Journal of Approximation Theory **97**, 158–180 (1999)
Article ID jath.1997.3252, available online at http://www.idealibrary.com on IDEAL®

Best Constants in Global Smoothness Preservation Inequalities for Some Multivariate Operators*

Jesús de la Cal† and Ana M. Valle‡

Departamento de Matemática Aplicada y Estadística e Investigación Operativa, Facultad de Ciencias, Universidad del País Vasco, Apartado 644, 48080 Bilbao, Spain
E-mail: †mepcaagj@lg.ehu.es and ‡mepvamaa@lg.ehu.es

Communicated by Dany Leviatan

Received March 25, 1997; accepted in revised form January 29, 1998

DEDICATED TO THE MEMORY OF PROFESSOR E. APARICIO

The Bernstein operator on the standard k-simplex and other analogous k-variate operators allow for a probabilistic representation in terms of the successive increments of a real valued superstationary stochastic process (a notion introduced in the paper) starting at the origin and having nondecreasing paths. For this class of operators, we obtain estimates of the best constants in preservation of the first modulus of continuity corresponding to the l_1-norm, and in preservation of classes of functions defined by concave moduli of continuity. We also show that, in some special cases, such best constants do not depend upon the dimension k. To show our results, we use probabilistic tools such as couplings and Wasserstein distances for multivariate probability distributions. The general results are applied to the computation of the aforementioned constants for several classical multivariate operators.

1. INTRODUCTION AND MAIN RESULTS

Throughout this paper, $I^{\langle k \rangle}$ $(k=1, 2, ...)$ will denote the convex subset of $\mathbb{R}^k$ given by

$$I^{\langle k \rangle} := \{\mathbf{x} = (x_1, ..., x_k) \in I^k : x_1 + \cdots + x_k \in I\},$$

where I is either the interval $[0, 1]$ or the interval $[0, \infty)$ (obviously, in the first case, $I^{\langle k \rangle}$ is the standard k-simplex, whilst, in the second case, $I^{\langle k \rangle}$

* Research supported by Grant PB95-0809 of the Spanish DGICYT and by the University of the Basque Country.

0021-9045/99 $30.00

coincides with I^k), and $M^{\langle k \rangle}$ will be the set of all real nonconstant measurable functions defined on $I^{\langle k \rangle}$ such that

$$\omega(f;\delta) < \infty, \qquad \delta \geqslant 0,$$

where $\omega(f;\cdot)$ stands for the usual modulus of continuity of f with respect to the l_1-norm on $\mathbb{R}^k$ (the only norm to be used here), i.e.,

$$\omega(f;\delta) := \sup\{|f(\mathbf{x}) - f(\mathbf{y})|: \mathbf{x}, \mathbf{y} \in I^{\langle k \rangle}, \|\mathbf{x}-\mathbf{y}\| \leqslant \delta\},$$

with

$$\|\mathbf{x}\| := \sum_{i=1}^{k} |x_i|, \qquad \mathbf{x} := (x_1, ..., x_k) \in \mathbb{R}^k.$$

On the other hand, let ω be a nonnull real function defined on $[0, \infty)$ which vanishes at zero and is nondecreasing and concave (these conditions imply that ω is continuous on $(0, \infty)$ and subadditive). Such a function will be called a concave modulus of continuity (c.m.c., for brevity). We shall denote by $M_\omega^{\langle k \rangle}$ the class of functions

$$M_\omega^{\langle k \rangle} := \{f \in M^{\langle k \rangle}: \omega(f;\cdot) \leqslant \omega(\cdot)\}.$$

The class $M_\omega^{\langle k \rangle}$ was introduced by Nikol'skii [18, 19] (see also the book by Korneichuk [16]). In the important particular case corresponding to the c.m.c. ω_α given by

$$\omega_\alpha(\delta) := \delta^\alpha, \qquad \delta \geqslant 0 \quad (\alpha \in (0, 1]),$$

$M_{\omega_\alpha}^{\langle k \rangle}$ is the Lipschitz class of order α (and constant 1).

Let T be a nonempty set (usually the set of natural numbers, the interval $(0, \infty)$, or a suitable set of multiindices (see Section 5 below)) and let $\{L_t^{\langle k \rangle}: t \in T\}$ be a family of positive linear operators assigning a real function on $I^{\langle k \rangle}$ to each $f \in M^{\langle k \rangle}$. The problem of investigating the preservation of global smoothness by the operators $L_t^{\langle k \rangle}$ can be described as the problem of finding sharp estimates (even better, the exact value) of the C-constants and the D-constants defined by

$$C_t^{\langle k \rangle}(\delta) := \sup_{f \in M^{\langle k \rangle}} \frac{\omega(L_t^{\langle k \rangle} f;\delta)}{\omega(f;\delta)}, \qquad \delta > 0, \quad t \in T,$$

$$D_t^{\langle k \rangle}(\omega) := \sup_{\delta > 0} \sup_{f \in M_\omega^{\langle k \rangle}} \frac{\omega(L_t^{\langle k \rangle} f;\delta)}{\omega(\delta)}, \qquad t \in T,$$

and of the corresponding uniform constants (in δ, t, etc.).

Problems of this kind have been considered by several authors in the last decade (see, for instance, [1, 2, 5–8, 15] and the references therein). However, until now, the problem of estimating the D-constants has been only discussed for Lipschitz classes, that is, for the moduli of continuity ω_α above.

In the present paper, we obtain estimates of the C-constants and the D-constants for families $\{L_t^{\langle k\rangle}: t\in T\}$ of k-dimensional operators allowing for a probabilistic representation of the following type: There exists a double-indexed stochastic process $Z:=\{Z_t(x): x\in I, t\in T\}$ defined on a complete probability space $(\Omega, \mathscr{F}, P)$ which fulfills the three conditions

(A) All the random variables $Z_t(x)$ are I-valued and integrable.

(B) For each $t\in T$, we have $Z_t(0)=0$ a.s.

(C) For all $t\in T$ and $x, y\in I$ with $x\leqslant y$, we have $Z_t(x)\leqslant Z_t(y)$ a.s.,

and such that

$$L_t^{\langle k\rangle} f(\mathbf{x})=Ef(\mathbf{Z}_t^{\langle k\rangle}(\mathbf{x})), \qquad \mathbf{x}\in I^{\langle k\rangle}, \quad t\in T, \tag{1}$$

where E denotes mathematical expectation, $\mathbf{Z}_t^{\langle k\rangle}(\mathbf{x})$ is the $I^{\langle k\rangle}$-valued random vector

$$\mathbf{Z}_t^{\langle k\rangle}(\mathbf{x}):=(Z_{t,1}^{\langle k\rangle}(\mathbf{x}), ..., Z_{t,k}^{\langle k\rangle}(\mathbf{x})),$$

whose components are given by the increments

$$Z_{t,j}^{\langle k\rangle}(\mathbf{x}):=Z_t\left(\sum_{i=1}^{j} x_i\right)-Z_t\left(\sum_{i=1}^{j-1} x_i\right), \qquad j=1, ..., k \tag{2}$$

($\sum_{i=1}^{0} x_i$ being understood as 0) and f is any real function on $I^{\langle k\rangle}$ for which the right-hand side in (1) is well defined. In particular, it is easy to check that $L_t^{\langle k\rangle} f$ is well defined for every $f\in M^{\langle k\rangle}$.

This mathematical framework is general enough to include many families of multivariate positive linear operators usually considered in the literature on approximation theory. Take, for instance, $I=[0,1]$, $T:=\{1, 2, ...\}$, and let $Z:=\{Z_n(x): x\in I, n\geqslant 1\}$ be the process defined by

$$Z_n(x):=\frac{1}{n}\sum_{i=1}^{n} 1_{[0,x]}(X_i), \tag{3}$$

where $\{X_i: i\geqslant 1\}$ is a sequence of independent and on the interval $[0,1]$ uniformly distributed random variables, and 1_A denotes the indicator function of the set A. Then, the random vector $n\mathbf{Z}_n^{\langle k\rangle}(\mathbf{x})$ has the multinomial distribution with parameters n, $\mathbf{x}$ (see [13]), and $L_n^{\langle k\rangle}$ becomes the classical Bernstein operator over the standard k-simplex $I^{\langle k\rangle}$, i.e.,

$$L_n^{\langle k\rangle} f(\mathbf{x}) := \sum_{\mathbf{u} \in U_n} f\left(\frac{\mathbf{u}}{n}\right) \frac{n!}{\prod_{i=0}^k u_i!} \prod_{i=0}^k x_i^{u_i}, \tag{4}$$

where

$$x_0 := 1 - \sum_{i=1}^k x_i, \qquad u_0 := n - \sum_{i=1}^k u_i,$$

and U_n is the set of all k-tuples of nonnegative integers $\mathbf{u} := (u_1, ..., u_k)$ such that $\sum_{i=1}^k u_i \leqslant n$. Additional examples can be found in Section 5 below.

