axioms

MDPI

Article

Exact Remainder and Two-Sided Bounds for the Harmonic Jensen Functional with Applications to f-Divergences

Josipa Barić [1], Slavica Ivelić Bradanović [2,*] and Neda Lovričević [2]

[1] Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture, University of Split, 21000 Split, Croatia; jbaric@fesb.hr
[2] Faculty of Civil Engineering, Architecture and Geodesy, University of Split, 21000 Split, Croatia; nblagaic@gradst.hr
* Correspondence: sivelic@gradst.hr

Abstract

We study the harmonic Jensen functional of harmonically convex functions. We give lower and upper bounds obtained from the support inequality and an identity for the difference between the functional and its lower bound. This identity characterizes the equality cases. In particular, equality holds for monotone functions. Its integral form gives a refinement of the harmonic Hermite–Hadamard inequality. Further, we obtain sharp two-sided bounds under the assumption that $tf(t) - \frac{m}{2}t^2$ is convex for a real parameter m. For $m < 0$, this assumption covers some functions which are not harmonically convex. For twice differentiable harmonically convex functions, the upper bound improves the bound of Dragomir. The results are applied to Csiszár f-divergences, where the harmonic mean of the ratios is equal to $(1 + \chi^2(\mathbf{p}, \mathbf{q}))^{-1}$. The bounds are illustrated for the Kullback–Leibler, squared Hellinger and Jeffreys divergences and compared with the Lah–Ribarič bound. An application to differential entropy is also given.

Keywords: harmonic convexity; Jensen functional; Hermite–Hadamard inequality; Bullen inequality; harmonic expectation; Csiszár f-divergence; χ^2-divergence; differential entropy

MSC: 26A51; 26D15; 26D20; 60E15; 62B10; 94A15; 94A17

1. Introduction and Preliminaries

Convex functions appear in almost every part of analysis and numerous generalizations of them have been introduced in the literature. We start by recalling that $f : I \to \mathbb{R}$, where $I \subseteq \mathbb{R}$ is an interval, is called convex if

$$f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y)$$

holds for arbitrary $x, y \in I$ and $0 \leq \lambda \leq 1$. Among the tools most often used when working with convex functions, two inequalities stand out (see [1,2]). According to Jensen's inequality, whenever $x_1, \ldots, x_n \in I$ and the nonnegative weights $w_1, \ldots, w_n$ sum up to one, every convex $f : I \to \mathbb{R}$ satisfies

$$f\left(\sum_{i=1}^{n} w_i x_i\right) \leq \sum_{i=1}^{n} w_i f(x_i). \tag{1}$$

Check for updates

Academic Editors: Alina Sintamarian, Ovidiu Furdui and Cristina Flaut

Received: 9 October 2026
Revised: 6 November 2026
Accepted: 11 November 2026
Published: 17 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

The other one, due to Hermite and Hadamard, compares the integral average of a convex $f : I \to \mathbb{R}$ on $[a, b] \subseteq I$ ($a < b$) with its value at the midpoint and with the arithmetic mean of $f(a)$ and $f(b)$:

$$f\left(\frac{a+b}{2}\right) \leq \frac{1}{b-a}\int_a^b f(x)\,\mathrm{d}x \leq \frac{f(a)+f(b)}{2}. \tag{2}$$

Given $x, y > 0$ and $\lambda \in [0, 1]$, the numbers

$$H_\lambda(x, y) = \left(\frac{1-\lambda}{x} + \frac{\lambda}{y}\right)^{-1} = \frac{xy}{(1-\lambda)y + \lambda x}$$

and

$$A_\lambda(x, y) = (1-\lambda)x + \lambda y$$

are called the weighted harmonic and the weighted arithmetic mean of x and y, respectively. By Jensen's inequality for the convex function $\varphi(t) = \frac{1}{t}$, $t > 0$, one always has $H_\lambda(x, y) \leq A_\lambda(x, y)$. If we put $H_\lambda(x, y)$ in place of $A_\lambda(x, y)$ in the argument of f in the convexity condition, we arrive at another important class of functions.

It was İşcan [3] who introduced this class. Let $I \subset (0, \infty)$ be an interval. We call $f : I \to \mathbb{R}$ harmonically convex provided that

$$f\big(H_\lambda(x, y)\big) \leq (1-\lambda)f(x) + \lambda f(y) \tag{3}$$

whenever $x, y \in I$ and $0 \leq \lambda \leq 1$.

An important feature of harmonic convexity is that the substitution $u = \frac{1}{x}$ reduces it to ordinary convexity. This fact allows us to transfer many results from the classical theory of convex functions to the harmonic setting. For an interval $I \subset (0, \infty)$, we put $I^{-1} = \left\{\frac{1}{x} : x \in I\right\}$.

Lemma 1. *Let f be a real function on an interval $I \subset (0, \infty)$. Then the following properties of f are equivalent:*

(i) *harmonic convexity of f on I,*

(ii) *convexity of $g : I^{-1} \to \mathbb{R}$, $g(u) = f\left(\frac{1}{u}\right)$,*

(iii) *convexity of $\psi : I \to \mathbb{R}$, $\psi(x) = xf(x)$.*

For $f \in C^2(I)$, each of them is also equivalent to

$$xf''(x) + 2f'(x) \geq 0, \qquad x \in I. \tag{4}$$

The equivalence (i)⇔(iii) is due to Dragomir (Theorem 4 in [4]). That (i) and (ii) describe the same property is clear from (3), while (4) is obtained from the second derivative of $g(u) = f\left(\frac{1}{u}\right)$.

Lemma 1 is used throughout the paper. By (ii), the harmonic Jensen functional $\mathcal{J}(f)$ introduced in (11) below is the ordinary Jensen functional of the convex function g at the points $u_i = \frac{1}{x_i}$. The arithmetic mean of these points is $\sum_{i=1}^n w_i u_i = \frac{1}{\bar{x}_H}$, i.e., it corresponds to the harmonic mean of the points x_i. Therefore the supporting lines of g at $\frac{1}{\bar{x}_H}$ and at $\frac{1}{x_i}$ give the lower and the upper bounds in Section 2. The equality $\sum_{i=1}^n w_i\left(\frac{1}{x_i} - \frac{1}{\bar{x}_H}\right) = 0$ makes their linear parts cancel. Condition (iii) and its differential form (4) provide a practical test of harmonic convexity. It leads to the parameter condition in Theorem 4 and to the intervals of harmonic convexity in the examples of Section 3. Figure 1 illustrates this correspondence.

https://doi.org/10.3390/axioms15110858

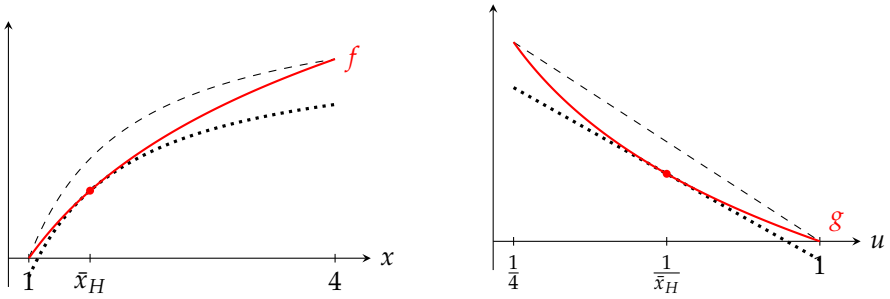


Figure 1. (**Left**): the graph of the harmonically convex function $f(x) = \ln x$ on $[1,4]$ (solid) lies below the harmonic chord $x \mapsto \frac{\alpha}{x} + \beta$ through its endpoints (dashed) and above the support curve $x \mapsto f(\bar{x}_H) + C_{\bar{x}_H}\left(\frac{1}{x} - \frac{1}{\bar{x}_H}\right)$ from (5) (dotted), where $\bar{x}_H = 1.6$ is the harmonic mean of 1 and 4. (**Right**): after the substitution $u = \frac{1}{x}$, the same objects become the convex function $g(u) = f\left(\frac{1}{u}\right) = -\ln u$ on $\left[\frac{1}{4}, 1\right]$, its chord and its tangent line at $\frac{1}{\bar{x}_H}$.

A related notion, reciprocal convexity, was studied earlier by Merkle [5]. He called $f : (0,\infty) \to \mathbb{R}$ reciprocally convex when

(a) f is a concave function and

(b) $k(x) = f\left(\frac{1}{x}\right)$, $x > 0$, defines a convex function.

In view of Lemma 1, (b) is nothing else but harmonic convexity of f. Consequently, reciprocally convex functions are just those harmonically convex functions which are, in addition, concave. However, the requirement of concavity excludes many important functions. Since harmonic convexity does not impose such a requirement, harmonically convex functions form a larger family. A simple example is $f(x) = \frac{1}{x^2}$. It is not concave, but it is harmonically convex on $(0,\infty)$ because $f\left(\frac{1}{x}\right) = x^2$ is convex.

Harmonically convex functions are also important in applications. The harmonic mean appears when rates or ratios are averaged, for example in the calculation of average speeds, average prices per unit or in comparing two probability distributions through the ratios of their probabilities. In such situations, it is natural to put the harmonic mean instead of the arithmetic one in the argument of f. Moreover, some important functions are concave, but harmonically convex. For example, the logarithmic function is concave, while the function $\psi(x) = x \ln x$ is convex, so by Lemma 1 the logarithm is harmonically convex. For such functions, Jensen's inequality (1) holds in the reverse direction, while its harmonic analogue gives an estimate of the usual type.

In the literature, harmonically convex functions are often called HA-convex (harmonic-arithmetic convex) functions (see [3,6]). In the last ten years, a large number of papers dealt with this notion and proposed its generalizations. We mention the strong [7], the s- and the h-variants [8,9] of harmonic convexity, reciprocal (s,m)-convexity [10], strong reciprocal (p,h)-convexity [11], uniform harmonic convexity [12] and several other related concepts [13–15].

Many of the corresponding inequalities were also transferred to the stochastic, quantum, fuzzy and fractional settings and some of them found applications in the estimation of entropy (see [16–22]).

The next lemma collects known facts about harmonically convex functions, which we use later (see [3–5,23,24]).

Lemma 2. *Let f be harmonically convex on I, where $I \subset (0,\infty)$ is an open interval. At every $x \in I$, the left and the right derivative of f are finite. Moreover, to each $x \in I$ corresponds a number $C_x \in \mathbb{R}$ with the property*

$$f(y) \geq f(x) + C_x\left(\frac{1}{y} - \frac{1}{x}\right), \qquad y \in I. \tag{5}$$

Writing $\partial_H f(x)$ for the collection of all such numbers C_x, we have

$$\partial_H f(x) = \left[-x^2 f'_+(x), -x^2 f'_-(x)\right]. \tag{6}$$

At points of differentiability, $\partial_H f(x)$ reduces to the single number

$$C_x = -x^2 f'(x). \tag{7}$$

Finally, assume that f is differentiable on an open interval $J \subset (0,\infty)$ and harmonically convex on a subinterval $I_0 \subseteq J$. Then

$$f(y) \geq f(x) - x^2 f'(x)\left(\frac{1}{y} - \frac{1}{x}\right), \qquad x, y \in I_0. \tag{8}$$

Inequality (5) was stated for differentiable functions in Lemma 10 in [4] (see also [23,24]). In general, one uses a supporting line of $g(u) = f\left(\frac{1}{u}\right)$, which is convex, at $u = \frac{1}{x}$. To prove (8), note that g is differentiable on J^{-1} and convex on I_0^{-1}. Since the slopes of the chords of a convex function are nondecreasing, we have

$$g(u) \geq g(s) + g'(s)(u - s)$$

for all $s, u \in I_0^{-1}$, including the endpoints of I_0^{-1}.

Our considerations start from the analogue of Jensen's inequality in the harmonic setting, obtained by Dragomir [6]. Let the points $x_1, \ldots, x_n$ belong to an interval $I \subset (0,\infty)$ and let the weights $w_1, \ldots, w_n \geq 0$ satisfy $\sum_{i=1}^n w_i = 1$. We denote their weighted harmonic mean by

$$\bar{x}_H = \left(\sum_{i=1}^n \frac{w_i}{x_i}\right)^{-1}. \tag{9}$$

If f is harmonically convex on I, then

$$f(\bar{x}_H) \leq \sum_{i=1}^n w_i f(x_i). \tag{10}$$

Clearly, (10) is (1) for the convex function $g(u) = f\left(\frac{1}{u}\right)$ at the points $\frac{1}{x_i}$. Equality in (10) occurs, e.g., for $f(x) = \frac{\alpha}{x} + \beta$ with real α, β. The difference

$$\mathcal{J}(f) = \sum_{i=1}^n w_i f(x_i) - f(\bar{x}_H) \tag{11}$$

is called the harmonic Jensen functional. By (10), $\mathcal{J}(f) \geq 0$ for every harmonically convex function f. Finding good estimates for $\mathcal{J}(f)$ is the main goal of this paper.

