# Classicalization and unitarization of wee partons in QCD and gravity: The CGC-black hole correspondence

Gia Dvali<sup>1,2</sup> and Raju Venugopalan<sup>3</sup>

<sup>1</sup>Arnold Sommerfeld Center, Ludwig-Maximilians-Universität, Theresienstrasse 37, 80333 München, Germany <sup>2</sup>Max-Planck-Institut für Physik, Föhringer Ring 6, 80805 München, Germany <sup>3</sup>Physics Department, Brookhaven National Laboratory, Bldg. 510A, Upton, New York 11973, USA

(Received 23 January 2022; accepted 11 March 2022; published 29 March 2022)

We discuss a remarkable correspondence between the description of black holes as highly occupied condensates of N weakly interacting gravitons and that of color glass condensates (CGCs) as highly occupied gluon states. In both cases, the dynamics of "wee partons" in Regge asymptotics is controlled by emergent semihard scales that lead to perturbative unitarization and classicalization of  $2 \rightarrow N$  particle amplitudes at weak coupling. In particular, they attain a maximal entropy permitted by unitarity, bounded by the inverse coupling  $\alpha$  of the respective constituents. Strikingly, this entropy is equal to the area measured in units of the Goldstone constant corresponding to the spontaneous breaking of Poincaré symmetry by the corresponding graviton or gluon condensate. In gravity, the Goldstone constant is the Planck scale, and gives rise to the Bekenstein-Hawking entropy. Likewise, in the CGC, the corresponding Goldstone scale is determined by the onset of gluon screening. We point to further similarities in black hole formation, thermalization and decay, to that of the glasma matter formed from colliding CGCs in ultrarelativistic nuclear collisions, which decays into a quark-gluon plasma.

DOI: 10.1103/PhysRevD.105.056026

#### I. INTRODUCTION

A topic of fundamental interest in physics is the possible existence of deep connections between the infrared behavior of the strong interaction and gravity, going back to the formulations of QCD and string theory. In the last couple of decades, the paradigmatic example has been the AdS/CFT correspondence, albeit both the AdS and the CFT sides of this correspondence are not the theories of phenomenological interest [1]. A striking recent example of a concrete connection between gravity and QCD is the so-called BCJ "double copy" whereby perturbative gravity amplitudes are constructed out of QCD amplitudes and additional kinematic factors [2].

In this work, we will discuss a possible novel correspondence between specific states in the two theories respectively arising from our observations of their remarkable similarities at an infrared fixed point characterized by high occupancy. In the case of gravity, these are high occupancy graviton states which provide a microscopic "quantum *N*-portrait" of a black hole (BHNP) [3,4]. In QCD, the corresponding quantum portrait of these states in the wave function of a hadron at high energies is a saturated gluon state commonly called a color glass condensate (CGC) [5]. In the collisions of two nuclei at ultrarelativistic energies, the collision of the two CGC gluon shock waves generates an overoccupied glasma nonequilibrium state [6,7] which subsequently evolves into a quark-gluon plasma (QGP) [8].

As we shall elaborate, the similarities between the two quantum portraits strongly hint at universal behavior at a quantum critical point that may be independent of the details of their microscopic dynamics or the initial conditions; the physics of these objects is defined by the fact that they represent saturated states corresponding to highly occupied soft quanta at a "critical packing" sufficient to unitarize the cross section for their formation. The critical packing of the quanta (gravitons or gluons) is defined by a characteristic saturation momentum scale, denoted by  $Q_{\rm S}$ , with the associated de Broglie wavelength  $R_S = 1/Q_S$ . Other key quantities are (1) The occupation number of quanta per de Broglie volume  $(R_s^3 = Q_s^{-3})$ , to be denoted by N, and (2) the running coupling constant  $\alpha(Q_S)$ , determining the strength of the interaction among these quanta. Critical packing is reached when the occupation number and the coupling satisfy,

$$N = \frac{1}{\alpha(Q_S)}.$$
 (1)

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

As suggested by this expression, the running of  $\alpha(Q_S)$  is evaluated at the scale  $Q_S$ . The physical meaning of critical packing is that it represents the point of optimal balance between the kinetic energies of the individual constituents and their potential energies. In particular, in the BHNP, this is the regime in which the graviton condensate forms a selfsustained long-lived bound state, a black hole.

In both systems (BHNP or CGC), the critical packing results in the emergence of a large number of gapless modes corresponding to a specific microstate entropy *S*. An important ingredient in the correspondence between the systems is the relation between this microstate entropy and the coupling  $\alpha(Q_S)$ . This relation, proposed and worked out in a series of papers [9–11] constitutes an upper bound imposed by unitarity on the entropy:

$$S_{\max} = \frac{1}{\alpha(Q_S)}.$$
 (2)

According to [11], the correlation between the saturation of unitarity and the bound in Eq. (2) is nonperturbative in nature and cannot be avoided by resummations in perturbation theory. It was argued further that any system saturating this entropy bound shares certain universal properties with a black hole. Perhaps the most striking is the relation between the entropy and surface area and the role of a Goldstone boson in this relation. The bottom line is that the maximal entropy permitted by unitarity is equal to the area ( $\sim R_S^2$ ) of the system measured in units of a Goldstone decay constant f [9–11],

$$S_{\max} = \operatorname{Area} \times f^2. \tag{3}$$

This Goldstone boson is a universal consequence of the spontaneous breaking of Poincaré symmetry. For a saturated system, the expression in Eq. (3) is equivalent to that in Eq. (2). To summarize Ref. [11], the information capacity of a unitary system is bounded by the area measured in units of the Goldstone scale.

In this respect, black holes are not special and their entropy exhibits the area-law of a generic saturated system. In the references noted, the connection between black holes and similar saturated objects in other physical systems was demonstrated for a number of systems, including states in gauge theories with large number of colors. Note that in the black hole case the Goldstone is a graviton; the expression in Eq. (3) gives the Bekenstein-Hawking entropy [12,13]. Other analogous universal features of such states include decay timescales and that of information recovery [11], which shall be discussed later.

A unique, and arguably the most striking, feature of the correspondence we wish to highlight here is that both the CGC state probed in QCD at high energies and the BHNP description of black holes occur in physical systems in nature. While QCD and gravity differ fundamentally away from the quantum critical point, the commonality in critical packing and as we shall discuss in the relations in Eqs. (2) and(3) in this regime, determines the similarities in important features of the two systems. The connections extend beyond critical packing and cover several other key features of their dynamics. Both black hole and CGC/glasma states are consequences of the "classicalization" and unitarization of  $2 \rightarrow N$  amplitudes in gravity and QCD respectively in the high energy Regge limit where the naive  $N!/N^N \sim \exp(-N)$  suppression of cross sections is compensated by an  $\exp(N)$  factor from counting the number of microstates in the system. In the case of black holes, the logarithm of this number is the black hole entropy factor. In the CGC scattering picture, it is commonly understood to arise from the exponentiation of soft gluon bremsstrahlung.

The idea of the saturation of trans-Planckian amplitudes by black hole production goes back to the papers by 't Hooft [14,15], Gross and Mende [16,17] and by Amati, Ciafaloni and Veneziano [18] (ACV), followed by a large body of literature we shall not attempt to fully review here. The step that is crucial for our CGC/BHNP correspondence is the realization [3] that the relevant amplitude describing black hole formation in the scattering of two trans-Planckian gravitons is that of the  $2 \rightarrow N$  Regge scattering process in which the bulk of the center of mass energy,  $\sqrt{s}$ , is redistributed among N quanta of roughly equal softness, corresponding to momenta  $|\mathbf{p}| \sim \frac{\sqrt{s}}{N}$  [19]. A two-dimensional effective field theory for  $2 \rightarrow N$  scattering amplitudes in this regime of "multi-Regge" kinematics in QCD and gravity was developed by Lipatov [20], who also observed early-on that gravitational effective vertices could be represented as "squares" of so-called Lipatov vertices in QCD [21].

Explicit computations [19] of this process both in string theory as well as in field theory show that when the momenta of final-state gravitons are given by  $|\mathbf{p}| \sim \frac{M_P^2}{\sqrt{s}}$ , the scattering rate delivers precisely the right exponential suppression factor  $e^{-N}$  that compensates the microstate degeneracy factor  $e^N$  of a black hole of mass  $M_{\rm BH} = \sqrt{s}$ . Note that this typical momentum is equal to the inverse Schwarzschild radius of such a black hole,  $|\mathbf{p}| = \frac{1}{R_s}$ . This momentum also corresponds to the energy of Hawking quanta (with temperature  $T_H \sim 1/R_s$ ) produced during its evaporation. At the same time, as noted, N is equal to black hole entropy. Thus without any additional input, the saturation of the  $2 \rightarrow N$  scattering amplitude in the Regge limit delivers key ingredients of the black hole N-portrait.

This result was further refined in an explicit computation of  $2 \rightarrow N$  Reggeized graviton amplitudes by Addazi, Bianchi and Veneziano [22] within the ACV approach. In this approach, developed over several decades, the contributions of two- and higher- graviton exchanges to single graviton exchange at large impact parameters are shown to become increasingly important with decreasing impact parameter *b*, with all order contributions becoming important at  $b \sim R_S$ . For a detailed discussion, and relevant references, see for instance [23]. Perturbative unitarization in this computation is imposed through the eikonalization of the Reggeon amplitudes at the unitarization boundary  $R_S$  or equivalently at  $T_H \sim 1/R_S$ . An essential feature in demonstrating the equivalence of this picture to the quantum *N*-portrait [3] is the exponentiation of real and virtual emissions of soft gravitons in the computation of [22].

In QCD, the cross section for  $2 \rightarrow N$  scattering in multi-Regge kinematics is given by the BFKL equation [24,25]. As shown by Lipatov and collaborators, infrared t- channel divergences cancel for this "BFKL Pomeron" in multi-Regge kinematics, giving a result for the cross section that corresponds to a resummation of leading logarithms  $\alpha_{\rm S} \ln(x)$ , where x denotes the longitudinal momentum fraction carried by wee partons in multi-Regge kinematics.<sup>1</sup> This resummation leads to a cross section that grows rapidly with the center-of-mass energy. However when gluon occupancies  $N \gg 1$ , many-body processes corresponding to the screening and recombination of small xgluons become important and counter soft gluon bremsstrahlung, a phenomenon known as gluon saturation [26,27]. As understood in this work, these many-body processes can be shown to arise, precisely as employed by ACV in their computations of gravitational amplitudes, from the "AGK cutting rules" for multi-Reggeon exchanges in  $2 \rightarrow 2$  scattering amplitudes [28]. Though derived in the particular context of Pomeron and Reggeon exchanges, the AGK rules are a general consequence of the combinatorics of producing a given number of particles from connected sub-graphs where some are cut and the rest are not; for a discussion in the language of quantum field theories with strong sources, applicable in our context, see [29,30].

Gluon saturation leads to the classicalization of  $2 \rightarrow N$ amplitudes [31,32] in the Regge limit of QCD and the wee gluon matter characterizing this saturated state is the CGC. A key feature of the CGC is that the typical momenta of wee gluons resolved by a probe with resolution Q is given by  $Q \equiv Q_S(x)$ , where  $Q_S(x)$  is an emergent energydependent scale whose inverse, given by  $R_S \sim 1/Q_S$ , denotes the close packing radius of saturated gluons at maximal occupancy. Our choice of  $R_S$  for this scale is deliberate since, as we will discuss shortly, the dynamics of wee partons in both QCD and gravity is governed by this scale. We note that parton saturation and the saturation scale were not considered in early discussions of the similarities between wee parton dynamics in QCD and gravity [33-35]. Taking this physics into account qualitatively modifies that discussion.<sup>2</sup>

This paper is organized as follows. In Sec. II, we will begin by noting the common features of saturated wee parton states in QCD. We will also discuss a fundamental difference in the UV (ultraviolet) completeness of the two theories, and discuss the sense in which the UV regime of gravity is actually probing the deep infrared (IR) regime of the theory. We will then establish the concrete dictionary of the BHNP-CGC correspondence. With this in hand, in Sec. III we discuss how the entropy bounds Eqs. (2) and (3) are saturated at the unitarization boundary both for black holes and the CGC. In particular, we discuss how this entropy can be expressed as an area law in units of the Goldstone constant characterizing the spontaneous breaking of Poincaré invariance by both saturated states. In Sec. IV, we discuss the timescales characterizing the decay of black holes in the BHNP and relate these to those in the CGC/ glasma; the relevant scattering rates are controlled by the dynamical screening of scattering between the wee partons mediated by collective modes of the N particle state. While the thermalization process is analogous in the semiclassical evolution of the two systems, their evolution differs qualitatively after a quantum break time when the occupancies in the two systems are no longer large. We also observe similarities in the very early-stage behavior of the system whereby the rapid scrambling of information leads to a quasiparticle picture with the subsequent generation of entanglement generated by kinetic scattering processes. We end with an outlook on further aspects of the BHNP-CGC correspondence and outline future directions of work to establish this correspondence on a quantitative footing. The Appendix outlines a general argument relating the Goldstone decay constant of collective modes to the area-law entropy in the two systems.

