Impact of the HERA I+II combined data on the CT14 QCD global analysis

S. Dulat[1,3], T.-J. Hou[2], J. Gao[3], M. Guzzi[4,*] J. Huston[5], P. Nadolsky[2], J. Pumplin[5], C. Schmidt[5], D. Stump[5], and C.-P. Yuan[5]

[1] *School of Physics Science and Technology, Xinjiang University, Urumqi, Xinjiang 830046, China*

[2] *Department of Physics, Southern Methodist University, Dallas, TX 75275-0181, U.S.A.*

[3] *High Energy Physics Division, Argonne National Laboratory, Argonne, Illinois 60439, U.S.A.*

[4] *The University of Manchester, Oxford Road, M13 9PL Manchester, United Kingdom*

[5] *Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, U.S.A.*

Abstract. A brief description of the impact of the recent HERA run I+II combination of inclusive deep inelastic scattering cross-section data on the CT14 global analysis of PDFs is given. The new CT14$_{\text{HERA2}}$ PDFs at NLO and NNLO are illustrated. They employ the same parametrization used in the CT14 analysis, but with an additional shape parameter for describing the strange quark PDF. The HERA I+II data are reasonably well described by both CT14 and CT14$_{\text{HERA2}}$ PDFs, and differences are smaller than the PDF uncertainties of the standard CT14 analysis. Both sets are acceptable when the error estimates are calculated in the CTEQ-TEA (CT) methodology and the standard CT14 PDFs are recommended to be continuously used for the analysis of LHC measurements.

1 Introduction

The CT14 parton distribution functions (PDFs) [1] are obtained in a global analysis of a variety of hadronic scattering experimental data and are suitable for general-purpose QCD calculations at the Large Hadron Collider (LHC) and in other experiments.

In global analyses to determine the PDFs of the proton, inclusive DIS measurements impose the most important constraints. In 2015, the H1 and ZEUS collaborations released a novel combination, "HERA2", of measurements of inclusive deep-inelastic scattering cross sections [2] at the HERA $e^{\pm}p$ collider. The two collaborations employed different experimental techniques and used different detectors and methods for kinematic reconstruction. As a result, the new HERA2 combined measurements exhibit reduced systematic uncertainties with respect to the previous combination of DIS data sets, "HERA1", published in 2009 [3].

In the study of Ref. [4], the impact of the HERA2 measurements on the CT14 global analysis is investigated by replacing the combined HERA1 data set used in the published CT14 PDFs [1], with the HERA2 set and by examining the resulting changes in PDF central values and uncertainties. Moreover, the dependence of the goodness-of-fit upon kinematic cuts on Q and x is also examined, as in Ref. [2] it was suggested that the low Q^2 HERA2 data are not well fitted by the CT10 and CT14

---

*Presenter, e-mail: marco.guzzi@manchester.ac.uk

PDFs. Related studies of the impact of HERA2 data in the context of MMHT14 and NNPDF3.0 analyses can be found in Refs. [5–7].

The CTEQ-TEA PDFs have been refitted at NLO and NNLO by using the global CT14 data ensemble, but with the HERA2 measurements in place of HERA1. The new PDFs obtained after the refitting procedure are named CT14$_{\text{HERA2}}$, to distinguish from CT14.

2 Features of the CT14$_{\text{HERA2}}$ analysis

The HERA2 data set has 1120 data points in the fitted region with $Q > 2$ GeV and $W^2 > 12.5$ GeV^2. There are 162 correlated systematic errors, and 7 procedural uncertainties, in addition to the luminosity uncertainty. When HERA2 is included in the global fit, there are in total 3287 data points in the CT14$_{\text{HERA2}}$ data ensembles, compared to 2947 in the original CT14 fits. This is because two other changes have been made in the data analysis: 1) the NMC muon-proton inclusive DIS data on F_2^p [8] have been dropped because this set is influenced by some unknown or underestimated systematic errors [1], in concordance with findings in our earlier studies [9]; 2) the data table for the CMS 7 TeV 5 fb^{-1} inclusive jet experiment [10], which became available after the completion of the CT14 study, has been updated and it produces no appreciable effects on the PDFs.

