

Light quark and antiquark constraints from new electroweak data

A. Accardi,^{1,2} X. Jing³, J. F. Owens,⁴ and S. Park^{2,5,6}

¹Hampton University, Hampton, Virginia 23668, USA

²Jefferson Lab, Newport News, Virginia 23606, USA

³Southern Methodist University, Dallas, Texas 75205, USA

⁴Florida State University, Tallahassee, Florida 32306, USA

⁵Mississippi State University, Mississippi State, Mississippi 39762, USA

⁶Center for Frontiers in Nuclear Science, Stony Brook, New York 11794, USA



(Received 22 March 2023; accepted 12 June 2023; published 29 June 2023)

We present a new parton distribution function analysis which includes new data for W boson production in proton-proton collisions and lepton pair production in proton-proton and proton-deuteron collisions. The new data provide strong constraints on the light antiquark parton distribution functions in the proton. We identify an interesting correlation between the d/u ratio and the \bar{d}/\bar{u} ratio which leads to a modification of our previous results for the d/u ratio as the parton momentum fraction $x \rightarrow 1$.

DOI: [10.1103/PhysRevD.107.113005](https://doi.org/10.1103/PhysRevD.107.113005)

I. INTRODUCTION

Predictions for high energy lepton-hadron and hadron-hadron hard collisions rely on perturbative QCD-based calculations for the parton-parton scattering cross sections. These are then convoluted with the appropriate parton distribution functions (PDFs) to obtain predictions for experimentally measured observables. A recent evaluation of PDF determinations can be found in Ref. [1]. Global fits for PDFs by the CJ Collaboration have focused on simultaneously extending the reach in x toward $x \approx 1$ and reducing the minimum value of the squared four-momentum transfer, Q^2 , included in the fitting process [2,3]. The focus of this analysis is, instead, on the flavor dependence of the light-quark sea, specifically the behavior in x of the ratio \bar{d}/\bar{u} . One important source of information on the behavior of this ratio is lepton pair production using a proton beam on both proton and deuteron targets. The most precise such data available previously came from the E866 experiment [4]. These data suggested that \bar{d}/\bar{u} initially rose from a value of unity to a maximum near $x \approx 0.15$ followed by a fall-off to a value below unity by $x \approx 0.30$. However, by this region in x the data were statistically limited. The CJ12 PDFs [2] were parametrized in such a way that the ratio could follow the data to values below one. This led to a rapid falloff of the \bar{d} PDF below 0 as x increased beyond about 0.3. The CJ15 PDFs [3] employed an alternative parametrization which

constrained \bar{d}/\bar{u} to approach one from above at large values of x .

New data from the SeaQuest experiment [5], the successor to the E866 experiment, have a greater reach in x as well as increased statistics. Additionally, new data [6] from the STAR Collaboration on W boson production in proton-proton collisions have become available. These data also offer additional constraints on \bar{d}/\bar{u} . It is the purpose of this analysis to assess the effects of these new datasets on the behavior of \bar{d}/\bar{u} over the x range out to $x \approx 0.4$. At the same time, we expose the effects of the (anti)correlation between the light-antiquark and light-quark ratios inherent in the available lepton pair production data and in the mid-rapidity weak boson production data, that affects the extrapolation of the d/u ratio to values of x approaching 1. In particular, we examine different parametrizations to see their effects on the extracted ratios. Preliminary results have been presented at DIS 2021 [7], and analyses of the new data have also been performed by the CT [8] and JAM [9] collaborations.

The plan of the paper is as follows. In Sec. II the framework for the global fits is described, including the parametrizations used for the various PDFs and the higher-twist and nucleon off-shell corrections. Section III contains a discussion of the datasets used with special attention paid to the new data, while Sec. IV presents the results of this analysis. The conclusions are summarized in Sec. V.

II. LIGHT QUARKS AND ANTIQUARKS IN THE CJ GLOBAL ANALYSIS

The new CJ22 global fit we report in this paper combines elements of the CJ15 [3] and CJ15-a [10]

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP³.

analyses in order to provide sufficient flexibility in the determination of the mid- x \bar{d}/\bar{u} ratio after the inclusion of the STAR and SeaQuest data. The new fit will also allow us to properly analyze the correlation of the mid- x \bar{d}/\bar{u} ratio and the large- x d/u ratio induced by weak boson production data and how these impact the extrapolation of d/u to $x \rightarrow 1$. In this section we focus on the methodological and numerical aspects of our fits, and in the next one we will discuss the global dataset we utilize.

A. PDF parametrization and theoretical setup

The latest CTEQ-JLab global fit (CJ15) was performed using the world deep-inelastic scattering (DIS) dataset, as well as a variety of jet and electroweak boson production measurements [3]. Among these, lepton pair production measurements by the E866 experiment at Fermilab provided the strongest constraints on the light antiquark sea, covering the $0.015 \lesssim x \lesssim 0.3$ parton momentum fraction region, with additional sensitivity provided by fixed target DIS data, in particular from the NMC experiment.

At the input scale of $Q_0^2 = 1.69 \text{ GeV}^2$, a standard five-parameter functional form was used for most of parton species, including the $\bar{u} + \bar{d}$ combination:

$$xf(x, Q_0^2) = a_0 x^{a_1} (1-x)^{a_2} (1 + a_3 \sqrt{x} + a_4 x) \quad (1)$$

The valence d quark was however allowed to mix with the valence u quark parametrization at large x , as to allow a finite limit for the d/u ratio:

$$d_v(x, Q_0^2) \rightarrow a_0^{d_v} \left(\frac{d_v(x, Q_0^2)}{a_0^{d_v}} + b x^c u_v(x, Q_0^2) \right), \quad (2)$$

with b and c as two additional parameters. As a result, the ratio d_v/u_v could tend to a nonzero value as $x \rightarrow 1$, provided that $a_2^{d_v} > a_2^{u_v}$, which is usually the case. As in the CJ15 and CJ15-a fits, only the b parameter was left free and $c = 2$ kept fixed, since the new data do not provide additional constraints on the valence quark ratio at large x .

Turning to the light antiquarks, the \bar{d}/\bar{u} ratio in the original CJ15 fit was parametrized as

$$\bar{d}/\bar{u} = a_0 x^{a_1} (1-x)^{a_2} + 1 + a_3 x (1-x)^{a_4} \quad (3)$$

due to the limited x coverage of the E866 data and the sharp downturn these required of the \bar{d}/\bar{u} ratio. With this parametrization we also enforced the theoretical expectation from most modeling efforts that \bar{d}/\bar{u} remains greater than or equal to one all the way up to $x \rightarrow 1$, and tends to 1 in that limit. This assumption was however revisited in Ref. [10], where the $\bar{d} - \bar{u}$ difference was considered instead of \bar{d}/\bar{u} and parametrized as in Eq. (1),

$$x(\bar{d} - \bar{u}) = \bar{a}_0 x^{\bar{a}_1} (1-x)^{\bar{a}_2} (1 + \bar{a}_4 x), \quad (4)$$

with the resulting fit called CJ15-a. Even if the new parametrization allowed for it, no strong indication of a sign change in the $\bar{d} - \bar{u}$ asymmetry in the $x \lesssim 0.3$ region measured by E866 was found.

