Structure functions

C. Diaconu

Centre de Physique des Particules de Marseille and Deutsches Elektronen Synchrotron Hamburg

Recent progress in the understanding of the nucleon is presented. The unpolarised structure functions are obtained with unprecedented precision from the combined H1 and ZEUS data and are used to extract proton parton distribution functions via NLO QCD fits. The obtained parametrisation displays an improved precision, in particular at low Bjorken x, and leads to precise predictions of cross sections for LHC phenomena. Recent data from proton-antiproton collisions at Tevatron indicate further precise constraints at large Bjorken x. The flavour content of the proton is further studied using final states with charm and beauty in DIS ep and $p\bar{p}$ collisions. Data from polarised DIS or proton-proton collisions are used to test the spin structure of the proton and to constrain the polarised parton distributions.

1. Introduction

The progress in understanding the structure of baryonic matter has been driven by the continuous improvement in the experimental ability to probe the matter with an increasing precision. Particle collisions in which one of the particles is preserved after the interaction is one of the most precise methods. For each collision the quantity of momentum transfered from the probe to the composed particle can be measured. This momentum, denoted by Q, reflects the virtuality of the exchanged boson and corresponds to a wavelegth given by $\lambda[\text{fm}] = 0.2/Q[\text{MeV}]$. It gives the order of magnitude of the resolution with which the composed particle (e.g. the nucleon) is scrutinised in the collision - the higher the Q the finer the details probed in the interaction. A representation of the historical developments associated with the search for matter structure, from the discovery of the atomic nucleus in the classical Geiger-Marsden-Rutherford experiment, to the unique electron–proton collider HERA, is shown in figure 1 (left). For a momentum transfer of around 1 GeV the probed distances are comparable with the size of the proton. The inner structure of the proton was revealed in the first deep-inelastic scattering experiments at SLAC, 40 years ago. The "partons" were found and later identified with quarks and gluons within the framework of quantum chromodynamics (QCD). The "state of the art" is defined by the unique electron-proton collider HERA, where the centre-of-mass energy of 319 GeV allows the investigation of the proton with resolutions corresponding to 10^{-18} m.

The status of the proton structure investigations can also be represented in the plane spaned by Q^2 and the variable x, representing the fraction of the proton momentum carried by the interacting parton, as in figure 1 (right). The unique reach of HERA allows for the first time the investigation in the perturbative domain of the proton structure at very low x, down to 10^{-5}. This regime is particularly important for LHC physics, since the production of particles with moderate masses (a few hundred GeV's) may involve partons at low x. The measurements are complemented by the proton-antiproton collider Tevatron, where the investigations of W boson production constrain the proton structure at high x. Finally, the picture of the structure of the proton is also revealed in spin studies, performed using collisions of polarised beams. These studies aim to measure the various components contributing to the proton's spin and also provide independent tests of the strong interaction mechanisms.

2. Proton structure and partons distribution functions

The proton can be described in the infinite momentum frame as a collection of charged constituents (partons) that are "frozen" during the interaction time. In other words, in lepton-proton collisions, the time scale of the interaction between the incoming lepton and the proton's charged constituents is much shorter than the typical re-arrangement time of partons within the proton. Each interaction is characterised by the negative momentum transferred squared Q^2 and by the Bjorken x variable, defined as the fraction of the proton momentum carried by the interacting parton.
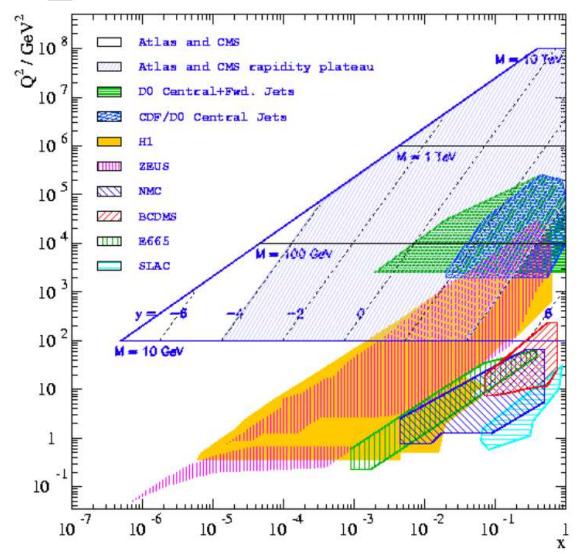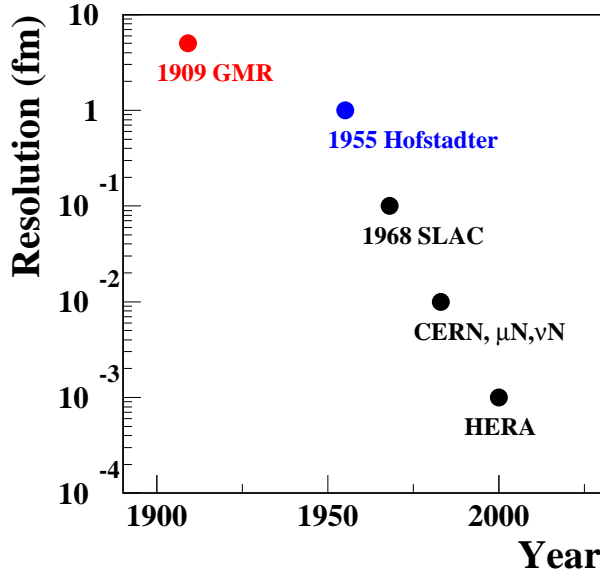
Figure 1: Left: The historical developments of the matter investigations. Right: The kinematic plane x-Q^2, with the regions covered by the fixed target experiments, and the HERA and Tevatron colliders. The domain probed at LHC is also indicated.

The understanding of high energy collisions involving protons within the QCD–improved quark parton model relies on the factorisation theorem, which states that the cross section is a convolution of a perturbative part, which describes the lepton-parton interaction at high energy, with a nonperturbative part that describes the momentum distributions of the proton inside the proton [1]. In the DIS case:

$$\sigma_{DIS}(x, Q^2) = \mathcal{C}_{DIS} \otimes f(x, Q^2) \tag{1}$$

The coeficient functions $\mathcal{C}_{DIS}$ are calculable at any order in QCD, while the parton distribution functions $f(x, Q^2)$ (PDFs) encapsulate the parton content of the proton and are non-perturbative by definition. Their form cannot be calculated within QCD, nevertheless the evolution in Q^2 (i.e. their derivatives as a function of Q^2) are predicted by QCD via the DGLAP equations [2]. If the PDFs are parametrised as a function of x at a given $Q^2 = Q_0^2$, the evolution equations allow a recalculation at any Q^2 and therefore a prediction for any (x, Q^2) point of the DIS cross section that depends only on the PDF x-parameterisation at Q_0^2.

The factorisation ansatz is usually also applied, though not completely demonstrated, for proton-proton or proton-antiproton collisions. In this case, the perturbative coefficient functions, describing the partonic interactions (between quarks and gluons) at high energies, are convoluted twice with the PDFs, corresponding to the two interacting (anti)protons. The interaction scale (the equivalent of Q^2 in DIS) can be given by the measurement of the final state. For instance, the cross section for W boson production in $p\bar{p}$ collisions at Tevatron is expressed as:

$$\sigma_{p\bar{p} \to WX} = \int dx_1 dx_2 \mathcal{C}_{partons \to W} \otimes f(x_1, M_W^2) \otimes f(x_2, M_W^2) \tag{2}$$

The level of experimental precision achieved, together with a precise calculation of the perturbative coefficients allow these measurements to further constrain the PDFs in the domain at large x, as will be described below.

