Available online at www.sciencedirect.com
ScienceDirect
Journal of Approximation Theory 159 (2009) 97–108
Journal of Approximation Theory
www.elsevier.com/locate/jat

Approximation of Sobolev classes by polynomials and ridge functions

V.N. Konovalov[a,1,†], D. Leviatan[b,*], V.E. Maiorov[c]

[a] Institute of Mathematics, NAS of Ukraine, 01601 Kyiv, Ukraine
[b] Raymond and Beverly Sackler School of Mathematical Sciences, Tel Aviv University, 69978 Tel Aviv, Israel
[c] Technion I.I.T., 32000 Haifa, Israel

Received 2 February 2008; received in revised form 16 October 2008; accepted 18 October 2008
Available online 19 November 2008

Communicated by C.K. Chui and H.N. Mhaskar

Dedicated to the memory of G.G. Lorentz

Abstract

Let $W_p^r(\mathbb{B}^d)$ be the usual Sobolev class of functions on the unit ball $\mathbb{B}^d$ in $\mathbb{R}^d$, and $W_p^{\circ,r}(\mathbb{B}^d)$ be the subclass of all radial functions in $W_p^r(\mathbb{B}^d)$. We show that for the classes $W_p^{\circ,r}(\mathbb{B}^d)$ and $W_p^r(\mathbb{B}^d)$, the orders of best approximation by polynomials in $L_q(\mathbb{B}^d)$ coincide. We also obtain exact orders of best approximation in $L_2(\mathbb{B}^d)$ of the classes $W_p^{\circ,r}(\mathbb{B}^d)$ by ridge functions and, as an immediate consequence, we obtain the same orders in $L_2(\mathbb{B}^d)$ for the usual Sobolev classes $W_p^r(\mathbb{B}^d)$.

Keywords: Sobolev classes; Radial functions; Polynomials; Ridge functions

1. Introduction and the main results

Let $d \in \mathbb{N}$ and let $\mathbb{B}^d$ be the open unit ball in the space $\mathbb{R}^d$. A function $x : \mathbb{B}^d \mapsto \mathbb{R}$ is called a radial function if $x(t) = y(|t|)$, $t = (t_1, \ldots, t_d) \in \mathbb{B}^d$, where $|t| := (t_1^2 + \cdots + t_d^2)^{1/2}$. For

* Corresponding author.
E-mail addresses: vikono@imath.kiev.ua (V.N. Konovalov), leviatan@post.tau.ac.il (D. Leviatan), maiorov@tx.technion.ac.il (V.E. Maiorov).
[1] Part of this work was done while the first author visited Tel Aviv University and Technion I.I.T. Haifa in June 2006.
† Victor Konovalov died on Nov. 1st, 2008.

0021-9045/$ - see front matter
doi:10.1016/j.jat.2008.10.009

$r \in \mathbb{N}$ and $1 \le p \le \infty$, we denote by $W_p^r(\mathbb{B}^d)$ the usual Sobolev class of r-times differentiable functions $x : \mathbb{B}^d \mapsto \mathbb{R}$ such that

$$\sum_{|k|=r} \|D^k x\|_{L_p(\mathbb{B}^d)} \le 1,$$

where $k = (k_1, \dots, k_d) \in \mathbb{Z}_+^d$ and $D^k x$ is the partial derivative, in the Sobolev sense, of order $|k| := k_1 + \dots + k_d$. By $W_p^{\circ,r}(\mathbb{B}^d)$ we denote the subclass of all radial functions in $W_p^r(\mathbb{B}^d)$.

The main purpose of this paper is to estimate the orders of best approximation by polynomials and ridge functions of the classes $W_p^{\circ,r}(\mathbb{B}^d)$ in the spaces $L_q(\mathbb{B}^d)$.

Let $\mathcal{P}_n(\mathbb{B}^d)$, be the space of polynomials

$$P_n(t) := \sum_{|k| \le n} a_k t^k, \quad t \in \mathbb{B}^d,$$

where $k = (k_1, \dots, k_d) \in \mathbb{Z}_+^d$, $|k| := k_1 + \dots + k_d$, $a_k \in \mathbb{R}$, and $t^k := t_1^{k_1} \dots t_d^{k_d}$. For $r - d(1/p - 1/q) > 0$, which guarantees the compact embedding of the classes $W_p^{\circ,r}(\mathbb{B}^d)$ and $W_p^r(\mathbb{B}^d)$ into the space $L_q(\mathbb{B}^d)$, we denote

$$E\left(W_p^{\circ,r}\left(\mathbb{B}^d\right), \mathcal{P}_n\left(\mathbb{B}^d\right)\right)_{L_q(\mathbb{B}^d)} := \sup_{x \in W_p^{\circ,r}(\mathbb{B}^d)} \inf_{P_n \in \mathcal{P}_n(\mathbb{B}^d)} \|x - P_n\|_{L_q(\mathbb{B}^d)},$$

and

$$E\left(W_p^{r}\left(\mathbb{B}^d\right), \mathcal{P}_n\left(\mathbb{B}^d\right)\right)_{L_q(\mathbb{B}^d)} := \sup_{x \in W_p^{r}(\mathbb{B}^d)} \inf_{P_n \in \mathcal{P}_n(\mathbb{B}^d)} \|x - P_n\|_{L_q(\mathbb{B}^d)}.$$

For a subset K of $\mathbb{R}$, let $M(K)$ denote the space of all real-valued functions on K. For $d > 1$, we denote by $r(a \cdot t)$, $a \in \mathbb{R}^d$, a ridge function defined as a function of $t \in \mathbb{R}^d$, where $a \cdot t$ is the usual inner product and $r \in M(\mathbb{R})$.

