First Monte Carlo global QCD analysis of pion parton distributions

P. C. Barry,1 N. Sato,2 W. Melnitchouk,3 and Chueng-Ryong Ji1

1North Carolina State University, Raleigh, North Carolina 27607, USA
2University of Connecticut, Storrs, Connecticut 06269, USA
3Jefferson Lab, Newport News, Virginia 23606, USA

Jefferson Lab Angular Momentum (JAM) Collaboration

(Dated: April 6, 2018)

We perform the first global QCD analysis of parton distribution functions (PDFs) in the pion, combining $\pi A$ Drell-Yan data with leading neutron electroproduction from HERA within a Monte Carlo approach based on nested sampling. Inclusion of the HERA data allows the pion PDFs to be determined down to much lower values of $x$, with relatively weak model dependence from uncertainties in the chiral splitting function. The combined analysis reveals that gluons carry a significantly higher pion momentum fraction, $\sim 30\%$, than that inferred from Drell-Yan data alone, with sea quarks carrying a somewhat smaller fraction, $\sim 15\%$, at the input scale. Within the same effective theory framework, the chiral splitting function and pion PDFs can be used to describe the $d-u$ asymmetry in the proton.

As the lightest QCD bound state, the pion has historically played a central role in the study of the strong nuclear interactions. On one hand, it has been the critical ingredient for understanding the consequences of dynamical chiral symmetry breaking in QCD, and how this dictates the nature of hadronic interactions at low energies. On the other hand, its quark and gluon (or parton) substructure has been revealed through high energy scattering experiments, such as Drell-Yan (DY) lepton-pair creation in inclusive pion–nucleon scattering [1]. In some cases, both aspects are on display, as in the role of the pion cloud of the proton in generating a flavor asymmetry in its light antiquark sea, $\bar{d} \neq \bar{u}$ [2].

As the simplest $q\bar{q}$ state, the structure of the pion is relatively more straightforward to compute theoretically than baryons, but the absence of fixed pion targets has made it difficult to determine the pion’s structure experimentally. Most information on the partonic structure of pions has come from pion–nucleus scattering with prompt photon or dilepton production at CERN [3, 4] and Fermilab [5]. These data have been used in several global QCD analyses [6–9] to fit the momentum dependence of the pion’s parton distribution functions (PDFs) for parton fractions $x_\pi \gtrsim 0.1$ of the pion’s light-cone momentum.

While the DY data constrain reasonably well the pion’s valence PDFs, the sea quark and gluon PDFs at small $x_\pi$ values have remained essentially unknown. More recently, leading neutron (LN) electroproduction from HERA [10, 11], which at forward angles is expected to be dominated by pion exchange, has been used to study the pion structure function down to very low values of $x_\pi \sim 10^{-3}$. The interpretability of the LN data in terms of pion PDFs is limited, however, by the model dependence inherent in this process, in which the cross section is given as a product of a proton to neutron + pion “chiral splitting function” and the structure function of the (nearly on-shell) exchanged pion. Consequently the LN data have never been used in global QCD analyses, although recently the first steps toward their inclusion were taken by McKenney et al. [12], who studied the impact of the model dependence on the extracted pion structure function by constraining the $p \rightarrow n\pi^+$ splitting function empirically.

In the present work, we combine the strategy of global QCD analysis with an empirical approach to using the DY and LN data in the same fit to determine the pion PDFs in both the high-$x_\pi$ and low-$x_\pi$ regions. We use for the first time a Monte Carlo (MC) approach, based on the nested sampling algorithm [13–15], to perform the global analysis at next-to-leading order (NLO) in the strong coupling. In contrast to the previous single-fit analyses based on maximum likelihood methods [6–9], the MC approach allows a systematic exploration of the parameter space by computing the likelihood function directly, providing a rigorous determination of PDF uncertainties.

An important feature of our analysis is the ability of the MC fit to quantify the uncertainty on the dependence of the extracted pion PDFs on the chiral splitting function model. To test the robustness of the chiral framework, we also perform a simultaneous fit to the pion DY + LN data together with the E866 pd/pp DY data [16], from which the $d/u$ ratio was extracted, using the same MC methodology. Such an analysis provides the most comprehensive study of pion PDFs and their impact on different observables.

