

Essential Conditions for the Full Synergy of Probability of Occurrence Distributions

Rubem P. Mondaini *  and Simão C. de Albuquerque Neto

COPPE, Centre of Technology, Federal University of Rio de Janeiro, Rio de Janeiro 21941-901, Brazil;
simao.albuquerque.neto@coppe.ufrj.br

* Correspondence: rubem.mondaini@ufrj.br

Abstract: In this contribution, we specify the conditions for assuring the validity of the synergy of the distribution of probabilities of occurrence. We also study the subsequent restriction on the maximal extension of the strict concavity region on the parameter space of Sharma–Mittal entropy measures, which has been derived in a previous paper in this journal. The present paper is then a necessary complement to that publication. Some applications of the techniques introduced here are applied to protein domain families (Pfam databases, versions 27.0 and 35.0). The results will show evidence of their usefulness for testing the classification work performed with methods of alignment that are used by expert biologists.

Keywords: entropy measures; synergy; probabilistic distributions



Citation: Mondaini, R.P.; de Albuquerque Neto, S.C. Essential Conditions for the Full Synergy of Probability of Occurrence Distributions. *Entropy* **2022**, *24*, 993. <https://doi.org/10.3390/e24070993>

Academic Editors: Nicola Cappetti, Yusif S. Gasimov and Francesco Villecco

Received: 14 June 2022

Accepted: 12 July 2022

Published: 18 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

We have been working since 2015 on the problem of testing the alignment of protein domain families which are proposed by expert biologists and bioinformaticians. We have found that the use of selected entropy measures is very proficient for testing the results published by those professionals and they favour a rigorous ANOVA statistical analysis [1]. In order to reduce the search space for admissible values of entropy measures, we have emphasized the need for work in the region related to strict concavity of these entropies. This study has been undertaken in a previous work, and we present in Section 2 a summary of those developments. In the present work, we aim to complement the results of a previous publication [2], and a subsequent restriction on the parameter space has to be performed in order to guarantee the synergy of the probability distributions to be tested. Non-synergetic distributions are not worthwhile for working because they will not preserve the fundamental property of getting more information of amino acids into t -sets of columns than to sum up the information obtained from individual columns. In Section 3, a brief digression is then made for introducing the Sharma–Mittal class of entropy measures. Section 4 emphasizes the aspects of synergy of the distributions and their consequences for the reduction of the parameter space of Sharma–Mittal entropies. In Section 5, we treat the analysis of the maximal extension of the parameter space, and we repeat the reduction process imposed by the requirement of fully synergetic distributions of Section 4. We conclude the paper in Section 6 by studying the relation of Hölder and generalized Khinchin–Shannon (GKS) inequalities.

2. The Construction of the Probabilistic Space

Let us consider a set of m_f domains (m_f rows) from a chosen family of protein domains. In order to associate a rectangular array with this family, to be taken as its representative in the probabilistic space we are constructing, we specify its number of columns as $n_f = n$. This means that among m_f rows, we disregard all rows such that the number of their amino acids satisfies $n_f < n$ and preserve m'_f rows whose number of amino acids satisfies $n_f \leq n$,

3. The Sharma–Mittal Class of Entropy Measures

As emphasized in Ref. [2], the introduction of random variable functions such as entropy measures associated with the probabilities of occurrence, is suitable to provide an analysis of the evolution of these probabilities through the regions of the parameter space of entropies. The class of Sharma–Mittal entropy measures seems to be particularly adapted to this task when related to the occurrence of amino acids in the objects $p_{j_1 \dots j_t}(a_1, \dots, a_t)$. The thermodynamic interpretation of the notion of entropy greatly helps to classify the distribution of its values associated with protein domain databases and to interpret its evolution through the Fokker–Planck equations to be treated in forthcoming articles in this line of research.

The two-parameter Sharma–Mittal class of entropy measures is usually given by

$$(SM)_{j_1 \dots j_t}^{(r,s)} = \frac{\left(\alpha_{j_1 \dots j_t}^{(s)}\right)^{\frac{1-r}{1-s}} - 1}{1-r}; \quad (SM)_{j_t}^{(r,s)} = \frac{\left(\alpha_{j_t}^{(s)}\right)^{\frac{1-r}{1-s}} - 1}{1-r}, \quad (6)$$

where

$$\alpha_{j_1 \dots j_t}^{(s)} = \sum_{a_1, \dots, a_t} [p_{j_1 \dots j_t}(a_1, \dots, a_t)]^s; \quad \alpha_{j_t}^{(s)} = \sum_{a_t} [p_{j_t}(a_t)]^s. \quad (7)$$

The parameters r, s must bound a region corresponding to a strict concavity in the parameter space. A necessary requirement to be satisfied [3] is

$$\frac{\partial^2 (SM)_{j_1 \dots j_t}^{(r,s)}}{\partial (p_{j_1 \dots j_t}(a_1, \dots, a_t))^2} = s \left(\alpha_{j_1 \dots j_t}^{(s)}\right)^{\frac{s-r}{1-s}} \cdot [p_{j_1 \dots j_t}(a_1, \dots, a_t)]^{s-2} \cdot \left[\frac{s(s-r)}{(1-s)^2} \hat{p}_{j_1 \dots j_t}(a_1, \dots, a_t) - 1\right] < 0, \quad (8)$$

where $\hat{p}_{j_1 \dots j_t}(a_1, \dots, a_t)$ stands for the escort probability associated with the joint probability $p_{j_1 \dots j_t}(a_1, \dots, a_t)$, or,

$$\hat{p}_{j_1 \dots j_t}(a_1, \dots, a_t) = \frac{[p_{j_1 \dots j_t}(a_1, \dots, a_t)]^s}{\alpha_{j_1 \dots j_t}^{(s)}}; \quad \hat{p}_{j_t}(a_t) = \frac{[p_{j_t}(a_t)]^s}{\alpha_{j_t}^{(s)}}. \quad (9)$$

Equation (8) leads to

$$r \geq s > 0. \quad (10)$$

Some special cases of one-parameter entropies are commonplace in the scientific literature [3–9]:

The $r = s$ region is the domain of the Havrda–Charvat [6] entropy measure $H_{j_1 \dots j_t}^{(s)}$,

$$H_{j_1 \dots j_t}^{(s)} = \frac{\alpha_{j_1 \dots j_t}^{(s)} - 1}{1-s}. \quad (11)$$

The $r = 2 - s, 0 \leq s \leq 1$, region will stand for the domain of the Landsberg–Vedral [7] entropy measure, $L_{j_1 \dots j_t}^{(s)}$,

$$L_{j_1 \dots j_t}^{(s)} = \frac{\alpha_{j_1 \dots j_t}^{(s)} - 1}{(1-s)\alpha_{j_1 \dots j_t}^{(s)}} \equiv \frac{H_{j_1 \dots j_t}^{(s)}}{\alpha_{j_1 \dots j_t}^{(s)}}. \quad (12)$$

