

Insights into nucleon structure from parton distributions

W. Melnitchouk*†

Jefferson Lab, Newport News, Virginia 23606, USA

E-mail: wmelnitc@jlab.org

We review recent progress in understanding the substructure of the nucleon from global QCD analysis of parton distribution functions (PDFs). New high-precision data on W -boson production in $p\bar{p}$ collisions have significantly reduced the uncertainty on the d/u PDF ratio at large values of x , indirectly constraining models of the medium modification of bound nucleons. Drell-Yan data from pp and pd scattering reveal new information on the $\bar{d} - \bar{u}$ asymmetry, clarifying the role of chiral symmetry breaking in the nucleon. In the strange sector, a new chiral SU(3) analysis finds a valence-like component of the strange-quark PDF, giving rise to a nontrivial $s - \bar{s}$ asymmetry at moderate x values. We also review recent analyses of charm in the nucleon, which have found conflicting indications of the size of the nonperturbative charm component.

*The 26th International Nuclear Physics Conference
11-16 September, 2016
Adelaide, Australia*

*Speaker.

†Plenary talk at INPC2016.

1. Historical perspective

The quest to understand nature in terms of fundamental building blocks dates back to the 5th century BC and the philosophical ideas of Leucippus and Democritus in ancient Greece. Fast forward some 23 centuries, and the experimental investigation of what lies at the heart of matter began in earnest with Rutherford's discovery of the atomic nucleus, followed shortly by his discovery of the proton. The field of nuclear physics can be said to have begun with Chadwick's observation of the neutron a decade and a half later. Understanding of the forces that hold the atomic nucleus together was advanced with the prediction and subsequent verification in the following decade of the π mesons as the mediators of the strong inter-nuclear force.

So more than 100 years after Rutherford's classic experiments, what do we now know about the nucleon? One of the seminal observations, made by Hofstadter *et al.* [1] in the study of the angular dependence of elastic electron-proton scattering cross sections, was that the proton is not pointlike but rather has a finite size. From the slope of the associated form factor versus the momentum transfer squared Q^2 , the charge radius was extracted to be 0.75×10^{-15} m, which remarkably is within 10% of the current best value [2]!

In the 1960s electron scattering experiments at larger Q^2 revealed the surprising feature that the total cross section, in the deep-inelastic scattering (DIS) region, does not fall with increasing Q^2 but remains approximately constant [3]. The natural explanation for this is that the scattering takes place from pointlike constituents of the nucleon — termed “partons” by Feynman [4] and identified with the quarks of Gell-Mann and Zweig [5]. In Feynman's parton model, the DIS cross sections, or structure functions, could be interpreted in terms of PDFs, which give probabilities for finding constituents in the nucleon with specific momentum fractions x .

With the development of Quantum Chromodynamics in the 1970s, the formal establishment of PDFs as universal, process-independent functions was made possible through the factorisation theorems of Collins, Soper and Sterman [6], which allow high-energy cross sections to be factorised into hard scattering partonic cross sections and nonperturbative matrix elements. The universality feature permits data from many different scattering processes to be analysed simultaneously, and parametrised by sets of spin-averaged $f = f^\uparrow + f^\downarrow$ and spin-dependent $\Delta f = f^\uparrow - f^\downarrow$ PDFs in global QCD analyses [7], where $f = q, \bar{q}, g$ denotes quarks, antiquarks or gluons, and the $\uparrow\downarrow$ represent the spin of the parton parallel or antiparallel to the spin of the nucleon.

In the 21st century, there are two main reasons to study PDFs. The first is the basic desire to understand the detailed quark and gluon structure of QCD bound states. In addition, precision PDFs are essential for the computation of backgrounds in searches for physics beyond the Standard Model in collider experiments — through the Q^2 evolution equations information from experiments at high x and low Q^2 , where the majority of existing data lie, feeds to the low- x and high- Q^2 region that characterizes the kinematics at the LHC. In the following, we will summarize the current knowledge of PDFs of the valence u and d quarks in the proton and neutron, as well as the light-quark, strange, and charm quark content of the quark-antiquark sea.

2. Valence quarks

Valence quarks accord the nucleon some of its essential global properties, such as its charge, isospin and baryon number. Knowledge of the valence quark PDFs in the nucleon, especially in the large- x region where they dominate, is vital for better understanding the nonperturbative flavor and spin dynamics of quarks, as well as testing predictions from perturbative and nonperturbative QCD models for the behavior of PDFs in the limit $x \rightarrow 1$ [8].

