Mathematical
Inequalities
& Applications
Volume 21, Number 2 (2018), 455–467 doi:10.7153/mia-2018-21-31

ON CSISZÁR AND TSALLIS TYPE f–DIVERGENCES INDUCED BY SUPERQUADRATIC AND CONVEX FUNCTIONS

PAWEŁ A. KLUZA AND MAREK NIEZGODA

(*Communicated by M. Praljak*)

Abstract. In this paper, Csiszár and Tsallis f-divergences are studied for superquadratic and convex functions. Some comparison theorems for two divergences are provided. The obtained results, when used for nonnegative superquadratic functions, give some refinements of the original inequalities corresponding to nonnegative convex functions. Some majorization assumptions for the involved matrix are simplified from column stochasticity to entrywise-nonnegativity.

1. Introduction

Throughout $\mathbb{R}_+$ and $\mathbb{R}_{++}$ denote the sets of nonnegative and positive numbers, respectively, i.e., $\mathbb{R}_+ = [0,\infty)$ and $\mathbb{R}_{++} = (0,\infty)$.

Let $f : \mathbb{R}_+ \to \mathbb{R}$ be a convex function, and $\mathbf{p} = (p_1,\dots,p_n)$ and $\mathbf{q} = (q_1,\dots,q_n)$ with $p_i, q_i \geqslant 0$, $i = 1,\dots,n$. The *Csiszár f-divergence* is defined by

$$C_f(\mathbf{p},\mathbf{q}) = \sum_{i=1}^{n} p_i f\left(\frac{q_i}{p_i}\right) \tag{1}$$

with the conventions $0f\left(\frac{0}{0}\right) = 0$ and $0f\left(\frac{c}{0}\right) = c\lim\limits_{t\to\infty}\frac{f(t)}{t}$, $c > 0$ (see [8, 9, 10]).

The Csiszár-Körner inequality asserts that

$$\sum_{i=1}^{n} p_i f\left(\frac{\sum_{i=1}^{n} q_i}{\sum_{i=1}^{n} p_i}\right) \leqslant C_f(\mathbf{p},\mathbf{q}) \tag{2}$$

(see [9, 17]). If in addition $\sum_{i=1}^{n} q_i = \sum_{i=1}^{n} p_i$ and $f(1) = 0$, then (2) implies

$$0 \leqslant C_f(\mathbf{p},\mathbf{q}). \tag{3}$$

For other inequalities for f-divergence, consult [10, 12].

A function $f : [0,\infty) \to \mathbb{R}$ is said to be *superquadratic* provided that for each $x \geqslant 0$ there exists a constant $c(x) \in \mathbb{R}$ such that

$$f(y) - f(x) - f(|y-x|) \geqslant c(x)(y-x) \tag{4}$$

Mathematics subject classification (2010): 26D15, 15B51, 94A17.

Keywords and phrases: convex function, superquadratic function, Csiszár f-divergence, Tsallis f_u-divergence, column stochastic matrix.

for all $y \in [0,\infty)$ (see [1, 3]). We say that f is *subquadratic* if $-f$ is superquadratic.

As noted in [1, p. 3], the functions x^p with $p \geqslant 2$, $x^2 \log x$ and $\sinh x$ are superquadratic. Moreover, if $f(0) = f'(0) = 0$ and f' is convex, then f is superquadratic (see [1, Lemma 2.2]).

The next result shows the Jensen type inequality for superquadratic functions.

THEOREM A. (Abramovich et al. [3]) *Suppose that f is a superquadratic function on $[0,\infty)$, $x_i \in [0,\infty)$, and $a_i \geqslant 0$, $i = 1,\ldots,n$, are such that $A_n = \sum_{i=1}^n a_i > 0$. Then*

$$f\left(\frac{1}{A_n}\sum_{i=1}^n a_i x_i\right) + \frac{1}{A_n}\sum_{i=1}^n a_i f\left(\left|x_i - \frac{1}{A_n}\sum_{i=1}^n a_i x_i\right|\right) \leqslant \frac{1}{A_n}\sum_{i=1}^n a_i f(x_i). \tag{5}$$

We say that an $n \times m$ real matrix $S = (s_{ij})$ is *column stochastic* (resp. *row stochastic*) if $s_{ij} \geqslant 0$ for $i = 1,\ldots,n$, $j = 1,\ldots,m$, and all column sums (resp. row sums) of S are equal to 1, i.e., $\sum\limits_{i=1}^n s_{ij} = 1$ for $j = 1,\ldots,m$ (resp. $\sum\limits_{j=1}^m s_{ij} = 1$ for $i = 1,\ldots,n$).

As usual, by S^T we denote the transpose matrix of S.

The following theorem describes Sherman's inequality (6) for convex functions.

THEOREM B. (Sherman [20]) *Let $f : \mathbb{R}_+ \to \mathbb{R}$ be convex on $\mathbb{R}_+$ and $\mathbf{x} = (x_1,\ldots,x_n) \in \mathbb{R}_+^n$, $\mathbf{y} = (y_1,\ldots,y_m) \in \mathbb{R}_+^m$, $\mathbf{a} = (a_1,\ldots,a_n) \in \mathbb{R}_+^n$ and $\mathbf{b} = (b_1,\ldots,b_m) \in \mathbb{R}_+^m$.*

Assume $\mathbf{y} = \mathbf{x}S$ and $\mathbf{a} = \mathbf{b}S^T$ for some $n \times m$ column stochastic matrix $S = (s_{ij})$. Then

$$\sum_{j=1}^m b_j f(y_j) \leqslant \sum_{i=1}^n a_i f(x_i). \tag{6}$$

THEOREM C. ([18]) *Let $f : \mathbb{R}_+ \to \mathbb{R}$ be superquadratic on $\mathbb{R}_+$ and $\mathbf{x} = (x_1,\ldots,x_n) \in \mathbb{R}_+^n$, $\mathbf{y} = (y_1,\ldots,y_m) \in \mathbb{R}_+^m$, $\mathbf{a} = (a_1,\ldots,a_n) \in \mathbb{R}_+^n$ and $\mathbf{b} = (b_1,\ldots,b_m) \in \mathbb{R}_+^m$.*

Assume $\mathbf{y} = \mathbf{x}S$ and $\mathbf{a} = \mathbf{b}S^T$ for some $n \times m$ column stochastic matrix $S = (s_{ij})$. Then

$$\sum_{j=1}^m b_j f(y_j) + \sum_{i=1}^n \sum_{j=1}^m b_j s_{ij} f(|x_i - y_j|) \leqslant \sum_{i=1}^n a_i f(x_i). \tag{7}$$

Csiszár f-divergence (1) has various applications in statistical physics, biology, economics, Markov processes, population dynamics etc. (see [5, 7]). In particular, specifications of (1) for some functions f lead to some physical notions [7]. For instance, the case $f(t) = -\ln t$ gives the Kullback-Leibler divergence, which can be viewed as the free energy difference [7]. Likewise, the Rényi entropy for $f(t) = t^{1-\alpha} - 1$, $\alpha \in (0,1)$, is connected with the free energy change [4]. Furthermore, the function $f(t) = \frac{t^{1-\alpha}-1}{\alpha-1}$, $\alpha \geqslant 0$, $\alpha \neq 1$, induces the Tsallis entropy [19].