The Renyi $R_{j_1 \dots j_t}^{(s)}$ [8] and the “non-extensive” Gaussian [9] $G_{j_1 \dots j_t}^{(r)}$ entropy measures are obtained from limit processes:

$$\begin{aligned}
 R_{j_1 \dots j_t}^{(s)} &\equiv \lim_{r \rightarrow 1} \frac{\left(\alpha_{j_1 \dots j_t}^{(s)}\right)^{\frac{1-r}{1-s}} - 1}{1-r} = \lim_{r \rightarrow 1} \frac{\frac{d}{dr} \left[\left(\alpha_{j_1 \dots j_t}^{(s)}\right)^{\frac{1-r}{1-s}} - 1 \right]}{\frac{d}{dr} (1-r)} = \lim_{r \rightarrow 1} \frac{e^{\frac{1-r}{1-s} \cdot \log \alpha_{j_1 \dots j_t}^{(s)}} \cdot \left(\frac{-1}{1-s}\right) \log \alpha_{j_1 \dots j_t}^{(s)}}{-1} \\
 &= \frac{\log \alpha_{j_1 \dots j_t}^{(s)}}{1-s}.
 \end{aligned} \tag{13}$$

$$\begin{aligned}
 G_{j_1 \dots j_t}^{(r)} &\equiv \lim_{s \rightarrow 1} \frac{\left(\alpha_{j_1 \dots j_t}^{(s)}\right)^{\frac{1-r}{1-s}} - 1}{1-r} = \frac{1}{1-r} \left[\exp \left((1-r) \cdot \lim_{s \rightarrow 1} \frac{\frac{d}{ds} \log \alpha_{j_1 \dots j_t}^{(s)}}{\frac{d}{ds} (1-s)} \right) - 1 \right] \\
 &= \frac{1}{1-r} \left[\exp \left(-(1-r) \cdot \lim_{s \rightarrow 1} \frac{\frac{d}{ds} \alpha_{j_1 \dots j_t}^{(s)}}{\alpha_{j_1 \dots j_t}^{(s)}} \right) - 1 \right].
 \end{aligned}$$

After using the definition of $\alpha_{j_1 \dots j_t}^{(s)}$, Equation (7), and $\lim_{s \rightarrow 1} \alpha_{j_1 \dots j_t}^{(s)} = 1$ from Equations (1) and (2), we get:

$$G_{j_1 \dots j_t}^{(r)} = \frac{e^{(1-r)S_{j_1 \dots j_t}} - 1}{1-r}, \tag{14}$$

where $S_{j_1 \dots j_t}$ is the Gibbs–Shannon entropy measure

$$S_{j_1 \dots j_t} = - \sum_{a_1, \dots, a_t} p_{j_1 \dots j_t}(a_1, \dots, a_t) \log p_{j_1 \dots j_t}(a_1, \dots, a_t). \tag{15}$$

The Gibbs–Shannon entropy measure, Equation (15), is also obtained by taking the convenient limits of the special cases of Sharma–Mittal entropies, Equations (11)–(14):

$$\lim_{s \rightarrow 1} H_{j_1 \dots j_t}^{(s)} = \lim_{s \rightarrow 1} L_{j_1 \dots j_t}^{(s)} = \lim_{s \rightarrow 1} R_{j_1 \dots j_t}^{(s)} = \lim_{r \rightarrow 1} G_{j_1 \dots j_t}^{(r)} = S_{j_1 \dots j_t}. \tag{16}$$

We shall analyse in the next section the structure of the two-parameter space of Sharma–Mittal entropy by taking into consideration these special cases.

We are now reminded that for the limit of Gibbs–Shannon entropy, a conditional entropy measure is defined [3] by

$$S_{j_1 \dots j_{t-1} | j_t} = - \sum_{a_1, \dots, a_t} p_{j_t}(a_t) p_{j_1 \dots j_t}(a_1, \dots, a_{t-1} | a_t) \log p_{j_1 \dots j_t}(a_1, \dots, a_{t-1} | a_t). \tag{17}$$

We then have analogously for the conditional Sharma–Mittal entropy measure [3]

$$(SM)_{j_1 \dots j_{t-1} | j_t} = \frac{\left(\sum_{a_1, \dots, a_t} \hat{p}_{j_t}(a_t) [p_{j_1 \dots j_t}(a_1, \dots, a_{t-1} | a_t)]^s \right)^{\frac{1-r}{1-s}} - 1}{1-r}. \tag{18}$$

It is easy to show by trivial calculation that, analogously to Equation (16), we will have

$$\lim_{s \rightarrow 1} \lim_{r \rightarrow s} (SM)_{j_1 \dots j_{t-1} | j_t} = \lim_{s \rightarrow 1} \lim_{r \rightarrow 2-s} (SM)_{j_1 \dots j_{t-1} | j_t} = \lim_{s \rightarrow 1} \lim_{r \rightarrow 1} (SM)_{j_1 \dots j_{t-1} | j_t} = \lim_{r \rightarrow 1} \lim_{s \rightarrow 1} (SM)_{j_1 \dots j_{t-1} | j_t} = S_{j_1 \dots j_{t-1} | j_t}. \tag{19}$$

From Equations (6), (7) and (18) and the application of the Bayes’ law, Equation (4), we can write

$$(SM)_{j_1 \dots j_t} = (SM)_{j_t} + (SM)_{j_1 \dots j_{t-1} | j_t} + (1-r)(SM)_{j_t} (SM)_{j_1 \dots j_{t-1} | j_t}. \tag{20}$$

4. Aspects of Synergy and the Reduction of the Parameter Space for Fully Synergetic Distributions

For the Gibbs–Shannon entropy measure, the inequality written by A. Y. Khinchin [3,10] is

$$S_{j_1 \dots j_{t-1} | j_t} \leq S_{j_1 \dots j_{t-1}}. \tag{21}$$

This inequality would be described by Khinchin as: “On the average, the knowledge a priori of the distribution on the column j_t can only decrease the uncertainty of the distribution on the j_1, \dots, j_{t-1} columns”. We can write an analogous inequality for the Sharma–Mittal class of entropies

$$(SM)_{j_1 \dots j_{t-1} | j_t} \leq (SM)_{j_1 \dots j_{t-1}}. \tag{22}$$

We then get from Equations (20) and (22)

$$(SM)_{j_1 \dots j_t} \leq (SM)_{j_t} + (SM)_{j_1 \dots j_{t-1}} + (1 - r)(SM)_{j_t} (SM)_{j_1 \dots j_{t-1}}. \tag{23}$$

After iteration of this equation, $t \rightarrow t - 1 \rightarrow t - 2 \rightarrow \dots$, we can also write

$$(SM)_{j_1 \dots j_t} \leq \frac{\prod_{l=1}^t [1 + (1 - r)(SM)_{j_l}] - 1}{1 - r}. \tag{24}$$

The inequalities in (21)–(24) are associated with what are called “synergetic conditions”. In this section, we also derive the fully synergetic conditions as GKS inequalities.

After using Equations (7) and (9) in Equation (23), we get

$$\frac{\left(\alpha_{j_1 \dots j_t}^{(s)}\right)^{\frac{1-r}{1-s}}}{1-r} \leq \frac{\left(\alpha_{j_t}^{(s)}\right)^{\frac{1-r}{1-s}} \cdot \left(\alpha_{j_1 \dots j_{t-1}}^{(s)}\right)^{\frac{1-r}{1-s}}}{1-r}, \tag{25}$$

and after iteration and use of Equation (24)

$$\frac{\left(\alpha_{j_1 \dots j_t}^{(s)}\right)^{\frac{1-r}{1-s}}}{1-r} \leq \frac{\prod_{l=1}^t \left(\alpha_{j_l}^{(s)}\right)^{\frac{1-r}{1-s}}}{1-r}. \tag{26}$$

The hatched region of strict concavity in the parameter space of Sharma–Mittal entropies, $C = \{(s, r) \mid r \geq s > 0\}$, is depicted in Figure 1. The special cases corresponding to Havrda–Charvat’s ($r = s$), Landsberg–Vedral’s ($r = 2 - s$), Renyi’s ($r = 1$), and “non-extensive” Gaussian’s ($s = 1$) entropies are also represented.

