


Article

Refined Young Inequality and Its Application to Divergences

Shigeru Furuichi ¹  and Nicușor Minculete ^{2,*}

¹ Department of Information Science, College of Humanities and Sciences, Nihon University, 3-25-40, Sakurajyousui, Setagaya-ku, Tokyo 156-8550, Japan; furuichi@chs.nihon-u.ac.jp

² Faculty of Mathematics and Computer Science, Transilvania University of Brașov, 500091 Brașov, Romania

* Correspondence: minculete.nicusor@unitbv.ro

Abstract: We give bounds on the difference between the weighted arithmetic mean and the weighted geometric mean. These imply refined Young inequalities and the reverses of the Young inequality. We also studied some properties on the difference between the weighted arithmetic mean and the weighted geometric mean. Applying the newly obtained inequalities, we show some results on the Tsallis divergence, the Rényi divergence, the Jeffreys–Tsallis divergence and the Jensen–Shannon–Tsallis divergence.

Keywords: Young inequality; arithmetic mean; geometric mean; Heinz mean; Cartwright–Field inequality; Tsallis divergence; Rényi divergence; Jeffreys–Tsallis divergence; Jensen–Shannon–Tsallis divergence

MSC: 26D15; 26E60; 94A17



Citation: Furuichi, S.; Minculete, N. Refined Young Inequality and Its Application to Divergences. *Entropy* **2021**, *23*, 514. <https://doi.org/10.3390/e23050514>

Academic Editor: Antonio M. Scarfone

Received: 29 March 2021
Accepted: 21 April 2021
Published: 23 April 2021

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



Copyright: © 2021 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

The Young integral inequality is the source of many basic inequalities. Young [1] proved the following: suppose that $f : [0, \infty) \rightarrow [0, \infty)$ is an increasing continuous function such that $f(0) = 0$ and $\lim_{x \rightarrow \infty} f(x) = \infty$. Then:

$$ab \leq \int_0^a f(x)dx + \int_0^b f^{-1}(x)dx, \quad (1)$$

with equality if $b = f(a)$. Such a gap is often used to define the Fenchel–Legendre divergence in information geometry [2,3]. For $f(x) = x^{p-1}$, ($p > 1$), in inequality (1), we deduce the classical Young inequality:

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}, \quad (2)$$

for all $a, b > 0$ and $p, q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$. The equality occurs if and only if $a^p = b^q$.

Minguzzi [4] proved a reverse Young inequality in the following way:

$$0 \leq \frac{a^p}{p} + \frac{b^q}{q} - ab \leq (b - a^{p-1})(b^{q-1} - a), \quad (3)$$

for all $a, b > 0$ and $p, q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$.

The classical Young inequality (2) is rewritten as

$$a^{1/p}b^{1/q} \leq \frac{a}{p} + \frac{b}{q} \quad (4)$$

by putting $a := a^{1/p}$ and $b := b^{1/q}$. Putting again:

$$a := \frac{a_j^p}{\sum_{j=1}^n a_j^p}, \quad b := \frac{b_j^q}{\sum_{j=1}^n b_j^q}$$

in the inequality (4), we obtain the famous Hölder inequality:

$$\sum_{j=1}^n a_j b_j \leq \left(\sum_{j=1}^n a_j^p \right)^{1/p} \left(\sum_{j=1}^n b_j^q \right)^{1/q}, \quad \left(p, q > 1, \frac{1}{p} + \frac{1}{q} = 1 \right)$$

for $a_1, \dots, a_n > 0$ and $b_1, \dots, b_n > 0$. Thus, the inequality (2) is often reformulated as

$$a^p b^{1-p} \leq pa + (1-p)b, \quad a, b > 0, \quad 0 \leq p \leq 1 \tag{5}$$

by putting $1/p =: p$ (then $1/q = 1 - p$) in the inequality (4). It is notable that α -divergence is related to the difference between the weighted arithmetic mean and the weighted geometric mean [5]. For $p = 1/2$, we deduce the inequality between the geometric mean and the arithmetic mean, $G(a, b) := \sqrt{ab} \leq \frac{a+b}{2} =: A(a, b)$. The Heinz mean ([6], Equation (3)) (see also [7]) is defined as $H_p(a, b) = \frac{a^p b^{1-p} + a^{1-p} b^p}{2}$ and $G(a, b) \leq H_p(a, b) \leq A(a, b)$.

Especially, when we discuss Young inequality, we will refer to the last form. We consider the following expression:

$$d_p(a, b) := pa + (1-p)b - a^p b^{1-p} \tag{6}$$

which implies that $d_p(a, b) \geq 0$ and $d_p(a, a) = d_0(a, b) = d_1(a, b) = 0$. We remark the following properties:

$$d_p(a, b) = b \cdot d_p\left(\frac{a}{b}, 1\right), \quad d_p(a, b) = d_{1-p}(b, a), \quad d_p\left(\frac{1}{a}, \frac{1}{b}\right) = \frac{1}{ab} \cdot d_p(b, a).$$

Cartwright–Field inequality (see, e.g., [8]) is often written as follows:

$$\frac{1}{2}p(1-p) \frac{(a-b)^2}{\max\{a, b\}} \leq d_p(a, b) \leq \frac{1}{2}p(1-p) \frac{(a-b)^2}{\min\{a, b\}} \tag{7}$$

for $a, b > 0$ and $0 \leq p \leq 1$. This double inequality gives an improvement of the Young inequality, and at the same time, gives a reverse inequality for the Young inequality.

Kober proved in [9] a general result related to an improvement of the inequality between arithmetic and geometric means, which for $n = 2$ implies the inequality:

$$r(\sqrt{a} - \sqrt{b})^2 \leq d_p(a, b) \leq (1-r)(\sqrt{a} - \sqrt{b})^2 \tag{8}$$

where $a, b > 0, 0 \leq p \leq 1$ and $r = \min\{p, 1-p\}$. This inequality was rediscovered by Kittaneh and Manasrah in [10] (See also [11]).

Finally, we found, in [12], another improvement of the Young inequality and a reverse inequality, given as

$$r(\sqrt{a} - \sqrt{b})^2 + A(p)\log^2\left(\frac{a}{b}\right) \leq d_p(a, b) \leq (1-r)(\sqrt{a} - \sqrt{b})^2 + B(p)\log^2\left(\frac{a}{b}\right) \tag{9}$$

where $a, b \geq 1, 0 < p < 1$ and $r = \min\{p, 1-p\}$ with $A(p) = \frac{p(1-p)}{2} - \frac{r}{4}, B(p) = \frac{p(1-p)}{2} - \frac{1-r}{4}$. It is remarkable that the inequalities (9) give a further refinement of (8), since $A(p) \geq 0$ and $B(p) \leq 0$.

In [13], we also presented two inequalities which give two different reverse inequalities for the Young inequality:

$$0 \leq d_p(a, b) \leq a^p b^{1-p} \exp \left\{ \frac{p(1-p)(a-b)^2}{\min^2\{a, b\}} \right\} - a^p b^{1-p} \tag{10}$$

and:

$$0 \leq d_p(a, b) \leq p(1-p) \log^2 \left(\frac{a}{b} \right) \max\{a, b\} \tag{11}$$

where $a, b > 0, 0 \leq p \leq 1$. See ([14], Chapter 2) for recent advances on refinements and reverses of the Young inequality.

