On the Uniqueness Theorem for Pseudo-Additive Entropies

Petr Jizba [1,†,*] **and Jan Korbel** [1,2,3,†] (iD)

1   Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague, Břehová 7, Prague 115 19, Czech Republic
2   Section for Science of Complex Systems, CeMSIIS, Medical University of Vienna, Spitalgasse 23, Vienna 1090, Austria
3   Complexity Science Hub Vienna, Josefstädterstrasse 39, Vienna 1090, Austria; korbeja2@fjfi.cvut.cz
*   Correspondence: p.jizbal@fjfi.cvut.cz; Tel.: +420-224-358-295
†   These authors contributed equally to this work.

Abstract: The aim of this paper is to show that the Tsallis-type (q-additive) entropic chain rule allows for a wider class of entropic functionals than previously thought. In particular, we point out that the ensuing entropy solutions (e.g., Tsallis entropy) can be determined uniquely only when one fixes the prescription for handling conditional entropies. By using the concept of Kolmogorov–Nagumo quasi-linear means, we prove this with the help of Darótzy's mapping theorem. Our point is further illustrated with a number of explicit examples. Other salient issues, such as connections of conditional entropies with the de Finetti–Kolmogorov theorem for escort distributions and with Landsberg's classification of non-extensive thermodynamic systems are also briefly discussed.

Keywords: pseudo-additive entropy; entropic chain rule; conditional entropy; Darótzy's mapping

1. Introduction

During the last two decades, complex dynamical systems have undergone an important conceptual shift. The catalyst was the infusion of new ideas from the theory of critical phenomena (scaling laws), (multi-)fractals and trees, the renormalization group, random matrix theory and information theory. On the other hand, the usual Boltzmann–Gibbs statistics, while successful in describing stationary systems characterized by ergodicity or metric transitivity, fails to reproduce the statistical behavior of many real-world complex systems in biology, astrophysics, geology and economic and social sciences. In recent years, the use of a new paradigm known as generalized statistics has become very popular in complex systems. The notion "generalized statistics" refers to statistical systems that are described via broad, (semi-)heavy-tailed distributions. Examples include generalized hyperbolic distributions, Meixner distributions, Weibull distributions and various power-law tail distributions (e.g., Zipf–Pareto, Lévy or Tsallis-type distributions). In particular, the statistics associated with power-law tails accounts for a rich class of phenomena observed in many real-world systems ranging from financial markets, physics and biology to geoscience (see, e.g., [1,2] and the citations therein).

Currently, there is a number of key techniques for dealing with generalized-statistics-based complex systems. Among these one can especially mention: (a) innate general statistics approaches such as the q-deformed thermostatistics of Tsallis [1], the superstatistics of Beck et al. [3] and various generalized non-extensive entropies [4–8]; (b) information-theoretic concepts such as the information theory of Rényi [9,10], transfer entropies [11,12], complexity theory [1,13] or information geometry [14]. Particularly the concept of entropy, both in its information-theoretic and combinatorial disguise, plays a central role here. It is, however, clear that entropy can provide meaningful and technically-sound modus operandi only if there is a good guiding principle behind its formulation. Admittedly, from a

mathematical point of view, the most satisfactory prescriptions are based on axiomatization of entropy or, preferably, on some operationally-viable coding theorem.

In the present paper, we focus on the axiomatic underpinning of pseudo-additive entropies that are currently being actively investigated in the context of non-equilibrium statistical physics and information theory (both classical and quantum) [1,2,15,16]. Among pseudo-additive entropies, we select and put under a deeper scrutiny a subclass of the so-called q-additive entropies. In this framework, we show that commonly-used axioms for q-additive entropies are prone to have multiplicity of distinct solutions with a common cause residing in the way that conditional entropies are handled. To do so, we extend arguments from [17] and demonstrate that the existence of different solutions is related to the Kolmogorov–Nagumo quasi-linear means and Darótzy's mapping theorem. This formal side of our exposition is further bolstered by a number of explicit examples of entropic functionals that all satisfy the same q-additive entropic chain rules.

The structure of the paper is as follows. After this Introduction, we present in Section 2 a concise overview of some standard (pseudo-)additive entropic rules, which yield degenerate solutions for entropic functionals. This will be important in Section 3 where the same phenomenon will be observed in the context of entropic chain rules. Despite their higher restrictiveness, the chain rules still lead to degenerate solutions, and this fact is present even in the (more sophisticated) q-additive entropic chain rules as demonstrated in Section 4. In Section 5, we identify the root cause of this situation in the way that conditional entropies are handled. In particular, with the help of Darótzy's mapping, we prove that any potential degeneracy at the level of additive entropic rules is automatically reflected also at the level of q-additivity. Additionally, we show that the analogous situation holds between normal and q-additive entropic chain rules whenever conditional entropies involved are defined via the Kolmogorov–Nagumo quasi-linear means. We close Section 5 by proving the de Finetti–Kolmogorov theorem for escort distributions. This allows demonstrating how more restrictive the q-additive entropic chain rules are in comparison with simple q-additive rules. Finally, Section 6 summarizes our results and discusses some further consequences.

