Nuclear effects in the deuteron and constraints on the d/u ratio

S. I. Alekhin, S. A. Kulagin, and R. Petti

Nuclear Effects in the Deuteron and Constraints on the d/u Ratio

S. I. Alekhin[*]

II. Institut für Theoretische Physik, Universität Hamburg
Luruper Chaussee 149, D–22761 Hamburg, Germany and

Institute of High Energy Physics, 142281 Protvino, Moscow Region, Russia

S. A. Kulagin[†]

Institute for Nuclear Research of the Russian Academy of Sciences,
60-letiya Oktyabrya prospekt 7a, 117312 Moscow, Russia

R. Petti[‡]

Department of Physics and Astronomy,
University of South Carolina, Columbia SC 29208, USA

Abstract

We present a detailed study of nuclear corrections in the deuteron (D) by performing an analysis of data from deep inelastic scattering off proton and D, dilepton pair production in pp and pD interactions, and $W^{\pm}$ and Z boson production in pp and $p\bar{p}$ collisions. In particular, we discuss the determination of the off-shell function describing the modification of the parton distribution functions in bound nucleons in the context of global QCD fits. Our results are consistent with the ones obtained independently from the study of data on deep inelastic scattering off heavy nuclei with mass number $A \geq 4$, further confirming the universality of the off-shell function of the bound nucleon. We also study the sensitivity to the modeling of the deuteron wave function. As an important application we discuss the impact of nuclear corrections to the deuteron on the determination of the d quark distribution.

PACS numbers: 13.60.Hb, 12.38.Qk

---
[*] sergey.alekhin@ihep.ru
[†] kulagin@ms2.inr.ac.ru
[‡] Roberto.Petti@cern.ch

I. INTRODUCTION

Parton distribution functions (PDFs) are universal process-independent characteristics of the target, which determine leading contributions to the cross sections of various hard processes involving leptons and hadrons [1]. The PDF content of both the proton and the neutron is extracted from global fits [2–5] to experimental data at large momentum transfer, including lepton deep inelastic scattering (DIS), lepton-pair production (Drell-Yan process), jet production, and W and Z boson production in hadron collisions. In order to disentangle the content of different parton flavors, global fits must include complementary data which are flavor dependent. Traditionally, the most efficient separation between d and u quark distributions is obtained by comparing charged-lepton DIS data for proton and deuterium targets, the latter being considered as an "effective" neutron target. Since the deuteron is a weakly bound nucleus with a binding energy of about $2.2\,\mathrm{MeV}$ – accounting only for about 0.1% of its mass – it is often assumed to be well approximated by the sum of a quasi-free proton and a quasi-free neutron in the PDF analyses.

However, charged-lepton DIS data from various nuclear targets demonstrate significant nuclear effects with a rate that is more than one order of magnitude larger than the ratio of the nuclear binding energy to the nucleon mass (for a review see Refs. [6, 7]). These observations indicate that the nuclear environment plays an important role even at energies and momenta much higher than those involved in typical nuclear ground state processes [6–8]. Similar considerations can be drawn for the DIS off the deuteron [9–29]. In spite of the broad range of predictions, such studies indicate that nuclear effects in the deuteron are non negligible and rise rapidly in the region of large Bjorken x. A recent direct measurement of nuclear effects in the deuteron [30] indicates a few-percent negative correction at $x \sim 0.5 - 0.6$, with a steep rise at large x. Therefore, if neglected or treated incorrectly, these nuclear effects can potentially introduce significant uncertainties and/or biases in the extraction of the neutron structure functions and of the d quark distribution from the DIS data [31].

A microscopic model for nuclear structure functions and PDFs accounting for a number of different nuclear effects was developed in Refs.[25, 32–34]. It includes the smearing with the energy-momentum distribution of bound nucleons (Fermi motion and binding, FMB), the off-shell correction (OS) to bound nucleon structure functions, the contributions from meson exchange currents and the propagation of the hadronic component of the virtual intermediate boson in the nuclear environment. This model has been successfully used to quantitatively explain the observed x, Q^2 and A dependence of the nuclear DIS data in a wide range of targets from $^3\mathrm{He}$ to $^{207}\mathrm{Pb}$ [25, 32, 33], the magnitude, the x and mass dependence of the nuclear Drell-Yan (DY) data [34], as well as the data on the differential cross sections and asymmetries for $W^{\pm}, Z$ production in $p + \mathrm{Pb}$ collisions at the LHC [35].

A consistent description of the scattering off bound nucleons not only involves the smearing due to the nuclear momentum distribution, but also requires the knowledge of the off-shell (OS) scattering amplitudes. The model of Ref. [25] exploits the observation that the nucleus is a weakly bound system and thus it is sufficient to evaluate the OS correction to the bound nucleon PDFs in the vicinity of the mass shell. The shape of this correction is defined by a universal function $\delta f(x)$ of the Bjorken variable x, while its nuclear dependence is driven by the average virtuality of the nucleon (off-shelness) inside the nucleus. The OS function δf can be regarded as a special nucleon structure function, which does not contribute to the cross section of the physical nucleon, but is relevant only for the bound nucleon and describes its response to the interaction with the nuclear environment. The off-shell cor-

rection proved to be an important contribution to explain nuclear effects at large x. The function δf was determined from the analysis of data on ratios of DIS structure functions in different nuclei [25]. It was also shown that in a simple single-scale model, in which the quark momentum distributions in the nucleon are functions of the nucleon core radius, the observed behaviour of δf can be interpreted in terms of an increase of the confinement radius in the bound nucleon in the nuclear environment [25].

The deuteron is a weakly bound state of two nucleons with peculiar attributes. Its dynamics is better understood than the dynamics of many-particle nuclei, making it an ideal benchmark tool for the study of different nuclear effects. However, it is also considerably different with respect to even a three-body nucleus like ^{3}He. For this reasons one can not rely on simple extrapolations of nuclear effects from heavy targets based upon nuclear density or atomic weight, as it is often assumed in phenomenological analyses. In contrast, the model of Ref. [25] suggests a unified treatment of the deuteron and heavier nuclei on the basis of common underlying physics mechanisms.

In this paper we discuss an independent determination of the off-shell function δf, together with the proton PDFs, from a global QCD analysis of proton and deuterium data. In Sec.II we review the model of nuclear corrections in the deuteron, while in Sec. III we discuss the details of the data analysis. In Sec. IV we compare our results with the one obtained from heavy nuclear targets and discuss the impact on the uncertainties related to the d/u ratio from global QCD fits. We summarize our results in Sec. V.

II. MODEL OF NUCLEAR CORRECTIONS

The nuclear corrections to the inelastic structure functions involve a number of different contributions. For the deuteron we can write (for simplicity we summarize the structure function F_2 here) [25, 34]:

$$F_2^D = F_2^{N/D} + \delta_{\mathrm{MEC}} F_2^D + \delta_{\mathrm{coh}} F_2^D, \tag{1}$$

where the first term in the right-hand side stands for the incoherent scattering off the bound isoscalar nucleon N including the off-shell correction, and $\delta_{\mathrm{MEC}} F_2^D$ and $\delta_{\mathrm{coh}} F_2^D$ are the corrections due to nuclear meson exchange currents (MEC) and coherent interactions of the intermediate virtual boson with nuclear target, respectively.

A. Incoherent Scattering off Bound Nucleons

The first term in Eq.(1) dominates at $x > 0.2$ and can be written as follows [25]:

$$\gamma^2 F_2^{N/D}(x, Q^2) = \int \frac{\mathrm{d}^3 \boldsymbol{p}}{(2\pi)^3} \, |\Psi_D(\boldsymbol{p})|^2 \left(1 + \frac{\gamma p_z}{M}\right) \left(\gamma'^2 + \frac{6x'^2 \boldsymbol{p}_\perp^2}{Q^2}\right) F_2^N(x', Q^2, p^2), \tag{2}$$

where the integration is over the momentum of the bound nucleon $\boldsymbol{p}$, $\Psi_D(\boldsymbol{p})$ is the deuteron wave function, $M = \frac{1}{2}(M_p + M_n)$ and $F_2^N = \frac{1}{2}(F_2^p + F_2^n)$ are respectively the mass and the structure function of the bound nucleon with four-momentum $p = (M + \varepsilon, \boldsymbol{p})$, where $\varepsilon = \varepsilon_D - \boldsymbol{p}^2/(2M)$ and $\varepsilon_D = M_D - 2M$ is the deuteron binding energy. The integration in Eq.(2) requires the structure function of the bound nucleon in the off-shell region and F_2^N depends on the Bjorken variable $x' = Q^2/(2pq)$, the momentum transfer square Q^2, and also

on the nucleon invariant mass squared p^2. In Eq.(2) we use a coordinate system such that the momentum transfer $\boldsymbol{q}$ is antiparallel to the z axis, $\boldsymbol{p}_\perp$ is the transverse component of the nucleon momentum, and $\gamma^2 = 1 + 4x^2 M^2/Q^2$ and $\gamma'^2 = 1 + 4x'^2 p^2/Q^2$.

The integrand in Eq.(2) factorizes into two independent terms involving the contribution from two different scales: i) the wave function $\Psi_D(\boldsymbol{p})$ describing the deuteron properties in momentum space, and ii) the nucleon structure function F_2^N describing processes at the parton level in the nucleon. In the following we will consider several deuteron wave functions $\Psi_D(\boldsymbol{p})$ corresponding to different models for the nucleon-nucleon potential: Paris [36], CD-Bonn [37], AV18 [38], WJC1 and WJC2 [39, 40]. These wave functions are constrained by high-precision fits to nucleon-nucleon scattering data at low energies. However, these models for $\Psi_D(\boldsymbol{p})$ can differ by more than a factor of 10 in the high momentum tail, as shown in Fig. 1. Table I summarizes the salient kinematic parameters associated with each deuteron wave function. To be consistent with the weak binding approximation of Ref.[25], we perform the integration over the nucleon momentum in Eq.(2) up to $|\boldsymbol{p}| < 1\,\text{GeV/c}$.
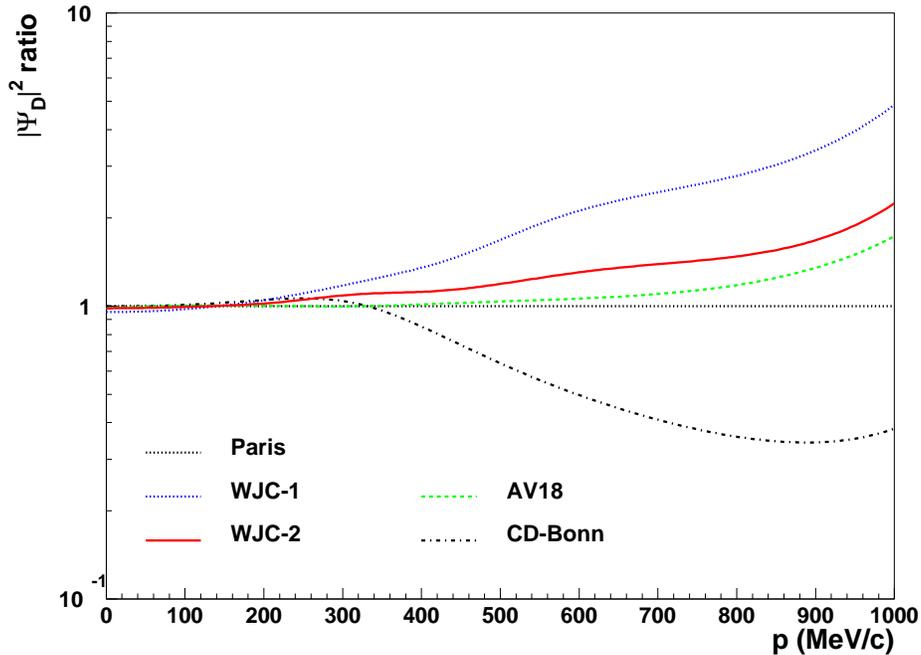

FIG. 1. Ratio of the deuteron momentum distribution $|\Psi_D(\boldsymbol{p})|^2$ calculated according to various models with respect to the corresponding value in the Paris model. See text for details.

