Physics of the Color Glass Condensate

Edmond Iancu [1]

Service de Physique Théorique, CEA/DSM/SPhT, Unité de recherche associée au CNRS (URA D2306), CE Saclay, F-91191 Gif-sur-Yvette, France

th2

PRELIMINARY VERSION !

Abstract

The Color Glass Condensate (CGC) is a high density form of hadronic matter which controls the interactions in QCD at high energies, or small values of Bjorken x. Predicted by theoretical considerations, this matter might have been seen in the recent experiments at HERA and RHIC, but the experimental conditions should be even more favorable for its observation at LHC. The CGC is *colored* since made of gluons which carry the 'color' charge of QCD; it is a *glass* since its internal dynamics is frozen over the natural time scales for high energy scattering; it is finally a *condensate* since it is characterized by high occupation numbers, or strong classical color fields. High density together with asymptotic freedom imply that the CGC is *weakly coupled*. Based on this observation, an *effective theory* has been constructed within perturbative QCD, which describes the evolution of a hadronic wavefunction with increasing energy in the presence of the non–linear effects associated with the high gluon density. This theory predicts the *saturation* of the gluon distribution for transverse momenta below some characteristic 'saturation scale' which grows rapidly with the energy. It further implies the *unitarization* of the scattering amplitudes at high energy and provides a natural explanation for some remarkable regularities seen in the experimental data, like the 'geometric scaling' in the HERA data for deep inelastic scattering at small x, or the 'high–p_T suppression' in the deuteron–gold collisions at RHIC.

This review paper is devoted to a pedagogical and self–contained presentation of the physical and mathematical foundations of the effective theory for the CGC, with emphasis on modern tools like the renormalization group analysis of the high–energy evolution in QCD and the ensuing Balitsky–JIMWLK equations. We discuss selected applications to the phenomenology at HERA and RHIC. We describe recent theoretical developments concerning the importance of particle number fluctuations in the dilute regime, in relation with modern problems in statistical physics, and also with the 'pomeron loops' of the diagrammatic approach to high–energy QCD.

---

[1] Membre du Centre National de la Recherche Scientifique (CNRS), France.

Contents

1 Introduction: High energy scattering in QCD

1.1 General introduction

Understanding the high energy behaviour of the hadronic scattering amplitudes is a long-standing problem, which emerged long before the advent of the fundamental theory of the strong interactions, the Quantum Chromodynamics (QCD). Back in 1952, Heisenberg [1] has given a qualitative argument combining the Yukawa meson theory with an assumption about the maximal energy rise of the scattering amplitude and the unitarity bound on the S–matrix to conclude that, in the high–energy limit, the total cross–section should rise like $\ln^2 s$, with $s = E_{\text{c.m.}}^2$ (the total center–of–mass energy squared). It was still Heisenberg who initiated the axiomatic approach to the theory of the S–matrix, an ambitious program trying to develop a complete description of the strong interactions without any reference to quantum field theory. Using general principles (like Lorentz symmetry, unitarity, and analiticity) within this approach, one has been able to deduce powerful results about the high energy behaviour, like the celebrated Froissart bound [2–4] stating that the total cross–sections must rise no faster than $\ln^2 s$ in the limit $s \to \infty$. When combined with the theory of the Regge poles [5–7], the S–matrix formalism allowed for a successful description of the phenomenology, which is still in use today for those problems, like the calculation of total cross–sections, which lie beyond the scope of perturbative QCD. In particular, it was in the context of the Regge phenomenology that the concepts of 'Pomeron' [6, 7] and 'Odderon' [8] have been introduced for the first time, although with a somewhat different meaning from that which will appear later in this report.

It was still before the advent of QCD that the *parton picture* has been first formulated, by Bjorken [9] and Feynman [10], and soon after confirmed by the spectacular identification of 'Bjorken scaling' in the electron–proton deep inelastic scattering (DIS) at SLAC [11]. In a critical analysis of the parton picture from the perspective of quantum field theory, Gribov and Lipatov [12] have found that, for an Abelian vector theory like QED, the 'naive' parton model should receive substantial radiative corrections leading to strong violations of the Bjorken scaling, which would contradict the experimental results at SLAC. This paradox has been solved by the recognition of the *asymptotic freedom* in QCD [13], which in particular implies that a quark looks like a quasi–free 'parton' when probed over relatively short times (so like in DIS). Some of the first applications of perturbative QCD [14] went to the field–theoretical justification of the parton picture, and to its improvement through the fecund concept of *quantum evolution* which culminated in the linear evolution equations in QCD [15, 16].

The modern era of (perturbative) QCD at high energy starts with the Great Schism between the DGLAP [15] and the BFKL [16] evolutions. The DGLAP equation [15] describes the evolution of the parton (quark and gluon) distributions with increasing virtuality Q^2 — to lowest order in the QCD coupling α, it resums the radiative corrections enhanced by factors of $(\alpha \ln Q^2)^n$ for any $n \geq 1$ — and thus proceeds through regions in phase–space which are more an more perturbative; therefore, this evolution is naturally well defined in perturbative QCD. The predictions of the DGLAP equation — including its subsequent refinements at next–to–leading–order (NLO) and, more recently, at NNLO — appear to

be very well confirmed by the experimental data at high Q^2 (in particular, by the HERA data for the DIS structure functions [17]).

The BFKL equation [16] describes the evolution of the gluon distribution, and of the scattering amplitudes mediated by gluon exchanges, with increasing energy — by resumming terms of order $(\alpha \ln s)^n$ for any $n \geq 1$ —, but it is not protected from deviations towards the non–perturbative domain at low virtuality $Q^2 \lesssim \Lambda_{\mathrm{QCD}}^2$. And as a matter of facts, the corresponding solution *does* receive contributions from soft Q^2, which become more and more important with increasing energy. This 'infrared diffusion' is a first indication of the inadequacy of the BFKL equation in the high–energy limit. The second indication in that sense comes from the behaviour of the respective solution with increasing energy: the gluon distribution, and also the scattering amplitudes computed from the BFKL equation, grow like a power of s, namely like $s^{\omega_{\mathbb{P}}}$, with $\omega_{\mathbb{P}} = (12/\pi \ln 2)\alpha$ (the 'BFKL Pomeron'). When extrapolated to the high energy limit, this behaviour violates the unitarity bound on the S–matrix, and also the Froissart bound on the total cross–sections.

The NLO corrections to the BFKL equation have recently became available [18], after several years of efforts. When properly interpreted [19–21], they considerably reduce both the value of the BFKL intercept (roughly by a factor of two) and the infrared diffusion. But this is not enough to cure the conceptual problems of the BFKL evolution in the high–energy limit. Even when amended by higher order corrections, the BFKL equation cannot be the fundamental equation in QCD at high energy. (Rather, it applies only in an intermediate range of energies.) This is so since, as we explain now, there is important physical information which is in fact missed by the BFKL equation, which is *linear*.

This is the physics of *parton saturation*, or of the *Color Glass Condensate*. In brief, the high–energy evolution is an evolution towards increasing gluon density, but the feedback of this high density environment on the evolution itself is not included in the linear, BFKL, equation. The *gluon distribution* (i.e., the total number of gluons in the hadron wavefunction) increases with s because a rise in the energy opens up the phase–space for the radiation of 'wee' gluons, with relatively small longitudinal momenta. This growth is *explosive* (exponential in $\ln s$, which is the natural 'evolution time') because of the coherency effects in the evolution: the gluons emitted in the previous steps live long enough to coherently act as sources for the emission of new gluons, with even smaller longitudinal momenta. This evolution leads to a rapid growth in the *gluon density* since the radiated gluons have comparable transverse sizes, and thus are bound to overlap with each other in the transverse plane when their total number becomes large enough. An occupation number larger than one is the hallmark of a high density system.

A high density system in QCD is genuinely *weakly coupled*. Therefore one should be able to consistently study the high–energy evolution within perturbative QCD, provided the effects associated with the high gluon density are properly taken into account.

In particular, the high density should favor *gluon recombination* processes (like two gluons merging into one), which would oppose to a further growth in the gluon distribution. This leads to the idea of *parton saturation*, as introduced by L. Gribov, Levin and Ryskin [22]. They conjectured that, when the density is large enough, the recombination processes, which are of higher order in α but are enhanced by the square of the density, should

compensate the radiation processes, which are linear in the density, in such a way that the gluon occupation number saturate at a value of order $1/\alpha$. The saturation should occur only at sufficiently low transverse momenta, below some characteristic scale, the *saturation momentum* Q_s, which rises very fast with the energy, and thus becomes *hard* ($Q_s \gg \Lambda_{\mathrm{QCD}}$) for sufficiently large s. Thus, the saturation has the potential to cure both the unitarity problem and the infrared problem of the BFKL equation. At the level of the evolution equations, this mechanism should translate into non–linear corrections to the BFKL equation; but the construction of such non–linear equations from QCD looked for longtime as an insurmountable task (see Refs. [22–24] for some early attempts).

The next major theoretical progress come with the idea of *classical coherency*, due to McLerran and Venugopalan [25] : By the principle of correspondence, a system with large occupation numbers is semi–classical; thus, it should be possible to describe the high–density gluonic system as a *coherent state* or, simpler, as a *classical color field*, with a strength of order $1/g$ at saturation. McLerran and Venugopalan (MV) introduced this idea in the form of a model for the gluon distribution of a large nucleus at low or moderate energy (as relevant for the heavy ion collisions at RHIC [26, 27]). In this model, the gluons are represented by the classical color fields radiated by sources associated with the valence quarks. It has been subsequently shown [28–30] that the gluon distribution in the MV model saturates (in the sense of growing only logarithmically with the atomic number A in the limit $A \to \infty$) via non–linear effects in the classical Yang–Mills equations.

But the classical coherency specific to the MV model is in fact preserved by the quantum evolution with increasing energy, because of the strong separation in time scales between partons with different longitudinal momenta. This observation [29] led to the development of a new approach to high–energy evolution in QCD [31–34], which is genuinely non–linear, and which culminated in the effective theory for the *Color Glass Condensate* (CGC) [34]. This new approach consists in a Wilsonian renormalization group analysis in which gluons are integrated out in layers of longitudinal momenta and in the presence of the strong color fields radiated by the gluons integrated out in the previous steps. The emission of new quantum gluons at each evolution step modifies the sources for the classical fields, and the non–linear effects induced by the latter provide the necessary feedback from the high–density environment on the evolution. The central result of this analysis is a functional evolution equation, the JIMWLK equation [31, 32, 34, 35], which describes the simultaneous evolution of all the n–point correlation functions of the classical fields (which mix with each under evolution because of the non–linear effects).

In the limit where the classical fields are weak, which corresponds to relatively low energies, the non–linear effects can be neglected and then the JIMWLK equation reduces to the BFKL evolution (not only for the gluon distribution, which is a particular 2–point function, but also for more general correlation functions, like the 3–point function describing odderon exchanges [36]). In the general case, the presence of non–linear effects entails the saturation of the gluon occupation number [37, 38] and leads to evolution equations for the scattering amplitudes which respect the unitarity constraints [34, 35]. These equations have been originally derived by Balitsky through a different approach [39].

The *Color Glass Condensate* is the high–density matter made with the saturated gluons. We have named it like that [34] to better emphasize its physical properties, and also the

mathematical structure of the effective theory which describes it :

- **Color** since the gluons carry the SU(3) 'color' charge of QCD.
- **Glass** since the associated color fields evolve very slowly relative to natural time scales, and are disordered.
- **Condensate** since the occupation numbers at saturation are of order $1/\alpha$, which is the largest value permitted by the gluon repulsive interactions. So, the saturated state is in fact a Bose condensate.

Like other glasses, the CGC is described by a *classical, stochastic, effective theory* : The color fields are classical random variables distributed according to a (functional) weight function, which is gauge invariant. Physical observables like the scattering amplitudes with external projectiles are first evaluated for a given configuration of the classical fields in the target, and then averaged over all such configurations, with the weight function alluded to above. When increasing the energy, the weight function evolves according to the JIMWLK equation, which is recognized as a functional Fokker–Planck equation [34, 35, 40].

Like other Bose condensates, the CGC is produced through an instability: the explosive, BFKL, growth of the gluon distribution with increasing energy is compensated by the gluon repulsive interactions at short distances.

Note that the glassy structure is in fact more general than saturation: It is a consequence of the separation of (temporal and longitudinal momentum) scales inherent in the high energy evolution, and thus it extends in the relatively dilute regime at transverse momenta above Q_s, that is, in the regime where the BFKL evolution applies.

Within this picture, the saturation momentum has a natural interpretation as the scale for perturbative *color neutrality* in the color glass: To minimize their repulsive interactions, the gluons arrange themselves in the transverse plane in such a way to mutually shield their color charges. Then color neutrality is achieved over a distance of order $1/Q_s$ [41, 42]. As a consequence of that, the gluon spectrum becomes softer at transverse momenta below Q_s, and there is no sensitivity to the non–perturbative, infrared, physics (unlike for the BFKL equation). The hard saturation scale Q_s acts effectively as an infrared cutoff. The physics of the CGC is consistently perturbative, the more so the higher the energy.

The importance of fluctuations for the high energy QCD evolution has been originally recognized within the dipole picture, by Mueller [43], and it became manifest in the associated Monte–Carlo simulations by Salam [44, 45]. But the profound influence that the fluctuations have on the evolution towards saturation, in particular, on the saturation scale itself, has been understood only recently [46–48], largely through a fruitful correspondence between high–energy QCD and modern problems in statistical physics. This has revived the interest in the dilute regime of QCD at high energy, and has triggered new developments leading to improved evolution equations [48–50], which include the effects of fluctuations in the large–N_c limit. It turns out that the combined effect of fluctuations and saturation corresponds to "Pomeron loops" in the diagrammatic representation of perturbative QCD.

Besides such conceptual clarifications, there is another — perhaps even more important — motivation for the nowadays interest in the CGC : this is the hope to observe this form of matter in laboratory experiments, at least indirectly. The physics of saturation can

convincingly explain some striking regularities observed in the HERA data in the 'semi–hard' domain at small Bjorken x ($x \leq 0.01$) and relatively small Q^2. Chiefly among these phenomena, there is a new scaling law at low x, known as *geometric scaling*, which has been identified in the data by Stasto, Golec-Biernat and Kwiecinski (themselves inspired by saturation), and for which the CGC formalism provides a natural, qualitative [51, 52] and even quantitative [53, 54], interpretation. We shall return to the phenomenology of gluon saturation at HERA in Sects. 1.2 and 5.1.

Furthermore, one expects the CGC to be liberated in the high–energy nucleus–nucleus collisions at RHIC (Brookhaven, USA) and LHC (CERN), and thus to determine the properties of the partonic system created right after a collision. This poses the challenge of identifying and measuring observables which can carry out information about the initial conditions, in spite of the violent evolution at later stages and the strong 'final state' interactions. Yet, it seems that some gross features of the RHIC data for multiparticle production, so like the total multiplicity or the spectrum of the hadrons produced in deuteron–gold collisions (including the remarkable phenomenon of "high–$k_\perp$ suppression" [27]) can find a natural interpretation in the framework of the CGC [55]. This will be further discussed in Sects. 1.3 and 5.2.

In view of such ongoing experimental efforts, it is of utmost importance to try and specify theoretically the expected properties of the CGC. Part of our motivations in writing this report is to contribute to this effort. Namely, the purpose of this report is to give a self contained, comprehensive, and updated discussion of the physical ideas behind the concept of color glass condensate, of the phenomenological evidence supporting this physical picture, and of the mathematical formalism used for its theoretical description.

In the recent years, several review papers have been already devoted to this topics [56], but the field is evolving so fast that some important, recent, developments (at both conceptual and phenomenological levels) are not at all covered by the previous reviews. For instance, the RHIC data on deuteron–gold collisions at forward rapidities become available during 2004 [27, 57], and their considerable impact on the CGC ideas had not been anticipated in the previous review papers. Also, on the side of the theory, significant progress has been triggered by the observation [48] that the effects of gluon number fluctuations in the dilute regime, which are known to play an important role in the QCD evolution towards high density [44, 46, 47, 58, 59], were in fact missed by the JIMWLK equation. This led to important developments, chiefly among them the recognition [47] of a profound correspondence between high–energy QCD and problems in statistical physics, and the construction of more complete evolution equations in QCD, including the 'Pomeron loops' [48–50, 60–63]. Such recent developments will be discussed at length in the present report.

Another difference with respect to previous review papers refers to the stronger emphasis that we shall put here on the mathematical aspects of the formalism and on the explicit construction of the CGC effective theory from QCD. The general formalism is quite technical and cumbersome to describe in detail, and the original papers are often lengthy and tedious, because of the need to carefully justify all the points. To cope with that, the strategy followed in the previous reviews was to avoid the construction of the formalism altogether, but only present the resulting effective theory. However, it turns out that, in order to gain more intuition about the formalism and better master its practical applica-

tions, a deeper understanding of its physical and mathematical foundations is necessary, and this is best achieved by following its derivation. The purpose of the discussion here will be to *explain and motivate* the formalism and its derivation, while still avoiding too lengthy calculations and the unessential technical details.

1.2 Motivation: The rise of the gluon distribution at HERA

Anticipated by theoretical developments within perturbative QCD already before the end of the seventies [16] (see the discussion in Sect. 1.4), the physics of high parton densities has found its first experimental motivation at the beginning of the nineties, with the advent of the HERA data for deep inelastic electron–proton scattering (DIS) at 'small–x' [64, 65] (that is, at relatively high energy; see below for more precise definitions). This was for the first time that the fine structure of a high energy hadron was revealed by an experiment, and the most striking phenomenon observed on this occasion is the rapid rise of the proton 'structure function' F_2 with decreasing x (i.e., with increasing energy): Essentially, F_2 rises as a power of $1/x$, which physically is interpreted as a rapid rise in the parton densities. This experimental fact, which was in a qualitative agreement with contemporary expectations from perturbative QCD [16], has revived the interest in the dynamics of QCD at high energy and (re)opened the debate about the fate of the hadronic matter and its interactions in the high energy limit [22, 23, 25, 43].

In this section we shall discuss the results at HERA which demonstrate that *the high energy evolution in QCD is an evolution towards high parton densities.* The theory of DIS is extensively discussed in the literature (see, e.g., Ref. [66–68] for recent presentations), and our succinct discussion below will not aim to completeness. Rather, we shall use DIS as a pretext to develop the physical picture underlying the *QCD parton description* [9, 10], which will be central to all subsequent developments, and to introduce the idea of *gluon saturation* [22]. In particular, we shall focus our discussion on the high–energy, or 'small–x', regime of DIS, thus departing from more conventional presentations which privilege the regime of high transferred momentum and the associated *DGLAP equation* [15]. This discussion will also allow us to introduce, in an intuitive way, some theoretical concepts like the *quantum evolution*, the *infinite momentum frame*, or the *eikonal approximation*, which will play a prominent role in the subsequent developments in this report.

Also, in discussing the HERA data, we shall emphasize a few phenomena which have a natural interpretation in terms of gluon saturation at small–x. One should warn at this point that, in view of the kinematical limitations of the experiment and of the present theoretical uncertainties, none of this phenomena can be taken as a conclusive evidence for saturation. Indeed, as we shall see, most of the interesting results are concentrated in the transition region from weak to strong coupling, where the analysis in terms of perturbative QCD is expected to require substantial corrections. Besides, each of these phenomena, when considered separately, could possibly be explained by some different model or mechanism. It is nevertheless remarkable that the simple idea of saturation provides a natural explanation of such a variety of seemingly different phenomena.
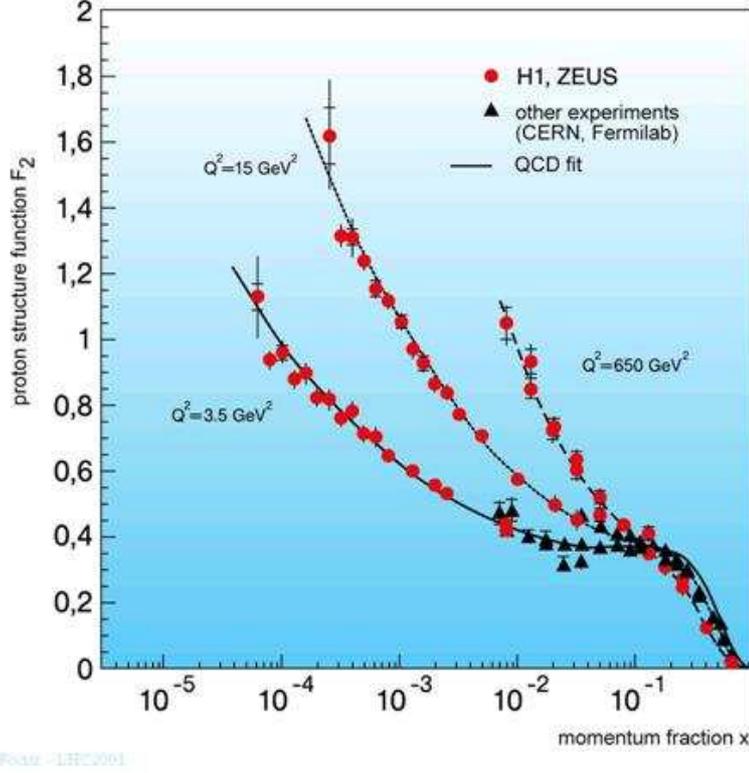
Fig. 1. *H1 and ZEUS data on the F_2 structure function shown in three bins of Q^2 as a function of x. The steep rise of the structure function at low x is clearly apparent. From Ref. [69].* HERA-F2

SECT_DIS

1.2.1 The QCD parton picture: From DIS to gluon saturation

As already mentioned, the main result at HERA as far as the small–x physics is concerned is the rise of the proton structure function $F_2(x, Q^2)$ when decreasing x at fixed Q^2, as illustrated in Fig. 1. To understand the implications of this phenomenon, one needs to know a little bit more about the kinematics of DIS and the partonic interpretation of its results. The DIS process is illustrated in Fig. 2: an electron with incident momentum k scatters off a quark inside the proton (P) with the exchange of a space–like photon γ^* with virtuality $Q^2 \equiv -q^\mu q_\mu > 0$. The scattered electron is detected, and from the measurement of its final momentum k' one can deduce the momentum $q = k' - k$ transferred by the virtual photon to the proton. The genuine "deep inelastic scattering" regime is when the invariant momentum transfer Q^2 is much larger than the proton mass squared M^2, but the kinematical range at HERA also covers the "semi–hard regime" where the energy is high, but Q^2 is of the order of, or even smaller than, M^2.

In what follows, we shall focus on the *inclusive* deep inelastic scattering, where the final hadronic system X can be anything. The quantity which is then measured in the experiment is the differential cross–section for the electron scattering $d\sigma/dQ^2$. But for the present purposes it is more convenient to take off the upper vertex in Fig. 2, which describes a purely electromagnetic process (the virtual photon splitting from the electron), and focus on the lower part of the diagram, which contains the interesting QCD process.

10
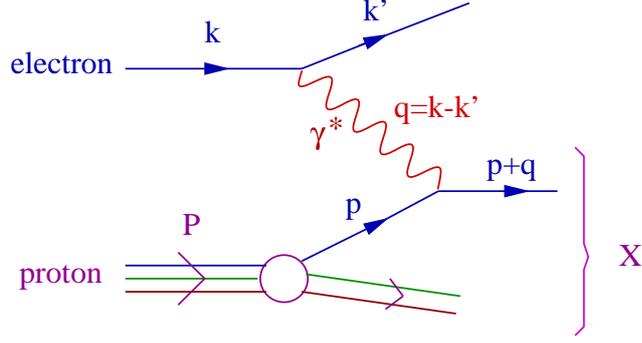
Fig. 2. *The kinematics of electron–proton deep inelastic scattering.* `DISgen`

The corresponding observable is the total cross–section $\sigma_{\gamma^* p}$ for the absorbtion of γ^* by the proton. This depends upon two independent kinematical invariants, conventionally chosen as Q^2 and the *Bjorken–x* variable:

$$x \equiv \frac{Q^2}{2P \cdot q} = \frac{Q^2}{s + Q^2 - M^2}, \tag{1.1}$$

where $s \equiv (P + q)^2$ is the invariant energy squared of the $\gamma^* p$ system. The kinematical conditions at HERA are such that at s (and generally also Q^2) is much larger than M^2. Then, Eq. (1.1) implies that $x \simeq Q^2/(s + Q^2)$ is comprised in between 0 and 1. In particular, the "small–x regime" ($x \ll 1$) is the same as the *high energy regime* ($s \gg Q^2$), which is our main interest here:

$$x \simeq \frac{Q^2}{s} \ll 1 \qquad \text{when} \qquad s \gg Q^2. \tag{1.2}$$

The measured photoabsorbtion cross–section $\sigma_{\gamma^* p}$ is conveniently parameterized as:

$$\sigma_{\gamma^* p}(x, Q^2) = \frac{4\pi^2 \alpha_{\mathrm{em}}}{Q^2} F_2(x, Q^2), \tag{1.3}$$

where $F_2(x, Q^2)$ is the *proton structure function* alluded to above. To appreciate the physical interpretation of the latter, some more preparation is necessary :

i) If the proton was a *pointlike* charged particle, so like the electron, the virtual photoabsorbtion cross section would read simply

$$\hat{\sigma}_{\gamma^* e}(x, Q^2) = \frac{4\pi^2 \alpha_{\mathrm{em}}}{Q^2} \delta(x - 1), \tag{1.4}$$

where the δ–function is the expression of the energy–momentum conservation: Namely, since both the incident and the final 'electrons' are on–shell (with negligible masses), we have $P^2 = 0$ and $(P + q)^2 = 0$, that is, $2P \cdot q - Q^2 = 0$, or $x = 1$. This argument shows that $F_2(x, Q^2)$ is a measure of the structure of the proton relative to that of a pointlike electron.

11

ii) In a conveniently chosen Lorentz frame, in which the *parton interpretation* makes sense, $F_2(x, Q^2)$ is related to the *quark and antiquark distribution functions* inside the proton. The QCD parton picture is a description of a hadron in terms of quasireal excitations (the *"partons"*) whose lifetimes are much larger than the characteristic times for the collisions. Whenever such a picture is applicable, the theoretical description of the scattering is greatly simplified because, first, one can represent the hadronic target as a kind of quantum–mechanical wavefunction (or, even simpler, as a set of classical phase–space parton distributions) and, second, one can decouple (or "factorize") the interactions with the external projectile from the internal dynamics in the target (as taken into account in the construction of the target wavefunction).

The very existence of a parton picture is quite non–trivial, as it can be appreciated from the following argument: In quantum field theory, the elementary constituents of a hadron can radiate virtual quanta (quarks and gluons in QCD) and interact with each other by exchanging such quanta. *A priori*, these virtual excitations can have arbitrarily high energies and momenta, and thus the interactions associated with them can occur on arbitrarily short time scales. In seems thus impossible to decouple the interactions within the target from those with an external projectile (like the virtual photon in DIS). Yet, in QCD, the *asymptotic freedom* ensures that virtual fluctuations with very large momenta occur with a low probability; thus, their influence on the structure of the hadron and its interactions is relatively small and computable in perturbation theory. This makes it possible to develop the parton picture (as amended by radiative corrections, or 'quantum evolution'; see below) for the short–time, or large–momentum transfer, processes. But this also requires a judicious choice of the Lorentz frame, as we explain now:

Consider first the proton rest frame, where $P^\mu = (M, 0, 0, 0)$. The characteristic time over which the partons interact with each other is of order $1/M$ (or, more precisely, $\sim 1/\Lambda_{\text{QCD}}$), since this is the scale at which the QCD interactions become strong. This internal time scale has to be compared with the duration of the collision with the virtual photon, which can be estimated from the uncertainty principle as:

$$\Delta t_{\text{coll}} \sim \frac{1}{q_0} = \frac{M}{P \cdot q} = \frac{2Mx}{Q^2}. \tag{1.5}$$

In the deep inelastic regime where $Q^2 \gg M^2$, we also have $\Delta t_{\text{coll}} \ll 1/M$, showing that the scattering is indeed rapid compared to the internal time scales of the proton.

However, this condition is not *sufficient* for the existence of a parton description, and in fact the latter does *not* exist in the proton rest frame, since in this frame the hadron excitations cannot be distinguished from *vacuum fluctuations* having similar momenta. In Fig. 3 we illustrate some virtual excitations of a fermion inside the hadron (say, a valence quark) together with a virtual fluctuation of the vacuum. The typical momenta are of the same order for all such fluctuations, namely, of order Λ_{QCD}; therefore, they all have a lifetime of order $1/\Lambda_{\text{QCD}}$ and can mix with each other. In particular, the interaction with a virtual photon, which is also shown in Fig. 3 (and which takes a comparatively short time), cannot distinguish between partons and vacuum excitations.

But there exists a way to disentangle the hadron fluctuations from those of the vacuum: This consists in performing a Lorentz boost to a frame in which the hadron has a very
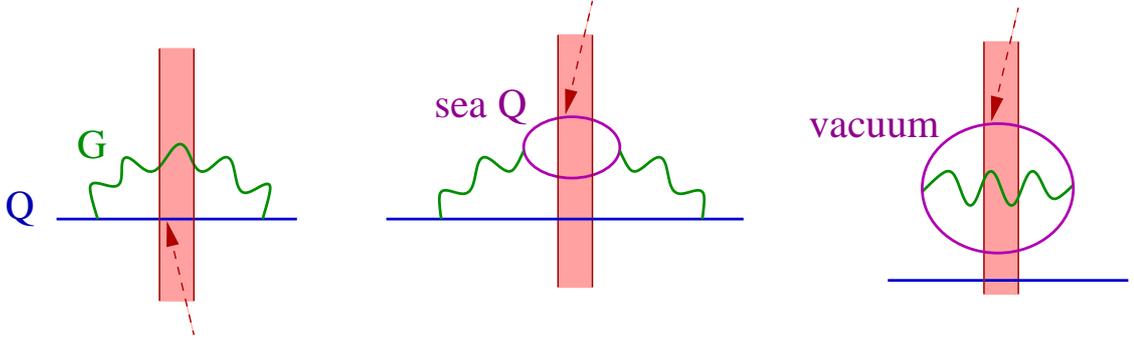
Fig. 3. *Gluon (G), 'sea quark' (sea Q), and vacuum fluctuations dressing a valence quark (Q) in the hadron wavefunction in its rest frame. The dotted arrow is meant to represent the interaction with the virtual photon in DIS, and the width of the shaded area around the arrow is suggestive of the duration of the scattering.* `RFfluct`

large longitudinal momentum $P_z \gg M$ — an *infinite momentum frame* (IMF) — and where the lifetime of its virtual excitations is enhanced by Lorentz *time dilation*. Note that the four–momentum of the hadron in such an IMF reads $P^\mu = (E, 0, 0, P_z)$ with $E \simeq P_z + M^2/2P_z \simeq P_z$. In what follows, we shall write simply $P_z \equiv P$. Consider then a fluctuation of a quark inside the hadron as seen in the IMF (see e.g. Fig. 5). All the quanta involved in this fluctuation have large longitudinal momenta, since they carry a non–zero fraction of the corresponding total momentum P. Let us assume, for simplicity, that the longitudinal momentum of the original fermion is exactly P (this is the 'hadron'), and compute the lifetime of the gluon fluctuation with four–momentum $k^\mu = (k_0, k_\perp, k_z)$, where $k_z = \xi P$ with $0 \le \xi \le 1$ and $k_\perp \ll \xi P$. This can be estimated by the uncertainty principle as $\Delta t \sim 1/\Delta E$, where ΔE is the energy difference between the final and initial states at the gluon emission vertex. We have (we neglect the parton masses) :

$$
\begin{aligned}
\Delta E &= -P + \sqrt{\xi^2 P^2 + k_\perp^2} + \sqrt{(1-\xi)^2 P^2 + k_\perp^2} \\
&\approx \frac{k_\perp^2}{2\xi P} + \frac{k_\perp^2}{2(1-\xi)P} \approx \frac{k_\perp^2}{2\xi P},
\end{aligned}
\tag{1.6}
$$

where in writing the last approximate equality we have assumed that $\xi \ll 1$; as we shall shortly see, this is the interesting case at high energy. This implies $\Delta t \sim 2\xi P/k_\perp^2$, which should be compared to the lifetime $\Delta t_{\rm vac} \sim 1/k_z$ of a vacuum fluctuation having a similar longitudinal momentum. (Indeed, for a vacuum bubble like that in Fig. 3, the energy denominator involves the sum of the on–shell energies of the particles in the intermediate state, so $\Delta E_{\rm vac} \sim k_z$.) So long as the condition $k_\perp \ll \xi P$ remains satisfied, we have

$$
\Delta t \sim \frac{2\xi P}{k_\perp^2} \;\gg\; \Delta t_{\rm vac} \sim \frac{1}{\xi P},
\tag{1.7}
$$

which confirms that, in the IMF, the typical hadron fluctuations are well separated from the vacuum fluctuations, and therefore the parton picture makes sense indeed.

The above considerations further imply that a parton in the IMF is *quasi–real* — its

13

virtuality $k^2 \simeq (\Delta E)^2$ is much smaller than the on–shell energy squared $\xi^2 P^2$ — and also *nearly collinear* with the hadron: $k_0 \simeq k_z = \xi P \gg k_\perp$, and therefore $k^\mu \approx \xi P^\mu$.

Let us now discuss the physical picture of DIS when viewed in the IMF. We shall shortly check that the lifetime of the struck quark is much larger than the collision time. Thus, this quark is nearly on–shell both before and after the scattering. If $k^\mu \approx (\xi P, 0, 0, \xi P)$ denotes its original four–momentum, then energy–momentum conservation implies

$$0 \approx (k + q)^2 = -Q^2 + 2\xi P \cdot q = -2P \cdot q(x - \xi) \quad \Longrightarrow \quad \xi \approx x. \tag{1.8}$$

Thus, the scattering selects a quark with a longitudinal momentum fraction ξ equal to Bjorken's x. This confirms the fact that, in DIS at high energy, we probe the proton constituents having $\xi \ll 1$. From now on, we shall not distinguish between x and ξ anymore, but denote both these quantities by x.

To further develop the physical picture of DIS, let us choose a specific IMF, known as the 'Breit frame', in which the virtual photon has zero longitudinal momentum: $q^\mu = (q_0, q_\perp, 0)$. Since:

$$q_0 = \frac{P \cdot q}{P} \to 0 \qquad \text{as} \qquad P \to \infty, \tag{1.9}$$

we deduce that, in this frame, the γ^* momentum is mainly transverse: $q_0 \ll q_\perp$ and $Q^2 \simeq q_\perp^2$. Thus, by experimentally varying the virtuality Q^2, we can control the resolution of the virtual photon in the transverse plane. The collision time in this particular frame can be estimated from the energy uncertainty ΔE at the absorption vertex, or simply as the reciprocal of the frequency q_0. Both methods yield

$$\Delta t_{\text{coll}} \sim \frac{2Px}{Q^2}. \tag{1.10}$$

For a quark fluctuation to be able to absorb the virtual photon, its lifetime must be larger than the collision time. This condition, together with Eqs. (1.7) and (1.10), imply that the partons which participate in DIS have transverse momenta $k_\perp^2 \lesssim Q^2$. By the uncertainty principle, such partons are localized within an area $\sim 1/Q^2$ in the transverse plane.

We have thus arrived at the important conclusion that the DIS experiment is a very specific analyzer of the *quark distribution* in the proton, as measured in the Breit frame: The virtual photon 'counts' the number of quarks and antiquarks having a longitudinal momentum fraction equal to x and which occupy a transverse area of order $1/Q^2$. These considerations motivate the following formula for the F_2 structure function:

$$F_2(x, Q^2) = \sum_f e_f^2 \left[x q_f(x, Q^2) + x \bar{q}_f(x, Q^2) \right], \tag{1.11}$$

where the sum runs over all the quark flavors, and $q_f(x, Q^2)\mathrm{d}x$ is the number of quarks of flavor f with longitudinal fraction between x and $x + \mathrm{d}x$ and which are localized in the transverse space within a size $1/Q$; similarly, $\bar{q}_f(x, Q^2)$ refers to the antiquarks. The quantity $x q_f(x, Q^2)$ is conventionally referred to as the *quark distribution function*; it

can be related to the gluon number density in Fock space[1] $\mathrm{d}N/\mathrm{d}k_z\mathrm{d}^2k_\perp$ via (note that $\mathrm{d}x/x = \mathrm{d}k_z/k_z$):

$$xq_f(x, Q^2) \,=\, \int^{Q^2} \mathrm{d}^2k_\perp \; x\frac{\mathrm{d}N_f}{\mathrm{d}x\,\mathrm{d}^2k_\perp}. \tag{1.12}$$

In the conventional (pre–QCD) 'parton model', the transverse momenta of the partons are restricted to $k_\perp \ll M$, so for $Q^2 > M^2$ the corresponding quark distribution is independent of Q^2. Then, Eq. (1.11) implies that F_2 'scales' as a function of x : when plotted as a function of Bjorken's x, the structure function is independent of Q^2. This prediction of the parton model, known as *Bjorken scaling*, has been indeed confirmed by the first DIS experiment at SLAC, but only as an *approximate* property of the data, and within a limited kinematical range (concentrated at rather large values of x). This is illustrated in Fig. 4 with the more recent (and also more accurate) data at HERA.
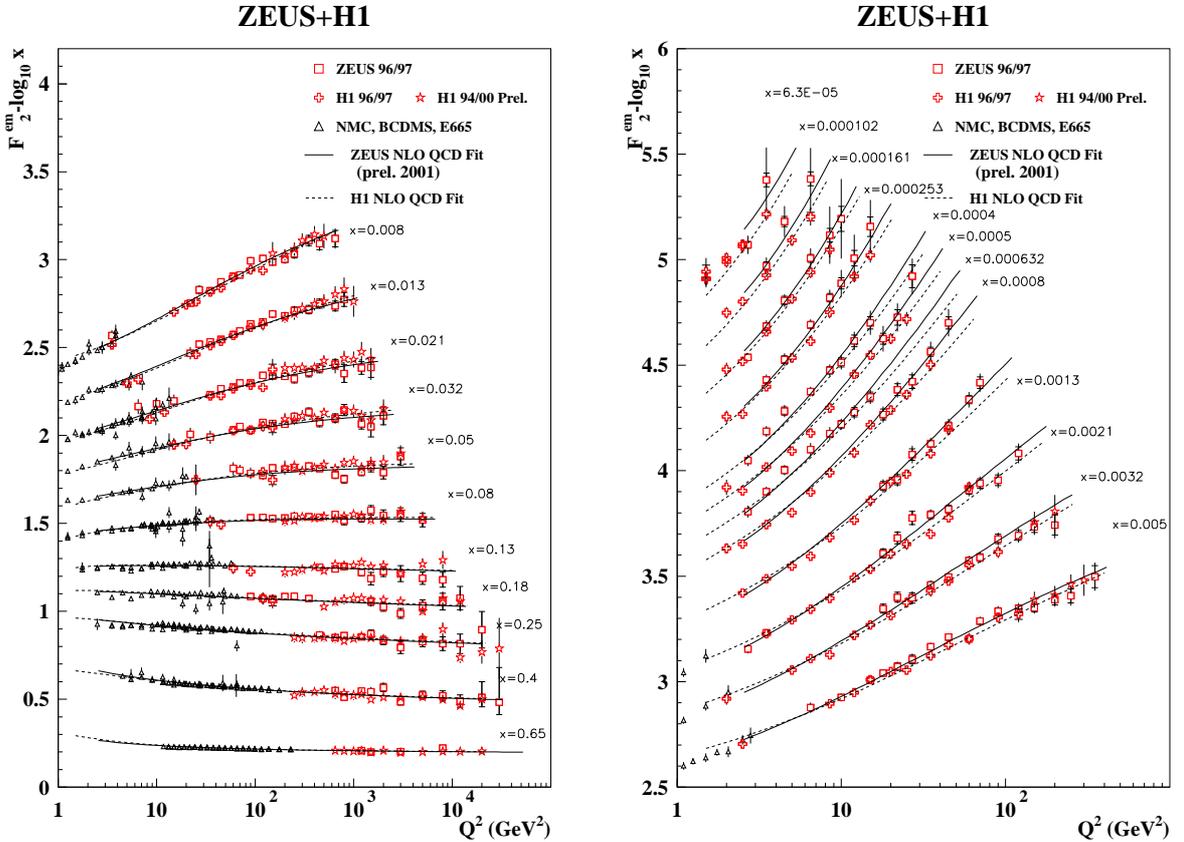
Fig. 4. The F_2 structure function as measured by the H1 and ZEUS experiments in bins of x as a function of Q^2. Left: The bins at relatively high x (in particular, it was around $x = 0.25$ that Bjorken scaling was originally observed in the SLAC experiments). Right: The bins at low x; the violations of Bjorken scaling are now obvious. HERA-BJ

However, already the qualitative discussion in this section shows that, in QCD, all the

---

[1] This quantity is defined in a specific gauge, the *light–cone gauge* ; see Sect. 2.4 for details.
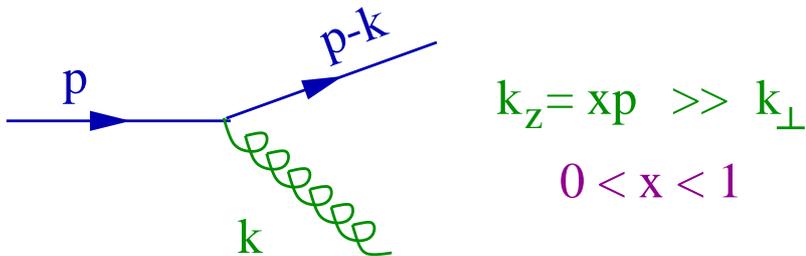
Fig. 5. An elementary radiation process in the proton wavefunction: the original quark with momentum p splits into a quark $p - k$ and a gluon k. SPLITT

virtual quanta having $k_\perp^2 \lesssim Q^2$ live long enough to contribute to the scattering. Thus, when increasing Q^2, we expect the number of the intervening partons to increase as well (because of the enlarged phase–space for radiation), which should translate into a rise of F_2 with Q^2 at fixed x. Such a rise is clearly visible in Fig. 4 for the data corresponding to small values of x (say, $x \leq 0.03$). Physically, increasing Q^2 is like improving the resolution of a microscope (here, in both space and time) : By doing so, we are able to discriminate partons which are more and more localized in the transverse space, and whose lifetime is shorter and shorter.

The violation of Bjorken scaling is a first example of *quantum evolution* in QCD — here, the evolution with increasing Q^2 — by which one generally means the variation of the structure functions with Q^2 and x induced through radiative processes in the underlying quantum field theory. In what follows, we shall be mostly interested in the high energy (or "small–x") evolution, which is responsible for the rapid rise of F_2 with decreasing x, cf. Fig. 1. To understand the latter, and also the fact that the corresponding rise with Q^2, i.e., the logarithmic slope $dF_2/d\ln Q^2$, is more pronounced at small values of x, cf. Fig. 4, it is necessary to consider also the *gluon distribution*, although this is not directly measured by DIS. Indeed, quarks and gluons mix under the evolution, and it is in fact the rapid rise of the gluon density which is the driving force for the evolution at small x.

Indeed, as we shall argue now, most of the partons produced by the evolution with decreasing x are in fact gluons. This is the consequence of the infrared singularity of the amplitude for bremsstrahlung. The differential probability for an on–shell parton to radiate a soft gluon which carries a small fraction $x \ll 1$ of the parton longitudinal momentum and a small transverse momentum $k_\perp \ll k_z = x p_z$ is given by (see e.g. Fig. 5)

$$dP_{\text{Brem}} \simeq \frac{\alpha C_R}{\pi^2} \frac{d^2 k_\perp}{k_\perp^2} \frac{dx}{x}, \tag{1.13}$$

where C_R is the Casimir for the SU(3) representation appropriate to the parent parton; that is, $C_F = (N_c^2 - 1)/N_c$ for a parent quark, or $C_A = N_c$ for a parent gluon. Note first the logarithmic enhancement in Eq. (1.13) at small transverse momenta: this is true for any kind of (soft) emitted quanta — quarks or gluons — and implies that the partons which participate to DIS are *logarithmically* distributed within the range $0 < k_\perp^2 < Q^2$, so that their typical momenta satisfy $k_\perp^2 \ll Q^2$. This observation is at the heart of the

DGLAP evolution [15] of the structure functions with increasing Q^2.

But here we are more interested in the evolution with decreasing x, and the crucial observation in that respect is the logarithmic enhancement in Eq. (1.13) at small x. This is related to the fact that the soft emitted particle is a gluon. (By contrast, the probability for the emission of a soft quark is non–singular at $x \to 0$.) This implies that gluon radiation is enhanced as compared to quark radiation in the quantum evolution with decreasing x. Therefore the quark 'seen' by the virtual photon in the small–x regime of DIS is typically produced by the dissociation of a gluon. In mathematical terms, and to lowest order in QCD perturbation theory (at small x and relatively high Q^2), the latter statement is encoded in the following formula:

$$\frac{\partial F_2(x, Q^2)}{\partial \ln Q^2} \simeq \frac{\alpha}{3\pi} \Big(\sum_f e_f^2 \Big) xG(x, Q^2) \,, \tag{1.14}$$

where $xG(x, Q^2)$ is the *gluon distribution function*, defined similarly to the quark distribution (see Sect. 2.4 below for a more precise definition), and by the previous arguments is expected to rise with decreasing x. Eq. (1.14) shows that, at small x, the gluon distribution is directly related to the Bjorken scaling violation in the structure function F_2.

By using Eq. (1.14) (or its more accurate versions including next–to–leading order perturbative corrections in QCD), one has been able to extract the gluon distribution from the F_2 structure function measured at HERA. Some results of this analysis are exhibited in Figs. 6 and 7 [17], and they show indeed a rapid rise of $xG(x, Q^2)$ with decreasing x at fixed Q^2. In particular, Fig. 7 — in which the gluon distribution is compared to the (valence and sea) quark distributions — makes it manifest that gluons dominate the proton wavefunction at small x.

From Fig. 7, we also see that the rise of the gluon distribution with $1/x$ is more pronounced at larger values of Q^2, as expected from the DGLAP evolution [15]. Still, for a given (small) value of x, the gluon *occupation factor* is larger at *smaller* values of Q^2, as we now explain: Recall indeed that in DIS we measure partons with a typical transverse area $1/Q^2$ and which carry a longitudinal momentum fraction equal to x. At small x (say $x < 0.01$), such partons have a low longitudinal momentum $k_z = xP \ll P$, and by the uncertainty principle they are delocalized over a large distance $\sim 1/xP$ in the z direction; that is, they extend well outside the Lorentz contracted disk of width $\sim 1/P$ which is occupied by the valence quarks. Then, the relevant density for the dynamics of the 'wee' gluons is their density in the *transverse plane* :

$$n_g(x, Q^2, b_\perp) \equiv x \frac{\mathrm{d}N_g}{\mathrm{d}x\mathrm{d}^2 b_\perp} \approx \frac{xG(x, Q^2)}{\pi R^2} \,, \tag{1.15}$$

where $b_\perp$ is the position of the measured gluon in the transverse plane, and is a meaningful quantity so long as $1/Q^2$ is much smaller than the typical distance scale for transverse inhomogeneity in the proton. In writing the last estimate in Eq. (1.15) we have treated the proton as a homogeneous disk of radius R. Each gluon has an area $\sim 1/Q^2$, so the total area occupied by the gluons probed in DIS is $\sim xG(x, Q^2)/Q^2$. Clearly, if this area is larger than the proton transverse area πR^2, then the gluons must overlap with each
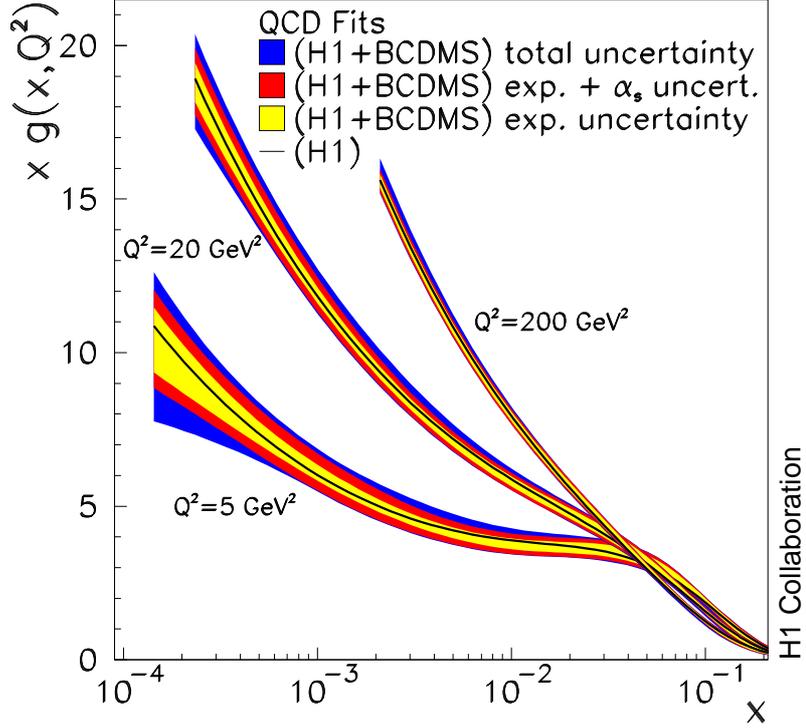
Fig. 6. *Gluon distribution as extracted at HERA, as a function of x in three bins of Q^2. Most of the total error band originates in the theoretical error concerning the relation between F_2 and $xG(x, Q^2)$.* `HERA-gluon`

other. This is measured by the dimensionless *gluon overlapping factor*

$$\varphi(x, Q^2, b_\perp) \approx \frac{1}{N_c^2 - 1} \frac{xG(x, Q^2)}{Q^2 \, \pi R^2}, \tag{1.16}$$

(for gluons of a given color), which is a coarse–grained version of the gluon phase–space occupation factor that we shall more precisely define later. The gluons overlap with each other when $\varphi > 1$.

Now, assume that we increase the energy at fixed Q^2 (so that we decrease x): Then, $xG(x, Q^2)$ increases very fast, cf. Fig. 6, meaning that we observe more and more gluons which all occupy more or less the same area. Since, in the same process, the transverse extent of the proton rises only slowly (the proton radius R increases at most logarithmically with the energy; see Sect. 4.4), we conclude that the system evolves towards higher and higher occupation factors, i.e., towards a *high density regime*.

If, on the other hand, we increase Q^2 at fixed x, then the number $xG(x, Q^2)$ of gluons is rising again, but the *total area* occupied by them is in fact decreasing, and rather fast, because the increase in the gluon number is more than compensated by the rapid decrease in the area $\sim 1/Q^2$ occupied by each gluon. (According to the DGLAP equation, the rise of $xG(x, Q^2)$ with Q^2 is slower than a power of Q^2.) Therefore, when increasing Q^2, the gluon distribution evolves towards lower occupation factors, i.e., towards a *dilute regime*.

These general considerations on the QCD evolution are illustrated in Fig. 8. As indicated
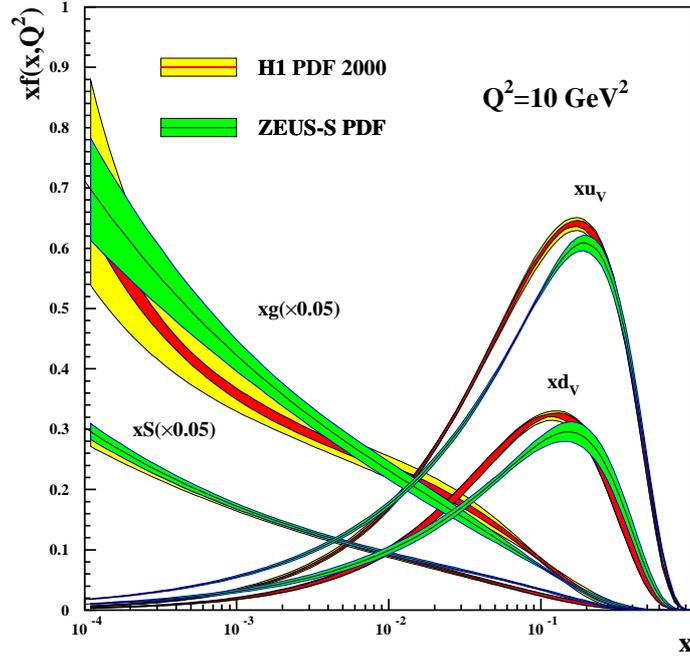
Fig. 7. The valence quark (xu_v, xd_v), gluon (xG) and sea quark (xS) distributions in a proton, as measured at HERA. Note that the gluon distribution is scaled down by a factor of 20. HERA-GQ

in this figure, the natural variables to describe the evolution are the logarithmic variables $\tau \equiv \ln 1/x$ (this is generally referred to as the *rapidity*) and $\rho \equiv \ln Q^2/Q_0^2$ (with Q_0^2 an arbitrary reference scale). Indeed, as it should be clear from the bremsstrahlung law, Eq. (1.13), the evolution is logarithmic in both x and Q^2.

The fact that the high energy evolution in QCD leads to a system with high gluon density is of fundamental importance: It is intuitively clear that a high–density system of mutually interacting particles (so like the gluons) must reach some equilibrium state in which the typical momenta carried by the particles should scale like a fractionary power of the density. Here, the gluons make up a two–dimensional distribution, so we expect $k_\perp \sim n^{1/2}$. Thus, if the density is sufficiently high, the typical momenta are *hard* ($k_\perp \gg \Lambda_{\mathrm{QCD}}$), and by asymptotic freedom the gluonic matter must be *weakly coupled*. We thus arrive at the important conclusion that *the high–energy regime of QCD is characterized by weak coupling* ($\alpha \ll 1$), and should be therefore accessible to theoretical investigations from first principles. But in spite of the weak coupling, the study of the high energy gluonic matter remains a formidable problem, because of the complexity of the collective phenomena entailed by the high density.

The following considerations should clarify the physical mechanism leading to such a high–density equilibrium state [22]. We have previously argued that (i) the gluon density rises rapidly with decreasing x, through radiation from 'color sources' (quarks or gluons) at larger values of x, and (ii) the small–x gluons overlap with each other, a situation which favors their mutual interactions. Since suppressed by powers of $\alpha \ll 1$, the gluon
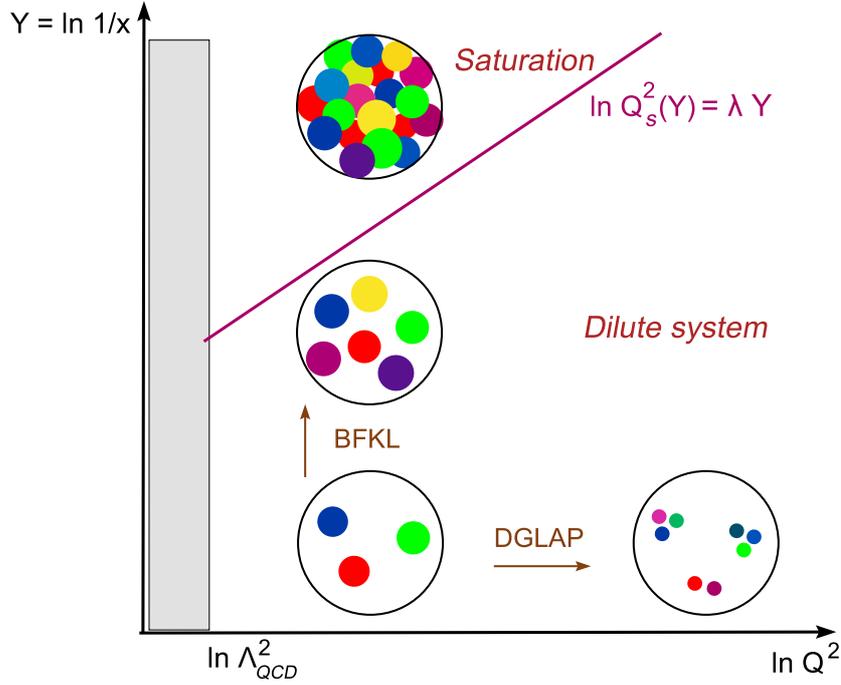
Fig. 8. *A cartoon of the QCD evolution in the kinematical plane for deep inelastic scattering. The BFKL evolution and the concept of Color Glass Condensate will be explained later.* `QCD-map`

interactions remain negligible so long as the density is not high enough; this is the case at relatively large transverse momenta, where the quantum evolution is dominated by radiative processes leading to a rapid increase in the gluon density with $1/x$ (see Sect. 1.4 below). However, with increasing density, *recombination processes* like two gluons merging into one become increasingly important, and start to oppose to the further growth of the gluon distribution. Since the efficiency of the recombination rises faster with $1/x$ than that of the radiation, an equilibrium state must be eventually reached, in which the two processes compensate each other. When this happens, the gluon density *saturates*, i.e., it does not grow with $1/x$ anymore. With this physical picture in mind, it is rather easy to estimate the critical occupation number at which the recombination processes start to become important, and saturation is expected.

Namely, starting in the dilute regime and increasing the energy at fixed Q^2, one should eventually approach a critical regime where the recombination rate for a gluon with longitudinal fraction x and transverse size $1/Q^2$ becomes of order one. The recombination probability is the same as the probability that the respective gluon be absorbed by any other gluon in the hadron. This process is similar to the virtual photon absorbtion in DIS, and the corresponding cross–section can be estimated by analogy to Eq. (1.3). The recombination rate $\Gamma(x, Q^2)$ is then obtained by dividing the cross section through the 'volume', here, the hadron transverse area. This finally yields:

$$\Gamma(x, Q^2) \sim \frac{\alpha N_c}{Q^2} \frac{x G(x, Q^2)}{(N_c^2 - 1)\pi R^2}, \tag{1.17}$$

20

which becomes of order one when the occupation factor (1.16) becomes of order $1/\alpha$:

$$\Gamma(x, Q^2) \sim 1 \quad \Longleftrightarrow \quad \varphi(x, Q^2, b_\perp) \sim \frac{1}{\alpha N_c}. \tag{1.18}$$

This *saturation condition* defines a 'critical line' relating Q^2 and x, namely $Q^2 = Q_s^2(x, b_\perp)$ with the *saturation momentum* $Q_s(x, b_\perp)$ implicitly defined by [22–25]

$$Q_s^2(x, b_\perp) \approx \frac{\alpha N_c}{N_c^2 - 1} \frac{xG(x, Q_s^2)}{\pi R^2}. \tag{1.19}$$

(A more precise definition for Q_s will be given later.) The saturation momentum is the fundamental scale in QCD at high energy. It separates between a dilute regime at $Q^2 \gg Q_s^2$, where the gluon occupation factor is relatively low, $\varphi \ll 1/\alpha$, and a high–density regime at $Q^2 \lesssim Q_s^2$, where φ saturates at a value of order $1/\alpha$.

Since the gluon distribution rises rapidly with $1/x$, a similar behaviour is expected also for the saturation momentum. If x is low enough, one should have $Q_s^2(x) \gg \Lambda_{\mathrm{QCD}}^2$, thus justifying the use of weak coupling techniques. We shall see later that, within perturbative QCD, Q_s^2 grows like a power of $1/x$, i.e., like an exponential of τ. Accordingly, the *saturation line* $\rho_s(\tau) \equiv \ln[Q_s^2(x)/Q_0^2]$ in the $(\tau - \rho)$–plane is a *straight line*, as illustrated in Fig. 8. It will be a main objective for us in what follows to justify the simple physical picture of saturation introduced above within perturbative QCD, and compute the slope of the saturation line (also known as the *saturation exponent*). But before doing that, let us return to the experimental situation at HERA and search for traces of saturation.

1.2.2 Dipole factorization and geometric scaling

As discussed in Sect. 1.2.1, the quark probed by the virtual photon in DIS at small x is typically a sea quark which is emitted by the small–x gluons in the proton. It is then convenient to disentangle this final quark emission from the small–x dynamics of the hadron wavefunction, which involves mostly gluons. This can be done by performing a Lorentz boost in such a way to pull the $\gamma^* q\bar{q}$ vertex out of the hadron (see Fig. 9). That is, instead of the hadron Breit frame, it is preferable to use the so-called *dipole frame* [70] in which most of the energy is still carried by the hadron (so that the high density effects are associated with the hadron wavefunction), but the virtual photon has enough energy to dissociate long before scattering into a quark–antiquark pair in a color singlet state — a *color dipole* — which then scatters off the hadron. This sequential picture of DIS is appropriate at high energy, since the lifetime of the $q\bar{q}$ pair is much larger than the interaction time between this pair and the hadron. It is furthermore useful since it offers a natural framework for the description of *multiple scattering*, which becomes important when the gluon density in the proton is high enough.

More precisely, the dipole frame is defined by $P^\mu \simeq (P, 0_\perp, P)$ and $q^\mu = (\sqrt{q^2 - Q^2}, 0_\perp, -q)$, where $q \gg Q$ but such that $\alpha \ln(q/Q) \ll 1$. In this frame, the virtual photon wavefunction can be written as the sum of a 'bare' photon plus hadronic fluctuations:

$$|\gamma^*\rangle = c_0|\gamma\rangle_0 + c_1|q\bar{q}\rangle_0 + c_2|q\bar{q}g\rangle_0 + \cdots, \tag{1.20}$$
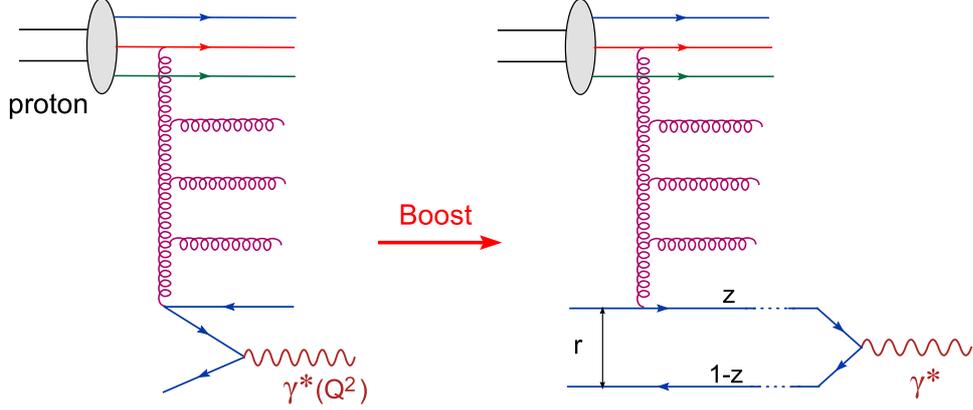
Fig. 9. *Illustration of the Lorentz boost leading from the Breit frame to the dipole frame.* `DIPOLEframe`

with $c_0 \sim 1$, $c_1 \sim \alpha_{\rm em}$, $c_2 \sim \alpha_{\rm em} \times \alpha \ln(q/Q)$, etc. To lowest order in $\alpha_{\rm em}$, the γ^*–hadron scattering is controlled by the 'color dipole' fluctuation $q\bar{q}$. The condition $q \gg Q$ ensures that this fluctuation has a relatively long lifetime (γ is the Lorentz factor) :

$$\Delta t_{q\bar{q}} \sim \frac{1}{E_{q\bar{q}} - E_{\gamma^*}} \gg \Delta t_{\rm coll} \sim R/\gamma \sim 1/P, \qquad (1.21)$$

and therefore the transverse coordinates $\boldsymbol{x}$ and $\boldsymbol{y}$ of the quark and, respectively, the antiquark which compose the dipole are 'frozen' during the interaction with the hadron. This makes it convenient to develop a coordinate–space picture of the collision. On the other hand, the condition $\alpha \ln(q/Q) \ll 1$ ensures that the QCD quantum evolution of the dipole itself is negligible, so that one can ignore higher Fock states like $|q\bar{q}g\rangle_0$: the dipole is a 'bare' $q\bar{q}$ pair without additional gluons.

This physical picture translates into the following factorization formula for σ_{γ^*p} in the dipole frame [71, 72] :

$$\sigma_{\gamma^*p}(\tau, Q^2) = \int_0^1 {\rm d}z \int {\rm d}^2\boldsymbol{r}\, |\Psi(z, \boldsymbol{r}; Q^2)|^2\, \sigma_{\rm dipole}(\tau, \boldsymbol{r}). \qquad (1.22)$$

Here, $|\Psi(z, \boldsymbol{r}; Q^2)|^2 \sim \alpha_{\rm em}$ is the probability density for the photon to split into a $q\bar{q}$ pair with transverse size $\boldsymbol{r} = \boldsymbol{x} - \boldsymbol{y}$ and a fraction z of the photon's longitudinal momentum carried by the quark. ($\Psi(z, \boldsymbol{r}; Q^2)$ is a light–cone wavefunction which can be computed within perturbative QED [71, 72].) Furthermore, $\sigma_{\rm dipole}(\tau, \boldsymbol{r})$ is the dipole total cross section, and can obtained via the optical theorem as

$$\sigma_{\rm dipole}(\tau, \boldsymbol{r}) = 2 \int {\rm d}^2\boldsymbol{b} \left[1 - {\rm Re}\, S_\tau(\boldsymbol{r}, \boldsymbol{b}) \right], \qquad (1.23)$$

where $S_\tau(\boldsymbol{r}, \boldsymbol{b})$ is the S–matrix element for the dipole–hadron scattering at a given impact parameter $\boldsymbol{b} = (\boldsymbol{x} + \boldsymbol{y})/2$. The integration in Eq. (1.23) covers the hadron disk.

As we explain now, there are two important physical regimes for dipole–hadron scattering, depending upon the size $r_\perp \equiv |\boldsymbol{r}|$ of the dipole relative to the hadron saturation length

22

$1/Q_s(\tau)$. Here, $Q_s(\tau)$ is the target saturation momentum averaged over all the impact parameters which contribute significantly to scattering.

i) A *small* dipole, with size $r_\perp \ll 1/Q_s(\tau)$, 'sees' the tail of the gluon distribution at relatively high transverse momenta $Q_s(\tau) \ll k_\perp \ll 1/r_\perp$. Since the gluon density is low in this tail, the dipole undergoes only *single scattering*, via two–gluon exchange with any of the available gluons. The corresponding cross–section is simply proportional to the gluon distribution evaluated at $Q^2 \sim 1/r_\perp^2$:

$$\sigma_{\text{dipole}}(\tau, r_\perp) \simeq \alpha r_\perp^2 \frac{\pi^2}{N_c} xG(x, 1/r_\perp^2) \qquad \text{for} \qquad r_\perp \ll 1/Q_s(\tau)\,. \tag{1.24}$$

The vanishes like $r_\perp^2$ when $r_\perp \to 0$, a property known as *color transparency* which reflects the fact a zero–size dipole cannot interact, as it carries no color charge. From Eq. (1.24) we conclude that a small dipole is a very direct probe of the gluon distribution in the target: it essentially 'counts' the gluons having a longitudinal momentum fraction $x = \mathrm{e}^{-\tau}$ and a transverse size $\sim r_\perp$.

ii) A *large* dipole with $r_\perp \gtrsim 1/Q_s(\tau)$ 'feels' the high–density system of gluons with momenta $k_\perp \lesssim Q_s(\tau)$, off which it undergoes *multiple scattering*. For sufficiently large $r_\perp$, one expects the dipole to be completely absorbed, so that the corresponding cross–section saturates the geometric, or 'black disk', limit (R is the target radius):

$$\sigma_{\text{dipole}}(\tau, r_\perp) \simeq 2\pi R^2 \qquad \text{for} \qquad r_\perp \gg 1/Q_s(\tau)\,. \tag{1.25}$$

According to Eq. (1.23), this implies that the scattering amplitude approaches the *unitarity limit* $S_\tau(\boldsymbol{r}, \boldsymbol{b}) = 0$ at all the points $\boldsymbol{b}$ which contribute to scattering. (We shall argue later that the 'black disk' expands in transverse directions when increasing the energy, but this expansion is only logarithmic in $1/x$, and thus much slower than the rapid, power–like, increase with $1/x$ of the cross–section for single scattering, Eq. (1.24).)

By comparing the above equations, one can check that the transition between the two regimes takes place indeed at dipole sizes of the order of the saturation length: for $r_\perp \sim 1/Q_s(\tau)$ with $Q_s(\tau)$ obeying Eq. (1.19), the cross–section for single scattering, Eq. (1.24), becomes comparable to the geometric cross–section (1.25). Thus, by studying the onset of unitarity corrections in the dipole–hadron scattering, one can determine the characteristic scale for gluon saturation in the hadron wavefunction. This reflects the fact that saturation and multiple scattering are different aspects of the same non–linear physics.

Later on, we shall describe the calculation of the dipole scattering amplitude within perturbative QCD (in the presence of unitarity corrections) and thus justify equations like (1.24) and (1.25). Here, however, we shall discuss a simple phenomenological model, due to Golec-Biernat and Wüsthoff [73], which incorporates the basic physics alluded to above, and which has played an important role in the evolution of the ideas about saturation in that it has provided the first compelling interpretation of the HERA data in terms of saturation. The original "saturation model" in Ref. [73] is encoded in the following parametrization for the dipole cross–section:

$$\sigma_{\text{dipole}}(x, r_\perp) = \sigma_0 \Big(1 - \mathrm{e}^{-r_\perp^2 Q_s^2(x)/4}\Big), \tag{1.26}$$

where $Q_s^2(x) = Q_0^2(x_0/x)^\lambda$, $Q_0 = 1$ GeV, and σ_0, x_0 and λ are free parameters, to be extracted from a fit to the HERA data for F_2 at small x.

As anticipated, Eq. (1.26) encodes the basic physics of saturation: it interpolates between 'color transparency' ($\sigma_{\text{dipole}} \propto r_\perp^2$) at small dipole sizes and the geometric cross–section ($\sigma_{\text{dipole}} \approx \sigma_0$) at large dipole sizes, with the transition between the two regimes taking place at a scale $r_\perp \sim 1/Q_s(x)$ which decreases as a power of x. At the same time, Eq. (1.26) has some obvious shortcomings: It ignores the subtleties of the QCD evolution both in the dilute regime, where it fails to describe the Bjorken–scaling violation, and in the transition region towards saturation, where we know that 'color transparency' is modified by the BFKL evolution already for small dipole sizes $r_\perp < 1/Q_s(x)$ (see Sect. 1.4 below). Besides, as we shall see in Sect. 4, the approach towards the unitarity limit within perturbative QCD is much slower than suggested by Eq. (1.26).

But in spite of these conceptual limitations — which have been subsequently corrected in more refined analyses (see below) — the simple formula (1.26) has been shown [73] to provide a rather good fit, via Eq. (1.22), to the original HERA data [65] for $\sigma_{\gamma^*p}(x, Q^2)$ at $x < 10^{-2}$ and for low and intermediate Q^2 (up to about 100 GeV^2; after adding a fourth parameter, the fit has been extended down to $Q^2 = 0$). The values of the three parameters extracted from the original fit are: $\sigma_0 = 23$ millibarns, $\lambda = 0.288$ and $x_0 = 3.04 \cdot 10^{-4}$ [73]. Remarkably, the saturation scale thus obtained is moderately hard ($Q_s^2(x) > 1$ GeV^2 for $x \leq 10^{-4}$), which suggests that a perturbative QCD approach may work rather well in the regime where saturation effects become important.

With the advent of new, more accurate, HERA data [17], the limitations of the original 'saturation model' become however obvious. This situation has stimulated new attempts, better rooted in QCD, towards improving the dipole cross–section [54, 74, 75]. In Ref. [74] the behaviour of the fit at large Q^2 has been considerably improved by using

$$\sigma_{\text{dipole}}(x, r_\perp) \,=\, \sigma_0 \left[1 - \exp\left(- r_\perp^2 \frac{\pi^2 \alpha}{N_c} \frac{xG(x, 1/r_\perp^2)}{\sigma_0} \right) \right], \qquad (1.27)$$

which amounts to a Glauber–like resummation of the dominant 'leading–twist' result at high Q^2, Eq. (1.24). In Ref. [75], the Glauber resummation has been supplemented with a model for the impact–parameter dependence of the dipole scattering amplitude. As we shall discover in Sect. 4, Eq. (1.27) is still inaccurate in the transition region around $r_\perp \sim 1/Q_s(x)$, where it fails, e.g., to describe the 'geometric scaling' property to which we shall shortly turn. More recent phenomenological analyses, which include more of the QCD dynamics in the approach towards saturation [54, 76] and lead to satisfactory fits to the new HERA data at small x, will be discussed in Sect. 5.

The phenomenological successes of the saturation models must be considered with great caution (see, e.g., the recent discussion by Thorne [77]). First, the experimental points at HERA are correlated in such a way that the smallest values of x correspond also to rather small values of Q^2 ($\lesssim 1\,\text{GeV}^2$), where the use of perturbation theory becomes questionable. Second, it is not clear whether the interpretation of the data requires saturation. Indeed, there are other phenomenological analyses which do not include saturation, but achieve a good description of the small–x data at HERA, at least within limited domains of Q^2.

For instance, the fits using the DGLAP equation [15] work remarkably well for all $Q^2 > 2$ GeV^2 (see, e.g., [78, 79]), while analyses inspired by the Regge theory [80, 81] lead to rather successful global fits. Furthermore, more complicated, 2–component, dipole models [82] manage to provide a good description of the transition region from high to low Q^2.

What is however even more remarkable than the apparent success of the saturation models in fitting the data is their ability to provide a natural interpretation for some striking, *qualitative*, phenomena observed at HERA and at RHIC (see Sect. 1.3), for which no other convincing physical interpretation has been proposed so far.

One such a phenomenon refers to the *diffractive* DIS at HERA. (This is the DIS process in which the proton emerges unscathed from the interaction, and a large 'rapidity gap' — a region in rapidity essentially devoid of particles — is produced between the fragmentation region of the electron and that of the proton.) Namely, the ratio $\sigma_{diff}/\sigma_{\gamma^*p}$ measured in the experiment turns out to be quite large (~ 0.1) and, especially, nearly independent of the energy, a feature that was not anticipated in the models based on perturbative QCD. The saturation models, through a subtle interplay of scales, reproduce this distinctive feature [73, 74] in a rather convincing way. Their success on this point is particularly suggestive since diffraction is expected to be very sensitive to saturation effects (because the corresponding cross–section receives a significant contribution from relatively large dipole configurations; see, e.g., the discussion in Ref. [73]).

But the most striking among the phenomena observed at HERA which are naturally associated with saturation is the property known as *"geometric scaling"* [83]: The Golec-Biernat–Wüsthoff dipole cross–section, Eq. (1.26), has the remarkable feature to depend upon the two kinematical variables x and $r_\perp$ only via the dimensionless combination $\mathcal{T} \equiv r_\perp^2 Q_s^2(x)$ (the "scaling variable"). Through the factorization formula (1.22), this scaling property then transmits to the photoabsorbtion cross–section : in the limit where the quark masses are negligible, σ_{γ^*p} is a function of the ratio $Q^2/Q_s^2(x)$ alone. Inspired by this observation, Stasto, Golec-Biernat and Kwiecinski performed a model–independent analysis of the data [83] and found that the measured cross–section $\sigma_{\gamma^*p}(x, Q^2)$ for $x \leq 10^{-2}$ shows indeed *approximate scaling* as a function of the variable $Q^2/Q_s^2(x)$, with $Q_s^2(x) \propto 1/x^\lambda$ and $\lambda \sim 0.3$, within the whole range available in Q^2 (namely, $Q^2 \leq 450\,\text{GeV}^2$). The quality of this scaling can be appreciated by inspection of Fig. 10.

It should be emphasized that, although natural within the framework of the simple 'saturation model' (1.26), the quality of this scaling in the high Q^2 regime has posed a real challenge to perturbative QCD. Whereas at low momenta $Q^2 < Q_s^2(x)$ such a scaling is indeed natural in the context of saturation, for larger $Q^2 > Q_s^2(x)$ one would expect the scaling to be violated by quantum evolution, similarly to what happens for the Bjorken–scaling: The radiative processes in QCD should introduce a dependence upon the infrared scale Λ_{QCD}, and thus break down the geometric scaling. (E.g., the gluon distribution in the Glauber–type cross–section (1.27) depends upon $Q^2/\Lambda_{\text{QCD}}^2$.) This paradox triggered new theoretical developments [51–53, 84], with the conclusion that geometric scaling *is* in fact consistent with perturbative QCD: The perturbative (QCD) evolution with saturation boundary conditions along the saturation line preserves the geometric scaling property characteristic of saturation up to relatively large transverse momenta $\sim Q_s^4(\tau)/\Lambda_{\text{QCD}}^2$ [51]. This is in qualitative agreement with the observation at HERA, and will be further dis-
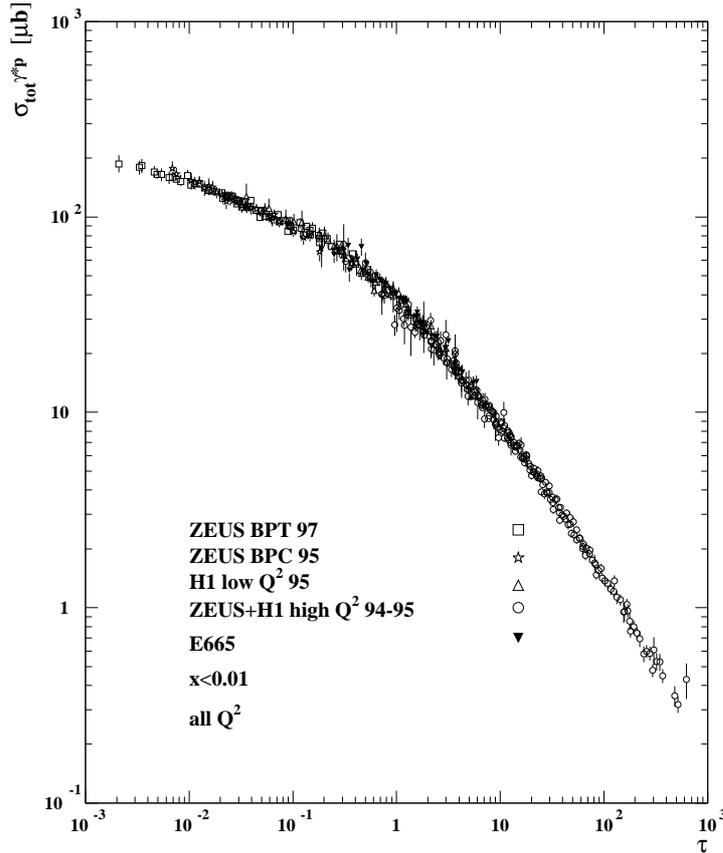
Fig. 10. *HERA data on the cross section for γ^*p DIS from the region $x < 0.01$ and $Q^2 < 450\text{GeV}^2$ plotted versus the scaling variable $\mathcal{T} = Q^2/Q_s^2(x)$ (from Ref. [83]).*

cussed in Sects. 4.3 and 5.

Note finally that indications of geometric scaling have been found also in the data for DIS off nuclei [85], and that this property plays an important role in the analysis of the RHIC data for particle production in deuteron–gold collisions [86, 87] (see Sects. 1.3 and 5).

CT_EIKONAL

1.2.3 Dipole scattering in the eikonal approximation

At high energy, multiple scattering in the dipole–hadron collision can be resummed to all orders in the *eikonal approximation.* This amounts to neglecting the recoil of the quark (or the antiquark) during its scattering off the color field in the target; the whole effect of the scattering then consists in a phase factor. In QCD, this phase factor is a *Wilson line* and describes color precession [39, 71, 88].

To describe the eikonal approximation and, more generally, the kinematics of a high–energy collision, it is convenient to use *light–cone variables* (coordinates and momenta). Let z be the longitudinal axis of the collision. For an arbitrary 4–vector $v^\mu = (v^0, v^1, v^2, v^3)$ ($v^3 = v_z$, etc.), we define its light–cone (LC) coordinates as

26

$$v^+ \equiv \frac{1}{\sqrt{2}}(v^0 + v^3), \qquad v^- \equiv \frac{1}{\sqrt{2}}(v^0 - v^3), \qquad \boldsymbol{v} \equiv (v^1, v^2). \tag{1.28}$$

We shall often write $v_\perp = |\boldsymbol{v}| = \sqrt{(v^1)^2 + (v^2)^2}$.

To appreciate the utility of such definitions, consider a particle propagating at nearly the speed of light in the positive z direction ('right–mover'). Its trajectory lies along the positive light–cone ($z \simeq t$) and its 4–momentum reads $P^\mu \simeq (P, 0, 0, P)$. In LC variables, the particle 'sits' at $x^- = 0$ while the 'LC time' $x^+ \simeq \sqrt{2}t$ goes by; its LC momentum has just one component, that along the light–cone: $P^\mu \simeq \delta^{\mu+}P^+$, with $P^+ = \sqrt{2}P$. If the particle carries some (electrical or color) charge, then the associated LC current has only a plus component: $j^\mu \simeq \delta^{\mu+}j^+$, and it generates a gauge field which, with a suitable choice of the gauge, can be described by a vector potential with a single component: $A^\mu \simeq \delta^{\mu+}A^+$. We see that, when expressed in terms of LC variables, the kinematics of the infinite momentum frame becomes formally similar to that of the particle rest frame.

For a fast particle moving in the negative z direction ('left–mover' : $z \simeq -t$), the physical roles of x^+ and x^- are interchanged: $x^- \simeq \sqrt{2}t$ plays now the role of time, where $x^+ \simeq 0$ is the 'longitudinal coordinate'. To avoid any confusion, in what follows we shall systematically use a terminology adapted to the *right mover* : x^+ will be thus referred to as the 'LC time', while x^- will be the 'LC longitudinal coordinate' (and, of course, $\boldsymbol{x} = (x, y)$ is the transverse coordinate). Also, following the example of DIS in the Breit frame, we shall often refer to the right mover as the 'target' and to the left mover as the 'projectile'. Note that, in LC variables the invariant dot product reads:

$$k \cdot x = k^- x^+ + k^+ x^- - \boldsymbol{k} \cdot \boldsymbol{x}, \tag{1.29}$$

which shows that k^- (the momentum variable conjugate to the 'LC time' x^+) should be interpreted as the *LC energy*, and k^+ as the *LC longitudinal momentum*. In particular, for particles on the mass–shell: $k^\pm = (E \pm k_z)/\sqrt{2}$, with $E = (m^2 + \boldsymbol{k}^2)^{1/2}$, and therefore:

$$k^+ k^- = \frac{1}{2}(E^2 - k_z^2) = \frac{1}{2}(k_\perp^2 + m^2) \equiv \frac{1}{2}m_\perp^2 . \tag{1.30}$$

We shall also need the *rapidity* (for an on–shell particle):

$$\mathrm{y} \equiv \frac{1}{2}\ln\frac{k^+}{k^-} = \frac{1}{2}\ln\frac{2k^{+2}}{m_\perp^2}. \tag{1.31}$$

Note that, under a longitudinal Lorentz boost ($k^+ \to \beta k^+$, $k^- \to (1/\beta)k^-$, with constant β), the rapidity is just shifted by a constant: $\mathrm{y} \to \mathrm{y} + \beta$.

For a parton inside a right–moving hadron, we introduce the boost–invariant longitudinal momentum fraction x as:

$$x \equiv \frac{k^+}{P^+} \qquad \text{(right mover)}. \tag{1.32}$$

(For a left mover, one would rather define $x \equiv k^-/P^-$.) Clearly, in the hadron IMF, this reduces to our previous definition $x = k_z/P_z$. Then the parton rapidity reads (assuming the parton to be nearly on–shell):

$$y = \frac{1}{2}\ln\frac{k^+}{k^-} = \frac{1}{2}\ln\frac{2k^{+2}}{m_\perp^2} = Y - \ln\frac{1}{x} + \ln\frac{M}{m_\perp}, \tag{1.33}$$

where $Y = \ln(\sqrt{2}P^+/M)$ is the rapidity of the fast moving hadron. We see that the quantity $\tau \equiv \ln(1/x)$ introduced before, in the context of DIS, should be more properly referred to as the rapidity *difference* between the hadron and the parton: $\tau \simeq Y - y$.

We now close this parenthesis on LC kinematics and return to the problem of the eikonal scattering. Consider a quark projectile (which by our previous conventions is a left mover) which propagates in the color field of the hadronic target (the right mover). As explained above, the color current of the quark has only one large component, namely j_a^-, which couples to the large component A_a^+ of the color field in the target. In a frame in which one can neglect gluon radiation from the quark itself, the quark dynamics is governed by the Dirac equation describing its propagation in the 'background field':

$$(\partial_- - igA_a^+ t^a)\psi = 0. \tag{1.34}$$

Here $\partial_- = \partial/\partial x^- = \partial^+$, and we have used the fact that neither the quark transverse coordinate, nor the helicity of the associated Dirac spinor, are changed by the high–energy scattering. The only non–trivial dynamics is that in the longitudinal direction, and this is described by Eq. (1.34). The S–matrix for the quark–background field scattering is then obtained as $S_q(\boldsymbol{x}; A^+) = \langle \text{out}|\text{in}\rangle$, where $\boldsymbol{x}$ is the transverse coordinate of the quark, and the outgoing and incoming quark states are related by $\psi_i^{out} = (V_{\boldsymbol{x}}^\dagger)_{ij}\psi_j^{in}$, with $V_{\boldsymbol{x}}^\dagger \equiv V^\dagger(\boldsymbol{x})$ a color matrix in the fundamental representation defined by:

$$(\partial^+ - igA_a^+ t^a)V_{\boldsymbol{x}}^\dagger = 0 \quad \Longrightarrow \quad V^\dagger(\boldsymbol{x}) = \text{P}\exp\left(ig\int dx^- A_a^+(x^-,\boldsymbol{x})t^a\right). \tag{1.35}$$

In the last equation, the symbol P denotes the x^- ordering of the color matrices $A_a^+(x^-,\boldsymbol{x})t^a$ in the exponent, from right to left in increasing order of their x^- arguments. The integration runs formally over all the values of x^-, but in reality this is restricted to the longitudinal extent of the hadron (recall that, for the right–moving hadron, x^- is the longitudinal coordinate !), which is localized near $x^- = 0$ because of *Lorentz contraction*. The x^+ coordinate is not explicitly shown since the corresponding dependence is unimportant: by Lorentz contraction again, but this time applied to the left–moving projectile, the latter is localized near $x^+ = 0$ (since x^+ is the longitudinal coordinate for the projectile); but in the frame in which we view the scattering, the internal dynamics of the hadron is slowed down by *Lorentz time dilation*, so its color field A^+ is essentially independent of x^+ over the duration of the scattering. Thus, the field in $A_a^+(x^-,\boldsymbol{x})$ in Eq. (1.35) can be interpreted as $A_a^+(x^+ \simeq 0, x^-, \boldsymbol{x})$.

The color matrix in Eq. (1.35) is recognized as a *Wilson line* (a path–ordered exponential of the gauge field, with the 'path' represented here by the trajectory of the quark). Consider similarly an *antiquark*: Since an antiquark moving in the negative z direction (i.e., towards increasing x^-) is equivalent to a quark propagating backwards in z *and* in time (i.e., towards decreasing x^-), the corresponding S–matrix will involve the Hermitian conjugate Wilson line $V_{\boldsymbol{y}}$, with $\boldsymbol{y}$ the transverse coordinate of the antiquark. Finally, for a left–moving *color dipole*, the S–matrix is obtained as the color average of the product of the

individual S–matrices for the quark and the antiquark respectively:

$$S(\boldsymbol{x}, \boldsymbol{y}; A^+) \;=\; \frac{1}{N_c} \operatorname{tr}(V_{\boldsymbol{x}}^\dagger V_{\boldsymbol{y}}). \tag{1.36}$$

The physical S–matrix for the dipole–hadron scattering is finally obtained after averaging over the color fields A^+ in the target:

$$\langle S(\boldsymbol{x}, \boldsymbol{y}) \rangle_\tau \;=\; \frac{1}{N_c} \Big\langle \operatorname{tr}(V_{\boldsymbol{x}}^\dagger V_{\boldsymbol{y}}) \Big\rangle_\tau. \tag{1.37}$$

It will be a main objective of the subsequent developments in this review to better characterize the average over the target wavefunction and to construct approximation schemes allowing its computation in the high energy regime. As indicated in Eq. (1.37), this averaging introduces the dependence upon the total energy of the scattering, i.e., upon the rapidity difference $\tau = \mathrm{Y}_{\text{hadron}} - \mathrm{Y}_{\text{dipole}}$ (indeed, $\tau \simeq \mathrm{Y}_{\text{hadron}}$ in the dipole frame).

Since explicitly constructed in terms of unitary matrices (the Wilson lines), the dipole S–matrix in Eqs. (1.36)–(1.37) respects the unitarity bound $|S| \le 1$, as it should. Clearly, the unitarization mechanism at work here is *multiple scattering* : the Wilson lines are non–linear in A^+ to all orders, and thus describe multiple gluon exchanges between the projectile and the target. The unitarity corrections become important when the color field is strong enough: $g \int dx^- A^+ \sim 1$; as we shall shortly discover, this is the same as the saturation condition for the gluon distribution in the target. This is physically reasonable: when the gluon occupation number is large enough for the associated non–linear effects to be important, these non–linearities can manifest themselves both as gluon recombination, thus leading to saturation in the target wavefunction, or as multiple interactions with an external projectile, then leading to the unitarization of the collision process.

A cartoon of high–energy DIS in the dipole frame and in the presence of non–linear effects is displayed in Fig. 11. Because of its conceptual simplicity and of its direct relevance for the phenomenology of DIS, the dipole–hadron scattering will represent a privileged laboratory for testing all the theoretical developments to be presented later on.

1.3   Motivation: Initial conditions for heavy ion collisions at RHIC

T_DATARHIC

In the previous section, we have seen that lepton–proton deep inelastic scattering at small Bjorken x is a formidable tool for studying the high parton density (or strong color field) phase of QCD at high energy. We have furthermore seen that the kinematical conditions at HERA are only marginally consistent with a theoretical analysis of gluon saturation based on perturbative QCD. But besides increasing the energy, there is also a different way to enhance the parton density and thus improve the case for perturbative QCD: this consists in replacing the proton with a large nucleus, with atomic number $A \gg 1$.

Indeed, as we shall explain in more detail in Sect. 2.4, the *nuclear* gluon overlapping factor $\varphi_A(x, Q^2)$ for gluons localized within a small transverse area $1/Q^2$ with $Q^2 \gg \Lambda_{\text{QCD}}$ and having a small longitudinal momentum fraction $x \ll A^{-1/3}$ can be estimated by replacing in Eq. (1.16) the gluon distribution $xG(x, Q^2) \equiv xG_N(x, Q^2)$ of a single nucleon
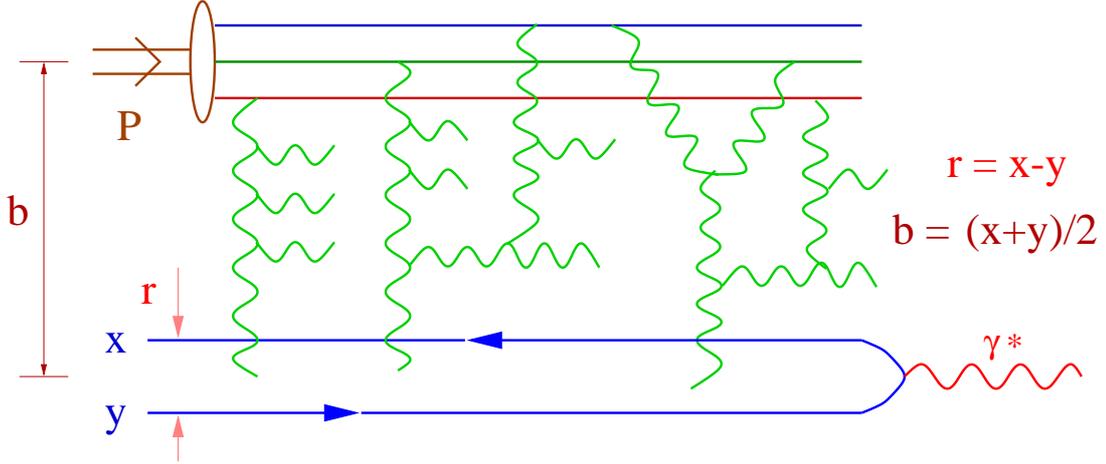
Fig. 11. *A cartoon of DIS at high energy and in the dipole frame. The non–linear effects associated with the high gluon density (gluon recombination in the hadron wavefunction and multiple scattering with the external dipole) are indicated by small circles. The horizontal gluon lines correspond to the BFKL evolution to be described in Sect. 1.4.*

by $xG_A(x, Q^2) \approx AxG_N(x, Q^2)$, and the proton radius R by the nuclear one $R_A \approx A^{1/3}A$. This yields $\varphi_A(x, Q^2) \sim A^{1/3}\varphi_N(x, Q^2)$, which shows that increasing A is like increasing the energy: it drives the system towards higher gluon density. In particular, Eq. (1.19) shows that the saturation scale $Q_s^2(A, x)$ for a nucleus should approximately scale like $A^{1/3}$. With the empirical parametrization $Q_s^2(A, x) \sim A^{1/3}x^{-\lambda}$ where $\lambda \approx 0.3$ from the fits to the HERA data (and also from perturbative QCD [53]), we deduce that replacing the proton by a gold nucleus should be tantamount to reducing x by a factor $A^{1/3\lambda} \approx 200$.

So far, there has been no experimental study of DIS off such a heavy nucleus (although the opportunity of an electron–ion collider is currently under study [89]). But the Relativistic Heavy Ion Collider (RHIC) is operating at BNL since 2000, and in this experiment nuclei as heavy as gold enter in collision with each other at energies as large as $\sqrt{s} = 200$ GeV *per nucleon pair*. With the advent of the Large Hadron Collider (LHC) at CERN, expected for 2007, the corresponding energy will be increased up to 5500 GeV.

1.3.1 Ultrarelativistic heavy ion collisions at RHIC

At such ultrarelativistic energies, the nuclei appear as two–dimensional sheets because of Lorentz contraction (the Lorentz γ factor at RHIC is ~ 100), and each of these sheets is essentially a 'color glass condensate' (i.e., a system of saturated gluons whose internal time scales are frozen by Lorentz time dilation). In the early stages of the collision, the original sheets pass through one another, but in their wake they leave melting colored glass, which eventually materializes as quarks and gluons. The 'melting' (or formation) proper time is set by the typical internal momentum scale before the collision, which is the saturation scale: $\tau_0 \sim 1/Q_s$, which at RHIC is estimated as $\tau_0 \sim 0.2$ fm/c $\sim 5 \times 10^{-25}$ sec. (Note that 0.2 fm/c is also comparable to the natural crossing time of two 10 fm nuclei, each contracted by a γ factor 100, in the center of mass frame.) For particles with a large momentum or rapidity along the beam axis, this time scale is Lorentz dilated. This

large p   small p   large p
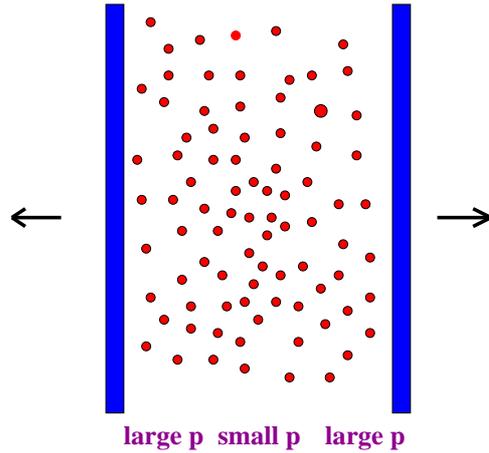
Fig. 12. A nucleus–nucleus collision. The produced particles are shown as circles.

collis

means that the slow (smaller rapidity) particles are produced first towards the center of the collision regions and the fast (larger rapidity) particles are produced later further away from the collision region. This Bjorken "inside–out" correlation [90] between space and momentum is similar to what happens to matter in Hubble expansion in cosmology: The stars which are further away have larger outward velocities. This means that the matter produced at RHIC, like the universe in cosmology, is born expanding. But unlike the expansion of the universe after the Big-Bang which is three dimensional, that of the matter produced in the "mini-bang" at RHIC is one–dimensional, along the collision axis. This is shown in Fig. 12.

The subsequent evolution of the matter produced after the collision is expected to pass through the stages depicted in Fig. 13, where the various phases are characterized according to their energy density. In turn, the latter is estimated by extrapolating backwards in time the measured energy density of the final particles. This extrapolation involves various theoretical scenarios resulting in the uncertainty bound manifest in Fig. 13. But even the lowest estimate for the energy density $\epsilon(\tau_0)$ at the formation time, namely $\epsilon(\tau_0) \sim 3$ GeV/fm^3, is still much higher than the energy density 0.15 GeV/fm^3 of the nuclear matter. Thus the matter produced by color glass melting is clearly in the deconfined phase of QCD: this is a high density system of quarks and gluons, which cools down as the system expands, and possibly thermalizes in a *quark–gluon plasma* [55], before eventually hadronizing into the multitude of hadrons (a few thousands) that are finally captured by the detectors at RHIC [26, 27]. It appears therefore as a challenge for the theorists to imagine observables which could survive (almost) unchanged to such a violent evolution, and thus carry out information about the initial conditions (in particular, about the color glass). It is furthermore a challenge for the experimentalists to extract and measure such observables with a significant accuracy.

A theory of high parton densities and gluon saturation should provide us with a description of the partonic system produced in the early stages of the collision and which represents the *initial condition* for the subsequent evolution illustrated in Fig. 13. In turn, such a description should involve two ingredients: (i) a theory for the high–density wavefunctions

31
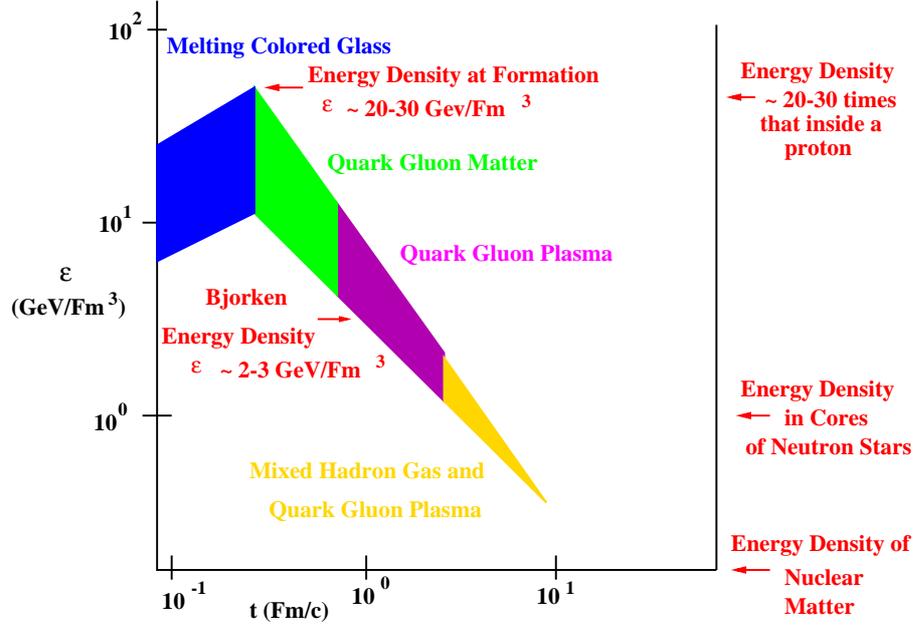
Fig. 13. Bounds on the energy density as a function of time in heavy ion collisions.

of the incoming nuclei, and (ii) a formalism allowing one to compute parton production in the nucleus–nucleus (AA) collision. As we shall explain at length in the subsequent sections, the *theory for the color glass condensate* provides us with both these ingredients (at least, within a limited theoretical accuracy), and thus with a description of the state of the system at time τ_0. There are at least two ways how one could test the predictions of such a theory for the initial conditions:

a) First, one can compute, and measure, observables which are sensitive to the state of the system at early times, and which are not strongly affected by the interactions at later times (generally referred to as emerging as a hadronic jet (see Fig. 14). Still as in DIS, this multiple scattering is a probe of the non–linear effects in the gluon distribution of the dense target (here, the nucleus). Thus from measurements of the particle yield in deuteron–gold collisions at RHIC one should be able to pinpoint saturation effects in the gold nucleus wavefunction [30, 86, 91–95]. These measurements represent one of the highlights of the experimental activity at RHIC 'final state interactions'). The 'hard' (high energy and large transverse momentum) electromagnetic and leptonic signals have these properties [96, 97], but they are rather difficult to measure, because of the large background. Also, one can reasonably argue that some gross properties of the produced hadrons, like their 'total multiplicity' $dN/d\eta$ (the total particle yield per unit of pseudo–rapidity), should be controlled by gluon production in the early stages. And indeed, it turns out that simple calculations within the CGC framework satisfactorily describe the global features of the total multiplicity at RHIC, like its dependencies upon the energy $\sqrt{s}$, the pseudo–rapidity η, and the centrality of the collision (or the "number of participants") [98]. Finally, one expects the 'final state' interactions to be negligible for the asymmetric *proton–nucleus* (pA) collisions, which in practice are realized as deuteron–gold collisions at RHIC. Indeed, in a frame in which the nucleus carries most of the total energy (and where the proton

32
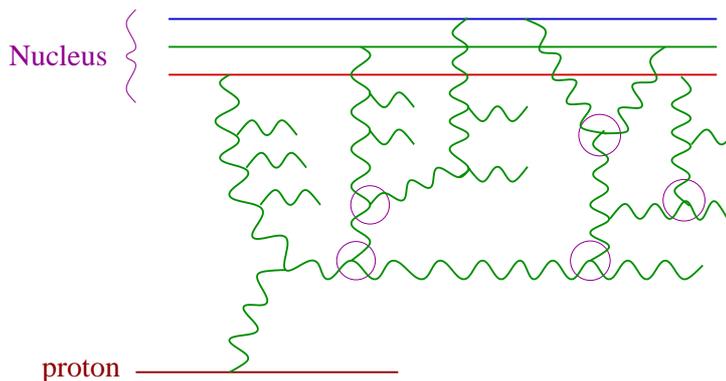
Fig. 14. pA collision at forward rapidity: the (relatively large–x) gluon emitted by the incoming proton scatters off the highly–evolved gluon distribution in the nuclear wavefunction at small–x. Non–linear effects (which correspond to either gluon recombination in the nuclear wavefunction, or multiple scattering of the 'external' gluon) are indicated by small cercles.

wavefunction is therefore dilute), a pA collision is quite similar to DIS : one parton from the proton undergoes multiple scattering off the color fields in the nucleus before so far, and their results will be described in some detail in the remaining part of this section.

b) Second, one can try to incorporate our knowledge about the initial conditions into a coherent theoretical description of the evolution of the partonic system at all times $\tau > \tau_0$, and then check whether the predictions of such a global theory are indeed consistent with the ensemble of observables measured at RHIC. Clearly, such a global understanding is the ultimate goal of the physics at RHIC. In particular, this should allow us to establish whether a *thermalized quark–gluon plasma* (QGP) is indeed produced at the intermediate stages of a ultrarelativistic heavy ion collision. This issue turns out to be highly non–trivial: Although the initial energy density, at a time $\tau_0 \sim 1/Q_s$, is significantly greater than the energy density required (according to the lattice calculations) to form a QGP, the initial partonic system is highly off–equilibrium, and it is not clear whether the subsequent interactions among the partons will be able to achieve thermalization over the finite amount of time (~ 10 fm/c) before hadronization. In fact, if the collective flow observed in the AA–data at RHIC is to be explained by perfect fluid hydrodynamics, as generally accepted nowadays, the equilibration time should be even shorter, $\tau_{eq} \lesssim 1$ fm/c [55]. However, it appears to be difficult to reconcile our present understanding of perturbative processes in QCD with such a short thermalization time.

A necessary condition for thermalization is that momentum distributions be isotropic. The CGC initial conditions are however very anisotropic, with $\langle k_\perp \rangle \sim Q_s$ and $\langle k_z \rangle \sim 0$. One expects isotropization to proceed via parton rescattering, but all the pQCD–based estimates of the corresponding equilibration time, which include both $2 \to 2$ processes [99–101] and inelastic $2 \to 3$ processes [102, 103], yield significantly larger times $\tau_{eq} \sim 2 - 3$ fm/c [104]. Recently, it has been suggested that collective instabilities, the non–Abelian analog of the well known Weibel instabilities in plasma physics, can speed up thermalization [105–107, ?–107]. Starting from very anisotropic (CGC–like) initial conditions, these instabilities drive the system to isotropy on very short time scales, of order $1/Q_s$ in some

33

estimates. But more studies are necessary before a final conclusion can be reached.

1.3.2   The R_{pA}–ratio

Let us now return to the proton–nucleus (or d+Au) collisions, which as mentioned before represent the cleanest measurement of initial state effects in the experiment at RHIC, and discuss in more detail a particular observable known as the *nuclear modification factor* (or the 'R_{pA}–ratio'), which so far has brought the strongest evidence for gluon saturation in the RHIC data. This observable is meant to single out nuclear, collective, effects — which are necessarily initial state effects in the nuclear wavefunction — in the particle production in pA collisions. To that aim, the spectrum $\mathrm{d}N_{pA}/\mathrm{d}^2 p_\perp \mathrm{d}\eta$ of the hadrons produced in a pA collision is normalized through the spectrum $A \times \mathrm{d}N_{pp}/\mathrm{d}^2 p_\perp \mathrm{d}\eta$ of A independent proton–proton (pp) collisions, at the same energy per nucleon pair. The ensuing ratio

$$R_{pA}(\eta, p_\perp) \equiv \frac{1}{A} \frac{(\mathrm{d}N_{pA}/\mathrm{d}\eta \mathrm{d}^2 p_\perp)}{(\mathrm{d}N_{pp}/\mathrm{d}\eta \mathrm{d}^2 p_\perp)}, \tag{1.38}$$

would be equal to one if the pA collision were an incoherent superposition of nucleon–nucleon collisions; conversely, its deviation from one is a measure of the collective effects in the wavefunction of the incoming nucleus. In this equation, $p_\perp$ is the hadron transverse momentum, and η is its pseudo–rapidity :

$$\eta \equiv \frac{1}{2} \ln \frac{p + p_z}{p - p_z} = \frac{1}{2} \ln \frac{1 + \cos\theta}{1 - \cos\theta} = -\ln \tan \frac{\theta}{2}, \tag{1.39}$$

with $p = \sqrt{p_\perp^2 + p_z^2}$ and θ the angle between the direction of the produced hadron and the collision axis. For a massless particle, η coincides with the standard rapidity y, Eq. (1.31); but in the experimental setting, it is easier to measure the deflection angle, and thus η.

Our conventions are such that positive values for η ('forward rapidities') correspond to particles produced in the proton fragmentation region. What is peculiar about the kinematics of particle production in pA collisions, as illustrated in Fig. 15, is that by increasing η we are probing gluons with smaller values of x in the nuclear wavefunction. Indeed, let us consider the scattering in the center–of–mass frame, where the nucleus is a left–mover, with LC 4–momentum $P_1^\mu = (0, P_1, 0_\perp)$, whereas the proton is a right–mover with $P_2^\mu = (P_2, 0, 0_\perp)$ (and $P_1 = P_2 = \sqrt{s/2}$). The produced gluon emerges with $p^\mu = \left(\frac{p_\perp}{\sqrt{2}} \mathrm{e}^\eta, \frac{p_\perp}{\sqrt{2}} \mathrm{e}^{-\eta}, \mathbf{p}_\perp\right)$. If $k_{1,2}^\mu$ refer to the two gluons which fuse with each other (see Fig. 15), then energy–momentum conservation implies $x_1 \equiv k_1^-/P_1 = (p_\perp/\sqrt{s}) \mathrm{e}^{-\eta}$ and $x_2 \equiv k_2^+/P_2 = (p_\perp/\sqrt{s}) \mathrm{e}^\eta$. Thus, as anticipated, larger positive values for η correspond to smaller values for the longitudinal fraction x_1 of the gluon from the nucleus.

