Journal of Computer Science and Cybernetics, V.35, N.2 (2019), 119–134
DOI 10.15625/1813-9663/35/2/13578

NONLINEAR APPROXIMATIONS OF FUNCTIONS HAVING MIXED SMOOTHNESS

NGUYEN MANH CUONG

Department of Natural Sciences, Hong Duc University
cuongnv.hdu@gmail.com

Abstract. For multivariate Besov-type classes $U^a_{p,\theta}$ of functions having nonuniform mixed smoothness $a \in \mathbb{R}^d_+$, we obtain the asumptotic order of entropy numbers $\epsilon_n(U^a_{p,\theta}, L_q)$ and non-linear widths $\rho_n(U^a_{p,\theta}, L_q)$ defined via pseudo-dimension. We obtain also the asymptotic order of optimal methods of adaptive sampling recovery in L_q-norm of functions in $U^a_{p,\theta}$ by sets of a finite capacity which is measured by their cardinality or pseudo-dimension.

Keywords. Besov-Type Spaces; Linear Sampling Recovery; Nonlinear Adaptive Sampling Recovery.

1. INTRODUCTION

We are interested in nonlinear approximations of multivariate functions having a given mixed smoothness and their optimality in terms of entropy numbers $\epsilon_n(W, L_q)$ and non-linear widths $\rho_n(W, L_q)$ defined via pseudo-dimension. The problem of $\epsilon_n(W, L_q)$ has a long history and there have been many papers devoted to it. We refer the reader to the book [7] for a survey and bibliography therein. The non-linear widths $\rho_n(W, L_q)$ has been introduced in [12, 13] and investigated there for classical Sobolev classes of functions. In [3], Dinh Dũng has investigated optimal non-linear approximations by sets of a finite capacity which is measured by their cardinality or pseudo-dimension, of multivariate periodic functions having uniform Besov mixed smoothness $r > 0$. In the present paper, we extend these results to multivariate Besov-type classes $U^a_{p,\theta}$ of functions having nonuniform mixed smoothness $a \in \mathbb{R}^d_+$ and the problems of entropy numbers $\epsilon_n(U^a_{p,\theta}, L_q)$ and non-linear widths $\rho_n(U^a_{p,\theta}, L_q)$. Moreover, generalizing the results in [1, 4, 5, 6] on adaptive sampling recovery, we obtain the asymptotic order of optimal methods of adaptive sampling recovery of functions in $U^a_{p,\theta}$ by sets of a finite capacity which is measured by their cardinality or pseudo-dimension.

We begin with a setting of the problems. Let $\mathbb{T}^d$ be the d-dimensional torus which is defined as the cross product of d copies of the interval $[0, 2\pi]$ with the identification of the end points. For $0 < q \le \infty$, let $L_q := L_q(\mathbb{T}^d)$ be the quasi-normed space of all functions on $\mathbb{T}^d$ with the integral quasi-norm $\| \cdot \|_q$ for $0 < q < \infty$, and the normed space $C(\mathbb{T}^d)$ of all continuous functions on $\mathbb{T}^d$ with the max-norm $\| \cdot \|_\infty$ for $q = \infty$. Let B and W be subsets in L_q. We approximate the elements in W by B via the deviation of W from B

$$E(W,B,L_q) := \sup_{f\in W}\inf_{\varphi\in B}\|f-\varphi\|_q.$$

Definition 1. Given a family $\mathcal{B}$ of subsets in L_q, we consider the best approximation by $B\in\mathcal{B}$ in terms of the quantity

$$d(W,\mathcal{B},L_q) := \inf_{B\in\mathcal{B}} E(W,B,L_q). \tag{1}$$

If $\mathcal{B}$ in (1) is the family of all subsets B of L_q which satisfy $|B|\le 2^n$, then $d(W,\mathcal{B},L_q)$ is the well known entropy number which is denoted by $\epsilon_n(W,L_q)$. If $\mathcal{B}$ in (1) is the family of all subsets B of L_q such that $\dim_p(B)\le n$, then $d(W,\mathcal{B},L_q)$ is denoted by $\rho_n(W,L_q)$. Here, $|B|$ denotes the cardinality of the finite set B and $\dim_p(B)$ denotes the pseudo-dimension of set B.

The pseudo-dimension of a set B of real-valued functions on a set Ω, is defined as the largest integer n such that there exist points $a^1,a^2,\dots,a^n$ in Ω and $b=(b_1,\dots,b_n)\in\mathbb{R}^n$, such that the cardinality of the set

$$\left\{\ \mathrm{sgn}\ (y) : y = \left(f(a^1)+b_1, f(a^2)+b_2,\dots,f(a^n)+b_n\right), f\in B\right\}$$

is 2^n, where $\mathrm{sgn}\ (t)=1$ for $t>0$, $\mathrm{sgn}\ (t)=-1$ for $t\le 0$, and for $x\in\mathbb{R}^n$,

$$\mathrm{sgn}\ (x) = (\ \mathrm{sgn}\ (x_1),\ \mathrm{sgn}\ (x_2),\dots,\ \mathrm{sgn}\ (x_n)).$$

We are also interested in the problem of adaptive sampling recovery by functions from W, of periodic functions in W. The error of sampling recovery is measured in the quasi-norm of L_q. We define a sampling recovery method with free choice of sample points and recovering function from B as follows. For each $f\in W$ we choose n of sample points $x^1,\dots,x^n$, and a function $g=S_n^B(f)\in B$ to recover f based on the information of sampled values $f(x^1),\dots,f(x^n)$. Then S_n^B is an adaptive recovering method which can be defined as follows.

Denote by I^n the set of subsets ξ in $\mathbb{T}^d$ of cardinality at most n. Let V^n be the set whose elements are collections of real numbers $a_\xi=\{a(x)\}_{x\in\xi}$, $\xi\in I^n$, $a(x)\in\mathbb{R}$. Let I_n be a mapping from W into I^n and P a mapping from V^n into B. Then the pair (I_n,P) generates the mapping S_n^B from W into B by the formula

$$S_n^B(f) := P\left(\{f(x)\}_{x\in I_n(f)}\right). \tag{2}$$

We want to choose a sampling recovery method S_n^B so that the error of this recovery $\|f-S_n^B(f)\|_q$ is as small as possible. Clearly, such an efficient choice should be adaptive to f.

