Canad. Math. Bull. Vol. **51** (2), 2008 pp. 236–248

Kolmogorov, Linear and Pseudo-Dimensional Widths of Classes of s-Monotone Functions in $\mathbb{L}_p$, $0 < p < 1$

Victor N. Konovalov and Kirill A. Kopotun

Abstract. Let B_p be the unit ball in $\mathbb{L}_p$, $0 < p < 1$, and let Δ_+^s, $s \in \mathbb{N}$, be the set of all s-monotone functions on a finite interval I, *i.e.*, Δ_+^s consists of all functions $x\colon I \mapsto \mathbb{R}$ such that the divided differences $[x; t_0, \dots, t_s]$ of order s are nonnegative for all choices of $(s+1)$ distinct points $t_0, \dots, t_s \in I$. For the classes $\Delta_+^s B_p := \Delta_+^s \cap B_p$, we obtain exact orders of Kolmogorov, linear and pseudo-dimensional widths in the spaces $\mathbb{L}_q$, $0 < q < p < 1$:

$$d_n(\Delta_+^s B_p)^{\mathrm{psd}}_{\mathbb{L}_q} \asymp d_n(\Delta_+^s B_p)^{\mathrm{kol}}_{\mathbb{L}_q} \asymp d_n(\Delta_+^s B_p)^{\mathrm{lin}}_{\mathbb{L}_q} \asymp n^{-s}.$$

1 Introduction, Preliminaries, and the Main Result

The general theory of widths deals with approximation of infinite-dimensional function classes by finite-dimensional manifolds which are optimal in a certain sense. It has numerous applications in numerical analysis (what is the optimal rate of convergence attainable by any numerical method for a specific problem?), learning theory (among various hypothesis classes what can be considered optimal?), image compression (what is the best compression ratio?), and many other areas.

Let X be a real linear space of vectors x with (quasi)norm $\|x\|_X$, and W and M be nonempty subsets of X. The *deviation* of W from M is defined by

$$E(W, M)_X := \sup_{x \in W} \inf_{y \in M} \|x - y\|_X .$$

The *Kolmogorov n-width* of W is defined by

$$d_n(W)^{\mathrm{kol}}_X := \inf_{M^n} E(W, M^n)_X, \quad n \geq 0,$$

where the infimum is taken over all affine subsets $M^n \subseteq X$ of dimension $\leq n$.

The *linear n-width* of W is defined by

$$d_n(W)^{\mathrm{lin}}_X := \inf_{M^n} \inf_{A} \sup_{x \in W} \|x - Ax\|_X, \quad n \geq 0,$$

Received by the editors December 12, 2005; revised February 10, 2006.
Supported by NSERC of Canada.
AMS subject classification: Primary: 41A46; secondary: 46E35, 41A29.
Keywords: Kolmogorov width, linear width, pseudo-dimensional widths, s-monotone functions.

where the left-hand infimum is taken over all affine subsets $M^n \subseteq X$ of dimension at most n, and the middle infimum is taken over all continuous affine maps A from affine subsets of X containing W into M^n.

Finally, we will also have estimates for yet another width, the *pseudo-dimensional width* which was introduced by Maiorov and Ratsaby [11,12,15] using the concept of pseudo-dimension due to Pollard [13]. Namely, let $M = M(T)$ be a set of real-valued functions $x(\cdot)$ defined on the set T, and denote

$$\operatorname{Sgn}(a) := \begin{cases} 1 & \text{if } a > 0, \\ 0 & \text{if } a \le 0. \end{cases}$$

The *pseudo-dimension* $\dim_{ps}(M)$ of the set M is the largest integer n such that there exist points $t_1, \dots, t_n \in T$ and a vector $(y_1, \dots, y_n) \in \mathbb{R}^n$ for which

$$\operatorname{card}\left\{ \left(\operatorname{Sgn}(x(t_1) + y_1), \dots, \operatorname{Sgn}(x(t_n) + y_n)\right) \;\middle|\; x \in M \right\} = 2^n.$$

If n can be arbitrarily large, then $\dim_{ps}(M) := \infty$.

The *pseudo-dimensional n-width* of W is defined by

$$d_n(W)_X^{\mathrm{psd}} := \inf_{M^n} \sup_{x \in W} \inf_{y \in M^n} \|x - y\|_X,$$

where the left-hand infimum is taken over all subsets $M^n \subseteq X$ such that $\dim_{ps}(M^n) \le n$. It is known (see [6]) that if M is an arbitrary affine subset in a space X of real-valued functions and $\dim M < \infty$, then $\dim_{ps}(M) = \dim(M)$. Clearly,

$$d_n(W)_X^{\mathrm{psd}} \le d_n(W)_X^{\mathrm{kol}} \le d_n(W)_X^{\mathrm{lin}}. \tag{1.1}$$

Let $s \in \mathbb{N}$, and $\Delta_+^s := \Delta_+^s(I)$ be the set of functions $x \colon I \mapsto \mathbb{R}$ on a finite interval I such that the divided differences $[x; t_0, \dots, t_s]$ of order s of x are nonnegative for all choices of $(s+1)$ distinct points $t_0, \dots, t_s \in I$. We call functions $x \in \Delta_+^s$ *s-monotone* on I.

It is well known (see [1, 14, 16]) that if x is s-monotone on $[a, b]$, $s \ge 2$, then $x^{(\nu)}$ exists on (a, b) for $\nu \le s - 2$, and, in fact, $x^{(\nu)} \in \Delta_+^{s-\nu}(a, b)$. In particular, $x^{(s-2)}$ exists, is convex, and therefore is locally absolutely continuous in (a, b), and has left and right (nondecreasing) derivatives $x_-^{(s-1)}$ and $x_+^{(s-1)}$ there. Moreover, the set E where $x^{(s-1)}$ fails to exist is countable, and $x^{(s-1)}$ is continuous on $(a, b) \setminus E$. Throughout this paper, if a function $x \colon I \mapsto \mathbb{R}$ possesses both the left-hand and the right-hand derivatives $x_-^{(k)}(t)$ and $x_+^{(k)}(t)$, of order $k \in \mathbb{N}$, at a point $t \in I$, then we denote $x^{(k)}(t) := \left(x_-^{(k)}(t) + x_+^{(k)}(t)\right)/2$. Evidently, this notation is compatible with that of the derivative $x^{(k)}(t)$ if it exists. We also write $x^{(0)}(t) := x(t)$, $t \in I$.