The nucleon structure function in Eq.(2) includes the target mass and the higher-twist corrections represented as follows

$$F_2^N(x, Q^2, p^2) = F_2^{\text{TMC}}(x, Q^2, p^2) + H_2^{(4)}(x)/Q^2 + \mathcal{O}(Q^{-4}) \tag{3}$$

where F_2^{TMC} is the leading-twist (LT) structure function corrected for the target mass effects (TMC) and $H_2^{(4)}$ describes the twist-4 contribution (we suppress any explicit notation to higher order terms for brevity). The LT structure function is computed using the proton

4

Wave function	Reference	$\langle v \rangle$	$\langle \varepsilon \rangle$ [MeV]	$\langle \boldsymbol{p}^2 \rangle / 2M$ [MeV]
WJC1	[39, 40]	-0.062	-16.16	13.94
WJC2	[39, 40]	-0.049	-12.89	10.67
AV18	[38]	-0.045	-11.93	9.71
Paris	[36]	-0.043	-11.55	9.33
CD-Bonn	[37]	-0.037	-9.96	7.74

TABLE I. Values of the average nucleon virtuality $v = (p^2 - M^2)/M^2$, bound nucleon energy ε, and kinetic energy $\boldsymbol{p}^2/2M$ for each deuteron wave function shown in Fig. 1.

and the neutron PDFs extracted from the analysis of data as described in Sec.III. The target mass correction is computed following the prescription of Ref. [41]:

$$F_2^{\mathrm{TMC}}(x, Q^2, p^2) = \frac{x^2}{\xi^2 \gamma^2} F_2^{\mathrm{LT}}(\xi, Q^2, p^2) + \frac{6 x^3 p^2}{Q^2 \gamma^4} \int_\xi^1 \frac{dz}{z^2} F_2^{\mathrm{LT}}(z, Q^2, p^2) + \mathcal{O}(Q^{-4}) \qquad (4)$$

where $\xi = 2x/(1 + \gamma)$ is the Nachtmann variable, we substitute p^2 for the mass of the bound nucleon squared M^2, and $\gamma^2 = 1 + 4x^2 p^2/Q^2$. Note that the second term in Eq.(4) is suppressed as $1/Q^2$ and therefore can be formally considered as a kinematic HT contribution. Recent phenomenology suggests that Eq.(3) with twist-4 contributions provides a good description of data down to $Q \sim 1\,\mathrm{GeV}$ [42]. It is also worth noting that this model is consistent with duality principle and on average describes the resonance data with $W < 1.8\,\mathrm{GeV}$ [42, 43].

B. Off-shell Correction

The structure function of the bound nucleon $F_2^N(x, Q^2, p^2)$ appearing in the calculation of the nuclear correction in Eq.(2) explicitly depends on the nucleon invariant mass squared p^2. The p^2-dependence of the structure function has two different sources [25, 44]: (i) the dynamic off-shell dependence of the LT structure function; (ii) the kinematic target mass correction, which generates terms of the order of p^2/Q^2. We evaluate the off-shell dependence of the target mass correction by replacing $M^2 \to p^2$ in Eq.(4). Since the characteristic momenta of a bound nucleon are small compared to its mass, the integration in Eq.(2) mainly covers a region in the vicinity of the mass shell. The nucleon virtuality $v = (p^2 - M^2)/M^2$ can then be treated as a small parameter, so that we can expand the structure function in series of v keeping only the leading term:

$$F_2^{\mathrm{LT}}(x, Q^2, p^2) = F_2^{\mathrm{LT}}(x, Q^2) \left[1 + \delta f(x, Q^2)\, v \right], \qquad (5)$$

$$\delta f = \partial \ln F_2^{\mathrm{LT}} / \partial \ln p^2 \big|_{p^2 = M^2}, \qquad (6)$$

where the first term on the r.h.s. in Eq.(5) is the structure function of the on-mass-shell nucleon. The off-shell (OS) function δf can be regarded as a special nucleon structure function, which describes the relative modification of the nucleon structure functions and PDFs in the vicinity of the mass shell. This function does not contribute to the cross section of the physical nucleon, but it is relevant only for the bound nucleon and describes its response to the interaction in a nucleus.

5

In general, the function δf might be flavor dependent and different for protons and neutrons. However, the study of nuclear DIS and DY data [25, 33, 34] supports the hypothesis of the OS function universality, with no significant Q^2 dependence suggested by the data, i.e. $\delta f(x, Q^2) = \delta f(x)$. Although we assume that δf is only function of x, the overall OS correction to the nuclear structure functions also depends on Q^2, as a result of the integration of Eq.(2). It is important to note that this Q^2 dependence is different from the ones of both the LT and HT contributions to the structure functions in Eq.(3). This difference allows a simultaneous extraction of PDFs, HTs and δf from global QCD fits (c.f. Sec. III).
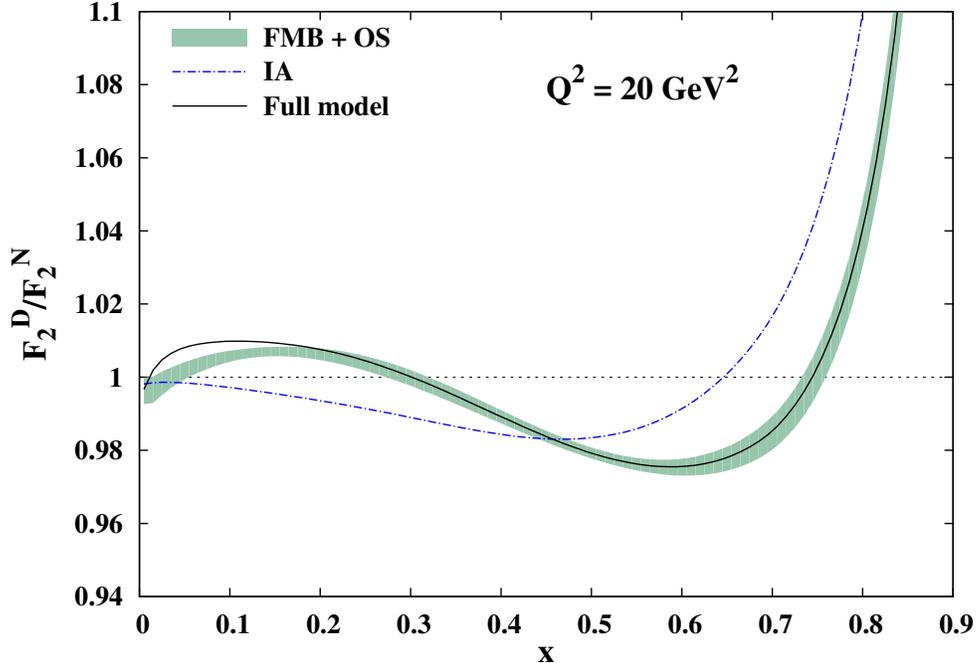

FIG. 2. Ratio of the deuteron and the isoscalar nucleon structure functions F_2^D/F_2^N calculated at $Q^2 = 20\,\mathrm{GeV}^2$ using different approximations. The solid line is the full model of Ref.[25], while the dashed line is the result of Ref.[25] with no off-shell ($p^2 = M^2$), nuclear pion and nuclear shadowing correction (impulse approximation). The shaded area represents the $\pm 1\sigma$ uncertainty band on the impulse approximation supplemented by the off-shell correction only.

C. Other Nuclear Corrections

The remaining corrections appearing in Eq.(1), i.e. the nuclear meson exchange current $\delta_{\mathrm{MEC}} F_2^D$ and the nuclear shadowing (NS) $\delta_{\mathrm{coh}} F_2^D$, are relevant only at small x. For the details of the treatment of these terms we refer to Ref.[25]. Here we only emphasize that nuclear effects present in different kinematical regions of x are related by the DIS sum rules and normalization constraints [34]. For example, the light-cone momentum sum rule links the FMB and the MEC corrections. We use this relation to constrain the mesonic contributions to the nuclear structure functions. Similarly, the baryon number sum rule links the shadowing and the OS corrections. In our approach, the OS effect provides the

mechanism to cancel a negative nuclear-shadowing contribution to the normalization of the nuclear valence quark distributions.

D. Model Predictions and Phenomenology

The model described above was used to perform a detailed analysis of data on the ratio $\mathcal{R}(A/B) = F_2^A/F_2^B$ between the nuclear targets A and B [25]. The data sets analyzed stem from a variery of electron and muon DIS experiments (CERN NMC, EMC, and BCDMS, SLAC E139 and E140, Fermilab E665), with targets ranging from ^{4}He to ^{208}Pb in a wide region of x and Q^2. In this way we tested the hypothesis that the OS modification of bound nucleons is responsible for the difference between the data and all known nuclear effects in Eq.(1), including the FMB [45, 46], the nuclear shadowing and the nuclear MEC. In turn, this OS correction is controlled by the universal off-shell function δf in Eq.(5), which was determined from DIS data with the corresponding uncertainty [25].

Such an approach leads to an excellent agreement with the available DIS data on the x, Q^2, and A dependence of $\mathcal{R}(A/B)$. The model predictions are also in good agreement [33] with the recent DIS data by the HERMES experiment at HERA [47] and the E03-103 experiment at JLab [48] down to ^{3}He. Furthermore, the same model allows a calculation of nuclear PDFs, which can describe well the magnitude, the x and mass dependence of the available data on Drell-Yan production off various nuclear targets [34], as well as the differential cross sections and asymmetries for $W^\pm, Z$ production in $p + $Pb collisions at the LHC [35].

In this paper we perform an independent analysis of deuterium and proton data in the context of global QCD fit. In Fig.2 we show the predictions of Ref. [25] for the ratio $\mathcal{R}(D/N) = F_2^D/F_2^N$ of the deuteron and the isoscalar nucleon structure functions at $Q^2 = 20$ GeV^2. The region of $0.35 \leq x \leq 0.55$ is characterized by an almost linear dependence from x, with a slope $\mathrm{d}\mathcal{R}(D/N)/\mathrm{d}x = -0.099 \pm 0.006$, including model uncertainties. This slope is often used in the analysis of experimental data [48] since it is less affected by the experimental uncertainties (especially the overall normalization) than the absolute value of the nuclear correction. At large $x > 0.1$ nuclear corrections are dominated by the FMB and OS effects, as shown in Fig. 2. In particular, the off-shell correction is a crucial contribution in this kinematic region, which is studied in more details in the present analysis.

III. OFF-SHELL CORRECTION FROM GLOBAL QCD FIT

In this paper we discuss the impact of nuclear effects in the deuteron in the context of global QCD fits. Our goals are twofold: i) an independent determination of the off-shell correction preferred by the deuteron data; ii) an estimate of the PDF uncertainties (in particular for the d/u ratio at large x) introduced by the nuclear corrections to deuterium data. The analysis framework and the main data sets used are common to the ABMP16 fit [2].

A. Analysis Framework

In our analysis we use the Next-to-Next-to-Leading-Order (NNLO) approximation in the QCD perturbation theory to calculate the partonic cross sections entering the LT terms

	Experiment	Reference	Beam	Target(s)	Final states	Data points
DIS collider	HERA I+II	[49]	e	p	eX	1168
					νX	
	HERA I+II	[50]	e	p	ecX	52
	H1	[51]	e	p	ebX	12
	ZEUS	[52]	e	p	ebX	17
DIS fixed target	BCDMS	[53, 54]	μ	p, D	μX	605
	NMC	[55]	μ	p, D	μX	490
	SLAC E49a	[56]	e	p, D	eX	118
	SLAC E49b	[56]	e	p, D	eX	299
	SLAC E87	[56]	e	p, D	eX	218
	SLAC E89b	[57]	e	p, D	eX	162
	SLAC E139	[19]	e	D	eX	17
	SLAC E140	[58]	e	D	eX	26
	JLab BONuS	[30]	e	D	eX	5
	NOMAD	[59]	ν	Fe	$\mu^+\mu^- X$	48
	CHORUS	[60]	ν	$Emul.$	$\mu c X$	6
	CCFR	[61]	ν	Fe	$\mu^+\mu^- X$	89
	NuTeV	[61]	ν	Fe	$\mu^+\mu^- X$	89
Drell-Yan fixed target	FNAL E866	[62]	p	p, D	$\mu^+\mu^-$	39
	FNAL E605	[63]	p	Cu	$\mu^+\mu^-$	119
W, Z collider	D0	[64]	p	$\bar{p}$	$W^+ \to \mu^+\nu$	10
					$W^- \to \mu^-\nu$	
	D0	[65]	p	$\bar{p}$	$W^+ \to e^+\nu$	13
					$W^- \to e^-\nu$	
	ATLAS	[66, 67]	p	p	$W^+ \to l^+\nu$	36
					$W^- \to l^-\nu$	
					$Z \to l^+l^-$	
	CMS	[68, 69]	p	p	$W^+ \to \mu^+\nu$	33
					$W^- \to \mu^-\nu$	
	LHCb	[70, 71]	p	p	$W^+ \to \mu^+\nu$	63
					$W^- \to \mu^-\nu$	
					$Z \to \mu^+\mu^-$	
	LHCb	[72]	p	p	$Z \to e^+e^-$	17

TABLE II. List of the various data sets used in the present analysis.

for the hard interaction processes considered. We set the renormalization and factorization scales to $\mu_r = \mu_f = \mu$ and we identify this scale μ with the relevant kinematics of each process, e.g. $\mu = Q$ for DIS. The individual PDFs are parameterized as in Ref. [2] at the starting scale $\mu^2 = Q_0^2 = 9$ GeV^2. The PDFs are subject to sum rule constraints due to the conservation of the quark number and the momentum in the nucleon.

The splitting functions controlling the scale dependence of the PDFs in the evolution equations are evaluated at NNLO in perturbation theory [73, 74]. The Wilson coefficients entering the massless DIS structure functions are calculated at NNLO [75–81]. Similarly, we use NNLO calculations for the partonic cross sections of the Drell-Yan process and the hadronic W and Z boson production [82–86].