Definition 2. Given a family $\mathcal{B}$ of subsets in L_q, then the error of optimal sampling recovery methods S_n^B with $B\in\mathcal{B}$ is defined by

$$R_n(W,\mathcal{B})_q := \inf_{B\in\mathcal{B}}\inf_{S_n^B}\sup_{f\in W}\|f-S_n^B(f)\|_q. \tag{3}$$

Denote $R_n(W,\mathcal{B})_q$ by $e_n(W)_q$ if $\mathcal{B}$ in (3) is the family of all subsets B in L_q such that $|B| \le 2^n$, and by $r_n(W)_q$ if $\mathcal{B}$ (3) is the family of all subsets B in L_q such that $\dim_p(B) \le n$.

The quantities $e_n(W)_q$ and $r_n(W)_q$ which are similar to $\epsilon_n(W)_q$ and $\rho_n(W)_q$, respectively, are related to the problem of optimal adaptive storage of data of a signal. The difference between them is that the quantities $\epsilon_n(W)_q$ and $\rho_n(W)_q$ are based on any information, while the quantities $e_n(W)_q$ and $r_n(W)_q$ are based on standard information, i.e., the sampling values of a signal.

The concept of ε-entropy introduced by Kolmogorov and Tikhomirov [9], comes from Information Theory. It expresses the necessary number of binary signs for approximate recovery of a signal from a certain set with accuracy ε.

The concept of pseudo-dimension of a real-valued functions set was introduced by Pollard [11] and later Haussler [8] as an extention of the Vapnik Chervonekis [14] dimension of an indicator function set. The pseudo-dimension and Vapnik Chervonekis dimension measure the capacity of a set of functions. They play an important role in theory of pattern recognition and regression estimation, empirical processes and Computational Learning Theory (see also [3, 12, 13] for details).

We define Besov-type space $B^a_{p,\theta} = B^a_{p,\theta}(\mathbb{T}^d)$. For univariate functions f on $\mathbb{T}$ the lth difference operator Δ^l_h is defined by

$$\Delta^l_h(f,x) := \sum_{j=0}^{l} (-1)^{l-j} \binom{l}{j} f(x+jh).$$

For $f \in L_p(\mathbb{T}^d)$. If e is any subset of $[d]$, for multivariate functions f on $\mathbb{T}^d$ the mixed (r,e)th difference operator $\Delta^{l,e}_h$ is defined by

$$\Delta^{l,e}_h := \prod_{i\in e} \Delta^l_{h_i},\ \Delta^{l,\varnothing}_h = I,$$

where the univariate operator $\Delta^l_{h_i}$ is applied to the univariate function f by considering f as a function of variable x_i with the other variables held fixed.

Let

$$\omega^e_l(f,t)_p := \sup_{|h_i|<t_i, i\in e} \left\|\Delta^{l,e}_h f\right\|_p,\ t \in \mathbb{T}^d,$$

be the mixed (r,e)th modulus of smoothness of f. In particular, $\omega^{\varnothing}_l(f,t)_p = \|f\|_p$.

Let $1 \le p \le \infty,\ 0 < \theta \le \infty,\ a = (a_1, a_2, \ldots, a_d) \in \mathbb{R}^d_+$. We introduce the quasi-semi-norm $|f|_{B^{a,e}_{p,\theta}}$ for a set $e \subset \{1, ..., d\}$ and a function $f \in L_p$ by

$$|f|_{B^{a,e}_{p,\theta}} := \begin{cases} \left(\int\limits_{\mathbb{T}^d} \left\{ \prod\limits_{i\in e} t_i^{-a_i} \omega^e_l(f,t)_p \right\}^{\theta} \prod\limits_{i\in e} t_i^{-1} dt \right)^{1/\theta}, & \theta < \infty, \\ \sup\limits_{t\in\mathbb{T}^d} \left\{ \prod\limits_{i\in e} t_i^{-a_i} \omega^e_l(f,t)_p \right\}, & \theta = \infty, \end{cases}$$

in particular, $|f|_{B^{a,\varnothing}_{p,\theta}} = \|f\|_p$, where l is a fixed integer such that $l > \max\limits_{1\le i\le d} a_i$. The Besov-type space $B^a_{p,\theta} = B^a_{p,\theta}(\mathbb{T}^d)$ is defined as the set of all functions $f \in L_p$ such that the Besov-type quasi-norm

$$\|f\|_{B^a_{p,\theta}} := \sum_{e\subset[d]} |f|_{B^{a,e}_{p,\theta}}$$

is finite.

It is well known that different admissible values of l define equivalent Besov-type quasi-norm. Denote by $U^a_{p,\theta} = U^a_{p,\theta}(\mathbb{T}^d)$ the unit ball in the space $B^a_{p,\theta}$, i. e.,

$$U^a_{p,\theta} := \big\{f \in B^a_{p,\theta} : \|f\|_{B^a_{p,\theta}} \le 1\big\}.$$

We denote by $A_n(f) \ll B_n(f)$ if $A_n(f) \le C.B_n(f)$, where C is a constant independent of n and $f \in W$; $A_n(f) \asymp B_n(f)$ if $A_n(f) \ll B_n(f)$ and $B_n(f) \ll A_n(f)$.

Through this paper we assume that the mixed smoothness $a = (a_1, a_2, \dots, a_d) \in \mathbb{R}^d_+$ of the space $B^a_{p,\theta}$ is fixed and such that

$$0 < r = a_1 = a_2 = \ldots = a_s = a_{s+1} < a_{s+2} \le \ldots \le a_d,\ 0 \le s \le d-1.$$

Let us briefly formulate the main results of the present paper. Let $1 < p,\ q < \infty$, $0 < \theta \le \infty$ and $r > 1/p$. We establish the asymptotic orders

$$\epsilon_n(U^a_{p,\theta}, L_q) \asymp \rho_n(U^a_{p,\theta}, L_q) \asymp n^{-r}(\log n)^{s(r+1/2-1/\theta)} \tag{4}$$

which exends the results in [3] for the case of uniformed mixed smoothness a, i. e., for the case $s = d-1$, and

$$e_n(U^a_{p,\theta}, L_q) \asymp r_n(U^a_{p,\theta}, L_q) \asymp n^{-r}(\log n)^{s(r+1/2-1/\theta)}. \tag{5}$$

To prove (4) and (5) we develop the method and technique in [3] with overcoming certain difficulties. The proof of the upper bounds, in particular, is based on a trigonometric sampling representations in the space $B^a_{p,\theta}$ with a discrete equivalent quasi-norm, and a special decomposition of functions $f \in B^a_{p,\theta}$ into a series corresponding to the non-uniformed mixed smoothness a (see (18) and (19)).