In recent years there has been a variety of research on bounding a given divergence [6, 10, 13]. In this paper our aim is to present some comparison results for Csiszár and Tsallis type f-divergences induced by superquadratic and/or convex functions f.

Taking into account some additional leading coefficients in the definition of f-divergence leads to the new notion of *generalized f-divergence*. By making use of

Theorems A, B, C, we construct a transformation of a given generalized f-divergence to obtain a smaller one. It turns that the provided coefficients play an important role in the transformation. Likewise, an involved matrix in the Sherman type condition is essential, too.

In contrast to the standard approach, we *do not assume* that the matrix is column stochastic. We relax that property to the usual entrywise-nonnegativity of the matrix. Thus the demonstrated theory is simplified.

In Section 2 we obtain some estimations for an f-divergence with a superquadratic function f. Section 3 is devoted to corresponding results for convex f. In general, the convex case is simpler than the superquadratic one. However, the bounds derived for superquadratic functions are more delicate. In fact, if f is nonnegative and superquadratic then it must be convex. In this situation, the obtained estimations are refinements of those obtained for the convex case.

In Section 4, by following some ideas from [11], we introduce a parametrized family of functions f_u and define a Tsallis type entropy (divergence) T_{f_u}. Next, we apply the results of the previous sections to T_{f_u}. Thus we get some estimations for a Tsallis type entropy from a general point of view. Here we also consider two cases: for convex f_u and for superquadratic f_u.

In summary, the present work uses both superquadratic functions and convex functions to produce some comparison theorems for generalized Csiszár and Tsallis type divergences.

2. Results for superquadratic functions

We extend definition (1) as follows.

Let $f:\mathbb{R}_+\to\mathbb{R}$ be a superquadratic function on $\mathbb{R}_+$, and $\mathbf{p}=(p_1,\dots,p_n)\in\mathbb{R}^n_{++}$, $\mathbf{q}=(q_1,\dots,q_n)\in\mathbb{R}^n_{+}$, $\mathbf{r}=(r_1,\dots,r_n)\in\mathbb{R}^n_{+}$. Then the *generalized Csiszár f-divergence* is defined by

$$C_f(\mathbf{p},\mathbf{q};\mathbf{r})=\sum_{i=1}^{n} r_i p_i f\left(\frac{q_i}{p_i}\right). \tag{8}$$

It is clear that $C_f(\mathbf{p},\mathbf{q};\mathbf{e})=C_f(\mathbf{p},\mathbf{q})$, where $\mathbf{e}=(1,\dots,1)\in\mathbb{R}^n$.

As usual, the notation $\langle\cdot,\cdot\rangle$ means the standard inner product on $\mathbb{R}^n$.

THEOREM 1. *Let $f:\mathbb{R}_+\to\mathbb{R}$ be a superquadratic function on $\mathbb{R}_+$. Let $\mathbf{p}=(p_1,\dots,p_n)\in\mathbb{R}^n_{++}$, $\mathbf{q}=(q_1,\dots,q_n)\in\mathbb{R}^n_{+}$, $\mathbf{r}_j=(r_{1j},\dots,r_{nj})\in\mathbb{R}^n_{+}$, $j=1,\dots,m$, $\mathbf{b}=(b_1,\dots,b_m)\in\mathbb{R}^m_{+}$. Then*

$$\sum_{j=1}^{m} b_j f\left(\frac{\langle\mathbf{q},\mathbf{r}_j\rangle}{\langle\mathbf{p},\mathbf{r}_j\rangle}\right)+\sum_{i=1}^{n}\sum_{j=1}^{m} b_j\frac{p_i r_{ij}}{\langle\mathbf{p},\mathbf{r}_j\rangle} f\left(\left|\frac{q_i}{p_i}-\frac{\langle\mathbf{q},\mathbf{r}_j\rangle}{\langle\mathbf{p},\mathbf{r}_j\rangle}\right|\right)\leqslant\sum_{i=1}^{n} a_i f\left(\frac{q_i}{p_i}\right), \tag{9}$$

where $a_i=\sum_{j=1}^{m} b_j\frac{p_i r_{ij}}{\langle\mathbf{p},\mathbf{r}_j\rangle}$, $i=1,\dots,n$.

If f is subquadratic then the inequality (9) is reversed.

Proof. We introduce the matrix

$$R = (r_{ij})_{i=1,\dots,n;j=1,\dots,m} \tag{10}$$

with columns $\mathbf{r}_j$ for $j = 1,\dots,m$. The following equality holds

$$\frac{\langle \mathbf{q}, \mathbf{r}_j \rangle}{\langle \mathbf{p}, \mathbf{r}_j \rangle} = \frac{p_1 r_{1j}}{\sum\limits_{i=1}^{n} p_i r_{ij}} \frac{q_1}{p_1} + \ldots + \frac{p_n r_{nj}}{\sum\limits_{i=1}^{n} p_i r_{ij}} \frac{q_n}{p_n} \quad \text{for } j = 1,\dots,m. \tag{11}$$

Hence the following identity is valid

$$\left[\frac{\langle \mathbf{q}, \mathbf{r}_1 \rangle}{\langle \mathbf{p}, \mathbf{r}_1 \rangle}, \dots, \frac{\langle \mathbf{q}, \mathbf{r}_m \rangle}{\langle \mathbf{p}, \mathbf{r}_m \rangle} \right] = \left[\frac{q_1}{p_1}, \dots, \frac{q_n}{p_n} \right] \begin{bmatrix} \frac{p_1 r_{11}}{\langle \mathbf{p}, \mathbf{r}_1 \rangle} & \cdots & \frac{p_1 r_{1m}}{\langle \mathbf{p}, \mathbf{r}_m \rangle} \\ \vdots & \ddots & \vdots \\ \frac{p_n r_{n1}}{\langle \mathbf{p}, \mathbf{r}_1 \rangle} & \cdots & \frac{p_n r_{nm}}{\langle \mathbf{p}, \mathbf{r}_m \rangle} \end{bmatrix}$$

with the column stochastic $n \times m$ matrix $S = (s_{ij})$, where $s_{ij} = \frac{p_i r_{ij}}{\langle \mathbf{p}, \mathbf{r}_j \rangle}$. That is, $\mathbf{y} = \mathbf{x}S$, where $\mathbf{x} = (x_1,\dots,x_n)$, $\mathbf{y} = (y_1,\dots,y_n)$, $x_i = \frac{q_i}{p_i}$ and $y_j = \frac{\langle \mathbf{q}, \mathbf{r}_j \rangle}{\langle \mathbf{p}, \mathbf{r}_j \rangle}$, $i = 1,\dots,n$, $j = 1,\dots,m$. Also, $\mathbf{a} = \mathbf{b}S^T$, where $a_i = \sum\limits_{j=1}^{m} b_j \frac{p_i r_{ij}}{\langle \mathbf{p}, \mathbf{r}_j \rangle}$, $i = 1,\dots,n$, is satisfied.