The experimental results at RHIC for the ratio R_{dAu} in deuteron–gold collisions show a remarkable structure, which is better appreciated by comparison with the corresponding results for the similar ratio R_{AuAu} for gold–gold collisions: As manifest in Fig. 16, at central rapidity ($\eta = 0$) the ratio R_{AuAu} is significantly smaller than one, especially at high transverse momenta ('*high–$p_\perp$ suppression*'); by contrast, the ratio R_{dAu} shows an
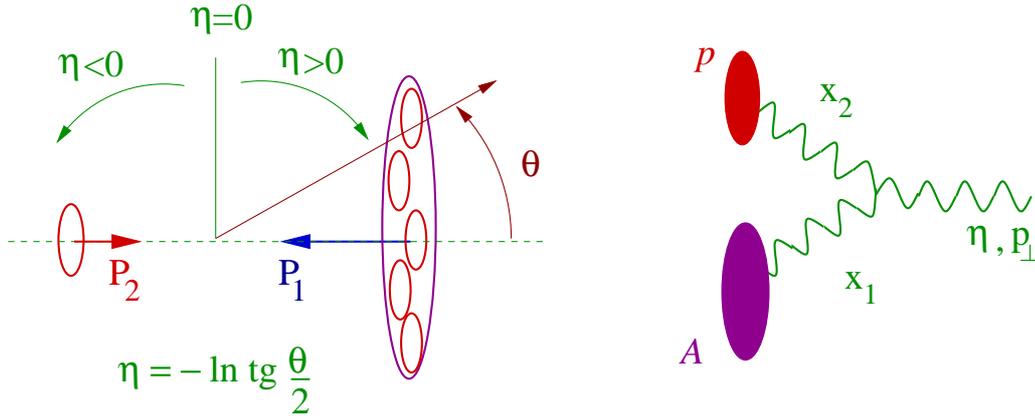
Fig. 15. *Kinematics for gluon production via gluon fusion in proton(p)–nucleus(A) collisions.* | KINEM |
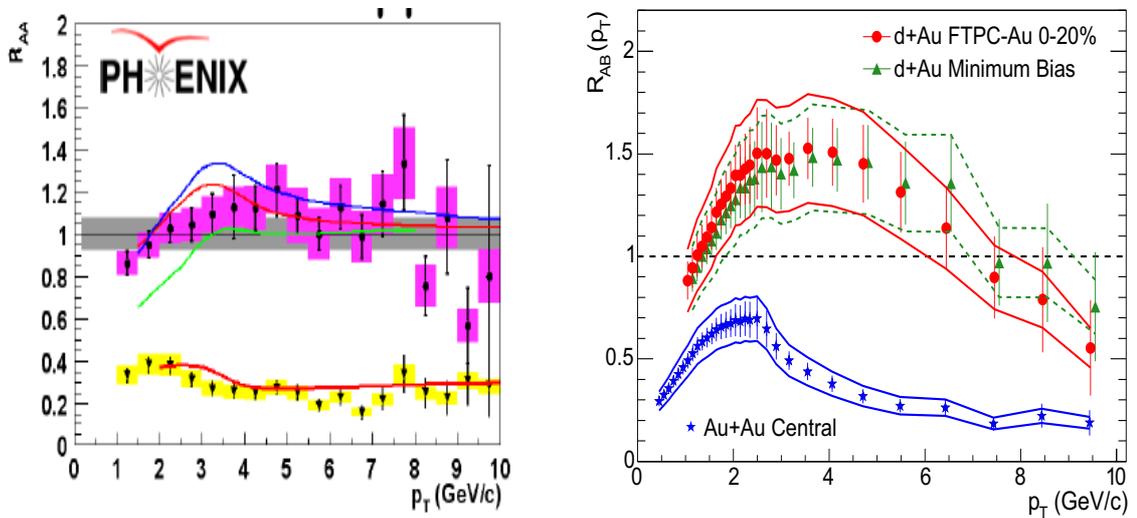
Fig. 16. Central rapidity ($\eta = 0$) PHENIX π^0 and STAR $h^\pm$ data comparing R_{dAu} to R_{AuAu} [109]. These results, together with similar data from BRAHMS and PHOBOS [109], prove that jet quenching in Au+Au collisions at central rapidity is a final state effect. | dAu0 |

enhancement at intermediate momenta, generally referred to as a *'Cronin peak'* (because of the analogy with a similar phenomenon observed in *pp* collisions at lower energies [108]).

The Cronin peak is generally attributed to the multiple interactions suffered by the colliding parton from the deuteron on its way through the nucleus. At small (nuclear) $x_1 \ll A^{-1/3}$ — a condition which is indeed satisfied for the $\eta = 0$ data, since then $x_1 \sim 10^{-2}$ for an emerging hadron with $p_\perp = 2$ GeV (for $\sqrt{s} = 200$ GeV) —, these multiple interactions are predominantly rescattering off the strong gluon fields in the nucleus, and thus are tantamount of *gluon saturation*. We shall see later, in Sect. 5.2, that models of saturation lead indeed to a Cronin enhancement at central rapidity. By contrast, the suppression observed in gold–gold collisions is not an initial state effect, rather this is to be attributed to 'jet quenching' through final–state interactions: a parton initially produced with a hard momentum will loose energy via rescattering and medium–induced

35

gluon radiation while propagating through the dense and 'colored' surrounding medium, before finally escaping as a jet [55, 110–113].

But the general pattern of the RHIC data for R_{dAu} changes dramatically when increasing the pseudo–rapidity η. The data obtained by the BRAHMS experiment at 'forward rapidities' $0 \leq \eta \leq 3.2$ are shown in Fig. 17 [57]: the Cronin peak is seen to disappear after just one unit of rapidity, while for $\eta > 1$ we rather have *suppression* at all $p_\perp$. Moreover, the trend of the centrality dependence gets also reversed when increasing η, as manifest in Fig. 18. This behaviour, which has been confirmed by the other collaborations at RHIC [27], may at a first sight look counterintuitive: As argued before, the Cronin enhancement is due to multiple scattering within the nucleus. When increasing η, and also when increasing centrality, we are probing a region of higher gluon density (e.g., $\eta = 3.2$ corresponds to $x_1 \sim 10^{-4}$ for $p_\perp = 2$ GeV), which should favour rescattering and thus enhance the Cronin peak. But the experimental results show precisely the opposite trend!
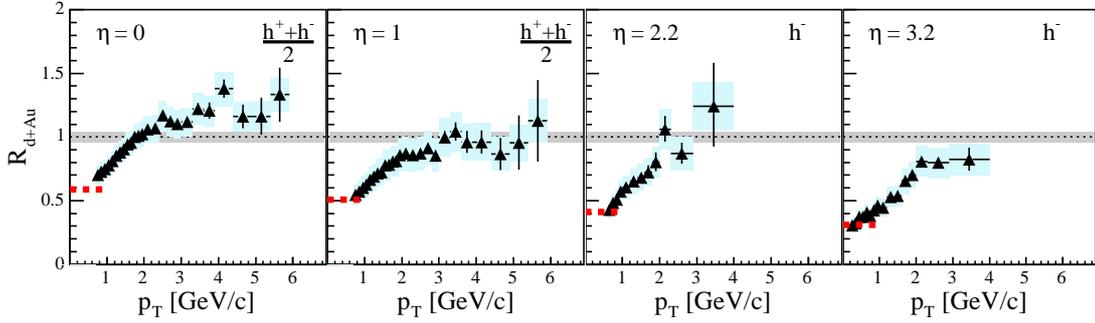
Fig. 17. BRAHMS results for d+Au collisions at 200 GeV/nucleon [57] : The ratio R_{dAu} for charged hadrons as a function of $p_\perp$ for central and forward pseudorapidities.   `brahmsR`
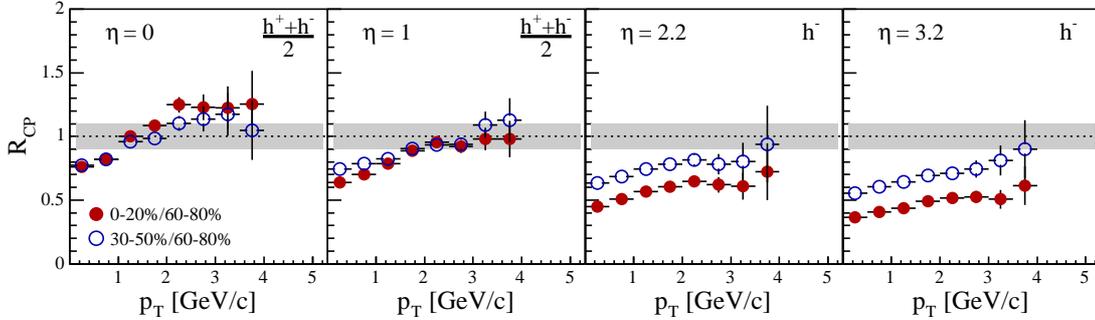
Fig. 18. d+Au collisions at BRAHMS [57] : The ratio R_{CP} of yields from collisions of a given centrality class (0-20% or 30-50%) to yields from more peripheral collisions (60-80%), scaled by the mean number of binary collisions in each sample, as a function of $p_\perp$ for central and forward pseudorapidities.   `brahmsCP`

But as counterintuitive as it might seem, this behaviour has been in fact anticipated, on the basis of the CGC ideas [86, 114, 115] : As we shall explain in Sect. 5.2, the R_{pA} ratio (1.38) is related, through '$k_\perp$–factorization', to the ratio between the gluon occupation factors in the nucleus and in the proton measured at the same value of $x = x_1$ and $k_\perp$.

When decreasing x_1 (or increasing η), the occupation factors rise both in the proton and in the nucleus, but this rise is much faster for the proton, which has a larger phase–space available for its evolution. Indeed, as manifest on Eq. (1.13) and also from Fig. 8, the transverse phase–space available for the linear QCD evolution is the logarithmic difference $\rho - \rho_s \equiv \ln k_\perp^2 / Q_s^2$ w.r.t. the saturation momentum. But the nuclear saturation scale $Q_s^2(A) \propto A^{1/3}$ is considerably larger than the proton one, so $\rho - \rho_s(A) < \rho - \rho_s(p)$, as anticipated. In Sect. 5.2, we shall see that quantum evolution in the CGC can indeed explain (qualitatively and even quantitatively) the trend of the data observed in Figs. 17 and 18 [87, 114, 115].

1.3.3 The McLerran-Venugopalan model

The simplest version of the theory for the color glass condensate has been originally introduced, by McLerran and Venugopalan [25] (see also Ref. [28]), as a model (a classical effective theory) for the gluon distribution of a large nucleus. This model is conceptually interesting in that it allows one to study in a simple setting the importance of non–linear effects in the gluon dynamics for phenomena like gluon saturation or the unitarization through multiple scattering. Also, the MV model provides a physically motivated initial condition for the quantum evolution with increasing energy that we shall describe in Sect. 3. On the phenomenological side, one should stress that this simple model — which applies for a very large nucleus $A \gg 1$ at not too high energy (for the effects of the quantum evolution to remain negligible) — has found a wealthy of applications to the physics at RHIC, including the lattice calculation of the energy density and the multiplicity of the gluons released in the initial stages of the collision [116–119] and of the original 'elliptic flow' [120]. Besides, the MV model can explain the Cronin enhancement seen in the ratio $R_{\rm dAu}$ at $\eta = 0$ (see Sect. 5.2), and it has inspired simple parametrizations of the nuclear gluon distribution including both saturation and energy dependence which have been extensively used for phenomenological studies at RHIC [98].

To introduce this model, consider a nucleus in the infinite momentum frame (IMF) with momentum $P^+ \to \infty$. We will assume that the nucleus is of nearly infinite transverse extent with a uniform nuclear matter distribution (but the model can be extended to include realistic nuclear density profiles [117]). In the IMF, partons which carry large fractions of the nuclear momentum ("valence" partons), are Lorentz contracted to a distance $\sim 2R_A/\gamma$, with $\gamma = P^+/M = p^+/m_N$ (here, $p^+ = P^+/A$ is the longitudinal momentum of a single nucleon, and m_N is its mass, with $M = Am_N$). The 'wee' partons with longitudinal momentum fractions $x \ll 1$ are delocalized in the x^- direction over relatively large distances $\sim 1/xp^+$, which for $x \ll A^{-1/3}$ are much larger than the Lorentz–contracted nuclear diameter $2R_A/\gamma$ (recall that $R_A m_N \sim A^{1/3}$). These partons then "see" the valence partons as a very thin layer (essentially, a sheet) of color charge. The model assumes a simple kinematic distinction between 'wee' and valence partons. This is correct so long as the relevant values of x are not *too* small, namely such that the condition $\ln A^{1/3} \lesssim \ln(1/x) \ll 1/\alpha$ remains satisfied. Indeed, under this condition, one can neglect gluon radiation from partons at intermediate x' with $x < x' < 1$. For instance, for $A = 200$ and $\alpha \simeq 0.2$, one can consider values of x in the range $x \simeq 10^{-2} \div 10^{-1}$.

As discussed in Sect. 1.2.1, a small–x excitation has a relatively short lifetime, proportional
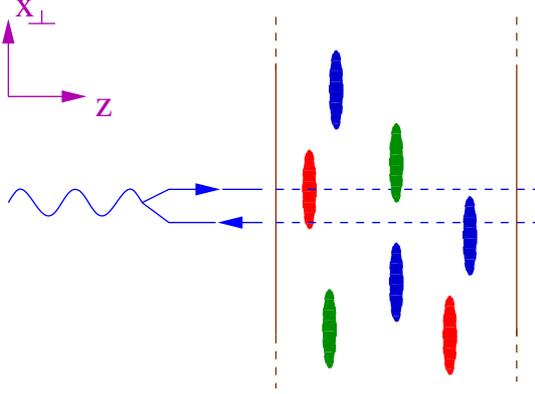
Fig. 19. *The nucleus scattering with a small external dipole serves as a microscope to measure to color charge of the valence quarks.* `MVvalence`

to x. This has been estimated in Eq. (1.7), with a result that can be easily recovered by using the uncertainty principle on the light–cone: $\Delta x^+ \sim 1/k^- \sim 2xp^+/k_\perp^2$. (We have used the LC dispersion relation for massless excitations, $k^- = m_\perp^2/2k^+ = k_\perp^2/2xp^+$, cf. Eq. (1.30).) Thus, the wee parton lifetime is much shorter than that of the valence partons: on the characteristic time scale for the former, the latter appear to live forever. The valence parton sources are thus *static sources of color charge*. Since their momenta are large, they are unaffected by absorbing or emitting soft quanta: they are recoilless sources of color charge. In this *eikonal approximation*, the wee parton cloud couples only to the "plus" component of the LC current, which, from the discussion here, can be written as :

$$J^{\mu,a}(x) = \delta^{\mu+}\delta(x^-)\rho^a(\boldsymbol{x})\,, \tag{1.40}$$

where $\rho^a(\boldsymbol{x})$ is the valence quark color charge density in the transverse plane. The δ–function in x^- assumes an infinitely thin sheet of color charge. The assumption must in fact be relaxed, for reasons to be discussed later; namely, $\delta(x^-)\rho^a(\boldsymbol{x}) \to \rho^a(x^-,\boldsymbol{x})$, where $\rho^a(x^-,\boldsymbol{x})$ is localized near $x^- = 0$, within a distance $\Delta x^- \sim 1/p^+$ which is again fixed by the uncertainty principle (this time applied to the valence partons). Note that ρ^a is *static*, i.e., independent of the LC time x^+, for the reasons explained previously.

We now turn to the color charge density ρ^a and show how it is generated. It is useful to imagine the nucleus interacting with an external probe like the virtual photon in DIS which has $x \ll A^{-1/3}$ and which can resolve distances $\Delta\boldsymbol{x}$ in the transverse plane that are much smaller than the nucleon size $\sim \Lambda_{\rm QCD}$ (see Fig. 19). This probe simultaneously couples to valence quarks from nucleons all along the nuclear diameter. On the small transverse scale $\Delta\boldsymbol{x}$, the confinement effects are unimportant, so the probe 'sees' the partons as sources of color charge. The number ΔN of these sources can be estimated as the product $n\Delta S_\perp$ between the density $n \equiv N_c A/\pi R_A^2$ of valence quarks in the transverse plane and the area $\Delta S_\perp \sim (\Delta\boldsymbol{x})^2$ covered by the external probe. That is:

$$\Delta N \approx n\,\Delta S_\perp \;=\; \Delta S_\perp\,\frac{AN_c}{\pi R_A^2} \;\sim\; \frac{\Lambda_{\rm QCD}^2}{Q^2}\,N_c A^{1/3}\,, \tag{1.41}$$

where $Q^2 \equiv 1/\Delta S_\perp$ is the external resolution, and we have used the fact that $R_A = R_0 A^{1/3}$ with $R_0 \sim 1/\Lambda_{\text{QCD}}$. Let us assume that, while being relatively hard, the transverse resolution Q^2 is still small enough to satisfy $\Lambda_{\text{QCD}}^2 \ll Q^2 \ll \Lambda_{\text{QCD}}^2 N_c A^{1/3}$. Then, the area $\Delta S_\perp$ is large enough to cover a large number $\Delta N \gg 1$ of valence quarks. These partons belong typically to different nucleons, so they are uncorrelated with each other, because of confinement in the longitudinal direction. Thus they are *randomly* distributed, in such a way that the total color charge $\mathcal{Q}^a$ seen by the external probe is the incoherent sum of the color charges of the individual partons. This implies

$$\langle \mathcal{Q}^a \rangle = 0, \qquad \langle \mathcal{Q}^a \mathcal{Q}^a \rangle = g^2 C_f \Delta N = \Delta S_\perp \frac{g^2 C_f N_c A}{\pi R_A^2}, \tag{1.42}$$

where we have used the fact that the color charge squared of a single quark is $g^2 t^a t^a = g^2 C_f$. One can treat this charge as *classical* since, when ΔN is large enough, we can ignore commutators of charges:

$$| \, [\mathcal{Q}^a, \mathcal{Q}^b] \, | = | \, i f^{abc} \mathcal{Q}^c \, | \ll \mathcal{Q}^2 \, . \tag{1.43}$$

In order to take the continuum limit (i.e., the limit where the transverse area $\Delta S_\perp$ of the tube is small), it is convenient to introduce the color charge densities $\rho^a(x^-, \boldsymbol{x})$ and

$$\rho^a(\boldsymbol{x}) \equiv \int dx^- \rho^a(x^-, \boldsymbol{x}). \tag{1.44}$$

Then,

$$\mathcal{Q}^a = \int_{\Delta S_\perp} d^2\boldsymbol{x} \, \rho^a(\boldsymbol{x}) = \int_{\Delta S_\perp} d^2\boldsymbol{x} \int dx^- \, \rho^a(x^-, \boldsymbol{x}), \tag{1.45}$$

which is consistent with Eq. (1.42) provided (recall that $C_f = (N_c^2 - 1)/2N_c$) :

$$\langle \rho_a(\boldsymbol{x})\rho_b(\boldsymbol{y}) \rangle_A = \delta_{ab}\delta^{(2)}(\boldsymbol{x} - \boldsymbol{y}) \, \mu_A, \qquad \mu_A \equiv \frac{g^2 A}{2\pi R_A^2} \, ,$$
$$\langle \rho_a(x^-, \boldsymbol{x})\rho_b(y^-, \boldsymbol{y}) \rangle_A = \delta_{ab}\delta^{(2)}(\boldsymbol{x} - \boldsymbol{y})\delta(x^- - y^-) \, \lambda_A(x^-),$$
$$\int dx^- \lambda_A(x^-) = \mu_A \, . \tag{1.46}$$

Here, $\mu_A \sim A^{1/3}$ is the average color charge squared of the valence quarks per unit transverse area and per color, and $\lambda_A(x^-)$ is the corresponding density per unit volume. The latter has some dependence upon x^-, whose precise form is, however, not important since the final formulae will involve only the integrated density μ_A. There is no explicit dependence upon $\boldsymbol{x}$ in μ_A or $\lambda_A(x^-)$ since we assume transverse homogeneity within the nuclear disk of radius R_A. Finally, the correlations are local in x^- since, as argued before, color sources at different values of x^- belong to different nucleons, so they are uncorrelated. All the higher-point, connected, correlation functions of ρ_a are assumed to vanish. The non–zero correlators (1.46) are generated by the following weight function [25] (with $\vec{x} = (x^-, \boldsymbol{x})$ and $\mathrm{d}^3x = \mathrm{d}x^-\mathrm{d}^2\boldsymbol{x}$) :

$$W_A[\rho] = \mathcal{N}\exp\left\{-\frac{1}{2}\int d^3x\,\frac{\rho_a(\vec{x})\rho_a(\vec{x})}{\lambda_A(x^-)}\right\}, \tag{1.47}$$

which is a Gaussian in ρ_a, with a local kernel. This is gauge-invariant (since local), so the variable ρ_a in this expression can be the color source in any gauge. The choice of a gauge will however soon become an issue when we shall study the dynamics of the gluons radiated by this random distribution of color charges.

The local Gaussian form of the weight function in Eq. (1.47) is valid, by construction, for a large nucleus, and within some restricted kinematical range that we spell here again, for more clarity. Specifically, this is correct for a transverse resolution Q^2 within the range $\Lambda_{\mathrm{QCD}}^2 \ll Q^2 \ll \Lambda_{\mathrm{QCD}}^2 N_c A^{1/3}$, and for relatively small $x \ll A^{-1/3}$, but such that $\alpha\ln(1/x) \ll 1$. For even smaller values of x or larger values of Q^2, the QCD quantum evolution cannot be neglected anymore, and the gluon distribution at the scale of interest is then dominated by the *quantum* color sources (as opposed to the valence quarks). Besides, for sufficiently high Q^2, the system becomes very dilute, and the classical approximation breaks down. Finally, the assumption that the valence quarks are uncorrelated must fail for transverse separations $\gtrsim 1/\Lambda_{\mathrm{QCD}}$, since the $N_c = 3$ valence quarks within the same nucleon are confined in a color singlet state. Thus, the total color charge, together with its higher multipolar moments, must vanish when measured over distances of the order of the nucleon size R_0, or larger. As emphasized by Lam and Mahlon [121] (see also Ref. [122]), the requirement of color neutrality can be included in the Gaussian weight function by replacing the δ–function in Eq. (1.46) with $\langle\rho^a(x^-,\boldsymbol{x})\rho^b(0)\rangle = \lambda(x^-,\boldsymbol{x})\delta^{ab}$, where $\lambda(x^-,\boldsymbol{x})$ is such that its Fourier transform $\lambda(x^-,\boldsymbol{k})$ vanishes rapidly at momenta $k_\perp \lesssim \Lambda_{\mathrm{QCD}}$.

Given the weight function (1.47) for the color charge density in the nucleus, observables related to the nuclear gluon distribution can be computed as follows [25, 28]: First, one needs to solve the classical field equations (the Yang–Mills equations) with the current (1.40), that is:

$$(D_\nu F^{\nu\mu})_a(x) = \delta^{\mu+}\rho_a(\vec{x}), \tag{1.48}$$

where $D_\nu = \partial_\nu - igA_\nu^a T^a$ with $(T^a)_{bc} = -if_{abc}$. The corresponding solution, to be denoted as $\mathcal{A}_a^\mu(x)$, is in general a non–linear functional of the charge density $\rho_a(\vec{x})$, whose precise form depends also upon the choice of a gauge. A rather general discussion of this classical solution will be given below, in Sect. 2.3. Then, the observable of interest — typically represented by some operator $\mathcal{O}[A^\mu]$ — is evaluated for $A^\mu = \mathcal{A}^\mu[\rho]$. This yields the value of the observable for a given configuration of the color sources in the nucleus (or for a *given event*). Finally, the physical observable is obtained by averaging $\mathcal{O}[\mathcal{A}^\mu]$ over all the configurations of the random source ρ, with the weight function (1.47). This specific averaging procedure reflects the separation of time scales in the problem — namely the fact that the changes in the distribution of the color sources occur over time scales much larger than the characteristic time scales for probing the system at high energy — and is reminiscent of a *glass* (see Sect. 2 below). As we shall discover in the subsequent sections, the separation in time scales and the corresponding 'glassy' structure is preserved by the dominant quantum evolution at high energy, which only modifies the functional form of the weight function $W_A[\rho]$. In general, this will depend also upon the energy, and will not

be a Gaussian anymore.

In what follows, we shall often return to the MV model (1.47), in particular in a study of gluon saturation through non–linear effects in the classical field equations, in Sect. 2.5, and in the discussion of the Cronin peak at central rapidity, in Sect. 5.2.

1.4   Motivation: The BFKL evolution and its small-x problem

We now return to the study of the energy dependence of the gluon distribution that we have begun in Sect. 1.2.1, and discuss this problem from the perspective of perturbative QCD. This will give us the opportunity to introduce the celebrated BFKL equation [16], which is the equation which governs the high energy behaviour of the gluon distribution (and, more generally, of scattering amplitudes in QCD) in the absence of saturation effects. This equation is relevant for the present discussion for several reasons: First, at a historical level, the BFKL equation marks the first attempt towards computing the high energy limit of QCD within perturbation theory. As we shall see, this amounts to the resummation of an infinite set of radiative corrections which are enhanced by the large energy logarithm $\ln s \sim \ln 1/x$. Second, this resummation results in conceptual problems whose analysis points towards the need for non–linear effects and saturation. In fact, much of the recent progress in understanding gluon saturation and unitarity corrections in high–energy QCD has been triggered by the 'small–x problem' of the BFKL equation, that we shall shortly discuss. Third, the BFKL resummation is a central ingredient of the modern theory for gluon saturation, the CGC theory, that we shall describe starting with the next section. Finally, the BFKL equation emerges as a particular limit (namely, the low density, or intermediate energy, limit) of the general evolution equations for the CGC. It turns out that a good comprehension of the BFKL solution is essential to be able to make progress with the more general, but also more complicated, equations.

In this section, our approach to the BFKL equation will be rather qualitative. We shall not explicitly derive the equation (since we shall do so, in at least a couple of different ways, in the subsequent sections), but rather we shall try to motivate its structure, and then discuss its solution, through successive iterations, by which we shall progressively refine our analysis.

1.4.1   The BFKL equation

We start by explaining why the resummation of the dominant radiative corrections at high energy leads necessarily to an *unstable growth* in the absence of non–linearities. To that aim, it suffices to rely on the bremsstrahlung law, Eq. (1.13), which shows that the radiation of small–x gluons is logarithmically enhanced. (The physical meaning of this enhancement will be discussed afterwards.) Using Eq. (1.13), let us compute the gluon distribution produced by an original color source, say a 'valence quark', at small x and in the presence of radiative corrections. The relevant diagrams are depicted in Fig. 20.

Fig. 20.a shows the tree–level process, that is, the direct emission of a soft (in the sense of 'small–x') gluon with longitudinal momentum $k^+ = xp^+ \ll p^+$ from the valence quark with momentum p^+. Fig. 20.b shows a diagram counting for the lowest–order radiative

41

corrections, and which describes the emission of an additional, quantum, gluon, with longitudinal momentum $k^+ < p_1^+ < p^+$. (Clearly, there are also other diagrams, like self-energy and vertex corrections, which enter at the same perturbative order, but their inclusion will not change our argument on the power counting at high energy.) From Eq. (1.13), the *relative* contribution of the process in Fig. 20.b is estimated as

$$\frac{\alpha N_c}{\pi} \int_{k^+}^{p^+} \frac{dp_1^+}{p_1^+} = \frac{\alpha N_c}{\pi} \ln \frac{p^+}{k^+} = \bar{\alpha} \ln \frac{1}{x} \tag{1.49}$$

where $\bar{\alpha} \equiv \alpha N_c / \pi$, and the color factor in Eq. (1.13) has been now taken as $C_A = N_c$, as appropriate for a gluon source. This quantum correction is enhanced by the large rapidity interval $\Delta \tau = \ln(1/x)$ available for the emission of the additional gluon.
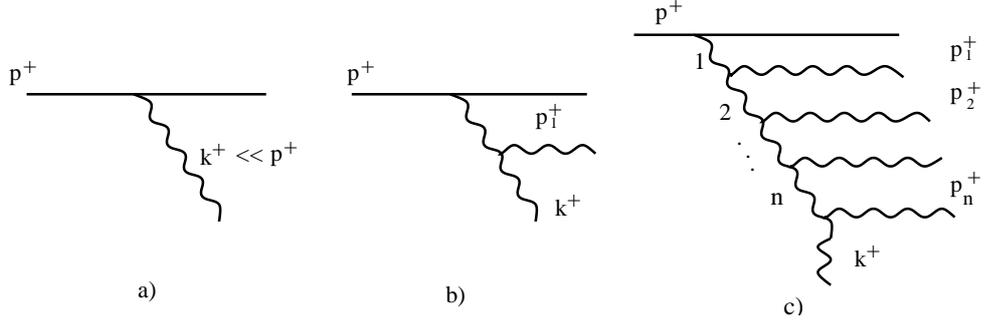
Fig. 20. a) Small–x gluon emission by a fast parton; b) the lowest-order radiative correction; c) a gluon cascade. The longitudinal momenta in this cascade are strongly ordered, and therefore so are also the lifetimes of the intermediate gluons, since we know that $\Delta x^+ \sim 2xp^+/k_\perp^2$.

A similar enhancement holds for the gluon cascade in Fig. 20.c, in which the successive gluons are strongly ordered in longitudinal momenta: $p^+ \gg p_1^+ \gg p_2^+ \gg \cdots \gg p_n^+ \gg k^+$. This gives a contribution of relative order

$$\frac{1}{n!} \left(\bar{\alpha} \ln \frac{1}{x} \right)^n , \tag{1.50}$$

where the factorial comes from the ordering in p^+. Clearly, when x is so small that $\bar{\alpha} \ln(1/x) \sim 1$, all such "higher–order corrections" become truly of order one, and must be resummed for consistency. A calculation which includes effects of order $(\alpha \ln(1/x))^n$ to all orders in n is said to be valid to "leading logarithmic accuracy" (LLA).

All the gluon cascades depicted in Fig. 20 contribute to the production of the final gluon with longitudinal fraction $x \ll 1$. Thus, in order to compute the number of produced gluons per unit rapidity — which is the same as the gluon distribution — one needs to resum such processes to all orders. Given the form of the general term in Eq. (1.50), this resummation will clearly produce an exponential:

$$\frac{dN}{d\tau} \equiv xG(x, Q^2) \sim e^{\omega \bar{\alpha} \tau} = x^{-\omega \bar{\alpha}} , \tag{1.51}$$

with ω a pure number which cannot be determined from our present, qualitative, argument. We have tacitly assumed above that all the gluons in the cascade have transverse momenta of the same order, namely of order Q. As we shall shortly see, a more refined treatment, based on the BFKL equation [16], allows one to compute ω and specifies the Q^2–dependence, and also the subleading τ–dependence (beyond the exponential behaviour shown in Eq. (1.51)) of the gluon distribution.

As anticipated, the gluon distribution in the BFKL approximation rises exponentially with the rapidity, that is, as a power of s. Let us try to understand this rapid growth in more physical terms. The crucial observation is that in a gluon cascade like that in Fig. 20.c, in which the longitudinal momenta are strongly ordered, all the gluons above the final one have a much longer lifetime than the latter, and thus act as a frozen color charge distribution for the emission of the last gluon. Gluons within a same cascade are coherent with each other (since they live all at the same time), but different cascades are statistically independent. Thus, the total color charge in the system after a rapidity evolution τ is the incoherent sum of the color charges of the individual cascades. Therefore, the average color charge squared $\langle \mathcal{Q}^a \mathcal{Q}^a \rangle_\tau$ — which represents the source for the emission of a new gluon — is proportional to the number $N(\tau)$ of gluons produced in the previous steps. Thus, the probability for the emission of a new gluon is proportional to the number of preexisting gluons, which is clearly the condition for an exponential instability.

To deduce the general form of the BFKL equation, the bremsstrahlung formula, Eq. (1.13), is not enough; indeed, the latter assumes that both x and $k_\perp$ are small, whereas in the BFKL evolution there is no ordering in transverse momenta, that is, the $k_\perp$–dependence is treated exactly. For this reason, the ensuing equation is non–local in transverse momenta, and applies to the so–called 'unintegrated gluon distribution' $f(x, k_\perp^2)$, defined through:

$$xG(x, Q^2) \equiv \int^{Q^2} \frac{\mathrm{d}k_\perp^2}{k_\perp^2} \, f(x, k_\perp^2) \iff f(x, k_\perp^2) = \frac{\partial \, xG(x, k_\perp^2)}{\partial \ln k_\perp^2} \,. \tag{1.52}$$

It is furthermore useful to relate this quantity to the gluon spectrum

$$xG(x, Q^2) = \int^{Q^2} \mathrm{d}^2\boldsymbol{k} \, \frac{\mathrm{d}N}{\mathrm{d}\tau \mathrm{d}^2\boldsymbol{k}} \iff f(x, k_\perp^2) = \pi \boldsymbol{k}^2 \, \frac{\mathrm{d}N}{\mathrm{d}\tau \mathrm{d}^2\boldsymbol{k}} \,. \tag{1.53}$$

The BFKL equation for $f(x, k_\perp^2) \equiv f(\tau, \boldsymbol{k}^2)$ reads then [16]

$$\frac{\partial f(\tau, \boldsymbol{k}^2)}{\partial \tau} = \bar{\alpha} \int \frac{\mathrm{d}^2\boldsymbol{p}}{\pi} \frac{\boldsymbol{k}^2}{\boldsymbol{p}^2(\boldsymbol{k} - \boldsymbol{p})^2} \left\{ f(\tau, \boldsymbol{p}^2) - \frac{1}{2} f(\tau, \boldsymbol{k}^2) \right\} . \tag{1.54}$$

The diagrammatic interpretation of this equation is illustrated in Fig. 21, which should be read as one step in the evolution: the straight line there represents a relatively fast gluon created at the previous step in the evolution, and which acts as a source for the final gluon at small x. When increasing the rapidity from τ to $\tau + \mathrm{d}\tau$, this last emission is dressed by real and virtual quantum corrections of order $\bar{\alpha}\mathrm{d}\tau$. The three diagrams on the left describe real gluon emission, and are represented in Eq. (1.54) by the first term, $f(\tau, \boldsymbol{p}^2)$, within the braces. The three diagrams on the right describe vertex and self–energy corrections, and are represented by the second term, $-(1/2)f(\tau, \boldsymbol{k}^2)$, within the braces. Note that,
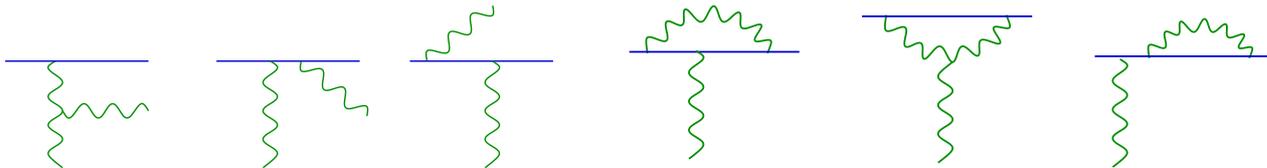
Fig. 21. *Real (left) and virtual (right) contributions to the first step in the BFKL evolution of the gluon distribution.* `BFKLfig`

taken separately, the real and virtual contributions to Eq. (1.54) have poles giving rise to *infrared singularities* : at $\boldsymbol{p} = 0$ and $\boldsymbol{p} = \boldsymbol{k}$ for the virtual term, but only at $\boldsymbol{p} = \boldsymbol{k}$ for the real one (since, as we shall shortly see, $f(\tau, \boldsymbol{p}^2) \to 0$ as $\boldsymbol{p} \to 0$). However, these divergences compensate between the two types of contributions, so the overall integral is indeed well defined, both in the infrared and in the ultraviolet (where the kernel decreases fast enough, as $1/\boldsymbol{p}^4$).

1.4.2 The BFKL solution

The solution to Eq. (1.54) valid at high energies will be explicitly constructed in Sect. 4.2. Here, we shall simply exhibit the corresponding result and comment on its properties. But before doing so, it is instructive to also exhibit an initial condition for this equation at low energies, that we shall choose as the gluon distribution of a single quark (this will facilitate appreciating the effects of the evolution). Namely, the differential probability for bremsstrahlung formula, Eq. (1.13), can be directly identified as the spectrum of the soft gluons emitted by a single quark:

$$\frac{\mathrm{d}N}{\mathrm{d}x\,\mathrm{d}\boldsymbol{k}^2} \equiv \frac{\mathrm{d}P_{\mathrm{Brem}}}{\mathrm{d}x\,\mathrm{d}\boldsymbol{k}^2} \simeq \frac{\alpha C_F}{\pi}\frac{1}{\boldsymbol{k}^2}\frac{1}{x} \qquad \text{for} \qquad x \ll 1\,, \tag{1.55}$$

which together with Eq. (1.53) immediately implies:

$$f_0(x, \boldsymbol{k}^2) \simeq \frac{\alpha C_F}{\pi}\,. \tag{1.56}$$

This estimate is valid so long as x is small, but not *too* small, such that $\alpha \ln(1/x) \ll 1$, and for perturbative transverse momenta $k_\perp \gg \Lambda_{\mathrm{QCD}}$. (At smaller momenta $k_\perp \lesssim \Lambda_{\mathrm{QCD}}$, one expects f_0 to vanish rapidly with decreasing $k_\perp$, because of confinement.) Within its validity range, the distribution (1.56) is independent upon both x and $\boldsymbol{k}$. However, such dependencies will be induced by the quantum evolution. Specifically, when adapted to the unintegrated gluon distribution $f(x, k_\perp^2)$, the BFKL solution of Sect. 4.2 reads:

$$f(\tau, \boldsymbol{k}^2) \simeq \left(\frac{\boldsymbol{k}^2}{Q_0^2}\right)^{1/2} \frac{\mathrm{e}^{\omega_0\bar{\alpha}\tau}}{\sqrt{2\pi\beta_0\bar{\alpha}\tau}} \exp\left\{-\frac{\ln^2\left(\boldsymbol{k}^2/Q_0^2\right)}{2\beta_0\bar{\alpha}\tau}\right\}, \tag{1.57}$$

where $\tau = \ln 1/x$ and Q_0 is a non–perturbative scale of order Λ_{QCD} introduced by the initial condition (roughly, the scale below which f_0 changes from the constant behaviour

shown in Eq. (1.56) to a rapidly decreasing function of $k_\perp$); furthermore, $\omega_0 = 4 \ln 2 \approx 2.77$ and $\beta_0 \approx 33.67$ are pure numbers determined by the BFKL equation.

Eq. (1.57) is valid in the *formal* high energy limit $\bar{\alpha}\tau \gg \rho$, with $\rho \equiv \ln(k_\perp^2/Q_0^2)$. This limit is formal in the sense that it neglects saturation effects, i.e., it assumes that the linear BFKL equation (1.54) continues to apply up to arbitrarily high energies. But from the discussion at the end of Sect. 1.2.1, we know that the linear approximation is valid only for transverse momenta above the saturation scale $Q_s(\tau)$, which is itself exponentially rising with τ. Therefore, Eq. (1.54) is strictly correct only for $\rho > \rho_s(\tau)$, with $\rho_s(\tau) \equiv \ln(Q_s^2/Q_0^2) \propto \tau$, which shows that the only meaningful 'high energy limit' of the BFKL equation is the limit in which τ and ρ are simultaneously increased, in such a way that the condition $\rho > \rho_s(\tau)$ remains satisfied at any τ (that is, the evolution must remain on the right side of the saturation line in Fig. 8). This being said, the approximation (1.57) is still interesting, for the following reasons: First, even in the presence of saturation, this is a correct approximation within some intermediate range of energies[2]. Second, if formally extrapolated at high energy, this equation illustrates the would–be conceptual difficulties of perturbative QCD in the absence of saturation.

First of all, Eq. (1.57) confirms that the dominant behaviour at high energy is an exponential increase with τ, as anticipated in Eq. (1.51), and provides us with the value of the relevant exponent: $\omega = \omega_0 \equiv 4 \ln 2$. (The quantity $\alpha_\mathbb{P} = 1 + (4 \ln 2)\bar{\alpha}$ is sometimes referred to as the 'BFKL intercept'.) Thus, the BFKL equation predicts a gluon distribution rising at small x as $(1/x)^{\lambda_g}$, with $\lambda_g = \alpha_\mathbb{P} - 1 \approx 0.5$ (for $\alpha \simeq 0.2$). This is in a qualitative agreement with the behaviour observed at HERA, but the above value of λ_g is substantially larger than that extracted from the data. In fact, if one tries a parametrization $F_2(x, Q^2) \sim x^{-\lambda_{\text{eff}}(Q^2)}$ for the rise of the HERA data at $x < 0.01$, then one finds that $\lambda_{\text{eff}}(Q^2)$ falls from 0.35 to 0.1 when Q^2 decreases from 200 GeV^2 to 1 GeV^2.

But by itself, this inability of the BFKL equation to predict the right value for λ_g should not be seen as a serious problem: So far, this is only a leading–order (LO) calculation in QCD, but higher order effects (in particular, the running of the coupling) are in fact known to be important, and to lower the value of λ_g. Indeed, in recent years the next–to–leading–order (NLO) corrections to the BFKL equation become available [18]. When properly interpreted [19], the effect of these corrections is to reduce the gluon exponent λ_g roughly by a factor of two, thus yielding a value in qualitative agreement with the HERA data for intermediate values of Q^2 (from 10 to 100 GeV^2). Although the full calculation of DIS at NLO–BFKL level has not been yet completed [**?**], one can quite safely conclude that the small–x evolution in perturbative QCD is not excluded by the data.

What *would be* a serious problem on the other hand, would be to forget about saturation and assume that the exponential increase with τ (even with a smaller value for the intercept) persists up to arbitrarily high energies. First, this would contradict our qualitative argument at the end of Sect. 1.2.1 about the importance of recombination effects

---

[2] Note indeed that the saturation line in Fig. 8 does not start at the origin, but rather at some large rapidity τ_0: $\rho_s(\tau) \approx c\bar{\alpha}(\tau - \tau_0)$, with $c \sim 1$ and $\bar{\alpha}\tau_0 \gg 1$. Thus, so long as τ is not much larger than τ_0, there exists a range in ρ in which the conditions $\bar{\alpha}\tau \gg \rho$ and $\rho > \rho_s(\tau)$ are both satisfied. Eq. (1.57) applies within this range.

in the high density regime. Second this would lead to unitarity violations in scattering at high energy. Let us consider the unitarity issue in the context of DIS. We have seen previously that the small–x increase of the structure function F_2 is driven by the gluon distribution (see, e.g., Eq. (1.14)). More precisely, within the dipole factorization of DIS, the unitarity problem refers to the dipole–hadron scattering. In Sect. 2.4 we shall show that in the *single–scattering* approximation (as obtained after expanding the Wilson lines in Eq. (1.37) to second order in the field gA^+ in their exponent), the dipole cross–section is linearly related to the unintegrated gluon distribution of Eq. (1.52):

$$\sigma_{\text{dipole}}(x, \boldsymbol{r}) \approx \frac{4\pi}{N_c}\, \alpha \int \frac{\mathrm{d}^2 \boldsymbol{k}}{\boldsymbol{k}^4}\, f(x, \boldsymbol{k}^2) \left(1 - \mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{r}}\right). \tag{1.58}$$

(For $r_\perp \to 0$ and by using Eq. (1.52), the equation above immediately reduces to Eq. (1.24), which is in turn equivalent to Eq. (1.14). Also, by inserting Eq. (1.58) into the dipole factorization (1.22) and using the known expression for the virtual photon wavefunction $|\Psi(z, \boldsymbol{r}; Q^2)|^2$, one obtains the DIS cross–section in '$k_\perp$–factorized' form [123].)

Clearly, if $f(x, \boldsymbol{k}^2)$ rises as a power of $1/x \sim s$, so does also the dipole cross–section, in violation of the *Froissart bound* [2], which stipulates that in the high enery limit a total cross–section cannot grow faster than $\ln^2 s$. At this point, one may remark that, in QCD, the issue of the Froissart bound cannot be fully addressed within perturbation theory, and thus goes beyond the applicability of Eq. (1.58) (see the discussion in Sect. 4.4). But the power–like increase of Eq. (1.58) with the energy leads also to a more severe form of unitarity violation, which should remain within the realm of perturbative QCD: The dipole cross–section at high energy can be obtained by integrating the *scattering amplitude*[3] $T_\tau(\boldsymbol{r}, \boldsymbol{b}) \equiv 1 - \mathrm{Re}\, S_\tau(\boldsymbol{r}, \boldsymbol{b})$ over all the impact parameters $\boldsymbol{b}$ within the hadron disk (cf. Eq. (1.23)). With increasing s, the hadron area can rise, at most, as a power of $\ln s$. Thus, the power–law increase of Eq. (1.58) implies a similar increase for the scattering amplitude, which would eventually violate the unitarity bound $T \leq 1$ at fixed impact parameter. But for sufficiently large energy, the gluon density at $\boldsymbol{b}$ should be large enough for perturbation theory to apply.

Now, in view of the previous discussion in Sects. 1.2.2 and 1.2.3, it is clear that this violation of the unitarity bound is a consequence of our restriction to a single–scattering approximation: Before expanding the Wilson lines, the dipole S–matrix (1.37) was manifestly unitary ! So, one may be tempted to argue that unitarity arises through multiple scattering independently of saturation. (For instance, the Glauber–like resummation in Eq. (1.27) would preserve unitarity independently of the fate of the gluon distribution at high energy.) Recall however that it is the same parameter, namely the strength gA^+ of the gluon field in the target, which controls both multiple scattering and recombination effects. Thus, in general[4], it would be inconsistent to take into account multiple scattering for the purpose of restoring the unitarity and at the same time neglect non–linear effects in the target wavefunction.

---

[3] The amplitude T_τ thus defined is in fact the *imaginary part* of the more conventional scattering amplitude $\mathcal{A}$, which is defined through $S \equiv 1 + i\mathcal{A}$.

[4] This is nevertheless possible in special frames, as we shall discuss in Sect. 6.

Returning to Eq. (1.57), let us focus on a different aspect of the BFKL approximation, namely the $k_\perp$–dependence of the gluon distribution at high energy. Recall first that the low–energy distribution (1.56) is independent of $k_\perp$, which reflects the bremsstrahlung spectrum in Eq. (1.55). If one inserts f_0 from Eq. (1.56) into the dipole cross–section (1.58), one finds a logarithmic infrared divergence due to the too rapid rise of the bremsstrahlung spectrum at small $k_\perp$. Clearly, when $k_\perp \lesssim \Lambda_{\rm QCD}$, we expect this rise to be tamed by confinement, but it is interesting to see that the BFKL resummation is already effective in reducing the infrared sensitivity of the gluon distribution: The dominant $k_\perp$–dependence of Eq. (1.57) is encoded in the overall factor $\sqrt{\boldsymbol{k}^2}$, showing that the BFKL spectrum $f(x, \boldsymbol{k}^2)/\boldsymbol{k}^2$ has a slower rise — as $1/\sqrt{\boldsymbol{k}^2}$ — in the infrared. In particular, when evaluated with the BFKL distribution, the dipole cross–section (1.58) is infrared–finite. Physically, this improved infrared behaviour comes from the partial *screening* of the original color charge (a quark in the case of Eq. (1.56)) by the gluons radiated in the evolution (which are correlated with each other, so they also screen themselves). As we shall explain in Sect. 4.1, this BFKL screening is only *partial*, in the sense that there is no finite correlation length (smaller than $1/\Lambda_{\rm QCD}$) over which color neutrality is achieved. Yet, from the above example of the dipole scattering, one may conclude that the BFKL physics is infrared–safe.

But this conclusion is naive, as shown by the following argument due to Bartels: Consider also the sub–dominant $k_\perp$ and τ–dependencies of the BFKL solution, as encoded in the function $\psi(\tau, \rho)$ defined by rewriting Eq. (1.57) as

$$f(\tau, \rho) \,=\, {\rm e}^{\rho/2 + \omega_0 \bar\alpha \tau}\, \psi(\tau, \rho)\,, \qquad \psi(\tau, \rho) \,\equiv\, \frac{1}{\sqrt{2\pi \beta_0 \bar\alpha \tau}}\, \exp\left\{ -\frac{\rho^2}{2\beta_0 \bar\alpha \tau} \right\}. \qquad (1.59)$$

The function $\psi(\tau, \rho)$ obeys the diffusion equation (with $D_0 \equiv \bar\alpha \beta_0$)

$$\frac{\partial}{\partial \tau}\, \psi(\tau, \rho) \,=\, \frac{D_0}{2}\, \frac{\partial^2}{\partial \rho^2}\, \psi(\tau, \rho), \qquad (1.60)$$

with initial condition $\psi(0, \rho) = \delta(\rho)$. The solution has the 'completeness' property:

$$\psi(\tau, \rho) = \int {\rm d}\rho_1\, \psi(\tau - \tau_1, \rho - \rho_1)\, \psi(\tau_1, \rho_1), \qquad (1.61)$$

for arbitrary τ_1 within the range $0 < \tau_1 < \tau$. Moreover, as it is clear from Eq. (1.59), the above integral over ρ_1 receives significant contributions from all the points which are located around ρ within a distance of the order of the 'diffusion radius' $\sqrt{2D_0(\tau - \tau_1)}$. Thus, when both τ and τ_1 are large enough, the intermediate point ρ_1 can diffuse far away from both the original point $\rho_0 = 0$ (i.e., $k_\perp = Q_0$) and the final point ρ. This diffusion is symmetric — it proceeds through points which are either larger or smaller than ρ and ρ_0 —, and for large enough τ it will ineluctably penetrate the non–perturbative domain at $\rho_1 \lesssim \rho_{\rm QCD} \equiv \ln \Lambda_{\rm QCD}^2 / Q_0^2$. That is, even if one starts with a hard scale $Q_0^2 \gg \Lambda_{\rm QCD}^2$ at low energy, and one measures the evolved gluon distribution at a hard scale as well ($k_\perp^2 \gg \Lambda_{\rm QCD}^2$), after a sufficiently large evolution 'time' $\bar\alpha \tau$ the measured distribution will receive significant contributions from the non–perturbative, soft, sector, where the BFKL approximation is not under control. This property is generally referred to as the *'infrared diffusion problem'* of the BFKL equation.

The power–like increase of the gluon distribution with the energy and the infrared diffusion are conceptual difficulties of the BFKL approximation, often referred to as the 'small–x problem of perturbative QCD'. But this problem is an artifact of having neglected non–linear effects in the evolution towards high energy: For sufficiently small x and/or small $k_\perp$, the gluon occupation factors become so large that recombination processes are favoured and prevent the further growth of the distribution. This mechanism introduces an intrinsic hard scale — the saturation momentum Q_s — which implies that the coupling is weak and at the same time suppresses the infrared diffusion (since the intermediate point ρ_1 in the argument above cannot penetrate the saturated region at $\rho < \rho_s$ where the occupation factor has already reached its maximally allowed value $\sim 1/\alpha$). Therefore, saturation makes a compelling situation for the applicability of perturbative QCD at high energy.

1.4.3 Dipole scattering and the BK equation

In the previous discussion, we have mentioned that the dipole cross–section in the single scattering approximation is proportional to the unintegrated gluon distribution, cf. Eq. (1.58). This implies that the dipole amplitude itself should satisfy (within this approximation) a corresponding version of the BFKL equation. It turns out that the respective equation can be rather easily derived provided we change our perspective over the evolution. So far, we have always assumed that the rapidity of the projectile dipole was fixed, whereas the rise of the scattering amplitude with the energy was due to the evolution of the gluon distribution in the target. Alternatively, one can use the rapidity increment $d\tau$ to accelerate the dipole, which then evolves by emitting one small–x virtual gluon, and then study the scattering between the ensuing dipole–gluon system and the (unevolved) target. This point of view of *projectile evolution* turns out to be extremely powerful, as it allows one to give a quick derivation not only of the BFKL equation for the dipole amplitude, but also of a non–linear generalization to this equation, the Balitsky–Kovchegov (BK) equation [39, 124], which takes into account the unitarity corrections.

Specifically, in Sect. 6 we shall estimate the differential probability for the emission of a small–x gluon from a dipole. This is most conveniently written in coordinate space, where it reads:

$$dP = \frac{\alpha N_c}{2\pi^2}\, \mathcal{M}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})\, d^2\boldsymbol{z}\, d\tau \,, \qquad \mathcal{M}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \equiv \frac{(\boldsymbol{x} - \boldsymbol{y})^2}{(\boldsymbol{x} - \boldsymbol{z})^2 (\boldsymbol{y} - \boldsymbol{z})^2} \,, \qquad (1.62)$$

where $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{z}$ are the transverse coordinates of the quark, the antiquark, and the emitted gluon, respectively. This probability vanishes when $r \equiv |\boldsymbol{x} - \boldsymbol{y}| \to 0$, as expected (since a zero–size dipole is non–interacting), and it becomes singular when the emitted gluon is collinear with either the quark or the antiquark. The gluon can be effectively replaced [5] by a zero–size $q\bar{q}$ pair, and the gluon emission then appears as the splitting of

---

[5] In general, this replacement is correct only in the limit in which the number of colors N_c is large (see Sect. 6); however, it turns out that the would–be suppressed terms at large N_c cancel exactly when computing the emission of a soft gluon from an elementary color dipole. Thus the equation to be established here holds for arbitrary N_c.

the original dipole $(\boldsymbol{x}, \boldsymbol{y})$ into two new dipoles: $(\boldsymbol{x}, \boldsymbol{z})$ and $(\boldsymbol{z}, \boldsymbol{y})$.

If the emitted gluon is in the wavefunction of the dipole at the time it scatters on the target, then what scatters is a system of two dipoles. If the gluon is not in the wavefunction at the time of the scattering, it can be viewed as the "virtual" term which decreases the probability that the original quark–antiquark pair remain a simple dipole, thus compensating the probability for the two–dipole state. The whole process can be summarized into the following evolution equation for the dipole S–matrix $\langle S(\boldsymbol{x}, \boldsymbol{y})\rangle_\tau$ [39]:

$$\frac{\partial}{\partial \tau}\langle S(\boldsymbol{x}, \boldsymbol{y})\rangle_\tau = \frac{\bar{\alpha}_s}{2\pi}\int \mathrm{d}^2\boldsymbol{z}\,\mathcal{M}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})\left\{-\langle S(\boldsymbol{x}, \boldsymbol{y})\rangle_\tau + \langle S(\boldsymbol{x}, \boldsymbol{z})S(\boldsymbol{z}, \boldsymbol{y})\rangle_\tau\right\}, \qquad (1.63)$$

where $\langle S(\boldsymbol{x}, \boldsymbol{z})S(\boldsymbol{z}, \boldsymbol{y})\rangle_\tau$ stands for the scattering of the two–dipole system on the target. Note that, for a *fixed* configuration of the color fields in the target (that is, for a given event) the two child dipoles scatter independently from each other, so that the two–dipole scattering operator is simply the product of the corresponding operators for the individual dipoles (in turn given by Eq. (1.36)): $S^{(2)}(\boldsymbol{x}, \boldsymbol{z}; \boldsymbol{z}, \boldsymbol{y}) = S(\boldsymbol{x}, \boldsymbol{z})S(\boldsymbol{z}, \boldsymbol{y})$. However, the color fields from different configurations are generally correlated with each other (in particular, because of the correlations built in the course of the evolution); because of that, the average over the target wavefunction will generally introduce a *correlation* between the scattering of the two dipoles: $\langle S(\boldsymbol{x}, \boldsymbol{z})S(\boldsymbol{z}, \boldsymbol{y})\rangle_\tau \neq \langle S(\boldsymbol{x}, \boldsymbol{z})\rangle_\tau\langle S(\boldsymbol{z}, \boldsymbol{y})\rangle_\tau$. Thus Eq. (1.63) is not a closed equation, but only the first equation in an hierarchy which turns out to be infinite and to couple scattering operators with an increasingly complicated (color and spatial) structure, but which are all expressed in terms of Wilson lines. This hierarchy has been first derived by Balitsky [39] — precisely by studying the projectile evolution — and then rederived from target evolution within the CGC formalism (see Sect. 3.5).

In order to recognize the BFKL equation, it is preferable to rewrite Eq. (1.63) in terms of the *scattering amplitude* $T = 1 - S$. The corresponding equation reads

$$\frac{\partial}{\partial \tau}\langle T(\boldsymbol{x}, \boldsymbol{y})\rangle_\tau = \frac{\bar{\alpha}_s}{2\pi}\int \mathrm{d}^2\boldsymbol{z}\,\mathcal{M}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \qquad (1.64)$$
$$\left\{-\langle T(\boldsymbol{x}, \boldsymbol{y})\rangle_\tau + \langle T(\boldsymbol{x}, \boldsymbol{z})\rangle_\tau + \langle T(\boldsymbol{z}, \boldsymbol{y})\rangle_\tau - \langle T(\boldsymbol{x}, \boldsymbol{z})T(\boldsymbol{z}, \boldsymbol{y})\rangle_\tau\right\},$$

and is illustrated with a few Feynman graphs in Fig. 22. For simplicity, in this figure we represent the scattering between an elementary dipole and the target in the two–gluon exchange approximation (single scattering). But in the high energy (or strong field) regime, the number of exchanged gluons can be arbitrary, as shown by the definition (1.36) of the dipole S–matrix in terms of Wilson lines. In Eq. (1.64), the multiple scattering is encoded in the last term, which describes the simultaneous scattering of the two child dipoles. When interpreted as target evolution, this non–linear terms corresponds to saturation effects (gluon merging in the target wavefunction), as it will be explained in Sect. 3.5.

The BFKL equation emerges as the weak scattering limit of Eq. (1.64): For relatively low energies, such that the target is dilute, the scattering amplitude is small: $T \ll 1$ (recall that $|S|^2 = |1-T|^2$ represents the probability that no interaction take place in the collision; for a dilute target, this should be close to one), and the term describing the scattering of two dipoles in the r.h.s. of Eq. (1.64) should be even smaller: $\langle T\,T\rangle \ll \langle T\rangle \ll 1$. We can then neglect this term and thus obtain a linear equation for the dipole amplitude:
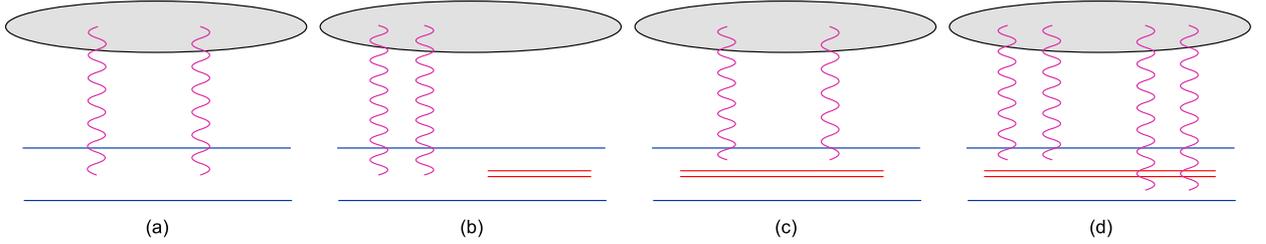
Fig. 22. *Diagrams for the evolution of the dipole scattering amplitude, cf. Eq. (1.64): (a) the tree–level contribution; (b) the virtual correction* $-\langle T(\boldsymbol{x},\boldsymbol{y})\rangle$; *(c) the scattering of one child dipole,* $\langle T(\boldsymbol{x},\boldsymbol{z})\rangle$ *or* $\langle T(\boldsymbol{z},\boldsymbol{y})\rangle$; *(d) the simultaneous scattering of both child dipoles,* $\langle T(\boldsymbol{x},\boldsymbol{z})T(\boldsymbol{z},\boldsymbol{y})\rangle$. `FIG_ODIP`

$$\frac{\partial}{\partial\tau}\langle T(\boldsymbol{x},\boldsymbol{y})\rangle_\tau = \frac{\bar{\alpha}_s}{2\pi}\int_{\boldsymbol{z}}\mathcal{M}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})\left\{-\langle T(\boldsymbol{x},\boldsymbol{y})\rangle_\tau + \langle T(\boldsymbol{x},\boldsymbol{z})\rangle_\tau + \langle T(\boldsymbol{z},\boldsymbol{y})\rangle_\tau\right\}, \quad (1.65)$$

which is recognized as the coordinate–space version of the BFKL equation. (In this form, the BFKL equation has been first derived by Mueller, within the dipole picture [58]; see Sect. 6.) One can check that this equation is indeed consistent, via a Fourier transform, with the BFKL equation (1.54) for the gluon distribution together with the relation (1.58) between σ_{dipole} and $f(x,\boldsymbol{k}^2)$. By linearizing Eq. (1.64) we have lost all the information about multiple scattering: Eq. (1.65) describes single scattering with the exchange of a 'BFKL pomeron' (a two–gluon exchange in the t–channel dressed by the BFKL evolution).

A closed, but *non–linear*, approximation to Eq. (1.64), which does preserve the unitarity corrections (while loosing some correlations due to fluctuations, however), is obtained by assuming the factorization property

$$\langle T(\boldsymbol{x},\boldsymbol{z})T(\boldsymbol{z},\boldsymbol{y})\rangle_\tau \approx \langle T(\boldsymbol{x},\boldsymbol{z})\rangle_\tau\,\langle T(\boldsymbol{z},\boldsymbol{y})\rangle_\tau, \quad (1.66)$$

which is a *mean field approximation* (MFA) for the gluon fields in the target. This yields

$$\frac{\partial}{\partial\tau}\langle T(\boldsymbol{x},\boldsymbol{y})\rangle_\tau = \frac{\bar{\alpha}_s}{2\pi}\int_{\boldsymbol{z}}\mathcal{M}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}) \quad (1.67)$$
$$\left\{-\langle T(\boldsymbol{x},\boldsymbol{y})\rangle_\tau + \langle T(\boldsymbol{x},\boldsymbol{z})\rangle_\tau + \langle T(\boldsymbol{z},\boldsymbol{y})\rangle_\tau - \langle T(\boldsymbol{x},\boldsymbol{z})\rangle_\tau\,\langle T(\boldsymbol{z},\boldsymbol{y})\rangle_\tau\right\},$$

an equation originally derived by Kovchegov [124] (within Mueller's dipole picture [43, 58]) and commonly referred nowadays as the *Balitsky–Kovchegov (BK) equation*. Remarkably, this equation preserves the unitarity constraint $T \leq 1$ during the evolution and predicts that the upper bound $T = 1$ (the 'black disk limit') will be eventually approached at sufficiently high energy. The corresponding solution will be discussed in Sect. 4.2.

The MFA (1.66) is tantamount to neglecting fluctuations in the gluon distribution in the target. As we have seen in the discussion of the MV model in Sect. 1.3.3, this is a good approximation if the target is a large nucleus at not so high energy. More generally, Eq. (1.66) should work reasonably well when the scattering is sufficiently strong, that is, when $\langle T\rangle_\tau$ is not much smaller than one, because in that case the external dipoles scatter off a high–density gluonic system, for which the *density fluctuations* are relatively unimportant. On the other hand, we know that, with increasing energy, the BFKL evolution

induces non–trivial correlations in the gluon distribution, and we expect such correlations to play an essential role in the dilute regime, where the scattering is relatively weak.

Deriving non–linear evolution equations like Eq. (1.64) and understanding the validity range and also the limitations of approximations like Eq. (1.66) will be one of the main goals of the forthcoming developments in this work.

1.4.4   Modern developments of the BFKL theory

To be rewritten

Since the advent of the Balitsky–Fadin–Kuraev–Lipatov (BFKL) equation [16, 125] in the mid seventies, there has been significant progress in our comprehension of high–energy QCD, and several theoretical approaches have been proposed which aim at a resummation of the energy–enhanced radiative corrections to high–energy processes in perturbative QCD. The BFKL equation is a *leading logarithmic approximation* (LLA), which allows one to resum to all orders corrections of the form $(\alpha \ln s)^n$ to the scattering between two colorless objects via the exchange of two gluons in the t–channel. As a result of this resummation, the bare two–gluon exchange is replaced by the *BFKL pomeron* (the sum of an infinite series of ladder diagrams of ordinary perturbation theory), or, equivalently, by two *reggeized* gluons which interact with each other. All the subsequent theoretical approaches proposed within perturbative QCD encompass the BFKL equation, and can be viewed as extensions of the latter towards increasing the complexity of the objects exchanged in the t–channel, and also towards enlarging the limits of the LLA.

The simplest object beyond the BFKL pomeron within perturbative QCD is the exchange of three interacting (reggeized) gluons in a symmetric color state. This object, which is negative (or "C–odd") under charge conjugation ($C = -1$), and has also negative parity ($P = -1$), represents the lowest order perturbative contribution to the *odderon*, the C–odd exchange which dominates the difference between the hadronic cross sections for direct and crossed channel processes at very high energies [8]. The evolution of the three–gluon odderon exchange with increasing energy in the LLA is described by the BKP equation, established by Bartels [126] and Kwiecinski and Praszalowicz [127] (see also the work by Jaroszewicz, Ref. [128]), which amounts to a pairwise iteration of the BFKL kernel. This equation can be immediately extended to describe the exchange of an arbitrary number $n \geq 3$ of reggeized gluons with pairwise BFKL interactions [126, 129–132]. The resulting formalism, also known as the *generalized leading logarithmic approximation* (GLLA), resums all radiative corrections that involve the maximally possible number of energy logarithms $\ln s$ for a given number of exchanged gluons. At the moment, two exact solutions of the BKP equation for odderon evolution are available [133, 134], and the subject continues to be under intensive debate [135, 136] (see also the recent review paper [137] and the discussion below).

In the formalisms described so far, the number of gluons in the t-channel remains *fixed* in the course of the evolution. This is probably a good approximation in some intermediate kinematical region, but it fails to describe two interesting physical situations: First, it does not incorporate correctly the *fluctuations in the number of gluons*, as resulting from processes in which one (reggeized) gluon splits into two, or, more generally, a n–gluon

state evolves into a $(n + m)$–one, with $m \geq 1$. Such processes are especially important in the dilute regime at relatively large transverse momenta (for a given energy), where gluon splitting is the main process through which higher–point correlations get built [48]. Second, the approximation in which the number of t–channel gluons is fixed cannot describe *recombination processes* in which (reggeized) gluons merge with each other, thus reducing the gluon density. Such processes are important in the high–energy regime where the gluon density becomes large enough (due to BFKL evolution and to the splitting processes alluded to above) to enhance recombination processes, which are then expected to lead to *gluon saturation* [22, 23, 25]. The inclusion of saturation is also necessary, for consistency, in studies of the *unitarization* of the scattering amplitudes, except for some exceptional kinematical configurations [58].

The simplest approach including *gluon splitting* in the framework of BFKL evolution is the *color dipole picture* developed by Mueller [58, 138]. This picture is valid at large N_c, and describes pomeron multiplication via vertices at which one (BFKL) pomeron splits into two. A more ambitious program, which is not restricted to the large–N_c approximation, is the *extended generalized leading logarithmic approximation* (EGLLA), initiated by Bartels [126, 129], in which the gluon number changing vertices are explicitly computed in perturbative QCD (see Refs. [139–147] for further developments along this line and Ref. [137] for a review). By using such vertices, evolution equations allowing for gluon splitting have been written down in Refs. [140, 143, 145]. Also, the equivalence between the triple pomeron vertex in the dipole picture [138, 148] and the one generated by EGLLA at large N_c [139, 141] has been verified in Refs. [140, 146].

So far, the only formalism allowing for the systematic inclusion of *gluon merging* in the high–energy evolution is the *Color Glass Condensate* (CGC) [56], in which the reggeized gluons are replaced by *classical color fields* whose correlations get built in the course of the evolution. But the corresponding evolution is *non–linear* : the new gluons radiated at one step in the evolution (the analog of the 'rungs' in the BFKL ladders) are allowed to scatter off the classical color fields generated in the previous steps, and this is the mechanism leading to gluon merging. Because of the non–linear effects, the evolution couples n–point functions with various values of n, and can be most compactly summarized as a *functional Fokker–Planck equation* for the weight function describing the correlations: the Jalilian-Marian–Iancu–McLerran–Weigert–Leonidov–Kovner (JIMWLK) equation [31, 34, 35]. Alternatively, and equivalently [40], the evolution can be formulated as an hierarchy of equations for scattering amplitudes — the Balitsky equations [39] —, in which unitarity is manifest. Note however that gluon *splittings* are not included in the JIMWLK equation [48]; this is obvious from the fact that, in the dilute, or *weak–field*, limit, this equation reduces to an evolution in which the number of gluons in the t–channel stays *constant* [34, 38]. An extension of the JIMWLK–Balitsky evolution which includes pomeron splitting has been proposed only very recently [48–50].

2 The Color Glass Condensate effective theory

With this section, we begin the study of the effective theory for the Color Glass Condensate (CGC), that we shall then pursue up to the end of this work. Following the common practice in the literature, the concept of "color glass condensate" [34] will be used to denote both a *physical reality* — the high–density gluonic matter which is produced at small x and low enough $k_\perp$ through the mechanism of gluon saturation — and the *mathematical formalism* which is used to describe this reality, and which is derived from QCD within a controlled scheme of approximations. As we shall see, the CGC formalism encompasses most of the theoretical approaches previously proposed within perturbative QCD at high energy — chiefly among them, the BFKL equation and the dipole picture — but it goes further than such previous approaches by including the non–linear effects responsible for gluon saturation and the unitarization of the scattering amplitudes.

The essential physical argument underlying the CGC picture is a kinematical one, and has been already discussed at length in the introductory section. This is the fundamental separation in (longitudinal and temporal) scales between partons with different values of x in the hadron wavefunction in the infinite momentum frame: The large–x partons are Lorentz contracted and time dilated, and coherently act as color sources for the emission of small–x gluons which then participate in the scattering with an external projectile. This separation of scales creates the premises for a *classical description*, which for a given value of x is similar to the MV model (cf. Sect. 1.3.3), but which *evolves* with decreasing x, because of the BFKL–like emission of new quantum gluons which act as additional 'color sources' at the lower value of x. Thus, the classical description is an *effective theory*, whose evolution with increasing energy can be computed within perturbative QCD. The structure and the construction of this effective theory will be described in detail in this and the nextcoming sections.

2.1 The effective theory

Since the effective theory depends upon the value of x at which we probe the hadron wavefunction, it is convenient to introduce the longitudinal momentum scale $\Lambda^+ = xP^+$ (with P^+ the total longitudinal momentum of the hadron) and distinguish between 'fast' ($k^+ > \Lambda^+$) and 'soft' ($k^+ \leq \Lambda^+$) degrees of freedom. The effective theory will be a theory for gluon correlations at the soft scale Λ^+ as obtained after having 'integrated' out the fast modes with $k^+ > \Lambda^+$, within approximations to be shortly specified.

From the discussion of the BFKL evolution in Sect. 1.4 we know that the separation of scales is *logarithmic*: the contribution of the quantum modes with momenta k^+ in the range $\Lambda^+ < k^+ < p^+$ to the correlations that we measure at the scale Λ^+ is of order $\alpha \ln(p^+/\Lambda^+)$. Therefore, in the *leading logarithmic approximation* in which we preserve only the radiative corrections enhanced by the large logarithm $\ln 1/x$ we can neglect the quantum gluons with momenta just above Λ^+, but focus on the truly fast modes with $p^+ \gg \Lambda^+$. By the same argument, we can ignore the genuine quantum fluctuations with momenta $k^+ \sim \Lambda^+$ as compared to the corresponding gluons radiated by the fast partons

(whose density and correlations are enhanced by $\ln 1/x$).

This strong separation in longitudinal momenta between the soft gluons and their sources entails a similar separation between the associated time scales, which in turn implies that the small–x and large–x dynamics decouple from each other and can be treated separately: Over the natural time scale $\Delta x^+ \sim 2\Lambda^+/k_\perp^2$ for the dynamics at the scale Λ^+, the dynamics of the fast color sources is essentially frozen. Therefore, any observable pertinent to a small–x process can be first computed for a *fixed configuration* of the color sources, and then averaged over all the possible configurations with some *classical* probability distribution, which reflects the dynamics of the fast partons. The classical approximation is appropriate because the high energy scattering probes only an instantaneous configuration of the color sources, and thus it is not sensitive to the quantum interference between different configurations. But the color sources are themselves produced through quantum evolution, thus the computation of the corresponding probability distribution requires in general a quantum calculation.

This decoupling between a slow and a rapid dynamics lies at the basis of the *color glass* picture to be explained towards the end of this subsection. This is also reminiscent of the Born–Oppenheimer approximation in the study of a molecule, in which the dynamics of the fast moving electrons is treated separately from that of the slow nuclei. This is furthermore similar to the parton picture and the associated factorization of a high–energy process into a partonic cross–section times a classical parton distribution (itself obeying quantum evolution). But in order to describe non–linear phenomena like multiple scattering or gluon saturation, we need a formalism which is more general than the QCD parton picture. The CGC effective theory to be described here is such a formalism.

Specifically, the effective theory at scale Λ^+ must involve the two ingredients alluded to above: the classical calculation of the small–x observables for a given distribution of the color sources, and the quantum calculation of the probability distribution for the latter.

i) The structure of the *classical theory* is fixed by the kinematics, and thus is the same as in the MV model [25]: The fast color sources are represented by a color current $J_a^\mu = \delta^{\mu+}\rho_a$, where ρ_a is static (i.e., independent of the LC time x^+) and random. The small–x gluons are the color fields generated by this current according to the Yang–Mills equation:

$$(D_\nu F^{\nu\mu})_a(x) = \delta^{\mu+}\rho_a(\vec{x}),\tag{2.1}$$

where $\vec{x} = (x^-, \boldsymbol{x})$ and $D_\nu = \partial_\nu - igA_\nu^a T^a$ with $(T^a)_{bc} = -if_{abc}$.

All the interesting correlations are included in the functional *weight function* $W_{\Lambda^+}[\rho]$ which describes the distribution of ρ. This is assumed to be positive semi–definite ($W_{\Lambda^+}[\rho] \geq 0$), gauge invariant, and normalized to unity:

$$\int \mathcal{D}[\rho]\, W_{\Lambda^+}[\rho] = 1,\qquad \mathcal{D}[\rho] \equiv \prod_a \prod_{x^-} \prod_{\boldsymbol{x}} \mathrm{d}\rho_a(x^-, \boldsymbol{x}).\tag{2.2}$$

Still as in the MV model, the observables in the effective theory are first evaluated on the solution $\mathcal{A}_a^\mu \equiv \mathcal{A}_a^\mu[\rho]$ to the Yang-Mills equations, and then averaged over ρ :

$$\langle \mathcal{O}[A^\mu]\rangle_{\Lambda^+} = \int \mathcal{D}[\rho]\, W_{\Lambda^+}[\rho]\, \mathcal{O}[\mathcal{A}^\mu].\tag{2.3}$$

This averaging restores the gauge symmetry which was *a priori* broken by the presence of a non–trivial color charge density, and also by the choice of a gauge in the classical solution. The solution $\mathcal{A}_a^\mu$ to Eq. (2.1) in various gauges will be constructed in Sect. 2.3.

Note that all the correlations computed within the effective theory are independent of time (x^+). This is manifest in the classical theory — since for a static source ρ, the solution $\mathcal{A}_a^\mu[\rho]$ to the classical equations (2.1) is static as well —, and reflects the fact that the small–x gluons inherit the time scales of their sources, which are time dilated. Also, the small–x gluons have, by definition, low longitudinal momenta (the lowest ones among all the modes included in the effective theory), therefore they 'see' an integrated version of the hadron in longitudinal direction. This is visible, e.g., in the expression (1.37) for the dipole scattering amplitudes: this involves integrations over x^- which are however *path–ordered* (see Eq. (1.35)), to account for the non–commutativity of the color matrices at different points x^-. Thus, in order to compute such correlations, it would be not enough to dispose of a simplified information concerning just the two–dimensional distribution of the color sources in the transverse plane (such a simplified description would involve integrated quantities like $\rho_a(\boldsymbol{x}) \equiv \int \mathrm{d}x^- \rho_a(x^-, \boldsymbol{x})$). Rather, one also needs to control their distribution in the longitudinal direction.

ii) The *quantum calculation* refers to the evolution of the weight function with decreasing Λ^+. As explained before, $W_{\Lambda^+}[\rho]$ includes the effects of all the color sources with momenta $k^+ \gg \Lambda^+ = xP^+$. With decreasing x, new sources are generated through quantum evolution, and for $x \ll 1$ these sources are predominantly gluons. (The valence quarks, which were the only color sources in the MV model, act now merely as the initiators of the gluon cascades.) Thus, the correlations of ρ are built by the evolution, and for sufficiently small x they should be largely independent of the initial conditions at $x_0 \sim 1$ (in the same way as the BFKL solution (1.57) at high energy is essentially independent of the initial condition: the latter only specifies the momentum scale Q_0). Since the evolution can be computed within perturbative QCD, as we shall shortly see, the expected *universality* of the high energy behaviour opens the way towards a calculation of the gluon distribution at small x from first principles.

For the purpose of computing the quantum evolution one needs to reintroduce in the problem the quantum gluons with momenta $k^+ < \Lambda^+$. The complete theory constructed after the evolution down to Λ^+ can be summarized into the following generating functional for gluon correlations at momenta $k^+ \leq \Lambda^+$ [31, 34] :

$$Z[j] \;=\; \int \mathcal{D}[\rho]\, W_{\Lambda^+}[\rho] \left\{ \frac{\int^{\Lambda^+} \mathcal{D}[\delta A]\, \delta(A^+)\, \mathrm{e}^{iS[A,\rho] - \int j \cdot A}}{\int^{\Lambda^+} \mathcal{D}[\delta A]\, \delta(A^+)\, \mathrm{e}^{iS[A,\rho]}} \right\}. \tag{2.4}$$

Here, the external current j_a^μ is not a dynamical current, but merely a formal device to generate Green's functions via differentiations. For instance, the 2–point function is obtained by taking two functional derivatives in Eq. (2.4) w.r.t. j and then letting $j = 0$:

$$\langle \mathrm{T} A^\mu(x) A^\nu(y) \rangle_{k^+} = \int \mathcal{D}[\rho]\, W_{\Lambda^+}[\rho] \left\{ \frac{\int^{\Lambda^+} \mathcal{D}[\delta A]\, \delta(A^+)\, A^\mu(x) A^\nu(y)\, \mathrm{e}^{iS[A,\rho]}}{\int^{\Lambda^+} \mathcal{D}[\delta A]\, \delta(A^+)\, \mathrm{e}^{iS[A,\rho]}} \right\}. \tag{2.5}$$

(The T–symbol denotes ordering in x^+.) In the above equations $S[A, \rho]$ is the action that describes the dynamics of the soft gluons in the presence of the classical color charge ρ. In the stationary phase approximation $\delta S/\delta A^\mu = 0$, this action must reproduce the classical field equations (2.1). The action will be presented in Sect. 3.

The averaging in Eqs. (2.4) and (2.5) involves two types of functional integrals: (a) a *classical* integral over the color charge density ρ (this represents the fast partons with $k^+ \gg \Lambda^+$ which have been already integrated out in the previous steps), and (b) a *quantum* path–integral over the fields δA^μ which represent gluon fluctuations with momenta $k^+ \leq \Lambda^+$ (this constraint is indicated by the upper cutoff Λ^+ in the integrals over δA^μ).

Note that the path integral over the gluon fields is written in the light–cone (LC) gauge $A_a^+ = 0$. This choice is motivated by the fact that this gauge allows for the most direct partonic interpretation [30], and is also correlated with our separation of scales in longitudinal momentum: As well known, a momentum cutoff violates gauge symmetry, and this is in particular true for our cutoff Λ^+. However, a mode separation in k^+ is invariant under the gauge transformations which are independent of x^-. But these are precisely the gauge transformations which preserve the LC–gauge condition $A_a^+ = 0$ (one says that they are the *residual gauge transformations* in the LC gauge). Thus, in this restricted sense, the separation of scales Eq. (2.4) has a gauge invariant meaning within the LC gauge.

The total field A^μ at momenta $k^+ \leq \Lambda^+$ is the sum $A^\mu = \mathcal{A}^\mu[\rho] + \delta A^\mu$ between the classical field $\mathcal{A}^\mu[\rho]$ radiated by ρ and the quantum fluctuations. To leading–log accuracy, the correlations $\langle T A^\mu A^\nu \cdots \rangle_{k^+} \equiv \langle T (\mathcal{A}^\mu + \delta A^\mu)(\mathcal{A}^\nu + \delta A^\nu) \cdots \rangle_{k^+}$ at the scale k^+ involve both the classical correlations of ρ, as encoded in the weight function $W_{\Lambda^+}[\rho]$, and the quantum correlations induced by the 'semi–fast' fluctuations with intermediate momenta $k^+ \leq p^+ \leq \Lambda^+$. On the other hand, the softer fluctuations with $p^+ < k^+$ are irrelevant for this purpose. Thus, for a given external scale $k^+ \leq \Lambda^+$, the quantum integral in equations like (2.5) can be restricted to the 'semi–fast' modes within the strip $k^+ \leq p^+ \leq \Lambda^+$.

In particular, when $k^+ \sim \Lambda^+$, the effects of the quantum fluctuations are negligible, and the path–integral in Eq. (2.5) can be evaluated in the saddle point approximation:

$$\frac{\delta S}{\delta A^\mu} = 0 \qquad \Longrightarrow \qquad A^\mu = \mathcal{A}^\mu[\rho], \tag{2.6}$$

where the r.h.s. follows since $\mathcal{A}^\mu$ is by definition the solution to the classical field equations. Then Eq. (2.5) reduces to

$$\langle A_a^\mu(x) A_b^\nu(y) \rangle_{\Lambda^+} = \int \mathcal{D}[\rho] \, W_{\Lambda^+}[\rho] \, \mathcal{A}_a^\mu(\vec{x}) \mathcal{A}_b^\nu(\vec{y}), \tag{2.7}$$

in agreement with Eq. (2.3). As compared to Eq. (2.5), we have now omitted the T–symbol, since the time–ordering becomes irrelevant for the time–independent classical correlations.

The evolution of $W_{\Lambda^+}[\rho]$ with decreasing Λ^+ is computed through a *renormalization group analysis* in which 'semi–fast' ($k^+ \leq p^+ \leq \Lambda^+$) quantum gluons are integrated out in layers of p^+ and the correlations thus induced at the scale $k^+ \ll \Lambda^+$ are absorbed into a redefinition of the weight function for the classical theory: $W_{\Lambda^+}[\rho] \rightarrow W_{k^+}[\rho]$. This procedure, to be detailed in Sect. 3, leads to a functional evolution equation for $W_{\Lambda^+}[\rho]$ [31,
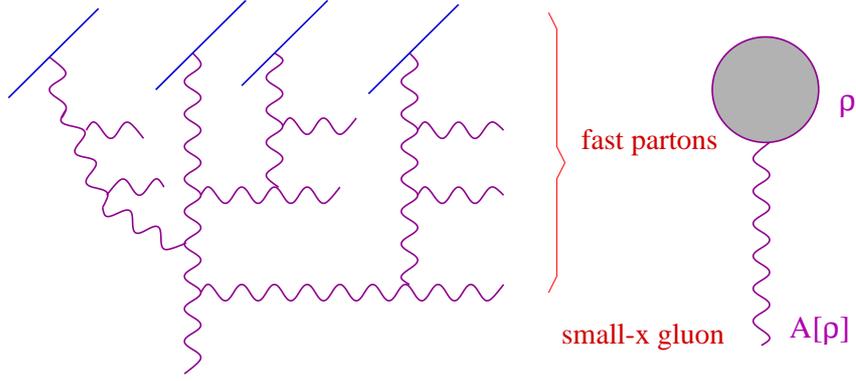
Fig. 23. *A typical gluon cascade which contributes to a classical color field configuration in the color glass condensate* CASC

34]. Once supplemented with some initial condition at low energy, so like the MV model, this equation determines the effective theory up to arbitrarily high energies. The general structure of the effective color source ρ which emerges from this analysis is illustrated in Fig. 23, in which we recognize both the BFKL ladders and the gluon mergings leading to saturation. As suggested by this picture, the quantum evolution of the CGC goes beyond the BFKL evolution by including the non–linear effects associated with the strong classical color fields. This will be discussed at length in Sect. 3.

2.2 Why "Color Glass Condensate" ?

SECT_WCGC

Let us now explain why the effective theory synthesized by Eq. (2.4) can be naturally interpreted as describing a *color glass*, and under which circumstances this becomes also a *condensate* [34, 56].

Note the special structure of the averaging in Eq. (2.5). This is *not* the same as[6] :

$$\frac{\int \mathcal{D}[\rho]\, W_\Lambda[\rho]\ \int^\Lambda \mathcal{D}[A]\, A^\mu(x) A^\nu(y)\ \mathrm{e}^{iS[A,\rho]}}{\int \mathcal{D}[\rho]\, W_\Lambda[\rho]\ \int^\Lambda \mathcal{D}[A]\ \mathrm{e}^{iS[A,\rho]}}\ . \tag{2.8}$$

In Eq. (2.8), both the colour source ρ_a and the gauge fields A_a^μ are dynamical variables that are summed over on the same footing. They are free to take on values which extremize the total "effective action" :

$$S_{eff}[A,\,\rho] \;=\; S[A,\,\rho] \;-\; i\ln W_\Lambda[\rho]. \tag{2.9}$$

By contrast, in Eq. (2.5) the average over A^μ is taken at fixed ρ : the gauge fields can vary in response to ρ, but ρ cannot vary in response to the gauge fields. That is, ρ is not a dynamical variable, but rather an "external" source. Giving a color charge distribution $\rho_a(\vec{x})$ specifies a medium in which propagate the quantum gluons. But this medium is, by itself, random, so after performing the quantum analysis at fixed ρ, one must also perform

---

[6] As compared to Eq. (2.5), we now use some slightly simplified notations and omit, e.g., the gauge fixing condition.

an average over ρ. The reason for treating ρ and A^μ so differently lies in the separation of scales in the problem: the changes in ρ happens on time scales much larger than the natural times for the dynamics of the fields A^μ. This situation is typical for amorphous materials called "glasses".

The prototype of a glass is the "spin glass" – a collection of magnetic impurities randomly distributed on a non–magnetic lattice. The dynamical degrees of freedom, which are rapidly varying, are the magnetic moments of the impurities (the "spins"), while the slowly varying disorder refers to the positions of these spins in the host lattice. To study the thermodynamics of such a system, one first computes the free–energy (= the logarithm of the partition function) of the spin system for a *fixed* disorder (namely, for a given spatial configuration of the impurities), and subsequently makes an average over all such configurations, with some weight function. The final average over the configurations is not a thermal one: what is averaged is the free–energy computed separately for each configuration.

Similarly, the *connected* correlation functions of the small–x gluons in the present effective theory are obtained from the following generating functional:

$$F[j] \,=\, \int \mathcal{D}[\rho] \; W_\Lambda[\rho] \; \ln \left(\int^\Lambda \mathcal{D}[A] \; \mathrm{e}^{iS[A,\rho] - i\int j \cdot A} \right) , \qquad (2.10)$$

where the logarithm is taken *inside* the integral over ρ. That is, the 'free–energy' reaches its extremum as a function of the external source j for a fixed distribution of the color sources. The *measured* free–energy (or correlation function) is finally obtained by also averaging over ρ.

We are thus naturally led to interpret the small–x component of the hadron wavefunction as a *glass*, with the color charge density playing the role of the magnetization for spin glasses. Thus, this is a *color* glass. Unlike what happens for spin glasses, which may have a non–zero value for the average magnetization (at least locally, i.e., at a given site), the *average* color charge density must be zero,

$$\langle \rho_a(\vec{x}) \rangle \,=\, 0 \quad \text{at any } \vec{x}, \qquad (2.11)$$

to comply with gauge symmetry. In practice, this is insured by the fact that we sum over all the possible configurations of $\rho_a(\vec{x})$ with a gauge–invariant weight function.

From the previous arguments, we also see that the 'color glass' interpretation is appropriate within the whole applicability range of the effective theory; this includes transverse momenta both below and above the saturation scale, up to the transition region towards the DGLAP dynamics at very high Q^2. By contrast, the *condensate* interpretation that we shall discuss now is appropriate only in the saturation region at $k_\perp \lesssim Q_s(x)$. We have already explained, in the qualitative discussion at the end of Sect. 1.2.1, that at saturation the gluon occupation factor φ becomes of order $1/\alpha$, which is the largest value allowed by the repulsive interactions of the high density gluons. Furthermore, in Sect. 1.4 we have seen that the evolution in the dilute regime is *explosive*: the occupation factor increases exponentially with the 'evolution time' τ. We recognize here the characteristic ingredients of the mechanism for Bose condensation: an unstable growth in the low density regime,

which is eventually stabilized by repulsive interactions when the density becomes large enough. We conclude that, in the saturation regime, the gluons form a *Bose condensate*.

2.3 The classical color field

In this subsection, we shall construct the solution to the classical field equations (2.1). Since the color current is static and has just a "+" component, it is always possible to construct a solution with the following properties [56]:

$$F_a^{ij} = 0, \qquad A_a^- = 0, \qquad A_a^+,\ A_a^i\ :\ \text{independent of } x^+\,. \tag{2.12}$$

Once such a solution is found in a given gauge, then its structure (2.12) is preserved by any time–independent gauge transformation (where by 'time' we mean x^+). Since $F^{ij} = 0$, the transverse fields A^i form a two-dimensional pure gauge; that is, there exists a group element $U(x^-, \boldsymbol{x}) \in \mathrm{SU}(N)$ such that:

$$A^i(x^-, \boldsymbol{x}) \,=\, \frac{i}{g}\, U(x^-, \boldsymbol{x})\, \partial^i U^\dagger(x^-, \boldsymbol{x})\,. \tag{2.13}$$

(in matrix notations appropriate for the adjoint representation: $A^i = A_a^i T^a$, etc). Thus, the requirements (2.12) leave just two independent field degrees of freedom, $A^+(\vec{x})$ and $U(\vec{x})$, which are further reduced to one (either A^+ or U) by imposing a gauge-fixing condition. Note also that the only non–trivial field strength for a gauge field obeying Eq. (2.12) is the *electric field* F_a^{+i}.

We consider first the *covariant gauge* (COV–gauge) $\partial_\mu A^\mu = 0$, where the solution takes its simplest form. Together with Eqs. (2.12) and (2.13), the gauge condition implies $\partial_i A^i = 0$, or $U = 0$. Thus, in this gauge, $\tilde{\mathcal{A}}_a^\mu(x) = \delta^{\mu+}\alpha_a(x^-, \boldsymbol{x})$, with $\alpha_a(\vec{x})$ linearly related to the color source $\tilde{\rho}_a$ in the COV–gauge :

$$-\nabla_\perp^2 \alpha_a(\vec{x}) \,=\, \tilde{\rho}_a(\vec{x})\,. \tag{2.14}$$

The electric field in this gauge is obtained as $\tilde{\mathcal{F}}_a^{+i} = -\partial^i \alpha_a$. (Note that we use a tilde to denote the classical source and fields in the COV–gauge.) Eq. (2.14) has the solution :

$$\alpha_a(x^-, \boldsymbol{x}) = \int d^2\boldsymbol{y}\, \Delta(\boldsymbol{x} - \boldsymbol{y})\, \tilde{\rho}_a(x^-, \boldsymbol{y})\,, \tag{2.15}$$

where $\Delta(\boldsymbol{x} - \boldsymbol{y})$ denotes the Coulomb propagator in two dimensions:

$$\Delta(\boldsymbol{x} - \boldsymbol{y}) \equiv \langle \boldsymbol{x}| \frac{1}{-\nabla_\perp^2} |\boldsymbol{y}\rangle = \frac{1}{4\pi}\, \ln \frac{1}{(\boldsymbol{x} - \boldsymbol{y})^2 \mu^2}\,. \tag{2.16}$$

Here, μ is an infrared cutoff which is necessary in order to invert the Laplacian operator in two dimensions, but which will generally disappear from the final, physical, results.

But the quantum effective theory in Eq. (2.4) is written in the LC–gauge $A^+ = 0$, so we also need the classical solution in this gauge. This is of the form $\mathcal{A}_a^\mu = \delta^{\mu i} \mathcal{A}_a^i$ with $\mathcal{A}_a^i(x^-, \boldsymbol{x})$ a "pure gauge", cf. Eq. (2.13). The group–valued function $U(\vec{x})$ is most simply
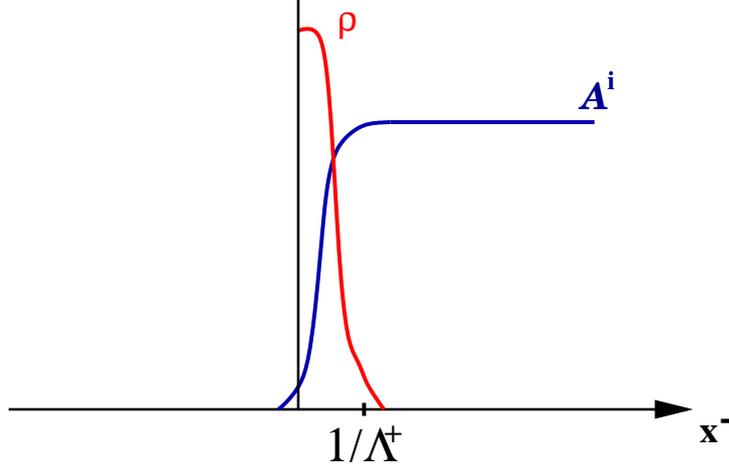
Fig. 24. The longitudinal structure of the color source ρ and of the classical field solution $\mathcal{A}^i$ for the effective theory at the scale Λ^+. As functions of x^-, α and $\mathcal{F}^{+i}$ are as localized as ρ.

related to the color field α_a in the COV–gauge. Indeed, $U(\vec{x})$ is precisely the gauge rotation which transforms the solution from the COV–gauge to the LC–one:

$$\mathcal{A}^\mu = U\left(\tilde{\mathcal{A}}^\mu + \frac{i}{g}\partial^\mu\right)U^\dagger. \tag{2.17}$$

With $\tilde{\mathcal{A}}^i = 0$, this transformation ensures that $\mathcal{A}^i$ has the expected form, cf. Eq. (2.13). Besides, Eq. (2.17) must lead to $\mathcal{A}^+ = 0$, which implies

$$U^\dagger(x^-, \boldsymbol{x}) = \mathrm{P}\exp\left\{ig\int_{-\infty}^{x^-}dz^-\,\alpha_a(z^-, \boldsymbol{x})T^a\right\}. \tag{2.18}$$

The lower limit $x_0^- \to -\infty$ in the integral over x^- in Eq. (2.18) has been chosen such as to impose the "retarded" boundary condition:

$$\mathcal{A}_a^i(x) \to 0 \quad \text{as} \quad x^- \to -\infty, \tag{2.19}$$

which will be useful in what follows. (The "retardation" property refers here to x^-, and not to time.) Note that the choice of a value for x_0^- amounts to a complete gauge–fixing (it eliminates the residual, $\boldsymbol{x}$–dependent, gauge transformations which would preserve both the LC–gauge condition and the general properties (2.12) of the classical field).

Together, Eqs. (2.13), (2.15) and (2.18) provide an explicit expression for the LC–gauge solution $\mathcal{A}^i$ in terms of the color source $\tilde{\rho}$ in the COV–gauge. This is sufficient for the purpose of computing correlation functions since the average in Eq. (2.7) can be always re–expressed as a functional integral over the covariant gauge color source $\tilde{\rho}$ by a change of variables. For instance (with $\mathcal{A}_x^i \equiv \mathcal{A}^i(\vec{x})$) :

$$\langle A^i(x^+, \vec{x})A^j(x^+, \vec{y})\cdots\rangle_{\Lambda^+}\Big|_{\mathrm{LC}} = \int \mathcal{D}\tilde{\rho}\, W_{\Lambda^+}[\tilde{\rho}]\, \mathcal{A}_x^i[\tilde{\rho}]\, \mathcal{A}_y^j[\tilde{\rho}]\cdots. \tag{2.20}$$

Besides physical observables are gauge–invariant, so they can be computed in any gauge.

60

So far, the longitudinal structure of the source has been arbitrary: the solutions written above hold for any function $\rho^a(x^-)$. For what follows, however, it is useful to recall that, by Lorentz contraction, the charge density ρ of the fast partons is localized near $x^- = 0$. More precisely, the quantum analysis in Sect. 3.1 will demonstrate that the classical source in the effective theory at the scale Λ^+ has support at x^- within the range $0 \lesssim x^- \lesssim 1/\Lambda^+$. From Eq. (2.15), it is clear that this is also the longitudinal support of the "Coulomb"– field α. Thus, integrals over x^- as that in Eq. (2.18) receive their whole contribution from x^- in this limited range. An external probe with momentum $k^+ \ll \Lambda^+$ (and thus with a comparatively low longitudinal resolution) will not be able to discriminate the internal structure of the source. Rather, it will see a source/field structure which is quasi–singular at $x^- = 0$ (see Fig. 24) : $\rho_a(x^-, \boldsymbol{x}) \approx \delta(x^-)\rho_a(\boldsymbol{x})$ and

$$\mathcal{A}^i(x^-, \boldsymbol{x}) \approx \theta(x^-)\frac{i}{g}V(\partial^i V^\dagger)(\boldsymbol{x}), \tag{2.21}$$

where V and $V^\dagger$ are obtained by letting $x^- \to \infty$ in Eq. (2.18) :

$$V^\dagger(\boldsymbol{x}) \equiv \mathrm{P}\exp\left\{ig\int_{-\infty}^{\infty}dz^-\,\alpha(z^-, \boldsymbol{x})\right\}. \tag{2.22}$$

2.4   The gluon distribution

SECT_XGX

The *gluon distribution* $xG(x, Q^2)$ is defined as the number of gluons per unit rapidity which are localized in transverse space to a size $\Delta\boldsymbol{x} \sim 1/Q$ (or, equivalently, which have transverse momenta $k_\perp \lesssim Q$) [56, 70]:

$$\begin{aligned} xG(x, Q^2) &= \int^{Q^2} \mathrm{d}^2 k_\perp\, k^+ \frac{\mathrm{d}N}{\mathrm{d}k^+\mathrm{d}^2\boldsymbol{k}}\bigg|_{k^+=xP^+} \\ &= \int \mathrm{d}^3 k\, \Theta(Q^2 - k_\perp^2)\, x\delta(x - k^+/P^+)\frac{\mathrm{d}N}{\mathrm{d}^3 k}, \end{aligned} \tag{2.23}$$

where $\Theta(x)$ is the step function and

$$\frac{\mathrm{d}N}{\mathrm{d}^3 k} = \sum_{\lambda=\pm}\langle a_\lambda^{c\dagger}(x^+, \vec{k})\, a_\lambda^c(x^+, \vec{k})\rangle = \frac{2k^+}{(2\pi)^3}\langle A_c^i(x^+, \vec{k})A_c^i(x^+, -\vec{k})\rangle, \tag{2.24}$$

with $\vec{k} \equiv (k^+, \mathbf{k}_\perp)$ is the Fock–space gluon density, i.e., the number of gluons per unit of volume in momentum space. We have used here the following relation between the gauge fields A_c^i and the Fock–space creation and annihilation operators (for gluons with momentum $\vec{k}$, color index c and polarization state λ), valid within the LC–gauge quantization of the Yang–Mills theory [149]:

$$A_c^\mu(x^+, \vec{x}) = \sum_{\lambda=\pm}\int_{k^+>0}\frac{\mathrm{d}^3 k}{\sqrt{(2\pi)^3 2k^+}}\left[a_\lambda^c(x^+, \vec{k})\epsilon_\lambda^\mu(\vec{k})\mathrm{e}^{i\vec{k}\cdot\vec{x}} + a_\lambda^{c\dagger}(x^+, \vec{k})\epsilon_\lambda^{\mu\dagger}(\vec{k})\mathrm{e}^{-i\vec{k}\cdot\vec{x}}\right], \tag{2.25}$$

with $\vec{x}\cdot\vec{k} = x^- k^+ - \boldsymbol{x}\cdot\boldsymbol{k}$ and the LC–gauge polarization vector

61

$$\epsilon_\lambda^\mu(\vec{k}) = (0, \epsilon_\lambda^-, \boldsymbol{\epsilon}_\lambda) \qquad \text{with} \qquad \epsilon_\lambda^- = \boldsymbol{\epsilon}_\lambda \cdot \boldsymbol{k}/k^+, \tag{2.26}$$

which satisfies $k_\mu \epsilon_\lambda^\mu = 0$, as it should. The Fock–space creation and annihilation operators obey the following commutation relation at equal LC time x^+ :

$$[a_\lambda^b(x^+, \vec{k}), a_{\lambda'}^{c\dagger}(x^+, \vec{q})] = \delta_{\lambda\lambda'}\delta^{bc}\delta^{(3)}(\vec{k} - \vec{q}). \tag{2.27}$$

A priori, the physical interpretation of the Fock–space 'gluon number' is hindered by the fact that, in general, this is not a gauge–invariant quantity. However, this is not true for its LC–gauge definition above, which within the CGC effective theory can be given a gauge-invariant meaning, as we explain now: Note first that, in this gauge, the electric field F_a^{i+} and the vector potential A_a^i are linearly related, $F_a^{i+}(k) = ik^+ A_a^i(k)$, so that Eq. (2.23) can be rewritten as a two–point Green's function of the (gauge–covariant) electric fields. After performing the integral over k^+, one obtains (with $k^+ = xP^+$ from now on) :

$$xG(x, Q^2) = \frac{1}{\pi} \int \frac{\mathrm{d}^2 k_\perp}{(2\pi)^2} \Theta(Q^2 - k_\perp^2) \left\langle F_a^{i+}(x^+, \vec{k}) F_a^{i+}(x^+, -\vec{k}) \right\rangle. \tag{2.28}$$

This does not look gauge invariant as yet: in coordinate space,

$$F_a^{i+}(x^+, \vec{k}) F_a^{i+}(x^+, -\vec{k}) = \int \mathrm{d}^3 x \int \mathrm{d}^3 y \; e^{i(\vec{x}-\vec{y})\cdot\vec{k}} F_a^{i+}(x^+, \vec{x}) F_a^{i+}(x^+, \vec{y}) \tag{2.29}$$

involves the electric fields at different spatial points $\vec{x}$ and $\vec{y}$. A manifestly gauge invariant operator can be constructed by inserting Wilson lines:

$$\mathrm{Tr}\left\{ F^{i+}(\vec{x}) \, U_\gamma(\vec{x}, \vec{y}) \, F^{i+}(\vec{y}) \, U_\gamma(\vec{y}, \vec{x}) \right\}, \tag{2.30}$$

where (with $\vec{A} \equiv (A^+, \boldsymbol{A})$)

$$U_\gamma(\vec{x}, \vec{y}) = \mathrm{P} \exp\left\{ ig \int_\gamma d\vec{z} \cdot \vec{A}(\vec{z}) \right\}, \tag{2.31}$$

and the temporal coordinates x^+ are omitted (they are the same for all the fields). In Eq. (2.31), γ is an arbitrary oriented path going from $\vec{y}$ to $\vec{x}$. For any such a path, Eq. (2.30) defines a gauge–invariant operator.

We now show that, with appropriate choices for the path, the gauge, and the boundary conditions, the gauge–invariant operator (2.30) reduces to the simple 2–point function (2.29). Specifically, consider the path shown in Fig. 25, with the following three elements: two "horizontal" pieces going along the x^- axis from $(y^-, \boldsymbol{y})$ to $(-\infty, \boldsymbol{y})$ and, respectively, from $(-\infty, \boldsymbol{x})$ to $(x^-, \boldsymbol{x})$, and a "vertical" piece from $(-\infty, \boldsymbol{y})$ to $(-\infty, \boldsymbol{x})$. Along the horizontal pieces, $d\vec{z} \cdot \vec{A} = dz^- A^+$, so these pieces do not matter in the LC gauge. Along the vertical piece, $d\vec{z} \cdot \vec{A} = d\boldsymbol{z} \cdot \boldsymbol{A}(-\infty, \boldsymbol{z})$, and the path γ between $\boldsymbol{y}$ and $\boldsymbol{x}$ is still arbitrary. But the contribution of any such a path vanishes for the classical solution constructed in Sect. 2.3, which obeys the "retarded" boundary condition (2.19). Thus, when evaluated in the effective theory and in the LC gauge, the operator (2.29) has a gauge–invariant meaning, as anticipated. Therefore, the definition (2.28) for the gluon distribution is physically meaningful indeed.
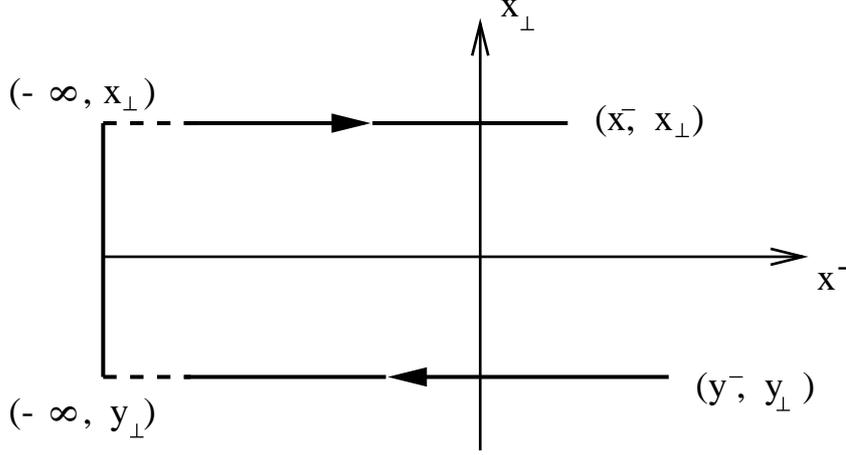
Fig. 25. The path γ used for the evaluation of the gauge-invariant operator (2.30).

In particular, by comparing (2.29) with the expressions in Eqs. (1.52) and (1.53), one can deduce the operator definitions for the gluon spectrum and the more standard 'unintegrated gluon distribution' $f(\tau, k_\perp^2)$:

$$\frac{\mathrm{d}N}{\mathrm{d}\tau\mathrm{d}^2\boldsymbol{k}} = \frac{1}{4\pi^3} \langle F_a^{i+}(\vec{k})F_a^{i+}(-\vec{k})\rangle_\tau, \qquad f(\tau, k_\perp^2) = \frac{k_\perp^2}{4\pi^2} \langle F_a^{i+}(\vec{k})F_a^{i+}(-\vec{k})\rangle_\tau . \qquad (2.32)$$

In what follows, we shall need also the *gluon occupation number*, which as explained in Sect. 1.2.1 is the most accurate indicator of the degree of overlapping (and thus of the density effects) in the system. This is defined as the number of gluons of given spin and color per unit rapidity and per unit of transverse phase–space :

$$\varphi(\tau, \boldsymbol{k}) \equiv \frac{(2\pi)^3}{2(N_c^2 - 1)} \frac{\mathrm{d}N}{\mathrm{d}\tau\, \mathrm{d}^2\boldsymbol{k}\, \mathrm{d}^2\boldsymbol{b}} . \qquad (2.33)$$

Note that, since the longitudinal dynamics is fully quantum, the rapidity interval

$$\Delta\tau \equiv \frac{\Delta k^+}{k^+} \simeq \Delta k^+ \Delta x^-$$

(we have used the uncertainty principle to estimate the delocalization in x^-) is the closest analog of the longitudinal phase–space interval. In that sense, Eq. (2.33) defines indeed a *physical, three–dimensional, occupation factor*. In particular, for a hadron which is quasi–homogeneous in the transverse plane (like a large nucleus), we can evaluate Eq. (2.33) as

$$\varphi(\tau, \boldsymbol{k}) \simeq \frac{4\pi^3}{N_c^2 - 1} \frac{1}{\pi R^2} \frac{\mathrm{d}N}{\mathrm{d}\tau\mathrm{d}^2\boldsymbol{k}} = \frac{1}{\pi R^2} \frac{\langle F_a^{i+}(\vec{k})F_a^{i+}(-\vec{k})\rangle}{N_c^2 - 1} . \qquad (2.34)$$

For illustration, let us use these formulae to compute the gluon distribution of a nucleus in the MV model. In this subsection, we shall consider just the low density regime, corresponding to a relatively large transverse momentum $k_\perp$. (The general case will be discussed in the next subsection.) For $k_\perp \gg Q_s$, the classical field is weak and can be

63

computed in the linear approximation. By expanding the general solution (2.13) to linear order in ρ, or, equivalently, by directly solving the linearized version of Eq. (2.1), one easily obtains:

$$\mathcal{A}_a^i(k) \simeq -\frac{k^i}{k^+ + i\varepsilon}\frac{\rho_a(k^+, k_\perp)}{k_\perp^2}, \qquad \mathcal{F}_a^{+i}(k) \simeq i\frac{k^i}{k_\perp^2}\,\rho_a(\vec{k})\,, \tag{2.35}$$

which together with Eq. (1.46) implies:

$$\langle \mathcal{F}_a^{i+}(\vec{k})\mathcal{F}_a^{i+}(-\vec{k})\rangle_A \simeq \frac{1}{k_\perp^2}\,\langle\rho_a(\vec{k})\rho_a(-\vec{k})\rangle_A \,=\, \pi R_A^2(N_c^2 - 1)\frac{\mu_A}{k_\perp^2}\,. \tag{2.36}$$

By inserting this approximation in Eqs. (2.34) and (2.28), one obtains the following estimates for the gluon occupation factor

$$\varphi_A(k_\perp) \simeq \frac{\mu_A}{k_\perp^2}\,, \tag{2.37}$$

and for the corresponding distribution function (cf. Eq. (1.46)) :

$$xG_A(x, Q^2) \simeq \frac{(N_c^2 - 1)R_A^2}{4\pi}\,\mu_A \int_{\Lambda_{\mathrm{QCD}}^2}^{Q^2} \frac{\mathrm{d}k_\perp^2}{k_\perp^2} \,=\, \frac{\alpha A N_c C_F}{\pi}\,\ln\frac{Q^2}{\Lambda_{\mathrm{QCD}}^2}\,. \tag{2.38}$$

Eq. (2.37) is simply $A \times N_c$ times the corresponding contribution of a single quark, as previously computed in Eq. (1.56). Similarly, $xG_A(x, Q^2) \simeq AN_c\,xG_q(x, Q^2)$, where $xG_q(x, Q^2)$ refers to a single quark. Thus, in this linear approximation and in the MV model, the classical effective theory reproduces, as expected, the bremsstrahlung spectrum of the independent color sources (the valence quarks). The integral over $k_\perp$ in Eq. (2.38) has a logarithmic infrared divergence which has been cut off by hand at the scale Λ_{QCD} where we expect color neutrality through confinement [121]. We will argue later that, for sufficiently small valiues of x and as a result of the quantum evolution with $1/x$, the actual scale for the screening of the infrared physics is not Λ_{QCD}, but the (relatively hard) saturation scale $Q_s(x)$.

