PHYSICAL REVIEW D **88,** 114020 (2013)

Single inclusive particle production at high energy from HERA data to proton-nucleus collisions

T. Lappi[1,2] and H. Mäntysaari[1]

[1]*Department of Physics, P.O. Box 35, 40014 University of Jyväskylä, Finland*
[2]*Helsinki Institute of Physics, P.O. Box 64, 00014 University of Helsinki, Finland*

(Received 11 October 2013; published 5 December 2013)

We study single inclusive hadron production in proton-proton and proton-nucleus collisions in the CGC framework. The parameters in the calculation are determined solely by standard nuclear geometry and by electron-proton deep inelastic scattering data, which are fit using the running coupling BK equation. We show that it also is possible to obtain a good fit of the HERA inclusive cross section without an anomalous dimension in the initial condition. We argue that one must consistently also use the proton transverse area as measured by a high virtuality probe in DIS for the single inclusive cross section in proton-proton and proton-nucleus collisions. We show that this leads to a midrapidity nuclear modification ratio R_{pA} that approaches unity at high transverse momentum independently of $\sqrt{s}$, in contrast to most CGC calculations in the literature. We also present predictions for future forward R_{pA} measurements at the LHC.

DOI: 10.1103/PhysRevD.88.114020 PACS numbers: 13.85.Ni, 13.85.Hd, 24.85.+p

I. INTRODUCTION

The color glass condensate (CGC) provides a convenient way to describe strongly interacting systems in the high energy limit, where nonlinear phenomena, such as gluon recombination, become important. Because the gluon density scales as $\sim A^{1/3}$, these nonlinearities are enhanced when the target is changed from a proton to a heavy nucleus. The p + Pb run at the LHC provides access to a kinematical region not yet explored and makes it possible to study QCD phenomena in a region where gluon densities are large and nonlinear effects become important.

The structure of a hadron can be studied accurately in deep inelastic scattering (DIS) where a (virtual) photon scatters off the hadron. A large amount of precise high energy electron-proton data measured at HERA have shown that the gluon density inside a proton grows rapidly at small Bjorken x or, equivalently, at high energy. These accurate measurements have also been a crucial test for the CGC, and recent analyses have confirmed that the CGC description is consistent with all the available small-x DIS data [1–3]. The CGC makes it possible to also consistently describe high-energy hadronic interactions other than inclusive deep inelastic scattering within the same unified framework. These include, for example, single [4–6] and double inclusive [7–10] particle production in proton-proton and proton-nucleus collisions, diffractive DIS [11,12] and the initial state for the hydrodynamical modeling of a heavy ion collision [13–15]. First steps beyond the leading order calculations have also been taken recently [16–18]. Comparing calculations fit to DIS data to particle production results for proton-proton and proton-nucleus collisions at different energies provides a nontrivial test of the universality of the CGC description and makes predictions for future LHC pA measurements.

In this work we study to what extent it is possible to compute single inclusive hadron production in proton-proton and proton-nucleus collisions in a consistent CGC framework. As an input we use only the HERA data for the inclusive DIS cross section, which is fitted using the running coupling Balitsky-Kovchegov (BK) equation [19–21]. The resulting initial condition for the dipole cross is extended to nuclei using only the standard Woods-Saxon nuclear density, without any additional nuclear parameters. Unlike in some of the recent works using dipole cross sections measured in DIS to compute particle production in hadronic collisions, we also consistently use the transverse area of the small-x gluonic degrees of freedom in a nucleon (σ_0) determined from the DIS fits. This is to be contrasted with the conceptually separate and numerically much larger ''soft'' area of the nucleus given by the total inelastic nucleon-nucleon cross section σ_{inel}, which is needed in the Glauber modeling of a nuclear collision. Treating the area factor consistently also makes it possible to interpret the ''K factor'' between the normalization of the data and our LO calculation as a real indication of the magnitude of higher order effects and not only as a completely uncontrolled fit parameter. In particular, we will argue that this leads to a much more controlled result for the nuclear modification factor R_{pA} than in the existing literature.

This work is structure as follows. In Sec. II we discuss how one fits the deep inelastic scattering data in order to obtain an initial condition for the BK evolution. In Sec. III we discuss how single inclusive hadron production can be computed in proton-proton collisions, and we generalize the discussion to the proton-nucleus case in Sec. IV. Finally, we show our numerical results in Sec. V before concluding in Sec. VI.

II. ELECTRON-PROTON BASELINE

Deep inelastic scattering provides a precision measurement of proton structure. The H1 and ZEUS collaborations

have measured the proton structure functions F_2 and F_L, and very precise combined results for the reduced cross section σ_r were published recently [22].

The reduced cross section is a function of the proton structure functions:

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

Here $y = Q^2/(sx)$ and $\sqrt{s}$ is the center-of-mass energy. The structure functions are related to the virtual photon-proton cross sections $\sigma_{T,L}^{\gamma^* p}$ for transverse (T) and longitudinal (L) photons:

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

and

$$F_L(x, Q^2) = \frac{Q^2}{4\pi^2 \alpha_{\mathrm{em}}} \sigma_L^{\gamma^* p}. \quad (3)$$

In this work we perform a fit to combined HERA σ_r data [22] for $Q^2 < 50$ GeV^2 and $x < 0.01$.

The virtual photon-proton cross sections in the color glass condensate framework can be computed as

$$\sigma_{T,L}^{\gamma^* p}(x, Q^2) = 2 \sum_f \int \mathrm{d}z \int \mathrm{d}^2\mathbf{b}_T |\Psi_{T,L}^{\gamma^* \to f\bar{f}}|^2 \mathcal{N}(\mathbf{b}_T, \mathbf{r}_T, x), \quad (4)$$

where $\Psi_{T,L}^{\gamma^* \to f\bar{f}}$ is the photon light cone wave function describing how the photon fluctuates to a quark-antiquark pair, computed from light cone QED [23]. The QCD dynamics is inside the function $\mathcal{N}(\mathbf{b}_T, \mathbf{r}_T, x)$, which is the imaginary part of the scattering amplitude for the process where a dipole (quark-antiquark pair with transverse separation $\mathbf{r}_T$) scatters off the color field of a hadron with impact parameter $\mathbf{b}_T$. It cannot be computed perturbatively, but its energy (or equivalently Bjorken x) dependence satisfies the BK equation, for which we use the running coupling corrections derived in Ref. [24]. The explicit factor 2 in Eq. (4) comes from the optical theorem, which tells us that the total cross section is given by twice the imaginary part of the forward scattering amplitude.

We assume here that the impact parameter dependence of the proton factorizes and one make the replacement

$$2 \int \mathrm{d}^2\mathbf{b}_T \to \sigma_0. \quad (5)$$

Note that with the usual convention adopted here, the factor 2 from the optical theorem is absorbed into the constant σ_0; thus, the transverse area of the proton is now $\sigma_0/2$. The proton area $\sigma_0/2$ could, in principle, be obtained from diffractive vector meson production measurements where one can parametrize the transverse momentum dependence of the production cross section as $\sim e^{-B_D \Delta_T^2}$, where Δ_T is the momentum transfer in the process. From the HERA J/Ψ data [25,26] one obtains $B_D \approx 4$ GeV. For a Gaussian impact parameter profile (exponential in t) these are related by $\sigma_0 = 4\pi B_D$. This would correspond to $\sigma_0 \approx 19.5$ mb, which is approximately two-thirds of the value obtained from the fit to the inclusive data. For comparison, a theta-function profile in b with the same t slope at $t = 0$ leads to $\sigma_0 = 8\pi B_D$, which is larger than the value from the inclusive fit. While the observed t distribution of diffractive J/Ψ does not favor the theta function profile, it is not known well enough to determine the profile precisely. We conclude that, in practice, the measured t distribution is insufficient to precisely determine σ_0 alone, and we need to include it as a free parameter in the fit.

As a nonperturbative input one also needs the dipole-proton amplitude at the initial $x = x_0$. For the dipole amplitude we use the following parametrization, based on the McLerran-Venugopalan model [27]:

$$\mathcal{N}(\mathbf{r}_T) = 1 - \exp\left[-\frac{(\mathbf{r}_T^2 Q_{s0}^2)^\gamma}{4} \ln\left(\frac{1}{|\mathbf{r}_T|\Lambda_{\mathrm{QCD}}} + e_c \cdot e \right) \right], \quad (6)$$

where we have generalized the AAMQS [1] form by also allowing the constant inside the logarithm to be different from e. This constant plays the role of an infrared cutoff in the MV model, and its value cannot be fixed by a weak coupling calculation; it is therefore natural to leave it as a free parameter. The other fit parameters are anomalous dimension γ and initial saturation scale Q_{s0}^2.

The BK equation with running coupling requires the strong coupling constant α_s as a function of the transverse separation $r = |\mathbf{r}_T|$. In order to obtain a slow enough evolution to be compatible with the data (see discussion in Ref. [28]) we include, as in [1], an additional fit parameter C^2 such that

$$\alpha_s(r) = \frac{12\pi}{(33 - 2N_f) \log\left(\frac{4C^2}{r^2 \Lambda_{\mathrm{QCD}}^2} \right)}, \quad (7)$$

with Λ_{QCD} fixed to the value 0.241 GeV. It has been argued that this encodes the uncertainty for the scale at which the coordinate space strong coupling constant should be evaluated. On the other hand, it has been justified analytically [29] and confirmed numerically [30] [for a slightly different running coupling prescription and for the Jalilian-Marian-Iancu-McLerran-Weigert-Leonidov-Kovner (JIMWLK) equation] that performing the Fourier transform would lead to $C^2 = e^{-2\gamma_e}$. With this interpretation the fit result $C^2 \sim 6$ corresponds to the QCD scale taking the value $\Lambda_{\mathrm{QCD}}/C \sim 50$ MeV. In numerical solutions we freeze α_s to 0.7 in the infrared. The quark mass is fixed to $m = 0.14$ GeV, as it was found in Ref. [1] that taking it to be a fit parameter does not improve the fit quality

TABLE I. Parameters from fits to HERA reduced cross section data at $x < 10^{-2}$ and $Q^2 < 50$ GeV^2 for different initial conditions. Also, the corresponding initial saturation scales Q_s^2 defined via equation $\mathcal{N}(r^2 = 2/Q_s^2) = 1 - e^{-1/2}$ are shown. The parameters for the MV$^\gamma$ initial condition are obtained by the AAMQS Collaboration [1].