Measurements in a large domain in x and Q^2, as indicated in figure 1(right), from DIS ep in fixed target and collisions at HERA, and from $p\bar{p}$ collisions at Tevatron are used to constrain the PDF parametrisation in a global fit. Each measurement at a given point (x, Q^2) is compared with the prediction obtained from a parameterisation of

the PDFs at a given scale Q_0^2, then evolved at the actual Q^2 using the DGLAP equations, and combined with the perturbative coefficient functions.

3. DIS cross sections and PDFs

The kinematics of the lepton-nucleon reaction is conveniently described by the variables x, Q^2 , already introduced, and y, the inelasticity, representing the energy transfer between the lepton and the hadron system. The inelasticity y is directly related to the lepton-parton scattering angle, and can be used to study the helicity structure of the interaction. The kinematic variables are related by $Q^2 = sxy$, where s is the lepton-proton centre-of-mass energy.

Neutral current (NC) and charged current (CC) processes are measured most recently by the H1 and ZEUS experiments in $e^\pm p$ collisions at HERA. The NC events contain a prominent electron and a jet of particles measured in the calorimeter, while in CC events only the jet is visible since the outgoing neutrino in not detected.

Since a large domain in x and Q^2 is accessed, the NC cross section become sensitive to weak effects, besides the photon exchange, which dominates the cross section. The Z^0 boson exchange can be incorporated into the so-called generalised structure functions. The cross section is parameterised as:

$$\frac{\mathrm{d}^2\sigma_{\text{NC}}^\pm}{\mathrm{d}x\mathrm{d}Q^2} = \frac{2\pi\alpha^2}{xQ^4}(Y_+\tilde{F}_2 \mp Y_- x\tilde{F}_3 - y^2\tilde{F}_L) \ , \tag{3}$$

The helicity dependence of the electroweak interactions is given by the terms $Y_\pm = 1 \pm (1 - y^2)$. The generalised structure functions $\tilde{F}_2$ and $x\tilde{F}_3$ can be further decomposed [3] as:

$$\tilde{F}_2 \equiv \ F_2 \ - v_e \ \frac{\kappa Q^2}{(Q^2 + M_Z^2)}F_2^{\gamma Z} \ + (v_e^2 + a_e^2)\left(\frac{\kappa Q^2}{Q^2 + M_Z^2}\right)^2 F_2^Z ,$$

$$x\tilde{F}_3 \equiv \quad - a_e \ \frac{\kappa Q^2}{(Q^2 + M_Z^2)}xF_3^{\gamma Z} + \ (2v_e a_e) \ \left(\frac{\kappa Q^2}{Q^2 + M_Z^2}\right)^2 xF_3^Z ,$$

with $\kappa^{-1} = 4\frac{M_W^2}{M_Z^2}(1 - \frac{M_W^2}{M_Z^2})$ in the on-mass-shell scheme. The quantities v_e and a_e are the vector and axial-vector weak couplings of the electron or positron to the Z^0. The electromagnetic structure function F_2 originates from photon exchange only and dominates in most of the accessible phase space.

The functions F_2^Z and xF_3^Z are the contributions to $\tilde{F}_2$ and $x\tilde{F}_3$ from Z^0 exchange and the functions $F_2^{\gamma Z}$ and $xF_3^{\gamma Z}$ are the contributions from γZ interference. These contributions are significant only at high Q^2. For longitudinally unpolarised lepton beams the $\tilde{F}_2$ contribution is the same for e^- and for e^+ scattering, while the $x\tilde{F}_3$ contribution changes sign as can be seen in eq. 3. The longitudinal structure function may be decomposed in a manner similar to $\tilde{F}_2$. Its contribution is significant only at high y.

In the quark parton model (QPM) the structure functions F_2, $F_2^{\gamma Z}$ and F_2^Z are related to the sum of the quark and anti-quark momentum distributions, $xq(x, Q^2)$ and $x\overline{q}(x, Q^2)$:

$$[F_2, F_2^{\gamma Z}, F_2^Z] = x\sum_q [e_q^2, 2e_q v_q, v_q^2 + a_q^2]\{q + \overline{q}\} \tag{4}$$

and the structure functions $xF_3^{\gamma Z}$ and xF_3^Z to the difference, which determines the valence quark distributions, $xq_v(x, Q^2)$:

$$[xF_3^{\gamma Z}, xF_3^Z] = 2x\sum_q [e_q a_q, v_q a_q]\{q - \overline{q}\} = 2x \sum_{q=u,d} [e_q a_q, v_q a_q]q_v \ . \tag{5}$$

In equations 4 and 5, v_q and a_q are the vector and axial-vector weak coupling constants of the quarks to the Z^0, respectively.
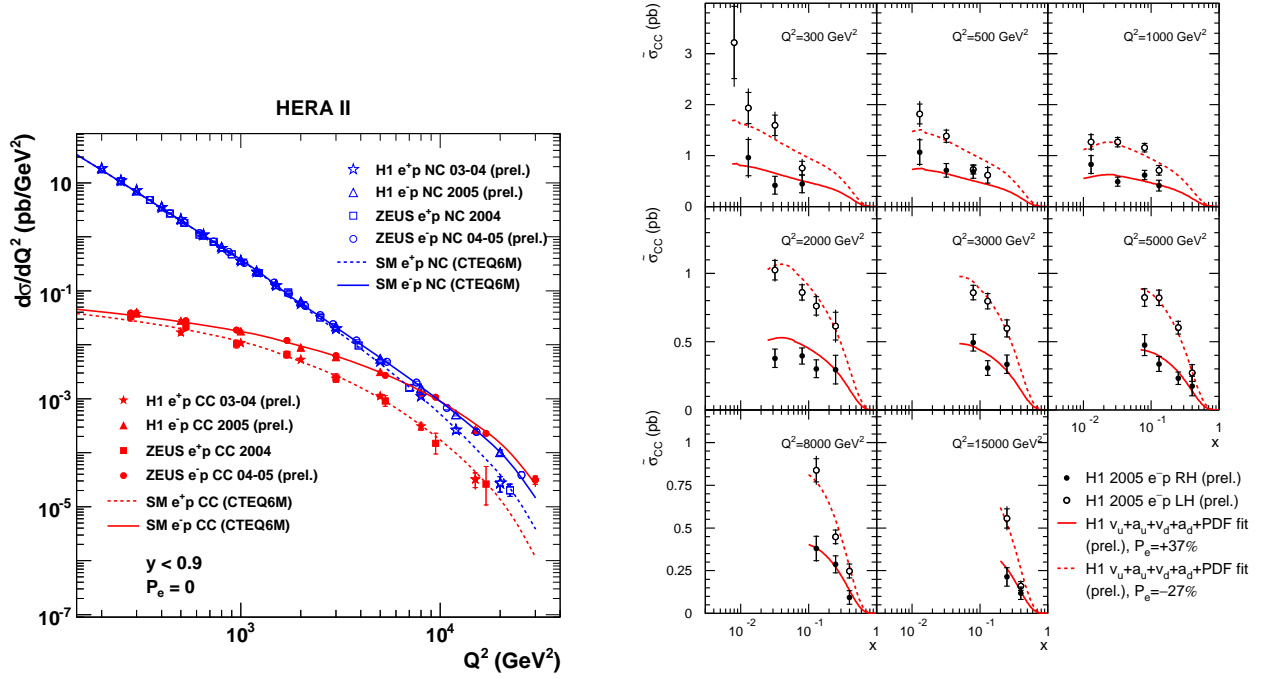
Figure 2: Left: the neutral current and charged current cross section as a function of Q^2 measured in $e^\pm p$ collisions at HERA. Right: the charged current reduced cross section $\tilde{\sigma}_{CC}$ as a function of x for various Q^2 values measured in electron–proton collisions with different electron beam polarisations.

