Signatures of gluon saturation from structure-function measurements

Nestor Armesto[,1,*] Tuomas Lappi[,2,3,†] Heikki Mäntysaari[,2,3,‡] Hannu Paukkunen,[2,3,§] and Mirja Tevio[,2,3,‖]

[1]*Instituto Galego de Física de Altas Enerxías IGFAE, Universidade de Santiago de Compostela,*
15782 Santiago de Compostela, Galicia, Spain
[2]*University of Jyvaskyla, Department of Physics, P.O. Box 35, FI-40014 University of Jyvaskyla, Finland*
[3]*Helsinki Institute of Physics, P.O. Box 64, FI-00014 University of Helsinki, Finland*

We study experimentally observable signals for nonlinear QCD dynamics in deep inelastic scattering (DIS) at small Bjorken variable x and moderate virtuality Q^2, by quantifying differences between the linear Dokshitzer-Gribov-Lipatov-Altarelli-Parisi evolution and nonlinear evolution with the Balitsky-Kovchegov equation. To remove the effect of the parametrization freedom in the initial conditions of both equations, we first match the predictions for the DIS structure functions F_2 and F_L from both frameworks in a region in x, Q^2 where both frameworks should provide an accurate description of the relevant physics. The differences in the dynamics are then quantified by the deviations when one moves away from this matching region. For free protons we find that the differences in F_2 remain at a few-percent level, while in F_L the deviations are larger, up to 10% at the Electron Ion Collider (EIC) and 40% at the Large Hadron-electron Collider (LHeC) kinematics. With a heavy nucleus the differences are up to 10% in F_2, and can reach 20% and 60% in F_L for the EIC and the LHeC, respectively.

I. INTRODUCTION

The high-energy regime of quantum chromodynamics (QCD) is characterized by high partonic densities where nonlinear phenomena play a dominant role [1]. In this regime, the predictions of usual linear evolution equations for parton densities like the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equation [2–5] which provides an accurate description of the QCD dynamics at moderate to large values of the momentum fractions x of the probed parton and virtualities $Q^2 \gg \Lambda_{QCD}^2$, or the Balitsky-Fadin-Kuraev-Lipatov (BFKL) equation [6,7] which accurately describes the partonic dynamics at small x, are expected to become less reliable. High-density effects should become relevant and unitarity corrections or, alternatively, gluon recombination processes, tame the growth of parton densities towards small x and make them saturate—thus nonlinear dynamics is also referred to as saturation. Nonlinear modifications to DGLAP evolution equations were first proposed in Refs. [8,9]. Later on, a formulation suitable for QCD at such high parton densities and energies (or small x) was developed as a weak coupling but nonperturbative effective field theory, the color glass condensate (CGC), see Refs. [1,10]. Such an approach resums all higher-twist corrections that are neglected in linear DGLAP evolution. In this framework, the BFKL equation for evolution in x is generalized to the Balitsky-Kovchegov (BK) one [11–13]. The BK equation includes a nonlinear term needed to preserve the unitarity of the (dipole) scattering amplitude.

Finding such a novel regime of QCD would be of uttermost importance for understanding QCD, for developing more realistic initial conditions for simulations of heavy ion collisions at high energy and for computing the main backgrounds for signals of physics beyond the Standard Model in hadronic collisions, especially at the high energies of the Future Circular Collider [14,15]. Linear, fixed-order DGLAP evolution equations are known to provide a very good description of the structure of the proton at the available values of x and Q^2, see e.g., Ref. [16]. However, there have also been claims that the hadron-electron ring accelerator (HERA) data [17,18] at small x and moderate Q^2 necessitates the inclusion of small-x resummation [19,20] or nonlinear evolution [21].[1] Note also that while DGLAP-based fits provide a good

*nestor.armesto@usc.es
†tuomas.v.v.lappi@jyu.fi
‡heikki.mantysaari@jyu.fi
§hannu.paukkunen@jyu.fi
‖mirja.h.tevio@jyu.fi

---

[1]The impact of the former on Higgs production at hadronic colliders on Higgs production has been examined in [22].

description of available data constraining parton densities inside nuclei (nPDFs), see, e.g., [23–28], nonlinear effects are density effects that should be enhanced by the nuclear size. Therefore, searches for saturation with heavier nuclei are crucial not only because the expected effects are stronger, and thereby more easily detectable, but also to eventually check the theoretical explanation of such nonlinear dynamics, if found. For the moment, however, even the D-meson production at the LHC that can probe the nuclei down to $x \sim 10^{-6}$ at low Q^2 shows no visible deviation from the DGLAP approach [28,29].

Searches of saturation are ongoing in hadronic colliders [30]—the Relativistic Heavy Ion Collider (RHIC) at BNL and the Large Hadron Collider (LHC) at CERN, see e.g., [31] for future prospects at the LHC. However, with a cleaner environment and direct experimental access to the kinematical variables x, Q^2, deep inelastic scattering (DIS) in lepton-proton/nucleus collisions offers the ideal environment to study the partonic structure of protons and heavier nuclei. The currently projected and proposed accelerators, the Electron Ion Collider (EIC) [32,33] at BNL, and the Large Hadron-electron Collider (LHeC) [34,35] and the Future Circular Collider in electron-hadron mode (FCC-he) [14,15] at CERN, will provide lepton-hadron collisions in the center-of-mass range from a few tens of GeV to a few TeV per nucleon. Together with the use of new detector techniques, they promise to revolutionize our understanding of hadron structure and partonic dynamics.

In DIS, several observables have been proposed as being sensitive to nonlinear dynamics: particle correlations [34,36], diffraction [32,37], and deviations from linear evolution in inclusive observables, in particular the total cross section. Concerning the latter, it is expected that the EIC and the LHeC will be sensitive to saturation, i.e., to deviations from DGLAP due to higher-twist effects. However, conclusively distinguishing these two pictures using experimental data can, in practice, be difficult. This is due to the fact that both nonlinear and DGLAP-based calculations require nonperturbative inputs that are obtained by fitting experimental data. When the initial conditions of the evolution are fit to the same data, also the predictions will not deviate dramatically. Thus different theoretical frameworks might be equally capable of describing experimental data *a posteriori*. Additionally, the DGLAP and BK equations predict the evolution of cross sections in different directions: increasing Q^2 and decreasing x, respectively. Thus one cannot solve both starting from the same initial condition: the perturbative QCD prediction of one equation is the nonperturbative initial condition of the other. Therefore, it is important to be careful to distinguish between genuine effects of the different evolution dynamics on one hand, and the ability of the initial condition parametrizations to adjust to data on the other hand.

Several works have investigated the possibility to disentangle DGLAP linear dynamics from saturation in ep collisions at the LHeC [34,35,38] and eA collisions at the EIC [39]. The basic technique employed in these studies is the inclusion of pseudodata from the projected experiments in a DGLAP-based global fit. Then, some measure of the fit quality is computed. Generically, a worsening is observed when the introduced pseudodata are generated using a model that contains nonlinear dynamics in comparison to the case when the corresponding pseudodata are produced using linear DGLAP dynamics. Focusing on the difference between the generated pseudodata and existing DGLAP-evolved parton distributions, as in Ref. [39], addresses potential tensions between predictions from nonlinear dynamics and *existing* PDF parametrizations. The latter are determined not only by the QCD dynamics, but by existing data and to some extent by parametrization choices. Thus the question addressed by such a comparison has many more facets than the simple one we pose here, namely: how different is the actual evolution dynamics once one allows both fits to adjust themselves to a common set of data?

The purpose of this paper is to evaluate the difference between DGLAP and nonlinear evolution in the kinematic regions corresponding to the EIC and the LHeC, both for ep and eA collisions, with a slightly different emphasis than previous studies. Our procedure is the following. We first produce results for the F_2 and F_L structure functions using nonlinear evolution equations. Then, employing reweighting techniques, a DGLAP evolved set of parton densities is produced that results in structure functions coinciding with the nonlinear evolution ones in a line of the x, Q^2 plane, *as precisely as possible*. Technically this is achieved by a Bayesian reweighting process. Here one takes an existing collection of "replica" PDF sets that all satisfy the DGLAP equation, and calculates for each of them a weight defined by how close it is to the given F_2 and F_L structure functions. The reweighted structure functions and PDFs are then defined as a weighted sum over all replicas. We will explain this procedure in more detail below in Sec. II C. Now the difference between the nonlinear results and the DGLAP ones for F_2 and F_L when moving away from this line provides a quantification of the signatures of saturation—of the difference between linear and nonlinear dynamics. Although we are interested in the experimentally accessible kinematical regimes, we do not use estimated experimental errors in this procedure.

Our approach can be compared to the one used in the previous studies [34,35,38,39], where a DGLAP fit to pseudodata from BK evolution in the whole kinematical region was performed. There one fits to the pseudodata in a wide region in the kinematical plane, and the weight given to specific kinematical regions is determined by the projected errors. Our goal, in contrast, is to force the two approaches to agree quite precisely on just a line in the kinematical plane, preferably in a region where both theoretical frameworks should be valid. The deviations as

one moves away from this line provide a clearer picture of the genuine differences in the evolution dynamics, independently of experimental errors, kinematical ranges or PDF parametrizations.

The structure of the manuscript is as follows. In Sec. II A we review the method to compute the structure functions in collinear factorization that will be employed to produce the DGLAP results. Section II B addresses the same point in the framework of nonlinear, saturation dynamics. The PDF reweighting technique to produce a DGLAP fit coincident with the saturation results in a line of the x, Q^2 plane is presented in Sec. II C. Results for proton and nucleus are shown and discussed in Sec. III. Finally, we present our conclusions in Sec. IV.

