RESEARCH　　　　　　　　　　　　　　　　　　　　　　　　**Open Access**

Approximations of Jensen divergence for twice differentiable functions

Eder Kikianty[1*], Sever S Dragomir[1,2], Isia T Dintoe[1] and David Sherwell[1]

*Correspondence:
eder.kikianty@wits.ac.za
[1] School of Computational and
Applied Mathematics, University of
the Witwatersrand, Private Bag 3,
Wits 2050, Johannesburg, South
Africa
Full list of author information is
available at the end of the article

Abstract

The Jensen divergence is used to measure the difference between two probability distributions. This divergence has been generalised to allow the comparison of more than two distributions. In this paper, we consider some bounds for generalised Jensen divergence of twice differentiable functions with bounded second derivatives. Evidently, these bounds provide approximations for the Jensen divergence of twice differentiable functions by the Jensen divergence of simpler functions such as the power functions and the paired entropies associated to the Harvda-Charvát functions.
MSC: 26D15; 94A17

Keywords: divergence measure; Jensen divergence; inequality for real numbers

1 Introduction

One of the more important applications of probability theory is finding an appropriate measure of distance (or difference) between two probability distributions [1]. A number of these divergence measures have been widely studied and applied by a number of mathematicians such as Burbea and Rao [2], Havrda and Charvát [3], Lin [4] and others.

In Burbea and Rao [2], a generalisation of Jensen divergence is considered to allow the comparison of more than two distributions. If Φ is a function defined on an interval I of the real line $\mathbb{R}$, the (*generalised*) *Jensen divergence* between two elements $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ in I^n (where $n \geq 1$) is given by the following equation (*cf.* Burbea and Rao [2]):

$$\mathcal{J}_{n,\Phi}(x, y) := \sum_{i=1}^{n} \left[\frac{1}{2} \big[\Phi(x_i) + \Phi(y_i) \big] - \Phi \left(\frac{x_i + y_i}{2} \right) \right] \tag{1}$$

for all $x, y \in I^n \times I^n$. Several measures have been proposed to quantify the difference (also known as the *divergence*) between two (or more) probability distributions. We refer to Grosse *et al.* [5], Kullback and Leibler [6] and Csiszar [7] for further references.

We denote by S_n

$$S_n = \left\{ (x_1, \ldots, x_n) \in I^n, \sum_{i=1}^{n} x_i = 1 \right\}, \quad I = [0, 1].$$

Utilising the family of functions, for $\alpha \in \mathbb{R}_+$,

$$\Phi_\alpha(t) := \begin{cases} (\alpha - 1)^{-1}(t^\alpha - t), & \alpha \neq 1, \\ t \log t, & \alpha = 1, \end{cases}$$

by Havrda and Charvát in [3] to introduce their entropies of degree α, Burbea and Rao [2] introduced the following family of Jensen divergences:

$$\mathcal{J}_{n,\alpha} := \begin{cases} (\alpha - 1)^{-1} \sum_{i=1}^n [\frac{1}{2}(x_i^\alpha + y_i^\alpha) - (\frac{x_i + y_i}{2})^\alpha], & \alpha \neq 1, \\ \frac{1}{2} \sum_{i=1}^n [x_i \log x_i + y_i \log y_i - (x_i + y_i) \log(\frac{x_i + y_i}{2})], & \alpha = 1, \end{cases}$$

that can be defined on $S_n \times S_n$ with the convention that $0 \log 0 = 0$ for $\alpha \in \mathbb{R}_+$. We note that the divergence $\mathcal{J}_{n,1}$ is also known as the *Jensen-Shannon divergence* [8].

These measures have been applied in a variety of fields, for example, in information theory [9]. The *Jensen divergence* introduced in Burbea and Rao [2] has its applications in bioinformatics [10, 11], where it is usually utilised to compare two samples of healthy population (control) and diseased population (case) in detecting gene expression for a certain disease. We refer the readers to Dragomir [1] for the applications in other areas.

In a recent paper by Dragomir *et al.* [12], the authors found sharp upper and lower bounds for the Jensen divergence for various classes of functions Φ, including functions of bounded variation, absolutely continuous functions, Lipschitzian continuous functions, convex functions and differentiable functions. We recall some of these results in Section 2, which motivates the new results we obtain in this paper.

In this paper, we provide bounds for Jensen divergence of twice differentiable function Φ whose second derivative Φ'' satisfies some boundedness conditions. These bounds provide approximations of the Jensen divergence $\mathcal{J}_{n,\Phi}$ (*cf.* (1)) by the divergence of simpler functions such as the power functions (*cf.* Section 3) and the above mentioned family of Jensen divergences $\mathcal{J}_{n,\alpha}$ (*cf.* Section 4). Finally, we apply these bounds to some elementary functions in Section 5.

2 Definitions, notation and previous results

In this section, we provide definitions and notation that will be used in the paper. We also provide some results regarding sharp bounds for the generalised Jensen divergence as stated in Dragomir *et al.* [12].

2.1 Definitions and notation

Throughout the paper, for any real number $r > 1$, we define r' to be its Hölder conjugate, that is, $1/r + 1/r' = 1$.

Definition 1 (Bullen [13]) If s is an extended real number, the generalised logarithmic mean of order s of two positive numbers x and y is defined by

$$\mathfrak{L}^{[s]}(x,y) = \begin{cases} \left[\frac{1}{s+1}\left(\frac{y^{s+1}-x^{s+1}}{y-x}\right)\right]^{\frac{1}{s}} & \text{if } s \neq -1, 0, \pm\infty, \\ \frac{y-x}{\log y - \log x} & \text{if } s = -1, \\ \frac{1}{e}\left(\frac{y^y}{x^x}\right)^{\frac{1}{y-x}} & \text{if } s = 0, \\ \max\{x,y\} & \text{if } s = +\infty, \\ \min\{x,y\} & \text{if } s = -\infty, \end{cases} \tag{2}$$

and $\mathfrak{L}^{[s]}(x,x) = x$.

