Mathematical
Inequalities
& Applications
Volume 22, Number 4 (2019), 1427–1449 doi:10.7153/mia-2019-22-100

MONOTONICITY OF THE JENSEN FUNCTIONAL FOR $f-$ DIVERGENCES WITH APPLICATIONS TO THE ZIPF–MANDELBROT LAW

NEDA LOVRIČEVIĆ*, ĐILDA PEČARIĆ AND JOSIP PEČARIĆ

(*Communicated by J. Jakšetić*)

Abstract. The Jensen functional in its discrete form is brought in relation to the Csiszár divergence functional via its monotonicity property. Thus deduced general results branch into specific forms for some of the well known $f-$ divergences, e.g. the Kullback-Leibler divergence, the Hellinger distance, the Bhattacharyya coefficient, χ^2- divergence, total variation distance. Obtained comparative inequalities are also interpreted in the environment of the Zipf and the Zipf-Mandelbrot law.

1. Introduction and preliminaries

In the monograph [13, p. 717] J. E. Pečarić investigated the monotonicity property of the Jensen functional which is derived by subtracting the left from the right hand side of the discrete Jensen inequality. Using Jensen's inequality and its reverse he proved that for a convex function f on an interval $I \subset \mathbb{R}$ and for $\mathbf{x} = (x_1, \ldots, x_n) \in I^n$, $n \geqslant 2$, and positive n-tuples $\mathbf{p}$ and $\mathbf{q}$ such that $\mathbf{p} \geqslant \mathbf{q}$, (*i.e.* $p_i \geqslant q_i$, $i = 1, \ldots, n$) the following relation holds:

$$\sum_{i=1}^{n} p_i f(x_i) - P_n f\left(\frac{1}{P_n}\sum_{i=1}^{n} p_i x_i\right) \geqslant \sum_{i=1}^{n} q_i f(x_i) - Q_n f\left(\frac{1}{Q_n}\sum_{i=1}^{n} q_i x_i\right) \geqslant 0, \tag{1}$$

where $P_n = \sum_{i=1}^{n} p_i$, $Q_n = \sum_{i=1}^{n} q_i$.

Inequalities (1) serve as a starting point when presenting new comparative inequalities for some of the most common $f-$ divergences which measure the distance between two probability distributions. Although $f-$ divergences were studied by several mathematicians, here we focus on Csiszár's approach [1, 2]. The Csiszár divergence functional is defined by

$$D_f(\mathbf{r}, \mathbf{s}) = \sum_{i=1}^{n} s_i f\left(\frac{r_i}{s_i}\right), \tag{2}$$

Mathematics subject classification (2010): 94A15, 94A17, 26D15, 26A51.
Keywords and phrases: Monotonicity property, Jensen's functional, Zipf's and Zipf-Mandelbrot's law, Csiszár divergence functional, $f-$divergences.
This publication was supported by the Ministry of Education and Science of the Russian Federation (the Agreement number No.02.a03.21.0008).
* Corresponding author.

where $\mathbf{r} = (r_1, \dots, r_n)$ and $\mathbf{s} = (s_1, \dots, s_n)$ are positive real $n-$tuples and $f\colon \langle 0, \infty \rangle \to \mathbb{R}$ is a convex function.
The Csiszár divergence functional (2) may also be defined for nonnegative real $n-$tuples $\mathbf{r}$ and $\mathbf{s}$ with undefined expressions interpreted as

$$f(0) := \lim_{t \to 0+} f(t); \qquad 0f\left(\frac{0}{0}\right) := 0; \qquad 0f\left(\frac{a}{0}\right) := \lim_{t \to 0+} tf\left(\frac{a}{t}\right), \qquad a > 0,$$

or even in a more general setting where $f\colon I \to \mathbb{R}$, $I \subseteq \mathbb{R}$. Still, in all of the results in the sequel we focus on positive real $n-$tuples $\mathbf{r}$ and $\mathbf{s}$ in definition (2).
Furthermore, Csiszár divergence functional (2) can be interpreted for special choices of the kernel function f. Thus in the case of positive probability distributions $\mathbf{r}$ and $\mathbf{s}$, that is $r_i, s_i \in \langle 0, 1]$, for $i = 1, \dots, n$ with $\sum_{i=1}^n r_i = \sum_{i=1}^n s_i = 1$ it assumes special forms which we recognize as some well known divergences.

The Kullback-Leibler divergence (see [5], [6], [7]) for positive probability distributions $\mathbf{r} = (r_1, \dots, r_n)$ and $\mathbf{s} = (s_1, \dots, s_n)$ is defined by

$$KL(\mathbf{r}, \mathbf{s}) := \sum_{i=1}^n r_i \log \frac{r_i}{s_i}. \tag{3}$$

In the sequel we analyze results for the logarithm function for different positive bases and distinguish the cases for the bases greater and less than 1.

The Hellinger distance between positive probability distributions $\mathbf{r} = (r_1, \dots, r_n)$ and $\mathbf{s} = (s_1, \dots, s_n)$ is defined by

$$h(\mathbf{r}, \mathbf{s}) := \frac{1}{\sqrt{2}} \sqrt{\sum_{i=1}^n (\sqrt{r_i} - \sqrt{s_i})^2}. \tag{4}$$

The Hellinger distance is a metric and is often used in its squared form, *i.e.* as $h^2(\mathbf{r}, \mathbf{s}) := \frac{1}{2} \sum_{i=1}^n \left(\sqrt{r_i} - \sqrt{s_i}\right)^2$.

The Bhattacharyya coefficient is an approximate measure of the amount of overlapping between two positive probability distributions and as such can be used to determine their relative closeness. It is defined as

$$B(\mathbf{r}, \mathbf{s}) := \sum_{i=1}^n \sqrt{r_i s_i}. \tag{5}$$

Furthermore, χ^2 (chi-square) divergence is defined as

$$\chi^2(\mathbf{r}, \mathbf{s}) := \sum_{i=1}^n \frac{(r_i - s_i)^2}{s_i} \tag{6}$$

and the total variation distance or statistical distance is given by

$$V(\mathbf{r}, \mathbf{s}) := \sum_{i=1}^n |r_i - s_i|. \tag{7}$$

One of the overviews of $f-$ divergences is given e.g. in [3].

Generalized comparative inequalities on $f-$ divergences that we obtained here are also observed in the context of the Zipf-Mandelbrot law and then specified for the Zipf law.
Philologist George Kingsley Zipf (1902-1950) studied statistical occurrences in different languages and concluded that, if words of a language were sorted in the order of decreasing frequencies of usage, a word's frequency was inversely proportional to its rank or sequence number in the list [12]. Thus the most frequent word will occur approximately twice as often as the second most frequent word, three times as often as the third most frequent word etc. It was one of the first academic studies of word frequency and was originally prescribed only for linguistics. It was only later that many other disciplines took credit of it: the Pareto law in economy reveals another aspect of it and the "Zipfian distribution" is present in other fields as well: information science, bibliometrics, social sciences etc.
Benoit Mandelbrot (1924-2010) generalized the Zipf law in 1966 [10, 11] and gave its improvement for the count of the low-rank words [14]. It is also used in information sciences for the purpose of indexing [4, 16], in ecological field studies [15] and has its role in art when determining the esthetics criteria in music [9]. The Zipf-Mandelbrot law is a discrete probability distribution and is defined by the following probability mass function:

$$f(i;N,v,w)=\frac{1}{(i+w)^v H_{N,v,w}},\quad i=1,\ldots,N, \tag{8}$$

where

$$H_{N,v,w}=\sum_{k=1}^{N}\frac{1}{(k+w)^v} \tag{9}$$

is a generalization of a harmonic number and $N\in\{1,2,\ldots\}$, $v>0$ and $w\in[0,\infty)$ are parameters.
For finite N and for $w=0$ the Zipf-Mandelbrot law is simply called the Zipf law. (In particular, if we observe the infinite N and $w=0$ we actually have the Zeta distribution.)
According to the expressions above, the probability mass function referring to the Zipf law is

$$f(i;N,v)=\frac{1}{i^v\cdot H_{N,v}},\quad \text{where}\quad H_{N,v}=\sum_{k=1}^{N}\frac{1}{k^v}, \tag{10}$$

that is, out of population of N elements the frequency of elements of rank i is $f(i;N,v)$, where v is the value of the exponent that characterizes the distribution.

Our paper provides the main inequalities via (1) for the Csiszár divergence functional (2) and its derived special divergences (in Section 2) and the further analysis of the results obtained therein in the light of the Zipf-Mandelbrot law and the Zipf law (in Section 3). Furthermore, results that are presented here generalize for the most part the results previously obtained in [8]. Therefore the bounds for the divergences provided in [8] become the special cases of the more general results obtained here, which is repeatedly accentuated throughout the paper.

2. Main results on $f-$ divergences

Monotonicity property (1) of the discrete Jensen functional and the Csiszár divergence functional (2) are integrated in the following two theorems.

THEOREM 1. *Let $f\colon \langle 0,\infty\rangle \to \mathbb{R}$ be a convex function and $\mathbf{r}=(r_1,\dots,r_n)$ and $\mathbf{s}=(s_1,\dots,s_n)$ be positive real $n-$tuples such that $R_n=\sum_{i=1}^n r_i$, $S_n=\sum_{i=1}^n s_i$. Suppose $\mathbf{t}=(t_1,\dots,t_n)$ and $\mathbf{u}=(u_1,\dots,u_n)$ are positive real $n-$tuples such that $T_n=\sum_{i=1}^n t_i$ and $U_n=\sum_{i=1}^n u_i$. If $s_i\geqslant u_i$, for $i=1,\dots,n$, then*

$$D_f(\mathbf{r},\mathbf{s})\geqslant S_n f\left(\frac{R_n}{S_n}\right)+\sum_{i=1}^n u_i f\left(\frac{r_i}{s_i}\right)-U_n f\left(\frac{1}{U_n}\sum_{i=1}^n u_i\frac{r_i}{s_i}\right). \tag{11}$$

If $s_i\leqslant t_i$, for $i=1,\dots,n$, then

$$D_f(\mathbf{r},\mathbf{s})\leqslant S_n f\left(\frac{R_n}{S_n}\right)+\sum_{i=1}^n t_i f\left(\frac{r_i}{s_i}\right)-T_n f\left(\frac{1}{T_n}\sum_{i=1}^n t_i\frac{r_i}{s_i}\right). \tag{12}$$

If f is a concave function, then reverse inequalities hold in (11) and (12).