So, taking into account above and applying Theorem C, we get the inequality

$$\sum_{j=1}^{m} b_j f\left(\frac{\langle \mathbf{q}, \mathbf{r}_j \rangle}{\langle \mathbf{p}, \mathbf{r}_j \rangle} \right) + \sum_{j=1}^{m} b_j \sum_{i=1}^{n} \frac{p_i r_{ij}}{\langle \mathbf{p}, \mathbf{r}_j \rangle} f\left(\left| \frac{q_i}{p_i} - \frac{\langle \mathbf{q}, \mathbf{r}_j \rangle}{\langle \mathbf{p}, \mathbf{r}_j \rangle} \right| \right) \leqslant \sum_{i=1}^{n} a_i f\left(\frac{q_i}{p_i} \right), \tag{12}$$

what we need to prove. □

The forthcoming corollary includes the comparison result for two f-divergences (see (13) and (14)).

COROLLARY 1. *Let the assumptions of Theorem 1 be satisfied. Let us denote*

$$\widetilde{\mathbf{p}} = \mathbf{p}R, \quad \widetilde{\mathbf{q}} = \mathbf{q}R \quad \textit{and} \quad \mathbf{R} = (R_1,\dots,R_n),$$

where the $n \times m$ matrix $R = (r_{ij})$ is given by (10), and $R_i = \sum\limits_{j=1}^{m} r_{ij}$ is the ith row sum of R, $i = 1,\dots,n$. Then

$$C_f(\widetilde{\mathbf{p}}, \widetilde{\mathbf{q}}) + \sum_{i=1}^{n} \sum_{j=1}^{m} r_{ij} p_i f\left(\left| \frac{q_i}{p_i} - \frac{\langle \mathbf{q}, \mathbf{r}_j \rangle}{\langle \mathbf{p}, \mathbf{r}_j \rangle} \right| \right) \leqslant C_f(\mathbf{p}, \mathbf{q}; \mathbf{R}). \tag{13}$$

In particular, if the matrix R is row stochastic, then

$$C_f(\widetilde{\mathbf{p}}, \widetilde{\mathbf{q}}) + \sum_{i=1}^{n} \sum_{j=1}^{m} r_{ij} p_i f\left(\left| \frac{q_i}{p_i} - \frac{\langle \mathbf{q}, \mathbf{r}_j \rangle}{\langle \mathbf{p}, \mathbf{r}_j \rangle} \right| \right) \leqslant C_f(\mathbf{p}, \mathbf{q}). \tag{14}$$

If f is subquadratic then the inequalities (13)–(14) are reversed.

Proof. If we substitute $b_j := \langle \mathbf{p}, \mathbf{r}_j \rangle$ for $j = 1, \ldots, m$ in Theorem 1, then $a_i = p_i \sum_{j=1}^{m} r_{ij} = R_i p_i$ for $i = 1, \ldots, n$. In this case, inequality (9) takes the form

$$\sum_{j=1}^{m} \langle \mathbf{p}, \mathbf{r}_j \rangle f\left(\frac{\langle \mathbf{q}, \mathbf{r}_j \rangle}{\langle \mathbf{p}, \mathbf{r}_j \rangle}\right) + \sum_{i=1}^{n} \sum_{j=1}^{m} r_{ij} p_i f\left(\left|\frac{q_i}{p_i} - \frac{\langle \mathbf{q}, \mathbf{r}_j \rangle}{\langle \mathbf{p}, \mathbf{r}_j \rangle}\right|\right) \leqslant \sum_{i=1}^{n} \sum_{j=1}^{m} r_{ij} p_i f\left(\frac{q_i}{p_i}\right) \tag{15}$$

with

$$\sum_{j=1}^{m} r_{ij} p_i f\left(\frac{q_i}{p_i}\right) = p_i f\left(\frac{q_i}{p_i}\right) \sum_{j=1}^{m} r_{ij} = p_i f\left(\frac{q_i}{p_i}\right) R_i,$$

which is equivalent to (13).

If in addition the matrix R is row stochastic, then $\mathbf{R} = (1, \ldots, 1)$. For this reason (13) reduces to (14). □

COROLLARY 2. *Let $f : \mathbb{R}_+ \to \mathbb{R}$ be a superquadratic function on $\mathbb{R}_+$. Let $\mathbf{p} = (p_1, \ldots, p_n) \in \mathbb{R}^n_{++}$, $\mathbf{q} = (q_1, \ldots, q_n) \in \mathbb{R}^n_+$, $\mathbf{r} = (r_1, \ldots, r_n) \in \mathbb{R}^n_+$. Then*

$$\langle \mathbf{p}, \mathbf{r} \rangle f\left(\frac{\langle \mathbf{q}, \mathbf{r} \rangle}{\langle \mathbf{p}, \mathbf{r} \rangle}\right) + \sum_{i=1}^{n} r_i p_i f\left(\left|\frac{q_i}{p_i} - \frac{\langle \mathbf{q}, \mathbf{r} \rangle}{\langle \mathbf{p}, \mathbf{r} \rangle}\right|\right) \leqslant C_f(\mathbf{p}, \mathbf{q}; \mathbf{r}), \tag{16}$$

where $\langle \mathbf{p}, \mathbf{r} \rangle = \sum_{i=1}^{n} p_i r_i > 0$ and $\langle \mathbf{q}, \mathbf{r} \rangle = \sum_{i=1}^{n} q_i r_i$.

If f is subquadratic then the inequality (16) is reversed.

Proof. Taking $m = 1$ in Corollary 1 and $\mathbf{r}_1 = (r_1, \ldots, r_n)$, we obtain $R_i = r_i$ for $i = 1, \ldots, n$, and therefore inequality (15) becomes (16). □

As a special case of the previous result, choosing $\mathbf{r} = (1, \ldots, 1)$ we get the following result.