The charged current (CC) interactions, $e^\pm p \to \overset{(-)}{\nu}_e X$, are mediated by the exchange of a W boson in the t channel. The cross section is parametrised as:

$$\frac{d^2\sigma^{\mathrm{CC}}(e^\pm p)}{dx dQ^2} = \frac{G_F^2}{2\pi x} \left[\frac{M_W^2}{M_W^2 + Q^2} \right]^2 \tilde{\sigma}_{CC}^\pm(x, Q^2),$$

$$\text{with} \quad \tilde{\sigma}_{CC}^\pm(x, Q^2) = \frac{1}{2} \left[Y_+ W_2^\pm(x, Q^2) \mp Y_- xW_3^\pm(x, Q^2) - y^2 W_L^\pm(x, Q^2) \right].$$

The reduced cross section is denoted by $\tilde{\sigma}$ is the reduced cross section, G_F is the Fermi constant, M_W, the mass of the W boson, and W_2, xW_3 and W_L, CC structure functions. In the QPM the structure functions $W_2^\pm$ and $xW_3^\pm$ may be interpreted as lepton charge dependent sums and differences of quark and anti-quark distributions:

$$W_2^+ = x(\overline{U} + D), \quad xW_3^+ = x(D - \overline{U}), \quad W_2^- = x(U + \overline{D}), \quad xW_3^- = x(U - \overline{D}), \tag{6}$$

whereas $W_L^\pm = 0$. The terms xU, xD, $x\overline{U}$ and $x\overline{D}$ are defined as the sum of up-type, of down-type and of their anti-quark-type distributions, i.e. below the b quark mass threshold: $xU = x(u + c)$, $xD = x(d + s)$, $x\overline{U} = x(\overline{u} + \overline{c})$, and $x\overline{D} = x(\overline{d} + \overline{s})$.

The differential NC and CC cross sections as a function of Q^2 are shown in figure 2 (left) for $e^- p$ and $e^+ p$ collisions at HERA. At low Q^2 the NC cross section, driven by the electromagnetic interaction, is two orders of magnitude larger than the CC cross section and corresponds to a pure weak interaction. At large $Q^2 \sim M_{W,Z}^2$ the two cross sections are similar. The largest Q^2 measurement corresponds to a resolution of $\delta \simeq 10^{-18}$ m, i.e. 1/1000 of the proton size. The agreement between the measurement and the prediction based on QCD improved parton model suggests no evidence for quark substructure.

The double differential reduced cross section $\tilde{\sigma}_{CC}(x, Q^2)$ is shown in figure 2 (right). Two data sets are shown, with different polarisations of the electron beam. The enhancement of the cross section as a function of the beam
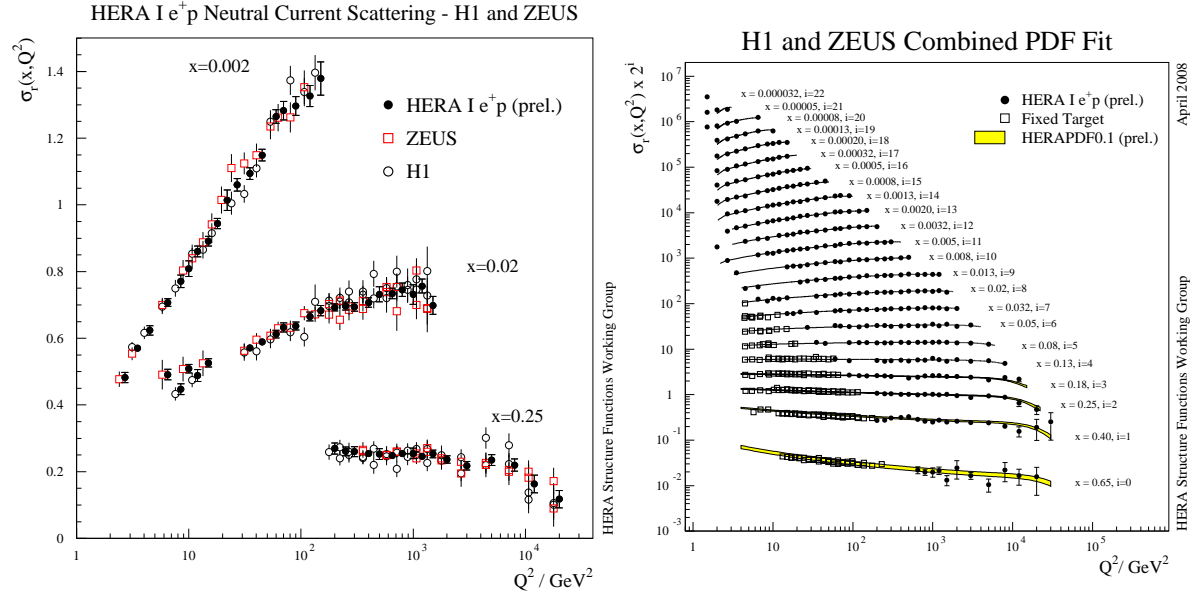
Figure 3: Left: Examples of cross section combinations for three bins in x as a function of Q^2. The precision of the combined experimental points is dramatically improved versus the individual measurements. Right: The combined HERA I data set compared with the HERAPDF 0.1 fit.

polarisation $\sigma_P^{e\pm p} = (1 \pm P_e)\sigma_{nopol}$ is visible. In addition, the CC processes are sensitive to individual quark flavours, which is especially visible at large Q^2: the e^+p collisions probe the $d(x)$ quark distribution, while e^-p are more sensitive to the $u(x)$. This is a very useful feature of the CC processes compared to the NC, where the flavour separation is weaker and occurs only at high Q^2 where the Z boson exchange is important.

4. H1 and ZEUS cross section combinations

The cross section measurements are generally affected by statistical and systematic uncertainties. The latter can also be correlated among various measurements, leading to common shifts or coherent change of shapes in the determined PDFs. These uncertainties and their correlations should be properly taken into account in the fit procedures, in order the propagate the experimental uncertainties of the PDFs determinations into predictions for cross sections, for instance at LHC. The fit procedure effectively combines the information of the same measurements (for instance the cross section in the same phase space or bin). However, this combination is done via the regular assumptions of such fits: input parameterisations, tolerances among different data sets and so on. Alternatively, the measurements of the same observables can be combined in model independent way using a coherent procedure [4] that takes into account the statistical and sytematic errors. The main advantage of such procedure is the proper treatment of the correlated and uncorrelated errors. The obtained data set therefore reflects the experimental knowledge in the given measurement phase space and is factorised from the fit procedures, which can be therefore better understood and optimised.

The average procedure is based on the minimisation of a χ^2 function, whose free parameters are not only the average values in each measurement point but also the correlated systematics shifts from each data set. The minimisation procedure is designed to find the best estimation of the local average measurements, also taking into account also long range systematic shifts that may occur due to the correlated systematic errors of the measurements. This averaging procedure achieve a significant reduction of the long-range correlated systematic errors, as illustrated in figure 3 for three sets of measurements at low values of x. The gain in precision is significant and better than the simple combination of errors in each point. This can be explained by the fact that for each measurement point a
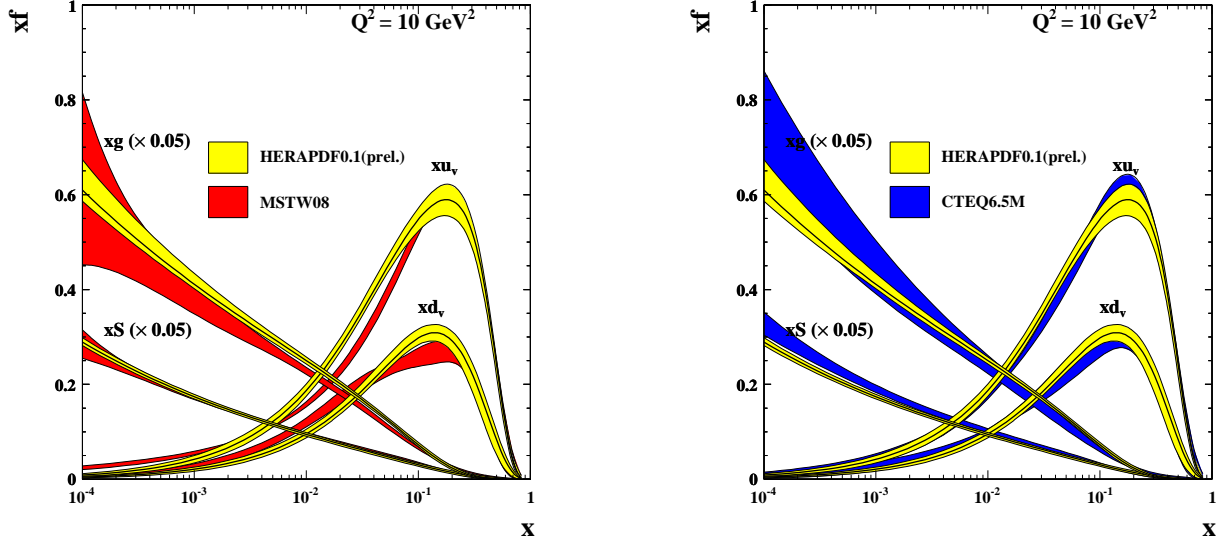
Figure 4: HERAPDF0.1 fit compared with MSTW and CTEQ fits.

