Imprints of QCD cascades in hadron multiplicity distributions

Maciej Rybczyński[*] and Zbigniew Włodarczyk[†]
Institute of Physics, Jan Kochanowski University, 25-406 Kielce, Poland

(Received 18 April 2025; accepted 13 May 2025; published 30 May 2025)

The relation between parton and hadron multiplicity distributions is discussed. To obtain parton multiplicity distribution, we propose decomposition of the multiplicity distributions of final-state hadrons. Such a procedure offers hope for experimental testing of physical probes of quantum complexity.

DOI: 10.1103/9mm7-rnxx

I. INTRODUCTION

In recent years, quantum information encompassing concepts such as entanglement and complexity has offered a novel perspective on high-energy physics. In a series of works (see Ref. [1] and references therein), it has been argued that understanding entanglement in quantum chromodynamics (QCD) could shed new light on the gluon-dominated structure and properties of high-energy hadrons. In high-energy physics, the investigation of quantum correlations is of paramount importance. Quantum correlations, often manifested as entanglement, play an essential role in describing the behavior of particles and their interactions. These correlations extend beyond the classical domain, enabling particles to exhibit nonlocal connections that challenge conventional notions of reality [2].

Entanglement, a fundamental phenomenon of quantum mechanics, has evolved from a conceptual puzzle into a central focus of high-energy physics. The principle of superposition, wherein particles can occupy multiple states simultaneously, and entanglement, wherein particles are intrinsically connected across vast distances, form the core of the quantum paradigm [3]. By transcending classical limits, entanglement challenges traditional views of particle separability. In high-energy physics, this nonlocal interdependence is crucial for understanding phenomena such as confinement, where quarks and gluons remain bound within hadrons [4]. Moreover, entanglement entropy not only characterizes the degree of entanglement but also provides insights into the structure of quantum states, revealing connections among particles and their collective behavior.

In the experiments involving multiparticle production, the observable quantities pertain to the final-state hadron distributions, while their connection to the underlying partonic distributions remains an open and unresolved problem. Within the framework of compound distributions, such as those encompassing the geometric distribution, we propose a method to infer properties of the parton-level dynamics from the experimentally accessible hadronic multiplicity distributions. The proposed approach offers a novel tool for experimental investigations of quantum complexity in strongly interacting systems.

II. PARTON DISTRIBUTION FROM A TOY MODEL OF QCD EVOLUTION

It will be convenient for us to describe the parton evolution using the dipole representation. In this representation, a set of partons is represented by a set of color dipoles. In this model, the Balitsky-Fadin-Kuraev-Lipatov (BFKL) equation for the dipole scattering cross section at a rapidity y reproduces the powerlike increase of the cross section with energy $\exp(\Lambda y) = x^{-\Lambda}$ where Λ is the BFKL intercept.

Let us now introduce $P(n, y)$, which is the probability to find n dipoles (of a fixed size in our model) at rapidity y. For this probability, we can write the following recurrent equation [5,6]:

$$\frac{dP(n,y)}{dy} = -\Lambda n P(n,y) + \Lambda(n-1)P(n-1,y). \quad (1)$$

This equation represents a standard form of a cascade process. The first term accounts for the reduction in the probability of finding n dipoles as a result of their splitting into $n+1$ dipoles, whereas the second term describes the corresponding increase arising from the splitting of $n-1$ dipoles into n dipoles. By employing the generating function method to solve this equation, one obtains a

[*]Contact author: maciej.rybczynski@ujk.edu.pl
[†]Contact author: zbigniew.wlodarczyk@ujk.edu.pl

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

geometric distribution, as detailed in [7]

$$P(n) = \Theta(1-\Theta)^{n-1}, \tag{2}$$

where $\Theta = \exp(-\Lambda y)$ determines the mean multiplicity $\langle n\rangle = (1-\Theta)/\Theta$. The geometric distribution is the only memoryless discrete probability distribution.[1] The resulting Shanon entropy $S = -\sum_n P(n)\ln[P(n)]$ is

$$S = -\ln\Theta - \frac{1-\Theta}{\Theta}\ln(1-\Theta) = (\langle n\rangle + 1)\ln(\langle n\rangle + 1) - \langle n\rangle \ln\langle n\rangle. \tag{3}$$

