

Revealing a deep connection between factorization and saturation: New insight into modeling high-energy proton-proton and nucleus-nucleus scattering in the EPOS4 framework

K. Werner 

SUBATECH, Nantes University, IN2P3/CNRS, IMT Atlantique, 44300 Nantes, France



(Received 16 March 2023; revised 10 July 2023; accepted 19 September 2023; published 6 December 2023)

It is known that multiple partonic scatterings in high-energy proton-proton (pp) collisions must happen in parallel. However, a rigorous parallel scattering formalism, taking energy sharing properly into account, fails to reproduce factorization, which on the other hand is the basis of almost all pp event generators. In addition, binary scaling in nuclear scatterings is badly violated. These problems are usually “solved” by simply not considering strictly parallel scatterings, which is not a solution. I will report on new ideas (leading to EPOS4), which allow recovering perfectly factorization, and also binary scaling in AA collisions, in a rigorous unbiased parallel scattering formalism. In this new approach, dynamical saturation scales play a crucial role, and this seems to be the missing piece needed to reconcile parallel scattering with factorization. From a practical point of view, one can compute within the EPOS4 framework parton distribution functions (EPOS PDFs) and use them to compute inclusive pp cross sections. So, for the first time, one may compute inclusive jet production (for heavy or light flavors) at very high transverse momentum (p_T) and at the same time in the same formalism study flow effects at low p_T in high-multiplicity pp events, making EPOS4 a full-scale “general purpose event generator”. I discuss applications, essentially multiplicity dependencies (of particle ratios, mean p_T , charm production) which are very strongly affected by the saturation issues discussed in this paper.

DOI: [10.1103/PhysRevC.108.064903](https://doi.org/10.1103/PhysRevC.108.064903)

I. SOME INTRODUCTORY REMARKS ABOUT FACTORIZATION, PARALLEL SCATTERING, AND ENERGY SHARING

Two major discoveries made it possible to reliably compute cross sections in high-energy proton-proton (pp) scattering. There is first of all the fact that the coupling constant α_s of strong interactions becomes weaker with increasing scale, referred to as “asymptotic freedom” [1,2], which allows the use of perturbation theory to compute parton-parton cross sections. The other crucial issue is called “factorization” [3,4], which amounts to separating short- and long-distance physics at some “factorization scale” μ , which allows one to write the inclusive pp cross section as a convolution of two parton distribution functions (PDFs) and a (calculable) elementary parton-parton cross section. The PDFs contain all the long-distance physics, below the scale μ . Factorization in connection with asymptotic freedom turned out to be an extremely powerful concept, with numerous important applications. Extended to collisions of two nuclei, composed of A and B nucleons, factorization means that the cross section for rare processes is given as AB times the pp cross section. This is usually referred to as “binary scaling”.

Factorization is an impressive tool, being very useful when it comes to studying inclusive particle production, but there are very interesting cases not falling into this category, like high-multiplicity events in proton-proton scattering in the TeV energy range, where a very large number of parton-parton scatterings contribute. Such events are particularly interesting, since the CMS Collaboration observed long-range near-side angular correlations for the first time in high-multiplicity proton-proton collisions [5], which was before considered to be a strong signal for collectivity in heavy ion collisions. And studying such high-multiplicity events (and multiplicity dependencies of observables) goes much beyond the frame covered by factorization. Here one needs an appropriate tool, able to deal with multiple scatterings.

The most important fact about multiple parton-parton scatterings is that they must occur in parallel, and not sequentially, as I am going to justify in the following. It is known that parton-parton scatterings are preceded by a series of successive parton emissions according to Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) evolution equations [6–8]. In particular, the first emitted partons carry a large momentum corresponding to a large γ factor, so they are “long-lived” particles. Correspondingly, the whole reaction takes a long time, which makes it impossible to have two (or more) successive parton-parton scatterings. Multiple scattering must therefore happen in parallel. In the case of nucleus-nucleus scattering, the nucleon-nucleon collisions also happen in parallel, and this is simply due to the fact that at very high energies, the “reaction time” (the time it takes for the two nuclei to pass through each other) is much shorter than the particle

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) 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 SCOAP³.

formation time. So first all the interactions are realized (instantaneously), and particle production comes later. One has a “double parallel scattering” scenario: the nucleon-nucleon scatterings happen in parallel, and for each nucleon-nucleon scattering the parton-parton collisions occur in parallel.

In the case of multiple scatterings, energy-momentum conservation is an important issue. Of course, everybody agrees on that and all event generators do conserve it. But this has to be seen in the light of the underlying theory. Multiple scatterings have been incorporated in an S-matrix approach in the Gribov-Regge (GR) theory a long time ago [6,9–11]. All the scatterings “are equal” in the sense that there is no sequence, nothing like “a first scattering” and a “second scattering” and so on. However, as discussed in [12,13], there is an inconsistency: the “energy-momentum sharing” is simply not taken into account. In a strictly parallel scenario, the initial energy-momentum has to be shared among n parallel scatterings and a projectile and a target remnant in an unbiased way, in the following referred to as “rigorous parallel scattering scenario”, which amounts to using integrands (for pp scattering) like

$$\prod_{i=1}^{n+2} f_i(p_i) \times \delta\left(p_{\text{initial}} - \sum_{i=1}^{n+2} p_i\right) \quad (1)$$

(with p referring to four-momentum). In the case of AA scattering, one has products of δ -functions of the type Eq. (1). I insist on the fact that the theoretical basis (S-matrix theory) on one hand, and the Monte Carlo realization in event generators on the other hand, should deal with energy conservation in the same way, i.e., they should be 100% compatible with each other; this is also what I mean by a “rigorous parallel scattering scenario”. This is usually not the case, as in some early work of the author [14], where the underlying theory has no energy sharing, but the Monte Carlo realization does, and this is even in recent event generators the usual method.

Employing the “rigorous parallel scattering scenario”, one encounters highly multidimensional integrals that cannot be separated. In [15], nevertheless, energy sharing in the sense of Eq. (1) and its generalization to AA scattering could be implemented, and the technical difficulties (using Markov chain techniques) could be handled. I am not aware of any other attempt in this direction.

Let me discuss the fundamental differences between the “standard QCD generators” and the “rigorous parallel scattering scenario”. All “standard QCD generators” such as Pythia [16], Herwig [17], or Sherpa [18] take as starting point the factorization formula, sketched in Fig. 1. In this plot and all the following ones, I show for simplicity only gluons; in the real calculations all kinds of partons are considered. The two light blue thick lines represent the projectile and the target protons. The proton structure and the so-called spacelike parton cascade are taken care of by using parton distribution functions (PDFs) f , which allows writing the jet cross section as a convolution of these PDFs and an elementary QCD cross section for the Born process in the middle. This formula (still based on Fig. 1) serves as a probability distribution, which allows one to generate a sequence of hard

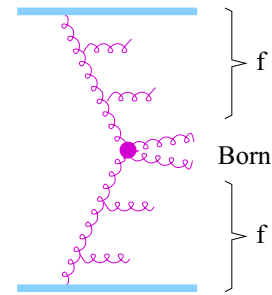


FIG. 1. Factorization formula.

processes, which are ordered in “hardness”. This is the method to introduce multiple parton scattering.

In the “rigorous parallel scattering scenario”, the starting point is a multiple scattering diagram as shown in Fig. 2 for $n = 3$ scatterings, where the corresponding mathematical formula contains in the case of pp scattering a δ function as in Eq. (1) for energy-momentum conservation. Here one also considers parton evolutions from both sides, but for each of the n interactions, so one cannot use the usual proton PDFs. Instead, one considers $2n$ evolutions, starting always from the first perturbative parton on the projectile side and the target side. But one is nevertheless able to define evolution functions E , which are based on the same DGLAP partial differential equations (see for example [3,4]), but in our case the initial condition is not a parton distribution $f(Q_0^2, x)$ in the proton at some initial scale Q_0^2 , but a parton carrying the full momentum fraction $x = 1$. The Monte Carlo procedure to generate partons will be done in two steps:

- (i) Step 1: The multiscattering formalism allows generating a number of scatterings n and in addition for each of the n scatterings its energy (expressed in terms of light-cone momentum fractions x_i^\pm), with 100% energy-momentum conservation (the cross section formulas contain a δ function to assure it, so energy-momentum violating configurations are never proposed).
- (ii) Step 2: With n and all the x_i^\pm known, one generates for each of the n scatterings the hard process based on a convolution $E_{\text{proj}} \otimes \text{Born} \otimes E_{\text{targ}}$, and then the parton emissions via backward evolution.

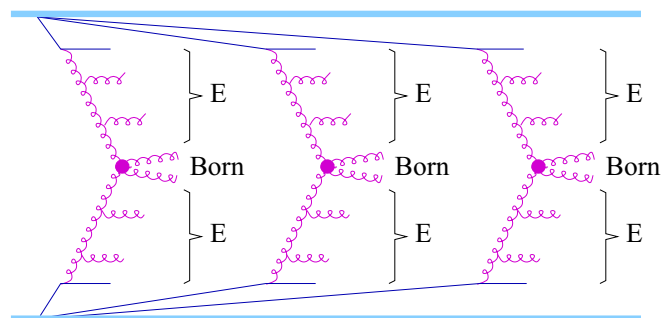


FIG. 2. Rigorous parallel scattering scenario, for $n = 3$ parallel scatterings.

The technical problems in the “rigorous parallel scattering scenario” can be handled, but there are conceptual problems. In the classical GR approach, it is known that in the case of inclusive cross sections the multiscattering contributions cancel, referred to as Abramovsky-Gribov-Kancheli (AGK) cancellations [11], and considering a Pomeron to be a parton ladder, one may deduce factorization and binary scaling in AA, as discussed in more detail in Sec. IV.

In the parallel scattering scenario with energy-momentum sharing, imposed in an unbiased fashion via a delta function as in Eq. (1), one does not get factorization (which requires a single Pomeron contribution), and one violates terribly binary scaling for AA scattering, as I am going to discuss in Sec. VI.

The solution to the problem is related to the treatment of saturation, as I discuss in a very qualitative fashion in Sec. II, and in detail in Sec. VII.

II. SOME INTRODUCTORY REMARKS ABOUT SATURATION AND ITS RELATION WITH ENERGY SHARING

The above sketched “rigorous parallel scattering scenario” is an elegant way to introduce unbiased parallel scattering, but in the end it does not work: one violates factorization (and binary scaling in AA). So something is still missing.

There is actually another important issue in high-energy scattering: with increasing energy, partons with very small momentum fractions $x \ll 1$ become increasingly important, since the PDFs at small x become large. This means that the parton density becomes large, and therefore the linear DGLAP evolution scheme is no longer valid, and nonlinear evolution takes over, considering explicitly gluon-gluon fusion. These phenomena are known as “small x physics” or “saturation” [19–32], the main effect being a screening of low transverse momentum (p_t) particle production (below a “saturation scale”). Saturation effects are expected to be even stronger in nucleus-nucleus collisions [20,21], simply because parton ladders emitted from different nucleons may fuse. At high energies, the diagrams for each scattering actually look more like the one shown in Fig. 3 (I do not consider, for simplicity, timelike parton emissions, but in the real EPOS4 simulations, they are of course taken care of). At least for scatterings carrying a large value of x^+x^- , one expects “nonlinear effects”, which means that two ladders which evolve first independently and in parallel, finally fuse. And only after that is the (linear) DGLAP evolution realized.

As mentioned above, such nonlinear effects lead to strong destructive interference, which may be summarized in terms of a saturation scale [20,21]. This is the motivation to treat these “saturation phenomena” not explicitly, but by introducing a saturation scale as the lower limit of the virtualities for the DGLAP evolutions, as sketched in Fig. 4. So the diagrams inside the red ellipses are replaced by two scales $Q_{\text{sat,proj}}^2$ and $Q_{\text{sat,targ}}^2$, and in pp scattering the two are equal. So the final version of the “rigorous parallel scattering scenario” in EPOS4 is sketched in Fig. 5. One still has DGLAP evolution, for each of the scatterings, but one introduces saturation scales. Most importantly, as discussed in great detail in Sec. VII, these scales are not constants: they depend on the number of scat-

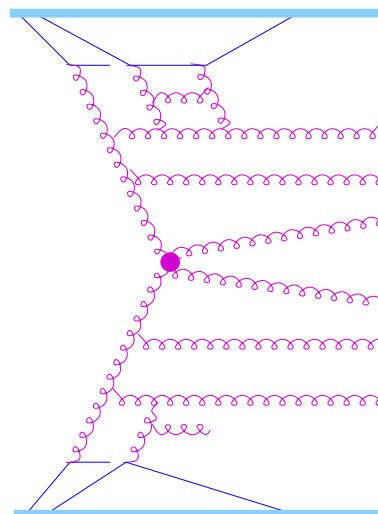


FIG. 3. Nonlinear effects: ladders which evolve first independently and in parallel finally fuse.

terings, and they depend as well on x^+ and x^- . A smart choice of these dependencies allows finally to recover factorization and binary scaling. One understands that there is a (so far unknown) very strong relation between factorization, energy conservation (or better energy sharing), parallel scattering, and saturation; see Fig. 6. Let me summarize the reasoning for this statement:

- (i) at high energies, multiple scatterings must happen in parallel, and there is nothing like a sequence or an ordering of elementary collisions;
- (ii) ignoring energy sharing (as in the GR approach), factorization and binary scaling are obtained (see Sec. IV);
- (iii) implementing energy sharing in the sense of a “rigorous parallel scattering scenario”, not only do the technical difficulties increase enormously, but there are conceptual problems: it spoils factorization and binary scaling (see Sec. VI);

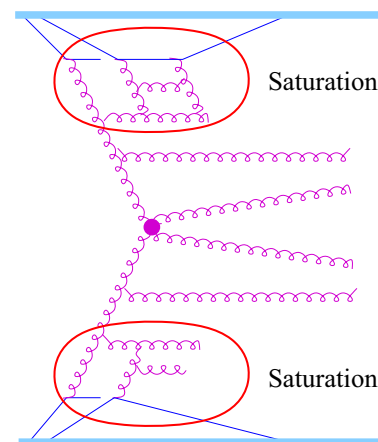


FIG. 4. Nonlinear effects (inside the red ellipses) are “summarized” in the form of saturation scales.

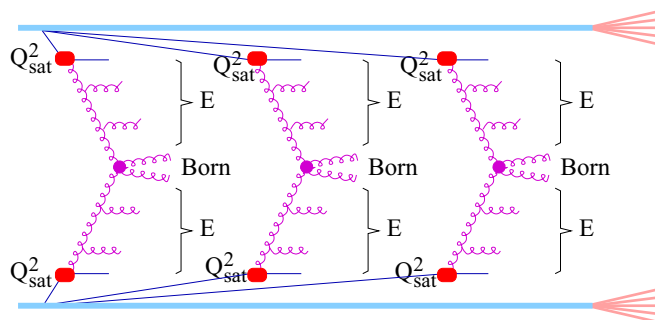


FIG. 5. Rigorous parallel scattering scenario, for $n = 3$ parallel scatterings, including nonlinear effects via saturation scales. The red symbols should remind one that the parts of the diagram representing nonlinear effects are replaced by simply using saturation scales.

- (iv) the only way out (it seems) is that one introduces saturation scales in a particular way, which recovers factorization and binary scaling (see Sec. VII).

Having solved the “factorization and binary scaling problem”, one may consider the “low p_t domain”—like studying high multiplicity (collective) phenomena (see Sec. X)—within a framework that has (finally) been proven to be compatible with the factorization approach, allowing one to do “high p_t physics” as well (as the generators that are based on factorization). In this sense EPOS4 is meant to be a “general purpose event generator”.

All this discussion about saturation and factorization is of fundamental importance since both are considered to be very important issues, but usually they are discussed independently. But the message of this paper is that they are connected, they affect each other (see Sec. VII), and they are just two aspects in a common approach.

Since saturation is so important in this approach, what about the other models? The above-mentioned “standard QCD generators” (based on factorization) do not explicitly deal with saturation (apart of a constant low virtuality cutoff), but certain features have similar effects. Let me consider AA scattering in the Pythia/Angantyr model [33,34]. As in EPOS4, there is first the “basic AA model” for $t = 0$, and in a second step there are string interactions, happening later. Concerning the basics AA model, the total S-matrix is given as product of sub-S-matrices. But in contrast to EPOS4, there are no energy-momentum arguments, and therefore no energy sharing. But one needs to introduce some kind of “sequence”, i.e., one loops over all the NN scatterings, and treats the interaction in two different ways: if a nucleon is already wounded

(having had already a scattering before) the current scattering is realized as diffractive scattering, called secondary scattering; otherwise a “normal” scattering according to Pythia happens, called primary scattering. The latter is in general a multiple parton scattering process. Here, one needs to introduce some ordering. The first sub-scattering is “normal”, whereas subsequent ones are not connected to projectile/target remnants as the first one, but they are connected to the parton of the previous sub-scatterings.

