



Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

ScienceDirect

Nuclear Physics A 1036 (2023) 122670



[www.elsevier.com/locate/nuclphysa](http://www.elsevier.com/locate/nuclphysa)

# Information entropy and fragmentation functions

Guillermo Benito-Calviño, Javier García-Olivares,  
Felipe J. Llanes-Estrada\*

*Univ. Complutense de Madrid, Fac. Ciencias Físicas, Dept. Física Teórica e IPARCOS, Plaza de las Ciencias 1,  
Madrid, 28040, Spain*

Received 1 October 2022; received in revised form 26 February 2023; accepted 28 April 2023  
Available online 9 May 2023

---

## Abstract

Several groups have recently investigated the flow of information in high-energy collisions, from the entanglement entropy of the proton yielding classical Shannon entropy of its parton distribution functions (pdfs), through jet splitting generating entropy, to the entropy distribution in hadron decays.

Lacking in the literature is a discussion of the information entropy of fragmentation functions (FFs) in the instances where they can be considered as probability distributions, and we here provide it. We find that this entropy is a single, convenient number to characterize future progress in the extraction of fragmentation functions.

We also deploy the related Kullback-Leibler divergence between two distributions to assess existing relations among FFs and parton distribution functions (pdfs) such as that of Barone, Drago and Ma. From a couple of current parametrizations of FFs, we do not find supporting empirical evidence for the relation, although it is possible that FFs and pdfs have similar power-laws near the  $x = 1$  endpoint.

© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

*Keywords:* Fragmentation functions; Parton distribution functions; Shannon entropy; Classical information entropy; Kullback-Leibler divergence; Barone-Drago-Ma relation

---

---

\* Corresponding author.

*E-mail address:* [fllanes@fis.ucm.es](mailto:fllanes@fis.ucm.es) (F.J. Llanes-Estrada).

<https://doi.org/10.1016/j.nuclphysa.2023.122670>

0375-9474/© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

### 1.1. Information entropy in particle physics

The rapid expansion of quantum information theory [1] has made groups in high-energy physics aware of its possible uses to think about particle accelerator data. Several applications include the distinction among conventional hadrons and exotica [2], the exploration of supersymmetric models [3], predictions on exotic decays of the Higgs boson [4] based on the observation that the Higgs-boson mass at 125 GeV plays a special role by maximizing the product of decay probabilities [5], which lead to the formulation of a maximum-entropy hypothesis for the Higgs boson [6] or the analysis of heavy-ion collisions [7].

A case of particular interest is information entropy in high-energy collisions. The initial states of collider experiments are usually prepared in pure quantum states where a few particles (for example, an electron and a proton or two hadrons) have well defined momenta. However, when they collide, cross-sections can often be described in collinear factorization in terms of parton distribution functions that have a classical probability distribution. The conversion of the entanglement entropy in the initial state wavefunction to the final pdf classical information entropy has been authoritatively described a short time ago [8], and studies continue [9]. If a jet forms in such collisions, entropy evolves [10] with the falling  $Q^2$  through successive splittings (in an evolution that reminds us of out-of-equilibrium physics with  $Q^2$  playing the role of time). Entropy is also generated as the hadrons produced in the final state, often unstable, successively decay to lighter, more stable ones [11,12].

What seems to be missing in the literature is a discussion of the fragmentation functions leading to an identified final-state hadron after the collision at the parton-level has run its course and hadronization occurs, that is, a study of the entropy of Fragmentation Functions. In this work we will carry out a small exploration thereof. We will examine the Shannon entropy,

$$S = - \sum_i p_i \log p_i \quad (1)$$

in section 2, where we report calculations on the known fragmentation functions, that we briefly recall in subsection 1.2; our estimates provide upper bounds for this  $S$ , and we try to model how the future analysis of fragmentation hadrons may change those bounds by educated guessing. In section 3 we turn to a comparison of FFs and pdfs, both direct and via the Kullback-Leibler divergence, related to the entropy and defined in subsec. 2.1.

In brief, the entropy of Eq. (1) is the expectation value of the information in a probability distribution. Given a set of  $I$  objects with two possible states (*e.g.* open/closed boxes), or bits of information, the probability of a configuration scales as  $p = (\frac{1}{2})^I$ , so that the information gained by measuring one of the bits can be quantified by the number  $I = -\log_2(p_i)$ . The average of this information gauge over the entire distribution, upon changing the base-2 logarithm to base  $e$  (with information measured in *nats* instead of bits) is hence Eq. (1), the classical Shannon entropy, whose quantum version, the Von Neumann entropy,  $S = -\text{Tr}(\rho \log \rho)$ , is probably more familiar to subatomic physicists.

### 1.2. Fragmentation functions and parton distribution functions

Parton distribution functions are extracted from deeply inelastic scattering (DIS) experiments via the structure functions, for example

$$\frac{d\sigma(e^-h \rightarrow e^-X)}{dx dy_1} = \frac{2\pi\alpha^2}{Q^4} s[1 + (1+y_1)^2 F_2(x)] \quad (2)$$

that are given in terms of the pdfs  $f_i$  for the particles and  $\bar{f}_i$  for the antiparticles by

$$F_2(x) = \sum_i e_i^2 x [f_i(x) + \bar{f}_i(x)], \quad (3)$$

with  $x$  being the same as  $x_i = p_i/P$  the momentum fraction of the hadron carried by the  $i^{\text{th}}$  parton, and  $e_i$  its charge. An important sum rule that pdfs satisfy is that induced by the addition of each parton's momentum to yield the hadron's total,

$$\sum_i \int x f_i(x) dx = 1. \quad (4)$$

This sum rule can be interpreted as the normalization of a continuous probability distribution given by  $p(i, x) := x f_i(x)$ . The flavor-sum rules could be chosen instead, but in comparing with fragmentation functions, this momentum sum rule seems more straightforward to us.

Analogously, fragmentation functions [13] can be extracted, for example, from the cross-section for electron-positron collisions

$$\frac{d\sigma(e^-e^+ \rightarrow hX)}{dz} = \frac{4\pi\alpha^2}{3Q^2} z^2 \hat{F}_2(z) \quad (5)$$

that depend on structure functions analogous to those of Eq. (3), given in terms of the fragmentation functions,

$$\hat{F}_2(z) = -3 \sum_q e_q^2 \frac{D^q(z) + D^{\bar{q}}(z)}{z^2}. \quad (6)$$

Here,  $D^q(z)$  represents, in the parton model, the probability that a parton  $q$  fragments into the hadron being considered in the final state,  $h$ , which carries a momentum fraction  $z = p_h/p_q$ . These fragmentation functions also satisfy a momentum sum rule,

$$\sum_h \int_0^1 z D_h^q(z) dz = 1. \quad (7)$$

The sum runs over all possible hadrons in which the parton can fragment, and the integral over all possible momentum fractions must yield the original parton momentum, that is, 1 upon normalizing by  $p_q$ .