In our PDF analysis we use a fixed flavor number scheme (FFNS) with $n_f = 3$ light flavors from which heavy quark PDFs are generated. The heavy quark masses m_q are defined in the $\overline{\text{MS}}$ renormalization scheme as running masses $m_q(\mu)$ depending upon the scale μ of the hard scattering in analogy to the running coupling $\alpha_s(\mu)$. As discussed in Ref. [87], the use of the $\overline{\text{MS}}$-mass allows better convergence properties and greater perturbative stability at higher orders. The heavy quark Wilson coefficients entering the DIS structure functions for heavy quark production are known exactly only to the Next-to-Leading-Order (NLO) for both the charged current (CC) [88, 89] and neutral current (NC) [90] processes. For the NC case approximate NNLO coefficients are employed [2].

In the kinematic range of our analysis the twist-6 terms give negligible contributions to the structure functions [42]. Therefore, in addition to the leading twist we only include two twist-4 contributions H_2 and H_T – as defined in Eq.(3) – to the structure functions F_2 and F_T, respectively. We also considered the target dependence of the HT parameterization. The isospin asymmetry in H_T is consistent with zero within uncertainties [23] and therefore is neglected in our analysis. The isospin asymmetry in H_2 is also small [23]. Although the values of this latter have a better statistical significance, we set it to zero as well in order to avoid a potential bias in the nuclear corrections extracted from the global QCD fits. In summary, we fit two twist-4 coefficients for the isoscalar nucleon, H_2^N and H_T^N. These power corrections are parameterized as cubic spline functions of x.

The nuclear corrections for the deuteron are calculated according to the model described in Sec. II. We do not include meson exchange currents and coherent nuclear effects (shadowing) for the deuteron, since their impact is negligible in the kinematic coverage of our analysis (see Fig. 2) and we are mainly focused on the study of the off-shell correction. The only free parameters entering the nuclear corrections are the ones describing the off-shell function $\delta f(x)$, which are extracted simultaneously with the PDFs and HT terms. To this end, we use a parameterization with generic second- and third- order polynomials for $\delta f(x)$. We verified that there is no statistically significant difference between these two options within the accuracy of the data samples used in our analysis.

B. Data Samples

In our analysis most of the information about the deuteron is provided by the inclusive DIS data off deuterium from the SLAC E49, E87, E89, E139, E140 [19, 56–58, 91] and the CERN BCDMS [54], NMC [55] experiments, as well as by the ratio of Drell-Yan production in pD and pp collisions from the Fermilab E866 experiment [62]. In addition, the recent direct measurement of the ratio F_2^D/F_2^N [30] by the BONuS experiment [92] at Jefferson laboratory allows a better disentanglement of the nuclear corrections in the deuteron from possible variations of the d/u ratio in the nucleon. Since most of the BONuS data either have low values of Q^2 or are in the resonance region, we only include the BONuS points with $Q^2 > 1.5$ GeV^2 and $W > 1.6$ GeV. Although these cuts are less stringent than the ones we apply for the other data sets, they are justified by a partial cancellation of HT effects in the ratio F_2^D/F_2^N and by the relevance of the direct BONuS measurement for our study.
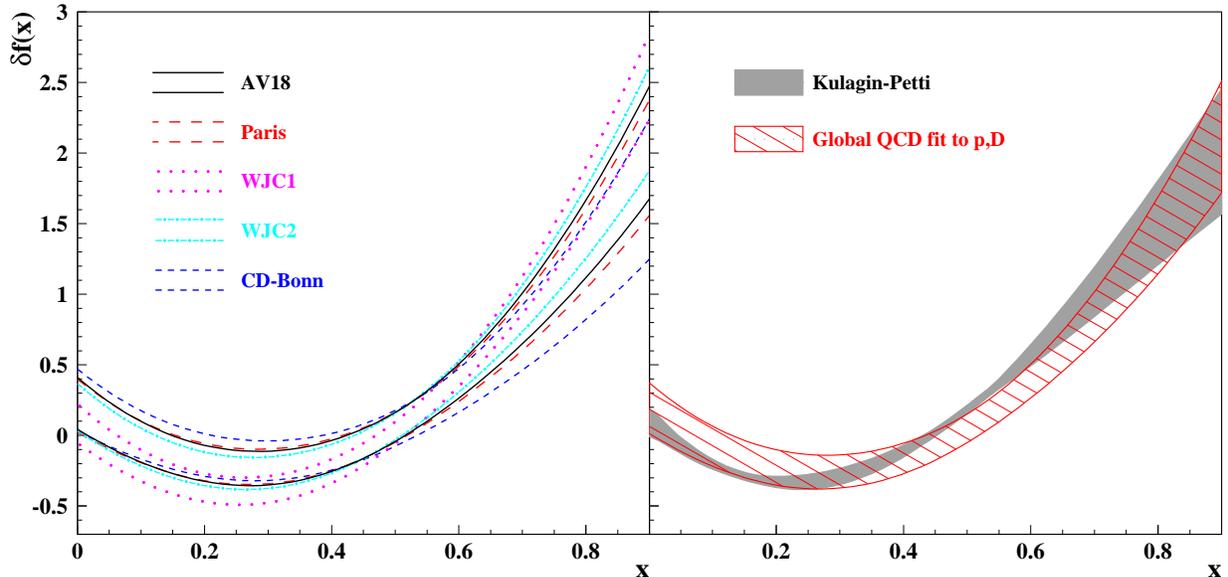
FIG. 3. Left panel: Comparison of the off-shell functions $\delta f(x)$ ($\pm 1\sigma$ uncertainty bands) extracted within our global QCD fit including all data sets in Table II by using different models for the deuteron wave functions. Right panel: Summary $\pm 1\sigma$ uncertainty band on $\delta f(x)$ obtained from this analysis, including the statistical uncertainty from the fit and the one related to the choice of the deuteron wave function. The corresponding uncertainty band obtained from heavy target data ($A \geq 4$) in Ref. [25] is shown for comparison.

For consistency with the ABMP16 fit [2] and to better constrain the sea quark distributions we include the Drell-Yan data in pCu interactions by the E605 experiment, as well as charm production data in (anti)neutrino interactions off heavy targets by the CCFR [61], NuTeV [61], NOMAD [59], and CHORUS [60] experiments. We verified that these data using nuclear targets do not affect our results on the deuteron by performing dedicated fits with and without such data. It is also worth noting that NOMAD and CHORUS measured the ratios of charm to inclusive charged-current cross sections and it was shown that the corresponding nuclear corrections cancel out at the sub-percent level in such ratios [93].

Nuclear corrections in the deuteron can be determined by comparing the available data for deuterium targets with the ones originated from interactions on the free nucleons. However, inclusive DIS data off protons do not allow to disentangle the d and u quark distributions because of the lack of a corresponding free neutron target.

The limitations of DIS data can be partially overcome with the addition of Drell-Yan, $W^{\pm}$, and Z production at Tevatron and the LHC [2, 64–72]. In particular, the data on W^+ and W^- production allow a d/u separation independent from the deuterium data. We note that the same $W^{\pm}$ production data sets collected at Tevatron and the LHC may result in two distinct (but correlated) measurements: i) the $l^{\pm}$ lepton asymmetry from the $W^{\pm}$ decays, ii) the actual $W^{\pm}$ asymmetry. The former is more closely related to the experimental observables, while the latter requires model-dependent acceptance corrections to account for the kinematics of the $W^{\pm}$ decay. As discussed in Ref. [94], some inconsistencies between the $l^{\pm}$ lepton asymmetries and the $W^{\pm}$ asymmetries obtained from the same experimental

data sets are observed. For this reason whenever both measurements are available, we only consider the $l^\pm$ lepton asymmetry data in our analysis.

Table II summarizes all the data sets used in the present analysis. In order to exclude the region of resonance production and to reduce the impact of HT corrections we require $Q^2 > 2.5 \text{ GeV}^2$ and $W > 1.8 \text{ GeV}$ for DIS data.

IV. RESULTS AND DISCUSSION

The general features of the PDFs extracted from the global fits described above as well as a detailed discussion of the individual data samples considered were presented elsewhere [2]. In this paper we focus on the nuclear correction extracted from deuterium data and on the corresponding impact for the d/u ratio.

A. Off-shell Function δf

The results of our determination of the off-shell function $\delta f(x)$ from the global QCD fits described in Sec. III are shown in Fig. 3. A simultaneous extraction of the off-shell function with both PDFs and HTs is possible because of the different Q^2 dependence of these three contributions and the wide Q^2 coverage of the data sets listed in Table II. In general, nuclear corrections to the deuteron data are partially correlated to the d-quark distribution. In order to reduce this correlation, the role of Drell-Yan and $W^\pm$ production at pp and $p\bar{p}$ colliders is crucial. In particular, the recent combined D0 data and the LHC data from LHCb, reaching values of $x \sim 0.8$ due to the wide rapidity coverage, offer precisions comparable to the ones of older fixed-target DIS experiments.

Figure 3 illustrates the dependence of the fitted $\delta f(x)$ function upon the choice of the deuteron wave function $\Psi_D(\boldsymbol{p})$ among the models listed in Sec. II A. The main differences are related to the high momentum component of the wave function, as shown in Fig. 1. Since this high momentum tail controls the region of large nucleon virtuality v, the off-shell correction in the large x region is in principle sensitive to the corresponding nuclear smearing in Eq.(2), which modifies the x and Q^2 dependence of the structure functions. A general trend can be observed from Fig. 3, with the harder wave function resulting in a slightly higher off-shell function at large x. Since the overall off-shell correction has opposite sign with respect to $\delta f(x)$ in Eq.(5), this trend implies an anti-correlation between $\Psi_D(\boldsymbol{p})$ and $v\delta f(x)$ in global QCD fits. From Fig. 3 we note that our results obtained with the Paris, CD-Bonn, AV18, WJC1, and WJC2 wave functions are all consistent within the corresponding uncertainties and indicate a relatively limited spread. This robustness against the modeling of the deuteron wave function can be explained by the use of data samples which can reduce the correlation between the nuclear correction and the d-quark distribution. In this context the recent direct measurement of the ratio F_2^D/F_2^N from the BONuS experiment constrains the overall normalization of the nuclear corrections in our fits.

A more precise determination of the off-shell function $\delta f(x)$ was obtained in Ref [25] from heavy nuclei with $A \geq 4$, as described in Sec. II D. In order to further study the sensitivity to the nuclear smearing in Eq.(2), we repeat the standalone extraction of $\delta f(x)$ in Ref [25] after rescaling the nuclear spectral function describing the properties of heavy nuclei [25] by the ratios of the various deuteron wave functions shown in Fig. 1. The results summarized in Fig. 4 demonstrate a small sensitivity to the choice of the nuclear spectral function and/or
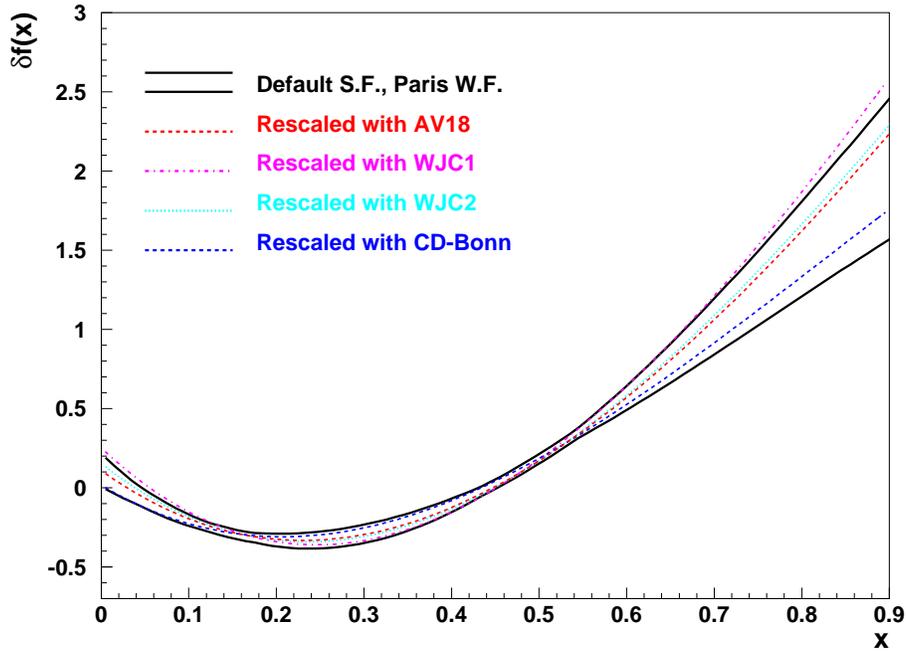
FIG. 4. Comparison of the off-shell functions $\delta f(x)$ extracted from the analysis of the ratios of nuclear structure functions with $A \geq 4$ [25] by varying the nuclear spectral function and the deuteron wave function. The nuclear spectral function has been rescaled by the ratios of the various models for the deuteron wave functions shown in Fig. 1. The solid band represents the overall $\pm 1\sigma$ uncertainty on $\delta f(x)$ from Ref. [25], including model systematics.

of the deuteron wave function, as well as a dramatic reduction of the uncertainties with respect to Fig. 3. This reduction can be explained by the different observables considered in the two independent extractions. In the global QCD fits we use the absolute DIS cross sections off the deuteron, while in the standalone determination of Ref [25] we consider only ratios $\mathcal{R}(A/B) = F_2^A/F_2^B$ between two nuclear targets A and B. Many model uncertainties largely cancel out in such ratios. For the same reason the data sets used in Ref [25] are more accurate than the deuteron ones, making them an ideal tool to study the off-shell function $\delta f(x)$. The $\pm 1\sigma$ uncertainty band shown in Fig. 4 includes model systematics due to the spectral and wave functions, the functional form, the PDFs, as well as corrections due to meson exchange currents and nuclear shadowing.