This mean is homogeneous and symmetric [13, p.385]. In particular, there is no loss in generality by assuming $0 < x < y$. Note also that

$$\mathfrak{L}^{[s]}(x,y) = \left(\int_x^y \left[(1-t)x + ty\right]^s dt\right)^{1/s}$$

for $0 < x < y$ and $s \in [1, \infty)$. This mean generalises not only logarithmic mean (when $s = -1$), which is particularly useful in distribution of electrical charge of a conductor, but also arithmetic mean (when $s = 1$) and geometric mean (when $s = -2$).

We use the following notations for Lebesgue integrable functions: for any Lebesgue integrable function g on $[a,b]$, we define, for $a \leq x \leq y \leq b$,

$$\|g\|_{[x,y],p} := \left|\int_x^y |g(s)|^p \, ds\right|^{1/p} \quad \text{if } p \geq 1 \text{ and } g \in L_p[a,b];$$

and for $g \in L_\infty[a,b]$, we denote $\|g\|_{[x,y],\infty} := \operatorname{ess\,sup}_{s \in [x,y]} |g(s)|$.

We recall that a function $f : [a,b] \to \mathbb{R}$ is absolutely continuous on $[a,b]$ if and only if it is differentiable almost everywhere in $[a,b]$, the derivative f' is Lebesgue integrable on this interval and $f(y) - f(x) = \int_x^y f'(t) \, dt$ for any $x, y \in [a,b]$.

2.2 Previous results

In a recent paper by Dragomir *et al.* [12], the authors provide sharp upper and lower bounds for the Jensen divergence for various classes of functions Φ. Some results are stated in the following.

Theorem 2 (Dragomir *et al.* [12]) *Assume that $\Phi : [a,b] \to \mathbb{R}$ is absolutely continuous on $[a,b]$. Then we have the bounds*

$$\left|\mathcal{J}_{n,\Phi}(x,y)\right| \leq \frac{1}{2} \times \begin{cases} \sum_{i=1}^n |y_i - x_i| \|\Phi'\|_{[x_i,y_i],\infty} & \text{if } \Phi' \in L_\infty[a,b], \\ \sum_{i=1}^n |y_i - x_i|^{\frac{p-1}{p}} \|\Phi'\|_{[x_i,y_i],p} & \text{if } \Phi' \in L_p[a,b], p > 1, \\ \sum_{i=1}^n \|\Phi'\|_{[x_i,y_i],1} \end{cases}$$

$$\leq \frac{1}{2} \times \begin{cases} \|\Phi'\|_{[a,b],\infty} \sum_{i=1}^n |y_i - x_i| & \text{if } \Phi' \in L_\infty[a,b], \\ \|\Phi'\|_{[a,b],p} \sum_{i=1}^n |y_i - x_i|^{\frac{p-1}{p}} & \text{if } \Phi' \in L_p[a,b], p > 1, \\ n\|\Phi'\|_{[a,b],1} \end{cases} \tag{3}$$

for any $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in [a,b]^n$.

Moreover, if the modulus of the derivative is convex, then we have the inequality

$$\left| \mathcal{J}_{n,\Phi}(x,y) \right| \leq \frac{1}{4} \sum_{i=1}^{n} |y_i - x_i| \left[\left| \Phi'\left(\frac{x_i + y_i}{2} \right) \right| + \frac{|\Phi'(x_i)| + |\Phi'(y_i)|}{2} \right]$$

$$\leq \frac{1}{4} \sum_{i=1}^{n} |y_i - x_i| \left[\left| \Phi'(x_i) \right| + \left| \Phi'(y_i) \right| \right]$$

$$\left(\leq \left\| \Phi' \right\|_{[a,b],\infty} \delta(x,y) \right) \tag{4}$$

for any $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in [a,b]^n$, *where* $\delta(x,y) = \frac{1}{2} \sum_{i=1}^{n} |y_i - x_i|$.
The constant $1/4$ *is best possible in both inequalities.*

Some more assumptions for Φ lead to the following results.

Theorem 3 (Dragomir *et al.* [12]) *Let* $\Phi : [a,b] \to \mathbb{R}$ *be a differentiable function on the interval* $[a,b]$ *of real numbers* $\mathbb{R}$.
  (i) *If the derivative* Φ' *is of bounded variation on* $[a,b]$, *then*

$$\left| \mathcal{J}_{n,\Phi}(x,y) \right| \leq \frac{1}{4} \sum_{i=1}^{n} |y_i - x_i| \left| \bigvee_{x_i}^{y_i} (\Phi') \right|$$

$$\leq \frac{1}{4} \bigvee_{a}^{b} (\Phi') \sum_{i=1}^{n} |y_i - x_i|$$

$$= \frac{1}{2} \bigvee_{a}^{b} (\Phi') \delta(x,y) \tag{5}$$

for any $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in [a,b]^n$.
  The constant $1/4$ *is best possible in both inequalities* (5).
  (ii) *If the derivative* Φ' *is* K-*Lipschitzian on* $[a,b]$ *with the constant* $K > 0$, *then*

$$\left| \mathcal{J}_{n,\Phi}(x,y) \right| \leq \frac{1}{8} K \sum_{i=1}^{n} (y_i - x_i)^2 = \frac{1}{2} K \mathcal{J}_{n,2}(x,y) \tag{6}$$

for any $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in [a,b]^n$, *where*

$$\mathcal{J}_{n,2}(x,y) = \frac{1}{4} \sum_{i=1}^{n} (y_i - x_i)^2.$$

The constant $1/8$ *is best possible in* (6).