A harmonic counterpart of (2) is due to İşcan [3]. If f is harmonically convex on I and $a < b$ are points of I, then

$$f\left(\frac{2ab}{a+b}\right) \leq \frac{ab}{b-a}\int_a^b \frac{f(x)}{x^2}\,\mathrm{d}x \leq \frac{f(a)+f(b)}{2}. \tag{12}$$

To obtain (12), it suffices to apply (2) to the convex function $g(u) = f\left(\frac{1}{u}\right)$ on $\left[\frac{1}{b}, \frac{1}{a}\right]$.

In [20] (see also [25]), Sayyari et al. considered HA-convex functions and their uniform variant, proved estimates of Jensen and Simić kind and used them to bound Shannon entropy. According to Kian [26], properties of the Csiszár f-divergence are closely related to the convexity type of f. HA-convexity was one of the types discussed there. Our setting differs from [20], since the Csiszár functionals we work with depend on two probability distributions. Our main tool is inequality (5), which we apply to the harmonic mean as well as to each of the given points. A similar approach was used in [27] for strongly convex functions.

Related results for differentiable convex functions, namely a refinement of Jensen's inequality and its converse, were obtained in [28] and [29], respectively. Theorems 1 and 2 give these estimates for harmonically convex functions without the assumption of differentiability. To the best of our knowledge, the identity in Theorem 3 is new also for convex functions (Remark 2). Its integral form is given in Theorem 7. In [27,30], strongly convex functions and the arithmetic Jensen functional were considered, while here the condition in Theorem 4 allows also negative values of m.

We point out the main results of the paper. First, we find the exact difference between the harmonic Jensen functional and one of its lower bounds. This allows us to describe precisely when the lower bound is equal to the functional. In particular, this is the case for monotone functions. Second, we obtain two-sided bounds under a condition which depends on a real parameter. For positive values of the parameter, this condition implies strong harmonic convexity, while for negative values it also includes functions which are not harmonically convex. For twice differentiable harmonically convex functions, the upper bound obtained in this way is never worse than the upper bound of Dragomir [6]. It is also compared with the Lah–Ribarič bound. Third, in the application to Csiszár divergences, the harmonic mean which appears in our estimates is determined by the chi-square divergence. Therefore, the chi-square divergence appears in all the obtained bounds. With the optimal value of the parameter, both bounds for the divergence are asymptotically sharp as $\mathbf{q} \to \mathbf{p}$ (Remark 10), which is not the case for the upper bound of Dragomir.

We now describe the content of the remaining sections. Section 2 deals with estimates of the harmonic Jensen functional from below and from above. Section 3 contains applications to Csiszár f-divergences, together with numerical examples. In Section 4, we refine the harmonic Hermite–Hadamard inequality and give the integral form of the identity from Section 2. As an illustration, we apply this refinement to the logarithmic function and compare the obtained bounds for differential entropy with its exact value. Section 5 contains concluding remarks.

2. Bounds for the Harmonic Jensen Functional

Throughout Section 2, the points $x_1, \dots, x_n$ lie in an open interval $I \subset (0, \infty)$ and the weights $w_1, \dots, w_n$ are nonnegative with $\sum_{i=1}^{n} w_i = 1$. Clearly, points with zero weight do not affect the results and they can be omitted.

As in Section 1, $\bar{x}_H$ is their weighted harmonic mean (9) and $\mathcal{J}(f)$ is the harmonic Jensen functional (11).

First we estimate the harmonic Jensen functional from below by using (5) at the point $\bar{x}_H$ and then from above by using the same inequality at the points x_i. The choice of $C_x \in \partial_H f(x)$ in the results below is arbitrary.

Theorem 1. *Let $C_{\bar{x}_H} \in \partial_H f(\bar{x}_H)$, where $f : I \to \mathbb{R}$ is harmonically convex. Then*

$$\begin{aligned} 0 &\le \left| \sum_{i=1}^{n} w_i \left(|f(x_i) - f(\bar{x}_H)| - \left| C_{\bar{x}_H} \left(\frac{1}{x_i} - \frac{1}{\bar{x}_H} \right) \right| \right) \right| \\ &\le \sum_{i=1}^{n} w_i \left| |f(x_i) - f(\bar{x}_H)| - \left| C_{\bar{x}_H} \left(\frac{1}{x_i} - \frac{1}{\bar{x}_H} \right) \right| \right| \\ &\le \mathcal{J}(f). \end{aligned} \tag{13}$$

Proof. For $i = 1, \ldots, n$, let

$$A_i = f(x_i) - f(\bar{x}_H), \qquad B_i = C_{\bar{x}_H} \left(\frac{1}{x_i} - \frac{1}{\bar{x}_H} \right).$$

By (5) at the point $\bar{x}_H$, we have $A_i \ge B_i$. Hence the reverse triangle inequality gives

$$\big| |A_i| - |B_i| \big| \le |A_i - B_i| = A_i - B_i. \tag{14}$$

Now we multiply (14) by w_i and add the resulting inequalities. This gives

$$\sum_{i=1}^{n} w_i \big| |A_i| - |B_i| \big| \le \sum_{i=1}^{n} w_i (A_i - B_i).$$

Since

$$\sum_{i=1}^{n} \frac{w_i}{x_i} = \frac{1}{\bar{x}_H},$$

the right-hand side is equal to

$$\begin{aligned} \sum_{i=1}^{n} w_i (A_i - B_i) &= \sum_{i=1}^{n} w_i f(x_i) - f(\bar{x}_H) - C_{\bar{x}_H} \sum_{i=1}^{n} w_i \left(\frac{1}{x_i} - \frac{1}{\bar{x}_H} \right) \\ &= \mathcal{J}(f). \end{aligned}$$

Thus the rightmost inequality in (13) is proved. For the middle one, it suffices to use the triangle inequality

$$\left| \sum_{i=1}^{n} w_i z_i \right| \le \sum_{i=1}^{n} w_i |z_i|,$$

and the first one is trivial. □

Remark 1. *Let $u_i = \frac{1}{x_i}$, $i = 1, \ldots, n$, $\bar{u} = \sum_{i=1}^{n} w_i u_i = \frac{1}{\bar{x}_H}$ and $g(u) = f\left(\frac{1}{u}\right)$. Since $\left| C_{\bar{x}_H} \left(\frac{1}{x_i} - \frac{1}{\bar{x}_H} \right) \right| = |C_{\bar{x}_H}||u_i - \bar{u}|$, the weaker lower bound in (13) takes the form*

$$\left| \sum_{i=1}^{n} w_i |g(u_i) - g(\bar{u})| - |C_{\bar{x}_H}| \sum_{i=1}^{n} w_i |u_i - \bar{u}| \right| \le \sum_{i=1}^{n} w_i g(u_i) - g(\bar{u}).$$

Theorem 1 holds for every $C_{\bar{x}_H} \in \partial_H f(\bar{x}_H)$ and also gives the stronger lower bound. A related lower bound for differentiable convex functions was given in Theorem 1 in [28]. We include the weaker lower bound since its integral form is calculated explicitly for the logarithmic function in Section 4.

Theorem 2. *Let $C_{x_i} \in \partial_H f(x_i)$, $i = 1, \ldots, n$, where $f : I \to \mathbb{R}$ is harmonically convex. Then*

$$\mathcal{J}(f) \le \sum_{i=1}^{n} w_i C_{x_i} \left(\frac{1}{x_i} - \frac{1}{\bar{x}_H} \right). \tag{15}$$

Proof. Let $i \in \{1, \dots, n\}$. Inequality (5) at the point x_i, with $y = \bar{x}_H$, gives

$$f(\bar{x}_H) \geq f(x_i) + C_{x_i}\left(\frac{1}{\bar{x}_H} - \frac{1}{x_i}\right),$$

i.e.,

$$f(x_i) - f(\bar{x}_H) \leq C_{x_i}\left(\frac{1}{x_i} - \frac{1}{\bar{x}_H}\right).$$

We multiply this inequality by w_i and add up for $i = 1, \dots, n$. Since $\sum_{i=1}^{n} w_i = 1$, this gives (15). □

Combining Theorems 1 and 2, we obtain a chain of inequalities for the harmonic Jensen functional $\mathcal{J}(f)$. It refines the lower bound $\mathcal{J}(f) \geq 0$ from (10) and complements it with an upper bound. For $C_{\bar{x}_H} \in \partial_H f(\bar{x}_H)$ and $C_{x_i} \in \partial_H f(x_i)$, let

$$\mathcal{L}(f) = \left|\sum_{i=1}^{n} w_i\left(|f(x_i) - f(\bar{x}_H)| - \left|C_{\bar{x}_H}\left(\frac{1}{x_i} - \frac{1}{\bar{x}_H}\right)\right|\right)\right|, \tag{16}$$

$$\mathcal{R}(f) = \sum_{i=1}^{n} w_i\left||f(x_i) - f(\bar{x}_H)| - \left|C_{\bar{x}_H}\left(\frac{1}{x_i} - \frac{1}{\bar{x}_H}\right)\right|\right|, \tag{17}$$

$$\mathcal{U}(f) = \sum_{i=1}^{n} w_i C_{x_i}\left(\frac{1}{x_i} - \frac{1}{\bar{x}_H}\right). \tag{18}$$

Corollary 1. *For a harmonically convex $f : I \to \mathbb{R}$, choose $C_{\bar{x}_H} \in \partial_H f(\bar{x}_H)$ and $C_{x_i} \in \partial_H f(x_i)$, $i = 1, \dots, n$. Then*

$$0 \leq \mathcal{L}(f) \leq \mathcal{R}(f) \leq \mathcal{J}(f) \leq \mathcal{U}(f). \tag{19}$$

Proof. Theorem 1 yields the first three inequalities, while the last one is a consequence of Theorem 2. □

If f is differentiable, then $C_{x_i} = -x_i^2 f'(x_i)$, $i = 1, \dots, n$ and

$$\mathcal{U}(f) = \sum_{i=1}^{n} w_i x_i^2 f'(x_i)\frac{1}{\bar{x}_H} - \sum_{i=1}^{n} w_i x_i f'(x_i) = \sum_{i=1}^{n} w_i x_i f'(x_i)\left(\frac{x_i}{\bar{x}_H} - 1\right).$$

Hence, for differentiable f, the estimate $\mathcal{J}(f) \leq \mathcal{U}(f)$ is exactly the discrete case of the upper estimate proved by Dragomir [6]. The lower bounds $\mathcal{L}(f)$ and $\mathcal{R}(f)$ in (19) complement this result.

Our next aim is to compute the difference $\mathcal{J}(f) - \mathcal{R}(f)$, i.e., to find exactly how much the harmonic Jensen functional exceeds its lower bound $\mathcal{R}(f)$ from (19). This shows when the lower bound $\mathcal{R}(f)$ is sharp. We use the positive and the negative part of a real number t:

$$t^+ = \max\{t, 0\} \qquad \text{and} \qquad t^- = \max\{-t, 0\}.$$

As in the proof of Theorem 1, let

$$A_i = f(x_i) - f(\bar{x}_H), \qquad B_i = C_{\bar{x}_H}\left(\frac{1}{x_i} - \frac{1}{\bar{x}_H}\right), \qquad i = 1, \dots, n.$$

Theorem 3. *In the setting of Corollary 1, the following identity is true:*

$$\mathcal{J}(f) - \mathcal{R}(f) = 2\sum_{i=1}^{n} w_i \min\{A_i^+, B_i^-\}. \tag{20}$$

Hence

$$\mathcal{R}(f) = \mathcal{J}(f) \quad \textit{if and only if} \quad A_i B_i \geq 0 \textit{ for every } i \in \{1, \ldots, n\} \textit{ with } w_i > 0.$$

In particular, if f is monotone on the smallest closed interval which contains all points x_i with $w_i > 0$, then

$$\mathcal{R}(f) = \mathcal{J}(f). \tag{21}$$

Proof. By (5), $A_i \geq B_i$. Let A and B be real numbers with $A \geq B$. A direct check shows that

$$\big|\,|A| - |B|\,\big| = A - B - 2\min\{A^+, B^-\}. \tag{22}$$

Namely, if $B \geq 0$ or $A \leq 0$, then A and B have the same sign and the minimum in (22) is zero. If $A > 0 > B$, then

$$\big|\,|A| - |B|\,\big| = |A + B| = A - B - 2\min\{A, -B\}.$$

We use (22) with $A = A_i$ and $B = B_i$, multiply by w_i and add the obtained equalities. This yields

$$\mathcal{R}(f) = \sum_{i=1}^{n} w_i (A_i - B_i) - 2\sum_{i=1}^{n} w_i \min\{A_i^+, B_i^-\}.$$

Since

$$\sum_{i=1}^{n} w_i B_i = C_{\bar{x}_H}\left(\sum_{i=1}^{n} \frac{w_i}{x_i} - \frac{1}{\bar{x}_H}\right) = 0 \quad \text{and} \quad \sum_{i=1}^{n} w_i A_i = \mathcal{J}(f),$$

we have

$$\sum_{i=1}^{n} w_i (A_i - B_i) = \mathcal{J}(f),$$

and (20) follows.