According to Eq. (2.37), the gluon occupation factor is proportional to $A^{1/3}$, and can become arbitrarily large when A increases. This is however an artifact of our previous approximations which have neglected the interactions among the radiated gluons, i.e., the non-linear effects in the classical field equations. The importance of the non–linear effects at large A will be discussed in the next subsection.

2.5 Gluon saturation in the MV model

SECT_MVSAT

In what follows, we shall compute the gluon distribution in the MV model by using the exact solution to the Yang–Mills equations, as obtained in Sect. 2.3. This will enable us to study the phenomenon of gluon saturation through non–linear effects in the classical field dynamics [29, 30].

From the previous subsection, we know that the gluon distribution is related to a field correlator in the LC–gauge (cf. Eq. (2.28)). On the other hand, the classical solution in the

LC–gauge has been explicitly computed in Sect. 2.3 in terms of the color charge density $\tilde{\rho}$ in the COV–gauge. It is therefore preferable to reexpress the LC–gauge operator (2.29) in terms of the color fields in the COV–gauge. This involves the gauge–rotation (2.17):

$$\mathcal{F}_a^{+i} = U_{ab}(-\partial^i \alpha^b) \quad \Longrightarrow \quad \langle \mathcal{F}_a^{+i}(\vec{x})\mathcal{F}_a^{+i}(\vec{y})\rangle_A = \left\langle \left(U_{ab}\partial^i \alpha^b\right)_{\vec{x}} \left(U_{ac}\partial^i \alpha^c\right)_{\vec{y}}\right\rangle_A. \tag{2.39}$$

Thus, our objective will be to evaluate the expectation value in the r.h.s. of the above equation knowing that, in the MV model, $\alpha_a(\vec{x})$ is a Gaussian random variable with zero average and 2–point correlation function

$$\langle \alpha_a(\vec{x})\alpha_b(\vec{y})\rangle_A = \delta_{ab}\delta(x^- - y^-)\,\gamma_A(x^-, \boldsymbol{x} - \boldsymbol{y}),$$
$$\gamma_A(x^-, k_\perp) \equiv \frac{1}{k_\perp^4}\,\lambda_A(x^-)\,. \tag{2.40}$$

We have used here $\tilde{\rho}_a(x^-, k_\perp) = k_\perp^2 \alpha_a(x^-, k_\perp)$ together with Eq. (1.46).

Because of the locality of the 2–point function (2.40) in x^-, the relevant correlator in Eq. (2.39) can be factorized as [29, 56]

$$\left\langle \left(U_{ab}\partial^i \alpha^b\right)_{\vec{x}} \left(U_{ac}\partial^i \alpha^c\right)_{\vec{y}}\right\rangle = \left\langle \partial^i \alpha^b(\vec{x})\partial^i \alpha^c(\vec{y})\right\rangle \left\langle U_{ab}(\vec{x})U_{ca}^\dagger(\vec{y})\right\rangle$$
$$= \delta(x^- - y^-) \left\langle \text{Tr}\, U(\vec{x})U^\dagger(\vec{y})\right\rangle \left(-\nabla_\perp^2 \gamma_A(x^-, \boldsymbol{x} - \boldsymbol{y})\right), \tag{2.41}$$

where in writing the second line we have also used $U_{ac}^\dagger = U_{ca}$ in the adjoint representation, together with the first equation (2.40). The Wilson line correlator in Eq. (2.41), that is,

$$S_A(x^-, \boldsymbol{x} - \boldsymbol{y}) \equiv \frac{1}{N_c^2 - 1} \left\langle \text{Tr}\, U(x^-, \boldsymbol{x})U^\dagger(x^-, \boldsymbol{y})\right\rangle_A, \tag{2.42}$$

will be shortly computed, with the following result [29]

$$S_A(x^-, r_\perp) = \exp\left\{-g^2 N_c \int_{-\infty}^{x^-} \mathrm{d}z^- \left[\gamma_A(z^-, 0_\perp) - \gamma_A(z^-, r_\perp)\right]\right\}, \tag{2.43}$$

where (cf. Eq. (2.40))

$$\gamma_A(x^-, 0_\perp) - \gamma_A(x^-, r_\perp) = \lambda_A(x^-)\int \frac{\mathrm{d}^2 \boldsymbol{k}}{(2\pi)^2}\frac{1}{k_\perp^4}\left[1 - \mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{r}}\right]. \tag{2.44}$$

Since correlation functions of the Wilson lines will play an important role in what follows, it is instructive to follow the calculation of (2.43) in some detail. To that aim, it is useful to discretize the longitudinal coordinate axis, by writing $x^- \equiv n\epsilon$ with ϵ the length of an infinitesimal step and n the number of steps. (Physically, ϵ should be much smaller than the longitudinal extent of the Lorentz contracted nucleus.) The Wilson line $U(x^- = n\epsilon, \boldsymbol{x}) \equiv U_n(\boldsymbol{x})$, cf. Eq. (2.18), can then be seen as the result of n successive, infinitesimal, gauge rotations, the last one being:

$$U_n(\boldsymbol{x}) = U_{n-1}(\boldsymbol{x})\,\mathrm{e}^{-ig\epsilon\alpha_n^a(\boldsymbol{x})T^a} \simeq U_{n-1}(\boldsymbol{x})\left\{1 - ig\epsilon\alpha_n - \frac{1}{2}(g\epsilon\alpha_n)^2\right\}, \tag{2.45}$$

where $\alpha_n \equiv \alpha_n^a(\boldsymbol{x})T^a$. Note that the path–ordering on x^- plays no role within a small interval of length ϵ since the field $\alpha(x^-)$ is assumed to be uniform within such an interval. The reason why we had to keep terms up to second order in the expansion of the exponential in Eq. (2.45) should become clear in a moment.

By using Eq. (2.45) for $U_n(\boldsymbol{x})$ together with the corresponding formula for $U_n^\dagger(\boldsymbol{y})$, one can evaluate the change in the correlator (2.42) corresponding to one additional step in x^-. To that aim, we also need the discretized version of the 2–point function (2.40) of the gauge fields. Using $\delta(x^- - y^-) \to \delta_{nm}/\epsilon$, one finds

$$\langle \alpha_n^a(\boldsymbol{x})\alpha_m^b(\boldsymbol{y}) \rangle \;=\; \frac{1}{\epsilon}\,\delta_{nm}\delta^{ab}\,\gamma_n(\boldsymbol{x} - \boldsymbol{y}), \tag{2.46}$$

which makes it clear that the typical size of the fluctuation α_n^a increases like $1/\sqrt{\epsilon}$ when $\epsilon \to 0$. This is a consequence of the locality of the 2–point function (2.40) in x^-, and is the reason for keeping terms up to second order in the expansion in Eq. (2.45): $(\epsilon\alpha_n)^2$ is truly an $\mathcal{O}(\epsilon)$ effect. After a straightforward calculation in which one keeps systematically terms up to $\mathcal{O}(\epsilon)$, one finds [7]:

$$\mathrm{Tr}\left\langle U_n(\boldsymbol{x})U_n^\dagger(\boldsymbol{y}) \right\rangle \simeq \mathrm{Tr}\left\langle 1 - \frac{(g\epsilon)^2}{2}\Big(\alpha_n^2(\boldsymbol{x}) + \alpha_n^2(\boldsymbol{y}) - 2\alpha_n(\boldsymbol{x})\alpha_n(\boldsymbol{y})\Big)\right\rangle\!\left\langle U_{n-1}(\boldsymbol{x})U_{n-1}^\dagger(\boldsymbol{y}) \right\rangle$$

$$= \left\{ 1 - \epsilon g^2 N_c\big[\gamma_n(0) - \gamma_n(\boldsymbol{x} - \boldsymbol{y})\big] \right\} \mathrm{Tr}\left\langle U_{n-1}(\boldsymbol{x})U_{n-1}^\dagger(\boldsymbol{y}) \right\rangle. \tag{2.47}$$

By iterating this procedure or, alternatively, by translating Eq. (2.47) into an equation for the evolution of $S_A(x^-, \boldsymbol{x} - \boldsymbol{y})$ with x^- and then solving this equation, one immediately obtains the result shown in Eq. (2.43).

Returning to Eq. (2.44), we note that the integral there has a logarithmic infrared divergence, and hence is dominated by relatively low momenta, such that $\boldsymbol{k} \cdot \boldsymbol{r} \ll 1$. In the present context, this divergence should be screened by confinement at the scale Λ_{QCD}. To leading *transverse* log accuracy, i.e., by keeping only terms enhanced by the large logarithm $\ln(1/r_\perp^2 \Lambda_{\mathrm{QCD}}^2)$, the precise value of the infrared cutoff is not important, and we can also expand the integrand as:

$$\int \frac{\mathrm{d}^2\boldsymbol{k}}{(2\pi)^2}\,\frac{1 - \mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{r}}}{k_\perp^4} \simeq \int^{1/r_\perp^2} \frac{\mathrm{d}^2\boldsymbol{k}}{(2\pi)^2}\,\frac{1}{k_\perp^4}\,\frac{(\boldsymbol{k}\cdot\boldsymbol{r})^2}{2} \simeq \frac{r_\perp^2}{16\pi}\,\ln\frac{1}{r_\perp^2\Lambda_{\mathrm{QCD}}^2}\,. \tag{2.48}$$

This gives, with $\mu_A(x^-) \equiv \int_{-\infty}^{x^-} \mathrm{d}z^- \lambda_A(z^-)$ (compare to Eq. (1.46)),

$$S_A(x^-, r_\perp) \simeq \exp\left\{ -\frac{\alpha N_c}{4}\,r_\perp^2\,\mu_A(x^-)\,\ln\frac{1}{r_\perp^2\Lambda_{\mathrm{QCD}}^2} \right\}, \tag{2.49}$$

which together with Eq. (2.41) can be used to evaluate the gluon occupation factor in Eq. (2.34). We first obtain

---

[7] Note that expectation values involving α_n and, respectively, U_{n-1} factorize from each other, because of the factor δ_{nm} in Eq. (2.46).

$$\varphi_A(\boldsymbol{k}) = \int \mathrm{d}^2\boldsymbol{r}\, \mathrm{e}^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \int \mathrm{d}x^-\, S_A(x^-, r_\perp) \left(-\nabla_\perp^2 \gamma_A(x^-, \boldsymbol{r}) \right), \tag{2.50}$$

where one of the longitudinal integrations (cf. Eq. (2.29)) has been trivially evaluated by using the δ–function in Eq. (2.41). The remaining one can be also explicitly computed, because (cf. Eq. (2.40)) :

$$-\nabla_\perp^2 \gamma_A(x^-, r_\perp) = \lambda_A(x^-) \int \frac{\mathrm{d}^2\boldsymbol{k}}{(2\pi)^2} \frac{\mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{r}}}{k_\perp^2} = \frac{\partial \mu_A(x^-)}{\partial x^-} \frac{1}{4\pi} \ln \frac{1}{r_\perp^2 \Lambda_{\mathrm{QCD}}^2} \tag{2.51}$$

is essentially the x^-–derivative of the exponent in Eq. (2.49). Therefore [29, 30]

$$\varphi_A(k_\perp) = \int \mathrm{d}^2\boldsymbol{r}\, \mathrm{e}^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \frac{1 - \exp\left\{ -\frac{1}{4} r_\perp^2 Q_A^2 \ln \frac{1}{r_\perp^2 \Lambda_{\mathrm{QCD}}^2} \right\}}{\pi \alpha N_c r_\perp^2}, \tag{2.52}$$

where (with $\mu_A \equiv \int dx^- \lambda_A(x^-)$, cf. Eq. (1.46))

$$Q_A^2 \equiv \alpha N_c \mu_A = \frac{2\alpha^2 A N_c}{R_A^2} \sim A^{1/3}. \tag{2.53}$$

The integration in Eq. (2.52) must be restricted to $r_\perp < 1/\Lambda_{\mathrm{QCD}}$, for consistency with the present perturbative approach, and also to avoid that the logarithm in the exponent changes sign. But so long as $k_\perp \gg \Lambda_{\mathrm{QCD}}$, the value of the integral is very little sensitive to the precise value of the upper cutoff [87].

To study the $k_\perp$–dependence of the gluon spectrum, one must still perform the Fourier transform in Eq. (2.52). Exact, analytic, results for this quantity can be found in Ref. [87]. Here, we shall only study the limiting behaviors at large and, respectively, small momenta, and emphasize the phenomenon of *saturation*. The *saturation momentum* $Q_s(A)$ is defined via the condition that, for $r_\perp = 2/Q_s(A)$, the exponent in Eq. (2.52) becomes of order one (since this is the scale which separates between the linear and the non–linear regimes; see below). This condition yields:

$$Q_s^2(A) \simeq Q_A^2 \ln \frac{Q_s^2(A)}{\Lambda_{\mathrm{QCD}}^2} \sim A^{1/3} \ln A^{1/3}. \tag{2.54}$$

Note that $Q_s(A)$ is parametrically larger than Q_A, Eq. (2.53), since, by assumption, A is so large that $Q_s(A) \gg \Lambda_{\mathrm{QCD}}$ (which in turn requires $A^{1/3} \gg 1$). The above expression for Q_s is in fact consistent with the original GLR estimate [22], Eq. (1.19). Indeed, after using Eq. (2.53) together with the expression (2.38) for the nuclear gluon distribution in the high–momentum regime, one can recognize the similarity between the r.h.s.'s of Eqs. (2.54) and (1.19), respectively. We shall distinguish between two physical regimes:

i) At high momenta $k_\perp \gg Q_s(A)$, the integral is dominated by small dipole sizes $r_\perp \ll 1/Q_s(A)$, and can be evaluated by expanding out the exponential. This is essentially an expansion in powers of $Q_s^2/k_\perp^2$, and is generally referred to as the *twist expansion*. We show here the first two terms in this expansion:

$$\varphi_A(k_\perp) \simeq \frac{1}{\alpha N_c} \frac{Q_A^2}{k_\perp^2} \left\{ 1 + \frac{Q_A^2}{k_\perp^2} \left[\ln \frac{k_\perp^2}{\Lambda_{\rm QCD}^2} + 2\gamma - 2 \right] \right\} \quad \text{for} \quad k_\perp \gg Q_s(A). \tag{2.55}$$

The first order term corresponds to the linear approximation in the classical field equations, and coincides with the bremsstrahlung spectrum, Eq. (2.37), as expected. The second term is positive showing that, at high–$k_\perp$, the 'leading twist' result is approached from the above.

ii) At small momenta, $k_\perp \ll Q_s(A)$ (with $k_\perp \gg \Lambda_{\rm QCD}$ though), the dominant contribution comes from large distances $r_\perp \gg 1/Q_s(A)$, where one can simply neglect the exponential in the numerator and recognize $1/r_\perp^2$ as the Fourier transform of $\ln k_\perp^2$:

$$\varphi_A(k_\perp) \approx \frac{1}{\alpha N_c} \ln \frac{Q_s^2(A)}{k_\perp^2} \qquad \text{for} \qquad k_\perp \ll Q_s(A). \tag{2.56}$$

Note that the saturation scale provides the upper cutoff for the logarithm in Eq. (2.56); this is so since the contribution of the short distances $r_\perp \ll 1/Q_s(A)$ to the Fourier transform is cut off by the exponential in Eq. (2.52). The behaviour in Eqs. (2.55) and (2.56) is qualitatively illustrated in Fig. 26.

Unlike the linear distribution (2.55), which grows rapidly with A (like $A^{1/3}$) and would rise very fast (like $1/k_\perp^2$) at low momenta, the distribution in Eq. (2.56), which takes into account the non–linear effects in the classical Yang–Mills equations, rises only logarithmically as a function of both A and $1/k_\perp^2$. This is *gluon saturation*. At saturation, the gluon occupation factor is parametrically of order $1/\alpha N_c$, which is the maximum value allowed by the repulsive interactions between the strong color fields. When increasing the atomic number A, new gluons are preponderantly radiated at large transverse momenta $k_\perp > Q_s(A)$, where this repulsion is less important.

Note also the *scaling behaviour* in Eq. (2.56) : At saturation, the gluon occupation number depends upon the variables A and $k_\perp$ only via the dimensionless ratio $k_\perp^2/Q_s^2$. This reflects the fact that, in this regime, the saturation momentum is the only intrinsic scale in the nuclear wavefunction: The soft QCD scale $\Lambda_{\rm QCD}$, which would be naturally introduced by radiative processes in the absence of non–linear effects, cf. Eq. (2.48), plays in fact no role at saturation ($k_\perp < Q_s$), since gluon radiation is suppressed there. This scaling behaviour is a simple manifestation of a more general, *'geometric scaling'*, property to be extensively discussed later (see Sect. 4.3).

There is still another way to see that the saturation scale acts effectively as an infrared cutoff: If one computes the 'integrated gluon distribution' $xG_A(x, Q^2)$ (i.e., the total number of gluons having $k_\perp^2 \leq Q^2$ in the nuclear wavefunction), then one finds that, due to the softening of the spectrum at low $k_\perp$, cf. Eq. (2.56), the integral in Eq. (2.28) is now convergent in the infrared. In particular, for $Q^2 \ll Q_s^2(A)$, Eqs. (2.28) and (2.56) yield

$$xG_A(x, Q^2) \simeq \frac{N_c^2 - 1}{4\pi N_c} \frac{1}{\alpha} R_A^2 Q^2 \ln \frac{Q_s^2(A)}{Q^2} \qquad \text{for} \qquad Q^2 \ll Q_s^2(A), \tag{2.57}$$

which shows *strong nuclear shadowing* : Unlike the 'leading–twist approximation', Eq. (2.38), which grows like the nuclear *volume* and thus scales like A, the saturation distribution
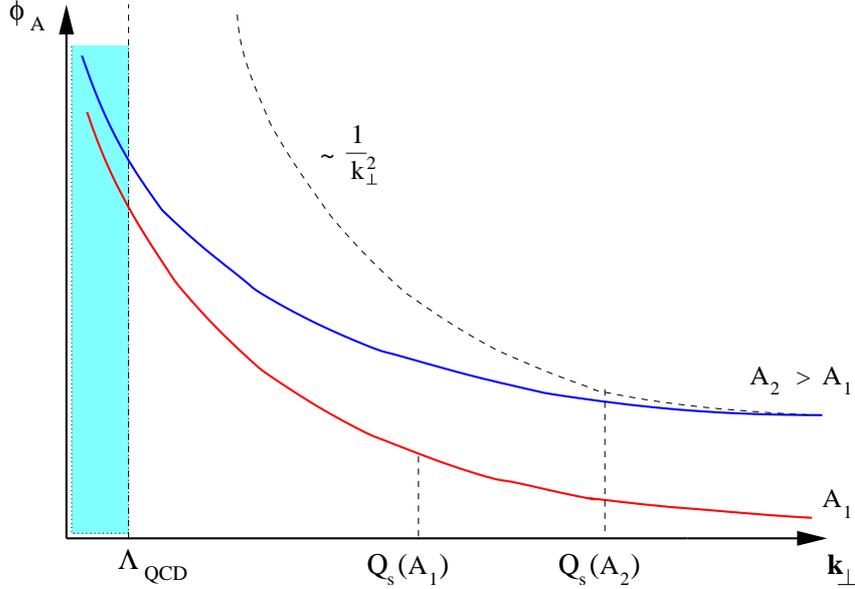
Fig. 26. The gluon phase-space density $\varphi_A(k_\perp)$ of a large nucleus (as described by the MV model) plotted as a function of $k_\perp$ for two values of A. Notice the change from a $1/k_\perp^2$ behaviour at large momenta $k_\perp > Q_s$ to a logarithmic behaviour at small momenta $k_\perp < Q_s$.

(2.57) is proportional to the nuclear *transverse area* and thus grows merely like $A^{2/3} \ln A$. But, of course, for sufficiently large $k_\perp^2 \gg Q_s^2(A)$, the leading–twist approximation becomes appropriate for the gluon spectrum, cf. Eq. (2.55), and as a matter of facts this is also a good approximation [8] for the *integrated* gluon distribution $xG_A(x,Q^2)$ at very large $Q^2 \gg Q_s^2(A)$ [87, 114] :

$$xG_A(x,Q^2) \simeq \frac{\alpha A N_c C_F}{\pi} \ln \frac{Q^2}{Q_s^2(A)} \qquad \text{for} \qquad Q^2 \gg Q_s^2(A) . \tag{2.58}$$

Except for the replacement of $\Lambda_{\rm QCD}$ by $Q_s(A)$ as the IR cutoff in the logarithm, this is the same as the bremsstrahlung result in Eq. (2.38). Eq. (2.58) shows that the *total number of gluons* in the nuclear spectrum (as measured by counting the gluons in $k_\perp$ up to a finite, but large, value $Q^2 \gg Q_s^2(A)$) is in fact the *same* in the presence of saturation as it would be in its absence.

This last observation sheds a new light on the mechanism of saturation in the context of the MV model: The non–linear effects responsible for saturation merely *redistribute* the gluons in momentum space, while intact leaving their overall number [87, 114]. The gluons which in the bremsstrahlung spectrum $1/k_\perp^2$ would be accumulated at low momenta, are in fact pushed up in $k_\perp$ by the non–linear effects, in order to minimize their repulsion. Some of these 'displaced' gluons are responsible for the 'higher–twist' contributions to the tail of the distribution at $k_\perp \gg Q_s(A)$, as visible on Eq. (2.55). But actually most of them are quasi–uniformly redistributed at momenta below $Q_s(A)$, thus giving a *saturation*

---

[8] More precisely, the difference between the distribution $xG_A(x,Q^2)$ given by the MV model and its leading–twist approximation in Eq. (2.58) dies off like $1/Q^2$ when $Q^2 \to \infty$ [87].
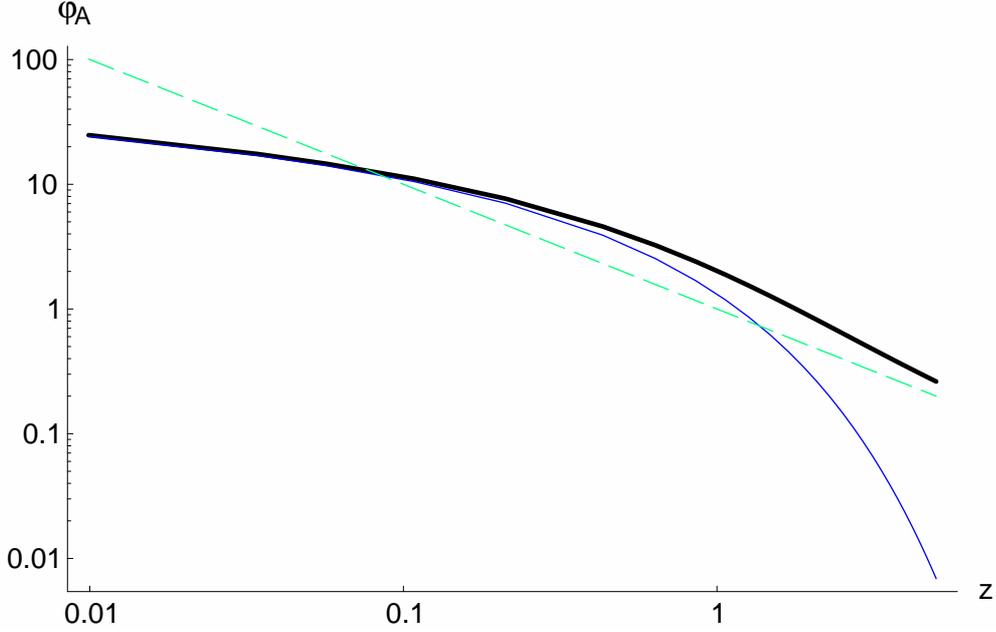
Fig. 27. *The gluon occupation factor* $\varphi_A(z)$ *as a function of the scaled momentum variable* $z = k_\perp^2/Q_s^2(A)$ *in the MV model (logarithmic units). The thick (black) line corresponds to* $\varphi_A(z)$; *the solid (blue) line shows the saturation plateau, which decreases exponentially when* $z > 1$; *the dashed (green) line represents the bremsstrahlung spectrum. When* $z \gg 1$, *the total areas enclosed below the thick line and, respectively, the dashed line are essentially the same.* [From Ref. [87].]

plateau at $k_\perp \le Q_s(A)$ (cf. Eq. (2.56)). This plateau ends up quite abruptly at momenta $k_\perp \sim Q_s(A)$, as demonstrated by the calculation in Ref. [87]: when increasing $k_\perp$ above $Q_s(A)$, $\varphi_A(k_\perp)$ starts by decreasing *exponentially* with $z = k_\perp^2/Q_s^2(A)$, before eventually relaxing to the power law decay displayed in Eq. (2.55). This behaviour is illustrated in Fig. 27, obtained via an exact calculation of Eq. (2.52) [87].

To conclude this discussion, let us briefly indicate how the previous results can be used to also understand the *unitarization of the dipole–nucleus scattering* in the framework of the MV model. Namely, after comparison with Eq. (1.37), the Wilson line correlator (2.42) with $x^- \to \infty$ is recognized as the S–matrix for an incoming color dipole in the adjoint representation (i.e., a two–gluon system in a color singlet state). In the MV model, this has been computed in Eq. (2.49). The corresponding S–matrix for a quark–antiquark pair is obtained by simply replacing $N_c \equiv T^a T^a$ by $C_F \equiv t^a t^a = (N_c^2 - 1)/2N_c$ in Eq. (2.49).

After also using Eq. (2.38), the dipole S–matrix in Eq. (2.49) can be rewritten as (for a $q\bar{q}$ dipole, for convenience)

$$S_A(r_\perp) = \exp\left\{ - r_\perp^2 \frac{\pi^2 \alpha}{2N_c} \frac{xG_A(x, 1/r_\perp^2)}{\pi R_A^2} \right\}, \tag{2.59}$$

which is recognized as the Glauber formula (1.27): the resummation of the multiple scattering series amounts to the exponentiation of the amplitude for a single scattering via two–gluon exchange (cf. Eq. (1.22)). This makes it manifest that the unitarity corrections become important for dipole sizes $r_\perp \gtrsim 1/Q_s(A)$. But the validity of the Glauber exponen-

tiation is in fact related to the simplicity of the MV model, which neglects all correlations among the color sources. As we shall see, in the presence of correlations induced by the quantum evolution such a simple exponentiation does not hold anymore.

3 Quantum evolution of the CGC

In this section, we shall explain how to construct the CGC effective theory at small x by integrating out the gluons with $x' > x$ in perturbation theory, in the presence of high density effects. The central result of this analysis will be a renormalization group equation for the weight function $W_{\Lambda^+}[\rho]$ in Eq. (2.4), which generalizes the BFKL equation by including non–linear effects, and has important physical consequences among which gluon saturation. The main observation is that, to LLA, all the quantum corrections described previously — both the exponentially developing BFKL cascade, and the gluon recombination which tames this rapid growth — can be incorporated into a change of the classical color charge and its correlations, namely, into a renormalization of the weight function $W_{\Lambda^+}[\rho]$ in Eq. (2.4) [31, 34].

To see this at an intuitive level, let us reconsider the first radiative correction, the one–gluon emission in Fig. 20.b, and note that, to LLA, the typical contributions to the integral in Eq. (1.49) come from momenta p_1^+ such that $p^+ \gg p_1^+ \gg k^+$. That is, the condition of separation of scales is indeed satisfied for the intermediate gluon with momentum p_1^+ to be treated as a 'frozen' color source for the final gluon with momentum k^+. The effect of this quantum correction is therefore simply to renormalize the *effective* color source at scale k^+, as pictorially illustrated in Fig. 28.
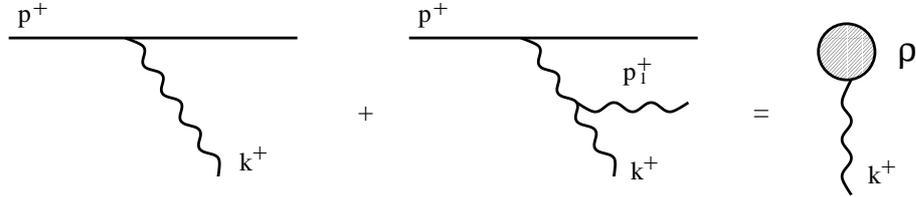
Fig. 28. *Effective color source after including the lowest-order radiative correction.*

By iterating this argument, it is quite clear that a whole BFKL cascade (see Fig. 20.c) can be included in the definition of the classical color source at the scale $\Lambda^+ = xP^+$ of interest. It is furthermore clear that the fusion between two gluon cascades, as illustrated in the l.h.s. of Fig. 29, can be represented as a non-linear effect in the *classical* dynamics of the color fields generated by this effective source (see the r.h.s. of Fig. 29).

But non–linear effects are important also in the *quantum* evolution, and actually interfere with it, as illustrated in Fig. 30. Fig. 30.a is an immediate generalization of the one–gluon emission in Fig. 20.b. It is clear that what is renormalized by the scattering off the "semi-fast" ($\Lambda^+ \gg p^+ \gg k^+$) quantum fluctuation is the classical field $\mathcal{A}^i[\rho]$ at scale Λ^+, which in turn is non–linear in ρ. Fig. 30.b shows an additional source of non–linearity, arising from the propagation of the radiated gluon in the classical 'background' field $\mathcal{A}^i[\rho]$. If $\Lambda^+ = xP^+$ is small enough ($x \ll 1$), the classical field is strong, $\mathcal{A}^i \sim 1/g$, and gluon

71
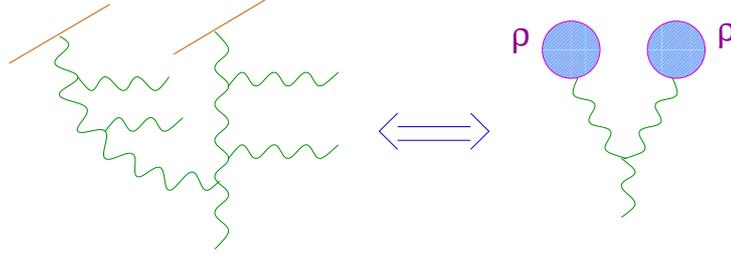
Fig. 29. *The fusion of two gluon cascades and its interpretation in the CGC theory.*

rescattering must be included to all orders in $\mathcal{A}^i$. The diagrams in Figs. 30.a and b can be both taken into account as the cut of the diagram in Fig. 30.c. The classical field that enters the vertices is the fully non–linear solution $\mathcal{A}^i[\rho]$ constructed in Sect. 2.3, and the propagator of the quantum gluon is computed to *all* orders in this background field, the resummation indicated here by a blob.
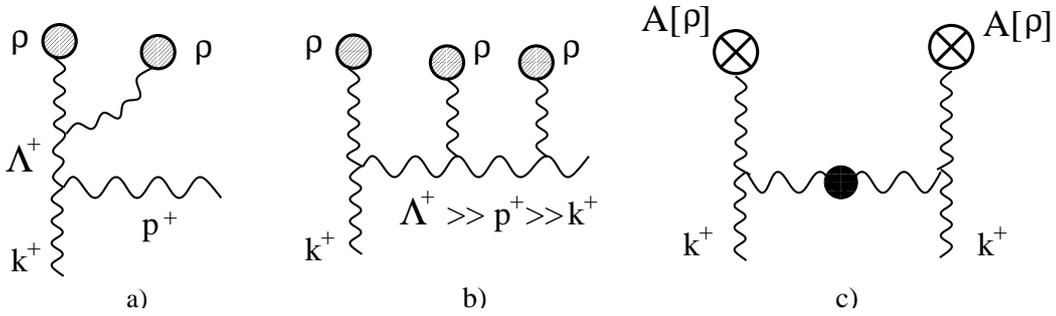
Fig. 30. *Some typical non-linear effects in the quantum evolution*

The diagram in Fig. 30.c is manifestly a quantum correction to the 2–point function of the gauge fields at scale k^+, and is of order $\alpha \ln(\Lambda^+/k^+)$. Thus, for this to be computable in perturbation theory, the separation of scales between Λ^+ and k^+ should not be too large: $k^+ \ll \Lambda^+$, but such that $\alpha \ln(\Lambda^+/k^+) \ll 1$. If this condition is satisfied, the *'semi–fast'* quantum gluons, i.e. the modes with momenta p^+ within the range $k^+ \ll p^+ \ll \Lambda^+$, can be integrated out in perturbation theory. Then this step can be iterated, resulting in a systematic *renormalization group procedure* at small–x [31, 34]. At each step in this proce-dure, one has to perform a one–loop quantum calculation, but with the exact background field propagator for the 'semi–fast' gluons. Such an all–order inclusion of the classical field effects allows one to resum not only the *large energy logarithms*, namely the terms $\sim (\alpha \ln 1/x)^n$, but also the dominant *high density effects* — the non–linear effects (like gluon recombination) which become of order one at saturation.

3.1   Renormalization group at small–x

Assume we have been able to construct the effective theory at the scale Λ^+, but we would like to compute correlations at some lower scale $b\Lambda^+$ with $b \ll 1$. This means that the d.o.f. with $k^+ > \Lambda^+$ have been already 'integrated out' and replaced, to the approximations of interest, by a classical color source with charge density $\rho_a(\vec{x})$ and correlations encoded in the weight function $W_{\Lambda^+}[\rho]$. To the same approximations, the correlations at scale $b\Lambda^+$

are determined by the color sources at all momenta $k^+ \gg b\Lambda^+$, that is, those which have been already replaced by ρ, but also the quantum gluons with longitudinal momenta in the intermediate strip $b\Lambda^+ < p^+ < \Lambda^+$. In what follows we will show how to integrate out the latter and thus construct the new effective theory at scale $b\Lambda^+$. We anticipate that, to LLA, the dominant contributions arise from the 'semi–fast' quantum gluons with momenta *deeply* within the strip $(b\Lambda^+ \ll p^+ \ll \Lambda^+)$, for which we can apply the same arguments relying on the separation of (longitudinal and time) scales as in the construction of the effective theory. Namely, one can study the quantum dynamics of the 'semi–fast' gluons for a given configuration $\rho_a(\vec{x})$ of the classical color source, and then average over the latter. As explained in Sect. 2.1, this involves two types of averaging: a *quantum* average over the semi–fast gluons in the presence of the classical color source ρ, and then a *classical* average over ρ, with weight function $W_{\Lambda^+}[\rho]$. The final result for the correlations at the lower scale $b\Lambda^+$ that we shall find after this quantum plus classical calculation should be the same as the correlations encoded in the *renormalized* weight function $W_{b\Lambda^+}[\rho]$. By imposing this *matching condition*, we shall determine the evolution of the weight function.

In this process, there will be several questions that we will have to answer: (i) First, we have to specify how the semi–fast gluons couple to the classical 'background' source ρ_a. As we shall see, the answer will uniquely follow from gauge symmetry together with the requirement that the action $S(A, \rho)$ generates the classical Yang–Mills equations with source ρ. (ii) Next, we should show that all the correlations induced by the semi–fast gluons at scale $b\Lambda^+$ can be indeed embedded in a change of the classical source and of its correlations: $\{\rho_a, W_{\Lambda^+}[\rho]\} \rightarrow \{\bar{\rho}_a, W_{b\Lambda^+}[\bar{\rho}]\}$. (For more clarity, we shall use a bar to denote the classical source at the lower scale.) This has been demonstrated in Ref. [34], but it can be understood from the fact that the separation of scales which has been invoked in the construction of the effective theory in Sect. 2.1 is preserved by the quantum evolution that we describe now. (iii) The third question refers to the calculation of the induced correlations. This is described too in detail in Ref. [34], and here will be presented as an extension of the more standard BFKL calculation (cf. Sect. 1.4). Namely, we would like to resum the same ladder diagrams as in the BFKL formalism (the 'real corrections'), together with the corresponding 'virtual corrections', except that, now, we shall perform our calculation in the presence of a strong classical 'background field', which represents the high density effects associated with 'fast' partons above the scale Λ^+. Thus, in computing the quantum corrections, we shall treat the non–linear effects associated with the background field as effects of order one (both in the vertices and in the propagators). This will require a *background field one–loop calculation*. (iv) Finally, once the quantum effects will be computed, we will have to show how to absorb these effects into a modification of the classical source and its correlations. This is the *matching procedure* alluded to above, and we shall start by presenting this procedure, since this is a logical way to organize the calculation.

We are interested in the correlation functions $\langle \bar{\rho}(1)\bar{\rho}(2) \cdots \bar{\rho}(n) \rangle_{b\Lambda^+}$ at the lower scale $b\Lambda^+$, as obtained after integrating out the semi–fast gluons and averaging over the classical source in the original effective theory at scale Λ^+. The general formula yielding these correlations reads (e.g., for the 2-point function)

$$\langle \bar{\rho}_a(\vec{x}) \bar{\rho}_b(\vec{y}) \rangle_{b\Lambda^+} = \int \mathcal{D}\rho \, W_{\Lambda^+}[\rho] \, < (\rho + \delta\rho)_a(\vec{x}) \, (\rho + \delta\rho)_b(\vec{y}) >_\rho \,, \qquad (3.1)$$

where (with the simplified notation $\rho_x \equiv \rho_a(\vec{x})$):

$$< (\rho + \delta\rho)_x \, (\rho + \delta\rho)_y >_\rho \equiv \frac{\int_{b\Lambda^+}^{\Lambda^+} \mathcal{D}a \, (\rho + \delta\rho)_x \, (\rho + \delta\rho)_y \, \mathrm{e}^{iS[\mathcal{A}+a,\,\rho]}}{\int_{b\Lambda^+}^{\Lambda^+} \mathcal{D}a \, \mathrm{e}^{iS[\mathcal{A}+a,\,\rho]}}$$

$$(3.2)$$

represents the quantum average over the semi–fast modes in the presence of the background source. The other notations are as follows: the total field $A^\mu = \mathcal{A}^\mu + a^\mu$ which appears in the action is the sum of the classical field $\mathcal{A}^\mu$ created by the source ρ (cf. Sect. 2.3) and the quantum field a^μ describing the semi–fast fast. The latter have longitudinal momenta restricted to the strip $b\Lambda^+ < p^+ < \Lambda^+$, as indicated by the lower and upper limits on the functional integral. Furthermore, $\bar{\rho}_a(\vec{x})$ is the color charge operator of the semi–fast gluons, to be specified shortly. Finally, the LC–gauge condition $A_a^+ = 0$ (i.e., $\mathcal{A}^+ = 0$ and $a^+ = 0$) is understood throughout the quantum calculation.

In order to integrate out the semi–fast gluons in a single step in the renormalization group analysis, the separation of scales between Λ^+ and $b\Lambda^+$ should not be too large: b is small, but such that $\alpha \ln(1/b) \ll 1$. Then we can perform a one loop calculation in order to compute the effects induced by the quantum gluons to $\mathcal{O}(\alpha d\tau)$, with $d\tau = \ln(1/b)$ the corresponding step in rapidity. To that aim, we also need to specify the action $S[A, \rho]$. By definitions, this must be such as to reproduce the classical field equations (2.1) at the tree-level. The simple guess $S[A, \rho] = S_{\mathrm{YM}}[A] + \int_x \rho_a A_a^-$ cannot be right since the second term, involving ρ, is not gauge–invariant. This reflects the fact that, written as in Eq. (2.1), the classical EOM is not manifestly gauge–covariant either. In fact, Eq. (2.1) is correct as written only for field configurations having $A_a^- = 0$. This is not a limitation for *classical* calculations, since it is always possible to construct a classical solution having this property. But Eq. (2.1) is not sufficient to determine $S[A, \rho]$, which is explicitly needed for the *quantum* calculation.

To find the general equation which replaces Eq. (2.1) in some arbitrary gauge (where $A_a^- \neq 0$), notice that, in general, the current J_a^μ in the r.h.s. of the Yang-Mills equations must satisfy the *covariant* conservation law $D_\mu J^\mu = 0$ (since we also have $D_\mu D_\nu F^{\nu\mu} = 0$). For the eikonal current (1.40), this implies $D^- J^+ \equiv (\frac{\partial}{\partial x^+} - ig A^-) J^+ = 0$, which reduces indeed to $\partial^- \rho = 0$ when $A^- = 0$. But in general, this is satisfied by $J^+(x^+, \vec{x}) = W(x^+, \vec{x}) \, \rho(\vec{x}) \, W^\dagger(x^+, \vec{x})$. We have introduced here the temporal Wilson line:

$$W[A^-](x^+, \vec{x}) \equiv \mathrm{T} \exp \left\{ ig \int_{-\infty}^{x^+} dz^+ A^-(z^+, \vec{x}) \right\}, \qquad (3.3)$$

with T denoting the ordering of the color matrices in the exponent from right to left in increasing order of their x^+ arguments. Thus, in general, the non–Abelian current J_a^+ can be static only up to a color precession. The action generating the classical field equations with current (3.3) reads [31, 34]

$$S[A, \rho] = -\int d^4x \, \frac{1}{4} \, F^a_{\mu\nu} F^{\mu\nu}_a + \frac{i}{gN_c} \int d^3\vec{x} \operatorname{Tr} \left\{ \rho(\vec{x}) \, W[A^-](\vec{x}) \right\}, \tag{3.4}$$

where $W[A^-](\vec{x})$ is given by Eq. (3.3) with $x^+ \to \infty$.

Note that the Wilson–line piece of the action is *a priori* non–linear in A^- to all orders, but in what follows only the first few terms in its expansion will be actually needed: Indeed, since the background field has $\mathcal{A}^- = 0$, it follows that $A^- = a^-$ is a weak field. For the purposes of the one–loop calculation, the action $S[\mathcal{A} + a, \rho]$ in Eq. (3.4) can be expanded to second order in a^μ:

$$S[\mathcal{A} + a, \rho] \approx \frac{1}{2} \int d^4x \int d^4y \, a^\mu_a(x) \left. \frac{\delta^2 S}{\delta A^\mu_a(x) \delta A^\nu_b(y)} \right|_{\mathcal{A}} a^\nu_b(y). \tag{3.5}$$

(The first two terms in this expansion vanish since $\mathcal{A}^i$ is a solution of the classical equations of motion $\delta S/\delta A^\mu = 0$, and, moreover, the action itself vanishes on this solution: $S[\mathcal{A}, \rho] = 0$.) Thus, the propagator for the semi–fast fields is obtained by inverting the differential operator which represents the kernel in the above integral, in the LC gauge and in background of the classical field.

The final object from Eq. (3.2) that we have to specify is the color charge density $\delta\rho_a$ of the semi–fast gluons. This is the operator through which these gluons couple to the slower fields with momenta $k^+ \lesssim b\Lambda^+$ for which the new effective theory is built. To identify $\delta\rho_a$, it is useful to temporarily introduce the field δA^μ which represent these slow gluons — such that, now, $A^\mu = \mathcal{A}^\mu + a^\mu + \delta A^\mu$ — and notice that in the eikonal approximation it is sufficient to retain the linear coupling $\delta A^-_a \delta\rho_a$ within the action $S[\mathcal{A} + a + \delta A, \rho]$. We thus deduce that:

$$\delta\rho_a(x) \equiv \left. \frac{\delta S}{\delta A^-_a(x)} \right|_{\mathcal{A}+a}, \tag{3.6}$$

where it is understood that only terms linear and quadratic in a^μ must be kept in the expansion of the expression in the r.h.s. in powers of the quantum field.

To gain some more intuition, we use as an example the contributions to $\delta\rho_a$ coming from the Yang-Mills piece of the action, $S_{YM} = \int d^4x(-F^2_{\mu\nu}/4)$:

$$\left. \delta\rho_a(x) \right|_{YM} = 2g f^{abc} \mathcal{F}^{+i}_b(\vec{x}) a^i_c(x) + g f^{abc} (\partial^+ a^i_b(x)) a^i_c(x). \tag{3.7}$$

The first term in the r.h.s., which is linear in a^i, is the only one to contribute to the *induced charge–charge correlation* $< \delta\rho_a \, \delta\rho_b >_\rho$ to leading order in α. It generates the tree-like diagram in Fig. 31.a, where the internal line with a blob represents the background field propagator $G^{ij}(x,y)$ of the semi-fast gluons. Physically, Fig. 31.a describes the emission of an on-shell (or "real") semi-fast gluon by the classical source.

Since $< a^i >_\rho = 0$, it is only the second, quadratic term in the r.h.s. of Eq. (3.7) which contributes to the *induced color charge density* $< \delta\rho_a >_\rho$. In Fig. 31.b we show such a contribution of lowest order in ρ. (This involves also vertices from the Wilson line piece of the action, Eq. (3.4).) Obviously, this represents a vertex correction to the tree-level emission in Fig. 20.a.
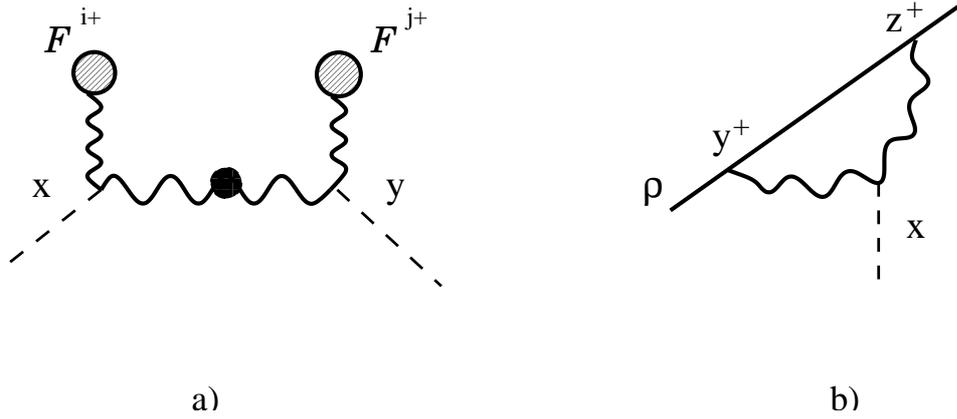
Fig. 31. *Typical Feynman diagrams contributing to χ (a) and σ (b). The internal wavy lines are propagators of the semi–fast gluons in the background field; the external dotted lines carry soft momenta, and couple to the fields δA^-. In Fig. (a), the external blobs denote insertions of the electric field $\mathcal{F}^{+i}$. In Fig. (b), the continuous line represents the source ρ.*

The structures illustrated by Figs. 31.a and b are generic: the induced charge–charge correlator $<\delta\rho_a\,\delta\rho_b>_\rho$ is the "real" correction, whose iteration generates the gluon cascades; the induced charge $<\delta\rho_a>_\rho$ is the "virtual" correction, which provides one–loop corrections to the emission vertices in these cascades, and also self–energy corrections. These quantities will play an important role in what follows, so it is useful to introduce specific names for them: we shall refer to them as χ and σ, respectively. Both χ and σ include terms non–linear in ρ which ultimately describe interactions among gluons at different rapidities and from different cascades. In general, real and virtual corrections are related by gauge symmetry, and this is also the case for χ and σ, as we shall discuss later.

Note that so far we have not considered correlations of $\delta\rho_a$ higher than the 2–point function: to the accuracy of interest, this is actually not needed. This can be checked by explicit power counting using equations like (3.7), but it also follows by analogy with the more standard BFKL resummation, that we generalize here merely by including the non–linear effects associated with the background source. In terms of Feyman diagrams, these non–linearities amount to $n \to 2$ *gluon mergings*, that is, vertices in which the number of the incoming gluon legs (corresponding to fast fields with higher values of x) is arbitrary, but there are just two outgoing legs (for the coupling to the slow fields) — see, e.g., Fig. 30.b; such diagrams induce a 2–point correlation $<\delta\rho_a\,\delta\rho_b>_\rho$, but not also higher n–point correlations with $n \geq 3$. We shall return to the diagrammatic interpretation after presenting the final evolution equation.

Consider now the space–time structure of the induced correlations. Because of Lorentz time dilation, the semi–fast gluons appear as frozen to the slow fields δA^μ, so the latter can only probe *equal–time correlations* of the former. This was already implicit in Eqs. (3.1)–(3.2), where the induced sources $\delta\rho_a$ were taken at equal times x^+. Because of that, and also by time homogeneity (which is not broken by the background source, which is itself static), we conclude that both χ and σ are independent of x^+.

By using similarly the argument of Lorentz contraction, one is tempted to argue that the slow fields δA^μ cannot discriminate the longitudinal structure of the induced source

76

— and of the total source $\rho_a + \delta\rho_a$ in general — but only 'see' it as a distribution concentrated at $x^- = 0$. This would be indeed the case in the BFKL approximation, where a complete separation holds between the longitudinal and transverse dynamics (the "$k_\perp$–factorization"; see, e.g., [123]). However, in the presence of non–linear effects, this is not true any longer: Recall that these non–linear effects describe the *multiple scattering* of the semi–fast gluons off the background field $\mathcal{A}$ created by ρ. Since the p^+ momenta of the semi–fast gluons are much lower than those of the fast partons responsible for ρ, this scattering can be described in the eikonal approximation, as a Wilson line directed along x^-. (Indeed, when seen in the rest frame of the background field, the semi–fast gluons appear to move in the negative z direction, that is, towards increasing x^-.) This eikonal line would be easier to visualize in a gauge other than the LC gauge, e.g., in the covariant gauge in which $\mathcal{A}^+ \equiv \alpha \neq 0$, but this is also manifest in the LC gauge, where it enters the structure of the transverse components $\mathcal{A}^i$ (cf. Eqs. (2.13) and (2.18)). Now, this eikonal line is restricted to the longitudinal extent of the source, namely $\Delta x^- \lesssim 1/\Lambda^+$, which is relatively small on the resolution scale of the semi–fast gluons (with momenta $p^+ \ll \Lambda^+$), yet this cannot be simply treated as a δ–function — i.e., one cannot replace $\rho(x^-, \boldsymbol{x})$ by $\delta(x^-)\rho(\boldsymbol{x})$ — because of the *color ordering* of the successive scatterings along the trajectory (the path–ordering in Eq. (2.18)). That is, in QCD, the successive collisions do not commute with each other, and to keep trace of that within the formalism, one needs the preserve some information about the x^-–dependence of the classical source and its correlations: namely, that information which is encoded in the Wilson lines. We have already met with a similar situation in the description of gluon saturation within the MV model, in Sect. 2.5.

More precisely, what one needs to know is that, when integrating out a new layer in p^+ of quantum fluctuations, the support of $\rho_a(\vec{x})$ extends to larger values of x^-. This is simply a consequence of the uncertainty principle: Since the new modes that are included in the effective theory have momenta within the strip $b\Lambda^+ < p^+ < \Lambda^+$, it is clear that these modes are delocalized in x^-, over a typical distance $\Delta x^- \lesssim 1/\Lambda^+$. In fact, it turns out that the dominant part (in the sense of the LLA) of the support of $\delta\rho_a$ lies in the region

$$1/\Lambda^+ \;\lesssim\; x^- \;\lesssim\; 1/(b\Lambda^+)\,. \tag{3.8}$$

Indeed, the calculation of the induced charge yields a result of the form [34]

$$<\delta\rho_a(x^-, \boldsymbol{x})>_\rho = \; F_\Lambda(x^-)\,\sigma_a(\boldsymbol{x})\,, \tag{3.9}$$

where $\sigma_a(\boldsymbol{x})$ starts at order $\alpha\rho$, as expected for a one–loop correction, but in general involves also non–linear terms $\sim \alpha\rho(g\rho)^n$, $n \geq 1$, whereas

$$F_\Lambda(x^-) \equiv \theta(x^-) \frac{e^{-ib\Lambda x^-} - e^{-i\Lambda x^-}}{x^-}\,, \quad \int dx^- F_\Lambda(x^-) \;=\; \ln\frac{1}{b} \;=\; d\tau, \tag{3.10}$$

have its support typically at[9] $1/\Lambda^+ \lesssim x^- \lesssim 1/(b\Lambda^+)$ and produces the expected logarithmic enhancement after integration over x^-.

---

[9] Indeed, $F_\Lambda(x^-) \approx 0$ both for small $x^- \ll 1/\Lambda^+$ (since in this case the two exponentials mutually cancel), and for large $x^- \gg 1/b\Lambda^+$ (where the two exponentials are individually small).

This enables us to draw two important lessons: First, that by integrating the quantum fluctuations in layers of p^+, one builds the classical source ρ in layers of x^-, with a one-to-one correspondence between the longitudinal momenta of the modes that have been integrated out and the the x^- coordinate of the new layer that has been generated in that way. Second, that the precise longitudinal structure of a particular layer is not important to LLA, but only the fact that this layer is located on top (in the sense of having a larger value of x^-) of the source created in the previous steps. By induction, we deduce that $\rho_a(\vec{x})$ ($\equiv$ the total colour charge generated by the quantum evolution down to Λ^+) has support at $0 \leq x^- \lesssim 1/\Lambda^+$, as anticipated in Sect. 2.3. Thus, the quantum evolution with τ proceeds by adding new layers to the color source in x^-, whereas the color source generated in the previous steps is not modified (see Fig. 32). For each new layer, its internal longitudinal structure is unimportant, but the relative ordering of the various layers must be correctly preserved, because the color matrices $\delta\rho(x^-) \equiv \delta\rho_a(x^-)T^a$ at different positions in x^- do not commute with each other. As we shall see, this ordering is naturally incorporated into the formalism with the help of Wilson lines.

The previous discussion also shows that the quantum evolution to LLA leads to a *coarse–graining* in longitudinal direction, with the natural width of the bin in x^- fixed by the rapidity step $d\tau$, as manifest on Eqs. (3.8)–(3.10). Thus, when discussing the high energy evolution, τ should be properly treated as a *discrete* variable, with a typical value for the step $d\tau$ of the order $1/\alpha$ (since when $\alpha\tau \sim 1$ the resummation becomes necessary). In view of the tight correspondence between p^+ and x^- alluded to above, and of the fact that the evolution is logarithmic in p^+ — the *evolution time* is $\tau = \ln P^+/\Lambda^+$ —, it is natural to introduce logarithmic units also for x^-. Namely, let us define the *space–time rapidity* y via (for positive x^- and $x_0^- = 1/P^+$):

$$\mathrm{y} \equiv \ln(x^-/x_0^-), \qquad -\infty < \mathrm{y} < \infty. \tag{3.11}$$

By the renormalization group analysis, y and τ are identified with each other, so y is a discrete variable as well, with step $d\mathrm{y} = d\tau$. It is then natural to replace the density $\delta\rho_a(x^-, \boldsymbol{x})$ whose explicit x^-–dependence goes beyond LLA by the corresponding quantity integrated over x^-:

$$\delta\rho_\tau^a(\boldsymbol{x})\,d\tau \equiv \int dx^- \delta\rho^a(x^-, \boldsymbol{x}), \tag{3.12}$$

where the subscript τ is introduced to remind that this is the color charge induced in the momentum rapidity step $(\tau, \tau + d\tau)$, and, at the same time, is localized in the space–time rapidity bin $(\mathrm{y}, \mathrm{y} + d\mathrm{y})$, with $\mathrm{y} = \tau$ and $d\mathrm{y} = d\tau$. Then, clearly,

$$\sigma^a(\boldsymbol{x}) = <\delta\rho_\tau^a(\boldsymbol{x})>_\rho \equiv \sigma_\tau^a(\boldsymbol{x}). \tag{3.13}$$

From now on, by τ we shall understand both the usual rapidity ($\tau = \ln 1/x$), and the space–time rapidity of the color charge distribution generated in the corresponding step of the evolution.

Consider similarly the induced 2–point correlation $\chi = <\delta\rho_a(x^+, \vec{x})\delta\rho_b(x^+, \vec{y})>_\rho$. As discussed on the example of Eq. (3.7), the component of $\delta\rho_a$ relevant for computing χ is linear in the quantum field a^μ, and also proportional to the background color source ρ_a (or the associated field $\mathcal{F}^{+i}$). This component is therefore localized within the relatively
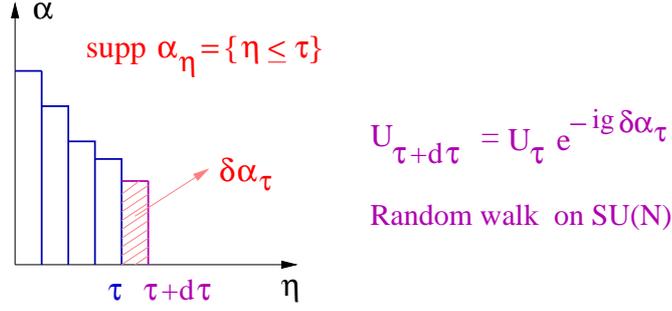
Fig. 32. *The longitudinal structure of the source built through the evolution (η denotes the space–time rapidity).* HISTY

small interval $0 \leq x^- \lesssim 1/\Lambda^+$, so $<\delta\rho_a(\vec{x})\delta\rho_b(\vec{y})>_\rho$ is *quasi–local* in x^- (i.e., it is non–vanishing only for small separations $x^- - y^-$). After coarse–graining in x^- as in Eq. (3.12), the ensuing distribution is local in space–time rapidity. That is,

$$<\delta\rho_\tau^a(\boldsymbol{x})\delta\rho_\tau^b(\boldsymbol{y})>_\rho = \frac{1}{d\tau}\,\chi_\tau^{ab}(\boldsymbol{x},\boldsymbol{y})$$

$$\chi_\tau^{ab}(\boldsymbol{x},\boldsymbol{y})d\tau \equiv \int dx^- \int dy^- <\delta\rho^a(x^-,\boldsymbol{x})\,\delta\rho^b(y^-,\boldsymbol{y})>_\rho, \tag{3.14}$$

where the quantity $\chi_\tau^{ab}(\boldsymbol{x},\boldsymbol{y})$ starts at order $\alpha\rho^2$.

We can summarize the previous discussion in this section as follows: The color source

$$\rho_{\mathrm{y}}^a(\boldsymbol{x}) \equiv \rho^a(x^-,\boldsymbol{x})\frac{dx^-}{d\mathrm{y}} = x^-\rho^a(x^-,\boldsymbol{x}), \tag{3.15}$$

(we have set $x^- = x_0^- e^{\mathrm{y}}$) generated by the quantum evolution from $\tau_0 = 0$ up to τ has support in the interval $0 \leq \mathrm{y} \leq \tau$ and is distributed according to the weight function $W_\tau[\rho] \equiv W_\Lambda[\rho]$. When a new layer of quantum modes, with rapidities in the interval $(\tau, \tau + d\tau)$, are integrated out, the preexisting colour source at $\mathrm{y} \leq \tau$ is not changed, but a new contribution $\delta\rho_\tau^a(\boldsymbol{x})$ is added to it, in the rapidity bin $\tau < \mathrm{y} < \tau + d\tau$. For a given distribution of the background source $\rho_{\mathrm{y}}^a(\boldsymbol{x})$, the induced source $\delta\rho_\tau^a(\boldsymbol{x})$ is a Gaussian random variable, with probability distribution

$$\mathcal{W}_{d\tau}[\delta\rho|\rho] = e^{-\frac{1}{2}\mathrm{Tr}\ln\chi_\tau}\,e^{-\mathcal{A}_\tau[\delta\rho]}$$

$$\mathcal{A}_\tau[\delta\rho] \equiv \frac{d\tau}{2}\int_{\boldsymbol{x},\boldsymbol{y}}(\delta\rho_\tau - \sigma_\tau)_{\boldsymbol{x}}^a[\chi_\tau^{-1}]_{\boldsymbol{x},\boldsymbol{y}}^{ab}(\delta\rho_\tau - \sigma_\tau)_{\boldsymbol{y}}^b. \tag{3.16}$$

The induced correlations σ and χ are functionals of ρ (generally, non–linear).

We are now in a position to deduce the formal structure of the evolution equation for $W_\tau[\rho]$. The correlations $\langle\bar{\rho}(1)\bar{\rho}(2)\cdots\rangle_{\tau+d\tau}$ of the total color source $\bar{\rho} = \rho + \delta\rho$ generated by the evolution up to rapidity $\tau + d\tau$ can be computed in two ways:

i) As classical correlations in the effective theory with weight function $W_{\tau+d\tau}[\bar{\rho}]$:

$$\langle\bar{\rho}(1)\bar{\rho}(2)\cdots\rangle_{\tau+d\tau} = \int \mathcal{D}\bar{\rho}\,W_{\tau+d\tau}[\bar{\rho}]\,\bar{\rho}(1)\bar{\rho}(2)\cdots, \tag{3.17}$$

ii) As the correlations of $\rho + \delta\rho$, where ρ is distributed according to the original weight function $W_\tau[\rho]$, whereas for fixed ρ, $\delta\rho$ has the Gaussian weight function shown in Eq. (3.16):

$$\langle \bar{\rho}(1)\bar{\rho}(2)\cdots\rangle_{\tau+d\tau} = \int \mathcal{D}\rho \, W_\tau[\rho] \int \mathcal{D}\delta\rho \, \mathcal{W}_{d\tau}[\delta\rho \,|\rho] \,(\rho+\delta\rho)(1)(\rho+\delta\rho)(2)\cdots. \quad (3.18)$$

By identifying the r.h.s.'s of the two above equations, we shall soon deduce the difference $W_{\tau+d\tau}[\bar{\rho}] - W_\tau[\rho]$, and then the evolution equation for $W_\tau[\rho]$. But in order to do so, we have to cope with the fact that the fields ρ and $\bar{\rho}$ have different supports in y (since, as compared to ρ, $\bar{\rho}$ involves an additional layer at $\tau < y < \tau + d\tau$), so the very definition of the Hilbert space for the effective theory evolves with τ. To take that into account, we shall extend the Hilbert space for ρ to arbitrarily large values of y, but include the information about the physical support of ρ in the structure of the weight function. That is, one should think about $W_\tau[\rho]$ as having the following general structure:

$$W_\tau[\rho] = \mathcal{W}_\tau[\rho_<] \, \delta_\tau[\rho_>] , \quad (3.19)$$

where $\rho_<$ ($\rho_>$) is the restriction of ρ to $0 \le y \le \tau$ (respectively, $\tau < y < \infty$). The function $\mathcal{W}_\tau[\rho_<]$ describes the correlations of ρ within the range of y where the color charge is non–trivial, while the functional δ–function

$$\delta_\tau[\rho_>] \equiv \prod_{y>\tau} \prod_a \prod_{\boldsymbol{x}} \delta\!\left(\rho_y^a(\boldsymbol{x})\right) \quad (3.20)$$

enforces the field variable ρ_y to be identically zero for any $y > \tau$.

With this interpretation for $W_\tau[\rho]$, there is no difference between the field variables ρ in Eq. (3.18) and, respectively, $\bar{\rho}$ in Eq. (3.17), so one can renote $\bar{\rho}$ simply as ρ, then make a change of variables $\rho \to \rho - \delta\rho$ in the integral over ρ in Eq. (3.18), and finally use the equality between the two equations to deduce the following condition on the new weight function $W_{\tau+d\tau}$:

$$W_{\tau+d\tau}[\rho] = \int \mathcal{D}\nu_\tau \, W_\tau[\rho - \nu_\tau] \, \mathcal{W}_{d\tau}[\nu_\tau | \rho - \nu_\tau] , \quad (3.21)$$

where we have temporarily renoted $\delta\rho_\tau \equiv \nu_\tau$, for more clarity [10] . Eq. (3.21) is a functional recurrence formula for $W[\rho]$, whose r.h.s. can be expanded as

$$W_\tau[\rho - \nu_\tau] \, \mathcal{W}_{d\tau}[\nu_\tau | \rho - \nu_\tau] \approx \left(1 - d\tau \int_{\boldsymbol{x}} \nu_\tau^a(\boldsymbol{x}) \frac{\delta}{\delta\rho_\tau^a(\boldsymbol{x})}\right.$$
$$\left. + \frac{(d\tau)^2}{2} \int_{\boldsymbol{x},\boldsymbol{y}} \nu_\tau^a(\boldsymbol{x})\,\nu_\tau^b(\boldsymbol{y}) \frac{\delta^2}{\delta\rho_\tau^a(\boldsymbol{x})\,\delta\rho_\tau^b(\boldsymbol{y})}\right) W_\tau[\rho]\,\mathcal{W}_{d\tau}[\nu_\tau | \rho] . \quad (3.22)$$

The factors of $d\tau$ appear in this expansion because $\rho \equiv \rho_y^a(\boldsymbol{x})$ is a function of both y and $\boldsymbol{x}$, whereas ν_τ is rather a function of $\boldsymbol{x}$ alone and occupies a single bin in y, namely the bin of width $d\tau$ located at $y = \tau$. Thus, e.g.,

---

[10] Namely, in order to avoid a possible confusion between the new charge induced in one step of the evolution and the argument of the functional derivatives introduced in Eq. (3.22).

$$W_\tau[\rho - \nu_\tau] - W_\tau[\rho] \approx - \int d\mathrm{y} \int_{\boldsymbol{x}} \nu_\tau^a(\boldsymbol{x}) \frac{\delta W_\tau}{\delta \rho_\mathrm{y}^a(\boldsymbol{x})} = -d\tau \int_{\boldsymbol{x}} \nu_\tau^a(\boldsymbol{x}) \frac{\delta W_\tau}{\delta \rho_\tau^a(\boldsymbol{x})} \,. \tag{3.23}$$

It is now possible to perform the Gaussian integration over ν_τ in Eq. (3.21), by using Eq. (3.16) and after extracting the functional derivatives outside the integration:

$$W_{\tau+d\tau} - W_\tau = -d\tau \int_{\boldsymbol{x}} \frac{\delta}{\delta \rho_\tau^a(\boldsymbol{x})} \big[W_\tau \sigma_\tau^a(\boldsymbol{x}) \big] + \frac{d\tau}{2} \int_{\boldsymbol{x},\boldsymbol{y}} \frac{\delta^2}{\delta \rho_\tau^a(\boldsymbol{x}) \, \delta \rho_\tau^b(\boldsymbol{y})} \big[W_\tau \chi_\tau^{ab}(\boldsymbol{x}, \boldsymbol{y}) \big] \,.$$
$$\tag{3.24}$$

One can now understand why a limited expansion to second order in ν has been sufficient in Eq. (3.22) : According to Eq. (3.16), the 2–point function $< \nu_\tau \nu_\tau >_\rho = \chi_\tau / d\tau$ is of $\mathcal{O}(1/d\tau)$, so the terms of higher order in ν in the expansion (3.22) would give contributions which vanish faster than $d\tau$ as $d\tau \to 0$. Clearly, such terms do not contribute to the *renormalization group equation* (RGE) for $W_\tau[\rho]$, which is obtained by dividing both sides of Eq. (3.24) by $d\tau$ and then letting $d\tau \to 0$:

$$\frac{\partial W_\tau[\rho]}{\partial \tau} = \frac{1}{2} \int_{\boldsymbol{x},\boldsymbol{y}} \frac{\delta^2}{\delta \rho_\tau^a(\boldsymbol{x}) \delta \rho_\tau^b(\boldsymbol{y})} [W_\tau \chi_{xy}^{ab}] - \int_{\boldsymbol{x}} \frac{\delta}{\delta \rho_\tau^a(\boldsymbol{x})} [W_\tau \sigma_x^a] \,. \tag{3.25}$$

We have introduced here the compact notations $\sigma_x^a \equiv \sigma_a(\boldsymbol{x})$, $\chi_{xy}^{ab} \equiv \chi_{ab}(\boldsymbol{x}, \boldsymbol{y})$.

Eq. (3.25) is a second–order functional differential equation, similar to a diffusion equation, but with the particularity that the functional derivatives are taken with respect to the random variable of the last rapidity bin, i.e. ρ_τ, which, according to the previous discussion, is the only one involved in going from $W_\tau[\rho]$ to $W_{\tau+d\tau}[\rho]$. This particularity has not been recognized in early attempts to derive such an equation [31–33, 150], in which the longitudinal extent of the classical source has been overlooked (that is, one has implicitly assumed $k_\perp$–factorization), so the corresponding Hilbert space was not suitable for describing the non–commutativity of the color matrices associated with fields at different values of y (or x^-).

3.2 The JIMWLK equation

For Eq. (3.25) to be complete and useful, one still has to compute the functional coefficients σ_x^a and χ_{xy}^{ab} which enter there. As discussed in the previous section, these are obtained via a one–loop calculation in the presence of a generally strong background field. The corresponding calculation is too lengthy and technical to be presented here in detail (this can be found in Ref. [34]). Rather, we shall explain some of the difficulties and subtle technical points that we have had to clarify during this construction, and then present its final results. Following an original suggestion by Mueller [151], and as a recognition of the main contributions [31, 32, 34, 35] leading to the establishment of this important result, the ensuing RGE is often referred to as the "JIMWLK equation".

Here are some of the difficulties that have been encountered and clarified in the calculation in Ref. [34] together with the related work in the literature:

81

i) *The construction of the background field propagator.* This is the propagator of the semi–fast gluons in the presence of the classical field of the fast partons, and has been obtained by inverting the differential operator in the quadratic action (3.5) in the LC gauge $\mathcal{A}^+ = a^+ = 0$. The difficulty consists in including the non–linear effects associated with the classical field $\mathcal{A}^i$ to all orders, and has been overcome in Refs. [34, 152, 153] by exploiting the separation of scales in the problem (here, between fast and semi–fast gluons) to simplify the longitudinal structure of the background field to the form shown in Eqs. (2.21)–(2.22); that is, the whole dependence upon x^- reduces to θ–functions. The ensuing propagator describes the eikonal scattering of a semi–fast gluon off the color field produced by a classical source located near $x^- = 0$.

ii) *The choice of a gauge–fixing prescription for the "axial pole".* As explained in Sect. 2.1, the choice of the LC gauge is not sufficient to completely fix the gauge: there is a residual gauge symmetry, in the form of x^-–independent gauge transformations, which manifests itself through unphysical zero modes with $p^+ = 0$ (the "axial pole" in the LC–gauge propagator). To remove these modes, and thus completely fix the gauge, one needs a prescription for the pole at $p^+ = 0$. This is a tricky issue since a light–cone partonic wavefunction is not gauge–invariant, and in fact it exists only in suitable gauges. In the present context, the correct prescription for the axial pole is the *retarded* one (cf. Eq. (2.35)), for consistency with the boundary condition (2.19) used for the classical solution. This is in agreement with previous studies of the role of the axial prescription on the light–cone wavefunction [28, 30, 151]. But it is important to stress that, when applied to the physical scattering amplitudes, the JIMWLK equation turns out to be equivalent [34, 35, 40] with an approach by Balitsky [39], which is formulated in the covariant gauge (see Sect. 3.5 for details). Also, one may notice in this context that some of the difficulties encountered in previous attempts [32, 33, 150] to compute the coefficients in the RGE can be traced back to the issue of the gauge fixing.

iii) *Performing the one–loop calculation.* Once the background field propagator is known and the axial prescription has been chosen, one is in a position to perform the one–loop calculation yielding the coefficients σ_x^a and χ_{xy}^{ab}. These are computed according to Eqs. (3.12)–(3.14), with $\delta\rho^a(\vec{x})$ determined from the action (3.4) via Eq. (3.6) expanded to second order in a^μ. For illustration, we display in Fig. 33 the Feynman diagrams contributing to χ. The first of these diagrams (which has been already presented in Fig. 31) involves only vertices from the Yang–Mills piece of the action (3.4), whereas the three other diagrams involve also vertices from the Wilson line piece. These diagrams look formally like *tree graphs*, but the loop is actually closed in *time*, because of the equal–time condition $x^+ = y^+$ implicit in Eq. (3.14). In the weak field approximation, χ reduces to the real piece of the BFKL kernel [31, 34], and we shall often refer to it as the *"real correction"*. The corresponding diagrams for σ (the *"virtual correction"*) are genuine one–loop graphs, and can be found in Ref. [34] (an example has been shown in Fig. 31.b). Note that, because of the spatial inhomogeneity of the background field (and thus also of the propagator), the one–loop calculation must be performed in *coordinate* space.

iv) *Restoring the longitudinal dimension of the evolution.* The coefficients σ_x^a and χ_{xy}^{ab} obtained as a result of the one–loop calculation are functions of the transverse coordinate $\boldsymbol{x}$ and functionals of the Wilson lines V and $V^\dagger$, which are themselves functions of $\boldsymbol{x}$,
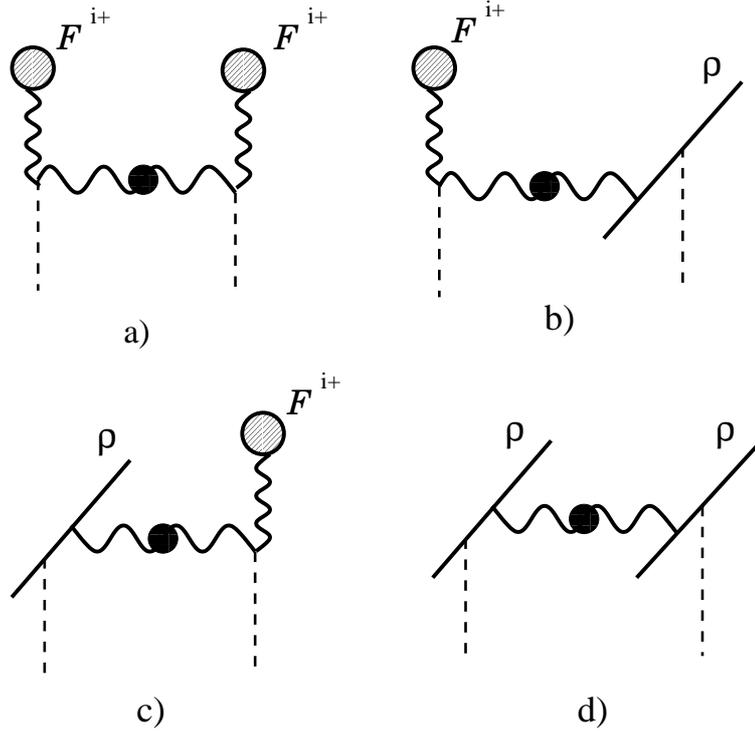
Fig. 33. *Feynman diagrams for the four contributions to χ arising in the effective theory with action (3.4). We use the same graphical symbols as in Fig. 31.*

cf. Eq. (2.22). If the effective theory was written directly in terms of Wilson lines, rather than in terms of ρ, one could satisfy oneself with such a two–dimensional description, in which the longitudinal coordinate x^- is not explicitly present. This is indeed the case for the effective theory written down by Weigert [35], which is an elegant reformulation of Balitsky equations [35], and which turns out to be equivalent to JIMWLK equation [40] (see Sect. 3.5). But even in that case, the problem of properly ordering the color matrices is an important one, and is solved with the help of Lie derivatives [35, 40].

But in the CGC effective theory, the functional space is spanned by the field $\rho^a(x^-, \boldsymbol{x})$, and the Wilson lines (2.22) are explicitly constructed as path–ordered exponentials of this field. Thus, the longitudinal direction x^- cannot be ignored. As explained in the previous section, the mathematical way to deal with this issue is to ascribe σ_x^a and χ_{xy}^{ab} to the space–time rapidity bin $(\tau, \tau + d\tau)$, and to interpret the functional derivatives in the RGE (3.25) as derivatives with respect to the 2–dimensional field $\rho_\tau^a(\boldsymbol{x})$ inside this last bin.

v) *The rotation to the background field covariant gauge.* Since the RGE is supposed to act on the Hilbert space spanned by ρ^a, its coefficients σ_x^a and χ_{xy}^{ab} must be known explicitly as functionals of the classical source, which in the quantum calculation has been taken in the *light–cone* gauge. Yet, the resulting coefficients are obtained as functionals of the Wilson lines (2.22), which are themselves expressed in terms of the source $\tilde{\rho}$ in the *covariant* gauge (the only gauge in which we know the solution to the Yang–Mills equations explicitly; cf. Sect. 2.3). To make the RGE tractable, it is therefore necessary to perform a rotation to the covariant gauge. This is a non–linear transformation on the Hilbert space, since the gauge rotation itself depends upon $\tilde{\rho}$. This is further complicated by the fact that the very

83

definition of the 'covariant gauge' depends upon the color charge content in the problem, and thus evolves with τ. This evolution brings in a supplementary contribution to $\tilde{\sigma}_a(\boldsymbol{x})$ (the induced charge in the COV–gauge) in addition to the expected piece $V_{ab}^\dagger(\boldsymbol{x})\sigma_b(\boldsymbol{x})$. As for the induced 2–point correlation, we have simply $\tilde{\chi}_{ab}(\boldsymbol{x}, \boldsymbol{y}) = V_{ac}^\dagger(\boldsymbol{x})\chi_{cd}(\boldsymbol{x}, \boldsymbol{y})\,V_{db}(\boldsymbol{y})$.

vi) *The α–representation.* Finally, it turns out that the RGE looks simpler if the COV–gauge field α_a (with $-\nabla_\perp^2 \alpha = \tilde{\rho}$, cf. Eq. (2.14)), rather than the corresponding source $\tilde{\rho}$, is used as a functional variable. This is so since the Wilson lines are most directly expressed in terms of α_a, cf. Eq. (2.22). Thus, from now on, we shall work with the weight function $W_\tau[\alpha]$. Also, to avoid a proliferation of symbols, we shall use the same notations as before, namely σ_x^a and χ_{xy}^{ab}, for the coefficients in the corresponding RGE; but from now on, these coefficients will be understood as functionals of α_a.

We are now prepared to present the RGE. It reads [34]

$$\frac{\partial W_\tau[\alpha]}{\partial \tau} = \frac{1}{2} \int_{\boldsymbol{x}, \boldsymbol{y}} \frac{\delta^2}{\delta \alpha_\tau^a(\boldsymbol{x})\delta\alpha_\tau^b(\boldsymbol{y})}\left[W_\tau \chi_{xy}^{ab}\right] - \int_{\boldsymbol{x}} \frac{\delta}{\delta\alpha_\tau^a(\boldsymbol{x})}\left[W_\tau \sigma_x^a\right], \tag{3.26}$$

where:

$$\chi^{ab}(\boldsymbol{x}, \boldsymbol{y}) = \frac{1}{\pi} \int \frac{d^2\boldsymbol{z}}{(2\pi)^2}\,\mathcal{K}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})\,\left(1 + \widetilde{V}_{\boldsymbol{x}}^\dagger \widetilde{V}_{\boldsymbol{y}} - \widetilde{V}_{\boldsymbol{x}}^\dagger \widetilde{V}_{\boldsymbol{z}} - \widetilde{V}_{\boldsymbol{z}}^\dagger \widetilde{V}_{\boldsymbol{y}}\right)^{ab}, \tag{3.27}$$

and

$$\sigma^a(\boldsymbol{x}) = \frac{ig}{2\pi} \int \frac{d^2\boldsymbol{z}}{(2\pi)^2}\,\frac{1}{(\boldsymbol{x} - \boldsymbol{z})^2}\,\mathrm{Tr}\left(T^a \widetilde{V}_{\boldsymbol{x}}^\dagger \widetilde{V}_{\boldsymbol{z}}\right). \tag{3.28}$$

Eq. (3.27) involves the following transverse kernel:

$$\mathcal{K}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \equiv \frac{(x^i - z^i)(y^i - z^i)}{(\boldsymbol{x} - \boldsymbol{z})^2(\boldsymbol{z} - \boldsymbol{y})^2}, \tag{3.29}$$

The above equations involves the Wilson lines $\widetilde{V}_{\boldsymbol{x}}^\dagger \equiv \widetilde{V}^\dagger(\boldsymbol{x})$ and $\widetilde{V}_{\boldsymbol{x}} \equiv \widetilde{V}(\boldsymbol{x})$ with, e.g.,

$$\widetilde{V}^\dagger(\boldsymbol{x}) = \mathrm{P}\exp\left(ig \int dx^- \alpha^a(x^-, \boldsymbol{x})T^a\right) \equiv \mathrm{P}\exp\left(ig \int dy\, \alpha_\mathrm{y}(\boldsymbol{x})\right), \tag{3.30}$$

(we use a tilde to denote the adjoint representation) where the support of the integration over y is effectively cut off at $\mathrm{y} = \tau$ because of structure (3.19) of the weight function. Note that, since the information about the longitudinal support of $\alpha_\mathrm{y}(\boldsymbol{x})$ is encoded in the weight function, there is no need to carry the indices τ on the coefficients σ and χ.

According to Eq. (3.26), we need to know the action of the functional derivatives on the coefficients σ and χ, and thus on the Wilson lines (3.30). Since the derivatives act on the color field created in the last step of the evolution, and which is therefore located in the highest bin in y (namely, in the bin $(\tau, \tau + d\tau)$), the action of the derivatives on the Wilson lines reads as follows:

$$\frac{\delta\widetilde{V}_{\boldsymbol{x}}^\dagger}{\delta\alpha_\tau^a(\boldsymbol{y})} = ig\delta^{(2)}(\boldsymbol{x} - \boldsymbol{y})T^a\widetilde{V}_{\boldsymbol{x}}^\dagger, \quad \frac{\delta\widetilde{V}_{\boldsymbol{x}}}{\delta\alpha_\tau^a(\boldsymbol{y})} = -ig\delta^{(2)}(\boldsymbol{x} - \boldsymbol{y})\widetilde{V}_{\boldsymbol{x}}T^a. \tag{3.31}$$

As a simple application of these formulae, note the following relation between the coefficients σ_x^a and χ_{xy}^{ab}, whose deep meaning should become clear later:

$$\frac{1}{2}\int d^2\boldsymbol{y}\,\frac{\delta\chi^{ab}(\boldsymbol{x},\boldsymbol{y})}{\delta\alpha_\tau^b(\boldsymbol{y})} \;=\; \sigma^a(\boldsymbol{x})\,. \tag{3.32}$$

It is easy to prove this relation by using Eq. (3.31) to act with $\delta/\delta\alpha_\tau^b(\boldsymbol{y})$ on $\chi_{ab}(\boldsymbol{x},\boldsymbol{y})$, Eq. (3.27). This yields, e.g., (with $\delta_{\boldsymbol{xy}}\equiv\delta^{(2)}(\boldsymbol{x}-\boldsymbol{y})$)

$$\frac{\delta}{\delta\alpha_\tau^b(\boldsymbol{y})}\,(\widetilde{V}_{\boldsymbol{x}}^\dagger\widetilde{V}_{\boldsymbol{y}})^{ab} = \frac{\delta\widetilde{V}_{\boldsymbol{x}}^{\dagger\,ac}}{\delta\alpha_\tau^b(\boldsymbol{y})}\,\widetilde{V}_{\boldsymbol{y}}^{cb} \;+\; \widetilde{V}_{\boldsymbol{x}}^{\dagger\,ac}\,\frac{\delta\widetilde{V}_{\boldsymbol{y}}^{cb}}{\delta\alpha_\tau^b(\boldsymbol{y})} \tag{3.33}$$
$$= ig\delta_{\boldsymbol{xy}}\big(T^b\widetilde{V}_{\boldsymbol{y}}^\dagger\big)_{ac}\widetilde{V}_{\boldsymbol{y}}^{cb} - ig\delta^{(2)}(0)\widetilde{V}_{\boldsymbol{x}}^{\dagger\,ac}\big(\widetilde{V}_{\boldsymbol{y}}T^b\big)_{cb} \;=\; 0,$$

where both terms in the second line — including the potentially singular one — vanish because of the antisymmetry of the colour group generators in the adjoint representation (e.g., $(T^b)_{ab}=0$). The only nonvanishing contribution is

$$-\frac{\delta}{\delta\alpha_\tau^b(\boldsymbol{y})}\,(\widetilde{V}_{\boldsymbol{x}}^\dagger\widetilde{V}_{\boldsymbol{z}})^{ab} \;=\; -ig\delta_{\boldsymbol{xy}}\big(T^b\widetilde{V}_{\boldsymbol{x}}^\dagger\widetilde{V}_{\boldsymbol{z}}\big)_{ab} \;=\; ig\delta_{\boldsymbol{xy}}\mathrm{Tr}\big(T^a\widetilde{V}_{\boldsymbol{x}}^\dagger\widetilde{V}_{\boldsymbol{z}}\big), \tag{3.34}$$

which reproduces indeed Eq. (3.28) after integration over $\boldsymbol{y}$, since $\mathcal{K}(\boldsymbol{x},\boldsymbol{x},\boldsymbol{z})=1/(\boldsymbol{x}-\boldsymbol{z})^2$ (see Eq. (3.29)).

Further properties of the RGE (3.26), also known as the *JIMWLK equation*, will be explored in the nextcoming sections.

3.3 Stochastic interpretation: From Fokker–Planck to Langevin equation

The renormalization group analysis in Sect. 3.1 is already suggestive in that it depicts the CGC evolution towards small x as a *stochastic process* in which the rapidity τ plays the role of an *evolution time*. It is the purpose of this section to make this interpretation more precise and thus clarify the stochastic nature of the evolution. This will give us an intuitive picture of the evolution, as a *random walk on an* SU(3) *manifold*, and will also allow us to draw important conclusions about the behaviour of the gluon distribution and of the scattering amplitudes in the high–energy limit.

Note first that, as a consequence of Eq. (3.32), the RGE (3.26) can be rewritten in the *Hamiltonian form*

$$\frac{\partial}{\partial\tau}W_\tau[\alpha] = -HW_\tau[\alpha] \equiv \frac{1}{2}\int_{\boldsymbol{xy}}\frac{\delta}{\delta\alpha_\tau^a(\boldsymbol{x})}\chi^{ab}(\boldsymbol{x},\boldsymbol{y})\frac{\delta}{\delta\alpha_\tau^b(\boldsymbol{y})}W_\tau[\alpha], \tag{3.35}$$

in which the 'virtual correction' σ is not manifest anymore (of course, this would be generated when commuting one functional derivative through χ). The Hamiltonian in the above equation is Hermitian (since η_{xy}^{ab} is real and symmetric), and also positive semi–definite, since it can be written in the factorized form

$$H = \int \frac{d^2 z_\perp}{2\pi} \, J_a^i(z_\perp) J_a^i(z_\perp),$$

$$J_a^i(z_\perp) \equiv \int \frac{d^2 \boldsymbol{x}}{2\pi} \, \frac{z^i - x^i}{(z_\perp - \boldsymbol{x})^2} \, (1 - \widetilde{V}_z^\dagger \widetilde{V}_x)_{ab} \, \frac{i\delta}{\delta \alpha_\tau^b(\boldsymbol{x})} \,, \qquad (3.36)$$

where the "current" $J_a^i(z_\perp)$ is itself Hermitian.

Eq. (3.35) may be formally seen as a 'Schrödinger equation in imaginary time', but this is properly a diffusion equation for a probability distribution, that is, a *Fokker–Planck equation* (see, e.g., [154–156]). It is easy to check that the structure of this equation is indeed consistent with the probabilistic interpretation of the weight function: (i) Its r.h.s. is a total derivative w.r.t. α; thus, the evolution preserves the correct normalization (2.2) of W_τ. (ii) The properties of the JIMWLK Hamiltonian mentioned above guarantee that the weight function remains positive semi–definite in the course of the evolution.

3.3.1 Brief review of the Brownian motion

At this stage, it becomes useful to recall a few facts about more standard versions of the Fokker–Planck equation, and their probabilistic interpretation. Let us consider the simplest example of a stochastic process, namely the Brownian motion of a small particle in a viscous liquid and in the presence of some external force, like gravitation. The particle is so small that it can feel the collisions with the molecules in the liquid; after each such a collision, the velocity of the particle changes randomly. And the liquid is so viscous that, after each collision, the particle enters immediately a constant velocity regime in which the friction force $\propto v^\alpha$ (with v^α, $\alpha = 1, 2, 3$ the velocity of the particle) is equilibrated by the random force due to collisions together with the external force $f^\alpha(x)$. In these conditions, the particle executes a random walk whose description is necessary statistical. There exist two different mathematical ways to describe this process:

i) The Langevin equation

The particle position x^α obeys the stochastic equation:

$$\dot{x}^\alpha(\tau) \,=\, f^\alpha(\boldsymbol{x}) + e^{\alpha\beta}(\boldsymbol{x})\, \nu^\beta(\tau), \qquad (3.37)$$

where τ is the time, the dot denotes time derivative, and the last term in the r.h.s. is the random change in the velocity of the particle at time τ due to collisions in the liquid. In turn, this term is the product of a smooth function $e^{\alpha\beta}(\boldsymbol{x})$ which characterizes the medium and the 'Gaussian white noise' $\nu^\alpha(\tau)$, a random variable with zero expectation value and local 2–point function (below, the brackets denote the average over the noise) :

$$\langle \nu^\alpha(\tau)\nu^\beta(\tau') \rangle \,=\, \delta^{\alpha\beta}\delta(\tau - \tau'). \qquad (3.38)$$

Since the function $e^{\alpha\beta}(\boldsymbol{x})$ which controls the strength of the noise is dependent upon the random trajectory $\boldsymbol{x}(\tau)$, the noise in Eq. (3.37) is called *multiplicative.*

Strictly speaking, Eq. (3.37) is formal as written — the trajectory $\boldsymbol{x}(\tau)$ of the random walk is not differentiable (see below) — and gets meaningful only with a discretization prescription. Here, we shall focus on the following discretization:

$$\frac{\boldsymbol{x}_i - \boldsymbol{x}_{i-1}}{\epsilon} \;=\; \boldsymbol{f}_{i-1} + e_{i-1}\,\boldsymbol{\nu}_i, \tag{3.39}$$

where ϵ is the length of the time step, $\boldsymbol{x}_i \equiv \boldsymbol{x}(\tau_i)$ with $\tau_i = i\epsilon$, $i = 0,1,\cdots,n$, and $\boldsymbol{f}_i \equiv \boldsymbol{f}(\boldsymbol{x}_i)$, $e_i \equiv e(\boldsymbol{x}_i)$. We use vector notations: $\boldsymbol{x}$, $\boldsymbol{f}$ and $\boldsymbol{\nu}$ are 3–dimensional vectors and e is a 3×3 matrix. The $\boldsymbol{\nu}_i$'s are Gaussian random variables:

$$\langle \nu_i^\alpha \rangle \;=\; 0, \qquad \langle \nu_i^\alpha \nu_j^\beta \rangle \;=\; \frac{1}{\epsilon}\,\delta^{\alpha\beta}\delta_{ij}. \tag{3.40}$$

In other words, the probability law for the noise variables $\boldsymbol{\nu}_i$ is a normalized Gaussian distribution:

$$d\mathcal{P}(\boldsymbol{\nu}_i) \equiv \left(\frac{\epsilon}{2\pi}\right)^{3/2} \mathrm{e}^{-\frac{\epsilon}{2}\,\boldsymbol{\nu}_i\cdot\boldsymbol{\nu}_i}\, d^3\boldsymbol{\nu}_i\,, \tag{3.41}$$

the same for each time step, and for n time steps it is the product of n such distributions. Eqs. (3.40)–(3.41) imply that the noise term scales like $\nu_i^\alpha \sim 1/\sqrt{\epsilon}$, and therefore $|\boldsymbol{x}_i - \boldsymbol{x}_{i-1}| \sim \sqrt{\epsilon}$, which confirms that the path is not differentiable. Because of that, the results are dependent upon our specific choice of a discretization — Eq. (3.39) would lead to different physical predictions if we were to replace $\boldsymbol{x}_{i-1}$ by $\boldsymbol{x}_i$ or $\bar{\boldsymbol{x}}_i \equiv (\boldsymbol{x}_i + \boldsymbol{x}_{i-1})/2$ as an argument in the functions $\boldsymbol{f}(\boldsymbol{x})$ and $e(\boldsymbol{x})$ —, and our choice in Eq. (3.39) is actually motivated by the analogy with the QCD problem that we are eventually interested in: For *fixed* $\boldsymbol{x}_{i-1}$, the *velocity*

$$\boldsymbol{v}_i \equiv \frac{\boldsymbol{x}_i - \boldsymbol{x}_{i-1}}{\epsilon} \;=\; \boldsymbol{f}_{i-1} + e_{i-1}\boldsymbol{\nu}_i\,, \tag{3.42}$$

at the ith step is a Gaussian random variable whose correlations:

$$\langle v_i^\alpha \rangle = f_{i-1}^\alpha, \qquad \langle v_i^\alpha v_i^\beta \rangle = \frac{1}{\epsilon}\, e_{i-1}^{\alpha\gamma} e_{i-1}^{\beta\gamma} \equiv \frac{1}{\epsilon}\, D_{i-1}^{\alpha\beta}, \tag{3.43}$$

depend solely upon the state of the particle after the previous $i-1$ steps, so like the correlations of the color charged $\delta\rho_\tau$ induced by the quantum evolution over an additional step in rapidity (cf. Eqs. (3.13)–(3.14)). In the statistical physics literature [154–156], the discretization prescription used in Eqs. (3.39) and (3.43) and which appears naturally in a renormalization group analysis, is known as the *Ito prescription*.

The coordinate $\boldsymbol{x}_n$ of the particle after n time step is clearly a random variable. The average of any function of $\boldsymbol{x}_n$, say $O(\boldsymbol{x}_n)$, can be obtained by first solving Eq. (3.39) for a given realization $\{\boldsymbol{\nu}_1, \boldsymbol{\nu}_2, \cdots, \boldsymbol{\nu}_n\}$ of the random variables, and then averaging over all such realizations:

$$\langle O(\boldsymbol{x}) \rangle_\tau = \langle O(\boldsymbol{x}_n) \rangle \equiv \int \prod_{i=1}^{n} d\mathcal{P}(\boldsymbol{\nu}_i)\, O(\boldsymbol{x}_n[\boldsymbol{\nu}])\,, \tag{3.44}$$

where $\tau = n\epsilon$, and $\boldsymbol{x}_n[\boldsymbol{\nu}]$ is the solution of Eq. (3.39) for n time steps, and depends on all the $\boldsymbol{\nu}_i$'s with $i \le n$.

ii) The Fokker–Planck equation

Alternatively, and equivalently, one may first determine the (conditional) *probability density* $P(\boldsymbol{x}, \tau) \equiv P(\boldsymbol{x}, \tau | \boldsymbol{x}_0, \tau_0)$ to find the particle at point $\boldsymbol{x}$ at time $\tau = n\epsilon$ knowing that it was at $\boldsymbol{x}_0$ at time $\tau_0 = 0$. This is given by

$$P(\boldsymbol{x}, \tau) \equiv \int \prod_{i=1}^{n} d\mathcal{P}(\boldsymbol{\nu}_i) \, \delta^D\big(\boldsymbol{x} - \boldsymbol{x}_n[\boldsymbol{\nu}]\big), \tag{3.45}$$

from which the average of $O(\boldsymbol{x})$ may then be obtained as

$$\langle O(\boldsymbol{x}) \rangle_\tau = \int d^3\boldsymbol{x} \, O(\boldsymbol{x}) \, P(\boldsymbol{x}, \tau). \tag{3.46}$$

The probability density (3.45) is properly normalized:

$$\int d^3\boldsymbol{x} \, P(\boldsymbol{x}, \tau) = 1, \tag{3.47}$$

and obeys the Chapman–Kolmogorov equation [156] :

$$P(\boldsymbol{x}, \tau | \boldsymbol{x}_0, \tau_0) = \int d^3 x_1 \, P(\boldsymbol{x}, \tau | \boldsymbol{x}_1, \tau_1) \, P(\boldsymbol{x}_1, \tau_1 | \boldsymbol{x}_0, \tau_0), \tag{3.48}$$

(τ_1 obeys $\tau_0 < \tau_1 < \tau$, but otherwise is arbitrary) which is similar to the recurrence formula (3.21) that we have derived previously in QCD.

By using Eq. (3.48) for an infinitesimal time step $\tau - \tau_1 = \epsilon$, together with Eq. (3.45) restricted to such a small step, one can deduce that, in the limit $\epsilon \to 0$, $P(\boldsymbol{x}, \tau)$ obeys a second–order differential equation of the diffusion type (see, e.g., [40] for details on the derivation) :

$$\dot{P}(\boldsymbol{x}, \tau) = \partial^\alpha \Big[\partial^\beta \Big(\frac{1}{2} D^{\alpha\beta}(\boldsymbol{x}) P(\boldsymbol{x}, \tau) \Big) - f^\alpha(\boldsymbol{x}) P(\boldsymbol{x}, \tau) \Big], \tag{3.49}$$

where (cf. Eq. (3.43))

$$D^{\alpha\beta}(\boldsymbol{x}) \equiv e^{\alpha\gamma}(\boldsymbol{x}) e^{\beta\gamma}(\boldsymbol{x}). \tag{3.50}$$

is the (generally, medium dependent and anisotropic) *diffusion tensor*, and is positive–semidefinite by construction. Eq. (3.49) is the most general form of the Fokker–Planck (FP) equation [154–156].

We conclude this brief review of the Brownian motion with a discussion of the solutions to Eq. (3.49) in some simple cases. We assume that the diffusion tensor is point–independent and isotropic, $D^{\alpha\beta}(\boldsymbol{x}) = \delta^{\alpha\beta} D$ with constant D, and consider two cases: (a) free motion ($f^\alpha = 0$) and (b) potential force ($f^\alpha = -\partial^\alpha V$).

(a) If $f^\alpha = 0$, the solution is immediate, and reads (with the initial condition $P(\boldsymbol{x}, 0) = \delta^{(3)}(\boldsymbol{x})$):

$$P(\boldsymbol{x}, \tau) = \frac{1}{(2\pi D\tau)^{3/2}} \exp\left\{ -\frac{\boldsymbol{x}^2}{2D\tau} \right\}. \tag{3.51}$$

For any fixed $\boldsymbol{x}$, this probability density goes smoothly to zero when $\tau \to \infty$ ('*runaway solution*'). Thus, for sufficiently large times, the probability density is quasi–homogeneous in space, and is small everywhere: The particle simply diffuses in all the available space. In particular, the average square displacement

$$\langle \boldsymbol{x}^2 \rangle(\tau) \equiv \int d^3\boldsymbol{x}\, \boldsymbol{x}^2\, P(\boldsymbol{x}, \tau) \,=\, 6D\tau, \tag{3.52}$$

grows indefinitely with τ.

(b) If the motion of the particle is biased by an external potential $V(\boldsymbol{x})$, such that $f^\alpha = -\partial^\alpha V$, then Eq. (3.49) admits the *stationary solution* $P_0(\boldsymbol{x}) \propto \exp[-\beta V(\boldsymbol{x})]$ with $\beta = 2/D$. This equilibrium distribution is a (non–trivial) fixed point for the evolution: Once this is reached, all the correlations become independent of time.

3.3.2 High–energy evolution in QCD as a random walk

A brief comparison between Eqs. (3.26) and (3.49) reveals that the RGE for high–energy evolution in QCD is in fact a *functional Fokker–Planck equation*, with χ playing the role of a 'point'–dependent diffusion tensor (here, the 'point' means, of course, a point in the functional space, i.e., a function), and σ that of a force. Although this interpretation is essentially correct, a subtle difference remains between Eqs. (3.26) and (3.49): the argument α_τ of the functional derivatives in the RGE depends upon the evolution time τ, a feature which is not seen on the FP equation (3.49). As we shall explain now, this difference reflects the fact that the Brownian motion analog of the field α_τ is not the position $\boldsymbol{x}$ of the particle, but rather its velocity $\boldsymbol{v}_n$, with $\tau = n\epsilon$. The position $\boldsymbol{x}$ corresponds instead to the Wilson line built with $\alpha_{\rm y}$, cf. Eq. (3.30).

This can be anticipated from the fact that α_τ, so like $\boldsymbol{v}_n$, is the elementary random ingredient which is brought in by a single step in the evolution, whereas $V \equiv V_\tau$, and also $\boldsymbol{x} \equiv \boldsymbol{x}_\tau$, are the 'integrated' results of this evolution, from τ_0 up to τ. More precisely, we have seen in Sect. 3.1 that, when integrating out gauge field fluctuations in the rapidity strip $(\tau, \tau + d\tau)$, a new layer $d\tau \alpha_\tau \equiv \epsilon \alpha_n$ (with $d\tau \equiv \epsilon$ and $\tau = n\epsilon$) is added on top to the previous field, so that the Wilson lines evolve according to (for a generic representation of the color group)

$$V_n \,=\, V_{n-1}\, {\rm e}^{-{\rm i}\epsilon \alpha_n^a t^a}, \qquad V_n^\dagger \,=\, {\rm e}^{{\rm i}\epsilon \alpha_n^a t^a}\, V_{n-1}^\dagger, \tag{3.53}$$

where we omit transverse coordinates and temporarily set $g = 1$, to simplify writing. By iterating this elementary step, a *discretized path* is generated on the manifold of V fields, namely

$$V_n^\dagger \,=\, {\rm e}^{{\rm i}\epsilon \alpha_n} V_{n-1}^\dagger \,=\, \cdots \,=\, {\rm e}^{{\rm i}\epsilon \alpha_n} {\rm e}^{{\rm i}\epsilon \alpha_{n-1}} \cdots {\rm e}^{{\rm i}\epsilon \alpha_1}\, V_0^\dagger. \tag{3.54}$$

In continuous notations, and with spatial coordinates included, the corresponding path is given by Eq. (3.30). Note that, at each step along this path, $\epsilon \alpha_i$ is an infinitesimal gauge rotation, so α_i^a plays indeed the role of a 'velocity' (here, an element of the SU(3) Lie algebra). Moreover, according to the discussion in Sect. 3.1, α_i^a is a random field, whose stochastic properties are very similar to those of the velocity $\boldsymbol{v}_i$ of the particle performing

the random walk. Indeed, Eqs. (3.13)–(3.14) can be rewritten in terms of α_i as follows:

$$\langle \alpha_i^a(\boldsymbol{x}) \rangle = \sigma_{i-1}^a(\boldsymbol{x})\,, \qquad \langle \alpha_i^a(\boldsymbol{x})\, \alpha_i^b(\boldsymbol{y}) \rangle = \frac{1}{\epsilon}\, \chi_{i-1}^{ab}(\boldsymbol{x}, \boldsymbol{y})\,, \tag{3.55}$$

`fluct-`

where σ_{i-1} and χ_{i-1} are calculated in terms of the Wilson lines obtained at the step $i-1$ (i.e., $V_{i-1}^\dagger$ and V_{i-1}), according to Eqs. (3.27) and (3.28). The similarity with the corresponding equations Eq. (3.43) for $\boldsymbol{v}_i$ is now manifest. In fact, the analogy goes even further: like in Eq. (3.43), the 2–point correlation function χ_{xy}^{ab} can be written in a factorized form, as obvious from Eqs. (3.27):

$$\chi_{\boldsymbol{xy}}^{ab}[V] = \int d^2\boldsymbol{z}\, \mathrm{e}^{ac,l}(\boldsymbol{x}, \boldsymbol{z})\, \mathrm{e}^{bc,l}(\boldsymbol{y}, \boldsymbol{z}), \tag{3.56}$$

`chidef`

where

$$e^{ab,l}(\boldsymbol{x}, \boldsymbol{z}) = \frac{1}{\sqrt{4\pi^3}} \frac{(\boldsymbol{x} - \boldsymbol{z})^l}{(\boldsymbol{x} - \boldsymbol{z})^2} \Big(1 - \widetilde{V}_{\boldsymbol{x}}^\dagger \widetilde{V}_{\boldsymbol{z}}\Big)^{ab} \tag{3.57}$$

`e`

is a matrix in color indices and a vector in transverse coordinates. Thus, the Gaussian stochastic process for α_i^a defined by Eq. (3.55) is equivalent to solving the following Langevin equation (with $e_{i-1}^{ab,l}$ given by Eq. (3.57) with $V \equiv V_{i-1}$):

$$\alpha_i^a(\boldsymbol{x}) = \sigma_{i-1}^a(\boldsymbol{x}) + \int_{\boldsymbol{z}} e_{i-1}^{ab,l}(\boldsymbol{x}, \boldsymbol{z})\, \nu_i^{b,l}(\boldsymbol{z})\,, \tag{3.58}$$

`alpha-`

where $\nu_i^a(\boldsymbol{x})$ is an auxiliary Gaussian white noise:

$$\langle \nu_i^{a,l}(\boldsymbol{x}) \rangle = 0\,, \qquad \langle \nu_i^{a,l}(\boldsymbol{x})\nu_j^{b,k}(\boldsymbol{y}) \rangle = \frac{1}{\epsilon}\, \delta_{ij}\delta^{ab}\delta^{lk}\delta^{(2)}(\boldsymbol{x} - \boldsymbol{y}), \tag{3.59}$$

`eq:no`

By averaging over the noise the solution of Eq. (3.58) one indeed recovers the expectation value and correlator of the random variables $\alpha_i^a(\boldsymbol{x})$ as given by Eq. (3.55).

We conclude that the quantum evolution in QCD with increasing rapidity amounts to a *random walk on the two–dimensional* SU(3) *manifold spanned by the Wilson lines*. But in contrast to what happens in the simple example of the Brownian motion in a flat space, the relation between the "coordinate" V_i and the "velocity" α_i is here non trivial, because of the non-trivial geometry of the group manifold. However, as long as we characterize the random walk by small displacements in the tangent space (the Lie algebra), the special geometry of the group manifold plays essentially no role (it only enters through the dependence of σ and χ on the Wilson lines). Eqs. (3.53)–(3.59) represent the Langevin description of this stochastic process, while the RGE (3.26) is a rather unusual version of the Fokker–Planck equation, namely its *velocity representation* [40]. The corresponding equation for the ordinary Brownian motion had never been written down; inspired by the QCD problem, we have done so in Ref. [40]. It is interesting to note that the probability distribution $W_\tau[\boldsymbol{v}]$ in the velocity representation (the analog of our QCD weight function $W_\tau[\alpha]$) remains a *functional* even for the ordinary random walk, because it characterizes a given *path* (as specified by the velocity function $\boldsymbol{v}_\tau$), rather than merely the final point $\boldsymbol{x}$ of this path. Thus, the relation between $W_\tau[\boldsymbol{v}]$ and the more standard probability $P(\boldsymbol{x}, \tau)$

introduced before (cf. Eq. (3.45)) reads

$$P(\boldsymbol{x}, \tau) = \int [\mathrm{d}\boldsymbol{v}] \, \delta^{(D)}(\boldsymbol{x} - \boldsymbol{x}[\boldsymbol{v}]) W_\tau[\boldsymbol{v}] \qquad (3.60)$$

<div style="text-align:right">PW</div>

with (compare to Eq. (3.30))

$$\boldsymbol{x}[\boldsymbol{v}] = \boldsymbol{x}_0 + \int_0^\infty \mathrm{d}\tau' \, \boldsymbol{v}(\tau'). \qquad (3.61)$$

<div style="text-align:right">xv0</div>

As in the corresponding QCD problem, the physical paths relevant for computing $P(\boldsymbol{x}, \tau)$ terminate at τ, but this information is encoded in the weight function $W_\tau[\boldsymbol{v}]$. Then, $W_\tau[\boldsymbol{v}]$ obeys an equation similar to the RGE (3.26).

