

**This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.**

**Author(s):** Dumitru, Adrian; Mäntysaari, Heikki; Paatelainen, Risto

**Title:** High-energy dipole scattering amplitude from evolution of low-energy proton light-cone wave functions

**Year:** 2023

**Version:** Published version

**Copyright:** © Funded by SCOAP3. Published by the American Physical Society

**Rights:** CC BY 4.0

**Rights url:** <https://creativecommons.org/licenses/by/4.0/>

**Please cite the original version:**

Dumitru, A., Mäntysaari, H., & Paatelainen, R. (2023). High-energy dipole scattering amplitude from evolution of low-energy proton light-cone wave functions. *Physical Review D*, 107, Article 114024. <https://doi.org/10.1103/PhysRevD.107.114024>

# High-energy dipole scattering amplitude from evolution of low-energy proton light-cone wave functions

Adrian Dumitru<sup>\*</sup>

*Department of Natural Sciences, Baruch College, CUNY, 17 Lexington Avenue,  
New York, New York 10010, USA  
and The Graduate School and University Center, The City University of New York,  
365 Fifth Avenue, New York, New York 10016, USA*

Heikki Mäntysaari<sup>†</sup>

*Department of Physics, University of Jyväskylä, P.O. Box 35, 40014 University of Jyväskylä, Finland  
and Helsinki Institute of Physics, P.O. Box 64, 00014 University of Helsinki, Finland*

Risto Paatelainen<sup>‡</sup>

*Helsinki Institute of Physics and Department of Physics, FI-00014 University of Helsinki, Finland*



(Received 6 April 2023; accepted 30 May 2023; published 16 June 2023)

The forward scattering amplitude of a small dipole at high energies is given in the mean field approximation by the Balitsky-Kovchegov (BK) evolution equation. It requires an initial condition  $N(r; x_0)$  describing the scattering of a dipole with size  $r$  off the target that is probed at momentum fraction  $x_0$ . Rather than using *ad hoc* parametrizations tuned to high-energy data at  $x \ll x_0$ , here we attempt to construct an initial scattering amplitude that is consistent with low-energy, large- $x$  properties of the proton. We start from a nonperturbative three quark light-cone model wave function from the literature. We add  $\mathcal{O}(g)$  corrections due to the emission of a gluon, and  $\mathcal{O}(g^2)$  virtual corrections due to the exchange of a gluon, computed in light-cone perturbation theory with exact kinematics. We provide numerical data as well as analytic parametrizations of the resulting  $N(r; x_0)$  for  $x_0 = 0.01$ – $0.05$ . Solving the BK equation in the leading logarithmic approximation towards lower  $x$ , we obtain a fair description of the charm cross section in deeply inelastic scattering measured at HERA by fitting one parameter, the coupling constant  $\alpha_s \simeq 0.2$ . However, without the option to tune the initial amplitude at  $x_0$ , the fit of the high precision data results in  $\chi^2/N_{\text{dof}} = 2.3$  at  $N_{\text{dof}} = 38$ , providing clear statistical evidence for the need of *systematic* improvement, e.g., of the photon wave function, evolution equation, and initial condition.

DOI: 10.1103/PhysRevD.107.114024

## I. INTRODUCTION

In deep inelastic scattering (DIS) a pointlike virtual photon probes the rich QCD dynamics taking place inside the proton or a nucleus. At high energies, where the small Bjorken- $x$  part of the target wave function is probed, one observes very large gluon densities [1]. When the gluon densities become of the same order as inverse coupling, nonlinear QCD dynamics start to dominate and multiple scattering effects are important [2]. In the high-energy

limit, the scattering process is most conveniently described in the dipole picture in a frame where the virtual photon has a large momentum [3], and its partonic Fock states, such as  $|q\bar{q}\rangle$  at leading order (LO), have a long lifetime as they scatter from the color field of the target.

Describing the QCD dynamics in this high-density domain is natural in the color glass condensate (CGC) [4] framework. Here the center-of-mass energy or Bjorken- $x$  dependence of various observables (and as such the target structure) is described in the large- $N_c$  limit by the perturbative Balitsky-Kovchegov (BK) renormalization group equation [5,6]. It describes how the dipole-target scattering amplitude, which contains information about the target structure, changes with increasing energy. The dipole amplitude (a correlator of two Wilson lines) is actually a convenient degree of freedom at high energies: all cross sections computed at high energy in the CGC framework are expressed in terms of the dipole amplitude or higher-point correlators which can be written, in a

<sup>\*</sup>adrian.dumitru@baruch.cuny.edu

<sup>†</sup>heikki.mantysaari@jyu.fi

<sup>‡</sup>risto.paatelainen@helsinki.fi

*Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP<sup>3</sup>.*

Gaussian approximation, in terms of the two-point dipole amplitude [7].

The initial condition for the dipole-proton scattering amplitude depends on nonperturbative properties of the proton. A typical approach in the field has been to assume an intuitive functional form at an initial  $x_0 \ll 1$  and fit various unknown parameters to the HERA total cross section data; see, e.g., Refs. [8–10] where a very good description of small- $x$  HERA data is obtained at leading order, resumming powers of  $\alpha_s \ln 1/x$  via BK evolution with running coupling corrections [11]. Recent developments to full next-to-leading order (NLO) accuracy have also allowed for a simultaneous description of total and heavy quark production data [12,13]. The drawback of this approach is that one is sensitive to the assumed functional form of the initial dipole amplitude and that the model parameters need to be refitted if the evolution is initialized at different  $x_0$ . Furthermore, there is no relation to the low energy (or “large- $x$ ”) proton structure.

In this work, we take a complementary approach aiming to *compute* the initial dipole-proton scattering amplitude at moderate  $x_0$ . As we will discuss in more detail next, the necessary nonperturbative input consists in a proton valence quark wave function that is constrained by low-energy data. The  $x_0$ -dependent initial condition is then obtained by computing the dipole-target scattering amplitude including one perturbative gluon emission in the target, with the gluon longitudinal momentum fraction regulated by  $x_0$  [14,15]. The advantages of this approach are that we do not assume an *ad hoc* functional form of the scattering amplitude and that the initial condition can be computed and the BK evolution initialized at any (moderate)  $x_0$  without a need to perform new fits. Also, this approach largely eliminates the freedom of tuning initial conditions in order to optimally match the evolution equation to the small- $x$  data. This may reveal quantitative evidence for the need for improvements beyond leading log, or even running coupling BK evolution.

Finally, we would like to point out that light-cone Hamiltonian calculations of wave functions have been employed previously to set initial conditions for QCD scale evolution to high virtuality  $Q^2$ , in order to describe DIS in the “dipole approach” using correlators of eikonal Wilson lines as degrees of freedom [16–19]. Our approach is similar in spirit although here the goal is to determine initial conditions for evolution to small  $x$ .

## II. DIPOLE-PROTON SCATTERING AT MODERATE $x$

We first provide an overview of our approach to the light-cone structure of the proton. We employ a truncated Fock space description which starts with a three quark state. The corresponding Fock space amplitude (wave function)  $\Psi_{qqq}$  corresponds to a nonperturbative solution of the QCD light-front Hamiltonian. To date, exact solutions for the light-cone

wave functions are not available. In the future, lattice gauge theory may provide numerical solutions for moderate parton momentum fractions  $x_i$  and transverse momenta  $\vec{k}_i$  via a large momentum expansion of equal-time Euclidean correlation functions in instant quantization [20,21]; see Ref. [22] for a recent lattice computation of the wave function of the leading  $q\bar{q}$  state of the pion. Also, the MAP collaboration [23] has recently extracted the wave functions of the first four Fock states of the pion from fits to its parton distribution functions and electromagnetic form factor.

Here, we rely on solutions of effective light-cone Hamiltonians for guidance on the low-energy and low virtuality  $Q^2$  structure of the proton. Specifically, we shall employ the HO wave function of Refs. [24,25]. In these references, the authors fixed the parameters of the three quark wave function to the proton “radius,” or Dirac form factor at  $Q^2 \rightarrow 0$ , to the anomalous magnetic moments of the proton and neutron, and to the axial vector coupling  $g_A$ . The wave function also matches reasonably well the empirical knowledge of the longitudinal and transverse momentum distribution of single quarks in the valence quark regime. Finally, the wave function of Refs. [24,25] also provides predictions for quark *momentum correlations*.