As in CT14, the theoretical predictions for the majority of processes in the CT14$_{\text{HERA2}}$ fit are calculated at the NNLO level of accuracy. In particular, a NNLO treatment [11] of heavy-quark mass effects in neutral-current DIS is realized in the S-ACOT-χ scheme [12–15] and is essential for obtaining correct predictions for LHC electroweak cross sections [16–19]. However, the calculations for charged-current DIS and inclusive jet production are included at NLO only; in both cases, the complete NNLO contributions are not yet available. In Sec. II of Ref. [1], various arguments are presented suggesting that the expected impact of the missing NNLO effects in jet production on the PDFs is small relatively to current experimental errors. Similarly, the NNLO contribution to charged-current DIS, including massive charm scattering contributions [20], is modest compared to the experimental uncertainties.

2.1 Parametrization of the PDFs

The CT14$_{\text{HERA2}}$ analysis adopts the CT14 PDF functional forms [1] at initial scale $Q_0 = 1.3$ GeV

$$x f_a(x, Q_0) = x^{a_1} (1 - x)^{a_2} P_a(x), \tag{1}$$

where the $P_a(x)$ functions are linear combinations of Bernstein polynomials. One change is made relative to the standard CT14. In Ref. [1] the strange quark PDF is parametrized according to Eq. (1), with $P_s(x)$ being a constant. The parameter a_1 is tied to the common a_1 of $\bar{u}$ and $\bar{d}$, and $s(x) = \bar{s}(x)$ is assumed in the analysis. Therefore, in the standard CT14 analysis, one has only two parameters for the strange quark PDF: a_2 and normalization. With this limitation on $s(x, Q_0)$, it is necessary to extend the strange quark uncertainty by adding two "extreme strange" PDFs to the set of Hessian error PDFs. In the CT14$_{\text{HERA2}}$ PDFs, a different technique is used to avoid underestimating the strangeness uncertainty provided by the Hessian error PDF set: while in the published CT14 PDFs the a_1 parameters are such that $a_1(s) = a_1(\bar{s}) = a_1(\bar{d}) = a_1(\bar{u})$, in CT14$_{\text{HERA2}}$ $a_1(s) = a_1(\bar{s})$ is allowed to differ from $a_1(\bar{d}) = a_1(\bar{u})$. By freeing the parameter $a_1(s)$, it is found that it is not necessary to construct additional extreme strange PDFs. So, whereas the CT14 error PDFs include two "extreme strange"

---

[1]In the CT14$_{\text{HERA2}}$ analysis of Ref. [4] it has been observed that even if they were included in the fit, they would induce a negligible change in the total χ^2

and two "extreme gluon" PDFs, the CT14$_{\text{HERA2}}$ error PDFs include only two "extreme gluon" PDFs to model the uncertainty of gluon PDFs in the very large x and small x regions, respectively. Thus the total number of error PDFs is the same for CT14 and CT14$_{\text{HERA2}}$.