We can identify three subregions in Figure 1. They will correspond to

$$R_I = \{(s, r) \mid 1 > r \geq s > 0\} \Rightarrow \alpha_{j_1 \dots j_t}^{(s < 1)} \leq \prod_{l=1}^t \alpha_{j_l}^{(s < 1)}, \tag{27}$$

$$R_{II} = \{(s, r) \mid r \geq s > 1\} \Rightarrow \alpha_{j_1 \dots j_t}^{(s > 1)} \geq \prod_{l=1}^t \alpha_{j_l}^{(s > 1)}, \tag{28}$$

$$R_{III} = \{(s, r) \mid r > 1 > s > 0\} \Rightarrow \alpha_{j_1 \dots j_t}^{(s < 1)} \leq \prod_{l=1}^t \alpha_{j_l}^{(s < 1)} \tag{29}$$

where the ordering of α -symbols has been obtained from Equation (26). The subregions R_I and R_{III} are what we call fully synergetic subregions, and the corresponding inequalities are the GKS inequalities [2].

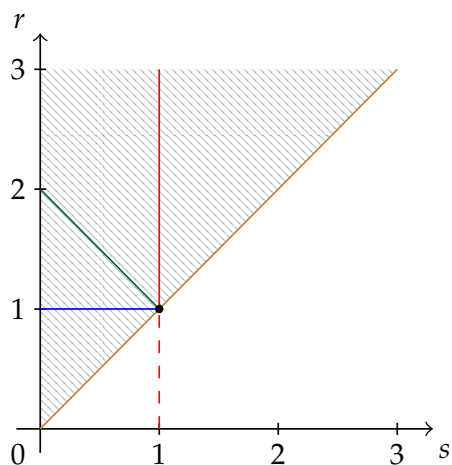


Figure 1. The strict concavity region $C = \{(s, r) \mid r \geq s > 0\}$ of Sharma–Mittal class of entropy measures. It is the epigraph of the curve $(r = s)$, and this corresponds to the Havrda–Charvat entropy, which is depicted in brown. The Landsberg–Vedral’s $(r = 2 - s)$, Renyi’s $(r = 1)$, and “non-extensive” Gaussian’s $(s = 1)$ are depicted in green, blue, and red, respectively.

The subregions R_I , R_{II} , and R_{III} are depicted in Figure 2a–c, respectively. The union of subregions R_I and R_{III} is the fully synergetic Khinchin–Shannon restriction to be imposed on the strict concavity region of Figure 1 and it is depicted in Figure 2d below.

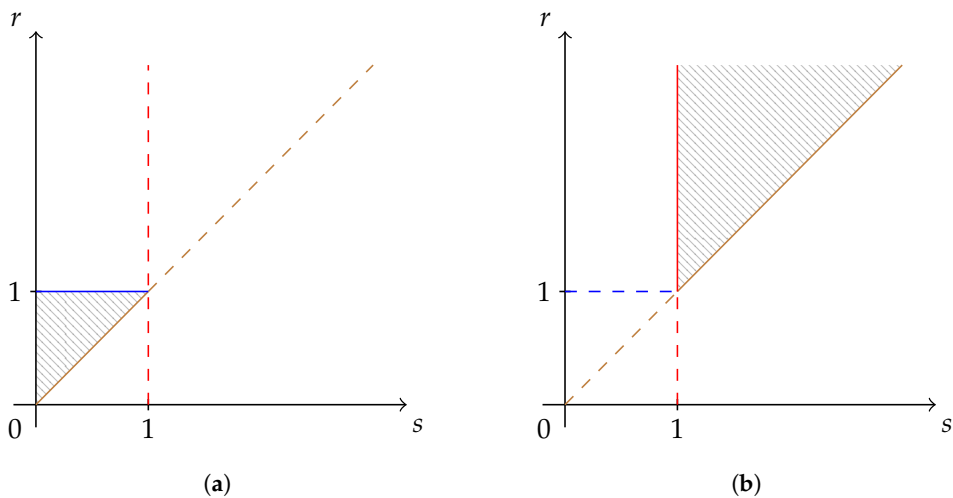


Figure 2. Cont.

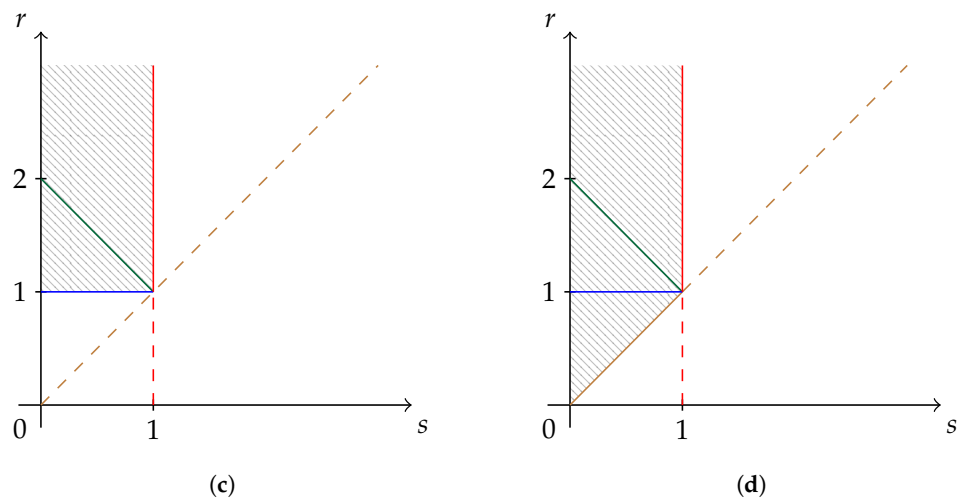


Figure 2. Subregions of the strict concavity region $C = \{(s, r) \mid r \geq s > 0\}$ of the Sharma–Mittal class of entropy measures. (a) Khinchin–Shannon subregion R_I —fully synergetic.; (b) The non-synergetic subregion R_{II} ; (c) Khinchin–Shannon subregion R_{III} —fully synergetic; (d) Khinchin–Shannon subregion $R_I \cup R_{III}$ —fully synergetic. The reduced subregion $R_I \cup R_{III}$ of Figure 1 is obtained by taking into consideration fully synergetic distributions only.

5. The Maximal Extension of the Parameter Space and Its Reduction for Fully Synergetic Distribution

In Figures 1 and 2d, we have depicted the structure of the strict concavity region for Sharma–Mittal entropy measures and its reduction to a subregion by the application of the requirement of fully synergetic distributions, respectively. Our analysis has used a coarse-grained approach to concavity given by Equations (8) and (10). We now introduce some necessary refinements for characterizing the probability of occurrence in subarrays of m rows and t columns, $m \times t$. For t columns, there are $(20)^t$ possibilities of occurrence of amino acids, which could be a large number, but we could count not individual amino acids, but groups of t -sets of amino acids (μ -groups) which appear on the m rows of the $m \times t$ array. We characterize these μ -groups by $\mu = 1, \dots, m$, from all equal μ -groups ($\mu = 1$) to m different μ -groups ($\mu = m$). We also call q_μ , the number of equal t -sets of a given μ -group.