Motivated by these results, we state bounds for $\mathcal{J}_{n,\Phi}$ for twice differentiable functions Φ with some boundedness conditions for the second derivative in the next sections.

3 Approximating with Jensen divergence for power functions
In this section we provide some bounds for the generalised Jensen divergence for twice differentiable function $\Phi : I \subset \mathbb{R} \to \mathbb{R}$, whose second derivative Φ'' is bounded above and

below in the following sense:

$$\gamma \leq \frac{t^{2-p}}{p(p-1)} \Phi''(t) \leq \Gamma \tag{7}$$

for some $\gamma < \Gamma$ and $p \in (-\infty, 0) \cup (1, \infty)$ and all $t \in I$; and

$$\delta \leq \frac{t^{2-q}}{q(q-1)} \Phi''(t) \leq \Delta \tag{8}$$

for some $\delta \leq \Delta$, some $q \in (0, 1)$ and all $t \in I$. These conditions enable us to provide approximations of the Jensen divergence for Φ via the functions $f(t) = t^p$ for $p \neq 0, 1$ and $t \in \mathbb{R}_+$, i.e.

$$\mathcal{J}_{n,(\cdot)^p} = \sum_{i=1}^{n} \left[\frac{1}{2}(x_i^p + y_i^p) - \left(\frac{x_i + y_i}{2} \right)^p \right].$$

Lemma 4 (Dragomir *et al.* [12]) *Let* $\Phi : [a, b] \to \mathbb{R}$ *be a differentiable function and let the derivative* Φ' *be absolutely continuous. Then*

$$\left| \mathcal{J}_{n,\Phi}(x, y) \right|$$

$$\leq \begin{cases} \frac{1}{8} \| \Phi'' \|_{[a,b],\infty} \sum_{i=1}^{n} (y_i - x_i)^2 & \text{if } \Phi'' \in L_\infty[a, b], \\ \frac{\| \Phi'' \|_{[a,b],r}}{(r'+1)^{1/r'} 2^{1+1/r'}} \sum_{i=1}^{n} |y_i - x_i|^{1+1/r'} & \text{if } \Phi'' \in L_r[a, b], r > 1. \end{cases} \tag{9}$$

We refer to [12] for the proof of the above lemma.

Lemma 5 *Let* $\Phi : [a, b] \to \mathbb{R}$ *be a twice differentiable function and* $0 < a < b < \infty$. *If* Φ'' *satisfies* (7), *then*

$$\left\| \left(\Phi - \frac{\gamma + \Gamma}{2}(\cdot)^p \right)'' \right\|_{[a,b],\infty} \leq p(p-1) \frac{\Gamma - \gamma}{2} \max\{a^{p-2}, b^{p-2}\}; \tag{10}$$

and

$$\left\| \left(\Phi - \frac{\gamma + \Gamma}{2}(\cdot)^p \right)'' \right\|_{[a,b],r} \leq p(p-1) \frac{\Gamma - \gamma}{2} \left(\mathfrak{L}^{[(p-2)r]}(a, b) \right)^{p-2}, \quad r > 1, \tag{11}$$

where $\mathfrak{L}^{[s]}$ *is the sth generalised logarithmic mean.*

Proof Note that condition (7) is equivalent to

$$\gamma p(p-1) t^{p-2} \leq \Phi''(t) \leq \Gamma p(p-1) t^{p-2}$$

since $p(p-1) > 0$. This is also equivalent to

$$\left| \Phi''(t) - p(p-1) \frac{\gamma + \Gamma}{2} t^{p-2} \right| \leq p(p-1) \frac{\Gamma - \gamma}{2} t^{p-2}. \tag{12}$$

We take the supremum of both sides to obtain (10). For $r > 1$, we note that (12) is equivalent to

$$
\left\| \Phi''(t) - p(p-1)\frac{\gamma+\Gamma}{2}t^{p-2} \right\|_{[a,b],r} = \left(\int_a^b \left| \Phi''(t) - p(p-1)\frac{\gamma+\Gamma}{2}t^{p-2} \right|^r dt \right)^{1/r}
$$
$$
\leq p(p-1)\frac{\Gamma-\gamma}{2} \left(\int_a^b t^{r(p-2)} dt \right)^{1/r}
$$
$$
= p(p-1)\frac{\Gamma-\gamma}{2} \left(\mathfrak{L}^{[r(p-2)]}(a,b) \right)^{p-2},
$$

which proves (11). □

Theorem 6 *Let $\Phi : [a,b] \to \mathbb{R}$ be a twice differentiable function and $0 < a < b < \infty$. If Φ'' satisfies* (7), *then*

$$
\left| \mathcal{J}_{n,\Phi}(x,y) - \frac{\gamma+\Gamma}{2}\mathcal{J}_{n,(\cdot)^p}(x,y) \right|
$$
$$
\leq \begin{cases} \frac{1}{16}p(p-1)(\Gamma-\gamma)\max\{a^{p-2},b^{p-2}\}\sum_{i=1}^n(y_i-x_i)^2 & \text{if } \Phi'' \in L_\infty[a,b], \\ \frac{p(p-1)(\Gamma-\gamma)}{(r'+1)^{1/r'}2^{2+1/r'}}\mathfrak{L}^{[(p-2)r]}(a,b)\sum_{i=1}^n|y_i-x_i|^{1+1/r'} & \text{if } \Phi'' \in L_r[a,b], r > 1. \end{cases}
$$

