IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received April 25, 2020, accepted May 18, 2020, date of publication May 25, 2020, date of current version June 5, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2997397

A New Bound for the Jensen Gap With Applications in Information Theory

MUHAMMAD ADIL KHAN[1], SHAHID KHAN[1], AND YUMING CHU[2,3]
[1]Department of Mathematics, University of Peshawar, Peshawar 25000, Pakistan
[2]Department of Mathematics, Huzhou University, Huzhou 313000, China
[3]Hunan Provincial Key Laboratory of Mathematical Modeling and Analysis in Engineering, Changsha University of Science and Technology, Changsha 410114, China

Corresponding author: Yuming Chu (chuyuming2005@126.com)

This work was supported by the Natural Science Foundation of China under Grant 61673169, Grant 11301127, Grant 11701176, Grant 11626101, and Grant 11601485.

ABSTRACT In this manuscript, we adopt a novel approach to present a new bound for the Jensen gap for functions whose double derivatives in absolute function, are convex. We demonstrate two numerical experiments to verify the main result and to discuss the tightness of the bound. Then we utilize the bound for deriving two new converses of the Hölder inequality and a bound for the Hermite-Hadamard gap. Finally, we demonstrate applications of the main result for various divergences in information theory. Also, we present a numerical example to verify the bound for Shannon entropy.

INDEX TERMS Jensen inequality, Hölder inequality, Hermite-Hadamard inequality, convex function, Green function, Csiszár divergence.

I. INTRODUCTION AND PRELIMINARIES

The field of mathematical inequalities and their applications has recorded an exponential and significant growth in the last three decades with considerable impact in various areas of Science such as Engineering [12], Economics [25], Mathematical Statistics [24], Qualitative Theory of Integral and Differential Equations [21], Information Theory and Coding [16], [18] etc. It is noteworthy that many innovative ideas about mathematical inequalities and their applications in various areas of Science can be developed by convexity [3], [4], [9], [11], [14], [27], [29], [32], [33]. One of the most important inequality for convex functions is Jensen inequality, which generalizes the classical convexity. This inequality is of pivotal importance, because other classical inequalities for example Hermite-Hadamard, Hölder, Ky-Fan, Beckenbach-Dresher, Minkowski's, arithmetic-geometric and Young's inequalities etc can be deduced from this inequality. An extensive literature exists regarding estimates for the Jensen gap and their applications in many branches of Science [1]–[8], [12], [15], [20], [24]–[26], [28]. In this manuscript, we present a new bound as an estimate for the Jensen gap.

In the following theorem, Jensen integral inequality has been presented [17]:

Theorem 1: Let $[\alpha_1, \alpha_2] \subset R$ *and* $h_1, h_2 : [c_1, c_2] \to R$ *be two functions such that* $h_1(t) \in [\alpha_1, \alpha_2]$, $\forall\, t \in [c_1, c_2]$. *Let the function* $\psi : [\alpha_1, \alpha_2] \to R$ *be convex and* h_2, h_1h_2, $(\psi \circ h_1).h_2$ *are integrable functions on* $[c_1, c_2]$. *Also suppose that* $h_2(t) \geq 0$ *for all* $t \in [c_1, c_2]$ *and* $\int_{c_1}^{c_2} h_2(t)dt > 0$, *then*

$$\psi\left(\frac{\int_{c_1}^{c_2} h_1(t)h_2(t)dt}{\int_{c_1}^{c_2} h_2(t)dt}\right) \leq \frac{\int_{c_1}^{c_2}(\psi \circ h_1)(t)h_2(t)dt}{\int_{c_1}^{c_2} h_2(t)dt}. \tag{1}$$

For deriving the main result, we need the following Green function defined on $[\omega_1, \omega_2] \times [\omega_1, \omega_2]$ [22]:

$$G(s,x) = \begin{cases} \dfrac{(s-\omega_2)(x-\omega_1)}{\omega_2-\omega_1}, & \omega_1 \leq x \leq s, \\ \dfrac{(x-\omega_2)(s-\omega_1)}{\omega_2-\omega_1}, & s \leq x \leq \omega_2. \end{cases} \tag{2}$$

This is a convex function with respect to both the variables s and x. Also, the following identity for the function $T \in C^2[\omega_1, \omega_2]$ holds, which is related to the Green function (2) [22]:

$$T(s) = \frac{\omega_2 - s}{\omega_2-\omega_1}T(\omega_1) + \frac{s-\omega_1}{\omega_2-\omega_1}T(\omega_2) + \int_{\omega_1}^{\omega_2} G(s,x)T''(x)dx. \tag{3}$$

We organize the remaining paper as: In Section II, we present a new main result following by a remark, two numerical experiments, a proposition, two corollaries and an another remark which completes the section. Numerical experiments give the surety of the tightness of the bound which is presented as the main result, Proposition 1 presents a converse

The associate editor coordinating the review of this manuscript and approving it for publication was Khmaies Ouahada.

of the Hölder inequality, Corollary 1 demonstrate another converse of the Hölder inequality while Corollary 2 presents a bound for the Hermite-Hadamard gap. In Section III, we give applications of the main result for Csiszár divergence, Shannon entropy, Kullback-Leibler divergence, Jeffrey's divergence, Bhattacharyya coefficient, Hellinger distance and Triangular discrimination.

II. MAIN RESULT

In the following theorem, we present a new bound for the Jensen gap by using functions whose double derivatives in the absolute function, are convex.

Theorem 2: Let $T \in C^2[\omega_1, \omega_2]$ be a function such that $|T''|$ is convex. Let h, f be two real valued functions defined on $[a_1, a_2]$ such that $h(y) \in [\omega_1, \omega_2]$ for all $y \in [a_1, a_2]$ with $f, hf, (T \circ h)f$ as integrable functions on $[a_1, a_2]$. Also suppose that $f(y) \geq 0$ on $[a_1, a_2]$ with $\int_{a_1}^{a_2} f(y)dy := F > 0$, then

$$\begin{aligned}&\left|\frac{1}{F}\int_{a_1}^{a_2}(T\circ h)(y)f(y)dy - T\left(\frac{1}{F}\int_{a_1}^{a_2}h(y)f(y)dy\right)\right| \\ &\leq \frac{|T''(\omega_2)| - |T''(\omega_1)|}{6(\omega_2-\omega_1)}\left(\frac{1}{F}\int_{a_1}^{a_2}h^3(y)f(y)dy\right. \\ &\quad \left.-\left(\frac{1}{F}\int_{a_1}^{a_2}h(y)f(y)dy\right)^3\right) + \frac{\omega_2|T''(\omega_1)| - \omega_1|T''(\omega_2)|}{2(\omega_2-\omega_1)} \\ &\quad \times\left(\frac{1}{F}\int_{a_1}^{a_2}h^2(y)f(y)dy - \left(\frac{1}{F}\int_{a_1}^{a_2}h(y)f(y)dy\right)^2\right). \quad (4)\end{aligned}$$

