FERMILAB-PUB-10-732-T

Nuclear Corrections in Neutrino-Nucleus Deep Inelastic Scattering and their Compatibility with Global Nuclear Parton-Distribution-Functions Analyses

K. Kovařík,[1,2] I. Schienbein,[1] F. I. Olness,[3] J. Y. Yu,[3]
C. Keppel,[4,5] J. G. Morfín,[6] J.F. Owens,[7] and T. Stavreva[1]

[1] *LPSC, Université Joseph Fourier/CNRS-IN2P3/INPG, UMR5821, Grenoble, F-38026, France*
[2] *Karlsruhe Institute of Technology, Karlsruhe, D-76128,Germany*
[3] *Southern Methodist University, Dallas, Texas 75275, USA*
[4] *Thomas Jefferson National Accelerator Facility, Newport News, Virginia 23602, USA*
[5] *Hampton University, Hampton, Virginia, 23668, USA*
[6] *Fermilab, Batavia, Illinois 60510, USA*
[7] *Florida State University, Tallahassee, Florida 32306-4350, USA*

We perform a global χ^2-analysis of nuclear parton distribution functions using data from charged current neutrino–nucleus (νA) deep-inelastic scattering (DIS), charged-lepton–nucleus ($\ell^\pm A$) DIS, and the Drell-Yan (DY) process. We show that the nuclear corrections in νA DIS are not compatible with the predictions derived from $\ell^\pm A$ DIS and DY data. We quantify this result using a hypothesis-testing criterion based on the χ^2 distribution which we apply to the total χ^2 as well as to the χ^2 of the individual data sets. We find that it is not possible to accommodate the data from νA and $\ell^\pm A$ DIS by an acceptable combined fit. Our result has strong implications for the extraction of both nuclear and proton parton distribution functions using combined neutrino and charged-lepton data sets.

PACS numbers: 12.38.-t,13.15.+g,13.60.-r,24.85.+p

arXiv:1012.0286v3 [hep-ph] 25 Mar 2011

High statistics neutrino deep-inelastic scattering (DIS) experiments have generated significant interest in the literature as they provide crucial information for global fits of parton distribution functions (PDFs). The neutrino DIS data provide the most stringent constraints on the strange quark distribution in the proton, and allow for flavor decomposition of the PDFs which is *essential* for precise predictions of the benchmark gauge boson production processes at the LHC. Moreover, the neutrino experiments have been used to make precision tests of the Standard Model (SM). A prominent example is the extraction of the weak mixing angle θ_W in a Paschos-Wolfenstein type analysis [1]. A good knowledge of the neutrino DIS cross sections is also very important for long baseline experiments of the next generation which aim at measuring small parameters of the mixing matrix such as the mixing angle θ_{13} and eventually the CP violating phase δ.

Because of the weak nature of neutrino interactions the use of heavy nuclear targets is unavoidable, and this complicates the analysis of the precision physics discussed above since model-dependent nuclear corrections must be applied to the data. Our present understanding of the nuclear corrections is mainly based on charged-lepton–nucleus (ℓA) DIS data. In the early 80s, the European Muon Collaboration (EMC) [2] found that the nucleon structure functions F_2 for iron and deuterium differ. This discovery triggered a vast experimental program to investigate the nuclear modifications of the ratio $R[F_2^{\ell A}] = F_2^{\ell A}/(A\,F_2^{\ell N})$ for a wide range of nuclear targets with atomic number A, see Table I. By now, such modifications have been established in a kinematic range from relatively small Bjorken x ($x \sim 10^{-2}$) to large x ($x \sim 0.8$) in the deep-inelastic region with squared momentum transfer $Q^2 > 1$ GeV^2. The behavior of the ratio $R[F_2^{\ell A}]$ can be divided into four regions: (i) $R > 1$ for $x \gtrsim 0.8$ (Fermi motion region), (ii) $R < 1$ for $0.25 \lesssim x \lesssim 0.8$ (EMC region), (iii) $R > 1$ for $0.1 \lesssim x \lesssim 0.25$ (antishadowing region), and (iv) $R < 1$ for $x \lesssim 0.1$ (shadowing region), with different physics mechanisms explaining the nuclear modifications. The shadowing suppression at small x occurs due to coherent multiple scattering inside the nucleus of a $q\bar{q}$ pair coming from the virtual photon [3] with destructive interference of the amplitudes [4]. The antishadowing region is theoretically less well understood but might be explained by the same mechanism with constructive interference of the multiple scattering amplitudes [4] or by the application of momentum, charge, and/or baryon number sum rules. Conversely, the modifications at medium and large x are usually explained by nuclear binding and medium effects and the Fermi motion of the nucleons [5].

