Proton PDFs with nonlinear corrections from gluon recombination

P. Duwentäster,* V. Guzey,† I. Helenius,‡ and H. Paukkunen§

University of Jyvaskyla, Department of Physics, P.O. Box 35, FI-40014 University of Jyvaskyla, Finland and Helsinki Institute of Physics, P.O. Box 64, FI-00014 University of Helsinki, Finland

(Received 8 January 2024; accepted 7 March 2024; published 2 May 2024)

We present numerical studies of the leading nonlinear corrections to the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi evolution equations of parton distribution functions (PDFs) resulting from gluon recombination. The effect of these corrections is to reduce the pace of evolution at small momentum fractions x while slightly increasing it at intermediate x. By implementing the nonlinear evolution in the XFITTER framework, we have carried out fits of proton PDFs using data on lepton-proton deep inelastic scattering from HERA, BCDMS, and NMC. While we find no evidence for the presence of nonlinearities, they cannot be entirely excluded either, and we are able to set limits for their strength. In terms of the recombination scale Q_r, which plays a similar role as the saturation scale in the dipole picture of the proton, we find that $Q_r \lesssim 2.5$ GeV. We also quantify the impact of the nonlinear terms for the longitudinal structure function at the Electron-Ion Collider and the Large Hadron Electron Collider and find that these measurements could provide stronger constraints on the projected effects.

DOI: 10.1103/PhysRevD.109.094004

I. INTRODUCTION

In the framework of quantum chromodynamics (QCD) and collinear factorization [1], the cross sections for hard processes involving initial-state hadrons are calculated as convolutions of perturbative, process-dependent coefficient functions and nonperturbative, process-independent parton distribution functions (PDFs) $f_i(x, Q^2)$. The PDFs describe the density of partons (gluons and quarks) of flavor i carrying a fraction x of the hadron's momentum at a resolution scale Q. While the x dependence cannot be calculated by the perturbative means of QCD, its Q^2 dependence is governed by the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) evolution equations [2–5]. These evolution equations originate from the renormalization of divergences induced by collinear radiation from the initial-state partons. As a result of the radiation, the density of partons that carry a momentum smaller than that from which the radiation tree started grows, leading to a rise of PDFs at low momentum fractions x as the scale Q^2 increases. This effect is especially pronounced for the gluon distribution. However, if the parton densities become sufficiently high, the inverse processes in which several partons recombine will also begin to play a part, moderating the growth of PDFs at small x as Q^2 increases. These contributions are formally of higher twist; i.e., they are suppressed by inverse powers of Q^2. Since the Q^2 dependence of PDFs is only logarithmic, the effects of recombination should disappear as Q^2 grows.

In the language of PDFs, the first recombination terms in the evolution can be derived from diagrams, where two initial-state gluons merge into collinear gluons/quarks that eventually participate in the hard scattering. The leading contributions at small x were originally calculated in Refs. [6–9], giving rise to the Gribov-Levin-Ryskin-Mueller-Qiu (GLR-MQ) corrections to the DGLAP evolution equations. Phenomenological applications of this approach have been studied in a number of publications, including its impact on observables in deep inelastic scattering (DIS) [10], comparisons with HERA data [11,12], its relation to the Balitsky-Kovchegov (BK) evolution equation [13,14], and the evolution of nuclear PDFs [15]. However, the GLR-MQ equations violate the momentum conservation. This issue was fixed in the calculation of Zhu and Ruan [16,17], which is based on the same class of diagrams but also keeps the nonleading contributions; see also Ref. [18]. The resulting evolution equation is the one whose implications we study here within the framework of global fits of PDFs. The effect of the nonlinear corrections is to moderate the Q^2 growth of PDFs at small x, while somewhat increasing the speed of evolution at moderate values of x such that the total momentum is conserved. This allows us to speak about

*pitduwen@jyu.fi
†vadim.a.guzey@jyu.fi
‡ilkka.m.helenius@jyu.fi
§hannu.t.paukkunen@jyu.fi

dynamically generated shadowing and antishadowing—terms which are perhaps more familiar from the context of nuclear PDFs [19]. Earlier phenomenological studies of these nonlinear corrections can be found, e.g., in Refs. [20,21].

For dimensional reasons, the higher-twist effects imply the presence of a dimensionful parameter R, which controls the strength of the nonlinear effects and which can be associated with the size of the area in the transverse plane, where the partonic overlap leading to the nonlinear corrections becomes important. The inverse of this parameter $1/R = Q_r$ gives, correspondingly, the momentum scale—the recombination scale as we will call it—below which the effects of recombination will be considerable. Once Q^2 is so low that the nonlinear terms in the evolution are as important as the linear ones, terms suppressed by higher inverse powers of Q^2 also become presumably important, and the resummation of all these terms can be thought to give rise to the phenomenon of partonic saturation [9]. This regime is traditionally discussed in the framework of the color glass condensate (CGC) effective field theory [22,23] and is implemented in the dipole picture of DIS. In this case, the saturation effects are characterized by the saturation momentum scale Q_s, which can be related to the critical dipole transverse size. It would be tempting to link the two emergent scales Q_r and Q_s. However, in the dipole language, Q_s also depends on x roughly as $Q_s \sim x^{-b}$, whereas in the picture discussed here, it is a constant, with the x dependence being entirely dynamical. Thus, the two cannot be compared directly, and the comparison is at best only indicative.

In the presented work, we implement the first momentum-conserving nonlinear corrections into the HOPPET evolution code [24] and interface it with the XFITTER framework [25,26]. This allows us to perform global fits of proton PDFs with a range of different values for R. Scanning through various values of R and inspecting the resulting goodness of fits and comparisons with the data, we are able to place lower limits for R or, equivalently, upper limits for the recombination scale Q_r. We use BCDMS [27], HERA [28], and NMC data [29] on lepton-proton DIS in our analysis. Apart from searching for signatures of and placing limits on the effects of recombination from these DIS data, our motivation is to qualitatively compare the systematics of nonlinearities with those resulting from the Balitsky-Fadin-Kuraev-Lipatov (BFKL) resummation [30–33] of $\log x$-type terms relevant at small x. The BFKL resummation was incorporated into the DGLAP equations and the coefficient functions for the DIS structure functions in Ref. [34], and evidence in favor of these dynamics has thereafter been seen in fits to the HERA data [35,36]. In particular, the implementation of the BFKL resummation allows for a larger gluon distribution at the input scale, which leads to a good description of the HERA data on the longitudinal structure function $F_L(x, Q^2)$. One might expect similar features from the gluon recombination as it slows down the scale evolution.

While in the present paper we discuss the recombination in the case of free protons, we note that the nonlinear corrections and partonic saturation are expected to be enhanced in the case of heavier nuclear targets [6,8,15,37,38], the recombination scale increasing naively as $A^{1/6}$ as a function of the nuclear mass number A. As of today, however, no clear signal has been seen. Discovering this enhancement is also one of the cornerstones of the envisioned physics programs of the Electron-Ion Collider (EIC) [39,40] and the Large Hadron Electron Collider (LHeC) [41,42], as well as a strong motivation to run the colliders with both light and heavy nuclei. The construction of a Future Circular Collider (FCC) [43], which would reach kinematics even further beyond those of the LHeC in electron-ion collisions, would naturally be able to put even stronger constraints on the nature and strength of nonlinear corrections to the evolution of PDFs.

The paper is organized as follows: First, Sec. II provides an overview of the theoretical background and the implementation of the Zhu and Ruan nonlinear corrections to the evolution of PDFs, including numerical comparisons between the linear and nonlinear evolution equations. In Sec. III, we present our new global fits of proton PDFs with nonlinear corrections. Section IV discusses the impact on the longitudinal structure function $F_L(x, Q^2)$ in the HERA, EIC, and LHeC kinematics. Finally, we summarize our findings and outline the way forward in Sec. V.

II. NONLINEAR EVOLUTION EQUATIONS

The DGLAP evolution equations are a set of renormalization group equations arising from the renormalization of collinear divergences in ladder-type Feynman graphs describing parton emission in QCD. They take the following standard form [2–5],

$$Q^2\frac{d}{dQ^2}\begin{pmatrix} q_i(x,Q^2) \\ G(x,Q^2)\end{pmatrix} = \frac{\alpha_s(Q^2)}{2\pi}\sum_j\int_x^1\frac{dy}{y}\begin{pmatrix} P_{q_iq_j}\left(\frac{x}{y}\right) & P_{q_ig}\left(\frac{x}{y}\right) \\ P_{gq_j}\left(\frac{x}{y}\right) & P_{gg}\left(\frac{x}{y}\right)\end{pmatrix}\begin{pmatrix} q_j(y,Q^2) \\ G(y,Q^2)\end{pmatrix}, \tag{1}$$

where $q_i(x, Q^2)$ are the PDFs for quarks of flavor i, $G(x, Q^2)$ is the gluon PDF, and $P_{ij}(x/y)$ are the partonic splitting functions. These equations are clearly linear in PDFs. The first nonlinear corrections stem from diagrams like the one in Fig. 1 showing a contribution to the DIS process with two gluons from the proton merging into one, which eventually splits into a quark-antiquark pair coupling to the virtual photon. Here, we will consider the corrections to the evolution of the gluon and quark-singlet distributions,

$$\Sigma(x,Q^2) = \sum_i [q_i(x,Q^2) + \bar{q}_i(x,Q^2)], \quad \text{with} \quad i \in \{u,d,s,c,b,t\}, \tag{2}$$

which can be written as [16,17]

$$x\frac{dG(x,Q^2)}{d\ln Q^2} = \text{linear terms} + \frac{9}{32\pi^2}\left(\frac{1}{RQ}\right)^2 \int_{x/2}^{1/2} dx_1 x x_1 G^2(x_1,Q^2)\sum_i P_i^{GG\to G}(x_1,x) - \frac{9}{16\pi^2}\left(\frac{1}{RQ}\right)^2 \int_{x}^{1/2} dx_1 x x_1 G^2(x_1,Q^2)\sum_i P_i^{GG\to G}(x_1,x), \equiv \text{linear terms} + \Delta G(x,Q^2), \tag{3}$$

$$x\frac{d\Sigma(x,Q^2)}{d\ln Q^2} = \text{linear terms} + \frac{9}{32\pi^2}\left(\frac{1}{RQ}\right)^2 \int_{x/2}^{1/2} dx_1 x x_1 G^2(x_1,Q^2)\sum_i P_i^{GG\to q\bar{q}}(x_1,x) - \frac{9}{16\pi^2}\left(\frac{1}{RQ}\right)^2 \int_{x}^{1/2} dx_1 x x_1 G^2(x_1,Q^2)\sum_i P_i^{GG\to q\bar{q}}(x_1,x), \equiv \text{linear terms} + \Delta\Sigma(x,Q^2), \tag{4}$$

where "linear terms" refer to the right-hand side of the usual DGLAP evolution; see Eq. (1). The square of the gluon PDF $G^2(x_1,Q^2)$ arises due to the modeling of the four-gluon correlation function; see the lower part of Fig. 1. The parameter R with the dimension of length is introduced on the basis of dimensional analysis. In a sense, the factor $1/(RQ)^2$ can be thought to be proportional to the overlap probability of two partons and R regarded as the length scale on which it takes place. A smaller value for the parameter R corresponds to stronger nonlinear effects. The recombination functions in Eqs. (3) and (4) are given by

$$\sum_i P_i^{GG\to G}(x_1,x) = \frac{3\alpha_s^2}{8}\frac{C_A^2}{N^2-1}\frac{(2x_1-x)(-136xx_1^3-64x_1x^3+132x_1^2x^2+99x_1^4+16x^4)}{xx_1^5}, \tag{5}$$