Parton distribution functions are well known, but for reference we plot in Fig. 1 one of the recent determinations of the less widely discussed fragmentation functions, obtained from neural network fits [14–16] to various experimental data, showing their current large uncertainty bands. Further interesting sets have been reported in the literature [17, 18] but we limit ourselves to relatively low-orders in QCD corrections to avoid the Wilson coefficients of the perturbative expansion forcing the nonperturbative distribution function to become negative, hiding its nature as a parton probability density, which is asymptotically a correct interpretation, but can be problematic at finite  $Q^2$  when QCD corrections are included.

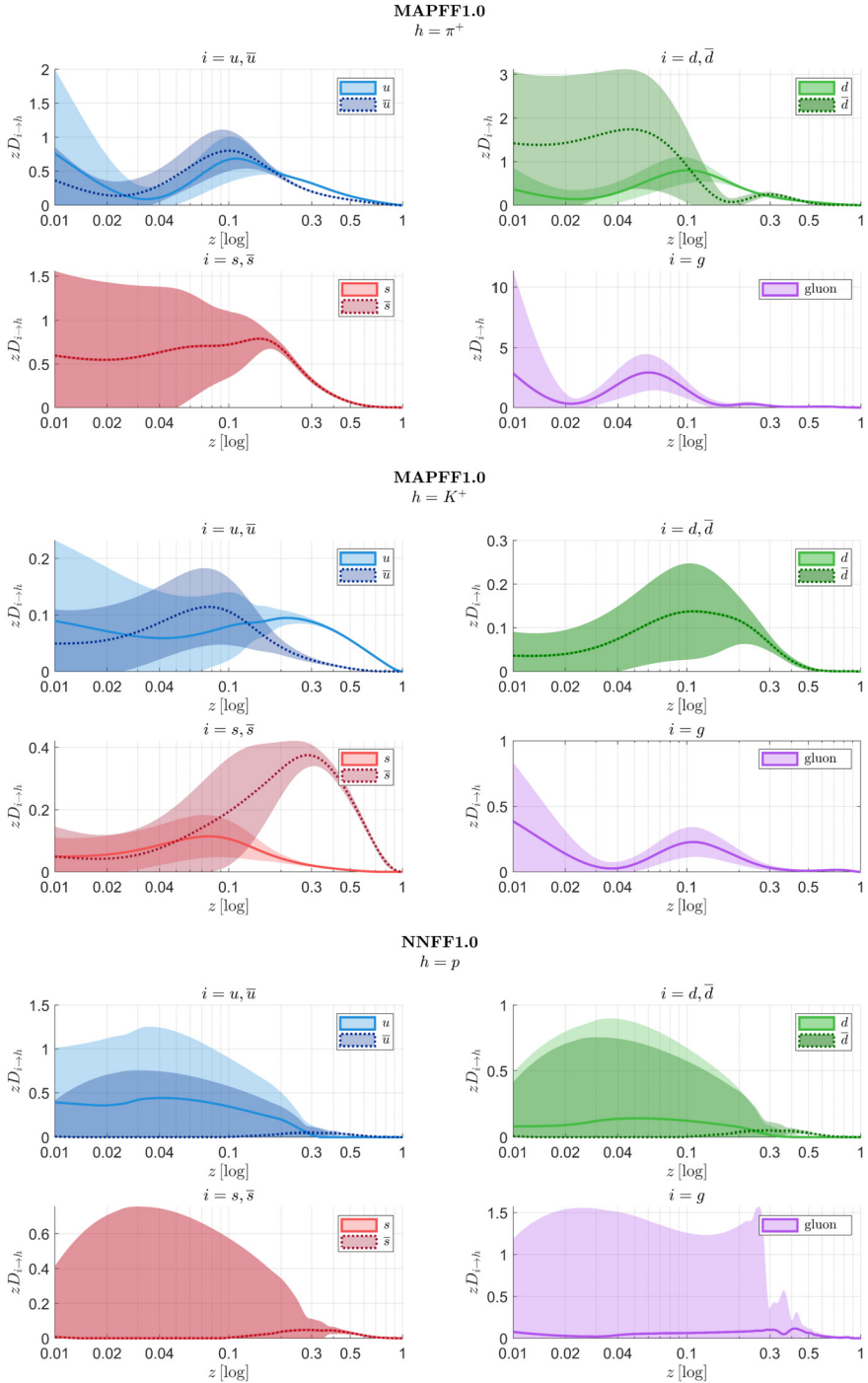


Fig. 1. Fragmentation functions for a pion (top panel), a kaon (middle panel) and a proton (lowest panel) fragmenting from the indicated parton, plot from NNPDF and MAPFF public fits [14–16].

### 1.3. Relations among pdfs and FFs

The similarity between pdfs and FFs is due to the basic physics processes that they represent,  $h \rightarrow q + X$  and  $q \rightarrow h + X$ , that are clearly related by crossing. This did not escape early authors [19], who noted that the two processes, deep inelastic scattering and fragmentation in  $e^-e^+$ , were related by analytic continuation in the kinematic variables.

In current notation, the resulting relation of Drell-Levy-Yan that relates the fragmentation function  $D$  with the parton distribution function  $f$  (for the same hadron and the same parton) reads

$$D(z) = z f\left(\frac{1}{z}\right). \quad (8)$$

This relation is not particularly useful, since for  $z \in [0, 1]$ ,  $\frac{1}{z} \in [1, \infty)$  is unphysical except at  $z \rightarrow 1$ , so that it relates each of the functions with an analytic extension of the other. (Perhaps this could be tested on the deuteron or with CLAS data [20], where  $z > 1$  might be at hand, but we have not investigated it as this has not entered the mainstream, where  $z \in [0, 1]$ .)

As discussed by Barone, Drago and Ma [21], Gribov and Lipatov [22] found a relation between splitting kernels that is usually taken to imply a relation between pdfs and FFs, the so called Gribov-Lipatov “reciprocity” relation

$$3D(z) = f(z) \quad (9)$$

but whose theoretical status should rather be considered that of a model.

Barone, Drago and Ma [21] presented then a more suitable expression

$$D(z) \simeq z f\left(2 - \frac{1}{z}\right) \quad (10)$$

that relates the pdf and the FF within their physical domains, that should become exact in the limit of  $z \rightarrow 1$ , and that apparently should be a good approximation for, say,  $z > 0.6$ . Two decades back the scarce quality data did not allow for a good test of this relation, and we use this opportunity to employ modern fits of the involved functions to see whether any progress has been made in empirically establishing this relation.

## 2. Entropy of parton distribution or fragmentation functions

### 2.1. Entropy of continuous distributions

A technical problem that arises in directly applying Eq. (1) for the Shannon information entropy to a continuous probability distribution is a failure of the simplest discretization. When taking the continuous probability distribution (here  $D(z) = \sum_i D_i(z)$  for a given parton but summed over final-state hadrons) and mapping it into an approximate, discrete probability set,  $zD(z) \rightarrow \{p_1, p_2, \dots, p_i, \dots, p_N\}$ ;  $\sum_i p_i = 1$  we find a divergence with the cardinal  $N$  of that set (the number of discretization points),

$$S = - \sum_{i=1}^N (p_i \log p_i) \underset{N \rightarrow \infty}{\sim} \log N \xrightarrow{N \rightarrow \infty} \infty. \quad (11)$$

There is a simple generalization of Eq. (1) to a continuum distribution  $F(x)$  (that can represent either  $xf(x)$  for a pdf, or  $zD(z)$  for an FF),

$$S[F] := - \int F(x) \ln F(x) dx \tag{12}$$

but because continuous probability densities can exceed the value  $F = 1$  in a small-enough interval, there can be contributions to this entropy that are negative. This means that there can be occasional sets that for whatever reason give a negative overall value, which is very counterintuitive for a variable representing any sort of entropy.