Recently, high-precision $D\bar{O}$ data on W^\pm production in $p\bar{p}$ collisions at Fermilab have led to a significant reduction in the uncertainties on the d/u ratio at large values of x . This is illustrated in Fig. 1 from the CJ15 global QCD analysis [9], which demonstrates the shrinking of the d/u uncertainty bands with the successive addition of various data sets. Compared with the fit to DIS only data, in which the d/u ratio has very large uncertainties beyond $x \approx 0.4$, the addition of the lepton asymmetries leads to more than a factor 2 reduction in d/u at $x \lesssim 0.4$, with more limited impact at higher x values due to the PDF smearing caused by the lepton decay vertex. Subsequent inclusion of the W asymmetries leads to a further halving of the uncertainty at $x \approx 0.6 - 0.8$, while having minimal effect on the errors at $x \lesssim 0.4$.

Independent of the W asymmetry data, a significant reduction of the d/u uncertainty at moderate x values is already provided by the Jefferson Lab F_2^n/F_2^d data from the BONuS experiment. While these have little at $x \lesssim 0.3$, the reduction in the d/u error at $x \sim 0.5 - 0.6$ is almost as large as that from the lepton asymmetries. Using all the available data from DIS and W boson production, the value of the d/u ratio extrapolated to $x = 1$ is [9]

$$d/u \xrightarrow{x \rightarrow 1} 0.09 \pm 0.03 \quad (2.1)$$

at the 90% CL at a scale $Q^2 = 2 \text{ GeV}^2$. This is somewhat smaller than the prediction from the helicity conservation model, in which the d/u ratio is expected to be ≈ 0.2 , but is larger than the zero limit found traditionally in analyses that do not take into account nuclear effects in the deuteron [8].

The discovery and determination of properties of new particles beyond the Standard Model at high-energy colliders depends on accurate knowledge of PDFs. This is especially true for the forward production of particles of mass m at large rapidities y , whose cross sections are given by products of PDFs with one evaluated at small $x_1 \approx (m/\sqrt{s})e^{-y}$ and the other at large $x_2 \approx (m/\sqrt{s})e^y$. The production of heavy W' and Z' bosons, for example, is sensitive to d -quark PDF uncertainties at high rapidities, exceeding 100% in the W'^- channel, which places limits on the accuracy of cross section measurements for masses near the kinematic thresholds [10].

Cross sections sensitive to large- x parton distributions typically fall rapidly with increasing x values, leading to limitations in the quantity and precision of experimental data and the kinematic range over which they can be obtained. This is illustrated by considering the production of a heavy W' boson as a function of the W' rapidity $y_{W'}$ [10]. Assuming Standard Model couplings, the parton luminosity for a produced negatively charged W'^- boson is given by

$$\mathcal{L}_{W'^-} = \frac{2\pi G_F}{3\sqrt{2}} x_1 x_2 \left[\cos^2 \theta_C (\bar{u}(x_2)d(x_1) + \bar{c}(x_2)s(x_1)) + \sin^2 \theta_C (\bar{u}(x_2)s(x_1) + \bar{c}(x_2)d(x_1)) \right] \\ + (x_1 \leftrightarrow x_2), \quad (2.2)$$

