ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 40, Number 3, 2010

CONVEX POLYNOMIAL AND RIDGE APPROXIMATION OF LIPSCHITZ FUNCTIONS IN $\mathbf{R}^d$

V.N. KONOVALOV†, K.A. KOPOTUN AND V.E. MAIOROV

ABSTRACT. We consider classes of uniformly bounded convex Lipschitz functions defined on convex bodies in $\mathbf{R}^d$, $d \geq 1$, and obtain exact orders of approximation of these classes by convex polynomials in $\mathbf{L}_p$, $1 \leq p \leq \infty$. If $d > 1$, we also find exact orders of approximation by convex ridge functions in $\mathbf{L}_2$.

1. Introduction and main results. Shape Preserving Approximation (SPA) is a branch of approximation theory dealing with approximation of classes of functions having certain shape by elements from some (simpler) manifold (e.g., polynomials, splines, etc.) having the same shape. The word "shape" usually means positivity, monotonicity, convexity, k-monotonicity, etc., but SPA also deals with more general shape constraints (one-sided, intertwining, co-monotone, co-convex approximation and so on). We refer interested readers to [**9**, Chapter 2] for discussions of some of the earlier results on SPA. At present, most of the known results in this area deal with approximation of univariate functions. Very little is known about SPA in the multivariate settings. Perhaps, the most important result in this area up to date is due to Shvedov [**14**] who managed to estimate the rate of convex polynomial approximation (in the uniform norm) of convex functions defined on a convex body in $\mathbf{R}^d$, $d > 1$, in terms of the first modulus of continuity. It is still an open problem if the first modulus of continuity can be replaced by the second or third modulus, and if similar estimates are valid for convex approximation in the $\mathbf{L}_p$ norm.

Additionally, ridge approximation of convex functions is a rather effective tool as far as the degree of approximation is concerned. This can be seen from recent results on approximation of convex bodies by simpler sets (see [**1, 2**], for example).

† Vitya Konovalov unexpectedly passed away on November 1, 2008. This paper is dedicated to his memory.

Received by the editors on January 31, 2008.

DOI:10.1216/RMJ-2010-40-3-957 Copyright ©2010 Rocky Mountain Mathematics Consortium

Our main goal in this paper is to obtain exact orders of approximation of classes of convex Lipschitz functions defined on arbitrary convex bodies in $\mathbf{R}^d$ by convex polynomials and ridge functions in the $\mathbf{L}_p$ metrics for $1 \le p \le \infty$. Recall that a "convex body" in $\mathbf{R}^d$ is a closed bounded convex set with nonempty interior, and let $\mathcal{K}^d$ denote the set of all convex bodies in $\mathbf{R}^d$.

As usual, we use $|x|$ to denote the Euclidean norm of $x \in \mathbf{R}^d$. By $\mathrm{Lip}_\lambda(\mathfrak{B})$, $\lambda > 0$, we denote the class of all functions on $\mathfrak{B}$ satisfying the Lipschitz condition

$$|f(x) - f(y)| \le \lambda |x - y|, \text{ for all } x, y \in \mathfrak{B}. \tag{1}$$

Function $f : \mathfrak{B} \mapsto \mathbf{R}$, $\mathfrak{B} \in \mathcal{K}^d$, is convex if and only if its epigraph is a convex set in $\mathbf{R}^{d+1}$. By $\widehat{\mathcal{L}}_\lambda(\mathfrak{B})$, $\lambda > 0$, we denote the class of all convex functions f on $\mathfrak{B}$ such that $f \in \mathrm{Lip}_\lambda(\mathfrak{B})$ and $|f(x)| \le \lambda$, $x \in \mathfrak{B}$. If $\lambda = 1$, then $\widehat{\mathcal{L}}(\mathfrak{B}) := \widehat{\mathcal{L}}_1(\mathfrak{B})$.

We note that the condition $f \in \mathrm{Lip}_\lambda(\mathfrak{B})$ is not very restrictive if f is assumed to be convex. In fact, if f is convex on an open set S containing $\mathfrak{B}$ then f is Lipschitz on $\mathfrak{B}$ (see [**13**, Theorem 41D]). Also, boundedness of f is added in the definition of the class $\widehat{\mathcal{L}}_\lambda(\mathfrak{B})$ for convenience only since, if f satisfies (1) on $\mathfrak{B}$, then $|f(x)| \le |f(x_0)| + \lambda \mathrm{diam}\,(\mathfrak{B})$, $x \in \mathfrak{B}$, for any fixed $x_0 \in \mathfrak{B}$.

As usual, if Ω is a measurable subset of $\mathbf{R}^d$, we denote by $\mathbf{L}_p(\Omega)$, $1 \le p \le \infty$, the linear space of all Lebesgue measurable functions $f : \Omega \mapsto \mathbf{R}$ equipped with the finite norm

$$\|f\|_{\mathbf{L}_p(\Omega)} := \begin{cases} \left(\int_\Omega |f(x)|^p \, dx\right)^{1/p} & 1 \le p < \infty, \\ \operatorname{ess\,sup}_{x\in\Omega} |f(x)| & p = \infty. \end{cases}$$

By $\mathcal{P}_n(\mathbf{R}^d)$, $n \in \mathbf{N}$, we denote the linear space of all algebraic polynomials $P_n(x)$, $x = (x_1, \dots , x_d) \in \mathbf{R}^d$, of total degree $\le n$, i.e.,

$$P_n(x) := \sum_{0 \le |k| \le n} a_k x^k,$$

where $k = (k_1, \dots , k_d) \in \mathbf{Z}^d_+$, $|k| := k_1 + \cdots + k_d$, $a_k \in \mathbf{R}$, $x^k := x_1^{k_1} \cdots x_d^{k_d}$, and $0^0 := 1$. Also, $\mathcal{P}_n(S)$ is the restriction of the space $\mathcal{P}_n(\mathbf{R}^d)$ to the set $S \subset \mathbf{R}^d$, and $\widehat{\mathcal{P}}_n(\mathfrak{B})$ denotes the subset of polynomials from $\mathcal{P}_n(\mathfrak{B})$ which are convex on $\mathfrak{B}$.

Besides algebraic polynomials we also consider ridge functions. Let $\mathbf{L}_2^{\mathrm{loc}}(\mathbf{R})$ denote the space of all locally 2nd-power summable univariate functions, i.e., Borel measurable functions such that $\|f\|_{\mathbf{L}_2(K)} < \infty$ for every compact subset K of $\mathbf{R}$. If $\{f_\nu\}_{\nu=1}^n$ is a collection of univariate functions f_ν from $\mathbf{L}_2^{\mathrm{loc}}(\mathbf{R})$ and $\{e_\nu\}_{\nu=1}^n$ is a collection of unit vectors from $\mathbf{R}^d$, then a function of type

$$R_n(x) := \sum_{\nu=1}^{n} f_\nu(e_\nu \cdot x), \quad x \in \mathbf{R}^d,$$

is called a ridge function. Here, $a \cdot b := a_1 b_1 + \cdots + a_d b_d$ is the dot product of vectors $a = (a_1, \ldots, a_d)$ and $b = (b_1, \ldots, b_d)$. The set of all ridges is denoted by $\mathcal{R}_n(\mathbf{R}^d)$, and by $\mathcal{R}_n(S)$ we denote the restriction of $\mathcal{R}_n(\mathbf{R}^d)$ to $S \subset \mathbf{R}^d$. For $\mathfrak{B} \in \mathcal{K}^d$, $\widehat{\mathcal{R}}_n(\mathfrak{B})$ denotes the subset of all ridges from $\mathcal{R}_n(\mathfrak{B})$ which are convex on $\mathfrak{B}$.