Conversely, the analog of the standard version, Eq. (3.49), of the Fokker–Planck equation exists also in QCD [35]. It applies to the following quantity (compare to Eq. (3.60); below, $\delta_G(U, V)$ is the group-invariant δ–function [157])

$$Z_\tau[V] \;=\; \int [\mathrm{d}\alpha] \, \delta_G\!\left(V^\dagger, \, \mathrm{P}\, \mathrm{e}^{ig \int \mathrm{dy}\, \alpha_\mathrm{y}}\right) W_\tau[\alpha]\,, \qquad (3.62)$$

<div style="text-align:right">ZW</div>

which represents the probability density that a given SU(3) group element V (actually, a group–valued field in the transverse space) be produced by the quantum evolution up to rapidity τ. Like $P(\boldsymbol{x}, \tau)$, the probability distribution (3.62) refers solely to the final 'point' of the evolution. The contributions of the various paths leading to this point are summed over in Eqs. (3.60) and (3.62). Then, $Z_\tau[V]$ obeys a RGE obtained from Eq. (3.26) after replacing the functional derivatives w.r.t. α_τ by (functional) *Lie derivatives* [157] w.r.t. the Wilson lines. This representation of the high energy evolution has been originally proposed by Weigert [35] independently of the CGC formalism: rather, Weigert managed to show that the infinite hierarchy of equations for scattering amplitudes originally derived by Balitsky [39], and that we shall discuss in the next section, can be compactly summarized as a single functional equation for the probability distribution of the Wilson lines (the scattering operators at high energy). In this description, the gauge field α_y is not present any longer, in the same way as the velocity field $\boldsymbol{v}(\tau)$ is not visible in the FP equation (3.49). As we shall see in the next sections, this representation is well suited to study the evolution of observables built with Wilson lines (so like the scattering amplitudes), but is perhaps less convenient for perturbative studies which require the expansion of the Wilson lines.

Let us finally consider a general implication of the RGE for the high energy limit of the effective theory (and thus of QCD). So far, the special relation (3.32) between the coefficients σ and χ in the RGE has not played any role: The identification with the FP equation and the probabilistic interpretation discussed above hold for generic σ and χ. In particular, we have seen that σ^a plays the role of a 'color force', and, moreover, according to the discussion at the end of Sect. 3.3.1, the presence of such a force may imply non–trivial 'fixed points' (i.e., stationary distributions) in the evolution. In the context of QCD, such a fixed point would mean that, at sufficiently high energy, all the correlations of α become independent of the energy. It is tantalizing to consider this scenario as the solution to the unitarity problem in QCD at high energy. As we explain now, the evolution

described by Eq. (3.26) solves indeed the unitarity problem, but in a different way, which is at the same time more subtle and simpler than the above scenario:

Note first that, in the presence of a point–dependent diffusion tensor, the derivatives acting on $D^{\alpha\beta}(\boldsymbol{x})$ in Eq. (3.49) generate terms which play a similar role as the 'external force'. In particular, one can show that when considering *free Brownian motion* (no external potential) on a *curved manifold,* an equation similar to Eq. (3.49) arises, in which however f^α and $D^{\alpha\beta}$ are related by the analog of Eq. (3.32) [155] (in that case, the diffusion tensor also involves the metric on the manifold). Coming back to QCD, the previous remark shows that the RGE (3.26) describes a *free random walk on a functional manifold with a complicated metric* (as characterized by χ). In view of this, the Hamiltonian in Eq. (3.35) is a purely kinetic operator, which describes diffusion on the group manifold. Thus, we do not expect non–trivial fixed points for this evolution, just runaway solutions [34], and this will be explicitly seen below (in Sect. 4.1) : The correlations of α (in particular, the gluon distribution) keep growing with the energy — so, strictly speaking, there is *no complete saturation at high energy* —, but for large s this growth is only linear in τ (so like for the ordinary Brownian motion, cf. Eq. (3.52)), that is, logarithmic in s [37, 38]. On the other hand, since the gauge field α enters as a phase in the Wilson lines, and at high energy this phase is large and has a typical variation scale $1/Q_s$, it follows that the Wilson lines are rapidly oscillating when probed over transverse distances larger than $1/Q_s$, and thus *their contribution vanish in any gauge–invariant correlator* [38]. In that average sense, $V = 0$ is an asymptotic fixed point for the evolution [35]. As we shall see in what follows, this implies that the scattering amplitudes computed in the CGC formalism respect unitarity, and saturate the black disk limit for sufficiently high energies.

3.4   Evolution equations for observables

SECT_OBS

As shown via explicit examples in the previous sections, the interesting observables in QCD at high energy can be always expressed as operators built with the A^+ component of the gauge field in the target (by working in a suitable gauge). Within the effective theory for the CGC, A_a^+ is identified with the classical random field α_a in the Coulomb gauge. Then, for any interesting operator, say $\mathcal{O}[\alpha]$, the corresponding expectation value is obtained by averaging over α, as in Eq. (2.7) :

$$\langle \mathcal{O} \rangle_\tau \equiv \int \mathcal{D}\alpha\, W_\tau[\alpha]\, \mathcal{O}[\alpha] \,. \tag{3.63}$$

The energy dependence in this equation is entirely encoded in the weight function $W_\tau[\alpha]$ for the classical field. Thus, by taking a τ-derivative in Eq. (3.63), using the Hamiltonian version (3.35) of the RGE, and performing some integrations by parts in the functional integral, one can deduce

$$\frac{\partial}{\partial \tau} \langle \mathcal{O} \rangle_\tau = - \int \mathrm{D}[\alpha]\, W_\tau[\alpha]\, H^\dagger\!\left[\alpha, \frac{\delta}{i\delta\alpha}\right] \mathcal{O}[\alpha] \,. \tag{3.64}$$

DOgen

This can be interpreted as describing the evolution of the scattering operator $\mathcal{O}_\tau[\alpha]$, with 'Hamiltonian' $H^\dagger$:

92

$$\frac{\partial \mathcal{O}_\tau}{\partial \tau} = -H^\dagger \left[\alpha, \frac{\delta}{i\delta\alpha} \right] \mathcal{O}_\tau. \tag{3.65}$$

(The JIMWLK Hamiltonian (3.36) is Hermitian, but we keep the Hermitian conjugation sign explicitly, to anticipate for more general situations.) Both points of view, somewhat reminiscent of, respectively, the Schrödinger and the Heisenberg pictures of quantum mechanics, will be used in the following discussion (although we shall refrain from introducing explicitly rapidity dependent operators). In the Schrödinger picture, one puts emphasis on the evolution of the state vector, whose role is played here by the weight functional $W_\tau[\alpha]$. In the Heisenberg picture, the state vector is a constant reference vector involved in the calculation of all expectation values, and one puts all the evolution in the operators, here $\mathcal{O}_\tau[\alpha]$. The Schrödinger picture corresponds to evolution equations which aim at providing a detailed microscopic description of the color field in the target, together with its complicated correlations. This is what the JIMWLK equation does. The Heisenberg picture rather describes how the 'test particles' (i.e., the projectile probing the CGC) get dressed by color field fluctuations as they are boosted to higher rapidities. In this approach, the complicated color correlations in the target wavefunction are not immediately visible, and indeed the resulting equation of motion are established somewhat more easily. This second approach is essentially the one used by Balitsky to obtain his hierarchy of equations [39] that we shall recover in Sect. 3.5 from the perspective of target evolution.

In the case of the JIMWLK evolution, Eqs. (3.35) and (3.64) imply :

$$\frac{\partial}{\partial \tau} \langle \mathcal{O} \rangle_\tau = \left\langle \frac{1}{2} \int_{\boldsymbol{xy}} \frac{\delta}{\delta\alpha_\tau^a(\boldsymbol{x})} \chi^{ab}(\boldsymbol{x}, \boldsymbol{y}) \frac{\delta}{\delta\alpha_\tau^b(\boldsymbol{y})} \mathcal{O} \right\rangle_\tau. \tag{3.66}$$

In what follows, we shall investigate the structure of Eq. (3.66) in more detail, starting with an analysis of its potential singularities. To that aim, we shall use the expression (3.27) for χ, that we conveniently rewrite here as

$$\chi^{ab}(\boldsymbol{x}, \boldsymbol{y}) = \frac{1}{\pi} \int \frac{d^2\boldsymbol{z}}{(2\pi)^2} \mathcal{K}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \left(1 - \widetilde{V}_{\boldsymbol{x}}^\dagger \widetilde{V}_{\boldsymbol{z}} \right)^{fa} \left(1 - \widetilde{V}_{\boldsymbol{z}}^\dagger \widetilde{V}_{\boldsymbol{y}} \right)^{fb}. \tag{3.67}$$

As manifest on Eq. (3.29), the transverse kernel $\mathcal{K}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ has simple poles at $\boldsymbol{z} = \boldsymbol{x}$ and $\boldsymbol{z} = \boldsymbol{y}$, which could give rise to short–range (or 'ultraviolet') divergences in the integration over $\boldsymbol{z}$. However, a brief inspection of Eq. (3.67) reveals that these poles are in fact innocuous since they have zero residue; e.g., $1 - \widetilde{V}_{\boldsymbol{x}}^\dagger \widetilde{V}_{\boldsymbol{z}} \to 0$ when $\boldsymbol{z} \to \boldsymbol{x}$.

The cancelation of the long–range, or 'infrared', singularities, on the other hand, is less obvious — at large distances $z \gg x, y$, the transverse kernel (3.29) decays only like $\mathcal{K}_{\boldsymbol{xyz}} \sim 1/z^2$, so Eq. (3.66) may develop logarithmic divergences ($\sim \int d^2\boldsymbol{z}/z^2$ for $\boldsymbol{z} \to \infty$) —, and not so general: it only holds for *gauge–invariant* observables, as is often the case in QCD. To understand how this works out, it is instructive to consider first the *weak field* limit of Eq. (3.66), corresponding to the dilute regime in which the gluon occupation factor is relatively small ($\ll 1/\alpha$). In this regime, the classical field is typically weak, $g\alpha \ll 1$, so one can expand the Wilson lines (within the Hamiltonian and the various observables) in powers of α. As we shall see, when restricted to lowest nontrivial order, this expansion generates the *BFKL approximation*.

3.4.1 The weak field approximation: BFKL equation

The weak–field expansion of a Wilson line reads (in a generic representation of $\mathrm{SU}(N_c)$)

$$
\begin{aligned}
V_{\boldsymbol{x}}^{\dagger}[\alpha] = 1 \,&+\, \mathrm{i}g \int dx^{-}\, \alpha^{a}(x^{-}, \boldsymbol{x}) t^{a} \\
&- \frac{g^{2}}{2} \int dx^{-} \int dy^{-}\, \alpha^{a}(x^{-}, \boldsymbol{x}) \alpha^{b}(y^{-}, \boldsymbol{x}) \Big[\theta(x^{-} - y^{-}) t^{a} t^{b} + \theta(y^{-} - x^{-}) t^{b} t^{a}\Big] \\
&+\, \cdots\, .
\end{aligned}
\tag{3.68}
$$

To compute the kernel χ in Eq. (3.67) to lowest order, i.e., to order $(g\alpha)^2$, it is sufficient to restrict ourselves to the term linear in α in the above expansion. Indeed, to the accuracy of interest we simply need:

$$
(1 - \tilde{V}_{z}^{\dagger}\tilde{V}_{x})^{fa} \approx \mathrm{i}g(\alpha^{c}(\boldsymbol{x}) - \alpha^{c}(\boldsymbol{z}))(T^{c})_{fa},
\tag{3.69}
$$

where

$$
\alpha^{a}(\boldsymbol{x}) \equiv \int dx^{-}\, \alpha^{a}(x^{-}, \boldsymbol{x}) \equiv \alpha_{\boldsymbol{x}}^{a}
\tag{3.70}
$$

is the effective color field in the transverse plane, as obtained after integrating over the longitudinal profile of the hadron. Then, Eq. (3.66) reduces to

$$
\begin{aligned}
\frac{\partial}{\partial\tau}\langle\mathcal{O}\rangle_{\tau} = -\langle H_{\mathrm{BFKL}}\mathcal{O}\rangle_{\tau} &\equiv \frac{g^{2}}{16\pi^{3}} \int_{\boldsymbol{xyz}} \mathcal{K}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \\
&\times \left\langle f^{acf}f^{bfd}(\alpha_{\boldsymbol{x}} - \alpha_{\boldsymbol{z}})^{a}\frac{\delta}{\delta\alpha_{\boldsymbol{x}}^{c}}(\alpha_{\boldsymbol{y}} - \alpha_{\boldsymbol{z}})^{b}\frac{\delta}{\delta\alpha_{\boldsymbol{y}}^{d}}\mathcal{O}\right\rangle_{\tau}.
\end{aligned}
\tag{3.71}
$$

We have anticipated here that, within the same approximation, any interesting observable $\mathcal{O}[\alpha]$ reduces to a function of the 2–dimensional field $\alpha^{a}(\boldsymbol{x})$ alone, so that the original functional derivatives in Eq. (3.66) can be replaced by $\delta/\delta\alpha_{\boldsymbol{x}}^{a}$. We see that, within the weak field, or BFKL, approximation, *the longitudinal structure of the color field becomes irrelevant.* This reflects the fact that, in this regime, the multiple scattering can be neglected (cf. Sect. 3.1).

In order for $\mathcal{O}[\alpha]$ to represent a physical observable, it must be *gauge–invariant*. As a simple example, consider the elastic scattering between the CGC and an external 'color dipole' (a quark–antiquark pair in a colorless state). In Sect. 1.2.3 we have seen that the corresponding scattering amplitude reads

$$
\langle T(\boldsymbol{x}, \boldsymbol{y})\rangle_{\tau} = 1 - \frac{1}{N_c}\Big\langle \mathrm{tr}(V_{\boldsymbol{x}}^{\dagger}V_{\boldsymbol{y}})\Big\rangle_{\tau},
\tag{3.72}
$$

which is gauge invariant indeed. The weak field approximation of this amplitude is obtained after expanding each of the Wilson lines to second order in α (the linear terms vanish after taking the color trace). Note that, to this order, the x^{-}–ordering of the color matrices starts to play a role in Eq. (3.68). Still, this ordering is irrelevant for the computation of the dipole amplitude to lowest order, because of the symmetry of the color trace: $\mathrm{tr}(t^{a}t^{b}) = \frac{1}{2}\delta^{ab} = \mathrm{tr}(t^{b}t^{a})$. Namely, one finds:

$$\langle T(\boldsymbol{x}, \boldsymbol{y}) \rangle_\tau \simeq \frac{g^2}{4N_c} \langle \alpha_{\boldsymbol{x}}^a \alpha_{\boldsymbol{x}}^a + \alpha_{\boldsymbol{y}}^a \alpha_{\boldsymbol{y}}^a - 2\alpha_{\boldsymbol{x}}^a \alpha_{\boldsymbol{y}}^a \rangle_\tau = \frac{g^2}{4N_c} \langle (\alpha_{\boldsymbol{x}}^a - \alpha_{\boldsymbol{y}}^a)^2 \rangle_\tau \,, \tag{3.73}$$

which involves only the integrated field $\alpha^a(\boldsymbol{x})$, as anticipated. We are thus led to study the evolution of the 2–point function $\langle \alpha_{\boldsymbol{x}}^a \alpha_{\boldsymbol{y}}^a \rangle_\tau$ in the dilute regime. By using $\mathcal{O} = \alpha_{\boldsymbol{x}}^a \alpha_{\boldsymbol{y}}^a$ in Eq. (3.71), one immediately obtains:

$$\frac{\partial}{\partial \tau} \left\langle \alpha_{\boldsymbol{x}}^a \alpha_{\boldsymbol{y}}^a \right\rangle_\tau = \frac{g^2 N_c}{2\pi} \int_{\boldsymbol{z}} \left\langle 2\mathcal{K}_{\boldsymbol{xyz}} \left(\alpha_{\boldsymbol{x}}^a - \alpha_{\boldsymbol{z}}^a \right) \left(\alpha_{\boldsymbol{y}}^a - \alpha_{\boldsymbol{z}}^a \right) \right. $$
$$\left. - \mathcal{K}_{\boldsymbol{xxz}} \alpha_{\boldsymbol{y}}^a \left(\alpha_{\boldsymbol{x}}^a - \alpha_{\boldsymbol{z}}^a \right) - \mathcal{K}_{\boldsymbol{yyz}} \alpha_{\boldsymbol{x}}^a \left(\alpha_{\boldsymbol{y}}^a - \alpha_{\boldsymbol{z}}^a \right) \right\rangle_\tau , \tag{3.74}$$

where $\mathcal{K}_{\boldsymbol{xyz}} \equiv \mathcal{K}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ and we have used $(T^c T^d)_{ab} \delta^{cd} = N_c \delta^{ab}$.

Consider the convergence properties of the above integral over $\boldsymbol{z}$. As expected from the general discussion, there is no singularity at short distances : the three terms within the integrand are separately ultraviolet finite.

To study the large distance behavior ($z \gg x, y$), it is convenient to group separately the terms involving $\alpha_{\boldsymbol{z}}^a$, and those without it. The latter combine into $\mathcal{M}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \langle \alpha_{\boldsymbol{x}}^a \alpha_{\boldsymbol{y}}^a \rangle_\tau$, where (cf. Eq. (3.29)) :

$$\mathcal{M}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \equiv \frac{(\boldsymbol{x} - \boldsymbol{y})^2}{(\boldsymbol{x} - \boldsymbol{z})^2 (\boldsymbol{z} - \boldsymbol{y})^2} = \mathcal{K}_{\boldsymbol{xxz}} + \mathcal{K}_{\boldsymbol{yyz}} - 2\mathcal{K}_{\boldsymbol{xyz}} \,, \tag{3.75}$$

is recognized as the *dipole kernel* [58], and is rapidly decaying, like $1/z^4$, at large z; thus, the corresponding integral is convergent. Note that the potentially troublesome terms behaving like $1/z^2$ have cancelled in the linear combination in the r.h.s. of Eq. (3.75). If one returns to the original formulation, Eq. (3.26), of the RGE, one can recognize this as a cancellation between "real" and "virtual" contributions (i.e., contributions generated by the χ piece and, respectively, the σ piece of the JIMWLK Hamiltonian).

But for the terms in Eq. (3.74) involving one or two factors of $\alpha_{\boldsymbol{z}}^a$, there is no such a cancellation. For instance:

$$\mathcal{K}_{\boldsymbol{xyz}} \left\langle \alpha_{\boldsymbol{x}}^a \alpha_{\boldsymbol{z}}^a \right\rangle_\tau \propto \frac{1}{z^2} \left\langle \alpha_{\boldsymbol{x}}^a \alpha_{\boldsymbol{z}}^a \right\rangle_\tau \qquad \text{for} \qquad z \gg x, y \,, \tag{3.76}$$

which leaves the place for a potential infrared problem.

However, the dangerous terms cancel in the equation satisfied by the linear combination $\langle \alpha_{\boldsymbol{x}}^a \alpha_{\boldsymbol{x}}^a + \alpha_{\boldsymbol{y}}^a \alpha_{\boldsymbol{y}}^a - 2\alpha_{\boldsymbol{x}}^a \alpha_{\boldsymbol{y}}^a \rangle_\tau$, which according to Eq. (3.73) is proportional to the physical amplitude. One finds indeed:

$$\frac{\partial}{\partial \tau} \langle (\alpha_{\boldsymbol{x}}^a - \alpha_{\boldsymbol{y}}^a)^2 \rangle_\tau = \frac{\bar{\alpha}_s}{2\pi} \int d^2\boldsymbol{z} \, \frac{(\boldsymbol{x} - \boldsymbol{y})^2}{(\boldsymbol{x} - \boldsymbol{z})^2 (\boldsymbol{y} - \boldsymbol{z})^2} \tag{3.77}$$
$$\times \left\langle -(\alpha_{\boldsymbol{x}}^a - \alpha_{\boldsymbol{y}}^a)^2 + (\alpha_{\boldsymbol{x}}^a - \alpha_{\boldsymbol{z}}^a)^2 + (\alpha_{\boldsymbol{z}}^a - \alpha_{\boldsymbol{y}}^a)^2 \right\rangle_\tau,$$

which after also using Eq. (3.73) is recognized as the BFKL equation for $\langle T(\boldsymbol{x}, \boldsymbol{y}) \rangle_\tau$, as anticipated. This simple example illustrates a general feature of the evolution generated

by the JIMWLK equation, namely, the infrared divergences cancel between 'real' and 'virtual' contributions, but only in the evolution equations for *gauge–invariant* quantities. The general connection between gauge symmetry and infrared finiteness will be clarified below, in Sect. 3.4.2.

Note also that in the weak–field limit we have obtained a *closed* equation for the 2–point function $\langle \alpha^a_{\boldsymbol{x}} \alpha^a_{\boldsymbol{y}} \rangle_\tau$. This is so since the corresponding version of the JIMWLK Hamiltonian (see Eq. (3.71)) is quadratic both in α and in the functional derivative $\delta/\delta\alpha$. Thus, the evolution generated by this Hamiltonian is diagonal in the number of fields: the n–point correlation functions $\langle \alpha(x_1)\alpha(x_2) \cdots \alpha(x_n) \rangle_\tau$ with different values of n evolve independently from each other. For each of them, one step of the evolution consists in the exchange of one gluon (together with the corresponding virtual corrections) between any pair of fields. In applications to scattering, this means that the number of gluons exchanged in the t-channel remains *fixed* in the course of the evolution. This is similar to the 'multi–reggeon' version of the BFKL evolution [126–130], and in fact it is equivalent to it, as we shall check explicitly for the case of the "odderon" 3–point function in Sect. 3.6.

3.4.2   Gauge symmetry and the dipole JIMWLK Hamiltonian

Returning to the general, strong–field ($g\alpha \sim 1$), case, as described by the full JIMWLK Hamiltonian, it is clear that the non–linear effects encoded in the latter correspond to *gluon mergings* in the hadron wavefunction. This was already manifest from the diagrammatic interpretation of the renormalization group analysis in Sect. 3.1 (see, e.g., Figs. 30 and 33), and to which we shall return in Sect. 3.5. This can be also checked on the structure of the Hamiltonian, as explicit in Eqs. (3.66)–(3.67): when expanding the kernel χ in powers of $g\alpha$, one generates terms like $(g\alpha)^n (\delta/\delta\alpha)^2$ which for $n > 2$ represent vertices in which n gluons merge into two. For instance, the r.h.s. of the evolution equation for the 2–point function $\langle \alpha(x_1)\alpha(x_2) \rangle_\tau$ will involve n–point functions with all values of $n \geq 2$.

Note that the structure of these vertices is quite peculiar since they are coming all from expanding the bilinears of Wilson lines in Eq. (3.27); thus, for any value of n, there will be only three independent transverse coordinates, but n longitudinal ones; besides, for $n \geq 3$, the ordering of the color matrices in x^- will be essential, unlike in the BFKL approximation ($n = 2$). The general vertices are certainly complicated, but there is no need to isolate them out: When the fields are so strong that the non–linear effects play a role, then the expansion of the Wilson lines makes no sense since all the terms in this expansion would be of the same order. Moreover, as anticipated by our previous discussion of eikonal scattering in Sect. 1.2.3, the relevant operators at high energy are themselves built with Wilson lines. In the next section, we shall see that the general evolution equations have a simpler structure when applied to Wilson line operators, while such a simplification does not hold for the n–point functions of the gauge fields themselves. In other terms, in the non–linear regime at high energy, it is crucial to correctly include non–linear effects not only in the *evolution Hamiltonian*, but also in the *definition of the observables*.

Such non–linear effects are in fact *required by gauge symmetry*, and in this section we shall demonstrate that the evolution equations are well defined (in the sense of being

free of infrared singularities) only for such properly defined observables, which are gauge–invariant. This will also give us the opportunity to deduce a simpler, *'dipolar'*, form of the JIMWLK Hamiltonian, valid on the Hilbert space of gauge–invariant operators, in which the cancellation of infrared singularities between the 'real' and 'virtual' corrections has been already performed [36]. This will greatly simplify the subsequent manipulations with this Hamiltonian.

Since the CGC effective theory is already written in a fixed gauge, namely the 'covariant gauge' of Sect. 2.3, it is important to clarify first what we mean by "gauge symmetry" in this context. A given operator $O[\alpha]$ is said to be *gauge–invariant* provided it can be recognized as the expression of a manifestly gauge–invariant functional of A_a^μ (e.g., a Wilson loop) restricted to a gauge field with the simple structure $A_a^\mu(x) = \delta^{\mu+}\alpha_a(x^-, \boldsymbol{x})$. We have seen on specific examples (like the gluon distribution or the dipole scattering amplitude) that such an identification is generally straightforward once the operator $O[\alpha]$ is given. But to be able to characterize the action of the JIMWLK Hamiltonian on a generic operator, we still need an operational criterion for gauge–invariance. This can be inferred by observing that, even within CGC, the gauge is not yet *completely* fixed: The structure $A_a^\mu(x) = \delta^{\mu+}\alpha_a(x^-, \boldsymbol{x})$ of the classical field is preserved by the *'residual'* gauge transformations:

$$\alpha(x^-, \boldsymbol{x}) \to \Omega(x^-)\left(\alpha(x^-, \boldsymbol{x}) + \frac{i}{g}\partial^+\right)\Omega^\dagger(x^-), \quad \Omega(x^-) = \mathrm{e}^{ig\omega^a(x^-)t^a}, \tag{3.78}$$

where the gauge function Ω depends solely upon x^-. Clearly, in order to be gauge–invariant, an operator $O[\alpha]$ must not change under (3.78). In what follows, we shall be only interested in operators built with the Wilson lines (2.22), for which the above transformation amounts to

$$V^\dagger(\boldsymbol{x}) \; \to \; \Omega(x^- = \infty)\, V^\dagger(\boldsymbol{x})\, \Omega^\dagger(x^- = -\infty). \tag{3.79}$$

Eq. (3.79) involves only the gauge function at the extreme points $x^- = \pm\infty$ (by which we mean, of course, points which lie outside the longitudinal support of the hadron, as measured at the resolution scale k^+ of interest). Thus, from the perspective of the 2–dimensional field $V^\dagger(\boldsymbol{x})$, the residual gauge transformations (3.79) are tantamount to two independent *global color rotations*, one on the left, the other on the right:

$$V_{\boldsymbol{x}}^\dagger \to \Omega_{\mathrm{L}} V_{\boldsymbol{x}}^\dagger, \qquad V_{\boldsymbol{x}}^\dagger \to V_{\boldsymbol{x}}^\dagger \Omega_{\mathrm{R}}^\dagger, \tag{3.80}$$

where $\Omega_{\mathrm{L/R}} = \exp(ig\omega_{\mathrm{L/R}}^a t^a)$ is a constant $\mathrm{SU}(N_c)$ matrix. It is easy to check that the infinitesimal generators of these global rotations are the following differential operators:

$$\begin{aligned}
\mathcal{G}_{\mathrm{L}} &\equiv \int d^2\boldsymbol{x}\, \frac{\delta}{\delta\alpha_\tau^a(\boldsymbol{x})}\,, \\
\mathcal{G}_{\mathrm{R}} &\equiv -\int d^2\boldsymbol{x}\, (\widetilde{V}_{\boldsymbol{x}})^{ab}\, \frac{\delta}{\delta\alpha_\tau^b(\boldsymbol{x})}\,.
\end{aligned} \tag{3.81}$$

For instance, the dipole operator $\mathrm{tr}(V_{\boldsymbol{x}}^\dagger V_{\boldsymbol{y}})$ is manifestly invariant under the left and right color rotations (3.80), and is indeed annihilated by the differential operators (3.81), since,

e.g.,

$$\frac{\delta}{\delta\alpha_\tau^a(\boldsymbol{v})}\mathrm{tr}(V_{\boldsymbol{x}}^\dagger V_{\boldsymbol{y}}) = ig\left\{\delta^{(2)}(\boldsymbol{v}-\boldsymbol{x}) - \delta^{(2)}(\boldsymbol{v}-\boldsymbol{y})\right\}\mathrm{tr}(V_{\boldsymbol{x}}^\dagger V_{\boldsymbol{y}}t^a)$$

which vanishes after integration over $\boldsymbol{v}$. For $\mathcal{G}_{\mathrm{R}}$, the corresponding check also involves the formulae $(\widetilde{V})^{ab} = 2\mathrm{tr}(t^a V t^b V^\dagger)$ and $\mathrm{tr}(t^a A)\mathrm{tr}(t^a B) = \frac{1}{2}\mathrm{tr}(AB) - \frac{1}{2N_c}\mathrm{tr}(A)\mathrm{tr}(B)$. More generally, the following $2n$-point operators constructed from Wilson lines in the fundamental representation

$$\mathcal{O}_n \equiv \mathrm{tr}\left(M_1 M_2 \cdots M_n\right), \qquad M_i \equiv V_{x_i}^\dagger V_{y_i}, \tag{3.82}$$

and arbitrary linear combinations and products of them $(\mathcal{O}_{n_1}\mathcal{O}_{n_2}\cdots)$, are gauge invariant: $\mathcal{G}_{\mathrm{L}}\mathcal{O}_n = 0 = \mathcal{G}_{\mathrm{R}}\mathcal{O}_n$.

In particular, it is interesting to show the transformation properties appropriate to the weak–field regime (cf. Sect. 3.4.1), where the natural field variable is $\alpha^a(\boldsymbol{x}) = \int dx^- \alpha^a(x^-, \boldsymbol{x})$, and the gauge function must be restricted to be weak as well. Under an infinitesimal gauge transformation, $\delta\alpha(x^-, \boldsymbol{x}) = \partial^+\omega(x^-)$, so that

$$\alpha^a(\boldsymbol{x}) \longrightarrow \alpha^a(\boldsymbol{x}) + \int_{-\infty}^{\infty} dx^- \partial^+\omega^a(x^-) = \alpha^a(\boldsymbol{x}) + \xi^a, \tag{3.83}$$

with $\xi^a \equiv \omega^a(x^- = \infty) - \omega^a(x^- = -\infty)$ a pure number. It is now obvious that weak–field operators involving the difference between fields at different points, so like Eq. (3.73), are gauge invariant.

Thus, in what follows we shall characterize the gauge–invariant operators by their property to be annihilated by the differential operators (3.81). It is then easily to check that, when acting on any such an operator, the JIMWLK equation (3.66) can be written in the 'dipolar form' (cf. Eq. (3.75)) [36] :

$$\frac{\partial}{\partial\tau}\langle\mathcal{O}\rangle_\tau = \langle H_{\mathrm{dp}}\mathcal{O}\rangle_\tau \equiv -\frac{1}{16\pi^3}\int_{\boldsymbol{xyz}}\frac{(\boldsymbol{x}-\boldsymbol{y})^2}{(\boldsymbol{x}-\boldsymbol{z})^2(\boldsymbol{z}-\boldsymbol{y})^2} \tag{3.84}$$
$$\times\left\langle\left(1 + \widetilde{V}_{\boldsymbol{x}}^\dagger\widetilde{V}_{\boldsymbol{y}} - \widetilde{V}_{\boldsymbol{x}}^\dagger\widetilde{V}_{\boldsymbol{z}} - \widetilde{V}_{\boldsymbol{z}}^\dagger\widetilde{V}_{\boldsymbol{y}}\right)^{ab}\frac{\delta}{\delta\alpha_\tau^a(\boldsymbol{x})}\frac{\delta}{\delta\alpha_\tau^b(\boldsymbol{y})}\mathcal{O}\right\rangle_\tau,$$

where infrared finiteness is now manifest.

Note that in the r.h.s. of Eq. (3.84), the functional derivatives do not act on the Wilson lines in the structure of the Hamiltonian, but only on the external operator $\mathcal{O}$. This is to be contrasted with the original equation (3.66), where the action of the derivatives on η plays a crucial role, in that it generates the 'virtual' piece σ of the quantum corrections (cf. Eq. (3.32)). In fact, in Eq. (3.84), the relative position of the Wilson lines and of the functional derivatives is irrelevant, since *in the presence of the dipole kernel, they commute with each other*. (Recall indeed that, when acting on the Wilson lines inside χ, the functional derivative generates a factor $\delta_{\boldsymbol{xy}}$; see Eq. (3.34).) This is not an accident: In constructing the dipole kernel in Eq. (3.84) we have combined contributions from the 'real' piece η and from the 'virtual' piece σ, and left outside some potentially divergent

terms which vanish when acting on gauge–invariant operators. Thus, the 'virtual' piece is now effectively included as a part of the dipole kernel in Eq. (3.84), and there is no need to generate it through functional differentiation within H. One should remind in this context that a similar combination of real and virtual corrections into the well–behaved dipole kernel is seen to hold in Mueller's "color dipole picture" [58] (see also Sect. 6 below).

We conclude this section with a comment on the importance of Eq. (3.32) relating σ and χ : Since the derivative $\delta/\delta\alpha_\tau^a(\boldsymbol{x})$ acts as a generator for gauge transformations, Eq. (3.32) can be interpreted as a generalized *Ward identity*, thus emphasizing its connection to gauge symmetry. As we have seen, this relation has two crucial consequences for the high energy evolution: (a) It implies that there is no non–trivial fixed–point (cf. Sect. 3.3), and (b) it guarantees the cancellation of infrared singularities in the evolution of physical observables.

3.5   The Balitsky equations

SECT_BALIT

In this section we shall derive evolution equations for the scattering amplitudes describing the elastic scattering between the CGC and some simple external projectiles, like a color dipole or a set of dipoles. As we have seen in Sect. 1.2.3, the relevant operators are Wilson lines which resum multiple scattering in the eikonal approximation. In terms of the color field α of the target, these operators perform a formidable resummation: each amplitude includes an infinite series of n–point functions of α. But in spite of this (or, more correctly, *because* of this), the ensuing equations are relatively simple and intuitive (much simpler than the equations satisfied by the field correlations themselves !) . At a technical level, this simplicity appears because the JIMWLK Hamiltonian has a rather simple expression in terms of Wilson lines (cf. Eq. (3.84)), and the functional derivatives there have a simple action on the Wilson lines (cf. Eq. (3.31)). But at a conceptual level, such a simplicity was indeed expected, as it reflects an alternative interpretation of the evolution as *projectile evolution* : Since the scattering amplitudes are boost invariant, the same equations should be obtained when using the increment in rapidity to accelerate the projectile, and thus evolves its wavefunction. But the projectile is a dilute system, so its evolution remains relatively simple and, especially, *linear*, however complicated are the interactions between its elementary constituents and the target. In fact, this was the physical picture used to justify the Balitsky–Kovchegov (BK) equation in Sect. 1.4, and also the picture used by Balitsky in its original derivation [39] of the equations that we shall find here. We shall return to the interplay between projectile and target evolution after deriving the relevant equations.

Let us start with a single incoming dipole, with quark leg at $\boldsymbol{x}$ and antiquark leg at $\boldsymbol{y}$. The corresponding S–matrix operator reads

$$S(\boldsymbol{x},\boldsymbol{y}) \,=\, \frac{1}{N_c}\,\mathrm{tr}(V_{\boldsymbol{x}}^\dagger V_{\boldsymbol{y}}), \tag{3.85}$$

where the Wilson lines are in the fundamental representation. By using $\mathcal{O} \equiv S(\boldsymbol{x},\boldsymbol{y})$ in Eq. (3.84), together with

$$\frac{\delta}{\delta \alpha_{\boldsymbol{u}}^a} \frac{\delta}{\delta \alpha_{\boldsymbol{v}}^b} S_{\boldsymbol{xy}} = (ig)^2 (\delta_{\boldsymbol{xv}} - \delta_{\boldsymbol{yv}}) \left[\delta_{\boldsymbol{ux}} \mathrm{tr}(t^b t^a V_{\boldsymbol{x}}^\dagger V_{\boldsymbol{y}}) - \delta_{\boldsymbol{uy}} \mathrm{tr}(t^a t^b V_{\boldsymbol{x}}^\dagger V_{\boldsymbol{y}}) \right], \tag{3.86}$$

and the following identities (below, A and B are generic matrices in the fundamental representation):

$$\widetilde{V}_{ab}^\dagger t^b = V t^a V^\dagger, \qquad \mathrm{tr}(t^a A)\mathrm{tr}(t^a B) = \frac{1}{2}\mathrm{tr}(AB) - \frac{1}{2N_c}\mathrm{tr}(A)\mathrm{tr}(B), \tag{3.87}$$

one can easily deduce the first Balitsky equation:

$$\frac{\partial}{\partial \tau}\langle S_{\boldsymbol{xy}}\rangle_\tau = \frac{\bar{\alpha}_s}{2\pi}\int_{\boldsymbol{z}} \mathcal{M}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}) \left\{ -\langle S_{\boldsymbol{xy}}\rangle_\tau + \langle S_{\boldsymbol{xz}} S_{\boldsymbol{zy}}\rangle_\tau \right\}, \tag{3.88}$$

with $\bar{\alpha}_s = \alpha_s N_c/\pi$. This has the structure anticipated in Sect. 1.4 (cf. Eq. (1.63)) and, as discussed there, it is naturally interpreted in terms of projectile evolution, as the splitting of the incoming dipole $(\boldsymbol{x}, \boldsymbol{y})$ into two new dipoles $(\boldsymbol{x}, \boldsymbol{z})$ and $(\boldsymbol{z}, \boldsymbol{y})$, which then both interact with the target. This accounts for the second term, quadratic in S, within the braces in Eq. (3.88), whereas the first term there, which is linear and negative, describes the possibility that the original dipole survives without splitting after an evolution $d\tau$.

Interestingly, this dipole interpretation holds in spite of the fact that there was no large–N_c approximation involved in the derivation of Eq. (3.88): it so happens that the radiation of a small–x gluon off a single dipole can be exactly described as the splitting of the dipole into two dipoles, which are contiguous in transverse space [58]. But Eq. (3.88) is not a closed equation — it couples a 2–point function of the Wilson lines to a 4–point function (with two external legs identified) — but only the first equation in an hierarchy which turns out to be infinite. As we shall soon discover, for the higher equations in this hierarchy, the dipole interpretation holds only in the large–N_c limit, as expected.

When moving higher up in the hierarchy, the complexity of the equations is rapidly increasing, so here we shall present only the second equation, that is, the one satisfied by the 4–point function which enters the r.h.s. of Eq. (3.88):

$$\begin{aligned}
\frac{\partial}{\partial \tau}\langle S_{\boldsymbol{xz}} S_{\boldsymbol{zy}}\rangle_\tau = {}& \frac{\bar{\alpha}_s}{2\pi}\int_{\boldsymbol{w}} \mathcal{M}_{\boldsymbol{xzw}} \left\langle \left(-S_{\boldsymbol{xz}} + S_{\boldsymbol{xw}} S_{\boldsymbol{wz}}\right) S_{\boldsymbol{zy}} \right\rangle_\tau \\
& + \frac{\bar{\alpha}_s}{2\pi}\int_{\boldsymbol{w}} \mathcal{M}_{\boldsymbol{zyw}} \left\langle S_{\boldsymbol{xz}}\left(-S_{\boldsymbol{zy}} + S_{\boldsymbol{zw}} S_{\boldsymbol{wy}}\right) \right\rangle_\tau \\
& + \frac{1}{N_c^2}\frac{\bar{\alpha}_s}{2\pi}\int_{\boldsymbol{w}} \left[\mathcal{M}_{\boldsymbol{xyz}} - \mathcal{M}_{\boldsymbol{xzw}} - \mathcal{M}_{\boldsymbol{zyw}}\right]\left\langle Q_{\boldsymbol{xzwy}} + Q_{\boldsymbol{xwzy}} \right\rangle_\tau,
\end{aligned} \tag{3.89}$$

where in the last line we have introduced the *quadrupole operator* :

$$Q(\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{w}, \boldsymbol{y}) \equiv \frac{1}{N_c}\left[\mathrm{tr}(V_{\boldsymbol{x}}^\dagger V_{\boldsymbol{z}} V_{\boldsymbol{w}}^\dagger V_{\boldsymbol{y}} V_{\boldsymbol{z}}^\dagger V_{\boldsymbol{w}}) - \mathrm{tr}(V_{\boldsymbol{x}}^\dagger V_{\boldsymbol{y}})\right]. \tag{3.90}$$

In the weak field limit, this operator starts at order $(g\alpha)^4$. This equation too admits a relatively simple interpretation in terms of projectile evolution: When increasing the rapidity by $d\tau$, a two–dipole system evolves by radiating one small–x gluon from any of the two incoming dipoles. If the radiated gluon is eventually absorbed by the same dipole

which has emitted it, the process amounts to the splitting of that dipole into two new ones; thus, at the time of scattering, the system consists in three dipoles, and is described by the first two lines in the r.h.s. of Eq. (3.89). If, one the other hand, the gluon is absorbed by the other dipole, then no dipole survives after the evolution, and the system that scatters is rather a color quadrupole [11] (made with the quarks and the antiquarks of the original dipoles together with the radiated gluon). However, the exchange of a gluon in between different dipoles is suppressed by a factor $1/N_c^2$ at large N_c, and this suppression factor is indeed manifest in the last term of Eq. (3.89).

At this point, the general pattern of the projectile evolution should be clear: Through successive gluon emissions, the original dipole evolves into a complicated system, which in general involves several dipoles together with composite objects with higher multipolar moments. For the hierarchy to be complete, we also need the evolution equations satisfied by such multipolar operators. But a drastic simplification occurs in the large–N_c limit, in which the evolution can be restricted to the space of dipoles: a general configuration then consists in a set of dipoles, which with increasing τ evolves through dipole splitting. Thus, at large N_c, the whole hierarchy can be generated from the following "operator equation":

$$\frac{\partial S(\boldsymbol{x}, \boldsymbol{y})}{\partial \tau} = \frac{\bar{\alpha}}{2\pi} \int d^2\boldsymbol{z} \, \mathcal{M}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \left\{ -S(\boldsymbol{x}, \boldsymbol{y}) + S(\boldsymbol{x}, \boldsymbol{z}) S(\boldsymbol{z}, \boldsymbol{y}) \right\}, \qquad (3.91)$$

with the help of the Leibniz rule; e.g. :

$$\frac{\partial}{\partial \tau} \left\langle S(1) S(2) \right\rangle_\tau = \left\langle \frac{\partial S(1)}{\partial \tau} S(2) \right\rangle_\tau + \left\langle S(1) \frac{\partial S(2)}{\partial \tau} \right\rangle_\tau . \qquad (3.92)$$

Note, however, that in this limit, the dipolar evolution is *quasi–deterministic*. For instance if the initial conditions at τ_0 are chosen in *factorized* form, i.e., $\langle S(1) \cdots S(N) \rangle_0 = \langle S(1) \rangle_0 \cdots \langle S(N) \rangle_0$, then this factorized form will be preserved by the large–N_c evolution up to arbitrarily large τ : $\langle S^{(N)} \rangle_\tau = \langle S \rangle_\tau^N$, with the one–point function $\langle S \rangle_Y$ obeying the Kovchegov equation [124] (i.e., the equation obtained by replacing $S \to \langle S \rangle_\tau$ into Eq. (3.91)). More generally, it has been shown in Refs. [158, ?] that the hierarchy generated by Eq. (3.91) admits a one–parameter family of fully factorized exact solutions. Thus, in their simplified version valid at large N_c, the Balitsky equations do not generate new correlations, but only propagate those already encoded in the initial conditions.

Note also that, although the Balitsky equations have been obtained here by performing *target* evolution, their interpretation appears to be more natural in terms of the evolution of the *projectile*. This is so because of the Hamiltonian structure of the RGE (3.35) which allows for a dual ('Schrödinger vs. Heisenberg', or 'passive vs. active') interpretation of the results of the evolution (cf. the discussion in Sect. 3.4). In particular, some approximations like the large–N_c limit are easier to implement in terms of projectile evolution, since they depend upon the internal color structure of the operator that the Hamiltonian is acting on. In Sect. 6 we shall see that the large–N_c approximation can be also implemented in the target evolution, provided the color structure of the latter is explicitly under control.

---

[11] If the two incoming dipoles were not contiguous with each other, the scattering state in this case would be rather a color sextupole.
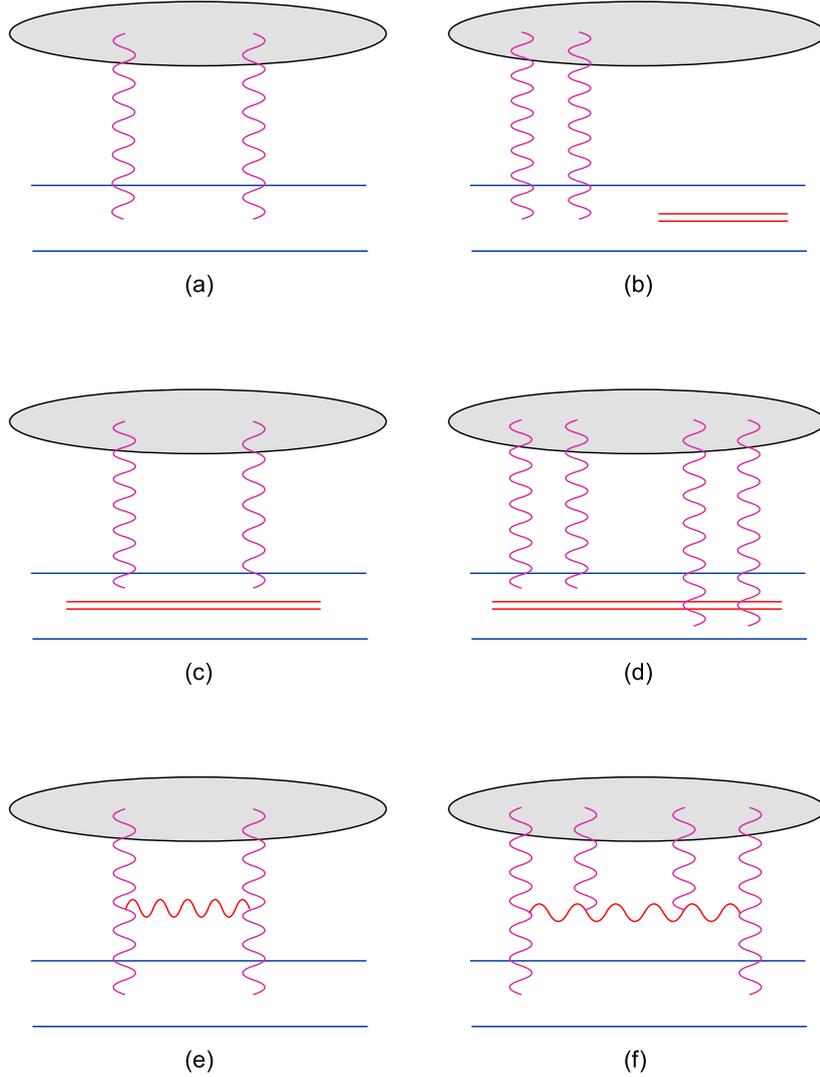
Fig. 34. *Diagrams for single dipole scattering: (a) the tree–level contribution; (b,c,d) one step evolution of the projectile; (e,f) one step evolution of the target.* `FIG_1DIP`

It is finally instructive to compare the digrammatic interpretations of Eq. (3.88) corresponding to projectile evolution, and to target evolution, respectively. This will make clear that non–linear effects appear as either multiple scattering, or gluon saturation, depending upon our perspective.

In Fig. 34, we display some of the relevant Feynman graphs. For more clarity, we show only one diagram contributing to each type of process, which is moreover taken at the lowest non–trivial order in perturbation theory. Thus, the scattering between a dipole and the CGC target starts at two gluon exchange, as shown in Fig. 34.a. The one–step evolution of the projectile (i.e., the dipole splitting) generates the diagrams in Figs. 34.b, c, d: Fig. 34.b is the 'virtual' term (the original dipole survives without splitting), Fig. 34.c shows the scattering of one of the two child dipoles (there is a similar diagram for the other dipole), while Fig. 34.d describes their simultaneous scattering.

The corresponding diagrams for target evolution are shown in Figs. 34.e and f. The first

one, Fig. 34.e, represents one step in the BFKL evolution of the 2–point function $\langle\alpha\alpha\rangle_\tau$, which yields the scattering amplitude to lowest order (cf. Eq. (3.73)); in terms of projectile evolution, it corresponds to the two diagrams shown in Figs. 34.b and c. The second diagram, Fig. 34.f, represents the merging of four gluons into two. This is a saturation effect encoded in the JIMWLK equation (the term of order $(g\alpha)^4(\delta/\delta\alpha)^2$ in the Hamiltonian) and corresponds to the double scattering process displayed in Fig. 34.d.

As we shall further discuss in Sect. 4.1, the non–linear effects encoded in the Balitsky hierarchy ensure the unitarization of the scattering amplitudes at high energy.

3.6 Odderon evolution in the CGC and the BKP equation

SECT_ODD

We have previously seen that, in the weak field regime, the dipole scattering amplitude (3.73) obeys the BFKL equation, and thus can be identified as the 'BFKL pomeron' [16] (that is, the two–gluon exchange dressed by the BFKL evolution). It is therefore tempting to identify the general dipole operator (3.72), which is built with Wilson lines and obeys the non–linear Balitsky equation Eq. (3.88), as the *non–linear generalization of the BFKL pomeron*, valid also in the strong field regime where unitarity corrections are important. As we shall argue now, this is essentially correct, except for the fact that the identification with the 'pomeron' should hold only for the *real* part of the dipole amplitude in Eq. (3.72) (which has indeed the quantum numbers of the pomeron). In general, this amplitude has also an *imaginary part*, which, as we shall shortly explain, is odd under charge conjugation, and therefore describes the exchange of an *odderon* [8] (see Ref. [137] for a recent review of the odderon in QCD, and more references).

The purpose of this section is to describe odderon exchanges in the CGC formalism, and to show that, in the weak field limit, they obey evolution equations which are equivalent to the corresponding equation in perturbative QCD, known as the Bartels–Kwiecinski–Praszalowicz (BKP) equation [126, 127]. As we shall see, the CGC approach sheds some new light on the physical relevance of the solutions to the BKP equation.

Before we consider the charge parity (C) of the dipole scattering operator in a more formal way, let us briefly discuss the physical need for odderon exchanges in dipole–CGC scattering. Consider the production of a of C–even meson like η_c in the diffractive scattering of a virtual photon on a hadronic target at high energy. In the LLA, the respective cross–section admits a factorization similar to that of DIS at small–x : the virtual photon dissociates into a $q\bar{q}$ pair (a 'color dipole'), which then elastically scatters off the color fields in the hadron before finally recombining in the outgoing meson. Since, however, a (virtual) photon has negative C–parity, the gluon exchange between the dipole and the hadron must be C–odd as well, in order for a C–even final state to be produced [134, 159]. In perturbative QCD, the lowest order t–channel exchange which is C–odd is the exchange of three gluons in a totally symmetric color state (see below). The eikonalized dipole amplitude in Eq. (3.72) contains n–gluon exchanges with any $n \geq 2$, and thus potentially C–odd exchanges, that we shall isolate now.

Let us first recall the transformation properties of the gauge fields under charge conjugation:

$$C \, A_\mu \, C^{-1} = -(A_\mu)^T, \tag{3.93}$$

which immediately implies that the three–gluon operator $d^{abc} A_\mu^a(\boldsymbol{x}) A_\nu^b(\boldsymbol{y}) A_\rho^c(\boldsymbol{z})$ (with $d^{abc} = 2\mathrm{tr}(\{t^a, t^b\} t^c)$ the totally symmetric tensor) is indeed C–odd. Moreover, Eq. (3.93) implies the following transformation law for Wilson lines:

$$C \, V \, C^{-1} = (V^\dagger)^T = V^*, \tag{3.94}$$

and therefore $C \, \mathrm{tr}(V_{\boldsymbol{x}}^\dagger V_{\boldsymbol{y}}) \, C^{-1} = \mathrm{tr}(V_{\boldsymbol{y}}^\dagger V_{\boldsymbol{x}})$. We thus conclude that the C–odd piece of the dipole S–matrix (3.85) *("the dipole odderon operator")* is given by

$$O(\boldsymbol{x}, \boldsymbol{y}) \equiv \frac{1}{2iN_c} \, \mathrm{tr}(V_{\boldsymbol{x}}^\dagger V_{\boldsymbol{y}} - V_{\boldsymbol{y}}^\dagger V_{\boldsymbol{x}}) \, = \, -O(\boldsymbol{y}, \boldsymbol{x}), \tag{3.95}$$

where the factor of i is introduced in order for this quantity to be real: indeed, since V and $V^\dagger$ are unitary matrices, we have $[\mathrm{tr}(V_{\boldsymbol{x}}^\dagger V_{\boldsymbol{y}})]^* = \mathrm{tr}(V_{\boldsymbol{y}}^\dagger V_{\boldsymbol{x}})$. Thus we shall write in general $S = 1 - T + iO$, with $T \equiv 1 - \Re\mathfrak{e}\, S$ identified as the 'pomeron'. Note the obvious boundary conditions which follow from $S(\boldsymbol{x}, \boldsymbol{x}) = 1$:

$$T(\boldsymbol{x}, \boldsymbol{x}) \, = \, O(\boldsymbol{x}, \boldsymbol{x}) \, = \, 0. \tag{3.96}$$

The lowest non–trivial contribution to Eq. (3.95) is obtained by expanding the Wilson lines there up to *cubic* order in the field α in the exponent:

$$O(\boldsymbol{x}, \boldsymbol{y}) \simeq \frac{-g^3}{24N_c} \, d^{abc} (\alpha_{\boldsymbol{x}}^a - \alpha_{\boldsymbol{y}}^a)(\alpha_{\boldsymbol{x}}^b - \alpha_{\boldsymbol{y}}^b)(\alpha_{\boldsymbol{x}}^c - \alpha_{\boldsymbol{y}}^c). \tag{3.97}$$

This expression is in agreement with the expectations from perturbative QCD, in the sense that it is a trilinear operator in α with the color indices contracted symmetrically by the d–symbol. Note that, because of the symmetry properties of this symbol, the path–ordering of the Wilson lines in x^- has been irrelevant for computing $O(\boldsymbol{x}, \boldsymbol{y})$. Furthermore, the linear combination of field operators in Eq. (3.97) is manifestly invariant under a residual gauge transformation, which consists in a constant shift of the field: $\alpha^a \to \alpha^a + \xi^a$ (cf. Eq. (3.83)).

The evolution equations for pomeron and odderon exchanges in the dipole–CGC scattering are immediately obtained by separating the real part and the imaginary part of the first Balitsky equation (3.88) [36] :

$$\frac{\partial}{\partial \tau} \langle O(\boldsymbol{x}, \boldsymbol{y}) \rangle_\tau = \frac{\bar{\alpha}}{2\pi} \int_{\boldsymbol{z}} \mathcal{M}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \Big\langle O(\boldsymbol{x}, \boldsymbol{z}) + O(\boldsymbol{z}, \boldsymbol{y}) - O(\boldsymbol{x}, \boldsymbol{y})$$
$$- O(\boldsymbol{x}, \boldsymbol{z})T(\boldsymbol{z}, \boldsymbol{y}) - T(\boldsymbol{x}, \boldsymbol{z})O(\boldsymbol{z}, \boldsymbol{y}) \Big\rangle_\tau, \tag{3.98}$$

$$\frac{\partial}{\partial \tau} \langle T(\boldsymbol{x}, \boldsymbol{y}) \rangle_\tau = \frac{\bar{\alpha}}{2\pi} \int_{\boldsymbol{z}} \mathcal{M}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \Big\langle T(\boldsymbol{x}, \boldsymbol{z}) + T(\boldsymbol{z}, \boldsymbol{y}) - T(\boldsymbol{x}, \boldsymbol{y})$$
$$- T(\boldsymbol{x}, \boldsymbol{z})T(\boldsymbol{z}, \boldsymbol{y}) + O(\boldsymbol{x}, \boldsymbol{z})O(\boldsymbol{z}, \boldsymbol{y}) \Big\rangle_\tau. \tag{3.99}$$

The mean field version of these equations, in which the non–linear terms are assumed to factorize (so like in the Kovchegov equation [124]), has been first proposed by Kovchegov, Szymanowski and Wallon [160].

Interestingly, the non–linear terms in the above equations couple the evolution of C–odd and C–even operators. For instance, the last term, quadratic in O, in the r.h.s. of Eq. (3.99) for $\langle N \rangle_\tau$ describes the merging of two odderons into one pomeron. The vertex connecting one pomeron to two odderons has been also computed in lowest order perturbation theory in Ref. [142], and it would be interesting to compare the respective result with the corresponding vertex in Eq. (3.99), which is essentially the dipole kernel. As we shall argue in Sect. 4.1, the odderon–pomeron coupling encoded in the last terms in Eq. (3.98) has the rather dramatic effect to suppress the odderon contributions in the high energy regime where unitarity corrections start to be important (i.e., where $\langle T \rangle_\tau \sim 1$).

In the remaining part of this section, we shall focus on the weak scattering regime, and the relation with perturbative QCD. For not too high energies, such that both T and O are small compared to one, one can neglect the non–linear terms in the equations above, which then reduce to the BFKL equation for *both* $\langle T \rangle_\tau$ and $\langle O \rangle_\tau$! That is, not only the pomeron amplitude obeys Eq. (1.65), which was to be expected, but a similar equation holds also for the odderon amplitude:

$$\frac{\partial}{\partial \tau} \langle O(\boldsymbol{x}, \boldsymbol{y}) \rangle_\tau = \frac{\bar{\alpha}}{2\pi} \int_{\boldsymbol{z}} \mathcal{M}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \big\langle O(\boldsymbol{x}, \boldsymbol{z}) + O(\boldsymbol{z}, \boldsymbol{y}) - O(\boldsymbol{x}, \boldsymbol{y}) \big\rangle_\tau. \tag{3.100}$$

(This equation has been first derived in Ref. [160], within the color dipole picture [58, 124].) In view of the difference between the respective operators, cf. Eqs. (3.73) and (3.97), this coincidence may look at the first sight surprising, but some further thinking reveals that this is actually not so: When seen from the perspective of the projectile, the (linear) evolution consists in the splitting of the original dipole into two, followed by the scattering between one of the child dipoles and the target. This process looks the same whatever are the quantum numbers of the exchanged object, and this is indeed the physical content of both Eqs. (1.65) and (3.100). Moreover, the structure of the BFKL equation is uniquely fixed by the singularities of the dipole kernel together with the conditions (3.96).

But the initial conditions corresponding to T and O are, of course, different (in particular, they have opposite C–parities), and they lead to very different behaviors for the corresponding solutions at high energy. For instance, in the case where the target itself is an elementary dipole at $\tau = 0$ (with the quark leg at $\boldsymbol{x}_0$ and the antiquark one at $\boldsymbol{y}_0$), the initial scattering amplitudes read (see Sect. 6)

$$\langle T(\boldsymbol{x}, \boldsymbol{y}) \rangle_{\tau=0} = \frac{\alpha^2}{2} \frac{N_c^2 - 1}{N_c^2} \ln^2 \frac{|\boldsymbol{x} - \boldsymbol{x}_0||\boldsymbol{y} - \boldsymbol{y}_0|}{|\boldsymbol{x} - \boldsymbol{y}_0||\boldsymbol{y} - \boldsymbol{x}_0|}, \tag{3.101}$$

for the pomeron and, respectively,

$$\langle O(\boldsymbol{x}, \boldsymbol{y}) \rangle_{\tau=0} = \frac{\alpha^3}{12} \frac{(N_c^2 - 4)(N_c^2 - 1)}{N_c^3} \ln^3 \frac{|\boldsymbol{x} - \boldsymbol{x}_0||\boldsymbol{y} - \boldsymbol{y}_0|}{|\boldsymbol{x} - \boldsymbol{y}_0||\boldsymbol{y} - \boldsymbol{x}_0|}. \tag{3.102}$$

for the odderon. At high energy (but within the linear regime though), the projection of the general BFKL solution onto C–even states has a maximal intercept $\alpha_{\mathbb{P}} = 1 + (4 \ln 2) \bar{\alpha}$ (the 'BFKL pomeron' [16]), whereas the respective projection on C–odd initial conditions has a maximal intercept $\alpha_{\mathbb{O}} = 0$ [160] (that is, the odderon solution rises as most as a logarithm

of the energy), and coincides with the Bartels–Lipatov–Vacca (BLV) solution [134] to the BKP equation. This brings us to the other point that we would like to discuss here, namely the relation between odderon evolution in the CGC formalism and in perturbative QCD.

The 'BKP equation' describing the BFKL evolution of the three–gluon odderon exchanges has been originally derived by Bartels [126] and Kwiecinski and Praszalowicz [127] (see also [128]). The BKP Hamiltonian is the sum of three BFKL Hamiltonians, one for each pair of gluons. In Ref. [160], the consistency between the BKP equation and the BFKL equation (3.100) describing C–odd exchanges in the dipole scattering has been verified via a direct comparison of the respective solutions [134]. In what follows, we shall follow a more general approach, proposed in Ref. [36], which demonstrates that *all* the odderon exchanges in the CGC formalism (and not only the one pertinent to dipole scattering) reduce to the BKP odderon in the weak field limit.

A direct comparison between the two formalisms is hindered by the fact that, *a priori*, they deal with different objects: The perturbative QCD approach exploits $k_\perp$–factorization to express a scattering amplitude as the convolution of a universal *odderon Green's function*, which is written in momentum space and obeys the BKP equation, with process–dependent *impact factors*, which connect the exchanged gluons to the external particles. On the other hand, in the CGC formalism, there is generally no $k_\perp$–factorization and only the overall scattering amplitudes make *a priori* sense; these are gauge–invariant operators built with Wilson lines, which include the projectile impact factor (the Wilson lines themselves !) and are written in *coordinate space*.

Yet, in the weak–field regime, the two formalisms are expected to become equivalent with each other, so $k_\perp$–factorization should emerge in this limit from the CGC formalism as well. As we shall show now, this is indeed the case: In the dilute regime, one can identify a universal odderon Green's function also in the CGC, and this turns out to obey the coordinate–space version of the BKP equation.

Specifically, the weak–field version of the dipolar odderon amplitude, Eq. (3.97), can be rewritten as:

$$\langle O(\boldsymbol{x}, \boldsymbol{y})\rangle_\tau \simeq \frac{-g^3}{24 N_c}\left\{3\Big(f_\tau(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{y}) - f_\tau(\boldsymbol{x}, \boldsymbol{x}, \boldsymbol{y})\Big) + f_\tau(\boldsymbol{x}, \boldsymbol{x}, \boldsymbol{x}) - f_\tau(\boldsymbol{y}, \boldsymbol{y}, \boldsymbol{y})\right\},$$

$$(3.103)$$

where

$$f_\tau(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \equiv d^{abc}\langle \alpha_{\boldsymbol{x}}^a \alpha_{\boldsymbol{y}}^b \alpha_{\boldsymbol{z}}^c\rangle_\tau \tag{3.104}$$

is a natural candidate for the *CGC odderon Green's function*. Indeed, after going to momentum space, via

$$f_\tau(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) \equiv \int \frac{d^2\boldsymbol{x}\, d^2\boldsymbol{y}\, d^2\boldsymbol{z}}{(2\pi)^6}\, \mathrm{e}^{-i\boldsymbol{k}_1\boldsymbol{x} - i\boldsymbol{k}_2\boldsymbol{y} - i\boldsymbol{k}_3\boldsymbol{z}}\, f_\tau(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}), \tag{3.105}$$

the scattering amplitudes (3.103) takes the $k_\perp$–factorized structure [12] :

---

[12] The impact factor of the target (the CGC) is still implicit in the definition (3.104) of the CGC

$$\langle O(\boldsymbol{x}, \boldsymbol{y}) \rangle_\tau = \int d^2 \boldsymbol{k}_1 d^2 \boldsymbol{k}_2 d^2 \boldsymbol{k}_3 f_\tau(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) \tag{3.106}$$
$$\times (e^{i\boldsymbol{k}_1 \boldsymbol{x}} - e^{i\boldsymbol{k}_1 \boldsymbol{y}})(e^{i\boldsymbol{k}_2 \boldsymbol{x}} - e^{i\boldsymbol{k}_2 \boldsymbol{y}})(e^{i\boldsymbol{k}_3 \boldsymbol{x}} - e^{i\boldsymbol{k}_3 \boldsymbol{y}}),$$

where the exponential terms within the parentheses correspond to all the possible attachments of the three exchanged gluons to the quark and antiquark lines in the dipole. This is the *dipole–odderon impact factor*. Note that this vanishes whenever $\boldsymbol{k}_i = 0$ for some i, as required by gauge symmetry: a zero–momentum gluon 'sees' the projectile as a whole, and the latter is globally color neutral.

To better appreciate the universality of the odderon Green's function (3.104), let us give another example where this quantity controls the weak field limit of odderon exchanges. Specifically, let us consider the eikonal scattering between a system of three quarks in a color singlet state and the CGC. For $N_c = 3$, a gauge–invariant "baryonic" operator which is sometimes used in the literature (see, e.g., [36, 161]) to model such a 3–quark system is $\epsilon^{ijk} \psi^i(\boldsymbol{x}) \psi^j(\boldsymbol{y}) \psi^k(\boldsymbol{z})$, where ϵ^{ijk} is the complete antisymmetric symbol, and the color indices i, j, k can take the values 1, 2, or 3. The corresponding S–matrix reads

$$S_\tau(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) = \frac{1}{3!} \epsilon^{ijk} \epsilon^{lmn} \langle V_{il}^\dagger(\boldsymbol{x}) V_{jm}^\dagger(\boldsymbol{y}) V_{kn}^\dagger(\boldsymbol{z}) \rangle_\tau, \tag{3.107}$$

where $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{z}$ are the transverse positions of the three quarks. This amplitude is symmetric under any permutations of the three coordinates, and is normalized as $S_\tau(\boldsymbol{x}, \boldsymbol{x}, \boldsymbol{x}) = 1$. By using Eq. (3.93), it is easy to check that the odderon contribution is given again by the imaginary part of the S-matrix :

$$\langle B(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \rangle_\tau = \Im m \, S_\tau(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}). \tag{3.108}$$

The 3–quark odderon operator $B(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ is totally symmetric too, and satisfies the boundary condition $B(\boldsymbol{x}, \boldsymbol{x}, \boldsymbol{x}) = 0$. In particular, when two of the coordinates are the same, the 3–quark odderon operator reduces to the dipole odderon operator, Eq. (3.95):

$$B(\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{z}) = O(\boldsymbol{x}, \boldsymbol{z}) = -B(\boldsymbol{x}, \boldsymbol{x}, \boldsymbol{z}). \qquad (N_c = 3) \tag{3.109}$$

In the weak field approximation, $B(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ reduces again to a gauge invariant linear combination of 3–point functions like Eq. (3.104) :

$$B(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \simeq \frac{g^3}{144} d^{abc} \left\{ (\alpha_{\boldsymbol{x}}^a - \alpha_{\boldsymbol{z}}^a) + (\alpha_{\boldsymbol{y}}^a - \alpha_{\boldsymbol{z}}^a) \right\} \tag{3.110}$$
$$\times \left\{ (\alpha_{\boldsymbol{y}}^b - \alpha_{\boldsymbol{x}}^b) + (\alpha_{\boldsymbol{z}}^b - \alpha_{\boldsymbol{x}}^b) \right\} \left\{ (\alpha_{\boldsymbol{z}}^c - \alpha_{\boldsymbol{y}}^c) + (\alpha_{\boldsymbol{x}}^c - \alpha_{\boldsymbol{y}}^c) \right\}.$$

This can be given the $k_\perp$–factorized form (compare to Eq. (3.106))

$$\langle B(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \rangle_\tau = \int d^2 \boldsymbol{k}_1 d^2 \boldsymbol{k}_2 d^2 \boldsymbol{k}_3 \, f_\tau(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) \, (2 \, e^{i\boldsymbol{k}_1 \boldsymbol{x}} - e^{i\boldsymbol{k}_1 \boldsymbol{y}} - e^{i\boldsymbol{k}_1 \boldsymbol{z}})$$
$$\times (2 \, e^{i\boldsymbol{k}_2 \boldsymbol{y}} - e^{i\boldsymbol{k}_2 \boldsymbol{z}} - e^{i\boldsymbol{k}_2 \boldsymbol{x}})(2 \, e^{i\boldsymbol{k}_3 \boldsymbol{z}} - e^{i\boldsymbol{k}_3 \boldsymbol{x}} - e^{i\boldsymbol{k}_3 \boldsymbol{y}}), \tag{3.111}$$

Green's function.

which reinforces our interpretation of $f_\tau(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)$ as a momentum–space odderon Green's function.

It remains to establish the evolution equation obeyed by the 3–point function (3.104). This quantity is not gauge–invariant by itself, so its evolution under the original JIMWLK Hamiltonian would be afflicted by infrared singularities. However, these singularities are physically harmless, as they cancel in the linear combination (3.103) for the scattering amplitude. Thus, we are allowed to *define* the odderon Green's function as the solution to the infrared–safe equation generated by the *dipolar* version of the Hamiltonian, Eq. (3.84). The use of H_{dp} in connection with a *non*–gauge-invariant quantity should be viewed as an infrared regularization, which is convenient at intermediate steps and has no incidence on the final results for gauge–invariant observables.

In the weak–field regime of interest here we need only the approximate form of Eq. (3.84) in which the Wilson lines are expanded to second order in α (like in Eq. (3.71)). By applying the ensuing Hamiltonian on $\mathcal{O} = d^{abc}\alpha_{\boldsymbol{x}}^a \alpha_{\boldsymbol{y}}^b \alpha_{\boldsymbol{z}}^c$, and after some simple algebra, one obtains

$$
\begin{aligned}
\frac{\partial}{\partial \tau} f_\tau(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \;=\; & \frac{\bar{\alpha}}{4\pi} \int d^2 \boldsymbol{w}\, \frac{(\boldsymbol{x} - \boldsymbol{y})^2}{(\boldsymbol{x} - \boldsymbol{w})^2 (\boldsymbol{y} - \boldsymbol{w})^2} \\
& \Big(f_\tau(\boldsymbol{x}, \boldsymbol{w}, \boldsymbol{z}) + f_\tau(\boldsymbol{w}, \boldsymbol{y}, \boldsymbol{z}) - f_\tau(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) - f_\tau(\boldsymbol{w}, \boldsymbol{w}, \boldsymbol{z}) \Big) \\
& + \big\{ 2 \text{ cyclic permutations} \big\}.
\end{aligned}
\tag{3.112}
$$

By using this equation, it is straightforward to check that the linear combination in Eq. (3.103) obeys indeed the BFKL equation (3.100), as it should. But our main focus in what follows will be directly on Eq. (3.112), which can be recognized as the coordinate–space version of the BKP equation [36]. Indeed, the Fourier transform of Eq. (3.112) — i.e., the equation satisfied by $f_\tau(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)$ — turns out to be the same as the standard BKP equation [126, 127] up to terms proportional to δ–functions $\delta^{(2)}(\boldsymbol{k}_i)$ $(i = 1, 2, 3)$, which are however irrelevant for the calculation of the scattering amplitudes in Eqs. (3.106) or (3.111). This finally confirms our identification of Eq. (3.104) as the odderon Green's function in the BFKL approximation. It also suggests that one could search for solutions to the BKP equation directly in coordinate space, by solving Eq. (3.112).

The structure of the last equation, and also the explicit examples of odderon amplitudes that have been discussed before, have interesting implications for the physical Hilbert space to be used in connection with the BKP equation and its generalizations. So far, most of the theoretical efforts aiming at solving these equations have concentrated on a particular Hilbert space, known as the *Möbius space*. This is the space of functions $f_\tau(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \dots)$ which have the property to vanish whenever two coordinates coincide with each other: $f_\tau(\boldsymbol{x}, \boldsymbol{x}, \boldsymbol{z}, \dots) = 0$, etc. For instance, the dipole scattering amplitude (3.73) belongs to this space. This property is important in the context of BFKL evolution, since it turns out that the BFKL Hamiltonian exhibits holomorphic separability when restricted to the Möbius space [130]. In turn, this has interesting mathematical consequences, which have been exploited in the recent years to construct exact solutions to the BKP equation (the Janik–Wosiek odderon [133]) and, more generally, to the equation which describes the exchange of n reggeized gluons (with $n \geq 3$) in the large–N_c limit [131, 132, 135–137].

Indeed, the restriction of the BKP equation to a given holomorphic sector describes an integrable system [131] which is in fact equivalent to the XXX Heisenberg model of spin $s = 0$ [132].

Although these constructions are mathematically elegant, one should keep in mind that the Möbius space is not the *complete* Hilbert space for the BFKL Hamiltonian, and it is not clear to which extent it overlaps with the *physical* Hilbert space (see also the related discussion in Ref. [146]). For instance, functions in the Möbius space do not contribute to the dipolar amplitude (3.103), yet we know by now that non–trivial BKP solutions exist, which describe odderon exchanges in the dipole scattering (and thus necessarily lie outside the Möbius space) [134, 160]. These are simply the solutions to the BFKL equation (3.100) with a C–odd initial condition like Eq. (3.102). Besides, although a Möbius function can in principle contribute to the C–odd amplitude (3.110) for the 3–quark system, the resulting amplitude has the rather curious property to vanish at equal points ($\langle B(\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{z})\rangle_\tau = 0$, etc.), for which there is no compelling physical justification.

Note finally that, unlike in the pomeron case, where the restriction to the Möbius space is natural and does not entail any loss of generality, for the odderon problem this is a highly nontrivial issue, as it can be appreciated when trying to project Eq. (3.112) on this particular space. Namely, one can check that, in order for a Möbius function to be a solution to Eq. (3.112), it must obey the following, strong, constraint (to see this, consider Eq. (3.112) for $\boldsymbol{y} = \boldsymbol{z}$) :

$$\int d^2\boldsymbol{w} \frac{(\boldsymbol{x} - \boldsymbol{y})^2}{(\boldsymbol{x} - \boldsymbol{w})^2(\boldsymbol{y} - \boldsymbol{w})^2} \, f_\tau(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{w}) \, = \, 0. \tag{3.113}$$

4 Some physical consequences

With this section, we begin the exploration of the physical consequences of the effective theory for the CGC. As se shall see in Sect. 4.1, it is rather straightforward to show that the JIMWLK evolution predicts the (quasi)saturation of the gluon distribution [37, 38] and that it respects the unitarity condition for the scattering amplitudes (which appear to approach the black disk limit for very high energy) [35, 38, 162, 163]. Then, in Sect. 4.2, we shall compute the energy dependence of the saturation momentum [22, 37]; to that aim, we shall rely on the linear, BFKL, equation supplemented with an absorptive boundary condition which mimics saturation at low momenta [51, 52]. In that context, we shall also discuss running coupling effects (in particular, the next–to–leading order calculation by Triantafyllopoulos [53]), as well as the atomic number dependence of Q_s for the case of a large nucleus. In this calculation, we shall discover a remarkable *geometric scaling* property for the dipole scattering amplitude [51, 52], which provides a natural explanation for the similar scaling observed at HERA [83] (cf. Sect. 1.2). A mathematically more elegant argument for scaling, due to Munier and Peschanski [164], will be also discussed, and this will give us the opportunity to introduce a far–reaching correspondence between the non–linear evolution equations in QCD and some modern problems in statistical physics [165, 166]. Finally, in Sect. 4.4 we shall discuss the impact parameter dependence of the scattering amplitude and the high energy behaviour of the total cross–section. Since

it lacks confinement, the JIMWLK evolution leads to total cross–sections which violate the Froissart bound, because of the too rapid expansion of the 'black disk' [167, 168]. Still, as we shall explain in Sect. 4.4, the Froissart bound can be restored provided the (perturbative) unitarization mechanism inherent in the Balitsky–JIMWLK equations is supplemented with a non–perturbative assumption about confinement [169].

4.1   Gluon saturation and unitarity from JIMWLK evolution

SECT_SAT

In the high energy regime and for relatively low transverse momenta, which is where the saturation effects are expected, the color fields are strong and the dynamics is fully non–linear, which hinders the use of perturbative techniques (like the expansion of the Wilson lines) in the analysis of the evolution equations. However, the general properties of the RGE, as described in the previous section, enable us to perform some drastic simplifications on the general equations, which makes it possible to study (at least, qualitatively) the general behaviour of the gluon correlations and of the scattering amplitudes in the high energy limit.

Specifically, as argued in Sect. 3.3 (see, especially, the discussion towards the end of that section), in the regime where the classical fields are strong [13] $g\alpha \sim 1$, the Wilson lines (3.30) — which are complex exponentials built with these fields — oscillate around zero over a transverse distance of the order of the wavelength $\sim 1/Q_s(\tau)$ of the fields α (we shall see indeed that the saturated gluons carry typical momenta $\sim Q_s$). This implies that Wilson lines which are separated by large distances $\gg 1/Q_s(\tau)$ are uncorrelated with each other, (their relative phases are random), so the respective correlation functions can be neglected, or treated as small quantities. This is the "random phase approximation" introduced in Refs. [35, 38].

4.1.1   Gluon saturation through quantum evolution

In the random phase approximation, the RGE (3.26) simplifies drastically [38]. Neglecting the Wilson lines, the kernel χ becomes independent of α, and the RGE reads (in momentum space)

$$\frac{\partial W_\tau[\alpha]}{\partial \tau} = \frac{1}{2} \int \frac{d^2\boldsymbol{k}}{(2\pi)^2} \frac{1}{\pi k_\perp^2} \frac{\delta^2 W_\tau[\rho]}{\delta\alpha_\tau^a(k_\perp)\delta\alpha_\tau^a(-k_\perp)} \,. \tag{4.1}$$

Being quadratic, this equation can be immediately integrated [38]

$$\mathcal{W}_\tau^{\text{sat}}[\rho] \approx \mathcal{N}_\tau \exp\left\{ -\frac{\pi}{2} \int\limits_{-\infty}^{\tau} dy \int^{Q_s(y)} \frac{d^2 k_\perp}{(2\pi)^2} \frac{\rho_y^a(k_\perp)\rho_y^a(-k_\perp)}{k_\perp^2} \right\}, \tag{4.2}$$

which for convenience has been written as a functional of $\rho_y^a(k_\perp) = k_\perp^2 \alpha_y^a(k_\perp)$. The expression in the exponent is recognized as the (colour) Coulomb energy. As indicated by

---

[13] More precisely, we mean here the fields integrated over all rapidities, i.e., $\alpha(\boldsymbol{x}) = \int dy\,\alpha_y(\boldsymbol{x})$.

the upper limit on the integration over $\boldsymbol{k}$, Eq. (4.2) is valid only at saturation: for a given rapidity y (with y $\leq \tau$), this is to be used for the modes with momenta $k_\perp < Q_s(\text{y})$.

We thus see an interesting *duality* emerging at saturation: The strong field regime allows for a description in terms of a Gaussian weight function, so like the MV model. However, the 2–point function encoded in Eq. (4.2) is different from that in the MV model and, in contrast to the latter, it describes *real correlations* among the color sources. We have indeed:

$$\langle \rho_\text{y}^a(\boldsymbol{x})\, \rho_{\text{y}'}^b(\boldsymbol{y}) \rangle_\tau = \delta^{ab} \delta(\text{y} - \text{y}') \theta(\tau - \text{y})\, \lambda_\text{y}(\boldsymbol{x}, \boldsymbol{y}),$$
$$\lambda_\text{y}(\boldsymbol{k}) \simeq \frac{1}{\pi}\, \boldsymbol{k}^2\,, \qquad \text{for} \qquad k \ll Q_s(\text{y}). \tag{4.3}$$

The fact that the charge–charge correlator vanishes like $\boldsymbol{k}^2$ at small momenta has a strong influence on the infrared behaviour of the effective theory: Quantities which were found to be logarithmically divergent in the MV model (cf. Eqs. (2.49) or (2.38)) now become finite. Moreover, the would–be infrared divergence is cut off at $k \sim Q_s(\tau)$, which is a *hard scale* — $Q_s^2(\tau) \gg \Lambda_{\text{QCD}}^2$ — for high enough energy. In fact, with the correlation function in Eq. (4.3), all the gauge–invariant quantities that we have considered so far become infrared safe, and thus are insensitive to non–perturbative physics of confinement.

This is a fundamental property of the perturbative QCD evolution in the presence of non–linear effects, and demonstrates the internal consistency of the whole approach. It is because the high gluon density introduces a hard intrinsic scale in the problem and, moreover, the non–linear effects promote this scale into an effective infrared cutoff, that the perturbative approach to high–energy QCD is self–consistent, and represent a correct approach at least for sufficiently high energies.

We shall return to the physical interpretation of this smooth infrared behaviour in the next subsection, where we shall see that Eq. (4.3) is the expression of *color neutrality* [41, 42]. But before that, let us show that Eq. (4.3) implies the *saturation of the gluon distribution* at high energy.

We first compute the total density of colour charge squared in the transverse plane corresponding to saturated gluons with a given momentum $k_\perp \ll Q_s(\tau)$:

$$\mu_\tau(\boldsymbol{k}) \equiv \frac{1}{N_c^2 - 1} \frac{\langle \rho^a(\boldsymbol{k})\, \rho^a(-\boldsymbol{k}) \rangle_\tau}{\pi R^2} \simeq \int\limits_{\tau_s(k)}^\tau d\text{y}\, \lambda_\text{y}(\boldsymbol{k}). \tag{4.4}$$

The lower limit $\tau_s(k)$ in the above integral represents the rapidity at which the saturation scale becomes equal to the momentum $k_\perp$ of interest : $Q_s^2(\tau) = k_\perp^2$ for $\tau = \tau_s(k)$. In Sect. 4.2, we shall find that, to the accuracy of interest:

$$Q_s^2(\tau) \simeq Q_0^2\, \text{e}^{c\bar{\alpha}\tau}\,, \qquad c = 4.883...\,, \tag{4.5}$$

and therefore

$$\tau_s(k) \simeq \frac{1}{c\bar{\alpha}}\, \ln \frac{\boldsymbol{k}^2}{Q_0^2}\,. \tag{4.6}$$

111

(These relations are illustrated in Fig. 38.) Eqs. (4.3) and (4.4) imply [38, 41]

$$\mu_\tau(\boldsymbol{k}) \simeq \left(\tau - \tau_s(k_\perp)\right)\frac{k_\perp^2}{\pi}\,, \tag{4.7}$$

which is valid for $k \ll Q_s(\tau)$ or, equivalently, $\tau \gg \tau_s(k_\perp)$. In this regime, $\mu_\tau(k)$ grows only *linearly* with τ, in sharp contrast with the exponential behaviour of the corresponding quantity at large momenta $k_\perp \gg Q_s(\tau)$, as obtained from the BFKL equation [14] (1.54).

We conclude that *the color charge density saturates* because of the non–linear effects in the quantum evolution [38]. By contrast, in the MV model, the corresponding quantity μ_A increases indefinitely as a power of the atomic number A (the analog of the energy in the MV model), since the respective sources are uncorrelated. Due to quantum evolution, correlations arise already in the linear regime, via the BFKL equation. However it is only after including the non–linear effects associated with gluon recombination that these correlations are such as to limit the growth of the color charge density with τ.

Since the density of the color sources saturate with τ, so does also the *gluon distribution* radiated by these sources. This is already obvious if we use the *linear* approximation, Eq. (2.35), for the classical field solution, which implies $\varphi_\tau(\boldsymbol{k}) \simeq \mu_\tau(\boldsymbol{k})/\boldsymbol{k}^2$, and therefore (cf. Eqs. (4.6)–(4.7))

$$\varphi_\tau(\boldsymbol{k}) \simeq \frac{\tau - \tau_s(k_\perp)}{\pi} \simeq \frac{1}{\pi c\bar{\alpha}} \ln\frac{Q_s^2(\tau)}{k_\perp^2}\,, \tag{4.8}$$

which is valid at low momenta $k_\perp \ll Q_s(\tau)$ [37, 38]. A more careful calculation, based on the non–linear solution (2.13), shows that the correct answer for $\varphi_\tau(\boldsymbol{k})$ differs from Eq. (4.8) only by a numerical factor $\delta \lesssim 1$ [38, 42]. This factor cannot be computed in the present approximations, as it is sensitive to the transition regime at $k_\perp \sim Q_s(\tau)$. This overall factor will not affect the salient features of Eq. (4.8), namely its dependence upon the energy and the transverse momentum.

Note that the saturation of the gluon distribution is *not complete* : $\varphi_\tau(\boldsymbol{k})$ still rises with τ, albeit only slowly (logarithmically in the energy) [37]. This slow increase can be attributed to the expansion of the hadron in longitudinal direction when including new modes with lower values of k^+. (Indeed, $\tau - \tau_s(k_\perp) \approx \tau$ is the longitudinal extent of the hadron in units of space–time rapidity; cf. Sect. 3.1 and Fig. 32.)

In addition to τ–*saturation* (the linear increase with τ), Eq. (4.8) shows also $k_\perp$–*saturation*: the $k_\perp$ spectrum is only logarithmic in $1/k_\perp$ at low momenta. Recall that, in the classical MV model, a similar spectrum emerged (see Eq. (2.56)) only after fully taking into account the non-linear effects in the classical Yang–Mills equations (2.1). By contrast, in the quantum case, the non-linear effects responsible for saturation have already been encoded in the distribution of the color sources. This is natural since, in the presence of evolution, the difference between "sources" and "small–x gluons" is only a matter of convention and depends upon the resolution scale in k^+.

---

[14] Note that, in the linear regime at high $k_\perp$, $\mu_\tau(\boldsymbol{k})$ is linearly related to the unintegrated gluon distribution $f(\tau, \boldsymbol{k})$.

To conclude, the gluon occupation factor at saturation is parametrically of order $1/\bar{\alpha}$ and is given by the same *universal* function of the ratio $k_\perp^2/Q_s^2(\tau)$ as in the MV model. This is *universal* in the sense that Eq. (4.8) is independent of either the initial condition at low energy, or the details of the evolution leading to saturation, except for the corresponding dependencies of the saturation scale itself. It furthermore exhibits a remarkable scaling property — it depends upon τ and $k_\perp$ only through the ratio $z \equiv k_\perp^2/Q_s^2(\tau)$ *(geometric scaling)* — which reflects the fact that the saturation momentum is the only intrinsic scale at low momenta in an infrared–safe theory.

4.1.2   Color neutrality through saturation

We have mentioned before the remarkable property of the charge–charge correlator in Eq. (4.3) to rapidly vanish as $k_\perp \to 0$. In what follows, we shall argue that this behaviour has a simple physical interpretation: the color sources are correlated over long distances in such a way that *color neutrality* is achieved over a transverse area of order $1/Q_s^2(\tau)$ [41, 42].

Recall first that, in a low energy hadron, color screening is due to confinement, and thus the typical correlation length among color charges — the color screening length — is $\Lambda_{\rm QCD}^{-1} \sim 1$ fm. At high energy (small x), partons are much more densely packed, so one may expect color neutralization to occur over shorter distances. As we shall explain now, the BFKL evolution already introduces correlations which reduce the total color charge within a small area $\Delta\Sigma \ll \Lambda_{\rm QCD}^{-2}$, without however providing a perturbative scale for color neutralization. Such a scale emerges only in the presence of saturation, and is of the order of $Q_s^{-1} \ll \Lambda_{\rm QCD}^{-1}$.

To see this, we shall consider the total color charge $\mathcal{Q}^a$ enclosed within a surface $\Delta\Sigma$, as given by Eq. (1.45). This is a random quantity with zero average (since $\langle \rho^a(\vec{x}) \rangle = 0$ at any point $\vec{x}$), so we shall rather compute the average color charge squared $\mathcal{Q}^2 = \mathcal{Q}^a\mathcal{Q}^a$. We have

$$\langle \mathcal{Q}^2 \rangle_\tau = (N_c^2 - 1) \int_{\Delta\Sigma} d^2\boldsymbol{x} \int_{\Delta\Sigma} d^2\boldsymbol{y}\, \mu_\tau(\boldsymbol{x}, \boldsymbol{y})\,. \tag{4.9}$$

In the MV model, where the color sources are uncorrelated (cf. Eq. (1.46)), we deduce

$$\langle \mathcal{Q}^2 \rangle_A = (N_c^2 - 1)\Delta\Sigma\, \mu_A \sim \frac{1}{\alpha}\Delta\Sigma\, Q_A^2\,, \tag{4.10}$$

which does not vanish not even when $\Delta\Sigma$ covers the whole nuclear area ($\Delta\Sigma = \pi R_A^2$), since the correlations due to confinement are not explicit in this model. Rather, non–perturbative color screening at the scale $\Lambda_{\rm QCD}^{-1}$ is introduced by hand, whenever a calculation is meeting with infrared divergences. Besides, the square charge *density* $\langle \mathcal{Q}^2 \rangle/\Delta\Sigma \sim Q_A^2$ increases rapidly with A, like $A^{1/3}$.

After including quantum evolution, the charge correlator acquires a non–trivial momentum dependence synonymous of correlations ($\mu_A \to \mu_\tau(k_\perp)$), and Eq. (4.9) can be estimated as (up to a color factor):

$$\langle \mathcal{Q}^2 \rangle_\tau \sim \Delta\Sigma\, \mu_\tau(Q^2) \qquad \text{with} \qquad Q^2 \sim 1/\Delta\Sigma\,. \tag{4.11}$$
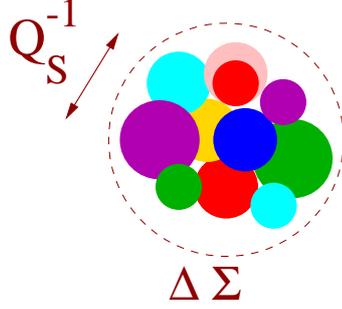
Fig. 35. A perturbative area over which color neutrality is achieved due to saturation. CNEUTRAL

So long as we look at distances smaller than the saturation length, $Q^2 \gg Q_s^2(\tau)$, and $\mu_\tau(Q^2)$ can be estimated from the BFKL equation (1.54). As we shall soon discover (see Sect. 4.2), the solution to the BFKL equation implies :

$$\langle \mathcal{Q}^2 \rangle_\tau \sim \frac{1}{\alpha} \left(\Delta\Sigma \, Q_s^2(\tau) \right)^\gamma \qquad \gamma = 0.627... . \tag{4.12}$$

This shows some color shielding — since $\gamma < 1$, so that the total color charge squared rises slower than the area —, which is however *incomplete* : $\langle \mathcal{Q}^2 \rangle$ still rises as a power of $\Delta\Sigma$ and, besides, the charge density $\langle \mathcal{Q}^2 \rangle / \Delta\Sigma$ increases exponentially with τ.

For larger surfaces, of the order of the saturation disk $1/Q_s^2(\tau)$ or larger, one should rather us Eq. (4.7), which gives:

$$\langle \mathcal{Q}^2 \rangle_\tau \sim \frac{1}{\alpha} \ln \left(\Delta\Sigma \, Q_s^2(\tau) \right). \tag{4.13}$$

Although, strictly speaking, non–zero, this total charge squared is only *logarithmically* increasing with both $1/x$ and $\Delta\Sigma$, that is, it remains essentially constant when adding new sources by either increasing the area or rising the energy. This means that the newly added color sources are almost totally screened by the other ones (*quasi–complete shielding*).

In particular, the total charge (4.13) is much smaller than it would be for a system of uncorrelated color sources, cf. Eq. (4.10), so we can indeed speak about *color neutrality.* A pictorial interpretation of this phenomenon is given in Fig. 35. Note, however, that unlike the non–perturbative mechanism of confinement, which completely washes out *all* the color correlations over distances $\gg \Lambda_{\mathrm{QCD}}^{-1}$, the perturbative screening associated with saturation neutralizes only the color charge, but not also its higher multipolar moments: the color field created by gluon sources at large distances $\gg 1/Q_s(\tau)$ is essentially a *dipolar* field [42].

SECT_BD

4.1.3 Unitarity: Towards the black disk limit

Consider now the high–energy limit of the scattering amplitudes, as predicted by the Balitsky equations (cf Sect. 3.5). These equations form an infinite hierarchy, so in general it is not clear how to approach their respective solutions. However, in the high energy regime of interest here, the color fields in the target are strong, and we can reasonably assume that the correlation functions factorize for gauge–invariant objects. In particular,

114

by assuming that

$$\langle \text{tr}(V_{\boldsymbol{x}}^{\dagger} V_{\boldsymbol{y}}) \text{tr}(V_{\boldsymbol{z}}^{\dagger} V_{\boldsymbol{y}}) \rangle_{\tau} \approx \langle \text{tr}(V_{\boldsymbol{x}}^{\dagger} V_{\boldsymbol{y}}) \rangle_{\tau} \langle \text{tr}(V_{\boldsymbol{z}}^{\dagger} V_{\boldsymbol{y}}) \rangle_{\tau}, \tag{4.14}$$

one immediately finds that the first Balitsky equation (3.88) reduces to a *closed* equation for the dipole–CGC S–matrix, Eq. (3.85). Namely:

$$\frac{\partial}{\partial \tau} \langle S_{\boldsymbol{xy}} \rangle_{\tau} = \frac{\bar{\alpha}_s}{2\pi} \int_{\boldsymbol{z}} \mathcal{M}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \left\{ - \langle S_{\boldsymbol{xy}} \rangle_{\tau} + \langle S_{\boldsymbol{xz}} \rangle_{\tau} \langle S_{\boldsymbol{zy}} \rangle_{\tau} \right\}, \tag{4.15}$$

which is the same as the BK equation (1.67) originally deduced by Kovchegov for the case of dipole–nucleus scattering [124]. Note that the mean field approximation (4.14) amounts to assuming that the two dipoles $(\boldsymbol{x}, \boldsymbol{z})$ and $(\boldsymbol{z}, \boldsymbol{y})$ scatter *independently* off the color fields in the target. This is indeed reasonable so long as the color sources in the target are uncorrelated with each other, as for a large nucleus in the MV approximation, or, more generally, when their only non–trivial correlation is a 2–point function, as for the saturated gluons (cf. Eq. (4.3)).

The latter case is appropriate for the study of the high energy limit. Namely, we shall assume that the energy is large enough for the saturation scale $Q_s(\tau)$ at the impact parameter of the dipole to be much larger than the typical transverse momentum $k_{\perp} \sim 1/r$, with $r \equiv |\boldsymbol{x} - \boldsymbol{y}|$, carried by the dipole. In other terms, the incoming dipole is large relative to the screening length in the target, $r \gg 1/Q_s(\tau)$, and therefore it scatters predominantly off saturated gluons. Then, by the 'random phase' argument discussed previously, we expect the corresponding S–matrix to be small: $\langle S_{\boldsymbol{xy}} \rangle_{\tau} \ll 1$. Moreover, a similar condition should hold for the two dipoles emerging after splitting, since their typical sizes are also much larger than $1/Q_s(\tau)$, as we shall shortly discover. The avove argument also implies that the target saturation momentum $Q_s(\tau)$ fixes the scale at which the dipole scattering makes the transition from weak ($r \ll 1/Q_s$) to strong ($r \gg 1/Q_s$).

Under these assumptions, we can neglect the second term, quadratic in $\langle S \rangle_{\tau}$, within the braces in Eq. (4.15), and thus obtain a simplified equation, which is linear :

$$\frac{\partial}{\partial \tau} \langle S_{\boldsymbol{xy}} \rangle_{\tau} \approx - \frac{\bar{\alpha}_s}{2\pi} \int_{\boldsymbol{z}} \frac{(\boldsymbol{x} - \boldsymbol{y})^2}{(\boldsymbol{x} - \boldsymbol{z})^2 (\boldsymbol{z} - \boldsymbol{y})^2} \langle S_{\boldsymbol{xy}} \rangle_{\tau} . \tag{4.16}$$

Note that the linear term which subsists in the r.h.s. of Eq. (4.16) is simply the probability that the incoming dipole survive without splitting after one step of evolution (also known as the "Sudakov factor"). The integral over $\boldsymbol{z}$ in the last equation has uncompensated poles at $\boldsymbol{z} = \boldsymbol{x}$ and $\boldsymbol{z} = \boldsymbol{y}$; this is an artifact of the above approximation (in the original equation (4.15), such singularities cancel in between the linear and the quadratic terms), which applies only so long as the emitted dipoles are much larger than $1/Q_s(\tau)$. Thus, the logarithmic singularities in Eq. (4.16) must be cut off at dipole sizes of the order of the saturation length. Then, the r.h.s. of Eq. (4.16) is dominated by the logarithmic regions of integration where one has either $1/Q_s^2 \ll (\boldsymbol{x} - \boldsymbol{z})^2 \ll r^2$ or $1/Q_s^2 \ll (\boldsymbol{z} - \boldsymbol{y})^2 \ll r^2$. One easily finds (with $S_{\tau}(r) \equiv \langle S_{\boldsymbol{xy}} \rangle_{\tau}$)

$$\frac{\partial}{\partial \tau} \ln S_{\tau}(r) \simeq -\bar{\alpha}_s \int_{Q_s^{-2}(\tau)}^{r^2} \frac{dz^2}{z^2} = -\bar{\alpha}_s \ln[Q_s^2(\tau) r^2], \tag{4.17}$$

LT00

115

which together with $\ln[Q_s^2(\tau)r^2] = c\bar{\alpha}_s(\tau - \tau_s)$ (cf. Eqs. (4.5) and (4.6)) immediately implies [38, 41, 42, 162]

$$S_\tau(r) = \exp\left\{-\frac{c}{2}\,\bar{\alpha}_s^2(\tau - \tau_s)^2\right\} S_{\tau_s}(r)\,. \tag{4.18}$$

In this equation, $\tau_s(r)$ is such that $Q_s(\tau_s) \sim 1/r$, and thus $S_{\tau_s}(r) \sim 1$. Eq. (4.18) can be alternatively rewritten as:

$$S_\tau(r) = \exp\left\{-\frac{1}{2c}\,\ln^2\left[Q_s^2(\tau)r^2\right]\right\} S_{\tau_s}(r)\,. \tag{4.19}$$

From either Eq. (4.18) or Eq. (4.19) it is clear that, for sufficiently large energies (such that $\tau \gg \tau_s(r)$) or, equivalently, for sufficiently large dipole sizes ($r \gg 1/Q_s(\tau)$), the S–matrix approaches rapidly to the "black–disk limit" $S = 0$. Equivalently, the scattering amplitude T approaches the unitarity limit $T = 1$. Since $|S|^2$ has the meaning of the probability for the dipole to survive after scattering (at a given impact parameter), it is clear that $S \to 0$ corresponds to the total absorbtion of the dipole by the CGC. Note also that, although the approach towards the black disk limit in Eqs. (4.18)–(4.19) is quite fast, this is nevertheless much slower than that predicted by the naive Glauber exponentiation of the single scattering amplitude, as, e.g., in the Golec-Biernat and Wüsthoff formula [73]. The difference between the above results and the corresponding ones in the MV model (where the Glauber exponentiation holds indeed; see Eq. (2.49)) stems from the presence of long–range correlations among the saturated color sources, as described by Eq. (4.3). It is interesting to note that the τ–dependence in Eq. (4.18) is consistent with a numerical calculation by Salam [44], based on the dipole picture [43, 45].

Consider similarly the high energy limit of the equation (3.98) describing the evolution of the odderon amplitude. Near the unitarity limit, $T \simeq 1$, Eq. (3.98) simplifies to (with $O_\tau(\boldsymbol{x}, \boldsymbol{y}) \equiv \langle O(\boldsymbol{x}, \boldsymbol{y})\rangle_\tau$)

$$\frac{\partial}{\partial\tau}\ln O_\tau(\boldsymbol{x}, \boldsymbol{y}) \simeq -\bar{\alpha}_s\int_{Q_s^{-2}(\tau)}^{r^2}\frac{dz^2}{z^2} = -\bar{\alpha}_s\ln[Q_s^2(\tau)r^2]\,, \tag{4.20}$$

which by the same manipulations as above implies:

$$O_\tau(\boldsymbol{x}, \boldsymbol{y}) \simeq \exp\left\{-\frac{c}{2}\,\bar{\alpha}_s^2(\tau - \tau_s)^2\right\} O_{\tau_s}(\boldsymbol{x}, \boldsymbol{y})\,. \tag{4.21}$$

Thus, unlike the pomeron exchanges, which appear to unitarize at the maximally allowed value $T = 1$, the odderon exchanges are rapidly suppressed with increasing energy.

4.2 Saturation momentum

In the previous section, we have investigated some properties of the non–linear regime at low transverse momenta $k_\perp \lesssim Q_s$, or relatively large dipole sizes, $r \gtrsim 1/Q_s$, but the value of the saturation momentum — which is the boundary of this regime — remained unspecified. In what follows, we shall rather focus on the linear regime, corresponding

to small dipoles, $r \ll 1/Q_s$, or relatively large momenta, $k_\perp \gg Q_s$, and perform an analysis of this linear regime based on the BFKL equation, from which we shall compute the saturation momentum — or, more precisely, its evolution with increasing energy. Essentially, we shall obtain the saturation momentum $Q_s(\tau)$ by 'unitarizing the single Pomeron exchange approximation to the dipole scattering' [22, 37, 51, 52], that is, from the condition that the BFKL solution for the dipole amplitude $\langle T(\boldsymbol{x}, \boldsymbol{y})\rangle_\tau$ become of order one when $r \equiv |\boldsymbol{x} - \boldsymbol{y}| \sim 1/Q_s(\tau)$. There are several implicit assumptions in this construction whose validity is not *a priori* obvious — like the possibility to extrapolate the linear solution towards the non–linear regime, or even the validity of the BFKL approximation for the (average) amplitude at weak scattering — and to which we shall return later on. Here, it suffices to say that our subsequent calculation of Q_s is equivalent [164] (within the limits of its accuracy) to solving the non–linear BK equation (1.67), whose direct resolution would be a considerably more complicated task, though.

The solution $\langle T(\boldsymbol{x}, \boldsymbol{y})\rangle_\tau$ to BK equation, that we shall simply denote as $T_\tau(\boldsymbol{x}, \boldsymbol{y})$ (the parentheses being redundant in this mean field approximation) is generally *inhomogeneous*: in addition to the dipole size $\boldsymbol{r} \equiv \boldsymbol{x} - \boldsymbol{y}$, it also depends upon its impact parameter $\boldsymbol{b} \equiv (\boldsymbol{x} + \boldsymbol{y})/2$. The inhomogeneity is introduced by the initial conditions at low energy and is further modified by the high–energy evolution. Accordingly, the saturation momentum is also a function of $\boldsymbol{b}$, $Q_s^2 = Q_s^2(\tau, \boldsymbol{b})$, as expected, since the gluon distribution in a hadron is generally inhomogeneous. However, an explicit study of this $\boldsymbol{b}$–dependence would be both cumbersome and delicate, and goes beyond the purpose of our analysis. This is cumbersome since it requires the use of the full solution to the 'non–forward' BFKL equation, including its $\boldsymbol{b}$–dependence, and is delicate since the BFKL appproximation (more generally, perturbation theory) can anyway not be trusted in the low–density regime towards the edge of the hadron disk, where the physical role of the inhomogeneity is particularly interesting (see the discussion in Sect. 4.4).

Fortunately, the inhomogeneity in $\boldsymbol{b}$ is not very important for the problem in which we are currently interested, namely the energy dependence of the saturation momentum at a given impact parameter. This is so since the evolution near saturation is biased towards small dipole sizes (or large transverse momenta) : when increasing τ along the saturation line, we at the same time decrease r (or increase $\rho \equiv \ln(1/r^2)$), in such a way to keep the amplitude (or the gluon occupation number) *constant*, cf. Fig. 8. To study the approach towards saturation at a given $\boldsymbol{b}$, we need the gluon distribution within a small disk with radius $r \sim 1/Q_s(\tau, \boldsymbol{b})$ centered at $\boldsymbol{b}$. The evolution being non–local, this distribution can be correlated with that at points $\boldsymbol{b}'$ far outside this disk. However, such long–range correlations are strongly reduced by saturation, via the phenomenon of 'colour neutrality' discussed in Sect. 4.1.2 : A region which, at some intermediate step τ' in the evolution, with $0 < \tau < \tau'$, has a saturation scale $Q_s(\tau', \boldsymbol{b}')$, will further evolve by predominantly emitting *smaller* gluons, with momenta $k_\perp \gg Q_s(\tau', \boldsymbol{b}')$. Hence, if the separation $|\boldsymbol{b}' - \boldsymbol{b}|$ is larger than $1/Q_s(\tau', \boldsymbol{b}')$, then this further evolution at $\boldsymbol{b}'$ cannot influence the physics at $\boldsymbol{b}$. Thus, due to saturation, domains which are relatively far away from each other do rapidly decouple in the course of the evolution, so that the approach towards saturation, or 'black disk' limit, can be characterized as *quasi–local* in $\boldsymbol{b}$ (the precise range of non–locality being fixed by the BFKL diffusion in the presence of saturation; see below).

So, in this section and in many of the subsequent ones, we shall ignore the b–dependence of the evolution equations, and write the dipole amplitude simply as $T_\tau(\boldsymbol{x}, \boldsymbol{y}) = T_\tau(r)$, where we also assume target *isotropy* (i.e., independence upon the orientations of the 2–dimensional vector $\boldsymbol{r}$). These assumptions are consistent with the structure of BK equation (1.67), since the (dipole) kernel in this equation is invariant under 2–dimensional translations and rotations.

In particular, the linearized version of this equation, cf. Eq. (1.65), is the 'forward' BFKL equation in coordinate space:

$$\frac{\partial}{\partial \tau} T_\tau(\boldsymbol{x} - \boldsymbol{y}) = \frac{\bar{\alpha}_s}{2\pi} \int_{\boldsymbol{z}} \frac{(\boldsymbol{x} - \boldsymbol{y})^2}{(\boldsymbol{x} - \boldsymbol{z})^2 (\boldsymbol{z} - \boldsymbol{y})^2} \\ \left\{ - T_\tau(\boldsymbol{x} - \boldsymbol{y}) + T_\tau(\boldsymbol{x} - \boldsymbol{z}) + T_\tau(\boldsymbol{z} - \boldsymbol{y}) \right\}, \tag{4.22}$$

which is valid (at least within the limitations of the mean field approximation: see Sect. 7 for a discussion of the role of fluctuations) so long as $T_\tau(r) \ll 1$. In what follows, we shall solve this equation and then deduce the saturation momentum from the condition

$$T_\tau(r) = \kappa \qquad \text{for} \qquad r = 1/Q_s(\tau), \tag{4.23}$$

where $\kappa < 1$ is a number of order one (its precise value does not matter to the accuracy of interest). As anticipated, this procedure requires an extrapolation of the BFKL solution towards the strong–scattering regime where $T_\tau(r) \sim \mathcal{O}(1)$, whose justification will be given later on.

4.2.1 The Mellin representation of the BFKL amplitude

The solution to Eq. (4.22) is most conveniently obtain by going to Mellin space, via the following transformation

$$T_\tau(r) = \int_C \frac{\mathrm{d}\gamma}{2\pi i} \left(\frac{r^2}{r_0^2}\right)^\gamma \mathcal{T}_\tau(\gamma), \tag{4.24}$$

where the complex variable γ runs along a contour C to be shortly specified and r_0 is an arbitrary reference scale, which drops out in the complete result (see below). This is convenient since the powers of r^2 are eigenfunctions of the BFKL kernel:

$$\mathcal{K}_{\mathrm{BFKL}} \otimes r^{2\gamma} = \bar{\alpha} \chi(\gamma) r^{2\gamma}, \tag{4.25}$$

where in the l.h.s. we use an operator notation for the convolution with the BFKL kernel (this is the same as the r.h.s. of Eq. (4.22) with $T_\tau(r) \to r^{2\gamma}$), and the eigenvalue $\chi(\gamma)$ reads ($\psi(\gamma)$ is known as the di-gamma function):

$$\chi(\gamma) = 2\psi(1) - \psi(\gamma) - \psi(1 - \gamma), \qquad \psi(\gamma) \equiv \mathrm{d} \ln \Gamma(\gamma)/\mathrm{d}\gamma, \tag{4.26}$$

and has simple poles at all the integer values of γ. Therefore, Eq. (4.22) becomes local in Mellin space, where it can be easily solved to give

118
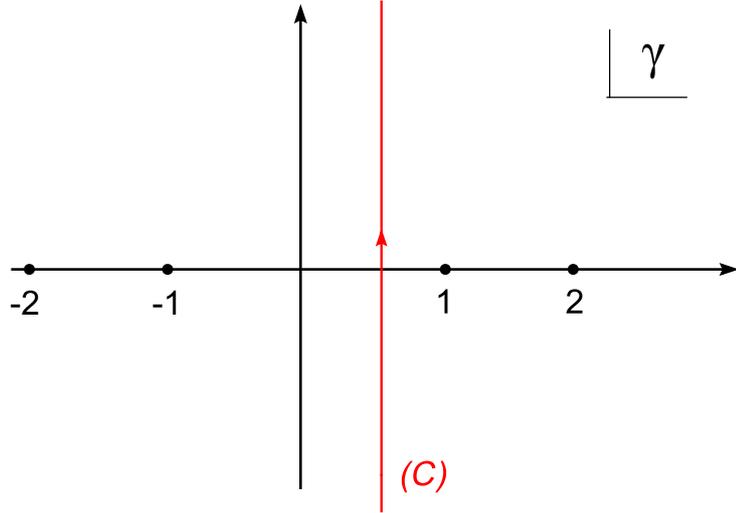
$$T_\tau(r) \;=\; \int_C \frac{\mathrm{d}\gamma}{2\pi i} \left(\frac{r^2}{r_0^2}\right)^\gamma \mathrm{e}^{\bar{\alpha}\tau\chi(\gamma)}\,\mathcal{T}_0(\gamma), \tag{4.27}$$

with $\mathcal{T}_0(\gamma)$ the Mellin transform of the initial condition at $\tau = 0$:

$$\mathcal{T}_0(\gamma) \;=\; \int_0^\infty \frac{\mathrm{d}^2 r}{r^2} \left(\frac{r_0^2}{r^2}\right)^\gamma T_0(r)\,. \tag{4.28}$$

Note that $\mathcal{T}_0(\gamma)$ is also a function of the arbitrary scale r_0, but this dependence goes away, as anticipated, when inserting Eq. (4.28) into the integrand of Eq. (4.27). The physical reference scale for r is rather introduced by the initial condition, as we now discuss.