So EPOS4 and Pythia/Angantyr are fundamentally different, but in the latter there are certain features that have similar effects as the saturation scale in EPOS:

In Pythia/Angantyr one needs some ordering of NN collisions in AA scattering, which is needed to distinguish primary and secondary scatterings. This is necessary to avoid an overproduction of charged particles in AA collisions. In EPOS4, the same effect is obtained by treating all NN scatterings equally, but introducing the dynamical saturation scale.

Concerning multiple parton scatterings in NN collisions, in Pythia/Angantyr the first and the subsequent scatterings are treated differently with regard to the color connections. This is needed to get the experimentally observed increase of the mean transverse momentum with multiplicity. In EPOS4, one treats all subscatterings equally, but one has a saturation scale, which increases with multiplicity (as will be discussed later), and this is the main mechanism that leads to the increase of the mean transverse momentum with multiplicity.

It is of course dangerous to generalize based on few examples, so let me take the following statement as conjecture. I believe, based on the work in this paper, compared to other approaches, that one has two possibilities:

- (i) either one considers subsequent subscatterings (parton-parton or nucleon-nucleon) as strictly equal, with appropriate energy-sharing, which requires a dynamical saturation scale as crucial element;
- (ii) or one does not consider saturation (other than a simple cutoff in the parton cascade), but one needs to distinguish between primary and secondary scatterings (the first one and subsequent ones), for both parton-parton and nucleon-nucleon collisions, which requires some ordering.

This paper is meant to be an overview, with a minimum of technical details. The latter can be found in separate publications, such as [35–37].

After these introductory remarks, I will

- (i) in Sec. III present the EPOS4 formalism,
- (ii) in Sec. IV show how factorization appears naturally if energy conservation is dropped,
- (iii) in Sec. V show how energy sharing deforms Pomeron energy distributions,
- (iv) in Sec. VI show how deformed Pomeron energy distributions spoil factorization and binary scaling in case of a “naive” Pomeron definition,
- (v) in Sec. VII show how a dynamical saturation scale allows recovering factorization and binary scaling,
- (vi) in Secs. IX to XI discuss results affected by saturation.

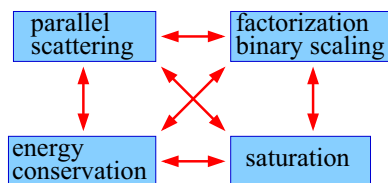


FIG. 6. Factorization, energy conservation, parallel scattering, and saturation: four concepts that are deeply connected.

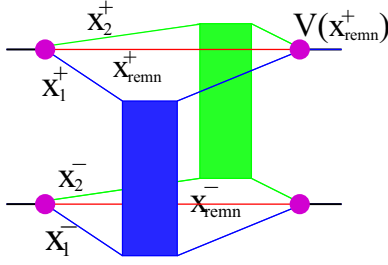


FIG. 7. Double scattering diagram.

III. EPOS4 S-MATRIX APPROACH TO REALIZE PARALLEL SCATTERINGS

I first consider pp scattering. An appropriate tool to implement parallel scatterings is provided by S-matrix theory (see [6,9–11,15]), where multiple parallel scatterings can be implemented in a simple and transparent fashion. Factorization and binary scaling are not “assumed”, they must come out. The S-matrix is by definition the representation $S_{ij} = \langle i|\hat{S}|j\rangle$ of the scattering operator \hat{S} using some basis of asymptotic states, and the corresponding T-matrix is defined via $S_{fi} = \delta_{fi} + i(2\pi)^4 \delta(p_f - p_i) T_{fi}$. Particularly important is the diagonal element, T_{ii} , representing elastic scattering, where the asymptotic state $|i\rangle$ corresponds to two incoming protons. Assuming purely transverse momentum transfer, one may Fourier transform T_{ii} with respect to the transverse momentum exchange \vec{k} and in addition divide by $2s$ to obtain some function $T(s, b)$, with the Mandelstam variable s and the impact parameter b , in the following simply named “T-matrix” T .

The EPOS4 S-matrix approach is based on the hypothesis that the T-matrix T can be written as a sum of products of “elementary” T-matrices $T_{\text{Pom}}^{(m)}$, the latter ones representing parton-parton scattering by exchanging a “Pomeron” (without specifying its nature for the moment), as

$$iT = \sum_{n=1}^{\infty} \int dX \frac{1}{n!} V^{(+)} \times \{iT_{\text{Pom}}^{(1)} \times \dots \times iT_{\text{Pom}}^{(n)}\} \times V^{(-)} \times \delta, \quad (2)$$

with the “vertices” $V^{(\pm)}$ representing the connection to the projectile and target remnants. The symbol X stands for all integration variables, to be specified in the following. In Fig. 7, I show a graphical representation of a double scattering ($n = 2$), where the blue and green boxes are the elementary T-matrices $T_{\text{Pom}}^{(m)}$, representing parton-parton scattering, and the magenta dots are the vertices $V^{(\pm)}$. The elementary T-matrices are characterized by the light-cone momentum fractions x_i^{\pm} of the incoming partons, in addition to s and the impact parameter b , so one has

$$T_{\text{Pom}}^{(i)} = T_{\text{Pom}}(x_i^+, x_i^-, s, b). \quad (3)$$

The precise content of the Pomerons (boxes) and the functional dependencies on these variables will be discussed later; the general discussion in this section does not depend on these details. The vertices depend on the light-cone momentum fractions of the remnants, x_{remn}^+ (projectile side) or x_{remn}^- (target

side), i.e.,

$$V^{(\pm)} = V(x_{\text{remn}}^{\pm}), \quad (4)$$

with a simple functional form (power law) of V . The “ δ ” in Eq. (2) stands for

$$\delta\left(1 - \sum_{i=1}^n x_i^+ - x_{\text{remn}}^+\right) \delta\left(1 - \sum_{i=1}^n x_i^- - x_{\text{remn}}^-\right), \quad (5)$$

to assure energy-momentum conservation, which will be crucial for the discussions in this paper. The integration $\int dX$ amounts to integrating over all light-cone momentum fractions. Each term (for $n > 1$) in the sum of Eq. (2) represents multiple scatterings happening in parallel, as it should.

The generalization of the multiple parallel scattering picture towards nucleus-nucleus (AA) collisions (including proton-nucleus) is trivial: one simply writes a product of pp expressions,

$$iT = \sum_{n_1, \dots, n_{AB}} \int dX \prod_{i=1}^A V^{(+i)} \times \prod_{j=1}^B V^{(-j)} \times \prod_{k=1}^{AB} \left\{ \frac{1}{n_k!} \{iT_{\text{Pom}}^{(k,1)} \times \dots \times iT_{\text{Pom}}^{(k,n_k)}\} \right\} \times \delta \quad (6)$$

for colliding two nuclei with mass numbers A and B , with at least one $n_k > 0$. Here, one has one vertex $V^{(m)}$ per remnant, and a sum of products of elementary T-matrices

$$T_{\text{Pom}}^{(k,v)} = T_{\text{Pom}}(x_{kv}^+, x_{kv}^-, s, b_k) \quad (7)$$

per nucleon-nucleon pair k . The “ δ ” in Eq. (6) stands for

$$\prod_{i=1}^A \delta\left(1 - \sum_{\substack{k,v \\ \pi(k)=i}} x_{kv}^+ - x_{\text{remn},i}^+\right) \prod_{j=1}^B \delta\left(1 - \sum_{\substack{k,v \\ \tau(k)=j}} x_{kv}^- - x_{\text{remn},j}^-\right), \quad (8)$$

where $\pi(k) = i$ amounts to summing Pomerons connected to projectile i and $\tau(k) = j$ to summing Pomerons connected to target j . This formula does *not* mean at all a sequence of pp collisions: they are perfectly happening in parallel; the crucial ingredient is the appearance of δ functions. The integration $\int dX$ here means integration over all light-cone momentum fractions and over all transverse position of the nucleons. The generalization Eq. (6) is conceptually trivial, but it should be noted that one has (for big nuclei) 10 000 000 dimensional nonseparable integrals.

So far I have discussed only elastic scattering for pp and AA, the connection with inelastic scattering provides the “optical theorem” (in b representation), which is at high energy given as

$$\sigma_{\text{tot}} = \int d^2b \text{cut } T, \quad (9)$$

with $\text{cut } T = \frac{1}{i} \text{disc } T$ (cut diagram), with disc T being the s -channel discontinuity $T(s + i\epsilon) - T(s - i\epsilon)$. So one needs to compute the “cut” of the complete diagram, cut T ; for example, for pp , one needs to evaluate expressions like

$$\text{cut}\{V^{(+)} \times \{iT_{\text{Pom}}^{(1)} \times \dots \times iT_{\text{Pom}}^{(n)}\} \times V^{(-)}\}. \quad (10)$$

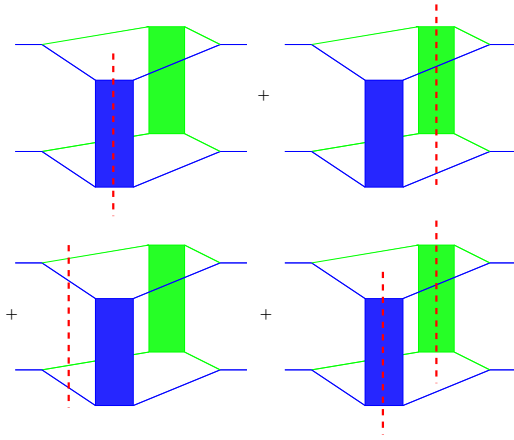


FIG. 8. Sum of all possible cuts of a two-Pomeron diagram.

Cutting a multi-Pomeron diagram corresponds to the sum of all possible cuts [38], considering, in particular, all possibilities of cutting or not any of the parallel Pomerons, so one has finally sums of products with some fraction of the Pomerons being cut (“cut Pomerons”), the others not (“uncut Pomerons”). I define G to be the cut of a single Pomeron,

$$G = G(x_i^+, x_i^-, s, b) = \text{cut } T_{\text{Pom}}(x_i^+, x_i^-, s, b). \quad (11)$$

An uncut Pomeron is by definition the sum of two contributions, the Pomeron being to the right or to the left of the cut, finally given as $-G$ (see the discussion in the next section). Since cut and uncut Pomerons have opposite signs, one gets big sums of positive and negative terms, with plenty of interference and cancellations. In Fig. 8, I show the sum of all possible cuts of a two-Pomeron diagram, not considering remnants for simplicity. The uncut Pomerons represent elastic scatterings; they are integrated out. So the expression for the cut multi-Pomeron diagram, Eq. (10), is finally a product of “ G ” terms (and in addition the vertex terms V), so G is the fundamental building block of the approach.

Let me consider a simple example of a realization of a Pomeron (the real one is much more complicated), namely a simple parton ladder with two gluon and two quark ladder rungs; see Fig. 9. The cut is represented by a vertical red

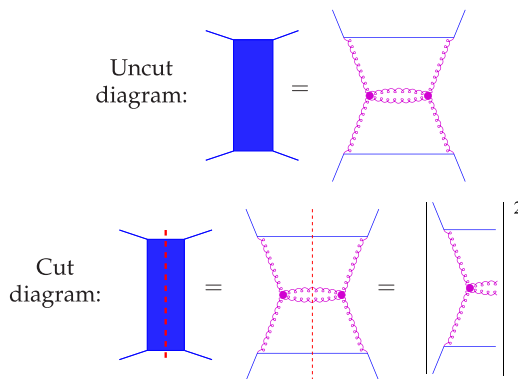


FIG. 9. A simple example of an uncut and the corresponding cut diagram.

dashed line. For a cut diagram, the Feynman rules are modified in the sense that all elements to the left of the cut are treated normally, for all elements to the right one takes the complex conjugate of the normal result, and all propagators crossing the cut line are replaced by a mass shell condition $\delta(p^2 - m^2)$. The cut diagram corresponds therefore to an inelastic amplitude squared, with all particles on the cut line being final on-shell particles. This is true not only for this simple example but always. So the notion of cut diagrams is very useful, in particular for multiple scattering scenarios.

Concerning nuclear scatterings, the total cross section is still given by Eq. (9), together with Eqs. (6)–(8), which gives for a collision of two nuclei with mass numbers A and B

$$\sigma_{\text{inel}} = \sum_{n_1, \dots, n_{AB}} \int dX \prod_{k=1}^{AB} \frac{1}{n_k!} \times W_{AB}(\{x_{\text{remn } i}^+\}, \{x_{\text{remn } j}^-\}) \prod_{k=1}^{AB} \prod_{v=1}^{n_k} G(x_{kv}^+, x_{kv}^-, s, b_k), \quad (12)$$

with at least one n_k being nonzero. Here, W_{AB} contains all the vertices and the integration over uncut Pomerons, with the symbol $\int dX$ explicitly given as $\int d^2b \int dT_{AB} \int dX_{AB}$, with

$$\int dT_{AB} = \int \prod_{i=1}^A d^2b_i^A T_A(b_i^A) \prod_{j=1}^B d^2b_j^B T_B(b_j^B), \quad (13)$$

representing the nuclear geometry, with the nuclear thickness functions $T_A(b)$ given as $\int dz \rho_A(\sqrt{b^2 + z^2})$, with ρ_A being the nuclear density, and with

$$\int dX_{AB} = \int \prod_{k=1}^{AB} \prod_{v=1}^{n_k} dx_{kv}^+ dx_{kv}^-. \quad (14)$$

The impact parameter is defined as $b_k = |\vec{b} + \vec{b}_{\pi(k)}^A - \vec{b}_{\tau(k)}^B|$, where $\pi(k)$ and $\tau(k)$ refer to the projectile and the target nucleons corresponding to pair k . There is no “ δ ” term, since here the remnant momentum fractions are no independent variables; they are expressed in terms of the momentum fractions x_{kv}^\pm as

$$x_{\text{remn}, i}^+ = 1 - \sum_{\substack{k, v \\ \pi(k)=i}} x_{kv}^+, \quad x_{\text{remn}, j}^- = 1 - \sum_{\substack{k, v \\ \tau(k)=j}} x_{kv}^-. \quad (15)$$

For completeness, and since it is needed in Sec. VI, let me note that W_{AB} in Eq. (12) can be written as [36]

$$W_{AB} = \prod_{i=1}^A V(x_{\text{remn } i}^+) \prod_{j=1}^B V(x_{\text{remn } j}^-) \times \prod_{k=1}^{AB} \exp[-\tilde{G}(x_{\text{remn } \pi(k)}^+ x_{\text{remn } \tau(k)}^-)], \quad (16)$$

with some known (simple) function \tilde{G} .

Let me close this section with some technical remarks, concerning the impact parameter dependence and the energy

dependence of the T-matrices and of the G functions:

- (i) As discussed in [35], the (Mandelstam) t dependence of the original T-matrices is given (in all cases) as factors of the form $\exp(R^2 t)$, with parameters R^2 . Considering purely transverse momentum exchange, one has $t = -k_\perp^2$, and the two-dimensional Fourier transform with respect to the transverse momentum exchange \vec{k}_\perp gives a factor $\exp[-b^2/(4R^2)]$.
- (ii) In this paper, all the “ G ” and “ T ” expressions refer to “impact parameter representation”, so all b dependencies are simply Gaussian factors. Here, I do not specify the precise structure of the Pomeron; this is done in very much detail in [35], where is shown that the “real” Pomerons are convolutions of several parts (soft preevolution, hard part), but the b dependencies are always Gaussians, giving always a final b dependence of the form of a factor $\exp[-b^2/(4R^2)]$. So the b dependencies are trivial, and easy to handle. In the following, I will not write the b dependencies explicitly.
- (iii) All T-matrices and G functions depend on s , with s referring in all cases to the nucleon-nucleon center-of-mass squared energy, because using the explicit arguments x^+ and x^- , the transverse mass of a Pomeron is $x^+ x^- s$. In the following, I will not write this s dependency explicitly.

IV. A SIMPLE CASE: FACTORIZATION AND BINARY SCALING IN A SCENARIO WITHOUT ENERGY CONSERVATION

Before further developing the full EPOS4 S-matrix approach, in order to understand the real importance of energy conservation (or energy sharing), I will discuss in this section the S-matrix approach without energy sharing.