Model	χ^2/d.o.f	Q_{s0}^2 (GeV^2)	Q_s^2 (GeV^2)	γ	C^2	e_c	$\sigma_0/2$ (mb)
MV	2.76	0.104	0.139	1	14.5	1	18.81
MV$^\gamma$	1.17	0.165	0.245	1.135	6.35	1	16.45
MV^e	1.15	0.060	0.238	1	7.2	18.9	16.36

significantly. We also consider only the three light quarks in this work.

The unknown parameters are obtained by performing a fit to small-x DIS data. The first parametrization considered in this work, denoted by MV$^\gamma$, is obtained by setting $e_c \equiv 1$ in Eq. (6) but keeping the anomalous dimension γ as a fit parameter. A global fit to HERA, NMC and E665 deep inelastic scattering data for this parametrization was performed by the AAMQS Collaboration in Ref. [1], resulting in a very good fit, χ^2/d.o.f ≈ 1.17. The fit parameters are listed in Table I.

We fit two other initial conditions for the dipole amplitude to the HERA reduced cross section data. First, we consider a parametrization where we do not include an anomalous dimension ($\gamma \equiv 1$) but let e_c be a free fit parameter. The second model studied for comparison is the MV model without modifications, where $\gamma \equiv 1$ and $e_c \equiv 1$. As we do not include E665 or NMC data, the fit procedure is not exactly the same as in Ref. [1]. However, due to their very small errors the HERA data dominate the AAMQS fit, and this results in only very minor differences between our result and that of Ref. [1].[1] Our fit result for the MV$^\gamma$ model is $Q_{s0}^2 = 0.159$ GeV^2, $\gamma = 1.129$, $C^2 = 7.05$ and $\sigma_0/2 = 16.35$ fm. We have also confirmed numerically that the two different MV$^\gamma$ model parametrizations give basically the same result when computing quantities considered in this work.

Our second parametrization, denoted by MV^e here, has an infrared cutoff e_c as a fit parameter but no anomalous dimension ($\gamma \equiv 1$). The fit quality is essentially as good as for the MV$^\gamma$ model, with the best fit giving χ^2/d.o.f. ≈ 1.15. The fit parameters are listed in Table I. Viewed in momentum space this parametrization provides a smoother interpolation between the small-k saturation region (where it resembles the Gaussian GBW form) and the power law behavior at high k. This is demonstrated in Fig. 1, where we show the Fourier transform of $S(\mathbf{r}_T) = 1 - \mathcal{N}(\mathbf{r}_T)$, which is proportional to the "dipole" gluon distribution.

[1]The other difference between our approach and Ref. [1] is that, for simplicity, we do not use the redefinition of $x \to x(1 + 4m_f^2/Q^2)$, where m_f is the quark mass.

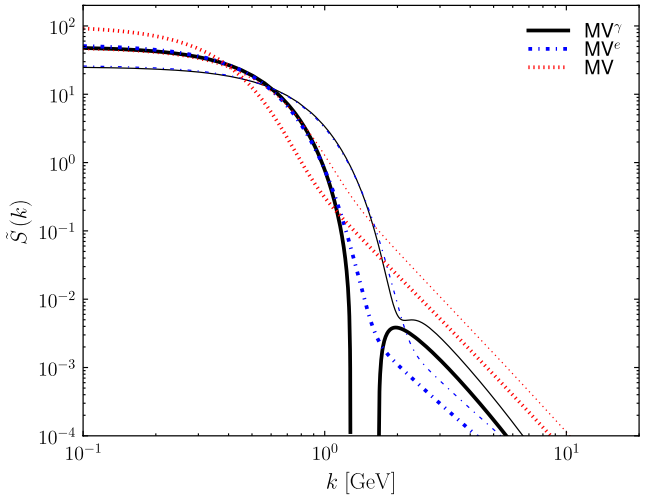

FIG. 1 (color online). Two-dimensional Fourier transform of $S(r) = 1 - \mathcal{N}(r)$ in fundamental (thick lines) and adjoint (thin lines) representations for MV$^\gamma$, MV^e and MV models.

For comparison, we also study a pure MV model initial condition fixing $\gamma = 1$ and $e_c = 1$, resulting in the parameters listed in Table I. The fit quality is not as good as with the modified MV model, (χ^2/d.o.f. ~ 2.8), but as one can see from Fig. 2, the description of the small-x DIS data is still reasonable.

The dipole amplitudes at initial Bjorken x obtained from the MV$^\gamma$ and MV^e models are close to each other, and they both deviate significantly from the pure MV model. To demonstrate this, we show in Fig. 3 the dipole amplitudes $\mathcal{N}(\mathbf{r}_T, x)$ at initial $x = x_0 = 0.01$ and at $x = 10^{-5}$. Note that one cannot directly compare the values of the parameters Q_{s0}^2 in the initial condition, as the functional form in different parametrizations is different. To perform a model-independent comparison of the values of the saturation

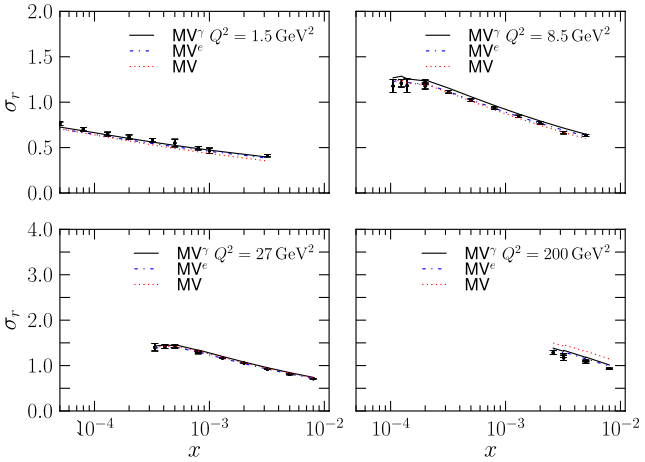

FIG. 2 (color online). Reduced cross section σ_r computed using the MV$^\gamma$, MV^e and MV model initial conditions for the dipole amplitude compared with combined HERA (H1 and ZEUS) data [22].

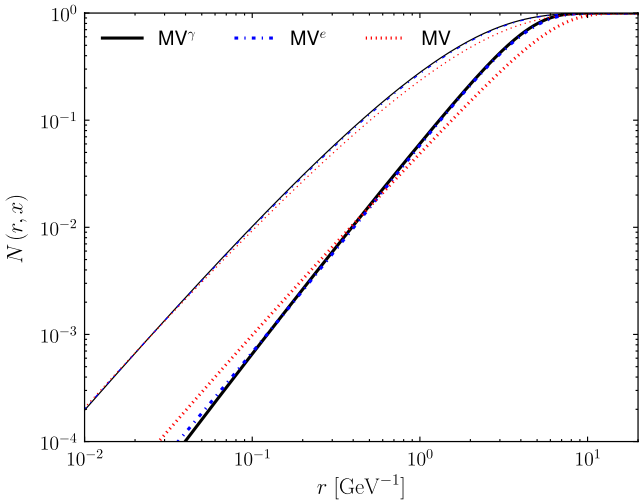

FIG. 3 (color online). Dipole amplitude at initial $x = 10^{-2}$ (thick lines) and after BK evolution at $x = 10^{-5}$ (thin lines) from MV$^\gamma$, MV^e and MV models.

scale in different fits, we define the saturation scale Q_s^2 as a solution to the equation $\mathcal{N}(\mathbf{r}_T^2 = 2/Q_s^2) = 1 - e^{-1/2}$. The initial saturation scales with this model-independent definition are also summarized in Table I.

III. SINGLE INCLUSIVE HADRON PRODUCTION IN CGC

The gluon spectrum in heavy ion collisions can be obtained by solving the classical Yang-Mills equations of motion for the color fields. For $k_T \gtrsim Q_s$ it has been shown numerically [31] that this solution is well approximated by the following k_T-factorized formula [32]:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}y\mathrm{d}^2\mathbf{k}_T\mathrm{d}^2\mathbf{b}_T} = \frac{2\alpha_s}{C_F\mathbf{k}_T^2}\int \mathrm{d}^2\mathbf{q}_T\mathrm{d}^2\mathbf{s}_T \frac{\varphi_p(\mathbf{q}_T,\mathbf{s}_T)}{\mathbf{q}_T^2} \times \frac{\varphi_p(\mathbf{k}_T-\mathbf{q}_T,\mathbf{b}_T-\mathbf{s}_T)}{(\mathbf{k}_T-\mathbf{q}_T)^2}. \quad (8)$$

Here φ_p is the dipole unintegrated gluon distribution (UGD) of the proton [33–35] and $\mathbf{b}_T$ is the impact parameter. Note that, unlike for DIS, there is no formal proof for the k_T-factorization formula in hadron-hadron collisions. For the proton we assume that the impact parameter dependence factorizes and

$$\varphi_p(\mathbf{k}_T) = \int \mathrm{d}^2\mathbf{b}_T\varphi_p(\mathbf{k}_T,\mathbf{b}_T) = \frac{C_F\sigma_0/2}{8\pi^3\alpha_s}\mathbf{k}_T^4\tilde{S}^p(\mathbf{k}_T). \quad (9)$$

Here $\tilde{S}^p(k)$ is the two-dimensional Fourier transform of the dipole-proton scattering matrix $S^p(r) = 1 - \mathcal{N}_A^p(r)$, where $\mathcal{N}_A^p$ is the dipole-proton scattering amplitude in the adjoint representation: $\mathcal{N}_A = 2\mathcal{N} - \mathcal{N}^2$. For the proton DIS area $\sigma_0/2$ we use the value from the fits to DIS data; see Sec. II.

