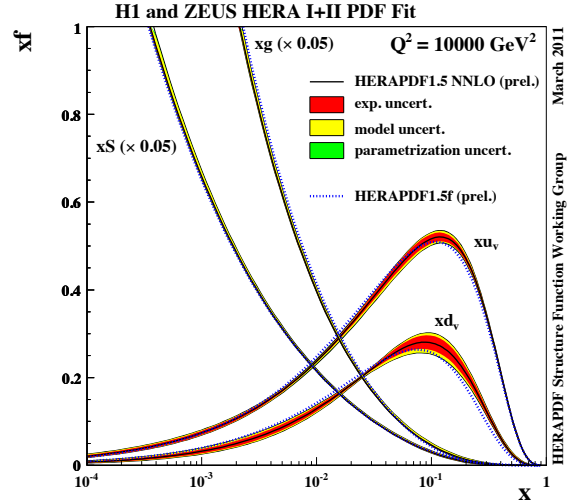


Figure 4: The parton distribution functions from HERAPDF1.5 NNLO at $Q^2 = 10000 \text{ GeV}^2$, i.e. a region relevant for the hadron colliders TEVATRON and LHC. The gluon and sea distributions are scaled down by a factor 20. The experimental, model and parametrisation uncertainties are shown separately. For comparison, the central values of HERAPDF1.0 NNLO are also shown.

2.4. Comparisons to recent LHC and TEVATRON results

The prediction of the Z boson rapidity distribution, based on three different PDFs, are compared to the CDF measurement in figure 5. The predictions of the

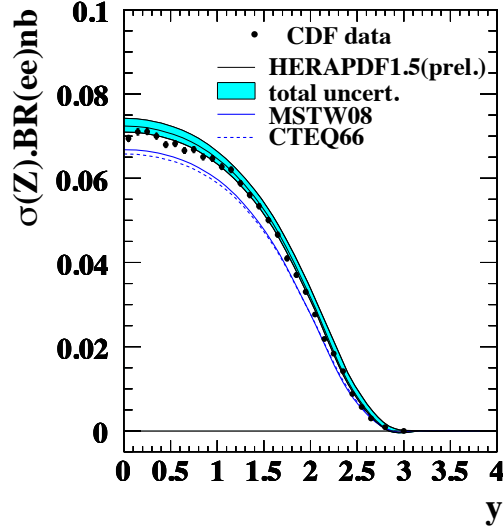


Figure 5: Z boson cross section times the branching ratio to electrons as a function of the rapidity y , as measured by the CDF collaboration. The cross sections are compared to the theory predictions based on three different PDFs. The total PDF uncertainties on the prediction on HERAPDF1.5 is shown as a blue band.

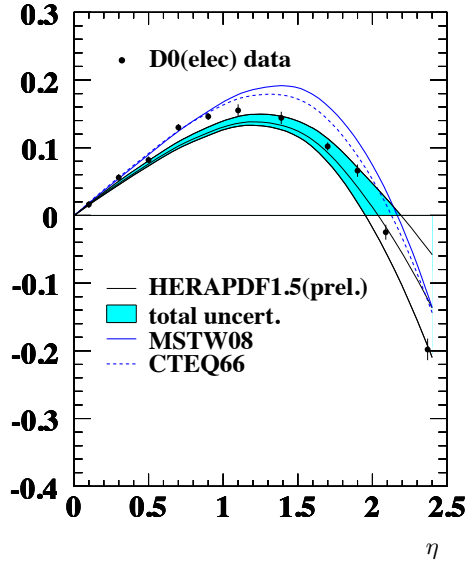


Figure 6: The electron charge asymmetry as a function of the lepton pseudorapidity η , as measured by the D0 collaboration. The measurement is compared to the theory predictions based on three different PDFs. The total PDF uncertainties on the prediction on HERAPDF1.5 is shown as a blue band.

W boson lepton asymmetry, based on the same PDFs are compared to the D0 measurement in figure 6. As can be seen from these figures, the predictions with the HERAPDF sets are in good agreement with data and the agreement is comparable to that obtained for the CTEQ66 and MRSTW08 PDF sets.

Recently many new physics results are released by the LHC experiments. The high precision LHC data can be used not only to compare with theory predictions to differentiate between different PDF sets but can also be used to put constraints on them instead. For example, the W lepton charge asymmetry A can help to constrain the u to d quark ratio ($A_w = (W^+ - W^-)/(W^+ + W^-)$ can be approximated to the ratio $A_W \approx (u_v - d_v)/(u_v + d_v + 2u_{sea})$ and thus is sensitive to the valence u and d quark ratio), the top quark measurement would have impact on gluons in the high- x region, inclusive jet measurements can be used to improve constraints on the gluon (the gg initial state has dominant contribution to the cross section at the low jet transverse momenta, P_T).