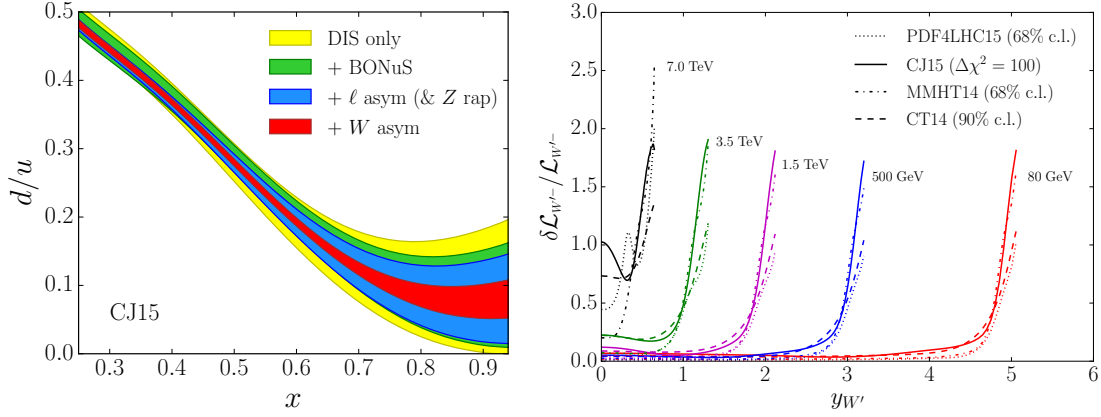


Figure 1: (Left) Impact of various data sets on the d/u ratio at $Q^2 = 10 \text{ GeV}^2$ [9]. The 90% CL uncertainty band is largest for the DIS only data, and decreases with the successive addition of JLab BONuS F_2^n/F_2^d data, lepton asymmetry (and Z rapidity) data, and W -boson asymmetry data. (Right) Relative uncertainty $\delta\mathcal{L}_{W'^-}/\mathcal{L}_{W'^-}$ in the W'^- luminosity as a function of rapidity $y_{W'}$ for the combined PDF4LHC15, CJ15, MMHT14 and CT14, for various W' masses from 80 GeV (Standard Model) to 7.0 TeV.

where G_F is the Fermi constant and θ_C the Cabibbo angle. The uncertainty $\delta\mathcal{L}_{W'^-}$ in the luminosity is shown in Fig. 1 for various PDF sets as a function of $y_{W'}$, for several fixed values of the boson mass from the Standard Model W up to 7 TeV.

Interestingly, at large $y_{W'}$ the relative uncertainties from the CJ15, MMHT14, CT14 and PDF4LHC15 PDF sets displayed in Fig. 1 almost coincide, even though the quoted confidence levels for each vary. For instance, the MMHT14 and PDF4LHC15 sets are for the 68% c.l., while CT14 corresponds to 90% c.l., and CJ15 assumes a $\Delta\chi^2 = 100$. Although this analysis is restricted to heavy vector bosons with Standard Model couplings, and the quantitative effects of the PDF uncertainties would be different in other models, it illustrates the point that caution must be exercised when using PDFs in regions where these are not directly constrained, or their uncertainties underestimated, as is the case at large x .

3. Light quark sea

In the sea-quark sector, at smaller values of x , PDFs offer important glimpses into nonperturbative QCD phenomena, such as chiral symmetry breaking and the role of the pion cloud of the nucleon. One of the most dramatic consequences of the nucleon's pion cloud has been in the flavor structure of the proton sea, first anticipated by Thomas [11] in the 1980s. The empirical observation of a \bar{d} excess over \bar{u} quarks a decade later firmly established the relevance of pions for understanding the partonic structure of the nucleon.

Building on these earlier observations, more recent studies have sought to develop the phenomenology of nonperturbative PDFs in the context of chiral effective theory, not just in terms of moments but also as a function of the parton momentum fraction x [12, 13]. While much of the attention has been focused on exploring the consequences of chiral symmetry breaking for the $\bar{d} - \bar{u}$ asymmetry in the proton, a complementary effort to reveal the dynamics of pion exchange

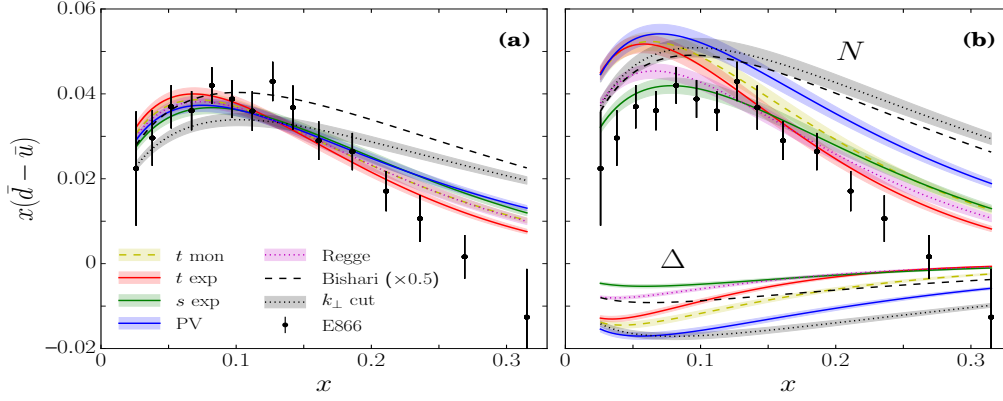


Figure 2: Comparison of the flavor asymmetry $x(\bar{d} - \bar{u})$ for (a) pion model fits for various regularization prescriptions with the empirical asymmetry extracted from the E866 Drell-Yan experiment [14], and (b) the individual (positive) nucleon and (negative) Δ contributions to the asymmetry. From McKenney *et al.* [15].