Let $\mathbb{S}^{d-1} := \partial\mathbb{B}^d$ be the unit sphere in $\mathbb{R}^d$ and $I := (-1, 1)$. We denote by $\mathcal{R}_n$, the (nonlinear) manifold consisting of all possible linear combinations of n ridge functions

$$R_n(t) := \sum_{k=1}^{n} r_k(a_k \cdot t), \quad a_k \in \mathbb{S}^{d-1}, r_k \in M(I).$$

If $r_k \in L_q(I)$, $1 \le k \le n$, then we write $\mathcal{R}_{n,q}(\mathbb{B}^d) := \mathcal{R}_n(\mathbb{B}^d)$, and under the condition $r - d(1/p - 1/q) > 0$, we denote by

$$E\left(W_p^{\circ,r}\left(\mathbb{B}^d\right), \mathcal{R}_{n,q}\left(\mathbb{B}^d\right)\right)_{L_q(\mathbb{B}^d)} := \sup_{x \in W_p^{\circ,r}(\mathbb{B}^d)} \inf_{R_n \in \mathcal{R}_{n,q}(\mathbb{B}^d)} \|x - R_n\|_{L_q(\mathbb{B}^d)},$$

and

$$E\left(W_p^{r}\left(\mathbb{B}^d\right), \mathcal{R}_{n,q}\left(\mathbb{B}^d\right)\right)_{L_q(\mathbb{B}^d)} := \sup_{x \in W_p^{r}(\mathbb{B}^d)} \inf_{R_n \in \mathcal{R}_{n,q}(\mathbb{B}^d)} \|x - R_n\|_{L_q(\mathbb{B}^d)},$$

the deviations, in $L_q(\mathbb{B}^d)$, of $W_p^{\circ,r}(\mathbb{B}^d)$ and $W_p^r(\mathbb{B}^d)$, respectively, from $\mathcal{R}_{n,q}(\mathbb{B}^d)$.

Finally, let

$$E\left(x, \mathcal{P}_n\left(\mathbb{B}^d\right)\right)_{L_q(\mathbb{B}^d)} := \inf_{P_n \in \mathcal{P}_n(\mathbb{B}^d)} \|x - P_n\|_{L_q(\mathbb{B}^d)},$$

and

$$E\left(x, \mathcal{R}_{n,q}\left(\mathbb{B}^d\right)\right)_{L_q(\mathbb{B}^d)} := \inf_{R_n \in \mathcal{R}_{n,q}(\mathbb{B}^d)} \|x - R_n\|_{L_q(\mathbb{B}^d)}$$

be the distances of $x \in L_q(\mathbb{B}^d)$ from the space $\mathcal{P}_n(\mathbb{B}^d)$, and from the manifold $\mathcal{R}_{n,q}(\mathbb{B}^d)$, respectively.

An important special case of elements in the manifold of ridge functions is neural-network functions of the form

$$H_n(t) = \sum_{k=1}^{n} c_k h(a_k \cdot t + b_k), \quad a_k \in \mathbb{R}^d,\ c_k, b_k \in \mathbb{R},$$

where $h : \mathbb{R} \to \mathbb{R}$ is a given activation function and n is the number of hidden units.

Approximation by ridge functions in general and by the neural networks in particular, has been extensively investigated in recent years (see, e.g., [17,12–14]). Ridge functions frequently appear in mathematics and its applications. Various aspects of the approximation by ridge functions are natural components in many theoretical and applied problems such as the Radon transform, tomography [3,18], mathematical physics equations, and geometry [5].

As an illustration we take the question of quadrature. Let f have an integral representation of the form $f(t) = \int h(a \cdot t + b)\, \mathrm{d}\mu(a, b)$, with an activation function $h : \mathbb{R} \to \mathbb{R}$ and a unit measure $\mu(a, b)$ defined on $\mathbb{R}^d \times \mathbb{R}$. The question of finding a quadrature formula for the integral is closely connected (see [11,6]) to the question of approximation of the function f by a linear combination of n neural-network units $h(a_k \cdot t + b_k)$, $k = 1, \ldots, n$. A similar connection exists between the approximation of the integral $f(t) = \int_{\mathbb{S}^{d-1}} g(a, a \cdot t)\, \mathrm{d}a$, where $g(a, a \cdot t)$ is the ridge function related to the Radon transform of the function f, and the approximation of f by linear combination of ridge functions $g(a_k, a_k \cdot t)$.

A series of results about estimating the degree of approximation of special functions by the ridge-manifold $\mathcal{R}_n(B^d)$ in the two-dimensional case were established by Oskolkov [15,16]. In particular, Oskolkov showed [15] that in the case $d = 2$ orders of approximation of every radial function by the ridge-manifold $\mathcal{R}_n(B^d)$ and by the space $\mathcal{P}_n$ of algebraic polynomials of degree n coincide, and also (see [16]) orders of approximation of every harmonic function by $\mathcal{R}_n(B^d)$ are (roughly) twice more than by the space $\mathcal{P}_n$. The result of Oskolkov [15] for radial functions was extended to the general case, $d \geq 2$, in the work of the authors [7].

Results about approximation of the Sobolev classes $W_2^r(\mathbb{B}^d)$ functions (worst case setting) by the ridge-manifold $\mathcal{R}_n(B^d)$ were obtained by Maiorov [8], namely,

$$E\left(W_2^r\left(\mathbb{B}^d\right), \mathcal{R}_{n,2}\left(\mathbb{B}^d\right)\right)_{L_2(\mathbb{B}^d)} \asymp n^{-r/(d-1)},$$

where for sequences $\{a_n\}_{n\geq 1}$ and $\{b_n\}_{n\geq 1}$, of positive numbers we write $a_n \asymp b_n$, $n \geq 1$, if there exist constants $0 < c_1 \leq c_2$ such that $c_1 \leq a_n/b_n \leq c_2$, for all $n \geq 1$. In addition, Maiorov, Meir and Ratsaby [9] show that the set of functions for which the estimate $n^{-r/(d-1)}$ holds is of large measure. In other words, this is not simply a worst case setting. Other interesting results may be found in papers [20,10,4].

Throughout the paper $c := c(\alpha, \beta, \ldots, \gamma)$ denotes a constant which depends on the given parameters, but may differ from one occasion to another even if it appears in the same line. Also, as usual, $(a)_+ := \max\{a, 0\}$, $a \in \mathbb{R}$.

We are ready to state the main results.

