

RESEARCH

Open Access



# A new bound for the Jensen gap pertaining twice differentiable functions with applications

Shahid Khan<sup>1</sup>, Muhammad Adil Khan<sup>1</sup>, Saad Ihsan Butt<sup>2</sup> and Yu-Ming Chu<sup>3,4\*</sup>

\*Correspondence:

chuyuming2005@126.com

<sup>3</sup>Department of Mathematics, Huzhou University, Huzhou, China

<sup>4</sup>Hunan Provincial Key Laboratory of Mathematical Modeling and Analysis in Engineering, Changsha University of Science & Technology, Changsha, China

Full list of author information is available at the end of the article

## Abstract

In this paper, we present a new bound for the Jensen gap with the help of a Green function. Using the bound, we deduce a converse of the Hölder inequality as well. Finally, we present some applications of the main result in information theory.

**MSC:** 26A51; 26D15

**Keywords:** Jensen inequality; Hölder inequality; Convex function; Green function; Csiszár divergence

## 1 Introduction

In the last few decades, the classical mathematical inequalities and their generalized versions for convex functions have recorded an exponential growth with significant impact in modern analysis [9, 16, 21, 22, 26–28, 30, 31, 33]. They have many applications in numerical quadrature, transform theory, probability, and statistical problems. Specially, they help to establish the uniqueness of the solutions of boundary value problems [8]. In the applied literature of mathematical inequalities, the Jensen inequality is a well-known, paramount and extensively used inequality [2–4, 7, 13, 32]. This inequality is of pivotal importance, because other classical inequalities, such as Hermite–Hadamard’s, Ky–Fan’s, Beckenbach–Dresher’s, Levinson’s, Minkowski’s, arithmetic–geometric, Young’s and Hölder’s inequalities, can be deduced from this inequality. The Jensen inequality and its generalizations, refinements, extensions and converses etc. have many applications in different fields of science, for example electrical engineering [11], mathematical statistics [23], financial economics [24], information theory, guessing and coding [1, 5, 6, 10, 12–15, 17–19, 25]. The discrete Jensen inequality can be found in [20], which states that:

If  $T : [\gamma_1, \gamma_2] \rightarrow \mathbb{R}$  is a convex function and  $s_i \in [\gamma_1, \gamma_2]$ ,  $u_i \geq 0$  for  $i = 1, \dots, n$  with  $\sum_{i=1}^n u_i = U_n > 0$ , then

$$T\left(\frac{1}{U_n} \sum_{i=1}^n u_i s_i\right) \leq \frac{1}{U_n} \sum_{i=1}^n u_i T(s_i).$$

If the function  $T$  is concave then the reverse inequality holds in the above expression.

© The Author(s) 2020. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

To derive the main result, we need the following Green function defined on  $[\gamma_1, \gamma_2] \times [\gamma_1, \gamma_2]$  [29]:

$$G(z, x) = \begin{cases} \frac{(z-\gamma_2)(x-\gamma_1)}{\gamma_2-\gamma_1}, & \gamma_1 \leq x \leq z, \\ \frac{(x-\gamma_2)(z-\gamma_1)}{\gamma_2-\gamma_1}, & z \leq x \leq \gamma_2. \end{cases} \tag{1.1}$$

This function  $G$  is continuous and convex with respect to the two variables  $z$  and  $x$ . Also, the following identity for the function  $T \in C^2[\gamma_1, \gamma_2]$  holds, which is related to the Green function (1.1) [29]:

$$T(z) = \frac{\gamma_2 - z}{\gamma_2 - \gamma_1} T(\gamma_1) + \frac{z - \gamma_1}{\gamma_2 - \gamma_1} T(\gamma_2) + \int_{\gamma_1}^{\gamma_2} G(z, x) T''(x) dx. \tag{1.2}$$

We organize the remaining paper as follows: In Sect. 2, we present a new bound for the Jensen gap for functions whose absolute values of the second derivative are convex, followed by a remark and a proposition presenting a converse of the Hölder inequality. In Sect. 3, we give applications of the main result for the Csiszár  $f$ -divergence functional, the Kullback–Leibler divergence, the Bhattacharyya coefficient, the Hellinger distance, the Rényi divergence, the  $\chi^2$ -divergence, the Shannon entropy and triangular discrimination. Section 4 is devoted to the conclusion of the paper.

## 2 Main result

We begin by presenting our main result.

**Theorem 2.1** *Let  $T \in C^2[\gamma_1, \gamma_2]$  be a function such that  $|T''|$  is convex and  $s_i \in [\gamma_1, \gamma_2]$ ,  $u_i \geq 0$  for  $i = 1, \dots, n$  with  $\sum_{i=1}^n u_i = U_n > 0$ , then*

$$\begin{aligned} & \left| \frac{1}{U_n} \sum_{i=1}^n u_i T(s_i) - T\left(\frac{1}{U_n} \sum_{i=1}^n u_i s_i\right) \right| \\ & \leq \frac{|T''(\gamma_2)| - |T''(\gamma_1)|}{6(\gamma_2 - \gamma_1)} \left( \frac{1}{U_n} \sum_{i=1}^n u_i s_i^3 - \left(\frac{1}{U_n} \sum_{i=1}^n u_i s_i\right)^3 \right) \\ & \quad + \frac{\gamma_2 |T''(\gamma_1)| - \gamma_1 |T''(\gamma_2)|}{2(\gamma_2 - \gamma_1)} \left( \frac{1}{U_n} \sum_{i=1}^n u_i s_i^2 - \left(\frac{1}{U_n} \sum_{i=1}^n u_i s_i\right)^2 \right). \end{aligned} \tag{2.1}$$

