Available online at www.sciencedirect.com
ScienceDirect
Journal of Approximation Theory 162 (2010) 1964–1981
JOURNAL OF Approximation Theory
www.elsevier.com/locate/jat

The best m-term approximations on generalized Besov classes $MB^{\Omega}_{q,\theta}$ with regard to orthogonal dictionaries☆

Liqin Duan

Institute of Mathematics, Hangzhou Dianzi University, Hangzhou 310018, China

Received 7 October 2009; received in revised form 23 March 2010; accepted 14 May 2010
Available online 24 May 2010

Communicated by Przemyslaw Wojtaszczyk

Abstract

In this paper, we investigate nonlinear m-term approximation with regard to orthogonal dictionaries. We consider this problem in the periodic multivariate case for generalized Besov classes $MB^{\Omega}_{q,\theta}$ under the condition $\Omega(\mathbf{t}) = \omega(t_1 \cdot \cdots \cdot t_d)$ where $\omega(t) \in \mathbf{\Psi}^*_l$ is a univariate function. We prove that the well-known dictionary U^d which consists of trigonometric polynomials (shifts of the Dirichlet kernels) is nearly optimal among orthogonal dictionaries. Moreover, it is established that for these classes near-best m-term approximation, with regard to U^d, can be achieved by simple greedy-type algorithms.

Keywords: Best m-term approximation; Orthogonal dictionary; Generalized Besov class; Greedy algorithm

1. Introduction

Nonlinear approximation, namely, m-term approximation, has been widely researched recently. Nonlinear m-term approximation is important in applications in image and signal processing (see, for instance, the survey [2]). One of the major questions in approximation (theoretical and numerical) is: What is an optimal method? Here, we discuss this question in a theoretical setting, with the only criterion of the quality of the approximating method its accuracy. One more important point in the setting of the optimization problem is specifying a

☆ Project supported by the Natural Science Foundation of China (Grants No. 10926056 and No. 10671009) and the Research Foundation of Hangzhou Dianzi University (Grant No. KYS091507089).
E-mail address: qinandfeng@163.com.

0021-9045/$ - see front matter
doi:10.1016/j.jat.2010.05.005

set of methods over which we are going to optimize—for example, when we are solving the problem on Kolmogorov's n-width for a given function class that we are optimizing, in the sense of accuracy for a given class over all subspaces of dimension n.

In what follows, we give the formulation of best m-term approximation. For the best m-term approximation of a given function with regard to a given system of functions (dictionary), we are optimizing over all m-dimensional subspaces spanned by elements from a given dictionary. In this paper, we want to solve the problem of the best m-term approximation for a given function class with regard to orthogonal dictionaries.

Denote by $\mathcal{D}$ a dictionary in a Banach space X and by

$$\sigma_m(f,\mathcal{D})_X := \inf_{g_i\in\mathcal{D},c_i,i=1,\ldots,m}\left\| f-\sum_{i=1}^m c_i g_i\right\|_X$$

the best m-term approximation of f with regard to $\mathcal{D}$. For a function class $F\subset X$ and a collection $\mathbf{D}$ of dictionaries we consider

$$\sigma_m(F,\mathcal{D})_X := \sup_{f\in F}\sigma_m(f,\mathcal{D})_X,$$
$$\sigma_m(F,\mathbf{D})_X := \inf_{\mathcal{D}\in\mathbf{D}}\sigma_m(F,\mathcal{D})_X.$$

Thus the quantity $\sigma_m(F,\mathbf{D})_X$ gives the sharp lower bound for the best m-term approximation of a given function class F with regard to any dictionary $\mathcal{D}\in\mathbf{D}$.

Denote by $\mathbf{O}$ the set of all orthonormal dictionaries defined on a given domain. Kashin [5] proved that for the class $H^{r,\alpha}$, $r=0,1,\ldots,$ $\alpha\in(0,1]$, of univariate functions such that

$$\|f\|_\infty+\|f^{(r)}\|_\infty\le 1 \quad\text{and}\quad |f^{(r)}(x)-f^{(r)}(y)|\le|x-y|^\alpha,\quad x,y\in[0,1],$$

we have

$$\sigma_m(H^{r,\alpha},\mathbf{O})_{L_2}\ge C(r,\alpha)m^{-r-\alpha}. \tag{1}$$

It is interesting to remark that estimates like (1) with L_2 replaced by L_p, $p<2$, cannot be obtained. Kashin and Temlyakov in [6] proved that there exists $\Phi\in\mathbf{O}$ such that for any $f\in L_1(0,1)$ one can obtain $\sigma_1(f,\Phi)_{L_1}=0$. The proof from [6] also works for L_p, $p<2$, instead of L_1. This remark means that to obtain nontrivial lower bounds for $\sigma_m(f,\Phi)_{L_p}$, $p<2$, we need to impose additional restrictions on $\Phi\in\mathbf{O}$. In [14] Temlyakov considered the best m-term approximation of classes of functions with bounded mixed derivative MW^r_q and classes with a restriction of Lipschitz type on the mixed difference MH^r_q with regard to orthogonal dictionaries and obtained the following results:

$$\sigma_m(MH^r_q,\mathbf{O})_{L_2}\gg m^{-r}(\log m)^{(d-1)(r+1/2)},\quad 1\le q<\infty, \tag{2}$$

$$\sigma_m(MW^r_q,\mathbf{O})_{L_2}\gg m^{-r}(\log m)^{(d-1)r},\quad 1\le q<\infty. \tag{3}$$

He also proved that the orthogonal basis U^d provides optimal upper estimates (like (2) and (3)) in the best m-term approximation of the classes MH^r_q and MW^r_q in the L_p-norm, $2\le p<\infty$. Moreover, he proved that for all $1<q,p<\infty$ the order of the best m-term approximation $\sigma_m(MH^r_q,U^d)_{L_p}$ and $\sigma_m(MW^r_q,U^d)_{L_p}$ can be achieved by a greedy-type algorithm $G^p(\cdot,U^d)$. In [15], he also studied universal bases and greedy algorithms for anisotropic function classes. Moreover, some authors also investigated the best m-term approximation of classes of functions with regard to the dictionaries with special structure. In [16], Wang studied the greedy algorithm for functions with low mixed smoothness with respect to the wavelet-type basis.

Temlyakov gave the definition of the greedy algorithm $G^p(\cdot, \Psi)$ as follows (see [14]). Assume that a given system Ψ of functions ψ_I indexed by dyadic intervals can be enumerated in such a way that $\{\psi_{I^j}\}_{j=1}^{\infty}$ is a basis for L_p. Let

$$f = \sum_{j=1}^{\infty} c_{I^j}(f, \Psi)\psi_{I^j}$$

and

$$c_I(f, p, \Psi) := \|c_I(f, \Psi)\psi_I\|_p.$$

Then $c_I(f, p, \Psi) \to 0$ as $|I| \to 0$. Denote Λ_m as a set of m dyadic intervals I such that

$$\min_{I\in\Lambda_m} c_I(f, p, \Psi) \geq \max_{J\notin\Lambda_m} c_J(f, p, \Psi).$$

Define $G^p(\cdot, \Psi)$ by the formula

$$G_m^p(f, \Psi) := \sum_{I\in\Lambda_m} c_I(f, \Psi)\psi_I.$$

The above defined "greedy algorithm" $G_m^p(f, \Psi)$ gives a procedure for constructing an approximant which turns out to be a good approximant, while the procedure for constructing $G_m^p(f, \Psi)$ is not a numerical algorithm ready for computational implementation. Therefore it would be more precise to call this procedure a "theoretical greedy algorithm" or "stepwise optimizing process". In [14], Temlyakov also obtained the following results: for $1 < q, p < \infty$ and big enough r,

$$\sup_{f\in MH_q^r} \|f - G_m^p(f, U^d)\|_p \asymp \sigma_m(MH_q^r, U^d)_p \asymp m^{-r}(\log m)^{(d-1)(r+1/2)}, \tag{4}$$

$$\sup_{f\in MW_q^r} \|f - G_m^p(f, U^d)\|_p \asymp \sigma_m(MW_q^r, U^d)_p \asymp m^{-r}(\log m)^{(d-1)r}, \tag{5}$$

where $\|\cdot\|_p := \|\cdot\|_{L_p}$.

Comparing (4) with (2) and (5) with (3), it can be concluded that the dictionary U^d is the best (in the sense of order) among all orthogonal dictionaries for m-term approximation of the classes MH_q^r and MW_q^r in L_p where $1 < q < \infty$ and $2 \leq p < \infty$. The near-best m-term approximation of functions from MH_q^r and MW_q^r in the L_p-norm can be realized by the simple greedy-type algorithm $G^p(\cdot, U^d)$ for all $1 < q, p < \infty$.

