

# High energy factorization in nucleus-nucleus collisions. I

 François Gelis,<sup>1</sup> Tuomas Lappi,<sup>2</sup> and Raju Venugopalan<sup>3</sup>
<sup>1</sup>*Theory Division, PH-TH, Case C01600, CERN, CH-1211, Geneva 23, Switzerland*
<sup>2</sup>*Institut de Physique Théorique (URA 2306 du CNRS) CEA/DSM/Saclay, Bâtiment 774 91191, Gif-sur-Yvette Cedex, France*
<sup>3</sup>*Physics Department, Brookhaven National Laboratory Upton, New York 11973, USA*

(Received 6 May 2008; published 19 September 2008)

We derive a high energy factorization theorem for inclusive gluon production in  $A + A$  collisions. Our factorized formula resums (i) all-order leading logarithms  $(g^2 \ln(1/x_{1,2}))^n$  of the incoming parton momentum fractions, and (ii) all contributions  $(g\rho_{1,2})^n$  that are enhanced when the color charge densities in the two nuclei are of order of the inverse coupling— $\rho_{1,2} \sim g^{-1}$ . The resummed inclusive gluon spectrum can be expressed as a convolution of gauge invariant distributions  $W[\rho_{1,2}]$  from each of the nuclei with the leading order gluon number operator. These distributions are shown to satisfy the JIMWLK equation describing the evolution of nuclear wave functions with rapidity. As a by-product, we demonstrate that the JIMWLK Hamiltonian can be derived entirely in terms of retarded light-cone Green's functions without any ambiguities in their pole prescriptions. We comment on the implications of our results for understanding the Glasma produced at early times in  $A + A$  collisions at collider energies.

 DOI: [10.1103/PhysRevD.78.054019](https://doi.org/10.1103/PhysRevD.78.054019)

PACS numbers: 12.39.St, 12.38.Cy, 13.60.Hb, 24.85.+p

## I. INTRODUCTION

Collinear factorization theorems [1] that isolate long distance nonperturbative parton distribution functions from perturbatively calculable short distance matrix elements are central to the predictive power and success of QCD. These theorems can be applied to compute inclusive cross sections of the form  $A + B \rightarrow I(M) + X$ , where  $I(M)$  is a set of heavy particles or jets with invariant mass  $M$ , and  $X$  corresponds to the sum over all possible states (including soft and collinear hadrons) that can accompany the object  $I(M)$ . This cross section, for center-of-mass energy  $\sqrt{s}$ , can be expressed as [2–6]

$$\sigma_{AB} = \sum_{ab} \int dx_a dx_b f_{a/A}(x_a, \mu^2) f_{b/B}(x_b, \mu^2) \times \hat{\sigma}_{ab} \left( \frac{M^2}{x_a x_b s}, \frac{M}{\mu}, \alpha_s(\mu) \right) \left( 1 + \mathcal{O} \left( \frac{1}{M^n} \right) \right). \quad (1)$$

In this equation,  $f_{a(b)/A(B)}(x_{a(b)}, \mu^2)$  are the nonperturbative “leading twist” parton distribution functions which give the distribution of a parton  $a(b)$  in the hadron  $A(B)$ , as a function of the longitudinal momentum fraction  $x_{a(b)}$  evolved up to the factorization scale  $\mu^2$ , while the hard scattering matrix element  $\hat{\sigma}_{ab}$  can be computed systematically in a perturbative expansion in powers of  $\alpha_s = g^2/4\pi$ , where  $g$  is the QCD coupling constant. Higher twist contributions to this formula are suppressed by powers  $n$  of the hard scale  $M$ . This factorization formula is valid in the Bjorken limit when  $M^2 \sim s \gg \Lambda_{\text{QCD}}^2$  (where  $\Lambda_{\text{QCD}} \sim 200$  MeV is the intrinsic QCD scale).

Our interest here is, instead, in a different regime of high energy scattering where, for fixed invariant mass  $M \gg \Lambda_{\text{QCD}}$ , one takes  $\sqrt{s} \rightarrow \infty$  and thus  $x_{a,b} \rightarrow 0$ . We shall call this the Regge-Gribov limit of QCD. An important insight is that in this limit the field strengths squared can

become very large [ $\mathcal{O}(\frac{1}{\alpha_s})$ ] corresponding to the saturation of gluon densities [7,8]. The onset of saturation is characterized by a saturation scale  $Q_s(x)$ , which opens a kinematic window  $M^2 \sim Q_s^2 \gg \Lambda_{\text{QCD}}^2$  accessible at very high energies. The physics of the Regge-Gribov regime is quite different from that of the Bjorken limit discussed previously. The typical momenta of partons are  $\sim Q_s \gg \Lambda_{\text{QCD}}$ , and higher twist contributions are not suppressed. These considerations are especially relevant for the scattering of large nuclei because the large transverse density of partons in the nuclear wave functions (proportional to the nuclear radius  $\sim A^{1/3}$ ) provides a natural enhancement of the saturation scale,  $Q_s^2(x, A) \propto A^{1/3}$ . Our goal is to derive a formula similar to Eq. (1) for inclusive gluon production in the Regge-Gribov limit.

The dynamics of large parton phase-space densities in the Regge-Gribov limit can be described in the color glass condensate (CGC) effective field theory where small  $x$  partons in hadrons and nuclei are described by a classical field, while the large  $x$  partons act as color sources for the classical field [9–11]. The lack of dependence of physical observables on the (arbitrary) separation between large  $x$  color sources and small  $x$  dynamical fields is exploited to derive a renormalization group (RG) equation, known as the JIMWLK equation [12–19]. This equation is a functional RG equation describing the change in the statistical distribution of color sources  $W_Y[\rho]$  with rapidity  $Y (= \ln(1/x))$ . It can be expressed as

$$\frac{\partial W_Y[\rho]}{\partial Y} = \mathcal{H} W_Y[\rho], \quad (2)$$

where  $\mathcal{H}$  is the JIMWLK Hamiltonian.<sup>1</sup> For a physical

<sup>1</sup>The explicit form of this Hamiltonian will be given later in the text.

observable defined by an average over all the source configurations,

$$\langle \mathcal{O} \rangle_Y \equiv \int [D\rho] W_Y[\rho] \mathcal{O}[\rho], \quad (3)$$

one obtains

$$\frac{\partial \langle \mathcal{O} \rangle_Y}{\partial Y} = \langle \mathcal{H} \mathcal{O} \rangle_Y. \quad (4)$$

We have used here Eq. (2) and integrated by parts (using the Hermiticity of  $\mathcal{H}$ ). The structure of  $\mathcal{H}$  is such that  $\langle \mathcal{H} \mathcal{O} \rangle_Y$  is an object distinct from  $\langle \mathcal{O} \rangle_Y$ , so that one obtains, in principle, an infinite hierarchy of evolution equations for operator expectation values  $\langle \mathcal{O} \rangle_Y$  [20]. In the large  $N_c$  and large  $A$  mean-field limit, this hierarchy simplifies greatly. When  $\mathcal{O}$  is the ‘‘dipole’’ operator, corresponding to the forward scattering amplitude in deep inelastic scattering, the resulting closed evolution equation is known as the Balitsky-Kovchegov (BK) equation [21,22].

In Refs. [23–25], we developed a formalism to compute observables related to multiparticle production in field theories with strong time-dependent sources. This formalism is naturally applicable to the CGC description of high energy scattering,<sup>2</sup> albeit, for simplicity, we considered only a scalar  $\phi^3$  field theory. (The corresponding QCD framework was briefly considered in Ref. [26].) In these papers, the formalism for multiparticle production was developed for a fixed distribution of sources, with the assumption that the final results could be averaged over, as in Eq. (3), with unspecified distributions of sources  $W_{Y_1}[\rho_1]$  and  $W_{Y_2}[\rho_2]$  (one for each of the projectiles). However, we did not discuss, in these papers, the validity of such a factorization formula.

In the formalism of Refs. [23–25], one can formally arrange the perturbative expansion of an observable like the single inclusive gluon spectrum as

$$\mathcal{O}[\rho_1, \rho_2] = \frac{1}{g^2} [c_0 + c_1 g^2 + c_2 g^4 + \dots], \quad (5)$$

where each term corresponds to a different loop order. Each of the coefficients  $c_n$  is itself an infinite series of terms involving arbitrary orders in  $(g\rho_{1,2})^p$ . We call ‘‘leading order’’ the contribution that comes from the first coefficient  $c_0$ :

$$\mathcal{O}_{\text{LO}}[\rho_1, \rho_2] \equiv \frac{c_0}{g^2}. \quad (6)$$

In the case of the single gluon spectrum, the first term  $c_0/g^2$  has been studied extensively. In [24] we developed tools to calculate the next term  $c_1$ . Following this terminology, we denote

<sup>2</sup>Although the color sources of each nucleus are independent of the corresponding light-cone time, their sum constitutes a time-dependent current.

$$\mathcal{O}_{\text{NLO}}[\rho_1, \rho_2] \equiv c_1, \quad \mathcal{O}_{\text{NNLO}}[\rho_1, \rho_2] \equiv c_2 g^2, \dots \quad (7)$$

However, this strict loop expansion ignores the fact that large logarithms of the momentum fractions  $x_{1,2}$  can appear in the higher order coefficients  $c_{1,2,\dots}$  when  $\sqrt{s}$  is very large. The term  $c_n$  can have up to  $n$  powers of such logarithms, and a more precise representation of these coefficients is

$$c_n = \sum_{i=0}^n d_{ni} \ln^i \left( \frac{1}{x_{1,2}} \right). \quad (8)$$

The ‘‘leading log’’ terms are defined as those terms that have as many logarithms as their order in  $g^2$ ,

$$\mathcal{O}_{\text{LLog}}[\rho_1, \rho_2] \equiv \frac{1}{g^2} \sum_{n=0}^{\infty} d_{nn} g^{2n} \ln^n \left( \frac{1}{x_{1,2}} \right). \quad (9)$$

In this work, we will go significantly further than the leading order result, and resum the complete series of leading log terms. We will prove that, after averaging over the sources  $\rho_{1,2}$ , all the leading log corrections are automatically resummed by the JIMWLK evolution of the distribution of sources, and that the event averaged leading log result is given by the factorized expression

$$\begin{aligned} \langle \mathcal{O} \rangle_{\text{LLog}} &= \int [D\rho_1][D\rho_2] W_{Y_{\text{beam}}-Y}[\rho_1] \\ &\times W_{Y_{\text{beam}}+Y}[\rho_2] \mathcal{O}_{\text{LO}}[\rho_1, \rho_2; Y]. \end{aligned} \quad (10)$$

In this formula,  $Y$  is the rapidity at which the gluon is measured, and the subscripts  $Y_{\text{beam}} \pm Y$  indicate the amount of rapidity evolution<sup>3</sup> of the source distributions of the two projectiles, starting in their respective fragmentation regions. Note that we have added a rapidity argument  $Y$  to the factor  $\mathcal{O}_{\text{LO}}$  under the integral in order to stress the fact that our factorization result is valid only for observables that depend on a single rapidity. Our recurring example in the present paper will be the single inclusive gluon spectrum, and in a sequel to this paper [27] we generalize this result to the inclusive multigluon spectrum in the special case where all the measured gluons are in a small range of rapidity. In the rest of this paper, the rapidity argument in  $\mathcal{O}_{\text{LO}}$  will be implicit in order to keep the notations as compact as possible.

The expressions  $W_{Y_{\text{beam}} \pm Y}[\rho_{1,2}]$  in Eq. (10) are gauge invariant functionals describing the source distributions in each of the nuclei. In analogy to the parton distribution functions  $f_{a(b)/A(B)}(x_{a(b)}, \mu^2)$  we introduced previously, they contain nonperturbative information on the distribution of sources at rapidities close to the beam rapidities.

<sup>3</sup>In terms of the center-of-mass energy  $\sqrt{s}$  of the collision (for a nucleon-nucleon pair) and the longitudinal momentum components  $p^\pm$  of the measured gluon, one has also—at leading log— $Y_{\text{beam}} - Y = \ln(\sqrt{s}/p^+)$  and  $Y_{\text{beam}} + Y = \ln(\sqrt{s}/p^-)$ .

Just as the latter evolve in  $\mu^2$  with the DGLAP [28–30] evolution equations, the former, as suggested by Eq. (2), obey the JIMWLK evolution equation in rapidity which evolves them up to the rapidities  $Y_{\text{beam}} - Y$  and  $Y_{\text{beam}} + Y$  from the nuclei  $A_1$  and  $A_2$ , respectively. As we will discuss in detail, the leading order inclusive gluon spectrum, for given sources  $\rho_{1,2}$ , can be computed by solving the classical Yang-Mills equations with simple retarded boundary conditions. Equation (10) indicates that the resummation of the leading logarithms of the collision energy is equivalent to averaging the leading result with evolved (from the beam rapidity to the rapidity  $Y$  of the produced gluon) weight functionals  $W$ .

In the Regge-Gribov limit, Eq. (10) is the analog of the factorization formula Eq. (1) proved in the Bjorken limit. While we will prove that Eq. (10) holds for leading logarithmic contributions at all orders in perturbation theory, we have not attempted to show that it is valid for subleading logarithms. There is currently intense activity in computing subleading logarithmic contributions in the high parton density limit [31–37], so an extension of our results beyond leading logs is feasible in the future. There is another aspect of  $A + A$  collisions that we have not discussed thus far. Our power counting does not account for the so-called “secular divergences” [38–40]. These are contributions that diverge at least as powers of the time elapsed after the collision. Including these contributions will not alter our factorization theorem; it does affect how “observables” defined at finite times *after* the nuclear collisions are related to quantities measured in  $A + A$  experiments. We will address this issue briefly. A fuller treatment requires more work.

The paper is organized as follows. In Sec. II, we derive an important formula for the next-to-leading order corrections to the inclusive gluon spectrum. This formula will play a crucial role later, in disentangling the initial state effects from the rest of the collision process. In Sec. III, we will derive the expressions stated in Eqs. (2)–(4) for JIMWLK evolution of a single nucleus. Albeit the result is well known, our derivation is quite different from those existing in the literature [12–19,41–45]. We will obtain our result entirely in terms of retarded light-cone (LC) Green’s functions without any recourse to time-ordered propagators. We will show that there are no ambiguities in specifying the pole prescriptions in this approach. More importantly, our derivation allows us to straightforwardly extend our treatment of the JIMWLK equation to the case of the collision of two nuclei. This is discussed separately in Sec. IV, where we show explicitly that nonfactorizable terms are suppressed and our key result, stated in Eq. (10), is obtained. In the following section, we will relate our work to previous work in this direction and briefly explore some of the connections between the different approaches. In Sec. VI, we will discuss how one can relate our result for the Glasma produced at early times in heavy ion collisions

[25,46] and its subsequent evolution into the quark gluon plasma. We conclude with a brief summary and discussion of open issues. There are three appendixes dealing with properties of Green’s functions in light-cone gauge relevant to the discussion in the main text of the paper.

## II. NLO CORRECTIONS TO INCLUSIVE OBSERVABLES

Before studying the logarithmic divergences that arise in loop corrections to observables, let us derive a formula that expresses the 1-loop corrections to inclusive observables in terms of the action of a certain operator acting on the same observable at leading order. As we shall see, this formula—albeit quite formal—can be used to separate the physics of the initial state from the collision itself.

We have in mind an operator made of elementary color fields, which probes multigluon correlations. To be specific, for a given source distribution, we shall consider the quantum expectation value

$$\mathcal{O}(x, y) \equiv \langle A^i(x) A^j(y) \rangle, \quad (11)$$

in the limit where the time arguments of the two fields go to  $+\infty$ . We chose this particular operator because we wish to study the single gluon spectrum—the first moment of the multiplicity distribution—in the collision of two nuclei; it is obtained by Fourier transforming this bilinear combination of fields. Note that the two fields are not time ordered. The expectation value of such a product can be calculated in the Schwinger-Keldysh formalism [47–49], by considering that  $A^i(x)$  lies on the  $-$  branch of the contour and  $A^j(y)$  on the  $+$  branch. (A representation of the Schwinger-Keldysh contour is shown in Fig. 1.)

This section is organized as follows. We first recall the expression of Eq. (11) at leading order in terms of retarded solutions of the classical Yang-Mills equations. This result is well known and has been derived in a number of different ways. We will then discuss the next-to-leading order (NLO) computation of this quantity in the CGC framework. There are two sorts of NLO corrections; these are the virtual corrections arising from 1-loop corrections to the classical fields and the real corrections which are obtained by computing the  $\mathcal{G}_{-+}$  propagator of a small fluctuation in light-cone gauge. We will show that  $\mathcal{O}_{\text{NLO}}$  can be expressed as a linear operator with real and virtual pieces acting on  $\mathcal{O}_{\text{LO}}$ , plus an unimportant (as far as the resummation of logs of  $1/x_{1,2}$  is concerned) additional term.

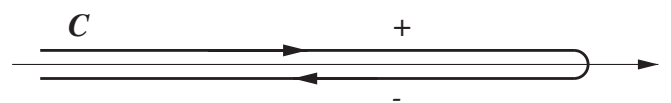


FIG. 1. The closed time path used in the Schwinger-Keldysh formalism.

**A. Leading order result**

We showed in [23] that, at leading order,  $\mathcal{O}$  is the product of two classical solutions of the Yang-Mills equations, with null retarded boundary conditions,<sup>4</sup>

$$\mathcal{O}_{\text{LO}}(x, y) = \mathcal{A}^i(x)\mathcal{A}^j(y), \quad (12)$$

with

$$[\mathcal{D}_\nu, \mathcal{F}^{\mu\nu}] = J^\nu, \quad \lim_{x^0 \rightarrow -\infty} \mathcal{A}^\mu(x) = 0. \quad (13)$$

Here,  $\mathcal{A}$  denotes the classical field, and  $J^\mu$  is the color current corresponding to a fixed configuration of the color sources. The current is comprised of one or two sources depending on whether we consider only one nucleus or the collision of two nuclei—this distinction is not important in this section. It is important to note that this current, which has support only on the light cone, must be covariantly conserved,

$$[\mathcal{D}_\mu, J^\mu] = 0. \quad (14)$$

This means that, in general, there is a feedback of the gauge field on the current itself, unless one chooses a gauge condition such that the gauge field does not couple to the nonzero components of the current on the light cone.

Although one can solve analytically the Yang-Mills equations with these boundary conditions in the case of a single nucleus [12,51], this is not possible in the case of two nuclei, and one must resort to numerical methods to obtain results in this case. Fortunately, as we shall see, the discussion of factorization in the case of two nuclei does not require that we know this solution analytically.

Because the solution of the Yang-Mills equations we need is defined with retarded boundary conditions, its value at the points  $x$  and  $y$  (where the observable is measured) is fully determined if we know its value<sup>5</sup> on an initial surface  $\Sigma$ —which is locally spacelike<sup>6</sup>—located below the points  $x$  and  $y$ , as illustrated in Fig. 2.

Therefore, we will write

$$\mathcal{O}_{\text{LO}}(x, y) \equiv \mathcal{O}_{\text{LO}}[\mathcal{A}], \quad (15)$$

which means that the observable is considered as a func-

<sup>4</sup>The retarded nature of the boundary conditions is intimately related to the inclusiveness of the observable under consideration. For instance, if instead of the single inclusive gluon spectrum, one wanted to calculate at leading order the probability of producing a fixed number of gluons, one would have to solve the classical Yang-Mills equations with boundary conditions both at  $x^0 = -\infty$  and at  $x^0 = +\infty$  (see [50]).

<sup>5</sup>Since the Yang-Mills equations contain second derivatives with respect to time, one must also know the value of the first time derivative of the field on this initial surface.

<sup>6</sup>This means that at every point  $u \in \Sigma$ , the vector  $n^\mu$  normal to  $\Sigma$  at the point  $u$  ( $n^\mu dx_\mu = 0$  for any displacement  $dx_\mu$  on  $\Sigma$  around the point  $u$ ) must be timelike. This condition prevents a signal emitted at the point  $u \in \Sigma$ , propagating at the speed of light, from encountering the surface  $\Sigma$  again.

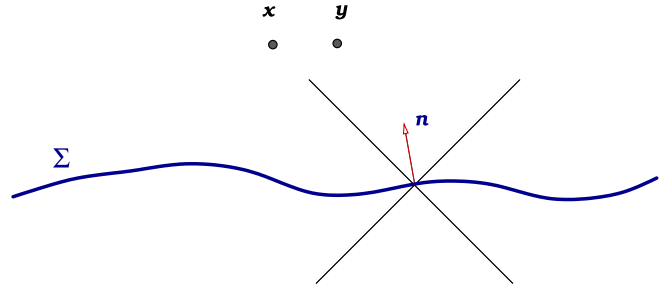


FIG. 2 (color online). A locally spacelike surface  $\Sigma$  used to define the initial value of the color field.

tional of the value of the color field on the initial surface  $\Sigma$ . Note that we use the same symbol for the color field and for its initial value on  $\Sigma$ , although mathematically these objects depend on a different number of variables and are therefore different functions.

**B. Next-to-leading order corrections**

A detailed discussion of the power counting for moments of the inclusive multiplicity distribution can be found in Ref. [23]. The leading order contribution to  $\mathcal{O}(x, y)$  involves only tree diagrams, which explains why it can be obtained from classical solutions of the Yang-Mills equations. As mentioned previously, this leading order contribution is of order  $\mathcal{O}(\alpha_s^{-1})$  but includes all orders in  $g\rho$ . In the rest of this section, we shall study the 1-loop corrections to this quantity, that are of order  $\mathcal{O}(1)$  in the coupling and to all orders in  $g\rho$ .

The framework to compute these 1-loop corrections (hereafter called “NLO”) to quantities such as Eq. (11) has been developed for a scalar theory in Ref. [24]. Much of this analysis can be carried over to QCD. To avoid complications such as ghost loops, we shall work in a gauge such as the light-cone gauge  $A^+ = 0$ . Following the discussion for the scalar case, we obtain at NLO

$$\mathcal{O}_{\text{NLO}}(x, y) = \mathcal{A}^i(x)\beta^j(y) + \beta^i(x)\mathcal{A}^j(y) + \mathcal{G}_{-+}^{ij}(x, y). \quad (16)$$

In this equation,  $\mathcal{G}_{-+}^{ij}(x, y)$  is the  $-+$  component of the small fluctuation Schwinger-Keldysh propagator in the presence of the classical background field  $\mathcal{A}^i$  and the field  $\beta^i$  is the 1-loop correction to the classical field. It is obtained by solving the small fluctuation equation of motion,

$$\left[ \square_x g^{\mu\nu} - \partial_x^\mu \partial_x^\nu - \frac{\partial^2 U(\mathcal{A})}{\partial \mathcal{A}_\mu(x) \partial \mathcal{A}_\nu(x)} \right] \beta_\nu(x) = \frac{1}{2} \frac{\partial^3 U(\mathcal{A})}{\partial \mathcal{A}_\mu(x) \partial \mathcal{A}^\nu(x) \partial \mathcal{A}^\rho(x)} \mathcal{G}_{++}^{\nu\rho}(x, x), \quad (17)$$

with null retarded boundary conditions:

$$\lim_{x^0 \rightarrow -\infty} \beta^\mu(x) = 0. \quad (18)$$

Here  $U(\mathcal{A})$  is the potential term in the Yang-Mills Lagrangian,<sup>7</sup> obtained by writing

$$\mathcal{L} = \mathcal{L}_{\text{quad}} - U(\mathcal{A}), \quad (19)$$

where  $\mathcal{L}_{\text{quad}}$  is defined in Eq. (A2) of Appendix A. We refer the reader to Appendix A for more details. The source term in this small fluctuation equation includes the closed loop formed by the Schwinger-Keldysh propagator  $\mathcal{G}_{++}(x, x)$  to be defined shortly, the third derivative corresponds to the 3-gluon vertex in the presence of a background field, and  $1/2$  is a symmetry factor.