Given a function space X and $W \subseteq X$, we denote by $\Delta_+^s W$ the subset of s-monotone functions $x \in W$, *i.e.*, $\Delta_+^s W := \Delta_+^s \cap W$.

Let $\mathbb{L}_p := \mathbb{L}_p(I)$, $0 < p \le \infty$, be the space of all functions x on I with (quasi)norm $\|x\|_{\mathbb{L}_p(I)} := \left(\int_I |x(t)|^p \, dt\right)^{1/p}$. By $B_p := B_p(I)$ we denote the unit ball in $\mathbb{L}_p$. Evidently

$\Delta^s_+ B_p \subset \mathbb{L}_q$, $0 < q \le p \le \infty$, but $\Delta^s_+ B_p \not\subset \mathbb{L}_q$, $0 < p < q \le \infty$. For the integers $r \in \mathbb{N}$, we denote the Sobolev classes

$$\mathbb{W}^r_p := \mathbb{W}^r_p(I) := \{x \mid x^{(r-1)} \in AC_{\mathrm{loc}}, \|x^{(r)}\|_{\mathbb{L}_p} \le 1\}, \quad 0 < p \le \infty.$$

Recall that for $1 \le p, q \le \infty$, the orders of most widths of the classical Sobolev classes $\mathbb{W}^r_p$ in $\mathbb{L}_q$ are well known. They are asymptotically $n^{-r+\alpha(p,q)}$, where $0 \le \alpha(p,q) \le 1/2$. In contrast, for $0 < p < 1$ the behavior of the widths differs essentially. Let

$$\mathbb{W}^r_{p,\infty} := \{x \mid x \in \mathbb{W}^r_p, \|x\|_{\mathbb{L}_\infty} \le 1\}, \quad 0 < p < 1.$$

It was proved in [3] that if $0 < p < 1$, then

$$d_n(\mathbb{W}^r_{p,\infty})^{\mathrm{psd}}_{\mathbb{L}_q} \asymp d_n(\mathbb{W}^r_{p,\infty})^{\mathrm{kol}}_{\mathbb{L}_q} \asymp d_n(\mathbb{W}^r_{p,\infty})^{\mathrm{lin}}_{\mathbb{L}_q} \asymp 1, \quad 0 < q \le \infty.$$

The aim of this paper is to show that s-monotone functions $x \in \Delta^s_+ B_p$, $0 < p < 1$, can be approximated well in $\mathbb{L}_q$, $0 < q < p < 1$. Our main result is:

Theorem 1 *If $s, n \in \mathbb{N}$ and $0 < q < p < 1$ then*

$$d_n(\Delta^s_+ B_p)^{\mathrm{psd}}_{\mathbb{L}_q} \asymp d_n(\Delta^s_+ B_p)^{\mathrm{kol}}_{\mathbb{L}_q} \asymp d_n(\Delta^s_+ B_p)^{\mathrm{lin}}_{\mathbb{L}_q} \asymp n^{-s}, \quad n \ge s. \tag{1.2}$$

2 Upper Bounds

This section is devoted to proving that a function $x \in \Delta^s_+ B_p$, $0 < p < 1$, can be well approximated in $\mathbb{L}_q$, $0 < q < p$, by piecewise polynomials associated with it, in a linear fashion. Specifically, we will show that for $x \in \Delta^s_+ \mathbb{L}_p$ there is a piecewise polynomial $\sigma_{s,n}(\cdot\,; x; I)$ with $2n - 2$ prescribed knots (see construction below), such that

$$\|x(\cdot) - \sigma_{s,n}(\cdot\,; x; I)\|_{\mathbb{L}_q(I)} \le c\|x\|_{\mathbb{L}_p(I)} n^{-s}, \quad n \ge 1,\ 0 < q < p < 1, \tag{2.1}$$

where $c = c(s, p, q)$.

2.1 Auxiliary Results

The following lemma is due to Bullen [1] (see also [9, Lemma 8.3] for discussion of the cases when some or all interpolation points coincide).

Lemma 1 *Let $s \in \mathbb{N}$, $f \in \Delta^s_+(a,b)$, and let $L_{s-1}(f, t; t_1, \dots, t_k)$ be the Lagrange (Hermite–Taylor) polynomial of degree $\le s-1$ interpolating f at the points t_i, $1 \le i \le s$, where $a =: t_0 < t_1 \le \cdots \le t_s < t_{s+1} := b$. Then*

$$(-1)^{s-i}\big(f(t) - L_{s-1}(f, t; t_1, \dots, t_s)\big) \ge 0, \quad t \in (t_i, t_{i+1}),\ i = 0, \dots, s.$$

In other words, $f - L_{s-1}$ changes sign at $t_1, \dots, t_s$.