A comparison between the two independent determinations of the off-shell function δf is given in Fig. 3. Since the five individual determinations of the off-shell function δf from deuterium data using different wave functions are characterized by a comparable fit quality, we combine them by taking an average of both the central values and the corresponding uncertainties. The resulting $\pm 1\sigma$ uncertainty band is shown in the right panel of Fig. 3. This band summarizes our determination of δf from deuteron data and is consistent with the more precise determination of δf from the analysis of the ratios of nuclear structure functions for $A \geq 4$ [25]. Since we are using a generic polynomial to parameterize δf (Sec. III A), no functional form bias is present in this comparison. The agreement between
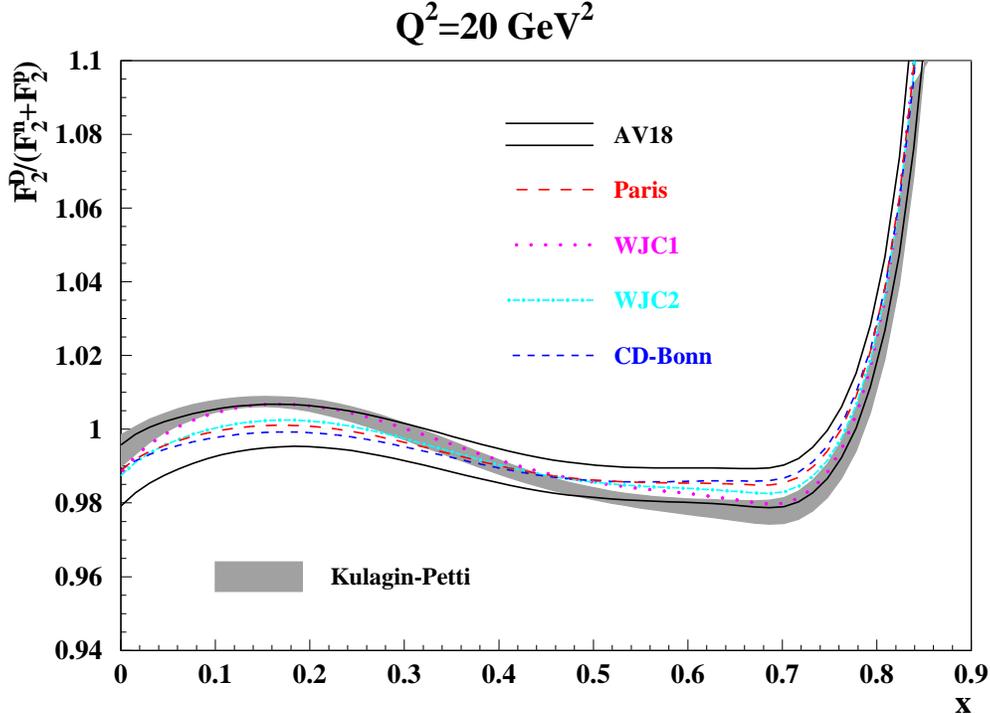
FIG. 5. Ratio F_2^D/F_2^N between the deuteron and the isoscalar nucleon structure functions and its $\pm 1\sigma$ uncertainty band obtained using the off-shell functions in Fig. 3 convoluted with the corresponding models for the deuteron wave function. The $\pm 1\sigma$ band for the same ratio obtained with the off-shell function δf from Ref. [25] is displayed as a shaded area for comparison.

the two independent determinations supports the interpretation of the off-shell function δf as a universal structure function of the nucleon, validating the unified treatment of the deuteron and heavier nuclei developed in Ref [25].

B.   Nuclear Corrections to F_2^D/F_2^N

The nuclear correction stemming from the FMB and OS effects on the ratio F_2^D/F_2^N is given in Fig. 5. This ratio is particularly interesting because it represents the overall nuclear corrections in the deuteron. The variation due to the choice of the deuteron wave function in the global QCD fit appears to be even smaller in the ratio F_2^D/F_2^N than in the off-shell functions δf given in Fig. 3. This behavior can be explained by the anti-correlation between $\Psi_D(\boldsymbol{p})$ and $v\delta f(x)$ discussed in Sec. IV A: a larger off-shell function partially compensates a reduced strength of the high momentum component of the wave function so that the observable structure function remains consistent with the fitted data. In addition, the recent BONuS measurement significantly constrains the ratio F_2^D/F_2^N, as mentioned in Sec. IV A. The results obtained from deuteron data with such a constraint agree with the predictions from Ref [25] based upon a standalone analysis of heavy nuclei with $A \geq 4$, as shown in Fig. 5 [1].

---

[1] The result with the off-shell function δf from Ref [25] shown in Fig. 5 is slightly different with respect to the calculation in Fig. 2. The differences are mainly at large x values and appear due to the fact that the
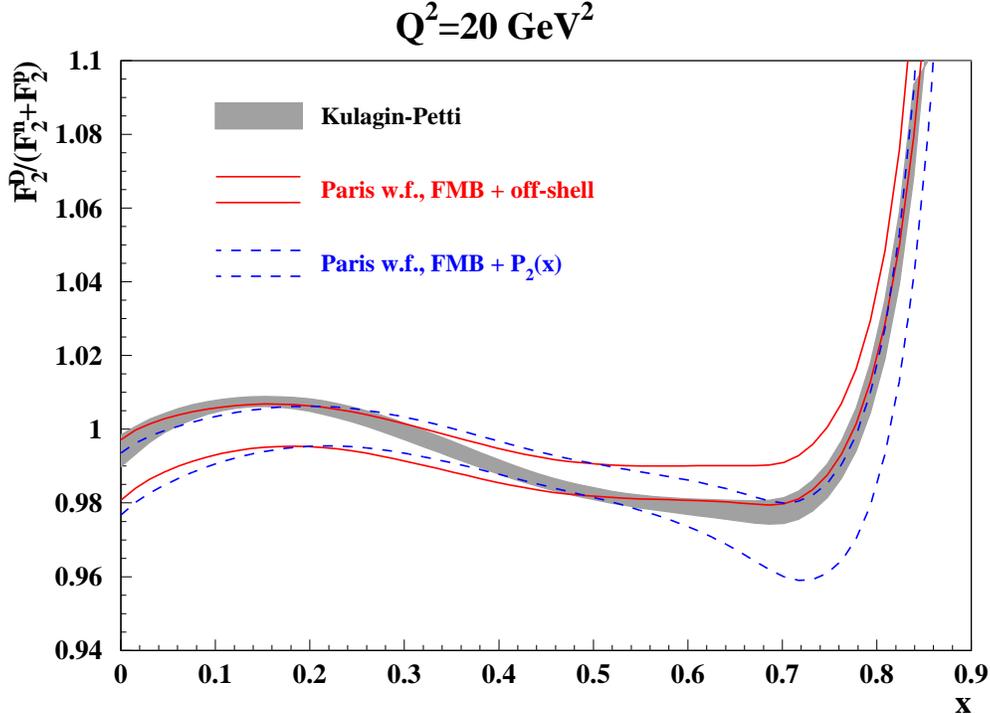
FIG. 6. Test of model dependence in the extraction of the off-shell function. The dashed lines represent the $\pm 1\sigma$ uncertainty band obtained by fitting a generic polynomial as off-shell correction to F_2^D/F_2^N, instead of using the nuclear convolution with δf in Eq.(2) (solid lines). The $\pm 1\sigma$ uncertainty band obtained from heavy target data ($A \geq 4$) in Ref. [25] is displayed as shaded area for comparison.

Although the off-shell function δf is extracted in our analysis as a generic polynomial, we are still calculating the nuclear correction to the structure functions using the nuclear convolution in Eq.(2), following the prescriptions of the model of Ref. [25]. In order to verify verify wether this procedure introduces any indirect model dependence in our results, we perform a separate fit in which we parameterize the overall off-shell correction to the structure function F_2^D as a generic polynomial added to the standard FMB correction. In this approach the fitted off-shell correction is model-independent as it is not part of the nuclear convolution in Eq.(2). The results obtained with such a parameterization shown in Fig. 6 are in good agreement with the corresponding fits based upon the nuclear convolution with the off-shell function δf. We can thus conclude that the functional form we are using in our fits for δf is flexible enough to reproduce the data and that our modeling of nuclear effects does not introduce any significant bias.

C. Systematic Studies

As we discuss in Sec. IV A, the uncertainty on the off-shell function $\delta f(x)$ related to the modeling of the deuteron wave function turns out to be negligible compared to the statistical

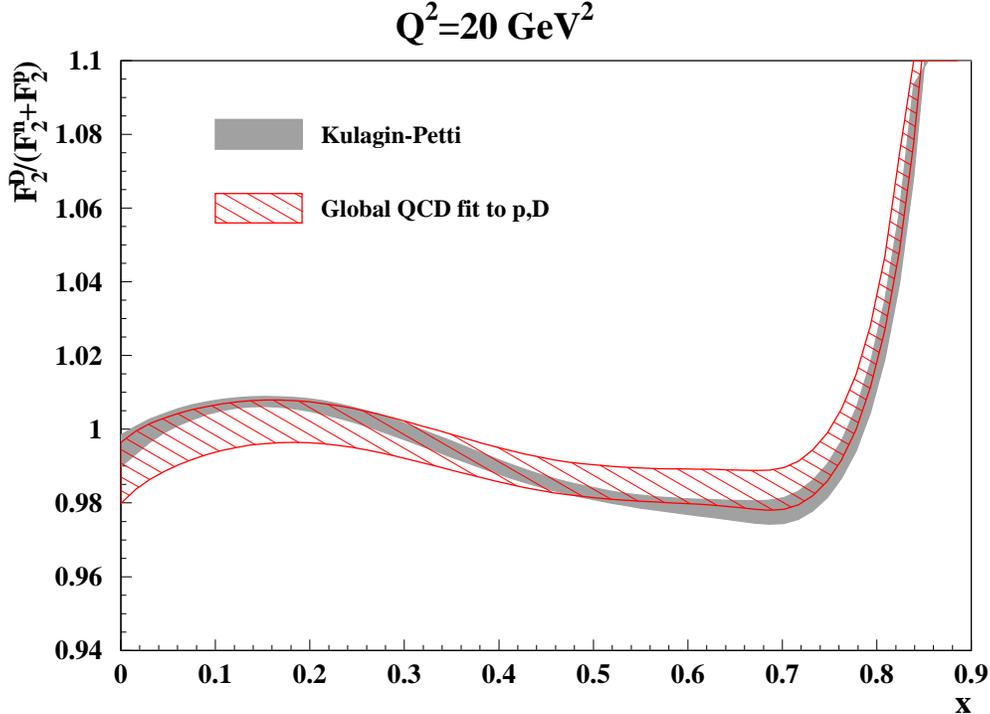results shown in Fig. 5 are obtained with the PDFs and HT terms extracted from our global QCD fit.

FIG. 7. Summary of the ratio F_2^D/F_2^N obtained from this analysis with the corresponding total $\pm 1\sigma$ band including the fit uncertainty and the one related to the choice of the deuteron wave function. The $\pm 1\sigma$ uncertainty band obtained from heavy target data ($A \geq 4$) in Ref. [25] is also shown as a shaded area.

uncertainty from our global QCD fits. Our final $\pm 1\sigma$ band on the ratio F_2^D/F_2^N, computed by averaging the results from the five individual fits with different wave functions in Fig. 5, is given in Fig. 7, including the statistical uncertainty from the fit and the one related to the choice of the deuteron wave function added in quadrature.