In Equation (2), the sum is over all the amino acids that make up the geometric object defined in Equation (1), the probability of occurrence. We can now perform the sum over μ -groups and write

$$\sum_{a_1^{q_\mu}, \dots, a_t^{q_\mu}} p_{j_1 \dots j_t}(a_1^{q_\mu}, \dots, a_t^{q_\mu}) = \sum_{a_1^{q_\mu}, \dots, a_t^{q_\mu}} \frac{n_{j_1 \dots j_t}(a_1^{q_\mu}, \dots, a_t^{q_\mu})}{m} = \sum_{\mu=1}^m \frac{q_\mu}{m} = 1, \tag{30}$$

where $(a_1^{q_\mu}, \dots, a_t^{q_\mu})$ are the t -sets of a μ -group. We also have from Equation (7)

$$\sum_{a_1^{q_\mu}, \dots, a_t^{q_\mu}} [p_{j_1 \dots j_t}(a_1^{q_\mu}, \dots, a_t^{q_\mu})]^s = \sum_{a_1^{q_\mu}, \dots, a_t^{q_\mu}} \left[\frac{n_{j_1 \dots j_t}(a_1^{q_\mu}, \dots, a_t^{q_\mu})}{m} \right]^s = \sum_{\mu=1}^m \left(\frac{q_\mu}{m} \right)^s = \alpha_{j_1 \dots j_t}^{(s)}. \tag{31}$$

From Equations (30) and (31), we can now proceed to the calculation of the Hessian matrix for Sharma–Mittal entropy measures. We have for the first derivative of $(SM)_{j_1 \dots j_t}$

$$\frac{\partial (SM)_{j_1 \dots j_t}}{\partial p_{j_1 \dots j_t}(a_1^{q_\mu}, \dots, a_t^{q_\mu})} = \frac{s}{1-s} \left(\alpha_{j_1 \dots j_t}^{(s)} \right)^{\frac{s-r}{1-s}} [p_{j_1 \dots j_t}(a_1^{q_\mu}, \dots, a_t^{q_\mu})]^{s-1}. \tag{32}$$

We then have for a generic element of the Hessian matrix [2]

$$\begin{aligned}
 H_{q_\mu q_\nu} &= \frac{\partial^2(SM)_{j_1 \dots j_t}}{\partial p_{j_1 \dots j_t}(a_1^{q_\mu}, \dots, a_t^{q_\mu}) \partial p_{j_1 \dots j_t}(a_1^{q_\nu}, \dots, a_t^{q_\nu})} \\
 &= s \left(\alpha_{j_1 \dots j_t}^{(s)} \right)^{\frac{s-r}{1-s}} \left[p_{j_1 \dots j_t}(a_1^{q_\mu}, \dots, a_t^{q_\mu}) \right]^{s-2} \left[\frac{s(s-r)}{(1-s)^2} \frac{p_{j_1 \dots j_t}(a_1^{q_\mu}, \dots, a_t^{q_\mu})}{p_{j_1 \dots j_t}(a_1^{q_\nu}, \dots, a_t^{q_\nu})} \cdot \hat{p}_{j_1 \dots j_t}(a_1^{q_\nu}, \dots, a_t^{q_\nu}) - \delta_{\mu\nu} \right], \tag{33}
 \end{aligned}$$

where $\hat{p}_{j_1 \dots j_t}(a_1^{q_\mu}, \dots, a_t^{q_\mu})$ is the escort probability associated to $p_{j_1 \dots j_t}(a_1^{q_\mu}, \dots, a_t^{q_\mu})$, or

$$\hat{p}_{j_1 \dots j_t}(a_1^{q_\mu}, \dots, a_t^{q_\mu}) \equiv \frac{\left[p_{j_1 \dots j_t}(a_1^{q_\mu}, \dots, a_t^{q_\mu}) \right]^s}{\sum_{b_1^{q_\nu}, \dots, b_t^{q_\nu}} p_{j_1 \dots j_t}(b_1^{q_\nu}, \dots, b_t^{q_\nu})}. \tag{34}$$

The principal minors are given by

$$\begin{aligned}
 \det H_{q_\mu q_\nu}(\mu, \nu = 1, \dots, k) &= (-1)^{k-1} s^k \left(\alpha_{j_1 \dots j_t}^{(s)} \right)^{\frac{k(s-r)}{1-s}} \left[\prod_{\mu=1}^k p_{j_1 \dots j_t}(a_1^{q_\mu}, \dots, a_t^{q_\mu}) \right]^{s-2} \\
 &\quad \cdot \left[\frac{s(s-r)}{(1-s)^2} \sum_{\mu=1}^k \hat{p}_{j_1 \dots j_t}(a_1^{q_\mu}, \dots, a_t^{q_\mu}) - 1 \right], \quad k = 1, \dots, m, \tag{35}
 \end{aligned}$$

and we have

$$\sum_{\mu=1}^k \hat{p}_{j_1 \dots j_t}(a_1^{q_\mu}, \dots, a_t^{q_\mu}) = \sum_{\mu=1}^k \frac{\left[p_{j_1 \dots j_t}(a_1^{q_\mu}, \dots, a_t^{q_\mu}) \right]^s}{\sum_{\mu=1}^m \left(\frac{q_\mu}{m} \right)^s} = \frac{\sum_{\mu=1}^k \left(\frac{q_\mu}{m} \right)^s}{\sum_{\mu=1}^m \left(\frac{q_\mu}{m} \right)^s} \equiv \sigma_k(s) \tag{36}$$

according to Equation (31).

From Equations (35) and (36), the requirement of strict concavity will lead to

$$\frac{s(s-r)}{(1-s)^2} \sigma_k(s) - 1 < 0. \tag{37}$$

We then have

$$\det H_{q_\mu q_\nu}(\mu, \nu = 1, \dots, k) < 0, \quad k \text{ odd}; > 0, \quad k \text{ even}. \tag{38}$$

This does correspond to the criterion of negative definiteness of the Hessian matrix for strict concavity of multivariate functions [11].

Each k -value is associated with the k -epigraph region, which is the k -extension of the strict concavity region presented in Figure 1. These regions are given by

$$C_{k_{\max}} = \{(s, r) \mid r \geq s > 0\} \cup \left\{ (s, r) \mid s > r > s - \frac{(1-s)^2}{s \sigma_k(s)} \right\}, \quad k = 1, \dots, m. \tag{39}$$

The greatest lower bound of the sequence of k -curves is given by $\sigma_m(s) = 1$. We then have

$$r_m(s) = 2 - \frac{1}{s}. \tag{40}$$

We can then write for the maximal extended region of strict concavity

$$C_{\max} = \{(s, r) \mid r \geq s > 0\} \cup \left\{ (s, r) \mid s > r > 2 - \frac{1}{s} \right\}. \tag{41}$$

The region corresponding to Equation (41) is depicted in Figure 3 below.

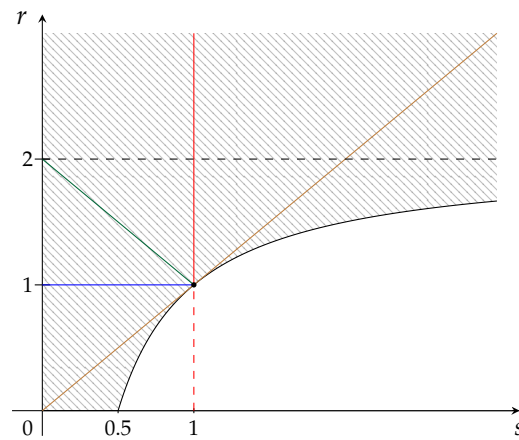


Figure 3. The maximal strict concavity region of the Sharma–Mittal class of entropy measures. The hatched region is the epigraph of the curve $r = 2 - 1/s$ which is depicted in black. The Havrda–Charvat ($r = s$) region is in brown. The Landsberg–Vedral ($r = 2 - s$), Renyi ($r = 1$), and “non-extensive” Gaussian ($s = 1$) regions are depicted in green, blue, and red, respectively.