For univariate operators of the above type, the problem of global smoothness preservation has been considered in several works [1, 2, 5]. In a recent paper, Adell and Pérez-Palomares [5] provide interesting characterizations of the constants $C_t^{\langle 1\rangle}(\delta)$ and $D_t^{\langle 1\rangle}(\omega_\alpha)$ in terms of the stochastic process Z, in the case in which it satisfies the following additional assumption:

(D) For each $t \in T$, the process $\{Z_t(x) : x \in I\}$ has stationary increments, i.e., for all $x, y \in I$ with $x \leqslant y$, the random variables $Z_t(y) - Z_t(x)$ and $Z_t(y-x) - Z_t(0)$ have the same probability distribution.

They concretely show the following theorem, in which $\lceil \cdot \rceil$ denotes the standard "ceiling" function, i.e.,

$$\lceil x \rceil := \text{the smallest integer not less than } x, \qquad x \in \mathbb{R}. \tag{5}$$

THEOREM A. *Assume that the process Z satisfies conditions* (A)–(D). *Then, for every $t \in T$, we have*

(a) $C_t^{\langle 1\rangle}(\delta) = E\lceil Z_t(\delta)/\delta \rceil = \sum_{i=0}^\infty P(Z_t(\delta) > i\delta)$, $\delta \in I - \{0\}$.

(b) $D_t^{\langle 1\rangle}(\omega_\alpha) = \sup_{\delta \in I - \{0\}} E(Z_t(\delta)/\delta)^\alpha$, $\alpha \in (0, 1]$.

In the same work, the authors succeed in finding the exact values of $C_t^{\langle 1\rangle}(\delta)$, $D_t^{\langle 1\rangle}(\omega_\alpha)$, $\sup_{\delta>0} C_t^{\langle 1\rangle}(\delta)$, etc., for a number of well known univariate Bernstein-type operators, by combining the preceding formulae with the particular properties of the processes used to represent the operators.

Our first result extends Theorem A(b) to general concave moduli of continuity. Although the proof is essentially the same as for ω_α, we include the details for the sake of completeness.

THEOREM 1. *Under the same assumptions as in Theorem A, we have, for every $t \in T$ and every c.m.c. ω,*

$$D_t^{\langle 1 \rangle}(\omega) = \sup_{\delta \in I - \{0\}} E\left(\frac{\omega(Z_t(\delta))}{\omega(\delta)}\right).$$

Proof. Let $t \in T$ and ω be fixed. Let $f \in M_\omega^{\langle 1 \rangle}$, $\delta \in I - \{0\}$ and $x, y \in I$ with $0 < y - x \leqslant \delta$. Using successively (1) together with (B), the hypothesis on f, assumptions (C), (D), (B) and again assumption (C), we have

$$\begin{aligned} |L_t^{\langle 1 \rangle} f(y) - L_t^{\langle 1 \rangle} f(x)| &\leqslant E\, |f(Z_t(y)) - f(Z_t(x))| \leqslant E\omega(Z_t(y) - Z_t(x)) \\ &= E\omega(Z_t(y-x) - Z_t(0)) \\ &= E\omega(Z_t(y-x)) \leqslant E\omega(Z_t(\delta)). \end{aligned}$$

We conclude that

$$\omega(L_t^{\langle 1 \rangle} f; \delta) \leqslant E\omega(Z_t(\delta)) = E\left(\frac{\omega(Z_t(\delta))}{\omega(\delta)}\right) \omega(\delta),$$

and, therefore,

$$D_t^{\langle 1 \rangle}(\omega) \leqslant \sup_{\delta \in I - \{0\}} E\left(\frac{\omega(Z_t(\delta))}{\omega(\delta)}\right).$$

The converse inequality follows from the fact that $\omega \in M_\omega^{\langle 1 \rangle}$, and we have, for all $\delta \in I - \{0\}$,

$$\omega(L_t^{\langle 1 \rangle} \omega; \delta) \geqslant |L_t^{\langle 1 \rangle} \omega(\delta) - L_t^{\langle 1 \rangle} \omega(0)| = E\omega(Z_t(\delta)) - \omega(0) = E\omega(Z_t(\delta)).$$

This completes the proof of Theorem 1. ▮

In the investigation of the multivariate operators $L_t^{\langle k \rangle}$ given in (1), new technical problems arise, and we have found it necessary, on the one hand, to distinguish the two cases $I = [0, 1]$ and $I = [0, \infty)$, and, on the other hand, to replace assumption (D) on Z by the more restrictive condition

(D′) For each $t \in T$, the process $\{Z_t(x) \colon x \in I\}$ is superstationary.

(On these points, see Lemma 2 and Remark 3 below.) To the best of our knowledge, the notion of a superstationary process, as it is defined in Section 2 below, is new in the literature on probability theory. Section 2 also contains some results concerning superstationary processes which will be used in the paper.

Using the above notations and concepts, our main results can be summarized as follows.

THEOREM 2. *Let $t \in T$, $\delta > 0$ and let ω be a concave modulus of continuity. Assume that the process Z satisfies conditions* (A)–(C). *Then*:

(I) *The sequences $\{C_t^{\langle k\rangle}(\delta): k \geqslant 1\}$ and $\{D_t^{\langle k\rangle}(\omega): k \geqslant 1\}$ are non-decreasing.*

(II) *Assume, further, that Z satisfies condition* (D′). *We have*

(IIa) *If $EZ_t(x) \leqslant \kappa_t x$ (with $\kappa_t > 0$), then, for all $k \geqslant 1$,*

$$C_t^{\langle k\rangle}(\delta) \leqslant 1 + \kappa_t, \qquad D_t^{\langle k\rangle}(\omega) \leqslant \sup_{\delta > 0} \frac{\omega(\kappa_t \delta)}{\omega(\delta)}.$$

(IIb) *If $I = [0, \infty)$, then, for all $k \geqslant 1$,*

$$C_t^{\langle k\rangle}(\delta) = C_t^{\langle 1\rangle}(\delta), \qquad D_t^{\langle k\rangle}(\omega) = D_t^{\langle 1\rangle}(\omega).$$

The proof of Theorem 2 is given in Section 4. It is based on Lemma 2 (in Section 3) concerning couplings and Wasserstein distances for multivariate probability distributions associated with superstationary processes. Finally, Section 5 contains applications of the preceding results to some well-known families of multivariate operators.

2. SUPERSTATIONARY PROCESSES

The following definition introduces the notion of a superstationary stochastic process.

DEFINITION. A real-valued stochastic process $\{S(x): x \in I\}$ is called "superstationary" if it satisfies the following condition: For each $k = 1, 2, ...$ and each finite family $\mathscr{I}$ of pairwise disjoint subintervals of I having the form

$$\mathscr{I} := \{(a_{ij}, b_{ij}]: 1 \leqslant i \leqslant k, 1 \leqslant j \leqslant r_i\},$$

where $r_1, ..., r_k$ are arbitrary natural numbers, the k-dimensional random vector $(S_1, ..., S_k)$ given by

$$S_i := \sum_{j=1}^{r_i} (S(b_{ij}) - S(a_{ij})), \qquad 1 \leqslant i \leqslant k,$$

has the same probability distribution as $(S'_1, ..., S'_k)$, where, for $1 \leqslant i \leqslant k$,

$$S'_i := S\left(\sum_{j=1}^{i} x_j\right) - S\left(\sum_{j=1}^{i-1} x_j\right)$$

and

$$x_i := \sum_{j=1}^{r_i} (b_{ij} - a_{ij}).$$

In other words, the distribution of $(S_1, ..., S_k)$ depends on the family $\mathscr{I}$ only through the vector $(x_1, ..., x_k)$.

Remark 1. Obviously, every superstationary process has stationary increments. Therefore, condition (D′) implies condition (D).