The α -divergence is related to the difference of a weighted arithmetic mean with a geometric mean [5]. We mention that the gap is used in information geometry to define the Fenchel–Legendre divergence [2,3]. We give bounds on the difference between the weighted arithmetic mean and the weighted geometric mean. These imply refined Young inequalities and the reverses of the Young inequality. We also studied some properties on the difference between the weighted arithmetic mean and the weighted geometric mean. Applying the newly obtained inequalities, we show some results on the Tsallis divergence, the Rényi divergence, the Jeffreys–Tsallis divergence and the Jensen–Shannon–Tsallis divergence [15,16]. The parametric Jensen–Shannon divergence can be used to detect unusual data, and this one can also use it as a means to perform the relevant analysis of fire experiments [17].

2. Main Results

We give estimates on $d_p(a, b)$ and also study the properties of $d_p(a, b)$. We give the following estimates of $d_p(a, b)$ first.

Theorem 1. For $0 < a, b \leq 1$ and $0 \leq p \leq 1$, we have:

$$r(\sqrt{a} - \sqrt{b})^2 + A(p)ab \cdot \log^2 \left(\frac{a}{b} \right) \leq d_p(a, b) \leq (1-r)(\sqrt{a} - \sqrt{b})^2 + B(p)ab \cdot \log^2 \left(\frac{a}{b} \right) \tag{12}$$

where $r = \min\{p, 1-p\}$ and $A(p) = \frac{p(1-p)}{2} - \frac{r}{4}, B(p) = \frac{p(1-p)}{2} - \frac{1-r}{4}$.

Proof. For $p = 0$ or $p = 1$ or $a = b$, we have equality. We assume $a \neq b$ and $0 < p < 1$. Because $0 < a, b \leq 1$, we have $\frac{1}{a}, \frac{1}{b} \geq 1$, so, applying inequality (9), we deduce the following relation:

$$r\left(\frac{1}{\sqrt{a}} - \frac{1}{\sqrt{b}}\right)^2 + A(p)\log^2 \left(\frac{b}{a} \right) \leq d_p\left(\frac{1}{a}, \frac{1}{b}\right) \leq (1-r)\left(\frac{1}{\sqrt{a}} - \frac{1}{\sqrt{b}}\right)^2 + B(p)\log^2 \left(\frac{b}{a} \right). \tag{13}$$

We know that $d_p\left(\frac{1}{a}, \frac{1}{b}\right) = \frac{1}{ab} \cdot d_{1-p}(a, b)$ and if we replace p by $1-p$ in relation (13) and because $A(p) = A(1-p), B(1-p) = B(p)$, then we proved the inequality from the statement. \square

Theorem 2. For $a \geq b > 0$ and $0 < p \leq 1$, we have:

$$\frac{p(a-b)(a^{1-p} - b^{1-p})}{2a^{1-p}} \leq d_p(a, b) \leq \frac{p(a-b)(a^{1-p} - b^{1-p})}{a^{1-p}}.$$

Proof. For $p = 1$ or $a = b$, we have equality. We assume $a > b$ and $0 < p < 1$. It is easy to see that:

$$\int_1^x (1 - t^{p-1}) dt = x - 1 - \frac{x^p - 1}{p}. \tag{14}$$

We take $x = a/b$ in (14) and then obtain:

$$pb \int_1^{a/b} (1 - t^{p-1}) dt = d_p(a, b), \quad 0 < p < 1$$

Then, we take the function $f : [1, a/b] \rightarrow \mathbb{R}$ defined by $f(t) := 1 - t^{p-1}$. By simple calculations we have:

$$\frac{df(t)}{dt} = (1 - p)t^{p-2} \geq 0, \quad \frac{d^2f(t)}{dt^2} = (1 - p)(p - 2)t^{p-3} \leq 0.$$

So the function f is concave so that we can apply Hermite–Hadamard inequality [18]:

$$\frac{1}{2}(f(1) + f(a/b)) \leq \frac{1}{a/b - 1} \int_1^{a/b} (1 - t^{p-1}) dt \leq f\left(\frac{1 + a/b}{2}\right).$$

The left-hand side of the inequalities above shows:

$$\frac{p(a - b)(a^{1-p} - b^{1-p})}{2a^{1-p}} \leq d_p(a, b).$$

Since the function $f(t) := 1 - t^{1-p}$ is increasing, we have:

$$1 - t^{p-1} \leq 1 - x^{p-1}, \quad (t \leq x, \quad 0 < p < 1).$$

Integrating the above inequality by t from 1 to x , we obtain:

$$\int_1^x (1 - t^{p-1}) dt \leq (x - 1)(1 - x^{p-1})$$

which implies:

$$d_p(a, b) = bp \int_1^{a/b} (1 - t^{p-1}) dt \leq bp(a/b - 1) \left(1 - (a/b)^{p-1}\right) = \frac{p(a - b)(a^{1-p} - b^{1-p})}{a^{1-p}}.$$

□

Theorem 3. For $a, b > 0$ and $0 \leq p \leq 1$, we have:

$$p(1 - p) \frac{(a - b)^2}{\max\{a, b\}} \leq d_p(a, b) + d_{1-p}(a, b) \leq p(1 - p) \frac{(a - b)^2}{\min\{a, b\}} \tag{15}$$

Proof. We give two different proofs (I) and (II).

- (I) For $a = b$ or $p \in \{0, 1\}$, we obtain equality in the relation from the statement. Thus, we assume $a \neq b$ and $p \in (0, 1)$. It is easy to see that $d_p(a, b) + d_{1-p}(a, b) = a + b - a^p b^{1-p} - a^{1-p} b^p = (a^p - b^p)(a^{1-p} - b^{1-p})$. Using the Lagrange theorem, there exists c_1 and c_2 between a and b such that $(a^p - b^p)(a^{1-p} - b^{1-p}) = p(1 - p)(a - b)^2 c_1^{p-1} c_2^{-p}$. However, we have the inequality $\frac{1}{\max\{a, b\}} \leq \frac{1}{c_1^{1-p} c_2^p} \leq \frac{1}{\min\{a, b\}}$. Therefore, we deduce the inequality of the statement.
- (II) Using the Cartwright–Field inequality, we have:

$$\frac{1}{2}p(1 - p) \frac{(a - b)^2}{\max\{a, b\}} \leq d_p(a, b) \leq \frac{1}{2}p(1 - p) \frac{(a - b)^2}{\min\{a, b\}}$$

and if we replace p by $1 - p$, we deduce:

$$\frac{1}{2}p(1 - p) \frac{(a - b)^2}{\max\{a, b\}} \leq d_{1-p}(a, b) \leq \frac{1}{2}p(1 - p) \frac{(a - b)^2}{\min\{a, b\}}$$

for $a, b > 0$ and $0 \leq p \leq 1$. By summing up these inequalities, we proved the inequality of the statement:

□

Remark 1. (i) From the proof of Theorem 3, we obtain $A(a, b) - H_p(a, b) = \frac{d_p(a,b)+d_{1-p}(a,b)}{2}$, we deduce an estimation for the Heinz mean:

$$A(a, b) - \frac{1}{2}p(1-p) \frac{(a-b)^2}{\min\{a, b\}} \leq H_p(a, b) \leq A(a, b) - \frac{1}{2}p(1-p) \frac{(a-b)^2}{\max\{a, b\}}. \tag{16}$$

(ii) Since $d_p(a, b) + d_{1-p}(a, b) = (a^p - b^p)(a^{1-p} - b^{1-p})$ and $d_{1-p}(a, b) \geq 0$, we have $0 \leq d_p(a, b) \leq (a^p - b^p)(a^{1-p} - b^{1-p})$ which is in fact the inequality given by Minguzzi (3).