# II. DICTIONARY BETWEEN THE SEMICLASSICAL PORTRAIT OF A BLACK HOLE AND A CGC

We will establish here a dictionary between the semiclassical UV/IR portraits of black holes and CGCs, by first identifying the correspondence in the emergent scales of the two theories.

# A. Black holes

According to the black hole *N*-portrait (BHNP) [3,4], the black hole represents a long lived state of soft gravitons

<sup>&</sup>lt;sup>1</sup>In the Regge limit,  $s \to \infty$ ,  $s/t \to -\infty$ , where -t = fixed is the momentum transfer squared in the gluon scattering amplitude; multi-Regge denotes Regge kinematics applied to each rung of the  $2 \to N$  ladder corresponding to bremsstrahlung emission. Equivalently, in the language of deeply inelastic scattering (DIS), the Regge limit corresponds to Bjorken  $x_{Bj} \approx Q^2/s \to 0$ , where  $Q^2$  is the ultraviolet resolution scale of the DIS probe. In the parton model,  $x_{Bj} \approx x$ , where x is the light cone momentum fraction of a hadron carried by a parton.

<sup>&</sup>lt;sup>2</sup>We note that while a saturation scale has been discussed previously in the AdS/CFT correspondence, it does not correspond to a critical packing scale of gluons and gravitons as in the CGC and the black hole N portrait.

with very high occupation number  $N \gg 1$ ; it can therefore be viewed as a graviton condensate that is approximately classical. For any given N, the graviton wavelength  $R_S$  (and thus momentum  $Q_S = R_S^{-1}$ ) is determined by the saturation condition,

$$\alpha_{\rm gr}(Q_S)N = 1,\tag{4}$$

where  $\alpha_{gr}(Q_S)$  is the graviton-graviton coupling evaluated at the scale  $Q_S$ ,

$$\alpha_{\rm gr}(Q_S) \equiv \frac{Q_S^2}{M_P^2} = (L_P Q_S)^2.$$
 (5)

Here (in speed of light units)  $M_P = \sqrt{\hbar/G}$  and  $L_P = \sqrt{\hbar G}$  denote the Planck mass and length respectively, with *G* denoting Newton's constant.

The physical meaning of Eq. (4) can be understood from several perspectives. It defines the condition for the *N*-graviton state to be self-sustaining; when this relation is saturated, the kinetic energy of each graviton is balanced by the attractive potential energy from the rest of the gravitons—the self-bound gravitons form the long-lived black hole bound state. Since the mass of this bound state is  $M_{\rm BH} = NQ_s$ , it is clear from Eqs. (4) and (5) that the wavelength of the constituent gravitons is equal<sup>3</sup> to the Schwarzschild radius of a classical black hole of mass  $M_{\rm BH}$ ,

$$Q_S^{-1} = R_S = \frac{M_{\rm BH}}{M_P^2}.$$
 (6)

This graviton bound state decays slowly through the loss of its constituent gravitons as a result of their rescattering. As we will soon discuss, their rate of emission matches Hawking's thermal evaporation rate in the semiclassical limit  $N \to \infty$ ,  $R_S = Q_S^{-1} =$  finite.

Another physical interpretation of Eq. (4) is that it represents the value of the coupling for which the *N*-graviton state unitarizes the scattering amplitude for the center-of-mass energy  $\sqrt{s} = M_{\rm BH}$ . Thus the formation of a black hole in the high energy scattering of gravitons can be understood as the creation of a long lived state of highly occupied ( $N \gg 1$ ) albeit weakly coupled ( $\alpha_{\rm gr} \sim 1/N$ ) gravitons. In this quantum portrait of a classical black hole, unitarization occurs at IR wavelengths  $R_S = \sqrt{N}L_P$ , much larger than the UV Planck scale  $L_P$ . Thus the Schwarzschild radius  $R_S$  can be understood as the emergent scale characterizing dynamical unitarization at a fixed point given by Eq. (4), for any  $N \gg 1$ .

#### B. Color glass condensate in QCD

Likewise, for high energy scattering in QCD, the perturbative unitarization line

$$\alpha_S(Q_S)N = 1, \tag{7}$$

corresponds to an emergent saturation scale  $Q_S$ . A physical interpretation of this relation is that with the emission of N soft gluons, the phase space density in the Regge limit of QCD becomes large due to bremsstrahlung. The further emission of soft gluons is compensated by the recombination and screening of soft gluons, leading to a saturation of their growth when the occupancy is given by Eq. (7). This gluon saturation phenomenon [26,27] is equivalently expressed as

$$N \equiv \frac{xG_A(x, Q_S^2)}{2(N_c^2 - 1)\pi R_A^2 Q_S^2} = \frac{1}{\alpha_S(Q_S)},$$
(8)

where  $R_A$  is the nuclear radius,  $xG_A(x, Q_S^2)$  is the corresponding gluon parton distribution, and  $Q_S^2 \equiv Q^2$ , is the resolution scale of the DIS probe that saturates Eq. (7). In the language of the operator product expansion (OPE), this saturation condition is satisfied when "all-twist" contributions to the growth of parton distributions are of equal magnitude to the leading twist contribution.

The equivalence between perturbative unitarization and close packing is manifest in the CGC EFT, where the dynamics of small gluons is represented by classical gluon field configurations with maximal occupancy  $A_{\mu} =$  $O(1/\sqrt{\alpha_s})$  corresponding to solutions of the QCD Yang-Mills equations [31,32,36]. In this framework, Eq. (7) emerges from a computation of the S-matrix for inclusive DIS scattering off a large nucleus at high energies. This Smatrix satisfies a nonlinear renormalization group equation,<sup>4</sup> the Balitsky-Kovchegov (BK) equation [41,42], describing its evolution in rapidity Y (=  $\ln(1/x)$ ) with the emission of small x gluons in the dense gluon environment. For a given impact parameter, the infrared fixed point of the BK equation, corresponding to the unitarization of the cross section, defines the emergent scale  $Q_S \gg \Lambda_{\text{OCD}}$  satisfying Eq. (7), where  $\Lambda_{\text{OCD}}$  is the strong coupling scale. From asymptotic freedom, this of course implies  $\alpha_{\rm S}(Q_{\rm S}) \ll 1$  in the Regge limit, where  $Q_S(x)$  is the largest scale associated with the  $2 \rightarrow N$ process of interest. An important point to note is that

<sup>&</sup>lt;sup>3</sup>There are factors of  $2\pi$  and O(1) constants that can be accounted for in computations in a specific framework. Since the focus here is on the universal features of the CGC-BHNP correspondence, independent of the details of the particular frameworks, we will not attempt to carefully account for these constant factors.

<sup>&</sup>lt;sup>4</sup>This JIMWLK equation describes a hierarchy of n-point Wilson line correlators [37–40] evolving with rapidity. The BK equation is the closed form equation for the lowest two point correlator in this hierarchy, obtained in the limit of large  $N_c$  and large nuclear size  $\alpha_s^2 A^{1/3} \gg 1$ .

"perturbative unitarization" of the  $2 \rightarrow N$  cross section, for fixed impact parameter, occurs at much shorter distances  $\sim 1/Q_S \ll 1/\Lambda_{\text{QCD}}$  than those corresponding to the nonperturbative confining dynamics of QCD.

### C. A digression on the UV/IR nature of gravity

Before we outline the concrete dictionary mapping the CGC to the BHNP, we need to clarify the properties that are intrinsic to the scale dependence of the gravitational interaction. It is often assumed that the high energy theory probes short distances. This is certainly the case in a class of quantum field theories (QFTs) in which "Wilsonian" UV-completion at any scale can be achieved by integrating in weakly interacting degrees of freedom corresponding to shorter and shorter wavelengths, with QCD being a paradigmatic example.

In such Wilsonian QFTs, there exists a clear connection between the energy scale and the distance probed by it. Namely, if we wish to probe a distance *L*, we must perform a scattering experiment with a center of mass energy  $E \sim 1/L$ . Therefore in a theory with Wilsonian UV behavior, if the particles that scatter have sufficiently high energies, we can probe arbitrarily short distance scales.

However, evidently there exists a class of theories [43] that differ dramatically from Wilsonian ones, of which gravity is a prominent representative [44]. In such theories, above certain scale, the scattering is dominated by formation of classical objects, consisting of high occupation number of "soft" quanta. This dynamics effectively prevents the theory from entering into a short distance regime. This concept was referred to [43] as the "UV-completion by classicalization."

In gravity, increasing the center of mass energy is no guarantee for probing short distances because of the existence of black holes. Instead, the formation of black holes suggests that scales shorter than the Planck length cannot be probed [44]. This can be understood already at the semiclassical level and, more profoundly, at the level of multigraviton scattering amplitudes.

The semiclassical argument is simple [45]. As soon as the initial center of mass energy gets localized within its Schwarzschild radius, a black hole forms. In this way, effectively, the scattering experiment with a center of mass energy  $\sqrt{s}$ , probes distances  $L \sim \sqrt{s}/M_P^2$  rather than  $L \sim 1/\sqrt{s}$ . That is, deep UV gravity, via black holes, effectively is converted into a deep IR gravity [44].

The black hole *N*-portrait [3] offers a microscopic understanding of this phenomenon. In field theoretic language, the process of black hole creation means that the initial center of mass energy of colliding hard quanta gets distributed among a large number *N* of soft gravitons, with wavelengths  $Q_S^{-1} \sim \sqrt{s}/M_P^2$ . In this way, the momentum transfer in each elementary vertex is minuscule. Since the graviton coupling [see Eq. (5)] probed in such an experiment is weak, it is easy to see that it satisfies the saturation relation in Eq. (4). Because weak coupling is accompanied by saturation, the physics of this regime is highly nontrivial; in particular, the occupation number of quantum and the entropy of the N-graviton state have to be just right for this collective effect to be realized.

Note that the above applies equally to effective field theories of quantum gravity as well as its embedding in string theory. String theory fully accommodates the non-Wilsonian asymptotic behavior of gravity. As long as the Schwarzschild radius corresponding to the center of mass energy is much larger than the string length, black hole formation in string theory will proceed exactly as in effective field theories of gravity. This expectation is confirmed by computations of multigraviton amplitudes—supposedly describing black hole formation—which agree when computed in both QFT and in stringy regimes [19].

Our above discussion suggests that while one can say that a black hole state probes gravity at very high energies, it would be a mistake to call it a UV-regime of gravity. The relevant scale (which is an IR scale) is set by the frequencies of the constituents and not by the overall energy of the system.

It is also important to not confuse the role of the Planck scale, which represents the strength of graviton coupling. In other words, as we will discuss further, it has to be viewed as the graviton decay constant. As pointed out in [11], when performing the mapping between a black hole and a generic saturated system, the Planck mass must be mapped on the decay constant f of the Goldstone boson of the spontaneously broken Poincaré symmetry in that system. This Goldstone boson must be universal since any saturated system spontaneously breaks Poincaré symmetry.

#### **D.** Dictionary mapping gravity in the UV to QCD in the IR

It is now straightforward to establish the dictionary between the BHNP and the CGC. Both states are uniquely determined by the corresponding saturation equations in Eqs. (4) and (7) respectively. In both cases, the saturation scale  $Q_s$  determines the typical momenta of constituent gravitons and gluons respectively. With this dictionary, a region of space of radius  $R_s = Q_s^{-1}$  of the CGC can be mapped on to a black hole of radius determined by the saturation of the graviton condensate.

Key features of the dynamics mapped between the two systems exhibit UV/IR relations that become apparent when we consider the behavior of the underlying theory in these two regimes. In gravity, the UV scale is given by the Planck scale  $M_P$  and in QCD the corresponding IR scale is  $\Lambda_{\text{OCD}}$ .

