Journal of
Mathematical
Inequalities
Volume 18, Number 1 (2024), 287–314
doi:10.7153/jmi-2024-18-17

INEQUALITIES FOR FUNCTIONS CONVEX ON THE COORDINATES WITH APPLICATIONS TO JENSEN AND HERMITE–HADAMARD TYPE INEQUALITIES, AND TO NEW DIVERGENCE FUNCTIONALS

LÁSZLÓ HORVÁTH

(*Communicated by M. Klaričić Bakula*)

Abstract. In this paper we show that inequalities for functions convex on the coordinates can be derived from inequalities for convex functions defined on real intervals, and essentially only this method works. As applications, we show how our result works for the Jensen's and Hermite-Hadamard inequalities for functions convex on the coordinates. Finally, we extend the classical notion of f-divergence functional to functions convex on the coordinates, and as a further application of our main result, we study the refinement of a basic inequality corresponding to the new divergence.

1. Introduction

In this paper we work with functions convex on the coordinates, a concept introduced by Dragomir in paper [5].

DEFINITION 1. Let us consider the bidimensional interval $\Delta := I \times J$ in $\mathbb{R}^2$, where I and J are compact intervals in $\mathbb{R}$ with nonempty interior. A function $f : \Delta \to \mathbb{R}$ is said to be convex on the coordinates if the partial mappings $f_q : I \to \mathbb{R}$, $f_q(s) := f(s,q)$ and $f_p : J \to \mathbb{R}$, $f_p(t) := f(p,t)$ are convex for all $q \in J$ and $p \in I$. The set of all functions convex on the coordinates on Δ will be denoted by F_Δ^{co}.

The notion is interesting because the set of convex functions defined on Δ is a proper subset of the class of functions convex on the coordinates on Δ, but the basic inequalities for convex functions (Jensen type inequalities, Hermite-Hadamard type inequalities) can also be formulated for functions convex on the coordinates (see e.g. the papers [1], [5], [13], [14], [17], [18] and [21]).

Definition 1 is obviously meaningful even if the intervals I and J are not compact, and most of the results could be formulated and verified under such conditions, but in general in a less transparent way. This is also true for the main result of this paper, so we shall work with compact intervals except the last section.

Mathematics subject classification (2020): 26D15, 26A51, 94A17.

Keywords and phrases: Functions convex on the coordinates, measures, Jensen's and Hermite-Hadamard type inequalities, refinement, f-divergence functional.

Research supported by the Hungarian National Research, Development and Innovation Office grant no. K139346.

The main result of this paper is that inequalities for functions convex on the coordinates can be derived from inequalities for convex functions defined on real intervals, and essentially only this method works. This is important because it shows that inequalities for functions convex on the coordinates can be constructed in a simple way. The result also explains why some of the inequalities associated with convex functions can be extended to functions convex on the coordinates, even though the latter are not convex in general. As applications, we show how our result works for the Jensen's and Hermite-Hadamard inequalities for functions convex on the coordinates. On the one hand, we give new proofs of these results under general conditions (Borel measures are considered on the intervals I and J), and on the other hand, we provide methods for generating refinements of these inequalities and give some new concrete refinements. Most of the results of previous papers on the refinement of Jensen's and Hermite-Hadamard inequalities for functions convex on the coordinates can be obtained, and even generalized, using our method. Some new and interesting results are also given on the Hermite-Hadamard inequality and its refinements for convex functions defined on real intervals, based on the recent papers [6] and [8]. Finally, we extend the classical notion of f-divergence functional to functions convex on the coordinates. We prove that the new notion has the basic properties of f-divergence, so that it can also be considered as a divergence functional. Then, as a further application of our main result, we study the refinement of a basic inequality corresponding to the new divergence.

2. Preliminary results

Let $(X, \mathscr{A}, \mu)$ be a measure space, where $\mu(X) < \infty$. The set of all μ-integrable real functions on X is denoted by $L(\mu)$.

Let $C \subset \mathbb{R}$ be an interval with nonempty interior. We denote by F_C the set of all convex functions on C.

If C is a compact interval, $\varphi : X \to C$ is a measurable function and $f \in F_C$, then $f \circ \varphi$ is also measurable, and since φ and f are bounded and μ is finite, φ, $f \circ \varphi \in L(\mu)$.

We need the following easily verifiable result (see Theorem 1.1.2 in [15]).

LEMMA 1. *Let $C \subset \mathbb{R}$ be an interval with nonempty interior, and let $f \in F_C$. If t_0 is an interior point of C and $\delta > 0$ such that $[t_0 - \delta, t_0 + \delta] \subset C$, then*

$$|f(t_0 \pm \lambda\delta) - f(t_0)| \leqslant \lambda \max(|f(t_0 - \delta) - f(t_0)|, |f(t_0 + \delta) - f(t_0)|)$$

for all $\lambda \in [0,1]$.

In the first lemma, we study the boundedness, continuity and measurability properties of functions convex on the coordinates.

LEMMA 2. *Let $\Delta := I \times J$ be an interval in $\mathbb{R}^2$, where I and J are compact intervals in $\mathbb{R}$ with nonempty interior. If $f \in F_\Delta^{co}$, then*

(a) It is bounded.

(b) It is continuous on the interior of the interval Δ.
(c) It is Borel-measurable.

Proof. Assume $I=[a,b]$ and $J=[c,d]$.
(a) By the convexity of the functions f_p $(p\in I)$ and f_q $(q\in J)$,

$$f(s,t)\leqslant \max\left(f(s,c),f(s,d),f(a,t),f(b,t)\right)$$

$$\leqslant \max\left(f(a,c),f(b,c),f(a,d),f(b,d)\right),\quad (s,t)\in\Delta,$$

and therefore the function f is bounded above.
Let $(p,q):=\left(\frac{a+c}{2},\frac{b+d}{2}\right)$ be the centre of Δ.
Then for all $(s,t)\in\Delta$ we have that

$$(p,q)=\frac{(p,t)}{2}+\frac{(p,2q-t)}{2}$$

and

$$(p,t)=\frac{(s,t)}{2}+\frac{(2p-s,t)}{2},$$

where the points

$$(p,q),\quad (p,t),\quad (p,2q-t),\quad (2p-s,t)$$

also belong to Δ.
Since the functions f_p and f_t are convex,

$$f(p,q)\leqslant \frac{f(s,t)}{4}+\frac{f(2p-s,t)}{4}+\frac{f(p,2q-t)}{2},$$

and hence

$$f(s,t)\geqslant 4f(p,q)-f(2p-s,t)-2f(p,2q-t)\geqslant 4f(p,q)-3K,$$

where K is an upper bound of f. This implies that the function f is bounded below.
(b) By part (a), there is $M>0$ such that $|f(s,t)|\leqslant M$ for all $(s,t)\in\Delta$.
Let (s_0,t_0) be an interior point of Δ and choose $\delta>0$ such that

$$\Delta_\delta:=[s_0-\delta,s_0+\delta]\times[t_0-\delta,t_0+\delta]\subset\Delta.$$

If $(s,t)\in\Delta_\delta$, then

$$|f(s,t)-f(s_0,t_0)|\leqslant |f(s,t)-f(s_0,t)|+|f(s_0,t)-f(s_0,t_0)|$$

$$=|f(s_0\pm\lambda_1\delta,t)-f(s_0,t)|+|f(s_0,t_0\pm\lambda_2\delta)-f(s_0,t_0)|,$$

where λ_1, $\lambda_2\in[0,1]$. By using Lemma 1, the convexity of the functions f_t and f_{s_0} implies that

$$|f(s,t)-f(s_0,t_0)|\leqslant 2M(\lambda_1+\lambda_2),$$

and hence f is continuous at (s_0,t_0).

(c) Let $\alpha \in \mathbb{R}$ be fixed, let $H_\alpha := f^{-1}([\alpha,\infty[)$, and let

$$\Delta_a := \{(a,t) \mid t \in [c,d]\}, \quad \Delta_b := \{(b,t) \mid t \in [c,d]\},$$
$$\Delta_c := \{(s,c) \mid s \in [a,b]\}, \quad \Delta_d := \{(s,d) \mid s \in [a,b]\}.$$

Then

$$H_\alpha = H_\alpha \bigcap (]a,b[\times]c,d[) \bigcup \left(H_\alpha \bigcap \Delta_a\right)$$
$$\bigcup \left(H_\alpha \bigcap \Delta_b\right) \bigcup \left(H_\alpha \bigcap \Delta_c\right) \bigcup \left(H_\alpha \bigcap \Delta_d\right). \tag{1}$$

Since f is continuous on the interior of Δ, $H_\alpha \bigcap]a,b[\times]c,d[$ is a Borel set. Since

$$H_\alpha \bigcap \Delta_a = f_a^{-1}([\alpha,\infty[)$$

and f_a is convex, $H_\alpha \bigcap \Delta_a$ is also a Borel set. It can be seen similarly that the sets in (1) are Borel sets too.

The proof is complete. □

REMARK 1. (a) Even among convex functions on Δ, it is easy to construct one that is not continuous at any point on the boundary of Δ.

(b) It was pointed out in the introduction that Definition 1 makes sense without the compactness of the intervals I and J. For example, in our previous result, only part (a) would change as follows: f is locally bounded.

The next lemma will be important in the proof of the main result.

LEMMA 3. *Let $\Delta := I \times J$ be an interval in $\mathbb{R}^2$, where I and J are compact intervals in $\mathbb{R}$ with nonempty interior, and let $(X,\mathscr{A},\mu)$ be a measure space, where $\mu(X) < \infty$. If $\varphi : X \to I$ is a measurable function, and $f \in F_\Delta^{co}$, then the function*

$$l_\varphi : J \to \mathbb{R}, \quad l_\varphi(t) := \int_X f(\varphi(x),t)\, d\mu(x)$$

is well defined and convex.

Proof. Since $f_q \circ \varphi \in L(\mu)$ for all $q \in J$, the definition of l_φ is correct.

Let t_1, $t_2 \in J$ and $\lambda \in [0,1]$. Since the function f_p is convex for all $p \in I$, and the integral is monotone,

$$l_\varphi(\lambda t_1 + (1-\lambda)t_2) = \int_X f(\varphi(x), \lambda t_1 + (1-\lambda)t_2)\, d\mu(x)$$
$$\leqslant \int_X (\lambda f(\varphi(x),t_1) + (1-\lambda) f(\varphi(x),t_2))\, \mu(x)$$
$$= \lambda l_\varphi(t_1) + (1-\lambda) l_\varphi(t_2),$$

which implies that l_φ is convex, that is $l_\varphi \in F_J$.

The proof is complete. □

Among inequalities for convex functions, integral Jensen inequalities of different types play a fundamental role.

