

Article

On Some Properties of Tsallis Hypoentropies and Hypodivergences

Shigeru Furuichi ^{1,*}, Flavia-Corina Mitroi-Symeonidis ² and Eleutherius Symeonidis ³

¹ Department of Information Science, College of Humanities and Sciences, Nihon University, 3-25-40, Sakurajyousui, Setagaya-ku, Tokyo, 156-8550, Japan

² Faculty of Engineering Sciences, LUMINA - University of South-East Europe, Șos. Colentina 64b, Bucharest, RO-021187, Romania; E-Mail: fcmitroi@yahoo.com

³ Mathematisch-Geographische Fakultät, Katholische Universität Eichstätt-Ingolstadt, 85071 Eichstätt, Germany; E-Mail: e.symeonidis@ku-eichstaett.de

* Author to whom correspondence should be addressed; E-Mail: furuichi@chs.nihon-u.ac.jp; Tel.:+81-3-3329-1151; Fax:+81-3-5317-9770.

External Editor: Antonio M. Scarfone

Received: 15 September 2014; in revised form: 8 October 2014 / Accepted: 8 October 2014 / Published: 15 October 2014

Abstract: Both the Kullback–Leibler and the Tsallis divergence have a strong limitation: if the value zero appears in probability distributions (p_1, \dots, p_n) and (q_1, \dots, q_n) , it must appear in the same positions for the sake of significance. In order to avoid that limitation in the framework of Shannon statistics, Ferreri introduced in 1980 hypoentropy: “such conditions rarely occur in practice”. The aim of the present paper is to extend Ferreri’s hypoentropy to the Tsallis statistics. We introduce the Tsallis hypoentropy and the Tsallis hypodivergence and describe their mathematical behavior. Fundamental properties, like nonnegativity, monotonicity, the chain rule and subadditivity, are established.

Keywords: mathematical inequality; Tsallis entropy; Tsallis hypoentropy; Tsallis hypodivergence; chain rule; subadditivity

MSC classifications: 26D15; 94A17

1. Preliminaries

Throughout this paper, X , Y and Z denote discrete random variables taking on the values $\{x_1, \dots, x_{|X|}\}$, $\{y_1, \dots, y_{|Y|}\}$ and $\{z_1, \dots, z_{|Z|}\}$, respectively, where $|A|$ denotes the number of the values of the discrete random variable A . We denote the discrete random variable following a uniform distribution by U . We set the probabilities as $p(x_i) \equiv Pr(X = x_i)$, $p(y_j) \equiv Pr(Y = y_j)$ and $p(z_k) \equiv Pr(Z = z_k)$. If $|U| = n$, then $p(u_k) = \frac{1}{n}$ for all $k = 1, \dots, n$. In addition, we denote by $p(x_i, y_j) = Pr(X = x_i, Y = y_j)$, $p(x_i, y_j, z_k) = Pr(X = x_i, Y = y_j, Z = z_k)$ the joint probabilities, by $p(x_i|y_j) = Pr(X = x_i|Y = y_j)$, $p(x_i|y_j, z_k) = Pr(X = x_i|Y = y_j, Z = z_k)$ the conditional probabilities, and so on.

The notion of entropy was used in statistical thermodynamics by Boltzmann [1] in 1871 and Gibbs [2] in 1902, in order to quantify the diversity, uncertainty and randomness of isolated systems. Later, it was seen as a measure of “information, choice and uncertainty” in the theory of communication, when Shannon [3] defined it by:

$$H(X) \equiv - \sum_{i=1}^{|X|} p(x_i) \log p(x_i). \quad (1)$$

In what follows, we consider $|X| = |Y| = |U| = n$, unless otherwise specified.

Making use of the concavity of the logarithmic function, one can easily check that the equiprobable states are maximizing the entropy, that is:

$$H(X) \leq H(U) = \log n. \quad (2)$$

The right-hand side term of this inequality has been known since 1928 as Hartley entropy [4].

For two random variables X and Y following distributions $\{p(x_i)\}$ and $\{p(y_i)\}$, the Kullback–Leibler [5] discrimination function (divergence or relative entropy) is defined by:

$$D(X||Y) \equiv \sum_{i=1}^n p(x_i)(\log p(x_i) - \log p(y_i)) = - \sum_{i=1}^n p(x_i) \log \frac{p(y_i)}{p(x_i)}. \quad (3)$$

(We note that the relative entropy is usually defined for two probability distributions $P = \{p_i\}$ and $Q = \{q_i\}$ as $D(P||Q) \equiv - \sum_{i=1}^n p_i \log \frac{q_i}{p_i}$ in the standard notation of information theory. $D(P||Q)$ is often rewritten by $D(X||Y)$ for random variables X and Y following the distributions P and Q . Throughout this paper, we use the style of Equation (3) for relative entropies to unify the notation with simple descriptions.) Here, the conventions $a \cdot \log \frac{0}{a} = -\infty$ ($a > 0$) and $0 \cdot \log \frac{b}{0} = 0$ ($b \geq 0$) are used. (We also note that the convention is often given in the following way with the definition of $D(X||Y)$. If there exists i , such that $p(x_i) \neq 0 = p(y_i)$, then we define $D(X||Y) \equiv +\infty$ (in this case, $D(X||Y)$ is no longer significant as an information measure). Otherwise, $D(X||Y)$ is defined by Equation (3) with the convention $0 \cdot \log \frac{0}{0} = 0$. This fact has been mentioned in the abstract of the paper.) In what follows, we use such conventions in the definitions of the entropies and divergences. However, we do not state them repeatedly.

It holds that:

$$H(U) - H(X) = D(X||U). \quad (4)$$

Moreover, the cross-entropy (or inaccuracy):

$$H^{(cross)}(X, Y) \equiv - \sum_{i=1}^n p(x_i) \log p(y_i) \tag{5}$$

satisfies the identity:

$$D(X||Y) = H^{(cross)}(X, Y) - H(X). \tag{6}$$

Many extensions of Shannon entropy have been studied. The Rényi entropy [6] and α -entropy [7] are famous. The mathematical results until the 1970s are well written in the book [8]. In the present paper, we focus on the hypoentropy introduced by Carlo Ferreri and the Tsallis entropy introduced by Constantino Tsallis.

The hypoentropy at the level λ (λ -entropy) was introduced in 1980 by Ferreri [9] as an alternative measure of information in the following form:

$$F_\lambda(X) \equiv \frac{1}{\lambda}(\lambda + 1) \log(\lambda + 1) - \frac{1}{\lambda} \sum_{i=1}^n (1 + \lambda p(x_i)) \log(1 + \lambda p(x_i)) \tag{7}$$

for $\lambda > 0$. According to Ferreri [9], the parameter λ can be interpreted as a measure of the information inaccuracy of economic forecast. As we will show that $F_\lambda(X) \leq H(X)$ in the second section, the name hypoentropy comes from this property.

On the other hand, Tsallis introduced a one-parameter extension of the entropy in 1988 in [10], for handling systems that appear to deviate from standard statistical distributions. It plays an important role in the nonextensive statistical mechanics of complex systems, being defined as:

$$T_q(X) \equiv - \sum_{i=1}^n p(x_i)^q \ln_q p(x_i) = \sum_{i=1}^n p(x_i) \ln_q \frac{1}{p(x_i)} \quad (q \geq 0, q \neq 1). \tag{8}$$

Here, the q -logarithmic function for $x > 0$ is defined by $\ln_q(x) \equiv \frac{x^{1-q}-1}{1-q}$, which converges to the usual logarithmic function $\log(x)$ in the limit $q \rightarrow 1$. The Tsallis divergence (relative entropy) [11] is given by:

$$S_q(X||Y) \equiv \sum_{i=1}^n p(x_i)^q (\ln_q p(x_i) - \ln_q p(y_i)) = - \sum_{i=1}^n p(x_i) \ln_q \frac{p(y_i)}{p(x_i)}. \tag{9}$$

Note that some important properties of the Tsallis relative entropy were given in the papers [12–14].

2. Hypoentropy and Hypodivergence

For nonnegative real numbers, a_i and b_i ($i = 1, \dots, n$), we define the generalized relative entropy (for incomplete probability distributions):

$$D^{(gen)}(a_1, \dots, a_n || b_1, \dots, b_n) \equiv \sum_{i=1}^n a_i \log \frac{a_i}{b_i}. \tag{10}$$

Then, we have the so-called “log-sum” inequality:

$$\sum_{i=1}^n a_i \log \frac{a_i}{b_i} \geq \left(\sum_{i=1}^n a_i \right) \log \frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n b_i}, \tag{11}$$

with equality if and only if $\frac{a_i}{b_i} = const.$ for all $i = 1, \dots, n$.

If we impose the condition:

$$\sum_{i=1}^n a_i = \sum_{i=1}^n b_i = 1,$$

then $D^{(gen)}(a_1, \dots, a_n || b_1, \dots, b_n)$ is just the relative entropy,

$$D(a_1, \dots, a_n || b_1, \dots, b_n) \equiv \sum_{i=1}^n a_i \log \frac{a_i}{b_i}. \tag{12}$$

We put $a_i = \frac{1}{\lambda} + p(x_i)$ and $b_i = \frac{1}{\lambda} + p(y_i)$ with $\lambda > 0$ and $\sum_{i=1}^n p(x_i) = \sum_{i=1}^n p(y_i) = 1$, $p(x_i) \geq 0, p(y_i) \geq 0$. Then, we find that it is equal to the hypodivergence (λ -divergence) introduced by Ferreri in [9],

$$K_\lambda(X || Y) \equiv \frac{1}{\lambda} \sum_{i=1}^n (1 + \lambda p(x_i)) \log \frac{1 + \lambda p(x_i)}{1 + \lambda p(y_i)}. \tag{13}$$

Clearly, we have:

$$\lim_{\lambda \rightarrow \infty} K_\lambda(X || Y) = D(X || Y). \tag{14}$$

Using the ‘‘log-sum’’ inequality, we have the nonnegativity:

$$K_\lambda(X || Y) \geq 0, \tag{15}$$

with equality if and only if $p(x_i) = p(y_i)$ for all $i = 1, \dots, n$.