Following [24], we can write the propagator  $\mathcal{G}_{-+}^{ij}(x, y)$  in Eq. (16) as a bilinear combination of small fluctuations of the gauge field *whose initial conditions are plane waves*,

$$\mathcal{G}_{-+}^{ij,bc}(x, y) = \sum_{\lambda, a} \int \frac{d^3 \mathbf{k}}{(2\pi)^3 2E_{\mathbf{k}}} a_{-k\lambda a}^{ib}(x) a_{+k\lambda a}^{jc}(y), \quad (20)$$

where

$$\left[ \square_x g^{\mu\nu} - \partial_x^\mu \partial_x^\nu - \frac{\partial^2 U(\mathcal{A})}{\partial \mathcal{A}_\mu(x) \partial \mathcal{A}_\nu(x)} \right] a_{\pm k\lambda a, \nu}(x) = 0, \quad (21)$$

$$\lim_{x^0 \rightarrow -\infty} a_{\pm k\lambda a}^\mu(x) = \epsilon_\lambda^\mu(\mathbf{k}) T^a e^{\pm i\mathbf{k}\cdot\mathbf{x}}.$$

The sum over  $\lambda$  is over the two physical polarizations for the initial plane wave, and the index  $a$  represents the initial color carried by the small fluctuation field. In Eq. (20), our notation is such that the lower color index ( $a$ ) represents the initial color of the fluctuation, while the upper color index ( $b$  or  $c$ ) refers to its color after it has evolved on top of the classical background field.<sup>8</sup> It is important to stress that this decomposition of  $\mathcal{G}_{-+}^{ij}$  is valid only if one uses small fluctuations that are plane waves in the remote past. Using other solutions of the small fluctuation equation of motion (21) would lead to a propagator that obeys incorrect boundary conditions.

The  $++$  propagator at equal points can be written in a similar fashion as<sup>9</sup>

$$\mathcal{G}_{++}^{ij,bc}(x, x) = \sum_{\lambda, a} \int \frac{d^3 \mathbf{k}}{(2\pi)^3 2E_{\mathbf{k}}} a_{-k\lambda a}^{ib}(x) a_{+k\lambda a}^{jc}(x). \quad (23)$$

<sup>7</sup>Unless one chooses a nonlinear gauge condition,  $U(\mathcal{A})$  is made of the usual 3-gluon and 4-gluon couplings.

<sup>8</sup>For future reference, note that quantities with only the lower color index are matrices in the adjoint representation of  $SU(N)$  defined by

$$a_{\pm k\lambda a}^\mu(x) \equiv a_{\pm k\lambda a}^{\mu b}(x) T^b. \quad (22)$$

<sup>9</sup>When the two endpoints are separated by a timelike interval, there can be an additional term contributing to this propagator—see [24] for more general formulas.

We note that, in a generic gauge, covariant current conservation may require the incoming field fluctuation to induce a color precession of the classical current  $J^\mu$ . This modification of the current will, in turn, induce an additional contribution to the field fluctuation. Our strategy [52–54] to avoid this complication will be to perform all intermediate calculations in a gauge where this phenomenon does not happen. For instance, on the line  $x^- = 0$  where the color charges moving in the  $+z$  direction are located, one should use a gauge in which  $\mathcal{A}^- = 0$ . Indeed, because the color current only has a  $+$  component, covariant conservation is trivial in this gauge. A gauge rotation of the final result is then performed to return to the light-cone gauge of interest. All effects due to current conservation are then taken care of by this final gauge transformation.

### C. Rearrangement of the NLO corrections—I

In this subsection, we will express the small fluctuation propagator  $\mathcal{G}_{-+}^{ij}(x, y)$  as the action of a differential operator on the classical fields  $\mathcal{A}^i(x)$  and  $\mathcal{A}^j(y)$ . This operator contains functional derivatives with respect to the initial value of the color field on  $\Sigma$ . In the following subsection, we will repeat the exercise for the 1-loop correction to the classical field  $\beta^\mu(x)$  and write it in terms of a similar operator acting on the classical field  $\mathcal{A}^\mu(x)$ . These identities, besides providing a transparent derivation of the JIMWLK equation for a single nucleus, will be especially powerful in our treatment of nucleus-nucleus collisions.

Let us begin from the Green's formula for the classical field  $\mathcal{A}^\mu$ ,

$$\mathcal{A}^\mu(x) = \int_{\Sigma^+} d^4 y D_{0,R}^{\mu\rho}(x, y) \frac{\partial U(\mathcal{A})}{\partial \mathcal{A}^\rho(y)} + \mathcal{B}_0^\mu[\mathcal{A}](x), \quad (24)$$

where  $D_{0,R}^{\mu\rho}(x, y)$  is the free retarded propagator (discussed in Appendix A in the case of the light-cone gauge) and  $\mathcal{B}_0^\mu[\mathcal{A}](x)$  is the boundary term that contains the initial value of the classical field on  $\Sigma$ . (Boundary terms for the classical and small fluctuation fields in light-cone gauge are discussed in detail in Appendix B.)  $\Sigma^+$  denotes the region of space-time above the surface  $\Sigma$ . Now, consider an operator  $\mathcal{T}$  (to be defined explicitly later) that acts on the initial value of the fields on the surface  $\Sigma$ , and assume that this operator is linear, which implies

$$\mathcal{T} \frac{\partial U(\mathcal{A})}{\partial \mathcal{A}^\rho(y)} = \frac{\partial^2 U(\mathcal{A})}{\partial \mathcal{A}^\rho(y) \partial \mathcal{A}^\nu(y)} \mathcal{T} \mathcal{A}^\nu(y). \quad (25)$$

Now applying this operator  $\mathcal{T}$  to both sides of Eq. (24), we get

$$\begin{aligned} \mathcal{T} \mathcal{A}^\mu(x) &= \int_{\Sigma^+} d^4 y D_{0,R}^{\mu\rho}(x, y) \frac{\partial^2 U(\mathcal{A})}{\partial \mathcal{A}^\rho(y) \partial \mathcal{A}^\nu(y)} \\ &\quad \times \mathcal{T} \mathcal{A}^\nu(y) + \mathcal{T} \mathcal{B}_0^\mu[\mathcal{A}](x). \end{aligned} \quad (26)$$

By comparing this equation with the Green's formula for a small fluctuation  $a^\mu$  (see Appendix B),

$$a^\mu(x) = \int_{\Sigma^+} d^4y D_{0,R}^{\mu\rho}(x,y) \frac{\partial^2 U(\mathcal{A})}{\partial \mathcal{A}^\rho(y) \partial A^\nu(y)} a^\nu(y) + \mathcal{B}_0^\mu[a](x), \quad (27)$$

we see that we can identify  $a^\mu(x) = \mathcal{T} \mathcal{A}^\mu(x)$  provided that we have

$$\mathcal{B}_0^\mu[a](x) = \mathcal{T} \mathcal{B}_0^\mu[\mathcal{A}](x). \quad (28)$$

Because  $\mathcal{B}_0$  is a *linear* functional of the initial value of the color fields on the surface  $\Sigma$ , it is easy to see that the operator  $\mathcal{T}$  that fulfills this goal is

$$\mathcal{T} \equiv \int_{\Sigma} d^3\vec{u} [a \cdot \mathbb{T}_u], \quad (29)$$

where  $\mathbb{T}_u$  is the generator of translations of the initial fields<sup>10</sup> at the point  $u \in \Sigma$ . We denote by  $d^3\vec{u}$  the measure on the surface  $\Sigma$  (for instance, if  $\Sigma$  is a surface defined by  $x^- = \text{const}$ , this measure reads  $d^3\vec{u} = du^+ d^2u_\perp$ .) The detailed expression of this operator can be obtained by writing explicitly the Green's formula for the retarded propagation of color fields above the surface  $\Sigma$ , and it usually depends both on the choice of the surface and on the choice of the gauge condition. An explicit expression of this operator will be given in the next section when the initial surface  $\Sigma$  is parallel to the light cone ( $u^- = \text{const}$ ) and when the fields are in the light-cone gauge  $\mathcal{A}^+ = 0$ . Therefore, we have established the following identity,

$$a^\mu(x) = \int_{\Sigma} d^3\vec{u} [a \cdot \mathbb{T}_u] \mathcal{A}^\mu(x). \quad (30)$$

Equation (30) provides a formal expression of a fluctuation at point  $x$  in terms of its value on some initial surface  $\Sigma$  [in the right-hand side (r.h.s.) of Eq. (30), only the value of the fluctuation  $a^\mu$  on  $\Sigma$  appears]. This formula is especially useful in situations where we can calculate analytically the initial value of the fluctuation on  $\Sigma$ , but where we do not know analytically the classical background field  $\mathcal{A}$  above this surface.

The single nucleus case is a bit academic in this respect because one can analytically compute the background gauge field and the fluctuation at any point in space-time. Rather, Eq. (30) will prove especially powerful for nuclear collisions because, in that case, one does not have an analytic expression for the classical background field after the collision.

Armed with Eq. (30), it is straightforward to write the third term of the right-hand side of Eq. (16) as

$$\mathcal{G}_{-+}^{ij,bc}(x,y) = \sum_{\lambda,a} \int \frac{d^3k}{(2\pi)^3 2E_k} \int_{\Sigma} d^3\vec{u} d^3\vec{v} [[a_{-k\lambda a} \cdot \mathbb{T}_u] \times \mathcal{A}^{ib}(x) [[a_{+k\lambda a} \cdot \mathbb{T}_v] \mathcal{A}^{jc}(y)]. \quad (31)$$

<sup>10</sup>For now, it is sufficient to think of this operator as an operator which is linear in first derivatives with respect to the color field on  $\Sigma$ .

In this equation, the brackets limit the scope of the operators  $\mathbb{T}_{u,v}$ .

#### D. Rearrangement of the NLO corrections—II

The terms involving the 1-loop correction  $\beta^\mu$  can also be written in terms of the operator  $\mathbb{T}_u$ , but this is not as straightforward as for  $\mathcal{G}_{-+}^{ij}$ . The first step is to write down the formal Green's function solution of Eq. (17). It is convenient to write it as

$$\beta^\mu(x) = \int_{\Sigma^+} d^4y D_R^{\mu\nu}(x,y) \underbrace{\frac{1}{2} \frac{\partial^3 U(\mathcal{A})}{\partial \mathcal{A}^\nu(y) \partial \mathcal{A}^\rho(y) \partial \mathcal{A}^\sigma(y)} \mathcal{G}_{++}^{\rho\sigma}(y,y)}_{\beta_1^\mu(x)} + \underbrace{\mathcal{B}^\mu[\beta]}_{\beta_2^\mu(x)}(x), \quad (32)$$

where  $\mathcal{B}^\mu[A](x)$  is identical to  $\mathcal{B}_0^\mu[A](x)$  except that all occurrences of the bare propagator  $D_{0,R}^{\mu\nu}$  in the latter are replaced in the former by the dressed propagator in the background field  $\mathcal{A}^\mu$ . This dressed propagator, denoted  $D_R^{\mu\nu}$ , satisfies the equation

$$\left[ \square_x g^{\mu\nu} - \partial_x^\mu \partial_x^\nu - \frac{\partial^2 U(\mathcal{A})}{\partial \mathcal{A}_\mu(x) \partial \mathcal{A}_\nu(x)} \right] D_{R,\mu}^\rho(x,y) = g^{\rho\nu} \delta(x-y), \quad (33)$$

plus a retarded boundary condition such that it vanishes if  $x^0 < y^0$ .

The second term on the right-hand side of Eq. (32) is the value  $\beta$  would have if one turns off the source term (proportional to  $\mathcal{G}_{++}$ ) in the domain  $\Sigma^+$  above the initial surface. It is therefore given by a formula identical to Eq. (30),

$$\beta_2^\mu(x) = \int_{\Sigma} d^3\vec{u} [\beta \cdot \mathbb{T}_u] \mathcal{A}^\mu(x). \quad (34)$$

To calculate  $\beta_1(x)$ , let us first make explicit the interactions with the background field by writing it as

$$\beta_1^\mu(x) = \int_{\Sigma^+} d^4y D_{0,R}^{\mu\nu}(x,y) \left[ \frac{\partial^2 U(\mathcal{A})}{\partial \mathcal{A}^\nu(y) \partial \mathcal{A}^\rho(y)} \beta_1^\rho(y) + \frac{1}{2} \frac{\partial^3 U(\mathcal{A})}{\partial \mathcal{A}^\nu(y) \partial \mathcal{A}^\rho(y) \partial \mathcal{A}^\sigma(y)} \mathcal{G}_{++}^{\rho\sigma}(y,y) \right]. \quad (35)$$

This expression is obtained by substituting the expression for the dressed retarded propagator in terms of the free retarded propagator in the definition of  $\beta_1^\mu$ .

Consider now the quantity

$$\zeta^\mu(x) \equiv \frac{1}{2} \sum_{\lambda,a} \int \frac{d^3k}{(2\pi)^3 2E_k} \int_{\Sigma} d^3\vec{u} d^3\vec{v} [a_{-k\lambda a} \cdot \mathbb{T}_u] \times [a_{+k\lambda a} \cdot \mathbb{T}_v] \mathcal{A}^\mu(x). \quad (36)$$

We shall prove that  $\beta_1^\mu$  and  $\zeta^\mu$  are identical. Using Eq. (30), we can write

$$\zeta^\mu(x) = \frac{1}{2} \sum_{\lambda,a} \int \frac{d^3k}{(2\pi)^3 2E_k} \int_{\Sigma} d^3\vec{u} [a_{-k\lambda a} \cdot \mathbb{T}_u] a_{+k\lambda a}^\mu(x). \quad (37)$$

Replace  $a_{+k\lambda a}^\mu(x)$  in this equation by the r.h.s. of Eq. (27).

$$\begin{aligned} \zeta^\mu(x) &= \frac{1}{2} \sum_{\lambda,a} \int \frac{d^3k}{(2\pi)^3 2E_k} \int_{\Sigma} d^3\vec{u} \int_{\Sigma^+} d^4y D_{0,R}^{\mu\nu}(x,y) \left\{ \frac{\partial^2 U(\mathcal{A})}{\partial \mathcal{A}^\nu(y) \partial \mathcal{A}^\rho(y)} [a_{-k\lambda a} \cdot \mathbb{T}_u] a_{+k\lambda a}^\rho(y) \right. \\ &\quad \left. + \frac{\partial^3 U(\mathcal{A})}{\partial \mathcal{A}^\nu(y) \partial \mathcal{A}^\rho(y) \partial \mathcal{A}^\sigma(y)} [[a_{-k\lambda a} \cdot \mathbb{T}_u] \mathcal{A}^\sigma(y)] a_{+k\lambda a}^\rho(y) \right\} \\ &= \int_{\Sigma^+} d^4y D_{0,R}^{\mu\nu}(x,y) \left[ \frac{\partial^2 U(\mathcal{A})}{\partial \mathcal{A}^\nu(y) \partial \mathcal{A}^\rho(y)} \zeta^\rho(y) + \frac{\partial^3 U(\mathcal{A})}{\partial \mathcal{A}^\nu(y) \partial \mathcal{A}^\rho(y) \partial \mathcal{A}^\sigma(y)} \mathcal{G}^{\rho\sigma}(y,y) \right], \end{aligned} \quad (38)$$

which is identical to Eq. (35). We therefore obtain  $\beta_1^\mu(x) = \zeta^\mu(x)$ . Combining the two contributions  $\beta_1$  and  $\beta_2$ , we finally arrive at the compact expression

$$\beta^\mu(x) = \left[ \int_{\Sigma} d^3\vec{u} [\beta \cdot \mathbb{T}_u] + \frac{1}{2} \sum_{\lambda,a} \int \frac{d^3k}{(2\pi)^3 2E_k} \int_{\Sigma} d^3\vec{u} d^3\vec{v} [a_{-k\lambda a} \cdot \mathbb{T}_u] [a_{+k\lambda a} \cdot \mathbb{T}_v] \right] \mathcal{A}^\mu(x). \quad (39)$$

We can now use Eqs. (31) and (39) to obtain a compact expression for NLO corrections to  $\mathcal{O}$  as

$$\begin{aligned} \mathcal{O}_{\text{NLO}}(x,y) &= \left[ \int_{\Sigma} d^3\vec{u} [\beta \cdot \mathbb{T}_u] \right. \\ &\quad \left. + \frac{1}{2} \sum_{\lambda,a} \int \frac{d^3k}{(2\pi)^3 2E_k} \int_{\Sigma} d^3\vec{u} d^3\vec{v} [a_{-k\lambda a} \cdot \mathbb{T}_u] [a_{+k\lambda a} \cdot \mathbb{T}_v] \right] \mathcal{O}_{\text{LO}}[\mathcal{A}] \\ &\quad + \Delta \mathcal{O}_{\text{NLO}}(x,y), \end{aligned} \quad (40)$$

where we recall that  $\mathcal{O}_{\text{LO}}[\mathcal{A}]$  is the same observable at *leading order*, considered as a functional of the value of the gauge fields on the initial surface  $\Sigma$ . The corrective term  $\Delta \mathcal{O}_{\text{NLO}}(x,y)$  is defined by

$$\begin{aligned} \Delta \mathcal{O}_{\text{NLO}}(x,y) &\equiv \frac{1}{2} \sum_{\lambda,a} \int \frac{d^3k}{(2\pi)^3 2E_k} \\ &\quad \times \int_{\Sigma} d^3\vec{u} d^3\vec{v} \{ [[a_{-k\lambda a} \cdot \mathbb{T}_u] \mathcal{A}^{ib}(x)] \\ &\quad \times [[a_{+k\lambda a} \cdot \mathbb{T}_v] \mathcal{A}^{jc}(y)] \\ &\quad - [[a_{+k\lambda a} \cdot \mathbb{T}_u] \mathcal{A}^{ib}(x)] \\ &\quad \times [[a_{-k\lambda a} \cdot \mathbb{T}_v] \mathcal{A}^{jc}(y)] \}. \end{aligned} \quad (41)$$

As we shall see later, this term  $\Delta \mathcal{O}_{\text{NLO}}$  does not contain any large logarithm. Only the terms in the first and second lines of Eq. (40) will be important for our later discussion of factorization.

### III. JIMWLK EVOLUTION FOR A SINGLE NUCLEUS

Equation (40) is central to our study of NLO corrections and of factorization. In the rest of this section, we will show how this formula is used to derive the JIMWLK evolution equation. In Sec. IV, we will show that it can be generalized to the collision of two nuclei. A very convenient choice of

Because the boundary term  $\mathcal{B}_0^\mu[a_{+k\lambda a}](x)$  does not depend on the initial value of the classical field  $\mathcal{A}$ , the action of  $[a_{-k\lambda a} \cdot \mathbb{T}_u]$  on this term gives zero. We thus obtain

initial surface  $\Sigma$  in the derivation of the JIMWLK equation is the surface defined by  $x^- = \epsilon$ . One should choose  $\epsilon$  so that all the color sources of the nucleus are located in the strip  $0 \leq x^- \leq \epsilon$ . An illustration of the objects involved in Eq. (40) and their localization in space-time is provided in Fig. 3.

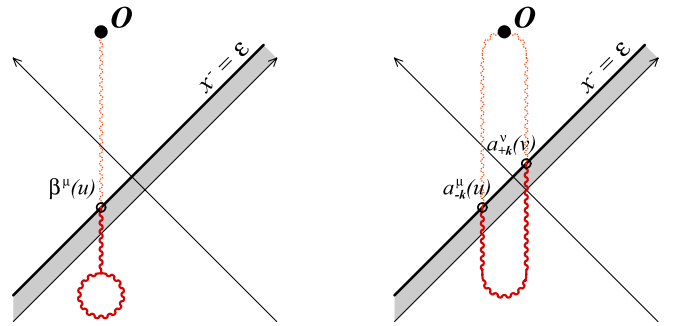


FIG. 3 (color online). NLO corrections in the single nucleus case, seen as an initial value problem on the surface  $x^- = \epsilon$ . The shaded area represents the domain where the nuclear color sources live ( $0 \leq x^- \leq \epsilon$ ). The field fluctuations represented by wavy lines continue to evolve in the region  $x^- > \epsilon$  until they hit the operator we want to evaluate. However, this evolution is entirely hidden in the dependence of the classical field upon its initial value at  $x^- = \epsilon$ , and we do not need to consider it explicitly.

### A. Gauge choice

We need first to choose the gauge in which to perform this calculation. Because the observable we wish to calculate and everything else in Eq. (40) are expressed in terms of light-cone gauge ( $A^+ = 0$ ) quantities, we need to obtain  $a_{\pm k\lambda a}$  and  $\beta$  in this gauge as well. However, as previously mentioned, covariant current conservation is most easily preserved in a gauge where the field fluctuations have no  $-$  component. This is because they do not induce a precession of the color current  $J^+$  while crossing the light cone. We are therefore going to adopt the strategy advocated in Refs. [17,19,52,53], which consists in performing intermediate calculations in a gauge where  $A^- = 0$  and then gauge transforming the final result to  $A^+ = 0$  gauge.

As discussed in detail in Appendix B, if one uses the LC gauge and the surface  $u^- = \epsilon$  as the initial surface, the linear differential operator  $a \cdot \mathbb{T}_u$  that appears in the identity (30) should be defined as<sup>11</sup>

$$\begin{aligned} a \cdot \mathbb{T}_u &= \partial^-(\Omega(u)a^i(u)) \frac{\delta}{\delta(\partial^-(\Omega(u)\mathcal{A}^i(u)))} \\ &+ \Omega(u)a^-(u) \frac{\delta}{\delta(\Omega(u)\mathcal{A}^-(u))} \\ &+ \partial^\mu(\Omega(u)a_\mu(u)) \frac{\delta}{\delta(\partial^\mu(\Omega(u)\mathcal{A}_\mu(u)))}, \end{aligned} \quad (42)$$

where  $\Omega$  is the adjoint color matrix<sup>12</sup> that will be defined in Eq. (46). Note that this operator in Eq. (42) contains a term for each of the field components that must be specified on the initial surface to know completely the field above this surface. This operator  $\mathbb{T}_u$  can therefore be interpreted as the generator of translations of the initial condition for a classical solution of the Yang-Mills equations. It is also important to note that the fluctuation field  $a^\mu(u)$  that multiplies this operator is evaluated just above the initial surface (at  $u^- = \epsilon$ ). Therefore, because one does not require its entire history beyond this surface, it can, in general, be calculated analytically.