Remark 2. It is easy to check that any stochastic process having stationary independent increments is superstationary. On the other hand, for each $n \geqslant 1$, the process $\{Z_n(x): x \in [0, 1]\}$ constructed in (3) to represent the Bernstein operator is superstationary. To see this, observe that if $X_1, ..., X_n$ are the same as in (3), and $A_1, ..., A_k$ are pairwise disjoint Borel subsets of $[0, 1]$ whose respective lengths (Lebesgue measures) are $x_1, ..., x_k$, then the random vector $(S_1, ..., S_k)$ with

$$S_i := \sum_{j=1}^{n} 1_{A_i}(X_j), \qquad i = 1, ..., k,$$

has the multinomial distribution with parameters n, $x_1, ..., x_k$. This argument applies, in particular, to the sets

$$A_i := \bigcup_{j=1}^{r_i} (a_{ij}, b_{ij}], \qquad i = 1, ..., k,$$

when

$$\mathscr{I} := \{(a_{ij}, b_{ij}]: 1 \leqslant i \leqslant k, 1 \leqslant j \leqslant r_i\},$$

is the same family as in the definition of a superstationary process.

The following proposition (in which J also denotes either the interval $[0, 1]$ or the interval $[0, \infty)$) provides a method to construct new superstationary processes from given ones. Both Proposition 1 and its subsequent corollary will be useful in Section 5 below.

PROPOSITION 1. *Let* $S := \{S(v): v \in J\}$ *and* $V := \{V(x): x \in I\}$ *be two independent superstationary processes defined on the same probability space. Assume further that* V *satisfies the following two conditions*:

(a) *For each* $x \in I$, *the random variable* $V(x)$ *takes values in* J.

(b) *For all* $x, y \in I$ *with* $x < y$, *we have* $V(x) \leqslant V(y)$ *a.s.*

Then, the process $W := \{S(V(x)): x \in I\}$ *is superstationary.*

Proof. Let $k \geqslant 1$, let

$$\mathscr{I} := \{(a_{ij}, b_{ij}] \colon 1 \leqslant i \leqslant k,\ 1 \leqslant j \leqslant r_i\},$$

be a finite family of pairwise disjoint subintervals of I ($r_1, ..., r_k$ being arbitrary natural numbers), and let $(W_1, ..., W_k)$ be the k-dimensional random vector given by

$$W_i := \sum_{j=1}^{r_i} (S(V(b_{ij})) - S(V(a_{ij}))), \qquad 1 \leqslant i \leqslant k.$$

From assumptions (a) and (b) on V, the random subintervals of J

$$(V(a_{ij}), V(b_{ij})], \qquad 1 \leqslant i \leqslant k, \qquad 1 \leqslant j \leqslant r_i,$$

are pairwise disjoint a.s. Therefore, by independence and the superstationarity of S, the probability distribution of $(W_1, ..., W_k)$ depends on $\mathscr{I}$ only through the distribution of the random vector $(X_1, ..., X_k)$, where

$$X_i := \sum_{j=1}^{r_i} (V(b_{ij}) - V(a_{ij})).$$

From the superstationarity of V, we conclude that the distribution of $(X_1, ..., X_k)$, and consequently that of $(W_1, ..., W_k)$, depends on the family $\mathscr{I}$ only through the vector $(x_1, ..., x_k)$, where

$$x_i := \sum_{j=1}^{r_i} (b_{ij} - a_{ij}).$$

This completes the proof. ∎

The following corollary is immediate.

COROLLARY 1. *Let $S := \{S(x) \colon x \in I\}$ be a superstationary process and let V be any* **I***-valued random variable independent of S and defined on the same probability space. Then, the process $\{S(xV) \colon x \in I\}$ is superstationary.*

3. COUPLINGS AND WASSERSTEIN DISTANCES

In what follows, we will denote by $\mathbf{U} \simeq \mathbf{V}$ the fact that $\mathbf{U}$ and $\mathbf{V}$ are random vectors having the same probability distribution.

Let $\mathbf{X}$ and $\mathbf{Y}$ be two integrable k-dimensional random vectors not necessarily defined on the same probability space. Each $\mathbb{R}^k \times \mathbb{R}^k$-valued random vector $(\mathbf{U}, \mathbf{V})$ such that $\mathbf{U} \simeq \mathbf{X}$ and $\mathbf{V} \simeq \mathbf{Y}$ is called a "coupling" for

$(\mathbf{X}, \mathbf{Y})$. The (l_1-)Wasserstein distance between (the probability distributions of) $\mathbf{X}$ and $\mathbf{Y}$ (relative to the l_1-norm $\|\cdot\|$) is defined by

$$d(\mathbf{X}, \mathbf{Y}) := \inf E\, \|\mathbf{U} - \mathbf{V}\|, \tag{6}$$

where the infimum is taken over all the couplings for $(\mathbf{X}, \mathbf{Y})$. It is well-known that the infimum in (6) is attained (see, for instance, [11]). If the infimum is attained at $(\mathbf{X}^*, \mathbf{Y}^*)$, then we say that this random vector is an "optimal coupling" for $d(\mathbf{X}, \mathbf{Y})$.

From the definition, it is clear that

$$d(\mathbf{X}, \mathbf{Y}) \geqslant \|E\mathbf{X} - E\mathbf{Y}\|,$$

where $E\mathbf{X} = (EX_1, ..., EX_k)$ (X_i being the ith component of $\mathbf{X}$). The following lemma characterizes the case in which the equality takes place. It generalizes [3, Lemma 1].

LEMMA 1. *Let $\mathbf{X}$ and $\mathbf{Y}$ be two integrable k-dimensional random vectors and let $(\mathbf{X}^*, \mathbf{Y}^*)$ be a coupling for $(\mathbf{X}, \mathbf{Y})$. Denote by X_i^* (resp. Y_i^*) the ith component of $\mathbf{X}^*$ (resp. $\mathbf{Y}^*$). Then, the following two conditions are equivalent:*

(a) *For each fixed $i \in \{1, 2, ..., k\}$, we have $X_i^* \leqslant Y_i^*$ a.s. or $Y_i^* \leqslant X_i^*$ a.s.*

(b) *$(\mathbf{X}^*, \mathbf{Y}^*)$ is an optimal coupling for $d(\mathbf{X}, \mathbf{Y})$ and $d(\mathbf{X}, \mathbf{Y}) = \|E\mathbf{X} - E\mathbf{Y}\|$.*

Proof. The conclusion is easily derived by using the following elementary fact: Given an integrable random variable X, the equality $E|X| = |EX|$ holds true if and only if we have either $X \geqslant 0$ a.s. or $X \leqslant 0$ a.s. Details are omitted. ∎

The following result will play a crucial role in the proof of Theorem 2(II).

LEMMA 2. *Let Z be the same process as in Theorem 2(II), and let $t \in T$ and $k \geqslant 1$ be fixed. For all $\mathbf{x}, \mathbf{y} \in I^{\langle k \rangle}$, we have*

$$d(\mathbf{Z}_t^{\langle k \rangle}(\mathbf{x}), \mathbf{Z}_t^{\langle k \rangle}(\mathbf{y})) = \|E\mathbf{Z}_t^{\langle k \rangle}(\mathbf{x}) - E\mathbf{Z}_t^{\langle k \rangle}(\mathbf{y})\| = \sum_{i=1}^{k} EZ_t(|x_i - y_i|), \tag{7}$$

where x_i (resp. y_i) is the ith coordinate of $\mathbf{x}$ (resp. $\mathbf{y}$). Furthermore, if $I = [0, \infty)$, there exists an optimal coupling $(\mathbf{X}^*, \mathbf{Y}^*)$ for $d(\mathbf{Z}_t^{\langle k \rangle}(\mathbf{x}), \mathbf{Z}_t^{\langle k \rangle}(\mathbf{y}))$ such that

$$\|\mathbf{X}^* - \mathbf{Y}^*\| \simeq Z_t(\|\mathbf{x} - \mathbf{y}\|). \tag{8}$$