II. THEORETICAL FRAMEWORK

A. Structure functions in collinear factorization

To compute the structure functions in collinear factorization we use APFEL [40] which accesses the PDFs through the LHAPDF library [41]. For the proton we use the next-to-leading order (NLO) NNPDF3.1 set NNPDF31_nlo_as_0118_1000 [42] which has $N_{\rm rep} = 1000$ replicas describing the PDF uncertainty, and for gold (our representative of a heavier nucleus) we use the NLO nNNPDF2.0 nNNPDF20_nlo_as_0118_Au197 [27] set which provides 1000 replicas as well.[2] The central value for any PDF-dependent quantity $\mathcal{O}$ is obtained as an average over the individual predictions $\mathcal{O}[f_k]$ for each individual replica,

$$\mathcal{O} = \frac{1}{N_{\rm rep}} \sum_{k=1}^{N_{\rm rep}} \mathcal{O}[f_k], \qquad (1)$$

and the uncertainty $\delta\mathcal{O}$ is computed as the variance,

$$\delta\mathcal{O} = \sqrt{\frac{1}{N_{\rm rep}} \sum_{k=1}^{N_{\rm rep}} (\mathcal{O}[f_k] - \mathcal{O})^2}. \qquad (2)$$

We calculate the neutral-current DIS structure functions at NLO accuracy in the FONLL-B general-mass scheme [44]. This scheme matches with the one used in the actual global analyses. With the NNPDF3.1 proton PDF set we ensure the consistent treatment of fitted charm quarks by setting the appropriate options in APFEL [45], in particular enabling intrinsic charm by APFEL:: EnableIntrinsicCharm(true). All the parameters (masses of heavy quarks, values of strong coupling α_s, etc.) are set to be the same as in the corresponding PDF fits.

---

[2]We note that a preprint of the updated nNNPDF3.0 set [43] became available when this work was being finalized.

We note that the NNPDF collaboration has also prepared PDF sets including small-x resummation effects at leading and next-to-leading logarithmic accuracy [19]. In particular the (linear) small-x resummation was found to be necessary to describe the HERA structure function data in the small-x, moderate Q^2 region at next-to-next-to-leading order fits. At NLO accuracy, on the other hand, this resummation does not have a numerically significant effect on structure functions at HERA kinematics and is not considered in this work.

B. Structure functions with gluon saturation

In order to calculate proton and nuclear structure functions taking into account nonlinear saturation effect we use the color glass condensate (CGC) framework [1,10,46]. Here it is convenient to describe DIS processes in the dipole picture in a frame where the target is at rest and the photon plus light-cone momentum is large, and write the leading order total photon-nucleus cross section as

$$\sigma_{T,L}^{\gamma^*A} = 2 \sum_f \int d^2\mathbf{b}\, d^2\mathbf{r}\, dz |\Psi^{\gamma^* \to q\bar{q}}(\mathbf{r}, z, Q^2)|^2 N(\mathbf{b}, \mathbf{r}, x). \quad (3)$$

Here the photon light front wave function $\Psi^{\gamma^* \to q\bar{q}}(\mathbf{r}, z, Q^2)$ describes the photon splitting to a quark (with flavor f) dipole with transverse size $\mathbf{r}$ and the quark carrying a fraction z of the photon plus momentum [1]. All information about the target is encoded in the dipole-target scattering amplitude $N(\mathbf{r}, \mathbf{b}, x)$, which describes the eikonal propagation of the quark-antiquark pair through the target color field. The subscripts T and L refer to the photon polarization states. The structure functions can then be written as

$$F_2(x, Q^2) = \frac{Q^2}{4\pi\alpha_{\rm em}} (\sigma_T^{\gamma^*A} + \sigma_L^{\gamma^*A}), \qquad (4)$$

$$F_L(x, Q^2) = \frac{Q^2}{4\pi\alpha_{\rm em}} \sigma_L^{\gamma^*A}. \qquad (5)$$

The energy (or Bjorken-x) evolution of the dipole-target scattering amplitude N (and thus the x dependence of the structure functions) can be obtained by solving the perturbative Balitsky-Kovchegov (BK) evolution equation [11,12] which resums contributions $\sim \alpha_s \ln 1/x$ to all orders. In contrast to the collinear factorization based approach described above, the Q^2 dependence is not calculated perturbatively—at least at leading order. Instead, it is effectively determined, along with other parameters, by the nonperturbative initial condition for the BK evolution. These initial conditions are obtained by fitting the HERA structure function data [17,18]. In this work we use the "MV^e" leading order fit with running coupling corrections [47] from Ref. [48] (see also a similar fit in Ref. [49]).

One now needs to generalize the dipole-proton scattering amplitude to the dipole-nucleus case in order to calculate nuclear structure functions. We again do this using the procedure of Ref. [48], where the squared nuclear saturation scale is taken to be proportional to the transverse density T_A as $Q_{s,A}^2 \sim T_A(\mathbf{b})$. This scaling is used to generate the initial condition for the BK evolution separately at every impact parameter $\mathbf{b}$, and different impact parameters are then evolved independently.

In principle it would be possible to include impact parameter dependence also in the BK evolution, but that would require us to include an additional phenomenological description of confinement effects, see e.g., Refs. [50–53]. As the structure functions are only sensitive to the integral over the impact parameter, we expect additional effects from the geometry evolution to only have a small effect on our results. In the region where the saturation scale of the nucleus would fall below that of the proton we use, again following Ref. [48], a dipole-proton scattering amplitude scaled such that the total dipole-nucleus cross section would be simply A times the dipole-proton cross section. By construction the nuclear modifications then vanish in the dilute region at the edge of the nucleus. This and similar setups have been successfully used to describe various LHC and RHIC measurements [54–60].

Equation (3) corresponds to the leading order contribution (with $\alpha_s \ln 1/x$ contributions resummed to all orders). Currently there is a rapid progress in the field towards next-to-leading order (NLO) accuracy, and necessary ingredients for NLO phenomenology are becoming available. These include, for example, the virtual photon wave function [61–66] and the BK evolution equation [67–70] at NLO. In the first phenomenological structure function calculations at NLO accuracy [62], very small differences were found between running coupling LO and NLO calculations in the HERA kinematics when the nonperturbative initial condition for the BK evolution is fitted to the data. This means that in leading order setups such as the one used in this work the fit parameters can effectively capture most of the higher order corrections. Consequently we expect our leading order framework to be sufficient for the purposes of this work, where the aim is to study differences between the DGLAP and BK dynamics.

C. Matching

As a technical method to match the structure functions in collinear factorization to those calculated with the BK evolution as discussed in Sec. II B, we use the Bayesian reweighting method [71,72]. The matching is achieved by reweighting the PDFs to reproduce structure functions calculated from the BK equation in a line in the x, Q^2 plane. The preferred region for this matching is the region where both theoretical frameworks should be valid. Thus one wants to be sufficiently above the saturation scale Q_s^2 to

be in a linear regime described by DGLAP, but not so high that the large logarithms of Q^2, not resummed by BK evolution, dominate. An additional technical requirement is that the matching points should be above the minimum scale Q_0^2 for which the PDF sets that we use are available. Based on these considerations we have chosen to match the BK and collinear factorization calculations in the region $10Q_s^2(x) < Q^2 < 11Q_s^2(x)$, where $Q_s^2(x)$ is the saturation scale of the proton as defined in Ref. [48]. This leads to the matching interval $x < 5.6 \times 10^{-3}$ for the proton and $x < 1 \times 10^{-2}$ for the gold nucleus. We perform the matching above a lower limit $x > 1.03 \times 10^{-5}$ that is chosen to be in the LHeC kinematical regime. In our matching region the saturation effects are expected to be weak, but Q^2 is still not too large. Consequently the DGLAP- and BK-evolution based frameworks are expected to deviate from each other less dramatically than if the matching was performed in some more extreme region of phase space. Our results naturally depend on the choice of the matching scale; the results obtained with a matching at a higher value of Q^2 are discussed in the Appendix.

Matching the DGLAP and BK predictions using reweighting amounts to forming an appropriate linear combination of the available PDF replicas such that the predictions given by this linear combination reproduce the structure functions from the BK framework. In practice, for each PDF replica f_k we define a figure of merit χ_k^2 by

$$\chi^2 = \sum_{i,j=1}^{N_{\mathrm{data}}} (y_i - y_i[f_k])\sigma_{ij}^{-1}(y_j - y_j[f_k]), \qquad (6)$$

where y_i denote the N_{data} matching values from BK and $y_i[f_k]$ are the corresponding values calculated by using a given PDF replica. The order of magnitude of χ^2 can be controlled by the entries of the covariance matrix σ_{ij} which is taken to be of the diagonal form,

$$\sigma_{ij} = (\delta_{\mathrm{BK}} y_i)^2 \delta_{ij}, \qquad (7)$$

where δ_{BK} is a constant. In a sense, we are thus assigning a constant relative uncertainty on the predictions ("pseudo-data") generated from the BK setup.