At next-to-leading order (NLO) in the Fock expansion we add the three quarks and one gluon state with amplitude  $\Psi_{qqqg}$ , as well as the virtual corrections to  $\Psi_{qqq}$  due to the exchange of a gluon by two quarks in the proton. These corrections are obtained via light-cone perturbation theory calculations [14,15]. The presence or exchange of the gluon extends the range of parton light-cone momentum fractions to lower  $x$ , and pushes their transverse momenta into the perturbative regime. It also affects their momentum correlations.

The central element of our analysis is the (imaginary part of the) eikonal scattering amplitude  $N$  of a small dipole of transverse size  $\mathbf{r}$ . The real part of  $N$  corresponds to two-gluon exchange,

$$\begin{aligned}
 N(\mathbf{r}, \mathbf{b}) = & -g^4 C_F \int \frac{d^2 \mathbf{K} d^2 \mathbf{q}}{(2\pi)^4} \frac{\cos(\mathbf{b} \cdot \mathbf{K})}{(\mathbf{q} - \frac{1}{2} \mathbf{K})^2 (\mathbf{q} + \frac{1}{2} \mathbf{K})^2} \\
 & \times \left( \cos(\mathbf{r} \cdot \mathbf{q}) - \cos\left(\frac{\mathbf{r} \cdot \mathbf{K}}{2}\right) \right) \\
 & \times G_2\left(\mathbf{q} - \frac{1}{2} \mathbf{K}, -\mathbf{q} - \frac{1}{2} \mathbf{K}\right). \quad (1)
 \end{aligned}$$

Here  $\mathbf{K}$  is the momentum transfer which is Fourier conjugate to the impact parameter  $\mathbf{b}$ . As explained below, we will eventually average  $N(\mathbf{r}, \mathbf{b})$  over a suitable range of impact parameters. We emphasize that the expression above accounts only for a single, perturbative two-gluon exchange (see its derivation in Ref. [26]), it does not resum the Glauber-Mueller multiple scattering series. This restricts its applicability to the regime of weak scattering,  $N(\mathbf{r}, \mathbf{b}) \ll 1$ . Furthermore,  $N(\mathbf{r}, \mathbf{b})$  actually acquires an imaginary part due to the perturbative exchange of three

gluons; its magnitude has been shown to be much smaller than its real part [27,28], and in practice it is of interest only for processes involving  $C$ -conjugation odd exchanges [29]. For the present purposes, it can be neglected.

The coupling of the two static gluons to the proton is described in terms of the color charge density correlator,

$$\langle \rho^a(\mathbf{q}_1) \rho^b(\mathbf{q}_2) \rangle \equiv \delta^{ab} g^2 G_2(\mathbf{q}_1, \mathbf{q}_2). \quad (2)$$

The color charge density operator corresponds to the light-cone plus component of the color current on the  $x^+ = 0$  light front, integrated over  $x^-$ ,  $\rho^a(\mathbf{q}) \equiv J^{+a}(\mathbf{q})$ , when the proton carries positive  $P_z$ . Equations (1) and (2) correspond to the leading twist contribution to the matrix element of the dipole operator in the proton. Dozens of diagrams contribute to this correlator at NLO, their explicit expressions are listed in Ref. [14]. We point out that  $G_2(\mathbf{q}_1, \mathbf{q}_2)$  satisfies a Ward identity due to the color neutrality of the proton; it vanishes when either  $q_1^2$  or  $q_2^2 \rightarrow 0$  so that  $N(\mathbf{r}, \mathbf{b})$  in Eq. (1) is free of IR divergences. However,  $G_2$  does exhibit a collinear singularity which is regularized by assigning a mass to the quarks in the light-cone energy denominators for the  $q \rightarrow qg$  and  $qg \rightarrow q$  vertices; see Ref. [14] for details. All the results presented here were obtained with  $m_{\text{coll}} = 0.2$  GeV. This is consistent with the quark mass and transverse momentum scales which appear in the nonperturbative three quark wave function of Refs. [24,25]. The color charge correlator also exhibits a soft singularity when the light-cone momentum fraction  $x_g$  of the gluon goes to zero. This is regularized with a cutoff  $x$  on  $x_g$ , and the resummation of yet softer gluons will be performed through the BK equation. Note that at  $x = 0.1$  the NLO contribution to  $G_2$  truly is a reasonably small  $\mathcal{O}(g^2)$  perturbative correction [15]. However, by  $x = 0.01$  its magnitude grows to essentially  $\mathcal{O}(1)$ , a leading-log correction. Hence, at such  $x$  resummation is required and it is justified to use the computed dipole as an initial condition for the leading order BK evolution.

We recall, also, that at the given order ultraviolet divergences cancel [14], so that  $G_2$  is independent of the renormalization scale, and that the coupling does not run. Lastly, let us mention that the angular dependence of the correlator  $G_2$ , as well as the dependence of its Fourier transform on impact parameter, has been analyzed numerically in detail in Ref. [15].

### III. SMALL- $x$ EVOLUTION OF THE PROTON LIGHT-CONE WAVE FUNCTION

In order to obtain an initial condition for  $\mathbf{b}$ -independent BK evolution<sup>1</sup> we average the dipole-target scattering amplitude obtained from Eq. (1) over the impact parameter  $\mathbf{b}$ ,

<sup>1</sup>We limit ourselves to the  $\mathbf{b}$ -independent evolution in order to avoid the need to effectively model confinement scale effects which has been attempted, e.g., in Refs. [30,31].

$$N(r, x_0) = \frac{1}{S_T} \int^{b_{\text{max}}} d^2 \mathbf{b} N(\mathbf{r}, \mathbf{b}, x_0). \quad (3)$$

Throughout this work, we denote the magnitudes of the transverse vectors as  $b = |\mathbf{b}|$  and  $r = |\mathbf{r}|$ . The resulting amplitude is dominated by perturbative contributions when the dipole size  $r$  is small. In this region there is a small  $\cos(2\phi)$  dependence on the angle  $\phi$  between  $\mathbf{r}$  and  $\mathbf{b}$  [15] which vanishes when we integrate over  $\mathbf{b}$ . Here  $S_T$  is the proton transverse area. Inclusive cross sections considered in this work are not sensitive to the actual shape of the target but only to the total transverse size. The proton geometry is most directly probed in exclusive vector meson production process where the total momentum transfer  $\mathbf{K}$  which is Fourier conjugate to the impact parameter is measurable. Parametrizing the  $J/\psi$  production cross section in HERA kinematics as  $e^{-B_D \mathbf{K}^2}$ , one obtains  $B_D = 4$  GeV<sup>-2</sup> [32]. Assuming a Gaussian impact parameter profile for the proton, this corresponds to a two-dimensional root-mean-square radius  $b_{\text{Gaussian}} = \sqrt{2B} \approx 0.56$  fm and a proton area  $S_T = 2\pi B$ . On the other hand, if we assume a step function (hard sphere) profile for the proton, the same diffractive slope is obtained when the proton radius is  $b_{\text{Hard sphere}} = 2\sqrt{B} \approx 0.79$  fm, which corresponds to  $S_T = 4\pi B$ .

Although exclusive vector meson data favors the Gaussian density profile over the hard sphere one (see e.g. [33]), the current data does not constrain the proton shape precisely. We also note that if the  $\mathbf{b}$ -dependent dipole amplitude from Eq. (1) is directly used to compute exclusive  $J/\psi$  production cross section, the resulting spectra differs from the Gaussian profile case only in the region where there are no experimental constraints [34]. In this work the results shown below by default correspond to the Gaussian density profile (with  $b_{\text{max}} = b_{\text{Gaussian}}$ ) unless otherwise stated, but we also study the dependence on the  $b_{\text{max}}$  cut by using a step function profile with  $b_{\text{max}} = b_{\text{Hard sphere}} = \sqrt{2} b_{\text{Gaussian}}$ .