Theorem 4. Let $a, b > 0$ and $0 \leq p \leq 1$.

- (i) For $1/2 \leq p \leq 1$, $a \geq b$ or $0 \leq p \leq 1/2$, $a \leq b$, we have $d_p(a, b) \geq d_{1-p}(a, b)$.
- (ii) For $0 \leq p \leq 1/2$, $a \geq b$ or $1/2 \leq p \leq 1$, $a \leq b$, we have $d_p(a, b) \leq d_{1-p}(a, b)$.

Proof. For $a = b$ or $p \in \{0, 1\}$, we obtain equality in the relations from the statement. Thus, we assume $a \neq b$ and $p \in (0, 1)$. However, we have:

$$\begin{aligned} d_p(a, b) - d_{1-p}(a, b) &= (2p - 1)(a - b) - a^p b^{1-p} + a^{1-p} b^p \\ &= b \left((2p - 1) \left(\frac{a}{b} - 1 \right) - \left(\frac{a}{b} \right)^p + \left(\frac{a}{b} \right)^{1-p} \right). \end{aligned}$$

We consider the function $f : (0, \infty) \rightarrow \mathbb{R}$ defined by $f(t) = (2p - 1)(t - 1) - t^p + t^{1-p}$. We calculate the derivatives of f , thus we have:

$$\begin{aligned} \frac{df(t)}{dt} &= (2p - 1) - pt^{p-1} + (1 - p)t^{-p}, \\ \frac{d^2f(t)}{dt^2} &= (1 - p)pt^{p-2} - p(1 - p)t^{-p-1} = p(1 - p)t^{-p-1} (t^{2p-1} - 1). \end{aligned}$$

For $t > 1$ and $1/2 \leq p < 1$, we have $\frac{d^2f(t)}{dt^2} > 0$, so, function $\frac{df}{dt}$ is increasing, so we obtain $\frac{df(t)}{dt} > \frac{df(1)}{dt} = 0$, which implies that function f is increasing, so we have $f(t) > f(1) = 0$, which means that $(2p - 1)(t - 1) - t^p + t^{1-p} > 0$. For $t = a/b > 1$, we find that $d_p(a, b) > d_{1-p}(a, b)$. For $t < 1$ and $0 < p \leq 1/2$, we have $\frac{d^2f(t)}{dt^2} > 0$, so, function $\frac{df}{dt}$ is increasing, so we obtain $\frac{df(t)}{dt} < \frac{df(1)}{dt} = 0$, which implies that function f is decreasing, so we have $f(t) > f(1) = 0$, which means that $(2p - 1)(t - 1) - t^p + t^{1-p} > 0$. For $t = a/b < 1$, we find that $d_p(a, b) > d_{1-p}(a, b)$. In the analogous way, we show the inequality in (ii). □

Remark 2. From (i) in Theorem 4 for $1/2 \leq p \leq 1$ and $a \geq b$, we have $d_p(a, b) \geq d_{1-p}(a, b)$, so we obtain:

$$\frac{1}{2}p(1-p) \frac{(a-b)^2}{\max\{a, b\}} \leq d_p(a, b), \tag{17}$$

which is just left hand side of Cartwright–Field inequality:

$$\frac{1}{2}p(1-p) \frac{(a-b)^2}{\max\{a, b\}} \leq d_p(a, b) \leq \frac{1}{2}p(1-p) \frac{(a-b)^2}{\min\{a, b\}}, \quad (a, b > 0, 0 \leq p \leq 1).$$

Therefore, it is quite natural to consider the following inequality:

$$d_p(a, b) \geq \frac{1}{2} \left\{ \frac{1}{2}p(1-p) \frac{(a-b)^2}{\max\{a, b\}} + \frac{1}{2}p(1-p) \frac{(a-b)^2}{\min\{a, b\}} \right\} = \frac{1}{4}p(1-p)(a-b)^2 \frac{a+b}{ab}$$

whether it holds or not for a general case $a, b > 0$ and $0 \leq p \leq 1$. However, this inequality does not hold in general. We set the function:

$$h_p(t) = pt + 1 - p - t^p - \frac{p(1-p)}{4} \frac{(t-1)^2(t+1)}{t}, \quad (t > 0, \quad 0 \leq p \leq 1).$$

Then, we have $h_{0.1}(0.3) \simeq -0.00434315$, $h_{0.1}(0.6) \simeq 0.000199783$ and also $h_{0.9}(1.8) \simeq 0.000352199$, $h_{0.9}(2.6) \simeq -0.00282073$.

Theorem 5. For $a, b \geq 1$ and $0 \leq p \leq 1$, we have:

$$\frac{1}{2}p(1-p) \frac{(a-b)^2}{\max\{a,b\}} \leq \frac{1}{2}E_p(a,b) \leq d_p(a,b), \tag{18}$$

where $E_p(a,b) := \min \left\{ \frac{p(a-b)(a^{1-p}-b^{1-p})}{(\max\{a,b\})^{1-p}}, \frac{(1-p)(a-b)(a^p-b^p)}{(\max\{a,b\})^p} \right\} = E_{1-p}(a,b)$.

Proof. For $p = 0$ or $p = 1$ or $a = b$, we have equality. We assume $a \neq b$ and $0 < p < 1$. If $b < a$, then using Theorem 2, we have:

$$\frac{p(a-b)(a^{1-p}-b^{1-p})}{2a^{1-p}} \leq d_p(a,b).$$

Using the Lagrange theorem, we obtain $a^{1-p} - b^{1-p} = (1-p)(a-b)\phi^{-p}$, where $b < \phi < a$. For $b \geq 1$, we deduce $a^{1-p} - b^{1-p} \geq (1-p)(a-b)a^{-p}$, which means that $\frac{1}{2}p(1-p) \frac{(b-a)^2}{a} \leq \frac{p(a-b)(a^{1-p}-b^{1-p})}{2a^{1-p}}$. If $b > a$ and we replace p by $1-p$, then Theorem 2 implies:

$$\frac{(1-p)(a-b)(a^p-b^p)}{2b^p} \leq d_p(a,b). \tag{19}$$

Using the Lagrange theorem, we obtain $b^p - a^p = p(b-a)\theta^{p-1}$, where $a < \theta < b$. For $a \geq 1$, we deduce $b^p - a^p \geq p(b-a)b^{p-1}$, which means that $\frac{1}{2}p(1-p) \frac{(b-a)^2}{b} \leq \frac{(1-p)(a-b)(a^p-b^p)}{2b^p}$. Taking into account the above considerations, we prove the statement. \square

Corollary 1. For $0 < a, b \leq 1$ and $0 \leq p \leq 1$, we have:

$$\frac{1}{2}p(1-p) \frac{(a-b)^2}{\max\{a,b\}} \leq \frac{ab}{2} E_p\left(\frac{1}{a}, \frac{1}{b}\right) \leq d_p(a,b), \tag{20}$$

where $E(\cdot, \cdot)$ is given in Theorem 5.