Proof: Using (3) in $\frac{1}{F}\int_{a_1}^{a_2}(T\circ h)(y)f(y)dy$ and $T\left(\frac{1}{F}\int_{a_1}^{a_2}h(y)f(y)dy\right)$, we get

$$\begin{aligned}&\frac{1}{F}\int_{a_1}^{a_2}(T\circ h)(y)f(y)dy \\ &= \frac{1}{F}\int_{a_1}^{a_2}\left(\frac{\omega_2-h(y)}{\omega_2-\omega_1}T(\omega_1) + \frac{h(y)-\omega_1}{\omega_2-\omega_1}T(\omega_2)\right. \\ &\quad \left.+\int_{\omega_1}^{\omega_2}G(h(y),x)T''(x)dx\right)f(y)dy. \quad (5)\end{aligned}$$

and

$$\begin{aligned}&T\left(\frac{1}{F}\int_{a_1}^{a_2}h(y)f(y)dy\right) \\ &= \frac{\omega_2 - \frac{1}{F}\int_{a_1}^{a_2}h(y)f(y)dy}{\omega_2-\omega_1}T(\omega_1) + \frac{\frac{1}{F}\int_{a_1}^{a_2}h(y)f(y)dy - \omega_1}{\omega_2-\omega_1} \\ &\quad \times T(\omega_2) + \int_{\omega_1}^{\omega_2}G\left(\frac{1}{F}\int_{a_1}^{a_2}h(y)f(y)dy, x\right)T''(x)dx. \quad (6)\end{aligned}$$

Subtracting (6) from (5), we obtain the following result

$$\begin{aligned}&\frac{1}{F}\int_{a_1}^{a_2}(T\circ h)(y)f(y)dy - T\left(\frac{1}{F}\int_{a_1}^{a_2}h(y)f(y)dy\right) \\ &= \int_{\omega_1}^{\omega_2}\left(\frac{1}{F}\int_{a_1}^{a_2}G(h(y),x)f(y)dy\right. \\ &\quad \left.-G\left(\frac{1}{F}\int_{a_1}^{a_2}h(y)f(y)dy, x\right)\right)T''(x)dx. \quad (7)\end{aligned}$$

Taking absolute of (7), we have

$$\begin{aligned}&\left|\frac{1}{F}\int_{a_1}^{a_2}(T\circ h)(y)f(y)dy - T\left(\frac{1}{F}\int_{a_1}^{a_2}h(y)f(y)dy\right)\right| \\ &= \left|\int_{\omega_1}^{\omega_2}\left(\frac{1}{F}\int_{a_1}^{a_2}G(h(y),x)f(y)dy\right.\right. \\ &\quad \left.\left.-G\left(\frac{1}{F}\int_{a_1}^{a_2}h(y)f(y)dy, x\right)\right)T''(x)dx\right| \\ &\leq \int_{\omega_1}^{\omega_2}\left|\left(\frac{1}{F}\int_{a_1}^{a_2}G(h(y),x)f(y)dy\right.\right. \\ &\quad \left.\left.-G\left(\frac{1}{F}\int_{a_1}^{a_2}h(y)f(y)dy, x\right)\right)\right||T''(x)|dx. \quad (8)\end{aligned}$$

Using the change of variable $x = t\omega_1 + (1-t)\omega_2$ for $t \in [0,1]$. Also as $G(s,x)$ is convex, so from (8), we get

$$\begin{aligned}&\left|\frac{1}{F}\int_{a_1}^{a_2}(T\circ h)(y)f(y)dy - T\left(\frac{1}{F}\int_{a_1}^{a_2}h(y)f(y)dy\right)\right| \\ &\leq (\omega_2-\omega_1)\int_0^1\left(\frac{1}{F}\int_{a_1}^{a_2}G(h(y), t\omega_1+(1-t)\omega_2)f(y)dy\right. \\ &\quad \left.-G(\bar{h}, t\omega_1+(1-t)\omega_2)\right)|T''(t\omega_1+(1-t)\omega_2)|dt, \quad (9)\end{aligned}$$

where

$$\bar{h} = \frac{1}{F}\int_{a_1}^{a_2}h(y)f(y)dy.$$

Since $|T''|$ is a convex function, therefore (9) takes the form

$$\begin{aligned}&\left|\frac{1}{F}\int_{a_1}^{a_2}(T\circ h)(y)f(y)dy - T\left(\frac{1}{F}\int_{a_1}^{a_2}h(y)f(y)dy\right)\right| \\ &\leq (\omega_2-\omega_1)\int_0^1\left(\frac{1}{F}\int_{a_1}^{a_2}G\Big(h(y), t\omega_1+(1-t)\omega_2\Big)f(y)dy\right. \\ &\quad \left.-G(\bar{h}, t\omega_1+(1-t)\omega_2)\right)\Big(t|T''(\omega_1)|+(1-t)|T''(\omega_2)|\Big)dt \\ &= (\omega_2-\omega_1) \\ &\quad \times\left(\int_0^1 t|T''(\omega_1)|\frac{1}{F}\int_{a_1}^{a_2}G\Big(h(y), t\omega_1+(1-t)\omega_2\Big)f(y)dydt\right. \\ &\quad +\int_0^1(1-t)|T''(\omega_2)|\frac{1}{F}\int_{a_1}^{a_2}G\Big(h(y), t\omega_1 \\ &\quad +(1-t)\omega_2\Big)f(y)dydt \\ &\quad -\int_0^1 t|T''(\omega_1)|G(\bar{h}, t\omega_1+(1-t)\omega_2)dt \\ &\quad \left.-\int_0^1(1-t)|T''(\omega_2)|G(\bar{h}, t\omega_1+(1-t)\omega_2)dt\right). \\ &= (\omega_2-\omega_1) \\ &\quad \times\left(\frac{|T''(\omega_1)|}{F}\int_{a_1}^{a_2}f(y)\left(\int_0^1 tG\Big(h(y), t\omega_1+(1-t)\omega_2\Big)dt\right)dy\right. \\ &\quad +\frac{|T''(\omega_2)|}{F}\int_{a_1}^{a_2}f(y)\left(\int_0^1 G\Big(h(y), t\omega_1+(1-t)\omega_2\Big)dt\right)dy \\ &\quad -\frac{|T''(\omega_2)|}{F}\int_{a_1}^{a_2}f(y)\left(\int_0^1 tG\Big(h(y), t\omega_1+(1-t)\omega_2\Big)dt\right)dy\end{aligned}$$

$$- |T''(\omega_1)| \int_0^1 tG(\bar{h}, t\omega_1 + (1-t)\omega_2)dt$$
$$- |T''(\omega_2)| \int_0^1 G(\bar{h}, t\omega_1 + (1-t)\omega_2)dt$$
$$+ |T''(\omega_2)| \int_0^1 tG(\bar{h}, t\omega_1 + (1-t)\omega_2)dt\Bigg). \tag{10}$$

Now by using the change of variable $x = t\omega_1 + (1-t)\omega_2$ for $t \in [0, 1]$, we obtain

$$\int_0^1 tG(h(y), t\omega_1 + (1-t)\omega_2)dt = \frac{1}{(\omega_1 - \omega_2)^3}\Bigg(\frac{\omega_1^3 h(y)}{6} - \frac{\omega_1 h^3(y)}{6} - \frac{5\omega_2 h^3(y)}{6} - \frac{\omega_2\omega_1^2 h(y)}{2} - \frac{\omega_2\omega_1^3}{6} + \frac{\omega_1\omega_2 h^2(y)}{2} - \frac{\omega_2^2 h^2(y)}{2} + \frac{\omega_2^2\omega_1^2}{2} + \frac{\omega_2^3 h(y)}{3} + \omega_2 h^3(y) - \frac{\omega_1\omega_2^3}{3}\Bigg). \tag{11}$$