We are now ready to undertake the application of restrictions for fully synergetic distributions (validity of GKS inequalities) to the maximal strict concavity region of Figure 3.

We start by identifying two regions included in Figure 3. They will be given by

$$R_{IV} = \left\{ (s, r) \mid 1 > s > r \geq 2 - \frac{1}{s} > 0 \right\} \Rightarrow \alpha_{j_1 \dots j_t}^{(s < 1)} \leq \prod_{l=1}^t \alpha_{j_l}^{(s < 1)}, \quad (42)$$

$$R_V = \left\{ (s, r) \mid s > r \geq 2 - \frac{1}{s} > 1 \right\} \Rightarrow \alpha_{j_1 \dots j_t}^{(s > 1)} \geq \prod_{l=1}^t \alpha_{j_l}^{(s > 1)}. \quad (43)$$

These regions are depicted in Figure 4a,b, respectively.

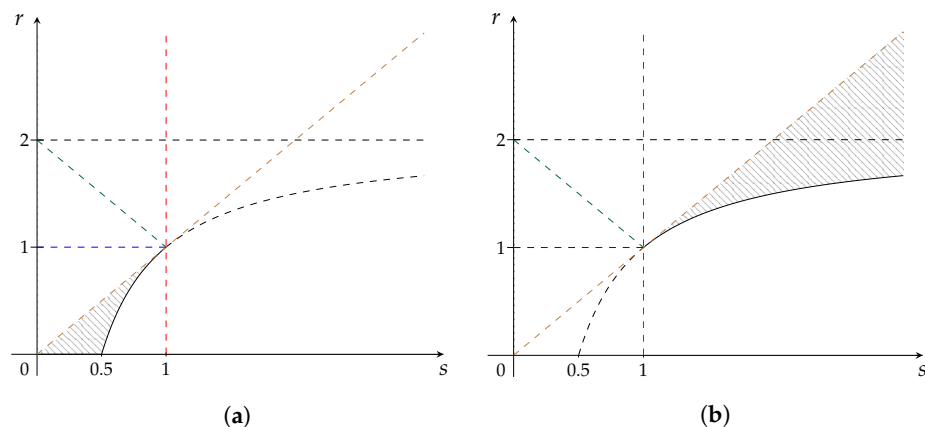


Figure 4. Subregions of the maximal strict concavity region of the Sharma–Mittal class of entropy measures (Figure 3). (a) Khinchin–Shannon subregion R_{IV} —fully synergetic; (b) The non-synergetic subregion R_V .

In order to find the reduced region corresponding to Figure 3, analogously to what has been done for Figure 1, we also need the subregions R_I , R_{III} , Equations (27) and (29): the resulting subregion of fully synergetic distributions is given by $R_{IV} \cup R_I \cup R_{III}$ and is depicted in Figure 5.

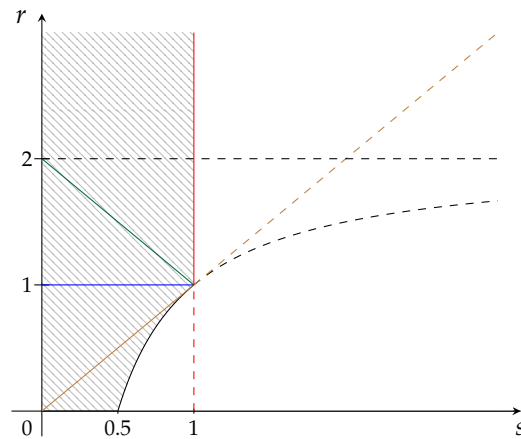


Figure 5. $R_{IV} \cup R_I \cup R_{III}$ is the reduction of the region of Figure 3 by taking into consideration the fully synergetic distributions only.

6. Hölder Inequalities and GKS Inequalities: A Possible Conjecture

In this section, we study the relation between GKS inequalities [2] and Hölder inequalities by using examples of distributions obtained from databases of protein domain families. In order to start, some definitions and properties of the probabilistic space are now in order.

Let us first introduce the definition of the conditional probability of occurrence of the escort probability of occurrence [12]. This is a simple application to escort probabilities of Equation (3):

$$\hat{p}_{j_1 \dots j_t}(a_1, \dots, a_{t-1} | a_t) = \frac{\hat{p}_{j_1 \dots j_t}(a_1, \dots, a_t)}{\hat{p}_{j_t}(a_t)}. \tag{44}$$

From the definitions of escort probabilities, Equation (9), we can write

$$\hat{p}_{j_1 \dots j_t}(a_1, \dots, a_t) = \frac{(p_{j_1 \dots j_t}(a_1, \dots, a_t))^s}{\sum_{b_1, \dots, b_t} (p_{j_1 \dots j_t}(b_1, \dots, b_t))^s}, \tag{45}$$

and

$$\hat{p}_{j_t}(a_t) = \frac{(p_{j_t}(a_t))^s}{\sum_{b_t} (p_{j_t}(b_t))^s}. \tag{46}$$

In Equations (44)–(46), the symbols $a_1, \dots, a_t; b_1, \dots, b_t$ assume the representative letters of the one-letter code for the 20 amino acids, $a_j; b_j (1 \leq j \leq t) \in \{A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y\}$.

After substituting Equations (45) and (46) into Equation (44), we get

$$\hat{p}_{j_1 \dots j_t}(a_1, \dots, a_{t-1} | a_t) = \frac{(p_{j_1 \dots j_t}(a_1, \dots, a_{t-1} | a_t))^s (p_{j_t}(a_t))^s}{\sum_{b_1, \dots, b_t} (p_{j_t}(b_t))^s (p_{j_1 \dots j_t}(b_1, \dots, b_{t-1} | b_t))^s \cdot \hat{p}_{j_t}(a_t)}, \tag{47}$$

and from Equation (46)

$$\hat{p}_{j_1 \dots j_t}(a_1, \dots, a_{t-1} | a_t) = \frac{(p_{j_1 \dots j_t}(a_1, \dots, a_{t-1} | a_t))^s}{\sum_{b_1, \dots, b_t} \hat{p}_{j_t}(b_t) (p_{j_1 \dots j_t}(b_1, \dots, b_{t-1} | b_t))^s}. \tag{48}$$

We also write the definition of escort probability of occurrence of the conditional probability of occurrence [12]

$$\overbrace{p_{j_1 \dots j_t}(a_1, \dots, a_{t-1} | a_t)} = \frac{(p_{j_1 \dots j_t}(a_1, \dots, a_{t-1} | a_t))^s}{\sum_{b_1, \dots, b_{t-1}} (p_{j_1 \dots j_t}(b_1, \dots, b_{t-1} | b_t))^s}. \tag{49}$$

We can check the definitions of Equations (48) and (49) from the equality of the two escort probabilities with the original conditional probability, for $s = 1$

$$s = 1 \Rightarrow \hat{p}_{j_1 \dots j_t}(a_1, \dots, a_{t-1} | a_t) = \overbrace{p_{j_1 \dots j_t}(a_1, \dots, a_{t-1} | a_t)} = p_{j_1 \dots j_t}(a_1, \dots, a_{t-1} | a_t). \tag{50}$$

We should note that the denominators of the right-hand sides of Equations (48) and (49), or,

$$Z \equiv \sum_{a_1, \dots, a_t} \hat{p}_{j_t}(a_t) [p_{j_1 \dots j_t}(a_1, \dots, a_{t-1} | a_t)]^s = \frac{\alpha_{j_1 \dots j_t}^{(s)}}{\alpha_{j_t}^{(s)}}, \tag{51}$$

and

$$X(a_t) \equiv \sum_{a_1, \dots, a_{t-1}} (p_{j_1 \dots j_t}(a_1, \dots, a_{t-1} | a_t))^s \tag{52}$$

will be equal if all amino acids in the j_t column are equal. If we have, for instance, the j_t column given by:

$$j_t \leftrightarrow (\underbrace{A, A, A, A, \dots, A}_m). \tag{53}$$

The unit vectors of probabilities \hat{p}_{j_t} and p_{j_t} will also be equal and given by

$$(\hat{p}_{j_t})^\top = (p_{j_t})^\top = (\underbrace{1, 0, 0, 0, \dots, 0}_{20}). \tag{54}$$

This means that for this special case of an event of rare occurrence, we also have the equality of the conditional of the escort probability and the escort probability of the conditional probability, or the left-hand sides of Equations (48) and (49), respectively.