THEOREM 1. *(a)* (discrete Jensen inequality, see [15]) *Let* $t_1,\ldots,t_n$ *be points from an interval* $C \subset \mathbb{R}$*, and let* $\alpha_1,\ldots,\alpha_n$ *be scalars in* $[0,1]$ *with* $\sum_{i=1}^{n}\alpha_i = 1$*. If* $f \in F_C$*, then*

$$f\left(\sum_{i=1}^{n}\alpha_i t_i\right) \leqslant \sum_{i=1}^{n}\alpha_i f(t_i).$$

(b) (integral Jensen inequality, see [15]) *Let* η *be an integrable function on a probability space* $(X,\mathscr{A},\mu)$ *taking values in an interval* $C \subset \mathbb{R}$*. Then* $\int_X \eta\, d\mu$ *lies in* C*. If* $f \in F_C$ *is such that* $f \circ \eta$ *is* μ*-integrable, then*

$$f\left(\int_X \eta\, d\mu\right) \leqslant \int_X f \circ \eta\, d\mu. \tag{2}$$

The following inequality will be used in the applications. In this result the intervals are not compact.

LEMMA 4. *Let* $(u_k)_{k=1}^m$*,* $(v_k)_{k=1}^m$ *and* $(z_l)_{l=1}^n$*,* $(w_l)_{l=1}^n$ *be positive sequences, where* k *and* l *are positive integers, and let* $\Delta :=]0,\infty[\times]0,\infty[$*. Then for all* $f \in F_\Delta^{co}$ *inequality*

$$\left(\sum_{k=1}^{m}u_k\right)\left(\sum_{l=1}^{n}z_l\right) f\left(\frac{\sum_{k=1}^{m}v_k}{\sum_{k=1}^{m}u_k}, \frac{\sum_{l=1}^{n}w_l}{\sum_{l=1}^{n}z_l}\right) \leqslant \sum_{k=1}^{m}\sum_{l=1}^{n}u_k z_l f\left(\frac{v_k}{u_k}, \frac{w_l}{z_l}\right)$$

holds.

Proof. Using the discrete Jensen inequality in both variables, we easily obtain

$$\begin{aligned}
&\left(\sum_{k=1}^{m}u_k\right)\left(\sum_{l=1}^{n}z_l\right) f\left(\frac{\sum_{k=1}^{m}v_k}{\sum_{k=1}^{m}u_k}, \frac{\sum_{l=1}^{n}w_l}{\sum_{l=1}^{n}z_l}\right)\\
&= \left(\sum_{k=1}^{m}u_k\right)\left(\sum_{l=1}^{n}z_l\right) f\left(\sum_{k=1}^{m}\frac{v_k}{u_k}\frac{u_k}{\sum_{k=1}^{m}u_k}, \sum_{l=1}^{n}\frac{w_l}{z_l}\frac{z_l}{\sum_{l=1}^{n}z_l}\right)\\
&\leqslant \sum_{k=1}^{m}\sum_{l=1}^{n}u_k z_l f\left(\frac{v_k}{u_k}, \frac{w_l}{z_l}\right).
\end{aligned}$$

The proof is complete. □

Let $C \subset \mathbb{R}$ be an interval with nonempty interior. The following notations are introduced for some special functions defined on C:

$$id_C(t) := t, \quad p_{C,w}(t) := (t-w)^+, \quad t,w \in C,$$

where a^+ means the positive part of $a \in \mathbb{R}$.

We shall also use the following assertion, which is a very simple special case of the main result in paper [6].

The interior of an interval $C \subset \mathbb{R}$ is denoted by C°.

THEOREM 2. *Let $(X, \mathscr{A}, \mu)$ and $(Y, \mathscr{B}, \nu)$ be measure spaces, where μ and ν are finite measures. Let $C \subset \mathbb{R}$ be a compact interval with nonempty interior, and let $\varphi : X \to C$, $\psi : Y \to C$ be measurable functions. Then for every $f \in F_C$ inequality*

$$\int_X f \circ \varphi d\mu \leqslant \int_Y f \circ \psi d\nu$$

holds if and only if

$$\mu(X) = \nu(Y), \quad \int_X \varphi d\mu = \int_Y \psi d\nu$$

and it is satisfied in the following special cases: the function f is $p_{C,w}$ $(w \in C^\circ)$.

Another very important inequality for convex functions is the Hermite-Hadamard inequality. Next, we give a general Hermite-Hadamard type inequality and a method to refine it.

The σ-algebra of Borel subsets of an interval $C \subset \mathbb{R}$ will be denoted by $\mathscr{B}_C$.

THEOREM 3. *Let $[a,b] \subset \mathbb{R}$ with nonempty interior, let μ, ν_i and ξ_i $(i = 1,2)$ be finite measures on $\mathscr{B}_{[a,b]}$ with $\mu([a,b]) > 0$, and let φ_1, φ_2, ψ_1, $\psi_2 : [a,b] \to [a,b]$ are Borel-measurable functions. Let*

$$x_\mu := \frac{1}{\mu([a,b])} \int_{[a,b]} t d\mu(t).$$

Then

(a) Inequalities

$$\int_{[a,b]} f \circ \varphi_2 d\nu_2 \leqslant \int_{[a,b]} f \circ \varphi_1 d\nu_1 \leqslant \int_{[a,b]} f d\mu \leqslant \int_{[a,b]} f \circ \psi_1 d\xi_1 \leqslant \int_{[a,b]} f \circ \psi_2 d\xi_2 \tag{3}$$

are satisfied for all $f \in F_{[a,b]}$ if and only if

$$\nu_i([a,b]) = \xi_i([a,b]) = \mu([a,b]), \quad i = 1,2, \tag{4}$$

and

$$\int_{[a,b]} \varphi_i d\nu_i = \int_{[a,b]} \psi_i d\xi_i = \int_{[a,b]} t d\mu(t), \quad i = 1,2$$

and

$$\int_{[a,b]} p_{[a,b],w} \circ \varphi_2 d\nu_2 \leqslant \int_{[a,b]} p_{[a,b],w} \circ \varphi_1 d\nu_1 \leqslant \int_{[w,b]} (t-w)\, d\mu(t)$$

$$\leqslant \int_{[a,b]} p_{[a,b],w} \circ \psi_1 d\xi_1 \leqslant \int_{[a,b]} p_{[a,b],w} \circ \psi_2 d\xi_2, \quad w \in \,]a,b[$$

hold.

(b) If $\nu_2([a,b]) = \mu([a,b])$ *and* $\int_{[a,b]} \varphi_2 d\nu_2 = \int_{[a,b]} t d\mu(t)$*, then for all* $f \in F_{[a,b]}$ *we have*

$$f(x_\mu)\, \mu([a,b]) \leqslant \int_{[a,b]} f \circ \varphi_2 d\nu_2.$$

(c) If $\xi_2([a,b]) = \mu([a,b])$ *and* $\int_{[a,b]} \psi_2 d\xi_2 = \int_{[a,b]} t d\mu(t)$*, then for all* $f \in F_{[a,b]}$ *we have*

$$\int_{[a,b]} f \circ \psi_2 d\xi_2 \leqslant \left(\frac{b-x_\mu}{b-a} f(a) + \frac{x_\mu - a}{b-a} f(b) \right) \mu([a,b]). \tag{5}$$

(d) If μ *is symmetric in the sense that*

$$\mu(A) = \mu(a+b-A), \quad A \in \mathscr{B}_{[a,b]}, \tag{6}$$

then

$$x_\mu = \frac{a+b}{2}, \quad \frac{b-x_\mu}{b-a} f(a) + \frac{x_\mu - a}{b-a} f(b) = \frac{f(a)+f(b)}{2}.$$

Proof. (a) This part of the result follows directly from Theorem 2.

(b) By using the integral Jensen inequality and the conditions,

$$\int_{[a,b]} f \circ \varphi_2 d\nu_2 \geqslant \nu_2([a,b])\, f\left(\frac{1}{\nu_2([a,b])} \int_{[a,b]} t d\nu_2(t) \right) = f(x_\mu)\, \mu([a,b]).$$

(c) It follows from the convexity of f that

$$f(\psi_2(t)) \leqslant \frac{b - \psi_2(t)}{b-a} f(a) + \frac{\psi_2(t) - a}{b-a} f(b), \quad t \in [a,b].$$

By integrating both sides of this inequality, and taking into account the conditions we obtain the result.

(d) Define the function $T : \left[a, \frac{a+b}{2}\right] \to \left[\frac{a+b}{2}, b\right]$ by $T(t) := a+b-t$. Let $T(\mu)$ be the image measure of the restriction of μ to $\mathscr{B}_{\left[a, \frac{a+b}{2}\right]}$ under the mapping T. If $B \in \mathscr{B}_{\left[\frac{a+b}{2}, b\right]}$, then by (6),

$$\mu\left(T^{-1}(B)\right) = \mu(a+b-B) = \mu(B),$$

and hence $T(\mu)$ is the restriction of μ to $\mathscr{B}_{\left[\frac{a+b}{2},b\right]}$. This implies that

$$\int\limits_{[a,b]} td\mu(t) = \int\limits_{\left[a,\frac{a+b}{2}\right]} td\mu(t) + \int\limits_{\left]\frac{a+b}{2},b\right]} tdT(\mu)(t)$$

$$= \int\limits_{\left[a,\frac{a+b}{2}\right]} td\mu(t) + \int\limits_{\left[a,\frac{a+b}{2}\right[} (a+b-t)d\mu(t) = \frac{a+b}{2}\mu\left(\left\{\frac{a+b}{2}\right\}\right)$$

$$+(a+b)\mu\left(\left[a,\frac{a+b}{2}\right[\right) = \frac{a+b}{2}\mu([a,b]).$$

The proof is complete. □

REMARK 2. Even for signed measures, necessary and sufficient conditions are known to satisfy either the first or the second inequality of

$$f(x_\mu)\mu([a,b]) \leqslant \int\limits_{[a,b]} fd\mu \leqslant \left(\frac{b-x_\mu}{b-a}f(a) + \frac{x_\mu - a}{b-a}f(b)\right)\mu([a,b]) \tag{7}$$

for any convex function f on $[a,b]$ (see the book [15] and the paper [8]). Our previous result is mainly interesting for refining the inequalities in (7). Hermite-Hadamard inequality for measures can also be found in paper [16].