For $A \geq B$, we have $\min\{A^+, B^-\} = 0$ if and only if $AB \geq 0$. So $\mathcal{R}(f) = \mathcal{J}(f)$ if and only if $A_i B_i \geq 0$ for every $i \in \{1, \ldots, n\}$ with $w_i > 0$.

Let K be the smallest closed interval which contains all x_i with $w_i > 0$. If all these points are equal, then $\mathcal{R}(f) = \mathcal{J}(f) = 0$. In the opposite case, $\bar{x}_H$ lies in the interior of K. Let f be nondecreasing on K. Then $f'_{\pm}(\bar{x}_H) \geq 0$ and by (6), we get

$$C_{\bar{x}_H} \leq 0.$$

Therefore,

$$A_i \geq 0, \quad B_i \geq 0 \qquad \text{if } x_i \geq \bar{x}_H,$$

while

$$A_i \leq 0, \quad B_i \leq 0 \qquad \text{if } x_i \leq \bar{x}_H.$$

Consequently, $A_i B_i \geq 0$ for every $i \in \{1, \ldots, n\}$ with $w_i > 0$ and (21) holds. The case of nonincreasing f is similar. □

Remark 2. *Identity (20) is not specific to harmonic convexity. Its proof uses only the inequalities $A_i \geq B_i$ and the equality $\sum_{i=1}^{n} w_i B_i = 0$. Let g be convex on an interval J, let $u_1, \ldots, u_n \in J$ and let $\bar{u} = \sum_{i=1}^{n} w_i u_i$ be an interior point of J. If $c \in \partial g(\bar{u})$, $a_i = g(u_i) - g(\bar{u})$ and $b_i = c(u_i - \bar{u})$, then the same proof gives*

$$\sum_{i=1}^{n} w_i g(u_i) - g(\bar{u}) = \sum_{i=1}^{n} w_i \big|\,|a_i| - |b_i|\,\big| + 2\sum_{i=1}^{n} w_i \min\{a_i^+, b_i^-\}.$$

For $g(u) = f\left(\frac{1}{u}\right)$, $u_i = \frac{1}{x_i}$ *and* $c = C_{\bar{x}_H}$, *we have* $a_i = A_i$ *and* $b_i = B_i$, *so we get back* (20). *The integral form of this identity is given in Theorem* 7.

Remark 3. *If* f *is monotone, then* (21) *holds, but* $\mathcal{L}(f)$ *can be smaller than* $\mathcal{J}(f)$. *For example, if* $f(x) = x$, $x_1 = 1$, $x_2 = 4$ *and* $w_1 = w_2 = \frac{1}{2}$, *then*

$$\bar{x}_H = 1.6, \qquad \mathcal{L}(f) = 0.54, \qquad \mathcal{R}(f) = \mathcal{J}(f) = 0.9.$$

Monotonicity is not a necessary condition for (21), *since* $C_{\bar{x}_H} = 0$ *also gives equality. Theorem* 4 *below provides bounds of a different kind, expressed through the difference of the arithmetic and the harmonic mean.*

In contrast to the preceding results, harmonic convexity of f is not assumed in the next theorem. Only the points with positive weights play a role in it. Let

$$\mathcal{I}_+ = \{i \in \{1, \dots, n\} \mid w_i > 0\},$$

$$x_{\min} = \min_{i \in \mathcal{I}_+} x_i, \qquad x_{\max} = \max_{i \in \mathcal{I}_+} x_i,$$

and

$$K = [x_{\min}, x_{\max}].$$

So K is the smallest closed interval which contains all points x_i with $w_i > 0$. We assume that $x_{\min} < x_{\max}$, i.e., that at least two points x_i with positive weights are different. Otherwise, $x_i = \bar{x}_H$ for all $i \in \mathcal{I}_+$ and all terms in Theorem 4 below are equal to zero. Let

$$\bar{x} = \sum_{i=1}^{n} w_i x_i \qquad \text{and} \qquad V = \frac{1}{\bar{x}_H} \sum_{i=1}^{n} w_i (x_i - \bar{x}_H)^2.$$

Theorem 4. *Let* $f : I \to \mathbb{R}$ *and* $m \in \mathbb{R}$. *If* $\psi_m(t) = tf(t) - \frac{m}{2}t^2$ *is a convex function on* K, *then*

$$\frac{m}{2}(\bar{x} - \bar{x}_H) \le \mathcal{J}(f).$$

In the case when f *is also differentiable on* I, *we have*

$$\frac{m}{2}(\bar{x} - \bar{x}_H) \le \mathcal{J}(f) \le \mathcal{U}(f) - \frac{m}{2}V. \tag{23}$$

Proof. Let $f_m(x) = f(x) - \frac{m}{2}x$. Since $tf_m(t) = \psi_m(t)$ is convex on K, Lemma 1 implies that f_m is harmonically convex on K. By inequality (10) applied to f_m on K, we get

$$\mathcal{J}(f) = \mathcal{J}(f_m) + \frac{m}{2}(\bar{x} - \bar{x}_H) \ge \frac{m}{2}(\bar{x} - \bar{x}_H).$$

Differentiability is not used in this part of the proof.

Assume now that f is differentiable on I. Applying (8) with $I_0 = K$ to f_m with $x = x_i$, $w_i > 0$ and $y = \bar{x}_H$ and using $f_m'(x) = f'(x) - \frac{m}{2}$, we get

$$f_m(x_i) - f_m(\bar{x}_H) \le -x_i^2 f'(x_i)\left(\frac{1}{x_i} - \frac{1}{\bar{x}_H}\right) + \frac{m}{2}\left(x_i - \frac{x_i^2}{\bar{x}_H}\right).$$

Since

$$f(x_i) - f(\bar{x}_H) = f_m(x_i) - f_m(\bar{x}_H) + \frac{m}{2}(x_i - \bar{x}_H),$$

it follows that

$$f(x_i) - f(\bar{x}_H) \le -x_i^2 f'(x_i)\left(\frac{1}{x_i} - \frac{1}{\bar{x}_H}\right) + \frac{m}{2}\left(2x_i - \bar{x}_H - \frac{x_i^2}{\bar{x}_H}\right).$$

Multiplication by w_i and summation over $i = 1, \dots, n$ lead to

$$\begin{aligned}\mathcal{J}(f) &\le \mathcal{U}(f) + \frac{m}{2}\left(2\bar{x} - \bar{x}_H - \frac{1}{\bar{x}_H}\sum_{i=1}^{n} w_i x_i^2\right)\\ &= \mathcal{U}(f) - \frac{m}{2}\left(\frac{1}{\bar{x}_H}\sum_{i=1}^{n} w_i x_i^2 - 2\bar{x} + \bar{x}_H\right).\end{aligned}$$

Since

$$\frac{1}{\bar{x}_H}\sum_{i=1}^{n} w_i (x_i - \bar{x}_H)^2 = \frac{1}{\bar{x}_H}\sum_{i=1}^{n} w_i x_i^2 - 2\bar{x} + \bar{x}_H,$$

the upper bound in (23) is proved. □

Remark 4. *Take $\alpha, \beta \in \mathbb{R}$ and a function f, differentiable on I, such that $f(x) = \frac{\alpha}{x} + \beta + \frac{m}{2}x$ for $x \in K$. Then both inequalities in (23) become equalities, so the bounds in Theorem 4 are sharp. Indeed, in this case $\psi_m(t) = tf(t) - \frac{m}{2}t^2 = \alpha + \beta t$ is affine on K, so the assumption of Theorem 4 holds and*

$$f_m(x) = f(x) - \frac{m}{2}x = \frac{\alpha}{x} + \beta, \qquad x \in K.$$

Since equality in (10) holds for f_m, we have $\mathcal{J}(f_m) = 0$, i.e.,

$$\mathcal{J}(f) = \frac{m}{2}(\bar{x} - \bar{x}_H).$$

Moreover, $-x^2 f_m'(x) = \alpha$ for $x \in K$ and

$$f_m(x_i) - f_m(\bar{x}_H) = \alpha\left(\frac{1}{x_i} - \frac{1}{\bar{x}_H}\right), \qquad i \in \mathcal{I}_+,$$

so the inequality obtained from (8) in the proof of Theorem 4 is an equality for every $i \in \mathcal{I}_+$. All other steps in that proof are identities, hence $\mathcal{J}(f) = \mathcal{U}(f) - \frac{m}{2}V$.

Remark 5. *For $m = 0$, the upper bound in (23) is the inequality $\mathcal{J}(f) \le \mathcal{U}(f)$ from Corollary 1 for differentiable HA-convex functions.*

Proposition 1. *Let $f \in C^2(I)$, let $K = [x_{\min}, x_{\max}]$ be as above and put*

$$m^* = \min_{t \in K}\big(2f'(t) + tf''(t)\big).$$

Then $\psi_m(t) = tf(t) - \frac{m}{2}t^2$ is convex on K if and only if $m \le m^*$. If $m > 0$ and ψ_m is convex, then f is strongly harmonically convex on K with modulus $\frac{m}{2}x_{\min}^3$. Conversely, if f is strongly harmonically convex on K with modulus $c > 0$, then ψ_m is convex for $m = \frac{2c}{x_{\max}^3}$.

Proof. Since

$$\psi_m''(t) = 2f'(t) + tf''(t) - m,$$

the first assertion follows. By [7], strong harmonic convexity with modulus $c > 0$ means that $g(u) = f(1/u)$ is strongly convex with modulus c on K^{-1}. But

$$g''(u) = t^3\big(2f'(t) + tf''(t)\big), \qquad t = \frac{1}{u}.$$

If ψ_m is convex and $m > 0$, then $g''(u) \geq m x_{\min}^3$ and $g(u) - \frac{m}{2} x_{\min}^3 u^2$ is convex. Conversely, strong harmonic convexity with modulus c gives $g''(u) \geq 2c$. Therefore,

$$2f'(t) + tf''(t) = \frac{g''(1/t)}{t^3} \geq \frac{2c}{x_{\max}^3}, \qquad t \in K,$$

which proves the last assertion. □

Remark 6. *For twice differentiable f, the choice $m = m^*$ gives the best bounds in (23), since $\bar{x} - \bar{x}_H \geq 0$ and $V \geq 0$. If $m < 0$, the function f need not be harmonically convex on K, as illustrated below.*

We point out that V differs from the variance of the points x_i. Indeed,

$$V = \frac{\sum_{i=1}^n w_i (x_i - \bar{x})^2 + (\bar{x} - \bar{x}_H)^2}{\bar{x}_H}.$$

For a simple example with $m < 0$, take $f(t) = (t-1)^2$ on $K = [\frac{1}{2}, \frac{3}{2}]$. Here,

$$2f'(t) + tf''(t) = 6t - 4, \qquad m^* = -1.$$

Thus $tf(t) + \frac{1}{2}t^2$ is convex on K, although f is not harmonically convex on this interval, since $2f'(t) + tf''(t) < 0$ for $\frac{1}{2} \leq t < \frac{2}{3}$.

For twice differentiable harmonically convex functions, Theorem 4 gives the following improvement of the upper bound $\mathcal{J}(f) \leq \mathcal{U}(f)$.

Corollary 2. *Let $f \in C^2(I)$ be harmonically convex on K and let*

$$m^* = \min_{t \in K} \big(2f'(t) + tf''(t)\big).$$

Then $m^* \geq 0$ and

$$\mathcal{J}(f) \leq \mathcal{U}(f) - \frac{m^*}{2} V \leq \mathcal{U}(f). \tag{24}$$

Moreover, if $m^* > 0$, then

$$\mathcal{U}(f) - \frac{m^*}{2} V < \mathcal{U}(f).$$

Proof. By (4), $m^* \geq 0$. Since

$$\psi''_{m^*}(t) = 2f'(t) + tf''(t) - m^* \geq 0, \qquad t \in K,$$

Theorem 4 with $m = m^*$ gives the first inequality in (24). Since $x_{\min} < x_{\max}$, not all points x_i with $w_i > 0$ are equal to $\bar{x}_H$. Hence $V > 0$ and the remaining assertions follow. □

Thus the upper bound of Dragomir [6] is improved by the nonnegative term $\frac{m^*}{2} V$. By Remark 4, the improved bound is sharp.

