New limits on intrinsic charm in the nucleon from global analysis of parton distributions

P. Jimenez-Delgado\textsuperscript{1}, T. J. Hobbs\textsuperscript{2}, J. T. Londergan\textsuperscript{2}, W. Melnitchouk\textsuperscript{1}

\textsuperscript{1}Jefferson Lab, 12000 Jefferson Avenue, Newport News, Virginia 23606, USA
\textsuperscript{2}Department of Physics and Center for Exploration of Energy and Matter, Indiana University, Bloomington, Indiana 47405, USA

(Dated: August 11, 2014)

We present a new global QCD analysis of parton distribution functions, allowing for possible nonperturbative or intrinsic charm (IC) contributions in the nucleon inspired by light-front models. The analysis makes use of the full range of available high energy scattering data for $Q^2 \gtrsim 1$ GeV\(^2\) and $W^2 \gtrsim 3.5$ GeV\(^2\), including SLAC proton and deuteron deep-inelastic scattering cross sections that were excluded in previously global analyses. The SLAC data in particular place more stringent constraints on the momentum carried by IC, with $\langle x \rangle_{IC} < 0.1\%$ at the 5\% level. We also critically assess the impact of older EMC measurements of $F_2^c$ at large $x$, which favor a nonzero IC, $\langle x \rangle_{IC} = 0.13 \pm 0.04\%$, but have a very large $\chi^2$ value.

There has been considerable interest recently in the nature of Fock states of the proton wave function involving five or more quarks, such as $|uudq\bar{q}\rangle$, where $q = u, d, s$ or $c$ [1–7]. This has arisen partly from attempts to understand flavor asymmetries observed in the nucleon sea, such as $d > u$ [8, 9] and $s \neq \bar{s}$ [10], which clearly point to a nonperturbative origin. In addition, there has been a long-standing debate about the existence of intrinsic charm (IC) quarks in the proton, associated with the nonperturbative $|uudc\bar{c}\rangle$ 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 lepton production of charm quarks is also rather important in providing information on the gluon distribution of the nucleon. In addition, a significant IC component in the nucleon wave function could influence observables measured at the LHC, either directly through enhanced cross sections at large $x$, or indirectly via the momentum sum rule leading to a decreased momentum fraction carried by gluons.

Following early indications from measurements of charm production in $pp$ scattering of an anomalous excess of $D$ mesons at large values of Feynman $x_F$ (see [11] and references therein), the proposal was made that the observed enhancement could be accounted for with the addition of intrinsic $c\bar{c}$ pairs in the nucleon that were not generated through perturbative gluon radiation [12]. Neglecting quark transverse momentum and assuming a charm mass much greater than other mass scales, Brodsky, Hoyer, Peterson and Sakai (BHPS) [12] derived an analytic approximation to the IC distribution that, unlike the perturbatively generated charm, was peaked at relatively large parton momentum fractions $x$.

A number of experimental and theoretical studies have since sought to elucidate this issue, although the evidence to date has been somewhat inconclusive. Measurements of the charm structure function $F_2^c$ by the European Muon Collaboration (EMC) [13] at CERN provided tantalizing evidence for an enhancement at large $x$; however, more recent experiments at HERA [14] focusing at small $x$ found significant tension with the EMC data in regions of overlapping kinematics.

Early theoretical analyses of the EMC charm lepton production data indicated an IC component with normalization $N_{IC} = \int_0^1 dx c(x) \sim 1\%$, although later, more sophisticated treatments incorporating the photon–gluon fusion (PGF) process, as well as quark and target mass corrections, argued for smaller IC level, $\sim 0.3\%$ [15]. A subsequent study by Harris, Smith and Vogt [16] which included $O(\alpha_s)$ corrections to the hard scattering cross section obtained a best fit to the highest-energy EMC data with $N_{IC} = (0.86 \pm 0.60)\%$. A follow-up analysis by Steffens et al. [17] employed a hybrid scheme to interpolate between massless evolution at large $Q^2$ and PGF at low $Q^2$, including quark and hadron mass and threshold effects. Using the BHPS model, as well as a meson-baryon model based on fluctuations of the nucleon to charmed baryon and virtual $D$ meson states [18–20], Steffens et al. found it difficult to fit the entire data set simultaneously in terms of a single IC framework, although a slight preference was found for IC in the meson-baryon model at a level of $N_{IC} \approx 0.4\%$.