Proof Since any differentiable function is absolutely continuous, we may employ Lemma 4. Combining this with Lemma 5, we have

$$
\left| \mathcal{J}_{n,\Phi}(x,y) - \frac{\gamma+\Gamma}{2}\mathcal{J}_{n,(\cdot)^p}(x,y) \right|
$$
$$
\leq \begin{cases} \frac{1}{8}\left\| (\Phi - \frac{\gamma+\Gamma}{2}(\cdot)^p)'' \right\|_{[a,b],\infty}\sum_{i=1}^n(y_i-x_i)^2, \\ \frac{1}{(r'+1)^{1/r'}2^{1+1/r'}}\left\| (\Phi - \frac{\gamma+\Gamma}{2}(\cdot)^p)'' \right\|_{[a,b],r}\sum_{i=1}^n|y_i-x_i|^{1+1/r'}, \end{cases}
$$
$$
\leq \begin{cases} \frac{1}{8}p(p-1)\frac{\Gamma-\gamma}{2}\max\{a^{p-2},b^{p-2}\}\sum_{i=1}^n(y_i-x_i)^2, \\ \frac{1}{(r'+1)^{1/r'}2^{1+1/r'}}p(p-1)\frac{\Gamma-\gamma}{2}\left(\mathfrak{L}^{[(p-2)r]}(a,b) \right)^{p-2}\sum_{i=1}^n|y_i-x_i|^{1+1/r'}, \end{cases}
$$

as desired. □

We omit the proofs for the next results as they follow similarly to those of Lemma 5 and Theorem 6.

Lemma 7 *Let $\Phi : [a,b] \to \mathbb{R}$ be a twice differentiable function and $0 < a < b < \infty$. If Φ'' satisfies* (8), *then*

$$
\left\| \left(\Phi - \frac{\delta+\Delta}{2}(\cdot)^q \right)'' \right\|_{[a,b],\infty} \leq q(1-q)\frac{\Delta-\delta}{2}\max\{a^{q-2},b^{q-2}\}; \tag{13}
$$

and

$$
\left\| \left(\Phi - \frac{\delta+\Delta}{2}(\cdot)^q \right)'' \right\|_{[a,b],q} \leq q(1-q)\frac{\Delta-\delta}{2}\left(\mathfrak{L}^{[r(q-2)]}(a,b) \right)^{p-2}, \quad r > 1, \tag{14}
$$

where $\mathfrak{L}^{[s]}$ is the sth generalised logarithmic mean.

Theorem 8 *Let* $\Phi : [a,b] \to \mathbb{R}$ *be a twice differentiable function and* $0 < a < b < \infty$. *If* Φ'' *satisfies* (8), *then*

$$\left| \mathcal{J}_{n,\Phi}(x,y) - \frac{\delta + \Delta}{2} \mathcal{J}_{n,(\cdot)^q}(x,y) \right|$$

$$\leq \begin{cases} \frac{1}{16} q(1-q)(\Delta - \delta) \max\{a^{q-2}, b^{q-2}\} \sum_{i=1}^{n} (y_i - x_i)^2 & \text{if } \Phi'' \in L_\infty[a,b], \\ \frac{q(1-q)(\Delta - \delta)}{(r'+1)^{1/r'} 2^{2+1/r'}} \left(\mathfrak{L}^{[(q-2)r']}(a,b) \right)^{p-2} \sum_{i=1}^{n} |y_i - x_i|^{1+1/r'} & \text{if } \Phi'' \in L_r[a,b], r > 1. \end{cases}$$

4 Further approximations

In this section, we present approximations for $\mathcal{J}_{n,\Phi}$ by utilising the family of the Jensen divergence

$$\mathcal{J}_{n,\alpha} := (\alpha - 1)^{-1} \sum_{i=1}^{n} \left[\frac{1}{2} (x_i^\alpha + y_i^\alpha) - \left(\frac{x_i + y_i}{2} \right)^\alpha \right], \quad \alpha \neq 1; \tag{15}$$

and

$$\mathcal{J}_{n,1}(x,y) := \frac{1}{2} \sum_{i=1}^{n} \left[x_i \log x_i + y_i \log y_i - (x_i + y_i) \log \left(\frac{x_i + y_i}{2} \right) \right]. \tag{16}$$

Although $\mathcal{J}_{n,\alpha}$ is defined for $\alpha \in \mathbb{R}_+$ in [2], we may let α to be negative in (15), and for $\alpha = 0$, we define

$$\mathcal{J}_{n,0}(x,y) := \sum_{i=1}^{n} \left[\log \left(\frac{x_i + y_i}{2} \right) - \frac{1}{2} (\log x_i + \log y_i) \right]. \tag{17}$$

Theorem 9 *Let* $\Phi : I \subset (0, \infty) \to \mathbb{R}$ *be a twice differentiable function on* I. *If* Φ'' *satisfies* (7), *then*

$$\gamma(p-1)\mathcal{J}_{n,p}(x,y) \leq \mathcal{J}_{n,\Phi}(x,y) \leq \Gamma(p-1)\mathcal{J}_{n,p}(x,y) \quad \text{for any } x,y \in I^n. \tag{18}$$

Furthermore, if Φ'' *satisfies* (8), *then*

$$\delta(q-1)\mathcal{J}_{n,q}(x,y) \geq \mathcal{J}_{n,\Phi}(x,y) \geq \Delta(q-1)\mathcal{J}_{n,q}(x,y) \quad \text{for any } x,y \in I^n. \tag{19}$$