If $\mathfrak{B} \subset \mathbf{R}^d$ is fixed, then in order to simplify the notation denote $\widehat{\mathcal{L}}^d := \widehat{\mathcal{L}}(\mathfrak{B})$, $\mathcal{P}_n^d := \mathcal{P}_n(\mathfrak{B})$, $\widehat{\mathcal{P}}_n^d := \widehat{\mathcal{P}}_n(\mathfrak{B})$, $\mathcal{R}_n^d := \mathcal{R}_n(\mathfrak{B})$, $\widehat{\mathcal{R}}_n^d := \widehat{\mathcal{R}}_n(\mathfrak{B})$ and $\mathbf{L}_p^d := \mathbf{L}_p(\mathfrak{B})$.

Now, the error of approximation of $f \in \mathbf{L}_p^d$ by elements from $W \subset \mathbf{L}_p^d$ is

$$E\big(f, W\big)_{\mathbf{L}_p^d} := \inf_{p \in W} \|f - p\|_{\mathbf{L}_p^d},$$

and the quantity

$$E\big(\widehat{\mathcal{L}}^d, W\big)_{\mathbf{L}_p^d} := \sup_{f \in \widehat{\mathcal{L}}^d} E\big(f, W\big)_{\mathbf{L}_p^d}$$

is called the deviation of $\widehat{\mathcal{L}}^d$ from W. In particular, $E(\widehat{\mathcal{L}}^d, \widehat{\mathcal{P}}_n^d)_{\mathbf{L}_p^d}$ and $E(\widehat{\mathcal{L}}^d, \widehat{\mathcal{R}}_n^d)_{\mathbf{L}_2^d}$ are the rates of best convex approximation of the class $\widehat{\mathcal{L}}^d$ by algebraic polynomials and ridges in the $\mathbf{L}_p^d$ and $\mathbf{L}_2^d$ norms, respectively.

For two sequences $\{a_n\}_{n=1}^\infty$ and $\{b_n\}_{n=1}^\infty$ of positive numbers, we write $a_n \asymp b_n$, $n \geq 1$, if and only if there exist positive constants c and C which are independent of n and such that $c \leq a_n / b_n \leq C$, $n \geq 1$.

We are now ready to state our main results.

Theorem 1. *Let* $\mathfrak{B} \in \mathcal{K}^d$, $d \in \mathbf{N}$, *and* $1 \le p \le \infty$. *Then*

$$E\big(\widehat{\mathcal{L}}^d, \mathcal{P}_n^d\big)_{\mathbf{L}_p^d} \asymp E\big(\widehat{\mathcal{L}}^d, \widehat{\mathcal{P}}_n^d\big)_{\mathbf{L}_p^d} \asymp n^{-1-1/p}, \quad n \ge 1. \tag{2}$$

Theorem 2. *If* $d > 1$, *then*

$$E\big(\widehat{\mathcal{L}}^d, \mathcal{R}_n^d\big)_{\mathbf{L}_2^d} \asymp E\big(\widehat{\mathcal{L}}^d, \widehat{\mathcal{R}}_n^d\big)_{\mathbf{L}_2^d} \asymp n^{-3/(2(d-1))}, \quad n \ge 1. \tag{3}$$

Note that, in the case $p = \infty$, the upper estimate in (2) immediately follows from [**14**, Theorem 1].

Everywhere below, $c = c(\alpha, \dots, \beta)$ denote positive constants that depend only upon parameters $\alpha, \dots, \beta$ which may be different even if they occur in the same line, and $c_i = c_i(\alpha, \dots, \beta)$, $i \ge 0$, denote constants which remain fixed throughout the paper.

Also, if $\mathbf{0} = (0, \dots, 0) \in \mathbf{R}^d$ and S is a set in $\mathbf{R}^d$, we denote by μS the dilation of S with the center $\mathbf{0}$ and dilation factor μ. Thus, for example, if $\mathbf{B}^d$ denotes the Euclidean unit ball in $\mathbf{R}^d$, then $\mu\mathbf{B}^d$ is the ball of radius μ and center at $\mathbf{0}$. Finally, we denote by

$$\omega\big(\delta; f, S\big) := \sup_{\substack{x,y \in S \\ |x-y| \le \delta}} |f(x) - f(y)|, \quad \delta > 0,$$

the (usual) modulus of continuity of f on $S \subset \mathbf{R}^d$, and note that $f \in \mathrm{Lip}_\lambda(\mathfrak{B})$ if and only if $\omega(\delta; f, \mathfrak{B}) \le \lambda\delta$, for all $\delta > 0$.

2. Proof of upper estimates in Theorem 1. The following lemma on convex extension of Lipschitz functions can be found in [**10**, Lemma 3].

Lemma 3. *Let* $\mathfrak{B} \in \mathcal{K}^d$ *and* $f \in \mathrm{Lip}_1(\mathfrak{B})$. *Then,* f *can be extended to a function* $\tilde{f}$ *(i.e.,* $\tilde{f}(x) = f(x)$ *for all* $x \in \mathfrak{B}$*) which is convex on* $\mathbf{R}^d$ *and satisfies the Lipschitz condition on* $\mathbf{R}^d$ *with the constant* $\sqrt{2}$*, i.e.,* $\tilde{f} \in \mathrm{Lip}_{\sqrt{2}}(\mathbf{R}^d)$.

We note that it was shown in [**10**] that $\tilde{f}$ can be defined as follows:

$$\tilde{f}(x) := \inf\big\{y \in \mathbf{R} : (x, y) \in CO\{(x, g(x)) \in \mathbf{R}^d \times \mathbf{R} : x \in \mathbf{R}^d\}\big\},$$

where $g(x) := f(P(x))+|x-P(x)|$, $P(x) := \arg\min_{y\in\mathfrak{B}} |y-x|$, $x \in \mathbf{R}^d$, and $CO(S)$ denotes the convex hull of S.

2.1. The case $d = 1$. In the univariate case, the set $\mathcal{K}^1$ consists of all closed intervals. Given $\mathfrak{B} = [a, b]$ and $f \in \widehat{\mathcal{L}}[a, b]$, we note that the function $\tilde{f}(x) := f((b-a)x/2+(a+b)/2)$ belongs to the class $\widehat{\mathcal{L}}_\lambda[-1, 1]$ with $\lambda := \max\{1, (b-a)/2\}$ (or we could assume that $\mathfrak{B} \subset [-1, 1]$ and use Lemma 3 to extend f to $[-1, 1]$ preserving its convexity). Hence, without loss of generality, we can assume that $\mathfrak{B} = \mathfrak{I} := [-1, 1]$.

Recall now that the second Ditzian-Totik modulus of $f \in \mathbf{L}_p(\mathfrak{I})$ is defined by

$$\omega_2^\varphi(f, t)_p := \sup_{0<h\le t} \|\Delta^2_{h\varphi(\cdot)}(f, \cdot)\|_p,$$

where $\varphi(x) := \sqrt{1-x^2}$, $\|\cdot\|_p := \|\cdot\|_{\mathbf{L}_p(\mathfrak{I})}$, and $\Delta^2_\mu(f, x) := f(x-\mu) - 2f(x) + f(x+\mu)$, if $x \pm \mu \in \mathfrak{I}$ and $\Delta^2_\mu(f, x) := 0$, otherwise. Denote $\omega(f, \delta)_\infty := \omega(\delta; f, \mathfrak{I})$ and note that $\omega_2^\varphi(f, t)_\infty \le c\omega(f, \delta)_\infty$.