$$\sum_i P_i^{GG\to q\bar{q}}(x_1,x) = \alpha_s^2\frac{T_f}{N(N^2-1)}(2x_1-x)^2\left[\frac{(4x^2+5x_1^2-6x_1x)}{x_1^5}+\frac{N^2}{(N^2-1)}\frac{(4x^2+4x_1^2-6x_1x)}{x_1^5}\right], \tag{6}$$

with $C_A = N = 3$ and $T_f = 1/2$. Unlike the GLR-MQ equations, these evolution equations conserve the PDF momentum sum rule,

$$\frac{d}{dQ^2}\int_0^1 dx x[G(x,Q^2)+\Sigma(x,Q^2)] = 0, \tag{7}$$

because the nonlinear terms satisfy the following conditions separately for the gluon and singlet quark distributions,

$$\frac{d}{d\ln Q^2}\int_0^1 dx\Delta G(x,Q^2) = 0, \tag{8}$$

$$\frac{d}{d\ln Q^2}\int_0^1 dx\Delta\Sigma(x,Q^2) = 0. \tag{9}$$

Note that the linear parts of the full evolution equations also conserve the total momentum but only in the sum of the gluon and quark-singlet PDFs. The method employed in Refs. [16,17] can also be used to derive the quark-antiquark and quark-gluon recombination, but these effects are omitted in the present analysis. While the quark-initiated recombination effects can be included in the nonlinear evolution equations, they are expected to be similarly numerically small as the quark-initiated parton splittings are suppressed compared to the gluon-initiated ones by the ratio of the color factors C_A/T_F or C_A/C_F, depending on the channel ($C_A = 3, C_F = 4/3, T_F = 1/2$). We also note that the gluon and quark-singlet PDFs form a closed set of evolution equations and that the quark flavor dependence enters through different nonsinglet combinations of quarks, where the gluon recombination terms cancel. It is therefore consistent to consider the nonlinear terms only in the evolution equations for gluon and quark-singlet PDFs, Eqs. (3) and (4), while still solving for the full flavor-dependent evolution of PDFs.

We have implemented the nonlinear corrections as an extension to the evolution toolkit HOPPET [24]. This code is optimized to rapidly solve the linear DGLAP equations by converting the numerical convolution of the PDFs with the splitting functions to multiplication of a matrix and a vector, which can be computed extremely fast on modern hardware. To do this, HOPPET first changes the variables from x and Q^2 to $y = \ln(1/x)$ and $t = \ln Q^2$, which helps with numerical accuracy. The PDF of each flavor can be represented numerically by the sum of a set of PDF values

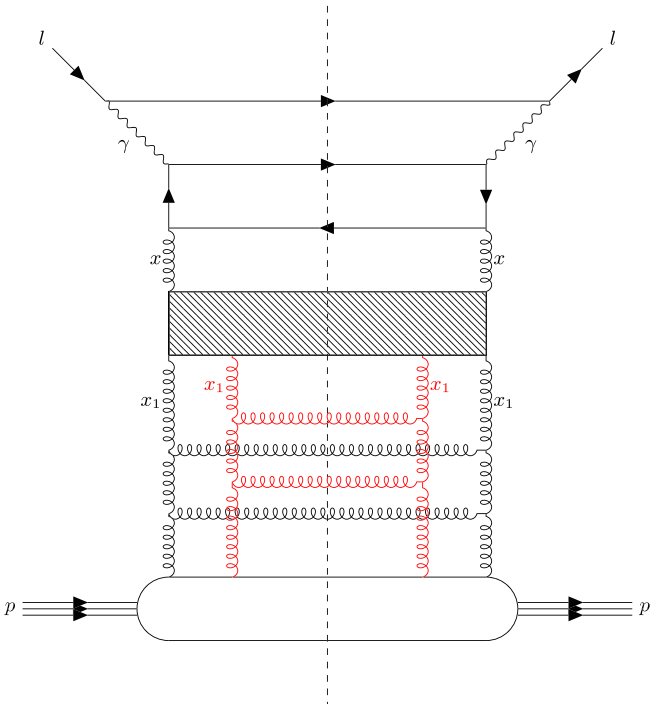

FIG. 1. Typical cut diagram representing the recombination of two gluon ladders into one. Note that the gluon ladders in black and red do not interact. The momentum fractions x and x_1 appearing in Eqs. (3) and (4) are indicated.

at fixed y values y_α, weighted with the interpolation weights $w_\alpha(y)$:

$$q(y,t) = \sum_\alpha w_\alpha(y) q(y_\alpha, t) \equiv \sum_\alpha w_\alpha(y) q_\alpha(t). \tag{10}$$

Using Eq. (10), the convolution $(P \otimes q)$ can then be written in an analogous way,

$$\begin{aligned}(P \otimes q)(y,t) &= \sum_\alpha w_\alpha(y)(P \otimes q)_\alpha(t) \\ &= \sum_\beta \sum_\alpha w_\alpha(y) P_{\alpha\beta}(t) q_\beta(t),\end{aligned} \tag{11}$$

where HOPPET precomputes the values of the matrix $P_{\alpha\beta}(t)$ once and reuses it for all future calculations,

$$P_{\alpha\beta}(t) = \int_{e^{-y_\alpha}}^1 dz P(z,t) w_\beta(y_\alpha + \ln z). \tag{12}$$

To apply this formalism to the nonlinear terms, we treat $G^2(x,Q^2)$ as an independent flavor, which then enters the evolution as a linear term and is updated to be equal to $[G(x,Q^2)]^2$ at each iteration. The integration limits are handled by substituting x with $x/2$ in the first nonlinear term and by using a modified integration method, which integrates only up to $y = \ln(2)$ for the second term. This method allows us to compute the evolution with nonlinear corrections while increasing the computing time by no more than ≈20%. We have also numerically checked our implementation by verifying the momentum sum rule.

In Fig. 2, we illustrate the effects of nonlinearities in comparison to the regular DGLAP evolution and show the ratios between the PDFs evolved using the nonlinear and linear evolution equations as a function of x at different values of Q. Here, we have initialized the evolution with the boundary condition given by the CJ15 proton PDFs [44] at $Q_0 = 1.3$ GeV and performed the evolution in Q^2 according to the linear and nonlinear equations up to different target values of Q. In both cases, the linear terms in the evolution are included up to next-to-next-to-leading order (NNLO) accuracy in the strong coupling constant α_s. The parameter R is set to 0.5, 1, and 2 GeV^{-1} in the top, center, and bottom panels, respectively—we will see in the next section that these are reasonable choices. The Q values are chosen as $Q_0 + \Delta Q$ with $Q_0 = 1.3$ GeV and ΔQ increasing by factors of 100 from 10^{-4} to 10^4 GeV. The dynamically generated suppression at small x ("shadowing") and an enhancement at intermediate x ("antishadowing") are clearly visible for both quark-singlet (right panels) and gluon (left panels) PDFs. At low Q, the effects are generally more significant for the gluon PDF than for the quark singlet, which, in part, results from the fact that the gluon distribution at the initial scale is smaller than the quark-singlet one—very close to zero at small x—and therefore even small changes are more significant in the ratios. As a result, even a tiny step to higher Q causes visible differences in the case of gluons, whereas for the quark singlet a much larger step has to be taken in order to see a difference. At higher values of Q, where both the gluon and quark-singlet PDFs are large at small x, the size of the relative effect becomes roughly equal.

While the relative effects of nonlinearities first increase as Q increases, there is eventually a turnaround, and the differences begin to decrease. This is to be expected as $G^2(x,Q^2)/Q^2 \to 0$ as $Q^2 \to \infty$, so the nonlinear terms die away at asymptotically large Q^2. For gluons, this turnaround happens already around $\Delta Q \approx 1$ GeV, but for quarks, it takes place only at the electroweak scale $\Delta Q \approx 100$ GeV. The dependence on the choice of R works as expected—a smaller R leads to more pronounced effects, although qualitatively there are no significant differences between different choices of R.

III. PDF FITS WITH RECOMBINATION EFFECTS

As was observed in Fig. 2, the effects of recombination can persist up to rather high values of Q^2, where the PDFs are constrained by experimental data. As a result, in order to be consistent with these data constraints, the initial condition at Q_0 has to be iterated. To get a more complete and consistent picture of the impact of nonlinearities, it is

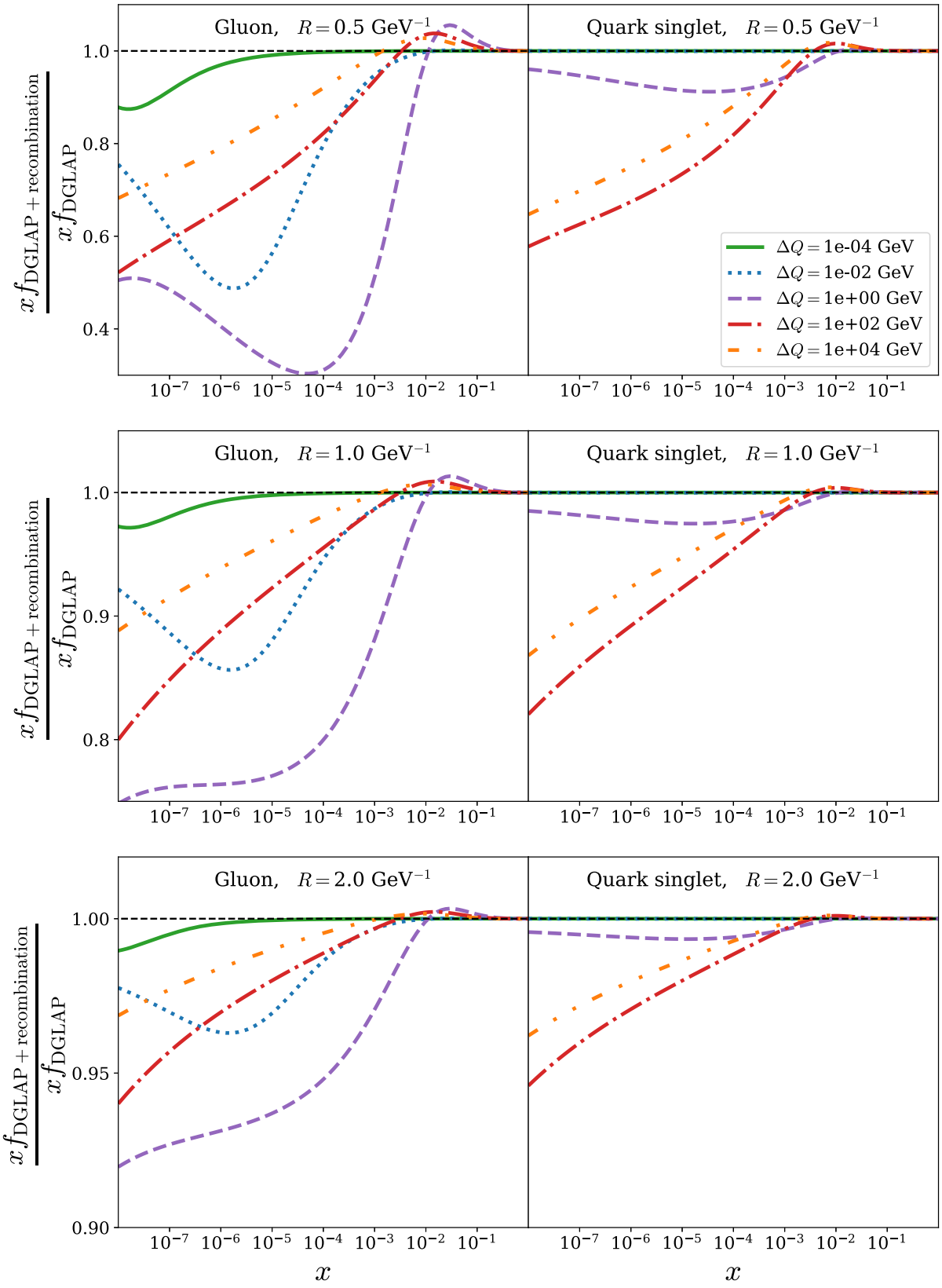

FIG. 2. Ratios between the PDFs evolved using the nonlinear and linear evolution equations with the CJ15 initial condition at $Q_0 = 1.3$ GeV as a function of x and at different values of $Q = Q_0 + \Delta Q$. The left column corresponds to the gluon PDF, while the right column shows the quark-singlet distribution. The top, middle, and bottom rows correspond to $R = 1/2$, 1, and 2 GeV^{-1}, respectively.

thus necessary to perform global PDF fits including the nonlinear terms in the evolution. In addition to the corrections to the evolution equations, higher-twist corrections also exist for the DIS coefficient functions like the ones for the structure function F_2 [18]. However, to the best of our knowledge, the complete expressions for all the needed structure functions consistent with the nonlinear evolution equations used in the present analysis are not available in the literature. Therefore, we are not able to directly assess their impact. In Ref. [18], the authors performed a numerical analysis of the twist-4 contributions to the Q^2 slope of F_2, finding that with $R \gtrsim 1.41/\mathrm{GeV}$ the effects were significant only at $x \lesssim 10^{-5}$. However, it is difficult to uniquely judge the relevance of higher-twist contributions based on that particular study only.