In Ref. [42] it was shown that there is some tension between the DIS data sets from the BCDMS and SLAC experiments, resulting in significant modifications of the extracted HT terms and PDFs. In order to mitigate the impact of this tension on our studies we allow the overall normalization of both the BCDMS proton and deuteron data sets to vary freely in our fits. This approach is justified by the use of separate normalizations for the deuterium and proton data sets in the BCDMS measurements [53, 54]. The normalization of the BCDMS proton data is essentially defined by the precise HERA data in the overlap region, resulting in an overall factor consistent with the corresponding normalization uncertainties quoted by the experiment, up to 3%. The only constrain on the normalization of the BCDMS deuteron data comes from the SLAC experiments. However, the partial correlation between this normalization and the determination of the deuteron nuclear correction can potentially introduce an additional uncertainty in the global fits. The recent direct measurement of the deuteron nuclear correction by the BONuS experiment substantially reduces this uncertainty by constraining the normalization of the overall nuclear corrections. As a result, the normalization factor for the BCDMS deuterium data obtained from our fits is stable against variations of the deuteron wave function and is very close to the one obtained for
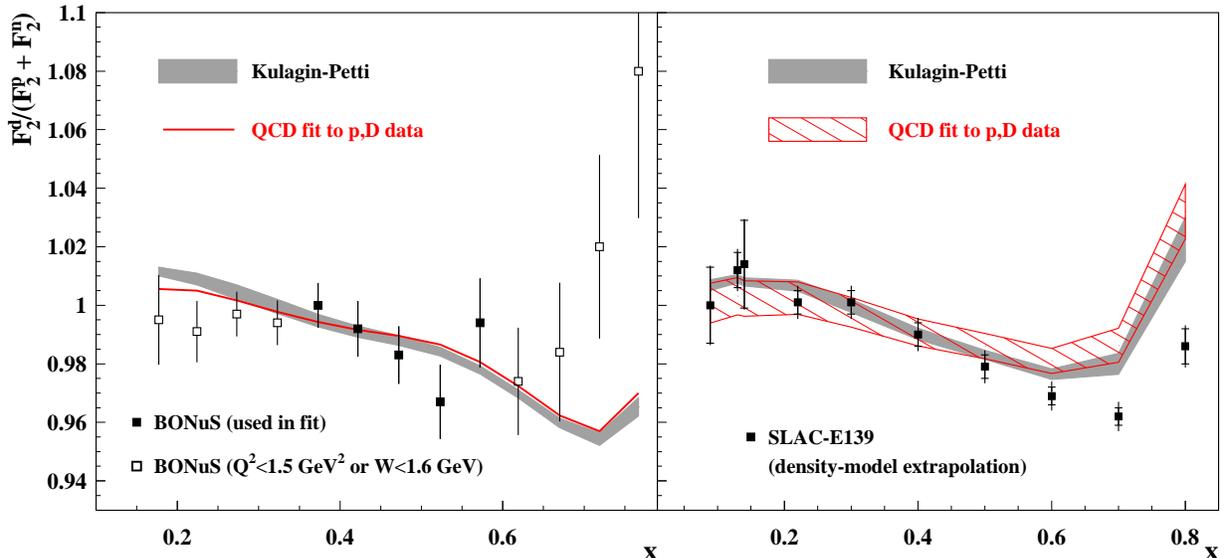
FIG. 8. Left panel: The ratio F_2^D/F_2^N measured by the BONuS experiment [30] compared to the central value obtained from this analysis (solid line), and to the $\pm 1\sigma$ uncertainty associated with the predictions from Ref. [25] (shaded area). Right panel: The model-dependent extrapolation of heavy target data from SLAC E139 [19] within the nuclear density model [14] compared to the corresponding $\pm 1\sigma$ uncertainty band from this analysis (hatched band) and to the predictions from Ref. [25] (shaded area).

the BCDMS proton data.

As an additional test of the robustness of our analysis, we perform separate fits to different subsets of the data listed in Table II. These variants of our analysis do not indicate any anomalous tension related to individual data samples, but rather suggest that our results on the off-shell correction are originated from the combined fit of all deuteron data sets.

D.   Discussion

The results of our analysis discussed in Sec. IV A support the predictions of Ref. [25] for nuclear effects in the deuteron and the unified treatment for all nuclei. We can then exploit the higher precision offered by the existing DIS data off heavier nuclear targets ($A \geq 4$) to fix the value of the off-shell function δf used in global QCD fits following Ref. [25]. The corresponding reduction of the overall uncertainties on the deuteron nuclear corrections is illustrated in Fig. 7. Within a simple single-scale model, relating the quark momentum distributions in the nucleon to the nucleon radius [25], the off-shell function δf obtained from nuclei with $A \geq 4$ suggests an increase of the nucleon core radius by about 2% in the deuteron, taking an average virtuality of -0.045 from Table I. This value is comparable to estimates obtained with a different model [95] in relation to an increase of the overlap of nucleons in nuclei with the nuclear density.

The predictions from both the present analysis and the ones from Ref. [25] are compared

with the recent direct measurement of the ratio F_2^D/F_2^N by the BONuS experiment [30] in Fig. 8. A good agreement with BONuS data is observed, although the current experimental accuracy is somewhat limited and many data points fall in the resonance region with $W < 1.6$ GeV or correspond to low $Q^2 < 1.5$ GeV^2. The model-dependent extrapolation of the F_2^A measurements with $A \geq 4$ performed by the SLAC E139 experiment [19] is also compared with our results and with the ones from Ref. [25] in Fig. 8. Although the basic assumption of Ref. [19] about the scaling of the magnitude of nuclear effects with the nuclear density [14] was excluded by the recent measurements on ^{9}Be target by the E03-103 experiment [48], the results of Ref. [19] with $x < 0.7$ are consistent with our predictions. It is worth noting that the typical Q^2 values of the E139 data are substantially larger than in the BONuS sample ($2-15$ GeV^2 vs. $1-4$ GeV^2). This difference allows to demonstrate the Q^2 dependence in the nuclear corrections, which appears mainly at large x due to the combined effect of the TMC and the off-shell corrections, c.f. Fig. 8.

Since nuclear corrections are almost linearly dependent from x in the region $0.35 < x < 0.55$, they are often quantified by the corresponding linear slope. The main advantage of the slope is that it can be measured more accurately since it is not affected by the normalization uncertainties. The empirical model-independent determination of the slope $\mathrm{d}\mathcal{R}(D/N)/\mathrm{d}x = -0.100\pm0.050$ [30] of the BONuS data agrees well with the value -0.099 ± 0.006 predicted by Ref. [25]. To this end, the model-dependent extrapolation of the SLAC E139 data [19] gives a consistent value of -0.098 ± 0.005, while the empirical extrapolation using the short range correlation scale factors from Ref. [28] results in a somewhat smaller slope -0.079 ± 0.006. However, while being useful for the analysis of experimental data, the slope $\mathrm{d}\mathcal{R}(D/N)/\mathrm{d}x$ describes the behavior of the nuclear corrections in a limited region only. Meanwhile, the microscopic model of Ref. [25] reproduces not only the measured slopes, but also the shape and magnitude of the nuclear corrections in the entire kinematic range covered by existing data.

Our results on the off-shell correction differ from the ones obtained using a similar formalism in Refs. [31, 96]. The analysis of Ref. [31] is based on a modified model of Ref. [25] to relate the off-shell function δf to an increase in the nucleon confinement radius in the nuclear medium. The analysis of Ref. [96] follows more closely the model of Ref. [25], with the off-shell function δf being determined from a global QCD fit to the deuteron and proton data. The differences in the results of those analyses with respect to our study can be attributed to the implementation of the deuteron model and the details of the calculations [97], as discussed in the Appendix.

As discussed in Sec. IV A, the ratio of DIS structure functions for two different nuclear targets, $\mathcal{R}(A/B) = F_2^A/F_2^B$, offers a good tool to study the off-shell function $\delta f(x)$, due to a large cancellation of both experimental and model uncertainties. In this respect we note that additional data on nuclear ratios $\mathcal{R}(A/B)$ should improve the determinetion of the nucleon off-shell function δf and of the d quark distribution in the proton.

E. Constraints on d/u and F_2^n/F_2^p

Correlations between the deuteron nuclear corrections and the d-quark distribution can substantially limit the PDF accuracy achievable in the PDF fits based on the proton and deuterium DIS data. In this context the data from flavor sensitive processes like $W^\pm$ production in $pp(\bar{p})$ collisions play a major role in reducing such a correlation. A possible approach to avoid the effects of the deuteron nuclear corrections is to avoid any DIS data off
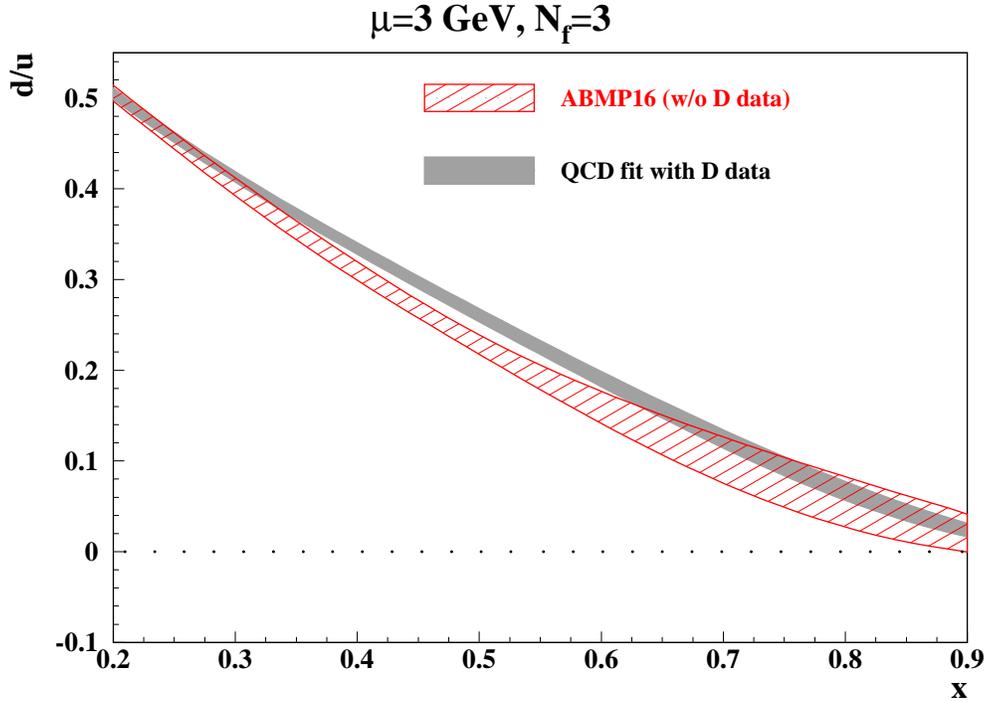
FIG. 9. Ratio d/u at a factorization scale $\mu = 3$ GeV as a function of x obtained from global QCD fits. The hatched $\pm 1\sigma$ error band corresponds to the ABMP16 fit [2], which does not include deuteron data. The shaded band shows the corresponding results obtained by including DIS deuteron data and by using the off-shell function $\delta f(x)$ and its uncertainty determined in Ref. [25].

the deuteron in global QCD fits, as in the recent ABMP16 analysis [2]. The corresponding results for the d/u ratio shown in Fig. 9 indicate that the recent precision data on $W^\pm$ boson production from D0 and the LHC experiments (c.f. Table II) provide a good sensitivity to the d-quark distribution. In particular, the d/u ratio at large $x > 0.7$ is well constrained, mainly due to the large rapidity data from the recent LHCb measurement of $W^\pm$ boson production [70, 71]. This sample indeed probes values of x up to 0.8 and its accuracy is comparable to the one of DIS experiments.

The universality of δf allows a further improvement of the accuracy in the determination of the d/u ratio, by using the deuteron DIS data in combination with the more precise off-shell function obtained from the analysis of the nuclear targets with $A \geq 4$ (see ec. IV A). In Fig. 9 we show the d/u ratio obtained in such a way in the ABMP16 fit. The impact of the DIS deuteron data on the d/u ratio is more evident in the region of $x > 0.4$, where the uncertainties are substantially reduced, as compared to the ABMP16 results obtained without the DIS deuteron data.