### B. Classical field

Let us recall the structure of the classical background field itself. As is well known, the field in the Lorenz gauge ( $\partial_\mu A^\mu = 0$ ) has no  $A^-$  component, and therefore fulfills the  $A^- = 0$  condition. Its explicit expression in terms of the color source<sup>13</sup>  $\tilde{\rho}$  is given by

$$\tilde{\mathcal{A}}^+(x) = -\frac{1}{\partial_\perp^2} \tilde{\rho}(x^-, \mathbf{x}_\perp), \quad \tilde{\mathcal{A}}^- = \tilde{\mathcal{A}}^i = 0. \quad (43)$$

The gauge transformation that relates the classical background fields in the  $A^+ = 0$  gauge and the corresponding fields in Lorenz gauge is<sup>14</sup>

$$\mathcal{A}^\mu = \Omega^\dagger \tilde{\mathcal{A}}^\mu \Omega + \frac{i}{g} \Omega^\dagger \partial^\mu \Omega, \quad (44)$$

where the tilde denotes fields in the Lorenz gauge; those without a tilde are in light-cone gauge. Using the light-cone gauge condition  $A^+ = 0$ , we get

$$\partial^+ \Omega = ig \tilde{\mathcal{A}}^+ \Omega, \quad (45)$$

which admits the Wilson line

$$\Omega(x^-, \mathbf{x}_\perp) \equiv \text{T exp} \left[ ig \int_{-\infty}^{x^-} dz^- \tilde{\mathcal{A}}_a^+(z^-, \mathbf{x}_\perp) T^a \right] \quad (46)$$

as a solution. Note that because the color sources do not depend on  $x^+$ ,  $\tilde{\mathcal{A}}^+$  and  $\Omega$  depend only on  $x^-$  and  $\mathbf{x}_\perp$ . The solution of the classical equations of motion in light-cone gauge is then

$$\begin{aligned} \mathcal{A}^+ &= \mathcal{A}^- = 0, \\ \mathcal{A}^i(x) &= \frac{i}{g} \Omega^\dagger(x^-, \mathbf{x}_\perp) \partial^i \Omega(x^-, \mathbf{x}_\perp). \end{aligned} \quad (47)$$

We should comment here on the residual gauge freedom of the classical solution. The most general solution of Eq. (45) is

$$\Omega(x^-, \mathbf{x}_\perp) \Theta(x^+, \mathbf{x}_\perp), \quad (48)$$

where  $\Theta$  is an arbitrary  $x^-$ -independent gauge transformation. With this more general choice, one obtains

$$\begin{aligned} \mathcal{A}^+ &= 0, \quad \mathcal{A}^- = \frac{i}{g} \Theta^\dagger \partial^- \Theta, \\ \mathcal{A}^i &= \Theta^\dagger \left[ \frac{i}{g} \Omega^\dagger \partial^i \Omega \right] \Theta + \frac{i}{g} \Theta^\dagger \partial^i \Theta. \end{aligned} \quad (49)$$

The arbitrariness in the solution is because the condition  $A^+ = 0$  does not completely fix the gauge, and  $x^-$ -independent  $\Theta$ 's span the residual gauge freedom. Requiring that the classical gauge field be of the form given in Eq. (47) amounts to the choice  $\Theta \equiv \mathbf{1}$ . This choice is assumed in the rest of this paper.

### C. Field fluctuations on the light cone

To readers familiar with the structure of the JIMWLK Hamiltonian, the structure of Eq. (40) is already sugges-

<sup>11</sup>We have omitted the color indices in this equation.  $\Omega$  should be understood as a matrix in the  $SU(N)$  group, and  $\mathcal{A}$  as a column vector.  $\Omega \mathcal{A}$  is therefore a column vector whose components are  $(\Omega \mathcal{A})_c \equiv \Omega_{cb} \mathcal{A}_b$ .

<sup>12</sup>At first sight,  $\Omega$  does not play any role in the definition of  $\mathbb{T}_u$ —the necessity to introduce this matrix  $\Omega$  in the definition of  $\mathbb{T}_u$  is also explained in Appendix B.

<sup>13</sup>The density  $\rho$  of color sources is a gauge-dependent quantity. When defined in the Lorenz gauge, we denote it with a tilde.

<sup>14</sup>In this expression,  $\Omega$  is a matrix in the group  $SU(N)$ , while  $\tilde{\mathcal{A}}$  is a matrix in the adjoint representation of the algebra  $SU(N)$ . The product  $\Omega^\dagger \tilde{\mathcal{A}} \Omega$  is a matrix in the  $SU(N)$  algebra. Note that, depending on the context, we use the same symbol for an element  $\mathcal{A}$  of the algebra (i.e. a matrix) and for the vector column made of its components  $\mathcal{A}_c$  on the basis of the algebra. The relation between the two is of course  $\mathcal{A} = \mathcal{A}_c T^c$ .



tive. In the rest of this section, we will show that the leading logarithmic contributions in this formula—terms that are linear in the rapidity differences between the projectile and target relative to the observed gluon—can be absorbed into a redefinition of the distribution of color sources of the nucleus. Our first task towards this conclusion is to compute the value of the field fluctuations  $a_{\pm k\lambda a}$  and  $\beta$  just above the light cone on the initial surface  $u^- = \epsilon$ .

Let us consider a small fluctuation  $a^\mu$  on top of the classical field  $\mathcal{A}^\mu$ . The relation between the two gauges must be modified,

$$\mathcal{A}^\mu + a^\mu = \bar{\Omega}^\dagger (\tilde{\mathcal{A}}^\mu + \tilde{a}^\mu) \bar{\Omega} + \frac{i}{g} \bar{\Omega}^\dagger \partial^\mu \bar{\Omega}, \quad (50)$$

with

$$\bar{\Omega} \equiv (1 + ig\omega)\Omega, \quad (51)$$

where  $\omega$  has components of order unity. Using this ansatz in Eq. (50), and keeping in mind that  $\mathcal{A}, \tilde{\mathcal{A}} \sim \mathcal{O}(g^{-1})$  while  $a, \tilde{a} \sim \mathcal{O}(1)$ , we obtain the relation

$$a^\mu = \Omega^\dagger (\tilde{a}^\mu - ig[\omega, \tilde{\mathcal{A}}^\mu] - \partial^\mu \omega) \Omega. \quad (52)$$

To determine  $\omega$ , as previously, apply the gauge condition  $a^+ = 0$ . This gives

$$\partial^+ \omega + ig[\omega, \tilde{\mathcal{A}}^+] = \tilde{a}^+, \quad (53)$$

the solution of which can be written as

$$\omega(x) = \Omega(x^-, \mathbf{x}_\perp) f(x^+, \mathbf{x}_\perp) + \int_{-\infty}^{x^-} dz^- \Omega(x^-, z^-; \mathbf{x}_\perp) \tilde{a}^+(z^-, x^+, \mathbf{x}_\perp). \quad (54)$$

In this equation  $f$  is an arbitrary function that does not depend on  $x^-$ , and  $\Omega(x^-, z^-; \mathbf{x}_\perp)$  is an “incomplete” Wilson line defined by

$$\Omega(x^-, z^-; \mathbf{x}_\perp) \equiv \text{Texp} \left\{ ig \int_{z^-}^{x^-} dz^- \tilde{\mathcal{A}}_a^+(z^-, \mathbf{x}_\perp) T^a \right\}. \quad (55)$$

The arbitrariness in the choice of the function  $f_b$  again means that there is a residual gauge freedom after we have imposed  $a^+ = 0$ .

A crucial point in our derivation is how the residual gauge freedom is fixed. We need small field fluctuations in order to represent the propagators as in Eqs. (20) and (23) as bilinear forms in these fluctuations. These equations are valid only if the initial values of the fluctuations  $a_{\pm k\lambda a}$  are plane waves with on-shell momenta; one can check easily that this is true for the free propagators. Thus Eq. (54) must give plane wave solutions for the field fluctuations in light-cone gauge when  $x^- < 0$ . This is simply achieved by taking plane waves for the fluctuation  $\tilde{a}^\mu$  in the original gauge and setting the function  $f$  to zero.<sup>15</sup> Therefore, the requirement that Eqs. (20) and (23) be valid leaves no residual gauge freedom.

<sup>15</sup>We note that it is also possible to choose  $\tilde{a}^\mu$ 's that are not plane waves, and a nonzero  $f$  to achieve our requirement that  $a^\mu$  be a plane wave. This, however, makes the intermediate calculations more tedious.

We only need to know  $\omega$  on our initial surface  $\Sigma$ —at  $x^- = \epsilon$ . Because the components of  $\Omega$  and of  $\tilde{a}$  are all of order unity, it is legitimate to neglect the values of  $z^-$  that are between 0 and  $\epsilon$  in the integration in Eq. (54). For  $x^- = \epsilon$  and  $z^- < 0$ , the incomplete Wilson line is equal to the complete Wilson line (which has the lower bound at  $-\infty$ ). We therefore obtain

$$\omega(x^- = \epsilon) = \Omega(\mathbf{x}_\perp) \int_{-\infty}^0 dz^- \tilde{a}^+(z^-, x^+, \mathbf{x}_\perp). \quad (56)$$

Note also that when  $\epsilon \leq x^-$ , the Wilson line becomes independent of  $x^-$  because all the color sources are in the strip  $0 \leq x^- \leq \epsilon$ . This explains why we only indicate  $\mathbf{x}_\perp$  in its list of arguments.

Once  $\omega$  has been determined, the  $-$  and  $i$  components of the fluctuation in light-cone gauge are determined from those in the  $A^- = 0$  gauge to be

$$a^- = \Omega^\dagger (-\partial^- \omega) \Omega, \quad a^i = \Omega^\dagger (\tilde{a}^i - \partial^i \omega) \Omega. \quad (57)$$

As we shall see shortly when we discuss the leading logarithmic divergences, the only quantity we need is<sup>16</sup>

$$\begin{aligned} \partial_\mu [\Omega a^\mu \Omega^\dagger] &= \partial_\mu [\tilde{a}^\mu - \partial^\mu \omega - ig[\omega, \tilde{\mathcal{A}}^\mu]] \\ &= -\partial^+ \partial^- \omega - \partial^i [\tilde{a}^i - \partial^i \omega], \end{aligned} \quad (58)$$

where we have used Eq. (53) and the fact that  $\tilde{a}^- = 0$  in order to eliminate a few terms. Using the equation for  $\partial^+ \omega$ , as well as the fact that  $\tilde{\mathcal{A}}^+$  is zero at  $x^- = \epsilon$ , we get

$$\partial_\mu [\Omega a^\mu \Omega^\dagger] = \partial_\perp^2 \omega - \partial^- \tilde{a}^+ - \partial^i \tilde{a}^i. \quad (59)$$

Let us now consider specifically the fluctuations  $a_{\pm k\lambda a}$ . In the gauge  $\tilde{a}^- = 0$ , their expression below the light-cone reads<sup>17</sup>

$$\tilde{a}_{\pm k\lambda a}^\mu(x) = \tilde{\epsilon}_\lambda^\mu(\mathbf{k}) T^a e^{\pm ik \cdot x}, \quad (60)$$

with

$$\begin{aligned} \tilde{\epsilon}_\lambda^-(\mathbf{k}) &= 0, \quad \sum_{\lambda=1,2} \tilde{\epsilon}_\lambda^i(\mathbf{k}) \tilde{\epsilon}_\lambda^j(\mathbf{k}) = -g^{ij}, \\ \tilde{\epsilon}_\lambda^+(\mathbf{k}) &= \frac{\mathbf{k}_\perp \cdot \tilde{\epsilon}_{\lambda\perp}(\mathbf{k})}{k^-}. \end{aligned} \quad (61)$$

The formulas that govern the light-cone crossing in this gauge have been worked out in [54]. Using these results, one finds the following expressions for the fluctuation fields just above the light cone:

<sup>16</sup>Note that  $(\Omega \mathcal{A} \Omega^\dagger)_c = \Omega_{cb} \mathcal{A}_b$ , from the definition of the adjoint representation. With the notation where  $\mathcal{A}$  is a column vector, this quantity would also be denoted by  $(\Omega \mathcal{A})_c$ .

<sup>17</sup>Therefore,  $\tilde{a}_{\pm k\lambda a}^{\mu b}(x) = \tilde{\epsilon}_\lambda^{\mu b}(\mathbf{k}) \delta^{ab} e^{\pm ik \cdot x}$ .

$$\begin{aligned}\tilde{a}_{\pm k\lambda a}^{ib}(x) &= \Omega_{ba}(\mathbf{x}_\perp)\tilde{\epsilon}_\lambda^i(\mathbf{k})e^{\pm ik\cdot x}, \\ \tilde{a}_{\pm k\lambda a}^{\pm b}(x) &= [\Omega_{ba}(\mathbf{x}_\perp)\tilde{\epsilon}_\lambda^\pm(\mathbf{k}) \pm (\partial^i\Omega_{ba}(\mathbf{x}_\perp))] \\ &\quad \times \frac{1}{ik^-}\tilde{\epsilon}_\lambda^i(\mathbf{k})e^{\pm ik\cdot x}.\end{aligned}\quad (62)$$

Note that for these field fluctuations, one has

$$\partial^- \tilde{a}_{\pm k\lambda a}^+ = \partial^i \tilde{a}_{\pm k\lambda a}^i. \quad (63)$$

Thus we have

$$\partial_\mu[\Omega a_{\pm k\lambda a}^\mu \Omega^\dagger] = \partial^i[\partial^i \omega - 2\tilde{a}_{\pm k\lambda a}^i]. \quad (64)$$

Substituting Eq. (60) in Eq. (56) gives the following expression for  $\omega$  just above the light cone,

$$\omega_b = \mp 2i\Omega_{ba}\frac{k^j}{k_\perp^2}\tilde{\epsilon}_\lambda^j(\mathbf{k})e^{\pm ik\cdot x}. \quad (65)$$

Therefore,

$$\partial_\mu[\Omega a_{\pm k\lambda a}^\mu \Omega^\dagger]_b = -2\partial^i[e^{\pm ik\cdot x}\alpha_{\pm ka}^{ilb}\epsilon_\lambda^l(\mathbf{k})], \quad (66)$$

where we have introduced the shorthand notation

$$\begin{aligned}\epsilon_\lambda^l(\mathbf{k}) &\equiv \left(\delta^{lm} - 2\frac{k^l k^m}{k_\perp^2}\right)\tilde{\epsilon}_\lambda^m(\mathbf{k}), \\ \alpha_{\pm ka}^{ilb} &\equiv \left(\delta^{il} - \frac{k^i k^l}{k_\perp^2}\right)\Omega_{ba} \mp i\frac{k^l}{k_\perp^2}\partial^i\Omega_{ba}.\end{aligned}\quad (67)$$

#### D. Logarithmic divergences

Let us recall that our objective is to isolate the leading logarithmic contributions to Eq. (40). From the structure of this equation, isolating these contributions requires that we examine Eq. (42) term by term. As we shall see later, the contribution in  $\beta \cdot \mathbb{T}$  (“virtual correction”) can be derived from the term bilinear in  $\mathbb{T}$  (“real correction”). Therefore, let us concentrate on the bilinear term for now.

To determine the leading logarithmic contributions in the real correction, we need to consider the integration over the on-shell momentum  $k^\mu$  as well. It involves an integral

$$\int_0^{+\infty} \frac{dk^+}{k^+}, \quad (68)$$

which potentially leads to logarithmic singularities both at  $k^+ \rightarrow 0$  and at  $k^+ \rightarrow +\infty$ . Note that wherever  $k^-$  appears in the integrand, it should be replaced by the on-shell value  $k^- = k_\perp^2/2k^+$ . Inspecting the integrand of Eq. (40), one sees that the  $k^+$  dependence contains exponential factors

$$e^{i(k_\perp^2/2k^+)(v^+ - u^+)}. \quad (69)$$

There is no factor depending on  $v^- - u^-$ , because the points  $u$  and  $v$  are both on the initial surface  $\Sigma$ , and thus have equal  $-$  coordinates. It is clear the integral converges at  $k^+ \rightarrow 0^+$  thanks to the oscillatory behavior of this exponential. On the other hand, when  $k^+ \rightarrow +\infty$ , the exponential goes to unity and one may have a logarithmic singularity there. However, to truly have a divergence, the other factors in the integrand should not have any power of  $1/k^+$ .

Let us now examine these. The coefficients in the operator  $a \cdot \mathbb{T}_u$  are the initial values of  $\Omega a^-$ ,  $\partial^-(\Omega a^i)$  and  $\partial_\mu(\Omega a^\mu)$ . We need only to keep the coefficients that have no power of  $1/k^+$ . One sees readily that this is not the case for  $\Omega a^-$  or  $\partial^-(\Omega a^i)$ : these two quantities [compare Eq. (57) to Eqs. (58) and (59)] contain a factor  $k^- \sim 1/k^+$ .

Thus, as previously anticipated, the only divergence arises when one picks up the term  $\partial_\mu(\Omega a^\mu)$  both in  $a \cdot \mathbb{T}_u$  and  $a \cdot \mathbb{T}_v$ .

In order to regularize the integral over  $k^+$ , we must introduce an upper bound  $\Lambda^+$ . Physically, this cutoff is related to the division of degrees of freedom one operates with in the CGC: the color sources describe the fast partons and thus correspond to modes  $k^+ > \Lambda^+$ , while the fields represent the slow degrees of freedom that have a longitudinal momentum  $k^+ < \Lambda^+$ . Therefore, when one performs a calculation in this effective description, the longitudinal momentum of all the fields and field fluctuations should not exceed  $\Lambda^+$ , in order not to overcount modes that are already represented as part of the color sources  $\rho$ . The lower scale in this logarithm is of the order of the longitudinal momentum  $p^+$  of the produced gluon. Therefore, the logarithm resulting from the  $k^+$  integration is a logarithm of  $\Lambda^+/p^+$ .

The determination of the lower bound in the  $k^+$  integral is the place where it is crucial to consider an observable that depends only on a single scale of longitudinal momentum (or rapidity)—in order to avoid an ambiguity in the value of this lower bound. Obviously, this condition is satisfied by the single inclusive gluon spectrum. But as one can see in [27], this restriction limits the generalization of this result to the multigluon spectrum to the case where all the measured gluons have similar rapidities.

To pick up the logarithm, one should approximate the exponential by unity. This implies that the coefficient of the logarithm is independent of  $u^+$  and  $v^+$  or, in other words, it is invariant under boosts in the  $+z$  direction. As we shall see, such perturbations of  $\partial_\mu(\Omega a^\mu)$  can be mapped to a change in the color source  $\tilde{\rho}$ , and these logarithms can be absorbed in a redefinition of the distribution  $W[\tilde{\rho}]$ .

#### E. Real corrections

Keeping only the term in  $\partial_\mu(\Omega a^\mu)$  in Eq. (42), and limiting ourselves to the divergent part of the real correction for now, we see that we must evaluate the operator

$$\begin{aligned} & \frac{1}{2\pi} \ln\left(\frac{\Lambda^+}{p^+}\right) \int \frac{d^2\mathbf{k}_\perp}{(2\pi)^2} \int d^2\mathbf{u}_\perp d^2\mathbf{v}_\perp \sum_a \partial^i (\alpha_{-ka}^{ilb}(\mathbf{u}) e^{ik_\perp \cdot \mathbf{u}_\perp}) \\ & \quad \times \partial^j (\alpha_{+ka}^{jlc}(\mathbf{v}_\perp) e^{-ik_\perp \cdot \mathbf{v}_\perp}) \int du^+ dv^+ \\ & \quad \times \frac{\delta^2}{\delta \partial_\mu \Omega(u)_{bd} \mathcal{A}_d^\mu(u^+, \mathbf{u}_\perp) \delta \partial_\mu \Omega(v)_{ce} \mathcal{A}_e^\mu(v^+, \mathbf{v}_\perp)}. \end{aligned} \quad (70)$$

Here, to avoid any confusion, we have written explicitly all the color indices. Note also that we have performed the sum over the two polarization states of the field fluctuation in this expression.<sup>18</sup>

The object on which this operator acts is the observable calculated at leading order, considered as a functional of the initial value of the fields  $\mathcal{A}^i$  in light-cone gauge. In this gauge, the initial values of  $\mathcal{A}^+$  and  $\mathcal{A}^-$  are zero (provided the residual gauge freedom is fixed as explained in Sec. III B). Moreover, from the setup of the problem, it turns out that these initial fields do not depend on  $x^+$ ,

$$\mathcal{A}^i(x^+, \mathbf{x}_\perp) = \mathcal{A}^i(\mathbf{x}_\perp), \quad (71)$$

and

$$\partial_\mu \Omega(u)_{bd} \mathcal{A}_d^\mu(u^+, \mathbf{u}_\perp) = -\partial^i \Omega(\mathbf{u}_\perp)_{bd} \mathcal{A}_d^i(\mathbf{u}_\perp). \quad (72)$$

When we restrict ourselves to functionals that depend only on  $x^+$ -independent initial fields, we can simply write<sup>19</sup>

$$\begin{aligned} & \int du^+ \frac{\delta}{\delta \partial_\mu \Omega(u)_{bd} \mathcal{A}_d^\mu(u^+, \mathbf{u}_\perp)} \\ & = -\frac{\delta}{\delta \partial^i \Omega(\mathbf{u}_\perp)_{bd} \mathcal{A}_d^i(\mathbf{u}_\perp)}. \end{aligned} \quad (73)$$

Our goal now is to relate the leading logarithmic contribution we have identified to the JIMWLK evolution of the distribution of color sources. As we have seen in the previous sections, the initial value of the field in light-cone gauge has a simple expression when expressed in terms of the sources  $\tilde{\rho}$  or fields  $\tilde{\mathcal{A}}^+$  in Lorenz gauge. Therefore, we will try to make the connection with the JIMWLK equation in this gauge. To do this, we must relate the functional derivative  $\delta/\delta \partial^i \Omega(\mathbf{u}_\perp)_{bd} \mathcal{A}_d^i(\mathbf{u}_\perp)$  to the functional deriva-

tive  $\delta/\delta \tilde{\mathcal{A}}^+$ . We begin by considering the light-cone gauge expression for the classical transverse gauge fields given by Eqs. (46) and (47). Rewriting  $\mathcal{A}^i(\mathbf{x}_\perp)$  more explicitly as

$$\begin{aligned} \mathcal{A}^i(x^-, \mathbf{x}_\perp) & = -\int_{-\infty}^{x^-} dz^- \Omega^\dagger(z^-, \mathbf{x}_\perp) \\ & \quad \times (\partial^i \tilde{\mathcal{A}}^+(z^-, \mathbf{x}_\perp)) \Omega(z^-, \mathbf{x}_\perp), \end{aligned} \quad (74)$$

one observes that a variation<sup>20</sup>  $\delta \tilde{\mathcal{A}}^+(\epsilon, \mathbf{x}_\perp)$  of the field in covariant gauge in the last  $x^-$  bin (of width  $dx^-$ ) leads to a change  $\delta \mathcal{A}^i(\mathbf{x}_\perp)$  of the initial value of the gauge field in light-cone gauge, given by

$$\delta \mathcal{A}^i(\mathbf{x}_\perp) = -\Omega^\dagger(\mathbf{x}_\perp) (\partial^i \delta \tilde{\mathcal{A}}^+(\epsilon, \mathbf{x}_\perp) dx^-) \Omega(\mathbf{x}_\perp). \quad (75)$$

From this formula, we get the variation of  $\partial^i \Omega(\mathbf{u}_\perp)_{bd} \mathcal{A}_d^i(\mathbf{u}_\perp)$ ,

$$\delta[\partial^i \Omega(\mathbf{u}_\perp)_{bd} \mathcal{A}_d^i(\mathbf{u}_\perp)] = -\partial_\perp^2 \delta \tilde{\mathcal{A}}^+(\epsilon, \mathbf{x}_\perp) dx^-. \quad (76)$$

Inverting this relation, one obtains

$$\begin{aligned} \frac{\delta}{\delta \partial^i \Omega(\mathbf{u}_\perp)_{bd} \mathcal{A}_d^i(\mathbf{u}_\perp)} & = -\int d^2\mathbf{x}_\perp G(\mathbf{u}_\perp - \mathbf{x}_\perp) \\ & \quad \times \frac{\delta}{\delta \tilde{\mathcal{A}}_b^+(\epsilon_Y, \mathbf{x}_\perp)}. \end{aligned} \quad (77)$$

Here  $G(\mathbf{u}_\perp - \mathbf{x}_\perp)$  is a two-dimensional propagator whose main properties are discussed in Appendix C.