I consider the general situation where the precise structure of Pomeron is not specified. All the diagrams which contribute to cut T (and therefore to the inelastic cross section) in pp represent an infinite series, composed of all possible cut and uncut Pomerons (boxes) as shown in Fig. 10 up to order $n = 3$. However, here energy sharing will be dropped, which is realized by removing the vertices and the δ term in Eq. (2), so one has

$$iT = \sum_{n=1}^{\infty} \int dX \frac{1}{n!} \{iT_{\text{Pom}}^{(1)} \times \dots \times iT_{\text{Pom}}^{(n)}\}. \quad (17)$$

This simplifies things enormously, since with $\int dX = \int dX_1 \dots dX_n$, and defining $\tilde{T}_{\text{Pom}} = \int dX_i T_{\text{Pom}}^{(i)}$ (actually not depending on i), one gets

$$iT = \sum_{n=1}^{\infty} \frac{1}{n!} \{i\tilde{T}_{\text{Pom}}\}^n, \quad (18)$$

which is precisely the expression used in the Gribov-Regge approach. The sub-T-matrix \tilde{T}_{Pom} depends only on s and b , as does the full S-matrix T (in both cases I do not write this dependence, for simplicity). Following Eqs. (9) and (10), including the subsequent discussion, and using Eq. (18), one

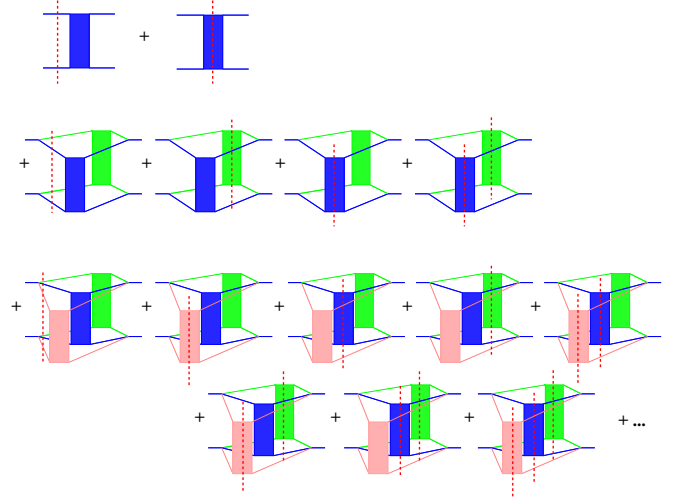


FIG. 10. All the diagrams which contribute to cut T in pp scattering up to order $n = 3$. Red dashed lines refer to cuts.

gets for the inelastic cross section

$$\sigma_{\text{inel}} = \int d^2b \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{\text{cuts}} \{i\tilde{T}_{\text{Pom}}\}^n, \quad (19)$$

with at least one cut. One usually assumes the sub-T-matrix to be purely imaginary, i.e., $\tilde{T}_{\text{Pom}} = i\frac{a}{2}$, with some real number a , and a factor $1/2$ for convenience. Then one gets for the cut Pomeron cut $\tilde{T}_{\text{Pom}} = 2\text{Im} \tilde{T}_{\text{Pom}} = a$. Concerning the uncut Pomerons, one sums up the contributions where the Pomeron is to the left or to the right of the cut, which gives $\{i\tilde{T}_{\text{Pom}}\} + \{i\tilde{T}_{\text{Pom}}\}^* = -a$. So cut and uncut Pomerons have opposite signs, and one gets

$$\sigma_{\text{inel}} = \int d^2b \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{m=1}^n \binom{n}{m} a^m (-a)^{n-m}, \quad (20)$$

where m refers to the number of cut Pomerons.

Let me consider inclusive cross sections, like jet cross sections, where m -cut-Pomeron events contribute m times more than single Pomeron events, so one gets

$$\sigma_{\text{incl}} = \int d^2b \sum_{n=1}^{\infty} \frac{1}{n!} \left\{ \sum_{m=1}^n m \binom{n}{m} a^m (-a)^{n-m} \right\}, \quad (21)$$

where the term in curly brackets represents the sum over all cuts.

For a given number n of Pomerons, an elementary calculation allows to compute the sum over all possible cuts, and one finds an amazing result,

$$\sum_{m=1}^n m \binom{n}{m} a^m (-a)^{n-m} = a^n \times \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{if } n > 1, \end{cases} \quad (22)$$

known as AGK cancellations [11]:

- (i) For a given number $n > 1$ of Pomerons, the sum of all cuts gives zero, i.e., one gets a complete cancellation.
- (ii) Only $n = 1$ contributes, which corresponds to the case of a single Pomeron.

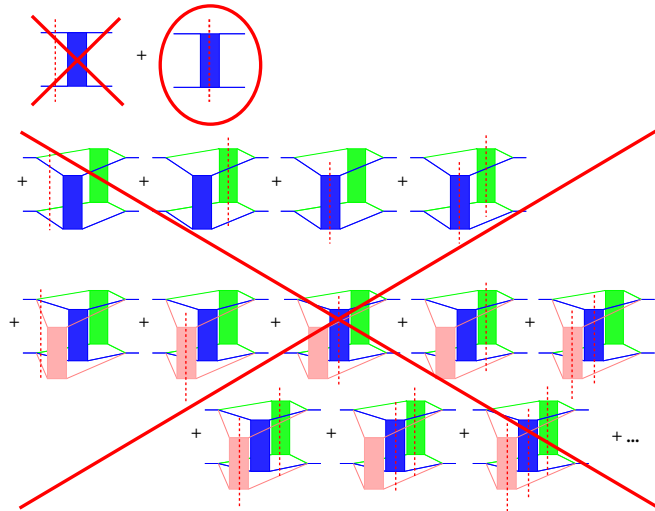


FIG. 11. Cancellations for inclusive cross sections, when energy-momentum conservation is ignored.

Therefore, for inclusive cross sections and only for those, only the single Pomeron events contribute, as indicated in Fig. 11.

If a Pomeron is considered to be a parton ladder, a single cut Pomeron looks like the graph in Fig. 12(a), with parton evolutions from both sides and a hard elementary parton-parton scattering in the middle, the corresponding inelastic process is shown in Fig. 12(b). So one can write the inclusive pp cross section as a convolution of two parton distribution functions (PDFs) and a (calculable) elementary parton-parton cross section (Born process),

$$\sigma_{\text{incl}} = \int d^2b \int dx^+ dx^- f_{\text{PDF}}(x^+) f_{\text{PDF}}(x^-) \sigma_{\text{QCD}}(x^+ x^- s), \quad (23)$$

which amounts to factorization. So factorization here is the result of a huge amount of cancellations.

For completeness, also for inclusive cross sections in AA scattering one observes this phenomenon of cancellations, such that finally only a single Pomeron contributes. Colliding two nuclei with mass numbers A and B , the cross section turns out to be AB times the proton-proton cross section (see for example [15]), which is nothing but “binary scaling”.

To summarize this section:

- (i) in a simplified picture, dropping energy conservation, one gets (easily) factorization for inclusive cross sections in pp scattering

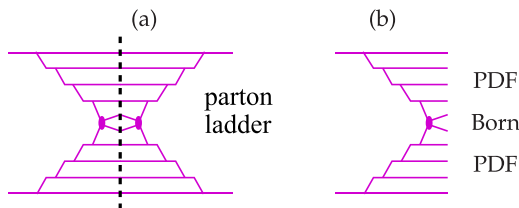


FIG. 12. A single cut Pomeron (a) and the corresponding inelastic process (b).

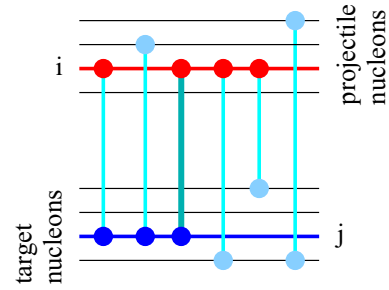


FIG. 13. A Pomeron connected to projectile nucleon i and target nucleon j , together with other Pomerons connected to one (or both) of these nucleons.

- (ii) and binary scaling for inclusive cross sections in AA scattering.

But dropping energy-momentum conservation is not really an acceptable solution, in particular since the Monte Carlo procedures eventually need the implementation of energy conservation, so one risks to introduce inconsistencies, in the sense that the theoretical basis and the Monte Carlo realization are not compatible.

In EPOS4, I insist on the compatibility of the theoretical basis and the Monte Carlo realization (this is an evidence, but widely ignored), so one must include energy conservation in the formulas representing the theoretical basis (S-matrix theory). In Secs. V and VI I will discuss why energy conservation spoils factorization and binary scaling, and in Sec. VII I discuss the solution of the problem, which depends strongly on saturation and which leads to “generalized AGK cancellations”.

V. HOW ENERGY SHARING DEFORMS POMERON ENERGY DISTRIBUTIONS

In this section, I come back to the full EPOS4 S-matrix formalism, including energy sharing, and try to understand why and how energy sharing affects Pomeron energy distributions, which will be crucial with respect to factorization and binary scaling.

Let me consider a particular Pomeron in AA collisions (including pp as a special case), connected to projectile nucleon i and target nucleon j ; see Fig. 13. There might be other Pomerons, connected to one (or both) of these nucleons. The corresponding Pomeron-nucleon connections are marked as red and blue dots. It is obvious that the additional Pomerons connected to the same nucleons i and j compete with each other: they have to share the initial energy-momentum of the two nucleons. The more Pomerons are connected, the less energy is available for one particular Pomeron.

To quantify this statement, I define the “connection number”

$$N_{\text{conn}} = \frac{N_p + N_T}{2}, \quad (24)$$

with N_p being the number of Pomerons connected to i , and with N_T being the number of Pomerons connected to j (the variable N_{conn} corresponds to half of the number of red and

blue points in Fig. 13). In the following, I will discuss the effect of energy sharing related to the connection number.

As discussed in Sec. III, the fundamental “building block” in EPOS4 is the cut single Pomeron expression $G = \text{cut } T_{\text{Pom}}$. As shown in Eq. (12), the inelastic cross section for the collision of two nuclei with mass numbers A and B is given in terms of expressions

$$W_{AB}(\{x_{\text{remn } i}^+\}, \{x_{\text{remn } j}^-\}) \prod_{k=1}^{AB} \left[\frac{1}{n_k!} \prod_{v=1}^{n_k} G(x_{kv}^+, x_{kv}^-) \right], \quad (25)$$

which represent particular configurations $\{n_k\}$ characterized by n_k cut Pomerons per nucleon-nucleon pair k (with k between 1 and AB). The indices k, v refer to the v th Pomeron associated with the pair k . Due to energy-momentum conservation, $1 - x_{\text{remn } i}^+$ is equal to the sum of all x_{kv}^+ with k connected to projectile i , and $1 - x_{\text{remn } j}^-$ is equal to the sum of all x_{kv}^- with k connected to target j . The expression Eq. (25) represents a multidimensional probability distributions for the light-cone momentum fractions x_{kv}^\pm for a given configuration $\{n_k\}$.

Let me consider, for a given configuration $\{n_k\}$, a particular Pomeron, which means a particular pair index k and a particular value v , with the associated momentum fractions x_{kv}^\pm . Let i and j be the projectile and target the Pomeron is connected to (see Fig. 13).

In the simplest case, one has $n_k = 1$ (only one Pomeron associated to pair k), and one has no other Pomeron connected to i and j , so one has $N_{\text{conn}} = 1$, the case of an isolated Pomeron. Using the known form of W_{AB} given in Eq. (16), one can see that the integration of Eq. (25) over all variables other than x_{kv}^\pm gives, up to a constant, the expression

$$W_{11}(x_{\text{remn } i}^+, x_{\text{remn } j}^-) G(x_{kv}^+, x_{kv}^-). \quad (26)$$

This is the probability distribution of x_{kv}^\pm for the case $N_{\text{conn}} = 1$, so I name it $f^{(1)}(x^+, x^-)$, using simply x^\pm instead of x_{kv}^\pm . Using the energy-moment conservation relations, one gets for the $N_{\text{conn}} = 1$ probability distribution

$$f^{(1)}(x^+, x^-) = W'_{11}(1 - x^+, 1 - x^-) G(x^+, x^-), \quad (27)$$

with $W'_{AB} = c W_{AB}$ with a normalization constant c . Since there may be more than one case with $N_{\text{conn}} = 1$, one averages over them, and one averages over all configurations $\{n_k\}$. But this does not change anything since they all have the same form, as given in Eq. (27).

This two-dimensional distribution Eq. (27) allows one to compute the distribution

$$f^{(1)}(x_{\text{PE}}) = \int dy_{\text{PE}} J \times f^{(1)}(x^+, x^-), \quad (28)$$

being the probability distribution with respect to the “Pomeron energy fraction”,

$$x_{\text{PE}} = x^+ x^- = \frac{M_{\text{Pom}}^2}{s}, \quad (29)$$

with M_{Pom} being the transverse mass of the Pomeron, with $y_{\text{PE}} = 0.5 \ln \frac{x^+}{x^-}$, and with J being the corresponding Jacobian determinant. Having very narrow y_{PE} distributions, one may

use $y_{\text{PE}} \approx 0$, $x^\pm = \sqrt{x_{\text{PE}}} e^{\pm y_{\text{PE}}} \approx \sqrt{x_{\text{PE}}}$, and one gets

$$f^{(1)}(x_{\text{PE}}) \propto W'_{11}(1 - x^+, 1 - x^-) G(x^+, x^-). \quad (30)$$

Let me now consider a more complicated situation, corresponding to $N_{\text{conn}} > 1$. The general formula for the probability distribution $f^{(N_{\text{conn}})}(x^+, x^-)$ is given as

$$f^{(N_{\text{conn}})}(x^+, x^-) = \sum_{k'=1}^{AB} \sum_{\{n_k\} \neq 0} \sum_{v'=1}^{n_{k'}} \delta_{N_{\text{conn}}}^{N_{\text{conn}}(k', v')} \times \int dX \{P(K) \delta(x^+ - x_{k'v'}^+) \delta(x^- - x_{k'v'}^-)\}, \quad (31)$$

where δ_a^b is the Kronecker delta, and where $P(K)$ is the expression Eq. (25) for particular multi-Pomeron configurations $K = \{\{n_k\}, \{x_{kv}^\pm\}\}$, with given energy-momentum sharing. The symbol $\sum_{\{n_k\} \neq 0}$ means summing over all possible choices of n_1, n_2, \dots, n_{AB} , excluding the case where all n_k are zero. The symbol $\int dX$ is explicitly given as $\int d^2b \int dT_{AB} \int dX_{AB}$, with $\int dT_{AB}$ and $\int dX_{AB}$ defined in Eqs. (13) and (14). I only consider Pomerons (k', v') with connection number $N_{\text{conn}}(k', v')$ equal to N_{conn} . In principle one does the same procedure as for the case $N_{\text{conn}} = 1$; namely, for a given configuration $\{n_k\}$, one chooses a particular pair index k' and a particular value v' , with the associated momentum fractions $x_{k'v'}^\pm$, which are replaced by x^\pm after integrating over these variables, because of the delta functions. Then one integrates Eq. (31) over all variables other than $x_{k'v'}^\pm$, which is always possible. But in this case, these integration variables and the “chosen variables” $x_{k'v'}^\pm$ (now x^\pm) can no longer be separated, and one needs to do the integration numerically (via Monte Carlo, in practice). Nevertheless, it is well defined, and one gets the x^\pm distribution $f^{(N_{\text{conn}})}(x^+, x^-)$. As for $N_{\text{conn}} = 1$, one integrates over y_{PE} , to get $f^{(N_{\text{conn}})}(x_{\text{PE}})$. In practice, one defines event classes (according to multiplicity or impact parameter) and computes the average N_{conn} values as well as the average x^\pm distributions per class, to finally get $f^{(N_{\text{conn}})}(x_{\text{PE}})$. Then one takes the obtained distributions as the basis to compute $f^{(N_{\text{conn}})}(x_{\text{PE}})$ for arbitrary values of N_{conn} via interpolation.

The distribution $f^{(N_{\text{conn}})}(x_{\text{PE}})$ or the two-dimensional distribution $f^{(N_{\text{conn}})}(x^+, x^-)$ are kind of “master” distributions for all kinds of “inclusive distributions”, for example the inclusive p_t distribution of partons, or of hadrons if one adds a fragmentation function. For computing just inclusive spectra, the knowledge of f is enough, whereas otherwise the full calculations are needed, using Monte Carlo methods based on Markov chains. In addition to $f^{(N_{\text{conn}})}(x_{\text{PE}})$ I also check the inclusive distribution of y_{PE} , which is narrow and strongly peaked at zero (note that the x^\pm refer to Pomeron momenta, not to those of outgoing partons). This is why I concentrate in the following on $f^{(N_{\text{conn}})}(x_{\text{PE}})$.

Although $f^{(N_{\text{conn}})}(x_{\text{PE}})$ for $N_{\text{conn}} > 1$ cannot be calculated analytically, one has some idea of how it should look compared to $f^{(1)}(x_{\text{PE}})$: the integrand of Eq. (31) contains in addition to $G(x_{k'v'}^+, x_{k'v'}^-)$ other G terms, and most importantly $W_{AB}(\{x_{\text{remn } i}^+\}, \{x_{\text{remn } j}^-\})$ contains factors of the form $(x_{\text{remn } i}^\pm)^\alpha$

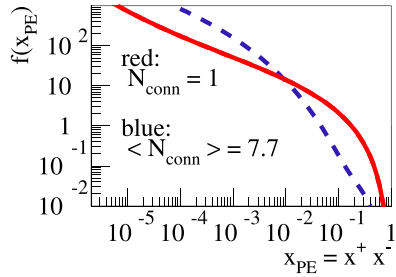


FIG. 14. Comparing $f^{(N_{\text{conn}})}(x_{\text{PE}})$ for the centrality class 0–5% with $f^{(1)}(x_{\text{PE}})$, in PbPb collisions at 5.02 TeV.

with $\alpha > 0$, with arguments

$$x_{\text{remn } i/j}^{\pm} = 1 - x_{k'v'}^{\pm} - \sum'' x_{k''v''}^{\pm}, \quad (32)$$

where \sum'' sums over all indices different from k', v' being connected to i or j . Due to the additional term $\sum'' x_{k''v''}^{\pm}$, not present for $N_{\text{conn}} = 1$, one gets, compared to $N_{\text{conn}} = 1$, a suppression of large values of x_{kv}^{\pm} . This is also what one expects without any calculation: energy sharing involving more than one Pomerons leads to a reduction of the energy of the Pomerons, compared to the case of an isolated Pomeron.