An interesting observation is that the d/u ratio tends to vanish as $x \to 1$ (see Fig. 9). In order to verify that this behavior is not biased by the functional form of the PDF parametrization, we multiply the d- and u- quark parameterizations by an additional free polynomial. We do not find any significant impact on the corresponding large-x behavior of the d/u ratio, thus confirming its stability. Furthermore, the value of d/u obtained in the present analysis is consistent with the one of the ABMP16 fit performed without using the
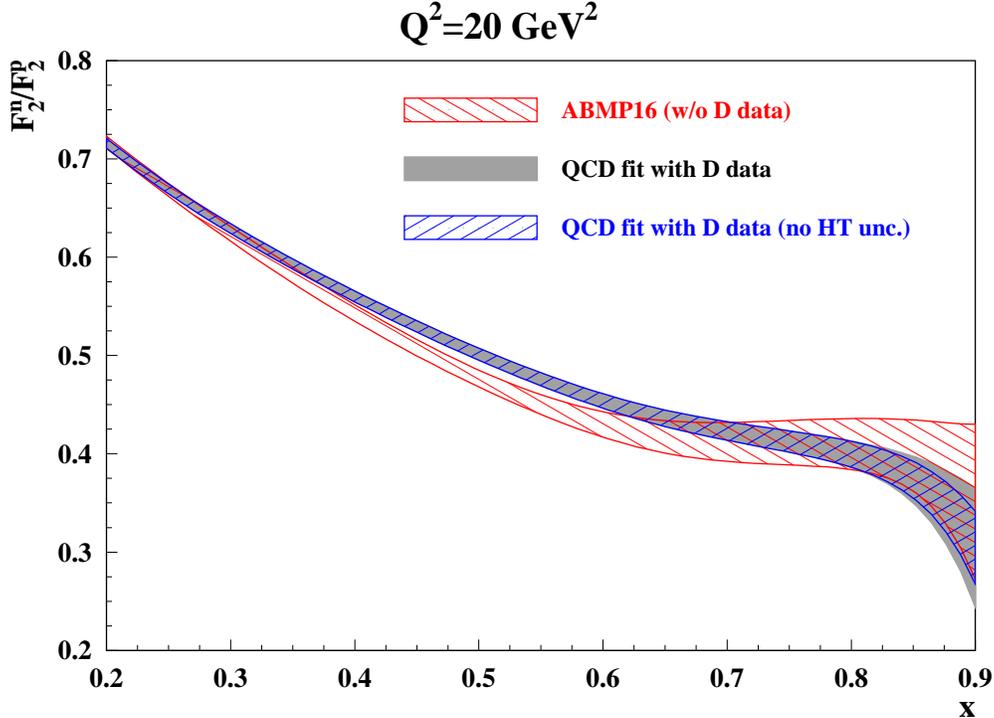
FIG. 10. The same as in Fig. 9, for the ratio F_2^n/F_2^p as a function of x at $Q^2 = 20\,\mathrm{GeV}^2$. The $\pm 1\sigma$ error band corresponding to the QCD fit with the deuteron data included and ho HT uncertainties taken into account is also given for comparison (right-tilted hatch).

deuteron data, c.f. Fig. 9. These results indicate that our d- and u- quark parameterizations are flexible enough to be driven by the data sets, rather than by the functional form used (c.f. Appendix).

The d/u ratio is related to the neutron to proton structure function ratio, F_2^n/F_2^p, displayed in Fig. 10. The impact of the deuteron nuclear correction on F_2^n/F_2^p is somewhat larger than on the d/u ratio. Note, however, that the behavior of F_2^n/F_2^p at $x \to 1$ is dominated by the HT contributions, which introduce a significant uncertainty on this ratio, as demonstrated in Fig. 10.

It is instructive to compare the nuclear corrections applied to the deuteron data in various PDF analyses. The CJ15 analysis [96] is based upon a formalism similar to the model of Ref. [25], with the corresponding off-shell correction in the deuteron determined from the global QCD fit (see discussion in Sec. IV D). The MMHT14 analysis [3] is based on an empirical parametrization of the nuclear correction to the ratio F_2^D/F_2^N, which is extracted from data. The CT14 [5] and NNPDF3.0 [4] do not use any nuclear correction arguing that nuclear corrections would introduce additional uncertainties in the analysis [4]. Furthermore, it was claimed that nuclear corrections can be neglected when using more stringent cuts in Q^2 and W^2 [5]. The nuclear corrections used by CJ15 and MMHT14, while mutually consistent, are characterized by relatively large uncertainties (c.f. Fig. 11). Since these two determinations are driven by the available deuterium data and are largely correlated with the d/u ratio, the possibility to improve such uncertainties in the context of global QCD fits appears to be limited. As discussed above, the use of the microscopic model of
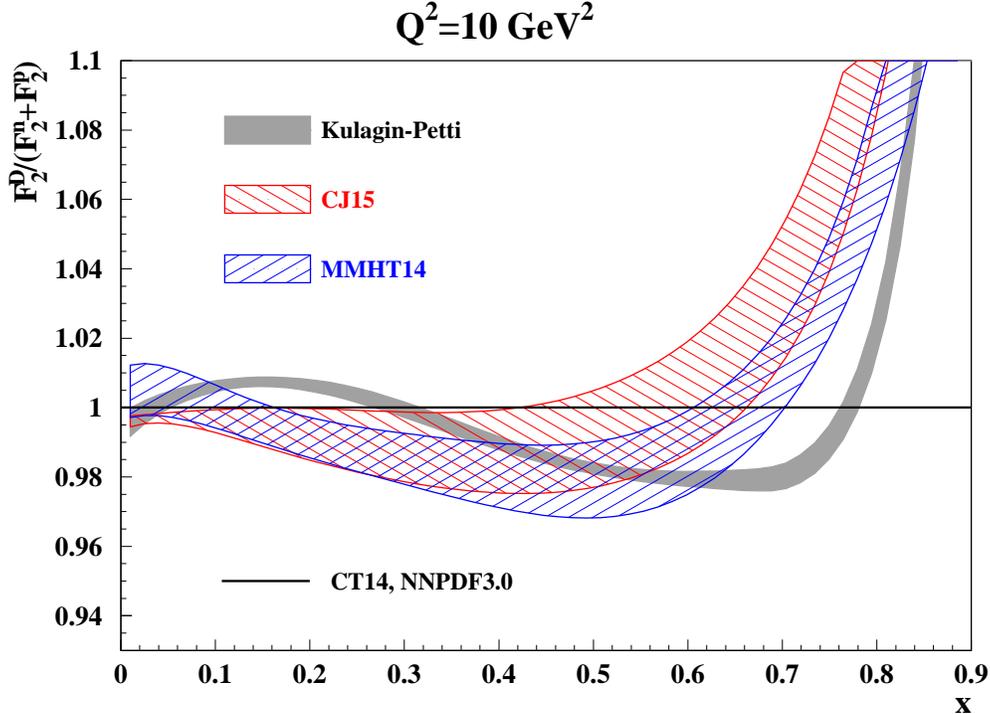
FIG. 11. Comparison of the F_2^D/F_2^N parameterization with the corresponding uncertainties used to correct for nuclear effects in the CJ15 [96], MMHT14 [3], CT14 [5], and NNPDF3.0 [4] analyses. The solid gray band gives the $\pm 1\sigma$ uncertainty associated with the predictions of Ref. [25].

Ref. [25] and of the corresponding results on the universal off-shell function δf, allows a substantial reduction of the uncertainties related to the nuclear correction in the deuteron. Such a reduction is illustrated by the comparison with the model of Ref. [25] given in Fig. 11. In turn, the uncertainty on the deuteron nuclear correction is a bottleneck in the overall accuracy of the d/u ratio obtained from global QCD fits. The differences in the treatment of the nuclear corrections in the deuteron illustrated in Fig. 11 would translate into a corresponding spread in the d-quark distribution. It is worth noting that these systematic effects can not be mitigated by more stringent Q^2 and W^2 cuts, since nuclear effects survive even at very large energy and momentum, as demonstrated in DIS experiments (for a review see Refs. [6, 7]) and by recent observations of nuclear modifications in $p + \text{Pb}$ and $\text{Pb} + \text{Pb}$ collisions at the LHC [98–104].

The largest deviations between the CJ15 and MMHT14 analyses and the model of Ref. [25] are observed in the intermediate region $x \sim 0.15$ and at large $x > 0.6$ (c.f. Fig. 11), although their significance is limited by the current uncertainties. Our analysis of deuteron data is consistent with Ref. [25], with an uncertainty band extending close to the CJ15 and MMHT14 ones (c.f. Fig. 7). The small enhancement present in the model of Ref. [25] for $0.05 < x < 0.3$ is the result of an interplay of the off-shell correction and the meson exchange currents (c.f. Fig. 2). Nuclear corrections at $x > 0.6$ are instead dominated by the FMB and OS corrections. This kinematic region is very sensitive to the treatment of both the bound nucleon momentum distribution and the target mass corrections to the nucleon structure functions (Section II). It is worth noting that the prescription of Ref. [41] for the

TMC is known to have an incorrect behavior for $x \to 1$, which can affect the calculations at very large x values. More detailed comparisons with the CJ15 and MMHT14 results can be found in the Appendix. Future DIS measurements from the BONuS experiment with the 12 GeV JLab upgrade [105], from neutrino and anti-neutrino scattering off free proton in the DUNE experiment [106], and from the electron-ion collider [107] can further improve our understanding of the ratio F_2^D/F_2^N.

V. SUMMARY

We performed a study of nuclear effects in the deuteron using the data from DIS off proton and deuterium, Drell-Yan production in pp and pD interactions, and $W^{\pm}$- and Z-boson production in pp and $p\bar{p}$ collisions in the context of global QCD fits. We found that it is possible to determine simultaneously PDFs, high twist terms, and the off-shell function describing the modification of PDFs in bound nucleons due to their different Q^2 dependence and the wide kinematic coverage of existing data. Flavor sensitive processes like $W^{\pm}$ production in $pp(\bar{p})$ collisions play an important role in disentangling the impact of the nuclear corrections in the deuteron from the d quark distribution function, allowing a more accurate determination of both contributions. We also evaluated the sensitivity of our results to various models of the deuteron wave function and found that the corresponding model dependence is reduced by the recent BONuS measurement of the ratio F_2^D/F_2^N.

The results on the off-shell function δf reported in this paper are in good agreement with the earlier determination obtained in the analysis of the ratio of nuclear structure functions with mass number $A \geq 4$ [25]. This result confirms the universality of δf, which can be regarded as a special structure function of the nucleon describing the modification of the bound nucleons in the nuclear medium. This study supports the unified treatment of the deuteron and heavier nuclei developed in Ref. [25].

We also studied the impact of deuteron nuclear corrections on the d/u ratio within global QCD fits. We found that the recent precision data on $W^{\pm}$ boson production from D0 and the LHC experiments allow a reduction of the uncertainties in the d/u ratio at large x. Our results indicate that the accuracy in the determination of the d/u ratio can be further substantially improved by including the DIS data off deuterium target corrected for nuclear effects using the model of Ref. [25] with the universal off-shell function δf.

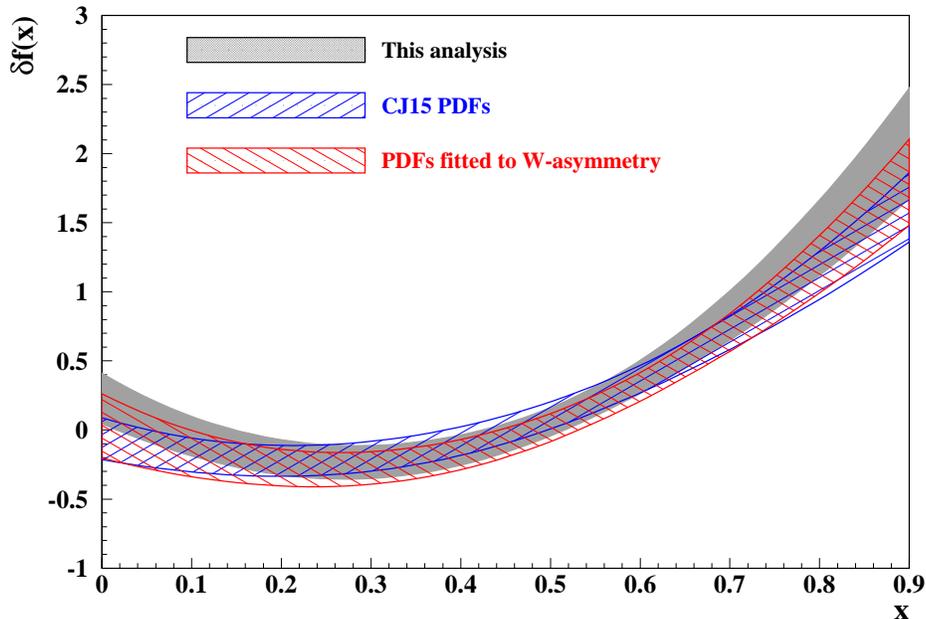## Appendix: Results from different analyses

FIG. 12. Comparison of the off-shell functions $\delta f(x)$ ($\pm 1\sigma$ uncertainty bands) extracted from different global QCD fits including deuteron data (shaded area: our results with the AV18 deuteron wave function [38]; left-tilted hatch: result obtained by using the D0 data on W-boson asymmetry [108]). The result from a NLO fit using fixed CJ15 proton PDFs from the LHAPDF library [109] is also displayed (right-tilted hatch).

A few phenomenological studies aimed to extract the nuclear corrections in the deuteron from global QCD fits are available in literature [3, 31, 96], including the analyses of Refs. [31, 96] based on a deuteron model similar to that used in our studies. In this Appendix we investigate the differences with respect to those studies [97] and benchmark the recent CJ15 analysis [96].