Proof We consider the auxiliary function $g_{\gamma,p} : I \to \mathbb{R}$ defined by $g_{\gamma,p}(t) = \Phi(t) - \gamma t^p$, where $p \in (-\infty, 0) \cup (1, \infty)$. We observe that $g_{\gamma,p}$ is twice differentiable on I and the second derivative is given by

$$g_{\gamma,p}''(t) = p(p-1)t^{p-2} \left[\frac{t^{2-p}}{p(p-1)} \Phi''(t) - \gamma \right] \quad \text{for any } t \in I.$$

Utilising condition (7) and since $p(p-1)t^{p-2} > 0$ for $t \in I$, we deduce that $g_{\gamma,p}''(t) \geq 0$ for any $t \in I$ which means that $g_{\gamma,p}$ is convex on I. Since for a convex function $g : I \to \mathbb{R}$ we

have that $\mathcal{J}_{n,g}(x,y) \geq 0$, then we can write that

$$
\begin{aligned}
0 \leq \mathcal{J}_{n,g_{\gamma,p}}(x,y) &= \sum_{i=1}^{n} \left[\frac{g_{\gamma,p}(x_i) + g_{\gamma,p}(y_i)}{2} - g_{\gamma,p}\left(\frac{x_i + y_i}{2} \right) \right] \\
&= \sum_{i=1}^{n} \left[\frac{\Phi(x_i) + \Phi(y_i)}{2} - \Phi\left(\frac{x_i + y_i}{2} \right) \right] - \gamma \sum_{i=1}^{n} \left[\frac{x_i^p + y_i^p}{2} - \left(\frac{x_i + y_i}{2} \right)^p \right] \\
&= \mathcal{J}_{n,\Phi}(x,y) - \gamma(p-1)\mathcal{J}_{n,p}(x,y),
\end{aligned}
$$

and the first inequality in (18) is proved. To prove the second inequality in (18), we consider the auxiliary function $g_{\Gamma,p}(x,y) : I \to \mathbb{R}$ with $g_{\Gamma,p}(x,y) = \Gamma t^p - \Phi(t)$, for which we perform a similar argument; and we omit the details.

Now, if $q \in (0,1)$ and if we consider the auxiliary function $\psi_{\delta,q}(x,y) : I \to \mathbb{R}$ with $\psi_{\delta,q}(x,y) = \Phi(t) - \delta t^q$, then ψ is twice differentiable and

$$
\psi_{\delta,q}''(x,y) = t^{q-2} q(q-1) \left[\frac{t^{2-q}\Phi''(t)}{q(q-1)} - \delta \right] \leq 0 \quad \text{for any } t \in I
$$

since $q \in (0,1)$. Therefore $\psi_{\delta,q}$ is concave on I, which implies that $\mathcal{J}_{n,\psi_{\delta,q}}(x,y) \leq 0$ for any $x,y \in I^n$ and, as above, we obtain

$$
\mathcal{J}_{n,\Phi}(x,y) \leq \delta \sum_{i=0}^{n} \left[\frac{x_i^q + y_i^q}{2} - \left(\frac{x_i + y_i}{2} \right)^q \right] = \delta(q-1)\mathcal{J}_{n,q}(x,y).
$$

The second inequality in (19) follows by considering the auxiliary function $\psi_{\Delta,q}(x,y) : I \to \mathbb{R}$ with $\psi_{\Delta,q}(x,y) = \Delta t^q - \Phi''(t)$, and we omit the details. This completes the proof. $\square$

Theorem 10 *Let $\Phi : I \subset (0,\infty) \to \mathbb{R}$ be a twice differentiable function on I. If there exist the constants $\omega < \Omega$ such that*

$$
\omega \leq t^2 \Phi''(t) \leq \Omega \quad \text{for any } t \in I, \tag{20}
$$

then we have the bounds

$$
\omega \mathcal{J}_{n,0}(x,y) \leq \mathcal{J}_{n,\Phi}(x,y) \leq \Omega \mathcal{J}_{n,0}(x,y) \quad \text{for any } x,y \in I^n. \tag{21}
$$

If there exist the constants $\lambda < \Lambda$ such that

$$
\lambda \leq t\Phi''(t) \leq \Lambda \quad \text{for any } t \in I, \tag{22}
$$

then we have the bounds

$$
\lambda \mathcal{J}_{n,1}(x,y) \leq \mathcal{J}_{n,\Phi}(x,y) \leq \Lambda \mathcal{J}_{n,1}(x,y) \quad \text{for any } x,y \in I^n. \tag{23}
$$

Proof Consider the auxiliary function $g_{\omega,0} : I \to \mathbb{R}$ with $g_{\omega,0}(t) = \Phi(t) + \omega \log t$. We observe that $g_{\omega,0}$ is twice differentiable, and by (20) we have $g_{\omega,0}''(t) = t^{-2}(t^2 \Phi''(t) - \omega) \geq 0$

for any $t \in I$, then we can conclude that $g_{\omega,0}$ is a convex function on I. Therefore we have $\mathcal{J}_{n,g_{\omega,0}}(x,y) \geq 0$ for any $x, y \in I^n$, which implies that

$$0 \leq \mathcal{J}_{n,\Phi}(x,y) + \omega \sum_{i=1}^{n} \left[\frac{\log x_i + \log y_i}{2} - \log\left(\frac{x_i + y_i}{2} \right) \right]$$

$$= \mathcal{J}_{n,\Phi}(x,y) - \omega \mathcal{J}_{n,0}(x,y)$$

and the first inequality in (21) is proved. Now, consider the auxiliary function $g_{\Omega,0} : I \to \mathbb{R}$ with $g_{\Omega,0}(t) = -\Omega \log t - \Phi(t)$. Then $g_{\Omega,0}''(t) = t^{-2}(\Omega - t^2 \Phi''(t))$ for any $t \in I$; and by (20) it is a convex function on I. By similar arguments, we deduce the second inequality in (21).