It is known (see, e.g., [**8, 15**]) that, for a convex function $f \in \mathbf{L}_p(\mathfrak{I})$, $1 \le p \le \infty$,

$$E(f, \widehat{\mathcal{P}}_n(\mathfrak{I}))_{\mathbf{L}_p(\mathfrak{I})} \le c\omega_2^\varphi(f, 1/n)_p, \quad n \ge 1.$$

Now, the inequality $\|g\|_p \le \|g\|_\infty^{1-1/p}\|g\|_1^{1/p}$, $1 \le p \le \infty$, implies

$$\omega_2^\varphi(f, \delta)_p \le \omega_2^\varphi(f, \delta)_\infty^{1-1/p}\omega_2^\varphi(f, \delta)_1^{1/p} \le c\omega(f, \delta)_\infty^{1-1/p}\omega_2^\varphi(f, \delta)_1^{1/p},$$

and therefore

$$E(f, \widehat{\mathcal{P}}_n(\mathfrak{I}))_{\mathbf{L}_p(\mathfrak{I})} \le cn^{-1+1/p}\omega_2^\varphi(f, 1/n)_1^{1/p},$$

for every $f \in \widehat{\mathcal{L}}(\mathfrak{I})$ since, for all functions from this class, $\omega(f, \delta)_\infty \le \delta$.

It was shown in [**7**, Theorem 1.1] that if $1 \le p < q \le \infty$ and $f \in \mathbf{L}_q(\mathfrak{I})$ is convex, then

$$\omega_2^\varphi(f, \delta)_p \le c\delta^{2/p-2/q}\|f\|_{\mathbf{L}_q(\mathfrak{I})},$$

and, in particular, the inequality $\omega_2^\varphi(f, \delta)_1 \le c\delta^2\|f\|_{\mathbf{L}_\infty(\mathfrak{I})}$ is satisfied for every convex function $f \in \mathbf{L}_\infty(\mathfrak{I})$. Hence,

$$E(f, \widehat{\mathcal{P}}_n(\mathfrak{I}))_{\mathbf{L}_p(\mathfrak{I})} \le cn^{-1-1/p}\|f\|^{1/p}_{\mathbf{L}_\infty(\mathfrak{I})},$$

and it remains to recall that $\|f\|_{\mathbf{L}_\infty(\mathfrak{I})} \le 1$ for $f \in \widehat{\mathcal{L}}(\mathfrak{I})$ in order to conclude that

$$E(\widehat{\mathcal{L}}(\mathfrak{I}), \widehat{\mathcal{P}}_n(\mathfrak{I}))_{\mathbf{L}_p(\mathfrak{I})} \le cn^{-1-1/p}.$$

The proof of upper estimates in (2) in the case $d = 1$ is now complete since

$$E(\widehat{\mathcal{L}}(\mathfrak{I}), \mathcal{P}_n(\mathfrak{I}))_{\mathbf{L}_p(\mathfrak{I})} \le E(\widehat{\mathcal{L}}(\mathfrak{I}), \widehat{\mathcal{P}}_n(\mathfrak{I}))_{\mathbf{L}_p(\mathfrak{I})}.$$

2.2. The case $d > 1$. Without loss of generality, we can assume that $\mathfrak{B} \subset Q^d := [-1,1]^d$. Using Lemma 3 we extend $f \in \widehat{\mathcal{L}}(\mathfrak{B})$ to the cube $4Q^d$ so that the resulting function $\tilde{f}$ will be from the class $\widehat{\mathcal{L}}_\lambda(4Q^d)$ for some $\lambda = \lambda(d, \mathfrak{B}) > 0$. We now smooth $\tilde{f}$ by considering its second Steklov mean (see also construction in [**14**, page 520])

$$\tilde{f}_\varepsilon(x) := (2\varepsilon)^{-2d} \int_{\varepsilon Q^d} \int_{\varepsilon Q^d} \tilde{f}(x + t_1 + t_2)\, dt_1\, dt_2, \quad x \in 3Q^d,$$

where $0 < \varepsilon < 1/2$ will be chosen later. For any $x \in 3Q^d$, we have

$$\begin{aligned} |\tilde{f}_\varepsilon(x) - \tilde{f}(x)| &\le (2\varepsilon)^{-2d} \int_{\varepsilon Q^d} \int_{\varepsilon Q^d} \left| \tilde{f}(x + t_1 + t_2) - \tilde{f}(x) \right| dt_1\, dt_2 \\ &\le \lambda (2\varepsilon)^{-2d} \int_{\varepsilon Q^d} \int_{\varepsilon Q^d} |t_1 + t_2|\, dt_1\, dt_2 \le 2\lambda\sqrt{d}\varepsilon, \end{aligned} \tag{4}$$

and it is not difficult to show that

$$\left\| \frac{\partial}{\partial x_i} \tilde{f}_\varepsilon \right\|_{\mathbf{L}_\infty(3Q^d)} \le (2\varepsilon)^{-1} \omega(2\varepsilon; \tilde{f}, 4Q^d) \le \lambda, \quad 1 \le i \le d. \tag{5}$$

Also, note that $\tilde{f}_\varepsilon$ is convex on $3Q^d$, and $\tilde{f}_\varepsilon \in \widehat{\mathcal{L}}_\lambda(3Q^d)$.

Let $T_n(x) := \cos(n \arccos t)$, $n \ge 0$, be the Chebyshev polynomial of degree n. We recall that $T_{n+1}(t) = 2tT_n(t) - T_{n-1}(t)$ and, in particular, if we let $\tau_n(t) := T_{2n+1}(t)/t$, then $\tau_n(t) = 2T_{2n}(t) - \tau_{n-1}(t)$, $n \in \mathbf{N}$, and $\tau_0(t) = 1$. Hence, τ_n is an even polynomial of degree $2n$. In order to construct approximating polynomials we use the following Jackson-type kernels

$$J_n(t) := j_n^{-1} \left[\tau_n(t/3)\right]^8, \quad j_n := \int_{-3}^{3} \left[\tau_n(t/3)\right]^8\, dt.$$

The kernels $J_n(t)$ have the following properties which can be easily verified by straightforward computations (see also [**4, 5**]):

(i) J_n is a polynomial of degree $16n$ which is non-negative and even on $[-3,3]$;

(ii) $\int_{-3}^{3} J_n(t)\,dt = 1$;

(iii) $j_n \asymp n^7$, $n \geq 1$;

(iv) $J_n(t) \leq cn(n|t|+1)^{-8}$, $t \in [-3,3]$;

(v) $\int_{-3}^{-1} J_n(t)\,dt = \int_{1}^{3} J_n(t)\,dt \leq cn^{-7}$;

(vi) $\|J_n\|_{\mathbf{L}_\infty[-3,-1]} = \|J_n\|_{\mathbf{L}_\infty[1,3]} \leq cn^{-7}$;

(vii) $\|J_n'\|_{\mathbf{L}_\infty[-3,-1]} = \|J_n'\|_{\mathbf{L}_\infty[1,3]} \leq cn^{-5}$.