Let us now consider a proton-proton collision. The cross section is obtained by integrating Eq. (8) over the impact parameter, which gives

$$\frac{\mathrm{d}\sigma}{\mathrm{d}y\mathrm{d}^2\mathbf{k}_T} = \frac{(\sigma_0/2)^2}{(2\pi)^2}\frac{C_F}{2\pi^2\mathbf{k}_T^2\alpha_s}\int\frac{\mathrm{d}^2\mathbf{q}_T}{(2\pi)^2}\mathbf{q}_T^2\tilde{S}^p(\mathbf{q}_T) \times(\mathbf{k}_T-\mathbf{q}_T)^2\tilde{S}^p(\mathbf{k}_T-\mathbf{q}_T). \quad (10)$$

The invariant yield is defined as the production cross section divided by the total inelastic cross section σ_{inel} and thus becomes

$$\frac{\mathrm{d}N}{\mathrm{d}y\mathrm{d}^2\mathbf{k}_T} = \frac{(\sigma_0/2)^2}{\sigma_{\mathrm{inel}}}\frac{C_F}{8\pi^4\mathbf{k}_T^2\alpha_s}\int\frac{\mathrm{d}^2\mathbf{q}_T}{(2\pi)^2}\mathbf{q}_T^2\tilde{S}^p(\mathbf{q}_T) \times(\mathbf{k}_T-\mathbf{q}_T)^2\tilde{S}^p(\mathbf{k}_T-\mathbf{q}_T). \quad (11)$$

Assuming that $|\mathbf{k}_T|$ is much larger than the saturation scale of one of the protons, we obtain the hybrid formalism result

$$\frac{\mathrm{d}N}{\mathrm{d}y\mathrm{d}^2\mathbf{k}_T} = \frac{\sigma_0/2}{\sigma_{\mathrm{inel}}}\frac{1}{(2\pi)^2}xg(x,\mathbf{k}_T^2)\tilde{S}^p(\mathbf{k}_T), \quad (12)$$

where

$$xg(x,\mathbf{k}_T^2) = \int_0^{\mathbf{k}_T^2}\frac{\mathrm{d}\mathbf{q}_T^2}{\mathbf{q}_T^2}\varphi_p(\mathbf{q}_T) \quad (13)$$

is the integrated gluon distribution function. This can then be replaced by the conventional parton distribution function, for which we can use the CTEQ LO [36] pdf. In Fig. 4 we show this function as resulting from the dipole fits to HERA data at $x = 10^{-2}$ and $x = 10^{-4}$ compared with the leading order CTEQ distribution. At the initial $x = 10^{-2}$ the gluon density grows much faster as a function of Q^2

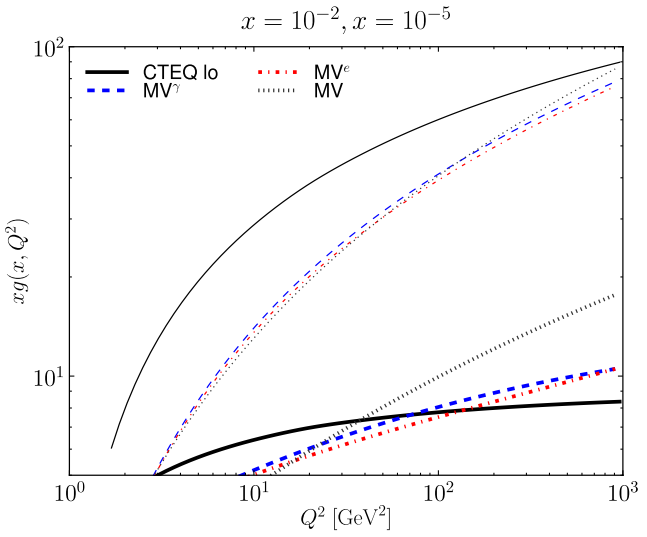

FIG. 4 (color online). Gluon distribution function at $x = 10^{-2}$ (lower thick lines) and at $x = 10^{-4}$ (upper thin lines) computed using MV, MV$^\gamma$ and MV^e initial conditions compared with the leading order CTEQ gluon distribution.

when the MV model is used, and the results obtained using the MV$^\gamma$ and MV^e models are close to each other. This observation suggests a new interpretation for why the experimental data seem to favor a steeper initial condition for BK evolution such as MV$^\gamma$ and MV^e: this form provides a better parametrization of the large logarithms of Q^2 that are resummed by DGLAP evolution but not included in the leading order BK equation. When moving from the initial condition to smaller x the gluon density from the CTEQ distribution grows faster than what is obtained from the BK evolution. At smaller x the difference between the initial conditions is relatively small as the solutions of the BK equation approach the universal form.

HERA measurements of diffractive vector meson electroproduction [37] indicate that the proton transverse area measured with a high virtuality probe is smaller than in soft interactions. In our case this shows up as a large difference in the numerical values of $\sigma_0/2$ and $\sigma_{\rm inel}$ and leads to an energy-dependent factor $\frac{\sigma_0/2}{\sigma_{\rm inel}} \sim 0.2 \ldots 0.3$ in the particle yield (12), in contrast with the treatment often used in CGC calculations. Physically this corresponds to a two-component picture of the transverse structure of the nucleon (see also Ref. [38] for a very similar discussion). The small-x gluons responsible for semihard particle production occupy a small area $\sim\sigma_0/2$ in the core of the nucleon. This core is surrounded by a nonperturbative edge that becomes larger with $\sqrt{s}$ but only participates in soft interactions that contribute to the large total inelastic cross section $\sigma_{\rm inel}$ (we use $\sigma_{\rm inel} = 42$ mb at RHIC [39], $\sigma_{\rm inel} = 44.4$ mb at Tevatron [40] and $\sigma_{\rm inel} = 70$ mb at LHC energies [41]). This description of the transverse profile in terms of only two numbers, an energy-independent σ_0 and an energy-dependent $\sigma_{\rm inel}$, is of course very simplistic, but we believe it captures a physical feature that has been neglected in many works on the subject. Note that Ref. [42] models the same physics by consistently using a b-dependent dipole cross section and incorporating the soft physics as an $\sqrt{s}$-dependent upper limit in the integration over b. In Sec. IV we will show that this separation between the two transverse areas brings much clarity to the extension of the calculation from protons to nuclei.

Now that the normalization ($\sigma_0/2$) from HERA data is also used in the calculation of the single inclusive spectrum, the result also represents the actual LO CGC prediction for the normalization of the spectrum. As is often the case for perturbative QCD, the LO result only agrees with data within a factor of $\sim$2. We therefore multiply the resulting spectrum with a "K factor" to bring it to the level of the experimental data. Now that the different areas σ_0 and $\sigma_{\rm inel}$ are properly included, this factor has a more conventional interpretation of the expected effect of NLO corrections on the result, although it depends quite strongly on the fragmentation function. In particular, we note that the numerical values cannot be directly compared with those given in Ref. [5].

In order to obtain a hadron spectrum from the parton spectrum we calculate convolution with the DSS LO fragmentation function [43] and, when using the hybrid formalism, also add the light quark-initiated channel to the gluonic one in Eq. (12). The momentum scale for the parton distribution functions, fragmentation functions and the strong coupling constant α_s is chosen as the transverse momentum of the produced hadron.

IV. FROM PROTON TO NUCLEUS

Due to a lack of small-x nuclear DIS data we cannot perform a similar fit to nuclear targets as what is done with the proton. Instead we use the optical Glauber model to generalize our dipole-proton amplitude to dipole-nucleus scattering.

First we observe that the total dipole (size r)-proton cross section reads

$$\sigma^p_{\rm dip} = \sigma_0 \mathcal{N}^p(r). \tag{14}$$

In the dilute limit of very small dipoles the dipole-nucleus cross section should be just an incoherent sum of dipole-nucleon cross sections, i.e. $\sigma^A_{\rm dip} = A\sigma^p_{\rm dip}$. On the other hand, for large dipoles we should have $\mathrm{d}\sigma^A_{\rm dip}/\mathrm{d}^2\mathbf{b}_T \equiv 2\mathcal{N}^A(\mathbf{r}_T, \mathbf{b}_T) \leq 2$. These requirements are satisfied with an exponentiated dipole-nucleus scattering amplitude

$$\mathcal{N}^A(\mathbf{r}_T, \mathbf{b}_T) = \left[1 - \exp\left(-\frac{AT_A(\mathbf{b}_T)}{2}\sigma^p_{\rm dip}\right)\right]. \tag{15}$$

This form is an average of the dipole cross section over the fluctuating positions of the nucleons in the nucleus (see e.g. [44]) and thus incorporates, in an analytical expression, the fluctuations discussed e.g. in Ref. [5].