Theorem 1. *Let $d, r \in \mathbb{N}$, and $1 \le p, q \le \infty$ be such that $r - d(1/p - 1/q) > 0$. Then*

$$E\left(W_p^{\circ,r}(\mathbb{B}^d), \mathcal{P}_n(\mathbb{B}^d)\right)_{L_q(\mathbb{B}^d)} \asymp E\left(W_p^r(\mathbb{B}^d), \mathcal{P}_n(\mathbb{B}^d)\right)_{L_q(\mathbb{B}^d)} \asymp n^{-r+d(1/p-1/q)_+},$$

and

Theorem 2. *Let $d, r \in \mathbb{N}$, $d > 1$, and $1 \le p \le \infty$ be such that $r - d(1/p - 1/2) > 0$. Then*

$$E\left(W_p^{\circ,r}(\mathbb{B}^d), \mathcal{R}_{n,2}(\mathbb{B}^d)\right)_{L_2(\mathbb{B}^d)} \asymp E\left(W_p^r(\mathbb{B}^d), \mathcal{R}_{n,2}(\mathbb{B}^d)\right)_{L_2(\mathbb{B}^d)} \asymp n^{-\frac{r-d(1/p-1/2)_+}{d-1}}.$$

It is well known that for $d > 1$, the space $\mathcal{P}_n(\mathbb{B}^d)$ may be embedded in the manifold $\mathcal{R}_{cn^{d-1}}(\mathbb{B}^d)$ where $c = c(d)$. More precisely (see, e.g., [17], Page 164), the space $\mathcal{P}_n(\mathbb{B}^d)$ is contained in the manifold $\mathcal{R}_N(\mathbb{B}^d)$, where $N = \binom{n+d-1}{d-1}$. Thus, an immediate consequence of Theorems 1 and 2 is:

Corollary 3. *Let $d, r \in \mathbb{N}$, $d > 1$, and $2 < q \le p \le \infty$. Then*

$$E\left(W_p^{\circ,r}(\mathbb{B}^d), \mathcal{R}_{n,q}(\mathbb{B}^d)\right)_{L_q(\mathbb{B}^d)} \asymp E\left(W_p^r(\mathbb{B}^d), \mathcal{R}_{n,q}(\mathbb{B}^d)\right)_{L_q(\mathbb{B}^d)} \asymp n^{-r/(d-1)}.$$

2. Auxiliary lemmas

The following Remez-type inequality is well known for $d = 1$ and $q = \infty$ (see, e.g., [1], p. 414, E21). For $d = 1$ and $1 \le q < \infty$ it was recently proved in [7]. We need to extend it to the case $d > 1$ and $1 \le q \le \infty$.

Lemma 4. *Let $d, n \in \mathbb{N}$, $1 \le q \le \infty$, and*

$$\mathbb{B}_n^d := \{t : |t| \le 1/(4n), t \in \mathbb{R}^d\}.$$

Then there exists $c^* = c^*(d, q) > 0$ such that for any polynomial $P_n \in \mathcal{P}_n(\mathbb{B}^d)$,

$$\|P_n\|_{L_q(\mathbb{B}^d)} \le c^* \|P_n\|_{L_q(\mathbb{B}^d \setminus \mathbb{B}_n^d)}.$$

Proof. Let $I := (-1, 1)$ and $I_n := (-1/(4n), 1/(4n))$. As mentioned above, for $d = 1$ the inequality

$$\|P_n\|_{L_q(I_n)} \le \bar{c} \|P_n\|_{L_q(I \setminus I_n)}, \tag{2.1}$$

where $\bar{c} = \bar{c}(q) > 0$, is known (see [1,7]). To prove such an inequality for $d > 1$ we use the spherical coordinates $(\rho, \varphi) := (\rho, \varphi_1, \dots, \varphi_{d-1})$, defined by $t_1 = \rho \cos\varphi_1$, $t_2 = \rho \sin\varphi_1 \cos\varphi_2, \dots, t_{d-2} = \rho \sin\varphi_1 \dots \sin\varphi_{d-3} \cos\varphi_{d-2}$, $t_{d-1} = \rho \sin\varphi_1 \dots \sin\varphi_{d-2} \cos\varphi_{d-1}$, $t_d = \rho \sin\varphi_1 \dots \sin\varphi_{d-2} \sin\varphi_{d-1}$, where $0 \le \rho < 1$; $0 \le \varphi_i \le \pi$, $i = 1, \dots, d-2$; $0 \le \varphi_{d-1} < 2\pi$. In these coordinates, the volume element $\mathrm{d}t$ of $\mathbb{B}^d$ becomes

$$\mathrm{d}\xi = J_d(\rho, \varphi)\mathrm{d}\rho\mathrm{d}\varphi,$$

where $J_d(\rho, \varphi)$ is the Jacobian given by

$$J_d(\rho, \varphi) := \rho^{d-1}(\sin\varphi_1)^{d-2}(\sin\varphi_2)^{d-3} \dots \sin\varphi_{d-2},$$

and for any $x \in L_1(\mathbb{B}^d)$, we have

$$\int_{\mathbb{B}^d} x(t)\mathrm{d}t = \int_0^1 \int_0^\pi \cdots \int_0^\pi \int_0^{2\pi} x(t(\rho,\varphi)) J_d(\rho,\varphi)\mathrm{d}\rho\mathrm{d}\varphi_1 \ldots \mathrm{d}\varphi_{d-2}\mathrm{d}\varphi_{d-1}.$$

Denote $\tilde{\varphi} := (\varphi_1, \ldots, \varphi_{d-2}, \varphi_{d-1})$, and let

$$J_{d-1}(\tilde{\varphi}) := (\sin\varphi_1)^{d-2}(\sin\varphi_2)^{d-3} \ldots \sin\varphi_{d-2}.$$

Then

$$J_d(\rho,\varphi) = \rho^{d-1} J_{d-1}(\tilde{\varphi}),$$

and it follows that

$$\begin{aligned}
&\int_0^1 \int_0^\pi \cdots \int_0^\pi \int_0^{2\pi} x(t(\rho,\varphi)) J_d(\rho,\varphi)\mathrm{d}\rho\mathrm{d}\varphi_1 \ldots \mathrm{d}\varphi_{d-2}\mathrm{d}\varphi_{d-1} \\
&\quad = \int_0^1 \int_{[0,\pi]^{d-2}} \int_0^{2\pi} x(t(\rho,\tilde{\varphi},\varphi_{d-1}))\rho^{d-1} J_{d-1}(\tilde{\varphi})\mathrm{d}\rho\mathrm{d}\tilde{\varphi}\mathrm{d}\varphi_{d-1} \\
&\quad = \int_{[0,\pi]^{d-2}} J_{d-1}(\tilde{\varphi}) \left(\int_0^{2\pi} \int_0^1 \rho^{d-1} x(t(\rho,\tilde{\varphi},\varphi_{d-1}))\mathrm{d}\rho\mathrm{d}\varphi_{d-1} \right) \mathrm{d}\tilde{\varphi}.
\end{aligned}$$