Proof. The second equality in (7) directly follows from assumptions (A)–(C) on Z and the fact that, by assumption (D′), the process $\{Z_t(x)\colon x \in I\}$

has stationary increments. By Lemma 1, to show the first equality in (7), we only need to construct a coupling for $(\mathbf{Z}_t^{\langle k\rangle}(\mathbf{x}), \mathbf{Z}_t^{\langle k\rangle}(\mathbf{y}))$ satisfying the requirements in Lemma 1(a). For $i=1, ..., k$, denote by

$$u_i := \min(x_i, y_i),$$

$$v_i := \begin{cases} x_i - y_i & \text{if} \quad x_i \geqslant y_i \\ 0 & \text{otherwise} \end{cases} \qquad w_i := \begin{cases} y_i - x_i & \text{if} \quad x_i < y_i \\ 0 & \text{otherwise.} \end{cases}$$

Let $\mathbf{U} := (U_1, ..., U_k)$, $\mathbf{V} := (V_1, ..., V_k)$ and $\mathbf{W} := (W_1, ..., W_k)$ be the k-dimensional random vectors whose respective jth components $(j=1, 2, ..., k)$ are given by

$$U_j := Z_t\left(\sum_{i=1}^{j} u_i\right) - Z_t\left(\sum_{i=1}^{j-1} u_i\right),$$

$$V_j := Z_t\left(\sum_{i=1}^{k} u_i + \sum_{i=1}^{j} v_i\right) - Z_t\left(\sum_{i=1}^{k} u_i + \sum_{i=1}^{j-1} v_i\right),$$

$$W_j := Z_t\left(\sum_{i=1}^{k} u_i + \sum_{i=1}^{j} w_i\right) - Z_t\left(\sum_{i=1}^{k} u_i + \sum_{i=1}^{j-1} w_i\right).$$

We claim that the random vector $(\mathbf{U}+\mathbf{V}, \mathbf{U}+\mathbf{W})$ is a coupling for $(\mathbf{Z}_t^{\langle k\rangle}(\mathbf{x}), \mathbf{Z}_t^{\langle k\rangle}(\mathbf{y}))$ satisfying the requirements in Lemma 1(a). To show this, we firstly observe that the $2k$ subintervals of I

$$\left(\sum_{i=1}^{j-1} u_i, \sum_{i=1}^{j} u_i\right], \qquad j=1, 2, ..., k, \tag{9}$$

$$\left(\sum_{i=1}^{k} u_i + \sum_{i=1}^{j-1} v_i, \sum_{i=1}^{k} u_i + \sum_{i=1}^{j} v_i\right], \qquad j=1, 2, ..., k, \tag{10}$$

are pairwise disjoint, and, moreover, for $j=1, 2, ..., k$, we have that $u_j + v_j = x_j$. We therefore conclude from assumption (D′) that $\mathbf{U}+\mathbf{V} \simeq \mathbf{Z}_t^{\langle k\rangle}(\mathbf{x})$. Analogously, we also have that $\mathbf{U}+\mathbf{W} \simeq \mathbf{Z}_t^{\langle k\rangle}(\mathbf{y})$, showing that $(\mathbf{U}+\mathbf{V}, \mathbf{U}+\mathbf{W})$ is a coupling for $(\mathbf{Z}_t^{\langle k\rangle}(\mathbf{x}), \mathbf{Z}_t^{\langle k\rangle}(\mathbf{y}))$. On the other hand, by assumption (C), the components of $\mathbf{U}$, $\mathbf{V}$ and $\mathbf{W}$ are (a.s.) nonnegative random variables and, for each fixed $j=1, 2, ..., k$, we have either $V_j=0$ or $W_j=0$ according to $x_j < y_j$ or $x_j \geqslant y_j$. This shows the claim and completes the proof of (7).

Finally, we show (8). If $I=[0, \infty)$, we can take the random vector $(\mathbf{X}^*, \mathbf{Y}^*)$ given by

$$\mathbf{X}^* := \mathbf{U}+\mathbf{V}, \qquad \mathbf{Y}^* := \mathbf{U}+\mathbf{W}',$$

where $\mathbf{U}$ and $\mathbf{V}$ are the same as above, and $\mathbf{W}' := (W'_1, ..., W'_k)$ is defined by

$$W'_j := Z_t\left(\sum_{i=1}^{k} (u_i + v_i) + \sum_{i=1}^{j} w_i\right) - Z_t\left(\sum_{i=1}^{k} (u_i + v_i) + \sum_{i=1}^{j-1} w_i\right),$$

for $j = 1, ..., k$. It is seen as before that $(\mathbf{X}^*, \mathbf{Y}^*)$ is an optimal coupling for $d(\mathbf{Z}_t^{\langle k \rangle}(\mathbf{x}), \mathbf{Z}_t^{\langle k \rangle}(\mathbf{y}))$. Moreover, since the $3k$ subintervals of $[0, \infty)$ given in (9), (10) and

$$\left(\sum_{i=1}^{k} (u_i + v_i) + \sum_{i=1}^{j-1} w_i, \sum_{i=1}^{k} (u_i + v_i) + \sum_{i=1}^{j} w_i\right], \qquad j = 1, 2, ..., k$$

are pairwise disjoint, we have from assumption (D′)

$$\begin{aligned} \|\mathbf{X}^* - \mathbf{Y}^*\| &= \sum_{j=1}^{k} (V_j + W'_j) \simeq Z_t\left(\sum_{j=1}^{k} (v_j + w_j)\right) - Z_t(0) \\ &= Z_t(\|\mathbf{x} - \mathbf{y}\|) - Z_t(0). \end{aligned} \tag{11}$$

By assumption (B), we obtain (8), and the proof of Lemma 3 is complete.

Remark 3. In the preceding proof, assumption (D′) is used to guarantee (11) and the facts that the random vectors $(\mathbf{U} + \mathbf{V}, \mathbf{U} + \mathbf{W})$ and $(\mathbf{X}^*, \mathbf{Y}^*)$ are couplings for $(\mathbf{Z}_t^{\langle k \rangle}(\mathbf{x}), \mathbf{Z}_t^{\langle k \rangle}(\mathbf{y}))$. When $k \geqslant 2$, these facts cannot be derived from the less restrictive assumption (D) which is basically a one-dimensional condition, and does not contain any information about the distribution of multidimensional random vectors whose components are increments of the process $\{Z_t(x) \colon x \in I\}$. On the other hand, the direct analogue of relation (8) does not remain valid if $[0, \infty)$ is replaced by $[0, 1]$, $k \geqslant 2$ and $\|\mathbf{x} - \mathbf{y}\| > 1$. Actually, in such a case, the right-hand side in (8) is not defined. In the same direction, observe that the construction of $\mathbf{W}'$, as done in the last part of the preceding proof, makes sense if and only if $\sum_{j=1}^{k} (u_j + v_j + w_j) = \sum_{j=1}^{k} \max(x_j, y_j) \in I$.

4. PROOF OF THEOREM 2

Let $t \in T$, $k \geqslant 1$, and $\delta > 0$ be fixed, and let $\omega(\cdot)$ be a c.m.c..

Proof of (I). To show that

$$C_t^{\langle k \rangle}(\delta) \leqslant C_t^{\langle k+1 \rangle}(\delta) \qquad \text{and} \qquad D_t^{\langle k \rangle}(\omega) \leqslant D_t^{\langle k+1 \rangle}(\omega), \tag{12}$$

we define a positive linear operator H in the following way: For any real function f on $I^{\langle k\rangle}$, Hf is the real function on $I^{\langle k+1\rangle}$ given by

$$(Hf)(x_1, ..., x_{k+1}) := f(x_1, ..., x_k).$$

It is clear that H is injective and fulfills the following conditions:

$$\begin{aligned}
&\omega(Hf; \cdot) = \omega(f; \cdot),\\
&H(M^{\langle k\rangle}) \subset M^{\langle k+1\rangle},\\
&H(M_\omega^{\langle k\rangle}) \subset M_\omega^{\langle k+1\rangle},\\
&L_t^{\langle k+1\rangle}(Hf) = H(L_t^{\langle k\rangle} f), \qquad f \in M^{\langle k\rangle}.
\end{aligned}$$

From these facts, the inequalities in (12) immediately follow.