in high-energy processes was undertaken at HERA with the study of leading neutron production in semi-inclusive DIS on the proton. Here a forward moving neutron is produced in coincidence with the scattered lepton in the high-energy reaction $ep \rightarrow enX$, carrying a large fraction of the proton's energy. As well as identifying the characteristic features of pion exchange in the leading neutron production cross sections, the HERA leading neutron data have also been analyzed in view of extracting the structure function of the exchanged pion in the small- x region. Analyses of the HERA leading neutron data have generally been able to extract the shape of the pion structure function F_2^π , but have been unable to fix the normalization because of large uncertainties in the pion flux (or pion light-cone momentum distribution in the nucleon). Since the pionic contributions to the leading neutron cross sections depend on both the pion structure function and the pion probability in the proton, the HERA data by themselves have been insufficient to disentangle information on F_2^π independently of assumptions about the pion flux.

Recently, McKenney *et al.* [15] addressed the question of whether one can reduce the model dependence of F_2^π extracted from the HERA leading neutron data by using additional constraints from other observables that are sensitive to the pion flux. In particular, the data on the SU(2) flavor asymmetry $\bar{d} - \bar{u}$, particularly those from the E866 Drell-Yan experiment [14], provide the strongest indication of significant pion cloud effects in the nucleon, see Fig. 2. Because the E866 data are at relatively high x values compared with the HERA measurements, within the pion exchange framework they are sensitive to the pion PDFs at large x , where the PDFs are well determined from pion-nucleon Drell-Yan data. In contrast, the HERA data are taken at very low x , $10^{-4} \lesssim x \lesssim 10^{-2}$, outside of the region where the pion PDFs have been constrained. Within the pion exchange framework, the same pion flux should be applicable for both observables, which should then reduce the uncertainty in the extracted F_2^π at small x .

McKenney *et al.* [15] performed a simultaneous fit to both the HERA leading neutron and E866 $\bar{d} - \bar{u}$ asymmetry data, for a wide range of prescriptions adopted in the literature for regularizing the pion-nucleon amplitudes. They found that both the HERA and E866 data could be described within a one-pion exchange framework, if the cutoff parameters in the πNN form factors

are fitted simultaneously with the pion structure function. The combined fits to both the HERA and E866 data suggest that models with a t -dependent exponential form factor give the best descriptions of the combined data sets over the largest range of kinematics, up to $y_{\text{cut}} \approx 0.3$, with reasonable values of the average pion multiplicity, $\langle n \rangle_{\pi N} = 0.34$. The fitted result for F_2^π is somewhat smaller than phenomenological parametrizations at $x \approx 0.1$, which may be due to the fit not being constrained at large x by the πN Drell-Yan data.

In the near future, the SeaQuest Drell-Yan experiment at Fermilab will measure the $\bar{d} - \bar{u}$ difference up to larger values of x , $x \approx 0.45$, which should allow improved constraints on the models of the pion distribution function in the nucleon. Beyond that, the tagged DIS experiment at Jefferson Lab [16] will provide precise information on pion exchange in leading proton production from an effective neutron target at kinematics complementary to the range covered by the HERA and Drell-Yan measurements.

4. Strange quarks

The role that strange quarks play in the nucleon has been the focus of attention in hadronic physics for nearly three decades. Early polarised DIS experiments suggested that a surprisingly large fraction of the proton's spin might be carried by strange quarks, in contrast to the naive quark model expectations. While the generation of $s\bar{s}$ pairs through perturbative gluon radiation typically produces symmetric s and \bar{s} distributions (at least up to two loop corrections), any significant difference between the momentum dependence of the s and \bar{s} PDFs would be a clear signal of nonperturbative effects.