It is well-known that Besov classes of functions have been widely and deeply researched by many authors. Many approximate characteristics of Besov classes have been obtained. For more information, we can refer the reader to [7,9–11] and the references therein. In [3], Dung considered nonlinear approximation of Besov classes with regard to the dictionary V which consists of trigonometric polynomials (shifts of the de la Vallee Poussin kernels) and obtained the asymptotic order of the best m-term approximation through a continuous algorithm. In addition, Wang [17] studied the best m-term approximation on Besov classes with respect to the tensor product periodic wavelet basis and gave the asymptotic order through a greedy-type algorithm. For more information about nonlinear approximation, one can refer to [14–17] and the papers given there. In 1994, Pustovoitov [8] introduced a function class $H_q^{\Omega}(\mathbf{T}^d)$. He first used a standard function $\Omega(\mathbf{t})$, a prototype of which is $\Omega(\mathbf{t}) = \mathbf{t}^{\mathbf{r}} := t_1^{r_1} \cdots t_d^{r_d}$ as a majorant function for the mixed modulus of smoothness of order l of functions $f \in L_q$ instead of the standard function $\mathbf{t}^{\mathbf{r}}$, and obtained the estimates of best approximations of classes H_q^{Ω} with some special $\Omega(t_1, \ldots, t_d)$. In 1997, Sun and Wang [13] introduced the Besov classes $B_{q,\theta}^{\Omega}(\mathbf{T}^d)$ by means of $\Omega(\mathbf{t})$, i.e., an extension of the Besov classes $S_{q,\theta}^{\mathbf{r}}(\mathbf{T}^d)$, which was introduced first by Amanov [1] and gave the asymptotic estimates for Kolmogorov n-widths of the classes under the condition $\Omega(\mathbf{t}) = \omega(t_1 \cdot \cdots \cdot t_d)$, where $\omega(t) \in \Psi_l^*$ (i.e., a univariate function) and Ψ_l^* will be given below. In addition, in [12,4], Stasyuk and Fedunyk studied the Kolmogorov and linear widths of $B_{q,\theta}^{\Omega}(\mathbf{T}^d)$ for some values of parameters p, q, θ, respectively. In this paper, we will consider the best m-term approximation of generalized Besov classes $MB_{q,\theta}^{\Omega}(\mathbf{T}^d)$ under the condition $\Omega(\mathbf{t}) = \omega(t_1 \cdot \cdots \cdot t_d)$, where $\omega(t) \in \Psi_l^*$ (i.e., a univariate function) with regard to orthogonal dictionaries and prove that the orthogonal basis U^d which consists of trigonometric polynomials (shifts of the Dirichlet kernels) is nearly optimal among orthogonal dictionaries. Moreover, we prove that the order of best m-term approximation $\sigma_m(MB_{q,\theta}^{\Omega}, U^d)_p$ can be achieved by a greedy-type algorithm $G^p(\cdot, U^d)$.

Throughout this paper, we will use the notation $\ll$ and $\asymp$. For two sequences $\{a_n\}_{n\in\mathbb{N}}$ and $\{b_n\}_{n\in\mathbb{N}}$ of positive real numbers, we write $a_n \ll b_n$ provided that $a_n \le c b_n$ for certain $c > 0$. If, furthermore, also $b_n \ll a_n$, then we write $a_n \asymp b_n$. Let $a_+ := \max\{0, a\}$ and $b_- := \min\{0, b\}$.

The paper is organized as follows: In Section 2, we give the orthogonal basis U^d constructed by Temlyakov and generalized Besov classes $MB_{q,\theta}^{\Omega}(\mathbf{T}^d)$ which will be studied in this paper. In Section 3, we will give the main results and their proofs.

2. Preliminary

In this section, we start with the construction of the orthogonal basis U^d and some properties. Then we will give the definition of the generalized Besov class $MB_{q,\theta}^{\Omega}(\mathbf{T}^d)$ which will be investigated in this paper.

We first recall the system $U := \{U_I\}$ in the univariate case. Define

$$U_n^+(x) := \sum_{k=0}^{2^n-1} \mathrm{e}^{\mathrm{i}kx} = \frac{\mathrm{e}^{\mathrm{i}2^n x} - 1}{\mathrm{e}^{\mathrm{i}x} - 1}, \quad n = 0, 1, \ldots,$$
$$U_{n,k}^+(x) := \mathrm{e}^{\mathrm{i}2^n x} U_n^+(x - 2\pi k 2^{-n}), \quad k = 0, 1, \ldots, 2^n - 1,$$
$$U_{n,k}^-(x) := \mathrm{e}^{-\mathrm{i}2^n x} U_n^+(-x + 2\pi k 2^{-n}), \quad k = 0, 1, \ldots, 2^n - 1.$$

It will be more convenient for us to normalize in L_2 the system of functions $\{U_{n,k}^+, U_{n,k}^-\}$ and to enumerate it using dyadic intervals. Write $U_{[0,1)}(x) = 1$,

$$U_I(x) := 2^{-n/2} U_{n,k}^+(x) \quad \text{with } I = [(k + 1/2)2^{-n}, (k+1)2^{-n}]$$

and

$$U_I(x) := 2^{-n/2} U_{n,k}^-(x) \quad \text{with } I = [k 2^{-n}, (k + 1/2)2^{-n}].$$

Define

$$D_n^+ := \{I : I = [(k + 1/2)2^{-n}, (k+1)2^{-n}],\ k = 0, 1, \ldots, 2^n - 1\}$$

and

$$D_n^- := \{I : I = [k2^{-n}, (k+1/2)2^{-n}],\ k = 0, 1, \dots, 2^n - 1\},$$
$$D_0^+ = D_0^- = D_0 := [0, 1), \qquad D := \bigcup_{n\ge 1}(D_n^+ \cup D_n^-) \cup D_0.$$

It is easy to check that for any $I, J \in D,\ I \neq J$ we have

$$\langle U_I, U_J\rangle = (2\pi)^{-1}\int_0^{2\pi} U_I(x)\bar{U}_J(x)\mathrm{d}x = 0,$$

and

$$\|U_I\|_2^2 = 1.$$

We use the following notation for $f \in L_1$:

$$f_I := \langle f, U_I\rangle = (2\pi)^{-1}\int_0^{2\pi} f(x)\bar{U}_I(x)\mathrm{d}x, \qquad \hat{f}(k) := (2\pi)^{-1}\int_0^{2\pi} f(x)\mathrm{e}^{-\mathrm{i}kx}\mathrm{d}x,$$

and

$$\delta_s^+(f) := \sum_{k=2^s}^{2^{s+1}-1}\hat{f}(k)\mathrm{e}^{\mathrm{i}kx}, \qquad \delta_s^-(f) := \sum_{k=-2^{s+1}+1}^{-2^s}\hat{f}(k)\mathrm{e}^{\mathrm{i}kx}, \qquad \delta_0(f) := \hat{f}(0).$$

Then, for each s and $f \in L_1$, we have

$$\delta_s^+(f) = \sum_{I\in D_s^+} f_I U_I, \qquad \delta_s^-(f) = \sum_{I\in D_s^-} f_I U_I, \qquad \delta_0(f) = f_{[0,1)}.$$

Moreover, the following important analog of the Marcinkiewicz theorem holds:

$$\|\delta_s^+(f)\|_p^p \asymp \sum_{I\in D_s^+}\|f_I U_I\|_p^p, \qquad \|\delta_s^-(f)\|_p^p \asymp \sum_{I\in D_s^-}\|f_I U_I\|_p^p, \tag{6}$$

for $1 < p < \infty$ with the constants depending only on p.

We remark that

$$\|U_I\|_p \asymp |I|^{1/p-1/2}, \quad 1 < p \le \infty, \tag{7}$$

which implies that for any $1 < q, p < \infty$

$$\|U_I\|_p \asymp \|U_I\|_q |I|^{1/p-1/q}. \tag{8}$$

In the multivariate case of $\mathbf{x} = (x_1, \dots, x_d)$, define the system U^d as the tensor product of the univariate systems U. Let $I = I_1 \times \cdots \times I_d,\ I_j \in D,\ j = 1, \dots, d$; then

$$U_I(\mathbf{x}) := \prod_{j=1}^d U_{I_j}(x_j).$$

In [14], Temlyakov proved that the system U^d can be enumerated in such a way that $\{U_{I^k}\}_{k=1}^\infty$ forms a basis for each $L_p,\ 1 < p < \infty$. In fact, for any $f \in L_p$, it can be uniquely represented as

$$f(\mathbf{x}) = \sum_{k=1}^{\infty} f_{I^k} U_{I^k}(\mathbf{x}), \qquad f_{I^k} := \int_{\mathbf{T}^d} f(\mathbf{x}) \bar{U}_{I^k}(\mathbf{x}) \mathrm{d}\mathbf{x},$$

in the sense of convergence in L_p.