Replacing $h(y)$ by $\bar{h}$ in (11), we get

$$\int_0^1 tG(\bar{h}, t\omega_1 + (1-t)\omega_2)dt = \frac{1}{(\omega_1 - \omega_2)^3}\Bigg(\frac{\omega_1^3 \bar{h}}{6} - \frac{\omega_1(\bar{h})^3}{6} - \frac{5\omega_2(\bar{h})^3}{6} - \frac{\omega_2\omega_1^2\bar{h}}{2} - \frac{\omega_2\omega_1^3}{6} + \frac{\omega_1\omega_2(\bar{h})^2}{2} - \frac{\omega_2^2(\bar{h})^2}{2} + \frac{\omega_2^2\omega_1^2}{2} + \frac{\omega_2^3\bar{h}}{3} + \omega_2(\bar{h})^3 - \frac{\omega_1\omega_2^3}{3}\Bigg). \tag{12}$$

Also,

$$\int_0^1 G(h(y), t\omega_1 + (1-t)\omega_2)dt = \frac{(h^2(y) - \omega_2 h(y) - \omega_1 h(y) + \omega_1\omega_2)}{2(\omega_2 - \omega_1)}. \tag{13}$$

Replacing $h(y)$ by $\bar{h}$ in (13), we get

$$\int_0^1 G(\bar{h}, t\omega_1 + (1-t)\omega_2)dt = \frac{((\bar{h})^2 - \omega_2\bar{h} - \omega_1\bar{h} + \omega_1\omega_2)}{2(\omega_2 - \omega_1)}. \tag{14}$$

Substituting the values from (11)–(14) in (10) and simplifying, we get the result (4). □

Remark 1: If we use the Green functions $G_2 - G_5$ as given in [22] instead of G in Theorem 2, we obtain the same result (4).

Now we demonstrate some numerical experiments to show the tightness of the bound (4).

Example 1: Let $T(y) = y^4, h(y) = y, f(y) = 1$ for all $y \in [0, 1]$ then $T''(y) = 12y^2 > 0$, $|T''|''(y) = 24 > 0$ for all $y \in [0, 1]$. Which shows that T and $|T''|$ are convex functions. Also, $h(y) \in [0, 1]$ for all $y \in [0, 1]$, therefore using inequality (4) for these facts with $[\omega_1, \omega_2] = [a_1, a_2] = [0, 1]$, we obtain $\int_0^1 T(h(y))dy - T\left(\int_0^1 h(y)dy\right) = 0.2 - 0.0625 = 0.1375$ and its corresponding right hand side gives 0.25. Thus from inequality (4) we conclude that

$$0.1375 < 0.25. \tag{15}$$

Now taking right hand side of inequality (5) in [13], we get

$$\frac{1}{2}\|T''\|_{L^\infty([0,1])}.\Big[\|h - c\|_{L^2}^2 + \|h - c\|_{L^1}^2\Big] = 6\Big[\frac{3c^2 - 3c + 1}{3} + \Big(c^2 - c + \frac{1}{2}\Big)^2\Big] = 6g(c). \tag{16}$$

It is important to note that $g(c)$ attains its minimum value at $c = 0.5$ which is $g(0.5) \approx 0.1458$ and thus from (16) we obtain

$$\frac{1}{2}\|T''\|_{L^\infty([0,1])}.\Big[\|h - c\|_{L^2}^2 + \|h - c\|_{L^1}^2\Big] \approx 0.8748.$$

Hence from inequality (5) in [13], we get

$$0.1375 < 0.8748. \tag{17}$$

Similarly taking right hand side of inequality (8) in [13], we get

$$\frac{1}{2}\|T''\|_{L^\infty([0,1])}.\|h - c\|_{L^2}^2 - \frac{1}{2}\inf_{[0,1]} T''\Big[\int_0^1 (h(y) - c)dy\Big]^2 = 6\Big[\frac{3c^2 - 3c + 1}{3}\Big] = 6l(c). \tag{18}$$

Now $l(c)$ attains its minimum value at $c = 0.5$ which is $l(0.5) \approx 0.0833$ and thus from (18) we deduce

$$\frac{1}{2}\|T''\|_{L^\infty([0,1])}.\|h - c\|_{L^2}^2 - \frac{1}{2}\inf_{[0,1]} T''\Big[\int_0^1 (h(y) - c)dy\Big]^2 \approx 0.4998.$$

Hence from inequality (8) in [13], we get

$$0.1375 < 0.4998. \tag{19}$$

From (15), (17) and (19), it can easily be concluded that the bound in (4) for the Jensen gap is better than the bounds in (5), (8) from [13]. Also, inequality (15) verifies the tightness of the bound in (4), towards the Jensen gap.

Example 2: Let $T(y) = e^y, h(y) = y^2, f(y) = 1$ for all $y \in [0, 1]$ then $T''(y) = e^y > 0$, $|T''|''(y) = e^y > 0$ for all $y \in [0, 1]$. Which shows that T and $|T''|$ are convex functions. Also, $h(y) \in [0, 1]$ for all $y \in [0, 1]$, therefore using inequality (4) for the above facts with $[\omega_1, \omega_2] = [a_1, a_2] = [0, 1]$ we obtain $\int_0^1 T(h(y))dy - T\left(\int_0^1 h(y)dy\right) = 1.4627 - 1.3956 = 0.0671$ and right hand side of inequality (4) gives 0.0748. Thus from inequality (4), we deduce the following result

$$0.0671 < 0.0748. \tag{20}$$

Now taking right hand side of inequality (5) in [13], we get

$$\frac{1}{2}\|T''\|_{L^\infty([0,1])}.\Big[\|h-c\|_{L^2}^2+\|h-c\|_{L^1}^2\Big]$$
$$=1.3592\Big[c^2-\frac{2}{3}c+\frac{1}{5}+\Big(\frac{4}{3}c^{\frac{3}{2}}-c+\frac{1}{3}\Big)^2\Big]$$
$$=1.3592g(c). \quad (21)$$

Now $g(c)$ attains its minimum value at $c \approx 0.31$ which is $g(0.31)\approx 0.1536$ and thus from (21) we get

$$\frac{1}{2}\|T''\|_{L^\infty([0,1])}.\Big[\|h-c\|_{L^2}^2+\|h-c\|_{L^1}^2\Big]\approx 0.2088.$$

Hence from inequality (5) in [13], we get

$$0.0671 < 0.2088. \quad (22)$$

Similarly taking right hand side of inequality (8) in [13], we obtain

$$\frac{1}{2}\|T''\|_{L^\infty([0,1])}.\|h-c\|_{L^2}^2-\frac{1}{2}\inf_{[0,1]}T''\Big[\int_0^1(h(y)-c)dy\Big]^2$$
$$=1.3592\Big[c^2-\frac{2}{3}c+\frac{1}{5}\Big]-\frac{1}{2}\Big[\frac{3c-1}{3}\Big]^2=l(c). \quad (23)$$

Now $l(c)$ attains its minimum value at $c\approx 0.33$ which is $l(0.33)\approx 0.1208$ and thus from (23) we get

$$\frac{1}{2}\|T''\|_{L^\infty([0,1])}.\|h-c\|_{L^2}^2-\frac{1}{2}\inf_{[0,1]}T''\Big[\int_0^1(h(y)-c)dy\Big]^2\approx 0.1208.$$

Hence from inequality (8) in [13], we get

$$0.0671 < 0.1208. \quad (24)$$

Now inequalities in (20), (22) and (24) show that the bound in (4) for the Jensen gap is better than the bounds in (5), (8) from [13]. Also, inequality (20) gives the surety of the tightness of the bound in (4), towards the Jensen gap.

In the following proposition, we present a converse of the Hölder inequality as an application of the above theorem.