Using the form (15) directly in computing particle production is, however, problematic. Because the forward S-matrix element $S = 1 - \mathcal{N}^A(\mathbf{r}_T, \mathbf{b}_T)$ approaches a limiting value $\exp\left(-\frac{AT_A(\mathbf{b}_T)}{2}\sigma_0\right) \sim \exp\left(-A^{1/3}\right)$ and not exactly zero at large r, the dipole gluon distribution develops unphysical oscillations as a function of k. We therefore expand the proton-dipole cross section in Eq. (15) and use the approximation

$$\begin{aligned}\sigma^p_{\rm dip} &= \sigma_0 \mathcal{N}^p(\mathbf{r}_T) \\ &\approx \sigma_0 \frac{(\mathbf{r}_T{}^2 Q_{s0}^2)^\gamma}{4} \ln\left(\frac{1}{|\mathbf{r}_T|\Lambda_{\rm QCD}} + e_c \cdot e\right)\end{aligned} \tag{16}$$

in the exponent of Eq. (15). The dipole-nucleus amplitude is then obtained by solving the rcBK evolution equation with an initial condition

$$\begin{aligned}\mathcal{N}^A(\mathbf{r}_T, \mathbf{b}_T) = 1 - \exp\Bigg[&-AT_A(\mathbf{b}_T)\frac{\sigma_0}{2}\frac{(\mathbf{r}_T{}^2 Q_{s0}^2)^\gamma}{4} \\ &\times \ln\left(\frac{1}{|\mathbf{r}_T|\Lambda_{\rm QCD}} + e_c \cdot e\right)\Bigg].\end{aligned} \tag{17}$$

We emphasize that besides the Woods-Saxon nuclear density $T_A(\mathbf{b}_T)$, all the parameters in this expression result from the fit to HERA data. Among recent works on the subject, this can be contrasted with e.g. Ref. [5], where the area corresponding to $\sigma_0/2$ in Eq. (17) is set by fiat to 42 mb, or with Ref. [2], where initial saturation scale is varied within a large range. The "optical Glauber" initial condition (17) also brings into evidence the advantage of the MV^e parametrization, which achieves a good fit to HERA data while imposing $\gamma = 1$. In contrast to the MV$^\gamma$ fit, this functional form avoids the ambiguity encountered in e.g. [5] of whether the factor $AT_A(\mathbf{b}_T)\sigma_0/2$ should be replaced by $(AT_A(\mathbf{b}_T)\sigma_0/2)^\gamma$ to achieve a natural scaling of Q_s^2 with the nuclear thickness.

The fully impact-parameter-dependent BK equation develops unphysical Coulomb tails which would need an additional screening mechanism at the confinement scale (see e.g. [45–48]). We therefore solve the scattering amplitudes for each $\mathbf{b}_T$ independently. Due to the rapid increase of the scattering amplitude at low densities (large $|\mathbf{b}_T|$), this effectively causes the nucleus to grow rapidly on the edges at large energies. To demonstrate this we plot in Fig. 5 the saturation scale Q_s^2 of the nucleus as a function of impact parameter $\mathbf{b}_T$, using again the model-independent definition of Q_s as the solution of $\mathcal{N}(\mathbf{r}_T^2 = 2/Q_s^2, \mathbf{b}_T) = 1 - e^{-1/2}$. The saturation scale of the lead nucleus falls below the proton saturation scale at $|\mathbf{b}_T| \gtrsim 6.3$ fm, which corresponds to centrality $\gtrsim 70\%$; see Table III. Due to the unphysical increase of the gluon density at very large $|\mathbf{b}_T|$, we do not consider this parametrization to be reliable in that region. Instead, for calculating minimum bias observables we simply scale up from proton-proton collisions by assuming that $R_{pA} = 1$ for very large impact parameters

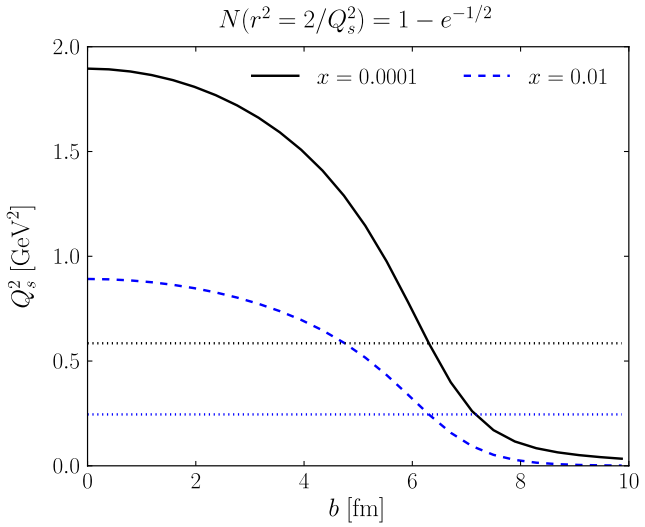

FIG. 5 (color online). Saturation scale of the lead nucleus at $x = 10^{-2}$ and at $x = 10^{-4}$ as a function of the impact parameter obtained using the MV$^\gamma$ initial condition. The dashed lines show the saturation scale of the proton at the same values of x.

(see below). A more refined treatment of the nuclear edge would be possible by replacing the optical Glauber framework by a Monte Carlo one along the lines of [5]. However, the edge region gives a very small contribution to the total inclusive cross sections, and this would have a negligible effect on the observables considered in this paper.

When computing proton-nucleus cross sections we compute convolutions of the nuclear and proton unintegrated gluon distributions. In terms of the dipole amplitudes, the k_T-factorization formula now reads

$$\frac{\mathrm{d}N(\mathbf{b}_T)}{\mathrm{d}y\mathrm{d}^2\mathbf{k}_T} = \frac{\sigma_0/2}{(2\pi)^2}\frac{C_\mathrm{F}}{2\pi^2\mathbf{k}_T^2\alpha_\mathrm{s}}\int\frac{\mathrm{d}^2\mathbf{q}_T}{(2\pi)^2}\mathbf{q}_T^2\tilde{S}^p(\mathbf{q}_T) \times(\mathbf{k}_T-\mathbf{q}_T)^2\tilde{S}^A(\mathbf{k}_T-\mathbf{q}_T), \tag{18}$$

where $\tilde{S}^p$ and $\tilde{S}^A$ are Fourier transforms of the dipole-proton and dipole-nucleus scattering matrices, respectively. Assuming, moreover. that the transverse momentum of the produced parton is much larger than the proton saturation scale, we get the hybrid formalism result

$$\frac{\mathrm{d}N(\mathbf{b}_T)}{\mathrm{d}y\mathrm{d}^2\mathbf{k}_T} = \frac{1}{(2\pi)^2}xg(x,\mathbf{k}_T^2)\tilde{S}^A(\mathbf{k}_T). \tag{19}$$

Notice that, in contrast to Eq. (12), in this case we do not get a factor $(\sigma_0/2)/\sigma_\mathrm{inel}$ in the yield.

Let us then show that with this parametrization we get $R_{pA} \to 1$ at large transverse momenta. First we observe that at large $|\mathbf{k}_T|$ (when x also approaches the initial condition value) the particle yield in proton-nucleus collisions is

$$\mathrm{d}N^{pA} \sim xg\tilde{S}^A(\mathbf{k}_T) \sim xg\int\mathrm{d}^2\mathbf{r}_T e^{i\mathbf{k}_T\cdot\mathbf{r}_T}\exp\left(-\frac{AT_A(b)}{2}\sigma_\mathrm{dip}^p\right) \sim xgAT_A(b)\frac{\sigma_0}{2}\mathcal{N}^p. \tag{20}$$

On the other hand, in proton-proton collisions we get

$$\mathrm{d}N^{pp} \sim \frac{\sigma_0/2}{\sigma_\mathrm{inel}}xg\mathcal{N}^p. \tag{21}$$

Now, as $N_\mathrm{bin} = AT_A\sigma_\mathrm{inel}$, the nuclear modification ratio is

$$R_{pA} = \frac{\mathrm{d}N^{pA}}{N_\mathrm{bin}\mathrm{d}N^{pp}} \to 1 \tag{22}$$

at all $\sqrt{s}$, even as σ_inel and thus N_bin are changing with $\sqrt{s}$ while the initial saturation scale Q_{s0} is not. This is in marked contrast to e.g. Refs. [4,5,49], where the physics of high energy evolution is basically the same, but the treatment of the transverse geometry is different, resulting in a variety of very different predictions for the high transverse momentum behavior of R_{pA}.

Once the dipole-nucleus amplitude is known, one immediately gets the unintegrated gluon distribution of the nucleus at fixed impact parameter $\mathbf{b}_T$ from Eq. (9):

$$\varphi^A(\mathbf{k}_T, \mathbf{b}_T) = \frac{C_F}{8\pi^3\alpha_s}\mathbf{k}_T{}^4\tilde{S}^A(\mathbf{k}_T, \mathbf{b}_T). \quad (23)$$

In order to obtain estimates for the dipole-nucleus scattering amplitude at large impact parameters, we assume that in the region where the saturation scale of the nucleus would fall below the corresponding scale of the proton, at $|\mathbf{b}_T| > b_0$, we obtain the differential particle production yield in proton-nucleus collisions by using the expanded form Eq. (20) at all r. This expansion is justified because at large impact parameters the combination $AT_A(b)$ is small. This gives, substituting $AT_A = N_{\rm bin}/\sigma_{\rm inel}$,

$$dN^{pA} \sim xgN_{\rm bin}\frac{\sigma_0/2}{\sigma_{\rm inel}}N^p = N_{\rm bin}dN^{pp}. \quad (24)$$

Notice that this parametrization is equivalent to imposing $R_{pA} = 1$ at large impact parameters.

In proton-nucleus collisions it is not possible to determine the impact parameter by measuring the total multiplicity as well as in heavy ion collisions, due to the large multiplicity fluctuations for a fixed impact parameter. The first LHC proton-lead results are divided into centrality classes based on multiplicity or energy deposit in forward calorimeters. This is a theoretically difficult quantity to handle, so we assume that we can obtain reasonable estimates for different centrality classes by using a standard optical Glauber model, described briefly in Appendix A.