Lemma 2 *Let* $n \in \mathbb{N}$, $0 < p < 1$, *and let* $(\alpha_{ij})_{i,j=1}^n$ *be an* $n \times n$ *matrix with nonnegative entries. Then*

$$\sum_{j=1}^n \Big(\sum_{i=1}^n \alpha_{ij}^p \Big)^{1/p} \le \Big(\sum_{i=1}^n \Big(\sum_{j=1}^n \alpha_{ij} \Big)^p \Big)^{1/p}.$$

Proof It is well known (see, *e.g.*, [5, (2.11.5)]) and easy to prove that, in the case $0 < p < 1$,

$$\|\mathbf{x}+\mathbf{y}\|_{l_p^n} \ge \|\mathbf{x}\|_{l_p^n} + \|\mathbf{y}\|_{l_p^n}, \quad \text{for any } \mathbf{x}, \mathbf{y} \in \mathbb{R}_+^n.$$

Therefore, for any $\mathbf{x}_j \in \mathbb{R}_+^n$, $1 \le j \le n$, we have

$$\Big\| \sum_{j=1}^n \mathbf{x}_j \Big\|_{l_p^n} \ge \sum_{j=1}^n \|\mathbf{x}_j\|_{l_p^n},$$

and hence choosing $\mathbf{x}_j = (\alpha_{1j}, \alpha_{2j}, \dots, \alpha_{nj})$, $1 \le j \le n$, we have

$$\begin{aligned}\Big(\sum_{i=1}^n \Big(\sum_{j=1}^n \alpha_{ij} \Big)^p \Big)^{1/p} &= \Big\| \Big(\sum_{j=1}^n \alpha_{1j}, \sum_{j=1}^n \alpha_{2j}, \dots, \sum_{j=1}^n \alpha_{nj} \Big) \Big\|_{l_p^n} = \Big\| \sum_{j=1}^n \mathbf{x}_j \Big\|_{l_p^n} \\ &\ge \sum_{j=1}^n \|\mathbf{x}_j\|_{l_p^n} = \sum_{j=1}^n \Big(\sum_{i=1}^n \alpha_{ij}^p \Big)^{1/p}.\end{aligned}$$

Lemma 3 *Let* $n \in \mathbb{N}$, $0 < p, q < 1$, *and suppose that* $\mathbf{a} = (a_1, \dots, a_n)$ *and* $\mathbf{b} = (b_1, \dots, b_n)$ *are n-tuples with nonnegative entries, and* $\mathbf{c} := (c_{ij})_{i,j=1}^n$ *is a nonnegative* $n \times n$ *matrix. Given a function*

$$f_{q,n}(\mathbf{w}; \mathbf{a}) := \Big(\sum_{i=1}^n (a_i \omega_i)^q \Big)^{1/q},$$

where $\mathbf{w} := (\omega_1, \dots, \omega_n) \in \mathbb{R}_+^n$, *and the set*

$$\Omega_p^n(\mathbf{b}; \mathbf{c}) := \mathbb{R}_+^n \cap \Big\{ \mathbf{w} : \sum_{i=1}^n \Big(b_i \sum_{j=1}^n c_{ij} \omega_j \Big)^p \le 1 \Big\},$$

we have

$$\max_{\mathbf{w} \in \Omega_p^n(\mathbf{b};\mathbf{c})} f_{q,n}(\mathbf{w}; \mathbf{a}) \le n^{1/q-1} \max_{1 \le j \le n} \Big\{ a_j \Big(\sum_{i=1}^n (b_i c_{ij})^p \Big)^{-1/p} \Big\}.$$

Proof By Jensen's inequality, for every $\mathbf{w} \in \mathbb{R}_+^n$, we have

$$f_{q,n}(\mathbf{w}\,;\mathbf{a}) = \Big(\sum_{i=1}^n (a_i\omega_i)^q\Big)^{1/q} \le n^{1/q-1}\sum_{i=1}^n a_i\omega_i =: g_{q,n}(\mathbf{w}\,;\mathbf{a}).$$

It follows from Lemma 2 that

$$\sum_{i=1}^n\Big(\sum_{j=1}^n \alpha_{ij}\Big)^p \le 1 \quad \Rightarrow \quad \sum_{j=1}^n\Big(\sum_{i=1}^n \alpha_{ij}^p\Big)^{1/p} \le 1.$$

Therefore, taking $\alpha_{ij} := b_i c_{ij}\omega_j$, we conclude that the set $\Omega_p^n(\mathbf{b}\,;\mathbf{c})$ is a subset of the simplex

$$S_p^n(\mathbf{b}\,;\mathbf{c}) := \mathbb{R}_+^n \cap \Big\{\mathbf{w} : \sum_{j=1}^n\Big(\sum_{i=1}^n (b_i c_{ij})^p\Big)^{1/p}\omega_j \le 1\Big\}.$$

Since $g_{q,n}(\cdot\,;\mathbf{a})$ is a linear function it achieves its maximum on $S_p^n(\mathbf{b}\,;\mathbf{c})$ at the vertices

$$\mathbf{z}_j = \Big(\sum_{i=1}^n (b_i c_{ij})^p\Big)^{-1/p}\mathbf{e}^j, \quad 1 \le j \le n,$$

of $S_p^n(\mathbf{b}\,;\mathbf{c})$ (here, $\mathbf{e}^j$, $1 \le j \le n$, is the standard basis of $\mathbb{R}^n$). Therefore,

$$\begin{aligned}\max_{\mathbf{w}\in\Omega_p^n(\mathbf{b};\mathbf{c})} f_{q,n}(\mathbf{w}\,;\mathbf{a}) &\le \max_{\mathbf{w}\in\Omega_p^n(\mathbf{b};\mathbf{c})} g_{q,n}(\mathbf{w}\,;\mathbf{a}) \le \max_{\mathbf{w}\in S_p^n(\mathbf{b};\mathbf{c})} g_{q,n}(\mathbf{w}\,;a)\\ &= n^{1/q-1}\max_{1\le j\le n}\Big\{a_j\Big(\sum_{i=1}^n (b_i c_{ij})^p\Big)^{-1/p}\Big\},\end{aligned}$$

and the proof is complete. ■

The following lemma is an immediate corollary of a stronger Theorem 1 in [8], taking into account [4] (see also [2, Theorem 4.6.3]).