With the new data from STAR sensitive to the smaller- x rise of the $\bar{d} - \bar{u}$ asymmetry, and the new SeaQuest data constraining the \bar{d}/\bar{u} ratio at $0.15 \lesssim x \lesssim 0.4$, well across the region where E866 indicated this would drop below 1, we can now revisit this whole issue. To allow for sufficient versatility in the description of the light quark sea, in this paper we will utilize the more flexible parametrization (4). Furthermore, we will leave the \bar{a}_2 parameter free instead of fixing it to 2.5 units larger than the corresponding parameter for the $\bar{u} + \bar{d}$ combination as done in the CJ15-a analysis, thus providing additional freedom to the \bar{d}/\bar{u} ratio in the limiting $x \rightarrow 1$ region.

Apart from this change in parametrization, we will adopt the same theoretical setup as in the CJ15 fits, as described in Ref. [3]. In particular: we perform fits at next-to-leading order accuracy in the ACOT- χ heavy quark scheme; include target mass corrections for DIS data according to the OPE prescription by Georgi and Politzer [11,12]; and adopt the “weak-binding approximation” to correct for nucleon binding and Fermi motion in DIS and DY cross sections on deuteron targets, with the AV18 deuteron wave function describing the nucleon dynamics inside the target. Higher-twist corrections for DIS structure functions and off-shell nucleon corrections in deuteron targets will be discussed in more detail in the next subsection.

From a numerical point of view, and at variance with the CJ15 family of analyses, NLO QCD corrections to the calculation of W and Z production cross sections were implemented by means of the APPLgrid [13] fast NLO interface. The necessary coefficient grids were calculated by means of the MCFM 6.8 event generator [14,15], and tested against the Tevatron weak boson production data already included in the CJ15 analysis. Details about the grids for weak boson production at the Relativistic Heavy Ion Collider (RHIC) will be discussed in Sec. III.

B. Higher-twist and off-shell corrections

In DIS at low Q^2 values, power suppressed corrections exist beyond target mass corrections, for example, genuine multiparton correlations; missing higher order perturbative corrections can also resemble power corrections at small scale values. Regardless of their origin, we account for these residual power suppressed contributions by using a phenomenological multiplicative factor to modify the proton and nucleon structure functions as in all earlier CJ fits,

$$F_2(x, Q^2) = F_2^{\text{LT}}(x, Q^2) \left(1 + \frac{C(x)}{Q^2} \right), \quad (5)$$

where F_2^{LT} denotes the leading twist structure function including target mass corrections, and C is assumed to be isospin independent due to the relatively weak constraining power of the adopted datasets [16–18]. Following common usage, we generically refer to the fitted $1/Q^2$ term as a “higher-twist” (HT) correction, and parametrize the coefficient function C by

$$C(x) = a_{\text{HT}} x^{b_{\text{HT}}} (1 + c_{\text{HT}} x). \quad (6)$$

For ease of notation, we collect the higher-twist parameters in the vector

$$\mathbf{a}_{\text{HT}} = (a_{\text{HT}}, b_{\text{HT}}, c_{\text{HT}}). \quad (7)$$

In deuteron targets, the nucleons are off their m_N^2 mass shell with a four-momentum squared $p_N^2 \neq m_N^2$. While the off-shell nucleon PDF, \tilde{f} , is not an observable *per se*, its dependence on the nucleon virtuality p_N^2 can be studied within a given theoretical framework. For weakly bound nucleons such as in the deuteron, for example, one may expand \tilde{f} to lowest order about its mass shell [19,20],

$$\tilde{f}(x, p_N^2, Q^2) = f(x, Q^2) \left(1 + \frac{p_N^2 - M^2}{M^2} \delta f(x, Q^2) \right), \quad (8)$$

and the off-shell correction function δf can be parametrized and fitted to data. In this work, we adopt the flavor-independent CJ15 parametrization

$$\delta f(x) = \mathcal{N}(x - x_0)(x - x_1)(1 + x_0 - x) \quad (9)$$

inspired by earlier work by Kulagin and Petti on off-shell PDF deformations in heavier nuclei [21]. The x_0 crossing and \mathcal{N} normalization parameters are simultaneously fitted with the PDF and HT parameters, and x_1 is determined by requiring that the off-shell correction does not modify the number of valence quarks in the nucleon,

$$\int_0^1 dx \delta f(x) [q(x) - \bar{q}(x)] = 0 \quad (10)$$

with $q = u, d$, see [3] for details. More flexible parametrizations have been studied [22] and will be reported elsewhere. Finally, for ease of discussion, we collect the off-shell parameters into the vector

$$\mathbf{a}_{\text{off}} = (\mathcal{N}, x_0, x_1). \quad (11)$$

C. Treatment of uncertainties

The full set of fit parameters, including the PDF parameters discussed in Sec. II A, the higher-twist parameters and the off-shell parameters reads

$$\mathbf{a} = (\mathbf{a}_{\text{PDF}}, \mathbf{a}_{\text{HT}}, \mathbf{a}_{\text{off}}) \quad (12)$$

for a total number n_{par} of parameters. The observables σ we are interested in (for example the PDFs themselves, or the DIS structure functions, or the lepton pair production cross section) depend on the fitting parameters via the PDFs f , the HT function C , and the off-shell function δf . Schematically,

$$\sigma[\mathbf{a}] = \sigma(f[\mathbf{a}_{\text{PDF}}], C[\mathbf{a}_{\text{HT}}], \delta f[\mathbf{a}_{\text{off}}]). \quad (13)$$

The uncertainty on these observables can be estimated in the Hessian formalism [23,24]. With a sufficiently precise dataset $m = \{m_1, \dots, m_{n_{\text{dat}}}\}$ and a suitably defined $\chi^2 = \chi^2(\mathbf{a}, \mathbf{m})$ chi-squared function, this method can approximate the parameter likelihood $\mathcal{L}(\mathbf{a}|\mathbf{m}) = \exp(-\frac{1}{2}\chi^2(\mathbf{a}, \mathbf{m}))$ as a multivariate Gaussian distribution in parameter space centered around the best-fit value, \mathbf{a}_0 , of the parameters [25]. Namely,

$$\mathcal{L}(\mathbf{a}|\mathbf{m}) \propto \exp\left(-\frac{1}{2} \Delta \mathbf{a}^T H \Delta \mathbf{a}\right), \quad (14)$$

where m represent the dataset being fitted, $\Delta \mathbf{a} = \mathbf{a} - \mathbf{a}_0$, and the Hessian matrix elements are given by

$$H_{ij} = \frac{1}{2} \frac{\partial^2 \chi^2(\mathbf{a})}{\partial a^i \partial a^j} \Big|_{\mathbf{a}=\mathbf{a}_0}, \quad i, j = 1, \dots, n_{\text{par}}. \quad (15)$$

The Hessian matrix can then be diagonalized, and reparametrized in terms of the eigendirections of the Hessian matrix via

$$\mathbf{a}(\mathbf{t}) = \mathbf{a}_0 + \sum_{k=1}^{n_{\text{par}}} t_k \frac{\mathbf{e}_k}{\sqrt{w_k}}, \quad (16)$$

where \mathbf{e}_k and w_k are the orthonormal eigenvectors and eigenvalues of the Hessian matrix, respectively, and $\mathbf{t} = \{t_1, \dots, t_{n_{\text{par}}}\}$ is a vector of scaling factors. In terms of these variables, the approximated likelihood (14) is a symmetric Gaussian with $t^2 = 1$ identifying the 68% confidence level on the fitted parameters.