But let me be quantitative, and discuss the real calculations. In Fig. 14, I plot $f^{(1)}(x_{\text{PE}})$ and $f^{(N_{\text{conn}})}(x_{\text{PE}})$ for the centrality class 0–5% in PbPb collisions at 5.02 TeV with an average value of N_{conn} of roughly 7.7, obtained after a full EPOS4 simulation. One observes (as expected) for $N_{\text{conn}} > 1$ a “deformation” of the x_{PE} distributions compared to $f^{(1)}(x_{\text{PE}})$, due to energy-momentum conservation. Therefore I define the ratio

$$R_{\text{deform}} = R_{\text{deform}}^{(N_{\text{conn}})}(x_{\text{PE}}) = \frac{f^{(N_{\text{conn}})}(x_{\text{PE}})}{f^{(1)}(x_{\text{PE}})}, \quad (33)$$

called the “deformation function”. In Fig. 15, I show the deformation function for the centrality class 0–5% in PbPb collisions at 5.02 TeV. The functional form is as one expects: large x_{PE} values (close to unity) are suppressed, and small x_{PE} values are increased. And the effect is big: the suppression in the interval $[0.1, 1]$ is roughly $1/20$.

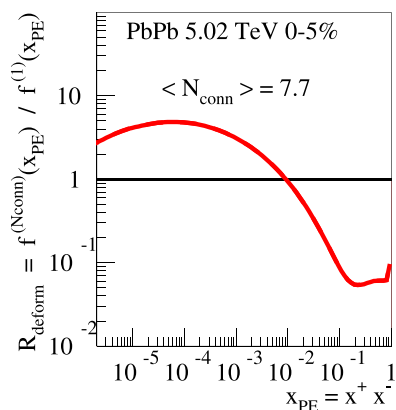


FIG. 15. Deformation function (see text) for the centrality class 0–5% in PbPb collisions at 5.02 TeV.

Let me quickly summarize the main results of this section:

- (i) Imposing energy sharing (as it should be) has a very important impact on the distribution of Pomeron energies.
- (ii) A useful variable to quantify the effect of energy sharing is the connection number N_{conn} , counting the number of “other Pomerons” connected to the same remnants as a “given Pomeron”. $N_{\text{conn}} = 1$ represents an isolated Pomeron.
- (iii) I define a variable $x_{\text{PE}} = x^+ x^-$, representing the squared energy per Pomeron, and the corresponding probability distribution.
- (iv) The probability distribution depends strongly on N_{conn} , so I use the notation $f^{(N_{\text{conn}})}(x_{\text{PE}})$. The results for $N_{\text{conn}} > 1$ show a suppression at large x_{PE} , as a consequence of energy sharing. This must be so; it is unavoidable, a fundamental feature.
- (v) I therefore define a “deformation function” R_{deform} as the ratio of $f^{(N_{\text{conn}})}(x_{\text{PE}})$ over $f^{(1)}(x_{\text{PE}})$, which drops below unity for large x_{PE} .

VI. HOW DEFORMED POMERON ENERGY DISTRIBUTIONS SPOIL FACTORIZATION AND BINARY SCALING IN CASE OF A “NAIVE” POMERON DEFINITION

In this section, the aim is to understand why and how energy sharing ruins factorization and binary scaling. I showed in the last section that energy sharing leads unavoidably to a “deformation” of the Pomeron energy distribution $f^{(N_{\text{conn}})}(x_{\text{PE}})$ compared to the reference $f^{(1)}(x_{\text{PE}})$, with N_{conn} being the connection number, counting the number of other Pomerons connected to the same remnants as a given Pomeron, which leads to the definition of a “deformation function” R_{deform} as the ratio of $f^{(N_{\text{conn}})}(x_{\text{PE}})$ over $f^{(1)}(x_{\text{PE}})$. In the following, I will show how this deformation spoils factorization.

Actually “the problem” related to factorization and binary scaling depends very much on the precise definition of G in terms of QCD. So far a Pomeron is a black box with all the QCD details hidden inside, but now one needs to be more specific. To do so, one introduces in [35] (with many details and all necessary formulas) a “QCD expression” called G_{QCD} representing a QCD calculation of parton-parton scattering. Every “ G function” which one uses, including G_{QCD} , is meant to be the cut of the Fourier transform of the T -matrix, divided by $2s$. The term G_{QCD} is a sum of several contributions, the most important one being the “sea-sea” contribution $G_{\text{QCD}}^{\text{sea-sea}}$; see Fig. 16. For a precise definition see [35]. The vertices $F_{\text{sea } 1}^i$ and $F_{\text{sea } 2}^j$ couple the parton ladder to the projectile and target nucleons. In addition, one has three blocks: the two soft blocks and in between a parton ladder, the latter being a DGLAP parton evolution from both sides, with a pQCD Born process in the middle. I define parton evolution functions E_{QCD} obeying the usual DGLAP equations, but in this case the evolution starts from a parton, not from a nucleon, since a Pomeron corresponds to parton-parton scattering. I compute and tabulate E_{QCD} , and then the convolution

$$E_{\text{soft}} \otimes E_{\text{QCD}} \otimes \text{Born} \otimes E_{\text{QCD}} \otimes E_{\text{soft}}, \quad (34)$$

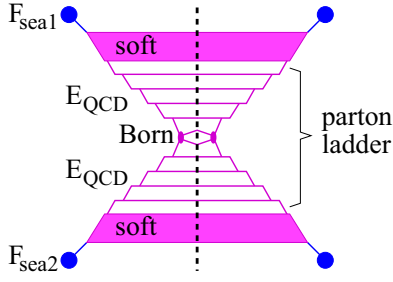


FIG. 16. The contribution $G_{\text{QCD}}^{\text{sea-sea}}$, which is the convolution $E_{\text{soft}} \otimes E_{\text{QCD}} \otimes \text{Born} \otimes E_{\text{QCD}} \otimes E_{\text{soft}}$.

with an elementary QCD cross section ‘‘Born’’ and a soft preevolution E_{soft} . In addition to ‘‘sea-sea’’, one has more contributions, named ‘‘val-val’’, ‘‘sea-val’’, ‘‘val-sea’’, ‘‘soft’’, and ‘‘psoft’’, as discussed in great detail in [35].

Like G , the QCD expression G_{QCD} depends as well on x^+ and on x^- , and in addition there is the crucial parameter Q_0^2 , which is the low virtuality cutoff in the DGLAP evolution, so I use the notation $G_{\text{QCD}}(Q_0^2, x^+, x^-)$. Whereas the cutoff is usually a constant of the order of 1 GeV², I consider it as a variable that may take any value, and I compute and tabulate $G_{\text{QCD}}(Q^2, x^+, x^-)$ for large ranges of discretized values of all arguments, such that G_{QCD} can be computed via interpolation for any choice of arguments. After these preparations, the functional form of $G_{\text{QCD}}(Q^2, x^+, x^-)$ is known. Actually, G and G_{QCD} also depend on s and b , not written explicitly as discussed earlier, so one should always consider ‘‘for given s and b ’’.

What is the relation between G (the Pomeron, the main building block of the multiple scattering theory) and G_{QCD} (which contains all the QCD part)? A first attempt might be (and this is what was actually used in [15]) to consider the two to be equal, i.e.,

$$G(x^+, x^-) = G_{\text{QCD}}(Q_0^2, x^+, x^-), \quad (35)$$

with a constant Q_0^2 . Then one gets for the Pomeron energy distribution for an isolated Pomeron corresponding to $N_{\text{conn}} = 1$ [see Eq. (30)]

$$f^{(1)}(x_{\text{PE}}) \propto W'_{11}(1 - x^+, 1 - x^-) G_{\text{QCD}}(Q_0^2, x^+, x^-), \quad (36)$$

with the Pomeron energy variable $x_{\text{PE}} = x^+ x^-$, and using $x^\pm = \sqrt{x_{\text{PE}}}$. In case of $N_{\text{conn}} > 1$, one has [see Eq. (33)]

$$f^{(N_{\text{conn}})}(x_{\text{PE}}) = R_{\text{deform}}^{(N_{\text{conn}})}(x_{\text{PE}}) \times f^{(1)}(x_{\text{PE}}), \quad (37)$$

which means that the x_{PE} distributions will get more and more ‘‘deformed’’, in particular suppressed at large x_{PE} . This is a general feature and is unavoidable, a direct consequence of energy sharing. What does this mean concerning transverse momentum (p_t) distributions of the outgoing particles from the Born process? Here one needs to consider the internal structure of G_{QCD} , first of all $G_{\text{QCD}}^{\text{sea-sea}}$ (actually similar arguments hold for the other contributions). The important element is the parton ladder, see Fig. [35], given as a convolution

$$E_{\text{QCD}} \otimes \text{Born} \otimes E_{\text{QCD}} \quad (38)$$

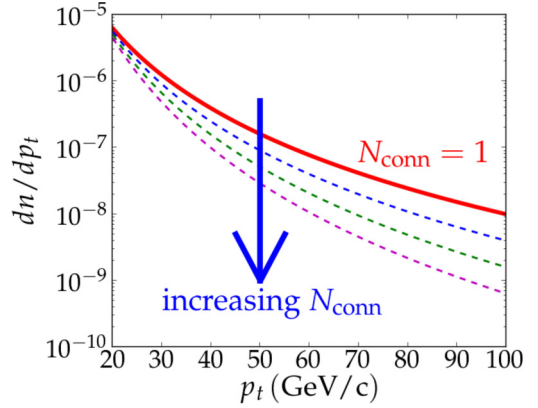


FIG. 17. Sketch of the suppression of parton yields at high p_t with increasing N_{conn} .

(for the formulas, see [35]). The p_t of the outgoing partons is related to the factorization scale μ_F^2 (one uses $\mu_F^2 = p_t^2$), which is the virtuality of the partons entering the Born process. Large values of p_t require large μ_F^2 and large squared energy \hat{s} of the Born process, and this requires a large Pomeron squared energy, and therefore a large value of x_{PE} . The essential points are

- (i) Large values of p_t of the outgoing partons are strongly correlated with large values of x_{PE} .
- (ii) A suppression of large x_{PE} values in $f^{(N_{\text{conn}} > 1)}(x_{\text{PE}})$ compared to $f^{(1)}(x_{\text{PE}})$ will therefore lead to a suppression of large p_t values in the case of $N_{\text{conn}} > 1$ compared to $N_{\text{conn}} = 1$.

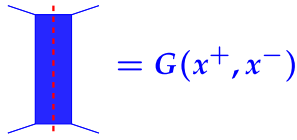
In Fig. 17, I sketch this situation of a suppression of parton yields at high p_t with increasing N_{conn} . Let me first discuss the consequences for pp scattering. In order to get factorization, as discussed in Sec. IV, one would need something like AGK cancellations, such that the full multiple scattering scenario is identical to the single Pomeron ($N_{\text{conn}} = 1$) case, which means eventually

$$\frac{dn^{(pp\text{MB})}}{dp_t} = \frac{dn^{(N_{\text{conn}}=1)}}{dp_t} \quad (39)$$

for the minimum bias (MB) inclusive particle production result. The latter may be written as a superposition of the different contribution for given values of N_{conn} [the latter being (for pp) identical to the number of cut Pomerons] with the corresponding weights $w^{(N_{\text{conn}})}$:

$$\frac{dn^{(pp\text{MB})}}{dp_t} = \sum_{N_{\text{conn}}=1}^{\infty} w^{(N_{\text{conn}})} \frac{dn^{(N_{\text{conn}})}}{dp_t}, \quad (40)$$

where the contributions $\frac{dn^{(N_{\text{conn}})}}{dp_t}$ show with increasing N_{conn} more and more suppression at large p_t (as indicated in Fig. 17). The average Pomeron number at, say, 7 TeV is around 2, so one has definitely an important contribution from terms with $N_{\text{conn}} > 1$. This means that also the MB result will be reduced at high p_t compared to $N_{\text{conn}} = 1$, and this means one cannot fulfill Eq. (39), so factorization is not achieved.

FIG. 18. The cut Pomeron G .

The discussion for the scattering of two nuclei with mass numbers A and B is similar to the pp case. To assure binary scaling, one expects

$$\frac{dn^{(ABMB)}}{dp_t} = AB \times \frac{dn^{(N_{\text{conn}}=1)}}{dp_t}, \quad (41)$$

for the minimum bias (MB) inclusive particle yields. For the contributions for different values of N_{conn} , one has a picture similar to that shown in Fig. 17, and also here one concludes that the MB inclusive yield will be reduced at high p_t compared to $N_{\text{conn}} = 1$, and therefore one cannot fulfill Eq. (41), and therefore binary scaling is violated.

Let me summarize this section:

- (i) I consider here the case where the cut Pomeron G is identical to G_{QCD} , the latter representing a pQCD result for parton-parton scattering.
- (ii) Considering the internal structure of G_{QCD} , one concludes that there is a strong correlation between the Pomeron energy variable x_{PE} and p_t of the outgoing partons (large p_t corresponds to large x_{PE}).
- (iii) Therefore the suppression of large x_{PE} with increasing N_{conn} amounts to a suppression of large p_t , and one can conclude a suppression of yields at large p_t for minimum bias results compared to $N_{\text{conn}} = 1$.
- (iv) This means one cannot obey Eqs. (39) and (41), which are necessary conditions for factorization (pp) and binary scaling (AA).

VII. HOW SATURATION ALLOWS ONE TO RECOVER FACTORIZATION AND BINARY SCALING (GENERALIZED AGK CANCELLATIONS)

In the following, I discuss the “key issue” of the EPOS4 approach, namely the appropriate definition of $G(x^+, x^-)$, the cut Pomeron, represented so far as a “cut box” as shown in Fig. 18, and used earlier (see for example Figs. 8 and 10), to develop the multiple scattering scheme. The latter is actually completely independent from the precise definition of G , which is very useful, so one can investigate different options concerning the internal structure of G .

I showed in the previous section that the “naive” assumption

$$G = G_{\text{QCD}} \quad (42)$$

(which was also adopted in [15] and [39]) completely spoils factorization and binary scaling. And from the discussion in the previous section, it is known that this is a fundamental, unavoidable problem, and not just a wrong parameter choice. So the assumption Eq. (42) seems to be simply wrong.

There is another serious problem with Eq. (42): as discussed somewhat in the previous section (and in detail in

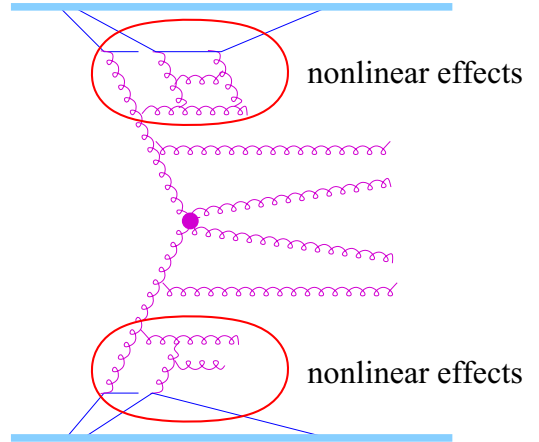


FIG. 19. Nonlinear effects (inside the red ellipses) also referred to as saturation effects are “summarized” in the form of saturation scales, which replace these non-linear parts.

[35]), the essential part of G_{QCD} is a cut parton ladder, based on DGLAP parton evolutions. But as already discussed in the introduction, this is certainly not the full story: with increasing energy, partons with very small momentum fractions $x \ll 1$ become increasingly important, since the parton density becomes large, and therefore the linear DGLAP evolution scheme is no longer valid and nonlinear evolution takes over, considering explicitly gluon-gluon fusion. These phenomena are known as “small x physics” or “saturation” [19–32].

At least for scatterings carrying a large value of x^+x^- , one expects “nonlinear effects”, which means that two ladders which evolve first independently and in parallel finally fuse. And only after that is the (linear) DGLAP evolution realized. Such nonlinear effects lead to strong destructive interference at low transverse momentum (p_t), which may be summarized in terms of a saturation scale [20,21]. This suggests treating these “saturation phenomena” not explicitly, but by introducing a saturation scale as the lower limit of the virtualities for the DGLAP evolutions, as sketched in Fig. 19.

So one has two problems:

- (i) a wrong identity $G = G_{\text{QCD}}$,
- (ii) a missing treatment of saturation.