It is important to observe that the functional derivatives on the left-hand side (l.h.s.) and right-hand side of this equation do not have the same dimensions. This is because they are defined with respect to fields that live in spaces with different dimensions. On the left-hand side, the initial transverse field in light-cone gauge does not depend on  $x^-$  as soon as we are outside the nucleus and is therefore a function of  $\mathbf{u}_\perp$  only. On the right-hand side, the field  $\tilde{\mathcal{A}}^+$  depends crucially on  $x^-$ . The  $\epsilon_Y$  argument in the right-hand side of Eq. (77) is not integrated over, and should be chosen as the value of  $x^-$  where the last layer of quantum evolution has produced its partons. This is the same as the location  $\epsilon$  of the surface  $\Sigma$  used for the initial conditions, but the subscript  $Y$  indicates that it may shift as the rapidity  $Y$  increases.

We can now rewrite the operator in Eq. (70) as follows:

$$\begin{aligned} & \frac{1}{2} \ln\left(\frac{\Lambda^+}{p^+}\right) \int d^2\mathbf{x}_\perp d^2\mathbf{y}_\perp \eta^{bc}(\mathbf{x}_\perp, \mathbf{y}_\perp) \\ & \quad \times \frac{\delta^2}{\delta \tilde{\mathcal{A}}_b^+(\epsilon_Y, \mathbf{x}_\perp) \delta \tilde{\mathcal{A}}_c^+(\epsilon_Y, \mathbf{y}_\perp)}, \end{aligned} \quad (78)$$

<sup>18</sup>A useful identity is

$$\left(\delta^{il} - 2\frac{k^i k^l}{k_\perp^2}\right) \left(\delta^{lj} - 2\frac{k^l k^j}{k_\perp^2}\right) = \delta^{ij}.$$

<sup>19</sup>It is useful to recall that the dimension of a functional derivative operator is  $\text{Mass}^{-d(\mathcal{A})-D}$  where  $d(\mathcal{A})$  is the mass dimension of the field with respect to which one is differentiating, and  $D$  the mass dimension of the space in which this field lives. For instance,

$$\frac{\delta}{\delta \mathcal{A}_b^i(u^+, \mathbf{u}_\perp)} \sim \text{Mass}^2, \quad \frac{\delta}{\delta \mathcal{A}_b^i(\mathbf{u}_\perp)} \sim \text{Mass}^1.$$

<sup>20</sup>It is natural that the size of the bin in which the field  $\tilde{\mathcal{A}}^+$  is changed plays a role here. Indeed, because  $\tilde{\mathcal{A}}^+$  is integrated over  $x^-$  in the expression of  $\mathcal{A}^i$ , a change in a bin of zero width produces no change in  $\mathcal{A}^i$ . Note also that the factor  $dx^-$  in Eq. (75) is necessary on dimensional grounds.

where we have defined<sup>21</sup>

$$\begin{aligned} \eta^{bc}(\mathbf{x}_\perp, \mathbf{y}_\perp) &\equiv \frac{1}{4\pi^3} \int \frac{d^2\mathbf{k}_\perp}{(2\pi)^2} \int d^2\mathbf{u}_\perp d^2\mathbf{v}_\perp \\ &\times \sum_a \alpha_{-ka}^{ilb}(\mathbf{u}_\perp) \alpha_{+ka}^{jlc}(\mathbf{v}_\perp) e^{ik_\perp \cdot (\mathbf{u}_\perp - \mathbf{v}_\perp)} \\ &\times \frac{\mathbf{u}_\perp^i - \mathbf{x}_\perp^i}{(\mathbf{u}_\perp - \mathbf{x}_\perp)^2} \frac{\mathbf{v}_\perp^j - \mathbf{y}_\perp^j}{(\mathbf{v}_\perp - \mathbf{y}_\perp)^2}. \end{aligned} \quad (79)$$

From Eq. (67),  $\alpha_{\pm ka}^{ilb}$  can naturally be broken in two terms. If we keep only the first term in each of the  $\alpha$ 's in Eq. (79), we obtain correspondingly

$$\begin{aligned} \eta_{(1)}^{bc}(\mathbf{x}_\perp, \mathbf{y}_\perp) &= -\frac{1}{8\pi^4} \int d^2\mathbf{u}_\perp d^2\mathbf{v}_\perp \frac{(\mathbf{x}_\perp^i - \mathbf{u}_\perp^i)(\mathbf{y}_\perp^j - \mathbf{v}_\perp^j)}{(\mathbf{x}_\perp - \mathbf{u}_\perp)^2 (\mathbf{y}_\perp - \mathbf{v}_\perp)^2} \\ &\times \Delta^{ij}(\mathbf{u}_\perp - \mathbf{v}_\perp) \\ &\times [\Omega(u)\Omega^\dagger(v) - 1]_{bc}. \end{aligned} \quad (80)$$

Here the function  $\Delta^{ij}$  is defined in Eq. (C11) of Appendix C. When we keep the first term in the first  $\alpha$  and the second term in the second  $\alpha$  (or vice versa), we get zero because the two terms in  $\alpha$  are mutually orthogonal. If we keep the second term in each of the  $\alpha$ 's, we obtain

$$\begin{aligned} \eta_{(2)}^{bc}(\mathbf{x}_\perp, \mathbf{y}_\perp) &= \frac{1}{\pi} \int \frac{d^2\mathbf{u}_\perp}{(2\pi)^2} \frac{(\mathbf{x}_\perp^i - \mathbf{u}_\perp^i)(\mathbf{y}_\perp^j - \mathbf{u}_\perp^j)}{(\mathbf{x}_\perp - \mathbf{u}_\perp)^2 (\mathbf{y}_\perp - \mathbf{u}_\perp)^2} \\ &\times [\Omega(x)\Omega^\dagger(y) - \Omega(x)\Omega^\dagger(u) \\ &- \Omega(u)\Omega^\dagger(y) + 1]_{bc} + \frac{1}{8\pi^4} \int d^2\mathbf{u}_\perp d^2\mathbf{v}_\perp \\ &\times \frac{(\mathbf{x}_\perp^i - \mathbf{u}_\perp^i)(\mathbf{y}_\perp^j - \mathbf{v}_\perp^j)}{(\mathbf{x}_\perp - \mathbf{u}_\perp)^2 (\mathbf{y}_\perp - \mathbf{v}_\perp)^2} \Delta^{ij}(\mathbf{u}_\perp - \mathbf{v}_\perp) \\ &\times [\Omega(u)\Omega^\dagger(v) - 1]_{bc}. \end{aligned} \quad (81)$$

When we add the two contributions, the terms involving  $\Delta^{ij}$  cancel, and we are finally left with

$$\begin{aligned} \eta^{bc}(\mathbf{x}_\perp, \mathbf{y}_\perp) &= \frac{1}{\pi} \int \frac{d^2\mathbf{u}_\perp}{(2\pi)^2} \frac{(\mathbf{x}_\perp^i - \mathbf{u}_\perp^i)(\mathbf{y}_\perp^j - \mathbf{u}_\perp^j)}{(\mathbf{x}_\perp - \mathbf{u}_\perp)^2 (\mathbf{y}_\perp - \mathbf{u}_\perp)^2} \\ &\times [\Omega(x)\Omega^\dagger(y) - \Omega(x)\Omega^\dagger(u) \\ &- \Omega(u)\Omega^\dagger(y) + 1]_{bc}. \end{aligned} \quad (82)$$

This function is precisely the function  $\eta^{bc}(\mathbf{x}_\perp, \mathbf{y}_\perp)$  that appears in the JIMWLK equation [17, 19].

At this point, a word must be said of the term  $\Delta\mathcal{O}_{\text{NLO}}$  in Eq. (40). It is given by the difference of two terms that can be obtained from each other by exchanging  $a_{+k\lambda a}$  and  $a_{-k\lambda a}$ . Going back to the calculation of  $\eta^{bc}(\mathbf{x}_\perp, \mathbf{y}_\perp)$ , it is easy to check that for the calculation of the leading log term these two terms give the same result and cancel. Physically this is due to charge conjugation symmetry—

<sup>21</sup>We performed along the way an integration by parts and used the identity in Eq. (C8).

because the classical field is real, we obtain the same result by exchanging the negative and positive energy asymptotic solutions for the quantum fluctuation;  $\Delta\mathcal{O}_{\text{NLO}}$  is the difference between these two and thus cancels out.

## F. Virtual corrections

In the previous subsection, we focused on the real contribution to Eq. (40). We now turn our attention to the term in  $\beta \cdot \mathbb{T}_u$  in Eq. (40). Recall that  $\beta^\mu$  is the 1-loop correction to the classical field in the LC gauge and is evaluated in Eq. (40) at  $u^- = \epsilon$ , just above the region occupied by the nuclear sources. Mimicking the evaluation of the real contribution, we can write directly<sup>22</sup>

$$\begin{aligned} &\int_{u^- = \epsilon} du^+ d^2\mathbf{u}_\perp [\beta \cdot \mathbb{T}_u] \\ &= \int d^2\mathbf{x}_\perp \int d^2\mathbf{u}_\perp G(\mathbf{x}_\perp - \mathbf{u}_\perp) \partial_\mu^\mu (\Omega(u)_{bd} \beta_d^\mu(u)) \\ &\quad \underbrace{\ln\left(\frac{\Lambda^+}{p^+}\right) \nu^b(\mathbf{x}_\perp)}_{\ln\left(\frac{\Lambda^+}{p^+}\right) \nu^b(\mathbf{x}_\perp)} \\ &\times \frac{\delta}{\delta \bar{\mathcal{A}}_b^+(\epsilon_Y, \mathbf{x}_\perp)}. \end{aligned} \quad (83)$$

We anticipate that a large logarithm in the  $k^+$  integral will show up in this quantity, and we have defined  $\nu^b(\mathbf{x}_\perp)$  as its coefficient. Note that in this definition of the function  $\nu^b(\mathbf{x}_\perp)$ , the value<sup>23</sup> of  $u^-$  is  $u^- = \epsilon$ .

We begin with Green's formula for the 1-point function  $\beta^\mu(u)$ , where the initial surface is taken at  $v^- = 0$  (instead of  $v^- = \epsilon$ ),

$$\begin{aligned} \beta^\mu(u) &= \int_{v^- > 0} d^4v D_{0,R}^{\mu\nu}(u, v) \left[ \frac{\partial^2 U(\mathcal{A})}{\partial \mathcal{A}^\nu(v) \partial \mathcal{A}^\rho(v)} \beta^\rho(v) \right. \\ &\left. + \frac{1}{2} \frac{\partial^3 U(\mathcal{A})}{\partial \mathcal{A}^\nu(v) \partial \mathcal{A}^\rho(v) \partial \mathcal{A}^\sigma(v)} \mathcal{G}_{++}^{\rho\sigma}(v, v) \right]. \end{aligned} \quad (84)$$

By this choice of the initial surface, we do not have a boundary term, because  $\beta^\mu$  is zero at  $u^- \leq 0$ . The propagator  $\mathcal{G}_{++}^{\rho\sigma}(v, v)$  can be expressed in terms of the field fluctuations  $a_{\pm k\lambda a}$  by using Eq. (23). Consider now the Green's formula for the fluctuation  $a_{+k\lambda a}$  we introduced in Eq. (27), but written this time for an initial surface at  $u^- = 0$ ,

$$\begin{aligned} a_{+k\lambda a}^\mu(x) &= \int_{y^- > 0} d^4y D_{0,R}^{\mu\nu}(x, y) \frac{\partial^2 U(\mathcal{A})}{\partial \mathcal{A}^\nu(y) \partial \mathcal{A}^\rho(y)} a_{+k\lambda a}^\rho(y) \\ &+ \mathcal{B}_0^\mu[a_{+k\lambda a}]. \end{aligned} \quad (85)$$

<sup>22</sup>One can confirm that  $(\Omega)_{bd} \beta_d^-$  and  $\partial^-(\Omega)_{bd} \beta_d^j$  are zero and therefore cannot appear in the operator  $[\beta \cdot \mathbb{T}_u]$ .

<sup>23</sup>The value of  $u^+$  is irrelevant because the 1-point function  $\beta^\mu(u)$  propagating over an  $x^+$ -independent background field (and with a vanishing initial condition in the past) is independent of  $u^+$ .

In this formula, both the fluctuation  $a_{+k\lambda a}$  and the derivative of the gauge potential depend on the background classical field in LC gauge. Let us apply to this equation the operator<sup>24</sup>  $[a_{-k\lambda a} \cdot \mathcal{T}]$  that substitutes one power of the background field by a power of  $a_{-k\lambda a}$ . By defining

$$\xi^\mu(u) \equiv \frac{1}{2} \sum_{\lambda,a} \int \frac{d^3\mathbf{k}}{(2\pi)^3 2E_k} \times \int_{v^- > 0} d^4v [a_{-k\lambda a} \cdot \mathcal{T}_v] a_{+k\lambda a}^\mu(u), \quad (86)$$

we obtain for this object Green's formula

$$\xi^\mu(u) = \int_{v^- > 0} d^4v D_{0,R}^{\mu\nu}(u, v) \left[ \frac{\partial^2 U(\mathcal{A})}{\partial \mathcal{A}^\nu(v) \partial \mathcal{A}^\rho(v)} \xi^\rho(v) + \frac{1}{2} \frac{\partial^3 U(\mathcal{A})}{\partial \mathcal{A}^\nu(v) \partial \mathcal{A}^\rho(v) \partial \mathcal{A}^\sigma(v)} \mathcal{G}_{++}^{\rho\sigma}(v, v) \right], \quad (87)$$

where we used Eq. (23) for the propagator that appears in the source term. We see that  $\xi^\mu$  and  $\beta^\mu$  are identical. Therefore, we have proved that

$$\beta_d^\mu(u) \equiv \frac{1}{2} \sum_{\lambda,a} \int \frac{d^3\mathbf{k}}{(2\pi)^3 2E_k} \times \int_{v^- > 0} d^4v [a_{-k\lambda a} \cdot \mathcal{T}_v] a_{+k\lambda a}^{\mu d}(u). \quad (88)$$

Inserting this expression into the definition of  $\nu^b(x_\perp)$ , we obtain

$$\ln\left(\frac{\Lambda^+}{p^+}\right) \nu^b(x_\perp) = \frac{1}{2} \sum_{\lambda,a} \int \frac{d^3\mathbf{k}}{(2\pi)^3 2E_k} \times \int_{v^- > 0} d^4v [a_{-k\lambda a} \cdot \mathcal{T}_v] \times \int d^2\mathbf{u}_\perp G(x_\perp - \mathbf{u}_\perp) \times \partial_\mu^u (\Omega(u)_{bd} a_{+k\lambda a}^{\mu d}(u)). \quad (89)$$

To obtain a divergence at large  $k^+$ , we need to tame the oscillations in this variable which exist because we now

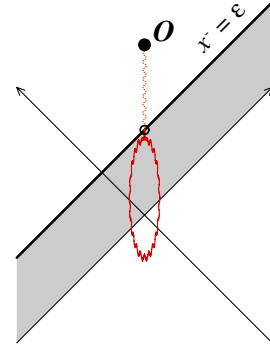


FIG. 4 (color online). Leading logarithmic contribution of the tadpole diagram.

have  $u^- = \epsilon$  while  $v^-$  can be anywhere in the range  $[0, \epsilon]$ . These oscillations are damped only if  $v^-$  is in the immediate vicinity of  $u^- = \epsilon$ . As a corollary, note that the left diagram in Fig. 3 is therefore a bit misleading because the tadpole contribution depicted vanishes when the upper vertex of the tadpole is below the light cone. In fact, to have a leading logarithmic contribution, this vertex of the tadpole must be very close to the surface  $u^- = \epsilon$ , as illustrated in Fig. 4.

For sufficiently small  $dx^-$ , we can use

$$\lim_{dx^- \rightarrow 0} \int_{\epsilon - dx^-}^{\epsilon} dv^- [a_{-k\lambda a} \cdot \mathcal{T}_v] = a_{-k\lambda a} \cdot \mathbb{T}_v, \quad (90)$$

namely, we recover the operator that substitutes the background field by the fluctuation in the last layer at  $v^- = \epsilon$ . Again, using Eqs. (73) and (77) from the previous subsection, we obtain the operator

$$\int_{v^- = \epsilon} dv^+ d^2\mathbf{v}_\perp [a_{-k\lambda a} \cdot \mathbb{T}_v] = \int_{\text{LLog}} d^2\mathbf{y}_\perp \int d^2\mathbf{v}_\perp G(\mathbf{y}_\perp - \mathbf{v}_\perp) \times \partial_\nu^v (\Omega(v)_{ce} a_{-k\lambda a}^{ve}(v)) \frac{\delta}{\delta \tilde{\mathcal{A}}_c^+(\epsilon_Y, \mathbf{y}_\perp)}. \quad (91)$$

When inserted in Eq. (89), this gives

$$\ln\left(\frac{\Lambda^+}{p^+}\right) \nu^b(x_\perp) = \frac{1}{2} \int d^2\mathbf{y}_\perp \sum_{\lambda,a} \int \frac{d^3\mathbf{k}}{(2\pi)^3 2E_k} \int d^2\mathbf{v}_\perp G(\mathbf{y}_\perp - \mathbf{v}_\perp) \partial_\nu^v (\Omega(v)_{ce} a_{-k\lambda a}^{ve}(v)) \times \frac{\delta}{\delta \tilde{\mathcal{A}}_c^+(\epsilon_Y, \mathbf{y}_\perp)} \int d^2\mathbf{u}_\perp G(x_\perp - \mathbf{u}_\perp) \partial_\mu^u (\Omega(u)_{bd} a_{+k\lambda a}^{\mu d}(u)). \quad (92)$$

Note that the products of the underlined terms, by themselves, are just

<sup>24</sup>This operator is similar to the operator  $a \cdot \mathbb{T}$  previously defined, but it performs the replacement of fields inside the region of the sources, instead of just on the surface of this region.

$$\ln\left(\frac{\Lambda^+}{p^+}\right)\eta^{bc}(\mathbf{x}_\perp, \mathbf{y}_\perp). \quad (93)$$

The final step in our derivation is to note that when  $\tilde{\mathcal{A}}$  shares a color index with  $\Omega$ , we have the identity [17,19,55]

$$\frac{\delta}{\delta\tilde{\mathcal{A}}_c^+(\epsilon_Y, \mathbf{y}_\perp)}\partial_\nu^v(\Omega(v)_{ce}a_{-k\lambda a}^{ve}(v)) = 0, \quad (94)$$

because of the antisymmetry of the adjoint generators of  $SU(N)$ . We can therefore move the operator  $\delta/\delta\tilde{\mathcal{A}}_c^+(\epsilon_Y, \mathbf{y}_\perp)$  immediately after the measure  $d^2\mathbf{y}_\perp$  to obtain

$$\nu^b(\mathbf{x}_\perp) = \frac{1}{2} \int d^2\mathbf{y}_\perp \frac{\delta}{\delta\tilde{\mathcal{A}}_c^+(\epsilon_Y, \mathbf{y}_\perp)} \eta^{bc}(\mathbf{x}_\perp, \mathbf{y}_\perp), \quad (95)$$

which is identical to the relation between  $\eta^{bc}$  and  $\nu^b$  in the JIMWLK equation.

### G. JIMWLK equation

We shall now combine the real and virtual corrections to write the JIMWLK equation. Using the real correction in Eq. (78) and the virtual one given by Eqs. (83) and (95), we can write the total NLO correction, Eq. (40), in the form

$$\mathcal{O}_{\text{NLO}} = \ln\left(\frac{\Lambda^+}{p^+}\right)\mathcal{H}\mathcal{O}_{\text{LO}}[\tilde{\mathcal{A}}^+] \quad (96)$$

where we have introduced the JIMWLK Hamiltonian,

$$\begin{aligned} \mathcal{H} \equiv & \frac{1}{2} \int d^2\mathbf{x}_\perp d^2\mathbf{y}_\perp \frac{\delta}{\delta\tilde{\mathcal{A}}_c^+(\epsilon_Y, \mathbf{y}_\perp)} \eta^{bc}(\mathbf{x}_\perp, \mathbf{y}_\perp) \\ & \times \frac{\delta}{\delta\tilde{\mathcal{A}}_b^+(\epsilon_Y, \mathbf{x}_\perp)}. \end{aligned} \quad (97)$$

Although the coupling does not appear explicitly in the Hamiltonian, it is of order  $\alpha_s$  because of the presence of two functional derivatives with respect to classical fields that are of order  $g^{-1}$ .