COROLLARY 3. *Let $f : \mathbb{R}_+ \to \mathbb{R}$ be a superquadratic function on $\mathbb{R}_+$ and $\mathbf{p} = (p_1, \ldots, p_n) \in \mathbb{R}^n_{++}$ and $\mathbf{q} = (q_1, \ldots, q_n) \in \mathbb{R}^n_+$. Then*

$$\sum_{i=1}^{n} p_i f\left(\frac{\sum_{i=1}^{n} q_i}{\sum_{i=1}^{n} p_i}\right) + \sum_{i=1}^{n} p_i f\left(\left|\frac{q_i}{p_i} - \frac{\sum_{i=1}^{n} q_i}{\sum_{i=1}^{n} p_i}\right|\right) \leqslant C_f(\mathbf{p}, \mathbf{q}). \tag{17}$$

If f is subquadratic then the inequality (17) is reversed.

REMARK 1. Inequality (16) includes Jensen inequality and Csiszár–Körner inequality for superquadratic functions as special cases. Namely, the substitution $\mathbf{p} = \mathbf{e} = (1, \ldots, 1)$ into (16) leads to Jensen inequality for superquadratic functions (see Theorem A with $x_i = q_i$, $a_i = r_i$, $i = 1, \ldots, n$). Likewise, the substitution $\mathbf{r} = (1, \ldots, 1)$ into (16) presents the Csiszár–Körner inequality for superquadratic functions (see Corollary 3, cf. also (2)).

3. Results for convex functions

In this section we compare two generalized Csiszár f-divergences with convex function f.

THEOREM 2. *Let* $f:\mathbb{R}_+\to\mathbb{R}$ *be a convex function on* $\mathbb{R}_+$. *Let* $\mathbf{p}=(p_1,\dots,p_n)\in\mathbb{R}^n_{++}$, $\mathbf{q}=(q_1,\dots,q_n)\in\mathbb{R}^n_+$, $\widetilde{\mathbf{p}}=(\widetilde{p}_1,\dots,\widetilde{p}_m)\in\mathbb{R}^m_{++}$, $\widetilde{\mathbf{q}}=(\widetilde{q}_1,\dots,\widetilde{q}_m)\in\mathbb{R}^m_+$, $\mathbf{c}=(c_1,\dots,c_n)\in\mathbb{R}^n_+$ *and* $\mathbf{d}=(d_1,\dots,d_m)\in\mathbb{R}^m_+$.

Let $R=(r_{ij})$ *be an* $n\times m$ *matrix with nonnegative entries such that*

$$\widetilde{\mathbf{p}}=\mathbf{p}R\,,\quad \widetilde{\mathbf{q}}=\mathbf{q}R \quad \text{and} \quad \mathbf{c}=\mathbf{d}R^T \tag{18}$$

is satisfied. Then

$$C_f\left(\widetilde{\mathbf{p}},\widetilde{\mathbf{q}};\mathbf{d}\right)\leqslant C_f\left(\mathbf{p},\mathbf{q};\mathbf{c}\right). \tag{19}$$

If f *is concave then the inequality* (19) *is reversed.*

Proof. According to (8) we have to prove that

$$\sum_{j=1}^{m} d_j\widetilde{p}_j f\left(\frac{\widetilde{q}_j}{\widetilde{p}_j}\right)\leqslant\sum_{i=1}^{n} c_i p_i f\left(\frac{q_i}{p_i}\right). \tag{20}$$

We denote $\mathbf{r}_j=(r_{1j},\dots,r_{nj})\in\mathbb{R}^n_+$. It follows from (18) that $\widetilde{p}_j=\langle\mathbf{p},\mathbf{r}_j\rangle=\sum\limits_{i=1}^{n}p_ir_{ij}$ and $\widetilde{q}_j=\langle\mathbf{q},\mathbf{r}_j\rangle=\sum\limits_{i=1}^{n}q_ir_{ij}$ for $j=1,\dots,m$, where $\langle\cdot,\cdot\rangle$ is the standard inner product in $\mathbb{R}^n$. Moreover, $c_i=\sum\limits_{j=1}^{m}d_jr_{ij}$ for $i=1,\dots,n$ (see (18)). Hence

$$a_i=\sum_{j=1}^{m} b_j\frac{p_ir_{ij}}{\langle\mathbf{p},\mathbf{r}_j\rangle}, \tag{21}$$

where $a_i=c_ip_i$ and $b_j=d_j\langle\mathbf{p},\mathbf{r}_j\rangle$ for $i=1,\dots,n,\ j=1,\dots,m$.

In a similar manner as in the proof of Theorem 1 we obtain the identity

$$\left[\frac{\langle\mathbf{q},\mathbf{r}_1\rangle}{\langle\mathbf{p},\mathbf{r}_1\rangle},\dots,\frac{\langle\mathbf{q},\mathbf{r}_m\rangle}{\langle\mathbf{p},\mathbf{r}_m\rangle}\right]=\left[\frac{q_1}{p_1},\dots,\frac{q_n}{p_n}\right]\begin{bmatrix}\frac{p_1r_{11}}{\langle\mathbf{p},\mathbf{r}_1\rangle} & \cdots & \frac{p_1r_{1m}}{\langle\mathbf{p},\mathbf{r}_m\rangle}\\ \vdots & \ddots & \vdots\\ \frac{p_nr_{n1}}{\langle\mathbf{p},\mathbf{r}_1\rangle} & \cdots & \frac{p_nr_{nm}}{\langle\mathbf{p},\mathbf{r}_m\rangle}\end{bmatrix}. \tag{22}$$

The matrix $S=(s_{ij})$, with $s_{ij}=\frac{p_ir_{ij}}{\langle\mathbf{p},\mathbf{r}_j\rangle}$, is column stochastic and with $\mathbf{x}=(x_1,\dots,x_n)$, $\mathbf{y}=(y_1,\dots,y_m)$, $x_i=\frac{q_i}{p_i}$ and $y_j=\frac{\langle\mathbf{q},\mathbf{r}_j\rangle}{\langle\mathbf{p},\mathbf{r}_j\rangle}$, $i=1,\dots,n,\ j=1,\dots,m$, satisfies condition $\mathbf{y}=\mathbf{x}S$ (see (22)). Furthermore, $\mathbf{a}=\mathbf{b}S^T$ for $\mathbf{a}=(a_1,\dots,a_n)$ and $\mathbf{b}=(b_1,\dots,b_m)$ (see (21)).