But fortunately the two problems are connected, and there is an amazingly simple solution that solves both problems. Instead of the “naive” assumption $G = G_{\text{QCD}}$, one postulates

$$G(x^+, x^-) = \frac{n}{R_{\text{deform}}^{(N_{\text{conn}})}(x_{\text{PE}})} G_{\text{QCD}}(Q_{\text{sat}}^2, x^+, x^-), \quad (43)$$

with

$$G \text{ independent of } N_{\text{conn}}. \quad (44)$$

Here, $R_{\text{deform}}^{(N_{\text{conn}})}(x_{\text{PE}})$ is the deformation function discussed in Sec. V, and n is a constant, not depending on x_{PE} . The independence of G on N_{conn} is absolutely crucial (as I will show later), and to achieve this one first parametrizes G based on the (very reasonable) assumption that G has a “Regge-pole structure” as $G \propto \alpha x_{\text{PE}}^\beta$, where the s and b dependences of α and β are parametrized with the parameters being fixed

by comparing simulation results to elementary experimental data, and then use Eq. (43) to determine Q_{sat}^2 . In this way, Q_{sat}^2 depends on N_{conn} and on x^\pm :

$$Q_{\text{sat}}^2 = Q_{\text{sat}}^2(N_{\text{conn}}, x^+, x^-), \quad (45)$$

which means that this Q_{sat}^2 , being the low virtuality cutoff for the DGLAP evolutions in G_{QCD} , is not a constant, but its value depends on the environment in terms of N_{conn} and on the energy of the Pomeron. I will refer to this as ‘‘dynamical saturation scale’’.

But why does Eq. (43) work? One gets for the Pomeron energy distribution for an isolated Pomeron, corresponding to $N_{\text{conn}} = 1$ [see Eq. (30), using $x^\pm = \sqrt{x_{\text{PE}}}$],

$$f^{(1)}(x_{\text{PE}}) \propto W'_{11}(1 - x^+, 1 - x^-) \times G_{\text{QCD}}(Q_{\text{sat}}^2(1, x^+, x^-), x^+, x^-), \quad (46)$$

where Eq. (43) with $N_{\text{conn}} = 1$ was used to replace G . In the case of $N_{\text{conn}} > 1$, one has [see Eq. (33)]

$$f^{(N_{\text{conn}})}(x_{\text{PE}}) = R_{\text{deform}}^{(N_{\text{conn}})}(x_{\text{PE}}) \times f^{(1)}(x_{\text{PE}}) \quad (47)$$

$$\propto R_{\text{deform}}^{(N_{\text{conn}})}(x_{\text{PE}}) \times W'_{11}(1 - x^+, 1 - x^-) \times G(x^+, x^-). \quad (48)$$

Here I will use again Eq. (43) to replace G , but with $N_{\text{conn}} > 1$ such that the $R_{\text{deform}}^{(N_{\text{conn}})}$ expressions cancel, and one gets

$$f^{(N_{\text{conn}})}(x_{\text{PE}}) \propto W'_{11}(1 - x^+, 1 - x^-) \times G_{\text{QCD}}(Q_{\text{sat}}^2(N_{\text{conn}}, x^+, x^-), x^+, x^-). \quad (49)$$

The crucial point here is the fact that, thanks to Eq. (43) and since G does not depend on N_{conn} , the $R_{\text{deform}}^{(N_{\text{conn}})}$ expressions disappear. Comparing now $f^{(N_{\text{conn}})}(x_{\text{PE}})$ and $f^{(1)}(x_{\text{PE}})$, one finds

$$\frac{f^{(N_{\text{conn}})}(x_{\text{PE}})}{f^{(1)}(x_{\text{PE}})} \propto \frac{G_{\text{QCD}}(Q_{\text{sat}}^2(N_{\text{conn}}, x^+, x^-), x^+, x^-)}{G_{\text{QCD}}(Q_{\text{sat}}^2(1, x^+, x^-), x^+, x^-)}. \quad (51)$$

This equation is very interesting: it means that the N_{conn} dependence of x_{PE} distributions is guided by the saturation scale, and nothing else. This is the only difference between the numerator and the denominator. The Eq. (51) also means that the partonic structure is given by G_{QCD} , and therefore also the p_t distribution of the outgoing partons is encoded in the single Pomeron expression G_{QCD} , for any N_{conn} . Only the saturation scales Q_{sat}^2 depend on N_{conn} , and these saturation scales suppress small p_t particle production, but do not affect high p_t results, as sketched in Fig. 20.

What does this mean concerning factorization? The minimum bias (MB) inclusive parton yield may be written as a superposition of the different contributions for given values of N_{conn} (for pp identical to the number of cut Pomerons) with weights $w^{(N_{\text{conn}})}$:

$$\frac{dn^{(pp\text{MB})}}{dp_t} = \sum_{N_{\text{conn}}=1}^{\infty} w^{(N_{\text{conn}})} \frac{dn^{(N_{\text{conn}})}}{dp_t}. \quad (52)$$

At large p_t , all contributions are equal, as indicated in Fig. 20, so one can replace $\frac{dn^{(N_{\text{conn}})}}{dp_t}$ by $\frac{dn^{(1)}}{dp_t}$ and one gets

$$\frac{dn^{(pp\text{MB})}}{dp_t} = \frac{dn^{(N_{\text{conn}}=1)}}{dp_t} \quad (\text{for large } p_t), \quad (53)$$

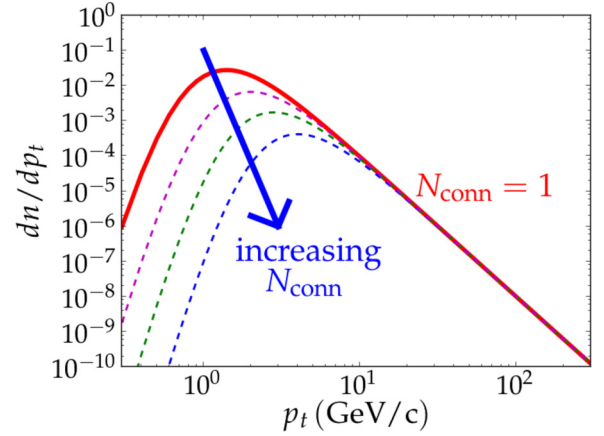


FIG. 20. Sketch of the suppression of low p_t partons with increasing N_{conn} .

so only a single Pomeron contributes. This will allow one to define parton distribution functions f_{PDF} and to compute cross sections as convolutions $f_{\text{PDF}} \otimes \text{Born} \otimes f_{\text{PDF}}$.

The discussion for the scattering of two nuclei with mass numbers A and B is similar to the pp case: the Pomeron connections do not affect high p_t , as shown in Fig. 20. The only difference compared to pp scattering is the fact that one has to sum over all possible nucleon-nucleon pairs, which gives

$$\frac{dn^{(AB\text{MB})}}{dp_t} = AB \times \frac{dn^{(N_{\text{conn}}=1)}}{dp_t} \quad (\text{for large } p_t), \quad (54)$$

for the minimum bias (MB) inclusive particle yields, which amounts to binary scaling.

Equations (53) and (54) state the following:

- (i) The inclusive pp cross section is equal to the one of a single Pomeron contribution, and the inclusive cross section of the scattering of two nuclei of mass numbers A and B is equal to AB times the single Pomeron contribution, leading to factorization and binary scaling.
- (ii) I refer to this as the ‘‘generalized AGK theorem’’, valid at high p_t , in a scenario with energy sharing.

One recalls that the classical AGK cancellations [11] are based on a scenario without energy sharing, as discussed in Sec. IV.

Let me summarize this section:

- (i) One tries to find the relation between G (the multiple scattering building block) and G_{QCD} which represents a QCD result concerning single parton-parton scattering.
- (ii) Two problems are identified: (1) the naive expectation $G = G_{\text{QCD}}$, having been used so far, does not work, and (2) an appropriate treatment of saturation is missing.
- (iii) Both problems are solved by postulating $G(x^\pm) \propto G_{\text{QCD}}(Q_{\text{sat}}^2, x^\pm)/R_{\text{deform}}$, which means that a saturation scale, depending on the Pomeron connection number N_{conn} and on x^\pm , replaces the virtuality cutoff

Q_0^2 usually used in DGLAP evolutions. In this way one incorporates saturation.

- (iv) A direct consequence of the above postulate is the fact that the Pomeron energy distribution $f^{(N_{\text{conn}})}(x_{\text{PE}})$ is for any value of N_{conn} given in terms of a single Pomeron expression G_{QCD} , with an only implicit N_{conn} dependence via Q_{sat}^2 .
- (v) As a consequence, N_{conn} affects low p_t (suppression) but not high p_t , and one recovers factorization and binary scaling (generalized AGK theorem).

As a final remark, within a rigorous parallel scattering scenario (which seems mandatory), and respecting energy conservation (which seems mandatory as well), the only way to not get in contradiction with factorization and binary scaling seems to be the consideration of saturation via $G = k \times G_{\text{QCD}}(Q_{\text{sat}}^2)$ with k being inversely proportional to the deformation function. In this sense, parallel scattering, energy conservation, saturation, and factorization are deeply connected.

VIII. REMARKS CONCERNING DEFORMATION FUNCTIONS AND SATURATION SCALES FOR GIVEN EVENT CLASSES

Let me come back to the deformation function $R_{\text{deform}}^{(N_{\text{conn}})}(x_{\text{PE}})$, which plays a fundamental role in the new approach. As explained earlier, this function can be computed based on Monte Carlo simulations. But to do so, one first needs to define the Pomerons. This is done using a parametrization of G in ‘‘Regge pole form’’ αs^β , and, based on this, one computes the deformation functions. Then one uses Eq. (43) to do full simulations, and compare with data. If needed, the initial parametrization of G , and as a consequence also the deformation functions, are changed, and one repeats the procedure. In practice, I have found a very simple functional form for the deformation functions which accommodates all systems, centrality classes, and energies [36]. I determine and tabulate the parameters, and then use parametrized deformation functions. In Fig. 21, I show as an example the deformation function for the centrality class 0–5% in PbPb collisions at 5.02 TeV, with an average value of N_{conn} of roughly 7.7. I plot the parametrized function (black dots) and the ‘‘computed distribution’’ which comes out after a simulation (red curve). The two curves agree.

One recalls that, in the above iterative procedure, Eq. (43) is used based on a G already known and parametrized in ‘‘Regge pole form’’ αs^β . As a historical side remark, including saturation effects via a Regge pole form αs^β was introduced first in EPOS1 [40], where the term ‘‘parton ladder splitting’’ was used rather than ‘‘saturation’’, but it refers to the same phenomenon. A dynamical saturation scale, assuming some functional form of $Q_{\text{sat}}^2(N_{\text{conn}}, x_{\text{PE}})$, was introduced first in EPOS3 [41] for proton-lead scattering, where also for the first time the expression ‘‘saturation’’ was used. Also in [41], real simulation results show (Fig. 3) the suppression of parton yields at high p_t , in case of the ‘‘naive’’ assumption $G = G_{\text{QCD}}$, and it is shown (Fig. 4) that the suppression can be avoided by introducing a saturation scale. Finally in [42] it was proposed

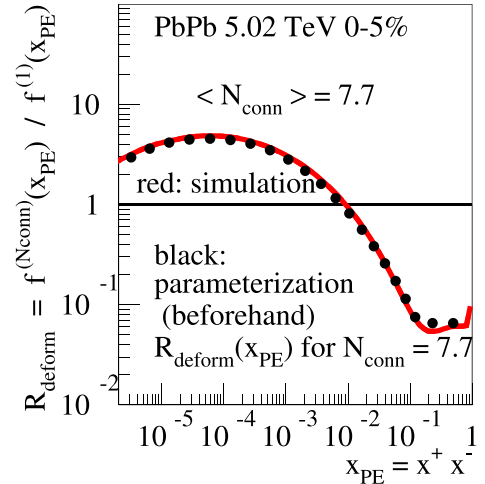


FIG. 21. Deformation function (see text) for the centrality class 0–5% in PbPb collisions at 5.02 TeV.

to use a parametrized G as in [40] and use it to determine $Q_{\text{sat}}^2(N_{\text{conn}}, x_{\text{PE}})$. However, at the time, the role of the deformation was not yet understood, but this is crucial to ensure the correct asymptotic behavior at large x_{PE} and large p_t .

As was done for the case of central PbPb in Fig. 21, one computes the deformation functions always for event classes and associates the obtained function to the mean N_{conn} for the corresponding class. There are several ways to define event classes; one possibility is to do it based on the number of cut Pomerons (or simply the Pomeron number) N_{Pom} , which is related to the multiplicity $dn/d\eta(0)$. I consider simulations for pp at 7 TeV and PbPb at 2.76 TeV (because I will later come back to these two systems to compare simulation results with data). In Fig. 22, I show the multiplicity $dn/d\eta(0)$ as a function of the Pomeron number N_{Pom} , for pp at 7 TeV (red line) and PbPb at 2.76 TeV (blue line), together with the dotted line representing the function $2.9(N_{\text{Pom}})^{0.9}$, which provides a simple conversion formula between these two quantities. This might be useful when I analyze later observables as a function

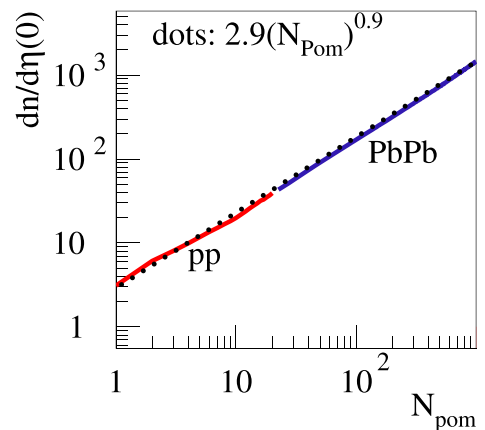


FIG. 22. The multiplicity $dn/d\eta(0)$ as a function of the Pomeron number N_{Pom} for pp (red line) and PbPb (blue line), together with the dotted line representing the function $2.9(N_{\text{Pom}})^{0.9}$.

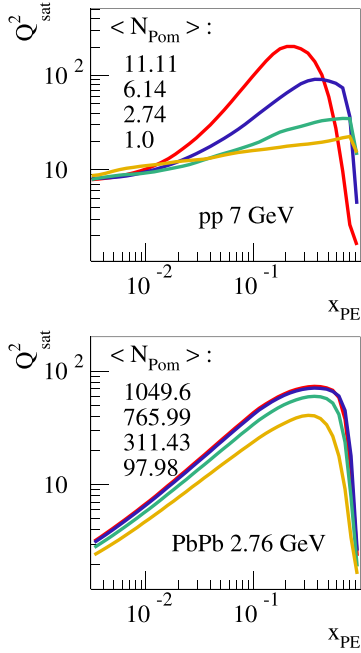


FIG. 23. The saturation scale Q_{sat}^2 as a function of x_{PE} , for several N_{Pom} event classes.

of $\langle dn/d\eta(0) \rangle$. One gets a continuous curve when going from pp to PbPb.

In Fig. 23, I plot the saturation scale Q_{sat}^2 as a function of x_{PE} , for several N_{Pom} event classes. The most striking result is the fact that in pp the Q_{sat}^2 values change very strongly with N_{Pom} , whereas for PbPb the variation is quite moderate, and towards central collisions Q_{sat}^2 even “saturates” (no variation anymore). This discussion will be important to understand the results in Sec. X.

As a first result of EPOS4 simulations, I am going to show that binary scaling really works in practice. In Fig. 24, I

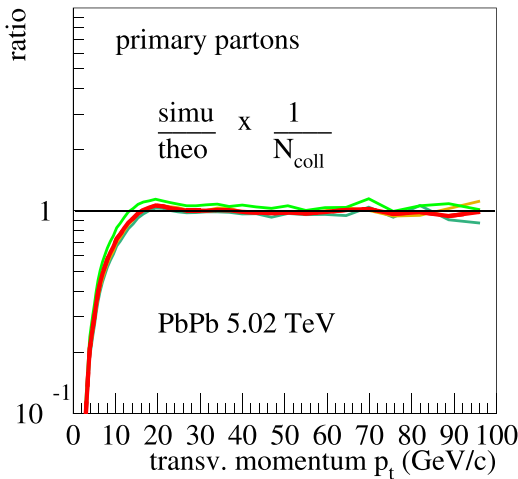


FIG. 24. The inclusive p_t distribution of partons for a full simulation (simu) divided by N_{coll} and the “reference curve” (theo), which is the corresponding distribution for a single Pomeron, calculated analytically. I show results for minimum bias PbPb collisions at 5.02 TeV (red curve) as well as results for different centrality classes.

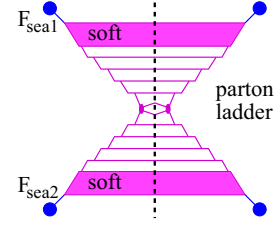


FIG. 25. Contribution $G_{\text{QCD}}^{\text{sea-sea}}$ to the cut single Pomeron.

show the inclusive p_t distribution of partons for a full simulation (simu) divided by N_{coll} and the “reference curve” (theo), which is the corresponding distribution for a single Pomeron, calculated analytically. I show results for minimum bias PbPb collisions at 5.02 TeV (red curve) as well as results for different centrality classes. One can see that the ratio is close to one for large values of p_t , whereas low p_t values are suppressed. Also in pp , the full simulation over the “reference curve” (single Pomeron) is close to unity at large p_t and, comparing pp and AA, one gets $R_{AA} \approx 1$.