We close this section with a probabilistic interpretation of the obtained results. Let $X > 0$ be a discrete random variable with $\mathbb{P}(X = x_i) = w_i$, $i = 1, \ldots, n$. We write

$$\mathbb{E}_H[X] = \left(\mathbb{E}\left[\frac{1}{X} \right] \right)^{-1}$$

for the harmonic expectation of X. Corollary 1 now gives

$$\begin{aligned} 0 &\le \left|\mathbb{E}\left[|f(X)-f(\mathbb{E}_H[X])|-\left|C_{\mathbb{E}_H[X]}\left(\frac{1}{X}-\frac{1}{\mathbb{E}_H[X]}\right)\right|\right]\right| \\ &\le \mathbb{E}[f(X)]-f(\mathbb{E}_H[X]) \le \mathbb{E}\left[C_X\left(\frac{1}{X}-\frac{1}{\mathbb{E}_H[X]}\right)\right]. \end{aligned} \tag{25}$$

Take $f(x)=x$ and denote $H=\mathbb{E}_H[X]$. Then $C_x=-x^2$ and since f is increasing, Theorem 3 gives $\mathcal{R}(f)=\mathcal{J}(f)$. Moreover,

$$\left||x_i-H|-H^2\left|\frac{1}{x_i}-\frac{1}{H}\right|\right| = |x_i-H|\left|1-\frac{H}{x_i}\right| = \frac{(x_i-H)^2}{x_i}.$$

Theorem 4 with $m=2$ gives equality in the upper bound. Therefore,

$$\begin{aligned} 0 &\le \left|\mathbb{E}\left[\frac{(X-H)|X-H|}{X}\right]\right| \le \mathbb{E}\left[\frac{(X-H)^2}{X}\right] \\ &= \mathbb{E}[X]-H = \frac{\mathbb{E}[X^2]}{H}-\mathbb{E}[X]-\frac{\mathbb{E}[(X-H)^2]}{H}. \end{aligned} \tag{26}$$

Thus the difference $\mathbb{E}[X]-H$ admits two different representations. In particular, $\mathbb{E}[(X-H)^2/X]=\mathbb{E}[X]-H$ follows directly by expanding the square and using $\mathbb{E}[1/X]=1/H$. The first inequality in (26) can be strict and the upper bound in (25) is larger than $\mathbb{E}[X]-H$ exactly by $\frac{\mathbb{E}[(X-H)^2]}{H}$.

3. Applications to Csiszár f-Divergences

Let $\mathcal{P}_n=\{\mathbf{p}=(p_1,\dots,p_n)\mid p_i>0,\ i=1,\dots,n,\ \sum_{i=1}^n p_i=1\}$ be the family of positive probability vectors with n coordinates. For $\mathbf{p},\mathbf{q}\in\mathcal{P}_n$ and a real function f on an interval that covers all quotients $\frac{q_i}{p_i}$, $i=1,\dots,n$, let

$$D_f(\mathbf{q},\mathbf{p})=\sum_{i=1}^n p_i f\left(\frac{q_i}{p_i}\right). \tag{27}$$

When f is convex, $D_f(\mathbf{q},\mathbf{p})$ is called the Csiszár f-divergence [31]. It measures the dissimilarity between the distributions $\mathbf{q}$ and $\mathbf{p}$ and in most applications the normalization $f(1)=0$ is assumed. Many divergences used in information theory and statistics are obtained from (27) for a suitable choice of f. For convex f, Jensen's inequality (1) and $\sum_{i=1}^n p_i\frac{q_i}{p_i}=\sum_{i=1}^n q_i=1$ give

$$D_f(\mathbf{q},\mathbf{p})\ge f(1), \tag{28}$$

so $D_f(\mathbf{q},\mathbf{p})\ge 0$ if $f(1)=0$. Equality in (28) is attained only for $\mathbf{p}=\mathbf{q}$ whenever, for instance, f is strictly convex at 1.

Sason and Verdú [32] compared numerous f-divergences with each other and with the χ^2-divergence, while in [30] (see also [27]) the generating function f was assumed to be strongly convex. The paper [26] explains how different convexity types of f, HA-convexity among them, influence D_f. In this section, we show that harmonic convexity of f leads to bounds for $D_f(\mathbf{q},\mathbf{p})$ in which the weighted harmonic mean of the ratios $\frac{q_i}{p_i}$ appears. Theorems 5 and 6 below follow directly from Corollary 1 and Theorem 4. What is specific for divergences is that this harmonic mean is equal to $(1+\chi^2(\mathbf{p},\mathbf{q}))^{-1}$ (see (30)). Hence the χ^2-divergence appears in all the obtained bounds.

Let

$$r_i = \frac{q_i}{p_i}, \qquad \tau = \min_{1\le i\le n} r_i, \qquad T = \max_{1\le i\le n} r_i,$$
$$\sigma = \left(\sum_{i=1}^{n} \frac{p_i}{r_i}\right)^{-1} = \left(\sum_{i=1}^{n} \frac{p_i^2}{q_i}\right)^{-1}. \tag{29}$$

So σ is the $\mathbf{p}$-weighted harmonic mean of the ratios r_i. Let

$$\chi^2(\mathbf{p},\mathbf{q}) = \sum_{i=1}^{n} \frac{(p_i - q_i)^2}{q_i} = \sum_{i=1}^{n} \frac{p_i^2}{q_i} - 1$$

denote the χ^2-divergence. It is the special case of D_f from (27) with $f(t) = (t-1)^2$. Consequently,

$$\sigma = \frac{1}{1+\chi^2(\mathbf{p},\mathbf{q})} \le 1. \tag{30}$$

We also use

$$\chi^2(\mathbf{q},\mathbf{p}) = \sum_{i=1}^{n} \frac{(q_i - p_i)^2}{p_i} = \sum_{i=1}^{n} \frac{q_i^2}{p_i} - 1$$

and put

$$\kappa = (1+\chi^2(\mathbf{p},\mathbf{q}))(1+\chi^2(\mathbf{q},\mathbf{p})) + \sigma - 2.$$

From $\sum_{i=1}^{n} p_i r_i = 1$ and $\sum_{i=1}^{n} p_i r_i^2 = 1 + \chi^2(\mathbf{q},\mathbf{p})$, it follows that

$$\kappa = \frac{1+\chi^2(\mathbf{q},\mathbf{p})}{\sigma} - 2 + \sigma = \frac{1}{\sigma}\sum_{i=1}^{n} p_i (r_i - \sigma)^2 \ge 0. \tag{31}$$

In particular, κ vanishes exactly when $\mathbf{p} = \mathbf{q}$.

We keep the convention $D_f(\mathbf{q},\mathbf{p})$ in (27), where $\mathbf{p}$ supplies the weights. To avoid confusing the order of the arguments, the lower and upper estimates below are denoted by $\mathcal{L}_f(\mathbf{q},\mathbf{p})$, $\mathcal{R}_f(\mathbf{q},\mathbf{p})$ and $\mathcal{U}_f(\mathbf{q},\mathbf{p})$.

For $C_x \in \partial_H f(x)$, let

$$\mathcal{L}_f(\mathbf{q},\mathbf{p}) = \left|\sum_{i=1}^{n} p_i \left(|f(r_i) - f(\sigma)| - \left| C_\sigma \left(\frac{1}{r_i} - \frac{1}{\sigma} \right) \right| \right) \right|, \tag{32}$$

$$\mathcal{R}_f(\mathbf{q},\mathbf{p}) = \sum_{i=1}^{n} p_i \left| |f(r_i) - f(\sigma)| - \left| C_\sigma \left(\frac{1}{r_i} - \frac{1}{\sigma} \right) \right| \right|, \tag{33}$$

$$\mathcal{U}_f(\mathbf{q},\mathbf{p}) = \sum_{i=1}^{n} p_i C_{r_i} \left(\frac{1}{r_i} - \frac{1}{\sigma} \right). \tag{34}$$

Theorem 5. *Let $\mathbf{p},\mathbf{q} \in \mathcal{P}_n$ and let the ratios $r_i = \frac{q_i}{p_i}$, $i = 1,\dots,n$, belong to an open interval $I \subset (0,\infty)$ on which f is harmonically convex. If $C_x \in \partial_H f(x)$, $x \in I$, then*

$$0 \le \mathcal{L}_f(\mathbf{q},\mathbf{p}) \le \mathcal{R}_f(\mathbf{q},\mathbf{p}) \le D_f(\mathbf{q},\mathbf{p}) - f(\sigma) \le \mathcal{U}_f(\mathbf{q},\mathbf{p}). \tag{35}$$

When, in addition, f is convex on I, we have

$$\max\{f(1),\, f(\sigma) + \mathcal{R}_f(\mathbf{q},\mathbf{p})\} \le D_f(\mathbf{q},\mathbf{p}) \le f(\sigma) + \mathcal{U}_f(\mathbf{q},\mathbf{p}). \tag{36}$$

If f is differentiable, then

$$\mathcal{U}_f(\mathbf{q},\mathbf{p}) = \sum_{i=1}^{n} p_i r_i f'(r_i) \left(\frac{r_i}{\sigma} - 1 \right). \tag{37}$$

Proof. We apply Corollary 1 with $w_i = p_i$ and $x_i = r_i$, $i = 1, \dots, n$. Then

$$\bar{x}_H = \sigma \qquad \text{and} \qquad \sum_{i=1}^{n} p_i f(r_i) = D_f(\mathbf{q}, \mathbf{p}),$$

so (35) holds. If f is also convex, then (28) holds and together with (35) it gives (36). Finally, putting $C_{r_i} = -r_i^2 f'(r_i)$ in (34), we get (37). □

Theorem 5 requires all ratios r_i to lie in an interval of harmonic convexity of f, e.g., $r_i > e^{-3/2}$ for $f(t) = t \ln t$ and $r_i > 9/16$ for the squared Hellinger divergence. It cannot be applied to a subset of indices, since the harmonic mean of the corresponding ratios differs from σ. In that case, Theorem 6 below can be used, since it holds for all positive ratios with a suitable m.

Remark 7. *If f is both convex and harmonically convex, the lower bound in (36) improves the classical bound (28) whenever $f(\sigma) + \mathcal{R}_f(\mathbf{q}, \mathbf{p}) > f(1)$. For example, if f is monotone on $[\tau, T]$, then by Corollary 3 below, $f(\sigma) + \mathcal{R}_f(\mathbf{q}, \mathbf{p}) = D_f(\mathbf{q}, \mathbf{p})$, i.e., the lower bound is equal to the divergence.*

For monotone f, Theorem 3 gives the following identity for $D_f(\mathbf{q}, \mathbf{p})$.

Corollary 3. *Let $\mathbf{p}, \mathbf{q} \in \mathcal{P}_n$ and let f be harmonically convex on I, where $I \supset [\tau, T]$ is an open subinterval of $(0, \infty)$. If, moreover, f is monotone on $[\tau, T]$, then for every $C_\sigma \in \partial_H f(\sigma)$,*

$$D_f(\mathbf{q}, \mathbf{p}) = f(\sigma) + \sum_{i=1}^{n} p_i \left| |f(r_i) - f(\sigma)| - |C_\sigma| \left| \frac{1}{r_i} - \frac{1}{\sigma} \right| \right|, \tag{38}$$

where $r_i = \frac{q_i}{p_i}$ and $\sigma = (1 + \chi^2(\mathbf{p}, \mathbf{q}))^{-1}$. If f is differentiable at σ, then $C_\sigma = -\sigma^2 f'(\sigma)$.

Proof. We apply Theorem 3 with $w_i = p_i$ and $x_i = r_i$, $i = 1, \dots, n$. In this case,

$$\mathcal{J}(f) = D_f(\mathbf{q}, \mathbf{p}) - f(\sigma) \qquad \text{and} \qquad \mathcal{R}(f) = \mathcal{R}_f(\mathbf{q}, \mathbf{p}),$$

so (21) gives (38). □

Theorem 4 leads to estimates of $D_f(\mathbf{q}, \mathbf{p})$ whenever $\psi_m(t) = tf(t) - \frac{m}{2}t^2$ is convex for a suitable $m \in \mathbb{R}$.

Theorem 6. *Let $\mathbf{p}, \mathbf{q} \in \mathcal{P}_n$, $f : I \to \mathbb{R}$ and $m \in \mathbb{R}$, where $I \subset (0, \infty)$ is an open interval with $[\tau, T] \subset I$. If $\psi_m(t) = tf(t) - \frac{m}{2}t^2$ is convex on $[\tau, T]$, then*

$$f(\sigma) + \frac{m}{2}(1 - \sigma) \le D_f(\mathbf{q}, \mathbf{p}).$$

In the case when f is also differentiable on I, we have

$$f(\sigma) + \frac{m}{2}(1 - \sigma) \le D_f(\mathbf{q}, \mathbf{p}) \le f(\sigma) + \mathcal{U}_f(\mathbf{q}, \mathbf{p}) - \frac{m}{2}\kappa, \tag{39}$$

where $\mathcal{U}_f(\mathbf{q}, \mathbf{p})$ is given by (37). If $\mathbf{p} = \mathbf{q}$, both inequalities in (39) become equalities.

Proof. We apply Theorem 4 with $w_i = p_i$ and $x_i = r_i$. Then $\bar{x} = 1$, $\bar{x}_H = \sigma$ and by (31),

$$V = \frac{1}{\sigma} \sum_{i=1}^{n} p_i (r_i - \sigma)^2 = \kappa.$$

Differentiability is needed only for the upper bound. If $\mathbf{p} = \mathbf{q}$, then $r_i = \sigma = 1$, $\kappa = 0$ and $\mathcal{U}_f(\mathbf{q}, \mathbf{p}) = 0$, so both inequalities become equalities. □

For comparison, assume that f is harmonically convex on $[\tau, T]$ and $\tau < T$. Put

$$\theta = \frac{1/\tau - 1/\sigma}{1/\tau - 1/T}.$$

Applying the classical Lah–Ribarič (chord) inequality [33] to $g(u) = f(1/u)$, $1/T \le u \le 1/\tau$, gives

$$D_f(\mathbf{q}, \mathbf{p}) \le \theta f(T) + (1-\theta) f(\tau). \tag{40}$$

Indeed, $1/\sigma = \theta/T + (1-\theta)/\tau$, so (40) follows by applying the chord bound for g at each $1/r_i$ and summing with the weights p_i. If every ratio is equal to τ or T, (40) is an equality.