Proposition 1: Let $p_1 \in R^+ - \{(2,3)\cup(0,1]\}$, $p_2 > 1$ such that $\frac{1}{p_1}+\frac{1}{p_2}=1$ and $\zeta_1, \zeta_2 : [a_1,a_2]\to R^+$ be two functions such that $\zeta_1^{p_1}(y)$, $\zeta_2^{p_2}(y)$, $\zeta_1(y)\zeta_2(y)$ and $\zeta_1^2(y)\zeta_2^{1-\frac{p_2}{p_1}}(y)$ are integrable on $[a_1,a_2]$. Also, let $[\omega_1,\omega_2]$ be a positive interval such that $\zeta_1(y)\zeta_2^{-\frac{p_2}{p_1}}(y)\in[\omega_1,\omega_2]$, then

$$\Big(\int_{a_1}^{a_2}\zeta_1^{p_1}(y)dy\Big)^{\frac{1}{p_1}}\Big(\int_{a_1}^{a_2}\zeta_2^{p_2}(y)dy\Big)^{\frac{1}{p_2}}-\int_{a_1}^{a_2}\zeta_1(y)\zeta_2(y)dy$$
$$\le\Bigg(\frac{p_1(p_1-1)(\omega_2^{p_1-2}-\omega_1^{p_1-2})}{6(\omega_2-\omega_1)}\Bigg(\frac{\int_{a_1}^{a_2}\zeta_1^3(y)\zeta_2^{1-2\frac{p_2}{p_1}}(y)dy}{\int_{a_1}^{a_2}\zeta_2^{p_2}(y)dy}-\Big(\frac{\int_{a_1}^{a_2}\zeta_1(y)\zeta_2(y)dy}{\int_{a_1}^{a_2}\zeta_2^{p_2}(y)dy}\Big)^3\Bigg)$$
$$+\frac{p_1(p_1-1)(\omega_2\omega_1^{p_1-2}-\omega_1\omega_2^{p_1-2})}{2(\omega_2-\omega_1)}\Bigg(\frac{\int_{a_1}^{a_2}\zeta_1^2(y)\zeta_2^{1-\frac{p_2}{p_1}}(y)dy}{\int_{a_1}^{a_2}\zeta_2^{p_2}(y)dy}-\Big(\frac{\int_{a_1}^{a_2}\zeta_1(y)\zeta_2(y)dy}{\int_{a_1}^{a_2}\zeta_2^{p_2}(y)dy}\Big)^2\Bigg)\Bigg)^{\frac{1}{p_1}}\int_{a_1}^{a_2}\zeta_2^{p_2}(y)dy. \quad (25)$$

Proof: Using (4) for $T(\varsigma)=\varsigma^{p_1}$, $f(y)=\zeta_2^{p_2}(y)$ and $h(y)=\zeta_1(y)\zeta_2^{-\frac{p_2}{p_1}}(y)$, we deduce

$$\Bigg(\Big(\int_{a_1}^{a_2}\zeta_2^{p_2}(y)dy\Big)^{p_1-1}\int_{a_1}^{a_2}\zeta_1^{p_1}(y)dy-\Big(\int_{a_1}^{a_2}\zeta_1(y)\zeta_2(y)dy\Big)^{p_1}\Bigg)^{\frac{1}{p_1}}$$
$$\le\Bigg(\frac{p_1(p_1-1)(\omega_2^{p_1-2}-\omega_1^{p_1-2})}{6(\omega_2-\omega_1)}\Bigg(\frac{\int_{a_1}^{a_2}\zeta_1^3(y)\zeta_2^{1-2\frac{p_2}{p_1}}(y)dy}{\int_{a_1}^{a_2}\zeta_2^{p_2}(y)dy}-\Big(\frac{1}{\int_{a_1}^{a_2}\zeta_2^{p_2}(y)dy}\int_{a_1}^{a_2}\zeta_1(y)\zeta_2(y)dy\Big)^3\Bigg)$$
$$+\frac{p_1(p_1-1)(\omega_2\omega_1^{p_1-2}-\omega_1\omega_2^{p_1-2}}{2(\omega_2-\omega_1)}\Bigg(\frac{\int_{a_1}^{a_2}\zeta_1^2(y)\zeta_2^{1-\frac{p_2}{p_1}}(y)dy}{\int_{a_1}^{a_2}\zeta_2^{p_2}(y)dy}-\Big(\frac{1}{\int_{a_1}^{a_2}\zeta_2^{p_2}(y)dy}\int_{a_1}^{a_2}\zeta_1(y)\zeta_2(y)dy\Big)^2\Bigg)\Bigg)^{\frac{1}{p_1}}\int_{a_1}^{a_2}\zeta_2^{p_2}(y)dy. \quad (26)$$

Utilizing the inequality $x^\beta-y^\beta\le(x-y)^\beta$, $0\le y\le x$, $\beta\in[0,1]$ for $\beta=\frac{1}{p_1}$, $x=\Big(\int_{a_1}^{a_2}\zeta_2^{p_2}(y)dy\Big)^{p_1-1}\Big(\int_{a_1}^{a_2}\zeta_1^{p_1}(y)dy\Big)$ and
$y=\Big(\int_{a_1}^{a_2}\zeta_1(y)\zeta_2(y)dy\Big)^{p_1}$, we obtain

$$\Big(\int_{a_1}^{a_2}\zeta_1^{p_1}(y)dy\Big)^{\frac{1}{p_1}}\Big(\int_{a_1}^{a_2}\zeta_2^{p_2}(y)dy\Big)^{\frac{1}{p_2}}-\int_{a_1}^{a_2}\zeta_1(y)\zeta_2(y)dy$$
$$\le\Bigg(\Big(\int_{a_1}^{a_2}\zeta_2^{p_2}(y)dy\Big)^{p_1-1}\int_{a_1}^{a_2}\zeta_1^{p_1}(y)dy-\Big(\int_{a_1}^{a_2}\zeta_1(y)\zeta_2(y)dy\Big)^{p_1}\Bigg)^{\frac{1}{p_1}}. \quad (27)$$

Now using (27) in (26), we get (25). □

In the following corollary, we demonstrate another converse of the Hölder inequality as an application of Theorem 2.

Corollary 1: Let $\zeta_1, \zeta_2 : [a_1,a_2]\to R^+$ be two functions such that $\zeta_1^{p_1}(y)$, $\zeta_2^{p_2}(y)$ and $\zeta_1(y)\zeta_2(y)$ are integrable on $[a_1,a_2]$. Also, let $[\omega_1,\omega_2]$ be a positive interval such that $\zeta_1^{p_1}(y)\zeta_2^{-p_2}(y)\in[\omega_1,\omega_2]$, then
(i): For $p_1>1$ and $p_2=\frac{p_1}{p_1-1}$, we have

$$\Big(\int_{a_1}^{a_2}\zeta_1^{p_1}(y)dy\Big)^{\frac{1}{p_1}}\Big(\int_{a_1}^{a_2}\zeta_2^{p_2}(y)dy\Big)^{\frac{1}{p_2}}-\int_{a_1}^{a_2}\zeta_1(y)\zeta_2(y)dy$$
$$\le\Bigg\{\frac{\Big(1-\frac{1}{p_1}\Big)\Big(\omega_2^{\frac{1}{p_1}-2}-\omega_1^{\frac{1}{p_1}-2}\Big)}{6p_1(\omega_2-\omega_1)}$$
$$\times\Bigg(\frac{\int_{a_1}^{a_2}\zeta_1^{3p_1}(y)\zeta_2^{-2p_2}(y)dy}{\int_{a_1}^{a_2}\zeta_2^{p_2}(y)dy}-\Big(\frac{\int_{a_1}^{a_2}\zeta_1^{p_1}(y)dy}{\int_{a_1}^{a_2}\zeta_2^{p_2}(y)dy}\Big)^3\Bigg)$$

$$+\frac{\left(1-\frac{1}{p_1}\right)\left(\omega_2\omega_1^{\frac{1}{p_1}-2}-\omega_1\omega_2^{\frac{1}{p_1}-2}\right)}{2p_1(\omega_2-\omega_1)} \times\left(\frac{\int_{a_1}^{a_2}\zeta_1^{2p_1}(y)\zeta_2^{-p_2}(y)dy}{\int_{a_1}^{a_2}\zeta_2^{p_2}(y)dy}-\left(\frac{\int_{a_1}^{a_2}\zeta_1^{p_1}(y)dy}{\int_{a_1}^{a_2}\zeta_2^{p_2}(y)dy}\right)^2\right)\Bigg\} \times\int_{a_1}^{a_2}\zeta_2^{p_2}(y)dy. \quad (28)$$