For $\mathbf{s} = (s_1, \ldots, s_d)$ and $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_d)$, $\varepsilon_j = +$ or $-$, define

$$D_{\mathbf{s}}^{\boldsymbol{\varepsilon}} := \{I : I = I_1 \times \cdots \times I_d,\ I_j \in D_{s_j}^{\varepsilon_j},\ j = 1, \ldots, d\}, \quad \text{and} \quad D^d := \bigcup_{\mathbf{s},\boldsymbol{\varepsilon}} D_{\mathbf{s}}^{\boldsymbol{\varepsilon}}.$$

It is easy to see that (7) and (8) are also true in the multivariate case. It is not difficult to derive from (6) that for any $\boldsymbol{\varepsilon}$ we have

$$\|\delta_{\mathbf{s}}^{\boldsymbol{\varepsilon}}(f)\|_p^p \asymp \sum_{I \in D_{\mathbf{s}}^{\boldsymbol{\varepsilon}}} \|f_I U_I\|_p^p, \quad 1 < p < \infty, \tag{9}$$

with constants depending on p and d. Here we define

$$\delta_{\mathbf{s}}^{\boldsymbol{\varepsilon}}(f) := \sum_{\mathbf{k} \in \rho(\mathbf{s},\boldsymbol{\varepsilon})} \hat{f}(k) \mathrm{e}^{\mathrm{i}(\mathbf{k},\mathbf{x})},$$

where

$$\rho(\mathbf{s}, \boldsymbol{\varepsilon}) := \varepsilon_1[2^{s_1}, 2^{s_1+1} - 1] \times \cdots \times \varepsilon_d[2^{s_d}, 2^{s_d+1} - 1].$$

We will often use the following inequality:

$$\left(\sum_{\mathbf{s},\boldsymbol{\varepsilon}} \|\delta_{\mathbf{s}}^{\boldsymbol{\varepsilon}}(f)\|_p^{p_l} \right)^{1/p_l} \ll \|f\|_p \ll \left(\sum_{\mathbf{s},\boldsymbol{\varepsilon}} \|\delta_{\mathbf{s}}^{\boldsymbol{\varepsilon}}(f)\|_p^{p_u} \right)^{1/p_u}, \quad 1 < p < \infty, \tag{10}$$

where $p_l := \max(2, p)$; $p_u := \min(2, p)$.

Now we introduce the space of functions which will be studied in this paper. Let $\mathbf{R}^d$ be the Euclidean space with dimension d. Denote by $L_q(\mathbf{T}^d)$, $1 < q < \infty$, the Lebesgue space of qth-power integrable functions defined on the d-dimensional torus $\mathbf{T}^d := [0, 2\pi)^d$, which are 2π-periodic with respect to each variable. Its norm is defined by

$$\|f\|_{L_q(\mathbf{T}^d)} = \|f\|_q := \left\{ (2\pi)^{-d} \int_{\mathbf{T}^d} |f(\mathbf{x})|^q \mathrm{d}\mathbf{x} \right\}^{1/q} < \infty, \quad 1 \le q < \infty;$$

$$\|f\|_{L_\infty(\mathbf{T}^d)} = \|f\|_\infty := \operatorname*{ess\,sup}_{\mathbf{x} \in \mathbf{T}^d} |f(\mathbf{x})| < \infty, \quad q = \infty.$$

In what follows, we assume that functions $f(\mathbf{x})$ belong to the space

$$L_q^0(\mathbf{T}^d) = \left\{ f : f \in L_q(\mathbf{T}^d),\ \int_{-\pi}^{\pi} f(\mathbf{x}) \mathrm{d}x_j = 0,\ j = 1, \ldots, d \right\}.$$

For $f \in L_q^0(\mathbf{T}^d)$, we set

$$\Omega^l(f, \mathbf{t})_q := \sup_{|\mathbf{h}| \le \mathbf{t}} \|\Delta_{\mathbf{h}}^l f(\mathbf{x})\|_q,$$

where $l \in \mathbf{Z}_+$ is a fixed positive integer, $\mathbf{t} = (t_1, \ldots, t_d) \geq 0$ (i.e., $t_j \geq 0,\ j = 1, \ldots, d$), $\mathbf{h} = (h_1, \ldots, h_d)$, $|\mathbf{h}| := (|h_1|, \ldots, |h_d|)$, and $|\mathbf{h}| \leq \mathbf{t}$ means $|h_j| \leq t_j,\ j = 1, \ldots, d$. Finally,

$$\Delta_{\mathbf{h}}^l f(\mathbf{x}) := \Delta_{h_d,d}^l(\Delta_{h_{d-1},d-1}^l \ldots (\Delta_{h_1,1}^l f(\mathbf{x})) \ldots),$$

where

$$\Delta_{h_i,i}^l f(\mathbf{x}) = \sum_{j=0}^{l} (-1)^{l-j} \binom{l}{j} f(x_1, \ldots, x_{i-1}, x_i + jh_i, x_{i+1}, \ldots, x_d), \quad i = 1, \ldots, d.$$

As we know, $\Omega^l(f, \mathbf{t})_q$ is the order l modulus of smoothness in $L_q(\mathbf{T}^d)$ norm (of mixed type).

In order to give the definition of the generalized Besov spaces $B_{q,\theta}^{\Omega}(\mathbf{T}^d)$, we need the following definitions.

Definition 1. Let $\phi : \mathbf{R}_+^d \to \mathbf{R}_+ = [0, \infty)$ be a non-negative function defined on $\mathbf{R}_+^d := \{(x_1, \ldots, x_d) |\, x_j \geq 0,\ j = 1, \ldots, d\}$. We say that $\phi(\mathbf{t}) = \phi(t_1, \ldots, t_d) \in \Phi_l^*$ if it satisfies

(1) $\phi(\mathbf{0}) = 0$, $\phi(\mathbf{t}) > 0$ for any $\mathbf{t} \in \mathbf{R}_+^d$, $\mathbf{t} > \mathbf{0}$ (i.e., $t_j > 0,\ j = 1, \ldots, d$);
(2) $\phi(\mathbf{t})$ is continuous;
(3) $\phi(\mathbf{t})$ is almost increasing, i.e., for any two points $\mathbf{t}, \boldsymbol{\tau} \in \mathbf{R}_+^d$ and $\mathbf{0} \leq \mathbf{t} \leq \boldsymbol{\tau}$ (i.e., $0 \leq t_i \leq \tau_j,\ j = 1, \ldots, d$), we have $\phi(\mathbf{t}) \leq C\phi(\boldsymbol{\tau})$, where $C \geq 1$ is a constant independent of $\mathbf{t}$;
(4) for any $\mathbf{n} := (n_1, \ldots, n_d) \in \mathbf{Z}_+^d$

$$\phi(n_1 t_1, n_2 t_2, \ldots, n_d t_d) \leq C \left(\prod_{j=1}^{d} n_j\right)^l \phi(t_1, \ldots, t_d),$$

where $l \geq 1$ is a fixed positive integer, and $C > 0$ is a constant independent of $\mathbf{n}$ and $\mathbf{t}$.

If $f \in L_q(\mathbf{T}^d)$, $f \neq \text{const}$, then $\Omega^l(f, \mathbf{t})_q \in \Phi_l^*$ with both constants $C = 1$. Throughout this paper, the capital letter C has different values in different places.

Definition 2. Let $\phi(\mathbf{t})$ be a non-negative function defined on $\mathbf{R}_+^d$ which satisfies conditions (1), (2) in Definition 1. We say that $\phi(\mathbf{t}) \in S^*$ provided that there exists a vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d) > \mathbf{0}$ such that $\phi(\mathbf{t})\mathbf{t}^{-\boldsymbol{\alpha}}$ is almost increasing ($\mathbf{t}^{\boldsymbol{\alpha}} := t_1^{\alpha_1} \cdots t_d^{\alpha_d}$).

It is easy to see that in this definition we can always assume $\mathbf{0} < \boldsymbol{\alpha} < \mathbf{1}$ (i.e., $0 < \alpha_j < 1,\ j = 1, \ldots, d$) without loss of generality.

Definition 3. Let $\phi(\mathbf{t})$ be a non-negative function defined on $\mathbf{R}_+^d$ satisfying (1), (2) in Definition 1. We say that $\phi(\mathbf{t}) \in S_l^*$ if there exist a $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_d)$ such that $\mathbf{0} < \boldsymbol{\gamma} < l \cdot \mathbf{1}$ (i.e., $0 < \gamma_j < l,\ j = 1, \ldots, d$) and a constant $C > 0$ such that for any two points $\mathbf{0} < \mathbf{t} \leq \boldsymbol{\tau}$ it always holds that

$$\phi(\mathbf{t}) \cdot \mathbf{t}^{\boldsymbol{\gamma} - l \cdot \mathbf{1}} \geq C\phi(\boldsymbol{\tau}) \cdot \boldsymbol{\tau}^{\boldsymbol{\gamma} - l \cdot \mathbf{1}}$$

(i.e., $\phi(\mathbf{t}) \cdot \mathbf{t}^{\boldsymbol{\gamma} - l \cdot \mathbf{1}}$ is almost decreasing).

