Limiting fragmentation as an initial-state probe in heavy ion collisions

Kayman J. Gonçalves,[1] Andre V. Giannini,[2,3] David D. Chinellato,[1] and Giorgio Torrieri[1]
[1]*IFGW, Unicamp, Campinas, Brazil*
[2]*Akita International University, Yuwa, Akita-city, Japan*
[3]*IFUSP, USP, São Paulo, Brazil*

(Received 5 June 2019; revised manuscript received 24 August 2019; published 4 November 2019)

We discuss limiting fragmentation within a few currently popular phenomenological models. We show that popular Glauber-inspired models of particle production in heavy ion collisions, such as the two-component model, generally fail to reproduce limiting fragmentation when all energies and system sizes experimentally available are considered. This is due to the energy dependence of number of participants and number of collisions. We quantify this violation in terms of the model parameters. We also make the same calculation within a color glass condensate scenario and show that the dependence of the saturation scale on the number of participants generally leads to violation of limiting fragmentation. We further argue that wounded parton models, provided the nucleon size and parton density vary predominantly with Bjorken x, could in principle reproduce both multiplicity dependence with energy and limiting fragmentation. We suggest, therefore, that an experimental measurement of deviation from limiting fragmentation in heavy ion collisions, for different system sizes and including the experimentally available range of energies, is a powerful test of initial-state models.

DOI: 10.1103/PhysRevC.100.054901

I. INTRODUCTION

The phenomenon of limiting fragmentation in hadronic collisions was originally both experimentally seen [1,2] and theoretically explained at the origins of the QCD theory of strong interactions [3].

The definition of limiting fragmentation is that

$$\left.\frac{d^2N}{d\kappa^2}\right|_{\kappa=y-y_0} = C, \quad \frac{\sqrt{s}}{C}\frac{dC}{d\sqrt{s}} \ll 1, \quad y_0 = \ln\frac{\sqrt{s}}{m_p}, \tag{1}$$

where $C \simeq -0.65$ is an energy-independent constant and m_p the proton mass. A qualitative illustration of limiting fragmentation is shown in panel (a) of Fig. 1. As it happens well away from midrapidity, distributions in both rapidity y and pseudorapidity η can be used for limiting fragmentation. (In fact the experimental data discussed later were collected in pseudorapidity. See the Appendix for a discussion of this issue.)

When plotted against $y - y_0$ (the difference between rapidity and beam rapidity), multiplicity distributions away from midrapidity should fall on a universal curve, independent of center-of-mass energy. Using the second derivative of the rapidity distribution, as defined in Eq. (1), provides a dimensionless quantitative estimator which should vanish within the error bar if limiting fragmentation holds, be $\gg 1$ if it breaks down, and is relatively insensitive to rapidity/pseudorapidity differences (see Appendix). Using directly C of Eq. (1) allows for a nice graphic comparison, while the logarithmic derivative $\frac{dC}{d\ln\sqrt{s}}$ makes large fluctuations blow up, but allows a quantitative comparison across different energy regimes.

The basic explanation for limiting fragmentation is to see it as a consequence of the parton model and Bjorken scaling [4]. The distribution functions of parton g, i.e., the parton amplitudes in the infinite momentum frame of the nucleon, depend on the nucleon momentum in a very particular way:

$$|\langle N(p)|g(q)\rangle|^2_{p\gg m_N} \sim f(x, \ln Q^2), \quad x = q/p, \tag{2}$$

where x is the momentum fraction and Q^2 the renormalization group momentum scale, corresponding to the momentum transfer of the process measuring $f(x, Q^2)$. As the parton mass is negligible, its momentum rapidity y_g is related to x very simply, as

$$y_g = \pm\ln(1/x). \tag{3}$$

Because of asymptotic freedom, the dependence on Q^2 as well as any hadronization effects do not change the final momentum much. As a result, away from midrapidity fragmentation does not change the rapidity of the hadron with respect to the parton. Thus, if one assumes parton interactions are local in rapidity, and the resulting hadron is not shifted in rapidity with respect to the parton, in other words (the delta function is simply the momentum conservation in the longitudinal axis)

$$\frac{dN}{dy} \sim \int dx_A dx_B f(x_A) f(x_B) \mathcal{N}(x_A\sqrt{s}/2, x_B\sqrt{s}/2) \times \delta(\sinh\ln(1/x_A) + \sinh\ln(1/x_B) - \sinh y), \tag{4}$$

then limiting fragmentation follows naturally, since asymptotic freedom and the longitudinal momentum conservation ensure that when $x_A \gg x_B$, corresponding to y very different from zero, a universal curve for dN/dy depending only on $x_A \simeq e^y$ emerges.

This reasoning is appropriate if the only dimensionful scale relevant to hadronic scattering is the nucleon size, i.e., if only

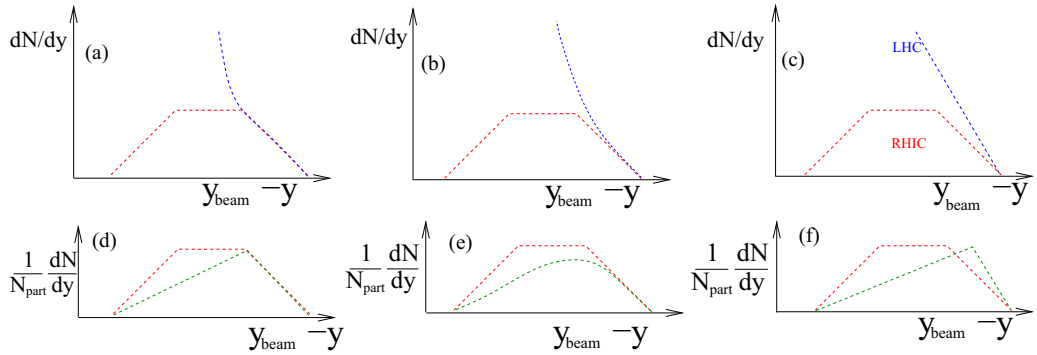

FIG. 1. Possible limiting fragmentation scenarios when LHC and lower energies are compared (top row) and an extended comparison with different system sizes (bottom row). Case (a),(d) presupposed an extended limiting fragmentation scenario up to the midrapidity plateau. Case (b),(e) is a smooth breaking of scaling interpolating between a universal fragmentation regime and the central rapidity plateau and case (c),(f) is wholesale violation, both at central rapidity and in the fragmentation region. The bottom row panels (d),(e),(f) show the three cases when a symmetric *AA* collision is compared to an asymmetric *pA* or *dA* collision.

one type of collision, in a not-too varied range of energies, is considered. In the last two decades, however, it has been seen in a wide range of hadronic [5–8] and even nuclear [9] collisions. Indeed, a variety of approaches—ranging from Landau hydrodynamics [10] to the color glass condensate [11–15] to phenomenological transport models such as AMPT [16] (whose limiting fragmentation is investigated in [17]), string percolation [18–20], and relativistic diffusion [21]—are at least roughly compatible with it.

However, limiting fragmentation turned out to work for a wider spectrum of energies, rapidities and system sizes than initially suspected. At Relativistic Heavy Ion Collider (RHIC) energies (5–200 GeV), it seems to hold [5,9,11] for all system sizes and all energies up to an energy-dependent rapidity interval of $O(1)$ or so. This can be surprising. Close to midrapidity, at the highest energy density, a lot of additional effects, from high energy fragmenting jets to soft gluon effects to viscous entropy production, are expected to contribute particles in a way that evolves nontrivially with energy. The reason why energy-specific dynamics should converge to a universal distribution away from midrapidity is not immediately clear, since Bjorken scaling means remnants of these processes should be present away from midrapidity.

At RHIC energies, there was no problem to incorporate this extended limiting fragmentation into phenomenological models [18], if not theories. A qualitative explanation links [22,23] limiting fragmentation to the close-to-logarithmic energy dependence of multiplicity dN/dy per participant $N_{\rm part}$ at midrapidity. The logarithmic dependence of the rapidity width (the base of the distribution) is fixed by kinematics. Then, "extra parton sources" at midrapidity [18,22] produce strings connected to rapidity edges, ensuring the self-similarity of the total multiplicity distribution.

The Large Hadron Collider (LHC) seems to have added some new twists to this story when energies of several TeV were reached. It has conclusively deviated from logarithmic scaling of multiplicity with respect to energy (see for example [24]), at the same time convincingly breaking the pure number-of-participants scaling [25,26] of particle production. This has motivated the development of more complicated characterizations of the Glauber initial state, generally based on "multi-component scenarios." For instance, core-corona models posit that the event is characterized both by a "soft medium" (participants that underwent multiple collisions) as well as a few hard collisions (a "corona" of single-hit participants). More phenomenologically [24,27,28] one could assume wounded nucleons and collisions provide different admixtures to the multiplicity.

In view of these new developments, it is worth revisiting limiting fragmentation at this energy. It should first be noted that it has *not* as yet been confirmed or falsified experimentally [29–31], partially because LHC experiments have not as yet explored rapidity regimes high enough to be compared even with RHIC energies. In principle, the three possible scenarios are described in the panels of Fig. 1. One can continue to have perfect limiting fragmentation in the whole region away from midrapidity [panel (a)], limiting fragmentation can smoothly be achieved as rapidity increases [panel (b)], or it can just totally break down [panel (c)]. The bottom panels (d,e,f) show the corresponding possible scenarios in *pA* collisions once the distribution is normalized by the number of participants (Note that only one side can exhibit limiting fragmentation since *pA* is rapidity asymmetric).

In this work, we argue that such an investigation is indeed necessary. We show that the most-commonly used model for the initial state of heavy ion collisions, the Glauber model, will generally break limiting fragmentation in the parameter space where it fits LHC data *provided all energies and system sizes are considered*. We show that the same is true for color glass condensate (CGC) models, as implemented in [14,15].

We then argue that a currently popular model capable of bringing the applicability of limiting fragmentation to LHC energies and varying system sizes, *provided it is indeed confirmed experimentally*, is a "wounded parton model," with Glauber wounded partons [32–36] smeared in rapidity space according to Bjorken scaling. We conclude with some experimental considerations regarding this observable.