*Proof* Using (1.2) in  $\frac{1}{U_n} \sum_{i=1}^n u_i T(s_i)$  and  $T(\frac{1}{U_n} \sum_{i=1}^n u_i s_i)$ , we get

$$\begin{aligned} & \frac{1}{U_n} \sum_{i=1}^n u_i T(s_i) \\ & = \frac{1}{U_n} \sum_{i=1}^n u_i \left( \frac{\gamma_2 - s_i}{\gamma_2 - \gamma_1} T(\gamma_1) + \frac{s_i - \gamma_1}{\gamma_2 - \gamma_1} T(\gamma_2) + \int_{\gamma_1}^{\gamma_2} G(s_i, x) T''(x) dx \right) \end{aligned} \tag{2.2}$$

and

$$T\left(\frac{1}{U_n} \sum_{i=1}^n u_i s_i\right) = \frac{\gamma_2 - \frac{1}{U_n} \sum_{i=1}^n u_i s_i}{\gamma_2 - \gamma_1} T(\gamma_1) + \frac{\frac{1}{U_n} \sum_{i=1}^n u_i s_i - \gamma_1}{\gamma_2 - \gamma_1} T(\gamma_2) + \int_{\gamma_1}^{\gamma_2} G\left(\frac{1}{U_n} \sum_{i=1}^n u_i s_i, x\right) T''(x) dx. \tag{2.3}$$

Subtracting (2.3) from (2.2), we obtain

$$\frac{1}{U_n} \sum_{i=1}^n u_i T(s_i) - T\left(\frac{1}{U_n} \sum_{i=1}^n u_i s_i\right) = \int_{\gamma_1}^{\gamma_2} \left(\frac{1}{U_n} \sum_{i=1}^n u_i G(s_i, x) - G\left(\frac{1}{U_n} \sum_{i=1}^n u_i s_i, x\right)\right) T''(x) dx. \tag{2.4}$$

Taking the absolute value of (2.4), we get

$$\begin{aligned} & \left| \frac{1}{U_n} \sum_{i=1}^n u_i T(s_i) - T\left(\frac{1}{U_n} \sum_{i=1}^n u_i s_i\right) \right| \\ &= \left| \int_{\gamma_1}^{\gamma_2} \left(\frac{1}{U_n} \sum_{i=1}^n u_i G(s_i, x) - G\left(\frac{1}{U_n} \sum_{i=1}^n u_i s_i, x\right)\right) T''(x) dx \right| \\ &\leq \int_{\gamma_1}^{\gamma_2} \left| \frac{1}{U_n} \sum_{i=1}^n u_i G(s_i, x) - G\left(\frac{1}{U_n} \sum_{i=1}^n u_i s_i, x\right) \right| |T''(x)| dx. \end{aligned} \tag{2.5}$$

Using a change of variable we write  $x = t\gamma_1 + (1 - t)\gamma_2, t \in [0, 1]$ . Also, as  $G(z, x)$  is convex, so from (2.5) we have

$$\begin{aligned} & \left| \frac{1}{U_n} \sum_{i=1}^n u_i T(s_i) - T(\bar{s}) \right| \\ &\leq (\gamma_2 - \gamma_1) \int_0^1 \left( \frac{1}{U_n} \sum_{i=1}^n u_i G(s_i, t\gamma_1 + (1 - t)\gamma_2) - G(\bar{s}, t\gamma_1 + (1 - t)\gamma_2) \right) \\ &\quad \times |T''(t\gamma_1 + (1 - t)\gamma_2)| dt, \end{aligned} \tag{2.6}$$

where  $\bar{s} = \frac{1}{U_n} \sum_{i=1}^n u_i s_i$ .

Since  $|T''|$  is a convex function, (2.6) becomes

$$\begin{aligned} & \left| \frac{1}{U_n} \sum_{i=1}^n u_i T(s_i) - T(\bar{s}) \right| \\ &\leq (\gamma_2 - \gamma_1) \int_0^1 \left( \frac{1}{U_n} \sum_{i=1}^n u_i G(s_i, t\gamma_1 + (1 - t)\gamma_2) - G(\bar{s}, t\gamma_1 + (1 - t)\gamma_2) \right) \\ &\quad \times (t|T''(\gamma_1)| + (1 - t)|T''(\gamma_2)|) dt \\ &= (\gamma_2 - \gamma_1) \int_0^1 \left( \frac{1}{U_n} \sum_{i=1}^n u_i G(s_i, t\gamma_1 + (1 - t)\gamma_2) \right) t|T''(\gamma_1)| \end{aligned}$$

$$\begin{aligned}
 & + \frac{1}{U_n} \sum_{i=1}^n u_i G(s_i, t\gamma_1 + (1-t)\gamma_2)(1-t) |T''(\gamma_2)| - G(\bar{s}, t\gamma_1 + (1-t)\gamma_2) \\
 & \times t |T''(\gamma_1)| - G(\bar{s}, t\gamma_1 + (1-t)\gamma_2)(1-t) |T''(\gamma_2)| \Big) dt \\
 = & (\gamma_2 - \gamma_1) \left( |T''(\gamma_1)| \frac{1}{U_n} \sum_{i=1}^n u_i \int_0^1 t G(s_i, t\gamma_1 + (1-t)\gamma_2) dt \right. \\
 & + |T''(\gamma_2)| \frac{1}{U_n} \sum_{i=1}^n u_i \int_0^1 (1-t) G(s_i, t\gamma_1 + (1-t)\gamma_2) dt \\
 & - |T''(\gamma_1)| \int_0^1 t G(\bar{s}, t\gamma_1 + (1-t)\gamma_2) dt \\
 & \left. - |T''(\gamma_2)| \int_0^1 (1-t) G(\bar{s}, t\gamma_1 + (1-t)\gamma_2) dt \right) \\
 = & (\gamma_2 - \gamma_1) \left( |T''(\gamma_1)| \frac{1}{U_n} \sum_{i=1}^n u_i \int_0^1 t G(s_i, t\gamma_1 + (1-t)\gamma_2) dt \right. \\
 & + |T''(\gamma_2)| \frac{1}{U_n} \sum_{i=1}^n u_i \int_0^1 G(s_i, t\gamma_1 + (1-t)\gamma_2) dt \\
 & - |T''(\gamma_2)| \frac{1}{U_n} \sum_{i=1}^n u_i \int_0^1 t G(s_i, t\gamma_1 + (1-t)\gamma_2) dt \\
 & - |T''(\gamma_1)| \int_0^1 t G(\bar{s}, t\gamma_1 + (1-t)\gamma_2) dt \\
 & - |T''(\gamma_2)| \int_0^1 G(\bar{s}, t\gamma_1 + (1-t)\gamma_2) dt \\
 & \left. + |T''(\gamma_2)| \int_0^1 t G(\bar{s}, t\gamma_1 + (1-t)\gamma_2) dt \right). \tag{2.7}
 \end{aligned}$$