Lemma 4 **([8])** *Let $s \in \mathbb{N}$, $0 < p \le \infty$, $I := (-1,1)$, and $x \in \Delta_+^s\mathbb{L}_p(I)$. Denote by*

$$T_{s-1}(t) := T_{s-1}(t\,;x\,;0) := \sum_{k=0}^{s-1}\frac{x^{(k)}(0)}{k!}t^k, \quad t \in I,$$

the Taylor polynomial of degree $\le s-1$ at $t = 0$. Then, there exists a constant $c = c(s,p)$ such that

$$\|x - T_{s-1}\|_{\mathbb{L}_p(I)} \le c\,\|x\|_{\mathbb{L}_p(I)}.$$

2.2 Proof of the Upper Bounds in Theorem 1

Fix $n \geq 1$, and let $\beta \in \mathbb{N}$, to be prescribed. For every $-n \leq i \leq n$, denote

$$t_i := \operatorname{sign}(i)\big(1 - \big((n-|i|)/n\big)^{\beta}\big).$$

Also,

$$I_i := \begin{cases} [t_{i-1}, t_i) & i = 1, \dots, n, \\ (t_i, t_{i+1}] & i = -1, \dots, -n. \end{cases}$$

Note that $t_{-i} = -t_i$ and $|I_{-i}| = |I_i|$ for all $-n \leq i \leq n$. Straightforward computations show that $\beta(n-i)^{\beta-1}n^{-\beta} \leq |I_i| \leq (2^\beta - 1)(n-i)^{\beta-1}n^{-\beta}$ for $1 \leq i \leq n-1$, $|I_n| = n^{-\beta}$, and $|I_{i+1}| \leq |I_i| \leq (2^\beta - 1)|I_{i+1}|$ for all $1 \leq i \leq n-1$.

Let $x \in \Delta^s_+(I)$. Recall that $x^{(s-1)}$ is nondecreasing, and so has left and right derivatives $x_-^{(s-1)}(\tau)$ and $x_+^{(s-1)}(\tau)$, $\tau \in I$. We recall our notation $x^{(s-1)}(\tau) := \big(x_+^{(s-1)}(\tau) + x_-^{(s-1)}(\tau)\big)/2$, and put

$$T_{s-1}(t\,;x\,;\tau) := \sum_{k=0}^{s-1} \frac{x^{(k)}(\tau)}{k!}(t-\tau)^k, \quad t \in I.$$

Finally, we denote

$$T_{s-1}(t\,;x\,;I_i) := \begin{cases} T_{s-1}(t\,;x\,;t_{i-1}) & i = 1, \dots, n, \\ T_{s-1}(t\,;x\,;t_{i+1}) & i = -1, \dots, -n, \end{cases}$$

and set

$$\sigma_{s,n}(t) := \sigma_{s,n}(t\,;x\,;I) := T_{s-1}(t\,;x\,;I_i), \quad t \in I_i, \quad i = \pm 1, \dots, \pm n. \tag{2.2}$$

We will estimate the distance of $\sigma_{s,n}(\cdot\,;x\,;I)$ from $x \in \Delta^s_+ B_p$. First we assume that x satisfies

$$x^{(k)}(0) = 0, \quad k = 0, \dots, s-1. \tag{2.3}$$

It follows from Lemma 1 (or may be proved directly) that in this case $x^{(k)}(t) \geq 0$, $t \in I_+ := [0,1)$, and $(-1)^{s-k}x^{(k)}(t) \geq 0$, $t \in I_- := (-1,0]$, $k = 0, \dots, s-1$. We restrict our discussion to I_+, the estimates for I_- follow by the observation that $y(t) := (-1)^s x(-t)$, $t \in I$, satisfies $y^{(k)}(t) = (-1)^{s-k}x^{(k)}(-t) \geq 0$, $t \in I_+$, $k = 0, \dots, s-1$, and that $\sigma_{s,n}(t\,;y\,;I) = (-1)^s\sigma_{s,n}(-t\,;x\,;I)$. Without loss of generality we assume that $\|x\|_{\mathbb{L}_p(I_+)} \neq 0$.

If $n = 1$, then $\sigma_{s,1}(t) \equiv 0$, $t \in I$, and by Hölder's inequality we have

$$\|x(\cdot) - \sigma_{s,1}\|_{\mathbb{L}_q(I)} = \|x\|_{\mathbb{L}_q(I)} \leq 2^{1/q-1/p}\,\|x\|_{\mathbb{L}_p(I)}.$$

From now on, we assume that $n > 1$, and denote

$$\omega_i := \omega_i(x^{(s-1)}\,;I_i) := x^{(s-1)}(t_i) - x^{(s-1)}(t_{i-1}), \quad i = 1, \dots, n-1.$$

First, if $s = 1$ and $t \in I_i$, $1 \le i \le n-1$, then it is readily seen that

$$\|x - \sigma_{1,n}\|_{\mathbb{L}_q(I_i)} \le |I_i|^{1/q}\omega_i. \tag{2.4}$$

If $s > 1$ and $t \in I_i$, $1 \le i \le n-1$, then Taylor's formula with integral remainder and integration by parts yield

$$\begin{aligned} x(t) - \sigma_{s,n}(t) &= x(t) - T_{s-1}(t\,;x\,;I_i) \\ &= \frac{1}{(s-2)!}\int_{t_{i-1}}^{t} \big(x^{(s-1)}(\tau) - x^{(s-1)}(t_{i-1})\big)\,(t-\tau)^{s-2}\,d\tau. \end{aligned}$$

Hence, using monotonicity of $x^{(s-1)}$ we conclude that

$$\|x - \sigma_{s,n}\|_{\mathbb{L}_q(I_i)} \le \frac{|I_i|^{s-1+1/q}}{(s-1)!}\omega_i, \quad i = 1, \dots, n-1. \tag{2.5}$$

Taking into account (2.4) we see that (2.5) is valid for all $s \ge 1$. It now remains to consider that case $t \in I_n$. For any $s \ge 1$, if $t \in I_n$ then by Lemma 1 and the fact that $x^{(k)}(t_{n-1}) \ge 0$, $k = 0, \dots, s-1$, we have $0 \le x(t) - \sigma_{s,n}(t) = x(t) - T_{s-1}(t\,;x\,;I_n) \le x(t)$. By Hölder's inequality we get

$$\|x - \sigma_{s,n}\|_{\mathbb{L}_q(I_n)} \le \|x\|_{\mathbb{L}_p(I_n)}\,|I_n|^{1/q-1/p}. \tag{2.6}$$