Define $\Psi_l^* = \Psi^* = \Psi_l^* \cap S^* \cap S_l^*$. A typical example of a function of type Ψ_l^* is $\Omega(\mathbf{t}) = \mathbf{t}^{\mathbf{r}} := t_1^{r_1} \ldots t_d^{r_d}$, $\mathbf{r} = (r_1, \ldots, r_d) > 0$. The generalized Besov spaces $B_{q,\theta}^{\Omega}(\mathbf{T}^d)$ are defined as follows. Let $e_d := \{1, \ldots, d\}$, $e \subset e_d$. If $e = \{j_1, \ldots, j_m\}$, $j_1 < j_2 < \cdots < j_m$, then we write $\mathbf{t}^e := (t_{j_1}, \ldots, t_{j_m})$, $(\mathbf{t}^e, 1^{\widehat{e}}) := (\bar{t}_1, \ldots, \bar{t}_d)$, where $\bar{t}_i = t_i$ for $i \in e$, $\bar{t}_i = 1$ for $i \in \widehat{e} = e_d \setminus e$.

Definition 4. For $\Omega(\mathbf{t}) \in \Psi_l^*$, we write $f \in B_{q,\theta}^{\Omega}(\mathbf{T}^d)$ if it satisfies
(1) $f \in L_q^0(\mathbf{T}^d)$;
(2) for any non-empty $e \subset e_d$,

$$\left\{\int_0^{2\pi} \cdots \int_0^{2\pi} \left(\frac{\Omega^{l^e}(f, \mathbf{t}^e)_q}{\Omega(\mathbf{t}^e, 1^{\widehat{e}})}\right)^\theta \prod_{j\in e} \frac{\mathrm{d}t_j}{t_j}\right\}^{1/\theta} < \infty, \quad 1 \le \theta < \infty,$$

and

$$\sup_{\mathbf{t}^e > \mathbf{0}} \frac{\Omega^{l^e}(f, \mathbf{t}^e)_q}{\Omega(\mathbf{t}^e, 1^{\widehat{e}})} < \infty, \quad \theta = \infty,$$

where

$$\Omega^{l^e}(f, \mathbf{t}^e)_q := \sup_{|\mathbf{h}^e| \le \mathbf{t}^e} \|\Delta_{\mathbf{h}^e}^{l^e}(f, \mathbf{x})\|_q, \quad \mathbf{h}^e := (h_{j_1}, \dots, h_{j_m}),$$

$$\Delta_{\mathbf{h}^e}^{l^e}(f, \mathbf{x}) = \Delta_{h_{j_m}, j_m}^{l}(\Delta_{h_{j_{m-1}}, j_{m-1}}^{l} \dots \Delta_{h_{j_1}, j_1}^{l} f(\dots, x_{j_1}, \dots, x_{j_m}, \dots) \dots).$$

We define

$$\|f\|_{B_{q,\theta}^{\Omega}(\mathbf{T}^d)} := \|f\|_q + \sum_{e\subset e_d}\left\{\int_0^{2\pi} \cdots \int_0^{2\pi} \left(\frac{\Omega^{l^e}(f, \mathbf{t}^e)_q}{\Omega(\mathbf{t}^e, 1^{\widehat{e}})}\right)^\theta \prod_{j\in e} \frac{\mathrm{d}t_j}{t_j}\right\}^{1/\theta}, \quad 1 \le \theta < \infty,$$

and

$$\|f\|_{B_{q,\theta}^{\Omega}(\mathbf{T}^d)} := \|f\|_{H_q^{\Omega}(\mathbf{T}^d)} := \|f\|_q + \sum_{e\subset e_d} \sup_{\mathbf{t}^e > \mathbf{0}} \frac{\Omega^{l^e}(f, \mathbf{t}^e)_q}{\Omega(\mathbf{t}^e, 1^{\widehat{e}})}, \quad \theta = \infty.$$

Moreover, by the result in [13], for $\Omega(\mathbf{t}) \in \Psi_l^*$, $1 < q < \infty$ and $f \in B_{q,\theta}^{\Omega}(\mathbf{T}^d)$, we have the following equivalent norms:

$$\|f\|_{B_{q,\theta}^{\Omega}(\mathbf{T}^d)} \asymp \left\{\sum_{\mathbf{s}>0} \|\delta_{\mathbf{s}}(f)\|_q^\theta \Omega(2^{-\mathbf{s}})^{-\theta}\right\}^{1/\theta}, \quad 1 \le \theta < \infty,$$

and

$$\|f\|_{B_{q,\infty}^{\Omega}(\mathbf{T}^d)} \asymp \sup_{\mathbf{s}>0} \frac{\|\delta_{\mathbf{s}}(f)\|_q}{\Omega(2^{-\mathbf{s}})}, \quad \theta = \infty.$$

It is not difficult to verify that the generalized Besov spaces $B_{q,\theta}^{\Omega}(\mathbf{T}^d)$ with the above norms are complete. In this paper, we mainly consider the case $\Omega(\mathbf{t}) = \omega(t_1 \cdot \dots \cdot t_d)$ where $\omega(t) \in \Psi_l^*$ (i.e., a univariate function) for some $0 < \alpha < 1$. Denote by $MB_{q,\theta}^{\Omega}(\mathbf{T}^d)$ the unit ball of space $B_{q,\theta}^{\Omega}(\mathbf{T}^d)$.

For convenience, we will suppress the domain $\mathbf{T}^d$ in the notation below.

3. Main results and proofs

In this section, we will give the main results of this paper and their proofs. For this, we first prove some auxiliary results. For the formulation and proofs of these lemmas, we mainly follow the ideas and methods given by Temlyakov in [14]. In the following, we always assume $\Omega(\mathbf{t}) = \omega(t_1 \cdot \cdots \cdot t_d)$, where $\omega(t) \in \Psi_l^*$ (i.e., a univariate function) for some $0 < \alpha < 1$ and $l \geq 1$ is a fixed positive integer.

Lemma 1. *For a fixed real number a and $\omega(t) \in \Psi_l^*$ for some $0 < \alpha < 1$, define*

$$h_n(\mathbf{s}) := \omega(2^{-n})2^{-n/2+a(\|\mathbf{s}\|_1-n)}n^{-(d-1)/\theta}$$

and for $f \in MB^{\Omega}_{q,\theta}$, $q \leq \theta \leq \infty$, consider the sets

$$A(f, n, a) := \{I : |f_I| \geq h_n(\mathbf{s}), \text{if } I \in D^{\boldsymbol{\varepsilon}}_{\mathbf{s}}\}, \quad n = 1, 2, \ldots.$$

Then if $\alpha > 1/q - 1/2 - a$ we have

$$\sharp A(f, n, a) \ll 2^n n^{d-1}$$

with a constant independent of n and f.

Proof. For convenience, we will omit $\boldsymbol{\varepsilon}$ in the notation $\delta^{\boldsymbol{\varepsilon}}_{\mathbf{s}}(f)$, $D^{\boldsymbol{\varepsilon}}_{\mathbf{s}}$, $N^{\boldsymbol{\varepsilon}}_{\mathbf{s}}$ (see below) meaning that we are estimating a quantity $\delta^{\boldsymbol{\varepsilon}}_{\mathbf{s}}(f)$ or $N^{\boldsymbol{\varepsilon}}_{\mathbf{s}}$ for a fixed $\boldsymbol{\varepsilon}$, and all the estimations that we are going to do are the same for all $\boldsymbol{\varepsilon}$.