Now by using the change of variable  $x = t\gamma_1 + (1-t)\gamma_2$  for  $t \in [0, 1]$ , we obtain

$$\begin{aligned}
 & \int_0^1 t G(s_i, t\gamma_1 + (1-t)\gamma_2) dt \\
 = & \frac{1}{(\gamma_1 - \gamma_2)^3} \left( \frac{\gamma_1^3 s_i}{6} - \frac{\gamma_1 s_i^3}{6} - \frac{5\gamma_2 s_i^3}{6} - \frac{\gamma_2 \gamma_1^2 s_i}{2} - \frac{\gamma_2 \gamma_1^3}{6} + \frac{\gamma_1 \gamma_2 s_i^2}{2} - \frac{\gamma_2^2 s_i^2}{2} \right. \\
 & \left. + \frac{\gamma_2^2 \gamma_1^2}{2} + \frac{\gamma_2^3 s_i}{3} + \gamma_2 s_i^3 - \frac{\gamma_1 \gamma_2^3}{3} \right). \tag{2.8}
 \end{aligned}$$

Replacing  $s_i$  by  $\bar{s}$  in (2.8), we get

$$\begin{aligned}
 & \int_0^1 t G(\bar{s}, t\gamma_1 + (1-t)\gamma_2) dt \\
 = & \frac{1}{(\gamma_1 - \gamma_2)^3} \left( \frac{\gamma_1^3 \bar{s}}{6} - \frac{\gamma_1 (\bar{s})^3}{6} - \frac{5\gamma_2 (\bar{s})^3}{6} - \frac{\gamma_2 \gamma_1^2 \bar{s}}{2} - \frac{\gamma_2 \gamma_1^3}{6} + \frac{\gamma_1 \gamma_2 (\bar{s})^2}{2} \right. \\
 & \left. - \frac{\gamma_2^2 (\bar{s})^2}{2} + \frac{\gamma_2^2 \gamma_1^2}{2} + \frac{\gamma_2^3 \bar{s}}{3} + \gamma_2 (\bar{s})^3 - \frac{\gamma_1 \gamma_2^3}{3} \right). \tag{2.9}
 \end{aligned}$$

Also,

$$\int_0^1 G(s_i, t\gamma_1 + (1-t)\gamma_2) dt = \frac{(s_i^2 - \gamma_2 s_i - \gamma_1 s_i + \gamma_1 \gamma_2)}{2(\gamma_2 - \gamma_1)}. \tag{2.10}$$

Replacing  $s_i$  by  $\bar{s}$  in (2.10), we get

$$\int_0^1 G(\bar{s}, t\gamma_1 + (1-t)\gamma_2) dt = \frac{((\bar{s})^2 - \gamma_2 \bar{s} - \gamma_1 \bar{s} + \gamma_1 \gamma_2)}{2(\gamma_2 - \gamma_1)}. \tag{2.11}$$

Substituting the values from (2.8)–(2.11) in (2.7) and simplifying, we get the required result (2.1). □

*Remark 2.2* If we use the Green functions  $G_1$ – $G_4$  as given in [29] instead of  $G$  in Theorem 2.1, we obtain the same result (2.1).

As an application of the above result, we derive a converse of the Hölder inequality in the following proposition.

**Proposition 2.3** *Let  $q > 1, p \in \mathbb{R}^+ - \{(2, 3) \cup (0, 1]\}$  such that  $\frac{1}{q} + \frac{1}{p} = 1$ . Also, let  $[\gamma_1, \gamma_2]$  be a positive interval and  $(a_1, \dots, a_n), (b_1, \dots, b_n)$  be two positive  $n$ -tuples such that  $\frac{\sum_{i=1}^n a_i b_i}{\sum_{i=1}^n b_i^q}, a_i b_i^{-\frac{q}{p}} \in [\gamma_1, \gamma_2]$  for  $i = 1, \dots, n$ , then*

$$\begin{aligned} & \left( \sum_{i=1}^n a_i^p \right)^{\frac{1}{p}} \left( \sum_{i=1}^n b_i^q \right)^{\frac{1}{q}} - \sum_{i=1}^n a_i b_i \\ & \leq \left( \frac{p(p-1)(\gamma_2^{p-2} - \gamma_1^{p-2})}{6(\gamma_2 - \gamma_1)} \left( \frac{1}{\sum_{i=1}^n b_i^q} \sum_{i=1}^n a_i^3 b_i^{1-2\frac{q}{p}} - \left( \frac{1}{\sum_{i=1}^n b_i^q} \sum_{i=1}^n a_i b_i \right)^3 \right) \right. \\ & \quad + \frac{p(p-1)(\gamma_2 \gamma_1^{p-2} - \gamma_1 \gamma_2^{p-2})}{2(\gamma_2 - \gamma_1)} \left( \frac{1}{\sum_{i=1}^n b_i^q} \sum_{i=1}^n a_i^2 b_i^{1-\frac{q}{p}} \right. \\ & \quad \left. \left. - \left( \frac{1}{\sum_{i=1}^n b_i^q} \sum_{i=1}^n a_i b_i \right)^2 \right) \right)^{\frac{1}{p}} \sum_{i=1}^n b_i^q. \end{aligned} \tag{2.12}$$