For each replica we assign the so-called Giele-Keller weight [73],

$$\omega_k = \frac{e^{-\frac{1}{2}\chi_k^2}}{\frac{1}{N_{\mathrm{rep}}} \sum_{k=1}^{N_{\mathrm{rep}}} e^{-\frac{1}{2}\chi_k^2}}, \qquad (8)$$

which always sums up to unity,

$$\frac{1}{N_{\mathrm{rep}}} \sum_{k=1}^{N_{\mathrm{rep}}} \omega_k = 1. \qquad (9)$$

The value for any quantity $\mathcal{O}$ using the reweighted PDFs is then obtained as a weighted average,

$$\mathcal{O}^{\text{Rew}} = \frac{1}{N_{\text{rep}}} \sum_{k=1}^{N_{\text{rep}}} \omega_k \mathcal{O}[f_k] \qquad (10)$$

with an uncertainty estimate

$$\delta\mathcal{O}^{\text{Rew}} = \sqrt{\frac{1}{N_{\text{rep}}} \sum_{k=1}^{N_{\text{rep}}} \omega_k (\mathcal{O}[f_k] - \mathcal{O}^{\text{Rew}})^2}. \qquad (11)$$

We note that in the case of reweighting to real experimental data, the appropriate weights to be used in conjunction with the NNPDF replicas include also a term $(\chi_k^2)^{(N_{\text{data}}-1)/2}$ [71,74]. While the Giele-Keller weights favor replicas with $\chi_k^2/N_{\text{data}} \approx 0$, the other weights favor replicas with $\chi_k^2/N_{\text{data}} \approx 1$, as discussed in Ref. [74]. Here we are matching PDFs to smooth values resulting from a theory calculation, without point-by-point statistical fluctuations, not reweighting to experimental data. Thus the preferred values of χ_k^2/N_{data} should be close to 0 and the Giele-Keller weights are better suited to our purpose.

It is useful to define the so-called effective number of replicas N_{eff}, which serves as a proxy to the number of replicas with a significant weight [71,74],

$$N_{\text{eff}} = \exp\left\{ \frac{1}{N_{\text{rep}}} \sum_{k=1}^{N_{\text{rep}}} \omega_k \ln\left(\frac{N_{\text{rep}}}{\omega_k}\right) \right\}. \qquad (12)$$

In the present analysis we use N_{eff} to choose an appropriate value for the error parameter δ_{BK}: If δ_{BK} is too low, $N_{\text{eff}} \approx 1$ (for large N_{rep}) and the procedure picks up only a single

replica irrespectively of how well it fits with the matching values. On the other hand, if $N_{\text{eff}} \sim N_{\text{rep}}$ all replicas have the equal weight and the reweighting does nothing, i.e., δ_{BK} is too high. We have found that iteratively adjusting δ_{BK} such that $N_{\text{eff}} \approx 10$ is a good strategy for finding a set of PDFs that matches the given boundary conditions, i.e., structure functions from the BK framework. In order to obtain $N_{\text{eff}} \approx 10$ we have fixed $\delta_{\text{BK}} = 4.5$ for the proton F_2, $\delta_{\text{BK}} = 11.5$ for the proton F_L, $\delta_{\text{BK}} = 39.5$ for the nuclear F_2, and $\delta_{\text{BK}} = 46$ for the nuclear F_L. For the nuclear reweighting we use $N_{\text{data}} = 138$, and for the proton $N_{\text{data}} = 125$.

III. RESULTS

A. Proton

The structure functions F_2 and F_L for the proton before and after the reweighting on the $Q^2 = 10Q_s^2(x)$ line are shown in Figs. 1(a) and 1(b). The reweighting is done separately for F_2 and F_L, as also in reality these two quantities will be measured in different kinematical domains and with a different experimental precision. The structure functions obtained after the reweighting can be seen to match very well to the BK results. This was to be expected since the proton PDFs and the initial condition for the BK evolution are fitted to the same precise HERA data at $x \gtrsim 10^{-4}$, and the central NNPDF3.1 results are already very close to the BK values to begin with in this domain. However, a nearly perfect agreement with the BK results is obtained also at $x \lesssim 10^{-4}$. All in all, the matching procedure is thus found to work extremely well here.

Next we study how the differences in the BK vs DGLAP dynamics become visible when we move away from the $Q^2 \approx 10Q_s^2(x)$ line. In Figs. 2(a) and 2(b) we show the relative difference
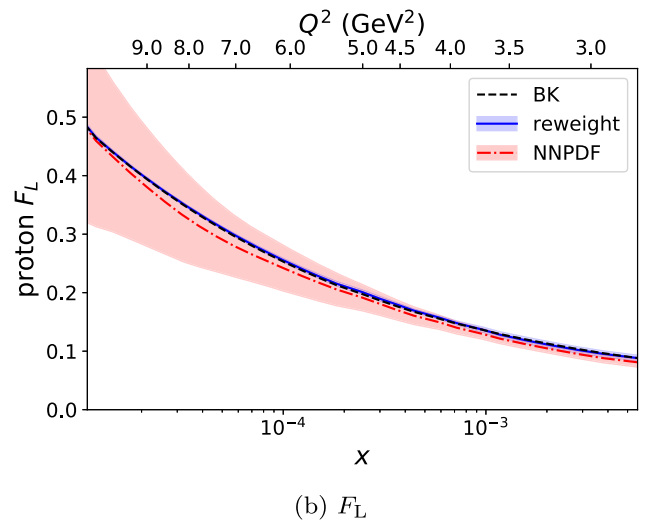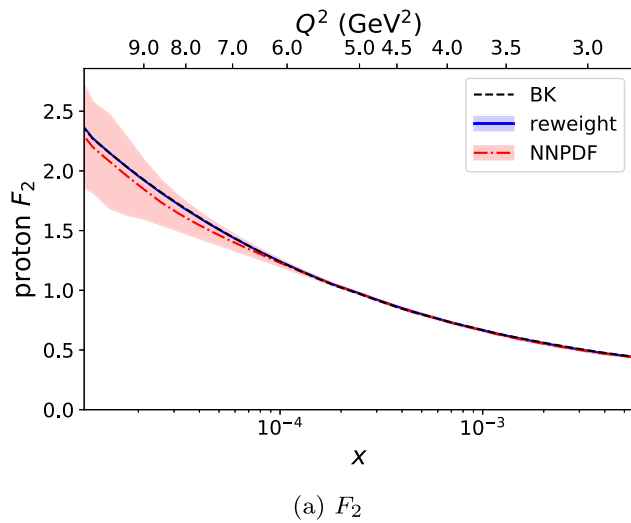

FIG. 1.   The F_2 (a) and F_L (b) structure functions for proton as a function of x at $Q^2 = 10Q_s^2(x)$. The black dashed curve shows the BK predictions, the red dash-dotted curve with the red error band the original NNPDF3.1 PDF predictions, and the blue solid curve with a light-blue error band (too narrow to be visible) the PDF predictions after the matching.
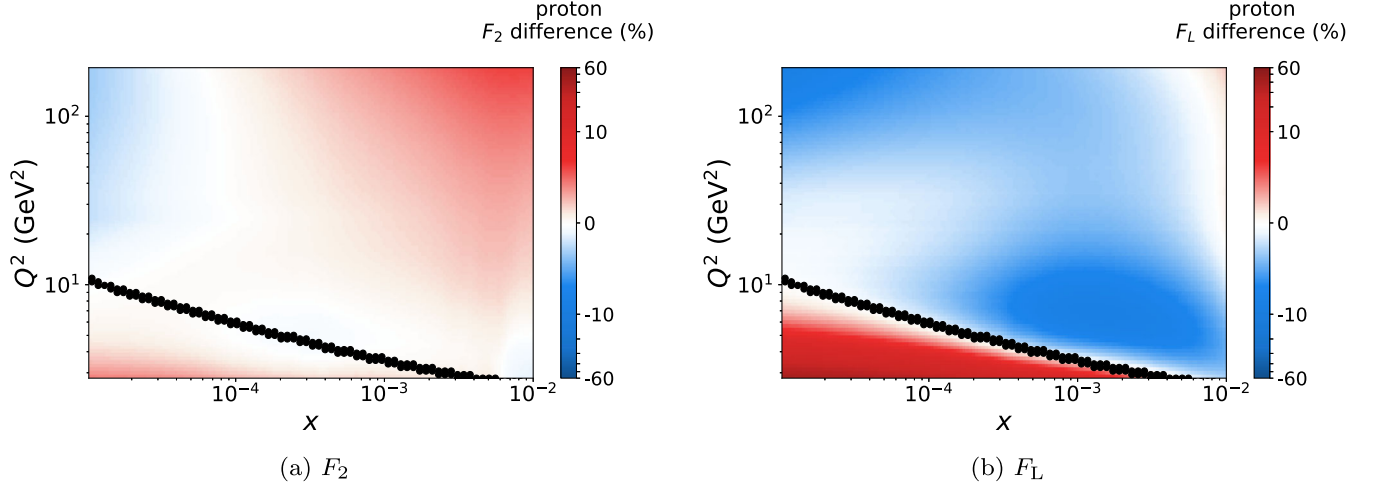
FIG. 2. Relative difference $(F_{2,L}^{BK} - F_{2,L}^{Rew})/F_{2,L}^{BK}$ between the BK structure functions and the matched F_2 (a) and F_L (b) for proton as a function of x and Q^2. The color scale/axis goes in a linear scale from -10% to 10% and in a logarithmic scale outside that range. The black dots indicate the matching points.

$$\frac{F_{2,L}^{BK} - F_{2,L}^{Rew}}{F_{2,L}^{BK}} \qquad (13)$$

as a function of both x and Q^2, where $F_{2,L}^{Rew}$ refers to the corresponding structure function calculated using the reweighted PDFs. The points used in the reweighting are also indicated in these figures. One-dimensional projections of the same quantity are plotted at fixed values of x in Fig. 3.