The standard CJ PDF error sets are then obtained by uniformly scaling each eigenvector by a “tolerance factor” $t_k = T$ to nominally produce an increase of T^2 above the minimum in the χ^2 function. In both the CJ15 and the CJ15-a analyses $T = 1.645$ was chosen, corresponding to a 90% Gaussian confidence level. In other analyses different choices are made to also account for tensions between the chosen datasets: for example, $T = 10$ in the CT10 global fit [26].

However, in global QCD fits including CJ15, a few Hessian eigenvectors are typically not constrained enough by the available data and the likelihood can deviate from a

Gaussian shape even within t_k variations of order $O(1)$. This can happen, in particular: when data are scarce for a particular flavor combination, such as for \bar{d}/\bar{u} at $x \gtrsim 0.3$; or closer to a kinematic threshold, such as for the d/u ratio as $x \rightarrow 1$, where one expects the ratio to decrease toward 0, but not necessarily reaching that value [27].

A better approximation to the likelihood function [25,28] can be obtained by scanning the χ^2 function along each eigenvector starting from the best-fit parameters \mathbf{a}_0 , until parameters \mathbf{a}_i are found in the plus- and minus-directions such that the χ^2 function increases above its best-fit value by an amount T^2 :

$$\Delta\chi^2(\mathbf{a}_{2i+1}) = \Delta\chi^2(\mathbf{a}_{2i}) = T^2 \quad \forall i = 1, \dots, n_{\text{par}}, \quad (17)$$

where $\Delta\chi^2(\mathbf{a}) = \chi^2(\mathbf{a}) - \chi^2(\mathbf{a}_0)$. These parameter vectors correspond to a set of t_k^\pm values that are close to T wherever the Gaussian approximation holds, but can substantially deviate from this value along a few eigendirections. In other words, we adopt a local and asymmetric tolerance criterion instead of assuming $t_k = T$ globally. In practice, this scheme deforms the Hessian approximation of the likelihood in order to account in an approximate way for departures from a purely Gaussian behavior. It is also suitable for large T values, for which the Hessian approximation cannot be *a priori* assumed to hold. As with other global QCD analyses using local tolerance criteria [29,30], the price to be paid is that, while the chosen T value legitimately defines a confidence region in parameter space, this cannot be readily and unambiguously associated with a confidence level figure as one can do with the pure Hessian approximation.

In practical terms, we define parameter sets corresponding to variations along each eigendirection in the plus and minus directions, respectively, as

$$\mathbf{a}_{2k} = \mathbf{a}_0 + t_k^+ \frac{\mathbf{e}_k}{\sqrt{w_k}} \quad (18)$$

$$\mathbf{a}_{2k+1} = \mathbf{a}_0 - t_k^- \frac{\mathbf{e}_k}{\sqrt{w_k}}, \quad (19)$$

such that $\Delta\chi^2[\mathbf{a}_i] = T$ for all $i = 1, \dots, 2n_{\text{par}}$. Then, the upper and lower $\delta\sigma_+$ and $\delta\sigma_-$ uncertainties on an observable σ can be calculated using the expressions

$$\delta\sigma_+^2 = \sum_{i=1}^{n_{\text{par}}} [\max(\sigma[\mathbf{a}_{2i-1}] - \sigma[\mathbf{a}_0], \sigma[\mathbf{a}_{2i}] - \sigma[\mathbf{a}_0], 0)]^2 \quad (20a)$$

$$\delta\sigma_-^2 = \sum_{i=1}^{n_{\text{par}}} [\max(\sigma[\mathbf{a}_0] - \sigma[\mathbf{a}_{2i-1}], \sigma[\mathbf{a}_0] - \sigma[\mathbf{a}_{2i}], 0)]^2. \quad (20b)$$

Alternatively, a symmetrized uncertainty can be obtained via

$$\delta\sigma^2 = \frac{1}{4} \sum_{i=1}^{n_{\text{par}}} (\sigma[\mathbf{a}_{2i-1}] - \sigma[\mathbf{a}_{2i}])^2. \quad (21)$$

Note that the choice of tolerance T value ($T = 1.645$ in this paper) is already incorporated in Eqs. (20a)–(21). The effect of alternative T' tolerance choices can be approximately obtained by rescaling these uncertainties by a T'/T factor as long as the two tolerance values are not too different from each other. However, care must be exercised for observables sensitive to the non-Gaussian regions of the parameter space.

III. DATASET

The CJ15 analysis included DIS data from fixed target electron-hadron scattering experiments at Jefferson Lab [31,32], HERMES [33], SLAC [34], BCDMS [35,36], and NMC [37,38], and from the HERA ep collider [39]; W [40–44] and Z [45,46] asymmetries as well as jet [47–49] and $\gamma + \text{jet}$ [50] data from CDF and D0 experiments at Tevatron; lepton pair production (LPP) from the E866 experiment at Fermilab [4].

Previously, the antiquark PDFs in the mid- x region were mainly constrained by the lepton pair production data from the E866 experiment. In the CJ22 fits we have now included recent data that are sensitive to the antiquarks from the new lepton pair production measurements by the E906/SeaQuest experiment [5] and the rapidity distribution of the W^+/W^- ratio in pp collisions by the STAR experiment [6]. The datasets used in the CJ22 fit are listed in Table I.

The SeaQuest data covers the kinematic range $0.1 < x < 0.45$ extending the large x reach from the E866 experiment at different Q^2 . The cross section ratio of the lepton pair production in pp and pd interactions attracts particular interest as it can be directly related to antiquark \bar{d}/\bar{u} and quark d/u ratios. This can be appreciated by writing the cross section ratio at leading order in collinear factorization; in particular, in the forward region, this reads

$$\frac{\sigma_{pd}}{\sigma_{pp}} \approx \frac{4 + \frac{d(x_b)}{u(x_b)}}{4 + \frac{d(x_b)}{u(x_b)} \frac{\bar{d}(x_t)}{\bar{u}(x_t)}} \left(1 + \frac{\bar{d}(x_t)}{\bar{u}(x_t)} \right). \quad (22)$$

When $x_b \rightarrow 1$, the d/u ratio tends to 0 and can be neglected, so that the cross section ratio becomes sensitive only to the \bar{d}/\bar{u} ratio. However, neither for the E866 experiment ($x_b = 0.3\text{--}0.5$) nor for the SeaQuest experiment ($x_b = 0.5\text{--}0.7$) is this condition satisfied, and the data are sensitive both to the \bar{d}/\bar{u} quark ratio and, subdominantly, to the d/u ratio. While the cross section ratio

TABLE I. Datasets and corresponding number of data points and χ^2 values from the CJ22 analysis.

