Eur. Phys. J. C (2023) 83:829
https://doi.org/10.1140/epjc/s10052-023-11999-6

THE EUROPEAN PHYSICAL JOURNAL C

Regular Article - Theoretical Physics

Impact of SeaQuest data on PDF fits at large x

S. Alekhin[1,a], **M. V. Garzelli**[1,b], **S. Kulagin**[2,c], **S.-O. Moch**[1,d]

[1] II. Institut für Theoretische Physik, Universität Hamburg, Luruper Chaussee 149, 22761 Hamburg, Germany
[2] Institute for Nuclear Research of the Russian Academy of Sciences, Moscow 117312, Russia

Received: 13 June 2023 / Accepted: 21 August 2023 / Published online: 20 September 2023

Abstract We evaluate the impact of recent SeaQuest (FNAL-E906 experiment) data on dimuon production in proton-deuteron and proton-proton collisions on parton distribution functions (PDFs). We find these data in a good agreement with the QCD predictions based on PDFs fitted to the Tevatron and LHC data on forward production of W and Z bosons. As a basis for this study we use the ABMP16 PDF fits and show that they turn out to be compatible with the SeaQuest data, and that these data have constraining power, allowing to reduce the uncertainties on the isospin asymmetry of the light-sea-quark distribution at large longitudinal momentum fraction x. We discuss the nuclear corrections needed to describe the deuteron and show that they affect the theoretical description of the proton-deuteron Drell–Yan cross section at the level of $\mathcal{O}(0.5-1)\%$. We also comment on the compatibility of the SeaQuest results with other state-of-the-art PDF fits and show that these data are in clear disagreement with models proposing an SU(3)-flavor symmetric quark sea. Finally, we perform a comparison between the second Mellin moments of the light-quark PDFs and recent results from various lattice QCD computations, which demonstrates good compatibility, albeit limited by the uncertainties inherent in current lattice QCD simulations.

1 Introduction

The knowledge of parton distribution functions (PDFs) at large values of longitudinal momentum fraction x is one of the most urgent open questions [1,2] concerning proton and nuclear structure to which not only theoretical but even experimental efforts are going to be dedicated in the future. While in the long term data from the Electron Ion Collider (EIC) [3,4] are expected to play a very important constraining role, as also emphasized in the Snowmass 2021 EIC-dedicated whitepaper [5], in the near future further experiments might also offer promising opportunities. Among those at the Large Hadron Collider (LHC), we mention here the fixed-target (FT) configurations exploiting one of the LHC beams [6], a possibility already realized by complementing the LHCb detector with the SMOG and SMOG2 apparata [7,8], and also conceptually studied, although not realized, by the ALICE collaboration with the ALICE-FT experiment [9], as well as perspective projects still under discussion, like the Forward Physics Facility [10,11]. In particular, the LHCb + SMOG system has already delivered the first data using proton and Pb beams impinging on gaseous nuclei like ^{4}He, ^{20}Ne and ^{40}Ar, at different nucleon-nucleon center-of-mass energies $\sqrt{s_{NN}} \sim \mathcal{O}(50-100)$ GeV, corresponding to various past LHC runs. The LHCb + SMOG2 system, active during Run 3 and 4, can make use of even lighter gases, like deuterium ^{2}H as well as hydrogen, with increased statistics.[1] These experiments allow to probe the longitudinal momentum fraction interval $0.1 < x < 1$ for target partons, on which the constraints from the sets of HERA data [12] which traditionally form the backbone of PDF fits, are quite loose and mostly indirect.[2]

For the time being, constraints on PDFs at large x are imposed by legacy measurements from inclusive deep-inelastic scattering (DIS) experiments at fixed-targets (SLAC, BCDMS, NMC, etc.), semi-inclusive DIS experiments using ν beams and capable of measuring heavy-quark production in DIS (CCFR, NuTeV, CHORUS, NOMAD, etc.) and fixed-

[a] e-mail: sergey.alekhin@desy.de
[b] e-mail: maria.vittoria.garzelli@desy.de (corresponding author)
[c] e-mail: sergey.kulagin@gmail.com
[d] e-mail: sven-olaf.moch@desy.de

[1] The data released so far concern open and hidden charm production. We expect that in the future even data on Drell–Yan production will become available.

[2] HERA has also delivered some sets of experimental cross-section data which could constrain PDFs at large x, up to ~ 1, see e.g. Ref. [13], but these data have not been used in most of the PDF fits.

target Drell–Yan (DY) experiments (CERN-NA51, FNAL-E605, FNAL-E866, etc.), complemented by measurements of cross sections for DY (+ jets) production and other specific processes in the main detectors at the Tevatron and the LHC in the standard collider-mode configuration (for an overview, see e.g. Ref. [1] and references therein).

The valence quark distributions are constrained by DIS HERA data, up to $x < 0.1$, and in fixed-target experiments, up to $x \sim 1$. A large-x and relatively low-Q domain is also probed at JLab [14]. The DY data from the Tevatron and the LHC (both inclusive cross sections and charge asymmetries) as well as from fixed-target experiments have also been used to probe up (u) and down (d) quark distributions and their differences (isospin asymmetries). Single-top quark production data have allowed to probe the u/d ratio at $x \sim 0.1$, where $u = u_{\rm val} + \bar{u}$ and $d = d_{\rm val} + \bar{d}$, notwithstanding the big systematic uncertainties still accompanying the experimental cross sections for this channel of top-quark production [15]. The (anti)strange sea quark distributions (s, $\bar{s}$) has been constrained by DY (+ jets) LHC data and older (anti)-neutrino-nuclear DIS data, with large uncertainties [1,16,17], and improving their determination remains one of the pressing issues in PDF analysis. The $s(x) - \bar{s}(x)$ asymmetry [18] can be constrained by semi-inclusive DIS data on dimuon production distinguishing neutrino and antineutrino beams (as discussed e.g. in Refs. [19] and [20]), by $W^+ + \bar{c}$ and $W^- + c$ data at the LHC [21] and by future DIS experiments using separate beams of neutrinos and anti-neutrinos (e.g. at the Forward Physics Facility). The up and down sea quark distributions are constrained by DY data. Finally, the gluon PDF at large-x is mostly constrained by measurements of heavy-quark and jet production at the LHC [22].

Recently, the SeaQuest collaboration (FNAL-E906 experiment) has released fixed-target data on dimuon production on ^{2}H and proton targets through DY, which allow to constrain the difference between down and up sea quarks, i.e. $\bar{d}(x) - \bar{u}(x)$ as the crucial ingredient in the violation of the Gottfried sum rule [23], and the $\bar{d}(x)/\bar{u}(x)$ ratio [24]. This experiment can be considered as a continuation of previous ones, FNAL-E866 [25] (NuSea) and FNAL-E605 [26], lowering the center-of-mass energy $\sqrt{s}$ and extending the kinematic coverage in x towards large x values.[3]

The new experimental results motivate the present study, where we focus on the light-quark distributions, within the framework of collinear factorization in QCD and with particular emphasis on the sea antiquark case. General reviews on the flavour structure of the nucleon sea, including a broad range of hadron models, have been provided, e.g., in Refs. [29,30]. In Sect. 2, we show the impact of the SeaQuest results on the ABMP16 fits, considering both versions, at next-to-leading order (NLO) and at next-to-NLO (NNLO) in perturbative QCD, published in Refs. [31] and [32], respectively. This leads to new PDF fits, dubbed as ABMP16 + SeaQuest NLO and NNLO, performed ab-initio using the same statistical methodology and inputs as for the ABMP16 fits, plus the most recent SeaQuest data. In Sect. 3 we comment on the compatibility of other state-of-the-art PDF fits with these data and in Sect. 4 we discuss nuclear corrections. In Sect. 5 we compare the second moments of the light-flavor quark PDFs with recent lattice QCD results. Our conclusions are delivered in Sect. 6. Specific comparisons of our theory predictions with NuSea data and discussion of the compatibility with SeaQuest data are presented in the Appendix.

2 Constraining power of the SeaQuest data on the ABMP16 NLO and NNLO PDF fits

The study extends the ABMP16 PDF fits (NLO and NNLO), which have used the combined data from HERA for inclusive DIS, data from the fixed-target experiments NOMAD and CHORUS for neutrino-induced DIS, as well as data from Tevatron and the LHC for the DY process and the hadroproduction of single-top and top-quark pairs. The ABMP16 approach uses a fixed-flavor-number scheme for $n_f = 3, 4, 5$ and simultaneously determines the PDFs, the value of the strong coupling $\alpha_s(M_z)$ and all masses of heavy quarks, fully preserving the correlations among these quantities.

For illustrative purpose, we summarize in Fig. 1 the (x_1, x_2) coverage of most of the DY data used in constraining the up and down sea quark distributions at large x in these fits[4], together with the (x_1, x_2) coverage of the recently released SeaQuest data. The variables x_1 and x_2 represent the momentum fractions carried by the incident (anti)quarks in beam and target, respectively, which roughly characterize the region of x probed by a particular experiment. Since x_1 and x_2 are not observables and cannot be measured, we detail here how we reconstruct them, assuming leading order (LO) kinematics. For the SeaQuest experiment $x_{1,2}$ are computed as follows:

$$x_{1,2} = \frac{P_{1,2} \cdot Q}{P_{1,2} \cdot P}, \tag{1}$$

where Q is the four-momentum of the virtual photon from the quark–antiquark annihilation in the non-resonant produc-

[3] Another past experiment providing insights on isospin symmetry breaking through ratios of pd and pp DY cross sections, following the idea of Ref. [27] and using a CERN-SPS proton beam of 450 GeV/c, was NA51 [28]. However, they gave results in the form of only one data point around $x \sim 0.18$. Their results are compatible within experimental uncertainties with those of E866 that covers a wider x range.

[4] We leave out the LHCb $pp \to W^\pm + X \to l^\pm + \overset{(-)}{\nu} + X$ production data [33,34], whose kinematical coverage is similar to the one from LHCb $pp \to Z + X \to l^+l^- + X$ data shown in the plot.

tion process, $P_{1,2}$ are the four-momenta of the projectile and target hadron, respectively, and $P = P_1 + P_2$. Considering $\gamma^* \to \mu^+\mu^-$ decays, the average values for $x_{1,2}$ in the bins of the muon-pair average Feynman variable $\langle x_F \rangle$ are reported in Ref. [24]. These values are plotted in Fig. 1 in comparison with the kinematics of other DY data included in the ABMP16 fits. In particular, for the E605 Fermilab fixed-target data [26], given in the form of a double differential distribution in $\sqrt{\tau} = M/\sqrt{s}$ and y, where M and y are the invariant mass and rapidity of the $\mu^+\mu^-$-pair, respectively, and $\sqrt{s}$ is the collision center-of-mass energy, the values of $x_{1,2}$ are computed according to the relation

$$x_{1,2} = \sqrt{\tau} e^{\pm y} . \tag{2}$$

For the E866 experiment [25] the same relation is employed. However, since the muon-pair rapidity is not tabulated in Ref. [25], it is computed from the muon-pair x_F and transverse momentum p_T using basic definition as follows:

$$y = \frac{1}{2} \ln \left(\frac{E + p_L}{E - p_L} \right) , \tag{3}$$

where $p_L = x_F \, p_{L,\,\mathrm{max}}$ and $E = \sqrt{p_L^2 + p_T^2 + M^2}$ are the muon-pair longitudinal momentum and energy, respectively, in the center-of-mass frame of the colliding hadrons, with $p_{L,max}$ the maximum longitudinal momentum of the muon-pair, depending on $\sqrt{s}$ according to the formula $p_{L,max} = \sqrt{s}\,(1 - M^2/s)/2$.