Proof. Inequality (11) follows directly from (1) if x_i is replaced by $\frac{r_i}{s_i}$ and p_i by s_i, regarding definition (2) of the Csiszár functional $D_f(\mathbf{r},\mathbf{s})$.
Inequality (12) follows directly from (1) if x_i is replaced by $\frac{r_i}{s_i}$ and q_i by s_i, regarding definition (2) of the Csiszár functional $D_f(\mathbf{r},\mathbf{s})$.
Inequalities change their signs in case of concavity of the function f as a consequence of the Jensen inequality implicitly included. □

REMARK 1. Inequalities (11) and (12) are a generalization of specific bounds for the Csiszár functional (2) that were previously obtained in [8]. Namely, by means of simultaneous inserting the constant $n-$tuples $\mathbf{u}$ and $\mathbf{t}$ into inequalities (11) and (12), where components $u_1=u_2=\cdots=u_n=\min_{i=1,\dots,n}\{s_i\}$ and $t_1=t_2=\cdots=t_n=\max_{i=1,\dots,n}\{s_i\}$, we get the following bounds as in [8]:

$$S_n f\left(\frac{R_n}{S_n}\right)+\max_{i=1,\dots,n}\{s_i\}\left(\sum_{i=1}^n f\left(\frac{r_i}{s_i}\right)-nf\left(\frac{1}{n}\sum_{i=1}^n\frac{r_i}{s_i}\right)\right)\geqslant D_f(\mathbf{r},\mathbf{s})$$
$$\geqslant S_n f\left(\frac{R_n}{S_n}\right)+\min_{i=1,\dots,n}\{s_i\}\left(\sum_{i=1}^n f\left(\frac{r_i}{s_i}\right)-nf\left(\frac{1}{n}\sum_{i=1}^n\frac{r_i}{s_i}\right)\right). \tag{13}$$

The (convex) function in the second general theorem is adjusted to some specific purposes in the sequel.

THEOREM 2. *Let $f\colon \langle 0,\infty\rangle \to \mathbb{R}$ be such that $t \mapsto tf(t)$ is a convex function. Assume $\mathbf{r} = (r_1,\ldots,r_n)$ and $\mathbf{s} = (s_1,\ldots,s_n)$ to be positive real $n-$tuples such that $R_n = \sum_{i=1}^n r_i$, $S_n = \sum_{i=1}^n s_i$. Suppose $\mathbf{t} = (t_1,\ldots,t_n)$ and $\mathbf{u} = (u_1,\ldots,u_n)$ are positive real $n-$tuples such that $T_n = \sum_{i=1}^n t_i$ and $U_n = \sum_{i=1}^n u_i$. If $s_i \geqslant u_i$, for $i = 1,\ldots,n$, then*

$$D_{id\cdot f}(\mathbf{r},\mathbf{s}) \geqslant R_n f\left(\frac{R_n}{S_n}\right) + \sum_{i=1}^n u_i \frac{r_i}{s_i} f\left(\frac{r_i}{s_i}\right) - \left(\sum_{i=1}^n u_i \frac{r_i}{s_i}\right) f\left(\frac{1}{U_n}\sum_{i=1}^n u_i \frac{r_i}{s_i}\right). \tag{14}$$

If $s_i \leqslant t_i$, for $i = 1,\ldots,n$, then

$$D_{id\cdot f}(\mathbf{r},\mathbf{s}) \leqslant R_n f\left(\frac{R_n}{S_n}\right) + \sum_{i=1}^n t_i \frac{r_i}{s_i} f\left(\frac{r_i}{s_i}\right) - \left(\sum_{i=1}^n t_i \frac{r_i}{s_i}\right) f\left(\frac{1}{T_n}\sum_{i=1}^n t_i \frac{r_i}{s_i}\right), \tag{15}$$

where $D_{id\cdot f}(\mathbf{r},\mathbf{s}) := \sum_{i=1}^n r_i f\left(\frac{r_i}{s_i}\right)$.

If $t \mapsto tf(t)$ is a concave function, then reverse inequalities hold in (14) *and* (15).

Proof. Inequality (14) follows directly from (1) for convex function $t \mapsto tf(t)$ if x_i is replaced by $\frac{r_i}{s_i}$ and p_i by s_i, regarding definition (2) of the Csiszár functional.
Inequality (15) follows directly from (1) for convex function $t \mapsto tf(t)$ if x_i is replaced by $\frac{r_i}{s_i}$ and q_i by s_i, regarding definition (2) of the Csiszár functional.
Inequalities change their signs in case of concavity of the function f as a consequence of the Jensen inequality implicitly included. □

REMARK 2. Inequalities (14) and (15) are a generalization of specific bounds for the functional $D_{id\cdot f}(\mathbf{r},\mathbf{s})$ that were previously obtained in [8]. Namely, by means of simultaneous inserting the constant $n-$tuples $\mathbf{u}$ and $\mathbf{t}$ into inequalities (14) and (15), where components $u_1 = u_2 = \cdots = u_n = \min_{i=1,\ldots,n}\{s_i\}$ and $t_1 = t_2 = \cdots = t_n = \max_{i=1,\ldots,n}\{s_i\}$, we get the following bounds as in [8]:

$$\begin{aligned} R_n f\left(\frac{R_n}{S_n}\right) + \max_{i=1,\ldots,n}\{s_i\}\left(\sum_{i=1}^n \frac{r_i}{s_i} f\left(\frac{r_i}{s_i}\right) - \sum_{i=1}^n \frac{r_i}{s_i} f\left(\frac{1}{n}\sum_{i=1}^n \frac{r_i}{s_i}\right)\right) \geqslant D_{id\cdot f}(\mathbf{r},\mathbf{s}) \\ \geqslant R_n f\left(\frac{R_n}{S_n}\right) + \min_{i=1,\ldots,n}\{s_i\}\left(\sum_{i=1}^n \frac{r_i}{s_i} f\left(\frac{r_i}{s_i}\right) - \sum_{i=1}^n \frac{r_i}{s_i} f\left(\frac{1}{n}\sum_{i=1}^n \frac{r_i}{s_i}\right)\right). \end{aligned} \tag{16}$$

In the following corollary functional $D_{id\cdot f}(\mathbf{r},\mathbf{s})$ assumes a specific role which precedes that of the Kullback-Leibler divergence (3).

COROLLARY 1. *Let* $\mathbf{r}$, $\mathbf{s}$, $\mathbf{t}$ *and* $\mathbf{u}$ *be as in Theorem* 2. *If* $s_i \geqslant u_i$, *for* $i = 1, \ldots, n$, *then*

$$\sum_{i=1}^{n} r_i \log \frac{r_i}{s_i} \geqslant R_n \log \frac{R_n}{S_n} + \sum_{i=1}^{n} u_i \frac{r_i}{s_i} \log \frac{r_i}{s_i} - \left(\sum_{i=1}^{n} u_i \frac{r_i}{s_i} \right) \log \left(\frac{1}{U_n} \sum_{i=1}^{n} u_i \frac{r_i}{s_i} \right), \tag{17}$$

where the logarithm base is greater than 1.
If $s_i \leqslant t_i$, *for* $i = 1, \ldots, n$, *then*

$$\sum_{i=1}^{n} r_i \log \frac{r_i}{s_i} \leqslant R_n \log \frac{R_n}{S_n} + \sum_{i=1}^{n} t_i \frac{r_i}{s_i} \log \frac{r_i}{s_i} - \left(\sum_{i=1}^{n} t_i \frac{r_i}{s_i} \right) \log \left(\frac{1}{T_n} \sum_{i=1}^{n} t_i \frac{r_i}{s_i} \right), \tag{18}$$

where the logarithm base is greater than 1.
If the logarithm base is less than 1, then reverse inequalities hold in (17) *and* (18).

Proof. Follows from Theorem 2 for the function $t \mapsto t \log t$ which is convex when the logarithm base is greater than 1 and is concave when the logarithm base is less than 1. □

REMARK 3. When we put constant $n-$tuples $\mathbf{u}$ and $\mathbf{t}$ into (17) and (18) with components $u_1 = u_2 = \cdots = u_n = \min_{i=1,\ldots,n} \{s_i\}$ and $t_1 = t_2 = \cdots = t_n = \max_{i=1,\ldots,n} \{s_i\}$, we get as a special case the inequalities previously obtained in [8]:

$$R_n \log \frac{R_n}{S_n} + \max_{i=1,\ldots,n} \{s_i\} \left(\sum_{i=1}^{n} \frac{r_i}{s_i} \log \frac{r_i}{s_i} - \sum_{i=1}^{n} \frac{r_i}{s_i} \log \left(\frac{1}{n} \sum_{i=1}^{n} \frac{r_i}{s_i} \right) \right) \geqslant \sum_{i=1}^{n} r_i \log \frac{r_i}{s_i}$$
$$\geqslant R_n \log \frac{R_n}{S_n} + \min_{i=1,\ldots,n} \{s_i\} \left(\sum_{i=1}^{n} \frac{r_i}{s_i} \log \frac{r_i}{s_i} - \sum_{i=1}^{n} \frac{r_i}{s_i} \log \left(\frac{1}{n} \sum_{i=1}^{n} \frac{r_i}{s_i} \right) \right), \tag{19}$$

where the logarithm base is greater than 1. If the logarithm base is less than 1, then the inequality signs are reversed.

REMARK 4. In case of positive probability distributions $\mathbf{r}$ and $\mathbf{s}$, *i.e.* $r_i,\ s_i \in \langle 0, 1]$, $\sum_{i=1}^{n} r_i = \sum_{i=1}^{n} s_i = 1$, where $\mathbf{t}$ and $\mathbf{u}$ are as in Theorem 2, we actually deal with the Kullback-Leibler divergence $KL(\mathbf{r}, \mathbf{s})$ defined by (3). Hence if $s_i \geqslant u_i$, for $i = 1, \ldots, n$ we have

$$KL(\mathbf{r}, \mathbf{s}) \geqslant \sum_{i=1}^{n} u_i \frac{r_i}{s_i} \log \frac{r_i}{s_i} - \left(\sum_{i=1}^{n} u_i \frac{r_i}{s_i} \right) \log \left(\frac{1}{U_n} \sum_{i=1}^{n} u_i \frac{r_i}{s_i} \right), \tag{20}$$

where the logarithm base is greater than 1.
If $s_i \leqslant t_i$, for $i = 1, \ldots, n$, then

$$KL(\mathbf{r}, \mathbf{s}) \leqslant \sum_{i=1}^{n} t_i \frac{r_i}{s_i} \log \frac{r_i}{s_i} - \left(\sum_{i=1}^{n} t_i \frac{r_i}{s_i} \right) \log \left(\frac{1}{T_n} \sum_{i=1}^{n} t_i \frac{r_i}{s_i} \right), \tag{21}$$

where the logarithm base is greater than 1. If the logarithm base is less than 1, then reverse inequalities hold in (20) and (21).
Furthermore, inequalities (20) and (21) generalize specific bounds for the Kullback-Leibler divergence which were previously obtained in [8]. Namely, by means of simultaneous inserting the constant $n-$tuples $\mathbf{u}$ and $\mathbf{t}$ into inequalities (20) and (21), where components $u_1 = u_2 = \cdots = u_n = \min_{i=1,\dots,n}\{s_i\}$ and $t_1 = t_2 = \cdots = t_n = \max_{i=1,\dots,n}\{s_i\}$, we get the following bounds as presented in [8]:

$$\max_{i=1,\dots,n}\{s_i\}\left(\sum_{i=1}^{n}\frac{r_i}{s_i}\log\frac{r_i}{s_i}-\sum_{i=1}^{n}\frac{r_i}{s_i}\log\left(\frac{1}{n}\sum_{i=1}^{n}\frac{r_i}{s_i}\right)\right)\geqslant KL(\mathbf{r},\mathbf{s})$$
$$\geqslant \min_{i=1,\dots,n}\{s_i\}\left(\sum_{i=1}^{n}\frac{r_i}{s_i}\log\frac{r_i}{s_i}-\sum_{i=1}^{n}\frac{r_i}{s_i}\log\left(\frac{1}{n}\sum_{i=1}^{n}\frac{r_i}{s_i}\right)\right), \tag{22}$$

where the logarithm base is greater than 1. If the logarithm base is less than 1, then the inequality signs are reversed.