For the F_2 structure function shown in Fig. 2(a) the differences remain very small, at most at a few-percent level almost everywhere in the studied x, Q^2 range, except in the high-x, high Q^2 and low-x, low Q^2 corners. This is better visible in Fig. 3(a) where we show the relative differences

as a function of virtuality Q^2 at four different x values from $x = 5.6 \times 10^{-3}$ [largest x for which $Q^2 = 10Q_s^2(x) \geq Q_0^2$, where Q_0^2 is the initial scale in the NNPDF3.1 PDF set] to $x = 10^{-5}$. The smallest x values in our plots are beyond reach for the EIC, which will collide electrons with energies 5–18 GeV on protons and nuclei with energies 250 and 100 GeV/nucleon respectively, resulting in a kinematic reach (at $Q^2 = 10$ GeV^2) down to $x \sim 10^{-3}$ [33]. Smaller x values could be probed at the LHeC (50 GeV electrons on $Z/A \times 7$ TeV/nucleon protons and nuclei) whose kinematic reach goes down to $x \sim 10^{-5}$ [35] and at the future circular collider with the electron-hadron collider configuration (FCC-he) [14] (60 GeV electrons on $Z/A \times 50$ TeV/nucleon protons and nuclei) whose kinematic coverage extends to even lower x. We see that
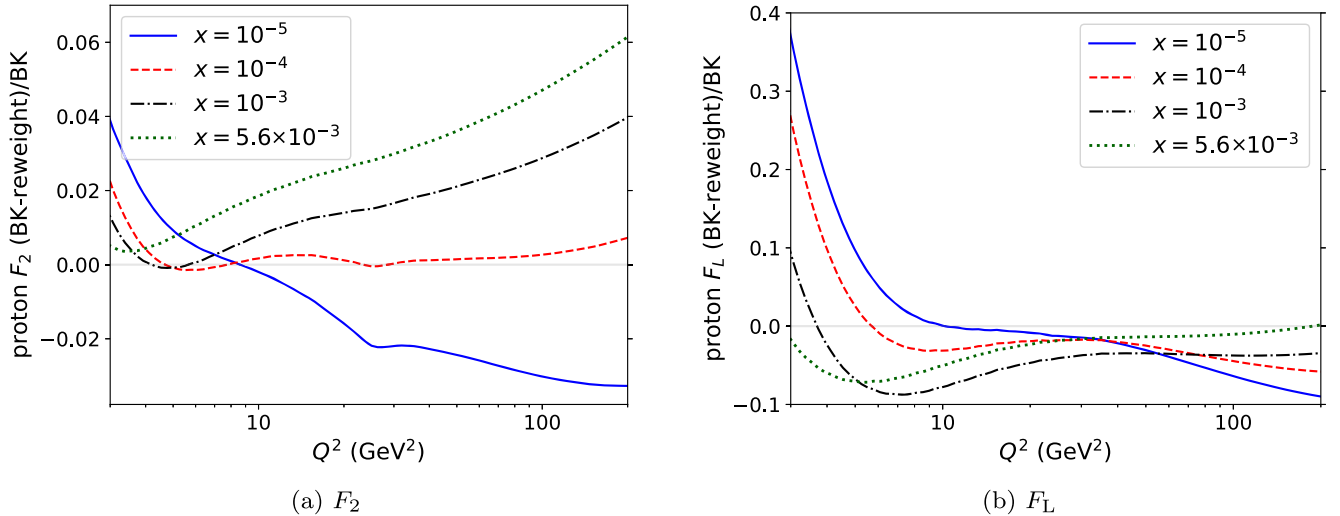

FIG. 3. The relative difference $(F_{2,L}^{BK} - F_{2,L}^{Rew})/F_{2,L}^{BK}$ between the BK predictions and the matched PDF predictions for F_2 (a) and F_L (b) for proton shown as a function of Q^2 for four different x values.
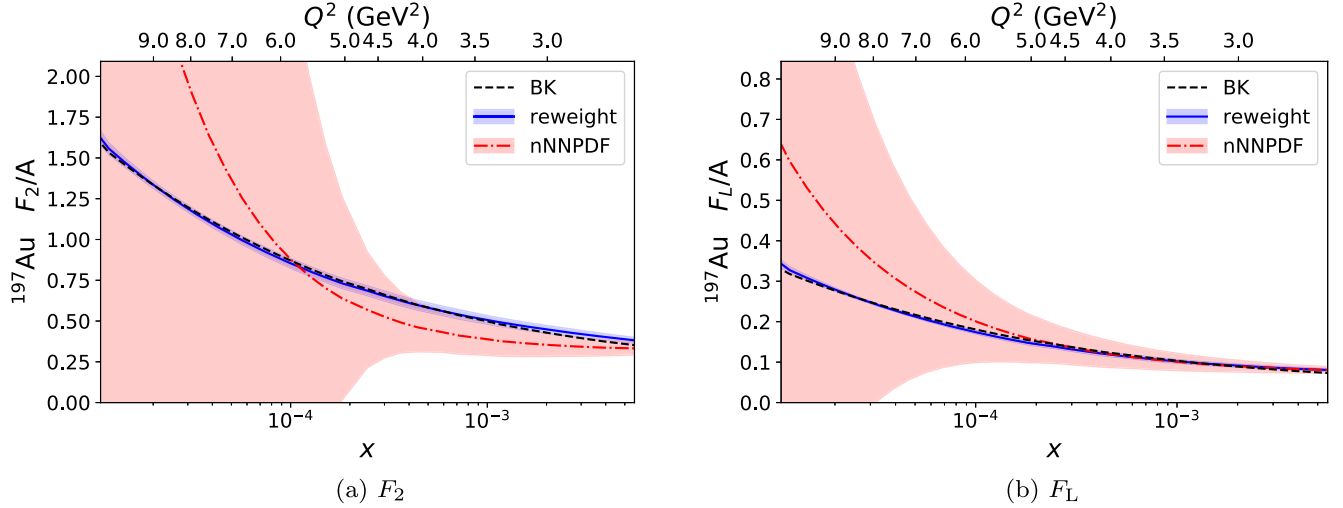
FIG. 4. The F_2 (a) and F_L (b) structure functions for ^{197}Au as a function of x at $Q^2 = 10Q_s^2(x)$. The black dashed curve shows the BK predictions, the red dash-dotted curve with the red error band the original NNPDF3.1 PDF predictions, and the blue solid curve with a light-blue error band the PDF predictions after the matching.

around $x \sim 10^{-4}$ the Q^2 dependencies are nearly equal in both frameworks. In the higher-x region the BK equation predicts a stronger Q^2 dependence than the DGLAP equation, while in the $x \lesssim 10^{-4}$ region the BK dynamics results with a weaker Q^2 dependence than what the DGLAP equation predicts. As a result, at fixed $Q^2 \sim$ 10 GeV^2 the relative difference changes sign as a function of x. Since the relative differences remain at a few-percent level, a very precise determination of the proton F_2 is required in order to distinguish between the two physical pictures in a statistically meaningful manner.

The differences between the BK and DGLAP dynamics are more clearly visible in the case of the structure function F_L. This can be seen from Figs. 2(b) and 3(b) which show the analogous plots for F_L that were above discussed for F_2. There are now larger differences even within the HERA kinematics as the F_L data from HERA are rather scarce. The DGLAP evolved F_L shows generically a significantly stronger Q^2 dependence in comparison to the BK evolved F_L. The x dependencies at fixed Q^2 are now also clearly different, particularly at $Q^2 \lesssim 10$ GeV^2 where the BK evolution predicts a stronger x dependence than what the reweighted F_L with DGLAP dynamics has. In the EIC kinematics the relative differences are maximally $\sim$10%. In the LHeC kinematics differences can be up to $\sim$40% which would be easier to resolve.

Note that for the proton F_2 in Fig. 3(a) there is a small b quark threshold effect in the FONLL-B scheme at $Q^2 = m_b^2 \approx 24$ GeV^2.

B. Heavy nucleus

Next we study the structure functions F_2 and F_L for the ^{197}Au nucleus. The procedure here is the same as for

protons above: first we match the nuclear PDF predictions with the BK predictions in the region $10Q_s^2(x) < Q^2 < 11Q_s^2(x)$, where Q_s is the same proton saturation scale as what was used above. As the saturation scale of a large nucleus is generically just a few times larger than that of the proton (see e.g., Ref. [48]), at $Q^2 = 10Q_s^2(x)$ the nonlinear effects should still be relatively small and both frameworks are expected to be applicable.

The nuclear structure functions $F_{2,L}$ at $Q^2 = 10Q_s^2(x)$ before and after reweighting are shown in Figs. 4(a) and 4(b). As one can see, this time the reweighting is not as successful as it was in the case of proton. The proton and nuclear PDF sets have exactly the same amount of Monte Carlo replicas but, as there are much fewer experimental constraints for the heavy-nucleus structure functions, there are larger variations between the different replicas (as illustrated by the large uncertainty band before reweighting). Consequently, the remaining uncertainty band calculated with $N_{\text{eff}} = 10$ effective replicas can be much larger than with protons, and a more precise matching would require an even larger number of replicas. Thus the larger error band after reweighting is mostly just a technical limitation, not a physics effect in itself. A partial explanation for the difference between the central values before and after reweighting is that the BK prediction is not constrained by experimental data on hard processes involving heavy nuclei, which do affect the nuclear-PDF predictions.