Instead of trying to address the origin of the nuclear effects, the data on nuclear structure functions can be analyzed in terms of nuclear PDFs (NPDFs) which are modified as compared to the free nucleon PDFs. Relying on factorization theorems in the same spirit as in the free nucleon case, the advantage of this approach is that the universal NPDFs can be used to make *predictions* for a large variety of processes in ℓA, νA, pA, and AA collisions. In addition, the nuclear correction factors required for the interpretation of the neutrino experiments can be calculated in a flexible way, taking into account the precise observable, the nuclear A, and the scale Q^2. The

Table I. The charged-lepton DIS data sets together with DY and with neutrino DIS data sets used in the fit. The table details the specific nuclear targets, and the number of data points with kinematical cuts ($Q^2 > 4\ GeV^2$, $W > 3.5\ GeV$). References for the data sets are cited in Refs. [6, 7]

ID	Observable	$A/A'(A)$	Experiment	# data
1	$F_2^A/F_2^{A'}$	He/D	SLAC-E139, NMC-95 re	15
2	$F_2^A/F_2^{A'}$	Li/D	NMC-95	11
3	$F_2^A/F_2^{A'}$	Be/D	SLAC-E139	3
4	$F_2^A/F_2^{A'}$	C/D	EMC-88,90, SLAC-E139	
			NMC-95,95 re, FNAL-E665-95	38
5	$F_2^A/F_2^{A'}$	N/D	BCDMS-85	9
6	$F_2^A/F_2^{A'}$	Al/D	SLAC-E049,E139	3
7	$F_2^A/F_2^{A'}$	Ca/D	EMC-90, SLAC-E139	
			NMC-95,re, FNAL-E665-95	17
8	$F_2^A/F_2^{A'}$	Fe/D	BCDMS-85,87	
			SLAC-E049,E139,E140	24
9	$F_2^A/F_2^{A'}$	Cu/D	EMC-88,93	27
10	$F_2^A/F_2^{A'}$	Ag/D	SLAC-E139	2
11	$F_2^A/F_2^{A'}$	Sn/D	EMC-88	8
12	$F_2^A/F_2^{A'}$	Xe/D	FNAL-E665-92	2
13	$F_2^A/F_2^{A'}$	Au/D	SLAC-E139	3
14	$F_2^A/F_2^{A'}$	Pb/D	FNAL-E665-95	3
15	$F_2^A/F_2^{A'}$	Be/C	NMC-96	14
16	$F_2^A/F_2^{A'}$	Al/C	NMC-96	14
17	$F_2^A/F_2^{A'}$	Ca/C	NMC-95,96	29
18	$F_2^A/F_2^{A'}$	Fe/C	NMC-95	14
19	$F_2^A/F_2^{A'}$	Pb/C	NMC-96	14
20	$F_2^A/F_2^{A'}$	C/Li	NMC-95	7
21	$F_2^A/F_2^{A'}$	Ca/Li	NMC-95	7
22	$F_2^A/F_2^{A'}$	He/D	Hermes	17
23	$F_2^A/F_2^{A'}$	Kr/D	Hermes	12
24	$F_2^A/F_2^{A'}$	Sn/C	NMC-96	111
25	$F_2^A/F_2^{A'}$	N/D	Hermes	19
32	$F_2^A/F_2^{A'}$	D	NMC-97	201
26	$\sigma_{DY}^{pA}/\sigma_{DY}^{pA'}$	C/D	FNAL-E772	9
27	$\sigma_{DY}^{pA}/\sigma_{DY}^{pA'}$	Ca/D	FNAL-E772	9
28	$\sigma_{DY}^{pA}/\sigma_{DY}^{pA'}$	Fe/D	FNAL-E772	9
29	$\sigma_{DY}^{pA}/\sigma_{DY}^{pA'}$	W/D	FNAL-E772	9
30	$\sigma_{DY}^{pA}/\sigma_{DY}^{pA'}$	Fe/Be	FNAL-E866	28
31	$\sigma_{DY}^{pA}/\sigma_{DY}^{pA'}$	W/Be	FNAL-E866	28
			$l^{\pm}A$ **DIS & DY Total:**	**708**
33	$d\sigma^{\nu A}/dx\,dy$	Pb	CHORUS ν	824
34	$d\sigma^{\nu A}/dx\,dy$	Pb	CHORUS $\bar{\nu}$	412
35	$d\sigma^{\nu A}/dx\,dy$	Fe	NuTeV ν	1170
36	$d\sigma^{\nu A}/dx\,dy$	Fe	NuTeV $\bar{\nu}$	966
37	$d\sigma^{\nu A}/dx\,dy$	Fe	CCFR di-μ	44
38	$d\sigma^{\nu A}/dx\,dy$	Fe	NuTeV di-μ	44
39	$d\sigma^{\nu A}/dx\,dy$	Fe	CCFR di-μ	44
40	$d\sigma^{\nu A}/dx\,dy$	Fe	NuTeV di-μ	42
			νA **Total:**	**3134**