Proof of (IIa). Let $\mathbf{x}, \mathbf{y} \in I^{\langle k\rangle}$ with $\|\mathbf{x}-\mathbf{y}\| \leqslant \delta$, and let $(\mathbf{X}^*, \mathbf{Y}^*)$ be an optimal coupling for $d(\mathbf{Z}_t^{\langle k\rangle}(\mathbf{x}), \mathbf{Z}_t^{\langle k\rangle}(\mathbf{y}))$. Then, for any $f \in M^{\langle k\rangle}$,

$$\begin{aligned}
|L_t^{\langle k\rangle} f(\mathbf{x}) - L_t^{\langle k\rangle} f(\mathbf{y})| &= |Ef(\mathbf{X}^*) - Ef(\mathbf{Y}^*)|\\
&\leqslant E\,|f(\mathbf{X}^*) - f(\mathbf{Y}^*)|\\
&\leqslant E\omega(f; \|\mathbf{X}^* - \mathbf{Y}^*\|).
\end{aligned} \tag{13}$$

Using successively the standard inequality

$$\omega(f; a\delta) \leqslant \{1 + a\}\ \omega(f; \delta), \qquad a, \delta \geqslant 0,$$

formula (7) and the hypothesis $EZ_t(x) \leqslant \kappa_t x$, we have

$$\begin{aligned}
|L_t^{\langle k\rangle} f(\mathbf{x}) - L_t^{\langle k\rangle} f(\mathbf{y})| &\leqslant \left\{1 + \frac{E\,\|\mathbf{X}^* - \mathbf{Y}^*\|}{\delta}\right\} \omega(f; \delta)\\
&\leqslant \left\{1 + \frac{\kappa_t\,\|\mathbf{x} - \mathbf{y}\|}{\delta}\right\} \omega(f; \delta)\\
&\leqslant \{1 + \kappa_t\}\ \omega(f; \delta),
\end{aligned}$$

implying that

$$C_t^{\langle k\rangle}(\delta) \leqslant 1 + \kappa_t.$$

On the other hand, if $f \in M_\omega^{\langle k\rangle}$, we have from (13) and Jensen's inequality

$$\begin{aligned}
|L_t^{\langle k\rangle} f(\mathbf{x}) - L_t^{\langle k\rangle} f(\mathbf{y})| &\leqslant E\omega(\|\mathbf{X}^* - \mathbf{Y}^*\|)\\
&\leqslant \omega(E\,\|\mathbf{X}^* - \mathbf{Y}^*\|)\\
&= \frac{\omega(E\,\|\mathbf{X}^* - \mathbf{Y}^*\|)}{\omega(\delta)}\,\omega(\delta).
\end{aligned}$$

Combining this fact with formula (7), the assumption on the expectation of $Z_t(x)$ and the nondecreasing character of $\omega(\cdot)$, we conclude that

$$D_t^{\langle k\rangle}(\omega) \leqslant \sup_{\delta>0} \frac{\omega(\kappa_t\,\delta)}{\omega(\delta)}.$$

Proof of (IIb). In view of Theorem A(a), Theorem 1 and part (I) above, we only need to show that

$$C_t^{\langle k\rangle}(\delta) \leqslant E\left\lceil \frac{Z_t(\delta)}{\delta} \right\rceil \qquad \text{and} \qquad D_t^{\langle k\rangle}(\omega) \leqslant \sup_{\delta>0} E\left(\frac{\omega(Z_t(\delta))}{\omega(\delta)}\right), \tag{14}$$

where $\lceil\cdot\rceil$ is the function defined in (5). To do this, let $\mathbf{x}, \mathbf{y} \in I^{\langle k\rangle}$ with $\|\mathbf{x}-\mathbf{y}\| \leqslant \delta$, and, recalling that $I=[0, \infty)$, let $(\mathbf{X}^*, \mathbf{Y}^*)$ be a coupling for $(\mathbf{Z}_t^{\langle k\rangle}(\mathbf{x}), \mathbf{Z}_t^{\langle k\rangle}(\mathbf{y}))$ satisfying (8). Proceeding as in the proof of part (IIa), we can write for any $f \in M^{\langle k\rangle}$

$$\begin{aligned} |L_t^{\langle k\rangle} f(\mathbf{x}) - L_t^{\langle k\rangle} f(\mathbf{y})| &\leqslant E\omega(f; \|\mathbf{X}^* - \mathbf{Y}^*\|) \\ &= E\omega(f; Z_t(\|\mathbf{x}-\mathbf{y}\|)) \\ &\leqslant E\omega(f; Z_t(\delta)), \end{aligned}$$

the last inequality by assumption (C) on Z. Therefore

$$\omega(L_t^{\langle k\rangle} f; \delta) \leqslant E\omega(f; Z_t(\delta)), \tag{15}$$

and, using the estimate

$$\omega(f; a\delta) \leqslant \lceil a \rceil\, \omega(f; \delta), \qquad a, \delta \geqslant 0,$$

we conclude that

$$\omega(L_t^{\langle k\rangle} f; \delta) \leqslant E\left\lceil \frac{Z_t(\delta)}{\delta} \right\rceil \omega(f; \delta),$$

showing the first inequality in (14). On the other hand, if $f \in M_\omega^{\langle k\rangle}$, we have from (15)

$$\omega(L_t^{\langle k\rangle} f; \delta) \leqslant E\left(\frac{\omega(Z_t(\delta))}{\omega(\delta)}\right) \omega(\delta),$$

which establishes the second inequality in (14). The proof of Theorem 2 is complete.

Remark 4. Let $C(I^{\langle k\rangle})$ be the set of all real continuous functions on $I^{\langle k\rangle}$ and define $\hat{C}_t^{\langle k\rangle}(\delta)$ in the same way as $C_t^{\langle k\rangle}(\delta)$ with $M^{\langle k\rangle}$ replaced by $M^{\langle k\rangle} \cap C(I^{\langle k\rangle})$. It is shown in [5] that

$$\hat{C}_t^{\langle 1\rangle}(\delta) = C_t^{\langle 1\rangle}(\delta), \qquad t \in T, \quad \delta > 0.$$

From this fact and the above proof, it becomes apparent that all the assertions in Theorem 2 concerning the constants $C_t^{\langle k\rangle}(\delta)$ remain true when each $C_t^{\langle k\rangle}(\delta)$ is replaced by the corresponding $\hat{C}_t^{\langle k\rangle}(\delta)$.

5. APPLICATIONS

In this section, we obtain the values of $C_t^{\langle k\rangle}(\delta)$, $D_t^{\langle k\rangle}(\omega)$, $\sup_{\delta>0} C_t^{\langle k\rangle}(\delta)$, etc., for seven well-known families of multivariate operators. Examples (A)–(D) correspond to operators on the standard k-simplex ($I = [0, 1]$), while examples (E)–(G) refer to operators on $[0, \infty)^k$ ($I = [0, \infty)$). In examples (A), (B), and (E)–(G), we make use of the results given in [5] for the respective univariate operators.

(A) *Bernstein Operators on Standard Simplices.* As it is said in Section 1, the operator $L_n^{\langle k\rangle}$ given in (4) is associated to the process $Z := \{Z_n(x) \colon x \in [0, 1], n \geqslant 1\}$, where $Z_n(x)$ is the uniform empirical distribution function described in (3). This process fulfills conditions (A)–(C) and (D′) (Remark 2). Moreover, we have

$$EZ_n(x) = x, \qquad x \in [0, 1], \qquad n \geqslant 1. \tag{16}$$

It is shown in [5] that, for every $n = 1, 2, ...,$

$$\sup_{\delta>0} C_n^{\langle 1\rangle}(\delta) = 2.$$

We therefore conclude from Theorem 2 that

$$\sup_{\delta>0} C_n^{\langle k\rangle}(\delta) = 2, \qquad n, k \geqslant 1. \tag{17}$$

We claim that, for all $n, k \geqslant 1$ (and any c.m.c. ω), we have

$$D_n^{\langle k\rangle}(\omega) = 1. \tag{18}$$

Actually, from (16) and Theorem 2(IIa), we obtain

$$D_n^{\langle k\rangle}(\omega) \leqslant 1.$$

On the other hand, by Theorem 2(I) and Theorem 1, we have

$$D_n^{\langle k\rangle}(\omega) \geqslant D_n^{\langle 1\rangle}(\omega) \geqslant E\left(\frac{\omega(Z_n(1))}{\omega(1)}\right) = 1,$$

the last equality by the fact that $Z_n(1) = 1$ (a.s.). The claim is shown.