(ii): For $0<p_1<1$ and $p_2=\frac{p_1}{p_1-1}$ with $\frac{1}{p_1}\in(0,\infty)\backslash(2,3)$, we have

$$\int_{a_1}^{a_2}\zeta_1(y)\zeta_2(y)dy-\left(\int_{a_1}^{a_2}\zeta_1^{p_1}(y)dy\right)^{\frac{1}{p_1}}\left(\int_{a_1}^{a_2}\zeta_2^{p_2}(y)dy\right)^{\frac{1}{p_2}}$$
$$\leq\Bigg\{\frac{\left(\frac{1}{p_1}-1\right)\left(\omega_2^{\frac{1}{p_1}-2}-\omega_1^{\frac{1}{p_1}-2}\right)}{6p_1(\omega_2-\omega_1)} \times\left(\frac{\int_{a_1}^{a_2}\zeta_1^{3p_1}(y)\zeta_2^{-2p_2}(y)dy}{\int_{a_1}^{a_2}\zeta_2^{p_2}(y)dy}-\left(\frac{\int_{a_1}^{a_2}\zeta_1^{p_1}(y)dy}{\int_{a_1}^{a_2}\zeta_2^{p_2}(y)dy}\right)^3\right)$$
$$+\frac{\left(\frac{1}{p_1}-1\right)\left(\omega_2\omega_1^{\frac{1}{p_1}-2}-\omega_1\omega_2^{\frac{1}{p_1}-2}\right)}{2p_1(\omega_2-\omega_1)} \times\left(\frac{\int_{a_1}^{a_2}\zeta_1^{2p_1}(y)\zeta_2^{-p_2}(y)dy}{\int_{a_1}^{a_2}\zeta_2^{p_2}(y)dy}-\left(\frac{\int_{a_1}^{a_2}\zeta_1^{p_1}(y)dy}{\int_{a_1}^{a_2}\zeta_2^{p_2}(y)dy}\right)^2\right)\Bigg\} \times\int_{a_1}^{a_2}\zeta_2^{p_2}(y)dy. \quad (29)$$

(iii): For $p_1<0$ and $p_2=\frac{p_1}{p_1-1}$, we have

$$\int_{a_1}^{a_2}\zeta_1(y)\zeta_2(y)dy-\left(\int_{a_1}^{a_2}\zeta_1^{p_1}(y)dy\right)^{\frac{1}{p_1}}\left(\int_{a_1}^{a_2}\zeta_2^{p_2}(y)dy\right)^{\frac{1}{p_2}}$$
$$\leq\Bigg\{\frac{\left(\frac{1}{p_2}-1\right)\left(\omega_2^{\frac{1}{p_2}-2}-\omega_1^{\frac{1}{p_2}-2}\right)}{6p_2(\omega_2-\omega_1)} \times\left(\frac{\int_{a_1}^{a_2}\zeta_2^{3p_2}(y)\zeta_1^{-2p_1}(y)dy}{\int_{a_1}^{a_2}\zeta_1^{p_1}(y)dy}-\left(\frac{\int_{a_1}^{a_2}\zeta_2^{p_2}(y)dy}{\int_{a_1}^{a_2}\zeta_1^{p_1}(y)dy}\right)^3\right)$$
$$+\frac{\left(\frac{1}{p_2}-1\right)\left(\omega_2\omega_1^{\frac{1}{p_2}-2}-\omega_1\omega_2^{\frac{1}{p_2}-2}\right)}{2p_2(\omega_2-\omega_1)} \times\left(\frac{\int_{a_1}^{a_2}\zeta_2^{2p_2}(y)\zeta_1^{-p_1}(y)dy}{\int_{a_1}^{a_2}\zeta_1^{p_1}(y)dy}-\left(\frac{\int_{a_1}^{a_2}\zeta_2^{p_2}(y)dy}{\int_{a_1}^{a_2}\zeta_1^{p_1}(y)dy}\right)^2\right)\Bigg\} \times\int_{a_1}^{a_2}\zeta_1^{p_1}(y)dy. \quad (30)$$

Proof: (i): Let $T(\varsigma)=\varsigma^{\frac{1}{p_1}}$, then it can easily be shown that the function T is concave and $|T''|$ is convex. Therefore using (4) for $h(y)=\zeta_1^{p_1}(y)\zeta_2^{-p_2}(y)$, $f(y)=\zeta_2^{p_2}(y)$, we get (28).

(ii): In this case, the functions $T(\varsigma)=\varsigma^{\frac{1}{p_1}}$ and $|T''|$ are convex. Therefore using (4) for $h(y)=\zeta_1^{p_1}(y)\zeta_2^{-p_2}(y)$, $f(y)=\zeta_2^{p_2}(y)$, we get (29).

(iii): If $p_1<0$, we have $0<p_2<1$, which shows that this case is the reflection of case (ii). Therefore, replacement of $p_1, p_2, \zeta_1, \zeta_2$ by $p_2, p_1, \zeta_2, \zeta_1$ in (29) will lead us towards the result (30). ☐

The following corollary proposes a bound for the Hermite-Hadamard gap as an application of Theorem 2.

Corollary 2: Let $\psi\in C^2[a_1,a_2]$ be a function such that $|\psi''|$ is convex, then

$$\left|\frac{1}{a_2-a_1}\int_{a_1}^{a_2}\psi(y)dy-\psi\left(\frac{a_1+a_2}{2}\right)\right| \leq\frac{(|\psi''(a_1)|+|\psi''(a_2)|)(a_2-a_1)^2}{48}. \quad (31)$$

Proof: Using (4) for $\psi=T$, $[\omega_1,\omega_2]=[a_1,a_2]$ and $f(y)=1$, $h(y)=y$ for all $y\in[a_1,a_2]$, we get (31). ☐

Remark 2: The inequality (31) has also been proved by Sarikaya et al. in [30].

III. APPLICATIONS IN INFORMATION THEORY

Information theory is a branch of Science, which scientifically deals with the communication, quantification and storage of different kinds of information. Information is an abstract entity, therefore it cannot be quantified easily. A probability density function can be used for quantification of information about a certain event. A divergence can measure the difference between two probability densities. Csiszár [22] introduced a divergence known as Csiszár divergence, which is the base for other divergences for example Kullback-Leibler divergence, χ^2-divergence, Jeffrey's divergence etc. Divergences have many applications in various fields of Science, Technology and Art for example Pattern recognition [23], Genetics [10], Applied Statistics [19], Signal processing and Coding [31] etc. Jensen inequality plays a vital role to deduce the estimates for various divergences [16], [18], [20], [28]. In this section, we present some applications of our main result for various divergences.