Among all the discrete distributions with a specified mean value $\langle n\rangle$, the maximum entropy distribution has the shape $P(n) = C\cdot r^n$. The constants C and r were determined by the requirements that the sum of all the probabilities must be 1, and the expected value must be $\langle n\rangle$. This gives $C = 1/(\langle n\rangle - 1)$ and $r = (\langle n\rangle - 1)/\langle n\rangle$, where the respective maximum entropy distribution is the geometric distribution.

At large Λy ($x \sim 10^{-3}$), the relation between the entanglement entropy and the gluon structure function $G(x) = \langle n(x)\rangle \sim x^{-\Lambda}$ becomes very simple

$$S = \ln[G(x)]. \tag{4}$$

It signals that all $\exp(\Lambda y)$ partonic states have about equal probabilities $\exp(-\Lambda y)$. In this case, the entanglement entropy is maximal, and the proton is a maximally entangled state [6]. In the maximally entangled regime at small x, it appears that the behavior of the gluon structure function becomes universal. In analogy to statistical mechanics, in thermal equilibrium (maximal entropy), the equation of state is determined by the temperature $(1/x)$ and the effective number of degrees of freedom.

III. DECOMPOSITION OF THE FINAL STATE

Multiplicity distributions $P(N)$, where N denotes the number of produced particles, are among the primary observables measured in multiparticle production experiments. They have long been recognized as a crucial source of information regarding the underlying dynamics of the particle production mechanism [8]. The precise relationship between partonic and hadronic multiplicity distributions remains an open question. A common working assumption has been the hypothesis of parton-hadron duality [9], which posits that the partonic and hadronic distributions are essentially identical. Under this assumption, the hadron multiplicity distribution would be expected to follow the geometric form given in Eq. (2) [10,11]. However, this assumption is highly idealized, particularly for hadrons produced in the central rapidity region, since the experimentally measured distributions $P(N)$ often exhibit significant deviations from the geometric distribution (2).

The negative binomial distribution (NBD) is typically the first choice for fitting multiplicity distributions $P(N)$ in high-energy collision data. However, as the collision energy and the number of produced secondaries increase, the single NBD increasingly fails to accurately describe the data, particularly in the high-multiplicity tail. This discrepancy has motivated the use of more sophisticated models, including two-, three-, or multicomponent NBDs, as well as alternative forms of $P(N)$ [12]. Since neither a single NBD nor a binomial distribution (BD) provides a satisfactory description of the data across the full multiplicity range, we explore the framework of compound distributions. These are particularly suitable in scenarios where the particle production process proceeds via an intermediate stage involving the generation of M sources, such as clusters, fireballs, or other composite objects, distributed according to some function $f(M)$ with the generating function $F(z)$. Each of these sources subsequently decays independently into a number of secondaries $n_{i=1,\ldots,M}$ governed by an identical for all M distribution $g(n)$ characterized by the generating function $G(z)$. The resultant multiplicity distribution

$$h\left(N = \sum_{i=0}^{M} n_i\right) \tag{5}$$

is a compound distribution of $f(M)$ and $g(n)$ with generating function

$$H(z) = F(G(z)). \tag{6}$$

By summing the contributions from k independent geometric distributions, where each individual multiplicity $n_i \in (0,\infty)$, one obtains the well-known NBD. The corresponding generating function is given by

$$H(z) = \left(\frac{\Theta'}{1-(1-\Theta')z}\right)^k, \tag{7}$$

where Θ' is a parameter related to the geometric distribution.[2]

When the number of particle-emitting sources M fluctuates according to a BD with mean λ and maximal value K, the resulting compound distribution is known as the geometric-binomial distribution [13] with generating function

$$H(z) = \left(1 + \frac{z-1}{1-(1-\Theta')z}\cdot\frac{\lambda}{K}\right)^K. \tag{8}$$

[1]Geometric distribution is the discrete version of the same property as in the exponential distribution. An approximate version of Eq. (2) for a continuous variable $z = n/\langle n\rangle$, and the relation to Koba-Nielsen-Olesen (KNO) scaling is discussed in Appendix A.