The significance of these scales is of course that the respective theories enter into the strong coupling regimes when the physics is sensitive to them. On the other hand, since the unitarization relations in Eqs. (4) and (7) are

saturated at weak coupling, the saturation scale  $Q_S$  is very far at large N from the corresponding strong coupling scale:

for BHNP: 
$$Q_S^{-1} \gg M_P^{-1}$$
,  
for CGC:  $Q_S \gg \Lambda_{\text{QCD}}$ . (9)

in the UV and IR respectively. As we will soon discuss, the saturation scale in both cases exhibits the same relation with the decay constant f of the Goldstone boson of the broken Poincaré symmetry resulting from the creation of these high occupancy states:

$$f = \sqrt{N}Q_S \tag{10}$$

The inverse of this scale  $f^{-1}$  in both theories is equal to  $\sqrt{\alpha}$  times the screening length of the corresponding degree of freedom (graviton or gluon). For a black hole, this corresponds to  $f = M_P$ .

The lesson here is that the form of the correspondence between BHNP and CGC depends on the resolution scale Qat which we probe these states. When  $Q = Q_S$ , the correspondence is direct, and driven by the dynamics of saturation at weak coupling in both theories, with  $\alpha = 1/N \ll 1$ .

However UV/IR type differences emerge whenever we discuss the relations between the saturation momentum  $Q_s$  and the strong coupling scales in the two theories, once we move to much shorter distances from  $Q_s^{-1}$  in gravity or much longer distances in QCD. Hence in order to maintain their qualitative similarities while changing the scale, we must synchronize the signs in the variations of the couplings in the two theories. That is, a motion toward the UV in the BHNP must be mapped on the motion toward the IR in the CGC and vice-versa.

Of course, away from the saturation point, the correspondence is lost in general. However it is conceivable that, on the weak coupling side, features of this correspondence may benefit from Lipatov's EFT for Reggeized gravitons and gluons [20,21], the previously noted BCJ relations [2] and their extension to classical "double copy" relations [46,47].

# **III. CONNECTING ENTROPIES IN THE BHNP AND IN THE CGC**

The connection between the microstate entropies of CGC and BHNP can be summarized by the fact that both satisfy Eqs. (2) and (3). According to [11], (see also, [9,10]) the above equalities are universally satisfied by critically packed states [defined by Eq. (1)] that saturate the unitary bound on entropy, resulting into the combined saturation relations:

$$S = N = \frac{1}{\alpha(Q)} = \operatorname{Area} \times f^2, \qquad (11)$$

where, as before, f is the decay constant of the Goldstone boson of broken Poincaré symmetry. We will first reproduce the general argument of [11] for the validity of this relation and then specialize the discussion to black holes and the CGC.

#### A. General argument

Consider a critically packed state of occupation number  $N = \alpha^{-1}$  of quanta with momentum  $Q_s$  and volume  $R_s^3 = Q_s^{-3}$ . We shall assume that the entropy is given by Eq. (2). Let us now show that the same entropy is also equal to the area of the region (=  $4\pi R_s^2$ ), in units of a Goldstone decay constant f. The Goldstone modes in question originate from the fact that any critically packed condensate spontaneously breaks symmetries both in coordinate and in relevant internal spaces. As discussed in Appendix A, the order parameter of this breaking is given by  $NQ_s^2$  and the decay constant of the canonically normalized Goldstone mode is

$$f = \sqrt{N}Q_S. \tag{12}$$

From this relation,<sup>5</sup> it is clear that the entropy in Eq. (11) of the region filled with the condensate is nothing but the area of that region (Area =  $4\pi R_s^2 = 4\pi Q_s^{-2}$ ) measured in units of  $f^2$ .

Notice, as a by-product, the above relation implies the saturation of the Bekenstein bound on entropy. Indeed, using the saturation relation in Eq. (2) and the critical packing relation in Eq. (1) we can write,

$$S = \alpha^{-1} = NQ_S R_S = ER, \tag{13}$$

This last expression is nothing but the energy of the condensate,  $E = NQ_S$  times its size  $R_S$ . We have therefore the well-known expression [48] for the Bekenstein entropy.<sup>6</sup>

#### **B.** Entropy of a black hole

We shall now apply the above reasoning to the BHNP. The observation that the Bekentein-Hawking entropy of a black hole is equal to the inverse graviton coupling evaluated at the momentum transfer  $Q_S = 1/R_S$  (or equivalently, to the number N of black hole constituents) was made in [3]. Indeed, it is obvious from Eq. (5) that the coupling  $\alpha_{gr}$  is nothing but the inverse of the black hole area in Planck units. Thus the black hole entropy is not just an area but also the inverse of the gravitational coupling:

<sup>&</sup>lt;sup>5</sup>This behavior was explicitly demonstrated in [10,11] for several field theory examples of extended objects such as solitons, instantons and baryons in large- $N_c$  QCD.

<sup>&</sup>lt;sup>6</sup>As per our remarks in footnote 3 we omit the factor of  $2\pi$  in this expression.

$$S_{\rm BH} = \frac{1}{\alpha_{\rm gr}} = \frac{4\pi R_S^2}{4L_P^2}.$$
 (14)

This is a particular form of the more general relation in Eq. (11). This way of understanding the black hole entropy connects *the geometrical notion of an area to the quantum field theoretic concept of the graviton coupling*.

Further, the BHNP reveals the microscopic meaning of the simultaneous saturation of the entropy bound and unitarity. With regard to the latter, as noted previously, the S-matrix for black hole formation in the scattering of two trans-Planckian quanta with center of mass energy  $\sqrt{s} \gg M_P$  is that of  $2 \rightarrow N$  scattering, where the final Ngraviton quanta have momenta given by  $Q_S = \frac{\sqrt{s}}{N} = R_S^{-1}$ . A detailed computation of this process, both in quantum field theory as well as in string theory, was performed in [19]. At large N, both computations at the black hole formation threshold give the following rate for the  $2 \rightarrow N$  process:

$$\Gamma_{2 \to BH} \sim e^{-\frac{1}{\alpha_{gr}} + S_{BH}}$$
(15)

where  $\alpha_{\rm gr} = \frac{Q_s^2}{M_P^2} = (R_S M_P)^{-2}$ . The exponential factor  $e^{S_{\rm BH}}$  comes from the summation over black hole microstates. The form of the rate in Eq. (15) highlights the role of Eq. (14) in the saturation of unitarity by black holes. The bottomline is that the black hole entropy is equal to the inverse graviton-graviton coupling at the point where unitarity is saturated, in full accord with Eq. (11).

Finally, as discussed in [10,11], the black hole entropy given by Eq. (14), is a particular example of the much more general expression in Eq. (3) for the microstate entropy of a saturated system. As already noted, this expression states that the maximal entropy of an object is equal to the area in units of the Goldstone decay constant.

In the case of a black hole, as discussed in [11] and in Appendix A, the Goldstone mode is the graviton itself. Indeed, the saturated state of the graviton condensate breaks Poincaré invariance spontaneously. The order parameter of this breaking is  $NQ_S^2 = N/R_S^2$  and the decay constant of the corresponding Goldstone mode in Eq. (12) is equal to  $f = \sqrt{N}/R_S = M_P$ , which is nothing but the Planck mass. This result makes physical sense because on the one hand, the Planck mass is the graviton decay constant, and on the other, the black hole is a graviton state of maximal occupation number so that the expectation value of the graviton field is  $M_P$ . The last statement is clear already from the classical theory, as this is the value attained by the Schwarzschild metric tensor relative to the Minkowski metric.

#### C. Entropy of the CGC

We now wish to argue that Eq. (11) is also satisfied by the CGC. We shall apply the generic reasoning of [11] which shows that states saturating unitarity must also saturate the entropy bounds Eqs. (2) and (3), and therefore Eq. (11).

As we noted previously, the CGC is a state of maximal occupancy in the proton that is excited in DIS or hadronhadron scattering in the high energy Regge limit of QCD. In this  $2 \rightarrow N$  scattering, a high energy probe effectively interacts with a lumpy deconfined configuration of gauge fields  $(A_{\mu} \sim O(1/\sqrt{\alpha_S}))$  with a typical correlation length  $R_S \sim 1/Q_S \ll 1/\Lambda_{\rm QCD}$  [5]. From the "perspective" of the probe, the wee gluons in the boosted proton are Lorentz contracted to a distance<sup>7</sup>  $1/Q_S$ , in contrast to the high energy valence parton modes that are contracted by  $2R_p/\gamma$ , with  $\gamma \sim \sqrt{s}/m_p$  the Lorentz factor of the valence parton modes, and  $R_p$ ,  $m_p$  the proton mass and radius respectively.

With this in mind, let us consider the CGC spatial volume  $V_{\text{CGC}} = R_S^3 = Q_S^{-3}$  of saturated wee gluon modes. The energy contained in this region, as seen by the high energy probe, is then simply

$$E_{\rm CGC} = NQ_S. \tag{16}$$

Now just as in the black hole case, we can think of the CGC as an approximately classical macrostate representing a large multiplicity of quantum "parton" microstates. In other words, these microstates are indistinguishable classically but represent distinct quantum states, which constitute the entropy  $S_{\text{CGC}}$  of the CGC in a region of rapidity  $Y = \ln(1/x)$  relative to the probe.

As we noted previously, explicit weak coupling computations show that the S-matrix for  $2 \rightarrow N$  scattering is saturated for fixed impact parameter in the Regge limit of QCD. The physical meaning of this key piece of information is as follows. The statement that the CGC saturates unitarity (for a given impact parameter) means that its microstates account for the *entire portion* of the Hilbert space<sup>8</sup> describing the complete set of QCD states localized within the region  $V_{CGC}$  with energy  $E_{CGC}$ . Indeed, if QCD matter could exist in other possible states within the same region  $V_{CGC}$ , this would imply that our statement about the CGC saturating unitarity would be false since such additional non-CGC states would be formed in the same scattering process with nonzero probability, thereby violating the unitarity of the CGC state.

The scaling of the CGC entropy as in Eq. (2) can be understood as follows. When *N* is very large, the CGC can be treated as a classical condensate. Thus when forming CGC, say, in the scattering of gluons in the Regge limit, we

<sup>&</sup>lt;sup>7</sup>Note that since  $Q_S \equiv Q_S(x)$ , this picture is invariant under boosts.

<sup>&</sup>lt;sup>8</sup>A caveat is that this statement is true up to power law suppressed IR modes (in gravity) and UV modes (in QCD) at a given energy scale. As the energy of the collision increases, these modes are absorbed into the respective evolution of the saturation scale in both theories.

The first way is to think about the process semiclassically. That is, we assume that the bulk of the suppression in forming the state comes from a classically forbidden Euclidean trajectory. This aspect is similar to any textbook semiclassical computation of tunneling, a classic example being Coleman's treatment of false vacuum decay [50].

However in the present case there is a slight technical complication. Ordinarily, in a false vacuum decay, both initial and final configurations are classical. Correspondingly, the classical Euclidean trajectory, a so-called "bounce," is relatively straightforward to find by solving the classical equations of motion in Euclidean space. Such a luxury does not exist in the present case, since the initial two-gluon state cannot be described classically. However this is not a serious obstacle, since the physics of the suppression remains unaltered and can be reliably estimated even without having a full-fledged saddle point classical solution in Euclidean time.

That is, even in the absence of such a solution, we can still think of the process as "happening" in the Euclidean time and thus being described by some intermediate Euclidean trajectory. This virtual trajectory, close to the initial two-gluon state, cannot be described classically. However, it is reasonable to assume that this initial quantum portion contributes very little in the Euclidean action, as compared to the part of the trajectory that is well described classically. The minimal action of this classical portion can be estimated without having an exact solution at hand.

Since  $E_{CGC} = NQ_S$  is the energy of the CGC, and  $R_S$  is the typical "duration" of the Euclidean time required for its formation, a reasonable guess for the minimal action of the relevant Euclidean trajectory is  $NQ_SR_S = N$ .

The alternative way [49] to arriving at the same expression is by thinking of the Euclidean action needed to create N particles of momentum  $Q_S$  from an initial two-particle state. The action per particle is the product of the de Broglie wavelength and energy,  $Q_S R_S \sim 1$ . Therefore, for N particles this gives the total minimal action  $\sim N$ .