Proof. For $p = 0$ or $p = 1$ or $a = b$, we have the equality. We assume $a \neq b$ and $0 < p < 1$. If in inequality (18), we replace $a, b \leq 1$ by $\frac{1}{a}, \frac{1}{b} \geq 1$, we deduce:

$$\frac{1}{2}p(1-p) \frac{(a-b)^2}{ab \max\{a,b\}} \leq \frac{1}{2}E_p\left(\frac{1}{a}, \frac{1}{b}\right) \leq d_p\left(\frac{1}{a}, \frac{1}{b}\right) = \frac{1}{ab}d_p(a,b).$$

Consequently, we prove the inequalities of the statement. \square

Theorem 6. For $a, b > 0$ and $0 \leq p \leq 1$, we have:

$$d_p(a,b) \leq (1-p) \frac{(a-b)^2}{b}. \tag{21}$$

Proof. For $p = 0$ or $p = 1$ or $a = b$, we have equality in the relation from the statement. We assume $a \neq b$ and $0 < p < 1$. We consider function $f : (0, \infty) \rightarrow \mathbb{R}$ defined by $f(t) =$

$1 - t^{p-1} - (1-p)(t-1)$, $p \in [0, 1]$. For $t \in (0, 1]$, we have $\frac{df(t)}{dt} = (1-p)(t^{p-2} - 1) \geq 0$, which implies that f is increasing, so we deduce $f(t) \leq f(1) = 0$. For $t \in [1, \infty)$, we have $\frac{df(t)}{dt} \leq 0$, which implies that f is decreasing, so we obtain $f(t) \leq f(1) = 0$. Therefore, we find the following inequality:

$$1 - t^{p-1} \leq (1-p)(t-1).$$

Multiplying the above inequality by $t > 0$, we have:

$$t - t^p \leq (1-p)(t^2 - t),$$

which is equivalent to the inequality:

$$pt + (1-p) - t^p \leq (1-p)(t-1)^2,$$

for all $t > 0$ and $p \in [0, 1]$. Therefore, if we take $t = \frac{a}{b}$ in the above inequality and after some calculations, we deduce the inequality of the statement. \square

Corollary 2. For $a, b > 0$ and $0 \leq p \leq 1$, we have:

$$d_p(a, b) + d_{1-p}(a, b) \leq (1-p) \frac{(a-b)^2(a+b)}{ab}. \quad (22)$$

Proof. For $p = 0$ or $p = 1$ or $a = b$, we have the equality. We assume $a \neq b$ and $0 < p < 1$. If in inequality (21), we exchange a with b , we deduce:

$$d_p(b, a) \leq (1-p) \frac{(a-b)^2}{a}.$$

However, $d_p(b, a) = d_{1-p}(a, b)$, so we have:

$$d_p(a, b) + d_{1-p}(a, b) \leq (1-p) \frac{(a-b)^2}{b} + (1-p) \frac{(a-b)^2}{a} = (1-p) \frac{(a-b)^2(a+b)}{ab}.$$

Consequently, we prove the inequality of the statement. \square

3. Applications to Some Divergences

The Tsallis divergence (e.g., [19,20]) is defined for two probability distributions $\mathbf{p} := \{p_1, \dots, p_n\}$ and $\mathbf{r} := \{r_1, \dots, r_n\}$ with $p_j > 0$ and $r_j > 0$ for all $j = 1, \dots, n$ as

$$D_q^T(\mathbf{p}|\mathbf{r}) := \sum_{j=1}^n \frac{p_j - p_j^q r_j^{1-q}}{1-q}, \quad (q > 0, q \neq 1).$$

The Rényi divergence (e.g., [21]) is also denoted by

$$D_q^R(\mathbf{p}|\mathbf{r}) := \frac{1}{q-1} \log \left(\sum_{j=1}^n p_j^q r_j^{1-q} \right).$$

We see in (e.g., [22]) that:

$$D_q^R(\mathbf{p}|\mathbf{r}) = \frac{1}{q-1} \log \left(1 + (q-1) D_q^T(\mathbf{p}|\mathbf{r}) \right). \quad (23)$$

It is also known that:

$$\lim_{q \rightarrow 1} D_q^T(\mathbf{p}|\mathbf{r}) = \lim_{q \rightarrow 1} D_q^R(\mathbf{p}|\mathbf{r}) = \sum_{j=1}^n p_j \log \frac{p_j}{r_j} =: D(\mathbf{p}|\mathbf{r}),$$

where $D(\mathbf{p}|\mathbf{r})$ is the standard divergence (KL information, relative entropy). The Jeffreys divergence (see [22,23]) is defined by $J_1(\mathbf{p}|\mathbf{r}) := D(\mathbf{p}|\mathbf{r}) + D(\mathbf{r}|\mathbf{p})$ and the Jensen–Shannon divergence [15,16] is defined by

$$JS_1(\mathbf{p}|\mathbf{r}) := \frac{1}{2}D\left(\mathbf{p} \middle| \frac{\mathbf{p} + \mathbf{r}}{2}\right) + \frac{1}{2}D\left(\mathbf{r} \middle| \frac{\mathbf{p} + \mathbf{r}}{2}\right).$$

In [24], the Jeffreys and the Jensen–Shannon divergence are extended to biparametric forms. In [23], Furuichi and Mitroi generalizes these divergences to the Jeffreys–Tsallis divergence, which is given by $J_q(\mathbf{p}|\mathbf{r}) := D_q^T(\mathbf{p}|\mathbf{r}) + D_q^T(\mathbf{r}|\mathbf{p})$ and to the Jensen–Shannon–Tsallis divergence, which is defined as

$$JS_q(\mathbf{p}|\mathbf{r}) := \frac{1}{2}D_q^T\left(\mathbf{p} \middle| \frac{\mathbf{p} + \mathbf{r}}{2}\right) + \frac{1}{2}D_q^T\left(\mathbf{r} \middle| \frac{\mathbf{p} + \mathbf{r}}{2}\right).$$

Several properties of divergences can be extended in the operator theory [25].

For the Tsallis divergence, we have the following relations.

Theorem 7. For two probability distributions $\mathbf{p} := \{p_1, \dots, p_n\}$ and $\mathbf{r} := \{r_1, \dots, r_n\}$ with $p_j > 0$ and $r_j > 0$ for all $j = 1, \dots, n$, we have:

$$q \sum_{j=1}^n \frac{(p_j - r_j)^2}{\max\{p_j, r_j\}} \leq J_q(\mathbf{p}|\mathbf{r}) \leq q \sum_{j=1}^n \frac{(p_j - r_j)^2}{\min\{p_j, r_j\}}, \quad (0 < q < 1). \tag{24}$$

Proof. From the definition of the Tsallis divergence, we deduce the equality:

$$J_q(\mathbf{p}|\mathbf{r}) = \sum_{j=1}^n \frac{p_j + r_j - p_j^q r_j^{1-q} - p_j^{1-q} r_j^q}{1 - q} = \frac{1}{1 - q} \sum_{j=1}^n \{d_q(p_j, r_j) + d_{1-q}(p_j, r_j)\},$$

where $d(\cdot, \cdot)$ is defined in (6). Applying Theorem 3, we obtain:

$$q(1 - q) \sum_{j=1}^n \frac{(p_j - r_j)^2}{\max\{p_j, r_j\}} \leq \sum_{j=1}^n \{d_q(p_j, r_j) + d(p_j, r_j)\} \leq q(1 - q) \sum_{j=1}^n \frac{(p_j - r_j)^2}{\min\{p_j, r_j\}}$$

and combining with the above equality, we deduce the inequalities (24). \square

Remark 3. (i) In the limit of $q \rightarrow 1$ in (24), we then obtain:

$$\sum_{j=1}^n \frac{(p_j - r_j)^2}{\max\{p_j, r_j\}} \leq J_1(\mathbf{p}|\mathbf{r}) \leq \sum_{j=1}^n \frac{(p_j - r_j)^2}{\min\{p_j, r_j\}}$$

for the standard divergence.