The proton transverse area  $S_T$  has also been extracted by fitting a parametrized initial condition for the BK evolution equation to the HERA structure function data. Leading order analyses [9,10] typically obtain  $S_T \sim 16$  mb. In recent fits at NLO accuracy [12,13] proton areas  $S_T \sim 10$ – $20$  mb were obtained depending on the details of the analysis setup. We test this uncertainty in the proton small- $x$  transverse profile by showing some results for both the Gaussian and hard sphere profiles with transverse areas 9.8 and 19.6 mb, respectively.

Before performing the impact parameter average we first study the impact parameter profile from the NLO light-cone wave function seen by a perturbative probe:

$$T(\mathbf{b}) = C \int^{r_{\text{max}}} d^2 \mathbf{r} N(\mathbf{r}, \mathbf{b}, x). \quad (4)$$

The normalization condition  $\int^{b_{\max}} d^2\mathbf{b}T(\mathbf{b}) = 1$  is used to fix the constant  $C$ . We will refer to  $T(\mathbf{b})$  as the transverse “density” profile to match standard terminology from the literature. As the dipole amplitude is a rapidly increasing function of the dipole size  $r$ , this integral is dominated by dipoles of size  $r \sim r_{\max}$ , as long as  $r_{\max}$  is in the perturbative domain.

The extracted density profiles up to  $b_{\max} = 0.8$  fm for different  $r_{\max}$  are shown in Fig. 1. For reference, a Gaussian profile, as used e.g. in the popular IPSat parametrization [35] for the dipole amplitude with the slope  $B = 4$  GeV<sup>-2</sup>, is also shown. We observe a similar transverse profile except for very central  $b \lesssim 0.2$  fm where the computed profile is more steeply falling. This region can only be probed at high momentum transfer  $|t| \gtrsim 1$  GeV<sup>2</sup> [34], which is not covered in the currently available coherent vector meson production data. The high- $b$  tails of  $T(\mathbf{b})$  resulting from the light cone perturbation theory (LCPT) one gluon emission corrections are exponential rather than Gaussian. However, in all we conclude that for the present purposes the Gaussian profile used to match  $S_T$  to  $b_{\max}$  is a reasonable approximation.

The  $\mathbf{b}$ -averaged dipole amplitudes (using a Gaussian profile) are shown in Fig. 2 (linear scale) and Fig. 3 (logarithmic scale) at  $x_0 = 0.05$ ,  $x_0 = 0.025$  and  $x_0 = 0.01$ . Here we also show the dependence on the diffractive slope  $B$ : the bands correspond to varying  $B$  by  $\pm 10\%$  which changes both  $S_T$  and  $b_{\max}$ . The results depend weakly on this cut especially in the perturbative small- $r$  domain. The dipole amplitude increases with  $r$ , approximately proportional to  $r^2$ , as expected. For  $r \gtrsim 0.4$  fm the color neutrality of the proton, and the fact that the dipole scatters from a target of finite transverse extent, begin to slow the growth of  $N(r)$ ; a model that does not account for the finite size of the proton in impact parameter space would attribute this to power corrections. Finally, when the size of the dipole becomes comparable to that of the target the amplitude is found to decrease again (not shown) as the end points of the

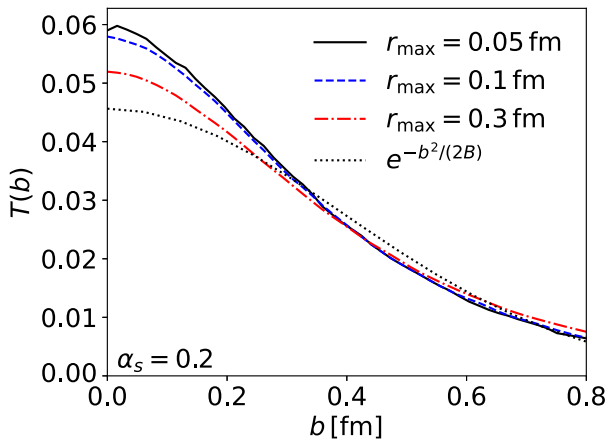


FIG. 1. Effective normalized proton density profiles at  $x = 0.01$  extracted from  $N(\mathbf{r}, \mathbf{b})$  with  $b_{\max} = 0.8$  fm up to NLO in the Fock expansion of the light-cone wave function.

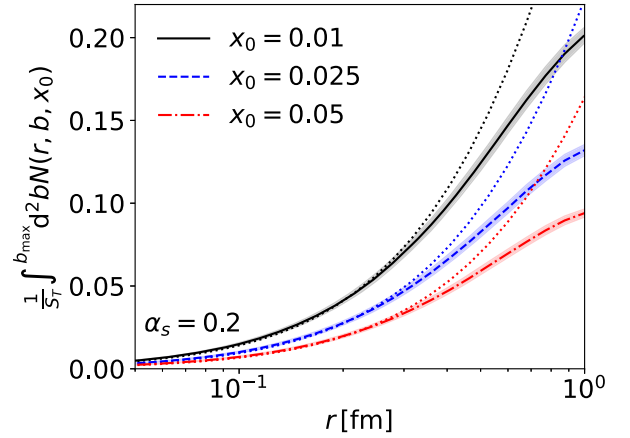


FIG. 2. Impact-parameter averaged dipole as a function of dipole size  $r$  at two different momentum fractions  $x$ . The bands correspond to varying the proton shape parameter  $B$  by 10%. The dotted lines show best fits to the central values with the modified McLerran-Venugopalan (MV) model parametrization of Eq. (5).

dipole essentially “miss” the target. However, we emphasize that this behavior occurs at large  $r \sim$  few fm where in any case the perturbative calculation of the scattering amplitude is not valid.

Figures 2 and 3 confirm that down to  $x = 0.01$  scattering of small dipoles with  $r$  significantly less than 1 fm remains quite weak, at least for  $\alpha_s = 0.2$  which we determine below from a fit to the charm cross section in DIS. Therefore, it appears reasonable to start small- $x$  evolution with this initial condition at  $x$  in the range 0.01–0.05.

To obtain analytic parametrizations of the dipole amplitude we fit our numerical data for the  $\mathbf{b}$ -averaged scattering amplitude to the following expression which is inspired by the McLerran-Venugopalan (MV) model [36]:

$$N(r) = 1 - \exp \left[ -\frac{(r^2 Q_{s,0}^2)^\gamma}{4} \ln \left( \frac{1}{r\Lambda} + e_c \cdot e \right) \right], \quad (5)$$

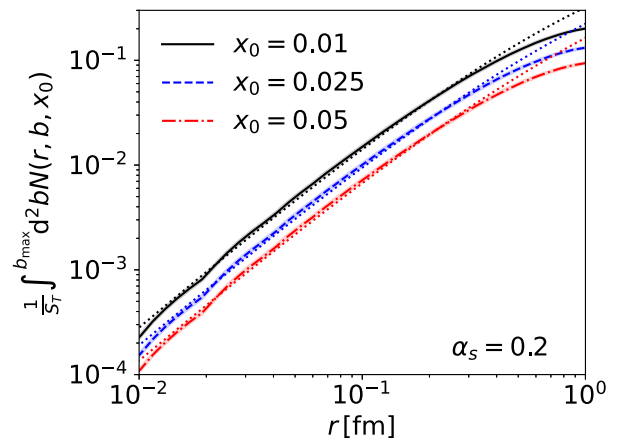


FIG. 3. Same as Fig. 2 but on a double logarithmic scale in order to better exhibit the behavior at small  $r$ .

where  $\Lambda = 0.241$  GeV is a fixed infrared scale. Such a parametrization has been used previously, e.g., in Refs. [9,10,12] to fit the initial condition for BK evolution to the HERA data. While our fit is restricted to  $r < 0.5$  fm, the parametrization forces  $N(r) \rightarrow 1$  in the large- $r$  region. Of course, the behavior at large  $r$  cannot be trusted, and other extrapolations would be possible. It is important, however, that the large- $r$  extrapolation is such that the Fourier transform of  $1 - N(r)$  at high  $k$  (and, consequently, the forward particle production cross section, for example) will be insensitive to the assumed form. We will also demonstrate below that perturbative observables, in our case the charm production cross section, are only sensitive to the perturbative regime of small dipoles where our calculation should apply, and not to the extrapolation to large  $r$ .

