

CHARGED-PARTICLE MULTIPLICITY DISTRIBUTIONS DERIVED FROM THE PRINCIPLE OF MAXIMAL ENTROPY*

SÁNDOR LÖKÖS 

Institute of Nuclear Physics Polish Academy of Sciences
Radzikowskiego 152, 31-342 Kraków, Poland

Received 26 February 2025, accepted 13 May 2025,
published online 26 June 2025

Recent theoretical results renewed the interest in charged-particle multiplicity distributions. The Shannon entropy of such distributions is conjectured to be related to the entanglement or von Neumann entropy of the partonic quantum system. In this paper, we show that the measured charged-particle multiplicities can be derived from the principle of maximum entropy (POME or MAXENT) without any *a priori* physical assumption. The approach provides a natural explanation for the well-known negative binomial shape of the measured distributions.

DOI:10.5506/APhysPolBSupp.18.5-A19

1. Introduction

It has been theorized that considering the proton as a (maximally) entangled partonic quantum system could explain many puzzling theoretical and experimental observations [1–3]. The proton’s wave function is *partially* probed by another proton or electron, leading to entanglement between the probed and unprobed regions. The proposed observable is the entanglement or von Neumann entropy, which quantifies the degree of “entanglement” within the quantum system.

A series of papers aimed at describing the phenomenon and provides evidence to quantum entanglement between the sampled and non-sampled parts of the proton in the data by utilizing a partonic dipole picture of pQCD [3–8]. The calculations were compared to multiplicity measurements from proton–proton [9–11] and electron–proton [12] collisions, showing convincing agreement between the data and theoretical predictions.

Inspired by the success of the assumption of maximal entropy of the initial partonic system, it is worth investigating the problem from the final-state perspective. The ultimate goal of this paper is to describe the measured

* Presented at the 31st Cracow Epiphany Conference on the *Recent LHC Results*, Kraków, Poland, 13–17 January, 2025.

charged-particle distributions using the general principle of maximal Shannon entropy (POME), which has proven to be a powerful tool for addressing classical statistical physics problems in classical [13] as well as in quantum systems [14]. The essential idea is that Shannon entropy serves as the primer quantity, subject to certain constraints dictated by the data itself. Then the entropy is maximized subject to the constraints and a distribution is obtained that can be regarded as the least bias choice and containing the least amount of *a priori* assumptions about the physical system. Thus, the use of POME is not the application of physical laws, but a method of reasoning that ensures that no tacit arbitrary assumptions are introduced [13]. A similar approach was adopted elsewhere, see *e.g.* Refs. [15, 16] and references therein.

2. Charged-particle multiplicities from POME

The multiplicity of charged particles is a statistical observable by its nature. The distribution of multiplicities can be interpreted as a probability distribution and is therefore normalized to unity. This interpretation can be regarded as constraint. With no further constraints, an assumption of a previous work can be recovered [4]. The final-state distributions can be directly derived from POME introducing an extra constraint that fixes the mean of the distribution.

2.1. Initial state parton distributions

If the initial partonic system is maximally entangled, it can consist of N , equally probable states, *i.e.*, the probability of the partonic configurations is uniformly distributed: $p(n) = \frac{1}{N}$. Substituting this probability into Shannon's entropy formula, the entropy of the system can be derived. The number of states is N , therefore the entropy S_{parton} , too, can be expressed with gluon ($g(x, Q)$) and quark ($\Sigma(x, Q)$) distribution function as

$$S_{\text{parton}} = - \sum_n p(n) \ln p(n) = \ln(N) = \ln(xg(x, Q) + x\Sigma(x, Q)), \quad (1)$$

where x is the Bjorken- x and Q is the virtuality. Hence, S_{parton} can be obtained as described in Refs. [3–8]. Let us emphasise that the assumption of the maximal entanglement is expressed through the probability $p(n) = \frac{1}{N}$ which leads to the result in Eq. (1).

Now, instead of assuming this distribution, let us derive it from the principle of maximum entropy (POME) using the variational method. The functional to be maximized is given by

$$\mathcal{F}_0[p, \alpha] = - \int_0^\infty p(n) \ln(p(n)) dn + \alpha \left(\int_0^\infty p(n) dn - 1 \right), \quad (2)$$

where the first term is Shannon entropy and the second term, with α being a Lagrange multiplier, enforces the normalization condition, ensuring that $p(n)$ is a probability distribution. Since it is the least constrained scenario, we shall denote it with the subscript ‘0’. By performing the variation and utilizing the constraint, the probability can be expressed as

$$\frac{\delta \mathcal{F}}{\delta p} \Rightarrow p(n) = e^{\alpha-1}, \quad e^{\alpha-1} \int_0^\infty dn = 1 \Rightarrow p(n) = \frac{1}{N}. \quad (3)$$

Entropy is chosen as the fundamental concept, and by maximizing it, the probability distribution (the uniform distribution) of the initial-state partonic system can be derived. Now, let us see how the final-state distributions can be obtained with an additional constraint.