The W boson lepton asymmetry as measured by ATLAS, CMS and LHCb is shown in figure 7. The measured asymmetry is compared to predictions obtained with MSTW08, CTEQ66 and HERAPDF1.0 PDFs. The PDF uncertainties are shown as bands for each prediction. There is already a more precise W asymmetry preliminary measurement available from the CMS collaboration [11].

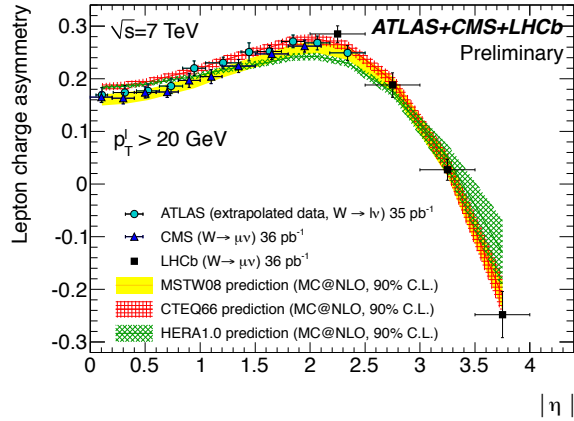


Figure 7: The W lepton charge asymmetry measured by ATLAS, CMS and LHCb (symbols) compared to the predictions based on different PDFs shown as lines with uncertainty bands.

The inclusive jet cross section is sensitive to the gluon density and can help to constrain it. In figure 8 the inclusive jet cross section measured by the ATLAS collaboration is shown as a function of the jet transverse momentum in different rapidity regions. The theoretical predictions based on different PDF

sets are describing the data equally well.

A similar measurement was performed by the CMS collaboration [12]. Figure 9 shows the inclusive jet cross section measured by CMS as a function of jet transverse momentum in the rapidity bin $0.0 \leq y \leq 0.5$. The measurement is compared to predictions calculated using different PDFs which are agreeing similarly well with the data as in the case of ATLAS measurement.

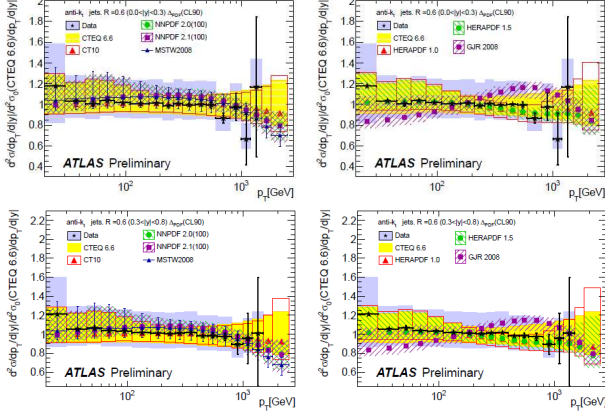


Figure 8: Inclusive double-differential jet cross sections as a function of jet p_T in different rapidity regions for jets identified using the anti- k_t algorithm with $R = 0.6$ (ATLAS). The plot shows the ratio of the data to the theoretical predictions calculated with different PDFs (CTEQ 6.6 is used as a reference).

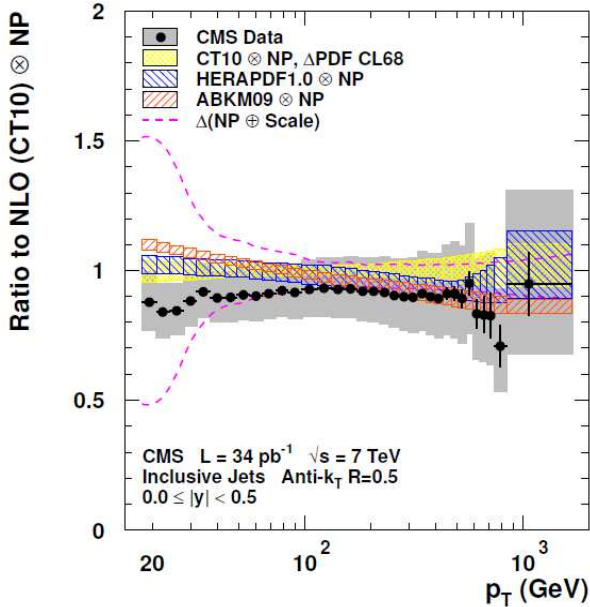


Figure 9: The ratio of inclusive jet cross section (CMS) for $|y| < 0.5$ as a function of jet p_T to NLO predictions based on the CT10 PDF set. Additional predictions are shown using HERAPDF1.0 and ABKM09 PDFs. PDF uncertainties are shown as coloured bands.