In the sequel we provide similar results for other divergences mentioned in Preliminaries: Hellinger distance (4), Bhattacharyya coefficient (5), chi-square distance (6) and total variation distance (7).

COROLLARY 2. *Let* $\mathbf{r}$, $\mathbf{s}$, $\mathbf{t}$ *and* $\mathbf{u}$ *be as in Theorem* 1. *If* $s_i \geqslant u_i$, *for* $i = 1,\dots,n$, *then*

$$\frac{1}{2}\sum_{i=1}^{n}(\sqrt{r_i}-\sqrt{s_i})^2 \geqslant \frac{S_n}{2}\left(\sqrt{\frac{R_n}{S_n}}-1\right)^2+\frac{1}{2}\sum_{i=1}^{n}u_i\left(\sqrt{\frac{r_i}{s_i}}-1\right)^2-\frac{U_n}{2}\left(\sqrt{\frac{1}{U_n}\sum_{i=1}^{n}u_i\frac{r_i}{s_i}}-1\right)^2. \tag{23}$$

If $s_i \leqslant t_i$, *for* $i = 1,\dots,n$, *then*

$$\frac{1}{2}\sum_{i=1}^{n}(\sqrt{r_i}-\sqrt{s_i})^2 \leqslant \frac{S_n}{2}\left(\sqrt{\frac{R_n}{S_n}}-1\right)^2+\frac{1}{2}\sum_{i=1}^{n}t_i\left(\sqrt{\frac{r_i}{s_i}}-1\right)^2-\frac{T_n}{2}\left(\sqrt{\frac{1}{T_n}\sum_{i=1}^{n}t_i\frac{r_i}{s_i}}-1\right)^2. \tag{24}$$

Proof. Follows from Theorem 1 for the convex function $t \mapsto \frac{1}{2}\left(\sqrt{t}-1\right)^2$. □

REMARK 5. Certain bounds formerly obtained in [8] can now be deduced from more general inequalities (23) and (24):

$$\frac{S_n}{2}\left(\sqrt{\frac{R_n}{S_n}}-1\right)^2+\frac{1}{2}\max_{i=1,\dots,n}\{s_i\}\left(\sum_{i=1}^{n}\left(\sqrt{\frac{r_i}{s_i}}-1\right)^2-n\left(\sqrt{\frac{1}{n}\sum_{i=1}^{n}\frac{r_i}{s_i}}-1\right)^2\right)$$
$$\geqslant \frac{1}{2}\sum_{i=1}^{n}(\sqrt{r_i}-\sqrt{s_i})^2$$
$$\geqslant \frac{S_n}{2}\left(\sqrt{\frac{R_n}{S_n}}-1\right)^2+\frac{1}{2}\min_{i=1,\dots,n}\{s_i\}\left(\sum_{i=1}^{n}\left(\sqrt{\frac{r_i}{s_i}}-1\right)^2-n\left(\sqrt{\frac{1}{n}\sum_{i=1}^{n}\frac{r_i}{s_i}}-1\right)^2\right), \tag{25}$$

by means of constant $n-$tuples $\mathbf{u}$ and $\mathbf{t}$, where $u_1 = u_2 = \cdots = u_n = \min_{i=1,\dots,n} \{s_i\}$ and $t_1 = t_2 = \cdots = t_n = \max_{i=1,\dots,n} \{s_i\}$.

REMARK 6. If $\mathbf{r}$ and $\mathbf{s}$ are observed as positive probability distributions and $\mathbf{t}$ and $\mathbf{u}$ are as in Theorem 1, then inequalities (23) and (24) concern the Hellinger distance $h^2(\mathbf{r},\mathbf{s})$. Thus for $s_i \geqslant u_i,\ i = 1,\dots,n$

$$h^2(\mathbf{r},\mathbf{s}) \geqslant \frac{1}{2}\sum_{i=1}^{n} u_i \left(\sqrt{\frac{r_i}{s_i}} - 1\right)^2 - \frac{U_n}{2}\left(\sqrt{\frac{1}{U_n}\sum_{i=1}^{n} u_i \frac{r_i}{s_i}} - 1\right)^2. \tag{26}$$

If $s_i \leqslant t_i,\ i = 1,\dots,n$, then

$$h^2(\mathbf{r},\mathbf{s}) \leqslant \frac{1}{2}\sum_{i=1}^{n} t_i \left(\sqrt{\frac{r_i}{s_i}} - 1\right)^2 - \frac{T_n}{2}\left(\sqrt{\frac{1}{T_n}\sum_{i=1}^{n} t_i \frac{r_i}{s_i}} - 1\right)^2. \tag{27}$$

Furthermore, inequalities (26) and (27) generalize by means of the constant $n-$ tuples $\mathbf{u}$ and $\mathbf{t}$, with $u_1 = u_2 = \cdots = u_n = \min_{i=1,\dots,n} \{s_i\}$ and $t_1 = t_2 = \cdots = t_n = \max_{i=1,\dots,n} \{s_i\}$ specific bounds for the Hellinger distance which were previously obtained in [8]:

$$\begin{aligned} &\frac{1}{2}\max_{i=1,\dots,n}\{s_i\}\left(\sum_{i=1}^{n}\left(\sqrt{\frac{r_i}{s_i}} - 1\right)^2 - n\left(\sqrt{\frac{1}{n}\sum_{i=1}^{n}\frac{r_i}{s_i}} - 1\right)^2\right) \geqslant h^2(\mathbf{r},\mathbf{s}) \\ &\geqslant \frac{1}{2}\min_{i=1,\dots,n}\{s_i\}\left(\sum_{i=1}^{n}\left(\sqrt{\frac{r_i}{s_i}} - 1\right)^2 - n\left(\sqrt{\frac{1}{n}\sum_{i=1}^{n}\frac{r_i}{s_i}} - 1\right)^2\right). \end{aligned} \tag{28}$$

COROLLARY 3. *Let* $\mathbf{r}$, $\mathbf{s}$, $\mathbf{t}$ *and* $\mathbf{u}$ *be as in Theorem 1. If* $s_i \geqslant u_i$, *for* $i = 1,\dots,n$, *then*

$$-\sum_{i=1}^{n}\sqrt{r_i s_i} \geqslant -\sqrt{R_n S_n} - \sum_{i=1}^{n} u_i \sqrt{\frac{r_i}{s_i}} + \sqrt{U_n \sum_{i=1}^{n} u_i \frac{r_i}{s_i}}. \tag{29}$$

If $s_i \leqslant t_i$, *for* $i = 1,\dots,n$, *then*

$$-\sum_{i=1}^{n}\sqrt{r_i s_i} \leqslant -\sqrt{R_n S_n} - \sum_{i=1}^{n} t_i \sqrt{\frac{r_i}{s_i}} + \sqrt{T_n \sum_{i=1}^{n} t_i \frac{r_i}{s_i}}. \tag{30}$$

Proof. Follows from Theorem 1 for the convex function $t \mapsto -\sqrt{t}$. □

REMARK 7. In case of $u_1 = u_2 = \cdots = u_n = \min_{i=1,\dots,n}\{s_i\}$ and $t_1 = t_2 = \cdots = t_n = \max_{i=1,\dots,n}\{s_i\}$ inserted into (29) and (30), constant $n-$tuples $\mathbf{u}$ and $\mathbf{t}$ yield the following bounds:

$$\max_{i=1,\dots,n}\{s_i\}\left(\sqrt{n\sum_{i=1}^{n}\frac{r_i}{s_i}} - \sum_{i=1}^{n}\sqrt{\frac{r_i}{s_i}}\right) - \sqrt{R_n S_n} \geqslant -\sum_{i=1}^{n}\sqrt{r_i s_i}$$

$$\geqslant \min_{i=1,\dots,n}\{s_i\}\left(\sqrt{n\sum_{i=1}^{n}\frac{r_i}{s_i}}-\sum_{i=1}^{n}\sqrt{\frac{r_i}{s_i}}\right)-\sqrt{R_nS_n}, \tag{31}$$

which were obtained in [8] due to a less general approach.

REMARK 8. For positive probability distributions $\mathbf{r}$ and $\mathbf{s}$, where $\mathbf{t}$ and $\mathbf{u}$ are as in Theorem 1, we actually deal with the Bhattacharyya coefficient $B(\mathbf{r},\mathbf{s})$. Hence if $s_i \geqslant u_i$, for $i=1,\dots,n$ we have

$$-B(\mathbf{r},\mathbf{s}) \geqslant -1-\sum_{i=1}^{n}u_i\sqrt{\frac{r_i}{s_i}}+\sqrt{U_n\sum_{i=1}^{n}u_i\frac{r_i}{s_i}}. \tag{32}$$

If $s_i \leqslant t_i$, for $i=1,\dots,n$, then

$$-B(\mathbf{r},\mathbf{s}) \leqslant -1-\sum_{i=1}^{n}t_i\sqrt{\frac{r_i}{s_i}}+\sqrt{T_n\sum_{i=1}^{n}t_i\frac{r_i}{s_i}}. \tag{33}$$

If we make use of constant $n-$tuples $\mathbf{u}$ and $\mathbf{t}$ with components $u_1=u_2=\cdots=u_n=\min_{i=1,\dots,n}\{s_i\}$ and $t_1=t_2=\cdots=t_n=\max_{i=1,\dots,n}\{s_i\}$ and insert them into inequalities (32) and (33), we get :

$$1-\min_{i=1,\dots,n}\{s_i\}\left(\sqrt{n\sum_{i=1}^{n}\frac{r_i}{s_i}}-\sum_{i=1}^{n}\sqrt{\frac{r_i}{s_i}}\right)\geqslant B(\mathbf{r},\mathbf{s})\geqslant 1-\max_{i=1,\dots,n}\{s_i\}\left(\sqrt{n\sum_{i=1}^{n}\frac{r_i}{s_i}}-\sum_{i=1}^{n}\sqrt{\frac{r_i}{s_i}}\right), \tag{34}$$

that is, bounds from [8] which are a special case in this more general setting.