Next we again study the relative differences between the DGLAP- and BK-evolved structure functions, defined by Eq. (13), on the x, Q^2 plane. The results are shown in Fig. 5(a) for F_2 and in Fig. 5(b) for F_L. In comparison to the proton results shown in Figs. 2(a) and 2(b) we find much larger differences with a heavy nucleus. This is expected, as
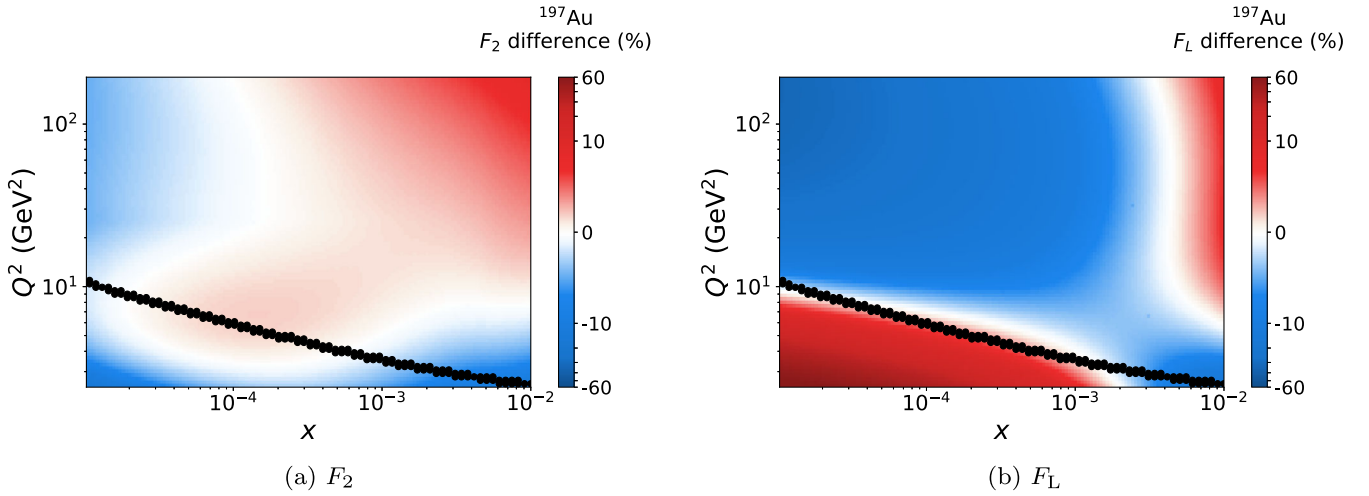
(a) F_2

(b) F_L

FIG. 5. Relative difference $(F_{2,L}^{BK} - F_{2,L}^{Rew})/F_{2,L}^{BK}$ between the BK structure functions and the matched F_2 (a) and F_L (b) for ^{197}Au as a function of x and Q^2. The color scale/axis goes in a linear scale from -10% to 10% and in a logarithmic scale outside that range. The black dots indicate the matching points.
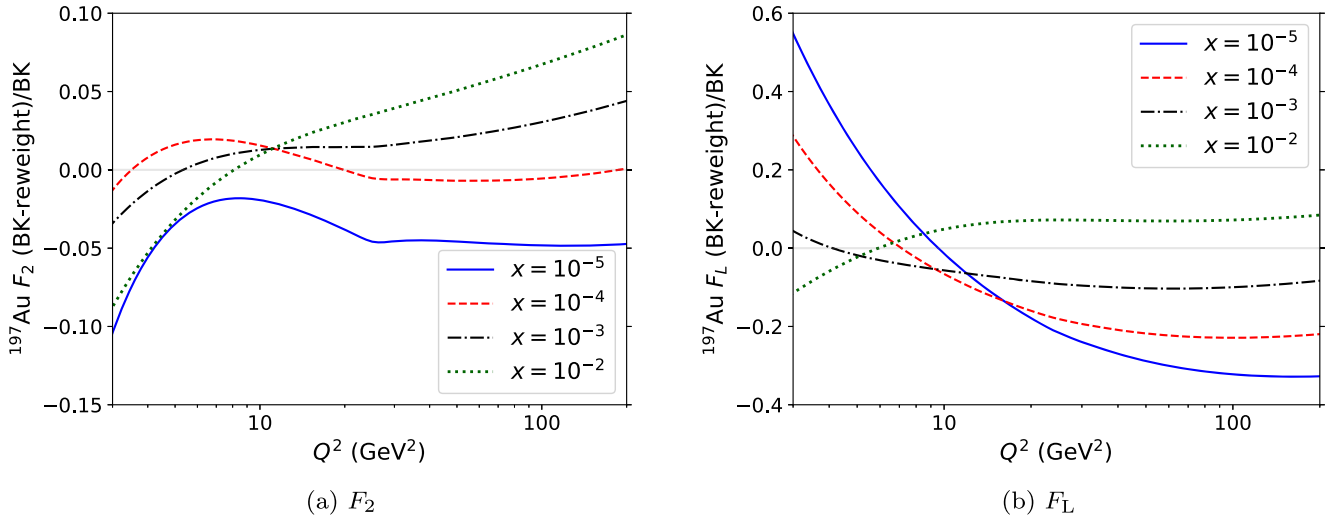

(a) F_2

(b) F_L

FIG. 6. The relative difference $(F_{2,L}^{BK} - F_{2,L}^{Rew})/F_{2,L}^{BK}$ between the BK predictions and the matched PDF predictions for F_2 (a) and F_L (b) for ^{197}Au shown as a function of Q^2 for four different x values.

nonlinear dynamics not present in the DGLAP evolution is enhanced in heavy nuclei by a factor $\sim A^{1/3}$. The fact that the results obtained with the matched PDFs do not exactly agree with the BK prediction results leads to the line of agreement between the two approaches (white in the figure) not exactly aligning with the points where the matching is done. However, the differences between the two evolution equations when moving away from the $Q^2 = 10Q_s^2$ line are still very clear.

Again, we see more dramatic differences in F_L compared to F_2. For F_2 the cleanest systematical effect is that the BK evolution predicts slower x evolution at all Q^2, and this difference in the evolution speed is much larger than with proton targets. In the case of F_L the DGLAP evolution

results in a significantly faster Q^2 evolution except at the highest $x \sim 10^{-2}$, and the x dependence from BK evolution is again generically slower. The fact that BK evolution generically results in a slower x dependence with heavy nuclei is expected, as the effect of the nonlinear contributions in the BK evolution is to slow down the evolution speed.

To quantify more accurately the expected magnitude of the deviations, we show in Figs. 6(a) and 6(b) the relative difference from Figs. 5(a) and 5(b) for fixed x values in the EIC and LHeC/FCC-he kinematics as a function of Q^2. The reweighted F_2 values differ from the BK-evolved results less than 10% in the studied kinematical domain, which implies that also with heavy nuclei the F_2 should be measured at a few-percent precision in order to be sensitive

to the differences in the evolution. For F_L the differences are again significantly larger and up to 60% in the LHeC/FCC-he kinematics and up to 15% in the EIC kinematics. Similarly to protons, a bottom quark mass threshold effect in the FONLL-B scheme is seen in the nuclear F_2 in Fig. 6(a) at $Q^2 \approx 24$ GeV^2.

C. Effects on PDFs

The matching process results in a new PDF set that is constrained by the BK evolved structure functions on the $Q^2 = 10\ldots11 Q_s^2(x)$ line. In this section we study how these constraints originating from the nonlinear dynamics included in the BK evolution alter the parton distribution functions. The coefficients ω_k defined in Eq. (8) determined in the reweighting process for the applied PDF sets are made available in the Supplemental Material [75]. Using these coefficients and the publicly available NNPDF3.1 or nNNPDF2.0 sets it is then possible to construct DGLAP evolved parton distribution functions that match the saturation model predictions at moderate Q^2 and are compatible with the neutral-current HERA data.

First we illustrate here the effect of reweighting on the proton parton distribution functions. The results are shown in Figs. 7(a) and 7(b) for the gluon and up-quark distributions at two different virtualities. In order to study how complementary the observables F_2 and F_L are, we separately show the results with matching to F_2 and matching to F_L. The central set values from the NPDF3.1 are also shown for comparison. Comparing Figs. 7(a) and 7(b) we find the reweighting to have a slightly stronger effect on the gluon distribution compared to up-quarks, but in general the reweighting has only a modest effect on the proton PDFs. This is expected, as the PDFs are fitted to the same HERA data that is used to constrain the BK boundary

conditions. Whether F_2 or F_L is used in reweighting has only a small effect on the determined reweighted PDFs. Thus, we do not expect to see strong tensions when measurements from the EIC or LHeC/FCC-he are eventually used to disentangle the BK and DGLAP dynamics.

The reweighted nuclear up-quark and gluon distributions are shown in Figs. 8(a) and 8(b). Comparing to the proton results shown in Figs. 7(a) and 7(b) we see that nuclear PDFs are affected much more by the reweighting already in the $x \lesssim 10^{-3}$ region, which is expected, as in nNNPDF2.0 there are only few data constraints in this region. The reweighted nuclear PDFs are suppressed by a large factor compared to the central values from the nNNPDF2.0 set. Again both F_2 and F_L pseudodata have similar effects and as such no strong tensions with already existing data included in the nuclear PDF fits are expected in global analyses. In Fig. 8(a) the nuclear gluon distribution, reweighted with F_2 data, becomes negative at small $x \lesssim 2 \times 10^{-5}$ and at $Q^2 = 3.1$ GeV^2. However, the gluon distribution is not an observable, and structure functions remain positive.