2.5. Recent Updates of HERAPDFs

The HERAPDF1.5 analysis has been extended to include HERA inclusive jet data. The new PDF set with jet data is called HERAPDF1.6 [13]. The jet data are sensitive to gluons and α_s and therefore reduce the correlation between the gluon density and α_s in the PDF fits. The impact of the jet data can be seen if $\alpha_s(M_Z)$ is treated as a free parameter in the fit (see figure 10 and 11). In case of fitting simultaneously PDFs and α_s to the inclusive DIS

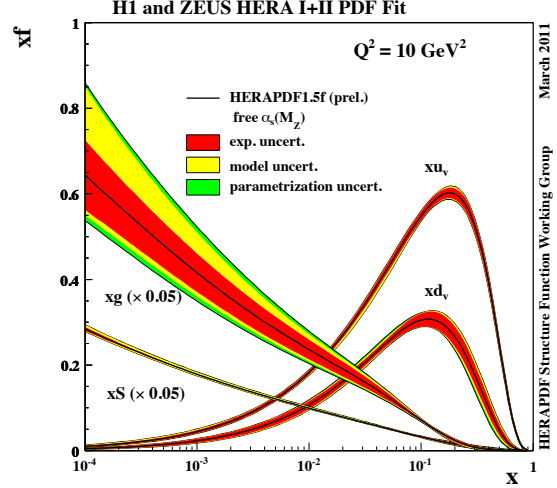


Figure 10: The parton distribution functions from HERAPDF1.5 at $Q^2 = 10 \text{ GeV}^2$ with $\alpha_s(M_Z)$ treated as a free parameter in the fit.

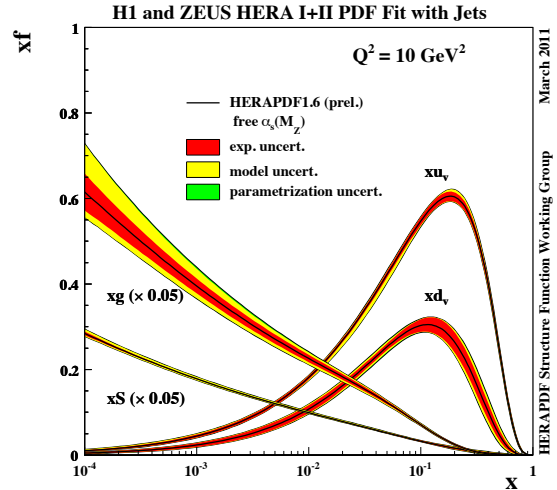


Figure 11: The parton distribution functions from HERAPDF1.6 (with HERA jet data included in the fit) at $Q^2 = 10 \text{ GeV}^2$ with $\alpha_s(M_Z)$ treated as a free parameter in the fit.

data the uncertainties on the gluon PDF becomes large at low x but as soon as the jet data are included (figure 11), the correlation between the gluon PDF and $\alpha_s(M_Z)$ is reduced, resulting in significantly reduced uncertainties on the gluon PDF similar to the case of using inclusive DIS data only and fixing $\alpha_s(M_Z)$. In the fit with jet data the value of α_s was extracted:

$$\alpha_s(M_Z) = 0.1202 \pm 0.0013(\text{exp}) \pm 0.0007(\text{mod/param}) \pm 0.0012(\text{hadronisation})^{+0.0045}_{-0.0036}(\text{scale}) \quad [13].$$

More details about QCD analysis with jets at HERA were presented in the talk by T. Schoerner-Sadenius [14].

The effect of including the HERA charm data in the QCD PDF fits was also studied at HERA. In this study, the charm data are used together with different implementations of the variable flavour number schemes (VFNS), which have different approaches for the interpolation function and counting of orders in α_s . Different VFN schemes can have different impacts on the charm contribution to the sea quark distribution and thus affect the composition of $x\bar{U}(x)$ from the $x\bar{u}(x)$ and the $x\bar{c}(x)$ contributions. Therefore the accuracy of the charm data influence the uncertainties on the W^\pm and Z cross section predictions at LHC. Figure 12 summarises the study by showing the charm quark mass m_c^{mod} (which is an effective parameter in the fit) scanning results for all schemes used in the fits. It is interesting to observe that first, different schemes have different optimal charm quark mass parameter m_c^{mod} , and second, the χ^2 minimum values are comparable for all schemes despite of different optimal values of m_c^{mod} .