The approach of Eq. (2) is also used for the LHCb data on Z-boson production [35–37] released in the form of lepton-pair pseudorapidity distributions.[5] The data on W-boson production evidently probe the same kinematics. However, the use of Eq. (2) is impossible in this case due to the neutrino escaping detection. Therefore, for W-boson production in the D0 experiment [38,39], we use the following approximate estimate:

$$x_{1,2} = \frac{M_W}{\sqrt{s}} e^{\pm y_l} , \tag{4}$$

where M_W is the W-boson pole mass and y_l is the lepton rapidity.

Both DY data at the Tevatron and the LHC and DY data in fixed-target experiments play a role in constraining the sea quark PDFs at large x. They allow to reach similarly large-x values, although in the case of fixed-target data, relatively large-x partons from both the projectile and the target participate in the same hard interaction, whereas in the case of the LHC, a large-x parton is typically probed simultaneously

[5] For the lepton energies of LHCb the numerical difference between pseudorapidity and rapidity is negligible.

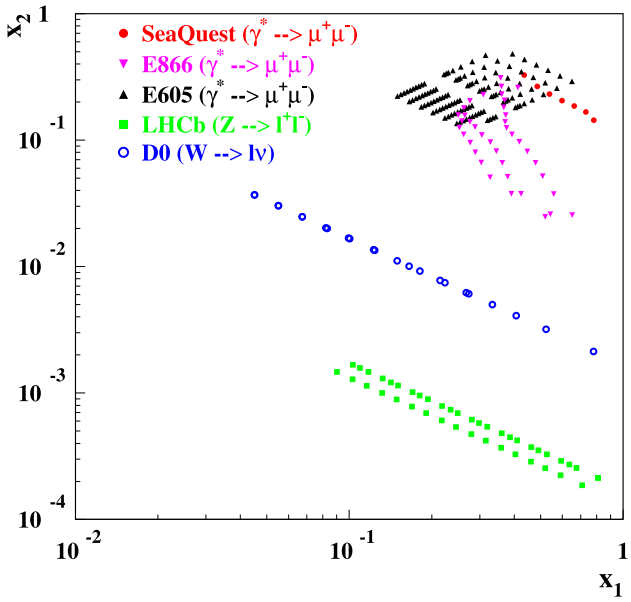

Fig. 1 The (x_1, x_2) coverage for the SeaQuest experiment [24] (full circles), with $x_{1,2}$ given by Eq. (1), in comparison to the coverages from DY data of other experiments used in the ABMP16 PDF fits (down-oriented triangles: E866 Fermilab fixed-target experiment [25]; up-oriented triangles: E605 [26] with $x_{1,2}$ computed from the lepton-pair rapidity using Eq. (2) for both the data sets; squares: LHCb, the LHC experiment [35–37] with $x_{1,2}$ computed from the Z-boson rapidity using Eq. (2); open circles: D0, the Tevatron collider experiment [38,39], with $x_{1,2}$ estimated from the charged-lepton rapidity using Eq. (4))

with a low-x one, as exemplified in the (x_1, x_2) correlation in Fig. 1. The correlation is quite evident for the LHC data and is related to the exchange of heavy bosons in the DY process. In the case of LHC, the largest x_1 values are probed by the LHCb detector with data at large positive rapidity, which covers the interval $2 < y < 4.5$. On the other hand, the fixed-target experiments E866 and E605, which have much lower center-of-mass energies than the LHC, probe larger x_2 values and present a less evident (x_1, x_2) correlation, related to the exchange of a γ^* with a broad range of mass values in the DY process. In the case of SeaQuest, the (x_1, x_2) correlation is again evident, considering that the invariant mass of the observed γ^* decay products is fixed to approximately $M \sim 5$ GeV. SeaQuest covers x_2 values higher than the LHC due to the use of a beam with much lower center-of-mass energy ($\sqrt{s} = 15.1$ GeV). The x_2 region covered by SeaQuest extends up to ≤ 0.45. The E605 experiment has a coverage extending even up to slightly higher x_2 values. However, the E605 experiment used a Copper target, thus requiring an evaluation of nuclear corrections (see the end of Sect. 4). Having only one target material, they could not provide data on cross-section ratios, unlike SeaQuest that has both a deuteron and a proton target. Also, given that Copper is a heavy nucleus close to isoscalarity, the E605 data are much less sensitive to the isospin asymmetry effects, that we

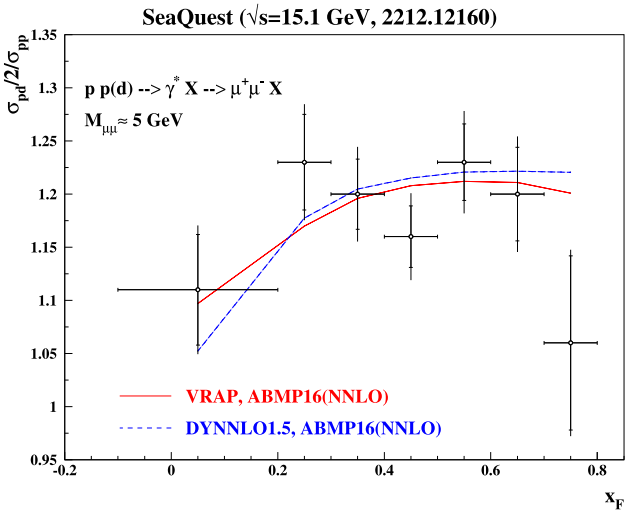

Fig. 2 The SeaQuest data [24] on the ratio of pd and pp DY distributions over Feynman variable x_F with respect to the NNLO predictions obtained using the code VRAP [40] (solid line) and DYNNLO [41] (dashes). For the DYNNLO predictions we have used a fixed $\langle M\rangle = 5$ GeV in all bins, to simplify the computation

investigate in this work. We explicitly verified the very small impact of E605 data, by removing them from our fits where they are included as default.

The green dots along two parallel lines in Fig. 1 refer to the cases of Z-boson production at the LHC at $\sqrt{s} = 7$ and 8 TeV, given that data at these center-of-mass energies were included in the ABMP16 PDF fits.

In order to compute predictions for the DY cross sections, we use the FEWZ2.1 code [42] for the collider cases and the VRAP code [40] for the fixed-target cases. In particular, the present analysis of SeaQuest data is based on the x_F-distribution, that was directly measured in the experiment and could also be computed to NNLO QCD accuracy using a Monte-Carlo code, like e.g. FEWZ or DYNNLO [41].[6] However, in the fit we employ the VRAP code, which is based on 2-dimensional integration that allows to greatly improve the code performance. To compute VRAP predictions for the SeaQuest data on the x_F-distribution we perform a mapping of x_F to the rapidity using the basic relation Eq. (2) and taking $P_L = \langle x_F\rangle\sqrt{s}/2(1-\langle M\rangle^2/s)$, $E = \sqrt{\langle M\rangle^2 + P_L^2 + \langle P_T\rangle^2}$, where s is the center-of-mass energy squared and $\langle x_F\rangle$, $\langle M\rangle$ and $\langle P_T\rangle$ are the averages of muon-pair Feynman variable x_F, invariant mass and transverse momentum, respectively, over the bins in x_F. $\langle M\rangle$ varies bin by bin, ranging from 4.9 Gev in the bin with largest $\langle x_F\rangle$ to 5.6 GeV in the one with smallest $\langle x_F\rangle$. These averaged quantities are all given in Ref. [24] for each bin in x_F. To validate such an approach, we compare its predictions with those obtained with the methodology used in Ref. [24], where the DYNNLO [41] code is employed, instead of VRAP, and the exact information concerning the transverse momentum and the invariant mass of the $\mu^+\mu^-$-pair on an event-by-event basis is considered to build the x_F distributions, instead of the average value of these quantities per x_F bin. We find that the difference is mostly well below the data uncertainties, cf. Fig. 2, where we compare predictions obtained with VRAP using the approximations outlined above with the DYNNLO predictions based on the exact values for P_T and M as input, as in Ref. [24], and applying their same cut $M > 4.5$ GeV. The latter suppresses the $\mu^+\mu^-$-background contribution from J/ψ and ψ' production and decay. We build the x_F distributions using Eq. (4) of Ref. [24]. From Fig. 2 it is evident that only in the smallest x_F bin the difference between VRAP and DYNNLO is comparable to the data uncertainty, an observation that might be related to the width of the bin, which is much larger for this bin, than for the other ones. Obviously, such a difference cannot have relevant impact on fit results. Therefore, considering that the approximated procedure with the use of VRAP allows for NNLO simulations much faster than the exact procedure using DYNNLO, and given that the results turn out to be very well compatible, we use VRAP for the analyses and all other plots presented in the rest of this work.

We observe that the corrections related to spectrometer acceptance as a function of x_1 and x_2 reported in Ref. [44] do not impact distributions depending on measured quantities, like e.g., x_F, that we consider in this work. On the other hand, their inclusion is relevant for the extraction of the $\bar{d} - \bar{u}$ asymmetry and the $\bar{d}/\bar{u}$ ratio from the SeaQuest data, as performed by the SeaQuest collaboration in approximated form as described in their papers [24,44].

The constraints from the SeaQuest experiment turn out to be compatible with those already imposed by collider data, as shown by the fact that the χ^2/NDP of the analyses including also the SeaQuest data does not change significantly with respect to the χ^2/NDP of the original ABMP16 analyses. Here NDP indicates the number of data points and the differences are well within the χ^2 statistical uncertainties, as shown in Table 1. The χ^2/NDP for the NNLO analyses turns out to be 1.18, slightly closer to 1 than the χ^2/NDP of the NLO analyses, which is equal to 1.20. We also observe that incorporating SeaQuest data in the fits has a negligible impact on the values of $\alpha_s(M_Z)$ and heavy-quark masses, extracted simultaneously to PDFs in all the fits considered in Table 1.

Separate χ^2 values for various data sets included in our NNLO QCD analyses are reported in Table 2. We have considered four variants: (I) the ABMP16 analysis, (II) the ABMP16 + SeaQuest analysis, as well as (III) an analysis, where we consider all data of (II), except the D0 DY data and (IV) an analysis, where we consider all data of (II), except the

[6] Note that NNLO predictions for the DY process obtained using non-local subtraction methods, may differ by power corrections, whose size varies depending on the experimental cuts on final-state leptons [43]. These corrections are, however, negligible in the context of this work.

Table 1 The total values of χ^2 obtained for the NLO and NNLO ABMP16 fits in comparison with the ones of the present analyses, including all data already considered in the ABMP16 fits plus SeaQuest data. See text for more detail

Fit	NDP	χ^2	
		NLO	NNLO
ABMP16	2861	3428.9	3377.6
Present analysis (ABMP16 + SeaQuest)	2868	3438.4	3384.7

LHCb DY data. We include variants (III) and (IV) due to the fact that in the past we have observed some tension between the D0 and LHCb DY data. By comparing (I) and (II), we find that, for each considered data set, the addition of SeaQuest data does not introduce significant modifications of the χ^2. Thus SeaQuest data are well compatible with both the LHCb and the D0 DY data. On the other hand, by comparing (II) and (III), we find that the elimination of the D0 DY datasets from the fit allows to improve the χ^2 of the analysis of the 7 TeV LHCb DY dataset by several units, beyond the statistical χ^2 uncertainty. Vice versa, the elimination of the LHCb DY datasets allows to improve the description of D0 data, as can be understood by comparing (II) and (IV). The role of the NuSea data, for which χ^2 values are also presented in Table 2 for the four variants of our fit, and their compatibility with other data in our fits is discussed in Appendix A.

The χ^2 values were computed accounting for statistical and systematic uncertainties of the SeaQuest data, assuming that systematic uncertainties are fully correlated bin-by-bin. Detailed information concerning correlations among the uncertainties characterizing the SeaQuest data are, however, not available. Therefore, we also consider a variant of the fit, where the systematic uncertainties are considered as fully uncorrelated. We have found that the χ^2 values related to the analysis of the SeaQuest data in both analyses are compatible within statistical fluctuations ($\chi^2_{corr.} = 7.3$ vs. $\chi^2_{uncorr.} = 5.9$, for NDP = 7). This implies that more details on the precise degree of bin-by-bin correlations of the systematic uncertainties in the SeaQuest data, when available, will not modify the main conclusions of our study.