It is worth mentioning two special cases of the previous theorem. In the first case

$$\varphi_i(t) = \psi_i(t) = t, \quad t \in [a,b], \quad i = 1,2,$$

while in the second case

$$\nu_i = \xi_i = \mu, \quad i = 1,2.$$

PROPOSITION 1. *Let $[a,b] \subset \mathbb{R}$ with nonempty interior, let μ, ν_i and ξ_i $(i = 1,2)$ be finite measures on $\mathscr{B}_{[a,b]}$ with $\mu([a,b]) > 0$.*

Then

(a) Inequalities

$$\int\limits_{[a,b]} fd\nu_2 \leqslant \int\limits_{[a,b]} fd\nu_1 \leqslant \int\limits_{[a,b]} fd\mu \leqslant \int\limits_{[a,b]} fd\xi_1 \leqslant \int\limits_{[a,b]} fd\xi_2$$

are satisfied for all $f \in F_{[a,b]}$ if and only if (4) holds,

$$\int\limits_{[a,b]} td\nu_i(t) = \int\limits_{[a,b]} td\xi_i(t) = \int\limits_{[a,b]} td\mu(t), \quad i = 1,2 \tag{8}$$

and for all of the signed measures

$$\eta_1 := \xi_2 - \xi_1, \quad \eta_2 := \xi_1 - \mu, \quad \eta_3 := \mu - \nu_1, \quad \eta_4 := \nu_1 - \nu_2$$

we have

$$\int_{[w,b]} (t-w)\,d\eta_i(t) \geqslant 0, \quad i=1,2,3,4, \quad w \in]a,b[. \tag{9}$$

(b) Under the conditions (4)*,* (8) *and* (9)

$$f(x_\mu)\,\mu([a,b]) \leqslant \int_{[a,b]} f\,d\nu_2$$

and

$$\int_{[a,b]} f\,d\xi_2 \leqslant \left(\frac{b-x_\mu}{b-a} f(a) + \frac{x_\mu - a}{b-a} f(b)\right)\mu([a,b]).$$

Proof. It is an immediate consequence of Theorem 3. □

REMARK 3. It is easy to check that (4), (8) and (9) imply

$$\int_{[a,w]} (w-t)\,d\eta_i(t) \geqslant 0, \quad i=1,2,3,4, \quad w \in]a,b[,$$

so in this case the measures η_i $(i=1,2,3,4)$ are essentially Steffensen-Popoviciu measures (see [15]). They would be true Steffensen-Popoviciu measures if $\eta_i([a,b]) > 0$ $(i=1,2,3,4)$ were satisfied.

In the case of equality of measures, we give only a sufficient condition, which is Theorem 8 in the recent paper [8].

PROPOSITION 2. *Let $[a,b] \subset \mathbb{R}$ with nonempty interior, and let μ be a finite measure on $\mathscr{B}_{[a,b]}$ such that $\mu([a,b]) > 0$. Assume φ_1, φ_2, ψ_1, $\psi_2 : [a,b] \to [a,b]$ are increasing functions such that*

$$\int_{[a,x]} \psi_2\,d\mu \leqslant \int_{[a,x]} \psi_1\,d\mu \leqslant \int_{[a,x]} t\,d\mu(t) \leqslant \int_{[a,x]} \varphi_1\,d\mu \leqslant \int_{[a,x]} \varphi_2\,d\mu, \quad x \in [a,b] \tag{10}$$

and

$$\int_{[a,b]} \varphi_i\,d\mu = \int_{[a,b]} \psi_i\,d\mu = \int_{[a,b]} t\,d\mu(t), \quad i=1,2 \tag{11}$$

are satisfied. Then for all $f \in F_{[a,b]}$ we have

$$f(x_\mu)\,\mu([a,b])$$

$$\leqslant \int_{[a,b]} f\circ\varphi_2\,d\mu \leqslant \int_{[a,b]} f\circ\varphi_1\,d\mu \leqslant \int_{[a,b]} f\,d\mu \leqslant \int_{[a,b]} f\circ\psi_1\,d\mu \leqslant \int_{[a,b]} f\circ\psi_2\,d\mu$$

$$\leqslant \left(\frac{b-x_\mu}{b-a} f(a) + \frac{x_\mu - a}{b-a} f(b)\right)\mu([a,b]).$$

Depending on the measure μ, it is easier or harder to construct functions that satisfy conditions (10) and (11), but it is straightforward under stronger conditions on the measure and the functions. The following result is derived from Theorem 9 (c) in [8].

PROPOSITION 3. *Let $[a,b] \subset \mathbb{R}$ with nonempty interior, and let μ be a finite measure on $\mathscr{B}_{[a,b]}$ such that $\mu([a,b]) > 0$. Assume φ_1, φ_2, ψ_1, $\psi_2 : [a,b] \to [a,b]$ are increasing functions. If the measure μ and the functions φ_1, φ_2 ψ_1, ψ_2 satisfy the symmetry properties*

$$\mu(A) = \mu(a+b-A), \quad A \in \mathscr{B}_{[a,b]} \tag{12}$$

and

$$\varphi(a+b-t) = a+b-\varphi(t), \quad t \in [a,b]$$

respectively, and

$$\psi_2(t) \leqslant \psi_1(t) \leqslant t \leqslant \varphi_1(t) \leqslant \varphi_2(t), \quad t \in \left[a, \frac{a+b}{2}\right],$$

then conditions (10) *and* (11) *hold.*

Moreover,

$$\int\limits_{[a,b]} f \circ \psi_2 d\mu \tag{13}$$

$$\leqslant \frac{f(a)+f(b)}{2}\mu([a,b]) + \left(f\left(\frac{a+b}{2}\right) - \frac{f(a)+f(b)}{2}\right)\mu\left(\left\{\frac{a+b}{2}\right\}\right).$$

REMARK 4. Inequality (13) is sharper than inequality (5).

The following result is a majorization type inequality which is contained in Theorem 9 of [6].

THEOREM 4. *Let $X := \{1,\ldots,m\}$ for some $m \in \mathbb{N}_+$, let $Y := \{1,\ldots,n\}$ for some $n \in \mathbb{N}_+$, and let $C \subset \mathbb{R}$ be an interval with nonempty interior. Assume $(p_i)_{i=1}^m$ and $(q_j)_{j=1}^n$ are real sequences, and $\mathbf{s} := (s_1,\ldots,s_m) \in C^m$ and $\mathbf{t} := (t_1,\ldots,t_n) \in C^n$. Let $u_1 > u_2 > \ldots > u_o$ be the different elements of $\mathbf{s}$ and $\mathbf{t}$ in decreasing order $(1 \leqslant o \leqslant m+n)$. Then for every $f \in F_C$ inequality*

$$\sum_{i=1}^m p_i f(s_i) \leqslant \sum_{j=1}^n q_j f(t_j) \tag{14}$$

holds if and only if

$$\sum_{i=1}^m p_i = \sum_{j=1}^n q_j, \quad \sum_{i=1}^m p_i s_i = \sum_{j=1}^n q_j t_j \tag{15}$$

and

$$\sum_{\{i\in X|s_i\geqslant u_l\}} p_i s_i - \sum_{\{j\in Y|t_j\geqslant u_l\}} q_j t_j$$
$$\leqslant u_l\left(\sum_{\{i\in X|s_i\geqslant u_l\}} p_i - \sum_{\{j\in Y|t_j\geqslant u_l\}} q_j\right), \quad l=1,\ldots,o \tag{16}$$

are satisfied.

3. Main result

Our main result shows that we can always derive inequalities for functions convex on the coordinates from inequalities for convex functions defined on intervals, and essentially only this method works.

THEOREM 5. *Let $\Delta := I\times J$ be an interval in $\mathbb{R}^2$, where I and J are compact intervals in $\mathbb{R}$ with nonempty interior, and let $(X_i,\mathscr{A}_i,\mu_i)$ and $(Y_i,\mathscr{B}_i,\nu_i)$ be measure spaces $(i=1,2)$, where $0<\mu_i(X_i)<\infty$ and $0<\nu_i(Y_i)<\infty$ $(i=1,2)$. Furthermore, let $\varphi_1 : X_1\to I$, $\varphi_2 : X_2\to J$, $\psi_1 : Y_1\to I$, $\psi_2 : Y_2\to J$ be measurable functions $(i=1,2)$, and let $\varphi := (\varphi_1,\varphi_2) : X_1\times X_2\to\Delta$ and $\psi := (\psi_1,\psi_2) : Y_1\times Y_2\to\Delta$.*

(a) If either $\mu_1(X_1)=\nu_1(Y_1)$ or $\mu_2(X_2)=\nu_2(Y_2)$ and for all $f\in F_\Delta^{co}$ we have

$$\int_{X_1\times X_2} f\circ\varphi d(\mu_1\times\mu_2)\leqslant\int_{Y_1\times Y_2} f\circ\psi d(\nu_1\times\nu_2), \tag{17}$$

then

$$\int_{X_1} g\circ\varphi_1 d\mu_1\leqslant\int_{Y_1} g\circ\psi_1 d\nu_1 \tag{18}$$

is satisfied for all $g\in F_I$ and inequality

$$\int_{X_2} h\circ\varphi_2 d\mu_2\leqslant\int_{Y_2} h\circ\psi_2 d\nu_2 \tag{19}$$

is satisfied for all $h\in F_J$.

(b) Conversely, if (18) is satisfied for all $g\in F_I$ and (19) is satisfied for all $h\in F_J$, then $\mu_i(X_i)=\nu_i(Y_i)$ $(i=1,2)$ and (17) holds for all $f\in F_\Delta^{co}$.

Proof. By Lemma 2 (c), f is Borel-measurable, and hence $f\circ\varphi$ is $\mathscr{A}_1\times\mathscr{A}_2$-measurable and $f\circ\psi$ is $\mathscr{B}_1\times\mathscr{B}_2$-measurable. Lemma 2 (a) implies that f is bounded. It now follows that the integrals in (17) exist, since $\mu_1\times\mu_2$ and $\nu_1\times\nu_2$ are finite.

(a) We consider the case $\mu_1(X_1)=\nu_1(Y_1)$, the other case can be discussed in an equivalent way.

Since the constant functions $f:\Delta\to\mathbb{R}$ by $f(s,t):=\pm 1$ belong to F_{Δ}^{co}, it follows from (17) that $\mu_1\times\mu_2(X_1\times X_2)=\nu_1\times\nu_2(Y_1\times Y_2)$, and hence $\mu_1(X_1)=\nu_1(Y_1)>0$ implies $\mu_2(X_2)=\nu_2(Y_2)$.

We show that inequality (18) is satisfied for all $g\in F_I$, a similar argument can be applied to inequality (19).

Let $g\in F_I$, and define the function $f:\Delta\to\mathbb{R}$ by $f(s,t):=g(s)$. Then $f\in F_{\Delta}^{co}$, and hence (17) and $\mu_2(X_2)=\nu_2(Y_2)>0$ imply that (18) holds.

(b) Inequality (18) holds for the constant functions $g:I\to\mathbb{R}$, $g(s):=\pm 1$, and hence $\mu_1(X_1)=\nu_1(Y_1)$. Similarly, we can obtain from (19) that $\mu_2(X_2)=\nu_2(Y_2)$.

Let $f\in F_{\Delta}^{co}$ be fixed.

Since inequality (18) is satisfied for all $g\in F_I$, it follows that

$$\int_{X_1} f_q\circ\varphi_1 d\mu_1 \leqslant \int_{Y_1} f_q\circ\psi_1 d\nu_1, \quad q\in J. \tag{20}$$

By Lemma 3, the function

$$l_{\varphi_1}:J\to\mathbb{R}, \quad l_{\varphi_1}(t):=\int_{X_1} f(\varphi_1(x_1),t)\,d\mu_1(x_1)$$

belongs to F_J. Since inequality (19) is satisfied for all $h\in F_J$, we have that

$$\int_{X_2}\left(\int_{X_1} f(\varphi_1(x_1),\varphi_2(x_2))\,d\mu_1(x_1)\right)d\mu_2(x_2)$$

$$\leqslant \int_{Y_2}\left(\int_{X_1} f(\varphi_1(x_1),\psi_2(y_2))\,d\mu_1(x_1)\right)d\nu_2(y_2). \tag{21}$$

Another application of Lemma 3 gives that the function

$$l_{\psi_1}:J\to\mathbb{R}, \quad l_{\psi_1}(t):=\int_{Y_1} f(\psi_1(y_1),t)\,d\nu_1(y_1)$$

also belongs to F_J.