Combining (2.5) and (2.6) we obtain

$$\begin{aligned} \|x - \sigma_{s,n}\|^q_{\mathbb{L}_q(I_+)} &= \sum_{i=1}^{n} \|x - \sigma_{s,n}\|^q_{\mathbb{L}_q(I_i)} \\ &\le \frac{1}{\big((s-1)!\big)^q}\sum_{i=1}^{n-1}(|I_i|^{s-1+1/q}\omega_i)^q + \|x\|^q_{\mathbb{L}_p(I_n)}|I_n|^{1-q/p}. \end{aligned}$$

Suppose now that $t \in I_i$, $2 \le i \le n$, is fixed. Since by (2.3), $T_{s-1}(\cdot\,;x\,;I_1) \equiv 0$, we have

$$x(t) = x(t) - T_{s-1}(t\,;x\,;I_i) + \sum_{j=2}^{i}\big(T_{s-1}(t\,;x\,;I_j) - T_{s-1}(t\,;x\,;I_{j-1})\big)$$

and note that

$$\begin{aligned} T_{s-1}(t\,;x\,;I_j) - T_{s-1}(t\,;x\,;I_{j-1}) &= \sum_{k=0}^{s-1}\big(x^{(k)}(t_{j-1}) - T^{(k)}_{s-1}(t_{j-1}\,;x\,;I_{j-1})\big)\,\frac{(t-t_{j-1})^k}{k!} \\ &= \sum_{k=0}^{s-1}\big(x^{(k)}(t_{j-1}) - T_{s-k-1}(t_{j-1}\,;x^{(k)}\,;I_{j-1})\big)\,\frac{(t-t_{j-1})^k}{k!}. \end{aligned}$$

Lemma 1 implies (this is also not difficult to show directly) that $x(t) \geq T_{s-1}(t\,;x\,;I_i)$ and $x^{(k)}(t_{j-1}) \geq T_{s-k-1}(t_{j-1}\,;x^{(k)}\,;I_{j-1})$, $0 \leq k \leq s-2$. Therefore,

$$\begin{aligned} &T_{s-1}(t\,;x\,;I_j) - T_{s-1}(t\,;x\,;I_{j-1}) \\ &\qquad \geq \left(x^{(s-1)}(t_{j-1}) - T_0(t_{j-1}\,;x^{(s-1)}\,;I_{j-1})\right)\frac{(t-t_{j-1})^{s-1}}{(s-1)!} \\ &\qquad = \frac{(t-t_{j-1})^{s-1}}{(s-1)!}\omega_{j-1}, \end{aligned}$$

and, hence, for any $t \in I_i$, $2 \leq i \leq n$,

$$x(t) \geq \sum_{j=2}^{i} \frac{(t-t_{j-1})^{s-1}}{(s-1)!}\omega_{j-1}. \tag{2.7}$$

Denoting $\bar{t}_i := (t_i + t_{i-1})/2$, for $2 \leq j \leq i \leq n$ and $t \in [\bar{t}_i, t_i)$, we have

$$\begin{aligned} t - t_{j-1} &= (t-\bar{t}_i) + (\bar{t}_i - t_{i-1}) + \sum_{k=j}^{i-1}(t_k - t_{k-1}) \geq \frac{1}{2}\sum_{k=j}^{i}(t_k - t_{k-1}) \\ &= \frac{1}{2}\sum_{k=j}^{i}|I_k| \geq \frac{1}{2}(i-j+1)|I_i|, \end{aligned}$$

since $|I_1| \geq |I_2| \geq \cdots \geq |I_n|$.

Combining this estimate with (2.7) we obtain

$$x(t) \geq \frac{|I_i|^{s-1}}{2^{s-1}(s-1)!}\sum_{j=2}^{i}(i-j+1)^{s-1}\omega_{j-1}, \quad t \in [\bar{t}_i, t_i), \quad 2 \leq i \leq n,$$

which implies

$$\begin{aligned} \|x\|^p_{\mathbb{L}_p(I_+)} &\geq \sum_{i=2}^{n}\|x\|^p_{\mathbb{L}_p[\bar{t}_i,t_i]} \\ &\geq \frac{1}{2^{(s-1)p+1}\left((s-1)!\right)^p}\sum_{i=2}^{n}|I_i|^{(s-1)p+1}\left(\sum_{j=2}^{i}(i-j+1)^{s-1}\omega_{j-1}\right)^p \\ &= \sum_{i=1}^{n-1}\left(\frac{|I_{i+1}|^{s-1+1/p}}{2^{s-1+1/p}(s-1)!}\sum_{j=1}^{i}(i-j+1)^{s-1}\omega_j\right)^p \\ &\geq \sum_{i=1}^{n-1}\left(\bar{c}|I_i|^{s-1+1/p}\sum_{j=1}^{i}(i-j+1)^{s-1}\omega_j\right)^p, \end{aligned} \tag{2.8}$$

where $\bar{c} = 2^{-(\beta+1)(s-1+1/p)}\big((s-1)!\big)^{-1}$.