Note that

$$\begin{aligned}
&\int_0^{2\pi} \int_0^1 \rho^{d-1} x(t(\rho,\tilde{\varphi},\varphi_{d-1}))\mathrm{d}\rho\mathrm{d}\varphi_{d-1} \\
&\quad = \int_0^\pi \left(\int_0^1 \rho^{d-1} x(t(\rho,\tilde{\varphi},\varphi_{d-1}))\mathrm{d}\rho + \int_0^1 \rho^{d-1} x(t(\rho,\tilde{\varphi},\varphi_{d-1}+\pi))\mathrm{d}\rho \right) \mathrm{d}\varphi_{d-1}.
\end{aligned}$$

Thus we let $\varphi = (\tilde{\varphi}, \varphi_{d-1})$ be such that $\tilde{\varphi} \in [0,\pi]^{d-2}$ and $\varphi_{d-1} \in [0,\pi]$, and denote by I_φ the chord connecting the two points $(1,\tilde{\varphi},\varphi_{d-1})$ and $(1,\tilde{\varphi},\varphi_{d-1}+\pi)$ on the unit sphere $\mathbb{S}^{d-1}$. Then, the interval I_φ can be represented as the set of all points $t = \tau e(\varphi)$, $\tau \in (-1,1) =: I$, where $e(\varphi) := (1,\tilde{\varphi},\varphi_{d-1}) \in \mathbb{R}^d$.

Let $x(\tau; I_\varphi)$, $\tau \in I$, be the univariate function that is the restriction of the function $x(\cdot)$ to the interval I_φ. Then, we have

$$\int_0^1 \rho^{d-1} x(t(\rho,\tilde{\varphi},\varphi_{d-1}))\mathrm{d}\rho + \int_0^1 \rho^{d-1} x(t(\rho,\tilde{\varphi},\varphi_{d-1}+\pi))\mathrm{d}\rho = \int_I |\tau|^{d-1} x(\tau; I_\varphi)\mathrm{d}\tau.$$

Hence, we get

$$\int_{\mathbb{B}^d} x(t)\mathrm{d}t = \int_{[0,\pi]^{d-2}} J_{d-1}(\tilde{\varphi}) \left(\int_{[0,\pi]} \left(\int_I |\tau|^{d-1} x(\tau; I_\varphi)\mathrm{d}\tau \right) \mathrm{d}\varphi_{d-1} \right) \mathrm{d}\tilde{\varphi}. \tag{2.2}$$

For $1 \le q < \infty$, we replace x in (2.2) with $|P_n|^q$, where P_n is any fixed polynomial $P_n \in \mathcal{P}_n(\mathbb{B}^d)$, and we obtain

$$\int_{\mathbb{B}^d} |P_n(t)|^q \mathrm{d}t = \int_{[0,\pi]^{d-2}} J_{d-1}(\tilde{\varphi}) \left(\int_{[0,\pi]} \left(\int_I |\tau|^{d-1} |P_n(\tau; I_\varphi)|^q \mathrm{d}\tau \right) \mathrm{d}\varphi_{d-1} \right) \mathrm{d}\tilde{\varphi}. \tag{2.3}$$

Similarly,

$$\int_{\mathbb{B}_n^d} |P_n(t)|^q \mathrm{d}t = \int_{[0,\pi]^{d-2}} J_{d-1}(\tilde{\varphi}) \left(\int_{[0,\pi]} \left(\int_{I_n} |\tau|^{d-1} |P_n(\tau; I_\varphi)|^q \mathrm{d}\tau \right) \mathrm{d}\varphi_{d-1} \right) \mathrm{d}\tilde{\varphi}. \tag{2.4}$$

Now,

$$\int_{I_n} |\tau|^{d-1} |P_n(\tau; I_\varphi)|^q \mathrm{d}\tau \le (4n)^{-d+1} \int_{I_n} |P_n(\tau; I_\varphi)|^q \mathrm{d}\tau$$
$$= (4n)^{-d+1} \| P_n(\cdot; I_\varphi)\|^q_{L_q(I_n)}, \tag{2.5}$$

and

$$\int_{I\setminus I_n} |\tau|^{d-1} |P_n(\tau; I_\varphi)|^q \mathrm{d}\tau \ge (4n)^{-d+1} \int_{I\setminus I_n} |P_n(\tau; I_\varphi)|^q \mathrm{d}\tau$$
$$= (4n)^{-d+1} \| P_n(\cdot; I_\varphi)\|^q_{L_q(I\setminus I_n)}. \tag{2.6}$$

By virtue of (2.1), we have

$$\| P_n(\cdot; I_\varphi)\|^q_{L_q(I_n)} \le \bar{c}^q \| P_n(\cdot; I_\varphi)\|^q_{L_q(I\setminus I_n)}.$$

Hence, by (2.5) and (2.6),

$$\int_I |\tau|^{d-1} |P_n(\tau; I_\varphi)|^q \mathrm{d}\tau \le (\bar{c}^q + 1) \int_{I\setminus I_n} |\tau|^{d-1} |P_n(\tau; I_\varphi)|^q \mathrm{d}\tau,$$

and substituting in (2.3), we obtain

$$\| P_n \|_{L_q(B^d)} \le c^* \| P_n \|_{L_q(B^d \setminus B^d_n)}, \quad 1 \le q < \infty,$$

where $c^* := c^*(d, q) > 0$. For $q = \infty$ the proof is similar. This completes the proof of Lemma 4. □

The next result generalizes for $d > 2$, the corresponding result, for $d = 2$, by Oskolkov ([15], Theorem 1). For the proof see [7].