Since l_{φ_1}, $l_{\psi_1}\in F_J$, it follows from (20) that

$$\int_{Y_2}\left(\int_{X_1} f(\varphi_1(x_1),\psi_2(y_2))\,d\mu_1(x_1)\right)d\nu_2(y_2)$$

$$\leqslant \int_{Y_2}\left(\int_{Y_1} f(\psi_1(y_1),\psi_2(y_2))\,d\nu_1\right)d\nu_2(y_2). \tag{22}$$

The result now follows from inequalities (21) and (22) by taking Fubini's theorem. The proof is complete. □

REMARK 5. (a) The preceding statement allows the simple generation of inequalities for functions convex on the coordinates from known inequalities for convex functions on real intervals.

(b) It comes from the proof that if inequality (17) is satisfied for all $f \in F_{\Delta}^{co}$, then $\mu_1(X_1)\mu_2(X_2) = \nu_1(Y_1)\nu_2(Y_2)$. But it is straightforward to construct concrete examples to show that inequality (17) can be satisfied for any $f \in F_{\Delta}^{co}$ such that $\mu_1(X_1) \neq \nu_1(Y_1)$ and $\mu_2(X_2) \neq \nu_2(Y_2)$. In these cases, neither inequality (18) is satisfied for all $g \in F_I$, nor inequality (19) is satisfied for all $h \in F_J$.

4. Jensen type inequalities for functions convex on the coordinates and their refinements

Let $(X, \mathscr{A})$ be a measurable space. The unit mass at $x \in X$ (the Dirac measure at x) will be denoted by ε_x.

In the following two results, we obtain a version of the integral Jensen inequality for functions convex on the coordinates, and give a refinement of it. The main objective is to illustrate the following phenomenon: by applying Theorem 5 to suitable refinements of the integral Jensen inequality, we can almost automatically obtain refinements to the integral Jensen inequality for functions convex on the coordinates.

To achieve this goal, we first recall two refinements of the integral Jensen inequality, far from being completely general.

THEOREM 6. *Let $(X, \mathscr{A})$ be a measurable space, let $\eta : X \to \mathbb{R}$ be a measurable function taking values in an interval $C \subset \mathbb{R}$, and let $f \in F_C$.*

(a) If μ is a probability measure on $\mathscr{A}$ and η, $f \circ \eta \in L(\mu)$, then

$$f\left(\int_X \eta \, d\mu\right) \leqslant \int_X f\left(\alpha \int_X \eta \, d\mu + (1-\alpha)\eta\right) d\mu \leqslant \int_X f \circ \eta \, d\mu$$

for all $\alpha \in [0,1]$.

(b) Let ν be a measure on $\mathscr{A}$ with $\nu(X) > 0$, and let w be a positive function on X such that $\int_X w\,d\nu = 1$. If ηw, $(f \circ \eta)w \in L(\nu)$ and $w_1, \ldots, w_n$ $(n \geqslant 2)$ are positive and measurable functions on X such that $\sum_{i=1}^{n} w_j = w$, then

$$f\left(\int_X \eta w \, d\nu\right) \leqslant \sum_{i=1}^{n} f\left(\frac{\int_X \eta w_i \, d\nu}{\int_X w_i \, d\nu}\right) \int_X w_i \, d\nu \leqslant \int_X (f \circ \eta) w \, d\nu.$$

Proof. (a) Using the definition of convexity and then the integral Jensen inequality twice, we obtain (a).

(b) It is an elementary case of the main result of paper [7].

The proof is complete. □

After these preparations, we can make the following two statements.

THEOREM 7. (Integral Jensen inequality for functions convex on the coordinates) *Let* $\Delta := I \times J$ *be an interval in* $\mathbb{R}^2$, *where* I *and* J *are compact intervals in* $\mathbb{R}$ *with nonempty interior, and let* $(X_1, \mathscr{A}_1, \mu_1)$ *and* $(X_2, \mathscr{A}_2, \mu_2)$ *be probability spaces. Furthermore, let* $\eta_1 : X_1 \to I$ *and* $\eta_2 : X_2 \to J$ *be measurable functions. Then*

$$\left(\int_{X_1} \eta_1 d\mu_1, \int_{X_2} \eta_2 d\mu_2 \right)$$

lies in Δ, *and for all* $f \in F_{\Delta}^{co}$

$$f\left(\int_{X_1} \eta_1 d\mu_1, \int_{X_2} \eta_2 d\mu_2 \right) \leqslant \int_{X_1 \times X_2} f \circ (\eta_1, \eta_2)\, d(\mu_1 \times \mu_2).$$

Proof. It follows directly from Theorem 5 (b) by using the integral Jensen inequality and

$$f\left(\int_{X_1} \eta_1 d\mu_1, \int_{X_2} \eta_2 d\mu_2 \right) = \int_{X_1 \times X_2} f\left(\int_{X_1} \eta_1 d\mu_1, \int_{X_2} \eta_2 d\mu_2 \right) d(\mu_1 \times \mu_2).$$

The proof is complete. $\square$

REMARK 6. (a) Theorem 7 is not new, it has already been proved in paper [14]. The main interest of the claim is the method of proof, which is based on Theorem 5. A special case of Theorem 7 has also been proved by using differentiation in [21] only for convex functions on Δ.

(b) The notion of a co-ordinate convex function is naturally generalizable to n-dimensional intervals, and Theorem 7 can also be easily formulated for this case. This variant can be found in paper [20], the form of the inequality and the proof are closely related to the analogous result in [14].

THEOREM 8. *Let* $\Delta := I \times J$ *be an interval in* $\mathbb{R}^2$, *where* I *and* J *are compact intervals in* $\mathbb{R}$ *with nonempty interiors. Let* $(X_1, \mathscr{A}_1, \mu_1)$ *be a probability space, and let* $\eta_1 : X_1 \to I$ *be a measurable function. Let* $(X_2, \mathscr{A}_2, \nu)$ *be a measure space with* $\nu(X_2) > 0$, *let* w *be a positive function on* X_2 *such that* $\int_{X_2} w d\nu = 1$, *let* $w_1, \ldots, w_n$ $(n \geqslant 2)$ *be positive and measurable functions on* X_2 *such that* $\sum_{i=1}^{n} w_i = w$, *and let* $\eta_2 :$

$X_2 \to J$ be a measurable function. Then for all $f \in F_{\Delta}^{co}$

$$f\left(\int\limits_{X_1}\eta_1 d\mu_1, \int\limits_{X_2}\eta_2 w d\nu\right)$$
$$\leqslant \sum_{i=1}^{n}\left(\int\limits_{X_1} f\circ\left(\alpha\int\limits_{X_1}\eta_1 d\mu_1 + (1-\alpha)\,\eta_1, \frac{\int\limits_{X_2}\eta_2 w_i d\nu}{\int\limits_{X_2} w_i d\nu}\right) d\mu_1 \cdot \int\limits_{X_2} w_i d\nu\right)$$
$$\leqslant \int\limits_{X_1\times X_2} f\circ(\eta_1,\eta_2)\, wd\,(\mu_1\times\nu),$$

where $\alpha \in [0,1]$ can be chosen arbitrarily.

Proof. By Theorem 6 (a),

$$g\left(\int\limits_{X_1}\eta_1 d\mu_1\right) \leqslant \int\limits_{X_1} g\left(\alpha\int\limits_{X_1}\eta_1 d\mu_1 + (1-\alpha)\,\eta_1\right) d\mu_1 \leqslant \int\limits_{X_1} g\circ\eta_1 d\mu_1$$

for all $g \in F_I$.

Let the probability measure μ_2 be defined on $\mathscr{A}_2$ by

$$\mu_2(A) := \int\limits_{A} w d\nu.$$

By introducing the notations

$$\omega_i := \int\limits_{X_2} w_i d\nu, \quad t_i := \frac{\int\limits_{X_2}\eta_2 w_i d\nu}{\int\limits_{X_2} w_i d\nu}, \quad i = 1,\ldots,n,$$

we define the discrete probability measure ω on $\mathscr{B}_J$ by

$$\omega := \sum_{i=1}^{n}\omega_i\cdot\varepsilon_{t_i}.$$

Theorem 6 (b) shows that

$$h\left(\int\limits_{X_2}\eta_2 w d\nu\right) \leqslant \sum_{i=1}^{n} h\left(\frac{\int\limits_{X_2}\eta_2 w_i d\nu}{\int\limits_{X_2} w_i d\nu}\right)\int\limits_{X_2} w_i d\nu \leqslant \int\limits_{X_2}(h\circ\eta_2)\, wd\nu,$$

or in another form

$$h\left(\int\limits_{X_2}\eta_2 d\mu_2\right) \leqslant \int\limits_{J} hd\omega \leqslant \int\limits_{X_2}(h\circ\eta_2)\, d\mu_2$$

for all $h \in F_J$.

Now by applying Theorem 5 (b) to $(Y_1,\mathscr{B}_1,\nu_1) := (X_1,\mathscr{A}_1,\mu_1)$, $(Y_2,\mathscr{B}_2,\nu_2) := (J,\mathscr{B}_J,\omega)$ and

$$\varphi_1 := \int_{X_1} \eta_1 d\mu_1, \quad \varphi_2 := \int_{X_2} \eta_2 d\mu_2$$

and

$$\psi_1 := \alpha \int_{X_1} \eta_1 d\mu_1 + (1-\alpha)\eta_1, \quad \psi_2 := id_J,$$

we obtain

$$f\left(\int_{X_1} \eta_1 d\mu_1, \int_{X_2} \eta_2 d\mu_2\right) = \int_{X_1\times X_2} f\left(\int_{X_1} \eta_1 d\mu_1, \int_{X_2} \eta_2 d\mu_2\right) d(\mu_1\times\mu_2)$$

$$\leqslant \int_{X_1\times J} f\circ\left(\alpha\int_{X_1} \eta_1 d\mu_1 + (1-\alpha)\eta_1, id_J\right) d(\mu_1\times\omega),$$

which is exactly the first inequality.

The same technique works for the second inequality.

The proof is complete. □

REMARK 7. We stress again that we have obtained a refinement of the integral Jensen inequality for functions convex on the coordinates in an almost elementary way from two different types of refinements of the integral Jensen inequality. Since many refinements of the integral Jensen inequality are known, by applying Theorem 5 (b) we have obtained a very efficient method for refining the integral Jensen inequality for functions convex on the coordinates. And it follows from Theorem 5 (a) that this is essentially the only option.

5. Hermite-Hadamard type inequalities for functions convex on the coordinates and their refinements

Combining our main result with either Theorem 3 or Proposition 1 or Proposition 2 or Proposition 3, we can obtain Hermite-Hadamard inequalities for functions convex on the coordinates and a refinement method for them. We formulate only the version based on Proposition 2.