2. Entropic (Pseudo-)Additivity Rules

2.1. Some (Pseudo-)Additivity Rules

Suppose X and Y are two events, e.g., messages, having W_X and W_Y as the possible number of states, respectively. The two events taken together form a joint event that we denote as XY. If X and Y are independent of each other, every combination of X and Y's values is a possible joint event, and so, $W_{XY} = W_X W_Y$. In this case, the entropy for a composite bipartite system XY, the joint entropy $H(X,Y)$, is additive, i.e.,

$$H(X,Y) = H(X) + H(Y). \tag{1}$$

On the other hand, if the events are not independent (and do not interact), it may be that some combinations of X and Y are not allowed (i.e., $W_{XY} < W_X W_Y$), so that the joint entropy $H(X,Y)$ may be less than $H(X) + H(Y)$, a property known as the subadditivity of the entropy. This would mean that a system increases its entropy by splitting up into separate systems. This so-called fragmentation is typical, e.g., in entangled quantum systems [18]. We will have more to say about this in Section 5.

In practice, the entropy behavior in a number of independent bipartite systems cannot be easily grasped via simple additivity prescription (1). The reasons behind this might be the fact that some marginal states (e.g., surface states) are neglected or that the long-range correlations are invalidating the assumption of independence, and the concept of independence is used only as a convenient approximation. Whatever the reason, the modus operandi in these cases is to resort to some simple one-parameter deformation of the additivity rule that grasps in one way or another

the non-additive contributions. Such deformations are known as pseudo-additive entropic forms. The most commonly-used versions are; Tsallis-type additivity [1]:

$$H(X, Y) = H(X) + H(Y) + (1 - q)H(X)H(Y),$$
(2)

Landsberg-type additivity [19]:

$$H(X, Y) = H(X) + H(Y) + (q - 1)H(X)H(Y),$$
(3)

δ-additivity [1,20] (used, e.g., in anti-de Sitter/conformal field theory (AdS/CFT) correspondence):

$$H(X, Y)^{1/\delta} = H(X)^{1/\delta} + H(Y)^{1/\delta},$$
(4)

Masi–Czachor-type supra-additivity [21], etc.

Let us point out that it might happen that the above (pseudo-)additivity rules hold also for more generic joint systems for which $W_{XY} \neq W_X W_Y$. In such cases, one speaks about (pseudo-)extensivity [6]. Since the (pseudo-)extensivity is more a characteristic property of the actual interaction/correlation between X and Y rather than entropy itself, we shall not dwell on this issue here. Let us just mention that the actual form of non-extensivity and the value of parameters can be often connected to some specific physical phenomena [8,22]. In Section 3, we will see that, from a strictly mathematical standpoint, it is logically more satisfactory to deal with dependent, but non-interacting systems. In such cases, the above (pseudo-)additivity rules will be replaced with more restrictive entropy chain rules.

2.2. Degeneracy in Solutions

It is well known that there is a number of "logically consistent", but form-inequivalent entropic functionals satisfying the simple additivity rule (1). Examples are provided by Shannon entropy (and ensuing Gibbs and von Neumann entropies) or Rényi entropy. In this connection, we should emphasize perhaps a less known fact that a similar situation holds also for pseudo-additivity. For instance, by assuming that with X is associated probability $\mathcal{P}_X = \{p_i\}_{i=1}^n$, then the Tsallis-type pseudo-additivity rule (2) is not only satisfied with Tsallis entropy:

$$S_q(X) \equiv S_q(\mathcal{P}_X) = \frac{\sum_{i=1}^n p_i^q - 1}{1 - q},$$
(5)

as one would expect, but also with Landsberg's entropy [19]:

$$H_L(X) = \frac{\left(\sum_{i=1}^n p_i^{2-q}\right)^{-1} - 1}{1 - q},$$
(6)

or with Behara–Chawla γ-entropy [23]:

$$H_{BC}(X) = \frac{1 - \left(\sum_{i=1}^n p_i^{1/\gamma}\right)^\gamma}{1 - 2^{\gamma-1}}, \quad q = 2 - 2^{\gamma-1}.$$
(7)

On the same footing, one may show that the (AdS/CFT) δ-additivity is satisfied with a class of the so-called Segal entropies [24]. Let us finally stress that the actual range of admissible q's in (5)–(7) could be restricted by imposing some further conditions (going beyond Tsallis-type additivity rule), e.g., expansibility, concavity [1], Schur concavity [25], Lesche stability [26], robustness [27], etc. Here, we shall not dwell on this issue.