COROLLARY 4. *Let* $\mathbf{r}$, $\mathbf{s}$, $\mathbf{t}$ *and* $\mathbf{u}$ *be as in Theorem* 1. *If* $s_i \geqslant u_i$, *for* $i=1,\dots,n$, *then*

$$\sum_{i=1}^{n}\frac{(r_i-s_i)^2}{s_i} \geqslant S_n\left(\frac{R_n}{S_n}-1\right)^2+\sum_{i=1}^{n}u_i\left(\frac{r_i}{s_i}-1\right)^2-U_n\left(\frac{1}{U_n}\sum_{i=1}^{n}u_i\frac{r_i}{s_i}-1\right)^2. \tag{35}$$

If $s_i \leqslant t_i$, *for* $i=1,\dots,n$, *then*

$$\sum_{i=1}^{n}\frac{(r_i-s_i)^2}{s_i} \leqslant S_n\left(\frac{R_n}{S_n}-1\right)^2+\sum_{i=1}^{n}t_i\left(\frac{r_i}{s_i}-1\right)^2-T_n\left(\frac{1}{T_n}\sum_{i=1}^{n}t_i\frac{r_i}{s_i}-1\right)^2. \tag{36}$$

Proof. Follows from Theorem 1 for the convex function $t \mapsto (t-1)^2$. □

REMARK 9. When we put constant $n-$tuples $\mathbf{u}$ and $\mathbf{t}$ into (35) and (36) with components $u_1=u_2=\cdots=u_n=\min_{i=1,\dots,n}\{s_i\}$ and $t_1=t_2=\cdots=t_n=\max_{i=1,\dots,n}\{s_i\}$ we get as a special case the inequalities previously obtained in [8]:

$$S_n\left(\frac{R_n}{S_n}-1\right)^2+\max_{i=1,\dots,n}\{s_i\}\left(\sum_{i=1}^{n}\left(\frac{r_i}{s_i}-1\right)^2-n\left(\frac{1}{n}\sum_{i=1}^{n}\frac{r_i}{s_i}-1\right)^2\right)\geqslant\sum_{i=1}^{n}\frac{(r_i-s_i)^2}{s_i}$$

$$\geqslant S_n\left(\frac{R_n}{S_n}-1\right)^2+\min_{i=1,\dots,n}\{s_i\}\left(\sum_{i=1}^{n}\left(\frac{r_i}{s_i}-1\right)^2-n\left(\frac{1}{n}\sum_{i=1}^{n}\frac{r_i}{s_i}-1\right)^2\right). \tag{37}$$

REMARK 10. In case of positive probability distributions $\mathbf{r}$ and $\mathbf{s}$, where $\mathbf{t}$ and $\mathbf{u}$ are as in Theorem 1, we actually deal with the chi-square divergence $\chi^2(\mathbf{r},\mathbf{s})$ defined by (6). Hence if $s_i \geqslant u_i$, for $i=1,\dots,n$ we have

$$\chi^2(\mathbf{r},\mathbf{s}) \geqslant \sum_{i=1}^{n} u_i\left(\frac{r_i}{s_i}-1\right)^2-U_n\left(\frac{1}{U_n}\sum_{i=1}^{n}u_i\frac{r_i}{s_i}-1\right)^2. \tag{38}$$

If $s_i \leqslant t_i,\ i=1,\dots,n$, then

$$\chi^2(\mathbf{r},\mathbf{s}) \leqslant \sum_{i=1}^{n} t_i\left(\frac{r_i}{s_i}-1\right)^2-T_n\left(\frac{1}{T_n}\sum_{i=1}^{n}t_i\frac{r_i}{s_i}-1\right)^2. \tag{39}$$

Furthermore, inequalities (38) and (39) generalize specific bounds for the chi-square divergence which were previously obtained in [8]. Namely, by means of simultaneous inserting the constant $n-$tuples $\mathbf{u}$ and $\mathbf{t}$ into inequalities (38) and (39), where components $u_1=u_2=\cdots=u_n=\min_{i=1,\dots,n}\{s_i\}$ and $t_1=t_2=\cdots=t_n=\max_{i=1,\dots,n}\{s_i\}$, we get the following bounds as presented in [8]:

$$\max_{i=1,\dots,n}\{s_i\}\left(\sum_{i=1}^{n}\left(\frac{r_i}{s_i}-1\right)^2-n\left(\frac{1}{n}\sum_{i=1}^{n}\frac{r_i}{s_i}-1\right)^2\right)\geqslant\chi^2(\mathbf{r},\mathbf{s})$$
$$\geqslant\min_{i=1,\dots,n}\{s_i\}\left(\sum_{i=1}^{n}\left(\frac{r_i}{s_i}-1\right)^2-n\left(\frac{1}{n}\sum_{i=1}^{n}\frac{r_i}{s_i}-1\right)^2\right). \tag{40}$$

COROLLARY 5. *Let* $\mathbf{r}$, $\mathbf{s}$, $\mathbf{t}$ *and* $\mathbf{u}$ *be as in Theorem* 1. *If* $s_i \geqslant u_i$, *for* $i=1,\dots,n$, *then*

$$\sum_{i=1}^{n}|r_i-s_i|\geqslant S_n\left|\frac{R_n}{S_n}-1\right|+\sum_{i=1}^{n}u_i\left|\frac{r_i}{s_i}-1\right|-U_n\left|\frac{1}{U_n}\sum_{i=1}^{n}u_i\frac{r_i}{s_i}-1\right|. \tag{41}$$

If $s_i \leqslant t_i$, *for* $i=1,\dots,n$, *then*

$$\sum_{i=1}^{n}|r_i-s_i|\leqslant S_n\left|\frac{R_n}{S_n}-1\right|+\sum_{i=1}^{n}t_i\left|\frac{r_i}{s_i}-1\right|-T_n\left|\frac{1}{T_n}\sum_{i=1}^{n}t_i\frac{r_i}{s_i}-1\right|. \tag{42}$$

Proof. Follows from Theorem 1 for the convex function $t\mapsto|t-1|$. □

REMARK 11. Certain bounds formerly obtained in [8] can now be deduced from more general inequalities (41) and (42):

$$S_n\left|\frac{R_n}{S_n}-1\right|+\max_{i=1,\dots,n}\{s_i\}\left(\sum_{i=1}^{n}\left|\frac{r_i}{s_i}-1\right|-n\left|\frac{1}{n}\sum_{i=1}^{n}\frac{r_i}{s_i}-1\right|\right)\geqslant\sum_{i=1}^{n}|r_i-s_i|$$

$$\geqslant S_n\left|\frac{R_n}{S_n}-1\right|+\min_{i=1,\dots,n}\{s_i\}\left(\sum_{i=1}^{n}\left|\frac{r_i}{s_i}-1\right|-n\left|\frac{1}{n}\sum_{i=1}^{n}\frac{r_i}{s_i}-1\right|\right), \tag{43}$$

by means of constant $n-$tuples $\mathbf{u}$ and $\mathbf{t}$, where components $u_1=u_2=\cdots=u_n=\min_{i=1,\dots,n}\{s_i\}$ and $t_1=t_2=\cdots=t_n=\max_{i=1,\dots,n}\{s_i\}$.

REMARK 12. If $\mathbf{r}$ and $\mathbf{s}$ are observed as positive probability distributions and $\mathbf{t}$ and $\mathbf{u}$ are as in Theorem 1, then inequalities (41) and (42) concern the total variation distance $V(\mathbf{r},\mathbf{s})$. Thus for $s_i\geqslant u_i,\ i=1,\dots,n$

$$V(\mathbf{r},\mathbf{s})\geqslant\sum_{i=1}^{n}u_i\left|\frac{r_i}{s_i}-1\right|-U_n\left|\frac{1}{U_n}\sum_{i=1}^{n}u_i\frac{r_i}{s_i}-1\right|. \tag{44}$$

If $s_i\leqslant t_i$, for $i=1,\dots,n$, then

$$V(\mathbf{r},\mathbf{s})\leqslant\sum_{i=1}^{n}t_i\left|\frac{r_i}{s_i}-1\right|-T_n\left|\frac{1}{T_n}\sum_{i=1}^{n}t_i\frac{r_i}{s_i}-1\right|. \tag{45}$$

Furthermore, inequalities (44) and (45) generalize by means of the constant $n-$ tuples $\mathbf{u}$ and $\mathbf{t}$, with $u_1=u_2=\cdots=u_n=\min_{i=1,\dots,n}\{s_i\}$ and $t_1=t_2=\cdots=t_n=\max_{i=1,\dots,n}\{s_i\}$ specific bounds for the total variation distance which were previously obtained in [8]:

$$\max_{i=1,\dots,n}\{s_i\}\left(\sum_{i=1}^{n}\left|\frac{r_i}{s_i}-1\right|-n\left|\frac{1}{n}\sum_{i=1}^{n}\frac{r_i}{s_i}-1\right|\right)\geqslant V(\mathbf{r},\mathbf{s})$$
$$\geqslant\min_{i=1,\dots,n}\{s_i\}\left(\sum_{i=1}^{n}\left|\frac{r_i}{s_i}-1\right|-n\left|\frac{1}{n}\sum_{i=1}^{n}\frac{r_i}{s_i}-1\right|\right). \tag{46}$$

3. Results for $f-$ divergences via the Zipf-Mandelbrot law

If we define $s_i=f(i;N,v,w)$, for $i=1,\dots,N$ as the Zipf-Mandelbrot law probability mass functions (8) we can use a new environment to observe the previously obtained results. When observed with the Zipf-Mandelbrot $N-$tuple $\mathbf{s}$ included, the Csiszár functional $D_f(\mathbf{r},\mathbf{s})$ defined by (2) becomes

$$D_f(\mathbf{r},i,N,v_2,w_2)=\sum_{i=1}^{N}\frac{1}{(i+w_2)^{v_2}H_{N,v_2,w_2}}f\left(r_i\,(i+w_2)^{v_2}H_{N,v_2,w_2}\right), \tag{47}$$

where $f\colon\langle 0,\infty\rangle\to\mathbb{R}$ and $N\in\mathbb{N},\ v_2,w_2>0$ are parameters.
Csiszár functional (2) assumes the following form when $\mathbf{r}$ and $\mathbf{s}$ are both defined as Zipf-Mandelbrot law $N-$tuples:

$$D_f(i,N,v_1,w_1,v_2,w_2)=\sum_{i=1}^{N}\frac{1}{(i+w_2)^{v_2}H_{N,v_2,w_2}}f\left(\frac{(i+w_2)^{v_2}H_{N,v_2,w_2}}{(i+w_1)^{v_1}H_{N,v_1,w_1}}\right), \tag{48}$$

where $f\colon \langle 0,\infty\rangle \to \mathbb{R}$ and $N\in\mathbb{N}$, $v_1,v_2,w_1,w_2>0$ are parameters.