Let us give a brief outline of the present paper. In Section 2, we introduce a notion of Besov-type spaces $B^a_{p,\theta}$ of functions having a mixed smoothness $a \in \mathbb{R}^d_+$ and describe a trigonometric sampling representations in the space $B^a_{p,\theta}$ with a discrete equivalent quasi-norm. In Section 3, we prove the asymptotic orders (4) and (5) and construct corresponding asymptotically optimal methods of nonlinear approximations.

2. TRIGONOMETRIC SAMPLING REPRESENTATIONS IN BESOV SPACES

In this section, we describe a trigonometric sampling representations in the space $B^a_{p,\theta}$ with a discrete equivalent quasi-norm.

As usual, $\widehat{f}(k)$ denotes the kth Fourier coefficient of $f \in L_p$ for $1 \le p \le \infty$. Let $k = (k_1, k_2, \dots, k_d) \in \mathbb{Z}^d_+$ and $P_k := \{s \in \mathbb{Z}^d : \lfloor 2^{k_j-1}\rfloor \le |s_j| < 2^{k_j}, j = 1, ..., d\}$, where $\lfloor a\rfloor$ denotes the integer part of $a \in \mathbb{R}_+$. We define the operator δ_k as

$$\delta_k(f) := \sum_{s\in P_k} \widehat{f}(s)e^{i(s,\cdot)}.$$

The well known Littlewood-Paley theorem (see [10]) states that for $1 < p < \infty$ there holds the norm equivalence

$$\|f\|_p \asymp \left\| \left(\sum_{k \in \mathbb{Z}_+^d} |\delta_k(f)|^2 \right)^{1/2} \right\|_p .$$

We next recall some known equivalences of quasi-norms (see [2]). If $x = (x_1, x_2, \dots, x_d)$, $y = (y_1, y_2, \dots, y_d) \in \mathbb{R}^d$, denote $(x, y) = \sum_{i=1}^{d} x_i y_i$. For $1 < p < \infty$ and $\theta < \infty$ we have that

$$\|f\|_{B^a_{p,\theta}} \asymp \left(\sum_{k \in \mathbb{Z}_+} \left\{ 2^{(a,k)} \|\delta_k(f)\|_p \right\}^\theta \right)^{1/\theta},$$

with the right side changed to a supremum for $\theta = \infty$.

For a positive integer m, the de la Vallée Poussin kernel V_m of order m is defined as

$$V_m(t) := \frac{1}{m} \sum_{k=m}^{2m-1} D_k(t) = \frac{\sin(mt/2)\sin(3mt/2)}{m \sin^2(t/2)},$$

where

$$D_m(t) := \sum_{|k| \le m} e^{ikt}$$

is the univariate Dirichlet kernel of order m. For completeness we put $V_0 = 1$.

For univariate functions $f \in L_p(\mathbb{T})$, we define the function $U_m(f)$ as

$$U_m(f) := f * V_m = \frac{1}{2\pi} \int_{\mathbb{T}} f(t)\ V_m(\cdot - t) dt,$$

and the function $V_m(f)$ as

$$V_m(f) := \frac{1}{3m} \sum_{k \in P_m} f(hk)\ V_m(\cdot - hk), \tag{6}$$

where $h = 2\pi/3m$ and $P_m := \{k \in \mathbb{Z} : 0 \le k < 3m\}$. If $m \in \mathbb{Z}_+^d$, the mixed operator V_m is defined for multivariate functions $f \in L_p(\mathbb{T}^d)$ by

$$V_m(f) := \prod_{j=1}^{d} V_{m_j}(f),$$

where the univariate operator V_{m_j} is applied to the function f by considering f as a function of variable x_j with the other variables held fixed. Notice that $V_m(f)$ is a trigonometric polynomial of order at most $2m_j - 1$ in the variable x_j, and

$$V_m(f, hk) = f(hk), \quad k \in P_m^d,$$

where $h = (2\pi/3)(m_1^{-1}, \dots, m_d^{-1}),\ P_m^d := \{k \in \mathbb{Z}^d : 0 \le k_j < 3m_j,\ j = 1, \dots, d\}$. We get

$$\|V_m(f)\|_p \asymp \prod_{j=1}^{d} m_j^{-1/p} \|\{f(hk)\}\|_{l_p^\nu}, \quad 1 \le p \le \infty,$$

where $\nu = |P_m^d| = 3^d \prod_{j=1}^d m_j$. Denote by $\mathcal{T}_m$ the space of all trigonometric polynomials of order at most m_j in the variable x_j for $j = 1, \dots, d$. It is easy to check that

$$V_m(f) = f, \quad \forall f \in \mathcal{T}_m. \tag{7}$$

Next, for univariate functions $f \in L_p(T)$, we define

$$v_0(f) := V_1(f),$$

$$v_k(f) := V_{2^k}(f) - V_{2^{k-1}}(f), \quad k = 1, 2, \dots.$$

For $k \in \mathbb{Z}_+^d$, the definition of the mixed operator v_k for multivariate functions in L_p is similar to the mixed operator V_m. The mixed operators $u_k,\ k \in \mathbb{Z}_+^d$ are defined in a similar way by replacing $V_m(f)$ by $U_m(f)$.

Note that $v_k(f)$ and $u_k(f)$ are a trigonometric polynomial of order at most $2^{k_j+1} - 1$ in the variable x_j for $j = 1, \dots, d$.

To prove the main results (4) and (5), we need the following two lemmas. Put $|k|_1 = \sum_{i=1}^{d} |k_i|$ for $k \in \mathbb{Z}^d$.

Lemma 2.1. *Let $\Lambda_a := \{\xi :\ \xi = (a,k),\ k \in \mathbb{Z}_+^d\}$, $D_\xi := \{k \in \mathbb{Z}_+^d : (a,k) = \xi\}$. Then we have*

$$\sum_{k \in D_\xi} 2^{|k|_1} \asymp 2^{\xi/r} \xi^s, \ \forall\ \xi \in \Lambda_a.$$

Lemma 2.2. *Let $1 \le p \le \infty$, $0 < \theta \le \infty$ and $r > 0$. Then for $\theta < \infty$, we have*

$$\|f\|_{B_{p,\theta}^a} \asymp \left(\sum_{k \in \mathbb{Z}_+} \left\{ 2^{(a,k)} \|u_k(f)\|_p \right\}^\theta \right)^{1/\theta},$$

and if in addition $r > 1/p$,

$$\|f\|_{B_{p,\theta}^a} \asymp \left(\sum_{k \in \mathbb{Z}_+} \left\{ 2^{(a,k)} \|v_k(f)\|_p \right\}^\theta \right)^{1/\theta},$$

with the right side changed to a supremum for $\theta = \infty$.