factorization assumption in the nuclear environment is therefore a question of considerable theoretical and practical importance and global analyses of NPDFs based on ℓA DIS and fixed target Drell-Yan (DY) data confirm its validity in the presently explored kinematic range.

However, in a recent analysis [6] of νFe DIS data from the NuTeV collaboration we found that the nuclear correction factors are surprisingly different from the predictions based on the $\ell^{\pm}Fe$ charged-lepton results with important implications for global analyses of proton PDFs. This is not completely unexpected since the structure functions in charged current (CC) neutrino DIS and neutral current (NC) electron or muon DIS are distinct observables with different parton model expressions. From this perspective it is clear that the nuclear correction factors will not be exactly the same even for a *universal* set of NPDFs. Note also that some models in the literature predict differences between reactions in CC and NC DIS [8]. What is, however, unexpected is the degree to which the R factors differ between the structure functions $F_2^{\nu Fe}$ and $F_2^{\ell Fe}$. In particular the lack of evidence for shadowing in neutrino scattering down to $x \sim 0.02$ is quite surprising.

Table II. Summary table of a family of compromise fits.

w	$l^{\pm}A$	χ^2 (/pt)	νA	χ^2 (/pt)	total χ^2(/pt)
0	708	638 (0.90)	-	-	638 (0.90)
1/7	708	645 (0.91)	3134	4710 (1.50)	5355 (1.39)
1/2	708	680 (0.96)	3134	4405 (1.40)	5085 (1.32)
1	708	736 (1.04)	3134	4277 (1.36)	5014 (1.30)
∞	-	-	3134	4192 (1.33)	4192 (1.33)

The study in Ref. [6] left open the question, whether the neutrino DIS data could be reconciled with the charged-lepton DIS data by a better flavor separation of the NPDFs. In this letter, we address this question in the A-dependent framework of Ref. [7] by performing a global χ^2-analysis of the combined data from νA DIS, ℓA DIS and the DY process listed in Table I.

When combining neutrino and charged-lepton+DY data into a compromise fit, we introduce a weight parameter w into the χ^2 via:

$$\chi^2 = \sum_{l^{\pm}A\ \mathrm{data}} \chi_i^2 + \sum_{\nu A\ \mathrm{data}} w\chi_i^2 \quad . \tag{1}$$

The w factor allows us to adjust for the different number of points in the separate data sets, and provides a parameter that interpolates between the νA and the $\ell^{\pm}A$+DY data. We should stress that the χ^2 cited in Table II and also in the text is the standard χ^2; Eq. (1) is only used internally in the fitting procedure. We construct a set of compromise fits with weights $w = \{0, \frac{1}{7}, \frac{1}{2}, 1, \infty\}$ and study the dependence of the result on this weight. The fit to only neutrino data, denoted $w = \infty$ in Table II, is compatible with the results in [6]. Similarly, the fit to only charged-lepton+DY data, denoted $w = 0$, agrees well with the analysis in [7].

We first compute the nuclear correction factors $R[F_2^{\ell Fe}] \simeq F_2^{\ell Fe}/F_2^{\ell N}$ and $R[F_2^{\nu Fe}] \simeq F_2^{\nu Fe}/F_2^{\nu N}$ in the QCD parton model at next-to-leading order employing the NPDF fits in Table II for the numerator and free nucleon PDFs for the denominator. [9] The x-dependence

of $R[F_2^{\ell Fe}]$ and $R[F_2^{\nu Fe}]$ is shown in Fig. 1 a) and b), respectively, at $Q^2 = 5$ GeV^2. Similar results hold at $Q^2 = 20$ GeV^2 which we do not present here. We observe that the fit to only ℓA DIS+DY data ($w = 0$) well describes the SLAC and BCDMS points in Fig. 1 a). The same is true for the fit to only νA DIS data ($w = \infty$) which is compatible with the results from the NuTeV experiment [10] exemplified in Fig. 1 b). However, comparing the results obtained with the $w = 0$ and the $w = \infty$ fits one can see that they predict considerably different x-shapes.