To prove the second part of the theorem, consider the auxiliary function $g_{\lambda,1} : I \to \mathbb{R}$, $g_{\lambda,1}(t) = \Phi(t) - \lambda t \log t$. We observe that $g_{\lambda,1}$ is twice differentiable and $g_{\lambda,1}''(t) := \Phi''(t) - \frac{1}{t}\lambda$, for $t \in I$. Since by (22) we have $g_{\lambda,1}''(t) = t^{-1}(t\Phi''(t) - \lambda) \geq 0$ for all $t \in I$, then we can conclude that $g_{\lambda,1}''$ is a convex function on I. The proof now follows along the lines outlined above and the first part of (23) is proved. The second part of (23) also follows by employing the auxiliary function $g_{\Lambda,1} : I \to \mathbb{R}$, $g_{\Lambda,1}(t) = \Lambda t \log t - \Phi(t)$; and this completes the proof. □

5 Applications to some elementary functions

We consider the approximations mentioned in Section 4 for some elementary functions.

We consider the function $\Phi(t) = e^{-t}$ for $t \in [a,b] \subset [0,1]$ and have the following bounds for all $x, y \in [a,b]^n$:

$$a^2 e^{-a} \mathcal{J}_{n,0}(x,y) \leq \mathcal{J}_{n,\Phi}(x,y) \leq b^2 e^{-b} \mathcal{J}_{n,0}(x,y),$$

$$a e^{-a} \mathcal{J}_{n,1}(x,y) \leq \mathcal{J}_{n,\Phi}(x,y) \leq b e^{-b} \mathcal{J}_{n,1}(x,y),$$

$$\frac{1}{2} e^{-b} \mathcal{J}_{n,2}(x,y) \leq \mathcal{J}_{n,\Phi}(x,y) \leq \frac{1}{2} e^{-a} \mathcal{J}_{n,2}(x,y).$$

In what follows, we apply these bounds to the above function on the interval $[0.1, 1]$, where $x = (0.2, 0.25, 0.3, \ldots, 1)$ and $y = (1, \ldots, 1)$ (*cf.* Figure 1).
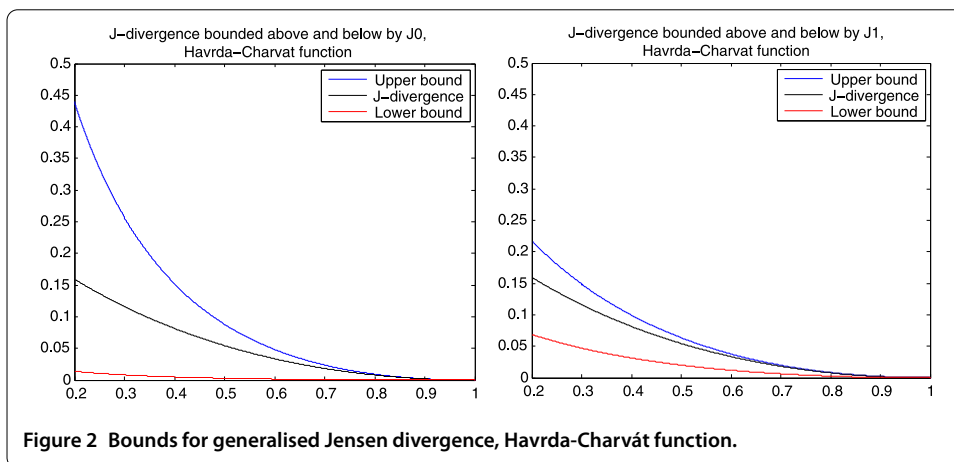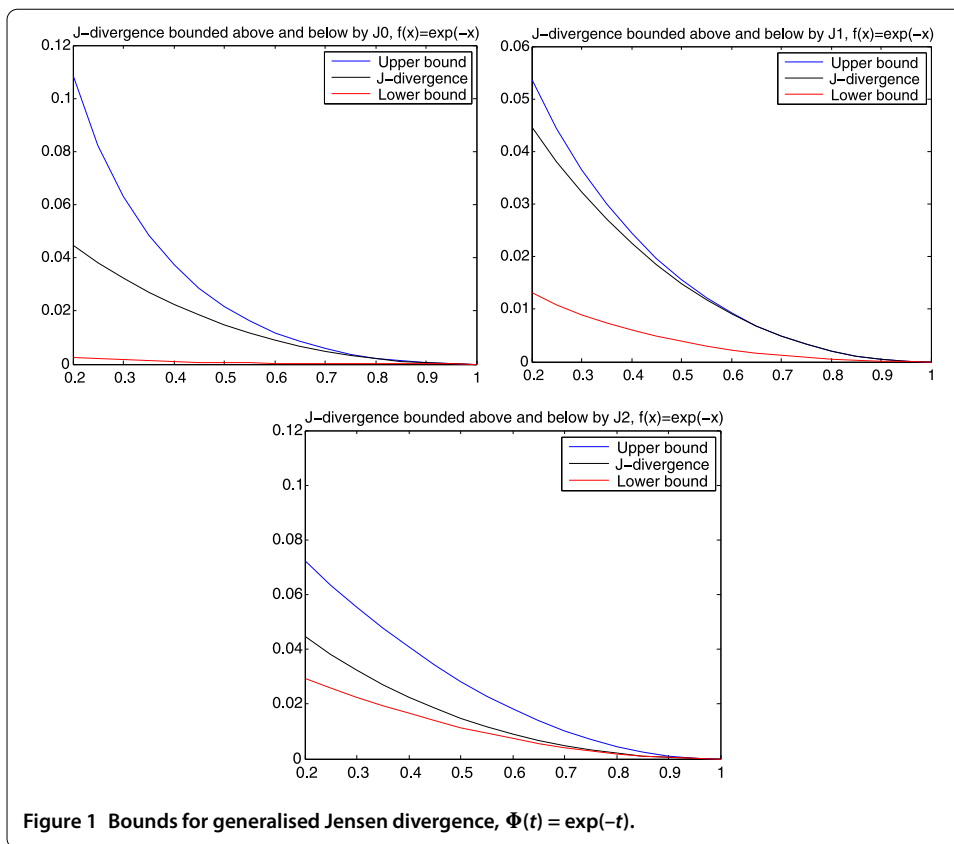
Discussion In this example, the best lower approximation (amongst the three) is given by $\frac{1}{2} e^{-1} \mathcal{J}_{n,2}(x,y)$, and the best upper approximation is given by $e^{-1} \mathcal{J}_{n,1}(x,y)$, where $x = (0.2, 0.25, 0.3, \ldots, 1)$ and $y = (1, \ldots, 1)$. However, it remains an open question whether this is true in general.