To place the study of IC on a more robust statistical footing, Pumplin et al. [21] used the framework of the CTEQ global QCD fit [22] to parton distribution functions (PDFs) to determine the level of IC that could be accommodated by the high-energy data. Utilizing the BHPS model, a $p \rightarrow \Lambda^+D^0$ fluctuation model with scalar couplings, and a sea-like ansatz in which the charm distribution is proportional to the $\bar{u}$ and $\bar{d}$ PDFs, the analysis found the allowed range of IC to be from zero to a level 2–3 times larger than earlier estimates [21].

An updated NNLO fit by Dulat et al. [23], based on the more recent CT10 global analysis [24] and the BHPS and sea-like IC models, found the momentum fraction...
carried by charm quarks,

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

to be \(\lesssim 2.5\%\) for the BHPS distribution, and \(\lesssim 1.5\%\) for the sea-like shape at the 90\% confidence level. Note that for the BHPS distribution with a 1\% normalization, the corresponding momentum fraction is \(\langle x \rangle_{IC} = 0.57\%\). The Dulat et al. analysis therefore suggests that the existing experimental data may tolerate rather significant momentum carried by IC.

In this letter, we revisit the question of the magnitude of IC that is allowed by the world’s \(F_2\) and other high-energy scattering data, by performing a new global QCD analysis, along the lines of the recent JR14 fit [25]. Unlike the previous global analyses [21, 23] which placed more stringent cuts on the data \((Q^2 \gtrsim 4 \text{ GeV}^2\) and \(W^2 \gtrsim 12 \text{ GeV}^2\), excluding for instance all fixed target cross section measurements from SLAC [26], we include all available data sets with \(Q^2 \gtrsim 1 \text{ GeV}^2\) and \(W^2 \gtrsim 3.5 \text{ GeV}^2\). Since most IC models predict this effect to be most prominent at large values of \(x\), excluding the largest-\(x\) data may seriously reduce the sensitivity of the global fit to any IC that may be present. Indeed, we find a dramatic effect of the SLAC proton and deuteron structure function data on the allowed range of IC, with significantly tighter bounds than in previous global analyses. In addition, we assess the consistency of the earlier EMC \(F_2\) data [13], which have often been cited as providing the strongest evidence for IC in high-energy processes.

Of course, inclusion of lower-\(Q^2\) data requires careful treatment of possible finite-\(Q^2\) and nuclear corrections at intermediate and large \(x\). Following Refs. [25, 27–29], we account for target mass corrections explicitly, using the moment space results for \(F_2\) and \(F_L\) from Ref. [30], and allow for \(1/Q^2\) higher twist contributions which are determined phenomenologically. For nuclear smearing and nucleon off-shell corrections in the deuteron, which are known to be most important at large values of \(x\), we adopt the method used in the CJ global analysis [28, 29], while for data on heavier nuclei the nuclear PDFs from de Florian and Sassot [31] are employed. Also, whenever possible, we fit the original cross section data rather than structure functions derived from the cross sections using an assumed ratio of longitudinal to transverse cross sections. This reduces the systematic uncertainties in the analysis by avoiding model-dependent separations of the cross sections into \(F_2\) and \(F_L\) structure functions with data for which actual longitudinal–transverse separations were not available. (Further details about the QCD analysis can be found in Ref. [25].)