The fits with weights $w = \{\frac{1}{7}, \frac{1}{2}, 1\}$ interpolate between these two incompatible solutions. As can be seen in Fig. 1 a) and b), with increasing weight the description of the ℓFe data is worsened in favor of a better agreement with the νFe points. This trend clearly demonstrates that the ℓFe and the νFe data pull in opposite directions. We identify the fits with $w = 1/2$ or $w = 1$ as the best candidates for a possible compromise. To be able to decisively accept or reject the compromise fits, we apply a statistical goodness-of-fit criterion [11–13] based on the probability distribution for the χ^2 given that the fit has N degrees of freedom:

$$P(\chi^2, N) = \frac{(\chi^2)^{N/2-1} e^{-\chi^2/2}}{2^{N/2}\Gamma(N/2)} \,. \qquad (2)$$

This allows us to define the percentiles ξ_p via $\int_0^{\xi_p} P(\chi^2, N) d\chi^2 = p\%$ where $p = \{50, 90, 99\}$. Here, ξ_{50} serves as an estimate of the mean of the χ^2 distribution and ξ_{90}, for example, gives us the value where

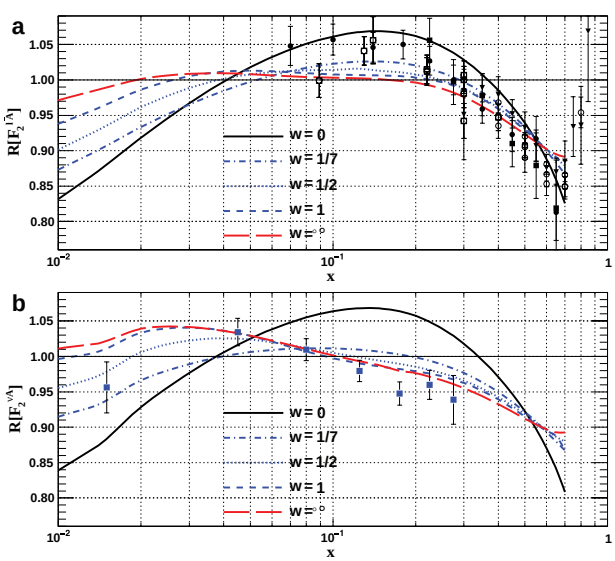

Figure 1. Predictions from the compromise fits for the nuclear correction factors $R[F_2^{\ell Fe}] \simeq F_2^{\ell Fe}/F_2^{\ell N}$ (a) and $R[F_2^{\nu Fe}] \simeq F_2^{\nu Fe}/F_2^{\nu N}$ (b) as a function of x for $Q^2 = 5\,\mathrm{GeV}^2$. The data points displayed in (a) are from BCDMS and SLAC experiments (for references see [7]) and those displayed in (b) come from the NuTeV experiment [10].

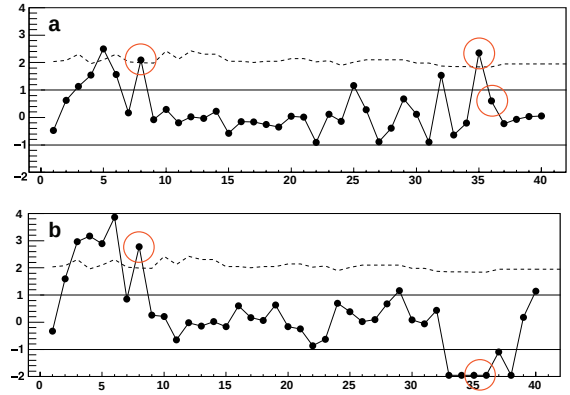

Figure 2. $\Delta\chi^2/\Delta C_{90}$ as defined in Eq. (3) for the 40 individual data sets. Results are shown for the $w = \frac{1}{2}$-fit (a) and the fit 'Ucor5' (b) with $w = 1$. The solid and dashed lines indicate the 90% and 99% confidence limits. The highlighted data sets correspond to DIS $\ell^\pm Fe$ (ID=8), νFe (ID=35), and $\bar{\nu} Fe$ (ID=36).