We noted that the observable  $\mathcal{O}$  at leading order can be expressed as a functional of the classical gauge field  $\tilde{\mathcal{A}}^+$  in covariant gauge. The average of this observable over all the configurations of the field  $\tilde{\mathcal{A}}^+$ , up to NLO, can be expressed as

$$\langle \mathcal{O}_{\text{LO}} + \mathcal{O}_{\text{NLO}} \rangle \equiv \int [D\tilde{\mathcal{A}}^+] W[\tilde{\mathcal{A}}^+] [\mathcal{O}_{\text{LO}} + \mathcal{O}_{\text{NLO}}]. \quad (98)$$

At the leading logarithmic level, this can be rewritten as

$$\begin{aligned} \langle \mathcal{O}_{\text{LO}} + \mathcal{O}_{\text{NLO}} \rangle = & \int [D\tilde{\mathcal{A}}^+] \{ [1 + \Delta Y \mathcal{H}] W[\tilde{\mathcal{A}}^+] \} \\ & \times \mathcal{O}_{\text{LO}}[\tilde{\mathcal{A}}^+], \end{aligned} \quad (99)$$

where we denote  $\Delta Y \equiv \ln(\Lambda^+/p^+)$ . Note that  $\Delta Y$  is also the rapidity interval between the slowest incoming sources

(that have  $k^+ \sim \Lambda^+$ ) and the measured gluon. To obtain this equation, one uses the Hermiticity of  $\mathcal{H}$  with respect to the functional integration over  $\tilde{\mathcal{A}}^+$ . In writing this equation, we have absorbed all the leading logarithms of  $k^+$  into a redefinition of the distribution  $W[\tilde{\mathcal{A}}^+]$ ,

$$W[\tilde{\mathcal{A}}^+] \rightarrow [1 + \Delta Y \mathcal{H}] W[\tilde{\mathcal{A}}^+]. \quad (100)$$

This suggests that the distribution  $W[\tilde{\mathcal{A}}^+]$  should depend on the scale  $\Lambda^+$  that separates the modes described as static sources from the modes described as dynamical fields in the CGC description. Of course, this is not surprising in an effective theory based on such a separation of the degrees of freedom. For this reason, it should be denoted as  $W_{\Lambda^+}[\tilde{\mathcal{A}}^+]$ . Therefore Eq. (99) can be written as

$$\begin{aligned} \langle \mathcal{O}_{\text{LO}} + \mathcal{O}_{\text{NLO}} \rangle = & \int [D\tilde{\mathcal{A}}^+] \{ [1 + \ln\left(\frac{\Lambda^+}{p^+}\right)\mathcal{H}] \\ & \times W_{\Lambda^+}[\tilde{\mathcal{A}}^+] \} \mathcal{O}_{\text{LO}}[\tilde{\mathcal{A}}^+]. \end{aligned} \quad (101)$$

Because  $\Lambda^+$  is an unphysical separation scale, the expectation value of observables should not depend on this parameter. Differentiating the previous equation with respect to  $\Lambda^+$  and requiring that the r.h.s. be zero, we get<sup>25</sup>

$$\frac{\partial}{\partial \ln(\Lambda^+)} W_{\Lambda^+}[\tilde{\mathcal{A}}^+] = -\mathcal{H} W_{\Lambda^+}[\tilde{\mathcal{A}}^+]. \quad (102)$$

Equivalently, if  $Y \equiv \ln(P^+/\Lambda^+)$  denotes the rapidity separation between the fragmentation region of the nucleus (located at  $k^+ \sim P^+$ ) and the rapidity down to which partons are described as static color sources, we have

$$\frac{\partial}{\partial Y} W_Y[\tilde{\mathcal{A}}^+] = \mathcal{H} W_Y[\tilde{\mathcal{A}}^+], \quad (103)$$

which is the JIMWLK equation that drives the  $Y$  dependence of the distribution  $W_Y[\tilde{\mathcal{A}}^+]$ .

The above considerations also indicate that the distribution  $W_Y[\tilde{\mathcal{A}}^+]$  must be evolved to a scale  $\Lambda^+$  comparable to the typical longitudinal momentum in the observable to avoid large residual logs contributing to the latter. Therefore, at leading logarithmic accuracy, the expectation value of the observable is given by

$$\langle \mathcal{O} \rangle_{\text{LLog}} = \int [D\tilde{\mathcal{A}}^+] W_Y[\tilde{\mathcal{A}}^+] \mathcal{O}_{\text{LO}}[\tilde{\mathcal{A}}^+], \quad (104)$$

with  $Y = \ln(P^+/p^+)$  the rapidity separation between the beam and the observable and  $W_Y[\tilde{\mathcal{A}}^+]$  given by the solution of Eq. (103).

<sup>25</sup>To avoid confusion, recall that  $\mathcal{H}$ , and hence  $\partial W/\partial \Lambda^+$ , is of order  $\alpha_s$ . Therefore, for consistency, one should not keep the term proportional to  $\mathcal{H}(\partial W/\partial \Lambda^+)$  because it is of order  $\alpha_s^2$  and therefore beyond the accuracy of the present calculation.

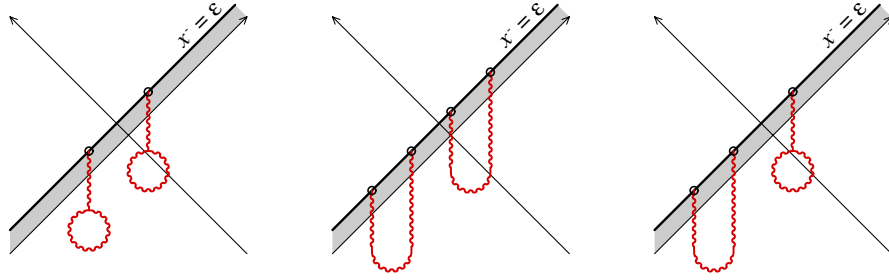


FIG. 5 (color online). The 2-loop contributions made of products of pieces already encountered at 1 loop. Although we do not make this distinction in the figure, one of the factors is attached at a slightly smaller value of  $x^-$ , because the two Hamiltonians in Eq. (110) are at different rapidities.

### H. All-order resummation of leading logs

Thus far, we only considered 1-loop corrections that generate one power of the large logarithm of  $P^+$ . On this basis, we deduced an evolution equation for  $W[\tilde{\mathcal{A}}^+]$  using renormalization group arguments. However, the solution of the RG equation is equivalent to a resummation of all  $n$ -loop diagrams that have  $n$  powers of large logarithms of  $p^+$ . We shall analyze here the structure of higher loop contributions to confirm whether the all-loop resummation performed by the RG equation is justified.

Here we will not perform a detailed analysis of these leading  $n$ -loop graphs to show that we indeed recover the solution of Eq. (103). More modestly, we will work *a posteriori* by examining the solution of the JIMWLK equation to see what the  $n$ -loop graphs that it resums are. Before proceeding, it is useful to recall a crucial property of the JIMWLK Hamiltonian defined in Eq. (97). The operator  $\mathcal{H}$  contains derivatives with respect to the field  $\tilde{\mathcal{A}}^+(\epsilon_Y, \mathbf{x}_\perp)$ , and its coefficients depend on all the fields  $\tilde{\mathcal{A}}^+(x^-, \mathbf{x}_\perp)$  for  $0 \leq x^- \leq \epsilon_Y$ . For this reason, we will denote it  $\mathcal{H}(Y)$ , where the endpoint  $\epsilon_Y$  at which the derivatives act is related to  $Y$  by  $Y \sim \ln(\epsilon_Y)$ . It is important to note that in a product  $\mathcal{H}(y_1)\mathcal{H}(y_2)$ , the derivatives in  $\mathcal{H}(y_1)$  do not act on the coefficients of  $\mathcal{H}(y_2)$  if  $y_1 > y_2$ .

The JIMWLK equation should now be written as

$$\frac{\partial}{\partial Y} W_Y[\tilde{\mathcal{A}}^+] = \mathcal{H}(Y)W_Y[\tilde{\mathcal{A}}^+], \quad (105)$$

and its solution reads

$$W_Y[\tilde{\mathcal{A}}^+] = \mathcal{U}(Y)W_0[\tilde{\mathcal{A}}^+], \quad (106)$$

with

$$\mathcal{U}(Y) \equiv T_Y \left[ \exp \int_0^Y dy \mathcal{H}(y) \right]. \quad (107)$$

In this equation,  $T_Y$  denotes a ‘‘rapidity ordering’’ such that products of  $\mathcal{H}$ ’s in the Taylor expansion of the exponential are ordered from left to right in order of decreasing  $y$ .  $W_0[\tilde{\mathcal{A}}^+]$  is a nonperturbative initial condition.  $\mathcal{U}(Y)$  is the evolution operator for the Hamiltonian  $\mathcal{H}(Y)$ . Inserting Eq. (106) into Eq. (104), we get

$$\langle \mathcal{O} \rangle_{\text{LLLog}} = \int [D\tilde{\mathcal{A}}^+] W_0[\tilde{\mathcal{A}}^+] \mathcal{U}^\dagger(Y) \mathcal{O}_{\text{LO}}[\tilde{\mathcal{A}}^+]. \quad (108)$$

Because  $\mathcal{H}(y)$  is Hermitian, the Hermitian conjugate of the evolution operator  $\mathcal{U}(Y)$  is the same operator with the rapidity ordering reversed:

$$\mathcal{U}^\dagger(Y) \equiv \bar{T}_Y \left[ \exp \int_0^Y dy \mathcal{H}(y) \right], \quad (109)$$

where  $\bar{T}_Y$  denotes the antirapidity ordering. The expansion of  $\mathcal{U}^\dagger$  to order one in  $\mathcal{H}$  gives the leading logarithmic 1-loop contributions that we have evaluated earlier in this section. [See Eq. (99), for instance.]

If one expands it to second order, we see that the leading logarithmic contributions in the observable at 2 loops should be given by

$$\mathcal{O}_{\text{NNLO}}^{\text{LLLog}} = \int_0^Y dy_1 \int_0^{y_1} dy_2 \mathcal{H}(y_2) \mathcal{H}(y_1) \mathcal{O}_{\text{LO}}[\tilde{\mathcal{A}}^+]. \quad (110)$$

Because  $y_2 < y_1$ , the derivatives in  $\mathcal{H}(y_2)$  can act on the coefficients  $\eta$  and  $\nu$  of  $\mathcal{H}(y_1)$ . Let us first consider the terms where this does not happen, namely, where the derivatives in  $\mathcal{H}(y_2)$  act directly on  $\mathcal{O}_{\text{LO}}[\tilde{\mathcal{A}}^+]$ . These terms correspond to the graphs depicted in Fig. 5. If we look only at what happens below the line  $x^- = \epsilon$ , these contributions are just disconnected products of terms we had already at 1 loop. The analysis we performed of the logarithmic contributions at 1 loop extends trivially to these terms, and it is easy to see that they have two powers of the logarithms.

In addition, Eq. (110) also contains terms in which at least one of the derivatives in  $\mathcal{H}(y_2)$  acts on the coefficients of  $\mathcal{H}(y_1)$ . This corresponds to topologies of the type displayed in Fig. 6. Such terms, that have a gluon vertex inside the region where the sources live, have a large logarithm for the same reason that the tadpole has a logarithm in the 1-loop terms. Thus one can see that it is crucial to properly order the powers of the Hamiltonian  $\mathcal{H}$  in rapidity to avoid losing these terms.<sup>26</sup>

<sup>26</sup>For instance, if the ordering of the two Hamiltonians in Eq. (110) is reversed, we get only the terms of Fig. 5.

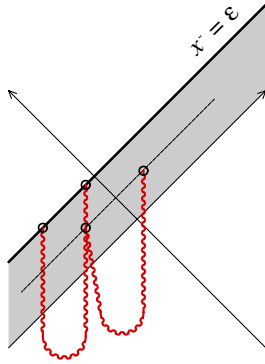


FIG. 6 (color online). Example of the term obtained when the derivatives in  $\mathcal{H}(y_2)$  can act on the coefficients of  $\mathcal{H}(y_1)$ . Here, one of the derivatives in  $\mathcal{H}(y_2)$  acts of the function  $\eta$  of  $\mathcal{H}(y_1)$ , and the second derivative in  $\mathcal{H}(y_2)$  acts on  $\mathcal{O}_{\text{LO}}$ .

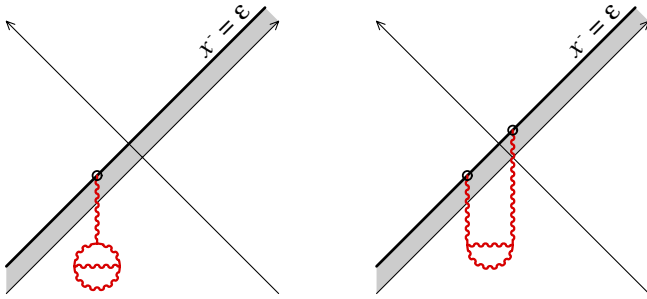


FIG. 7 (color online). Some of the 2-loop corrections to the observable  $\mathcal{O}$  that do not appear at leading log.

Finally, there also exist at 2 loops some topologies that never appear in Eq. (110), such as those of Fig. 7. The contributions in this figure are 1-loop corrections to the coefficients of the operators  $\mathbb{T}_{u,v}$  in Eq. (40). In other words, these terms generate corrections of order  $\alpha_s$  to the coefficients in the JIMWLK equation, and do not have double logs of  $\Lambda^+$ . This explains why they are not generated by the leading log formula in Eq. (110).

#### IV. NUCLEUS-NUCLEUS COLLISIONS

In the previous section, we obtained an expression for resummed leading logarithmic inclusive gluon observables in a single nucleus in terms of the equivalent leading order observable. Along the way, we presented a novel derivation of the JIMWLK evolution equation. In this section, we will extend our analysis to the case of nuclear collisions. We will show that the leading logarithms of  $k^\pm$  that arise in the calculation of loop corrections to the single inclusive gluon spectrum can be factored out in the distributions  $W[\tilde{\mathcal{A}}_1^+]$  and  $W[\tilde{\mathcal{A}}_2^-]$  that describe the two incoming nuclei. This result will complete a proof of factorization of leading logarithms of  $1/x_{1,2}$  for inclusive observables in nucleus-nucleus collisions.

#### A. Inclusive observables at leading order

As in the single nucleus case, our discussion is valid for an inclusive multigluon operator  $\mathcal{O}$  (corresponding to a moment of the multiplicity or energy distribution produced in nucleus-nucleus collisions), but for simplicity, we will focus on the first moment of the multiplicity distribution—the inclusive gluon spectrum. As we discussed in [24,26], the inclusive single particle spectrum in nucleus-nucleus collisions can be expressed as

$$E_p \frac{dN}{d^3\mathbf{p}} = \frac{1}{16\pi^3} \lim_{x_0 \rightarrow +\infty} \int d^3\mathbf{x} d^3\mathbf{y} e^{ip \cdot (x-y)} (\partial_x^0 - iE_p) \times (\partial_y^0 + iE_p) \sum_\lambda \epsilon_\lambda^\mu(\mathbf{p}) \epsilon_\lambda^\nu(\mathbf{p}) \langle A_\mu(x) A_\nu(y) \rangle. \quad (111)$$

Unsurprisingly, the operator  $\langle A_\mu(x) A_\nu(y) \rangle$  is identical to what we considered previously in the single nucleus case. In particular, at leading order, the single gluon spectrum is evaluated by replacing the two gauge operators in the right-hand side of the previous equation by classical solutions of the Yang-Mills equations. These classical solutions are obtained by imposing retarded boundary conditions that vanish in the remote past. The only difference with the previous section and with Eqs. (13) is that the current  $J^\nu$  that drives the solutions of the Yang-Mills equations is now comprised of two contributions corresponding to each of the nuclei. This is a significant complication in that, unlike the single nucleus case, analytical solutions do not exist. However, the classical fields and the inclusive spectrum have been computed numerically [56–64].

Formally, the single inclusive gluon spectrum at leading order is a functional of the LC gauge fields  $\mathcal{A}_{1,2}$  of the two nuclei on the surfaces  $x^- = \epsilon$  and  $x^+ = \epsilon$ , respectively, or of the covariant gauge fields  $\tilde{\mathcal{A}}_{1,2}^\pm$  in the strips  $0 \leq x^- < \epsilon$  and  $0 \leq x^+ < \epsilon$  (see Fig. 8),

$$E_p \frac{dN}{d^3\mathbf{p}} \Big|_{\text{LO}} \equiv \mathcal{O}_{\text{LO}}[\mathcal{A}_1, \mathcal{A}_2] \equiv \mathcal{O}_{\text{LO}}[\tilde{\mathcal{A}}_1^+, \tilde{\mathcal{A}}_2^-]. \quad (112)$$

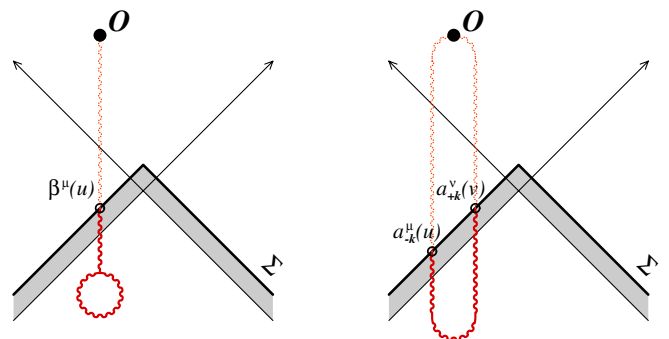


FIG. 8 (color online). NLO corrections in the collision of two nuclei. The thick solid line is the initial surface where the functions  $\beta^\mu$  and  $a_{\pm k\lambda a}$  are evaluated. The precise shape of the small portion of this surface located above the forward light cone is not important because its contribution is power suppressed.



This quantity does not depend on the rapidity  $y \sim \ln(p^+/p^-)$  because of the boost invariance of the classical equations of motion [65–67].

### B. One-loop corrections

At 1 loop, Eq. (16) can be used again to compute the inclusive spectrum. The manipulations in Secs. II C and II D were not specific to the case of a single nucleus. Indeed, we did not specify the detailed content of the current  $J^\mu$  in Sec. II. The only requirement for the validity of the final formula is that one chooses an initial surface  $\Sigma$  which is locally spacelike (or lightlike at worst).

We can now exploit this freedom in the choice of  $\Sigma$  in order to take a surface that treats the two nuclei on the same footing. A convenient choice is a surface  $\Sigma$  with the two branches

$$\begin{aligned} x^- &= \epsilon, & x^+ &< \epsilon, \\ x^+ &= \epsilon, & x^- &< \epsilon, \end{aligned} \quad (113)$$

as illustrated by the thick solid line in Fig. 8. We shall denote the measure on this initial surface as  $d\Sigma_u$ . It is simply  $du^+ d^2\mathbf{u}_\perp$  on the first branch and  $du^- d^2\mathbf{u}_\perp$  on the second branch. Similarly, the definition of the operator  $[a \cdot \mathbb{T}_u]$  depends on the branch on which it is evaluated, because Green's formula for the classical fields depends on a different set of initial field components on the two branches.<sup>27</sup> It is also important to note that the functional derivatives with respect to the initial gauge fields are derivatives with respect to the field  $\mathcal{A}_1$  of the first nucleus on the first branch, and likewise the field  $\mathcal{A}_2$  of the second nucleus on the second branch.

We also need to say a few words about the gauge in which the initial fields on  $\Sigma$  are expressed. On the left branch of  $\Sigma$  (i.e. on the branch  $u^- = \epsilon$ ), we use the  $A^+ = 0$  gauge, while we use the  $A^- = 0$  gauge on the other branch. Using different gauge conditions on these two branches is possible because they are not causally connected. Similarly, for the propagation of the small fluctuations  $a_{\pm k\lambda a}$  and  $\beta$ , we use the  $A^+ = 0$  gauge if their endpoint is on the left branch of  $\Sigma$ , and the  $A^- = 0$  gauge if it is on the other side.

Modulo these obvious changes, Eq. (40) is valid in the case of two nuclei, and we can now express it as

$$\begin{aligned} \mathcal{O}_{\text{NLO}} &= \left[ \int_{\Sigma} d\Sigma_u [\beta \cdot \mathbb{T}_u] \right] + \frac{1}{2} \sum_{\lambda, a} \int \frac{d^3\mathbf{k}}{(2\pi)^3 2E_k} \\ &\times \int_{\Sigma} d\Sigma_u d\Sigma_v [a_{-k\lambda a} \cdot \mathbb{T}_u] [a_{+k\lambda a} \cdot \mathbb{T}_v] \\ &\times \mathcal{O}_{\text{LO}}[\mathcal{A}_1, \mathcal{A}_2] + \Delta \mathcal{O}_{\text{NLO}}. \end{aligned} \quad (114)$$

The first two terms in this formula are illustrated in Fig. 8.

<sup>27</sup>This result is evident from the derivation of Green's formula in LC gauge discussed at length in Appendix B.

As in the case of a single nucleus, the leading logs will cancel in  $\Delta \mathcal{O}_{\text{NLO}}$  because of the charge conjugation symmetry discussed previously.

The leading log piece of the term involving  $[\beta \cdot \mathbb{T}_u]$  can be mapped into the corresponding term of the JIMWLK equation in the same way as in the case of a single nucleus. Depending on whether we are on the first or second branch of the initial surface  $\Sigma$ , we get two terms which can be expressed together as

$$\begin{aligned} &\left[ \ln\left(\frac{\Lambda^+}{p^+}\right) \int d^2\mathbf{x}_\perp \nu_1^b(\mathbf{x}_\perp) \frac{\delta}{\delta \tilde{\mathcal{A}}_{1,b}^+(\epsilon_Y, \mathbf{x}_\perp)} + \ln\left(\frac{\Lambda^-}{p^-}\right) \right. \\ &\times \left. \int d^2\mathbf{x}_\perp \nu_2^b(\mathbf{x}_\perp) \frac{\delta}{\delta \tilde{\mathcal{A}}_{2,b}^-(\epsilon_Y, \mathbf{x}_\perp)} \right] \mathcal{O}_{\text{LO}}[\tilde{\mathcal{A}}_1^+, \tilde{\mathcal{A}}_2^-], \end{aligned} \quad (115)$$

where  $\nu_{1,2}^b(\mathbf{x}_\perp)$  are, respectively, the one-point functions from the JIMWLK Hamiltonian for the two nuclei, and likewise,  $\tilde{\mathcal{A}}_1^+, \tilde{\mathcal{A}}_2^-$  are classical gauge fields in Lorenz gauge of the first and second nuclei, respectively. We have also introduced a cutoff  $\Lambda^-$  that separates the color sources of the second nucleus from the dynamical fields.

There is a subtlety in generalizing the single nucleus derivation to obtain this result. In Eq. (73), the integration over  $u^+$  runs from  $-\infty$  to  $+\infty$ . Now, because of the choice of the surface  $\Sigma$ , this integration runs only from  $-\infty$  to 0, and we must justify that this difference is irrelevant. To simplify the notations in this argument, let us use the shorthand  $\alpha(u^+, \mathbf{u}_\perp) \equiv \partial_\mu(\Omega(u) \mathcal{A}^\mu(u^+, \mathbf{u}_\perp))$ . In our problem, the functional derivative with respect to  $\alpha(u^+, \mathbf{u}_\perp)$  is only applied to functionals that depend solely on the  $u^+$ -independent mode of  $\alpha(u^+, \mathbf{u}_\perp)$ ,

$$\alpha(\mathbf{u}_\perp) \equiv \frac{1}{L} \int du^+ \alpha(u^+, \mathbf{u}_\perp), \quad (116)$$

where  $L$  is the length of the  $u^+$  interval.<sup>28</sup> When this is the case, we have

$$\frac{\delta}{\delta \alpha(u^+, \mathbf{u}_\perp)} F[\alpha(\mathbf{u}_\perp)] = \frac{1}{L} \frac{\delta}{\delta \alpha(\mathbf{u}_\perp)} F[\alpha(\mathbf{u}_\perp)]. \quad (117)$$

Moreover, the result of this differentiation does not depend on the value of  $u^+$  in the l.h.s. Therefore, the subsequent integration over  $u^+$  merely generates a factor  $L$  equal to the length of the integration range. We have therefore proven that

$$\int du^+ \frac{\delta}{\delta \alpha(u^+, \mathbf{u}_\perp)} F[\alpha(\mathbf{u}_\perp)] = \frac{\delta}{\delta \alpha(\mathbf{u}_\perp)} F[\alpha(\mathbf{u}_\perp)], \quad (118)$$

regardless of the integration range for the variable  $u^+$ .