To interpret Eq. (12), we need to resort to the Kullback-Leibler divergence [23]. This divergence takes as argument a second distribution  $Q$  and tries to quantify how much does the first,  $P$ , differ from this  $Q$ , taken as reference:

$$D_{\text{KL}}(P \parallel Q) := \sum_x p(x) \log \frac{p(x)}{q(x)} . \tag{13}$$

It readily admits the continuum generalization

$$D_{\text{KL}}(F \parallel G) = \int f(x) \log \frac{f(x)}{g(x)} dx \tag{14}$$

and then, a typical continuum expression such as Eq. (12) can be understood as the Kullback-Leibler divergence of the  $F(x)$  distribution to the uniform one (which is that with least information).

### 2.2. Evolution equations and the choice of distribution and associated entropy

In Quantum Chromodynamics, scaling violations introduce a  $Q^2$  scale dependence in fragmentation functions, that become  $D_i^h(z, Q^2)$ . The  $Q^2$  dependence follows from an evolution equation that reads, analogously to the Altarelli-Parisi equations but with transposed splitting kernels  $P_{ji}$  to deal with fragmentation functions, and omitting hadron indices for notation conciseness,

$$\frac{\partial}{\partial \log Q^2} D_j(x, Q^2) = \sum_j \int_x^1 \frac{dz}{z} P_{ji}(z, \alpha_s(Q^2)) D_i(x/z, Q^2) . \tag{15}$$

The simplest of such equations is that for a gluon at low  $x \simeq 0$ , which suppresses the quark splitting, reading

$$\frac{\partial D_g(x, Q^2)}{\partial \log Q^2} = \frac{2N_c \alpha_s}{2\pi} \int_x^1 \frac{dz}{z} \left( \frac{1}{z} \right) D_g(x/z, Q^2) . \tag{16}$$

This evolution is linear in  $D_g$ , that yields exponential growth in  $\log Q^2$ , that is, power-law growth in  $Q^2$  at low  $x$  with anomalous dimension in the literature [24–26] for further details. Because of this growth at low  $x$  and the momentum sum rule of Eq. (7), the function  $zD(z)$  becomes peaked at low  $x$  for large  $Q^2$ , just as parton distribution functions do.

This evolution forces the entropy to also depend on  $Q^2$  (as if it were a nonequilibrium evolution in a real-time variable  $t$  instead of  $Q^2$ ), as can be seen by taking a derivative of Eq. (12) respect to  $\log Q^2$ , and employing the momentum sum rule, to obtain the evolution of the entropy with the scale,

$$\frac{\partial S}{\partial \log Q^2} = - \sum_h \int_0^1 dz (zD(z)) \log(zD(z)) \left( \frac{\partial \log D(z)}{\partial \log Q^2} \right) \quad (17)$$

with  $D(z) = D(z, Q^2)$ . It is obvious that as long as  $D(z, Q^2)$  keeps running, the entropy cannot become extremal over the  $Q^2$  trajectory.

Thus,  $S(Q^2 \rightarrow \infty)$  is not, either for asymptotic  $Q^2$  nor at any value of  $Q^2$  that we can think of, maximal respect to other possible distributions. The function that would maximize the value of  $S$  would be  $zD_q(z) = 1$  in the case of only one hadron species; (or the relative multiplicity of a given hadron in the case that more than one was considered,  $zD_q^h(z) = N_h/N_{\text{total}}$ ). Setting the parton as a gluon, to be specific, this would be

$$D_g^h(z) = \frac{c}{z} \quad (18)$$

with  $c$  a constant. Now Eq. (18) is neither a solution nor a limit of the QCD evolution equation (16).

In fact, a function that increasingly peaks at low- $x$  will evolve to have arbitrarily negative continuum entropy from Eq. (12). This is easily seen by setting  $zD(z) \sim \frac{1}{x_0} \theta(x_0 - x)$ , a top-hat function that peaks towards  $x = 0$  with decreasing  $x_0$ , with entropy  $-\log(1/x_0)$ . Ultimately this is due to the certainty that at high  $Q^2$  the fragmenting parton will carry an infinitesimal share of the total hadron momentum, so that there is no surprise upon measuring it.

An alternative probability function can be constructed from the observation that, once the momentum fraction has been chosen, the parton has to fragment into one among the existing hadrons, with probability  $p_i$ . One could think of constructing an entropy based on the fragmented hadron probability distribution. However, the theory here is less clear than the reciprocal one with a fixed hadron, the nucleon, and a sum over partons extracted from it in ([8]): Kharzeev and Levin were able to show that the proton is in a maximally entangled state at low- $x$ .

The most we could do with fragmentation functions is to note that at high energy, perhaps the low- $x$  sea presents universality properties since all hadrons appear in the expansion  $|N\rangle + |N\pi\rangle + |N\pi\pi\rangle + |K\Lambda\rangle + \dots$  so that the parton can equally well fragment into any of them. The sum over equidistributed hadrons could then be a maximum of entropy. But given the fragility of such arguments, and the fact that the entropy coming from a sum over hadrons does not convey much information about the shape of the fragmentation function itself (of interest for phenomenological applications), we find it more convenient to focus on the continuum momentum distribution instead of the hadron-flavor one.

### 2.3. Limitation to an upper bound on the entropy when data is insufficient

To saturate the momentum sum rule in Eq. (7), the sum needs to be extended over *all hadrons*  $h$  that can fragment out of the parton  $q$  under consideration.

But experimental efforts have been at best identifying a part of the low-lying hadrons (in practice,  $\pi^\pm$ ,  $K^\pm$ ,  $p$  and little more), let alone all the unstable ones that decay onto these and never reach the detector, so how can an appropriately normalized distribution function be built with insufficient data?

