Received 21/01/13

Journal of Nonlinear and Convex Analysis
Volume 14, Number , 2013, 1–

DIRECT AND CONVERSE RESULTS FOR GENERALIZED MULTIVARIATE JENSEN-TYPE INEQUALITIES

ALLAL GUESSAB

Abstract. Mercer's 'a variant of Jensen's inequality' for functions of one variable in [22] is shown to be a special case of a refinement of Jensen's inequality, available for any multivariate convex function defined on a *convex* polytope Ω in the d-dimensional Euclidean space. In addition, we also examine the converse inequality for Mercer's result under appropriate conditions. The key to prove these results is two general composition formulae obtained for a class of linear approximation operators, that are *nonnegative* for *nonnegative affine* functions (Theorem 4.1 and 4.2). Moreover, as a result, we may be able to provide similar inequalities but with an extra term which makes them tighter for smooth (nonconvex twice continuously differentiable) functions (Theorem 6.1). We will show with examples that by following this approach we may consequently obtain direct and converse some important inequalities. Thus the present study unifies and extends a number of Jensen-type inequalities existing in the literature.

1. Introduction and motivation for the problem

Our motivation for the problem solved in this paper arose from a recent paper by Mercer [22, Theorem 1.2], which is connected with a remarkable variant of the classical Jensen's inequality for convex functions of one variable. His result says: if f is a real convex function on an interval containing numbers x_i for $i = 0, \dots, n$, and $0 < a \leq x_0 \leq x_1 \leq \cdots \leq x_n \leq b$, then

$$f\left(a + b - \sum_{i=0}^{n} w_i x_i\right) \leq f(a) + f(b) - \sum_{i=0}^{n} w_i f(x_i), \tag{1.1}$$

where $\sum_{i=0}^{n} w_i = 1$ with $w_i > 0$.
Recently, there has been considerable interest to look for refined inequalities of the type (1.1). For more details, we refer the interested reader to [1, 5, 8, 17, 18, 19, 20] and the references therein, which have been devoted to generalizations, refinements and applications of (1.1).
It is the purpose of this paper to point out that this interesting variant of Jensen's inequality, properly interpreted, holds in a most general form for any multivariate convex function defined on a convex polytope in the d-dimensional Euclidean space, thus extending previous results presented in [1, 5, 8, 19] for functions of one variable. In addition, we also examine the converse inequality for Mercer's result under appropriate conditions. Afterward as a result, we may be able to provide an extra

2010 *Mathematics Subject Classification.* Primary 41A36, 41A63, 41A80, 47A58; Secondary 26B25, 47B65.
Key words and phrases. Barycentric coordinates, convexity, nonnegative linear operators, Jensen's inequality, polytopes, pointwise inequality, linear precision, multivariate approximation, partition of unity.

term which makes the inequalities tighter for smooth (nonconvex twice continuously differentiable) functions.
The key to prove these results is two general composition formulae available for a class of linear approximation operators, that are *nonnegative* for *nonnegative affine* functions (see Theorem 4.1 and 4.2). On account of the most general nature of the linear operators used in our composition formulae, the main results of this paper are unified in nature and, by suitably specializing the coefficients and the parameters in these formulae, capable of yielding many Jensen-type inequalities existing in the literature, or generate new ones. We illustrate this point by significant examples in Section 5.
Four of main tools to handle the multidimensional case are the supporting hyperplane theorem, Farkas' lemma, the *continuous* barycentric coordinates for convex polytopes and some basic notions and results from convex analysis.

2. Generalized Barycentric Coordinates on polytopes

Since perhaps not every reader of this paper is familiar with these coordinates, we wish to give a brief overview of the basic elements of barycentric coordinates in d dimensions, see, e. g., [15, pp. 132-135] for more details. Throughout this paper, we assume that d is a positive integer. Let us quickly recall how these so-called coordinates are defined. Fix an integer $n \geq 1$ and let $W := \{\boldsymbol{x}_0, \dots, \boldsymbol{x}_n\}$ be a finite subset of distinct but otherwise arbitrary points in R^d. The following linear combination,

$$\boldsymbol{b} = \sum_{i=0}^{n} \alpha_i \boldsymbol{x}_i \tag{2.1}$$

is called a convex combination if the coefficients α_i are all nonnegative. All convex combinations of points of the set W define the convex hull of the set W. The resulting set is a convex set $conv(W)$, i. e., the smaller convex set containing W. Following the terminology of [26], a *convex* polytope Ω, or simply a polytope, we mean a set which is the convex hull of a non-empty finite set of points $W \subset \mathrm{R}^d$.

From now on let $\Omega \subset \mathrm{R}^d$ be a (convex) polytope generated from a finite subset of points in R^d, $W := \{\boldsymbol{x}_0, \dots, \boldsymbol{x}_n\}$, i. e. $\Omega = conv(W)$.
A vector $\boldsymbol{x} \in \mathrm{R}^d$ is an extreme point of Ω if $\boldsymbol{x} \in \Omega$ and $\boldsymbol{x}$ cannot be expressed as a convex combination of two vectors of Ω, both of which are different from $\boldsymbol{x}$. The set of extreme points of the polytope Ω shall be denoted by $Vert(\Omega)$. It is well know that the convex hull of a finite set W is compact, and its set of extreme points is nonempty and included in W. That is, $Vert(\Omega) \neq \emptyset$ and $Vert(\Omega) \subset W$. We assume throughout this paper that the number of vertices of Ω is greater than 2.
Introduced by Möbius in 1827 as mass points to define a coordinate-free geometry [23], barycentric coordinates over polytopes are a very common tool in many computations, has many useful applications, ranging from Gouraud and Phong shading, rendering of quadrilaterals, image warping, mesh deformation and finite element applications, see, e. g., [14, 27]. Given a polytope $\Omega = conv(\{\boldsymbol{x}_0, \dots, \boldsymbol{x}_n\})$, we wish to construct one coordinate function $\lambda_i(\boldsymbol{x})$ per point $\boldsymbol{x}_i$ for all $\boldsymbol{x} \in \Omega$. These functions

are called barycentric coordinates with respect to $\{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_n\}$ (or Ω) if they satisfy three properties. First, the coordinate functions are nonnegative on Ω,

$$\lambda_i(\boldsymbol{x}) \geq 0 \tag{2.2}$$

for all $\boldsymbol{x} \in \Omega$. Second, the functions form a partition of unity, which means that the equation

$$\sum_{i=0}^{n} \lambda_i(\boldsymbol{x}) = 1 \tag{2.3}$$

is obtained for all $\boldsymbol{x} \in \Omega$. Finally, the functions act as coordinates in that, given a value of $\boldsymbol{x}$, weighting each point $\boldsymbol{x}_i$ by $\lambda_i(\boldsymbol{x})$ returns back $\boldsymbol{x}$, i.e.,

$$\boldsymbol{x} = \sum_{i=0}^{n} \lambda_i(\boldsymbol{x})\boldsymbol{x}_i. \tag{2.4}$$

This last property is also sometimes referred to as linear precision since the coordinate functions can reproduce linear functions. For most potential applications, it is also preferable that these coordinate functions are as smooth as possible. Constructing the barycentric coordinates of a point $\boldsymbol{x}$ with respect to some given points in a polytope Ω is often not a trivial task. If Ω is a simplex, then $n = d$, (e.g., a triangle in $2D$ or a tetrahedron in $3D$), with vertices $\boldsymbol{x}_0, \ldots, \boldsymbol{x}_d \in \mathrm{R}^d$ that are affinely independent, then each point $\boldsymbol{x}$ of their convex hull Ω has a (unique) representation, that is there exist unique nonnegative real numbers $\{\lambda_i, i = 0, \ldots, d\}$ so that $\sum_{i=0}^{d} \lambda_i(\boldsymbol{x}) = 1$, and $\boldsymbol{x} = \sum_{i=0}^{d} \lambda_i(\boldsymbol{x})\boldsymbol{x}_i$. The barycentric coordinates $\lambda_0, \ldots, \lambda_d$ are nonnegative affine functions (linear polynomials) on Ω, see [7, p. 288]. Note that a d-simplex is a special polytope given as the convex hull of $d + 1$ vertices in d dimensions, each pair of which is joined by an edge. For $n > d$, which is the case of interest in this paper, the linear constraints form an under-determined system. Barycentric coordinates also exist for more general types of polytopes, and will be a crucial ingredient in what follows. Indeed, we have, see [16, Theorem 2]:

Theorem 2.1. *Let $W = \{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_n\}$ be a set of finite points of R^d and let the polytope $\Omega = conv(W)$. Then there exist nonnegative real-valued continuous functions $\lambda_0, \lambda_1, \ldots, \lambda_n$ defined on Ω such that*

$$\boldsymbol{x} = \sum_{i=0}^{n} \lambda_i(\boldsymbol{x})\boldsymbol{x}_i \quad \text{and} \quad \sum_{i=0}^{n} \lambda_i(\boldsymbol{x}) = 1 \tag{2.5}$$

for each $\boldsymbol{x} \in \Omega$.