THEOREM 9. *Let $\Delta := [a,b]\times[c,d]$ be an interval in $\mathbb{R}^2$, where $[a,b]$ and $[c,d]$ are compact intervals in $\mathbb{R}$ with nonempty interiors, let μ_1 be a finite measure on $\mathscr{B}_{[a,b]}$, and let μ_2 be a finite measure on $\mathscr{B}_{[c,d]}$ such that $\mu_i([a,b]) > 0$ $(i=1,2)$. Assume φ_1, φ_2, ψ_1, $\psi_2 : [a,b]\to[a,b]$ and η_1, η_2, ϑ_1, $\vartheta_2 : [c,d]\to[c,d]$ are increasing functions such that*

$$\int_{[a,x]} \psi_2 d\mu_1 \leqslant \int_{[a,x]} \psi_1 d\mu_1 \leqslant \int_{[a,x]} s\,d\mu_1(s) \leqslant \int_{[a,x]} \varphi_1 d\mu_1 \leqslant \int_{[a,x]} \varphi_2 d\mu_1, \quad x\in[a,b], \quad (23)$$

$$\int_{[a,b]} \varphi_i d\mu_1 = \int_{[a,b]} \psi_i d\mu_1 = \int_{[a,b]} s d\mu_1(s), \quad i = 1,2 \tag{24}$$

and

$$\int_{[c,y]} \eta_2 d\mu_2 \leqslant \int_{[c,y]} \eta_1 d\mu_2 \leqslant \int_{[c,y]} t d\mu_2(t) \leqslant \int_{[c,y]} \vartheta_1 d\mu_2 \leqslant \int_{[c,y]} \vartheta_2 d\mu_2, \quad y \in [c,d],$$

$$\int_{[c,d]} \eta_i d\mu_2 = \int_{[c,d]} \vartheta_i d\mu_2 = \int_{[c,d]} t d\mu_2(t), \quad i = 1,2$$

are satisfied. Then for all $f \in F_\Delta^{co}$ *we have*

$$f\left(x_{\mu_1}, y_{\mu_2}\right) \mu_1([a,b]) \mu_2([c,d]) \tag{25}$$

$$\leqslant \iint_\Delta f \circ (\varphi_2, \vartheta_2) d(\mu_1 \times \mu_2) \leqslant \iint_\Delta f \circ (\varphi_1, \vartheta_1) d(\mu_1 \times \mu_2) \leqslant \iint_\Delta f d(\mu_1 \times \mu_2) \tag{26}$$

$$\leqslant \iint_\Delta f \circ (\psi_1, \eta_1) d(\mu_1 \times \mu_2) \leqslant \iint_\Delta f \circ (\psi_2, \eta_2) d(\mu_1 \times \mu_2) \tag{27}$$

$$\leqslant \left(\frac{b - x_{\mu_1}}{b-a} \left(\frac{d - y_{\mu_2}}{d-c} f(a,d) + \frac{y_{\mu_2} - c}{d-c} f(a,c) \right) \right. \tag{28}$$

$$\left. + \frac{x_{\mu_1} - a}{b-a} \left(\frac{d - y_{\mu_2}}{d-c} f(b,d) + \frac{y_{\mu_2} - c}{d-c} f(b,c) \right) \right) \mu_1([a,b]) \mu_2([c,d]),$$

where

$$x_{\mu_1} := \frac{1}{\mu_1([a,b])} \int_{[a,b]} s d\mu_1(s), \quad y_{\mu_2} := \frac{1}{\mu_2([c,d])} \int_{[c,d]} t d\mu_2(t).$$

Proof. The first, the second, the third, the fourth and the fifth inequalities in (25–28) are immediate consequences of Proposition 2 and Theorem 5. The sixth inequality is also a consequence of these two theorems, where on the right hand side we use the discrete measure

$$\left(\frac{b - x_{\mu_1}}{b-a} \varepsilon_a + \frac{x_{\mu_1} - a}{b-a} \varepsilon_b \right) \mu_1([a,b])$$

on $\mathscr{B}_{[a,b]}$ and the discrete measure

$$\left(\frac{d - y_{\mu_2}}{d-c} \varepsilon_c + \frac{y_{\mu_2} - c}{d-c} \varepsilon_d \right) \mu_2([c,d])$$

on $\mathscr{B}_{[c,d]}$.

The proof is complete. $\square$

REMARK 8. (a) Many authors have dealt with Hermite–Hadamard type inequalities for functions convex on the coordinates and their refinements using the classical Lebesgue integral (see e.g. the papers [1], [5], [13], [14], [17] and [18]). We emphasize that our result gives the Hermite-Hadamard inequality for functions convex on the coordinates using Borel measures, and a method for refining the obtained inequality. All of the refinements in the aforementioned papers can be obtained from Theorem 9.

(b) Assume the measure μ_1 has density $u_1 : [a,b] \to [0,\infty[$ with respect to the classical Lebesgue measure on $\mathscr{B}_{[a,b]}$ and μ_2 has density $u_2 : [c,d] \to [0,\infty[$ with respect to the classical Lebesgue measure on $\mathscr{B}_{[c,d]}$. If

$$u_1(s) = u_1(a+b-s), \quad s \in [a,b]$$

and

$$u_2(t) = u_2(c+d-t), \quad t \in [c,d],$$

then the symmetry property (12) is true for both measures. This is satisfied, for example, in paper [17], where

$$u_1(s) := h_1\left(\frac{b-s}{b-a}\right) + h_1\left(\frac{s-a}{b-a}\right)$$

and

$$u_2(t) := h_2\left(\frac{d-t}{d-c}\right) + h_2\left(\frac{t-c}{d-c}\right),$$

where $h_1,\ h_2 : [0,1] \to]0,\infty[$ are Lebesgue-integrable functions.

(c) It is worth noting that from known refinements, new refinements can be obtained by different processes. We illustrate this for the functions ψ_1 and ψ_2. Define the functions $\psi_\lambda : [a,b] \to [a,b]$ $(0 \leqslant \lambda \leqslant 1)$ by

$$\psi_\lambda(t) := (1-\lambda)\psi_2(t) + \lambda\psi_1(t).$$

Then it is easy to check that for each $\lambda \in [0,1]$ the function ψ_λ is also increasing. By the first inequality in (23),

$$\int\limits_{[a,x]} \psi_2 d\mu_1 \leqslant \int\limits_{[a,x]} \psi_\lambda d\mu_1 \leqslant \int\limits_{[a,x]} \psi_1 d\mu_1, \quad x \in [a,b], \quad \lambda \in [0,1],$$

and by (24),

$$\int\limits_{[a,b]} \psi_\lambda d\mu_1 = \int\limits_{[a,b]} t d\mu_1(t), \quad \lambda \in [0,1].$$

Now, by applying Theorem 9 and the convexity of f_q $(q \in [c,d])$ we have

$$\iint\limits_{\Delta} f \circ (\psi_2, \eta_2)\, d(\mu_1 \times \mu_2) \leqslant \iint\limits_{\Delta} f \circ (\psi_\lambda, \eta_2)\, d(\mu_1 \times \mu_2)$$

$$\leqslant (1-\lambda) \iint\limits_{\Delta} f \circ (\psi_1, \eta_1)\, d(\mu_1 \times \mu_2) + \lambda \iint\limits_{\Delta} f \circ (\psi_2, \eta_1)\, d(\mu_1 \times \mu_2)$$

$$\leqslant \iint\limits_{\Delta} f \circ (\psi_2, \eta_2)\, d(\mu_1 \times \mu_2).$$

Our last results are intended to illustrate the applicability of the previous theorem.

First, by using Theorem 9, we give an extension of Theorem 2.2. in paper [13] with a simple proof.

THEOREM 10. *Let* $\Delta := [a,b] \times [c,d]$ *be an interval in* $\mathbb{R}^2$*, where* $[a,b]$ *and* $[c,d]$ *are compact intervals in* $\mathbb{R}$ *with nonempty interiors. If* $f \in F_{\Delta}^{co}$*, then*

$$f\left(\frac{a+b}{2},\frac{c+d}{2}\right) \leqslant \frac{\alpha}{b-a}\int_a^b f\left(s,\frac{c+d}{2}\right)ds + \frac{1-\alpha}{d-c}\int_c^d f\left(\frac{a+b}{2},t\right)dt$$

$$\leqslant \frac{1}{(b-a)(d-c)}\iint_{\Delta} f \leqslant \frac{\beta}{4(b-a)}\int_a^b \left(f(s,c)+f(s,d)+2f\left(s,\frac{c+d}{2}\right)\right)ds$$

$$+\frac{1-\beta}{4(d-c)}\int_c^d \left(f(a,t)+f(b,t)+2f\left(\frac{a+b}{2},t\right)\right)dt$$

$$\leqslant \frac{f(a,c)+f(b,c)+f(a,d)+f(b,d)}{16}+\frac{1}{4}f\left(\frac{a+b}{2},\frac{c+d}{2}\right)$$

$$+\frac{f\left(\frac{a+b}{2},c\right)+f\left(\frac{a+b}{2},d\right)+f\left(a,\frac{c+d}{2}\right)+f\left(b,\frac{c+d}{2}\right)}{8}$$

for all $\alpha,\ \beta \in [0,1]$.

Proof. We first show that

$$f\left(\frac{a+b}{2},\frac{c+d}{2}\right) \leqslant \frac{1}{b-a}\int_a^b f\left(s,\frac{c+d}{2}\right)ds \leqslant \frac{1}{(b-a)(d-c)}\iint_{\Delta} f \tag{29}$$

$$\leqslant \frac{1}{4(b-a)}\int_a^b \left(f(s,c)+f(s,d)+2f\left(s,\frac{c+d}{2}\right)\right)ds \tag{30}$$

$$\leqslant \frac{f(a,c)+f(b,c)+f(a,d)+f(b,d)}{16}+\frac{1}{4}f\left(\frac{a+b}{2},\frac{c+d}{2}\right) \tag{31}$$

$$+\frac{f\left(\frac{a+b}{2},c\right)+f\left(\frac{a+b}{2},d\right)+f\left(a,\frac{c+d}{2}\right)+f\left(b,\frac{c+d}{2}\right)}{8}. \tag{32}$$

Define the functions $\varphi_1,\ \psi_1,\ \psi_2 : [a,b] \to [a,b]$ by

$$\varphi_1(s) = \psi_1(s) := s, \quad \psi_2(s) := \begin{cases} a, & s \in \left[a,\frac{3a+b}{4}\right[\\ \frac{a+b}{2}, & s \in \left[\frac{3a+b}{4},\frac{a+3b}{4}\right] \\ b, & s \in \left]\frac{a+3b}{4},b\right] \end{cases},$$

and define the functions ϑ_1, η_1, $\eta_2 : [c,d] \to [c,d]$ by

$$\vartheta_1(t) := \frac{c+d}{2}, \quad \eta_1(t) = \eta_2(t) := \begin{cases} c, & t \in \left[c, \frac{3c+d}{4}\right[\\ \frac{c+d}{2}, & t \in \left[\frac{3c+d}{4}, \frac{c+3d}{4}\right] \\ d, & t \in \left]\frac{c+3d}{4}, d\right] \end{cases}.$$

Some elementary calculations show that the conditions of Theorem 9 are satisfied with these functions, and hence inequalities (29-32) follow.