The free parameters in Eq. (5),  $Q_{s,0}^2$ ,  $\gamma$  and  $e_c$ , are fit to the calculated dipole amplitude in the region  $0.01 \text{ fm} < r < 0.5 \text{ fm}$  (we actually fit the logarithm of the dipole in order to give equal weight to small and intermediate  $r$ ). The upper limit restricts to the perturbative domain, and the lower limit is imposed in order to give some weight to the region of intermediate  $r$  as well. The resulting dipole amplitudes are shown in Figs. 2 and 3 as dotted lines. The fit parameters are listed in Table I. In the fit we require that  $e_c > e^{-1}$  in order to enforce positivity of the logarithm in Eq. (5), and all fit results give  $e_c = e^{-1}$  within numerical accuracy, i.e. they require as small an infrared cutoff as allowed. The MV-model inspired parametrization is found to describe the dipole-proton scattering amplitude quite well, for all dipole sizes in the perturbative  $r \lesssim 0.5$  fm region. Here, of course, the linearized version of Eq. (5) is sufficient, as it should be: recall that Eq. (1) does not resum multiple scattering.

The momentum scale  $Q_{s,0}^2$  remains nonperturbative down to  $x = 0.01$ ; see below for an extraction of a “saturation scale” at lower  $x$ . However, it increases approximately as  $Q_{s,0}^2 \sim 1/x^{0.47}$ . The “anomalous dimension” of the dipole amplitude is  $\gamma = 1$  within numerical accuracy. It appears reasonable to us that the initial condition for the evolution equation admits a power series expansion in  $r^2$ , starting at its first power. On the other hand, leading order fits to HERA total cross section data [9,10] require  $\gamma \sim 1.1$ – $1.2$  in order to obtain as slow a  $Q^2$  dependence of the cross section as required by the HERA data [1,37]; recent fits at next-to-leading order accuracy performed in Refs. [12,13] also prefer  $\gamma \gtrsim 1$  when the heavy quark production data is included. A problem with  $\gamma > 1$  is that it renders the (dipole) unintegrated gluon distribution function [7,38,39] and the forward particle production cross section negative [10,40] in some range of transverse momentum  $\mathbf{k}_T$ . The dipole amplitude obtained here does not display this issue.

Next we solve the leading order BK equation with fixed coupling, using the numerical data for  $N(r, x_0)$  as an initial condition at  $x_0 = 0.01$ . Note that at this order in  $\alpha_s$  the

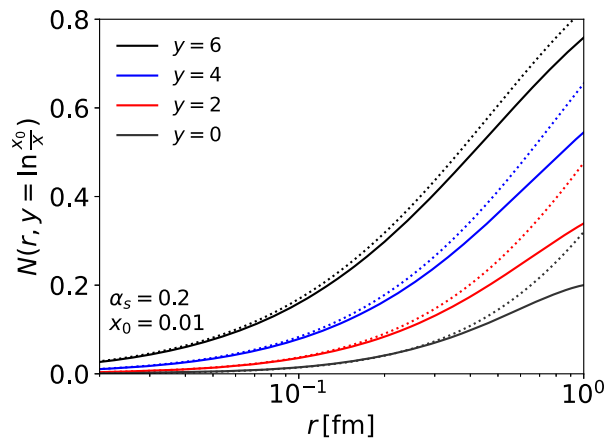


FIG. 4. Leading-log BK evolution at  $\alpha_s = 0.2$  starting at  $x_0 = 0.01$ . From bottom to top, the lines correspond to evolution rapidity 0, 2, 4, and 6. The dashed lines are obtained with the fitted MV-model like parametrization from Table I as an initial condition; the solid lines evolve the actual numerical data for  $N(|\mathbf{r}|, x_0)$ .

coupling constant does not run in the LCPT calculation of the initial condition, and consequently we also limit ourselves to the fixed coupling case here. Evolution over six units of rapidity is shown in Fig. 4. For comparison, we also solve the BK equation using the modified MV-model initial condition with parameters as shown in Table I. This parametrized initial condition has a completely different behavior in the infrared region with  $N(r) \rightarrow 1$  at large  $r$  whereas the numerical data gives a decreasing  $N(r)$  when  $r$  exceeds a few fm, as already mentioned above. However, as can be seen in Fig. 4 the resulting BK-evolved dipole amplitudes are basically identical in the perturbative  $r \lesssim 0.5$  fm domain. In fact, due to the approach to the fixed point of the BK equation [41–43], at high rapidity the difference between the scattering amplitudes evolved with the two initial conditions diminishes. This demonstrates that the BK-evolved amplitude at small  $r$  is not affected by the uncontrolled large- $r$  extrapolation of the initial condition.

One may define a saturation radius  $r_s$ , and a corresponding saturation momentum  $Q_s = \sqrt{2}/r_s$  from the condition that  $N(r_s) = 1 - \exp(-\frac{1}{2}) \simeq 0.4$ . For this to be a perturbative scale requires about six units of rapidity evolution, as can also be seen from Fig. 4. This corresponds to  $x \simeq 2.5 \times 10^{-5}$ , where  $r_s \simeq 0.3$  fm, and  $Q_s \simeq 1$  GeV. These values are not very far from the first “saturation model” fit to HERA DIS data by Golec-Biernat and Wüsthoff [44] from 25 years ago. Many more recent fits mentioned above have since confirmed that reaching the strong scattering regime with a small dipole and a proton target requires deep evolution to rather small  $x$ . Also, some studies [17] of diffractive small- $x$  scattering of a  $q\bar{q} - g$  state from the proton have indicated that the regime of “color transparency” sets in when the typical transverse distance of the gluon from the  $q\bar{q}$  is around 0.2 fm.

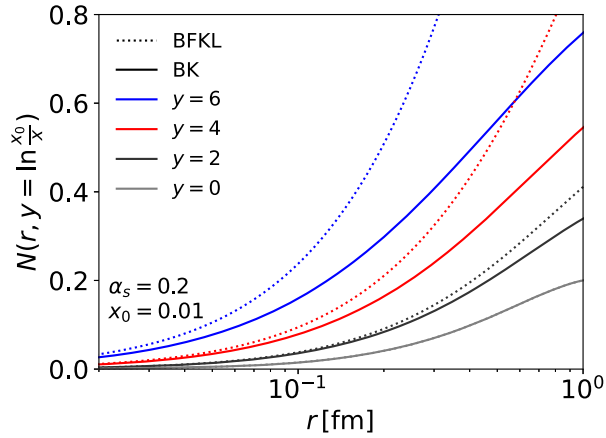


FIG. 5. Leading-log BK (solid lines) vs BFKL (dotted lines) evolution starting at  $x_0 = 0.01$ .

Since scattering at  $x = 0.01$  is fairly weak we have also evolved our initial condition with the linear Balitsky-Fadin-Kuraev-Lipatov (BFKL) equation [45–47], see Fig. 5. After a few units of rapidity evolution, the linear equation begins to violate unitarity,  $N(r) \leq 1$ , at large  $r$ . However, this regime of large dipoles is not under control in any case. More importantly though, at  $y = 2$ – $4$  the absence of the nonlinear correction begins to affect the solution significantly even at  $r$  substantially less than 1 fm. With BFKL we also noticed a greater difference between evolving the actual numerical data vs the analytic modified MV-model parametrization (not shown), which differ in their large- $r$  extrapolation. Therefore, for accurate results it appears to be rather important to evolve with the nonlinear BK equation even if one restricts to  $r < 1$  fm.