Thus, from now on, it proves useful to work with barycentric coordinates. Therefore, unless otherwise indicated, throughout the paper it is assumed that $\lambda_i(\boldsymbol{x}), i = 0, \ldots, n$, are the barycentric coordinates of $\boldsymbol{x}$ with respect to a set of finite fixed points $\{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_n\}$ of the polytope

$$\Omega = conv(\{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_n\}).$$

We shall not always trouble to repeat this at each stage. Furthermore, they need not be the vertices of Ω, of course, the polytope Ω may be generated by another different set of points $\{\boldsymbol{y}_0, \ldots, \boldsymbol{y}_k\}$ on Ω.

3. Some properties of a class of linear operators

In this section we present a general class of linear operators, which are crucial tool in our context and appear in natural examples. We investigate some of their general properties, that will be summarized in Theorem 3.1. We also discuss the conditions that we impose them. Although the nonnegative restriction condition for affine functions may seem too restrictive, it is often satisfied by many approximation operators, see [9, 10, 12, 13].

For the sake of completeness, we recall some frequently used notions and definitions for nonnegative linear operators. Let $C(\Omega)$ be the set of all real-valued continuous functions in the polytope Ω. The class of all linear operators that map $C(\Omega)$ into itself will be denoted by $\mathcal{L}(C(\Omega))$. For the analysis of finite element it is natural and sometimes instrumental to build up approximation operators, which are also nonnegative, at least for some nonnegative elementary functions. The vector space of affine functions $a : \Omega \to \mathrm{R}$, denoted $A(\Omega)$, is isomorphic to R^{n+1}. The cone of all nonnegative real-valued affine functions on Ω is denoted $A_+(\Omega)$, we refer to [4, Section 2.6] for a detailed analysis of such a cone. $\mathcal{L}_+(A_+(\Omega))$ will design the cone of all elements of $\mathcal{L}(C(\Omega))$ that are *nonnegative* for *nonnegative affine* functions. In other words, an operator from $\mathcal{L}_+(A_+(\Omega))$ is guaranteed to take nonnegative values for nonnegative *affine* functions. We are able to completely characterize such operators, see Theorem 3.2.

To describe some fundamental properties of $\mathcal{L}_+(A_+(\Omega))$, we require a bit more material. By e_i, we denote the ith projection

$$e_i \,:\, \boldsymbol{x} = (x_1, \ldots, x_d) \longmapsto x_i$$

and write $\boldsymbol{e} := (e_1, \ldots, e_d)$ for the identity on R^d, that is,

$$\boldsymbol{e}(x_1, \ldots, x_d) \;=\; (x_1, \ldots, x_d).$$

When $L \in \mathcal{L}(C(\Omega))$ and

$$\boldsymbol{f} = (f_1, \ldots, f_d) \,\in\, C(\Omega)^d\,,$$

where $C(\Omega)^d$ is the Cartesian product of d copies of $C(\Omega)$, we define

$$L[\boldsymbol{f}] \;:=\; \big(L[f_1], \ldots, L[f_d]\big)\,.$$

In this way, L is extended to an operator

$$L \,:\, C(\Omega)^d \longrightarrow C(\Omega)^d.$$

The class of the normalized operators belonging to $\mathcal{L}_+(A_+(\Omega))$ enjoys certain approximation properties, which we collect in the following theorem. These properties will help to better understand geometric properties of any operator $T \in \mathcal{L}_+(A_+(\Omega)$, especially when T has more structure than $T \in \mathcal{L}_+(A_+(\Omega))$.

Theorem 3.1. *Let $W = \{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_n\}$ a set of finite points of R^d and let the polytope $\Omega = conv(W)$. Let $\lambda_i, i = 0, \ldots, n$ be the barycentric coordinates defined by $\boldsymbol{x}_0, \ldots, \boldsymbol{x}_n$. Let $T \in \mathcal{L}_+(A_+(\Omega))$ such that $T[1] = 1$. Then the following statements hold*

(i) $\sum_{i=0}^n T[\lambda_i] = 1$;
(ii) $\sum_{i=0}^n \boldsymbol{x}_i T[\lambda_i] = T[\boldsymbol{e}]$;

(iii) $\sum_{i=0}^{n} l(\boldsymbol{x}_i)T[\lambda_i] = T[l]$, *for all* $l \in A(\Omega)$*;*
(iv) $T[l] = l \circ T[\boldsymbol{e}]$, *for all* $l \in A(\Omega)$*;*
(v) $T[\boldsymbol{e}]$ *maps* Ω *into itself;*
(vi) *Moreover, if* T *satisfies for every affine function* l *on* $\Omega, T[l] \geq 0$ *implies* $l \in A_+(\Omega)$*, then*

$$conv(T[\boldsymbol{e}](\Omega)) = \Omega. \tag{3.1}$$

Proof. To prove identities (i)-(iv) we do not require that the operator T belongs to $\mathcal{L}_+(A_+(\Omega))$: it is enough to assume that T is normalized and $T \in \mathcal{L}(C(\Omega))$. It is first observed that since $\lambda_i, i = 0, \ldots, n$, are the barycentric coordinates of $\boldsymbol{x}$ with respect to the points $\boldsymbol{x}_0, \ldots, \boldsymbol{x}_n$, the first two parts follow directly from equation (2.5) and then making use of the linearity of T. Also, by (2.4), we have for every $j = 1, \ldots, d$,

$$e_j = \sum_{i=0}^{n} e_j(\boldsymbol{x}_i)\lambda_i. \tag{3.2}$$

Fix an arbitrary affine function $l \in A(\Omega)$ then there exist $a = (a_1, \ldots, a_d) \in \mathrm{R}^d$ and $b \in \mathrm{R}$ such that

$$\begin{aligned} l: &= \sum_{j=1}^{d} a_j e_j + b \\ &= \sum_{j=1}^{d} a_j \left(\sum_{i=0}^{n} e_j(\boldsymbol{x}_i)\lambda_i \right) + b \\ &= \sum_{i=0}^{n} \left(\sum_{j=1}^{d} a_j e_j(\boldsymbol{x}_i) + b \right) \lambda_i \\ &= \sum_{i=0}^{n} l(\boldsymbol{x}_i)\lambda_i \end{aligned}$$

where the second step comes from (3.2), and the third equality follows from an interchange sums and using the fact that $\sum_{i=0}^{n} \lambda_i = 1$. Now applying the operators T on the both sides of the above equality, we conclude that (iii) holds.
To prove the identity (iv), it suffices to apply T to $l := \sum_{j=1}^{d} a_j e_j + b$.
We now show that the operator $T[\boldsymbol{e}]$ sends Ω into itself, assume the contrary that there exists a $\boldsymbol{y} \in \Omega$ such that $T[\boldsymbol{e}](\boldsymbol{y}) \notin \Omega$. Then, due to the Separation Theorem for closed convex sets (see, e.g., [28, p. 65, Theorem 2.4.1]), there exists a point $\boldsymbol{x}^* \in \Omega$ such that the affine function

$$l(\boldsymbol{x}) := \langle T[\boldsymbol{e}](\boldsymbol{y}) - \boldsymbol{x}^*, \boldsymbol{x} - \boldsymbol{x}^* \rangle \tag{3.3}$$

satisfies $l(\boldsymbol{x}) \leq 0$, for all $\boldsymbol{x} \in \Omega$. Here $\langle ., . \rangle$ denotes the usual scalar product in R^d. Hence $T[l] \leq 0$, since $l \leq 0$ and T is nonnegative for every nonnegative affine function. We know by (iv) the composition formula $T[l] = l \circ T[\boldsymbol{e}]$, and consequently,

$$T[l](\boldsymbol{y}) = l \circ T[\boldsymbol{e}](\boldsymbol{y}) := \|T[\boldsymbol{e}](\boldsymbol{y}) - \boldsymbol{x}^*\|^2 \leq 0.$$