We now define

$$\mathbf{J}_n(x) := \prod_{i=1}^{d} J_n(x_i), \quad x = (x_1, \dots, x_d) \in 3Q^d,$$

and note that the properties of J_n imply that

$$\int_{3Q^d} |x|^k \mathbf{J}_n(x)\,dx \leq cn^{-k}, \quad 0 \leq k \leq 8.$$

We are now ready to define polynomials $P_n(\cdot; f)$ which we use to approximate functions from the class $\widehat{\mathcal{L}}^d(\mathfrak{B})$. Let

$$P_n(x; f) := \int_{2Q^d} \tilde{f}_\varepsilon(y)\mathbf{J}_n(x-y)\,dy + \alpha n^{-2}|x|^2, \quad x \in Q^d, \tag{6}$$

where $\varepsilon := n^{-2}$, and $\alpha = \alpha(d, \mathfrak{B}) > 0$ is a parameter which will be chosen below so that to make polynomials $P_n(\cdot; f)$ convex on Q^d.

2.2.1. Convexity of $P_n(f)$. Suppose that $x \in Q^d$, and $e = (e_1, \dots, e_d)$ is an arbitrary fixed unit vector from $\mathbf{R}^d$. Differentiating (6) twice in the direction e we get

$$\mathcal{D}_e^2 P_n(x; f) = \int_{2Q^d} \tilde{f}_\varepsilon(y)\mathcal{D}_e^2 \mathbf{J}_n(x-y)dy + \alpha n^{-2}\mathcal{D}_e^2 |x|^2.$$

Since

$$\mathcal{D}_e^2 = \sum_{i=1}^d \left(\frac{\partial^2}{\partial x_i^2}\right) e_i^2 + \sum_{i\neq j} \left(\frac{\partial^2}{\partial x_i \partial x_j}\right) e_i e_j,$$

where $\sum_{i\neq j}$ is the summation over all $1 \leq i,\, j \leq d$ such that $i \neq j$, we get

$$\mathcal{D}_e^2 |x|^2 = \mathcal{D}_e^2 \bigg(\sum_{k=1}^d x_k^2\bigg) = 2\sum_{i=1}^d e_i^2 = 2$$

and

$$\begin{aligned}\mathcal{D}_e^2 \mathbf{J}_n(x-y) &= \sum_{i=1}^d \left(\frac{\partial^2}{\partial x_i^2}\,\mathbf{J}_n(x-y)\right) e_i^2 \\ &\quad + \sum_{i\neq j} \left(\frac{\partial^2}{\partial x_i \partial x_j}\,\mathbf{J}_n(x-y)\right) e_i e_j \\ &= \sum_{i=1}^d \left(\frac{\partial^2}{\partial y_i^2}\,\mathbf{J}_n(x-y)\right) e_i^2 \\ &\quad + \sum_{i\neq j} \left(\frac{\partial^2}{\partial y_i \partial y_j}\,\mathbf{J}_n(x-y)\right) e_i e_j.\end{aligned}$$

It now follows from the definition of $\mathbf{J}_n$ that

$$\frac{\partial^2}{\partial y_i^2}\mathbf{J}_n(x-y) = \Bigg(\prod_{\substack{k=1\\k\neq i}}^d J_n(x_k-y_k)\Bigg) J_n''(x_i-y_j),$$

and

$$\frac{\partial^2}{\partial y_i \partial y_j}\mathbf{J}_n(x-y) = \Bigg(\prod_{\substack{k=1\\k\neq i,j}}^d J_n(x_k-y_k)\Bigg) J_n'(x_i-y_i) J_n'(x_j-y_j).$$

Integrating by parts twice with respect to the ith variable we get

$$\begin{aligned}\int_{-2}^{2} \tilde f_\varepsilon(\dots, y_i, \dots) J_n''(x_i - y_i)\, dy_i \\ &= \int_{-2}^{2} J_n(x_i - y_i) \frac{\partial^2}{\partial y_i^2} \tilde f_\varepsilon(\dots, y_i, \dots)\, dy_i \\ &\quad - \tilde f_\varepsilon(\dots, 2, \dots) J_n'(x_i - 2) + \tilde f_\varepsilon(\dots, -2, \dots) J_n'(x_i + 2) \\ &\quad - \frac{\partial}{\partial y_i} \tilde f_\varepsilon(\dots, 2, \dots) J_n(x_i - 2) \\ &\quad + \frac{\partial}{\partial y_i} \tilde f_\varepsilon(\dots, -2, \dots) J_n(x_i + 2) \\ &=: \int_{-2}^{2} J_n(x_i - y_i) \frac{\partial^2}{\partial y_i^2} \tilde f_\varepsilon(\dots, y_i, \dots)\, dy_i \\ &\quad + \sigma_i(x; \tilde f_\varepsilon, J_n).\end{aligned}$$

Similarly, integration by parts with respect to the ith and jth variables yields

$$\begin{aligned}&\int_{-2}^{2}\int_{-2}^{2} \tilde f_\varepsilon(\dots, y_i, \dots, y_j, \dots) J_n'(x_i - y_i) J_n'(x_j - y_j)\, dy_i\, dy_j \\ &= \int_{-2}^{2}\int_{-2}^{2} J_n(x_i - y_i) J_n(x_j - y_j) \frac{\partial^2}{\partial y_i \partial y_j} \tilde f_\varepsilon(\dots, y_i, \dots, y_j, \dots)\, dy_i\, dy_j \\ &\quad + \sigma_{i,j}(x; \tilde f_\varepsilon, J_n),\end{aligned}$$

where

$$\begin{aligned}\sigma_{i,j}(x; \tilde f_\varepsilon, J_n) \\ &:= \tilde f_\varepsilon(\dots, 2, \dots, 2, \dots) J_n(x_i - 2) J_n(x_j - 2) \\ &\quad - \tilde f_\varepsilon(\dots, 2, \dots, -2, \dots) J_n(x_i - 2) J_n(x_j + 2) \\ &\quad - \tilde f_\varepsilon(\dots, -2, \dots, 2, \dots) J_n(x_i + 2) J_n(x_j - 2) \\ &\quad + \tilde f_\varepsilon(\dots, -2, \dots, -2, \dots) J_n(x_i + 2) J_n(x_j + 2) \\ &\quad - J_n(x_i - 2) \int_{-2}^{2} J_n(x_j - y_j) \frac{\partial}{\partial y_j} \tilde f_\varepsilon(\dots, 2, \dots, y_j, \dots) dy_j \\ &\quad + J_n(x_i + 2) \int_{-2}^{2} J_n(x_j - y_j) \frac{\partial}{\partial y_j} \tilde f_\varepsilon(\dots, -2, \dots, y_j, \dots) dy_j\end{aligned}$$

$$- J_n(x_j - 2) \int_{-2}^{2} J_n(x_i - y_i) \frac{\partial}{\partial y_i} \tilde{f}_\varepsilon(\dots, y_i, \dots, 2, \dots)\, dy_i$$
$$+ J_n(x_j + 2) \int_{-2}^{2} J_n(x_i - y_i) \frac{\partial}{\partial y_i} \tilde{f}_\varepsilon(\dots, y_i, \dots, -2, \dots)\, dy_i.$$

Estimate (5) and properties of the kernels J_n imply that

$$|\sigma_i(x; \tilde{f}_\varepsilon, J_n)| \le c(d, \lambda) n^{-2}$$

and

$$|\sigma_{i,j}(x; \tilde{f}_\varepsilon, J_n)| \le c(d, \lambda) n^{-2}, \text{for all } x \in Q^d.$$