Let us finally study how the  $x$  dependence obtained from the direct, fixed order NLO LCPT calculation differs to the one obtained by solving the BK equation. We note that in the LCPT calculation  $x$  is an explicit cutoff for the longitudinal momentum of the emitted gluon, and this gluon emission is calculated in exact kinematics. On the other hand, in BK evolution multiple soft gluon emissions are resummed. This comparison is done by calculating the dipole amplitude at  $x = 0.01$  directly from the LCPT using Eq. (1), and comparing that to the dipole amplitude obtained by solving the BK equation with the initial

TABLE I. Best fit parameters to the modified MV model parametrization, Eq. (5), for a fit over  $0.01 \text{ fm} < r < 0.5 \text{ fm}$ . The upper and lower limits are obtained by varying the proton shape parameter  $B$  by  $\pm 10\%$ . All fit results give  $e_c = e^{-1}$  within numerical accuracy.

$x$	$Q_{s,0}^2 [\text{GeV}^2]$	$\gamma$	$e_c$
0.01	$0.100^{+0.004}_{-0.004}$	$1.001^{+0.001}_{-0.001}$	$e^{-1}$
0.025	$0.066^{+0.003}_{-0.003}$	$0.998^{+0.001}_{-0.001}$	$e^{-1}$
0.05	$0.047^{+0.002}_{-0.002}$	$0.997^{+0.001}_{-0.001}$	$e^{-1}$

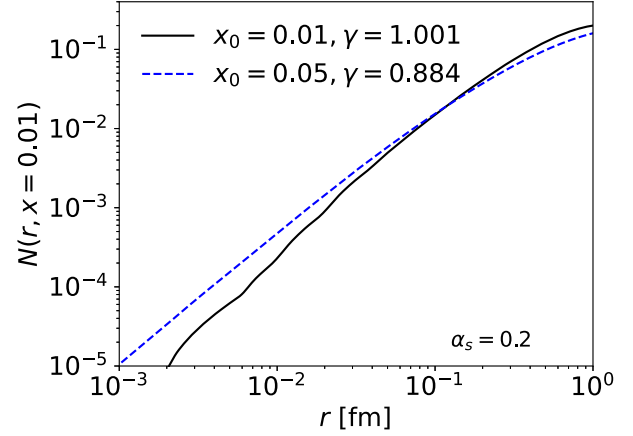


FIG. 6. The dipole scattering amplitude at  $x = 0.01$  with and without prior BK evolution. The parameter  $\gamma$  denotes the resulting anomalous dimension fitted in the region  $0.01 \text{ fm} < r < 0.5 \text{ fm}$ .

condition computed at  $x_0 = 0.05$ . The resulting dipole amplitudes are shown in Fig. 6. The most significant difference between the fixed order  $\mathcal{O}(g^2)$  LCPT amplitude and the BK evolved dipole is that the evolution equation decreases the anomalous dimension  $\gamma$  towards the asymptotic value  $\gamma \sim 0.6$ . On the other hand, the emission of one gluon in the direct LCPT calculation does not modify the extracted anomalous dimension, as can also be seen from Table. I. This is, of course, the expected behavior. As already mentioned above, DIS phenomenology does not appear to support  $\gamma < 1$  at  $x = 0.01$  or greater, so it seems reasonable to treat at least the emission of the first gluon with  $x_g > 0.01$  explicitly in fixed order light-cone perturbation theory with exact kinematics.<sup>2</sup>

#### IV. TOTAL CROSS SECTION AT SMALL $x$

Next, we consider the DIS structure functions at small Bjorken- $x$ . The overall normalization of the dipole amplitude depends on the strong coupling constant  $\alpha_s = g^2/(4\pi)$ , see Eq. (1). The same coupling constant also affects the Bjorken- $x$  dependence of the dipole amplitude via the BK evolution. In this work, our strategy is to fix the value of  $\alpha_s$  by calculating the total charm production cross section, and comparing it to the HERA reduced cross section data from Ref. [48]. We set the initial condition for the BK evolution at  $x_0 = 0.01$ , and compare it to the HERA data in the region  $x < 0.01$ ,  $Q^2 < 100 \text{ GeV}^2$  (note that the smallest  $Q^2$  bin in the data is  $Q^2 = 2.5 \text{ GeV}^2$ ). In this region, there are  $N = 39$  data points. The experimental data is reported as reduced cross section

<sup>2</sup>Also, the emission of the first gluon actually increases the imaginary part due to  $C$ -odd three gluon exchange [28], which provides another indication that small- $x$  evolution should not be started much before  $x_0 \simeq 0.01$ .

$$\sigma_r(x, y, Q^2) = F_2(x, Q^2) - \frac{y^2}{1 + (1 - y)^2} F_L(x, Q^2). \quad (6)$$

Here  $y = Q^2/(sx)$  is the inelasticity variable, not to be confused with the evolution rapidity. The proton structure functions  $F_2$  and  $F_L$  are expressed in terms of the total cross section for the virtual photon-proton cross section  $\sigma^{\gamma^* p}$ :

$$F_2(x, Q^2) = \frac{Q^2}{4\pi\alpha_{\text{em}}} (\sigma_T^{\gamma^* A} + \sigma_L^{\gamma^* A}), \quad (7)$$

$$F_L(x, Q^2) = \frac{Q^2}{4\pi\alpha_{\text{em}}} \sigma_L^{\gamma^* A}. \quad (8)$$

In the dipole picture, the total cross section for the virtual photon-proton scattering can be expressed as a convolution of the photon wave function and the dipole amplitude as [2]

$$\sigma_{T,L}^{\gamma^* A} = 2 \sum_f \int \int d^2\mathbf{b} d^2\mathbf{r} dz |\Psi^{\gamma^* \rightarrow q\bar{q}}(\mathbf{r}, z, Q^2)|^2 N(\mathbf{r}, \mathbf{b}, \bar{x}). \quad (9)$$

Here  $f$  is the quark flavor,  $Q^2$  the photon virtuality and  $\Psi^{\gamma^* \rightarrow q\bar{q}}$  is the leading order wave function for the  $q\bar{q}$  Fock state of the virtual photon. In this equation we replace  $N(\mathbf{r}, \mathbf{b}, \bar{x})$  by the impact parameter averaged dipole amplitude  $N(r, \bar{x})$ , as described above, and  $\int d^2\mathbf{b} \rightarrow S_T$ . We fix the mass of the  $c$  quark to 1.4 GeV. The dipole amplitude in Eq. (9) is evaluated at  $\bar{x} = x(1 + 4m_q^2/Q^2)$ , where  $m_q$  is the quark mass which enforces a smooth approach to the photoproduction limit [9,44].

In order to confirm that the charm production cross section is not sensitive to nonperturbatively large dipoles, we show in Fig. 7 the fraction of the total cross section at  $x = 0.0056$ ,  $Q^2 = 10 \text{ GeV}^2$  as a function of the upper limit  $r_{\text{max}}$  for the  $r$  integral in Eq. (9). It is evident that the charm cross section is saturated by small dipoles whereas the inclusive cross section (calculated using  $m_q = 0.14 \text{ GeV}$  for the light quarks) at  $Q^2 = 10 \text{ GeV}^2$  is sensitive to larger dipoles beyond sizes where we may trust our calculation. When using a modified MV-model parametrization as an initial condition for the evolution with different extrapolation in the infrared region, one needs to integrate up to even larger  $r$  in order to recover the full result for  $F_2$ . The charm production cross section is not affected by the different large- $r$  extrapolation (not shown). Qualitatively similar results have been obtained with the commonly used IPsat parametrization for the dipole-proton amplitude where, typically, even larger dipole sizes contribute as compared to the setup with factorized impact parameter dependence applied here [35,49]. For these reasons we shall focus on charm production. In the future, our approach could be applied to other hard, perturbative

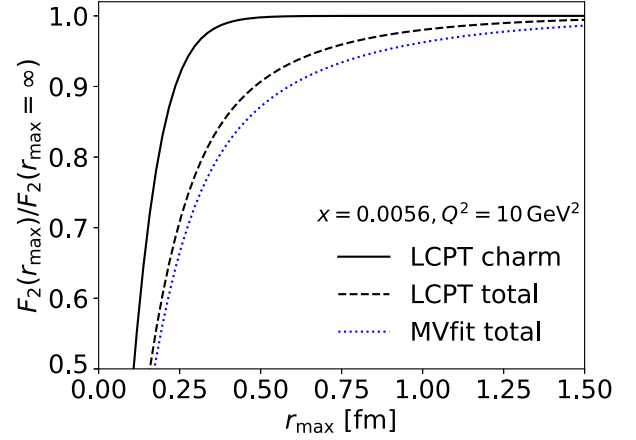


FIG. 7. The fraction of the charm and inclusive DIS structure functions at  $x = 0.0056$ ,  $Q^2 = 10 \text{ GeV}^2$  (corresponding to  $\bar{x} = 0.01$  in the case of charm production) as a function of the cutoff on the dipole size imposed in Eq. (9). The total cross section is the sum of light quark (mass 0.14 GeV) and charm quark (mass 1.4 GeV) production contributions.

processes such as, for example, single-inclusive particle production at high enough transverse momentum.