This clearly implies $T[\boldsymbol{e}](\boldsymbol{y}) = \boldsymbol{x}^*$, and contradicts the fact that $T[\boldsymbol{e}](\boldsymbol{y}) \notin \Omega$. This yields the assertion (v).
To complete the proof, we must now show that, under the hypothesis on T, $conv(T[\boldsymbol{e}](\Omega)) = \Omega$. By (v) we have $T[\boldsymbol{e}](\Omega) \subset \Omega$, then $conv(T[\boldsymbol{e}](\Omega)) \subset \Omega$. Hence, it remains to prove the inverse inclusion $\Omega \subset conv(T[\boldsymbol{e}](\Omega))$. Since $T[\boldsymbol{e}]$ is continuous on Ω, then $T[\boldsymbol{e}](\Omega)$ is closed. Therefore $conv(T[\boldsymbol{e}](\Omega))$ is a nonempty closed convex set. Let us now assume the contrary, there exists $\boldsymbol{y} \in \Omega$, such that $\boldsymbol{y} \notin conv(T[\boldsymbol{e}](\Omega))$. Therefore, applying one more time the Separation Theorem, there exists $\boldsymbol{x}^* \in conv(T[\boldsymbol{e}](\Omega))$ such that the affine function

$$h(\boldsymbol{x}) := \langle \boldsymbol{y} - \boldsymbol{x}^*, \boldsymbol{x} - \boldsymbol{x}^* \rangle \,, \tag{3.4}$$

is nonpositive for all $\boldsymbol{x} \in conv(T[\boldsymbol{e}](\Omega))$. This means in particular that we would have $h \circ T[\boldsymbol{e}](\boldsymbol{x}) \leq 0$, for all $\boldsymbol{x} \in \Omega$. Then we must have, by (iv), $T[h] \leq 0$, and therefore we can, under hypothesis on T, get that $h \leq 0$. But if we take $\boldsymbol{x} = \boldsymbol{y}$ in (3.4), we conclude that $\|\boldsymbol{y} - \boldsymbol{x}^*\|^2 \leq 0$, contradicting the fact that $\boldsymbol{y} \neq \boldsymbol{x}^*$. Hence, we get the desired property and complete the proof of Theorem 3.1. □

The following key result characterizes all normalized operators, which belong to $\mathcal{L}_+(A_+(\Omega))$. Indeed we shall show that the condition (v) given in Theorem 3.1 is also a necessary one.

Theorem 3.2. *Let Ω be a convex polytope in R^d and let $T \in \mathcal{L}(C(\Omega))$ such that $T[1] = 1$. Then the following statements are equivalent:*
(i) *$T \in \mathcal{L}_+(A_+(\Omega))$;*
(ii) *$T[\boldsymbol{e}]$ maps Ω into itself.*

Proof. By Theorem 3.1, (v), it only remains to show that (ii) implies (i). Let us recall that the polytope Ω may also be defined by p inequalities:

$$\Omega = \left\{ \boldsymbol{x} \in \mathrm{R}^d : \langle \boldsymbol{a}_k, \boldsymbol{x} \rangle + b_k \geq 0, k = 1, \dots, p \right\},$$

here $\boldsymbol{a}_k \in \mathrm{R}^d$ and $b_k \in \mathrm{R}$, see, e. g., [26]. Fix now a nonnegative affine function l on Ω. Then, by the so-called affine form of Farkas' lemma, l is a nonnegative affine combination of the affine forms used to define the polytope Ω, see [25]. Therefore, there exist some nonnegative values $\alpha_k \geq 0$, for any $k = 0, \dots, p$ such that

$$l(\boldsymbol{x}) = \alpha_0 + \sum_{k=1}^{p} \alpha_k \left(\langle \boldsymbol{a}_k, \boldsymbol{x} \rangle + b_k \right), \quad (\boldsymbol{x} \in \Omega). \tag{3.5}$$

Then, applying T to both sides of the last equality and using the fact that $T[1] = 1$ yields for every $\boldsymbol{x} \in \Omega$,

$$T[l](\boldsymbol{x}) = \alpha_0 + \sum_{k=1}^{p} \alpha_k \left(\langle \boldsymbol{a}_k, T[\boldsymbol{e}](\boldsymbol{x}) \rangle + b_k \right).$$

This shows that $T[l](\boldsymbol{x}) = l(T[\boldsymbol{e}](\boldsymbol{x}))$, for any $\boldsymbol{x}$ in Ω. Since l takes nonnegative values on Ω and $T[\boldsymbol{e}]$ sends Ω into itself, then we obtain that $T[l]$ is nonnegative on Ω. Thus $T \in \mathcal{L}_+(A_+(\Omega))$ and we have completed the proof of the theorem. □

We now give an example showing that, in general, the polytope Ω is not a subset of $T[\boldsymbol{e}](\Omega)$, even in the one-dimensional context, if Ω is a closed interval of R.

Example 3.3. As a simple example, take $\Omega = [0,1]$ and consider the following operator

$$T : C([0,1] \to C\,[0,1]\,,$$

defined by

$$T[f](x) = \frac{1}{4}(2-x)f(0) + \frac{1}{4}(2+x)f(1). \tag{3.6}$$

In the present context, the function e is given by $e(x) = x, \forall x \in [0,1]$. It is immediate to realize that such an operator is linear, normalized, nonnegative for nonnegative affine functions and transfers any function from $C(\Omega)$ to an affine function. However, linear functions are not reproduced. Moreover, we have $T[e](x) = \frac{1}{4}(2+x)$, hence $T[e]([0,1]) = [1/2, 3/4]$. Therefore, $T[e]([0,1])$ does not contain $[0,1]$. Thus, for this operator, the assertion $\Omega \subset T[\boldsymbol{e}](\Omega)$ fails.

We would like to mention that, imposing the following normalization condition $T[1] = c$, where c is any fixed real in the open interval $(0,1)$. For this class of operators the following observation is valid.

Remark 3.4. The same argument can be applied to such operators to show that, under the assumption that $\mathbf{0} \in \Omega$, the statements of Theorem 3.1 still hold. Thus, we shall be concerned exclusively with the case where $c = 1$, but there is no difficulty in extending our result to the general case $T[1] \in (0,1)$.

When the operator satisfies some natural conditions, the following results give a new and very simple characterization of the normalized operators, which belong to $\mathcal{L}_+(A_+(\Omega))$. We denote by the symbol δ_{ij}, the Kronecker delta function evaluating to 1 for $i = j$ and 0 otherwise.

Proposition 3.5. Let Ω be a convex polytope with vertices $\{\boldsymbol{v}_0, \ldots, \boldsymbol{v}_n\}$. Let $\lambda_i, i = 0, \ldots, n$ be the barycentric coordinates defined by the vertices of Ω. Assume that $T \in \mathcal{L}(C(\Omega))$ and satisfies the following conditions:
(C1) $T[1] = 1$;
(C2) $T[\lambda_i] \geq 0, i = 0, \ldots, n$;
(C3) $T[\lambda_i](\boldsymbol{v}_j) = \delta_{ij}, i, j = 0, \ldots, n$.
Then for every affine function f, the following statements are equivalent:

- (i) f is nonnegative on Ω;
- (ii) f is nonnegative on the set of the vertices of Ω;
- (iii) $T[f]$ is nonnegative on the set of the vertices of Ω;
- (iv) $T[f]$ is nonnegative on Ω.

Before proving Proposition 3.5 we give a lemma of independent interest.

Lemma 3.6. *Let f be an affine function on a polytope Ω with vertices $\{\boldsymbol{v}_0, \ldots, \boldsymbol{v}_n\}$. If $m \leq f(\boldsymbol{v}_i) \leq M$ for $i = 0, \ldots, n$, then*

$$m \leq f(\boldsymbol{x}) \leq M \text{ for all } \boldsymbol{x} \in \Omega. \tag{3.7}$$

Proof. An affine function f on a polytope Ω attains its maximum at an extreme point of Ω. Therefore, since $f(\boldsymbol{v}_i) \leq M$ for $i = 0, \ldots, n$, then $f(\boldsymbol{x}) \leq M$ for all $\boldsymbol{x} \in \Omega$. Consider the set $E = \{\boldsymbol{x} \in \Omega, f(\boldsymbol{x}) \geq m\}$. The left inequality holds in (3.7) if and only if $E = \Omega$. Observe that E is a nonempty, closed, convex set and that $\{\boldsymbol{v}_0, \ldots, \boldsymbol{v}_n\} \subset E \subset \Omega$. This shows that $E = \Omega$, since $\Omega = conv(\{\boldsymbol{v}_0, \ldots, \boldsymbol{v}_n\})$. ☐

Now we are ready to prove Proposition 3.5.