Definition 1 (Csiszár divergence): Let $[\omega_1,\omega_2]\subseteq R$ and $\phi:[\omega_1,\omega_2]\to R$ be a function. Also let $X:[a_1,a_2]\to[\omega_1,\omega_2]$, $Z:[a_1,a_2]\to(0,\infty)$ be two functions such that $\frac{X(y)}{Z(y)}\in[\omega_1,\omega_2]$ for all $y\in[a_1,a_2]$, then the Csiszár divergence is defined by [22]

$$D^c(X,Z)=\int_{a_1}^{a_2}Z(y)\phi\left(\frac{X(y)}{Z(y)}\right)dy.$$

Theorem 3: Let $\phi\in C^2[\omega_1,\omega_2]$ be a function such that $|\phi''|$ is convex. Also $X:[a_1,a_2]\to[\omega_1,\omega_2]$, $Z:[a_1,a_2]\to(0,\infty)$ be two functions such that $\frac{\int_{a_1}^{a_2}X(y)dy}{\int_{a_1}^{a_2}Z(y)dy}$, $\frac{X(y)}{Z(y)}\in[\omega_1,\omega_2]$, for all $y\in[a_1,a_2]$, then

$$\left|\frac{1}{\int_{a_1}^{a_2}Z(y)dy}D^c(X,Z)-\phi\left(\frac{\int_{a_1}^{a_2}X(y)dy}{\int_{a_1}^{a_2}Z(y)dy}\right)\right| \leq\frac{|\phi''(\omega_2)|-|\phi''(\omega_1)|}{6(\omega_2-\omega_1)} \times\left(\frac{\int_{a_1}^{a_2}\frac{X^3(y)}{Z^2(y)}dy}{\int_{a_1}^{a_2}Z(y)dy}-\left(\frac{\int_{a_1}^{a_2}X(y)dy}{\int_{a_1}^{a_2}Z(y)dy}\right)^3\right)$$

$$+\frac{\omega_2|\phi''(\omega_1)|-\omega_1|\phi''(\omega_2)|}{2(\omega_2-\omega_1)} \times\left(\frac{\int_{a_1}^{a_2}\frac{X^2(y)}{Z(y)}dy}{\int_{a_1}^{a_2}Z(y)dy}-\left(\frac{\int_{a_1}^{a_2}X(y)dy}{\int_{a_1}^{a_2}Z(y)dy}\right)^2\right). \quad (32)$$

Proof: Using (4) for $T=\phi$, $h(y)=\frac{X(y)}{Z(y)}$ and $f(y)=Z(y)$, we obtain (32). □

Definition 2 (Shannon entropy): For a positive probability density function $Z(y)$ defined on $[a_1,a_2]$, the Shannon entropy is defined by [22]

$$E^s(Z)=-\int_{a_1}^{a_2}Z(y)\log Z(y)dy.$$

Corollary 3: Let $[\omega_1,\omega_2]\subseteq R^+$ and $Z(y)$ be a positive probability density function defined on $[a_1,a_2]$ such that $\frac{1}{Z(y)}\in[\omega_1,\omega_2]$ for all $y\in[a_1,a_2]$, then

$$\log(a_2-a_1)-E^s(Z) \leq \frac{\omega_1^2+\omega_1\omega_2+\omega_2^2}{2\omega_1^2\omega_2^2}\left(\int_{a_1}^{a_2}\frac{1}{Z(y)}dy-(a_2-a_1)^2\right) -\frac{\omega_1+\omega_2}{6\omega_1^2\omega_2^2}\left(\int_{a_1}^{a_2}\frac{1}{Z^2(y)}dy-(a_2-a_1)^3\right). \quad (33)$$

Proof: Using (32) for the function $\phi(\varsigma)=-\log\varsigma$, $\varsigma\in[\omega_1,\omega_2]$ and $X(y)=1$, we get (33). □

In the following example we verify the inequality (33).

Example 3: Let $Z(y)=\frac{1}{y\ln 2}$ for all $y\in[a_1,a_2]=[1,2]$ be a reciprocal distribution, then $\frac{1}{Z(y)}\in[\frac{1}{2},\frac{3}{2}]=[\omega_1,\omega_2]$ for all $y\in[1,2]$. Therefore using (33) for these facts we obtain

$$\log(a_2-a_1)-E^s(Z) =0.0199 <0.0429=\frac{\omega_1^2+\omega_1\omega_2+\omega_2^2}{2\omega_1^2\omega_2^2}\left(\int_{a_1}^{a_2}\frac{1}{Z(y)}dy-(a_2-a_1)^2\right) -\frac{\omega_1+\omega_2}{6\omega_1^2\omega_2^2}\left(\int_{a_1}^{a_2}\frac{1}{Z^2(y)}dy-(a_2-a_1)^3\right).$$

Definition 3 (Kullback-Leibler divergence): For two positive probability densities $X(y)$ and $Z(y)$ defined on $[a_1,a_2]$, the Kullback-Leibler divergence is defined by [22]

$$D^{kl}(X,Z)=\int_{a_1}^{a_2}X(y)\log\left(\frac{X(y)}{Z(y)}\right)dy.$$

Corollary 4: Let $[\omega_1,\omega_2]\subseteq R^+$ and $X:[a_1,a_2]\to[\omega_1,\omega_2]$, $Z:[a_1,a_2]\to(0,\infty)$ be two positive probability density functions such that $\frac{X(y)}{Z(y)}\in[\omega_1,\omega_2]$, for all $y\in[a_1,a_2]$, then

$$D^{kl}(X,Z)\leq\frac{\omega_1+\omega_2}{2\omega_1\omega_2}\left(\int_{a_1}^{a_2}\frac{X^2(y)}{Z(y)}dy-1\right) -\frac{\int_{a_1}^{a_2}\frac{X^3(y)}{Z^2(y)}dy-1}{6\omega_1\omega_2}. \quad (34)$$

Proof: Using (32) for the function $\phi(\varsigma)=\varsigma\log\varsigma$, we get (34). □

Definition 4 (Jeffrey's divergence): For two positive probability densities $X(y)$ and $Z(y)$ defined on $[a_1,a_2]$, the Jeffrey's divergence is defined by [22]

$$D^j(X,Z)=\int_{a_1}^{a_2}(X(y)-Z(y))\log\left(\frac{X(y)}{Z(y)}\right)dy.$$

Corollary 5: Let $[\omega_1,\omega_2]\subseteq R^+$ and $X:[a_1,a_2]\to[\omega_1,\omega_2]$, $Z:[a_1,a_2]\to(0,\infty)$ be two positive probability densities such that $\frac{X(y)}{Z(y)}\in[\omega_1,\omega_2]$, for all $y\in[a_1,a_2]$, then

$$D^j(X,Z)\leq\frac{\omega_1^2+\omega_2^2+\omega_1\omega_2+\omega_1\omega_2^2+\omega_2\omega_1^2}{2\omega_1^2\omega_2^2} \times\left(\int_{a_1}^{a_2}\frac{X^2(y)}{Z(y)}dy-1\right)-\frac{\omega_1+\omega_2+\omega_1\omega_2}{6\omega_1^2\omega_2^2} \times\left(\int_{a_1}^{a_2}\frac{X^3(y)}{Z^2(y)}dy-1\right). \quad (35)$$

Proof: Using the function $\phi(\varsigma)=(\varsigma-1)\log\varsigma$ in (32), we get (35). □