II. LIMITING FRAGMENTATION AND THE GLAUBER MODEL

The two-component model is a physically intuitive parametrization of multiplicity [25,26]. The basic idea is that in a nucleus-nucleus collision some particles are produced in hard scatterings as a result of fragmentation of parton-parton interactions. The rest of the particles, typically the majority, arise from "wounded nucleons": nuclei that underwent a collision, emitted color charge, and then emitted particles due to nonperturbative color neutralization.[1] An alternative but related picture is the "core-corona" model, where wounded nucleons form "the medium" (presumably a quark-gluon plasma), while individual nucleon-nucleon collisions (and nuclei which underwent just one collision) can be thought of as being similar to proton-proton collisions [28].

Thus, quantitatively, the two sources of particles are related by the parameter f, controlling the input of the number of collisions $N_{\rm coll}$ vs number of participants $N_{\rm part}$ to the total number of "ancestor particle sources" N_{anc}:

$$N_{anc} = fN_{\rm part} + (1-f)N_{\rm coll}. \tag{5}$$

Physically, $N_{\rm coll}$ scaling reflects the fragmentation of partons produced in a QCD collision. $N_{\rm part}$ reflects the products "wounded nucleons" which emit partons as they color neutralize.

To describe multiplicity we need another parameter α, controlling how many particles are produced per ancestor, interpolating between logarithmic ($\alpha = 0$) and power-law ($\alpha > 0$) limits:

$$\left.\frac{dN}{dy}\right|_{y=0} = \mathcal{N}_0 N_{anc} \int_1^{\sqrt{s}} \zeta^{\alpha-1} d\zeta, \tag{6}$$

where $\mathcal{N}_0$ is an overall normalization parameter, uncorrelated with f, α.

In [27] and preceding literature, it was assumed that α is the same for pp, pA, and AA collisions. This means one knows α from pp fits and can obtain f from AA fits. In other words, The different exponents seen experimentally in these systems are due entirely from the fluctuation of $N_{\rm part}$ and $N_{\rm coll}$ of a two-component model.

This, while possible, is not guaranteed. It is well known (the "EMC effect") from eA scattering that parton distribution functions significantly change in a nuclear medium [37]. Given that different energy regimes probe different x, it is plausible that the difference between proton and nuclear medium could give rise of different excitation functions of dN/dy on top of geometric effects. This, however, means that the f parameter and the α parameter need to be understood together within a fit to experimental data, since the Hessian matrix relating them will contain a diagonal term.

A phenomenology tool comes from the fact that the cross-sectional area depends on energy, with its dependence roughly

$$\sigma = A_1 \ln\left(\frac{\sqrt{s}}{m_p}\right) + A_2\left[\ln\left(\frac{\sqrt{s}}{m_p}\right)\right]^2 \tag{7}$$

with $A_1 = 25$ and $A_2 = 0.146$ fitted from data such as [38].[2]

It then becomes clear that in a general Glauber model the ratio of $N_{\rm coll}/N_{\rm part}$ will depend on energy as well. Indeed, a calculation using the model described in [5,27] is shown in Fig. 2. While the number of participants varies very little with energy (as is obvious, since only variations in higher order moments are permitted by definition), the number of collisions varies systematically with the increase of the inelastic scattering cross section.

The best fit, also shown in the figure, can be parameterized by

$$N_{cp}(\sqrt{s}) = T\left[\ln\left(\frac{\sqrt{s}}{m_p}\right)\right]^2 + G\ln\left(\frac{\sqrt{s}}{m_p}\right) + J, \tag{8}$$

Where T, G, J are fit parameters (see footnote 1; the χ^2 here is negligible, $\approx 10^{-5}$). For the top centrality bin of Au-Au, used later in multiplicity fits, their numerical values are $T = 0.043$, $G = -0.045$, and $J = 2.1$

While, as panel (a) shows, the event-by-event $N_{\rm part}$ distribution is relatively constant in energy (unsurprisingly, since it is bounded between zero and twice the mass number globally and to a good approximation is related to the local transverse density at each point in the transverse area), the ratio of $N_{\rm coll}/N_{\rm part}$ has a monotonic increase. It is intuitively clear that in a general two-component model this will introduce an energy dependence on dN/dy which does not appear in rapidity, since both $N_{\rm part}$ and $N_{\rm coll}$ are parameters characterizing the whole event, rapidity independently.

It is clear that α and f are strongly correlated by experimental data. We therefore made a χ^2 contour using the data of [24] to fit them. The results are shown in Fig. 3(b). We highlight the curve going through the major axis of the ellipse in f-α space, parametrizeable by

$$f_{\rm fit}(\alpha) = F_1\alpha^4 + F_2\alpha^3 + F_3\alpha^2 + F_4\alpha + F_5, \tag{9}$$

where the fitted values are $F_{1,\ldots,5} = -250.1$, 175.8, -53.6, 9.9, 0.14 (see footnote 1; $\chi^2 \approx 10^{-2}$).

The $\mathcal{N}_0$ parameter of course is also correlated with f, α. In terms of α its minimum can be parametrized as

$$\mathcal{N}_0 \simeq N_1\alpha^2 + N_2\alpha + N_3, \tag{10}$$

[1] We note in passing that event-by-event fluctuations and higher cumulants of these two mechanisms are expected to be very different. $N_{\rm coll}$ reflects particles produced within hard scatterings, suggesting that fluctuations are a convolution of binomial fluctuations and inherent random variations in fragmentation functions. On the other hand, if wounded nuclei emit hadrons via color neutralization, the negative binomial distribution would be natural, as this is the distribution of the "number of throws" required to get an outcome. If parton generation is "locally" not color neutral, one needs a negative binomial number of generations before you get a color-neutral state. Thus, an energy and system size investigation of fluctuation scaling, in particular the so-called Koba-Nielsen-Olsen scaling, would parallel the limiting fragmentation investigation outlined here. This is left to a forthcoming work.

[2] Note that this form was chosen because it provides a good interpolation of the experimental data ($\chi^2 = 0.97$) and should not be taken as physically motivated.

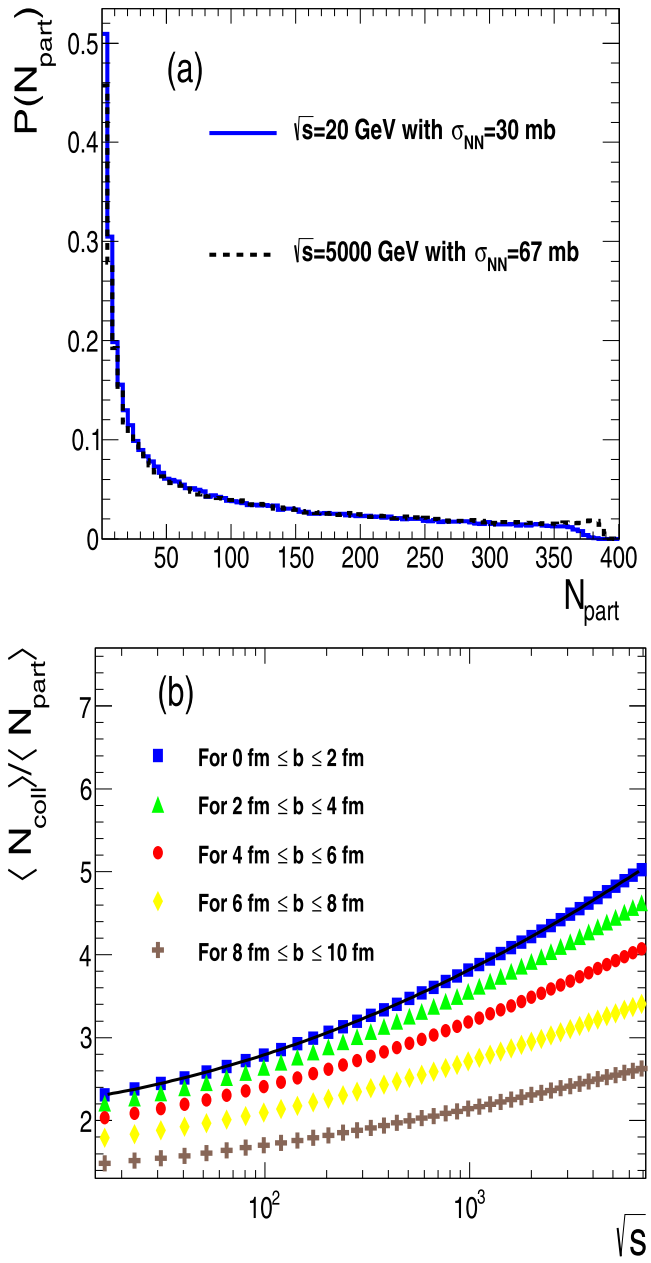

FIG. 2. (a) The distribution of $N_{\rm part}$ as a function of energy. (b) The subsequent evolution of $N_{\rm coll}/N_{\rm part}$ in a Glauber model, with the best fit to the most central collisions added as a line

where $N_{1,\ldots,3} = -3.35,\ 1.32,\ 0.27$ (see footnote 1; $\chi^2 \approx 10^{-3}$).

We also note that the inclusion of both in a simultaneous fit greatly widens the space of f-α allowed, and in fact the fit seems to prefer a value of f considerably lower ($N_{\rm coll}$ dominated ancestors) than most of the work of this kind.

Figure 3(a), showing the fit quality of the opposite ends compared to experimental data, however, demonstrates one must not take this too seriously (although it can also give a decent fit of eccentricities; see for instance [39]) since the correlation between the fit parameters (as well as the overall normalization) is such that the two extremes do a very similar job of fitting the data and in fact generate a nearly identical curve. This and the disagreements at lower energy also make the χ^2 contour deviate significantly from the Gaussian.

Given these ambiguities, our strategy, rather than focusing on inferring a tighter bound on f-α space, is to use this fit to define a relation between the two variables, based on $f_{\rm fit}(\alpha)$, that fixes the midrapidity curve, and investigate limiting fragmentation given this relation. We therefore use the graph where Eq. (9) was used to parametrize $f_{\rm fit}(\alpha)$: we now scan through α to check how limiting fragmentation behaves for the class of Glauber models fitting midrapidity.