(ii) From (23), we have:

$$\begin{aligned} 2 + (q - 1) \{D_q^T(\mathbf{p}|\mathbf{r}) + D_q^T(\mathbf{r}|\mathbf{p})\} &= \exp\left((q - 1)D_q^R(\mathbf{p}|\mathbf{r})\right) + \exp\left((q - 1)D_q^R(\mathbf{r}|\mathbf{p})\right) \\ &\geq 2 + (q - 1) \{D_q^R(\mathbf{p}|\mathbf{r}) + D_q^R(\mathbf{r}|\mathbf{p})\}, \end{aligned}$$

where we used the inequality $e^x \geq x + 1$ for all $x \in \mathbb{R}$. Thus, we deduce the inequalities:

$$D_q^T(\mathbf{p}|\mathbf{r}) + D_q^T(\mathbf{r}|\mathbf{p}) \leq D_q^R(\mathbf{p}|\mathbf{r}) + D_q^R(\mathbf{r}|\mathbf{p}), \quad (0 < q < 1) \tag{25}$$

and:

$$D_q^T(\mathbf{p}|\mathbf{r}) + D_q^T(\mathbf{r}|\mathbf{p}) \geq D_q^R(\mathbf{p}|\mathbf{r}) + D_q^R(\mathbf{r}|\mathbf{p}), \quad (q > 1).$$

Combining (25) with Theorem 7, we therefore have the following result for the Rényi divergence:

$$q \sum_{j=1}^n \frac{(p_j - r_j)^2}{\max\{p_j, r_j\}} \leq D_q^R(\mathbf{p}|\mathbf{r}) + D_q^R(\mathbf{r}|\mathbf{p}), \quad (0 < q < 1).$$

We give the relation between the Jeffreys–Tsallis divergence and the Jensen–Shannon–Tsallis divergence:

Theorem 8. For two probability distributions $\mathbf{p} := \{p_1, \dots, p_n\}$ and $\mathbf{r} := \{r_1, \dots, r_n\}$ with $p_j > 0$ and $r_j > 0$ for all $j = 1, \dots, n$, we have:

$$JS_q(\mathbf{p}|\mathbf{r}) \leq \frac{1}{4}J_q(\mathbf{p}|\mathbf{r}), \tag{26}$$

where $q \geq 0$ with $q \neq 1$.

Proof. We consider the function $g : (0, \infty) \rightarrow \mathbb{R}$ defined by $g(t) = t^{1-q}$, which is concave for $q \in [0, 1)$. Therefore, we have $\left(\frac{p_j+r_j}{2}\right)^{1-q} \geq \frac{p_j^{1-q}+r_j^{1-q}}{2}$, which implies the following inequalities:

$$p_j - p_j^q \left(\frac{p_j+r_j}{2}\right)^{1-q} \leq \frac{p_j - p_j^q r_j^{1-q}}{2}, r_j - r_j^q \left(\frac{p_j+r_j}{2}\right)^{1-q} \leq \frac{r_j - r_j^q p_j^{1-q}}{2},$$

From the definition of the Tsallis divergence, we deduce the inequality:

$$D_q^T\left(\mathbf{p} \middle| \frac{\mathbf{p}+\mathbf{r}}{2}\right) + D_q^T\left(\mathbf{r} \middle| \frac{\mathbf{p}+\mathbf{r}}{2}\right) \leq \frac{1}{2}\left(D_q^T(\mathbf{p}|\mathbf{r}) + D_q^T(\mathbf{r}|\mathbf{p})\right),$$

which is equivalent to the relation of the statement. For the case of $q > 1$, the function $g(t) = t^{1-q}$ is convex in $t > 0$. Similarly, we have the statement, taking into account that $1 - q < 0$. \square

Remark 4. In the limit of $q \rightarrow 1$ in (26), we then obtain:

$$JS_1(\mathbf{p}|\mathbf{r}) \leq \frac{1}{4}J_1(\mathbf{p}|\mathbf{r}).$$

We give the bounds on the Jeffreys–Tsallis divergence by using the refined Young inequality given in Theorem 1. In [26], we found the *Battacharyya coefficient* defined as

$$B(\mathbf{p}|\mathbf{r}) := \sum_{j=1}^n \sqrt{p_j r_j},$$

which is a measure of the amount of overlapping between two distributions. This can be expressed in terms of the Hellinger distance between the probability distributions $\mathbf{p} := \{p_1, \dots, p_n\}$ and $\mathbf{r} := \{r_1, \dots, r_n\}$, which is given by

$$B(\mathbf{p}|\mathbf{r}) = 1 - h^2(\mathbf{p}|\mathbf{r}),$$

where the *Hellinger distance* ([26,27]) is a metric distance and defined by

$$h(\mathbf{p}|\mathbf{r}) := \frac{1}{\sqrt{2}} \sqrt{\sum_{j=1}^n (\sqrt{p_j} - \sqrt{r_j})^2}.$$

Theorem 9. For two probability distributions $\mathbf{p} := \{p_1, \dots, p_n\}$ and $\mathbf{r} := \{r_1, \dots, r_n\}$ with $p_j > 0$ and $r_j > 0$ for all $j = 1, \dots, n$, and $0 \leq q < 1$, we have:

$$\begin{aligned} \frac{4r}{1-q} h^2(\mathbf{p}|\mathbf{r}) + \frac{2A(q)}{1-q} \sum_{j=1}^n p_j r_j \cdot \log^2 \left(\frac{p_j}{r_j} \right) &\leq J_q(\mathbf{p}|\mathbf{r}) \\ &\leq \frac{4(1-r)}{1-q} h^2(\mathbf{p}|\mathbf{r}) + \frac{2B(q)}{1-q} \sum_{j=1}^n p_j r_j \cdot \log^2 \left(\frac{p_j}{r_j} \right). \end{aligned} \tag{27}$$

where $r = \min\{q, 1 - q\}$ and $A(q) = \frac{q(1-q)}{2} - \frac{r}{4}$, $B(q) = \frac{q(1-q)}{2} - \frac{1-r}{4}$.