Now, applying Theorem B, we get

$$\sum_{j=1}^{m} b_jf\left(\frac{\langle\mathbf{q},\mathbf{r}_j\rangle}{\langle\mathbf{p},\mathbf{r}_j\rangle}\right)\leqslant\sum_{i=1}^{n} a_if\left(\frac{q_i}{p_i}\right), \tag{23}$$

where $a_i = c_i p_i$ and $b_j = d_j \langle \mathbf{p}, \mathbf{r}_j \rangle$, which is equivalent to (20). This completes the proof. □

REMARK 2. In contrast to Theorem B, we do not assume in Theorem 2 that the matrix R is column stochastic. Instead of that, the matrix $S = \left(\frac{p_i r_{ij}}{\langle \mathbf{p}, \mathbf{r}_j \rangle} \right)$ in equation (22) is column stochastic.

COROLLARY 4. *Let the assumptions of Theorem 2 be satisfied. Let* $\mathbf{R} = (R_1, \ldots, R_n)$, *where* $R_i = \sum_{j=1}^{m} r_{ij}$, $i = 1, \ldots, n$, *is the* i*th row sum of* R. *Then*

$$C_f(\widetilde{\mathbf{p}}, \widetilde{\mathbf{q}}) \leqslant C_f(\mathbf{p}, \mathbf{q}; \mathbf{R}). \tag{24}$$

In particular, if the matrix R *is row stochastic, then*

$$C_f(\widetilde{\mathbf{p}}, \widetilde{\mathbf{q}}) \leqslant C_f(\mathbf{p}, \mathbf{q}). \tag{25}$$

If f *is concave then the inequalities* (24) *and* (25) *are reversed.*

Proof. By making use of Theorem 2, we take $\mathbf{d} = (d_1, \ldots, d_m) = (1, \ldots, 1)$, i.e., $d_j = 1$ for $j = 1, \ldots, m$. Then $c_i = \sum_{j=1}^{m} r_{ij} = R_i$ for $i = 1, \ldots, n$ (see (18)). Therefore inequalities (19)–(20) imply

$$\sum_{j=1}^{m} \langle \mathbf{p}, \mathbf{r}_j \rangle f\left(\frac{\langle \mathbf{q}, \mathbf{r}_j \rangle}{\langle \mathbf{p}, \mathbf{r}_j \rangle} \right) \leqslant \sum_{i=1}^{n} R_i p_i f\left(\frac{q_i}{p_i} \right), \tag{26}$$

which is equivalent to (24).

If in addition the matrix R is row stochastic, then $\mathbf{R} = (1, \ldots, 1) \in \mathbb{R}^n$. In this case (24) becomes (25). □

COROLLARY 5. *Let* $f : \mathbb{R}_+ \to \mathbb{R}$ *be a convex function on* $\mathbb{R}_+$. *Let* $\mathbf{p} = (p_1, \ldots, p_n) \in \mathbb{R}^n_{++}$, $\mathbf{q} = (q_1, \ldots, q_n) \in \mathbb{R}^n_+$, $\mathbf{r} = (r_1, \ldots, r_n) \in \mathbb{R}^n_+$. *Then*

$$\langle \mathbf{p}, \mathbf{r} \rangle f\left(\frac{\langle \mathbf{q}, \mathbf{r} \rangle}{\langle \mathbf{p}, \mathbf{r} \rangle} \right) \leqslant C_f(\mathbf{p}, \mathbf{q}; \mathbf{r}), \tag{27}$$

where $\langle \mathbf{p}, \mathbf{r} \rangle = \sum_{i=1}^{n} p_i r_i > 0$ *and* $\langle \mathbf{q}, \mathbf{r} \rangle = \sum_{i=1}^{n} q_i r_i$.

If f *is concave then the inequality* (27) *is reversed.*

Proof. It is sufficient to apply Corollary 4 for $m = 1$, $\mathbf{r}_1 = (r_1, \ldots, r_n)$ and $R = \mathbf{r}_1^T$. Then $R_i = r_i$ for $i = 1, \ldots, n$. So, inequality (26) leads to (27). □

We finish this section with some concluding remarks.

REMARK 3. Inequality (27) includes the classical Csiszár–Körner inequality (see (2)) and Jensen inequality for convex functions as special cases. To see the former, use the substitution $\mathbf{r} = (1,\ldots,1)$ in (27). For the latter case, set $\mathbf{r} = (1,\ldots,1)$, $\langle \mathbf{p},\mathbf{r}\rangle = \sum_{i=1}^{n} p_i > 0$, $x_i = \frac{q_i}{p_i}$, $\lambda_i = \frac{p_i}{\sum_{i=1}^{n} p_i}$, $i = 1,\ldots,n$.

REMARK 4. It follows from [1, Lemma 2.1] that nonnegative superquadratic functions must be convex. Thus the results demonstrated in Section 2, when applied to nonnegative superquadratic functions, give refinements of the corresponding inequalities obtained for nonnegative convex functions in Section 3. For instance, compare the Jensen inequalities for both convex functions and nonnegative superquadratic functions (see (5)). Similarly, see Csiszár–Körner type inequalities (2) and (17).

4. Results for Tsallis type entropy

Throughout this section we consider a one parameter family of functions

$$f_u(t) : [0,\infty) \to \mathbb{R}, \quad u \in [0,\infty).$$

For a given f_u, a function $g_u : [0,\infty) \to \mathbb{R}$ is defined by

$$g_u(t) = \frac{f_u(t) - f_0(t)}{u} \quad \text{for } u > 0, \tag{28}$$

$$g_0(t) = \lim_{u\to 0^+} \frac{f_u(t) - f_0(t)}{u} \tag{29}$$

(see [11, p. 854]).

EXAMPLE 1. Let $f_u : [0,\infty)$ be defined by $f_u(t) = -t^u$, $u \geqslant 0$. Then $g_u(t) = -\frac{t^u-1}{u}$ for $u > 0$, and $g_0(t) = \lim_{u\to 0^+}\left(-\frac{t^u-1}{u}\right) = -\ln t$.

LEMMA 1. *Let f_0 be a constant function on $\mathbb{R}_+$. Let $u > 0$.*

(i) If f_u is superquadratic (resp. subquadratic) on $[0,\infty)$ and $f_0 \geqslant 0$ (resp. $f_0 \leqslant 0$) then g_u is superquadratic (resp. subquadratic) on $[0,\infty)$.

(ii) If f_u is convex (resp. concave) on $[0,\infty)$ then g_u is convex (resp. concave) on $[0,\infty)$.