Hence, we conclude that

$$\mathcal{D}_e^2 P_n(x; f) = \int_{2Q^d} \mathbf{J}_n(x - y) \mathcal{D}_e^2 \tilde{f}_\varepsilon(y)\, dy + \sigma(x; \tilde{f}_\varepsilon, J_n, e) + 2\alpha n^{-2},$$

where $\sigma(x; \tilde{f}_\varepsilon, J_n, e)$ satisfies the inequality

$$|\sigma(x; \tilde{f}_\varepsilon, J_n, e)| \le c_0(d, \lambda) n^{-2}, \quad x \in Q^d.$$

We now choose $\alpha := c_0/2$, and note that this implies that $\sigma(x; \tilde{f}_\varepsilon, J_n, e) + 2\alpha n^{-2} \ge 0$, for all $x \in Q^d$. Hence, since $\tilde{f}_\varepsilon$ is convex and the kernel $\mathbf{J}_n$ is nonnegative, we conclude that the second derivative of $P_n(x; f)$ in any direction e at any point $x \in Q^d$ is nonnegative. Therefore, $P_n(\cdot; f)$ is convex on Q^d.

2.2.2. Approximation properties of $P_n(f)$. Given a set $S \subset R^d$, we denote by $S(x) := x + S$ its shift by the vector $x \in \mathbf{R}^d$.

We now rewrite polynomials $P_n(f)$ in the form which will make it more convenient to estimate their deviation from f. Changing variables we get

$$
\begin{aligned}
P_n(x;f) - \alpha n^{-2}|x|^2 &= \int_{2Q^d} \tilde{f}_\varepsilon(y)\mathbf{J}_n(x-y)\,dy \\
&= \int_{2Q^d(x)} \tilde{f}_\varepsilon(x-y)\mathbf{J}_n(y)\,dy \\
&= \int_{2Q^d} \tilde{f}_\varepsilon(x-y)\mathbf{J}_n(y)\,dy \\
&\quad + \int_{2Q^d(x)\setminus 2Q^d} \tilde{f}_\varepsilon(x-y)\mathbf{J}_n(y)\,dy \\
&\quad - \int_{2Q^d\setminus 2Q^d(x)} \tilde{f}_\varepsilon(x-y)\mathbf{J}_n(y)\,dy.
\end{aligned}
$$

Additionally, using property (ii) of kernels J_n, we have

$$
\begin{aligned}
\tilde{f}_\varepsilon(x) &= \int_{3Q^d} \tilde{f}_\varepsilon(x)\mathbf{J}_n(y)\,dy \\
&= \int_{2Q^d} \tilde{f}_\varepsilon(x)\mathbf{J}_n(y)\,dy \\
&\quad + \tilde{f}_\varepsilon(x)\int_{3Q^d\setminus 2Q^d} \mathbf{J}_n(y)\,dy.
\end{aligned}
$$

Since $f(x) = f(x) \mp \tilde{f}_\varepsilon(x)$, we have

(7)

$$
f(x) - P_n(x;f) = -\int_{2Q^d} (\tilde{f}_\varepsilon(x-y) - \tilde{f}_\varepsilon(x))\mathbf{J}_n(y)\,dy + \sigma(x;f,\tilde{f}_\varepsilon,\mathbf{J}_n),
$$

where

$$
\begin{aligned}
\sigma(x;f,\tilde{f}_\varepsilon,\mathbf{J}_n) := \tilde{f}_\varepsilon(x)&\int_{3Q^d\setminus 2Q^d} \mathbf{J}_n(y)\,dy \\
&- \int_{2Q^d(x)\setminus 2Q^d} \tilde{f}_\varepsilon(x-y)\mathbf{J}_n(y)\,dy \\
&+ \int_{2Q^d\setminus 2Q^d(x)} \tilde{f}_\varepsilon(x-y)\mathbf{J}_n(y)\,dy \\
&+ f(x) - \tilde{f}_\varepsilon(x) - \alpha n^{-2}|x|^2.
\end{aligned}
$$

Recalling that $\tilde{f} \in \widehat{\mathcal{L}}_\lambda(4Q^d)$ which, in particular, implies that $|\tilde{f}_\varepsilon(x)| \le \lambda$, $x \in 4Q^d$, we estimate the norm of $\sigma(x; f, \tilde{f}_\varepsilon, \mathbf{J}_n)$ in $\mathbf{L}_p(\mathfrak{B})$. Using property (vi) of the kernels J_n and the fact that $Q^d \subset 2Q^d(x) \subset 3Q^d$, $x \in Q^d$, we have for any $x \in Q^d$,

$$\begin{aligned} |\tilde{f}_\varepsilon(x)| \int_{3Q^d \setminus 2Q^d} \mathbf{J}_n(y)\, dy &\le \lambda \int_{3Q^d \setminus Q^d} \mathbf{J}_n(y)\, dy \\ &\le c(d, \mathfrak{B}) n^{-7}, \\ \int_{2Q^d(x) \setminus 2Q^d} |\tilde{f}_\varepsilon(x-y)| \mathbf{J}_n(y)\, dy &\le \lambda \int_{3Q^d \setminus Q^d} \mathbf{J}_n(y)\, dy \\ &\le c(d, \mathfrak{B}) n^{-7}, \end{aligned}$$

and

$$\begin{aligned} \int_{2Q^d \setminus 2Q^d(x)} |\tilde{f}_\varepsilon(x-y)| \mathbf{J}_n(y)\, dy &\le \lambda \int_{3Q^d \setminus Q^d} \mathbf{J}_n(y)\, dy \\ &\le c(d, \mathfrak{B}) n^{-7}. \end{aligned}$$

Now, the fact that $\tilde{f} \equiv f$ on $\mathfrak{B}$ and inequality (4) imply

$$\max_{x \in \mathfrak{B}} |f(x) - \tilde{f}_\varepsilon(x)| \le \max_{x \in Q^d} |\tilde{f}(x) - \tilde{f}_\varepsilon(x)| \le c(d, \mathfrak{B}) n^{-2}.$$

Finally, it is obvious that $\alpha |x|^2 n^{-2} \le \alpha\, d n^{-2}$, $x \in Q^d$, and so

$$\left\| \sigma(\cdot; f, \tilde{f}_\varepsilon, \mathbf{J}_n) \right\|_{\mathbf{L}_p(\mathfrak{B})} \le c(d, \mathfrak{B}) n^{-2}, \quad 1 \le p \le \infty. \tag{8}$$

It remains to estimate the norm of

$$\mathcal{I}(x; \tilde{f}_\varepsilon, \mathbf{J}_n) := \int_{2Q^d} \big(\tilde{f}_\varepsilon(x-y) - \tilde{f}_\varepsilon(x)\big) \mathbf{J}_n(y)\, dy.$$

We need to show that, for $1 \le p \le \infty$,

$$\left\| \mathcal{I}(\cdot; \tilde{f}_\varepsilon, \mathbf{J}_n) \right\|_{\mathbf{L}_p(Q^d)} \le c n^{-1-1/p}. \tag{9}$$

Since

$$\left\|\mathcal{I}(\cdot;\tilde{f}_\varepsilon,\mathbf{J}_n)\right\|_{\mathbf{L}_p(Q^d)} \le \left\|\mathcal{I}(\cdot;\tilde{f}_\varepsilon,\mathbf{J}_n)\right\|_{\mathbf{L}_\infty(Q^d)}^{1-1/p} \left\|\mathcal{I}(\cdot;\tilde{f}_\varepsilon,\mathbf{J}_n)\right\|_{\mathbf{L}_1(Q^d)}^{1/p}, \quad 1 \le p \le \infty,$$

it is sufficient to verify (9) only in the cases $p=\infty$ and $p=1$.