In the pion-induced DY process [1], partons from the pion and target nucleus A annihilate to produce a dimuon pair in the final state, $\pi A \rightarrow \mu^+\mu^-X$, with cross section

$$\frac{d^2\sigma}{dQ^2dY} = \frac{4\pi\alpha^2}{9Q^2S} \sum_{ij} \int_{x_\pi}^1 \frac{d\hat{x}_\pi}{\hat{x}_\pi} \int_{x_A}^1 \frac{d\hat{x}_A}{\hat{x}_A} \times C_{ij}(\hat{x}_\pi, \hat{x}_A, x_\pi, x_A, Q/\mu) f_i^\pi(\hat{x}_\pi, \mu) f_j^A(\hat{x}_A, \mu),$$

where $f_i^\pi(f_j^A)$ is the PDF for parton flavor $i$ in the pion
(flavor \(j\) in the nucleus) as a function of the parton momentum fraction \(x_A\), and \(C_{ij}\) is the hard scattering kernel [17, 18], with \(\mu\) the renormalization scale. The cross section is differential in the dilepton invariant mass squared \(Q^2\) and rapidity \(Y\), in terms of which one defines \(x_{\pi,A} = \sqrt{\tau} e^{-Y}\), where \(\tau = Q^2/S\) and \(S\) is the \(\pi\)-target invariant mass squared. At lowest order the parton momentum fractions are given by \(x_{\pi,A} = x_{\pi,A}\). Typically, the experimental DY cross section cuts are analyzed in terms of the Feynman variable \(x_F = x_x - x_N\), where \(x_N = Ax_A\) is the nuclear Bjorken variable scaled per nucleon [19]. The available pion DY data from the CERN NA10 [4] and Fermilab E615 [5] experiments were taken on tungsten nuclei, and for the nuclear PDFs in our analysis we use the parametrization from Eskola et al. [20].

For the LN cross section, \(\sigma_{\pi\pi}^{LN}\), the differential LN structure function, \(F_2^{LN}(x,Q^2,x_L)\), which for \(\pi^+\) exchange is given by [12]

\[
F_2^{LN}(x,Q^2,x_L) = 2f_{\pi N}(x_L)F_2^\pi(x,Q^2). \tag{2}
\]

Here \(f_{\pi N}(x_L)\) is the chiral splitting function for fraction \(x_L \equiv 1-xL = x/x_\pi\) of the proton’s light-cone momentum carried by the pion, and \(F_2^\pi(x,Q^2)\) is the pion structure function, evaluated at NLO. The splitting function is evaluated from chiral effective theory [21–23], and for \(x_L > 0\) is given by

\[
f_{\pi N}(x_L) = \frac{g_\pi^2 M^2}{4\pi f_{\pi\pi}} \int dk_1^2 \frac{x_L[k_1^2 + x_L^2 M^2]}{x_L D_{\pi N}^2} |\mathcal{F}|^2, \tag{3}
\]

where \(D_{\pi N} \equiv t-m_\pi^2 = -[k_1^2 + x_L^2 M^2 + x_L m_\pi^2]/x_L\), with \(M\) and \(m_\pi\) the nucleon and pion masses, \(g_A\) the axial charge, and \(f_\pi\) the pion decay constant. The form of the splitting function in Eq. (3) is constrained by chiral symmetry [21, 24, 25], and its infrared or leading nonanalytic behavior is model independent [26–29]. The ultraviolet behavior, however, is dependent on the regularization procedure, represented in Eq. (3) by the function \(\mathcal{F}\). In the literature various forms have been advocated, including cutoff regularization, Pauli-Villars, and phenomenological \(\pi N\) form factors, and following Ref. [12] we consider several forms,

\[
\mathcal{F} = \begin{cases} 
(i) \exp\left[(M^2 - s)/\Lambda^2\right] & \text{s-dep. exponential} \\
(ii) \exp\left(D_{\pi N}^2/\Lambda^2\right) & \text{t-dep. exponential} \\
(iii) (A^2 - m_\pi^2)/(A^2 - t) & \text{t-dep. monopole} \\
(iv) \frac{x_L^{\alpha_\pi(t)}\exp\left(D_{\pi N}^2/\Lambda^2\right)}{\left[1 - D_{\pi N}^2/(A^2 - t)^2\right]^{1/2} \text{Pauli-Villars}} & \text{Regge} \\
\end{cases}
\]