A. Fitting framework

The framework used to perform our global PDF fits with nonlinear corrections is based on the XFITTER package [25,26], which we have extended by adding our modified version of HOPPET as a new module; see Sec. II. We have verified that this reproduces the results obtained with the default QCDNUM [45] evolution in the limit $R \to \infty$. The fits are performed at the NNLO accuracy in both evolution and scattering coefficients. We use the RTOPT general-mass variable-flavor-number scheme [46] for the heavy quark coefficient functions. In order to study the dependence of the nonlinear effects on the assumed parametrization, we have prepared PDF fits with two different parametrizations for gluons and also varied the initial scale. We refer to these as parametrizations 1 and 2. For parametrization 1, the PDFs are parametrized at the initial scale $Q_0 = 1.0$ GeV using the same parametric form as employed in the HERAPDF2.0 PDF analysis [28]: For the valence distributions u_v, d_v, and the sea-quark densities $\bar{u}, \bar{d}$, we use the form

$$xf(x) = A_f x^{B_f}(1-x)^{C_f}(1 + D_f x + E_f x^2), \quad (13)$$

and the gluon PDF is parametrized as

$$xg(x) = A_g x^{B_g}(1-x)^{C_g} - A'_g x^{B'_g}(1-x)^{C'_g}. \quad (14)$$

When using the parametrization in Eq. (14), the gluon distribution tends to become negative at small values of x and Q^2. In the case of parametrization 2, we apply the same parametrization for gluons as we have for quarks in Eq. (13), which guarantees that the small-x gluons will remain positive at the initial scale. To partly compensate for this restriction, we increase the initial scale to $Q_0 = 1.3$ GeV for parametrization 2. The strange quark is set to $xs = x\bar{s} = (2/3)x\bar{d}$ at Q_0 in both cases.

The normalization parameters A_{u_v}, A_{d_v}, and A_g are fixed through the quark-number and momentum sum rules. For simplicity and for lack of constraints for the flavor separation, we fix $A_{\bar{u}} = A_{\bar{d}}$ and $B_{\bar{u}} = B_{\bar{d}}$. With parametrization 1, we fix $C'_g = 25$, where the exact value is unimportant as long as $C'_g \gg C_g$ such that the negative term does not contribute beyond low x. The full set of 14 open PDF parameters in the case of parametrization 1 is

$$\{B_g, C_g, A'_g, B'_g, B_{u_v}, C_{u_v}, E_{u_v}, B_{d_v}, C_{d_v}, C_{\bar{u}}, D_{\bar{u}}, A_{\bar{d}}, B_{\bar{d}}, C_{\bar{d}}\}, \quad (15)$$

and in the case of parametrization 2,

$$\{B_g, C_g, D_g, E_g, B_{u_v}, C_{u_v}, E_{u_v}, B_{d_v}, C_{d_v}, C_{\bar{u}}, D_{\bar{u}}, A_{\bar{d}}, B_{\bar{d}}, C_{\bar{d}}\}. \quad (16)$$

Any parameters which are not explicitly mentioned are set to zero as in the HERAPDF2.0 analysis because the additional BCDMS and NMC data do not provide qualitatively new constraints on the fitted PDFs. The parameter R is kept fixed in each fit, and the procedure is repeated for R values from $0.2\ \mathrm{GeV}^{-1}$ to $3.0\ \mathrm{GeV}^{-1}$ in steps of $0.01\ \mathrm{GeV}^{-1}$. For $R \gtrsim 3.0\ \mathrm{GeV}^{-1}$, the nonlinear modifications were found to be negligible.

We find the optimal set of parameters by minimizing the χ^2 function defined as

$$\chi^2 = \sum_i \frac{[T_i - \sum_\alpha \gamma^i_\alpha D_i b_\alpha - D_i]^2}{(\delta_{i,\mathrm{stat}}\sqrt{D_i T_i})^2 + (\delta_{i,\mathrm{uncorr}} T_i)^2} + \sum_\alpha b^2_\alpha. \quad (17)$$

The theoretical predictions for data point i and a given set of PDF parameters are represented by T_i. The measured data values are given by D_i, and $\delta_{i,\mathrm{stat}}, \delta_{i,\mathrm{uncorr}}$, and γ^i_α are the statistical, uncorrelated, and correlated uncertainties associated with each data point. Minimizing this χ^2 function with respect to the 14 fit parameters and the nuisance parameters b_α, we determine the central set of PDFs and systematic shifts, $-\sum_\alpha \gamma^i_\alpha D_i b_\alpha$, which give the best description of the data. The uncertainties of the obtained PDFs are then determined using the Hessian method [47], which is based on a quadratic expansion of the χ^2 function around its minimum, $\chi^2 = \chi^2_{\mathrm{min}} + \Delta\chi^2$. The choice of the maximum displacement $\Delta\chi^2$, the tolerance, depends on the definition of uncertainties, and in PDF fits, it is usually conservatively chosen to be $\Delta\chi^2 \gg 1$. Indeed, assuming that the data points follow a Gaussian distribution around the "truth," the expected χ^2/N_{dof} should be less than $\sim$1.05 in 90% of the cases with $\sim$1600 data points, which is the amount of points we are now considering. However, the minimum values we find are $\chi^2/N_{\mathrm{dof}} \sim 1.18$; i.e., the ideal choice $\Delta\chi^2 = 1$ is, perhaps, an underestimate. Assuming a Gaussian probability density for the fit parameters results in $\Delta\chi^2_{\mathrm{max}} \approx 20$ at the 90% confidence level; see, e.g., Sec. A.3 of Ref. [48] or Ch. 15.6 of Ref. [49]. In practice, we determine the PDFs with the tolerance set to $\Delta\chi^2 = 1$ and

then rescale the resulting uncertainties by $\sqrt{20} \approx 4.5$ to match the actual tolerance since this leads to a better convergence of the XFITTER algorithms.

B. Data selection

Our analysis, like the ones presented in Refs. [50] and [51], includes the lepton-proton DIS data from the BCDMS, HERA, and NMC experiments, which are summarized in Table I. The data from BCDMS and NMC are given in terms of the structure function $F_2^p(x, Q^2)$ measured in neutral-current (NC) deep inelastic muon scattering. The HERA data sets provide the reduced cross section of DIS in electron-proton and positron-proton collisions, both for charged-current (CC) and NC processes. We do not include the HERA data on the longitudinal structure function F_L [52] since it is derived from the other HERA data sets, and including it would, to some extent, double-count some constraints. We will, however, discuss F_L separately in Sec. IV. The data on heavy-quark production from HERA are omitted as well [53–55].

In the first round of fits, we apply the same $Q^2 > 3.5$ GeV^2 cut as, e.g., in the HERAPDF2.0 analysis, but with parametrization 1, we also perform a second round of fits with the cut lowered to $Q^2 > 1.0$ GeV^2 to see whether the inclusion of the $1/Q^2$ suppressed terms in the evolution equation can provide an improved description of the data in the low-Q^2 region. Indeed, it is known that in NNLO fits, $\chi^2/N_{\rm dof}$ tends to get systematically worse as the minimum Q^2 is lowered from 15 GeV^2 to 2 GeV^2 [35,36]—a tension that the BFKL resummation appears to ease. The two rightmost columns of Table I indicate the number of data points per experiment which pass these cuts. The total number of data points is 1568 for the higher-Q^2 cut and 1636 for the lower-Q^2 one.

C. Fit results

The resulting values of $\chi^2/N_{\rm dof}$ minima for all the fits with different R values are shown in Fig. 3, and the contributions of different data sets to the total $\chi^2/N_{\rm dof}$ are presented in Appendix A—Tables II–IV. In Fig. 3, the orange (parametrization 1) and green (parametrization 2) crosses correspond to the $Q^2 > 3.5$ GeV^2 cut, and the blue ones (parametrization 1) are for the fits with $Q^2 > 1.0$ GeV^2. We have not performed a fit with parametrization 2 with the lower cut, as it lies below the initial scale of the PDF evolution in this case.

In the case of parametrization 1, the resulting values of $\chi^2/N_{\rm dof}$ generally decrease with increasing R, i.e., with a decreasing strength of the nonlinear corrections, but remain constant for $R > 1.5$ GeV^{-1}. Lowering the value of R down to ≈ 0.7 GeV^{-1} causes only a small increase, but below that, $\chi^2/N_{\rm dof}$ rapidly starts diverging. Comparing the curves between the two different cuts, there is no qualitative difference in the R dependence, but the lower cut shifts $\chi^2/N_{\rm dof}$ upwards by approximately 0.04 units across all values of R. While this increase in $\chi^2/N_{\rm dof}$ from the lowered cut might not seem too significant at first sight, one has to consider the small number of added data points that causes it. Namely, the 68 added data points increase the total χ^2 by 145, i.e., more than 2 units per point. Therefore, the nonlinear corrections do not help to alleviate the tension with lower-Q^2 data. In the case of parametrization 2 (green crosses), i.e., with a positive-defined gluon distribution at the initial scale, the best descriptions are obtained in the

TABLE I. Summary of DIS data sets used in out fits of proton PDFs with nonlinear corrections. The last two columns show the number of data points passing the $Q^2 > 3.5$ GeV^2 and $Q^2 > 1$ GeV^2 cuts, respectively. The energies listed for the HERA measurements refer to the proton beam energy, while those listed for BCDMS refer to the muon beam energy. NMC uses the same energies as BCDMS but combines them into one data set.