An iterative procedure is employed that relies very much on experimental data: One starts with a parametrization of G in “Regge pole form”, already constrained by basic experimental data like energy dependence of cross sections. Then one computes the deformation functions, which allows finally to determine saturation scales, which allows in this way including saturation effects. Based on G_{QCD} , with a (now) known saturation scale, one can generate partons, and then make very detailed comparisons with all kinds of data, and if needed redo the procedure with an improved parametrization of G . So to some extent one has a data-driven method to obtain saturation scales, based on a fully self-consistent pQCD based multiple scattering framework, which is complementary to efforts of computing saturation scales.

IX. EPOS4 FACTORIZATION MODE (SINGLE POMERON) AND EPOS4 PDFS

Since in the case of inclusive spectra at large p_t everything can be reduced to the single Pomeron case (generalized AGK cancellations; see Sec. VII), one may use “a shortcut” and compute inclusive particle production simply by using a single Pomeron, without any need to use complicated Monte Carlo procedures. This is referred to as “EPOS4 factorization mode”. In this case, one simply needs to evaluate the cut single Pomeron, corresponding to G_{QCD} , which is composed of several contributions; see [35]. The most important one is $G_{\text{QCD}}^{\text{sea-sea}}$; see Fig. 25. It is composed of a parton ladder with parton evolutions (E_{QCD}) from both sides and an elementary QCD Born process in the middle. In addition, the QCD parton evolution is preceded by a soft evolution (E_{soft}). The vertices F_{sea1} and F_{sea2} couple the parton ladder to the projectile and target nucleons. The complete expression is a convolution of several elements, $F_{\text{sea1}} \otimes E_{\text{soft}} \otimes E_{\text{QCD}} \otimes \text{Born} \otimes E_{\text{QCD}} \otimes E_{\text{soft}} \otimes F_{\text{sea2}}$, which in addition needs to be convoluted with the vertices V as $V \otimes \dots \otimes V$. This expression may be regrouped in several ways. One possibility is to convolute first the vertices, the soft evolution, and the QCD evolution on the projectile side, representing the parton distribution function

(PDF) of the projectile, and correspondingly on the target side. The two PDFs represent actually the upper and lower parts of the graph in Fig. 25, plus a vertex V , but excluding the Born process.

So far I only consider the so-called “sea-sea” contribution $G_{\text{QCD}}^{\text{sea-sea}}$, with a sea quark (after a soft evolution) being the first parton entering the partonic cascade on both sides. But as shown in [35] there is, in addition, a “val-val” contribution, where valence quarks enter the partonic cascade, and correspondingly “val-sea” and “sea-val” contributions. Since the parton distribution function is just half of the Pomeron diagram, there are two contributions: the “sea” and a “val” one. For a precise definition of the PDFs, see [35].

One computes (and tabulates) the PDFs $f_{\text{PDF}}^k(x, \mu_F^2)$, with x being the light-cone momentum fraction of the parton of flavor k entering the Born process, and μ_F^2 being the factorization scale. After this preparation, one may express the dijet cross section (where dijet simply refers to the two outgoing partons of the Born process) in terms of the PDFs, as

$$E_3 E_4 \frac{d^6 \sigma_{\text{dijet}}}{d^3 p_3 d^3 p_4} = \sum_{kl} \int \int dx_1 dx_2 f_{\text{PDF}}^k(x_1, \mu_F^2) f_{\text{PDF}}^l(x_2, \mu_F^2) \times \frac{1}{32s\pi^2} \sum_{\bar{m}} |\mathcal{M}^{kl \rightarrow mn}|^2 \delta^4 \times (p_1 + p_2 - p_3 - p_4), \quad (55)$$

with $p_{1/2}$ and $p_{3/4}$ being the four-momenta of the incoming and outgoing partons, and $\mathcal{M}^{kl \rightarrow mn}$ being the corresponding matrix element. In order to get the complete expression corresponding to Fig. 25, one needs to integrate $\int \frac{d^3 p_3 d^3 p_4}{E_3 E_4}$ over the differential cross section, Eq. (55), whereas to obtain the inclusive jet (=parton) cross section one needs to integrate $\frac{d^3 p_4}{E_4}$. In any case, thanks to the four-dimensional δ function, the remaining numerical integration can be done, as discussed in detail in [35].

At least the quark parton distribution functions can be tested and compared with experimental data from deep inelastic electron-proton scattering. The structure function F_2 is given as $F_2 = \sum_k e_k^2 x f_{\text{PDF}}^k(x, Q^2)$ with $x = x_B = Q^2/(2pq)$, with p being the momentum of the proton, q the momentum of the exchanged photon, and $Q^2 = -q^2$. In Fig. 26, I plot F_2 as a function of x for different values of Q^2 . The red curve refers to EPOS PDFs, the green one to CTEQ PDFs [43], and the black points are data from ZEUS [44] and H1 [45–47]. The two PDFs give very similar results, and both are close to the experimental data.

Having checked the EPOS PDFs, I will use these functions to compute the jet (parton) cross section, using Eq. (55), integrating out the momentum of the second parton and the azimuthal angle of the first parton, for pp at 13 TeV. I define the parton yield $dn/dp_t dy$ as the cross section, divided by the inelastic pp cross section, showing the result in Fig. 27. I show results based on EPOS PDFs (red full line), CTEQ PDFs [43] (green dashed line), the full EPOS simulation (blue circles), and experimental data from ATLAS [48] (black triangles). At large values of p_t , all the different distribution agree, whereas at low p_t the EPOS Monte Carlo simulation results (using the

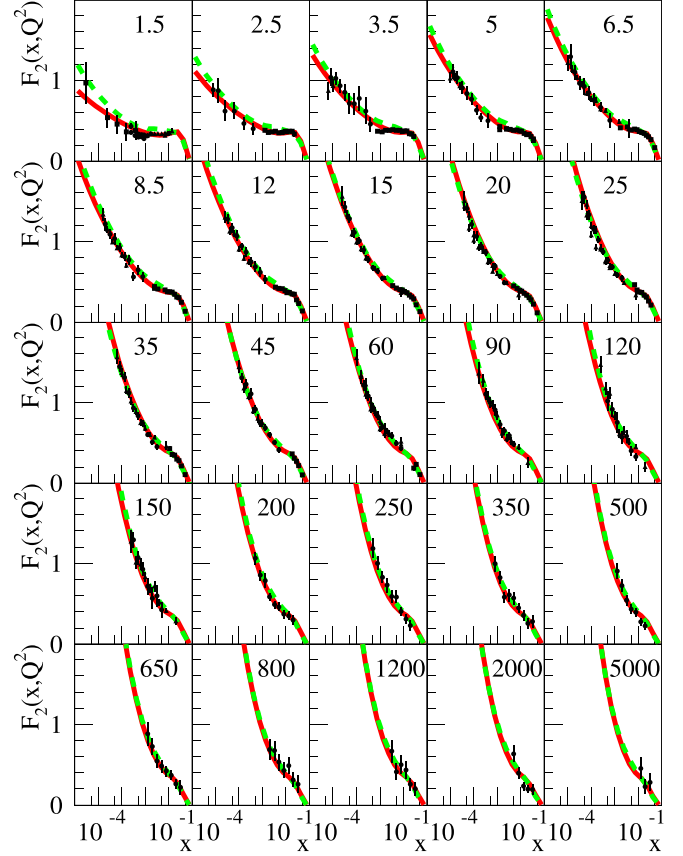


FIG. 26. F_2 as a function of x for different values of Q^2 , the latter one indicated (in units of GeV^2) in the upper right corners of each subplot. The red curve refers to EPOS PDFs, the green one to CTEQ PDFs, and the black points are data from ZEUS and H1.

full multiple scattering scenario) are significantly below the PDF results, as expected due to screening effects.

X. FULL EPOS4 (CORE + CORONA, HYDRO, MICROCANONICAL DECAY): CHECKING MULTIPLICITY DEPENDENCIES

The “factorization mode” as discussed in the last section is very useful to investigate inclusive cross sections at high p_t . But this represents only a very small fraction of all possible applications, and there are very interesting cases outside the applicability of that approach. A prominent example, one of the highlights of the past decade, concerns “collective phenomena in small systems”, following many discoveries showing that high-multiplicity pp events show very similar “collective” features as earlier observed in heavy ion collisions [5]. High multiplicity means automatically “multiple parton scattering”, and as discussed earlier this means that one has to employ the full parallel scattering machinery developed earlier, based on S-matrix theory.

But this is not the full story. The S-matrix part concerns “primary scatterings”, happening instantaneously at $t = 0$. As a result, in the case of a large number of Pomerons, one has correspondingly a large number of strings, which may overlap and “fuse”. In the EPOS framework, a core-corona procedure

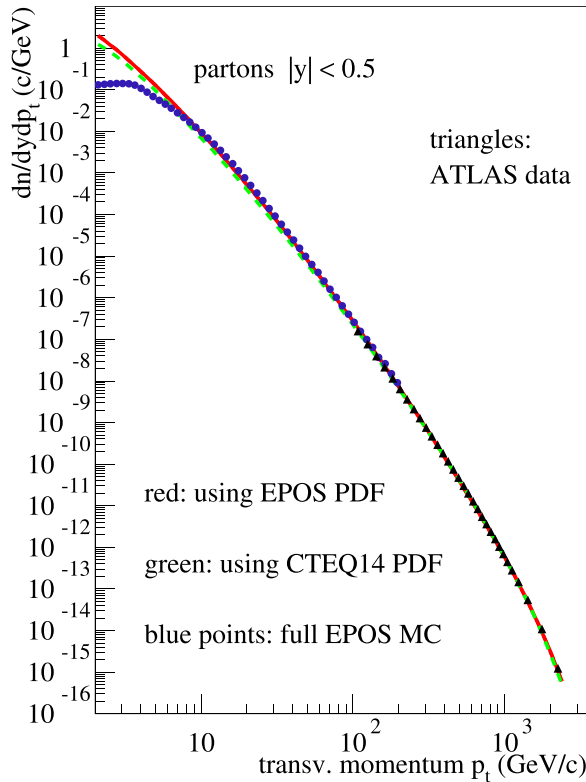


FIG. 27. Parton yield $dn/dp_t dy$ for pp at 13 TeV. I show results based on EPOS PDFs (red full line), CTEQ PDFs (green dashed line), the full EPOS simulation (blue circles), and experimental data from ATLAS (black triangles).

[49–51] is employed, where the strings at a given proper time τ_0 are first cut into “string segments”, which are separated into “core” and “corona” segments, depending on the energy loss of each segment when traversing the “matter” composed of all the other segments. Corona segments (per definition) can escape, whereas core segments lose all their energy and constitute what is called the “core”, which acts as an initial condition for a hydrodynamic evolution [51]. The evolution of the core ends whenever the energy density falls below some critical value ϵ_{FO} , which marks the point where the fluid “decays” into hadrons. It is not a switch from fluid to particles; it is a sudden decay, called “hadronization”.

In EPOS4, as discussed in detail in [37], a new procedure was developed concerning the energy-momentum flow through the “freeze-out (FO) hypersurface” defined by ϵ_{FO} , which allows defining an effective invariant mass, which decays according to microcanonical phase space into hadrons, which are then Lorentz boosted according to the flow velocities computed at the FO hypersurface. Also new and very efficient methods for the microcanonical procedure [37] were developed. Also in the full scheme, including primary and secondary interactions, energy-momentum and flavors are conserved. All the technical details about the new hadronization procedures can be found in [37]; the aim of this paper is to present an overview and some important results.

To summarize the above discussion, one has the “full” EPOS4 scheme, composed of

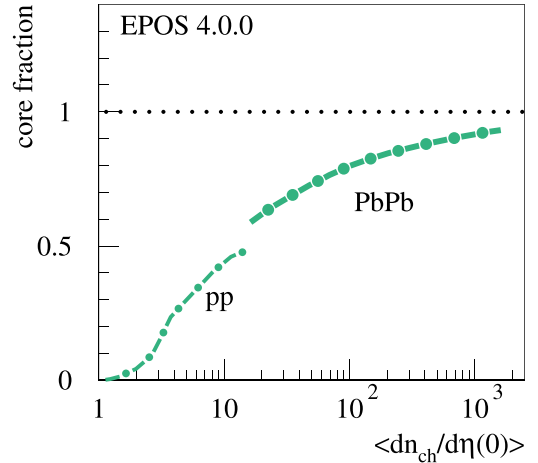


FIG. 28. The “core fraction” for pp (thin line) and $PbPb$ scattering (thick line).

- (i) primary interactions, based on an S-matrix approach for parallel scatterings,
- (ii) secondary interactions, composed of
 - (a) core-corona separation procedure,
 - (b) hydrodynamic evolution and microcanonical hadronization,
 - (c) hadronic afterburner (UrQMD [52,53]).

As an alternative, in order to better understand the different components, I also consider

- (i) the “core + corona” (“co-co”) contribution, i.e., primary interactions + secondary interactions but *without* hadronic afterburner;
- (ii) the “core” contribution, i.e., primary interactions + secondary interactions but *without* hadronic afterburner, only considering core particles;
- (iii) the “corona” contribution, i.e., primary interactions + secondary interactions but *without* hadronic afterburner, only considering corona particles.

One needs to exclude in these cases the hadronic afterburner, because the latter affects both core and corona particles, so in the full approach the core and corona contributions are no longer visible.

In the following, I will present particle ratios, always relative to pion yields, as well as mean p_t results, for the different contributions (“core”, “corona” etc), in pp and $PbPb$ collisions at LHC energies. In all cases, the results depend strongly on the relative weight of core to corona. It is clear that for low multiplicity pp scattering corona will dominate, whereas for central $PbPb$ collisions the core will dominate. To be more quantitative, I compute the “core fraction”, defined as the ratio of core to core + corona for pion production (with pions being the most frequent particle species). In Fig. 28, I show results for pp (thin lines) and $PbPb$ (thick lines), and one sees an almost continuous curve, going from zero (for low multiplicity pp) up to unity (for central $PbPb$). There is no overlap at intermediate multiplicity because one is running out of statistics for both pp and $PbPb$ collisions. I consider

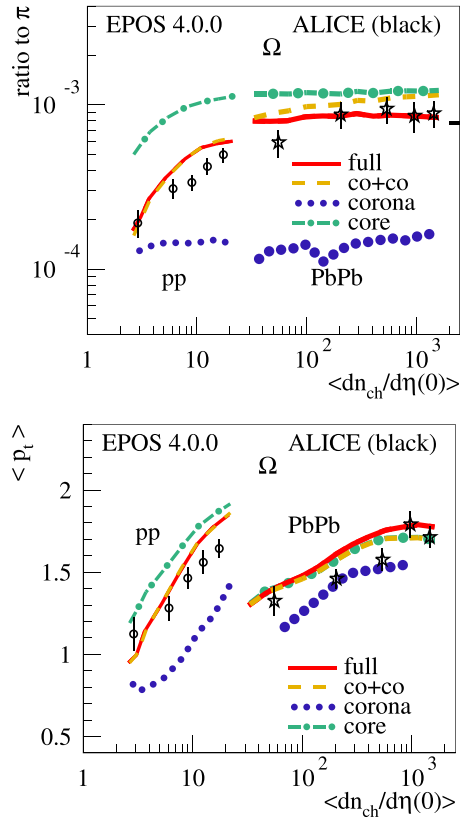


FIG. 29. Ratio of Ω over π (upper panel) and average transverse momentum (lower panel) versus multiplicity, for pp at 7 TeV (thin lines) and PbPb at 2.76 TeV (thick lines), compared to ALICE data.

actually pp at 7 TeV and PbPb at 2.76 TeV, since the particle ratios in the following refer to these energies.