The bottom line in the above estimates is that the probability of creating each microstate of CGC is suppressed by

$$\mathrm{e}^{-N} = \mathrm{e}^{-\frac{1}{\alpha_S}}.\tag{17}$$

This expression is equivalent to the universal suppression factor for a unitary  $2 \rightarrow N$  transition to a generic *N*-particle

microstate, derived from general consistency considerations. $^{9}$ 

Explicit computations [38,39,51,52] of small fluctuations on top of the classical CGC solution demonstrate that the  $\alpha_s$  suppression is compensated by the phase space in rapidity for gluon emission; such contributions are  $\alpha_s Y \sim$ O(1) for  $Y \sim 1/\alpha_s$ . The resummation of such contributions, in the classical background, to all orders, leads to the BK/ JIMWLK equations, which as noted previously, unitarizes the  $2 \rightarrow N$  scattering amplitude.

The fact that the CGC saturates unitarity then implies that the number of microstates ( $\sim e^{S_{CGC}}$ ) must compensate the suppression in Eq. (17). Hence it follows that the entropy of the CGC in the volume  $V_{CGC}$  must be

$$S_{\rm CGC} = \frac{1}{\alpha_s} \equiv N. \tag{18}$$

This demonstrates that the entropy of the CGC satisfies the first equality in the universal relation in Eq. (11), or equivalently, saturates the bound in Eq. (2).

Of course, by the generic reasoning already given above, as an added bonus, Eq. (18) automatically implies the saturation of the Bekenstein bound,

$$S_{\rm CGC} = E_{\rm CGC} R_S. \tag{19}$$

We will now argue that (just as for the black hole case) the entropy does not scale with the volume of the CGC but instead as its area in units of a decay constant f of the Goldstone boson corresponding to the breaking of Poincaré invariance by the CGC. This follows from the generic argument of [11] leading to Eq. (11), also discussed in Appendix A. Here we shall apply this reasoning to the specifics of the CGC.

In QCD, partons are characterized by hard transverse momenta  $p_{\perp} \sim Q$ , where Q is the (UV) momentum resolution scale of the probe. The longitudinal momenta on the other hand (depending on the x values probed) can range from  $\Lambda_{\rm QCD}$  to Q in QCD's Bjorken limit, where the parton concept is robust.

In the Regge limit, the picture changes qualitatively due to the emergent saturation scale  $Q_s$ . For momentum modes below this scale, the parton concept is not viable and the corresponding microstates are indistinguishable from a classical field. A simple computation in the CGC shows that the typical transverse momentum is of order  $Q_s$ . Interestingly, the softest longitudinal momentum modes that carry any field strength also typically have momenta  $O(Q_s)$ . In other words, as the proton is boosted, the valence modes seen by a probe shrink by the Lorentz factor;

<sup>&</sup>lt;sup>9</sup>That derivation [11], which will not be reproduced here, does not rely on a particular assumption about the structure of the theory. The matching with it therefore provides a consistency check of the above estimate.

however, the wee modes shrink as well, albeit more slowly and with a rate computable in the CGC, and their field strengths in the longitudinal direction are localized on the lightcone at  $x^- \sim 1/Q_s$ , or  $p^+ \sim Q_s$ .

There is however an intermediate hard scale between  $Q_S$ and Q which characterizes the transition from a gapped dilute gas "parton gas" picture to that of the overoccupied CGC classical field. Partons below this scale are screened or, in QCD language, are "shadowed"; in other words, many-body (or higher twist, in OPE language) effects become important. Just as for a thermal gas, the screening of parton modes breaks the translational invariance of the parton gas; the Goldstone scale f corresponding to this breaking of Poincaré invariance is exactly analogous to that in the BHNP. Indeed, as argued in [11], on general grounds, this scale is of order  $f \sim \sqrt{N}Q_S$  for any classical "solitonlike" lump with characteristic momentum  $Q_S$  and the occupation number N. We shall discuss these arguments further in Appendix A. Thus,

$$S_{\rm CGC} = \frac{1}{\alpha_s} = N = f^2 \text{Area}, \tag{20}$$

since Area  $\propto R_s^2$ . We thus arrive at a universal relation (11) characteristic of a generic saturates system [11]. As we shall discuss further, it is important to note that this entropy is not the entanglement entropy (which is generically expected to obey an area law in units of the UV cutoff in the theory) but the microstate entropy with an area counted in units of f. This is highly nontrivial since dynamical screening is how the CGC restores unitarity; the breaking of Poincaré invariance leads to area scaling of the entropy as opposed to volume scaling, with the scale set by the Goldstone constant f. As is evident, the above expressions are identical to the ones satisfied by the black hole entropy and possess the same microscopic meaning as revealed by the quantum *N*-portrait.<sup>10</sup>

For QCD dynamics in the Regge limit, from solutions of the "mean field" CGC BK equation,<sup>11</sup> one can deduce this scale to be closely related to  $p_{\perp,\text{shdw}}^2 \sim Q_s^4 / \Lambda_{\text{QCD}}^2$  [57]. The latter can be expressed as  $p_{\perp,\text{shdw}}^2 \sim f^2 \alpha_s n_D$ , where  $n_D = Q_s^2 / \Lambda_{\text{QCD}}^2$  is the number of domains in the proton.

Since the entropy of the CGC is given by its area and not its volume, the maximal total entropy of wee partons released from scattering off the proton is given by the ratio of the transverse area of the proton  $\propto \frac{1}{\Lambda_{QCD}^2}$  to that of the CGC,

$$S_{\text{tot}} = \frac{1}{\alpha_S} \frac{1}{(R_S \Lambda_{\text{QCD}})^2} = n_D S_{\text{CGC}},$$
 (21)

where we employed Eq. (18). We assumed here that, since  $R_S$  dynamically screens color charge at larger distances, the microstates of each of the  $n_D$  domains, given by the ratio of the area of the proton ( $\sim 1/\Lambda_{\rm QCD}^2$ ) to that of a domain ( $R_S^2$ ), are not correlated. Hence the total entropy of the proton is given by the number of domains times the entropy of each.

Note that the occupation number N (implicitly through the running of the coupling) as well as  $n_D$  and  $Q_S$  (explicitly) depend on the intrinsic transverse size of the proton  $\Lambda_{QCD}$ , so the relation between N and  $n_D$  is an additional input that must self-consistently satisfy the relations in Eqs. (7), (18), and (21). In other words, the universality we posit is between the black hole state and a lump of CGC and therefore in principle independent of the number of saturated lumps  $n_D$  that can fit within the proton. Of course, the running of the coupling in QCD ensures that the two are not entirely independent of each other.

When entropy is released in the lab frame of the scattering, it is appropriate to describe it in terms of the rapidity variable, which adds additively under boosts between frames.<sup>12</sup> In the bremsstrahlung process, on average, one parton is released in  $1/\alpha_s$  units of rapidity:  $\alpha_s dY \sim 1$ . Hence the boost invariant distribution of entropy in the scattering process is given by<sup>13</sup>

$$\frac{dS_{\text{tot}}}{dY} = n_D(\alpha_S). \tag{22}$$

There have been several discussions in the literature on the entropy of the CGC and the proton at small x [60–68]. Comparisons of our results to these are not straightforward since several employ different definitions of the entropy or address different kinematic regions in high energy

<sup>&</sup>lt;sup>10</sup>This unitarity limit of maximal screening in the high energy QCD literature is very appropriately referred to as the "black disc" limit—for a nice discussion of the different regimes, see [53].

limit—for a nice discussion of the different regimes, see [53]. <sup>11</sup>The parton distributions corresponding to the solutions of this equation indeed have the structure of a propagating solitonlike wavefront [54,55]. The kinematic region between  $Q_s$  and  $p_{\perp,shdw}$  is often referred to as the "geometrical scaling" or "leading twist shadowing" regime in the small *x* QCD literature, and there are strong arguments that small *x* DIS data from HERA can be interpreted in this light [56,57].

<sup>&</sup>lt;sup>12</sup>Indeed, as argued in [58,59], it is only meaningful to rigorously define this entanglement entropy for a finite interval of rapidity.

of rapidity. <sup>13</sup>It would be nice to compare our estimates of the microstate entropy of CGC (which is determined by the number of nearly degenerate pure microstates) with the entanglement entropy in the CGC using the techniques discussed in [59–62]. Since the typical scaling of the entanglement entropy in local field theories is given by the area in units of the cutoff of the theory, we expect that the two must match, at least when entanglement entropy is computed in the EFT of the Goldstone modes. Indeed, the natural cutoff of this effective theory is the Goldstone decay constant  $f = \sqrt{NQ_s}$ . Thus, the entanglement entropy of the CGC in a volume  $Q_s^{-3}$ , must scale as area in units of f, thereby matching the general bound in Eq. (3).

scattering.<sup>14</sup> The closest treatment to ours is that of [67] which likewise addresses the Von Neumann entropy obtained from the counting of the quantum microstates of the CGC. In deeply inelastic scattering (DIS), the relevant quantity is the entropy of a lump of CGC  $(n_D = 1)$  given in Eq. (18) that unitarizes DIS cross section in Regge asymptotics:

$$S^{\text{DIS}} = \frac{1}{\alpha_S(Q_S)} \to \frac{dS^{\text{DIS}}}{dY} = \text{constant for } Y \le \frac{1}{\alpha_S},$$
 (23)

to leading logarithmic accuracy. This result, which is valid in a finite rapidity interval, is consistent with that of [67]. In contrast, in hadron-hadron (h-h)collisions, since several lumps of CGC are produced, the entropy released is given by

$$S^{\mathrm{h-h}} = n_D \frac{1}{\alpha_S(Q_S)} \to \frac{dS^{\mathrm{h-h}}}{dY} = n_D.$$
 (24)

This estimate is consistent with previous estimates of the glasma entropy [6,69,70] at  $t = 1/Q_s$  if there is a solution to the self-consistency relation<sup>15</sup>  $n_D \sim N = 1/\alpha_s$ . For collisions of heavy-nuclei, for rapidities  $Y \leq$ 

 $\ln^2(A^{1/3})$ , this estimate should be multiplied by the atomic number A since  $Q_{S,A}^2 = \Lambda_{\text{QCD}}^2 A^{1/3}$  [26,27,31] and  $R_A \propto A^{1/3}$ . However at asymptotically high energies corresponding to  $Y \gg \ln^2(A^{1/3})$ ,  $Q_S^2$  becomes independent of A [71]; in this asymptotics, the entropy will scale as the area of the nucleus  $R_A^2 \propto A^{2/3}$ .

# IV. GOLDSTONE ORIGIN OF ENTROPY AND THE TIMESCALE OF INFORMATION-RETRIEVAL

In this section, following [11] [9,10], we shall give a Goldstone interpretation of the quantum information carrying modes that are responsible for the microstate entropy of a generic system. This provides a deeper understanding of the appearance of the Goldstone decay constant f in the area law scaling of the entropy in Eq. (3). The Goldstone interpretation of entropy was originally introduced in the context of BHNP and the key essence was also illustrated for a critically packed condensate of bosons in [72].

As a further step, it was shown in [9–11] that the microstate entropy of systems in ordinary gauge theories, and its connection with unitarity, can be understood and explained in terms of the Goldstone phenomenon.

This Goldstone formulation allows one to derive a universal expression [11], presented below, for the timescale required for retrieving quantum information stored in microstates of the system. In particular, this knowledge was applied to saturated multiparticle states such as lumps of gluons or baryons in QCD with a large number of colors. Due to the universality of the Goldstone phenomenon, the same should be applicable to the CGC.

Let us consider the critically packed state [Eq. (1)] of some gauge degrees of freedom. We shall assume that the state is self-sustained, meaning that its "life-expectancy" is a growing function of N. In certain cases, the configuration may be exactly stable due to conserved quantum numbers. For example, this can be the case for topological solitons and baryons (Skyrmions). For such a system, the infinite-Nlimit is well-defined and, in this limit, corresponds to a classical background configuration which we shall generically refer to as a "lump."

The first key point of [10,11] is that such a field configuration necessarily breaks spontaneously a set of global symmetries. As already discussed, among the broken symmetries is Poincaré invariance, as the condensate transforms nontrivially under it. However in addition, there exist broken global symmetries that parametrize the embedding of the lump into the group space.