Using the following two properties of the system $\{U_I\}$:

$$\|\delta^{\boldsymbol{\varepsilon}}_{\mathbf{s}}(f)\|_q^q \asymp \sum_{I \in D^{\boldsymbol{\varepsilon}}_{\mathbf{s}}} \|f_I U_I\|_q^q, \quad 1 < q < \infty, \tag{11}$$

$$\|U_I\|_q \asymp 2^{\|\mathbf{s}\|_1(1/2-1/q)}, \quad I \in D^{\boldsymbol{\varepsilon}}_{\mathbf{s}}, \tag{12}$$

we get

$$\sum_{I \in D^{\boldsymbol{\varepsilon}}_{\mathbf{s}}} |f_I|^q \ll 2^{-\|\mathbf{s}\|_1(q/2-1)}\|\delta^{\boldsymbol{\varepsilon}}_{\mathbf{s}}(f)\|_q^q. \tag{13}$$

Define $N^{\boldsymbol{\varepsilon}}_{\mathbf{s}} := \sharp(A(f, n, a) \cap D^{\boldsymbol{\varepsilon}}_{\mathbf{s}})$. Then (13) implies

$$N^{\boldsymbol{\varepsilon}}_{\mathbf{s}} h_n(\mathbf{s})^q \ll 2^{-\|\mathbf{s}\|_1(q/2-1)}\|\delta^{\boldsymbol{\varepsilon}}_{\mathbf{s}}(f)\|_q^q$$

and

$$\begin{aligned} \sum_{\|\mathbf{s}\|_1=l} N^{\boldsymbol{\varepsilon}}_{\mathbf{s}} &\ll \sum_{\|\mathbf{s}\|_1=l} h_n(\mathbf{s})^{-q}2^{-}\|\mathbf{s}\|_1(q/2-1)\|\delta^{\boldsymbol{\varepsilon}}_{\mathbf{s}}(f)\|_q^q \\ &\ll \sum_{\|\mathbf{s}\|_1=l} h_n(\mathbf{s})^{-q}2^{-}\|\mathbf{s}\|_1(q/2-1)\|\delta_{\mathbf{s}}(f)\|_q^q. \end{aligned} \tag{14}$$

It is known that for $f \in MB^{\Omega}_{q,\theta}$, we have

$$\|f\|_{B^{\Omega}_{q,\theta}} \asymp \left\{\sum_{\mathbf{s}>0} \|\delta_{\mathbf{s}}(f)\|_q^{\theta}\Omega(2^{-\mathbf{s}})^{-\theta}\right\}^{1/\theta} \leq C, \quad 1 \leq \theta < \infty, \tag{15}$$

$$\|f\|_{B^{\Omega}_{q,\infty}} \asymp \sup_{\mathbf{s}>0} \frac{\|\delta_{\mathbf{s}}(f)\|_q}{\Omega(2^{-\mathbf{s}})} \leq C, \quad \theta = \infty,$$

where the constant C is independent of f. By the Hölder inequality and (15), it follows from (14) that

$$\sum_{\|\mathbf{s}\|_1=l} N_{\mathbf{s}}^{\boldsymbol{\varepsilon}} \ll \omega(2^{-n})^{-q} 2^{(1/2+a)nq} 2^{-l(aq+q/2-1)} n^{(d-1)q/\theta} \times \left(\sum_{\|\mathbf{s}\|_1=l} \|\delta_{\mathbf{s}}(f)\|_q^{\theta}\right)^{q/\theta} \left(\sum_{\|\mathbf{s}\|_1=l} 1\right)^{1-q/\theta}$$
$$\ll \omega(2^{-n})^{-q} 2^{(1/2+a)nq} 2^{-l(aq+q/2-1)} n^{(d-1)q/\theta} \omega(2^{-l})^q l^{(d-1)(1-q/\theta)}.$$

Using the almost increasing property of $\omega(t)/t^{\alpha}$ and the assumption $\alpha > 1/q - 1/2 - a$, we get

$$\sum_{\boldsymbol{\varepsilon}} \sum_{l \ge n} \sum_{\|\mathbf{s}\|_1=l} N_{\mathbf{s}}^{\boldsymbol{\varepsilon}} \ll \omega(2^{-n})^{-q} 2^{(1/2+a)nq} n^{(d-1)q/\theta} \sum_{\boldsymbol{\varepsilon}} \sum_{l \ge n} \frac{\omega(2^{-l})^q}{2^{-\alpha l q}} 2^{-l(\alpha q + aq + q/2 - 1)} l^{(d-1)(1-q/\theta)}$$
$$\ll 2^{(1/2+a+\alpha)nq} n^{(d-1)q/\theta} \sum_{\boldsymbol{\varepsilon}} \sum_{l \ge n} 2^{-l(\alpha q + aq + q/2 - 1)} l^{(d-1)(1-q/\theta)}$$
$$\ll 2^n n^{d-1}. \tag{16}$$

It remains to remark that for $\|\mathbf{s}\|_1 < n$ we have the following trivial estimates:

$$\sum_{\boldsymbol{\varepsilon}} \sum_{\|\mathbf{s}\|_1<n} N_{\mathbf{s}}^{\boldsymbol{\varepsilon}} \le \sum_{\boldsymbol{\varepsilon}} \sum_{\|\mathbf{s}\|_1<n} \sharp D_{\mathbf{s}}^{\boldsymbol{\varepsilon}} \ll 2^n n^{d-1}. \tag{17}$$

Combining (16) and (17) we obtain the required estimate. □

For $\theta = \infty$, we can finish the proof along the same lines as above. Here we omit the details.

Lemma 2. *Let $h_n(\mathbf{s})$, $A(f,n,a)$ be from Lemma 1 and $\omega(t) \in \Psi_l^*$ for some $0 < \alpha < 1$ and let $a > -1/2$. For each n define*

$$g_n(f) := \sum_{I \in A(f,n,a)} f_I U_I, \qquad f^n := f - g_n(f).$$

Then for any $f \in MB_{q,\theta}^{\Omega}$, and $p \ge 2$ satisfying $1 < q \le p < \infty$, we have for $\alpha > (a+1/2)(p/q-1)$, $\max(2,q) \le \theta \le \infty$,

$$\|f^n\|_p \ll \omega(2^{-n}) n^{(d-1)(1/2-1/\theta)}$$

with a constant independent of n and f.

Proof. By the relation $MB_{q_1,\theta}^{\Omega} \subset MB_{q_2,\theta}^{\Omega}$, for $q_1 > q_2$, it is sufficient to prove Lemma 2 for $1 < q \le 2$.

For $2 \le p < \infty$, we have, by a corollary to the Littlewood–Paley inequalities,

$$\|f^n\|_p^2 \ll \sum_{\boldsymbol{\varepsilon}} \left(\sum_{\mathbf{s}} \|\delta_{\mathbf{s}}^{\boldsymbol{\varepsilon}}(f^n)\|_p^2\right)$$
$$= \sum_{\boldsymbol{\varepsilon}} \left(\sum_{\|\mathbf{s}\|_1<n} \|\delta_{\mathbf{s}}^{\boldsymbol{\varepsilon}}(f^n)\|_p^2 + \sum_{\|\mathbf{s}\|_1 \ge n} \|\delta_{\mathbf{s}}^{\boldsymbol{\varepsilon}}(f^n)\|_p^2\right) := \sum_{\boldsymbol{\varepsilon}} (\Sigma' + \Sigma'').$$

We first estimate Σ'. By the definition of $A(f, n, a)$ we have, for all I,

$$|f_I^n| < h_n(\mathbf{s}), \quad I \in D_{\mathbf{s}}^{\boldsymbol{\varepsilon}}.$$

Therefore,

$$\|\delta_{\mathbf{s}}^{\boldsymbol{\varepsilon}}(f^n)\|_p^p \ll h_n(\mathbf{s})^p \sum_{I \in D_{\mathbf{s}}^{\boldsymbol{\varepsilon}}} \|U_I\|_p^p \ll \omega(2^{-n})^p 2^{-np(1/2+a)} n^{-(d-1)p/\theta} 2^{\|\mathbf{s}\|_1 p(1/2+a)}$$

and

$$\sum_{\|\mathbf{s}\|_1 < n} \|\delta_{\mathbf{s}}^{\boldsymbol{\varepsilon}}(f^n)\|_p^2 \ll \omega(2^{-n})^2 2^{-n(1+2a)} n^{-(d-1)2/\theta} \sum_{\|\mathbf{s}\|_1 < n} 2^{\|\mathbf{s}\|_1 (1+2a)}$$
$$\ll \omega(2^{-n})^2 n^{(d-1)(1-2/\theta)}. \tag{18}$$

We proceed to estimate Σ'' now. We have

$$\begin{aligned}
\|\delta_{\mathbf{s}}^{\boldsymbol{\varepsilon}}(f^n)\|_p^p &\ll \sum_{I \in D_{\mathbf{s}}^{\boldsymbol{\varepsilon}}} \|f_I^n U_I\|_p^p \\
&\ll \left(h_n(\mathbf{s}) 2^{\|\mathbf{s}\|_1(1/2-1/p)}\right)^{p-q} \sum_{I \in D_{\mathbf{s}}^{\boldsymbol{\varepsilon}}} \|f_I^n U_I\|_p^q \\
&\ll \left(h_n(\mathbf{s}) 2^{\|\mathbf{s}\|_1(1/2-1/p)}\right)^{p-q} \sum_{I \in D_{\mathbf{s}}^{\boldsymbol{\varepsilon}}} \|f_I^n U_I\|_q^q 2^{\|\mathbf{s}\|_1(1/q-1/p)q} \\
&\ll \left(h_n(\mathbf{s}) 2^{\|\mathbf{s}\|_1(1/2-1/p)}\right)^{p-q} \sum_{I \in D_{\mathbf{s}}^{\boldsymbol{\varepsilon}}} \|f_I U_I\|_q^q 2^{\|\mathbf{s}\|_1(1/q-1/p)q}.
\end{aligned}$$