First, we use the opposite ends of (f, α) parameter space to investigate the behavior away from midrapidity of dN/dy, assuming Gaussian distributions in a manner similar to [17], with widths given kinematically and heights by the observed midrapidity. As can be seen in Fig 4, at both extremes limiting fragmentation is broken in a similar manner.

To quantify this violation further, we used a Gaussian, a trapezoidal, and a triangular distribution in η, with the bottom base given by kinematics $\Delta\eta_{\rm bottom} = \ln(\sqrt{s}/m_p)$, while the top base was universally a fraction, either $\Delta\eta_{\rm top} = 0$ (for a triangle and a Gaussian) or $\Delta\eta_{\rm top} = \Delta\eta_{\rm bottom}/2$ of the bottom base (for a trapezium). The height is chosen by the two-component model calculation.

The results are shown in Fig. 5. As can be seen, all cases exhibit practically constant and sizable violations of limiting fragmentation for all values of (f, α). The constancy is not a surprise, given the good fit of the central rapidity for all values of α, $f_{\rm fit}(\alpha)$. Furthermore, the quantitative amount for the violation of limiting fragmentation is to a good extent independent of both the value of α and the model used. We can therefore conclude with some confidence that this violation is approximately what can be expected in any distribution respecting midrapidity multiplicity and kinematic constraints. Note that, while very simple, these distributions have the same functional form as the pseudorapidity distributions measured in [29]. The triangular distribution also has the advantage of generating limiting fragmentation analytically at $\alpha \to 0$, $f \to 1$, while the Gaussian one reproduces it to a good approximation for $\alpha = 1/4$, $f = 1$ [10].

This limit, which is not captured in our plot since it is very far from the best fit line $f \to 1$, prefers higher numbers of α, which breaks limiting fragmentation. $\alpha \to 0$ prefers very low f (collision-dominated multiplicity) which again breaks limiting fragmentation. We can therefore make the qualitative statement that a general Glauber model cannot realistically fit midrapidity multiplicity without a significant break in limiting fragmentation. This is experimentally testable.

It is very simple to understand the universality of the violation of the limiting fragmentation, and generalize from Gaussians and trapeziums (where one can get a quantitative analytically calculable answer) with respect to any rapidity shape, where the difference should be within a factor $O(1)$ of the one calculated here.

All we need to assume are the following crucial assumptions for the Glauber model of particle production as described here:

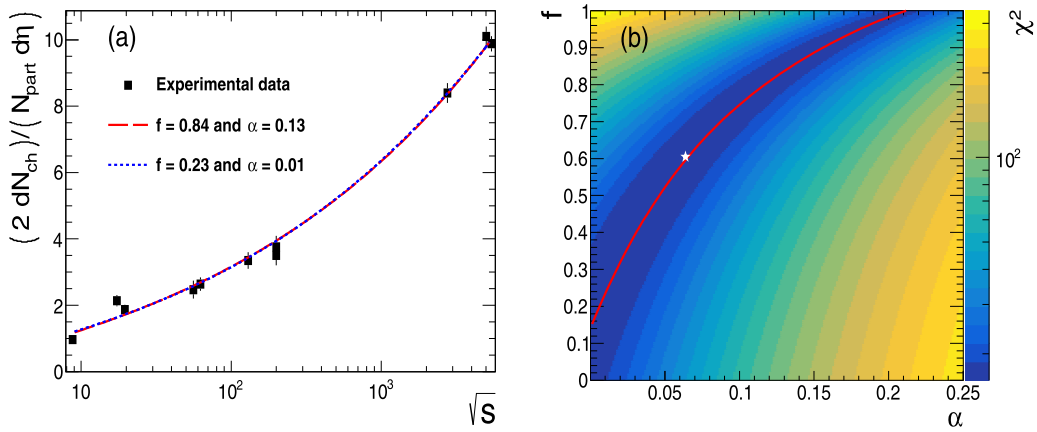

FIG. 3. (a) Experimental data, referenced in [24] with the opposite ends of the fitted curve. (b) Contour of fit to midrapidity experimental data in α-f, with best fit line highlighted and a star denoting the χ^2 minimum.

(i) An emission function $F(y)$ which is specific to each ancestor. Different emission functions $F_{\rm part}(y)$ and $F_{\rm coll}(y)$ would not change the result

$$\frac{dN}{dy} = \sum_{N_{anc}} F(y) \sim N_{\rm part} F_{\rm part}(y) + N_{\rm coll} F_{\rm coll}(y). \quad (11)$$

(ii) The base of the emission function is kinematically determined to be $\sim \ln(\frac{\sqrt{s}}{m_p})$.

This section makes it clear that, as long as $N_{\rm coll}/N_{\rm part}$ evolves nonlogarithmically and $f \neq 0$ (as is mandated by experimental data), $N_{\rm part}^{-1} d^2N/d\eta^2$ cannot be universal. These statements will hold as long as assumptions (i) and (ii) are correct, and putting in reasonable distributions shows violations big enough to be qualitative rather than small corrections. In particular, the middle scenario in Fig. 1 (smooth approach to limiting fragmentation) is excluded by assumption (i) and by the inability of $f = 1$ models to fit midrapidity data, since for this to happen the rapidity distributions of individual ancestors must be aware of how the number of ancestors (producing the midrapidity peak) increases.

Let us illustrate this by extending our analytically solvable model: We take the trapezium distribution [Fig. 5(b)] but impose an arbitrary variation with $\sqrt{s}$ of the top rapidity plateau $\Delta(\sqrt{s})$ (Since this variation is arbitrary, this includes the case of $N_{\rm part}$ and $N_{\rm coll}$ having different plateaus (Fig. 6). Calling $N_{cp}(\sqrt{s})$ the ratio of participants to collisions [shown in Fig. 2 and Eq. (8)], we will have

$$\left.\frac{d^2N}{dy^2}\right|_{|y-y_0|\ll y_0} = \frac{Z_1}{Z_2},$$
$$Z_1 = -4\mathcal{N}_0 N_{\rm part}[f + (1-f)N_{cp}(\sqrt{s})]\left(\sqrt{s}^{\alpha} - 1\right),$$
$$Z_2 = \alpha[1 - \Delta(\sqrt{s})] \ln\left(\frac{\sqrt{s}}{m_p}\right). \quad (12)$$

Since the only things that are permitted to vary with energy are N_{cp} and Δ, this means that for limiting fragmentation to occur we must have

$$F_1(f, \alpha, N_{cp}(\sqrt{s})) = CF_2(\Delta(\sqrt{s})), \quad (13)$$

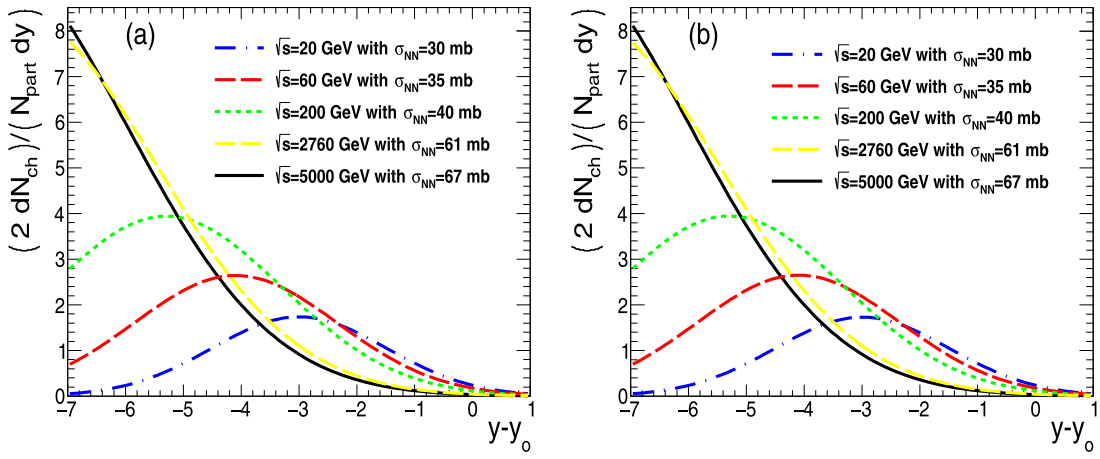

FIG. 4. Gaussian rapidity dependencies with the width given by a Landau-like kinematic parametrization, and the height extrapolated from pairs of parameters f, α adjusted to reproduce data at midrapidity