Following the prediction by Thomas of a $\bar{d} - \bar{u}$ asymmetry, a similar mechanism was suggested [17] as a natural way for a nonzero $s - \bar{s}$ asymmetry to emerge from the breaking of chiral SU(3) symmetry in QCD. The nonperturbative dissociation of the proton to a kaon (containing an antistrange quark) and a hyperon (containing a strange quark) necessarily induces different momentum distributions for the s and \bar{s} quarks. Apart from its intrinsic interest, the possible $s - \bar{s}$ asymmetry is of great importance in connection with its contribution to the Paschos-Wolfenstein ratio and the NuTeV anomaly, which suggested a surprisingly large value for the weak mixing angle, $\sin^2 \theta_W$. A positive value of the integrated difference

$$S^- \equiv \langle x(s - \bar{s}) \rangle = \int_0^1 dx x [s(x) - \bar{s}(x)], \quad (4.1)$$

of the order $S^- \sim 2 \times 10^{-3}$, along with other corrections such as charge symmetry violation, was found to significantly reduce the excess and bring the NuTeV $\sin^2 \theta_W$ measurement closer to the Standard Model value [18].

Unfortunately, a reliable estimate of the strange asymmetry has been very difficult to obtain. Analysis of dimuon production in neutrino-nucleus reactions found comparable values, $S^- = (1.96 \pm 1.43) \times 10^{-3}$ at $Q^2 = 16 \text{ GeV}^2$ [19], although uncertainties remain in the treatment of bound state effects and the propagation of the D meson through the nucleus. More recently, Wang *et al.* [20] used chiral SU(3) effective theory together with phenomenological constraints to estimate the shape of the $s - \bar{s}$ PDF. In contrast to earlier chiral model studies, the analysis of Ref. [20] considered both the rainbow and tadpole diagram contributions, in addition to Kroll-Ruderman

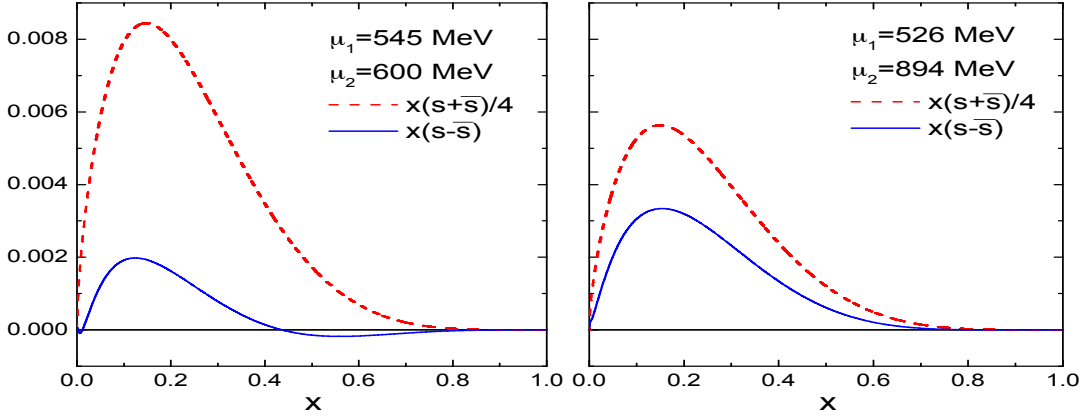


Figure 3: Total sum $x(s + \bar{s})$ (scaled down by a factor 1/4) and difference $x(s - \bar{s})$ of the strange and anti-strange PDFs from kaon loops at $Q^2 = 1 \text{ GeV}^2$ for two sets of fit parameters for the ultraviolet cutoffs, from Wang *et al.* [20].

terms needed to ensure gauge invariance. In particular, the δ -function contributions that arise from the Weinberg-Tomozawa contact interaction give rise to a valence-like component of the strange PDF, which cannot be generated from gluon radiation in perturbative QCD alone, and provide greater flexibility in the allowed phenomenological parametrization of $s - \bar{s}$.

Using experimental constraints from inclusive Λ production in pp scattering for the ultraviolet cutoff parameters, the s and \bar{s} quarks were found to contribute up to $\sim 1\%$ of the total momentum of the nucleon, or $\sim 30\% - 50\%$ of the phenomenological strange sea of the nucleon at a scale of $Q^2 = 1 \text{ GeV}^2$ [9]. In contrast, the magnitude of the strange asymmetry, $s - \bar{s}$, is about a factor of 10 smaller than the sum. Compared with other possible corrections to the NuTeV anomaly [18], this is a relatively minor effect, reducing the discrepancy by less than 0.5σ . The sign is, however, such as to reduce the anomaly, which in itself answers a long-standing uncertainty.