For the QCD analysis we use the framework of the JR14 global analysis [25], in which the total \(F_2\) structure function is given by

$$F_2 = F_2^{\text{light}} + F_2^{\text{heavy}},$$

where \(F_2^{\text{light}}\) denotes the light-quark \((u, d, s)\) contributions, and \(F_2^{\text{heavy}}\) includes contributions from the heavy \(c\) and \(b\) quarks. The charm structure function is further decomposed into a perturbative part, \(F_2^{c}\), and a nonperturbative (IC) component, \(F_2^{IC}\),

$$F_2 = F_2^{c} + F_2^{IC}.$$
to simulate the effects of confinement [11, 33]. In the present analysis we fix the shapes of the various IC distributions computed in Ref. [11] by the respective best fit cut-off parameters, but allow the overall normalization to vary. Provided the variation of the cut-off parameters is not dramatic, the effect on the shape of the IC distribution is minor. Note that, unlike the BHPS model, the meson-baryon fluctuation model naturally accommodates asymmetric $c$ and $\bar{c}$ distributions as a function of $x$ [18]. Finally, from the IC distribution in a given model, the IC contribution to the charm structure function is computed using the framework introduced by Hoffman and Moore [15] at $\mathcal{O}(\alpha_s)$, where

$$F_2^{IC}(x, Q^2) = F_2^{IC(0)}(x, Q^2) + \alpha_s F_2^{IC(1)}(x, Q^2), \quad (5)$$

with the coefficients $F_2^{IC(0,1)}$ given in Ref. [15].

The results of our global analysis are summarized in Fig. 1, where the total $\chi^2$ values for each of the 26 data sets used in the fit are shown (relative to the value for no IC, $\chi^2_0$) as a function of the momentum fraction $\langle x \rangle_{IC}$ carried by IC quarks. The total $\chi^2$ has its minimum for zero IC, and rises rapidly with increasing $\langle x \rangle_{IC}$. The largest contributions to the $\chi^2$ arise from the SLAC deep-inelastic proton and deuteron structure function measurements [26], with smaller contributions from HERA charm production cross sections at low $x$ [14], and NMC proton and deuteron cross sections in the medium-$x$ region [34]. All other data sets have little or no sensitivity to IC, as evidenced by the rather shallow $\chi^2$ profiles. The total $\chi^2$ for the global fit gives $\chi^2/N_{\text{dat}} = 1.25$ for $N_{\text{dat}} = 4296$ data points.

Because of the more restrictive $Q^2$ and $W^2$ cuts employed in previous global IC studies [21, 23], which were based on analyses tuned particularly to collider data (for which the most relevant kinematics are small $x$ and large $Q^2$), the lower-$Q^2$ fixed-target data from SLAC were excluded from the fits. This produced rather weak limits on the IC momentum fraction $\langle x \rangle_{IC} \lesssim 2 - 3\%$. Including the SLAC data in our analysis, we find that the total $\chi^2$ leads to a much more stringent constraint on the momentum carried by IC, with $\langle x \rangle_{IC} < 0.1\%$ at the $5\sigma$ level. The rest of the $\chi^2$ profile allows slightly larger IC values, as illustrated in Fig. 2(a), with $\langle x \rangle_{IC} < 0.1\%$ at the $1\sigma$ level. Higher momentum fraction limits of $0.2\%$ are excluded at the $2\sigma$ level, and $0.3\%$ excluded at the $3\sigma$ level.

The difference in the $\langle x \rangle_{IC}$ values between our analysis
without the SLAC data and those in Refs. [21, 23] are partly explained by the different tolerance criteria used: in our fits the PDF errors refer to variations of $\Delta \chi^2 = 1$ around the minimum, whereas the previous analyses [21, 23] assumed a tolerance of $\Delta \chi^2 = 100$. There is no unique criterion for selecting the correct $\Delta \chi^2$ interval, and other groups make different choices [35]. Following Refs. [25, 36] we use the traditional $\Delta \chi^2 = 1$ interval, based on statistical considerations alone. If one chose to define the 90% confidence level at $\Delta \chi^2 = 100$, then from our global analysis one would conclude that the limit on intrinsic charm was $(x)_{IC} \leq 0.4\%$.