IV. CONCLUSIONS

Understanding the importance of nonlinear effects is a major motivation for future high-energy scattering experiments. Experimental measurements have, however, a limited precision and kinematical range. It is therefore essential to quantify in a clear way the magnitude of the effects that one is looking for. In this paper we have done this by comparing the DIS structure functions F_2 and F_L calculated using the running coupling BK equation, which includes saturation effects, to collinear factorization at NLO. The approaches require a set of nonperturbative initial conditions in some initial line in the (x, Q^2) plane and,
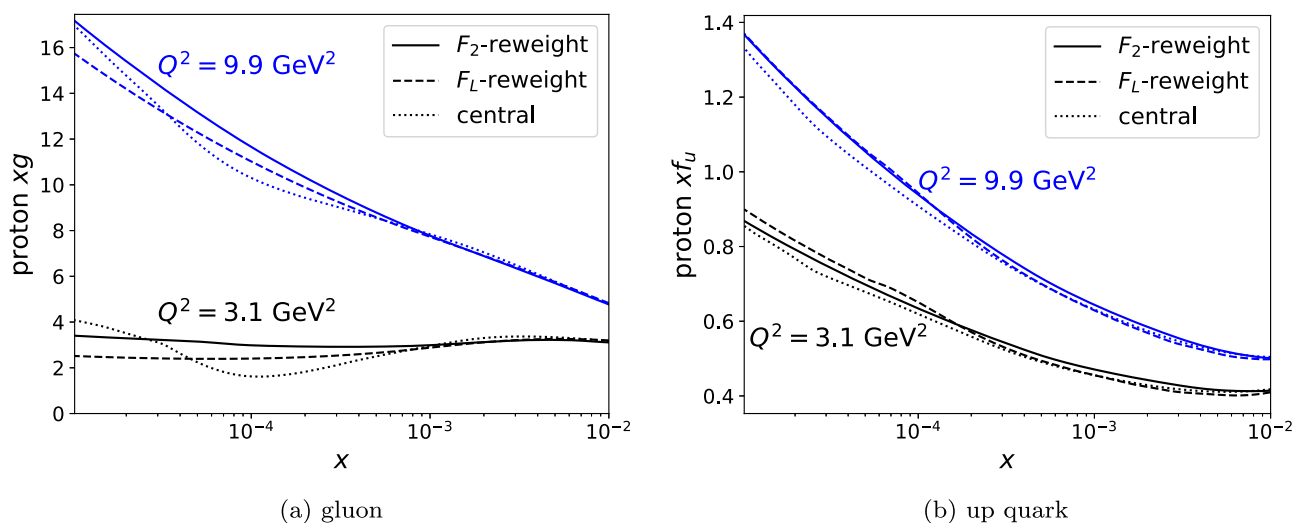

(a) gluon

(b) up quark

FIG. 7. Reweighted (solid and dashed curves) and central (dotted curves) gluon (a) and up-quark (b) NNPDF3.1 proton PDFs as a function of x at low (black) and moderate Q^2 (blue). Results matched with the BK predictions for F_2 (solid curves) and F_L (dashed curves) are shown separately.
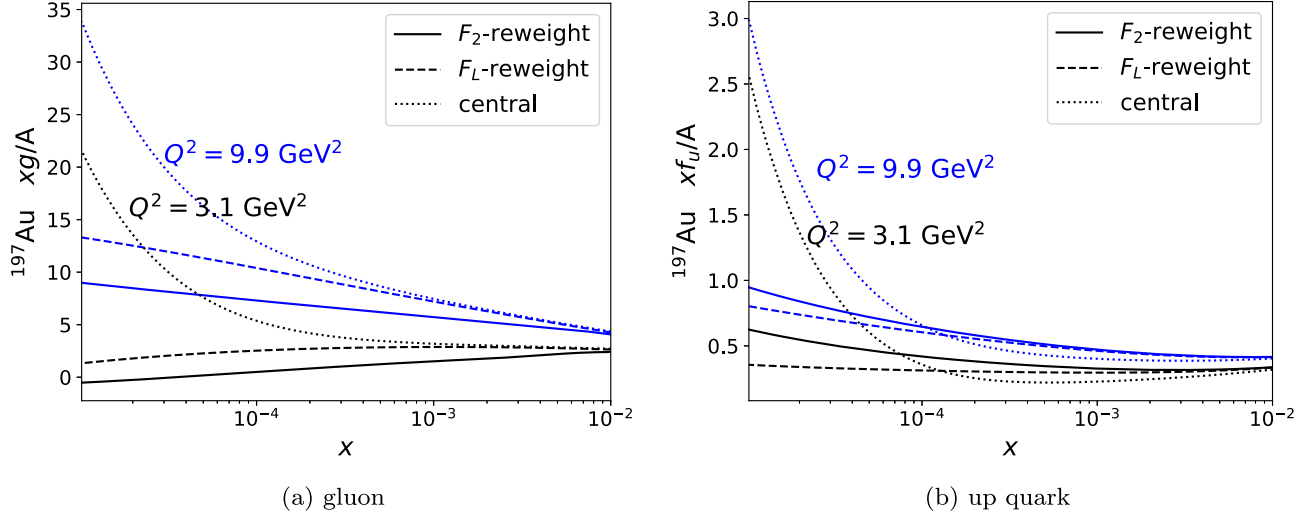
FIG. 8. Reweighted (solid and dashed curves) and central (dotted curves) gluon (a) and up-quark (b) nNNPDF2.0 ^{197}Au PDFs as a function of x at low (black) and moderate Q^2 (blue). Results matched with the BK predictions for F_2 (solid curves) and F_L (dashed curves) are shown separately.

given these initial conditions, make predictions in other regions. In order to compare the two approaches to evolution it is essential to differentiate between the effects of these initial conditions, and the actual dynamical predictions of the evolution. Our focus here is on the evolution dynamics. Thus we perform this comparison by choosing the non-perturbative initial conditions so that they match in a regime where both frameworks should be valid, at virtualities somewhat above the saturation scale. Technically this is achieved by profiting from large samples of DGLAP-evolved Monte-Carlo replicas of PDFs available from the NNPDF collaboration, and using reweighting to match them to the structure functions from BK evolution. There is already a significant agreement due to the fact that both calculations have been (for protons) fit to the same HERA data. The deviations of the two calculations, when moving away from the matching line, provide a clear assessment of the quantitative difference in the dynamics between the two.

As a result of this analysis, we see that F_2 has a quite weak sensitivity to the difference between the evolution equations, at least in the region where Q^2 is large enough for the PDF sets to be available. The relative differences in F_2 are at a percent level for protons, and somewhat higher for nuclei. The longitudinal structure function F_L is significantly more sensitive, but unfortunately also harder to measure as precisely. Overall the difference, as expected, is strongest at the lowest Q^2 values. Pushing collinear factorization to lower scales would clearly make the difference more manifest. In finer details, the behavior of the difference in the x, Q^2-plane also has more unexpected aspects.

We have here focused on providing a comparison of the predictions on a purely theoretical level, independently of any assumptions about a specific experiment. To turn this calculation into a more controlled estimate relevant to a specific measurement requires, in addition, a more detailed inclusion of estimated experimental errors, their correlations, the expected kinematical coverage and so on. This we leave to future work. Another interesting avenue for the future would be to use this as a starting point for using the two approaches together. It should be possible to use BK evolution to predict the behavior at relatively low Q^2 values and very low x and combine this with DGLAP evolution towards higher Q^2. Such a combined framework would significantly extend the predictive power of weak coupling QCD by replacing a part of what are now initial conditions fitted to data by predictions from a first-principles calculation. It would be interesting to perform the matching required in such a calculation analytically. This is, however, beyond the scope of this work. Instead we take a first step in this direction by providing, as Supplemental Material to this paper, the DGLAP-evolved PDF sets for protons and gold nuclei that do match the BK-evolved cross section at our chosen matching scale.

APPENDIX: MATCHING AT A HIGHER SCALE

Figures 9–11 show the result of applying our matching procedure in the kinematical range $Q^2 = 25...28Q_s^2(x)$

for the gold nucleus. These figures should be contrasted with Figs. 4–6, where the same quantities are plotted with our default matching scale. One can see that an equally good matching at this higher scale can be achieved. In this case there is much more phase space available below the matching scale, and at the lower Q^2 values the disagreement between the two calculations becomes quite substantial. This indicates that a factor ~ 2.5 in Q^2 is already sufficient to reveal a substantial difference between what is in some sense a backward DGLAP evolution towards lower scales, and the BK dynamics. This is, however, not relevant for an actual determination based on experimental data.
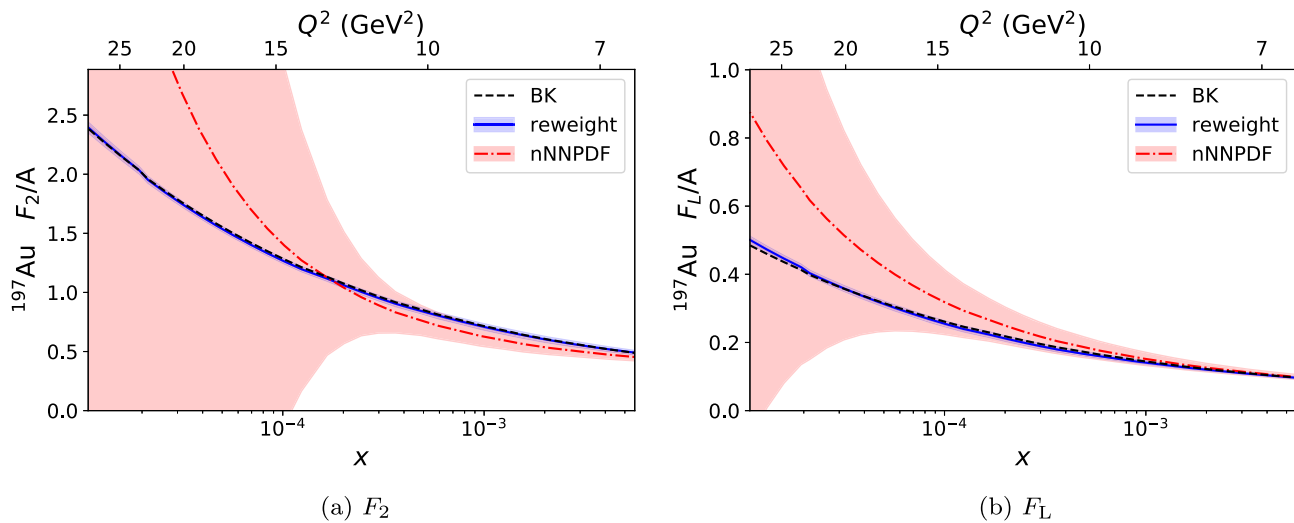