5. Charm in the nucleon

There has been a long-standing debate about the existence of intrinsic charm (IC) in the proton, associated with the 5-quark ($uudc\bar{c}$) component of the proton wave function. Aside from the intrinsic interest in the role of nonperturbative dynamics in the structure of the nucleon sea, the leptoproduction of charm quarks is also important in providing information on the gluon distribution in the nucleon. A decades-old measurement of the charm structure function F_2^c by the European Muon Collaboration (EMC) [21] provides tantalizing evidence for an enhancement at large x relative to what would be expected from perturbatively radiated $c\bar{c}$ pairs. More recent experiments, however, at small x from HERA found significant tension with the EMC data in regions of overlapping kinematics.

Several recent global QCD analyses have addressed the question of how much IC is allowed by the world's high-energy data set (see Ref. [22] and references therein). These have found the

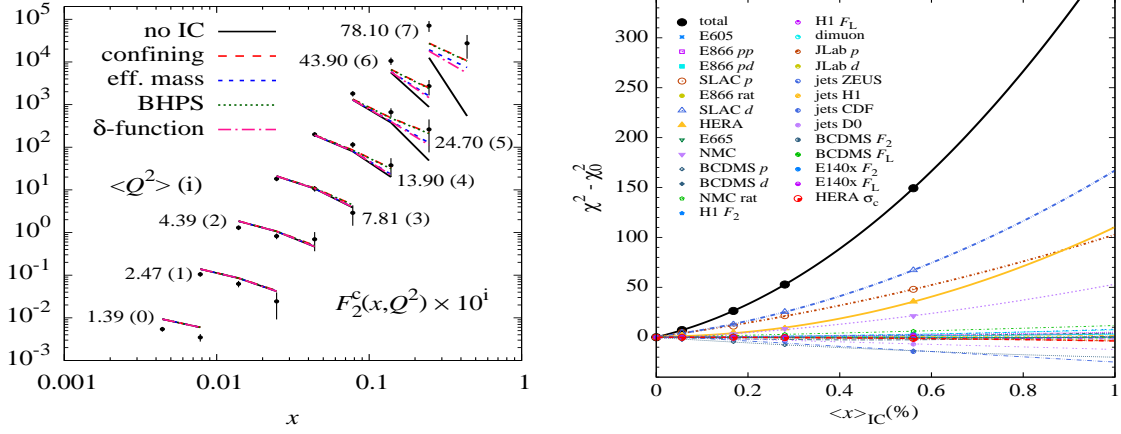


Figure 4: (Left) Comparison of the total fitted F_2^c structure function with the full set of EMC data [21], with the results for no IC compared with those using several nonperturbative models for IC [22]. (Right) Contributions to the total χ^2 relative to the value χ_0^2 for no IC, of various data sets as a function of the momentum fraction $\langle x \rangle_{\text{IC}}$ (in percent).

momentum fraction carried by IC quarks,

$$\langle x \rangle_{\text{IC}} \equiv \int_0^1 dx x [c(x) + \bar{c}(x)], \quad (5.1)$$

to be generally $\lesssim 1 - 2\%$. While most of these analyses do not include the old EMC F_2^c data, and typically make rather stringent kinematic cuts on the data fitted ($Q^2 \gtrsim 4 \text{ GeV}^2$ and $W^2 \gtrsim 12 \text{ GeV}^2$), a recent study [22] included all available data sets with $Q^2 \geq 1 \text{ GeV}^2$ and $W^2 \geq 3.5 \text{ GeV}^2$, and assessed the consistency of the EMC F_2^c data with the global PDF fit.

The comparison with the full set of F_2^c data from EMC is shown in Fig. 4 for several models of IC, as well as for a fit without IC. At small x values ($x \lesssim 0.02$) the global fits generally overestimate the data, regardless of whether IC (which is negligible in this region) is included or not. At intermediate x , where the IC contributions are still small, the agreement improves, while at the largest x values ($x \gtrsim 0.2$) the fit with no IC clearly lies below the data. The overall description of the EMC data is relatively poor, with a χ^2 value of 4.3 per datum for 19 data points. Better agreement with the EMC data would require significantly larger IC at high x , together with some additional suppression mechanism at low x values, neither of which appear very probable.