Note that while the global fits in Fig. 1 incorporate the charm production cross sections from HERA [14], they do not include the earlier charm structure function data from EMC [13]. Since the HERA cross sections are predominantly measured at small $x$, they have less sensitivity to the presence of IC than the SLAC data at larger $x$, as the $\chi^2$ profile in Fig. 1 illustrates. However, the EMC $F_2^c$ measurements include data points also at large $x$ values, and do have greater impact on the IC determination. In Fig. 2(b) the $\chi^2$ values for the global fits including the EMC data, $\chi^2/N_{\text{dat}} = 1.27$ for $N_{\text{dat}} = 4315$ data points, indicate a slight preference for a nonzero IC, with $(x)_{IC} = 0.13 \pm 0.04\%$. The EMC data alone favor a value $\sim 0.3-0.4\%$, while the SLAC data in particular pull this down considerably, with other data sets mostly insensitive to IC. Note, however, that the description of the EMC data is far from satisfactory, giving a $\chi^2$ value of 4.3 per datum for 19 data points.

The comparison with the full set of $F_2^c$ data from EMC is shown in Fig. 3, for several models of IC from Refs. [11, 12], 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 $x$ ($0.02 \lesssim x \lesssim 0.1$), where the IC contributions are still small, the agreement improves, while at the largest $x$ values ($x \gtrsim 0.2$) the fits with no IC clearly lie below the data points. Here the addition of IC improves the agreement for all models considered, with the meson-baryon model for the confining $c$ quark-diquark interaction [11] and the BHPS model [12] resulting in the biggest enhancement. On the other hand, the experimental uncertainties at the high $x$ values are rather large compared with those in the small-$x$ region, where the fit to the EMC $F_2^c$ data is worse. This then results in a poor overall $\chi^2$ per degree of freedom for the total EMC data set. Because of the significant tension with the other global data sets, the EMC data are usually not included in the standard global PDF analyses [22–25, 27–29, 35].

These conclusions are more consistent with those reached in the MSTW global PDF analysis [35], which found reasonable global fits including the EMC data for a value $N_{IC} = 0.3\%$ using the BHPS model for IC. On the other hand, the analysis [35] also utilized more stringent cuts ($Q^2 \geq 2$ GeV$^2$ and $W^2 \geq 15$ GeV$^2$) than those used in our fit, which removed much of the SLAC data [26] at large $x$, and did not consider higher twist corrections to the structure functions — both of which we find to be important in the region where IC is expected to contribute.

In summary, we have performed a comprehensive global QCD analysis of the world’s high energy scattering data, synthesizing the latest developments in global fitting technology and nonperturbative studies of charm production to fully exploit all of the available data that may have bearing on the question of IC in the nucleon. By relaxing the cuts on $Q^2$ and $W^2$ used in earlier global fits [21, 23, 35], while systematically accounting for finite-$Q^2$ and other hadronic and nuclear corrections [25, 27, 29], we found that the low-$Q^2$, high-$x$ data from SLAC [26] in particular place stronger constraints on the magnitude of IC than found previously. Excluding the older $F_2^c$ measurements from the EMC [13], the best fit to the data is obtained for zero IC, with the momentum carried by IC $(x)_{IC} < 0.1\%$ at the $5\sigma$ level. Including the EMC data, our fits favor a small but nonzero value, $(x)_{IC} = 0.13 \pm 0.04\%$, but with a very large $\chi^2$ for those data. The tension between the EMC data and the more precise measurements of the charm structure function at HERA at low $x$ [14] has prompted many global PDFs analyses to omit these data from their fits. 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 facility [37].

T.J.H. and J.T.L. were supported in part by the U.S. National Science Foundation under Grant No. NSF-PHY-1205019. The work of T.J.H. was also supported in part by DOE Grant No. DE-FG02-87ER40365. W.M. was supported by the DOE Contract No. DE-AC05-06OR23177, under which Jefferson Science Associates, LLC operates Jefferson Lab.