Lemma 5. *Let $d \in \mathbb{N}$ and $d > 1$. There exist $\bar{c} = \bar{c}(d) > 0$, and integers $\hat{c} = \hat{c}(d)$ and $\check{c} = \check{c}(d)$, such that for any radial function $x \in L_2(\mathbb{B}^d)$,*

$$\bar{c} E\left(x, \mathcal{P}_{\hat{c}n}\left(\mathbb{B}^d\right)\right)_{L_2(\mathbb{B}^d)} \le E\left(x, \mathcal{R}_{n^{d-1},2}\left(\mathbb{B}^d\right)\right)_{L_2(\mathbb{B}^d)} \le E\left(x, \mathcal{P}_{\check{c}n}\left(\mathbb{B}^d\right)\right)_{L_2(\mathbb{B}^d)}.$$

3. Proofs

Proof of Theorem 1. The upper bound

$$E\left(W^r_p(\mathbb{B}^d), \mathcal{P}_n(\mathbb{B}^d)\right)_{L_q(\mathbb{B}^d)} \le c^* n^{-r+d(1/p-1/q)_+}$$

is essentially known as this estimate may be easily obtained from well-known extension theorems, the Jackson-type theorem for periodic functions and Nikol'skii's inequalities comparing various norms of trigonometric polynomials. However, we could find no reference which gives it in the above form, therefore we sketch a proof.

Let $W^r_p(\mathbb{T}^d)$ be the Sobolev class of 2π-periodic functions on the d-dimensional torus $\mathbb{T}^d$, and let $\mathcal{T}_n(\mathbb{T}^d)$ be the space of trigonometric polynomials T_n of order $\le n$. If $1 \le p = q \le \infty$ then the estimate

$$E\left(W^r_p(\mathbb{T}^d), \mathcal{T}_n(\mathbb{T}^d)\right)_{L_p(\mathbb{T}^d)} \le \acute{c} n^{-r}, \quad 1 \le p \le \infty, \tag{3.1}$$

is an immediate consequence of [21], Thm 5.3.1. This in turn immediately implies applying Hölder's inequality

$$E\left(W_p^r(\mathbb{T}^d), \mathcal{T}_n(\mathbb{T}^d)\right)_{L_q(\mathbb{T}^d)} \le \acute{c}n^{-r}, \quad 1 \le q < p \le \infty.$$

From (3.1) and Nikol'skii's inequalities (see, e.g., [21], 4.9.4 (18))

$$\|T_n\|_{L_q(\mathbb{T}^d)} \le \grave{c}n^{d(1/p-1/q)}\|T_n\|_{L_p(\mathbb{T}^d)}, \quad 1 \le p < q \le \infty,$$

it is readily seen that under the condition $r - d(1/p - 1/q) > 0$ we have

$$E\left(W_p^r(\mathbb{T}^d), \mathcal{T}_n(\mathbb{T}^d)\right)_{L_q(\mathbb{T}^d)} \le cn^{-r+d(1/p-1/q)}, \quad 1 \le p < q \le \infty.$$

Fix $0 < \varepsilon < 1$, and a function $v \in C_0^\infty(\mathbb{R}^d)$, supported on $\mathbb{B}_\varepsilon^d := \{t : |t| \le \varepsilon\}$, such that

$$\int_{\mathbb{R}^d} v(t)\mathrm{d}t = 1.$$

Set

$$P_{r-1}(t; x) := \sum_{|k|<r} \frac{1}{k!} \int_{\mathbb{B}_\varepsilon^d} x(\tau) D_\tau^k \left((t-\tau)^k v(\tau)\right) \mathrm{d}\tau,$$

where D_τ^k is the partial derivative with respect to τ of order $|k|$. Then any function $x \in W_p^r(\mathbb{B}^d)$ may be represented (see [2]) a.e. in $t \in \mathbb{B}^d$ by

$$x(t) = P_{r-1}(t; x) + r \sum_{|k|=r} \frac{1}{k!} \int_{V_{t,\varepsilon}^d} \frac{(t-\tau)^k}{|t-\tau|^d} w(t, \tau) D^k x(\tau)\mathrm{d}\tau,$$

where $V_{t,\varepsilon}^d$ is the convex hull of the set $\{t, \mathbb{B}_\varepsilon^d\}$, and

$$w(t, \tau) := \int_{|t-\tau|}^\infty v\left(t + \theta(\tau - t)/|\tau - t|\right) \theta^{d-1}\mathrm{d}\theta. \tag{3.2}$$

Evidently, $P_{r-1}(\cdot; x)$ is a polynomial of total degree $\le r - 1$.

Let

$$x_r(t) := x(t) - P_{r-1}(t; x), \quad t \in \mathbb{B}^d.$$

Then, it follows easily by (3.2) that

$$\sum_{|k|\le r} \|D^k x_r\|_{L_p(\mathbb{B}^d)} \le c_\circ \sum_{|k|=r} \|D^k x\|_{L_p(\mathbb{B}^d)},$$

where $c_\circ = c_\circ(d, r, p, \varepsilon, v)$.

The function x_r can be extended to the cube $\mathbb{Q}^d := \{t : |t_i| \le 2, i = 1, \ldots, d\}$ preserving the smoothness class (see, e.g. [19], Thm 6.3.5). More precisely, there exists a function $\bar{x}_r$ on $\mathbb{Q}^d$ such that $\bar{x}_r(t) = x_r(t)$ for $t \in \mathbb{B}^d$, and

$$\sum_{|k|\le r} \|D^k \bar{x}_r\|_{L_p(\mathbb{Q}_\pi^d)} \le \bar{c} \sum_{|k|\le r} \|D^k x_r\|_{L_p(\mathbb{B}^d)},$$

where $\bar{c} = \bar{c}(d, r, p, \varepsilon, v)$.

Fix $\omega \in C_0^\infty(\mathbb{R}^d)$ such that $\omega(t) \equiv 1$ for $t \in \mathbb{B}^d$, and its support is contained in the interior of $\mathbb{Q}^d$, and write $y_r(t) := \bar{x}_r(t)\omega(t)$, $t \in \mathbb{Q}^d$. Evidently, $y_r(t) \equiv x_r(t)$ for $t \in \mathbb{B}^d$. Moreover, it is easy to check that

$$\sum_{|k|\le r} \|D^k y_r\|_{L_p(\mathbb{Q}^d)} \le \hat{c} \sum_{|k|=r} \|D^k x\|_{L_p(\mathbb{B}^d)},$$

where $\hat{c} = \hat{c}(d, r, p, \varepsilon, v)$.