Our approach [11] is to split the fragmentation function into a piece that sums over all those hadrons that are specifically identified in a given experiment, and one that is supposed to represent the unidentified ones, perhaps even unknown, and that completes the normalization of the momentum sum rule to be unity,



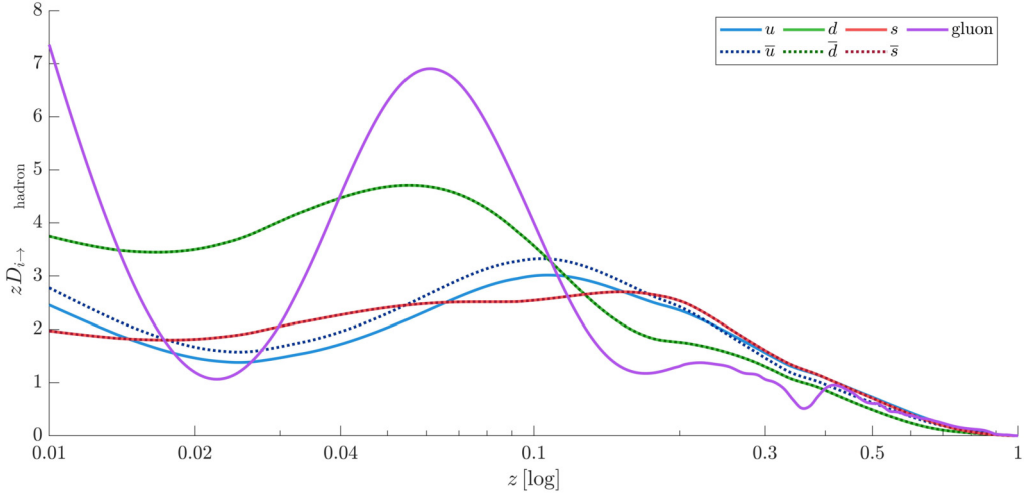


Fig. 3. Reconstructed aggregated fragmentation functions summing over all low-lying hadrons following the rules of Table 1. This is an educated guess of what future experimental progress could bring; upon summing over hadron species, the remaining information is of course only about a momentum distribution, and the bound over the entropy in Fig. 2 should hopefully be converging to the actual entropy of that distribution.

The contribution to the entropy of Eq. (12) due to the identified charged hadrons, with NNPDF fits for  $D_q^h$ , is shown as the first step to the left of the figure. Each possible fragmenting parton is distinguished by a different type of line as specified in the caption. For this first step, Eq. (20) has been used.

The figure is then organized so that hadrons from the lowest-mass baryon octet and meson nonet (both  $\eta$  and  $\eta'$ , mixed with  $\theta(\eta, \eta') = -15.5^\circ$ , are included [27]) appear towards the right, broadly ordered by increasing mass.

In simulating how the upper bound on the entropy might descend with future knowledge, we have coarsely represented each successive function  $D_h^q$  by a known fragmentation function, assigning them employing  $SU(3)$  relations linking the various flavors of hadrons and quarks, and employing the customary valence/sea separation. For example, the as of yet unmeasured fragmentation of a down quark to a neutral Kaon  $D_{d \rightarrow K^0}$ , is set to be equal to that of an up quark fragmenting to a positively charged kaon  $D_{u \rightarrow K^+}$  (already extracted). An extensive list of how these relations were applied can be found in Table 1.

In view of Fig. 2, we therefore propose the upper bound on the entropy  $S$  as a salient number that quantifies progress on each of the fragmentation functions.

For completeness, we plot in Fig. 3 the “reconstructed” aggregate fragmentation function that sums over the hadron species just discussed.<sup>1</sup> That aggregated reconstruction eliminates (sums over) hadron flavor information to display only the distribution of the hadron momentum fraction taken from the fragmenting parton.

<sup>1</sup> How each of them contributes to that distribution can be seen in the sequence of graphs reported to the XVth International Conference on Quark Confinement and the Hadron Spectrum, [downloadable at its website](#).

Table 1

Relations used to recreate the fragmentation functions into light mesons and baryons, starting from a few ones that are already measured, those of  $\pi^+$ ,  $K^+$  and  $p$ . Those with a shaded background are the input ones, whereas the majority without a background are related to those as an educated guess to compute the previsible improvement to the upper bound on the entropy. Color online: we employ red for valence quark combinations, blue for sea quarks, orange for partons with additional suppression, and grey is reserved for gluons. We omit the  $\pi^-$ ,  $K^-$ ,  $\bar{K}^0$  antiparticles and all antibaryons, that are immediately obtained from charge conjugation such as  $D_{q \rightarrow h} = D_{\bar{q} \rightarrow \bar{h}}$ .