We can rewrite (2.8) in the following equivalent form

$$\sum_{i=1}^{n-1}\Big(\frac{\bar{c}|I_i|^{s-1+1/p}}{\|x\|_{\mathbb{L}_p(I_+)}}\sum_{j=1}^{i}(i-j+1)^{s-1}\omega_j\Big)^p \le 1. \tag{2.9}$$

We are interested in estimating the sum in (2.2) from above subject to (2.9). In other words, we want to find (estimate) the maximum value of the function

$$\big(f_{q,n-1}(\mathbf{w};\mathbf{a})\big)^q := \sum_{i=1}^{n-1}(a_i\omega_i)^q,$$

on the set

$$\Omega_p^{n-1}(\mathbf{b};\mathbf{c}) := \mathbb{R}_+^{n-1}\cap\Big\{\mathbf{w}:\sum_{i=1}^{n-1}\Big(b_i\sum_{j=1}^{n-1}c_{ij}\omega_j\Big)^p\le 1\Big\},$$

where $a_i := |I_i|^{s-1+1/q}$, $b_i := \bar{c}|I_i|^{s-1+1/p}\,\|x\|_{\mathbb{L}_p(I_+)}^{-1}$ and $c_{ij} := (i-j+1)_+^{s-1} := (\max\{i-j+1,0\})^{s-1}$, $i,j=1,\ldots,n-1$. We now estimate $\sum_{i=1}^{n-1}(b_ic_{ij})^p$, and then apply Lemma 3 in order to estimate this maximum value.

$$\begin{aligned}
\sum_{i=1}^{n-1}(b_ic_{ij})^p &\le c\,\|x\|_{\mathbb{L}_p(I_+)}^{-p}\sum_{i=1}^{n-1}|I_i|^{(s-1)p+1}(i-j+1)_+^{(s-1)p}\\
&\le c\,\|x\|_{\mathbb{L}_p(I_+)}^{-p}\sum_{i=j}^{n-1}|I_i|^{(s-1)p+1}(i-j+1)^{(s-1)p}\\
&\le c\,\|x\|_{\mathbb{L}_p(I_+)}^{-p}\sum_{i=j}^{n-1}(n-i)^{(\beta-1)(sp-p+1)}n^{-\beta(sp-p+1)}(i-j+1)^{(s-1)p}\\
&\le c\,\|x\|_{\mathbb{L}_p(I_+)}^{-p}\Big(\frac{n-j}{n}\Big)^{\beta(sp-p+1)}.
\end{aligned}$$

We now take $\beta\in\mathbb{N}$ to be such that

$$\beta \ge p(sq-q+1)/(p-q). \tag{2.10}$$

Then, Lemma 3 (with $n-1$ instead of n) implies

$$\begin{aligned}
&\max_{\mathbf{w}\in\Omega_p^{n-1}(\mathbf{b};\mathbf{c})} \left(f_{q,n-1}(\mathbf{w};\mathbf{a})\right)^q \\
&\qquad \le n^{1-q} \max_{1\le j\le n-1}\left\{a_j^q\Big(\sum_{i=1}^{n-1}(b_i c_{ij})^p\Big)^{-q/p}\right\} \\
&\qquad \le c\,\|x\|^q_{\mathbb{L}_p(I_+)}\, n^{1-q} \max_{1\le j\le n-1}\left\{|I_j|^{sq-q+1}\Big(\frac{n-j}{n}\Big)^{-\beta q(sp-p+1)/p}\right\} \\
&\qquad \le c\,\|x\|^q_{\mathbb{L}_p(I_+)}\, n^{1-q-\beta+\beta q/p} \max_{1\le j\le n-1}\left\{(n-j)^{\beta(1-q/p)-(sq-q+1)}\right\} \\
&\qquad \le c\,\|x\|^q_{\mathbb{L}_p(I_+)}\, n^{-sq}.
\end{aligned}$$

Therefore, using the above as well as the observation that, if β satisfies (2.10), then $|I_n|^{1-q/p} = n^{-\beta(1-q/p)} < n^{-sq}$, we immediately get from (2.2)

$$\|x-\sigma_{s,n}\|^q_{\mathbb{L}_q(I_+)} \le c\,\|x\|^q_{\mathbb{L}_p(I_+)}\, n^{-sq} + \|x\|^q_{\mathbb{L}_p(I_n)}\; |I_n|^{1-q/p} \le c\,\|x\|^q_{\mathbb{L}_p(I_+)}\, n^{-sq},$$

and so (2.1) is proved for all $x \in \Delta^s_+\mathbb{L}_p$ satisfying (2.3).

In general, if $x \in \Delta^s_+\mathbb{L}_p$, then the function

$$\tilde{x}(t) := x(t) - T_{s-1}(t\,;x\,;0), \quad t \in I,$$

satisfies (2.3), and by Lemma 4 $\|\tilde{x}\|_{\mathbb{L}_p(I)} \le c\,\|x\|_{\mathbb{L}_p(I)}$, where c depends only on s and p. Therefore, it is enough to set

$$\sigma_{s,n}(t\,;x\,;I) := \sigma_{s,n}(t\,;\tilde{x}\,;I) + T_{s-1}(t\,;x\,;0), \quad t \in I,$$

in order to complete the proof of (2.1) for all $x \in \Delta^s_+\mathbb{L}_p$.

Denote by $\Sigma_{\beta,s,n} := \Sigma_{\beta,s,n}(I)$, where β satisfies (2.10), the space of piecewise polynomials $\sigma\colon I \mapsto \mathbb{R}$, of order s (of degree $\le s-1$), with knots at t_i, $i = \pm 1, \dots, \pm(n-1)$. Then, for $x \in \Delta^s_+B_p$, clearly $\sigma_{s,n}(\cdot\,;x\,;I) \in \Sigma_{\beta,s,n}$. Also, by our construction, the mapping $A\colon\ \mathrm{span}(\Delta^s_+B_p) \mapsto \Sigma_{\beta,s,n}$ defined by (2.2) is linear. Since $\dim(\Sigma_{\beta,s,n}) = s(2n-1)$, it follows by (1.1) that

$$d_n(\Delta^s_+B_p)^{\mathrm{psd}}_{\mathbb{L}_q} \le d_n(\Delta^s_+B_p)^{\mathrm{kol}}_{\mathbb{L}_q} \le d_n(\Delta^s_+B_p)^{\mathrm{lin}}_{\mathbb{L}_q} \le cn^{-s}, \quad n \ge s, \quad 0<q<p<1,$$

where $c = c(s,p,q)$. This proves the upper bound in (1.2).

3 Lower Bounds

3.1 Auxiliary Results

The following lemma can be proved in exactly the same way as [10, Lemma 2.2, p. 489] (also see [12, Claim 1]).