For a j_t -column with a generic distribution of amino acids, the denominators Z and $X(a_t)$ on the right-hand sides of Equations (48) and (49) will no longer be equal. An ordering of these denominators should be decided from the probabilities of amino acid occurrence in a chosen protein domain family.

This study is undertaken with the help of the functions Z and $X(a_t)$ of Equations (51) and (52) and with the functions J and U , defined below:

$$J \equiv \sum_{a_1, \dots, a_{t-1}} \left[\sum_{a_t} \hat{p}_{j_t}(a_t) p_{j_1 \dots j_t}(a_1, \dots, a_{t-1} | a_t) \right]^s, \tag{55}$$

$$U \equiv \sum_{a_1, \dots, a_{t-1}} (p_{j_1 \dots j_{t-1}}(a_1, \dots, a_{t-1}))^s \equiv \alpha_{j_1 \dots j_{t-1}}^{(s)}. \tag{56}$$

Our method will then be the comparison of pairs of functions in order to proceed with the search for the effect of fully synergetic distributions of amino acids.

There are six comparisons to study:

(I) $X(a_t) \geq Z$

or

$$\frac{1}{(p_{j_t}(a_t))^s} \sum_{a_1, \dots, a_{t-1}} (p_{j_1 \dots j_t}(a_1, \dots, a_t))^s \geq \frac{\alpha_{j_1 \dots j_t}^{(s)}}{\alpha_{j_t}^{(s)}}; p_{j_t}(a_t) \neq 0. \tag{57}$$

(II) $X(a_t) \geq J$
or

$$\frac{1}{(p_{j_t}(a_t))^s} \sum_{a_1, \dots, a_{t-1}} (p_{j_1 \dots j_t}(a_1, \dots, a_t))^s \geq \frac{1}{\alpha_{j_t}^{(s)}} \sum_{a_1, \dots, a_{t-1}} \left[\sum_{a_t} (p_{j_t}(a_t))^{s-1} p_{j_1 \dots j_t}(a_1, \dots, a_t) \right]^s; \tag{58}$$

$p_{j_t}(a_t) \neq 0.$

(III) $X(a_t) \geq U$
or

$$\frac{1}{(p_{j_t}(a_t))^s} \sum_{a_1, \dots, a_{t-1}} (p_{j_1 \dots j_t}(a_1, \dots, a_t))^s \geq \alpha_{j_1 \dots j_{t-1}}^{(s)}. \tag{59}$$

(IV) $U \geq J$
or

$$\sum_{a_1, \dots, a_{t-1}} (p_{j_1 \dots j_{t-1}}(a_1, \dots, a_{t-1}))^s \equiv \alpha_{j_1 \dots j_{t-1}}^{(s)} \geq \frac{\mathcal{H}}{(\alpha_{j_t}^{(s)})^s}, \tag{60}$$

where \mathcal{H} is defined by,

$$\mathcal{H} \equiv \sum_{a_1, \dots, a_{t-1}} \left[\sum_{a_t} (p_{j_t}(a_t))^{s-1} p_{j_1 \dots j_t}(a_1, \dots, a_t) \right]^s. \tag{61}$$

(V) $J \geq Z$
or

$$\sum_{a_1, \dots, a_{t-1}} \left[\sum_{a_t} \hat{p}_{j_t}(a_t) p_{j_1 \dots j_t}(a_1, \dots, a_{t-1} | a_t) \right]^s \equiv \frac{\mathcal{H}}{(\alpha_{j_t}^{(s)})^s} \geq \frac{\alpha_{j_1 \dots j_t}^{(s)}}{\alpha_{j_t}^{(s)}}. \tag{62}$$

(VI) $U \geq Z$
or

$$\sum_{a_1, \dots, a_{t-1}} (p_{j_1 \dots j_{t-1}}(a_1, \dots, a_{t-1}))^s \equiv \alpha_{j_1 \dots j_{t-1}}^{(s)} \geq \frac{\alpha_{j_1 \dots j_t}^{(s)}}{\alpha_{j_t}^{(s)}}. \tag{63}$$

Equations (57)–(59) should be multiplied by $(p_{j_t}(a_t))^s$ and after that, each one has to be summed over a_t . We then have, respectively,

$$\alpha_{j_1 \dots j_t}^{(s)} \geq \alpha_{j_1 \dots j_t}^{(s)}, \tag{64}$$

$$(\alpha_{j_t}^{(s)})^{s-1} \alpha_{j_1 \dots j_t}^{(s)} \geq \mathcal{H}, \tag{65}$$

$$\alpha_{j_1 \dots j_t}^{(s)} \geq \alpha_{j_t}^{(s)} \cdot \alpha_{j_1 \dots j_{t-1}}^{(s)}. \tag{66}$$

Equations (60), (62) and (63) can be written, respectively, as

$$(\alpha_{j_t}^{(s)})^s \cdot \alpha_{j_1 \dots j_{t-1}}^{(s)} \geq \mathcal{H} \tag{67}$$

$$\mathcal{H} \geq (\alpha_{j_t}^{(s)})^{s-1} \cdot \alpha_{j_1 \dots j_t}^{(s)}, \tag{68}$$

$$\alpha_{j_t}^{(s)} \cdot \alpha_{j_1 \dots j_{t-1}}^{(s)} \geq \alpha_{j_1 \dots j_t}^{(s)}, \tag{69}$$

The Hölder’s inequality as applied to probabilities of occurrence [3] is written as

$$\frac{1}{\left(\alpha_{j_t}^{(s)}\right)^s} \left[\sum_{a_t} \left(p_{j_t}(a_t)\right)^{s-1} p_{j_1 \dots j_t}(a_1, \dots, a_t) \right]^s \geq \frac{\sum_{a_t} \left[p_{j_1 \dots j_t}(a_1, \dots, a_t)\right]^s}{\alpha_{j_t}^{(s)}}. \tag{70}$$

After multiplying by $\left(\alpha_{j_t}^{(s)}\right)^s$ and summing over a_1, \dots, a_{t-1} , we get

$$\mathcal{H} \geq \left(\alpha_{j_t}^{(s)}\right)^{s-1} \alpha_{j_1 \dots j_t}^{(s)}, \quad s \leq 1. \tag{71}$$

We also define

$$O \equiv \left(\alpha_{j_t}^{(s)}\right)^{s-1} \alpha_{j_1 \dots j_t}^{(s)}, \tag{72}$$

$$B \equiv \left(\alpha_{j_t}^{(s)}\right)^s \alpha_{j_1 \dots j_{t-1}}^{(s)}. \tag{73}$$

We then summarize the results obtained:

1. Equation (64) is only an identity: $\alpha_{j_1 \dots j_t}^{(s)} = \alpha_{j_1 \dots j_t}^{(s)}$.
2. Equations (65) and (68) can be ordered by Hölder’s inequality, Equations (70) and (71).
3. Equations (66) and (69) can be ordered by GKS inequalities, corresponding to fully synergetic distributions of amino acids, $\alpha_{j_1 \dots j_t}^{(s < 1)} \leq \alpha_{j_t}^{(s < 1)} \cdot \alpha_{j_1 \dots j_{t-1}}^{(s < 1)}$.
4. Equation (67) cannot be ordered without additional experimental/phenomenological information on the probabilities of occurrence to be obtained from updated versions of protein domain family databases [13].