2.3. Uniqueness Issue: Part I

At this stage, there are two points that should be noted. First, the above (pseudo-)additive forms (1)–(4) are true only for independent events. Second, one could entertain the idea that when the entropic chain rules (which employ also the conditional entropy $H(Y|X)$) are used instead, then their higher restrictiveness could remove the unwanted degeneracy and lead to a single unique solution. This might seem as a good strategy particularly because the entropic chain rule with ensuing $H(Y|X)$ is of practical relevance in a number of areas. In this connection, one can mention applications in:

1. information theory, where it describes input-output information of the (quantum) communication channel and enters in the data processing inequality [28],
2. non-equilibrium and complex dynamics, where it enters in the May–Wigner criterion for the stability of dynamical systems and helps to estimate the connectivity of the network of system exchanges [29],
3. time series and data analysis, where it describes transfer entropies in bivariate time series and the degree of synchronization between two signals [12].

In the following two sections, we will see what the entropic chain rules and their pseudo-additive extensions can really say about the uniqueness of entropic functionals.

3. Entropic Chain Rule

Additive Entropic Chain Rule: Some Fundamentals

Let us consider two generally dependent events. The elementary (Shannon-type) entropic chain rule for ensuing random events X and Y reads:

$$H(X,Y) \;=\; H(X) \;+\; H(Y|X), \tag{8}$$

where $H(X)$ is a single-random event entropy, $H(X,Y)$ is joint entropy and $H(Y|X)$ represents the related conditional entropy of Y given X. The meaning of the chain rule (8) is depicted in Figure 1.
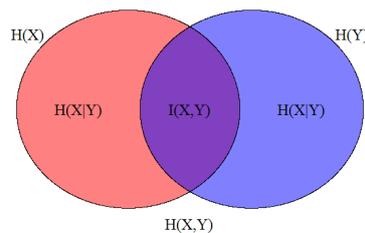

Figure 1. Entropy Venn diagram for two random events. The symbol $I(X,Y)$ denotes the mutual information, i.e., $I(X,Y) = H(X) + H(Y) - H(X,Y)$.

By induction, one can generalize the previous relation (8) to:

$$
\begin{aligned}
H(X_1, X_2, \ldots, X_n) \;&=\; H(X_1) \;+\; H(X_2|X_1) \;+\; H(X_3|X_2, X_1) \;+\; H(X_4|X_3, X_2, X_1) \;+\; \ldots \\
&=\; \sum_{i=1}^{n} H(X_i | X_{i-1}, X_{i-2}, \ldots, X_1).
\end{aligned}
\tag{9}
$$

At this stage, we can ask ourselves the following question. By imposing simple consistency conditions, such as:

1. continuity, i.e., when $H(\mathcal{P}_X)$ is a continuous function of all arguments,

2.   maximality, i.e., when for given n, $H(\mathcal{P}_X)$ is maximal for $\mathcal{P}_X = (1/n, \ldots, 1/n)$ and
3.   expansibility, i.e., when $H(p_1, \ldots, p_n, 0) = H(p_1, \ldots, p_n)$,

to what extent is $H(X)$ (with the above chain rule) unique? We might notice in passing that the above consistency conditions are nothing but Kolmogorov Axioms 1–3; see, e.g., [5]. It might perhaps come as a surprise that the uniqueness is not guarantied, at least not unless one prescribes how to handle conditional entropies.

   To understand this point better, let us state the basic defining properties of $H(X|Y)$:

1. $H(Y|X) = 0$ iff Y is completely determined by X, e.g., $H(g(X)|X) = 0$ (g is some function),
2. $H(Y|X) = H(X|Y) - H(X) + H(Y)$ entropic Bayes' rule,
3. $H(X|Y) \leq H(X)$ "second law of thermodynamics",
4. $H(X|Y) = H(X)$ if Y and X are independent (sometimes "if and only if" is required).