Lemma 1 *Let $m \in \mathbb{N}$ and $V^m := \{(v_1, \dots, v_m) : v_i = \pm 1, i = 1, \dots, m\}$. Then there exists a subset $V^{(m)} \subset V^m$ of cardinality $\geq 2^{m/16}$ such that for any $\mathbf{v}, \mathbf{u} \in V^{(m)}$, $\mathbf{v} \neq \mathbf{u}$, the distance $\|\mathbf{v} - \mathbf{u}\|_{l_1^m} \geq m/2$.*

The following property of the pseudo-dimension is well known (see [17], [6]) and immediately follows from the definition.

Lemma 2 *Let $T \neq \varnothing$ and $M := M(T)$ be a family of functions $x\colon T \mapsto \mathbb{R}$. Fix any function $y\colon T \mapsto \mathbb{R}$,*

$$\dim_{ps}\{z : z = y + x, x \in M\} = \dim_{ps}(M).$$

Lemma 3 **([3, Lemma 1])** *Let $I := (0,1)$, and let $a > 0$, $\varepsilon > 0$, and $m \in \mathbb{N}$, such that $m \geq 16\big(8 + \log_2(a/\varepsilon)\big)$, be given. Suppose that a set $\Phi^{(m)}$ of functions $\varphi \in \mathbb{L}_\infty(I)$ is such that*

$$\operatorname{card}(\Phi^{(m)}) \geq 2^{m/16},$$

$$\|\varphi\|_{\mathbb{L}_\infty(I)} \leq a, \quad \varphi \in \Phi^{(m)},$$

and for some $0 < q < 1$,

$$\|\phi_1 - \phi_2\|_{\mathbb{L}_q(I)} \geq \varepsilon, \quad \phi_1 \neq \phi_2, \quad \phi_1, \phi_2 \in \Phi^{(m)}.$$

Then for any $n \in \mathbb{N}$ such that $n \leq \Big(16\big(8 + \log_2(a/\varepsilon)\big)\Big)^{-1} m$ we have

$$d_n(\Phi^{(m)})^{\mathrm{psd}}_{\mathbb{L}_q(I)} \geq 2^{-2-1/q}(2^q - 1)^{1/q}\varepsilon.$$

3.2 Proof of the Lower Bounds in Theorem 1

Let $\varphi \in C^\infty(\mathbb{R})$ be nonnegative with $\operatorname{supp}\varphi = I_+ = [0,1]$, $\|\varphi\|_{\mathbb{L}_\infty(I)} = 1$, and $\varphi(t) = 1$ if $t \in [1/4, 3/4]$. Denote $\vartheta_s := \|\varphi^{(s)}\|^{-1}_{\mathbb{L}_\infty(I)}$.

For $s \in \mathbb{N}$, let

$$\phi_s(t) := \vartheta_s \varphi(t), \quad t \in \mathbb{R},$$

and for $m \in \mathbb{N}$ to be prescribed, take $t_i^* := t_{m,i}^* := i/m$, $0 \leq i \leq m$, and $I_i^* := I_{m,i}^* := [t_{i-1}^*, t_i^*]$, $1 \leq i \leq m$. Denote

$$\kappa := (sp+1)^{1/p} 2^{-1/p} s!,$$

and, for each $1 \leq i \leq m$, set

$$\phi_{s,m,i}(t) := \kappa m^{-s} \phi_s\big(m(t - t_{i-1}^*)\big), \quad t \in \mathbb{R}.$$

Then, $\operatorname{supp}\phi_{s,m,i} = I_i^*$, $\|\phi^{(s)}_{s,m,i}\|_{\mathbb{L}_\infty(I)} = \kappa$, $0 \leq \phi_{s,m,i}(t) \leq \kappa\vartheta_s m^{-s}$, $t \in I$, and $\phi_{s,m,i}(t) = \kappa\vartheta_s m^{-s}$, $t \in [t_{i-1}^* + 1/(4m), t_i^* - 1/(4m)]$.

Write

$$\Phi_s^{(m)} := \left\{ \phi : \phi = \sum_{i=1}^{m} v_i \phi_{s,m,i}, (v_1, \ldots, v_m) \in V^{(m)} \right\},$$

where $V^{(m)}$ is the class of sign-vectors defined in Lemma 1. Then, for all $\phi \in \Phi_s^{(m)}$,

$$\|\phi\|_{\mathbb{L}_\infty(I)} \le \kappa \vartheta_s m^{-s} \quad \text{and} \quad \|\phi^{(s)}\|_{\mathbb{L}_\infty(I)} \le \kappa.$$

Set

$$\psi_s(t) := \frac{\kappa}{s!} t_+^s, \quad t \in I.$$

It is easy to check that $\|\psi_s\|_{\mathbb{L}_p(I)} = 2^{-1/p}$, and $\psi_s^{(s)}(t) = \kappa$, $t \in I_+$.

Clearly, for all $\phi \in \Phi_s^{(m)}$, $\psi_s^{(s)}(t) + \phi^{(s)}(t) \ge 0$ a.e. on I, and

$$\|\psi_s + \phi\|_{\mathbb{L}_p(I)}^p \le \|\psi_s\|_{\mathbb{L}_p(I)}^p + \|\phi\|_{\mathbb{L}_p(I)}^p \le 1/2 + \kappa \vartheta_s m^{-s}.$$

Thus, for $m \ge (2\kappa)^{1/s} \vartheta_s^{1/s}$ we have $\psi_s + \Phi_s^{(m)} \subset \Delta_+^s B_p$, and so applying Lemma 2 we have

$$d_n(\Delta_+^s B_p)_{\mathbb{L}_q(I)}^{\text{psd}} \ge d_n(\psi_s + \Phi_s^{(m)})_{\mathbb{L}_q(I)}^{\text{psd}} = d_n(\Phi_s^{(m)})_{\mathbb{L}_q(I)}^{\text{psd}}, \quad n \ge 1. \tag{3.1}$$