FIG. 9. The F_2 (a) and F_L (b) structure functions for ^{197}Au as a function of x at $Q^2 = 25...28Q_s^2(x)$. The black dashed curve shows the BK predictions, the red dash-dotted curve with the red error band the original NNPDF3.1 PDF predictions, and the blue solid curve with a light-blue error band the PDF predictions after the matching.
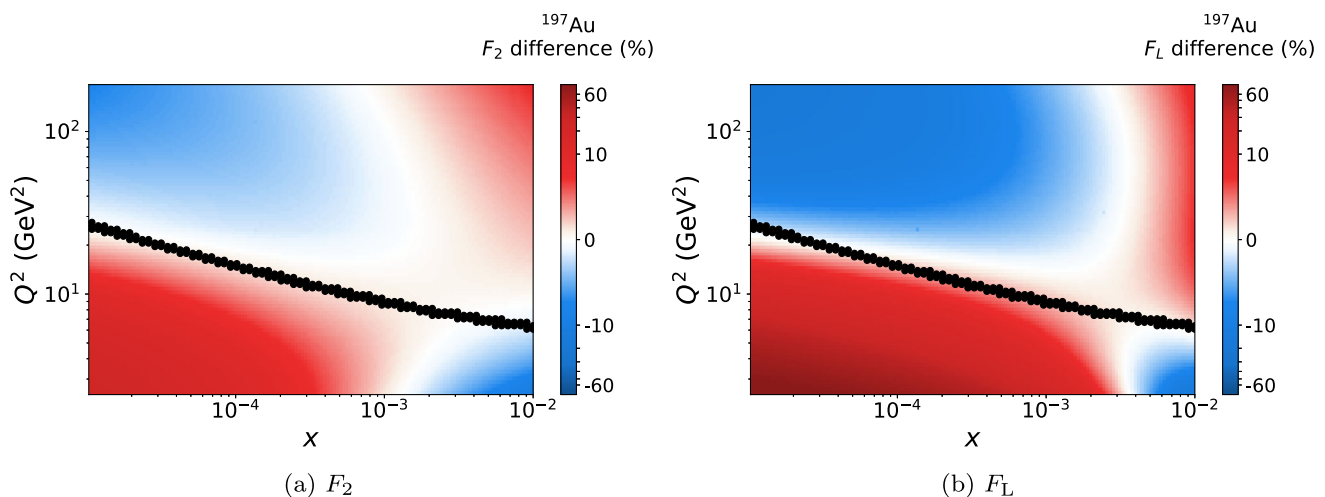

FIG. 10. Relative difference $(F_{2,L}^{BK} - F_{2,L}^{Rew})/F_{2,L}^{BK}$ between the BK structure functions and the matched F_2 (a) and F_L (b) for ^{197}Au as a function of x and Q^2. The color scale/axis goes in a linear scale from -10% to 10% and in a logarithmic scale outside that range. The black dots indicate the matching points.
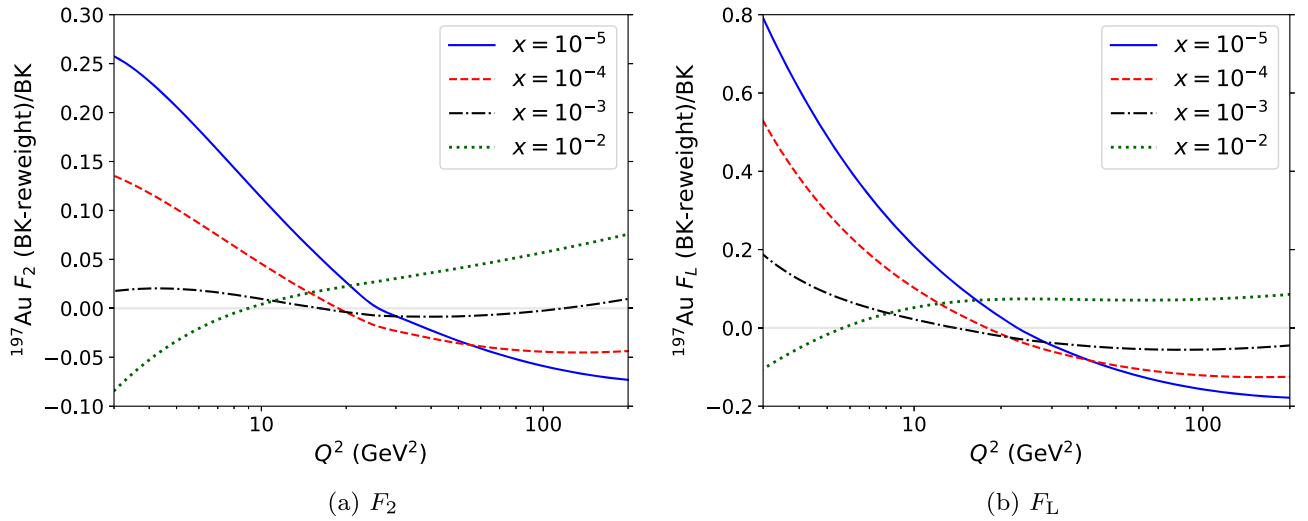
FIG. 11.   The relative difference $(F_{2,L}^{\text{BK}} - F_{2,L}^{\text{Rew}})/F_{2,L}^{\text{BK}}$ between the BK predictions and the matched PDF predictions for F_2 (a) and F_L (b) for ^{197}Au shown as a function of Q^2 for four different x values.

First, experimentally there is more phase space and thus more data available at the lowest scales. Second, and more important, allowing both calculations to adjust nonperturbative initial conditions to the same data will tend to match them at a scale which makes them differ less, which in this case is at smaller values of Q^2.

[1] Y. V. Kovchegov and E. Levin, *Quantum Chromodynamics at High Energy* (Cambridge University Press, Cambridge, England, 2012), Vol. 33, 10.1017/CBO9781139022187.

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

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

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

[5] 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).

[6] E. A. Kuraev, L. N. Lipatov, and V. S. Fadin, The Pomeranchuk singularity in non-Abelian gauge theories, Sov. Phys. JETP **45**, 199 (1977).

[7] I. I. Balitsky and L. N. Lipatov, The Pomeranchuk singularity in quantum chromodynamics, Sov. J. Nucl. Phys. **28**, 822 (1978).

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

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

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

[11] I. Balitsky, Operator expansion for high-energy scattering, Nucl. Phys. **B463**, 99 (1996).

[12] Y. V. Kovchegov, Small x F_2 structure function of a nucleus including multiple Pomeron exchanges, Phys. Rev. D **60**, 034008 (1999).

[13] Y. V. Kovchegov, Unitarization of the BFKL Pomeron on a nucleus, Phys. Rev. D **61**, 074018 (2000).

[14] A. Abada *et al.* (FCC Collaboration), FCC physics opportunities: Future circular collider conceptual design report volume 1, Eur. Phys. J. C **79**, 474 (2019).

[15] A. Abada *et al.* (FCC Collaboration), FCC-hh: The Hadron Collider: Future circular collider conceptual design report volume 3, Eur. Phys. J. Special Topics **228**, 755 (2019).

[16] J. Gao, L. Harland-Lang, and J. Rojo, The structure of the proton in the LHC precision era, Phys. Rep. **742**, 1 (2018).

[17] F. D. Aaron *et al.* (H1, ZEUS Collaborations), Combined measurement and QCD analysis of the inclusive $e^{\pm}p$ scattering cross sections at HERA, J. High Energy Phys. 01 (2010) 109.

[18] H. Abramowicz *et al.* (H1, 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).

[19] 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).

[20] H. Abdolmaleki *et al.* (xFitter Developers' Team Collaboration), Impact of low-x resummation on QCD analysis of HERA data, Eur. Phys. J. C **78,** 621 (2018).

[21] J. L. Albacete, J. G. Milhano, P. Quiroga-Arias, and J. Rojo, Linear vs nonlinear QCD evolution: From HERA data to LHC phenomenology, Eur. Phys. J. C **72,** 2131 (2012).

[22] M. Bonvini, Small-x phenomenology at the LHC and beyond: HELL3.0 and the case of the Higgs cross section, Eur. Phys. J. C **78,** 834 (2018).

[23] K. Kovarik *et al.,* nCTEQ15—Global analysis of nuclear parton distributions with uncertainties in the CTEQ framework, Phys. Rev. D **93,** 085037 (2016).

[24] H. Khanpour and S. Atashbar Tehrani, Global analysis of nuclear parton distribution functions and their uncertainties at next-to-next-to-leading order, Phys. Rev. D **93,** 014026 (2016).