[2]The geometric distribution in Eq. (2) is defined for $n > 0$. For the case $n = 0, 1, 2, \ldots$, the parameter transforms as $1/\Theta' = 1/\Theta - 1$.

Alternatively, if the number of sources follows a Poisson distribution with mean λ, the resulting distribution is referred to as the geometric-Poisson distribution (also known as the Pólya-Aeppli distribution), whose generating function takes the form

$$H(z) = \exp\left(\frac{\lambda(z-1)}{1-(1-\Theta)z}\right). \tag{9}$$

Multiplicity distributions $P(N)$ of the compound geometric distributions (7)–(9) are given in Appendix B.

IV. EXPERIMENTAL TESTS

Reference [14] demonstrated that the multiplicity distribution of inclusive photons—predominantly originating from π^0 decays—produced in inelastic proton-proton collisions at center-of-mass energies $\sqrt{s} = 900$ GeV, and 2.76 and 7 TeV, and measured in the forward hemisphere $2.3 < \eta < 3.9$ by the ALICE experiment at the LHC, is well described by the geometric-Poisson distribution. This distribution is a specific case of a compound Poisson distribution suitable for modeling particle production in clustered systems: The number of clusters follows a Poisson distribution, while the number of particles within each cluster is governed by a geometric distribution. The extracted values of the distribution parameter Θ plotted as a function of the kinematic variable x are presented in Fig. 1.[3]

It is noteworthy that the multiplicity distribution arising from QCD parton cascades generally takes the form of the geometric distribution given by Eq. (2), where the parameter $\Theta = 1/(\langle n\rangle + 1)$ is directly related to the average parton multiplicity $\langle n\rangle$ [5,6]. In the regime of small Feynman x and large rapidity y, the parameter Θ exhibits an exponential dependence

$$\Theta = \exp\left[-4\ln 2\bar{\alpha}_S y\right], \tag{10}$$

where $\bar{\alpha}_S = \alpha_S N_c/\pi$ denotes the rescaled strong coupling constant, and N_c is the number of QCD colors. Empirically, the extracted values of Θ correspond to relatively large values of the effective coupling: $\bar{\alpha}_S = 0.12$, 0.16, and 0.19 for $\sqrt{s} = 0.9$, 2.76, and 7 TeV, respectively. This observation suggests that the hadronic final state may not be adequately described by a dilute system of partons. Instead, it may be more appropriate to treat the system as a dense partonic medium. In such a scenario, one expects a scaling relation of the form $1/\Theta = \langle n\rangle \sim Q_S^2$, where Q_S denotes the saturation scale [5]. Our analysis reveals that Q_S^2 exhibits a power-law dependence on the Feynman-x variable,

$$Q_S^2 \sim x^{-\Lambda} \tag{11}$$

[3]For a given rapidity y, the variable x is defined as $x = 2m_T\cosh(y)/\sqrt{s}$, where the transverse mass is given by $m_T = \sqrt{p_T^2 + m^2}$.