Proof. For $q = 0$, we obtain the equality. Now, we consider $0 < q < 1$. Using Theorem 1 for $a = p_j < 1$ and $b = r_j < 1$, $j \in \{1, 2, \dots, n\}$, we deduce:

$$\begin{aligned} r(\sqrt{p_j} - \sqrt{r_j})^2 + A(q)p_j r_j \cdot \log^2 \left(\frac{p_j}{r_j} \right) &\leq d_q(p_j, r_j) \\ &\leq (1-r)(\sqrt{p_j} - \sqrt{r_j})^2 + B(q)p_j r_j \cdot \log^2 \left(\frac{p_j}{r_j} \right), \end{aligned}$$

where $r = \min\{q, 1 - q\}$. If we replace q by $1 - q$ and taking into account that $A(q) = A(1 - q)$ and $B(q) = B(1 - q)$, then we have:

$$\begin{aligned} 2r(\sqrt{p_j} - \sqrt{r_j})^2 + 2A(q)p_j r_j \cdot \log^2 \left(\frac{p_j}{r_j} \right) &\leq d_q(p_j, r_j) + d_{1-q}(p_j, r_j) \\ &\leq 2(1-r)(\sqrt{p_j} - \sqrt{r_j})^2 + 2B(q)p_j r_j \cdot \log^2 \left(\frac{p_j}{r_j} \right). \end{aligned}$$

Taking the sum on $j = 1, 2, \dots, n$, we find the inequalities:

$$\begin{aligned} 2r \sum_{j=1}^n (\sqrt{p_j} - \sqrt{r_j})^2 + 2A(q) \sum_{j=1}^n p_j r_j \cdot \log^2 \left(\frac{p_j}{r_j} \right) &\leq \sum_{j=1}^n (d_q(p_j, r_j) + d_{1-q}(p_j, r_j)) \\ = (1-q) \left(D_q^T(\mathbf{p}|\mathbf{r}) + D_q^T(\mathbf{r}|\mathbf{p}) \right) &\leq 2(1-r) \sum_{j=1}^n (\sqrt{p_j} - \sqrt{r_j})^2 + 2B(q) \sum_{j=1}^n p_j r_j \cdot \log^2 \left(\frac{p_j}{r_j} \right), \end{aligned}$$

which is equivalent to the inequalities in the statement. \square

Remark 5. In the limit of $q \rightarrow 1$ in (27), we then obtain:

$$4h^2(\mathbf{p}|\mathbf{r}) + \frac{1}{2} \sum_{j=1}^n p_j r_j \cdot \log^2 \left(\frac{p_j}{r_j} \right) \leq J_1(\mathbf{p}|\mathbf{r}),$$

since $\lim_{q \rightarrow 1} \frac{r}{1-q} = 1$, $\lim_{q \rightarrow 1} \frac{A(q)}{1-q} = \frac{1}{4}$ and $\lim_{q \rightarrow 1} \frac{1-r}{1-q} = \infty$.

We give the further bounds on the Jeffreys–Tsallis divergence by the use of Theorem 5 and Corollary 2:

Theorem 10. For two probability distributions $\mathbf{p} := \{p_1, \dots, p_n\}$ and $\mathbf{r} := \{r_1, \dots, r_n\}$ with $p_j > 0$ and $r_j > 0$ for all $j = 1, \dots, n$, and $0 \leq q < 1$, we have:

$$q \sum_{j=1}^n \frac{(p_j - r_j)^2}{\max\{p_j, r_j\}} \leq \frac{1}{1-q} \sum_{j=1}^n p_j r_j E_q \left(\frac{1}{p_j}, \frac{1}{r_j} \right) \leq J_q(\mathbf{p}|\mathbf{r}) \leq \sum_{j=1}^n \frac{(p_j - r_j)^2 (p_j + r_j)}{p_j r_j}, \quad (28)$$

where $E(\cdot, \cdot)$ is given in Theorem 5.

Proof. Putting $a := p_j, b := r_j$ and $p := q$ in (20), we deduce:

$$\frac{1}{2} q(1-q) \frac{(p_j - r_j)^2}{\max\{p_j, r_j\}} \leq \frac{p_j r_j}{2} E_q \left(\frac{1}{p_j}, \frac{1}{r_j} \right) \leq d_q(p_j, r_j),$$

and:

$$\frac{1}{2} q(1-q) \frac{(p_j - r_j)^2}{\max\{p_j, r_j\}} \leq \frac{p_j r_j}{2} E_{1-q} \left(\frac{1}{p_j}, \frac{1}{r_j} \right) \leq d_{1-q}(p_j, r_j).$$

Taking into account that:

$$E_q \left(\frac{1}{p_j}, \frac{1}{r_j} \right) = E_{1-q} \left(\frac{1}{p_j}, \frac{1}{r_j} \right),$$

and by taking the sum on $j = 1, 2, \dots, n$, we have:

$$q(1-q) \sum_{j=1}^n \frac{(p_j - r_j)^2}{\max\{p_j, r_j\}} \leq \sum_{j=1}^n p_j r_j E_q \left(\frac{1}{p_j}, \frac{1}{r_j} \right) \leq \sum_{j=1}^n (d_q(p_j, r_j) + d_{1-q}(p_j, r_j))$$

we prove the lower bounds of $J_q(\mathbf{p}|\mathbf{r})$. To prove the upper bound of $J_q(\mathbf{p}|\mathbf{r})$, we put $a := p_j, b := r_j$ and $p := q$ in inequality (22). Then, we deduce:

$$d_q(p_j, r_j) + d_{1-q}(p_j, r_j) \leq (1-q) \frac{(p_j - r_j)^2 (p_j + r_j)}{p_j r_j}.$$

By taking the sum on $j = 1, 2, \dots, n$, we find:

$$\sum_{j=1}^n (d_q(p_j, r_j) + d_{1-q}(p_j, r_j)) \leq (1-q) \sum_{j=1}^n \frac{(p_j - r_j)^2 (p_j + r_j)}{p_j r_j}.$$

Consequently, we prove the inequalities of the statement. \square

We also give the further bounds on the Jeffreys–Tsallis divergence by the use of Cartwright–Field inequality given in (7).

Theorem 11. For two probability distributions $\mathbf{p} := \{p_1, \dots, p_n\}$ and $\mathbf{r} := \{r_1, \dots, r_n\}$ with $p_j > 0$ and $r_j > 0$ for all $j = 1, \dots, n$, and $0 \leq q < 1$, we have:

$$\begin{aligned} & \frac{q}{8} \sum_{j=1}^n (p_j - r_j)^2 \left(\frac{1}{p_j + \max\{p_j, r_j\}} + \frac{1}{r_j + \max\{p_j, r_j\}} \right) \\ & \leq JS_q(\mathbf{p}|\mathbf{r}) \leq \frac{q}{8} \sum_{j=1}^n (p_j - r_j)^2 \left(\frac{1}{p_j + \min\{p_j, r_j\}} + \frac{1}{r_j + \min\{p_j, r_j\}} \right) \end{aligned} \quad (29)$$