At infinite N, the notion of a spontaneous breaking of symmetry is exact and the lumps with different internal orientations can be viewed as different "vacua." They are connected by the broken global symmetry transformations and are strictly degenerate in energy. Correspondingly, there exist associated Goldstone modes. These modes are gapless at  $N = \infty$ . At finite N, the Goldstones remain gapless, at least to leading order in 1/N.

The lumps corresponding to different orientations form different *microstates* of the same *macrostate*. The resulting microstate entropy is defined as the log from the number of independent orientations  $n_{\rm st}$ ,

$$S = \ln(n_{\rm st}). \tag{25}$$

The different orientations are obtained from one another by the excitations of the Goldstone modes. This fact makes it transparent that the Goldstone modes are the carriers of the microstate entropy.

We next present a universal expression for the timescale required for the retrieval of the quantum information, derived in [11]. Since quantum information is stored within the Goldstone modes, the retrieval of this information requires the read-out of this Goldstone state which in turn necessarily involves interactions with the Goldstone modes. Thus the processing of quantum information cannot happen on timescales shorter than the Goldstone interaction time given by

$$t_{\text{Goldstone}} = \frac{1}{\Gamma_{\text{Goldstone}}} \sim f^2 R_S^3 = N R_S, \qquad (26)$$

where  $\Gamma_{Goldstone}$  is the Goldstone interaction rate (the cross section times the density of such modes) for the characteristic

<sup>&</sup>lt;sup>14</sup>We would like to thank V. Skokov and Z. Tu for discussions

pertaining to these issues. <sup>15</sup>Note that this assumption would also give us  $p_{\perp,\text{shdw}}^2 = f^2 \alpha_S n_D \sim f^2$ .

momentum transfer scale  $Q_s = 1/R_s$ . This expression therefore sets an universal lower bound on the timescale required for retrieving quantum information stored within the system. The Goldstone argument explains why this timescale is macroscopic and is determined by the number of constituents N. From the saturation relation [Eq. (11)], this timescale can be expressed as

$$t_{\rm Goldstone} = \frac{R_S}{\alpha} \tag{27}$$

As noted in [11] for the particular case of the BHNP, the expression in Eqs. (26) and (27) reproduce the timescale proposed by Page [73] for the time required to retrieve information from a black hole. These timescales also agree with the half-decay time of a black hole and, as we shall now discuss, with its quantum break-time [74].

# V. COMPARING TIMESCALES FOR BLACK HOLE DECAY AND FOR GLASMA EVOLUTION TO A QUARK-GLUON PLASMA

## A. Semi-classical and quantum portraits of black hole life time

In black hole physics, an extremely important timescale can be expressed as follows,

$$t_{\rm BH} = R_S \frac{R_S^2}{L_P^2} = NR_S = \omega^{-1} = \alpha_{\rm gr}^{-1} R_S, \qquad (28)$$

where, as we will now discuss, the four equivalent ways of writing this timescale will be convenient for making transparent the physical meaning of each.

The first relation  $t_{\rm BH} = R_S \frac{R_S^2}{L_P^2}$  is physically meaningful for two reasons, neither of which requires any knowledge of the microscopic theory. The first macroscopic interpretation of  $t_{\rm BH}$  is that of the half-decay time of a black hole due to its semiclassical evaporation via Hawking radiation [75]. The latter can easily be obtained from the Stefan-Boltzmann law applied to Hawking's thermal evaporation relation,

$$\frac{dM_{\rm BH}}{dt} = -T_H^2,\tag{29}$$

where  $T_{\rm H} = R_{\rm S}^{-1} = M_P^2 M_{\rm BH}^{-1}$  is the Hawking temperature. We should note that the microscopic theory of the BHNP reveals that Eq. (29) is valid until the half-decay time. Beyond this point, the semiclassical analysis can no longer be trusted.

The second macroscopic interpretation of  $t_{BH}$ , as noted previously, is that of Page [73], who argued that  $t_{BH}$  is the minimal time required for an observer to be able to resolving outgoing information at an order one rate.

The other three ways of writing  $t_{\rm BH}$  in Eq. (28) correspond to three different microscopic interpretations of the

same physics that is provided by the BHNP. Two of these clarify our microscopic understanding of the above interpretations  $\dot{a}$  la Hawking and Page respectively and reveal underlying structure not captured by the semiclassical theory. The final microscopic relation is noteworthy because it provides an interpretation of  $t_{\rm BH}$  that has no equivalent counterpart within the semiclassical picture of a black hole.

We shall now consider each of these three microscopic relations separately. When we write Eq. (28) as  $t_{BH} = NR_S$ , it describes the time during which the black hole graviton condensate loses an O(1) fraction of its initial N constituents. Indeed according to the BHNP, for each Hawking emission time of  $\sim R_S$ , the condensate on average loses one constituent, significantly depleting over  $\sim N$  emissions. Hence this form of  $t_{BH}$  provides a microscopic explanation to the half-life of a black hole obtained in Hawking's semiclassical theory.

The second microscopic relation in Eq. (28),  $t_{BH} = \omega^{-1}$ , equates the left-hand side to the inverse of  $\omega \sim \frac{1}{S_{BH}R_s} \sim \frac{1}{NR_s}$ , which represents the characteristic energy gap of information-carrying Bogoliubov-Goldstone modes of the graviton condensate [4,72]. This relation then gives a transparent microscopic interpretation of Page's time: Information encoded in a quantum degree of freedom cannot be resolved faster than the inverse of its excitation energy. As we discussed already, Eq. (28) is particular example of Eq. (26), stating that information-retrieval time from an arbitrary system is given the interaction time of information carrying Goldstone modes [11].

The last equality in Eq. (28),  $t_{BH} = \alpha_{gr}^{-1} R_S \equiv t_Q$ , provides us with a qualitatively different microscopic meaning of  $t_{BH}$  being the *quantum break time*. This notion has no analog in semiclassical theory and describes a timescale beyond which the quantum evolution of the black hole departs fully from the classical one [74]. As argued in [76], this timescale is characteristic of critically packed systems for which it has the following simple form,

$$t_Q = \frac{1}{\alpha Q_S},\tag{30}$$

where  $Q_S = 1/R_S$  is the characteristic momentum scale and  $\alpha$  is the coupling. For the BHNP,  $\alpha = \alpha_{\rm gr}$  and  $R_S$  is the gravitational radius.

Again, the similarity between the quantum break-time Eq. (30) and minimal time of information retrieval Eq. (27) has a deep physical meaning as both processes are controlled by the strength of quantum interactions [11].

However the phenomenon is very general and takes place even when more than one scale is present. In such cases, the parametric dependence on  $\alpha$  may change. For example, it was shown in [74] that in the presence of classical time evolution, or Lyapunov instabilities, the quantum break time can be significantly shorter relative to that given in Eq. (30).

The origin of quantum breaking time can be understood in general terms by viewing the semiclassical BHNP as a many-body system [76]. In essence, it arises from the internal rescattering of the constituents of the condensate which leads to its depletion as well as to the generation of entanglement. It also results in an effective "eigenstate thermalization" [77,78]. Although the quantum state of the system is always pure, it evolves close to a thermal state in the sense that distinguishing it from a true thermal state requires precision measurements over macroscopically long timescales. We will discuss these points further shortly.

In the BHNP framework, the equivalence of Eqs. (28)–(30) can be understood as follows. The rate of rescattering of a pair of constituents can be estimated as

$$\Gamma \sim Q_S \,\alpha_{\rm gr}^2 N^2 \sim R_S^{-1},\tag{31}$$

where the characteristic momentum transfer in our case is  $Q_S$  and the Bose enhancement factor  $\sim N^2$  comes from the number of pairs. Taking into account Eq. (11), during the time  $t_Q \sim N\Gamma^{-1} \sim NR_S$ , an O(1) fraction of constituents rescatters and the condensate half depletes. This gives Eq. (30) and simultaneously establishes its equivalence to Eq. (28). Most of the depleted gravitons leave the condensate and escape in the form of Hawking quanta. With each scattering and emission, the remaining constituents become more and more entangled among each other. The entanglement reaches a maximal value after most of the gravitons have experienced collisions, with timescale given by Eq. (28). Beyond this point, the semiclassical description of the system breaks down completely and it "quantum breaks."

# B. Quantum breaking of glasma formed in collisions of CGCs

In the QCD case, the thermalization process is best studied in the ultrarelativistic collisions of large nuclei, which can be represented as colliding CGC shockwaves.<sup>16</sup> In these collisions, a macroscopic deconfined region of high occupancy gluon fields called the glasma [6,7] is formed on a timescale  $\tau \sim 1/Q_s$  corresponding to the longitudinal size of the wee gluon modes in the shockwaves. How this glasma evolves and the thermalization

process of the system into a quark-gluon plasma (QGP) is an outstanding problem [79].

In the CGC picture, the collision of the shockwaves, and the subsequent evolution of the glasma, can be studied numerically in a classical-statistical approximation at early times when the occupancies  $N \gg 1$  [80]. As we will discuss further later, Weibel-like plasma instabilities [81] play a significant role in the decoherence of the glasma, which then subsequently evolves, just as in the BHNP, through the rescattering of its constituents. Detailed 3 + 1 - D numerical simulations [82,83] of the glasma strongly suggest that this scattering occurs through a "bottom-up" thermalization scenario [84] which we will now outline below.

We shall closely follow [84] with a slight change to their notation, with  $N_h$  of [84] denoting here the occupation number instead of the number density, with the latter denoted instead by the lowercase  $n_h$ . The evolution in [84] starts from an initial saturated state of gluons that are referred to as "hard," corresponding to momenta  $p \sim Q_s$ . Their number density in the initial glasma stage decreases due to two effects. The first is the one-dimensional "classical" expansion<sup>17</sup> of the system which changes the number density as,

$$n_h \sim \frac{Q_S^3}{\alpha_S} \frac{1}{(Q_S t)},\tag{32}$$

where *t* must be understood as the proper time. (In the absence of interactions, the occupation number is of course conserved and is given by  $N_h = \frac{n_h}{Q_s^2 p_z} = \frac{1}{\alpha_s}$ .) The other effect influencing the decrease in the number density is the quantum scattering of gluons.

The only difference in the evolution of the glasma to that of a black hole is the absence of the 1-D expansion. As explained above, the sole mechanism for a black hole to lose its quantum constituents is through the depletion of the graviton condensate by their rescattering. The gravitons that leave the condensate in this way are responsible for Hawking radiation. The depletion of the graviton condensate due to rescattering is strikingly similar to the initial time evolution of the CGC/glasma, which we shall now review. Of course, unlike Hawking quanta, the rescattered gluons cannot escape to infinity unscathed, but form a QGP first. This is an inessential difference, as we shall discuss, from the "early-time" perspective of our UV/IR correspondence. The rest of the early-stage dynamics in the two systems are qualitatively identical.

When the occupancies are large, the net effect of number changing processes is small, and the scattering of the hard gluons is described by  $2 \leftrightarrow 2$  kinetic processes with the lowest momentum exchanged given by the Debye mass

<sup>&</sup>lt;sup>16</sup>One can also formulate the issues here for hadron-hadron scattering at sufficiently high energies where  $R_S \ll R$ , the hadron radius. While in the asymptotic Regge limit this system should thermalize to a quark-gluon plasma, whether it does so in high energy proton-proton collisions is still open because its not clear that the required parametric separation of timescales is cleanly achieved even at the highest LHC energies.

<sup>&</sup>lt;sup>17</sup>This is a consequence of the geometry of the collision which leads to a "boost-invariant" expansion [85].

$$m_D^2 \sim \frac{\alpha_S n_h}{Q_S} \sim \frac{Q_S^2}{(Q_S t)}.$$
 (33)

The frequency of collisions, per unit volume, is

$$\frac{d\mathcal{N}_{\rm col}}{dt} \sim \sigma n_h N_h \sim \frac{\alpha_S n_h}{m_D^2 p_z t},\tag{34}$$

where  $\sigma \sim \alpha_s^2 m_D^{-2}$  is the scattering cross section and the occupancy  $N_h = \frac{n_h}{Q_s^2 p_s}$ .