Proof of Proposition 3.5. By Theorem 2.1, a polytope is defined by its vertices, and any point of the polytope is a (continuous) nonnegative barycentric combination of the polytope vertices. The equivalence between the first two parts is an immediate consequence of Lemma 3.6 applied with $m = 0$. Now, from the equation

$$l = \sum_{i=0}^{n} l(\boldsymbol{v}_i)\lambda_i, \tag{3.8}$$

which holds for all $l \in A(\Omega)$, it is clear that

$$T[f] = \sum_{i=0}^{n} f(\boldsymbol{v}_i)T[\lambda_i]. \tag{3.9}$$

Consequently, under condition (C3), every linear operator must satisfy the interpolation conditions at the vertices:

$$T[f](\boldsymbol{v}_i) = f(\boldsymbol{v}_i), i = 1, \ldots, n, \tag{3.10}$$

which shows the equivalence between (ii) and (iii).
The equivalence between (iii) and (iv) is a direct consequence of equations (3.9) and (3.10). We have thus completed the proof of Proposition 3.5. ☐

We remark that in Proposition 3.5, T can be taken the identity operator. We also note that in the special case that T is the identity operator all conditions (C1-C3) are automatically satisfied.

We are now interested in the circumstances under which the statements (vi) of Theorem 3.1 become equivalent, thus the natural question to ask is: how to characterize the equality $conv(T[\boldsymbol{e}](\Omega)) = \Omega$?
The following observation, which is quite easy to check, characterizes the operators for which we have $conv(T[\boldsymbol{e}](\Omega)) = \Omega$.

Remark 3.7. Subject to conditions (C1-C3) of Proposition 3.5, the following statements are equivalent for every normalized linear operator $T \in \mathcal{L}_+(A_+(\Omega))$.

(i) $conv(T[\boldsymbol{e}](\Omega)) = \Omega$;
(ii) If T satisfies $\forall l \in A(\Omega), T[l] \geq 0$ implies $l \geq 0$.

We are unable to decide how much the hypotheses of Proposition 3.5 below can be weakened, but it cannot be omitted entirely, as can be seen by considering the following example.

Example 3.8. We will use Example 3.3 to show that, subject to the only conditions (C1) and (C2), Proposition 3.5 is false. Indeed, it is easily verified that the operator T defined by (3.6) satisfies both conditions (C1) and (C2). On the other hand, for the function f defined by $f(x) = 1 - \frac{4}{3}x$, we have $T[f](x) = \frac{1}{3}(1-x)$. Therefore, $T[f]$ is nonnegative however f is not of constant sign in the interval $[0, 1]$. This makes that proposition not true but false. Obviously, such an operator cannot also satisfies condition (C3) as this would violate Proposition 3.5. In fact, a simple calculation will give that for $\lambda_1(x) = x$, we have $T[\lambda_1](0) = 1/2$. Thus, this example shows that condition (C3) may not be omitted from the statement of Proposition 3.5.

Now, if in Proposition 3.5, instead of conditions (C1) and (C3), we assume, under condition (C2), that the operator T reproduces exactly all affine functions (and therefore T would automatically belong to $\mathcal{L}_+(A_+(\Omega))$, we can show that both conditions (C1) and (C3) hold.
In fact, this observation may be derived from the following more general result which is of some independent interest.

Theorem 3.9. *Let Ω be a convex polytope with vertices $\{\boldsymbol{v}_0, \ldots, \boldsymbol{v}_n\}$. Let $\lambda_i, i = 0, \ldots, n$ be the barycentric coordinates defined by the vertices of Ω. Let $T \in \mathcal{L}(C(\Omega))$, which satisfies condition (C2) and reproduces all affine functions. Define the operator $\widetilde{T}$ by*

$$\widetilde{T}[f] = \sum_{i=0}^{n} f(\boldsymbol{v}_i) T[\lambda_i]. \tag{3.11}$$

Then, $\widetilde{T}$ preserves affine functions and satisfies

$$f(\boldsymbol{v}_i) = \widetilde{T}(\boldsymbol{v}_i), (i = 0, \ldots, n) \tag{3.12}$$

for every $f \in C(\Omega)$.

Proof. Fix an affine function l, then $l = \sum_{i=0}^{n} l(\boldsymbol{v}_i)\lambda_i$. Since T preserves affine functions then we have

$$l = T[l] = \sum_{i=0}^{n} l(\boldsymbol{v}_i) T[\lambda_i] := \widetilde{T}[l]. \tag{3.13}$$

This shows that the operator $\widetilde{T}$ preserves affine functions too.
Under condition (C2), the operator $\widetilde{T}$ is nonnegative on $C(\Omega)$, so that $\widetilde{T}[f] \geq 0$, for all *nonnegative* functions f of $C(\Omega)$. We may now apply [11, Theorem 4.5] to conclude

$$\widetilde{T}[f](\boldsymbol{v}_i) = f(\boldsymbol{v}_i), (i = 0, \ldots, n), \tag{3.14}$$

and complete the proof of the theorem. □

Thus, we may now derive the following easy implication.

Remark 3.10. If T satisfies condition (C2) and preserves affine function then, in this situation, we have by (3.13)

$$\boldsymbol{e} = T[\boldsymbol{e}] = \widetilde{T}[\boldsymbol{e}] = \sum_{i=0}^{n} \boldsymbol{v}_i T[\lambda_i].$$

Hence, by (3.12) and the fact that $\boldsymbol{v}_i$ are the vertices of the polytope Ω, T must satisfy condition (C3).

4. Generalized multivariate Jensen-type inequalities

We are going to present generalized Jensen-type inequalities in the multi-variable context. In this section, we derive two general composition formulae from which, by suitably specializing their coefficients as parameters, we may get a large number of (new and known) Jensen-type inequalities for functions of one variable [1, 5, 8, 19, 22]. We first examine the case of a multivariate convex function $f : \Omega \to \mathrm{R}$, where smoothness is not required; only continuity is assumed.
The main difficulty here is how to construct practical and effective linear approximation normalized operators G, such that $G[\boldsymbol{e}]$ sends Ω into itself. Theorem 3.2 says that it is necessary and sufficient that G satisfies $G \in \mathcal{L}_+(A_+(\Omega))$. The latter condition is sometimes difficult to check. Therefore it is worthwhile to seek conditions sufficient to guarantee its validity. Thus, it is natural for us to build up new operators with such property from simpler ones, via operations preserving or even yielding it. To this end, let us first introduce some linear approximation operators, which will be used throughout the rest of the paper and help to provide the necessary motivation. Given a function $f \in C(\Omega)$, and $n+1$ linear operators $L_i : C(\Omega) \to C(\Omega)$, as a general form of the approximation operator $L[f]$ to the function f, we can select the averaging approximation operator

$$L[f] = \sum_{i=0}^{n} \beta_i L_i[f], \tag{4.1}$$

where $\beta_i, i = 0, \dots, n$ are some given nonnegative continuous functions in $C(\Omega)$, which form a partition of unity, that is

$$\sum_{i=0}^{n} \beta_i(\boldsymbol{x}) = 1 \text{ on } \Omega. \tag{4.2}$$

One possible choice of a partition of unity is the barycentric coordinates given by Theorem 2.1 with respect to $(n+1)$-points $\{\boldsymbol{x}_0, \dots, \boldsymbol{x}_n\}$, or a collection of finite element functions (as we shall discuss), see [2, 21]. It should be noted that we do not impose the affine functions reproduction property. Equations (4.1) and (4.2) in conjunction with the nonnegativity property of $\beta_i, i = 0, \dots, n$ ensure that the approximation operator L is bounded between the minimum and maximum of $\min_{0 \le i \le n} L_i$ and $\max_{0 \le i \le n} L_i$.
It will be convenient to have a brief description of the operator given by (4.1). The applications of (4.1) involve various requirements from the operators L_j. Often, in many practical situations, we do not know the operators L_i but only some kind of approximations of it. It will be more reasonable to take, for instance, some convex combination of the values of the nearest points around a specific point $\boldsymbol{x}$, as an approximation to $L_i[f](\boldsymbol{x})$ in regions of interest, and interpolate to it. Hence, if for each $i, 0 \le i \le n$, we assume that for some integer $m \ge 2$, there are nonnegative continuous functions $\{\psi_{ij}, j = 1, \dots, m\}$, such that

$$\sum_{j=1}^{m} \psi_{ij}(\boldsymbol{x}) = 1 \text{ on } \Omega,$$

a most general form of an approximation operator $G : C(\Omega) \to C(\Omega)$ of f is given by

$$G[f] = \sum_{i=0}^{n} \left(\sum_{j=1}^{m} \frac{1 - \psi_{ij}}{m-1} L_j[f] \right) \beta_i. \tag{4.3}$$