Lemma 2.1 and Lemma 2.2 have been proved in [2].

Lemma 2.3.

(i) *Let* $G_\xi := \{k \in \mathbb{U}_+^d : (a,k) \leq \xi\}$, $\xi > 0$. *Then there exist positive constants* C_1 *and* C_2 *such that*

$$C_2 2^{\xi/r}\xi^s \leq \sum_{k\in G_\xi} 2^{|k|_1} \leq C_1 2^{\xi/r}\xi^s. \tag{8}$$

(ii) *For a fixed number* $\lambda > r\log_2 C_1/C_2$, *let* $\{\xi_j\}_{j=1}^\infty$ *be any positive sequence of numbers such that* $\xi_{j+1} - \xi_j \geq \lambda$, $j \geq 1$. *Then we have that*

$$\sum_{k\in G_{\xi_{j+1}}\setminus G_{\xi_j}} 2^{|k|_1} \asymp 2^{\xi_j/r}\xi_j^s. \tag{9}$$

Proof. (i) This assertion follows from Lemma 2.1.

(ii) From (8), we have

$$\begin{aligned}\sum_{k\in G_{\xi_{j+1}}\setminus G_{\xi_j}} 2^{|k|_1} &= \sum_{k\in G_{\xi_{j+1}}} 2^{|k|_1} - \sum_{k\in G_{\xi_j}} 2^{|k|_1}\\ &\geq C_2 2^{\xi_{j+1}/r}\xi_{j+1}^s - C_1 2^{\xi_j/r}\xi_j^s\\ &\geq C_2 2^{(\xi_j+\lambda)/r}(\xi_j+\lambda)^s - C_1 2^{\xi_j/r}\xi_j^s\\ &\geq (C_2 2^{\lambda/r} - C_1)2^{\xi_j/r}\xi_j^s.\end{aligned}$$

Hence

$$\sum_{k\in G_{\xi_{j+1}}\setminus G_{\xi_j}} 2^{|k|_1} \asymp 2^{\xi_j/r}\xi_j^s.$$

□

Let $\varphi_{k,s} := V_{m^k}(\cdot - sh^k)$, and

$$Q_k := \{s \in \mathbb{Z}^d : 0 \leq s_j < 3.2^{k_j},\ j = 1,\dots,d\}.$$

where $m^k := (2^{k_1},\dots,2^{k_d})$, $h^k := (2\pi/3)(2^{-k_1},\dots,2^{-k_d})$.

From Lemma 2.2 and (6)-(7) we derive the following trigonometric sampling representation in spaces $B^a_{p,\theta}$. Let $1 \leq p \leq \infty$, $0 < \theta \leq \infty$, and $r > 0$. Then every $f \in B^a_{p,\theta}$ can be represented as the series

$$f = \sum_{k\in\mathbb{Z}^d_+}\sum_{s\in Q_k} f_{k,s}\varphi_{k,s} \tag{10}$$

for which there holds the quasi-norm equivalence

$$\|f\|_{B^a_{p,\theta}} \asymp \left(\sum_{k\in\mathbb{Z}_+}\left\{2^{(a,k)-|k|_1/p}\|\,\{f_{k,s}\}\,\|_{l_p^{|Q_k|}}\right\}^\theta\right)^{1/\theta} \tag{11}$$

for $\theta < \infty$, with the sum replaced by a supremum for $\theta = \infty$. Based on the representation (10)-(11), we can extend the definition of Besov space of mixed smoothness a for $a \in \mathbb{R}^d$ and

$0 < p,\ \theta \le \infty$, as the space of all functions f on $\mathbb{T}^d$ which can be represented by the series (10) for which the discrete quasi-norm in the right-hand side of (11) is finite. We also use the notation $B_{p,\theta} = B^a_{p,\theta}$ for $a = (0, \dots, 0)$.

Let $1 < q < \infty$. From these quasi-norm equivalences, it is easy to verify the inequalities

$$\|f\|_{B_{q,\max\{q,2\}}} \le \|f\|_q \le \|f\|_{B_{q,\min\{q,2\}}}. \tag{12}$$

Let $0 < p \le \infty$, we define l^m_p as the quasi-normed space of all real number sequences $x = \{x_k\}_{k=1}^m$ equipped with the quasi-norm

$$\| \{x_k\} \|_{l^m_p} = \|x\|_{l^m_p} := \left(\sum_{k=1}^m |x_k|^p \right)^{1/p},$$

with the change to max norm when $p = \infty$.

Let $0 < p, \theta \le \infty$ and $N = \{N_k\}_{k \in Q}$ be a sequence of natural numbers, with Q a finite set of indices. Denote by $b^N_{p,\theta}$ a the space of all such sequences $x = \{x^k\}_{k\in Q} = \{\{x^k_j\}_{j=1}^{N_k}\}_{k \in Q}$ for which the mixed quasi-norm $\|\{\{x^k_j\}\}\|_{b^N_{p,\theta}} = \|x\|_{b^N_{p,\theta}}$ is finite. Here, the mixed quasi-norm $\|.\|_{b^N_{p,\theta}}$ is defined as

$$\|x\|_{b^N_{p,\theta}} := \left(\sum_{k \in Q} \|x^k\|^\theta_{l^{N_k}_p} \right)^{1/\theta}$$

for finite θ, the sum is replaced by a supremum when $\theta = \infty$. Let $S^N_{p,\theta}$ be the unit ball in $b^N_{p,\theta}$.

3. ASYMPTOTIC ORDER FOR ENTROPY NUMBERS

In this section, we give the asymptotic order of entropy numbers $\epsilon_n(U^a_{p,\theta}, L_q)$, non-linear widths $\rho_n(U^a_{p,\theta}, L_q)$ and $e_n(U^a_{p,\theta}, L_q)$, $r_n(U^a_{p,\theta}, L_q)$.