Finally, both $N-$tuples $\mathbf{r}$ and $\mathbf{s}$ may be defined via the Zipf law (10) where $w_1 = w_2 = 0$ and thus the Csiszár functional (2) assumes the form:

$$D_f(i,N,v_1,v_2) = \sum_{i=1}^{N} \frac{1}{i^{v_2} H_{N,v_2}} f\left(i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}\right). \tag{49}$$

In the first case, that is for the Csiszár functional $D_f(i,N,v_2,w_2,\mathbf{r})$ given as in (47) we transform Theorem 1 and Theorem 2 in the following way.

COROLLARY 6. *Let $f\colon \langle 0,\infty\rangle \to \mathbb{R}$ be a convex function, $v_2,w_2>0$ and let $\mathbf{r} = (r_1,\ldots,r_N)$ be an $N-$tuple of positive real numbers with $R_N=\sum_{i=1}^N r_i$. Suppose $\mathbf{t} = (t_1,\ldots,t_N)$ and $\mathbf{u}=(u_1,\ldots,u_N)$ are positive real $N-$tuples such that $T_N=\sum_{i=1}^N t_i$ and $U_N=\sum_{i=1}^N u_i$. If $\dfrac{1}{(i+w_2)^{v_2}H_{N,v_2,w_2}} \geqslant u_i$, for $i=1,\ldots,N$, then*

$$D_f(\mathbf{r},i,N,v_2,w_2) \geqslant f(R_N) + \sum_{i=1}^{N} u_i f\left(r_i(i+w_2)^{v_2}H_{N,v_2,w_2}\right) - U_N f\left(\frac{1}{U_N}\sum_{i=1}^{N} u_i r_i (i+w_2)^{v_2} H_{N,v_2,w_2}\right). \tag{50}$$

If $\dfrac{1}{(i+w_2)^{v_2}H_{N,v_2,w_2}} \leqslant t_i$, for $i=1,\ldots,N$, then

$$D_f(\mathbf{r},i,N,v_2,w_2) \leqslant f(R_N) + \sum_{i=1}^{N} t_i f\left(r_i(i+w_2)^{v_2}H_{N,v_2,w_2}\right) - T_N f\left(\frac{1}{T_N}\sum_{i=1}^{N} t_i r_i (i+w_2)^{v_2} H_{N,v_2,w_2}\right). \tag{51}$$

If f is a concave function, then reverse inequalities hold in (50) and (51).
Suppose $t\mapsto tf(t)$ is a convex function.
If $\dfrac{1}{(i+w_2)^{v_2}H_{N,v_2,w_2}} \geqslant u_i$, for $i=1,\ldots,N$, then

$$\begin{aligned} D_{id\cdot f}(\mathbf{r},i,N,v_2,w_2) \geqslant{}& R_N f(R_N) + \sum_{i=1}^{N} u_i r_i (i+w_2)^{v_2} H_{N,v_2,w_2} f\left(r_i(i+w_2)^{v_2}H_{N,v_2,w_2}\right) \\ &- \left(\sum_{i=1}^{N} u_i r_i (i+w_2)^{v_2} H_{N,v_2,w_2}\right) \\ &\times f\left(\frac{1}{U_N}\sum_{i=1}^{N} u_i r_i (i+w_2)^{v_2} H_{N,v_2,w_2}\right), \end{aligned} \tag{52}$$

where $D_{id\cdot f}(\mathbf{r},i,N,v_2,w_2) := \sum_{i=1}^{N} r_i f\left((i+w_2)^{v_2}H_{N,v_2,w_2}\right)$.

If $\frac{1}{(i+w_2)^{v_2}H_{N,v_2,w_2}} \leqslant t_i$, *for* $i=1,\dots,N$, *then*

$$D_{id\cdot f}(\mathbf{r},i,N,v_2,w_2) \leqslant R_N f(R_N) + \sum_{i=1}^{N} t_i r_i (i+w_2)^{v_2} H_{N,v_2,w_2} f\left(r_i (i+w_2)^{v_2} H_{N,v_2,w_2}\right) - \left(\sum_{i=1}^{N} t_i r_i (i+w_2)^{v_2} H_{N,v_2,w_2}\right) \times f\left(\frac{1}{T_N}\sum_{i=1}^{N} t_i r_i (i+w_2)^{v_2} H_{N,v_2,w_2}\right). \tag{53}$$

If $t \mapsto tf(t)$ *is a concave function, then reverse inequalities hold in* (52) *and* (53).

Proof. Inequalities (50) and (51) lean on the proof of Theorem 1 wherein we insert for s_i the expression $\frac{1}{(i+w)^{v_2}H_{N,v_2,w2}}$ and $S_N = 1$ by the definition (8) of the Zipf-Mandelbrot law. Inequalities (52) and (53) follow analogously after the proof of Theorem 2. Inequalities change their signs in the case of concavity of functions f or $t \mapsto tf(t)$ as a consequence of the Jensen inequality implicitly included. □

REMARK 13. If we put $u_1 = u_2 = \cdots = u_N = \min_{i=1,\dots,N}\left\{\frac{1}{(i+w_2)^{v_2}H_{N,v_2,w_2}}\right\} = \frac{1}{(N+w_2)^{v_2}H_{N,v_2,w_2}}$ and simultaneously $t_1 = t_2 = \cdots = t_N = \max_{i=1,\dots,N}\left\{\frac{1}{(i+w_2)^{v_2}H_{N,v_2,w_2}}\right\} = \frac{1}{(1+w_2)^{v_2}H_{N,v_2,w_2}}$ into inequalities (50) and (51) we get the following bounds as a special case of Corollary 6. These were obtained earlier in [8]:

$$f(R_N) + \frac{1}{(1+w_2)^{v_2}H_{N,v_2,w_2}}\Lambda_1 \geqslant D_f(\mathbf{r},i,N,v_2,w_2) \geqslant f(R_N) + \frac{1}{(N+w_2)^{v_2}H_{N,v_2,w_2}}\Lambda_1, \tag{54}$$

where $\Lambda_1 = \sum_{i=1}^{N} f\left(r_i (i+w_2)^{v_2} H_{N,v_2,w_2}\right) - Nf\left(\frac{1}{N}\sum_{i=1}^{N} r_i (i+w_2)^{v_2} H_{N,v_2,w_2}\right)$.
If we repeat the similar procedure with inequalities (52) and (53), we get the analogous bounds for $D_{id\cdot f}(\mathbf{r},i,N,v_2,w_2)$, previously obtained in [8], as well:

$$R_N f(R_N) + \frac{1}{(1+w_2)^{v_2}H_{N,v_2,w_2}}\widetilde{\Lambda}_1 \geqslant D_{id\cdot f}(\mathbf{r},i,N,v_2,w_2) \geqslant R_N f(R_N) + \frac{1}{(N+w_2)^{v_2}H_{N,v_2,w_2}}\widetilde{\Lambda}_1, \tag{55}$$

where

$$\widetilde{\Lambda}_1 = \sum_{i=1}^{N} r_i (i+w_2)^{v_2} H_{N,v_2,w_2} f\left(r_i (i+w_2)^{v_2} H_{N,v_2,w_2}\right)$$

$$-\left(\sum_{i=1}^{N} r_i (i+w_2)^{v_2} H_{N,v_2,w_2}\right) f\left(\frac{1}{N}\sum_{i=1}^{N} r_i (i+w_2)^{v_2} H_{N,v_2,w_2}\right). \tag{56}$$

In the second case, that is for the Csiszár functional $D_f(i,N,v_1,w_1,v_2,w_2)$ as in (48), we transform Theorem 1 and Theorem 2 as follows.

COROLLARY 7. *Let $f\colon \langle 0,\infty\rangle \to \mathbb{R}$ be a convex function and $v_1,v_2,w_1,w_2>0$. Suppose $\mathbf{t}=(t_1,\ldots,t_N)$ and $\mathbf{u}=(u_1,\ldots,u_N)$ are positive real $N-$tuples such that $T_N=\sum_{i=1}^{N} t_i$ and $U_N=\sum_{i=1}^{N} u_i$. If $\frac{1}{(i+w_2)^{v_2}H_{N,v_2,w_2}} \geqslant u_i$, for $i=1,\ldots,N$, then*

$$D_f(i,N,v_1,w_1,v_2,w_2) \geqslant f(1)+\sum_{i=1}^{N} u_i f\left(\frac{(i+w_2)^{v_2}H_{N,v_2,w_2}}{(i+w_1)^{v_1}H_{N,v_1,w_1}}\right) -U_N f\left(\frac{1}{U_N}\sum_{i=1}^{N} u_i \frac{(i+w_2)^{v_2}H_{N,v_2,w_2}}{(i+w_1)^{v_1}H_{N,v_1,w_1}}\right). \tag{57}$$

If $\frac{1}{(i+w_2)^{v_2}H_{N,v_2,w_2}} \leqslant t_i$, for $i=1,\ldots,N$, then

$$D_f(i,N,v_1,w_1,v_2,w_2) \leqslant f(1)+\sum_{i=1}^{N} t_i f\left(\frac{(i+w_2)^{v_2}H_{N,v_2,w_2}}{(i+w_1)^{v_1}H_{N,v_1,w_1}}\right) -T_N f\left(\frac{1}{T_N}\sum_{i=1}^{N} t_i \frac{(i+w_2)^{v_2}H_{N,v_2,w_2}}{(i+w_1)^{v_1}H_{N,v_1,w_1}}\right). \tag{58}$$