Considering only the strong coupling constant  $\alpha_s$  as a free parameter, we obtain a reasonably good description of the charm production data. The value of  $\chi^2/(N - 1)$  as a function of  $\alpha_s$  is shown in Fig. 8 using two different density profiles (Gaussian and hard sphere) for the proton. These two setups have different upper limits for the impact parameter  $b$  and correspondingly different transverse areas for the proton. The extracted optimal values for the strong coupling constant are  $\alpha_s = 0.200$  for the Gaussian proton

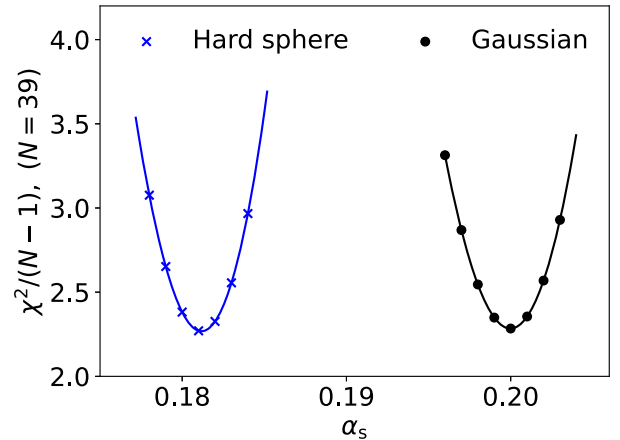


FIG. 8. Determining the strong coupling constant by fitting the HERA charm production data in the region  $Q^2 < 100 \text{ GeV}^2$ ,  $x < 0.01$ , with leading-log BK evolution started at  $x_0 = 0.01$ . The solid lines are polynomial fits to the computed values used to extract the optimal values for  $\alpha_s$ . The optimal values are  $\alpha_s = 0.200$ ,  $\chi^2/(N - 1) = 2.27$  for the Gaussian proton and  $\alpha_s = 0.181$ ,  $\chi^2/(N - 1) = 2.28$  for the hard sphere proton.



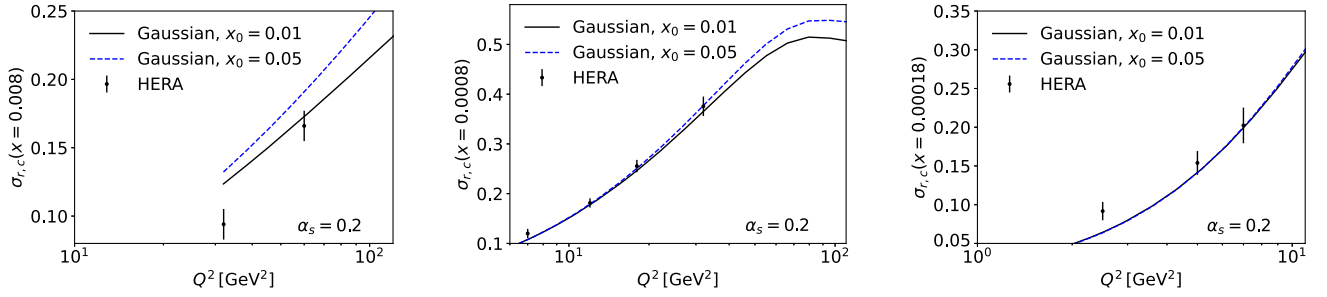


FIG. 9. Charm production reduced cross section at  $\sqrt{s} = 318$  GeV compared to the HERA data. Results are shown in the region where  $\bar{x} \leq 0.01$ .

and  $\alpha_s = 0.181$  for the hard sphere profile. These values are used throughout this paper. We note that fits of the total (rather than charm) cross section with tuned initial conditions [9,10] and running coupling corrections to the BK equation have achieved lower  $\chi^2/N \approx 1$ , without being able to simultaneously describe the charm data [9]. However, with our *calculated* initial condition there is room for the expected improvements of the photon wave function, evolution equation, and, of course, of the initial condition.

In this analysis, we have fixed the collinear regulator to  $m_{\text{coll}} = 0.2$  GeV in the LCPT calculation of the initial condition. As the charm cross section is dominated by small dipoles, our results are not highly sensitive to the actual value of this regulator: changing  $m_{\text{coll}}$  by a factor of 2 changes  $\chi^2/(N-1)$  to HERA data by only 2%–5% when using the optimal  $\alpha_s$ . We also keep the charm mass fixed to 1.4 GeV. The optimal value for  $\alpha_s$  and the achieved  $\chi^2/(N-1)$  naturally will depend on this choice. We choose to work with fixed quark mass and collinear regulator and do not attempt to fit these simultaneously with  $\alpha_s$ , as the purpose of this work is to demonstrate the feasibility of *computing* the initial condition for the BK equation, and we emphasize that numerically potentially important higher order effects such as running coupling are still missing from the setup.

A comparison to the HERA charm production data in different Bjorken- $x$  bins is shown in Fig. 9 as a function of the photon virtuality. We have checked that these results remain the same if we use the analytic parametrization (5), with parameters from Table I, as the initial condition; this confirms the insensitivity of the charm cross section to the large- $r$  extrapolation of the scattering amplitude.

At  $x = 0.008$  there is only very little ( $\leq 0.2$  units of rapidity when  $x_0 = 0.01$ ) evolution, so the dipole amplitude is almost completely determined by our initial condition. On the other hand, we also show results at lower  $x$  which is dominated by the BK evolution. In addition to our standard setup where the initial condition for the BK evolution is set at  $x_0 = 0.01$ , we also show results using an initial condition computed at larger  $x_0 = 0.05$ . Note the weak dependence of this observable, at least, on where the “hand-off” from the  $x$ -dependent initial condition to the

evolution equation occurs. In contrast, *ad hoc* initial condition parametrizations have to be retuned when  $x_0$  is changed.

Figure 9 shows a fair agreement of the reduced cross section obtained from our light-cone wave function with the HERA charm data. Close to the initial condition we obtain a slightly slower  $Q^2$  dependence than seen in the data. As a result of the evolution this changes into faster virtuality dependence at very small  $x$ . This is because the BK evolution at fixed coupling develops a small anomalous dimension  $\gamma \approx 0.6$  for the dipole amplitude and a smaller anomalous dimension results in faster  $Q^2$  dependence.

Lastly, in Fig. 10 we study how sensitive the charm production cross section is on the chosen proton density profile, and as such on the maximum impact parameter  $b_{\text{max}}$  used in Eq. (3). The cross section is calculated at  $x = 0.008$  which is close to the initial condition for the BK evolution, again set at  $x_0 = 0.01$ . In both cases, we use the optimal value for the strong coupling constant extracted above. The cross section increases only slightly when the hard sphere profile with larger  $b_{\text{max}}$  is used, but the dependence on the virtuality is not affected. This weak dependence on the selected proton profile confirms that our

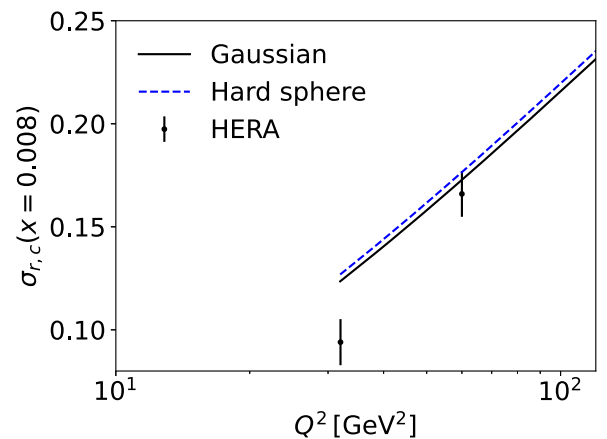


FIG. 10. Reduced cross section at  $\sqrt{s} = 318$  GeV calculated close to the initial condition using the Gaussian and Hard sphere density profiles, respectively.

results are not sensitive to nonperturbatively large impact parameters.