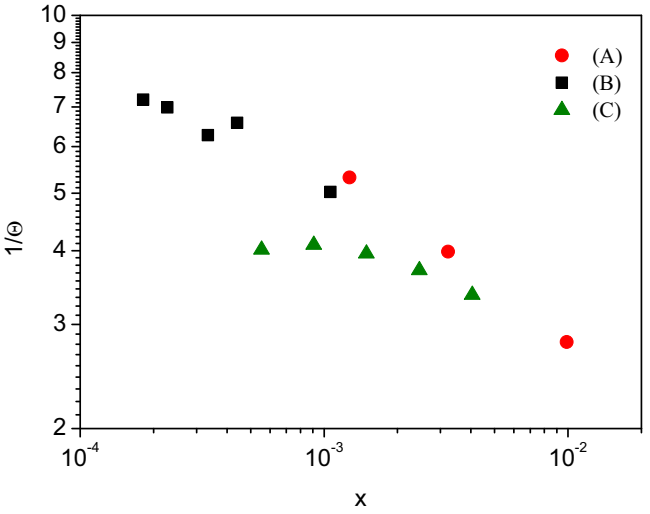


FIG. 1. Parameter Θ derived from multiplicity distributions as a function of x. (A) Circles: $\sqrt{s} = 0.9$, 2.76, and 7 TeV, $2.3 < \eta < 3.9$ (photons). (B) Squares: $\sqrt{s} = 8$ TeV, $|\eta| < 2$, $|\eta| < 2.4$, $|\eta| < 3$, $|\eta| < 3.4$, and $-2.4 < \eta < 5$ (charged particles). (C) Triangles: $\sqrt{s} = 7$ TeV, $2 < \eta < 2.5$, $2.5 < \eta < 3$, $3 < \eta < 3.5$, $3.5 < \eta < 4$, and $4 < \eta < 4.5$ (charged particles).

with an extracted exponent $\Lambda = 0.31$. This result is in close agreement with previous determinations from deep inelastic scattering and diffractive processes, which report $\Lambda = 0.29$–0.30 [15,16], as well as with theoretical predictions based on nonlinear QCD evolution [17].

In the midrapidity region, charged-particle multiplicity distributions in proton-proton collisions have typically been described using the NBDs [18]. For $\sqrt{s} = 8$ TeV collisions, the ALICE Collaboration performed fits to the multiplicity distributions measured within various pseudorapidity windows—specifically, $|\eta| < 2$, $|\eta| < 2.4$, $|\eta| < 3$, $|\eta| < 3.4$, and $-3.4 < \eta < 5$—using a two-component NBD model [18]. For the soft component, we extract the corresponding Θ parameter shown in Fig. 1. In contrast, the hard component exhibits a distinct behavior, suggesting that its dynamics are potentially governed by the quark structure function $x\Sigma(x)$, which does not exhibit the same scaling behavior as the gluon distribution.

Additionally, Fig. 1 displays the values of Θ extracted from the multiplicity distributions $P(N)$ of charged particles measured in the forward pseudorapidity region ($2 < \eta < 2.5$, $2.5 < \eta < 3$, $3 < \eta < 3.5$, $3.5 < \eta < 4$, and $4 < \eta < 4.5$) in $\sqrt{s} = 7$ TeV proton-proton collisions in the LHCb experiment [19]. In this kinematic regime, a single NBD provides an adequate description of the data.

The associated hadron entropy calculated using Eq. (3) is presented in Fig. 2. The obtained entropy values are found to be in qualitative agreement with the von Neumann entanglement entropy associated with gluon distributions, approximated by $\ln(n_{\text{gluon}})$ [11]. In Fig. 3, we compare our estimates of the gluon number density at fixed x, defined as $G(x) = \exp(S)$, with several widely used parton

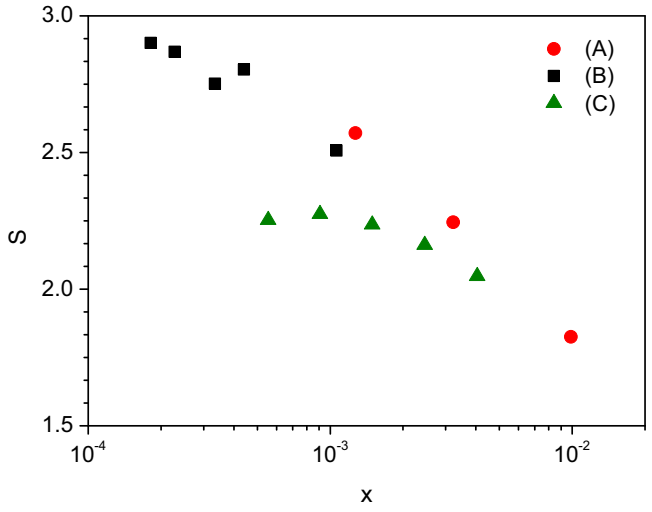


FIG. 2. Hadron entropy derived from multiplicity distributions as a function of x. Symbols as in Fig. 1.