<sup>28</sup>Since here this interval is semi-infinite, it is best to consider  $u^+ \in [-L, 0]$  in all the intermediate steps, and to take  $L \rightarrow \infty$  only at the end.

Another possible concern is whether there is a contribution to  $[\beta \cdot \mathbb{T}_u]$  from the small portion of the initial surface  $\Sigma$  that lies above the forward light cone in the region where *both*  $u^\pm$  are positive. It is easy to convince oneself that the contribution from this region does not lead to stronger singularities than the rest of the initial surface. Furthermore, contributions from this region are phase-space suppressed due to its small size of order  $\epsilon$ .

The leading log contribution of the terms of Eq. (114) that are bilinear in  $[a \cdot \mathbb{T}]$  is equally simple when the two points  $u$  and  $v$  belong to the same branch of the initial surface  $\Sigma$ . If this is so, it is straightforward to reproduce what we did for a single nucleus, and we find the two separate contributions

$$\begin{aligned} & \left[ \ln\left(\frac{\Lambda^+}{p^+}\right) \int d^2x_\perp d^2y_\perp \eta_1^{bc}(x_\perp, y_\perp) \right. \\ & \times \frac{\delta^2}{\delta \tilde{\mathcal{A}}_{1,b}^+(\epsilon_Y, \mathbf{x}_\perp) \delta \tilde{\mathcal{A}}_{1,c}^+(\epsilon_Y, \mathbf{y}_\perp)} + \ln\left(\frac{\Lambda^-}{p^-}\right) \\ & \times \int d^2x_\perp d^2y_\perp \eta_2^{bc}(x_\perp, y_\perp) \\ & \left. \times \frac{\delta^2}{\delta \tilde{\mathcal{A}}_{2,b}^-(\epsilon_Y, \mathbf{x}_\perp) \delta \tilde{\mathcal{A}}_{2,c}^-(\epsilon_Y, \mathbf{y}_\perp)} \right] \mathcal{O}_{\text{LO}}[\tilde{\mathcal{A}}_{1,2}^\pm]. \quad (119) \end{aligned}$$

Summing Eqs. (115) and (119), and expressing  $\nu$  in terms of  $\eta$ , we obtain the leading log 1-loop expression for the single inclusive gluon spectrum to be

$$\mathcal{O}_{\text{NLO}} =_{\text{LLog}} \left[ \ln\left(\frac{\Lambda^+}{p^+}\right) \mathcal{H}_1 + \ln\left(\frac{\Lambda^-}{p^-}\right) \mathcal{H}_2 \right] \mathcal{O}_{\text{LO}}[\tilde{\mathcal{A}}_1^+, \tilde{\mathcal{A}}_2^-], \quad (120)$$

where  $\mathcal{H}_{1,2}$  are the JIMWLK Hamiltonians of the first and second nuclei, respectively. This equation—assuming we can prove that there are no other terms at leading log—is the generalization of Eq. (96) to the case of the collision of two nuclei. In the next subsection, we will demonstrate that indeed there are no other contributions.

### C. Absence of precollision mixings

Thus far, we did not discuss the contribution to the bilinear  $[a \cdot \mathbb{T}]$  terms where the coordinates  $u$  and  $v$  belong to different branches of the initial surface. This contribution is illustrated in Fig. 9. If it contains leading log contributions, such a term would spoil Eq. (120), because it would generate a term that mixes derivatives with respect to  $\tilde{\mathcal{A}}_1^+$  and  $\tilde{\mathcal{A}}_2^-$ , thereby precluding any possibility of factorization.

Fortunately, this possibility is not realized because terms where  $u$  and  $v$  are on different branches contain the phases

$$e^{ik^+(u^- - v^-)} e^{ik^-(u^+ - v^+)} \quad (121)$$

in the integral over  $d^3k$ . For generic points  $u$  and  $v$  in this configuration, neither  $u^- - v^-$  nor  $u^+ - v^+$  are vanishing

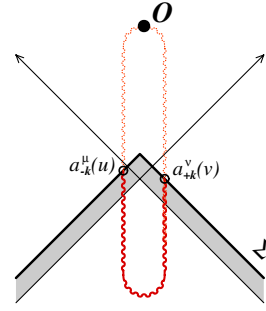


FIG. 9 (color online). Contribution that mixes the two nuclei and may lead to a violation of factorization.

and these exponentials oscillate rapidly when either  $k^+ \rightarrow +\infty$  or  $k^- \rightarrow +\infty$ . Therefore, the integral over  $k^+$  (or  $k^-$ ) is completely finite, and we do not get a large logarithm from this configuration of  $u$ 's and  $v$ 's.

The only potential danger might come from the configuration where  $u$  or  $v$  (or both) lies in the small portion of  $\Sigma$  above the tip of the light cone. Again, such a configuration can, at most, produce a logarithmic singularity, but is suppressed by a small phase-space prefactor of order  $\epsilon$  due to the small size of this region. Therefore, Eq. (120) contains all the leading log terms that show up in the 1-loop corrections to the single inclusive gluon spectrum.

### D. Factorization

Finally, integrating over all the configurations of the nuclear fields  $\tilde{\mathcal{A}}_{1,2}^\pm$  with weights  $W[\tilde{\mathcal{A}}_1^+]$  and  $W[\tilde{\mathcal{A}}_2^-]$ , and using the fact that the JIMWLK Hamiltonian is Hermitian, we can write the sum of the LO and NLO (leading logs only) of the single inclusive gluon spectrum as

$$\begin{aligned} \langle \mathcal{O}_{\text{LO}} + \mathcal{O}_{\text{NLO}} \rangle =_{\text{LLog}} & \int [D\tilde{\mathcal{A}}_1^+] [D\tilde{\mathcal{A}}_2^-] \\ & \times \{ [1 + \Delta Y_1 \mathcal{H}_1] W[\tilde{\mathcal{A}}_1^+] \} \\ & \times \{ [1 + \Delta Y_2 \mathcal{H}_2] W[\tilde{\mathcal{A}}_2^-] \} \\ & \times \mathcal{O}_{\text{LO}}[\tilde{\mathcal{A}}_1^+, \tilde{\mathcal{A}}_2^-]. \quad (122) \end{aligned}$$

In this equation, we denote  $\Delta Y_1 \equiv \ln(\Lambda_1^+/p^+)$  and  $\Delta Y_2 \equiv \ln(\Lambda_2^-/p^-)$ , where  $\Lambda^+$  is the cutoff in the CGC description of the first nucleus,  $\Lambda^-$  of the second nucleus, and  $p^\pm$  the longitudinal momentum components of the produced gluon. We can now choose the (arbitrary) cutoffs as  $\Lambda^\pm = p^\pm$  and express, as anticipated in Eq. (10), the leading log part of the NLO result in terms of the LO operator convoluted with the appropriately evolved weight functions as

$$\begin{aligned} \langle \mathcal{O} \rangle_{\text{LLog}} = & \int [D\tilde{\mathcal{A}}_1^+] \\ & \times [D\tilde{\mathcal{A}}_2^-] W_{Y_1}[\tilde{\mathcal{A}}_1^+] W_{Y_2}[\tilde{\mathcal{A}}_2^-] \mathcal{O}_{\text{LO}}[\tilde{\mathcal{A}}_1^+, \tilde{\mathcal{A}}_2^-], \quad (123) \end{aligned}$$

where each of the  $W[\tilde{\mathcal{A}}^\pm]$ 's obeys the JIMWLK equation (possibly with different initial conditions if the two nuclei are not identical) and  $Y_1 = \ln(P_1^+/p^+)$  and  $Y_2 = \ln(P_2^-/p^-)$ .

## V. HIGH ENERGY FACTORIZATION RESULT IN CONTEXT

It is useful to consider our result in Eq. (123) in the context of related work in the high energy limit. Factorization, in the specific sense of our work, was proven previously for proton-nucleus collisions in the large  $N_c$  limit of dipole scattering off a large nucleus [68–71]. In the case of nucleus-nucleus collisions, there has been recent work by Braun, computing single and double inclusive gluon production in a Reggeon field theory approach [72]. At present, it is unclear how to relate these results to the JIMWLK evolution. A first attempt at establishing such a dictionary between cut disconnected diagrams in the CGC effective theory and cut Pomerons was discussed in Ref. [23]; see also Refs. [73,74].

It is important to note that the factorization theorem proven here is valid only for inclusive quantities such as moments of the multiplicity or energy distributions. In fact, it seems unlikely that these results will extend to discussions of total cross sections and exclusive final states [75–77]. Indeed, it is known [24,50] that the retarded nature of the boundary conditions for the fields and field fluctuations has a close connection with the inclusiveness of an observable, and we have seen in the present paper that the retarded nature of these objects plays an essential role in our proof of factorization. Whether the Pomeron loops that may play a role in those computations are suppressed for the observables we consider is also unclear. Our results certainly suggest that these contributions are not important for inclusive moments in nucleus-nucleus collisions, provided the densities  $\rho_{1,2}$  of color sources are large.<sup>29</sup>

Another important trend in the literature is computing next-to-leading order contributions to high energy evolution. In the Reggeon field theory approach [78], we note the very significant work on multi-Regge factorization at NLO by Fadin and collaborators [79] which builds on the extension of the Balitsky-Fadin-Kuraev-Lipatov equation to NLO [80–82]. In the CGC effective theory, there has been significant recent work to include running coupling corrections [31,33–37] culminating in the recent NLO extension [32] of the Balitsky-Kovchegov equation. As our result is valid for JIMWLK factorization at leading log, these NLO results will be useful in attempts to extend our

proof of high energy factorization to next-to-leading logarithmic accuracy.

Finally, we should emphasize that JIMWLK factorization proven here is far more general than the  $k_\perp$  factorization often discussed in the literature. The latter is a high energy factorization in terms of unintegrated  $k_\perp$ -dependent parton distributions [83–85] and can be obtained in the low density limit of JIMWLK factorization [23,86,87]. In order to see this explicitly, one should take the limit of low gluon density in the observable  $\mathcal{O}_{\text{LO}}[\tilde{\mathcal{A}}_1^+, \tilde{\mathcal{A}}_2^-]$  that enters in the integrand of our factorized formula (123). Formally, this amounts to letting the source densities  $\rho_{1,2}$  tend to zero and keeping only the first nonzero contribution. In the particular case where the observable is the single inclusive gluon spectrum, the lowest order term is of order  $\rho_1^2 \rho_2^2$ . The integral of  $\rho_1^2$  over  $\tilde{\mathcal{A}}_1^+$  weighted by the distribution  $W_{Y_1}[\tilde{\mathcal{A}}_1^+]$  gives the nonintegrated gluon distribution of the first nucleus, evolved to the rapidity of the produced gluon—and likewise for the second nucleus. If one also takes the limit of low density in the JIMWLK equation that drives the evolution of the distributions  $W_{Y_{1,2}}$ , these nonintegrated gluon distributions will obey the Balitsky-Fadin-Kuraev-Lipatov equation, and one will automatically recover the usual  $k_\perp$ -factorization result for single gluon production in the low density regime.  $k_\perp$  factorization also holds for single inclusive gluon production at leading order in proton-nucleus collisions [68,86,88–92]—i.e. when one takes the low density limit only for one of the two nuclei. From this way of recovering the known  $k_\perp$ -factorized results as limits of our Eq. (123), it should be obvious that JIMWLK factorization is far more general and robust than  $k_\perp$  factorization, since it contains the latter and remains valid in the description of the collision of two dense projectiles.

$k_\perp$  factorization was, however, shown to be broken explicitly for quark pair production even at leading order [87], albeit it is restored [23] for large momenta  $k_\perp \gg Q_s$ . Likewise, this breaking of factorization is also seen for gluon pair production [69,93]. In a subsequent work [27], we have proven that the JIMWLK factorization formula (123) is also valid in the case of the multigluon inclusive spectra, provided all the measured gluons are in a small window in rapidity. Some work is under way in order to extend this result to the general case of gluons located at arbitrary rapidities.

To a large extent, factorization in hadronic collisions is merely a consequence of causality: two fast projectiles cannot interact before they collide. Thus the objects that describe their content must be universal—independent of the other projectile, and of the observable that one is going to measure *after* the collision. However, this general argument does not tell us what information should be included in the objects describing the projectiles; indeed, this depends on the observable under consideration, and on whether we are in the saturation regime or not. In the

<sup>29</sup>If  $\rho_{1,2}$  are not of order  $g^{-1}$ , then the power counting on which our considerations are based may be modified. Since it has been argued that Pomeron loops play a role in the dilute regime, this leaves open the possibility that these effects may alter our conclusions close to the fragmentation region of the projectiles.

saturated regime, a given observable will generally be produced via the coherent interaction of many partons of the projectiles, which means that one will need to know the probability of these multiparton configurations in the wave function of the projectiles. In contrast, in the dilute regime, since only one parton of each projectile interacts, one needs only to know the probabilities for 1-parton configurations. This is why JIMWLK factorization is more general than  $k_\perp$  factorization: the distribution  $W[\rho]$  contains enough information<sup>30</sup> to calculate the nonintegrated gluon distribution, but the converse is certainly not true.<sup>31</sup> Similar considerations suggest that JIMWLK factorization may not work in the case of exclusive observables. Indeed, inclusive observables usually require less detailed information about the projectiles than exclusive ones.<sup>32</sup>

The factorization theorem that we have proved here is a necessary first step before a full NLO computation of gluon production in the Glasma. Equation (123) includes only the NLO terms that are enhanced by a large logarithm of  $1/x_{1,2}$ , while the complete NLO calculation would also include the nonenhanced terms. This would be of the same order in  $\alpha_s$  as the production of quark-antiquark pairs [94,95] from the classical field. Note that, to be really useful, this complete NLO calculation would probably have to be promoted to a next-to-leading log result by resumming all the terms in  $\alpha_s(\alpha_s \ln(1/x_{1,2}))^n$ . Now that evolution equations in the dense regime are becoming available at NLO, work in this direction is a promising prospect.

## VI. FACTORIZATION, THE GLASMA, AND THERMALIZATION

The Glasma is the nonequilibrium hot and dense matter formed immediately in the aftermath of a high energy heavy ion collision [25,46,96]. How this matter thermalizes is of great importance for a quantitative understanding of the phenomenology of heavy ion collisions.<sup>33</sup> We will discuss here the relevance of our factorization theorem, present qualitative ideas about its generalization, and discuss their importance in quantifying the properties of the Glasma.

<sup>30</sup>It provides information about multiparton correlations such as  $\langle \rho(x_1)\rho(x_2)\cdots\rho(x_n) \rangle$ .

<sup>31</sup>Nonintegrated gluon distributions depend only on 2-parton correlations  $\langle \rho(x_1)\rho(x_2) \rangle$ .

<sup>32</sup>For instance, in order to study single diffractive processes, one would need “conditional” probabilities of multiparton configurations, where one imposes the condition that no parton has been radiated between the rapidity of the projectile and the rapidity where the gap ends. This information is not provided by the distributions  $W[\rho]$  that are the basis of JIMWLK factorization.

<sup>33</sup>Another important aspect is how jets propagate inside this matter, in order to assess issues such as leading parton quenching in jets.

At leading order, the Glasma is described by the solution of the Yang-Mills equations in the forward light cone with retarded boundary conditions (given by the classical fields of the two nuclei before the collision). The produced fields have large occupation numbers of order  $\alpha_s^{-1}$  and are boost invariant [65,66]. This boost invariance of fields implies that the classical dynamics can be described by the proper time evolution of gauge fields that live in the transverse plane. An interesting consequence of the classical field dynamics is that the chromo-electric and magnetic fields are purely longitudinal after the collision [46,65], leading to the generation of Chern-Simons charge density in the collision [96]. The Glasma fields at this order generate only transverse pressure at proper times  $\tau \gtrsim Q_s^{-1}$ , so it seems impossible that a treatment of the Glasma at this order leads to thermalization.

This is where the small quantum fluctuations of the color field (of order 1, compared to the classical field of order  $g^{-1}$ ) become relevant. In an observable such as the inclusive gluon spectrum, these quantum fluctuations lead to corrections that are  $\alpha_s$  smaller than the leading order classical contribution. As we have discussed at length in the previous sections, some contributions of these small fluctuations—those that are enhanced by leading powers of  $\ln(1/x_{1,2})$ —can be resummed and absorbed into universal distributions  $W[\rho]$  that describe the high energy evolution of the nuclear wave functions.

But what about the remaining part of these small fluctuation terms, that are purely of order  $\alpha_s$  relative to the classical fields? Our resummation of leading logs corresponds to a well-controlled approximation provided the coefficients  $d_{ni}$  in the expansion of Eqs. (5) and (8) are truly numbers of order unity. Indeed, we have disregarded thus far the terms  $d_{ni}$  for  $i < n$ , on the basis that they do not have as many logs as powers of  $\alpha_s$ . However, numerical simulations of the classical Yang-Mills equations with initial conditions that break boost invariance show the existence of an instability of the rapidity-dependent fluctuations [97–99]. In these simulations, it is observed that the small rapidity-dependent perturbations superimposed to the boost invariant classical field grow exponentially with the square root of time as<sup>34</sup>

$$a^\mu \sim e^{\sqrt{\mu}\tau}, \quad (124)$$

where  $\mu$  is a quantity of the order of  $Q_s$  (its precise value depends on the wavelength of the fluctuation in the rapidity direction). This growth has variously been interpreted as either a Weibel-type [98,100] or Nielsen-Olesen-type [101,102] instability. The former mechanism, in particular, has been discussed extensively as a possible mechanism for thermalization in heavy ion collisions [103–110]. The

<sup>34</sup>The fact that the square root of the proper time, rather than the proper time itself, controls the growth of the instability is due to the longitudinal expansion of the system. This has also been observed analytically in the study of the Weibel instability [100].

existence of these unstable modes suggests that our assumption that the coefficients  $d_{ni}$  for  $i < n$  are of order unity is incorrect.

Our present understanding is that there are three classes among the small field fluctuations, which can be organized according to the momentum  $p_\eta$  they have in the  $\eta$  direction:

- (i) Zero modes ( $p_\eta = 0$ ) that generate a leading log. That the leading logs come solely from zero modes is obvious from the fact that the coefficients of the leading logs do not depend on  $x^\pm$ . These terms are already included in the resummation we have discussed at length in this paper.
- (ii) Zero modes that do not contribute at leading log because they have an extra power of  $k^-$  that prevents the divergence when  $k^+ \rightarrow \infty$  (see the discussion in Sec. III D). These terms have not been resummed in our scheme, and they do not seem to trigger the instability either. They would only become relevant in a full NLO calculation, and in resummation of next-to-leading log terms [32].
- (iii) Nonzero modes ( $p_\eta \neq 0$ ). These terms do not contribute large logarithms of  $1/x_{1,2}$ , but they are unstable and grow exponentially as  $\exp(\sqrt{\mu\tau})$ .

It is the latter boost noninvariant terms that are potentially dangerous. While also suppressed by a power of  $\alpha_s$ , they can be enhanced by exponentials of the proper time after the collision. Terms that diverge with time are called ‘‘secular divergences,’’ and some techniques for resumming these divergences are well known<sup>35</sup> in the literature [38].

Based on the above considerations, let us refine the expansion we wrote in Eqs. (5) and (8), in order to keep track also of powers of  $\exp(\sqrt{\mu\tau})$ . We should now write

$$\mathcal{O}[\rho_1, \rho_2] = \frac{1}{g^2} [c_0 + c_1 g^2 + c_2 g^4 + \dots], \quad (125)$$

with

$$c_n \equiv \sum_{p=i}^n \sum_{i=0}^p f_{npi} e^{(p-i)\sqrt{\mu\tau}} \ln^i\left(\frac{1}{x_{1,2}}\right). \quad (126)$$

In other words, the coefficients  $d_{ni}$  that we have introduced in Eq. (8), and assumed to be of order unity, are in fact

$$d_{ni} = \sum_{p=i}^n f_{npi} e^{(p-i)\sqrt{\mu\tau}}, \quad (127)$$

and can thus grow exponentially in time after the collision. In Eq. (126), the sum of the number of logs and of factors  $\exp(\sqrt{\mu\tau})$  (this sum is the index  $p$ ) cannot exceed  $n$  at  $n$  loops. This is because a fluctuation mode cannot be at the

same time a zero mode (required to generate a log) and a nonzero mode (required to generate an instability). In this new language, the leading log resummation that we have performed so far amounts to keeping only the term  $f_{nnn}$  in every  $c_n$ .

At first sight, one may expect a complete breakdown of the leading log description when the time

$$\tau_{\max} \sim Q_s^{-1} \ln^2\left(\frac{1}{\alpha_s}\right) \quad (128)$$

is reached. This is the time at which 1-loop corrections become as large as the LO contribution. This conclusion can be avoided if one can resum these divergent contributions leading to a resummed result that is better behaved for  $\tau \rightarrow +\infty$ . Indeed, it is possible to improve upon the leading log approximation, by keeping at every loop order all the terms where  $p = n$ : this corresponds to all the terms where every power of  $\alpha_s$  is accompanied by either a log or an  $\exp(\sqrt{\mu\tau})$ . Thus, let us define

$$\begin{aligned} \mathcal{O}_{\text{LLog+LInst}}[\rho_1, \rho_2] &\equiv \frac{1}{g^2} \\ &\times \sum_{n=0}^{\infty} g^{2n} \sum_{i=0}^n f_{nni} e^{(n-i)\sqrt{\mu\tau}} \ln^i\left(\frac{1}{x_{1,2}}\right). \end{aligned} \quad (129)$$

The subscript ‘‘LInst’’ means ‘‘leading instability.’’