[25] K. J. Eskola, P. Paakkinen, H. Paukkunen, and C. A. Salgado, EPPS16: Nuclear parton distributions with LHC data, Eur. Phys. J. C **77,** 163 (2017).

[26] I. Helenius, M. Walt, and W. Vogelsang, TUJU21: NNLO nuclear parton distribution functions with electroweak-boson production data from the LHC, arXiv:2112.11904.

[27] R. Abdul Khalek, J. J. Ethier, J. Rojo, and G. van Weelden, nNNPDF2.0: Quark flavor separation in nuclei from LHC data, J. High Energy Phys. 09 (2020) 183.

[28] K. J. Eskola, P. Paakkinen, H. Paukkunen, and C. A. Salgado, EPPS21: A global QCD analysis of nuclear PDFs, Eur. Phys. J. C **82,** 413 (2022).

[29] K. J. Eskola, I. Helenius, P. Paakkinen, and H. Paukkunen, A QCD analysis of LHCb D-meson data in p + Pb collisions, J. High Energy Phys. 05 (2020) 037.

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

[31] Z. Citron *et al.,* Report from Working Group 5: Future physics opportunities for high-density QCD at the LHC with heavy-ion and proton beams, CERN Yellow Rep. Monogr. **7,** 1159 (2019).

[32] 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).

[33] R. Abdul Khalek *et al.,* Science requirements and detector concepts for the electron-ion collider: EIC yellow report, arXiv:2103.05419.

[34] J. L. Abelleira Fernandez *et al.* (LHeC Study Group Collaboration), A large hadron electron collider at CERN: Report on the physics and design concepts for machine and detector, J. Phys. G **39,** 075001 (2012).

[35] P. Agostini *et al.* (LHeC, FCC-he Study Group Collaboration), The large hadron-electron collider at the HL-LHC, J. Phys. G **48,** 110501 (2021).

[36] C. Marquet, B.-W. Xiao, and F. Yuan, Semi-inclusive deep inelastic scattering at small x, Phys. Lett. B **682,** 207 (2009).

[37] H. Mäntysaari and R. Venugopalan, Systematics of strong nuclear amplification of gluon saturation from exclusive vector meson production in high energy electron-nucleus collisions, Phys. Lett. B **781,** 664 (2018).

[38] J. Rojo and F. Caola, in *17th International Workshop on Deep-Inelastic Scattering and Related Subjects* (Science Wise Publishing, Berlin, Germany, 2009), p. 222.

[39] C. Marquet, M. R. Moldes, and P. Zurita, Unveiling saturation effects from nuclear structure function measurements at the EIC, Phys. Lett. B **772,** 607 (2017).

[40] V. Bertone, S. Carrazza, and J. Rojo, APFEL: A PDF evolution library with QED corrections, Comput. Phys. Commun. **185,** 1647 (2014).

[41] 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).

[42] R. D. Ball *et al.* (NNPDF Collaboration), Parton distributions from high-precision collider data, Eur. Phys. J. C **77,** 663 (2017).

[43] R. A. Khalek, R. Gauld, T. Giani, E. R. Nocera, T. R. Rabemananjara, and J. Rojo, nNNPDF3.0: Evidence for a modified partonic structure in heavy nuclei, arXiv: 2201.12363.

[44] S. Forte, E. Laenen, P. Nason, and J. Rojo, Heavy quarks in deep-inelastic scattering, Nucl. Phys. **B834,** 116 (2010).

[45] R. D. Ball, V. Bertone, M. Bonvini, S. Forte, P. Groth Merrild, J. Rojo, and L. Rottoli, Intrinsic charm in a matched general-mass scheme, Phys. Lett. B **754,** 49 (2016).

[46] E. Iancu and R. Venugopalan, The color glass condensate and high-energy scattering in QCD, *Quark Gluon Plasma 3* (World Scientific, Singapore, 2003), p. 249.

[47] I. Balitsky, Quark contribution to the small-x evolution of color dipole, Phys. Rev. D **75,** 014001 (2007).

[48] T. Lappi and H. Mäntysaari, Single inclusive particle production at high energy from HERA data to proton-nucleus collisions, Phys. Rev. D **88,** 114020 (2013).

[49] J. L. Albacete, N. Armesto, J. G. Milhano, P. Quiroga-Arias, and C. A. Salgado, AAMQS: A nonlinear QCD analysis of new HERA data at small-x including heavy quarks, Eur. Phys. J. C **71,** 1705 (2011).

[50] J. Berger and A. M. Stasto, Exclusive vector meson production and small-x evolution, J. High Energy Phys. 01 (2013) 001.

[51] J. Berger and A. M. Stasto, Small x nonlinear evolution with impact parameter and the structure function data, Phys. Rev. D **84,** 094022 (2011).

[52] J. Berger and A. Stasto, Numerical solution of the nonlinear evolution equation at small x with impact parameter and beyond the LL approximation, Phys. Rev. D **83,** 034015 (2011).

[53] H. Mäntysaari and B. Schenke, Confronting impact parameter dependent JIMWLK evolution with HERA data, Phys. Rev. D **98,** 034013 (2018).

[54] H. Mäntysaari and H. Paukkunen, Saturation and forward jets in proton-lead collisions at the LHC, Phys. Rev. D **100,** 114029 (2019).

[55] B. Ducloué, T. Lappi, and H. Mäntysaari, Forward J/ψ production at high energy: Centrality dependence and mean transverse momentum, Phys. Rev. D **94,** 074031 (2016).

[56] B. Ducloué, T. Lappi, and H. Mäntysaari, Forward J/ψ production in proton-nucleus collisions at high energy, Phys. Rev. D **91,** 114005 (2015).

[57] T. Lappi and H. Mäntysaari, Forward dihadron correlations in deuteron-gold collisions with the Gaussian approximation of JIMWLK, Nucl. Phys. **A908,** 51 (2013).

[58] Y.-Q. Ma, R. Venugopalan, K. Watanabe, and H.-F. Zhang, $\psi(2S)$ versus J/ψ suppression in proton-nucleus collisions from factorization violating soft color exchanges, Phys. Rev. C **97**, 014909 (2018).

[59] H. Fujii and K. Watanabe, Heavy quark pair production in high energy pA collisions: Quarkonium, Nucl. Phys. **A915**, 1 (2013).

[60] P. Tribedy and R. Venugopalan, Saturation models of HERA DIS data and inclusive hadron distributions in p + p collisions at the LHC, Nucl. Phys. **A850**, 136 (2011); Erratum, Nucl. Phys. **A859**, 185 (2011).

[61] G. Beuf, T. Lappi, and R. Paatelainen, Massive quarks at one loop in the dipole picture of deep inelastic scattering, arXiv:2112.03158.

[62] G. Beuf, H. Hänninen, T. Lappi, and H. Mäntysaari, Color glass condensate at next-to-leading order meets HERA data, Phys. Rev. D **102**, 074028 (2020).

[63] G. Beuf, T. Lappi, and R. Paatelainen, Massive quarks in NLO dipole factorization for DIS: Longitudinal photon, Phys. Rev. D **104**, 056032 (2021).

[64] H. Hänninen, T. Lappi, and R. Paatelainen, One-loop corrections to light cone wave functions: The dipole picture DIS cross section, Ann. Phys. (Amsterdam) **393**, 358 (2018).

[65] G. Beuf, Dipole factorization for DIS at NLO: Combining the $q\bar{q}$ and $q\bar{q}g$ contributions, Phys. Rev. D **96**, 074033 (2017).

[66] G. Beuf, Dipole factorization for DIS at NLO: Loop correction to the $\gamma^*_{T,L} \to q\bar{q}$ light-front wave functions, Phys. Rev. D **94**, 054016 (2016).

[67] T. Lappi, H. Mäntysaari, and A. Ramnath, Next-to-leading order Balitsky-Kovchegov equation beyond large N_c, Phys. Rev. D **102**, 074027 (2020).

[68] T. Lappi and H. Mäntysaari, Next-to-leading order Balitsky-Kovchegov equation with resummation, Phys. Rev. D **93**, 094004 (2016).

[69] T. Lappi and H. Mäntysaari, Direct numerical solution of the coordinate space Balitsky-Kovchegov equation at next to leading order, Phys. Rev. D **91**, 074016 (2015).

[70] I. Balitsky and G. A. Chirilli, Next-to-leading order evolution of color dipoles, Phys. Rev. D **77**, 014019 (2008).

[71] R. D. Ball, V. Bertone, F. Cerutti, L. Del Debbio, S. Forte, A. Guffanti, J. I. Latorre, J. Rojo, and M. Ubiali (NNPDF Collaboration), Reweighting NNPDFs: The W lepton asymmetry, Nucl. Phys. **B849**, 112 (2011); Erratum, Nucl. Phys. **B854**, 926 (2012); Erratum, Nucl. Phys. **B855**, 927 (2012).

[72] R. D. Ball, V. Bertone, F. Cerutti, L. Del Debbio, S. Forte, A. Guffanti, N. P. Hartland, J. I. Latorre, J. Rojo, and M. Ubiali, Reweighting and unweighting of parton distributions and the LHC W lepton asymmetry data, Nucl. Phys. **B855**, 608 (2012).

[73] W. T. Giele and S. Keller, Implications of hadron collider observables on parton distribution function uncertainties, Phys. Rev. D **58**, 094023 (1998).

[74] H. Paukkunen and P. Zurita, PDF reweighting in the Hessian matrix approach, J. High Energy Phys. 12 (2014) 100.

[75] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevD.105.114017, for the coefficients determined in the reweighting process.