Remark 5. The equality (17) was already obtained in [7] by using a different approach. In [5], it is shown that $D_n^{\langle 1\rangle}(\omega_\alpha) = 1$ $(n \geqslant 1, \alpha \in (0, 1])$. The inequality $D_n^{\langle k\rangle}(\omega_\alpha) \leqslant 1$ $(n, k \geqslant 1, \alpha \in (0, 1])$ was established in [15] (see also [8] for a different proof), but it should be observed that the probabilistic proof in [15, p. 313] is not correct. Actually, it is based upon the following claim: For $\mathbf{x}, \mathbf{y} \in I^{\langle k\rangle}$, there exists a $2k$-dimensional random vector having the multinomial distribution with parameters n, $u_1, ..., u_k$, $|x_1 - y_1|, ..., |x_k - y_k|$, where $u_i := \min(x_i, y_i)$ $(i = 1, ..., k)$. However, such a probability distribution does not exist if $\sum_{i=1}^k u_i + \sum_{i=1}^k |x_i - y_i| > 1$ (in particular, if $\|\mathbf{x} - \mathbf{y}\| > 1$).

(B) *Multivariate Beta Operators.* Let $Z := \{Z_t(x) \colon x \in [0, 1], t > 0\}$ be the double-indexed stochastic process given by

$$Z_t(x) := \frac{U_{tx}}{U_t},$$

where $\{U_t \colon t \geqslant 0\}$ is a standard gamma process, i.e., a stochastic process starting at the origin, having stationary independent increments, and such that, for $t > 0$, the random variable U_t has the gamma distribution with density

$$g_t(\theta) := \frac{\theta^{t-1} e^{-\theta}}{\Gamma(t)}, \qquad \theta > 0. \tag{19}$$

The process Z obviously fulfills conditions (A)–(C) and, using the properties of gamma distributions, it is not hard to check that it also satisfies condition (D′) (see, for instance, [20, pp. 458 and 481]). In this case, $I^{\langle k\rangle}$ is the standard k-simplex, as in the preceding example. Denote by $I_\circ^{\langle k\rangle}$ and $\partial I^{\langle k\rangle}$ the interior and the boundary of $I^{\langle k\rangle}$, respectively. If $t > 0$ and $\mathbf{x} = (x_1, ..., x_k) \in I_\circ^{\langle k\rangle}$, then the k-dimensional random vector $\mathbf{Z}_t^{\langle k\rangle}(\mathbf{x})$ given in (2) has the Dirichlet distribution with density

$$d_t(\mathbf{x}; \theta_1, ..., \theta_k) := \frac{\Gamma(t)}{\prod_{i=0}^k \Gamma(tx_i)} \prod_{i=0}^k \theta_i^{tx_i - 1}, \qquad (\theta_1, ..., \theta_k) \in I_\circ^{\langle k\rangle},$$

where

$$x_0 := 1 - \sum_{i=1}^{k} x_i, \qquad \theta_0 := 1 - \sum_{i=1}^{k} \theta_i$$

(see [14, 20]), and we have

$$L_t^{\langle k \rangle} f(\mathbf{x}) = \int \cdots \int_{I^{\langle k \rangle}} f(\theta_1, ..., \theta_k)\, d_t(\mathbf{x}; \theta_1, ..., \theta_k)\, d\theta_1 \cdots d\theta_k,$$

while, for $\mathbf{x} \in \partial I^{\langle k \rangle}$, the distribution of $\mathbf{Z}_t^{\langle k \rangle}(\mathbf{x})$ is singular with respect to the k-dimensional Lebesgue measure. In particular, if $\mathbf{x}$ is a vertex of $I^{\langle k \rangle}$, then $\mathbf{Z}_t^{\langle k \rangle}(\mathbf{x}) = \mathbf{x}$ a.s., so that $L_t^{\langle k \rangle} f$ interpolates f at the vertices of $I^{\langle k \rangle}$. The one-dimensional operator $L_t^{\langle 1 \rangle}$ is the beta operator introduced by Mühlbach [17]

$$L_t^{\langle 1 \rangle} f(x) = \begin{cases} \int_0^1 f(\theta)\, d_t(x; \theta)\, d\theta, & x \in (0, 1) \\ f(x), & x = 0, 1. \end{cases}$$

As in example (A), we have $EZ_t(x) = x$, $(t > 0,\ x \in [0, 1])$, and the facts

$$\sup_{\delta > 0} C_t^{\langle k \rangle}(\delta) = 2, \qquad t > 0, \quad k \geqslant 1,$$

and

$$D_t^{\langle k \rangle}(\omega) = 1, \qquad t > 0, \quad k \geqslant 1,$$

can be shown by using the same arguments as in the preceding example.

(C) *Multivariate Stancu Operators.* The multivariate Stancu operator [21, 22] $L_{n,a}^{\langle k \rangle}$ (a being a nonnegative parameter) on the standard k-simplex $I^{\langle k \rangle}$ is defined by

$$L_{n,a}^{\langle k \rangle} f(\mathbf{x}) := \sum_{\mathbf{u} \in U_n} f\left(\frac{\mathbf{u}}{n}\right) \frac{n!}{\prod_{i=0}^{k} u_i!} \frac{\prod_{i=0}^{k} x_i^{\langle u_i, a \rangle}}{1^{\langle n, a \rangle}},$$

where U_n, x_0 and u_0 are the same as in (4), and

$$v^{\langle l, a \rangle} := \begin{cases} \prod_{j=0}^{l-1} (v + ja) & \text{if} \quad l \geqslant 1 \\ 1 & \text{if} \quad l = 0. \end{cases}$$

For $a = 0$, this operator becomes the Bernstein operator (example (A)). In the case $a > 0$, we have that $L_{n,a}^{\langle k \rangle}$ is the composition

$$L_{n,a}^{\langle k \rangle} = \beta_{a^{-1}}^{\langle k \rangle} \circ L_{n,0}^{\langle k \rangle},$$

where $\beta_t^{\langle k\rangle}$ denotes the k-variate beta operator in example (B), and a suitable probabilistic representation for it is constructed as follows: Let $\{X_n: n\geqslant 1\}$ be a sequence of independent and on the interval $[0, 1]$ uniformly distributed random variables, and let $\{U_t: t>0\}$ be a standard gamma process independent of the preceding sequence and defined on the same probability space. Set

$$S_n(x) := \frac{1}{n}\sum_{i=1}^{n} 1_{[0, x]}(X_i), \qquad x\in[0, 1],$$

$$V_{a^{-1}}(x) := \frac{U_{xa^{-1}}}{U_{a^{-1}}}, \qquad x\in[0, 1],$$

and

$$Z_{n, a}(x) := S_n(V_{a^{-1}}(x)), \qquad x\in[0, 1].$$

By examples (A) and (B), and Proposition 1, the process $\{Z_{n, a}(x): x\in[0, 1]\}$ fulfills conditions (A)–(C) and (D′). It is not hard to check that, for all $n\geqslant 1$ and $\mathbf{x}\in I^{\langle k\rangle}$, the k-dimensional random vector $n\mathbf{Z}_{n, a}^{\langle k\rangle}(\mathbf{x})$ (where $\mathbf{Z}_{n, a}^{\langle k\rangle}(\mathbf{x})$ is given by (2)) has the appropriate Dirichlet compound distribution [13] in order that

$$L_{n, a}^{\langle k\rangle} f(\mathbf{x}) = Ef(\mathbf{Z}_{n, a}^{\langle k\rangle}(\mathbf{x})).$$

We claim that

$$\sup_{\delta>0} C_{n, a}^{\langle k\rangle}(\delta) = 2, \qquad n, k\geqslant 1, \tag{20}$$

and

$$D_{n, a}^{\langle k\rangle}(\omega) = 1, \qquad n, k\geqslant 1. \tag{21}$$

Since

$$EZ_{n, a}(x) = x, \qquad n\geqslant 1, \quad x\in[0, 1] \tag{22}$$

(and $Z_{n, a}(1)=1$ a.s.), (21) follows in the same way as (18). To show (20), observe that, from (22) and Theorem 2(IIa), we obtain

$$\sup_{\delta>0} C_{n, a}^{\langle k\rangle}(\delta) \leqslant 2, \qquad n, k\geqslant 1.$$

On the other hand, for all $n \geqslant 1$, $\delta \in (0, 1)$, we have

$$E\left[\frac{Z_{n,a}(\delta)}{\delta}\right] \geqslant P(Z_{n,a}(\delta) > 0) + P(Z_{n,a}(\delta) = 1)$$

$$= 1 - \prod_{i=0}^{n-1} \frac{1 - \delta + ia}{1 + ia} + \prod_{i=0}^{n-1} \frac{\delta + ia}{1 + ia}.$$

Letting $\delta \uparrow 1$, we conclude from Theorem 2(I) and Theorem A(a) that

$$\sup_{\delta > 0} C_{n,a}^{\langle k \rangle}(\delta) \geqslant \sup_{\delta > 0} C_{n,a}^{\langle 1 \rangle}(\delta) \geqslant 2, \qquad n, k \geqslant 1,$$

completing the proof of the claim.