Observables	Experiment	References	# Points	χ^2
DIS	JLab (p)	[31]	136	161.0
	JLab (d)	[31]	136	119.1
	JLab (n/d)	[32]	191	213.2
	HERMES (p)	[33]	37	29.1
	HERMES (d)	[33]	37	29.5
	SLAC (p)	[34]	564	469.8
	SLAC (d)	[34]	582	412.1
	BCDMS (p)	[35]	351	472.2
	BCDNS (d)	[36]	254	321.8
	NMC (p)	[37]	275	416.5
	NMC (d/p)	[38]	189	199.6
	HERA (NC e^-p)	[39]	159	249.7
	HERA (NC e^+p 1)	[39]	402	598.9
	HERA (NC e^+p 2)	[39]	75	98.8
	HERA (NC e^+p 3)	[39]	259	250.0
	HERA (NC e^+p 4)	[39]	209	229.1
	HERA (CC e^-p)	[39]	42	45.6
HERA (CC e^+p)	[39]	39	52.5	
LPP	E866 (pp)	[4]	121	144.1
	E866 (pd)	[4]	129	157.4
	SeaQuest (d/p)	[5]	6	7.5
W	CDF (e)	[40]	11	12.6
	D0 (e)	[41]	13	28.8
	D0 (μ)	[42]	10	17.5
	CDF (W)	[43]	13	18.0
	D0 (W)	[44]	14	14.5
	STAR (e^+/e^-)	[6]	9	25.3
	(less η_{\max} point)		(8)	(15.4)
Z	CDF	[45]	28	29.2
	D0	[46]	28	16.1
jet	CDF	[47]	72	14.0
	D0	[48,49]	110	14.0
γ + jet	D0 1	[50]	16	8.7
	D0 2	[50]	16	19.3
	D0 3	[50]	12	25.0
	D0 4	[50]	12	12.2
Total			4557	4936.6
<i>Total</i> + norm			4573	4948.6

receives corrections at higher perturbative orders, Eq. (22) clearly indicates an anticorrelation between the quark and antiquark PDF ratios, that will be important to understand, in particular, the behavior of the large- x d/u ratio in the CJ22 fit compared to the CJ15 analysis.

In pp collisions, W bosons are produced from quark-antiquark fusion and therefore provide clean access to quark and antiquark distributions inside the proton at a large momentum scale $Q^2 = M_W^2$. The STAR experiment at RHIC has recently reported the unpolarized W and Z boson cross sections at $\sqrt{s} = 510$ GeV via $W^\pm \rightarrow e^\pm + \nu$ and $Z \rightarrow e^+e^-$ decays, respectively in the pseudorapidity range $-1.0 < \eta < 1.5$. An observable that is particularly sensitive to the $\bar{d}(x)/\bar{u}(x)$ ratio is the W^+/W^- ratio of W boson cross

sections. Looking again at the leading order calculation, one can approximate

$$\frac{\sigma_{W^+}}{\sigma_{W^-}} \approx \frac{u(x_1)\bar{d}(x_2) + \bar{d}(x_1)u(x_2)}{d(x_1)\bar{u}(x_2) + \bar{u}(x_1)d(x_2)} \approx \frac{\bar{d}/\bar{u}}{d/u}, \quad (23)$$

where $x_{1,2} = (M_W/\sqrt{s}) \exp(\pm y_W)$ is the fractional momentum carried by the scattering partons, and y_W the rapidity of the produced boson. At midrapidity, where $x_1 = x_2 \approx 0.16$, the cross section ratio directly accesses both the antiquark and the quark ratios. At larger rapidity, the accessible $x_{1,2}$ range is somewhat limited by the boson decay kinematics as well as by the statistical precision of the data, and the measured lepton asymmetry effectively probes light quarks and antiquarks with fractional momenta x in the $0.05 \lesssim x \lesssim 0.25$ range.

Of the STAR measurements, the W boson charge ratio is the most sensitive to the quark and antiquark ratios, and has been included in the CJ22 analysis. We have not included either of the charged separated W or the Z measurements because these do not provide significant additional constraints on the PDF determination, but we will discuss in the next section how well the new fit describes those data. As already mentioned, for the STAR W and Z cross section calculations we use fast NLO interpolation grids that were created using APPLgrid [13] interfaced with the MCFM event generator [14,15]. The events were generated using the experimental cuts for electron transverse momentum ($p_e > 15$ GeV/ c) and energy ($25 < E_e < 50$ GeV). The STAR $W^\pm \rightarrow e^\pm + \nu$ measurements also require cuts to suppress jet background. To reproduce a similar condition, we excluded events with produced jets, obtaining approximately a 20% reduction in the calculated cross section. For $Z \rightarrow e^+e^-$ events, the generated events are collected within the experimental invariant mass range of $70 \text{ GeV} < M_{e^+e^-} < 110 \text{ GeV}$ for electron pairs. During the fit, the grids so obtained are then convoluted with the PDFs to calculate the needed NLO cross sections.

IV. RESULTS

In this section, the results of the new analysis are presented and compared with the previously published CJ15 results [3].

Figure 1 compares the lepton pair production cross section ratios from E866 and SeaQuest with our calculations before and after including SeaQuest data into the CJ22 fit. The comparison is done separately for the kinematics of each experiment, with E866 using a 800 GeV proton beam and SeaQuest a 120 GeV beam, and each experiment accessing a different range of lepton pair mass M . For the SeaQuest data, we include the spectrometer acceptance matrix provided in Ref. [5], which has a relatively small effect on an already relatively flat observable. The figure also demonstrates the role of the quark / antiquark ratio anticorrelation discussed earlier.