given correlated uncertainty of one set of measurements is constrained by the independent measurements from other data sets. This cross-calibration leads to a precise and coherent HERA data set to be used to determine the PDFs.

5. QCD fits

As explained above, a variety of data from DIS and hadronic collisions can be used to constrain the PDFs in the so-called global fits. The approach is based on the factorisation assumption and provides PDF parameterisations that can be used for further predictions, for instance at LHC. The consistency of this procedure provides also a stringent test of the QCD. Recent progress has been reported in including Tevatron data and refining the error evaluations [5] and in detailed studies of the precision of the observables and its correlation with the PDFs [6]. Such approaches are the main source for PDFs parameterisations since they incorporate all available information on the proton structure.

However, it is well known that the global fits encouter and reveal tensions among various experimental data sets and, in absence of further experimental information, approximate statistical methods are used in order to cope with these inconsistencies. The most striking issue in the global fit procedures is the estimation of PDF errors, which has to be done with an somewhat arbitrary prescription of $\Delta\chi^2 \simeq 50 \div 100$ in order to accomodate systematic differences between various data sets. The PDF parametrisations obtained in the global fits using various approaches are in good agreement most of the time, but there are also some cases where the predictions differ by more than the quoted PDF uncertainty [7, 8]. In addition, the various improvements of the treatment of the theoretical input have lead in the past to variations of the predicted cross sections at LHC by significantly more than the previously cited uncertainties [9]. This situation raises the question of the theoretical errors associated with the extracted PDFs.

In these conditions, a useful approach is to use one coherent data set covering a significant area of the phase space, in order to study the QCD fit behavior, to compare to other data sets and to calculate predictions for observables at LHC. The newly obtained HERA combined data set has the advantage of very small systematic errors, which in fact no longer dominate. This makes the QCD fit procedure more robust and allows a test of the influence of the theoretical assumptions. The HERA combined preliminary data, described above, has been used by the H1 and ZEUS structure functions working group to perform a stand alone NLO QCD fit, HERAPDF 0.1 [10]. The consistency of the input data set simplifies the determination of the experimental uncertainties on the PDFs, performed using the
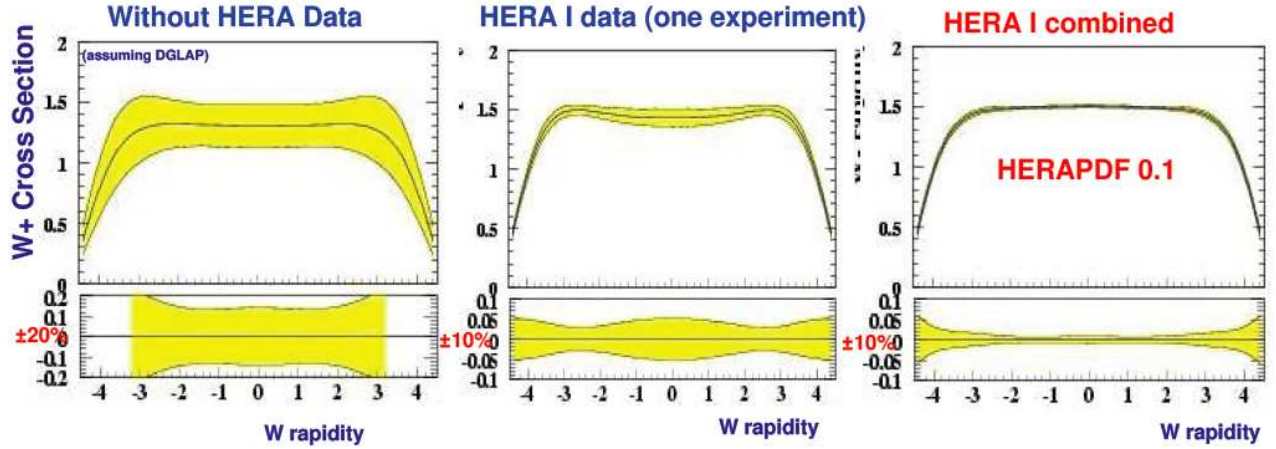
Figure 5: The improvement in precision obtained using HERA data: prior to HERA data (left), using one HERA experiment (center) and using HERAPDF 0.1 parametrisation (right). Note that pre-HERA data do not constrain W production at LHC and the precision shown in the left figure fully relies on parameterisation extrapolation from large x measurements from fixed target experiments.

more standard tolerance condition $\Delta\chi^2 = 1$.

In figure 3 (right), the NC DIS combined reduced cross section is compared to the HERAPDF 0.1 preliminary fit. The impressive precision of the combined data is reflected in the precision of the fit, which, extrapolated to lower Q^2, is found to be in very good agreement with the fixed target DIS data. The precision of the new data lead to precise PDFs, in particular at low x, where most of the cross-calibration that occurs during the averaging procedure dramatically improves the systematic errors of the combined cross section. The obtained parameterisations together with the estimated errors are compared in figure 4 to the global fits based on separate data sets.
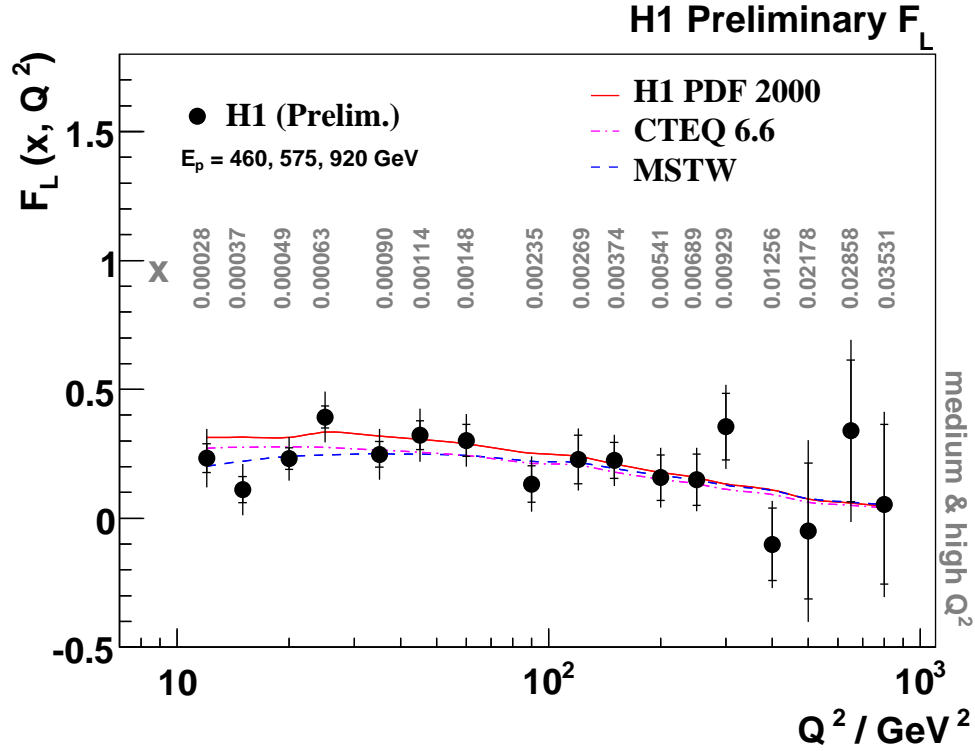
Although the observed improvement in precision is mostly due to the improved combined data, the fit procedure itself, which is similar to the global fits, contains an intrinsic artificial constraint of the estimated error due to the functional form assumed for the PDFs at the initial scale Q_0^2. Recent progress in the field of PDF parametrisation opens the way to an model independent assesement of the PDFs [11].