We now collect the formulae obtained from the analysis performed on this section. Equations (65) and (68) are ordered by Hölder’s inequality. We write

$$\mathcal{H} - O \geq 0, \quad s < 1. \tag{74}$$

Equations (66) and (69) are ordered by GKS inequality. We write

$$B - O \geq 0, \quad s < 1. \tag{75}$$

After using Equation (73), we can write Equation (67) as

$$B - \mathcal{H} \geq 0. \tag{76}$$

In Figure 6a,b we have depicted the curves corresponding to functions $\mathcal{H} - O$ and $B - O$ for seven 3-sets of contiguous columns and 80 rows, chosen from databases Pfam 27.0 and Pfam 35.0, respectively. There are also inset figures in order to show the curves for $s \geq 1$.

In Figure 7a,b, we do the same for the differences $B - \mathcal{H}$. We emphasize that for the 3-sets such that $B - \mathcal{H} \geq 0, 0 \leq s \leq 1$, the GKS inequalities $B - O \geq 0$ will result from the validity of Hölder’s inequality. We have worked with the PF01926 protein domain family to perform all the calculations.

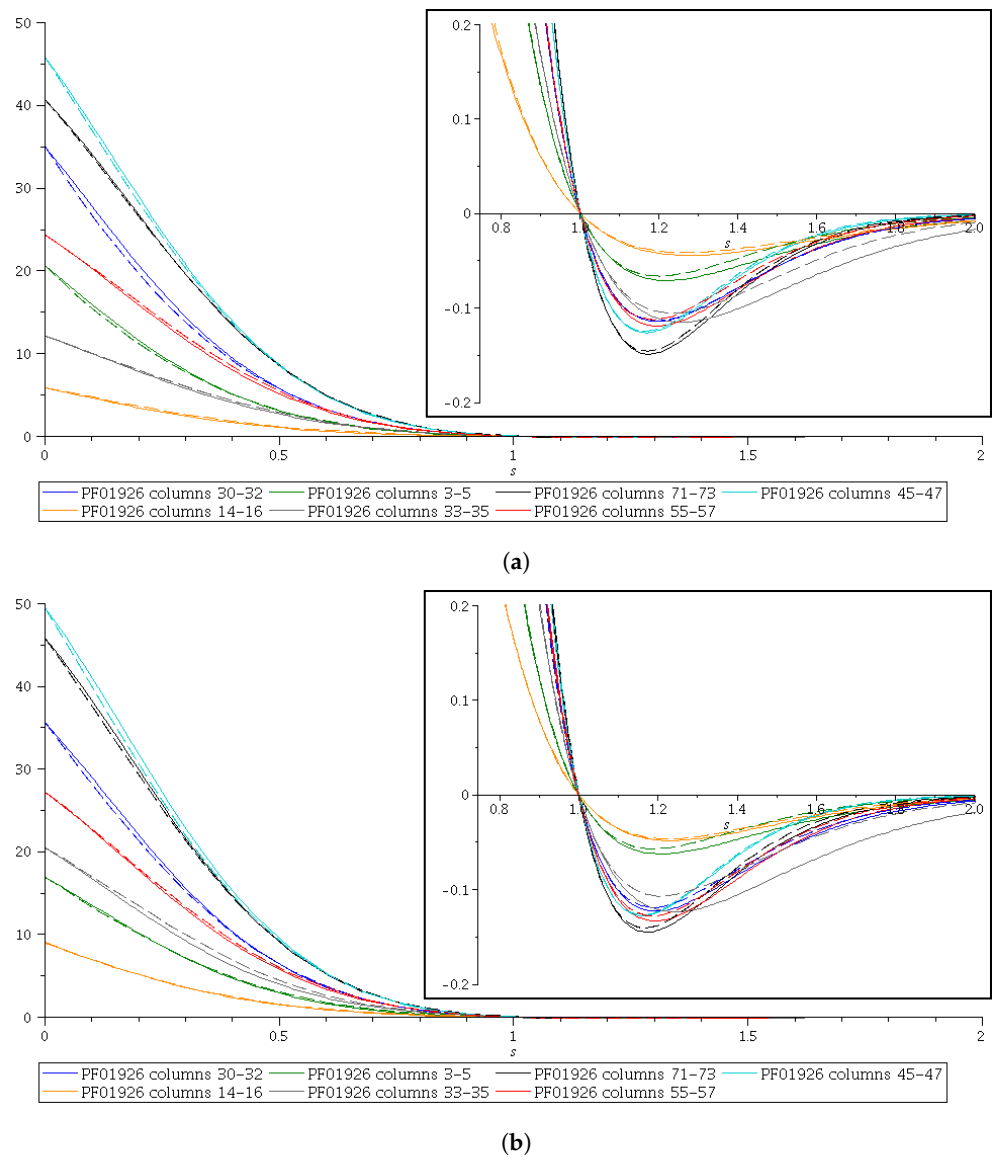


Figure 6. Hölder ($\mathcal{H} - O \geq 0, 0 \leq s \leq 1$) distributions (dashed curves) and Khinchin–Shannon ($B - O \geq 0, 0 \leq s \leq 1$) distributions (continuous curves) of the PF01926 protein domain family from (a) protein domain family PF01926 obtained from Pfam 27.0 and (b) protein domain family PF01926 obtained from Pfam 35.0. The top-right inset shows details of the curves for $s \geq 1$.