The precise form is the initial amplitude $T_0(r)$ is not very important, so long as it satisfies two important physical requirements: it vanishes like r^2 when $r \to 0$ ('color transparency') and it obeys the unitarity bound $T_0(r) \leq 1$ — meaning that there is some intrinsic scale Q_0, characteristic of the target, such that the amplitude stops growing with r when $r \gtrsim 1/Q_0$. For instance, for a nuclear target, Q_0 is the nuclear saturation momentum $Q_s(A)$, cf. Eq. (2.54), and is moreover a 'hard' scale, $Q_s(A) \gg \Lambda_{\mathrm{QCD}}$, when A is large enough. This is convenient, as it justifies to use of perturbation theory all the way from low top high energies. Another example of that type is when the target itself is a small dipole, of size $r_t \ll 1/\Lambda_{\mathrm{QCD}}$; then, $Q_0 = 1/r_t$ and the dipole–dipole scattering is never strong : the amplitude $T_0(r, r_t)$ takes a maximum value of $\mathcal{O}(\alpha^2)$ when $r \sim r_t$ (see Sect. 6). Given these general properties of T_0, it can be easily checked that the integral in Eq. (4.28) is absolutely convergent for any complex γ such that $0 < \mathrm{Re}\,\gamma < 1$ and is moreover dominated by $r \sim 1/Q_0$. Hence, after the integration over γ, Q_0 will supplant r_0 as the natural sale for r in the solution, as anticipated.

We choose the contour C in Eq. (4.27) as a straight line parallel to the imaginary axis which crosses the real axis at $0 < \mathrm{Re}\,\gamma < 1$ (see Fig. 4.2.1). Then we can write

$$T_\tau(r) \;=\; \int_C \frac{d\gamma}{2\pi i} \left(r^2 Q_0^2\right)^\gamma \mathrm{e}^{\bar{\alpha}\tau\chi(\gamma)}\,\mathcal{T}_0(\gamma) \;=\; \int_C \frac{d\gamma}{2\pi i}\, \mathrm{e}^{F(\gamma|\rho,\tau)}\,\mathcal{T}_0(\gamma), \tag{4.29}$$

119
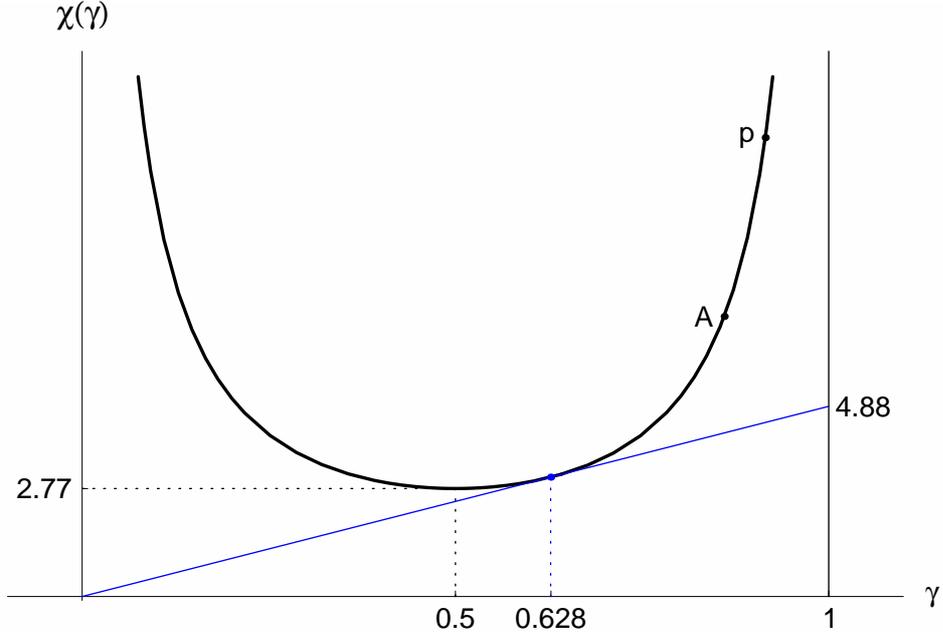
Fig. 36. *The BFKL eigenvalue $\chi(\gamma)$ and the graphical solution to the saturation problem* The value $\gamma_s = 0.627$ corresponds to the saturation saddle point and the value $\chi(\gamma_s)/\gamma_s = 4.88$ determines the asymptotic energy dependence of the saturation momentum (for comparison, the "hard pomeron" saddle point $\gamma_0 = 1/2$ and its intercept $\omega_0 = 4\ln 2 = 2.77$ are shown). The points p and A correspond to the proton and the nucleus respectively, for the same transverse momentum and rapidity, when they are far above the saturation scale (see Sect. 5.2). CHI

where $\mathcal{T}_0(\gamma)$ has poles [15] at $\gamma = 0$ and $\gamma = 1$ and in writing the second equality we introduced the logarithmic variable $\rho \equiv \ln(1/r^2 Q_0^2)$ and the function

$$F(\gamma|\rho, \tau) \equiv -\gamma\rho + \bar{\alpha}\tau\chi(\gamma). \tag{4.30}$$

The weak scattering regime corresponds to $r \ll 1/Q_s(\tau)$, with $Q_s(\tau) \gg Q_0$ (for sufficiently large τ), as we shall shortly see. Hence, $r^2 Q_0^2 \ll 1$ and ρ is typically much larger than one. For what follows it is also convenient to introduce the 'evolution time' $t \equiv \bar{\alpha}\tau$.

4.2.2 The BFKL Pomeron

At large energy $t \equiv \bar{\alpha}\tau \gg 1$, one can evaluate the integral in Eq. (4.29) via a saddle point approximation. The saddle point depends upon the value of ρ, but it always lies along the real axis, so it is useful to notice the behaviour of the function $\chi(\gamma)$ for real γ in between 0 and 1, as illustrated in Fig. 36. In the *formal* high energy limit $t \to \infty$ at fixed ρ (i.e., $\rho/t \to 0$), the saddle point lies at $\gamma = \gamma_0 \equiv 1/2$ (the 'Pomeron' saddle point), where the function $\chi(\gamma)$ has a minimum. We can then expand around this minimum,

---

[15] Indeed, when $\gamma = 0$ (respectively, $\gamma = 1$), the integral in Eq. (4.28) develops logarithmic divergences coming from the integration over large (respectively, small) values of $r_\perp$.

$$\chi(\gamma) \approx \omega_0 + \frac{\beta_0}{2}\left(\gamma - \frac{1}{2}\right)^2, \quad \omega_0 = 4\ln 2 \approx 2.77, \quad \beta_0 = 28\zeta(3) \approx 33.67, \qquad (4.31)$$

and perform the ensuing Gaussian integration, to finally obtain

$$T(\rho, t) \simeq \mathrm{e}^{\omega_0 t - \frac{1}{2}\rho} \, \frac{\exp\left\{-\frac{\rho^2}{2\beta_0 t}\right\}}{\sqrt{2\pi\beta_0 t}} \, \mathcal{T}_0(1/2). \qquad (4.32)$$

This is the asymptotic form of the BFKL solution at high energy. As already discussed in Sect. 1.4 (cf. Eq. (1.57)), this solution develops conceptual problems in the high–energy limit that it is supposed to describe: it violates unitarity for $\tau > \tau_0(\rho)$ with, e.g.,

$$\tau_0(0) \sim \frac{1}{\omega_0 \bar{\alpha}} \ln \frac{1}{\mathcal{T}_0}, \qquad (4.33)$$

(with $\mathcal{T}_0 \equiv \mathcal{T}_0(1/2)$) and, moreover, it exhibits 'infrared diffusion', meaning that the solution involves an excursion through soft momenta, or large dipole sizes, at intermediate stages, for which the unitarity bound is violated and perturbation theory fails to apply. As we understand by now, such problems reflect the fact that the 'Pomeron' approximation (4.32) applies, at most, within a limited rapidity range $\tau \lesssim \tau_0$, where the amplitude remains much smaller than one. In order to stay in the linear regime when increasing τ we have to ensure that $r \ll 1/Q_s(\tau)$, or $\rho > \rho_s(\tau)$, for any τ. Here, $\rho_s(\tau) \equiv \ln[Q_s^2(\tau)/Q_0^2]$ is the value of the saturation momentum in these logarithmic units. Thus, the ratio ρ/t cannot be smaller than the corresponding value at saturation $\rho_s(t)/t$, which, as we shall shortly see, is $\mathcal{O}(1)$ for sufficiently large $\tau \gg \tau_0$. Accordingly, when increasing energy, we must simultaneously increase ρ, in such a way that the condition $\rho > \rho_s(\tau)$ remain satisfied (see Fig. 4.2.2). In particular, in order to determine the locus of the saturation line, we must follow an evolution which remains close to this line, with ρ always greater than ρ_s though. The saddle point describing this evolution will be determined in the next subsection.

4.2.3   The saturation exponent

The saturation line is, by definition, a line in the (ρ, τ) plane along which the dipole amplitude $T_\tau(r) \equiv T(\rho, \tau)$ is constant and of $\mathcal{O}(1)$. As we shall shortly see, for sufficiently high energy, the lines of constant amplitude $T(\rho, \tau) = \kappa$ for the solution $T(\rho, \tau)$ to the BFKL equation are *straight lines*, whose slope is independent of κ. Although this result will be here derived only for the weak–scattering regime where $\kappa \ll 1$, we shall eventually extrapolate it to values $\kappa \sim \mathcal{O}(1)$ and thus identify the common slope of these lines with the saturation exponent: $\rho_s(\tau) \simeq \lambda_s\tau$ for $\tau \gg \tau_0$. A justification for such an extrapolation will be presented later on, in Sect. 4.3.2. Here, τ_0 is the rapidity at which unitarity corrections become important on the resolution scale fixed by the original dipole 'size' $r_0 = 1/Q_0$, i.e., for $\rho = 0$. As manifest on Eq. (4.33), this critical rapidity is independent of r_0 and of the target size, but it depends upon the value of the original amplitude at low energy. For a dipole target, we have $\mathcal{T}_0 \sim \alpha^2$ and hence $\tau_0 \sim (1/\omega_0\bar{\alpha})\ln(1/\alpha^2)$.

Let us return to the Mellin representation of the BFKL amplitude, cf. Eq. (4.29), and
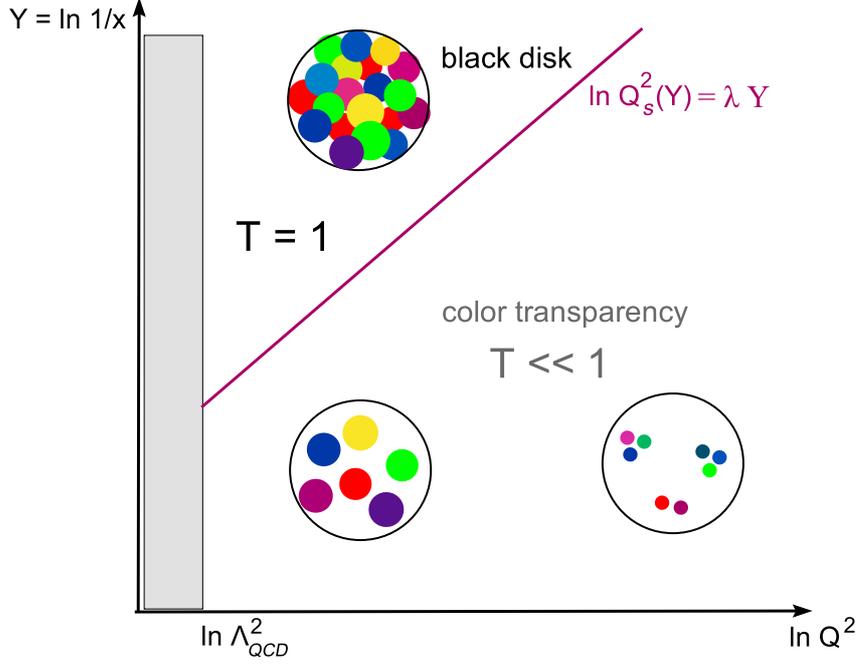
specialize to dipole–dipole scattering ($\mathcal{T}_0 \sim \alpha^2$), for definiteness. In the saddle point approximation, valid for $t \gg 1$ (and hence also $\rho \gg 1$), the line $\rho_\kappa(t)$ of constant amplitude $T(\rho_\kappa(t), t) = \kappa$ is approximately determined by the two following conditions

$$
F'(\gamma_\kappa | \rho_\kappa, t) = 0 \quad \Longrightarrow \quad \chi'(\gamma_\kappa) = \frac{\rho_\kappa}{t},
$$

$$
F(\gamma_\kappa | \rho_\kappa, t) = \ln \frac{\kappa}{\alpha^2} \quad \Longrightarrow \quad \chi(\gamma_\kappa) = \frac{1}{t}\left(\gamma_\kappa \rho_\kappa + \ln \frac{\kappa}{\alpha^2}\right). \tag{4.34}
$$

The first line (where the prime on F denotes a derivative w.r.t. γ) is the condition that γ_κ be the relevant saddle point. The second line is the condition that $\ln T \approx \ln \kappa$ along the line $\rho_\kappa(t)$. Notice that the dependence upon κ is very weak, since merely logarithmic. Thus one can let $\kappa \to 1$ in the previous formulæand thus deduce the *saturation saddle point* γ_s and the *saturation line* $\rho_s(t)$ [22, 51, 52] :

$$
\gamma_s = \frac{\chi(\gamma_s)}{\chi'(\gamma_s)} \simeq 0.627...,
$$

$$
\rho_s(t) \simeq c_s t - \frac{\ln(1/\alpha^2)}{\gamma_s} \quad \text{with} \quad c_s \equiv \frac{\chi(\gamma_s)}{\gamma_s} \approx 4.883. \tag{4.35}
$$

We have assumed here that τ, and hence ρ_s, are large enough for the term $\ln(1/\alpha^2)$ to be treated as a small perturbation; this essentially amounts to $\tau \gg \tau_0$. Note that γ_s is a pure number, independent of either τ or $\bar{\alpha}$, corresponding to the fact that $\rho_s(\tau)$ is a *straight* line, as anticipated, with slope $\lambda_s = c_s \bar{\alpha}$. This in turn implies that the saturation momentum increases exponentially with rapidity, i.e., as a power of s :

$$
Q_s^2(\tau) \simeq (\alpha^2)^{1/\gamma_s} Q_0^2 \, e^{c_s \bar{\alpha} \tau}, \qquad c_s = 4.883.... \tag{4.36}
$$

The prefactor α^{2/γ_s} is specific to dipole–dipole scattering and, more generally, to a situation in which the scattering started by being weak ($\mathcal{T}_0 \sim \alpha^2$) at low energy. In such a case, Eq. (4.36) shows that it takes, roughly, a rapidity evolution $\tau \sim [1/\chi(\gamma_s)\bar{\alpha}]\ln(1/\alpha^2) \sim \tau_0$ before saturation is reached on the resolution scale $1/Q_0$ of the initial target dipole. Eq. (4.36) is valid, strictly speaking, for very large rapidities $\tau \gg \tau_0$. We shall later compute the first sub–asymptotic correction to this result [52]. But before doing that, it is interesting to deduce a simplified expression for the scattering amplitude valid in the vicinity of the saturation line.

4.2.4 The scattering amplitude near the saturation line

To that aim, we shall expand $F(\gamma|\rho, t)$ to second order in $\gamma - \gamma_s$, by using Eq. (4.34) :

$$F(\gamma|\rho, t) \simeq -(\rho - \rho_s)(\gamma - \gamma_s) + \frac{t}{2}\chi''(\gamma_s)(\gamma - \gamma_s)^2\,, \tag{4.37}$$

and then perform the Gaussian integration over γ, to deduce

$$T(\rho, \tau) \simeq \frac{e^{-\gamma_s(\rho - \rho_s)}}{\sqrt{2\pi\beta\bar{\alpha}\tau}}\,\exp\left\{-\frac{(\rho - \rho_s)^2}{2\beta\bar{\alpha}\tau}\right\}\,, \tag{4.38}$$

where $\beta \equiv \chi''(\gamma_s) = 48.518...$ (β measures the curvature of the function $\chi(\gamma)$ at the saturation saddle point γ_s; see Fig. 36). This formula can be more suggestively rewritten in terms of the original variables r and τ, as

$$T_\tau(r) \simeq \frac{(r^2 Q_s^2(\tau))^{\gamma_s}}{\sqrt{2\pi\beta\bar{\alpha}\tau}}\,\exp\left\{-\frac{(\ln r^2 Q_s^2(\tau))^2}{2\beta\bar{\alpha}\tau}\right\}\,. \tag{4.39}$$

Note that this expression is not truly of $\mathcal{O}(1)$ when $r \sim 1/Q_s(\tau)$, because in our previous calculation of the saturation momentum we did not take into account the (relatively slowly varying) prefactor coming from the integration over the Gaussian fluctuations. A better treatment of this point will be presented in the next subsection.

Since obtained via an expansion around the saturation saddle point, the approximation (4.39) is valid so long as ρ is close to (but larger than) ρ_s. In fact, Eq. (4.38) can be recognized as the expansion of $\ln[T(\rho, t)]$ to second order in $(\rho - \rho_s)/t$, which is a priori justified so long as $(\rho - \rho_s)/t \ll 1$. However, since $\beta \sim 50$ is a rather large number, the second order term in this expansion remains small as compared to the first one even for $\rho - \rho_s \sim t$. By also using the fact that $\rho_s(t) \sim t$, we conclude that Eq. (4.38) (or (4.39)) is a good approximation for any ρ within the following range [51]

$$0 < \rho - \rho_s(\tau) \lesssim C\rho_s(\tau)\,, \tag{4.40}$$

with $C \sim \mathcal{O}(1)$. Within this range, the functional form of the scattering amplitude is given by an universal function of $z \equiv r^2 Q_s^2(\tau)$ and τ, which is determined solely by the quantum (BFKL) evolution together with the saturation boundary condition. The initial condition at $\tau = 0$ fixes only the relevant normalization factors, namely $\mathcal{T}_0$ in Eq. (4.38) and the overall scale Q_0 for the saturation momentum in Eq. (4.36).

It is instructive to compare in this respect Eq. (4.38) with the more familiar expression of the amplitude corresponding to a very small dipole, such that $\rho \gg \alpha\tau$. (This limit is known as the "double logarithmic approximation", or DLA.) For $\rho \gg t$, the saddle point in Eq. (4.29) is close to one (where the function $\chi(\gamma)$ becomes very large, cf. Fig. 36), and it reads $\gamma \approx 1 - \sqrt{t/\rho}$. Then the saddle point approximation yields (in mixed notations involving both r and ρ)

$$T(\rho, \tau) \simeq \alpha^2 r^2 Q_0^2 \, \frac{e^{\sqrt{4\bar{\alpha}\tau\rho}}}{\sqrt{2\pi\sqrt{4\bar{\alpha}\tau\rho}}} \,, \tag{4.41}$$

where the dominant behaviour upon r^2 (i.e., the overall factor $r^2 Q_0^2$) is inherited from the initial condition. This behaviour exhibits 'color transparency' : the scattering amplitude for a very small dipole is proportional to r^2, up to logarithmic corrections due to the DGLAP evolution [15], i.e., the evolution with $\rho \sim \ln Q^2$, which here has been considered only at DLA level. However, this power–like dependence upon r^2 is strongly affected by the high–energy, BFKL, evolution, which reduces the dominant power of r^2 from 1 to $\gamma(\rho)$ — the saddle point appropriate to the interesting value of ρ. In the (formal) high energy limit $t \gg \rho$, the *BFKL anomalous dimension* $1 - \gamma$ reaches its maximal value, equal to $1 - \gamma_0 = 1/2$. But as mentioned before, this 'Pomeron' exponent is not the one which governs the actual high–energy behaviour (in the sense of the approach towards unitarity): In the vicinity of the saturation line, the power–law behaviour is rather controlled by the anomalous dimension at saturation, $1 - \gamma_s \approx 0.37$.

4.2.5   The approach towards asymptotics

The law (4.36) for the energy dependence of the saturation momentum is valid only for sufficiently large values of τ, such that $\tau \gg \tau_0 \gtrsim 1/\bar{\alpha}$. Because of that, we shall refer to the exponent $c_s\bar{\alpha}$ appearing in that equation as the *asymptotic saturation exponent*. The first subasymptotic correction (as predicted by the BK equation with fixed coupling) is also known [52, 170], and will be succinctly derived here following an intuitive method due to Mueller and Triantafyllopoulos [52].

For this purpose, we need also the *subleading* behaviour of the dipole amplitude at high energy, that is, the τ–dependent 'prefactor' coming from the integration over small fluctuations around the saturation saddle point, that we shall denote as $\Delta^{-1}(\tau)$. Assuming this prefactor to be correctly given by Eq. (4.38) — as we shall shortly see, this is not the case ! —, an improved estimate for the saturation momentum can be easily deduced be requiring this expression to be of $\mathcal{O}(1)$ when $\rho = \rho_s'$ (our improved estimate for the saturation line). Then, to first order in $\rho_s' - \rho_s$, one easily finds

$$\rho_s' - \rho_s \simeq -\frac{\ln\Delta(\tau)}{\gamma_s} \,, \qquad \ln\Delta(\tau) = \frac{1}{2}\ln(2\pi\beta\bar{\alpha}\tau) \,, \tag{4.42}$$

and therefore

124

$$\lambda_s(\tau) \equiv \frac{\mathrm{d}\rho_s(\tau)}{\mathrm{d}\tau} \simeq c_s\bar{\alpha} - \frac{1}{\gamma_s}\frac{\mathrm{d}\ln\Delta}{\mathrm{d}\tau} = \bar{\alpha}\left(c_s - \frac{1}{2\gamma_s\bar{\alpha}\tau}\right). \tag{4.43}$$

However, this estimate is not full right: the corrective term shown above has the right sign and also the right τ–dependence, but is misses the correct normalization by a factor of 3. The reason for this failure is the fact that the effects of the fluctuations around the saddle point are over–estimated in Eq. (4.38) as compared to the correct solution to the non–linear (BK) equation. In turn, this is related to the diffusive nature of the BFKL equation, as manifest on the saddle point solution in Eq. (4.38). With $\xi \equiv \rho - \rho_s$, the latter can be rewritten as

$$T(\xi, \tau) = \mathrm{e}^{-\gamma_s\xi}\,\psi(\xi, \tau), \tag{4.44}$$

where $\psi(\xi, \tau)$ is a particular solution to the diffusion equation (with $D \equiv \bar{\alpha}\beta$):

$$\frac{\partial}{\partial\tau}\,\psi(\xi, \tau) = \frac{D}{2}\frac{\partial^2}{\partial\xi^2}\,\psi(\xi, \tau). \tag{4.45}$$

Namely, this solution is such that it explores (in the sense of receiving contributions from) a region in ξ around $\xi = 0$ within a distance given by the 'diffusion radius' $\xi_{\mathrm{diff}} = \sqrt{2D\tau}$. Accordingly, even if one tries to follow a trajectory of weak amplitude $T(\xi, \tau) = \kappa$ with $\kappa \ll 1$ (meaning $\xi \gg 1$), the respective solution will receive contributions from points ξ' which, with increasing τ, lie further and further away from this trajectory (within a distance $|\xi' - \xi| \lesssim \sqrt{2D\tau}$). For sufficiently large values of τ, such contributing points will enter the saturation region ($\xi < 0$), where the BFKL amplitude becomes unphysical, since larger than one. In fact, for negative ξ, the BFKL solution (4.44) blows up exponentially, thus giving a very large weight to such unphysical evolutions.

This diffusion is a real property of the high–energy evolution near saturation, but its range should be limited by saturation. Namely, if one uses Eq. (4.44) to *define* $\psi(\xi, \tau)$ in the whole physical phase–space, including the non–linear region at saturation, then it is clear that the physical function $\psi(\xi, \tau)$ should exponentially die away when $\xi < 0$ (since $T = 1$ there). Motivated by this, Mueller and Triantafyllopoulos [52] have observed that a better approximation in the linear regime ($\xi \gg 1$) can be obtained by solving the diffusion equation (4.45) with an absorptive boundary condition at saturation: $\psi(\xi, \tau) = 0$ for $\xi = -\xi_0$, where ξ_0 is a number of order one. The corresponding solution is easily obtained through the "image source" method, and reads

$$\psi(\xi, \tau) = \frac{1}{\sqrt{2\pi D\tau}}\left\{\mathrm{e}^{-\frac{\xi^2}{2D\tau}} - \mathrm{e}^{-\frac{(\xi+2\xi_0)^2}{2D\tau}}\right\} \simeq \frac{2\xi_0}{\sqrt{2\pi}(D\tau)^{3/2}}\,\xi\,\mathrm{e}^{-\frac{\xi^2}{2D\tau}}, \tag{4.46}$$

where the final, approximate, expansion is appropriate for ρ within the scaling window (4.40), since there $\xi \gg \xi_0$ and at the same time $\xi \ll D\tau$. The unwanted contributions coming from diffusion inside the saturation region cancel in the difference between the two exponentials in Eq. (4.46).

With $\psi(\xi, \tau)$ in Eq. (4.46) replacing the diffusion piece of the original amplitude (4.38), it is clear that the prefactor $\Delta(\tau)$ behaves like $\ln\Delta(\tau) \sim (D\tau)^{3/2}$. Therefore, the correct

version of Eq. (4.43) reads

$$\lambda_s(\tau) \simeq c_s \bar{\alpha} - \frac{3}{2\gamma_s} \frac{1}{\tau} .\qquad (4.47)$$

This is indeed the right result, as subsequently confirmed through a more rigorous analysis by Munier and Peschanski [170] (see the next section). Remarkably, this last method has also allowed one to obtain the third term in the asymptotic expansion of $\lambda(\tau)$ at large τ, the last one to be universal [171].

4.2.6 Generalizations: running coupling, NLO, large nucleus

We shall now make a brief excursion beyond the leading–order formalism and discuss the consequences of the running of the coupling (and, more generally, of the full NLO corrections to the BFKL equation [18]) on the calculation of the saturation exponent λ_s [51–53]. As we shall see, these consequences are both qualitatively and quantitatively important, and in particular they affect the interplay between the energy dependence and the atomic number dependence of the saturation momentum of a large nucleus [172].

Whereas the NLO version of the BFKL kernel is known for almost a decade [18], a full NLO formalism including saturation and unitarity corrections (say, the corresponding generalization of the Balitsky–JIMWLK hierarchy) is still missing, and its construction is currently an active topics of research. In particular, the proper way to modify the BK equation in such a way to include the effects of the running coupling (which represents a part of the NLO corrections) has been recently clarified [173–177]. However, the respective calculation of the saturation exponent has preceded these recent developments by several years [51–53]. This was possible because, as previously noticed, the saturation momentum is effectively determined by the BFKL evolution with an appropriate boundary condition, and this remains true at NLO level.

In what follows we shall use the one–loop version of the QCD running coupling, which reads (for N_c colors and N_f flavors of quarks)

$$\bar{\alpha}(Q^2) \equiv \frac{g^2(Q^2)N_c}{4\pi^2} = \frac{b}{\ln(Q^2/\Lambda_{\text{QCD}}^2)}, \qquad \text{with} \qquad b \equiv \frac{12 N_c}{11 N_c - 2 N_f}, \qquad (4.48)$$

where Q^2 is the relevant virtuality scale; in the present context, this is either the gluon transverse momentum $k_\perp^2$, or the inverse dipole size, $Q^2 \sim 4/r^2$. Also, it will be convenient to choose the scale Λ_{QCD} which enters the running coupling as the reference scale for the various logarithmic variables; thus we shall write

$$\rho \equiv \ln \frac{1}{r^2 \Lambda_{\text{QCD}}^2}, \qquad \rho_0 \equiv \ln \frac{Q_0^2}{\Lambda_{\text{QCD}}^2}, \qquad \rho_s(\tau) \equiv \ln \frac{Q_s^2(\tau)}{\Lambda_{\text{QCD}}^2}, \qquad (4.49)$$

and hence $\bar{\alpha}(\rho) = b/\rho$, etc.

The introduction of a running coupling into the leading–order (LO) BFKL, or BK, equations is *a priori* ambiguous. As above mentioned, the correct respective prescription is

126

known by now [174, 175, 177], but this is quite complicated and thus renders the analytic calculations prohibitive. It is therefore useful to keep in mind that, so long as one is interested in the saturation line alone, this prescription can be considerably simplified. Indeed, at least for sufficiently high energy, the saturation line is determined by the evolution within a limited distance $\rho - \rho_s(\tau)$ above ρ_s, namely, within the window for geometric scaling. (The saturation region at $\rho < \rho_s(\tau)$ is essentially inert under the evolution, since there T has already reached its maximal value $T = 1$.) The width $\xi_{\mathrm{diff}}(\tau)$ of this window is increasing with τ, but this increase is quite slow and, moreover, it is further slowed down by running coupling effects, as we shall see. Accordingly, the relative width $(\rho - \rho_s(\tau))/\rho_s(\tau)$ of the active region decreases with τ, so for large enough τ it is a good approximation to expand the running coupling around its value at ρ_s :

$$\bar{\alpha}(\rho) \, = \, \frac{b}{\rho} \, \approx \, \bar{\alpha}(\rho_s(\tau)) \left(1 - \frac{\rho - \rho_s(\tau)}{\rho_s(\tau)} \right) . \tag{4.50}$$

The first term $\bar{\alpha}(\rho_s(\tau))$ in this expansion is in fact enough for computing the asymptotic behaviour of $\rho_s(\tau)$ at $\tau \to \infty$ [22, 51], as verified through more sophisticated calculations (both analytic [52, 53, 178] and numerical [176, 179–182]). Namely, assuming that the rate of growth is determined *locally* by the corresponding fixed coupling result, cf. Eqs. (4.35) and (4.36), one easily finds (for large enough τ)

$$\frac{\mathrm{d}\rho_s}{\mathrm{d}\tau} \, \simeq \, c_s \bar{\alpha}(\rho_s) \, = \, \frac{c_s b}{\rho_s} \, \Longrightarrow \, \rho_s(\tau) \simeq \sqrt{2 c_s b \tau} \, . \tag{4.51}$$

This result exhibits an important, general, consequence of the running of the coupling, which is to considerably slow down the high-energy evolution. In particular, with running coupling, the effective 'saturation exponent' $\lambda_s \equiv \mathrm{d}\rho_s/\mathrm{d}\tau$ decreases like $1/\sqrt{\tau}$ (at least, for sufficiently large τ), rather than being a constant (as it was in the fixed coupling case).

Although valid only for very large τ, Eq. (4.51) can be used to construct an interpolation between low and high energies, which is sometimes useful. Namely, one can write

$$Q_s^2(\tau) \, \simeq \, \Lambda_{\mathrm{QCD}}^2 \, \mathrm{e}^{\sigma(\tau)} \, , \qquad \sigma(\tau) \, = \, \sqrt{2 c_s b \tau + \rho_s^2(0)} \, . \tag{4.52}$$

This has the expected behaviour both for large τ (i.e., for $2 c_s b \tau \gg \rho_s^2(0)$), where it reduces to Eq. (4.51), and for relatively small τ (with $2 c_s b \tau \ll \rho_s^2(0)$), where it takes the same form as at fixed coupling, cf. Eq. (4.36), but with $\bar{\alpha}$ now evaluated at the initial saturation scale: $\bar{\alpha} \to b/\rho_s(0)$.

To have a better control on the saturation exponent at intermediate values of τ, one however needs to include in the calculation also the second term in the expansion in Eq. (4.50) together with the other NLO corrections (one–loop corrections of relative order α) to the BFKL equation [18]. These corrections become available in the recent years and, after some initial confusion caused by the fact that they appeared to be unreasonably large, new resummation techniques have been developed (in relation with collinear radiation) which allow one to keep these corrections under control. By using one of these techniques, known as the 'renormalization group (RG) improvement' [19], Triantafyllopoulos [53] has
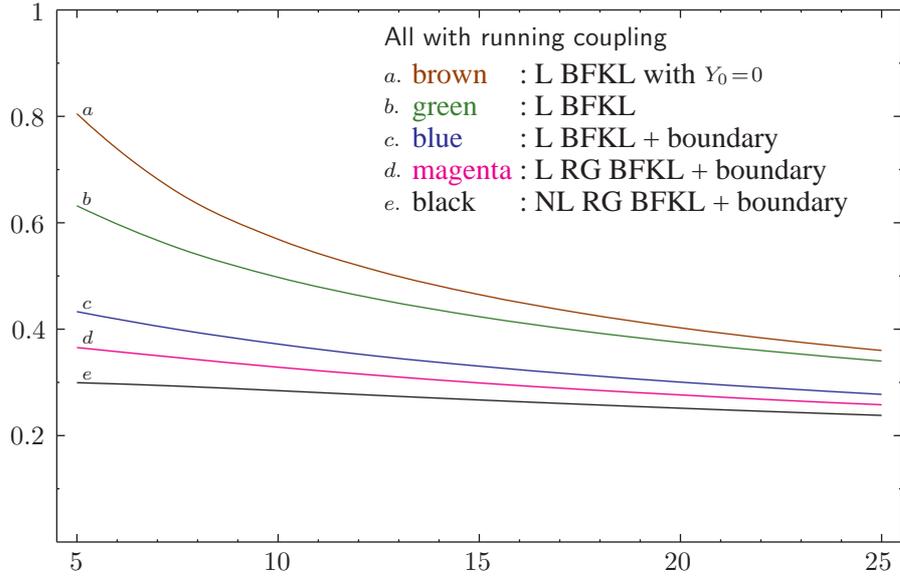
Fig. 37. The saturation exponent $\lambda_s(\tau) = \mathrm{d}\ln Q_s^2/\mathrm{d}\tau$ with running coupling and for successive approximations to the NLO corrections is represented as a function of the rapidity τ (here denoted as Y). Plot taken from Ref. [53].

given a calculation of the saturation exponent at NLO accuracy. His results which are in a qualitative agreement with the phenomenology at HERA [54, 73, 83] and, moreover, they appear to smoothly approach the corresponding results at leading order and *running* coupling when increasing the energy. This last feature was expected since, with increasing τ, the relevant coupling $\bar{\alpha}(\rho_s(\tau))$ is monotonously decreasing, so the influence of the other NLO corrections (beyond the running of the coupling) should become less and less important.

The calculation in Ref. [53] is too complex to be described here, but its results are illustrated in Fig. 37, where the function $\lambda_s(\tau)$ — as numerically obtained from the full NLO calculation — is compared with the simpler result in Eq. (4.52) and also with some intermediate results corresponding to various degrees of sophistication in the inclusion of NLO effects and/or of the effects of the saturation boundary condition. The simple estimate (4.52) is referred in this figure as 'L BFKL' and it is rewritten as $\rho_s(Y) = \sqrt{2c_s b(Y + Y_0)}$, with $Y \equiv \tau$ being a different notation for the rapidity variable. As one can see from this plot, whereas the 'L BFKL' estimate shows a rather pronounced decrease with τ, as $1/\sqrt{\tau}$, this decrease is slowed down, and at the same time the absolute value of λ_s diminishes, when other NLO corrections (including the RG improvement of Ref. [19]) are successively taken into account. The complete result is a rather flat curve — so in that respect it resembles the corresponding result for fixed coupling, cf. Eq. (4.36) — but the corresponding value is sensibly smaller than at fixed coupling: namely, one finds $\lambda_s(\tau) = 0.30 \div 0.29$ when τ is increased from 5 to 9 (which corresponds to the phenomenologically range of interest at HERA: $10^{-2} \gtrsim x \gtrsim 10^{-4}$). As anticipated, this value of λ_s is roughly consistent with that extracted from the 'saturation' fits to the HERA data [54, 73, 83] (see Sect. 5.1).

Consider finally the case where the target is a large nucleus with atomic number $A \gg 1$.

What is then the A–dependence of the saturation scale and how does this change when increasing the energy ? Clearly, the A–dependence enters via the initial conditions at low energy. For a sufficiently large nucleus, saturation is expected already for relatively large values of x, or small values of τ, namely for $\tau_0 \sim \ln A^{1/3}$ (which yields $\tau_0 \sim 2$ for a Gold nucleus). Hence, if one starts the high–energy evolution at this rapidity τ_0, then the relevant 'infrared cutoff' (the scale Q_0 introduced in Eq. (4.29)) is the *nuclear saturation momentum* $Q_s(A)$, which is expected to scale like $Q_s^2(A) \sim A^{1/3}$ (cf. Sect. 2.5). For instance, this scaling is explicit in the MV model, cf. Eq. (2.54).

In the *fixed coupling case*, this initial scale $Q_0 \equiv Q_s(A)$ is the only intrinsic scale in the problem, so the saturation momentum at any $\tau > \tau_0$ must be proportional to it. Accordingly, $Q_s^2(A, \tau)$ scales like $A^{1/3}$ for any τ. In fact, the previous results allow us to deduce (for $\tau \gg \tau_0$) :

$$
Q_s^2(A, \tau) \simeq Q_s^2(A)\, \frac{\mathrm{e}^{c_s \bar{\alpha}\tau}}{(\bar{\alpha}\tau)^{3/2\gamma_s}}\,, \qquad c = 4.883...,
\tag{4.53}
$$

where we have also included the first sub–asymptotic correction, cf. Eq. (4.47).

For a *running coupling*, on the other hand, Eq. (4.52) involves two momentum scales: Λ_{QCD} and $Q_s(A)$ (the latter enters through $\rho_s(0) = \ln(Q_s^2(A)/\Lambda_{\mathrm{QCD}}^2) \equiv \rho_A$), and this allows for an interesting change in the A–dependence with increasing τ [172] : For small τ, such that $2cb\tau \ll \rho_A^2$, we have $Q_s^2(\tau) \propto Q_s^2(A)$, and thus the usual $A^{1/3}$–scaling holds. But with increasing τ, the dependence on A becomes weaker and weaker [16] :

$$
Q_s^2(A, \tau) \simeq \Lambda_{\mathrm{QCD}}^2\, \mathrm{e}^{\sqrt{2c_s b\tau}}\, \exp\left\{ \frac{\rho_A^2}{2\sqrt{2c_s b\tau}} \right\} \qquad \text{for} \qquad 2c_s b\tau \gg \rho_A^2 .
\tag{4.54}
$$

By using similar formulae (with $\rho_A \to \rho_p \sim 1$) also for the proton, it becomes clear that, unlike in the fixed coupling case, where the ratio $Q_s^2(A, \tau)/Q_s^2(p, \tau) = Q_s^2(A)/Q_s^2(p)$ is independent of τ and of $\mathcal{O}(A^{1/3})$, with a running coupling this ratio decreases with τ and approaches to one at sufficiently large τ [172] :

$$
\frac{Q_s^2(A, \tau)}{Q_s^2(p, \tau)} \simeq \exp\left\{ \frac{\rho_A^2 - \rho_p^2}{2\sqrt{2c_s b\tau}} \right\} \simeq 1 \qquad \text{for} \qquad 2c_s b\tau \gtrsim \rho_A^4 .
\tag{4.55}
$$

Thus, due to running coupling effects in the evolution, the initial difference between the proton and the nucleus is washed out after a rapidity evolution $c_s b\tau \sim \rho_A^4$ [172]. This will play an important role in Sect. 5.2, in the theoretical analysis of the ratio R_{pA} at forward rapidities (cf. Sect. 1.3).

---

[16] Choosing $\rho_A = 4$ (a reasonable value at RHIC) and using $2c_s b \sim 10$, one finds that the running effects become important — in the sense that the two terms under the square root in Eq. (4.52) become comparable — already for $\tau - \tau_0$ of order one [87].

4.3 Geometric scaling

In the previous section we have used BFKL equation supplemented with a saturation boundary condition as an approximation to BK equation, and on its basis we computed the saturation exponent $\lambda_s(\tau) = \mathrm{d}\ln Q_s^2(\tau)/\mathrm{d}\tau$ and the dipole scattering amplitude in the vicinity of the saturation line (see Eqs. (4.38)–(4.40)). In what follows, we shall emphasize a crucial property of the amplitude in Eq. (4.38), known as *geometric scaling* [51–53], which is a manifestation of the physics of saturation in the weak scattering regime at $Q^2 \gg Q_s^2(\tau)$. Furthermore, we shall describe a remarkable physics and mathematical correspondence — that between high–energy evolution in QCD and a classical problem in statistical physics [47, 170] — which in particular helps explaining why we can trust the results of the linear BFKL equation even close to the saturation line.

For simplicity, throughout this section we shall consider the case of a fixed coupling, but similar results hold also for a running coupling [52, 53, 178].

4.3.1 Geometric scaling above the saturation line

In the subsequent discussion, we shall rely on the more refined calculation in Sect. 4.2.5, which permits a better control of the approach towards saturation and asymptotics. The resulting saturation exponent is displayed in Eq. (4.47), while the corresponding expression for the dipole amplitude can be inferred from Eqs. (4.44)–(4.46) as

$$T(\rho, \tau) \simeq \kappa_0 \, (\rho - \rho_s) \, \mathrm{e}^{-\gamma_s(\rho - \rho_s)} \, \exp\left\{-\frac{(\rho - \rho_s)^2}{2\beta\bar{\alpha}\tau}\right\}, \tag{4.56}$$

which is valid within the window in Eq. (4.40). In the above equation, κ_0 is an undetermined normalization factor, and the original prefactor $\propto \tau^{-3/2}$ in Eq. (4.46) has been absorbed into the definition of the saturation momentum, which therefore reads

$$\rho_s(\tau) \equiv \ln \frac{Q_s^2(\tau)}{Q_0^2} = c_s \bar{\alpha}\tau - \frac{3}{2\gamma_s} \ln \tau + \text{const.}, \tag{4.57}$$

where the last, constant, term (which fixes the absolute normalization of Q_s) depends upon the initial conditions at low energy and generally involves powers of α (see, e.g., Eq. (4.36)). Eq. (4.56) appears to vanish when $\rho = \rho_s$, but this property merely reflects the fact that the physical amplitude must approach a value of order one (rather than grow up exponentially) when $\rho - \rho_s$ becomes negative; in reality, Eq. (4.56) is valid only for $\rho - \rho_s \gg 1$.

In particular, let us consider the situation where the difference $\rho - \rho_s(\tau)$ is much smaller than the *diffusion radius* $\xi_{\text{diff}} \sim \sqrt{2\beta\bar{\alpha}\tau}$. Then, the Gaussian in Eq. (4.56) can be replaced by one, and the amplitude becomes purely a function of $\rho - \rho_s(\tau)$:

$$T(\rho, \tau) \simeq \kappa_0 \, (\rho - \rho_s(\tau)) \, \mathrm{e}^{-\gamma_s(\rho - \rho_s(\tau))} \qquad \text{for} \qquad 1 \ll \rho - \rho_s \ll \sqrt{2\beta\bar{\alpha}\tau}, \tag{4.58}$$

that is, a *scaling function* of $z \equiv r^2 Q_s^2(\tau)$:

$$T_\tau(r) \simeq \kappa_0 \left[\ln \frac{1}{r^2 Q_s^2(\tau)} \right] \left(r^2 Q_s^2(\tau) \right)^{\gamma_s} . \tag{4.59}$$

The kinematical window where this is valid, that is,

$$1 \ll \rho - \rho_s(\tau) \ll \sqrt{2\beta\bar{\alpha}\tau} \qquad \text{(geometric scaling for } T(\rho, \tau)), \tag{4.60}$$

is generally referred to as the *window for extended geometric scaling*. The word 'extended' refers to the fact that this scaling is now found to hold *above* the saturation line, i.e., in the relatively dilute region at $\rho > \rho_s$ where the evolution is *a priori* linear, yet it is influenced — as we know by now — by the boundary condition of saturation. In Sect. 4.1 we have seen that a similar scaling holds in the saturation region at $\rho < \rho_s$ (i.e., for $k_\perp \lesssim Q_s(\tau)$, or $r \gtrsim 1/Q_s(\tau)$), cf. Eqs. (4.8) and (4.19); but in that context, such a scaling looks less surprising, because the only evolution which remains possible at saturation is that of the saturation momentum itself. We now see that this scaling gets transmitted across the saturation boundary, up to relatively large values of ρ.

If, moreover, we are interested in the scaling behaviour of the *logarithm* of the amplitude, so like in the phenomenological studies at HERA [83] (see Fig. 10), then from Eq. (4.56) we deduce

$$\ln T(\rho, \tau) \approx -\gamma_s(\rho - \rho_s) - \frac{(\rho - \rho_s)^2}{2\beta\bar{\alpha}\tau}, \tag{4.61}$$

which shows that for $\ln T$ the scaling holds in an even wider region in $\rho - \rho_s$ — namely, so long as the second, diffusive, term in the r.h.s. of Eq. (4.61) can be neglected next to the first term there. This requires $\rho - \rho_s \ll 2\beta\gamma_s\bar{\alpha}\tau$, an inequality which given the relatively large value of $\beta \approx 48$ is well satisfied for any ρ within the larger window (4.40) [51]. In terms of $Q^2 \equiv 1/r^2$ (which roughly corresponds to the virtuality of the exchanged photon in DIS), the window (4.40) amounts to

$$Q_s^2(\tau) \ll Q^2 \ll \frac{Q_s^4(\tau)}{Q_0^2} \qquad \text{(geometric scaling for } \ln T(\rho, \tau)). \tag{4.62}$$

For the typical 'small–x' kinematics at HERA we have $Q_s \sim 1$ GeV, which together with $Q_0 \sim \Lambda_{\text{QCD}} \approx 200$ MeV implies that the upper boundary of the window (4.62) is in the ballpark of 10 to 100GeV^2, in qualitative agreement with the observed geometric scaling at HERA [83] (see Fig. 10). The various regions for QCD evolution identified so far are illustrated in Fig. 38.

Let us finally comment on geometric scaling with running coupling: in that case, analytic approximations are more difficult to construct and they are moreover restricted to asymptotically high energies. Yet, the respective studies in Refs. [52, 53, 178] found that geometric scaling persists in that case too, although in a much narrower window than at fixed coupling. Specifically, in Refs. [53, 87], this window was estimated as

$$1 \ll \rho - \rho_s(\tau) \lesssim [2c_s b\tau + \rho_s^2(0)]^{1/6} , \tag{4.63}$$
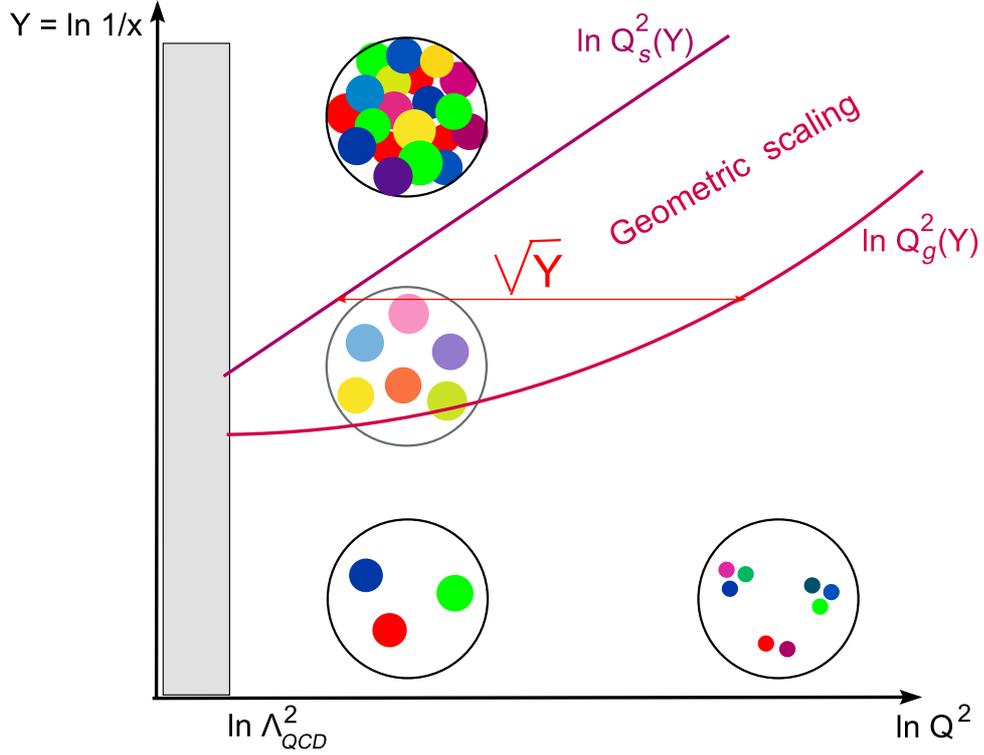
Fig. 38. A map of the quantum evolution in the $\tau - \ln Q^2$ plane.

where now $\rho_s(\tau) \simeq \sqrt{2c_s b\tau}$, cf. Eq. (4.51). But precisely because this window is so narrow, the analytic form of the approximate solution becomes ambiguous (different functions being able to describe the same behaviour over such a limited interval in ρ), and a more recent analysis [183] has actually found a different scaling behaviour: instead of being a function of $\rho - \sqrt{2c_s b\tau}$, so like in the running–coupling version of Eq. (4.58), the approximate solution found in Ref. [183] rather scales as a function of the variable $\rho - 2c_s b\tau/\rho$. So far, it seems difficult to distinguish between these different scalings via numerical analyses of the BK equation. What is however certain is that the running of the coupling considerably slows down the high–energy evolution, thus reducing both the value of the saturation momentum and the phase–space for geometric scaling.

4.3.2 The FKPP equation and the traveling wave

So far, all our explicit results concerning saturation and geometric scaling have been obtained from a study of the BFKL equation (which is the linearized version of the BK equation) with appropriate boundary conditions at saturation. One may legitimately wonder whether these results are indeed consistent with the actual, non–linear, dynamics described by BK equation. In this section, we shall give a general argument in that sense [170].

For the present purposes, the BK equation (4.15) is more conveniently rewritten in terms of the scattering amplitude (with impact parameter dependence neglected) :

$$\frac{\partial}{\partial \tau} T_\tau(\boldsymbol{x} - \boldsymbol{y}) = \frac{\bar{\alpha}_s}{2\pi} \int_{\boldsymbol{z}} \frac{(\boldsymbol{x} - \boldsymbol{y})^2}{(\boldsymbol{x} - \boldsymbol{z})^2 (\boldsymbol{z} - \boldsymbol{y})^2} \Big\{ - T_\tau(\boldsymbol{x} - \boldsymbol{y}) \tag{4.64}$$

$$+ T_\tau(\boldsymbol{x} - \boldsymbol{z}) + T_\tau(\boldsymbol{z} - \boldsymbol{y}) - T_\tau(\boldsymbol{x} - \boldsymbol{z}) T_\tau(\boldsymbol{z} - \boldsymbol{y}) \Big\}.$$

So far, there are no exact analytic solutions known for this equation, nor systematic approximations valid in the transition region from the linear to the non–linear regime (see however below). But BK equation has been extensively studied via numerical calculations, sometimes performed directly in coordinate space [180], but more often in momentum space (since the non–local structure of the equation gets simplified after a Fourier transform $T_\tau(\boldsymbol{r}) \rightarrow T_\tau(\boldsymbol{k})$) [115, 162, 168, 176, 179, 181, 184, 185]. All these studies confirmed the existence of geometric scaling in a wide kinematical region, and some of them have also checked some fine details of the approximate analytic solutions, like the values of the anomalous dimension and of the saturation exponent. In particular, the studies involving a running coupling [176, 179–181] have demonstrated the change in the τ–dependence of the saturation momentum, cf. Eq. (4.52), and also the fact that the evolution is drastically slowed down as compared with the fixed–coupling case.

An important step towards understanding geometric scaling directly in the context of the non–linear BK equation was made in Ref. [170], where a correspondence has been identified between BK equation and another non–linear equation — the *FKPP equation* (from Fisher Kolmogorov, Petrovsky, and Piscounov [186]; see below) —, which is well known in the context of statistical physics and was intensively studied over several decades (see e.g. [165] for a review paper and references). This correspondence allowed one to translate some rigorous results known about the FKPP equation to the corresponding problem in QCD, and thus gain new insight about properties like geometric scaling or the energy dependence (4.47) of the saturation exponent.

Proposed back in the thirties [186], the FKPP equation is most succinctly written as

$$\partial_t u(x,t) = \partial_x^2 u(x,t) + u(x,t) - u^2(x,t), \tag{4.65}$$

and appears as a mean field approximation (typically, in the limit of a large number of particles) in a variety of stochastic problems in chemistry, physics, and biology [165], the prototype of which is the *reaction–diffusion problem* $A \rightleftharpoons A + A$. This problem can be formulated as follows:

Consider a one–dimensional lattice in which the site x_i, with $i = 1, 2, \ldots$, is occupied by n_i particles ('molecules') of type A. They can be involved in the following three processes:

• A particle can locally split into two, at a rate α, i.e.

$$A_i \xrightarrow{\alpha} A_i + A_i. \tag{4.66}$$

• Two particles can recombine locally into one at a rate 2β, i.e.

$$A_i + A_i \xrightarrow{2\beta} A_i. \tag{4.67}$$

• A particle can diffuse to a neighboring site at a rate μ, i.e.

$$A_i \xrightarrow{\mu} A_{i+1}, \qquad A_i \xrightarrow{\mu} A_{i-1}. \tag{4.68}$$

As we shall shortly see, the maximal occupation number is $N = \alpha/\beta$, and corresponds to a stable state (since for $n_i = N$, the splitting and recombination processes compensate each other). Introducing the *concentration* $u(x_i, t) \equiv n_i(t)/N$ at site i and time t, then in limit $N \to \infty$, the function $u(x, t)$ obeys Eq. (4.65) (in continuum notations and after an appropriate rescaling of the diffusion coefficient). This equation admits two fixed points: an unstable one at $u = 0$ and a stable one at $u = 1$ (the latter corresponds to the maximal occupation number alluded to above). It has been rigorously established [187] that for initial conditions $u(x, 0)$ such that $u(x, 0) \to 1$ as $x \to -\infty$ and $u(x, 0) \sim e^{-\gamma x}$ as $x \to \infty$, there exists a one–parameter family F_v of solutions (indexed by the velocity v) which behave as *traveling waves* at large times:

$$u(x, t) \simeq F_v(x - vt), \tag{4.69}$$

with $F_v(z) \to 1$ as $x \to -\infty$ and $F_v(z)$ decreasing exponentially at large z. The approximate equality in Eq. (4.69) should be understood as the large t limit for fixed $z \equiv x - vt$. The analytic expression of the shape $F_v(z)$ is generally not known, but one can determine the range of velocities v for which traveling wave solutions exist. Namely, by assuming $F_v(z) \simeq e^{-\gamma z}$ at large z, where $F_v \ll 1$, and by inserting this *Ansatz* in the linearized version of Eq. (4.65), one immediately finds:

$$v(\gamma) = \gamma + 1/\gamma, \tag{4.70}$$

which shows that $v \geq 2$, with the minimal velocity $v_0 = 2$ being reached for a slope $\gamma_0 = 1$.

What is remarkable about Eq. (4.65) is that the asymptotic velocity of a traveling wave is given by Eq. (4.70) only so long as the parameter γ in the initial condition is sufficiently small: $\gamma \leq \gamma_0 = 1$. But for a steeper initial condition $u(x, 0)$, which decays faster than $e^{-\gamma_0 x}$, the solution selected at large time by the non–linear dynamics is the one corresponding to the minimal velocity v_0. In that case, the *front* of the traveling wave decays like $e^{-\gamma_0 z}$, while the more abrupt decrease of the initial condition persists only at very large x (for fixed t), far ahead of the front. This *selection mechanism* is related to an important property of the FKPP equation (and, by extension, of the BK equation; see below), namely the fact that the propagation of the front is driven by the linear dynamics in the tail of the front at large z, where u (or T) is small: The front is *pulled* by the growth and spread of small perturbations around the unstable state (*"pulled front"*).

Because of this property, the characteristics of the front, like its velocity (including at intermediate times) or the shape $F_{v_0}(z)$ for $z \gg 1$, can be obtained from an analysis of the linearized equation valid at large z. For instance, one readily finds (A is a normalization constant) :

$$F_{v_0}(z) \simeq A z \, e^{-\gamma_0 z} \qquad \text{for} \qquad z \gg 1, \tag{4.71}$$

where the factor of z occurs because the equation $v(\gamma) = v_0$ has two degenerate solutions (cf. Eq. (4.70)) : $\gamma_1 = \gamma_2 \equiv \gamma_0$ (with $\gamma_0 = 1$, of course). Furthermore, the position of the front $X(t)$ can be defined as a level curve for $u(x, t)$: $u(X(t), t) = \kappa$, with $\kappa \ll 1$. We already know that, for asymptotically large times, $X(t) \simeq v_0 t$, but the first two sub–asymptotic corrections have also been computed (we only show here the first such a

correction) [165, 187, 188]:

$$X(t) \simeq v_0 t - \frac{3}{2\gamma_0} \ln t .$$

(4.72)

What is the relation between the FKPP equation and the BK equation of interest for us in QCD at high energy ? There is an obvious parallel between the above equations (4.71) and (4.72) and the previous equations (4.58) and (4.57) in QCD, and this is not an accident [164]: The linear part of the FKPP equation is essentially the same as the BFKL equation in the saddle point approximation[17], whereas the non–linear term in Eq. (4.65) for u corresponds to the non–linear term in the BK equation (4.64) for T (this correspondence becomes particularly transparent when using the momentum–space version of the BK equation; see Sect. 7.2). This strongly suggests that the BK equation should be in the *same universality class* as the FKPP equation, at least in so far as the large–τ behaviour is concerned (since this is controlled by the saddle point approximation to the BFKL equation).

This correspondence can be summarized in the following intuitive picture of the solution $T(\rho, \tau)$ to the BK equation (4.64): At large 'time' τ, the function $T(\rho, \tau)$ is a *front* which interpolates between the *stable* state $T = 1$ behind the front, where the unitarity limit has been reached, and the *unstable* state $T \to 0$ ahead of the front, where the amplitude is still controlled by the initial condition and thus it exhibits 'color transparency' (which is strictly speaking modified by the DLA evolution, cf. Eq. (4.41)). With increasing τ, the *position* $\rho_s(\tau)$ of the front moves towards larger values of ρ, according to Eq. (4.57), and the *velocity* of the front approaches its asymptotic value $\lambda_{as} = c\bar{\alpha}$, cf. Eq. (4.47). The most interesting region of the front is the intermediate region at $\rho_s(\tau) \ll \rho \ll \rho_g(\tau)$, with $\rho_g(\tau) \approx 2\rho_s(\tau)$, in which the solution takes the universal shape (4.56), and which extends with τ. This includes the *extended scaling window* at $\rho - \rho_s \simeq \sqrt{2\beta\bar{\alpha}\tau}$, in which the solution takes the shape of a traveling wave, cf. Eq. (4.58).

Within this interpretation, the *geometric scaling* property of the scattering amplitude, cf. Eq. (4.58), is simply the property (4.69) of the front to propagate as a traveling wave, i.e., to undergo a uniform translation without distortion [164]. For $T(\rho, \tau)$, the scaling holds within the width of the front, which grows like $\sqrt{\tau}$. For $\ln T$, approximate scaling holds within the much larger scaling window $\rho_g - \rho_s \sim \tau$.

In the above discussion, the correspondence between high–energy evolution in QCD and the propagation of pulled fronts in statistical physics has helped us better understanding the mean field approximation in QCD, that is, the BK equation. As we shall see in Sect. 7, this also helps us understanding the *limitations* of the latter in the presence of fluctuations.

---

[17] Indeed, Eqs. (4.44) and (4.45) imply that $T(\rho, \tau)$ obeys the following equation in this approximation: $\partial_Y T = \bar{\alpha} \left(-c\partial_\rho + \frac{\beta}{2}(-\partial_\rho - \gamma_s)^2 \right) T$.

4.4 Total cross–section and the problem of the Froissart bound

In our previous applications of the BK equation to physical problems, we have always ignored the impact parameter dependence, that is, we have considered solutions $T_\tau(\boldsymbol{r})$ which depend upon the size $\boldsymbol{r} = \boldsymbol{x} - \boldsymbol{y}$ of the dipole, but not also upon its impact parameter $\boldsymbol{b} = (\boldsymbol{x} + \boldsymbol{y})/2$. As explained before, this approximation is reasonable so long as one is interested in the evolution towards saturation and the unitarity limit at a fixed impact parameter and, more generally, in a given region in the impact parameter space (say, the central region of a large nucleus, which is quasi–homogeneous). On the other hand, such an approximation is generally not legitimate (not even in the high–energy limit !) for the computation of quantities like the total cross–section

$$\sigma_{\text{tot}}(\tau, r_\perp) = 2 \int \mathrm{d}^2\boldsymbol{b}\, T_\tau(\boldsymbol{r}, \boldsymbol{b}), \qquad (4.73)$$

which involve an integration over all the impact parameters. The inhomogeneity in the target is generally determined by non–perturbative physics, and thus one needs to rely on some model or on a phenomenological parametrization in order to establish the $\boldsymbol{b}$–dependence of the initial condition $T_0(\boldsymbol{r}, \boldsymbol{b})$ at low energy ($\tau = 0$). By itself, this is hardly a novel problem: for a realistic target, the initial conditions at low energy are always non–perturbative and require a model. What is, however, new is that not even the *evolution* of the impact parameter dependence of T_τ with increasing energy is not fully under control in perturbation theory, and this even at arbitrarily high energy. This should be contrasted with the evolution towards saturation at fixed impact parameter, which becomes perturbative as soon as the energy is high enough for the local saturation scale $Q_s(\tau, \boldsymbol{b})$ to be hard: $Q_s(\tau, \boldsymbol{b}) \gg \Lambda_{\text{QCD}}$. Thus, whereas it is possible to study the unitarization of the scattering amplitude for a small dipole (with $r_\perp \ll 1/\Lambda_{\text{QCD}}$) in perturbation theory, on the other it is not possible to perturbatively follow the evolution of the associated cross–section up to very high energy.

The problems originates in the *long–range non–locality* of the perturbative evoluation at high–energy : in dipole language, an original dipole $(\boldsymbol{x}, \boldsymbol{y})$ can dissociate into two dipoles $(\boldsymbol{x}, \boldsymbol{z})$ and $(\boldsymbol{z}, \boldsymbol{y})$ with a probability density $\mathcal{M}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$, Eq. (1.62), which decreases only as a power of $z_\perp$ (namely, $\mathcal{M} \sim r_\perp^2/z_\perp^4$) when $z_\perp \to \infty$. Therefore, even if at $\tau = 0$ one starts with a gluon distribution which is well localized in the impact parameter space (say, within a disk of radius R_0), the subsequent evolution with τ generates a *power–like tail* at large $b_\perp$, which is then rapidly amplified by the BFKL evolution, with the result that the saturation front expands *exponentially fast* in $b_\perp$.

This 'radial expansion' should be compared to the front propagation in the ρ–direction (with $\rho \sim \ln 1/r_\perp^2$), as discussed in Sect. 4.3.2. In that context, the competition between the exponential decrease of the front with ρ and its exponential increase with τ leads to the formation of a 'traveling wave' which propagates with constant velocity. Here, the front has only a power–law decrease with $b_\perp$, which together with the exponential increase with τ implies that the velocity in the $b_\perp$–direction increases exponentially with τ (see below for a simple estimate). Accordingly, the (perturbatively computed) total cross–section grows like a power of the energy [167], in violation of the Froissart bound [2].
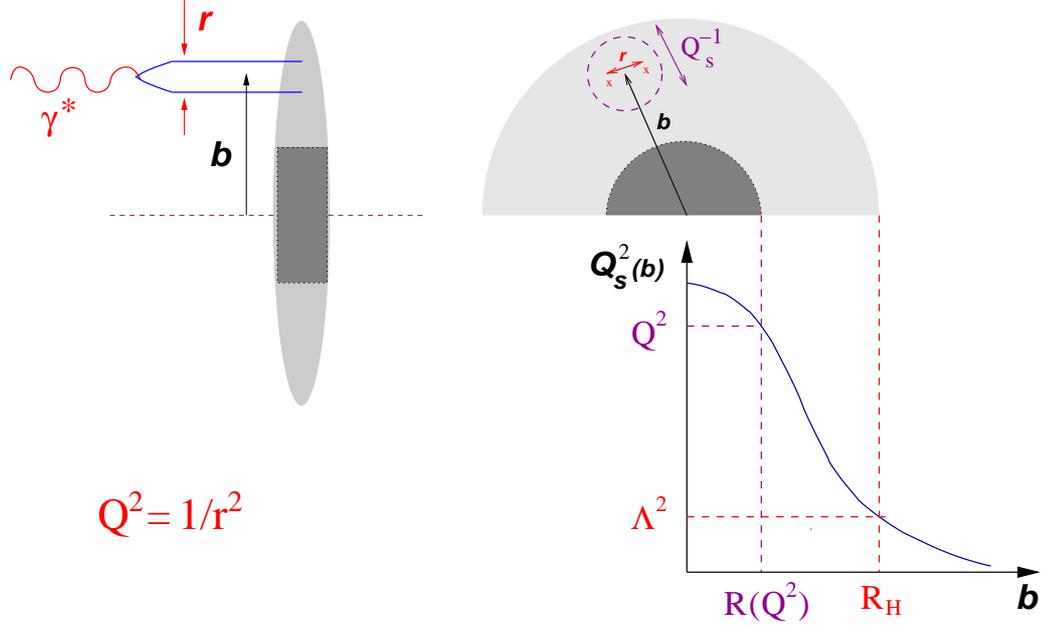
Fig. 39. Dipole-hadron scattering in longitudinal (left) and transverse (right) projections. In the right figure, we also show the b–profile of the saturation momentum (at fixed τ), as corresponding to a 'realistic' hadron, with an exponential tail at large impact parameters (cf. Eq. (4.80)). The black disk corresponds to $b \leq R(Q^2)$ and the perturbative grey corona to $R(Q^2) < b < R_H$.

The fact that the unitarization of the scattering amplitude at fixed impact parameter is not sufficient to guarantee the fulfillment of the Froissart bound should not come as a surprise: As is should be clear from the previous considerations, and is also manifest in the original derivations of the Froissart bound [2–4]), the latter requires some 'screening mass' to eliminate the long–range power tails. In the absence of any such a mass, a theory may very well be unitary, yet not fulfill the Froissart bound; this is, for instance, the case for 'pure QED' — that is, QED seen as an individual theory, taken outside from the Standard Model [18]. But although the gluons are massless in perturbation theory, we know that in real QCD the power tails of the color fields are removed by *confinement* and replaced at large distances by *exponential tails*, with the scale in the exponent of order Λ_{QCD}. Hence, after a rapid expansion at intermediate energies, the saturation front should eventually enter the exponential tail at very high energy, and then slow down and further expand with a constant speed. This in turn implies that the Froissart bound should not only be respected, but actually *saturated* [1, 169].

Let us illustrate the above considerations with a few simple formulae. For an incoming dipole with transverse size $r_\perp$ and at sufficiently high energy, the central region of the target will look 'black', since the local saturation scale $Q_s^2(\tau, b_\perp)$ is larger than $1/r_\perp^2$ at the central points. Let $R(\tau, r_\perp)$ denote the corresponding *black disk radius*, i.e., the position of the saturation front in the transverse plane at rapidity τ (see Fig. 18; we assume isotropy,

---

[18] Within the Standard Model, the cross–sections for processes involving electrons and photons do obey Froissart–like bounds, coming from their coupling to quarks and thus to the strong interactions; see Sect. 5.1 for a discussion of the corresponding bound on the DIS cross–section.

for simplicity):

$$T_\tau(r_\perp, b_\perp) \simeq 1 \qquad \text{for} \qquad b_\perp \leq R(\tau, r_\perp). \tag{4.74}$$

With increasing τ, the black disk expands towards larger values of $b_\perp$, and for sufficiently high energy, it dominates the total cross section:

$$\sigma_{\text{tot}}(\tau, r_\perp) \approx 2\pi R^2(\tau, r_\perp) \qquad \text{(at large } \tau\text{)}. \tag{4.75}$$

To compute $R(\tau, Q^2)$ from the BK equation, it is enough to consider the dynamics at impact parameters well ahead of the front, $b_\perp \gg R(\tau, r_\perp)$, where the scattering amplitude is weak, $T_\tau(r_\perp, b_\perp) \ll 1$, and can be estimated from the linear, BFKL, equation (1.65). One should keep in mind, however, that for a given τ, perturbation theory is valid only within a limited 'grey corona' outside of the 'black disk', namely, within the region where the local saturation scale $Q_s(\tau, b_\perp)$ remains much larger than Λ_{QCD} (see Fig. 18).

For definiteness, let us assume that the target itself has started by being a dipole at low energy, with an original size[19] $R_0 \gg r$. This situation is in fact quite generic, since at high energy and large N_c, any *dilute* target can be described as a collection of color dipoles [58] (see Sect. 6). Thus, the dipole R_0 can be viewed as one of the dipoles making up the dilute part of the wavefunction of a generic target at the original rapidity. The total contribution of these dipoles to the scattering with the external dipole r is simply additive, and here we shall focus on the evolution of a single such a contribution, that of the dipole with an original size R_0.

In perturbative QCD, the dipole–dipole scattering amplitude at low energy $T_0(r, R_0, b)$ can be computed in the two–gluon exchange approximation, with the result shown in Eq. (6.47). Here we shall need this amplitude only for large impact parameters $b \gg R_0$, where it can be easily estimated as follows: the dipole r couples to the color electric field $E_a^i \sim g t^a R_0^i / b^2$ of the target dipole R_0, with an amplitude $T_0 \simeq |g\boldsymbol{r} \cdot \boldsymbol{E}|^2 \sim \alpha^2 r^2 R_0^2 / b^4$. Thus, the original amplitude dies off as $1/b^4$ at large b.

This behaviour will be modified by the BFKL evolution, because of the non–locality of the dipole kernel. At high energy (but still within the weak scattering regime), the corresponding solution can be estimated in the saddle point approximation, with the saddle point $\gamma_0 = 1/2$, cf. Fig. 36. One thus finds the following dominant behaviour at large τ and b [138, 189] (compare to the homogeneous solution in Eq. (1.57)):

$$T_\tau(r, R_0, b) \sim \alpha^2 \frac{rR_0}{b^2} \exp\left\{ \omega_0 \bar\alpha \tau - \frac{2\ln^2(16b^2/rR_0)}{\beta_0 \bar\alpha \tau} \right\}, \tag{4.76}$$

where $\omega_0 \equiv \chi(1/2) = 4\ln 2$ and $\beta_0 = \chi''(1/2) \approx 33.67$ (cf. Eq. (4.27)). As compared to the initial condition T_0, the prefactor rR_0/b^2 in Eq. (4.76) decays twice as slow at large b, which reflects the BFKL 'anomalous dimension' $1 - \gamma_0 = 1/2$.

By supplementing Eq. (4.76) with the saturation conditions (4.23) or (4.74), one can estimate the saturation momentum and the black disk radius, respectively. One thus finds [167, 168, 190]

---

[19] For simplicity, $r_\perp$ and $b_\perp$ will be denoted as r and b in the remaining part of this section.

$$Q_s^2(\tau, b) \simeq \frac{R_0^2}{256\, b_\perp^4}\, \mathrm{e}^{c\bar{\alpha}\tau}\,, \qquad \sigma_{\mathrm{tot}}(\tau, Q^2) \approx 2\pi R^2(\tau, r) \simeq \frac{\pi}{8}\, r R_0\, \mathrm{e}^{(c/2)\bar{\alpha}\tau}\,, \qquad (4.77)$$

where

$$c = \frac{1}{2}\left(-\beta_0 + \sqrt{\beta_0(\beta_0 + 8\omega_0)}\right) \approx 4.8 \qquad (4.78)$$

is essentially the same saturation exponent as in Eq. (4.36), and has been reobtained here via an expansion [20] around the high–energy BFKL saddle point $\gamma_0 = 1/2$.

As we have anticipated, the evolution described by the BK equations leads to power–like tails in the dipole scattering amplitude and the target gluon distribution (as measured, e.g., by the respective saturation momentum), and thus to a dipole cross–section growing exponentially with τ. These semi–quantitative estimates are in agreement with the results of a numerical resolution of the BK equation in impact parameter space, in Ref. [168].

However, as also emphasized before, this perturbative dynamics becomes physically unacceptable at high energy, as it violates an essential, but non–perturbative, property of QCD: the color confinement. To understand the origin of the problem, it is instructive to consider in more detail the evolution leading to strong scattering at large impact parameters. This is most intuitively described in the dipole language valid at large N_c, in which the original dipole R_0 evolves into a collection of dipoles via successive $1 \to 2$ dipole splitting (see Sect. 6 for details). The average evolution is quasi–local, as explained above Eq. (4.22) : the typical child dipoles have transverse sizes and impact parameters comparable to those of the parent dipole R_0. But such *typical* child dipoles contribute very little to the scattering of the external dipole r, since their respective contributions are suppressed as $1/b^4$. Rather, the dominant contributions to $T_\tau(r, R_0, b)$ come from the *rare fluctuations* in which a gluon is radiated at impact parameters close to b, that is, far outside the original dipole R_0. The probability of such a fluctuation is suppressed as $1/b^4$ too, because of the large–distance behaviour of the dipole kernel, but its contribution to the scattering amplitude $T_\tau(r, R_0, b)$ is rapidly amplified by the subsequent gluon evolution around b, for which this fluctuation acts as a seed. This pattern for the evolution is clearly seen in the numerical simulations of the dipole picture by Salam [44, 45].

But in QCD, color fields cannot propagate through the vacuum over distances larger than the confinement scale, of order $1/\Lambda_{\mathrm{QCD}}$. Thus, the above perturbative picture makes physical sense only so long as $b \ll 1/\Lambda_{\mathrm{QCD}}$. Presently, we do not know how to modify the evolution equations in order to systematically account for the effects of confinement; but one can at least heuristically do so, e.g., by modifying the splitting kernel by hand in such a way to suppress the radiation of very large dipoles, with sizes $\gtrsim 1/\Lambda_{\mathrm{QCD}}$. On physical grounds, one expects this suppression to be *exponential* [169] : the probability for emitting a large dipole with size b should decrease like $\mathrm{e}^{-\mu b}$, with μ a 'screening mass' (or infrared cutoff) of order Λ_{QCD}. A similar suppression must be introduced also in the

---

[20] This approximate calculation yields relatively good estimates for both the saturation exponent c and the anomalous dimension γ_s [51] because γ_s is not very different from γ_0 and the function $\chi(\gamma)$ is only slowly varying around its minimum at $\gamma_0 = 1/2$; see Fig. 36.

initial condition T_0 : the physical scattering amplitude between a dipole and a $q\bar{q}$ 'meson' (a true bound state of QCD) should decay exponentially at large distances.

With this in mind, let us return to our analysis of the scattering amplitude at large impact parameters: A relevant, rare, fluctuation, resulting in a dipole with a leg at b, occurs with a small probability $\sim \mathrm{e}^{-\mu b}$, but once that this happens, the subsequent evolution of the dipole distribution around b is quasi–local, since governed by the typical BFKL dynamics. (We implicitly assume here that the point b of interest lies within the perturbative 'grey corona'; see Fig. 18.) Therefore, the 'soft' inhomogeneity over the scale $1/\mu$ does not interfere with the subsequent BFKL evolution around b, so that the b–dependence of the final amplitude factorizes out:

$$T_\tau(r,b) \simeq T_\tau(r)\,\mathrm{e}^{-\mu b} \sim \frac{r}{R_0}\,\exp\left\{-\mu b + \omega_0 \bar{\alpha}\tau\right\}, \tag{4.79}$$

where $T_\tau(r)$ is the solution to the homogeneous BFKL equation (4.22), and we have kept only the dominant behavior of the latter at high energy (cf. Eq. (1.57)). From this equation and the saturation conditions (4.23) and (4.74), one can now estimate the saturation momentum and the black disk radius. One obtains (with $Q_0 \equiv 1/R_0$) [169]

$$Q_s^2(\tau,b) \simeq Q_0^2\,\mathrm{e}^{2\omega_0\bar{\alpha}\tau}\mathrm{e}^{-2\mu b}, \qquad R(\tau,r) \simeq \frac{1}{\mu}\left(\omega_0\bar{\alpha}\tau - \ln\frac{R_0}{r}\right), \tag{4.80}$$

which in turn implies the following result for the total cross section at high energy [169]

$$\sigma_{\mathrm{tot}}(s,r) \approx 2\pi\left(\frac{\omega_0\bar{\alpha}}{\mu}\right)^2 \ln^2 s \qquad \text{when} \qquad s \to \infty. \tag{4.81}$$

These results apply in the high energy limit $s \to \infty$, when the black disk is already much larger than the original target size R_0. In this limit, and within the present assumptions, the Froissart bound appears to be *saturated*, with a proportionality coefficient which is *universal* (i.e., the same for all hadrons), and which reflects the combined role of perturbative (BFKL) and non–perturbative (confinement) physics in controlling the asymptotic behaviour of the total cross–section at high energy. Numerical simulations of modified versions of the BK equation in which the infrared cutoff μ is introduced by hand confirm the conclusions in Eqs. (4.80)–(4.81) [191, 192].

In Ref. [169] it has been further argued that the characteristic screening mass μ should be equal to twice the pion mass: $\mu = 2m_\pi$. This is because the pion is the lightest bound state in QCD and because of isospin conservation: The long–range scattering should be controlled by pion exchange, but pions have isospin one, while gluons have isospin zero, so the scattering must proceed via the exchange of (at least) two pions.

At this point, one should recall that the Froissart bound $\sigma_{\mathrm{tot}} \leq \sigma_0 \ln^2 s$ [2] is a consequence of general principles (unitarity, crossing, analiticity, and finite range of the interactions), but does not rely on detailed dynamical information. So, in reality, this bound may very well not be saturated. It so happens, however, that the measured cross–sections (e.g., for pp and $p\bar{p}$ scattering) show a slow, but monotonous, increase with s, which are indeed well fitted by a $\ln^2 s$ behaviour [193]. This suggests that the Froissart bound is actually

saturated in nature. Then the simple mechanism described above provides a physical picture for such a saturation, with a definite prediction for the scale σ_0. Note however that this is merely a crude estimate, because of the use of the leading–order perturbative QCD formalism and, especially, because of the heuristic treatment of confinement in the study of the expansion of the black disk.

To conclude, while perturbation theory alone appears to be sufficient to describe unitarization at fixed impact parameter, one still needs some information about the finite range of the strong interactions in order to be able to compute total cross-sections. This is reminiscent of an old argument by Heisenberg [1] which combines unitarity and short–rangeness (as modelled by a Yukawa potential) to deduce cross sections which saturate the Froissart bound. Fifty years later, our progress in understanding high energy strong interactions allows us to confirm Heisenberg's intuition, and identify short–rangeness with confinement and unitarization with gluon saturation.

5 Some consequences for the phenomenology at HERA & RHIC

In this section, we shall discuss two applications of the theory for the CGC to the present–day phenomenology: a fit to the HERA data for deep inelastic scattering (DIS) at small x, and an interpretation of the RHIC data for particle production in deuteron–gold (d–Au) collisions. These applications have been selected because of their potential to test some very non–trivial predictions of the theory, which have the merit to be robust and qualitative. For DIS, these predictions are the approximate geometric scaling (including the pattern of its violations) and the tame of the growth of the structure function F_2 at sufficiently low values of x and/or Q^2. For the d–Au collisions at RHIC, these are the multiple scattering in a large nucleus (leading to a Cronin peak in the particle production at central rapidity) and the evolution of the nuclear gluon distribution with $1/x$ in the presence of saturation (which can explain the 'high–$k_\perp$ suppression' in the particle production at forward rapidities). The experimental evidence for the phenomena to be discussed here has been presented in Sects. 1.2 and 1.3, for HERA and RHIC respectively.

5.1 Geometric scaling and its violations at HERA

In our precedent discussion of the small–x experimental situation at HERA, in Sect. 1.2, we have already emphasized the difficulty of extracting compelling evidences for saturation from the data. This is related to the kinematical limitations of the present experiments, and also to our limited understanding of QCD at small–x.

On the other, we have also mentioned there that simple models inspired by the idea of saturation — namely the Golec-Biernat–Wüsthoff "saturation model" [73] and its various extensions [74, 75] — were successful in explaining, or even predicting, some remarkable regularities of the small–x data, like the geometric scaling [83], the weak energy dependence of the ratio between the diffractive and the inclusive structure functions, and the qualitative changes in the behaviour of F_2 at very small x and/or Q^2 (the progressive taming of its rise with $1/x$, and the approach towards the 'photoabsorption limit'

141

$Q^2 \to 0$) [74, 75]. These simple models were also able to provide reasonable fits, which involved only few parameters (three to five), to the original HERA data for F_2 [65], but the quality of these fits (in terms of χ^2 = variance per degree of freedom) became unsatisfactory with the advent of the new HERA data [17], which are much more precise.

However, the parametrizations used in these original models for the dipole cross–section (cf. Sect. 1.2.2) were rather *ad–hoc*, and certainly too crude to capture the full complexity of the QCD dynamics even within the limited kinematical range in which they were used. In the mean time, our theoretical understanding of gluon saturation and unitarity corrections has substantially progressed, as explained in the previous sections, and the conceptual limitations of the original "saturation models" have become even more obvious. It is therefore interesting to check whether one can use our improved knowledge of QCD at small x in order to achieve a more satisfactory description of the respective HERA data. The answer turns out to be positive, as demonstrated by several recent analyses [54, 76, 194, 195]. In what follows, we shall focus on the 'CGC model' of Ref. [54], which is closest in spirit to the original 'saturation model' by Golec-Biernat and Wüsthoff : it involves the same set of (three) free parameters and it includes the minimal modifications necessary to render the dipole cross–section consistent with the CGC analysis in Sect. 4. But before that, it is interesting to understand in more detail the interplay between the virtual photon wavefunction and the dipole cross–section in determining the high–energy DIS cross–section according to Eq. (1.22).

5.1.1   Dipole factorization: a qualitative analysis

In this subsection, we shall perform an analytic but qualitative study of the DIS cross–section with the purpose of identifying the typical dipole sizes contributing to the convolution in Eq. (1.22) and estimating their contribution. We shall thus see how some important properties like geometric scaling or the unitarization transmit from the scattering of the dipole to that of the virtual photon.

To start with, let us be more specific about the dipole factorization of the inclusive DIS cross–section at high energy, as expressed by Eq. (1.22). This can be decomposed as

$$\sigma_{\gamma^* p} = \sigma_T + \sigma_L$$
$$\sigma_{T/L}(x, Q^2) = \int_0^1 dz \int d^2\boldsymbol{r} \, |\Psi_{T/L}(z, r; Q^2)|^2 \, \sigma_{\text{dipole}}(x, r), \tag{5.1}$$

where Ψ_α, with $\alpha = T$ or L, is the LC wavefunction which describes the $q\bar{q}$ fluctuation of a virtual photon with transverse, or longitudinal, polarization. Furthermore, $r = |\boldsymbol{r}|$ is the dipole size and z is the fraction of the longitudinal momentum of the virtual photon which is taken away by the quark component of the dipole. To lowest order in the electromagnetic coupling, one has [71, 72]

$$|\Psi_T(z, r; Q^2)|^2 = \frac{\alpha_{\text{em}} N_c}{2\pi^2} \sum_f e_f^2 \left\{ \left(z^2 + (1-z)^2 \right) \bar{Q}_f^2 K_1^2(\bar{Q}_f r) + m_f^2 K_0^2(\bar{Q}_f r) \right\}, \tag{5.2}$$

$$|\Psi_L(z, r; Q^2)|^2 = \frac{\alpha_{\text{em}} N_c}{2\pi^2} \sum_f e_f^2 \left\{ 4Q^2 z^2 (1-z)^2 K_0^2(\bar{Q}_f r) \right\}. \tag{5.3}$$

In these equations, $\bar{Q}_f^2 \equiv z(1-z)Q^2 + m_f^2$ (e_f and m_f are the electric charge and, respectively, the mass for the quark with flavor f) and K_0 and K_1 are the modified Bessel functions. The factor of N_c comes from summing over the color degrees of freedom of the quarks. To get a simple estimate for the DIS cross–section in the high–energy regime at $x \ll 1$ and $Q^2 \gg \Lambda_{\mathrm{QCD}}^2$, we shall ignore the quark masses [21] (a reasonable approximation so long as $Q^2 \gg m_f^2$) and use simple approximations for the Bessel functions. Specifically, since $K_\nu(x)$ decreases exponentially at large values of x, one can restrict the integrations in Eq. (5.1) to values z and r such that $\bar{Q}_f r < 1$. Besides, for $x \ll 1$, $K_1(x) \sim 1/x$ and $K_0(x) \sim (-\ln x)$, so the Bessel functions in Eqs. (5.2)–(5.3) can be approximately replaced by:

$$K_0(x) \sim \Theta(1-x), \quad K_1(x) \sim \frac{1}{x}\Theta(1-x), \tag{5.4}$$

where we ignore the overall normalization as well as the logarithmic singularity of K_0 as $x \to 0$, since the latter is innocuous within the relevant convolutions. (Note that we shall be merely interested here in order–of–magnitude estimates.) A final simplification can be performed when evaluating the integral over z: the constraint $z(1-z)Q^2 r^2 < 1$ together with the fact that $z(1-z) \leq 1/4$ for $0 \leq z \leq 1$ makes it natural to distinguish between

(i) *symmetric* configurations, for which $Q^2 r^2 < 4$ ("small dipoles"), and

(ii) *aligned jet* configurations, with $4 < Q^2 r^2$ ("large dipoles").

In the first case, there is no restriction on the v integral, which is then dominated by symmetric values $v \sim 1/2$: the total longitudinal momentum is "democratically" divided among the quark and the antiquark. In the second case, the dominant contributions correspond to the situation in which one of the two dipoles carries most of the total longitudinal momentum, that is, v is either close to zero, or close to one, which allows us to further simplify the respective integrand.

As a result of such simplifications, the integrations involving the transverse and longitudinal photon wavefunctions are estimated as

$$\int \mathrm{d}z\, \mathrm{d}^2\boldsymbol{r}\, |\Psi_T(z,r;Q^2)|^2\, f(r) \sim \alpha_{\mathrm{em}} N_c \sum_f e_f^2 \left[\int_0^{2/Q} \frac{\mathrm{d}r}{r}\, f(r) + \frac{1}{Q^2} \int_{2/Q}^{\infty} \frac{\mathrm{d}r}{r^3}\, f(r) \right], \tag{5.5}$$

and, respectively,

$$\int \mathrm{d}z\, \mathrm{d}^2\boldsymbol{r}\, |\Psi_L(z,r;Q^2)|^2\, f(r) \sim \alpha_{\mathrm{em}} N_c \sum_f e_f^2 \left[Q^2 \int_0^{2/Q} \mathrm{d}r\, r f(r) + \frac{1}{Q^4} \int_{2/Q}^{\infty} \frac{\mathrm{d}r}{r^5}\, f(r) \right], \tag{5.6}$$

where in writing the r.h.s.'s we have kept the various parametric dependencies, but ignored all numerical factors. For each of these expressions, the first term within the square brack-

---

[21] For the phenomenology at HERA, quark masses are important though, and will be reintroduced in the respective analysis later on.

ets corresponds to the symmetric configurations, and the second one to the aligned–jet configurations.

Turning now to the dipole cross–section, this will be estimated as $\sigma_{\text{dipole}}(x, r) \approx \sigma_0 T(r, x)$, where σ_0 is a geometric cross–section: $\sigma_0 \equiv 2\pi R^2$ with R the 'proton radius' (more precisely, the radius of the central area of the hadron disk, which contributes significantly to scattering) and $T(r, x)$ is the dipole–hadron scattering amplitude averaged over impact parameters within the relevant area. For the dipole amplitude, we shall use a simple, piecewise, approximation suggested by the calculations in the previous section (i.e., by the solution to the BK equation). Specifically, we shall use

$$\sigma_{\text{dipole}}(x, r) \approx \begin{cases} \sigma_0 \, r^2 Q_s^2(x) & \text{for} \quad r \lesssim 1/Q_g(x) \\ \sigma_0 \, (r^2 Q_s^2(x))^{\gamma_s} & \text{for} \quad 1/Q_g(x) \ll r \ll 1/Q_s(x) \\ \sigma_0 & \text{for} \quad r \gg 1/Q_s(x) \,, \end{cases} \quad (5.7)$$

where $Q_s(x)$ is the target saturation momentum (averaged over impact parameters in the central area) and $Q_g^2(x) \equiv Q_s^4(x)/\Lambda_{\text{QCD}}^2$ is the high–Q^2 boundary of the geometric scaling window (cf. Fig. 38). In writing Eq. (5.7) we have distinguished between the 'color transparency' regime at very small r, the 'geometric scaling window' at intermediate values of r, and the 'saturation' region at very large r. More subtle phenomena, like the DGLAP evolution at very small r, the violation of geometric scaling through BFKL diffusion, or the expansion of the black disk at very high energy, have been ignored, since unimportant for a qualitative discussion. (Some of these effects will be reintroduced later.)

We are now equipped to evaluate the convolutions appearing in Eq. (5.1). To simplify writing, in the remaining part of this section we shall omit the overall factor $\alpha_{\text{em}} N_c \sum_f e_f^2$, which is common to all the cross–sections.

(L) Consider the *longitudinal* sector first: It is then easily verified that, within the whole kinematic range in which the present approximations are justified, the cross–section $\sigma_L(x, Q^2)$ is dominated by dipole sizes $r \sim 2/Q$, and thus receives lowest–order contributions from both the symmetric and the aligned–jet configurations. Let us check this for the large–Q^2 regime at $Q^2 \gg Q_s^2(x)$. One can then write (below, γ is either equal to one, or equal to γ_s, depending upon the value of Q^2) :

$$\frac{\sigma_L}{\sigma_0} \sim Q^2 \int_0^{2/Q} \mathrm{d}r \, r(r^2 Q_s^2)^\gamma + \frac{1}{Q^4} \int_{2/Q}^{1/Q_s} \frac{\mathrm{d}r}{r^5} (r^2 Q_s^2)^\gamma + \frac{1}{Q^4} \int_{1/Q_s}^\infty \frac{\mathrm{d}r}{r^5} \sim \left(\frac{Q_s^2(x)}{Q^2} \right)^\gamma, \quad (5.8)$$

where the aligned–jet integral has been split in two: $2/Q < r < 1/Q_s$ and $1/Q_s < r < \infty$. The first two integrals in Eq. (5.8) are dominated by dipole sizes of order $2/Q$ — i.e., of the order of the upper/lower cutoff for the first/second integral — and contribute both to the final result shown in the r.h.s. The third integral, with lower cutoff $1/Q_s$, gives a 'higher–twist' contribution of order Q_s^4/Q^4, which is power–suppressed in the high–Q^2 regime at $Q^2 \gg Q_s^2$. A similar analysis of the low–Q^2 regime at $Q^2 \ll Q_s^2(x)$ yields

$$\sigma_L(x, Q^2) \sim \sigma_0 \qquad \text{for} \qquad Q^2 \ll Q_s^2(x) \,. \quad (5.9)$$

By comparing the above results to Eq. (5.7), we deduce that $\sigma_L(x, Q^2)$ is proportional to the dipole cross–section for a dipole with size $2/Q$, and thus is a direct probe of the gluon distribution in the target (so like the dipole itself). Unfortunately, it turns out that $\sigma_L(x, Q^2)$ (or $F_L(x, Q^2)$) is rather difficult to extract from the data — the quantity which is most directly measured in the experiment is the complete structure function $F_2 = F_T + F_L$ — and the accuracy of the current measurements of F_L at HERA is not good enough to impose strong constraints on the saturation physics.

(T) While unessential in the longitudinal sector, the distinction between symmetric and aligned–jet configurations becomes important in the *transverse* sector, as we show now:

At large Q^2, it is further necessary to distinguish between values of Q^2 above, or within, the geometric scaling window, since the cross–section $\sigma_T(x, Q^2)$ shows different behaviors in these two regimes. Namely, for Q^2 within the range $Q_s^2 \ll Q^2 \ll Q_g^2$, the situation is similar to that in the longitudinal sector: the dominant contribution to σ_T is of order $(Q_s^2/Q^2)^{\gamma_s}$ and comes from dipole sizes $r \sim 2/Q$, i.e. from both symmetric and aligned–jet configurations. But when $Q^2 \gg Q_g^2$, the situation changes : the dominant term is then produced by the *aligned–jet* configurations alone and, more precisely, by dipole sizes r which are logarithmically distributed within the range $2/Q < r < 1/Q_g$. Indeed, in this regime we can decompose the integrals in Eq. (5.5) as follows

$$
\frac{\sigma_T}{\sigma_0} \sim \int_0^{2/Q} \frac{\mathrm{d}r}{r}\, r^2 Q_s^2 + \frac{1}{Q^2} \int_{2/Q}^{1/Q_g} \frac{\mathrm{d}r}{r^3}\, r^2 Q_s^2 + \frac{1}{Q^2} \int_{1/Q_g}^{1/Q_s} \frac{\mathrm{d}r}{r^3}\, (r^2 Q_s^2)^{\gamma} + \frac{1}{Q^2} \int_{1/Q_s}^{\infty} \frac{\mathrm{d}r}{r^3}
$$

$$
\sim \frac{Q_s^2}{Q^2} \ln\left(\frac{Q^2}{Q_g^2}\right) \qquad \text{for} \qquad Q^2 \gg Q_g^2(x)\,, \tag{5.10}
$$

where the dominant term, as isolated in the second line, is enhanced with respect to the subleading ones, as coming from all the other integrals shown above, by the large logarithm $\ln(Q^2/Q_g^2) \simeq \ln(Q^2/Q_s^2)$. Note that the subleading terms in this calculation are not power–suppressed anymore — they are all of order Q_s^2/Q^2 —, so a complete calculation of σ_T to 'leading–twist' order (i.e., to leading order in $1/Q^2$) must include all such terms.

A similar analysis of the low–Q^2 regime at $Q^2 \ll Q_s^2(x)$ shows that, in that case, σ_T is dominated by the *symmetric* dipole configurations with dipole sizes r logarithmically distributed within the range $1/Q_s < r < 2/Q$. This leads to a logarithmic enhancement once again, as visible in the third line of Eq. (5.11) below. To summarize, within the present approximations we obtain the following parametric estimates for σ_T :

$$
\sigma_T(x, Q^2) \approx \begin{cases} \sigma_0 (Q_s^2/Q^2) \ln(Q^2/Q_s^2) & \text{for } Q^2 \gg Q_g^2 \\ \sigma_0 (Q_s^2/Q^2)^{\gamma_s} & \text{for } Q_s^2 \ll Q^2 \ll Q_g^2 \\ \sigma_0 \ln(Q_s^2/Q^2) & \text{for } Q^2 \ll Q_s^2\,. \end{cases} \tag{5.11}
$$

At this level, several comments are in order:

(i) The previous results show that, for all Q^2 except in the intermediate range for geomet-

ric scaling, the transverse cross–section σ_T is parametrically larger than the longitudinal one σ_L (by a large logarithm in our estimates). This is in agreement with the experimental observation that $F_2 = F_T + F_L$ is dominated by its transverse piece F_T [69].

(ii) For large values of Q^2, the transverse cross–section σ_T receives 'leading–twist' contributions from the relatively large dipoles, with sizes of the order of the unitarization scale $1/Q_s$ (see the last integral in Eq. (5.10)). This shows that the very *calculability* of σ_T, and thus of F_2, within perturbative QCD relies crucially on the existence of perturbative gluon saturation, and thus it becomes possible only for sufficiently high energies, or small values of x, such that $Q_s^2(x) \gg \Lambda_{\mathrm{QCD}}^2$. By contrast, at lower energies, where the physics of unitarization is non–perturbative, an integral like the last one [22] in Eq. (5.10) is effectively cut–off (by the physics of confinement) at the 'soft' scale $R \sim 1/\Lambda_{\mathrm{QCD}}$, and thus generates a non–calculable 'leading–twist' term $\sim \Lambda_{\mathrm{QCD}}^2/Q^2$. Such a term does not contribute to the 'evolution' of $F_2(x, Q^2) \propto Q^2 \sigma_{\gamma^* p}$ with either x or Q^2 — hence, the *evolution* of the proton structure function at high Q^2 can be computed in perturbative QCD, as well known [15, 16] —, however it affects its *absolute value* already at leading–twist order.

(iii) The DIS cross–sections in Eqs. (5.8), (5.9) and (5.11) show *geometric scaling*, that is, they depend upon x and Q^2 only via the dimensionless ratio $Q^2/Q_s^2(x)$. This property looks very natural within the present calculation: after having neglected the quark masses and the impact–parameter dependence, and in the absence of any infrared problem, the scaling behavior of the dipole cross–section, Eq. (5.7), automatically transmits to the virtual–photon cross–section, as clear from Eqs. (5.1)–(5.3) via dimensional analysis. Note, however, that this behavior is natural only in the context of saturation, that is, for sufficiently small values of x: Both the scaling form of the dipole cross–section (5.7) and the fact that Q_s acts effectively as an infrared cut–off in the relevant integrals (see, e.g., Eq. (5.10)) are essential consequences of the mechanism of saturation. In fact, the previous results show perfect scaling up to *arbitrarily* high values of Q^2, but this is just an artifact of our oversimplified form for the dipole cross–section (5.7): From Sect. 4.3, we know that the geometric scaling starts to be violated already within the scaling window, by the BFKL diffusion (cf. Eq. (4.38)), and it is totally broken at $r \lesssim 1/Q_g$, by the DGLAP evolution (cf. Eq. (4.41)). The effects of the BFKL diffusion on the DIS cross–section will be included in the phenomenological analysis in the next subsection.

(iv) The unitarization of the dipole amplitude for large dipole sizes $r \gtrsim 1/Q_s(x)$ transmits into a corresponding property for the DIS cross–sections at relatively low virtualities $Q^2 < Q_s^2(x)$. This implies an important turnover in the Q^2 and x–dependencies of the proton structure function F_2: whereas at large $Q^2 \gg Q_s^2(x)$, $F_2 \propto Q_s^2(x) \ln(Q^2/Q_s^2)$ shows a rapid, power–like, increase with $1/x$, but only a weak, logarithmic, dependence upon Q^2 (the 'Bjorken–scaling violation', attributed here to the DGLAP–like evolution of the virtual photon, cf. Eq. (5.10)), at lower values of $Q^2 < Q_s^2(x)$, $F_2 \propto Q^2 \ln(Q_s^2/Q^2)$ vanishes like Q^2 and increases only slowly, logarithmically, with $1/x$. Note that the strong deviation from Bjorken scaling, in the sense of a power–like dependence upon Q^2, starts already at momenta $Q^2 > Q_s^2(x)$, namely, within the geometric scaling window of Fig. 38, where

---

[22] Similarly, the logarithmic divergence of the second integral there would be cutoff at the scale R, this yielding a 'leading–twist' contribution enhanced by the logarithm $\ln(Q^2/\Lambda_{\mathrm{QCD}}^2)$.

$F_2 \propto Q_s^2(x)(Q^2/Q_s^2)^{1-\gamma_s}$ with $1 - \gamma_s \approx 0.37$.

(v) From Eq. (5.11) one sees that, even when σ_{dipole} saturates at a fixed value σ_0 (as assumed in Eq. (5.7)), the virtual photon cross–section σ_T has still a slow, logarithmic, increase with $1/x$ in the high–energy regime: $\sigma_T \propto \ln(Q_s^2/Q^2) \sim \ln(1/x)$ when $Q_s^2(x) \gg Q^2$. This is due to the increase in the transverse phase–space available for the dissociation of the (transverse) virtual photon into a $q\bar{q}$ pair with size r within the range $1/Q_s < r < 2/Q$. It is then easily to understand what happens if, instead of using Eq. (5.7), one rather assumes the dipole cross–section to *saturate* the Froissart bound at very high energy (cf. the discussion in Sect. 4.4). One then finds

$$\sigma_{\text{dipole}}(s, r) \approx \sigma_0 \ln^2 s \quad \text{when} \quad s \to \infty \quad \Longrightarrow \quad F_2(x, Q^2) \propto \ln^3(1/x) \quad \text{when} \quad x \to 0,$$
(5.12)

where the third factor of $\ln(1/x)$ in F_2 has been generated, as expected, by the transverse phase–space $\ln(Q_s^2/Q^2)$ for the dissociation of γ^*. The fact that the virtual–photon cross–section appears to grows faster with the energy than permitted by the Froissart bound is not contradictory: Froissart bound applies only to properly normalized hadronic states, so like the dipole. But γ^* is a virtual *superposition* of dipole states, with a weight function $|\Psi|^2$ which is not properly normalized. (Indeed, the integral of $|\Psi_T|^2$ over r has a logarithmic singularity at $r \to 0$, which becomes manifest when replacing $f = 1$ in Eq. (5.5). This means that there is an infinite probability for emitting very small dipoles, which is however innocuous since such small dipoles do not contribute significantly to physical quantities like σ_T — as obvious in Eq. (5.10).) Thus, there is *a priori* no analog of the Froissart bound for F_2 ; the one in Eq. (5.12) is merely induced by the corresponding bound on the dipole scattering together with the properties of the superposition yielding γ^*.

5.1.2 The CGC fit to F_2

The CGC–inspired phenomenological analysis in Ref. [54] has focused on the F_2 data within the following kinematical range: $x \leq 10^{-2}$ and $Q^2 < 50 \, \text{GeV}^2$. This is the range where one expects important high density effects (in particular, it includes the transition region from high to low Q^2). The upper limit on Q^2 has been chosen large enough to include a significant number of "perturbative" data points, but low enough to justify the emphasis on BFKL, rather than DGLAP, evolution (for the small–x values of interest). In fact, $50 \, \text{GeV}^2$ is roughly consistent with the estimated upper bound $Q_s^4(x)/\Lambda_{\text{QCD}}^2$ of the kinematical window for BFKL behaviour (cf. Eq. (4.62)).

The analysis is based on a model for the dipole cross–section which is inspired by the solution to the non–linear BK equation discussed in Sect. 4 and which involves the same set of three free parameters as the GBW "saturation model" [73] (recall Eq. (1.26)): namely, the geometric cross–section $\sigma_0 \equiv 2\pi R^2$ (with R the 'proton radius'), the value x_0 at which $Q_s^2(x_0) = 1 \, \text{GeV}^2$, and the saturation exponent λ. The need for the first two parameters reflects the fact that, even in a first principle calculation including saturation, some aspects of the calculation remain genuinely non–perturbative: the impact parameter dependence of the scattering amplitude and the initial condition at low energy. (A more realistic impact parameter dependence can be implemented along the lines of Refs. [75, 76].) We thus write

$\sigma_{\text{dipole}}(x, r) = 2\pi R^2 T(x, r)$ with $T(x, r)$ the approximate solution to the homogeneous BK equation constructed in Sect. 4.2. Given the universal form of this solution at high energy (cf. the discussion after Eq. (4.40)), the initial condition enters through a single parameter: the overall normalization scale Q_0^2 of the saturation momentum (cf. Eq. (4.36)) or, equivalently, the value x_0 of x at which Q_s equals 1GeV.

The parameter λ, on the other hand, *is* computable in perturbation theory, and presently this is known to NLO, as discussed in Sect. 4.2. Namely, the NLO calculation of Ref. [53], which is summarized by the plots in Fig. 37, shows that $\lambda(\tau) \equiv d \ln Q_s^2(\tau)/d\tau$ (with $\tau = \ln 1/x$) is not simply a constant, but rather a slowly varying function, which decreases from $\lambda \approx 0.30$ for $\tau = 5$ to $\lambda \approx 0.27$ for $\tau = 15$, with a theoretical uncertainty of about 15%. Unfortunately, this uncertainty is still too large to permit a good description of the data: λ is the exponent which controls the growth of F_2 with τ at large Q^2, so the data are very sensitive to its precise value. To cope with that, we shall treat λ as a free parameter, as announced. Remarkably, the value of λ that will come out from the fit, namely $\lambda \approx 0.25 - 0.29$, is consistent with the theoretical prediction of Ref. [53], within the theoretical uncertainty alluded to above.

The function $T(x, r)$ to be used in the fit is constructed by smoothly joining together the limiting behaviours of the BK solution, which are analytically under control: These are encoded in Eq. (4.39) for small dipole sizes, $r \ll 1/Q_s(x)$, and, respectively, in Eq. (4.19) for larger dipoles, $r \gg 1/Q_s(x)$. We implicitly assume here that the smallest values of r to be relevant for the fit are those within the geometric scaling window at $1/Q_g \lesssim r \ll 1/Q_s$, where Eq. (4.39) applies. This is a reasonable approximation for the not–too–large values of Q^2 of interest, as shown by the qualitative analysis in the previous subsection. We thus write: $T(x, r) \equiv \mathcal{T}(rQ_s, \tau)$, with

$$\mathcal{T}(rQ_s, \tau) = \mathcal{T}_0 \left(\frac{rQ_s}{2} \right)^{2\left(\gamma_s + \frac{\ln(2/rQ_s)}{\kappa \lambda \tau} \right)} \qquad \text{for} \quad rQ_s \leq 2,$$
$$\mathcal{T}(rQ_s, \tau) = 1 - \mathrm{e}^{-a \ln^2(b \, rQ_s)} \qquad \qquad \text{for} \quad rQ_s > 2, \qquad\qquad (5.13)$$

where $Q_s(x) = (x_0/x)^{\lambda/2}$ GeV, $\gamma_s \approx 0.63$ and $\kappa \equiv \beta/c \approx 9.9$. The first line in Eq. (5.13) is essentially a rewriting of Eq. (4.39) (in particular, the coefficients γ_s and κ are fixed to their LO BFKL values). Also, Q_s has been defined in such a way that $\mathcal{T}(rQ_s, \tau) = \mathcal{T}_0$ for $rQ_s = 2$, so like in the corresponding GBW formula, Eq. (1.26). The coefficients a and b in the second line are uniquely determined from the condition that $\mathcal{T}(rQ_s, \tau)$ and its slope be continuous at $rQ_s = 2$. The value $\mathcal{T}_0$ of the scattering amplitude at saturation is ambiguous, reflecting an ambiguity in the definition of Q_s. The fit has been repeated for various values of $\mathcal{T}_0$ between 0.5 and 0.9. It has been checked in Ref. [54] that the quality of the fits is not very sensitive to the details of the interpolation, nor to the precise form of the approach towards the black–disk limit $T = 1$. On the other hand, the data are quite sensitive to the details of the scattering amplitude at smaller sizes $r \leq 1/Q_s(x)$, and thus provide a test of the BFKL dynamics in the presence of saturation.

As compared to the GBW parametrization (1.26), the amplitude in Eq. (5.13) exhibits the 'anomalous dimension' $1 - \gamma_s \approx 0.37$, which describes deviations from 'color transparency' at small r, and involves geometric scaling violations through the 'BFKL diffusion term'.

Both effects are hallmarks of the BFKL dynamics within the extended scaling window (cf. Fig. 38). Furthermore, the approach towards the unitarity limit $T = 1$ is much slower in Eq. (5.13) than in the original GBW formula (1.26), or in any other approach assuming Glauber–like exponentiation [74, 75]. The importance of the scaling violations for the quality of the fit will be emphasized below.