For the hypoentropy $F_\lambda(X)$ defined in Equation (7), we firstly show the fundamental relations. To do so, we prepare with the following lemma.

Lemma 1. For any $a > 0$ and $0 \leq x \leq 1$, we have

$$x(1 + a) \log(1 + a) \geq (1 + ax) \log(1 + ax). \tag{16}$$

Proof. We set $f(x) \equiv x(1 + a) \log(1 + a) - (1 + ax) \log(1 + ax)$. For any $a > 0$, we then have $\frac{d^2 f(x)}{dx^2} = \frac{-a^2}{1+ax} < 0$ and $f(0) = f(1) = 0$. Thus, we have the inequality. \square

Proposition 1. For $\lambda > 0$, we have the following inequalities:

$$0 \leq F_\lambda(X) \leq F_\lambda(U). \tag{17}$$

The equality in the first inequality holds if and only if $p(x_j) = 1$ for some j (then $p(x_i) = 0$ for all $i \neq j$). The equality in the second inequality holds if and only if $p(x_i) = 1/n$ for all $i = 1, \dots, n$.

Proof. From the nonnegativity of the hypodivergence Equation (15), we get:

$$0 \leq K_\lambda(X || U) \tag{18}$$

$$= \frac{1}{\lambda} \sum_{i=1}^n (1 + \lambda p(x_i)) \log(1 + \lambda p(x_i)) - \frac{1}{\lambda} (n + \lambda) \log \left(1 + \frac{\lambda}{n} \right). \tag{19}$$

Thus, we have:

$$- \frac{1}{\lambda} \sum_{i=1}^n (1 + \lambda p(x_i)) \log(1 + \lambda p(x_i)) \leq - \frac{1}{\lambda} (n + \lambda) \log \left(1 + \frac{\lambda}{n} \right). \tag{20}$$

Adding $\frac{1}{\lambda}(\lambda + 1) \log(\lambda + 1)$ to both sides, we have:

$$F_\lambda(X) \leq F_\lambda(U), \tag{21}$$

with equality if and only if $p(x_i) = 1/n$ for all $i = 1, \dots, n$.

For the first inequality, it is sufficient to prove:

$$(1 + \lambda) \log(1 + \lambda) - \sum_{i=1}^n (1 + \lambda p(x_i)) \log(1 + \lambda p(x_i)) \geq 0. \tag{22}$$

Since $\sum_{i=1}^n p(x_i) = 1$, the above inequality is written as:

$$\sum_{i=1}^n \{p(x_i)(1 + \lambda) \log(1 + \lambda) - (1 + \lambda p(x_i)) \log(1 + \lambda p(x_i))\} \geq 0, \tag{23}$$

so that we have only to prove:

$$p(x_i)(1 + \lambda) \log(1 + \lambda) - (1 + \lambda p(x_i)) \log(1 + \lambda p(x_i)) \geq 0, \tag{24}$$

for any $\lambda > 0$ and $0 \leq p(x_i) \leq 1$. Lemma 1 shows this inequality and the equality condition. \square

It is a known fact [9] that $F_\lambda(X)$ is monotonically increasing as a function of λ and:

$$\lim_{\lambda \rightarrow \infty} F_\lambda(X) = H(X), \tag{25}$$

whence its name, as we noted in the Introduction. Thus, the hypoentropy appears as a generalization of Shannon’s entropy. One can see that the hypoentropy also equals zero as the entropy does, in the case of certainty (*i.e.*, for a so-called pure state when all probabilities vanish, but one).

It also holds that:

$$F_\lambda(U) - F_\lambda(X) = K_\lambda(X||U). \tag{26}$$

It is of some interest for the reader to look at the hypoentropy that arises for equiprobable states,

$$F_\lambda(U) = \left(1 + \frac{1}{\lambda}\right) \log(1 + \lambda) - \left(1 + \frac{n}{\lambda}\right) \log\left(1 + \frac{\lambda}{n}\right). \tag{27}$$

Seen as a function of two variables, n and λ , it increases in each variable [9]. Since:

$$\lim_{\lambda \rightarrow \infty} F_\lambda(U) = \log n, \tag{28}$$

We shall call it Hartley hypoentropy. (Throughout the paper, we add the name Hartley to the name of mathematical objects whenever they are considered for the uniform distribution. In the same way, we proceed with the name Tsallis, which we add to the name of some mathematical objects that we define, to emphasize that they are used in the framework of Tsallis statistics. This means that we will have Tsallis hypoentropies, Tsallis hypodivergences, and so on.) We have the cross-hypoentropy:

$$F_\lambda^{(cross)}(X, Y) \equiv \left(1 + \frac{1}{\lambda}\right) \log(1 + \lambda) - \frac{1}{\lambda} \sum_{i=1}^n (1 + \lambda p(x_i)) \log(1 + \lambda p(y_i)). \tag{29}$$

It holds:

$$K_\lambda(X||Y) = F_\lambda^{(cross)}(X, Y) - F_\lambda(X) \geq 0, \tag{30}$$

therefore, we have $F_\lambda^{(cross)}(X, Y) \geq F_\lambda(X)$.

We can show an upper bound for $F_\lambda(X)$ as a direct consequence.

Proposition 2. *The following inequality holds.*

$$F_\lambda(X) \leq (1 - p_{max}) \log(1 + \lambda),$$

for all $\lambda > 0$, where $p_{max} \equiv \max \{p(x_1), \dots, p(x_n)\}$.

Proof. In the inequality (30), if for a fixed k , one takes the probability of the k -th component of Y to be $p(y_k) = 1$, then:

$$-\sum_{i=1}^n (1 + \lambda p(x_i)) \log(1 + \lambda p(x_i)) \leq -(1 + \lambda p(x_k)) \log(1 + \lambda). \tag{31}$$

This implies that:

$$F_\lambda(X) \leq \left(1 + \frac{1}{\lambda}\right) \log(1 + \lambda) - \frac{1}{\lambda} (1 + \lambda p(x_k)) \log(1 + \lambda) \tag{32}$$

$$= (1 - p(x_k)) \log(1 + \lambda). \tag{33}$$

Since k is arbitrarily fixed, the conclusion follows. \square

Remark 1. *It is of interest to notice now that, for the particular case $X = U$, we have:*

$$F_\lambda(U) \leq \left(1 - \frac{1}{n}\right) \log(1 + \lambda). \tag{34}$$

We add here one more detail: the inequality (34) can be verified using Bernoulli's inequality.

3. Tsallis Hypoentropy and Hypodivergence

Now, we turn our attention to the Tsallis statistics. We extend the definition of hypodivergences as follows:

Definition 1. *The Tsallis hypodivergence (q -hypodivergence, Tsallis relative hypoentropy) is defined by:*

$$D_{\lambda,q}(X||Y) \equiv -\frac{1}{\lambda} \sum_{i=1}^n (1 + \lambda p(x_i)) \ln_q \frac{1 + \lambda p(y_i)}{1 + \lambda p(x_i)} \tag{35}$$

for $\lambda > 0$ and $q \geq 0$.

Then, we have the relation:

$$\lim_{\lambda \rightarrow \infty} D_{\lambda,q}(X||Y) = S_q(X||Y) \tag{36}$$

which is the Tsallis divergence, and:

$$\lim_{q \rightarrow 1} D_{\lambda,q}(X||Y) = K_\lambda(X||Y) \tag{37}$$

which is the hypodivergence.

Remark 2. This definition can be also obtained from the generalized Tsallis relative entropy (for incomplete probability distributions $\{a_1, \dots, a_n\}$ and $\{b_1, \dots, b_n\}$):

$$D_q^{(gen)}(a_1, \dots, a_n || b_1, \dots, b_n) \equiv - \sum_{i=1}^n a_i \ln_q \frac{b_i}{a_i}, \tag{38}$$

by putting $a_i = \frac{1}{\lambda} + p(x_i)$ and $b_i = \frac{1}{\lambda} + p(y_i)$ for $\lambda > 0$.

The generalized relative entropy (10) and the generalized Tsallis relative entropy (38) can be written as the generalized f -divergence (for incomplete probability distributions):

$$D_f^{(gen)}(a_1, \dots, a_n || b_1, \dots, b_n) \equiv \sum_{i=1}^n a_i f\left(\frac{b_i}{a_i}\right) \tag{39}$$

for a convex function f on $(0, \infty)$ and $a_i \geq 0, b_i \geq 0$ ($i = 1, \dots, n$). See [15] and [16] for details.

By the concavity of the q -logarithmic function, we have the following “ \ln_q -sum” inequality:

$$- \sum_{i=1}^n a_i \ln_q \frac{b_i}{a_i} \geq - \left(\sum_{i=1}^n a_i \right) \ln_q \left(\frac{\sum_{i=1}^n b_i}{\sum_{i=1}^n a_i} \right), \tag{40}$$

with equality if and only if $\frac{a_i}{b_i} = const.$ for all $i = 1, \dots, n$. Using the “ \ln_q -sum” inequality, we have the nonnegativity of the Tsallis hypodivergence:

$$D_{\lambda,q}(X||Y) \geq 0, \tag{41}$$

with equality if and only if $p(x_i) = p(y_i)$ for all $i = 1, \dots, n$ (the equality condition comes from the equality condition of the “ \ln_q -sum” inequality and the condition $\sum_{i=1}^n p(x_i) = \sum_{i=1}^n p(y_i) = 1$).