*Proof* Let  $T(x) = x^p, x \in [\gamma_1, \gamma_2]$ , then  $T''(x) = p(p-1)x^{p-2} > 0$  and  $|T''|''(x) = p(p-1)(p-2)(p-3)x^{p-4} > 0$ . This shows that  $T$  and  $|T''|$  are convex functions, therefore, using (2.1) for  $T(x) = x^p, u_i = b_i^q$  and  $s_i = a_i b_i^{-\frac{q}{p}}$ , we derive

$$\begin{aligned} & \left( \left( \sum_{i=1}^n a_i^p \right) \left( \sum_{i=1}^n b_i^q \right)^{p-1} - \left( \sum_{i=1}^n a_i b_i \right)^p \right)^{\frac{1}{p}} \\ & \leq \left( \frac{p(p-1)(\gamma_2^{p-2} - \gamma_1^{p-2})}{6(\gamma_2 - \gamma_1)} \left( \frac{1}{\sum_{i=1}^n b_i^q} \sum_{i=1}^n a_i^3 b_i^{1-2\frac{q}{p}} - \left( \frac{1}{\sum_{i=1}^n b_i^q} \sum_{i=1}^n a_i b_i \right)^3 \right) \right. \end{aligned}$$

$$\begin{aligned}
 &+ \frac{p(p-1)(\gamma_2\gamma_1^{p-2} - \gamma_1\gamma_2^{p-2})}{2(\gamma_2 - \gamma_1)} \left( \frac{1}{\sum_{i=1}^n b_i^q} \sum_{i=1}^n a_i^2 b_i^{1-\frac{q}{p}} \right. \\
 &\left. - \left( \frac{1}{\sum_{i=1}^n b_i^q} \sum_{i=1}^n a_i b_i \right)^2 \right)^{\frac{1}{p}} \sum_{i=1}^n b_i^q. \tag{2.13}
 \end{aligned}$$

By utilizing the inequality  $x^\alpha - y^\alpha \leq (x - y)^\alpha$ ,  $0 \leq y \leq x$ ,  $\alpha \in [0, 1]$  for  $x = (\sum_{i=1}^n a_i^p) \times (\sum_{i=1}^n b_i^q)^{p-1}$ ,  $y = (\sum_{i=1}^n a_i b_i)^p$  and  $\alpha = \frac{1}{p}$ , we obtain

$$\begin{aligned}
 &\left( \sum_{i=1}^n a_i^p \right)^{\frac{1}{p}} \left( \sum_{i=1}^n b_i^q \right)^{\frac{1}{q}} - \sum_{i=1}^n a_i b_i \\
 &\leq \left( \left( \sum_{i=1}^n a_i^p \right) \left( \sum_{i=1}^n b_i^q \right)^{p-1} - \left( \sum_{i=1}^n a_i b_i \right)^p \right)^{\frac{1}{p}}. \tag{2.14}
 \end{aligned}$$

Now using (2.14) in (2.13), we get (2.12). □

### 3 Applications in information theory

**Definition 3.1** (Csiszár  $f$ -divergence) Let  $[\gamma_1, \gamma_2] \subset \mathbb{R}$  and  $f : [\gamma_1, \gamma_2] \rightarrow \mathbb{R}$  be a function, then, for  $\mathbf{r} = (r_1, \dots, r_n) \in \mathbb{R}^n$  and  $\mathbf{w} = (w_1, \dots, w_n) \in \mathbb{R}_+^n$  such that  $\frac{r_i}{w_i} \in [\gamma_1, \gamma_2]$  ( $i = 1, \dots, n$ ), the Csiszár  $f$ -divergence functional is defined as [17, 25]

$$\bar{D}_c(\mathbf{r}, \mathbf{w}) = \sum_{i=1}^n w_i f\left(\frac{r_i}{w_i}\right).$$

**Theorem 3.2** Let  $f \in C^2[\gamma_1, \gamma_2]$  be a function such that  $|f''|$  is convex and  $\mathbf{r} = (r_1, \dots, r_n) \in \mathbb{R}^n$ ,  $\mathbf{w} = (w_1, \dots, w_n) \in \mathbb{R}_+^n$  such that  $\frac{\sum_{i=1}^n r_i}{\sum_{i=1}^n w_i}, \frac{r_i}{w_i} \in [\gamma_1, \gamma_2]$  for  $i = 1, \dots, n$ , then

$$\begin{aligned}
 &\left| \frac{1}{\sum_{i=1}^n w_i} \bar{D}_c(\mathbf{r}, \mathbf{w}) - f\left(\frac{\sum_{i=1}^n r_i}{\sum_{i=1}^n w_i}\right) \right| \\
 &\leq \frac{|f''(\gamma_2)| - |f''(\gamma_1)|}{6(\gamma_2 - \gamma_1)} \left( \frac{1}{\sum_{i=1}^n w_i} \sum_{i=1}^n \frac{r_i^3}{w_i^2} - \left( \frac{\sum_{i=1}^n r_i}{\sum_{i=1}^n w_i} \right)^3 \right) \\
 &+ \frac{\gamma_2 |f''(\gamma_1)| - \gamma_1 |f''(\gamma_2)|}{2(\gamma_2 - \gamma_1)} \left( \frac{1}{\sum_{i=1}^n w_i} \sum_{i=1}^n \frac{r_i^2}{w_i} - \left( \frac{\sum_{i=1}^n r_i}{\sum_{i=1}^n w_i} \right)^2 \right). \tag{3.1}
 \end{aligned}$$