In exactly the same way, we can justify easily that

$$f\left(\frac{a+b}{2}, \frac{c+d}{2}\right) \leqslant \frac{1}{d-c} \int_c^d f\left(\frac{a+b}{2}, t\right) dt \leqslant \frac{1}{(b-a)(d-c)} \iint_\Delta f \tag{33}$$

$$\leqslant \frac{1}{4(d-c)} \int_c^d \left(f(a,t) + f(b,t) + 2f\left(\frac{a+b}{2}, t\right)\right) dt \tag{34}$$

$$\leqslant \frac{f(a,c) + f(b,c) + f(a,d) + f(b,d)}{16} + \frac{1}{4} f\left(\frac{a+b}{2}, \frac{c+d}{2}\right) \tag{35}$$

$$+ \frac{f\left(\frac{a+b}{2}, c\right) + f\left(\frac{a+b}{2}, d\right) + f\left(a, \frac{c+d}{2}\right) + f\left(b, \frac{c+d}{2}\right)}{8}. \tag{36}$$

The result is a trivial consequence of inequalities (29–32) and (33–36).

The proof is complete. □

REMARK 9. Theorem 2.2. in [13] is the special case of our previous result when $\alpha = \beta = \frac{1}{2}$. The proof also illustrates the method described in Remark 8 (c).

Second, a parametric refinement of the Hermite-Hadamard inequality is obtained. It is given under the conditions of Proposition 3 for the sake of simplicity and clarity. Even in this case, it is useful to introduce a few terms before making the claim.

Let $[a,b] \subset \mathbb{R}$ be an interval with nonempty interior. We shall say that $I_1, \ldots, I_m$, $I_{m+1}, I_{m+2}, \ldots, I_{2m+1}$ $(m \geqslant 1)$ is a symmetric partition of $[a,b]$ if they are adjacent pairwise disjoint intervals with union $[a,b]$, $I_1, \ldots, I_m$ are left-closed and right open, $I_{m+1} := \left\{\frac{a+b}{2}\right\}$, and I_i is symmetric to I_{2m+2-i} with respect to the point $\frac{a+b}{2}$ $(i = 1, \ldots, m)$. We shall also say the next: the points $s_1, \ldots, s_m, s_{m+1}, s_{m+2}, \ldots, s_{2m+1}$ are generated from below by the symmetric partition just described if $s_1, \ldots, s_m$ are the left-hand endpoints of $I_1, \ldots, I_m$, $s_{m+1} := \frac{a+b}{2}$, and $s_{m+2}, \ldots, s_{2m+1}$ are the right-hand endpoints of $I_{m+2}, \ldots, I_{2m+1}$; the points $\hat{s}_1, \ldots, \hat{s}_m, \hat{s}_{m+1}, \hat{s}_{m+2}, \ldots, \hat{s}_{2m+1}$ are generated from above by the same symmetric partition if $\hat{s}_1, \ldots, \hat{s}_m$ are the right-hand endpoints of $I_1, \ldots, I_m$, $\hat{s}_{m+1} := \frac{a+b}{2}$, and $\hat{s}_{m+2}, \ldots, \hat{s}_{2m+1}$ are the left-hand endpoints of $I_{m+2}, \ldots, I_{2m+1}$ (by the symmetry of the intervals, $s_{2m+2-i} = a+b-s_i$ and $\hat{s}_{2m+2-i} = a+b-\hat{s}_i$, $i = 1, \ldots, m$).

We are now ready to give the promised result.

PROPOSITION 4. *Let* $\Delta := [a,b] \times [c,d]$ *be an interval in* $\mathbb{R}^2$, *where* $[a,b]$ *and* $[c,d]$ *are compact intervals in* $\mathbb{R}$ *with nonempty interiors, let* μ_1 *be a finite measure on* $\mathscr{B}_{[a,b]}$, *and let* μ_2 *be a finite measure on* $\mathscr{B}_{[c,d]}$ *such that* $\mu_i([a,b]) > 0$ $(i = 1,2)$ *and the symmetry property* (12) *is true for both measures. Let* $\left(I_i^k\right)_{i=1}^{2m_k+1}$ $(k = 1,2)$ *be symmetric partitions of* $[a,b]$, *and let* $\left(J_j^k\right)_{j=1}^{2n_k+1}$ $(k = 1,2)$ *be symmetric partitions of* $[c,d]$. *Assume the points* $\left(s_i^1\right)_{i=1}^{2m_1+1}$ *are generated from below by the partition* $\left(I_i^1\right)_{i=1}^{2m_1+1}$ *and* $\left(\hat{s}_i^2\right)_{i=1}^{2m_2+1}$ *are generated from above by the partition* $\left(I_i^2\right)_{i=1}^{2m_2+1}$, *and* $\left(t_j^1\right)_{j=1}^{2n_1+1}$ *are generated from below by the partition* $\left(J_j^1\right)_{j=1}^{2n_1+1}$ *and* $\left(\hat{t}_j^2\right)_{j=1}^{2n_2+1}$ *are generated from above by the partition* $\left(J_j^2\right)_{j=1}^{2n_2+1}$. *Then*

$$f\left(\frac{a+b}{2}, \frac{c+d}{2}\right) \mu_1([a,b]) \mu_2([c,d]) \leqslant \sum_{i=1}^{2m_2+1} \sum_{j=1}^{2n_2+1} f\left(\hat{s}_i^2, \hat{t}_j^2\right) \mu_1\left(I_i^2\right) \mu_2\left(J_j^2\right)$$

$$\leqslant \iint_{\Delta} f d(\mu_1 \times \mu_2) \leqslant \sum_{i=1}^{2m_1+1} \sum_{j=1}^{2n_1+1} f\left(s_i^1, t_j^1\right) \mu_1\left(I_i^1\right) \mu_2\left(J_j^1\right)$$

$$\leqslant \iint_{\Delta} f d(\nu_1 \times \nu_2) \leqslant \frac{f(a,c) + f(b,c) + f(a,d) + f(b,d)}{4} \mu_1([a,b]) \mu_2([c,d]),$$

where

$$\nu_1 := \left(\frac{\mu_1([a,b]) - \mu_1\left(\left\{\frac{a+b}{2}\right\}\right)}{2}\right)(\varepsilon_a + \varepsilon_b) + \mu_1\left(\left\{\frac{a+b}{2}\right\}\right) \varepsilon_{\frac{a+b}{2}}$$

and

$$\nu_2 := \left(\frac{\mu_2([c,d]) - \mu_2\left(\left\{\frac{c+d}{2}\right\}\right)}{2}\right)(\varepsilon_c + \varepsilon_d) + \mu_2\left(\left\{\frac{c+d}{2}\right\}\right) \varepsilon_{\frac{c+d}{2}}.$$

Proof. Define the functions ψ_1, $\varphi_1 : [a,b] \to [a,b]$ and η_1, $\vartheta_1 : [c,d] \to [c,d]$ by

$$\psi_1(s) := s_i^1 \text{ if } s \in I_i^1, \quad i = 1, \ldots 2m_1 + 1,$$

$$\varphi_1(s) := \hat{s}_i^2 \text{ if } s \in I_i^2, \quad i = 1, \ldots 2n_1 + 1,$$

$$\eta_1(t) := t_i^1 \text{ if } t \in J_j^1, \quad j = 1, \ldots 2m_2 + 1,$$

and

$$\vartheta_1(t) := \hat{t}_i^2 \text{ if } t \in J_j^2, \quad j = 1, \ldots 2n_2 + 1.$$

Then φ_1, ψ_1, η_1 and ϑ_1 are increasing functions such that ψ_1 and η_1 satisfy

$$\psi_1(a+b-s) = a+b-\psi_1(s), \quad \varphi_1(a+b-s) = a+b-\varphi_1(s), \quad s \in [a,b],$$

and

$$\psi_1(s) \leqslant s \leqslant \varphi_1(s), \quad s \in \left[a, \frac{a+b}{2}\right]$$

while φ_1 and ϑ_1 satisfy

$$\eta_1(c+d-t) = c+d-\eta_1(t), \quad \vartheta_1(c+d-t) = c+d-\vartheta_1(t), \quad t \in [c,d]$$

and

$$\eta_1(t) \leqslant t \leqslant \vartheta_1(t), \quad t \in \left[c, \frac{c+d}{2}\right].$$

The result now follows from Theorem 9, taking Theorem 3 into account.

The proof is complete. □

REMARK 10. We mention one refinement, which is a special case of our previous result: Theorem 2.5 of [17].

6. Application to a new functional corresponding to f-divergence functional

The following notion was introduced by Csiszár in [2] and [3].

DEFINITION 2. Let $g:]0,\infty[\to]0,\infty[$ be a convex function, and let $\mathbf{p} := (p_1, \ldots, p_n)$ and $\mathbf{q} := (q_1, \ldots, q_n)$ be positive probability distributions. The g-divergence functional is

$$I_g(\mathbf{p},\mathbf{q}) := \sum_{i=1}^{n} q_i g\left(\frac{p_i}{q_i}\right).$$

It is possible to use nonnegative probability distributions in the g-divergence functional, by defining

$$g(0) := \lim_{t \to 0+} g(t); \quad 0g\left(\frac{0}{0}\right) := 0; \quad 0g\left(\frac{a}{0}\right) := \lim_{t \to 0+} tg\left(\frac{a}{t}\right), \quad a > 0.$$

The basic inequality (which comes from the discrete Jensen inequality)

$$I_g(\mathbf{p},\mathbf{q}) \geqslant g(1) \tag{37}$$

is one of the key properties of g-divergences.

The refinement of inequality (37) is the subject of several papers (for a non-exhaustive list, see book [12] and references therein, and papers [4], [9], [10] and [11]).

Starting from the concept of g-divergence, we introduce the following quantity, in which we use a function convex on the coordinates instead of a convex function. For clarity, only positive probability distributions are considered in this section.

DEFINITION 3. Let $o \in \{1,2\}$. Let $\mathbf{p}_o := \left(p_1^o, \ldots, p_{n_o}^o\right)$, $\mathbf{q}_o := \left(q_1^o, \ldots, q_{n_o}^o\right)$ be positive probability distributions for some $n_o \geqslant 1$, and let $f:]0,\infty[\times]0,\infty[\to]0,\infty[$ be a function convex on the coordinates. We introduce the following functional

$$I_f(\mathbf{p}_1,\mathbf{p}_2,\mathbf{q}_1,\mathbf{q}_2) := \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} q_i^1 q_j^2 f\left(\frac{p_i^1}{q_i^1}, \frac{p_j^2}{q_j^2}\right).$$

We show below that the basic properties of g-divergence (see [19]) are also satisfied for the introduced quantity $I_f(\mathbf{p}_1,\mathbf{p}_2,\mathbf{q}_1,\mathbf{q}_2)$.