Remark 8. *In Theorem 6, the function f need not be convex. For $m \ge 0$, f is harmonically convex on $[\tau, T]$, while for $m < 0$ it need not be (see Proposition 1 and the example following it). Let f be of class C^2 on some open interval including $[\tau, T]$ and let $\mathbf{p} \ne \mathbf{q}$, so that $\tau < T$. Then the largest admissible value of m is*

$$m^* = \min_{t \in [\tau, T]} \big(2f'(t) + t f''(t)\big).$$

As m grows, the lower bound in (39) becomes larger and the upper bound becomes smaller, so $m = m^*$ is the optimal choice. For $\mathbf{p} = \mathbf{q}$, equality holds for every m.

Since $1 - \sigma = \frac{\chi^2(\mathbf{p},\mathbf{q})}{1+\chi^2(\mathbf{p},\mathbf{q})}$, the lower bound in (39) can be written as

$$D_f(\mathbf{q}, \mathbf{p}) \ge f\left(\frac{1}{1+\chi^2(\mathbf{p},\mathbf{q})}\right) + \frac{m}{2}\,\frac{\chi^2(\mathbf{p},\mathbf{q})}{1+\chi^2(\mathbf{p},\mathbf{q})}.$$

So, for a fixed m, it depends on $\mathbf{p}$ and $\mathbf{q}$ only through $\chi^2(\mathbf{p}, \mathbf{q})$.

The lower bound in (39) is, in fact, inequality (10) applied to the function $g_m(t) = f(t) - \frac{m}{2}(t-1)$. Namely,

$$t g_m(t) = \psi_m(t) + \frac{m}{2} t$$

is convex by the assumption of Theorem 6 and $D_{g_m}(\mathbf{q}, \mathbf{p}) = D_f(\mathbf{q}, \mathbf{p})$, since adding a multiple of $t - 1$ does not change the divergence. If $m \ge 0$ and f is nonincreasing on $[\sigma, 1]$, this lower bound is not smaller than $f(1)$. If f is convex, we can also use $D_f(\mathbf{q}, \mathbf{p}) \ge f(1)$ and take the larger of the two lower bounds.

The term $\frac{m}{2}\kappa$ depends only on chi-square divergences, but $\mathcal{U}_f$ still depends on all ratios r_i.

In the following examples, we illustrate the obtained results by considering three divergences which are frequently used in information theory.

Example 1. *For $f(t) = t \ln t$, the functional (27) becomes the Kullback–Leibler divergence*

$$\mathrm{KL}(\mathbf{q}, \mathbf{p}) = \sum_{i=1}^{n} q_i \ln \frac{q_i}{p_i}.$$

In this case, $2f'(t) + tf''(t) = 2\ln t + 3$, *so by* (4) *the function* f *is harmonically convex on* $I_{\mathrm{KL}} = (e^{-\frac{3}{2}}, \infty)$. *Let* $r_i > e^{-\frac{3}{2}}$ *for all* i *and let*

$$L_{\mathrm{KL}} = \left| \sum_{i=1}^{n} p_i \left(|r_i \ln r_i - \sigma \ln \sigma| - \sigma^2 |\ln \sigma + 1| \left| \frac{1}{r_i} - \frac{1}{\sigma} \right| \right) \right|,$$
$$U_{\mathrm{KL}} = \sum_{i=1}^{n} p_i r_i (\ln r_i + 1) \left(\frac{r_i}{\sigma} - 1 \right).$$

By Theorem 5,

$$\max\{0, \sigma \ln \sigma + L_{\mathrm{KL}}\} \le \mathrm{KL}(\mathbf{q}, \mathbf{p}) \le \sigma \ln \sigma + U_{\mathrm{KL}}. \tag{41}$$

If $r_i \ge \frac{1}{e}$ *for all* i, *then* $t \ln t$ *is increasing on* $[\tau, T]$ *and Corollary* 3 *gives*

$$\mathrm{KL}(\mathbf{q}, \mathbf{p}) = \sigma \ln \sigma + \sum_{i=1}^{n} p_i \left| |r_i \ln r_i - \sigma \ln \sigma| - \sigma^2 |1 + \ln \sigma| \left| \frac{1}{r_i} - \frac{1}{\sigma} \right| \right|.$$

For $\mathbf{p} = (0.5, 0.5)$ *and* $\mathbf{q} = (0.6, 0.4)$, *which is case (A) in Table* 1, *we have* $\sigma = 0.96$ *and*

$$\mathrm{KL}(\mathbf{q}, \mathbf{p}) \approx -0.0391891 + 0.0593246 = 0.0201355.$$

Since $\sigma \le 1$, *we have* $\sigma \ln \sigma \le 0$. *So the lower bound in* (41) *is better than the classical bound* 0 *if* $\sigma \ln \sigma + L_{\mathrm{KL}} > 0$.

The function $2\ln t + 3$ *is increasing, so* $m^* = 3 + 2\ln \tau$. *For this* m, *Theorem* 6 *holds for all positive ratios* r_i *and gives*

$$\begin{aligned} \sigma \ln \sigma + \left(\frac{3}{2} + \ln \tau \right)(1 - \sigma) &\le \mathrm{KL}(\mathbf{q}, \mathbf{p}) \\ &\le \sigma \ln \sigma + U_{\mathrm{KL}} - \left(\frac{3}{2} + \ln \tau \right) \kappa. \end{aligned} \tag{42}$$

Example 2. *Taking* $f(t) = \frac{1}{2}(\sqrt{t} - 1)^2$ *in* (27) *yields the squared Hellinger divergence*

$$\mathrm{Hel}^2(\mathbf{q}, \mathbf{p}) = \frac{1}{2} \sum_{i=1}^{n} (\sqrt{q_i} - \sqrt{p_i})^2.$$

For this function,

$$2f'(t) + tf''(t) = 1 - \frac{3}{4\sqrt{t}},$$

so f *is harmonically convex on* $I_H = (\frac{9}{16}, \infty)$. *Let* $r_i > \frac{9}{16}$ *for all* i *and let*

$$L_H = \left| \sum_{i=1}^{n} p_i \left(\frac{1}{2} \left| (\sqrt{r_i} - 1)^2 - (\sqrt{\sigma} - 1)^2 \right| - \frac{1}{2} \sigma^{\frac{3}{2}} (1 - \sqrt{\sigma}) \left| \frac{1}{r_i} - \frac{1}{\sigma} \right| \right) \right|,$$
$$U_H = \frac{1}{2} \sum_{i=1}^{n} p_i (r_i - \sqrt{r_i}) \left(\frac{r_i}{\sigma} - 1 \right).$$

Since $f(\sigma) \ge 0$ *and* $L_H \ge 0$, *the lower bound* $f(\sigma) + L_H$ *from* (35) *is not smaller than* $f(1) = 0$. *Therefore Theorem* 5 *gives*

$$\frac{1}{2}(\sqrt{\sigma} - 1)^2 + L_H \le \mathrm{Hel}^2(\mathbf{q}, \mathbf{p}) \le \frac{1}{2}(\sqrt{\sigma} - 1)^2 + U_H. \tag{43}$$

By (30), *the first term in* (43) *depends only on* $\chi^2(\mathbf{p}, \mathbf{q})$.

For arbitrary $\mathbf{p}, \mathbf{q} \in \mathcal{P}_n$, *Theorem 6 with* $m = 1 - \frac{3}{4\sqrt{\tau}}$ *gives*

$$\begin{aligned} \frac{1}{2}(\sqrt{\sigma}-1)^2 + \frac{1}{2}\left(1 - \frac{3}{4\sqrt{\tau}}\right)(1-\sigma) &\le \mathrm{Hel}^2(\mathbf{q},\mathbf{p}) \\ &\le \frac{1}{2}(\sqrt{\sigma}-1)^2 + U_H - \frac{1}{2}\left(1 - \frac{3}{4\sqrt{\tau}}\right)\kappa. \end{aligned} \tag{44}$$

Example 3. *With* $f(t) = (t-1)\ln t$, *the functional (27) reduces to the Jeffreys divergence*

$$D_J(\mathbf{q},\mathbf{p}) = \sum_{i=1}^{n}(q_i - p_i)\ln\frac{q_i}{p_i}.$$

For this function, condition (4) reads

$$2\ln t + 3 - \frac{1}{t} \ge 0.$$

Let $\rho \approx 0.5520012554$ *be the unique positive solution of*

$$2\ln\rho + 3 - \frac{1}{\rho} = 0. \tag{45}$$

Then f *is harmonically convex on* $I_J = (\rho, \infty)$. *Let* $r_i > \rho$ *for all* i *and let*

$$\begin{aligned} C_\sigma &= -\sigma^2\left(\ln\sigma + 1 - \frac{1}{\sigma}\right), \\ L_J &= \left|\sum_{i=1}^{n} p_i\left(|(r_i-1)\ln r_i - (\sigma-1)\ln\sigma| - |C_\sigma|\left|\frac{1}{r_i} - \frac{1}{\sigma}\right|\right)\right|, \\ U_J &= \sum_{i=1}^{n} p_i(r_i\ln r_i + r_i - 1)\left(\frac{r_i}{\sigma} - 1\right). \end{aligned}$$

Since $(\sigma-1)\ln\sigma \ge 0$ *and* $L_J \ge 0$, *the lower bound* $(\sigma-1)\ln\sigma + L_J$ *from (35) is not smaller than* $f(1) = 0$. *Thus*

$$(\sigma-1)\ln\sigma + L_J \le D_J(\mathbf{q},\mathbf{p}) \le (\sigma-1)\ln\sigma + U_J. \tag{46}$$

The function $2\ln t + 3 - \frac{1}{t}$ *is increasing, so Theorem 6 with* $m = 3 + 2\ln\tau - \frac{1}{\tau}$ *gives the following bounds for all positive ratios* r_i*:*

$$\begin{aligned} (\sigma-1)\ln\sigma + \frac{1}{2}\left(3 + 2\ln\tau - \frac{1}{\tau}\right)(1-\sigma) &\le D_J(\mathbf{q},\mathbf{p}) \\ &\le (\sigma-1)\ln\sigma + U_J - \frac{1}{2}\left(3 + 2\ln\tau - \frac{1}{\tau}\right)\kappa. \end{aligned} \tag{47}$$

All bounds in Examples 1–3 become equalities for $\mathbf{p} = \mathbf{q}$.

Now we compare these bounds numerically for two pairs of distributions:

(A) $\mathbf{p} = (0.5, 0.5)$ and $\mathbf{q} = (0.6, 0.4)$,
(B) $\mathbf{p} = (0.1, 0.4, 0.5)$ and $\mathbf{q} = (0.3, 0.3, 0.4)$.

The bounds from Theorem 6 are calculated with $m = m^*$. In part (a) of Table 1, the KL value in the column $\mathcal{L}_f$ is the lower bound $\max\{0, \sigma\ln\sigma + L_{\mathrm{KL}}\}$ from (41). In case (A), $L_{\mathrm{KL}} \approx 0.014488$ and $\sigma\ln\sigma + L_{\mathrm{KL}} \approx -0.024701$, so this lower bound is 0.

Table 1. Numerical comparison of the lower and upper bounds.

(a) Lower bounds					
Case	Divergence	$D_f(\mathbf{q},\mathbf{p})$	$\mathcal{L}_f$-bound	$\mathcal{R}_f$-bound	Theorem 6
(A)	KL	0.02014	0.00000	0.02014	0.01189
	Hel^2	0.00506	0.00308	0.00308	0.00343
	D_J	0.04055	0.02471	0.02471	0.02771
(B)	KL	0.15402	0.14059	0.15402	0.04784
	Hel^2	0.03317	0.02763	0.02763	0.01430
	D_J	0.27081	0.22644	0.22644	0.11597
(b) Upper bounds					
Case	Divergence	$D_f(\mathbf{q},\mathbf{p})$	$\mathcal{U}_f$-bound	Theorem 6	
(A)	KL	0.02014	0.08637	0.03104	
	Hel^2	0.00506	0.01067	0.00717	
	D_J	0.04055	0.08552	0.05728	
(B)	KL	0.15402	1.43682	0.75657	
	Hel^2	0.03317	0.17034	0.13275	
	D_J	0.27081	1.42051	1.11433	

Remark 9. *Part (a) of Table 1 contains three lower bounds. The $\mathcal{R}_f$-bound is the better of the two lower bounds from Theorem 5. In cases (A) and (B), the function $t\ln t$ is increasing on $[\tau, T]$, so by Theorem 3 the $\mathcal{R}_f$-bound is equal to $\mathrm{KL}(\mathbf{q},\mathbf{p})$. For the squared Hellinger and Jeffreys divergences in cases (A) and (B), all terms in $\mathcal{L}_f$ and $\mathcal{R}_f$ have the same sign, so these two bounds coincide. Neither of their generating functions is monotone throughout $[\tau, T]$, so this does not imply $\mathcal{R}_f(\mathbf{q},\mathbf{p}) = D_f(\mathbf{q},\mathbf{p}) - f(\sigma)$.*

In case (A), the lower bounds from Theorem 6 for the squared Hellinger and Jeffreys divergences are better than the $\mathcal{R}_f$-bounds, since $0.00343 > 0.00308$ and $0.02771 > 0.02471$. Note that for all three functions the lower estimate from Theorem 6 is determined by σ and τ alone. In case (B), the $\mathcal{R}_f$-bounds for these two divergences are better than the lower bounds from Theorem 6. Consequently, the $\mathcal{R}_f$-bound from Theorem 5 and the lower bound from Theorem 6 complement each other.