## V. SUMMARY AND DISCUSSION

The present work represents a first attempt at relating the short-distance structure of the proton at high energies to its low-energy properties, covering several orders of magnitude in energy. We start from an effective three quark light-cone wave function which models the nonperturbative longitudinal and transverse momentum distributions of quarks at  $x \gtrsim 0.1$  as well as some of their correlations. The next step involves the computation, using exact kinematics, in light-cone perturbation theory of the  $\mathcal{O}(g)$  correction to the light-cone wave function due to the emission of a gluon, and the  $\mathcal{O}(g^2)$  virtual corrections due to the exchange of a gluon by two quarks. This provides a leading twist contribution to the dipole scattering amplitude,  $N(r) \sim r^{2\gamma}$  with  $\gamma = 1$  at small  $r$ . Optimistically, the LCPT correction extends the validity of the resulting light-cone wave function into the regime of perturbative transverse momenta, and parton momentum fractions  $x \gtrsim 0.01$ . The corresponding dipole scattering amplitude is then evolved to yet higher energies (lower  $x$ ) by solving the BK equation, which resums emissions of additional soft gluons, and generates power corrections and an anomalous dimension.

The convolution of the LO photon light-cone wave function with the impact parameter averaged BK dipole scattering amplitude at leading logarithmic accuracy provides a fair description of the reduced DIS charm cross section measured at HERA, for  $\alpha_s \simeq 0.2$ . This value of the strong coupling was obtained from a fit to the charm cross section at  $Q^2 < 100 \text{ GeV}^2$  and  $x < 0.01$ . None of the parameters of the low-energy three-quark model wave function were re-tuned to the high-energy data. Despite the fair description of the highly accurate data, the resulting  $\chi^2/N_{\text{dof}} \approx 2.27$  with  $N_{\text{dof}} = 38$  implies a very low statistical significance, i.e. a very low probability that the data represents statistical fluctuations about the model predictions: the integral over the  $\chi^2$  distribution from  $\chi^2 = 2.27 \times 38$  to infinity, commonly denoted as the “p value,” is  $1.3 \times 10^{-5}$ . However, the very low statistical significance of the fit should not be confused with a need for large corrections; Fig. 9 shows that this is clearly not the case. This is entirely expected since there are multiple *known* sources of corrections such as, for example, of the photon wave function [50–52], of the evolution equation [53–57], and, of course, of the initial condition for the evolution

equation (our proton light-cone wave function) which, e.g., may be improved with running coupling corrections. The data requires fairly moderate but *systematic* improvements of the model predictions across the relevant ranges of  $x$  and  $Q^2$ .

We have also provided analytic parametrizations of the impact parameter averaged dipole scattering amplitude for  $x = 0.01$ – $0.05$  which accurately fit the numerical data in the regime of perturbative dipoles,  $r \lesssim 0.5 \text{ fm}$ . These parametrizations can be used in practice to estimate the corrections predicted by more accurate evolution equations. In the Supplemental Material [58] we also provide the tabulated numerical data for  $N(r, x)$ . Their large- $r$  extrapolation differs from that of the analytic parametrizations which allows for tests of the (in)sensitivity to the uncontrolled nonperturbative regime of large dipoles. The quest for more accurate theoretical predictions at high energy for the upcoming electron-ion collider at Brookhaven National Laboratory [59] and the proposed LHeC/FCC-he at CERN [60] requires initial conditions for the evolution equations which do not absorb theoretical improvements into a retune of their parameters.

## ACKNOWLEDGMENTS

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics, within the framework of the Saturated Glue (SURGE) Topical Theory collaboration. A. D. acknowledges support by the DOE Office of Nuclear Physics through Grant No. DE-SC0002307, and The City University of New York for PSC-CUNY Research Grant No. 65079-00 53. This work was supported by the Academy of Finland, the Centre of Excellence in Quark Matter and Projects No. 338263 and No. 346567 (H. M.), and Projects No. 347499 and No. 353772 (R.P). H. M. is also supported under the European Union’s Horizon 2020 research and innovation program by the European Research Council (ERC, Grant Agreement No. ERC-2018-ADG-835105 YoctoLHC) and by the STRONG-2020 project (Grant Agreement No. 824093). The content of this article does not reflect the official opinion of the European Union and responsibility for the information and views expressed therein lies entirely with the authors. Computing resources from CSC–IT Center for Science in Espoo, Finland, and from the Finnish Grid and Cloud Infrastructure (persistent identifier [urn:nbn:fi:research-infrastructure-2016072533](https://nbn-resolving.org/urn:nbn:fi:research-infrastructure-2016072533)) were used in this work.