First, we formulate the analogue of inequality (37) for $I_f(\mathbf{p}_1,\mathbf{p}_2,\mathbf{q}_1,\mathbf{q}_2)$. This, like inequality (37), is almost obvious, we just need to apply the integral Jensen inequality for functions convex on the coordinates.

PROPOSITION 5. *Let $o\in\{1,2\}$. Let $\mathbf{p}_o:=\left(p_1^o,\ldots,p_{n_o}^o\right)$, $\mathbf{q}_o:=\left(q_1^o,\ldots,q_{n_o}^o\right)$ be positive probability distributions for some $n_o\geqslant 1$. Then for every function $f:\,]0,\infty[\times]0,\infty[\,\to\,]0,\infty[$ convex on the coordinates inequality*

$$I_f(\mathbf{p}_1,\mathbf{p}_2,\mathbf{q}_1,\mathbf{q}_2)\geqslant f(1,1) \tag{38}$$

holds.

Proof. Define the probability measures μ_1 and μ_2 on $\mathscr{B}_{]0,\infty[}$ by

$$\mu_1:=\sum_{i=1}^{n_1}q_i^1\varepsilon_{p_i^1/q_i^1}\text{ and }\mu_2:=\sum_{j=1}^{n_2}q_j^2\varepsilon_{p_j^2/q_j^2},$$

and let $\eta_1,\ \eta_2:\,]0,\infty[\,\to\,]0,\infty[$, $\eta_1(t)=\eta_2(t):=t$.

By Theorem 7,

$$I_f(\mathbf{p}_1,\mathbf{p}_2,\mathbf{q}_1,\mathbf{q}_2)=\int\limits_{]0,\infty[\times]0,\infty[}f\circ(\eta_1,\eta_2)\,d(\mu_1\times\mu_2)$$

$$\geqslant f\left(\int\limits_{]0,\infty[}\eta_1d\mu_1,\ \int\limits_{]0,\infty[}\eta_2d\mu_2\right)=f(1,1).$$

The proof is complete. $\square$

REMARK 11. Theorem 7 can be applied since there are compact intervals $I,\ J\subset\,]0,\infty[$ such that

$$\frac{p_i^1}{q_i^1}\in I,\quad i=1,\ldots,n_1\quad\text{and}\quad\frac{p_j^2}{q_j^2}\in J,\quad j=1,\ldots,n_2.$$

The next basic property of g-divergence corresponds to the perspective of g which is defined by

$$g^*:\,]0,\infty[\,\to\,]0,\infty[,\quad g^*(t):=tg\left(\frac{1}{t}\right).$$

It is well known that g^* is also convex and

$$I_{g^*}(\mathbf{q},\mathbf{p})=I_g(\mathbf{p},\mathbf{q}). \tag{39}$$

To formulate the equivalent of property (39) for the introduced quantity $I_f(\mathbf{p}_1,\mathbf{p}_2,\mathbf{q}_1,\mathbf{q}_2)$, we need to define the perspective of a function convex on the coordinates.

DEFINITION 4. Let $f :]0,\infty[\times]0,\infty[\to]0,\infty[$ be a function convex on the coordinates, and define the perspective of f by

$$f^* :]0,\infty[\times]0,\infty[\to]0,\infty[, \quad f^*(s,t) := st \cdot f\left(\frac{1}{s}, \frac{1}{t}\right).$$

We can then make the following statement.

PROPOSITION 6. *(a) If $f :]0,\infty[\times]0,\infty[\to]0,\infty[$ is a function convex on the coordinates, then f^* is also a function convex on the coordinates.*

(b) Let $o \in \{1,2\}$. Let $\mathbf{p}_o := \left(p_1^o, \ldots, p_{n_o}^o\right)$, $\mathbf{q}_o := \left(q_1^o, \ldots, q_{n_o}^o\right)$ be positive probability distributions for some $n_o \geqslant 1$. Then for every function $f :]0,\infty[\times]0,\infty[\to]0,\infty[$ convex on the coordinates

$$I_{f^*}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{p}_1, \mathbf{p}_2) = I_f(\mathbf{p}_1, \mathbf{p}_2, \mathbf{q}_1, \mathbf{q}_2).$$

Proof. (a) Consider the partial mappings $f_{1/q} :]0,\infty[\to]0,\infty[$, $f_{1/q}(s) := f\left(s, \frac{1}{q}\right)$ for all $q \in]0,\infty[$ and $f_{1/p} :]0,\infty[\to]0,\infty[$, $f_{1/p}(t) := f\left(\frac{1}{p}, t\right)$ for all $p \in]0,\infty[$.

Let $q \in]0,\infty[$ be fixed. Since $f_{1/q}$ is convex, the perspective of $f_{1/q}$

$$f_{1/q}^*(s) = sf\left(\frac{1}{s}, \frac{1}{q}\right), \quad s \in]0,\infty[$$

is also convex, and therefore the convexity of

$$f^*(s,q) = qf_{1/q}^*(s), \quad s \in]0,\infty[$$

follows.

We can prove similarly that the function

$$f^*(p,t) = pf_{1/p}^*(t), \quad t \in]0,\infty[$$

is convex for all $p \in]0,\infty[$.

(b) It can be obtained by elementary calculation that

$$I_{f^*}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{p}_1, \mathbf{p}_2) = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} p_i^1 p_j^2 \frac{q_i^1}{p_i^1} \frac{q_j^2}{p_j^2} f\left(\frac{1}{\frac{q_i^1}{p_i^1}}, \frac{1}{\frac{q_j^2}{p_j^2}}\right) = I_f(\mathbf{p}_1, \mathbf{p}_2, \mathbf{q}_1, \mathbf{q}_2).$$

The proof is complete. $\square$

The third important property of g-divergence is monotonicity. This means the next: Let $(A_i)_{i=1}^m$ be pairwise disjoint subsets of $\{1,\ldots,n\}$ with $A_i \neq \varnothing$ for all $i = 1,\ldots,m$ and $\bigcup_{i=1}^m A_i = \{1,\ldots,n\}$, and let $P_i := \sum_{j \in A_i} p_j$ and $Q_i := \sum_{j \in A_i} q_j$ $(i = 1,\ldots,m)$. Then

$$I_g(\mathbf{P}, \mathbf{Q}) \leqslant I_g(\mathbf{p}, \mathbf{q}),$$

where $\mathbf{P} := (P_1, \ldots, P_m)$ and $\mathbf{Q} := (Q_1, \ldots, Q_m)$.

The monotonicity can be formulated as follows for $I_f(\mathbf{p}_1, \mathbf{p}_2, \mathbf{q}_1, \mathbf{q}_2)$:

PROPOSITION 7. *Let $o \in \{1,2\}$. Let $\mathbf{p}_o := \left(p_1^o, \dots, p_{n_o}^o\right)$, $\mathbf{q}_o := \left(q_1^o, \dots, q_{n_o}^o\right)$ be positive probability distributions for some $n_o \geqslant 1$. Let $\left(A_k^1\right)_{k=1}^{m_1}$ be pairwise disjoint subsets of $\{1, \dots, n_1\}$ with $A_k^1 \neq \varnothing$ for all $k = 1, \dots, m_1$ and $\bigcup\limits_{k=1}^{m_1} A_k^1 = \{1, \dots, n_1\}$, and let $\left(A_l^2\right)_{l=1}^{m_2}$ be pairwise disjoint subsets of $\{1, \dots, n_2\}$ with $A_l^2 \neq \varnothing$ for all $l = 1, \dots, m_2$ and $\bigcup\limits_{l=1}^{m_2} A_l^2 = \{1, \dots, n_2\}$. Define $P_k^1 := \sum\limits_{i \in A_k^1} p_i^1$, $Q_k^1 := \sum\limits_{i \in A_k^1} q_i^1$ $(k = 1, \dots, m_1)$, and $P_l^2 := \sum\limits_{j \in A_l^2} p_j^2$, $Q_l^2 := \sum\limits_{j \in A_l^2} q_j^2$ $(l = 1, \dots, m_2)$. Then for every function $f :]0,\infty[\times]0,\infty[\to]0,\infty[$ convex on the coordinates*

$$I_f(\mathbf{P}_1, \mathbf{P}_2, \mathbf{Q}_1, \mathbf{Q}_2) \leqslant I_f(\mathbf{p}_1, \mathbf{p}_2, \mathbf{q}_1, \mathbf{q}_2),$$

where $\mathbf{P}_o := \left(P_1^o, \dots, P_{m_o}^o\right)$ and $\mathbf{Q}_o := \left(Q_1^o, \dots, Q_{m_o}^o\right)$.

Proof. Since

$$I_f(\mathbf{P}_1, \mathbf{P}_2, \mathbf{Q}_1, \mathbf{Q}_2) = \sum_{k=1}^{m_1} \sum_{l=1}^{m_2} Q_k^1 Q_l^2 f\left(\frac{P_k^1}{Q_k^1}, \frac{P_l^2}{Q_l^2}\right),$$

by applying Lemma 4 to each member of the sum, we obtain the statement.

The proof is complete. $\square$

REMARK 12. Let $o \in \{1,2\}$. Let $\mathbf{p}_o := \left(p_1^o, \dots, p_{n_o}^o\right)$, $\mathbf{q}_o := \left(q_1^o, \dots, q_{n_o}^o\right)$ be positive probability distributions for some $n_o \geqslant 1$, and let $g :]0,\infty[\to]0,\infty[$ be a convex function. The products of the probability distributions $\mathbf{p}_1$ and $\mathbf{p}_2$ and $\mathbf{q}_1$ and $\mathbf{q}_2$ are the probability distributions

$$\mathbf{p}_1 \otimes \mathbf{p}_2 = \left(p_i^1 p_j^2\right)_{i \in \{1,\dots,n_1\}}^{j \in \{1,\dots,n_2\}} \text{ and } \mathbf{q}_1 \otimes \mathbf{q}_2 = \left(q_i^1 q_j^2\right)_{i \in \{1,\dots,n_1\}}^{j \in \{1,\dots,n_2\}},$$

respectively.

If the function $f :]0,\infty[\times]0,\infty[\to]0,\infty[$ is defined by $f(s,t) := g(st)$, then it is convex on the coordinates and

$$I_f(\mathbf{p}_1, \mathbf{p}_2, \mathbf{q}_1, \mathbf{q}_2) = I_g(\mathbf{p}_1 \otimes \mathbf{p}_2, \mathbf{q}_1 \otimes \mathbf{q}_2),$$

which shows that the new quantity $I_f(\mathbf{p}_1, \mathbf{p}_2, \mathbf{q}_1, \mathbf{q}_2)$ contains the g-divergence.

This, Proposition 5, Proposition 6 and Proposition 7 show that the quantity $I_f(\mathbf{p}_1, \mathbf{p}_2, \mathbf{q}_1, \mathbf{q}_2)$ can be seen as a generalization of the g-divergence.