The 2π-periodic function $\tilde{y}_r(\tau_1, \ldots, \tau_d) := y_r(2\cos\tau_1, \ldots, 2\cos\tau_d)$ on the torus $\mathbb{T}^d$, is even and satisfies

$$\sum_{|k|\le r} \|D^k \tilde{y}_r\|_{L_p(\mathbb{T}^d)} \le \tilde{c} \sum_{|k|=r} \|D^k x\|_{L_p(\mathbb{B}^d)},$$

where $\tilde{c} = \tilde{c}(d, r, p, \varepsilon, v)$. Hence, there exists an even trigonometric polynomial $T_n(\cdot; \tilde{y}_r) \in \mathcal{T}_n(\mathbb{T}^d)$, such that

$$\|\tilde{y}_r(\cdot) - T_n(\cdot; \tilde{y}_r)\|_{L_q(\mathbb{T}^d)} \le \check{c} n^{-r+d(1/p-1/q)}.$$

Setting $P_n(t_1, \ldots, t_d; x_r) := T_n(\arccos(t_1/2), \ldots, \arccos(t_d/2); \tilde{y}_r)$, we obtain an (algebraic) polynomial that satisfies

$$\|x_r(\cdot) - P_n(\cdot; x_r)\|_{L_q(\mathbb{B}^d)} \le c^* n^{-r+d(1/p-1/q)}.$$

Clearly, the polynomial $P_n(\cdot; x) := P_{r-1}(\cdot; x) + P_n(\cdot; x_r)$ satisfies

$$\|x(\cdot) - P_n(\cdot; x)\|_{L_q(\mathbb{B}^d)} \le c^* n^{-r+d(1/p-1/q)},$$

and we have the required upper bound in the case $1 \le p < q \le \infty$.

Since $W_p^{\circ,r}(\mathbb{B}^d) \subset W_p^r(\mathbb{B}^d)$, it remains to prove the lower bound, namely,

$$E\left(W_p^{\circ,r}(\mathbb{B}^d), \mathcal{P}_n(\mathbb{B}^d)\right)_{L_q(\mathbb{B}^d)} \ge c_* n^{-r+d(1/p-1/q)_+}. \tag{3.3}$$

Let $r, n \in \mathbb{N}$ be fixed, and let $1 \le p < q \le \infty$ be such that $r - d(1/p-1/q) > 0$. Let $v \in C_0^\infty(\mathbb{R})$ be a nonnegative even function with $\operatorname{supp} v = [-1, 1]$ and $v(t) = 1$ for $t \in [-1/2, 1/2]$. Given α and $\beta > 0$, the function $\alpha v(\beta|t|)$, $t \in \mathbb{R}^d$, is obviously radial and belongs to the space $C_0^\infty(\mathbb{R}^d)$. Take $u_n(t) := v(4n|t|)$ where $t \in \mathbb{B}^d$ and denote

$$\tilde{\mathbb{B}}_n^d := \{t : |t| \le 1/(8n), t \in \mathbb{R}^d\}.$$

Then, obviously $u_n(t) = 1$ for $t \in \tilde{\mathbb{B}}_n^d$ and it is supported on $\mathbb{B}_n^d$. Hence it is readily seen that there exists $c_\circ = c_\circ(r, d, p, v) > 0$, such that

$$\omega_n(t) := c_\circ n^{-r+d/p} u_n \in W_p^{\circ,r}(\mathbb{B}^d).$$

Then $\omega_n(t) = c_\circ n^{-r+d/p}$ for $t \in \tilde{\mathbb{B}}_n^d$, it is supported on $\mathbb{B}_n^d$, and

$$\|\omega_n\|_{L_q(\tilde{\mathbb{B}}^d)} \ge c^\circ n^{-r+d(1/p-1/q)}, \tag{3.4}$$

where $c^\circ = c^\circ(r, d, p, q, v) > 0$.

We will show that there exists $c_* = c_*(r, d, p, q, v) > 0$, such that for every polynomial $P_n \in \mathcal{P}_n(\mathbb{B}^d)$,

$$\|\omega_n - P_n\|_{L_q(\mathbb{B}^d)} \ge c_* n^{-r+d(1/p-1/q)}, \quad 1 \le p < q \le \infty. \tag{3.5}$$

To this end, set $\bar{c} := c^\circ(c^*+2)^{-1}$, where c° is the constant from (3.4) and c^* is the constant from Lemma 4, and assume to the contrary that

$$\|\omega_n - P_n\|_{L_q(\mathbb{B}^d)} \le \bar{c}n^{-r+d(1/p-1/q)}. \tag{3.6}$$

Since ω_n is supported on $\mathbb{B}_n^d$, it follows that

$$\|P_n\|_{L_q(\mathbb{B}^d\setminus\mathbb{B}_n^d)} \le \bar{c}n^{-r+d(1/p-1/q)}.$$

By virtue of Lemma 4 we conclude that

$$\|P_n\|_{L_q(\mathbb{B}_n^d)} \le c^*\bar{c}n^{-r+d(1/p-1/q)},$$

which, in turn, implies

$$\|P_n\|_{L_q(\tilde{\mathbb{B}}_n^d)} \le c^*\bar{c}n^{-r+d(1/p-1/q)}.$$

Hence, by (3.4),

$$\begin{aligned}\|\omega_n - P_n\|_{L_q(\tilde{\mathbb{B}}_n^d)} &\ge \|\omega_n\|_{L_q(\tilde{\mathbb{B}}_n^d)} - \|P_n\|_{L_q(\tilde{\mathbb{B}}_n^d)} \\ &\ge (c^\circ - c^*\bar{c})n^{-r+d(1/p-1/q)} \\ &= 2\bar{c}n^{-r+d(1/p-1/q)},\end{aligned}$$

which contradicts (3.6). This proves (3.5), which readily implies

$$E\left(\omega_n, \mathcal{P}_n(\mathbb{B}^d)\right)_{L_q(\mathbb{B}^d)} \ge c_* n^{-r+d(1/p-1/q)}, \quad 1 \le p < q \le \infty.$$

Therefore,

$$E\left(W_p^{\circ,r}(\mathbb{B}^d), \mathcal{P}_n(\mathbb{B}^d)\right)_{L_q(\mathbb{B}^d)} \ge c_* n^{-r+d(1/p-1/q)}, \quad 1 \le p < q \le \infty, \tag{3.7}$$

and (3.3) is proved for $1 \le p < q \le \infty$.