Like in the treatment of Lemma 1, by virtue of the Hölder inequality and $f \in MB_{q,\theta}^{\Omega}$, we get

$$\begin{aligned}
\sum_{\|\mathbf{s}\|_1 = l} \|\delta_{\mathbf{s}}^{\boldsymbol{\varepsilon}}(f^n)\|_p^2 &\ll \omega(2^{-n})^{2(p-q)/p} 2^{-2n(1/2+a)(p-q)/p} n^{-2(d-1)(p-q)/p\theta} 2^{2l(p-q)(1/2+a)/p} \\
&\quad \times \sum_{\|\mathbf{s}\|_1 = l} \|\delta_{\mathbf{s}}(f)\|_q^{2q/p} \\
&\ll \omega(2^{-n})^{2(p-q)/p} 2^{-2n(1/2+a)(p-q)/p} n^{-2(d-1)(p-q)/p\theta} \\
&\quad \times 2^{2l(p-q)(1/2+a)/p} \omega(2^{-l})^{2q/p} l^{(d-1)(1-2q/p\theta)}.
\end{aligned}$$

Using the almost increasing property of $\omega(t)/t^{\alpha}$ and the assumption $\alpha > (a + 1/2)(p/q - 1)$, we get

$$\begin{aligned}
\sum_{l \geq n} \sum_{\|\mathbf{s}\|_1 = l} \|\delta_{\mathbf{s}}^{\boldsymbol{\varepsilon}}(f^n)\|_p^2 &\ll \omega(2^{-n})^2 2^{n[2\alpha q/p - (1/2+a)(2-2q/p)]} n^{-2(d-1)(p-q)/p\theta} \\
&\quad \times \sum_{l \geq n} 2^{-l[2\alpha q/p - (2-2q/p)(1/2+a)]} l^{(d-1)(1-2q/p\theta)} \\
&\ll \omega(2^{-n})^2 n^{(d-1)(1-2/\theta)}.
\end{aligned} \tag{19}$$

Combining (18) and (19) we finish the estimates for $\max(2, q) \leq \theta < \infty$.

For $\theta = \infty$, we can compute the estimate along the same lines as above with slight modification. Thus we finish the proof of Lemma 2. □

It is clear from the proof of Lemma 2 that the following statement holds:

Lemma 3. *Let $h_n(\mathbf{s})$ be from Lemma 1 and $\omega(t) \in \Psi_l^*$ for some $0 < \alpha < 1$ and let $a > -1/2$. Assume that a function f satisfies the restrictions*

$$\left(\sum_{\|\mathbf{s}\|_1 = l} \|\delta_{\mathbf{s}}^{\boldsymbol{\varepsilon}}(f)\|_q^\theta \right)^{1/\theta} \ll \omega(2^{-l}), \quad 1 < \theta < \infty,$$

$$\|\delta_{\mathbf{s}}^{\boldsymbol{\varepsilon}}(f)\|_q \ll \omega(2^{-l}), \quad \theta = \infty,$$

$$|f_I| \ll h_n(\mathbf{s}), \quad I \in D_{\mathbf{s}}^{\boldsymbol{\varepsilon}},$$

with constants independent of f, n and $\mathbf{s}$ for $1 < q < \infty$. Then for $\max(2, q) \le p < \infty$ and $\alpha > (a + 1/2)(p/q - 1)$, $\max(2, q) \le \theta \le \infty$ we have

$$\|f\|_p \ll \omega(2^{-n}) n^{(d-1)(1/2-1/\theta)}$$

with a constant independent of n and f.

Consider the following greedy-type algorithm $G^{c,a}$. Take a real number a and rearrange the sequence $|f_I|\,|I|^a$ in the decreasing order

$$|f_{I^1}|\,|I^1|^a \ge |f_{I^2}|\,|I^2|^a \ge \cdots.$$

Define

$$G_m^{c,a}(f, U^d) := \sum_{k=1}^m f_{I^k} U_{I^k}.$$

Theorem 1. *Let $\Omega(\mathbf{t}) = \omega(t_1 \cdot \cdots \cdot t_d)$, where $\omega(t) \in \Psi_l^*$ for some $0 < \alpha < 1$ and let $1 < q < \infty$, $\max(2, q) \le p < \infty$ and $\max(2, q) \le \theta \le \infty$. Then for any $a > -1/2$, $\alpha > \max\{(a + 1/2)(p/q - 1), 1/q - a - 1/2\}$ and any natural numbers m and n such that $m \asymp 2^n n^{d-1}$, we have*

$$\sup_{f \in MB_{q,\theta}^\Omega} \|f - G_m^{c,a}(f, U^d)\|_p \asymp \sigma_m(MB_{q,\theta}^\Omega, U^d)_p \asymp \omega(2^{-n}) n^{(d-1)(1/2-1/\theta)}.$$

Proof. Let m be given. We can choose the biggest n satisfying

$$\sup_{f \in MB_{q,\theta}^\Omega} \sharp A(f, n, a) \le m.$$

For $f \in MB_{q,\theta}^\Omega$, let

$$g := f - G_m^{c,a}(f, U^d).$$

It is not difficult to see that g satisfies

$$|g_I| \le h_n(\mathbf{s}), \quad I \in D_{\mathbf{s}}^{\boldsymbol{\varepsilon}},$$

and

$$\left(\sum_{\|\mathbf{s}\|_1 = l} \|\delta_{\mathbf{s}}^{\boldsymbol{\varepsilon}}(g)\|_q^\theta \right)^{1/\theta} \ll \omega(2^{-l}), \quad \text{or} \quad \|\delta_{\mathbf{s}}^{\boldsymbol{\varepsilon}}(g)\|_q \ll \omega(2^{-l}),$$

with a constant independent of $\mathbf{s}$ and g. Applying Lemma 3 to g we get

$$\|g\|_p \ll \omega(2^{-n})n^{(d-1)(1/2-1/\theta)},$$

which proves the upper estimate in Theorem 1

$$\sup_{f\in MB^{\Omega}_{q,\theta}} \|f - G^{c,a}_m(f, U^d)\|_p \ll \omega(2^{-n})n^{(d-1)(1/2-1/\theta)}.$$

The lower estimate

$$\sigma_m(MB^{\Omega}_{q,\theta}, U^d)_p \gg \omega(2^{-n})n^{(d-1)(1/2-1/\theta)}$$

follows from Theorem 4. The proof of Theorem 1 is complete. □

Consider now the L_b-greedy algorithm $G^b(\cdot, U^d)$. Take a number $1 \le b \le \infty$ and rearrange the sequence $\{\|f_I U_I\|_b\}$ in decreasing order:

$$\|f_{I_1} U_{I_1}\|_b \ge \|f_{I_2} U_{I_2}\|_b \ge \cdots.$$

Define

$$G^b_m(f, U^d) := \sum_{k=1}^{m} f_{I_k} U_{I_k}.$$

It is clear from the relation

$$\|f_I U_I\|_b \asymp |f_I|\,|I|^{1/b-1/2}$$

that the algorithm G^b and $G^{c,a}$ with $a = 1/b - 1/2$ are closely connected. The following theorem can be proved similarly to Theorem 1.

Theorem 2. *Let $\Omega(\mathbf{t}) = \omega(t_1 \cdot \cdots \cdot t_d)$, where $\omega(t) \in \Psi^*_l$ for some $0 < \alpha < 1$ and let $1 < q < \infty$ and $\max(2, q) \le p < \infty$, $\max(2, q) \le \theta \le \infty$. Then for any $1 < b < \infty$, $\alpha > \max\{(p/q - 1)/b, 1/q - 1/b\}$ and any natural numbers m and n such that $m \asymp 2^n n^{d-1}$, we have*

$$\sup_{f\in MB^{\Omega}_{q,\theta}} \|f - G^b_m(f, U^d)\|_p \asymp \sigma_m(MB^{\Omega}_{q,\theta}, U^d)_p \asymp \omega(2^{-n})n^{(d-1)(1/2-1/\theta)}.$$

We formulate now the corollary of Theorem 2 in the most interesting case $b = p$.