(D) (*Modified*) *Multivariate Bernstein–Durrmeyer Operators.* Denote by $B_n^{\langle k \rangle}$ and $\beta_t^{\langle k \rangle}$ the Bernstein operator (example (A)) and the multivariate beta operator (example (B)) on the standard k-simplex $I^{\langle k \rangle}$, respectively. The operator on $I^{\langle k \rangle}$ given by the composition

$$L_n^{\langle k \rangle} := B_n^{\langle k \rangle} \circ \beta_n^{\langle k \rangle}, \qquad n \geqslant 1,$$

is the modification introduced by Goodman and Sharma [12] of the multivariate Bernstein–Durrmeyer operator (see [9, 10] and the references therein). It turns out that $\{L_n^{\langle k \rangle} : n \geqslant 1\}$ is the sequence of k-variate operators associated (via (1)) to the stochastic process $\{Z_n(x) : x \in [0, 1], n \geqslant 1\}$ given by

$$Z_n(x) := V_n(S_n(x)),$$

where $V_n(x)$ and $S_n(x)$ are the same as in example (C). It is clear that the process fulfills conditions (A)–(C) and (D′). Moreover, for all $n \geqslant 1$ and $\delta \in [0, 1]$, we have

$$Z_n(1) = 1 \qquad \text{a.s.},$$

$$E(Z_n(\delta)) = \delta,$$

$$P(Z_n(\delta) > 0) = P(S_n(\delta) > 0) = 1 - (1 - \delta)^n,$$

and

$$P(Z_n(\delta) = 1) = P(S_n(\delta) = 1) = \delta^n.$$

Therefore, the same arguments as in example (C) show that

$$\sup_{\delta>0} C_n^{\langle k\rangle}(\delta)=2, \qquad \text{and} \qquad D_n^{\langle k\rangle}(\omega)=1,$$

for all $n, k \geqslant 1$ (and every c.m.c. ω).

(E) *Multivariate Gamma Operators.* Let $Z := \{Z_t(x): x \geqslant 0, t>0\}$ be the double-indexed stochastic process given by

$$Z_t(x) := \frac{xU_t}{t},$$

where $\{U_t: t \geqslant 0\}$ is the same standard gamma process as in example (B). It is immediate that the process Z satisfies the assumptions in Theorem 2(IIb). For $k \geqslant 1$, $t>0$ and $\mathbf{x}=(x_1, ..., x_k) \in I^{\langle k\rangle}=[0, \infty)^k$, we have from (2) and (1)

$$\mathbf{Z}_t^{\langle k\rangle}(\mathbf{x})=\frac{\mathbf{x}U_t}{t},$$

and

$$L_t^{\langle k\rangle} f(\mathbf{x})=\int_0^\infty f\left(\frac{x_1\theta}{t}, ..., \frac{x_k\theta}{t}\right) g_t(\theta)\, d\theta,$$

where $g_t(\theta)$ is defined in (19). It is shown in [5] that

$$C_t^{\langle 1\rangle}(\delta)=\int_0^\infty \left\lceil \frac{\theta}{t} \right\rceil g_t(\theta)\, d\theta, \qquad t>0, \quad \delta>0,$$

$$\sup_{t>0} C_t^{\langle 1\rangle}(\delta)=2, \qquad \delta>0,$$

$$D_t^{\langle 1\rangle}(\omega_\alpha)=\frac{\Gamma(t+\alpha)}{t^\alpha \Gamma(t)}, \qquad t>0, \quad \alpha \in (0, 1],$$

and

$$\sup_{t>0} D_t^{\langle 1\rangle}(\omega_\alpha)=1, \qquad \alpha \in (0, 1].$$

By Theorem 2(IIb), these results immediately extend to the multivariate operator $L_t^{\langle k\rangle}$.

We claim that (for any c.m.c. ω)

$$\sup_{t>0} D_t^{\langle 1\rangle}(\omega)=1,$$

and, therefore, by Theorem 2(IIb)

$$\sup_{t>0} D_t^{\langle k \rangle}(\omega) = 1, \qquad k \geqslant 1.$$

Using that $EZ_t(x) = x$ ($t > 0$, $x \geqslant 0$), we have by Theorem 1 and Jensen's inequality

$$D_t^{\langle 1 \rangle}(\omega) \leqslant 1, \qquad t > 0.$$

On the other hand, the strong law of large numbers for the standard gamma process implies that $Z_t(x) \to x$ a.s., as $t \to \infty$, ($x \geqslant 0$). Therefore, by Theorem 1, Fatou's lemma and the continuity of ω in $(0, \infty)$, we have

$$\sup_{t>0} D_t^{\langle 1 \rangle}(\omega) \geqslant \sup_{\delta>0} \liminf_{t \to \infty} E\left(\frac{\omega(Z_t(\delta))}{\omega(\delta)}\right) \geqslant \sup_{\delta>0} E\left(\liminf_{t \to \infty} \frac{\omega(Z_t(\delta))}{\omega(\delta)}\right) = 1.$$

(F) *Multivariate Szász–Mirakyan Operators.* Let $Z := \{Z_t(x) \colon x \geqslant 0, t > 0\}$ be the double-indexed stochastic process given by

$$Z_t(x) := \frac{N_{tx}}{t},$$

where $\{N_t \colon t \geqslant 0\}$ is a standard Poisson process. It is clear that Z satisfies conditions (A)–(C). Moreover, for each $t > 0$, the process $\{Z_t(x) \colon x \geqslant 0\}$ has stationary independent increments. Therefore, according to Remark 2 above, Z also satisfies condition (D′). The mentioned properties of $\{Z_t(x) \colon x \geqslant 0\}$ also imply that the components of the k-dimensional random vector $\mathbf{Z}_t^{\langle k \rangle}(\mathbf{x})$ given in (2) are independent random variables, so that the associated operator $L_t^{\langle k \rangle}$ is actually the tensor product of k copies of the univariate operator $L_t^{\langle 1 \rangle}$ which is analytically given by

$$L_t^{\langle 1 \rangle} f(x) = e^{-tx} \sum_{i=0}^{\infty} f(i/t) \frac{(tx)^i}{i!}, \qquad t > 0, \quad x \geqslant 0.$$

Combining Theorem 2(IIb) with the results in [5] for $L_t^{\langle 1 \rangle}$, we obtain

$$C_t^{\langle k \rangle}(\delta) = \int_0^{t\delta} e^{-x} \sum_{i=0}^{\infty} \frac{x^{\lfloor it\delta \rfloor}}{\lfloor it\delta \rfloor!} \, dx, \qquad t, \delta > 0, \quad k \geqslant 1$$

(where $\lfloor a \rfloor$ indicates the "floor" or integral part of a),

$$\sup_{t>0} C_t^{\langle k \rangle}(\delta) = \sup_{\delta>0} C_t^{\langle k \rangle}(\delta) = 2 - \frac{1}{e}, \qquad t, \delta > 0, \quad k \geqslant 1,$$

and

$$D_t^{\langle k\rangle}(\omega_\alpha)=1, \qquad t>0, \qquad \alpha\in(0,1], \quad k\geqslant 1.$$

On the other hand, since

$$EZ_t(x)=x, \qquad t>0, \quad x\geqslant 0,$$

and (by the law of large numbers)

$$\lim_{t\to\infty} Z_t(x)=x \qquad \text{a.s.}, \quad x\geqslant 0,$$

the same arguments as in example (E) yield

$$\sup_{t>0} D_t^{\langle k\rangle}(\omega)=1, \qquad k\geqslant 1,$$

for every c.m.c. ω.

Remark 6. Recalling that the multivariate Szász operator $L_t^{\langle k\rangle}$ is a tensor product, the inequality $D_t^{\langle k\rangle}(\omega_\alpha)\leqslant 1$ also follows from the results in [15, p. 314].