2.2. Final-state hadron distributions

The same calculation can be repeated to derive the final-state multiplicity distributions, however, an additional constraint needs to be introduced. In the measurements, the multiplicities are observed to have a mean (denoted by μ). Therefore, an additional constraint can be included using a Lagrange multiplier β , as follows:

$$\mathcal{F}[p, \alpha, \beta] = \mathcal{F}_0[\alpha, p] + \beta \left(\int_0^\infty np(n) dn - \mu \right), \quad (4)$$

where $\mathcal{F}_0[\alpha, p]$ is defined in Eq. (2). Using the variational method, the probability can be calculated as previously

$$p(n) = \lambda e^{-\lambda n} \Rightarrow S_{\text{hadron}} = 1 - \ln(\lambda), \quad (5)$$

where $\lambda = 1/\mu$ is the rate parameter of the distribution. This result can be considered to be a prediction and compared to the data with a free parameter to be fitted. Let us point out that the mean is constant, which in reality, is a good approximation for measurements performed in narrow rapidity ranges.

LHCb measured [17] charged-particle multiplicities in disjoint, fixed-width rapidity windows of 0.5 units, selected within the range of $2 < \eta < 4.5$. This approach provides the multiplicity distributions, hence their entropy as a function of rapidity. ALICE, ATLAS, CMS (see *e.g.* in Refs. [9–11]) as well as earlier experiments such as, *e.g.*, at UA5 [18] or E665 [19], measured multiplicities cumulatively over wider and wider rapidity intervals, typically centered at $\eta = 0$.

3. Comparison to data

If the charged-particle multiplicity distributions are measured in narrow rapidity windows such that the $\mu \approx \text{const.}$ approximation holds, then Eq. (5) can be fitted to the data. A comparison with the LHCb measurement [17] was performed, and an example fit is shown in the left panel of Fig. 1. The extracted Shannon entropy is presented in the right panel of Fig. 1. (Comparison with data measured using expanding rapidity windows is also possible but requires further developments of the presented methods, as discussed in Section 4.)

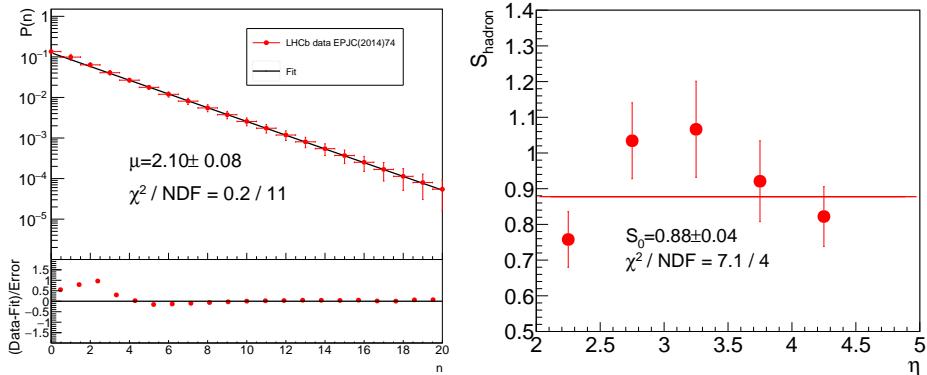


Fig. 1. An example fit of the multiplicity distributions (left) measured by LHCb [17] in $2 < \eta < 2.5$ range of rapidity. The fit excludes points $n < 3$. The extracted entropies as functions of rapidity (right).

The distribution derived from the POME can describe the data, and the extracted entropy is approximately constant. The parameter values and the χ^2/NDF are given in the plot. The POME-based description of the multiplicity distributions is statistically acceptable. The entropy is observed to be constant which is in agreement with the H1 result [12].

4. Towards large rapidity intervals

The above derivation assumed that the mean is constant which is only valid in a narrow rapidity interval. To describe distributions measured in wider rapidity intervals, two main approaches can be considered. One can compose multiple exponential distributions from Eq. (5) into a single distribution. Another way is to generalize the constraint so the mean depends on the variable as $\mu = \mu(n)$. In this paper, we follow the first approach, while the second is left for discussion in a future work.