If f is a concave function, then reverse inequalities hold in (57) and (58).
Suppose $t\mapsto tf(t)$ is a convex function. If $\frac{1}{(i+w_2)^{v_2}H_{N,v_2,w_2}} \geqslant u_i$, for $i=1,\ldots,N$, then

$$\begin{aligned} &D_{id\cdot f}(i,N,v_1,w_1,v_2,w_2)\\ &\geqslant f(1)+\sum_{i=1}^{N} u_i \frac{(i+w_2)^{v_2}H_{N,v_2,w_2}}{(i+w_1)^{v_1}H_{N,v_1,w_1}} f\left(\frac{(i+w_2)^{v_2}H_{N,v_2,w_2}}{(i+w_1)^{v_1}H_{N,v_1,w_1}}\right)\\ &-\left(\sum_{i=1}^{N} u_i \frac{(i+w_2)^{v_2}H_{N,v_2,w_2}}{(i+w_1)^{v_1}H_{N,v_1,w_1}}\right) f\left(\frac{1}{U_N}\sum_{i=1}^{N} u_i \frac{(i+w_2)^{v_2}H_{N,v_2,w_2}}{(i+w_1)^{v_1}H_{N,v_1,w_1}}\right), \end{aligned} \tag{59}$$

where $D_{id\cdot f}(i,N,v_1,w_1,v_2,w_2) := \sum_{i=1}^{N} \frac{1}{(i+w_1)^{v_1}H_{N,v_1,w_1}} f\left(\frac{(i+w_2)^{v_2}H_{N,v_2,w_2}}{(i+w_1)^{v_1}H_{N,v_1,w_1}}\right)$.
If $\frac{1}{(i+w_2)^{v_2}H_{N,v_2,w_2}} \leqslant t_i$, for $i=1,\ldots,N$, then

$$\begin{aligned} &D_{id\cdot f}(i,N,v_1,w_1,v_2,w_2)\\ &\leqslant f(1)+\sum_{i=1}^{N} t_i \frac{(i+w_2)^{v_2}H_{N,v_2,w_2}}{(i+w_1)^{v_1}H_{N,v_1,w_1}} f\left(\frac{(i+w_2)^{v_2}H_{N,v_2,w_2}}{(i+w_1)^{v_1}H_{N,v_1,w_1}}\right) \end{aligned}$$

$$-\left(\sum_{i=1}^{N} t_i \frac{(i+w_2)^{v_2} H_{N,v_2,w_2}}{(i+w_1)^{v_1} H_{N,v_1,w_1}}\right) f\left(\frac{1}{T_N}\sum_{i=1}^{N} t_i \frac{(i+w_2)^{v_2} H_{N,v_2,w_2}}{(i+w_1)^{v_1} H_{N,v_1,w_1}}\right). \tag{60}$$

If $t \mapsto tf(t)$ is a concave function, then reverse inequalities hold in (59) *and* (60).

Proof. Inequalities (57) and (58) follow from Theorem 1 when we insert the expressions $\frac{1}{(i+w_1)^{v_1} H_{N,v_1,w_1}}$ and $\frac{1}{(i+w_2)^{v_2} H_{N,v_2,w_2}}$ by the definition (8) of the Zipf-Mandelbrot law in r_i and s_i respectively. Proving inequalities (59) and (60) follows the similar procedure, only concerning the proof of Theorem 2. Inequalities change their signs in the case of concavity of functions f or $t \mapsto tf(t)$ as a consequence of the Jensen inequality implicitly included. □

REMARK 14. For the choice of $u_1 = u_2 = \cdots = u_N = \min_{i=1,\ldots,N}\left\{\frac{1}{(i+w_2)^{v_2} H_{N,v_2,w_2}}\right\} = \frac{1}{(N+w_2)^{v_2} H_{N,v_2,w_2}}$ and simultaneously $t_1 = t_2 = \cdots = t_N = \max_{i=1,\ldots,N}\left\{\frac{1}{(i+w_2)^{v_2} H_{N,v_2,w_2}}\right\} = \frac{1}{(1+w_2)^{v_2} H_{N,v_2,w_2}}$, inequalities (57) and (58) assume the form of the bounds that were obtained earlier in [8] :

$$\begin{aligned} f(1) + \frac{1}{(1+w_2)^{v_2} H_{N,v_2,w_2}} \Lambda_2 &\geqslant D_f(i,N,v_1,w_1,v_2,w_2) \\ &\geqslant f(1) + \frac{1}{(N+w_2)^{v_2} H_{N,v_2,w_2}} \Lambda_2, \end{aligned} \tag{61}$$

where $\Lambda_2 = \sum_{i=1}^{N} f\left(\frac{(i+w_2)^{v_2} H_{N,v_2,w_2}}{(i+w_1)^{v_1} H_{N,v_1,w_1}}\right) - Nf\left(\frac{1}{N}\sum_{i=1}^{N}\frac{(i+w_2)^{v_2} H_{N,v_2,w_2}}{(i+w_1)^{v_1} H_{N,v_1,w_1}}\right)$.
If we repeat the similar procedure with inequalities (59) and (60), we get the analogous bounds for $D_{id\cdot f}(i,N,v_1,w_1,v_2,w_2)$, as in [8]:

$$\begin{aligned} f(1) + \frac{1}{(1+w_2)^{v_2} H_{N,v_2,w_2}} \widetilde{\Lambda}_2 &\geqslant D_{id\cdot f}(i,N,v_1,w_1,v_2,w_2) \\ &\geqslant f(1) + \frac{1}{(N+w_2)^{v_2} H_{N,v_2,w_2}} \widetilde{\Lambda}_2, \end{aligned} \tag{62}$$

where

$$\begin{aligned} \widetilde{\Lambda}_2 = &\sum_{i=1}^{N} \frac{(i+w_2)^{v_2} H_{N,v_2,w_2}}{(i+w_1)^{v_1} H_{N,v_1,w_1}} f\left(\frac{(i+w_2)^{v_2} H_{N,v_2,w_2}}{(i+w_1)^{v_1} H_{N,v_1,w_1}}\right) \\ &-\left(\sum_{i=1}^{N} \frac{(i+w_2)^{v_2} H_{N,v_2,w_2}}{(i+w_1)^{v_1} H_{N,v_1,w_1}}\right) f\left(\frac{1}{N}\sum_{i=1}^{N}\frac{(i+w_2)^{v_2} H_{N,v_2,w_2}}{(i+w_1)^{v_1} H_{N,v_1,w_1}}\right). \end{aligned} \tag{63}$$

Finally, when the Csiszár functional $D_f(i,N,v_1,v_2)$ is defined as in (49), that is by means of the Zipf law $N-$tuples, Theorem 1 and Theorem 2 assume the following form.

COROLLARY 8. *Let $f\colon \langle 0,\infty\rangle \to \mathbb{R}$ be a convex function and $v_1, v_2 > 0$. Suppose $\mathbf{t} = (t_1,\dots,t_N)$ and $\mathbf{u} = (u_1,\dots,u_N)$ are positive real $N-$tuples such that $T_N = \sum_{i=1}^N t_i$ and $U_N = \sum_{i=1}^N u_i$. If $\frac{1}{i^{v_2} H_{N,v_2}} \geqslant u_i$, for $i = 1,\dots,N$, then*

$$D_f(i,N,v_1,v_2) \geqslant f(1) + \sum_{i=1}^N u_i f\left(i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}\right) - U_N f\left(\frac{1}{U_N}\sum_{i=1}^N u_i i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}\right). \tag{64}$$

If $\frac{1}{i^{v_2} H_{N,v_2}} \leqslant t_i$, for $i = 1,\dots,N$, then

$$D_f(i,N,v_1,v_2) \leqslant f(1) + \sum_{i=1}^N t_i f\left(i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}\right) - T_N f\left(\frac{1}{T_N}\sum_{i=1}^N t_i i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}\right). \tag{65}$$

If f is a concave function, then reverse inequalities hold in (64) and (65).

Suppose $t \mapsto tf(t)$ is a convex function. If $\frac{1}{i^{v_2} H_{N,v_2}} \geqslant u_i$, for $i = 1,\dots,N$, then

$$\begin{aligned} D_{id\cdot f}(i,N,v_1,v_2) \geqslant{}& f(1) + \sum_{i=1}^N u_i i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}} f\left(i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}\right) \\ &- \left(\sum_{i=1}^N u_i i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}\right) f\left(\frac{1}{U_N}\sum_{i=1}^N u_i i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}\right), \end{aligned} \tag{66}$$

where $D_{id\cdot f}(i,N,v_1,v_2) := \sum_{i=1}^N \frac{1}{i^{v_1} H_{N,v_1}} f\left(i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}\right)$.

If $\frac{1}{i^{v_2} H_{N,v_2}} \leqslant t_i$, for $i = 1,\dots,N$, then

$$\begin{aligned} D_{id\cdot f}(i,N,v_1,v_2) \leqslant{}& f(1) + \sum_{i=1}^N t_i i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}} f\left(i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}\right) \\ &- \left(\sum_{i=1}^N t_i i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}\right) f\left(\frac{1}{T_N}\sum_{i=1}^N t_i i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}\right), \end{aligned} \tag{67}$$

If $t \mapsto tf(t)$ is a concave function, then reverse inequalities hold in (66) and (67).

Proof. Similarly as in Corollary 7, inequalities (64), (65), (66) and (67) follow from Theorem 1 by analogous steps, if we observe the probability mass functions r_i and s_i as Zipf laws defined by (10). □

REMARK 15. For $u_1 = u_2 = \cdots = u_N = \min_{i=1,\dots,N}\left\{\frac{1}{i^{v_2} H_{N,v_2}}\right\} = \frac{1}{N^{v_2} H_{N,v_2}}$ and $t_1 = t_2 = \cdots = t_N = \max_{i=1,\dots,N}\left\{\frac{1}{i^{v_2} H_{N,v_2}}\right\} = \frac{1}{H_{N,v_2}}$ inequalities (64) and (65) assume the form of the bounds that were obtained earlier in [8] :

$$f(1)+\frac{1}{H_{N,v_2}}\Lambda_3 \geqslant D_f(i,N,v_1,v_2) \geqslant f(1)+\frac{1}{N^{v_2}H_{N,v_2}}\ \Lambda_3, \tag{68}$$

where $\Lambda_3 = \sum_{i=1}^{N} f\left(i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}\right) - Nf\left(\frac{1}{N}\sum_{i=1}^{N} i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}\right)$.
If we repeat the similar procedure with inequalities (66) and (67), we get the analogous bounds for $D_{id\cdot f}(i,N,v_1,v_2)$:

$$f(1)+\frac{1}{H_{N,v_2}}\ \widetilde{\Lambda}_3 \geqslant D_{id\cdot f}(i,N,v_1,v_2) \geqslant f(1)+\frac{1}{N^{v_2}H_{N,v_2}}\ \widetilde{\Lambda}_3, \tag{69}$$

where

$$\widetilde{\Lambda}_3 = \sum_{i=1}^{N} i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}} f\left(i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}\right) - \left(\sum_{i=1}^{N} i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}\right) f\left(\frac{1}{N}\sum_{i=1}^{N} i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}\right). \tag{70}$$

Following the steps from the previous section, we again accompany the general results on $f-$divergence functionals defined via the Zipf-Mandelbrot law with some special choices on kernel function f. In the sequel we observe the Kullback-Leibler divergence (3), starting with the case where only one of two $N-$tuples $\mathbf{r}$ and $\mathbf{s}$ is defined via the Zipf-Mandelbrot law (8).