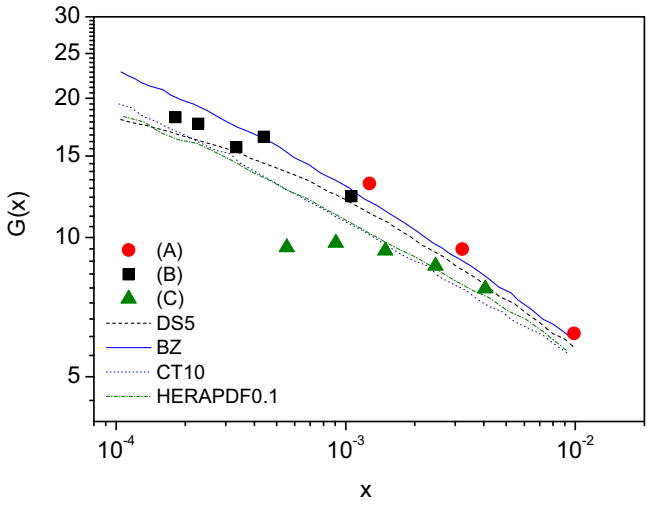


FIG. 3. Small-x behavior of gluon distribution $G(x) = \exp(S)$ in comparison with different parametrizations. Dashed curve represents the DS result, solid curve represents the BZ model, dotted curve represents CT10, and dash-dotted curve represents the HERAPDF0.1 result. Symbols as in Fig. 1.

distribution function (PDF) parametrizations, including DS [20] (dashed), BZ [21,22] (solid), CT10 [23,24] (dotted), and HERAPDF0.1 [25,26] (dash-dotted) (Abbreviations of the PDF parametrizations are derived from the authors' names or the names of research groups). The comparison is performed at a fixed scale of $Q^2 = 20$ GeV^2; however, the dependence on Q^2 primarily affects the normalization rather than the shape of the distributions. This behavior is consistent with the Regge-like ansatz used to solve the linear Dokshitzer-Gribov-Lipatov-Altarelli-Parisi evolution equation at small x, which leads to a factorized form

$$G(x, Q^2) = G'(Q^2)G(x) = G'(Q^2)x^{-\Lambda}, \qquad (12)$$

where $G'(Q^2)$ is a function of the hard scale Q^2, and Λ represents the Regge intercept of the gluon distribution [20].

The HERAPDF0.1 parametrization is based on a combined analysis of H1 and ZEUS datasets, which resolves discrepancies between the two through cross-calibration. It plays an important role in making precise predictions for Standard Model processes. The CT10 PDFs derived from global fits including HERA-1 neutral-current and charged-current deep-inelastic scattering (DIS) cross sections represent one of the most comprehensive modern PDF sets. They are widely employed in phenomenological studies at the Tevatron, LHC, and other high-energy experiments. Our results exhibit a power-law behavior for the gluon density of the form $G(x) = 1.5x^{-0.3}$, and the resulting shapes show good agreement with the trends predicted by the aforementioned parametrizations.

V. CONCLUSIONS

We highlight a compelling connection between hadronic observables and the underlying partonic dynamics involved in fragmentation and hadronization processes. Within the framework of compound multiplicity distributions for final-state hadrons, we identify a geometric distribution that effectively characterizes the gluon multiplicity emerging from QCD cascade evolution.

This study introduces novel tools for probing quantum information and complexity in QCD and opens the possibility for experimental investigations of quantum complexity through accessible hadronic observables. In particular, the entropy extracted from the hadron multiplicity distribution leads to a gluon density of the form $G(x) = \exp(S) \sim x^{-0.3}$, exhibiting small-x behavior consistent with predictions from various gluon density models.