Finally, as another application of our main result we give a necessary and sufficient condition for the inequality

$$I_f(\mathbf{p}_1, \mathbf{p}_2, \mathbf{q}_1, \mathbf{q}_2) \geqslant I_f(\mathbf{u}_1, \mathbf{u}_2, \mathbf{v}_1, \mathbf{v}_2)$$

to be satisfied, so we obtain a necessary and sufficient condition for refining inequality (38) by another divergence.

THEOREM 11. *Let $o\in\{1,2\}$. Let $X_o:=\{1,\dots,n_o\}$ for some $n_o\geqslant 1$, and let $Y_o:=\{1,\dots,m_o\}$ for some $m_o\geqslant 1$. Let $\mathbf{p}_o:=\left(p_1^o,\dots,p_{n_o}^o\right)$, $\mathbf{q}_o:=\left(q_1^o,\dots,q_{n_o}^o\right)$, $\mathbf{u}_o:=\left(u_1^o,\dots,u_{m_o}^o\right)$ and $\mathbf{v}_o:=\left(v_1^o,\dots,v_{m_o}^o\right)$ be positive probability distributions. Let $c_1^o>c_2^o>\ldots>c_{k_o}^o$ be the different elements of $\left(\frac{p_i^o}{q_i^o}\right)_{i=1}^{n_o}$ and $\left(\frac{u_i^o}{v_i^o}\right)_{i=1}^{m_o}$ in decreasing order $(1\leqslant k_o\leqslant m_o+n_o)$. For every function $f:\,]0,\infty[\times]0,\infty[\to]0,\infty[$ convex on the coordinates inequality*

$$I_f(\mathbf{p}_1,\mathbf{p}_2,\mathbf{q}_1,\mathbf{q}_2)=\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}q_i^1q_j^2f\left(\frac{p_i^1}{q_i^1},\frac{p_j^2}{q_j^2}\right)$$

$$\geqslant\sum_{i=1}^{m_1}\sum_{j=1}^{m_2}v_i^1v_j^2f\left(\frac{u_i^1}{v_i^1},\frac{u_j^2}{v_j^2}\right)=I_f(\mathbf{u}_1,\mathbf{u}_2,\mathbf{v}_1,\mathbf{v}_2)\tag{40}$$

holds if and only if

$$\sum_{\left\{j\in Y_o\mid\frac{u_j^o}{v_j^o}\geqslant c_l^o\right\}}u_j^o-\sum_{\left\{i\in X_o\mid\frac{p_i^o}{q_i^o}\geqslant c_l^o\right\}}p_i^o$$

$$\leqslant c_l^o\left(\sum_{\left\{j\in Y_o\mid\frac{u_j^o}{v_j^o}\geqslant c_l^o\right\}}v_j^o-\sum_{\left\{i\in X_o\mid\frac{p_i^o}{q_i^o}\geqslant c_l^o\right\}}q_i^o\right),\quad l=1,\dots,k_o,\quad o=1,2$$

are satisfied.

Proof. Define the probability measures μ_1, μ_2, ν_1 and ν_2 on $\mathscr{B}_{]0,\infty[}$ by

$$\mu_1:=\sum_{i=1}^{n_1}q_i^1\varepsilon_{p_i^1/q_i^1},\quad \mu_2:=\sum_{j=1}^{n_2}q_j^2\varepsilon_{p_j^2/q_j^2}$$

and

$$\nu_1:=\sum_{i=1}^{m_1}v_i^1\varepsilon_{u_i^1/v_i^1},\quad \nu_2:=\sum_{j=1}^{m_2}v_j^2\varepsilon_{u_j^2/v_j^2},$$

and let φ_1, φ_2, ψ_1, $\psi_2:\,]0,\infty[\to]0,\infty[$, $\varphi_1(t)=\varphi_2(t)=\psi_1(t)=\psi_2(t):=t$.

With these notations inequality (40) is equivalent to the following integral inequality:

$$\int_{]0,\infty[\times]0,\infty[}f\circ(\varphi_1,\varphi_2)\,d(\mu_1\times\mu_2)\geqslant\int_{]0,\infty[\times]0,\infty[}f\circ(\psi_1,\psi_2)\,d(\nu_1\times\nu_2).\tag{41}$$

We first show that (41) holds for every function $f:\,]0,\infty[\times]0,\infty[\to]0,\infty[$ convex on the coordinates if and only if

$$\int_{]0,\infty[}g\circ\varphi_o d\mu_o\geqslant\int_{]0,\infty[}g\circ\psi_o d\nu_o,\quad o=1,2\tag{42}$$

are satisfied for all $g \in F_{]0,\infty[}$.

Assume (42) holds for all $g \in F_{]0,\infty[}$. Since there are compact intervals I, $J \subset]0,\infty[$ such that

$$\frac{p_i^1}{q_i^1} \in I^\circ, \quad i = 1,\dots,n_1 \quad \text{and} \quad \frac{p_j^2}{q_j^2} \in J^\circ, \quad j = 1,\dots,n_2,$$

and

$$\frac{u_i^1}{v_i^1} \in I^\circ, \quad i = 1,\dots,m_1 \quad \text{and} \quad \frac{u_j^2}{q_j^2} \in J^\circ, \quad j = 1,\dots,m_2,$$

inequalities

$$\int_I g \circ \varphi_1 d\mu_1 \leqslant \int_I g \circ \psi_1 d\nu_1 \text{ and } \int_J h \circ \varphi_2 d\mu_2 \leqslant \int_J g \circ \psi_2 d\nu_2$$

are also satisfied for all $g \in F_I$ and $h \in F_J$. Then by Theorem 5 (b), inequality

$$\int_{I\times J} f \circ (\varphi_1, \varphi_2)\, d(\mu_1 \times \mu_2) \geqslant \int_{I\times J} f \circ (\psi_1, \psi_2)\, d(\nu_1 \times \nu_2)$$

holds for every function $f : I \times J \to]0,\infty[$ convex on the coordinates, and therefore (41) is true for every function $f :]0,\infty[\times]0,\infty[\to]0,\infty[$ convex on the coordinates.

Conversely, assume (41) holds for every function $f :]0,\infty[\times]0,\infty[\to]0,\infty[$ convex on the coordinates. Then, copying the proof of Theorem 5 (a), we can show that (42) holds for all $g \in F_{]0,\infty[}$.

Since (42) is equivalent to

$$\sum_{i=1}^{n_o} q_i^o g\left(\frac{p_i^o}{q_i^o}\right) \geqslant \sum_{j=1}^{m_o} v_j^o g\left(\frac{u_j^o}{v_j^o}\right), \quad o = 1,2,$$

the result follows from Theorem 4.

The proof is complete. □

REMARK 13. The previous result generalizes Theorem 10 in the recent paper [8].

REFERENCES

[1] H. BUDAK, F. USTA AND M. Z. SARIKAYA, *Refinements of the Hermite–Hadamard inequality for co-ordinated convex mappings*, J. Appl. Anal. **25** (1) (2019) 73–81.

[2] I. CSISZÁR, *Information measures: A critical survey*, in Trans. 7th Prague Conf. on Info. Th., Statist. Decis. Funct., Random Processes and 8th European Meeting of Statist., vol. B, pp. 73–86, Academia Prague 1978.

[3] I. CSISZÁR, *Information-type measures of difference of probability distributions and indirect observations*, Studia Sci. Math. Hungar **2** (1967) 299–318.

[4] S. S. DRAGOMIR, *A new refinement of Jensen's inequality in linear spaces with applications*, Math. Comput. Modelling **52** (2010) 1497–1505.

[5] S. S. DRAGOMIR, *On the Hadamard's inequality for convex functions on the co-ordinates in a rectangle from the plane*, Taiwanese J. Math. **5** (4) (2001) 775–778.

[6] L. HORVÁTH, *Integral inequalities using signed measures corresponding to majorization*, Rev. Real Acad. Cienc. Exactas Fis. Nat. Ser. A-Mat. 117, **80** (2) (2023).
[7] L. HORVÁTH, *Refinements of the integral Jensen's inequality generated by finite or infinite permutations*, J. Inequal. Appl. (2021) 2021:12 pp. 14.
[8] L. HORVÁTH, *Uniform treatment of integral majorization inequalities with applications to Hermite-Hadamard-Fejér type inequalities and f-divergences*, Entropy 2023 **25** (6), 954.
[9] L. HORVÁTH, D. PEČARIĆ AND J. PEČARIĆ, *Estimations of f- and Rényi divergences by using a cyclic refinement of the Jensen's inequality*, Bull. Malays. Math. Sci. Soc. **42** (2019) 933–946.
[10] M. A. KHAN, Z. M. AL-SAHWI AND Y. M. CHU, *New estimations for Shannon and Zipf–Mandelbrot entropies*, Entropy 2018 **20** (8), 608.
[11] M. A. KHAN, F. FAISAL AND S. KHAN, *Estimation of Jensen's gap through an integral identity with applications to divergence*, Innov. J. Math. **1** (2022) 99–114.
[12] M. A. KHAN, K. A. KHAN, D. PEČARIĆ AND J. PEČARIĆ, *Some New Improvements of Jensen's Inequality, Jensen's Type Inequalities in Information Theory*, Element, Zagreb, 2020. pp. 1–148.
[13] M. KLARIČIĆ BAKULA, *An improvement of the Hermite-Hadamard inequality for functions convex on the coordinates*, Aust. J. Math. Anal. Appl. **11** (1) (2014) 1–7.
[14] M. KLARIČIĆ BAKULA AND J. PEČARIĆ, *On the Jensen's inequality for convex functions on the co-ordinates in a rectangle from the plane*, Taiwanese J. Math. **10** (5) (2006) 1271–1292.
[15] C. NICULESCU AND L. E. PERSSON, *Convex Functions and Their Applications*, A Contemporary Approach (Springer, Berlin, 2006).
[16] Z. PAVIĆ, *Improvements of the Hermite–Hadamard inequality*, J. Inequal. Appl. (2015) 2015:222, pp. 11.
[17] M. Z. SARIKAYA AND D. KILIÇER, *On the extension of Hermite-Hadamard type inequalities for coordinated convex mappings*, Turkish J. Math. **45** (6) (2021) Article 23.
[18] KAI-CHEN SHU, *Refinements of Hermite-Hadamard type inequalities for differentiable co-ordinated convex functions and applications*, Taiwanese J. Math. **19** (1) (2015) 133–157.
[19] I. VAJDA, *On metric divergences of probability measures*, Kybernetika **45** (6) (2009) 885–900.
[20] J. M. VILORIA AND M. VIVAS-CORTEZ, *Jensen's inequality for convex functions on N-coordinates*, Appl. Math. Inf. Sci. **12** (5) (2018) 1–5.
[21] G. ZABANDAN AND A. KILIÇMAN, *A new version of Jensen's inequality and related results*, J. Inequal Appl. (2012) 2012:238, pp. 7.

(Received August 6, 2023)

László Horváth
Department of Mathematics
University of Pannonia
Egyetem u. 10., 8200 Veszprém, Hungary
e-mail: horvath.laszlo@mik.uni-pannon.hu