We consider the Havrda-Charvát function

$$\Phi_\alpha(t) = \begin{cases} (\alpha - 1)^{-1}(t^\alpha - t) & \text{if } \alpha \neq 1, \\ t \log(t) & \text{if } \alpha = 1. \end{cases}$$

For $\alpha = 1$, we have the following bounds for all $x, y \in [a,b]^n$:

$$a \mathcal{J}_{n,0}(x,y) \leq \mathcal{J}_{n,\Phi}(x,y) \leq b \mathcal{J}_{n,0}(x,y) \quad \text{for } [a,b] \subset [0,\infty);$$

$$a \mathcal{J}_{n,1}(x,y) \leq \mathcal{J}_{n,\Phi}(x,y) \leq b \mathcal{J}_{n,1}(x,y) \quad \text{for } 0 \leq a \leq 1 \leq b < \infty.$$

Figure 1 Bounds for generalised Jensen divergence, $\Phi(t) = \exp(-t)$.

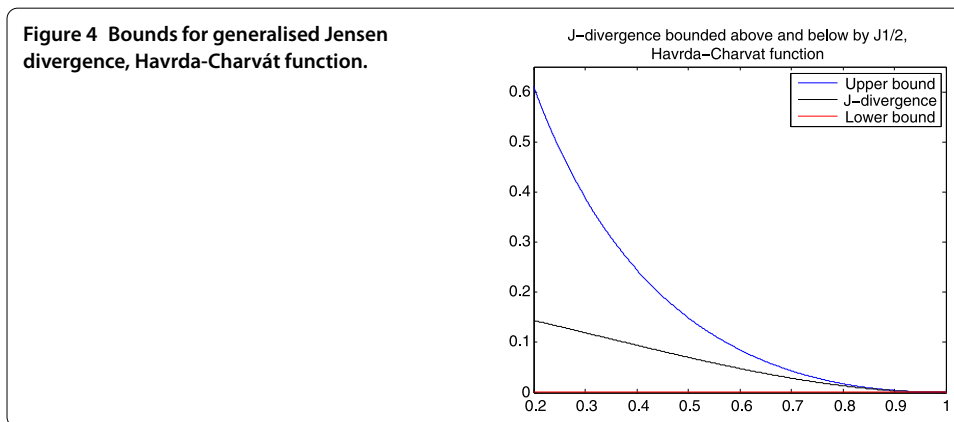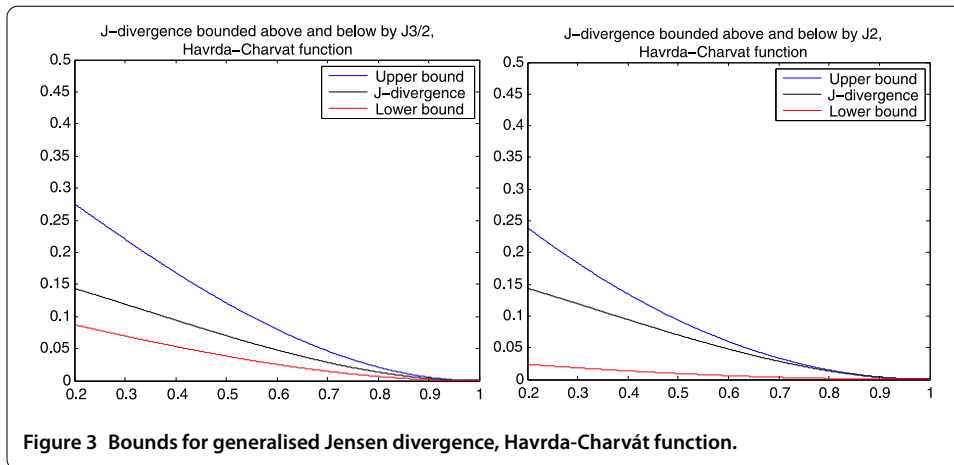**Figure 2 Bounds for generalised Jensen divergence, Havrda-Charvát function.**

We have the following bounds for all $x, y \in [a, b]^n$:

$$\alpha a^\alpha \mathcal{J}_{n,0}(x,y) \leq \mathcal{J}_{n,\Phi}(x,y) \leq \alpha b^\alpha \mathcal{J}_{n,0}(x,y) \quad \text{for } \alpha \geq 0, \alpha \neq 1 \text{ and } [a,b] \subset [0,\infty),$$

$$\alpha a^{\alpha-1} \mathcal{J}_{n,1}(x,y) \leq \mathcal{J}_{n,\Phi}(x,y) \leq \alpha b^{\alpha-1} \mathcal{J}_{n,1}(x,y) \quad \text{for } \alpha > 0 \text{ and } [a,b] \subset [0,\infty).$$

In Figure 2, we apply these bounds to the above function on the interval $[0.1, 1]$, where $x = (0.2, 0.201, 0.202, \ldots, 1)$, $y = (1, \ldots, 1)$, $\alpha = 3/2$.