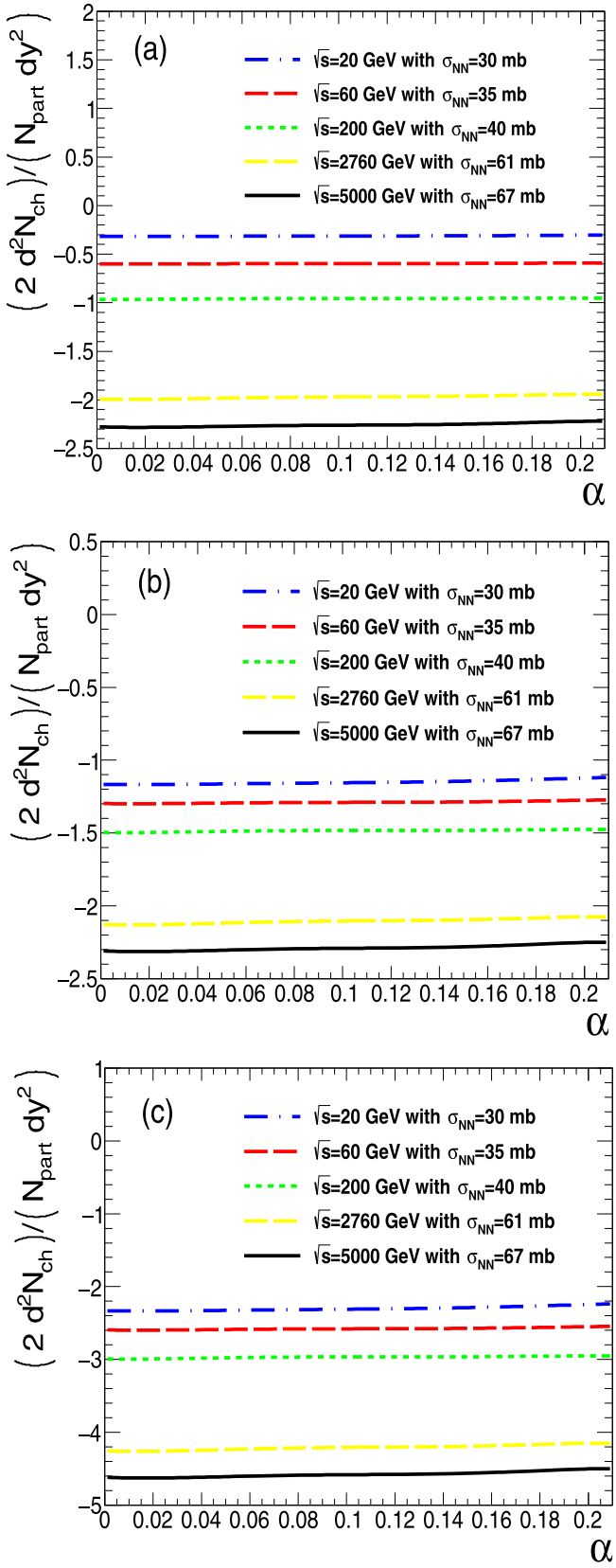

FIG. 5. The gradient of the rapidity distribution at $y - y_0 = 3$ assuming Gaussian (a), triangular (b), and trapezoidal (c) shapes respectively, against the α parameter, with the corresponding $f = f_{\rm fit}(\alpha)$ parameter adjusted to reproduce data at midrapidity (Fig. 3).

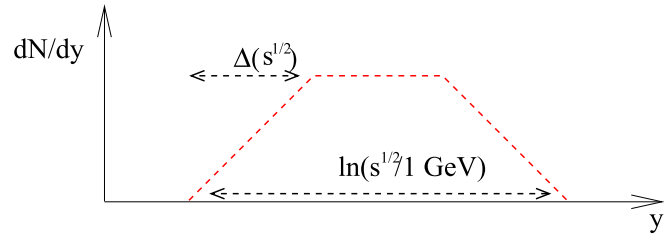

FIG. 6. A trapezoidal rapidity distribution with a varying width parameter

where C is the constant defined in Eq. (1) and

$$F_1(f,\alpha,\sqrt{s}) = \frac{\mathcal{N}_0 N_{\rm part}}{\alpha} \times [f + (1-f)N_{cp}(\sqrt{s})](\sqrt{s}^{\alpha} - 1), \quad (14)$$

$$F_2(\Delta(\sqrt{s})) = -\frac{1}{4}[1 - \Delta(\sqrt{s})]\ln\left(\frac{\sqrt{s}}{m_p}\right). \quad (15)$$

This is physically impossible, not only because the first function depends on nuclear geometry and the cross-section area and the second just on the partonic structure, but because the first depends on α and the second does not.

For instance, let us suppose that the Δ evolution arises from particle distributions produced in $N_{\rm part}$ and $N_{\rm coll}$ having different widths (i.e., the fragmentation region is different for wounded nuclei than for parton-parton collisions, a wholly reasonable scenario), parametrized as $\Delta_{1,2}\ln(\frac{\sqrt{s}}{m_p})$, where $\Delta_{1,2}$ are fixed coefficients. This would mean

$$\Delta(\sqrt{s}) = \frac{\mathcal{N}_0 N_{\rm part}(\sqrt{s}^{\alpha} - 1)}{\alpha} \times \left(\frac{f}{\Delta_1} + \frac{(1-f)}{\Delta_2}N_{cp}(\sqrt{s})\right). \quad (16)$$

This will lead to the supposed "constant" C of the form

$$C = \frac{-4\mathcal{N}_0 N_{\rm part}(\sqrt{s}^{\alpha} - 1)}{\alpha \ln\left(\frac{\sqrt{s}}{m_p}\right)} \times \frac{[f + (1-f)N_{cp}(\sqrt{s})]}{\left[1 - \left(\frac{\mathcal{N}_0 N_{\rm part}(\sqrt{s}^{\alpha}-1)}{\alpha}\right)W(\Delta_1,\Delta_2,f,\sqrt{s})\right]}, \quad (17)$$

where

$$W(\Delta_1,\Delta_2,f,\sqrt{s}) = \frac{f}{\Delta_1} + \frac{(1-f)}{\Delta_2}N_{cp}(\sqrt{s}).$$

A plot of $\sqrt{s}C^{-1}dC/d\sqrt{s}$, for $f = f_{\rm fit}(\alpha)$ and α adjusted according to Fig. 3, $\mathcal{N}_0$ given by the correlated values Eq. (10), and $\Delta_{1,2}$ scaled by $\ln(\frac{\sqrt{s}}{m_p})$ is shown in Fig. 7. It confirms the universality of breaking of limiting fragmentation within the Glauber model, by showing that any reasonable value of $\Delta_{1,2}$ does not satisfy Eq. (13) for an f, α given by the best fit parametrization (9). It is easy to see that similar equivalency, albeit with additional rapidity dependence, will arise for a Gaussian and any other shape. A comparison between pA and AA for the same number of participants (as in Fig. 1) would

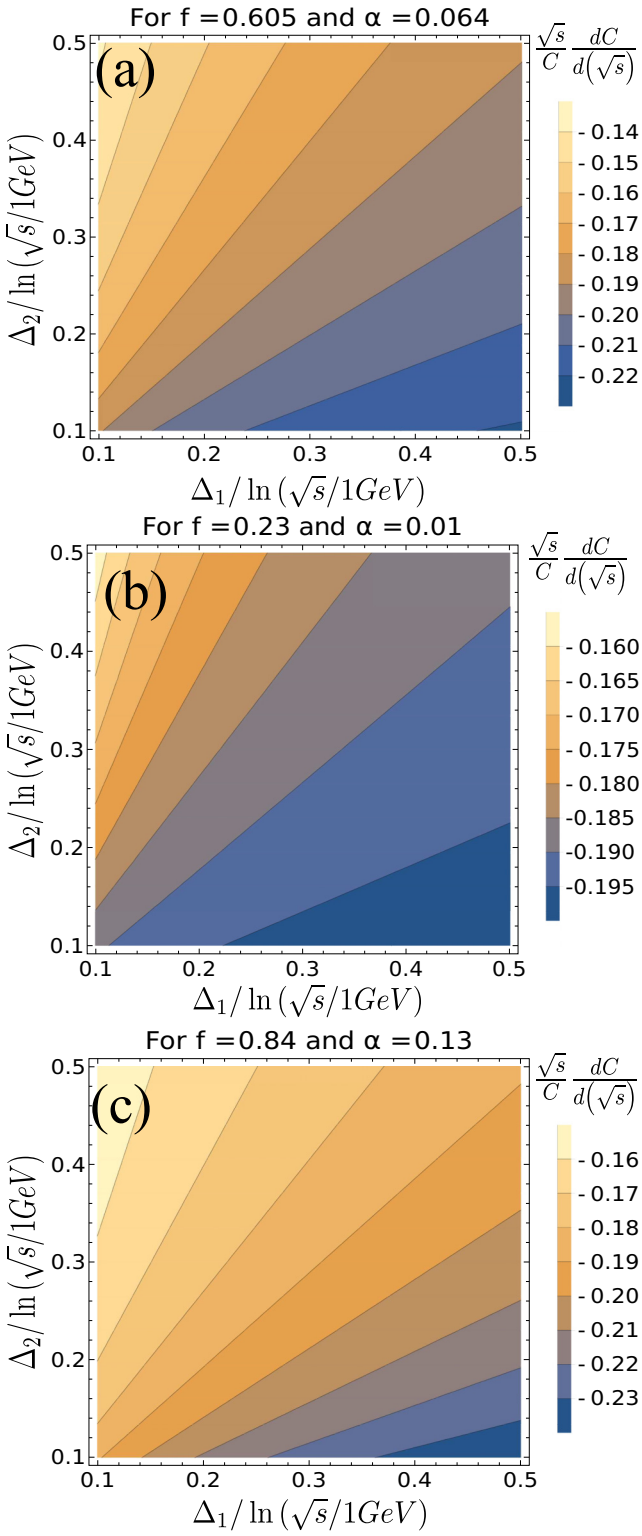

FIG. 7. A scan of the expected limiting fragmentation violation in conformal units, $\sqrt{s}C^{-1}dC/d\sqrt{s}$, as a function of the best fit parameters in rapidity data α, $f_{\rm fit}(\alpha)$ (kept fixed) as well as $\Delta_{1,2}$ (on the axes), the widths of $N_{\rm part}$ and $N_{\rm coll}$. The derivative was taken at $\sqrt{s} = 200$ GeV. The three panels show the choices of f, α corresponding to the extremes of Fig. 3 [panels (a) and (c)] and the χ^2 minimum [panel (b)]. We verified that intermediate values are well described by a smooth interpolation between these two

be even more problematic since for pA $N_{\rm coll} \simeq N_{\rm part} - 1$ while for AA no such constraint is present.

We also show (Fig. 8) a compendium of estimates of $d^2N/d^2\eta$ and the observable seen in Fig. 7, obtained with numerical derivatives of experimental data compiled in [9] from [6] and for energies of 130 and 200 GeV (where dN/dy per participant is still approximately logarithmic with $\sqrt{s}$). Figure 8(a) appears to show that limiting fragmentation is broken only weakly, if at all. Figure 8(b), where a common normalization with theory was included, shows a breaking comparable to Fig. 7 in magnitude, but, unlike Fig. 7, spread around zero.

This figure is included as a proof of concept, but extreme care should be taken to interpret it as a quantitative limit. The data it is based on do not include error bars ([9] and references therein), and the multiplication by $\sqrt{s}$, necessary to define a dimensionless observable quantifying breaking, magnifies fluctuations invisible to the naked eye in Fig. 8(a). The fact that the points oscillate chaotically around zero strongly suggests that Fig. 8(b) is in fact dominated by statistical fluctuations, and quantitatively the violation of limiting fragmentation until the top RHIC energies is parametrically smaller than that of Fig. 7. A reanalysis such as that of Fig. 8 by the members of the experimental collaborations, with access to raw data, together with LHC data could potentially verify this claim.