$h$	Quark model wavefunction	$D_{u \rightarrow h}$	$D_{\bar{u} \rightarrow h}$	$D_{d \rightarrow h}$	$D_{\bar{d} \rightarrow h}$	$D_{s \rightarrow h}$	$D_{\bar{s} \rightarrow h}$	$D_{g \rightarrow h}$
$\pi^0$	$\frac{1}{\sqrt{2}}(u\bar{u} + d\bar{d})$	$\frac{1}{\sqrt{2}}D_{u \rightarrow \pi^+}$	$\frac{1}{\sqrt{2}}D_{u \rightarrow \pi^+}$	$\frac{1}{\sqrt{2}}D_{\bar{d} \rightarrow \pi^+}$	$\frac{1}{\sqrt{2}}D_{\bar{d} \rightarrow \pi^+}$	$D_{s \rightarrow \pi^+}$	$D_{s \rightarrow \pi^+}$	$D_{g \rightarrow \pi^+}$
$\pi^+$	$u\bar{d}$	$D_{u \rightarrow \pi^+}$	$D_{d \rightarrow \pi^+}$	$D_{d \rightarrow \pi^+}$	$D_{\bar{d} \rightarrow \pi^+}$	$D_{s \rightarrow \pi^+}$	$D_{s \rightarrow \pi^+}$	$D_{g \rightarrow \pi^+}$
$K^+$	$u\bar{s}$	$D_{u \rightarrow K^+}$	$D_{s \rightarrow K^+}$	$D_{d \rightarrow K^+}$	$D_{d \rightarrow K^+}$	$D_{s \rightarrow K^+}$	$D_{\bar{s} \rightarrow K^+}$	$D_{g \rightarrow K^+}$
$K^0$	$d\bar{s}$	$D_{d \rightarrow K^+}$	$D_{d \rightarrow K^+}$	$D_{u \rightarrow K^+}$	$D_{s \rightarrow K^+}$	$D_{s \rightarrow K^+}$	$D_{\bar{s} \rightarrow K^+}$	$D_{g \rightarrow K^+}$
$\eta$	$c_1(u\bar{u} + d\bar{d})$ $+c_2s\bar{s}$	$\frac{c_1}{\sqrt{2}}D_{u \rightarrow \pi^+}$ $+c_2D_{d \rightarrow K^+}$	$\frac{c_1}{\sqrt{2}}D_{u \rightarrow \pi^+}$ $+c_2D_{d \rightarrow K^+}$	$\frac{c_1}{\sqrt{2}}D_{\bar{d} \rightarrow \pi^+}$ $+c_2D_{d \rightarrow K^+}$	$\frac{c_1}{\sqrt{2}}D_{\bar{d} \rightarrow \pi^+}$ $+c_2D_{d \rightarrow K^+}$	$c_1D_{s \rightarrow \pi^+}$ $+c_2D_{\bar{s} \rightarrow K^+}$	$c_1D_{s \rightarrow \pi^+}$ $+c_2D_{\bar{s} \rightarrow K^+}$	$D_{g \rightarrow \pi^+}$
$\eta'$	$c'_1(u\bar{u} + d\bar{d})$ $+c'_2s\bar{s}$	$\frac{c'_1}{\sqrt{2}}D_{u \rightarrow \pi^+}$ $+c'_2D_{d \rightarrow K^+}$	$\frac{c'_1}{\sqrt{2}}D_{u \rightarrow \pi^+}$ $+c'_2D_{d \rightarrow K^+}$	$\frac{c'_1}{\sqrt{2}}D_{\bar{d} \rightarrow \pi^+}$ $+c'_2D_{d \rightarrow K^+}$	$\frac{c'_1}{\sqrt{2}}D_{\bar{d} \rightarrow \pi^+}$ $+c'_2D_{d \rightarrow K^+}$	$c'_1D_{s \rightarrow \pi^+}$ $+c'_2D_{\bar{s} \rightarrow K^+}$	$c'_1D_{s \rightarrow \pi^+}$ $+c'_2D_{\bar{s} \rightarrow K^+}$	$D_{g \rightarrow \pi^+}$
$p$	$uud$	$2D_{\text{val} \rightarrow p}$	$D_{\text{sea} \rightarrow p}$	$D_{\text{val} \rightarrow p}$	$D_{\text{sea} \rightarrow p}$	$D_{\text{sea} \rightarrow p}$	$D_{\text{sea} \rightarrow p}$	$D_{g \rightarrow p}$
$n$	$udd$	$D_{\text{val} \rightarrow p}$	$D_{\text{sea} \rightarrow p}$	$2D_{\text{val} \rightarrow p}$	$D_{\text{sea} \rightarrow p}$	$D_{\text{sea} \rightarrow p}$	$D_{\text{sea} \rightarrow p}$	$D_{g \rightarrow p}$
$\Lambda^0$	$uds$	$D_{\text{val} \rightarrow p}$	$D_{\text{sea} \rightarrow p}$	$D_{\text{val} \rightarrow p}$	$D_{\text{sea} \rightarrow p}$	$D_{\text{val} \rightarrow p}$	$D_{\text{sea} \rightarrow p}$	$D_{g \rightarrow p}$
$\Sigma^+$	$uus$	$2D_{\text{val} \rightarrow p}$	$D_{\text{sea} \rightarrow p}$	$D_{\text{sea} \rightarrow p}$	$D_{\text{sea} \rightarrow p}$	$D_{\text{val} \rightarrow p}$	$D_{\text{sea} \rightarrow p}$	$D_{g \rightarrow p}$
$\Sigma^0$	$uds$	$D_{\text{val} \rightarrow p}$	$D_{\text{sea} \rightarrow p}$	$D_{\text{val} \rightarrow p}$	$D_{\text{sea} \rightarrow p}$	$D_{\text{val} \rightarrow p}$	$D_{\text{sea} \rightarrow p}$	$D_{g \rightarrow p}$
$\Sigma^-$	$dds$	$D_{\text{sea} \rightarrow p}$	$D_{\text{sea} \rightarrow p}$	$2D_{\text{val} \rightarrow p}$	$D_{\text{sea} \rightarrow p}$	$D_{\text{val} \rightarrow p}$	$D_{\text{sea} \rightarrow p}$	$D_{g \rightarrow p}$
$\Xi^0$	$uss$	$D_{\text{val} \rightarrow p}$	$D_{\text{sea} \rightarrow p}$	$D_{\text{sea} \rightarrow p}$	$D_{\text{sea} \rightarrow p}$	$2D_{\text{val} \rightarrow p}$	$D_{\text{sea} \rightarrow p}$	$D_{g \rightarrow p}$
$\Xi^-$	$dss$	$D_{\text{sea} \rightarrow p}$	$D_{\text{sea} \rightarrow p}$	$D_{\text{val} \rightarrow p}$	$D_{\text{sea} \rightarrow p}$	$2D_{\text{val} \rightarrow p}$	$D_{\text{sea} \rightarrow p}$	$D_{g \rightarrow p}$

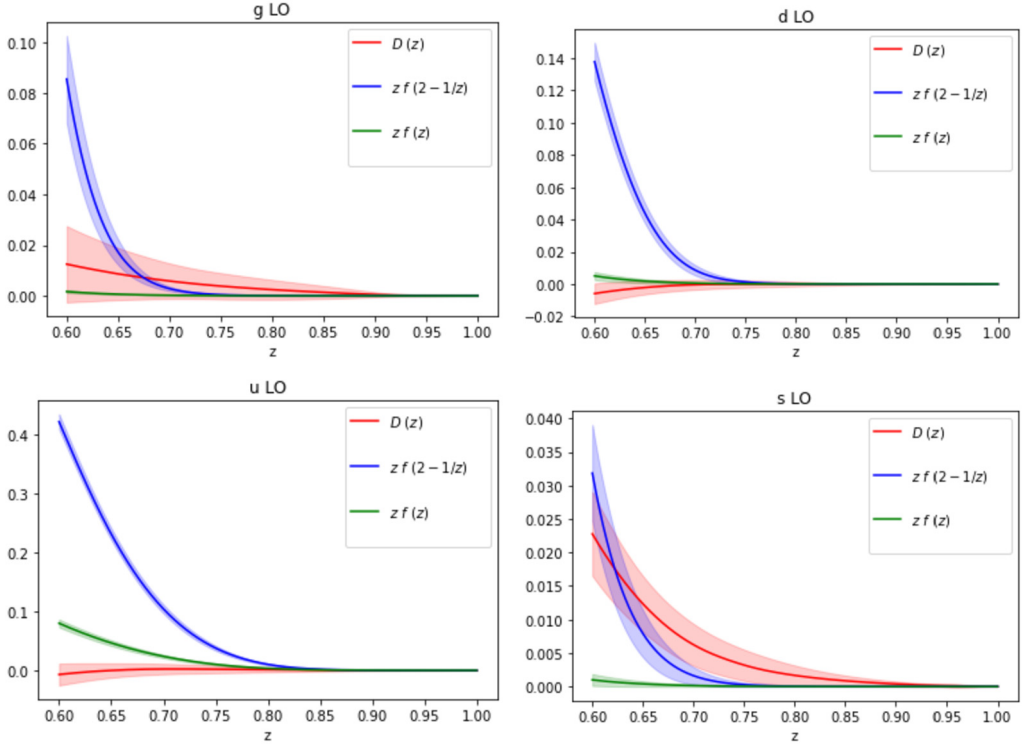


Fig. 4. Visualization of Eq. (10) and (22) for the proton.

### 3. Comparison between $D$ s and $FF$ s

#### 3.1. Graphical examination

Next, we turn to a cursory examination of Eq. (10), as well as that of its further simplification as  $z \rightarrow 1$

$$D(z) \simeq z f(z) \tag{22}$$

that comes about because

$$z = \frac{1}{1 - (1 - 1/z)} \simeq 2 - \frac{1}{z} \tag{23}$$

which holds up to two orders of the Taylor series.

In the first place, we plot all three functions for DIS on the proton and proton fragmentation (using the NNPDF parametrizations, with scale  $Q^2 = 100\text{GeV}^2$ ) in Fig. 4.

Then, in Fig. 5, we present the equivalent plots for the charged pion. Here, we have used the fit to  $\pi^+$  data, with scale  $Q^2 = 100\text{GeV}^2$ , obtained by the Jefferson Lab Angular Momentum collaboration [28] that has simultaneously fit parton and fragmentation densities, employing traditional template fits with power-law end-point behaviour  $\alpha(1 - z)^\beta$ .