By Definition 1 and Definition 2, we have inequalities

$$e_n(U^a_{p,\theta}, L_q) \ge \epsilon_n(U^a_{p,\theta}, L_q),\ r_n(U^a_{p,\theta}, L_q) \ge \rho_n(U^a_{p,\theta}, L_q). \tag{13}$$

Moreover, from the definitions we can see that $\dim_p(B) \le \log |B|$, and consequently, the pseudo-dimension of a set B of cardinality $\le 2^n$ is not greater than n, and therefore, there hold the inequalities

$$e_n(U^a_{p,\theta}, L_q) \ge r_n(U^a_{p,\theta}, L_q),\ \epsilon_n(U^a_{p,\theta}, L_q) \ge \rho_n(U^a_{p,\theta}, L_q). \tag{14}$$

Hence, the upper bounds of $r_n(U^a_{p,\theta}, L_q)$, $\epsilon_n(U^a_{p,\theta}, L_q)$ and $\rho_n(U^a_{p,\theta}, L_q)$ in (4) and (5) are implied from the upper bound of $e_n(U^a_{p,\theta}, L_q)$.

Let $\Phi = \{\varphi_k\}_{k \in Q}$ a family of elements in L_q. Denote by $M_n(\Phi)$ the nonlinear manifold of all linear combinations of the form $\varphi = \sum_{k\in K} a_k \varphi_k$, where K is a subset of Q having cardinality n. The n-term L_q-approximation of an element $f \in L_q$ with regard to the family Φ is called the L_q-approximation of f by elements from $M_n(\Phi)$. To establish the upper bound for the asymptotic orders of $\epsilon_n(U^a_{p,\theta}, L_q)$, we use the non-linear n-term L_q-approximation with respect to the family

$$V := \{\varphi_{k,s}\}_{s\in Q_k, k\in\mathbb{Z}^d_+}.$$

Note that the family V is formed from the integer translates of the mixed dyadic scales of the tensor product multivariate de la Vallée Poussin kernel.

Theorem 3.1. *Let $1 < p,\ q < \infty$, $0 < \theta \le \infty$ and $r > 1/p$. Then we have that*

$$\epsilon_n(U^a_{p,\theta}, L_q) \le e_n(U^a_{p,\theta}, L_q) \ll (n/\log^s n)^{-r}(\log n)^{s(1/2-1/\theta)}. \tag{15}$$

In addition, we can explicitly construct a finite subset V^* of V, a subset B in $M_n(V^*)$ having $|B| \le 2^n$, and a mapping $S^B_n : U^a_{p,\theta} \to B$ of the form (2) such that

$$E(U^a_{p,\theta}, B, L_q) \le \sup_{f\in U^a_{p,\theta}} \|f - S^B_n(f)\|_q \ll (n/\log^s n)^{-r}(\log n)^{s(1/2-1/\theta)}.$$

Theorem 3.1 is derived from the following theorem.

Theorem 3.2. *Let $0 < p,\ q,\ \theta \le \infty,\ 0 < \tau \le \theta$ and $r > 1/p$. Then, we have that*

$$\epsilon_n(U^a_{p,\theta},\ B_{q,\tau}) \le e_n(U^a_{p,\theta},\ B_{q,\tau}) \ll E_{\theta,\tau}(n), \tag{16}$$

where $E_{\theta,\tau}(n) = (n/\log^s n)^{-r}(\log n)^{s(1/\tau-1/\theta)}$.

In addition, we can explicitly construct a finite subset V^* in V, a subset B in $M_n(V^*)$ having $|B| \le 2^n$, and a mapping $S^B_n : U^a_{p,\theta} \to B$ of the form (2) such that

$$E(U^a_{p,\theta}, B, B_{q,\tau}) \le \sup_{f\in U^a_{p,\theta}} \|f - S^B_n(f)\|_{B_{q,\tau}} \ll E_{\theta,\tau}(n). \tag{17}$$

Proof. Obviously, (16) follows from (17), and consequently, it is enough to prove (17). Take $k = (k_1, k_2, ..., k_{s+1}, k_{s+2}, ..., k_d) \in \mathbb{U}^d_+$. Denote by $\Lambda = \{ \sum_{i=s+2}^{d} a_i k_i : k_i \in \mathbb{U}_+,\ i = s+2, ..., d \}$. We fix a subsequence $\Lambda' := \{\nu_{2,j}\}_{j=1}^\infty \subset \Lambda$ such that $\nu_{2,j} - \nu_{2,j-1} > \max\{a_d,\ \lambda\}$ (number λ is defined in Lemma 2.3).

Let $G_{\nu_{2,j}} := \{(k_{s+2}, ..., k_d) : \sum_{i=s+2}^{d} a_i k_i \le \nu_{2,j}\},\ D'_{\nu_{2,j}} = G_{\nu_{2,j}} \setminus G_{\nu_{2,j-1}},\ j \ge 2$ and $D'_{\nu_{2,1}} := G_{\nu_{2,1}}$.

By (10), (11) we can verify that every $f \in B^a_{p,\theta}$ is represented as the series

$$f = \sum_{\nu=(\nu_1,\nu_2)} f_\nu, \tag{18}$$

converging in the norm of $B_{q,\tau}$, any $\nu = (\nu_1, \nu_2) \in \mathbb{Z}_+ \times \Lambda$ and

$$f_\nu = \sum_{k\in D_\nu} \sum_{s\in Q_k} f_{k,s}\varphi_{k,s}, \tag{19}$$

where $D_\nu := D''_\nu \cap D'_{\nu_{2,j}}$, $D''_\nu := \{(k_1, k_2, ..., k_{s+1}) : \ k_1 + k_2 + \cdots + k_{s+1} = \nu_1\}$. Moreover, there hold the quasi-norm equivalences

$$\begin{aligned} &\|f_\nu\|_{B^a_{p,\theta}} \asymp 2^{r\nu_1+\nu_2}\|\{\{2^{-|k|_1/p} f_{k,s}\}\}\|_{b^{N^\nu}_{p,\theta}}, \\ &\|f_\nu\|_{B_{q,\tau}} \asymp \|\{\{2^{-|k|_1/q} f_{k,s}\}\}\|_{b^{N^\nu}_{q,\tau}}, \ N^\nu := \{N_k\}_{k\in D_\nu} = \{|Q_k|\}_{k\in D_\nu}. \end{aligned} \tag{20}$$

The representation (18) – (19) with the the quasi-norm equivalences (20) plays a basic role in the proof of the theorem. Notice that in the case of the uniform mixed smoothness it required a much simpler representation [3].