where \(s = (k_1^2 + M^2)/x_L + (k_2^2 + m_\pi^2)/x_L\), \(\alpha_\pi(t) = \alpha_\pi' t\), with \(\alpha_\pi' \approx 1 \text{ GeV}^{-2}\), and \(\Lambda\) is a cutoff parameter. We also considered a model based on a large-\(k_\perp\) cutoff [23], and the Bishari model [11] (which is analogous to the Regge form but with \(\Lambda \to \infty\), but found poor agreement between the fitted results and data, and do not consider them further.

In addition to the LN structure function data collected by H1 [10], the ZEUS collaboration measured the ratio [11]

\[
r(x,Q^2,x_L) = \frac{d^3\sigma_{LN}/dx^2dQ^2}{d^3\sigma_{\pi\pi}/dx^2dQ^2} \Delta x_L \tag{4}
\]

of LN to inclusive cross sections, where the latter is expressed in terms of the proton structure function, \(F_2^\pi\), and \(\Delta x_L\) is the bin size in \(x_L\). Consistent with expectations from earlier theoretical calculations [30, 31], at large \(x_L \sim 1\) pion exchange is the dominant contribution [12]. Other processes, such as absorption and the exchange of other mesons, play an increasingly important role at smaller \(x_L\). Instead of choosing a specific value of \(x_L\) below which one pion exchange is assumed, we fit the lower value of \(x_L\) for which the data can be described within this framework.

For the data analysis we use a Bayesian Monte Carlo method based on the nested sampling algorithm [13–15], which allows a faithful Monte Carlo representation of the probability distribution \(P(a|\text{data}) = \mathcal{L}(\text{data}|a)\pi(a)/\mathcal{Z}\), where \(\mathcal{Z}\) is an \(n\)-dimensional array of the pion PDFs shape parameters. Here \(\pi(a)\) is the Bayesian prior distribution for \(a\), which allows the parameter sampling to be restricted to physical regions, \(\mathcal{L}(\text{data}|a) = \exp[-\chi^2(a)]\) is the likelihood function, and \(Z = \int d^a a \mathcal{L}(\text{data}|a)\pi(a)\) is the Bayesian evidence parameter, which normalizes the probability distribution. We use a \(\chi^2\) function in the likelihood that takes into account correlated systematic shifts, as well as overall normalizations of the data sets [32]. For physical observables \(\mathcal{O}\), such as the pion PDFs and functions thereof, from the MC samples \(\{a_k\}\) one then obtains expectation values \(E[\mathcal{O}] = \sum_k w_k \mathcal{O}(a_k)\) and variances \(V[\mathcal{O}] = \sum_k w_k (\mathcal{O}(a_k) - E[\mathcal{O}])^2\), where \(\{w_k\}\) are the MC weights. Similar MC technology based on Bayesian statistics has also been applied recently to study nucleon PDFs [33, 34] and fragmentation functions [35], as well as the transverse momentum dependent transversity distribution [36].

For the pion valence PDFs we assume charge symmetry, \(q_\pi^\pm = u_\pi^\pm = u_\pi^\pm - d_\pi^\pm\), and invoke SU(3) symmetry for the pion sea, \(q_\pi^\pm = d_\pi^\pm = s_\pi^\pm = \bar{s}_\pi^\pm\). The valence, sea quark, and gluon PDFs are parameterized at the input scale of the charm quark mass \(Q_W^0 = m_c^2 = (1.3 \text{ GeV})^2\) by the form

\[
f(x, Q_0^2; a) = \frac{N}{B(2 + \alpha, \beta)} x_{\pi}^\alpha (1 - x_{\pi})^\beta, \tag{5}
\]

where \(a = \{N, \alpha, \beta\}\) are the fitting parameters and \(B\) is the Euler beta function. The valence PDFs are normal-
ized such that \( \int_0^1 dx_\pi q^0_\pi = 1 \), and the momentum sum rule gives the constraint \( \int_0^1 dx_\pi x_\pi (2q^0_\pi + 6q^+ + g^2) = 1 \).