These properties still allow for some additional flexibility. Examples might be provided with Shannon's solution and Rényi's solution that both share the same chain condition (9) and the above defining conditions for $H(X|Y)$. In particular, for Shannon's entropy, the conditional entropy of X given Y is constructed in two steps:

$$H(X|Y = y) = -\sum_{x \in X} p(x|y) \log_2 p(x|y),$$

$$H(X|Y) = \sum_{y \in Y} p(y) H(X|Y = y) = -\sum_{x \in X; y \in Y} p(x,y) \log_2 p(x|y), \tag{10}$$

i.e., $H(X|Y)$ is defined via the linear mean. In contrast, in the case of Rényi's entropy, the conditional entropy of X given Y is defined by means of the quasi-linear Kolmogorov–Nagumo (KN) mean [30,31] in the following two-step sequence [32]:

$$I_q(X|Y = y) = \frac{1}{1-q} \log_2 \sum_{x \in X} p^q(x|y),$$

$$I_q(X|Y) = \phi^{-1} \left[\sum_{y \in Y} \rho_q(y) \phi \left(I_q(X|Y = y) \right) \right] = \frac{1}{1-q} \log_2 \frac{\sum_{x \in X; y \in Y} p^q(x,y)}{\sum_{y \in Y} p^q(y)}. \tag{11}$$

Here, the KN function ϕ reads [9,31]:

$$\phi(x) = 2^{(1-q)x}. \tag{12}$$

Particularly noteworthy is the appearance of the so-called escort distribution [32,33]:

$$\rho_q(y) = \frac{p^q(y)}{\sum_{y \in Y} p^q(y)}, \tag{13}$$

in Equation (11). Interestingly, a need for $\rho_q(y)$ in the definition of $I_q(X|Y)$ was already observed by Rényi in the 1960s [32]; some 30 years before the escort distribution was officially introduced in [33].

4. q-Additive Entropic Chain Rule

Uniqueness Issue: Part II

   In analogy with the preceding section, one defines the q-additive entropic chain rule for two random generally dependent events X and Y as:

$$H(X,Y) = H(X) + H(Y|X) + (1-q)H(X)H(Y|X). \tag{14}$$

By induction, one can generalize this to an n-partite system as:

$$
\begin{aligned}
H(X_1, X_2, \ldots, X_n) \;=\;& \sum_{i=1}^{n} H(X_i|X_{i-1}, \ldots, X_1) \\
+\;& (1-q)\sum_{i<j} H(X_i|X_{i-1}, \ldots, X_1)H(X_j|X_{j-1}, \ldots, X_1) \\
+\;& (1-q)^2 \sum_{i<j<k} H(X_i|X_{i-1}, \ldots, X_1)H(X_j|X_{j-1}, \ldots, X_1)H(X_k|X_{k-1}, \ldots, X_1) \\
+\;& \ldots + (1-q)^{n-1}\prod_{i=1}^{n} H(X_i|X_{i-1}, X_{i-2}, \ldots, X_1).
\end{aligned}
\tag{15}
$$

We can ask again a similar question as before, namely: To what extent is $H(X) \equiv H(\mathcal{P}_X)$ (subject to the above chain rule) unique?

As before, we should not be surprised to learn that the uniqueness is not guaranteed unless $H(X|Y)$ is specified. Indeed, the q-additive entropic chain rule together with consistency conditions:

1. $H(Y|X) = 0$ iff Y is completely determined by X, e.g., $H(g(X)|X) = 0$,
2. $\Phi(X,q)H(Y|X) = H(Y) - H(X) + \Phi(Y,q)H(X|Y)$ q-entropic Bayes' rule (The explicit form of the function $\Phi(X,q)$ depends on the type of q deformation. For instance, for q additivity, we have $\Phi(X,q) = 1 + (1-q)H(X)$.),
3. $H(X|Y) \leq H(X)$ "second law of thermodynamics",
4. $H(X|Y) = H(X)$ if Y and X are independent,

are not enough to fix the unique form of $H(X)$. Examples include: (a) Tsallis entropy where the conditional entropy of X given Y is defined as [34]:

$$
\begin{aligned}
S_q(X|Y=y) \;&=\; \frac{1}{1-q}\left[\sum_{x\in X} p(x|y)^q - 1\right], \\
S_q(X|Y) \;&=\; \sum_{y\in Y} \rho_q(y) S_q(X|Y=y),
\end{aligned}
\tag{16}
$$

i.e., $S_q(A|B)$ is defined via the linear (escort) mean, (b) Frank–Daffertshofer entropy [35], where the conditional entropy is defined as:

$$
\begin{aligned}
S_{FD}(X|Y=y) \;&=\; \frac{1}{1-q}\left[\left(\sum_{x\in X} p(x|y)^r\right)^{\frac{1-q}{1-r}} - 1\right], \\
S_{FD}(X|Y) \;&=\; \phi^{-1}\left[\sum_{y\in Y} \rho_q(y)\phi\left(S_{FD}(X|Y=y)\right)\right],
\end{aligned}
\tag{17}
$$

with the KN function $\phi(x) = \log_r e_q^x$, i.e., $S_{FD}(X|Y)$ is obtained via quasi-linear KN mean, (c) Sharma–Mittal entropy [36], where the conditional entropy is given by:

$$
\begin{aligned}
S_{SM}(X|Y=y) \;&=\; \frac{1}{\delta}\left[\left(\sum_{x\in X} p(x|y)^r\right)^{\frac{1-q}{1-r}} - 1\right], \\
S_{SM}(X|Y) \;&=\; \phi^{-1}\left[\sum_{y\in Y} \rho_q(y)\phi\left(S_{SM}(X|Y=y)\right)\right].
\end{aligned}
\tag{18}
$$