Proof. For $q = 0$, we have the equality. We assume $0 < q < 1$. By direct calculations, we have:

$$\begin{aligned}
 JS_q(\mathbf{p}|\mathbf{r}) &= \frac{1}{2}D_q^T\left(\mathbf{p}|\frac{\mathbf{p}+\mathbf{r}}{2}\right) + \frac{1}{2}D_q^T\left(\mathbf{r}|\frac{\mathbf{p}+\mathbf{r}}{2}\right) \\
 &= \frac{1}{2(1-q)}\sum_{j=1}^n\left\{p_j - p_j^q\left(\frac{p_j+r_j}{2}\right)^{1-q} + r_j - r_j^q\left(\frac{p_j+r_j}{2}\right)^{1-q}\right\} \\
 &= \frac{1}{2(1-q)}\sum_{j=1}^n\left\{qp_j + (1-q)\frac{p_j+r_j}{2} - p_j^q\left(\frac{p_j+r_j}{2}\right)^{1-q} + qr_j + (1-q)\frac{p_j+r_j}{2} - r_j^q\left(\frac{p_j+r_j}{2}\right)^{1-q}\right\} \\
 &= \frac{1}{2(1-q)}\sum_{j=1}^n\left\{d_q\left(p_j, \frac{p_j+r_j}{2}\right) + d_q\left(r_j, \frac{p_j+r_j}{2}\right)\right\}.
 \end{aligned}$$

Using inequality (7), we deduce:

$$\frac{q(1-q)}{4}\frac{(p_j-r_j)^2}{p_j+\max\{p_j,r_j\}} \leq d_q\left(p_j, \frac{p_j+r_j}{2}\right) \leq \frac{q(1-q)}{4}\frac{(p_j-r_j)^2}{p_j+\min\{p_j,r_j\}}$$

and:

$$\frac{q(1-q)}{4}\frac{(p_j-r_j)^2}{r_j+\max\{p_j,r_j\}} \leq d_q\left(r_j, \frac{p_j+r_j}{2}\right) \leq \frac{q(1-q)}{4}\frac{(p_j-r_j)^2}{r_j+\min\{p_j,r_j\}}.$$

From the above inequalities, we have the statement, by summing on $j = 1, 2, \dots, n$. \square

It is quite natural to extend the Jensen–Shannon–Tsallis divergence to the following form:

$$JS_q^v(\mathbf{p}|\mathbf{r}) := vD_q^T(\mathbf{p}|v\mathbf{p} + (1-v)\mathbf{r}) + (1-v)D_q^T(\mathbf{r}|v\mathbf{p} + (1-v)\mathbf{r}),$$

where $0 \leq v \leq 1, q > 0, q \neq 1$. We call this the v -weighted Jensen–Shannon–Tsallis divergence. For $v = 1/2$, we find that $JS_q^{1/2}(\mathbf{p}|\mathbf{r}) = JS_q(\mathbf{p}|\mathbf{r})$ which is the Jensen–Shannon–Tsallis divergence. For this quantity $JS_q^v(\mathbf{p}|\mathbf{r})$, we can obtain the following result in a way similar to the proof of the Theorem 11.

Proposition 1. For two probability distributions $\mathbf{p} := \{p_1, \dots, p_n\}$ and $\mathbf{r} := \{r_1, \dots, r_n\}$ with $p_j > 0$ and $r_j > 0$ for all $j = 1, \dots, n, 0 \leq q < 1$ and $0 \leq v \leq 1$, we have:

$$\begin{aligned}
 &\frac{qv(1-v)}{2}\sum_{j=1}^n(p_j-r_j)^2\left(\frac{1-v}{vp_j+(1-v)\max\{p_j,r_j\}} + \frac{v}{(1-v)r_j+v\max\{p_j,r_j\}}\right) \\
 &\leq JS_q^v(\mathbf{p}|\mathbf{r}) \\
 &\leq \frac{qv(1-v)}{2}\sum_{j=1}^n(p_j-r_j)^2\left(\frac{1-v}{vp_j+(1-v)\min\{p_j,r_j\}} + \frac{v}{(1-v)r_j+v\min\{p_j,r_j\}}\right).
 \end{aligned}$$

Proof. We calculate as

$$\begin{aligned}
JS_q^v(\mathbf{p}|\mathbf{r}) &= \frac{v}{1-q} \sum_{j=1}^n \left\{ p_j - p_j^q (vp_j + (1-v)r_j)^{1-q} \right\} + \frac{1-v}{1-q} \sum_{j=1}^n \left\{ r_j - r_j^q (vp_j + (1-v)r_j)^{1-q} \right\} \\
&= \frac{1}{1-q} \sum_{j=1}^n \left\{ vp_j + (1-v)r_j - vp_j^q (vp_j + (1-v)r_j)^{1-q} - (1-v)r_j^q (vp_j + (1-v)r_j)^{1-q} \right\} \\
&= \frac{v}{1-q} \sum_{j=1}^n \left\{ qp_j + (1-q)(vp_j + (1-v)r_j) - p_j^q (vp_j + (1-v)r_j)^{1-q} \right\} \\
&+ \frac{1-v}{1-q} \sum_{j=1}^n \left\{ qr_j + (1-q)(vp_j + (1-v)r_j) - r_j^q (vp_j + (1-v)r_j)^{1-q} \right\} \\
&= \frac{1}{1-q} \sum_{j=1}^n \left\{ vd_q(p_j, vp_j + (1-v)r_j) + (1-v)d_q(r_j, vp_j + (1-v)r_j) \right\}.
\end{aligned}$$

Using inequality (7), we deduce:

$$\begin{aligned}
\frac{q(1-q)}{2} \frac{(1-v)^2(p_j - r_j)^2}{vp_j + (1-v)\max\{p_j, r_j\}} &\leq d_q(p_j, vp_j + (1-v)r_j) \\
&\leq \frac{q(1-q)}{2} \frac{(1-v)^2(p_j - r_j)^2}{vp_j + (1-v)\min\{p_j, r_j\}}
\end{aligned}$$

and:

$$\begin{aligned}
\frac{q(1-q)}{2} \frac{v^2(p_j - r_j)^2}{(1-v)r_j + v\max\{p_j, r_j\}} &\leq d_q(r_j, vp_j + (1-v)r_j) \\
&\leq \frac{q(1-q)}{2} \frac{v^2(p_j - r_j)^2}{(1-v)r_j + v\min\{p_j, r_j\}}.
\end{aligned}$$

Multiplying v and $1-v$ by the above inequalities, respectively, and then taking the sum on $j = 1, 2, \dots, n$, we obtain the statement. \square

4. Conclusions

We obtained new inequalities which improve classical Young inequality by analytical calculations with known inequalities. We also obtained some bounds on the Jeffreys–Tsallis divergence and the Jensen–Shannon–Tsallis divergence. At this point, we do not clearly know whether the obtained bounds will play any role in the information theory. However, if there exists a purpose to find the meaning of the parameter q in divergences based on Tsallis divergence, then we may state that almost all theorems (except for Theorem 8) hold for $0 \leq q < 1$. In the first author's previous studies [19,28], some results related to Tsallis divergence (relative entropy) are still true for $0 \leq q < 1$, while some results related to Tsallis entropy are still true for $q > 1$. In this paper, we treated the Tsallis type divergence so it is shown that almost all results are true for $0 \leq q < 1$. This insight may give a rough meaning of the parameter q .