Proof. (i). Fix any $x \geqslant 0$. Since f_u is superquadratic, then by definition there exists a real constant $c_u(x)$ such that

$$f_u(y) - f_u(x) \geqslant c_u(x)(y-x) + f_u(|y-x|) \quad \text{for all } y \geqslant 0.$$

Simultaneously $f_0(x) = f_0(y)$, because f_0 is a constant function. Therefore, by the nonnegativity of f_0, we get

$$\begin{aligned}(f_u(y)-f_0(y))-(f_u(x)-f_0(x)) &\geqslant c_u(x)(y-x)+f_u(|y-x|)\\ &\geqslant c_u(x)(y-x)+f_u(|y-x|)-f_0(|y-x|) \quad \text{for all } y \geqslant 0.\end{aligned}$$

Hence

$$\frac{f_u(y)-f_0(y)}{u}-\frac{f_u(x)-f_0(x)}{u} \geqslant \frac{c_u(x)}{u}(y-x)+\frac{f_u(|y-x|)-f_0(|y-x|)}{u},$$

which means

$$g_u(y)-g_u(x) \geqslant \frac{c_u(x)}{u}(y-x)+g_u(|y-x|) \quad \text{for all } y \geqslant 0.$$

Thus g_u is superquadratic (with the constant $\frac{c_u(x)}{u}$), as claimed.

The case when f_u is subquadratic and $f_0 \leqslant 0$ requires an analogous proof, and therefore is left for the reader.

(ii). The proof of the part (ii) is straightforward, and therefore omitted. □

Let $\mathbf{p} = (p_1,\dots,p_n) \in \mathbb{R}^n_{++}$, $\mathbf{q} = (q_1,\dots,q_n) \in \mathbb{R}^n_{+}$, $\mathbf{r} = (r_1,\dots,r_n) \in \mathbb{R}^n_{+}$ be n-tuples of nonnegative numbers. According to (8), the *generalized Csiszár f_u-divergence* is given by

$$C_{f_u}(\mathbf{p},\mathbf{q};\mathbf{r}) = \sum_{i=1}^{n} r_i p_i f_u\left(\frac{q_i}{p_i}\right). \tag{30}$$

We now update (30) by providing the *generalized Tsallis f_u-divergence* defined by

$$T_{f_u}(\mathbf{p},\mathbf{q};\mathbf{r}) = C_{g_u}(\mathbf{p},\mathbf{q};\mathbf{r}) = \sum_{i=1}^{n} r_i p_i g_u\left(\frac{q_i}{p_i}\right) = \sum_{i=1}^{n} r_i p_i \frac{f_u\left(\frac{q_i}{p_i}\right)-f_0\left(\frac{q_i}{p_i}\right)}{u}. \tag{31}$$

For example, for $f_u(t) = -t^u$ for $u \geqslant 0$, we get $g_u(t) = -\frac{t^u-1}{u}$ for $u > 0$, and $g_0(t) = \lim_{u\to 0^+}\left(-\frac{t^u-1}{u}\right) = -\ln t$. Therefore the generalized Tsallis f_u-divergence reduces to the *generalized Tsallis relative entropy* as follows:

$$T_u(\mathbf{p},\mathbf{q};\mathbf{r}) = -\sum_{i=1}^{n} r_i p_i \frac{\left(\frac{q_i}{p_i}\right)^u-1}{u}. \tag{32}$$

The case $\mathbf{r} = (1,\dots,1) \in \mathbb{R}^n$ of (32) gives the *Tsallis relative entropy* [6, p. 12]:

$$T_u(\mathbf{p},\mathbf{q}) = -\sum_{i=1}^{n} p_i \frac{\left(\frac{q_i}{p_i}\right)^u-1}{u}. \tag{33}$$

In the sequel we utilize the results of the previous sections for the Tsallis type divergences (entropies).

THEOREM 3. *Let $f_u : \mathbb{R}_+ \to \mathbb{R}$ be a convex function on $\mathbb{R}_+$ for some $u > 0$. Let $f_0 : \mathbb{R}_+ \to \mathbb{R}$ be a constant function. Let $\mathbf{p} = (p_1, \ldots, p_n) \in \mathbb{R}^n_{++}$, $\mathbf{q} = (q_1, \ldots, q_n) \in \mathbb{R}^n_+$, $\widetilde{\mathbf{p}} = (\widetilde{p}_1, \ldots, \widetilde{p}_m) \in \mathbb{R}^m_{++}$, $\widetilde{\mathbf{q}} = (\widetilde{q}_1, \ldots, \widetilde{q}_m) \in \mathbb{R}^m_+$, $\mathbf{c} = (c_1, \ldots, c_n) \in \mathbb{R}^n_+$ and $\mathbf{d} = (d_1, \ldots, d_m) \in \mathbb{R}^m_+$.*

Let $R = (r_{ij})$ be an $n \times m$ matrix with nonnegative entries such that

$$\widetilde{\mathbf{p}} = \mathbf{p}R\,, \quad \widetilde{\mathbf{q}} = \mathbf{q}R \quad \text{and} \quad \mathbf{c} = \mathbf{d}R^T \tag{34}$$

is satisfied. Then

$$T_{f_u}(\widetilde{\mathbf{p}}, \widetilde{\mathbf{q}}; \mathbf{d}) \leqslant T_{f_u}(\mathbf{p}, \mathbf{q}; \mathbf{c})\,. \tag{35}$$

If f_u is concave then the inequality (35) is reversed.

Proof. Since the function g_u is convex on $I = [0, \infty)$ (see Lemma 1), applying Theorem 2 we get the inequality

$$T_{f_u}(\widetilde{\mathbf{p}}, \widetilde{\mathbf{q}}; \mathbf{d}) = C_{g_u}(\widetilde{\mathbf{p}}, \widetilde{\mathbf{q}}; \mathbf{d}) \leqslant C_{g_u}(\mathbf{p}, \mathbf{q}; \mathbf{c}) = T_{f_u}(\mathbf{p}, \mathbf{q}; \mathbf{c})\,, \tag{36}$$

what we need to prove.

The case of concave f_u is analogous. □

COROLLARY 6. *Let $\mathbf{p} = (p_1, \ldots, p_n) \in \mathbb{R}^n_{++}$, $\mathbf{q} = (q_1, \ldots, q_n) \in \mathbb{R}^n_+$, $\widetilde{\mathbf{p}} = (\widetilde{p}_1, \ldots, \widetilde{p}_m) \in \mathbb{R}^m_{++}$, $\widetilde{\mathbf{q}} = (\widetilde{q}_1, \ldots, \widetilde{q}_m) \in \mathbb{R}^m_+$, $\mathbf{c} = (c_1, \ldots, c_n) \in \mathbb{R}^n_+$ and $\mathbf{d} = (d_1, \ldots, d_m) \in \mathbb{R}^m_+$.*

Let $R = (r_{ij})$ be an $n \times m$ matrix with nonnegative entries and

$$\widetilde{\mathbf{p}} = \mathbf{p}R\,, \quad \widetilde{\mathbf{q}} = \mathbf{q}R \quad \text{and} \quad \mathbf{c} = \mathbf{d}R^T \tag{37}$$

is satisfied. Then for $u \geqslant 1$

$$T_u(\widetilde{\mathbf{p}}, \widetilde{\mathbf{q}}; \mathbf{d}) \geqslant T_u(\mathbf{p}, \mathbf{q}; \mathbf{c})\,. \tag{38}$$

If $0 < u < 1$ then the inequality (38) is reversed.