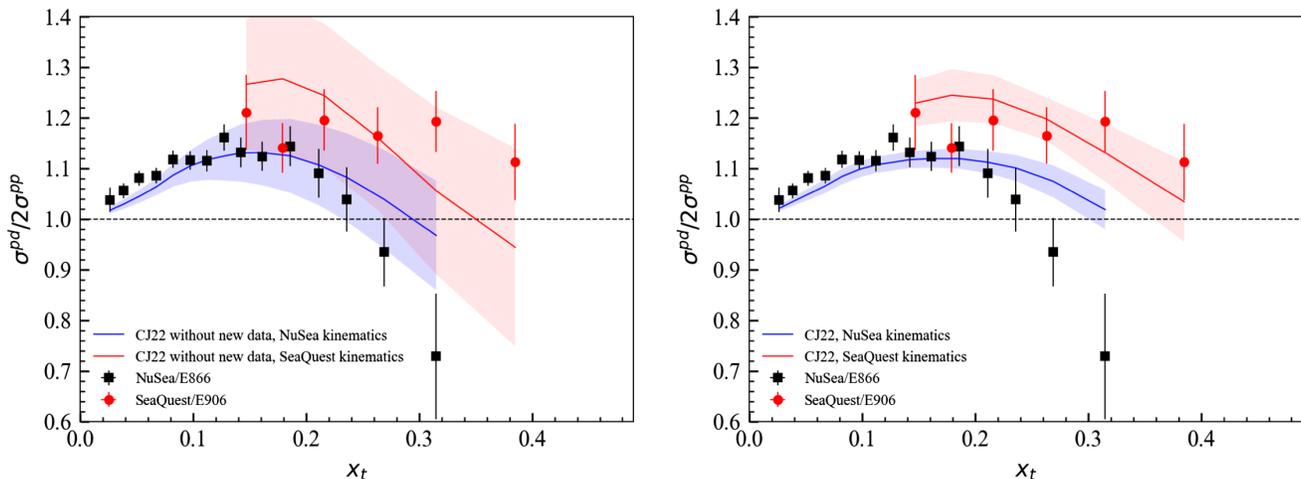


FIG. 1. Comparison of the measured cross section ratio for lepton pair production in pd and pp collisions from the E866 [4,51] and SeaQuest [5] experiments with NLO calculations. The solid blue (red) curve with $T = 1.645$ uncertainty band represents the ratio calculated at the E866 (SeaQuest) kinematics before (left) and after (right) including the new weak boson production data from SeaQuest and STAR in the fit.

Indeed, looking at Eq. (22), if $d(x_b)/u(x_b) \ll 1$ was really negligible in the measured kinematic range, one would expect that the ratio of the proton to deuteron cross section data should be approximately independent of M^2 . Instead, the ratio measured by the experiments is different, which is a direct reflection of the role of the $d(x_b)/u(x_b)$ ratio in Eq. (22). In more detail, at the higher x_b probed by SeaQuest that ratio is smaller than at the lower x_b values of the corresponding E866 measurement. Therefore, one should expect the cross section ratio to be higher for SeaQuest than for E866, which is confirmed in Fig. 1. Furthermore, Eq. (22) shows that an increase in the antiquark $\bar{d}(x_t)/\bar{u}(x_t)$ ratio can be compensated in the PDF fit by a decrease in the $d(x_b)/u(x_b)$ quark ratio and vice versa. The resulting, non-negligible anticorrelation will be important to understand the behavior of the fitted PDF ratios that will be discussed later.

In the left panel of Fig. 1, the calculations from the fit that only includes the E866 data show a steeper downturn in the cross section ratio than allowed by the SeaQuest data. With the new data added to the fit, however, the CJ22 PDFs bring the ratio plotted in the right panel distinctly above 1 in the large- x region where E866 has limited kinematic coverage and statistical precision, with substantially reduced PDF uncertainty. While the new cross section calculation lies higher than the last two E866 data points, only a minor increase in χ^2/datum from 1.63 to 1.93 is observed for the E866 data because of the relatively large uncertainty of the last few data points. Conversely the χ^2/datum value sharply reduces from 3.19 to 1.25 for the SeaQuest data after including the new data, reflecting both the enhanced kinematic range and the precision of the new data.

The CJ22 fit also included the $W^+ \rightarrow e^+/W^- \rightarrow e^-$ cross section ratio measured by the STAR collaboration,

which, as discussed in the previous section, provide complementary information on the \bar{d}/\bar{u} ratio around a smaller $x \approx 0.16$ value, overlapping with the E866 data but at a higher scale. The quality of the fit to the charge ratio data is shown in the upper left panel of Fig. 2. The remaining panels show a comparison of NLO calculations using the new CJ22 PDFs to the unfitted STAR data on the Z , $W^+ \rightarrow e^+$ and $W^- \rightarrow e^-$ rapidity distributions. Differences with CJ15 calculations are minor, and the corresponding curves not shown in the plots.

Overall, CJ22 describes reasonably well the W and Z measurements. However, there is a suggestion of more structure in the $W^- \rightarrow e^-$ channel than shown by the theory. One can also note that the highest rapidity point of the W^+ cross section is much lower than the corresponding theoretical calculation, which has PDFs strongly constrained by the rest of the world dataset and cannot accommodate such a small measurement. Similar features were also observed in other PDF analyses of the STAR data [6,8,9]. In fact, reducing the calculated W^+ cross section in this region would require substantially increasing the $d(x_1)/u(x_1)$ ratio at large values of x_1 and/or decreasing the value of the $\bar{d}(x_2)/\bar{u}(x_2)$ ratio at small values of x_2 . Both possibilities would cause the fits to the lepton pair production data to be worse. Hence, it has not proven possible to get a good description of this one data point.

Figure 3 shows the impact of the new data on the d/u and \bar{d}/\bar{u} ratios at a scale of $Q^2 = 10 \text{ GeV}^2$. The CJ22 results are compared with the CJ15 light quark and antiquark ratios. In the CJ15 analysis, the E866 data provided the strongest constraints for the light-antiquarks and the larger- x region ($x > 0.3$) was essentially left unconstrained by data. As a result, the CJ15 analysis was performed with a more rigid parametrization of the