Theorem 3. *Let $\Omega(\mathbf{t}) = \omega(t_1 \cdot \cdots \cdot t_d)$, where $\omega(t) \in \Psi^*_l$ for some $0 < \alpha < 1$ and let $1 < q, p < \infty$. Then for all $\alpha > \alpha(q, p)$, $\max(2, q) \le \theta \le \infty$ and any natural numbers m and n such that $m \asymp 2^n n^{d-1}$, we have*

$$\sup_{f\in MB^{\Omega}_{q,\theta}} \|f - G^p_m(f, U^d)\|_p \asymp \sigma_m(MB^{\Omega}_{q,\theta}, U^d)_p \asymp \omega(2^{-n})n^{(d-1)(1/2-1/\theta)},$$

with

$$\alpha(q, p) := \begin{cases} (1/q - 1/p)_+, & \text{for } p \ge 2 \\ (\max(2/q, 2/p) - 1)/p, & \text{otherwise.} \end{cases}$$

Proof. We first prove the upper estimates. Consider the case $2 \le p < \infty$. If $1 < q \le p$ we use Theorem 2 with $b = p$ and get a restriction $\alpha > 1/q - 1/p$. If $p < q < \infty$ we use the inequality

$$\sup_{f \in MB^{\Omega}_{q,\theta}} \|f - G^p_m(f, U^d)\|_p \le \sup_{f \in MB^{\Omega}_{p,\theta}} \|f - G^p_m(f, U^d)\|_p \tag{20}$$

and reduce this case to the case $q = p$ which has already been considered. It remains to consider the case $1 < p < 2$. If $1 < q \le p$ we use Theorem 2 with $p = 2$ and $b = p$ and get

$$\sup_{f \in MB^{\Omega}_{q,\theta}} \|f - G^p_m(f, U^d)\|_p \le \sup_{f \in MB^{\Omega}_{q,\theta}} \|f - G^p_m(f, U^d)\|_2 \ll \omega(2^{-n}) n^{(d-1)(1/2-1/\theta)}$$

provided $\alpha > (2/q - 1)/p$. If $p < q < \infty$, using the following inequality and Theorem 2 with $p = 2$ and $b = p$, we obtain

$$\begin{aligned} \sup_{f \in MB^{\Omega}_{q,\theta}} \|f - G^p_m(f, U^d)\|_p &\le \sup_{f \in MB^{\Omega}_{p,\theta}} \|f - G^p_m(f, U^d)\|_p \\ &\le \sup_{f \in MB^{\Omega}_{p,\theta}} \|f - G^p_m(f, U^d)\|_2 \\ &\ll \omega(2^{-n}) n^{(d-1)(1/2-1/\theta)}. \end{aligned}$$

In this case, we get a restriction $\alpha > (2/p - 1)/p$.

The lower estimates follow from Theorem 4. The proof of Theorem 3 is complete. □

In what follows, we will give the lower estimates of the best m-term approximation $\sigma_m(MB^{\Omega}_{q,\theta}, U^d)_p$ with the orthogonal basis U^d and $\sigma_m(MB^{\Omega}_{q,\theta}, \mathbf{O})_p$ with regard to all orthogonal dictionaries, i.e., $\mathbf{O}$. To deal with the lower estimates, we will adopt different methods. For the lower estimates of $\sigma_m(MB^{\Omega}_{q,\theta}, U^d)_p$, i.e., Theorem 4 below, we mainly follow the idea given by Wang in [17], while for the lower estimates of $\sigma_m(MB^{\Omega}_{q,\theta}, \mathbf{O})_p$, i.e., Theorem 5, we will use the idea given by Temlyakov in [14] and the following auxiliary lemma.

Lemma 4 ([5]). *There exists an absolute constant $c_0 > 0$ such that for any orthonormal basis Φ and any N-dimensional cube*

$$B_N(\Psi) := \left\{ \sum_{j=1}^{N} a_j \psi_j,\ |a_j| \le 1,\ j = 1, \dots, N;\ \Psi := \{\psi_j\}_{j=1}^N \text{ an orthonormal system} \right\}$$

we have

$$\sigma_m(B_N, \Phi)_2 \ge 3N^{1/2}/4$$

if $m \le c_0 N$.

Theorem 4. *Let $\Omega(\mathbf{t}) = \omega(t_1 \cdot \dots \cdot t_d)$, where $\omega(t) \in \Psi^*_l$ for some $0 < \alpha < 1$ and let $1 < q, p < \infty$, $2 \le \theta \le \infty$. Then for $\alpha > (1/q - 1/p)_+$ and any natural numbers m and n such that $m \asymp 2^n n^{d-1}$, we have*

$$\sigma_m(MB^{\Omega}_{q,\theta}, U^d)_p \gg \omega(2^{-n}) n^{(d-1)(1/2-1/\theta)}.$$

Proof. By the monotonicity of the L_p-norm, it is sufficient to estimate the lower bounds for $1 < p \le 2$.

For any natural number m, we can choose a natural number n such that

$$\sum_{\|\mathbf{s}\|_1=n} \sharp D_{\mathbf{s}}^{+} \geq 2m, \quad \text{and} \quad \sum_{\|\mathbf{s}\|_1=n} \sharp D_{\mathbf{s}}^{+} \asymp 2^n n^{d-1} \asymp m,$$

where $D_{\mathbf{s}}^{+} := \{I : I = I_1 \times \cdots \times I_d,\ I_j \in D_{s_j}^{+},\ j = 1, \ldots, d\}$.

Consider the function

$$\psi_n(\mathbf{x}) = \sum_{\|\mathbf{s}\|_1=n} \sum_{I \in D_{\mathbf{s}}^{+}} U_I(\mathbf{x}).$$

Noting that

$$\left\| \sum_{I \in D_{\mathbf{s}}^{+}} U_I \right\|_q \asymp \left(\sum_{I \in D_{\mathbf{s}}^{+}} \|U_I\|_q^q \right)^{1/q} \asymp 2^{\|\mathbf{s}\|_1/2},$$

we have

$$\begin{aligned} \|\psi_n\|_{B_{q,\theta}^{\Omega}} &\asymp \left(\sum_{\|\mathbf{s}\|_1=n} \left\| \sum_{I \in D_{\mathbf{s}}^{+}} U_I \right\|_q^{\theta} \Omega(2^{-\mathbf{s}})^{-\theta} \right)^{1/\theta} \\ &\ll \left(\sum_{\|\mathbf{s}\|_1=n} 2^{\|\mathbf{s}\|_1\theta/2} \omega(2^{-n})^{-\theta} \right)^{1/\theta} \\ &\ll \omega(2^{-n})^{-1} 2^{n/2} n^{(d-1)/\theta}. \end{aligned}$$

Therefore, there exists a positive constant $c > 0$ such that

$$c \cdot \omega(2^{-n}) 2^{-n/2} n^{-(d-1)/\theta} \psi_n \in M B_{q,\theta}^{\Omega}. \tag{21}$$

For any set $I_m \subset D^d$ satisfying $\sharp I_m \leq m$, let $M_{\mathbf{s}} := \sharp(D_{\mathbf{s}}^{+} \setminus I_m)$ and then we have

$$\sum_{\|\mathbf{s}\|_1=n} M_{\mathbf{s}} \geq \sum_{\|\mathbf{s}\|_1=n} \sharp D_{\mathbf{s}}^{+} - m \geq 2m - m = m.$$

For any $c_I,\ I \in I_m$, by (10) and (12) we get

$$\begin{aligned} &\left\| \psi_n - \sum_{I \in I_m} c_I U_I \right\|_p \gg \left\| \sum_{\|\mathbf{s}\|_1=n} \sum_{I \in D_{\mathbf{s}}^{+} \setminus I_m} U_I \right\|_p \\ &\gg \left(\sum_{\|\mathbf{s}\|_1=n} \left\| \sum_{I \in D_{\mathbf{s}}^{+} \setminus I_m} U_I \right\|_p^2 \right)^{1/2} \gg \left(\sum_{\|\mathbf{s}\|_1=n} \left(\sum_{I \in D_{\mathbf{s}}^{+} \setminus I_m} \|U_I\|_p^p \right)^{2/p} \right)^{1/2} \\ &\gg \left(\sum_{\|\mathbf{s}\|_1=n} M_{\mathbf{s}}^{2/p} 2^{\|\mathbf{s}\|_1(1-2/p)} \right)^{1/2} \gg 2^{n(1/2-1/p)} \left(\sum_{\|\mathbf{s}\|_1=n} M_{\mathbf{s}}^{2/p} \right)^{1/2}. \end{aligned} \tag{22}$$

By virtue of

$$2^n n^{d-1} \ll \sum_{\|\mathbf{s}\|_1=n} M_{\mathbf{s}} \leq \left(\sum_{\|\mathbf{s}\|_1=n} M_{\mathbf{s}}^{2/p} \right)^{p/2} \left(\sum_{\|\mathbf{s}\|_1=n} 1 \right)^{1-p/2}$$

$$\ll n^{(d-1)(1-p/2)} \left(\sum_{\|\mathbf{s}\|_1 = n} M_{\mathbf{s}}^{2/p} \right)^{p/2},$$

it follows from (22) that

$$\sigma_m(\psi_n, U^d)_p = \inf_{I_m, c_I} \left\| \psi_n - \sum_{I \in I_m} c_I U_I \right\|_p \gg 2^{n/2} n^{(d-1)/2}. \tag{23}$$