Here, $\delta = 2^{1-q} - 1$ and $\phi(x) = \log_r e_{\delta-1}^x$, i.e., $S_{SM}(X|Y)$ is again defined the via quasi-linear KN mean. For more details on the applications of S_{FD} and S_{SM}, see [35,37,38]

5. Theoretical Justification of Degeneracy

From the previous examples, one may observe that the multiplicity of possible entropic functionals that are compatible with ordinary entropic chain rules and/or q-additive chain rules can be associated (at least in part) with a specific choice of the KN average, which enters the definition of $H(X|Y)$. In this connection, the reader should note that KN averages do not spoil any of the defining properties of $H(X|Y)$ listed in Sections 3 and 4.

The aim of this section is to show that whenever a degeneracy exists at the level of ordinary entropic chain rules (as it always does in accordance with Section 3), then one can systematically generate degeneracy at the level of q-additive chain rules. To this end, we prove the following theorem:

Theorem 1. *Darótzy's mapping is a monotonic mapping $h_\gamma : \mathbb{R} \to \mathbb{R}$ such that:*

$$h_\gamma(x + y) = h_\gamma(x) + h_\gamma(y) + \gamma h_\gamma(x) h_\gamma(y) \equiv h_\gamma(x) \oplus_\gamma h_\gamma(y),$$

h_γ is parameterized by a, λ and γ, so that:

$$h_\gamma(x) = \begin{cases} ax, & a > 0 \text{ for } \lambda = 0 \\ (2^{\lambda x} - 1)/\gamma, & \lambda\gamma > 0 \text{ for } \lambda \neq 0. \end{cases}$$

Statement: Let us have the additive chain rule with $H(X|Y)$ defined with KN function $\phi(x)$. Darótzy's mapping then generates the q-additive chain rule (provided we set $\gamma = 1 - q$) where new $H(X|Y)$ is defined via KN function $\phi \circ h_\gamma^{-1}$.

Proof of Theorem 1. First, proof of the validity of Darótzy's mapping is by a simple inspection. Second, let us assume that certain entropic functional $H(X)$ with $H(Y|X)$ defined via some KN function ϕ satisfies additive chain rule:

$$H(X, Y) = H(X) + H(Y|X).$$

Darótzy's mapping then ensures that the entropic functional $h_\gamma(H(X))$ satisfies:

$$h_\gamma(H(X, Y)) = h_\gamma(H(X)) \oplus_\gamma h_\gamma(H(X|Y)),$$

with:

$$h_\gamma(H(X|Y)) = h_\gamma \circ \phi^{-1} \left[\sum_{y \in Y} \rho_r(y) \phi \circ h_\gamma^{-1}(h_\gamma(H(X|Y = y))) \right].$$

The actual values of a and λ are immaterial for the q-additivity as the Darótzy mapping generates q-additivity only via parameter γ. The last line shows that the new conditional entropy is defined via KN function $\phi \circ h_\gamma^{-1}$, while the escort parameter r is generally unrelated to q. Consequently, the ensuing q-additive entropy represents a three-parametric functional with parameters λ, r and q (apart from parameters originating from the KN function ϕ). □

The core idea of Theorem 1 can be phrased as follows: if there exist two different entropy functionals for a given additive chain rule with two different KN-related conditional entropies, then one can use Darótzy's mapping to map these two solutions to another two solutions that solve the q-additive chain rule. Simple examples of this mechanism are provided when we apply Darótzy's mapping

$h_{1-q}(x) = [2^{(1-q)x} - 1]/(1-q)$ to both Shannon entropy and Rényi entropy. In the first case, we obtain the entropy:

$$S(X) \equiv h_{1-q}(H(X)) = \frac{1}{1-q}\left(2^{-(1-q)\sum_{x \in X} p(x) \log_2 p(x)} - 1\right) = \frac{1}{1-q}\left(\prod_{x \in X} p(x)^{(q-1)p(x)} - 1\right), \quad (19)$$

while in the second case we have:

$$S_R(X) \equiv h_{1-q}(I_q(X)) = \frac{1}{1-q}\left(\sum_{x \in X} p(x)^q - 1\right), \quad (20)$$

which is nothing but the Tsallis entropy $S_q(X)$. Both $S(X)$ and $S_R(X)$ satisfy the q-additive chain rules provided the corresponding conditional entropies are defined via KN functions $\phi = h_{1-q}^{-1}$ and $\phi = 2^{(1-q)x} \circ h_{1-q}^{-1} = (1-q)x + 1$, respectively.