The case $p=\infty$. Using the fact that $\tilde{f}_\varepsilon \in \widehat{\mathcal{L}}_\lambda(3Q^d)$, we have

$$\left\|\mathcal{I}(\cdot;\tilde{f}_\varepsilon,\mathbf{J}_n)\right\|_{\mathbf{L}_\infty(Q^d)} \le \lambda \int_{2Q^d} |y|\mathbf{J}_n(y)\,dy \le cn^{-1}.$$

The case $p=1$. Using the fact that the kernel $\mathbf{J}_n(y) = \mathbf{J}_n(y_1,\dots,y_d)$ is even with respect to every variable we get

$$\int_{2Q^d} \tilde{f}_\varepsilon(x-y)\mathbf{J}_n(y)\,dy = \int_{2Q^d} \tilde{f}_\varepsilon(x+y)\mathbf{J}_n(y)\,dy, \quad x \in Q^d,$$

and therefore

$$\begin{aligned}\mathcal{I}(x;\tilde{f}_\varepsilon,\mathbf{J}_n) &= \frac{1}{2}\int_{2Q^d} \big(\tilde{f}_\varepsilon(x-y) - 2\tilde{f}_\varepsilon(x) + \tilde{f}_\varepsilon(x+y)\big)\mathbf{J}_n(y)\,dy \\ &=: \frac{1}{2}\int_{2Q^d} \Delta_y^2\tilde{f}_\varepsilon(x)\mathbf{J}_n(y)\,dy,\end{aligned}$$

where

$$\Delta_y^2\tilde{f}_\varepsilon(\cdot) := \tilde{f}_\varepsilon(\cdot+y) - 2\tilde{f}_\varepsilon(\cdot) + \tilde{f}_\varepsilon(\cdot-y).$$

Therefore,

$$\begin{aligned}\left\|\mathcal{I}(\cdot;\tilde{f}_\varepsilon,\mathbf{J}_n)\right\|_{\mathbf{L}_1(Q^d)} &\le c\left\|\int_{2Q^d} \Delta_y^2\tilde{f}_\varepsilon(\cdot)\mathbf{J}_n(y)\,dy\right\|_{\mathbf{L}_1(Q^d)} \\ &\le c\int_{Q^d}\left|\int_{2Q^d} \Delta_y^2\tilde{f}_\varepsilon(x)\mathbf{J}_n(y)\,dy\right|dx \\ &\le c\int_{Q^d}\int_{2Q^d}\left|\Delta_y^2\tilde{f}_\varepsilon(x)\right|\mathbf{J}_n(y)\,dy\,dx \\ &= c\int_{Q^d}\int_{2Q^d} \Delta_y^2\tilde{f}_\varepsilon(x)\mathbf{J}_n(y)\,dy\,dx \\ &\le c\int_{2Q^d}\left(\int_{Q^d} \Delta_y^2\tilde{f}_\varepsilon(x)\,dx\right)\mathbf{J}_n(y)\,dy,\end{aligned} \tag{10}$$

since convexity of $\tilde{f}_\varepsilon$ on $3Q^d$ implies that $\Delta_y^2 \tilde{f}_\varepsilon(x) \geq 0$ for all $y \in 2Q^d$ and $x \in Q^d$. Now, recalling that $S(x) := x + S$ denotes the shift of the set S by the vector x and changing variables we have, for every $y \in 2Q^d$,

$$\begin{aligned}\int_{Q^d} \Delta_y^2 \tilde{f}_\varepsilon(x)\,dx &= \left(\int_{Q^d(-y)} + \int_{Q^d(y)} - 2\int_{Q^d}\right)\tilde{f}_\varepsilon(x)\,dx \\ &= \left(\int_{Q^d(-y)\setminus Q^d} - \int_{Q^d\setminus Q^d(-y)}\right. \\ &\qquad \left. + \int_{Q^d(y)\setminus Q^d} - \int_{Q^d\setminus Q^d(y)}\right)\tilde{f}_\varepsilon(x)\,dx \\ &\leq \left(\int_{Q^d(-y)\setminus Q^d} - \int_{Q^d\setminus Q^d(y)}\right)\tilde{f}_\varepsilon(x)\,dx \\ &\quad + \left(\int_{Q^d(y)\setminus Q^d} - \int_{Q^d\setminus Q^d(-y)}\right)\tilde{f}_\varepsilon(x)\,dx.\end{aligned}$$

Now, using the fact that

$$\left(Q^d(-y) \setminus Q^d\right)(y) = Q^d \setminus Q^d(y)$$

and

$$\left(Q^d \setminus Q^d(-y)\right)(y) = Q^d(y) \setminus Q^d$$

and changing variables again we have

$$\begin{aligned}\int_{Q^d} \Delta_y^2 \tilde{f}_\varepsilon(x)\,dx \leq \int_{Q^d\setminus Q^d(y)} &\left(\tilde{f}_\varepsilon(x-y) - \tilde{f}_\varepsilon(x)\right)\,dx \\ + \int_{Q^d(y)\setminus Q^d} &\left(\tilde{f}_\varepsilon(x) - \tilde{f}_\varepsilon(x-y)\right)\,dx.\end{aligned}$$

Hence, since $\tilde{f}_\varepsilon \in \widehat{\mathcal{L}}_\lambda(3Q^d)$, for every $y \in 2Q^d$, we have

$$\int_{Q^d} \Delta_y^2 \tilde{f}_\varepsilon(x)\,dx \leq \lambda|y| \left(\operatorname{meas}\left(Q^d \setminus Q^d(y)\right) + \operatorname{meas}\left(Q^d(y) \setminus Q^d\right)\right).$$

We now note that

$$Q^d(y) \setminus Q^d \subset \cup_{j=1}^d S_j,$$

where

$$S_j := \{x \in \mathbf{R}^d : -1 + y_i \le x_i \le 1 + y_i,\ 1 \le i \le d,\\ x_j \in (-\infty, -1) \cup (1, \infty)\}$$

and meas $(S_j) \le 2^{d-1}|y_j|$, $1 \le j \le d$. This implies that

$$\text{meas}\ \big(Q^d(y) \setminus Q^d\big) \le 2^{d-1} \sum_{j=1}^{d} |y_j|$$

and, similarly,

$$\text{meas}\ \big(Q^d \setminus Q^d(y)\big) \le 2^{d-1} \sum_{j=1}^{d} |y_j|.$$

Therefore, using Jensen's inequality we have

$$\int_{Q^d} \Delta_y^2 \tilde{f}_\varepsilon(x)\, dx \le \lambda 2^d |y| \sum_{j=1}^{d} |y_j| \le \lambda 2^d \sqrt{d} \sum_{j=1}^{d} y_j^2 = \lambda 2^d \sqrt{d} |y|^2,$$

and using this in (10) we have

$$\Big\| \mathcal{I}(\cdot; \tilde{f}_\varepsilon, \mathbf{J}_n) \Big\|_{\mathbf{L}_1(Q^d)} \le c \int_{2Q^d} |y|^2 \mathbf{J}_n(y)\, dy \le c n^{-2}$$

as needed.