V. RESULTS

In Fig. 6 we show the single inclusive π^0 and negative hadron yields computed using the hybrid formalism at $\sqrt{s} = 200$ GeV and compare them with the experimental data from RHIC [39,50,51]. As initial conditions for the

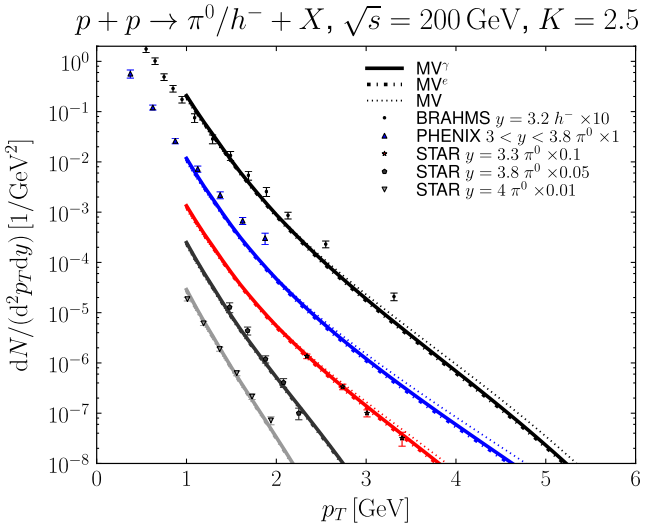

FIG. 6 (color online). Single inclusive π^0 and negative hadron production computed using MV, MV^e and MV$^\gamma$ initial conditions compared with RHIC data from STAR [39], PHENIX [50] and BRAHMS [51] collaborations.

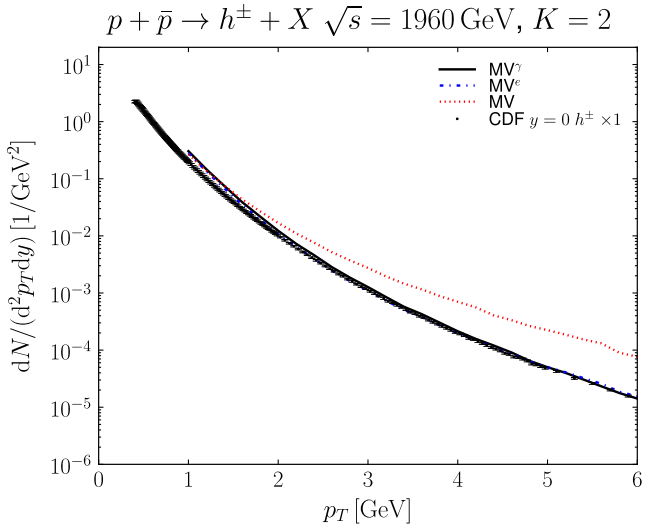

FIG. 7 (color online). Single inclusive charged hadron production computed using MV, MV^e and MV$^\gamma$ initial conditions compared with CDF data [40].

BK evolution we use MV, MV$^\gamma$ and MV^e fits. We recall that all fits, especially MV$^\gamma$ and MV^e, give good descriptions of the HERA DIS data (see Fig. 2). We observe that all initial conditions yield very similar particle spectra, and in particular, the STAR π^0 spectra work very well, using $K = 2.5$. The agreement with BRAHMS and PHENIX data is still reasonably good, even though the p_T slope is not exactly correct.

Figure 7 shows a comparison with the CDF charged hadron data [40] at $\sqrt{s} = 1960$ GeV computed using k_T factorization. The Tevatron data seem to require a K factor $\sim$2, and our calculation slightly overestimates the yield at small p_T. The standard MV model does not give a reasonable description of the data anymore, whereas the MV$^\gamma$ and MV^e models are in good agreement, still giving a slightly wrong slope at small p_T.

The single inclusive π^0 and charged hadron yields at $\sqrt{s} = 7000$ GeV computed using k_T factorization and compared with ALICE [52] and CMS [53] data are shown in Fig. 8. Both MV$^\gamma$ and MV^e models describe the data well without any additional K factor.[2] The pure MV model gives a spectrum that is too hard, similar to the Tevatron data.

In order to study the sensitivity to different ingredients, we also compute the neutral pion spectrum at LHC energies using both k_T factorization and the hybrid formalism using the MV^e initial condition. The results are shown in Fig. 9. With the hybrid formalism we use both CTEQ parton distribution functions and a gluon distribution

[2]As the Tevatron data seem to require a K factor $\sim$2 and no such factor is needed for the LHC, it seems that the required K factor decreases as a function of energy. A similar result for LO pQCD calculations was found in Ref. [54].

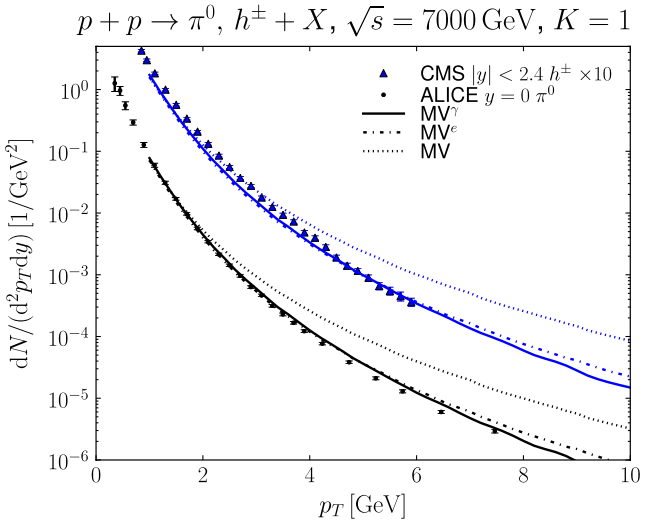

FIG. 8 (color online). Single inclusive π^0 production computed using MV, MV$^\gamma$ and MV^e initial conditions at $\sqrt{s} =$ 7000 GeV compared with ALICE π^0 [52] and CMS charged hadron data [53]. The CMS yield is computed at $y = 0$.

obtained by integrating the unintegrated gluon distribution; see Eq. (13). We also compare the DSS and KKP fragmentation functions (we consistently use only LO distributions in this work). The results are scaled by a K factor which is chosen to fit the data around $\mathbf{p}_T \approx 2$ GeV. We observe that, apart from the different overall normalization, the different model combinations give very similar spectra, the hybrid formalism with unintegrated gluon distribution PDFs deviating slightly from the data at large $|\mathbf{p}_T|$. Notice

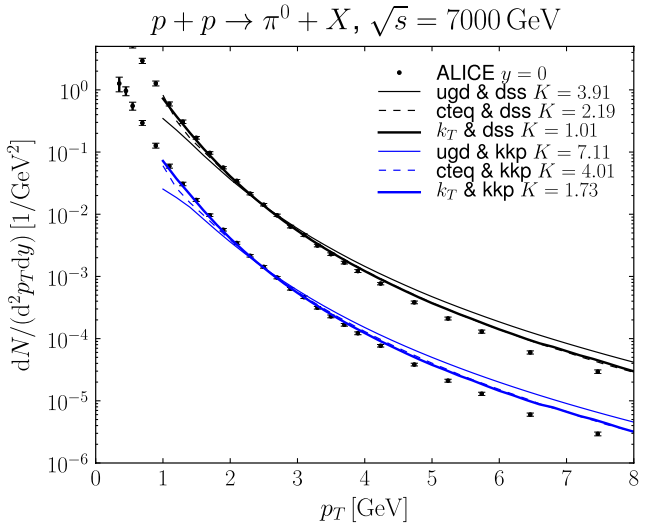

FIG. 9 (color online). Single inclusive π^0 production at $\sqrt{s} =$ 7000 GeV compared with ALICE π^0 [52] data computed using k_T-factorization and hybrid formalisms with CTEQ and UGD parton distribution functions and DSS (upper curves, multiplied by 10) and KKP (lower curves) fragmentation functions. The initial condition for the BK evolution is MV^e.

that when using the UGD parton distribution function we can only compute the gluon channel, as we have no straightforward way to obtain the quark distribution from the gluon density. Similar conclusions are obtained when the analysis is performed with NLO PDFs and FFs, when one generally needs slightly larger K factors and the $\mathbf{p}_T$ slope obtained using the hybrid formalism is slightly worse.

We can conclude from Fig. 9 that the absolute normalization depends strongly on the choice of hybrid vs k_T-factorized formalisms and also the fragmentation function set used. The $\mathbf{p}_T$ slope, on the other hand, is a more solid prediction of the BK evolution. This is easily understood since the $\mathbf{p}_T$ is directly related to the probed Bjorken x of the target, as (neglecting fragmentation) $x \sim p_T e^{-y}/\sqrt{s}$; thus, the $\mathbf{p}_T$ dependence is given by the BK equation.

In order to study the differences between the k_T factorization and the hybrid formalisms, we plot in Fig. 10 the parton-level gluon production yield at rapidities $y = 0, 1, 2$ and 3 at LHC energies $\sqrt{s} = 7000$ GeV computed using the hybrid formalism (with both CTEQ and UGD parton distribution functions) and normalized by the corresponding yield obtained using k_T factorization. We observe that at more forward rapidities, where the saturation scale of one of the protons is relatively small, the hybrid formalism and the k_T factorization are relatively close to each other. This is especially clear when the gluon distribution is computed from the same dipole amplitude (called a UGD parton distribution function). We conclude that as the

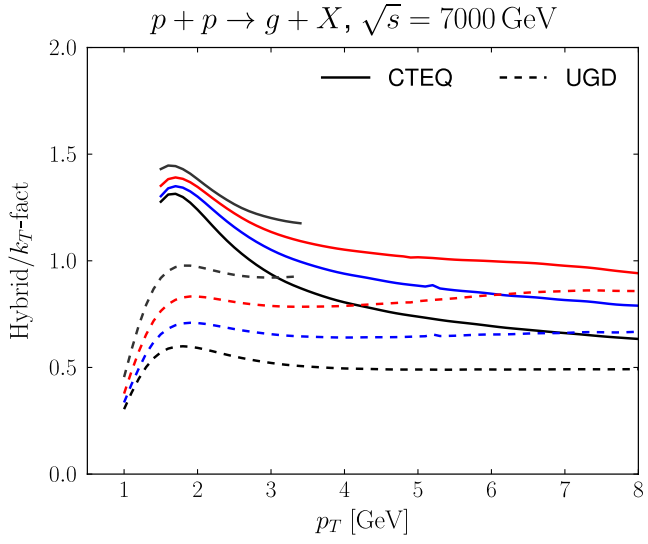

FIG. 10 (color online). Single inclusive gluon production spectrum at $\sqrt{s} = 7000$ GeV obtained by using the hybrid formalism and CTEQ (solid lines) or UGD (dashed lines) parton gluon distribution functions normalized by the corresponding spectrum obtained by using the k_T factorization. The rapidities are, from bottom to top, $y = 0$, 1, 2, 3. The results are shown in the kinematical region where $x < 10^{-2}$.

k_T-factorization and hybrid formalisms are relatively close to each other around $y \sim 2$, and as the k_T factorization cannot be used in kinematics when the x of one of the protons is large, it is reasonable to switch to the hybrid formalism around $y \sim 2$.