Approximation by means of (4.3) constitutes the core of the partition of unity finite element methods [2]. Suitable forms for the operators L_j can usually be taken to be a weighted integral, which includes point evaluation of a partial derivative of f. This can best be illustrated by the classical Lagrange interpolation operator, which will be recovered if we take $L_j[f](\boldsymbol{x}) = f(\boldsymbol{v}_j)$, $j = 1, \dots, m$. Here, $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_m\}$ are the set of vertices of Ω. Then for each i, we take $\psi_{ij} := \lambda_j(\boldsymbol{y}_i), j = 1, \dots, m$, the barycentric coordinates with respect to $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_m\}$ of the finite subset of distinct but otherwise arbitrary $(n+1)$-points $\boldsymbol{y}_i \in \Omega$. While for the functions $\beta_i, i = 0, \dots, n$, we may take the barycentric coordinates with respect to $\{\boldsymbol{x}_0, \dots, \boldsymbol{x}_n\}$.
We note that the value of n used in (4.3) can be taken equal to -1 (therefore only the second sum appears) or arbitrarily large.
In order to avoid all unnecessary repetition, we will assume that the operators $L_j \in \mathcal{L}_+(A_+(\Omega))$, and they are normalized, that is $L_j[1] = 1$. We note here that the operator G will automatically 'inherit' these properties. To see this it suffices to remark that by interchanging the order of summation in the double-sum in equation (4.3), it follows that

$$G[f] = \sum_{j=1}^{m} \left(\sum_{i=0}^{n} \frac{1 - \psi_{ij}}{m-1} \beta_i \right) L_j[f]. \tag{4.4}$$

Thus, the operator G defined in this manner takes the form of a weighted combination with nonnegative weights of the operators L_j, with the sum of weights is 1, and such that G preserves the belonging to the cone $\mathcal{L}_+(A_+(\Omega))$, the normalization condition $G[1] = 1$, and therefore by Theorem 3.1, (v), $G[\boldsymbol{e}]$ sends Ω into itself.

Note also that, the expression (4.3) defining G can immediately be rewritten in the following form

$$G[f] = \frac{m}{m-1} U[f] - \frac{1}{m-1} V[f], \tag{4.5}$$

with

$$U := \sum_{j=1}^{m} \frac{1}{m} L_j, \tag{4.6}$$

$$V := \sum_{i=0}^{n} \left(\sum_{j=1}^{m} \psi_{ij} L_j \right) \beta_i. \tag{4.7}$$

Since each operator $L_j \in \mathcal{L}_+(A_+(\Omega))$, it is easy to see that the operators U and V belong to $\mathcal{L}_+(A_+(\Omega))$. Now, if we define, for every i, j, the functions $\boldsymbol{y}_i$ and $\boldsymbol{w}_j$ as follows:

$$\boldsymbol{y}_i = \sum_{j=1}^{m} \psi_{ij} L_j[\boldsymbol{e}], \tag{4.8}$$

$$\boldsymbol{w}_j = L_j[\boldsymbol{e}], \tag{4.9}$$

then, by Theorem 3.1, (v), for every i, j, the functions $\boldsymbol{y}_i$, $\boldsymbol{w}_j$ will send Ω into itself, since we have represented them as a convex combination of elements of $\mathcal{L}_+(A_+(\Omega))$. An even more simplified version of the expression of $G[\boldsymbol{e}]$ can be obtained if we now substitute equations (4.8) and (4.9) into (4.5) we then have

$$\begin{aligned} G[\boldsymbol{e}] &= \frac{1}{m-1}\sum_{j=1}^{m} \boldsymbol{w}_j - \frac{1}{m-1}\sum_{i=0}^{n} \beta_i \boldsymbol{y}_i, \\ &:= \frac{m}{m-1} U[\boldsymbol{e}] - \frac{1}{m-1} V[\boldsymbol{e}]. \end{aligned} \tag{4.10}$$

Therefore, the operators $G[\boldsymbol{e}]$ is always regular, i.e., it can be represented as the difference of two operators from $\mathcal{L}_+(A_+(\Omega))$.

The next Theorem provides a natural generalization of Mercer's inequality (1.1) to a multivariate setting, these inequalities are to be understood in the usual sense in the space $C(\Omega)$, e.g., for all $\boldsymbol{x} \in \Omega$. The reader should observe that, $\{\beta_i, i = 0, \ldots, n\}$ is any fixed (but otherwise arbitrary) partition of unity defined the approximation operator (4.4).

Theorem 4.1. *Let $\boldsymbol{w}_j$ and $\boldsymbol{y}_i$ be the functions given by equations* (4.8) *and* (4.9) *respectively. Then, for any convex function f the following inequality holds on Ω.*

$$f \circ \left(\frac{1}{m-1} \left(\sum_{j=1}^{m} \boldsymbol{w}_j - \sum_{i=0}^{n} \beta_i \boldsymbol{y}_i \right) \right) \le \frac{1}{m-1} \left(\sum_{j=1}^{m} f \circ \boldsymbol{w}_j - \sum_{i=0}^{n} \beta_i f \circ \boldsymbol{y}_i \right). \tag{4.11}$$

Proof. Starting from (4.10), we already know

$$G[\boldsymbol{e}] = \frac{1}{m-1} \left(\sum_{j=1}^{m} \boldsymbol{w}_j - \sum_{i=0}^{n} \beta_i \boldsymbol{y}_i \right).$$

But, taking into account (4.3), we also have

$$G[\boldsymbol{e}] = \sum_{i=0}^{n} \left(\sum_{j=1}^{m} \frac{1-\psi_{ij}}{m-1} L_j[\boldsymbol{e}] \right) \beta_i, \tag{4.12}$$

then by interchanging the order of summation in the double-sum in equation (4.12) it follows that $G[\boldsymbol{e}]$ takes the form

$$G[\boldsymbol{e}] = \sum_{j=1}^{m} \left(\sum_{i=0}^{n} \frac{1-\psi_{ij}}{m-1} \beta_i \right) L_j[\boldsymbol{e}]. \tag{4.13}$$

This permits us to rewrite $G[\boldsymbol{e}]$ as a convex combination of $L_j[\boldsymbol{e}]$, and therefore f is well-defined and since f is a convex function, we conclude that

$$f \circ G[\boldsymbol{e}] \leq \sum_{j=1}^{m} \left(\sum_{i=0}^{n} \frac{1-\psi_{ij}}{m-1} \beta_i \right) f \circ L_j[\boldsymbol{e}]. \tag{4.14}$$

By transforming the last term at the right-hand side of the above equation gives

$$f \circ G[\boldsymbol{e}] \leq \frac{1}{m-1} \left(\sum_{j=1}^{m} f \circ L_j[\boldsymbol{e}] - \sum_{i=0}^{n} \left(\sum_{j=1}^{m} \psi_{ij} f \circ L_j[\boldsymbol{e}] \right) \beta_i \right). \tag{4.15}$$

The inequality (4.15) actually implies the desired result (4.11), because the classical Jensen's inequality says

$$f \circ \left(\sum_{j=1}^{m} \psi_{ij} L_j[\boldsymbol{e}] \right) \leq \sum_{j=1}^{m} \psi_{ij} f \circ L_j[\boldsymbol{e}]$$

which may be rewritten

$$-\sum_{j=1}^{m} \psi_{ij} f \circ L_j[\boldsymbol{e}] \leq -f \circ \boldsymbol{y}_i$$

taken together with (4.15) and (4.12) yields (4.11) after a little simplification. This concludes the proof of Theorem 4.1. □

We now examine the converse inequality for approximation operator (4.4). Our results will make use of the remarkable property (4.5) satisfied by the operator G, which can be rewritten as

$$U = \frac{m-1}{m} G + \frac{1}{m} V, \tag{4.16}$$

then U can be represented as a convex combination of two operators on $\mathcal{L}_+(A_+(\Omega))$.

The following converse of inequality (4.11) holds.

Theorem 4.2. *Let U and V be the two operators defined by equations* (4.6) *and* (4.7) *respectively. Then, for any convex function f the following inequality holds on Ω.*

$$\frac{m}{m-1} f \circ U[\boldsymbol{e}] - \frac{1}{m-1} f \circ V[\boldsymbol{e}] \leq f \circ \left(\frac{m}{m-1} U[\boldsymbol{e}] - \frac{1}{m-1} V[\boldsymbol{e}] \right). \tag{4.17}$$

Proof. The desired result may be obtained by using the convexity of f, with the help of the key identity (4.16). □

Let us observe that, the right-hand side of inequality (4.17) is exactly the left-hand side of in inequality (4.11). Therefore, (4.17) is a converse inequality of (4.11).

5. New inequalities related to the Jensen-type inequalities

In this section, we are now going to go through some examples to further illustrate the results of Theorem 4.1 and 4.2. For convenience, we restrict ourselves to the case of linear operators for the evaluation of functions, since in this context most of the ideas underlying the general setting are present. Our goal in the next examples is to demonstrate this by inserting suitably chosen operators L_j in (4.11) and (4.17). More precisely we have the following direct consequence of Theorem 4.1, which we state it as a theorem.