*Proof* The result (3.1) can easily be deduced from (2.1) by choosing  $T = f$ ,  $s_i = \frac{r_i}{w_i}$ ,  $u_i = \frac{w_i}{\sum_{i=1}^n w_i}$ . □

**Definition 3.3** (Rényi divergence) For two positive probability distributions  $\mathbf{r} = (r_1, \dots, r_n)$ ,  $\mathbf{w} = (w_1, \dots, w_n)$  and a nonnegative real number  $\mu$  such that  $\mu \neq 1$ , the Rényi divergence is defined as [17, 25]

$$D_{re}(\mathbf{r}, \mathbf{w}) = \frac{1}{\mu - 1} \log \left( \sum_{i=1}^n r_i^\mu w_i^{1-\mu} \right).$$

**Corollary 3.4** Let  $[\gamma_1, \gamma_2] \subseteq \mathbb{R}^+$ . Also let  $\mathbf{r} = (r_1, \dots, r_n)$ ,  $\mathbf{w} = (w_1, \dots, w_n)$  be positive probability distributions and  $\mu > 1$  such that  $\sum_{i=1}^n w_i \left(\frac{r_i}{w_i}\right)^\mu, \left(\frac{r_i}{w_i}\right)^{\mu-1} \in [\gamma_1, \gamma_2]$  for  $i = 1, \dots, n$ . Then

$$\begin{aligned}
 D_{re}(\mathbf{r}, \mathbf{w}) &= \frac{1}{\mu - 1} \sum_{i=1}^n r_i \log \left(\frac{r_i}{w_i}\right)^{\mu-1} \\
 &\leq \frac{\gamma_1 + \gamma_2}{6(1 - \mu)\gamma_1^2\gamma_2^2} \left( \sum_{i=1}^n r_i \left(\frac{r_i}{w_i}\right)^{3(\mu-1)} - \left( \sum_{i=1}^n r_i^\mu w_i^{1-\mu} \right)^3 \right) \\
 &\quad + \frac{\gamma_1^2 + \gamma_1\gamma_2 + \gamma_2^2}{2(\mu - 1)\gamma_1^2\gamma_2^2} \left( \sum_{i=1}^n r_i \left(\frac{r_i}{w_i}\right)^{2(\mu-1)} - \left( \sum_{i=1}^n r_i^\mu w_i^{1-\mu} \right)^2 \right). \tag{3.2}
 \end{aligned}$$

*Proof* Let  $T(x) = -\frac{1}{\mu-1} \log x$ ,  $x \in [\gamma_1, \gamma_2]$ , then  $T''(x) = \frac{1}{(\mu-1)x^2} > 0$  and  $|T'''|(x) = \frac{6}{(\mu-1)x^4} > 0$ . This shows that  $T$  and  $|T''|$  are convex functions, therefore using (2.1) for  $T(x) = -\frac{1}{\mu-1} \log x$ ,  $u_i = r_i$  and  $s_i = \left(\frac{r_i}{w_i}\right)^{\mu-1}$ , we derive (3.2).  $\square$

**Definition 3.5** (Shannon entropy) For a positive probability distribution  $\mathbf{w} = (w_1, \dots, w_n)$ , the Shannon entropy (information divergence) is defined as [17, 25]

$$E_s(\mathbf{w}) = - \sum_{i=1}^n w_i \log w_i.$$

**Corollary 3.6** Let  $[\gamma_1, \gamma_2] \subseteq \mathbb{R}^+$  and  $\mathbf{w} = (w_1, \dots, w_n)$  be a positive probability distribution such that  $\frac{1}{w_i} \in [\gamma_1, \gamma_2]$  for  $i = 1, \dots, n$ . Then

$$\log n - E_s(\mathbf{w}) \leq \frac{\gamma_1^2 + \gamma_1\gamma_2 + \gamma_2^2}{2\gamma_1^2\gamma_2^2} \left( \sum_{i=1}^n \frac{1}{w_i} - n^2 \right) - \frac{\gamma_1 + \gamma_2}{6\gamma_1^2\gamma_2^2} \left( \sum_{i=1}^n \frac{1}{w_i^2} - n^3 \right). \tag{3.3}$$

*Proof* Let  $f(x) = -\log x$ ,  $x \in [\gamma_1, \gamma_2]$ , then  $f''(x) = \frac{1}{x^2} > 0$  and  $|f'''|(x) = \frac{6}{x^4} > 0$ . This shows that  $f$  and  $|f''|$  are convex functions, therefore using (3.1) for  $f(x) = -\log x$  and  $(r_1, \dots, r_n) = (1, \dots, 1)$ , we get (3.3).  $\square$

**Definition 3.7** (Kullback–Leibler divergence) For two positive probability distributions  $\mathbf{r} = (r_1, \dots, r_n)$  and  $\mathbf{w} = (w_1, \dots, w_n)$ , the Kullback–Leibler divergence is defined as [17, 25]

$$D_{kl}(\mathbf{r}, \mathbf{w}) = \sum_{i=1}^n r_i \log \frac{r_i}{w_i}.$$

**Corollary 3.8** Let  $[\gamma_1, \gamma_2] \subseteq \mathbb{R}^+$  and  $\mathbf{r} = (r_1, \dots, r_n)$ ,  $\mathbf{w} = (w_1, \dots, w_n)$  be positive probability distributions such that  $\frac{r_i}{w_i} \in [\gamma_1, \gamma_2]$  for  $i = 1, \dots, n$ . Then

$$D_{kl}(\mathbf{r}, \mathbf{w}) \leq \frac{\gamma_1 + \gamma_2}{2\gamma_1\gamma_2} \left( \sum_{i=1}^n \frac{r_i^2}{w_i} - 1 \right) - \frac{1}{6\gamma_1\gamma_2} \left( \sum_{i=1}^n \frac{r_i^3}{w_i^2} - 1 \right). \tag{3.4}$$