The parametrisation HERAPDF 0.1 is used to predict the W boson production in pp collisions at LHC. The improvement of the precision of this observable is shown in figure 5. The figure [12] displays the unique role of HERA for LHC physics and the significant gain in precision obtained recently from the H1 and ZEUS combined data. More HERA data, in particular at high x and Q^2 and from high statistics jet production analysis, are expected in the near future. They will further improve the precision of PDFs and are therefore an essential input for the LHC physics.

6. The direct measurement of the structure function F_L at low x

The longitudinal structure function F_L is a fundamental form factor of the proton, as can be seen from equation 3. Its contribution to the cross section is usually small, and only visible at large y where the weighting factor approaches the unity. Nevertheless, the structure function F_L is of particular importance. In the naive quark-parton model, the longitudinal structure function $F_L = F_2 - 2xF_1 \equiv 0$ (the Callan-Gross relation) reflecting the fact that longitudinal photons with helicity ± 1 cannot absorb spin 1/2 quarks. This is possible only if a gluon radiation allows a spin flip of the interacting quarks. It can be shown that F_L is directly related to the gluon density in the proton [13, 14] $xg(x) = 1.8[\frac{3\pi}{2\alpha_s}F_L(0.4x) - F_2(0.8x)] \simeq \frac{8.3}{\alpha_s}F_L$ meaning that at low x, to a good approximation F_L is a direct measure for the gluon distribution.

A direct measurement of F_L can be performed if the cross section $\sigma(E_p) \sim F_2(x, Q^2) + f(y) F_L(x, Q^2)$ is measured

Figure 6: The measurement of the longitudinal structure function F_L at HERA.

at fixed x and Q^2 but variable y. This can only be performed if the collision energy $\sqrt{s}$ is varied, for instance by reducing the proton beam energy at HERA. Then $F_L(x, Q^2)$ can be directly measured with reduced uncertainties from the difference of cross sections: $F_L \sim C(y) * (\sigma(E_p^1) - \sigma(E_p^2))$. The data taking at lower proton energies took place at the end of the HERA run in 2007. The data collected at proton energies of 460, 575 and 920 GeV have been used to perform the first direct measurement of F_L in the low x regime. The first measurements in the medium Q^2 range from 12 GeV^2 to 90 GeV^2 were published recently by the H1 collaboration [15]. Preliminary measurements have been performed by H1 [16] in an extended range of Q^2 up to 800 GeV^2. The H1 measurements averaged in x for each bin in Q^2 are shown in figure 6. The ZEUS collaboration presented measurements in the Q^2 range from 24 to 110 GeV^2 [17]. The measured F_L is significantly different from 0 at lowest Q^2 and are in agreement with the theoretical predictions based on various PDF parameterisations. It should be noted that in the presented phase space the gluon density and therefore the F_L predictions are well (but indirectly) constrained by other (inclusive) measurements. The measurements at lower $Q^2 < 12$ GeV^2 are experimentally challenging, but will be explored soon at HERA. They will shed light on various theoretical approaches predicting different values for F_L at low x [18] and together with other measurement in this kinematic regime will constraint even further the gluon component of the proton.

7. Constraints from measurements in $p\bar{p}$ collisions at Tevatron

Precise measurements of cross sections in $p\bar{p}$ collisions for processes well understood theoretically, can also be used to constrain the PDFs, in particular in the high x regime. Two such recent measurements are described here: the weak boson and jet production.

Events with jets at Tevatron are produced by colliding quarks and gluons with high values of x. A recent measurement of the inclusive jet production was performed using a large statistics sample of CDF run II data [19],
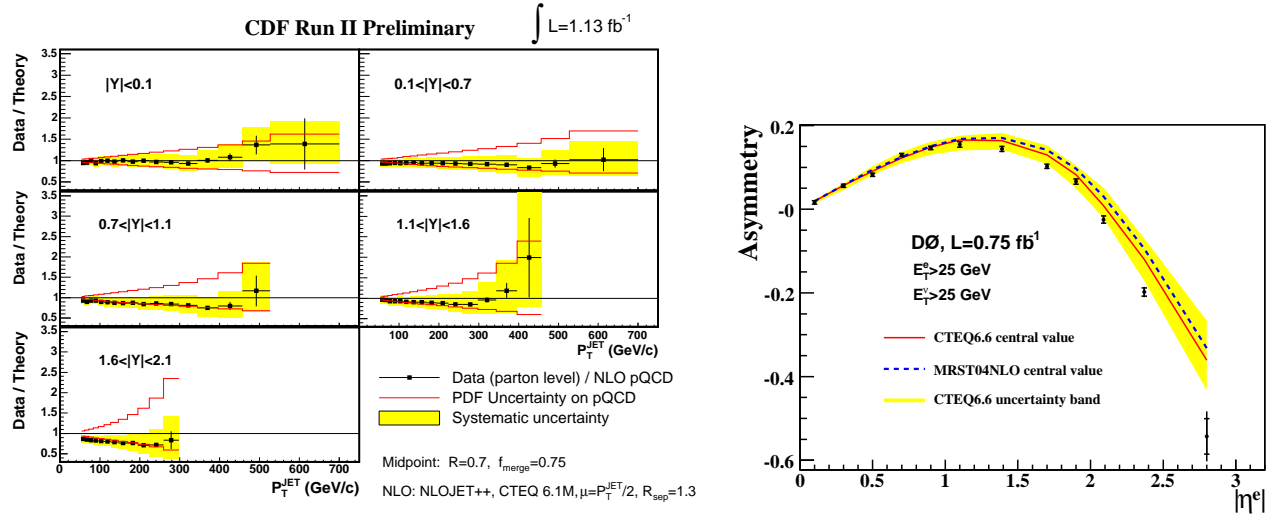
Figure 7: Left: Jet production cross section. Right: W production asymmetry.

corresponding to 1.1 fb^{-1}, using the cone-based midpoint jet clustering algorithm in the rapidity region of $|y| < 2.1$ for jets with transverse momenta in the range from 60 GeV to 700 GeV. A similar measurement, using a k_T jet clustering algorithm, was published by the D0 collaboration [20] for jet rapidities in the range -2.4 to 2.4 in the transverse momentum range from 50 GeV to 600 GeV. The results are consistent with next-to-leading-order perturbative QCD predictions based on recent parton distribution functions. The dominant experimental uncertainty is related to the hadronic energy scale. Cross sections are measured as a function of the jet transverse momentum and in various regions of rapidity. The measurements are in good agreement with the NLO QCD prediction, as can be seen in figure 7, which shows results from D0. The main theoretical uncertainties originate from the PDFs and exceed the experimental precision.