Theorem 5.1. *Let $\Omega = conv\left(\{\boldsymbol{t}_1, \ldots, \boldsymbol{t}_m\}\right)$ with $m \geq 2$, and $\{\boldsymbol{z}_0, \ldots, \boldsymbol{z}_n\}$ a set of $(n+1)$-points in Ω. Then, for any convex function f the following inequality holds for every $\boldsymbol{x} \in \Omega$.*

$$f\left(\frac{1}{m-1}\left(\sum_{j=1}^{m} \boldsymbol{t}_j - \sum_{i=0}^{n} \beta_i(\boldsymbol{x})\boldsymbol{z}_i\right)\right) \leq \frac{1}{m-1}\left(\sum_{j=1}^{m} f(\boldsymbol{t}_j) - \sum_{i=0}^{n} \beta_i(\boldsymbol{x}) f(\boldsymbol{z}_i)\right).$$

Proof. Simply take $L_j[f](\boldsymbol{x}) = f(\boldsymbol{t}_j)$, $j = 1, \ldots, m$, then by (4.8) and (4.9) we have $\boldsymbol{w}_j = \boldsymbol{t}_j$ and $\boldsymbol{y}_i = \sum_{j=1}^{m} \psi_{ij}\boldsymbol{t}_j$. For each $i = 0, \ldots, n$, we take $\psi_{ij} = \lambda_j(\boldsymbol{z}_i), j = 1, \ldots, m$, the barycentric coordinates with respect to $\{\boldsymbol{t}_1, \ldots, \boldsymbol{t}_m\}$ of the point $\boldsymbol{z}_i$. Hence $\boldsymbol{y}_i = \boldsymbol{z}_i, i = 0, \ldots, n$ and therefore by Theorem 4.1 the statement of Theorem 5.1 is true. ☐

In the special case that the points $\{\boldsymbol{t}_1, \ldots, \boldsymbol{t}_m\}$ are chosen the vertices of the polytope Ω, as an application of Theorem 5.1 we obtain the following Corollary:

Corollary 5.2. *Let Ω be a polytope with m vertice $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m\}$ with $m \geq 2$, and let $\boldsymbol{z}_0, \ldots, \boldsymbol{z}_n$ be $(n+1)$-points in Ω. Then, for any convex function f the following inequality holds on Ω*

$$f\left(\frac{1}{m-1}\left(\sum_{j=1}^{m} \boldsymbol{v}_i - \sum_{i=0}^{n} \beta_i(\boldsymbol{x})\boldsymbol{z}_i\right)\right) \leq \frac{1}{m-1}\left(\sum_{j=1}^{m} f(\boldsymbol{v}_i) - \sum_{i=0}^{n} \beta_i(\boldsymbol{x}) f(\boldsymbol{z}_i)\right).$$

The case where the polytope is a simplex appears already in [19] with some constant coefficients, and so the general inequality of Theorem 5.1 covers the constant coefficient case of [19, Theorem 1, inequality (4)]. Finally, when Ω is the closed interval $\Omega = [a, b]$ of R (hence m=2), and some points

$$x_0 \leq x_2 \leq \cdots \leq x_n \tag{5.1}$$

contained in the interval $[a, b]$. Corollary 5.2 yields in particular Mercer's inequality (1.1), and therefore, as we mentioned in the introduction, we have in hand Theorem 1.2 of [22]. However, it is important to observe that in order to obtain this result, the author in [22] has assumed that all points x_i are nonnegative, as a part of whole assumptions. But in his proof, we can see that this condition is needless to this theorem, condition (5.1) is enough. We note, in passing, that the resulting weighted inequality holds in our case with not necessarily constant coefficients.

Fix a positive integer $m \geq 2$ and let $\Omega = conv\,(W)\,,$ where $W = \{\boldsymbol{t}_1, \ldots, \boldsymbol{t}_m\}$ is a set of a sequence of points in R^d. We define the barycenter of W to be

$$\boldsymbol{t}_W = \frac{1}{m}\sum_{j=1}^{m} \boldsymbol{t}_j.$$

We now apply our approach to derive direct and converse multidimensional extensions of a one-dimensional inequality due to Bougoffa [3, Theorem 1.4].

Remark 5.3. Recall that the partition of unity $\{\beta_i, i = 0, \ldots, n\}$ used in Theorem 5.1 is arbitrary. We now present a natural choice of such a partition of unity to obtain some new inequalities. To this end we use Theorem 5.1. Let us first fix $n = m-1$ and the functions $\beta_i = \lambda_{i+1}, i = 0, \ldots, m-1$, the usual piecewise linear basis functions of the P_1 finite element, defined on a simplex mesh subdivision of Ω using the points $\boldsymbol{t}_1, \ldots, \boldsymbol{t}_m$ (so every point $\boldsymbol{t}_i$ must be a vertex of some simplex of the triangulation). It is well known that the P_1 finite element basis functions form a partition of unity with the Kronecker delta property at the nodes $\boldsymbol{t}_1, \ldots, \boldsymbol{t}_m$, see [6]. We now fix $\boldsymbol{z}_i = \boldsymbol{t}_{i+1}, i = 0, \ldots, m-1$ and define the vectors

$$\widetilde{\boldsymbol{t}}_i = \frac{m\boldsymbol{t}_W - \boldsymbol{t}_i}{m-1}, i = 1, \ldots, m.$$

We then let $\boldsymbol{x} = \boldsymbol{t}_j$ and use the Kronecker delta property of the partition of unity to get immediately the following consequence of Theorem 5.1

$$f\left(\widetilde{\boldsymbol{t}}_j\right) \leq \frac{1}{m-1}\left(\sum_{i=1}^{m} f(\boldsymbol{t}_i) - f(\boldsymbol{t}_j)\right).$$

Summing over $j = 1\ldots, m$, yields

$$\sum_{j=1}^{m} f\left(\widetilde{\boldsymbol{t}}_j\right) \leq \frac{m}{m-1}\sum_{i=1}^{m} f(\boldsymbol{t}_i) - \frac{1}{m-1}\sum_{j=1}^{m} f(\boldsymbol{t}_j),$$

and hence the inequality

$$(m-1)\sum_{j=1}^{m} f\left(\widetilde{\boldsymbol{t}}_j\right) \leq m\left(\sum_{i=1}^{m} f(\boldsymbol{t}_i) - f(\boldsymbol{t}_W)\right) \tag{5.2}$$

is valid. This is exactly the multidimensional version of Bougoffa's inequality [3, Theorem 1.4].

Here we present our converse to the inequality of Corollary 5.2.

Corollary 5.4. *Let $\boldsymbol{z}_0, \ldots, \boldsymbol{z}_n$ be n-points in Ω. Then, for any convex function f the following inequality holds on Ω*

$$\frac{m}{m-1} f\left(\boldsymbol{t}_W\right) - \frac{1}{m-1} f\left(\sum_{i=0}^{n} \beta_i(\boldsymbol{t})\boldsymbol{z}_i\right) \leq f\left(\frac{m}{m-1}\boldsymbol{t}_W - \frac{1}{m-1}\sum_{i=0}^{n} \beta_i(\boldsymbol{t})\boldsymbol{z}_i\right). \tag{5.3}$$

Proof. This result can be obtained as a direct and immediate consequence of Theorem 4.2. □

Now we are in the position to give a converse of the multidimensional version (5.2) of Bougoffa's inequality. Indeed, with the same notation as we did for the direct inequality given by formula (5.2) below, an easy calculation shows, when we use again P_1 finite element basis functions, that inequality (5.3) in Corallary 5.4 can be simplified to

$$m^2 f(\boldsymbol{t}_W) - \sum_{i=1}^{m} f(\boldsymbol{t}_i) \leq (m-1) \sum_{i=1}^{m} f\left(\widetilde{\boldsymbol{t}}_i\right). \tag{5.4}$$

We summarize the two above multidimensional extensions (5.2) and (5.4) in the following corollary.

Corollary 5.5. *Assume that* $\Omega = conv(\{\boldsymbol{t}_1, \dots, \boldsymbol{t}_m\})$. *Define*

$$\widetilde{\boldsymbol{t}}_i = \frac{m\boldsymbol{t}_\Omega - \boldsymbol{t}_i}{m-1}, i = 1, \dots, m,$$

where $\boldsymbol{t}_\Omega = \frac{1}{m}\sum_{j=1}^{m} \boldsymbol{t}_j$. *Then for every convex function* f *on* Ω, *the inequalities*

$$m^2 f(\boldsymbol{t}_\Omega) - \sum_{i=1}^{m} f(\boldsymbol{t}_i) \leq (m-1) \sum_{i=1}^{m} f\left(\widetilde{\boldsymbol{t}}_i\right) \leq m\left(\sum_{i=1}^{m} f(\boldsymbol{t}_i) - f(\boldsymbol{t}_\Omega)\right) \tag{5.5}$$

are valid.