Additionally, as was already noted in [9], extrapolating to LHC energies (Figs. 8 and 10 of that work) means that any quantitative disagreement with limiting fragmentation seen at RHIC would be amplified by the departure from the logarithmic energy dependence per participant at LHC energies (predicted by a number of models in [9] and confirmed). A logarithmic derivative, such as the one we defined ($\sqrt{s}d/d\sqrt{s} = d/d\ln\sqrt{s}$), allows us to highlight the quantitative violation of limiting fragmentation inherent in this effect despite the large energy gap between top RHIC and bottom LHC energies.

Given these considerations, Figs. 7 and 8 strongly motivate a precise quantitative estimate of the violation of limiting fragmentation across both LHC and RHIC energy regimes.

The persistence of limiting fragmentation, especially in the LHC energy regime, would preclude the system from being a superposition of nucleon-nucleon collisions. However, such a picture was called into doubt in the literature by models where partonic degrees of freedom appear at the nuclear level, such as the color glass and the wounded quark models. In the next section we describe what such models imply for limiting fragmentation.

III. PARTONIC NUCLEAR DYNAMICS AND LIMITING FRAGMENTATION

It is obvious from the discussion in the previous section that any $N_{\rm coll}$ dependence on multiplicity is inherently incompatible with limiting fragmentation. Ultimately, the last section can be summarized in a simple statement: It is an experimental fact that the cross-sectional area changes with energy, and this by itself will introduce an energy-dependent

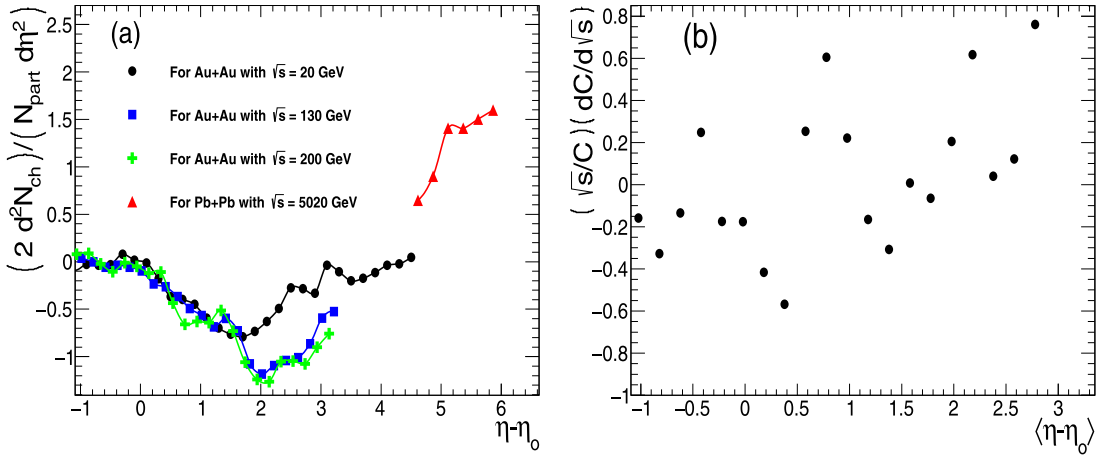

FIG. 8. (a) $d^2N/d\eta^2$ and the observable in Fig. 7 obtained from experimental data, from the compilation in [9] (data from [6]) and [29]. (b) The numerical derivative calculated between 130 and 200 GeV.

contribution which, unless unnecessary fine tuning is introduced, will never seep into rapidity dependence.

All Glauber model analyses so far, however, show that some $N_{\rm coll}$ dependence is unavoidable if nucleon-nucleon collisions are extrapolated from pp to AA. One answer to this is the "wounded quark model" [32–36] (more accurately described as wounded *parton* model), which also could be justified from a dynamics such as that of [40].

It is unsurprising that wounded parton models have more leeway in fitting multiplicities with only participant scaling, since by definition they have a greater number of parameters: Very little is known about configuration space parton distribution of the nucleons (the "three constituent quarks" picture is an obvious simplification at high energies). Thus, one can substitute the effect of varying f, α in the two-component model with parameters describing the parton transverse and longitudinal distribution within the nucleon. The works cited above, in their own way, essentially accomplished this.

One can, however, also note that the fundamental difference here is that the "size of the nucleon," i.e., the Fourier transform of the nucleon form factor, is something that is allowed to vary with Bjorken x. Boost invariance, built into the parton model, means that space-time rapidity of the parton's trajectory and momentum rapidity are equal:

$$y = -\ln(x^{-1}) = \frac{1}{2}\ln\left(\frac{z+t}{z-t}\right). \tag{18}$$

As Fig. 9 shows, one can imagine a nucleon-nucleon collision within such a model as a superposition of a "train" of collisions, of transverse shapes located in different bins in Bjorken x.

Each participant starts from that Bjorken x, is shifted by one or more collisions to a different Bjorken x, and finally

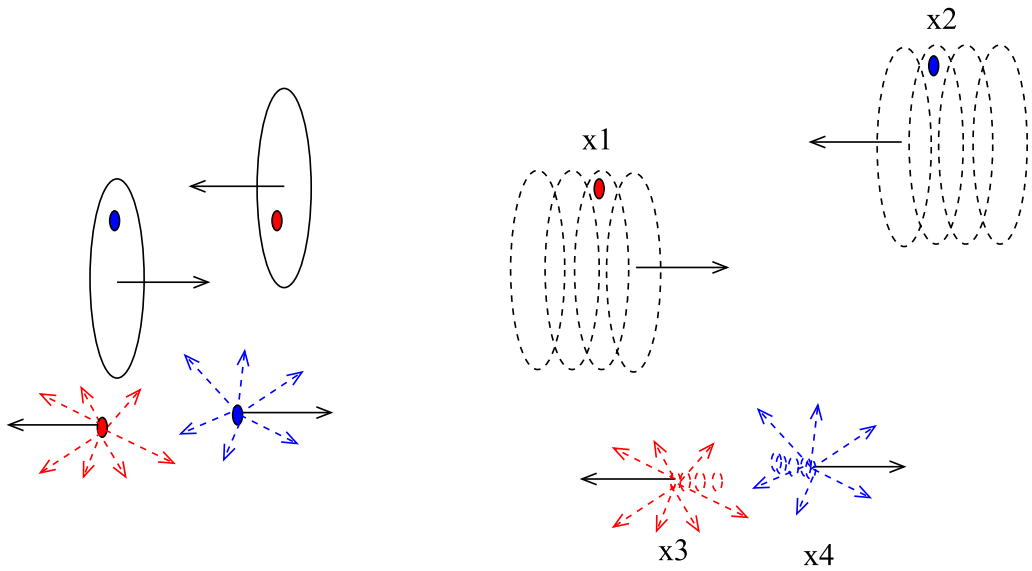

FIG. 9. An illustration of the Glauber model based on wounded nucleons vs wounded partons. The latter has the distinction that each initial degree of freedom is at a different Bjorken x and shifts to a different x (i.e., final rapidity bin) once "wounded"

emits hadrons according to some distribution in rapidity determined by this final Bjorken x and fragmentation dynamics. The transverse nucleon shape at each x can also vary, and hence can mimic the variation of the "nucleon-nucleon cross section" with energy via the growth of the typical x with $\sqrt{s}$. Crucially, *the same variation* will occur in rapidity.

We investigate this idea with a toy model, combining a Glauber-type wounded parton model with a transverse quark size parametrized by $Q_s(x)$. Note that "s" in Q_s here might stand for size, not for saturation, i.e., it might represent a nucleon-specific rather than a transverse space-specific scale. All that we need is that Q_s is allowed to vary with Bjorken x. Basically, we assume "a tube of pancakes," each at a momentum and spacetime rapidity $e^{\pm x}$ of transverse size $Q_s(x)^{-1}$.

Let us briefly recap as to how this setup generates limiting fragmentation. The key assumptions relating limiting fragmentation to the parton model are

(a) All transverse momentum scale (denoted here by k) dependence is via $[k/Q_s(x)]^2$, a scale (reflecting saturation or size) depending only on x, usually $\sim x^{-\lambda}$.
(b) Normalization of $\tilde{T}(k,x)$ depends only on x [possibly via $Q_s(x)$] or is constant:
$$\int^{Q^2} \frac{d^2p_T}{p_T^2}\tilde{T}(p_T,x) \sim \mathcal{N}(x).$$
(c) Asymptotic freedom. Most partons are produced with little shifts in momenta. As a result, if $-\ln x$ is not $\gg 1$, one can assume $x \sim e^{-y}$ of the produced parton. In simple language, asymptotic freedom prevents partons away from midrapidity from fragmenting away from their rapidity bin

Let us illustrate how these assumptions *generically* give limiting fragmentation in pp collisions. Extending the Glauber model formalism to such a system is simple:

$$T_A(\mathbf{s}) = \int \rho_A(\mathbf{s},z_A)dz_A \to T_A(\mathbf{s},x) \sim F(x,Q_s(x)) \quad (19)$$

with $Q_s(x)$ characterizing the nucleon size at that x.

Now the usual definition for T_{AB} can be duly updated to

$$T_{AB}(\hat{\mathbf{b}},y) = \int T_A(\mathbf{s},x_A)T_B(\mathbf{s}-\hat{\mathbf{b}},x_B)\delta(x_A+x_B-e^{-y})d^2s. \quad (20)$$

If one uses the standard formula for the number of collisions,

$$\langle N_{\rm coll}(\hat{\mathbf{b}})\rangle = ABT_{AB}(\hat{\mathbf{b}})\sigma^{qq}_{\rm inel}, \quad (21)$$

with Bjorken scaling and conformal cross-section

$$\sigma^{qq}_{\rm inel} = \frac{1}{p_T^2}, \quad x_{A,B} = \frac{p_T e^{\pm y}}{\sqrt{s}} \quad (22)$$

parallels, the k_T factorization formula (derived through somewhat different physics [41]), reminiscent of a simplification of Eq. (4), is

$$\frac{dN}{dy} \sim \int d^2p_T \frac{1}{p_T^2}\int k\,dk\Big[A(x_A)\tilde{T}_A(x_A,k)\int dx_{A,B}B(x_B) \times \tilde{T}_B(x_B,p_T-k)\delta(x_A+x_B-e^{-y})\Big]\Big|_{x_{A,B}=\frac{p_T e^{\pm y}}{\sqrt{s}}}, \quad (23)$$

where $A(x_A)$, $B(x_B)$ are the absolute numbers of partons sitting in that x bin.