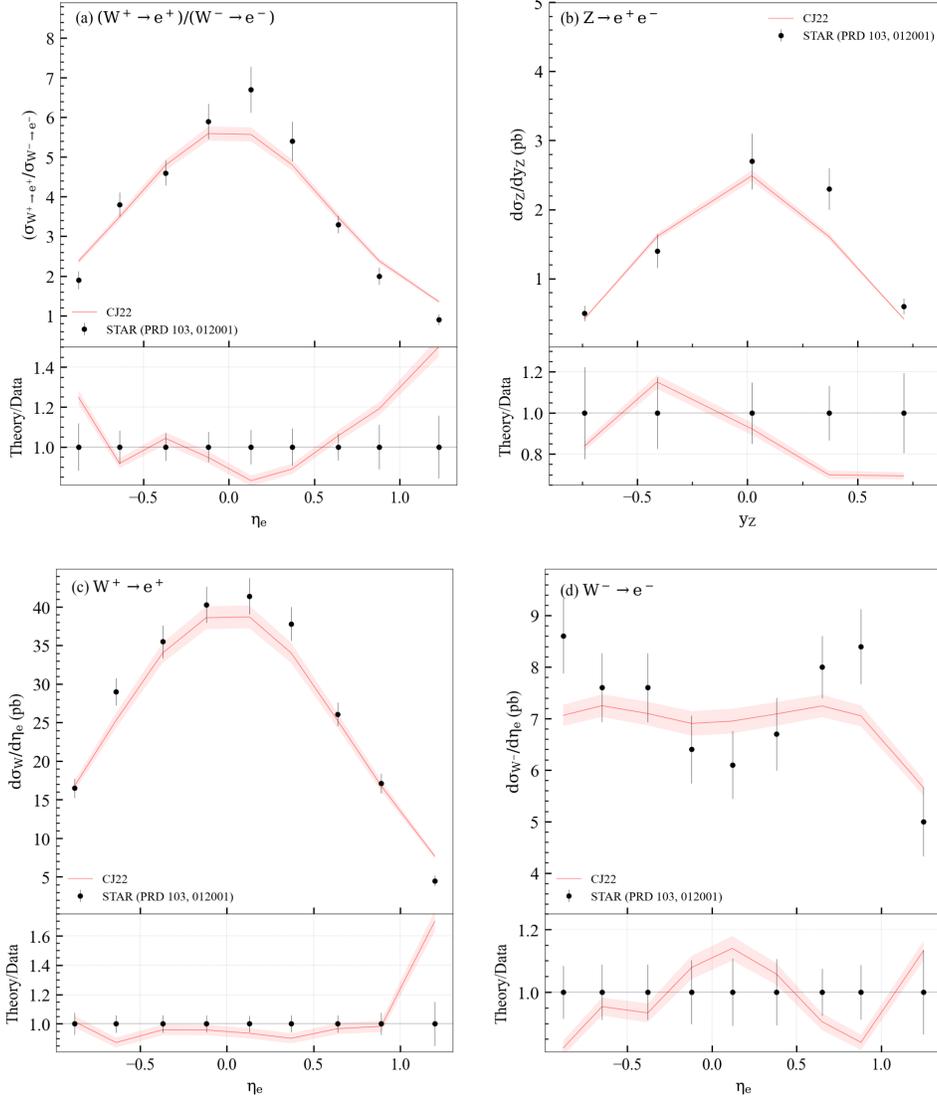


FIG. 2. The measured (a) $\sigma_{W^+}/\sigma_{W^-}$, (b) $d\sigma_Z/dy_Z$, (c) $d\sigma_{W^+}/d\eta_e^+$, and (d) $d\sigma_{W^-}/d\eta_e^-$ are compared with the CJ22 calculations. The statistical and the total systematic uncertainties are added in quadrature and shown as the solid error bars for the data points. The solid red lines show the central values from our fit. The red bands correspond to the $T = 1.645$ PDF uncertainty. The differences with respect to CJ15 calculations are minor and the corresponding curves are omitted for visual clarity.

light-antiquarks, and \bar{d}/\bar{u} ratio was forced to approach 1 as $x \rightarrow 1$. The new data from SeaQuest adds significant constraints on the \bar{d}/\bar{u} with a larger reach in x and allowed us to relax the parametrization used in CJ22, which does not prescribe the large x behavior of the \bar{d}/\bar{u} ratio. Fits where only the parametrization was changed but no new data added were performed as part of the CJ15-a analysis [10] (but are not shown here for visual clarity) and will help disentangling the effects of the change in parametrization and of the new experimental constraints.

The CJ22 fit obtains a \bar{d}/\bar{u} ratio that keeps increasing until $x \approx 0.25$ in the region where the E866 data alone would have required a sharp drop, as evidenced by the CJ15-a fit that left the ratio free to drop below 1 as rapidly as needed (see the green curve in Fig. 5 of Ref. [10]).

At $x \gtrsim 0.25$, the CJ22 antiquark ratio also naturally starts decreasing (in order to accommodate both the E866 and the new SeaQuest data, which by itself would be also consistent with a flatter result) but remains above 1 within uncertainties because of the pull exerted by the new SeaQuest data. This is also the case for the \bar{d}/\bar{u} ratio obtained by the CT [8] and JAM [9] collaborations after inclusion of the SeaQuest data in their global fit. At $x \lesssim 0.2$ the antiquark ratio is driven by the STAR data slightly below the CJ15 result, but remains compatible with the latter.

Turning to the light quark ratio displayed in the left panel of Fig. 3, one can see that the CJ15 d/u ratio remains decidedly above 0 at large x , with a central value extrapolated to $x = 1$ of 0.09 ± 0.03 . On the contrary, in the CJ22

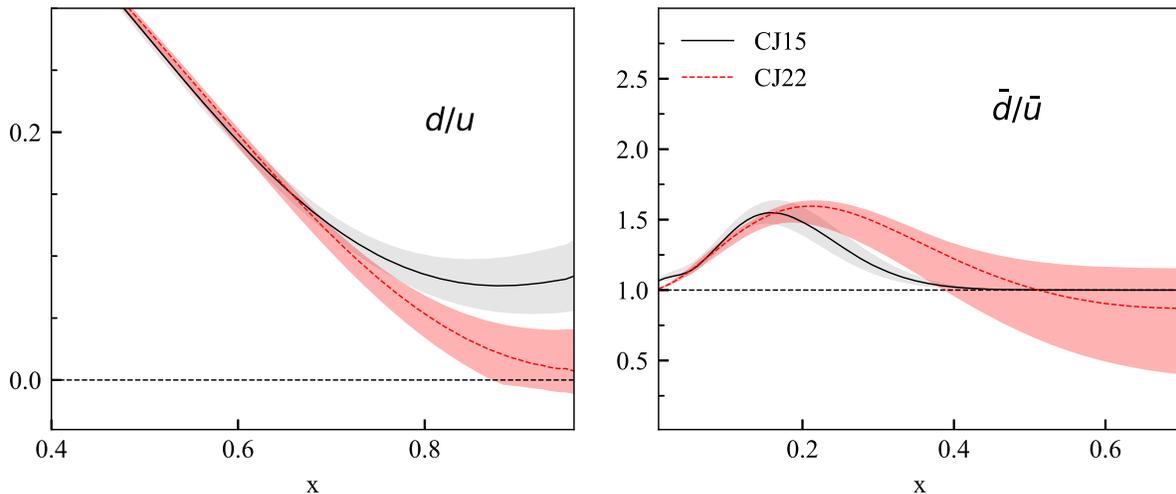


FIG. 3. Comparison of the d/u (left) and \bar{d}/\bar{u} (right) at $Q^2 = 10 \text{ GeV}^2$ between CJ15 (black solid curve) and CJ22 (red dashed curve) with 90% CL uncertainty bands.

analysis, the ratio approaches 0 within uncertainties as $x \rightarrow 1$. This is due to the anticorrelation between the \bar{d}/\bar{u} and d/u induced by the lepton pair production data, and evidenced in Eq. (22), that was discussed earlier: the increase of \bar{d}/\bar{u} in the medium x region, which is allowed by the more flexible CJ15-a and CJ22 parametrizations, and further driven by the increased kinematic reach of the SeaQuest data in the latter, has caused a decrease in d/u in the large x region. This is purely an effect of the more flexible parametrization adopted here, as confirmed by the d/u ratio obtained in the CJ15-a analysis, which is only very marginally smaller than in the CJ22 fit and is not displayed in the figure for visual clarity. The CJ15 non-zero d/u limit was thus the result of a parametrization bias that also underestimated the nominal uncertainty band. With this bias removed, the new result is compatible with the recent d/u fits performed by Alëkhin, Kulagin and Petti [52,53], that have a similar large- x theoretical setup and data coverage (except for the SeaQuest and STAR data).