Let us now discuss proton-nucleus and deuteron-nucleus collisions. In Fig. 11 we present the single inclusive π^0 and negative charged hadron production yields in minimum bias deuteron-nucleus collisions at forward rapidities computed using the hybrid formalism with different initial conditions for the BK evolution and compared with the RHIC data [39,51,55]. We use here the same K factor, $K = 2.5$. that was required to obtain the correct normalization with the RHIC pp data. The p_T slopes agree roughly with the data, the agreement being very good with the STAR data. The absolute normalization now works relatively well with the BRAHMS and PHENIX data (which were underestimated in the proton-proton case), whereas the STAR yield is overestimated by a factor of $\sim$2.

In Fig. 12 we show the single inclusive charged hadron yield in proton-lead collisions compared with the ALICE data [56]. The conclusion is very similar to that for proton-proton collisions (see Fig. 8): the pure MV model gives a completely wrong p_T slope, but both MV$^\gamma$ and MV^e models describe the data well.

In Fig. 13 we compare our result for midrapidity minimum bias R_{pA} with ALICE charged hadron measurements [56]. Even though the p_T spectra obtained using the MV$^\gamma$ and MV^e initial conditions are very different from the pure MV model, all three initial conditions yield a very similar R_{pA}. Recall that we get exactly $R_{pA} \to 1$ at large p_T, consistent with the ALICE result. The centrality dependence of midrapidity R_{pA} is shown in Fig. 14. Here we only

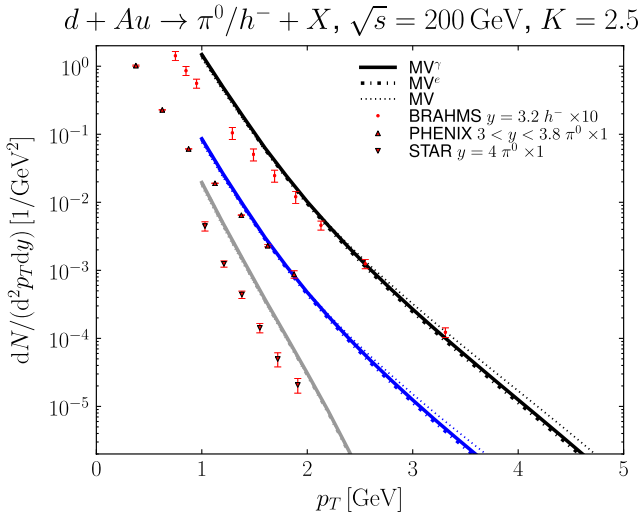

FIG. 11 (color online). Single inclusive π^0 and negative hadron production at $\sqrt{s} = 200$ GeV d + Au collisions compared with BRAHMS [51], STAR [39] and PHENIX [55] data.

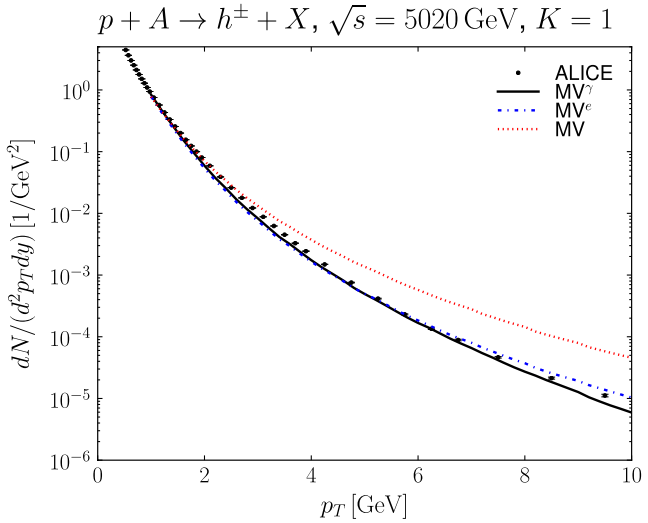

FIG. 12 (color online). Single inclusive charged hadron production in minimum bias p + Pb collisions at $\sqrt{s} =$ 5020 GeV computed using k_T factorization and compared with ALICE data [56].

compute results using the MV^e initial condition, as the results obtained using different initial conditions are basically the same. The centrality dependence is relatively weak; the results start to differ only for the most peripheral classes where centrality $\gtrsim 60\%$. Notice that in our calculation we explicitly set $R_{pA} = 1$ at centralities $\gtrsim 70\%$; see discussion in Sec. IV.

Let us now present predictions for the future R_{pA} measurements. In Fig. 15 we show R_{pA} at $\sqrt{s} = 5020$ GeV in minimum bias collisions at forward rapidities. As can be seen from Fig. 10, at more forward rapidities the hybrid

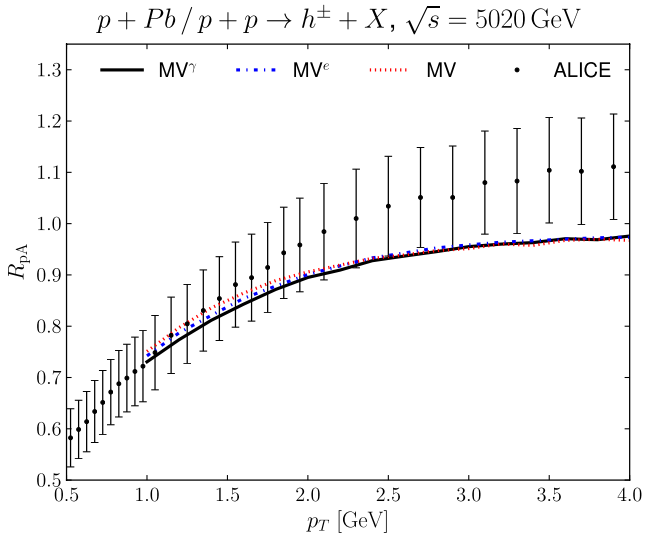

FIG. 13 (color online). Minimum bias nuclear suppression factor $R_{pA}(y = 0)$ at different centrality classes computed using k_T factorization and MV$^\gamma$, MV^e and MV model initial conditions compared with the minimum bias ALICE data [56] at the smallest p_T region.

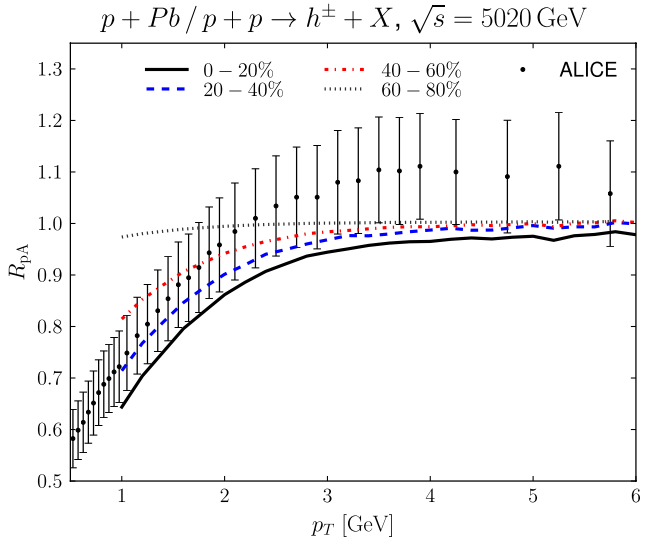

FIG. 14 (color online). Centrality dependence of $R_{pA}(y=0)$ computed using k_T factorization and an MV^e initial condition compared with the ALICE data [56].

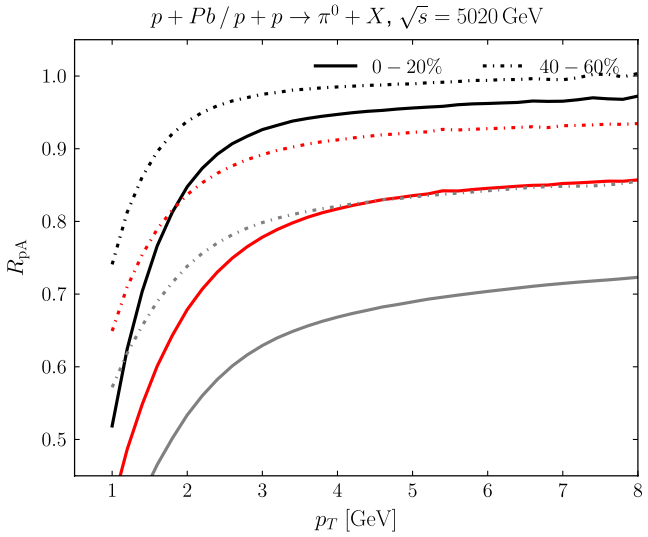

FIG. 16 (color online). Nuclear modification factor at rapidities $y=2$, 4, 6 (from top to bottom) using the MV$^\gamma$ initial condition at centrality classes 0%–20% and 40%–60%.

formalism and the k_T factorization are closer to each other, and we compute the nuclear suppression factors using the hybrid formalism at $y \geq 2$. Note that it is not possible to use k_T factorization in this kinematical region, as one also has a relatively large component from the large-x part of the proton. We compute R_{pA} using the MV, MV$^\gamma$ and MV^e initial conditions, and we see that their difference remains small at all rapidities. The centrality dependence of R_{pA} is shown in Fig. 16. At $y=2$ most central and peripheral collisions give a similar R_{pA}, and the difference between the two centrality classes increases when we move to more forward rapidities.