Definition 2. For $\lambda > 0$ and $q \geq 0$, the Tsallis hypoentropy (q -hypoentropy) is defined by:

$$H_{\lambda,q}(X) \equiv \frac{h(\lambda, q)}{\lambda} \left\{ -(1 + \lambda) \ln_q \frac{1}{1 + \lambda} + \sum_{i=1}^n (1 + \lambda p(x_i)) \ln_q \frac{1}{1 + \lambda p(x_i)} \right\} \tag{42}$$

where the function $h(\lambda, q) > 0$ satisfies two conditions,

$$\lim_{q \rightarrow 1} h(\lambda, q) = 1 \tag{43}$$

and:

$$\lim_{\lambda \rightarrow \infty} \frac{h(\lambda, q)}{\lambda^{1-q}} = 1. \tag{44}$$

These conditions are equivalent to:

$$\lim_{q \rightarrow 1} H_{\lambda,q}(X) = F_\lambda(X) = \text{Hypoentropy} \tag{45}$$

and, respectively,

$$\lim_{\lambda \rightarrow \infty} H_{\lambda,q}(X) = T_q(X) = \text{Tsallis entropy}. \tag{46}$$

Some interesting examples are $h(\lambda, q) = \lambda^{1-q}$ and $h(\lambda, q) = (1 + \lambda)^{1-q}$.

Remark 3. It may be remarkable to discuss the Tsallis cross-hypoentropy. The first candidate for the definition of the Tsallis cross-hypoentropy is:

$$H_{\lambda,q}^{(cross)}(X, Y) \equiv \frac{h(\lambda, q)}{\lambda} \left\{ -(1 + \lambda) \ln_q \frac{1}{1 + \lambda} - \sum_{i=1}^n (1 + \lambda p(x_i))^q \ln_q (1 + \lambda p(y_i)) \right\} \quad (47)$$

which recovers the cross-hypoentropy defined in Equation (29) in the limit $q \rightarrow 1$. Then, we have:

$$\begin{aligned} H_{\lambda,q}^{(cross)}(X, Y) - H_{\lambda,q}(X) &= \frac{h(\lambda, q)}{\lambda} \sum_{i=1}^n (1 + \lambda p(x_i))^q \{ \ln_q (1 + \lambda p(x_i)) - \ln_q (1 + \lambda p(y_i)) \} \\ &= -\frac{h(\lambda, q)}{\lambda} \sum_{i=1}^n (1 + \lambda p(x_i)) \ln_q \frac{1 + \lambda p(y_i)}{1 + \lambda p(x_i)} \\ &= h(\lambda, q) D_{\lambda,q}(X||Y) \geq 0. \end{aligned}$$

The last inequality is due to the nonnegativity given in Equation (41). Since $\lim_{q \rightarrow 1} h(\lambda, q) = 1$ by the definition of the Tsallis hypoentropy (see Equation (43)), the above relation recovers the inequality (30) in the limit $q \rightarrow 1$.

The second candidate for the definition of the Tsallis cross-hypoentropy is:

$$\tilde{H}_{\lambda,q}^{(cross)}(X, Y) \equiv \frac{h(\lambda, q)}{\lambda} \left\{ -(1 + \lambda) \ln_q \frac{1}{1 + \lambda} + \sum_{i=1}^n (1 + \lambda p(x_i)) \ln_q \frac{1}{1 + \lambda p(y_i)} \right\}, \quad (48)$$

which also recovers the cross-hypoentropy defined in Equation (29) in the limit $q \rightarrow 1$. Then, we have:

$$\begin{aligned} \tilde{H}_{\lambda,q}^{(cross)}(X, Y) - H_{\lambda,q}(X) &= -\frac{h(\lambda, q)}{\lambda} \sum_{i=1}^n (1 + \lambda p(x_i)) \left\{ \ln_q \frac{1}{1 + \lambda p(x_i)} - \ln_q \frac{1}{1 + \lambda p(y_i)} \right\} \\ &= h(\lambda, q) \tilde{D}_{\lambda,q}(X||Y), \end{aligned}$$

where the alternative Tsallis hypodivergence has to be defined by:

$$\tilde{D}_{\lambda,q}(X||Y) \equiv -\frac{1}{\lambda} \sum_{i=1}^n (1 + \lambda p(x_i)) \left\{ \ln_q \frac{1}{1 + \lambda p(x_i)} - \ln_q \frac{1}{1 + \lambda p(y_i)} \right\}.$$

We have $\tilde{D}_{\lambda,q}(X||Y) \neq D_{\lambda,q}(X||Y)$ and $\lim_{q \rightarrow 1} \tilde{D}_{\lambda,q}(X||Y) = K_\lambda(X||Y)$. However, the nonnegativity of $\tilde{D}_{\lambda,q}(X||Y)$ ($q \geq 0$) does not hold in general, as the following counter-examples show. Take $\lambda = 1$, $n = 2$, $p(x_1) = 0.9$, $p(y_1) = 0.8$, $q = 0.5$, then $\tilde{D}_{\lambda,q}(X||Y) \simeq -0.0137586$. In addition, take $\lambda = 1$, $n = 3$, $p(x_1) = 0.3$, $p(x_2) = 0.4$, $p(y_1) = 0.2$, $p(y_2) = 0.7$ and $q = 1.9$, then $\tilde{D}_{\lambda,q}(X||Y) \simeq -0.0195899$. Therefore, we may conclude that Equation (47) is to be given the preference over Equation (48).

We turn to show the nonnegativity and maximality for the Tsallis hypoentropy.

Lemma 2. For any $a > 0$, $q \geq 0$ and $0 \leq x \leq 1$, we have:

$$x(1 + a) \ln_q \frac{1}{1 + a} \leq (1 + ax) \ln_q \frac{1}{1 + ax}. \quad (49)$$

Proof. We set $g(x) \equiv x(1 + a) \ln_q \frac{1}{1+a} - (1 + ax) \ln_q \frac{1}{1+ax}$. For any $a > 0$ and $q \geq 0$, we then have $\frac{d^2g(x)}{dx^2} = qa^2 \left(\frac{1}{1+ax}\right)^{2-q} \geq 0$ and $g(0) = g(1) = 0$. Thus, we have the inequality. \square

Proposition 3. For $\lambda > 0$, $q \geq 0$ and $h(\lambda, q) > 0$ satisfying (43) and (44), we have the following inequalities:

$$0 \leq H_{\lambda,q}(X) \leq H_{\lambda,q}(U). \tag{50}$$

The equality in the first inequality holds if and only if $p(x_j) = 1$ for some j (then $p(x_i) = 0$ for all $i \neq j$). The equality in the second inequality holds if and only if $p(x_i) = 1/n$ for all $i = 1, \dots, n$.

Proof. In a similar way to the proof of Proposition 1, for the first inequality, it is sufficient to prove:

$$-\sum_{i=1}^n \left\{ p(x_i)(1 + \lambda) \ln_q \frac{1}{1 + \lambda} - (1 + \lambda p(x_i)) \ln_q \frac{1}{1 + \lambda p(x_i)} \right\} \geq 0, \tag{51}$$

so that we have only to prove:

$$p(x_i)(1 + \lambda) \ln_q \frac{1}{1 + \lambda} \leq (1 + \lambda p(x_i)) \ln_q \frac{1}{1 + \lambda p(x_i)} \tag{52}$$

for any $\lambda > 0$, $q \geq 0$ and $0 \leq p(x_i) \leq 1$. Lemma 2 shows this inequality with the equality condition.

The second inequality is proven by the use of the nonnegativity of the Tsallis hypodivergence in the following way:

$$0 \leq D_{\lambda,q}(X||U) = -\frac{1}{\lambda} \sum_{i=1}^n (1 + \lambda p(x_i)) \ln_q \frac{1 + \frac{\lambda}{n}}{1 + \lambda p(x_i)} \tag{53}$$

which implies (by the use of the formula $\ln_q \frac{b}{a} = b^{1-q} \ln_q \frac{1}{a} + \ln_q b$):

$$\frac{1}{\lambda} \sum_{i=1}^n (1 + \lambda p(x_i)) \ln_q \frac{1}{1 + \lambda p(x_i)} \leq \frac{n + \lambda}{\lambda} \ln_q \frac{n}{n + \lambda}. \tag{54}$$

The equality condition of the second inequality follows from the equality condition of the nonnegativity of the Tsallis hypodivergence (41). \square

We may call:

$$H_{\lambda,q}(U) = \frac{h(\lambda, q)}{\lambda} \left\{ -(1 + \lambda) \ln_q \frac{1}{1 + \lambda} + (n + \lambda) \ln_q \frac{1}{1 + \frac{\lambda}{n}} \right\}$$

the Hartley–Tsallis hypoentropy. We study the monotonicity for n or λ of the Hartley–Tsallis hypoentropy $H_{\lambda,q}(U)$ and the Tsallis hypoentropy $H_{\lambda,q}(X)$. (Throughout the present paper, the term “monotonicity” means the monotone increasing/decreasing as a function of the parameter λ . We emphasize that it does not mean the monotonicity for some stochastic maps.)

Lemma 3. The function:

$$f(x) = (x + 1) \ln_q \frac{x}{x + 1} \quad (x > 0)$$

is monotonically increasing in x , for any $q \geq 0$.

Proof. By direct calculations, we have:

$$\frac{df(x)}{dx} = \frac{1}{1 - q} \left\{ \left(1 + \frac{1}{x}\right)^{q-1} \left(1 + \frac{1 - q}{x}\right) - 1 \right\}$$

and:

$$\frac{d^2 f(x)}{dx^2} = -qx^{-3} \left(1 + \frac{1}{x}\right)^{q-2} \leq 0.$$

Since $\lim_{x \rightarrow \infty} \frac{df(x)}{dx} = 0$, we have $\frac{df(x)}{dx} \geq 0$. \square

Proposition 4. *The Hartley–Tsallis hypoentropy:*

$$H_{\lambda,q}(U) = \frac{h(\lambda, q)}{\lambda} \left\{ -(1 + \lambda) \ln_q \frac{1}{1 + \lambda} + (n + \lambda) \ln_q \frac{1}{1 + \frac{\lambda}{n}} \right\}$$

is a monotonically increasing function of n , for any $\lambda > 0$ and $q \geq 0$.