Excluding the EMC F_2^c measurements, the analysis [22] rules out large values of IC, with $\langle x \rangle_{\text{IC}}$ at most 0.5% at the 4σ level. The strongest constraint comes from the SLAC DIS data, as illustrated in Fig. 4, with smaller contributions from HERA charm production and NMC DIS cross sections. All other data sets have little or no sensitivity to IC, as evidenced by the rather shallow χ^2 profiles. The total χ^2 for the global fit gives $\chi^2/N_{\text{dat}} = 1.25$ for $N_{\text{dat}} = 4296$ data points.

Given that the signal for IC relies so heavily on charm production data at large values of x , it would be essential to obtain new, more precise data on F_2^c to determine limits (upper or lower) on the nonperturbative charm content of the nucleon with greater confidence. Such measurements could be feasible at a future Electron-Ion Collider.

6. Outlook

Improvements in the determination of PDFs are expected on several fronts over the next few years. A new generation of experiments at Jefferson Lab will map out the difficult to explore large- x region with dedicated measurements involving DIS from the ^3He and ^3H mirror nuclei and tagged DIS from the deuteron, which will provide significantly better constraints on the d -quark PDF, in particular, up to $x \approx 0.8$. Drell-Yan data from the SeaQuest experiment at Fermilab will fix \bar{d}/\bar{u} up to $x \approx 0.45$, answering the question of whether there is a sign change in the $\bar{d} - \bar{u}$ asymmetry at high x . At higher energies, data from the LHC on lepton-pair and W -boson production will continue to provide constraints on sea-quark PDFs, while a future Electron-Ion Collider is expected to open up new horizons for precision studies of both unpolarised and polarised PDFs.

References

- [1] R. Hofstadter and R. W. McAllister, Phys. Rev. **98**, 217 (1955).
- [2] R. Pohl *et al.*, Nature **466**, 213 (2010).
- [3] M. Breidenbach *et al.*, Phys. Rev. Lett. **23**, 935 (1969).
- [4] R. P. Feynman, Phys. Rev. Lett. **23**, 1415 (1969).
- [5] M. Gell-Mann, Phys. Lett. **8**, 214 (1964); G. Zweig, CERN preprint CERN-TH-401 (1964).
- [6] J. C. Collins, D. E. Soper and G. F. Sterman, Adv. Ser. Direct. High Energy Phys. **5**, 1 (1988).
- [7] P. Jimenez-Delgado, W. Melnitchouk and J. F. Owens, J. Phys. G **40**, 093102 (2013).
- [8] W. Melnitchouk and A. W. Thomas, Phys. Lett. B **377**, 11 (1996).
- [9] A. Accardi, L. T. Brady, W. Melnitchouk, J. F. Owens and N. Sato, Phys. Rev. D **93**, 114017 (2016).
- [10] L. T. Brady, A. Accardi, W. Melnitchouk and J. F. Owens, JHEP **1206**, 019 (2012).
- [11] A. W. Thomas, Phys. Lett. **126 B**, 97 (1983).
- [12] C.-R. Ji, W. Melnitchouk and A. W. Thomas, Phys. Rev. D **88**, 076005 (2013).
- [13] Y. Salamu, C.-R. Ji, W. Melnitchouk and P. Wang, Phys. Rev. Lett. **114**, 122001 (2015).
- [14] R. S. Towell *et al.*, Phys. Rev. D **64**, 052002 (2001).
- [15] J. McKenney, N. Sato, W. Melnitchouk and C.-R. Ji, Phys. Rev. D **93**, 054011 (2016).
- [16] C. E. Keppel, these proceedings.
- [17] A. I. Signal and A. W. Thomas, Phys. Lett. B **191**, 205 (1987).
- [18] W. Bentz, I. C. Cloët, J. T. Londergan and A. W. Thomas, Phys. Lett. B **693**, 462 (2010).
- [19] D. Mason *et al.*, Phys. Rev. Lett. **99**, 192001 (2007).
- [20] X. Wang, C.-R. Ji, W. Melnitchouk, Y. Salamu, A. W. Thomas and P. Wang, Phys. Lett. B **762**, 52 (2016); Phys. Rev. D **94**, 094035 (2016),
- [21] J. J. Aubert *et al.*, Nucl. Phys. **B213**, 31 (1983).
- [22] P. Jimenez-Delgado, T. J. Hobbs, J. T. Londergan and W. Melnitchouk, Phys. Rev. Lett. **114**, 082002 (2015); Phys. Rev. Lett. **116**, 019102 (2016).