In both cases we do not perceive by eyeball that either of the relations in Eq. (10) or Eq. (22) are satisfied better than the fact that all three bands smoothly converge to zero as  $z \rightarrow 1$ .

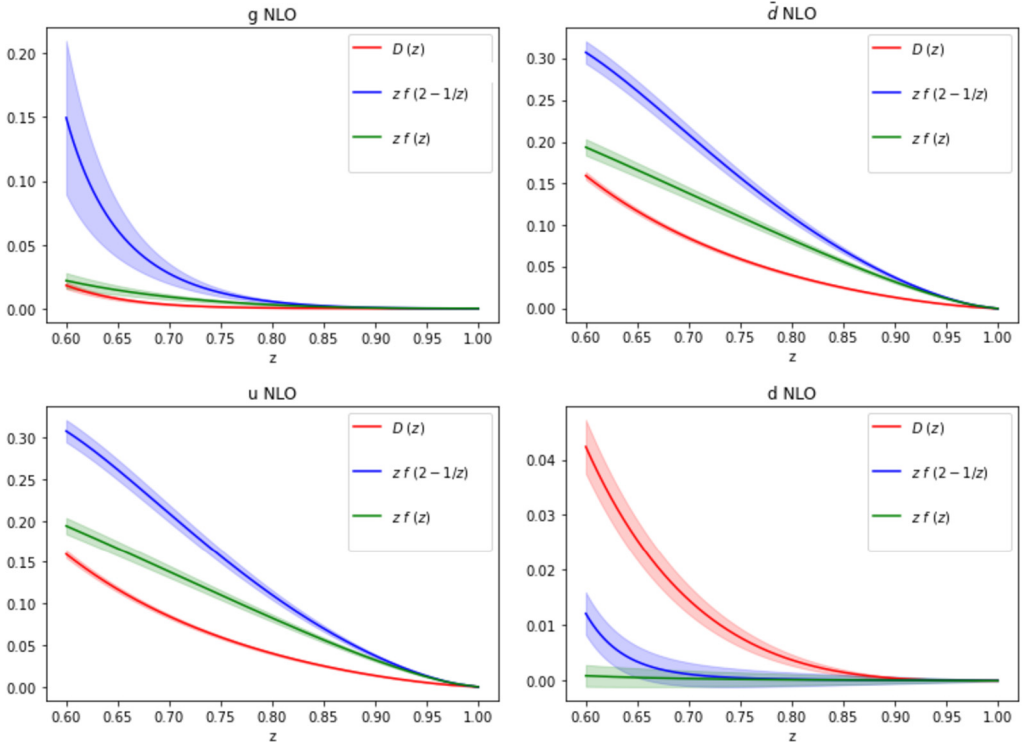


Fig. 5. Visualization of Eq. (10) and (22) for  $\pi^+/\pi^-$ .

This is corroborated by taking ratios between  $D$ s and pdfs (plot in Fig. 6). There is no definite evidence that the ratios are near unity: two of the plots have huge errors and although compatible with unity, they are inconclusive; and two are clearly incompatible, even as a limit as  $z \rightarrow 1$ , where the ratio seems to overshoot.

### 3.2. Quantification by KL divergence

To quantify this perceived separation between the fragmentation function and the corresponding small modification of the parton distribution function, we make use of the Kullback-Leibler divergence of Eq. (14). To get a feeling for its typical value, we first carry out an exercise with two simple pairs of normal probability distributions (unrelated to physical FFs) with different parameters.

First, the Kullback-Leibler divergence among

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \quad g_1(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-1)^2} \tag{24}$$

which are two distributions that overlap at  $1\sigma$ , and thus relatively likely, turns out to be  $D_{KL}(f, g_1) = 0.12$ . On the other hand, if we substitute  $g_1$  by

$$g_2(x) = \frac{1}{5\sqrt{2\pi}} e^{-\frac{1}{2}(x-5)^2}, \tag{25}$$

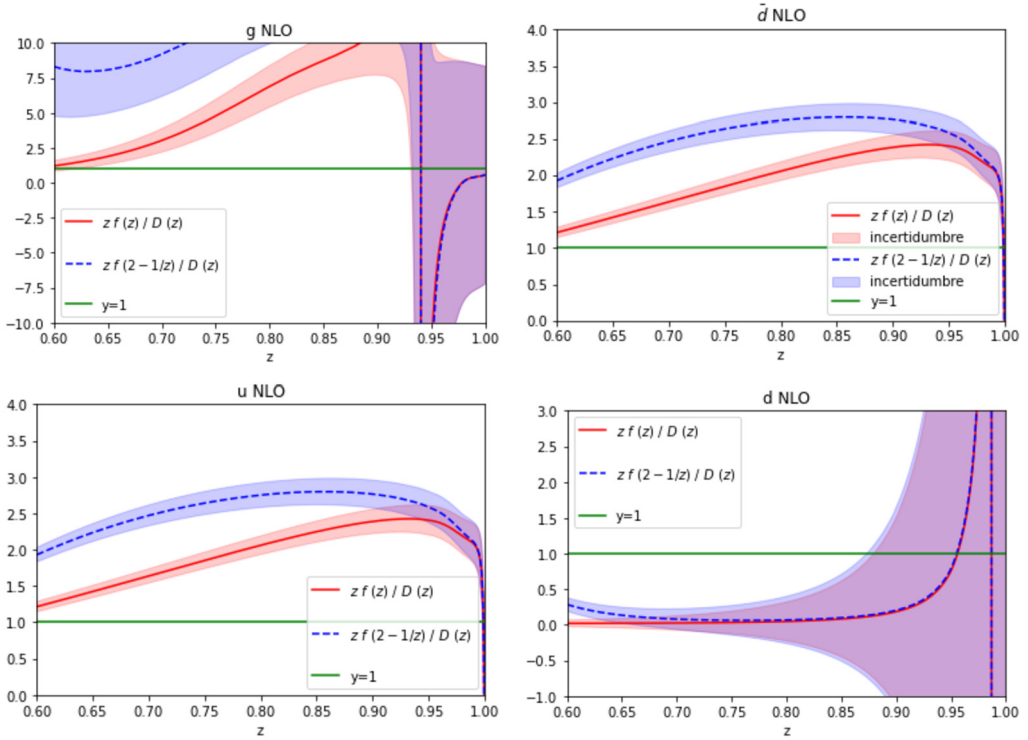


Fig. 6. Ratios of the adequate evaluation (Eqs. (10) and (22)) of the parton distribution functions and the fragmentation functions do not seem to be very near unity in the region near the endpoint  $z \rightarrow 1$ .