Along the same lines, one can apply two different Darótzy's mappings to the same seeding entropy to get two distinct solutions for the same q-additive chain rule. For instance, when we apply to Rényi entropy $I_r(X)$ Darótzy's mappings $h_\delta(x) = [2^{(1-q)x} - 1]/\delta$ with $\delta = 2^{1-q} - 1$, we obtain the Frank–Daffertshofer entropy (17). When $\delta = 1 - q$, we obtain the Sharma–Mittal entropy (18). For some further details on this issue, see, e.g., [4].

Darótzy's mapping can also be directly used to generate from (simple) additive entropies new q-additivity entropies. However, many of such newly generated entropic functionals are inconsistent as they do not satisfy the entropy chain rule in the sense that there is no consistent definition of conditional entropy for them. This is the content of our second theorem, which highlights some tricky points related to joint escort distributions.

Theorem 2. *de Finetti–Kolmogorov relation and escort distributions: when working with escort distributions, the joint distributions derived from the original joint distributions $\{r_{kl}\}$ do not satisfy the de Finetti–Kolmogorov relation, i.e., the analogue of $r_{kl} = p_l r_{k|l}$ (here $\{p_l\}$ and $\{r_{k|l}\}$ are marginal and conditional distributions, respectively).*

Proof of Theorem 2. By using the standard de Finetti–Kolmogorov relation for original distributions, i.e., $r_{kl} = p_l r_{k|l}$, we may write the following chain of relations:

$$
\begin{aligned}
r_{kl} = p_l r_{k|l} &\Leftrightarrow r_{kl}^q = p_l^q r_{k|l}^q \Leftrightarrow R(q)_{kl} = \frac{r_{kl}^q}{\sum_{mn} r_{mn}^q} = \frac{p_l^q r_{k|l}^q}{\sum_{mn} r_{mn}^q} \\
&= \frac{p_l^q r_{k|l}^q}{\sum_n (p_n^q \sum_m r_{m|n}^q)} \neq \frac{p_l^q}{\sum_n p_n^q} \frac{r_{k|l}^q}{\sum_m r_{m|l}^q} \equiv \tilde{R}(q)_{kl}.
\end{aligned}
$$

Here $\tilde{R}(q)_{kl}$ is the correct would-be joint escort distribution. Note that $\tilde{R}(q)_{kl}$ is not the escort of r_{kl} and $\tilde{R}(q)_{kl} = R(q)_{kl}$ iff events are independent [17]. □

An important upshot of this theorem is that some entropies (particularly those born out of escort distributions) might satisfy simple q-additivity, but not necessarily the q-additive entropic chain rule. The reason is that the more restrictive chain rule is sensitive to the failure of the de Finetti–Kolmogorov relation for joint escort distributions. A typical example is provided by the Jizba–Arimitsu entropy [17,25,39]:

$$S_{JA,q}(X) = \frac{1}{1-q}\left(2^{-(1-q)\sum_k \rho_k(q) \log_2 p_k} - 1\right) = h_{1-q}\left(-\sum_k \rho_k(q) \log_2 p_k\right),$$

where $h_{1-q}(x)$ is the KN function $h_{1-q}(x) = [2^{(1-q)x} - 1]/(1-q)$. Because $-\sum_k \rho_k(q) \log_2 p_k$ satisfies the simple additivity rule for independent events, Darótzy's mapping ensures that $S_{JA,q}(X)$ satisfies

the q-additivity. On the other hand, $-\sum_k \rho_k(q) \log_2 p_k$ does not satisfy the simple chain rule due to failure of the de Finetti–Kolmogorov relation for joint events. Consequently, $S_{JA,q}(X)$ is not q-extensive.

In this connection, we might recall Landsberg's classification of non-extensive thermodynamics. In 1990s P.T. Landsberg classified types of thermodynamics according to all possible functional properties of the entropy [19]. By defining entropy classes:

1. Superadditivity, S: $H(\mathcal{P}_{X+Y}) \geq H(\mathcal{P}_X) + H(\mathcal{P}_Y)$,
2. Homogeneity, H: $H(\mathcal{P}_{\lambda X}) = \lambda H(\mathcal{P}_X)$,
3. Concavity, C: $H(\mathcal{P}_{\lambda X+(1-\lambda)Y}) \geq \lambda H(\mathcal{P}_X) + (1-\lambda)H(\mathcal{P}_Y)$.