- [1] 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).
- [2] Y. V. Kovchegov and E. Levin, *Quantum Chromodynamics at High Energy* (Cambridge University Press, Cambridge, England, 2012), Vol. 33.
- [3] A. H. Mueller and B. Patel, Single and double BFKL Pomeron exchange and a dipole picture of high-energy hard processes, *Nucl. Phys.* **B425**, 471 (1994).
- [4] F. Gelis, E. Iancu, J. Jalilian-Marian, and R. Venugopalan, The color glass condensate, *Annu. Rev. Nucl. Part. Sci.* **60**, 463 (2010).
- [5] Y. V. Kovchegov, Small- $x$   $F_2$  structure function of a nucleus including multiple Pomeron exchanges, *Phys. Rev. D* **60**, 034008 (1999).
- [6] I. Balitsky, Operator expansion for high-energy scattering, *Nucl. Phys.* **B463**, 99 (1996).
- [7] F. Dominguez, C. Marquet, B.-W. Xiao, and F. Yuan, Universality of unintegrated gluon distributions at small  $x$ , *Phys. Rev. D* **83**, 105005 (2011).
- [8] J. L. Albacete, N. Armesto, J. G. Milhano, and C. A. Salgado, Nonlinear QCD meets data: A global analysis of lepton-proton scattering with running coupling BK evolution, *Phys. Rev. D* **80**, 034031 (2009).
- [9] J. L. Albacete, N. Armesto, J. G. Milhano, P. Quiroga-Arias, and C. A. Salgado, AAMQS: A non-linear QCD analysis of new HERA data at small- $x$  including heavy quarks, *Eur. Phys. J. C* **71**, 1705 (2011).
- [10] 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).
- [11] I. Balitsky, Quark contribution to the small- $x$  evolution of color dipole, *Phys. Rev. D* **75**, 014001 (2007).
- [12] 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).
- [13] H. Hänninen, H. Mäntysaari, R. Paatelainen, and J. Penttala, Proton Structure Functions at NLO in the Dipole Picture with Massive Quarks, *Phys. Rev. Lett.* **130**, 192301 (2023).
- [14] A. Dumitru and R. Paatelainen, Sub-femtometer scale color charge fluctuations in a proton made of three quarks and a gluon, *Phys. Rev. D* **103**, 034026 (2021).
- [15] A. Dumitru, H. Mäntysaari, and R. Paatelainen, Color charge correlations in the proton at NLO: Beyond geometry based intuition, *Phys. Lett. B* **820**, 136560 (2021).
- [16] F. Hautmann, Z. Kunszt, and D. E. Soper, Hard scattering factorization and light cone Hamiltonian approach to diffractive processes, *Nucl. Phys.* **B563**, 153 (1999).
- [17] F. Hautmann and D. E. Soper, Color transparency in deeply inelastic diffraction, *Phys. Rev. D* **63**, 011501 (2001).
- [18] F. Hautmann, Power-like corrections and the determination of the gluon distribution, *Phys. Lett. B* **643**, 171 (2006).
- [19] F. Hautmann and D. E. Soper, Parton distribution function for quarks in an s-channel approach, *Phys. Rev. D* **75**, 074020 (2007).
- [20] X. Ji, Y.-S. Liu, Y. Liu, J.-H. Zhang, and Y. Zhao, Large-momentum effective theory, *Rev. Mod. Phys.* **93**, 035005 (2021).
- [21] X. Ji and Y. Liu, Computing light-front wave functions without light-front quantization: A large-momentum effective theory approach, *Phys. Rev. D* **105**, 076014 (2022).
- [22] M.-H. Chu *et al.*, Transverse-momentum-dependent wave functions of pion from lattice QCD, [arXiv:2302.09961](https://arxiv.org/abs/2302.09961).
- [23] B. Pasquini, S. Rodini, and S. Venturini, The valence quark, sea, and gluon content of the pion from the parton distribution functions and the electromagnetic form factor, [arXiv:2303.01789](https://arxiv.org/abs/2303.01789).
- [24] F. Schlumpf, Relativistic constituent quark model of electro-weak properties of baryons, *Phys. Rev. D* **47**, 4114 (1993); **49**, 6246(E) (1994).
- [25] S. J. Brodsky and F. Schlumpf, Wave function independent relations between the nucleon axial coupling  $g_A$  and the nucleon magnetic moments, *Phys. Lett. B* **329**, 111 (1994).
- [26] A. Dumitru, G. A. Miller, and R. Venugopalan, Extracting many-body color charge correlators in the proton from exclusive DIS at large Bjorken  $x$ , *Phys. Rev. D* **98**, 094004 (2018).
- [27] A. Dumitru, H. Mäntysaari, and R. Paatelainen, Cubic color charge correlator in a proton made of three quarks and a gluon, *Phys. Rev. D* **105**, 036007 (2022).
- [28] A. Dumitru, H. Mäntysaari, and R. Paatelainen, Stronger C-odd color charge correlations in the proton at higher energy, *Phys. Rev. D* **107**, L011501 (2023).
- [29] A. Dumitru and T. Stebel, Multiquark matrix elements in the proton and three gluon exchange for exclusive  $\eta_c$  production in photon-proton diffractive scattering, *Phys. Rev. D* **99**, 094038 (2019).
- [30] J. Berger and A. M. Stasto, Small  $x$  nonlinear evolution with impact parameter and the structure function data, *Phys. Rev. D* **84**, 094022 (2011).
- [31] H. Mäntysaari and B. Schenke, Confronting impact parameter dependent JIMWLK evolution with HERA data, *Phys. Rev. D* **98**, 034013 (2018).
- [32] A. Aktas *et al.* (H1 Collaboration), Elastic  $J/\psi$  production at HERA, *Eur. Phys. J. C* **46**, 585 (2006).
- [33] H. Kowalski, L. Motyka, and G. Watt, Exclusive diffractive processes at HERA within the dipole picture, *Phys. Rev. D* **74**, 074016 (2006).
- [34] A. Dumitru, H. Mäntysaari, and R. Paatelainen, Impact parameter dependence of color charge correlations in the proton, *SciPost Phys. Proc.* **8**, 058 (2022).
- [35] H. Kowalski and D. Teaney, An impact parameter dipole saturation model, *Phys. Rev. D* **68**, 114005 (2003).
- [36] L. D. McLerran and R. Venugopalan, Computing quark and gluon distribution functions for very large nuclei, *Phys. Rev. D* **49**, 2233 (1994).
- [37] F. D. Aaron *et al.* (H1 and 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.
- [38] Y. V. Kovchegov and A. H. Mueller, Gluon production in current nucleus and nucleon-nucleus collisions in a quasiclassical approximation, *Nucl. Phys.* **B529**, 451 (1998).
- [39] D. Kharzeev, Y. V. Kovchegov, and K. Tuchin, Cronin effect and high  $p_T$  suppression in pA collisions, *Phys. Rev. D* **68**, 094013 (2003).

- [40] B. Ducloué, T. Lappi, and Y. Zhu, Implementation of NLO high energy factorization in single inclusive forward hadron production, *Phys. Rev. D* **95**, 114007 (2017).
- [41] A. M. Stasto, K. J. Golec-Biernat, and J. Kwiecinski, Geometric Scaling for the Total  $\gamma^*p$  Cross Section in the Low  $x$  Region, *Phys. Rev. Lett.* **86**, 596 (2001).
- [42] S. Munier and R. B. Peschanski, Geometric Scaling as Traveling Waves, *Phys. Rev. Lett.* **91**, 232001 (2003).
- [43] S. Munier and R. B. Peschanski, Traveling wave fronts and the transition to saturation, *Phys. Rev. D* **69**, 034008 (2004).
- [44] K. J. Golec-Biernat and M. Wusthoff, Saturation effects in deep inelastic scattering at low  $Q^2$  and its implications on diffraction, *Phys. Rev. D* **59**, 014017 (1998).
- [45] L. N. Lipatov, Reggeization of the vector meson and the vacuum singularity in non-Abelian gauge theories, *Sov. J. Nucl. Phys.* **23**, 338 (1976).
- [46] E. A. Kuraev, L. N. Lipatov, and V. S. Fadin, The Pomeron singularity in non-Abelian gauge theories, *Sov. Phys. JETP* **45**, 199 (1977).
- [47] I. I. Balitsky and L. N. Lipatov, The Pomeron singularity in quantum chromodynamics, *Sov. J. Nucl. Phys.* **28**, 822 (1978).
- [48] H. Abramowicz *et al.* (H1 and ZEUS Collaborations), Combination and QCD analysis of charm and beauty production cross-section measurements in deep inelastic  $ep$  scattering at HERA, *Eur. Phys. J. C* **78**, 473 (2018).
- [49] 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).
- [50] G. Beuf, T. Lappi, and R. Paatelainen, Massive Quarks at One Loop in the Dipole Picture of Deep Inelastic Scattering, *Phys. Rev. Lett.* **129**, 072001 (2022).
- [51] G. Beuf, T. Lappi, and R. Paatelainen, Massive quarks in NLO dipole factorization for DIS: Longitudinal photon, *Phys. Rev. D* **104**, 056032 (2021).
- [52] G. Beuf, T. Lappi, and R. Paatelainen, Massive quarks in NLO dipole factorization for DIS: Transverse photon, *Phys. Rev. D* **106**, 034013 (2022).
- [53] T. Lappi and H. Mäntysaari, Next-to-leading order Balitsky-Kovchegov equation with resummation, *Phys. Rev. D* **93**, 094004 (2016).
- [54] I. Balitsky and G. A. Chirilli, Next-to-leading order evolution of color dipoles, *Phys. Rev. D* **77**, 014019 (2008).
- [55] E. Iancu, J. D. Madrigal, A. H. Mueller, G. Soyez, and D. N. Triantafyllopoulos, Resumming double logarithms in the QCD evolution of color dipoles, *Phys. Lett. B* **744**, 293 (2015).
- [56] B. Ducloué, E. Iancu, A. H. Mueller, G. Soyez, and D. N. Triantafyllopoulos, Non-linear evolution in QCD at high-energy beyond leading order, *J. High Energy Phys.* **04** (2019) 081.
- [57] E. Iancu, J. D. Madrigal, A. H. Mueller, G. Soyez, and D. N. Triantafyllopoulos, Collinearly-improved BK evolution meets the HERA data, *Phys. Lett. B* **750**, 643 (2015).
- [58] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevD.107.114024> for tabulated numerical data for the dipole-proton scattering amplitude  $N(r, x)$ .
- [59] R. Abdul Khalek *et al.*, Science requirements and detector concepts for the electron-ion collider: EIC yellow report, *Nucl. Phys.* **A1026**, 122447 (2022).
- [60] 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).