Experiment	Data set	Year	Reference	$N_{\rm data}^{Q^2>3.5\ {\rm GeV}^2}$	$N_{\rm data}^{Q^2>1.0\ {\rm GeV}^2}$
BCDMS	NC μ F_2^p 100 GeV	1996	[27]	83	83
	NC μ F_2^p 120 GeV			91	91
	NC μ F_2^p 200 GeV			79	79
	NC μ F_2^p 280 GeV			75	75
HERA	NC e^+ 920 GeV	2015	[28]	377	425
	NC e^+ 820 GeV			70	78
	NC e^+ 575 GeV			254	260
	NC e^+ 460 GeV			204	210
	NC e^- 920 GeV			159	159
	CC e^+ 920 GeV			39	39
	CC e^- 920 GeV			42	42
NMC	NC μ F_2^p	1997	[29]	95	95
	Total			1568	1636

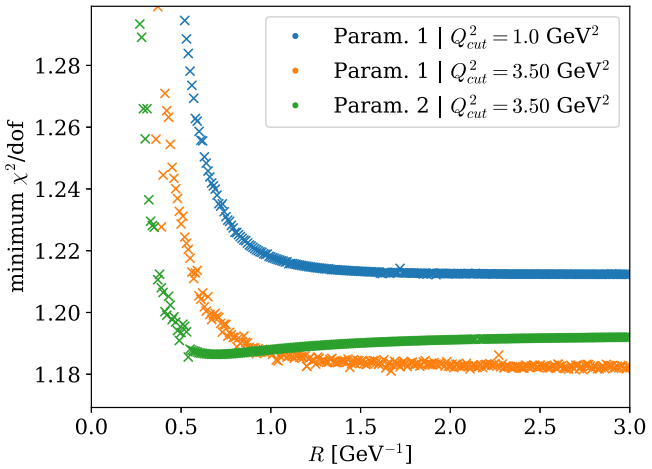

FIG. 3. Minimum $\chi^2/N_{\rm dof}$ as a function of R for parametrization 1 with $Q^2 > 3.5$ GeV^2 (orange) and $Q^2 > 1$ GeV^2 (blue), and parametrization 2 with $Q^2 > 3.5$ GeV^2 (green).

interval $0.5 < R < 0.7$ GeV^{-1}, although the difference with respect to $R = 3$ GeV^{-1} is small. The fact that the $\chi^2/N_{\rm dof}$ obtained in this region is lower than at high R, but is still higher than parametrization 1 at $R \to \infty$, implies that it is not due to the physical effect of recombination allowing for a better description of the data. Instead, it means that parametrization 2 is lacking sufficient flexibility to give an optimal description of the data, which is then alleviated by the introduction of another parameter, i.e., R.

Tables II–IV in Appendix A break down the obtained χ^2 into contributions from different data sets. Tables II and IV, which correspond to the fits with parametrization 1 at $Q^2_{\rm cut} = 3.5$ GeV^2 and 1 GeV^2, respectively, show that the preference towards higher values of R is a feature generally shared by all the data sets. At the same time, some data sets are more sensitive to R than others, with the NMC data set showing the largest difference in $\chi^2/N_{\rm dof}$. Table III shows that with parametrization 2, the $\chi^2/N_{\rm dof}$ profiles as a function of R are flatter than in the case of parametrization 1.

In the following, we discuss proton PDFs determined from the fits with the $Q^2 > 3.5$ GeV^2 cut for a representative selection of R values of $\{0.4, 0.6, 0.8, 1, 2, 3\}$ GeV^{-1}. Figures 4 and 5 show the resulting gluon (left) and quark-singlet (right) distributions at $Q = \{Q_0, 2$ GeV, 10 GeV, 100 GeV$\}$ for parametrizations 1 and 2, respectively. They are compared to the HERAPDF2.0 results given by the black dashed curve. To avoid visual clutter, the uncertainties (90% confidence level) are only shown for the HERAPDF2.0 and $R = 3.0$ GeV^{-1} cases, but the remaining uncertainties of the rest closely resemble the latter one. The 90% confidence level for HERAPDF2.0 is obtained by multiplying the uncertainties, which are given at 68% confidence level with a factor of 1.645.

In the case of parametrization 1, the fit with the highest value of R, i.e., the one with smallest nonlinear effects, behaves similarly to the HERAPDF2.0 reference but, despite the added data from BCDMS and NMC, has significantly larger uncertainties due to the higher Hessian tolerance—20 for our fit vs. 2.71 for HERAPDF2.0 at the 90% confidence level. The differences in the central values can, at least partly, be attributed to the fact the we included these additional data, which may affect the PDFs even at low x due to the sum rules. As the value of R is lowered, the shapes of the fitted PDFs change progressively further away from the baseline with $R = 3.0$ GeV^{-1}. In particular, one observes a significant suppression of the gluon and quark-singlet PDFs at small x and $Q \geq 2$ GeV as R decreases. This suppression becomes weaker as the scale Q is increased due to the weakening of the nonlinear effects by a factor of $1/Q^2$ in the evolution equations, but it is still visible even at $Q = 100$ GeV. As expected, the largest differences between the PDFs corresponding to different values of R are visible at the initial scale $Q_0 = 1$ GeV, though they carry little meaning since there are no data that constrain this region due to the $Q^2_{\rm cut} > 3.5$ GeV^2 cut.

A comparison of the results presented in Figs. 4 and 5 shows that the systematics of the R and Q dependence of the PDFs obtained using parametrizations 1 and 2 are very similar even though there are large differences around the initial scale Q_0. This gives us confidence that our conclusions are not very sensitive to the details of the parametrization. An interesting detail in the fits with parametrization 2 is that the gluon PDF of the $R = 0.4$ GeV^{-1} fit turns negative at $Q = 2$ GeV even though it is positive at the initial scale. While it is counterintuitive that recombination could decrease the PDFs from positive to negative values, this behavior can be explained by the convolution integral in Eq. (3). Since the upper bound of the integral is $1/2$, the PDF at a given x value is affected by larger x values, where the PDF is still positive and can therefore be further reduced even when it is zero at this x. In this case, the nonlinear terms dominate the evolution and, as discussed in the Introduction, this is an indication of exceeding the limits of the considered approach terminating the power-corrected expression at a twist-4 level.

The parameters of the fits are summarized in Tables V–VII in Appendix B.

IV. IMPACT ON THE LONGITUDINAL STRUCTURE FUNCTION

The longitudinal structure function F_L carries a direct sensitivity to the gluon PDFs and should therefore have an increased sensitivity to the nonlinear effects. We note that since the HERA data entered our analysis as reduced cross sections $\sigma_r = F_2 - (y^2/Y_+)F_L$, where $Y_+ = 1 + (1-y)^2$, $y = s/xQ^2$, s being the squared center-of-mass energy of the lepton-proton collision, they already carry information of F_L. In addition, the Q^2 dependence of F_2 is directly sensitive to F_L [56], particularly at small x. In this sense,

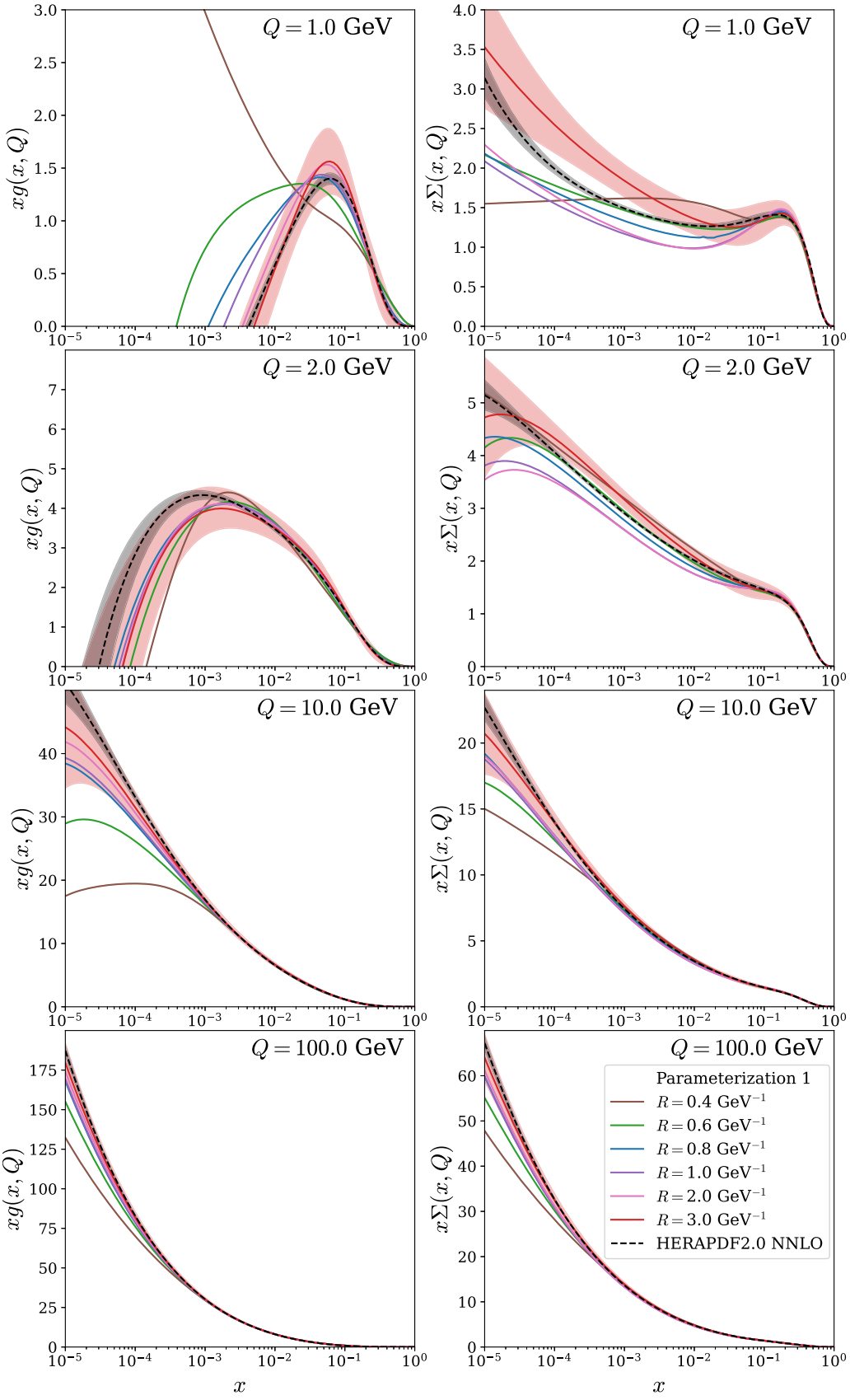

FIG. 4. PDFs resulting from fits with parametrization 1 at various R values. The left column shows the gluon PDF, while the right side shows the quark-singlet PDF. The rows correspond to different scales Q.