Now, according to assumption (a) $\tilde{T}$ is a Fourier-transformable function (the Fourier transform of the nuclear transverse size, which serves as a proxy for the scattering cross section) characterized by a size parameter Q_s:

$$\tilde{T}(x,k) \sim \tilde{F}\left(x,\frac{k}{Q_s(x)}\right), \quad Q_s(x) \sim x^{-\lambda}. \quad (24)$$

Performing all momentum integrals and using the unitarity assumption (b), we will get, up to a constant,

$$\frac{dN}{dy} \sim x_A f(x_A) \simeq e^{-y} f(e^{-y}),$$

where the last approximation, due to (c), gives rise to limiting fragmentation.

This is more or less how [12,13] derived limiting fragmentation in the color glass. However, the $\simeq$ signs should be examined in more detail when different system sizes as well as energies are compared, since generally in the CGC scenario $f(x,Q_s(x,N_{\rm part}))$ will maintain a residual $N_{\rm part}$ dependence, breaking the scaling of $N_{\rm part}^{-1}d^2N/d\eta^2$.

To quantify this effect within the saturation scenario, we have plotted this variable for the running coupling Balitsky-Kovchegov (rcBK) model developed in [15], where the nonlinear gluon evolution is solved numerically, as well as the Kharzeeva-Levina-Nardia (KLN) model [42] which parametrizes the qualitative features of this evolution, extrapolating from midrapidity. The normalization was performed with the data of the centrality dependence of charged particle multiplicity at midrapidity at 2.76 TeV [43]. The result is plotted in Fig. 10. As can be seen, the dynamics is qualitatively similar to scenario (b) of Fig. 1, with the normalized rapidity density smoothly reaching the universal limit. In rcBK this limit is reached slightly sooner than in KLN, although the scaling is not perfect in either. This confirms earlier results published on this topic [12,13], as well as predictions in [9] (Fig. 8).

However, one should keep in mind that the reason is that the rcBK results reach the $\eta_{\rm beam}$ region is conditional on the numerical limitation in the rcBK case, since the gluon distribution function has to be extrapolated for $x > 0.01$, which is exactly the kinematical regime where one of the hadrons is probed in the fragmentation region, where a significant quark admixture is present. It is reasonable to believe, however, that an additional admixture of sources not present at midrapidity would weaken the scaling. In addition, of course, the k_T factorization scenario of Eq. (23) and saturation in general are, by an order-of-magnitude estimate, likely to break down at lower RHIC energies, where experimentally limiting fragmentation works quite well (as Fig. 8 and the accompanying discussion showed), and, for that matter, the calculation of Fig. 10, if taken "literally," seems to reproduce this fact.

In this spirit, the rcBK calculation done here should be considered as a "lower limit" to limiting fragmentation violation within the CGC scenario, in the sense that this calculation is likely to be rather inaccurate at many energies and rapidities, but deviations from it are likely to break limiting fragmentation. And, comparing different system sizes at LHC and

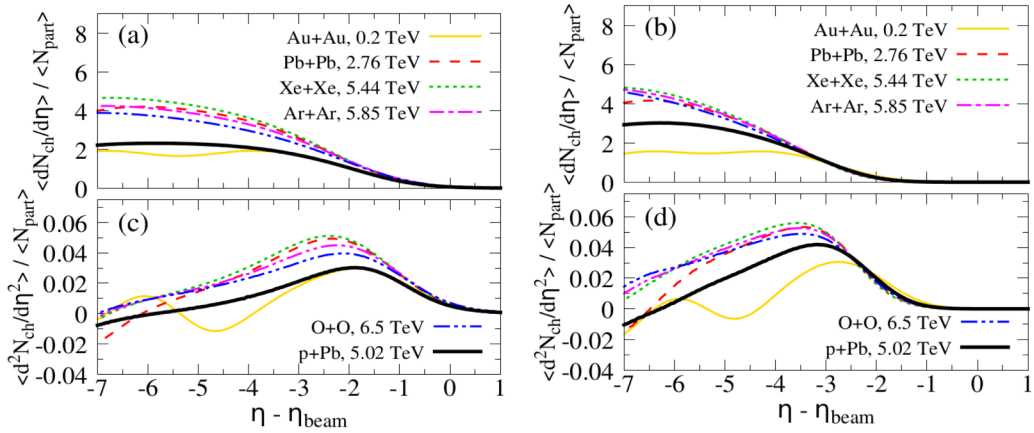

FIG. 10. $dN/d\eta$ and $N_{\rm part}^{-1}d^2N/d\eta^2$ calculated in KLN [panels (a),(c)] and rcBK [panels (b),(d)] scenarios. The normalization was adjusted to reproduce the data in [43] using the usual parton-hadron duality parameters

bottom RHIC energies of Fig. 10, even in the ideal scenario the breaking of limiting fragmentation would be substantial. Putting the results of Fig. 10 into the form of Figs. 7 and 8 is shown in Fig. 11. We confirm that the deviation from limiting fragmentation in these models is higher than the data, and surprisingly also of the Glauber model when normalized. So, what conclusions would we draw if it is found to hold, quantitatively, for pA, dA, and AA collisions up to LHC energies? Given the success of the wounded quark model, we could combine it with partonic limiting fragmentation insights to calculate rapidity-dependent "wounded partons": the number of partons that experienced at least one collision. Without resort to nonperturbative physics, it is difficult to justify this model, since particle production is thought to be perturbative in the regime where partons are good degrees of freedom, and perturbative dynamics does not scale with participants. That said, since wide rapidity intervals correspond to wide spacetime rapidity intervals, the idea that confinement effects seep into the longitudinal evolution of partons, and hence "wounded partons" become relevant, is not so unreasonable. Since, phenomenologically, a wounded quark model appears to work [32–35] well enough that fusing this with the geometrical deep inelastic scattering warrants a try.

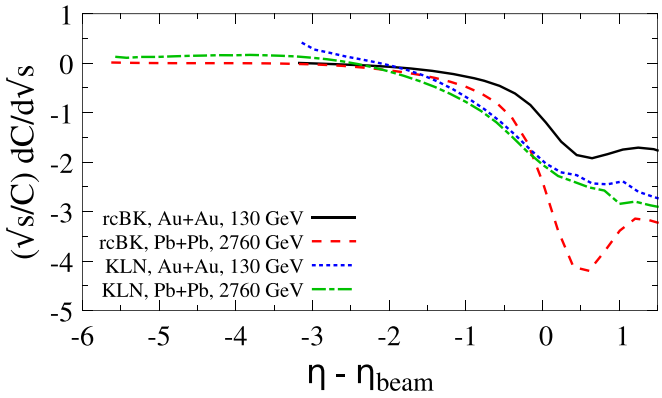

FIG. 11. Violation of limiting fragmentation quantified in the way of Figs. 7 and 8 in the KLN and rcBK models.

We use the usual formula for the number of wounded nucleons adapted to the wounded parton picture, with the Fourier transform of the transverse participant density of nucleus A,

$$\tilde{N}^{qA}_{\rm part}(x_A,k_A,A,B)\simeq A(x_A)\tilde{T}_A(k_A,x_A)\int d^2k_B dx_B \times\left(1-\left(1-\frac{\tilde{T}_B(k_A-k_B,x_B)}{(k_A-k_B)^2}\right)^{B(x_B)}\right). \tag{25}$$

If one introduces the emission function [36] for a wounded parton characterized by x_A, k_A to emit a particle of p_T, y as $F(p_T,y,k_A,x_A)$, one can get the multiplicity in terms of the wounded partons.

For each nucleon collision we will have

$$\frac{dN}{dy}=\int p_T dp_T dk dx F(p_T,y,k,x) \times\left(\tilde{N}^{qA}_{\rm part}(x,k,A,B)+\tilde{N}^{qB}_{\rm part}(x,k,B,A)\right). \tag{26}$$

Expanding to first order in $\tilde{T}_B(k)/k^2$ it is possible to reduce this expression into a sum of terms of the form

$$\frac{dN}{dy}\sim\int p_T dp_T\int dx_A k_A dk_A A(x_A)\tilde{T}(k_A,x_A) \times F(p_T,y,k_A,x_A)\int dx_B k_B dk_B\frac{B(x_B)\tilde{T}_B(k_A-k_B,x_B)}{(k_A-k_B)^2}.$$

Assumption (b), together with the reasonable requirement that $B(x\to 0)$ does not diverge, will change the second integral to an approximately constant value for each nucleon. Assumption (c) ensures $F(p_T,y,k,x)\simeq F(p_T,y,k,e^y)$ and also that $k_A-k_B\simeq k_A$. Additionally assuming the equivalence of (b) and (c) for $F(\cdots)$, i.e., that it is only a narrowly peaked and normalized function of y differences, will ensure that the resulting function will only depend on e^{-y}, analogously with Eq. (23). Hence, limiting fragmentation is recovered.

Unlike the color glass, however, one can straightforwardly extend limiting fragmentation to AA. What prevented limiting fragmentation in the wounded nucleon model was the fact that $f=1$ in Eq. (5) is incompatible with any choice of

α in Eq. (6) that fits multiplicity at midrapidity. Physically wounded partons, by allowing participants to be spread across x and nuclear size to vary across x, make it possible to combine a purely participant $f = 1$ scaling (which does not depend on $\sqrt{s}$) with strongly nonlogarithmic dependence of multiplicity per participant, obtained through variation with x of $A(x)$ and $\tilde{T}(x)$. The difference with the color glass is that $\tilde{T}(x)$ would still be individual nuclear collisions, albeit with a size that effectively depends in Bjorken x. Thus, one expects $Q_s(x)$ ("s" is size and not saturation here) to be independent of the number of participants, leading to a universal $N_{\rm part}^{-1} d^2N/d\eta^2$.