Since our results in Section 3 are based on the inequalities in Section 2, we summarized the tightness for our obtained inequalities in Section 2. The double inequality (12) is a counterpart of the double inequality (9) for $a, b \in (0, 1]$. Therefore, they can not be compared with each other from the point of view on the tightness, since the conditions are different. The double inequality (12) was used to obtain Theorem 9. The double inequality (15) is essentially a Cartwright–Field inequality in itself, and it was used to obtain Theorem 7 as the first result in Section 3. The results in Theorem 4 are mathematical properties on $d_p(a, b)$. The inequalities given in (18) gave an improvement of the left-hand side in the inequality (7) for the case $a, b \geq 1$ and we obtained Theorem 10 by (18). We obtained the upper bound of $d_p(a, b)$ as a counterpart of (18) for a general $a, b > 0$. This is

used to prove Corollary 2 which was used to prove Theorem 10. However, we found that the upper bound of $d_p(a, b) + d_{1-p}(a, b)$ given in (22) is not tighter than the one in (15).

Finally, Theorem 8 can be obtained from the convexity/concavity of the function t^{1-q} . This study will be continued in order to obtain much sharper bounds. We extend the Jensen–Shannon–Tsallis divergence to the following:

$$JS_q^v(\mathbf{p}|\mathbf{r}) := vD_q^T(\mathbf{p}|v\mathbf{p} + (1-v)\mathbf{r}) + (1-v)D_q^T(\mathbf{r}|v\mathbf{p} + (1-v)\mathbf{r}), \quad (0 \leq v \leq 1, q > 0, q \neq 1),$$

and we call this the v -weighted Jensen–Shannon–Tsallis divergence. For $v = 1/2$, we find that $JS_q^{1/2}(\mathbf{p}|\mathbf{r}) = JS_q(\mathbf{p}|\mathbf{r})$ which is the Jensen–Shannon–Tsallis divergence. For this quantity, as an information-theoretic divergence measure $JS_q^v(\mathbf{p}|\mathbf{r})$, we obtained several characterizations.

Author Contributions: Conceptualization, S.F. and N.M.; investigation, S.F. and N.M.; writing—original draft preparation, S.F. and N.M.; writing—review and editing, S.F.; funding acquisition, S.F. and N.M. All authors have read and agreed to the published version of the manuscript.

Funding: The author (S.F.) was partially supported by JSPS KAKENHI Grant Number 21K03341.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to thank the referees for their careful and insightful comments to improve our manuscript.

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

References

- Young, W.H. On classes of summable functions and their Fourier series. *Proc. R. Soc. Lond. Ser. A* **1912**, *87*, 225–229.
- Blondel, M.; Martins, A.F.T.; Niculae, V. Learning with Fenchel–Young Losses. *J. Mach. Learn. Res.* **2020**, *21*, 1–69.
- Nielsen, F. On Geodesic Triangles with Right Angles in a Dually Flat Space. In *Progress in Information Geometry: Theory and Applications*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 153–190.
- Minguzzi, E. An equivalent form of Young’s inequality with upper bound. *Appl. Anal. Discrete Math.* **2008**, *2*, 213–216. [[CrossRef](#)]
- Nielsen, F. The α -divergences associated with a pair of strictly comparable quasi-arithmetic means. *arXiv* **2020**, arXiv:2001.09660.
- Bhatia, R. Interpolating the arithmetic–geometric mean inequality and its operator version. *Linear Alg. Appl.* **2006**, *413*, 355–363. [[CrossRef](#)]
- Furuichi, S.; Ghaemi, M.B.; Gharakhanlu, N. Generalized reverse Young and Heinz inequalities. *Bull. Malays. Math. Sci. Soc.* **2019**, *42*, 267–284. [[CrossRef](#)]
- Cartwright, D.I.; Field, M.J. A refinement of the arithmetic mean–geometric mean inequality. *Proc. Am. Math. Soc.* **1978**, *71*, 36–38. [[CrossRef](#)]
- Kober, H. On the arithmetic and geometric means and Hölder inequality. *Proc. Am. Math. Soc.* **1958**, *9*, 452–459.
- Kittaneh, F.; Manasrah, Y. Improved Young and Heinz inequalities for matrix. *J. Math. Anal. Appl.* **2010**, *361*, 262–269. [[CrossRef](#)]
- Bobylev, N.A.; Krasnoselsky, M.A. *Extremum Analysis (Degenerate Cases)*; Institute of Control Sciences: Moscow, Russia, 1981; 52p. (In Russian)
- Minculete, N. A refinement of the Kittaneh–Manasrah inequality. *Creat. Math. Inform.* **2011**, *20*, 157–162.
- Furuichi, S.; Minculete, N. Alternative reverse inequalities for Young’s inequality. *J. Math. Inequal.* **2011**, *5*, 595–600. [[CrossRef](#)]
- Furuichi, S.; Moradi, H.R. *Advances in Mathematical Inequalities*; De Gruyter: Berlin, Germany, 2020.
- Lin, J. Divergence measures based on the Shannon entropy. *IEEE Trans. Inform. Theory* **1991**, *37*, 145–151. [[CrossRef](#)]
- Sibson, R. Information radius. *Z. Wahrscheinlichkeitstheorie Verw Gebiete* **1969**, *14*, 149–160. [[CrossRef](#)]
- Mitroi-Symeonidis, F.C.; Anghel, I.; Minculete, N. Parametric Jensen–Shannon Statistical Complexity and Its Applications on Full-Scale Compartment Fire Data. *Symmetry* **2020**, *12*, 22. [[CrossRef](#)]
- Niculescu, C.P.; Persson, L.-E. *Convex Functions and Their Applications*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2018.
- Furuichi, S.; Yanagi, K.; Kuriyama, K. Fundamental properties of Tsallis relative entropy. *J. Math. Phys.* **2004**, *45*, 4868–4877. [[CrossRef](#)]
- Tsallis, C. Generalized entropy-based criterion for consistent testing. *Phys. Rev. E* **1998**, *58*, 1442–1445. [[CrossRef](#)]
- Aczél, J.; Daróczy, Z. *On Measures of Information and Their Characterizations*; Academic Press: Cambridge, MA, USA, 1975.
- Furuichi, S.; Minculete, N. Inequalities related to some types of entropies and divergences. *Physica A* **2019**, *532*, 121907. [[CrossRef](#)]
- Furuichi, S.; Mitroi, F.-C. Mathematical inequalities for some divergences. *Physica A* **2012**, *391*, 388–400. [[CrossRef](#)]

24. Mitroi, F.C.; Minculete, N. Mathematical inequalities for biparametric extended information measures. *J. Math. Ineq.* **2013**, *7*, 63–71. [[CrossRef](#)]
25. Moradi, H.R.; Furuichi, S.; Minculete, N. Estimates for Tsallis relative operator entropy. *Math. Ineq. Appl.* **2017**, *20*, 1079–1088. [[CrossRef](#)]
26. Lovričević, N.; Pečarić, D.; Pečarić, J. Zipf-Mandelbrot law, f-divergences and the Jensen-type interpolating inequalities. *J. Inequal. Appl.* **2018**, *2018*, 36. [[CrossRef](#)] [[PubMed](#)]
27. Van Erven, T.; Harremoës, P. Rényi Divergence and Kullback -Leibler Divergence. *IEEE Trans. Inf. Theory* **2014**, *60*, 3797–3820. [[CrossRef](#)]
28. Furuichi, S. Information theoretical properties of Tsallis entropies. *J. Math. Phys.* **2006**, *47*, 023302. [[CrossRef](#)]