$\mathcal{T}_0$/model	0.5	0.6	0.7	0.8	0.9	GBW
χ^2	146.43	129.88	123.63	125.61	133.73	243.87
χ^2/d.o.f	0.96	0.85	0.81	0.82	0.87	1.59
x_0 ($\times 10^{-4}$)	0.669	0.435	0.267	0.171	0.108	4.45
λ	0.252	0.254	0.253	0.252	0.250	0.286
R (fm)	0.692	0.660	0.641	0.627	0.618	0.585

Table 1
The CGC fits for different values of $\mathcal{T}_0$ and 3 quark flavors with mass $m_q = 140$ MeV. Also shown is the fit obtained by using the GBW model, Eq. (1.26).

The fit has been performed for the F_2 data at ZEUS (the first two references in [17]) with $x \leq 10^{-2}$ and Q^2 between 0.045 and 45 GeV^2 (156 data points). The values obtained for the three free parameters and the χ^2/d.o.f. for the fits corresponding to different choices for $\mathcal{T}_0$ are shown in Table 1. (The photon wavefunctions in Eqs. (5.2)–(5.3) have been evaluated with three active quark flavors with equal masses $m_q = 140$ MeV.) Note the good quality of the fit (χ^2/d.o.f. $\simeq 0.8 \div 0.9$), and also the stability of the result for λ, which changes only by 15% (within the range $\lambda = 0.25 \div 0.29$) when varying $\mathcal{T}_0$ (and also when varying the quark masses [54]). As anticipated, this value of λ is in agreement with the theoretical calculation in Ref. [53]. In Table 1, we also show, for comparison, the results obtained when performing a fit to the same set of data with the GBW cross–section (1.26). Clearly, the corresponding χ^2/d.o.f. = 1.59 is relatively poor. Note also that the value of $Q_s^2(x)$ emerging from the CGC fit is smaller — roughly, by a factor of 2 for $\mathcal{T}_0 = 0.7$ — than the corresponding value for the GBW fit (since x_0 is smaller).

In Figs. 40 and 41, the results of the fit are plotted against the HEAR data. Fig. 41 shows also the extrapolation of the fit towards larger values of x and Q^2 (outside the range of the fit), together with the corresponding HERA data, in order to emphasize that the deviations from BFKL dynamics — due notably to the presence of the valence quarks and to the DGLAP evolution [15] — do eventually show up, as expected.

In both figures, it also shown (with dashed line) the prediction of the BFKL calculation without saturation, as obtained by extrapolating the formula in the first line of Eq. (5.13) to arbitrarily large rQ_s (with the same values for the parameters as in Table 1, and the BFKL diffusion term switched off for $rQ_s > 2$). This pure BFKL fit shows a too strong increase with $1/x$ at small Q^2, as expected from similar analyses in the literature [196].

On the other hand, the complete fit, including saturation, works remarkably well down to the lowest values of Q^2 that we have included. Since, a priori, such low values of Q^2
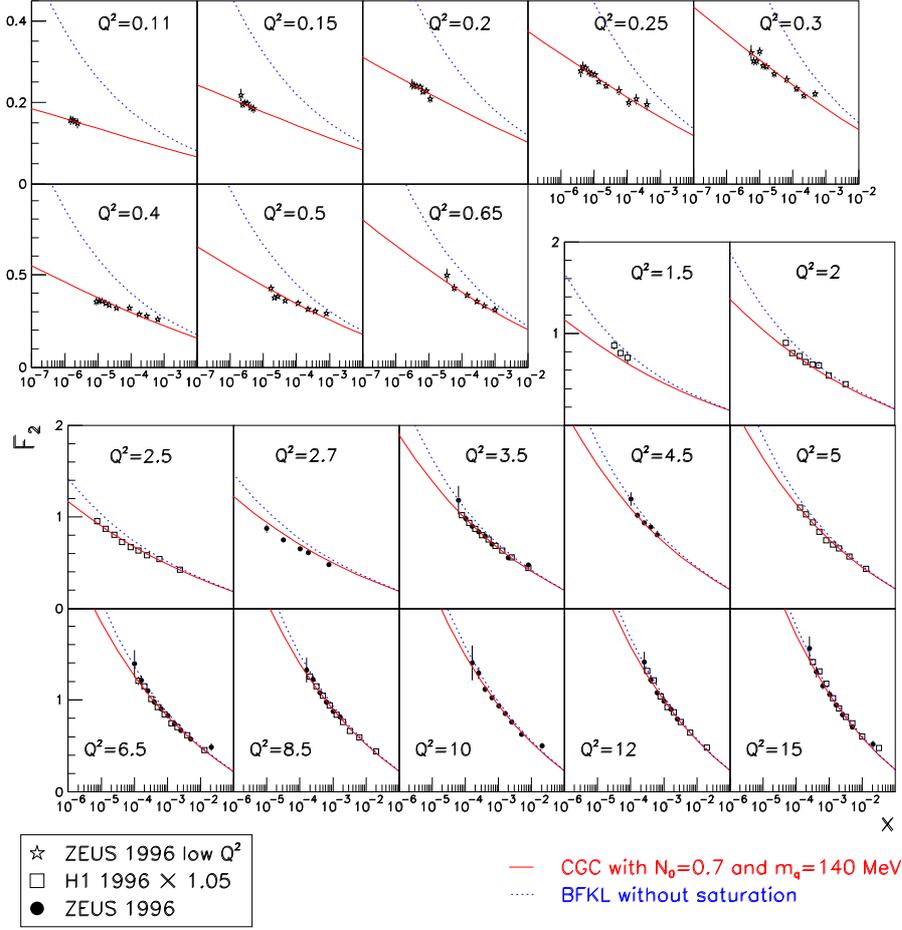
Fig. 40. The F_2 structure function in bins of Q^2 for small (upper part) and moderate (lower part) values of Q^2. The experimental points are taken from Refs. [17]. (The H1 data have been rescaled by a factor 1.05 which is within the normalization uncertainty.) The few data points at the lowest available Q^2 (0.045, 0.065 and 0.085 GeV^2) are not displayed although they are included in the fit. The full line is the result of the CGC fit, whereas the dashed line shows the predictions of the pure BFKL part of the fit (no saturation). (Figure taken from Ref. [54].) | F2low |

seem to be completely out of the reach of perturbation theory, it is important to explain why it is nevertheless meaningful to approach these data within a saturation fit:

The reason is that, for the low Q^2 data at HERA, the associated values of x are also very small, so the corresponding saturation momentum is relatively large (much larger than Q^2) and marginally perturbative : $Q_s^2(x) \sim 1.3 \div 2.3 \, \mathrm{GeV}^2$ for $x \sim 10^{-5} \div 10^{-6}$. From the discussion prior to Eq. (5.11) we know that, when $Q^2 \ll Q_s^2$, the strength of the convolution yielding the DIS cross–section, Eq. (5.1), is pushed towards relatively small dipole sizes, $r \sim 1/Q_s$. Thus, a large fraction of the dipole configurations relevant at low Q^2 is actually perturbative. These perturbative configurations are responsible for the dominant logarithmic behaviour exhibited in the third line of Eq. (5.11). In turn, this behaviour naturally explains the turnover at low Q^2 observed in the HERA data for F_2
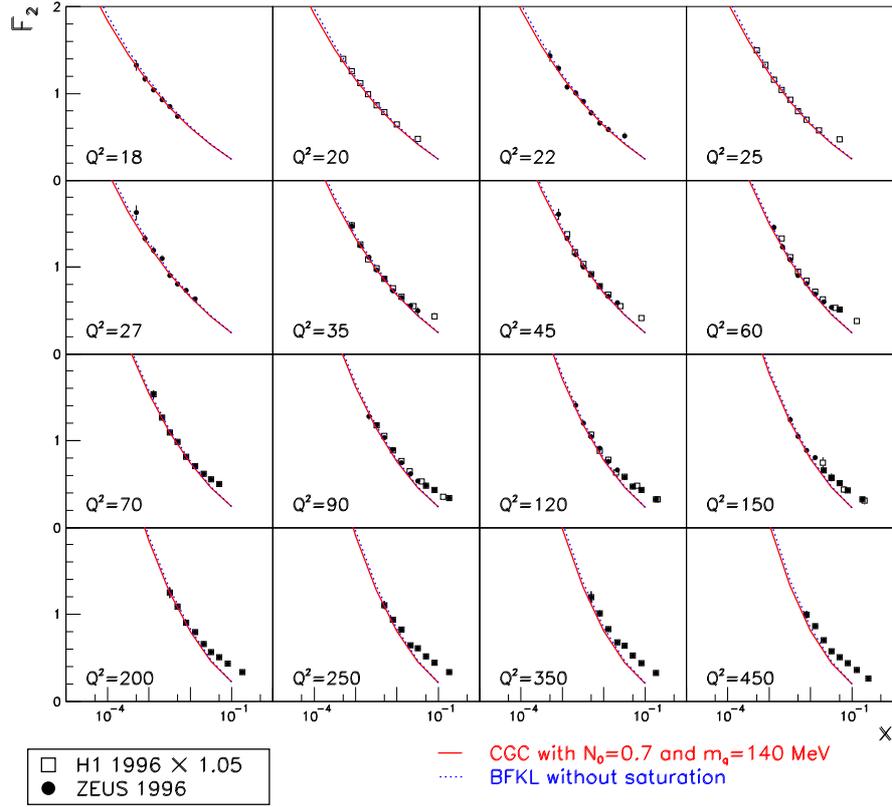
Fig. 41. The same as in Fig. 40, but for large Q^2. Note that in the bins with $Q^2 \geq 60\,\text{GeV}^2$, the CGC fit is extrapolated outside the range of the fit ($Q^2 \leq 50\,\text{GeV}^2$ and $x \leq 10^{-2}$), to better emphasize its limitations. (Figure taken from Ref. [54].) F2high

when these data are plotted as a function of Q^2 for small values of x (cf. Fig. 4).

Let us finally emphasize that, within the kinematical range considered in the fit, the HERA data show clear evidence for geometric scaling *violation*, and the success of the CGC fit relies crucially on its ability to describe this violation, via the BFKL diffusion. To clearly demonstrate this point, the authors of Ref. [54] have also tried a different fit, from which the diffusion term has been excluded. That is, the dipole amplitude has been assumed to be a *pure scaling function* : $T(x, r) = \mathcal{T}(rQ_s)$, with

$$
\mathcal{T}(rQ_s) = \mathcal{T}_0 \left(\frac{rQ_s}{2} \right)^{2\gamma} \qquad \text{for} \quad rQ_s \leq 2,
$$
$$
\mathcal{T}(rQ_s) = 1 - e^{-a \ln^2(b\, rQ_s)} \qquad \text{for} \quad rQ_s > 2, \tag{5.14}
$$

where the exponent γ is now treated as a *free parameter*. But in spite of involving one more parameter, the best 'scaling fit' has a poor quality: $\chi^2/\text{d.o.f.} \approx 1.4$. Besides, the value of γ which emerges from this fit is $\gamma \approx 0.84$, which has no fundamental meaning:

this value is significantly smaller than one (at variance with the GBW "saturation model", and also with the DGLAP evolution), but also significantly larger than the BFKL value $\gamma_s = 0.63$. This clearly shows that the data 'prefer' an *effective* 'anomalous dimension', which varies with x and Q^2 in such a way to violate geometric scaling. Within the CGC fit based on Eq. (5.13) this effective 'anomalous dimension' can be estimated as

$$\gamma_{\mathrm{eff}}(rQ_s, \tau) \equiv -\frac{\mathrm{d} \ln \mathcal{T}(rQ_s, \tau)}{\mathrm{d} \ln(4/r^2 Q_s^2)} = \gamma_s + 2 \frac{\ln(2/rQ_s)}{\kappa \lambda \tau} \quad \text{for} \quad rQ_s \leq 2, \tag{5.15}$$

and is larger than γ_s. The difference $\gamma_{\mathrm{eff}} - \gamma_s$ decreases with τ, but increases with the deviation $\rho - \rho_s = 2 \ln(2/rQ_s)$ from the saturation line. For $rQ_s \ll 1$, γ_{eff} is close to one and thus mimics the DGLAP behaviour; this may explain why the CGC fit works reasonably well up to relatively large values of $Q^2 \sim 50 \, \mathrm{GeV}^2$, where the DGLAP dynamics is expected to be important.

The CGC parametrization (5.13) for the dipole cross–section (with the values of the parameters fixed as above) has been subsequently used for various phenomenological applications, both to the HERA physics [194, 195, 197, 198] and to the RHIC physics [199–201], in particular to the study of particle production in d+Au collisions at forward rapidities, a topic that we shall discuss in a model–independent way in the next subsection.

5.2   Cronin effect and high–$p_\perp$ suppression from the CGC

In our discussion of the experimental situation at RHIC, in Sect. 1.3, we have emphasized that the '$R_{\mathrm{pA}}(\eta, p_\perp)$–ratio' defined in Eq. (1.38) is a good indicator of collective nuclear effects like gluon saturation. (In the experiment at RHIC, one is actually performing deuteron–gold collisions, but these are similar to the pA collisions that we shall consider here, for simplicity.) From Sect. 1.3, we recall that η and $p_\perp$ denote respectively the pseudo–rapidity and the transverse momentum of the particles (hadrons) produced in the collision, with positive values for η corresponding to the proton fragmentation region. At high energy, the particles are predominantly produced via gluon–gluon scattering, so by measuring the particle yield one has access to the gluon distributions in the incoming hadrons (the proton and the nucleus). In particular, when increasing η, one is probing gluons with smaller and smaller values of x in the nuclear wavefunction, which is interesting as it allows us to follow the evolution of the nuclear gluon distribution in the presence of high density effects. Indeed, in the experimental conditions at RHIC, the nuclear gluon density is quite large already for the intermediate values of $x \sim 10^{-1} \div 10^{-2}$ corresponding to central rapidity ($\eta = 0$), and this density can only increase when decreasing x ($=$ increasing η). But for the same values of η, the pp collisions which serve as a level of reference in Eq. (1.38) are not sensitive to saturation effects, but only to the standard, 'perturbative' (in the sense of *linear*), evolution in QCD. Therefore, by dividing out the particle yield in pA collisions by the corresponding one in pp collisions multiplied by A, as in Eq. (1.38), one should be able to single out the high density effects in the evolution.

As also discussed in Sect. 1.3, the experimental results for this ratio at RHIC [57, 202–204] show a very interesting structure, which evolves quite fast with increasing η, from *enhancement* at central rapidity ($\eta = 0$) — where the data exhibit a *Cronin peak* ($R_{\mathrm{pA}} > 1$)

at intermediate momenta (cf. Fig. 16) — to *suppression* at forward rapidities ($1 < \eta < 4$), where the ratio is significantly smaller than one at all the measured momenta, and exhibits an 'abnormal' centrality dependence (cf. Figs. 17 and 18). It is then natural to ask: can one understand this complex behaviour within the scenario of saturation ?

The answer turns out to be 'yes', as substantiated by an abundant literature [86, 87, 114, 115, 205–211]. And as a matter of facts, the pattern of the evolution of the $R_{\rm pA}$–ratio with increasing η has been theoretically predicted [86] before the advent of the first respective data at RHIC. In this subsection, we shall give a brief theoretical discussion of this phenomenon within the CGC formalism, following the very detailed analysis in Ref. [87]. Although only qualitative, the latter has the virtue to provide a conceptual explanation of the main features seen in the RHIC data from first principles, without additional assumptions or free parameters. More quantitative analyses, which introduce additional parameters in order to actually fit the data, can be found in Refs. [208, 210, 211].

The first question is, how to relate the measured ratio $R_{\rm pA}$ to the 'theoretical' gluon distributions inside the hadrons participating in the collision ? Or, equivalently, how to compute gluon and quark production[23] in pA or AA collisions ? For the case of pp collisions at not too high energies (namely, so long as the saturation effects remain negligible in both the incoming protons), it was since long known that "$k_\perp$–factorization" applies: the spectrum of the partons produced in the collision can be computed in terms of the (unintegrated) gluon distributions in the target and the projectile together with the elementary cross–section for gluon–gluon scattering (see, e.g., [123]). More recent studies [30, 91, 207, 212, 213] have demonstrated that the $k_\perp$–factorization extends to *gluon* production in pA collisions as well, provided a generalized 'unintegrated gluon distribution' is used for the nucleus: this is truly a Fourier transform of the dipole–nucleus cross–section, and includes the effects of the multiple scattering between the produced gluons and the strong color fields in the nuclear wavefunction. No similar factorization is expected for *quark pair* production though [207, 214], nor for parton (quark or gluon) production in AA collisions (although $k_\perp$–factorization is sometimes used, for convenience, in phenomenological analyses [98, 215] of the Au+Au data at RHIC). So far, the only first–principle approach [216] to gluon production in AA collisions consists in solving the Yang–Mills equations describing the scattering between two classical color charges distributed according to the MV model (and hence representing the valence quarks of the two incoming nuclei). The lattice implementation of these equations has been constructed by Krasnitz and Venugopalan [116] and also by Lappi [118], and has been used for extensive numerical studies of gluon production at central rapidity [116–119, 205, 217]. There is an ongoing effort towards extending these calculations to quark pair production [218] and also to forward rapidities (by including quantum evolution in the distribution of the color charges).

Returning to the case of the ('dilute–dense') pA collisions, which is our main focus here, the respective gluon yield can be computed as [30, 91, 207, 213, 219]

---

[23] Note that we restrict ourselves here to *parton* (quark or gluon) production and gloss over the complicated problem of the 'parton fragmentation' into the final hadrons.

$$\frac{\mathrm{d}N_{\mathrm{pA}}}{\mathrm{d}\eta\,\mathrm{d}^2p_\perp\mathrm{d}^2b_\perp} = \frac{N_c^2-1}{4\pi^3}\,\frac{\bar{\alpha}}{p_\perp^2}\int\mathrm{d}^2\boldsymbol{k}\,\Phi_A(\boldsymbol{k},\boldsymbol{b},\tau_1)\,\varphi_p(\boldsymbol{p}-\boldsymbol{k},\tau_2)\,, \tag{5.16}$$

where $\tau_{1,2} = \tau_0\pm\eta$, with $\tau_0 \equiv \ln(\sqrt{s}/k_\perp)$ (see kinematics in Fig. 15). In the above equation, $\varphi_p(\boldsymbol{k},\tau)$ is the gluon occupation factor in the proton, cf. Eq. (2.33), here integrated over the proton area: $\varphi_p(\boldsymbol{k},\tau) = \int\mathrm{d}^2\boldsymbol{b}\,\varphi_p(\boldsymbol{k},\boldsymbol{b},\tau)$. Furthermore, the quantity $\Phi_A(\boldsymbol{k},\boldsymbol{b},\tau)$, which plays a formally similar role — that of the gluon occupation factor in the nucleus —, is really a *scattering amplitude*, namely [212]

$$\Phi_A(\boldsymbol{k},\boldsymbol{b},\tau) \equiv \frac{1}{\alpha N_c}\int\frac{\mathrm{d}^2\boldsymbol{r}}{4\pi}\,\mathrm{e}^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\,\nabla^2 T_{gg}(\boldsymbol{r},\boldsymbol{b},\tau)\,, \tag{5.17}$$

where $T_{gg}(\boldsymbol{r},\boldsymbol{b},\tau)$ is the scattering amplitude for a *gluonic* (gg) dipole, with size $\boldsymbol{r}$ and impact parameter $\boldsymbol{b}$, as computed in the eikonal approximation: $T_{gg} \equiv 1 - S_{gg}$, with S_{gg} defined as in Eq. (2.42). This 'dipole' is merely a mathematical construction — it arises in the calculation of the gluon production cross–section as the product between the gluon line in the direct amplitude (see Fig. 14) and the corresponding line in the complex conjugate amplitude — and its amplitude T_{gg} encodes the effects of multiple scattering, as anticipated.

The quantity (5.17) is not the same as the 'canonical' gluon occupation factor, Eq. (2.33), although it essentially reduces to the latter at high momenta $k_\perp \gg Q_s(A,\tau)$. For generic momenta, one rather has $\Phi_A \propto k^2\nabla_k^2\varphi_A$ (this follows, e.g., by comparing Eqs. (5.17) and (2.52)). In fact, there is *a priori* no reason why the canonical gluon distribution should enter the calculation of an observable at high energy: in the high–density environment characteristic for a high–energy target, the scattering operators are non–linear in the color fields; so, unlike what usually happens at low energy, they are not anymore proportional to the 2–point function which defines the gluon distribution.

From its definition (5.17), it is straightforward to evaluate the generalized 'gluon distribution' Φ_A in the kinematical regions of interest, via manipulations similar to those already encountered in this report (see, e.g., Ref. [114] for analytic estimates of Φ_A). However, it turns out that, in order to perform a *qualitative* study of the R_{pA}–ratio, one can simply replace $\Phi_A \to \varphi_A$ in Eq. (5.16) and also ignore the convolution with the 'spectator' proton — the one evolved up to the lower rapidity τ_2. That is, all the salient features of the ratio (1.38) are preserved when this is simply evaluated as[24]

$$\mathcal{R}_{pA}(\eta,k_\perp) \equiv \frac{\varphi_A(\tau_1,k_\perp)}{A^{1/3}\,\varphi_p(\tau_1,k_\perp)}\,, \tag{5.18}$$

where we recall that $\tau_1 \equiv \tau_0 + \eta$ is the rapidity of the participating gluon in the nuclear wavefunction. Eq. (5.18) measures the difference between the gluon distribution in the nucleus and that in the proton (scaled up by $A^{1/3}$) at the same value of x, and thus is the most direct expression of the nuclear effects which should be responsible for the deviation

---

[24] More quantitative analyses [114, 115] have confirmed that the simplified ratio, Eq. (5.18), and the original one, cf. Eqs. (1.38) and (5.16), have indeed a similar behaviour.

of the experimentally measured ratio R_{pA} from one. In what follows, we shall estimate Eq. (5.18), separately for the central rapidity and for the forward ones.

5.2.1 Central rapidity: Cronin peak in the MV model

At central pseudo–rapidity ($\eta = 0$) and for the kinematics at RHIC, $\tau_1 = \tau_0 \sim 3$ is rather small, so one can ignore quantum evolution towards small x in the nuclear wavefunction, and describe the corresponding gluon distribution as the result of classical radiation from the valence quarks. (From experience with the phenomenology at HERA, we expect small–x evolution effects to become important only $x < 10^{-2}$, i.e. $\tau > 5$; cf. Sect. 5.1).

Under these assumptions, the gluon distribution of a large nucleus ($A \gg 1$) can be described within the McLerran–Venugopalan model introduced in Sect. 1.3.3. As further explained in Sect. 2.5, this model encodes non–linear effects associated with the classical dynamics of strong color fields. When computed in the LC–gauge in which the gluon distribution makes sense, these non–linear effects are interpreted as *gluon saturation* at low transverse momenta $k_\perp \lesssim Q_s(A)$, cf. Eq. (2.56), and, respectively, as 'higher twist' corrections to the bremsstrahlung spectrum at large momenta $k_\perp \gg Q_s(A)$, cf. Eq. (2.55). The intermediate scale $Q_s(A)$ is the nuclear saturation momentum defined in Eq. (2.54); it scales with A like $Q_s^2(A) \sim A^{1/3} \ln A^{1/3}$. Alternatively, when viewed in the COV–gauge in which the classical field equations remain linear even for strong color fields, the non–linear effects encoded in Eq. (2.49) are interpreted as the *multiple scattering* of a gluonic dipole crossing the nucleus. This last interpretation, which is perhaps more suggestive for the problem at hand (gluon production in pA collisions), is illustrated in Fig. 42.
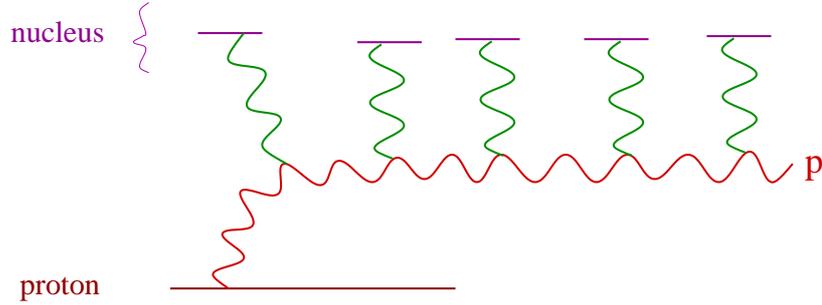

Fig. 42. Multiple scattering in a proton–nucleus collision. In the experiments at RHIC, the proton is actually replaced by a deuteron.

At the same (small) rapidities at which the large nucleus can be described within the MV model, the proton is in a dilute regime, so the associated gluon distribution $\varphi_p(k_\perp)$ can be computed in the 'leading–twist' approximation. To the accuracy of interest, this is simply the bremsstrahlung spectrum (compare to Eqs. (2.53) and (2.55)) :

$$\varphi_p(k_\perp) \simeq \frac{1}{\alpha N_c}\frac{Q_p^2}{k_\perp^2}, \qquad Q_p^2 \approx \frac{2\alpha^2 N_c}{R_p^2}, \tag{5.19}$$

with $Q_p \sim \Lambda_{\mathrm{QCD}}$. Eq. (5.19) is valid in the perturbative regime at $k_\perp \gg Q_p$.

We are now in a position to compute the $\mathcal{R}_{pA}$–ratio, Eq. (5.18). In view of Eq. (5.19), it is clear that the quantity $A^{1/3}\varphi_p(k_\perp)$ which appears in the denominator of Eq. (5.18)
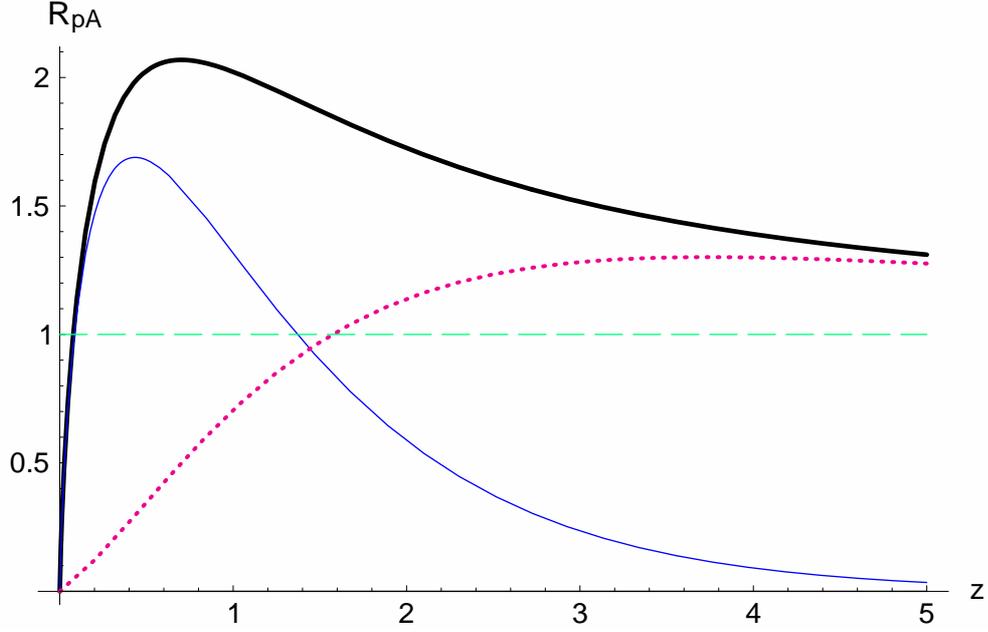
Fig. 43. The ratio $\mathcal{R}_{pA}(z)$ as a function of the scaled momentum variable $z = k^2/Q_s^2(A)$ in the McLerran-Venugopalan model with $\rho_A = 6$. The thick (black) line is the complete result; the solid (blue) line shows the contribution $\mathcal{R}_{pA}^{\mathrm{sat}}(z)$ of the saturation plateau alone; the dotted (magenta) line shows the difference $\mathcal{R}_{pA}^{\mathrm{twist}}(z) \equiv \mathcal{R}_{pA} - \mathcal{R}_{pA}^{\mathrm{sat}}$ due to higher–twist effects at $z \ll 1$. (From Ref. [87].)

coincides with the nuclear gluon distribution in the high momentum limit — i.e., the dominant term at high–$k_\perp$ in the r.h.s. of Eq. (2.55). Thus, within the MV model at least, the ratio $\mathcal{R}_{pA}(k_\perp)$ is also a measure of the deviation of the actual nuclear gluon distribution from the corresponding prediction of the 'leading–twist' (or linear) perturbation theory. As explained in Sect. 2.5, this deviation is associated with non–linear effects which 'push' the gluons towards the modes with higher transverse momenta (of the order of the saturation scale, or larger). Accordingly, when compared to the bremsstrahlung spectrum, the MV spectrum shows a strong depletion at very low momenta $k_\perp \ll Q_s(A)$, but a relative excess for larger momenta $k_\perp \gtrsim Q_s(A)$ (see Fig. 27). This immediately implies that the ratio (5.18) should be smaller than one for $k_\perp \ll Q_s(A)$ and larger than one for $k_\perp \gtrsim Q_s(A)$, and moreover it must approach one from the above at sufficiently high momenta :

$$\mathcal{R}_{pA}(k_\perp) \searrow 1 \qquad \text{when} \qquad k_\perp \to \infty. \tag{5.20}$$

This last property follows since, at high $k_\perp$, the nuclear spectrum is amended by 'higher twist' corrections which are positive, cf. Eq. (2.55). From these considerations, it becomes obvious that the function $\mathcal{R}_{pA}(k_\perp)$ must have a *maximum* at some intermediate value of $k_\perp$, of the order of $Q_s(A)$. This maximum is the *Cronin peak*.

To study the emergence and the properties of this maximum in more detail, let us consider increasing $k_\perp$ from very small values, $k_\perp \sim Q_p$. (See also Fig. 43.) So long as $k_\perp \lesssim Q_s(A)$, the nuclear distribution $\varphi_A(k_\perp)$ remains nearly constant, cf. Eq. (2.56), whereas the proton distribution $\varphi_p(k_\perp)$ decreases very fast, like $1/k_\perp^2$. Accordingly, the ratio $\mathcal{R}_{pA}(k_\perp)$ increases

rapidly, $\mathcal{R}_{pA}(k_\perp) \propto k_\perp^2$, from a value $\mathcal{R}_{pA} \sim A^{-1/3} \ll 1$ when $k_\perp \sim Q_p$ up to a value

$$\mathcal{R}_{pA}(k_\perp) \sim \rho_A > 1 \qquad \text{for} \qquad k_\perp \sim Q_s(A). \tag{5.21}$$

We have defined here (cf. Eq. (2.54)) :

$$\rho_A \equiv \frac{Q_s^2(A)}{Q_A^2} = \ln \frac{Q_s^2(A)}{\Lambda_{\text{QCD}}^2} \sim \ln A^{1/3}, \tag{5.22}$$

and we have used the fact that $\varphi_A \sim 1/\alpha N_c$ for $k_\perp \sim Q_s(A)$, together with Eq. (5.19) and the estimate $Q_A^2 \approx A^{1/3} Q_p^2$, cf. Eq. (2.53).

But when $k_\perp$ is further increased above $Q_s(A)$, the nuclear saturation plateau ends up quite abruptly (cf. Fig. 27), and the gluon occupation factor $\varphi_A(k_\perp)$ decreases very fast — exponentially in $z \equiv k_\perp^2/Q_s^2(A)$ — down to the power law tail given by the 'twist expansion' in Eq. (2.55). Correspondingly, the ratio $\mathcal{R}_{pA}(k_\perp)$ decreases quite fast for $k_\perp > Q_s(A)$, down to a value of order one (but strictly larger than one).

Thus, the maximum is located at $k_\perp \sim Q_s(A)$, as anticipated, and has a height $\mathcal{R}_{\max}(A) \sim \rho_A$ which is parametrically enhanced for large A. Still for large A, this maximum is rather well peaked: the width/height ratio at the peak can be estimated as $(\ln \rho_A/\rho_A) \ll 1$. Note that, in these estimates, we have implicitly treated ρ_A (and not only A) like a large number. To make contact with RHIC phenomenology, it is therefore useful to notice that, for a gold nucleus at RHIC energies, one expects $Q_s^2(A) \simeq 2$ GeV^2 [55], which together with $\Lambda_{\text{QCD}} \simeq 200$ MeV implies $\rho_A \simeq \ln 50 \simeq 4$. This number is reasonably large.

The qualitative analysis above is confirmed by an exact (analytic) calculation [87] of the nuclear gluon spectrum in the MV model, Eq. (2.52), which together with Eq. (5.19) for the corresponding spectrum in the proton leads to the ratio $\mathcal{R}_{pA}$ displayed in Fig. 43. In particular, the location z_0 of the maximum and its height $\mathcal{R}_{\max}(A) \equiv \mathcal{R}_{pA}(z_0)$ have been computed in an expansion in powers of $1/\rho_A$ [87], with the following results:

$$z_0 = 0.435 + \frac{0.882}{\rho_A} + \mathcal{O}\left(\rho_A^{-2}\right), \qquad \mathcal{R}_{\max}(A) = 0.281\,\rho_A + 0.300 + \mathcal{O}\left(\rho_A^{-1}\right). \tag{5.23}$$

To conclude, the experimental observation of a rather well–defined Cronin peak in d+Au collisions at central rapidity at RHIC (cf. Fig. 16) may be naturally interpreted as an evidence of the non–linear effects responsible for gluon saturation in the nuclear wavefunction. A closely related effect is that the spectrum of the produced gluons gets 'harder' as a result of their rescattering inside the nucleus (cf. Fig. 42). To see this, let us return to Eq. (5.16) for the gluon yield and notice that, in the absence of saturation effects in the nucleus (i.e., with $\Phi_A(k_\perp) \propto 1/k_\perp^2$), the integral there would yield

$$\frac{dN_{\text{pA}}}{d\eta d^2 p_\perp d^2 b_\perp} \propto \frac{1}{p_\perp^4} \qquad \text{when} \qquad p_\perp \gg Q_s(A). \tag{5.24}$$

However, as emphasized in writing the r.h.s. of Eq. (5.24), the 'leading–twist' spectrum $\propto 1/p_\perp^4$ holds only for very high momenta, well above the nuclear saturation scale. But it is easily to check that, for lower momenta $p_\perp \lesssim Q_s(A)$, the spectrum becomes less infrared divergent, because of gluon saturation inside the nucleus:

$$\frac{\mathrm{d}N_{\mathrm{pA}}}{\mathrm{d}\eta \mathrm{d}^2 p_\perp \mathrm{d}^2 b_\perp} \propto \frac{1}{p_\perp^2} \qquad \text{when} \qquad p_\perp \lesssim Q_s(A) \,. \tag{5.25}$$

5.2.2 Forward rapidities: High–$p_\perp$ suppression

From the experimental results at RHIC, we know that the physical picture changes dramatically when moving from central to forward rapidities. As we shall explain in what follows, such a dramatic change can be understood as the result of the dissymmetry between the small–x evolutions of the nuclear and, respectively, proton gluon distributions which form the ratio (5.18).

At forward, and sufficiently large, pseudo–rapidities ($\eta = 2 - 3$), the gluons with $\tau_1 = \tau_0 + \eta$ in either the nucleus or the proton [25] wavefunctions are predominantly produced by quantum evolution, that is, they are emitted by gluons with larger values of x which are themselves radiated by the valence quarks. In the kinematical range relevant for d+Au collisions at RHIC, one can assume the proton to be always in a dilute regime, so its evolution is linear. But the nucleus develops a saturation region at sufficiently low transverse momenta (since it already did so in the initial condition at $\tau_1 = \tau_0$), hence its evolution must include the non–linear effects discussed in Sect. 3. It is the difference between the *linear* evolution of the proton and that *non–linear* of the nucleus which leads to the rapid suppression of the $\mathcal{R}_{pA}$–ratio with increasing η, as we shall shortly see.

To be consistent with our previous discussion of the $\eta = 0$ case (cf. Sect. 5.2.1), where we have assumed that the effects of the evolution are negligible for $\tau_1 = \tau_0$, we need to consider that the rapidity interval which is really available for the small–x evolution is not τ_1, but rather $\tau_1 - \tau_0 = \eta$. To keep in line with the previous notations in this report, where this rapidity interval has been systematically denoted as τ, we shall often rewrite $\eta \equiv \tau$. That is, in the remaining part of this section, τ and η will be just different names for the same physical quantity — the rapidity of the particles produced in pA collisions.

The gluon occupation factor predicted by the CGC effective theory has been either explicitly computed in Sect. 4 (in some kinematical regions), or it can be easily inferred from the results presented there (in the other regions). For convenience, we summarize here the relevant results:

(1) The gluon distribution at saturation has been computed in Sect. 4.1 with the following result (cf. Eq. (4.8), which is here extended to a nucleus) :

$$\varphi_A(k_\perp, \tau) \approx \frac{1}{\alpha N_c} \left\{ \ln \frac{Q_s^2(A, \tau)}{k_\perp^2} + \mathcal{O}(1) \right\}, \qquad \text{for} \qquad k_\perp^2 \ll Q_s^2(A, \tau) \,. \tag{5.26}$$

Here, $Q_s(A, \tau)$ is the nuclear saturation momentum, which has been discussed in Sect. 4.2.6. The LO approximation with fixed coupling predicts that $Q_s(A, \tau)$ rises very fast with both τ and A (cf. Eq. (4.53)) : $Q_s^2(A, \tau) \simeq Q_s^2(A) \, e^{c\bar{\alpha}\tau}$ with $c \simeq 4.88$, $\bar{\alpha} \equiv \alpha N_c/\pi$, and the initial condition $Q_s^2(A) \propto A^{1/3}$ provided by the MV model, cf. Eq. (2.54). But this growth is slowed down by running coupling effects, which become more and more

---

[25] Note that the 'proton' here is that one which appears in the denominator of the ratio $\mathcal{R}_{pA}$– ratio, Eq. (5.18), and which has the same rapidity τ as the nucleus.

important with increasing τ (cf. Eq. (4.52)) and ultimately wash out any A–dependence in the saturation scale [172] (see Eq. (4.55)).

(2) For transverse momenta well above the saturation scale, the occupation factor is given by the solution to the BFKL equation with the saturation boundary condition[26] (??). This can be solved via the same manipulations described in Sect. 4.1 for the case of the dipole scattering amplitude, with similar results. Specifically, for momenta within the 'extended scaling region' (5.27), which for a nucleus reads

$$Q_s^2(A, \tau) \ll k_\perp^2 \ll Q_g^2(A, \tau), \qquad \text{with} \qquad Q_g^2(A, \tau) \equiv Q_s^4(A, \tau)/Q_s^2(A), \qquad (5.27)$$

the BFKL solution is well approximated by the scaling form (compare with Eq. (4.58))

$$\varphi_A(k_\perp, \tau) \approx \frac{1}{\alpha N_c} \left\{ \ln \frac{k_\perp^2}{Q_s^2(A, \tau)} + \mathcal{O}(1) \right\} \left(\frac{Q_s^2(A, \tau)}{k_\perp^2} \right)^\gamma, \qquad \gamma \approx 0.63. \qquad (5.28)$$

Scaling violations in the form of the diffusion term can be easily added (cf. Eq. (4.56)), but they are omitted here for simplicity.

(3) For even higher momenta $k_\perp^2 \gg Q_g^2(A, \tau)$, the occupation factor must be computed by resumming small–x *and* collinear emissions. To the present accuracy, the corresponding result is given by the "double–logarithmic approximation" of Eq. (4.41), that is

$$\varphi_A(k_\perp, y) \simeq \frac{1}{\alpha N_c} \frac{Q_A^2}{k_\perp^2} \exp \left\{ \sqrt{4\bar{\alpha} y \rho(A, k_\perp)} \right\}, \qquad (5.29)$$

where $\rho(A, k_\perp) \equiv \ln(k_\perp^2/Q_s^2(A))$ is recognized as the phase–space available for the (linear) $k_\perp^2$–evolution:

$$\rho(A, k_\perp) = \int_{Q_s^2(A)}^{k_\perp^2} \frac{\mathrm{d}k_\perp^2}{k_\perp^2} = \ln \frac{k_\perp^2}{Q_s^2(A)}. \qquad (5.30)$$

Notice that, in agreement with the discussion in Sect. 4.2.1, it is the infrared cutoff in the initial conditions at low energy — here, the nuclear saturation momentum $Q_s(A)$ — which sets the lower cutoff for the transverse phase–space available for the evolution. For the proton, this scale would be Q_p, with $Q_p \ll Q_s(A)$. Hence, $\rho(p, k_\perp) > \rho(A, k_\perp)$ for any $k_\perp$, which explains why a pronounced difference persists between the gluon spectrum in the nucleus and that in the proton even at very high momenta $k_\perp \gg Q_s(A, \tau)$.

We are now prepared for a study of the evolution of the ratio $\mathcal{R}_{pA}(k_\perp, \tau)$ with increasing τ, starting with the initial condition provided by the MV model. The most distinguished features of this evolution are summarized below (we refer to [87] for more details) :

i) The main effect of the evolution is a *rapid suppression* of the ratio $\mathcal{R}_{pA}$ with increasing τ at any value of $k_\perp$ (except for the asymptotic ones), which is due to the *different evolution rates* of the gluon distributions in the *nucleus* (the numerator in Eq. (5.18))

---

[26] This is tantamount to solving the BK equation in momentum space. Indeed, the gluon occupation factor can be alternatively computed as a specific Fourier transform of the dipole scattering amplitude [51, 87, 124].

and respectively the *proton* (the denominator there). Roughly speaking, the proton distribution grows faster because, for a given $k_\perp$, the transverse phase–space available for the proton evolution is larger than that for the nucleus (cf. Eq. (5.30)) :

$$\rho(p, k_\perp) \equiv \ln \frac{k_\perp^2}{Q_p^2} = \ln \frac{k_\perp^2}{Q_s^2(A)} + \ln \frac{Q_s^2(A)}{Q_p^2} \simeq \rho(A, k_\perp) + \rho_A, \qquad (5.31)$$

where $\rho_A \gg 1$ has been generated according to Eq. (5.22). Correspondingly, $\mathcal{R}_{pA}(k_\perp, \tau)$ decreases very fast with τ, and already after a short evolution $\tau_c \sim 1$ it becomes smaller than one at all but the asymptotic momenta (τ_c will be estimated below).

In fact, the *suppression rate* $\mathrm{d}\ln \mathcal{R}_{pA}/\mathrm{d}\tau$ is largest at small τ and for not so large transverse momenta — namely, for $k_\perp \lesssim Q_g(A, \tau)$ —, since in this regime the dissymmetry between the evolution of the proton and that of the nucleus is most pronounced: The proton is in the DLA regime, and thus evolves very fast, whereas the nucleus is either at saturation, or in the geometric scaling window, and thus it evolves quite slowly (because, so long as $\bar\alpha\tau < 1$, the nuclear saturation momentum rises very slowly with τ; see Eq. (4.53)). This explains, in particular, the rapid suppression in $\mathcal{R}_{pA}$ observed in the early stages of the evolution in the numerical study in Ref. [115], which is based on the BK equation. The results obtained there are shown in Fig. 44.

To better appreciate how fast is the evolution and estimate τ_c, let us study the evolution of the ratio $\mathcal{R}_{pA}(k_\perp, \tau)$ for momenta $k_\perp \sim Q_s(A, \tau)$; that is, τ and $k_\perp$ are simultaneously increased, in such a way to remain in the vicinity of the nuclear saturation line, where the Cronin peak is expected to be located. We thus consider the following function of τ :

$$\mathcal{R}_{\mathrm{sat}}(A, \tau) \equiv \mathcal{R}_{pA}(k_\perp = Q_s(A, \tau), \tau), \qquad (5.32)$$

which measures the height of the Cronin peak, so long as the latter exists. (For $\tau = 0$, $\mathcal{R}_{\mathrm{sat}} \sim \rho_A$, cf. Eq. (5.21).) By definition, τ_c is the rapidity evolution after which $\mathcal{R}_{\mathrm{sat}}(A, \tau)$ decreases from its initial value of $\mathcal{O}(\rho_A)$ to a value of $\mathcal{O}(1)$.

For $k_\perp \sim Q_s(A, \tau)$, the gluon occupation factor in the nucleus is independent of τ and of order $1/\alpha N_c$, while that in the proton is given by the DLA formula (5.29) (the DGLAP evolution of the proton prevails over the BFKL one so long as $\bar\alpha\tau \ll \rho_A$). We then have

$$\mathcal{R}_{\mathrm{sat}}(A, \tau) \sim \rho_A \exp\left\{c\bar\alpha\tau - \sqrt{4\bar\alpha\tau\rho_A}\right\}. \qquad (5.33)$$

For $\tau = 0$, this is of $\mathcal{O}(\rho_A)$, as expected. But when increasing τ, the ratio decreases very fast — the DGLAP increase of the proton distribution being faster than the BFKL increase of the nuclear saturation momentum, cf. Eq. (4.36) —, and becomes parametrically of $\mathcal{O}(1)$ already after the very short rapidity evolution

$$\bar\alpha\tau_c \sim \frac{\ln^2 \rho_A}{4\rho_A} \ll 1. \qquad (5.34)$$

This value τ_c is in fact so small that one can in fact ignore the corresponding evolution of the nucleus: *The rapid decrease in the height of the peak in the very early stages of the evolution is entirely due to the DGLAP evolution of the proton.*
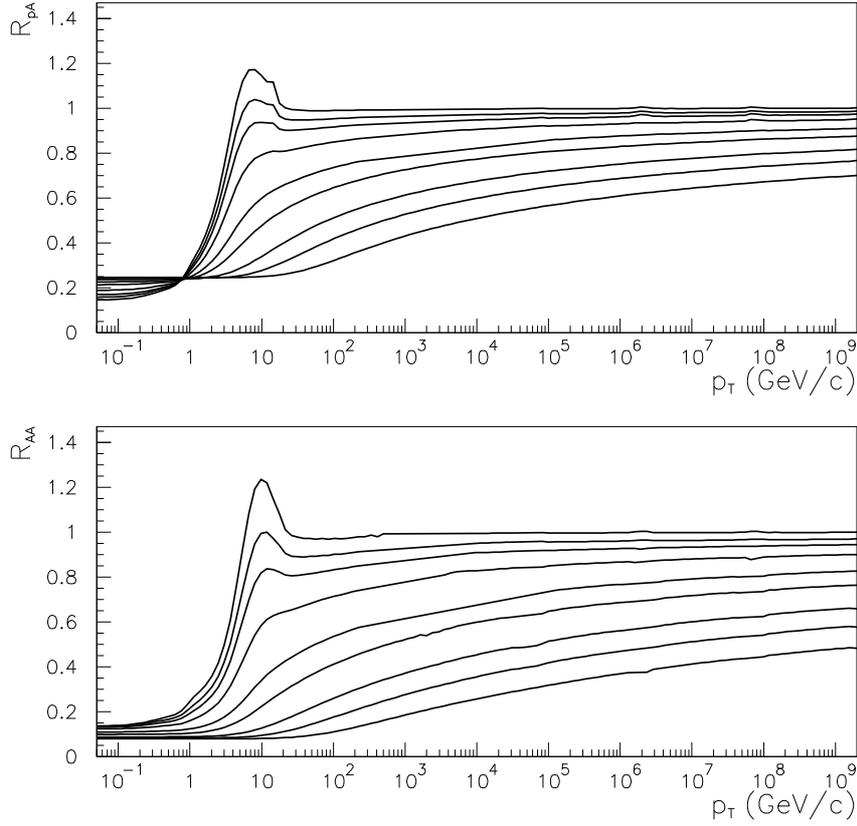
Fig. 44. Ratios R_{pA} and R_{AA} of gluon yields in p–A (upper plot) and A–A (lower plot) for BK evolution, with MV as initial condition with $Q_s^2 = 0.1$ GeV^2 for p and 2 GeV^2 for A. Lines from top to bottom correspond to τ=0, 0.25, 0.5, 1, 2, 3, 5, 7 and 10. From Ref. [115].

Note that, parametrically, $\tau_c \sim 1/(\bar{\alpha}\rho_A)$ is indeed of order one, as anticipated, since the coupling $\bar{\alpha}$ is to be evaluated at the nuclear saturation scale, which implies $\bar{\alpha}(Q_s^2(A)) = b/\rho_A$ with $b = \mathcal{O}(1)$.

Insert the double-scaling regime.

ii) By the same argument, the suppression goes away at extremely large momenta, where the difference between $Q_s(A)$ and Q_p becomes unimportant in computing the phase–space. In fact, when $k_\perp \gg Q_s(A, \tau)$, one can use the DLA formula (5.29) for both the proton and the nucleus, and thus deduce (for fixed coupling) :

$$\mathcal{R}_{pA}(k_\perp, \tau) \simeq \mathrm{e}^{-\rho_A \sqrt{\bar{\alpha}\tau/\rho(A, k_\perp)}} \qquad \text{for} \qquad k_\perp \gg Q_s(A, \tau), \tag{5.35}$$

which approaches one *from below* when $k_\perp \to \infty$.

iii) For larger τ such that $\bar{\alpha}\tau \gtrsim 1$, the ratio $\mathcal{R}_{pA}(k_\perp, \tau)$ is *monotonously increasing* with $k_\perp$. That is, the Cronin peak has flattened out during the first $1/\bar{\alpha}$ units of rapidity.

The *flattening* of the Cronin peak cannot be attributed to the proton evolution alone

161

(the latter produces a quasi–uniform suppression in $\mathcal{R}_{pA}$ at momenta around $Q_s(A, \tau)$, so, by itself, it would preserve a local structure like a peak), rather this must be related to the evolution of the nucleus. As demonstrated in Ref. [87], it is indeed the *nuclear evolution* which washes out that distinguished feature of the initial distribution which was responsible for the existence of a well–pronounced peak at $\tau = 0$: the exponential fall off of the gluon occupation factor at momenta just above the saturation plateau.

iv) Whereas the *generic* features of the evolution, as described above, are qualitatively similar for both fixed and running coupling, important differences persist between these two scenarios as far as the *details* of the evolution, and also the precise structure of the final results, are concerned. Specifically, after including running coupling effects, the evolution appears to be *slower* (one needs a larger increase in rapidity to achieve a given suppression in $\mathcal{R}_{pA}$), but eventually *stronger* (the final value for $\mathcal{R}_{pA}$ which is obtained after a very large evolution in τ is significantly smaller with running coupling than with fixed coupling). Let us be more specific on these two points:

• To appreciate how *fast* is the evolution, one can use the rapidity τ_c introduced above. For fixed coupling, this is given by Eq. (5.34), whereas for running coupling one finds [87]

$$\tau_c \simeq \frac{1}{4b} \ln \rho_A \sim \ln A^{1/3} \qquad \text{(running coupling)}. \tag{5.36}$$

For large A, the above estimate is parametrically larger than the previous one in Eq. (5.34); thus, the running of the coupling slows down the evolution indeed.

• Furthermore, to characterize the *strength* of the suppression after a very large rapidity evolution, consider the limit of $\mathcal{R}_{pA}$ when $\tau \to \infty$ with fixed $z \equiv k_\perp^2/Q_s^2(A, \tau)$. For $z = \mathcal{O}(1)$, one finds [87]:

$$\mathcal{R}_{pA}(z \sim 1, \tau \to \infty) \sim \frac{1}{(A^{1/3}\rho_A)^{1-\gamma}} \qquad \text{(fixed coupling)}, \tag{5.37}$$

and, respectively,

$$\mathcal{R}_{pA}(z \sim 1, \tau \to \infty) = \frac{1}{A^{1/3}} \qquad \text{(running coupling)}. \tag{5.38}$$

As anticipated, for large A, the running coupling result (5.38) is much smaller than the corresponding one for fixed coupling, Eq. (5.37) (recall that $1 - \gamma \simeq 0.37$).

In fact, the power of $A^{1/3}$ in the r.h.s. of Eq. (5.38) is simply the factor introduced by hand in the definition (5.18) of $\mathcal{R}_{pA}$. That is, the result (5.38) arises directly from the observation that, with a running coupling and for sufficiently large τ, the nuclear and proton saturation scales become equal with each other, cf. Eq. (4.55), so the corresponding occupation factors will be equal too, in the whole kinematic range where geometric scaling applies (which includes the saturation domain at $z \leq 1$ and the BFKL regime (5.27)).

v) The dependence of the ratio $\mathcal{R}_{pA}$ upon A is also interesting, since this corresponds to the *centrality dependence* of the ratio R_{dAu} measured at RHIC [27, 57]. Consider the A–dependence for momenta around the Cronin peak: Whereas at $\tau = 0$, the ratio $\mathcal{R}_{pA}(k_\perp \sim$

$Q_s(A)$) is logarithmically *increasing* with A (recall Eq. (5.21)), this tendency is rapidly reversed by the evolution, as manifest by inspection of Eq. (5.33) : After only a small rapidity increase $\tau \sim \tau_c \sim 1$, $\mathcal{R}_{pA}(k_\perp, \tau)$ becomes a *decreasing* function of A for any $k_\perp$, in qualitative agreement with the corresponding change in the centrality dependence observed in the data, cf. Fig. 18.

6 The color dipole picture

In this section, we shall give an introduction to the color dipole picture developed by A. Mueller [43, 58], which is an intuitive and very economical description of the lightcone wavefunction in the BFKL approximation and for large N_c. As we shall see, this formalism allows one to study the onset of unitarity corrections in dipole–dipole scattering at high energy, and is also well suited for numerical calculations, as demonstrated by the explicit Monte–Carlo simulations performed by Salam [44, 45]. Our main emphasis will be on the relation between this picture and the CGC formalism discussed previously. As we shall see, the color dipole picture includes at the same time *less* and *more* than the CGC effective theory as developed so far: It contains *less* in the sense that it does not include saturation effects in the wavefunction, and thus applies only in the dilute regime, where such effects are unimportant. But it contains *more*, in the sense that it allows for a faithful description of the wavefunction and its evolution in the dilute regime, including the effects of *particle number fluctuations* which have been neglected in the JIMWLK evolution of the CGC. Such fluctuations are a priori important at low density; but, somehow unexpectedly, they turn out to play an important role also in the approach towards saturation, as we shall discover in Sect. 7. Despite of these dissimilarities, the color dipole picture can be equivalently reformulated as the theory of a (non–saturated) *color glass* [163], which however evolves according to a different renormalization group equation [50] as compared to JIMWLK.

6.1 A Markovian picture for the BFKL evolution at large N_c

Let us start with an elementary color dipole (a quark–antiquark pair in a color singlet state) at low energy and study the evolution of the associated lightcone wavefunction when increasing the energy. We shall work in the leading logarithmic approximation at high energy, as before, but we shall also assume that the number of colors is large: $N_c \gg 1$. Also, we shall assume that the final energy is not too high, so that the saturation effects within the evolved wavefunction remain negligible: the evolution is linear, of the BFKL type. Under these assumptions, a mathematically simple and physically intuitive description emerges for the evolution of the *norm* (i.e., the modulus squared) of the wavefunction — the quantity which acts as a weight function in the calculation of expectation values over the dipole wavefunction. (This is the analog of the CGC weight function and plays a similar role, e.g., when computing the elastic S–matrix or the total cross–section for the scattering between the evolved dipole and some external target.) Namely, it turns out that the relevant evolution is a classical stochastic process in which a system of (elementary)
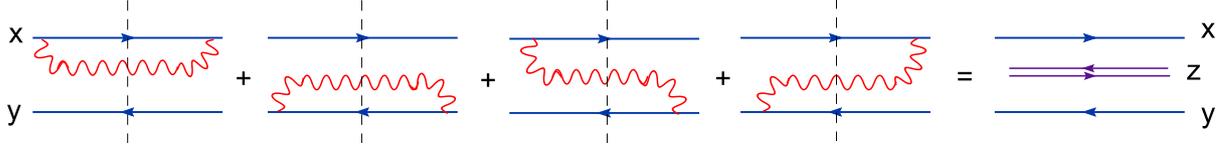
Fig. 45. `dipole_evol`

dipoles evolves through $1 \to 2$ dipole splitting [58, 70].

To avoid any confusion between the evolved dipole and an elementary one (i.e., a $q\bar{q}$ pair without additional gluons), we shall refer to the former as the 'onium' [58]. Let us examine the first few steps in the evolution of the onium. As usual, there are two different, but equivalent, ways to think about this evolution — either as boosting the onium to higher rapidities, or as including gluons with smaller and smaller values of x in the onium wavefunction —, and in what follows we shall adopt the second perspective.

The first step is illustrated in Fig. 45. The norm of the wavefunction is depicted as the product between the direct amplitude (on the left) and the complex conjugate (c.c.) amplitude (on the right), separated by a dashed vertical line. In this first step, a small–x gluon is emitted, in the direct amplitude, from either the original quark or the original antiquark, and then it is absorbed back again, in the c.c. amplitude, by any of these two fermions. This generates the four diagrams displayed in the l.h.s. of Fig. 45. Since the transverse coordinates of the quark and the antiquark are not modified by the emission of the small–x gluon, it is convenient to use the transverse coordinates to indicate the various fields, as shown in Fig. 45.

The key simplification arising at large N_c is that a gluon can be replaced by a pointlike quark–antiquark pair in a color octet state. This can be seen with the help of the Fierz identity:

$$t_{ij}^a \, t_{kl}^a \; = \; \frac{1}{2} \, \delta^{il} \delta^{jk} \; - \; \frac{1}{2N_c} \, \delta^{ij} \delta^{kl} \,, \tag{6.1}$$

where the two color matrices in the fundamental representation correspond to the vertices for gluon emission and, respectively, absorbtion in the diagrams in Fig. 45. When $N_c \to \infty$, one can neglect the second term in the r.h.s. of Eq. (6.1), and the remaining term describes indeed the propagation of a $q\bar{q}$ pair (in so far as color indices are concerned). Accordingly, the radiated gluon can be effectively replaced by a $q\bar{q}$ pair at $\boldsymbol{z}$, and the color indices are such that the 'quark' component of the gluon (at $\boldsymbol{z}$) together with the original antiquark (at $\boldsymbol{y}$) make up a new color dipole, and similarly for the 'antiquark' at $\boldsymbol{z}$ and the original quark at $\boldsymbol{x}$. That is, as a result of the evolution, the original dipole $(\boldsymbol{x}, \boldsymbol{y})$ has effectively split into two new dipoles $(\boldsymbol{x}, \boldsymbol{z})$ and $(\boldsymbol{z}, \boldsymbol{y})$, as illustrated in the r.h.s. of Fig. 45.

In the second rapidity step, another gluon with an even smaller value of x can be emitted from either the original $q\bar{q}$ pair $(\boldsymbol{x}, \boldsymbol{y})$ or from the first gluon at $\boldsymbol{z}$. In the latter case, it is convenient to think about the second gluon as being emitted from either the 'quark' or the 'antiquark' components of the first gluon.

164

Consider, e.g., the case where the second gluon is emitted, in the direct amplitude, from the quark $\boldsymbol{x}$ or the 'antiquark' $\boldsymbol{z}$. Then, there are two possibilities for the c.c. amplitude: (i) The second gluon is absorbed back again by the quark $\boldsymbol{x}$ or the 'antiquark' $\boldsymbol{z}$; in such a case, the process can be represented as the splitting of the dipole $(\boldsymbol{x}, \boldsymbol{z})$ into two new dipoles $(\boldsymbol{x}, \boldsymbol{w})$ and $(\boldsymbol{w}, \boldsymbol{z})$. (Note that the dipole $(\boldsymbol{z}, \boldsymbol{y})$ is merely a spectator in this scenario.) (ii) The second gluon is rather absorbed by the 'quark' $\boldsymbol{z}$ or the antiquark $\boldsymbol{y}$; this process does not admit a dipolar interpretation, but on the other hand it is suppressed at large N_c (by a power $1/N_c^2$ with respect to the first one).

A similar discussion holds when the second gluon is emitted off the 'quark' $\boldsymbol{z}$ or the antiquark $\boldsymbol{y}$, and, more generally, for higher steps in this evolution: In the large–N_c limit, the evolution can be restricted to dipole configurations and proceeds via dipole splitting. After being produced, the dipoles evolve independently from each other, yet the ensuing dipole branches are correlated with each other, because they have a common ancestor. In order to follow the evolution of the dipole distribution with increasing τ, we need the probability density for the elementary splitting, that we shall now compute.

For definiteness, we take the onium to be a right–mover, in which case it is convenient to use the LC gauge $A^+ = 0$. Consider first the soft gluon emission from the quark component of the dipole alone. The amplitude for a quark with longitudinal momentum p^+ and transverse momentum $p_\perp = 0$ to emit a small–x gluon with longitudinal momentum $k^+ \ll p^+$ and transverse momentum $k_\perp$ is easily computed as [70]

$$\mathcal{M}_\lambda^a(k^+, \boldsymbol{k}) \simeq gt^a \frac{\epsilon_\lambda^-}{\sqrt{(2\pi)^3 2k^+}} \frac{2k^+}{k_\perp^2} = \frac{2gt^a}{\sqrt{(2\pi)^3 2k^+}} \frac{\boldsymbol{\epsilon}_\lambda \cdot \boldsymbol{k}}{k_\perp^2}, \tag{6.2}$$

where a and λ represent color and polarization indices for the emitted gluon, and the quark color indices are implicit (the amplitude is a color matrix in the fundamental representation). Furthermore, $gt^a \epsilon_\lambda^-$ is the quark–gluon vertex in the eikonal approximation, $2k^+/k_\perp^2 = 1/k^-$ is the energy denominator for the radiation process (within the present approximations), and the factor involving a square–root arises from the normalization of the gluon wavefunction. In writing Eq. (6.2), we have also used the fact that, for a gluon in the gauge $A^+ = 0$, the transversality condition $k_\mu \epsilon_\lambda^\mu = 0$ implies $\epsilon_\lambda^- = \boldsymbol{\epsilon}_\lambda \cdot \boldsymbol{k}/k^+$.

As a first application of Eq. (6.2), let us use it to deduce the familiar expression of the differential probability for bremsstrahlung, which appeared already in Sect. 1.4. We have:

$$dP_{\text{Brem}} \equiv \sum_{a,\lambda} |\mathcal{M}_\lambda^a(k^+, \boldsymbol{k})|^2 \, d^2\boldsymbol{k} \, dk^+ \simeq \frac{\alpha C_F}{\pi^2} \frac{d^2\boldsymbol{k}}{k_\perp^2} \frac{dk^+}{k^+}, \tag{6.3}$$

where we have used $\sum_\lambda \epsilon_\lambda^i (\epsilon_\lambda^j)^* = \delta^{ij}$ and the Casimir $C_F = (N_c^2 - 1)/2N_c$ of the fundamental representation has been generated via the quark color algebra, i.e., by taking the sum over the final color states and the average over the initial ones: $(1/N_c)t_{ij}^a (t_{ij}^a)^* = C_F$. At large N_c, one can approximate $C_F \approx N_c/2$.

For the problem of the high–energy evolution, it is preferable to use the Fourier transform of the amplitude (6.2) to coordinate space, obtained as ($\boldsymbol{x}$ denotes the transverse coordinate of the quark, and $\boldsymbol{z}$ that of the emitted gluon)

$$\mathcal{M}_\lambda^a(k^+, \boldsymbol{x}-\boldsymbol{z}) \,=\, \int \frac{\mathrm{d}^2\boldsymbol{k}}{(2\pi)^2}\, \mathrm{e}^{i\boldsymbol{k}\cdot(\boldsymbol{x}-\boldsymbol{z})} \mathcal{M}_\lambda^a(k^+, \boldsymbol{k}) = \frac{2\mathrm{i}gt^a}{\sqrt{(2\pi)^3 2k^+}}\, \frac{\boldsymbol{\epsilon}_\lambda \cdot (\boldsymbol{x}-\boldsymbol{z})}{(\boldsymbol{x}-\boldsymbol{z})^2}\,. \tag{6.4}$$

To deduce the corresponding amplitude for a *dipole*, one must add a similar contribution for the emission from the antiquark at $\boldsymbol{y}$, and also enforce the quark and the antiquark to be in a same color state before the gluon is emitted (because the original $q\bar{q}$ pair was in a color singlet state). To ensure the proper counting of the color states, a factor $1/\sqrt{N_c}$ must then be introduced in the amplitude. One has:

$$\mathcal{M}_\lambda^{a,ij}(k^+, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \,=\, \frac{t_{ji}^a}{\sqrt{N_c}}\, \frac{2\mathrm{i}g}{\sqrt{(2\pi)^3 2k^+}} \left\{ \frac{\boldsymbol{\epsilon}_\lambda \cdot (\boldsymbol{x}-\boldsymbol{z})}{(\boldsymbol{x}-\boldsymbol{z})^2} \,-\, \frac{\boldsymbol{\epsilon}_\lambda \cdot (\boldsymbol{y}-\boldsymbol{z})}{(\boldsymbol{y}-\boldsymbol{z})^2} \right\}, \tag{6.5}$$

where i and j denote the color indices (for the quark and, respectively, the antiquark) in the final state, *after* the gluon has been emitted. The differential probability for dipole splitting is finally obtained by taking the modulus squared of this dipole amplitude and summing over the final color indices. One thus obtains [58] :

$$\mathrm{d}P_{\text{split}} \,=\, \frac{\alpha N_c}{2\pi^2}\, \frac{(\boldsymbol{x}-\boldsymbol{y})^2}{(\boldsymbol{x}-\boldsymbol{z})^2(\boldsymbol{y}-\boldsymbol{z})^2}\, \mathrm{d}^2\boldsymbol{z}\, \mathrm{d}\tau \,\equiv\, \frac{\bar{\alpha}}{2\pi}\, \mathcal{M}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})\, \mathrm{d}^2\boldsymbol{z}\, \mathrm{d}\tau\,, \tag{6.6}$$

where we have set $\mathrm{d}k^+/k^+ = \mathrm{d}\tau$ and the *dipole kernel* has been generated as in Eq. (3.75); that is, $\mathcal{M}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) = \mathcal{K}_{\boldsymbol{xxz}} + \mathcal{K}_{\boldsymbol{yyz}} - \mathcal{K}_{\boldsymbol{xyz}} - \mathcal{K}_{\boldsymbol{xyz}}$, with the four terms in the r.h.s. corresponding to the four diagrams in Fig. 45.

Note the emergence of the coupling constant scaled by the number of colors, $\bar{\alpha} \equiv \alpha N_c/\pi$, which is characteristic for a large–N_c approximation. This shows that the meaningful way to take the large–N_c limit is to let $N_c \to \infty$ and simultaneously $\alpha \to 0$ in such a way that $\bar{\alpha}$ remains constant. Besides, $\bar{\alpha} \ll 1$ in the perturbative regime. As we shall see, in the presence of unitarity corrections or saturation effects, the definition of the 'large–N_c limit' needs to be adjusted in order to accommodate the relevant physics.

As anticipated, under the present, high–energy and large–N_c, approximations, the quantum evolution of the onium can be described as a *classical stochastic process* for a system of dipoles living in the two transverse dimensions and 'time' τ; this is the 'color dipole' picture [58]. At each step in this process, one of the preexisting dipoles can split into two new dipoles with a differential probability given by Eq. (6.6). The latter is independent of the history of the evolution in the previous steps, so the process is *Markovian* [154, 156]. The result of this evolution can then be described as a *statistical ensemble* of dipole configurations endowed with a probability distribution evolving with τ.

Specifically, a given configuration is specified by the number of dipoles N and by $N-1$ transverse coordinates $\{\boldsymbol{z}_i\} = \{\boldsymbol{z}_1, \boldsymbol{z}_2, ...\boldsymbol{z}_{N-1}\}$, such that the coordinates of the N dipoles are $(\boldsymbol{z}_0, \boldsymbol{z}_1), (\boldsymbol{z}_1, \boldsymbol{z}_2),...,(\boldsymbol{z}_{N-1}, \boldsymbol{z}_N)$, with $\boldsymbol{z}_0 \equiv \boldsymbol{x}_0$ and $\boldsymbol{z}_N \equiv \boldsymbol{y}_0$. (Note that, from now on, we shall denote the transverse coordinates of the original dipole as $(\boldsymbol{x}_0, \boldsymbol{y}_0)$, for more clarity.) We therefore introduce the *probability density* $P_N(\{\boldsymbol{z}_i\}|\tau)$ to find a given configuration (the dependence upon the original coordinates $(\boldsymbol{x}_0, \boldsymbol{y}_0)$ is kept implicit), and study its evolution with τ. Physically, the intermediate coordinates $\boldsymbol{z}_i$ represent the positions of the emitted

gluons, and P_N is the norm of the Fock space component of the LC wavefunction which involves $N-1$ soft gluons.

The dipole probabilities are normalized according to

$$\sum_{N=1}^{\infty} \int d\Gamma_N \, P_N(\{\boldsymbol{z}_i\}|\tau) = 1, \tag{6.7}$$

where the phase space integration is $d\Gamma_N = d^2\boldsymbol{z}_1 d^2\boldsymbol{z}_2 \ldots d^2\boldsymbol{z}_{N-1}$. Then expectation values over the onium wavefunction are obtained as

$$\langle \mathcal{O}(Y) \rangle = \sum_{N=1}^{\infty} \int d\Gamma_N \, P_N(\{\boldsymbol{z}_i\}; Y) \, \mathcal{O}_N(\{\boldsymbol{z}_i\}), \tag{6.8}$$

`opeav`

where the operator $\mathcal{O}$ represents some physical observable which depends upon the dipole coordinates alone.

From Eq. (6.6), it is straightforward to deduce the evolution law for the dipole probabilities (the *'Master equations'* in the language of statistical physics [154, 156]). When increasing rapidity from τ to $\tau + d\tau$, the probability for a given configuration $\{\boldsymbol{z}_i\}$ of N dipoles can increase via splitting within a previous configuration of $N-1$ dipoles and decrease because of the splitting of one of the N dipoles in the configuration at study. This leads to a set of coupled evolution equations for the probabilities P_N [163, 220]

$$\frac{\partial P_N(\{\boldsymbol{z}_i\}|\tau)}{\partial \tau} = -\frac{\bar{\alpha}}{2\pi} \left[\sum_{i=1}^{N} \int d^2\boldsymbol{z} \, \mathcal{M}(\boldsymbol{z}_{i-1}, \boldsymbol{z}_i, \boldsymbol{z}) \right] P_N(\{\boldsymbol{z}_i\}|\tau) \tag{6.9}$$

`evolP`

$$+ \frac{\bar{\alpha}}{2\pi} \sum_{i=1}^{N-1} \mathcal{M}(\boldsymbol{z}_{i-1}, \boldsymbol{z}_{i+1}, \boldsymbol{z}_i) \, P_{N-1}(\boldsymbol{z}_1, ..., \boldsymbol{z}_{i-1}, \boldsymbol{z}_{i+1}, ..., \boldsymbol{z}_{N-1}|\tau),$$

where in the r.h.s. one can easily identify the loss and gain terms characteristic of a Master equation [156]. Note that, from the point of view of perturbative QCD, the loss term in Eq. (6.9) corresponds to the decrease in the norm of the $(N-1)$–gluon Fock state, as associated with *virtual–gluon* loop corrections to the LC wavefunction. The corresponding diagrams have not been explicitly considered above, rather their overall contribution has been directly inferred — via the condition of probability conservation — from the respective contribution of the *real–gluon* corrections that we have computed indeed (and which are responsible for the gain terms in Eq. (6.9)).

The hierarchy represented by Eq. (6.9) must be solved with the initial condition that, at $\tau = 0$, there is only one dipole: $P_N(\tau = 0) = \delta_{N1}$. Thus, the normalization condition, Eq. (6.7), is trivially satisfied at $\tau = 0$ and is further preserved at any τ by the evolution according to Eq. (6.9) (as one can check by replacing $\mathcal{O}_N \to 1$ in Eq. (6.11) below).

Note a subtle point concerning the dipole probabilities: As it stands, the integral over $\boldsymbol{z}$ yielding the loss term in Eq. (6.9) has logarithmic divergencies due to the poles of the dipole kernel at $\boldsymbol{z} = \boldsymbol{z}_i$ and $\boldsymbol{z} = \boldsymbol{z}_{i-1}$. These singularities reflect the fact that, in the construction of the wavefunction, one cannot forbid the radiation of dipoles of arbitrarily small sizes. To regularize these divergences at intermediate steps, one must introduce some ultraviolet cutoff, e.g., a minimal size $r_{\min}$ for the radiated dipoles. The probabilities P_N

will logarithmically depend upon this cutoff; for instance, the survival probability P_1 for the original dipole $(\boldsymbol{x}_0, \boldsymbol{y}_0)$, or 'Sudakov factor', is easily obtained as

$$P_1(\tau) \,=\, \exp\left\{-\bar{\alpha}\tau \ln \frac{r_0^2}{r_{\min}^2}\right\}, \tag{6.10}$$

where $r_0 \equiv |\boldsymbol{x}_0 - \boldsymbol{y}_0|$. But the cutoff dependence cancels out between the loss and gain terms in the calculation of physical quantities according to Eq. (6.8), as we shall see below on explicit examples.

From Eqs. (6.9) and (6.8), one can deduce the following evolution equation for the expectation value of an observable [163]

$$\frac{\partial \langle \mathcal{O}(\tau) \rangle}{\partial \tau} = \frac{\bar{\alpha}}{2\pi} \sum_{N=1}^{\infty} \int \mathrm{d}\Gamma_N \, P_N(\{\boldsymbol{z}_i\}|\tau) \sum_{i=1}^{N} \int \mathrm{d}^2 z \, \mathcal{M}(\boldsymbol{z}_{i-1}, \boldsymbol{z}_i, \boldsymbol{z}) \tag{6.11}$$

$$\times \Big[-\mathcal{O}_N(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_{N-1}) + \mathcal{O}_{N+1}(\boldsymbol{x}_1, \ldots, \boldsymbol{z}_{i-1}, \boldsymbol{z}, \boldsymbol{z}_i, \ldots, \boldsymbol{z}_{N-1}) \Big],$$

where the first term, $(-\mathcal{O}_N)$, comes from the loss terms in the Master equation, whereas the second one, $\mathcal{O}_{N+1}$, is generated by the gain terms there. By inspection of Eq. (6.11), one can anticipate the mechanism for the compensation of the ultraviolet singularities: Consider, e.g., the pole of the dipole kernel at $\boldsymbol{z} = \boldsymbol{z}_i$ and notice that, within the 'gain' term $\mathcal{O}_{N+1}$, $\boldsymbol{z}$ is a common leg for the two adjacent dipoles $(\boldsymbol{z}_{i-1}, \boldsymbol{z})$ and $(\boldsymbol{z}, \boldsymbol{z}_i)$. Thus, when $\boldsymbol{z} \to \boldsymbol{z}_i$, the dipole $(\boldsymbol{z}, \boldsymbol{z}_i)$ becomes of zero size and cannot have any physical influence; hence, in that limit, $\mathcal{O}_{N+1}$ must effectively reduce to $\mathcal{O}_N$, so the residue of the would–be pole $\boldsymbol{z} = \boldsymbol{z}_i$ appears to vanish. This will be later checked on explicit expressions for the observable $\mathcal{O}$.

The hierarchy in Eq. (6.9) has a relatively simple structure, which reflects the fact that the onium evolution as described above is *linear*: For $N = 1$, we have a closed equation for P_1 that we have solved in Eq. (6.10), and for $N \geq 2$ the evolution couples P_N only to P_{N-1}. Hence, the second equation determines P_2 in terms of P_1, etc., and the equations can be solved one after the other, without any need for a truncation (in contrast to the Balitsky hierarchy, which describes *non–linear* evolution in the presence of saturation). In practice, it turns out that the most convenient way to compute the dipole distribution is via Monte Carlo simulations [44, 45]. In the next subsection, we shall see that the evolution described by Eq. (6.9) is of the BFKL type, as expected.

The manifest linearity of the evolution reflects our assumption that saturation effects can be neglected within the onium wavefunction — the gluons can split in the course of the evolution, but they cannot merge with each other —, an approximation that we know to fail at sufficiently high energy. Incidentally, it might be interesting to notice that the evolution would remain strictly linear (at any energy) within the *formal* large–N_c limit: $N_c \to \infty$ at fixed energy. Indeed, as we have seen in the construction of the dipole picture, the gluon exchanges between different dipoles correspond to 'non–planar' diagrams, which are suppressed by powers of $1/N_c$. But, of course, the physically interesting "large–N_c approximation" is that one where N_c is large ($N_c \gg 1$) but *finite*, and then it is easy to see that the growth of the dipole distribution with τ rapidly compensates for the $1/N_c$

168

suppression of the saturation effects. Let us make this argument more precise and use it to estimate the rapidity range for the validity of the dipole picture.

We shall soon discover that the characteristic amplitude for dipole–dipole interactions is of order $\alpha^2 \equiv \bar{\alpha}^2/N_c^2$ (see Sect. 6.5.1), and thus is suppressed at large N_c, as anticipated. On the other hand, the average dipole number density $\langle n \rangle_\tau$ increases very fast with the energy, like the BFKL pomeron (see Sect. 6.2) : $\langle n \rangle_\tau \sim \mathrm{e}^{\omega_0 \tau}$ with $\omega_0 = (4 \ln 2)\bar{\alpha}$. Thus, for any finite value of N_c, there exists a value τ_c at which the $1/N_c^2$ suppression of the elementary amplitude is compensated by the large density of dipoles available for scattering. This happens, roughly, when the total amplitude for the interaction between a given dipole and any other dipole in the system becomes of order one:

$$
\alpha^2 \, \mathrm{e}^{\omega_0 \tau} \sim 1 \qquad \text{for} \qquad \tau = \tau_c \simeq \frac{1}{\omega_0} \ln \frac{N_c^2}{\bar{\alpha}^2} . \tag{6.12}
$$

For even larger rapidities, the dipole picture must be abandoned and the $1/N_c$ expansion must be modified to account for the high–density effects: quantities of order $\alpha N_c \tau$ and $\alpha^2 \mathrm{e}^{\omega_0 \tau}$ must be treated on equal footing. Note, however, that the saturation effects cannot be accommodated in the dipole basis, since the interaction between two dipoles inside the wavefunction leads in general to more complicated color configurations, like quadrupoles, sextupoles, etc. In particular, one cannot treat the saturation effects as "dipole recombination" [221], although, as we shall see in Sect. 7, one can arrive at an effective description [49, 61, 222] which looks extremely (even confusingly !) close to it.

6.2   Dipole densities and their evolution

By using Eq. (6.11), we shall now deduce evolution equations for the *dipole number densities* [43, 48, 58, 163, 223]. The resulting equations will make it clear that the onium evolution as described above is a special representation of the BFKL evolution (at large N_c), which is however more general than the BFKL equation by itself, in the sense of including *correlations* induced through dipole splitting. The equations to be established in this subsection will play an essential role in the construction of improved evolution equations (going beyond the Balitsky–JIMWLK hierarchy) in Sect. 7.

Consider first the average *dipole number density*. The corresponding operator for an N–dipole configuration is

$$
n_N(\boldsymbol{x}, \boldsymbol{y}) = \sum_{j=1}^{N} \delta^{(2)}(\boldsymbol{z}_{j-1} - \boldsymbol{x}) \delta^{(2)}(\boldsymbol{z}_j - \boldsymbol{y}). \tag{6.13}
$$

densit

By using Eq. (6.11) with $\mathcal{O}_N \equiv n_N$, and after simple manipulations we arrive at the

169
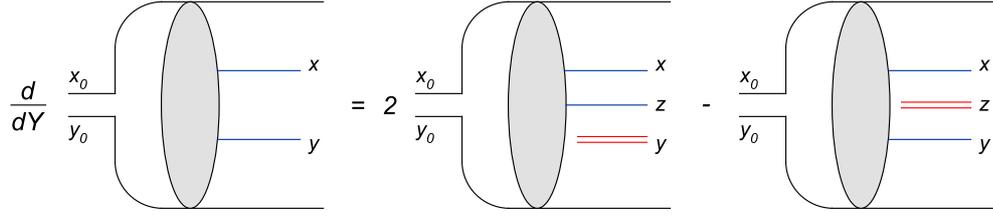
Fig. 46. The one–step evolution of the average dipole number density as described by Eq. (6.14).

Fig_T1

following evolution equation for the average dipole number density $n_\tau(\boldsymbol{x}, \boldsymbol{y}) \equiv \langle n(\boldsymbol{x}, \boldsymbol{y}) \rangle_\tau$

$$
\begin{aligned}
\frac{\partial n_\tau(\boldsymbol{x}, \boldsymbol{y})}{\partial \tau} &= \frac{\bar{\alpha}}{2\pi} \int \mathrm{d}^2 \boldsymbol{z} \Big[- \mathcal{M}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})\, n_\tau(\boldsymbol{x}, \boldsymbol{y}) \\
&\qquad\qquad + \mathcal{M}(\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{y})\, n_\tau(\boldsymbol{x}, \boldsymbol{z}) + \mathcal{M}(\boldsymbol{z}, \boldsymbol{y}, \boldsymbol{x})\, n_\tau(\boldsymbol{z}, \boldsymbol{y}) \Big] \\
&\equiv \frac{\bar{\alpha}}{2\pi} \int \mathrm{d}^2 \boldsymbol{z}\, \mathcal{N}_{\boldsymbol{xyz}} \otimes n_\tau(\boldsymbol{x}, \boldsymbol{y}),
\end{aligned}
\tag{6.14}
$$

evolnu

which is illustrated in Fig. 46. This is recognized as the BFKL equation in coordinate space [58, 163]. Note that the poles in the various dipole kernels are innocuous, as their residues cancel in the integrand: e.g., when $\boldsymbol{z} \to \boldsymbol{x}$, the first term cancels against the third one, whereas the second term is not singular in that limit. Thus, as anticipated in the previous subsection, the ultraviolet singularities cancel between the gain and the loss terms, and the dipole number density grows with the energy like the BFKL pomeron.

Now we turn to the *dipole pair density*. The corresponding operator for a given N–dipole configuration is defined as [48]:

$$
n_N^{(2)}(\boldsymbol{x}_1, \boldsymbol{y}_1; \boldsymbol{x}_2, \boldsymbol{y}_2) \equiv n_N(\boldsymbol{x}_1, \boldsymbol{y}_1) n_N(\boldsymbol{x}_2, \boldsymbol{y}_2) - \delta_{\boldsymbol{x}_1 \boldsymbol{x}_2} \delta_{\boldsymbol{y}_1 \boldsymbol{y}_2} n_N(\boldsymbol{x}_1, \boldsymbol{y}_1),
\tag{6.15}
$$

where the subtracted term eliminates the "pairs" made of the same dipole. With this subtraction, $n^{(2)}$ measures non–trivial correlations between distinct pairs.

The construction of the equation obeyed by the *average* pair density $n_\tau^{(2)} \equiv \langle n^{(2)} \rangle_\tau$ is a bit more tedious (see Ref. [48] for details), but its final structure can be easily anticipated via physical considerations. To that aim, let us consider the construction of the dipole pair density via evolution, starting with a single dipole at low energy. That is, consider the following initial conditions at $\tau = 0$:

$$
n_0(\boldsymbol{x}, \boldsymbol{y}) = \delta^{(2)}(\boldsymbol{x} - \boldsymbol{x}_0) \delta^{(2)}(\boldsymbol{y} - \boldsymbol{y}_0) \qquad \text{and} \qquad n_0^{(2)} = 0 \,.
\tag{6.16}
$$

After increasing the rapidity by $\mathrm{d}\tau$, the original dipole can decay into two dipoles $(\boldsymbol{x}_0, \boldsymbol{z})$ and $(\boldsymbol{z}, \boldsymbol{y}_0)$, with the differential probability (6.6). Thus, if one measures the pair density (6.15) at $\tau = \mathrm{d}\tau$, one can find a non–zero result provided the coordinates $(\boldsymbol{x}_1, \boldsymbol{y}_1)$ and $(\boldsymbol{x}_2, \boldsymbol{y}_2)$ of the measured dipoles match onto the coordinates $(\boldsymbol{x}_0, \boldsymbol{z})$ and $(\boldsymbol{z}, \boldsymbol{y}_0)$ of the two child dipoles. This requires the measured dipoles to have a common leg: $\boldsymbol{x}_2 = \boldsymbol{y}_1$ or $\boldsymbol{y}_2 = \boldsymbol{x}_1$ (see Fig. 47) — we shall say that they are *contiguous* with each other. But even
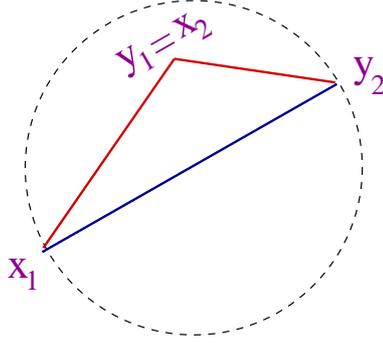
Fig. 47. *The geometry of dipole splitting* $\boxed{\texttt{SPLITG}}$

when this condition is satisfied, the measured pair density $n_{\mathrm{d}\tau}^{(2)}$ is parametrically small, of order $\bar{\alpha}\mathrm{d}\tau$, to be compared to the one–body density $n_{\mathrm{d}\tau}$, which is of order one.

We see that, so long as $\bar{\alpha}\mathrm{d}\tau \lesssim 1$, the two–body correlations remain weak, but they get built via dipole splitting in the course of the evolution. In this process, the one–body dipole density n_τ acts as a *source* for the two–body correlations. Given the form (6.6) of the differential probability for one–step evolution, and the geometry of splitting in Fig. 47, it is clear that the rate of change in the dipole pair density due to the splitting of one of the dipoles in n_τ reads (with $\boldsymbol{r}_1 = \boldsymbol{x}_1 - \boldsymbol{y}_1$ and $\boldsymbol{r}_2 = \boldsymbol{x}_2 - \boldsymbol{y}_2$) :

$$
\frac{\partial}{\partial \tau}\, n_\tau^{(2)}(\boldsymbol{x}_1, \boldsymbol{y}_1\,;\boldsymbol{x}_2, \boldsymbol{y}_2)\Big|_{\mathrm{fluct.}} = \frac{\bar{\alpha}}{2\pi}\, \frac{(\boldsymbol{r}_1 + \boldsymbol{r}_2)^2}{\boldsymbol{r}_1^2 \boldsymbol{r}_2^2}\Big\{ \delta^{(2)}(\boldsymbol{x}_2 - \boldsymbol{y}_1) n_\tau(\boldsymbol{x}_1, \boldsymbol{y}_2)
$$
$$
+ \delta^{(2)}(\boldsymbol{y}_2 - \boldsymbol{x}_1) n_\tau(\boldsymbol{x}_2, \boldsymbol{y}_1)\Big\}. \tag{6.17}
$$

As indicated in the l.h.s. of Eq. (6.17), this is just the "fluctuating" contribution to the evolution of $n_\tau^{(2)}$, in which the two measured dipoles arise from the splitting of the same parent dipole (and thus are contiguous with each other). The general equation is easily obtained after adding in the r.h.s. the terms linear in $n_\tau^{(2)}$ which describe the usual BFKL evolution of any of the two measured dipoles. Thus, the complete equation reads:

$$
\frac{\partial\, n_\tau^{(2)}(\boldsymbol{x}_1, \boldsymbol{y}_1; \boldsymbol{x}_2, \boldsymbol{y}_2)}{\partial \tau} = \frac{\bar{\alpha}}{2\pi}\Bigg[\int \mathrm{d}^2\boldsymbol{z}\, \mathcal{N}_{\boldsymbol{x}_1 \boldsymbol{y}_1 \boldsymbol{z}} \otimes n_\tau^{(2)}(\boldsymbol{x}_1, \boldsymbol{y}_1; \boldsymbol{x}_2, \boldsymbol{y}_2) \tag{6.18}
$$
$$
+ \mathcal{M}(\boldsymbol{x}_1, \boldsymbol{y}_2, \boldsymbol{x}_2)\, n_\tau(\boldsymbol{x}_1, \boldsymbol{y}_2)\, \delta^{(2)}(\boldsymbol{x}_2 - \boldsymbol{y}_1)\Bigg] + \big\{1 \leftrightarrow 2\big\},
$$

$\boxed{\texttt{eqn2}}$

(we use the notation introduced in Eq. (6.14)) and is illustrated in Fig. 48.

It is now manifest in the r.h.s. of Eq. (6.18) that the fluctuation term, proportional to n_τ, controls the growth of $n_\tau^{(2)}$ in the dilute regime where $n_\tau \gg n_\tau^{(2)}$. Although the previous discussion has focused, for simplicity, on the low–energy regime $\bar{\alpha}\tau \lesssim 1$, one should keep in mind that the 'dilute regime' is not restricted to the early stages of the evolution: However large is τ, the dipole distribution will always have a tail corresponding to small dipole sizes $r \ll 1/Q_s(\tau)$ in which the density is low and the fluctuations are important. (Recall the discussion of the saturation front in Sect. 4.3.2.) In particular, such a tail will
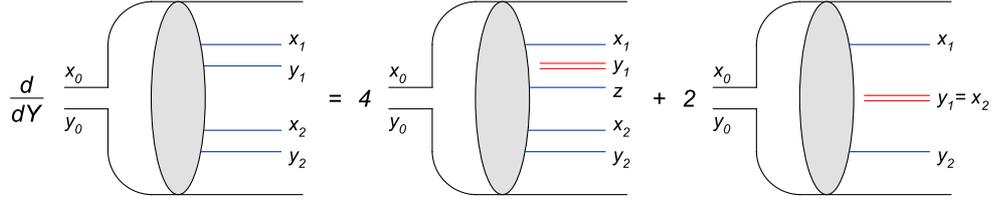
Fig. 48. *The one–step evolution of the dipole pair density $n_\tau^{(2)}$ as described by Eq. (6.18). The virtual BFKL term is not shown.* Fig_T2

be produced by the high–energy evolution even if one starts with a relatively dense system at $\tau = 0$, so like a large 'nucleus' composed with many dipoles.

On the other hand, once that a non–zero contribution to the dipole pair density has been generated via fluctuations, this contribution is then rapidly amplified by the standard BFKL evolution, which proceeds twice as fast for $n_\tau^{(2)}$ as compared to n_τ — because the two dipoles making the pair can simultaneously evolve. This too is manifest on the r.h.s. of Eq. (6.18) and it implies that, for sufficiently high energy, $n_\tau^{(2)}$ rises with τ as a 'double BFKL Pomeron' : $n_\tau^{(2)} \sim e^{2\omega_0\tau}$ for $\bar{\alpha}\tau \gtrsim 1$.

A similar discussion applies to the generic κ–body dipole density $n_\tau^{(\kappa)}$, with $\kappa \geq 2$, as defined via the straightforward generalization of Eq. (6.15), that is,

$$ n_{u_1...u_\kappa}^{(\kappa)} \equiv n_{u_1}(n_{u_2} - \delta_{u_1 u_2}) \ldots (n_{u_\kappa} - \delta_{u_1 u_\kappa} - \cdots - \delta_{u_{\kappa-1} u_\kappa}), \tag{6.19} $$

where $u_i = (\boldsymbol{x}_i, \boldsymbol{y}_i)$ is a 4–dimensional variable representing in a compact way the coordinates of the quark and the antiquark of the i–th dipole. The corresponding evolution equation involves the standard BFKL terms — a total of κ terms linear in $n_\tau^{(\kappa)}$, which describe the independent BFKL evolution of each of the dipoles —, together with *fluctuation terms* — a total of $\kappa(\kappa-1)/2$ terms linear in $n_\tau^{(\kappa-1)}$, which account for the possibility that a pair of dipoles with sizes $\boldsymbol{r}_i$ and, respectively, $\boldsymbol{r}_j$ from $n_\tau^{(\kappa)}$, which are contiguous with each other, be generated in one step of the evolution via the splitting of a unique dipole of size $\boldsymbol{r}_i + \boldsymbol{r}_j$ from $n_\tau^{(\kappa-1)}$ [48, 223].

Let us conclude this discussion with the following observation, which will play an important role in what follows: Although a dilute system, the onium does not evolve according to the weak–field limit of the JIMWLK Hamiltonian, which is the BFKL Hamiltonian (cf. Sect. 3.4.1), but rather according to the more general Master equations (6.9), which lead to a non–trivial hierarchy for the κ–body dipole densities, as we have just seen. Whereas the BFKL equation (6.14) describes the change in the *average* dipole number density alone, the higher equations in this hierarchy — which are not accommodated by the BFKL Hamiltonian, nor by the JIMWLK one — describe the construction of *correlations* in the evolving dipole distribution. To summarize, the dipole picture goes beyond the JIMWLK equation in the sense of providing a better description of the correlations produced by the evolution in the dilute regime, although only for large N_c.

6.3 Onium–CGC scattering: a lesson from boost–invariance

Now that we know how to construct the 'wavefunction' of the onium (more precisely, its modulus squared, or the probability density), let us consider some physical applications of the formalism. We start with the elastic scattering between the onium and some energetic hadron (a CGC). As we shall see, this calculation is interesting for, at least, two reasons: First, it reveals some limitations of the previously established Balitsky–JIMWLK (or BK) equations. Second, it will naturally lead us — in the next subsection — to a generalized 'color glass' picture (that of the onium), which transcends the standard picture provided by the JIMWLK evolution.

Let $\mathcal{S}(\boldsymbol{x}_0, \boldsymbol{y}_0; \tau)$ denote the S–matrix element for the elastic scattering between an onium (the 'projectile') and the hadronic target corresponding to a rapidity separation τ. At the time of scattering, the projectile consists in a statistical ensemble of dipoles which elastically scatter off the color fields in the target. This picture suggests the following 'factorization formula', which can more formally derived within the lightcone wavefunction formalism (see below for a sketch of its proof):

$$\mathcal{S}(\boldsymbol{x}_0, \boldsymbol{y}_0; \tau) = \sum_{N=1}^{\infty} \int d\Gamma_N \, P_N(\{\boldsymbol{z}_i\}|\tau - \tau_0) \left\langle S(1)S(2)\cdots S(N) \right\rangle_{\tau_0} . \qquad (6.20)$$

This formula is written in the Lorentz frame where the target rapidity is τ_0 and that of the projectile is $\tau - \tau_0$, with some restriction on the value of τ_0 (see below). In the r.h.s., $S(i) \equiv S(\boldsymbol{z}_{i-1}, \boldsymbol{z}_i)$ is the S–matrix for the scattering between the ith dipole in the projectile and a given configuration of the target fields. Accordingly, $\langle S(1)S(2)\cdots S(N)\rangle_{\tau_0}$, where the average refers to the target wavefunction, represents the S–matrix element for the elastic scattering of a system of N dipoles in a given transverse configuration. After some simple algebra, Eq. (6.20) can be equivalently rewritten as an expression for the 'forward' (i.e., elastic) scattering amplitude $\mathcal{T} \equiv 1 - \mathcal{S}$:

$$\mathcal{T}(\boldsymbol{x}_0, \boldsymbol{y}_0; \tau) = \sum_{\kappa=1}^{\infty} \frac{(-1)^{\kappa+1}}{\kappa!} \int du_1...du_\kappa \, n_{\tau-\tau_0}^{(\kappa)}(u_1...u_\kappa) \left\langle T^{(\kappa)}(u_1...u_\kappa)\right\rangle_{\tau_0} , \qquad (6.21)$$

where $n_{\tau-\tau_0}^{(\kappa)}$ is the κ–body dipole density in the projectile, cf. Eq. (6.19), and $T^{(\kappa)} = T(1)T(2)\cdots T(\kappa)$, with $T(i) \equiv 1 - S(i)$, is the amplitude for the simultaneous scattering of κ dipoles.

A priori, the choice of a specific frame should not influence the final result of the calculation, since the elastic S–matrix is boost invariant. However, the above expressions, Eqs. (6.20) and (6.21), are not valid in *any* frame: they assume that the projectile is dilute enough for the saturation effects to be negligible, which in turn implies a restriction on its rapidity (cf. Eq. (6.12)) : $\tau - \tau_0 \ll \tau_c$. But there is no corresponding restriction on the total rapidity τ: however large τ is, one can always choose the frame in such a way that $\tau - \tau_0$ be small enough. In particular, with the choice $\tau_0 \approx \tau$, the projectile reduces to a single dipole and then the problem is tantamount to the dipole–CGC scattering that we have already considered in some previous chapters. (The precise condition for neglecting the evolution of the projectile reads $\bar{\alpha}(\tau - \tau_0) \ll 1$.) It is nevertheless interesting to allow

173

τ_0 to take generic values (modulo the restriction on $\tau - \tau_0$ alluded to above), because by varying τ_0 one can study the interplay between the condition of boost–invariance and the high–energy evolution. Ultimately, this study will unveil some important limitations of the CGC formalism introduced so far.

Specifically, the constraint that $\tau - \tau_0 \ll \tau_c$ leaves a rather large range of variation for τ_0; within that range, Eqs. (6.20) and (6.21) must be independent of τ_0, to the accuracy of interest — i.e., up to corrections related to saturation effects in the projectile wavefunction (which are small indeed within the considered range for τ_0). By requiring that

$$\frac{\mathrm{d}\mathcal{S}(\tau)}{\mathrm{d}\tau_0} = 0, \tag{6.22}$$

within Eq. (6.20) and then using the Master equation (6.9) for the dipole probabilities (or, equivalently, from $\mathrm{d}\mathcal{T}/\mathrm{d}\tau_0 = 0$ together with the equations in Sect. 6.2 for the dipole densities), it is straightforward to deduce [221, 223] that the dipole amplitudes $\langle T^{(\kappa)} \rangle_{\tau_0}$ obey the large–N_c version of the Balitsky hierarchy [27] (cf. the discussion after Eq. (3.91)). In particular, the non–linear terms in these equations (the 'unitarity corrections') arise from the fluctuation terms in the equations for $n_\tau^{(\kappa)}$ with $\kappa \geq 2$: one dipole splits into two, and both child dipoles are then measured via scattering off the target. This finding is in agreement with the interpretation of the Balitsky equations in terms of projectile evolution, as explained in Sects. 1.4.3 and 3.5.

But although natural at a first sight, this result as a rather curious consequence, which sheds doubts on the internal consistency of the Balitsky equations (and implicitly also on that of the JIMWLK evolution of the target) : It implies that, at a mathematical level, Eqs. (6.20) and (6.21) are *strictly* independent off τ_0 so long as the projectile evolves according to the dipole picture and the target expectation values evolve according to the (large–N_c version of the) Balitsky–JIMWLK hierarchy. This is paradoxical since it implies that, if one uses the Balitsky–JIMWLK equations to compute the dipole amplitudes which enter Eq. (6.20), then the latter provides exactly the same result for $\mathcal{S}(\tau)$ in *any* frame, including the frames like the target rest frame ($\tau_0 \simeq 0$) where this formula is bound to be wrong. The solution to this paradox consists in recognizing that the Balitksy–JIMWLK equations are in fact incomplete [46–48].

This failure will be discussed at length in Sect. 7, where we shall also propose a remedy to it, but the origin of the problem can be already recognized here: This is the profound *dissymmetry* between the evolutions of the projectile and, respectively, the target as encoded in the Balitksy equations. We have just argued that the non–linear terms in these equations correspond to dipole (or gluon) number *fluctuations* in the evolution of the *projectile*. Alternatively, from the discussion in Sect. 3, we now that, from the perspective of the *target* evolution, the same terms corresponds to *saturation* effects. This 'duality' in our physical interpretation is quite natural, as it reflects the boost–invariance of the scattering amplitudes: by varying τ_0, one can reinterpret the same elementary evolution step as either a gluon splitting ($1 \rightarrow 2$) inside the projectile, or as a gluon merging ($2 \rightarrow 1$)

---

[27] The original derivation of the BK equation by Kovchegov [124] uses precisely this argument, further supplemented by the mean field assumption $\langle T^{(\kappa)} \rangle \approx \langle T(1) \rangle \langle T(2) \rangle \cdots \langle T(\kappa) \rangle$.

inside the target. But this discussion also shows that the physical information encoded in the Balitsky equations is necessarily *incomplete* : it misses the saturation effects in the projectile and also (by 'duality') the gluon–number fluctuations in the target.

One may hope that, by working in a well chosen frame, in which the projectile is dilute and the target is dense, all this missing information becomes irrelevant, so the Balitksy– JIMWLK equations can there be used as written. But the previous discussion of Eq. (6.20) shows that this cannot be true: By boost–invariance, the very same results would be obtained by working in a different frame where the physical situation is reversed (the projectile becomes dense, whereas the target is dilute), and where equations like (6.20) do not apply anymore — so the respective results are necessarily wrong, and then they are wrong in *any* frame. This 'duality' argument may look a little formal, or at least indirect, since based on boost–invariance alone; but in the subsequent discussion this will be corroborated by explicit dynamical considerations, which will shed more light on the 'missing physics' and eventually drive us, in Sect. 7, towards an improved set of evolution equations.

6.4 The 'color glass' description of the onium

The discussion in the previous subsection points towards some limitations of the CGC formalism in the description of the dilute regime. On the other hand, we know that the dipole picture provides a faithful description of this regime (at large N_c), including the essential correlations induced via gluon–number fluctuations. To better understand the limitations of the CGC picture, it would be interesting to compare the predictions of the dipole picture and of the JIMWLK equation at low density. A priori, these two formalisms are quite different from each other — one is formulated in terms of the (colorless) dipole number density, the other one in terms of the color charge density — and these differences prevent a direct comparison. One possibility would be to compare their respective predictions for the same physical observables, like the scattering amplitudes for given projectiles. This is the strategy that we shall follow in Sect. 7, where we shall combine these two approaches in a unified picture. Another possibility, that we shall develop in this section, is to reformulate the dipole picture as a generalized 'color glass' description [50, 163], valid at low density, which encompasses and extends (within this low density regime) the standard picture based on the JIMWLK equation, and thus directly demonstrates the limitations of the latter.

To that aim, let us return to the factorization formula (6.20) and show now that this can be interpreted as the scattering between two 'color glasses', which represent the hadronic target and the onium, respectively. For definiteness, we shall take the target to be a right–mover. Then, the projectile is a left–mover and the S–matrix $S(\boldsymbol{z}_{i-1}, \boldsymbol{z}_i)$ for the scattering of an elementary dipole is given by Eq. (1.36) with the Wilson lines $V^\dagger$ and V built with the color field $A_a^+ \equiv \alpha_R^a$ of the target, as shown in Eq. (1.35). The following mathematical identity will be useful in what follows (in matrix notations appropriate for the fundamental representation) :

$$\mathrm{P}\exp\left(-g\int \mathrm{d}x^- \frac{\delta}{\delta\rho_L^a(x^-,\boldsymbol{z})}\,t^a\right)\,\mathrm{e}^{-\mathrm{i}\int \mathrm{d}x^-\mathrm{d}^2\boldsymbol{x}\,\rho_L^a(x^-,\boldsymbol{x})\alpha_R^a(x^-,\boldsymbol{x})}\;=$$

$$=\;\mathrm{P}\exp\left(ig\int \mathrm{d}x^-\alpha_R^a(x^-,\boldsymbol{z})\,t^a\right)\,\mathrm{e}^{-\mathrm{i}\int \mathrm{d}x^-\mathrm{d}^2\boldsymbol{x}\,\rho_L^a(x^-,\boldsymbol{x})\alpha_R^a(x^-,\boldsymbol{x})}\,. \qquad (6.23)$$

This is easily verified by expanding the x^-–ordered differential operator in the first line, acting with the derivatives on the exponential function, and finally reexponentiating the result. In the second line of Eq. (6.23), one recognizes the Wilson line describing the eikonal scattering of a left–moving quark with transverse coordinate $\boldsymbol{z}$:

$$V^\dagger(\boldsymbol{z})\;=\;\mathrm{P}\exp\left(ig\int \mathrm{d}x^-\alpha_R^a(x^-,\boldsymbol{z})\,t^a\right)\,. \qquad (6.24)$$

A similar identity holds, of course, for the anti–chronological Wilson line V which describes the scattering of an antiquark. By using such identities together with some integrations by parts, one can check that the S–matrix $S(\boldsymbol{z}_{i-1},\boldsymbol{z}_i)$ for the ith dipole in the projectile admits the following integral representation

$$S(\boldsymbol{z}_{i-1},\boldsymbol{z}_i)[\alpha_R] = \int \mathrm{D}[\rho_L]\,\mathrm{e}^{-\mathrm{i}\int \mathrm{d}x^-\mathrm{d}^2\boldsymbol{x}\,\rho_L^a\alpha_R^a}\,\frac{1}{N_c}\,\mathrm{tr}\left(V^\dagger(\boldsymbol{z}_{i-1})V(\boldsymbol{z}_i)\right)\left[\mathrm{i}\frac{\delta}{\delta\rho_L}\right]\delta[\rho_L]\,, \quad (6.25)$$

where $\rho_L \equiv \rho_L^a(x^-,\boldsymbol{x})$, $\delta[\rho_L]$ is the functional δ–function :

$$\delta[\rho_L]\;\equiv\;\prod_a\prod_{x^-}\prod_{\boldsymbol{x}}\delta\left(\rho_L^a(x^-,\boldsymbol{x})\right), \qquad (6.26)$$

and the Wilson lines inside the integrand must be read as 'derivative Wilson lines', i.e., all–orders, functional, differential operators obtained by replacing $\alpha_R^a \to \mathrm{i}\,\delta/\delta\rho_L$ in expressions like Eq. (6.24); e.g.,

$$V^\dagger(\boldsymbol{z})\left[\mathrm{i}\frac{\delta}{\delta\rho_L}\right]\;=\;\mathrm{P}\exp\left(-g\int \mathrm{d}x^-\frac{\delta}{\delta\rho_L^a(x^-,\boldsymbol{z})}\,t^a\right). \qquad (6.27)$$

The exponential factor $\exp\{-\mathrm{i}\int \rho_L^a\alpha_R^a\}$ is recognized as the *eikonal coupling* between the color charge of the left–mover and the color field created by the target. Then, Eq. (6.25) is suggestive of a 'color glass'–like averaging over the wavefunction of the ith dipole, with the following 'color glass' weight function for a single, left–moving, dipole [163, 224] :

$$W_i[\rho_L]\;\equiv\;\frac{1}{N_c}\,\mathrm{tr}\left(V^\dagger(\boldsymbol{z}_{i-1})V(\boldsymbol{z}_i)\right)\left[\mathrm{i}\frac{\delta}{\delta\rho_L}\right]\delta[\rho_L]\;\equiv\;D^\dagger(\boldsymbol{z}_{i-1},\boldsymbol{z}_i)\,\delta[\rho_L]. \qquad (6.28)$$

To gain more intuition with this formula, notice that, in QED, where $gt^a \to e$ with e the electric charge and there is no path–ordering, the analog of Eq. (6.28) would read

$$W_i[\rho_L] = \exp\left(-e\int \mathrm{d}x^-\left[\frac{\delta}{\delta\rho_L(x^-,\boldsymbol{z}_{i-1})}-\frac{\delta}{\delta\rho_L(x^-,\boldsymbol{z}_i)}\right]\right)\delta[\rho_L]$$

$$=\delta\left[\rho_L(x^-,\boldsymbol{x})-e\left(\delta^{(2)}(\boldsymbol{x}-\boldsymbol{z}_{i-1})-\delta^{(2)}(\boldsymbol{x}-\boldsymbol{z}_i)\right)\right], \qquad (6.29)$$

which expresses the fact that ρ_L is the (x^-–independent) electric charge density of a dipole made with a positive charge at $\boldsymbol{z}_{i-1}$ and a negative charge at $\boldsymbol{z}_i$. In QCD, the physical interpretation is complicated by the non–commutativity of the color matrices, but even there one can understand Eq. (6.28) as follows: The differential operator $\mathrm{i}\,\delta/\delta\rho_L(x^-, \boldsymbol{x})$ can be viewed as the creation operator for a gluon with the indicated coordinates. Then, $\delta[\rho_L]$ is the vacuum state for this operator (no gluon), and the derivative Wilson line in Eq. (6.27) is the creation operator for a fast loving quark, seen as the source of arbitrarily many small–x gluons, whose emission is treated in the eikonal approximation. Correspondingly, $D^\dagger(\boldsymbol{z}_{i-1}, \boldsymbol{z}_i)$ is the creation operator for a left–moving [28] dipole, and the meaning of Eq. (6.28) is then transparent: a one–dipole state is produced by acting with the dipole creation operator on the vacuum.