there is only a 10% probability that a fit with $\chi^2 > \xi_{90}$ genuinely describes the given set of data. In a global PDF fit, the best fit χ^2 value often deviates from the mean value because the data come from different possibly incompatible experiments having unidentified, unknown errors which are not accounted for in the experimental systematic errors. For this reason we rescale the ξ_{90} and ξ_{99} percentiles relative to the best fit χ_0^2 [11] to define $C_{90} = \chi_0^2(\xi_{90}/\xi_{50})$and $C_{99} = \chi_0^2(\xi_{99}/\xi_{50})$. This defines our criterion: a fit with a given χ^2 is compatible with the best fit with χ_0^2 at 90% (99%) confidence if $\chi^2 < C_{90}$ ($\chi^2 < C_{99}$). We apply it to both the total χ^2 and the χ^2 of the individual data sets.

For the ℓA DIS+DY data we use the fit with $w = 0$ as benchmark with $\chi_0^2 = 638$ and $N = 677$ degrees of freedom (for 708 data points and 31 free parameters). The upper limits on the χ^2 at 90% and 99% confidence level (C.L.) are then $C_{90}^{l^\pm A} = 684$ and $C_{99}^{l^\pm A} = 722$. The benchmark fit for the νA DIS data ($w = \infty$) uses 3134 data points with 33 free parameters resulting in $N = 3101$ and one finds $C_{90}^{\nu A} = 4330$ and $C_{99}^{\nu A} = 4445$. We see that none of the compromise fits satisfies both limits at the 90% C.L. which is usually used in global analyses of PDFs to define the uncertainty bands. At the 99% C.L., there are two fits ($w = 1/2$, $w = 1$) which are below the $C_{99}^{\nu A}$ limit. However, only the $w = 1/2$ fit satisfies the corresponding constraint from the charged-lepton benchmark fit.

We now apply our criterion also to the individual data sets with IDs between 1 and 40 in Table I. For the ℓA DIS+DY data (ID=[1,31]) we determine the 31 C_{90} (C_{99}) limits by using the individual χ_i^2 of the $w = 0$ fit as $\chi_{0,i}^2$. For the νA DIS data (ID=[33,40]) we proceed in a similar manner using the individual χ_i^2 of the $w = \infty$ fit. The results of this detailed analysis are depicted in Fig. 2,

where we show the quantity

$$\frac{\Delta\chi^2}{\Delta C_{90}} = \frac{\chi_i^2 - \chi_{0,i}^2}{C_{90,i} - \chi_{0,i}^2} \quad (i = 1, \ldots, 40)\,, \tag{3}$$

where χ_i^2 represents the χ^2-value of the i'th data set. In cases where $\chi_i^2 > C_{90,i}$ the fit is not compatible with the best fit at the 90% level and $\Delta\chi^2/\Delta C_{90} > 1$. The exact 90% C.L. limit is shown as a constant solid line and the dotted line represents the 99% confidence limit. The local application of the χ^2 hypothesis-testing criterion reveals that even the compromise fit with weight $w = \frac{1}{2}$ which was considered acceptable at the 99% C.L. when looking at the nuclear correction factors and at the global change in χ^2, cannot be accepted as a compromise solution as both the charged-lepton and neutrino DIS data on iron exceed the 99% limit.

In conclusion, the tension between the $\ell^\pm Fe$ and νFe data sets leaves us with no possible compromise fit when investigating the results in detail, not even when using the 99% percentile as the limit as opposed to the more restrictive 90% limit which is usually used to construct the error PDFs. This detailed analysis confirms the preliminary conclusions of Refs. [6, 7] that there is no possible compromise fit which adequately describes the neutrino DIS data along with the charged-lepton data.

At face value, this conclusion differs from some results in the literature which argue the νA and $\ell^\pm A$ data are in accord [14]. Here, we believe an essential element in our analysis is the use of the correlated systematic errors of the νA data. To highlight this point, we now repeat our analysis, *but* we combine the statistical and all systematic errors in quadrature (thereby neglecting the information contained in the correlation matrix) for νA data for the $w = 1$ fit with $Q^2 > 4\,GeV$ (as before); we denote this the "Ucor4" fit, and we obtain χ^2/pt of 1.14 for $\ell^\pm A$ and 1.00 for νA. We also use a $Q^2 > 5\,GeV$ fit (denoted "Ucor5") to mimic the cuts of Ref. [14]; here we obtain χ^2/pt of 1.14 for $\ell^\pm A$ and 0.96 for νA.