Figure 3 shows the constraining power of the SeaQuest data on the $\bar{d}(x) - \bar{u}(x)$ difference, increasing towards large x values. At NLO, the uncertainty band of the analysis with SeaQuest data has a large overlap, but is not completely included within the band of the default ABMP16 analysis (not including these data). Additionally, for $0.1 < x < 0.2$, the band of the analysis with SeaQuest data turns out to be of the same size of the band of the one without these data. On the other hand, at NNLO, the uncertainty bands are in general smaller than at NLO and the one of the analysis with SeaQuest data is always included and smaller than the band of the analysis without SeaQuest data. These findings confirm that theory predictions at NNLO accuracy are in general more robust and consistent among each other than NLO ones, i.e., the theory description at NLO is still incomplete and hardly provides a simultaneous excellent description of all DY data, like is instead happening at NNLO. This is also reflected in the comparison of the χ^2/NDP values presented in Table 1. In any case, in Fig. 3, the constraining power of SeaQuest data is certainly evident for $x > 0.3$ for both the NLO and NNLO analyses. However, for large x values the difference between the distributions of $\bar{d}(x)$ and $\bar{u}(x)$ diminishes.

Analogous observations can be made when examining Fig. 4, whose left panel illustrates the variation of the ratio

Table 2 The values of χ^2 obtained for the data sets probing the large-x PDFs, which are included in various analyses (column I: NNLO ABMP16 PDF fit [32], column II: present analysis, column III: a variant of present analysis with D0 DY data excluded, column IV: a variant of present analysis with LHCb DY data excluded)

Experiment	Process	$\sqrt{s}$ (TeV)	References	NDP	χ^2			
					I	II	III	IV
SeaQuest	$pp \to \gamma^* X \to \mu^+\mu^- X$ $pd \to \gamma^* X \to \mu^+\mu^- X$	0.0151	[24]	7	–	7.3	8.1	7.6
NuSea	$pp \to \gamma^* X \to \mu^+\mu^- X$ $pd \to \gamma^* X \to \mu^+\mu^- X$	0.0388	[25]	39	52.8	54.3	52.5	53.0
D0	$\bar{p}p \to W^\pm X \to \mu^{\pm(-)}\nu X$	1.96	[45]	10	17.6	17.6	–	14.5
	$\bar{p}p \to W^\pm X \to e^{\pm(-)}\nu X$	1.96	[39]	13	19.0	19.0	–	15.9
LHCb	$pp \to W^\pm X \to \mu^{\pm(-)}\nu X$ $pp \to ZX \to \mu^+\mu^- X$	7	[33]	31	45.1	43.9	35.0	–
	$pp \to W^\pm X \to \mu^{\pm(-)}\nu X$ $pp \to ZX \to \mu^+\mu^- X$	8	[34]	32	40.0	39.6	38.2	–
	$pp \to ZX \to e^+e^- X$	8	[46]	17	21.7	21.9	21.9	–

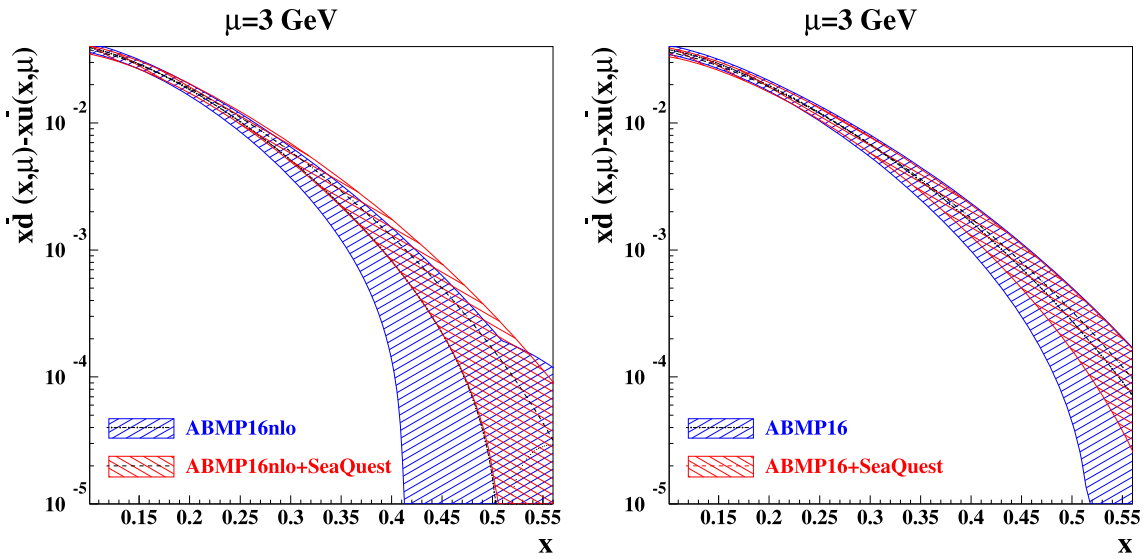

Fig. 3 The 1σ band for the $n_f = 3$-flavour isospin asymmetry of the sea distribution $x(\bar{d} - \bar{u})(x)$ at the scale μ =3 GeV obtained in the present analysis (central pdf: dashed line, uncertainty band: left-tilted hash) compared to the one of the ABMP16 fit (central pdf: dot-dashed line, uncertainty band: right-tilted hash). The left panel shows results of the NLO analysis, whereas the right panel refers to the NNLO one

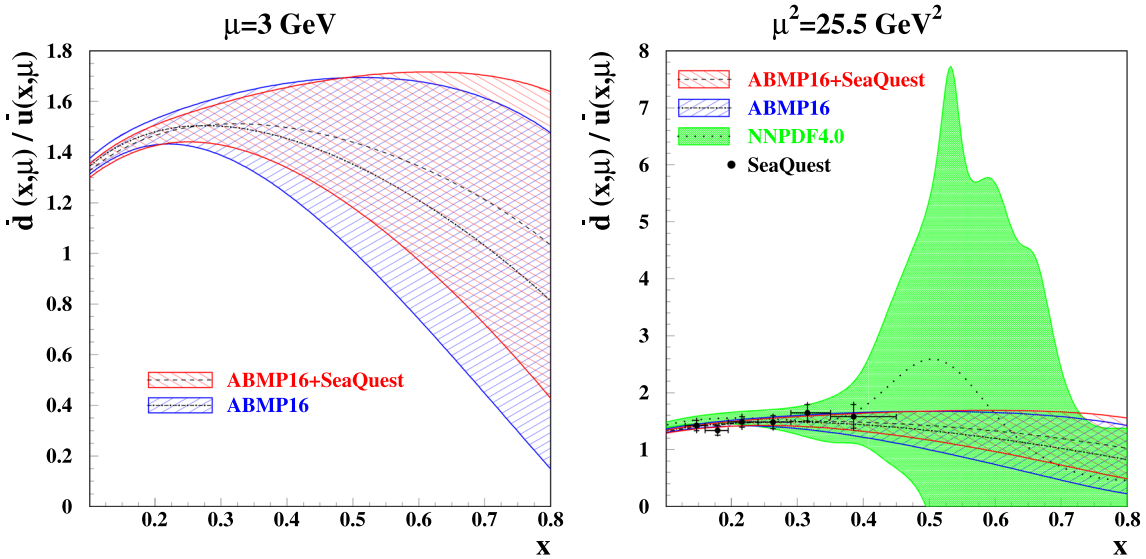

Fig. 4 Left panel: the 1σ band for the ratio of the $n_f = 3$-flavour sea distributions $\bar{d}/\bar{u}$ as a function of x at the scale $\mu = 3$ GeV obtained in the present NNLO analysis (central pdf: dashed line, uncertainty band: left-tilted hash) compared to the one of the ABMP16 fit (central pdf: dot-dashed line, uncertainty band: right-tilted hash). Right panel: same as in the left panel, but at the scale $\mu^2 = 25.5$ GeV^2 at which the SeaQuest collaboration extracted the $\bar{d}/\bar{u}$ ratio, that is also plotted. Also shown are the 1σ predictions with the NNPDF4.0 NNLO PDF fit (central pdf: dotted line, uncertainty band: solid), which has incorporated SeaQuest data

$\bar{d}(x)/\bar{u}(x)$ with respect to x for $\mu = 3$ GeV. The ratio is larger than unity within a large x interval, up to at least $x < 0.5$ - 0.6. At these x values, both the $x\bar{u}$ and $x\bar{d}$ sea distributions are tiny, of the order of 10^{-5}. The analysis incorporating SeaQuest data exhibits a high level of compatibility with the analysis that excludes them, and displays a smaller uncertainty band, especially for $x > 0.3$. This confirms the constraining role of the SeaQuest data. As shown in the right panel of Fig. 4, the results are also very well compatible with the $\bar{d}(x)/\bar{u}(x)$ ratio extracted by the SeaQuest collaboration at the scale $Q^2 = 25.5$ GeV^2, which is characteristic of the kinematics of the experiment, using as a starting point the experimentally measured cross-section ratio $\sigma_{pd}/(2\sigma_{pp})$ and Eqs. (8), (10) and (11) of Ref. [24]. Although

this extraction depends in principle on the PDFs used (the quoted SeaQuest values are those reported in Table 8 of Ref. [24], obtained using cross sections computed with the CT18 PDF fit as input of their Eq. (11)), this dependence is quite weak, i.e., it comes from subleading terms in Eq. (11) of Ref. [24], generating minor corrections to the leading result corresponding to the case $x_1 \sim x_2$. Therefore the extracted $\bar{d}/\bar{u}$ ratio can be considered as a robust quantity, as also already observed in Ref. [24].[7] We also note that the SeaQuest data cover target x values up to 0.45. The uncertainty band of the ABMP16 + SeaQuest PDFs remains small at even larger x values, which is a consequence of assumptions about the parameterization of these PDFs and their extrapolation to large x, performed under assumption of smoothness of the distributions. The same is true for the ABMP16 PDF fit. Only future experimental data in the large x region will be able to check the correctness of this extrapolated result shown here. Regardless, it is important to emphasize that the ABMP16 + SeaQuest fits rely on the identical PDF parameterization employed in the original ABMP16 fits. Remarkably, this parameterization already yielded a satisfactory fit to the new data, without necessitating any post-adjustments through the introduction of additional parameters. During the original ABMP16 fit, we employed a strategy that involved investigating the impact of various functional forms while minimizing the number of parameters used. Our aim was to avoid introducing any additional parameters that did not contribute significantly to an improved description of the data.

We also point out that the effect of SeaQuest data, when comparing the ABMP16 PDFs to the ABMP16 + SeaQuest PDFs, is not dramatic, because the ABMP16 fits already included the E866 data, capable of constraining the $\bar{d}/\bar{u}$ ratio up to slightly lower x values than SeaQuest. The main addition of SeaQuest has been to have provided reliable measurements in the interval $x \sim 0.24$ - 0.45, which have helped to further constrain PDFs with respect to the past.

3 Compatibility of SeaQuest data with other PDF fits

The compatibility of the SeaQuest data with a number of modern PDF fits is shown in Fig. 5. The SeaQuest data align well with the predictions based on the NNPDF4.0 fit [47], which is not surprising since the NNPDF collaboration incorporated these data into their fitting process. Nevertheless, the uncertainty range associated with this particular fit remains larger compared to our own uncertainty range, in contrast to the uncertainties accompanying the data. We argue that this behaviour can be ascribed to inefficiencies in the statistical estimators used in their analysis.