The measurement of W boson production at Tevatron also provides a probe of the proton structure. $W^{\pm}$ bosons are produced in $p\bar{p}$ collisions predominantly by the fusion of $u(d)$ quarks from the proton and $\bar{d}(\bar{u})$ quarks from the anti-proton. As a simple example of the kinematics probed in this reaction at Tevatron and to compare with LHC, it can be noted that in order to produce a W boson at rest by the fusion of a u quark from the proton with an $\bar{d}$ quark from the antiproton, the two quarks should have $x = 0.2$, while in the same configuration, the W will be produced at rest at LHC by partons with $x = 0.005$.

The W^+ or W^- rapidity distributions reflect the quark momentum distribution inside the proton. As the u quark tends to carry a higher fraction of the proton's momentum than the d quark, the $W^+(W^-)$ is boosted, on average, in the proton (anti-proton) direction. An asymmetry can be defined such that the charge of the W boson is correlated to the proton or anti-proton beam directions. The W charge asymmetry is defined as

$$A(y_W) = (d\sigma(W+)/dy_W - d\sigma(W^-)/dy_W)/(d\sigma(W^+)/dy_W + d\sigma(W^-)/dy_W), \tag{7}$$

where y_W is the rapidity of the W bosons. In the leading-order parton model, $A(y_W)$ is given approximately by

$$A(y_W) = [u(x1)d(x2) - d(x1)u(x2)]/[u(x1)d(x2) + d(x1)u(x2)] \tag{8}$$

where $x_{1,2} = x_0 e^{\pm y_W}$ and $x_0 = M_W/\sqrt{s}$. Since the W four-vector cannot be unambigously reconstructed, a reweighting procedure is performed for the two reconstruction solutions obtained in each event by imposing the W mass constraint. The measurements were performed recently by both D0 and CDF experiments [21, 22], each using each more than 1 fb^{-1} of run II data . The result obtained by D0 is shown in figure 7(right). The lepton charge asymmetry in Z^0 boson production leads also, in a similar way, to constraints on PDFs and has been measured recently with good precision at Tevatron [23].
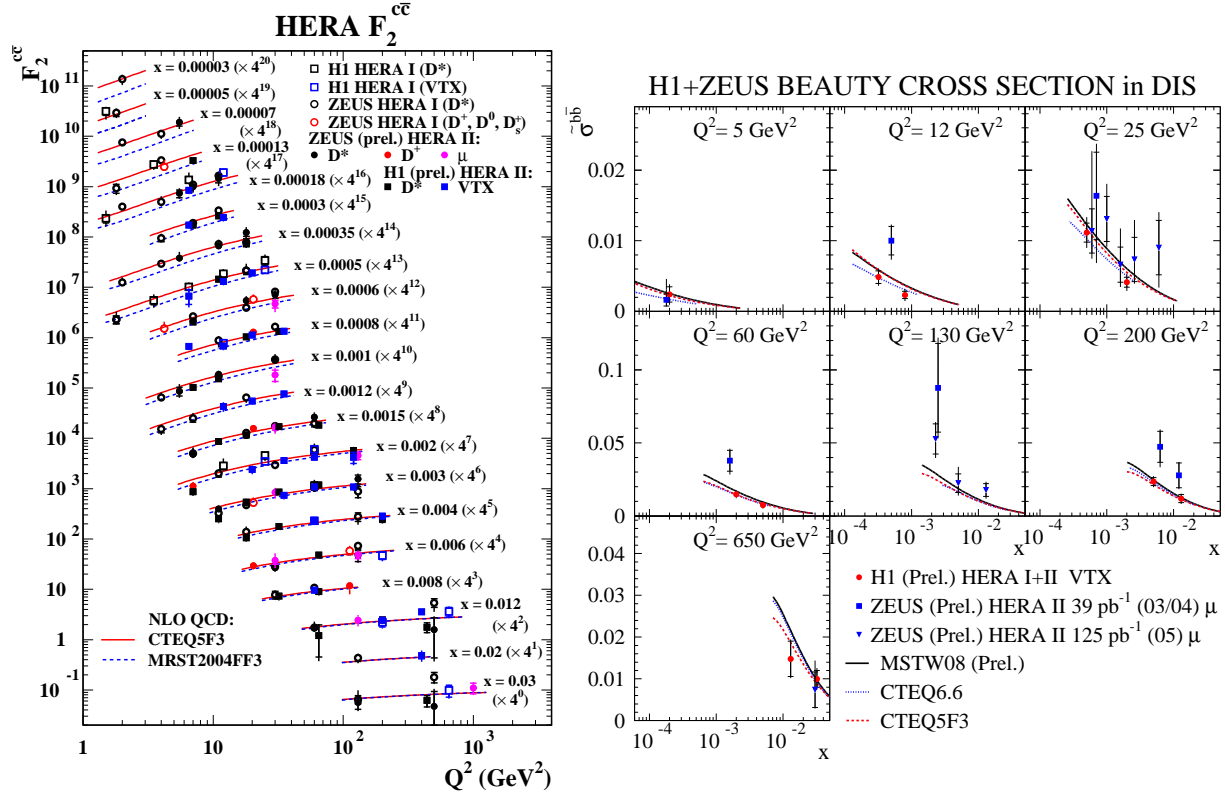
Figure 8: The charm structure function (left) and the reduced cross section for beauty production (right) measurements at HERA compared to theorectical prediction. The precision of the data, in particular that of the new measurements, is comparable or better with the difference among the various predictions and may lead to improvements in the proton structure function determination.

The data from Tevatron measurements of the inclusive jets and W and Z charge asymetries were used in the most recent MSTW global fit and lead to significant improvements, in particular for the gluon in the large x domain [5].

8. Charm and beauty from the proton

Of particular importance for the search of particles that couple to the mass at the LHC, especially for the Higgs studies, is the understanding of the heavy flavour content of the proton. In DIS, the interactions initiated by heavy quarks can be tagged by the presence of heavy flavour signals in the final state. Such interactions occur when gluons in the proton fluctuate to a pair of heavy quarks, one of which interacts with the photon (boson-gluon fusion), and therefore the study of heavy flavour interactions in DIS may also constrain the gluon content of the proton. The cross section for charmed tagged events can be written in a similar way as the inclusive DIS cross section, as expressed in equation 3. The main contribution to the charm (beauty) cross section originates from the structure function F_2^c (F_2^b). At leading order, the F_2^c and F_2^b structure functions are proportional to the momentum distribution of c and b quarks inside the proton.

Measurements of heavy flavour production in DIS are performed at HERA [24] using two established experimental techniques: the tagging of D^* mesons and the detection of particles emitted from long lived heavy hadrons. The measurements have been recently performed using the full available statistics from the HERA II run by the H1 collaboration. The production of D^* mesons has been mesured in the kinematic range of Q^2 from 5 GeV^2 to 1000 GeV^2 [25]. The measurements are in good agreement with the calculations based on NLO QCD and with LO Monte Carlo programs. An extrapolation of the measured cross section from the visible phase space to the full phase

10

space is needed in order to extract the structure functions and to constrain the heavy quarks PDFs. The extrapolation is typically larger than a factor two and is dependent on theoretical assumptions. Therefore, a consistent theoretical framework has to be used in order to obtain the structure functions. The F_2^c structure function based on those precise measurements has been extracted recently [27].

The measurement of processes with heavy quarks in the final state based on their lifetime use precise silicon tracking devices in order to identify tracks with significant distance of closest approach to the beam axis, potentially originating from heavy quarks decays. A recent measurement by the H1 collaboration [26] uses a multi-variate analysis and exploits the full data sets collected at HERA. A good agreement is obtained with the theoretical calculations, while the extrapolation factors to obtain the structure functions are closer to unity than in the case of D^* method.

A summary of the present status of the extraction of the heavy flavours structure functions F_2^c and F_2^b is presented in figure 8. The obtained precision is comparable to the theoretical calculation. The full potential of the HERA data will be obtained from the combination of H1 and ZEUS final data and may play an important role for contraints on the gluon and on the heavy flavour content of the proton.