3 Impact of data selection cuts

The HERA2 publication [2] found that both HERAPDF2.0 PDFs and χ^2 values depend significantly on the choice of Q_{cut}, the minimum value of the four-momentum-transfer Q in the HERA2 analysis. In order to explore the impact of variations of Q_{cut} on the CT14$_{\text{HERA2}}$ global analysis, multiple fits of CT14$_{\text{HERA2}}$ PDFs have been performed in which Q_{cut} is varied from 2 GeV to 6 GeV and the results are compared to those of the CT14 analysis. For every choice of Q_{cut}, the total χ^2, reduced-χ^2 (*i.e.*, χ^2_{re}), and the systematic shift penalty R^2 (defined respectively in Eq. 2 and 3 of Ref. [4]), are illustrated in Fig. 1. The χ^2/N_{pts} accounts for the sum of all four DIS subprocesses and the graphs show the
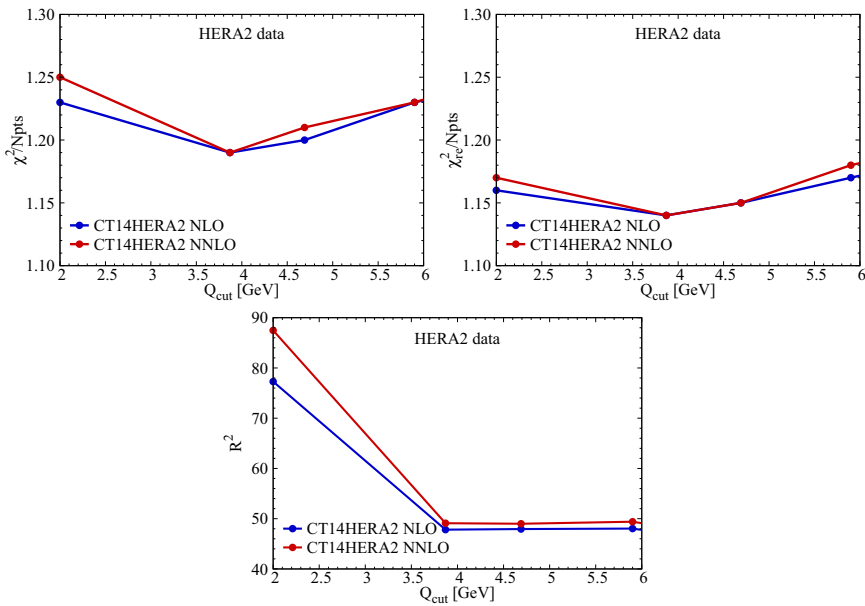
Figure 1. χ^2/N_{pts} (left), reduced-χ^2/N_{pts} (right), and R^2 (bottom) for the HERA2 data and CT14$_{\text{HERA2}}$ PDFs, as a function of Q_{cut}.

dependence on Q_{cut} in the CT14$_{\text{HERA2}}$ analysis at both NLO and NNLO. The values of χ^2/N_{pts} for the HERA2 data exhibit a *shallow minimum* for Q_{cut} in the range $3.5 \lesssim Q_{\text{cut}} \lesssim 4$ GeV. The reduction of χ^2 at $Q_{\text{cut}} \sim 4$ GeV, compared to the CT standard choice of $Q_{\text{cut}} = 2$ GeV, from 1.17 to 1.15, does not seem significant. An interesting feature of the graphs is that near the minimum the NNLO and NLO results are equal, whereas NNLO has larger χ^2 on either side of the minimum. The lower panel in Fig. 1 shows R^2, the total quadratic penalty for the systematic errors, as a function of Q_{cut}. The value of R^2 decreases significantly from $Q_{\text{cut}} = 2$ to 3.87 GeV, from 87 to 49 (for 170 systematic errors). For ideal Gaussian systematic errors one would expect $R^2 \sim 170$. When the low-Q data points are discarded by the cut, the systematic errors become less important. However, this reduction of R^2 is shared by 1120 total data points, so the overall net change in χ^2/N_{pts} is mild.

4 Results for the CT14$_{\text{HERA2}}$ PDFs

The central values and uncertainties of the CT14$_{\text{HERA2}}$ PDFs are compared to those of the CT14 global analysis in Fig. 2, where only the NNLO PDFs at the scale Q equal to the initial scale $Q_0 = 1.3$ GeV are shown. At this low scale, the PDF uncertainties are magnified, and they are reduced at electroweak scales as a consequence of DGLAP evolution. Additional plots can be found on the CTEQ public website [21]. The comparison is summarized below where a few comments are in order.