Obviously, $D_\nu \cap D_{\nu'} = \emptyset$ if $\nu \neq \nu'$ and $\mathbb{Z}^d_+ = \cup_{\nu\in\mathbb{Z}_+\times\Lambda} D_\nu$. We have

$$|D'_\nu| \asymp {\nu_2}^{d-s-2}, \ |D''_\nu| \asymp \nu_1^s$$

and consequently,

$$|D_\nu| = |D'_\nu||D''_\nu| \asymp {\nu_1}^s{\nu_2}^{d-s-2}.$$

Let $r' = a_{s+2} = \ldots = a_{s+s'+2} < a_{s+s'+3} \leq \ldots \leq a_d$. From (9) we get

$$m_\nu = 3^d \sum_{k\in D_\nu} 2^{|k|_1} \asymp \nu_1^s 2^{\nu_1} 2^{\nu_2/r'} \nu_2^{s'}, \tag{21}$$

where $m_\nu := \sum_{k\in D_\nu} |Q_k|$. Given a positive integer n, we take a positive integer $\xi = \xi(n)$ satisfying the condition

$$C2^\xi \xi^s \leq n \asymp 2^\xi \xi^s, \tag{22}$$

where C is an absolute constant whose value will be chosen below.

Notice that there hold the inequality $\|f\|_{B_{q,\tau}} \leq \|f\|_{B_{\infty,\tau}}$ and the inclusion $U^a_{p,\theta} \subset U^a_{p,\max\{p,\theta\}}$. Therefore, it suffices to treat the case $p \leq \theta$ and $q = \infty$. We choose fixed numbers δ, α, ε satisfying $0 < \delta < \min\{1, \ p(r-1/p)\}$, $\max\{r, \ (1+\delta)r'/pr\} < \alpha < r'$, $(1+\delta)/pr < \varepsilon < \alpha/r'$. Let the sequence $\{n_\nu\}_{\nu=0}^\infty$ be given by

$$n_\nu := \begin{cases} \lfloor m_\nu 2^{(1-\delta)(\xi-\nu_1-\nu_2/\alpha)} \rfloor + 1 & \text{if } \ 0 \leq \nu_1 + \nu_2/\alpha < \xi, \\ \lfloor m_\nu 2^{(1+\delta)(\xi-\nu_1-\nu_2/\alpha)} \rfloor & \text{if } \ \nu_1 + \nu_2/\alpha \geq \xi. \end{cases} \tag{23}$$

It is easy to check that $n_\nu > 0$ for $\nu_1 + \nu_2/\alpha \leq \xi(1+\delta)/(1+\delta-\varepsilon) - \nu_0$, where $\nu_0 = \nu_0(\delta, d)$ is a positive constant. Since $(1+\delta)/(1+\delta-\varepsilon) > r/(r-1/p)$, we can fix a number γ so that $r/(r-1/p) < \gamma < (1+\delta)/(1+\delta-\varepsilon)$. Put $\xi^* = \lfloor \gamma\xi \rfloor$. Then for ξ large enough, we have $n_\nu > 0, \ \ \forall \ \nu_1 + \nu_2/\alpha \leq \xi^*$.

Let $0 \leq \nu_1 + \nu_2/\alpha \leq \xi$. Then $n_\nu \geq m_\nu$. Take a number ρ such that $0 < \rho \leq \min\{1, p, \theta\}$ and $N_k = 2^{|k|_1} \leq 2^{\nu_1}2^{\nu_2/r'} := N_0, \ \forall k \in D_\nu$. From the inequalities

$$\|\cdot\|_{b^{N^\nu}_{\rho,\rho}} \leq |D_\nu|^{1/\rho-1/\theta} N_0^{1/\rho-1/p} \|\cdot\|_{b^{N^\nu}_{p,\theta}}$$

and

$$\|\cdot\|_{b^{N^\nu}_{\infty,\tau}} \leq |D_\nu|^{1/\tau} \|\cdot\|_{b^{N^\nu}_{\infty,\infty}},$$

it follows that for any subset $M_\nu \subset b^{N^\nu}_{\infty,\tau}$ and mapping $G_\nu : b^{N^\nu}_{p,\theta} \to M_\nu$ such that

$$\sup_{x \in S^{N^\nu}_{p,\theta}} \|x - G_\nu(x)\|_{b^{N^\nu}_{\infty,\tau}} \le |D_\nu|^{1/\rho - 1/\theta + 1/\tau} N_0^{1/\rho - 1/p} \sup_{x \in S^{N^\nu}_{\rho,\rho}} \|x - G_\nu(x)\|_{b^{N^\nu}_{\infty,\infty}}.$$

Considering $S^{N^\nu}_{\rho,\rho}$ and $b^{N^\nu}_{\infty,\infty}$ as $B^{m_\nu}_\rho$ and $l^{m_\nu}_\infty$ and applying the result proved in [3, Lemma 1], then for any positive integer n we can explicitly construct a subset M of l^m_∞ for $n \ge m$ having cardinality at most 2^n and a mapping $S : l^m_\rho \to M$ such that

$$\sup_{x \in B^m_p} \|x - S(x)\|_{l^m_\infty} \le C(p) m^{-1/\rho} 2^{-n/m}.$$

Hence, we obtain there exists a set $M_\nu \subset b^{N^\nu}_{\infty,\tau}$ of cardinality at most 2^{n_ν} and a mapping $G_\nu : b^{N^\nu}_{p,\theta} \to M_\nu$ such that

$$\sup_{x \in S^{N^\nu}_{p,\theta}} \|x - G_\nu(x)\|_{b^{N^\nu}_{\infty,\tau}} \le |D_\nu|^{1/\rho - 1/\theta + 1/\tau} N_0^{1/\rho - 1/p} m_\nu^{-1/\rho} 2^{-n_\nu/m_\nu}.$$