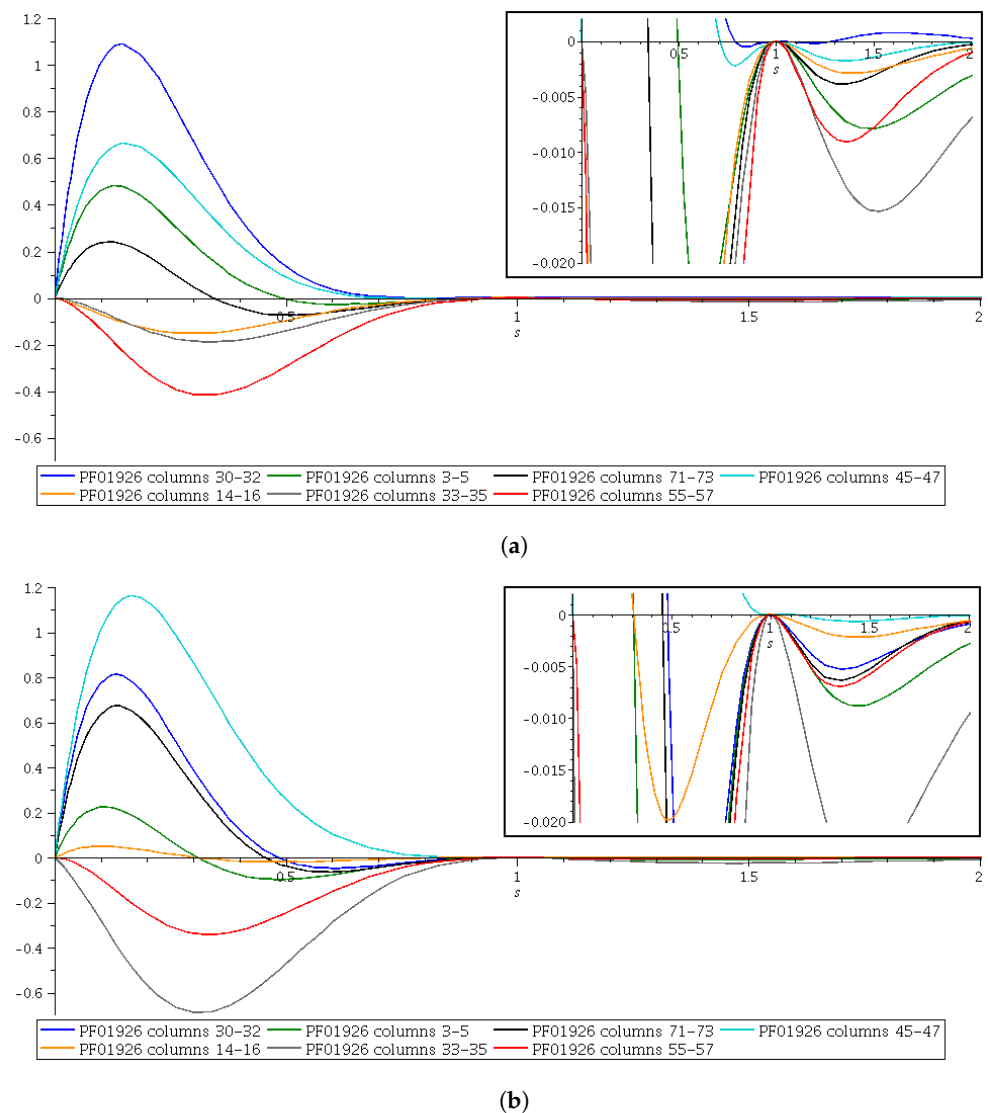


Figure 7. $B - \mathcal{H}$ difference of the PF01926 protein domain family from (a) protein domain family PF01926 obtained from Pfam 27.0 and (b) protein domain family PF01926 obtained from Pfam 35.0. The top-right inset shows details of the curves for $s \geq 1$. $B - \mathcal{H} \geq 0, \mathcal{H} - O \geq 0 \implies B - O \geq 0$.

7. Concluding Remarks

The first comment we want to make to the present work is about the possibility of working in a region of the parameter space that preserves the strict concavity and the fully synergetic structure of the Sharma–Mittal class of entropy measure distributions to be visited by solutions of a new successful statistical mechanics approach. The usual work with Havrda–Charvat distributions describes the evolution along the boundary ($r = s$) of the region ($r \geq s > 0$) that was correctly considered to correspond to strict concavity, but it is also known to be non-synergetic for $s > 1$. We now have the opportunity to develop this statistical mechanics approach along an extended boundary, preserving the strict concavity and providing the study of the evolution of fully synergetic entropy distributions. A first sketch of these developments will be presented in a forthcoming publication.

With respect to Figures 6 and 7, we could hypothesize that if the ordering of B and \mathcal{H} could not be obtained, this would be due to the poor alignment of some protein domain families we have been using, but we are not confident enough that we could do this, because we would need much more information “in silico” to be obtained from many other protein domain families. In other words, we expect that a good alignment of a protein domain family will result in the ordering of B and \mathcal{H} , but we need to verify this in a large

number of families from different Pfam versions before we proceed with a proposal of a method to improve the Pfam database. This looks promising for good scientific work in the line of research we have been aiming to introduce in Ref. [2] and in this contribution.

Author Contributions: Conceptualization, R.P.M. and S.C.d.A.N.; methodology, R.P.M. and S.C.d.A.N.; formal analysis, R.P.M. and S.C.d.A.N.; writing—original draft preparation, R.P.M.; writing—review and editing, R.P.M. and S.C.d.A.N.; visualization, R.P.M. and S.C.d.A.N.; supervision, R.P.M.; project administration, R.P.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviation is used in this manuscript:

GKS Generalized Khinchin-Shannon

References

1. Mondaini, R.P.; de Albuquerque Neto, S.C. The Statistical Analysis of Protein Domain Family Distributions via Jaccard Entropy Measures. In *Trends in Biomathematics: Modeling Cells, Flows, Epidemics, and the Environment*; Mondaini, R.P., Ed.; Springer: Cham, Switzerland, 2020; pp. 169–207.
2. Mondaini, R.P.; de Albuquerque Neto, S.C. Alternative Entropy Measures and Generalized Khinchin-Shannon Inequalities. *Entropy* **2021**, *23*, 1618. [[CrossRef](#)] [[PubMed](#)]
3. Mondaini, R.P.; de Albuquerque Neto, S.C. Khinchin-Shannon Generalized Inequalities for “Non-additive” Entropy Measures. In *Trends in Biomathematics: Mathematical Modeling for Health, Harvesting, and Population Dynamics*; Mondaini, R.P., Ed.; Springer: Cham, Switzerland, 2019; pp. 177–190.
4. Beck, C. Generalized Information and Entropy Measures in Physics. *Contemp. Phys.* **2009**, *50*, 495–510. [[CrossRef](#)]
5. Sharma, B.D.; Mittal, D.P. New Non-additive Measures of Entropy for Discrete Probability Distributions. *J. Math. Sci.* **1975**, *10*, 28–40.
6. Havrda, J.; Charvat, F. Quantification Method of Classification Processes. Concept of Structural α -entropy. *Kybernetika* **1967**, *3*, 30–35.
7. Landsberg, P.T.; Vedral, V. Distributions and Channel Capacities in Generalized Statistical Mechanics. *Phys. Lett. A* **1998**, *247*, 211–217. [[CrossRef](#)]
8. Rényi, A. On Measures of Entropy and Information. In *Contributions to the Theory of Statistics, Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, CA, USA, 20 June–30 July 1960*; Neyman, J., Ed.; University of California Press: Berkeley, CA, USA, 1961; Volume 1, pp. 547–561.
9. Oikonomou, T. Properties of the “Non-extensive Gaussian” Entropy. *Phys. A* **2007**, *381*, 155–163. [[CrossRef](#)]
10. Khinchin, A.Y. *Mathematical Foundations of Information Theory*; Dover Publications, Inc.: New York, NY, USA, 1957.
11. Marsden, J.E.; Tromba, A. *Vector Calculus*, 6th ed.; W. H. Freeman and Company Publishers: New York, NY, USA, 2012.
12. Mondaini, R.P.; de Albuquerque Neto, S.C. The Maximal Extension of the Strict Concavity Region on the Parameter Space for Sharma-Mittal Entropy Measures. In *Trends in Biomathematics: Stability and Oscillations in Environmental Social and Biological Models*; Mondaini, R.P., Ed.; Springer: Cham, Switzerland, 2022.
13. Mistry, J.; Chuguransky, S.; Williams, L.; Qureshi, M.; Salazar, G.A.; Sonnhammer, E.L.L.; Tosatto, S.C.E.; Paladin, L.; Raj, S.; Richardson, L.J.; et al. Pfam: The Protein Families Database in 2021. *Nucleic Acids Res.* **2021**, *49*, D412–D419. [[CrossRef](#)]