Proof. Note that:

$$H_{\lambda,q}(U) = h(\lambda, q) \left\{ -\left(1 + \frac{1}{\lambda}\right) \ln_q \frac{1}{1 + \lambda} + \left(1 + \frac{n}{\lambda}\right) \ln_q \frac{1}{1 + \frac{\lambda}{n}} \right\}.$$

Putting $x = \frac{n}{\lambda} > 0$ for $\lambda > 0$ fixed in Lemma 3, we get the function:

$$g(n) = \left(1 + \frac{n}{\lambda}\right) \ln_q \frac{1}{1 + \frac{\lambda}{n}},$$

which is a monotonically increasing function of n . Thus, we have the present proposition. \square

Remark 4. *We have the relation:*

$$\lim_{n \rightarrow \infty} H_{\lambda,q}(U) = h(\lambda, q) \left\{ -\left(1 + \frac{1}{\lambda}\right) \ln_q \frac{1}{1 + \lambda} - 1 \right\}.$$

We notice from the condition (44) that:

$$\begin{aligned} \lim_{\lambda \rightarrow \infty} \left(\lim_{n \rightarrow \infty} H_{\lambda,q}(U) \right) &= \lim_{\lambda \rightarrow \infty} \frac{h(\lambda, q)}{\lambda^{1-q}} \cdot \lambda^{1-q} \left\{ -1 - \left(1 + \frac{1}{\lambda}\right) \ln_q \frac{1}{1 + \lambda} \right\} \\ &= \frac{1}{1-q} \lim_{\lambda \rightarrow \infty} \frac{1+q\lambda-(1+\lambda)^q}{\lambda^q} = \begin{cases} 0 & (q = 0) \\ \infty & (0 < q < 1) \\ \frac{1}{q-1} & (q > 1), \end{cases} \end{aligned}$$

and conclude that the result is independent of the choice of $h(\lambda, q)$.

For the limit $\lambda \rightarrow 0$, we consider two cases.

(1) In the case of $h(\lambda, q) = \lambda^{1-q}$, we have:

$$\begin{aligned} \lim_{\lambda \rightarrow 0} \left(\lim_{n \rightarrow \infty} H_{\lambda,q}(U) \right) &= \lim_{\lambda \rightarrow 0} \lambda^{1-q} \left\{ -1 - \left(1 + \frac{1}{\lambda}\right) \ln_q \frac{1}{1 + \lambda} \right\} \\ &= \frac{1}{1-q} \lim_{\lambda \rightarrow 0} \frac{1+q\lambda-(1+\lambda)^q}{\lambda^q} = \begin{cases} \infty & (q > 2) \\ 1 & (q = 2) \\ 0 & (0 \leq q < 2), \end{cases} \end{aligned}$$

as one obtains using l'Hôpital's rule.

(2) In the case of $h(\lambda, q) = (1 + \lambda)^{1-q}$, we have for all $q \geq 0$:

$$\begin{aligned} \lim_{\lambda \rightarrow 0} \left(\lim_{n \rightarrow \infty} H_{\lambda, q}(U) \right) &= \lim_{\lambda \rightarrow 0} (1 + \lambda)^{1-q} \left\{ -1 - \left(1 + \frac{1}{\lambda}\right) \ln_q \frac{1}{1+\lambda} \right\} \\ &= \frac{1}{1-q} \lim_{\lambda \rightarrow 0} \frac{1+q\lambda-(1+\lambda)^q}{\lambda(1+\lambda)^{q-1}} = \frac{q}{1-q} \lim_{\lambda \rightarrow 0} \frac{1-(1+\lambda)^{q-1}}{(1+\lambda)^{q-1}+(q-1)\lambda(1+\lambda)^{q-2}} = 0. \end{aligned}$$

These results mean that our Hartley–Tsallis hypoentropy with $h(\lambda, q) = \lambda^{1-q}$ or $(1 + \lambda)^{1-q}$ has the same limits as the Hartley hypoentropy, $F_\lambda(U)$ (see also [9]), in the case $0 < q < 1$.

We study here the monotonicity of $H_{\lambda, q}(X)$ for $h(\lambda, q) = (1 + \lambda)^{1-q}$. The other case $h(\lambda, q) = \lambda^{1-q}$ is studied in the next section; see Lemma 5.

Proposition 5. We assume $h(\lambda, q) = (1 + \lambda)^{1-q}$. Then, $H_{\lambda, q}(X)$ is a monotone increasing function of $\lambda > 0$ when $0 \leq q \leq 2$.

Proof. Note that:

$$H_{\lambda, q}(X) = \sum_{i=1}^n S_{n, \lambda, q}(p(x_i)),$$

where:

$$S_{n, \lambda, q}(x) \equiv \frac{(1 + \lambda)^{1-q}}{\lambda(1 - q)} \left\{ (1 + \lambda x)^q - (1 + \lambda)^q x + x - 1 \right\}$$

is defined on $0 \leq x \leq 1, 0 \leq q \leq 2$ and $\lambda > 0$. Then, we have:

$$\frac{dH_{\lambda, q}(X)}{d\lambda} = \sum_{i=1}^n \frac{dS_{n, \lambda, q}(p(x_i))}{d\lambda} = \sum_{i=1}^n s_{\lambda, q}(p(x_i)),$$

where:

$$s_{\lambda, q}(x) \equiv \frac{q\lambda(1 - x) \left\{ 1 - (1 + \lambda x)^{q-1} \right\} + 1 - x + (1 + \lambda)^q x - (1 + \lambda x)^q}{(1 - q)\lambda^2(1 + \lambda)^q}$$

is defined on $0 \leq x \leq 1, 0 \leq q \leq 2$ and $\lambda > 0$. By some computations, we have:

$$\frac{d^2 s_{\lambda, q}(x)}{dx^2} = \frac{-q(1 + \lambda x)^{q-3} [1 + \lambda \{ (x - 1)(q - 1) + 1 \}]}{(1 + \lambda)^q} \leq 0,$$

since $(x - 1)(q - 1) + 1 \geq 0$ for $0 \leq x \leq 1$ and $0 \leq q \leq 2$. We easily find $s_{\lambda, q}(0) = s_{\lambda, q}(1) = 0$. Thus, we have $s_{\lambda, q}(x) \geq 0$ for $0 \leq x \leq 1, 0 \leq q \leq 2$ and $\lambda > 0$. Therefore, we have $\frac{dH_{\lambda, q}(X)}{d\lambda} \geq 0$ for $0 \leq q \leq 2$ and $\lambda > 0$. \square

This result agrees with the known fact that the usual (Ferreri) hypoentropy is increasing as a function of λ .

Closing this subsection, we give a q -extended version for Proposition 2.

Proposition 6. Let $p_{max} \equiv \max\{p(x_1), \dots, p(x_n)\}$. Then, we have the following inequality.

$$H_{\lambda, q}(X) \leq \frac{h(\lambda, q)}{\lambda} \left\{ (1 + \lambda)^q - (1 + \lambda p_{max})^q \right\} \ln_q(1 + \lambda) \tag{55}$$

for all $\lambda > 0$ and $q \geq 0$.

Proof. From the “ \ln_q -sum” inequality, we have $D_{\lambda,q}(X||Y) \geq 0$. Since $\lambda > 0$, we have:

$$-\sum_{i=1}^n (1 + \lambda p(x_i)) \ln_q \frac{1 + \lambda p(y_i)}{1 + \lambda p(x_i)} \geq 0 \tag{56}$$

which is equivalent to:

$$\sum_{i=1}^n (1 + \lambda p(x_i))^q \{ \ln_q(1 + \lambda p(x_i)) - \ln_q(1 + \lambda p(y_i)) \} \geq 0. \tag{57}$$

Thus, we have:

$$-\sum_{i=1}^n (1 + \lambda p(x_i))^q \ln_q(1 + \lambda p(x_i)) \leq -\sum_{i=1}^n (1 + \lambda p(x_i))^q \ln_q(1 + \lambda p(y_i)), \tag{58}$$

which extends the result given from the inequality (30). For arbitrarily fixed k , we set $p(y_k) = 1$ (and $p(y_i) = 0$ for $i \neq k$) in the above inequality; then, we have:

$$-\sum_{i=1}^n (1 + \lambda p(x_i))^q \ln_q(1 + \lambda p(x_i)) \leq -(1 + \lambda p(x_k))^q \ln_q(1 + \lambda). \tag{59}$$

Since $x^q \ln_q x = -x \ln_q \frac{1}{x}$, we have:

$$\sum_{i=1}^n (1 + \lambda p(x_i)) \ln_q \frac{1}{1 + \lambda p(x_i)} \leq -(1 + \lambda p(x_k))^q \ln_q(1 + \lambda). \tag{60}$$

Multiplying both sides by $\frac{h(\lambda,q)}{\lambda} > 0$ and then adding

$$-\frac{h(\lambda,q)}{\lambda} (1 + \lambda) \ln_q \frac{1}{1 + \lambda} = \frac{h(\lambda,q)}{\lambda} (1 + \lambda)^q \ln_q(1 + \lambda) \tag{61}$$

to both sides, we have:

$$H_{\lambda,q}(X) \leq \frac{h(\lambda,q)}{\lambda} \{ (1 + \lambda)^q - (1 + \lambda p(x_k))^q \} \ln_q(1 + \lambda). \tag{62}$$

Since k is arbitrary, we have this proposition. \square

Letting $q \rightarrow 1$ in the above proposition, we recover Proposition 2.

4. The Subadditivities of the Tsallis Hypoentropies

Throughout this section, we assume $|X| = n, |Y| = m, |Z| = l$. We define the joint Tsallis hypoentropy at the level λ by:

$$H_{\lambda,q}(X, Y) \equiv \frac{h(\lambda,q)}{\lambda} \left\{ -(1 + \lambda) \ln_q \frac{1}{1 + \lambda} + \sum_{i=1}^n \sum_{j=1}^m (1 + \lambda p(x_i, y_j)) \ln_q \frac{1}{1 + \lambda p(x_i, y_j)} \right\}. \tag{63}$$

Note that $H_{\lambda,q}(X, Y) = H_{\lambda,q}(Y, X)$.