If we examine the total χ^2 values, we find the $\chi^2/dof \sim 1$, and might be tempted to conclude we are able to fit both the νA and $\ell^\pm A$ data simultaneously. However, if we look at individual data sets and apply our hypothesis-testing criteria, the picture is quite different. Figure 2 (b) displays the results for the Ucor5 fit. The higher Q^2 cut of the Ucor5 fit removes some of the very precise NuTeV data at small-x, thus resulting in an improved χ^2 compared to Ucor4. Nevertheless, many of the $\ell^\pm A$ data sets (ID=3,4,5,6,8) still lie outside the 99% CL percentile. Thus, we still conclude that there is no compromise fit for the νA and $\ell^\pm A$ data even if we relax the constraints by using uncorrelated errors.

Consequently, the nuclear correction factor for the neutrino DIS data are indeed *incompatible* with that of the charged-lepton DIS and DY data, and this result depends crucially on the use of the precision correlated errors of the neutrino data. This result has important implications for both nuclear and proton PDFs. If we do not know the appropriate nuclear correction to relate different nuclear targets, our ability to extract PDFs is limited. For example, the CTEQ6.6 analysis [15] sidesteps these issues by removing most of the νA data from the fit; however, they retain the NuTeV dimuon data since this data is critical to constraining the strange quark PDF. This underscores the importance of the νA data for flavor differentiation. Although the NuTeV data provide the tightest constraints due to their statistics, we note that this issue cannot be tied to a single data set. For example, we find that NuTeV is generally compatible with CCFR and CDHSW[16]. The CHORUS νPb and $\bar{\nu}Pb$ data have larger uncertainties, so they can be compatible with both the $\ell^\pm A$ data and the NuTeV νFe data because the $\Delta\chi^2/\Delta C_{90} < 1$ for all weights. Compared to the theory predictions, NuTeV agrees well in the central x region, but exhibits differences both for low x at low Q^2, and also for very high x ($x \sim 0.65$).

We have demonstrated that the νA and $\ell^\pm A$ data prefer different nuclear correction factors, and that there is no single "compromise" result that will simultaneously satisfy both data sets. While we have focused on the phenomenological aspects for the present study, this result has strong implications for the extraction of both nuclear and proton PDFs using combined neutrino and charged-lepton data sets. Possibilities include unexpectedly large higher-twist effects, or even nonuniversal nuclear effects; we leave such questions for a future study.

[1] G. P. Zeller *et al.* (NuTeV), Phys. Rev. Lett. **88**, 091802 (2002).
[2] J. J. Aubert *et al.* (European Muon), Phys. Lett. **B123**, 275 (1983).
[3] N. Armesto, J. Phys. **G32**, R367 (2006).
[4] S. J. Brodsky and H. J. Lu, Phys. Rev. Lett. **64**, 1342 (1990).
[5] D. F. Geesaman, K. Saito, and A. W. Thomas, Ann. Rev. Nucl. Part. Sci. **45**, 337 (1995).
[6] I. Schienbein *et al.*, Phys. Rev. **D77**, 054013 (2008).
[7] I. Schienbein *et al.*, Phys. Rev. **D80**, 094004 (2009).
[8] S. J. Brodsky, I. Schmidt, and J. J. Yang, Phys. Rev. **D70**, 116003 (2004).
[9] We allow each individual partonic flavor to have different nuclear corrections, and the details are outlined in Refs. [6, 7]. While we focus on F_2 , we can consider other observables such as $\{F_1, F_3, d\sigma\}$ in a similar manner.
[10] M. Tzanov *et al.* (NuTeV), Phys. Rev. **D74**, 012008 (2006).
[11] D. Stump *et al.*, Phys. Rev. **D65**, 014012 (2001).
[12] A. D. Martin *et al.*, Eur. Phys. J. **C63**, 189 (2009).
[13] K. J. Eskola, H. Paukkunen, and C. A. Salgado, JHEP **04**, 065 (2009).
[14] H. Paukkunen and C. A. Salgado, JHEP **07**, 032 (2010).
[15] P. M. Nadolsky *et al.*, Phys. Rev. **D78**, 013004 (2008).

[16] For $x < 0.4$, NuTeV is compatible with both CCFR and CDHSW data; for larger x, NuTeV agrees with CDHSW, and the difference with CCFR has been reconciled.