This issue seems to be confirmed also by the predictions on the $\bar{d}/\bar{u}$ ratio shown in Fig. 4 (right), where the constraining power of the SeaQuest data seems to be only partially reflected in the NNPDF4.0 uncertainties. This is particularly visible in the region $x \sim 0.3$ - 0.45, where the NNPDF4.0 uncertainties become large, although this region is still covered by SeaQuest data. The inclusion of the FNAL-E605 and the SeaQuest data in the NNPDF4.0 fit should have imposed additional constraints on the $\bar{d}/\bar{u}$ ratio at these specific values of x. Consequently, one would expect the size of their 1σ band to be smaller compared to the result from their fit. The use of a large fixed number of parameters in the parameterization of these PDFs might be responsible for a relatively large uncertainty in the x region where SeaQuest data are present. The shape of the spike region around $x \sim 0.5$ in Fig. 4 (right) seems to be driven by the step functions used in the parameterization of these PDFs. A careful study from the NNPDF collaboration on this issue is warranted, also to delineate it from the impact of other datasets used in their fit.

The large uncertainties of NNPDF4.0 at larger x values in Fig. 4 (right), on the other hand, can be attributed to the lack of data. The smaller uncertainty of the ABMP16 fits (in comparison to NNPDF4.0) in the very large x region, not covered by the SeaQuest data, is neither related to the use of looser W cuts ($W > 2$ GeV) on the invariant mass of the hadronic system

$$W^2 = M_P^2 + Q^2(1-x)/x\,, \tag{5}$$

in DIS data[8] (M_P is the proton mass), nor to the inclusion of higher-twist corrections in the fit. Moreover, it is important to note that these uncertainties cannot be considered highly informative. This is due to the fact that the uncertainty arises solely from the extrapolation beyond the region where data are available, relying on assumptions of smoothness, as already mentioned in the previous section. It is however true that the sum rules play a role in constraining the shape of PDFs there. We have checked that a shape with even more spikes and larger uncertainties for the $\bar{d}/\bar{u}$ ratio occurs in the case of the predecessor of NNPDF4.0 fit, i.e. the NNPDF3.1 PDF fit [48] (not shown in the plot), not including SeaQuest data.

These concerns about the NNPDF4.0 PDFs at large x are also manifest in unusual predictions for the forward-backward asymmetry A_{FB}^* in the invariant mass of the dilepton final state at the LHC, quite different from those of many others PDF fits particularly for large invariant masses [49,50]. The measurement of this quantity and its

[7] We refrain from comparisons with the $(\bar{d}-\bar{u})(x)$ values also reported in Table 8 of Ref. [24], because these values are indeed more sensitive to the PDF used in their extraction, depending on additional assumptions on the $\bar{u}(x)$ distribution.

[8] In fact electron DIS data are sensitive to the total $q = q_{\rm val} + \bar{q}$ distributions, which can be separated into the valence and sea components only by adding DY data and/or neutrino DIS data to PDF fits.

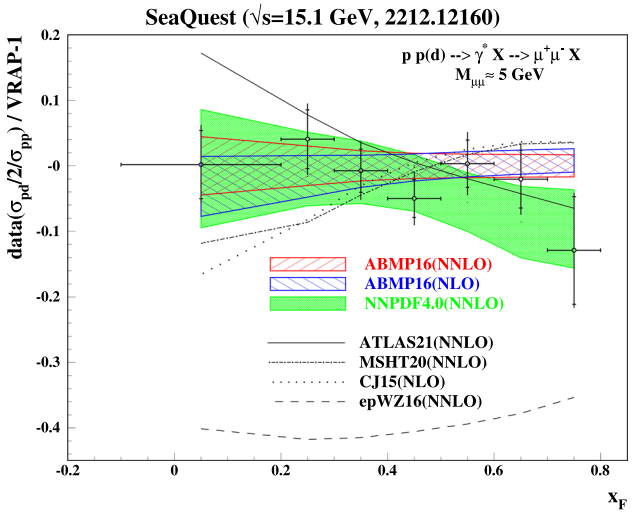

Fig. 5 The pulls for SeaQuest data [24] on the ratio of pd and pp DY distributions over x_F with respect to predictions obtained using the code VRAP [40] in combination with the NNLO ABMP16 PDFs. The 1σ band for prediction (right-tilted hash) is compared to the NLO ABMP16 [31] (left-tilted hash) and NNLO NNPDF4.0 [47] (shaded area) ones. The central values of predictions with other PDFs are shown for comparison (solid: NNLO ATLAS21 [53], long dashes: NNLO epWZ16 [54], suggesting the SU(3)-symmetric quark sea, dashed dots: NNLO MSHT20 [19], dots: NLO CJ15 [55])

comparison with theory predictions might be important for improved fits of large-x quark PDFs within the Standard Model (SM) and/or for discovering new physics associated to new gauge sectors beyond the SM, such as a heavy neutral Z'-boson, see, e.g. Refs. [51,52].

The SeaQuest data are also compatible with the CT18 fit [56] (not shown in Fig. 5), although the uncertainty of the latter looks particularly large, even due to the tolerance criterion used in this PDF fit ($\Delta\chi^2 = 100$ at 90% C.L. roughly corresponding to $\Delta\chi^2 \sim 30$ at 68% C.L., vs $\Delta\chi^2 = 1$ used in various other PDF fits adopting the Hessian approach, although not in all[9]), and this prevents any strong conclusion. The CT18 collaboration has investigated the impact of first SeaQuest data of Ref. [44] on their NNLO PDFs in Ref. [58] and they have compared their predictions even to the BNL STAR data on W-boson production [60]. Additionally, the CT18A variant of the fit, together with further variants incorporating lattice QCD data on the strangeness asymmetry distribution $s(x) - \bar{s}(x)$, have also been compared to first SeaQuest data of Ref. [44] in Ref. [20]. An advanced study aiming at separating the so-called connected and disconnected sea components, reflecting the topology of the quark lines in the four-point current–current correlator in the nucleon, under the CT18 parameterization, has led to the CT18CS fit [61], using as a basis the original CT18 data sets. The CT18CS PDFs have also been compared with the distributions extracted from the SeaQuest data of Ref. [44], and older E866 data of Ref. [25].

On the other hand, a comparison of the SeaQuest data with predictions obtained with the MSHT20 [19] and CJ15 [55] fits, both shown in Fig. 5, reveals that the $\bar{d}/\bar{u}$ ratio according to the latter has a trend compatible with the data only in part of the (x_1, x_2) range. The CJ collaboration has also investigated the impact of the first SeaQuest data of Ref. [44] plus the aforementioned STAR data, on the CJ15 PDFs in Ref. [59][10] and very recently proposed the new global PDF fit CJ22 in a follow-up paper [63], incorporating the SeaQuest data plus the aforementioned STAR data, including higher-twist effects and nucleon off-shell corrections. It would be interesting to study as well the modification of the MSHT20 fit, after inclusion of the SeaQuest data.

Finally, we remark that the behaviour of the $(\bar{d}/\bar{u})(x)$ ratio predicted by the ATLAS 2016 fit [54] turns out to be incompatible with the SeaQuest data, systematically underestimating the latter, pointing to issues in the whole technique to derive these PDFs and/or shortcomings during the fit. In particular, the comparison of the x_F distribution with the SeaQuest data in Fig. 5 confirms the point raised already in Ref. [16] that the assumptions concerning d-quark suppression with respect to the u-quarks in the ATLAS PDF parameterization adopted in that fit, using 14 parameters and now outdated, are problematic.[11] The considerations in Ref. [16] were based on the observation that these PDFs already exhibited disagreement with the E866 data, which were already accessible at that particular time. One should in any case not be surprised that this old PDF fit is not in agreement with SeaQuest data, considering that, by definition, it did not include typical non-ATLAS datasets constraining high-x PDFs. In turn, this lack of data required to make more constraining assumptions on the PDF form. Newer ATLAS PDF fits have added more ATLAS data, partially extending x coverage and allowing for more flexible parameterizations. However, we have verified that even the central PDF from a more recent ATLAS fit, ATLASepWZVjet20-EIG [66] (not shown in our plot), including 16 parameters and the W, Z/γ^* + jet data that are sensitive to partons at larger x's than the inclusive W, Z/γ^* data, turns out to be also incompatible with the SeaQuest data, overestimating the data up to several ten percent in the smallest x_F bin (corresponding to the largest x). On the other hand, the central PDF from the most recent publicly available ATLAS PDF fit, ATLASpdf21 [53], a fit

[9] MSHT, for instance, uses a dynamical tolerance procedure [57], while CT18 uses a combination of global and dynamic tolerance.

[10] Almost simultaneously, Ref. [62] has presented a global QCD analysis using these same data in the JAM Bayesian Monte Carlo framework.

[11] Also observe that in the CJ15 PDFs the d-quark content of the proton is parameterized in terms of the u-quark one, introducing a correlation that can affect results for the d/u ratios at large x's, as discussed in Refs. [64,65].

that includes 21 parameters, further data and also considers the role of scale uncertainties, largely overestimates the SeaQuest data in the first x_F bin, but is compatible with the latter in the other bins, i.e. for $0.2 < x_F < 0.8$, as shown in Fig. 5. The lack of agreement in the smallest x_F bin can be probably attributed to the fact that ATLAS does not have data constraining $\bar{u}(x)$ and $\bar{d}(x)$ for $x > 0.3$. On the other hand, the agreement visible at larger x_F, corresponding to $x < 0.3$, remarks the compatibility between SeaQuest and DY ATLAS and Tevatron data. In Ref. [53] the ATLAS collaboration has provided their own comparison of ATLASpdf21 $(\bar{d}/\bar{u})(x)$ ratio with that extracted by the NuSea and SeaQuest collaborations in Refs. [25,44]. Considering that the smallest x_F correspond to the largest x values, our results and conclusions on compatibility between fixed-target and collider DY datasets are compatible with their ones.

4 Impact of nuclear corrections

SeaQuest data discussed and used in previous sections have been collected for a deuteron target and for this reason the analysis should address the corresponding nuclear corrections. Here we discuss the effect of nuclear corrections on the DY cross section following Ref. [67]. This model addresses a number of mechanisms for nuclear corrections including the effect of nuclear momentum distribution (Fermi motion), nuclear binding, the off-shell modification of bound nucleon PDFs, as well as meson-exchange currents and nuclear shadowing corrections. For the kinematics of SeaQuest data the relevant corrections originate from nuclear momentum distribution, binding and off-shell effects on the PDFs. The deuteron PDFs $q_{i/d}$ of type $i = u, d, \ldots$ can be written as follows [67] (see also Appendix B of Ref. [64])

$$x q_{i/d}(x, Q^2) = \int d^3\boldsymbol{k}\, |\Psi_d(\boldsymbol{k})|^2\, (1 + k_z/M) x' \times \Big(q_{i/p}(x', Q^2, k^2) + q_{i/n}(x', Q^2, k^2) \Big), \tag{6}$$

where $q_{i/p(n)}$ is the corresponding proton (neutron) PDFs, the integration is performed over the nucleon momentum $\boldsymbol{k}$, $\Psi_d(\boldsymbol{k})$ is the deuteron wave function in the momentum space, which is normalized as $\int d^3\boldsymbol{k}\, |\Psi_d(\boldsymbol{k})|^2 = 1$, and M is the nucleon mass. We consider the deuteron in the rest frame and the z axis is chosen to be antiparallel to the momentum transfer. The four-momentum of the bound nucleon is $k = (M_d - \sqrt{M^2 + \boldsymbol{k}^2}, \boldsymbol{k})$, where M_d is the deuteron mass and $k^2 = k_0^2 - \boldsymbol{k}^2$ is the invariant mass squared (virtuality), while $x' = xM/(k_0 + k_z)$ is the Bjorken variable of the off-shell nucleon.