By using the representation in Eq. (6.25) for all the dipoles making up the onium (the projectile) together with the CGC representation for the average over the target, we are finally lead to the following rewriting of Eq. (6.20) [163] :

$$\mathcal{S}(\boldsymbol{x}_0, \boldsymbol{y}_0; \tau) = \int \mathrm{D}[\alpha_R] \int \mathrm{D}[\rho_L]\, \mathrm{e}^{-\mathrm{i} \int \mathrm{d}x^- \mathrm{d}^2 \boldsymbol{x}\, \rho_L^a(x^-, \boldsymbol{x}) \alpha_R^a(x^-, \boldsymbol{x})}\, W_{\tau_0}[\alpha_R]\, W^{\mathrm{D}}_{\tau-\tau_0}[\rho_L]\,, \quad (6.30)$$

where $W_{\tau_0}[\alpha_R]$ is the color–glass weight function for the target and $W^{\mathrm{D}}_{\tau-\tau_0}[\rho_L]$ is the corresponding weight function for the onium, and reads (cf. Eq. (6.28)) [225, 226]

$$W^{\mathrm{D}}_\tau[\rho] \;=\; \sum_{N=1}^\infty \int \mathrm{d}\Gamma_N\, P_N(\{\boldsymbol{z}_i\}|\tau) \prod_{i=1}^N D^\dagger(\boldsymbol{z}_{i-1}, \boldsymbol{z}_i)\, \delta[\rho]\,. \quad (6.31)$$

As anticipated, Eq. (6.30) depicts the elastic onium–hadron scattering as the collision between two color glasses, with a particularly simple form for the eikonal coupling between the respective color charges: a simple exponential, so like in QED. What is especially appealing about this formula is that it looks symmetric between the target and the projectile. (Recall that $-\nabla_\perp^2 \alpha_R^a = \rho_R^a$, so the integral giving the eikonal phase can be more symmetrically rewritten as $\int \rho_L^a \alpha_R^a \equiv \int \rho_L^a \frac{-1}{\nabla_\perp^2} \rho_R^a$.) However, given the approximations underlying its derivation, this symmetry is only formal: Eq. (6.30) is valid only in an asymmetric frame, where the projectile is dilute and the target is dense. It is presently not clear whether a similar factorization exists also for the more symmetric situation where *both* colliding systems are dense (see the discussion in [62, 63, 227–230]). But even for the *asymmetric* situation under consideration, the discussion in Sect. 6.3 shows that, in general, one cannot trust the JIMWLK equation for providing the weight function $W_{\tau_0}[\alpha_R]$ of the dense target. This is so because the JIMWLK evolution misses some correlations which are generated via gluon splitting in the dilute regime, and which are important for the scattering in the presence of unitarity corrections.

Remarkably, these correlations are explicitly encoded in the onium weight function in Eq. (6.31), which represents an explicit example of a color–glass description which transcends the JIMWLK evolution — albeit only in the dilute regime and for large–N_c. It is therefore interesting to understand the structure of this equation in more detail. Note

---

[28] Clearly, the corresponding formulæ for a right–moving dipole are obtained by simply replacing $x^- \to x^+$ in equations like (6.27).

first that, for the left–moving onium, the LC coordinates x^- and x^+ play the roles of 'time' and 'longitudinal' directions, respectively. The standard color glass description (as developed in the previous chapters) of such a dilute, left–moving, system would involve a color charge density ρ_L^a which is localized near $x^+ = 0$ (or $z = -t$) and is independent of x^-. The first condition is indeed satisfied for the color charge appearing in Eqs. (6.23)–(6.28), which should be actually viewed [29] as the coefficient of a δ–function in x^+; e.g., the eikonal phase in Eq. (6.23) should be understood as

$$\int \mathrm{d}^4 x \, \delta(x^+) \rho_L^a(x^-, \boldsymbol{x}) \, \alpha_R^a(x) \;=\; \int \mathrm{d}x^- \mathrm{d}^2\boldsymbol{x} \, \rho_L^a(x^-, \boldsymbol{x}) \, \alpha_R^a(x^+ = 0, x^-, \boldsymbol{x}). \qquad (6.32)$$

On the other hand, and unlike in the standard CGC description, the charge density ρ_L^a which enters Eqs. (6.23)–(6.28) is explicitly dependent upon the 'time' variable x^-. This dependence accounts for the color matrix ordering of the gluons exchanged in the *multiple scattering* between a (anti)quark from the onium and the target. Hence, this dependence is important within the interaction region, which is localized near $x^- = 0$ because of the Lorentz contraction of the right–moving target.

We have thus identified an important limitation of the CGC formalism: a purely 'glassy' description, which is strictly (LC) time–independent, cannot be used whenever one needs to take into account the multiple scattering of the individual constituents (quarks, gluons, dipoles) which compose the 'color glass'. This should not come as a surprise: We have seen already in Sect. 1.2.3 that the multiple scattering of a single–dipole projectile, which is a left–mover, probes the longitudinal (x^-) structure of the CGC target, which is a right–mover. Now, if the projectile itself is dense, then the elementary constituents of the target can undergo multiple scattering too, and thus probe the longitudinal (x^+) structure of the *projectile*; accordingly, the temporal (x^+) structure of the *target* is now important. The example of Eq. (6.31) together with our previous experience with JIMWLK equation suggests that a complete 'CGC' description should involve *two* types of Wilson lines: ordinary Wilson lines in the longitudinal direction (x^- for a right–mover) and derivative Wilson lines in the time direction (x^+). It appears to be quite difficult to construct a generalized CGC picture accommodating both types of Wilson lines and their high–energy evolution [60, 62, 63, 224, 228, 229]. We shall return to this issue in the next section.

But even when the multiple scattering can be ignored for the elementary constituents, and a genuine glass description is indeed possible, the corresponding description of the onium as emerging from Eq. (6.31) is still different from the JIMWLK evolution, as we explain now: With reference to Eq. (6.30), consider the situation where the right–moving target is dilute, so that we can restrict ourselves to the single–scattering approximation for the individual dipoles which compose the left–moving onium. This amounts to evaluating Eq. (6.25) in the two–gluon exchange approximation, which yields (cf. Eq. (3.73))

$$1 - \frac{g^2}{4N_c} \left(\alpha_R^a(\boldsymbol{z}_{i-1}) - \alpha_R^a(\boldsymbol{z}_i) \right)^2 = \int \mathrm{D}[\rho_L] \, \mathrm{e}^{-\mathrm{i}\int \mathrm{d}^2\boldsymbol{x} \, \rho_L^a \alpha_R^a} \, D_0^\dagger(\boldsymbol{z}_{i-1}, \boldsymbol{z}_i) \, \delta[\rho_L], \qquad (6.33)$$

---

[29] Alternatively, and equivalently, the functional derivative in equations like (6.27) can be replaced as $\delta/\delta\rho_L^a(x^-, \boldsymbol{x}) \to \delta/\delta\rho_L^a(x^+ = 0, x^-, \boldsymbol{x})$, and then the integral giving the eikonal phase is understood in the four–dimensional sense: $\int \mathrm{d}^4 x \, \rho_L^a(x) \alpha_R^a(x)$.

where now the fields ρ_L and α_R depend upon the transverse coordinate $\boldsymbol{x}$ alone, and

$$D_0^\dagger(\boldsymbol{x}, \boldsymbol{y}) \equiv 1 + \frac{g^2}{4N_c}\left(\frac{\delta}{\delta\rho_L^a(\boldsymbol{x})} - \frac{\delta}{\delta\rho_L^a(\boldsymbol{y})}\right)^2. \tag{6.34}$$

That is, any reference to the coordinate x^- has disappeared, as expected in a single–scattering approximation. Physics–wise, we can make the following identifications :

$$\alpha_R^a(\boldsymbol{x}) \equiv \int \mathrm{d}x^-\, \alpha_R^a(x^+ \simeq 0, x^-, \boldsymbol{x})\,, \qquad \rho_L^a(\boldsymbol{x}) \equiv \int \mathrm{d}x^+\, \rho_L^a(x^+, x^- \simeq 0, \boldsymbol{x})\,. \tag{6.35}$$

The operator $D_0^\dagger$ in Eq. (6.34) is recognized as the expansion of the dipole creation operator $D^\dagger$ introduced in Eq. (6.28) to second–order in the functional derivatives. That is, a 'dipole' created by $D_0^\dagger$ is a color source which can emit, or absorb, exactly two gluons.

Corresponding to Eq. (6.33), we obtain the following, "Dipole Model", approximation for the onium weight function [163] :

$$W_\tau^{\mathrm{DM}}[\rho] = \sum_{N=1}^\infty \int \mathrm{d}\Gamma_N\, P_N(\{\boldsymbol{z}_i\}|\tau) \prod_{i=1}^N D_0^\dagger(\boldsymbol{z}_{i-1}, \boldsymbol{z}_i)\, \delta[\rho]\,, \tag{6.36}$$

which is sufficient for all the physical situations in which one can limit oneself to the exchange of two gluons per dipole. As we shall later argue, this is indeed the most interesting situation at high energy (and large N_c).

As compared to Eq. (6.31), the dependence upon x^- has disappeared from Eq. (6.36), which therefore provides a truly 'glassy' description, as anticipated. Yet, this description is not the one that would be produced by the JIMWLK evolution of the original dipole. In fact, since the left–moving system is dilute, the corresponding JIMWLK evolution reduces to the BFKL evolution, which rather implies

$$W_\tau^{\mathrm{BFKL}}[\rho] = \left\{1 + \frac{g^2}{4N_c}\int \mathrm{d}^2\boldsymbol{x}\,\mathrm{d}^2\boldsymbol{y}\, n_\tau(\boldsymbol{x}, \boldsymbol{y})\left(\frac{\delta}{\delta\rho^a(\boldsymbol{x})} - \frac{\delta}{\delta\rho^a(\boldsymbol{y})}\right)^2\right\}\delta[\rho]\,, \tag{6.37}$$

where n_τ is the average dipole number density, as given by the solution to the BFKL equation (6.14). By using the latter, one can indeed check that Eq. (6.37) obeys the low–density (or BFKL) version of the JIMWLK evolution, Eq. (3.71), at large N_c.

Eq. (6.37) can be formally obtained from Eq. (6.36) by keeping only the terms of $\mathcal{O}(\}^\in)$ (or, equivalently, of second order in the functional derivatives) in the product $\prod_{i=1}^N D_0^\dagger(\boldsymbol{z}_{i-1}, \boldsymbol{z}_i)$. Hence, out of the dynamical information *a priori* encoded in Eq. (6.36), its BFKL approximation (6.37) retains just the information about the average dipole number density. By contrast, the more general weight function (6.36) also encompasses the higher many–body correlations $n_\tau^{(\kappa)}$ with $\kappa \geq 2$, that is, it encodes the complete information about the statistical distribution of the dipoles in transverse space. Thus, in applications to scattering, Eq. (6.37) can be used when only one of the dipoles composing the onium can scatter at a time (i.e., in the single–scattering approximation *for the onium as a whole*), whereas Eq. (6.36) is also appropriate in more general situations, where two or more dipoles can simultaneously scatter (with each such a dipole exchanging two gluons). Physical situations where this difference is relevant will be presented later, in Sects. 6.5 and 7.

Still another way to visualize the difference between the dipole picture and the dilute (or BFKL) version of the JIMWLK equation, is to construct the Dipole Model analog of the JIMWLK Hamiltonian. This is the Hamiltonian $H_{\rm DM}$ which generates the dipole–picture evolution of the onium weight function (6.36), via the following equation:

$$\frac{\partial}{\partial\tau}\,W^{\rm DM}_\tau \;=\; -\,H_{\rm DM}\,W^{\rm DM}_\tau\,, \tag{6.38}$$

where the τ–derivative in the l.h.s. is independently determined by the Master equation (6.9) for $P_N(\tau)$. The appropriate Hamiltonian is [50, 163]

$$H_{\rm DM} \;=\; -\,\frac{\bar\alpha}{2\pi}\int_{xyz}\mathcal{M}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})\Big[-D_0^\dagger(\boldsymbol{x},\boldsymbol{y}) + D_0^\dagger(\boldsymbol{x},\boldsymbol{z})D_0^\dagger(\boldsymbol{z},\boldsymbol{y})\Big]D_0(\boldsymbol{x},\boldsymbol{y}), \tag{6.39}$$

where $D_0^\dagger$ is the dipole creation operator introduced in Eq. (6.34) and [30]

$$D_0(\boldsymbol{x},\boldsymbol{y}) \;\equiv\; -\frac{1}{g^2N_c}\,\rho^a(\boldsymbol{x})\rho^a(\boldsymbol{y}) \quad \text{for} \quad \boldsymbol{x}\neq\boldsymbol{y}. \tag{6.40}$$

is the corresponding annihilation operator, as it can be seen from the following commutation relation, which holds at large N_c (with $\delta_{\boldsymbol{ux}}\equiv\delta^{(2)}(\boldsymbol{u}-\boldsymbol{x})$) :

$$[\,D_0(\boldsymbol{x},\boldsymbol{y}),D_0^\dagger(\boldsymbol{u},\boldsymbol{v})\,] \;=\; \frac{1}{2}\Big(\delta_{\boldsymbol{ux}}\delta_{\boldsymbol{vy}} + \delta_{\boldsymbol{uy}}\delta_{\boldsymbol{vx}}\Big). \tag{6.41}$$

By using this commutation relation, it is straightforward to verify that the action of $H_{\rm DM}$ in the r.h.s. of Eq. (6.38) is indeed consistent with the prediction of the Master equation (6.9) for the variation $\partial W^{\rm DM}_\tau/\partial\tau$ in the l.h.s.

The structure of the Hamiltonian (6.39) is physically transparent: when this is acting on the onium weight function (6.36), a dipole $(\boldsymbol{x},\boldsymbol{y})$ is removed from the original configuration and replaced by a new pair of dipoles $(\boldsymbol{x},\boldsymbol{z})$ and $(\boldsymbol{z},\boldsymbol{y})$, with the right probability density for this process to be recognized as a dipole splitting. In addition, there is a negative, 'virtual', term which accounts for the decrease in the probability that the original configuration remain unchanged after one step in the evolution.

The Hamiltonian (6.39) can be decomposed as $H_{\rm DM} = H_0 + H_{1\to 2}$, with H_0 describing the BFKL evolution of the average color charge density (or dipole number) [163]:

$$H_0 = -\frac{1}{2N_c^2}\frac{\bar\alpha}{2\pi}\int_{x,y,z}\mathcal{M}(\S,\dagger,\ddagger)\left[\frac{\delta}{\delta\rho^\perp(\S)} - \frac{\delta}{\delta\rho^\perp(\ddagger)}\right]\left[\frac{\delta}{\delta\rho^\perp(\ddagger)} - \frac{\delta}{\delta\rho^\perp(\dagger)}\right]\rho^\perp(\S)\rho^\perp(\dagger)\,, \tag{6.42}$$

HBFKLI

---

[30] The expression of $D_0(\boldsymbol{x},\boldsymbol{y})$ in the limit $\boldsymbol{x}\to\boldsymbol{y}$ is not needed, because the dipole kernel in Eq. (6.39) vanishes in that limit.

and $H_{1\to 2}$ describing the new correlations induced via $1 \to 2$ dipole splitting [50]:

$$H_{1\to 2} = \frac{g^2}{16 N_c^3} \frac{\bar{\alpha}}{2\pi} \int_{\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}} \mathcal{M}(\S, \dagger, \ddagger) \left[\frac{\delta}{\delta \rho^\dashv(\S)} - \frac{\delta}{\delta \rho^\dashv(\ddagger)} \right]^\in \left[\frac{\delta}{\delta \rho^\llcorner(\ddagger)} - \frac{\delta}{\delta \rho^\llcorner(\dagger)} \right]^\in \rho^\lrcorner(\S)\rho^\lrcorner(\dagger).$$

(6.43) HMSW

Whereas H_0 can be recognized [61] as the large–N_c version of the low–density JIMWLK Hamiltonian in Eq. (3.71), the second piece $H_{1\to 2}$ clearly goes beyond the JIMWLK formalism, as it involves four functional derivatives w.r.t. ρ. Recalling the interpretation of the derivative $\delta/\delta\rho$ as a creation operator for one gluon, it becomes clear that $H_{1\to 2}$ represents the vertex for $2 \to 4$ gluon splitting at large N_c.

To summarize, unlike the JIMWLK picture, which goes beyond the BFKL evolution by including the non–linear, gluon merging, effects at high density, the dipole picture deviates from the BFKL approximation by including the correlations associated with gluon splitting in the dilute regime. As we shall later argue, all such effects are in fact essential for the problem of high–energy scattering.

6.5   Onium–onium scattering

In this section we shall consider a few applications of the formalism that we have just developed — the (generalized) color glass description of an onium — to the problem of the elastic scattering between two dilute systems (dipoles or onia). This problem corresponds to low or intermediate energies, where the saturation effects remain negligible in both colliding systems (namely, this requires $\tau \ll \tau_c$, cf. Eq. (6.12)). This situation is interesting in that it allows us to study the effects of dipole correlations on the scattering amplitude in a simple setting, for which the formalism has already been developed. (This extension of this discussion towards higher energies will be the subject of Sect. 7.) Moreover, as we shall see in Sect. 6.5.4, when used in the right kinematics and pushed towards its limits, this set–up also allows us to study the onset of the unitarity corrections while staying within the dipole picture.

6.5.1   Dipole-onium scattering

We start with the simplest case, that of the elastic scattering between two elementary dipoles, with transverse coordinates $(\boldsymbol{x}_0, \boldsymbol{y}_0)$ for the left–mover and, respectively, $(\boldsymbol{x}, \boldsymbol{y})$ for the right–mover. To lowest order in perturbation theory, the scattering proceeds via the exchange of two gluons and can be computed according to Eq. (6.30) with the following weight functions for the incoming dipoles (cf. Eq. (6.28)) :

$$W_L[\rho_L] = D_{0,L}^\dagger(\boldsymbol{x}_0, \boldsymbol{y}_0)\, \delta[\rho_L], \qquad W_R[\rho_R] = D_{0,R}^\dagger(\boldsymbol{x}, \boldsymbol{y})\, \delta[\rho_R],$$

(6.44)

with $D_0^\dagger$ given by Eq. (6.34) (now applied to both dipoles). The functional integral over ρ_L is computed as in Eq. (6.33), so we are left with

181

$$T_0(\boldsymbol{x}, \boldsymbol{y}|\boldsymbol{x}_0, \boldsymbol{y}_0) = \left(\frac{g^2}{4N_c}\right)^2 \int \mathrm{D}[\rho_R]\left(\alpha_R^a(\boldsymbol{x}_0) - \alpha_R^a(\boldsymbol{y}_0)\right)^2 \left(\frac{\delta}{\delta\rho_R^b(\boldsymbol{x})} - \frac{\delta}{\delta\rho_R^b(\boldsymbol{y})}\right)^2 \delta[\rho_R],$$

$$(6.45)$$

with ρ_R and α_R related through $-\nabla_\perp^2 \alpha_R^a = \rho_R^a$. After some integrations by parts and using $C_F \approx N_c/2$, one finds the standard result (see, e.g., [189])

$$T_0(\boldsymbol{x}, \boldsymbol{y}|\boldsymbol{x}_0, \boldsymbol{y}_0) = \frac{g^4}{8}\Big[\Delta(\boldsymbol{x} - \boldsymbol{x}_0) - \Delta(\boldsymbol{y} - \boldsymbol{x}_0) - \Delta(\boldsymbol{x} - \boldsymbol{y}_0) + \Delta(\boldsymbol{y} - \boldsymbol{y}_0)\Big]^2$$
$$\equiv \alpha^2 \mathcal{A}_\prime(\S, \dagger|\S_\prime, \dagger_\prime),$$

$$(6.46)$$

where $\Delta(\boldsymbol{x} - \boldsymbol{y})$ denotes the Coulomb propagator, cf. Eq. (2.16), and

$$\mathcal{A}_\prime(\S, \dagger|\S_\prime, \dagger_\prime) \equiv \frac{\infty}{\forall}\ln^\epsilon\left[\frac{(\S - \dagger_\prime)^\epsilon(\dagger - \S_\prime)^\epsilon}{(\S - \S_\prime)^\epsilon(\dagger - \dagger_\prime)^\epsilon}\right].$$

$$(6.47)$$

The four terms inside the square brackets in Eq. (6.46) correspond to the four possible attachments of a gluon propagator in between a (anti)quark line in the first dipole and a corresponding line in the second dipole. Thus, Eq. (6.46) describes the exchange of two gluons, as anticipated.

SECT_DIPO

6.5.2 Dipole–onium scattering

Consider now increasing the rapidity of the left–mover, which consequently evolves into an onium, whose 'color glass' weight function is generally given by Eq. (6.31). But for the purpose of computing the dipole–onium scattering amplitude to lowest order, it is enough to use the BFKL weight function in Eq. (6.37) (since the onium scatters only once). Then one easily finds

$$\langle T(\boldsymbol{x}, \boldsymbol{y}|\boldsymbol{x}_0, \boldsymbol{y}_0)\rangle_\tau = \alpha^2 \int \mathrm{d}^2\boldsymbol{u}\,\mathrm{d}^2\boldsymbol{v}\,\mathcal{A}_\prime(\S, \dagger|\sqcap, \sqsubseteq)\,\backslash_\tau(\sqcap, \sqsubseteq),$$

$$(6.48)$$

Tconv

where the brackets on T refer to the average over the onium wavefunction. Eq. (6.48) tells us that the collision proceeds via a two–gluon exchange between the external dipole and any of the dipoles internal to the onium. In what follows, we shall often denote a dipole-onium amplitude simply as $T(\boldsymbol{x}, \boldsymbol{y})$; that is, the coordinates $(\boldsymbol{x}_0, \boldsymbol{y}_0)$ of the dipole which has initiated the onium will be kept implicit.

SECT_2DIPO

6.5.3 Two dipoles–onium scattering

For the same left–moving onium as above, consider now the situation where the right–moving system is made with two dipoles, with coordinates $(\boldsymbol{x}_1, \boldsymbol{y}_1)$ and $(\boldsymbol{x}_2, \boldsymbol{y}_2)$, respectively. We are interested in the amplitude $\langle T^{(2)}(\boldsymbol{x}_1, \boldsymbol{y}_1; \boldsymbol{x}_2, \boldsymbol{y}_2)\rangle_\tau \equiv \langle T(\boldsymbol{x}_1, \boldsymbol{y}_1)T(\boldsymbol{x}_2, \boldsymbol{y}_2)\rangle_\tau$ for the *simultaneous* scattering of the two external dipoles. This quantity is interesting in that it enters as the 'non–linear term' (or the 'unitarity correction') in the Balitsky equation (1.64) for the evolution of the single–dipole amplitude. Previously, this quantity has been generally evaluated in the mean field approximation $\langle TT\rangle \approx \langle T\rangle\langle T\rangle$, which neglects

the correlations in the target wavefunction, and under which Eq. (1.64) reduces to the Kovchegov equation (1.67). Here, we would like to study the *violation* of this MFA by the correlations built in the onium wavefunction by the high–energy evolution.

To lowest order in perturbation theory, meaning $\mathcal{O}(\alpha^4)$, each of the external dipoles scatters only once, via two–gluon exchange, so we have

$$\langle T^{(2)}(\boldsymbol{x}_1, \boldsymbol{y}_1; \boldsymbol{x}_2, \boldsymbol{y}_2)\rangle_\tau \simeq \left(\frac{g^2}{4N_c}\right)^2 \left\langle \left(\alpha_L^a(\boldsymbol{x}_1) - \alpha_L^a(\boldsymbol{y}_1)\right)^2 \left(\alpha_L^b(\boldsymbol{x}_2) - \alpha_L^b(\boldsymbol{y}_2)\right)^2 \right\rangle_\tau, \text{ (6.49)}$$

where α_L is the color field created by the onium, with $-\nabla_\perp^2 \alpha_L^a = \rho_L^a$. (Note that, in LC notations, this field is an A^- component, since the onium is a left mover.) The four color fields which appear inside the matrix element in the r.h.s. can be attached to either two *different* dipoles inside the onium (which then separately undergo single scattering), or to a *same* internal dipole (which therefore undergoes double scattering). *A priori*, both types of contributions are of the same order in α and N_c. Yet, as we shall demonstrate in what follows, the second type of contribution, involving the multiple scattering of *individual* dipoles, has a slower rise with the energy and thus can be neglected in the high–energy regime where the unitarity corrections are important [225, 226].

Note first that, since the onium as a whole undergoes a double scattering, the use of the BFKL approximation (6.37) for its weight function would be totally inappropriate. One should rather use the more general expressions in Eq. (6.31) or (6.36) which allow for multiple scattering. It is convenient to separately discuss the two types of contributions alluded to above:

(i) *Scattering off two different internal dipoles.* In that case, each of the internal dipoles partaking in the collision undergoes a single scattering, so one can use the approximation (6.36) for the onium weight function. Then the expectation value in Eq. (6.49) is easily computed via integration by parts (similarly to Eq. (6.45)), to yield

$$\left\langle T^{(2)}(\boldsymbol{x}_1, \boldsymbol{y}_1; \boldsymbol{x}_2, \boldsymbol{y}_2)\right\rangle_\tau^{2-\text{dipoles}} = \alpha^4 \sum_{N=1}^\infty \int \mathrm{d}\Gamma_N \, P_N(\{\boldsymbol{z}_i\}|\tau)$$
$$\sum_{1 \le i \ne j \le N} \mathcal{A}_{\prime}(\S_\infty, \dagger_\infty|\ddagger\rangle_{-\infty}, \ddagger\rangle) \, \mathcal{A}_{\prime}(\S_\in, \dagger_\in|\ddagger\rangle_{-\infty}, \ddagger_|). \quad \text{(6.50)}$$

We have also used the large–N_c approximation to avoid crossed exchanges in which one dipole exchanges two gluons with two different dipoles. After also recalling the definition (6.15) of the average dipole pair density, the above result is finally rewritten as

$$\left\langle T^{(2)}(\boldsymbol{x}_1, \boldsymbol{y}_1; \boldsymbol{x}_2, \boldsymbol{y}_2)\right\rangle_\tau^{2-\text{dipoles}} = \alpha^4 \int_{\boldsymbol{u}_i, \boldsymbol{v}_i} \mathcal{A}_{\prime}(\S_\infty, \dagger_\infty|\sqcap_\infty, \sqsubseteq_\infty) \mathcal{A}_{\prime}(\S_\in, \dagger_\in|\sqcap_\in, \sqsubseteq_\in) \, \backslash_\tau^{(\in)}(\sqcap_\infty, \sqsubseteq_\infty; \sqcap_\in, \sqsubseteq_\in),$$

$$\text{(6.51)} \qquad \boxed{\text{T2con\hspace{-0.3em}|}}$$

with a transparent physical interpretation: the two external dipoles scatter off two different internal ones, via two–gluon exchange per dipole–dipole scattering.

(ii) *Double scattering with only one internal dipole.* In this case, the calculation is more complicated (and, *a priori*, even ambiguous) because of the need to use the complete

183

expression, Eq. (6.31), for the onium weight function. An ambiguity arises because the functional derivatives which enter this expression, cf. Eq. (6.27), involve the color charge density $\rho_L^a(x^-, \boldsymbol{x})$ at different 'times' x^-, whereas in Eq. (6.49) the x^- coordinate has been suppressed: roughly speaking, the field α_L^a in that equation should be evaluated near $x^- = 0$, cf. Eq. (6.35), but its detailed structure in x^- cannot be better specified in the present, high–energy, approximation. As discussed in Refs. [225, 226], this ambiguity can be fixed by interpreting α_L as the field at the upper end of the interaction region in x^- : $\alpha_L^a(\boldsymbol{x}) \equiv \alpha_L^a(x^- \to \infty, \boldsymbol{x}) = \int_{\boldsymbol{u}} \Delta(\boldsymbol{x} - \boldsymbol{u})\rho_L^a(\infty, \boldsymbol{u})$, where 'infinity' truly means 'well above the support of the interaction region'. The calculation can now be done, by using (below, $V^\dagger(\boldsymbol{z})$ is the derivative Wilson line in Eq. (6.27))

$$ V^\dagger(\boldsymbol{z})\, \rho_L^a(\infty, \boldsymbol{x}) \equiv \mathrm{Pe}^{-g \int \mathrm{d}x^- t^b \frac{\delta}{\delta \rho_L^a(x^-, \boldsymbol{z})}}\, \rho_L^b(\infty, \boldsymbol{x}) = -gt^a \delta^{(2)}(\boldsymbol{x} - \boldsymbol{z})\, V^\dagger(\boldsymbol{z}), \qquad (6.52) $$

and similarly for $V(\boldsymbol{z})$. Note the similitude between the above equation and Eq. (3.31): the color charge density at infinity acts on the derivative Wilson lines in the same way as the functional derivative w.r.t. the field α_τ at infinity acts on the normal Wilson lines — that is, as an infinitesimal color rotation at the end point. By repeatedly using such identities, one can integrate by parts the derivative Wilson lines implicit in Eq. (6.49), and thus compute the contribution to $\langle T^{(2)} \rangle$ due to the double scattering of one of the internal dipoles. One thus finds (at large N_c) [226]

$$ \left\langle T^{(2)}(\boldsymbol{x}_1, \boldsymbol{y}_1; \boldsymbol{x}_2, \boldsymbol{y}_2) \right\rangle_\tau^{1-\mathrm{dipole}} = \alpha^4 \int_{\boldsymbol{u},\boldsymbol{v}} \mathcal{A}_{/}(\S_\infty, \dagger_\infty | \sqcap, \sqsubseteq) \mathcal{A}_{/}(\S_\in, \dagger_\in | \sqcap, \sqsubseteq) \setminus_\tau (\sqcap, \sqsubseteq), \qquad (6.53) $$

which has a very simple physical interpretation: the two external dipoles scatter *independently of each other* off a same internal dipole, chosen according to the density n_τ. This result is in fact natural at large N_c: non–planar diagrams, in which the gluons exchanged by the external dipoles cross with each other, are relatively suppressed at large N_c.

To summarize, to lowest order in α and $1/N_c$, the two–dipole amplitude $\langle T^{(2)} \rangle_\tau$ receives the two contributions shown in Eqs. (6.51) and (6.53). At relatively low energies, $\bar{\alpha}\tau \lesssim 1$, and for a target onium which has started as a single dipole at $\tau = 0$, the second contribution, which involves the single dipole density, dominates over the first one. But the situation rapidly changes with increasing energy, since $n_\tau^{(2)}$ rises much faster with τ than n_τ (cf. the discussion in Sect. 6.2). Thus, the first contribution, which describes the scattering off different target dipoles, becomes the dominant one as soon as $\bar{\alpha}\tau \gtrsim 1$, and it remains so in the approach towards unitrarization. A different argument leading to the same conclusion can be found in Ref. [225].

It should be further emphasized that both contributions computed above are in strong disagreement with the mean field approximation $\langle TT \rangle_{\mathrm{MFA}} = \langle T \rangle \langle T \rangle$, with $\langle T \rangle$ given (to the requested accuracy) by Eq. (6.48). This is obvious for Eq. (6.53), but it is also true for Eq. (6.51) since the dipole pair density $n_\tau^{(2)}(u_1, u_2)$ is predominantly built via gluon–number fluctuations at low density (cf. Sect. 6.2) and hence is very different in general from its mean field estimate, or 'disconnected piece', $n_\tau(u_1)n_\tau(u_2)$. A simple way to appreciate this difference is to consider the zero–transverse–dimensions version of the equations for the dipole densities, which read (compare to Eqs. (6.14) and (6.18))

184

$$\frac{\mathrm{d}n^{(1)}}{\mathrm{d}\tau} = \alpha\, n^{(1)}, \qquad \frac{\mathrm{d}n^{(2)}}{\mathrm{d}\tau} = 2\alpha\left[n^{(2)} + n^{(1)}\right], \quad \dots \tag{6.54}$$

where the dots stay for the higher equations in the hierarchy, and the local splitting rate α replaces the non–local dipole kernel of QCD. The solutions corresponding to the initial conditions $n^{(1)}(0) = n_0$ and $n^{(2)}(0) = 0$ (cf. Eq. (6.16)) read

$$\begin{aligned} n^{(1)}(\tau) &= n_0 \exp(\alpha\tau), \\ n^{(2)}(\tau) &= 2n_0 \exp(2\alpha\tau) - 2n_0 \exp(\alpha\tau), \ \dots \end{aligned} \tag{6.55}$$ toyso

They confirm the fact that, so long as $\alpha\tau \lesssim 1$, $n^{(1)}$ dominates over $n^{(2)}$ in the r.h.s. of the second equation (6.54), and thus it is the driving force for the growth in the number of pairs. For $\alpha\tau \gtrsim 1$, $n^{(2)}$ starts to dominate over $n^{(1)}$, as expected. However, the ratio

$$\frac{n^{(2)}(\tau)}{[n^{(1)}(\tau)]^2} = 2 - 2\exp(-\alpha\tau), \tag{6.56}$$ toyra

approaches 2 at large τ, and thus differs from the naive prediction $n^{(2)}(\tau) \approx [n^{(1)}(\tau)]^2$ of the MFA. In fact, one can easily check that the only way to satisfy the mean–field behaviour is to enforce it already in the initial conditions, by assuming $n_0 \gg 1$, $n_0^{(2)} \approx n_0^2$, etc. This might suggest that, in QCD, the MFA would be a good approximation at any τ provided one starts the evolution with a very large nucleus at $\tau = 0$. But this is actually not so: the QCD evolution is non–local in transverse momenta (or dipole sizes) and, independently of the initial conditions, it develops a tail at large momenta where the density is low and the MFA breaks down completely.

6.5.4 Onium–onium scattering near the unitarity limit

As a final application of the factorization formula (6.30), we shall now consider the elastic onium–onium scattering in the center–of–mass (COM) frame. The COM kinematics is advantageous in that it allows one to study the onset of unitarity corrections while neglecting saturation effects in any of the incoming wavefunctions [43]: Indeed, for $\tau_0 = \tau - \tau_0 = \tau/2$, the dipole picture applies (separately for the left–mover and the right–mover) so long as $\tau/2 \ll \tau_c$, cf. Eq. (6.12); this condition leaves a parametrically large window at $\tau_c \lesssim \tau \ll 2\tau_c$ within which both hadronic systems are non–saturated, yet their multiple scattering is important and leads to unitarization.

To compute the elastic S–matrix within this kinematical window, one can neglect the multiple scattering of the *individual* dipoles from both onia: indeed, as explained in the previous subsection, when κ dipoles simultaneously scatter off a (non–saturated) onium, then for $\bar{\alpha}\tau \gtrsim 1$ they predominantly scatter off κ *different* dipoles from that onium. (In Sect. 6.5.3 we have taken $\kappa = 2$, but the generalization to arbitrary κ is straightforward.) Hence, the onium–onium scattering in the kinematics of interest should be dominated by the scattering between κ dipoles from one onium against κ dipoles from the other onium, for any $\kappa \geq 1$ and arbitrary transverse configurations for the two sets of colliding dipoles. This is precisely the picture which emerges from Eq. (6.30), as we show now.

Namely, after using Eq. (6.36) for the weight functions of both incoming onia, Eq. (6.30) can be rewritten as

$$\mathcal{S}(\tau) = \sum_{N=1}^{\infty} \int \mathrm{d}\Gamma_N \, P_N(\tau/2) \sum_{N'=1}^{\infty} \int \mathrm{d}\Gamma_{N'} \, P_{N'}(\tau/2) \, S_{N \times N'}. \tag{6.57}$$

where $S_{N \times N'}$ is the S–matrix for the elastic scattering between two given dipole configurations, which involve N and N' dipoles, respectively:

$$S_{N \times N'} = \int \mathrm{D}[\alpha_R] \int \mathrm{D}[\rho_L] \, \mathrm{e}^{-\mathrm{i} \int \mathrm{d}^2 \boldsymbol{x} \, \rho_L^a(\boldsymbol{x}) \alpha_R^a(\boldsymbol{x})} \, W_N[\rho_L] \, W_{N'}[\rho_R] \, ,$$

$$W_N[\rho_L] \equiv \prod_{i=1}^{N} D_{0,L}^{\dagger}(\boldsymbol{z}_{i-1}, \boldsymbol{z}_i) \, \delta[\rho_L] \, , \qquad W_{N'}[\rho_R] \equiv \prod_{j=1}^{N'} D_{0,R}^{\dagger}(\boldsymbol{v}_{j-1}, \boldsymbol{v}_j) \, \delta[\rho_R]. \tag{6.58}$$

Via integrations by parts like in Eq. (6.33), it is straightforward to work out one of the functional integrals in Eq. (6.58), say, that over ρ_L. This yields

$$S_{N \times N'} = \int \mathrm{D}[\alpha_R] \prod_{i=1}^{N} \left[1 - \frac{g^2}{4N_c} \left(\alpha_R^a(\boldsymbol{z}_{i-1}) - \alpha_R^a(\boldsymbol{z}_i) \right)^2 \right] W_{N'}[\rho_R] \, . \tag{6.59}$$

It is furthermore easy to extract the single–scattering contribution to Eq. (6.59), that is, the piece of $\mathcal{O}(\alpha^2)$: by using Eqs. (6.45)–(6.47), one finds

$$S_{N \times N'}^{\mathrm{one-scatt}} = 1 - \alpha^2 \sum_{i=1}^{N} \sum_{j=1}^{N'} \mathcal{A}_{\prime}(\ddagger\rangle_{-\infty}, \ddagger\rangle|\sqsubseteq_{|-\infty}, \sqsubseteq_{|}) \, , \tag{6.60}$$

which immediately implies

$$\langle T \rangle_{1\mathbb{P}} = \alpha^2 \int_{\boldsymbol{uv}, \boldsymbol{xy}} \mathcal{A}_{\prime}(\S, \dagger|\sqcap, \sqsubseteq) \setminus_{\tau/\in}(\S, \dagger) \setminus_{\tau/\in}(\sqcap, \sqsubseteq). \tag{6.61}$$

This is the 'one–pomeron exchange' approximation and describes the scattering between two dipoles — one from each onium. Since the dipole densities rise very fast with τ in both onia, $n_{\tau/2} \sim \mathrm{e}^{\omega_0 \tau/2}$, it is clear that $\langle T \rangle_{1\mathbb{P}} \sim \alpha^2 \mathrm{e}^{\omega_0 \tau}$ will violate unitarity for $\tau \gtrsim \tau_c$. This means that, for such high energies, we need a better estimate for Eq. (6.59), which includes the unitarity corrections associated with multiple scattering. It turns out that, to the order of interest, Eq. (6.59) is equivalent to the following, *Glauber*, formula :

$$S_{N \times N'} = \exp\left\{ -\alpha^2 \sum_{i=1}^{N} \sum_{j=1}^{N'} \mathcal{A}_{\prime}(\ddagger\rangle_{-\infty}, \ddagger\rangle|\sqsubseteq_{|-\infty}, \sqsubseteq_{|}) \right\} , \tag{6.62}$$

on which unitarity is manifest: $S_{N \times N'} \leq 1$. This formula implies that the successive scatterings are independent from each other (in the *event–by–event* description, i.e., for given configurations of dipoles in the two onia), so that the multiple–scattering series simply exponentiates. To demonstrate this exponentiation, it is convenient to introduce the simpler notations:

$$t_i \equiv \frac{g^2}{4N_c}\left(\alpha_R^a(\boldsymbol{z}_{i-1}) - \alpha_R^a(\boldsymbol{z}_i)\right)^2, \qquad d_j \equiv \frac{g^2}{4N_c}\left(\frac{\delta}{\delta\rho_R^c(\boldsymbol{v}_{j-1})} - \frac{\delta}{\delta\rho_R^c(\boldsymbol{v}_j)}\right)^2, \qquad (6.63)$$

in terms of which Eq. (6.59) is compactly rewritten as

$$S_{N\times N'} = \int \mathrm{D}[\rho_R] \prod_{i=1}^{N}\left[1 - t_i\right] \prod_{j=1}^{N'}\left[1 + d_j\right] \delta[\rho_R] \equiv \left\langle \prod_{i=1}^{N}\left[1 - t_i\right] \prod_{j=1}^{N'}\left[1 + d_j\right] \right\rangle. \qquad (6.64)$$

When expanding the two products above, the only non–trivial terms are those involving an equal number of factors t and d. We then use the following 'contraction' (in compact notations; see Eqs. (6.45)–(6.47)) :

$$\langle t_i d_j \rangle = T_0(i|j) \equiv \alpha^2 \mathcal{A}_{\prime}(\rangle||) \qquad (6.65)$$

together with factorization properties valid at large N_c, e.g.,

$$\langle t_i t_m d_j d_n \rangle \approx T_0(i|j) T_0(m|n) + T_0(i|n) T_0(m|j), \qquad \text{etc.} \qquad (6.66)$$

to deduce that:

$$S_{N\times N'} = 1 - \alpha^2 \sum_{i=1}^{N} \sum_{j=1}^{N'} \mathcal{A}_{\prime}(\rangle||) + \frac{\alpha^{\triangle}}{\in} \sum_{\infty \leq \rangle \neq \Updownarrow \leq \mathcal{N}} \sum_{\infty \leq |\neq \backslash \leq \mathcal{N}'} \mathcal{A}_{\prime}(\rangle||)\, \mathcal{A}_{\prime}(\Updownarrow|\backslash) + \cdots \quad (6.67)$$

This is not exactly the same as the expansion of the exponential (6.62), because, e.g., of the restrictions $i \neq m$ and $j \neq n$ in the second–order term in Eq. (6.59). However, at high energy and for the typical configurations, both N and N' are very large, of order $e^{\omega_0 \tau/2}$, so the difference between Eqs. (6.59) and (6.62) is indeed negligible.

Note, however, that the Glauber exponentiation holds only *configuration by configuration*, but not also for the *average S–matrix*, Eq. (6.57). In turns out that, in the region where the unitarity corrections start to be important (i.e., for $\tau \gtrsim \tau_c$), the average amplitude is very different [43] from the naive exponentiation of the amplitude for 'one–pomeron exchange', Eq. (6.61). This has been clearly seen in the numerical calculations of the dipole picture by Salam [44, 45] : one has thus found that the S–matrix (6.57) has a rather slow approach towards the unitarity limit $S = 0$, namely, of the form

$$\mathcal{S}(\tau) \sim \exp\left\{-\kappa \bar{\alpha}_s^2 \tau^2\right\} \qquad \text{with} \qquad \kappa \approx 0.72, \qquad (6.68)$$

which should be contrasted with the much faster decay predicted by a simple exponentiation of the BFKL amplitude for single–scattering:

$$\mathcal{S}_{\text{Glauber}}(\tau) \sim \exp\left\{-\langle T\rangle_{1\mathbb{P}}\right\} \qquad \text{with} \qquad \langle T\rangle_{1\mathbb{P}} \sim \alpha^2\, e^{\omega_0 \tau}. \qquad (6.69)$$

In fact, the functional form in Eq. (6.68) appears to be consistent with the prediction (4.18) of the BK equation [162], although the coefficient $\kappa \approx 0.72$ which appears in Eq. (6.68) is somewhat smaller than the corresponding value $c/2 \approx 2.44$ in Eq. (4.18).

As explained in Ref. [59], the difference between the correct result, Eq. (6.68), and the corresponding predictions of either the naive Glauber formula, Eq. (6.69), or the BK equation, cf. Eq. (4.18), reflects the limitations of the mean field approximations underlying

the latter calculations. Namely, as discussed in Refs. [45, 59, 231], the approach of the (average) S–matrix $\mathcal{S}$ towards the black–disk limit $\mathcal{S} = 0$ with increasing energy is controlled by *very rare* configurations, which involve only few dipoles (much fewer than the typical configurations at the respective energy), and therefore yield relatively large contributions $S \sim 1$. Note indeed that the probability for suppressing evolution, although very small at high energy (see Eq. (6.10)), decreases with τ much slower than the S–matrix $S_{N \times N'}$ for the elastic scattering between two *typical* configurations, for which $N \sim N' \sim e^{\omega_0 \tau / 2}$. Hence, although very rare, the dilute configurations with $N \sim N' \sim \mathcal{O}(\infty)$ dominate the expectation value $\mathcal{S}$ at sufficiently high energy. In fact, the exponential result in Eq. (6.68) can be recognized [59] as the probability for the relevant, rare, configurations: one of the factors $\bar{\alpha}_s \tau$ in the exponent comes from the 'Sudakov factor', Eq. (6.10), and the other one, from the integral over the transverse size of the dipoles, so like in Eq. (4.17).

The fact that the prediction (4.18) of the BK equation appears to be much closer to the correct result (6.68) than the naive Glauber exponentiation, Eq. (6.69), is because the respective mean field approximations are performed in different frames: the rest frame of one of the incoming onia for the BK equation and, respectively, the COM frame for Eq. (6.69). Or, whereas in the COM frame, the relevant configurations have $N \sim N' \sim \mathcal{O}(\infty)$ — so, for them, the MFA is *very* bad —, on the other hand, in the rest frame of one dipole, $\mathcal{S}$ is dominated by relatively dense configurations in the wavefunction of the other dipole (the highly–evolved one), for which the MFA is better justified.

7 Beyond JIMWLK: Fluctuations & Pomeron loops

The importance of the correlations in the gluon distribution induced by the high–energy evolution have been one of the leitmotifs of the discussion in the previous sections. For instance, in Sect. 4.1 we have seen that the *color correlations* generated in this way modify the $k_\perp$–dependence of the unintegrated gluon distribution $\langle \rho_a(k_\perp) \rho_a(-k_\perp) \rangle_\tau$ from the bremsstrahlung spectrum $\sim 1/k_\perp^2$ at very large $k_\perp$ to a softer spectrum $\sim 1/k_\perp^{2\gamma}$ (with the 'anomalous dimension' $\gamma \simeq 0.63$) at intermediate momenta, and finally to an almost flat spectrum (logarithmic in $k_\perp$) at momenta below the saturation scale. Another example of color correlations is the coupling between scattering operators with different multipolar moments within the Balitsky equations, like the emergence of the quadrupole operator (3.90) in the r.h.s. of the evolution equation (3.89) for the two–dipole amplitude. As explained in Sect. 3.5, this coupling is suppressed in the large–N_c limit, where the Balitsky hierarchy essentially reduces to the BK equation.

Then, in Sect. 6 we have met with a new type of correlations, which refer to the distribution of the *colorless* dipole (or gluon) number operator, and thus survive at large N_c. These correlations are preponderantly produced via particle–number fluctuations in the dilute regime, which are then amplified by the BFKL evolution and eventually play an important role in the approach towards saturation and unitarity. This role has been partially anticipated by our discussion of multiple scattering within the dipole picture (cf. Sect. 6.5), and it should fully reveal itself through the subsequent analysis in this section. As we shall discover, the effects of the particle–number fluctuations on the high–energy

evolution in QCD are at the same time *dramatic* (in terms of physical consequences) and very *subtle* (in so far as the interplay between fluctuations and saturation is concerned). This subtlety, together with the technical challenges that these effects pose for the practical calculations (see below), may explain why they have not been properly appreciated for quite some time — in particular, they have not been included in the original Balitsky–JIMWLK equations, as explained in Ref. [48].

The conceptual understanding of these new effects has gradually developed over the last few years, starting with the accumulation of physical arguments (like the role of rare fluctuations in the approach towards unitarity [59]) and also of internal contradictions of the BK equation (like the lack of boost invariance [46]), which led to the recognition of a powerful correspondence between QCD and statistical physics [47] and, shortly after, to the construction of a new set of evolution equations [48–50] — the *'Pomeron loop' equations* —, which generalize the Balitsky–JIMWLK equations by including the effects of the particle–number fluctuations in the dilute regime, at the level of accuracy of the dipole picture. It is important to emphasize here that a complete leading–order formalism (valid for generic N_c and providing a faithful description of all physical regimes, including the dense–to–dilute transition region) is still lacking, in spite of significant progresses that we shall describe, or at least mention, in what follows. Besides, the known Pomeron loop equations are relatively complicated, due to their non–local and non–linear structure, and so far they resisted to all attempts to find explicit (analytic or numerical) solutions. So, in fact, most of the properties expected for these solutions have been inferred from the correspondence with statistical physics alluded to above, to which we shall devote a large part of the subsequent discussion.

7.1 Beyond the Balitsky–JIMWLK equations: a diagrammatic perspective

The discussion of the dipole picture in Sect. 6 gives us guidance about the physical effects which have been missed by the Balitsky–JIMWLK equations — namely, the particle–number fluctuations in the dilute regime —, and also about the modality to construct new, improved, equations which take these effects into account at large N_c. We shall present the improved equations in the next subsection. But before that, it is useful to gain more intuition about the role of the fluctuations in the high–energy evolution by inspecting the corresponding Feynman graphs.

As in the previous construction of the JIMWLK equation in Sect. 3.1, we shall not compute the relevant graphs explicitly (so, we do not need to specify the Feynman rule), but rather we shall rely on them to illustrate the perturbative content of the evolution equations. Then, the diagrams describing gluon–number fluctuations in perturbative QCD can be obtained by straightforwardly redrawing the corresponding diagrams of the dipole picture. From Sect. 6.2, we recall that, at large N_c, such a fluctuation corresponds to an evolution process in which one dipole splits into two and *both* child dipoles are subsequently measured — so that one keeps trace of the correlations among these dipoles which come from the fact that they have a common parent. Now, the dipoles correspond to gluons in the s–channel, while the *measurement* of a dipole proceeds by exchanging gluons in the t–channel (say, with an external projectile). This is manifest in the color

189

glass description of the onium, cf. Sect. 6.4, where a dipole is precisely represented as a source for gluons in the t–channel. Thus, a particle–number fluctuation, which in the dipole picture is depicted as in Fig. ??, can be alternatively represented by the CGC–like diagrams in Fig. ??, which involve gluons alone. In drawing Fig. ??, we have assumed that each dipole is measured via the exchange of only two gluons[31], which is indeed the relevant situation at high energy, cf. Sect. 6.5.3. Also, the parent dipole was implicitly assumed to belong to some general, 'onium', wavefunction, represented by the upper blob in Fig. ??. Note that, among all the quark and antiquark lines appearing in Fig. ??, only that pair has been represented in Fig. ?? which corresponds to the 'semi–hard' gluon emitted in the s–channel in one step of the evolution.

Clearly, the diagrams in Fig. ?? are not of the type discussed in Sect. 3.1 in relation with the JIMWLK equation. For more clarity, we have redrawn in Figs. ?? some diagrams which are representative for the latter: all such diagrams allow for the emission of *only two* new t–channel gluons in one step of the evolution. This corresponds to the fact that the JIMWLK Hamiltonian (3.36) is a functional operator of only second order in $\delta/\delta\alpha$, but of all orders in α; that is, this operator involves $n \rightarrow 2$–gluons vertices with any $n \geq 2$. By contrast, the 'fluctuation' effect in Fig. ?? is clearly a $2 \rightarrow 4$–gluons vertex, so it would correspond to a term $\alpha^2(\delta/\delta\alpha)^4$ in the evolution Hamiltonian. Such a term has been indeed observed in the 'color–glass' version of the Hamiltonian for the dipole picture, cf. Eq. (6.43). Via this vertex, the 2–point function $\langle\alpha\alpha\rangle_\tau$ acts as a source for the evolution of the 4–point function $\langle\alpha\alpha\alpha\alpha\rangle_\tau$. Recalling that the bilinear $\alpha\alpha$ corresponds, roughly speaking, to the gluon–number operator, we conclude that this vertex describes indeed a fluctuation in the gluon number, as expected.

To understand the reason for neglecting this vertex within the context of the JIMWLK evolution, one should compare its effects to those of the vertices which, in that context, contribute to the evolution of the 4–point function (see Fig. ??). At low density (the relevant regime for such a comparison), the BFKL diagram in Fig. ??.a gives the dominant contribution, which is of order $g^2\langle\alpha\alpha\alpha\alpha\rangle_\tau \sim g^2\langle\varphi\varphi\rangle_\tau$. Here, $\varphi \propto \alpha\alpha$ is the gluon occupation number, cf. Eq. (2.33), presently understood as an operator. By comparison, the contribution $\sim g^4\langle\varphi\rangle_\tau$ of the 'fluctuation' diagram in Fig. ?? is formally suppressed by a factor g^2. Thus, it is indeed legitimate to neglect this contribution *unless* the density is so low that $\langle\varphi\rangle_\tau \lesssim g^2$. (In the language of Sect. 6, this corresponds to an average dipole occupation number $\lesssim 1$.) When this happens, not only the saturation effects are completely negligible — this would only require the weaker condition $\langle\varphi\rangle_\tau \ll 1/g^2$ —, but in fact there is *on the average* 'of order $\mathcal{O}(\}^{\in})$ gluons' per unit cell in the transverse phase–space. Physically, this means that, when looking at a particular realization of the evolution, or 'event', one finds that most of these cells are actually empty. The fact that fluctuations are relatively important in such a dilute regime is, of course, not a surprise; however, this regime was *a priori* expected to be irrelevant for the physical problem of interest for us here, which is the evolution towards saturation and unitarity with increasing energy/gluon density. It is on the basis of this expectation that the particle–number

---

[31] In the general case, where a dipole can emit arbitrarily many gluons, so like in Eq. (6.28), one can show [224–226] that a $1 \rightarrow 2$ *dipole* splitting corresponds to $2 \rightarrow n$–gluons vertices with any $n \geq 2$ (evaluated at large N_c, of course).
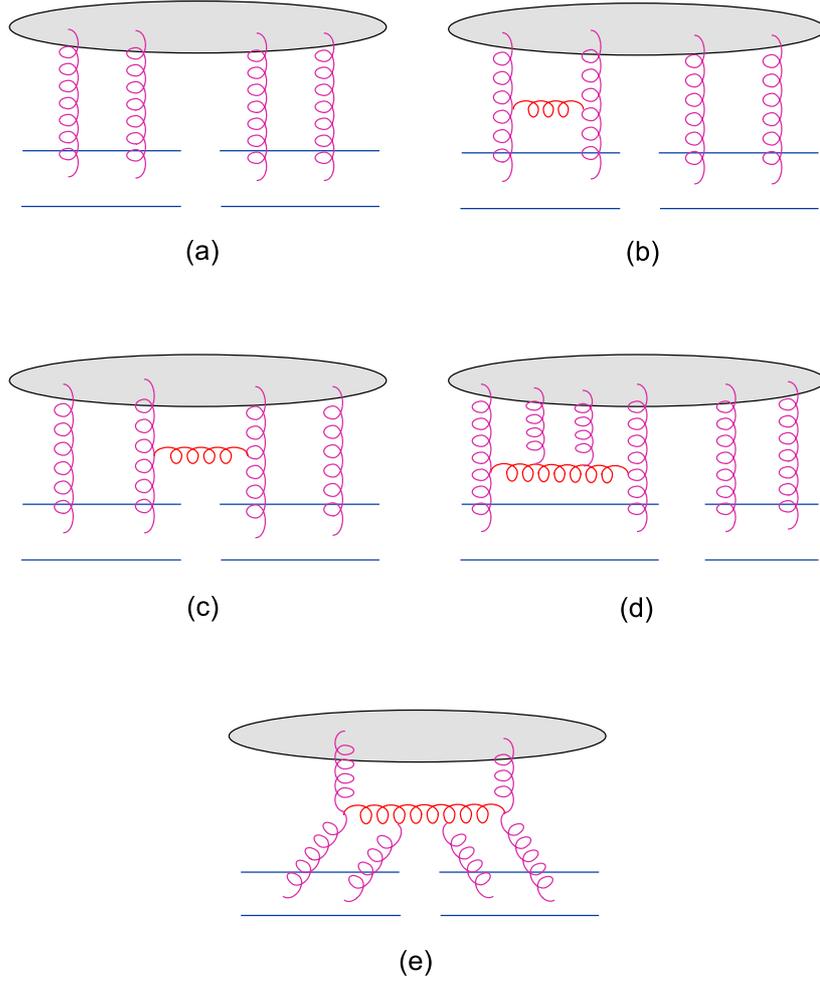
Fig. 49. *The one–step evolution of the scattering amplitude for a pair of dipoles: (a) tree–level contribution; (b, c) BFKL evolution; (d) saturation effects; (e) gluon–number fluctuations.* `FIG_2DIP`

fluctuations have been neglected in the construction of the JIMWLK equation [31–35]. However, this expectation turned out to be naive, for reasons to be explained in detail later on and which can be summarized here as follows: the dynamics in the dilute tail of the gluon distribution at high momenta, where fluctuations are important, is the driving force for the evolution towards high density with increasing energy.

Let us finally consider, still at the diagrammatic level, the consequences of the fluctuations on the scattering with external dipoles. This will allows us to make contact with the 'dual' point of view, in which is evolution is interpreted as *projectile* evolution. In order to probe the evolution of the 4–point function $\langle \alpha\alpha\alpha\alpha \rangle_\tau$, we consider the scattering with two external dipoles, as illustrated by the diagrams in Fig. 49. The first diagram, Fig. 49.a, represents the 'tree–level' process, while the other diagrams illustrate the contributions generated in one step of the evolution, that one should compare to the terms in the r.h.s. of the Balitsky equation for $\langle T^{(2)} \rangle_\tau$, Eq. (3.89). Namely, Figs. 49.b and c describe the BFKL evolution of the four–gluon exchange in Fig. 49.a (the second diagram, which corresponds to the quadrupole piece in the r.h.s. of Eq. (3.89), is suppressed at large N_c), while Fig. 49.d represents a saturation effect in the target wavefunction and corresponds to
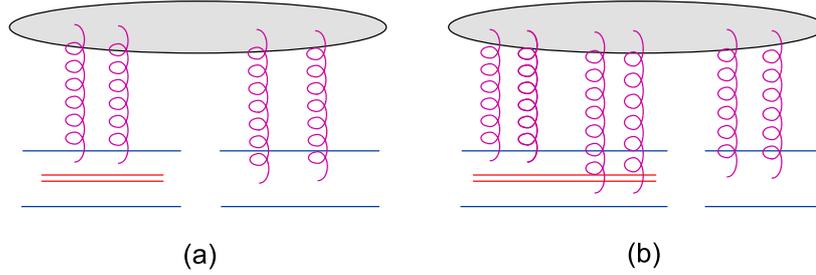
Fig. 50. *Alternative representations of the processes in Figs. 49.b and d, which are now inter-preted as the splitting of a dipole from the projectile, followed by the scattering of one of the child dipoles (a), or of both of them (b).* FIG_2DIPPROJ

the non–linear terms (the 'unitarity corrections') in Eq. (3.89). Finally, Fig. 49.e displays the contribution of the gluon–number fluctuations in the target, which is missing in the Balitsky–JIMWLK equations. From the structure of this diagram, it is suggestive that it should correspond to a term linear in $\langle T \rangle_\tau$ in the r.h.s. of the evolution equation for $\langle T^{(2)} \rangle_\tau$. Such a term will be explicitly constructed in the next subsection.

Of course, all these processes can be reinterpreted as the evolution of the projectile, and this interpretation is most suggestive for the diagrams in Figs. 49.b and d, which then correspond to the splitting of one of the projectile dipoles, as represented in Fig. 50. In particular, the diagram in Fig. 50.b can be characterized as a dipole–number fluctuation in the wavefunction of the projectile: one dipole splits into two, and then both child dipoles are measured through scattering. From this new perspective, Fig. 49.c describes a process in which the two projectile dipoles exchange a gluon with each other and thus evolve into a color quadrupole, which then interacts with the target via a four–gluon exchange. Finally, Fig. 49.e can be reinterpreted as a saturation effect in the projectile wavefunction — this amounts the reading the $2 \rightarrow 4$–gluon vertex there upside down. We thus see that, in the evolution of the scattering amplitudes, saturation effects in the target can be reinterpreted as particle–number fluctuations in the projectile, and vice–versa. This is just a consequence of boost invariance which allows one to transfer the evolution from the wavefunction of the target into that of the projectile, via a change of frame. As we shall later discover, this property introduces a constraint on the high–energy evolution, which can be formalized as the 'self–duality' of the corresponding Hamiltonian [61, 62].

7.2   The Pomeron loop equations

T_LANGEVIN

The previous analysis shows that in order to dispose of correct evolution equations for the scattering amplitudes which are complete to lowest order in perturbative QCD, one should generalize the previous approaches by including the effects of gluon–number fluctuations (or bremsstrahlung) in the JIMWLK evolution of the target wavefunction or, equivalently, by including saturation effects in the Balitsky approach to the evolution of the projectile.

This turns to be a formidable task, which in the framework of the JIMWLK–like renor-malization group analysis would require (for a single step in the evolution) the calculation of infinitely many $m \rightarrow n$–gluon vertices with arbitrary values for m and n and for an

infinite class of topologies. The first two topologies in this class are illustrated in Fig. ??. (As usual, we take the target to be a right–mover.) In the first diagram, Fig. ??.a, the 'semi–hard' gluon emitted in the s–channel is allowed to undergo multiple scattering off two types of fields: the fields A^+ radiated by the relatively fast color sources produced in the previous steps of the evolution (the corresponding rescatterings implement the saturation effects, as in the usual JIMWLK evolution) and the fields A^- which represent the additional t–channel gluons emitted in this step of the evolution, and which are arbitrarily many (and not just two, as in the JIMWLK evolution) since they must encode all the new correlations induced in the process (in particular, the effects of the gluon–number fluctuations). The second diagram, Fig. ??.a, deviates even further from the standard JIMWLK diagrams in the sense of including the self–interactions of the 'semi–hard' gluons (here, via a three–gluon vertex). Such a diagram would be negligible in the JIMWLK context, where A^+ is the only strong field ($A^+ \sim 1/g$), but now it becomes relevant since both A^+ and A^- are allowed to be strong, at least at a formal level, in order to encode all the relevant correlations produced by the evolution. (But, of course, A^- can be even *physically* strong, in the case of the scattering with a dense left–mover.)

As a matter of facts, the classes of diagrams represented in Fig. ?? have been explicitly computed [63, 228], for generic values of A^+ and A^-. However, conceptual problems appear when trying to promote the respective results into vertices in the evolution Hamiltonian: at a formal level, this requires replacing $A_a^- \to i\delta/\delta\rho_a$, with ρ_a the color charge density of the CGC, but in practice this procedure meets with ambiguities beyond the JIMWLK approximation. Besides, there are infinitely many other diagrams, involving all the possible *tree–level* self–interactions of the 'semi–hard' gluons, which so far have not been computed. We refer to Refs. [60, 63, 224–226, 228] for a more detailed account of such problems and further discussions.

Fortunately, at large N_c, there exists an alternative strategy [48] which, although not entirely rigorous, has the merit to provide a physically meaningful description while avoiding most of the complications alluded to above: This consists in putting together the dipole picture and the Balitsky–JIMWLK equations — which separately describe the particle–number fluctuations at low density and, respectively, the saturation/unitarization effects at high density — in order to obtain a unified description which is, by construction, piecewise correct and includes the most important physical ingredients. A priori, the transition region between the dense and the dilute regimes is not correctly covered by this description; still, as we shall later argue, this failure is expected to be irrelevant for the most interesting physical issues.

Consider first the weak–scattering regime, as corresponding to the dilute tail of the target wavefunction, or, equivalently, to a projectile made with relatively small dipoles. At large N_c, the relevant part of the target wavefunction can be described as a collection of dipoles which evolves according to the picture developed in Sect. 6. Then, the projectile dipoles scatter off the dipoles internal to the target, and the (forward) scattering amplitude can be computed according to formulæ like (6.48) or (6.51). By also using the evolution equations for the (target) dipole densities, as established in Sect. 6.2, one can easily deduce corresponding equations for the scattering amplitudes, valid within this weak–scattering regime. Finally, these equations are supplemented with the non–linear terms taken from

the Balitsky equations and thus extended into the non–linear, strong–scattering, regime. Specifically, by using Eq. (6.48) for the scattering amplitude of a single external dipole together with Eq. (6.14) for the dipole density $n_\tau(\boldsymbol{u}, \boldsymbol{v})$, one finds [58, 163] that $\langle T(\boldsymbol{x}, \boldsymbol{y})\rangle_\tau$ obeys the BFKL equation (1.65), as expected. This is the same as the linear part of the first Balitsky equation (3.88), and since this is also true for the non–linear term, by construction, we conclude that the first equation in the new hierarchy is formally the same as the corresponding Balitsky equation:

$$\frac{\partial \langle T(\boldsymbol{x}, \boldsymbol{y})\rangle_\tau}{\partial \tau} = \frac{\bar{\alpha}}{2\pi} \int_{\boldsymbol{z}} \left\{ \mathcal{M}_{\S\dagger\ddagger} \otimes \langle \mathcal{T}(\S, \dagger)\rangle_\tau - \mathcal{M}(\S, \dagger, \ddagger) \left\langle \mathcal{T}^{(\in)}(\S, \ddagger; \ddagger, \dagger)\right\rangle_\tau \right\}, \qquad (7.1)$$

where we have introduced the shorthand notation

$$\mathcal{M}_{\S\dagger\ddagger} \otimes \{(\S, \dagger) \equiv \mathcal{M}(\S, \dagger, \ddagger)[-\{(\S, \dagger) + \{(\S, \ddagger) + \{(\ddagger, \dagger)]. \qquad (7.2)$$

Note however that the meaning of the two–dipole amplitude which appears in the r.h.s. of Eq. (7.1) is now different, because of the fluctuation effects which enter the higher equations in the hierarchy, starting with the second one.

Namely, using Eq. (6.51) for a two–dipole projectile together with Eq. (6.18) for $n_\tau^{(2)}$, one finds an evolution equation for $\langle T^{(2)}\rangle_\tau$ which, in addition to the standard BFKL terms (one for each external dipole; see Eq. (3.89)), contains also a *fluctuation term*, as induced by the corresponding term in the equation for $n_\tau^{(2)}$ (cf. Eq. (6.17)):

$$\frac{\partial \left\langle T^{(2)}(\boldsymbol{x}_1, \boldsymbol{y}_1; \boldsymbol{x}_2, \boldsymbol{y}_2)\right\rangle_\tau}{\partial \tau}\bigg|_{\text{fluct}} = 2\alpha^4 \frac{\bar{\alpha}}{2\pi} \int_{\boldsymbol{u},\boldsymbol{v},\boldsymbol{z}} \mathcal{M}(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{z})\, n_\tau(\boldsymbol{u}, \boldsymbol{v})$$
$$\times \mathcal{A}_{/}(\S_\infty, \dagger_\infty | \sqcap, \ddagger)\, \mathcal{A}_{/}(\S_\in, \dagger_\in | \ddagger, \sqsubseteq). \qquad (7.3)$$

In this equation, one can eliminate the dipole density $n_\tau(\boldsymbol{u}, \boldsymbol{v})$ in favour of the scattering amplitude $\langle T(\boldsymbol{u}, \boldsymbol{v})\rangle_\tau$ by using the following relation, obtained from Eqs. (6.48) and (6.47) (recall that $\nabla_{\boldsymbol{x}}^2 \ln \boldsymbol{x}^2 = 4\pi \delta^{(2)}(\boldsymbol{x})$) :

$$n_\tau(\boldsymbol{x}, \boldsymbol{y}) + n_\tau(\boldsymbol{y}, \boldsymbol{x}) = \frac{4}{g^4} \nabla_{\boldsymbol{x}}^2 \nabla_{\boldsymbol{y}}^2 \langle T(\boldsymbol{x}, \boldsymbol{y})\rangle_\tau. \qquad (7.4)$$

This relation is valid for $\boldsymbol{x} \neq \boldsymbol{y}$, and can be understood as the amputation of the gluon lines attached to the target dipole. By using this, one finds:

$$\frac{\partial \left\langle T^{(2)}(\boldsymbol{x}_1, \boldsymbol{y}_1; \boldsymbol{x}_2, \boldsymbol{y}_2)\right\rangle_\tau}{\partial \tau}\bigg|_{\text{fluct}} = \left(\frac{\alpha}{2\pi}\right)^2 \frac{\bar{\alpha}}{2\pi} \int_{\boldsymbol{u},\boldsymbol{v},\boldsymbol{z}} \mathcal{M}(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{z})\, \nabla_{\boldsymbol{u}}^2 \nabla_{\boldsymbol{v}}^2 \langle T(\boldsymbol{u}, \boldsymbol{v})\rangle_\tau$$
$$\times \mathcal{A}_{/}(\S_\infty, \dagger_\infty | \sqcap, \ddagger)\, \mathcal{A}_{/}(\S_\in, \dagger_\in | \ddagger, \sqsubseteq). \qquad (7.5)$$

The overall evolution equation for $\langle T^{(2)}\rangle_\tau$ within the present scheme can be now obtained by adding the unitarity corrections at high energy, i.e. the terms involving $\langle T^{(3)}\rangle_\tau$ in the

194

corresponding Balitsky equation (3.89). One finally obtains[32] [48, 49]

$$
\frac{\partial \left\langle T^{(2)}(\boldsymbol{x}_1, \boldsymbol{y}_1; \boldsymbol{x}_2, \boldsymbol{y}_2) \right\rangle_\tau}{\partial \tau} = \frac{\bar{\alpha}}{2\pi} \int_{\boldsymbol{z}} \left\{ \left[\mathcal{M}_{\S_\infty \dagger_\infty \ddagger} \otimes \left\langle \mathcal{T}^{(\in)}(\S_\infty, \dagger_\infty; \S_\in, \dagger_\in) \right\rangle_\tau \right. \right.
$$
$$
\left. \left. - \mathcal{M}(\boldsymbol{x}_1, \boldsymbol{y}_1, \boldsymbol{z}) \left\langle T^{(3)}(\boldsymbol{x}_1, \boldsymbol{z}; \boldsymbol{z}, \boldsymbol{y}_1; \boldsymbol{x}_2, \boldsymbol{y}_2) \right\rangle_\tau \right] + [1 \leftrightarrow 2] \right\}
$$
$$
+ \left. \frac{\partial \left\langle T^{(2)}(\boldsymbol{x}_1, \boldsymbol{y}_1; \boldsymbol{x}_2, \boldsymbol{y}_2) \right\rangle_\tau}{\partial Y} \right|_{\text{fluct}}, \tag{7.6}
$$

with the last term given by Eq. (7.5). We have again used the notation in Eq. (7.2). As anticipated in the previous subsection, the fluctuation term is formally suppressed with respect to the BFKL terms by a factor of α^2, but it is truly a leading–order effect in the dilute regime where $\langle T \rangle_\tau \lesssim \alpha^2$ [48]. In that regime, this term acts as a *source* for two–body correlations in the scattering off external dipoles. Because of such correlations, the two–dipole amplitude $\langle T^{(2)} \rangle_\tau$ does not factorizes for generic Y, even if it did so in the initial conditions (e.g., if the target at $Y = 0$ was a large nucleus described by the MV model).

Eq. (7.6) is representative for the new hierarchy including both saturation and fluctuations [49] : the scattering amplitude $\langle T^\kappa \rangle_\tau$ for a projectile made with κ dipoles obeys an evolution equation which in addition to the BFKL terms, linear in $\langle T^\kappa \rangle_\tau$, and to the unitarity corrections, proportional to $\langle T^{(\kappa+1)} \rangle_\tau$, involves also terms linear in $\langle T^{(\kappa-1)} \rangle_\tau$, which describe the effects of gluon–number fluctuations and can be constructed via the straightforward generalization of the above procedure, starting with the respective equation for the κ–body dipole density $n_\tau^{(\kappa)}$, cf. Eq. (6.19). We shall not explicitly write down here the general equation (see Ref. [49]), but in the next subsection we shall present an alternative representation of the hierarchy as a whole, in terms of a special Langevin equation.

We have already presented, in Fig. 49.e, a diagrammatic interpretation for the fluctuation term. But now that we know the detailed structure of this term, cf. Eq. (7.5), it is instructive to give a more precise graphical representation for it, which exhibits the precise mechanism responsible for fluctuations. This is shown in Fig. 51, where we represent the target as a collection of dipoles (as appropriate in the dilute regime and at large N_c), one out of which splits into two dipoles which then interact both with the external ones.

It is furthermore interesting to consider *two* successive steps in the evolution of the single–dipole amplitude $\langle T(\boldsymbol{x}, \boldsymbol{y}) \rangle_\tau$, as described by Eqs. (7.1) and (7.6). This helps understanding why we refer to this new hierarchy as the *'Pomeron loop equations'*. Namely, the most interesting diagram generated by this two–step evolution is that in Fig. 52, which involves both fluctuations and saturation. The upper part of this diagram, which belongs to the dipole picture, is the same as the splitting process in Fig. 51. The lower part, which is characteristic for the CGC formalism (compare to Fig. 34.f), describes the recombination of the four gluons resulting from that splitting into two gluons, which then scatter with the projectile dipole. The overall diagram in the r.h.s. of Fig. 52 represents the simplest *Pomeron loop* generated by the present evolution; here, the 'Pomeron' is merely a two–

---

[32] Recall that the quadrupole operator $\langle Q \rangle$ in Eq. (3.89) has to be omitted at large N_c.
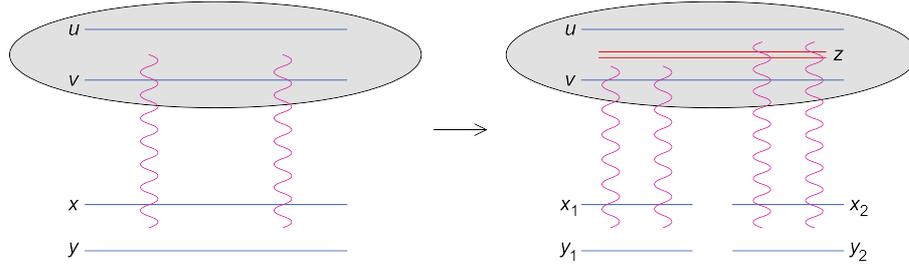
Fig. 51. *Diagrammatic illustration of the fluctuation term in Eq. (7.5) : the original dipole $(\boldsymbol{u}, \boldsymbol{v})$ within the target splits at the time of the interaction into two new dipoles $(\boldsymbol{u}, \boldsymbol{z})$ and $(\boldsymbol{z}, \boldsymbol{v})$, which then scatter off two external dipoles.* `Fig_split`
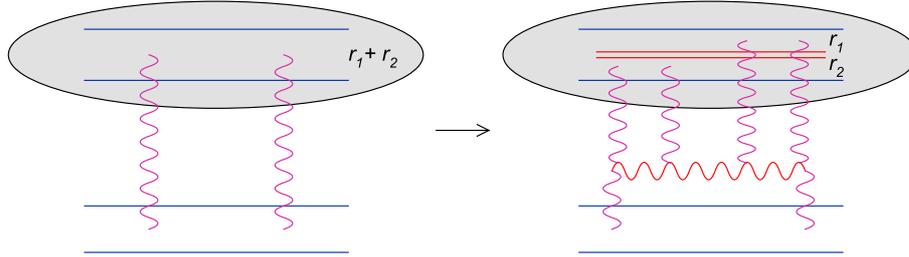

Fig. 52. *Two steps in the evolution of the average scattering amplitude of a single dipole: the original amplitude (left) and its evolution after two steps (right).* `Two_steps`

gluon exchange, but this will be dressed with BFKL ladders in the subsequent steps of the evolution, and thus become a genuine BFKL Pomeron.

Fig. 52 nicely illustrates an important characteristic of the evolution with Pomeron loops, that was missing in the Balitsky–JIMWLK equations, namely, the *feedback* of the evolution on the correlation functions: via particle–number fluctuations, the one–dipole amplitude $\langle T \rangle_\tau$ acts as a source for the two–dipole amplitude $\langle T^{(2)} \rangle_\tau$, which in turn influences the evolution of $\langle T \rangle_\tau$, via the saturation effects. In this way, $\langle T \rangle_\tau$ becomes sensitive to fluctuations, although the correspond evolution equation, Eq. (7.1), contains no explicit fluctuation term. The physical consequences of this feedback, together with other properties of the Pomeron loop equations and their solutions will be analyzed in the remaining part of this section.

Let us conclude this subsection with a word of warning concerning the interpretation of the fluctuation term, Eq. (7.5), from the perspective of *projectile* evolution. As already discussed in connection with Fig. 49.e, this term corresponds to a saturation effect in the projectile, that one might be tempted to interpret as *dipole recombination*: the two dipoles composing the projectile merge with each other into a single dipole, which then scatters off the target. Indeed, such a process would generate a term linear in $\langle T \rangle_\tau$ in the r.h.s. of the equation for $\langle T^{(2)} \rangle_\tau$, in agreement with the structure apparent in Eq. (7.5). Such an interpretation would be interesting since it would allow one to treat the ensemble of the evolution described by the Pomeron loop equations as a generalized 'reaction–diffusion' process occurring in the wavefunction of the projectile, in which dipoles evolve

196

via splitting and merging [33] . In turn, this would open the way for Monte–Carlo simulations of this evolution. Unfortunately, however, this interpretation is incorrect, as shown by the following argument [221]: after some integrations by parts, the r.h.s. of Eq. (7.5) can be rewritten in such a way to isolate the would–be $2 \to 1$–dipole recombination vertex:

$$\frac{\partial \left\langle T^{(2)}(\boldsymbol{x}_1, \boldsymbol{y}_1; \boldsymbol{x}_2, \boldsymbol{y}_2) \right\rangle_\tau}{\partial \tau}\bigg|_{\text{fluct}} = \left(\frac{\alpha}{2\pi}\right)^2 \frac{\bar{\alpha}}{2\pi} \int_{\boldsymbol{u}, \boldsymbol{v}} \mathcal{V}_{\ni\mathbb{P}}(\S_\infty, \dagger_\infty; \S_\in, \dagger_\in | \sqcap, \sqsubseteq) \left\langle \mathcal{T}(\sqcap, \sqsubseteq) \right\rangle, (7.7)$$

where

$$\mathcal{V}_{\ni\mathbb{P}}(\S_\infty, \dagger_\infty; \S_\in, \dagger_\in, \sqsubseteq) \equiv \int_{\ddagger} \nabla_{\sqcap}^{\in} \nabla_{\sqsubseteq}^{\in} \Big[\mathcal{M}(\sqcap, \sqsubseteq, \ddagger) \, \mathcal{A}_{\prime}(\S_\infty, \dagger_\infty | \sqcap, \ddagger) \, \mathcal{A}_{\prime}(\S_\in, \dagger_\in | \ddagger, \sqsubseteq) \Big] (7.8)$$

is recognized as the *triple–Pomeron vertex* which describes a $2 \to 6$–gluon transition in one step of the high–energy evolution at large N_c [138–141, 146, 148]. For this vertex to be identified with a probability rate for dipole merging, it should be positive–semidefinite, so like the corresponding vertex for dipole splitting — the 'dipole kernel' in Eq. (6.6). However, this condition is not satisfied for the vertex in Eq. (7.8), which is a total derivative with respect to its arguments $\boldsymbol{u}$ and $\boldsymbol{v}$, and hence it has not a fixed sign. Indeed, we have

$$\int_{\boldsymbol{u}, \boldsymbol{v}} \mathcal{R}(\sqcap_\infty, \sqsubseteq_\infty; \sqcap_\in, \sqsubseteq_\in | \sqcap, \sqsubseteq) = \prime \tag{7.9}$$

(one can check that the boundary terms vanish [221]), showing that this vertex must be negative in some regions of the transverse space, and positive in some others. Even though a little disappointing (as it prevents straightforward Monte–Carlo simulations of the evolution with Pomeron loops), this lack of a dipolar description for the effects of saturation should not come as a surprise: as emphasized in Sect. 6, the dipole picture cannot accommodate the physics of saturation, which involves gluon exchanges inside the wavefunction and hence goes beyond a strict large–N_c approximation. Rather, this physics must be described in terms of the fundamental, gluon, degrees of freedom, and at that level it can be given a probabilistic interpretation — as manifest in the context of the JIMWLK equation. However, the complicated nature of this phenomenon makes that the corresponding statistical description is generally more complicated than a simple reaction–diffusion process. Yet, in what follows, we shall argue that such a simple description can nevertheless be given under some additional assumptions.

7.3 The stochastic FKPP equation

In Sect. 3.3, we have already noticed the possibility to reformulate the JIMWLK evolution as a Langevin equation with multiplicative noise which describes a random walk in the

---

[33] The general structure of the evolution equations that would be generated by a reaction–diffusion process in the projectile can be found in Ref. [221].

functional space of the color charge density, or the associated color field. The noise term in that equation carries a color index showing that the corresponding fluctuations refer to the color degrees of freedom and thus die way, as we have seen, in the limit where $N_c \gg 1$. As furthermore mentioned in Sect. 3.3, the Langevin formulation was particularly convenient for numerical simulations of the JIMWLK evolution, as performed in Ref. [180].

It turns out that the Pomeron loop hierarchy introduced in the previous section can also be compactly encoded into a single Langevin equation [49] — now, with a colorless (multiplicative) noise term, as appropriate for fluctuations in the particle number —, whose practical utility (say, in view of numerical simulations) is however not obvious, becose of the rather complicated structure of the noise. Still, the ensuing equation is interesting, at least, at a conceptual level, since it allows us to make contact with the similar description of the reaction–diffusion process in statistical physics.

Let us therefore introduce the following Langevin equation ("the BK equation supplemented with a noise term")

$$\frac{\partial T_\tau(\boldsymbol{x}, \boldsymbol{y})}{\partial \tau} = \frac{\bar{\alpha}}{2\pi} \int_{\boldsymbol{z}} \left[\mathcal{M}_{\S\dagger\ddagger} \otimes \mathcal{T}_\tau(\S, \dagger) - \mathcal{M}(\S, \dagger, \ddagger) \, \mathcal{T}_\tau(\S, \ddagger) \, \mathcal{T}_\tau(\ddagger, \dagger) \right] + \frac{\partial \mathcal{T}_\tau(\S, \dagger)}{\partial \tau} \bigg|_{\text{noise}},$$

(7.10)   Tlang1

where the noise term reads

$$\frac{\partial T_\tau(\boldsymbol{x}, \boldsymbol{y})}{\partial \tau} \bigg|_{\text{noise}} = \frac{\alpha}{2\pi} \sqrt{\frac{\bar{\alpha}}{2\pi}} \int_{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{z}} \mathcal{A}_{\prime}(\S, \dagger | \sqcap, \ddagger) \frac{|\sqcap - \sqsubseteq|}{(\sqcap - \ddagger)^{\in}} \sqrt{\nabla_{\sqcap}^{\in} \nabla_{\sqsubseteq}^{\in} \, \mathcal{T}_\tau(\sqcap, \sqsubseteq)} \, \nu(\sqcap, \sqsubseteq, \ddagger, \tau),$$

(7.11)   Tlang

and $\nu(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{z}, \tau)$ is a Gaussian noise which is local in the evolution time τ and also in its third spatial variable $\boldsymbol{z}$, but which is non–local, and in a very special way, in its first two arguments: namely, it is *off–diagonal* in $\boldsymbol{u}$ and $\boldsymbol{v}$:

$$\langle \nu(\boldsymbol{u}_1, \boldsymbol{v}_1, \boldsymbol{z}_1, \tau) \nu(\boldsymbol{u}_2, \boldsymbol{v}_2, \boldsymbol{z}_2, \tau') \rangle = \delta^{(2)}(\boldsymbol{u}_1 - \boldsymbol{v}_2) \delta^{(2)}(\boldsymbol{v}_1 - \boldsymbol{u}_2) \delta^{(2)}(\boldsymbol{z}_1 - \boldsymbol{z}_2) \delta(\tau - \tau').$$

(7.12)   noise

This peculiar structure for the noise correlations is needed in order to properly account for the geometry of the dipole splitting (see below).

This stochastic equation is to be understood with the Ito prescription for discretizing rapidity (cf. Sect. 3.3), as usual for a renormalization group problem. One can then easily check that the equations for the correlators of T generated by this Langevin equation coincide indeed with the original Pomeron loop equations. Consider, e.g., the change in the 2–dipole amplitude $T_1 T_2$ under an infinitesimal increase $\Delta \tau$ in rapidity (in T_i, the index i serves as a shorthand for the coordinates of the i-th dipole). Eqs. (7.10)–(7.12) imply that the variation ΔT has two sources, $\Delta T = \Delta T|_{\text{BK}} + \Delta T|_{\text{noise}}$, where $\Delta T|_{\text{BK}} \sim \Delta \tau$ is the contribution induced by the BK–like terms in the r.h.s. of Eq. (7.10), whereas $\Delta T|_{\text{noise}} \sim \sqrt{\Delta \tau}$ is the contribution of the noise term there. (Recall that ν scales like $1/\sqrt{\Delta \tau}$; cf. Eq. (3.40).) Thus, in order to evaluate $\Delta(T_1 T_2)$ to $\mathcal{O}(\cdot \tau)$, one has to keep

terms up to second order in the ΔT's :

$$\Delta(T_1 T_2) = (T_1 + \Delta T_1)(T_2 + \Delta T_2) - T_1 T_2 = T_1\,\Delta T_2\,+\,\Delta T_1\,T_2\,+\,\Delta T_1\,\Delta T_2\,, \qquad (7.13)$$

where in the quadratic term $\Delta T_1\,\Delta T_2$ it is enough to keep the contribution of the noise. After averaging over the noise, cf. Eq. (7.12), we have, e.g., $\langle T_1\,\Delta T_2\rangle = \langle T_1\,\Delta T_2|_{\mathrm{BK}}\rangle$, so the first two terms, linear in ΔT_i, in the r.h.s. of Eq. (7.13) generate, as expected, the Balistky–like terms in the r.h.s. of Eq. (7.6). Furthermore, after using Eqs. (7.11)–(7.12) and some simple algebra, it is straightforward to verify that $\langle \Delta T_1\,\Delta T_2\rangle$ is indeed equal to the respective fluctuation term, cf. Eq. (7.5).

Eqs. (7.10)–(7.12) offer a synthetic representation of the Pomeron loop hierarchy, which is convenient for visualizing structural properties (e.g., is shows at a glance that the BK equation represents the *mean field approximation* in which one neglects the correlations induced by the noise term), but perhaps not also for direct numerical simulations. The off–diagonal nature of the noise correlations makes it tricky to simulate such a noise via Monte–Carlo techniques. Besides, the appearance of the square root factor in front of the noise, although physically natural (see below), may introduce further complications: The quantity under the square root is bound to be positive when evaluated *exactly* and in the *weak scattering regime* (where this quantity coincides with the dipole number density in the dilute target; cf. Eq. (7.4)), but it may be difficult to maintain its positivity in numerical calculations, where the discretization effects affect the smoothness of the solution T_τ. Besides, in the strong scattering regime, where $T_\tau \sim \mathcal{O}(\infty)$ and Eq. (7.4) is not valid anymore, the fluctuation term (7.11) is relatively unimportant (since suppressed by a power of α relative to the BK terms), yet it is not strictly zero, and the function under the square root need not be positive anymore. But perhaps the most serious obstacle against numerical simulations of either the above Langevin equation or the original Pomeron loop equations is the strong transverse non–locality of the triple–Pomeron vertex within the fluctuation terms, cf. Eqs. (7.5) or (7.8), which makes direct numerical integrations prohibitively difficult. This structure reflects the non–locality in impact parameter space of elementary processes like the $1 \to 2$–dipole splitting, cf. Fig. 47, and the elastic dipole–dipole scattering, cf. Eq. (6.47).

Via additional approximations — which amount to giving up the impact–parameter dependence of the scattering amplitudes —, it is possible to bring the fluctuation terms into a *local* form [48]. To that aim, one first notices that the elementary dipole–dipole scattering amplitude $T_0 \equiv \alpha^2 \mathcal{A}_\prime$, cf. Eq. (6.47), is *quasi–local* in transverse phase–space: we have indeed $T_0 \sim \alpha^2$ when the two dipoles have similar sizes and are relatively close to each other in impact parameter space (i.e., they overlap with each other), but T_0 decreases very fast for a large separation between the dipoles: $T_0 \sim \alpha^2 r_1^2 r_2^2/(\boldsymbol{b}_1 - \boldsymbol{b}_2)^4$ when $|\boldsymbol{b}_1 - \boldsymbol{b}_2| \gg r_>$, and also for a large disparity in between their respective sizes: $T_0 \sim \alpha^2 r_<^2/r_>^2$. In these formulæ, r_i, b_i, $i = 1, 2$, are the transverse sizes and, respectively, impact parameters of the dipoles partaking in the collision, and $r_< = \min(r_1, r_2)$, $r_> = \max(r_1, r_2)$. This discussion suggests the following, local, approximation for T_0 :

$$\alpha^2 \mathcal{A}_\prime(\S, \dagger | \sqcap, \sqsubseteq) \simeq \rfloor\,\alpha^\in (\S - \dagger)^\triangle\,\delta^{(\in)}(\S - \sqcap)\,\delta^{(\in)}(\dagger - \sqsubseteq), \qquad (7.14)$$

199

with c some numerical fudge factor. Within this approximation, the fluctuation term takes the simpler form

$$\frac{\partial \left\langle T^{(2)}(\boldsymbol{x}_1, \boldsymbol{y}_1; \boldsymbol{x}_2, \boldsymbol{y}_2) \right\rangle_\tau}{\partial \tau} \bigg|_{\text{fluct}} \approx c\, \alpha^2\, \frac{\bar{\alpha}}{2\pi}\, \frac{(\boldsymbol{x}_1 - \boldsymbol{y}_1)^2 (\boldsymbol{x}_2 - \boldsymbol{y}_2)^2}{(\boldsymbol{x}_1 + \boldsymbol{y}_2)^2}\, \left\langle T(\boldsymbol{x}_1, \boldsymbol{y}_2)\right\rangle_\tau\, \delta^{(2)}(\boldsymbol{x}_2 - \boldsymbol{y}_1)\,,$$

(7.15)   T2appr

where as compared to Eq. (7.5) all the convolutions have now disappeared: the coordinates of the external dipoles are identified to those of the child dipoles produced after a splitting in the target. This identification requires the incoming dipoles to be contiguous with each other (cf. the δ–function in the r.h.s. of Eq. (7.15)), which in turn calls for an off–diagonal noise term in the associated Langevin equation — so like in Eq. (7.12).

To obtain a fully local noise term, we shall now perform a *coarse–graining* which eliminates the dependence upon the impact parameter. This is physically motivated, since an external dipole of size r cannot distinguish details in the target on a transverse scale much smaller than r. That is, the gluon (or dipole) distribution measured by the external dipole is necessarily averaged in $\boldsymbol{b}$ over a disk of radius r around its impact parameter. As usual, we shall assume the target to be quasi–homogeneous over distances of $\mathcal{O}(r)$. It is convenient to go to momentum space: $T_\tau(\boldsymbol{x}, \boldsymbol{y}) \to \phi_\tau(\boldsymbol{k}, \boldsymbol{b})$, where

$$\phi_\tau(\boldsymbol{k}, \boldsymbol{b}) \equiv \int \frac{\mathrm{d}^2 \boldsymbol{r}}{2\pi \boldsymbol{r}^2}\, \mathrm{e}^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\, T_\tau(\boldsymbol{r}, \boldsymbol{b})\,,$$

(7.16)

with the standard notations $\boldsymbol{r} = \boldsymbol{x} - \boldsymbol{y}$, $\boldsymbol{b} = (\boldsymbol{x} + \boldsymbol{y})/2$. (Incidentally, the quantity $\varphi_\tau(\boldsymbol{k}, \boldsymbol{b}) = (1/\bar{\alpha})\phi_\tau(\boldsymbol{k}, \boldsymbol{b})$ can be interpreted as the gluon occupation number in the target [51].) The coarse–graining then consists in dividing the impact–parameter space into cells with area $\Sigma \sim 1/k^2$ and averaging the evolution equations over impact parameters within given cells. After doing this, different cells decouple from each other, and the ensuing hierarchy can be succinctly encoded into the following Langevin equation [48]:

$$\frac{1}{\bar{\alpha}}\frac{\partial \phi_\tau(\boldsymbol{k})}{\partial \tau} = \int \frac{\mathrm{d}^2 \boldsymbol{p}}{2\pi}\, \frac{\boldsymbol{k}^2}{\boldsymbol{p}^2 (\boldsymbol{k} - \boldsymbol{p})^2}\left(2 \frac{\boldsymbol{p}^2}{\boldsymbol{k}^2}\, \phi_\tau(\boldsymbol{p}) - \phi_\tau(\boldsymbol{k})\right) - \phi_\tau^2(\boldsymbol{k})$$
$$+ \sqrt{\kappa\, \alpha^2\, \phi_\tau(\boldsymbol{k})}\, \nu(\boldsymbol{k}, \tau)$$

(7.17)

where $\phi_\tau(\boldsymbol{k})$ is the cell–averaged gluon occupation number, κ is some fudge factor, and $\nu(\boldsymbol{k}, \tau)$ is a Gaussian noise :

$$\left\langle \nu(\boldsymbol{k}, \tau)\right\rangle = 0, \qquad \left\langle \nu(\boldsymbol{k}_1, \tau_1)\nu(\boldsymbol{k}_2, \tau_2)\right\rangle = \frac{1}{\bar{\alpha}}\, \delta(\tau_1 - \tau_2)\, \delta^{(2)}(\boldsymbol{k}_1 - \boldsymbol{k}_2)\, \boldsymbol{k}_1^2\,.$$

(7.18)

Once again, Eq. (7.17) must be understood with the Ito prescription for discretizing τ.

As anticipated, the $\boldsymbol{b}$–dependence has completely disappeared from Eq. (7.17). The first line in this equation is recognized as the momentum–space version of the BK equation [124], whereas the noise term in the second line is responsible for deviations from the mean field behaviour. For the sake of the comparison with the reaction–diffusion process, it is convenient to introduce the logarithmic variable $\rho \equiv \ln(k_\perp^2/k_0^2)$ and use the diffusion

approximation (which, we recall, it is tantamount to a second–order expansion around the saturation saddle point) to simplify the non–locality of the BFKL kernel. This leads to the following, final, form for the Langevin equation which reproduces the QCD evolution with Pomeron loops within the present approximations [48] ($t \equiv \bar{\alpha}\tau$):

$$\partial_t\phi(\rho,t) \;=\; -c\partial_\rho\phi + (\beta/2)(-\partial_\rho - \gamma_0)^2\phi \;-\; \phi^2 \;+\; \sqrt{\kappa\,\alpha^2\,\phi}\,\nu(\rho,t), \qquad (7.19)$$

with the Gaussian white noise :

$$\langle\nu(\rho,t)\rangle = 0, \qquad \langle\nu(\rho,t)\nu(\rho',t')\rangle = \delta(t-t')\,\delta(\rho-\rho')\,. \qquad (7.20)$$

Up to a linear change of variables and an appropriate rescaling of ϕ, the above equation is essentially the same as the *stochastic FKPP equation* (sFKPP) [165, 166] — the Langevin equation which describes a reaction–diffusion process [232], whose standard form reads

$$\partial_t u(x,t) = \partial_x^2 u \;+\; u(1-u) \;+\; \sqrt{\frac{2}{N}\,u(1-u)}\,\nu(x,t)\,, \qquad (7.21)$$

where $N \gg 1$ is the maximal occupation number per site and $\nu(x,t)$ is a Gaussian noise with $\langle\nu(x,t)\rangle = 0$ and $\langle\nu(x,t)\nu(x',t')\rangle = \delta(t-t')\delta(x-x')$.

The physical meaning of the multiplicative noise in Eq. (7.21) becomes clear if one remembers that, in the corresponding microscopic problem (cf. Sect. 4.3.2), $u = n/N$, with n the number of particles per site. For such a particle system, one expects typical fluctuations to scale like $\delta n \sim \sqrt{n}$, and hence $\delta u = \delta n/N \sim \sqrt{u/N}$. This is indeed consistent with the noise term in Eq. (7.21) in the dilute regime where $u \ll 1$. One the other hand, when $u \to 1$ (i.e., in the high–density regime near, or at, saturation), the noise term in the sFKPP equation is explicitly suppressed by the factor $1 - u$ under the square root, which appears to be absent in the corresponding QCD equation, Eq. (7.19). However, this factor is not really essential [34], since when $u \sim 1$ the fluctuation term is anyway suppressed by a factor $1/N \sim \alpha^2$. Moreover, if one returns to the original Langevin equation in QCD, Eq. (7.11), one sees that also in that equation the coefficient of the noise vanishes at saturation ($T \to 1$), because of the gradients acting on T.

In Sect. 4.3.2, we have already noticed the correspondence between the FKPP equation (4.65), which is obtained by neglecting the noise term in Eq. (7.21), and the BK equation. Now we see that this correspondence between the high–energy evolution in QCD and the reaction–diffusion process in statistical physics extends beyond this mean field approximation, that is, it also holds after including the effects of particle number fluctuations (at least, at the level of accuracy of the approximations underlying Eq. (7.17)). In this analogy, the (reduced) particle occupation number $u(x,t) = n(x,t)/N$ in the reaction–diffusion process correspond to (momentum–space) dipole scattering amplitude $\phi(\rho,\tau)$ in QCD (with $x \leftrightarrow \rho$ and $t \leftrightarrow \tau$), and N corresponds to $1/\alpha^2$. These identifications may look formal at a first sight, but they reflect the fact that the evolution described by the

---

[34] Numerical simulations within the context of statistical physics [232, 233] have confirmed that the sFKPP equation and the so–called 'one–dimensional Reggeon model', where the coefficient of the noise is simply $\sqrt{(2/N)u}$, give identical results to the accuracy of interest.

Pomeron loop equations can be loosely viewed as a reaction–diffusion process with dipoles in either the target, or the projectile, wavefunction [47]. This interpretation is imperfect, because of the sign problem of the triple–Pomeron vertex (7.8) alluded to before, but it is still useful as it provides an intuitive basis for the correspondence with statistical physics. In particular, the fact that $N \sim 1/\alpha^2$ reflects the fact that the saturation effects become important when the dipole occupation number becomes of $\mathcal{O}(\infty/\alpha^\in)$ (cf. Eq. (6.12)). Note that this estimate, $N \sim 1/\alpha^2$, for the dipole occupation number at saturation is parametrically different from the corresponding one for the *gluon* occupation number, which is $\varphi \sim 1/\bar{\alpha}$. This difference can be attributed to the fact that the dipoles act as sources for the gluons, hence the emission of two gluons from one dipole (the process which measures the gluon number) costs an additional factor g^2.

7.4   Fluctuating pulled fronts: The saturation exponent

As previously mentioned, the complexity of the non–linear and non–local structure of the Pomeron loop equations has so far prevented the obtention of general solutions to these equations, via either analytic and numerical techniques. On the other hand, the 'sFKPP equation', by which we generically mean any of the Langevin equations shown in Eqs. (7.17)–(7.21), looks considerably simpler — there is no impact–parameter dependence anymore, and the fluctuation term has been reduced to a local form —, and indeed much more is known about its properties and solutions.

In the framework of statistical physics, Eq. (7.21) has been intensively studied in relation with stochastic processes of the 'reaction–diffusion' type. Such processes appear in various contexts — from physics to chemistry, and from astrophysics to biology —, and admit various microscopic realizations, yet they are known to exhibit a universal behaviour at late times, which is in particular encoded in the sFKPP equation (see the review papers [165, 166] for detailed discussions and more references). The same universal behaviour is therefore expected in QCD at sufficiently high energy — at least, within the accuracy limits of Eq. (7.17). More recently, the reaction–diffusion process has been also studied in relation with the high–energy QCD problem [234–236]; in that context, some efforts have been made to go beyond the canonical sFKPP equation (7.21) by including more detailed features of the actual QCD dynamics, like the non–locality of the BFKL kernel [234], or the effects of multiple scattering [236].

Also, in a search for analytic solutions, one has considered 'zero–dimensional' versions of the reaction–diffusion process in which the whole dynamics (i.e., the particle splitting and merging) occurs at a single point in space [45, 231, 237, 238]. Such models shed light on particular aspects of the dynamics with Pomeron loops, like the role of the rare fluctuations in the approach towards unitarity. On the other hand, they cannot accommodate the most interesting aspects of this dynamics, like the emergence of a saturation scale and its evolution with increasing energy. For this reason, we shall not discuss such models any further, but rather focus on the one–dimensional dynamics generated by the sFKPP equation. For applications to QCD, one should keep in mind that the 'spatial dimension' in Eq. (7.21) truly corresponds to a transverse *momentum* (or *size*) in QCD: $x \leftrightarrow \rho$, with $\rho \equiv \ln(k_\perp^2/k_0^2)$, where $k_\perp$ is roughly the momentum of the gluons probed in the target

wavefunction, or the inverse size of the dipole projectile ($k_\perp^2 \sim 1/r^2$).

In order to describe the effects of fluctuations, it is convenient to use the traveling–wave picture, as introduced in Sect. 4.3.2 in the context of the mean field approximation (i.e., for the solutions to the BK, or FKPP, equations). This picture remains appropriate in the presence of fluctuations since the solution to, say, Eq. (7.21) is again a front which interpolates between the fixed points $u = 1$ and $u = 0$, and which propagates towards larger values of x with increasing t, but whose propagation is now perturbed by a random force — the noise term. This perturbation has two major physical consequences:

(i) The evolution becomes *stochastic* : Instead of the single front that would be generated by a deterministic equation, so like BK, we now have a *statistical ensemble* of fronts which encompasses all the possible realizations of the evolution — where a 'realization' refers to a particular front obtained by iterating the Langevin equation up to the rapidity of interest. To describe this evolution, one can either use the 'event–by–event picture', which focuses on the individual realizations and their fluctuations, or the 'statistical description', which deals with quantities averaged over the ensemble, so like the n–point correlation functions $\langle u(x_1)\cdots u(x_n)\rangle_t$, or the n–dipole scattering amplitudes in QCD.

(ii) The evolution *'slows down'*, by which we more precisely mean that the *average* velocity of the fronts is smaller than the corresponding prediction of the mean field approximation, as previously obtained in Sects. 4.2–4.3. As we shall later see, this property reflects the fact that, in the presence of fluctuations, the 'unitarity corrections' (the non–linear terms in the Pomeron loop equations) become important already in the regime where the *average* amplitude is significantly smaller than one, and thus they considerably limit the subsequent evolution.

Although relatively easily to understand at a qualitative level, the quantitative description of these effects turns out to be quite subtle, and to reserve some surprises: On physical grounds, and also by inspection of the noise term in any of the previous Langevin equations, it is quite clear that the effects of the fluctuations must die away in the limit where the particle occupation number at saturation N becomes arbitrarily large. That is, one expects the mean field approximation to become reliable when $N \to \infty$, which in QCD corresponds to the weak coupling limit [35] $\alpha^2 \to 0$. And this is indeed what is found by the various analyses of the sFKPP equation (or the reaction–diffusion process), via either numerical and analytic techniques. What came however as a surprise in these analyses, is the *extremely slow convergence* towards the mean field limit with increasing N : when $N \to \infty$, the deviations due to fluctuations vanish only *logarithmically*, i.e., as inverse powers of $\ln N \sim \ln(1/\alpha^2)$, rather than as inverse powers of N, as generally expected for stochastic particle systems. In the remaining part of this subsection, we shall try to motivate this behaviour via physical arguments, and even give a heuristic derivation for it (taken from an original analysis by Brunet and Derrida [188], in the context of statistical physics; see also Ref. [46] for a somewhat different argument, in the framework of the BK

---

[35] Note that, in the present discussion, α^2 and $\bar{\alpha} \equiv \alpha N_c/\pi$ play different physical roles, and thus should be treated on different footings: While α^2 measures the strength of fluctuations and it can be arbitrarily small, $\bar{\alpha}$ sets the scale for the 'evolution time' $t \equiv \bar{\alpha}\tau$, which is typically large: $t \gg 1$. Hence, the 'weak coupling limit' is here understood as the limit $\alpha^2 \to 0$ at fixed t.
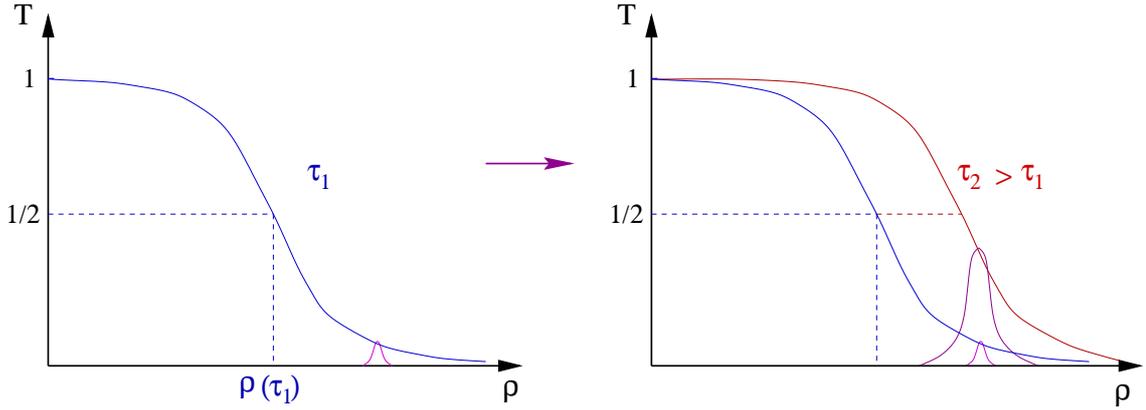
Fig. 53. *Evolution of the continuum front of the Balitsky–Kovchegov equation with increasing τ.* TWave5

equation), for the case of the *average front velocity* (i.e., the saturation exponent in QCD). Then, in the next subsection, we shall discuss the statistical properties of the ensemble, in particular, the *front dispersion*.

This anomalously slow convergence towards the mean field behaviour shows that the reaction–diffusion processes are more sensitive to fluctuations than the ordinary stochastic processes. This can be attributed to the *pulled* nature of the front, a property already emphasized in the discussion of the FKPP equation in Sect. 4.3.2: the front is driven by the dynamics in its tail, where the occupation numbers are relatively small and the fluctuations are correspondingly important. The following, qualitative, argument shows that the mechanism for front propagation is in fact different in the mean field approximation (i.e., for the solution to the BK, or FKPP, equation) and in the actual, microscopic, process, which includes fluctuations.

The propagation of the mean–field front is illustrated in Fig. 53, in notations appropriate for the QCD problem. The front has an exponential tail which extends up to arbitrarily large values of ρ; with increasing τ, this tail is rapidly amplified (at all the points $\rho \gg \rho_s$ where the amplitude is still small) by the BFKL instability. This rapid growth is predominantly *local* (the non–local, diffusion, term is relatively unimportant, as manifest at the level of the FKPP equation, where the diffusion term and the local growth term are isolated from each other), and it has the effect to 'pull' the front towards larger values of ρ. In Fig. 54, we similarly consider a particular realization of the front in the microscopic problem, where the particle occupation numbers are discrete [36] . The front looks now like a histogram and thus is necessarily *compact* — for any τ, there is only a finite number of bins in ρ ahead of ρ_s where T_τ is non–zero. In the empty bins on the right of the tip of the front, the local BFKL growth is not possible anymore, as this would require a seed. Thus, the only way for the front to progress towards these bins is via *diffusion*, i.e., via radiation from the occupied bins at $\rho < \rho_{\rm tip}$ (with $\rho_{\rm tip}$ = the 'tip of the front', i.e., the

---

[36] In QCD, the 'particles' can be assimilated with dipoles in the dilute part of the target wavefunction (since here we are interested only in the dynamics of the tail). Then, $T_\tau(\rho) \sim \alpha^2 n_\tau(\rho)$, with $n_\tau(\rho)$ = the dipole occupation number; accordingly, the event–by–event amplitude is a *discrete* quantity, which takes values in steps of α^2, as illustrated in Fig. 54.
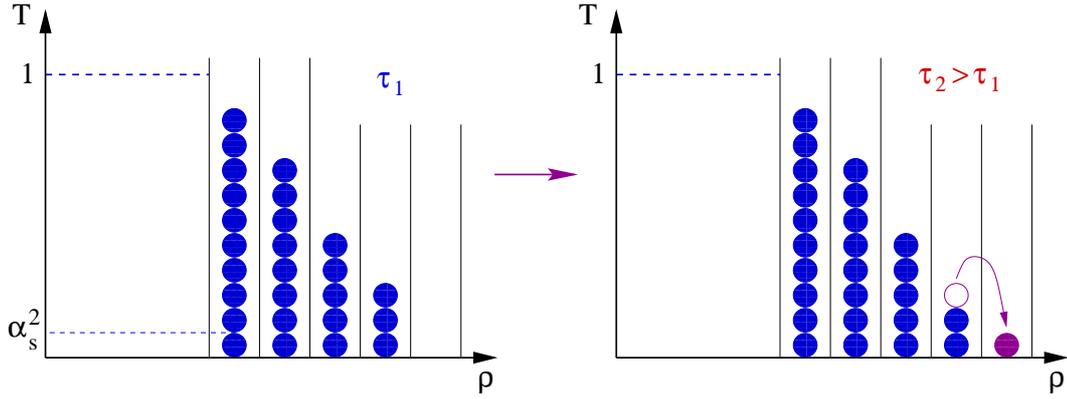
Fig. 54. *Evolution of the discrete front of a microscopic event with increasing rapidity τ. The small blobs are meant to represent the elementary quanta α^2 of T in a microscopic event.* TWave6

position of the foremost occupied bin). But since diffusion is less effective than the local growth, we expect the velocity v_N of the microscopic front to be reduced as compared to the respective prediction v_0 of the MFA.