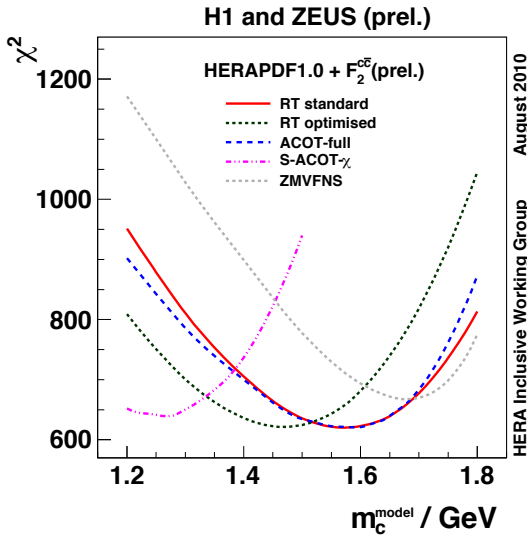


Figure 12: Comparison of the χ^2 of HERA I + F_2^{cc} fits using different heavy flavour schemes represented as lines of different styles.

The PDFs with optimal m_c^{mod} were then propagated to the calculation of W^\pm and Z boson cross section predictions for the LHC. As an example, the W^+ cross section as a function of m_c^{mod} for the different schemes is shown in figure 13. Good agreement between these predictions is observed at optimal m_c^{mod} which results in a reduction of the uncertainties due to the heavy flavour treatment to below 1.0%. More details about the study can be found in [15].

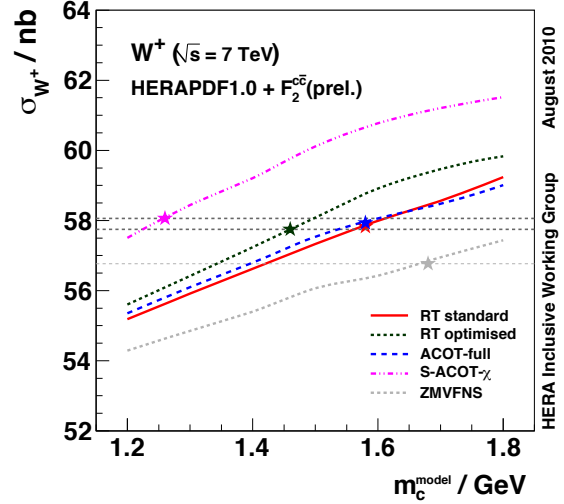


Figure 13: W^+ production cross section σ_{W^+} at the LHC for $\sqrt{s} = 7$ TeV as a function of m_c^{mod} . The lines show predictions for different VFN schemes as indicated by the legend. The stars show the predictions obtained with the optimal value of m_c^{mod} used in a given scheme. The thick dashed horizontal lines indicate the range of σ_{W^+} , determined for $m_c^{\text{mod}} = m_c^{\text{mod}}(\text{opt})$.

Finally, a QCD fit was performed which includes all currently available HERA NC and CC inclusive, charm, low energy and jet data (HERAPDF1.7). This fit prefers the same α_s value as obtained with a fit using only inclusive and jet data. Similarly, the same optimal m_C value was found in the HERAPDF1.7 fit as that obtained in a fit when only inclusive and charm data are fitted. The break-up of the HERAPDF1.7 PDFs into different flavours is illustrated in figure 14.

3. SUMMARY

Parton distribution functions of the proton are essential for any cross section prediction at hadron colliders. PDFs are generally determined by global QCD analyses fitting various sets of experimental data with the main knowledge about the proton structure coming from DIS experiments. Six modern PDF sets are

HERA I+II inclusive, jets, charm PDF Fit

 xf

$Q^2 = 10 \text{ GeV}^2$

- HERAPDF1.7 (prel.)
- exp. uncert.
- model uncert.
- parametrization uncert.
- HERAPDF1.6 (prel.)

 $xg (\times 0.05)$ $xS (\times 0.05)$ xu_v xd_v x

June 2011

HERAPDF Structure Function Working Group