The fits to the DY and LN data sets are shown in Fig. 1, where for clarity the E615 and HERA points are scaled by 3'. To avoid the \( J/\Psi \) and \( \Upsilon \) resonances, the DY data were restricted to the mass region 4.16 < \( Q < 8.34 \) GeV, covering the range 0.05 < \( x_F \) < 0.9. Generally very good agreement is found for the entire set of 250 data points. For the best fit, corresponding to model (i) for the LN cross section with a cutoff \( \Lambda = 1.31(4) \) GeV, the combined \( \chi^2/N_{\text{dat}} = 0.98 \) (244.8/250). The overall normalizations for the DY data are found to be 0.816, 0.758 and 0.985 for the NA10 (194 GeV), NA10 (252 GeV) and E615 data sets, and 1.17 and 0.964 for the H1 and ZEUS LN data, respectively.

For the LN data good fits were obtained for the cut \( x_L > 0.8 \); including smaller-\( x_L \) data deteriorated the fit due to larger non-pionic contributions away from the forward limit [12, 30, 31]. Fitting the DY data alone yields only marginally smaller \( \chi^2 \) values, with \( \chi^2/N_{\text{dat}} = 0.97 \) (55.5/70 for NA10 and 82.6/72 for E615). For the combined DY and LN data sets, the total \( \chi^2 \) per datum for other models are also close to 1.0, with \( \chi^2 = \{267.7, 266.0, 262.8, 273.8\} \) for models (ii)–(v), corresponding to cutoff parameters \( \Lambda = \{0.58(2), 0.52(2), 0.78(5), 0.25(1)\} \) GeV, respectively.

The resulting pion PDFs are shown in Fig. 2 for the valence, sea quark, and gluon distributions at \( Q^2 = 10 \) GeV.\(^2\). Compared with the DY-only fits, which constrain mainly the valence quark PDF, the simultaneous DY+LN analysis yields significantly reduced uncertainties on the pion sea and gluon distributions at \( Q^2 = 10 \) GeV\(^2\). While the addition of the LN data gives a small, \( \approx 10\% \) reduction of the valence PDF at intermediate \( x_\pi \), the impact on the sea is more dramatic, with the gluon PDF increasing twofold at \( x_\pi \sim 0.001–0.1 \) compared with the DY-only result, but with half of the uncertainty. The sea quark PDF \( q^0_\pi \) is reduced at \( x_\pi \geq 0.1 \), but is slightly larger at \( x_\pi \lesssim 0.1 \) for the full result. Importantly, the model dependence of the combined DY+LN fit (represented in Fig. 2 by the yellow bands) reveals a relatively small uncertainty, especially compared with the scale of the effect induced by the addition of the LN results.

For the valence PDF in the large-\( x_\pi \) region, our analysis finds a behavior \( \sim (1 - x) \) at the input scale, consistent with Nambu-Jona-Lasinio [37] and constituent quark models [38] and duality arguments [39], but harder than expectations based on pQCD [40] and Dyson-Schwinger calculations [41], which prefer a \( (1 - x)^2 \) fall-off. On the other hand, the present analysis does not include threshold resummation effects, which are known to be important at large \( x_\pi \) [9, 42], which will be examined in a separate analysis [43].

The inclusion of the LN data into the global analysis allows a more precise breakdown of the pion momentum into fractions carried by valence quarks, sea quarks, and gluons, shown in Fig. 3. The total valence, sea and gluon momentum fractions for the full analysis at the input scale \( Q^2 = m^2 \) are found to be \( \{x_v, x_s, x_g\} = \{54(1)\%, 16(2)\%, 30(2)\%\} \) for the best fit with model (i). Compared with the DY-only fit, where the respective momentum fractions are \{60(1)\%, 30(5)\%, 10(5)\%\}, the
fraction carried by gluons is about 3 times larger, but with less than half of the uncertainty. Since the valence fraction remains relatively unchanged, the momentum sum rule forces the sea quarks to carry about 1/2 of the momentum fraction compared with the DY-only fit.