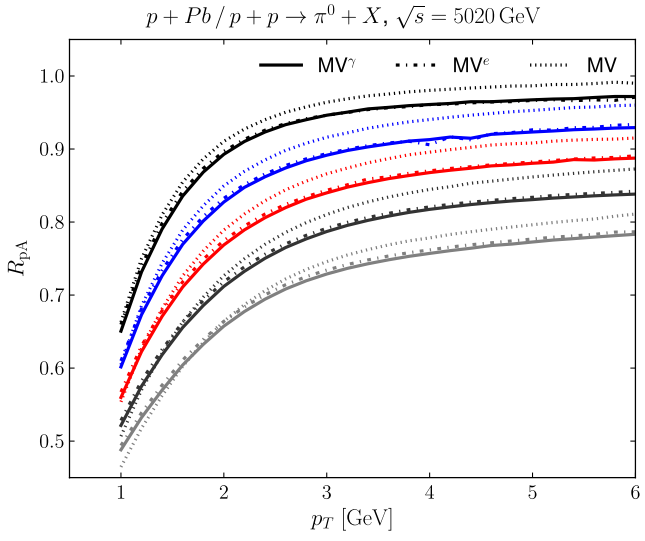

FIG. 15 (color online). Rapidity dependence of the nuclear modification factor at rapidities $y=2$, 3, 4, 5, 6 (from top to bottom) using the MV, MV$^\gamma$ and MV^e initial conditions.

We observe slightly less suppression than obtained in Ref. [5] where the saturation scale of the nucleus is computed in a Monte Carlo Glauber model. We conjecture that this is due to the fact that since $\sigma_0/2 < 42$ mb, the nuclear Q_s is smaller than assumed in Ref. [5]. The evolution speed obtained using the Monte Carlo method is very similar to what is obtained in this work.

In order to further demonstrate the evolution speed of the nuclear modification factor, we plot $R_{pA}(\mathbf{p}_T = 3$ GeV) for neutral pion production at LHC energies in Fig. 17 in

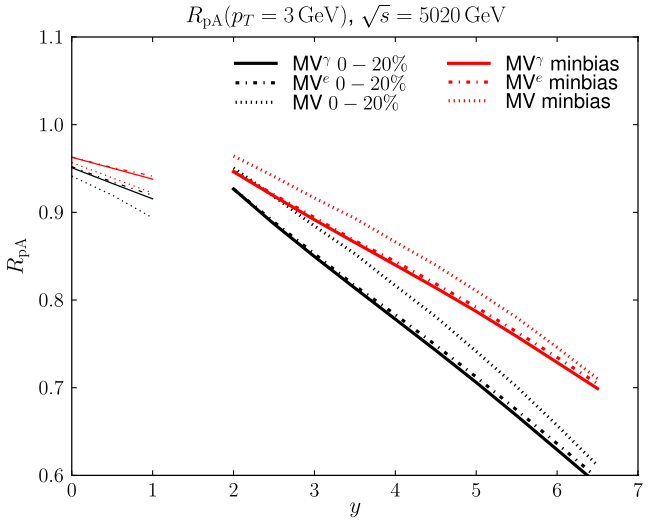

FIG. 17 (color online). Rapidity and centrality dependence of the nuclear modification factor in neutral pion production in 0%–20% most central (solid lines) and minimum bias collisions using MV, MV$^\gamma$ and MV^e initial conditions. Thin lines at $y \leq 1$ are computed using k_T factorization, and thick lines at $y \geq 2$ are computed using the hybrid formalism.

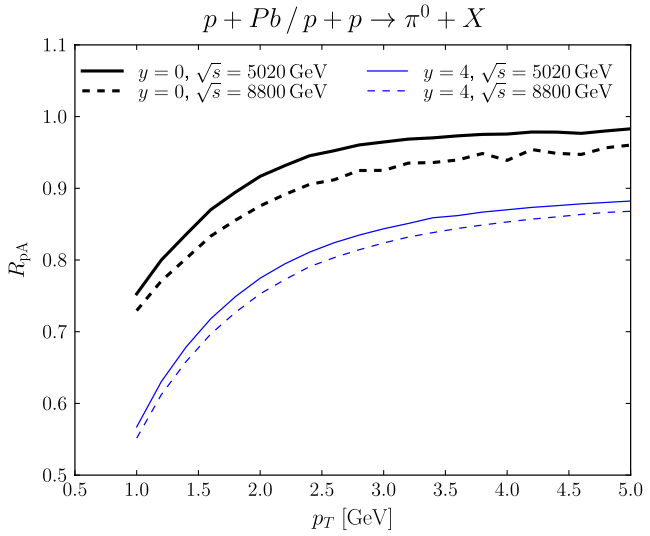

FIG. 18 (color online). Center-of-mass energy ($\sqrt{s}$) dependence of the nuclear modification factor in neutral pion production in minimum-bias p + Pb collisions computed using the MV^e initial condition. The results at midrapidity $y = 0$ are computed using the k_T factorization, and at $y = 4$ the hybrid formalism is used.

most central and minimum bias collisions. We compute R_{pA} close to midrapidity using k_T factorization and at forward rapidities using the hybrid formalism, where we use CTEQ parton distribution functions and also include the quark initiated channel. Thus, the obtained curve is not exactly continuous. The evolution speed close to midrapidity (where k_T factorization should be valid) is slightly slower than at more forward rapidities where the hybrid formalism is more reliable. The MV model initial condition gives a slightly different result than the MV$^\gamma$ and MV^e models, and all dipole models give basically the same evolution speed. Thus, R_{pA} is not sensitive to the details of the initial dipole amplitude, and the evolution speed of R_{pA} is driven by the BK evolution. The centrality and especially the rapidity evolution speed are significantly faster than in a NLO pQCD calculation using the EPS09s nuclear parton distribution functions [57,58].

Finally, we demonstrate the energy dependence of the nuclear modification factor by showing in Fig. 18 R_{pA} at midrapidity and at $y = 4$ in minimum-bias p + Pb collisions. Increasing the energy from the current $\sqrt{s} = 5020$ GeV to the design energy $\sqrt{s} = 8800$ GeV (where we use $\sigma_{\rm inel} = 75$ mb [41]) does not change R_{pA} significantly, and we get midrapidity $R_{pA} \to 1$ at large p_T at all $\sqrt{s}$, as discussed in Sec. IV. The result differs significantly from the corresponding prediction shown in Ref. [4] where a much faster energy evolution was predicted.

VI. CONCLUSIONS

Taking only input from electron-proton deep inelastic scattering and standard nuclear geometry, we compute single inclusive hadron production in proton-proton and proton-nucleus collisions from the color glass condensate framework. We observe that in order to obtain a consistent description of all the single inclusive data, one has to modify the MV model, which is used as an initial condition for the BK evolution. We show that while a modification is required, one does not have to introduce an anomalous dimension $\gamma > 1$, but instead it is enough to take the infrared cutoff in the MV model to be a fit parameter. Using the pure MV model (without an anomalous dimension or modification to the infrared cutoff) one also obtains a reasonably good description of the HERA and RHIC data, but Tevatron and LHC proton-proton data clearly favor models with an anomalous dimension or scaling of the infrared cutoff parameter.

We obtain a good description of the available proton-nucleus and deuteron-nucleus data, the absolute normalizations of the RHIC results being difficult to reproduce simultaneously, as is also the case with the RHIC forward proton-proton data. We obtain exactly $R_{pA} \to 1$ at large $\mathbf{p}_T$ which is a natural requirement and is consistent with the available ALICE data; it follows directly from our consistent treatment of the difference between the proton transverse areas measured in DIS and the inelastic proton-proton cross section. We present predictions for the future forward R_{pA} measurements. In particular, we find that the rapidity evolution of the R_{pA} at fixed $\mathbf{p}_T$ is a solid prediction of the CGC, given by the BK equation.

ACKNOWLEDGMENTS

We thank I. Helenius and L. Korkeala for discussions and M. Chiu for providing us with the PHENIX π^0 yield in proton-proton collisions. This work has been supported by the Academy of Finland, Projects No. 133005, No. 267321, and No. 273464 and by computing resources from CSC – IT Center for Science in Espoo, Finland. H. M. is supported by the Graduate School of Particle and Nuclear Physics.

APPENDIX: OPTICAL GLAUBER

Let us briefly specify the optical Glauber model used here. In a proton-nucleus collision at an impact parameter $\mathbf{b}_T$ the number of binary collisions is given by

$$N_{\rm bin} = AT_A(\mathbf{b}_T)\sigma_{\rm inel} \tag{A1}$$

where $\sigma_{\rm inel}$ is the total inelastic nucleon-nucleon cross section and $T_A(\mathbf{b}_T)$ is the transverse thickness function of the nucleus obtained by integrating the Woods-Saxon distribution

$$\rho_A(\mathbf{b}_T, z) = \frac{n}{1 + \exp\left[\frac{\sqrt{\mathbf{b}_T^2 + z^2} + R_A}{d}\right]} \tag{A2}$$

over the longitudinal distance z. The parameters are $d = 0.54$ fm and $R_A = (1.12A^{1/3} - 0.86A^{-1/3})$ fm. The distribution is normalized to unity, $\int \mathrm{d}^2\mathbf{b}_T \mathrm{d}z \rho_A(\mathbf{b}_T, z) = 1$.