Such a reduction has been indeed seen in all the numerical simulations of the reaction–diffusion dynamics, which moreover found that this reduction is quite dramatic, and that it dies away very slowly when increasing N : for $N \ggg 1$ (by which we mean a tremendously large value for N, so like $N \sim 10^{10}$), one finds $v_N \simeq v_0 - \mathcal{C}/\ln^2 N$, with $\mathcal{C}$ a constant factor [188, 232, 239, 240]. This logarithmic law appears to be *universal*, i.e., it is the same for all the reaction–diffusion processes, independently of their detailed microscopic dynamics. An analytic argument which explains this universal, slow, convergence and also provides an estimate for the coefficient $\mathcal{C}$ has been given by Brunet and Derrida [188].

Specifically, Brunet and Derrida argued that the dynamics of a given front in the presence of discreteness and for $N \ggg 1$ can be reasonably described by a modified version of the FKPP equation (4.65) in which the linear, growth, term is switched off when the concentration $u(x,t)$ becomes smaller than $1/N$ (its minimal value allowed when the particle number is taken to be discrete):

$$\partial_t u(x,t) = \partial_x^2 u(x,t) + \Theta(u - a/N)\, u(x,t)\big(1 - u(x,t)\big). \tag{7.22}$$

This equation captures in a simple way the physical ideas developed before, namely the fact that the MFA is expected to be reliable in the bulk of the front where $u \gg 1/N$, and that the local growth term should become ineffective towards the tip of the front, where $u \sim 1/N$. The fudge factor a is not under control, but this is not going to affect the dominant contribution to $v_N - v_0$, which is only logarithmically sensitive to the cutoff; thus, without loss in generality, we shall set $a = 1$ in what follows. Note that the form of Eq. (7.22) makes no reference to the detailed microscopic dynamics, in agreement with the expected universality at large N (and for large times).

Since we are interested in the dynamics of the tail of the front, where $u \ll 1$, we shall use the same strategy as for the BK equation in Sect. 4.2, that is, we shall solve the linearized version of Eq. (7.22) with saturation boundary conditions at large, negative, values of x. Anticipating that, for sufficiently large values of t, the solution will have a traveling–wave

pattern, we search for it in the form $u(x,t) = F(x - v_N t)$, where $F(z)$ represents the *shape* of the front and the velocity v_N is yet to be determined. As usual, we define the position $X(t)$ of the front (the would–be 'saturation line' in QCD) as a constant–level curve for $u(x,t)$: $u(X(t), t) = \kappa$, with $1 \gg \kappa \gg 1/N$. Clearly,

$$X(t) = X_0 + v_N t \qquad \text{for large } t, \tag{7.23}$$

with constant X_0. Also, we define the *width of the front* as the distance $L = x - X(t)$ ahead of the front where the amplitude falls down to a value $\sim \mathcal{O}(\infty/\mathcal{N}) : u(X(t) + L, t) = 1/N$. We expect L to increase with N, and this will be confirmed by the calculation to follow. For a given, and relatively large, t, we shall then distinguish between three physical regions:

- Region I : $x \ll X(t) \implies 1 \gtrsim u(x,t) \gg \kappa$.
  In this region, the dynamics is essentially non–linear.
- Region II : $X(t) < x < X(t) + L \implies \kappa > u(x,t) > 1/N$.
  In this region, the shape function $F(z)$, with $z \equiv x - X(t)$, obeys the linear equation:

$$F''(z) + v_N F'(z) + F(z) = 0. \tag{7.24}$$

- Region III : $x > X(t) + L \implies u(x,t) < 1/N$.
  In this region, the equation reduces to:

$$F''(z) + v_N F'(z) = 0. \tag{7.25}$$

Moreover, when $N \gg 1$, we expect the cutoff on u to have almost no effect on the dynamics in Region I. Hence, for $x \gtrsim X(t)$, which is the region of transition between Region I and Region II, we expect $F(z) \approx F_{v_0}(z) = A z \, \mathrm{e}^{-\gamma_0 z}$, cf. Eq. (4.71), where we recall that $v_0 = 2$ and $\gamma_0 = 1$ are the mean–field (or FKPP) estimates for the velocity and the exponential slope of the front, respectively. Clearly, v_N must reduce to v_0 as $N \to \infty$, and we shall write $v_N - v_0 \equiv \Delta$, where Δ becomes arbitrarily small for sufficiently large values of N, and it is expected to be positive according to our previous, physical, discussion.

The boundary condition

$$F(z) \simeq A z \, \mathrm{e}^{-\gamma_0 z} \qquad \text{for} \quad z \lesssim 1 \quad \text{and} \quad N \to \infty, \tag{7.26}$$

together with equations (7.24) and (7.25) completely determine the solution $F(z)$ in the linear regime, including the value of v_N, as we now show.

The general solution to Eq. (7.24) reads:

$$F(z) = C \, \mathrm{e}^{-\gamma_r z} \sin(\gamma_i z + D), \tag{7.27}$$

where γ_r and γ_i are, respectively, the real and the imaginary part of the solution γ_N to the dispersion equation $v_N = v(\gamma)$ (cf. Eq. (4.70)), that is,

$$v_N = v(\gamma) \equiv \gamma + \frac{1}{\gamma} \implies \gamma = \gamma_N \equiv \gamma_r + \mathrm{i}\gamma_i. \tag{7.28}$$

Clearly, we have $\gamma_N \to \gamma_0$ as $N \to \infty$, and it is easy to find the first corrections to γ_r and γ_i by expanding the above equation around its mean–field solution. One thus finds

$$v_N \simeq v_0 - \frac{\gamma_i^2}{2} v''(\gamma_0) \,, \tag{7.29}$$

which shows that $\gamma_i \sim \mathcal{O}(\cdot^{\infty/\in})$, and, respectively,

$$\gamma_r - \gamma_0 \simeq \frac{\gamma_i^2}{6} \frac{v'''(\gamma_0)}{v''(\gamma_0)} \,. \tag{7.30}$$

Note that $\gamma_r - \gamma_0 \sim \mathcal{O}(\cdot)$ is of higher order than γ_i, so in what follows we shall approximate $\gamma_r \simeq \gamma_0$ for $N \ggg 1$. Then, by comparing Eqs. (7.27) and (7.26) for $z \lesssim 1$ (and hence $\gamma_i z \ll 1$), one can deduce the values of the constants C and D: $C = A/\gamma_i$ and $D = 0$. Furthermore, to solution to Eq. (7.25) obeying the boundary condition $F(z) = 1/N$ for $z = L$ reads

$$F(z) = \frac{1}{N} e^{-v_N(z-L)} \,. \tag{7.31}$$

By imposing the continuity of F and of its first derivative at $z = L$, we are finally in a position to determine v_N and L. It is in fact easier to first deduce

$$L \simeq \frac{1}{\gamma_0} \ln N \qquad \text{and} \qquad \gamma_i \simeq \frac{\pi}{L} \simeq \frac{\pi\gamma_0}{\ln N} \,, \tag{7.32}$$

where the second result, namely $L\gamma_i \simeq \pi$, is the condition that the sine function in Eq. (7.27) has a node at $z = L$. This reflects the fact that the function $F(z)$ decays much faster at $z > L$ than at $z < L$, as obvious when comparing the respective exponents in Eqs. (7.27) and (7.31). When inserted into Eq. (7.29), this result yields [188]

$$v_N \simeq v_0 - \frac{\pi^2\gamma_0^2}{2} \frac{v''(\gamma_0)}{\ln^2 N} \qquad \text{for} \qquad N \ggg 1 \,. \tag{7.33}$$

The next–to–leading–order term in the large–N expansion of v_N has also been computed [241], and found to be of $\mathcal{O}(\infty/\ln^{\in}(\ln\mathcal{N}))$. Thus, the convergence towards the mean–field result with increasing N is extremely slow, as anticipated. The result (7.33), including the coefficient in front of the $1/\ln^2 N$–behaviour, has been checked through numerical simulations for various reaction–diffusion processes [188, 240, 241]. It is important to notice here that this result is *universal*, in the sense that it depends upon the microscopic process under consideration only via the dispersion relation $v = v(\gamma)$ for the corresponding mean field approximation.

In particular, the result corresponding to the high–energy evolution in QCD is obtained by replacing $1/N \to \alpha^2$ and $v(\gamma) \to \chi(\gamma)/\gamma$, where $\chi(\gamma)$ is the Mellin transform of the BFKL kernel, cf. Eq. (4.26). By also recalling that $t = \bar{\alpha}\tau$ and using the standard notation λ for the saturation exponent, we finally deduce the following prediction from statistical physics for the (asymptotic) saturation exponent in QCD at high energy [46, 47]:

$$\lambda \simeq \bar{\alpha} \left[c - \frac{\pi^2\gamma_s\beta}{2\ln^2(1/\alpha^2)} \right] \qquad \text{for} \qquad \alpha \ll 1 \,, \tag{7.34}$$

where $c \approx 4.88$, $\gamma_s \approx 0.63$, and $\beta \equiv \chi''(\gamma_s) \approx 48.52$ (cf. Sect. 4.2). Note that the second, corrective, term is particularly large, not only because it decreases very slowly with α^2, but also because its coefficient turns out to be numerically large: $\pi^2 \gamma_s \beta / 2 \approx 150$. Thus, Eq. (7.34) makes sense only for *extremely* small values of α, which are physically unrealistic. (In fact, its r.h.s. becomes negative as soon as $\alpha > 0.06$!) Yet, this result is conceptually interesting, in that it illustrates the importance of particle–number fluctuations for the dynamics towards saturation.

Let us conclude this presentation which a qualitative argument [47] which is able to reproduce the logarithmic behaviour in Eq. (7.34), if not also the respective coefficient, and thus sheds more light on the physical origin of this peculiar behaviour. This argument is based on an estimate of the *formation time* of the front, which is the time required for the front to reach the asymptotic, scaling, form in Eq. (7.27) starting from some arbitrary initial condition at $t = 0$, so like a Θ–function. Namely, from the discussion in Sect. 4.3, we know that the front sets in diffusively; thus, in order for the front to spread over a given distance $\rho - \rho_s$, one needs a rapidity evolution (cf. Eq. (4.56))

$$\bar{\alpha} \tau \sim \frac{(\rho - \rho_s)^2}{2\beta} \,. \tag{7.35}$$

During this evolution, the velocity of the front increases towards its asymptotic value according to Eq. (4.47). If the front is allowed to extend arbitrarily far away, as in the case of the MFA, then the velocity will asymptotically approach the value $\lambda_0 = c\bar{\alpha}$. However, when the front is compact, as is the case for the actual microscopic evolution, then the formation time is finite as well, namely of the order

$$\bar{\alpha} \Delta \tau \sim \frac{L^2}{2\beta} \sim \frac{\ln^2(1/\alpha^2)}{2\beta \gamma_s^2} \,, \tag{7.36}$$

where L is the width of front, as previously computed in Eq. (7.32). During such a limited evolution, the front velocity cannot increase beyond a maximal value

$$\lambda \simeq \lambda_0 - \frac{3}{2\gamma_s} \frac{1}{\Delta \tau} \simeq c\bar{\alpha} - \kappa \, \bar{\alpha} \frac{\gamma_s \beta}{\ln^2(1/\alpha^2)} \,, \tag{7.37}$$

with κ some unknown fudge factor. We thus recover the qualitative behaviour in Eq. (7.34).

7.5 *Front diffusion: From geometric scaling to diffusive scaling*

SECT_DIFF

As emphasized in the previous subsection, the stochastic evolution leads to a *statistical ensemble* of fronts, all of them corresponding to the same initial condition at $\tau = 0$, but which differ from each other because of the fluctuations inherent in the evolution process. From this perspective, the front velocity computed in Eq. (7.34) should be viewed as the *average* velocity, with the "average" understood as an expectation value over the statistical ensemble. More generally, all the relevant physical quantities, so like the dipole scattering amplitudes, should be computed as expectation values over this ensemble, and to that aim one needs to know the corresponding probability distribution. In what follows, we

shall infer this distribution from studies of the sFKPP equation [188, 241] and then use the result for some applications to high–energy QCD [47, 48].

The crucial observation is that geometric scaling is, to a good approximation, satisfied for the *individual* fronts. Indeed, the discussion in the previous subsection shows that, except in the vicinity of the tip of the front (where the fluctuations are important), a particular front can be viewed as a traveling–wave: $T(\rho, \tau) \approx \mathcal{T}(\rho - \rho_s(\tau))$, where the shape function $\mathcal{T}(z)$ with $z = \rho - \rho_s(\tau)$ admits the following piece–parametrization:

$$\mathcal{T}(z) = \begin{cases} 1 & \text{for } z < 0 \\ A\,e^{-\gamma_s z}\,\dfrac{L}{\pi}\,\sin\dfrac{\pi z}{L} & \text{for } 0 < z < L \\ B\,e^{-(z-L)} & \text{for } z > L, \end{cases} \tag{7.38}$$

and $L \approx (1/\gamma_s) \ln(1/\alpha^2)$ is the front width, cf. Eq. (7.32). Note that, throughout this section, we use notations adapted to QCD: $T(\rho, \tau)$ is the event–by-event scattering amplitude for a dipole with logarithmic size $\rho = \ln(1/r^2 Q_0^2)$, with Q_0 an arbitrary scale of reference, and for a rapidity gap τ between the dipole and the target. The coefficients A and B are undetermined constants, and the tail of amplitude at large $z > L$ is the *color–transparency* tail, $T \sim e^{-\rho} \sim r^2$, specific to QCD, and not the sFKPP tail in Eq. (7.31). This change in the tail appears because, for very small dipole sizes, the BFKL evolution matches onto the DGLAP one (in the DLA approximation) and the 'diffusion approximation' to the BFKL kernel (which, we recall, was used to reduce Eq. (7.17) to the sFKPP equation (7.19)) is not justified any longer. Strictly speaking, this dilute tail is further modified by the DLA evolution and also by fluctuations, but this needs not be specified here in more detail since, as we shall see, the tail does not contribute to the most interesting observables.

Because of geometric scaling, a front is characterized by a unique 'parameter' — the position ρ_s of the front, or the associated saturation momentum $Q_s^2 = Q_0^2 e^{\rho_s}$ —, which for the stochastic evolution is a *random variable* — it can vary from one event to another. Accordingly, the statistics of the fronts reduces to that of a single variable: ρ_s. Moreover, from studies of the reaction–diffusion process [241, 242] one knows that, within a wide range of values for ρ, which covers the most interesting physical situations (see below), and for sufficiently large values of τ (larger than the front 'formation time' in Eq. (7.36)), the corresponding probability distribution is simply a Gaussian, and thus is characterized by only two parameters: the expectation value $\langle \rho_s \rangle$, which is fixed by the discussion in the previous subsection as

$$\langle \rho_s \rangle_\tau \simeq \lambda \tau, \tag{7.39}$$

(we have chosen the arbitrary scale Q_0 as the original saturation scale at $\tau = 0$), and the *front dispersion*, which is found to increase linearly with τ :

$$\sigma^2(\tau) \equiv \langle \rho_s^2 \rangle_\tau - \langle \rho_s \rangle_\tau^2 \simeq D\,\tau. \tag{7.40}$$

One may say that, with increasing τ, the position ρ_s of the front executes a random walk superposed over an uniform progression with velocity λ. This random walk is, clearly, an
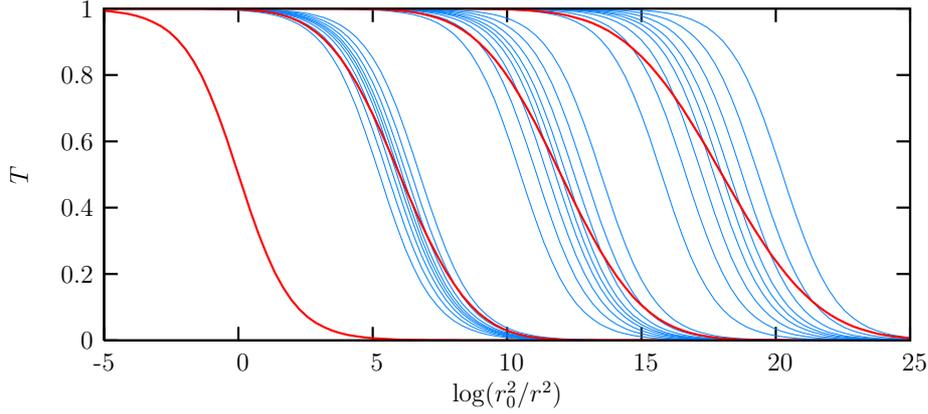
Fig. 55. | DISPERSION |

effect of the fluctuations, so the dispersion should vanish in the mean–field limit $N \equiv 1/\alpha^2 \to \infty$. And, indeed, one has found (first via numerical simulations [188, 240], which more recently have been confirmed via an analytic argument [241]) that the *front diffusion coefficient* D vanishes when $N \to \infty$, albeit (again !) very slowly, namely like $1/\ln^3 N$. Translating this result to QCD, one finds

$$D \simeq \frac{\bar{\alpha} d}{\ln^3(1/\alpha^2)} \qquad \text{for} \qquad \alpha \lll 1, \tag{7.41}$$

where the coefficient d is also known [241] (When specialized to QCD, the result for d in Ref. [241] turns out to be uncomfortably large.)

This diffusive wandering of the fronts is illustrated in Figs. **??** and 55. Fig. **??** has been obtained [47] via numerical simulations within the discrete particle model of Ref. [188], but a similar situation is expected in QCD as well. (In fact, similar numerical results have been obtained in Refs. [234, 236] for statistical models which include more of the actual QCD dynamics.) All the fronts represented there are different realizations of the same evolution; that is, they have been obtained by evolving the same initial condition over the same period of time. The dispersion of ρ_s is manifest in this picture, and so is also the universality of the shape of the individual fronts for T. However, precisely because of the dispersion, the shape of the *average* amplitude $\langle T(\rho) \rangle_\tau$ (represented by the thick line in Fig. **??**) is significantly different from the shape of the individual fronts. Besides, this average shape keeps changing with τ: when increasing τ, the dispersion increases, so the average front $\langle T(\rho) \rangle_\tau$ becomes flatter and flatter, as schematically illustrated in Fig. 55 (see also Fig. 57 below). This means that, although valid for the individual fronts, the geometric scaling is in fact violated at the level of the *average* amplitude [46, 47].

To study this violation in more detail, let us explicitly compute the average amplitude, which in the present context is obtained by averaging the event–by–event amplitude $T(\rho) \approx \mathcal{T}(\rho - \rho_s)$ over ρ_s, with the Gaussian probability distribution alluded to above :

$$\langle T(\rho) \rangle_\tau = \int\limits_{-\infty}^{\infty} \mathrm{d}\rho_s \, P(\rho_s, \tau) \, \mathcal{T}(\rho - \rho_s), \tag{7.42}$$

| averag

210

with $T(\rho - \rho_s)$ given by Eq. (7.38) and

$$P(\rho_s, \tau) \equiv \frac{1}{\sqrt{2\pi}\,\sigma(\tau)}\,\exp\left[-\frac{(\rho_s - \langle\rho_s\rangle_\tau)^2}{2\sigma^2(\tau)}\right]. \qquad (7.43)$$

One can simplify the calculation, while retaining its most salient features, by replacing Eq. (7.38) with

$$T(\rho - \rho_s) = \Theta(\rho_s - \rho) + \Theta(\rho - \rho_s)\,e^{-\gamma(\rho - \rho_s)}. \qquad (7.44)$$

For $\gamma = \gamma_s \approx 0.63$, this simulates the front region in the second line of Eq. (7.38), while for $\gamma = 1$ it rather corresponds to the 'color transparency' region in the third line there. By using Eqs. (7.44) and (7.43) it is straightforward to deduce

$$\langle T(z)\rangle_\tau = \frac{1}{2}\,\mathrm{Erfc}\left(\frac{z}{\sqrt{2}\sigma}\right) + \frac{1}{2}\exp\left(\frac{\gamma^2\sigma^2}{2} - \gamma z\right)\left[2 - \mathrm{Erfc}\left(\frac{z}{\sqrt{2}\sigma} - \frac{\gamma\sigma}{\sqrt{2}}\right)\right], \qquad (7.45)$$

where here and from now on $z \equiv \rho - \langle\rho_s\rangle_\tau$, and $\mathrm{Erfc}(x)$ is the complimentary error function:

$$\mathrm{Erfc}(x) \equiv \frac{2}{\sqrt{\pi}}\int_x^\infty dt\,e^{-t^2} \simeq
\begin{cases}
2 - \dfrac{\exp(-x^2)}{\sqrt{\pi}\,|x|} & \text{for } x \ll -1 \\[2mm]
1 & \text{for } |x| \ll 1 \\[2mm]
\dfrac{\exp(-x^2)}{\sqrt{\pi}x} & \text{for } x \gg 1.
\end{cases} \qquad (7.46)$$

The first term in the r.h.s. of Eq. (7.45) comes from the 'saturated' piece, $\Theta(\rho_s - \rho)$, of the individual front (7.44), while the second term comes from its exponential tail at $\rho > \rho_s$. The behaviour of the average amplitude depends upon the competition between σ (the width of the Gaussian distribution of the fronts) and $1/\gamma \sim \mathcal{O}(1)$, which characterizes the exponential decay of the individual fronts. We can thus distinguish between two types of behaviour, one at *intermediate* energies, the other one at (very) *high* energies:

i) When $\sigma \ll 1/\gamma$ (the situation in the early or the intermediate stages of the evolution), the Gaussian ensemble is strongly peaked around the average front, hence the average amplitude retains the single event profile (except in the short interval $|z| \lesssim \sigma$ where it gets smoothed), and thus shows *geometric scaling*:

$$\langle T(z)\rangle_\tau \approx T(z). \qquad (7.47)$$

In practice, the region for geometric scaling is limited by BFKL diffusion, as explained in Sect. 4.3, and is anyway restricted to the width L of the front, cf. Eq. (7.32).

ii) When $\sigma \gg 1/\gamma$ (the situation at sufficiently high energy), then the first term in Eq. (7.45) dominates up to very large values of z, namely for $z \ll \gamma\sigma^2$:

$$\langle T(z)\rangle_\tau \simeq \int_\rho^\infty d\rho_s\,P(\rho_s, \tau) = \frac{1}{2}\,\mathrm{Erfc}\left(\frac{z}{\sqrt{2}\sigma}\right) \qquad \text{for} \quad -\infty < z \ll \gamma\sigma^2, \qquad (7.48)$$

where the neglected terms are suppressed by, at least, one power of $1/(\gamma\sigma)$ and/or $z/(\gamma\sigma^2)$. In this regime, the average amplitude is dominated by gluon configurations in the target wavefunction which have $\rho_s \geq \rho$ and therefore appear as 'black' ($T = 1$) on the resolution scale of the projectile. That is, if one fixes a point ρ on the abscissa axis in Fig. **??**, then the average amplitude $\langle T(\rho)\rangle$ is dominated by those fronts in the ensemble which are ahead of ρ. This is true, in particular, for $\rho \gg \langle\rho_s\rangle$, i.e., for relatively small projectile dipoles, for which the *average* scattering amplitude is weak, $\langle T(\rho)\rangle \ll 1$, yet this average is dominated by the *rare configurations* with $\rho_s \geq \rho$, which are at saturation for that (large) value of ρ. (A configuration is 'rare' when $|\rho_s - \langle\rho_s\rangle| \gg \sigma$, since in that case $P(\rho_s, \tau) \ll 1$.) This is so because such rare–but–dense configurations yield large contributions, $T \sim \mathcal{O}(\infty)$, to the scattering amplitude, whereas the respective contributions of the *typical* configurations — those with $\rho_s \sim \langle\rho_s\rangle$ — are exponentially suppressed, cf. Eq. (7.44). This dominance of the 'black' configurations has important physical consequences, that we now discuss.

As manifest on Eq. (7.48), the high–energy amplitude takes a relatively simple form, which is *universal* [48] — i.e., independent of the initial conditions at low energy, and also of the details of the evolution (so like the BFKL anomalous dimension) —, and which *scales* as a function of z/σ [47] — it depends upon the two kinematical variables ρ and τ only via the ratio $(\rho - \langle\rho_s\rangle)/\sigma$, where $\langle\rho_s\rangle = \lambda\tau$ and $\sigma^2 = D\tau$. This new scaling law, known as *diffusive scaling* [243], emerges at sufficiently high energy, such that $\sigma \gg 1$, and holds within a wide region in ρ, which extends well above the average saturation line — namely, up to $\rho_D(\tau) \simeq \langle\rho_s\rangle + \gamma\sigma^2 = (\lambda + \gamma D)\tau$.

This discussion suggests the 'phase–diagram' for high–energy evolution depicted in Fig. 56, whose structure is considerably richer than it was in the MFA: in addition to the 'geometric scaling' window, which persists at not too high energies ($\sigma \ll 1$), and which is now limited by the finite width L of the front, there is an entirely new region which opens up when $\sigma \gtrsim 1$, where the average amplitude shows diffusive scaling, and which extends more and more in the 'dilute' regime with increasing τ. The quotation marks on 'dilute' are intended to remind that the large–ρ part of this diffusive scaling region is dilute *only on the average*, but in reality this region is characterized by *strong density fluctuations*, whose effects are measurable — e.g., the scattering of a small projectile dipole is controlled by the rare fluctuations which are at saturation.

Since determined by such a different physics, the amplitude (7.48) has other interesting properties, which are revealed by a more detailed analysis. (The subsequent points are also illustrated by Fig. 57, which represents the complete average amplitude, Eq. (7.45), as a function of z for various values of σ.) Eqs. (7.48) and (7.46) imply the following piecewise behaviour:

$$\langle T(z)\rangle_\tau \approx \begin{cases} 1 & \text{for } z \ll -\sigma \\[2mm] \dfrac{1}{2} & \text{for } -\sigma \ll z \ll \sigma \\[2mm] \dfrac{\sigma}{\sqrt{2\pi}z}\exp\left(-\dfrac{z^2}{2\sigma^2}\right) & \text{for } \sigma \ll z \ll \gamma\sigma^2\,, \end{cases} \qquad (7.49)$$

which is remarkable in several respects: (a) The approach towards the unitarity limit at
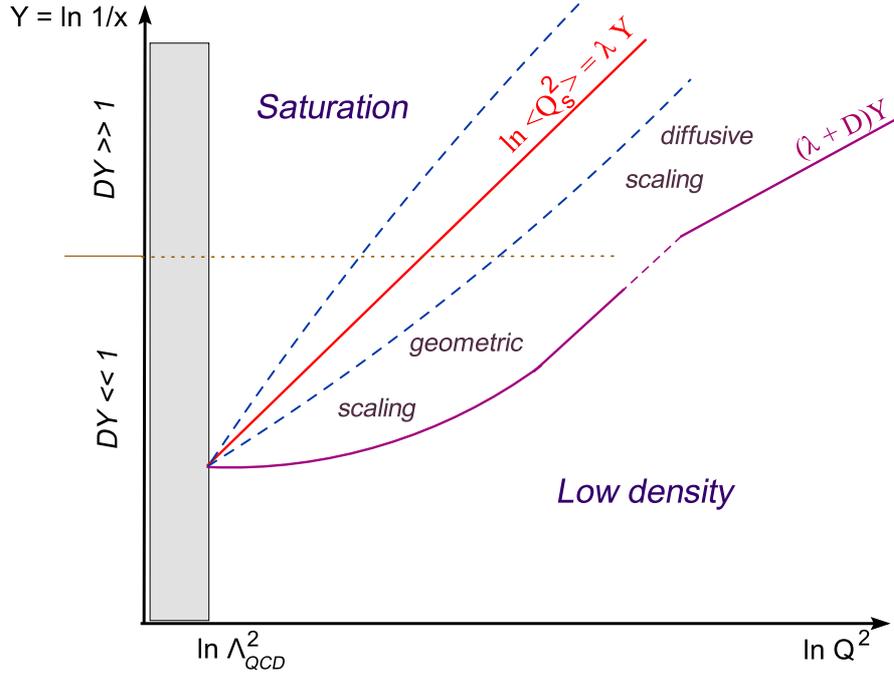
Fig. 56. *A 'phase–diagram' for the high–energy evolution with Pomeron loops. Shown are the average saturation line and the approximate boundaries of the scaling regions at large values of $\rho \sim \ln Q^2$. Note the gradual transition with increasing Y from geometric scaling at intermediate energies ($DY \ll 1$) to diffusive scaling at very high energies ($DY \gg 1$). This transition can be more quantitatively studied on Eq. (7.45).* $\boxed{\texttt{phase}}$

negative values of z (i.e., $\rho < \langle\rho_s\rangle$) is extremely slow, the more so the larger σ is. In fact, when $\sigma \gg 1$, there is a large interval, of width $\sim \mathcal{O}(\sigma)$, around the average saturation line $\langle\rho_s\rangle = \lambda\tau$ within which the amplitude is almost constant and equal to $1/2$. This interval is recognized as the 'diffusive radius' for the probability distribution (7.43), as represented in Fig. 56. (b) For relatively large values of ρ, or small incoming dipoles, there is a large window at $\sigma \ll \rho - \langle\rho_s\rangle \ll \gamma\sigma^2$ within which the average amplitude is very small, $\langle T(\rho)\rangle_\tau \ll 1$, yet it is radically different from the respective prediction of the BFKL equation, or of any other approximation assuming single scattering. For instance, there is no trace of 'color transparency' (not even modified by some 'anomalous dimension'), rather the amplitude decreases with $\rho \sim \ln(1/r^2)$ as a Gaussian [37] in $(\rho - \langle\rho_s\rangle)/\sigma$. Also, there is no Pomeron–like behaviour, i.e., no exponential growth of the amplitude when increasing τ at fixed ρ; rather, the expression in the last line of Eq. (7.49) rises quite slowly with τ.

Of course, the usual 'leading–twist' behaviour is eventually recovered, but only at values of ρ much larger than naively expected — namely, for $\rho \gtrsim \rho_D(\tau)$ in Fig. 56. Indeed, it is easy to check that for $z \gtrsim \gamma\sigma^2$, the dominant contribution to the average amplitude (7.45) comes from the second term in that equation, as generated by the tails of the individual

---

[37] Hence, this amplitude neither admits a 'twist–expansion', i.e., an expansion in powers of r^2, which would correspond to the series of multiple scattering.
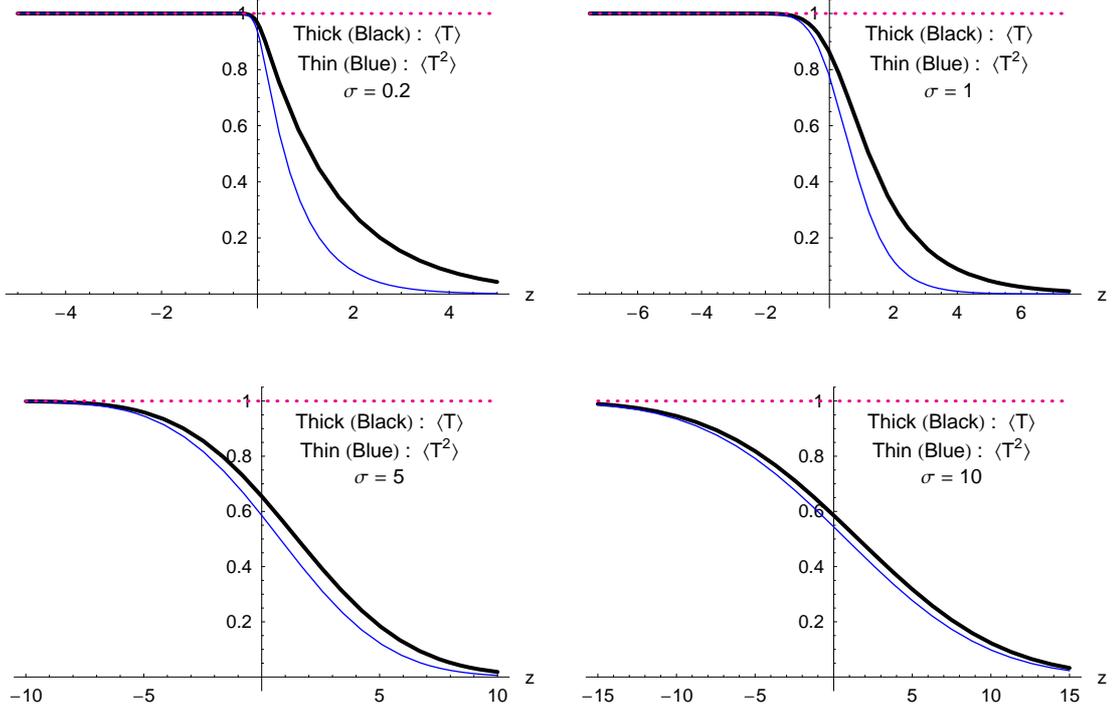
Fig. 57. *Evolution of $\langle T \rangle_\tau$ and $\langle T^2 \rangle_\tau$ with increasing σ. From Ref. [48].* `sigma`

fronts, and can be estimated as

$$\langle T(z) \rangle_\tau \simeq \mathrm{e}^{-\gamma z} \; \exp\left(\frac{\gamma^2 \sigma^2}{2}\right) \qquad \text{for} \qquad z \gg \gamma \sigma^2. \tag{7.50}$$

`TaveL`

As anticipated, this exhibits color transparency (for $\gamma = 1$) and also a Pomeron–like increase with τ: $\langle T(z) \rangle_\tau \sim \mathrm{e}^{\omega \tau}$, with $\omega = \lambda + D/2$.

Returning to 'diffusive scaling' region at $\rho < \rho_D(\tau)$, the dominance of the saturated configurations has another interesting consequence, which sheds some more light on the breakdown of the MFA: Since the dipole scattering is controlled by 'black' configurations, which individually yield $T \simeq 1$, it follows that the average *two*–dipole amplitude $\langle T^2(\rho) \rangle_\tau$ — i.e., the amplitude for the simultaneous scattering of two incoming dipoles with equal sizes and impact parameters — is the same as the single–dipole amplitude $\langle T(\rho) \rangle_\tau$ previously computed in Eq. (7.48). Indeed, event–by–event one has $T^2 \simeq T \simeq 1$ (for the relevant configurations), which after averaging yields $\langle T^2(\rho) \rangle_\tau \simeq \langle T(\rho) \rangle_\tau \simeq$ the total probability to find 'black' configurations for a given dipole resolution ρ, cf. Eq. (7.48). This is true, in particular, for relatively large values of ρ, for which $\langle T(\rho) \rangle_\tau \ll 1$, in which case it strongly contradicts the mean–field factorization $\langle T^2 \rangle \simeq \langle T \rangle^2$. We thus see that, due to fluctuations, the unitarity corrections (here, in the form of the two–dipole scattering) become important already in the weak scattering regime.

To illustrate this behaviour, we have displayed in Fig. 57 the functions $\langle T(z) \rangle_\tau$ and $\langle T^2(z) \rangle_\tau$, as computed from Eq. (7.45) together with its analogue for a two–dipole projectile (the latter is obtained by replacing $\gamma \to 2\gamma$ in Eq. (7.45)). As expected, the two

214
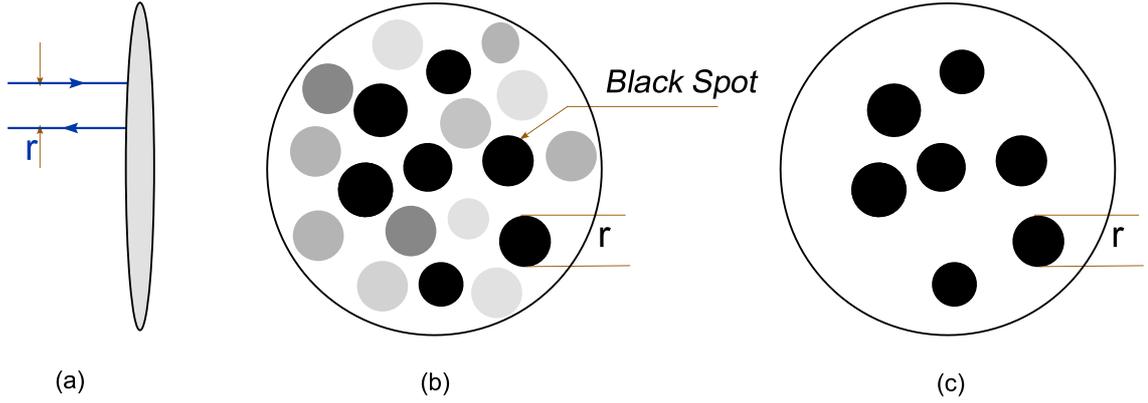
Fig. 58. *Dipole–hadron scattering in the fluctuation–dominated regime at $\sigma^2 \gg 1$. (a) a view along the collision axis; (b) a transverse view of the hadron, as 'seen' by a small dipole impinging at different impact parameters; (c) the simplified, black&white, picture of the hadron which is relevant for the average dipole amplitude.* `HotSpot`

curves converge towards each other with increasing σ. The above result can be easily generalized to N dipoles with different sizes (but the same impact parameter): under the same conditions as for Eq. (7.48), one finds

$$\langle T(\rho_1)T(\rho_2)\cdots T(\rho_N)\rangle_\tau \simeq \frac{1}{2}\operatorname{Erfc}\left(\frac{z_>}{\sqrt{2}\sigma}\right)\,, \tag{7.51}$$

where $z_> = \rho_> - \langle\rho_s\rangle$, with $\rho_> =$ the largest among the ρ_i's.

Although there was no explicit impact–parameter dependence in our discussion so far (recall that the $\boldsymbol{b}$–dependence has been washed out by the coarse–graining leading from the original Pomeron loop equations to the sFKPP equation), one can nevertheless rely on the previous arguments to deduce a qualitative picture of the hadron structure in the transverse space, as produced by the high–energy evolution [243, 244]. To that aim, we shall use the fact that, as argued in Sect. 7.3, the evolution is *quasi–local* in $\boldsymbol{b}$, so regions which are well separated in $\boldsymbol{b}$ evolve independently from each other, and randomly generate the possible realizations of the stochastic evolution. In other terms, in order to build–up the statistical ensemble of fronts, one can either repeat the evolution many times at a fixed impact parameter (our previous strategy in this subsection), or simultaneously consider the results produced by a single evolution at all the impact parameters — a point of view that we shall adopt in what follows. At sufficiently high energy, when the dispersion in the ensemble is large, $\sigma \gg 1$, we expect to see strong density fluctuations from one impact parameter to another: *the event–by–event picture of a high–energy hadron should be extremely inhomogeneous*. Note that these inhomogeneities occur on very short distances, of the order of the (local) saturation length $1/Q_s(\boldsymbol{b})$, and thus have nothing to do with the 'soft' (i.e., long–range) inhomogeneities introduced by the non–perturbative initial conditions at low energy. The 'granularity' of a high–energy hadron on short distance scales is rather a consequence of the (perturbative) evolution alone.

In order to explore this structure, we shall use, as usual, the scattering with a dipole

215

projectile (see also Fig. 58). For a fixed dipole size r and various values of its impact parameter $\boldsymbol{b}$, the dipole will 'see' spots of area r^2, which look either 'black' (when the local gluon density is so high that the dipole is completely absorbed: $T(r, \boldsymbol{b}) = 1$; this in turn requires $Q_s^2(\boldsymbol{b}) \gg 1/r^2$), or 'grey' (low gluon density, or weak scattering: $T(r, \boldsymbol{b}) \ll 1$), or even 'white' ($T(r, \boldsymbol{b}) \simeq 0$). The most interesting situation is that of a relatively small dipole, for which $\rho \equiv \ln(1/r^2) \gg \langle \rho_s \rangle_\tau$. Then, the target looks 'grey' (or 'white') at most impact parameters (since the typical gluon configurations have $\rho_s \ll \rho$), and the *average* scattering amplitude is small, $\langle T(\rho) \rangle_\tau \ll 1$, yet this average is entirely determined by the relatively rare 'black spots' which yield $T \sim 1$. Hence, in so far as the dipole scattering is concerned, the hadron looks effectively black or white, with most regions being white. This is illustrated by the comparison between Figs. 58.b and c.

Now, clearly, this qualitative picture does not tell us anything about the actual distribution of the density fluctuations in impact–parameter space. The underlying assumption, namely that different regions are totally independent from each other, cannot be exactly right: neither the high–energy evolution (e.g., the dipole splitting kernel), nor the dipole scattering, are exactly local, rather they are characterized by power–law tails in $\boldsymbol{b}$, which lead to long–range correlations between different domains. For instance, two 'spots' which had a common ancestor in the early stages of the evolution are *a priori* correlated with each other. Such correlations have been explicitly computed within the BFKL approximation [245], but it is not clear to which extent they survive the saturation effects.

Such issues could be clarified, in principle, by obtaining more complete solutions to the Pomeron loop equations, which include the dependence on $\boldsymbol{b}$. Since, however, this objective looks too ambitious for the time being, it is interesting to notice here an alternative strategy [246], which is to construct an *effective theory* describing the $\boldsymbol{b}$–dependence of the saturation momentum on the basis of symmetry arguments alone. The theory proposed in Ref. [246] is a 2–dimensional field theory for the field $\rho_s(\boldsymbol{b})$ (a generalization of the $\boldsymbol{b}$–independent probability distribution in Eq. (7.43)), with an action which is closely related to the Liouville action — itself proposed as an effective theory for quantum gravity in two dimensions [247] (see the recent review paper [248]).

7.6 Measuring the black spots: DIS, pp, pA,...

In this final subsection, we shall describe some physical consequences of the evolution with Pomeron loops which could be experimentally measured. They refer to observables which have been already introduced in the previous sections, namely the cross–sections for inclusive and diffractive (quasi–elastic) deep inelastic scattering (DIS), and that for inclusive gluon production in hadron–hadron collisions (with emphasis on forward rapidities). As we have seen in Sect. 5, these quantities have in common the fact they can be expressed in terms of the dipole–target forward scattering amplitude, and thus they can be computed within the formalism developed so far. Generalized "k_T–factorizations" of this type, which allows for the inclusion of multiple scattering effects through dipole amplitudes (generally, for a set of several dipoles) are known to hold for *asymmetric* collisions, between a system which is dense and one which is dilute. This is the case of DIS in the dipole frame, of the pA collisions at RHIC at LHC, and also of the pp collisions when

216

viewed in an asymmetric frame (say, in fixed target experiments). If, on the other hand, the *pp* collisions are viewed in the center–of–mass frame, so like at LHC, then the present formalism can be used only for the particle production at sufficiently forward rapidities, where one of the participating protons (the 'target') looks much denser than the other one (the 'projectile'), which is still dilute (cf. Sect. 5.2).

In the previous subsection, we have seen that, at sufficiently high energy, the dipole amplitude is strongly modified by gluon–number fluctuations in the target wavefunction. These modifications will naturally transmit to the related observables that we shall now consider. In particular, we shall found that, for high enough energies, the cross–sections are expected to exhibit *diffusive* (rather than *geometric*) scaling. Thus, the fact that the current data at HERA or RHIC appear to be consistent with geometric scaling (cf. Sect. 5) implies that these experiments correspond to *intermediate* (rather than *high*) energies in the terminology of Fig. 56. But the genuine high–energy regime could be accessible at LHC. In principle, the theory is able to predict the boundary between these two regimes: this occurs at $D\tau \sim \mathcal{O}(\infty)$, with D the diffusion coefficient which characterizes the dispersion of the fronts due to fluctuations. However, by lack of explicit solutions to the Pomeron loop equations, the value of D remains unknown, so like the value of the saturation exponent λ beyond the mean–field approximation. Besides, the present formalism being a leading–order one, the corresponding *quantitative* predictions are anyway not reliable for applications to the phenomenology. To cope with that, in what follows we shall treat the coefficients D and λ as *free parameters*, in terms of which we shall express the interesting cross–sections. Hence, our analysis will be, at most, *qualitative*. Yet, as we shall see, this analysis predicts some interesting phenomena (like the transition from geometric to diffusive scaling with increasing rapidity, or a total breakdown of the collinear factorization up to very large values of Q^2, well above the average saturation momentum), which could be well observed in the data.

7.6.1 DIS revisited

As explained in Sects. 1.2.2 and 5.1, DIS is most conveniently computed in the 'dipole frame', in which the proton carries most of the total energy — and hence the totality of the high–energy evolution, which now includes the Pomeron loops —, whereas the virtual photon can be replaced by a $q\bar{q}$ fluctuation, or 'color dipole'. Then, the 'inclusive' cross–section σ_{γ^*p} (the cross–section for the absorbtion of the virtual photon by the hadronic target) is related to the dipole total cross–section σ_{dipole} as shown in Eq. (5.1). In turn, the latter is determined by the (average) forward dipole amplitude $\langle T(r)\rangle_\tau$, which is now assumed to be a solution of the 'Pomeron loop' equations, and thus to include the effects of fluctuations. (The corresponding process is pictorially illustrated in Fig. 59.) These effects are most spectacular at relatively large values of Q^2, well above the average saturation momentum $\langle Q_s^2\rangle_\tau \equiv Q_0^2 e^{\langle \rho_s\rangle_\tau}$ (but inside the diffusive–scaling region), which is the range that we shall focus on in what follows. For simplicity, we shall ignore the details of the impact–parameter dependence, so like in Sect. 5.1, and write $\sigma_{\text{dipole}}(\tau, r) = \sigma_0\langle T(r)\rangle_\tau$, with $\sigma_0 \equiv 2\pi R^2$ a 'geometric' cross–section. Also, we shall use the approximate expression (7.45) (or even (7.48), when appropriate) for the average dipole amplitude.

In Sect. 5.1 we found that, for $Q^2 \gg Q_s^2(\tau)$, $\sigma_{\gamma^*p}(\tau, Q^2)$ is dominated by relatively small
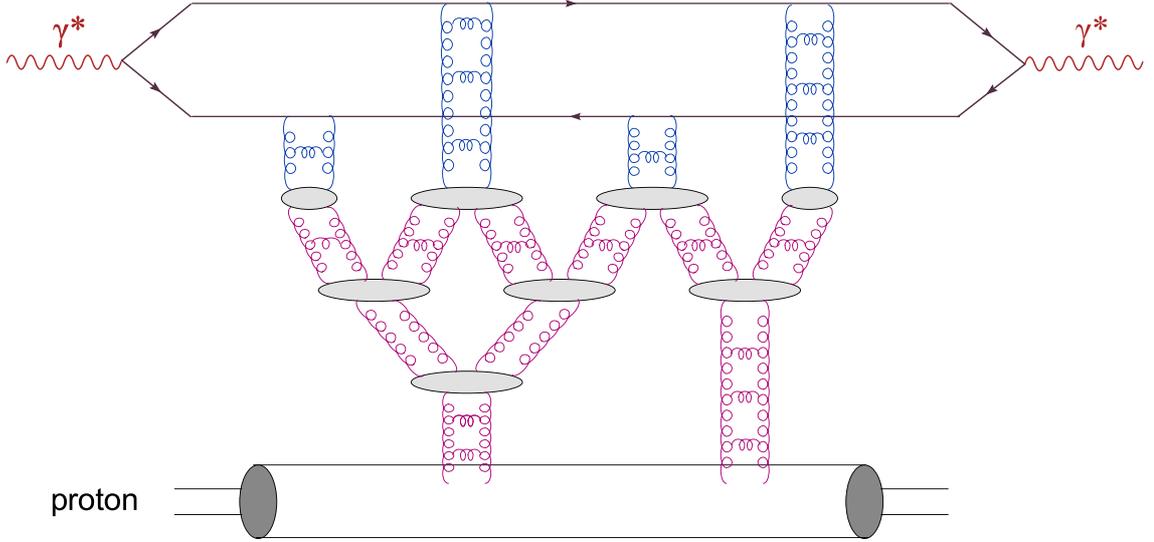
Fig. 59. | PloopDIS |

dipole sizes, $2/Q \lesssim r \lesssim 1/Q_s$, and as a consequence of that it shows approximate geometric scaling. In particular, in the 'leading–twist' regime at very large Q^2, the leading contribution — which is enhanced by a large logarithm $\ln(Q^2/Q_s^2)$ with respect to the subleading ones; see Eq. (5.11) — is generated by the scattering of *aligned–jet* dipole configurations in the *transverse* sector. These results did not include fluctuations, so they are expected to hold at moderately large energies, such that [38] $\sigma^2 \equiv D\tau \ll 1$. But for higher energies, where $D\tau \gg 1$, they are totally modified by fluctuations, as we show now:

It is convenient to denote $Z \equiv \ln(Q^2/\langle Q_s^2 \rangle)$, with $\langle Q_s^2 \rangle_\tau \equiv Q_0^2 e^{\lambda\tau}$. The most interesting regime at high energy and large Q^2 is the one characterized by $\sigma \ll Z \ll \sigma^2$. In this regime, one can check [243] that (i) the convolution in Eq. (5.1) is *strongly* dominated by dipole sizes $r \sim 1/Q$, (ii) the leading contribution comes from *symmetric* dipole configurations in the *transverse* sector, and (iii) the final result shows *diffusive scaling*, i.e., it depends upon Q^2 and τ only via the single variable Z/σ (after a trivial rescaling).

To emphasize the first point, we have displayed in Fig. 60 [see the curves denoted as "incl" in that figure; the "diff" curves will be discussed later] the integrand in Eq. (5.1) as a function of rQ_s, for $Q^2 = 100Q_s^2$ and for two scenarios for the scattering amplitude: (I) the "fluctuation" amplitude (7.45) with $\sigma = 5$, and (II) the "mean–field" amplitude (7.44). Case (II) can be seen as the limit of case (I) when $\sigma \to 0$. As manifest on this figure, the strength of the integration is pushed towards smaller dipole sizes after including the effects of fluctuations: the respective integrand is strongly peaked at $r \sim 2/Q$, while that for the mean field approximation is rather smoothly distributed at all sizes $r \gtrsim 2/Q$.

Since controlled by $r \sim 1/Q$, the cross–section (5.1) corresponding to $\sigma \ll Z \ll \sigma^2$ can be

---

[38] Note that, throughout this section, the notation σ, eventually accompanied by a subscript, will be used for two different types of quantities, namely, for the front dispersion due to fluctuations, cf. Eq. (7.40), and for the various cross–sections. The precise meaning should be clear from the context, and also by inspection of the accompanying subscript, when appropriate.

$$Q^2 = 100\,Q_s^2$$
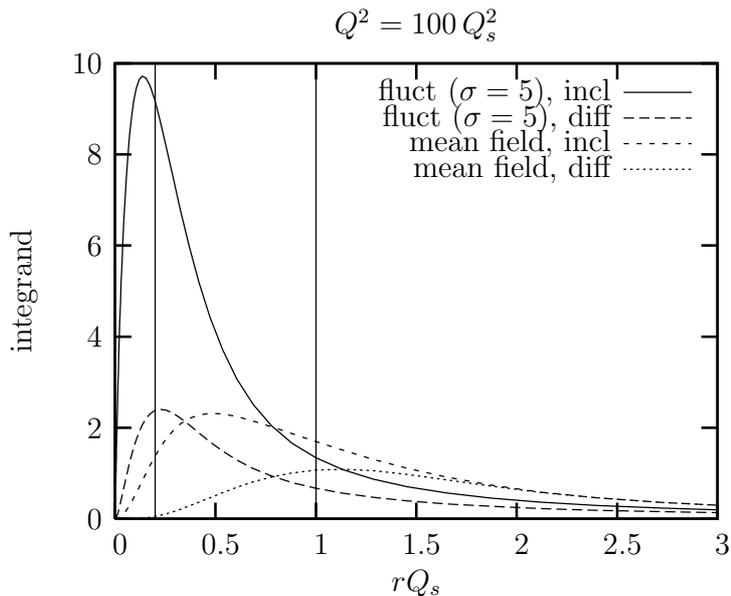
Fig. 60. *The integrands in Eqs. (5.1) and (??) after the integration over z and the azimuthal angle, as computed with two different expressions for the dipole amplitude — the mean–field amplitude (7.44) with $\gamma = 1$ and the 'fluctuation' amplitude (7.45) with $\sigma = 5$ [both with $\gamma = 1$] — are plotted as functions of rQ_s for $Q^2 = 100Q_s^2$. The vertical line ar $rQ_s = 0.2$ corresponds to $r = 2/Q$.* `integrand`

estimated by using the approximate form (7.48) for $\langle T(r)\rangle_\tau$, together with the simplified integrals in Eqs. (5.5)–(5.6). In the transverse sector, one finds [as in Sect. 5.1, we focus on order–of–magnitude estimates and omit the overall factor $\alpha_{\rm em} N_c \sum_f e_f^2$]

$$\frac{\sigma_T}{\sigma_0} \sim \frac{\sigma^3}{Z^2} \exp\left(-\frac{Z^2}{2\sigma^2}\right) + \frac{\sigma}{Z} \exp\left(-\frac{Z^2}{2\sigma^2}\right) \qquad \text{for} \qquad \sigma \ll Z \ll \sigma^2, \qquad (7.52)$$

where the first (second) term in the r.h.s. corresponds to symmetric (aligned–jet) configurations, respectively. As anticipated, the 'symmetric' contribution dominates, by a relative factor $\sigma^2/Z \gg 1$. The longitudinal sector yields a contribution similar to the 'aligned–jet' piece in Eq. (7.52), which is therefore subleading.

Up to prefactors, all the contributions to σ_{γ^*p} in this high–energy regime show the same dominant behaviour at high Q^2, namely a Gaussian decrease with Z/σ. This Gaussian is simply the probability for the virtual photon to find a 'black spot' inside the proton, i.e., a (rare) gluon configuration which is saturated on the photon resolution scale Q^2 (cf. Fig. 58). Since built up as an average over 'black' events, the cross–section (7.52) admits no expansion in powers of $1/Q^2$; this demonstrates the *breakdown of the twist–expansion* in DIS at high energy. (The standard leading–twist result will be eventually recovered, but only for Q^2 so large that $Z \gtrsim \sigma^2$.) Another consequence of this physical mechanism, which is manifest too on Eq. (7.52), is *diffusive scaling*: up to prefactors, all the contributions to σ_{γ^*p} scale as functions of Z/σ. If one keeps the dominant contribution alone, i.e., the first term in Eq. (7.52), then one finds exact diffusive scaling (including the prefactor) after dividing out a factor of σ : $\sigma_{\gamma^*p}/\sigma = f(Z/\sigma)$ when $Z \ll \sigma^2$ [243].

But the effects of Pomeron loops on the DIS dynamics at high energy become even more striking when one considers the respective *diffractive* cross–section. By a 'diffractive process' in the context of high–energy scattering, one generally means a process characterized by a substantial rapidity gap between the particles produced after the collision in the fragmentation regions of the projectile and, respectively, the target. Here, we shall limit ourselves to the simplest such a process in DIS, namely a $\gamma^* p$ collision in which the $q\bar{q}$ fluctuation of the virtual photon undergoes elastic scattering off the proton, and then emerges as a 'vector meson'. This process is indeed diffractive, as the two emerging systems — the 'vector meson' and the (elastically scattered) proton — are separated by a rapidity τ (the same as the original rapidity separation between the projectile and the target), which is the maximal possible value for a rapidity gap [39]. The generalization of the subsequent discussion to more general diffractive processes in DIS, for which $\tau_{\mathrm{gap}} < \tau$, requires the inclusion of higher Fock states, like $q\bar{q}g$, etc., in the wavefunction of the virtual photon, and has been given in Ref. [243] (within the high–energy approximation). See also Refs. [73, 74, 76, 88, 167, 194, 195, 198, 249–261] for related studies, which however do not include the effects of Pomeron loops.

What makes diffractive DIS especially appealing in the context of high–energy QCD is the fact that the corresponding cross–section is dominated by relatively large dipole sizes, of the order of the unitarization scale in the problem. In the past, this led to the general wisdom that diffractive DIS should be a 'soft' (i.e., non–perturbative) process even for large Q^2. More recently, it has been recognized [73] that, at x small enough, the phenomenon of gluon saturation introduces a relatively hard unitarization scale — the saturation momentum —, so diffractive DIS is rather a 'semi–hard' process. As already mentioned in Sects. 1.2.2 and 5.1 (see also below), this saturation scenario has led to a rather successful description of the ensemble of the HERA data (including the diffractive ones) at $x \leq 0.01$. Yet, as we shall now explain, the situation changes again when going to very high energies : then, up to relatively large values of Q^2, the DIS diffraction is controlled by dipole sizes $r \sim 1/Q$, and thus is a truly 'hard' process [243].

To see this, consider the cross–section for the simple diffractive process alluded to above, in which the dipole scatters elastically. This is given by

$$
\sigma_{\mathrm{diff}}(\tau, Q^2) = \int_0^1 \mathrm{d}z \int \mathrm{d}^2\boldsymbol{r} \sum_{\alpha=L,T} |\Psi_\alpha(z, r; Q^2)|^2 \; \sigma_{\mathrm{el}}(\tau, r),
$$
$$
\sigma_{\mathrm{el}}(\tau, r) = \sigma_0 \left| \langle T(r) \rangle_\tau \right|^2, \tag{7.53}
$$

where we have again ignored the impact–parameter dependence of the dipole amplitude. We shall focus on the transverse contribution (this turns out to be the dominant piece), that we estimate with the help of Eq. (5.5) :

---

[39] This process is sometimes referred to as the "diffractive vector meson production", or as "quasi–elastic scattering". Note that this process is strictly speaking not elastic, since the virtual photon is not present in the final state.

$$\frac{\sigma_{\text{diff}}^T}{\sigma_0} \sim \int\limits_0^{2/Q} \frac{dr}{r}\, |\langle T(r)\rangle_\tau|^2 \; + \; \frac{1}{Q^2} \int\limits_{2/Q}^{\infty} \frac{dr}{r^3}\, |\langle T(r)\rangle_\tau|^2 \, . \tag{7.54}$$

The standard analysis of this expression proceeds as follows [note that this analysis will change at high energy]: For large Q^2 and $r \lesssim 2/Q$, $\langle T(r)\rangle \propto r^2$, so the first, 'symmetric', integral in Eq. (7.54) gives a higher–twist contribution $\propto 1/Q^4$, which is negligible. (The same is true for the whole contribution of the longitudinal sector.) Rather, the dominant contribution, of leading–twist order, comes from the second, 'aligned–jet', integral. By itself, the measure dr/r^3 in this integral (as introduced by the virtual–photon wave-function) would favor relatively small dipoles, with $r \sim 2/Q$, but this tendency is more than compensated by the rapid increase of the dipole amplitude squared at small r : $\langle T(r)\rangle^2 \propto r^4$. Accordingly, the integral is dominated by large dipole sizes, of the order of the unitarization scale at which the amplitude ceases to grow.

Without gluon saturation, the unitarization scale would be non–perturbative, say, of the order of the proton size R. Then, by using the estimate $T(r) \approx \alpha(r^2/R^2)xG(x,Q^2)$ (cf. Eq. (1.24)) and cutting the integration of $r = R$, one obtains

$$\frac{\sigma_{\text{diff}}^T}{\sigma_0} \sim \frac{\alpha^2}{Q^2 R^2}\left[xG(x,Q^2)\right]^2 \sim x^{-2(\alpha_{\text{soft}\mathbb{P}}-1)}\,, \tag{7.55}$$

which is the standard, 'Regge'–like, estimate for diffractive DIS. In this scenario, σ_{diff} would grow with $1/x$ as a *double soft Pomeron*, with the soft Pomeron intercept estimated as $\alpha_{\text{soft}\mathbb{P}} - 1 \approx 0.08$ from fits to total cross–sections at high–energy.

However, for sufficiently small values of x, (perturbative) gluon saturation becomes important, and the above calculation should be modified. By using $T(r) \approx (r^2 Q_s^2)^\gamma$, with $1/2 < \gamma \leq 1$ for $2/Q < r < 1/Q_s$ and $T = 1$ for $r > 1/Q_s$, one finds

$$\frac{\sigma_{\text{diff}}^T}{\sigma_0} \sim \frac{1}{Q^2} \int\limits_{2/Q}^{1/Q_s} \frac{dr}{r^3}\, (r^2 Q_s^2)^{2\gamma} + \frac{1}{Q^2} \int\limits_{1/Q_s}^{\infty} \frac{dr}{r^3} \sim \frac{Q_s^2(x)}{Q^2} \sim x^{-\lambda}\,, \tag{7.56}$$

where the final result, which is now fully perturbative (and also insensitive to the precise value of the 'anomalous dimension' $1 - \gamma$), comes from dipole sizes $r \sim 1/Q_s$, and thus from both integration domains shown above. This result rises with $1/x$ in the same way as $Q_s^2(x)$, i.e., like a *single hard Pomeron*. By using $\lambda \simeq 0.25 \div 0.30$, as inferred from the fits to the inclusive cross–section at HERA (cf. Sect. 5.1), we see that this growth is in fact faster than the one predicted by Eq. (7.55). Such a faster growth seems to be preferred by the diffractive data at HERA, which moreover show a remarkable feature — the ratio between the diffractive and the inclusive cross–sections appears to be roughly independent of the total energy —, which is consistent with the saturation scenario [73, 74], but not also with the Regge–theory one: Indeed, Eq. (7.56) together with our previous estimate, Eq. (5.11), for $\sigma_{\gamma^* p}$ imply

$$\frac{\sigma_{\text{diff}}}{\sigma_{\gamma^* p}} \sim \frac{1}{\ln(Q^2/Q_s^2(\tau))}\,, \tag{7.57}$$

which varies only logarithmically when changing the energy of the $\gamma^* p$ system.

Let us now move up in the energy, such that $\sigma^2 \gg 1$, while staying in the kinematical region at large Q^2 where the fluctuations are expected to be important: $\sigma \ll Z \ll \sigma^2$, with $Z \equiv \ln(Q^2/\langle Q_s^2 \rangle)$. Then, one assists at a qualitative change in behaviour, which can be understood as follows: Even though Q^2 is high, the contribution $T(r)$ of small dipoles with $r \sim 1/Q$ is not necessarily suppressed anymore, since the target also involves dense configurations with $Q_s^2 > Q^2$, for which $T(r \sim 1/Q) \simeq 1$. Although relatively rare, such configurations are strongly favored by the integral over r with the virtual photon wavefunction, cf. Eq. (7.54). Accordingly, this integral is now controlled by dipole sizes $r \sim 1/Q$, with a slight preference (visible at the level of the prefactors) towards the 'symmetric' dipole configurations — in full analogy to what happens for the inclusive cross–section, as found in Eq. (7.52). One eventually finds [243]

$$\frac{\sigma_{\text{diff}}}{\sigma_0} \sim \frac{\sigma^4}{Z^3} \exp\left(-\frac{Z^2}{\sigma^2}\right) + \frac{\sigma^2}{Z^2} \exp\left(-\frac{Z^2}{\sigma^2}\right) \qquad \text{for} \qquad \sigma \ll Z \ll \sigma^2, \qquad (7.58)$$

where the first term in the r.h.s., which is the dominant one, corresponds to symmetric configurations in the transverse sector, while the second term receives contributions from both the aligned–jet configurations in the transverse sector, and from the longitudinal sector. These results exhibit *diffusive scaling*, a property which also transmits to the diffractive/inclusive ratio at high energy, now estimated as

$$\frac{\sigma_{\text{diff}}}{\sigma_{\gamma^* p}} \sim \frac{\sigma}{Z} \exp\left(-\frac{Z^2}{\sigma^2}\right) \qquad \text{for} \qquad \sigma \ll Z \ll \sigma^2. \qquad (7.59)$$

7.6.2 Forward gluon production revisited: Preparing for LHC

In Sect. 1.3.2 and 5.2.2 we have discussed at length the relevance of the particle production at 'forward rapidities' in proton–nucleus collisions as an indicator of the saturation effects in the nuclear target. As emphasized there, the larger is the rapidity η of the produced particles, the smaller is the value x_1 at which one probes the gluon distribution in the target: $x_1 = (p_\perp/\sqrt{s})\, e^{-\eta}$, where s is the COM energy squared for the hadronic collision and $p_\perp$ and η refer to the produced particle (see the kinematics in Fig. 15). In that context, it has been convenient to consider the ratio $R_{pA}(\eta, p_\perp)$, Eq. (1.38), between the particle yield in pA and, respectively, pp collisions, with the latter serving as a reference level and being computed in the linear, 'leading–twist', approximation (cf. Sect. 5.2.2). This was appropriate for the kinematics at RHIC since the smallest values of x_1 which are accessible in that experiment are still large enough for the saturation effects in the proton target to remain negligible.

However, when moving to LHC, where $\sqrt{s} = 14$ TeV in pp collisions and the rapidity coverage will extend up $\eta_{max} = 6.6$ (in the CMS experiment), one will be able to probe values of x as small as $x_1 \approx 10^{-6}$ (corresponding to produced jets with $p_\perp \sim 10\,\text{GeV}$) in the target hadron wavefunction, and then the saturation effects should be important even in the proton. Moreover, by increasing η (i.e., decreasing x_1), one may be able to observe the onset of the fluctuation–dominated regime in the target wavefunction, and thus directly probe the transition from geometric to diffusive scaling. Similar effects are also expected

222

in the pA or AA collisions at the LHC, except that, in those cases, the 'Pomeron loop' effects should be less important because, first, the corresponding energies will be somewhat lower ($\sqrt{s_{NN}} = 8.8$ TeV in pA and, respectively, $\sqrt{s_{NN}} = 5.5$ TeV in AA) and, second, the presence of a high gluon density in the nuclear wavefunction already at low energy should somehow "enhance the mean–field behaviour", i.e., delay the development of the fluctuations in the early stages of the evolution — an effect which has not been captured by our previous analysis in Sect. 7.5, valid at asymptotic τ, but which could in principle be studied by solving the full Pomeron loop equations with nuclear (e.g., MV–like) initial conditions at low energy. This discussion suggests that, at LHC, pp collisions should be a better laboratory (as compared to pA or AA) for a study of Pomeron loop effects in the high–energy evolution. In what follows we shall therefore focus on the forward gluon production in pp collisions and describe its qualitative evolution with decreasing x_1, as expected from the previous analysis throughout this report.

8   Conclusions

References

Heisenberg
[1]   W. Heisenberg, *Z. Phys.* **133** (1952) 65.

Froissart
[2]   M. Froissart, *Phys. Rev.* **123** (1961) 1053.

Martin
[3]   Y.S. Jin and A. Martin, *Phys. Rev.* **135** (1964) B 1375; A. Martin, *Nuovo Cimento* **42** (1966) 930.

LLMartin
[4]   L. Lukaszuk and A. Martin, *Nuovo Cimento* **A52** (1967) 122.

Regge
[5]   T. Regge, *Nuovo Cimento* **14** (1959) 951; *ibid.* **18** (1960) 947.

CF61
[6]   G.F.P. Chew and S.C. Frautschi, *Phys. Rev. Lett.* **7** (1961) 395.

Gribov61
[7]   V.N. Gribov, *Zh. Eksp. Teor. Fiz* **41**, 667 (1961) (*Sov. Phys. JETP* **14** (1962) 478).

Basarab
[8]   L. Lukaszuk and B. Nicolescu, Lett. Nuovo Cim. **8** (1973) 405; D. Joynson, E. Leader, B. Nicolescu and C. Lopez, Nuovo Cim. A **30** (1975) 345.

BJ69
[9]   J.D. Bjorken, *Phys. Rev.* **179** (1969) 1547; J.D. Bjorken and E.A. Paschos, *Phys. Rev.* **185** (1969) 1975.

nmanParton
[10] R. Feynman, *Photon–Hadron Interactions*, Benjamin, Reading, Massachusetts, 1972.

SLACDIS
[11] A summary of the early SLAC experiments appears in J.I. Friedman and H.W. Kendall, *Ann. Rev. Nucl. Sci.* **22** (1972) 203.

GL72
[12] V.N. Gribov and L.N. Lipatov, *Sov. Journ. Nucl. Phys.* **15** (1972), 438.

GWP73
[13] D.J. Gross and F. Wilczek, *Phys. Rev. Lett.* **30** (1973) 1343; H.D. Politzer, *ibidem* p. 1346.

GW74
[14] H. Georgi and H.D. Politzer, *Phys. Rev.* **D9** (1974) 416; D.J. Gross and F. Wilczek, *ibidem* p. 980.

DGLAP
[15] V.N. Gribov and L.N. Lipatov, *Sov. Journ. Nucl. Phys.* **15** (1972), 438; G. Altarelli and G. Parisi, *Nucl. Phys.* **B126** (1977), 298; Yu. L. Dokshitzer, *Sov. Phys. JETP* **46** (1977), 641.

BFKL
[16] L.N. Lipatov, *Sov. J. Nucl. Phys.* **23** (1976) 338;
E.A. Kuraev, L.N. Lipatov and V.S. Fadin, *Zh. Eksp. Teor. Fiz* **72**, 3 (1977) (*Sov. Phys. JETP* **45** (1977) 199);
Ya.Ya. Balitsky and L.N. Lipatov, *Sov. J. Nucl. Phys.* **28** (1978) 822.

New_HERA
[17] ZEUS Collaboration, J. Breitweg et al., *Phys. Lett.* **B487** (2000) 53; ZEUS Collaboration, S. Chekanov et al., *Eur. Phys. J.* **C21** (2001) 443; H1 Collaboration, C. Adloff et al., *Eur. Phys. J.* **C21** (2001) 33.

NLBFKL
[18] V.S. Fadin and L.N. Lipatov, *Phys. Lett.* **B429** (1998) 127; G. Camici and M. Ciafaloni, *Phys. Lett.* **B430** (1998) 349.

Salam99
[19] G.P. Salam, **JHEP 9807** (1998) 19; M. Ciafaloni, D. Colferai, *Phys. Lett.* **B452** (1999) 372; M. Ciafaloni, D. Colferai, and G.P. Salam, *Phys. Rev.* **D60** (1999) 114036.

FRSV99
[20] J. R. Forshaw, D. A. Ross and A. Sabio Vera, *Phys. Lett.* **B455** (1999) 273.

ABF00   [21] G. Altarelli, R.D. Ball and S. Forte, *Nucl. Phys.* **B575** (2000) 313; *ibidem* **B599** (2001) 383; *ibidem* **B621** (2002) 359; *ibidem* **B674** (2003) 459.

GLR   [22] L.V. Gribov, E.M. Levin, and M.G. Ryskin, *Phys. Rept.* **100** (1983) 1.

MQ85   [23] A.H. Mueller and J. Qiu, *Nucl. Phys.* **B268** (1986) 427.

BM87   [24] J.-P. Blaizot and A.H. Mueller, *Nucl. Phys.* **B289** (1987) 847.

MV   [25] L. McLerran and R. Venugopalan, *Phys. Rev.* **D49** (1994) 2233; *ibid.* **49** (1994) 3352; *ibid.* **50** (1994) 2225.

QM02   [26] "Quark Matter 2002", Proceedings of the 16th International Conference on Ultra-Relativistic Nucleus-Nucleus Collisions, (Nantes, France 18-24 July, 2002), *Nucl. Phys.* **A 715** (2003).

QM04   [27] "Quark Matter 2004", Proceedings of the 17th International Conference on Ultra-Relativistic Nucleus-Nucleus Collisions, (Oakland, USA, 11-17 January 2004), *J. Phys.* **G30** (2004) S633-S1429.

K96   [28] Yu.V. Kovchegov, *Phys. Rev.* **D54** (1996), 5463; *Phys. Rev.* **D55** (1997), 5445.

JKMW97   [29] J. Jalilian-Marian, A. Kovner, L. McLerran, H. Weigert, *Phys. Rev.* **D55** (1997) 5414.

KM98   [30] Yu.V. Kovchegov and A.H. Mueller, *Nucl. Phys.* **B529** (1998), 451.

JKLW97   [31] J. Jalilian-Marian, A. Kovner, A. Leonidov and H. Weigert, *Nucl. Phys.* **B504** (1997) 415; *Phys. Rev.* **D59** (1999) 014014.

JKLW99   [32] J. Jalilian-Marian, A. Kovner and H. Weigert, *Phys. Rev.* **D59** (1999), 014015.

KM00   [33] A. Kovner and G. Milhano, *Phys. Rev.* **D61** (2000), 014012; A. Kovner, J. G. Milhano and H. Weigert, *Phys. Rev.* **D62** (2000), 114005.

RGE   [34] E. Iancu, A. Leonidov and L. McLerran, *Nucl. Phys.* **A692** (2001) 583; *Phys. Lett.* **B510** (2001) 133; E. Ferreiro, E. Iancu, A. Leonidov and L. McLerran, *Nucl. Phys.* **A703** (2002) 489.

W   [35] H. Weigert, *Nucl. Phys.* **A703** (2002) 823.

ODDERON   [36] Y. Hatta, E. Iancu, K. Itakura, L. McLerran, *Nucl. Phys.* **A760** (2005) 172, arXiv:hep-ph/0501171.

AM99   [37] A. H. Mueller, *Nucl. Phys.* **B558** (1999) 285.

SAT   [38] E. Iancu and L. McLerran, *Phys. Lett.* **B510** (2001) 145.

B   [39] I. Balitsky, *Nucl. Phys.* **B463** (1996) 99; *High-energy QCD and Wilson lines*, arXiv:hep-ph/0101042.

PATH   [40] J.-P. Blaizot, E. Iancu and H. Weigert, *Nucl. Phys.* **A713** (2003) 441.

AM02   [41] A. H. Mueller, *Nucl. Phys.* **B643** (2002) 501.

GAUSS   [42] E. Iancu, K. Itakura, and L. McLerran, *Nucl. Phys.* **A724** (2003) 181.

225

AM95	[43] A.H. Mueller, *Nucl. Phys.* **B437** (1995) 107.
Salam95	[44] G.P. Salam, *Nucl. Phys.* **B449** (1995) 589; *Nucl. Phys.* **B461** (1996) 512.
AMSalam96	[45] A.H. Mueller and G.P. Salam, *Nucl. Phys.* **B475** (1996) 293.
MS04	[46] A.H. Mueller and A.I. Shoshi, *Nucl. Phys.* **B692** (2004) 175.
IMM04	[47] E. Iancu, A.H. Mueller, and S. Munier, *Phys. Lett.* **B606** (2005) 342.
IT04	[48] E. Iancu and D.N. Triantafyllopoulos, *Nucl. Phys.* **A756** (2005) 419, arXiv:hep-ph/0411405.
IT05	[49] E. Iancu and D.N. Triantafyllopoulos, *Phys. Lett.* **B610** (2005) 253 [arXiv:hep-ph/0501193].
MSW05	[50] A.H. Mueller, A.I. Shoshi and S.M.H. Wong, *Nucl. Phys.* **B715** (2005) 440 [arXiv:hep-ph/0501088].
SCALING	[51] E. Iancu, K. Itakura, and L. McLerran, *Nucl. Phys.* **A708** (2002) 327.
MT02	[52] A.H. Mueller and D.N. Triantafyllopoulos, *Nucl. Phys.* **B640** (2002) 331.
DT02	[53] D.N. Triantafyllopoulos, *Nucl. Phys.* **B648** (2003) 293.
IIM03	[54] E. Iancu, K. Itakura and S. Munier, *Phys. Lett.* **B590** (2004) 199 [arXiv:hep-ph/0310338].
GM04	[55] M. Gyulassy, L. McLerran, *New Forms of QCD Matter Discovered at RHIC*, nucl-th/0405013.
CGCreviews	[56] A.H. Mueller, *Parton Saturation–An Overview*, hep-ph/0111244; E. Iancu, A. Leonidov and L. McLerran, *The Colour Glass Condensate: An Introduction*, hep-ph/0202270. Published in *QCD Perspectives on Hot and Dense Matter*, Eds. J.-P. Blaizot and E. Iancu, NATO Science Series, Kluwer, 2002; E. Iancu and R. Venugopalan, *The Color Glass Condensate and High Energy Scattering in QCD*, hep-ph/0303204. Published in *Quark-Gluon Plasma 3*, Eds. R.C. Hwa and X.-N. Wang, World Scientific, 2003; H. Weigert, *Evolution at small x_{bj}: The Color Glass Condensate*, arXiv:hep-ph/0501087.
rahms-data	[57] I. Arsene [BRAHMS Collaboration], *Phys. Rev. Lett.* **93** (2004) 242303 [arXiv:nucl-ex/0403005].
AM94	[58] A.H. Mueller, *Nucl. Phys.* **B415** (1994) 373.
IM032	[59] E. Iancu and A.H. Mueller, *Nucl. Phys.* **A730** (2004) 494.
KL051	[60] A. Kovner and M. Lublinsky, *Phys. Rev.* **D71** (2005) 085004 [arXiv:hep-ph/0501198].
BIIT05	[61] J.-P. Blaizot, E. Iancu, K. Itakura, and D.N. Triantafyllopoulos, *Phys. Lett.* **B615** (2005) 221 [arXiv:hep-ph/0502221].
KL053	[62] A. Kovner and M. Lublinsky, *Phys. Rev. Lett.* **94** (2005) 181603 [arXiv:hep-ph/0502119].
BREM	[63] Y. Hatta, E. Iancu, L. McLerran, A.M. Stasto, and D.N. Triantafyllopoulos, *Nucl. Phys.* **A764** (2006) 423, arXiv:hep-ph/0504182.

226

First_HERA

[64] ZEUS Collaboration, M. Derrick et al., *Phys. Lett.* **B316** (1993) 412.

Old_HERA

[65] H1 Collaboration, S. Aid et al., *Nucl. Phys.* **B470** (1996) 3; H1 Collaboration, C. Adloff et al., *ibid.* **B497** (1997) 3; ZEUS Collaboration, M. Derrick et al., *Z. Phys.* **C72** (1996) 399; ZEUS Collaboration, J. Breitweg et al., *Eur. Phys. J.* **C7** (1999) 609.

Sterman

[66] G. Sterman, *An Introduction to Quantum Field Theory*, Cambridge University Press, Cambridge, 1993.

Peskin

[67] M.E. Peskin and D.V. Schroeder, *An Introduction to Quantum Field Theory*, Addison-Wesley, New York, 1995.

Stirling

[68] R.K. Ellis, W.J. Stirling and B.R. Webber, *QCD and Collider Physics*, Cambridge University Press, Cambridge, 1996.

Foster

[69] B. Foster, *"Lectures on HERA physics"*, EPJdirect **A1**, 1–11 (2003) [hep-ex/0206011].

AMCARGESE

[70] A.H. Mueller, *Parton Saturation–An Overview*, hep-ph/0111244. Published in *QCD Perspectives on Hot and Dense Matter*, Eds. J.-P. Blaizot and E. Iancu, NATO Science Series, Kluwer, 2002.

AM90

[71] A. H. Mueller, *Nucl. Phys.* **B335** (1990) 115.

NZ91

[72] N.N. Nikolaev and B.G. Zakharov, *Z. Phys.* **C49** (1991) 607, *ibid.* **C53** (1992) 331.

GBW99

[73] K. Golec-Biernat and M. Wüsthoff, *Phys. Rev.* **D59** (1999) 014017; *ibid.* **D60** (1999) 114023. 313.

BGBK

[74] J. Bartels, K. Golec-Biernat, and H. Kowalski, *Phys. Rev.* **D66** (2002) 014001.

KT

[75] H. Kowalski and D. Teaney, *Phys. Rev.* **D68** (2003) 114005.

LevinHERA

[76] E. Gotsman, E. Levin, M. Lublinsky, U. Maor, *Eur. Phys. J.* **C27** (2003) 411.

Thorne

[77] R.S. Thorne, *" Gluon Distributions and Fits using Dipole Cross-Sections"* , hep-ph/0501124.

CTEQ

[78] H. L. Lai *et al.* [CTEQ Collaboration], Eur. Phys. J. **C12** (2000) 375.

MRST

[79] A.D. Martin, R.G. Roberts, W.J. Stirling, R.S. Thorne, Eur. Phys. J. **C23** (2002) 73.

DL

[80] A. Donnachie and P. V. Landshoff, **B437** (1998) 408.

CFSK

[81] A. Capella, E. G. Ferreiro, C. A. Salgado and A. B. Kaidalov, Nucl. Phys. **B593** (2001) 336.

FKS

[82] J. R. Forshaw, G. Kerley and G. Shaw, Phys. Rev. **D60** (1999) 074012.

geometric

[83] A.M. Stasto, K. Golec-Biernat and J. Kwiecinski, Phys. Rev. Lett. **86** (2001) 596.

KS02

[84] J. Kwieciński and A.M. Staśto, *Geometric scaling and QCD evolution*, hep-ph/0203030.

WF02

[85] A. Freund, K. Rummukainen, H. Weigert and A. Schafer, inclusive e A arXiv:hep-ph/0210139.

KLM02

[86] D. E. Kharzeev, E. Levin, and L. McLerran, *Phys. Lett.* **B561** (2003) 93.

227

IIT04	[87] E. Iancu, K. Itakura, and D.N. Triantafyllopoulos, *Nucl. Phys.* **A742** (2004) 182.
BH	[88] W. Buchmuller, M.F. McDermott and A. Hebecker, *Nucl. Phys.* **B487** (1997) 283; Gehrmann, and A. Hebecker *Nucl. Phys.* **B537** (1999) 477.
RajDIS	[89] R. Venugopalan, Proc. **588**, 121 (2001) [arXiv:hep-ph/0102087].
Bjorken:uj	[90] J. D. Bjorken, Phys. Rev. D **27**, 140 (1983).
DM01	[91] A. Dumitru and L. McLerran, *Nucl. Phys.* **A700** (2002) 492.
DJM02	[92] A. Dumitru and J. Jalilian-Marian, Phys. Rev. Lett. **89**, 022301 (2002).
GJM1	[93] F. Gelis and J. Jalilian-Marian, Phys. Rev. D **66**, 014021 (2002).
GJM2	[94] F. Gelis and J. Jalilian-Marian, Phys. Rev. D **66**, 094014 (2002); Phys. Rev. D **66**, 014021 (2002).
GJM3	[95] F. Gelis and J. Jalilian-Marian, Phys. Rev. D **67**, 074019 (2003).
lisPeshier	[96] F. Gelis and A. Peshier, Nucl. Phys. A **697**, 879 (2002); Nucl. Phys. A **707**, 175 (2002).
Jamal	[97] J. Jalilian-Marian, Nucl. Phys. A **739**, 319 (2004).
KNL	[98] D. Kharzeev and M. Nardi, *Phys. Lett.* **B507** (2001) 121; D. Kharzeev and E. Levin, *Phys. Lett.* **B523** (2001) 79; D. Kharzeev, E. Levin and M. Nardi, *Phys. Rev.* **C71** (2005) 054903.
Mueller6	[99] A. H. Mueller, Phys. Lett. B **475**, 220 (2000).
BV	[100] J. Bjoraker and R. Venugopalan, Phys. Rev. C **63**, 024609 (2001).
reauSchiff	[101] J. Serreau and D. Schiff, JHEP **0111**, 039 (2001).
Wong	[102] S.M.H. Wong, Nucl. Phys. **A607** (1996) 442; Phys. Rev. **C54** (1996) 2588; *ibid.* **56** (1997) 1075.
BMSS	[103] R. Baier, A. H. Mueller, D. Schiff and D. T. Son, Phys. Lett. B **502**, 51 (2001).
BMSS2	[104] R. Baier, A. H. Mueller, D. Schiff and D. T. Son, Phys. Lett. B **539** 46 (2002).
Stan	[105] S. Mrowczynski, Phys. Lett. B **314**, 118 (1993).
AML	[106] P. Arnold, J. Lenaghan and G. D. Moore, JHEP **0308**, 002 (2003).
RS1	[107] P. Romatschke and M. Strickland, Phys. Rev. D **68**, 036004 (2003).
CroninExp	[108] J. W. Cronin, et al. *Phys. Rev.* **D11** (1975) 3105; D. Antreasyan, et al. *Phys. Rev. Lett.* **38** (1977) 112; D. Antreasyan, et al. *Phys. Rev.* **D19** (1979) 764.
IC-dAu-mid	[109] B. B. Back *et al.* [PHOBOS Collaboration], *Phys. Rev. Lett.* **91** (2003) 072302 [arXiv:nucl-ex/0306025]; S. S. Adler *et al.* [PHENIX Collaboration], *Phys. Rev. Lett.* **91** (2003) 072303 [arXiv:nucl-ex/0306021]; J. Adams *et al.* [STAR Collaboration], *Phys. Rev. Lett.* **91** (2003) 072304 [arXiv:nucl-ex/0306024]; I. Arsene *et al.* [BRAHMS Collaboration], *Phys. Rev. Lett.* **91** (2003) 072305 [arXiv:nucl-ex/0307003].
BDMPS	[110] R. Baier, Yu. L. Dokshitzer, A. H. Mueller, S. Peigne, and D. Schiff, Nucl. Phys. **B483** (1997) 291; *ibid.* **484** (1997) 265.

228

BSZ00   [111] For a review, see R. Baier, D. Schiff and B. G. Zakharov, perturbative QCD," Ann. Rev. Nucl. Part. Sci. **50**, 37 (2000).

tQuenching   [112] For a review, see M. Gyulassy, I. Vitev, X. N. Wang and B. W. Zhang, *"Jet quenching and radiative energy loss in dense nuclear matter,"* arXiv:nucl-th/0302077. Published in Quark Gluon Plasma 3, editors: R.C. Hwa and X.N. Wang, World Scientific, Singapore.

KW03   [113] For a review, see A. Kovner and U.A. Wiedemann, *"Gluon radiation and parton energy loss"*, arXiv:hep-ph/0304151. Published in Quark Gluon Plasma 3, editors: R.C. Hwa and X.N. Wang, World Scientific, Singapore.

KKT   [114] D. Kharzeev, Yu. V. Kovchegov, and K. Tuchin, *Phys. Rev.* **D66** (2003) 094013.

Nestor03   [115] J.L. Albacete, N. Armesto, A. Kovner, C.A. Salgado and U.A. Wiedemann, *Phys. Rev. Lett.* **92** (2004) 082001.

AR   [116] A. Krasnitz and R. Venugopalan, Nucl. Phys. B **557**, 237 (1999); Phys. Rev. Lett. **84**, 4309 (2000); Phys. Rev. Lett. **86**, 1717 (2001).

AYR   [117] A. Krasnitz, Y. Nara and R. Venugopalan, Phys. Rev. Lett. **87**, 192302 (2001); Nucl. Phys. A **717**, 268 (2003).

Lappi   [118] T. Lappi, Phys. Rev. C **67**, 054903 (2003).

AYR2   [119] A. Krasnitz, Y. Nara and R. Venugopalan, Nucl. Phys. A **727**, 427 (2003).

AYR3   [120] A. Krasnitz, Y. Nara and R. Venugopalan, Phys. Lett. B **554** 21 (2003).

LM00   [121] C. S. Lam and G. Mahlon, *Phys. Rev.* **D62** (2000) 114023; *ibid.* **D64** (2001) 016004.

GV96   [122] R. V. Gavai and R. Venugopalan, Phys. Rev. **D54**, 5795 (1996).

FR97   [123] J. R. Forshaw and D. A. Ross, *Quantum Chromodynamics and the Pomeron*, Cambridge University Press, Cambridge, 1997.

K   [124] Yu. V. Kovchegov, *Phys. Rev.* **D60** (1999), 034008; *ibid.* **D61** (2000) 074018.

lipatov2   [125] L. N. Lipatov, JETP **63** (1986) 904.

bartels   [126] J. Bartels, Nucl. Phys. B **175** (1980) 365.

KP80   [127] J. Kwiecinski and M. Praszalowicz, Phys. Lett. B **94** (1980) 413.

Jaro   [128] T. Jaroszewicz, Acta, Phys. Polon. B**11** (1980) 965.

bartels0   [129] J. Bartels, Nucl. Phys. B **151** (1979) 293.

lipatov   [130] L. N. Lipatov, Phys. Lett. B **251** (1990) 284; Phys. Lett. B **309** (1993) 394; JETP Lett. **59** (1994) 596; see also L. N. Lipatov, "Pomeron in Quantum Chromodynamics", in *Perturbative Quantum Chromodynamics*, Ed. A. H. Mueller, World Scientific, Singapore, 1989.

lipatov1   [131] L. N. Lipatov, arXiv:hep-th/9311037; DFPD/93/TH/70 (unpublished); JETP Lett. **59** (1994) 596.

FK95 [132] L. D. Faddeev and G. P. Korchemsky, Phys. Lett. B **342** (1995) 311; G. P. Korchemsky, Nucl. Phys. B **443** (1995) 255.

janik [133] R.A. Janik, *Phys. Lett.* **B604** (2004) 192.

BLV [134] J. Bartels, L. N. Lipatov and G. P. Vacca, Phys. Lett. B **477**, (2000) 178.

GK [135] G. P. Korchemsky, J. Kotanski and A. N. Manashov, Phys. Rev. Lett. **88** (2002) 122002; S. E. Derkachov, G. P. Korchemsky and A. N. Manashov, Nucl. Phys. B **617** (2001) 375; S. E. Derkachov, G. P. Korchemsky, J. Kotanski and A. N. Manashov, Nucl. Phys. B **645** (2002) 237.

DVL02 [136] H. J. De Vega and L. N. Lipatov, Phys. Rev. D **64** (2001) 114019; Phys. Rev. D **66** (2002) 074013.

ewerz [137] C. Ewerz, " *The odderon in quantum chromodynamics*", hep-ph/0306137.

MP94 [138] A.H. Mueller and B. Patel, *Nucl. Phys.* **B425** (1994) 471.

BW93 [139] J. Bartels and M. Wüsthoff, *Z. Phys.* **C66** (1995) 157.

BV99 [140] M. Braun and G.P. Vacca, *Eur. Phys. J.* **C6** (1999) 147.

B99 [141] M. Braun, *Eur. Phys. J.* **C6** (1999) 321.

BE99 [142] J. Bartels and C. Ewerz, JHEP **9909** (1999) 026.

ewerz0 [143] C. Ewerz, Phys. Lett. B **472** (2000) 135; *ibid.* **512** (2001) 239; C. Ewerz and V. Schatz, Nucl. Phys. A **736** (2004) 371.

ES04 [144] C. Ewerz and V. Schatz, *Nucl. Phys.* **A736** (2004) 371.

BE05 [145] C. Ewerz and S. Braunewell, "*The C-Odd Four-Gluon State in the Color Glass Condensate,*" hep-ph/0501110.

BLV05 [146] J. Bartels, L.N. Lipatov and G.P. Vacca, *Nucl. Phys.* **B706** (2005) 391.

BBV05 [147] J. Bartels, M. Braun and G.P. Vacca, "*Pomeron vertices in perturbative QCD in diffractive scattering*", arXiv:hep-ph/0412218.

RP97 [148] R. Peschanski, *Phys. Lett.* **B409** (1997) 491.

KS70 [149] J.B. Kogut and D.E. Soper, *Phys.Rev.* **D1** (1970) 2901.

JKLW99b [150] J. Jalilian-Marian, A. Kovner, A. Leonidov and H. Weigert, *Phys. Rev.* **D59** (1999), 034007; Erratum-ibid. **D59** (1999), 099903.

AM01 [151] A. H. Mueller, *Phys. Lett.* **B523** (2001) 243.

AJMV95 [152] A. Ayala, J. Jalilian-Marian, L. McLerran and R. Venugopalan, *Phys. Rev.* **D52** (1995), 2935.

HW98 [153] A. Hebecker and H. Weigert, *Phys. Lett.* **B432** (1998) 215.

Parisi [154] G. Parisi, *Statistical Field Theory*, Addison-Wesley, New York, 1988.

ZJ   [155] J. Zinn-Justin, *Quantum field theory and critical phenomena*, (International series of monographs on physics, 77), Clarendon, Oxford, 1989.

Gardiner   [156] C.W. Gardiner, *Handbook of Stochastic Methods*, (Springer series in synergetics), Springer, Berlin, 2004.

Miller   [157] W. Miller, *Symmetry groups and their applications*, Academic Press, New York, 1972.

JP04   [158] R.A. Janik, R. Peschanski, *Phys. Rev.* **D70** (2004) 094005.

bartels2   [159] J. Bartels, M. A. Braun, D. Colferai and G. P. Vacca, Eur. Phys. J. C **20** (2001) 323.

KSW04   [160] Y. V. Kovchegov, L. Szymanowski and S. Wallon, dipole model," Phys. Lett. B **586** (2004) 267.

dosch   [161] H.G. Dosch, C. Ewerz and V. Schatz, Eur. Phys. J. C **24** (2002) 561.

LT99   [162] E. Levin and K. Tuchin, *Nucl. Phys.* **B573** (2000) 833; *Nucl. Phys.* **A691** (2001) 779; *Nucl. Phys.* **A693** (2001) 787.

IM031   [163] E. Iancu and A.H. Mueller, *Nucl. Phys.* **A730** (2004) 460.

MP03   [164] S. Munier and R. Peschanski, *Phys. Rev. Lett.* **91** (2003) 232001; *Phys. Rev.* **D69** (2004) 034008; *ibid.* **D70** (2004) 077503.

Saar   [165] For a recent review, see W. Van Saarloos, *Phys. Rep.* **386** (2003) 29.

Panja   [166] For a recent review, see D. Panja, Phys. Rep. **393** (2004) 87 [arXiv:cond-mat/0307363].

KW02   [167] A. Kovner and U.A. Wiedemann, *Phys. Rev.* **D66** (2002) 034031; *ibidem* 051502.

GBS03   [168] K. Golec-Biernat and A.M. Stasto, *Nucl. Phys.* **B668** (2003) 345.

FB   [169] E. Ferreiro, E. Iancu, K. Itakura, and L. McLerran, *Nucl. Phys.* **A710** (2002) 373.

MP031   [170] S. Munier and R. Peschanski, *Phys. Rev. Lett.* **91** (2003) 232001.

MP033   [171] S. Munier and R. Peschanski, *Phys. Rev.* **D70** (2004) 077503.

AM03   [172] A. H. Mueller, *Nucl. Phys.* **A724** (2003) 223.

GKRW   [173] E. Gardi, J. Kuokkanen, K. Rummukainen and H. Weigert, *Nucl. Phys.* **A784** (2007) 282.

KWrun1   [174] Yu.V. Kovchegov and H. Weigert, *Nucl. Phys.* **A784** (2007) 188; *ibid.* **A789** *(2007) 260.*

Brun   [175] I. Balitsky, *Phys. Rev.* **D75** (2007) 014001.

Alb07   [176] J.L. Albacete and Yu.V. Kovchegov, arXiv:0704.0612 [hep-ph].

alitChir07   [177] I. Balitsky and G. A. Chirilli, *NLO evolution of color dipoles*, arXiv:0710.4330 [hep-ph].

MP032   [178] S. Munier and R. Peschanski, *Phys. Rev.* **D69** (2004) 034008.

Motyka   [179] K. Golec-Biernat, L. Motyka, and A.M. Staśto, *Phys. Rev.* **D65** (2002) 074037.

RW03   [180] K. Rummukainen and H. Weigert, *Nucl. Phys.* **A739** (2004) 183.

231

Nestor04	[181] J.L. Albacete, N. Armesto, J.G. Milhano, C.A. Salgado and U.A. Wiedemann, *"Numerical analysis of the Balitsky-Kovchegov equation with running coupling: dependence of the saturation scale on nuclear size and rapidity"*, arXiv:hep-ph/0408216, *Phys. Rev.* **D71** (2005) 014003.
RPLOOP	[182] A. Dumitru, E. Iancu, L. Portugal, G. Soyez, and D.N. Triantafyllopoulos, *JHEP* **0708** (2007) 062, arXiv:0706.2540 [hep-ph].
GB07	[183] G. Beuf, *"An alternative scaling solution for high-energy QCD saturation with running coupling,"* arXiv:0803.2167 [hep-ph].
AB01	[184] N. Armesto and M. Braun, *Eur. Phys. J.* **C20** (2001) 517; *ibid.* **C22** (2001) 351.
LL01	[185] E. Levin and M. Lublinsky, *Phys. Lett.* **B521** (2001) 233; *Eur. Phys. J.* **C22** (2002) 647; M. Lublinsky, *Eur. Phys. J.* **C21** (2001) 513.
FKPP	[186] R.A. Fisher, *Ann. Eugenics* **7** (1937) 355; A. Kolmogorov, I. Petrovsky, and N. Piscounov, *Moscou Univ. Bull. Math.* **A1** (1937) 1.
Bramson	[187] M. Bramson, *Mem. Am. Math. Soc.* **44** (285) 1983.
BD	[188] E. Brunet and B. Derrida, *Phys. Rev.* **E56** (1997) 2597; *Comp. Phys. Comm.* **121-122** (1999) 376; *J. Stat. Phys.* **103** (2001) 269.
NW97	[189] H. Navelet and S. Wallon *Nucl.Phys.* **B522** (1998) 237.
MW03	[190] S. Munier and S. Wallon, *Eur. Phys. J.* **C30** (2003) 359.
KL02	[191] M. Kozlov and E. Levin, *Eur. Phys. J.* **C28** (2003) 483; S.Bondarenko, M.Kozlov and E.Levin, *Acta Phys. Polon.* **B34** (2003) 3081.
Levin_BKB	[192] E. Gotsman, M. Kozlov, E. Levin, U. Maor, E. Naftali, *Nucl. Phys.* **A742** (2004) 79.
BNetal	[193] J.R. Cudell et al, *Phys. Rev.* **D65** (2002) 074024.
Forshaw	[194] J. R. Forshaw, R. Sandapen and G. Shaw, *Phys. Rev.* **D69** (2004) 094013; *Phys. Lett.* **B594** (2004) 283.
Forshaw2	[195] J. R. Forshaw and G. Shaw, JHEP **0412**, 052 (2004).
BFKL_fits	[196] H. Navelet, R. Peschanski, C. Royon, and S. Wallon, *Phys. Lett.* **B385** (1996) 357; S. Munier and R. Peschanski, *Nucl. Phys.* **B524** (1998) 377; A.I. Lengyel and M.V.T.Machado, hep-ph/0304195.
Machado	[197] V.P. Goncalves, M.V.T. Machado, *Eur. Phys.* **J.C37** (2004) 299.
ariaDiff05	[198] M.S. Kugeratski, V.P. Goncalves and F.S. Navarra, *"Saturation in diffractive deep inelastic eA scattering"*, arXiv:hep-ph/0511224.
JamalCGC	[199] J. Jalilian-Marian, *Nucl. Phys.* **A748** (2005) 2005; *"Production of Forward Rapidity Photons in High Energy Heavy Ion Collisions"*, hep-ph/0501222.
Ducati	[200] M.A. Betemps, M.B. Gay Ducati, *Phys. Rev.* **D70** (2004) 116005.
Dumitru05	[201] A. Dumitru, A. Hayashigaki, and J. Jalilian-Marian, *Geometric scaling violations in the central rapidity region of d+Au collisions at RHIC*, arXiv:hep-ph/0512129.

| Au-forward | [202] R. Debbe, BRAHMS Collaboration, J. Phys. G30 (2004) S759, nucl-ex/0403052; M. Murray, BRAHMS Collaboration, J. Phys. G30 (2004) S667, nucl-ex/0404007; (Proceedings for QM2004, Oakland, USA, January 2004). |

| ward-other | [203] A. Frawley, PHENIX Collaboration, J. Phys. G30 (2004) S675, nucl-ex/0404009; M.X. Liu, PHENIX Collaboration, J. Phys. G30 (2004) S1193, nucl-ex/0403047; P. Steinberg, PHOBOS Collaboration, J. Phys. G30 (2004) S683, nucl-ex/0405023; R. Nouicer, PHOBOS Collaboration, J. Phys. G30 (2004) S1133, nucl-ex/0403033; L.S. Barnby, STAR Collaboration, J. Phys. G30 (2004) S1121, nucl-ex/0404027; (Proceedings for QM2004, Oakland, USA, January 2004). |

| STAR-data | [204] J. Adams *et al.* [STAR Collaboration], [arXiv:nucl-ex/0602011]. |

| JNV | [205] J. Jalilian-Marian, Y. Nara and R. Venugopalan, *Phys. Lett.* **B577** (2003) 54. |

| Baier03 | [206] R. Baier, A. Kovner and U. A. Wiedemann, Phys. Rev. D **68** (2003) 054009. |

| BGV04 | [207] J.P. Blaizot, F. Gelis, R. Venugopalan, Nucl. Phys. A **743**, 13 (2004); *ibid.* 57 (2004). |

| KKT2 | [208] D. Kharzeev, Yu. V. Kovchegov, and K. Tuchin, *Phys. Lett.* **B599** (2004) 23. |

| BMTS06 | [209] R. Baier, Y. Mehtar-Tani and D. Schiff, Nucl. Phys. **A764** (2006) 515, hep-ph/0508026. |

| Dumitru1 | [210] A. Dumitru, A. Hayashigaki, and J. Jalilian-Marian, *Nucl. Phys.* **A765** (2006) 464. |

| Dumitru2 | [211] A. Dumitru, A. Hayashigaki, and J. Jalilian-Marian, *Nucl. Phys.* **A770** (2006) 57. |

| BraunG | [212] M. Braun, *Eur. Phys. J.* **C16** (2000) 337; *Phys. Lett.* **B483** (2000) 105. |

| KT02 | [213] Yu. V. Kovchegov and K. Tuchin, *Phys. Rev.* **D65** (2002) 074026. |

| FGV05 | [214] H. Fujii, F. Gelis and R. Venugopalan, *Phys. Rev. Lett.* **95** (2005) 162002, hep-ph/0504047; *Quark pair production in high energy pA collisions: General features*, hep-ph/0603099. |

| KLN02 | [215] D. Kharzeev, E. Levin, and M. Nardi, *Nucl. Phys.* **A730** (2004) 448; D. Kharzeev, E. Levin, and M. Nardi, hep-ph/0212316, Erratum, Nuclear Physics A in press. |

| KMW95 | [216] A. Kovner, L. D. McLerran and H. Weigert, Phys. Rev. D **52**, 3809 (1995). |

| EdiRaju | [217] E. Iancu and R. Venugopalan, *The Color Glass Condensate and High Energy Scattering in QCD*, hep-ph/0303204. Published in *Quark-Gluon Plasma 3*, Eds. R.C. Hwa and X.-N. Wang, World Scientific, 2003. |

| FTK04 | [218] F. Gelis, K. Kajantie and T. Lappi, classical fields in heavy ion collisions: 1+1 arXiv:hep-ph/0409058. |

| KTS99 | [219] B.Z. Kopeliovich, A.V. Tarasov, and A. Schafer, *Phys. Rev.* **C59** (1999) 1609. |

| LL03 | [220] E. Levin and M. Lublinsky, *Nucl. Phys.* **A730** (2004) 191. |

| IST05 | [221] E. Iancu, G. Soyez and D.N. Triantafyllopoulos, *Nucl. Phys.* **A768** (2006) 194. |

| LL05 | [222] E. Levin and M. Lublinsky, *Nucl. Phys.* **A763** (2005) 172, arXiv:hep-ph/0501173. |

| LL04 | [223] E. Levin and M. Lublinsky, *Phys. Lett.* **B607** (2005) 131, arXiv:hep-ph/0411121. |

KL054	[224] A. Kovner and M. Lublinsky, *Phys. Rev.* **D72** (2005) 074023, arXiv:hep-ph/0503155.
MMSW05	[225] C. Marquet, A.H. Mueller, A.I. Shoshi, and S.M.H. Wong *Nucl. Phys.* **A762** (2005) 252, arXiv:hep-ph/0505229.
HIMS05	[226] Y. Hatta, E. Iancu, L. McLerran, and A. Stasto, *Nucl. Phys.* **A762** (2005) 272, arXiv:hep-ph/0505235.
Balit98	[227] I. Balitsky, *Phys. Rev. Lett.* **81** (1998) 2024; *Phys. Lett.* **B518** (2001) 235.
Balit05	[228] I. Balitsky, *Phys. Rev.* **D72** (2005) 074027, arXiv:hep-ph/0507237.
YH06	[229] Y. Hatta, *Nucl. Phys.* **A768** (2006) 222; *Nucl. Phys.* **A781** (2007) 104.
Braun05	[230] M. Braun, *Phys. Lett.* **B483** (2000) 115; *ibid.* **B632** (2006) 297; *"Conformal invariant equations for nucleus-nucleus scattering in perturbative QCD with $N_c \to \infty$"*, arXiv:hep-ph/0504002.
BIT06	[231] J.-P. Blaizot, E. Iancu, D.N. Triantafyllopoulos, *Nucl. Phys.* **A784** (2007) 227.
PL99	[232] L. Pechenik and H. Levine, *Phys. Rev.* **E69** (1999) 3893.
Moro042	[233] E. Moro, *Phys. Rev.* **E70** (2004) 045102(R).
GS05	[234] G. Soyez, *Phys. Rev.* **D72** (2005) 016007.
EGBM05	[235] R. Enberg, K. Golec–Biernat, S. Munier, *Phys. Rev.* **D72** (2005) 074021.
ISST06	[236] E. Iancu, J.T. de Santana Amaral, G. Soyez, and D.N. Triantafyllopoulos, *Nucl. Phys.* **A786** (2007) 131 [arXiv:hep-ph/0611105].
SX05	[237] A. I. Shoshi and B.-W. Xiao, *Phys. Rev.* **D73** (2006) 094014.
KozlovL05	[238] M. Kozlov and E. Levin, *Nucl. Phys.* **A764** (2006) 498.
PS01	[239] D. Panja, W. Van Saarloos, *Phys. Rev.* **E 66** (2002) 036206 [arXiv:cond-mat/0109528].
Moro041	[240] E. Moro, *Phys. Rev.* **E69** (2004) 060101(R).
BDMM	[241] E. Brunet, B. Derrida, A. H. Mueller and S. Munier, *Phys. Rev.* **E73** (2006) 056126, arXiv:cond-mat/0512021.
MSX06	[242] C. Marquet, G. Soyez, Bo-Wen Xiao, *Phys. Lett.* **B639** (2006) 635, arXiv:hep-ph/0606233.
HIMST06	[243] Y. Hatta, E. Iancu, C. Marquet, G. Soyez, and D.N. Triantafyllopoulos, *Nucl. Phys.* **A773** (2006) 95, arXiv:hep-ph/0601150.
GLUON	[244] E. Iancu, C. Marquet, and G. Soyez, *Nucl. Phys.* **A780** (2006) 52, arXiv:hep-ph/0605174.
HM07	[245] Y. Hatta and A.H. Mueller, *"Correlation of small-x gluons in impact parameter space,* arXiv: hep-ph/0702023.
IM07	[246] E. Iancu and L. McLerran, *"Liouville field theory for gluon saturation in QCD at high energy"* arXiv:hep-ph/0701276.
Poly81	[247] A.Polyakov. *Phys. Lett.*, **B103** (1981) 207.

Tesch	[248] J. Teschner, *"Liouville theory revisited"*, in *Class. Quant. Grav.* **18** (2001) R153; e-Print Archive: hep-th/0104158.
LW94	[249] E. Levin and M. Wüsthoff, *Phys. Rev.* **D50** (1994) 4306.
BW95	[250] J. Bartels and M. Wüsthoff, *Z. Phys.* **C66** (1995) 157; *ibid.* **68** (1995) 121.
BP96	[251] A. Bialas, R. Peschanski, *Phys. Lett.* **B378** (1996) 302; *Phys. Lett.* **B387** (1996) 405; A. Bialas, R. Peschanski and C. Royon, *Phys. Rev.* **D57** (1998) 6899.
W97	[252] M. Wüsthoff, *Phys. Rev.* **D56** (1997) 4311.
Hebecker	[253] A. Hebecker, Phys. Rept. **331** (2000) 1.
KST00	[254] B. Z. Kopeliovich, A. Schaefer and A. V. Tarasov, *Phys. Rev.* **D62** (2000) 0540022.
Gotsman00	[255] E. Gotsman, E. Levin, M. Lublinsky, U. Maor and K. Tuchin, *Phys. Lett.* **B492** (2000) 47; *Nucl. Phys.* **A697** (2002) 521.
KM99	[256] Yu.V. Kovchegov and L. McLerran, *Phys. Rev. D* **60** (1999) 054025.
KL00	[257] Yu.V. Kovchegov and E. Levin, *Nucl. Phys.* **B577** (2000) 221.
MS03	[258] S. Munier and A. Shoshi, *Phys. Rev.* **D69** (2004) 074022.
CM04	[259] C. Marquet, *Nucl. Phys.* **B705** (2005) 319.
W05	[260] M. Hentschinski, H. Weigert and A. Schafer, "Extension of the color glass condensate approach to diffractive reactions," arXiv:hep-ph/0509272.
GBM05	[261] K. Golec-Biernat and C. Marquet, *Phys. Rev.* **D71** (2005) 114005.