The upper estimates are given in part (b). For all three divergences, in both cases (A) and (B), the ratios r_i lie in the interval of harmonic convexity of f and $m^* > 0$. So, by Corollary 2, the upper bound from Theorem 6 is smaller than the $\mathcal{U}_f$-bound. The improvement is considerable. For example, in case (A) the $\mathcal{U}_f$-bound for the Kullback–Leibler divergence is 0.08637, while Theorem 6 gives 0.03104. In case (B), the corresponding values are 1.43682 and 0.75657.

For $m < 0$, Theorem 6 applies also to some functions which are not harmonically convex on $[\tau, T]$ (see the example below).

The following rule of thumb summarizes the comparison, including the examples and Remark 10 below. For the upper bound, Theorem 6 with $m = m^*$ should always be preferred to the $\mathcal{U}_f$-bound when f is harmonically convex and of class C^2 (Corollary 2). It is most accurate when the ratios r_i are concentrated around 1 and τ is not small, whereas the Lah–Ribarič bound (40) can be better when most of the mass of $\mathbf{p}$ lies at ratios close to τ or T. For the lower bound, the $\mathcal{R}_f$-bound from Theorem 5 is preferable when f is monotone on $[\tau, T]$, since it is then exact. It is also preferable when the ratios are widely spread, as in case (B). The lower bound from Theorem 6 requires only σ and τ. It is the only one of the two available when some ratios lie outside the interval of harmonic convexity. It also becomes accurate as $\mathbf{q} \to \mathbf{p}$. For close distributions, either of the two lower bounds can be better (compare case (A) with case (C) below).

For comparison, the inequality

$$\mathrm{KL}(\mathbf{q},\mathbf{p}) \le \ln\bigl(1+\chi^2(\mathbf{q},\mathbf{p})\bigr)$$

from [32] gives the upper bounds 0.03922 and 0.36811 in cases (A) and (B), respectively. In case (A), Theorem 6 gives the better bound 0.03104, while in case (B), the bound from [32] is better than the value 0.75657 from Theorem 6.

The classical bound (40) gives the following comparison with Theorem 6:

Case	Divergence	Exact value	Theorem 6	Lah–Ribarič
(A)	KL	0.02014	0.03104	0.02014
	Hel^2	0.00506	0.00717	0.00506
	D_J	0.04055	0.05728	0.04055
(B)	KL	0.15402	0.75657	0.28171
	Hel^2	0.03317	0.13275	0.04566
	D_J	0.27081	1.11433	0.37301
(C)	KL	0.01182	0.02112	0.12902
	Hel^2	0.00289	0.00464	0.01327
	D_J	0.02315	0.03715	0.10643

In case (A), both ratios are endpoints of the interval, so the Lah–Ribarič bound is exact. In case (B), it is again smaller than the upper bound from Theorem 6. The large ratio $r_1 = 3$ in case (B) increases both $\mathcal{U}_f$ and κ. The correction $\frac{m^*}{2}\kappa$ does not compensate for the growth of $\mathcal{U}_f$. For case (C) we take $\mathbf{p} = (0.45, 0.05, 0.4, 0.1)$ and $\mathbf{q} = (0.405, 0.075, 0.44, 0.08)$, so that $\mathbf{r} = (0.9, 1.5, 1.1, 0.8)$. Here, the upper bound from Theorem 6 is better than (40) for all three divergences, since 85% of the mass of $\mathbf{p}$ lies at the interior ratios 0.9 and 1.1. Hence the two upper bounds complement each other.

We also consider $\mathbf{p} = (0.5, 0.5)$ and $\mathbf{q} = (0.275, 0.725)$. Here,

$$\tau = 0.55, \qquad T = 1.45, \qquad \sigma = 0.7975, \qquad \kappa \approx 0.305337.$$

For the squared Hellinger divergence, $m^* = 1 - 3/(4\sqrt{0.55}) \approx -0.011300$, so the generating function is not harmonically convex on $[\tau, T]$. Nevertheless, Theorem 6 yields

$$0.004577 \le \mathrm{Hel}^2(\mathbf{q},\mathbf{p}) \approx 0.027110 \le 0.072599.$$

For the Jeffreys divergence, $m^* \approx -0.013856$ and the lower bound is approximately 0.044417, while the exact value is 0.218115. Theorem 5 cannot be applied in this case, while Theorem 6 still gives positive lower bounds.

Concerning the practical range of m, the admissible values form the interval $(-\infty, m^*]$. Every $m < m^*$ gives weaker bounds than $m = m^*$. For the three divergences considered here, the function $2f'(t) + tf''(t)$ is increasing, so m^* is determined by τ alone and $m^* < 0$ exactly when τ is smaller than $e^{-3/2}$, $\frac{9}{16}$ and ρ, respectively. If m^* is only slightly negative, as above, the correction terms $\frac{m^*}{2}(1-\sigma)$ and $\frac{m^*}{2}\kappa$ are small, so the bounds differ only slightly from $f(\sigma)$ and $f(\sigma) + \mathcal{U}_f(\mathbf{q},\mathbf{p})$. In the example above, the lower and the upper bound for Hel^2 are about 0.17 and 2.68 times the exact value. For very small τ, m^* is large in absolute value and the bounds lose accuracy. Thus negative values of m are useful mainly in order to cover ratios slightly below the interval of harmonic convexity.

To see how the bounds behave when $\mathbf{q}$ approaches $\mathbf{p}$, we take $\mathbf{p}$ and $\mathbf{q}$ from case (B) and put $\mathbf{q}_\varepsilon = (1-\varepsilon)\mathbf{p} + \varepsilon\mathbf{q}$, $0 < \varepsilon \le 1$. Then $r_i(\varepsilon) = 1 + \varepsilon(r_i - 1)$ and all ratios lie in the

intervals of harmonic convexity. Table 2 gives the ratio of each bound to the exact value $D_f(\mathbf{q}_\varepsilon, \mathbf{p})$.

Table 2. Ratios of the bounds to $D_f(\mathbf{q}_\varepsilon, \mathbf{p})$ for $\mathbf{q}_\varepsilon = (1-\varepsilon)\mathbf{p} + \varepsilon\mathbf{q}$, with $\mathbf{p}, \mathbf{q}$ from case (B). Theorem 6 is used with $m = m^*$.

ε	Divergence	$\mathcal{R}_f$-Bound	Theorem 6, Lower	$\mathcal{U}_f$-Bound	Theorem 6, Upper	Lah–Ribarič
1	KL	1.000	0.311	9.329	4.912	1.829
	Hel^2	0.833	0.431	5.135	4.002	1.377
	D_J	0.836	0.428	5.246	4.115	1.377
0.5	KL	1.000	0.526	6.408	2.663	1.587
	Hel^2	0.847	0.625	3.454	2.254	1.243
	D_J	0.849	0.621	3.485	2.292	1.243
0.1	KL	1.000	0.856	4.434	1.295	1.413
	Hel^2	0.937	0.890	2.273	1.221	1.147
	D_J	0.937	0.890	2.274	1.223	1.147
0.01	KL	1.000	0.983	4.042	1.029	1.375
	Hel^2	0.992	0.988	2.027	1.022	1.126
	D_J	0.992	0.988	2.027	1.022	1.126

Both bounds from Theorem 6 approach the exact value as $\varepsilon \to 0$, while the $\mathcal{U}_f$-bound and the Lah–Ribarič bound do not. This is explained by the following remark.

Remark 10. *Let f be of class C^2 on an open interval containing 1, with $f(1) = 0$ and $f''(1) > 0$. Let $\mathbf{p}$ be fixed and let $\mathbf{q} \to \mathbf{p}$, $\mathbf{q} \neq \mathbf{p}$, so that $\delta = \max_i |r_i - 1| \to 0$. Put $d_i = r_i - 1$ and $s = \chi^2(\mathbf{q}, \mathbf{p}) = \sum_{i=1}^n p_i d_i^2$. Note that $\delta^2 \le s / \min_i p_i$, so $o(\delta)\sqrt{s} = o(s)$. Since $\sum_{i=1}^n p_i d_i = 0$, we have*

$$\chi^2(\mathbf{p}, \mathbf{q}) = \sum_{i=1}^n p_i \frac{d_i^2}{1+d_i} = s + o(s), \qquad 1 - \sigma = s + o(s), \qquad \kappa = \frac{s + (1-\sigma)^2}{\sigma} = s + o(s).$$

Taylor's formula gives $D_f(\mathbf{q}, \mathbf{p}) = \frac{f''(1)}{2} s + o(s)$ and $f(\sigma) = -f'(1)s + o(s)$. Moreover, $r_i f'(r_i) = f'(1) + \big(f'(1) + f''(1)\big) d_i + o(\delta)$ and $\frac{r_i}{\sigma} - 1 = d_i + s + o(s) + O(\delta s)$, so (37) yields

$$\mathcal{U}_f(\mathbf{q}, \mathbf{p}) = \big(2f'(1) + f''(1)\big) s + o(s).$$

By continuity, $m^* \to 2f'(1) + f''(1)$. Consequently, with $m = m^*$ both bounds in (39) are equal to $\frac{f''(1)}{2} s + o(s)$, i.e., their ratio to $D_f(\mathbf{q}, \mathbf{p})$ tends to 1. On the other hand, $f(\sigma) + \mathcal{U}_f(\mathbf{q}, \mathbf{p}) = \big(f'(1) + f''(1)\big) s + o(s)$, so the ratio of the $\mathcal{U}_f$-bound to $D_f(\mathbf{q}, \mathbf{p})$ tends to $2\big(f'(1) + f''(1)\big) / f''(1)$. This limit is 4 for KL *and 2 for* Hel^2 *and D_J, in accordance with Table 2. Hence the correction $\frac{m^*}{2}\kappa$ makes Dragomir's bound asymptotically sharp.*

4. Refinement of the Harmonic Hermite–Hadamard Inequality

This section is a continuous analogue of Section 2, in which the role of $\bar{x}_H$ is played by $H_{a,b}$ defined below. The upper bounds in (53) are contained in Remark 9 in [4] (see Remark 11), while the lower bounds and identity (54) are new. We fix $0 < a < b$ and a function f which is harmonically convex on some open subinterval of $(0, \infty)$ containing $[a, b]$. We put

$$H_{a,b} = \frac{2ab}{a+b}, \qquad M_f(a,b) = \frac{ab}{b-a} \int_a^b \frac{f(x)}{x^2}\, \mathrm{d}x. \tag{48}$$

With this notation, İşcan's inequality (12) takes the form

$$f(H_{a,b}) \le M_f(a,b) \le \frac{f(a)+f(b)}{2}. \tag{49}$$

For $C \in \partial_H f(H_{a,b})$ and $x \in [a,b]$, let

$$A(x) = f(x) - f(H_{a,b}), \qquad B(x) = C\left(\frac{1}{x} - \frac{1}{H_{a,b}}\right), \tag{50}$$

and define

$$\mathcal{L}_f^{HH}(a,b) = \left| \frac{ab}{b-a} \int_a^b \frac{|A(x)| - |B(x)|}{x^2}\, \mathrm{d}x \right|, \tag{51}$$

$$\mathcal{R}_f^{HH}(a,b) = \frac{ab}{b-a} \int_a^b \frac{\big|\, |A(x)| - |B(x)| \,\big|}{x^2}\, \mathrm{d}x. \tag{52}$$

Theorem 7. *Let $0 < a < b$, let $I \supset [a,b]$ be an open subinterval of $(0,\infty)$ and let $C \in \partial_H f(H_{a,b})$. If $f : I \to \mathbb{R}$ is harmonically convex, then*

$$\begin{aligned} f(H_{a,b}) + \mathcal{L}_f^{HH}(a,b) &\le f(H_{a,b}) + \mathcal{R}_f^{HH}(a,b) \\ &\le M_f(a,b) \\ &\le \frac{1}{2}\left(f(H_{a,b}) + \frac{f(a)+f(b)}{2} \right) \\ &\le \frac{f(a)+f(b)}{2}. \end{aligned} \tag{53}$$

Moreover,

$$M_f(a,b) - f(H_{a,b}) - \mathcal{R}_f^{HH}(a,b) = \frac{2ab}{b-a} \int_a^b \frac{\min\{A(x)^+, B(x)^-\}}{x^2}\, \mathrm{d}x. \tag{54}$$

In particular, $\mathcal{R}_f^{HH}(a,b) = M_f(a,b) - f(H_{a,b})$ if f is monotone on $[a,b]$. If f is differentiable at $H_{a,b}$, then $C = -H_{a,b}^2 f'(H_{a,b})$.