For all $i = 1, \dots, n$ for which $p(x_i) \neq 0$, we define the Tsallis hypoentropy of Y given $X = x_i$, at the level $\lambda p(x_i)$, by:

$$\begin{aligned} & H_{\lambda p(x_i),q}(Y|x_i) \\ \equiv & \frac{h(\lambda p(x_i), q)}{\lambda p(x_i)} \left\{ - (1 + \lambda p(x_i)) \ln_q \frac{1}{1 + \lambda p(x_i)} + \sum_{j=1}^m (1 + \lambda p(x_i)p(y_j|x_i)) \ln_q \frac{1}{1 + \lambda p(x_i)p(y_j|x_i)} \right\} \\ = & \frac{h(\lambda p(x_i), q)}{\lambda p(x_i)} \left\{ - (1 + \lambda p(x_i)) \ln_q \frac{1}{1 + \lambda p(x_i)} + \sum_{j=1}^m (1 + \lambda p(x_i, y_j)) \ln_q \frac{1}{1 + \lambda p(x_i, y_j)} \right\}. \end{aligned} \tag{64}$$

For $n = 1$, this coincides with the hypoentropy $H_{\lambda,q}(Y)$. As for the particular case $m = 1$, we get $H_{\lambda p(x_i),q}(Y|x_i) = 0$.

Definition 3. The Tsallis conditional hypoentropy at the level λ is defined by:

$$H_{\lambda,q}(Y|X) \equiv \sum_{i=1}^n p(x_i)^q H_{\lambda p(x_i),q}(Y|x_i). \tag{65}$$

(As a usual convention, the corresponding summand is defined as zero, if $p(x_i) = 0$.)

Throughout this section, we consider the particular function $h(\lambda, q) = \lambda^{1-q}$ for $\lambda > 0, q \geq 0$.

Lemma 4. We assume $h(\lambda, q) = \lambda^{1-q}$. The chain rule for the Tsallis hypoentropy holds:

$$H_{\lambda,q}(X, Y) = H_{\lambda,q}(X) + H_{\lambda,q}(Y|X). \tag{66}$$

Proof. The proof is done by straightforward computation as follows.

$$\begin{aligned} H_{\lambda,q}(X) + H_{\lambda,q}(Y|X) &= \frac{\lambda^{1-q}}{\lambda} \left\{ - (1 + \lambda) \ln_q \frac{1}{1+\lambda} + \sum_{i=1}^n (1 + \lambda p(x_i)) \ln_q \frac{1}{1+\lambda p(x_i)} \right\} \\ &+ \sum_{i=1}^n \frac{(\lambda p(x_i))^{1-q}}{\lambda p(x_i)} p(x_i)^q \left\{ - (1 + \lambda p(x_i)) \ln_q \frac{1}{1+\lambda p(x_i)} + \sum_{j=1}^m (1 + \lambda p(x_i, y_j)) \ln_q \frac{1}{1+\lambda p(x_i, y_j)} \right\} \\ = & \frac{\lambda^{1-q}}{\lambda} \left\{ - (1 + \lambda) \ln_q \frac{1}{1+\lambda} + \sum_{i=1}^n (1 + \lambda p(x_i)) \ln_q \frac{1}{1+\lambda p(x_i)} \right\} \\ &+ \frac{\lambda^{1-q}}{\lambda} \left\{ - \sum_{i=1}^n (1 + \lambda p(x_i)) \ln_q \frac{1}{1+\lambda p(x_i)} + \sum_{i=1}^n \sum_{j=1}^m (1 + \lambda p(x_i, y_j)) \ln_q \frac{1}{1+\lambda p(x_i, y_j)} \right\} \\ = & \frac{\lambda^{1-q}}{\lambda} \left\{ - (1 + \lambda) \ln_q \frac{1}{1+\lambda} + \sum_{i=1}^n \sum_{j=1}^m (1 + \lambda p(x_i, y_j)) \ln_q \frac{1}{1+\lambda p(x_i, y_j)} \right\} = H_{\lambda,q}(X, Y). \end{aligned}$$

□

In the limit $\lambda \rightarrow \infty$, the identity (66) becomes $T_q(X, Y) = T_q(X) + T_q(Y|X)$, where $T_q(Y|X) \equiv \sum_{i=1}^n p(x_i)^q T_q(Y|x_i) = - \sum_{i=1}^n \sum_{j=1}^m p(x_i, y_j)^q \ln_q p(y_j|x_i)$ is the Tsallis conditional entropy and $T_q(X, Y) \equiv \sum_{i=1}^n \sum_{j=1}^m p(x_i, y_j) \ln_q \frac{1}{p(x_i, y_j)}$ is the Tsallis joint entropy (see also p. 3 in [17]).

In the limit $q \rightarrow 1$ in Lemma 4, we also obtain the identity $F_\lambda(X, Y) = F_\lambda(X) + F_\lambda(Y|X)$, which naturally leads to the definition of $F_\lambda(Y|X)$ as conditional hypoentropy.

In order to obtain the subadditivity for the Tsallis hypoentropy, we prove the monotonicity in λ of the Tsallis hypoentropy.

Lemma 5. We assume $h(\lambda, q) = \lambda^{1-q}$. The Tsallis hypoentropy $H_{\lambda,q}(X)$ is a monotonically increasing function of $\lambda > 0$ when $0 \leq q \leq 2$ and a monotonically decreasing function of $\lambda > 0$ when $q \geq 2$ (or $q \leq 0$).

Proof. Note that:

$$H_{\lambda,q}(X) = \sum_{i=1}^n Ln_{\lambda,q}(p(x_i)), \tag{67}$$

where:

$$Ln_{\lambda,q}(x) \equiv \frac{(1 + \lambda x)^q - (1 + \lambda)^q x + x - 1}{\lambda^q(1 - q)} \tag{68}$$

is defined on $0 \leq x \leq 1$ and $\lambda > 0$. Then, we have:

$$\frac{dH_{\lambda,q}(X)}{d\lambda} = \sum_{i=1}^n \frac{dLn_{\lambda,q}(p(x_i))}{d\lambda} = \sum_{i=1}^n l_{\lambda,q}(p(x_i)), \tag{69}$$

where:

$$l_{\lambda,q}(x) \equiv \frac{q}{\lambda^2(1 - q)} \left\{ \left(\frac{1}{\lambda} + 1 \right)^{q-1} x - \left(\frac{1}{\lambda} + x \right)^{q-1} - \frac{(x - 1)}{\lambda^{q-1}} \right\} \tag{70}$$

is defined on $0 \leq x \leq 1$ and $\lambda > 0$. By elementary computations, we obtain:

$$\frac{d^2 l_{\lambda,q}(x)}{dx^2} = q(q - 2)\lambda^{1-q}(1 + \lambda x)^{q-3}. \tag{71}$$

Since we have $l_{\lambda,q}(0) = l_{\lambda,q}(1) = 0$, we find that $l_{\lambda,q}(x) \geq 0$ for $0 \leq q \leq 2$ and any $\lambda > 0$. We also find that $l_{\lambda,q}(x) \leq 0$ for $q \geq 2$ (or $q \leq 0$) and any $\lambda > 0$. Therefore, we have $\frac{dH_{\lambda,q}(X)}{d\lambda} \geq 0$ when $0 \leq q \leq 2$, and $\frac{dH_{\lambda,q}(X)}{d\lambda} \leq 0$ when $q \geq 2$ (or $q \leq 0$). □

This result also agrees with the known fact that the usual (Ferrerri) hypoentropy is increasing as a function of λ .

Theorem 1. We assume $h(\lambda, q) = \lambda^{1-q}$. It holds $H_{\lambda,q}(Y|X) \leq H_{\lambda,q}(Y)$ for $1 \leq q \leq 2$.

Proof. We prove this theorem by the method used in [18] with Jensen’s inequality. We note that $Ln_{\lambda,q}(x)$ is a nonnegative and concave function in x , when $0 \leq x \leq 1$, $\lambda > 0$ and $q \geq 0$. Here, we use the notation for the conditional probability as $p(y_j|x_i) = \frac{p(x_i,y_j)}{p(x_i)}$ when $p(x_i) \neq 0$. By the concavity of $Ln_{\lambda,q}(x)$, we have:

$$\sum_{i=1}^n p(x_i) Ln_{\lambda,q}(p(y_j|x_i)) \leq Ln_{\lambda,q} \left(\sum_{i=1}^n p(x_i)p(y_j|x_i) \right) \tag{72}$$

$$= Ln_{\lambda,q} \left(\sum_{i=1}^n p(x_i, y_j) \right) = Ln_{\lambda,q}(p(y_j)). \tag{73}$$

Summing both sides of the above inequality over j , we have:

$$\sum_{i=1}^n p(x_i) \sum_{j=1}^m Ln_{\lambda,q}(p(y_j|x_i)) \leq \sum_{j=1}^m Ln_{\lambda,q}(p(y_j)). \tag{74}$$

Since $p(x_i)^q \leq p(x_i)$ for $1 \leq q \leq 2$ and $Ln_{\lambda,q}(x) \geq 0$ for $0 \leq x \leq 1$, $\lambda > 0$ and $q \geq 0$, we have:

$$p(x_i)^q \sum_{j=1}^m Ln_{\lambda,q}(p(y_j|x_i)) \leq p(x_i) \sum_{j=1}^m Ln_{\lambda,q}(p(y_j|x_i)). \tag{75}$$

Summing both sides of the above inequality over i , we have:

$$\sum_{i=1}^n p(x_i)^q \sum_{j=1}^m Ln_{\lambda,q}(p(y_j|x_i)) \leq \sum_{i=1}^n p(x_i) \sum_{j=1}^m Ln_{\lambda,q}(p(y_j|x_i)). \tag{76}$$