Previous literature on wounded quarks [32–36] considers nucleon transverse substructure but does not separate it into x distributions. This allows the authors to fit midrapidity data or rapidity data at a given energy. Generalizing these models along the lines presented in this work while maintaining the current phenomenological agreements seems like a straightforward exercise, albeit one beyond the scope of this paper. Thus, if an eventual confirmation of a $N_{\rm part}^{-1} d^2N/d\eta^2$ universal across energies and system sizes will indeed be observed, the wounded parton model will be a promising avenue to model this.

IV. DISCUSSION AND SUMMARY

In this work, we have examined the behavior of $N_{\rm part}^{-1} d^2N/d\eta^2$, the multiplicity density per participant, at high rapidity, motivated by the near independence with energy of its first derivative. The universal behavior of $N_{\rm part}^{-1} d^2N/d\eta^2$ was found to generally break in a Glauber model, particularly in the two-component models usually used to fit data at midrapidity. Thus, this is a good observable to constrain such models.

The reason these models fail is physically very simple: Inasmuch as both collisions and participants contribute to multiplicity, the energy dependence of the two factors is nontrivial and generically different from the rapidity dependence. This means that the height of the distribution depends nontrivially on the energy while the width is constrained by kinematics. Absent unnatural cancellations, limiting fragmentation should be broken.

The only way to restore it is to make multiplicity dependence entirely driven by wounded nucleons at all energies. This is, however, not enough, since the price for this is to make the multiplicity per nucleon in AA collisions rise with energy much faster than logarithmically. This, as well as clashing with experimental data below LHC, will generally break limiting fragmentation as well. A "wounded parton model" might be able to evade such a constraint since there the "size of the wounded degree of freedom," instead of being encoded in the cross section, is allowed to vary with longitudinal x, which in this case is tightly correlated with momentum rapidity. Energy and rapidity distributions are therefore naturally correlated in this regime. The observation of a universal $N_{\rm part}^{-1} d^2N/d\eta^2$ could be indicative of "wounded quark" dynamics.

One can ask how universal is the class of models reproducing such universality. Models such as the color glass [13] share some similarities with the wounded parton scenario considered in the previous section, but there nuclei lose their individuality and Q_s is common to the same area of transverse space ("s" in color glass models is saturation instead of size). Also, Eq. (23) (k_T factorization) used to connect the gluon density to particle density in this regime leads to a universal slope in the fragmentation region in the same way as Bjorken scaling.

Indeed, as [13] finds, approximate limiting fragmentation is explained by the factorization of parton distributions in target and projectile at large rapidities together with the fact that the multiplicity distribution is directly proportional to the parton density in the target and relatively independent of the scales of the process. The wounded parton model conjectured in the previous section shares these characteristics. We note, however, that [13] predicts some violation of limiting fragmentation inasmuch as the assumptions above cease to have validity. Thus, the calculations of [13] point to some violation of $N_{\rm part}^{-1} d^2N/d\eta^2$, which turns out to be comparable, if not larger, than the Glauber model.

Looking at [44–46] a similar discussion can be made about anti—-de Sitter and conformal field theory (AdS-CFT) initial states, where the breaking of scaling appears even stronger as it is controlled both by a critical transparency and by the coupling constant. Thus, some breaking of limiting fragmentation, when all energies and systems sizes are concerned, appears likely in all models claiming connection to field theory (the wounded parton model so far does not).

We continue with experimental considerations. To our knowledge, so far measurements at high enough rapidity to compare to even top RHIC energy were not done, with the closest experimental measurements being [29–31]. Also, a result, seeming to confirm scenario (b) of Fig. 1, was obtained for $dE_T/d\eta$ by the CMS Collaboration using the CASTOR detector [47]. Since multiplicity and transverse energy have a nontrivial separate dependence which is also sensitive to system size [48], we hesitate at drawing conclusions there. Reference [49], in contrast to CASTOR, has reported a breaking of limiting fragmentation in inclusive photons. While QED processes are not expected to limit fragmentation [there is no hadronization, and $y \sim \ln(1/x)$ is not expected to hold], some 85% of photons in [49] are thought to come from π^0 decays.[3] Hence, this result makes a breaking of limiting fragmentation at LHC energies likely. In summary, until a direct measurement of both $dN/d\eta$ and $dE_T/d\eta$ is performed, relying on these data to make a conclusive statement is difficult.

Studies comparing the rapidity dependence of p-Pb and Pb-Pb are totally lacking. Comparing with smaller asymmetric systems, such as Pb-Pb and p-Pb collisions, where a deviation from purely wounded dynamics should be more pronounced, the observable can be studied on the "same side," as discussed in [9].

The fact that most experiments focus the detector on midrapidity of course makes this measurement problematic. We want to point out, however, that this is a bulk observable, not requiring particle identification or momentum measurement;

[3]We thank Mauro Cosentino for pointing this out.

it is problematic since particles at high rapidity are highly relativistic. The existing small experiments at rapidity comparable to lower RHIC energies [50,51] as well as LHCb [52] could take part in this investigation together with the larger collaborations.

Since the calculations here were focused on proof-of-concept estimates testing for violation, we will give a "cartoon" of what we expect in each scenario, with the the alternatives are summarized in Fig. 1. The key is to go to a high enough rapidity as to compare with a lower energy. If limiting fragmentation still holds, $N_{\rm part}^{-1}d^2N/d\eta^2$ will evolve to smoothly "touch" the corresponding value at midrapidity of that energy [scenario (b) of Fig. 1]. Otherwise, the slopes will be different.

In this paper, we have limited ourselves to multiplicity, which, given a low-viscosity nearly isentropic fluid evolution, can be considered to be an initial-state effect [53]. However, past experimental results also reported to have seen limiting fragmentation for elliptic flow at RHIC energies [7]. This is considered to be a final-state effect, sensitive primarily to the transport coefficients and freeze-out dynamics of the system [22]. Should limiting fragmentation of flow observables, or even of average transverse momentum, be confirmed at the LHC,[4] especially in events of same eccentricity but different size, one might have to rethink this paradigm and start exploring scenarios where "flow" arises as an initial-state effect [55–57] in the systems concerned.

In conclusion, we have examined limiting fragmentation in various phenomenological models, namely Glauber, color glass, and wounded quarks. Glauber models generally fail to reproduce limiting fragmentation at LHC energy once they are tuned to reproduce LHC data. The same is true for color glass models once different system sizes are considered. We have further argued that wounded parton scenarios have the potential to model limiting fragmentation also at LHC energies and for all system sizes. Since such limiting fragmentation has not been verified to date, no such model can be considered to have been ruled out. Rather, this paper motivates an experimental search for this observable, and generally for a comparison between low-energy bulk observables and the high rapidity limit of the same observable at high energy. We eagerly away these experimental results.

ACKNOWLEDGMENTS

G.T. acknowledges support from FAPESP proc. 2017/06508-7, participation in FAPESP thematic project 2017/05685-2, and CNPQ fund for productivity 301996/2014-8. K.J.G. was supported by CAPES CNPQ graduate Fellowship 88882.328988/2019-01. This work is a part of the project INCT-FNA Proc. No. 464898/2014-5. We thank Fernando Navarra, Radoslaw Ryblewski, Mauro Rogerio Cosentino, and Wit Busza for fruitful discussions and suggestions. A.V.G. acknowledges the Brazilian funding agency FAPESP for financial support through grants 2017/14974-8 and 2018/23677-0.

APPENDIX: A NOTE ON RAPIDITY AND PSEUDORAPIDITY

We note that the experimental observable usually measured here is not rapidity (the z direction, where momentum is p_z; the total momentum is p),

$$y = \tanh^{-1}\left(\frac{p_z}{\sqrt{p^2+m^2}}\right), \tag{A1}$$

but the pseudorapidity, a function of the angle with the beam axis which does not require particle identification:

$$\eta = \tanh^{-1}\left(\frac{p_z}{p}\right) \tag{A2}$$

The rapidity is the observable with the "nice" transformation property of being linear under boosts. It is also related to Bjorken x by Eq. (3) and equal to the space-time rapidity in the boost-invariant limit [Eq. (18)]. However, it is very difficult to measure for ultra-relativistic particles as it requires particle identification, and/or the simultaneous measurement of the energy and momentum.

The pseudorapidity has no "nice" properties, but it is very easy to measure as it is directly related to the longitudinal angle θ:

$$\eta = -\ln\tan\left(\frac{\theta}{2}\right) \tag{A3}$$

At midrapidity, binning the distribution in terms of η rather than y causes the near-gaussian distribution measured in [8] to aquire a "plateau" necessitating a double gaussian fit, as was done in [17].

Away for midrapidity, for the great majority of produced particles, the two are interchangeable:

$$\eta - y \sim -O\left(\frac{m^6 p_z^3}{p^9}\right) \simeq \frac{\cos\theta}{2}\left(\frac{m}{p_T}\right)^2. \tag{A4}$$

In this work we do not fit the rapidity distribution globally, but rather rather concentrate on observables local in rapidity sensitive to limiting fragmentation: the slope in the fragmentation region

$$|y| \geqslant O\left(\frac{1}{2}-\frac{3}{4}\right)|y_0| \tag{A5}$$

and its derivative with respect to $\sqrt{s}$. Hence, we assume differences between y and η to be negligible.

To estimate the goodness of this assumption, lacking particle identification and p_T measurements away from midrapidity, we shall take a limiting fragmentation inspired rapidity distribution [9],

$$\frac{dN}{p_T dp_T dy} = C\left(\ln\left[\frac{\sqrt{s}}{m_p}\right]-y\right)\exp\left[-\frac{p_T}{2\langle p_T\rangle}\right], \tag{A6}$$

[4]We note in passing that the rapidity dependence of elliptic flow in pA, dA, and AA collisions measured in [54] is qualitatively what you would expect given a universal $v_2(p_T)$ curve together with a dependence of average momentum on multiplicity and system size expected from [48].