Proof. We consider $f_u(t) = -t^u$ for $t \geqslant 0$ with $u \geqslant 1$. Then f_u is concave. We also take $f_0(t) = -1$ for $t \geqslant 0$. On account of Theorem 3 we obtain

$$T_u(\widetilde{\mathbf{p}}, \widetilde{\mathbf{q}}; \mathbf{d}) = T_{f_u}(\widetilde{\mathbf{p}}, \widetilde{\mathbf{q}}; \mathbf{d}) \geqslant T_{f_u}(\mathbf{p}, \mathbf{q}; \mathbf{c}) = T_u(\mathbf{p}, \mathbf{q}; \mathbf{c})\,, \tag{39}$$

completing the proof of (38).

If $0 < u < 1$ then the function $f_u(t) = -t^u$, $t \geqslant 0$, is convex, and therefore the inequality (38) is reversed (see Theorem 3). □

We now study Tsallis type divergences induced by superquadratic functions.

THEOREM 4. *Let $f_u : \mathbb{R}_+ \to \mathbb{R}$ be a superquadratic function on $\mathbb{R}_+$ for some $u > 0$. Let $f_0 : \mathbb{R}_+ \to \mathbb{R}$ be a nonnegative constant function. Let $\mathbf{p} = (p_1, \ldots, p_n) \in \mathbb{R}^n_{++}$, $\mathbf{q} = (q_1, \ldots, q_n) \in \mathbb{R}^n_+$, $\widetilde{\mathbf{p}} = (\widetilde{p}_1, \ldots, \widetilde{p}_m) \in \mathbb{R}^m_{++}$, $\widetilde{\mathbf{q}} = (\widetilde{q}_1, \ldots, \widetilde{q}_m) \in \mathbb{R}^m_+$, $\mathbf{c} = (c_1, \ldots, c_n) \in \mathbb{R}^n_+$ and $\mathbf{d} = (d_1, \ldots, d_m) \in \mathbb{R}^m_+$.*

Let $R=(r_{ij})$ be an $n\times m$ matrix with nonnegative entries such that

$$\widetilde{\mathbf{p}}=\mathbf{p}R\,,\quad \widetilde{\mathbf{q}}=\mathbf{q}R \quad \text{and} \quad \mathbf{c}=\mathbf{d}R^T \tag{40}$$

is satisfied. Then

$$T_{f_u}(\widetilde{\mathbf{p}},\widetilde{\mathbf{q}};\mathbf{d})+\sum_{i=1}^{n}\sum_{j=1}^{m}d_jp_ir_{ij}\,g_u\left(\left|\frac{q_i}{p_i}-\frac{\widetilde{q}_j}{\widetilde{p}_j}\right|\right)\leqslant T_{f_u}(\mathbf{p},\mathbf{q};\mathbf{c})\,. \tag{41}$$

If f_u is subquadratic on $\mathbb{R}_+$ and f_0 is nonpositive constant function on $\mathbb{R}_+$ then the inequality (41) is reversed.

Proof. We denote $\mathbf{r}_j=(r_{1j},\dots,r_{nj})$. By (40), $\widetilde{p}_j=\langle\mathbf{p},\mathbf{r}_j\rangle=\sum_{i=1}^{n}p_ir_{ij}$ and $\widetilde{q}_j=\langle\mathbf{q},\mathbf{r}_j\rangle=\sum_{i=1}^{n}q_ir_{ij}$ for $j=1,\dots,m$. Additionally, $c_i=\sum_{j=1}^{m}d_jr_{ij}$ for $i=1,\dots,n$. By denoting $a_i=c_ip_i$ and $b_j=d_j\langle\mathbf{p},\mathbf{r}_j\rangle$ for $i=1,\dots,n$, $j=1,\dots,m$, we get $a_i=\sum_{j=1}^{m}b_j\frac{p_ir_{ij}}{\langle\mathbf{p},\mathbf{r}_j\rangle}$.

Let us consider the function f_u defined on $\mathbb{R}_+$. Since f_u is superquadratic (see [1, p. 1]), then, by Lemma 1, g_u is also superquadratic. Applying Theorem 1 we get

$$\sum_{j=1}^{m}b_jg_u\left(\frac{\langle\mathbf{q},\mathbf{r}_j\rangle}{\langle\mathbf{p},\mathbf{r}_j\rangle}\right)+\sum_{i=1}^{n}\sum_{j=1}^{m}b_j\frac{p_ir_{ij}}{\langle\mathbf{p},\mathbf{r}_j\rangle}g_u\left(\left|\frac{q_i}{p_i}-\frac{\langle\mathbf{q},\mathbf{r}_j\rangle}{\langle\mathbf{p},\mathbf{r}_j\rangle}\right|\right)\leqslant\sum_{i=1}^{n}a_ig_u\left(\frac{q_i}{p_i}\right). \tag{42}$$

More precisely, we get

$$\sum_{j=1}^{m}d_j\widetilde{p}_jg_u\left(\frac{\widetilde{q}_j}{\widetilde{p}_j}\right)+\sum_{i=1}^{n}\sum_{j=1}^{m}d_jp_ir_{ij}\,g_u\left(\left|\frac{q_i}{p_i}-\frac{\widetilde{q}_j}{\widetilde{p}_j}\right|\right)\leqslant\sum_{i=1}^{n}c_ip_ig_u\left(\frac{q_i}{p_i}\right), \tag{43}$$

that is

$$C_{g_u}(\widetilde{\mathbf{p}},\widetilde{\mathbf{q}};\mathbf{d})+\sum_{i=1}^{n}\sum_{j=1}^{m}d_jp_ir_{ij}\,g_u\left(\left|\frac{q_i}{p_i}-\frac{\widetilde{q}_j}{\widetilde{p}_j}\right|\right)\leqslant C_{g_u}(\mathbf{p},\mathbf{q};\mathbf{c})\,. \tag{44}$$

Thus we establish the inequality (41). □

COROLLARY 7. *Let $\mathbf{p}=(p_1,\dots,p_n)\in\mathbb{R}^n_{++}$, $\mathbf{q}=(q_1,\dots,q_n)\in\mathbb{R}^n_{+}$, $\widetilde{\mathbf{p}}=(\widetilde{p}_1,\dots,\widetilde{p}_m)\in\mathbb{R}^m_{++}$, $\widetilde{\mathbf{q}}=(\widetilde{q}_1,\dots,\widetilde{q}_m)\in\mathbb{R}^m_{+}$, $\mathbf{c}=(c_1,\dots,c_n)\in\mathbb{R}^n_{+}$ and $\mathbf{d}=(d_1,\dots,d_m)\in\mathbb{R}^m_{+}$.*