We define a subset B_ν of $B_{\infty,\tau}$ and a mapping $S_\nu : B^a_{p,\theta} \to B_\nu$ as follows. From (11),

$$\|f\|_{B^a_{p,\theta}} = \left(\sum_{k \in \mathbb{Z}_+} \left\{ 2^{(a,k) - |k|_1/p} \| \{f_{k,s}\} \|_{l_p^{|Q_k|}} \right\}^\theta \right)^{1/\theta},$$

$$\|f_\nu\|_{B^a_{p,\theta}} = \left(\sum_{k \in D_\nu} \left\{ 2^{(a,k) - |k|_1/p} \| \{f_{k,s}\} \|_{l_p^{|Q_k|}} \right\}^\theta \right)^{1/\theta},$$

we obtain $\|f_\nu\|_{B^a_{p,\theta}} \le \|f\|_{B^a_{p,\theta}}$. Hence, if $f \in B^a_{p,\theta}$ then $f_\nu \in B^a_{p,\theta}$, and consequently $\{\{f_{k,s}\}_{s \in Q_k}\}_{k \in D_\nu}$ belongs to $b^{N^\nu}_{p,\theta}$. We put

$$S_\nu(f) = \sum_{k \in D_\nu} \sum_{s \in Q_k} f^*_{k,s} \varphi_{k,s}$$

and $B_\nu = S_\nu(M_\nu)$, where $\{\{f^*_{k,s}\}_{s \in Q_k}\}_{k \in D_\nu} = G_\nu(\{\{f_{k,s}\}_{s \in Q_k}\}_{k \in D_\nu})$. We can see that $|B_\nu| \le |M_\nu| \le 2^{n_\nu}$ and

$$\begin{aligned} \|f_\nu - S_\nu(f)\|_{B_{\infty,\tau}} &\asymp \|\{\{f_{k,s} - f^*_{k,s}\}\}\|_{b^{N^\nu}_{\infty,\tau}} \\ &\ll |D_\nu|^{1/\rho - 1/\theta + 1/\tau} N_0^{1/\rho - 1/p} m_\nu^{-1/\rho} 2^{-n_\nu/m_\nu} 2^{-r\nu_1 - \nu_2} N_0^{1/p} \|f_\nu\|_{B^a_{p,\theta}} \\ &\ll \nu_1^{s(1/\tau - 1/\theta)} 2^{-r\xi} 2^{r(\xi - \nu_1 - \nu_2/\alpha)} 2^{-2^{(1-\delta)(\xi - \nu_1 - \nu_2/\alpha)}} 2^{(r/\alpha - 1)\nu_2} \nu_2^\mu \|f_\nu\|_{B^a_{p,\theta}} \\ &\ll \xi^{s(1/\tau - 1/\theta)} 2^{-r\xi} 2^{r(\xi - \nu_1 - \nu_2/\alpha)} 2^{-2^{(1-\delta)(\xi - \nu_1 - \nu_2/\alpha)}} \|f_\nu\|_{B^a_{p,\theta}}, \end{aligned}$$

where $\mu = (d - s - 2)(1/\rho - 1/\theta + 1/\tau) - s'/\rho$.

Therefore

$$\|f_\nu - S_\nu(f)\|_{B_{\infty,\tau}} \ll A(\nu) \|f_\nu\|_{B^a_{p,\theta}}, \tag{24}$$

where $A(\nu) = \xi^{s(1/\tau-1/\theta)}2^{-r\xi}2^{r(\xi-\nu_1-\nu_2/\alpha)}2^{-2^{(1-\delta)(\xi-\nu_1-\nu_2/\alpha)}}$.

Let $\xi < \nu_1+\nu_2/\alpha \le \xi^*$. Then $n_\nu < m_\nu$. The following result was proved in [3, Lemma 4]. Let $0<p,\ \theta,\ \tau \le \infty$. Then for any positive integer $n < m = \sum_{k\in Q} N_k$, we can explicitly construct a subset $M \subset b^N_{\infty,\tau}$ of cardinality at most $2^n\binom{m}{n}$ and a mapping $S : b^N_{p,\theta} \to M$ such that

$$\sup_{x\in S^N_{p,\theta}} \|x - S(x)\|_{b^N_{\infty,\tau}} \le C(p)n^{-1/p}\ |Q|^{1/\tau+(1/p-1/\theta)_+}.$$

Therefore, we can construct a subset B_ν of $B_{\infty,\tau}$ having cardinality at most $2^{n_\nu}\binom{m_\nu}{n_\nu}$, as well as a mapping $S_\nu : B^a_{p,\theta} \to B_\nu$ such that

$$\|f_\nu - S_\nu(f)\|_{B_{\infty,\tau}} \asymp \|\{\{f_{k,s} - f^*_{k,s}\}\|_{b^{N^\nu}_{\infty,\tau}} \ll {n_\nu}^{-1/p}\ |D_\nu|^{1/\tau+(1/p-1/\theta)_+}\|\{\{f_{k,s}\}\}\|_{b^{N^\nu}_{p,\theta}}. \quad (25)$$

We have $|k|_1 \le \nu_1 + \nu_2/r'$, hence

$$\|f_\nu\|_{B^a_{p,\theta}} \asymp 2^{r\nu_1+\nu_2}\|\{\{2^{-|k|_1/p}f_{k,s}\}\}\|_{b^{N^\nu}_{p,\theta}} \ge 2^{r\nu_1+\nu_2}2^{-\nu_1/p}2^{-\nu_2/pr'}\|\{\{f_{k,s}\}\}\|_{b^{N^\nu}_{p,\theta}},$$

and consequently $\|\{\{f_{k,s}\}\}\|_{b^{N^\nu}_{p,\theta}} \ll 2^{-r\nu_1-\nu_2}2^{\nu_1/p}2^{\nu_2/pr'}\|f_\nu\|_{B^a_{p,\theta}}$. We continue the estimation (25),

$$\begin{aligned}
\|f_\nu - S_\nu(f)\|_{B_{\infty,\tau}} &\asymp \|\{\{f_{k,s} - f^*_{k,s}\}\}\|_{b^{N^\nu}_{\infty,\tau}} \\
&\ll {n_\nu}^{-1/p}\ |D_\nu|^{1/\tau+(1/p-1/\theta)_+}\|\{\{f_{k,s}\}\}\|_{b^{N^\nu}_{p,\theta}} \\
&\ll \{\nu_1^s 2^{\nu_1}2^{\nu_2/r'}\nu_2^{s'}2^{(1+\delta)\mu_1}\}^{-1/p}(\nu_1^s\nu_2^{d-s-2})^{\mu_2}\ 2^{-r\nu_1-\nu_2}2^{\nu_1/p}2^{\nu_2/pr'}\|f_\nu\|_{B^a_{p,\theta}} \\
&\ll 2^{-r\xi}\nu_1^{s(1/\tau-1/\theta)}2^{(r-(1+\delta)/p)\mu_1}\nu_2^{(d-s-2)\mu_2-s'/p}2^{-(1-r/\alpha)\nu_2}\|f_\nu\|_{B^a_{p,\theta}} \\
&\ll 2^{-r\xi}\nu_1^{s(1/\tau-1/\theta)}2^{(r-(1+\delta)/p)\mu_1}\|f_\nu\|_{B^a_{p,\theta}} \\
&\ll C(\nu)\|f_\nu\|_{B^a_{p,\theta}},
\end{aligned}$$