It is convenient to discuss the virtuality dependence of the nucleon PDFs in terms of the dimensionless variable $v = (k^2 - M^2)/M^2$. Since nuclei are weakly bound systems, the value of $|v|$ is small on average. For this reason the off-shell PDFs can be expanded in a power series in v about $v = 0$ [68]. Keeping the terms linear in v we have [69]

$$q(x, Q^2, k^2) = q(x, Q^2)[1 + \delta f(x, Q^2)\, v], \tag{7}$$

$$\delta f(x, Q^2) = \partial \ln q(x, Q^2, k^2)/\partial \ln k^2, \tag{8}$$

where the derivative is taken for $k^2 = M^2$. The function $\delta f(x, Q^2)$ measures the modification of the nucleon PDFs in the off-shell region. In Eq. (7), in order to simplify notations, we suppress the subscripts referring to the PDF type i. Also, we implicitly assume an average over the proton (p) and neutron (n), $q_i = (q_{i/p} + q_{i/n})/2$, since Eq. (6) for the deuteron depends only on this isoscalar PDF combination. Detailed studies of nuclear DIS, DY lepton-pair and W/Z boson production indicate that the data are consistent with an universal function $\delta f(x)$, independent of the parton type, and without significant scale and nucleon isospin dependencies [64,65,67,69–72].[12] In this work we use the results on the function $\delta f(x)$ from the recent analysis of some of us in Ref. [65].

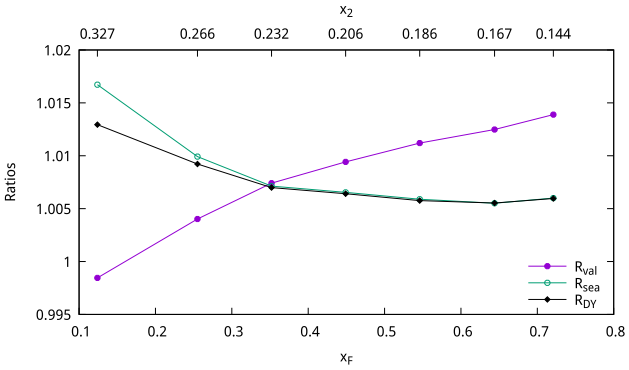

Fig. 6 Nuclear effects in the deuteron for the valence quark PDFs ($R_{\rm val}$), antiquark PDFs ($R_{\rm sea}$) and the DY cross sections ($R_{\rm DY}$) (see text for more detail on the definition of these ratios) vs x_F, computed using Table 5 of Ref. [24]. The upper horizontal axis indicates the corresponding x values of the deuteron target (x_2)

In Fig. 6 we illustrate the nuclear effects obtained for the valence quark PDFs, antiquark PDFs and the DY cross sections for the kinematics of the SeaQuest experiment. In particular, we show the ratios $R_{\rm val} = u_{{\rm val}/d}/(u_{{\rm val}/p} + u_{{\rm val}/n})$, $R_{\rm sea} = \overline{u}_d/(\overline{u}_p + \overline{u}_n)$ and $R_{\rm DY} = \sigma_{pd}/(\sigma_{pp} + \sigma_{pn})$ computed using Eq. (6), the NNLO proton PDFs of Ref. [32] and the values of kinematical variables from Table 6 of Ref. [24]. Note the different shapes of $R_{\rm val}$ and $R_{\rm sea}$ vs x_F. This is caused by different x dependencies of the valence and antiquark nucleon PDFs and the smearing effect in the nuclear

[12] The proton-neutron asymmetry $\delta f_p(x) - \delta f_n(x)$ was constrained in Ref. [65] in a global PDF fit using data on the proton, ^{2}H, ^{3}H, and ^{3}He targets. This asymmetry is consistent with zero within uncertainties.

convolution, Eq. (6). The shape and the magnitude of $R_{\rm DY}$ and $R_{\rm sea}$ are similar corresponding to the fact that the DY cross sections σ_{pd}, σ_{pp} and σ_{pn} for SeaQuest kinematics are dominated by the partonic contribution involving a proton beam valence u quark and a target $\bar{u}$, considering PDF x and flavour dependence. However, this dominance is violated for small values of x_F ($x_F < 0.3$), causing the different values of nuclear corrections for DY cross sections and the up-antiquark PDFs in this region. The magnitude of the nuclear corrections on the DY σ_{pd} is extremely modest, typically $\mathcal{O}(0.5-1)\%$, and has a practically negligible impact on the present analysis. This result is consistent with the claim of Ref. [24] that nuclear corrections can be neglected, on the basis of the results of Refs. [73,74].

Nuclear corrections should also be addressed when dealing with data from FNAL-E605 experiment on proton-copper collisions [26]. The corresponding corrections on the DY cross sections have been calculated in Ref. [67] (see Fig. 8 and Table 2 there). The rate of nuclear corrections depends on both the target x_2 and the mass of the muon pair as illustrated in Fig. 8 of Ref. [67]. Note, however, that the E605 experiment only provides data for copper target and did not take data for the proton target. Since copper is almost an isoscalar target with about 8% of the neutron excess, the E605 cross section data on copper target alone provides a little sensitivity to measuring the $(\bar{d}-\bar{u})(x)$ asymmetry of the sea distributions. We also verified that removing the E605 data from our fits does not essentially change the results presented in Figs. 3 and 4.

5 Second Mellin moments of quark distributions: comparisons with lattice QCD computations

Important information on PDFs can also be gained from lattice QCD, which gives access to some of their moments. Recalling that $q(x,Q^2)=q_{\rm val}(x,Q^2)+\bar{q}(x,Q^2)$, $\bar{q}(x,Q^2)=\bar{q}_{\rm sea}(x,Q^2)$, the following definition allows to summarize the moments of the $q^+\equiv q+\bar{q}$ (total) and $q^-\equiv q-\bar{q}$ (valence) quark combinations at a scale Q^2:

$$\langle x^{n-1}\rangle_{q^\pm}(Q^2) = \int_0^1 dx\, x^{n-1}\, q^\pm(x,Q^2), \tag{9}$$

where $n=1,2,3,\ldots$ refers to the first, second, third, etc. Mellin moment (equivalent to zeroth, first, second, etc. x moment), respectively. The first Mellin moments $\langle 1\rangle_{q^-}$ correspond to the quark number sum rules[13]. Lattice QCD computations have allowed to calculate the second Mellin moments $\langle x\rangle_{u^+-d^+}$ (isovector combination) and $\langle x\rangle_{q^+}$ for all individual light quarks, together with the third Mellin moments $\langle x^2\rangle_{u^--d^-}$ and $\langle x^2\rangle_{q^-}$ [75–77].

In Ref. [32] we compared the values of $\langle x^2\rangle_{u^-}$, $\langle x^2\rangle_{d^-}$, $\langle x^2\rangle_{u^--d^-}$ and $\langle x\rangle_{u^+-d^+}$ that we computed for various NNLO PDF fits with corresponding values extrapolated from lattice QCD computations.[14] In this work, we update and extend the comparison of Ref. [32]. On the one hand, in addition to several modern NLO and NNLO PDF fits, we incorporate the newly presented PDF fits from the previous sections, which take into account the SeaQuest data. As discussed in those sections, these data have minimal impact on the valence quark distributions. However, they play a crucial role in constraining the isospin asymmetry $(\bar{d}-\bar{u})(x)$ of the sea-quark distributions. On the other hand, we also consider updated evaluations from lattice QCD, utilizing new computational methods that yield reduced uncertainties compared to previous analyses. Additionally, we incorporate the recently released results on the moments of the u^+ and d^+ distributions, which were not available at the time of Ref. [32].

In Table 3 we compare our calculations of second Mellin moments using as a basis the NLO and NNLO quark distributions considered in the previous section, to the most recent results from lattice QCD [78–83]. In particular, the χQCD and ETMC collaborations have recently released data on the second moments of u^+, d^+ and u^+-d^+ combinations dependent on both valence and sea quarks, in Refs. [78] and [80], respectively, while the RQCD, NME, PNDME and Mainz collaborations have determined the second moments of the u^+-d^+ combination in Refs. [79,81,83], respectively.[15] The outcomes from Table 3 are also represented graphically in Fig. 7.

We observe that the present status of PDF fits is advanced to the point that the quoted uncertainties on the second Mellin moments, which represent the experimental data uncertainties propagated through the fit, are so small that the results from different fits are not always compatible among each other within their uncertainties. This is partly related to the theory assumptions made in those fits, but also due to the data sets considered, i.e., inclusion of DY data from colliders, see e.g., Ref. [1]. Across different orders of perturbation theory, the second Mellin moment values and uncertainties from NLO and NNLO PDF fits have almost comparable values, indicating very good perturbative stability.

The corresponding second Mellin moments from QCD lattice computations turn out to be significantly more uncertain and not yet able to discriminate between the various PDF fits. Taking into account the range of lattice results and the

[13] On the other hand, the moments $\langle 1\rangle_{q^+}$ are not constrained by symmetries and are divergent.

[14] See [20,61] for similar studies within the framework of the CT18 PDFs, also including lattice data in PDF fits.

[15] In the case where lattice collaborations have released values of second moments in more than one work, we only cite the most updated ones.

Table 3 Comparison of second Mellin moments for various combinations of light-quark distributions from different NLO and NNLO PDF fits, including those proposed in this work, with uncertainties due to PDF variations, to corresponding values extrapolated from n_f-flavour lattice QCD computations (Q = 2 GeV). In the case of the CT18 fit, the uncertainties refer to the 90% C.L. interval, instead of the 68% one used by other Hessian PDF fits

	$\langle x \rangle_{u^+}$	$\langle x \rangle_{d^+}$	$\langle x \rangle_{u^+-d^+}$
PDF fit			
ABMP16 + SeaQuest NLO	0.3523 ± 0.0010	0.1813 ± 0.0023	0.1711 ± 0.0029
ABMP16 + SeaQuest NNLO	0.3535 ± 0.0026	0.1858 ± 0.0028	0.1677 ± 0.0036
ABMP16 NLO	0.3522 ± 0.0026	0.1814 ± 0.0027	0.1708 ± 0.0036
ABMP16 NNLO	0.3532 ± 0.0027	0.1858 ± 0.0029	0.1673 ± 0.0037
NNPDF4.0 NNLO	0.3468 ± 0.0026	0.1934 ± 0.0032	0.1533 ± 0.0041
CT18 NNLO	$0.3498\ ^{+0.0078}_{-0.0085}$	$0.1934\ ^{+0.0083}_{-0.0103}$	$0.1564\ ^{+0.0123}_{-0.0120}$
MSHT20 NNLO	$0.3471\ ^{+0.0048}_{-0.0048}$	$0.1923\ ^{+0.0046}_{-0.0060}$	$0.1548\ ^{+0.0062}_{-0.0056}$
CJ15 NLO	$0.3480\ ^{+0.0009}_{-0.0012}$	$0.1962\ ^{+0.0015}_{-0.0014}$	$0.1518\ ^{+0.0019}_{-0.0024}$
epWZ16 NNLO	$0.3628\ ^{+0.0027}_{-0.0028}$	$0.1741\ ^{+0.0047}_{-0.0039}$	$0.1887\ ^{+0.0041}_{-0.0050}$
Lattice computation			
χQCD18 [78] ($n_f = 2+1$)	$0.307 \pm 0.030 \pm 0.018$	$0.160 \pm 0.027 \pm 0.040$	$0.151 \pm 0.028 \pm 0.029$
RQCD18 [79] ($n_f = 2$)	–	–	$0.195 \pm 0.007 \pm 0.015$
ETMC20 [80] ($n_f = 2+1+1$)	0.359 ± 0.030	0.188 ± 0.019	0.171 ± 0.018
PNDME20 [81] ($n_f = 2+1+1$)	–	–	$0.173 \pm 0.014 \pm 0.007$
NME20 [82] ($n_f = 2+1$)	–	–	$0.155 \pm 0.017 \pm 0.020$
Mainz21 [83] ($n_f = 2+1$)	–	–	0.139 ± 0.018 (stat)