6. A refinement for twice continuously differentiable functions

A related and important problem to approximation operators is their precision. This section is devoted to a useful elementary principle, to improve the quality of the general inequalities (4.11) and (4.17) for smooth (nonconvex) twice continuously differentiable functions. This is important to the error bounds that are going to be derived. In order to proceed in developing our argument, we need some additional necessary background and notation.
By S_d we denote the set of all $d \times d$ symmetric matrices in R. Let $A \in S_d$, and $\beta_i[A], i = 1, \dots, d$, the (real) eigenvalues of A, we define

$$\beta_{min}[A] := \min_{1\leq i\leq d} \beta_i[A] = \min_{\|\boldsymbol{y}\|=1} \langle A\boldsymbol{y}, \boldsymbol{y}\rangle.$$

We say $A \in S_d$ is positive semidefinite if $\langle A\boldsymbol{y}, \boldsymbol{y}\rangle \geq 0$, for every $\boldsymbol{y} \in \mathrm{R}^d$. The set of positive semidefinite symmetric matrices (all eigenvalues ≥ 0) is denoted by S_d^+.
By $D^2 f(\boldsymbol{x})$, we mean the $d \times d$ matrix whose entries are the second-order partial derivatives of f at $\boldsymbol{x}$. It is well known that when f is a $C^2(\Omega)$-function, its convexity is characterized by the fact that for all $\boldsymbol{x} \in \Omega, D^2 f(\boldsymbol{x}) \in S_d^+$ (see e.g. [24]). For every $x \in \Omega$, the Hessian matrix $D^2 f(\boldsymbol{x})$, as real-valued and symmetric matrix, has real-valued eigenvalues. Therefore, for every function f in $C^2(\Omega)$, we may define

$$\lambda_{min}[f] := \inf_{\boldsymbol{x}\in\Omega} \beta_{min}[D^2 f(\boldsymbol{x})].$$

We shall call $\lambda_{min}[f]$ the 'globally' smallest eigenvalue of the Hessian $D^2 f(\boldsymbol{x})$ on Ω. Now, let f be any $C^2(\Omega)-$function and set

$$g := f - \frac{\lambda_{min}[f]}{2} \|.\|^2, \tag{6.1}$$

($\|.\|^2$ denotes the Euclidean norm on R^d.) The Hessian matrix of g is

$$D^2g(\boldsymbol{x}) = D^2f(\boldsymbol{x}) - \lambda_{min}[f]I_d,$$

where I_d denotes the $d\times d$ identity matrix. Therefore, for $\boldsymbol{y} \in \mathrm{R}^d$ such that $\|\boldsymbol{y}\| = 1$, we have

$$\langle \boldsymbol{y}, D^2g(\boldsymbol{x})(\boldsymbol{y})\rangle = \langle \boldsymbol{y}, D^2f(\boldsymbol{x})(\boldsymbol{y})\rangle - \lambda_{min}[f].$$

It is clear from the definition of $\lambda_{min}[f]$ that, for every $\boldsymbol{x} \in \Omega$, the right-hand term in the above equation is nonnegative. This means that the Hessian matrix of g is positive semidefinite for all $\boldsymbol{x}$ of the set Ω, and consequently g is convex.
Hence, an arbitrary nonconvex twice continuously differentiable function is made convex after adding to it the quadratic $-\frac{\lambda_{min}[f]}{2}\|.\|^2$. It should be noted that for every $\alpha \leq \lambda_{min}[f]$, the function

$$g_\alpha := f - \frac{\alpha}{2}\|.\|^2,$$

is convex. Indeed, we have

$$g_\alpha = g + \frac{\lambda_{min}[f] - \alpha}{2}\|.\|^2,$$

where g is described by the formula (6.1), then g_α is convex as sum of two convex functions.
Note that $\lambda_{min}[f]$ is not necessarily zero if f is convex over Ω. On the other hand, if $\lambda_{min}[f] \geq 0$ then f is convex over Ω.
Let $f \in C^2(\Omega)$. Consider the operators

$$T[f] := f\circ\left(\frac{1}{m-1}\left(\sum_{j=1}^{m}\boldsymbol{w}_j - \sum_{i=0}^{n}\beta_i\boldsymbol{y}_i\right)\right) \tag{6.2}$$

$$T_+[f] := \frac{1}{m-1}\left(\sum_{j=1}^{m} f\circ\boldsymbol{w}_j - \sum_{i=0}^{n}\beta_i f\circ\boldsymbol{y}_i\right), \tag{6.3}$$

$$T_-[f] := \frac{m}{m-1}f\circ U[\boldsymbol{e}] - \frac{1}{m-1}f\circ V[\boldsymbol{e}]. \tag{6.4}$$

Theorem 4.1 and 4.2 say that T_+,(resp. T_-), is an overestimator (resp. an underestimator) of T , in the sense

$$T_-[g] \leq T[g] \leq T_+[g], \tag{6.5}$$

for all convex functions $g \in C(\Omega)$. Define

$$R_+[\|.\|^2] = T_+[\|.\|^2] - T[\|.\|^2],$$

and

$$R_-[\|.\|^2] = T[\|.\|^2] - T_-[\|.\|^2],$$

by $\|.\|$ we mean the Euclidean norm in R^d. Note that $R_+[\|.\|^2]$ and $R_-[\|.\|^2]$ are nonnegative (since $\|.\|^2$ is a convex function).

With this notation, the next result examines the case of estimating a given (non-convex) twice continuously differentiable function and shows that there exist better estimators than T_- and T_+ defined by (6.5). More precisely we have:

Theorem 6.1. *Let f be an arbitrary twice continuously differentiable function defined on Ω. Let T_- and T_+ be an underestimator (resp. overestimator) of T. Then the following estimates hold*

$$T_-[f] + \frac{\lambda_{min}[f]}{2} R_-[\|.\|^2] \leq T[f] \leq T_+[f] - \frac{\lambda_{min}[f]}{2} R_+[\|.\|^2]. \tag{6.6}$$

Let us emphasize before proving Theorem 6.1 that by comparing the estimates (6.5) and (6.6), we can see that if f is strictly convex (then $\lambda_{min}[f] > 0$), and therefore (6.6) always provides better bounds than (6.5). Note also that, since $R_-[\|.\|^2]$ and $R_+[\|.\|^2]$ are nonnegative, then the new bounding expressions are obtained, by subtracting (resp. adding) a nonnegative correction term from the right-hand (resp. left-hand) side of the estimate (6.5). These bounds will be used later to obtain a new class of general refined Jensen type inequalities.

Proof. Under the present assumption about the function f, it is evident that the auxiliary function

$$g := f - \frac{\lambda_{min}[f]}{2} \|.\|^2$$

is convex. Hence we can apply the inequality (6.5) to g and rearranging terms leads to the required inequality. □

6.1. **An application: A refinement of Jensen's discrete inequality.** As a simple example of illustration, we now see how using Theorem 6.1 on concrete situation yields a new refined form of Jenesen's inequality. We begin by introducing the starting inequality, that will be modified, using our proposed technique. Let Ω be a polytope with m vertices $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m\}$ with $m \geq 2$, and define $\boldsymbol{z}_i = \boldsymbol{v}_{i+1}, i = 0, \ldots, m-1$. Recall that the partition of unity, β_i used in Corollary 5.2 is arbitrary. Let $\lambda_i, i = 1, \ldots, n$, be the barycentric coordinates defined by $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$. Now, we may fix such a partition of unity in such a way that $\lambda_i = \frac{1-\beta_i}{m-1}, i = 1, \ldots, m$. Then with the help of this system of notation Corollary 5.2 gives

$$f\left(\sum_{j=1}^{m} \lambda_j(\boldsymbol{x})\boldsymbol{v}_j\right) \leq \sum_{j=1}^{m} \lambda_j(\boldsymbol{x}) f(\boldsymbol{v}_j), \tag{6.7}$$

for every convex function f. This is just the classical Jensen's inequality.
The next result, which is based on the inequality of Theorem 6.1, provides the following refined inequality of the classical Jensen's inequality for $C^2(\Omega)$-functions. Note that convexity is not required for our result.