By the two inequalities (74) and (76), we have:

$$\sum_{i=1}^n p(x_i)^q \sum_{j=1}^m Ln_{\lambda,q}(p(y_j|x_i)) \leq \sum_{j=1}^m Ln_{\lambda,q}(p(y_j)). \tag{77}$$

Here, we can see that $\sum_{j=1}^m Ln_{\lambda,q}(p(y_j|x_i))$ is the Tsallis hypoentropy for fixed x_i and the Tsallis hypoentropy is a monotonically increasing function of λ in the case $1 \leq q \leq 2$, due to Lemma 5. Thus, we have:

$$\sum_{j=1}^m Ln_{\lambda p(x_i),q}(p(y_j|x_i)) \leq \sum_{j=1}^m Ln_{\lambda,q}(p(y_j|x_i)). \tag{78}$$

By the two inequalities (77) and (78), we finally have:

$$\sum_{i=1}^n p(x_i)^q \sum_{j=1}^m Ln_{\lambda p(x_i),q}(p(y_j|x_i)) \leq \sum_{j=1}^m Ln_{\lambda,q}(p(y_j)), \tag{79}$$

which implies (since $p(y_j|x_i) = \frac{p(x_i,y_j)}{p(x_i)}$):

$$\sum_{i=1}^n p(x_i)^q H_{\lambda p(x_i),q}(Y|x_i) \leq \sum_{j=1}^m Ln_{\lambda,q}(p(y_j)), \tag{80}$$

since we have for all fixed x_i ,

$$H_{\lambda p(x_i),q}(Y|x_i) = \frac{1}{\lambda^q p(x_i)^q} \sum_{j=1}^m \left\{ -p(y_j|x_i) (1 + \lambda p(x_i)) \ln_q \frac{1}{1 + \lambda p(x_i)} \right. \\ \left. + (1 + \lambda p(x_i)p(y_j|x_i)) \ln_q \frac{1}{1 + \lambda p(x_i)p(y_j|x_i)} \right\} = \sum_{j=1}^m Ln_{\lambda p(x_i),q}(p(y_j|x_i)).$$

Therefore, we have $H_{\lambda,q}(Y|X) \leq H_{\lambda,q}(Y)$. \square

Corollary 1. We have the following subadditivity for the Tsallis hypoentropies:

$$H_{\lambda,q}(X, Y) \leq H_{\lambda,q}(X) + H_{\lambda,q}(Y) \tag{81}$$

in the case $h(\lambda, q) = \lambda^{1-q}$ for $1 \leq q \leq 2$.

Proof. The proof is easily done by Lemma 4 and Theorem 1. \square

We are now in a position to prove the strong subadditivity for the Tsallis hypoentropies. The strong subadditivity for entropy is one of interesting subjects in entropy theory [19]. For this purpose, we firstly give a chain rule for three random variables X, Y and Z .

Lemma 6. We assume $h(\lambda, q) = \lambda^{1-q}$. The following chain rule holds:

$$H_{\lambda,q}(X, Y, Z) = H_{\lambda,q}(X|Y, Z) + H_{\lambda,q}(Y, Z). \tag{82}$$

Proof. The proof can be done following the recipe used in Lemma 4.

$$\begin{aligned} & H_{\lambda,q}(X|Y, Z) + H_{\lambda,q}(Y, Z) \\ &= \sum_{j=1}^m \sum_{k=1}^l p(y_j, z_k)^q \frac{1}{(\lambda p(y_j, z_k))^q} \left\{ - (1 + \lambda p(y_j, z_k)) \ln_q \frac{1}{1 + \lambda p(y_j, z_k)} \right. \\ &\quad \left. + \sum_{i=1}^n \left(1 + \lambda p(y_j, z_k) \frac{p(x_i, y_j, z_k)}{p(y_j, z_k)} \right) \ln_q \frac{1}{1 + \lambda p(y_j, z_k) \frac{p(x_i, y_j, z_k)}{p(y_j, z_k)}} \right\} \\ &\quad + \frac{1}{\lambda^q} \left\{ - (1 + \lambda) \ln_q \frac{1}{1 + \lambda} + \sum_{j=1}^m \sum_{k=1}^l (1 + \lambda p(y_j, z_k)) \ln_q \frac{1}{1 + \lambda p(y_j, z_k)} \right\} \\ &= \frac{1}{\lambda^q} \left\{ - (1 + \lambda) \ln_q \frac{1}{1 + \lambda} + \sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^l (1 + \lambda p(x_i, y_j, z_k)) \ln_q \frac{1}{1 + \lambda p(x_i, y_j, z_k)} \right\} \\ &= H_{\lambda,q}(X, Y, Z). \end{aligned}$$

□

Theorem 2. We assume $h(\lambda, q) = \lambda^{1-q}$. The strong subadditivity for the Tsallis hypoentropies,

$$H_{\lambda,q}(X, Y, Z) + H_{\lambda,q}(Z) \leq H_{\lambda,q}(X, Z) + H_{\lambda,q}(Y, Z), \tag{83}$$

holds for $1 \leq q \leq 2$.

Proof. This theorem is proven in a similar way as Theorem 1. By the concavity of the function $Ln_{\lambda p(z_k),q}(x)$ in x and by using Jensen’s inequality, we have:

$$\sum_{j=1}^m p(y_j|z_k) Ln_{\lambda p(z_k),q}(p(x_i|y_j, z_k)) \leq Ln_{\lambda p(z_k),q} \left(\sum_{j=1}^m p(y_j|z_k) p(x_i|y_j, z_k) \right).$$

Multiplying both sides by $p(z_k)^q$ and summing over i and k , we have:

$$\begin{aligned} & \sum_{j=1}^m \sum_{k=1}^l p(z_k)^q p(y_j|z_k) \sum_{i=1}^n Ln_{\lambda p(z_k),q}(p(x_i|y_j, z_k)) \\ & \leq \sum_{k=1}^l p(z_k)^q \sum_{i=1}^n Ln_{\lambda p(z_k),q}(p(x_i|z_k)), \end{aligned} \tag{84}$$

since $\sum_{j=1}^m p(y_j|z_k) p(x_i|y_j, z_k) = p(x_i|z_k)$. By $p(y_j|z_k)^q \leq p(y_j|z_k)$ for all j, k and $1 \leq q \leq 2$, and by the nonnegativity of the function $Ln_{\lambda p(z_k),q}$, we have:

$$p(y_j|z_k)^q \sum_{i=1}^n Ln_{\lambda p(z_k),q}(p(x_i|y_j, z_k)) \leq p(y_j|z_k) \sum_{i=1}^n Ln_{\lambda p(z_k),q}(p(x_i|y_j, z_k)).$$

Multiplying both sides by $p(z_k)^q$ and summing over j and k in the above inequality, we have:

$$\begin{aligned} & \sum_{j=1}^m \sum_{k=1}^l p(z_k)^q p(y_j|z_k)^q \sum_{i=1}^n Ln_{\lambda p(z_k),q}(p(x_i|y_j, z_k)) \\ & \leq \sum_{j=1}^m \sum_{k=1}^l p(z_k)^q p(y_j|z_k) \sum_{i=1}^n Ln_{\lambda p(z_k),q}(p(x_i|y_j, z_k)). \end{aligned} \tag{85}$$

From the two inequalities (84) and (85), we have:

$$\sum_{j=1}^m \sum_{k=1}^l p(z_k)^q p(y_j|z_k)^q \sum_{i=1}^n Ln_{\lambda p(z_k),q}(p(x_i|y_j, z_k)) \leq \sum_{k=1}^l p(z_k)^q \sum_{i=1}^n Ln_{\lambda p(z_k),q}(p(x_i|z_k)),$$

which implies:

$$\sum_{j=1}^m \sum_{k=1}^l p(y_j, z_k)^q \sum_{i=1}^n Ln_{\lambda p(y_j, z_k),q}(p(x_i|y_j, z_k)) \leq \sum_{k=1}^l p(z_k)^q \sum_{i=1}^n Ln_{\lambda p(z_k),q}(p(x_i|z_k)),$$

since $p(y_j, z_k) \leq p(z_k)$ (because of $\sum_{j=1}^m p(y_j, z_k) = p(z_k)$) for all j and k and the function $Ln_{\lambda p(z_k),q}$ is monotonically increasing in $\lambda p(z_k) > 0$, when $1 \leq q \leq 2$. Thus, we have $H_{\lambda,q}(X|Y, Z) \leq H_{\lambda,q}(X|Z)$, which is equivalent to the inequality:

$$H_{\lambda,q}(X, Y, Z) - H_{\lambda,q}(Y, Z) \leq H_{\lambda,q}(X, Z) - H_{\lambda,q}(Z)$$

by Lemmas 4 and 6. \square

Remark 5. Passing to the limit $\lambda \rightarrow \infty$ in Corollary 1 and Theorem 2, we recover the subadditivity and the strong subadditivity [20] for the Tsallis entropy:

$$T_q(X, Y) \leq T_q(X) + T_q(Y) \quad (q \geq 1)$$

and:

$$T_q(X, Y, Z) + T_q(Z) \leq T_q(X, Z) + T_q(Y, Z) \quad (q \geq 1).$$

Thanks to the subadditivities, we may define the Tsallis mutual hypoentropies for $1 \leq q \leq 2$ and $\lambda > 0$.