Definition 5 (Bhattacharyya coefficient): For two positive probability densities $X(y)$ and $Z(y)$ defined on $[a_1,a_2]$, the Bhattacharyya coefficient is defined by [22]

$$C^b(X,Z)=\int_{a_1}^{a_2}\sqrt{X(y)Z(y)}dy.$$

Corollary 6: Let $[\omega_1,\omega_2]\subseteq R^+$ and $X:[a_1,a_2]\to[\omega_1,\omega_2]$, $Z:[a_1,a_2]\to(0,\infty)$ be two positive probability density functions such that $\frac{X(y)}{Z(y)}\in[\omega_1,\omega_2]$, for all $y\in[a_1,a_2]$, then

$$1-C^b(X,Z)\leq\frac{\omega_1^{\frac{3}{2}}-\omega_2^{\frac{3}{2}}}{24\omega_1^{\frac{3}{2}}\omega_2^{\frac{3}{2}}(\omega_2-\omega_1)}\left(\int_{a_1}^{a_2}\frac{X^3(y)}{Z^2(y)}dy-1\right) +\frac{\omega_2^{\frac{5}{2}}-\omega_1^{\frac{5}{2}}}{8\omega_1^{\frac{3}{2}}\omega_2^{\frac{3}{2}}(\omega_2-\omega_1)}\left(\int_{a_1}^{a_2}\frac{X^2(y)}{Z(y)}dy-1\right). \quad (36)$$

Proof: Using (32) for the function $\phi(\varsigma)=-\sqrt{\varsigma}$, we get (36). □

Definition 6 (Hellinger distance): For two probability density functions $X(y)$ and $Z(y)$ defined on $[a_1,a_2]$, the Hellinger distance is defined by [22]

$$D^h(X,Z)=\frac{1}{2}\int_{a_1}^{a_2}\left(\sqrt{X(y)}-\sqrt{Z(y)}\right)^2dy.$$

Corollary 7: Let $[\omega_1,\omega_2]\subseteq R^+$ and $X:[a_1,a_2]\to[\omega_1,\omega_2]$, $Z:[a_1,a_2]\to(0,\infty)$ be two probability densities such that $\frac{X(y)}{Z(y)}\in[\omega_1,\omega_2]$, for all $y\in[a_1,a_2]$, then

$$D^h(X,Z)\leq\frac{\omega_1^{\frac{3}{2}}-\omega_2^{\frac{3}{2}}}{24\omega_1^{\frac{3}{2}}\omega_2^{\frac{3}{2}}(\omega_2-\omega_1)}\left(\int_{a_1}^{a_2}\frac{X^3(y)}{Z^2(y)}dy-1\right) +\frac{\omega_2^{\frac{5}{2}}-\omega_1^{\frac{5}{2}}}{8\omega_1^{\frac{3}{2}}\omega_2^{\frac{3}{2}}(\omega_2-\omega_1)}\left(\int_{a_1}^{a_2}\frac{X^2(y)}{Z(y)}dy-1\right). \quad (37)$$

Proof: Using (32) for the function $\phi(\varsigma)=\frac{1}{2}(1-\sqrt{\varsigma})^2$, we get (37). □

Definition 7 (Triangular discrimination): Let $X(y)$ and $Z(y)$ be two positive probability density functions defined on

$[a_1, a_2]$, *then the Triangular discrimination is defined by [22]*

$$D^{\triangle}(X, Z) = \int_{a_1}^{a_2} \frac{(X(y) - Z(y))^2}{X(y) + Z(y)} dy.$$

Corollary 8: Let $[\omega_1, \omega_2] \subseteq R^+$ *and* $X : [a_1, a_2] \to [\omega_1, \omega_2]$, $Z : [a_1, a_2] \to (0, \infty)$ *be two positive probability density functions such that* $\frac{X(y)}{Z(y)} \in [\omega_1, \omega_2]$, *for all* $y \in [a_1, a_2]$, *then*

$$\begin{aligned} D^{\triangle}(X, Z) &\leq \frac{4((\omega_1+1)^3 - (\omega_2+1)^3)}{3(\omega_1+1)^3(\omega_2+1)^3(\omega_2-\omega_1)} \Big(\int_{a_1}^{a_2} \frac{X^3(y)}{Z^2(y)} dy - 1 \Big) \\ &\quad + \frac{4(\omega_2(\omega_2+1)^3 - \omega_1(\omega_1+1)^3)}{(\omega_1+1)^3(\omega_2+1)^3(\omega_2-\omega_1)} \Big(\int_{a_1}^{a_2} \frac{X^2(y)}{Z(y)} dy - 1 \Big). \end{aligned} \tag{38}$$

Proof: Using (32) for the function $\phi(\varsigma) = \frac{(\varsigma-1)^2}{\varsigma+1}$, we obtain (38). □

Remark 3: It is important to note that we can give the discrete version of the results presented in this manuscript.

IV. CONCLUDING REMARKS

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 pretty good in their applications. The Jensen inequality has generalized the concept of classical convexity. This inequality and the results around its gap, resolve some difficulties in the modeling of some physical phenomena. Thus, we have derived a new bound for the integral version of the Jensen gap involving functions whose absolute value of second derivative are convex. Based on this bound, we have deduced converses of the Hölder inequality as well. Also, a bound for the Hermite-Hadamard gap has been obtained. Finally, we have demonstrated some bounds for Csiszár, Jeffrey's and Kullback-Leibler divergences etc in information theory as applications of the main result. The numerical experiments, which are demonstrated in Section II not only confirm the sharpness of the Jensen inequality but also give the surety of the tightness of the bound in (4) towards the Jensen gap. An application of the main result for Shannon entropy has also been discussed through a numerical example, which verifies the bound of Shannon entropy. Also, it is important to note that the bounds around various divergences can be applied for signal processing, magnetic resonance image analysis, pattern recognition and image segmentation etc. The proposed idea may inculcate further research in the area of mathematical inequalities.

ACKNOWLEDGMENT

The authors would like to express their sincere thanks to anonymous reviewers for their valuable suggestions and comments which helped the authors to improve this article substantially.