Figure 3 Bounds for generalised Jensen divergence, Havrda-Charvát function.

Figure 4 Bounds for generalised Jensen divergence, Havrda-Charvát function.

We also have, for all $x, y \in [a, b]^n$,

$$\gamma(p-1)\mathcal{J}_{n,p}(x,y) \leq \mathcal{J}_{n,\Phi}(x,y) \leq \Gamma(p-1)\mathcal{J}_{n,p}(x,y)$$

for $p \in (-\infty, 0) \cup (1, \infty)$, where

$$\Gamma = \alpha \frac{b^{\alpha-p}}{p(p-1)} \quad \text{and} \quad \gamma = \alpha \frac{a^{\alpha-p}}{p(p-1)}$$

for $\alpha \geq p$ and $[a, b] \subset [0, \infty)$.

In Figure 3, we apply these bounds to the above function on the interval $[0.1, 1]$, where $x = (0.2, 0.201, 0.202, \ldots, 1)$, $y = (1, \ldots, 1)$, $\alpha = 3$ and $p = 3/2, 2$.

Similarly, we have, for all $x, y \in [a, b]^n$,

$$\delta(q-1)\mathcal{J}_{n,q}(x,y) \geq \mathcal{J}_{n,\Phi}(x,y) \geq \Delta(q-1)\mathcal{J}_{n,q}(x,y)$$

for $q \in (0, 1)$ and $\alpha > 1$, where

$$\Delta = \alpha \frac{a^{\alpha-q}}{q(q-1)} \quad \text{and} \quad \delta = \alpha \frac{b^{\alpha-q}}{q(q-1)}$$

for $[a, b] \subset [0, \infty)$. In Figure 4, we apply these bounds to the above function on the interval $[0.1, 1]$, where $x = (0.2, 0.201, 0.202, \ldots, 1)$, $y = (1, \ldots, 1)$, $q = 1/2$ and $\alpha = 3$.

Discussion In this example, the best lower approximation (amongst the five) is given by $2(0.1)^{3/2} \mathcal{J}_{n,3/2}$, and the best upper approximation is given by $(3/2)\mathcal{J}_{n,1}(x, y)$, where $x = (0.2, 0.201, 0.202, \ldots, 1)$, $y = (1, \ldots, 1)$. However, it remains an open question whether this is true in general.

Author details
[1]School of Computational and Applied Mathematics, University of the Witwatersrand, Private Bag 3, Wits 2050, Johannesburg, South Africa. [2]School of Engineering and Science, Victoria University, P.O. Box 14428, Melbourne, Victoria 8001, Australia.

References
1. Dragomir, SS: Some reverses of the Jensen inequality with applications. RGMIA Research Report Collection (Online) **14**, v14a72 (2011)
2. Burbea, J, Rao, CR: On the convexity of some divergence measures based on entropy functions. IEEE Trans. Inf. Theory **28**(3), 489-495 (1982)
3. Havrda, ME, Charvát, F: Quantification method of classification processes: concept of structural α-entropy. Kybernetika **3**, 30-35 (1967)
4. Lin, J: Divergence measures based on the Shannon entropy. IEEE Trans. Inf. Theory **37**(1), 145-151 (1991)
5. Grosse, I, Bernaola-Galvan, P, Carpena, P, Roman-Roldan, R, Oliver, J, Stanley, HE: Analysis of symbolic sequences using the Jensen-Shannon divergence. Phys. Rev. E, Stat. Nonlinear Soft Matter Phys. **65**(4), 041905 (2002). doi:10.1103/PhysRevE.65.041905
6. Kullback, S, Leibler, RA: On information and sufficiency. Ann. Math. Stat. **22**, 79-86 (1951)
7. Csiszar, I: Information-type measures of difference of probability distributions and indirect observations. Studia Sci. Math. Hung. **2**, 299-318 (1967)
8. Shannon, CE: A mathematical theory of communications. Bell Syst. Tech. J. **27**, 379-423, 623-565 (1948)
9. Menendez, ML, Pardo, JA, Pardo, L: Some statistical applications of generalized Jensen difference divergence measures for fuzzy information systems. Fuzzy Sets Syst. **52**, 169-180 (1992)
10. Arvey, AJ, Azad, RK, Raval, A, Lawrence, JG: Detection of genomic islands via segmental genome heterogeneity. Nucleic Acids Res. **37**(16), 5255-5266 (2009)
11. Gómez, RM, Rosso, OA, Berretta, R, Moscato, P: Uncovering molecular biomarkers that correlate cognitive decline with the changes of Hippocampus' gene expression profiles in Alzheimer's disease. PLoS ONE **5**(4), e10153 (2010). doi:10.1371/journal.pone.0010153
12. Dragomir, SS, Dragomir, NM, Sherwell, D: Sharp bounds for the Jensen divergence with applications. RGMIA Research Report Collection (Online) **14**, v14a47 (2011)
13. Bullen, PS: Handbook of Means and Their Inequalities. Mathematics and Its Applications, vol. 560. Kluwer Academic, Dordrecht (2003)