The probability for having an inelastic collision is

$$p(\mathbf{b}_T) \approx 1 - e^{-AT_A(\mathbf{b}_T)\sigma_{\mathrm{inel}}}, \tag{A3}$$

and the total inelastic proton-nucleus cross section is then

$$\sigma_{\mathrm{inel}}^{pA} = \int \mathrm{d}^2\mathbf{b}_T p(\mathbf{b}_T). \tag{A4}$$

A centrality class $(c_1 - c_2)\%$ corresponds to the impact parameter interval $[b_1, b_2]$ for which

$$(c_1 - c_2)\% = \frac{1}{\sigma_{\mathrm{inel}}^{pA}} \int_{b_1}^{b_2} \mathrm{d}^2\mathbf{b}_T p(\mathbf{b}_T). \tag{A5}$$

The $(0 - c)\%$ most central collisions give $c\%$ of the total inelastic proton-nucleus cross section.

The average number of binary collisions in a certain impact parameter class is

$$\langle N_{\mathrm{bin}} \rangle_{b_1,b_2} = \frac{\int_{b_1}^{b_2} \mathrm{d}^2\mathbf{b}_T N_{\mathrm{bin}}(\mathbf{b}_T)}{\int_{b_1}^{b_2} \mathrm{d}^2\mathbf{b}_T p(\mathbf{b}_T)}, \tag{A6}$$

where the denominator is $(c_1 - c_2)\%$ of the total inelastic proton-nucleus cross section $\sigma_{\mathrm{inel}}^{pA}$. The centrality classes and the corresponding values of N_{bin} for RHIC and LHC energies are shown in Tables II and III.

TABLE II. Impact parameter intervals and number of binary collisions in deuteron-gold collisions for different centrality classes at RHIC energies, where $\sqrt{s} = 200$ GeV and $\sigma_{\mathrm{inel}} = 42$ mb. We assume that deuterons consist of two independent nucleons, giving $N_{\mathrm{bin}}^{dA} = 2N_{\mathrm{bin}}^{pA}$.

Bin	b_1 (fm)	b_2 (fm)	$\langle N_{\mathrm{bin}}/2 \rangle$
0%–20%	0.0	3.26	8.45
20%–40%	3.26	4.62	6.95
40%–60%	4.62	5.66	5.03
60%–80%	5.66	6.61	2.89
0%–100%	0		4.95

TABLE III. Impact parameter intervals and number of binary collisions in proton-lead collisions at LHC energies, where $\sqrt{s} = 5020$ GeV and $\sigma_{\mathrm{inel}} = 70$ mb.

Bin	b_1 (fm)	b_2 (fm)	$\langle N_{\mathrm{bin}} \rangle$
0%–20%	0.0	3.47	14.24
20%–40%	3.47	4.91	11.41
40%–60%	4.91	6.01	7.66
60%–80%	6.01	6.99	3.68
0%–100%	0.0		7.69

The particle yield in a centrality class is computed as

$$\frac{\mathrm{d}N}{\mathrm{d}y\mathrm{d}^2\mathbf{k}_T} = \frac{\int \mathrm{d}^2\mathbf{b}_T \frac{\mathrm{d}N(\mathbf{b}_T)}{\mathrm{d}y\mathrm{d}^2\mathbf{k}_T}}{\int \mathrm{d}^2\mathbf{b}_T p(\mathbf{b}_T)}, \tag{A7}$$

where integration limits are set according to the corresponding centrality class.

[1] J. L. Albacete, N. Armesto, J. G. Milhano, P. Quiroga-Arias, and C. A. Salgado, Eur. Phys. J. C **71**, 1705 (2011).
[2] A. H. Rezaeian, M. Siddikov, M. Van de Klundert, and R. Venugopalan, Phys. Rev. D **87**, 034002 (2013).
[3] A. H. Rezaeian and I. Schmidt, Phys. Rev. D **88**, 074016 (2013).
[4] P. Tribedy and R. Venugopalan, Phys. Lett. B **710**, 125 (2012).
[5] J. L. Albacete, A. Dumitru, H. Fujii, and Y. Nara, Nucl. Phys. **A897**, 1 (2013).
[6] A. H. Rezaeian, Phys. Lett. B **718**, 1058 (2013).
[7] T. Lappi and H. Mäntysaari, Nucl. Phys. **A908**, 51 (2013).
[8] J. L. Albacete and C. Marquet, Phys. Rev. Lett. **105**, 162301 (2010).
[9] A. Stasto, B.-W. Xiao, and D. Zaslavsky, Phys. Rev. D **86**, 014009 (2012).
[10] J. Jalilian-Marian and A. H. Rezaeian, Phys. Rev. D **86**, 034016 (2012).
[11] H. Kowalski, L. Motyka, and G. Watt, Phys. Rev. D **74**, 074016 (2006).
[12] T. Lappi and H. Mäntysaari, Phys. Rev. C **87**, 032201 (2013).
[13] T. Lappi, Phys. Lett. B **703**, 325 (2011).
[14] B. Schenke, P. Tribedy, and R. Venugopalan, Phys. Rev. Lett. **108**, 252301 (2012).
[15] C. Gale, S. Jeon, B. Schenke, P. Tribedy, and R. Venugopalan, Phys. Rev. Lett. **110**, 012302 (2013).
[16] I. Balitsky and G. A. Chirilli, Phys. Rev. D **77**, 014019 (2008).
[17] G. A. Chirilli, B.-W. Xiao, and F. Yuan, Phys. Rev. D **86**, 054005 (2012).
[18] A. M. Stasto, B.-W. Xiao, and D. Zaslavsky, arXiv:1307.4057.
[19] I. Balitsky, Nucl. Phys. **B463**, 99 (1996).
[20] Y. V. Kovchegov, Phys. Rev. D **60**, 034008 (1999).
[21] Y. V. Kovchegov, Phys. Rev. D **61**, 074018 (2000).
[22] F. Aaron *et al.* (H1 and ZEUS Collaborations), J. High Energy Phys. 01 (2010) 109.

[23] Y. V. Kovchegov and E. Levin, *Quantum Chromodynamics at High Energy* (Cambridge University Press, Cambridge, England, 2012).
[24] I. Balitsky, Phys. Rev. D **75**, 014001 (2007).
[25] S. Chekanov *et al.* (ZEUS Collaboration), Nucl. Phys. **B695**, 3 (2004).
[26] A. Aktas *et al.* (H1 Collaboration), Eur. Phys. J. C **46**, 585 (2006).
[27] L. D. McLerran and R. Venugopalan, Phys. Rev. D **49**, 2233 (1994).
[28] J. Kuokkanen, K. Rummukainen, and H. Weigert, Nucl. Phys. **A875**, 29 (2012).
[29] Y. V. Kovchegov and H. Weigert, Nucl. Phys. **A784**, 188 (2007).
[30] T. Lappi and H. Mäntysaari, Eur. Phys. J. C **73**, 2307 (2013).
[31] J. P. Blaizot, T. Lappi, and Y. Mehtar-Tani, Nucl. Phys. **A846**, 63 (2010).
[32] Y. V. Kovchegov and K. Tuchin, Phys. Rev. D **65**, 074026 (2002).
[33] D. Kharzeev, Y. V. Kovchegov, and K. Tuchin, Phys. Rev. D **68**, 094013 (2003).
[34] J. P. Blaizot, F. Gelis, and R. Venugopalan, Nucl. Phys. **A743**, 13 (2004).
[35] F. Dominguez, C. Marquet, B.-W. Xiao, and F. Yuan, Phys. Rev. D **83**, 105005 (2011).
[36] J. Pumplin, D. R. Stump, J. Huston, H.-L. Lai, P. Nadolsky, and W.-K. Tung, J. High Energy Phys. 07 (2002) 012.
[37] F. Aaron *et al.* (H1 Collaboration), J. High Energy Phys. 05 (2010) 032.
[38] L. Frankfurt, M. Strikman, and C. Weiss, Phys. Rev. D **83**, 054012 (2011).
[39] J. Adams *et al.* (STAR Collaboration), Phys. Rev. Lett. **97**, 152302 (2006).
[40] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. D **79**, 112005 (2009).
[41] G. Antchev, P. Aspell, I. Atanassov, V. Avati, J. Baechler *et al.* (TOTEM Collaboration), Europhys. Lett. **96**, 21002 (2011).
[42] P. Tribedy and R. Venugopalan, Nucl. Phys. **A850**, 136 (2011).
[43] D. de Florian, R. Sassot, and M. Stratmann, Phys. Rev. D **75**, 114010 (2007).
[44] H. Kowalski, T. Lappi, and R. Venugopalan, Phys. Rev. Lett. **100**, 022303 (2008).
[45] K. J. Golec-Biernat and A. Stasto, Nucl. Phys. **B668**, 345 (2003).
[46] J. Berger and A. Stasto, Phys. Rev. D **83**, 034015 (2011).
[47] J. Berger and A. M. Stasto, Phys. Rev. D **84**, 094022 (2011).
[48] J. Berger and A. M. Stasto, J. High Energy Phys. 01 (2013) 001.
[49] J. L. Albacete and C. Marquet, Phys. Lett. B **687**, 174 (2010).
[50] A. Adare *et al.* (PHENIX Collaboration), Phys. Rev. Lett. **107**, 172301 (2011).
[51] I. Arsene *et al.* (BRAHMS Collaboration), Phys. Rev. Lett. **93**, 242303 (2004).
[52] B. Abelev *et al.* (ALICE Collaboration), Phys. Lett. B **717**, 162 (2012).
[53] V. Khachatryan *et al.* (CMS Collaboration), Phys. Rev. Lett. **105**, 022002 (2010).
[54] K. Eskola and H. Honkanen, Nucl. Phys. **A713**, 167 (2003).
[55] B. A. Meredith, Ph.D. thesis, University of Illinois at Urbana-Champaign, 2011.
[56] B. Abelev *et al.* (ALICE Collaboration), Phys. Rev. Lett. **110**, 082302 (2013).
[57] I. Helenius, K. J. Eskola, H. Honkanen, and C. A. Salgado, J. High Energy Phys. 07 (2012) 073.
[58] I. Helenius (private communication).