Other measurements directly test the flavour aspects of the proton interactions. The measurement of events containing strange particles in the final state were performed at HERA [28, 29]. These measurements provide a new test of the QCD and allow a thorough investigations of strangeness production models. The link to the strange content of the proton is however difficult to establish since in the majority of events the strange particles are produced via the fragmentation process. Recent measurement at Tevatron [30] may also shed light on the strange content of the proton. The associated production of a W boson and a charm particle has been measured using soft muon tagging in a data sample corresponding to an integrated luminosity of 1.8 fb^{-1}. The measured process is described at leading order by the scattering of a gluon predominanlty with a strange quark. The measurement is found to be in good agreement with the theoretical prediction, but is so far statistically limited. A precision of about 15% on the strange quark content of the proton can be achieved with $\sim$ 6 fb^{-1} of data, expected by the end of Tevatron run II.

9. Study of the nucleon spin in polarised ep collisions

The understanding of the nucleon structure cannot be complete without the understanding of the mechanism by which the quarks and the gluons compose the spin of the nucleon $S_z = 1/2$. The various contributions to the nucleon spin can be expressed as following:

$$S_z = \frac{1}{2} = \frac{1}{2}\Delta\Sigma(\mu^2) + \Delta g(\mu^2) + L_z^q(\mu^2) + L_z^g(\mu^2)\,. \tag{9}$$

Here $\Delta\Sigma$ (Δg) describes the integrated contribution of quark and anti-quark (gluon) helicities to the nucleon helicity and L_z^q (L_z^g) is the z component of the orbital angular momentum among all quarks (gluons) at a given scale μ^2. The main puzzle has been the observation that, contrary to the naive expectation, the quark contribution does not account for the nucleon spin.

The spin-dependent DIS cross section can be parametrised by two structure functions g_1 and g_2, where g_2 is negligible and g_1 is given by:

$$g_1^{p,n}(x,Q^2) = \frac{1}{2}\sum_q e_q^2\left[\Delta q^{p,n}(x,Q^2) + \Delta\bar{q}^{p,n}(x,Q^2)\right]\,.$$

Here $\langle e^2\rangle = \sum_q e_q^2/N_q$ is the average squared charge of all involved quark flavours, and $\Delta q(x,Q^2) = q_{\Rightarrow}(x,Q^2) - q_{\Leftarrow}(x,Q^2)$ is the quark helicity distribution for massless quarks of flavour q in a longitudinally polarised nucleon in the "infinite momentum frame".

The structure function g_1 is related directly to the cross section difference: $\sigma_{LL} \equiv \frac{1}{2}(\sigma^{\overleftarrow{=}} - \sigma^{\overrightarrow{=}})/2$, where longitudinally ($L$) polarised leptons ($\rightarrow$) scatter on longitudinally (L) polarised nuclear targets with polarisation direction either parallel or anti-parallel ($\overrightarrow{\Rightarrow}$, $\overleftarrow{\Leftarrow}$) to the spin direction of the beam. The relationship to spin structure functions
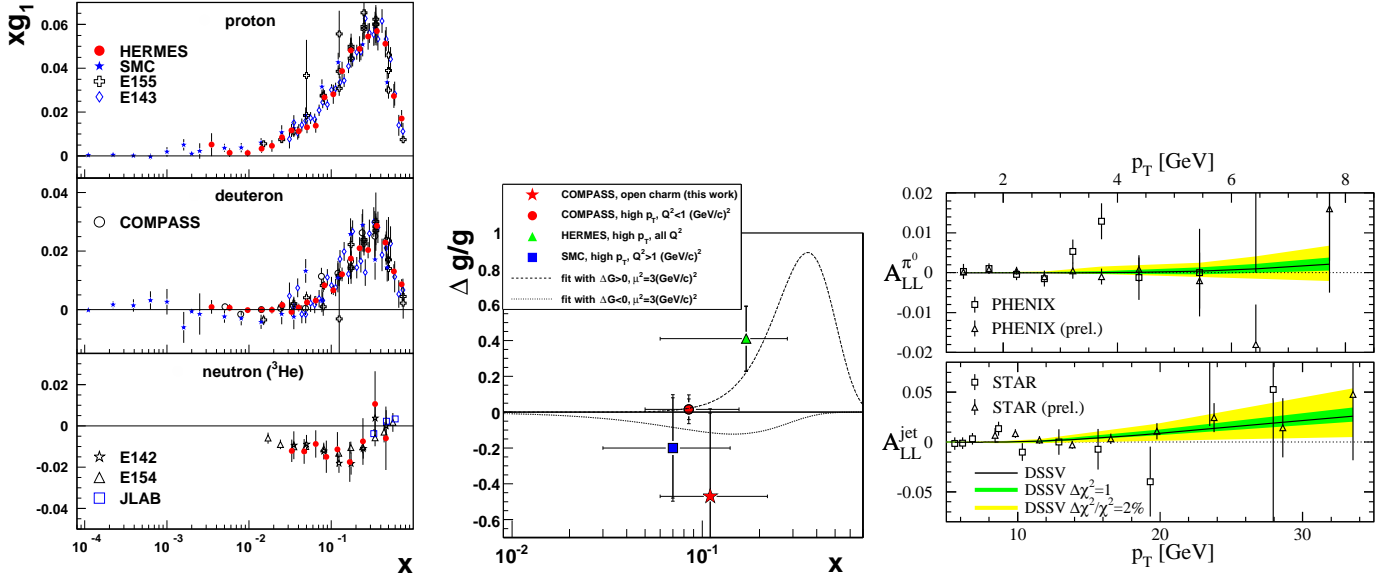
Figure 9: Left: The measured structure function g_1 for proton, deuteron and neutron. The precise measurement from HERMES is compared with measurements from other experiments. Center: The estimations of the gluon contributions $\Delta g/g$ (from [32]). Right: Cross sections spin asymmetries for π^0 and jet production in polarised pp collisions at RHIC, compared with a global fit.

is:

$$\frac{\mathrm{d}^2\sigma_{LL}(x,Q^2)}{\mathrm{d}x\,\mathrm{d}Q^2} = \frac{8\pi\alpha^2 y}{Q^4} \times \left[\left(1 - \frac{y}{2} - \frac{y^2}{4}\gamma^2\right) g_1(x,Q^2) - \frac{y}{2}\gamma^2\, g_2(x,Q^2) \right] , \tag{10}$$

where $\gamma^2 = Q^2/\nu^2$, with ν the energy of the virtual photon in the target rest frame.

Measurements of g_1 for the proton, deuteron and neutron are shown in figure 9(left) (from [31]). They can be used to extract the contribution of sea and valence quarks to the proton spin. Within some theoretical assumptions, this contribution is found to be $\Delta\Sigma_{(Q^2=5\ \mathrm{GeV}^2)} \simeq \int dx g_1(x,Q^2) \simeq 0.33$, which leaves a significant fraction for the gluon contribution to the proton spin.

The gluon contribution to the proton spin can in principle be extracted, similarly to the unpolarised case, from the scaling violations of the polarised structure functions. Unfortunately, in the phase space accessed by the fixed target experiments - relatively large x and low Q^2 - scaling violations are expected to be small and the precision of the data limit this approach. An alternative method is based on the identification of hadrons in the final state with significant transverse momentum, enriching the selected sample in events initiated by direct gluon-photon interactions and thereby enhancing the sensitivity of the analysis to the gluon polarisation $\Delta g/g$. A summary of the various analyses leading to estimations of the gluon contributions $\Delta g/g$ is presented in figure 9 (center, from [32]). The precision of these extractions are still limited.

The spin of the proton is also investigated in polarised pp collisions at RHIC [33]. Recent measurements of the cross section spin asymmetries for processes with inclusive jets in the final state open the way to new constraints, in particular for the gluon polarisation. The correlation between the observed jets allows for an approximate reconstruction of the kinematics. The measurement of the spin asymmetry as a function of the jet transverse momentum allows the investigation of the integrated gluon contribution to the proton spin $\Delta G(Q^2) = \int dx \Delta g(x,Q^2)$.