We anticipate that these findings will motivate broader theoretical investigations and contribute to a deeper understanding of multiparticle production mechanisms in ultra-relativistic hadronic collisions. While the simplicity of this scenario is appealing, its applicability to the full range of experimental data remains to be confirmed through further studies based on dynamical modeling of the collision process.

ACKNOWLEDGMENTS

In preparation of this publication, we used the resources of the Centre for Computation and Computer Modelling of the Faculty of Exact and Natural Sciences of the Jan Kochanowski University in Kielce, modernized from the funds of the Polish Ministry of Science and Higher Education in the "Regional Excellence Initiative" program under the Project No. RID/SP/00015/2024/01.

DATA AVAILABILITY

No data were created or analyzed in this study.

APPENDIX A: KNO SCALING

In high-energy particle physics, the KNO scaling introduced over half a century ago has played a foundational role in the empirical analysis of hadronic multiplicity distributions at high energies [27]. KNO scaling posits that the probability $P(N)$ of producing N particles in a given collision process exhibits a universal scaling behavior of the form

$$\langle N\rangle P(N) = \psi(z), \tag{A1}$$

where $\langle N\rangle$ is the mean multiplicity, and $z = N/\langle N\rangle$ is the scaling variable. Despite its empirical success, the derivation of KNO scaling directly from first principles in QCD remains nontrivial. In the context of perturbative QCD, theoretical studies suggest that KNO scaling naturally emerges in the asymptotic limit of high virtuality, particularly in jet fragmentation processes [28]. The potential universality of KNO scaling and its connection with quantum entanglement entropy has also been discussed in Ref. [29].

For small values of the geometric parameter Θ (corresponding to large average multiplicities $\langle n\rangle$), the approximation

$$\exp\left[-\Theta/(1-\Theta)\right] \simeq 1-\Theta \tag{A2}$$

holds, allowing the geometric distribution given by Eq. (2) to be expressed as an exponential distribution:

$$P(n) = \frac{\Theta}{1-\Theta}\exp\left(-n\frac{\Theta}{1-\Theta}\right). \tag{A3}$$

Introducing the scaling variable $z = n/\langle n\rangle$, we obtain the KNO scaling form:

$$\langle n\rangle P(z) = \exp(-z). \tag{A4}$$

Consider a total multiplicity $N = \sum_i^k n_i$ arising from k independent sources, where each contribution of the multiplicity n_i is taken from the exponential distribution (A4), and the total mean multiplicity becomes $\langle N\rangle = k\langle n\rangle$. The sum of k exponentially distributed random variables is known to follow a gamma distribution, leading to the following KNO-scaled form:

$$\langle N\rangle P(N) = \frac{k^k}{\Gamma(k)} z^{k-1}\exp(-kz), \tag{A5}$$

where $z = N/\langle N\rangle$.

Notably, the scaling form in Eq. (A5) has also been derived by various other authors, though their theoretical motivations differ from the approach presented here [30].

APPENDIX B: $P(N)$ OF THE COMPOUND GEOMETRIC DISTRIBUTIONS

The probability distribution $P(N)$ of the random variable N is given by the derivatives of the generating function $H(z)$,

$$P(N) = \frac{1}{N!}\frac{\partial^N H(z)}{\partial z^N}\bigg|_{z=0}. \tag{B1}$$

Generating function (7) leads to the well-known negative binomial distribution

$$P(N) = \binom{N+k-1}{N}(1-\Theta')^N(\Theta')^k. \tag{B2}$$

For the generating function (8), we have the geometric-binomial distribution

$$\begin{aligned} P(0) &= \left(1-\frac{\lambda}{K}+\frac{\lambda}{K}\Theta'\right)^K \\ P(N) &= (1-\Theta')^N\sum_{m=1}^{K}\binom{K}{m}\binom{N+m-1}{N} \\ &\quad\times\left(\frac{\lambda}{K}\Theta'\right)^m\left(1-\frac{\lambda}{K}\right)^{K-m}, \end{aligned} \tag{B3}$$

and for the generating function (9), we obtain the geometric-Poisson distribution

$$\begin{aligned} P(0) &= \exp(-\lambda) \\ P(N) &= \sum_{m=1}^{N}\frac{e^{-\lambda}\lambda^m}{m!}\binom{N-1}{m-1}\Theta^m(1-\Theta)^{N-m}. \end{aligned} \tag{B4}$$