From (21) and (23), we obtain

$$\sigma_m(MB^{\Omega}_{q,\theta}, U^d)_p \gg \omega(2^{-n}) 2^{-n/2} n^{-(d-1)/\theta} \cdot \sigma_m(\psi_n, U^d)_p \gg \omega(2^{-n}) n^{(d-1)(1/2-1/\theta)}.$$

The proof of Theorem 4 is complete. □

Theorem 5. *Let $\Omega(\mathbf{t}) = \omega(t_1 \cdot \cdots \cdot t_d)$, where $\omega(t) \in \Psi_l^*$ for some $0 < \alpha < 1$ and $2 \le \theta \le \infty$. For any orthonormal basis Φ, we have for $\alpha > (1/q - 1/2)_+$ and any natural numbers m and n such that $m \asymp 2^n n^{d-1}$,*

$$\sigma_m(MB^{\Omega}_{q,\theta}, \Phi)_2 \gg \omega(2^{-n}) n^{(d-1)(1/2-1/\theta)}, \quad 1 < q < \infty,$$

and

$$\sigma_m(MB^{\Omega}_{q,\theta}, \mathbf{O})_2 \gg \omega(2^{-n}) n^{(d-1)(1/2-1/\theta)}, \quad 1 < q < \infty.$$

Proof. For $1 < q < \infty$, let m be given. Define

$$D(n) = \bigcup_{\|\mathbf{s}\|_1 = n} D_{\mathbf{s}}^+$$

and choose a minimal n such that

$$m \le c_0 \sharp D(n), \quad \text{and} \quad m \asymp 2^n n^{d-1}.$$

We set $N := \sharp D(n)$ and choose the system $U(n) := \{U_I\}_{I \in D(n)}$. Then for any $f \in B_N(U(n))$ we have

$$\|\delta_{\mathbf{s}}(f)\|_q^q \asymp \sum_{I \in D_{\mathbf{s}}^+} \|f_I U_I\|_q^q \le \sum_{I \in D_{\mathbf{s}}^+} \|U_I\|_q^q \ll 2^{nq/2}, \tag{24}$$

and

$$\|f\|_{B^{\Omega}_{q,\theta}} \asymp \left(\sum_{\|\mathbf{s}\|_1 = n} \|\delta_{\mathbf{s}}(f)\|_q^{\theta} \Omega(2^{-\mathbf{s}})^{-\theta} \right)^{1/\theta}$$
$$\ll \omega(2^{-n})^{-1} 2^{n/2} n^{(d-1)/\theta}. \tag{25}$$

According to (24) and (25) we have for some positive $C(q, d)$,

$$C(q,d) \omega(2^{-n}) 2^{-n/2} n^{-(d-1)/\theta} B_N(U(n)) \subset MB^{\Omega}_{q,\theta}.$$

Therefore, by Lemma 4, we obtain

$$\sigma_m(MB^{\Omega}_{q,\theta}, \Phi)_2 \gg \omega(2^{-n}) n^{(d-1)(1/2-1/\theta)}.$$

Further, we have

$$\sigma_m(MB^{\Omega}_{q,\theta}, \mathbf{O})_2 \gg \omega(2^{-n}) n^{(d-1)(1/2-1/\theta)}.$$

The proof of Theorem 5 is complete. □

Combining Theorems 3 and 5, we can obtain the following result:

Theorem 6. *Let $\Omega(\mathbf{t}) = \omega(t_1 \cdot \cdots \cdot t_d)$, where $\omega(t) \in \Psi_l^*$ for some $0 < \alpha < 1$. For $1 < q < \infty$, $2 \le p < \infty$, $\max(2, q) \le \theta \le \infty$, $\alpha > (1/q - 1/p)_+$, and any natural numbers m and n such that $m \asymp 2^n n^{d-1}$, we have*

$$\sigma_m(MB_{q,\theta}^{\Omega}, \mathbf{O})_p \asymp \omega(2^{-n}) n^{(d-1)(1/2-1/\theta)}.$$

From Theorems 3 and 6, we can see that the orthogonal basis U^d is nearly optimal among orthogonal dictionaries. Moreover, it is established that for these classes the near-best m-term approximation, with regard to U^d, can be achieved by simple greedy-type algorithms. The best m-term approximation considered in this paper is a nonlinear analog of the Kolmogorov n-width. In [13] Sun and Wang obtained the following results on the Kolmogorov n-width of generalized Besov classes $MB_{q,\theta}^{\Omega}$:

Suppose $\Omega(\mathbf{t}) = \omega(t_1 \cdot \cdots \cdot t_d)$, $\omega(t) \in \Psi_l^*$, $\omega(t)t^{-\alpha}$ is almost increasing; then

(1) if $1 < q < 2 \le p < \infty$, $\alpha > 1/q$, $1 \le \theta \le \infty$,

$$d_N(MB_{q,\theta}^{\Omega}, L_p) \asymp \omega(2^{-n}) 2^{n(1/q-1/2)} n^{(d-1)(1/2-1/\theta)_+};$$

(2) if $2 \le q \le p < \infty$, $\alpha > 1/2$, $1 \le \theta \le \infty$, then

$$d_N(MB_{q,\theta}^{\Omega}, L_p) \asymp \omega(2^{-n}) n^{(d-1)(1/2-1/\theta)_+},$$

where $N \asymp 2^n n^{(d-1)}$.

From the above results and the results obtained in this paper, we can see that the best m-term approximation is better in approximate order than the Kolmogorov n-width for the values of parameters $1 < q < 2 \le p < \infty$ and $2 \le \theta \le \infty$. In addition, according to the definitions of Kolmogorov width and linear width, we know that Kolmogorov width gives a lower bound for linear width. Therefore, the nonlinear approximation, i.e., best m-term approximation is also superior to the linear approximation in approximate order for some certain values of parameters.

Acknowledgments

The author is extremely grateful to the referees for their excellent suggestions, comments and remarks which greatly helped to improve the presentation of this paper and make it more readable.

References

[1] T.I. Amanov, Representation and imbedding theorems for the function spaces $S_{p,\theta}^{r}B(\mathbf{R}_n)$, $S_{p,\theta}^{*r}B(0 \le x_j \le 2\pi, j = 1, \ldots, n)$, Trudy Mat. Inst. Akad. Nauk SSSR 77 (1965) 5–34.

[2] R.A. Devore, Nonlinear approximation, Acta Numer. (1998) 51–150.

[3] D. Dung, Continuous algorithms in n-term approximation and non-linear widths, J. Approx. Theory 102 (2000) 217–242.

[4] O.V. Fedunyk, Linear widths of the classes $B_{p,\theta}^{\Omega}$ of periodic functions of many variables in the space L_q, Ukrainian Math. J. 58 (2006) 103–117.

[5] B.S. Kashin, On approximation properties of complete orthonormal systems, Tr. Mat. Inst. Steklova 172 (1985) 187–191. English transl. in Proc. Steklov Inst. Math. (1987) 207–211.

[6] B.S. Kashin, V.N. Temlyakov, On best m-term approximation and the entropy of sets in the space L^1, Mat. Zametki 56 (1994) 57–86. English transl. in Math. Notes. 56(1994) 1137–1157.

[7] S.M. Nikolskii, Approximation of Functions of Several Variables and Imbeddings Theorems, Springer-Verlag, Berlin, 1975.
[8] N.N. Pustovoitov, Representation and approximation of multivariate periodic functions with a given mixed modulus of smoothness, Anal. Math. 20 (1994) 35–48.
[9] A.S. Romanyuk, On estimate of the Kolmogorov widths of the classes $B^r_{p,\theta}$ in the space L_q, Ukrainian Math. J. 53 (2001) 1189–1196.
[10] A.S. Romanyuk, Linear widths of the Besov classes of periodic functions of many variables. II, Ukrainian Math. J. 53 (2001) 965–977.
[11] A.S. Romanyuk, Approximation of classes $B^r_{p,\theta}$ by linear methods and best approximations, Ukrainian Math. J. 54 (2002) 825–838.
[12] S.A. Stasyuk, Best approximations and Kolmogorov and trigonometric widths of the classes $B^{\Omega}_{p,\theta}$ of periodic functions of many variables, Ukrainian Math. J. 56 (2004) 1849–1863.
[13] Y.S. Sun, H.P. Wang, Representation and approximation of multivariate periodic functions with bounded mixed moduli of smoothness, Proc. Steklov Inst. Math. 219 (1997) 350–371.
[14] V.N. Temlyakov, Greedy algorithms with regard to multivariate systems with special structure, Constr. Approx. 16 (2000) 399–425.
[15] V.N. Temlyakov, Universal bases and greedy algorithms for anisotropic function classes, Constr. Approx. 18 (2002) 529–550.
[16] H.P. Wang, Greedy algorithm for functions with low mixed smoothness, J. Approx. Theory 137 (2005) 264–276.
[17] H.P. Wang, Greedy algorithm for Besov classes, Acta Math. Sinica 47 (2004) 1079–1088 (in Chinese).