We believe that if $1 \le q \le p \le \infty$, then the lower bounds for the best approximation by polynomials of the classes $W_p^{\circ,r}(\mathbb{B}^d)$ are essentially known. However, we could find no reference which provides the lower bound that we need, so we provide a simple proof of (3.3) in this case, namely,

$$E\left(W_\infty^{\circ,r}(\mathbb{B}^d), \mathcal{P}_n(\mathbb{B}^d)\right)_{L_1(\mathbb{B}^d)} \ge c_* n^{-r}, \tag{3.8}$$

where $c_* = c_*(r, d, v) > 0$ depends on r, d and some fixed function v.

To this end, fix $n > 1$ and divide $[1/2, 1]$ into n subintervals,

$$I_{n,i} := [1/2 + (i-1)/2n, 1/2 + i/2n], \quad i = 1, \ldots, n,$$

and similarly, divide $[-1, -1/2]$ into n subintervals

$$I_{n,i} := [-1/2 - i/2n, -1/2 - (i+1)/2n], \quad i = -1, \ldots, -n.$$

Take a nonnegative function $v \in C_0^\infty(\mathbb{R})$ with $\operatorname{supp} v = [-1, 1]$ and such that $v(t) = 1$ for $t \in [-1/2, 1/2]$. Let

$$t_{n,i} := \begin{cases} 1/2 + (2i-1)/(4n), & i = 1, \ldots, n, \\ -1/2 + (2i+1)/(4n), & i = -1, \ldots, -n, \end{cases}$$

and denote

$$f_n(t) := \sum_{i=\pm 1}^{\pm n} (-1)^i v(4n(t - t_{n,i})), \quad t \in I := (-1, 1).$$

Note that f_n is even and alternates in sign between $I_{n,i}$ and $I_{n,i+1}$, $i = 1, \dots, n-1$. Finally, let $w_n(t) := c_\circ n^{-r} f_n(|t|)$, $t \in \mathbb{B}^d$, where $c_\circ = c_\circ(r, d, v) > 0$ is so chosen that $w_n(t) \in W_\infty^{\circ,r}(\mathbb{B}^d)$. By definition it clearly satisfies

$$w_n(t)\,\mathrm{sgn}(-1)^i = c_\circ n^{-r}, \quad \frac{1}{2} + \frac{4i-3}{8n} \le |t| \le \frac{1}{2} + \frac{4i-1}{8n}, \quad i = 1, \dots, n. \tag{3.9}$$

In order to prove (3.8) it suffices to prove that there exists $c_* = c_*(r, d, v)$, such that for every polynomial $P_n \in \mathcal{P}_n(\mathbb{B}^d)$ the inequality holds

$$\|w_n - P_n\|_{L_1(\mathbb{B}^d)} \ge c_* n^{-r}. \tag{3.10}$$

We recall the spherical coordinates defined above, namely, let $\varphi = (\tilde{\varphi}, \varphi_{d-1})$ be such that $\tilde{\varphi} \in [0, \pi]^{d-2}$, and $\varphi_{d-1} \in [0, \pi]$. Again, denote by I_φ the chord connecting the two points $(1, \tilde{\varphi}, \varphi_{d-1})$ and $(1, \tilde{\varphi}, \varphi_{d-1} + \pi)$ on the unit sphere $\mathbb{S}^{d-1}$, and recall that I_φ can be represented as the set of all points $t = \tau e(\varphi)$, $\tau \in (-1, 1) =: I$, where $e(\varphi) := (1, \tilde{\varphi}, \varphi_{d-1}) \in \mathbb{R}^d$.

For $\tau \in I$ we denote by $w_n(\tau; I_\varphi)$ and $P_n(\tau; I_\varphi)$ the restrictions, to I_φ, of the function $w_n(\cdot)$ and the polynomial $P_n(\cdot)$, respectively. Clearly, $P_n(\tau; I_\varphi)$ is the univariate polynomial in $\tau \in I$ of degree $\le n$.

By (3.9) and the fact that w_n is even, we have

$$w_n(\tau; I_\varphi)\,\mathrm{sgn}(-1)^i \ge c_\circ n^{-r}, \quad \tau \in I'_{n,i}, \quad i = \pm 1, \dots, \pm n, \tag{3.11}$$

where

$$I'_{n,i} := \left[\frac{1}{2} + \frac{4i-3}{8n}, \frac{1}{2} + \frac{4i-1}{8n}\right], i = 1, \dots, n,$$

and

$$I'_{n,i} := \left[-\frac{1}{2} + \frac{4i+1}{8n}, -\frac{1}{2} + \frac{4i+3}{8n}\right], \quad i = -1, \dots, -n.$$

Since the univariate polynomial $P_n(\cdot; I_\varphi)$ changes its sign at most n times in I, we conclude that there exist $(n-1)$ distinct indices i_k, $k = 1, \dots, n-1$ (recall that $n - 1 \ge 1$), such that

$$\mathrm{sgn}\, P_n(\tau; I_\varphi) \ne \mathrm{sgn}\, w_n(\tau; I_\varphi), \quad \tau \in I'_{n,i_k}, \quad k = 1, \dots, n-1,$$

which in turn implies by (3.11),

$$\begin{aligned} |w_n(\tau; I_\varphi) - P_n(\tau; I_\varphi)| &\ge |w_n(\tau; I_\varphi)| \\ &\ge c_\circ n^{-r}, \quad \tau \in I'_{n,i_k}, \quad k = 1, \dots, n-1. \end{aligned} \tag{3.12}$$