In the formalism we have developed in this paper, the growth of small fluctuations with time can be traced to the action of the linear operator in Eq. (114) on the classical field. The quantity

$$\mathbb{T}_u \mathcal{A}(x) \sim \frac{\delta \mathcal{A}(x)}{\delta \mathcal{A}(u)} \sim e^{\sqrt{\mu\tau}} \quad (130)$$

is a measure of how sensitive the classical field  $\mathcal{A}(x)$  is to the initial condition at the point  $u$  on the initial surface. If there is an instability, small perturbations of the initial conditions lead to exponentially large deviations in the classical solutions. We will assume for now that the improved resummation defined in Eq. (129) can be performed and leads to

$$\mathcal{O}_{\text{LLog+LInst}} = Z[\mathbb{T}_u] \mathcal{O}_{\text{LLog}}[\mathcal{A}], \quad (131)$$

where  $Z[\mathbb{T}_u]$  is a certain functional of the operator  $\mathbb{T}_u$ . In the r.h.s. we have emphasized the dependence of the observable on the initial value of the gauge field. This formula can be expressed more intuitively by performing a Laplace transform of  $Z[\mathbb{T}_u]$  which reads

$$Z[\mathbb{T}_u] \equiv \int [Da(\vec{u})] e^{\int_\Sigma d^3\vec{u} [a \cdot \mathbb{T}_u]} \tilde{Z}[a(\vec{u})]. \quad (132)$$

Given the structure of  $a \cdot \mathbb{T}_u$  in Eq. (42), the functional integration  $[Da(\vec{u})]$  is an integration over the initial fluctuation  $a^\mu(\vec{u})$  itself and over some of its first derivatives. Because  $\mathbb{T}_u$  is the generator of translations of the initial

<sup>35</sup>Indeed, one can think of the Boltzmann equation as an equation that effectively resums a certain class of secular divergences.

conditions on the light cone, the exponential in the previous formula is the translation operator itself. When this exponential acts on a functional of the initial classical field  $\mathcal{A}$ , it gives the same functional evaluated with a shifted initial condition  $\mathcal{A} + a$ . Therefore, we can write

$$\mathcal{O}_{\text{LLog}+\text{LInst}} = \int [Da(\vec{u})] \tilde{Z}[a(\vec{u})] \mathcal{O}_{\text{LLog}}[\mathcal{A} + a]. \quad (133)$$

The effect of the resummation is simply to add fluctuations to the initial conditions of the classical field, with a distribution that depends on the outcome of the resummation.<sup>36</sup> The resummation lifts the limited applicability of the CGC approach implied by Eq. (128). Indeed, after the resummation, the fluctuation  $a(\mathbf{u})$  enters only in the initial condition for the full Yang-Mills equations whose nonlinearities prevent the solution from blowing up. Combining our factorization formula in Eq. (123) with the *conjectured* result of the resummation of the leading instabilities, one obtains a generalization of Eq. (123) which reads

$$\begin{aligned} \langle \mathcal{O} \rangle_{\text{LLog}+\text{LInst}} &= \int [D\tilde{\mathcal{A}}_1^+] [D\tilde{\mathcal{A}}_2^-] W_{Y_1}[\tilde{\mathcal{A}}_1^+] W_{Y_2}[\tilde{\mathcal{A}}_2^-] \\ &\quad \times \int [Da(\vec{u})] \tilde{Z}[a(\vec{u})] \\ &\quad \times \mathcal{O}_{\text{LO}}[\tilde{\mathcal{A}}_1^+ + a, \tilde{\mathcal{A}}_2^- + a]. \end{aligned} \quad (134)$$

This formula resums the most singular terms at each order in  $\alpha_s$ . In comparison to the physics of the initial and final states, respectively, in the collinear factorization framework, the distributions  $W[\rho]$  are analogous to parton distributions while  $\tilde{Z}[a]$  plays a role similar to that of a fragmentation function.<sup>37</sup> To prove Eq. (134), and to extract the spectrum of fluctuations, one needs to compute the behavior of fluctuations on the forward light-cone wedge at  $x^\mp = \epsilon$ ,  $x^\pm \rightarrow +\infty$ .

Even after the resummations are performed in the initial and final states, Eq. (134) still suffers from the usual problem of collinear gluon splitting in the final state [36]. This, however, is not a serious concern in heavy ion collisions because collinear singularities occur only when one takes the  $\tau \rightarrow +\infty$  limit. In practice, we expect to have switched to a more efficient description like kinetic theory or hydrodynamics long before this becomes a problem. Indeed, the initial condition for hydrodynamics, which is specified in terms of the energy-momentum tensor  $T^{\mu\nu}$ , is free of any infrared and collinear divergences because it measures only the density and flow of energy and momentum. It is straightforward to reexpress our results for multiplicity moments in terms of  $T^{\mu\nu}$ .

<sup>36</sup>In a recent work, using a completely different approach, the spectrum of initial fluctuations was found to be Gaussian [111].

<sup>37</sup>Naturally, this functional has nothing to do with a gluon fragmenting into a hadron. Instead, it describes how classical fields become gluons.

A far more challenging problem, that has still not received a satisfactory answer, is to understand how the initial particle spectrum—or the local energy-momentum tensor—becomes isotropic and perhaps even thermal. Indeed, a very important question is whether this improved resummation, that includes the leading unstable terms, hastens the local thermalization of the system formed in heavy ion collisions.

## VII. SUMMARY AND OUTLOOK

In this paper, we have presented a novel derivation of the JIMWLK equation. We showed that in this approach the JIMWLK Hamiltonian can be determined entirely in terms of retarded propagators with no ambiguities related to light-cone pole prescriptions. Our approach generalizes easily to the case of nucleus-nucleus collisions, and we were able to derive the factorization formula in Eq. (123). This formula is valid to all orders for leading logs in  $x$  and to all orders in the color charge densities of the nuclei. For this factorization to work, it appears crucial to consider an observable that depends on a single rapidity, and that can be expressed in terms of retarded fields. Since we had previously linked retarded boundary conditions to the inclusiveness of an observable, this emphasizes the importance of inclusiveness for factorization, and the difficulties one may expect when considering exclusive observables.

In view of this, it seems interesting to study whether the factorization theorem proved here can be extended to less inclusive quantities. One such example is the production of two jets that are separated in rapidity by  $\Delta Y \gg 1/\alpha_s$ . In particular, can the evolution between the jets be factorized from JIMWLK evolution of the wave functions as in the case of inclusive gluon production? Answers to these questions will be of great importance in assessing whether the early time dynamics in heavy ion collisions leaves an imprint in the long range rapidity correlations at later stages.

We further conjectured the existence of the generalized factorization formula in Eq. (134). This expression also resums the leading exponentials in time arising from the instability of the classical fields to quantum fluctuations on the initial light-cone surface. The resulting spectrum of fluctuations is very important for determining the subsequent thermalization of the Glasma. Work in this direction is in progress.

## ACKNOWLEDGMENTS

We would like to thank Nestor Armesto, Edmond Iancu, Jamal Jalilian-Marian, Yuri Kovchegov, Alex Kovner, Misha Lublinsky, and Larry McLerran for very useful conversations. F.G. and R.V. thank the Yukawa Institute for Theoretical Physics of Kyoto University and the Yukawa International Program for Quark-Hadron Sciences for their support during the completion of this

work. R. V.'s work is supported by the U.S. Department of Energy under DOE Contract No. DE-AC02-98CH10886. F. G.'s work is supported in part by Agence Nationale de la Recherche via the programme ANR-06-BLAN-0285-01.

## APPENDIX A: GLUON PROPAGATOR IN LC GAUGE

Consider the QCD Lagrangian to which we add a gauge fixing term proportional to  $(\tilde{n} \cdot A)^2$ ,

$$\mathcal{L} \equiv -\frac{1}{4} F_{\mu\nu}^a F_a^{\mu\nu} + \frac{1}{2\alpha} (\tilde{n} \cdot A)^2. \quad (\text{A1})$$

We are mostly interested in the case where  $\tilde{n} \cdot A = A^+$ , but in fact most of the discussion is valid for any vector  $\tilde{n}^\mu$ . In order to determine the free propagator in this gauge, we need first to isolate the quadratic part of the Lagrangian,

$$\mathcal{L}_{\text{quad}} = \frac{1}{2} A_\mu^a \left[ \square g^{\mu\nu} - \partial^\mu \partial^\nu + \frac{1}{\alpha} \tilde{n}^\mu \tilde{n}^\nu \right] A_\nu^a. \quad (\text{A2})$$

The free propagator we are looking for is a Green's function of the operator in the square brackets. Its calculation is best performed in momentum space, where we need to invert

$$-g^{\mu\nu} k^2 + k^\mu k^\nu + \frac{1}{\alpha} \tilde{n}^\mu \tilde{n}^\nu. \quad (\text{A3})$$

Because this tensor is symmetric in  $(\mu, \nu)$ , its inverse must be a linear combination of  $g^{\mu\nu}$ ,  $k^\mu k^\nu$ ,  $\tilde{n}^\mu \tilde{n}^\nu$ , and  $k^\mu \tilde{n}^\nu + k^\nu \tilde{n}^\mu$ . Writing the most general linear combination of these elementary tensors, and multiplying it by Eq. (A3), we finally obtain the following expression for the propagator in momentum space:

$$D_{0,r}^{\mu\nu}(k) = -\frac{g^{\mu\nu}}{k^2} + \frac{k^\mu k^\nu}{(\tilde{n} \cdot k)^2} \left[ \alpha - \frac{\tilde{n}^2}{k^2} \right] + \frac{k^\mu \tilde{n}^\nu + k^\nu \tilde{n}^\mu}{k^2 (\tilde{n} \cdot k)}. \quad (\text{A4})$$

Note that this expression is still incomplete, because we need to add  $i\epsilon$ 's to the denominators in order to make the propagator regular on the real energy axis. Doing so amounts to choosing certain boundary conditions for the fields that evolve according to this propagator. In this paper, the central object is the retarded propagator, which has all its poles below the real energy axis. This amounts to writing

$$D_{0,r}^{\mu\nu}(k) = -\frac{g^{\mu\nu}}{k^2 + ik^0\epsilon} + \frac{k^\mu k^\nu}{(\tilde{n} \cdot k + i\epsilon)^2} \left[ \alpha - \frac{\tilde{n}^2}{k^2 + ik^0\epsilon} \right] + \frac{k^\mu \tilde{n}^\nu + k^\nu \tilde{n}^\mu}{(k^2 + ik^0\epsilon)(\tilde{n} \cdot k + i\epsilon)}. \quad (\text{A5})$$

(Our choice for the  $i\epsilon$  prescription of the  $\tilde{n} \cdot k$  denominators is indeed retarded if  $n^0 > 0$ . We will assume that this is the case.)

In the case of the light-cone gauge  $A^+ = 0$ , this amounts to choosing a vector  $\tilde{n}^\mu$  that has  $\tilde{n}^- = 1$  and all its other

components zero. Moreover, we work in the ‘‘strict’’ light-cone gauge, which corresponds to the limit  $\alpha \rightarrow 0$  for the gauge fixing parameter. The propagator simplifies somewhat in this particular case:

$$D_{0,r}^{\mu\nu}(k) = -\frac{1}{k^2 + ik^0\epsilon} \left[ g^{\mu\nu} - \frac{k^\mu \tilde{n}^\nu + k^\nu \tilde{n}^\mu}{\tilde{n} \cdot k + i\epsilon} \right]. \quad (\text{A6})$$

Note that this propagator is zero if any of its Lorentz indices is equal to +.

## APPENDIX B: GREEN'S FORMULA IN LC GAUGE

An essential ingredient in our discussion is the Green's formula that expresses a field fluctuation in terms of its value on some initial surface. In this appendix, this initial surface will be the lightlike plane defined by  $x^- = 0$ , but our derivation is more general than that and applies to any initial surface.

### 1. Green's formula for a small fluctuation in the vacuum

Consider first a small field fluctuation  $a^\mu$  propagating in the vacuum. In the strict light-cone gauge, it obeys

$$a^+(y) = 0, \quad [\square_y g^{\mu\nu} - \partial_y^\mu \partial_y^\nu] a_\nu(y) = 0. \quad (\text{B1})$$

Recall also that the free propagator  $D_{0,r}^{\rho\mu}(x, y)$  obeys

$$D_{\mu 0,r}^\rho(x, y) [\overleftarrow{\square}_y g^{\mu\nu} - \overleftarrow{\partial}_y^\mu \overleftarrow{\partial}_y^\nu] = g^{\rho\nu} \delta(x - y), \quad (\text{B2})$$

where the arrows indicate that the derivatives act on the left. Now, multiply Eq. (B1) by  $D_{0,r}^{\rho\mu}(x, y)$  on the left, Eq. (B2) by  $a_\nu(y)$  on the right, integrate  $y$  over all the domain defined by  $y^- > 0$ , and subtract the two equations. One obtains

$$a^\rho(x) = \int_{y^- > 0} d^4 y D_{\mu 0,r}^\rho(x, y) [\overrightarrow{\partial}_y^\mu \overrightarrow{\partial}_y^\nu - \overleftarrow{\square}_y g^{\mu\nu}] a_\nu(y), \quad (\text{B3})$$

where  $\overrightarrow{A} \equiv \overrightarrow{A} - \overleftarrow{A}$ . Using the relations

$$A \overleftarrow{\square} B = \partial^\mu [A \overrightarrow{\partial}_\mu B], \quad (\text{B4})$$

$$A \overrightarrow{\partial}^\mu \overrightarrow{\partial}^\nu B = \frac{1}{2} \partial^\mu [A \overrightarrow{\partial}^\nu B] + \frac{1}{2} \partial^\nu [A \overrightarrow{\partial}^\mu B],$$

we see that the integrand in Eq. (B3) is a total derivative. Therefore, we can rewrite this integral as an integral on the boundary of the integration domain. If the derivative we integrate by parts is a  $\partial^i$  or  $\partial^-$ , then the corresponding boundary is located at infinity in the direction  $y^i$  or  $y^+$ , respectively. We will assume that the field fluctuation under consideration has a compact enough support so that these contributions vanish. We are thus left with the terms coming from the derivative  $\partial^+$ . The contribution from the boundary at  $y^- = +\infty$  is zero, because of our choice of the retarded prescription for the propagator.

Therefore, the only contribution is from the boundary at  $y^- = 0$ ,

$$a^\rho(x) = \int_{y^- = 0} dy^+ d^2 y_\perp D_{\mu 0, R}^\rho(x, y) \left[ g^{\mu\nu} (n \cdot \overleftrightarrow{\partial}) - \frac{1}{2} (n^\mu \overleftrightarrow{\partial}_y^\nu + n^\nu \overleftrightarrow{\partial}_y^\mu) \right] a_\nu(y), \quad (\text{B5})$$

where  $n^\mu$  is a vector such that  $n \cdot A = A^-$  (it is the unit vector normal to the surface  $y^- = 0$ ). This formula indicates how the value of the fluctuation at the point  $x$  is related to its value on an initial surface located at  $y^- = 0$ . (Note that this dependence is linear since small fluctuations obey a linear equation of motion.) *A priori*, it involves the values of all the components of the fluctuation on this surface, as well as that of its first derivatives. However, some of this information is not necessary because the propagator vanishes when  $\mu = +$  and because of the gauge condition  $a^+(y) = 0$ . If one eliminates from the previous formula all the terms that are obviously zero and integrates some terms by parts,<sup>38</sup> we get  $a^\rho(x) \equiv \mathcal{B}_0^\rho[a](x)$ , where  $\mathcal{B}_0^\rho[a](x)$  is an integral that depends only on the value of the field and of some of its derivatives on the initial surface,

$$\begin{aligned} \mathcal{B}_0^\rho[a](x) = & \int_{y^- = 0} dy^+ d^2 y_\perp \{ [\partial_\mu^y D_{0, R}^{\rho\mu}(x, y)] a^-(y) \\ & - D_{0, R}^{\rho-}(x, y) [\partial_y^\mu a_\mu(y)] \\ & - D_{0, R}^{\rho i}(x, y) 2\partial_y^- a^i(y) \}. \end{aligned} \quad (\text{B6})$$

Therefore, it appears that in the light-cone gauge  $A^+ = 0$ , and for an initial surface  $x^- = 0$ , we need to know the initial value of  $a^-$ ,  $\partial^- a^i$  and  $\partial_\mu a^\mu$  in order to fully determine the value of the fluctuation at the point  $x$ . This fact is the reason why there are only three terms in the definition of the operator  $\mathbb{T}_\mu$  in Eq. (42) (but we postpone until the end of this section the explanation of why one needs to include the Wilson line  $\Omega$  in this definition).

Moreover, the first term in the right-hand side of Eq. (B6) can be simplified considerably by using the explicit expression of the free propagators in light-cone gauge:

$$\partial_\mu^y D_{0, R}^{\rho\mu}(x, y) = \delta^{\rho-} \theta(x^- - y^-) \delta(x^+ - y^+) \delta(\mathbf{x}_\perp - \mathbf{y}_\perp). \quad (\text{B7})$$

## 2. Green's formula for classical solutions

There is also a similar Green's formula for retarded classical solutions of the Yang-Mills equations. Contrary to the case of small fluctuations, we do not assume that the

<sup>38</sup>The antisymmetric derivatives  $\overleftrightarrow{\partial}_y^-$  and  $\overleftrightarrow{\partial}_y^+$  can be eliminated by integration by parts. This is not possible for  $\overleftrightarrow{\partial}_y^+$  since the boundary term does not contain an integral with respect to  $y^-$ . This is why we have a term involving the derivative  $\partial_y^+ D_{0, R}^{\rho-}$ .

gauge field is small, and we keep all the self-interactions as well as the interactions with some external source. Formally, we can write the Lagrangian as

$$\mathcal{L} = \mathcal{L}_{\text{quad}} - U(\mathcal{A}), \quad (\text{B8})$$

where  $U(\mathcal{A})$  is a local polynomial of the gauge field. It contains the 3- and 4-gluon couplings and the coupling to the external source. In the  $\mathcal{A}^+ = 0$  gauge, the corresponding classical equation of motion is

$$[\square_y g^{\mu\nu} - \partial_y^\mu \partial_y^\nu] \mathcal{A}_\nu(y) = \frac{\partial U(\mathcal{A})}{\partial \mathcal{A}_\mu(y)}. \quad (\text{B9})$$

Then one can follow the same procedure as in the case of small fluctuations, and we obtain

$$\mathcal{A}^\rho(x) = \int_{y^- > 0} d^4 y D_{0, R}^{\rho\mu}(x, y) \frac{\partial U(\mathcal{A})}{\partial \mathcal{A}^\mu(y)} + \mathcal{B}_0^\rho[\mathcal{A}](x). \quad (\text{B10})$$

Of course, the dependence of the classical field on its initial conditions is no longer linear because of the first term in the right-hand side; the self-interactions of the gauge fields lead to an involved dependence on the initial conditions.

## 3. Green's formula for $a^\mu$ in a background field

Finally, the Green's formula of Eq. (B6) can be extended to the situation where the fluctuation  $a^\mu(x)$  propagates on top of a classical background field  $\mathcal{A}^\mu$  rather than the vacuum. The only change is that the free propagator must be replaced by the propagator in a background field. The property that its  $\mu = +$  Lorentz component vanishes remains true, because it is a consequence of the choice of the gauge. For such a fluctuation, there is also a Green's formula that uses only the free gauge propagator, and where the interactions with the background field appear explicitly as the additional term

$$a^\rho(x) = \int_{y^- > 0} d^4 y D_{0, R}^{\rho\nu}(x, y) \frac{\partial^2 U(\mathcal{A})}{\partial \mathcal{A}_\nu(y) \partial A^\sigma(y)} a^\sigma(y) + \mathcal{B}_0^\rho[a](x). \quad (\text{B11})$$

The derivation of this formula is very similar to that for the classical field  $\mathcal{A}^\mu$ . We can also rewrite it in a form very similar to Eq. (B6), i.e.  $a^\rho(x) = \mathcal{B}[a](x)$  with

$$\begin{aligned} \mathcal{B}^\rho[a](x) = & \int_{y^- = 0} dy^+ d^2 y_\perp \{ [\partial_\mu^y D_R^{\rho\mu}(x, y)] a^-(y) \\ & - D_R^{\rho-}(x, y) [\partial_y^\mu a_\mu(y)] \\ & - D_R^{\rho i}(x, y) 2\partial_y^- a^i(y) \}. \end{aligned} \quad (\text{B12})$$

The boundary term  $\mathcal{B}[a]$  differs from  $\mathcal{B}_0[a]$  in the fact that it contains the retarded propagator  $D_R^{\mu\nu}$  dressed by the background field instead of the bare retarded propagator  $D_{0, R}^{\mu\nu}$ . A crucial difference between the dressed and bare



propagators is that the simplification of Eq. (B7) does not occur with the dressed propagator.

In the derivation of the JIMWLK equation, the fluctuations  $a^\mu(x)$  one considers are fluctuations whose initial conditions at  $x^0 \rightarrow -\infty$  are plane waves of momentum  $k$ . One can calculate explicitly their value on the initial surface, which means that we know analytically the quantities  $a^-$ ,  $\partial^- a^i$ , and  $\partial_\mu a^\mu$  in the r.h.s. of Eq. (B12). A crucial property is that the initial values of  $a^-$  and  $\partial^- a^i$  are suppressed by an extra factor  $1/k^+$ , and thus any term containing them cannot have a logarithmic divergence when  $k^+ \rightarrow +\infty$ . This argument is correct provided the prefactors of these quantities in Eq. (B12) do not bring factors of  $k^+$ . There is no problem with the second and third terms, since their prefactors are just propagators.

However, the coefficient of the first term can be large because it involves the derivative of the propagator. More specifically, the problem comes from the  $\partial_{y^+}$  derivative, which is large in the region where the  $y^-$  dependence is strong, where the boundary condition is set. There are two ways to proceed: keep track separately of the first and second terms of Eq. (B12), or try to combine them into a unique term. The second option is the simplest in the case where the propagator  $D_R^{\rho\mu}$  is dressed by a pure gauge background field: by rotating the fluctuation  $a^\mu$  by the Wilson line  $\Omega$  that parametrizes the pure gauge background,  $a^\mu \rightarrow \Omega a^\mu$ , we can rewrite the boundary term as

$$\begin{aligned} B^{\rho}[a](x) = & \Omega^\dagger(x) \int_{y^-=0} dy^+ d^2 y_\perp \{ [\partial_\mu^y (D_{0,R}^{\rho\mu}(x, y)) \\ & \times \Omega(y) a^-(y) - D_{0,R}^{\rho-}(x, y) [\partial_y^\mu \Omega(y) a_\mu(y)] \\ & - D_{0,R}^{\rho i}(x, y) 2\partial_y^- \Omega(y) a^i(y) \}, \end{aligned} \quad (\text{B13})$$

where we have now only bare propagators. (To obtain this expression, one should apply a gauge transformation  $\Omega$  that eliminates the background field, and perform the inverse gauge transform at the endpoint  $x$ .) This is why the most convenient definition of  $\mathbb{T}_u$  in Eq. (42) involves functional derivatives with respect to  $\Omega a^\mu$  rather than  $a^\mu$  itself.<sup>39</sup> Note that for this discussion to hold, it is only necessary that the background field is a pure gauge in the vicinity above the initial surface, since the derivative is with respect to a coordinate on this initial surface. Whether the background field is a pure gauge everywhere above the initial surface is not important.

<sup>39</sup>Of course, the two ways of defining  $\mathbb{T}_u$ —with and without the  $\Omega$ —are exactly equivalent. But if we did not include the  $\Omega$  in the definition, the logarithmic divergences would come from a combination of the second and third terms of Eq. (42), instead of being limited to the third term if we include the  $\Omega$  in the definition of  $\mathbb{T}_u$ .

## APPENDIX C: TWO-DIMENSIONAL FREE PROPAGATOR

In the derivation of the JIMWLK equation, one makes use of several formulas involving the bare two-dimensional propagators. These formulas are not new: all of them have already been used in one form or another in previous papers discussing the JIMWLK equation. We compile them in this appendix, with their derivation, as a convenient reference for the reader.