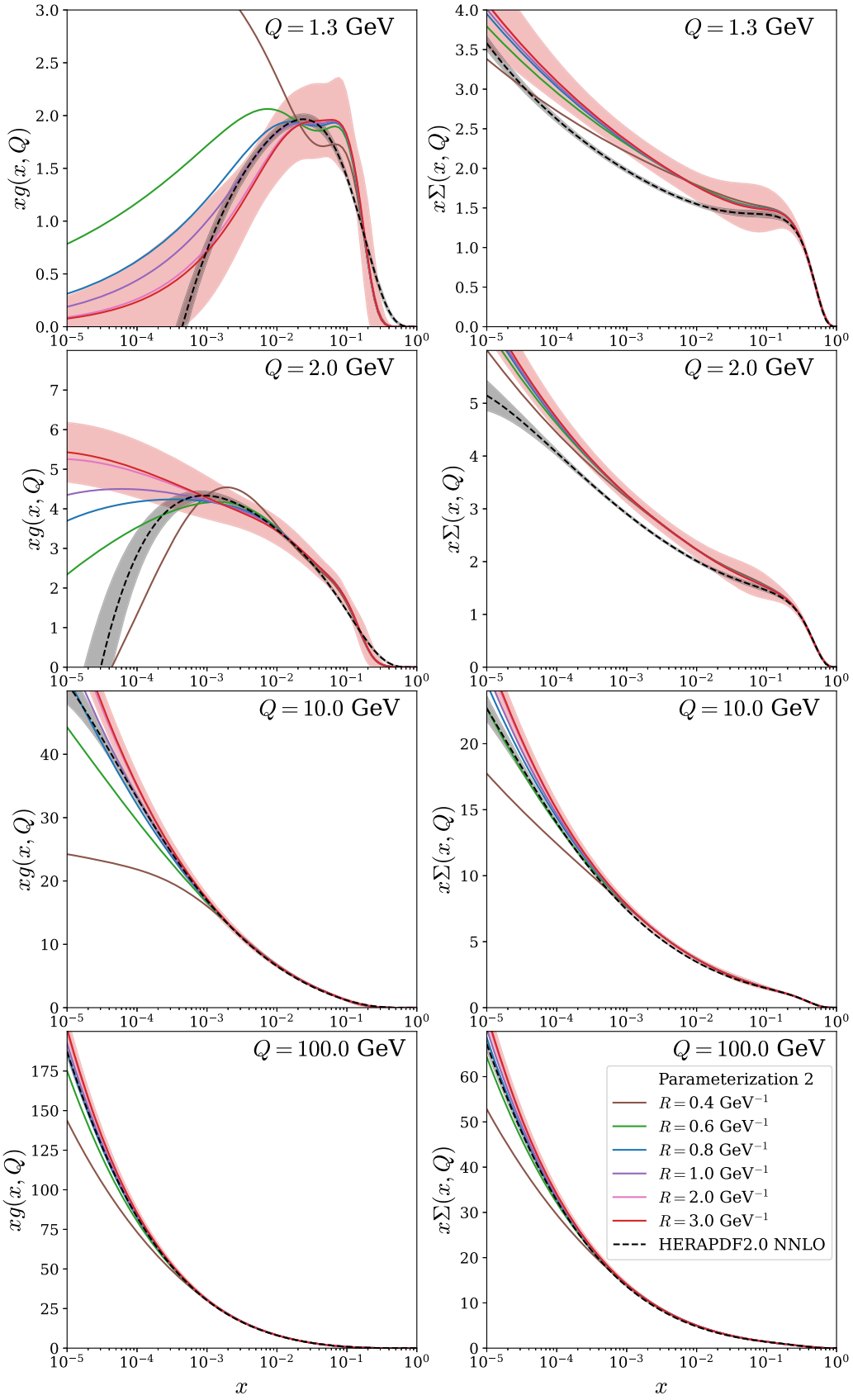

FIG. 5. Same as Fig. 4, but showing PDFs fitted with parametrization 2.

the discussion below is not entirely independent of the fits presented above.

A. Longitudinal structure function at HERA

To gain a deeper understanding of the sensitivity of the HERA data to nonlinear corrections, we compare the values for the longitudinal structure function $F_L(x, Q^2)$ that result from our PDFs with the data taken by H1 [52] and ZEUS [57]. Figure 6 presents the HERA F_L measurements at different values of Q^2 and integrated over x bins. The left panel shows the data compared with the predictions using the PDFs fitted at $Q^2_{\text{cut}} = 3.5$ GeV^2 with parametrization 1 and the right panel with parametrization 2. For $Q^2 \gtrsim 100$ GeV^2, the predictions converge and become indistinguishable. Towards lower Q^2 values, the predictions spread out with higher R values corresponding to slightly larger values of F_L. At low Q^2, the behavior depends on the parametrization. In the case of parametrization 2, the predictions converge again towards $F_L \to 0$ at the PDF parametrization scale $Q^2_0 = 1.69$ GeV^2. The predictions for parametrization 1 do not converge towards $F_L \to 0$ since the parametrization allows the PDFs to be negative and therefore yield negative F_L values for $Q^2 \lesssim 5$ GeV^2. At $Q^2 \lesssim 10$ GeV^2, the NNLO predictions generally lie below the values for F_L measured by H1, as has been observed in other studies [35,36]. On the other hand, the ZEUS measurements generally lie below the predictions but come with larger uncertainties.

Interestingly, the systematics of nonlinearities we find goes in the opposite direction in comparison to the effects of small-x BFKL resummation: While the recombination effects tend to decrease F_L, the small-x BFKL resummation has a tendency to increase it. This can be traced back to the small-x behavior of gluon and quark-singlet PDFs, which are increased in the fits applying the BFKL approach in comparison to the fixed-order NNLO result. In our case, parametrizations 1 and 2 also correspond to the gluon and quark-singlet distributions at the initial scale, which at small x are larger than the one corresponding to our reference PDFs with $R = 3$ GeV^{-1}. However, as we discussed above, the nonlinear Q^2 evolution rapidly decreases our PDFs, especially in the gluon case; see Figs. 4 and 5. It leads to predictions for F_L which lie somewhat below those made using the HERAPDF2.0 PDFs with parametrization 1, and similarly with parametrization 2, a smaller value for R results in a decrease in F_L at small values of x and Q^2.

Figure 7 shows F_L as a function of x at different values of Q^2. The theoretical calculations are performed with the PDFs corresponding to fits with different values of R using parametrization 2. Predictions for F_L made with PDFs from parametrization 1 result in negative values of F_L at $Q^2 \leq 10$ GeV^2 due to the gluon PDFs not being restrained to positive values at the initial scale. Since the negativity of F_L seen in the previous figure indicates that the corresponding theoretical setup is not entirely consistent, e.g., too low Q^2 or too low R, we omit the PDFs fitted with parametrization 1 from the following discussion. For clarity, we show the uncertainty band only for the $R = 3$ GeV^{-1} case. One can see from the figure that for $x \gtrsim 0.3 \times 10^{-2}$, the predictions with different R are almost indistinguishable. However, for $Q^2 > 5$ GeV^2 and towards lower x values, the calculations with a smaller R lead to significantly reduced F_L. This was to be expected as both gluon and quark-singlet PDFs at small x were found to be increasingly suppressed as R decreased; see Figs. 4 and 5. At lower values of Q^2, where the relative differences between predictions corresponding to different R values are expected to be the largest, the absolute values are too close to zero to see any meaningful differences. As already observed in Sec. II, the effects of nonlinearities die out rather slowly at small x as Q^2 increases. In addition, we see in Fig. 7 that the differences

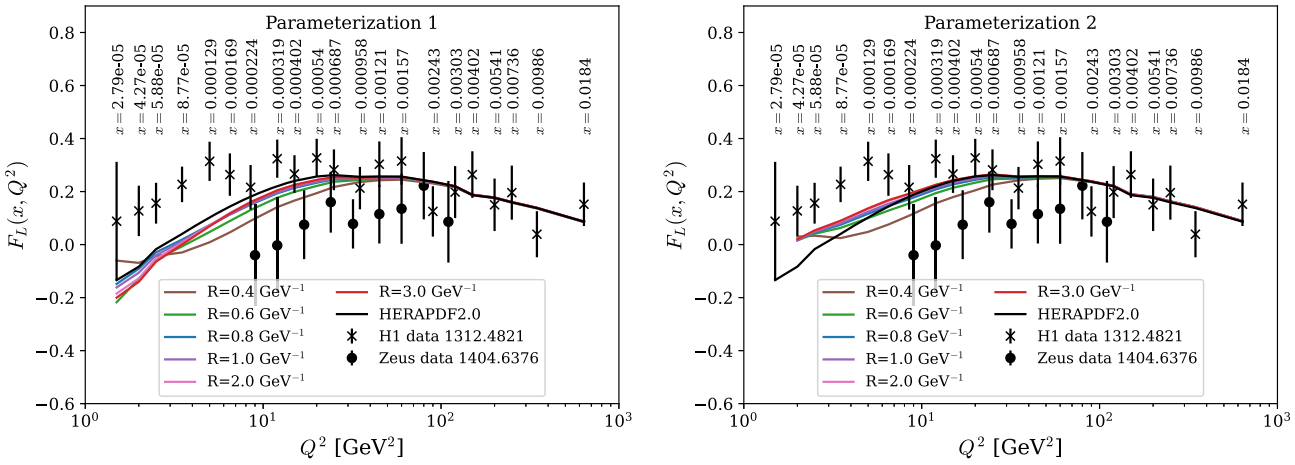

FIG. 6. Structure function F_L calculated with PDFs resulting from different choices of R compared with the measurements by H1 [52] and ZEUS [57]. We only show predictions for $Q^2 > Q^2_0$.

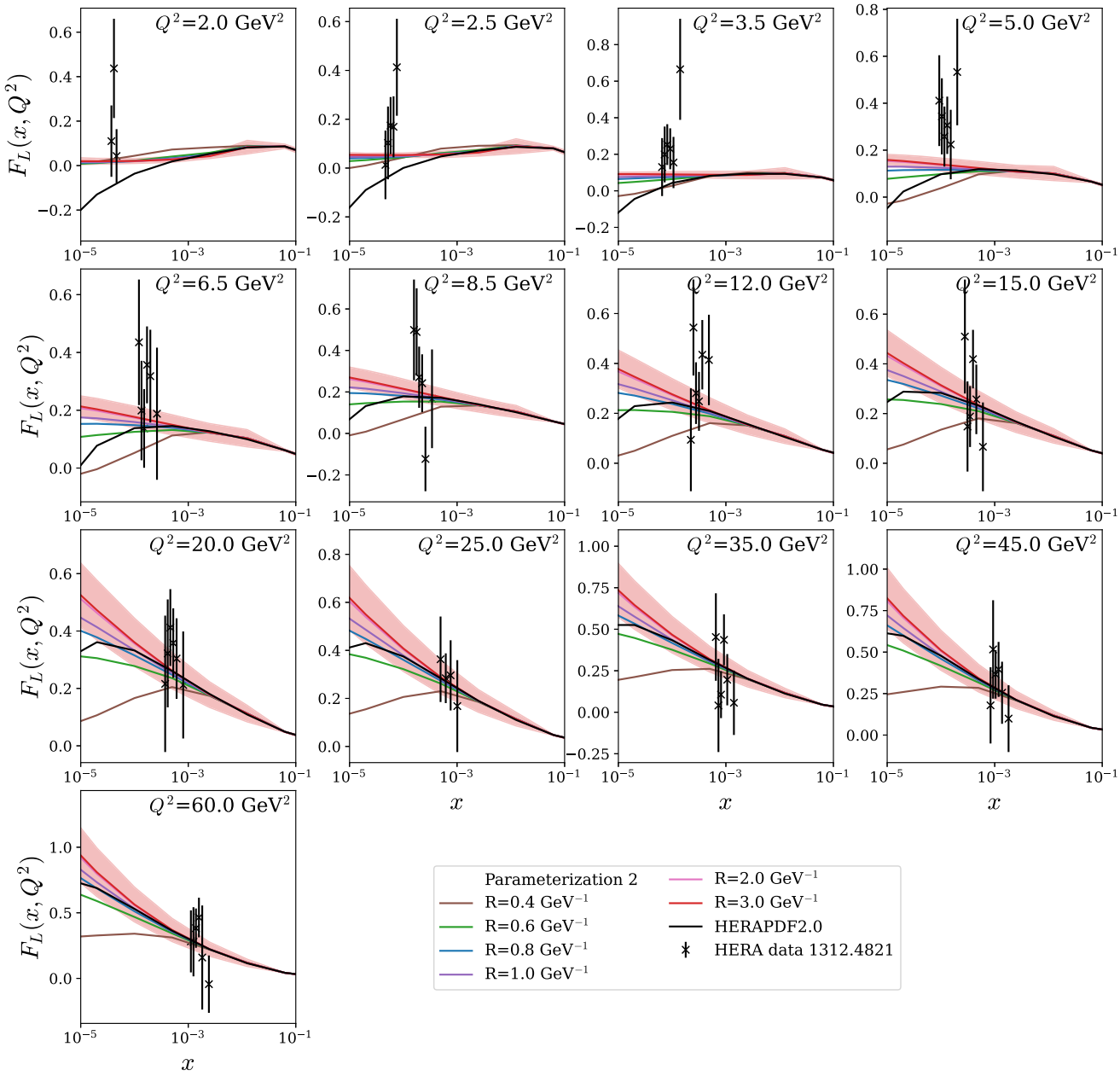

FIG. 7. Structure function F_L resulting from PDFs fitted with parametrization 2 and different R values compared with the HERA F_L measurements [52]. We only show predictions for $Q^2 > Q_0^2$.

between calculations with different R actually still increase at $x \approx 10^{-5}$ until the highest considered Q^2. We note that at $Q^2 > 8.5$ GeV^2, the F_L data are well described by most of the predictions. However, with $R \leq 0.4$ GeV^{-1}, the description begins to visibly deteriorate even at higher values of Q^2.