*Proof* Let  $f(x) = x \log x$ ,  $x \in [\gamma_1, \gamma_2]$ , then  $f''(x) = \frac{1}{x} > 0$  and  $|f'''|(x) = \frac{2}{x^3} > 0$ . This shows that  $f$  and  $|f''|$  are convex functions, therefore using (3.1) for  $f(x) = x \log x$ , we get (3.4).  $\square$

**Definition 3.9** ( $\chi^2$ -divergence) Let  $\mathbf{r} = (r_1, \dots, r_n)$ ,  $\mathbf{w} = (w_1, \dots, w_n)$  be positive probability distributions, then  $\chi^2$ -divergence is defined as [25]:

$$D_{\chi^2}(\mathbf{r}, \mathbf{w}) = \sum_{i=1}^n \frac{(r_i - w_i)^2}{w_i}.$$

**Corollary 3.10** If  $[\gamma_1, \gamma_2] \subseteq \mathbb{R}^+$  and  $\mathbf{r} = (r_1, \dots, r_n)$ ,  $\mathbf{w} = (w_1, \dots, w_n)$  are two positive probability distributions such that  $\frac{r_i}{w_i} \in [\gamma_1, \gamma_2]$  for  $i = 1, \dots, n$ , then

$$D_{\chi^2}(\mathbf{r}, \mathbf{w}) \leq \left( \sum_{i=1}^n \frac{r_i^2}{w_i} - 1 \right). \tag{3.5}$$

*Proof* Let  $f(x) = (x - 1)^2$ ,  $x \in [\gamma_1, \gamma_2]$ , then  $f''(x) = 2 > 0$  and  $|f'''(x)| = 0$ . This shows that  $f$  and  $|f''|$  are convex functions, therefore using (3.1) for  $f(x) = (x - 1)^2$ , we obtain (3.5).  $\square$

**Definition 3.11** (Bhattacharyya coefficient) Bhattacharyya coefficient for two positive probability distributions  $\mathbf{r} = (r_1, \dots, r_n)$  and  $\mathbf{w} = (w_1, \dots, w_n)$  is defined by [25]

$$C_b(\mathbf{r}, \mathbf{w}) = \sum_{i=1}^n \sqrt{r_i w_i}.$$

The Bhattacharyya distance is given by  $D_b(\mathbf{r}, \mathbf{w}) = -\log C_b(\mathbf{r}, \mathbf{w})$ .

**Corollary 3.12** Let  $[\gamma_1, \gamma_2] \subseteq \mathbb{R}^+$  and  $\mathbf{r} = (r_1, \dots, r_n)$ ,  $\mathbf{w} = (w_1, \dots, w_n)$  be two positive probability distributions such that  $\frac{r_i}{w_i} \in [\gamma_1, \gamma_2]$  for  $i = 1, \dots, n$ . Then

$$1 - C_b(\mathbf{r}, \mathbf{w}) \leq \frac{\gamma_1^{\frac{3}{2}} - \gamma_2^{\frac{3}{2}}}{24\gamma_1^{\frac{3}{2}}\gamma_2^{\frac{3}{2}}(\gamma_2 - \gamma_1)} \left( \sum_{i=1}^n \frac{r_i^3}{w_i^2} - 1 \right) + \frac{\gamma_2^{\frac{5}{2}} - \gamma_1^{\frac{5}{2}}}{8\gamma_1^{\frac{3}{2}}\gamma_2^{\frac{3}{2}}(\gamma_2 - \gamma_1)} \left( \sum_{i=1}^n \frac{r_i^2}{w_i} - 1 \right). \tag{3.6}$$

*Proof* Let  $f(x) = -\sqrt{x}$ ,  $x \in [\gamma_1, \gamma_2]$ , then  $f''(x) = \frac{1}{4x^{\frac{3}{2}}} > 0$  and  $|f'''(x)| = \frac{15}{16x^{\frac{7}{2}}} > 0$ . This shows that  $f$  and  $|f''|$  are convex functions, therefore using (3.1) for  $f(x) = -\sqrt{x}$ , we obtain (3.6).  $\square$

**Definition 3.13** (Hellinger distance) For two positive probability distributions  $\mathbf{r} = (r_1, \dots, r_n)$ ,  $\mathbf{w} = (w_1, \dots, w_n)$  the Hellinger distance is defined as [25]

$$D_h^2(\mathbf{r}, \mathbf{w}) = \frac{1}{2} \sum_{i=1}^n (\sqrt{r_i} - \sqrt{w_i})^2.$$

**Corollary 3.14** If  $[\gamma_1, \gamma_2] \subseteq \mathbb{R}^+$  and  $\mathbf{r} = (r_1, \dots, r_n)$ ,  $\mathbf{w} = (w_1, \dots, w_n)$  are two positive probability distributions such that  $\frac{r_i}{w_i} \in [\gamma_1, \gamma_2]$  for  $i = 1, \dots, n$ . Then

$$D_h^2(\mathbf{r}, \mathbf{w}) \leq \frac{\gamma_1^{\frac{3}{2}} - \gamma_2^{\frac{3}{2}}}{24\gamma_1^{\frac{3}{2}}\gamma_2^{\frac{3}{2}}(\gamma_2 - \gamma_1)} \left( \sum_{i=1}^n \frac{r_i^3}{w_i^2} - 1 \right) + \frac{\gamma_2^{\frac{5}{2}} - \gamma_1^{\frac{5}{2}}}{8\gamma_1^{\frac{3}{2}}\gamma_2^{\frac{3}{2}}(\gamma_2 - \gamma_1)} \left( \sum_{i=1}^n \frac{r_i^2}{w_i} - 1 \right). \tag{3.7}$$