Since  $m_D$  rapidly becomes smaller than  $Q_S$ , the net effect of the scatterings is to increase the longitudinal momentum of gluons (which otherwise would be rapidly depleted at the rate  $p_z \sim 1/t$ ) via the random walk relation,

$$p_z^2 \sim \mathcal{N}_{col} m_D^2 \sim \frac{\alpha_S n_h}{p_z},$$
 (35)

and therefore,

$$p_z \sim (\alpha_S n_h)^{\frac{1}{3}} \sim \frac{Q_S}{(Q_S t)^{\frac{1}{3}}}.$$
 (36)

Thus the occupation number of hard gluons diminishes in time as

$$N_h \sim (\alpha_S)^{-1} (Q_S t)^{-\frac{2}{3}}.$$
 (37)

The semiclassical picture of the glasma as a high occupancy state with  $N_h \sim \frac{1}{\alpha_s}$  is no longer tenable when  $N_h \sim 1$ , which occurs when

$$t_{\rm glasma} \sim Q_S^{-1} \alpha_S^{-\frac{3}{2}}.$$
 (38)

This timescale is the "quantum breaking" time of the CGC.

This quantum breaking time in the glasma should however be considered an upper bound, because there is an overlap between the classical and quantum descriptions [86,87] that begins parametrically sooner as long as the occupancy satisfies  $1 \ll N_h \ll 1/\alpha_s$ . As a lower bound, we can ask what the timescale is if rescattering of soft quantum is the only mechanism driving the depletion of the glasma, as in the case of a black hole. This timescale is

$$t_{\text{glasma}}^{\text{low}} \sim N/\Gamma = \frac{N}{(N^2 \alpha_S^2) Q_S} = \frac{1}{Q_S \alpha_S}, \quad (39)$$

which is not surprisingly the same as that in Eq. (30). One thus expects quantum breaking to occur in the glasma in the temporal window  $\alpha_s^{-1} < Q_s t < \alpha_s^{-3/2}$ .

The subsequent evolution and bottom-up thermalization of the glasma into a quark-gluon plasma is described by quantum kinetic  $2 \rightarrow 3$  processes captured by an effective kinetic theory [88]. A parametric estimate [84] of the thermalization time gives  $t_{\text{thermal}} \sim \alpha_s^{-13/5} Q_s^{-1}$ . Since  $\alpha_s$  only grows logarithmically, no power of the coupling can beat the  $Q_s^{-1}$  term, so asymptotically, in the Regge limit, thermalization in heavy-ion collisions is extremely rapid relative to the lifetime of the system  $t_{\text{QGP}} \sim R_A$ , the nuclear radius. This "late-stage" dynamics is of course qualitatively different from that of a black hole.

### C. Connecting the timescales

We shall now construct a dictionary between these timescales in the two systems (CGC/glasma versus BHNP). To do so, we will describe the analogous process in the graviton condensate picture of a black hole. There are two inessential differences. The first is that no decrease in the occupation number analogous to that in Eq. (32) takes place due to the absence of the classical expansion of the system.<sup>18</sup> So for gravitons, quantum rescattering is the only source for reducing their occupancy.

The other difference, of no great import to our dictionary, is that due to the isotropy of the system, all relevant momenta of constituent gravitons are always of order  $R_S^{-1}$ . As we discussed previously, this is also true in the CGC/ glasma for  $t \le 1/Q_S$ , but  $p_z$  depletes rapidly subsequently due to the 1-D expansion. In order to translate the CGC results into analogous results for graviton condensate, we must take

$$p_z \sim Q_S \sim R_S^{-1}. \tag{40}$$

Next, in order to make contact with Eq. (33), we need to evaluate the equivalent mass of the many-body graviton state analogous to the QCD Debye scale. In this context, the origin of this mass is a spontaneous breaking of space-time translation symmetries by the graviton condensate. As a result of this breaking, a graviton propagating through the condensate mixes with the corresponding Goldstone mode and becomes effectively massive. However, unlike the ordinary Higgs effect, the Goldstone is not an external field but comes from the collective excitations of the graviton condensate itself.<sup>19</sup>

The strength of the effect can be estimated from the coupling of the graviton  $h_{\mu\nu}$  with its own energy momentum tensor. It is sufficient to evaluate the latter quantity  $T^{\mu\nu} = \partial^{\mu}h_{\alpha\beta}\partial^{\nu}h^{\alpha\beta} + \dots$ , to bilinear order. The details of the tensorial structure are not important for our purposes and Lorentz indices will be suppressed henceforth. It is sufficient to remember that all the interaction terms contain only

<sup>&</sup>lt;sup>18</sup>Fundamentally, this is because all the scattered partons in the  $2 \rightarrow N$  high energy graviton scattering are "wee" partons; in QCD, a substantial fraction of the momentum is carried by hard "valence" partons.

<sup>&</sup>lt;sup>19</sup>This is analogous to the "polaron" concept in condensed matter physics, with the broken translational symmetry reflected in the localized polaron wave function.

two derivatives. The order parameter of spontaneous breaking of translation symmetry is represented by the expectation value of the derivative of a canonically normalized graviton field over the condensate state  $\langle \partial h \rangle \sim \sqrt{\frac{n}{R_s}} \sim \frac{\sqrt{N}}{R_s^2}$ . As a result, the various components of the graviton field derivatively mix with each other and this generates a Debye screening effect analogous to a thermal plasma,

$$\frac{1}{M_P}h_{\mu\nu}T^{\mu\nu} = \frac{\sqrt{n}}{M_P\sqrt{R_S}}h\partial h + \dots = m_Dh\partial h + \dots$$
(41)

Remembering that  $\alpha_{\rm gr} = (R_S M_P)^{-2}$  and taking the number density at the saturation point of critical packing,  $n = \frac{N}{R_S^3} = \frac{1}{\alpha_{\rm gr} R_S^3}$ , we can write,

$$m_D^2 = \alpha_{\rm gr} n R_S = R_S^{-2} = Q_S^2. \tag{42}$$

We see therefore an exact analogy between the Debye mass and the first relation in both Eqs. (33) and (42), and the equivalence of the second relation in the former to the latter at  $t = 1/Q_s$ , which is the minimum crossing time of two CGC shockwaves.

It is easy to see that the expressions for scattering rates in the two systems also match. Indeed, the scattering cross section of constituent gravitons is

$$\sigma \sim \alpha_{\rm gr}^2 R_S^2 \sim \alpha_{\rm gr}^2 m_D^{-2}, \tag{43}$$

where in the last equation we took Eq. (42) into account. The resulting interaction rate per unit volume is,

$$\frac{d\mathcal{N}_{\rm col}}{dt} \sim \sigma nN \sim \frac{\alpha_{\rm gr} n}{m_D^2},\tag{44}$$

which by Eq. (40) is the same as Eq. (34).

Taking into account critical packing, it is not hard to see that the above expression fully matches the depletion rate given by Eq. (31), which is the source of Hawking evaporation. The corresponding half-depletion time is therefore given by Eq. (28). The reason why the analogous time in Eq. (39) in the glasma is faster by a factor  $\sim \alpha_s^{1/3}$  is due to its one-dimensional expansion, which contributes to its dilution; as we argued previously, this factor should be considered an upper bound.

# D. Lyapunov exponents, quantum breaking, chaos and scrambling

We now turn to parallels between the two systems regarding chaos and information scrambling. The concept of black holes as information scramblers was introduced by Hayden and Preskill [89] who suggested that the thermalization time of perturbed black holes is bounded from below by

$$t_{\rm sc}^{\rm BH} = R_S \ln(S_{\rm BH}),\tag{45}$$

where  $R_s$  is the gravitational radius and  $S_{BH}$  is its entropy. Later, Sekino and Susskind [90] suggested that black holes actually saturate this bound. However, the underlying quantum mechanism behind black hole's fast scrambling remained unclear for some time due to the lack of a suitable microscopic theory of a black hole. The formulation of the black hole quantum *N*-portrait made it possible to establish such a connection. This was done in [74], where the physical meaning of fast scrambling was identified with the *quantum break-time* and chaos. That is, scrambling has been linked with a significant departure of the quantum evolution from the classical one. It was proposed in this work that the Lyapunov exponent  $\lambda$  must play the crucial role in making the system a fast scrambler. The scrambling time can therefore be expressed as [74],

$$t_{\rm sc}^{\rm BH} = \lambda^{-1} \ln(S_{\rm BH}) = \lambda^{-1} \ln(N).$$
 (46)

The very last equality uses the fact that in *N*-portrait the black hole entropy is equal to the number of its constituent gravitons  $S_{BH} = N$  and reflects the fact that the scrambling time can be represented solely in terms of a Lyapunov exponent and the number of constituents of the system, without reference to any of its other characteristics. Some related ideas about Lyapunov exponent were further suggested in [91–93].

As a result, the formula (46) for the scrambling time can be generalized to arbitrary systems beyond black holes. In particular, this equation tells us that any critically packed condensate with occupation number N and a Lyapunov exponent  $\lambda$  is a fast scrambler. In [74] this was explicitly demonstrated on a 1 + 1 dimensional Bose-Einstein condensate that was originally introduced in [4] as a prototype toy model for a black hole portrait. It was discovered that an initial classical state of an uniform condensate would develop a Lyapunov exponent for the regime  $N\alpha > 0$ . Then, it would undergo quantum breaking and chaos and would reach the state of maximal entanglement, after the time given by Eq. (46).

When applying the above knowledge to a black hole, we must keep in mind that there may exist several Lyapunov exponents that can have potentially different physical origins, depending on the situation in which the black hole is placed. For example, in case of a classically-perturbed black hole, a natural candidate(s) for  $\lambda$  is the inverse dissipation time(s) for a quasinormal mode(s). As an extreme form of the disturbance we can think of a black hole merger. Correspondingly, we have to be careful when drawing analogies with the similar timescales in the CGC/ glasma to which we shall turn now. As we shall see, the appearance of a  $\ln(N)$  factor in these timescales is universal there.

When the spacelike CGC fields generate the timelike glasma fields in nuclear collisions, the initial state of the matter after  $t \sim 1/Q_s$ , because of its occupancy of  $O(1/\alpha_s)$ , is that of a coherent classical field. This field develops subsequently to the bottom-up scattering and thermalization scenario through a process of information scrambling and entanglement. Because of its initial 1-D geometry, with matter exploding undergoing a Hubble-like expansion into the vacuum, the  $p_z \sim Q_S$  modes are rapidly redshifted relative to the transverse momentum modes with typical  $p_T \sim Q_S$ . This anisotropy triggers an instability whereby quantum fluctuations, initially  $\alpha_S$  suppressed relative to the coherent glasma condensate, grow nearly exponentially as  $\sim e^{\sqrt{Q_s t}}$  [81,94]. These quantum fluctuations therefore become of the order of the classical field on the timescale  $t_{\rm sc}^{\rm glasma} \sim Q_S^{-1} \ln^2(\alpha_S^{-1})$ —corresponding to a butterfly effect that qualitatively changes the physics. This differs from  $t_{sc}^{BH}$  only by the additional logarithmic factor a consequence of the expanding geometry of the former.<sup>20</sup>

In the framework of classical statistical field theory for nonequilibrium fields [98,99], these exponentially growing quantum fluctuations can be resummed to all orders in  $\alpha_{S}e^{\sqrt{Q_{S}t}}$  and expressed as a initial spectrum of fluctuations for the evolution of the classical fields [100,101]. Each time evolution of a classical field configuration  $A_{\rm cl}(t) \sim 1/\sqrt{\alpha_s}$ is seeded with a small quantum fluctuation  $a \sim O(1)$  at the initial time  $t = 1/Q_S$  drawn from this initial spectrum of fluctuations. Averaging observables over the spectrum of fluctuations leads to the randomization of the information contained in the glasma. This is seen very cleanly in a toy 0 + 1-dimensional  $\phi^4$  model where the field equations can be solved analytically and correspond to Jacobi elliptic functions; in this simple setup, the difference between two seeds can be mapped on to a small initial phase difference in classical trajectories in the Poincaré phase plane of the phase space density. With time evolution, the phase difference between the trajectories grows rapidly, leading to decoherence of the trajectories and the filling up the Poincaré plane [100].

This phenomenon, dubbed "prethermalization" [102], causes such high occupancy fields to flow to a nonthermal fixed point described by the quasistationary self-similar behavior of single particle distributions characteristic of weak wave turbulence [103]; for the 3 + 1 - D glasma, numerical simulations [82,83] show that the corresponding single particle distributions are precisely those that describe the early-stage of bottom-up thermalization we discussed earlier. Remarkably, this nonthermal fixed point appears universal, and is seen to be identical to that of overoccupied self-interacting scalar fields [104]. The subsequent scattering

generates further entanglement, whose entanglement entropy grows until the system thermalized completely.