Definition 4. Let $1 \leq q \leq 2$ and $\lambda > 0$. The Tsallis mutual hypoentropy is defined by:

$$I_{\lambda,q}(X; Y) \equiv H_{\lambda,q}(X) - H_{\lambda,q}(X|Y)$$

and the Tsallis conditional mutual hypoentropy is defined by:

$$I_{\lambda,q}(X; Y|Z) \equiv H_{\lambda,q}(X|Z) - H_{\lambda,q}(X|Y, Z).$$

From the chain rule given in Lemma 4, we find that the Tsallis mutual hypoentropy is symmetric, that is,

$$\begin{aligned} I_{\lambda,q}(X; Y) & \equiv H_{\lambda,q}(X) - H_{\lambda,q}(X|Y) \\ & = H_{\lambda,q}(X) + H_{\lambda,q}(Y) - H_{\lambda,q}(X, Y) \\ & = H_{\lambda,q}(Y) - H_{\lambda,q}(Y|X) = I_{\lambda,q}(Y; X). \end{aligned} \tag{86}$$

In addition, we have:

$$0 \leq I_{\lambda,q}(X; Y) \leq \min \{H_{\lambda,q}(X), H_{\lambda,q}(Y)\} \tag{87}$$

from the subadditivity given in Theorem 1 and the nonnegativity of the Tsallis conditional hypoentropy. We also find $I_{\lambda,q}(X; Y|Z) \geq 0$ from the strong subadditivity given in Theorem 2.

Moreover, we have the chain rule for the Tsallis mutual hypoentropies in the following.

$$\begin{aligned} I_{\lambda,q}(X; Y|Z) &= H_{\lambda,q}(X|Z) - H_{\lambda,q}(X|Y, Z) \\ &= H_{\lambda,q}(X|Z) - H_{\lambda,q}(X) + H_{\lambda,q}(X) - H_{\lambda,q}(X|Y, Z) \\ &= -I_{\lambda,q}(X; Z) + I_{\lambda,q}(X; Y, Z). \end{aligned} \tag{88}$$

From the strong subadditivity, we have $H_{\lambda,q}(X|Y, Z) \leq H_{\lambda,q}(X|Z)$; thus, we have:

$$I_{\lambda,q}(X; Z) \leq I_{\lambda,q}(X; Y, Z)$$

for $1 \leq q \leq 2$ and $\lambda > 0$.

5. Jeffreys and Jensen–Shannon Hypodivergences

In what follows, we indicate extensions of two known information measures.

Definition 5 ([21,22]). *The Jeffreys divergence is defined by:*

$$J(X||Y) \equiv D(X||Y) + D(Y||X) \tag{89}$$

and the Jensen–Shannon divergence is defined as:

$$JS(X||Y) \equiv \frac{1}{2} \left\{ D \left(X \parallel \frac{X+Y}{2} \right) + D \left(Y \parallel \frac{X+Y}{2} \right) \right\} \tag{90}$$

$$= H \left(\frac{X+Y}{2} \right) - \frac{1}{2} (H(X) + H(Y)). \tag{91}$$

The Jensen–Shannon divergence was introduced in 1991 in [23], but its roots can be older, since one can see some analogous formulae used in thermodynamics under the name entropy of mixing (p. 598 in [24]), for the study of gaseous, liquid or crystalline mixtures.

Jeffreys and Jensen–Shannon divergences have been extended to the context of Tsallis theory in [25]:

Definition 6. *The Jeffreys–Tsallis divergence is:*

$$J_q(X||Y) \equiv S_q(X||Y) + S_q(Y||X) \tag{92}$$

and the Jensen–Shannon–Tsallis divergence is:

$$JS_q(X||Y) \equiv \frac{1}{2} \left\{ S_q \left(X \parallel \frac{X+Y}{2} \right) + S_q \left(Y \parallel \frac{X+Y}{2} \right) \right\}. \tag{93}$$

Note that:

$$JS_q(X||Y) \neq T_q \left(\frac{X + Y}{2} \right) - \frac{1}{2} (T_q(X) + T_q(Y)).$$

This expression was used in [26] as Jensen–Tsallis divergence.

In accordance with the above definition, we define the directed Jeffreys and Jensen–Shannon q -hypodivergence measures between two distributions and emphasize the mathematical significance of our definitions.

Definition 7. *The Jeffreys–Tsallis hypodivergence is:*

$$J_{\lambda,q}(X||Y) \equiv D_{\lambda,q}(X||Y) + D_{\lambda,q}(Y||X) \tag{94}$$

and the Jensen–Shannon–Tsallis hypodivergence is:

$$JS_{\lambda,q}(X||Y) \equiv \frac{1}{2} \left\{ D_{\lambda,q} \left(X || \frac{X + Y}{2} \right) + D_{\lambda,q} \left(Y || \frac{X + Y}{2} \right) \right\}. \tag{95}$$

Here, we point out that again, one has:

$$JS_{\lambda}(X||Y) = \frac{1}{2} K_{\lambda} \left(X || \frac{X + Y}{2} \right) + \frac{1}{2} K_{\lambda} \left(Y || \frac{X + Y}{2} \right) \tag{96}$$

$$= F_{\lambda} \left(\frac{X + Y}{2} \right) - \frac{1}{2} (F_{\lambda}(X) + F_{\lambda}(Y)), \tag{97}$$

where:

$$JS_{\lambda}(X||Y) \equiv \lim_{q \rightarrow 1} JS_{\lambda,q}(X||Y).$$

Lemma 7. *The following inequality holds:*

$$D_{\lambda,q} \left(X || \frac{X + Y}{2} \right) \leq \frac{1}{2} D_{\lambda, \frac{1+q}{2}}(X||Y)$$

for $q \geq 0$ and $\lambda > 0$.

Proof. Using the inequality between the arithmetic and geometric mean, one has:

$$D_{\lambda,q} \left(X || \frac{X + Y}{2} \right) = -\frac{1}{\lambda} \sum_{i=1}^n (1 + \lambda p(x_i)) \ln_q \frac{\frac{(1+\lambda p(x_i))+(1+\lambda p(y_i))}{2}}{1 + \lambda p(x_i)} \tag{98}$$

$$\leq -\frac{1}{\lambda} \sum_{i=1}^n (1 + \lambda p(x_i)) \ln_q \sqrt{\frac{1 + \lambda p(y_i)}{1 + \lambda p(x_i)}} \tag{99}$$

$$= -\frac{1}{2\lambda} \sum_{i=1}^n (1 + \lambda p(x_i)) \frac{\left(\frac{1+\lambda p(y_i)}{1+\lambda p(x_i)} \right)^{1 - \frac{1+q}{2}} - 1}{1 - \frac{1+q}{2}} \tag{100}$$

$$= \frac{1}{2} D_{\lambda, \frac{1+q}{2}}(X||Y). \tag{101}$$

Thus, the proof is completed. \square

In the limit $\lambda \rightarrow \infty$, Lemma 7 recovers Lemma 3.4 in [25].

Lemma 8 ([25]). *The function:*

$$f(x) = -\ln_r \frac{1 + \exp_q x}{2}$$

is concave for $0 \leq r \leq q$.

The next two results of the present paper are stated in order to establish the counterpart of Theorem 3.5 in [25] for hypodivergences.

Proposition 7. *It holds:*

$$JS_{\lambda,q}(X||Y) \leq \frac{1}{4} J_{\lambda, \frac{1+q}{2}}(X||Y) \tag{102}$$

for $q \geq 0$ and $\lambda > 0$.

Proof. By the use of Lemma 7, one has:

$$2JS_{\lambda,q}(X||Y) = D_{\lambda,q} \left(X || \frac{X+Y}{2} \right) + D_{\lambda,q} \left(Y || \frac{X+Y}{2} \right) \tag{103}$$

$$\leq \frac{1}{2} D_{\lambda, \frac{1+q}{2}}(X||Y) + \frac{1}{2} D_{\lambda, \frac{1+q}{2}}(Y||X) \tag{104}$$

$$= \frac{1}{2} J_{\lambda, \frac{1+q}{2}}(X||Y). \tag{105}$$

This completes the proof. \square

Proposition 8. *It holds that:*

$$JS_{\lambda,r}(X||Y) \leq -\frac{n+\lambda}{\lambda} \ln_r \frac{1 + \exp_q \left(-\frac{1}{2} \cdot \frac{\lambda}{n+\lambda} \cdot J_{\lambda,q}(X||Y) \right)}{2} \tag{106}$$

for $0 \leq r \leq q$ and $\lambda > 0$.

Proof. According to Lemma 8,

$$\begin{aligned} JS_{\lambda,r}(X||Y) &= -\frac{n+\lambda}{2\lambda} \left\{ \sum_{i=1}^n \frac{1+\lambda p(x_i)}{n+\lambda} \ln_r \frac{1+\exp_q \ln_q \left(\frac{1+\lambda p(y_i)}{1+\lambda p(x_i)} \right)}{2} + \sum_{i=1}^n \frac{1+\lambda p(y_i)}{n+\lambda} \ln_r \frac{1+\exp_q \ln_q \left(\frac{1+\lambda p(x_i)}{1+\lambda p(y_i)} \right)}{2} \right\} \\ &\leq -\frac{n+\lambda}{2\lambda} \left\{ \ln_r \frac{1+\exp_q \sum_{i=1}^n \frac{1+\lambda p(x_i)}{n+\lambda} \ln_q \left(\frac{1+\lambda p(y_i)}{1+\lambda p(x_i)} \right)}{2} + \ln_r \frac{1+\exp_q \sum_{i=1}^n \frac{1+\lambda p(y_i)}{n+\lambda} \ln_q \left(\frac{1+\lambda p(x_i)}{1+\lambda p(y_i)} \right)}{2} \right\} \\ &= -\frac{n+\lambda}{2\lambda} \left\{ \ln_r \frac{1+\exp_q \left(-\frac{\lambda}{n+\lambda} D_{\lambda,q}(X||Y) \right)}{2} + \ln_r \frac{1+\exp_q \left(-\frac{\lambda}{n+\lambda} D_{\lambda,q}(Y||X) \right)}{2} \right\}. \end{aligned} \tag{107}$$

Then:

$$\begin{aligned} JS_{\lambda,r}(X||Y) &\leq -\frac{n+\lambda}{\lambda} \ln_r \frac{1 + \exp_q -\frac{\lambda}{n+\lambda} \left(\frac{D_{\lambda,q}(X||Y)+D_{\lambda,q}(Y||X)}{2} \right)}{2} \\ &= -\frac{n+\lambda}{\lambda} \ln_r \frac{1 + \exp_q \left(-\frac{1}{2} \cdot \frac{\lambda}{n+\lambda} \cdot J_{\lambda,q}(X||Y) \right)}{2}. \end{aligned} \tag{108}$$

Thus, the proof is completed. \square

We further define the dual symmetric hypodivergences.