*Proof* Let  $f(x) = \frac{1}{2}(1 - \sqrt{x})^2$ ,  $x \in [\gamma_1, \gamma_2]$ , then  $f''(x) = \frac{1}{4x^{\frac{3}{2}}} > 0$  and  $|f'''|(x) = \frac{15}{16x^{\frac{5}{2}}} > 0$ . This shows that  $f$  and  $|f''|$  are convex functions, therefore using (3.1) for  $f(x) = \frac{1}{2}(1 - \sqrt{x})^2$ , we deduce (3.7).  $\square$

**Definition 3.15** (Triangular discrimination) For two positive probability distributions  $\mathbf{r} = (r_1, \dots, r_n)$ ,  $\mathbf{w} = (w_1, \dots, w_n)$ , the triangular discrimination is defined as [25]

$$D_{\Delta}(\mathbf{r}, \mathbf{w}) = \sum_{i=1}^n \frac{(r_i - w_i)^2}{r_i + w_i}.$$

**Corollary 3.16** Let  $[\gamma_1, \gamma_2] \subseteq \mathbb{R}^+$  and  $\mathbf{r} = (r_1, \dots, r_n)$ ,  $\mathbf{w} = (w_1, \dots, w_n)$  be positive probability distributions such that  $\frac{r_i}{w_i} \in [\gamma_1, \gamma_2]$  for  $i = 1, \dots, n$ . Then

$$D_{\Delta}(\mathbf{r}, \mathbf{w}) \leq \frac{4((\gamma_1 + 1)^3 - (\gamma_2 + 1)^3)}{3(\gamma_1 + 1)^3(\gamma_2 + 1)^3(\gamma_2 - \gamma_1)} \left( \sum_{i=1}^n \frac{r_i^3}{w_i^2} - 1 \right) + \frac{4(\gamma_2(\gamma_2 + 1)^3 - \gamma_1(\gamma_1 + 1)^3)}{(\gamma_1 + 1)^3(\gamma_2 + 1)^3(\gamma_2 - \gamma_1)} \left( \sum_{i=1}^n \frac{r_i^2}{w_i} - 1 \right). \tag{3.8}$$

*Proof* Let  $f(x) = \frac{(x-1)^2}{(x+1)}$ ,  $x \in [\gamma_1, \gamma_2]$ , then  $f''(x) = \frac{8}{(x+1)^3} > 0$  and  $|f'''|(x) = \frac{96}{(x+1)^5} > 0$ . This shows that  $f$  and  $|f''|$  are convex functions, therefore using (3.1) for  $f(x) = \frac{(x-1)^2}{(x+1)}$ , we get (3.8).  $\square$

### 4 Conclusion

A growing interest in applying the notion of convexity to various fields of science has been recorded, in the last few decades. Convex functions have some rational properties such as differentiability, monotonicity and continuity, which help in their applications. The Jensen inequality generalizes and improves the notion of classical convexity. This inequality and its extensions, improvements, refinements and converses etc and bounds for its gap resolve some difficulties in the modeling of some physical phenomena. For such a purpose, in this paper we have derived a new bound for the Jensen gap for functions whose absolute value of second derivative are convex. Based on this bound, we have deduced a new converse of the Hölder inequality as well. Finally, we have demonstrated new bounds for the Csiszár, Rényi,  $\chi^2$  and Kullback–Leibler divergences etc. in information theory as applications of the main result. The idea and technique used in this paper may be extended to other inequalities to reduce the number of difficulties in the applied literature of mathematical inequalities.

#### Acknowledgements

The authors would like to express their sincere thanks to the editor and the anonymous reviewers for their helpful comments and suggestions.

#### Funding

This research was supported by the Natural Science Foundation of China (11701176, 61673169, 11301127, 11626101, 11601485), and the Science and Technology Research Program of Zhejiang Educational Committee (Y201635325).

#### Availability of data and materials

Not applicable.

### Competing interests

The authors declare that they have no competing interests.

### Authors' contributions

All authors contributed equally to the writing of this paper. All authors read and approved the final manuscript.

### Author details

<sup>1</sup>Department of Mathematics, University of Peshawar, Peshawar, Pakistan. <sup>2</sup>Department of Mathematics, Comsats Institute of Information Technology, Lahore, Pakistan. <sup>3</sup>Department of Mathematics, Huzhou University, Huzhou, China. <sup>4</sup>Hunan Provincial Key Laboratory of Mathematical Modeling and Analysis in Engineering, Changsha University of Science & Technology, Changsha, China.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 8 March 2020 Accepted: 22 June 2020 Published online: 06 July 2020