[1] P. Caputa and K. Kutak, Phys. Rev. D **110**, 085011 (2024).
[2] M. Nadir, Entanglement in high energy physics: An overview, in *Quantum Entanglement in High Enery Physics*, edited by O. K. Baker (IntechOpen, London, 2024), p. 187.
[3] A. Aspect, J. Dalibard, and G. Roger, Phys. Rev. Lett. **40**, 1804 (1982).
[4] F. Wilczek, Annu. Rev. Nucl. Part. Sci. **32**, 177 (1982).
[5] E. Gotsman and E. Levin, Phys. Rev. D **102**, 074008 (2020).
[6] D. E. Kharzeev and E. M. Levin, Phys. Rev. D **95**, 114008 (2017).
[7] S. Chaturvedi, V. Gupta, and S. K. Soni, Mod. Phys. Lett. A **09**, 3359 (1994).

[8] W. Kittel and E. A. De Wolf, *Soft Multihadron Dynamics* (World Scientific, Singapore, 2005).
[9] Y. L. Dokshitzer, V. A. Khoze, S. I. Troian, and A. H. Mueller, Rev. Mod. Phys. **60**, 373 (1988).
[10] V. Andreev *et al.* (H1 Collaboration), Eur. Phys. J. C **81**, 212 (2021).
[11] Z. Tu, D. Kharzeev, and T. Ullrich, Phys. Rev. Lett. **124**, 062001 (2020).
[12] M. Rybczyński, G. Wilk, and Z. Włodarczyk, Phys. Rev. D **99**, 094045 (2019).
[13] H. Eledum and A. R. El-Alosey, Mathematische Operationsforschung und Statistik **10**, 1218 (2022).
[14] R. R. Nair, G. Wilk, and Z. Włodarczyk, Eur. Phys. J. A **59**, 203 (2023).
[15] K. Golec-Biernat and M. Wüsthoff, Phys. Rev. D **59**, 014017 (1998).
[16] K. Golec-Biernat and M. Wüsthoff, Phys. Rev. D **60**, 114023 (1999).
[17] D. N. Triantafyllopoulos, Nucl. Phys. **B648**, 293 (2003).
[18] ALICE Collaboraion, Eur. Phys. J. C **77**, 852 (2017).
[19] R. Aaij *et al.* (LHCb Collaboration), Eur. Phys. J. C **74**, 2888 (2014).
[20] M. Devee and J. K. Sarma, Nucl. Phys. **B885**, 571 (2014).
[21] G. R. Boroun, Eur. Phys. J. A **42**, 251 (2009).
[22] G. R. Boroun and S. Zarrin, Eur. Phys. J. Plus **128**, 119 (2013).
[23] M. Guzzi *et al.*, arXiv:1101.0561v1.
[24] H.-Liang Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin, and C.-P. Yuan, Phys. Rev. D **82**, 074024 (2010).
[25] H1 Collaboration, Eur. Phys. J. C **30**, 1 (2003).
[26] ZEUS Collaboration, Eur. Phys. J. C **42**, 1 (2005).
[27] S. Hegyi, Nucl. Phys. B, Proc. Suppl. **92**, 122 (2001).
[28] I. M. Dremin and J. W. Gary, Phys. Rep. **349**, 301 (2001).
[29] Y. A. Liu, M. Nowak, and I. Zahed, Universality of Koba–Nielsen–Olesen scaling in QCD at high energy and entanglement, in *Quantum Entanglement in High Enery Physics*, edited by O. K. Baker (IntechOpen, London, 2024), p. 257, 10.5772/intechopen.1002706.
[30] S. Mrowczynski, Z. Phys. C **27**, 131 (1985).