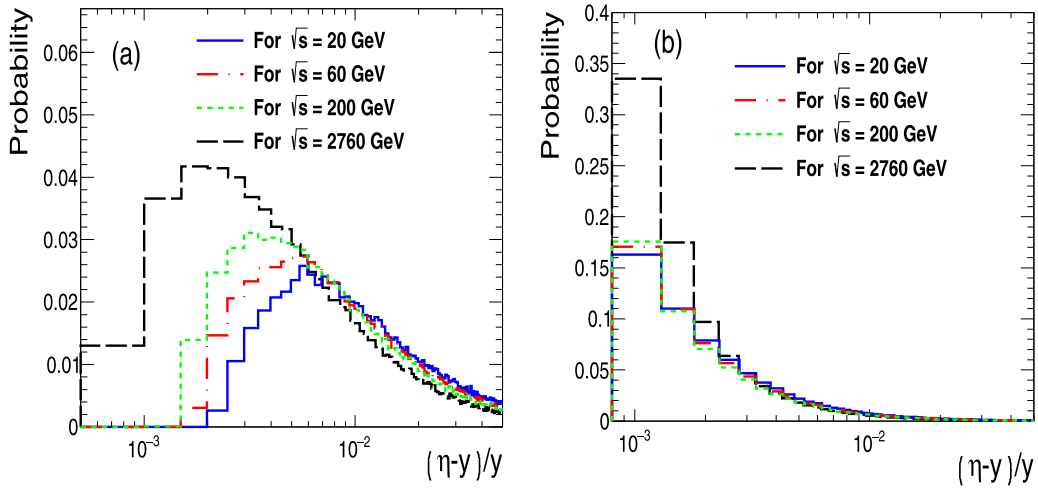

FIG. 12. Estimated distribution of $(y-\eta)/y$ for protons (a) and π (b) assuming $\langle p_T \rangle$ does not deviate from its midrapidity value.

where

$$\langle p_T \rangle = \begin{cases} 500\ \text{MeV}, & \pi, \\ 1\ \text{GeV}, & p, \end{cases} \quad C \simeq -0.65,$$

and histogram $(\eta - y)/y$. The results are shown in Fig. 12 for protons and pions, in the region of rapidity considered in the rest of the paper, $y_0 - y \leqslant 6$ or so (the figure should be symmetrized if both "target" and "projectile" are considered) and assuming a thermal transverse p_T distribution with $\langle p_T \rangle$ weakly dependent on rapidity (central values are as in [58]). As can be seen, any reasonable admixture between baryons and mesons will result in a systematic error of order of a percent if η and y. This error is significantly lower than other experimental errors, and hence is neglected in this work.

To correct this would require a rapidity as well as the energy dependence of the temperature, baryochemical potential, and $\langle p_T \rangle$, something currently unknown (measured very partially in [8] and [58]), but, in the spirit of the rest of the paper, any nontrivial variation of this is likely to spoil limiting fragmentation.

The observables treated in this paper, defined around the second derivative of the rapidity distribution away from midrapidity, are optimized to minimize the effect of the rapidity-pseudorapidity interchange. Alternatively, for example, analyzing the width of the rapidity distribution as a whole (as was done, for example, in [17]) is much more sensitive to this distinction, since the midrapidity plateau varies significantly between η and y, and this influences both the height the width of the distribution in a particle-dependent manner (see [59] for a discussion of this issue at low energies).

[1] J. E. Elias, W. Busza, C. Halliwell, D. Luckey, P. Swartz, L. Votta, and C. Young, Phys. Rev. D **22**, 13 (1980).
[2] W. Busza, J. E. Elias, D. F. Jacobs, P. A. Swartz, C. C. Young, and M. R. Sogard, Phys. Rev. Lett. **34**, 836 (1975).
[3] J. Benecke, T. T. Chou, C. N. Yang, and E. Yen, Phys. Rev. **188**, 2159 (1969).
[4] F. Halzen and A. D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics* (Wiley, New York, 1984).
[5] B. B. Back *et al.* (PHOBOS Collaboration), Phys. Rev. C **74**, 021901 (2006).
[6] B. B. Back *et al.*, Phys. Rev. Lett. **91**, 052303 (2003).
[7] B. B. Back *et al.*, Nucl. Phys. A **757**, 28 (2005).
[8] I. Arsene *et al.* (BRAHMS Collaboration), Nucl. Phys. A **757**, 1 (2005).
[9] S. Jeon, V. T. Pop, and M. Bleicher, Phys. Rev. C **69**, 044904 (2004).
[10] P. Steinberg, Acta Phys. Hung. A **24**, 51 (2005).
[11] W. Busza, Nucl. Phys. A **854**, 57 (2011).
[12] J. Jalilian-Marian, Phys. Rev. C **70**, 027902 (2004).
[13] F. Gelis, A. M. Stasto, and R. Venugopalan, Eur. Phys. J. C **48**, 489 (2006).
[14] F. O. Duraes, A. V. Giannini, V. P. Goncalves, and F. S. Navarra, Phys. Rev. C **89**, 035205 (2014).
[15] A. Dumitru, A. V. Giannini, M. Luzum, and Y. Nara, Phys. Lett. B **784**, 417 (2018).
[16] Z. W. Lin, C. M. Ko, B. A. Li, B. Zhang, and S. Pal, Phys. Rev. C **72**, 064901 (2005).
[17] P. Sahoo, P. Pareek, S. K. Tiwari, and R. Sahoo, Phys. Rev. C **99**, 044906 (2019).
[18] A. Capella, U. Sukhatme, C. I. Tan, and J. Tran Thanh Van, Phys. Rep. **236**, 225 (1994).
[19] P. Brogueira, J. Dias de Deus, and C. Pajares, Phys. Rev. C **75**, 054908 (2007).
[20] I. Bautista, C. Pajares, J. G. Milhano, and J. Dias de Deus, Phys. Rev. C **86**, 034909 (2012).
[21] B. Kellers and G. Wolschin, Prog. Theor. Exp. Phys. **2019**, 053D03 (2019).
[22] G. Torrieri, Phys. Rev. C **82**, 054906 (2010).
[23] G. Torrieri, EPJ Web Conf. **13**, 04002 (2011).

[24] S. Acharya *et al.* (ALICE Collaboration), Phys. Lett. **790**, 35 (2019).
[25] M. L. Miller, K. Reygers, S. J. Sanders, and P. Steinberg, Annu. Rev. Nucl. Part. Sci. **57**, 205 (2007).
[26] C. Loizides, J. Nagle, and P. Steinberg, SoftwareX **1-2**, 13 (2015).
[27] F. S. Borcsik, Maestrado thesis, University of Campinas, 2018, http://www.repositorio.unicamp.br/handle/REPOSIP/332829.
[28] K. Werner, Phys. Rev. Lett. **98**, 152301 (2007).
[29] J. Adam *et al.* (ALICE Collaboration), Phys. Lett. B **772**, 567 (2017).
[30] A. M. Sirunyan *et al.* (CMS Collaboration), Phys. Rev. Lett. B (2019), doi: 10.1016/j.physletb.2019.135049.
[31] G. Aad *et al.* (ATLAS Collaboration), Phys. Lett. B **710**, 363 (2012).
[32] J. T. Mitchell, D. V. Perepelitsa, M. J. Tannenbaum, and P. W. Stankus, Phys. Rev. C **93**, 054910 (2016).
[33] A. Adare *et al.* (PHENIX Collaboration), Phys. Rev. C **93**, 024901 (2016).
[34] R. Nouicer, Eur. Phys. J. C **49**, 281 (2007).
[35] P. Bożek, W. Broniowski, and M. Rybczyński, Phys. Rev. C **94**, 014902 (2016).
[36] M. Barej, A. Bzdak, and P. Gutowski, arXiv:1904.01435.
[37] R. G. Arnold *et al.*, Phys. Rev. Lett. **52**, 727 (1984).
[38] M. Aaboud *et al.* (ATLAS Collaboration), Phys. Rev. Lett. **117**, 182002 (2016).
[39] P. Romatschke and U. Romatschke, Phys. Rev. Lett. **99**, 172301 (2007).
[40] S. J. Brodsky, J. F. Gunion, and J. H. Kuhn, Phys. Rev. Lett. **39**, 1120 (1977).
[41] Y. V. Kovchegov and K. Tuchin, Phys. Rev. D **65**, 074026 (2002).
[42] D. Kharzeev, E. Levin, and M. Nardi, Nucl. Phys. A **747**, 609 (2005).
[43] K. Aamodt *et al.* (ALICE Collaboration), Phys. Rev. Lett. **105**, 252301 (2010).
[44] J. Casalderrey-Solana, M. P. Heller, D. Mateos, and W. van der Schee, Phys. Rev. Lett. **111**, 181601 (2013).
[45] J. Casalderrey-Solana, M. P. Heller, D. Mateos, and W. van der Schee, Phys. Rev. Lett. **112**, 221602 (2014).
[46] S. Grozdanov and W. van der Schee, Phys. Rev. Lett. **119**, 011601 (2017).
[47] A. M. Sirunyan *et al.* (CMS Collaboration), Eur. Phys. J. C **79**, 391 (2019).
[48] B. B. Abelev *et al.* (ALICE Collaboration), Phys. Lett. B **727**, 371 (2013).
[49] B. B. Abelev *et al.* (ALICE Collaboration), Eur. Phys. J. C **75**, 146 (2015).
[50] T. Csörgő (TOTEM Collaboration), EPJ Web Conf. **206**, 06004 (2019).
[51] O. Adriani *et al.* (LHCf Collaboration), J. High Energy Phys. 11 (2018) 073.
[52] Y. Zhang (LHCb Collaboration), arXiv:1605.07509.
[53] A. Dumitru, E. Molnar, and Y. Nara, Phys. Rev. C **76**, 024910 (2007).
[54] C. Aidala *et al.* (PHENIX Collaboration), Nat. Phys. **15**, 214 (2019).
[55] G. Gambini and G. Torrieri, Eur. Phys. J. A **53**, 17 (2017).
[56] C. Bierlich, Nucl. Phys. A **982**, 499 (2019).
[57] K. Dusling, M. Mace, and R. Venugopalan, Phys. Rev. Lett. **120**, 042002 (2018).
[58] L. Adamczyk *et al.* (STAR Collaboration), Phys. Rev. C **96**, 044904 (2017).
[59] F. Becattini, J. Cleymans, A. Keranen, E. Suhonen, and K. Redlich, Phys. Rev. C **64**, 024901 (2001).