Clear differences with the CJ15 fit appear in the data sets used and in the proton PDFs correspondingly obtained. From Table I of Ref. [96] we conclude that most of the sensitivity to the nuclear corrections in the CJ15 fit manifests for the D0 data on the $W^{\pm}$-boson production asymmetry and the DIS deuteron data by the SLAC experiments (c.f. reduction of χ^2 obtained by adding the nuclear corrections). While we include the SLAC DIS data (Table II), we use significantly different W, Z-boson collider data in our analysis. (see Sec. III B). More specifically, we include the D0 data on the lepton asymmetry from the W-boson decays rather than the actual W-boson asymmetry data, as well as the recent LHC DY data [66–72]. In order to test the impact of such differences, we perform a variant of our fit in which we drop all the W, Z-boson collider data, replacing them with the D0 W-asymmetry data [108]. The corresponding results for the function δf are consistent with the ones presented in Sec. IV A (c.f. Fig. 12). We also perform a separate NLO fit to all deuteron data sets using fixed CJ15 proton PDFs from the LHAPDF library [109]. The
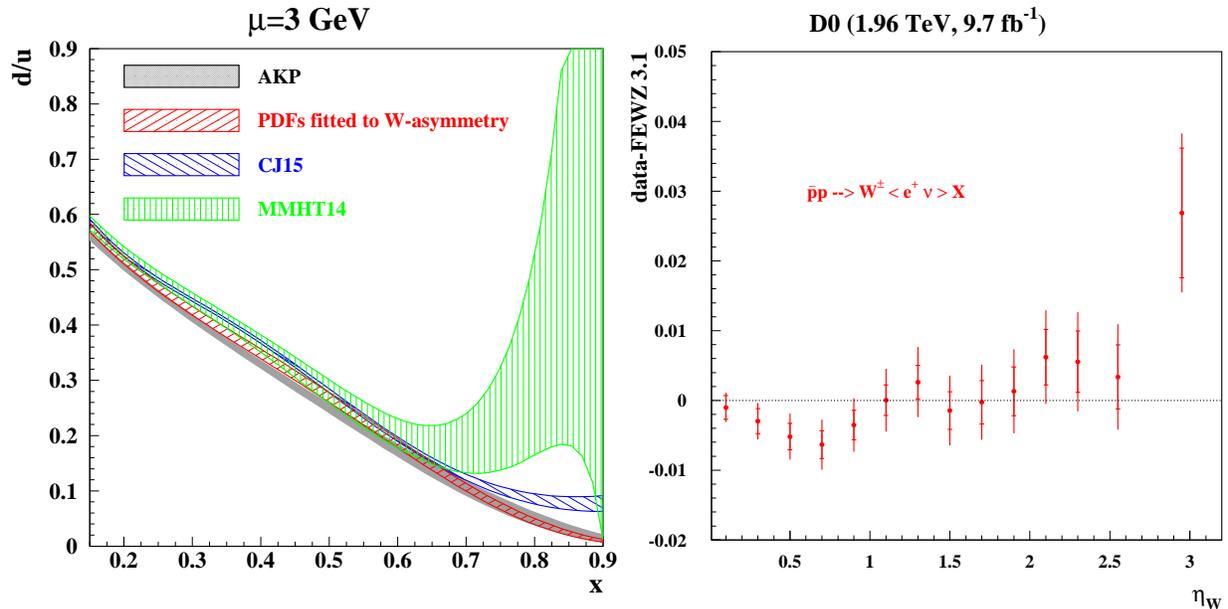
FIG. 13. Left panel: Comparison of the d/u ratio ($\pm 1\sigma$ uncertainty bands) obtained in different global QCD fits (shaded area: present analysis; right-tilted hatch: variant of the our analysis using the D0 data on W-boson asymmetry [108]; left-tilted hatch: NLO CJ15 fit [96]; vertical hatch: NNLO MMHT14 fit [3]). Right panel: Difference between the D0 W-asymmetry data [108] and the NLO predictions obtained with the FEWZ package [110, 111] using CJ15 PDFs from the LHAPDF library [109].

off-shell function δf obtained in this way is consistent with our results in Sec. IV A and Ref. [25] (c.f. Fig. 12). Therefore, we can conclude that the differences in the data samples and proton PDFs can not explain the different off-shell correction obtained in Ref. [96].

Meanwhile, it is instructive to compare the values of the d/u ratio obtained in the various fits. The d/u ratio obtained from the variant of our fit with the D0 W-asymmetry data is consistent within uncertainties with the results presented in Sec. IV E (c.f. Fig. 13). The differences between the central values suggest that our PDF parameterizations are flexible enough to describe different data sets without limitations from the functional form used. In particular, we do not explicitly constrain the d/u ratio to vanish for $x \to 1$. The flexibility of our PDF parameterization is confirmed by the fact that we obtain similar results by multiplying the d- and u- quark parameterizations by an additional free polynomial (c.f. Sec. IV E). The uncertainty of the d/u ratio determined in the MMHT14 analysis [3] at large x does not allow quantitative comparisons (c.f. Fig. 13), due to the lack of relevant experimental constraints. It is worth noting that the d/u ratio of the CJ15 analysis [96] displays a substantially different behavior at large x. Since this kinematics is largely controlled by the data on W-boson asymmetry at large rapidity, we compute this quantity at the NLO approximation using the FEWZ package [110, 111] with CJ15 proton PDFs from the LHAPDF library [109]. A comparison with the D0 data [108] (c.f. Fig. 13) shows deviations from these predictions at large values of the W-boson rapidity, in contrast with the corresponding results of the CJ15 analysis (c.f. Fig. 13 of Ref. [96]).

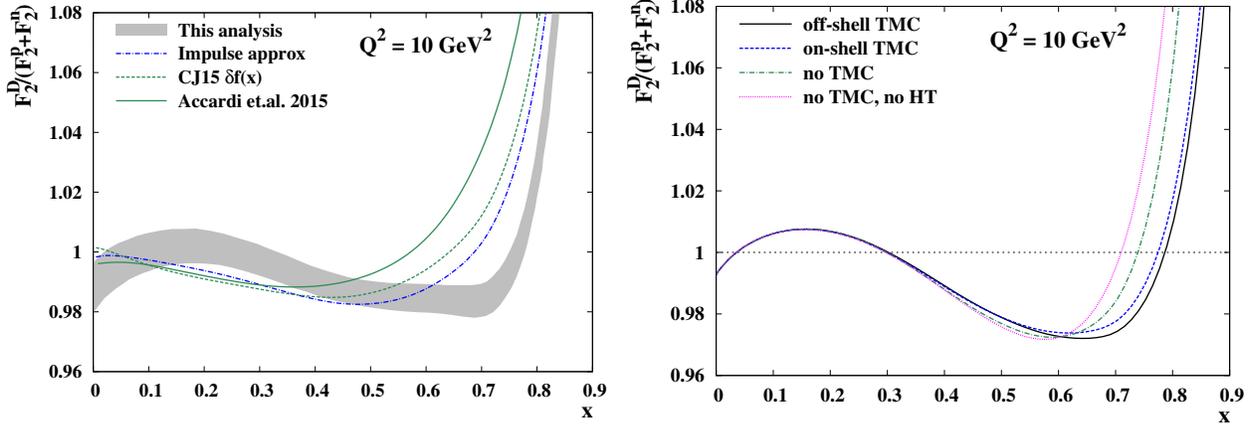A possible source of differences between this analysis and Ref.[96] may stem from the

FIG. 14. Left panel: Comparison of the ratio F_2^D/F_2^N from Ref. [96] (solid line) with our calculation using the same $\delta f(x)$ and AV18 deuteron wave function [38] (dashed line) and with the corresponding impulse approximation without OS correction (dashed-dotted line). The $\pm 1\sigma$ uncertainty band obtained from this analysis is also shown as a shaded area. Right panel: Same ratio computed with the off-shell correction from Ref. [25] using different approximations: off-shell TMC from Eq.(4) and HT (solid line), on-shell TMC [41] and HT (dashed line), no TMC (dashed-dotted line), no TMC nor HT (dotted line).

implementation of the convolution model, as well as from the treatment of the TMC and/or the HT contributions. In order to check the sensitivity to these effects we compute the ratio F_2^D/F_2^N using the $\delta f(x)$ function from Ref.[96] and the AV18 wave function [38]. The result shown in Fig. 14 (left panel) deviates significantly from both the one presented in Sec.IV A and the one of Ref. [96] using the same deuteron wave function, suggesting a different implementation of the convolution model in Ref.[96]. Figure 14 (right panel) also illustrates the sensitivity of the ratio F_2^D/F_2^N to various implementations of the TMC and HT corrections [2]. In particular, we compare the standard TMC scheme of Ref. [41] (on-shell TMC) with its off-shell continuation by Eq.(4). The TMC and HT corrections to F_2^D/F_2^N are treated differently in the CJ15 and MMHT14 analyses, and could contribute to the disagreement in the region of large x (c.f. left panel of Figure 14).

In summary, the present studies indicate that we can not reproduce the CJ15 results of Ref. [96] on the function δf. All our systematic checks are consistent with the determination presented in Sec. IV A. The differences with the results of Ref.[96] can not be explained by the different data samples nor by the PDFs used. Instead, we find indications pointing towards the implementation of the deuteron model, TMC and HT corrections in the CJ15 fit.

[1] J. C. Collins, D. E. Soper, and G. F. Sterman, Adv. Ser. Direct. High Energy Phys. **5**, 1 (1989), arXiv:hep-ph/0409313 [hep-ph].

[2] S. Alekhin, J. Blümlein, S. Moch, and R. Placakyte, arXiv:1701.05838 [hep-ph].

---

[2] Both TMC and HT corrections are characterized by a strong Q^2 dependence, increasing their impact at lower Q^2 values.

[3] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, Eur. Phys. J. **C75**, 204 (2015), arXiv:1412.3989 [hep-ph].

[4] R. D. Ball *et al.* (NNPDF), JHEP **04**, 040 (2015), arXiv:1410.8849 [hep-ph].

[5] S. Dulat, T.-J. Hou, J. Gao, M. Guzzi, J. Huston, P. Nadolsky, J. Pumplin, C. Schmidt, D. Stump, and C. P. Yuan, Phys. Rev. **D93**, 033006 (2016), arXiv:1506.07443 [hep-ph].

[6] M. Arneodo, Phys. Rept. **240**, 301 (1994).

[7] P. R. Norton, Rept. Prog. Phys. **66**, 1253 (2003).

[8] D. F. Geesaman, K. Saito, and A. W. Thomas, Ann. Rev. Nucl. Part. Sci. **45**, 337 (1995).

[9] W. B. Atwood and G. B. West, Phys. Rev. **D7**, 773 (1973).

[10] P. V. Landshoff and J. C. Polkinghorne, Phys. Rev. **D18**, 153 (1978).

[11] D. Kusno and M. J. Moravcsik, Nucl. Phys. **B184**, 283 (1981).

[12] A. Bodek and J. L. Ritchie, Phys. Rev. **D23**, 1070 (1981).

[13] G. V. Dunne and A. W. Thomas, Nucl. Phys. **A455**, 701 (1986).

[14] L. L. Frankfurt and M. I. Strikman, Phys. Rept. **160**, 235 (1988).

[15] L. P. Kaptar and A. Yu. Umnikov, Phys. Lett. **B259**, 155 (1991).

[16] K. Nakano and S. S. M. Wong, Nucl. Phys. **A530**, 555 (1991).

[17] L. N. Epele, H. Fanchiotti, C. A. Garcia Canal, and R. Sassot, Phys. Lett. **B287**, 247 (1992).

[18] M. A. Braun and M. V. Tokarev, Phys. Lett. **B320**, 381 (1994).

[19] J. Gomez *et al.*, Phys. Rev. **D49**, 4348 (1994).

[20] W. Melnitchouk, A. W. Schreiber, and A. W. Thomas, Phys. Lett. **B335**, 11 (1994), arXiv:nucl-th/9407007 [nucl-th].

[21] F. Abe *et al.* (CDF), Phys. Rev. Lett. **74**, 850 (1995), arXiv:hep-ex/9501008 [hep-ex].

[22] V. V. Burov and A. V. Molochkov, Nucl. Phys. **A637**, 31 (1998).

[23] S. I. Alekhin, S. A. Kulagin, and S. Liuti, Phys. Rev. **D69**, 114009 (2004), arXiv:hep-ph/0304210 [hep-ph].

[24] V. V. Burov, A. V. Molochkov, G. I. Smirnov, and H. Toki, Phys. Lett. **B587**, 175 (2004), arXiv:hep-ph/0311191 [hep-ph].

[25] S. A. Kulagin and R. Petti, Nucl. Phys. **A765**, 126 (2006), arXiv:hep-ph/0412425 [hep-ph].

[26] S. A. Kulagin and W. Melnitchouk, Phys. Rev. **C77**, 015210 (2008), arXiv:0710.1101 [nucl-th].

[27] J. Arrington, F. Coester, R. J. Holt, and T.-S. H. Lee, J. Phys. **G36**, 025005 (2009), arXiv:0805.3116 [nucl-th].

[28] L. B. Weinstein, E. Piasetzky, D. W. Higinbotham, J. Gomez, O. Hen, and R. Shneor, Phys. Rev. Lett. **106**, 052301 (2011), arXiv:1009.5666 [hep-ph].

[29] J. Arrington, J. G. Rubin, and W. Melnitchouk, Phys. Rev. Lett. **108**, 252001 (2012), arXiv:1110.3362 [hep-ph].