B. Prospects for the longitudinal structure function at EIC and LHeC

The observations made above naturally raise the question of how much better future DIS experiments such as EIC [39] and LHeC [41,42] could constrain the presence of nonlinearities. To this end, we compare the spread of F_L predictions by taking the ratio of the prediction resulting from each set of PDFs determined at a different value of R over that at $R = 3$ GeV^{-1}:

$$R_{F_L}(x, Q^2) \equiv \frac{F_L(x, Q^2)}{F_L^{R=3\ \mathrm{GeV}^{-1}}(x, Q^2)}. \tag{18}$$

These ratios are compared to the projected relative uncertainties for F_L taken from the EIC development documents [58] and the LHeC White Paper [41] in Figs. 8 and 9, respectively, which are shown as black error bars. Again,

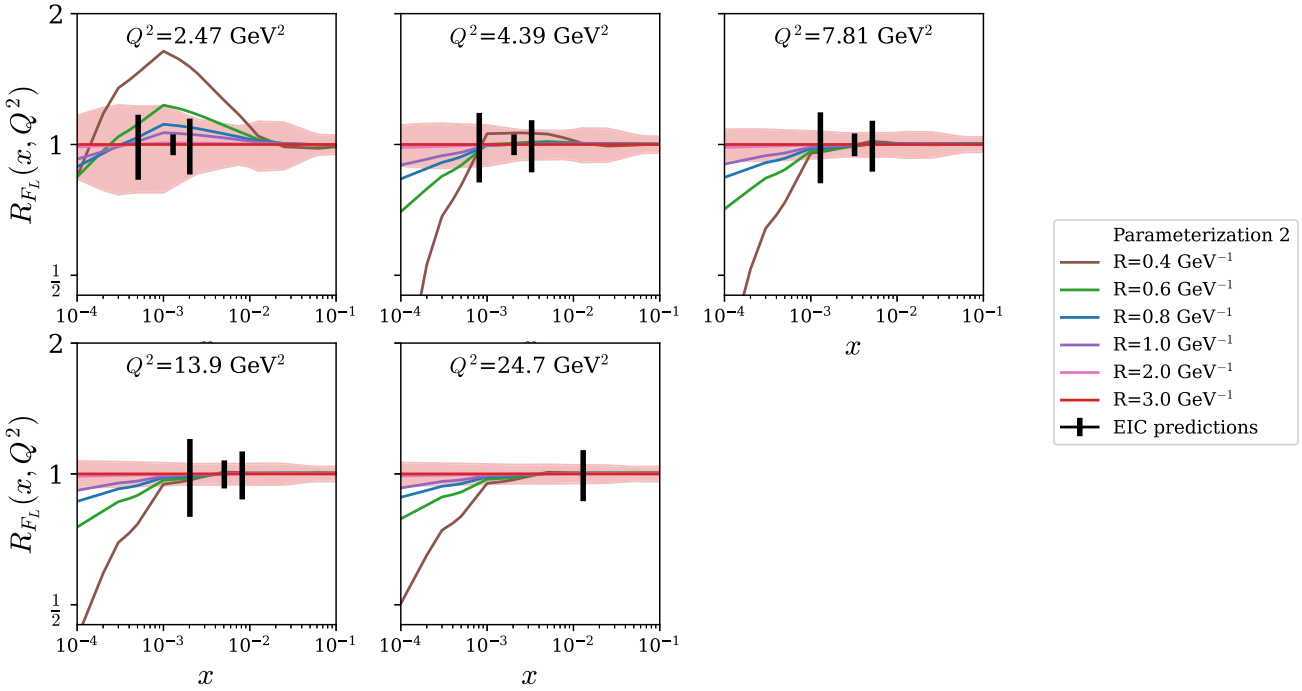

FIG. 8. Ratios of the longitudinal structure functions corresponding to different values of R to that with $R = 3$ GeV^{-1} [see Eq. (18)] as a function of x. The calculations use parametrization 2 and are carried out at different values of Q^2. The vertical bars represent the expected statistical uncertainty of the corresponding measurements at the EIC [58]. We only show predictions for $Q^2 > Q_0^2$.

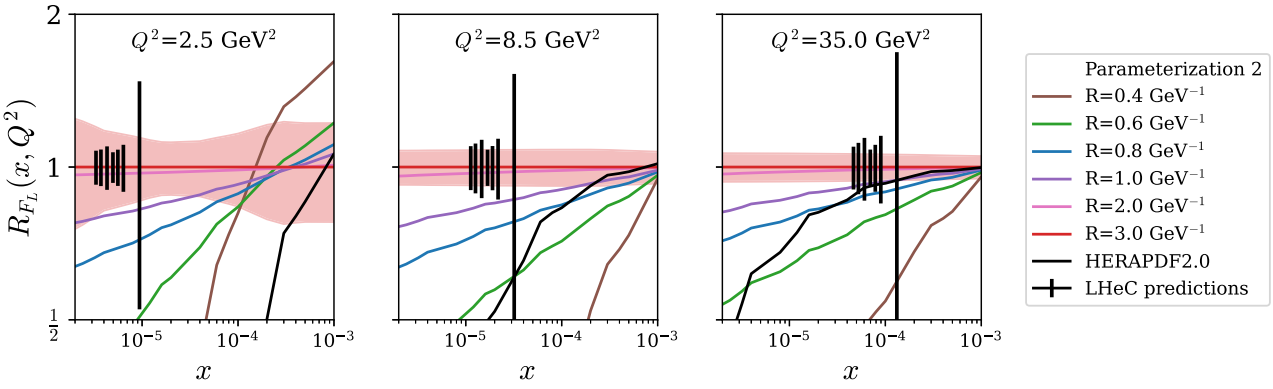

FIG. 9. Same as Fig. 8, but for LHeC [41].

we only show the results for parametrization 2 due to the negativity of F_L when calculated with the PDFs from parametrization 1. This allows us to estimate whether F_L measurements at these future experiments can help to constrain the size of nonlinear effects under the assumption that F_L can be accurately reproduced by the theory, e.g., by adding small-x resummation as in Refs. [35,36]. The projected EIC data only reach the x region, where the different R values are clearly distinguishable at $Q < 2.5$ GeV^2, but even at $Q = 2.47$ GeV^2, the uncertainties of the pseudodata are barely smaller than the spread between predictions with $R > 0.6$ GeV^{-1}. In the case of LHeC, however, the reach towards lower x values should lead to more stringent constraints for the nonlinear effects since the different R values lead to vastly different predictions in this kinematic region.

V. CONCLUSIONS AND OUTLOOK

We have presented numerical studies of the nonlinear gluon recombination corrections to the DGLAP evolution as derived by Zhu and Ruan, which improve upon the commonly used GLR-MQ equations and restore the PDF momentum sum rule. We have extended the HOPPET and XFITTER toolkits to account for these corrections. With these extensions, we performed several fits of proton PDFs to BCDMS, HERA, and NMC data on DIS at NNLO accuracy varying the dimensionful parameter R, which

controls the strength of the recombination effects. We examined two different parametrizations for the gluon PDF—one corresponding to HERAPDF2.0, which allows the gluon PDF to become negative at the parametrization scale, and another which is positive definite at the parametrization scale. We found that both parametrizations result in a similarly good description of the data, $\chi^2/N_{\rm dof} \sim 1.19$, for $R > 1$ GeV^{-1}, but that the description begins to eventually deteriorate towards small R. For the parametrization allowing for a negative gluon PDF, this deterioration sets in for $R \lesssim 0.7$ GeV^{-1}, but for the positive-definite parametrization at somewhat lower values, $R \lesssim 0.5$ GeV^{-1}. In both cases, $R < 0.4$ GeV^{-1} seems to be excluded, which translates to an upper limit for the recombination scale $Q_r = 1/R \lesssim 2.5$ GeV. We also studied the sensitivity of our fits to the choice of $Q^2_{\rm cut}$: Lowering $Q^2_{\rm cut}$ from 3.50 GeV^2 to 1 GeV^2 leads to an increase of $\chi^2/N_{\rm dof}$, which the nonlinear corrections are unable to tame. In other words, we do not find evidence for the presence of nonlinear effects in the used DIS data but can set an upper limit for the parameter controlling their strength, $R \gtrsim 0.4$ GeV^{-1} or $Q_r = 1/R \lesssim 2.5$ GeV. Our conclusions about HERA data not supporting nonlinear effects is in line with a recent study made in Ref. [59]. We also compared our PDF fits with the longitudinal structure function F_L available from HERA. While the ZEUS data lie below the predictions from our fits, the large uncertainties still make the two compatible. However, the H1 data, which reaches lower Q^2 values and has smaller uncertainties, lies well above the predictions at low Q^2. Nonlinear effects in the evolution do not help us to resolve this tension between NNLO calculations and the data, indicating that other theoretical ingredients, such as small-x resummation, are necessary to describe these data properly. Assuming that this tension can be resolved, we show that the improved statistics of the EIC, and particularly the reach towards lower x values at the LHeC, could provide significantly stronger constraints on nonlinear effects.

In general, after a short interval of Q^2 evolution, the nonlinear effects tend to reduce both gluon and quark PDFs at small values of x in comparison to the case with no recombination. This suppression becomes more pronounced as the parameter R decreases, and the effects can persist up to $Q \gtrsim 100$ GeV. This behavior was not sensitive to the form of the initial parametrization and thereby appears to be a general feature of the nonlinear effects. This is opposite to the BFKL resummation which tends to increase the small-x PDFs at large Q^2 when refitting PDFs to the HERA data. Nevertheless, the fact that the Q^2 evolution does not wash away all the possible remnants of nonlinearities even at $\mathrm{Q} = \mathcal{O}(100\ \mathrm{GeV})$ means that some effects could be seen in proton-proton collisions at the LHC as well. For example, the direct photon and heavy-flavor production in the forward direction are sensitive to PDFs at small x, and some visible differences could be induced from the fits with different R. Charting such effects is one possible way to make use of the results of the present analysis. One can also extend the analysis itself by, e.g., including the contributions of two-gluon initiated processes in the coefficient functions as well. For example, the recombination contributions enter the coefficient functions of F_L at $\mathcal{O}(\alpha_s^2)$; i.e., they are of the same perturbative order as the leading-twist NLO terms. As these leading-twist NLO terms turn out to be rather significant for F_L, the inclusion of higher-twist terms could potentially change the picture in the case of F_L. For F_2, the higher-twist terms are of the same perturbative order as the leading-twist NNLO terms. Future precision studies could also include the quark recombination terms that were omitted in this analysis. In addition, one could include a wider range of data sets in the fit, e.g., from the LHC, and/or generalize the framework to the case of heavier nuclei with the expectation of observing enhanced effects of recombination.