Definition 8. The dual symmetric Jeffreys–Tsallis hypodivergence is defined by:

$$J_{\lambda,q}^{(ds)}(X||Y) \equiv D_{\lambda,q}(X||Y) + D_{\lambda,2-q}(Y||X)$$

and the dual symmetric Jensen–Shannon–Tsallis hypodivergence is defined by:

$$JS_{\lambda,q}^{(ds)}(X||Y) \equiv \frac{1}{2} \left\{ D_{\lambda,q} \left(X \left\| \frac{X+Y}{2} \right. \right) + D_{\lambda,2-q} \left(Y \left\| \frac{X+Y}{2} \right. \right) \right\}.$$

Using Lemma 7, we have the following inequality.

Proposition 9. It holds:

$$JS_{\lambda,q}^{(ds)}(X||Y) \leq \frac{1}{4} J_{\lambda, \frac{1+q}{2}}^{(ds)}(X||Y)$$

for $0 \leq q \leq 2$ and $\lambda > 0$.

In addition, we have the following inequality.

Proposition 10. It holds:

$$JS_{\lambda,q}^{(ds)}(X||Y) \leq -\frac{n+\lambda}{\lambda} \ln_r \frac{1 + \exp_q \left(-\frac{\lambda}{2(n+\lambda)} J_{\lambda,q}(X||Y) \right)}{2}$$

for $1 < r \leq 2$, $r \leq q$ and $\lambda > 0$.

Proof. The proof can be done by similar calculations with Proposition 8, applying the facts (see Lemmas 3.9 and 3.10 in [25]) that $\exp_q(x)$ is a monotonically increasing function in q for $x \geq 0$ and the inequality $-\ln_{2-r} x \leq -\ln_r x$ holds for $1 < r \leq 2$ and $x > 0$. □

6. Concluding Remarks

In this paper, we introduced the Tsallis hypoentropy $H_{\lambda,q}(X)$ and studied some properties of $H_{\lambda,q}(X)$. We named $H_{\lambda,q}(X)$ Tsallis hypoentropy because of the relation $H_{\lambda,q}(X) \leq T_q(X)$, which follows from the monotonicity in λ given in Proposition 5 and Lemma 5 for the case $h(\lambda, q) = (1 + \lambda)^{1-q}$ and the case $h(\lambda, q) = \lambda^{1-q}$, respectively (this relation can be also proven directly). In this naming, we follow Ferreri, as he has termed $F_\lambda(X)$ hypoentropy due to the relation $F_\lambda(X) \leq H(X)$.

The monotonicity of the hypoentropy and the Tsallis hypoentropy for $\lambda > 0$, indeed, is an interesting feature. It may be remarkable to examine the monotonicity of the Tsallis entropy for the parameter $q \geq 0$. We find that the Tsallis entropy $T_q(X)$ is monotonically decreasing with respect to $q \geq 0$. Indeed, we find $\frac{dT_q(X)}{dq} = \sum_{j=1}^n \frac{p_j^q v_q(p_j)}{(1-q)^2}$, where $v_q(x) \equiv 1 - x^{1-q} + (1-q) \log x$ ($0 \leq x \leq 1$). Since $x^q v_q(x) = 0$ for $x = 0$ and $q > 0$, we prove $v_q(x) \leq 0$ for $0 < x \leq 1$. We find $\frac{dv_q(x)}{dx} = \frac{(1-q)(1-x^{1-q})}{x} \geq 0$ when $0 < x \leq 1$; thus, we have $v_q(x) \leq v_q(1) = 0$, which implies $\frac{dT_q(X)}{dq} \leq 0$. This monotonicity implies the relations $H(X) \leq T_q(X)$ for $0 \leq q < 1$ and $T_q(X) \leq H(X)$ for $q > 1$ (these relations also follow from the inequalities $\log \frac{1}{x} \leq \ln_q \frac{1}{x}$ for $0 \leq q < 1$, $x > 0$ and $\log \frac{1}{x} \geq \ln_q \frac{1}{x}$ for $q > 1$, $x > 0$).

As other important results, we also gave the chain rules, subadditivity and the strong subadditivity of the Tsallis hypoentropies in the case of $h(\lambda, q) = \lambda^{1-q}$. For the case of $h(\lambda, q) = (1+\lambda)^{1-q}$, we can prove

$H_{\lambda,q}(Y|X) \leq H_{\lambda,q}(X)$ and $H_{\lambda,q}(X|Y, Z) \leq H_{\lambda,q}(X|Z)$ for $1 \leq q \leq 2$ in a similar way to the proofs of Theorems 1 and 2, since the function $Sn_{\lambda,q}(x)$ defined in the proof of Proposition 5 is also nonnegative, monotone increasing and concave in $x \in [0, 1]$, and we have $H_{\lambda p(x_i),q}(Y|x_i) = \sum_{j=1}^m Sn_{\lambda p(x_i),q}(p(y_j|x_i))$ for all fixed x_i . However, we cannot obtain the inequalities:

$$H_{\lambda,q}(X, Y) \leq H_{\lambda,q}(X) + H_{\lambda,q}(Y) \quad (1 \leq q \leq 2),$$

$$H_{\lambda,q}(X, Y, Z) + H_{\lambda,q}(Z) \leq H_{\lambda,q}(X, Z) + H_{\lambda,q}(Y, Z) \quad (1 \leq q \leq 2)$$

for $h(\lambda, q) = (1 + \lambda)^{1-q}$, because the similar proof for the chain rules does not work well in this case.

Acknowledgments

The first author was partially supported by JSPS KAKENHI Grant Number 24540146.

Author Contributions

The work presented here was carried out in collaboration between all authors. The study was initiated by the second author. The first author played also the role of the corresponding author. All authors contributed equally and significantly in writing this article. All authors have read and approved the final manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Boltzmann, L.E. Einige allgemeine Sätze über Wärmegleichgewicht. *Wiener Berichte* **1871**, *63*, 679–711.
2. Gibbs, J.W. *Elementary Principles in Statistical Mechanics—Developed with Especial Reference to the Rational Foundation of Thermodynamics*; Charles Scribner's Sons: New York, NY, USA, 1902.
3. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423, 623–656.
4. Hartley, R.V.L. Transmission of Information. *Bell Syst. Tech. J.* **1928**, *7*, 535–563.
5. Kullback, S.; Leibler, R.A. On the information and sufficiency. *Ann. Math. Stat.* **1951**, *17*, 79–86.
6. Rényi, A. On measures of information and entropy. In Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, CA, USA, 20 June–30 July 1960; University of California Press: Berkeley, CA, USA, 1961; pp. 547–561.
7. Havrda, J.; Charvat, F. Quantification methods of classification processes: Concept of structural alpha-entropy. *Kybernetika* **1967**, *3*, 30–35.
8. Aczél, J.; Daróczy, Z. *On Measures of Information and Their Characterizations*; Academic Press: New York, NY, USA, 1975.
9. Ferreri, C. Hypoentropy and related heterogeneity, divergence and information measures. *Statistica* **1980**, *2*, 155–167.

10. Tsallis, C. Possible generalization of Boltzmann–Gibbs statistics. *J. Stat. Phys.* **1988**, *52*, 479–487.
11. Tsallis, C. Generalized entropy-based criterion for consistent testing. *Phys. Rev. E* **1998**, *58*, 1442–1445.
12. Borland, L.; Plastino, A.R.; Tsallis, C. Information gain within nonextensive thermostatics. *J. Math. Phys.* **1998**, *39*, 6490–6501.
13. Gilardoni, G. On Pinsker’s and Vajda’s type inequalities for Csiszár’s f-divergence. *IEEE Trans. Inf. Theory* **2010**, *56*, 5377–5386.
14. Rastegin, A.E. Bounds of the Pinsker and Fannes types on the Tsallis relative entropy. *Math. Phys. Anal. Geom.* **2013**, *16*, 213–228.
15. Csiszár, I. Information-type measures of difference of probability distributions and indirect observations. *Stud. Sci. Math. Hung.* **1967**, *2*, 299–318.
16. Csiszár, I. Axiomatic characterizations of information measures. *Entropy* **2008**, *10*, 261–273.
17. Furuichi, S. On uniqueness theorems for Tsallis entropy and Tsallis relative entropy. *IEEE Trans. Inf. Theory* **2005**, *47*, 3638–3645.
18. Daroczy, Z. Generalized information functions. *Inf. Control* **1970**, *16*, 36–51.
19. Petz, D.; Virosztek, D. Some inequalities for quantum Tsallis entropy related to the strong subadditivity. *Math. Inequal. Appl.* **2014**, in press.
20. Furuichi, S. Information theoretical properties of Tsallis entropies. *J. Math. Phys.* **2006**, *47*, doi:10.1063/1.2165744.
21. Dragomir, S.S.; Šunde, J.; Buşe, C. New inequalities for Jeffreys divergence measure. *Tamsui Oxf. J. Math. Sci.* **2000**, *16*, 295–309.
22. Jeffreys, H. An invariant form for the prior probability in estimation problems. *Proc. R. Soc. Lond. A* **1946**, *186*, 453–461.
23. Lin, J. Divergence measures based on the Shannon entropy. *IEEE Trans. Inf. Theory* **1991**, *37*, 145–151.
24. Tolman, R.C. *The Principles of Statistical Mechanics*; Clarendon Press: London, UK, 1938.
25. Furuichi, S.; Mitroi, F.-C. Mathematical inequalities for some divergences. *Physica A* **2012**, *391*, 388–400.
26. Hamza, A.B. A nonextensive information-theoretic measure for image edge detection. *J. Electron. Imaging* **2006**, *15*, 13011.1–13011.8.