### References

1. Adil Khan, M., Al-sahwi, Z.M., Chu, Y.-M.: New estimations for Shannon and Zipf–Mandelbrot entropies. *Entropy* **20**(8), Article ID 608 (2018)
2. Adil Khan, M., Ali Khan, G., Ali, T., Kilicman, A.: On the refinement of Jensen's inequality. *Appl. Math. Comput.* **262**, 128–135 (2015)
3. Adil Khan, M., Hanif, M., Hameed Khan, Z.A., Ahmad, K., Chu, Y.-M.: Association of Jensen's inequality for  $s$ -convex function with Csiszár divergence. *J. Inequal. Appl.* **2019**, Article ID 162 (2019)
4. Adil Khan, M., Khalid, S., Pečarić, J.: Improvement of Jensen's inequality in terms of Gâteaux derivatives for convex functions in linear spaces with applications. *Kyungpook Math. J.* **52**, 495–511 (2012)
5. Adil Khan, M., Pečarić, Đ., Pečarić, J.: Bounds for Shannon and Zipf–Mandelbrot entropies. *Math. Methods Appl. Sci.* **40**(18), 7316–7322 (2017)
6. Adil Khan, M., Pečarić, Đ., Pečarić, J.: Bounds for Csiszár divergence and hybrid Zipf–Mandelbrot entropy. *Math. Methods Appl. Sci.* **42**(18), 7411–7424 (2019)
7. Adil Khan, M., Pečarić, J., Chu, Y.-M.: Refinements of Jensen's and McShane's inequalities with applications. *AIMS Math.* **5**(5), 4931–4945 (2020)
8. Agarwal, R., Yadav, M.P., Baleanu, D., Purohit, S.D.: Existence and uniqueness of miscible flow equation through porous media with a nonsingular fractional derivative. *AIMS Math.* **5**(2), 1062–1073 (2020)
9. Bainov, D., Simeonov, P.: *Integral Inequalities and Applications*. Kluwer Academic, Boston (1992)
10. Budimir, I., Dragomir, S.S., Pečarić, J.: Further reverse results for Jensen's discrete inequality and applications in information theory. *J. Inequal. Pure Appl. Math.* **2**(1), Article 5 (2001)
11. Cloud, M.J., Drachman, B.C., Lebedev, L.P.: *Inequalities with Applications to Engineering*. Springer, Heidelberg (2014)
12. Dragomir, S.S.: A converse of the Jensen inequality for convex mappings of several variables and applications. *Acta Math. Vietnam.* **29**(1), 77–88 (2004)
13. Dragomir, S.S.: A new refinement of Jensen's inequality in linear spaces with applications. *Math. Comput. Model.* **52**, 1497–1505 (2010)
14. Dragomir, S.S.: A refinement of Jensen's inequality with applications for  $f$ -divergence measures. *Taiwan. J. Math.* **14**(1), 153–164 (2010)
15. Dragomir, S.S., Goh, C.J.: A counterpart of Jensen's discrete inequality for differentiable convex mappings and applications in information theory. *Math. Comput. Model.* **24**(2), 1–11 (1996)
16. Gujrati, P.D.: Jensen inequality and the second law. *Phys. Lett. A* (2020). <https://doi.org/10.1016/j.physleta.2020.126460>
17. Horváth, L., Pečarić, Đ., Pečarić, J.: Estimations of  $f$ - and Rényi divergences by using a cyclic refinement of the Jensen's inequality. *Bull. Malays. Math. Sci. Soc.* **42**(3), 933–946 (2019)
18. Khan, K.A., Niaz, T., Pečarić, Đ., Pečarić, J.: Refinement of Jensen's inequality and estimation of  $f$ - and Rényi divergence via Montgomery identity. *J. Inequal. Appl.* **2018**, Article ID 318 (2018)
19. Khan, S., Adil Khan, M., Chu, Y.-M.: Converses of the Jensen inequality derived from the Green functions with applications in information theory. *Math. Methods Appl. Sci.* **43**, 2577–2587 (2020)
20. Khan, S., Adil Khan, M., Chu, Y.-M.: New converses of Jensen inequality via Green functions with applications. *Rev. R. Acad. Cienc. Exactas Fis. Nat., Ser. A Mat.* **114**, 114 (2020). <https://doi.org/10.1007/s13398-020-00843-1>
21. Kian, M.: Operator Jensen inequality for superquadratic functions. *Linear Algebra Appl.* **456**, 82–87 (2014)
22. Kim, J.-H.: Further improvement of Jensen inequality and application to stability of time-delayed systems. *Automatica* **64**, 121–125 (2016)
23. Liao, J.G., Berg, A.: Sharpening Jensen's inequality. *Am. Stat.* **73**(3), 278–281 (2019)
24. Lin, Q.: Jensen inequality for superlinear expectations. *Stat. Probab. Lett.* **151**, 79–83 (2019)
25. Lovričević, N., Pečarić, Đ., Pečarić, J.: Zipf–Mandelbrot law,  $f$ -divergences and the Jensen-type interpolating inequalities. *J. Inequal. Appl.* **2018**, Article ID 36 (2018)
26. Mishra, A.M., Baleanu, D., Tchier, F., Purohit, S.D.: Certain results comprising the weighted Chebyshev functional using pathway fractional integrals. *Mathematics* **7**(10), Article ID 896 (2019)
27. Mishra, A.M., Kumar, D., Purohit, S.D.: Unified integral inequalities comprising pathway operators. *AIMS Math.* **5**(1), 399–407 (2020)
28. Niaz, T., Khan, K.A., Pečarić, J.: On refinement of Jensen's inequality for 3-convex function at a point. *Turk. J. Ineq.* **4**(1), 70–80 (2020)
29. Pečarić, Đ., Pečarić, J., Rodić, M.: About the sharpness of the Jensen inequality. *J. Inequal. Appl.* **2018**, Article ID 337 (2018)

30. Purohit, S.D., Jolly, N., Bansal, M.K., Singh, J., Kumar, D.: Chebyshev type inequalities involving the fractional integral operator containing multi-index Mittag-Leffler function in the kernel. *Appl. Appl. Math. Spec. Issue* **6**, 29–38 (2020)
31. Rashid, S., Noor, M.A., Noor, K.I., Safdar, F., Chu, Y.-M.: Hermite–Hadamard type inequalities for the class of convex functions on time scale. *Mathematics* **7**, Article ID 956 (2019)
32. Song, Y.-Q., Adil Khan, M., Zaheer Ullah, S., Chu, Y.-M.: Integral inequalities involving strongly convex functions. *J. Funct. Spaces* **2018**, Article ID 6595921 (2018)
33. Wang, M.-K., Zhang, W., Chu, Y.-M.: Monotonicity, convexity and inequalities involving the generalized elliptic integrals. *Acta Math. Sci.* **39B**(5), 1440–1450 (2019)

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

---

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)

---