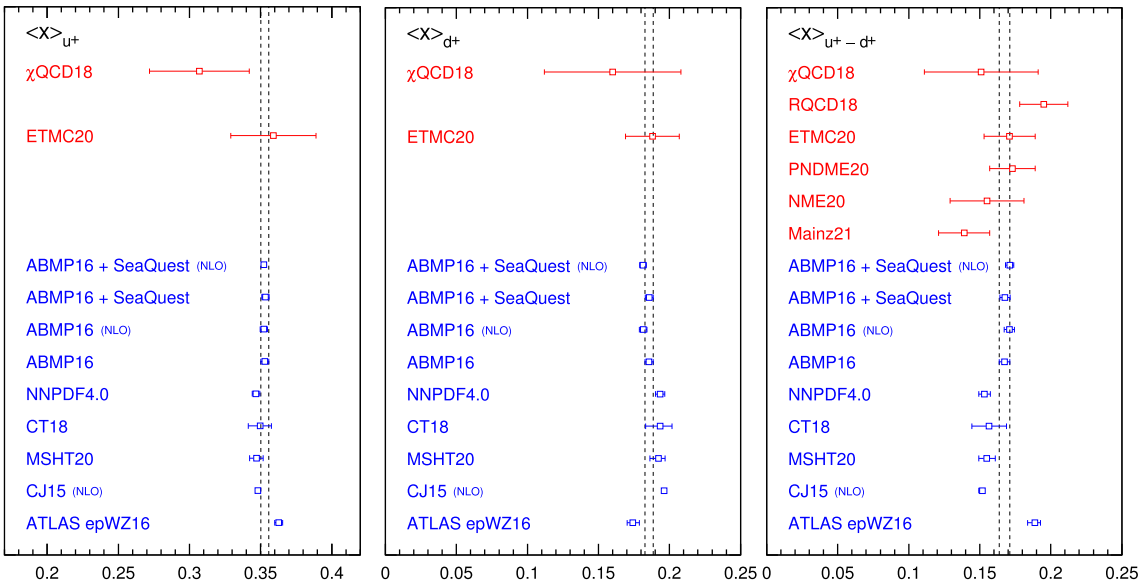

Fig. 7 Second Mellin moments of $u^+(x)$, $d^+(x)$ and the isovector combination $(u^+ - d^+)(x)$ and their uncertainties computed for a range of PDF fits and from lattice QCD. The corresponding numerical values are tabulated in the columns of Table 3 and reported in the panels of this plot for a more immediate visualization. The vertical band in each panel brackets the values from the ABMP16 NNLO fit

inherent uncertainty associated with each of them, they are presently exhibiting a high level of compatibility with nearly all the PDF fits. The lattice moments $\langle x \rangle_{u^+}$ by the χQCD collaboration and $\langle x \rangle_{u^+-d^+}$ by the RQCD collaborations, both computed in 2018, exhibit a slight tension, deviating from their 1σ range, when compared to the majority of PDF fits. Nonetheless, these results carry substantial uncertainties and align with the findings of PDF fits within a 2σ range. Most of the recent lattice results, in particular those obtained by the ETMC collaboration in 2020, turn out to agree very well with almost all PDF fits. The 2021 result on $\langle x \rangle_{u^+-d^+}$ of the Mainz collaboration agrees with moments of some of the global PDF fits, but is slightly smaller, although compatible within 2σ, with the second moments from the ABMP16 (+ SeaQuest) NLO and NNLO PDFs.

We also observe that the addition of SeaQuest data to the ABMP16 NNLO PDF fit has a tiny effect on the values of the considered moments, slightly decreasing the associated uncertainties, while the central values remain approximately the same. The improvement of the uncertainties turns out to be more pronounced in the case of the ABMP16 NLO PDF fit. Overall, the results from NLO and NNLO ABMP16 (+ SeaQuest) PDFs are consistent among each other and, as mentioned, the order of perturbation theory does not have a significant impact on the second Mellin moments, being rather inclusive quantities.

In light of the comparisons discussed here, it will be interesting to develop precise lattice calculations of higher Mellin moments (beyond the second/third ones), maybe exploiting concepts and techniques of Refs. [84,85], so as to enable similar comparisons for the fourth, etc. moments. Another valuable improvement would be the ability to distinguish between valence and sea quark PDFs in lattice results.

6 Conclusions

We have studied a variant of the ABMP16 NLO and NNLO fits, including the SeaQuest non-resonant data on $\sigma_{pd}/(2\sigma_{pp})$ as a function of x_F. We find that these data reduce uncertainties on the $(\bar{d} - \bar{u})(x)$ difference as well as on the $(\bar{d}/\bar{u})(x)$ ratio at large x, while leaving essentially unchanged the values of the other quantities, which are simultaneously constrained in these fits ($\alpha_s(M_Z)$ and heavy-quark masses). The χ^2/NDP for the fits including SeaQuest data are within statistical uncertainty of those previously obtained without these data. The simultaneous description of all DY data turns out to be slightly more consistent at NNLO than at NLO, as expected for the improved precision of the theoretical predictions. In particular, we observe the compatibility of SeaQuest data constraints on the $\bar{d} - \bar{u}$ asymmetry, with the corresponding constraints from collider DY data at both the Tevatron and the LHC. This confirms the presence of an asymmetric sea, ruling out PDF fits based on the assumption of (or leading to) a symmetric sea.

Our present results support using the SeaQuest data together with collider DY data in future updated PDF analyses, that would allow further reducing PDF uncertainties, as well as a cross-check of the compatibility with the data already included there. The inclusion of SeaQuest data is facilitated by the fact that nuclear corrections for the deuteron target, that we have explicitly computed in this work, turned out to be $\mathcal{O}(0.5-1)\%$ in all SeaQuest x_F bins, thus having a practically negligible effect on the final PDFs. The smallness of the observed nuclear effects can be attributed to the kinematics of the SeaQuest experiment itself. The experiment combines partons with relatively small but still significant x_2 values (specifically, x_{target}) and larger x_1 values (specifically, x_{beam}), with only the target experiencing nuclear corrections. The most substantial corrections occur in the bin with the smallest x_F, which corresponds to the largest x_2 values (≤ 0.45). It is worth noting that larger nuclear corrections would be anticipated at larger x_2 values, which correspond to backward kinematics that fall outside the scope of SeaQuest's current detector capabilities.

The second moments of various combinations of light-quark distributions from NLO and NNLO PDF fits are compatible with current lattice QCD results. Although lattice QCD is not yet competitive for distinguishing between different PDF fits, advancements in techniques and increased attention from the lattice community are expected to improve this limitation in the future.

We strongly encourage the SeaQuest collaboration to continue their efforts in reducing the uncertainties associated with their measurements, aiming for values below the current level of approximately 5%. Achieving this would significantly enhance the constraining power of the data on sea quark distributions. It is worth noting that only around one half of the experiment's data has been utilized for the published studies thus far, suggesting the potential for further improvements. Moreover, it would be highly beneficial if the SeaQuest collaboration releases separate data on pp and pd cross sections. Such separate data sets would enable more precise constraints to be obtained for the $\bar{u}$ and $\bar{d}$ quark distributions, facilitating a deeper understanding of their individual characteristics.

Acknowledgements We would like to thank R. Petti for stimulating disussions, and S. Collins and G. Bali for insightful comments concerning lattice QCD results. We thank A. Cooper-Sarkar and F. Giuli for remarks on the ATLAS PDF fits and the participants of the Les Houches Workshop PhysTeV 2023 for further comments. We are also grateful to the anonymous Referee for the constructive report and for having brought to our attention further studies closely related to the topic of this manuscript. The work of S.A. and S.-O.M. has been supported in part by the Deutsche Forschungsgemeinschaft through the Research Unit FOR 2926, *Next Generation pQCD for Hadron Structure: Preparing for the EIC*, project number 40824754. The work of M.V.G. and

S.-O.M. has been supported in part by the Bundesministerium für Bildung und Forschung under contract 05H21GUCCA.

Data Availability Statement This manuscript has no associated data or the data will not be deposited. [Authors' comment: The ABMP16+SeaQuest PDF fit will be released. Therefore the corresponding data will be deposited.].

Funded by SCOAP[3]. SCOAP[3] supports the goals of the International Year of Basic Sciences for Sustainable Development.

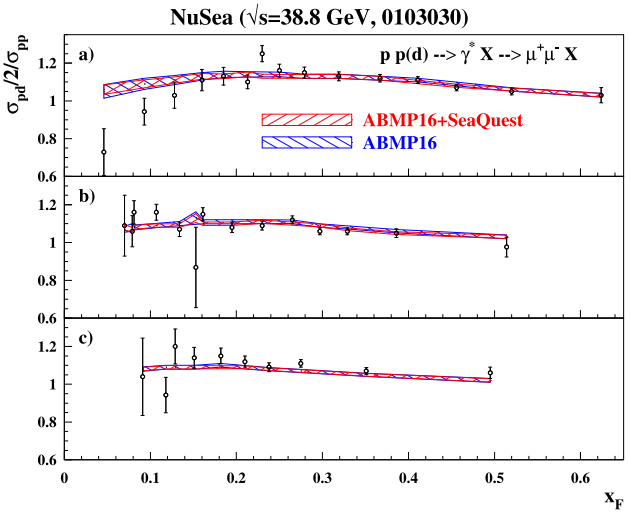

Fig. 8 The NuSea data [25] on the ratio of the pd and pp DY distributions over x_F w.r.t. the NNLO predictions obtained using the code VRAP in combination with the PDFs obtained in the present analysis (right-tilted hash) and in the ABMP16 one [32] (left-tilted hash). The three panels correspond to (a): the high-mass spectrometer setting, (b): the intermediate-mass one, (c): the low-mass one

Appendix A: Comparison of our theory predictions with NuSea data on the DY $\sigma_{pd}/(2\sigma_{pp})$ ratio

A comparison of predictions using as input the ABMP16 and the ABMP16 + SeaQuest fits with the experimental data on the $\sigma_{pd}/(2\sigma_{pp})$ ratio of cross sections for DY production as a function of x_F released by the NuSea collaboration in Ref. [25] is shown in Fig. 8. These data were included in both our fits, as well as in many predecessors of them, and their role was already investigated and discussed at the time these fits were published. The agreement of our predictions with the NuSea experimental data turn out to be of similar quality for both fits considered in this work. The global χ^2 for NuSea data (NDP = 39) differs by 1.5 units only: χ^2(ABMP16) = 52.8 vs. χ^2(ABMP16 + SeaQuest)= 54.3, as shown in Table 2. This difference is too small to be statistically relevant. Beyond considering them globally, data in the NuSea experimental analysis are organized in three groups, corresponding to different spectrometer settings, emphasizing the role of different dimuon invariant mass intervals, as shown in Fig. 2 of Ref. [25]. In particular, data from the high-invariant-mass settings cover x_F values in the range $0.046 < \langle x_F \rangle < 0.624$, with x_2 in the range $0.35 > x_2 > 0.015$, data from the intermediate-mass settings cover x_F values in the range $0.153 < \langle x_F \rangle < 0.514$, with x_2 in the range $0.3 > x_2 > 0.015$, whereas data in the low-mass settings cover x_F values in the range $0.091 < \langle x_F \rangle < 0.495$ with x_2 within the range $0.175 > x_2 > 0.015$.