Let $R=(r_{ij})$ be an $n\times m$ matrix with nonnegative entries such that

$$\widetilde{\mathbf{p}}=\mathbf{p}R\,,\quad \widetilde{\mathbf{q}}=\mathbf{q}R \quad \text{and} \quad \mathbf{c}=\mathbf{d}R^T \tag{45}$$

is satisfied. Then for $u\geqslant 2$

$$T_u(\widetilde{\mathbf{p}},\widetilde{\mathbf{q}};\mathbf{d})-\sum_{i=1}^{n}\sum_{j=1}^{m}d_jp_ir_{ij}\,\frac{\left|\frac{q_i}{p_i}-\frac{\widetilde{q}_j}{\widetilde{p}_j}\right|^u-1}{u}\geqslant T_u(\mathbf{p},\mathbf{q};\mathbf{c})\,. \tag{46}$$

Proof. We set $f_u(t) = -t^u$, $f_0(t) = -1$ and $g_u(t) = -\frac{t^u-1}{u}$ for $t \geqslant 0$ with $u \geqslant 2$. Then f_u is subquadratic on $\mathbb{R}_+$. By virtue of Theorem 4 we get

$$\begin{aligned} & T_u(\widetilde{\mathbf{p}}, \widetilde{\mathbf{q}}; \mathbf{d}) - \sum_{i=1}^{n} \sum_{j=1}^{m} d_j p_i r_{ij} \frac{\left| \frac{q_i}{p_i} - \frac{\widetilde{q}_j}{\widetilde{p}_j} \right|^u - 1}{u} \\ & = T_{f_u}(\widetilde{\mathbf{p}}, \widetilde{\mathbf{q}}; \mathbf{d}) - \sum_{i=1}^{n} \sum_{j=1}^{m} d_j p_i r_{ij} \frac{\left| \frac{q_i}{p_i} - \frac{\widetilde{q}_j}{\widetilde{p}_j} \right|^u - 1}{u}, \\ & \geqslant T_{f_u}(\mathbf{p}, \mathbf{q}; \mathbf{c}) = T_u(\mathbf{p}, \mathbf{q}; \mathbf{c}), \end{aligned} \tag{47}$$

as required. □

Acknowledgements. The authors would like to thank anonymous referee for his valuable remarks and suggestions that improved an earlier version of the manuscript.

REFERENCES

[1] S. ABRAMOVICH, G. JAMESON AND G. SINNAMON, *Inequalities for averages of convex and superquadratic functions*, J. Inequal. Pure Appl. Math. **5** (2004), Art. 91.
[2] S. ABRAMOVICH, G. JAMESON AND G. SINNAMON, *Refining Jensen's inequality*, Bull. Math. Soc. Sci. Math. Roumanie (N. S.), **47**, 95 (2004), 3–14.
[3] S. ABRAMOVICH, S. IVELIĆ AND J. E. PEČARIĆ, *Improvement of Jensen-Steffensen's inequality for superquadratic functions*, Banach J. Math. Anal. **4** (2010), 159–169.
[4] J. C. BAEZ, *Rényi entropy and free energy*, J. Stat. Mech. Theory E. (2011), arXiv:1102.2098.
[5] J. C. BAEZ AND B. S. POLLARD, *Relative entropy in biological systems*, Entropy **18**, 2 46 (2016), doi:10.3390/e18020046.
[6] G. E. CROOKS, *On measures of entropy and information*, Tech. Note **009**, v0.5, (2016), 2016-08-16, http://threeplusone.com/info.
[7] G. E. CROOKS AND D. A. SIVAK, *Measures of trajectory ensemble disparsity in nonequilibrium statistical dynamics*, J. Stat. Mech. Theory E., (2011) P06003.
[8] I. CSISZÁR, *Information-type measures of differences of probability distributions and indirect observations*, Studia Sci. Math. Hung. **2** (1967), 299–318.
[9] I. CSISZÁR AND J. KÖRNER, *Information Theory: Coding Theorems for Discrete Memory-less Systems*, Academic Press, New York, 1981.
[10] S. S. DRAGOMIR, *Upper and lower bounds for Csiszár f-divergence in terms of the Kullback-Leibler distance and applications*, in Inequalities for the Csiszár f-divergence in Information Theory, ed. S. S. Dragomir, 2000, http://rgmia.vu.edu.au/monographs/csiszar.htm.
[11] P. A. KLUZA AND M. NIEZGODA, *Inequalities for relative operator entropies*, Electron. J. Linear Algebra **27** (2014), 851–864.
[12] P. KLUZA AND M. NIEZGODA, *Generalizations of Crooks and Lin's results on Jeffreys–Csiszár and Jensen–Csiszár f-divergences*, Physica A, **463** (2016), 383–393.
[13] J. LIN, *Divergence measures based on the Shannon entropy*, IEEE Trans. Inf. Th. **37**, 1 (1991), 145–151.
[14] F.-C. MITROI-SYMEONIDIS AND N. MINCULETE, *On the Jensen functional and superquadraticity*, Aequationes Math. **90**, 4 (2016), 705–718.
[15] F.-C. MITROI-SYMEONIDIS AND N. MINCULETE, *On the Jensen functional and strong convexity*, Bull. Malays. Math. Sci. Soc., 2016, doi:10.1007/s40840-015-0293- z.
[16] M. NIEZGODA, *Shannon like inequalities for f-connections of positive linear maps and positive operators*, Linear Algebra Appl. **481** (2015), 186–201.
[17] M. NIEZGODA, *Vector joint majorization and generalization of Csiszár-Körner's inequality for f-divergence*, Discrete Appl. Math. **198** (2016), 195–205.

[18] M. NIEZGODA, *Inequalities for H-invex functions with applications for uniformly convex and superquadratic functions*, Filomat **31**, 15 (2017), 4781–4794.

[19] C. TSALLIS, *Possible generalization of Bolzmann-Gibbs statistics*, J. Stat. Phys. **52** (1988,) 479–487.

[20] S. SHERMAN, *On a theorem of Hardy, Littlewood, Pólya, and Blackwell*, Proc. Nat. Acad. Sci. USA, **37** (1957), 826–831.

(Received May 11, 2017)

Paweł A. Kluza
Department of Applied Mathematics and Computer Science
University of Life Sciences in Lublin
Akademicka 13, 20-950 Lublin, Poland
e-mail: pawel.kluza@up.lublin.pl

Marek Niezgoda
Department of Applied Mathematics and Computer Science
University of Life Sciences in Lublin
Akademicka 13, 20-950 Lublin, Poland
e-mail: marek.niezgoda@up.lublin.pl