V. SUMMARY

We have presented the results of the recent CJ122 global QCD analysis of parton distributions which included new electroweak data from SeaQuest and STAR. The SeaQuest data, in particular, extends the x coverage to

larger $x < 0.45$ compared to the previous measurement by E866, leading to significant constraints on the \bar{d}/\bar{u} ratio, allowing the use of the more flexible light-antiquark parametrization discussed in the text. In the CJ22 fit the \bar{d}/\bar{u} ratio remains near 1 as $x \rightarrow 1$ without having to build that into its parametrization, as was previously done in CJ15. The data are also sensitive to the d/u light quark ratio, and the interplay between this and \bar{d}/\bar{u} leads to a d/u ratio that lies below that found in CJ15 as $x \rightarrow 1$, and is compatible with 0 in that limit.

ACKNOWLEDGMENTS

We would like to thank J. Bane, S. Fazio, M. Posik, and A. Tadepalli for informative discussions on their experimental measurements, as well as C. Cocuzza and W. Melnitchouk for useful comments and criticism. This work was supported in part by the U.S. Department of Energy (DOE) Contract No. DE-AC05-06OR23177, under which Jefferson Science Associates LLC manages and operates Jefferson Lab. A. A. also acknowledges support from DOE Contract No. DE-SC0008791. X. Jing was partially supported by DOE Grant No. DE-SC0010129. S. P. acknowledges support from DOE contract No. DE-FG02-05ER41372 and the Center for Frontiers in Nuclear Science.

- [1] A. Accardi *et al.*, A critical appraisal and evaluation of modern PDFs, *Eur. Phys. J. C* **76**, 471 (2016).
- [2] J. F. Owens, A. Accardi, and W. Melnitchouk, Global parton distributions with nuclear and finite- Q^2 corrections, *Phys. Rev. D* **87**, 094012 (2013).
- [3] A. Accardi, L. T. Brady, W. Melnitchouk, J. F. Owens, and N. Sato, Constraints on large- x parton distributions from new weak boson production and deep-inelastic scattering data, *Phys. Rev. D* **93**, 114017 (2016).
- [4] E. A. Hawker *et al.* (NuSea Collaboration), Measurement of the Light Anti-Quark Flavor Asymmetry in the Nucleon Sea, *Phys. Rev. Lett.* **80**, 3715 (1998).
- [5] J. Dove *et al.* (SeaQuest Collaboration), The asymmetry of antimatter in the proton, *Nature (London)* **590**, 561 (2021).
- [6] Jaroslav Adam *et al.* (STAR Collaboration), Measurements of W and Z/γ^* cross sections and their ratios in $p+p$ collisions at RHIC, *Phys. Rev. D* **103**, 012001 (2021).
- [7] Sanghwa Park, Alberto Accardi, Xiaoxian Jing, and J. F. Owens, CJ15 global PDF analysis with new electroweak data from the STAR and SeaQuest experiments, in *Proceedings of the 28th International Workshop on Deep Inelastic Scattering and Related Subjects* (2021), arXiv:2108.05786.
- [8] Marco Guzzi *et al.*, NNLO constraints on proton PDFs from the SeaQuest and STAR experiments and other developments in the CTEQ-TEA global analysis, *SciPost Phys. Proc.* **8**, 005 (2022).
- [9] C. Cocuzza, W. Melnitchouk, A. Metz, and N. Sato (Jefferson Lab Angular Momentum (JAM) Collaboration), Bayesian Monte Carlo extraction of the sea asymmetry with SeaQuest and STAR data, *Phys. Rev. D* **104**, 074031 (2021).
- [10] A. Accardi, C. E. Keppel, S. Li, W. Melnitchouk, G. Niculescu, I. Niculescu, and J. F. Owens (CTEQ-Jefferson Lab (CJ) Collaboration), On the shape of the $\bar{d}-\bar{u}$ asymmetry, *Phys. Lett. B* **801**, 135143 (2020).
- [11] Howard Georgi and H. David Politzer, Freedom at moderate energies: Masses in color dynamics, *Phys. Rev. D* **14**, 1829 (1976).
- [12] Alvaro De Rujula, Howard Georgi, and H. David Politzer, Demythification of electroproduction, local duality and precocious scaling, *Ann. Phys. (N.Y.)* **103**, 315 (1977).
- [13] Tancredi Carli, Dan Clements, Amanda Cooper-Sarkar, Claire Gwenlan, Gavin P. Salam, Frank Siebert, Pavel Starovoitov, and Mark Sutton, *A posteriori* inclusion of parton density functions in NLO QCD final-state calculations at hadron colliders: The APPLGRID project, *Eur. Phys. J. C* **66**, 503 (2010).
- [14] John M. Campbell and R. Keith Ellis, Update on vector boson pair production at hadron colliders, *Phys. Rev. D* **60**, 113006 (1999).
- [15] John M. Campbell and R. Keith Ellis, Top-quark processes at NLO in production and decay, *J. Phys. G* **42**, 015005 (2015).
- [16] Marc Virchaux and Alain Milsztajn, A measurement of α_s and higher twists from a QCD analysis of high statistics F_2 data on hydrogen and deuterium targets, *Phys. Lett. B* **274**, 221 (1992).
- [17] S. I. Alekhin, Sergey A. Kulagin, and S. Liuti, Isospin dependence of power corrections in deep inelastic scattering, *Phys. Rev. D* **69**, 114009 (2004).
- [18] Johannes Blumlein and Helmut Bottcher, Higher twist contributions to the structure functions $F_2^p(x, Q^2)$ and $F_2^d(x, Q^2)$ at large x and higher orders, *Phys. Lett. B* **662**, 336 (2008).
- [19] Sergey A. Kulagin, G. Piller, and W. Weise, Shadowing, binding and off-shell effects in nuclear deep-inelastic scattering, *Phys. Rev. C* **50**, 1154 (1994).
- [20] Sergey A. Kulagin, W. Melnitchouk, G. Piller, and W. Weise, Spin-dependent nuclear structure functions: General approach with application to the deuteron, *Phys. Rev. C* **52**, 932 (1995).
- [21] Sergey A. Kulagin and R. Petti, Global study of nuclear structure functions, *Nucl. Phys. A* **765**, 126 (2006).
- [22] Shujie Li, Alberto Accardi, and Ishara Fernando, Tagged DIS and offshell dynamics in QCD fits, in *Exploring QCD with Tagged Processes* (Paris, 2021).
- [23] J. Pumplin, D. Stump, R. Brock, D. Casey, J. Huston, J. Kalk, H. L. Lai, and W.-K. Tung, Uncertainties of predictions from parton distribution functions. 2. The Hessian method, *Phys. Rev. D* **65**, 014013 (2001).
- [24] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, Uncertainties of predictions from parton distributions. 1: Experimental errors, *Eur. Phys. J. C* **28**, 455 (2003).
- [25] N. T. Hunt-Smith, A. Accardi, W. Melnitchouk, N. Sato, A. W. Thomas, and M. J. White, Determination of uncertainties in parton densities, *Phys. Rev. D* **106**, 036003 (2022).
- [26] Hung-Liang Lai, Marco Guzzi, Joey Huston, Zhao Li, Pavel M. Nadolsky, Jon Pumplin, and C. P. Yuan, New parton distributions for collider physics, *Phys. Rev. D* **82**, 074024 (2010).
- [27] Roy J. Holt and Craig D. Roberts, Nucleon and pion distribution functions in the valence region, *Rev. Mod. Phys.* **82**, 2991 (2010).
- [28] A. Accardi, T. J. Hobbs, X. Jing, and P. M. Nadolsky, Deuterium scattering experiments in CTEQ global QCD analyses: A comparative investigation, *Eur. Phys. J. C* **81**, 603 (2021).
- [29] Tie-Jiun Hou *et al.*, New CTEQ global analysis of quantum chromodynamics with high-precision data from the LHC, *Phys. Rev. D* **103**, 014013 (2021).
- [30] S. Bailey, T. Cridge, L. A. Harland-Lang, A. D. Martin, and R. S. Thorne, Parton distributions from LHC, HERA, Tevatron and fixed target data: MSHT20 PDFs, *Eur. Phys. J. C* **81**, 341 (2021).
- [31] S. P. Malace *et al.* (Jefferson Lab E00-115 Collaboration), Applications of quark-hadron duality in the F_2 structure function, *Phys. Rev. C* **80**, 035207 (2009).
- [32] S. Tkachenko *et al.* (CLAS Collaboration), Measurement of the structure function of the nearly free neutron using spectator tagging in inelastic $^3\text{H}(e, e'p_s)X$ scattering with CLAS, *Phys. Rev. C* **89**, 045206 (2014); *Phys. Rev. C* **90**, 059901(A) (2014).
- [33] A. Airapetian *et al.*, Inclusive measurements of inelastic electron and positron scattering from unpolarized hydrogen and deuterium targets, *J. High Energy Phys.* **05** (2011) 126.
- [34] L. W. Whitlow, E. M. Riordan, S. Dasu, Stephen Rock, and A. Bodek, Precise measurements of the proton and deuteron structure functions from a global analysis of the SLAC deep