In the case of spectrometer settings focused on high dimuon invariant masses (upper panel), the theory/data agreement is lost at low x_F in both fits, i.e. in the first two bins (the x_2 value corresponding to the ensemble of these two bins cover the interval $0.25 < x_2 < 0.35$). In the case of intermediate-mass settings (intermediate panel), the agreement is overall reasonable but it slightly deteriorates in the $x_F \sim 0.15$ bin (the x_2 value in this bin is $0.25 < x_2 < 0.3$), whereas in case of settings focused on low invariant masses (lowest panel) the agreement is also reasonable but so large x_2 values are not involved. One should first of all observe that the size of the uncertainty bands of our fits is not driven just by NuSea data, but is defined also by many other data points from different datasets. Additionally, one should also notice that the NuSea experimental uncertainties grow at low x_F, especially in case of high- and intermediate-invariant-mass settings, that makes our predictions still compatible with the experimental data within twice the errorbars of the latter at low x_F. Therefore the disagreement mentioned above is of no statistical significance for our fits. Furthermore, considering their larger uncertainties, these NuSea datapoints play a much less relevant role in the global NuSea χ^2 of our fits, when compared to the role played by other data. On the other hand, NuSea data at large x_F are very relevant for constraining the $(\bar{d} - \bar{u})(x)$ difference at low x, i.e. in the $0.015 < x_2 < 0.12$ range, not covered by SeaQuest data (see Fig. 1 for better visualizing the difference in the kinematical coverage of SeaQuest and NuSea data). DY data from collider experiments are complementary and also relevant on a wide x range, although they probe it at a larger scale. Our findings emphasize that the role of SeaQuest and NuSea data is partly complementary, considering the x_2 kinematical coverage, and point to no statistically relevant tension between NuSea and SeaQuest results, when considering their inclu-

sion in our PDF fits. Already in Ref. [44], it was commented that the NuSea and SeaQuest experiments refer to different scales and acceptances. On the other hand, the lack of smoothness of NuSea data when comparing the behaviour in some contiguous bins point to the fact that the 1σ uncertainties on the latter might have been underestimated. An experimental reanalysis of the data collected in 2001 or a repetition of the measurement would be welcome.

References

1. A. Accardi et al., A critical appraisal and evaluation of modern PDFs. Eur. Phys. J. C **76**, 471 (2016). https://doi.org/10.1140/epjc/s10052-016-4285-4. arXiv:1603.08906
2. M. Begel et al., Precision QCD, hadronic structure and forward QCD, heavy ions: report of energy Frontier topical groups 5, 6, 7 submitted to Snowmass 2021. arXiv:2209.14872
3. R. Abdul Khalek et al., Science requirements and detector concepts for the electron-ion collider: EIC Yellow Report. Nucl. Phys. A **1026**, 122447 (2022). https://doi.org/10.1016/j.nuclphysa.2022.122447. arXiv:2103.05419
4. R. Abir et al., The case for an EIC theory alliance: theoretical challenges of the EIC. arXiv:2305.14572
5. R. Abdul Khalek et al., Snowmass 2021 White Paper: electron ion collider for high energy physics. arXiv:2203.13199
6. C. Hadjidakis et al., A fixed-target programme at the LHC: physics case and projected performances for heavy-ion, hadron, spin and astroparticle studies. Phys. Rep. **911**, 1 (2021). https://doi.org/10.1016/j.physrep.2021.01.002. arXiv:1807.00603
7. A. Bursche et al., Physics opportunities with the fixed-target program of the LHCb experiment using an unpolarized gas target. https://cds.cern.ch/record/2649878/ [cds.cern.ch]
8. P. Di Nezza, V. Carassiti, G. Ciullo, P. Lenisa, L.L. Pappalardo, E. Steffens et al., The SMOG2 project. https://cds.cern.ch/record/2651269 [cds.cern.ch]
9. F. Galluccio, C. Hadjidakis, A. Kurepin, L. Massacrier, S. Porteboeuf, K. Pressard et al., Physics opportunities for a fixed-target programme in the ALICE experiment. http://cds.cern.ch/record/2671944 [cds.cern.ch]
10. L.A. Anchordoqui et al., The Forward Physics Facility: sites, experiments, and physics potential. Phys. Rep. **968**, 1 (2022). https://doi.org/10.1016/j.physrep.2022.04.004. arXiv:2109.10905
11. J.L. Feng et al., The forward physics facility at the high-luminosity LHC. J. Phys. G **50**, 030501 (2023). https://doi.org/10.1088/1361-6471/ac865e. arXiv:2203.05090
12. H1, ZEUS Collaboration, 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). https://doi.org/10.1140/epjc/s10052-015-3710-4. arXiv:1506.06042
13. Z.E.U.S. Collaboration, Measurement of neutral current e±p cross sections at high Bjorken x with the ZEUS detector. Phys. Rev. D **89**, 072007 (2014). https://doi.org/10.1103/PhysRevD.89.072007. arXiv:1312.4438
14. Jefferson Lab Hall A Tritium Collaboration, Measurement of the nucleon F_2^n/F_2^p structure function ratio by the Jefferson Lab MARATHON tritium/helium-3 deep inelastic scattering experiment. Phys. Rev. Lett. **128**, 132003 (2022). https://doi.org/10.1103/PhysRevLett.128.132003. arXiv:2104.05850
15. ATLAS Collaboration, Comprehensive measurements of t-channel single top-quark production cross sections at $\sqrt{s} = 7$ TeV with the ATLAS detector. Phys. Rev. D **90**, 112006 (2014). https://doi.org/10.1103/PhysRevD.90.112006. arXiv:1406.7844
16. S. Alekhin, J. Blumlein, S. Moch, Strange sea determination from collider data. Phys. Lett. B **777**, 134 (2018). https://doi.org/10.1016/j.physletb.2017.12.024. arXiv:1708.01067
17. F. Faura, S. Iranipour, E.R. Nocera, J. Rojo, M. Ubiali, The strangest proton? Eur. Phys. J. C **80**, 1168 (2020). https://doi.org/10.1140/epjc/s10052-020-08749-3. arXiv:2009.00014
18. S. Catani, D. de Florian, G. Rodrigo, W. Vogelsang, Perturbative generation of a strange-quark asymmetry in the nucleon. Phys. Rev. Lett. **93**, 152003 (2004). https://doi.org/10.1103/PhysRevLett.93.152003. arXiv:hep-ph/0404240
19. S. Bailey, T. Cridge, L.A. Harland-Lang, A.D. Martin, R.S. Thorne, Parton distributions from LHC, HERA, Tevatron and fixed target data: MSHT20 PDFs. Eur. Phys. J. C **81**, 341 (2021). https://doi.org/10.1140/epjc/s10052-021-09057-0. arXiv:2012.04684
20. T.-J. Hou, H.-W. Lin, M. Yan, C.P. Yuan, Impact of lattice strangeness asymmetry data in the CTEQ-TEA global analysis. Phys. Rev. D **107**, 076018 (2023). https://doi.org/10.1103/PhysRevD.107.076018. arXiv:2211.11064
21. G. Bevilacqua, M.V. Garzelli, A. Kardos, L. Toth, W + charm production with massive c quarks in PowHel. JHEP **04**, 056 (2022). https://doi.org/10.1007/JHEP04(2022)056. arXiv:2106.11261
22. S. Amoroso et al., Snowmass 2021 Whitepaper: proton structure at the precision Frontier. Acta Phys. Polon. B **53**, 1 (2022). https://doi.org/10.5506/APhysPolB.53.12-A1. arXiv:2203.13923
23. K. Gottfried, Sum rule for high-energy electron-proton scattering. Phys. Rev. Lett. **18**, 1174 (1967). https://doi.org/10.1103/PhysRevLett.18.1174
24. SeaQuest Collaboration, Measurement of flavor asymmetry of light-quark sea in the proton with Drell–Yan dimuon production in $p + p$ and $p + d$ collisions at 120 GeV. arXiv:2212.12160
25. FNAL E866/NuSea Collaboration, Improved measurement of the anti-d/anti-u asymmetry in the nucleon sea. Phys. Rev. D **64**, 052002 (2001). https://doi.org/10.1103/PhysRevD.64.052002. arXiv:hep-ex/0103030
26. G. Moreno, C. Brown, W. Cooper, D. Finley, Y. Hsiung et al., Dimuon production in proton–copper collisions at $\sqrt{s} = 38.8$-GeV. Phys. Rev. D **43**, 2815 (1991). https://doi.org/10.1103/PhysRevD.43.2815
27. S.D. Ellis, W.J. Stirling, Constraints on isospin breaking in the light quark sea from the Drell-Yan process. Phys. Lett. B **256**, 258 (1991). https://doi.org/10.1016/0370-2693(91)90684-I
28. NA51 collaboration, Study of the isospin symmetry breaking in the light quark sea of the nucleon from the Drell-Yan process. Phys. Lett. B **332**, 244 (1994). https://doi.org/10.1016/0370-2693(94)90884-2
29. D.F. Geesaman, P.E. Reimer, The sea of quarks and antiquarks in the nucleon. Rep. Prog. Phys. **82**, 046301 (2019). https://doi.org/10.1088/1361-6633/ab05a7. arXiv:1812.10372
30. W.-C. Chang, J.-C. Peng, Flavor structure of the nucleon sea. Prog. Part. Nucl. Phys. **79**, 95 (2014). https://doi.org/10.1016/j.ppnp.2014.08.002. arXiv:1406.1260
31. S. Alekhin, J. Blümlein, S. Moch, NLO PDFs from the ABMP16 fit. Eur. Phys. J. C **78**, 477 (2018). https://doi.org/10.1140/epjc/s10052-018-5947-1. arXiv:1803.07537
32. S. Alekhin, J. Blümlein, S. Moch, R. Placakyte, Parton distribution functions, α_s, and heavy-quark masses for LHC Run II. Phys. Rev. D **96**, 014011 (2017). https://doi.org/10.1103/PhysRevD.96.014011. arXiv:1701.05838
33. LHCb Collaboration, Measurement of the forward Z boson production cross-section in pp collisions at $\sqrt{s} = 7$ TeV. JHEP **08**, 039 (2015). https://doi.org/10.1007/JHEP08(2015)039. arXiv:1505.07024
34. LHCb Collaboration, Measurement of forward W and Z boson production in pp collisions at $\sqrt{s} = 8$ TeV. JHEP **01**, 155 (2016). https://doi.org/10.1007/JHEP01(2016)155. arXiv:1511.08039