# **VI. OUTLOOK**

In this work, we outlined some of the elements of a remarkable correspondence between the saturated wee gravitons that constitute the black hole quantum N-portrait [3,4] and saturated wee gluons that constitute a color glass condensate state [5] inside hadrons that can be probed at very high energies. Both systems can be understood as classical states generated in  $2 \rightarrow N$  scattering when  $N = 1/\alpha$ , which also corresponds to the unitarization boundary for the scattering. A key feature of the correspondence is an emergent saturation scale  $Q_S$  that characterizes the universal dynamics of the two systems on the unitarization boundary. This feature was missed by previous studies in the literature that attempted to draw parallels between wee parton dynamics in the two theories.

The universal feature shared between the two systems is the saturation of the unitarity bounds Eqs. (2) and (3) on microstate entropy and the attainment of the relation Eq. (11) for a generic saturated system [11], [9,10]. In particular, this implies that both objects exhibit the area form of the entropy in terms of respective Goldstone constants f of spontaneously broken Poincaré invariance. In the case of the BHNP, where the Goldstone boson comes from graviton with the decay constant  $M_P$ , Eq. (3) reproduces the Bekenstein-Hawking entropy. Analogously, the area law of entropy in the CGC is determined by the Goldstone scale that controls the gluon screening length.

The minimal timescales required for retrieving quantum information from the two systems saturate the general bounds in Eqs. (26) and (27) derived in [11]. As already noticed there for a black hole, both expressions reproduce Page time [73]. Further, the decay of these semiclassical states through scrambling and scattering also is strongly analogous (modulo differences due to the different geometries of the two systems) until a quantum breaking time, beyond which time the two systems differ significantly.

We showed that the observed universality provides deep insight into the complementary languages used to describe the two systems and their mapping into each other. Specifically, as a concrete consequence of this universality, we conjectured that the entropy of the CGC saturates the unitarity bounds Eqs. (2) and (3). It further allowed us to interpret the onset of screening and shadowing in this system in terms of a Goldstone scale for the breaking of Poincaré invariance; this understanding was crucial in establishing the nontrivial area law in terms of the Goldstone scale f, for the CGC entropy in a given rapidity interval.

From the CGC side of the CGC-BHNP correspondence, this universality may be useful in developing a quantitative understanding of unitarization in the BHNP, employing key features of a double copy between QCD and gravity

 $<sup>^{20}</sup>$ In a fixed box geometry, these instabilities indeed grow exponentially with t [95–97] as opposed to  $\sqrt{t}$  in the expanding case.

amplitudes. In particular, a classical double copy between the two theories, may allow one to demonstrate, employing the dynamics of saturation in QCD, the onset of classicalization and unitarization in gravity amplitudes. This work is in progress and will be reported separately.

An interesting question is whether our work sheds light on a suggestion in [105, 106] that there is an equality between the critical exponents, as observed by Choptuik [107], appearing in the self-similar collapse of classical fluids in D > 4 dimensions and the exponent governing the growth of the BFKL equation in D = 4 dimensions. The principal point of our work however is that the CGC and black holes in BHNP are structurally similar states at the full quantum level in D = 4 dimensions. This follows from the universal behavior of these objects saturating the entropy bounds in Eqs. (2) and (3). The saturation of the entropy bounds implies the unitarization of the corresponding  $2 \rightarrow N$  processes (gluon and graviton) in weak coupling in both QCD and gravity in D = 4 dimensions. It would be interesting to explore whether it is at all possible that the saturation of the entropy and unitarity bounds translates into values of classical critical exponents.

Another striking consequence of the BHNP-CGC correspondence that deserves to be explored further is is the relation between gravitational memory and color memory in this picture, and their connection to soft theorems governing the infrared structure of scattering amplitudes. Strominger has argued that an infrared "triangle" relates the asymptotic BMS symmetries of gravity to the soft theorems and the gravitational memory effect, and conjectured that this infrared triangle is a universal feature of gauge theories [108]. The connection of this picture to the BHNP was discussed in [109]. As argued there, the Goldstone modes on the black hole event horizon can be identified with those of spontaneously broken *horizon* supertranslations.

The mapping of gravitational memory to color memory was first discussed in [110] in the context of Yang-Mills fields on the celestial sphere at null infinity. It was shown in [111] that the classical equations describing the transition between degenerate vacua in the infinite momentum frame are precisely the equations describing the CGC classical field. Further, just as gravitational memory can be measured as a physical displacement of inertial gravitational wave detectors, the color memory of the CGC state is measurable due to the large transverse momentum kick  $p_{\perp} \sim Q_S$  delivered to a colored probe such as a quarkantiquark pair that experiences the passage of a CGC shock wave. The relation of color memory in the CGC to the color memory of gauge fields at null infinity has been argued on the basis of the CGC fields being static in lightcone time. However the mapping of the CGC to the BHNP suggests that there is a further interpretation of information carrying CGC modes as Goldstone modes of spontaneously broken Poincaré symmetry. This feature of the BHNP-CGC correspondence deserves further investigation.

## ACKNOWLEDGMENTS

This work of G. D. is supported in part by the Humboldt Foundation under Humboldt Professorship Award, by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy-No. EXC-2111-390814868, and Germany's Excellence Strategy under Excellence Cluster Origins. R. V. is supported by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics, under Contract No. DE- SC0012704. His work was also supported by the Humboldt Foundation through a Humboldt Prize and by the DFG Collaborative Research Centre SFB 1225 (ISOQUANT) at Heidelberg University. R. V. would like to thank Alan Caldwell, the Max Planck Institute and Ludwig-Maximilians University in Munich for their kind hospitality. He would also like to thank Himanshu Raj, Vladi Skokov and Kong Tu for valuable discussions.

## APPENDIX: GOLDSTONE MODES AND THE AREA LAW ENTROPY BOUND

We will briefly review here the arguments of [11] about the emergence of Goldstone modes and their role in the area-law form Eq. (3) of the entropy bound Eq. (2) for a generic critically packed system Eq. (1), thereby supporting the relation Eq. (11).

Consider a bosonic field  $\phi_j$  described by creation/ annihilation operators  $\hat{a}_i(\mathbf{k})^{\dagger}$ ,  $\hat{a}_i(\mathbf{k})$ ,

$$\hat{\phi}_j = \sum_{\boldsymbol{k}} \frac{1}{\sqrt{2V\omega_{\boldsymbol{k}}}} (\mathrm{e}^{i\boldsymbol{k}\boldsymbol{x}} \hat{a}_j(\boldsymbol{k}) + \mathrm{e}^{-i\boldsymbol{k}\boldsymbol{x}} \hat{a}_j(\boldsymbol{k})^{\dagger}). \quad (A1)$$

Here *V* is the volume and the label *k* refers to the threemomentum. The label j = 1, ..., n accounts for different spin-polarizations and for a representation content with respect to internal symmetries such as color or flavor. The modes obey the standard bosonic commutation relations,  $[\hat{a}_i(\mathbf{k}), \hat{a}_j(\mathbf{k}')^{\dagger}] = \delta_{ij}\delta_{kk'}, [\hat{a}_i(\mathbf{k}), \hat{a}_j(\mathbf{k}')] = 0$ . We shall assume that the field interacts via a coupling  $\alpha$ .

We shall use states in which modes of certain momenta k are macroscopically occupied,

$$|N\rangle = \prod_{j=1}^{n} \frac{(\hat{a}_j(\boldsymbol{k})^{\dagger})^{N_j}}{\sqrt{N_j!}} |0\rangle, \qquad (A2)$$

where N refers to the total occupation number,

$$N = \sum_{j=1}^{n} N_j. \tag{A3}$$

Analogously, we can form a coherent state that describes a classical field-configuration localized within a certain characteristic radius R,

$$|\mathrm{sol}\rangle = \mathrm{e}^{\sum_{\mathbf{k}}\sum_{j=1}^{n}\sqrt{N_{j}(\mathbf{k})}(\hat{a}_{j}(\mathbf{k})^{\dagger} - \hat{a}(\mathbf{k})_{j})}|0\rangle, \qquad (\mathrm{A4})$$

with

$$\sum_{j=1}^{n} \sum_{\boldsymbol{k}} N_j(\boldsymbol{k}) = N \gg 1, \qquad (A5)$$

where the  $N_j(\mathbf{k})$ 's are sharply peaked around the momentum  $|\mathbf{k}| = \frac{1}{R} = Q_s$ .

The classical field is described by the expectation value over this coherent state,

$$\phi_i = \langle \text{sol} | \hat{\phi}_i | \text{sol} \rangle. \tag{A6}$$

We shall refer to such a state as a soliton state.

We shall assume that the occupation number N does not exceed  $1/\alpha$ . In this case, the kinetic energy of a given quantum,  $E_{\rm kin} \sim \frac{1}{R}$ , is not subdominant to the potential energy of its interaction with the rest, given by  $E_{\rm pot} \sim \frac{aN}{R}$ . Under such conditions, the localized classical field configuration  $\phi_j$  represents a N-particle state of characteristic momenta  $\sim Q_S = 1/R_S$ , each contributing  $\sim Q_S$  toward the energy of the soliton. Therefore the total energy is,

$$E \sim \frac{N}{R_S} \sim NQ_S. \tag{A7}$$

Now according to [11], unitarity puts an upper bound on entropy given by Eq. (2). This also agrees with the argument in [9] that the entropy of a self-sustained *N*-particle state with  $N \leq 1/\alpha$  is bounded by *N*.

Let us now show that taking into account Eq. (1), the entropy bound in Eq. (2) is equal to the area in units of a Goldstone decay constant f. For this, we must determine f. The localized classical field configuration  $\phi_j$  always spontaneously breaks Poincaré symmetry (and internal symmetries), with the strength of the breaking of Poincaré invariance measured by the gradient of the field  $\sum_j (\nabla \phi_j)^2 \sim \frac{N}{R^4}$ . The corresponding Goldstone fields are collective modes emerging from the infinitesimal coordinate-dependent transformations. The scale f is the coefficient accompanying the dimensionless parameter of such transformations.

For example, consider translations along a coordinate x. For simplicity, in this case, we can drop the internal index j. The relevant infinitesimal variation of the solution is,

$$\delta\phi = \epsilon_x R \partial_x \phi = \epsilon_x f \tag{A8}$$

where  $\epsilon_x$  is dimensionless. Its coefficient  $f = R\partial_x \phi_j$ represents the decay constant of the Goldstone field. It is clear that

$$f = R\partial_x \phi = \frac{\sqrt{N}}{R}.$$
 (A9)

Now taking into account Eq. (1), it is obvious that we can rewrite Eq. (2), in form of the area law Eq. (3),

$$S_{\max} = \frac{1}{\alpha} = N = (Rf)^2, \qquad (A10)$$

thereby reproducing Eq. (11) in its entirety.

Due to the universality of the Goldstone language and unitarity, the above relations emerge as a generic property of the saturated systems, regardless of their underlying structure, as has been confirmed for several systems [9-11].

Now, for both the BHNP and the CGC, the Goldstone decay constant is given by Eq. (A9). For the BHNP, this scale is equal to the Planck mass  $f = \sqrt{N}Q_S = \sqrt{N}/R_S = M_P$ . It is easy to check that in terms of parameters N, f and  $\alpha$ , the effective 3-graviton vertex in BHNP and 3-gluon vertex in CGC take similar forms. We remind the reader that a generic three-graviton vertex is weighted by two derivatives ( $\partial^2$ ) and a single power of  $1/M_P$ . For the saturated state of BHNP, these translate as  $\partial^2 \rightarrow Q_S^2$  and  $1/M_P = 1/f = R_S/\sqrt{N}$  respectively. Thus we have,

three graviton effective vertex

$$= 1/(R_S^2 M_P) = \sqrt{\alpha_{\rm gr}} Q_S = Q_S^2/f.$$
 (A11)

This matches the expression for the strength of the 3-gluon vertex for CGC fields,

three gluon effective vertex = 
$$\sqrt{\alpha_s}Q_s = Q_s/\sqrt{N} = Q_s^2/f$$
.  
(A12)

Thus in units of the Goldstone scale, the strength of the coupling in the two saturated systems is identical.