Proof. Let $u = \frac{1}{x}$, $g(u) = f\left(\frac{1}{u}\right)$ and

$$\alpha = \frac{1}{b}, \qquad \beta = \frac{1}{a}, \qquad \vartheta = \frac{\alpha+\beta}{2} = \frac{1}{H_{a,b}}.$$

For every integrable function φ, we have

$$\frac{ab}{b-a} \int_a^b \frac{\varphi(x)}{x^2}\, \mathrm{d}x = \frac{1}{\beta-\alpha} \int_\alpha^\beta \varphi\left(\frac{1}{u}\right) \mathrm{d}u.$$

By (5), $g(u) \ge g(\vartheta) + C(u-\vartheta)$, i.e. C is a subgradient of g at ϑ. For $u \in [\alpha,\beta]$, we have

$$A\left(\frac{1}{u}\right) = g(u) - g(\vartheta) \ge C(u-\vartheta) = B\left(\frac{1}{u}\right),$$

so (22) gives

$$\left|\, \left|A\left(\frac{1}{u}\right)\right| - \left|B\left(\frac{1}{u}\right)\right| \,\right| = g(u) - g(\vartheta) - C(u-\vartheta) - 2\min\left\{ A\left(\frac{1}{u}\right)^+, B\left(\frac{1}{u}\right)^- \right\}.$$

We integrate this equality over $[\alpha, \beta]$. Since $\int_\alpha^\beta (u - \vartheta)\,\mathrm{d}u = 0$, this gives (54) and also the second inequality in (53). The triangle inequality for integrals yields the first inequality in (53).

For the upper bounds, we use the second inequality in (2) for g on $[\alpha, \vartheta]$ and on $[\vartheta, \beta]$:

$$\frac{2}{\beta - \alpha}\int_\alpha^\vartheta g(u)\,\mathrm{d}u \le \frac{g(\alpha) + g(\vartheta)}{2}, \qquad \frac{2}{\beta - \alpha}\int_\vartheta^\beta g(u)\,\mathrm{d}u \le \frac{g(\vartheta) + g(\beta)}{2}.$$

Adding these inequalities and dividing by 2, we get

$$\frac{1}{\beta - \alpha}\int_\alpha^\beta g(u)\,\mathrm{d}u \le \frac{1}{2}\left(g(\vartheta) + \frac{g(\alpha) + g(\beta)}{2}\right).$$

The last inequality in (53) follows from $g(\vartheta) \le \frac{g(\alpha)+g(\beta)}{2}$.

When f is nondecreasing, the function g is nonincreasing and $C \le 0$. Consequently, $A(x)B(x) \ge 0$ for all $x \in [a, b]$ and the minimum in (54) is zero. The case of nonincreasing f is similar. □

Remark 11. *Note that the upper bounds in* (53) *also follow from the refinement of* (49) *obtained by Dragomir Remark 9, inequality (17) in [4]:*

$$\begin{aligned} f(H_{a,b}) &\le \frac{1}{2}\left[f\left(\frac{4ab}{a+3b}\right) + f\left(\frac{4ab}{3a+b}\right)\right] \\ &\le M_f(a,b) \\ &\le \frac{1}{2}\left(f(H_{a,b}) + \frac{f(a)+f(b)}{2}\right) \\ &\le \frac{f(a)+f(b)}{2}. \end{aligned}$$

In particular, it contains

$$M_f(a,b) - f(H_{a,b}) \le \frac{f(a)+f(b)}{2} - M_f(a,b),$$

which may be regarded as a harmonic form of Bullen's inequality [34]. Theorem 7 gives the lower bounds $\mathcal{L}_f^{HH}$ and $\mathcal{R}_f^{HH}$ and identity (54).

Remark 12. *Let f be monotone on $[a, b]$. In this case, $\mathcal{R}_f^{HH}$ gives equality in* (53), *which in general is not true for $\mathcal{L}_f^{HH}$. Indeed, suppose that f is nondecreasing on $[a, b]$. Then $C \le 0$ and* (5) *implies*

$$0 \le B(x) \le A(x) \quad \text{for } x \in [H_{a,b}, b], \qquad B(x) \le A(x) \le 0 \quad \text{for } x \in [a, H_{a,b}].$$

Hence $\big|\,|A(x)| - |B(x)|\,\big| = A(x) - B(x)$ for all $x \in [a, b]$, while

$$|A(x)| - |B(x)| = \begin{cases} -\big(A(x) - B(x)\big), & x \in [a, H_{a,b}], \\ A(x) - B(x), & x \in [H_{a,b}, b]. \end{cases}$$

Let

$$\Delta_- = \frac{ab}{b-a}\int_a^{H_{a,b}} \frac{A(x) - B(x)}{x^2}\,\mathrm{d}x, \qquad \Delta_+ = \frac{ab}{b-a}\int_{H_{a,b}}^b \frac{A(x) - B(x)}{x^2}\,\mathrm{d}x.$$

Then $\Delta_- \ge 0$ and $\Delta_+ \ge 0$. Since

$$\frac{ab}{b-a}\int_a^b \frac{B(x)}{x^2}\,\mathrm{d}x = 0,$$

we get

$$\mathcal{R}_f^{HH}(a,b) = \Delta_- + \Delta_+ = M_f(a,b) - f(H_{a,b}) \qquad \text{and} \qquad \mathcal{L}_f^{HH}(a,b) = |\Delta_+ - \Delta_-|.$$

Therefore, $\mathcal{L}_f^{HH}(a,b) = \mathcal{R}_f^{HH}(a,b)$ *if and only if* $\Delta_- = 0$ *or* $\Delta_+ = 0$. *Since the function* $A - B$ *is continuous and nonnegative, this holds if and only if*

$$f(x) = f(H_{a,b}) + C\left(\frac{1}{x} - \frac{1}{H_{a,b}}\right)$$

for all $x \in [a, H_{a,b}]$ *or for all* $x \in [H_{a,b}, b]$. *Otherwise, the inequality* $\mathcal{L}_f^{HH}(a,b) < \mathcal{R}_f^{HH}(a,b)$ *is strict. Nonincreasing functions* f *can be treated analogously. The logarithmic function is an example with* $0 < \mathcal{L}_{\ln}^{HH}(a,b) < \mathcal{R}_{\ln}^{HH}(a,b)$ *(see Proposition 2).*

Next, we illustrate Theorem 7 by considering the logarithmic function and the differential entropy of a random variable whose reciprocal has the uniform distribution. For $0 < a < b$, we define the probability density

$$p_{a,b}(x) = \frac{ab}{(b-a)x^2}, \qquad a \le x \le b. \tag{55}$$

If U has the uniform distribution on $\left[\frac{1}{b}, \frac{1}{a}\right]$, a simple change of variables shows that $X = \frac{1}{U}$ has density (55). Moreover,

$$\mathbb{E}[f(X)] = \frac{ab}{b-a}\int_a^b \frac{f(x)}{x^2}\,\mathrm{d}x = M_f(a,b).$$

Therefore, Theorem 7 provides estimates of $\mathbb{E}[f(X)]$.

For $f(x) = \ln x$, Theorem 7 gives

$$\ln H_{a,b} + \mathcal{L}_{\ln}^{HH}(a,b) \le \mathbb{E}[\ln X] \le \frac{1}{2}\left(\ln H_{a,b} + \frac{1}{2}\ln(ab)\right), \tag{56}$$

where

$$\mathcal{L}_{\ln}^{HH}(a,b) = \left|\frac{ab}{b-a}\int_a^b \frac{\left|\ln\frac{x}{H_{a,b}}\right| - \left|1 - \frac{H_{a,b}}{x}\right|}{x^2}\,\mathrm{d}x\right|. \tag{57}$$

Since the logarithm is increasing, Theorem 7 also gives

$$\mathcal{R}_{\ln}^{HH}(a,b) = M_{\ln}(a,b) - \ln H_{a,b},$$

so the lower bound with $\mathcal{R}_{\ln}^{HH}$ gives the exact value of $\mathbb{E}[\ln X]$. We now calculate the weaker correction $\mathcal{L}_{\ln}^{HH}$ in (56) explicitly.

Proposition 2. *Let* $0 < a < b$. *Then*

$$\mathcal{L}_{\ln}^{HH}(a,b) = \frac{b\ln\left(\frac{2b}{a+b}\right) + a\ln\left(\frac{2a}{a+b}\right)}{b-a} - \frac{b-a}{2(a+b)}. \tag{58}$$

Equivalently, for $\mathbf{w} = \left(\frac{a}{a+b}, \frac{b}{a+b}\right)$ *and* $\mathbf{u} = \left(\frac{1}{2}, \frac{1}{2}\right)$,

$$\mathcal{L}_{\ln}^{HH}(a,b) = \frac{a+b}{b-a}\mathrm{KL}(\mathbf{w},\mathbf{u}) - \frac{b-a}{2(a+b)}. \tag{59}$$

Moreover,

$$0 < \mathcal{L}_{\ln}^{HH}(a,b) < \mathcal{R}_{\ln}^{HH}(a,b).$$

Proof. Let $\delta = \frac{b-a}{a+b} \in (0,1)$. The substitution $z = \frac{H_{a,b}}{x}$ in (57) gives

$$\mathcal{L}_{\ln}^{HH}(a,b) = \left| \frac{1}{2\delta} \int_{1-\delta}^{1+\delta} \left(|\ln z| - |1-z| \right) \mathrm{d}z \right|.$$

Splitting the integrals at $z = 1$, we get

$$\int_{1-\delta}^{1+\delta} |\ln z| \, \mathrm{d}z = (1+\delta)\ln(1+\delta) + (1-\delta)\ln(1-\delta),$$
$$\int_{1-\delta}^{1+\delta} |1-z| \, \mathrm{d}z = \delta^2.$$

Let

$$F(\delta) = (1+\delta)\ln(1+\delta) + (1-\delta)\ln(1-\delta) - \delta^2.$$

Since $F(0) = F'(0) = 0$ and $F''(\delta) = \frac{2\delta^2}{1-\delta^2} > 0$, we have $F(\delta) > 0$ for $0 < \delta < 1$. Therefore the absolute value can be removed and

$$\mathcal{L}_{\ln}^{HH}(a,b) = \frac{F(\delta)}{2\delta}.$$

Using $1 - \delta = \frac{2a}{a+b}$ and $1 + \delta = \frac{2b}{a+b}$, we get (58). Formula (59) follows from

$$(a+b)\mathrm{KL}(\mathbf{w},\mathbf{u}) = a\ln\left(\frac{2a}{a+b}\right) + b\ln\left(\frac{2b}{a+b}\right).$$

Finally, the function $|\ln z| - |1-z|$ is positive on $(1-\delta, 1)$ and negative on $(1, 1+\delta)$. Hence the triangle inequality for its integral is strict and $\mathcal{L}_{\ln}^{HH}(a,b) < \mathcal{R}_{\ln}^{HH}(a,b)$. □

Recall the definition $\mathsf{h}(X) = -\int_{\mathbb{R}} p_X(x) \ln p_X(x) \, \mathrm{d}x$ of the differential entropy, where p_X denotes the density of X. For the density (55), a simple calculation gives

$$\mathsf{h}(X) = -\ln\left(\frac{ab}{b-a}\right) + 2\,\mathbb{E}[\ln X]. \tag{60}$$

Since $\mathbb{E}[\ln X]$ can be computed directly (see (62) below), the following bounds are not needed for the calculation of $\mathsf{h}(X)$. We give them to show how accurate the bounds from Theorem 7 are in a case where the exact value is known.

Corollary 4. *Let $0 < a < b$ and let X have the density* (55)*. Then*

$$\begin{aligned} \ln\left(\frac{4ab(b-a)}{(a+b)^2}\right) + \frac{2}{b-a}\left(b\ln\left(\frac{2b}{a+b}\right) + a\ln\left(\frac{2a}{a+b}\right)\right) - \frac{b-a}{a+b} \\ \leq \mathsf{h}(X) \\ \leq \ln\left(\frac{2(b-a)\sqrt{ab}}{a+b}\right) < \ln(b-a). \end{aligned} \tag{61}$$

Proof. Combining the left inequality in (56) with (60) and (58), we obtain the lower bound. For the upper bound, we use the right inequality in (56) together with (60):

$$\mathsf{h}(X) \leq -\ln\left(\frac{ab}{b-a}\right) + \ln H_{a,b} + \frac{1}{2}\ln(ab) = \ln\left(\frac{2(b-a)\sqrt{ab}}{a+b}\right).$$

Since $2\sqrt{ab} < a + b$, the last term is smaller than $\ln(b - a)$. □

For comparison, direct integration gives

$$\mathsf{h}(X) = -\ln\left(\frac{ab}{b-a}\right) + \frac{2[b(\ln a + 1) - a(\ln b + 1)]}{b-a}. \tag{62}$$

Example 4. *Let $a = 1$ and $b = 4$. Then $H_{a,b} = 1.6$ and Proposition 2 gives*

$$\mathcal{L}_{\ln}^{HH}(1,4) \approx 0.021241, \qquad \mathcal{R}_{\ln}^{HH}(1,4) \approx 0.067898.$$

By Corollary 4,

$$0.694808 \le \mathsf{h}(X) \le 0.875469 < \ln 3 \approx 1.098612,$$

while (62) gives $\mathsf{h}(X) \approx 0.788122$. Without the term $\mathcal{L}_{\ln}^{HH}$, the lower bound is $\ln\left(\frac{48}{25}\right) \approx 0.652325$. So the relative errors of the lower and the upper bound are about 11.8% and 11.1%, while without $\mathcal{L}_{\ln}^{HH}$ the relative error of the lower bound is about 17.2%.