In Fig. 29 (upper panel), I plot the ratio of Ω baryon yields over π yields versus multiplicity. I show results for pp at 7 TeV (thin lines) and PbPb at 2.76 TeV (thick lines), compared to ALICE data [54,55]. The different line styles refer to different contributions: the yellow dashed line refers to “core + corona” (“co-co”), i.e., primary interactions + hydro but *without* hadronic afterburner, the blue dotted line refers to the “corona”, and the green dashed-dotted line refers to the “core” part. The red line is the “full” contribution, i.e., core + corona + hadronic afterburner. One sees an almost flat line for the corona contribution, similar for pp and PbPb, which is understandable, since “corona” means particle production from string fragmentation, which does not depend on the system. One observes also a flat curve for the “core” part at high multiplicity, which is again expected since the core hadronization is determined by the freeze-out energy density, which is system independent. However, when the system gets very small, one gets a reduction of heavy particle production due to the microcanonical procedure (with its energy and flavor conservation constraints), whereas a grand canonical treatment would give a flat curve down to small multiplicities. It is remarkable that the “core” curve is far above the “corona” one, which simply reflects the fact that Ω production is much more suppressed in string decay, compared to statistical (“thermal”) production. This explains why the “core + corona” contribution increases

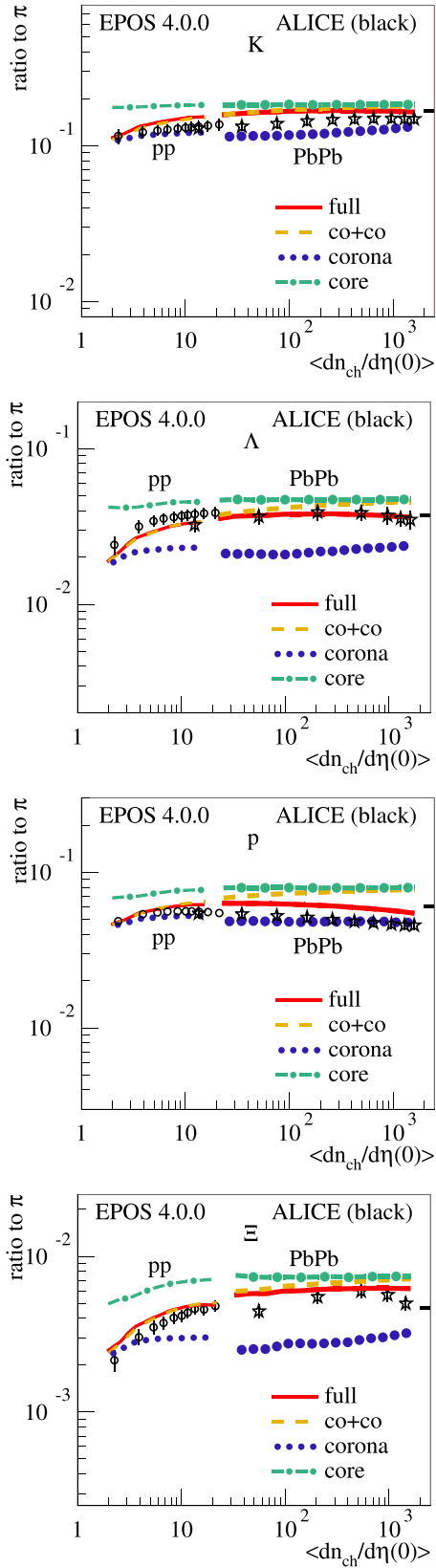
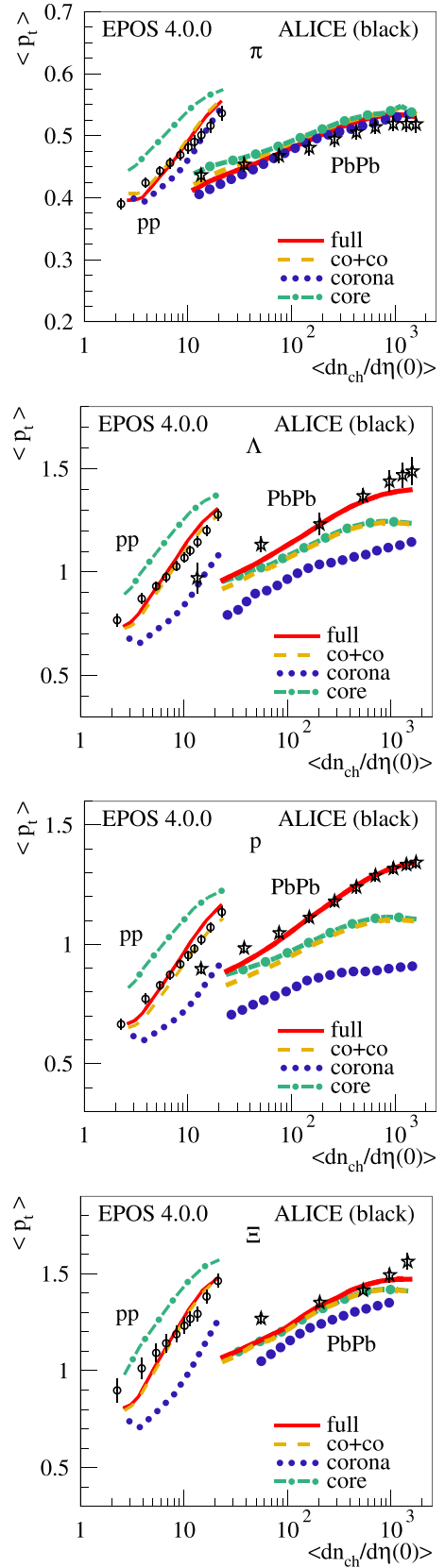
by one order of magnitude from low to high multiplicity, because simply the relative weight of the core fraction increases from zero to unity. The effect from hadronic rescattering (difference between “full” and “co-co”) is relatively small, some suppression due to baryon-antibaryon annihilation can be seen.

Whereas the Ω over π ratios are essentially smooth curves, from pp to PbPb, the situation changes completely when looking at the average transverse momentum $\langle p_t \rangle$ versus multiplicity, as shown in Fig. 29 (lower panel), where I also show results for pp (thin curves) and PbPb (thick curves), for the different contributions. Here one sees (for all curves) a significant discontinuity when going from pp to PbPb. The “corona” contributions are not flat (as the ratios), but they increase with multiplicity, in the case of pp being even more pronounced than for PbPb. This is a “saturation effect”: the saturation scale increases with multiplicity, which means with increasing multiplicity the events get harder, producing higher p_t . The situation is different for PbPb, where an increase of multiplicity is mainly due to an increase of the number of active nucleons, with a more modest increase of the saturation scale with multiplicity. Also, the “core” curves increase strongly with multiplicity, and here as well more pronounced in the case of pp , due to the fact that one gets for high-multiplicity pp high energy densities within a small volume, leading to strong radial flow. Again, the core + corona contribution is understood based on the continuous increase of the core fraction from low to high multiplicity.

In Figs. 30 and 31, I show the multiplicity dependencies of ratios and mean p_t for different hadrons, which are qualitatively similar to the Ω results, just the differences between the corona and the core curves are smaller. The data are from ALICE [54–58].

It is very useful (and necessary) to consider at the same time the multiplicity dependence of particle ratios and of mean p_t results, since their behavior is completely different (the former is continuous, the latter jumps). Despite these even qualitative differences between the two observables, the physics issues behind these results is the same, namely saturation, core-corona effects which mix flow (being very strong) and nonflow, and microcanonical hadronization of the core.

Another very important and useful variable is the multiplicity dependence of D meson production, where ‘ D ’ stands for the sum of D^0 , D^+ , and D^{*+} . This is much more than just “another particle”, since the D meson contains a charm quark, the latter one being created exclusively in the parton ladder and not during fragmentation or in the plasma. In Fig. 32, I plot the normalized D meson multiplicity $\frac{d^2N}{dydp_t} / \langle \frac{d^2N}{dydp_t} \rangle$ as a function of the normalized charged particle multiplicity $\frac{d^2N_c}{dydp_t} / \langle \frac{d^2N_c}{dydp_t} \rangle$ for different p_t ranges in pp scattering at 7 TeV, compared to ALICE data [59]. It is interesting to see in which way the simulations and the data deviate from the reference curve, which is the dashed black line representing identical multiplicity dependence for D mesons and charged particles. Considering the EPOS results without hydro (green lines), for low p_t (1–2 GeV/ c) the curve is slightly above the reference, but with increasing p_t the green curves get steeper, which is due to the fact that with increasing multiplicity the saturation


 FIG. 30. As Fig. 29 (upper panel), but for K , p , Λ , Ξ .

 FIG. 31. As Fig. 29 (lower panel), but for K , p , Λ , Ξ .

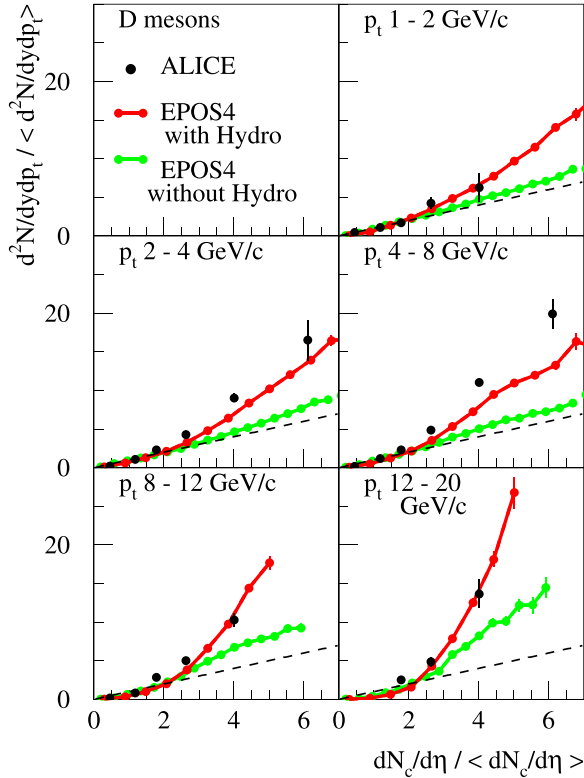


FIG. 32. Normalized D meson multiplicity as a function of the normalized charged particle multiplicity for different p_t ranges in pp scattering at 7 TeV. I show EPOS results with and without hydro, compared to ALICE data.

scale increases, and the events get harder, producing more easily both high p_t and charmed particles. Considering EPOS with hydro (red curves), the increase compared to the green curves is much stronger, which is due to the fact that “turning on hydro” will reduce the multiplicity (the available energy is partly transformed into flow). The red curves are close to the experimental data, both showing a much stronger increase compared to the reference curve, with the effect getting bigger with increasing p_t . So one may conclude this paragraph: to get these final results (the strong increase), two phenomena are crucial, namely, saturation which makes high multiplicity events harder, and the “hydro effect” which reduces multiplicity and “compresses” the multiplicity axis.

Concerning earlier EPOS versions, there are no “real publications” concerning these multiplicity dependencies, only plots based on “preliminary versions” shown at conferences or given to experimental colleagues. But none of the preliminary versions were able to fit reasonably well at the same time all the data shown in this section.

XI. CHARMED HADRONS

Having already discussed the multiplicity dependence of charm production in the last section, I will show here some basic charm results (a detailed discussion about charm production can be found in [35]). I consider here just primary interactions, no hydro and no hadronic cascade, so the charm quarks originate from cut Pomerons, more precisely from

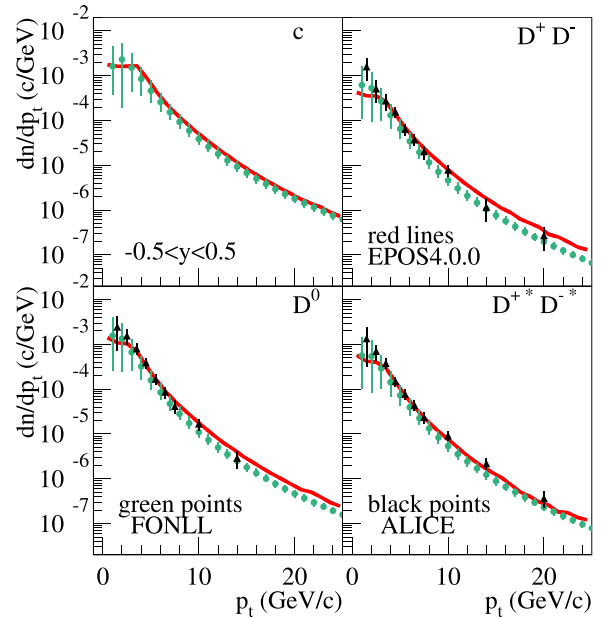


FIG. 33. Transverse momentum spectra of c quarks and charmed mesons in pp at 7 TeV.

the parton ladder. Cut parton ladders correspond in general to two chains of partons $q - g - \dots - g - \bar{q}$ identified as kinky strings, with q referring to light flavor quarks, and g to gluons. The Born process or branchings in the spacelike or the timelike cascade may lead to $Q\bar{Q}$ production, where Q refers to “heavy flavor” (HF) quarks, i.e., charm or bottom. In this case, one ends up with parton chains of the type $q - g - \dots - g - \bar{Q}$ and $Q - g - \dots - g - q$, which will decay (among others) into HF hadrons. In Fig. 33, I show transverse momentum spectra of c quarks (upper left), D^+ , D^- mesons (upper right), D^0 mesons (lower left), and D^{+*} , D^{-*} mesons (lower right) in pp collisions at 7 TeV. The red lines refer to EPOS simulations, the green points to FONLL calculations [60], and the black points to ALICE data [61]. In Fig. 34, I plot transverse momentum spectra of Λ_c and Ξ_c baryons (upper panel) and their ratio with respect to D^0 mesons in pp collisions at 5 TeV. The red lines refer to EPOS simulations, and the black points to ALICE data [62,63]. The production of charmed baryons is in principle straightforward: they are also coming from $q - g - \dots - g - \bar{Q}$ and $Q - g - \dots - g - q$ strings (with Q being a c quark in this case). The only difference compared to charmed meson production is the fact that here a diquark-antidiquark breakup occurs, which results in an essentially flat baryon/meson ratio, whereas the data show an increase towards small p_t . A similar “baryon/meson enhancement” in the region around 2–6 GeV/c has already been observed in the light flavor sector, where one possible explanation is collective flow.

Since charm is produced (as everything else) in an event-by-event manner, one produces for each charm quark the corresponding charm antiquark, and depending on the production details they have characteristic correlations, which are also visible in D meson pair correlations. In Fig. 35, I show two-hadron correlations for $D^0 D^0$ (upper left plot), $D^0 \bar{D}^0$ (upper right), $D^+ D^-$ (lower left), and $D^+ D_s^-$ (lower right) as

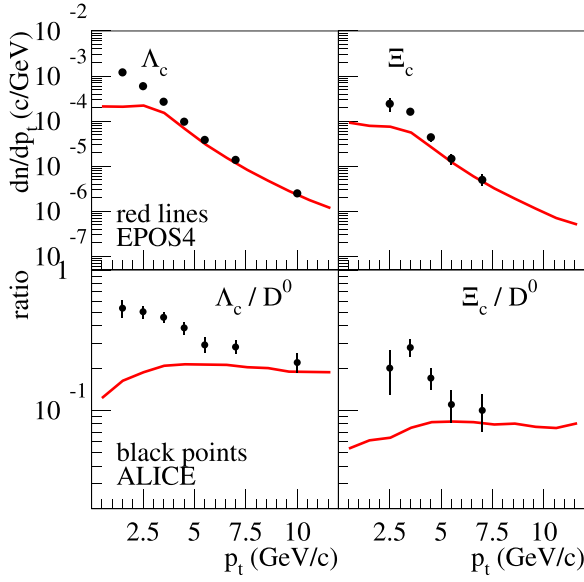


FIG. 34. Transverse momentum spectra of charmed baryons in pp at 5.02 TeV.

a function of the rapidity difference Δy in pp collisions at 7 TeV, with p_t values between 3 and 12 GeV/ c and rapidities between 2 and 4. Red lines represent EPOS4 simulations and black dots data from LHCb [64]. In Fig. 36, I plot the correlations of these pairs as a function of the azimuthal angle difference $\Delta\phi$, again compared to LHCb.

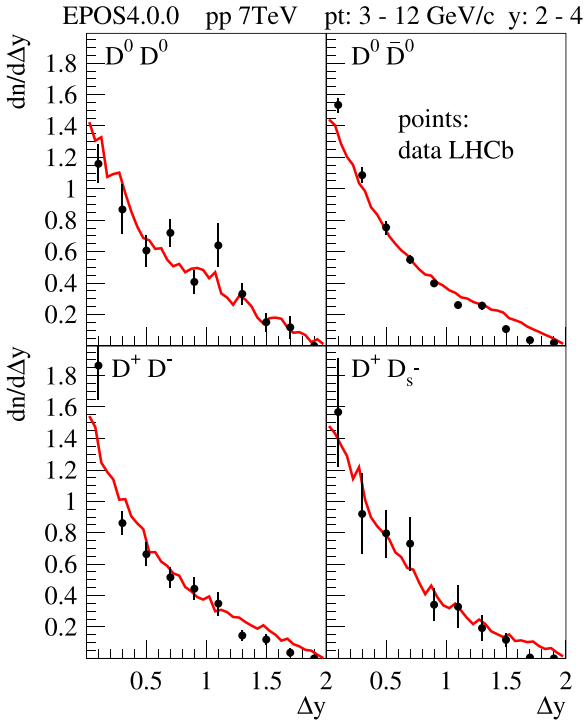


FIG. 35. Two hadron correlations for $D^0 D^0$, $D^0 \bar{D}^0$, $D^+ D^-$, and $D^+ D_s^-$ as a function of the rapidity difference Δy in pp collisions at 7 TeV. Red lines represent EPOS4 simulations and black dots data from LHCb.