- [2] Z. Bern, J. Carrasco, and H. Johansson, Phys. Rev. D 78, 085011 (2008).
- [3] G. Dvali and C. Gomez, Fortschr. Phys. 61, 742 (2013).
- [4] G. Dvali and C. Gomez, Eur. Phys. J. C 74, 2752 (2014).
- [5] F. Gelis, E. Iancu, J. Jalilian-Marian, and R. Venugopalan, Annu. Rev. Nucl. Part. Sci. 60, 463 (2010).
- [6] T. Lappi and L. McLerran, Nucl. Phys. A772, 200 (2006).
- [7] F. Gelis and R. Venugopalan, Acta Phys. Pol. B **37**, 3253 (2006).
- [8] D. H. Rischke, Prog. Part. Nucl. Phys. 52, 197 (2004).
- [9] G. Dvali, Fortschr. Phys. 69, 2000090 (2021).
- [10] G. Dvali, Fortschr. Phys. 69, 2000091 (2021).
- [11] G. Dvali, J. High Energy Phys. 03 (2021) 126.
- [12] J. D. Bekenstein, Phys. Rev. D 7, 2333 (1973).
- [13] S. W. Hawking, Phys. Rev. D 13, 191 (1976).
- [14] G. 't Hooft, Nucl. Phys. B256, 727 (1985).
- [15] G. 't Hooft, Int. J. Mod. Phys. A 11, 4623 (1996).
- [16] D. J. Gross and P. F. Mende, Nucl. Phys. B303, 407 (1988).
- [17] D. J. Gross and P. F. Mende, Phys. Lett. B 197, 129 (1987).
- [18] D. Amati, M. Ciafaloni, and G. Veneziano, Int. J. Mod. Phys. A 03, 1615 (1988).
- [19] G. Dvali, C. Gomez, R. Isermann, D. Lüst, and S. Stieberger, Nucl. Phys. B893, 187 (2015).
- [20] L. N. Lipatov, Nucl. Phys. B365, 614 (1991).
- [21] L. N. Lipatov, Phys. Lett. **116B**, 411 (1982).
- [22] A. Addazi, M. Bianchi, and G. Veneziano, J. High Energy Phys. 02 (2017) 111.
- [23] M. Ciafaloni, D. Colferai, F. Coradeschi, and G. Veneziano, Phys. Rev. D 93, 044052 (2016).
- [24] E. Kuraev, L. Lipatov, and V. S. Fadin, Sov. Phys. JETP 45, 199 (1977).
- [25] I. Balitsky and L. Lipatov, Sov. J. Nucl. Phys. 28, 822 (1978).
- [26] L. Gribov, E. Levin, and M. Ryskin, Phys. Rep. 100, 1 (1983).
- [27] A. H. Mueller and J.-w. Qiu, Nucl. Phys. **B268**, 427 (1986).
- [28] V. A. Abramovsky, V. N. Gribov, and O. V. Kancheli, Yad. Fiz. 18, 595 (1973).
- [29] F. Gelis and R. Venugopalan, Nucl. Phys. A776, 135 (2006).
- [30] F. Gelis and R. Venugopalan, Nucl. Phys. A782, 297 (2007).
- [31] L. D. McLerran and R. Venugopalan, Phys. Rev. D 49, 2233 (1994).
- [32] L. D. McLerran and R. Venugopalan, Phys. Rev. D 49, 3352 (1994).
- [33] H. L. Verlinde and E. P. Verlinde, Nucl. Phys. B371, 246 (1992).
- [34] H. L. Verlinde and E. P. Verlinde, arXiv:hep-th/9302104.
- [35] L. Susskind, J. Math. Phys. (N.Y.) 36, 6377 (1995).
- [36] J. Jalilian-Marian, A. Kovner, L. D. McLerran, and H. Weigert, Phys. Rev. D 55, 5414 (1997).
- [37] J. Jalilian-Marian, A. Kovner, A. Leonidov, and H. Weigert, Phys. Rev. D 59, 014014 (1998).
- [38] J. Jalilian-Marian, A. Kovner, and H. Weigert, Phys. Rev. D 59, 014015 (1998).

- [39] E. Iancu, A. Leonidov, and L. D. McLerran, Nucl. Phys. A692, 583 (2001).
- [40] E. Ferreiro, E. Iancu, A. Leonidov, and L. McLerran, Nucl. Phys. A703, 489 (2002).
- [41] I. Balitsky, Nucl. Phys. B463, 99 (1996).
- [42] Y. V. Kovchegov, Phys. Rev. D 60, 034008 (1999).
- [43] G. Dvali, G.F. Giudice, C. Gomez, and A. Kehagias, J. High Energy Phys. 08 (2011) 108.
- [44] G. Dvali and C. Gomez, arXiv:1005.3497.
- [45] G. 't Hooft, Conf. Proc. C 930308, 284 (1993).
- [46] R. Monteiro, D. O'Connell, and C. D. White, J. High Energy Phys. 12 (2014) 056.
- [47] W. D. Goldberger and A. K. Ridgway, Phys. Rev. D 95, 125010 (2017).
- [48] J. D. Bekenstein, Phys. Rev. D 23, 287 (1981).
- [49] G. Dvali, C. Gomez, and A. Kehagias, J. High Energy Phys. 11 (2011) 070.
- [50] S. R. Coleman and F. De Luccia, Phys. Rev. D 21, 3305 (1980).
- [51] A. Ayala, J. Jalilian-Marian, L. D. McLerran, and R. Venugopalan, Phys. Rev. D 53, 458 (1996).
- [52] F. Gelis, T. Lappi, and R. Venugopalan, Phys. Rev. D 78, 054019 (2008).
- [53] O. V. Kancheli, arXiv:2003.04654.
- [54] S. Munier and R. B. Peschanski, Phys. Rev. Lett. 91, 232001 (2003).
- [55] S. Munier and R. B. Peschanski, Phys. Rev. D 69, 034008 (2004).
- [56] A. M. Stasto, K. J. Golec-Biernat, and J. Kwiecinski, Phys. Rev. Lett. 86, 596 (2001).
- [57] E. Iancu, K. Itakura, and L. McLerran, Nucl. Phys. A708, 327 (2002).
- [58] J. Berges, S. Floerchinger, and R. Venugopalan, Phys. Lett. B 778, 442 (2018).
- [59] J. Berges, S. Floerchinger, and R. Venugopalan, J. High Energy Phys. 04 (2018) 145.
- [60] A. Kovner and M. Lublinsky, Phys. Rev. D 92, 034016 (2015).
- [61] H. Duan, C. Akkaya, A. Kovner, and V. V. Skokov, Phys. Rev. D 101, 036017 (2020).
- [62] N. Armesto, F. Dominguez, A. Kovner, M. Lublinsky, and V. Skokov, J. High Energy Phys. 05 (2019) 025.
- [63] K. Kutak, Phys. Lett. B 705, 217 (2011).
- [64] B. Muller and A. Schafer, Int. J. Mod. Phys. E 20, 2235 (2011).
- [65] R. Peschanski, Phys. Rev. D 87, 034042 (2013).
- [66] L. McLerran and B. Schenke, Nucl. Phys. A929, 71 (2014).
- [67] D. E. Kharzeev and E. M. Levin, Phys. Rev. D 95, 114008 (2017).
- [68] Y. Hagiwara, Y. Hatta, B.-W. Xiao, and F. Yuan, Phys. Rev. D 97, 094029 (2018).
- [69] A. H. Mueller, Nucl. Phys. **B572**, 227 (2000).
- [70] A. Krasnitz and R. Venugopalan, Phys. Rev. Lett. 86, 1717 (2001).
- [71] A. H. Mueller, Nucl. Phys. A724, 223 (2003).
- [72] G. Dvali, A. Franca, C. Gomez, and N. Wintergerst, Phys. Rev. D 92, 125002 (2015).
- [73] D. N. Page, Phys. Rev. D 13, 198 (1976).

- [74] G. Dvali, D. Flassig, C. Gomez, A. Pritzel, and N. Wintergerst, Phys. Rev. D 88, 124041 (2013).
- [75] S. W. Hawking, Commun. Math. Phys. 43, 199 (1975); 46, 206(E) (1976).
- [76] G. Dvali, C. Gomez, and S. Zell, J. Cosmol. Astropart. Phys. 06 (2017) 028.
- [77] M. Srednicki, Phys. Rev. E 50, 888 (1994).
- [78] C. Murthy and M. Srednicki, Phys. Rev. Lett. 123, 230606 (2019).
- [79] J. Berges, M. P. Heller, A. Mazeliauskas, and R. Venugopalan, Rev. Mod. Phys. 93, 035003 (2021).
- [80] A. Krasnitz and R. Venugopalan, Nucl. Phys. B557, 237 (1999).
- [81] P. Romatschke and R. Venugopalan, Phys. Rev. Lett. 96, 062302 (2006).
- [82] J. Berges, K. Boguslavski, S. Schlichting, and R. Venugopalan, Phys. Rev. D 89, 074011 (2014).
- [83] J. Berges, K. Boguslavski, S. Schlichting, and R. Venugopalan, Phys. Rev. D 89, 114007 (2014).
- [84] R. Baier, A. H. Mueller, D. Schiff, and D. Son, Phys. Lett. B 502, 51 (2001).
- [85] J. D. Bjorken, Phys. Rev. D 27, 140 (1983).
- [86] A. H. Mueller and D. T. Son, Phys. Lett. B 582, 279 (2004).
- [87] S. Jeon, Phys. Rev. C 72, 014907 (2005).
- [88] P. B. Arnold, G. D. Moore, and L. G. Yaffe, J. High Energy Phys. 11 (2000) 001.
- [89] P. Hayden and J. Preskill, J. High Energy Phys. 09 (2007) 120.
- [90] Y. Sekino and L. Susskind, J. High Energy Phys. 10 (2008) 065.
- [91] J. Maldacena, S. H. Shenker, and D. Stanford, J. High Energy Phys. 08 (2016) 106.
- [92] S. H. Shenker and D. Stanford, J. High Energy Phys. 05 (2015) 132.

- [93] D. Stanford, J. High Energy Phys. 10 (2016) 009.
- [94] P. Romatschke and R. Venugopalan, Phys. Rev. D 74, 045011 (2006).
- [95] S. Mrowczynski, Phys. Lett. B 314, 118 (1993).
- [96] P.B. Arnold, J. Lenaghan, and G.D. Moore, J. High Energy Phys. 08 (2003) 002.
- [97] J. Berges, S. Scheffler, and D. Sexty, Phys. Rev. D 77, 034504 (2008).
- [98] J. Berges, AIP Conf. Proc. 739, 3 (2004).
- [99] G. Aarts and J. Berges, Phys. Rev. Lett. **88**, 041603 (2002).
- [100] K. Dusling, F. Gelis, and R. Venugopalan, Nucl. Phys. A872, 161 (2011).
- [101] T. Epelbaum and F. Gelis, Phys. Rev. D 88, 085015 (2013).
- [102] J. Berges, S. Borsanyi, and C. Wetterich, Phys. Rev. Lett. 93, 142002 (2004).
- [103] V. Zakharov, V. L'vov, and G. Falkovich, *Kolmogorov Spectra of Turbulence I: Wave Turbulence*, Springer Series in Nonlinear Dynamics (Springer, New York, 2012).
- [104] J. Berges, K. Boguslavski, S. Schlichting, and R. Venugopalan, Phys. Rev. Lett. 114, 061601 (2015).
- [105] L. Alvarez-Gaume, C. Gomez, and M. A. Vazquez-Mozo, Phys. Lett. B 649, 478 (2007).
- [106] L. Alvarez-Gaume, C. Gomez, A. Sabio Vera, A. Tavanfar, and M. A. Vazquez-Mozo, Nucl. Phys. B806, 327 (2009).
- [107] M. W. Choptuik, Phys. Rev. Lett. 70, 9 (1993).
- [108] A. Strominger, arXiv:1703.05448.
- [109] A. Averin, G. Dvali, C. Gomez, and D. Lust, J. High Energy Phys. 06 (2016) 088.
- [110] M. Pate, A.-M. Raclariu, and A. Strominger, Phys. Rev. Lett. 119, 261602 (2017).
- [111] A. Ball, M. Pate, A.-M. Raclariu, A. Strominger, and R. Venugopalan, Ann. Phys. (Amsterdam) 407, 15 (2019).