(G) *Multivariate Baskakov Operators.* Let $Z:=\{Z_t(x): x\geqslant 0, t>0\}$ be the double-indexed stochastic process given by

$$Z_t(x):=\frac{N_{xU_t}}{t},$$

where $\{N_t: t\geqslant 0\}$ is a standard Poisson process and $\{U_t: t\geqslant 0\}$ is a standard gamma process independent of the former and defined on the same probability space. Then, Z satisfies conditions (A)–(C) and, by Corollary 1, it also satisfies condition (D′). It is not hard to check that, for $k\geqslant 1$, $t>0$ and $\mathbf{x}=(x_1, ..., x_k)\in I^{\langle k\rangle}=[0,\infty)^k$, the k-dimensional random vector $t\mathbf{Z}_t^{\langle k\rangle}(\mathbf{x})$ defined in (2) has the negative multinomial distribution with parameters t, $\mathbf{x}$ (see [13]), and the associated operator $L_t^{\langle k\rangle}$ is therefore given by

$$L_t^{\langle k\rangle} f(\mathbf{x})=\sum_{\mathbf{v}\in V} f\left(\frac{\mathbf{v}}{t}\right)\frac{\Gamma(t+\sum_{i=1}^k v_i)}{\Gamma(t)\prod_{i=1}^k v_i!}\,\frac{\prod_{i=1}^k x_i^{v_i}}{(1+\sum_{i=1}^k x_i)^{t+\sum_{i=1}^k v_i}},$$

where V is the set of all k-tuples of nonnegative integers $\mathbf{v}:=(v_1, ..., v_k)$. The univariate operator $L_t^{\langle 1\rangle}$ is the celebrated Baskakov operator. The multivariate operator $L_t^{\langle k\rangle}$ has been considered in [4] in connection with the problem of monotonic convergence under convexity. It is interesting to observe that $L_t^{\langle k\rangle}$ can be written as the composition of the operators in examples (E) and (F) above. As in these examples, Theorem 2(IIb), together with the results in [5] for $L_t^{\langle 1\rangle}$, yields

$$C_t^{\langle k\rangle}(\delta)=\int_0^\delta \sum_{i=0}^{\infty} \frac{1}{B(t,\lfloor it\delta\rfloor+1)}\,\frac{x^{\lfloor it\delta\rfloor}}{(1+x)^{\lfloor it\delta\rfloor+t+1}}\,dx, \qquad t,\delta>0, \quad k\geqslant 1$$

(where $B(\cdot,\cdot)$ is the standard beta function),

$$\sup_{t,\delta>0} C_t^{\langle k\rangle}(\delta)=2, \qquad k\geqslant 1,$$

$$D_t^{\langle k\rangle}(\omega_\alpha)=\frac{\Gamma(t+\alpha)}{t^\alpha\Gamma(t)}, \qquad t>0, \quad \alpha\in(0,1], \quad k\geqslant 1,$$

and

$$\sup_{t>0} D_t^{\langle k\rangle}(\omega_\alpha)=1, \qquad \alpha\in(0,1], \qquad k\geqslant 1.$$

Finally, using the same arguments as in examples (E) and (F), it can be readily shown that

$$\sup_{t>0} D_t^{\langle k\rangle}(\omega)=1, \qquad k\geqslant 1,$$

for every c.m.c. ω.

Remark 7. Lemma 2 is also of interest in probability theory since it gives the exact Wasserstein distance between two k-dimensional probability distributions associated with a superstationary process. In particular, from the examples (A), (B), (F) and (G) in this section, we have that

$$d(\mathbf{X},\mathbf{Y})=t\,\|\mathbf{x}-\mathbf{y}\|,$$

if $\mathbf{X}$ and $\mathbf{Y}$ have (negative) multinomial distributions with parameters t, $\mathbf{x}$ and t, $\mathbf{y}$, respectively, and

$$d(\mathbf{X},\mathbf{Y})=\|\mathbf{x}-\mathbf{y}\|,$$

if $\mathbf{X}$ and $\mathbf{Y}$ have Dirichlet (or Poisson) distributions with parameters $\mathbf{x}$ and $\mathbf{y}$, respectively.

ACKNOWLEDGMENTS

We are grateful to the referees for their comments and suggestions. Special thanks are given to the referee who pointed out Ref. [16] and suggested generalizing the problem of the D-constants from Lipschitz classes to classes of functions defined by concave moduli of continuity.

REFERENCES

1. J. A. Adell and J. de la Cal, Using stochastic processes for studying Bernstein-type operators, *Rend. Circ. Mat. Palermo (2) Suppl.* **33** (1993), 125–141.
2. J. A. Adell and J. de la Cal, Preservation of moduli of continuity for Bernstein-type operators, *in* "Approximation, Probability and Related Fields" (G. A. Anastassiou and S. T. Rachev, Eds.), pp. 1–18, Plenum, New York, 1994.
3. J. A. Adell and J. de la Cal, Optimal Poisson approximation of uniform empirical processes, *Stochastic Process. Appl.* **64** (1996), 135–142.
4. J. A. Adell, J. de la Cal, and M. San Miguel, On the property of monotonic convergence for multivariate Bernstein-type operators, *J. Approx. Theory* **80** (1995), 132–137.
5. J. A. Adell and A. Pérez-Palomares, Best constants in preservation inequalities concerning the first modulus and Lipschitz classes for Bernstein-type operators, *J. Approx. Theory* **93** (1998), 128–139.
6. J. A. Adell and A. Pérez-Palomares, Global smoothness preservation properties for generalized Szász-Kantorovich operators, preprint, 1996.
7. G. A. Anastassiou, C. Cottin, and H. H. Gonska, Global smoothness of approximating functions, *Analysis* **11** (1991), 43–57.
8. G. A. Anastassiou, C. Cottin, and H. H. Gonska, Global smoothness preservation by multivariate approximation operators, *in* "Israel Mathematical Conference Proceedings, Vol. IV" (S. Baron and D. Leviatan, Eds.), pp. 31–44, Bar-Ilan University, Ramat-Gan, 1991.
9. M. M. Derriennic, Sur l'approximation de fonctions intégrables sur $[0,1]$ par des polynômes de Bernstein modifiés, *J. Approx. Theory* **31** (1981), 325–343.
10. M. M. Derriennic, On multivariate approximation by Bernstein-type polynomials, *J. Approx. Theory* **45** (1985), 155–166.
11. C. R. Givens and R. M. Shortt, A class of Wasserstein metrics for probability distributions, *Michigan Math. J.* **31** (1984), 231–240.
12. T. N. T. Goodman and A. Sharma, A Bernstein type operator on the simplex, *Math. Balkanica* (*N.S.*) **5** (1991), 129–145.
13. N. L. Johnson and S. Kotz, "Discrete Distributions," Houghton Mifflin, New York, 1969.
14. N. L. Johnson and S. Kotz, "Continuous Multivariate Distributions," Wiley, New York, 1972.
15. M. K. Khan and M. A. Peters, Lipschitz constants for some approximation operators of a Lipschitz continuous function, *J. Approx. Theory* **59** (1989), 307–315.
16. N. Korneichuk, "Exact Constants in Approximation Theory," Cambridge Univ. Press, New York, 1991.
17. G. Mühlbach, Verallgemeinerungen der Bernstein- und der Lagrangepolynome, *Rev. Roumaine Math. Pures Appl.* **15** (1970), 1235–1252.
18. S. M. Nikol'skii, Approximation of periodic functions by trigonometric polynomials, *Trudy Mat. Inst. Stekhlov* **15** (1945), 1–76. [In Russian]
19. S. M. Nikol'skii, Fourier series of functions with a given modulus of continuity, *Dokl. Akad. Nauk SSSR* **52** (1946), 191–194. [In Russian]
20. S. C. Port, "Theoretical Probability for Applications," Wiley, New York, 1994.
21. D. D. Stancu, Approximation of functions by a new class of linear polynomial operators, *Rev. Roumaine Math. Pures Appl.* **13** (1968), 1173–1198.
22. D. D. Stancu, Probabilistic methods in the theory of approximation of functions of several variables by linear positive operators, *in* "Approximation Theory" (A. Talbot, Ed.), pp. 329–342, Academic Press, New York, 1970.