The central value of the CT14$_{\text{HERA2}}$ gluon in the range $10^{-2} \lesssim x \lesssim 0.2$ is almost unchanged compared to CT14; it is larger by about 30% at $x \approx 10^{-4}$, by a larger factor for $x > 0.5$, and it is smaller by about 10% at $x \approx 0.3$.

The u and d quarks are generally slightly larger than (but close to) CT14 in the range $10^{-2} \lesssim x \lesssim 0.5$ where the CT14$_{\text{HERA2}}$ uncertainty band is comparable to that of CT14; whereas they are both systematically larger by about 5% in the intermediate region of $10^{-4} \lesssim x \lesssim 10^{-2}$. The CT14$_{\text{HERA2}}$/CT14 ratio decreases at $x \lesssim 10^{-4}$ in both cases. The d quark increases at $x > 0.5$, while the u quark decreases slightly at $x \approx 0.5$. The slow oscillations in $d(x, Q_0)$ reflect the behaviour of Bernstein polynomials in Eq. (1).

The strange quark central prediction is reduced over the entire x range, mainly due to the change of adding one additional shape parameter for describing the strange quark PDF; but this reduction is statistically insignificant and completely within the uncertainty of the previous PDF ensemble. In particular a reduction of approximately -50% is observed at both $x \lesssim 10^{-3}$ and $x \gtrsim 0.5$.

The $\bar{u}$ and $\bar{d}$ quarks share similar features. They are almost unchanged for $10^{-2} \lesssim x \lesssim 0.2$ where the uncertainties are comparable to those of CT14. The $\bar{u}$ quark PDF increases by about 10% at x around 0.2, and the $\bar{d}$ quark PDF similarly at x around 0.3. Both of $\bar{u}$ and $\bar{d}$ quarks, similar to s quark, decrease by large factors for $x \gtrsim 0.4$, where both gluon and down quark PDFs increase, as a consequence of momentum sum rule. It is important to keep in mind that at $x > 0.5$ the antiquark PDFs take very small values, their behaviour is very uncertain and strongly depends on the parametrization form.

The individual PDF uncertainties do not change much, except in the unconstrained x regions of the strange quark PDF due to the change made in its parametrization form.

It has been verified that the change seen in gluon, up and down quark PDFs mainly arises from replacing the HERA1 data (in CT14 analysis) by the HERA2 data (in CT14$_{\text{HERA2}}$ analysis). This was explicitly checked by comparing CT14 PDFs to the the result of a new fit which uses exactly the same setup as that in the CT14 global analysis, but with the HERA1 data replaced by the HERA2 data.