[30] K. A. Griffioen *et al.*, Phys. Rev. **C92**, 015211 (2015), arXiv:1506.00871 [hep-ph].

[31] A. Accardi, W. Melnitchouk, J. F. Owens, M. E. Christy, C. E. Keppel, L. Zhu, and J. G. Morfin, Phys. Rev. **D84**, 014008 (2011), arXiv:1102.3686 [hep-ph].

[32] S. A. Kulagin and R. Petti, Phys. Rev. **D76**, 094023 (2007), arXiv:hep-ph/0703033 [hep-ph].

[33] S. A. Kulagin and R. Petti, Phys. Rev. **C82**, 054614 (2010), arXiv:1004.3062 [hep-ph].

[34] S. A. Kulagin and R. Petti, Phys. Rev. **C90**, 045204 (2014), arXiv:1405.2529 [hep-ph].

[35] P. Ru, S. A. Kulagin, R. Petti, and B.-W. Zhang, Phys. Rev. **D94**, 113013 (2016), arXiv:1608.06835 [nucl-th].

[36] M. Lacombe, B. Loiseau, J. M. Richard, R. Vinh Mau, J. Cote, P. Pires, and R. De Tourreil, Phys. Rev. **C21**, 861 (1980).

[37] R. Machleidt, Phys. Rev. **C63**, 024001 (2001), arXiv:nucl-th/0006014 [nucl-th].

[38] S. Veerasamy and W. N. Polyzou, Phys. Rev. **C84**, 034003 (2011), arXiv:1106.1934 [nucl-th].

[39] F. Gross and A. Stadler, Phys. Rev. **C78**, 014005 (2008), arXiv:0802.1552 [nucl-th].

[40] F. Gross and A. Stadler, Phys. Rev. **C82**, 034004 (2010), arXiv:1007.0778 [nucl-th].

[41] H. Georgi and H. D. Politzer, Phys. Rev. **D14**, 1829 (1976).

[42] S. Alekhin, S. A. Kulagin, and R. Petti, AIP Conf. Proc. **967**, 215 (2007), arXiv:0710.0124 [hep-ph].

[43] S. P. Malace *et al.* (Jefferson Lab E00-115), Phys. Rev. **C80**, 035207 (2009), arXiv:0905.2374 [nucl-ex].

[44] S. A. Kulagin, G. Piller, and W. Weise, Phys. Rev. **C50**, 1154 (1994), arXiv:nucl-th/9402015 [nucl-th].

[45] S. V. Akulinichev, S. A. Kulagin, and G. M. Vagradov, Phys. Lett. **B158**, 485 (1985).

[46] S. A. Kulagin, Nucl. Phys. **A500**, 653 (1989).

[47] K. Ackerstaff *et al.* (HERMES), Phys. Lett. **B475**, 386 (2000), [Erratum: Phys. Lett. B567, 339 (2003)], arXiv:hep-ex/9910071 [hep-ex].

[48] J. Seely *et al.*, Phys. Rev. Lett. **103**, 202301 (2009), arXiv:0904.4448 [nucl-ex].

[49] H. Abramowicz *et al.* (ZEUS, H1), Eur. Phys. J. **C75**, 580 (2015), arXiv:1506.06042 [hep-ex].

[50] H. Abramowicz *et al.* (ZEUS, H1), Eur. Phys. J. **C73**, 2311 (2013), arXiv:1211.1182 [hep-ex].

[51] F. D. Aaron *et al.* (H1), Eur. Phys. J. **C65**, 89 (2010), arXiv:0907.2643 [hep-ex].

[52] H. Abramowicz *et al.* (ZEUS), JHEP **09**, 127 (2014), arXiv:1405.6915 [hep-ex].

[53] A. C. Benvenuti *et al.* (BCDMS), Phys. Lett. **B223**, 485 (1989).

[54] A. C. Benvenuti *et al.* (BCDMS), Phys. Lett. **B237**, 592 (1990).

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

[56] A. Bodek *et al.*, Phys. Rev. **D20**, 1471 (1979).

[57] M. D. Mestayer, W. B. Atwood, E. D. Bloom, R. L. Cottrell, H. C. DeStaebler, C. Y. Prescott, L. S. Rochester, S. Stein, R. E. Taylor, and D. Trines, Phys. Rev. **D27**, 285 (1983).

[58] S. Dasu *et al.*, Phys. Rev. **D49**, 5641 (1994).

[59] O. Samoylov *et al.* (NOMAD), Nucl. Phys. **B876**, 339 (2013), arXiv:1308.4750 [hep-ex].

[60] A. Kayis-Topaksu *et al.*, New J. Phys. **13**, 093002 (2011), arXiv:1107.0613 [hep-ex].

[61] M. Goncharov *et al.* (NuTeV), Phys. Rev. **D64**, 112006 (2001), arXiv:hep-ex/0102049 [hep-ex].

[62] R. S. Towell *et al.* (NuSea), Phys. Rev. **D64**, 052002 (2001), arXiv:hep-ex/0103030 [hep-ex].

[63] G. Moreno *et al.*, Phys. Rev. **D43**, 2815 (1991).

[64] V. M. Abazov *et al.* (D0), Phys. Rev. **D88**, 091102 (2013), arXiv:1309.2591 [hep-ex].

[65] V. M. Abazov *et al.* (D0), Phys. Rev. **D91**, 032007 (2015), [Erratum: Phys. Rev. D91, 079901 (2015)], arXiv:1412.2862 [hep-ex].

[66] G. Aad *et al.* (ATLAS), Phys. Rev. **D85**, 072004 (2012), arXiv:1109.5141 [hep-ex].

[67] G. Aad *et al.* (ATLAS), Phys. Lett. **B759**, 601 (2016), arXiv:1603.09222 [hep-ex].

[68] S. Chatrchyan *et al.* (CMS), Phys. Rev. **D90**, 032004 (2014), arXiv:1312.6283 [hep-ex].

[69] V. Khachatryan *et al.* (CMS), Eur. Phys. J. **C76**, 469 (2016), arXiv:1603.01803 [hep-ex].

[70] R. Aaij *et al.* (LHCb), JHEP **08**, 039 (2015), arXiv:1505.07024 [hep-ex].

[71] R. Aaij *et al.* (LHCb), JHEP **01**, 155 (2016), arXiv:1511.08039 [hep-ex].

[72] R. Aaij *et al.* (LHCb), JHEP **05**, 109 (2015), arXiv:1503.00963 [hep-ex].

[73] S. Moch, J. A. M. Vermaseren, and A. Vogt, Nucl. Phys. **B688**, 101 (2004), arXiv:hep-ph/0403192 [hep-ph].

[74] A. Vogt, S. Moch, and J. A. M. Vermaseren, Nucl. Phys. **B691**, 129 (2004), arXiv:hep-ph/0404111 [hep-ph].

[75] W. L. van Neerven and E. B. Zijlstra, Phys. Lett. **B272**, 127 (1991).

[76] E. B. Zijlstra and W. L. van Neerven, Phys. Lett. **B273**, 476 (1991).

[77] E. B. Zijlstra and W. L. van Neerven, Nucl. Phys. **B383**, 525 (1992).

[78] E. B. Zijlstra and W. L. van Neerven, Phys. Lett. **B297**, 377 (1992).

[79] S. Moch and J. A. M. Vermaseren, Nucl. Phys. **B573**, 853 (2000), arXiv:hep-ph/9912355 [hep-ph].

[80] S. Moch, J. A. M. Vermaseren, and A. Vogt, Phys. Lett. **B606**, 123 (2005), arXiv:hep-ph/0411112 [hep-ph].

[81] J. A. M. Vermaseren, A. Vogt, and S. Moch, Nucl. Phys. **B724**, 3 (2005), arXiv:hep-ph/0504242 [hep-ph].

[82] R. Hamberg, W. L. van Neerven, and T. Matsuura, Nucl. Phys. **B359**, 343 (1991), [Erratum: Nucl. Phys. B644, 403 (2002)].

[83] R. V. Harlander and W. B. Kilgore, Phys. Rev. Lett. **88**, 201801 (2002), arXiv:hep-ph/0201206 [hep-ph].

[84] C. Anastasiou, L. J. Dixon, K. Melnikov, and F. Petriello, Phys. Rev. Lett. **91**, 182002 (2003), arXiv:hep-ph/0306192 [hep-ph].

[85] C. Anastasiou, L. J. Dixon, K. Melnikov, and F. Petriello, Phys. Rev. **D69**, 094008 (2004), arXiv:hep-ph/0312266 [hep-ph].

[86] S. Catani, L. Cieri, G. Ferrera, D. de Florian, and M. Grazzini, Phys. Rev. Lett. **103**, 082001 (2009), arXiv:0903.2120 [hep-ph].

[87] S. Alekhin and S. Moch, Phys. Lett. **B699**, 345 (2011), arXiv:1011.5790 [hep-ph].

[88] T. Gottschalk, Phys. Rev. **D23**, 56 (1981).

[89] M. Gluck, S. Kretzer, and E. Reya, Phys. Lett. **B380**, 171 (1996), [Erratum: Phys. Lett. B405, 391 (1997)], arXiv:hep-ph/9603304 [hep-ph].

[90] E. Laenen, S. Riemersma, J. Smith, and W. L. van Neerven, Nucl. Phys. **B392**, 162 (1993).

[91] W. B. Atwood, E. D. Bloom, R. L. Cottrell, H. C. DeStaebler, M. D. Mestayer, C. Y. Prescott, L. S. Rochester, S. Stein, R. E. Taylor, and D. Trines, Phys. Lett. **B64**, 479 (1976).

[92] S. Tkachenko *et al.* (CLAS), Phys. Rev. **C89**, 045206 (2014), [Addendum: Phys. Rev.C90,059901(2014)], arXiv:1402.2477 [nucl-ex].

[93] Alekhin, S. and Blümlein, J. and Caminadac, L. and Lipka, K. and Lohwasser, K. and Moch, S. and Petti, R. and Placakyte, R., Phys. Rev. **D91**, 094002 (2015), arXiv:1404.6469 [hep-ph].

[94] S. Alekhin, J. Blümlein, S. Moch, and R. Placakyte, Phys. Rev. **D94**, 114038 (2016), arXiv:1508.07923 [hep-ph].

[95] F. E. Close, R. L. Jaffe, R. G. Roberts, and G. G. Ross, Phys. Rev. **D31**, 1004 (1985).

[96] A. Accardi, L. T. Brady, W. Melnitchouk, J. F. Owens, and N. Sato, Phys. Rev. **D93**, 114017 (2016), arXiv:1602.03154 [hep-ph].

[97] A. Accardi, W. Melnitchouk, and J. Owens, Private communication.

[98] V. Khachatryan *et al.* (CMS), Phys. Lett. **B750**, 565 (2015), arXiv:1503.05825 [nucl-ex].

[99] V. Khachatryan *et al.* (CMS), Phys. Lett. **B759**, 36 (2016), arXiv:1512.06461 [hep-ex].

[100] G. Aad *et al.* (ATLAS), Phys. Rev. **C92**, 044915 (2015), arXiv:1507.06232 [hep-ex].

[101] R. Aaij *et al.* (LHCb), JHEP **09**, 030 (2014), arXiv:1406.2885 [hep-ex].

[102] K. Senosi (ALICE), *Proceedings, 53rd International Winter Meeting on Nuclear Physics (Bormio 2015): Bormio, Italy, January 26-30, 2015*, PoS **Bormio2015**, 042 (2015), arXiv:1511.06398 [hep-ex].

[103] G. Aad *et al.* (ATLAS), Phys. Rev. Lett. **110**, 022301 (2013), arXiv:1210.6486 [hep-ex].

[104] S. Chatrchyan *et al.* (CMS), JHEP **03**, 022 (2015), arXiv:1410.4825 [nucl-ex].

[105] S. Bueltmann *et al.* (BONuS), (2012), JLAB-PR12-06-113.

[106] R. Acciarri *et al.* (DUNE), (2015), arXiv:1512.06148 [physics.ins-det].

[107] A. Accardi *et al.*, Eur. Phys. J. **A52**, 268 (2016), arXiv:1212.1701 [nucl-ex].

[108] V. M. Abazov *et al.* (D0), Phys. Rev. Lett. **112**, 151803 (2014), [Erratum: Phys. Rev. Lett.114,no.4,049901(2015)], arXiv:1312.2895 [hep-ex].

[109] Buckley, Andy and Ferrando, James and Lloyd, Stephen and Nordström, Karl and Page, Ben and Rüfenacht, Martin and Schönherr, Marek and Watt, Graeme, Eur. Phys. J. **C75**, 132 (2015), arXiv:1412.7420 [hep-ph].

[110] Y. Li and F. Petriello, Phys. Rev. **D86**, 094034 (2012), arXiv:1208.5967 [hep-ph].

[111] R. Gavin, Y. Li, F. Petriello, and S. Quackenbush, Comput. Phys. Commun. **184**, 208 (2013), arXiv:1201.5896 [hep-ph].