COROLLARY 9. *Let* $\mathbf{r}$ *be a positive probability distribution,* $\mathbf{t}$ *and* $\mathbf{u}$ *be as in Corollary 6 and* $v_2, w_2 > 0$*. If* $\frac{1}{(i+w_2)^{v_2}H_{N,v_2,w_2}} \geqslant u_i$*, for* $i = 1,\dots,N$*, then*

$$\begin{aligned} KL(\mathbf{r},i,N,v_2,w_2) \geqslant &\sum_{i=1}^{N} u_i r_i\,(i+w_2)^{v_2}\,H_{N,v_2,w_2}\log\left(r_i\,(i+w_2)^{v_2}\,H_{N,v_2,w_2}\right) \\ &-\left(\sum_{i=1}^{N} u_i r_i\,(i+w_2)^{v_2}\,H_{N,v_2,w_2}\right) \\ &\times\log\left(\frac{1}{U_N}\sum_{i=1}^{N} u_i r_i\,(i+w_2)^{v_2}\,H_{N,v_2,w_2}\right), \end{aligned} \tag{71}$$

where the logarithm base is greater than 1.
If $\frac{1}{(i+w_2)^{v_2}H_{N,v_2,w_2}} \leqslant t_i$*, for* $i = 1,\dots,N$*, then*

$$\begin{aligned} KL(\mathbf{r},i,N,v_2,w_2) \leqslant &\sum_{i=1}^{N} t_i r_i\,(i+w_2)^{v_2}\,H_{N,v_2,w_2}\log\left(r_i\,(i+w_2)^{v_2}\,H_{N,v_2,w_2}\right) \\ &-\left(\sum_{i=1}^{N} t_i r_i\,(i+w_2)^{v_2}\,H_{N,v_2,w_2}\right) \end{aligned}$$

$$\times \log \left(\frac{1}{T_N} \sum_{i=1}^{N} t_i r_i \left(i+w_2\right)^{v_2} H_{N,v_2,w_2} \right), \tag{72}$$

where the logarithm base is greater than 1.
If the logarithm base is less than 1, then reverse inequalities hold in (71) *and* (72).

Proof. Considering definition (3) of the Kullback-Leibler divergence with $R_N = \sum_{i=1}^{N} r_i = 1$, inequalities (71) and (72) follow from inequalities (52) and (53) of Corollary 6. Namely, function $t \mapsto t \log t$ is convex when the logarithm base is greater than 1 and is concave when the logarithm base is less than 1. □

REMARK 16. If we put $u_1 = u_2 = \cdots = u_N = \min_{i=1,\ldots,N} \left\{ \frac{1}{(i+w_2)^{v_2} H_{N,v_2,w_2}} \right\} = \frac{1}{(N+w_2)^{v_2} H_{N,v_2,w_2}}$ and at the same time $t_1 = t_2 = \cdots t_N = \max_{i=1,\ldots,N} \left\{ \frac{1}{(i+w_2)^{v_2} H_{N,v_2,w_2}} \right\} = \frac{1}{(1+w_2)^{v_2} H_{N,v_2,w_2}}$ into inequalities (71) and (72) we get the following bounds as a special case of Corollary 9. These were obtained earlier in [8]:

$$\frac{1}{(1+w_2)^{v_2} H_{N,v_2,w_2}} \Lambda_{KL}^1 \geqslant KL(\mathbf{r}, i, N, v_2, w_2) \geqslant \frac{1}{(N+w_2)^{v_2} H_{N,v_2,w_2}} \Lambda_{KL}^1,$$

where

$$\Lambda_{KL}^1 = \sum_{i=1}^{N} r_i \left(i+w_2\right)^{v_2} H_{N,v_2,w_2} \log \left(r_i \left(i+w_2\right)^{v_2} H_{N,v_2,w_2}\right) - \left(\sum_{i=1}^{N} r_i \left(i+w_2\right)^{v_2} H_{N,v_2,w_2} \right) \log \left(\frac{1}{N} \sum_{i=1}^{N} r_i \left(i+w_2\right)^{v_2} H_{N,v_2,w_2} \right). \tag{73}$$

We still deal with the Kullback-Leibler divergence (3) and proceed with the case of both $N-$tuples $\mathbf{r}$ and $\mathbf{s}$ being interpreted via the Zipf-Mandelbrot law (8).

COROLLARY 10. *Let* $\mathbf{t}$ *and* $\mathbf{u}$ *be as in Corollary* 7 *and let* $v_1, v_2, w_1, w_2 > 0$.
If $\frac{1}{(i+w_2)^{v_2} H_{N,v_2,w_2}} \geqslant u_i$, *for* $i = 1, \ldots, N$, *then*

$$KL(i, N, v_1, w_1, v_2, w_2) \geqslant \sum_{i=1}^{N} u_i \frac{(i+w_2)^{v_2} H_{N,v_2,w_2}}{(i+w_1)^{v_1} H_{N,v_1,w_1}} \log \frac{(i+w_2)^{v_2} H_{N,v_2,w_2}}{(i+w_1)^{v_1} H_{N,v_1,w_1}} - \left(\sum_{i=1}^{N} u_i \frac{(i+w_2)^{v_2} H_{N,v_2,w_2}}{(i+w_1)^{v_1} H_{N,v_1,w_1}} \right) \times \log \left(\frac{1}{U_N} \sum_{i=1}^{N} u_i \frac{(i+w_2)^{v_2} H_{N,v_2,w_2}}{(i+w_1)^{v_1} H_{N,v_1,w_1}} \right), \tag{74}$$

where the logarithm base is greater than 1.
If $\frac{1}{(i+w_2)^{v_2} H_{N,v_2,w_2}} \leqslant t_i$, *for* $i = 1,\ldots,N$, *then*

$$KL(i,N,v_1,w_1,v_2,w_2) \leqslant \sum_{i=1}^{N} t_i \frac{(i+w_2)^{v_2} H_{N,v_2,w_2}}{(i+w_1)^{v_1} H_{N,v_1,w_1}} \log \frac{(i+w_2)^{v_2} H_{N,v_2,w_2}}{(i+w_1)^{v_1} H_{N,v_1,w_1}} - \left(\sum_{i=1}^{N} t_i \frac{(i+w_2)^{v_2} H_{N,v_2,w_2}}{(i+w_1)^{v_1} H_{N,v_1,w_1}} \right) \times \log \left(\frac{1}{T_N} \sum_{i=1}^{N} u_i \frac{(i+w_2)^{v_2} H_{N,v_2,w_2}}{(i+w_1)^{v_1} H_{N,v_1,w_1}} \right), \tag{75}$$

where the logarithm base is greater than 1.
If the logarithm base is less than 1, then reverse inequalities hold in (74) *and* (75).

Proof. Inequalities (74) and (75) follow from inequalities (59) and (60) of Corollary 7 when observing the function $t \mapsto t \log t$ which is convex when the logarithm base is greater than 1 and is concave when the logarithm base is less than 1. □

REMARK 17. If we again include in (74) and in (75) that $u_1 = u_2 = \cdots = u_N = \min_{i=1,\ldots,N} \left\{ \frac{1}{(i+w_2)^{v_2} H_{N,v_2,w_2}} \right\} = \frac{1}{(N+w_2)^{v_2} H_{N,v_2,w_2}}$ and at the same time $t_1 = t_2 = \cdots = t_N = \max_{i=1,\ldots,N} \left\{ \frac{1}{(i+w_2)^{v_2} H_{N,v_2,w_2}} \right\} = \frac{1}{(1+w_2)^{v_2} H_{N,v_2,w_2}}$, then the following bounds for the Kullback-Leibler divergence hold, as a special case of Corollary 10:

$$\frac{1}{(1+w_2)^{v_2} H_{N,v_2,w_2}} \Lambda_{KL}^2 \geqslant KL(i,N,v_1,w_1,v_2,w_2) \geqslant \frac{1}{(N+w_2)^{v_2} H_{N,v_2,w_2}} \Lambda_{KL}^2, \tag{76}$$

where

$$\Lambda_{KL}^2 = \sum_{i=1}^{N} \frac{(i+w_2)^{v_2} H_{N,v_2,w_2}}{(i+w_1)^{v_1} H_{N,v_1,w_1}} \log \frac{(i+w_2)^{v_2} H_{N,v_2,w_2}}{(i+w_1)^{v_1} H_{N,v_1,w_1}} - \left(\sum_{i=1}^{N} \frac{(i+w_2)^{v_2} H_{N,v_2,w_2}}{(i+w_1)^{v_1} H_{N,v_1,w_1}} \right) \log \left(\frac{1}{N} \sum_{i=1}^{N} \frac{(i+w_2)^{v_2} H_{N,v_2,w_2}}{(i+w_1)^{v_1} H_{N,v_1,w_1}} \right). \tag{77}$$

Bounds (76) were obtained earlier in [8], due to a less general approach.

The Kullback-Leibler divergence is observed once more in the case of both $N-$tuples $\mathbf{r}$ and $\mathbf{s}$ being interpreted via the Zipf law (10), *i.e.* for $w_1 = w_2 = 0$.

COROLLARY 11. *Let* $\mathbf{t}$ *and* $\mathbf{u}$ *be as in Corollary* 8 *and let* $v_1, v_2 > 0$.
If $\frac{1}{i^{v_2} H_{N,v_2}} \geqslant u_i$, *for* $i = 1,\ldots,N$, *then*

$$KL(i,N,v_1,v_2) \geqslant \sum_{i=1}^{N} u_i i^{v_2 - v_1} \frac{H_{N,v_2}}{H_{N,v_1}} \log \left(i^{v_2 - v_1} \frac{H_{N,v_2}}{H_{N,v_1}} \right)$$

$$-\left(\sum_{i=1}^{N} u_i i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}\right)\log\left(\frac{1}{U_N}\sum_{i=1}^{N} u_i i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}\right), \tag{78}$$

where the logarithm base is greater than 1.
If $\frac{1}{i^{v_2}H_{N,v_2}} \leqslant t_i$, *for* $i=1,\ldots,N$, *then*

$$KL(i,N,v_1,v_2) \leqslant \sum_{i=1}^{N} t_i i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}\log\left(i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}\right)$$
$$-\left(\sum_{i=1}^{N} t_i i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}\right)\log\left(\frac{1}{T_N}\sum_{i=1}^{N} u_i i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}\right), \tag{79}$$

where the logarithm base is greater than 1.
If the logarithm base is less than 1, then reverse inequalities hold in (78) *and* (79).