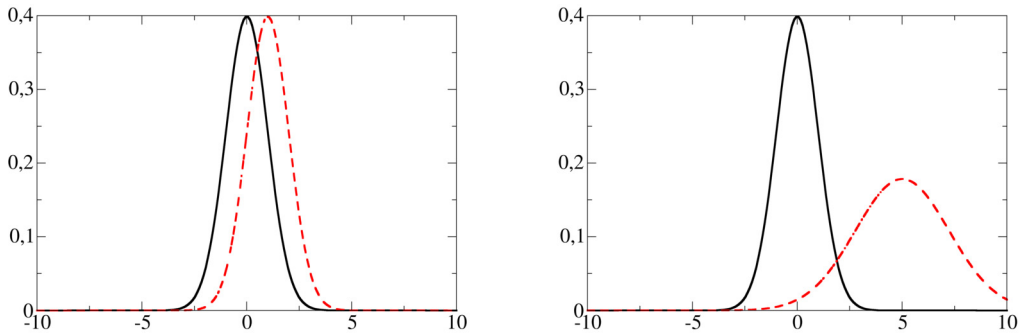


Fig. 7. Two pairs of normalized probability distributions, similar and dissimilar, as given in Eq. (24) and (25), which have respective Kullback-Leibler divergences of 0.12 and 1.63.

that is significantly different from  $f$  (both pairs of distributions are plot in Fig. 7), we have  $D_{KL}(f, g_2) = 1.63$ ; thus, we expect values of  $D_{KL}$  near 1 or above to signify that the distributions are not alike, while values well below 1 mean that the distributions are more similar.

Having a feeling for the divergence, we turn to the physical relation among fragmentation functions and parton distribution functions, with the result given in Table 2.

Table 2

Kullback-Leibler divergence for two physical relations among pdfs and FFs for the charged pion from Jefferson Lab Angular Momentum collaboration fits, considering only the integration interval  $z \in (0.7, 1)$  where the relations have been proposed to be valid. We find them to not be well satisfied.

Parton	$D_{KL}(zf_{\pi}(z)  D^{\pi}(z))$	$D_{KL}(zf_{\pi}(2-1/z)  D^{\pi}(z))$
gluon	2,28	7,70
up	1,19	2,36
anti-down	1,19	2,36
down	6,64	5,28

We see there that the Kullback-Leibler divergence gives values of order 1 or above for all considered partons (we are using the Jefferson Lab Angular Momentum (JAM) evaluations for the charged pion) so that none of these relations looks very promising at the present time.

### 3.3. Similar power laws

However, we do find, always looking at the JAM distributions [28], that the fragmentation functions do have a similarity to the parton distribution functions. This is observed in Fig. 8. In it, we once more turn to the  $\pi^+$  fragmentation functions as well as parton distribution functions from the NLO parametrizations of the JAM collaboration [28] at a scale  $Q^2 = 100\text{GeV}^2$ . We look, in particular, at the  $u$ -quark (which is a valence parton for the positively charged pion), so that we will be examining relations among  $f_{\pi}^u(x)$  and  $D_u^{\pi}(z)$ .

The top plot is a logarithmic plot of the three distribution functions discussed by Barone, Drago and Ma [21] against  $\log(z-1)$ , so that the functions run down towards the left of the plot. In such a log-log plot, a power-law appears as a straight line, and this is indeed the case. Moreover, all three are rather parallel lines, and that suggests that the failure of the Barone-Drago-Ma relations at least in certain fits of data can be due to overall multiplicative constants in the region  $z \simeq 1$ , since the distributions with the appropriate arguments seem to be coming into the end point with the same slope.

Both the original  $zf(2-1/z)$  as well as its further approximation  $zf(z)$  seem to track the fall-off of the fragmentation function as proposed by those authors. The slopes of the three quasi-straight lines shown (the power law exponent of  $\alpha_i(z-1)^{\beta_i}$ ) are, respectively, 1.48 for  $D(z)$ , against 1.52 for  $zf(2-\frac{1}{z})$ , which are rather close, and even a not too far 1.3 for  $zf(z)$ .

Thus, a multiplicative constant seems to be the difference between these distributions. This can be appreciated in plain sight in the lower plot of Fig. 8, where an *ad-hoc* constant multiplies the FF and puts the three lines on top of each other. The Kullback-Leibler divergence between pairs of functions sees a significant drop, as expected.

We presently ignore what the theoretical reason for this may be, but we hope it can be studied in the future by the collaborations addressing fragmentation functions and parton distribution functions in a unified manner.

## 4. Conclusions and outlook

In this contribution we have deployed two very well-known estimators of information theory, the Shannon information entropy and the Kullback-Leibler divergence, to assess fragmentation functions.

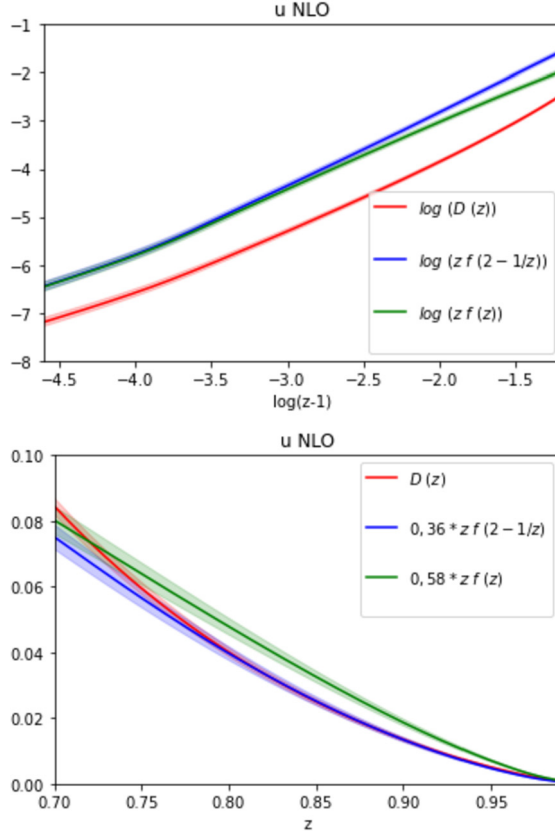


Fig. 8. Top plot: FFs vs pdfs in a log-log plot coming into the end-point  $z \rightarrow 1$  (notice this makes the  $OX$  axis opposite to its usual representation, so that the functions fall towards the left). This suggests that the power-law is the same for both types of functions, and that the reason that the Kullback-Leibler divergence in Table 2 perceives a difference is due to a multiplicative constant. Bottom plot: FFs vs pdfs multiplied by a fixed constant, as suggested by the top plot, now in the usual scale, showing great resemblance.

We have found that the entropy is one adequate number to quantify future progress on knowledge of the fragmentation functions, binding  $S$  from above with a uniform, maximum-entropy distribution for the unknown channels of fragmentation, as shown in Fig. 2.

One issue that we have not discussed in length but that may play a role in future analysis is the fact that fragmenting partons may yield excited hadron resonances that are now very well known to exist [29] but were not contemplated in early quark analysis of the sixties. These excited hadron resonances (such as  $N^*$ ,  $\Delta^* \rightarrow p \dots$ ) undergo one or more secondary decays to Gell-Mann’s quark-model basic hadrons, and add to their fragmentation function [30] (for example, here that of the proton), even when the time-scales involved may be different and it would be more convenient to distinguish them as separate fragmentation functions.