35. LHCb Collaboration, Measurement of the forward Z boson production cross-section in pp collisions at $\sqrt{s} = 7$ TeV. JHEP **08**, 039 (2015). https://doi.org/10.1007/JHEP08(2015)039. arXiv:1505.07024
36. LHCb Collaboration, Measurement of forward Z $\rightarrow e^+e^-$ production at $\sqrt{s} = 8$ TeV. JHEP **05**, 109 (2015). https://doi.org/10.1007/JHEP05(2015)109. arXiv:1503.00963
37. LHCb Collaboration, Measurement of forward W and Z boson production in pp collisions at $\sqrt{s} = 8$ TeV. JHEP **01**, 155 (2016). https://doi.org/10.1007/JHEP01(2016)155. arXiv:1511.08039
38. 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). https://doi.org/10.1103/PhysRevD.88.091102. arXiv:1309.2591
39. D0 Collaboration, Measurement of the electron charge asymmetry in $\boldsymbol{p\bar{p} \rightarrow W + X \rightarrow e\nu + X}$ decays in $\boldsymbol{p\bar{p}}$ collisions at $\boldsymbol{\sqrt{s} = 1.96}$ TeV. Phys. Rev. D **91**, 032007 (2015). https://doi.org/10.1103/PhysRevD.91.032007. arXiv:1412.2862
40. C. Anastasiou, L.J. Dixon, K. Melnikov, F. Petriello, High precision QCD at hadron colliders: electroweak gauge boson rapidity distributions at NNLO. Phys. Rev. D **69**, 094008 (2004). https://doi.org/10.1103/PhysRevD.69.094008. arXiv:hep-ph/0312266
41. S. Catani, L. Cieri, G. Ferrera, D. de Florian, M. Grazzini, Vector boson production at hadron colliders: a fully exclusive QCD calculation at NNLO. Phys. Rev. Lett. **103**, 082001 (2009). https://doi.org/10.1103/PhysRevLett.103.082001. arXiv:0903.2120
42. R. Gavin, Y. Li, F. Petriello, S. Quackenbush, W physics at the LHC with FEWZ 2.1. Comput. Phys. Commun. **184**, 208 (2013). https://doi.org/10.1016/j.cpc.2012.09.005. arXiv:1201.5896
43. S. Alekhin, A. Kardos, S. Moch, Z. Trócsányi, Precision studies for Drell–Yan processes at NNLO. Eur. Phys. J. C **81**, 573 (2021). https://doi.org/10.1140/epjc/s10052-021-09361-9. arXiv:2104.02400
44. SeaQuest Collaboration, The asymmetry of antimatter in the proton. Nature **590**, 561 (2021). https://doi.org/10.1038/s41586-022-04707-z. arXiv:2103.04024
45. D0 Collaboration, Measurement of the muon charge asymmetry in $p\bar{p} \rightarrow$ W+X $\rightarrow \mu\nu$ + X events at $\sqrt{s}$ = 196 TeV. Phys. Rev. D **88**, 091102 (2013). https://doi.org/10.1103/PhysRevD.88.091102. arXiv:1309.2591
46. LHCb Collaboration, Measurement of forward Z $\rightarrow$ e$^+$e$^-$ production at $\sqrt{s} = 8$ TeV. JHEP **05**, 109 (2015). https://doi.org/10.1007/JHEP05(2015)109. arXiv:1503.00963
47. NNPDF Collaboration, The path to proton structure at 1% accuracy. Eur. Phys. J. C **82**, 428 (2022). https://doi.org/10.1140/epjc/s10052-022-10328-7. arXiv:2109.02653
48. NNPDF Collaboration, Parton distributions from high-precision collider data. Eur. Phys. J. C **77**, 663 (2017). https://doi.org/10.1140/epjc/s10052-017-5199-5. arXiv:1706.00428
49. J. Fiaschi, F. Giuli, F. Hautmann, S. Moch, S. Moretti, Z'-boson dilepton searches and the high-x quark density. Phys. Lett. B **841**, 137915 (2023). https://doi.org/10.1016/j.physletb.2023.137915. arXiv:2211.06188
50. R.D. Ball, A. Candido, S. Forte, F. Hekhorn, E.R. Nocera, J. Rojo et al., Parton distributions and new physics searches: the Drell–Yan forward–backward asymmetry as a case study. Eur. Phys. J. C **82**, 1160 (2022). https://doi.org/10.1140/epjc/s10052-022-11133-y. arXiv:2209.08115
51. E. Accomando, J. Fiaschi, F. Hautmann, S. Moretti, Constraining parton distribution functions from neutral current Drell–Yan measurements. Phys. Rev. D **98**, 013003 (2018). https://doi.org/10.1103/PhysRevD.98.013003. arXiv:1712.06318
52. J. Fiaschi, F. Giuli, F. Hautmann, S. Moretti, Enhancing the Large Hadron Collider sensitivity to charged and neutral broad resonances of new gauge sectors. JHEP **02**, 179 (2022). https://doi.org/10.1007/JHEP02(2022)179. arXiv:2111.09698
53. ATLAS Collaboration, Determination of the parton distribution functions of the proton using diverse ATLAS data from pp collisions at $\sqrt{s} = 7$, 8 and 13 TeV. Eur. Phys. J. C **82**, 438 (2022). https://doi.org/10.1140/epjc/s10052-022-10217-z. arXiv:2112.11266
54. ATLAS Collaboration, Precision measurement and interpretation of inclusive W^+ , W^- and Z/γ^* production cross sections with the ATLAS detector. Eur. Phys. J. C **77**, 367 (2017). https://doi.org/10.1140/epjc/s10052-017-4911-9. arXiv:1612.03016
55. A. Accardi, L.T. Brady, W. Melnitchouk, J.F. Owens, N. Sato, Constraints on large-x parton distributions from new weak boson production and deep-inelastic scattering data. Phys. Rev. D **93**, 114017 (2016). https://doi.org/10.1103/PhysRevD.93.114017. arXiv:1602.03154
56. T.-J. Hou et al., New CTEQ global analysis of quantum chromodynamics with high-precision data from the LHC. Phys. Rev. D **103**, 014013 (2021). https://doi.org/10.1103/PhysRevD.103.014013. arXiv:1912.10053
57. A.D. Martin, W.J. Stirling, R.S. Thorne, G. Watt, Parton distributions for the LHC. Eur. Phys. J. C **63**, 189 (2009). https://doi.org/10.1140/epjc/s10052-009-1072-5. arXiv:0901.0002
58. M. 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). https://doi.org/10.21468/SciPostPhysProc.8.005. arXiv:2108.06596 [hep-ph]
59. S. Park, A. Accardi, X. Jing , J.F. Owens, CJ15 global PDF analysis with new electroweak data from the STAR and SeaQuest experiments, in *28th International Workshop on Deep Inelastic Scattering and Related Subjects*, 8 (2021). arXiv:2108.05786
60. 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). https://doi.org/10.1103/PhysRevD.103.012001. arXiv:2011.04708
61. T.-J. Hou, M. Yan, J. Liang, K.-F. Liu, C.P. Yuan, Connected and disconnected sea partons from the CT18 parametrization of PDFs. Phys. Rev. D **106**, 096008 (2022). https://doi.org/10.1103/PhysRevD.106.096008. arXiv:2206.02431
62. 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). https://doi.org/10.1103/PhysRevD.104.074031. arXiv:2109.00677
63. A. Accardi, X. Jing, J.F. Owens, S. Park, Light quark and antiquark constraints from new electroweak data. Phys. Rev. D **107**, 113005 (2023). https://doi.org/10.1103/PhysRevD.107.113005. arXiv:2303.11509
64. S.I. Alekhin, S.A. Kulagin, R. Petti, Nuclear effects in the deuteron and global QCD analyses. Phys. Rev. D **105**, 114037 (2022). https://doi.org/10.1103/PhysRevD.105.114037. arXiv:2203.07333
65. S.I. Alekhin, S.A. Kulagin, R. Petti, Off-shell effects in bound nucleons and parton distributions from H1, H2, H3, and He3 data. Phys. Rev. D **107**, L051506 (2023). https://doi.org/10.1103/PhysRevD.107.L051506. arXiv:2211.09514
66. ATLAS Collaboration, Determination of the parton distribution functions of the proton from ATLAS measurements of differential W$^{\pm}$ and Z boson production in association with jets. JHEP **07**, 223 (2021). https://doi.org/10.1007/JHEP07(2021)223. arXiv:2101.05095
67. S.A. Kulagin, R. Petti, Nuclear parton distributions and the Drell-Yan process. Phys. Rev. C **90**, 045204 (2014). https://doi.org/10.1103/PhysRevC.90.045204. arXiv:1405.2529
68. S.A. Kulagin, G. Piller, W. Weise, Shadowing, binding and off-shell effects in nuclear deep inelastic scattering. Phys. Rev. C **50**, 1154 (1994). https://doi.org/10.1103/PhysRevC.50.1154. arXiv:nucl-th/9402015

69. S.A. Kulagin, R. Petti, Global study of nuclear structure functions. Nucl. Phys. A **765**, 126 (2006). https://doi.org/10.1016/j.nuclphysa.2005.10.011. arXiv:hep-ph/0412425
70. S.A. Kulagin, R. Petti, Structure functions for light nuclei. Phys. Rev. C **82**, 054614 (2010). https://doi.org/10.1103/PhysRevC.82.054614. arXiv:1004.3062
71. P. Ru, S.A. Kulagin, R. Petti, B.-W. Zhang, Study of $W^{\pm}$ and Z boson production in proton-lead collisions at the LHC with Kulagin–Petti nuclear parton distributions. Phys. Rev. D **94**, 113013 (2016). https://doi.org/10.1103/PhysRevD.94.113013. arXiv:1608.06835
72. S.I. Alekhin, S.A. Kulagin, R. Petti, Nuclear effects in the deuteron and constraints on the d/u ratio. Phys. Rev. D **96**, 054005 (2017). https://doi.org/10.1103/PhysRevD.96.054005. arXiv:1704.00204
73. H. Kamano, T.S.H. Lee, Fermi motion and pion-exchange effects on Drell-Yan processes. Phys. Rev. D **86**, 094037 (2012). https://doi.org/10.1103/PhysRevD.86.094037. arXiv:1207.0181
74. P.J. Ehlers, A. Accardi, L.T. Brady, W. Melnitchouk, Nuclear effects in the proton-deuteron Drell-Yan process. Phys. Rev. D **90**, 014010 (2014). https://doi.org/10.1103/PhysRevD.90.014010. arXiv:1405.2039
75. H.-W. Lin et al., Parton distributions and lattice QCD calculations: a community white paper. Prog. Part. Nucl. Phys. **100**, 107 (2018). https://doi.org/10.1016/j.ppnp.2018.01.007. arXiv:1711.07916
76. M. Constantinou et al., Parton distributions and lattice-QCD calculations: toward 3D structure. Prog. Part. Nucl. Phys. **121**, 103908 (2021). https://doi.org/10.1016/j.ppnp.2021.103908. arXiv:2006.08636
77. Z. Davoudi et al., Report of the Snowmass 2021 Topical Group on Lattice Gauge Theory, in Snowmass 2021, 9 (2022). arXiv:2209.10758
78. Y.-B. Yang, J. Liang, Y.-J. Bi, Y. Chen, T. Draper, K.-F. Liu et al., Proton mass decomposition from the QCD energy momentum tensor. Phys. Rev. Lett. **121**, 212001 (2018). https://doi.org/10.1103/PhysRevLett.121.212001. arXiv:1808.08677
79. G.S. Bali, S. Collins, M. Göckeler, R. Rödl, A. Schäfer, A. Sternbeck, Nucleon generalized form factors from two-flavor lattice QCD. Phys. Rev. D **100**, 014507 (2019). https://doi.org/10.1103/PhysRevD.100.014507. arXiv:1812.08256
80. C. Alexandrou, S. Bacchio, M. Constantinou, J. Finkenrath, K. Hadjiyiannakou, K. Jansen et al., Complete flavor decomposition of the spin and momentum fraction of the proton using lattice QCD simulations at physical pion mass. Phys. Rev. D **101**, 094513 (2020). https://doi.org/10.1103/PhysRevD.101.094513. arXiv:2003.08486
81. S. Mondal, R. Gupta, S. Park, B. Yoon, T. Bhattacharya, H.-W. Lin, Moments of nucleon isovector structure functions in 2+1+1-flavor QCD. Phys. Rev. D **102**, 054512 (2020). https://doi.org/10.1103/PhysRevD.102.054512. arXiv:2005.13779
82. Nucleon Matrix Elements (NME) Collaboration, Nucleon momentum fraction, helicity and transversity from 2+1-flavor lattice QCD. JHEP **21**, 004 (2020). https://doi.org/10.1007/JHEP04(2021)044. arXiv:2011.12787
83. K. Ottnad, D. Djukanovic, T. Harris, H.B. Meyer, G. von Hippel, H. Wittig, Improved analysis of nucleon isovector charges and twist-2 matrix elements on CLS $N_f = 2 + 1$ ensembles. PoS LATTICE2021, 343 (2022). https://doi.org/10.22323/1.396.0343. arXiv:2110.10500
84. W. Detmold, C.J.D. Lin, Deep-inelastic scattering and the operator product expansion in lattice QCD. Phys. Rev. D **73**, 014501 (2006). https://doi.org/10.1103/PhysRevD.73.014501. arXiv:hep-lat/0507007
85. Z. Davoudi, M.J. Savage, Restoration of rotational symmetry in the continuum limit of lattice field theories. Phys. Rev. D **86**, 054505 (2012). https://doi.org/10.1103/PhysRevD.86.054505. arXiv:1204.4146