Let us consider that within a narrow rapidity window, the exponential distribution can describe the measured multiplicity. To extend the rapidity range to twice its width, one could calculate the convolution of the distribution with itself. Repeating this process N times would ultimately lead to the Gamma distribution, denoted by $\Gamma(k = N, \lambda)$, where $\lambda = 1/\mu$ is the rate parameter, and k is so-called the shape parameter. The Gamma distribution is known to be the scaling function of the negative binomial distribution. Indeed, with the Poisson–Gamma mixture or Poisson transformation, which transforms a continuous distribution into a discrete one, it can be shown that if $Y \sim \Gamma(k, \lambda)$ and $X \sim \text{Poisson}(\mu)$ as $\mu = Y$, then

$$P(X = x|\mu) = \text{NBD}(k, \bar{n} = k/\lambda) . \quad (6)$$

Consequently, it is shown that a natural explanation for the widely observed negative binomial distribution (NBD) shape of charged-particle multiplicities emerges from the POME-based derivation: the origin of the NBD shape lies in the convolution of maximal entropy distributions repeated $k = N$ times. This argument is more general than the well-known clan model [20]. Let us examine the relation between the parameters of the Gamma distribution and those of the NBD

$$k_\Gamma = k_{\text{NBD}} , \quad \lambda_\Gamma = \frac{k_{\text{NBD}}}{\bar{n}} . \quad (7)$$

These relations have three consequences. First, k should depend linearly on the width of the rapidity window; it is observed. Second, the entropy should saturate. It is expected that the entropy does not increase indefinitely as more phase space is covered — this saturation behaviour was already observed in the UA5 measurements [18, 21].

Finally, k should be a positive integer as it is the number of underlying or convoluted exponential distributions. However, this behaviour is not confirmed by the experimental results. The reason could be short- and long-range correlations that would reduce entropy but were not considered in the derivation. Also, by convoluting exponential distributions, tacitly, a resolution is defined, which may not be valid across all regions. It is also important to note that the negative binomial parameters are strongly correlated, which may be a consequence of the relationship shown in Eq. (7).

5. Summary

The derivation of the charged-particle multiplicity distributions was presented based on the principle of maximal entropy. The derived distributions have a well-known functional form that describes the data well in a narrow rapidity ranges, where the constant mean approximation can be considered

to be valid. The results were compared to the LHCb data and Shannon's entropy was extracted, found to be consistent with a constant value. This is in agreement with previous observations in deep inelastic e^-p scattering by H1. The approach was then extended to wide rapidity ranges and was shown to explain the negative binomial shape of the measured distributions.

S.L. is grateful for the support of the Miniatura 7 2023/07/X/ST2/00652 grant of the National Science Centre (NCN), Poland. S.L. thanks for the hospitality and insightful discussions with Mateusz Płoskoń and Jesus Guillermo Contreras Nuno. S.L. would like to thank to Krzysztof Kutak, Martin Hentschinski, and Jacek Otwinowski for the fruitful discussions.

REFERENCES

- [1] D.E. Kharzeev, E.M. Levin, *Phys. Rev. D* **95**, 114008 (2017).
- [2] O.K. Baker, D.E. Kharzeev, *Phys. Rev. D* **98**, 054007 (2018).
- [3] Z. Tu, D.E. Kharzeev, T. Ullrich, *Phys. Rev. Lett.* **124**, 062001 (2020).
- [4] M. Hentschinski, K. Kutak, *Eur. Phys. J. C* **82**, 111 (2022); *Erratum ibid.* **83**, 1147 (2023).
- [5] D.E. Kharzeev, E. Levin, *Phys. Rev. D* **104**, L031503 (2021).
- [6] K. Zhang *et al.*, *Phys. Rev. D* **105**, 014002 (2022).
- [7] M. Hentschinski, K. Kutak, R. Straka, *Eur. Phys. J. C* **82**, 1147 (2022).
- [8] M. Hentschinski *et al.*, *Phys. Rev. Lett.* **131**, 241901 (2023).
- [9] CMS Collaboration (V. Khachatryan *et al.*), *J. High Energy Phys.* **2011**, 079 (2011).
- [10] ATLAS Collaboration (G. Aad *et al.*), *Eur. Phys. J. C* **76**, 403 (2016).
- [11] ALICE Collaboration (S. Acharya *et al.*), *Eur. Phys. J. C* **77**, 852 (2017).
- [12] H1 Collaboration (V. Andreev *et al.*), *Eur. Phys. J. C* **81**, 212 (2021).
- [13] E.T. Jaynes, *Phys. Rev.* **106**, 620 (1957).
- [14] E.T. Jaynes, *Phys. Rev.* **108**, 171 (1957).
- [15] E.D. Malaza *et al.*, *Phys. Lett. B* **266**, 169 (1991).
- [16] F.S. Navarra, O.V. Utyuzh, G. Wilk, Z. Włodarczyk, *Physica A* **344**, 568 (2004).
- [17] LHCb Collaboration (R. Aaij *et al.*), *Eur. Phys. J. C* **74**, 2888 (2014).
- [18] UA5 Collaboration (R.E. Ansorge *et al.*), *Z. Phys. C* **43**, 357 (1989).
- [19] E665 Collaboration (M.R. Adams *et al.*), *Z. Phys. C* **61**, 179 (1994).
- [20] A. Giovannini, S. Lupia, R. Ugoccioni, *Nucl. Phys. B Proc. Suppl.* **25**, 115 (1992).
- [21] V. Šimák, M. Šumbera, I. Zborovský, *Phys. Lett. B* **206**, 159 (1988).