Global fits have been performed recently with the aim of determining the polarised PDFs [34] using measurements from polarised inclusive and semi-inclusive DIS and from polarised proton-proton collisions. The analysis shows that the RHIC data have a good sensitivity to Δg. The measurement of the double helicity asymmetry $A_{LL} =$

$(\sigma_{++} - \sigma_{+-})/\sigma_{++} - \sigma_{+-})$ is compared with the fit in figure 9(right). The fit indicates a positive Δg in a restricted region of x. The global fits are expected to also incorporate the analysis of final states with photons and W bosons. These analyses have a significant potential to enhance the sensitivity of polarised pp collision data to the gluon polarisation. In addition, the data collected from polarised pp collisions at higher energy $\sqrt{s_{pp}} = 500$ GeV and lower energy $\sqrt{s_{pp}} = 64$ GeV [35] will allow to explore a larger x domain, in particular at low x where Δg is unconstrained at present. The new measurements together with the expected refinements of the global fit procedures are expected to significantly improve the knowledge of polarised PDFs in the future.

10. Outlook

The research of nucleon structure is the scene of fast progress. Precise data from HERA constrain the parton distribution functions to an unprecedented level and provide solid support for predictions at the LHC in most of the phase space. The picture is complemented by Tevatron data at large x. The global fits exploit - and also prove - the universality of the parton distribution functions, combining the most precise results, the yield of which will increase in the years to come. Knowledge of the proton structure is profoundly enriched by spin studies in polarised collisions. Future proposed projects at higher energies like LHeC [38] and EIC [39], have the potential sharpen significantly the precision of nucleon structure investigations and to bring this fundamental research area into new paradigms.

References

[1] J. C. Collins, D. E. Soper and G. Sterman, Adv. Ser. Direct. High Energy Phys. **5** (1988) 1 [arXiv:hep-ph/0409313].

[2] V.N. Gribov and L.N. Lipatov, Sov. J. Nucl. Phys. 15 (1972) 438, 675; L.N. Lipatov, Sov. J. Nucl. Phys. **20** (1975) 94; Yu. L. Dokshitzer, Sov. Phys. JETP **46** (1977) 641; G. Altarelli and G. Parisi, Nucl. Phys. **B126** (1977) 298.

[3] M. Klein and T. Riemann, Z. Phys. C **24** (1984) 151.

[4] A. Glazov, AIP Conf. Proc. **792** (2005) 237.

[5] G. Watt, A. D. Martin, W. J. Stirling and R. S. Thorne, arXiv:0806.4890 [hep-ph].

[6] P. M. Nadolsky *et al.*, Phys. Rev. D **78** (2008) 013004 [arXiv:0802.0007 [hep-ph]].

[7] N. E. Adam, V. Halyo and S. A. Yost, JHEP **0805** (2008) 062 [arXiv:0802.3251 [hep-ph]].

[8] A. Djouadi and S. Ferrag, Phys. Lett. B **586** (2004) 345 [arXiv:hep-ph/0310209].

[9] W. K. Tung, H. L. Lai, A. Belyaev, J. Pumplin, D. Stump and C. P. Yuan, JHEP **0702** (2007) 053 [arXiv:hep-ph/0611254].

[10] H1 and ZEUS Collaborations, Preliminary Result, to be found at:
https://www-h1.desy.de/h1/www/publications/htmlsplit/H1prelim-08-045.long.html

[11] L. Del Debbio *et. al* [NNPDF Collaboration], JHEP **0503** (2005) 080 [arXiv:hep-ph/0501067;

R. D. Ball *et al.* [NNPDF Collaboration], Nucl. Phys. B **809** (2009) 1 [arXiv:0808.1231 [hep-ph]].

[12] A.M. Cooper-Sarkar, private communication.

[13] G. Altarelli and G. Martinelli, Phys. Lett. B **76** (1978) 89.

[14] A. M. Cooper-Sarkar *et al.*, Z. Phys. C **39** (1988) 281.

[15] F. D. Aaron *et al.* [H1 Collaboration], Phys. Lett. B **665** (2008) 139 [arXiv:0805.2809 [hep-ex]].

[16] H1 Collaboration, Preliminary Result, to be found at:
https://www-h1.desy.de/h1/www/publications/htmlsplit/H1prelim-08-042.long.html

[17] ZEUS Collaboration, Preliminary Results ZEUS-prel-08-001, to be found at:
http://www-zeus.desy.de/publications.php3

[18] R.S. Thorne, arXiv:0808.1845 [hep-ph]

[19] T. Aaltonen *et al.* [CDF Collaboration], Phys. Rev. D **78** (2008) 052006 [arXiv:0807.2204 [hep-ex]].

[20] V. M. Abazov *et al.* [D0 Collaboration], Phys. Rev. Lett. **101** (2008) 062001 [arXiv:0802.2400 [hep-ex]].

[21] V. M. Abazov *et al.* [D0 Collaboration], arXiv:0807.3367 [hep-ex].

[22] CDF Public Note, http://www-cdf.fnal.gov/physics/ewk/2007/WChargeAsym/PubNote.pdf

[23] CDF Public Note, http://www-cdf.fnal.gov/physics/ewk/2008/dszdy/public_note_022808.ps

[24] K. Kruger, for the H1 and ZEUS Collaborations, arXiv:0810.4019 [hep-ex].

[25] H1 Collaboration, Preliminary Results, to be found at
https://www-h1.desy.de/h1/www/publications/htmlsplit/H1prelim-08-072.long.html and
https://www-h1.desy.de/h1/www/publications/htmlsplit/H1prelim-08-074.long.html

[26] H1 Collaboration, Preliminary Result, to be found at
https://www-h1.desy.de/h1/www/publications/htmlsplit/H1prelim-08-172.long.html

[27] H1 Collaboration, Preliminary Results, to be found at
https://www-h1.desy.de/h1/www/publications/htmlsplit/H1prelim-08-173.long.html and
https://www-h1.desy.de/h1/www/publications/htmlsplit/H1prelim-08-174.long.html

[28] F.D. Aaron *et al.*, H1 Collaboration, arXiv:0810.4036 [hep-ex].

[29] S. Chekanov *et al.* [ZEUS Collaboration], Eur. Phys. J. C **51** (2007) 1 [arXiv:hep-ex/0612023].

[30] T. Aaltonen *et al.* [CDF Collaboration], Phys. Rev. Lett. **100** (2008) 091803 [arXiv:0711.2901 [hep-ex]].

[31] A. Airapetian *et al.* [HERMES Collaboration], Phys. Rev. D **75** (2007) 012007 [arXiv:hep-ex/0609039].

[32] M. Alekseev *et al.* [COMPASS Collaboration], arXiv:0802.3023 [hep-ex].

[33] F. Ellinghaus, for the PHENIX and STAR Collaborations, arXiv:0810.3178 [hep-ex].

[34] D. de Florian, R. Sassot, M. Stratmann and W. Vogelsang, Phys. Rev. Lett. **101** (2008) 072001 [arXiv:0804.0422 [hep-ph]].

[35] :. A. Adare [PHENIX Collaboration], arXiv:0810.0701 [hep-ex].

[36] A. Hillenbrand and f. t. H. Collaboration, arXiv:0810.3617 [hep-ex].

[37] M. Alekseev *et al.* [COMPASS Collaboration], arXiv:0802.2160 [hep-ex].

[38] J. B. Dainton, M. Klein, P. Newman, E. Perez and F. Willeke, JINST **1** (2006) P10001 [arXiv:hep-ex/0603016].

[39] A. Deshpande, R. Milner, R. Venugopalan and W. Vogelsang, Ann. Rev. Nucl. Part. Sci. **55** (2005) 165 [arXiv:hep-ph/0506148].