Proof. Inequalities (78) and (79) follow from inequalities (66) and (67) of Corollary 8 when observing the function $t \mapsto t\log t$ which is convex when the logarithm base is greater than 1 and is concave when the logarithm base is less than 1. □

REMARK 18. If $u_1 = u_2 = \cdots = u_N = \min_{i=1,\ldots,N}\left\{\frac{1}{i^{v_2}H_{N,v_2}}\right\} = \frac{1}{N^{v_2}H_{N,v_2}}$ and $t_1 = t_2 = \cdots = t_N = \max_{i=1,\ldots,N}\left\{\frac{1}{i^{v_2}H_{N,v_2}}\right\} = \frac{1}{H_{N,v_2}}$ in inequalities (78) and (79), then the following bounds for the Kullback-Leibler divergence hold, as a special case of Corollary 11:

$$\frac{1}{H_{N,v_2}}\Lambda_{KL}^3 \geqslant KL(i,N,v_1,v_2) \geqslant \frac{1}{N^{v_2}H_{N,v_2}}\Lambda_{KL}^3, \tag{80}$$

where

$$\Lambda_{KL}^3 = \sum_{i=1}^{N} i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}\log\left(i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}\right)$$
$$-\left(\sum_{i=1}^{N} i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}\right)\log\left(\frac{1}{N}\sum_{i=1}^{N} i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}\right). \tag{81}$$

Bounds (80) were obtained earlier in [8], due to a less general approach.

In the sequel we present similar results for other divergences that were observed in Section 2: the Hellinger distance, the Bhattacharyya coefficient, the chi-square divergence and the total variation distance. In comparison with the ones related to the Kullback-Leibler divergence, these are more concise with an accent put on the implementation of the Zipf law alone, that is for $w_1 = w_2 = 0$.

COROLLARY 12. *Let* $\mathbf{t}$ *and* $\mathbf{u}$ *be as in Corollary* 8 *and let* $v_1, v_2 > 0$. *If* $\frac{1}{i^{v_2} H_{N,v_2}} \geqslant u_i$, *for* $i = 1, \ldots, N$, *then*

$$h^2(i,N,v_1,v_2) \geqslant \frac{1}{2}\sum_{i=1}^{N} u_i\left(\sqrt{i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}}-1\right)^2 - \frac{U_N}{2}\left(\sqrt{\frac{1}{U_N}\sum_{i=1}^{N} u_i i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}}-1\right)^2, \quad (82)$$

$$B(i,N,v_1,v_2) \leqslant 1 + \sum_{i=1}^{N} u_i\sqrt{i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}} - \sqrt{U_N\sum_{i=1}^{N} u_i i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}}, \quad (83)$$

$$\chi^2(i,N,v_1,v_2) \geqslant \sum_{i=1}^{N} u_i\left(i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}-1\right)^2 - U_N\left(\frac{1}{U_N}\sum_{i=1}^{N} u_i i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}-1\right)^2, \quad (84)$$

$$V(i,N,v_1,v_2) \geqslant \sum_{i=1}^{N} u_i\left|i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}-1\right| - U_N\left|\frac{1}{U_N}\sum_{i=1}^{N} u_i i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}-1\right|. \quad (85)$$

If $\frac{1}{i^{v_2} H_{N,v_2}} \leqslant t_i$, *for* $i = 1, \ldots, N$, *then*

$$h^2(i,N,v_1,v_2) \leqslant \frac{1}{2}\sum_{i=1}^{N} t_i\left(\sqrt{i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}}-1\right)^2 - \frac{T_N}{2}\left(\sqrt{\frac{1}{T_N}\sum_{i=1}^{N} t_i i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}}-1\right)^2, \quad (86)$$

$$B(i,N,v_1,v_2) \geqslant 1 + \sum_{i=1}^{N} t_i\sqrt{i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}} - \sqrt{T_N\sum_{i=1}^{N} t_i i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}}, \quad (87)$$

$$\chi^2(i,N,v_1,v_2) \leqslant \sum_{i=1}^{N} t_i\left(i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}-1\right)^2 - T_N\left(\frac{1}{T_N}\sum_{i=1}^{N} t_i i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}-1\right)^2, \quad (88)$$

$$V(i,N,v_1,v_2) \leqslant \sum_{i=1}^{N} t_i\left|i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}-1\right| - T_N\left|\frac{1}{T_N}\sum_{i=1}^{N} t_i i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}-1\right|. \quad (89)$$

Proof. Inequalities (82)-(85) and (86)-(89) follow from inequalities (64) and (65) respectively of Corollary 8 when observing the convex functions: $t \mapsto \frac{1}{2}\left(\sqrt{t}-1\right)^2$ for (82) and (86), $t \mapsto -\sqrt{t}$ for (83) and (87), $t \mapsto (t-1)^2$ for (84) and (88) and $t \mapsto |t-1|$ for (85) and (89). □

REMARK 19. If $u_1 = u_2 = \cdots = u_N = \min_{i=1,\ldots,N}\left\{\frac{1}{i^{v_2} H_{N,v_2}}\right\} = \frac{1}{N^{v_2} H_{N,v_2}}$ in inequalities (82)-(85) and $t_1 = t_2 = \cdots = t_N = \max_{i=1,\ldots,N}\left\{\frac{1}{i^{v_2} H_{N,v_2}}\right\} = \frac{1}{H_{N,v_2}}$ in inequalities (86)-(89), then the following bounds for the divergences hold, as special cases of Corollary 12.

Thus we have for Hellinger distance (4):

$$\frac{1}{2H_{N,v_2}}\Lambda_h \geqslant h^2(i,N,v_1,v_2) \geqslant \frac{1}{2N^{v_2} H_{N,v_2}}\Lambda_h, \quad (90)$$

where

$$\Lambda_h = \sum_{i=1}^{N}\left(\sqrt{i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}} - 1\right)^2 - N\left(\sqrt{\frac{1}{N}\sum_{i=1}^{N} i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}} - 1\right)^2; \tag{91}$$

for Bhattacharyya coefficient (5):

$$1 - \frac{1}{N^{v_2}H_{N,v_2}}\Lambda_B \geqslant B(i,N,v_1,v_2) \geqslant 1 - \frac{1}{H_{N,v_2}}\Lambda_B, \tag{92}$$

where

$$\Lambda_B = N\sqrt{\frac{1}{N}\sum_{i=1}^{N} i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}} - \sum_{i=1}^{N}\sqrt{i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}}}; \tag{93}$$

for chi-square divergence (6):

$$\frac{1}{H_{N,v_2}}\Lambda_{chi} \geqslant \chi^2(i,N,v_1,v_2) \geqslant \frac{1}{N^{v_2}H_{N,v_2}}\Lambda_{chi}, \tag{94}$$

where

$$\Lambda_{chi} = \sum_{i=1}^{N}\left(i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}} - 1\right)^2 - N\left(\frac{1}{N}\sum_{i=1}^{N} i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}} - 1\right)^2; \tag{95}$$

and for total variation distance (7):

$$\frac{1}{H_{N,v_2}}\Lambda_V \geqslant V(i,N,v_1,v_2) \geqslant \frac{1}{N^{v_2}H_{N,v_2}}\Lambda_V, \tag{96}$$

where

$$\Lambda_V = \sum_{i=1}^{N}\left|i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}} - 1\right| - N\left|\frac{1}{N}\sum_{i=1}^{N} i^{v_2-v_1}\frac{H_{N,v_2}}{H_{N,v_1}} - 1\right|. \tag{97}$$

Bounds (90), (92), (94) and (96) were also obtained earlier in [8], due to a less general approach.

REFERENCES

[1] I. CSISZÁR, *Information-type measures of difference of probability functions and indirect observations*, Studia Sci. Math. Hungar., **2** (1967), 299–318.
[2] I. CSISZÁR, *Information measures: A critical survey*, Trans. 7th Prague Conf. on Info. Th. Statist. Decis. Funct., Random Processes and 8th European Meeting of Statist., Volume B, Academia Prague, 1978, 73–86.
[3] S. S. DRAGOMIR, *Some inequalities for the Csiszár Φ− divergence*, RGMIA.
[4] L. EGGHE, R. ROUSSEAU, *Introduction to Informetrics. Quantitative Methods in Library, Documentation and Information Scince*, Elsevier Science Publishers, New York, (1990).

[5] T. VAN ERVEN, P. HARREMOËS, *Rényi Divergence and Kullback-Leibler Divergence*, Journal of Latex Class Files, **6** 1 (2007); arXiv: 1206.2459v2 [cs.IT]24 April 2014.
[6] S. KULLBACK, *Information Theory and Statistics*, J.Wiley, New York, 1959.
[7] S. KULLBACK, R. A. LEIBLER, *On information and sufficiency*, Annals of Mathematical Statistics, **22** 1 (1951), 79–86.
[8] N. LOVRIČEVIĆ, Ð. PEČARIĆ, J. PEČARIĆ, *Zipf-Mandelbrot law, $f-$divergences and the Jensen-type interpolating inequalities*, J. Inequal. Appl. (2018) 2018:36 https://doi.org/10.1186/s13660-018-1625-y
[9] B. MANARIS, D. VAUGHAN, C. S. WAGNER, J. ROMERO, R. B. DAVIS, *Evolutionary Music and the Zipf-Mandelbrot Law: Developing Fitness Functions for Pleasant Music*, Proceedings of 1st European Workshop on Evolutionary Music and Art (EvoMUSART2003), 522–534.
[10] B. MANDELBROT, *An informational theory of the statistical structure of language*, In Jackson, W., editor, Communication Theory, New York, Academic Press, 1953., 486–502.
[11] B. MANDELBROT, *Information Theory and Psycholinguistics*, in Wolman, B.B., and Nagel, E. (eds.), *Scientific psychology*, Basic Books, 1965.
[12] D. YU. MANIN, *Mandelbrot's Model for Zipf's Law: Can Mandelbrot's Model Explain Zipf's Law for Language?*, Journal of Quantitative Linguistics, **16** 3 (2009), 274–285.
[13] D. S. MITRINOVIĆ, J. E. PEČARIĆ, A. M. FINK, *Classical and New Inequalities in Analysis*, Kluwer Academic Publishers, Dordrecht/Boston/London, 1993.
[14] M. A. MONTEMURRO, *Beyond the Zipf-Mandelbrot law in quantitative linguistics*, URL:arxiv:condmat/0104066v2, (2001).
[15] D. MOUILLOT, A. LEPRETRE, *Introduction of relative abundance distribution (RAD) indices, estimated from the rank-frequency diagrams(RFD), to assess changes in community diversity*, Environmental Monitoring and Assessment, Springer **63** (2) (2000), 279–295.
[16] Z. K. SILAGADZE, *Citations and the Zipf-Mandelbrot Law*, Complex Systems, **11** (1997), 487–499.

(Received November 1, 2018)

Neda Lovričević
Faculty of Civil Engineering, Architecture and Geodesy
University of Split
Matice hrvatske 15, 21000 Split, Croatia
e-mail: neda.lovricevic@gradst.hr

Ðilda Pečarić
Catholic University of Croatia
Ilica 242, 10000 Zagreb, Croatia
e-mail: gildapeca@gmail.com

Josip Pečarić
RUDN University
Miklukho-Maklaya str. 6, 117198 Moscow, Russia
e-mail: pecaric@element.hr