We have also examined the relations among parton distribution and fragmentation functions proposed in the literature for the same parton and hadron, and found them to not be well satisfied with current sets.

We have found at least in one case, that of  $D_{u \rightarrow \pi^+}$ , that the reason is not a power-law mismatch between the parton distribution and the fragmentation function, but rather a multiplicative constant (an upward shift in a logarithmic plot). While data sets tend to have larger errors near the  $z = 1$  endpoint, we believe that with modest effort, a future systematic analysis of this observation can be carried out as those uncertainty bands come somewhat down, to see whether it stands further tests. Measurement of the power-laws of fragmentation functions has long been known to offer a window into hadron dynamics [31] and we hope to see future progress.

## CRediT authorship contribution statement

Guillermo Benito Calviño and Javier García Olivares have conducted all numerical computations in the article and downloaded the data from the prior works that have been exploited (JAM, NNPDF, etc.). Felipe J. Llanes Estrada is responsible for directing the investigation and drafting the manuscript.

## Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Felipe J. Llanes-Estrada reports financial support was provided by EU Horizon Program. Felipe J. Llanes-Estrada reports financial support was provided by Spain Ministry of Science and Innovation.

## Data availability

Data is from the websites and publications of the NNPDF, MAPFF and JAM collaborations and should be obtained from them if needed.

## Acknowledgements

We thank Dr. Pia Zurita for useful comments.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824093; grants MICINN: PID2019-108655GB-I00, PID2019-106080GB-C21 (Spain); UCM research group 910309 and the IPARCOS institute.

## References

- [1] A. Galindo, M.A. Martin-Delgado, *Rev. Mod. Phys.* 74 (2002) 347–423, <https://doi.org/10.1103/RevModPhys.74.347>.
- [2] P. Colangelo, F. Loparco, *Phys. Lett. B* 788 (2019) 500–504, <https://doi.org/10.1016/j.physletb.2018.11.053>.
- [3] S. Gupta, S.K. Gupta, *Nucl. Phys. B* 965 (2021) 115336, <https://doi.org/10.1016/j.nuclphysb.2021.115336>.
- [4] A. Alves, A.G. Dias, R. da Silva, *Nucl. Phys. B* 959 (2020) 115137, <https://doi.org/10.1016/j.nuclphysb.2020.115137>.
- [5] D. d'Enterria, preprint, arXiv:1208.1993 [hep-ph].
- [6] A. Alves, A.G. Dias, R. da Silva, *Physica A* 420 (2015) 1–7, <https://doi.org/10.1016/j.physa.2014.10.084>.
- [7] C.W. Ma, Y.G. Ma, *Prog. Part. Nucl. Phys.* 99 (2018) 120–158, <https://doi.org/10.1016/j.pnpnp.2018.01.002>.
- [8] D.E. Kharzeev, E.M. Levin, *Phys. Rev. D* 95 (2017) 114008, <https://doi.org/10.1103/PhysRevD.95.114008>.
- [9] W. Kou, X. Wang, X. Chen, arXiv:2208.07521 [hep-ph].
- [10] D. Neill, W.J. Waalewijn, *Phys. Rev. Lett.* 123 (2019) 142001, <https://doi.org/10.1103/PhysRevLett.123.142001>.

- [11] P. Carrasco Millán, et al., Nucl. Phys. B 930 (2018) 583–596, <https://doi.org/10.1016/j.nuclphysb.2018.04.003>.
- [12] F.J. Llanes-Estrada, et al., PoS Hadron2017 (2017) 245.
- [13] J.A. Osborne, E. Wang, X.N. Wang, Phys. Rev. D 67 (2003) 094022, <https://doi.org/10.1103/PhysRevD.67.094022>.
- [14] V. Bertone, et al., NNPDF, Eur. Phys. J. C 77 (2017) 516, <https://doi.org/10.1140/epjc/s10052-017-5088-y>.
- [15] R.A. Khalek, V. Bertone, E.R. Nocera, Phys. Rev. D 104 (3) (2021) 034007, <https://doi.org/10.1103/PhysRevD.104.034007>.
- [16] R. Abdul Khalek, V. Bertone, A. Khoudli, E.R. Nocera, Phys. Lett. B 834 (2022) 137456, <https://doi.org/10.1016/j.physletb.2022.137456>.
- [17] D. de Florian, R. Sassot, M. Epele, R.J. Hernández-Pinto, M. Stratmann, Phys. Rev. D 91 (2015) 014035, <https://doi.org/10.1103/PhysRevD.91.014035>.
- [18] I. Borsa, D. de Florian, R. Sassot, M. Stratmann, W. Vogelsang, Phys. Rev. Lett. 129 (1) (2022) 012002, <https://doi.org/10.1103/PhysRevLett.129.012002>.
- [19] S.D. Drell, D.J. Levy, T.M. Yan, Phys. Rev. D 1 (1970) 1617–1639, <https://doi.org/10.1103/PhysRevD.1.1617>.
- [20] H. Paukkunen, P. Zurita, Eur. Phys. J. C 80 (2020) 381, <https://doi.org/10.1140/epjc/s10052-020-7971-1>.
- [21] V. Barone, A. Drago, B.Q. Ma, Phys. Rev. C 62 (2000) 062201, <https://doi.org/10.1103/PhysRevC.62.062201>.
- [22] V.N. Gribov, L.N. Lipatov, Sov. J. Nucl. Phys. 15 (1972) 675–684.
- [23] S. Kullback, R.A. Leibler, Ann. Math. Stat. 22 (1951) 79–86, <https://doi.org/10.1214/aoms/1177729694>.
- [24] M. Le Bellac, in: Proceedings of 19th Rencontres de Moriond: New Particle Production at High-Energy (Hadronic Session), 1984, pp. 487–508, preprint N TH 84/7.
- [25] M. Hirai, S. Kumano, Comput. Phys. Commun. 183 (2012) 1002–1013, <https://doi.org/10.1016/j.cpc.2011.12.022>, arXiv:1106.1553 [hep-ph].
- [26] C.H. Kom, A. Vogt, K. Yeats, J. High Energy Phys. 10 (2012) 033, [https://doi.org/10.1007/JHEP10\(2012\)033](https://doi.org/10.1007/JHEP10(2012)033).
- [27] A. Bramon, R. Escribano, M.D. Scadron, Eur. Phys. J. C 7 (1999) 271–278, <https://doi.org/10.1007/s100529801009>.
- [28] E. Moffat, et al., J.A.M. Jefferson Lab Angular Momentum, Phys. Rev. D 104 (1) (2021) 016015, <https://doi.org/10.1103/PhysRevD.104.016015>.
- [29] R.L. Workman, Particle Data Group, PTEP 2022 (2022) 083C01, <https://doi.org/10.1093/ptep/ptac097>.
- [30] A. Adamov, G.R. Goldstein, Phys. Rev. D 64 (2001) 014021, <https://doi.org/10.1103/PhysRevD.64.014021>.
- [31] S.J. Brodsky, J.F. Gunion, Phys. Rev. D 17 (1978) 848–857, <https://doi.org/10.1103/PhysRevD.17.848>.