It now follows from (7), (8) and (9) that

$$\|f(\cdot) - P_n(f)\|_{\mathbf{L}_p(\mathfrak{B})} \le c(d, \mathfrak{B}) n^{-1-1/p}.$$

Since degrees of convex polynomials $P_n(f)$ are $\le 16n$ and since the upper estimate in (2) is clearly true for $n = 1$ we conclude that

$$E\big(\widehat{\mathcal{L}}^d, \mathcal{P}_n^d\big)_{\mathbf{L}_p^d} \le E\big(\widehat{\mathcal{L}}^d, \widehat{\mathcal{P}}_n^d\big)_{\mathbf{L}_p^d} \le c(d, \mathfrak{B}) n^{-1-1/p},\\ n \ge 1, \quad 1 \le p \le \infty.$$

3. Proof of lower estimates in Theorem 1. For the proof of lower estimates we need the following result from [**6**, Lemma 10] (in the case $p = \infty$, this lemma follows from [**3**, E21 (page 414)]).

Lemma 4. *Let $n \in \mathbf{N}$, $1 \le p \le \infty$, $I := [-1/2, 1/2]$, and $I_n := [-1/(8n), 1/(8n)]$. Then, there exists a constant $c_1 = c_1(p)$ such that, for every polynomial P_n of degree $\le n$,*

$$\|P_n\|_{\mathbf{L}_p(I)} \le c_1 \|P_n\|_{\mathbf{L}_p(I \setminus I_n)}. \tag{11}$$

Recall that $\mathbf{B}^d$ denotes the Euclidean unit (closed) ball in $\mathbf{R}^d$, and that $3\mathbf{B}^d$ is the ball of radius 3. Without loss of generality we assume that $3\mathbf{B}^d \subset \mathfrak{B}$. Define

$$g_\alpha(t) := \alpha(t-2)_+ := \alpha \max\{0, t-2\},$$

where $t \in \mathbf{R}$ and $\alpha > 0$ is a parameter. We now show that, for any $1 \le p \le \infty$, $n \in \mathbf{N}$ and any polynomial P_n of degree $\le n$, the following inequality holds

$$\|g_\alpha - P_n\|_{\mathbf{L}_p[1,3]} > c_2 \alpha n^{-1-1/p}, \tag{12}$$

where $c_2 := 2^{-5-2/p}(p+1)^{-1/p}(c_1+2)^{-1}$ and c_1 is the constant from (11).

Indeed, suppose that $1 \le p \le \infty$ and $n \in \mathbf{N}$ are fixed, and that there exists a polynomial P_n which satisfies

$$\|g_\alpha - P_n\|_{\mathbf{L}_p[1,3]} \le c_2 \alpha n^{-1-1/p}. \tag{13}$$

Then, the function

$$g^*_\alpha(t) := g_\alpha(t + 2 + 1/(8n)) - 2g_\alpha(t+2) + g_\alpha(t + 2 - 1/(8n)),$$

and the polynomial

$$P^*_n(t) := P_n(t + 2 + 1/(8n)) - 2P_n(t+2) + P_n(t + 2 - 1/(8n)),$$

should satisfy the following inequality on $I := [-1/2, 1/2]$:

$$\|g^*_\alpha - P^*_n\|_{\mathbf{L}_p(I)} \le 4c_2 \alpha n^{-1-1/p}. \tag{14}$$

Since g^*_α is identically zero on $I \setminus I_n$ where $I_n := [-1/(8n), 1/(8n)]$, we conclude that

$$\|P^*_n\|_{\mathbf{L}_p(I \setminus I_n)} \le 4c_2 \alpha n^{-1-1/p}.$$

Lemma 4 now implies

$$\|P_n^*\|_{\mathbf{L}_p(I_n)} \le 4c_1c_2\alpha n^{-1-1/p}.$$

It is straightforward to show that

$$g_\alpha^*(t) = \alpha(t + 1/(8n)) - 2\alpha t_+ = \alpha(1/(8n) - |t|), \quad t \in I_n.$$

Hence

$$\|g_\alpha^*\|_{\mathbf{L}_p(I_n)} = \alpha 2^{-3-2/p}(p+1)^{-1/p} n^{-1-1/p},$$

and so

$$\begin{aligned}\|g_\alpha^* - P_n^*\|_{\mathbf{L}_p(I)} &\ge \|g_\alpha^*\|_{\mathbf{L}_p(I_n)} - \|P_n^*\|_{\mathbf{L}_p(I_n)} \\ &\ge \alpha(2^{-3-2/p}(p+1)^{-1/p} - 4c_1c_2)n^{-1-1/p} \\ &> 4c_2\alpha n^{-1-1/p},\end{aligned}$$

which contradicts inequality (14). Therefore, our assumption that (13) holds is not true, and so (12) is proved.

Now, the radial function $f_\alpha(x) := g_\alpha(|x|)$ is clearly convex on $\mathbf{R}^d$, and we choose α so that $f_\alpha \in \widehat{\mathcal{L}}(\mathfrak{B})$. Also, because of our assumption $3\mathbf{B}^d \subset \mathfrak{B}$, it is obvious that, for any polynomial $P_n \in \mathcal{P}_n(\mathbf{R}^d)$, we have

$$\|f_\alpha - P_n\|_{\mathbf{L}_p(\mathfrak{B})} \ge \|f_\alpha - P_n\|_{\mathbf{L}_p(3\mathbf{B}^d \setminus \mathbf{B}^d)}.$$

If $d = 1$, then together with (12), this immediately implies that

$$E\big(\widehat{\mathcal{L}}^1, \mathcal{P}_n^1\big)_{\mathbf{L}_p^1} \ge c(p, \mathfrak{B}) n^{-1-1/p}, \quad n \ge 1, \quad 1 \le p \le \infty.$$

If $d > 1$, then using hyperspherical coordinates $x = x(\rho, \varphi)$ in $\mathbf{R}^d$, where $\rho \in \mathbf{R}$, $\varphi = (\varphi_1, \dots, \varphi_{d-1}) \in \Phi := \prod_{i=1}^{d-2}[0,\pi] \times [0, 2\pi]$, and

$$x_i = \rho\left(\prod_{j=1}^{i-1} \sin\varphi_j\right)\cos\varphi_i, \quad 1 \le i \le d,$$

(with $\varphi_d := 0$), we have

$$\|f_\alpha - P_n\|_{\mathbf{L}_p(3\mathbf{B}^d\setminus\mathbf{B}^d)} = \left(\int_\Phi \int_1^3 \rho^{d-1}\mathfrak{J}_d(\varphi)|f_\alpha(\rho,\varphi) - P_n(\rho,\varphi)|^p \, d\rho \, d\varphi\right)^{1/p},$$

where $\mathfrak{J}_d(\varphi) := \prod_{i=1}^{d-2}(\sin\varphi_i)^{d-1-i}$ (and so $dV = \rho^{d-1}\mathfrak{J}_d(\varphi)\,d\rho\,d\varphi$ is the volume element in hyperspherical coordinates).