Now,

$$\begin{aligned} &\int_I |\tau|^{d-1} |w_n(\tau; I_\varphi) - P_n(\tau; I_\varphi)| \mathrm{d}\tau \\ &\quad \ge \int_{[-1,-1/2]\cup[1/2,1]} |\tau|^{d-1} |w_n(\tau; I_\varphi) - P_n(\tau; I_\varphi)| \mathrm{d}\tau \end{aligned}$$

$$
\begin{aligned}
&\geq 2^{-d+1}\int_{[-1,-1/2]\cup[1/2,1]}|w_n(\tau;I_\varphi)-P_n(\tau;I_\varphi)|\mathrm{d}\tau\\
&\geq 2^{-d+1}\sum_{k=1}^{n-1}\int_{I'_{n,i_k}}|w_n(\tau;I_\varphi)-P_n(\tau;I_\varphi)|\mathrm{d}\tau\\
&\geq 2^{-d-1}(1-1/n)c_\circ n^{-r}\\
&\geq 2^{-d-2}c_\circ n^{-r},
\end{aligned}
$$

where for the last inequality we have applied (3.12) and the fact that $|I'_{n,i_k}| = 1/(4n)$, $1 \leq k \leq n-1$, and $n > 1$.

Applying (2.2) we obtain

$$
\|w_n - P_n\|_{L_1(\mathbb{B}^d)} \geq \left(\int_{[0,\pi]^{d-2}} J_{d-1}(\tilde{\varphi})\mathrm{d}\tilde{\varphi}\right) 2^{-d-2}c_\circ n^{-r} =: c_* n^{-r},
$$

which concludes the proof of (3.10), and in turn implies (3.8) since $w_n \in W_\infty^{\circ,r}(\mathbb{B}^d)$. Hence the lower bounds

$$
E\left(W_p^{\circ,r}(\mathbb{B}^d), \mathcal{P}_n(\mathbb{B}^d)\right)_{L_q(\mathbb{B}^d)} \geq c_* n^{-r}, \quad 1 \leq q \leq p \leq \infty. \tag{3.13}
$$

Combining (3.7) and (3.13), the lower bounds (3.1) follow for all $1 \leq p, q \leq \infty$ such that $r - d(1/p - 1/q) > 0$. The proof of Theorem 1 is complete. □

Proof of Theorem 2. Since $\mathcal{P}_n(\mathbb{B}^d) \subset \mathcal{R}_N(\mathbb{B}^d)$ where $N \asymp n^{d-1}$, the upper bounds follow immediately from the upper bounds for $q = 2$ in Theorem 1.

The lower bounds are an immediate consequence of Lemma 5 combined with the lower bounds for $q = 2$ in Theorem 1. This completes the proof of Theorem 2. □

References

[1] P. Borwein, T. Erdélyi, Polynomials and Polynomial Inequalities, in: Graduate Texts in Mathematics, vol. 161, Springer Verlag, New York, 1995.

[2] V.I. Burenkov, Sobolev's integral representation and Taylor's formula, Proceedings of the Steklov Institute of Mathematics 131 (1974) 33–38.

[3] M.E. Davison, F.A. Grunbaum, Tomographic reconstruction with arbitrary directions, Communications on Pure and Applied Mathematics 34 (1981) 77–120.

[4] Y. Gordon, V. Maiorov, M. Meyer, S. Reisner, On Best Approximation by Ridge Functions in the Uniform Norm, Constructive Approximation 18 (2002) 61–85.

[5] H. Groemer, Geometric Applications of Fourier Series and Spherical Harmonics, Cambridge Univ. Press, 1996.

[6] P.C. Kainen, V. Kurkova, A. Vogt, A Sobolev-type upper bound for rates of approximation by linear combinations of Heaviside plane waves, Journal of Approximation Theory 147 (2007) 1–10.

[7] V.N. Konovalov, D. Leviatan, V.E. Maiorov, Approximation by polynomials and ridge functions of the classes of s-monotone radial functions, Journal of Approximation Theory 152 (2008) 20–51.

[8] V.E. Maiorov, On best approximation by ridge functions, JAT 99 (1999) 68–94.

[9] V.E. Maiorov, R. Meir, J. Ratsaby, On the approximation of functional classes equipped with a uniform measure using ridge functions, Journal of Approximation Theory 99 (1999) 95–111.

[10] V.E. Maiorov, K.I. Oskolkov, V.N. Temlyakov, Gridge approximation and Radon compass, Approximation Theory: A volume dedicated to B. Sendov. B. Bojanov (Ed.), DARBA, Sofia, (2002) pp. 284–309.

[11] H.M. Mhaskar, On the tractability of multivariate integration and approximation by neural networks, Journal of Complexity 20 (2004) 561–590.

[12] H.M. Mhaskar, When is approximation by Gaussian networks necessarily a linear process?, Neural Networks 17 (2004) 989–1001.

[13] H.M. Mhaskar, A Markov–Bernstein inequality for Gaussian networks, in: M.G. de Bruin, D.H. Mache, J. Szabados (Eds.), Trends and Applications in Constructive Approximation, in: ISNM, vol. 105, Birkhäuser Verlag, Bassel, 2005.
[14] H.M. Mhaskar, F.J. Narcowich, J.D. Ward, Approximation properties of zonal function networks using scattered data on the sphere, Advances in Computational Mathematics 11 (1999) 121–137.
[15] K.I. Oskolkov, Ridge approximation, Chebyshev-Fourier analysis and optimal quadrature formulas, Proceedings of the Steklov Institute of Mathematics 219 (1997) 265–280.
[16] K.I. Oskolkov, Linear and non-linear methods of relief approximation, Metric theory of functions and adjoining questions of analysis, Transaction of Papers: dedicated to P. L. Ulyianov, 1999, pp. 165–195 (in Russian).
[17] A. Pinkus, Approximation theory of the MLP model in neural networks, Acta Numerica 8 (1999) 143–195.
[18] A.G. Ramm, A.I. Katsevich, The Radon Transform and Local Tomography, Boca Raton, CA, CRC Press, 1996.
[19] E.M. Stein, Singular Integrals and Differentiability Properties of Functions, Princeton, New Jersey, 1970.
[20] V. Temlyakov, On Approximation by Ridge Functions, Department of Mathematics, University of South Carolina, 1996, p. 12.
[21] A.F. Timan, Theory of Approximation of Function of the real Variable, Macmillan Co., New York, 1963.