REFERENCES

[1] M. Adil Khan, T. Ali, A. Kiliçman, and Q. Din, "Refinements of Jensen's inequality for convex functions on the co-ordinates in a rectangle from the plane," *Filomat*, vol. 30, no. 3, pp. 803–814, 2016.
[2] M. Adil Khan, G. A. Khan, T. Ali, and A. Kilicman, "On the refinement of Jensen's inequality," *Appl. Math. Comput.*, vol. 262, pp. 128–135, Jul. 2015.
[3] M. Adil Khan, Z. Al-sahwi, and Y.-M. Chu, "New estimations for Shannon and Zipf–Mandelbrot entropies," *Entropy*, vol. 20, no. 8, p. 608, Aug. 2018.
[4] M. Adil Khan, M. Hanif, Z. Abdul Hameed Khan, K. Ahmad, and Y.-M. Chu, "Association of Jensen's inequality for s-convex function with Csiszár divergence," *J. Inequalities Appl.*, vol. 2019, no. 1, Dec. 2019, Art. no. 162.
[5] M. Adil Khan, J. Khan, and J. Pečarić, "Generalization of Jensen's and Jensen-Steffensen's inequalities by generalized majorization theorem," *J. Math. Inequalities*, vol. 11, no. 4, pp. 1049–1074, 2017.
[6] M. Adil Khan, Đ. Pečarić, and J. Pečarić, "Bounds for Csiszár divergence and hybrid Zipf–Mandelbrot entropy," *Math. Methods Appl. Sci.*, vol. 42, no. 18, pp. 7411–7424, Dec. 2019.
[7] M. Adil Khan, Đ. Pečarić, and J. Pečarić, "Bounds for Shannon and Zipf–Mandelbrot entropies," *Math. Methods Appl. Sci.*, vol. 40, no. 18, pp. 7316–7322, Dec. 2017.
[8] M. Adil Khan, Đ. Pečarić, and J. Pečarić, "On Zipf–Mandelbrot entropy," *J. Comput. Appl. Math.*, vol. 346, pp. 192–204, Jan. 2019.
[9] M. Adil Khan, S.-H. Wu, H. Ullah, and Y.-M. Chu, "Discrete majorization type inequalities for convex functions on rectangles," *J. Inequalities Appl.*, vol. 2019, no. 1, Dec. 2019, Art. no. 16.
[10] J. Burbea and C. Rao, "On the convexity of some divergence measures based on entropy functions," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 3, pp. 489–495, May 1982.
[11] Y. Chu, C. Zong, and G. Wang, "Optimal convex combination bounds of Seiffert and geometric means for the arithmetic mean," *J. Math. Inequalities*, no. 3, pp. 429–434, 2011.
[12] M. J. Cloud, B. C. Drachman, and L. P. Lebedev, *Inequalities: With Applications to Engineering*. Cham, Switzerland: Springer, 2014.
[13] D. Costarelli and R. Spigler, "How sharp is the Jensen inequality?" *J. Inequalities Appl.*, vol. 2015, no. 1, Dec. 2015, Art. no. 69.
[14] L. Dedić, C. E. M. Pearce, and J. Pečarić, "The euler formulæ and convex functions," *Math. Inequalities Appl.*, vol. 3, no. 2, pp. 211–221, 2000.
[15] S. S. Dragomir, "Inequalities in terms of the Gâteaux derivatives for convex functions on linear spaces with applications," *Bull. Austral. Math. Soc.*, vol. 83, no. 3, pp. 500–517, Jun. 2011.
[16] S. Sever Dragomir, M. Adil Khan, and A. Abathun, "Refinement of the Jensen integral inequality," *Open Math.*, vol. 14, no. 1, pp. 221–228, Jan. 2016.
[17] S. S. Dragomir, C. E. M. Pearce, and J. Pečarić, "Interpolations of Jensen's inequality," *Tamkang J. Math.*, vol. 34, no. 2, pp. 175–187, 2003.
[18] L. Horváth, Đ. Pečarić, and J. Pečarić, "Estimations of f- and Rényi divergences by using a cyclic refinement of the Jensen's inequality," *Bull. Malaysian Math. Sci. Soc.*, vol. 42, no. 3, pp. 933–946, May 2019.
[19] J. H. Justice, Ed., *Maximum Entropy and Bayesian Methods in Applied Statistics*. Cambridge, U.K.: Cambridge Univ. Press, 1986.
[20] S. Khan, M. Adil Khan, and Y. Chu, "Converses of the Jensen inequality derived from the green functions with applications in information theory," *Math. Methods Appl. Sci.*, vol. 43, no. 5, pp. 2577–2587, Mar. 2020.
[21] V. Lakshmikantham and A. S. Vatsala, "Theory of differential and integral inequalities with initial time difference and applications," in *Analytic and Geometric Inequalities and Applications*. Berlin, Germany: Springer, 1999.
[22] N. Latif, Đ. Pečarić, and J. Pečarić, "Majorization, 'useful' Csiszár divergence and 'useful' Zipf–Mandelbrot law," *Open Math.*, vol. 16, no. 1, pp. 1357–1373, 2018.
[23] R. Leahy and C. Goutis, "An optimal technique for constraint-based image restoration and reconstruction," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-34, no. 6, pp. 1629–1642, Dec. 1986.
[24] J. G. Liao and A. Berg, "Sharpening Jensen's inequality," *Amer. Statistician*, vol. 73, no. 3, pp. 278–281, Jul. 2019.
[25] Q. Lin, "Jensen inequality for superlinear expectations," *Statist. Probab. Lett.*, vol. 151, pp. 79–83, Aug. 2019.
[26] H. R. Moradi, M. E. Omidvar, M. Adil Khan, and K. Nikodem, "Around Jensen's inequality for strongly convex functions," *Aequat. Math.*, vol. 92, no. 1, pp. 25–37, 2018.

[27] C. P. Niculescu and L.-E. Persson, *Convex Functions and their Applications*. New York, NY, USA: Springer, 2006.

[28] Đ. Pečarić, J. Pečarić, and M. Rodić, "About the sharpness of the Jensen inequality," *J. Inequalities Appl.*, vol. 2018, no. 1, Dec. 2018, Art. no. 337.

[29] J. Pečarić, F. Proschan, and Y. L. Tong, *Convex Functions, Partial Orderings, and Statistical Applications*. New York, NY, USA: Academic, 1992.

[30] M. Z. Sarikaya, A. Saglam, and H. Yildirim, "New inequalities of Hermite-Hadamard type for functions whose second derivatives absolute values are convex and quasi-convex," *Int. J. Open Problems Comput. Sci. Math.*, vol. 5, no. 3, pp. 1–14, 2012.

[31] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, Jul./Oct. 1948.

[32] W.-F. Xia and Y. Chu, "Optimal inequalities for the convex combination of error function," *J. Math. Inequalities*, vol. 9, no. 1, pp. 85–99, 2015.

[33] S. Zaheer Ullah, M. Adil Khan, and Y.-M. Chu, "Majorization theorems for strongly convex functions," *J. Inequalities Appl.*, vol. 2019, no. 1, Dec. 2019, Art. no. 58.

MUHAMMAD ADIL KHAN received the Ph.D. degree in mathematics from the Abdus Salam School of Mathematical Sciences, GC University Lahore, Pakistan, in March 2012. He is currently working as an Assistant Professor at the Department of Mathematics, University of Peshawar, Pakistan. He has published several research articles in well-reputed journals. He also published one book on majorization. He supervised three Ph.D. and 20 M.Phil. students while he is currently supervising five Ph.D. students. His primary research interests include mathematical inequalities, operator theory, theory of convexity, and information theory. He is actively involved in research. He has participated in several national and international conferences and presented his research work in different countries such as Turkey, Bosnia and Herzegovina, Mongolia, Philippines, and Croatia.

SHAHID KHAN received the M.Phil. degree from the Department of Mathematics, University of Engineering and Technology, Peshawar, Pakistan, in 2017. He is currently pursuing the Ph.D. degree with the Department of Mathematics, University of Peshawar, Pakistan. He is also serving as a Lecturer in mathematics at the KPK Higher Education Department. His research interests include mathematical inequalities, theory of convex functions, information theory, and fractional calculus.

YUMING CHU was born in Huzhou, Zhejiang, China, in June 3, 1966. He received the B.S. degree from Hangzhou Normal University, Hangzhou, China, in 1988, and the M.S. and Ph.D. degrees from Hunan University, Changsha, China, in 1991 and 1994, respectively. He worked as an Assistant Professor, from 1994 to 1996, and an Associate Professor, from 1997 to 2002, with the Department of Mathematics, Hunan Normal University, Changsha. Since 2002, he has been a Professor and the Dean at the Department of Mathematics, Huzhou University, Huzhou, China. His current research interests include special functions, functional analysis, numerical analysis, operator theory, ordinary differential equations, partial differential equations, inequalities theory and applications, and robust filtering and control.