The tools developed in this work and LHAPDF files [60] for the fitted PDFs are available in [61].

ACKNOWLEDGMENTS

Our work was supported by the Academy of Finland, Projects No. 331545 and No. 330448, and was funded as a part of the Center of Excellence in Quark Matter of the Academy of Finland, Project No. 346326. This research is part of the European Research Council Project No. ERC-2018-ADG-835105 YoctoLHC. We acknowledge grants of computer capacity from the Finnish Grid and Cloud Infrastructure (persistent identifier urn:nbn:fi:research-infras-2016072533).

APPENDIX A: OBTAINED χ^2 VALUES FOR EACH DATA SET

Tables II–IV show the values of $\chi^2/N_{\rm dof}$ obtained in the different sets of PDF fits for a selection of R values $R = \{0.4, 0.6, 0.8, 1.0, 2.0, 3.0\}$ GeV^{-1}. Table II shows the values obtained in the fits with parametrization 1 and $Q^2_{\rm cut} = 3.5$ GeV^2. Table IV shows the same for $Q^2_{\rm cut} = 1.0$ GeV^2 and Table III for parametrization 2 with $Q^2_{\rm cut} = 3.5$ GeV^2.

TABLE II. Values of $\chi^2/N_{\rm dof}$ obtained in the fits with parametrization 1 and $Q^2_{\rm cut} = 3.5$ GeV^2 for various values of R for each individual data set. The final row lists the total $\chi^2/N_{\rm dof}$ for each R.

R [GeV^{-1}]		0.4	0.6	0.8	1.0	2.0	3.0
BCDMS	NC μ F_2^p 100 GeV	1.257	1.273	1.236	1.229	1.202	1.197
	NC μ F_2^p 120 GeV	0.770	0.766	0.753	0.754	0.748	0.746
	NC μ F_2^p 200 GeV	1.118	1.106	1.091	1.093	1.087	1.086
	NC μ F_2^p 280 GeV	0.949	0.947	0.946	0.951	0.948	0.948
HERA	NC e^+ 920 GeV	1.324	1.248	1.218	1.210	1.200	1.200
	NC e^+ 820 GeV	1.088	1.040	1.029	1.008	0.991	0.988
	NC e^+ 575 GeV	0.924	0.895	0.883	0.879	0.875	0.874
	NC e^+ 460 GeV	1.137	1.109	1.094	1.095	1.095	1.094
	NC e^- 920 GeV	1.495	1.469	1.457	1.458	1.454	1.455
	CC e^+ 920 GeV	1.330	1.421	1.349	1.412	1.373	1.365
	CC e^- 920 GeV	1.505	1.434	1.501	1.505	1.494	1.491
NMC	NC μ F_2^p	1.179	1.038	0.975	0.972	0.972	0.972
$\chi^2/N_{\rm dof}$		1.244	1.205	1.194	1.188	1.182	1.182

TABLE III. Same as Table II, but with parametrization 2.

R [GeV^{-1}]		0.4	0.6	0.8	1.0	2.0	3.0
BCDMS	NC μ F_2^p 100 GeV	1.045	1.006	1.012	1.013	1.016	1.017
	NC μ F_2^p 120 GeV	0.755	0.749	0.749	0.750	0.751	0.752
	NC μ F_2^p 200 GeV	1.090	1.083	1.082	1.082	1.082	1.082
	NC μ F_2^p 280 GeV	0.934	0.934	0.935	0.935	0.937	0.937
HERA	NC e^+ 920 GeV	1.211	1.188	1.188	1.191	1.197	1.198
	NC e^+ 820 GeV	1.066	1.041	1.029	1.024	1.021	1.021
	NC e^+575 GeV	0.883	0.878	0.877	0.877	0.878	0.878
	NC e^+ 460 GeV	1.074	1.066	1.068	1.070	1.072	1.073
	NC e^- 920 GeV	1.455	1.455	1.456	1.456	1.457	1.457
	CC e^+ 920 GeV	1.006	0.952	0.946	0.944	0.942	0.941
	CC e^- 920 GeV	1.218	1.263	1.262	1.261	1.262	1.262
NMC	NC μ F_2^p	1.092	1.082	1.081	1.083	1.086	1.087
$\chi^2/N_{\rm dof}$		1.206	1.187	1.187	1.188	1.191	1.192

TABLE IV. Same as Table II, but with the lower cut $Q^2_{\rm cut} = 1.0$ GeV^2.

R [GeV^{-1}]		0.4	0.6	0.8	1.0	2.0	3.0
BCDMS	NC μF_2^p 100 GeV	1.272	1.262	1.227	1.206	1.175	1.167
	NC μ F_2^p 120 GeV	0.797	0.762	0.755	0.751	0.747	0.746
	NC μ F_2^p 200 GeV	1.133	1.099	1.092	1.089	1.087	1.087
	NC μ F_2^p 280 GeV	0.955	0.949	0.952	0.952	0.952	0.952
HERA	NC e^+ 920 GeV	1.778	1.331	1.256	1.231	1.213	1.212
	NC e^+ 820 GeV	1.816	1.218	1.213	1.196	1.179	1.177
	NC e^+ 575 GeV	1.017	0.919	0.907	0.901	0.898	0.898
	NC e^+ 460 GeV	1.229	1.111	1.078	1.081	1.076	1.075
	NC e^- 920 GeV	1.560	1.495	1.472	1.466	1.463	1.463
	CC e^+ 920 GeV	1.420	1.332	1.362	1.361	1.363	1.361
	CC e^- 920 GeV	2.039	1.598	1.504	1.494	1.484	1.480
NMC	NC μ F_2^p	1.648	1.068	1.008	1.008	1.023	1.028
	$\chi^2/N_{\rm dof}$	1.507	1.259	1.226	1.218	1.213	1.212

APPENDIX B: OBTAINED PDF PARAMETERS

Tables V–VII show the parameters obtained in the different sets of PDF fits for a selection of R values $R = \{0.4, 0.6, 0.8, 1.0, 2.0, 3.0\}$ GeV^{-1}. Table V shows the parameters obtained in the fits with parametrization 1 and $Q^2_{\rm cut} = 3.5$ GeV^2. Table VI shows the same for $Q^2_{\rm cut} = 1.0$ GeV^2 and Table VII for parametrization 2 with $Q^2_{\rm cut} = 3.5$ GeV^2.

TABLE V. Parameters obtained in the fits with parametrization 1 and $Q^2_{\rm cut} = 3.5$ GeV^2 at various values of R. Parameters in parentheses are fixed at the given value, while parameters not shown are set to zero.

R [GeV^{-1}]	0.4	0.6	0.8	1.0	2.0	3.0
$A_{\bar{d}}$	0.404469	0.161368	0.104539	0.061467	0.056025	0.054686
A'_g	2.218085	0.435239	0.410244	0.201921	0.494136	0.428424
$B_{\bar{d}}$	0.010765	−0.085343	−0.108657	−0.128949	−0.144103	−0.145589
B_{d_v}	1.422865	1.454721	1.363376	1.323526	1.314441	1.313151
B_g	−0.211063	−0.342516	−0.174697	−0.042353	−0.123857	−0.069253
B'_g	0.256574	−0.449499	−0.309159	−0.302806	−0.260254	−0.235
B_{u_v}	0.914925	0.921869	0.934284	0.921569	0.923402	0.923452
$C_{\bar{d}}$	18.358827	13.091428	6.798707	2.759383	2.15534	2.020401
C_{d_v}	5.876824	5.987926	5.560861	5.694443	5.733379	5.735318
C_g	2.208999	2.120541	3.683178	4.354789	4.851641	5.230455
C'_g	(25.0)	(25.0)	(25.0)	(25.0)	(25.0)	(25.0)
$C_{\bar{u}}$	15.873966	17.198562	16.234897	16.742077	15.787451	15.521866
C_{u_v}	2.48766	2.496799	2.529641	2.489838	2.489735	2.487229
$D_{\bar{u}}$	16.866158	28.489481	22.125291	23.503082	22.000836	21.205419
E_{u_v}	−0.997742	−0.988268	−0.971517	−1.003805	−1.020271	−1.027135

TABLE VI. Same as Table V, but for PDFs fitted with the lower cut $Q^2_{\rm cut} = 1.0$ GeV^2.

R [GeV^{-1}]	0.4	0.6	0.8	1.0	2.0	3.0
$A_{\bar{d}}$	0.150472	0.105445	0.094215	0.082058	0.06272	0.061345
A'_g	−0.481786	0.104096	0.142766	0.234982	0.24655	0.268521
$B_{\bar{d}}$	−0.089444	−0.091097	−0.100384	−0.104039	−0.109834	−0.110594
B_{d_v}	1.026321	1.222286	1.314041	1.297563	1.266633	1.265011
B_g	0.112897	−0.090972	0.012384	0.019837	0.103054	0.1072
B'_g	−0.148734	−0.378519	−0.278353	−0.215802	−0.153492	−0.140128
B_{u_v}	0.872954	0.913072	0.927043	0.924927	0.919602	0.91878
$C_{\bar{d}}$	4.207006	5.524392	6.20659	5.084081	2.925174	2.766838
C_{d_v}	4.871045	5.154293	5.408388	5.359225	5.395968	5.409781
C_g	2.001991	2.939454	4.006964	4.490556	5.487986	5.642055
C'_g	(25.0)	(25.0)	(25.0)	(25.0)	(25.0)	(25.0)
$C_{\bar{u}}$	22.296374	16.795233	15.487592	14.595522	13.624455	13.376877
C_{u_v}	2.365049	2.476635	2.499119	2.481032	2.461876	2.458407
$D_{\bar{u}}$	30.234495	17.061131	15.205939	12.889824	10.462935	9.919356
E_{u_v}	−1.056546	−1.003855	−0.993606	−1.007731	−1.02491	−1.028614

TABLE VII. Same as Table V, but for PDFs fitted with parametrization 2.

R [GeV^{-1}]	0.4	0.6	0.8	1.0	2.0	3.0
$A_{\bar{d}}$	0.248371	0.238772	0.229188	0.226517	0.224377	0.224177
$B_{\bar{d}}$	−0.095279	−0.108632	−0.115726	−0.118081	−0.120341	−0.120637
B_{d_v}	0.989654	1.05047	1.050583	1.053705	1.057556	1.058181
B_g	−0.063721	0.176504	0.300771	0.368229	0.46931	0.489468
B_{u_v}	0.330037	0.10342	0.10933	0.108711	0.108755	0.108843
$C_{\bar{d}}$	9.978846	10.592777	10.523469	10.595957	10.745633	10.781037
C_{d_v}	4.363475	4.575691	4.5757	4.591326	4.61234	4.61603
C_g	19.131659	21.235982	21.791519	21.993852	22.228964	22.267116
$C_{\bar{u}}$	8.828193	9.34704	9.525786	9.554065	9.563123	9.560328
C_{u_v}	3.24987	2.501855	3.573471	3.572397	3.568641	3.567745
D_g	−9.671812	−8.386368	−7.830814	−7.681217	−7.689173	−7.720356
$D_{\bar{u}}$	9.759016	10.223122	11.255959	11.466205	11.578213	11.574223
D_{u_v}	31.344142	232.987381	215.930099	217.666728	217.330601	216.978654
E_g	465.912581	399.008994	335.978167	302.820564	257.896509	249.69638
E_{u_v}	−20.809607	−238.165353	11.060703	11.049585	9.585398	9.200095

[1] J. C. Collins, D. E. Soper, and G. F. Sterman, Factorization of hard processes in QCD, Adv. Ser. Dir. High Energy Phys. **5**, 1 (1989).