Conclusions

A short overview of the CT14 global analysis of proton's PDFs has been given. With rapid advancements in LHC measurements, the focus of the global analysis has shifted toward providing accurate predictions in the wide range of x and Q covered by the LHC data. This development requires long-term multi-prong effort in theoretical, experimental, and statistical areas. On the theory side, a more flexible parametrization has been introduced to better capture variations in the PDF dependence. A series of benchmark tests of NNLO cross sections, carried out in the run-up for the CT14 fit for all key fitted processes, have resulted in better agreement with most experiments and brought accuracy of most predictions to the truly NNLO level. The final CT14 PDFs are presented in the form of 1 central and 56 Hessian eigenvector sets at NLO and NNLO. The 90% C.L. PDF uncertainties for physical observables can be estimated from these sets using the symmetric [9] or asymmetric [22, 23] master formulae. These PDFs are determined for the central QCD coupling of $\alpha_s(M_Z) = 0.118$, consistent with the world-average α_s value.
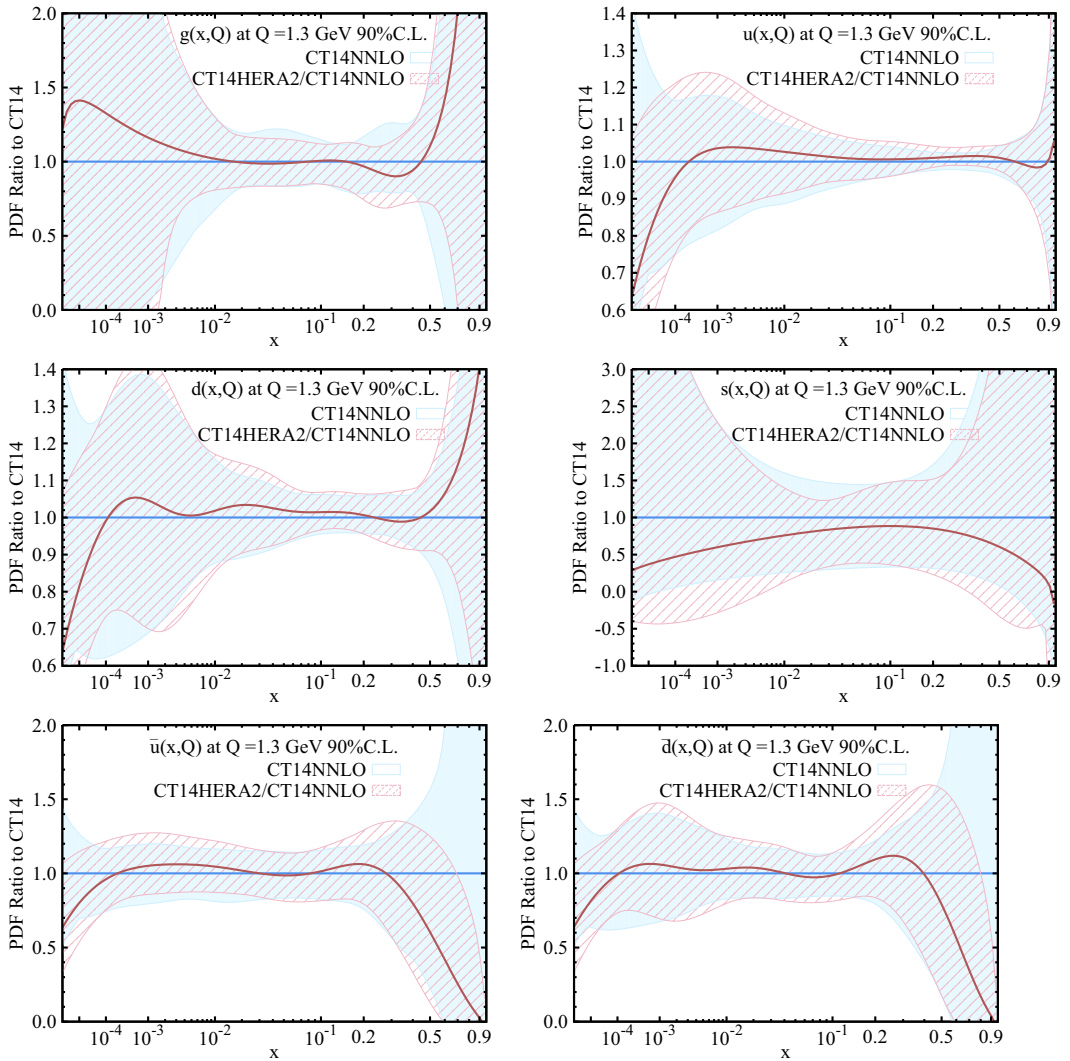
Figure 2. Comparison of CT14$_{HERA2}$ (red) and CT14 (blue) PDFs at Q = 1.3 GeV. Flavors $g, u, d, s, \bar{u}, \bar{d}$ are shown. The curves compare the central fits, plotted as ratios to CT14. The uncertainty bands are 90% C.L. uncertainties evaluated from the CT14 (shaded blue) and CT14$_{HERA2}$ (hatched red) error ensembles; both error bands are normalized to the corresponding central CT14 PDFs. All PDFs are from the NNLO QCD analysis.