This turns out to be similar to the result from the older SMRS analysis [6] of DY data, which considered several scenarios in which the momentum fraction carried by sea quarks at $Q^2 = 5$ GeV$^2$ varied from 10% to 20%. With the valence momentum fraction in [6] constrained to 46%, the momentum fraction carried by gluons varied from 43% to 34% over this range. At the same scale the analogous fractions from our analysis are (48(1)%, 17(1)%, 35(2)%), which is closer to the SMRS scenario with maximum sea and minimum glue. In contrast, the GRS analysis [8], which assumes a constituent quark model to constrain the sea quark and gluon PDFs, gives a gluon momentum fraction (44%) that is actually larger than the valence fraction (39%) at this scale.

As an application of our results and a test of the validity of the chiral framework for the LN data, we consider the contribution from the $p \rightarrow n\pi^+$ dissociation to the $\bar{d} - \bar{u}$ asymmetry in the proton sea. This asymmetry, predicted from pion loop effects in the nucleon [2], was conclusively established by the $pp$ and $pd$ Drell-Yan data from the E866 experiment at Fermilab [16]. While the effect of these data on PDFs is most rigorously quantified through global QCD fits, an approximate NLO analysis of the E866 data extracted the $\bar{d} - \bar{u}$ difference, shown in Fig. 4, assuming knowledge of the valence quark distributions in the proton.

The antiquark asymmetry can be represented within chiral effective theory as a convolution (represented by the symbol “⊗”) of the $p \rightarrow n\pi^+$ and $p \rightarrow \Delta^{++}\pi^−$ splitting functions and the valence pion PDF [44, 45].

The antiquark asymmetry is shown in Fig. 4, for all models (i)–(v) for the regulator $F$. Without compromising the description of the DY or LN data, a reasonably good fit to the E866 data points can be achieved in all cases for $x \lesssim 0.2$, beyond which all the fits overestimate the data. It is known that the apparent change of sign in $\bar{d} - \bar{u}$ at high $x$ is difficult to accommodate theoretically [46], and the new DY SeaQuest experiment [47] will allow a more precise determination of the asymmetry up to $x \approx 0.5$.

In the future, pion PDFs will be further constrained by new $\pi A$ DY data from COMPASS [48, 49], as well
as from the Tagged DIS (TDIS) [50] experiment at Jefferson Lab, which will study pion structure through the charge exchange mechanism in leading proton production form a quasi-free neutron in the deuteron, \( ed \to eppX \). Theoretically, effects from gluon resummation [9, 42] and higher twists [40] will be explored [43] systematically in order to unravel the behavior of pion PDFs at very high \( x_F \sim 1 \). Beyond this, an ultimate future goal will be a simultaneous fit of pion, proton and nuclear PDFs within the same MC global QCD analysis.

This work was supported by the U.S. Department of Energy (DOE) Contract No. DE-AC05-06OR23177, under which Jefferson Science Associates, LLC operates Jefferson Lab, by NSFC under Grant No. 11475186, CRC 110 by DFG and NSFC, and by DOE Contract No. DE-FG02-03ER41260.