To judge these errors, we compare them with general bounds for the differential entropy. Every density supported on $[a, b]$ satisfies $\mathsf{h}(X) \le \ln(b - a)$. By Jensen's inequality for the concave logarithm, $\mathsf{h}(X) = -\mathbb{E}[\ln p_X(X)] \ge -\ln \int_{\mathbb{R}} p_X^2(x)\,\mathrm{d}x$, which for the density (55) gives

$$\mathsf{h}(X) \ge \ln\left(\frac{3ab(b-a)}{a^2 + ab + b^2}\right).$$

Finally, the Gaussian maximum entropy bound gives $\mathsf{h}(X) \le \frac{1}{2}\ln(2\pi e \operatorname{Var} X)$. For $a = 1$ and $b = 4$, these bounds are approximately 1.098612, 0.538997 and 1.149536, with relative errors of about 39.4%, 31.6% and 45.9%. So, in this example, the relative errors of the bounds in Corollary 4 are less than half of those of these general estimates, since they take into account the form of the density (55).

As an example in which $M_f(a, b)$ has no elementary closed form, take $f(x) = e^x$, $a = 1$ and $b = 2$. Then

$$M_f(1,2) = 2\int_1^2 \frac{e^x}{x^2}\,\mathrm{d}x.$$

Since $2f'(t) + tf''(t) = (t + 2)e^t \ge 3e$ on $[1, 2]$, the function $f(t) - \frac{3e}{2}t$ is harmonically convex on $[1, 2]$ by Lemma 1. The left inequality in (49) for this function and the upper bounds in (53) give

$$e^{4/3} + \frac{3e}{2}\left(2\ln 2 - \frac{4}{3}\right) \le M_f(1,2) \le \frac{1}{2}\left(e^{4/3} + \frac{e + e^2}{2}\right),$$

i.e., $4.009612 \le M_f(1,2) \approx 4.165741 \le 4.423668$, while (49) gives the lower bound $e^{4/3} \approx 3.793668$.

5. Conclusions

The subject of this paper was the harmonic Jensen functional. Its difference from the lower bound $\mathcal{R}(f)$ was computed exactly and it turned out that the two quantities coincide for monotone functions. The corresponding identity holds for arbitrary convex functions g as well. Further, two-sided estimates were derived for functions f such that $\psi_m(t) = tf(t) - \frac{m}{2}t^2$ is convex for a real m. For $m > 0$, this assumption implies strong harmonic convexity of f, while for $m < 0$, it also includes functions which are not harmonically convex. For twice differentiable harmonically convex functions, the upper estimate obtained in this way

improves the upper bound of Dragomir. The comparison with the Lah–Ribarič bound shows that the two bounds complement each other.

In the application to Csiszár f-divergences, the weighted harmonic mean of the ratios $\frac{q_i}{p_i}$ is equal to $(1+\chi^2(\mathbf{p},\mathbf{q}))^{-1}$, so the obtained bounds contain the χ^2-divergence. These results were applied to the Kullback–Leibler, squared Hellinger and Jeffreys divergences. Numerical examples show that the lower estimates from Theorems 5 and 6 complement each other. With the optimal parameter m^*, both bounds from Theorem 6 are asymptotically sharp as $\mathbf{q} \to \mathbf{p}$. A refinement of the harmonic Hermite–Hadamard inequality was also obtained, together with the integral form of the identity. For $f(t) = \ln t$, the term $\mathcal{L}_{\ln}^{HH}$ was found in closed form and the obtained bounds were compared with the exact value of differential entropy. We believe that the method based on the support inequality (5) can be applied also to the integral form of the harmonic Jensen functional and to other information measures.

Author Contributions: Conceptualization, J.B., S.I.B. and N.L.; methodology, J.B., S.I.B. and N.L.; formal analysis, J.B., S.I.B. and N.L.; writing—original draft preparation, J.B., S.I.B. and N.L.; writing—review and editing, J.B., S.I.B. and N.L. All authors have read and agreed to the published version of the manuscript.

Funding: The authors gratefully acknowledge the support of the projects "Synergy of Mathematics and Physics in Engineering" (SYMPHA, IP-UNIST-11), "Geoinformatics Framework for Sustainable Coastal and Island Development" (GEORAZ, IP-UNIST-17) and "Resilience and Vulnerability Assessment of Diocletian's Palace" (RESILIO, IP-UNIST-19). All projects are funded through the Institutional Research Projects Programme of the University of Split, financed by the Recovery and Resilience Facility of the European Union and approved by the Ministry of Science, Education and Youth of the Republic of Croatia.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

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

References

1. Niculescu, C.P.; Persson, L.-E. *Convex Functions and Their Applications: A Contemporary Approach*, 2nd ed.; Springer: Cham, Switzerland, 2018.
2. Pečarić, J.; Proschan, F.; Tong, Y.L. *Convex Functions, Partial Orderings, and Statistical Applications*; Academic Press: New York, NY, USA, 1992.
3. İşcan, İ. Hermite-Hadamard type inequalities for harmonically convex functions. *Hacet. J. Math. Stat.* **2014**, *43*, 935–942.
4. Dragomir, S.S. New inequalities of Hermite-Hadamard type for HA-convex functions. *J. Numer. Anal. Approx. Theory* **2018**, *47*, 26–41. https://doi.org/10.33993/jnaat471-1119.
5. Merkle, M. Reciprocally convex functions. *J. Math. Anal. Appl.* **2004**, *293*, 210–218.
6. Dragomir, S.S. Inequalities of Jensen type for HA-convex functions. *An. Univ. Oradea Fasc. Mat.* **2020**, *27*, 103–124.
7. Bracamonte, M.; Giménez, J.; Medina, J. Hermite-Hadamard and Fejér-type inequalities for strongly harmonically convex functions. *MATUA* **2016**, *3*, 33–46.
8. İşcan, İ. Ostrowski-type inequalities for harmonically s-convex functions. *Konuralp J. Math.* **2015**, *3*, 63–74.
9. Noor, M.A.; Noor, K.I.; Awan, M.U.; Costache, S. Some integral inequalities for harmonically h-convex functions. *U.P.B. Sci. Bull. Ser. A* **2015**, *77*, 5–16.
10. Bracamonte, M.R.; Medina-Viloria, J.; Vivas-Cortez, M. Some inequalities for reciprocally (s,m)-convex functions in the second sense and applications to special means. *J. Math. Inequal.* **2020**, *14*, 483–500.
11. Li, H.; Saleem, M.S.; Ahmed, I.; Aslam, K.N. Hermite-Hadamard and Fejér-type inequalities for strongly reciprocally (p,h)-convex functions of higher order. *J. Inequal. Appl.* **2023**, *2023*, 57.
12. Vivas-Cortez, M.; Awan, M.U.; Javed, M.Z.; Noor, M.A.; Noor, K.I. A study of uniform harmonic χ-convex functions with respect to Hermite-Hadamard's inequality and its Caputo-Fabrizio fractional analogue and applications. *J. Funct. Spaces* **2021**, *2021*, 9911295.

13. Almalki, Y.; Afzal, W. Some new estimates of Hermite-Hadamard inequalities for harmonic *cr-h*-convex functions via a generalized fractional integral operator on set-valued mappings. *Mathematics* **2023**, *11*, 4041.
14. Ivelić Bradanović, S.; Sayyari, Y.; Barić, J.; Barsam, H. Unified approach to harmonic and reciprocal (h, g, m)-convexity with associated generalized inequalities. *Filomat*, in press.
15. Sayyari, Y.; Dehghanian, M. A new class of convex functions and applications in entropy and analysis. *Chaos Solitons Fractals* **2024**, *181*, 114677.
16. Khan, M.B.; Noor, M.A.; Abdeljawad, T.; Abdalla, B.; Althobaiti, A. Some fuzzy-interval integral inequalities for harmonically convex fuzzy-interval-valued functions. *AIMS Math.* **2022**, *7*, 349–370.
17. Latif, M.A. On symmetrized stochastic harmonically convexity and Hermite-Hadamard type inequalities. *Axioms* **2022**, *11*, 570. doi:10.3390/axioms11100570.
18. Mohsin, B.B.; Awan, M.U.; Noor, M.A.; Riahi, L.; Noor, K.I.; Almutairi, B. New quantum Hermite-Hadamard inequalities utilizing harmonic convexity of functions. *IEEE Access* **2019**, *7*, 20479–20483.
19. Okur, N.; İşcan, İ.; Yüksek Dizdar, E. Hermite-Hadamard type inequalities for harmonically convex stochastic processes. *Int. J. Econ. Adm. Stud.* **2018**, *18*, 281–292. https://doi.org/10.18092/ulikidince.353602.
20. Sayyari, Y.; Butt, S.I.; Dehghanian, M.; Agarwal, P.; Nieto, J.J. New Jensen's bounds for HA-convex mappings with applications to Shannon entropy. *Demonstr. Math.* **2025**, *58*, 20250122. https://doi.org/10.1515/dema-2025-0122.
21. Shah, A.F.; Boulaaras, S.M.; Wong, P.J.Y.; Saleem, M.S.; Shahzad, M.U. Quantum integral inequalities via different variants of parameterized harmonically convex functions on finite intervals with related applications. *Math. Methods Appl. Sci.* **2025**, *48*, 9194–9206.
22. Sun, W.; Wan, H. Hermite-Hadamard type inequalities involving several kinds of fractional calculus for harmonically convex functions. *Fractals* **2023**, *31*, 2350109. https://doi.org/10.1142/S0218348X23501098.
23. Baloch, I.A.; De La Sen, M.; İşcan, İ. Characterizations of classes of harmonic convex functions and applications. *Int. J. Anal. Appl.* **2019**, *17*, 722–733. https://doi.org/10.28924/2291-8639-17-2019-722.
24. Dragomir, S.S. Inequalities of Hermite-Hadamard type for HA-convex functions. *Moroccan J. Pure Appl. Anal.* **2017**, *3*, 83–101. https://doi.org/10.1515/mjpaa-2017-0008.
25. Sayyari, Y.; Furuichi, S.; Barsam, H. Some inequalities related to harmonically convexity with applications in Tsallis entropy and means. *J. Math. Inequal.* **2025**, *19*, 793–810. https://doi.org/10.7153/jmi-2025-19-51.
26. Kian, M. How the type of convexity of the core function affects the Csiszár f-divergence functional. *Math. Inequal. Appl.* **2021**, *24*, 559–570. https://doi.org/10.7153/mia-2021-24-39.
27. Ivelić Bradanović, S.; Lovričević, N. Generalized Jensen and Jensen-Mercer inequalities for strongly convex functions with applications. *J. Inequal. Appl.* **2024**, *2024*, 112. https://doi.org/10.1186/s13660-024-03189-z.
28. Dragomir, S.S.; Scarmozzino, F.P. A refinement of Jensen's discrete inequality for differentiable convex functions. *RGMIA Res. Rep. Coll.* **2002**, *5*.
29. Dragomir, S.S.; Ionescu, N.M. Some converse of Jensen's inequality and applications. *Rev. Anal. Numér. Théor. Approx.* **1994**, *23*, 71–78.
30. Ivelić Bradanović, S. Improvements of Jensen's inequality and its converse for strongly convex functions with applications to strongly f-divergences. *J. Math. Anal. Appl.* **2024**, *531*, 127866. https://doi.org/10.1016/j.jmaa.2023.127866.
31. Csiszár, I. Information-type measures of difference of probability distributions and indirect observations. *Studia Sci. Math. Hungar.* **1967**, *2*, 299–318.
32. Sason, I.; Verdú, S. f-divergence inequalities. *IEEE Trans. Inform. Theory* **2016**, *62*, 5973–6006. https://doi.org/10.1109/TIT.2016.2603151.
33. Jakšić, R.; Pečarić, J. New converses of the Jessen and Lah–Ribarič inequalities. *Math. Inequal. Appl.* **2014**, *17*, 197–216. https://doi.org/10.7153/mia-17-15.
34. Bullen, P.S. Error estimates for some elementary quadrature rules. *Univ. Beograd. Publ. Elektrotehn. Fak. Ser. Mat. Fiz.* **1978**, No. 602–633, 97–103.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.