Acknowledgments

References

[1] S. Dulat, T.J. Hou, J. Gao, M. Guzzi, J. Huston, P. Nadolsky, J. Pumplin, C. Schmidt, D. Stump, C.P. Yuan, Phys. Rev. **D93**, 033006 (2016), `1506.07443`.

[2] H. Abramowicz et al. (ZEUS, H1), Eur. Phys. J. **C75**, 580 (2015), `1506.06042`.

[3] F.D. Aaron et al. (ZEUS, H1), JHEP **01**, 109 (2010), `0911.0884`.

[4] T.J. Hou, S. Dulat, J. Gao, M. Guzzi, J. Huston, P. Nadolsky, J. Pumplin, C. Schmidt, D. Stump, C.P. Yuan (2016), `1609.07968`.

[5] L.A. Harland-Lang, A.D. Martin, P. Motylinski, R.S. Thorne, Eur. Phys. J. **C76**, 186 (2016), `1601.03413`.

[6] R.S. Thorne, L.A. Harland-Lang, A.D. Martin, P. Motylinski, PoS **EPS-HEP2015**, 505 (2015), `1508.06621`.

[7] J. Rojo (NNPDF), PoS **EPS-HEP2015**, 506 (2015), `1508.07731`.

[8] M. Arneodo et al. (New Muon), Nucl. Phys. **B483**, 3 (1997), `hep-ph/9610231`.

[9] J. Pumplin, D.R. Stump, J. Huston, H.L. Lai, P.M. Nadolsky, W.K. Tung, JHEP **07**, 012 (2002), `hep-ph/0201195`.

[10] S. Chatrchyan et al. (CMS), Phys. Rev. **D87**, 112002 (2013), [Erratum: Phys. Rev. D87 (2013) 119902], `1212.6660`.

[11] M. Guzzi, P.M. Nadolsky, H.L. Lai, C.P. Yuan, Phys. Rev. **D86**, 053005 (2012), `1108.5112`.

[12] M.A.G. Aivazis, F.I. Olness, W.K. Tung, Phys. Rev. **D50**, 3085 (1994), `hep-ph/9312318`.

[13] M.A.G. Aivazis, J.C. Collins, F.I. Olness, W.K. Tung, Phys. Rev. **D50**, 3102 (1994), `hep-ph/9312319`.

[14] J.C. Collins, Phys. Rev. **D58**, 094002 (1998), `hep-ph/9806259`.

[15] W.K. Tung, S. Kretzer, C. Schmidt, J. Phys. **G28**, 983 (2002), `hep-ph/0110247`.

[16] J. Gao, M. Guzzi, P.M. Nadolsky, Eur. Phys. J. **C73**, 2541 (2013), `1304.3494`.

[17] H.L. Lai, J. Huston, Z. Li, P. Nadolsky, J. Pumplin, D. Stump, C.P. Yuan, Phys. Rev. **D82**, 054021 (2010), `1004.4624`.

[18] P.M. Nadolsky, H.L. Lai, Q.H. Cao, J. Huston, J. Pumplin, D. Stump, W.K. Tung, C.P. Yuan, Phys. Rev. **D78**, 013004 (2008), `0802.0007`.

[19] W.K. Tung, H.L. Lai, A. Belyaev, J. Pumplin, D. Stump, C.P. Yuan, JHEP **02**, 053 (2007), `hep-ph/0611254`.

[20] E.L. Berger, J. Gao, C.S. Li, Z.L. Liu, H.X. Zhu, Phys. Rev. Lett. **116**, 212002 (2016), `1601.05430`.

[21] http://hep.pa.msu.edu/cteq/public/index.html, the CTEQ-TEA webpage.

[22] H.L. Lai, M. Guzzi, J. Huston, Z. Li, P.M. Nadolsky, J. Pumplin, C.P Yuan, Phys. Rev. **D82**, 074024 (2010), `1007.2241`.

[23] P.M. Nadolsky, Z. Sullivan, eConf **C010630**, P510 (2001), `hep-ph/0110378`.