[1] S. D. Drell and T.-M. Yan, Phys. Rev. Lett. 25, 316 (1970) [Erratum ibid. 25, 902 (1970)].
[2] A. W. Thomas, Phys. Lett. B 126, 97 (1983).
[3] J. Badier et al., Z. Phys. C 18, 281 (1983).
[4] B. Betev et al., Z. Phys. C 28, 9 (1985).
[18] C. Anastasiou, L. Dixon, K. Melnikov and F. Petriello, Eur. Phys. J. A 7, 109 (2000).
[19] B. Z. Kopelevich, I. K. Potashnikova, B. Povh and I. Schmidt, Phys. Rev. D 85, 114025 (2012).
[20] L. A. Harland-Lang, A. D. Martin, P. Motylinski and R. S. Thorne, Eur. Phys. J. C 75, (2015) 204.
[21] T. Shigetani, K. Suzuki and H. Toki, Phys. Lett. B 436, 1 (1998).
[22] M. Burkardt, K. S. Hendricks, C. R. Ji, W. Melnitchouk and A. W. Thomas, Phys. Rev. D 87, 054024 (2013).
[23] Y. Salam, C.-R. Ji, W. Melnitchouk and F. Petriello, Phys. Rev. D 90, 034022 (2014).
[24] C. Anastasiou, L. Dixon, K. Melnikov and F. Petriello, Phys. Rev. Lett. 110, 179101 (2013).
[25] A. W. Thomas, W. Melnitchouk and F. M. Steffens, Phys. Rev. Lett. 85, 25892 (2000).
[26] W. Detmold, W. Melnitchouk, J. W. Negele, D. B. Rennier and A. W. Thomas, Phys. Rev. Lett. 87, 172001 (2001).
[27] J.-W. Chen and X. Ji, Phys. Lett. B 523, 107 (2001).
[28] D. Arndt and M. J. Savage, Nucl. Phys. A697, 429 (2002).
[29] U. D’Alesio and H. J. Pirner, Eur. Phys. J. A 7, 109 (2000).
[30] B. Z. Kopelevich, I. K. Potashnikova, B. Povh and I. Schmidt, Phys. Rev. D 85, 114025 (2012).
[31] L. A. Harland-Lang, A. D. Martin, P. Motylinski and R. S. Thorne, Eur. Phys. J. C 75, (2015) 204.
[32] N. Sato, W. Melnitchouk, S. E. Kuhn, J. J. Ethier and A. Accardi, Phys. Rev. D 93, 074005 (2016).
[33] J. J. Ethier, N. Sato and W. Melnitchouk, Phys. Rev. Lett. 119, 132001 (2017).
[34] N. Sato, J. J. Ethier, W. Melnitchouk, M. Hirai, S. Kumar and A. Accardi, Phys. Rev. D 94, 114004 (2016).
[35] H. W. Lin, W. Melnitchouk, A. Prokudin, N. Sato and H. Shows, arXiv:1710.09858 [hep-ph], to appear in Phys. Rev. Lett.
[36] T. Shigetani, K. Suzuki and H. Toki, Phys. Lett. B 308, 383 (1993).
[37] A. Szczepaniak, C.-R. Ji and S. R. Cotanch, Phys. Rev. D 49, 3466 (1994).
[38] W. Melnitchouk, Eur. Phys. J. A 17, 223 (2003); Phys. Rev. D 67, 077502 (2003).
[39] G. R. Farrar and D. R. Jackson, Phys. Rev. Lett. 43, 246 (1979); E. L. Berger and S. J. Brodsky, Phys. Rev. Lett. 42, 940 (1979); S. J. Brodsky and F. Yuan, Phys. Rev. D 74, 094018 (2006); F. Yuan, Phys. Rev. D 69, 051501 (2004).
[40] M. B. Hecht, C. D. Roberts and S. M. Schmidt, Phys. Rev. C 63, 052123 (2001).
[41] D. Westmark and J. F. Owens, Phys. Rev. D 95, 056024 (2017).
[42] M. C. Barry et al., in preparation.
[43] A. W. Schreiber, P. J. Mulders, A. I. Signal and A. W. Thomas, Phys. Rev. D 45, 3069 (1992).
[44] J.-W. Chen and X. Ji, Phys. Rev. Lett. 85, 094018 (2000); F. Yuan, Phys. Rev. D 69, 051501 (2004).
[45] W. Melnitchouk, J. Speth and A. W. Thomas, Phys. Rev. Lett. 85, 2892 (2000).
[46] W. Melnitchouk, J. Speth and A. W. Thomas, Phys. Rev. D 88, 052002 (2003).
[47] Fermilab E906 Experiment (SeaQuest), Drell-Yan Measurements of Nucleon and Nuclear Structure with the Fermilab Main Injector, http://www.phy.anl.gov/mep/SeaQuest/index.html.
[48] The Drell-Yan project at COMPASS, http://wwwcompass.cern.ch/compass/future_physics/drellyan/.
[49] C. Franco, on behalf of the COMPASS collaboration, XIV Hadron Physics, Florianopolis, March 2018, Polarised Drell-Yan results from COMPASS.
[50] J. Annand, D. Dutta, C. E. Keppel, P. King and B. Wojtsekhowski (spokespersons), Jefferson Lab experiment E12-15-006.