- inelastic electron scattering cross sections, *Phys. Lett. B* **282**, 475 (1992).
- [35] A. C. Benvenuti *et al.* (BCDMS Collaboration), A high statistics measurement of the proton structure functions $F_2(x, Q^2)$ and R from deep inelastic muon scattering at high Q^2 , *Phys. Lett. B* **223**, 485 (1989).
- [36] A. C. Benvenuti *et al.* (BCDMS Collaboration), A high statistics measurement of the deuteron structure functions $F_2(x, Q^2)$ and R from deep inelastic muon scattering at high Q^2 , *Phys. Lett. B* **237**, 592 (1990).
- [37] M. Arneodo *et al.* (New Muon Collaboration), Measurement of the proton and deuteron structure functions, F_2^p and F_2^d , and of the ratio σ_L/σ_T , *Nucl. Phys.* **B483**, 3 (1997).
- [38] M. Arneodo *et al.* (New Muon Collaboration), Accurate measurement of F_2^d/F_2^p and $R^d - R^p$, *Nucl. Phys.* **B487**, 3 (1997).
- [39] H. Abramowicz *et al.* (H1 and ZEUS Collaborations), Combination of measurements of inclusive deep inelastic $e^\pm p$ scattering cross sections and QCD analysis of HERA data, *Eur. Phys. J. C* **75**, 580 (2015).
- [40] D. Acosta *et al.* (CDF Collaboration), Measurement of the forward-backward charge asymmetry from $W \rightarrow e\nu$ production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV, *Phys. Rev. D* **71**, 051104 (2005).
- [41] V. M. Abazov *et al.* (D0 Collaboration), Measurement of the electron charge asymmetry in $p\bar{p} \rightarrow W + X \rightarrow e\nu + X$ decays in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV, *Phys. Rev. D* **91**, 032007 (2015); *Phys. Rev. D* **91**, 079901(E) (2015).
- [42] V. M. Abazov *et al.* (D0 Collaboration), Measurement of the muon charge asymmetry in $p\bar{p} \rightarrow W + X \rightarrow \mu\nu + X$ events at $\sqrt{s} = 1.96$ TeV, *Phys. Rev. D* **88**, 091102 (2013).
- [43] T. Aaltonen *et al.* (CDF Collaboration), Direct Measurement of the W Production Charge Asymmetry in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV, *Phys. Rev. Lett.* **102**, 181801 (2009).
- [44] V. M. Abazov *et al.* (D0 Collaboration), Measurement of the W Boson Production Charge Asymmetry in $p\bar{p} \rightarrow W + X \rightarrow e\nu + X$ Events at $\sqrt{s} = 1.96$ TeV, *Phys. Rev. Lett.* **112**, 151803 (2014); *Phys. Rev. Lett.* **114**, 049901(E) (2015).
- [45] Timo Antero Aaltonen *et al.* (CDF Collaboration), Measurement of $d\sigma/dy$ of Drell-Yan e^+e^- pairs in the Z mass region from $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV, *Phys. Lett. B* **692**, 232 (2010).
- [46] V. M. Abazov *et al.* (D0 Collaboration), Measurement of the shape of the boson rapidity distribution for $p\bar{p} \rightarrow Z/\gamma^* \rightarrow e^+e^- + X$ events produced at \sqrt{s} of 1.96 TeV, *Phys. Rev. D* **76**, 012003 (2007).
- [47] T. Aaltonen *et al.* (CDF Collaboration), Measurement of the inclusive jet cross section at the Fermilab Tevatron $p\bar{p}$ collider using a cone-based jet algorithm, *Phys. Rev. D* **78**, 052006 (2008); *Phys. Rev. D* **79**, 119902(E) (2009).
- [48] B. Abbott *et al.* (D0 Collaboration), Inclusive Jet Production in $p\bar{p}$ Collisions, *Phys. Rev. Lett.* **86**, 1707 (2001).
- [49] V. M. Abazov *et al.* (D0 Collaboration), Measurement of the Inclusive Jet Cross-Section in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ -TeV, *Phys. Rev. Lett.* **101**, 062001 (2008).
- [50] V. M. Abazov *et al.* (D0 Collaboration), Measurement of the differential cross-section for the production of an isolated photon with associated jet in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV, *Phys. Lett. B* **666**, 435 (2008).
- [51] R. S. Towell *et al.* (NuSea Collaboration), Improved measurement of the \bar{d}/\bar{u} asymmetry in the nucleon sea, *Phys. Rev. D* **64**, 052002 (2001).
- [52] S. I. Alekhin, S. A. Kulagin, and R. Petti, Nuclear effects in the deuteron and constraints on the d/u ratio, *Phys. Rev. D* **96**, 054005 (2017).
- [53] S. I. Alekhin, S. A. Kulagin, and R. Petti, Nuclear effects in the deuteron and global QCD analyses, *Phys. Rev. D* **105**, 114037 (2022).