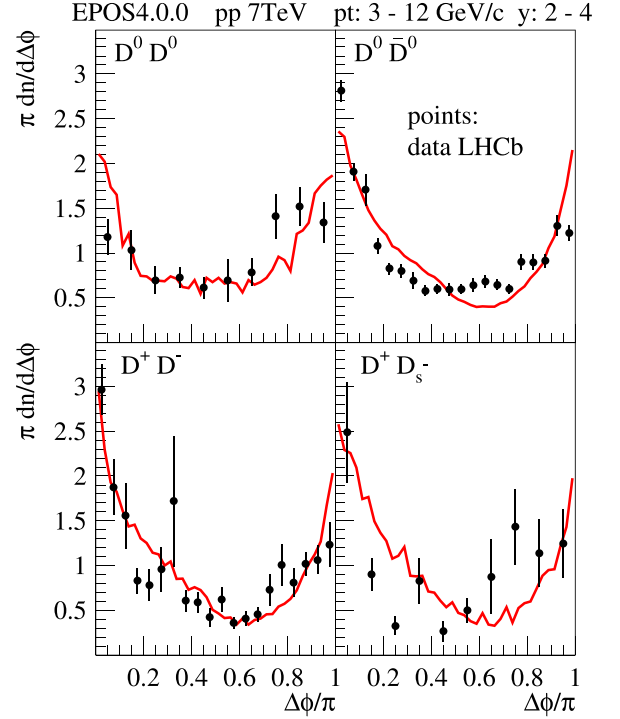


FIG. 36. Two hadron correlations for $D^0 D^0$, $D^0 \bar{D}^0$, $D^+ D^-$, and $D^+ D_s^-$ as a function of the azimuthal angle difference $\Delta\phi$ in pp collisions at 7 TeV. Red lines represent EPOS4 simulations and black dots data from LHCb.

It should be noted that $D^0 D^0$ represents c - c correlations, whereas the three other combinations $D^0 \bar{D}^0$, $D^+ D^-$, and $D^+ D_s^-$ represent c - \bar{c} correlations. For the latter ones, the situation is quite simple: the c and the \bar{c} are always produced as pair from the same process, and therefore one expects them to be close in rapidity, with a preference of $\Delta\phi = 0$ (in case of a timelike $g \rightarrow c\bar{c}$), or at $\Delta\phi = \pi$ in case of a Born process $gg \rightarrow c\bar{c}$. This is precisely what is seen: The rapidity correlations have maxima at $\Delta y = 0$ and then drop quickly; the $\Delta\phi$ correlations have maxima at $\Delta\phi = 0$ and $\Delta\phi = \pi$, observed in both EPOS4 simulations and data.

Surprisingly, the $D^0 D^0$ correlations (corresponding to a c - c pair) look very similar, which suggests that also c - c pairs originate from the same process, like a timelike $g \rightarrow gg \rightarrow c\bar{c} c\bar{c}$ or a Born process $gg \rightarrow gg$ followed by $g \rightarrow c\bar{c}$, $g \rightarrow c\bar{c}$.

Since EPOS4 creates charm always in terms of c - \bar{c} pairs, it is quite tempting to look into the possibility to produce charmonium. It is easy to implement the idea of the color evaporation model [65–67], where charmonium is created with a certain probability in the case of a c - \bar{c} pairs being in the appropriate mass range. So one considers all c - \bar{c} pairs from the same Pomeron (fully evaluated, including timelike emissions), and compute the invariant mass $M_{c\bar{c}}$. Whenever this mass is less than the sum of two D meson masses and bigger than the J/Ψ mass, the c - \bar{c} pair is with a certain probability $w_{J/\Psi}$ considered to be a J/Ψ . In Fig. 37, I plot prompt J/Ψ (not coming from beauty decays) from EPOS4 simulations compared to ATLAS data [68].

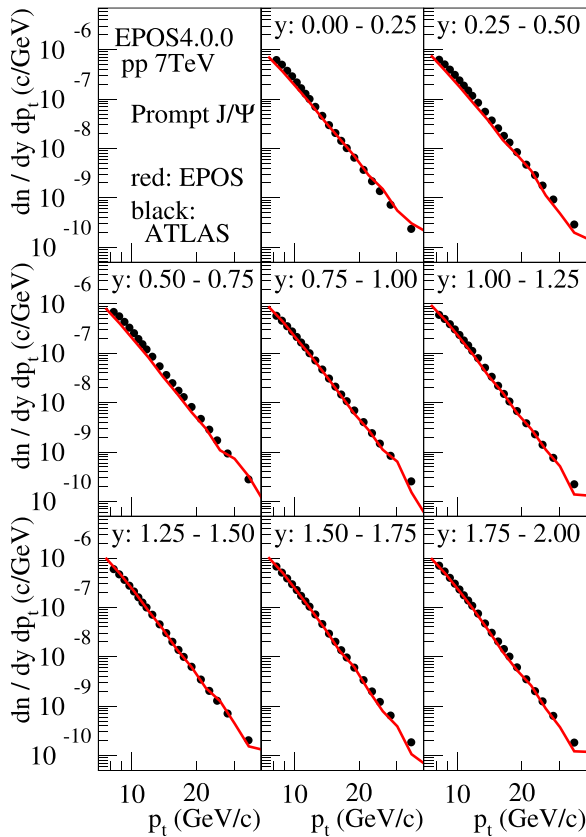


FIG. 37. Transverse momentum spectra of prompt J/Ψ in pp at 7 TeV.

XII. SUMMARY

I reported on new ideas, implemented in EPOS4, which provide a new understanding of a deep connection between four basic concepts in pp and AA collisions: rigorous parallel scattering, energy conservation, factorization, and saturation. It is mandatory to treat multiple scatterings in parallel, and a “natural” framework is S-matrix theory, with an S-matrix being given as a product of several entities representing individual scatterings referred to as Pomerons, and with energy-momentum conservation being implemented in an unbiased fashion via δ functions without imposing any ordering of collisions (this is what is meant by rigorous parallel scattering). The fundamental quantity of the multiple scattering approach is the cut single Pomerons expression G , representing inelastic parton-parton scattering. The fundamental question discussed in this paper is how to relate G to $G_{\text{QCD}}(Q_0^2)$, where the latter refers to parton-parton scattering in the framework of pQCD, having as basic elements parton evolutions with constant virtuality cutoff Q_0^2 and a hard $2 \rightarrow 2$ elementary QCD scattering. I refer to $G = G_{\text{QCD}}(Q_0^2)$ as “naive choice”.

One recalls that factorization and binary scaling, often mentioned in this paper, amount to reducing the inclusive cross sections for pp and AA scattering to single Pomeron results, although the underlying physical processes involve multiple parallel scatterings.

I showed in Sec. IV that neglecting energy conservation leads perfectly to factorization and binary scaling. But in the Monte Carlo procedures eventually one needs the implementation of energy conservation, so one introduces inconsistencies, in the sense that the theoretical basis and the Monte Carlo realization are not compatible.

On the other hand, as shown in Secs. V and VI, considering energy conservation (or energy sharing) and using the “naive choice” $G = G_{\text{QCD}}(Q_0^2)$, one completely spoils factorization for hard processes, contradicted by data. I have shown that the problem is due to a “deformation” of the inclusive energy distribution of Pomerons connected to many other Pomerons, compared to isolated Pomerons: the probability of carrying a large fraction of the total energy is reduced, which is unavoidable. These deformations can be quantified in terms of deformation functions R_{deform} depending on the number N_{conn} and the squared energy fraction x_{PE} .

In Sec. VII, one takes note of two problems: (1) spoiling factorization when using the naive choice $G = G_{\text{QCD}}$ in case of respecting energy conservation, and (2) not considering saturation effects which are known to be important. The solution of these two problems has been shown to be a dynamical saturation scale Q_{sat}^2 , defined via $G = k \times G_{\text{QCD}}(Q_{\text{sat}}^2)$ with k being inversely proportional to the deformation function, with a G which must be independent of the connection number N_{conn} . In that case, even having multiple scattering, all inclusive pp and AA cross sections are reduced to a single Pomeron result, but only for hard processes as it should be. This is referred to as “generalized AGK cancellations”, which holds at large p_t , even in a scenario respecting energy conservation. The dynamical saturation scale works, because even a large number of parallel scatterings will not affect high p_t particle production; it will only make the saturation scale big and thus suppress small p_t particle production.

Since in the new formalism, the full multiple scattering scenario converges to the single Pomeron result in case of inclusive cross sections (generalized AGK cancellations), one may use the single Pomeron (or factorization) mode, based on EPOS parton distribution functions. So one can now, with the same formalism, treat extremely high p_t particle production in factorization mode, and as well collective effects in high multiplicity events using the full simulation.

I discussed several examples, essentially multiplicity dependencies (of particle ratios, mean p_t , charm production) which are very strongly affected by the saturation issues discussed in this paper and core-corona effects mixing flow (being very strong) and nonflow contributions.

ACKNOWLEDGMENTS

I thank T. Pierog for many contributions during the past decade and in particular the important proposal (and first implementation) in 2015 of using a Pomeron definition as $G \propto G_{\text{QCD}}(Q_{\text{sat}}^2)$ with a parametrized G . He also did a first attempt in 2020 to take into account the “deformation”, by using a correction factor being a ratio of a “theoretical” G_{theo} and the same function G_{theo} but with rescaled longitudinal momentum arguments.

- [1] D. Gross and F. Wilczek, *Phys. Rev. Lett.* **30**, 1343 (1973).
- [2] H. Politzer, *Phys. Rev. Lett.* **30**, 1346 (1973).
- [3] J. Collins, D. Soper, and G. Sterman, in *Perturbative Quantum Chromodynamics*, edited by A. H. Mueller (World Scientific, Singapore, 1989).
- [4] R. Ellis, W. Stirling, and B. Webber, in *QCD and Collider Physics*, Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology (Cambridge, University Press, Cambridge, 1996).
- [5] V. Khachatryan *et al.* (CMS Collaboration), *J. High Energy Phys.* **09** (2010) 091.
- [6] V. N. Gribov and L. N. Lipatov, *Sov. J. Nucl. Phys.* **15**, 438 (1972).
- [7] G. Altarelli and G. Parisi, *Nucl. Phys. B* **126**, 298 (1977).
- [8] Y. L. Dokshitzer, *Sov. Phys. JETP* **46**, 641 (1977).
- [9] V. N. Gribov, *Zh. Eksp. Teor. Fiz.* **53**, 654 (1967).
- [10] V. N. Gribov, *Sov. Phys. JETP* **29**, 483 (1969).
- [11] V. A. Abramovsky, V. N. Gribov, and O. V. Kancheli, *Yad. Fiz.* **18**, 595 (1973).
- [12] M. Braun, *Yad. Fiz. (Rus)* **52**, 257 (1990).
- [13] V. A. Abramovskii and G. G. Leptoukh, *Sov. J. Nucl. Phys.* **55**, 903 (1992).
- [14] K. Werner, *Phys. Rep.* **232**, 87 (1993).
- [15] H. J. Drescher, M. Hladik, S. Ostapchenko, T. Pierog, and K. Werner, *Phys. Rep.* **350**, 93 (2001).
- [16] T. Sjöstrand, S. Mrenna, and P. Skands, *Comput. Phys. Commun.* **178**, 852 (2008).
- [17] M. Baehr *et al.*, *Eur. Phys. J. C* **58**, 639 (2008).
- [18] S. Schumann and F. Krauss, *J. High Energy Phys.* **03** (2008) 038.
- [19] L. V. Gribov, E. M. Levin, and M. G. Ryskin, *Phys. Rep.* **100**, 1 (1983).
- [20] L. D. McLerran and R. Venugopalan, *Phys. Rev. D* **49**, 2233 (1994).
- [21] L. D. McLerran and R. Venugopalan, *Phys. Rev. D* **49**, 3352 (1994).
- [22] A. Kovner, L. McLerran, and H. Weigert, *Phys. Rev. D* **52**, 3809 (1995).
- [23] Y. V. Kovchegov, *Phys. Rev. D* **54**, 5463 (1996).
- [24] Y. V. Kovchegov, *Phys. Rev. D* **55**, 5445 (1997).
- [25] Y. V. Kovchegov and D. H. Rischke, *Phys. Rev. C* **56**, 1084 (1997).
- [26] J. Jalilian-Marian, A. Kovner, L. McLerran, and H. Weigert, *Phys. Rev. D* **55**, 5414 (1997).
- [27] J. Jalilian-Marian, A. Kovner, A. Leonidov, and H. Weigert, *Nucl. Phys. B* **504**, 415 (1997).
- [28] Y. V. Kovchegov and A. H. Mueller, *Nucl. Phys. B* **529**, 451 (1998).
- [29] A. Krasnitz and R. Venugopalan, *Nucl. Phys. B* **557**, 237 (1999).
- [30] J. Jalilian-Marian, A. Kovner, A. Leonidov, and H. Weigert, *Phys. Rev. D* **59**, 034007 (1999).
- [31] J. Jalilian-Marian, A. Kovner, and H. Weigert, *Phys. Rev. D* **59**, 014015 (1998).
- [32] J. Jalilian-Marian, A. Kovner, A. Leonidov, and H. Weigert, *Phys. Rev. D* **59**, 014014 (1998).
- [33] C. Bierlich, G. Gustafson, L. Lönnblad, and H. Shah, *J. High Energy Phys.* **10** (2018) 134.
- [34] C. Bierlich, S. Chakraborty, G. Gustafson, and L. Lönnblad, *Phys. Lett. B* **835**, 137571 (2022).
- [35] K. Werner and B. Guiot, *Phys. Rev. C* **108**, 034904 (2023).
- [36] K. Werner, [arXiv:2310.09380](https://arxiv.org/abs/2310.09380).
- [37] K. Werner, [arXiv:2306.10277](https://arxiv.org/abs/2306.10277).
- [38] R. E. Cutcosky, *J. Math. Phys.* **1**, 429 (1960).
- [39] T. Pierog, I. Karpenko, J. M. Katzy, E. Yatsenko, and K. Werner, *Phys. Rev. C* **92**, 034906 (2015).
- [40] K. Werner, F.-M. Liu, and T. Pierog, *Phys. Rev. C* **74**, 044902 (2006).
- [41] K. Werner, B. Guiot, I. Karpenko, and T. Pierog, *J. Phys.: Conf. Ser.* **458**, 012020 (2013).
- [42] T. Pierog and K. Werner, *Acta Phys. Pol. Suppl.* **8**, 1031 (2015).
- [43] S. Dulat *et al.*, *Phys. Rev. D* **93**, 033006 (2016).
- [44] M. Derrick *et al.* (ZEUS Collaboration), *Z. Phys. C* **72**, 399 (1996).
- [45] I. Abt *et al.* (H1 Collaboration), *Z. Phys. C* **63**, 377 (1994).
- [46] S. Aid *et al.* (H1 Collaboration), *Nucl. Phys. B* **470**, 3 (1996).
- [47] Adloff *et al.* (H1 Collaboration), *Nucl. Phys. B* **485**, 3 (1997).
- [48] M. Aaboud *et al.* (ATLAS Collaboration), *J. High Energy Phys.* **05** (2018) 195.
- [49] K. Werner, *Phys. Rev. Lett.* **98**, 152301 (2007).
- [50] K. Werner, I. Karpenko, T. Pierog, M. Bleicher, and K. Mikhailov, *Phys. Rev. C* **82**, 044904 (2010).
- [51] K. Werner, B. Guiot, I. Karpenko, and T. Pierog, *Phys. Rev. C* **89**, 064903 (2014).
- [52] S. A. Bass *et al.*, *Prog. Part. Nucl. Phys.* **41**, 255 (1998).
- [53] M. Bleicher *et al.*, *J. Phys. G* **25**, 1859 (1999).
- [54] J. Adam *et al.* (ALICE Collaboration), *Nat. Phys.* **13**, 535 (2017).
- [55] B. B. Abelev *et al.* (ALICE Collaboration), *Phys. Lett. B* **728**, 216 (2014); **734**, 314(E) (2014).
- [56] B. Abelev *et al.* (ALICE Collaboration), *Phys. Rev. C* **88**, 044910 (2013).
- [57] B. B. Abelev *et al.* (ALICE Collaboration), *Phys. Rev. Lett.* **111**, 222301 (2013).
- [58] J. Adam *et al.* (ALICE Collaboration), *Eur. Phys. J. C* **75**, 226 (2015).
- [59] J. Adam *et al.* (ALICE Collaboration), *J. High Energy Phys.* **09** (2015) 148.
- [60] M. Cacciari *et al.*, *J. High Energy Phys.* **10** (2012) 137.
- [61] B. Abelev *et al.* (ALICE Collaboration), *J. High Energy Phys.* **01** (2012) 128.
- [62] S. Acharya *et al.* (ALICE Collaboration), *Phys. Rev. Lett.* **127**, 202301 (2021).
- [63] S. Acharya *et al.* (ALICE Collaboration), *J. High Energy Phys.* **10** (2021) 159.
- [64] R. Aaij *et al.* (LHCb Collaboration), *J. High Energy Phys.* **06** (2012) 141; Addendum: **03** (2014) 108.
- [65] V. Barger, W. Keung, and R. Phillips, *Phys. Lett. B* **91**, 253 (1980).
- [66] R. Gavai *et al.*, *Int. J. Mod. Phys. A* **10**, 3043 (1995).
- [67] Y.-Q. Ma and R. Vogt, *Phys. Rev. D* **94**, 114029 (2016).
- [68] G. Aad *et al.* (ATLAS Collaboration), *Eur. Phys. J. C* **76**, 283 (2016).