Now, (12) implies that, for every $\varphi \in \Phi$,

$$\left(\int_1^3 |f_\alpha(\rho,\varphi) - P_n(\rho,\varphi)|^p\,d\rho\right)^{1/p} \geq c_2\alpha n^{-1-1/p}.$$

Hence, taking into account that $1 \leq \rho \leq 3$, we have
(15)

$$\begin{aligned}\|f_\alpha - P_n\|_{\mathbf{L}_p(3\mathbf{B}^d\setminus\mathbf{B}^d)} &\geq \left(\int_\Phi \mathfrak{J}_d(\varphi)\int_1^3 |f_\alpha(\rho,\varphi) - P_n(\rho,\varphi)|^p\,d\rho\,d\varphi\right)^{1/p}\\ &\geq c_2\alpha n^{-1-1/p}\left(\int_\Phi \mathfrak{J}_d(\varphi)\,d\varphi\right)^{1/p}\\ &\geq c(d,p,\mathfrak{B})n^{-1-1/p}.\end{aligned}$$

Since polynomials P_n are arbitrary we conclude that

$$E\big(\widehat{\mathcal{L}}^d, \mathcal{P}_n^d\big)_{\mathbf{L}_p^d} \geq cn^{-1-1/p}, \quad n \geq 1,\ 1 \leq p \leq \infty.$$

The proof of lower estimates in Theorem 1 is now complete since

$$E\big(\widehat{\mathcal{L}}^d, \widehat{\mathcal{P}}_n^d\big)_{\mathbf{L}_p^d} \geq E\big(\widehat{\mathcal{L}}^d, \mathcal{P}_n^d\big)_{\mathbf{L}_p^d}.$$

4. Proof of Theorem 2. Recall that $\mathbf{B}^d$ denotes the Euclidean unit ball in $\mathbf{R}^d$, and that a function $f : \mathbf{B}^d \mapsto \mathbf{R}$ is called "radial" if $f(x) = g(|x|)$, $x \in \mathbf{B}^d$, where $g : [0,1] \mapsto \mathbf{R}$ is some univariate function.

It is well known (see, e.g., [**12**, page 164]) that the space $\mathcal{P}_n^d$ is contained in the manifold $\mathcal{R}_N^d$, where $N = \binom{n+d-1}{d-1} \asymp n^{d-1}$, $n \geq 1$, and so the estimate

$$E\big(\widehat{\mathcal{L}}^d, \widehat{\mathcal{R}}_n^d\big)_{\mathbf{L}_2^d} \leq c(d,\mathfrak{B})n^{-3/(2(d-1))}, \quad n \geq 1,$$

immediately follows from Theorem 1 (with $p = 2$).

Without loss of generality we now assume that $3\mathbf{B}^d \subset \mathfrak{B}$. It follows from (15) that, for some radial function f_α which is defined on $3\mathbf{B}^d$ and such that $f_\alpha \in \widehat{\mathcal{L}}^d = \widehat{\mathcal{L}}(\mathfrak{B})$, we have

$$\|f_\alpha - P_n\|_{\mathbf{L}_2(3\mathbf{B}^d)} \geq c(d, \mathfrak{B}) n^{-3/2},$$

for any polynomial $P_n \in \mathcal{P}_n(\mathbf{R}^d)$.

It was shown in [**6**, Theorem 4] (see also [**11**, Theorem 1] for the proof in the case $d = 2$) that, for any radial function $f \in \mathbf{L}_2^d$,

$$E\big(f, \mathcal{R}_{n^{d-1}}(\mathbf{B}^d)\big)_{\mathbf{L}_2(\mathbf{B}^d)} \geq c(d) E\big(f, \mathcal{P}_{mn}(\mathbf{B}^d)\big)_{\mathbf{L}_2(\mathbf{B}^d)}$$

for some $m \in \mathbf{N}$. This implies that

$$E\big(f_\alpha, \mathcal{R}_{n^{d-1}}(3\mathbf{B}^d)\big)_{\mathbf{L}_2(3\mathbf{B}^d)} \geq c(d, \mathfrak{B}) n^{-3/2},$$

and therefore,

$$E\big(\widehat{\mathcal{L}}^d, \mathcal{R}_n^d\big)_{\mathbf{L}_2^d} \geq c(d, \mathfrak{B}) n^{-3/(2(d-1))}, \quad n \geq 1.$$

The proof of lower estimates in Theorem 2 is now complete since

$$E\big(\widehat{\mathcal{L}}^d, \widehat{\mathcal{R}}_n^d\big)_{\mathbf{L}_2^d} \geq E\big(\widehat{\mathcal{L}}^d, \mathcal{R}_n^d\big)_{\mathbf{L}_2^d}.$$

REFERENCES

1. S. Artstein-Avidan, O. Friedland and V. Milman, *Geometric applications of Chernoff-type estimates and a zigzag approximation for balls*, Proc. Amer. Math. Soc. **134** (2006), 1735–1742 (electronic).

2. A.R. Barron, *Universal approximation bounds for superpositions of a sigmoidal function*, IEEE Trans. Inform. Theory **39** (1993), 930–945.

3. P. Borwein and T. Erdélyi, *Polynomials and polynomial inequalities*, Grad. Texts Math. **161**, Springer-Verlag, New York, 1995.

4. V.K. Dzjadyk, *Vvedenie v teoriyu ravnomernogo priblizheniya funktsii polinomami*, Izdat. "Nauka", Moscow, 1977 (in Russian).

5. V.N. Konovalov, *A Jackson-type approximation theorem for functions of several variables*, Ukrain. Mat. Zh. **33** (1981), 757–764, 860 (in Russian).

6. V.N. Konovalov, D. Leviatan and V.E. Maiorov, *Approximation by polynomials and ridge functions of the classes of s-monotone radial functions*, J. Approx. Theory **152** (2008), 20–51.

7. K.A. Kopotun, *On moduli of smoothness of k-monotone functions and applications*, Math. Proc. Camb. Phil. Soc. **146** (2009), 213–223.

8. D. Leviatan, *Pointwise estimates for convex polynomial approximation*, Proc. Amer. Math. Soc. **98** (1986), 471–474.

9. G.G. Lorentz, M. v. Golitschek, and Y. Makovoz, *Constructive approximation*, Grundl. Math. Wissen. (Fundamental Principles of Mathematical Sciences) **304** (1996), Springer-Verlag, Springer-Verlag, xii+649.

10. X.G. Lu, *Convex approximation by multivariate polynomials*, Math. Numer. Sinica **12** (1990), 186–193 (in Chinese, with English summary).

11. K.I. Oskolkov, *Ridge approximation, Fourier-Chebyshev analysis, and optimal quadrature formulas*, Tr. Mat. Inst. Steklova **219** (1997), Teor. Priblizh. Garmon. Anal., 269–285 (in Russian); Proc. Steklov Inst. Math. **4** (1997), 265–280 (English transl.).

12. A. Pinkus, *Approximation theory of the MLP model in neural networks*, Acta Numerica **8**, Cambridge Univ. Press, Cambridge, 1999.

13. A.W. Roberts and D.E. Varberg, *Convex functions*, Pure Appl. Math. **57** Academic Press, New York, 1973.

14. A.S. Shvedov, *Co-convex approximation of functions of several variables by polynomials*, Mat. Sb. (N.S.) **115** (1981), 577–589, 655 (in Russian).

15. X.M. Yu, *Convex polynomial approximation in L_p space*, Approx. Theory Appl. **3** (1987), 72–83.

Institute of Mathematics, National Academy of Sciences of Ukraine, Kyiv 01601, Ukraine
Email address: vikono@imath.kiev.ua

Department of Mathematics, University of Manitoba, Winnipeg, Manitoba, R3T 2N2, Canada
Email address: kopotunk@cc.umanitoba.ca

Technion, 32000 Haifa, Israel
Email address: maiorov@tx.technion.ac.il