Landsberg concluded that there are only six logically-consistent thermodynamic classes: SHC, $S\bar{H}C$, $S\bar{H}\bar{C}$, $\bar{S}\bar{H}C$, $\bar{S}\bar{H}C$, $\bar{S}\bar{H}\bar{C}$. Types $SH\bar{C}$ are $\bar{S}HC$ are logically impossible [19]; see Figure 2.
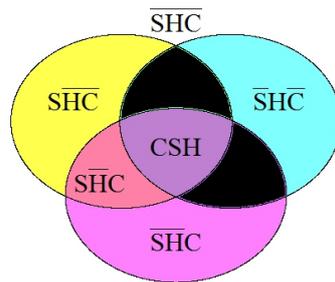

Figure 2. Venn diagram for Landsberg's thermodynamic types.

This classification is, however, based entirely on entropies for independent events. Since the entropic (pseudo-)additivity rule and chain rule for a given system do not need to follow the same pattern (e.g., they do not need to be both q-additive), the entropy class might change when conditional probabilities are included. For instance, according to Landsberg, $S_{JA,q}(X)$ should, for $q < 1$, belong to the $S\bar{H}C$ class, while the corresponding chain-rule generalization allows for flipping between S and $\bar{S}$; see [17].

6. Conclusions

In this paper, we have studied the uniqueness of entropic functionals under their composition laws. We have first recalled a well-known fact that many standard pseudo-additive entropic rules allow for a wide class of distinct entropic functionals. This is, of course, an unpleasant feature as the entropy should be (within a prescribed operational or axiomatic framework) a unique quantifier of ignorance, chaoticity, entanglement, etc. Since pseudo-additive laws employ only independent events, one might hope that the use of entropic chain rules (which deal also with dependent events) would restrict, or potentially even eliminate, the unwanted degeneracy. Here, we have specifically concentrated on a sub-class of pseudo-additive rules, the so-called q-additive entropic chain rules, and have shown that though some entropy functionals indeed cease to exist in this restrictive framework, one still finds degeneracy. The root cause behind this can be traced down to the way that a conditional entropy $H(X|Y)$ is handled. There is, in fact, a flexibility in the definition of permissible $H(X|Y)$'s. We have shown how such permissible $H(X|Y)$'s for any desired q-additive entropic chain rule can be systematically generated with the help of KN quasi-linear means and Darótzy's mapping theorem. This formal side of our exposition was further bolstered by a number of explicit examples of entropic functionals that all satisfy the same q-additive entropic chain rules.

In view of the results obtained, the natural question presents itself: What is a defining property of the currently popular concept of non-extensive statistics [1]? Is it, (a) pseudo-additivity (as indicated by Landsberg's classification), (b) the pseudo-additive chain rule (with generalized conditional entropies) or (c) power-law-type entropy maximizers. This question still awaits a satisfactory answer.

Author Contributions: Both P.J. and J.K. participated equally on theoretical calculations and on the writing of the manuscript.

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

Abbreviations

The following abbreviation is used in this manuscript:

KN   Kolmogorov–Nagumo

References

1. Tsallis, C. *Introduction to Nonextensive Statistical Mechanics; Approaching a Complex World*; Springer: New York, NY, USA, 2009.
2. Naudts, J. *Generalised Thermostatistics*; Springer: London, UK, 2011.
3. Beck, C.; Cohen, E.G.D. Superstatistics. *Phys. A* **2003**, *322*, 267–275.
4. Karmeshu, J. (Ed.) *Entropy Measures, Maximum Entropy Principle and Emerging Applications*; Springer: Berlin, Germany, 2003.
5. Hanel, R.; Thurner, S. A comprehensive classification of complex statistical systems and an axiomatic derivation of their entropy and distribution functions. *EPL Europhys. Lett.* **2011**, *93*, 50006-p1–50006-p6.
6. Hanel, R.; Thurner, S. When do generalized entropies apply? How phase space volume determines entropy. *EPL Europhys. Lett.* **2011**, *96*, 50003-p1–50005-p6.
7. Tempesta, P. Group entropies, correlation laws, and zeta functions. *Phys. Rev. E* **2011**, *84*, 021121-1–021121-10.
8. Biró, T.; Barnaföldi, G.; Ván, P. New entropy formula with fluctuating reservoir. *Phys. A* **2015**, *417*, 215–220.
9. Jizba, P.; Arimitsu, T. The world according to Rényi: Thermodynamics of multifractal systems. *Ann. Phys.* **2004**, *312*, 17–59.
10. Jizba, P.; Ma, Y.; Hayes, A.; Dunningham, J. One-parameter class of uncertainty relations based on entropy power. *Phys. Rev. E* **2016**, *93*, 060104(R)-1–060104(R)-5.
11. Schreiber, T. Measuring Information Transfer. *Phys. Rev. Lett.* **2000**, *85*, 461–464.
12. Jizba, P.; Kleinert, H.; Shefaat, M. Rényi's information transfer between financial time series. *Phys. A* **2012**, *391*, 2971–2989.
13. Eom, Y.-H.; Jo, H.-H. Using friends to estimate heavy tails of degree distributions in large-scale complex networks. *Sci. Rep.* **2015**, *5*, 09752-1–09752-9.
14. Bercher, J.-F. Some properties of generalized Fisher information in the context of nonextensive thermostatistics. *Phys. A* **2013**, *392*, 3140–3154.
15. Short, A.J.; Wehner, S. Entropy in general physical theories. *New J. Phys.* **2010**, *12*, 033023.
16. Majhi, A. Non-extensive statistical mechanics and black hole entropy from quantum geometry. *arXiv* **2017**, arXiv:1703.09355. doi:10.1016/j.physletb.2017.10.043.
17. Jizba, P.; Korbel, J. Remarks on "Comments on 'On q-non-extensive statistics with non-Tsallisian entropy'" [Physica A 466 (2017) 160]. *Phys. A* **2017**, *468*, 238–243.
18. Schumacher, B.; Westmoreland, M. *Quantum Processes, Systems, and Information*; Cambridge University Press: Cambridge, UK, 2010.
19. Landsberg, P.T. Entropies Galore! *Braz. J. Phys.* **1999**, *29*, 46–49.
20. Vos, G. Generalized additivity in unitary conformal field theories. *Nucl. Phys. B* **2015**, *899*, 91–111.
21. Masi, M. A step beyond Tsallis and Renyi entropies. *Phys. Lett. A* **2005**, *338*, 217–224.
22. Korbel, J. Rescaling the nonadditivity parameter in Tsallis thermostatistics. *Phys. Lett. A* **2017**, *381*, 2588–2592.
23. Behara, M.; Chawla, J.S. Generalized gamma entropy. *Sel. Stat. Can.* **1974**, *2*, 15–38.
24. Ochs, W.; Spohn, H.A. Characterization of the Segal entropy. *Rep. Math. Phys.* **1978**, *14*, 75–87.
25. Jizba, P.; Korbel, J. On q-non-extensive statistics with non-Tsallisian entropy. *Phys. A* **2016**, *444*, 808–827.
26. Lesche, B. Instabilities of Rényi entropies. *J. Stat. Phys.* **1982**, *27*, 419–422.
27. Hanel, R.; Thurner, S.; Tsallis, C. On the robustness of q-expectation values and Rényi entropy. *EPL Europhys. Lett.* **2009**, *85*, 20005-p1–20005-p6.
28. Campbell, L.L. A coding theorem and Rényi's entropy. *Inf. Control* **1965**, *8*, 423–429.
29. Hastings, H.M. The May-Wigner stability theorem. *J. Theor. Biol.* **1982**, *97*, 155–166.

30. Nagumo, M. Über eine Classe der Mittelwerte. *Jpn. J. Math.* **1930**, *7*, 71–79.

31. Aczél, J. *Lectures on Functional Equations an dtheir Applications*; Academic Press: New York, NY, USA, 1966.

32. Rényi, A. *Selected Papers of Alfred Rényi*; Akademia Kiado: Budapest, Hungary, 1976.

33. Beck, C.; Schlögl, F. *Thermodynamics of Chaotic Systems: An Introduction*; Cambridge University Press: Cambridge, UK, 1993.

34. Abe, S. Axioms and uniqueness theorem for Tsallis entropy. *Phys. Lett. A* **2000**, *271*, 74–79.

35. Frank, T.D.; Daffertshofer, A. Exact time-dependent solutions of the Renyi Fokker-Planck equation and the Fokker-Planck equations related to the entropies proposed by Sharma and Mittal. *Phys. A* **2000**, *285*, 351–366.

36. Sharma, B.D.; Mittal, D.P. New non-additive measures of entropy for discrete probability distributions. *J. Math. Sci.* **1975**, *10*, 28–40.

37. Aktürk, O.Ü.; Aktürk, E.; Tomak, M. Can Sobolev inequality be written for Sharma-Mittal entropy? *Int. J. Theor. Phys.* **2008**, *47*, 3310–3320.

38. Kosztołowicz, T.; Lewandowska, K.D. First-passage time for subd-diffusion: The nonadditive entropy approach versus the fractional model. *Phys. Rev. E* **2012**, *86*, 021108-1–021108-11.

39. Çankaya, M.N.; Korbel, J. On statistical properties of Jizba–Arimitsu hybrid entropy. *Phys. A* **2017**, *474*, 1–10.