[2] Y. L. Dokshitzer, Calculation of the structure functions for deep inelastic scattering and e^+e^- annihilation by perturbation theory in quantum chromodynamics, Sov. Phys. JETP **46**, 641 (1977).

[3] V. N. Gribov and L. N. Lipatov, Deep inelastic ep scattering in perturbation theory, Sov. J. Nucl. Phys. **15**, 438 (1972).

[4] V. N. Gribov and L. N. Lipatov, e^+e^- pair annihilation and deep inelastic ep scattering in perturbation theory, Sov. J. Nucl. Phys. **15**, 675 (1972).

[5] G. Altarelli and G. Parisi, Asymptotic freedom in parton language, Nucl. Phys. **B126**, 298 (1977).

[6] J.-w. Qiu, Nuclear shadowing at small values of x, Nucl. Phys. **B291**, 746 (1987).

[7] Y. V. Kovchegov and E. Levin, *Quantum Chromodynamics at High Energy* (Oxford University Press, New York, 2013), Vol. 33.

[8] A. H. Mueller and J.-w. Qiu, Gluon recombination and shadowing at small values of x, Nucl. Phys. **B268**, 427 (1986).

[9] L. V. Gribov, E. M. Levin, and M. G. Ryskin, Semihard processes in QCD, Phys. Rep. **100**, 1 (1983).

[10] J. Bartels, G. A. Schuler, and J. Blumlein, A numerical study of the small x behavior of deep inelastic structure functions in QCD, Z. Phys. C **50**, 91 (1991).

[11] K. Prytz, Signals of gluon recombination in deep inelastic scattering, Eur. Phys. J. C **22**, 317 (2001).

[12] K. J. Eskola, H. Honkanen, V. J. Kolhinen, J.-w. Qiu, and C. A. Salgado, Nonlinear corrections to the DGLAP equations in view of the HERA data, Nucl. Phys. **B660**, 211 (2003).

[13] G. R. Boroun and S. Zarrin, An approximate approach to the nonlinear DGLAP evaluation equation, Eur. Phys. J. Plus **128**, 119 (2013).

[14] G. R. Boroun, Shadowing correction to the gluon distribution behavior at small x, Eur. Phys. J. A **42**, 251 (2009).

[15] J. Rausch, V. Guzey, and M. Klasen, Numerical evaluation of the nonlinear Gribov-Levin-Ryskin-Mueller-Qiu evolution equations for nuclear parton distribution functions, Phys. Rev. D **107**, 054003 (2023).

[16] W. Zhu, A new approach to parton recombination in a QCD evolution equation, Nucl. Phys. **B551**, 245 (1999).

[17] W. Zhu and J.-h. Ruan, A new modified Altarelli-Parisi evolution equation with parton recombination in proton, Nucl. Phys. **B559**, 378 (1999).

[18] J. Blumlein, V. Ravindran, J.-h. Ruan, and W. Zhu, Twist four gluon recombination corrections for deep inelastic structure functions, Phys. Lett. B **504**, 235 (2001).

[19] M. Klasen and H. Paukkunen, Nuclear PDFs after the first decade of LHC data, arXiv:2311.00450.

[20] X. Chen, J. Ruan, R. Wang, P. Zhang, and W. Zhu, Applications of a nonlinear evolution equation I: The parton distributions in the proton, Int. J. Mod. Phys. E **23**, 1450057 (2014).

[21] R. Wang and X. Chen, Dynamical parton distributions from DGLAP equations with nonlinear corrections, Chin. Phys. C **41**, 053103 (2017).

[22] F. Gelis, E. Iancu, J. Jalilian-Marian, and R. Venugopalan, The color glass condensate, Annu. Rev. Nucl. Part. Sci. **60**, 463 (2010).

[23] A. Morreale and F. Salazar, Mining for gluon saturation at colliders, Universe **7**, 312 (2021).

[24] G. P. Salam and J. Rojo, A higher order perturbative parton evolution toolkit (HOPPET), Comput. Phys. Commun. **180**, 120 (2009).

[25] H. Abdolmaleki *et al.* (xFitter Collaboration), xFitter: An open source QCD analysis framework. A resource and reference document for the Snowmass study, arXiv:2206.12465.
[26] S. Alekhin *et al.*, HERAFitter, Eur. Phys. J. C **75**, 304 (2015).
[27] A. C. Benvenuti *et al.* (BCDMS Collaboration), A high statistics measurement of the proton structure functions $F_2(x, Q^2)$ and R from deep inelastic muon scattering at high Q^2, Phys. Lett. B **223**, 485 (1989).
[28] H. Abramowicz *et al.* (H1 and ZEUS Collaborations), Combination of measurements of inclusive deep inelastic $e^{\pm}p$ scattering cross sections and QCD analysis of HERA data, Eur. Phys. J. C **75**, 580 (2015).
[29] M. Arneodo *et al.* (New Muon Collaboration), Measurement of the proton and deuteron structure functions, F_2^p and F_2^d, and of the ratio σ_L/σ_T, Nucl. Phys. **B483**, 3 (1997).
[30] V. S. Fadin, E. A. Kuraev, and L. N. Lipatov, On the Pomeranchuk singularity in asymptotically free theories, Phys. Lett. **60B**, 50 (1975).
[31] E. A. Kuraev, L. N. Lipatov, and V. S. Fadin, Multi-Reggeon processes in the Yang-Mills theory, Sov. Phys. JETP **44**, 443 (1976).
[32] E. A. Kuraev, L. N. Lipatov, and V. S. Fadin, The Pomeranchuk singularity in nonabelian gauge theories, Sov. Phys. JETP **45**, 199 (1977).
[33] I. I. Balitsky and L. N. Lipatov, The Pomeranchuk singularity in quantum chromodynamics, Sov. J. Nucl. Phys. **28**, 822 (1978).
[34] M. Bonvini, S. Marzani, and C. Muselli, Towards parton distribution functions with small-x resummation: HELL 2.0, J. High Energy Phys. 12 (2017) 117.
[35] R. D. Ball, V. Bertone, M. Bonvini, S. Marzani, J. Rojo, and L. Rottoli, Parton distributions with small-x resummation: Evidence for BFKL dynamics in HERA data, Eur. Phys. J. C **78**, 321 (2018).
[36] H. Abdolmaleki *et al.* (xFitter Developers' Team), Impact of low-x resummation on QCD analysis of HERA data, Eur. Phys. J. C **78**, 621 (2018).
[37] K. J. Eskola, J.-w. Qiu, and X.-N. Wang, Perturbative gluon shadowing in heavy nuclei, Phys. Rev. Lett. **72**, 36 (1994).
[38] K. J. Eskola, H. Honkanen, V. J. Kolhinen, J.-w. Qiu, and C. A. Salgado, Nonlinear corrections to the DGLAP equations: Looking for the saturation limits, arXiv:hep-ph/0302185.
[39] A. Accardi *et al.*, Electron Ion Collider: The next QCD frontier: Understanding the glue that binds us all, Eur. Phys. J. A **52**, 268 (2016).
[40] R. Abdul Khalek *et al.*, Science requirements and detector concepts for the Electron-Ion Collider: EIC Yellow Report, Nucl. Phys. **A1026**, 122447 (2022).
[41] P. Agostini *et al.* (LHeC and FCC-he Study Group Collaborations), The Large Hadron–Electron Collider at the HL-LHC, J. Phys. G **48**, 110501 (2021).
[42] J. L. Abelleira Fernandez *et al.* (LHeC Study Group), A Large Hadron Electron Collider at CERN: Report on the physics and design concepts for machine and detector, J. Phys. G **39**, 075001 (2012).
[43] A. Abada *et al.* (FCC Collaboration), FCC physics opportunities: Future Circular Collider conceptual design report volume 1, Eur. Phys. J. C **79**, 474 (2019).
[44] A. Accardi, L. T. Brady, W. Melnitchouk, J. F. Owens, and N. Sato, Constraints on large-x parton distributions from new weak boson production and deep-inelastic scattering data, Phys. Rev. D **93**, 114017 (2016).
[45] M. Botje, QCDNUM: Fast QCD evolution and convolution, Comput. Phys. Commun. **182**, 490 (2011).
[46] R. S. Thorne, A variable-flavor number scheme for NNLO, Phys. Rev. D **73**, 054019 (2006).
[47] J. Pumplin, D. Stump, R. Brock, D. Casey, J. Huston, J. Kalk, H. L. Lai, and W. K. Tung, Uncertainties of predictions from parton distribution functions. 2. The Hessian method, Phys. Rev. D **65**, 014013 (2001).
[48] K. J. Eskola, H. Paukkunen, and C. A. Salgado, EPS09: A new generation of NLO and LO nuclear parton distribution functions, J. High Energy Phys. 04 (2009) 065.
[49] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in FORTRAN 77 The Art of Scientific Computing* (Cambridge University Press, 1986).
[50] M. Walt, I. Helenius, and W. Vogelsang, Open-source QCD analysis of nuclear parton distribution functions at NLO and NNLO, Phys. Rev. D **100**, 096015 (2019).
[51] I. Helenius, M. Walt, and W. Vogelsang, NNLO nuclear parton distribution functions with electroweak-boson production data from the LHC, Phys. Rev. D **105**, 094031 (2022).
[52] V. Andreev *et al.* (H1 Collaboration), Measurement of inclusive ep cross sections at high Q^2 at $\sqrt{s} = 225$ and 252 GeV and of the longitudinal proton structure function F_L at HERA, Eur. Phys. J. C **74**, 2814 (2014).
[53] F. D. Aaron *et al.* (H1 Collaboration), Measurement of the charm and beauty structure functions using the H1 vertex detector at HERA, Eur. Phys. J. C **65**, 89 (2010).
[54] H. Abramowicz *et al.* (H1 and ZEUS Collaborations), Combination and QCD analysis of charm production cross section measurements in deep-inelastic ep scattering at HERA, Eur. Phys. J. C **73**, 2311 (2013).
[55] H. Abramowicz *et al.* (ZEUS Collaboration), Measurement of beauty and charm production in deep inelastic scattering at HERA and measurement of the beauty-quark mass, J. High Energy Phys. 09 (**2014**) 127.
[56] T. Lappi, H. Mäntysaari, H. Paukkunen, and M. Tevio, Evolution of structure functions in momentum space, Eur. Phys. J. C **84**, 84 (2024).
[57] H. Abramowicz *et al.* (ZEUS Collaboration), Deep inelastic cross-section measurements at large y with the ZEUS detector at HERA, Phys. Rev. D **90**, 072002 (2014).
[58] E. Aschenauer *et al.*, The longitudinal structure function f_l at an EIC, https://www.phenix.bnl.gov/WWW/publish/elke/EIC/EIC-R&D-Tracking/Meetings/fl.pdf; accessed 03-May-2023.
[59] H. Mäntysaari and P. Zurita, In depth analysis of the combined HERA data in the dipole models with and without saturation, Phys. Rev. D **98**, 036002 (2018).
[60] A. Buckley, J. Ferrando, S. Lloyd, K. Nordström, B. Page, M. Rüfenacht, M. Schönherr, and G. Watt, LHAPDF6: Parton density access in the LHC precision era, Eur. Phys. J. C **75**, 132 (2015).
[61] https://research.hip.fi/qcdtheory/nuclear-pdfs/.