Let us denote  $G(\mathbf{x}_\perp - \mathbf{y}_\perp)$  a Green's function of the two-dimensional Laplacian operator,

$$\partial_\perp^2 G(\mathbf{x}_\perp - \mathbf{y}_\perp) = \delta(\mathbf{x}_\perp - \mathbf{y}_\perp). \quad (\text{C1})$$

It admits a simple Fourier representation,

$$G(\mathbf{x}_\perp - \mathbf{y}_\perp) = - \int \frac{d^2 \mathbf{k}_\perp}{(2\pi)^2} e^{i\mathbf{k}_\perp \cdot (\mathbf{x}_\perp - \mathbf{y}_\perp)} \frac{1}{k_\perp^2}. \quad (\text{C2})$$

Note that this object suffers from an infrared problem, which is obvious for dimensional reasons: this propagator is a dimensionless object in coordinate space, invariant under translations and rotations, and therefore it must be a function of  $\mu |\mathbf{x}_\perp - \mathbf{y}_\perp|$  where  $\mu$  is some mass scale that was not present in the previous equation.

Derivatives of this propagator do not suffer from this infrared ambiguity. Consider, for instance,<sup>40</sup>

$$\partial_x^i G(\mathbf{x}_\perp - \mathbf{y}_\perp) = i \int \frac{d^2 \mathbf{k}_\perp}{(2\pi)^2} e^{i\mathbf{k}_\perp \cdot (\mathbf{x}_\perp - \mathbf{y}_\perp)} \frac{k^i}{k_\perp^2}. \quad (\text{C3})$$

From its symmetries and dimension, it is obvious that this derivative can be written as

$$\partial_x^i G(\mathbf{x}_\perp - \mathbf{y}_\perp) = C \frac{\mathbf{x}_\perp^i - \mathbf{y}_\perp^i}{(\mathbf{x}_\perp - \mathbf{y}_\perp)^2}, \quad (\text{C4})$$

where the prefactor  $C$  is dimensionless. Because the derivative of the propagator is not infrared singular, the cutoff  $\mu$  cannot appear in its expression and  $C$  must be a pure number (otherwise it would have to be a function of  $\mu |\mathbf{x}_\perp - \mathbf{y}_\perp|$  to have the correct dimension). In order to determine the constant, take another derivative  $\partial_x^i$  and integrate over  $\mathbf{x}_\perp$  the resulting equation over some domain  $\Omega$  of the plane that contains the point  $\mathbf{y}_\perp$ . On the left-hand side, we get the integral of a delta function since  $G$  is a Green's function of  $\partial_\perp^2$ . We then get

$$1 = C \int_\Omega d^2 \mathbf{x}_\perp \partial_x^i \frac{\mathbf{x}_\perp^i - \mathbf{y}_\perp^i}{(\mathbf{x}_\perp - \mathbf{y}_\perp)^2}. \quad (\text{C5})$$

The right-hand side can be transformed by using the two-dimensional Stokes theorem, leading to an integral on the boundary of  $\Omega$  (oriented counterclockwise),

<sup>40</sup>Let us recall that  $\partial_x^i = \frac{\partial}{\partial x_i} = - \frac{\partial}{\partial x^i}$ .

$$1 = C \int_{\partial\Omega} \frac{\epsilon^{ij}(\mathbf{x}_\perp^i - \mathbf{y}_\perp^i) dx^j}{(\mathbf{x}_\perp - \mathbf{y}_\perp)^2}, \quad (\text{C6})$$

where  $\epsilon^{ij}$  is completely antisymmetric ( $\epsilon^{12} = 1$ ). The contour integral in this equation is a topological quantity that depends only on the winding number of the contour  $\partial\Omega$  around the point  $\mathbf{y}_\perp$ . Thus, it is best calculated by deforming  $\partial\Omega$  into the unit circle around the point  $\mathbf{y}_\perp$ . We easily get

$$1 = 2\pi C. \quad (\text{C7})$$

Thus we have

$$\partial_x^i G(\mathbf{x}_\perp - \mathbf{y}_\perp) = \frac{1}{2\pi} \frac{\mathbf{x}_\perp^i - \mathbf{y}_\perp^i}{(\mathbf{x}_\perp - \mathbf{y}_\perp)^2}. \quad (\text{C8})$$

The second derivative of the propagator is also useful in the derivation of the JIMWLK equation. By applying  $\partial_x^j$  to the previous equation, one obtains

$$\begin{aligned} \partial_x^i \partial_x^j G(\mathbf{x}_\perp - \mathbf{y}_\perp) &= \frac{1}{2\pi} \partial_x^j \frac{\mathbf{x}_\perp^i - \mathbf{y}_\perp^i}{(\mathbf{x}_\perp - \mathbf{y}_\perp)^2} \\ &= \frac{1}{2\pi(\mathbf{x}_\perp - \mathbf{y}_\perp)^2} \left[ \delta^{ij} \right. \\ &\quad \left. - 2 \frac{(\mathbf{x}_\perp^i - \mathbf{y}_\perp^i)(\mathbf{x}_\perp^j - \mathbf{y}_\perp^j)}{(\mathbf{x}_\perp - \mathbf{y}_\perp)^2} \right]. \end{aligned} \quad (\text{C9})$$

This formula, although perfectly correct for  $\mathbf{x}_\perp \neq \mathbf{y}_\perp$ , is incorrect at the point  $\mathbf{x}_\perp = \mathbf{y}_\perp$ . In order to see this, take the trace over the indices  $i$  and  $j$ . In the left-hand side, we have the Laplacian of the propagator, i.e.  $\delta(\mathbf{x}_\perp - \mathbf{y}_\perp)$ , while the right-hand side would give zero. Thus the full formula for the second derivative is

$$\partial_x^i \partial_x^j G(\mathbf{x}_\perp - \mathbf{y}_\perp) = \frac{\delta^{ij}}{2} \delta(\mathbf{x}_\perp - \mathbf{y}_\perp) + \frac{1}{2\pi} \Delta^{ij}(\mathbf{x}_\perp - \mathbf{y}_\perp), \quad (\text{C10})$$

with

$$\begin{aligned} \Delta^{ij}(\mathbf{x}_\perp - \mathbf{y}_\perp) &\equiv \frac{1}{(\mathbf{x}_\perp - \mathbf{y}_\perp)^2} \\ &\times \left[ \delta^{ij} - 2 \frac{(\mathbf{x}_\perp^i - \mathbf{y}_\perp^i)(\mathbf{x}_\perp^j - \mathbf{y}_\perp^j)}{(\mathbf{x}_\perp - \mathbf{y}_\perp)^2} \right]. \end{aligned} \quad (\text{C11})$$

This function  $\Delta^{ij}$  obeys an interesting identity. By integrating by parts, one can check that

$$\begin{aligned} &\int \frac{d^2\mathbf{u}_\perp d^2\mathbf{v}_\perp}{(2\pi)^2 (2\pi)^2} \frac{(\mathbf{x}_\perp^i - \mathbf{u}_\perp^i)(\mathbf{y}_\perp^j - \mathbf{v}_\perp^j)}{(\mathbf{x}_\perp - \mathbf{u}_\perp)^2 (\mathbf{y}_\perp - \mathbf{v}_\perp)^2} \partial_u^i \partial_v^j G(\mathbf{u}_\perp - \mathbf{v}_\perp) \\ &= -\frac{1}{(2\pi)^2} \int \frac{d^2\mathbf{u}_\perp}{(2\pi)^2} \frac{(\mathbf{x}_\perp^i - \mathbf{u}_\perp^i)(\mathbf{y}_\perp^j - \mathbf{u}_\perp^j)}{(\mathbf{x}_\perp - \mathbf{u}_\perp)^2 (\mathbf{y}_\perp - \mathbf{u}_\perp)^2} \\ &= -\int \frac{d^2\mathbf{u}_\perp d^2\mathbf{v}_\perp}{(2\pi)^2 (2\pi)^2} \frac{(\mathbf{x}_\perp^i - \mathbf{u}_\perp^i)(\mathbf{y}_\perp^j - \mathbf{v}_\perp^j)}{(\mathbf{x}_\perp - \mathbf{u}_\perp)^2 (\mathbf{y}_\perp - \mathbf{v}_\perp)^2} \delta^{ij} \delta(\mathbf{u}_\perp - \mathbf{v}_\perp). \end{aligned} \quad (\text{C12})$$

Using now Eq. (C10), we obtain the following identity,

$$\begin{aligned} &\int \frac{d^2\mathbf{u}_\perp d^2\mathbf{v}_\perp}{(2\pi)^2 (2\pi)^2} \frac{(\mathbf{x}_\perp^i - \mathbf{u}_\perp^i)(\mathbf{y}_\perp^j - \mathbf{v}_\perp^j)}{(\mathbf{x}_\perp - \mathbf{u}_\perp)^2 (\mathbf{y}_\perp - \mathbf{v}_\perp)^2} \\ &\times \left[ \frac{\delta^{ij}}{2} \delta(\mathbf{u}_\perp - \mathbf{v}_\perp) - \frac{1}{2\pi} \Delta^{ij}(\mathbf{u}_\perp - \mathbf{v}_\perp) \right] = 0. \end{aligned} \quad (\text{C13})$$

Let us also provide an alternate representation of the two-dimensional propagator that is sometimes helpful. Let us start with the integral

$$\begin{aligned} &\int \frac{d^2\mathbf{u}_\perp}{(2\pi)^2} \frac{\mathbf{u}_\perp^i - \mathbf{x}_\perp^i}{(\mathbf{u}_\perp - \mathbf{x}_\perp)^2} \frac{\mathbf{u}_\perp^j - \mathbf{y}_\perp^j}{(\mathbf{u}_\perp - \mathbf{y}_\perp)^2} \\ &= \int d^2\mathbf{u}_\perp [\partial_u^i G(\mathbf{u}_\perp - \mathbf{x}_\perp)] [\partial_u^j G(\mathbf{u}_\perp - \mathbf{y}_\perp)]. \end{aligned} \quad (\text{C14})$$

The integral in the right-hand side can be performed by parts, since it leads to the Laplacian of a propagator, which is a delta function. Thus, we obtain the identity

$$G(\mathbf{x}_\perp - \mathbf{y}_\perp) = -\int \frac{d^2\mathbf{u}_\perp}{(2\pi)^2} \frac{\mathbf{u}_\perp^i - \mathbf{x}_\perp^i}{(\mathbf{u}_\perp - \mathbf{x}_\perp)^2} \frac{\mathbf{u}_\perp^j - \mathbf{y}_\perp^j}{(\mathbf{u}_\perp - \mathbf{y}_\perp)^2}. \quad (\text{C15})$$

Note that the integral over  $\mathbf{u}_\perp$  suffers from the same infrared problems that we have already mentioned at the beginning of this appendix.

[1] A.H. Mueller, *Perturbative QCD* (World Scientific, Singapore, 1988).  
 [2] J.C. Collins, D.E. Soper, and G. Sterman, Nucl. Phys. **B250**, 199 (1985).  
 [3] J.C. Collins, D.E. Soper, and G. Sterman, Nucl. Phys. **B261**, 104 (1985).

[4] J.C. Collins, D.E. Soper, and G. Sterman, Nucl. Phys. **B263**, 37 (1986).  
 [5] J.C. Collins, D.E. Soper, and G. Sterman, Nucl. Phys. **B286**, 704 (1987).  
 [6] J.C. Collins, D.E. Soper, and G. Sterman, Nucl. Phys. **B308**, 833 (1988).

- [7] L. V. Gribov, E. M. Levin, and M. G. Ryskin, Phys. Rep. **100**, 1 (1983).
- [8] A. H. Mueller and J-W. Qiu, Nucl. Phys. **B268**, 427 (1986).
- [9] L. D. McLerran and R. Venugopalan, Phys. Rev. D **49**, 2233 (1994).
- [10] L. D. McLerran and R. Venugopalan, Phys. Rev. D **49**, 3352 (1994).
- [11] L. D. McLerran and R. Venugopalan, Phys. Rev. D **50**, 2225 (1994).
- [12] J. Jalilian-Marian, A. Kovner, L. D. McLerran, and H. Weigert, Phys. Rev. D **55**, 5414 (1997).
- [13] J. Jalilian-Marian, A. Kovner, A. Leonidov, and H. Weigert, Nucl. Phys. **B504**, 415 (1997).
- [14] J. Jalilian-Marian, A. Kovner, A. Leonidov, and H. Weigert, Phys. Rev. D **59**, 014014 (1998).
- [15] J. Jalilian-Marian, A. Kovner, A. Leonidov, and H. Weigert, Phys. Rev. D **59**, 034007 (1999).
- [16] J. Jalilian-Marian, A. Kovner, A. Leonidov, and H. Weigert, Phys. Rev. D **59**, 099903(E) (1999).
- [17] E. Iancu, A. Leonidov, and L. D. McLerran, Nucl. Phys. **A692**, 583 (2001).
- [18] E. Iancu, A. Leonidov, and L. D. McLerran, Phys. Lett. B **510**, 133 (2001).
- [19] E. Ferreiro, E. Iancu, A. Leonidov, and L. D. McLerran, Nucl. Phys. **A703**, 489 (2002).
- [20] I. Balitsky, Nucl. Phys. **B463**, 99 (1996).
- [21] I. Balitsky, Phys. Rev. D **70**, 114030 (2004).
- [22] Yu. V. Kovchegov, Phys. Rev. D **61**, 074018 (2000).
- [23] F. Gelis and R. Venugopalan, Phys. Rev. D **69**, 014019 (2004).
- [24] F. Gelis and R. Venugopalan, Nucl. Phys. **A776**, 135 (2006).
- [25] F. Gelis and R. Venugopalan, arXiv:hep-ph/0611157.
- [26] F. Gelis, T. Lappi, and R. Venugopalan, Int. J. Mod. Phys. E **16**, 2595 (2007).
- [27] F. Gelis, T. Lappi, and R. Venugopalan, following Article, Phys. Rev. D **78** 054020 (2008).
- [28] V. N. Gribov and L. N. Lipatov, Sov. J. Nucl. Phys. **15**, 675 (1972).
- [29] G. Altarelli and G. Parisi, Nucl. Phys. **B126**, 298 (1977).
- [30] Yu. Dokshitzer, Sov. Phys. JETP **46**, 641 (1977).
- [31] I. Balitsky, Phys. Rev. D **75**, 014001 (2007).
- [32] I. Balitsky and G. A. Chirilli, Phys. Rev. D **77**, 014019 (2008).
- [33] E. Gardi, J. Kuokkanen, K. Rummukainen, and H. Weigert, Nucl. Phys. **A784**, 282 (2007).
- [34] Yu. V. Kovchegov and H. Weigert, Nucl. Phys. **A784**, 188 (2007).
- [35] Yu. V. Kovchegov and H. Weigert, Nucl. Phys. **A789**, 260 (2007).
- [36] Yu. V. Kovchegov and H. Weigert, Nucl. Phys. **A807**, 158 (2008).
- [37] J. L. Albacete and Y. Kovchegov, Phys. Rev. D **75**, 125021 (2007).
- [38] N. Goldenfeld, *Lectures on Phase Transitions and the Renormalization Group*, Frontiers in Physics (Addison Wesley, Reading, MA, 1992).
- [39] F. Gelis, S. Jeon, and R. Venugopalan, arXiv:0706.3775.
- [40] J. Berges, AIP Conf. Proc. **739**, 3 (2004).
- [41] A. Kovner and G. Milhano, Phys. Rev. D **61**, 014012 (1999).
- [42] H. Weigert, Nucl. Phys. **A703**, 823 (2002).
- [43] A. H. Mueller, Phys. Lett. B **523**, 243 (2001).
- [44] A. Kovner and U. Wiedemann, Phys. Rev. D **64**, 114002 (2001).
- [45] A. H. Mueller, A. I. Shoshi, and S. M. H. Wong, Phys. Lett. B **632**, 257 (2006).
- [46] T. Lappi and L. D. McLerran, Nucl. Phys. **A772**, 200 (2006).
- [47] J. Schwinger, J. Math. Phys. (N.Y.) **2**, 407 (1961).
- [48] L. V. Keldysh, Sov. Phys. JETP **20**, 1018 (1964).
- [49] P. M. Bakshi and K. T. Mahanthappa, J. Math. Phys. (N.Y.) **4**, 1 (1963).
- [50] F. Gelis and R. Venugopalan, Nucl. Phys. **A779**, 177 (2006).
- [51] Yu. V. Kovchegov, Phys. Rev. D **54**, 5463 (1996).
- [52] A. Ayala, J. Jalilian-Marian, L. D. McLerran, and R. Venugopalan, Phys. Rev. D **52**, 2935 (1995).
- [53] A. Ayala, J. Jalilian-Marian, L. D. McLerran, and R. Venugopalan, Phys. Rev. D **53**, 458 (1996).
- [54] F. Gelis and Y. Mehtar-Tani, Phys. Rev. D **73**, 034019 (2006).
- [55] A. H. Mueller, arXiv:hep-ph/9911289.
- [56] A. Krasnitz and R. Venugopalan, Phys. Rev. Lett. **84**, 4309 (2000).
- [57] A. Krasnitz and R. Venugopalan, Phys. Rev. Lett. **86**, 1717 (2001).
- [58] A. Krasnitz and R. Venugopalan, Nucl. Phys. **B557**, 237 (1999).
- [59] A. Krasnitz, Y. Nara, and R. Venugopalan, Nucl. Phys. **A727**, 427 (2003).
- [60] A. Krasnitz, Y. Nara, and R. Venugopalan, Phys. Rev. Lett. **87**, 192302 (2001).
- [61] A. Krasnitz, Y. Nara, and R. Venugopalan, Phys. Lett. B **554**, 21 (2003).
- [62] T. Lappi, Phys. Rev. C **67**, 054903 (2003).
- [63] T. Lappi, Phys. Rev. C **70**, 054905 (2004).
- [64] T. Lappi, Phys. Lett. B **643**, 11 (2006).
- [65] A. Kovner, L. D. McLerran, and H. Weigert, Phys. Rev. D **52**, 3809 (1995).
- [66] A. Kovner, L. D. McLerran, and H. Weigert, Phys. Rev. D **52**, 6231 (1995).
- [67] Yu. V. Kovchegov and D. H. Rischke, Phys. Rev. C **56**, 1084 (1997).
- [68] Yu. V. Kovchegov and K. Tuchin, Phys. Rev. D **65**, 074026 (2002).
- [69] J. Jalilian-Marian and Y. Kovchegov, Prog. Part. Nucl. Phys. **56**, 104 (2006).
- [70] J. Jalilian-Marian and Y. Kovchegov, Phys. Rev. D **70**, 114017 (2004); **71**, 079901(E) (2005).
- [71] A. Kovner and M. Lublinsky, J. High Energy Phys. **11** (2006) 083.
- [72] M. A. Braun, Eur. Phys. J. C **55**, 377 (2008).
- [73] M. Kozlov, E. Levin, and A. Prygarin, arXiv:hep-ph/0606260.
- [74] M. Kozlov, E. Levin, and A. Prygarin, Nucl. Phys. **A792**, 122 (2007).
- [75] I. Balitsky, Phys. Rev. D **72**, 074027 (2005).
- [76] Y. Hatta, E. Iancu, L. McLerran, A. Stasto, and D. N. Triantafyllopoulos, Nucl. Phys. **A764**, 423 (2006).
- [77] Y. Hatta, Nucl. Phys. **A768**, 222 (2006).

- [78] L. N. Lipatov, Phys. Rep. **286**, 131 (1997).
- [79] V. S. Fadin, M. I. Kotsky, and R. Fiore, Phys. Lett. B **359**, 181 (1995).
- [80] V. S. Fadin and L. N. Lipatov, Phys. Lett. B **429**, 127 (1998).
- [81] G. Camici and M. Ciafaloni, Phys. Lett. B **412**, 396 (1997); **417**, 390(E) (1998).
- [82] M. Ciafaloni and G. Camici, Phys. Lett. B **430**, 349 (1998).
- [83] J. C. Collins and R. K. Ellis, Nucl. Phys. **B360**, 3 (1991).
- [84] S. Catani, M. Ciafaloni, and F. Hautmann, Nucl. Phys. **B366**, 135 (1991).
- [85] E. M. Levin, M. G. Ryskin, Yu. M. Shabelsky, and A. G. Shuvaev, Sov. J. Nucl. Phys. **53**, 657 (1991).
- [86] J. P. Blaizot, F. Gelis, and R. Venugopalan, Nucl. Phys. **A743**, 13 (2004).
- [87] J. P. Blaizot, F. Gelis, and R. Venugopalan, Nucl. Phys. **A743**, 57 (2004).
- [88] Yu. V. Kovchegov and L. D. McLerran, Phys. Rev. D **60**, 054025 (1999).
- [89] M. A. Braun, arXiv:hep-ph/0101070.
- [90] N. Armesto and M. A. Braun, Eur. Phys. J. C **20**, 517 (2001).
- [91] B. Z. Kopeliovich, A. Schafer, and A. V. Tarasov, Phys. Rev. D **62**, 054022 (2000).
- [92] D. Kharzeev, Yu. Kovchegov, and K. Tuchin, Phys. Rev. D **68**, 094013 (2003).
- [93] R. Baier, A. Kovner, M. Nardi, and U. A. Wiedemann, Phys. Rev. D **72**, 094013 (2005).
- [94] F. Gelis, K. Kajantie, and T. Lappi, Phys. Rev. C **71**, 024904 (2005).
- [95] F. Gelis, K. Kajantie, and T. Lappi, Phys. Rev. Lett. **96**, 032304 (2006).
- [96] D. Kharzeev, A. Krasnitz, and R. Venugopalan, Phys. Lett. B **545**, 298 (2002).
- [97] P. Romatschke and R. Venugopalan, Phys. Rev. Lett. **96**, 062302 (2006).
- [98] P. Romatschke and R. Venugopalan, Eur. Phys. J. A **29**, 71 (2006).
- [99] P. Romatschke and R. Venugopalan, Phys. Rev. D **74**, 045011 (2006).
- [100] A. K. Rebhan and P. Romatschke, Phys. Rev. Lett. **97**, 252301 (2006).
- [101] A. Iwazaki, arXiv:0803.0188.
- [102] H. Fujii and K. Itakura, Nucl. Phys. **A809**, 88 (2008).
- [103] P. Arnold, J. Lenaghan, and G. D. Moore, J. High Energy Phys. 08 (2003) 002.
- [104] A. K. Rebhan, P. Romatschke, and M. Strickland, Phys. Rev. Lett. **94**, 102303 (2005).
- [105] A. K. Rebhan, P. Romatschke, and M. Strickland, J. High Energy Phys. 09 (2005) 041.
- [106] A. Dumitru, Y. Nara, and M. Strickland, Phys. Rev. D **75**, 025016 (2007).
- [107] A. H. Mueller, A. I. Shoshi, and S. M. H. Wong, Nucl. Phys. **B760**, 145 (2007).
- [108] D. Bodeker and K. Rummukainen, J. High Energy Phys. 07 (2007) 022.
- [109] S. Mrowczynski, Proc. Sci., CPOD2006 (2006) 042 [arXiv:hep-ph/0611067].
- [110] A. K. Rebhan, M. Strickland, and M. Attems, arXiv:0802.1714 [Phys. Rev. D (to be published)].
- [111] K. Fukushima, F. Gelis, and L. McLerran, Nucl. Phys. **A786**, 107 (2007).