For any two different vectors $\mathbf{v} := (v_1, \ldots, v_m)$ and $\mathbf{u} := (u_1, \ldots, u_m)$ in $V^{(m)}$, let

$$\phi_1 := \sum_{i=1}^{m} v_i \phi_{s,m,i} \quad \text{and} \quad \phi_2 := \sum_{i=1}^{m} u_i \phi_{s,m,i}$$

be the two corresponding functions in $\Phi_s^{(m)}$. Since $\|\mathbf{v} - \mathbf{u}\|_{l_1^m} \ge m/2$, then there exist $\lceil m/4 \rceil$ indices $i_1, \ldots, i_{\lceil m/4 \rceil}$ such that $v_{i_k} = -u_{i_k}$, $k = 1, \ldots, \lceil m/4 \rceil$. Hence,

$$\begin{aligned}
\|\phi_1 - \phi_2\|_{\mathbb{L}_q(I)}^q &= \int_I \Big| \sum_{i=1}^{m} (v_i - u_i) \phi_{s,m,i}(t) \Big|^q dt = \sum_{i=1}^{m} \int_{I_i^*} |v_i - u_i|^q \big(\phi_{s,m,i}(t) \big)^q dt \\
&\ge \kappa^q \vartheta_s^q m^{-sq} \sum_{k=1}^{\lceil m/4 \rceil} |v_{i_k} - u_{i_k}|^q \int_{t_{i_k}^* - 1 + \frac{1}{4m}}^{t_{i_k}^* - \frac{1}{4m}} dt \\
&= \kappa^q \vartheta_s^q m^{-sq} (2m)^{-1} \sum_{k=1}^{\lceil m/4 \rceil} 2^q \\
&\ge \kappa^q 2^{q-3} \vartheta_s^q m^{-sq} =: \varepsilon^q.
\end{aligned}$$

If we set $a := \kappa \vartheta_s m^{-s}$, and given $n \in \mathbb{N}$, we take $m = \lceil 80(\kappa 2^{3/q-1} + 1) \rceil n$, then applying Lemma 3, we conclude that

$$d_n(\Phi_s^{(m)})_{\mathbb{L}_q(I)} \ge c n^{-s}, \quad n \ge 1,$$

where $c = c(s, p, q)$. By virtue of (1.1) and (3.1) this implies

$$d_n(\Delta_+^s B_p)_{\mathbb{L}_q}^{\text{lin}} \ge d_n(\Delta_+^s B_p)_{\mathbb{L}_q}^{\text{kol}} \ge d_n(\Delta_+^s B_p)_{\mathbb{L}_q}^{\text{psd}} \ge c n^{-s}, \quad n \ge s,$$

where $c = c(s, p, q)$. This completes the proof of the lower bound and so of Theorem 1. ■

References

[1] P. S. Bullen, *A criterion for n-convexity.* Pacific J. Math. **36**(1971), 81–98.
[2] R. A. DeVore and G. G. Lorentz, *Constructive approximation.* Grundlehren Math. Wiss. **303**, Springer–Verlag, Berlin, 1993.
[3] Z. Ditzian, V. N. Konovalov and D. Leviatan, *Approximation of Sobolev-type classes with quasi-seminorms.* Rocky Mountain J. Math. **35**(2005), 445–478.
[4] N. Dyn, G. G. Lorentz and S. D. Riemenschneider, *Continuity of the Birkhoff interpolation.* SIAM J. Numer. Anal. **19**(1982), 507–509.
[5] G. H. Hardy, J. E. Littlewood and G. Pólya, *Inequalities.* Second edition, Cambridge University Press, 1952.
[6] D. Haussler, *Decision-theoretic generalizations of the PAC model for neural net and other learning applications.* Inform. and Comput. **100**(1992), 78–150.
[7] V. N. Konovalov and D. Leviatan, *Shape preserving widths of Sobolev-type classes of s-monotone functions on a finite interval.* Israel J. Math. **133**(2003), 239–268.
[8] K. A. Kopotun, *Whitney theorem of interpolatory type for k-monotone functions.* Constr. Approx. **17**(2001), 307–317.
[9] K. Kopotun and A. Shadrin, *On k-monotone approximation by free knot splines.* SIAM J. Math. Anal. **34**(2003), 901–924 (electronic).
[10] G. G. Lorentz, M. v. Golitschek and Y. Makovoz, *Constructive approximation. Advanced problems.* Grundlehren Math. Wiss. **304**, Springer–Verlag, Berlin, 1996.
[11] V. Maiorov and J. Ratsaby, *The degree of approximation of sets in Euclidean space using sets with bounded Vapnik–Chervonenkis dimension.* Discrete Appl. Math. **86**(1998), 81–93.
[12] ———, *On the degree of approximation by manifolds of finite pseudo-dimension.* Constr. Approx. **15**(1999), 291–300
[13] D. Pollard, *Convergence of stochastic processes.* Springer Series in Statistics, Springer–Verlag, New York, 1984.
[14] J. E. Pečarić, F. Proschan and Y. L. Tong, *Convex functions, partial orderings, and statistical applications.* Mathematics in Science and Engineering **187**, Academic Press Inc., Boston, MA, 1992.
[15] J. Ratsaby and V. Maiorov, *Generalization of the PAC-model for learning with partial information (extended abstract).* In: Computational learning theory (Jerusalem, 1997), Lecture Notes in Comput. Sci. **1208**(1997), 51–65.
[16] A. W. Roberts and D. E. Varberg, *Convex functions.* Pure and Applied Mathematics **57**, Academic Press, New York–London, 1973.
[17] R. S. Wenocur and R. M. Dudley, *Some special Vapnik–Chervonenkis classes.* Discrete Math. **33**(1981), 313–318.

Institute of Mathematics, National Academy of Sciences of Ukraine, Kyiv 01601, Ukraine
e-mail: vikono@imath.kiev.ua

Department of Mathematics, University of Manitoba, Winnipeg, Manitoba, R3T 2N2
e-mail: kopotunk@cc.umanitoba.ca