[1] J. Speth and A. W. Thomas, Adv. Nucl. Phys. 24, 83 (1998).
[2] S. Kumano, Phys. Rep. 303, 183 (1998).
[3] R. Vogt, Prog. Part. Nucl. Phys. 45, S105 (2000).
[4] G. T. Garvey and J.-C. Peng, Prog. Part. Nucl. Phys. 47, 203 (2001).
[5] W.-C. Chang and J.-C. Peng, Phys. Rev. Lett. 106, 252002 (2011).
[6] W.-C. Chang and J.-C. Peng, arXiv:1406.1260 [hep-ph].
[7] J.-C. Peng and J.-W. Qiu, Prog. Part. Nucl. Phys. 76, 43 (2014).
[8] P. Amaudruz et al., Phys. Rev. Lett. 66, 2712 (1991).
[9] E. A. Hawker et al., Phys. Rev. Lett. 80, 3715 (1998).
[10] D. Mason et al., Phys. Rev. Lett. 99, 192001 (2007).
[11] T. J. Hobbs, J. T. Londergan and W. Melnitchouk, Phys. Rev. D 89, 074008 (2014).
[12] S. J. Brodsky, P. Hoyer, C. Peterson and N. Sakai, Phys. Lett. B 93, 451 (1980).
[13] J. J. Aubert et al., Nucl. Phys. B213, 31 (1983); Phys. Lett. B 94, 96 (1980); ibid. B 110, 73 (1982).
[14] H. Abramowicz et al., Eur. Phys. J. C 73, 2311 (2013).
[15] E. Hoffmann and R. Moore, Z. Phys. C 20, 71 (1983).
[16] B. W. Harris, J. Smith and R. Vogt, Nucl. Phys. B461, 181 (1996).
[17] F. M. Steffens, W. Melnitchouk and A. W. Thomas, Eur. Phys. J. C 11, 673 (1999).
[18] W. Melnitchouk and A. W. Thomas, Phys. Lett. B 414, 134 (1997).
[19] F. S. Navarra, M. Nielsen, C. A. A. Nunes and M. Teixeira, Phys. Rev. D 54, 842 (1996); F. S. Navarra and M. Nielsen, Phys. Lett. B 443, 285 (1998); F. O. Düräes, F. S. Navarra and M. Nielsen, Phys. Lett. B 498, 169 (2001).
[20] J. Pumplin, Phys. Rev. D 73, 114015 (2006).
[21] J. Pumplin, H. L. Lai and W.-K. Tung, Phys. Rev. D 75, 054029 (2007).
[22] W.-K. Tung et al., JHEP 0702, 053 (2007).
[23] S. Dulat et al., Phys. Rev. D 89, 073004 (2014).
[24] J. Gao et al., Phys. Rev. D 89, 073004 (2014).
[25] P. Jimenez-Delgado and E. Reya, Phys. Rev. D 89, 074049 (2014).
[26] L. W. Whitlow, E. M. Riordan, S. Dasu, S. Rock and A. Bodek, Phys. Lett. B 282, 475 (1992).
[27] S. Alekhin, J. Blümlein, S. Klein and S.-O. Moch, Phys. Rev. D 81, 014032 (2010).
[28] A. Accardi et al., Phys. Rev. D 84, 014008 (2011).
[29] J. F. Owens, A. Accardi and W. Melnitchouk, Phys. Rev. D 87, 094012 (2013).
[30] F. M. Steffens, M. D. Brown, W. Melnitchouk and S. Sanches, Phys. Rev. C 86, 065208 (2012).
[31] D. de Florian and R. Sassot, Phys. Rev. D 69, 074028 (2004).
[32] P. Jimenez-Delgado, arXiv:0902.3947 [hep-ph].
[33] W. Melnitchouk, A. W. Schreiber and A. W. Thomas, Phys. Rev. D 49, 1183 (1994).
[34] M. Arneodo et al., Nucl. Phys. B487, 3 (1997).
[35] A. D. Martin, W. J. Stirling, R. S. Thorne and G. Watt, Eur. Phys. J. C 63, 189 (2009).
[36] P. Jimenez-Delgado, A. Accardi and W. Melnitchouk, Phys. Rev. D 89, 034025 (2014).
[37] A. Accardi et al., arXiv:1212.1701 [nucl-ex].