Theorem 6.2. *Let Ω be a convex polytope with vertices $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m\}$. Then, for every $f \in C^2(\Omega)$, the following inequality holds on Ω*

$$f\left(\sum_{j=1}^{m} \lambda_j(\boldsymbol{x})\boldsymbol{v}_j\right) \leq \sum_{j=1}^{m} \lambda_j(\boldsymbol{x}) f(\boldsymbol{v}_j) - \frac{\lambda_{min}[f]}{2} \sum_{i,j=1, i<j}^{m} \lambda_i(\boldsymbol{x})\lambda_j(\boldsymbol{x}) \|\boldsymbol{v}_i - \boldsymbol{v}_j\|^2 .$$

Proof. Define, for every $g \in C^2(\Omega)$,

$$T[g] = g\left(\sum_{j=1}^{m} \lambda_j \boldsymbol{v}_j\right)$$

and

$$T_+[g] = \sum_{j=1}^{m} \lambda_j g(\boldsymbol{v}_j).$$

It is easy seen that T, T_+ are two linear operators from $\mathcal{L}(C(\Omega))$, and satisfy (6.7). Then, applying Theorem 6.1, we find that for all $\boldsymbol{x}$ in Ω

$$T[f](\boldsymbol{x}) \leq T_+[f](\boldsymbol{x}) - \frac{\lambda_{min}[f]}{2}\left(\sum_{j=1}^{m} \lambda_j(\boldsymbol{x}) \left\|\boldsymbol{v}_j\right\|^2 - \left\|\boldsymbol{x}\right\|^2\right).$$

Also, from the fact that $\boldsymbol{x} = \sum_{j=1}^{m} \lambda_j(\boldsymbol{x})\boldsymbol{v}_j$, we have

$$\sum_{j=1}^{m} \lambda_j(\boldsymbol{x}) \left\|\boldsymbol{x} - \boldsymbol{v}_j\right\|^2 = \sum_{j=1}^{m} \lambda_j(\boldsymbol{x}) \left\|\boldsymbol{v}_j\right\|^2 - \left\|\boldsymbol{x}\right\|^2.$$

Let us now write $\boldsymbol{x} - \boldsymbol{v}_j = \sum_{i=1}^{m} \lambda_i(\boldsymbol{x})(\boldsymbol{v}_i - \boldsymbol{v}_j)$, we then have

$$\sum_{i=1}^{m} \lambda_i(\boldsymbol{x}) \left\|\boldsymbol{x} - \boldsymbol{v}_i\right\|^2 = \sum_{i,j=1}^{m} \lambda_i(\boldsymbol{x})\lambda_j(\boldsymbol{x}) \left\langle \boldsymbol{v}_i - \boldsymbol{v}_j, \boldsymbol{x} - \boldsymbol{v}_j \right\rangle. \tag{6.8}$$

By using symmetry of indices i, j, we can write the above equation as:

$$\sum_{i=1}^{m} \lambda_i(\boldsymbol{x}) \left\|\boldsymbol{x} - \boldsymbol{v}_i\right\|^2 = \sum_{i,j=1}^{m} \lambda_i(\boldsymbol{x})\lambda_j(\boldsymbol{x}) \left\langle \boldsymbol{v}_j - \boldsymbol{v}_i, \boldsymbol{x} - \boldsymbol{v}_i \right\rangle. \tag{6.9}$$

Combining relations (6.8) and (6.9) completes the proof of Theorem 6.2. □

Remark 6.3. By examining the proof of Theorem 4.1 and 6.2 more carefully, we gain the following additional information. It can be shown that their conclusions should extend to classes of linear operators wider than those satisfying the normalization condition $G[1] = 1$. All we need is that $G[\boldsymbol{e}]$ sends Ω into itself (for which there is no restriction on L_j, see (4.4)). We remark in conclusion that, under the assumption that $\mathbf{0} \in \Omega$, all the inequalities of Theorem 4.1 and 6.2 still hold in case that the operator G belongs to $\mathcal{L}_+(A_+(\Omega))$ and satisfies $T[1] = c$, with c in the open interval $(0, 1)$, olny the coefficient $\frac{1}{m-1}$ may be replaced by any real value $\frac{\alpha}{m-1}$, with $\alpha \in (0, 1)$. The proof of this result can be developed by proceeding on the lines similar to result of Theorem 4.1.

We end with the following:

Remark 6.4. Let us finally mention that the work described by this paper was carried out by delivering a generalized multivariate *discrete* Jensen-type inequality. However, we have also extended such results to the continuous version, which we shall present in another paper.

References

[1] A. Abramovich M. Klaričić, M. Matić and J. Pečarić, *A variant of Jensen-Steffensen's inequality and quasi-arithmetic means*, J. Math. Anal. Appl. **307** (2005), 370–386.
[2] I. Babuška, G. Caloz and J. E. Osborn, *Special finite element methods for a class of second order elliptic problems with rough coefficients*, SIAM J. Numerical Analysis **31** (1994), 745–981.
[3] L. Bougoffa, *New inequalities about convex functions*, J. Inequal. Pure Appl. Math. **7** (2006), Art. 148.
[4] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, 2004.
[5] W. S. Cheung, A. Matković and J. Pečarić, *A variant of Jensen's inequality and generalized means*, J. Inequal. Pure Appl. Math. **7** (2006), Article 10.
[6] P. J. Ciarlet *The Finite Element Method for Elliptic Problems*, North Holland, , Amsterdam, 1979.
[7] G. Farin, *Curves and Surfaces for Computer Aided Geometric Design*, Academic Press, San Diego, 1993.
[8] I. Gavrea, *Some considerations on the monotonicity property of power means*, J. Inequal. Pure Appl. Math. **5** (2004), Article 93.
[9] A. Guessab and G. Schmeisser, *Negative definite cubature formulae, extremality and delaunay triangulation*, Constr. Approx. **31** (2010), 95–113. (2010)
[10] A. Guessab and G. Schmeisser, *Construction of positive definite cubature formulae and approximation of functions via Voronoi tessellations*, Adv Comput Math **32** (2010), 25–41.
[11] A. Guessab, O. Nouisser and G. Schmeisser, *Enhancement of the algebraic precision of a linear operator and consequences under positivity*, Positivity **13** (2009), 693–707.
[12] A. Guessab and G. Schmeisser, *A definiteness theory for cubature formulae of order two*, Constr. Approx. **24** (2006), 263–288.
[13] A. Guessab and G. Schmeisser, *Convexity results and sharp error estimates in approximate multivariate approximation*, Math. Comput. **73** (2004), 1365–1384.
[14] T. Ju, P. Liepa and J. Warren, *A general geometric construction of coordinates simplicial polytope*, Computer Aided Geometric Design, **24** (2007), 161–178.
[15] P. J. Kelly and M. L. Weiss, *Geometry and Convexity*, John Wiley & Sons, New York, 1979.
[16] J. A. Kalman, *Continuity and convexity of projections and barycentric coordinates in convex polyhedra*, Pacific J. of Math. **II** (1961), 1017–1022.
[17] S. Leorato, *A refined Jensen's inequality in Hilbert spaces and empirical approximations*, Journal of Multivariate Analysis **100** (2009), 1044–1060.
[18] M. Niezgoda, *A generalization of Mercer's result on convex functions*, Nonlinear Analysis **71** (2009), 2771–2779.
[19] A. Matković and J. Pečarić, *A variant of Jensen's inequality for convex functions of several variables*, J. Math. Inequal. **1** (2007), 45–51.
[20] A. Matković, J. Pečarić and I. Perić, *A variant of Jensen's inequality of Mercer's type for operators with applications*, Linear Algebra Appl. **418** (2006), 551–564.
[21] J. M. Melenk and I. Babuška, *The partition of unity finite element method: Basic theory and applications*, Computer Methods in Applied Mechanics and Engineering **139** (2006), 289–314.
[22] A. McD. Mercer, *A variant of Jensen's inequality*, JIPAM **4** (2003), Article 73.
[23] A. F. Möbius, *Der barycentrische Calcul*, Johann Ambrosius Barth, Leipzig, 1827.
[24] R. Schneider, *Convex bodies: the Brunn-Minkowski theory*, Encyclopedia of Mathematics and its Applications, 44, Cambridge University Press, Cambridge, 1993.
[25] A. Schrijver, *Theory of Linear and Integer Programming*, Wiley, Chichester, 1987.
[26] G. M. Ziegler, *Lectures on Polytopes*, Springer Verlag, New York, 1995.
[27] J. Warren, *Barycentric coordinates for convex polytopes*, Advances in Computational Mathematics **6** (1996), 97–108.
[28] R. Webster, *Convexity*, Oxford University Press, Oxford, 1994

Manuscript received ,

revised ,

Allal Guessab
Laboratoire de Mathématiques et de leurs Applications, UMR CNRS 4152,
Université de Pau et des Pays de l'Adour, 64000 Pau, France
E-mail address: allal.guessab@univ-pau.fr