where $C(\nu) = 2^{-r\xi}\nu_1^{s(1/\tau-1/\theta)}2^{-\beta(\nu_1+\nu_2/\alpha-\xi)}$, $\beta = r-(1+\delta)/p > 0$, $\mu_1 = \xi - \nu_1 - \nu_2/\alpha$, $\mu_2 = 1/\tau + 1/p - 1/\theta$. It is easy to check that

$$C(\nu) \le \begin{cases} 2^{-r\xi}\xi^{s(1/\tau-1/\theta)}2^{-\beta(\nu_1+\nu_2/\alpha-\xi)} & \text{if } \ \nu_1 \le \xi, \\ 2^{-r\xi}\xi^{s(1/\tau-1/\theta)}(\nu_1+\nu_2/\alpha-\xi)^{s(1/\tau-1/\theta)}2^{-\beta(\nu_1+\nu_2/\alpha-\xi)} & \text{if } \ \nu_1 > \xi. \end{cases}$$

Finally, let $\nu_1 + \nu_2/\alpha > \xi^*$. From (20) and the Holder inequality, it follows that for any $\nu_1+\nu_2/\alpha > \xi^*$. Put $\mu = r - 1/p$, we get

$$\begin{aligned}
\|f_\nu\|_{B_{\infty,\tau}} &\ll 2^{-(r\nu_1+\nu_2)}2^{\nu_1/p}2^{\nu_2/pr'}\|f_\nu\|_{B^a_{p,\tau}} \\
&\ll 2^{-(r\nu_1+\nu_2)}2^{\nu_1/p}2^{\nu_2/pr'}|D_\nu|^{1/\tau-1/\theta}\|f_\nu\|_{B^a_{p,\theta}} \\
&\ll 2^{-\mu\xi^*}(\xi^*)^{s(1/\tau-1/\theta)}(\nu_1+\nu_2/\alpha-\xi^*)^{s(1/\tau-1/\theta)}2^{-\mu(\nu_1+\nu_2/\alpha-\xi^*)}\|f_\nu\|_{B^a_{p,\theta}} \qquad (26) \\
&\ll 2^{-r\xi}\xi^{s(1/\tau-1/\theta)}(\nu_1+\nu_2/\alpha-\xi^*)^{s(1/\tau-1/\theta)}2^{-\mu(\nu_1+\nu_2/\alpha-\xi^*)}\|f_\nu\|_{B^a_{p,\theta}} \\
&\ll E(\nu)\|f_\nu\|_{B^a_{p,\theta}},
\end{aligned}$$

where $E(\nu) = 2^{-r\xi}\xi^{s(1/\tau-1/\theta)}(\nu_1+\nu_2/\alpha-\xi^*)^{s(1/\tau-1/\theta)}2^{-\mu(\nu_1+\nu_2/\alpha-\xi^*)}$.

For a function $f \in U^a_{p,\theta}$, we define the mapping S by

$$S(f) := \sum_{\nu\in\mathbb{Z}_+\times\Lambda} S_\nu(f).$$

We obtain

$$f - S(f) = \sum_{\nu_1+\nu_2/\alpha=0}^{\xi^*}(f - S_\nu(f)) + \sum_{\nu_1+\nu_2/\alpha>\xi^*} f_\nu.$$

Therefore, by (22), (24)–(26) and the inequalities $\|f_\nu\|_{B^a_{p,\theta}} \ll \|f\|_{B^a_{p,\theta}}$ we get the following estimates for any $f \in U^a_{p,\theta}$

$$\begin{aligned}
\|f - S(f)\|_{B_{\infty,\tau}} &\le \sum_{\nu_1+\nu_2/\alpha=0}^{\xi^*}\|f - S_\nu(f)\|_{B_{\infty,\tau}} + \sum_{\nu_1+\nu_2/\alpha>\xi^*}\|f_\nu\|_{B_{\infty,\tau}}\\
&\ll \sum_{0\le\nu_1+\nu_2/\alpha\le\xi} A(\nu) + \sum_{\xi<\nu_1+\nu_2/\alpha\le\xi^*} C(\nu) + \sum_{\nu_1+\nu_2/\alpha>\xi^*} E(\nu)\\
&\ll 2^{-r\xi}\xi^{s(1/\tau-1/\theta)}\sum_{0\le\nu_1+\nu_2/\alpha\le\xi} 2^{r(\xi-\nu_1-\nu_2/\alpha)}2^{-2^{(1-\delta)(\xi-\nu_1-\nu_2/\alpha)}}\\
&+ 2^{-r\xi}\xi^{s(1/\tau-1/\theta)}\sum_{\xi<\nu_1+\nu_2/\alpha\le\xi^*}(\nu_1+\nu_2/\alpha-\xi)^{s(1/\tau-1/\theta)}2^{-\beta(\nu_1+\nu_2/\alpha-\xi)}\\
&+ 2^{-r\xi}\xi^{s(1/\tau-1/\theta)}\sum_{\nu_1+\nu_2/\alpha>\xi^*}(\nu_1+\nu_2/\alpha-\xi^*)^{s(1/\tau-1/\theta)}2^{-\mu(\nu_1+\nu_2/\alpha-\xi^*)}\\
&\ll 2^{-r\xi}\xi^{s(1/\tau-1/\theta)} \asymp E_{\theta,\tau}(n).
\end{aligned}$$

This means that

$$\sup_{f\in U^a_{p,\theta}}\big\|f - S(f)\big\| \ll E_{\theta,\tau}(n). \tag{27}$$

Notice that S is a mapping from $U^a_{p,\theta}$ into $B := \sum_{\nu_1+\nu_2/\alpha=0}^{\xi^*} B_\nu$. Moreover, by (21), (23) we have

$$\begin{aligned}
\log|B| \le \sum_{\nu_1+\nu_2/\alpha=0}^{\xi^*}\log|B_\nu| &\ll \sum_{0\le\nu_1+\nu_2/\alpha\le\xi} 2^\xi\xi^s 2^{-\delta(\xi-\nu_1-\nu_2/\alpha)}2^{-\nu_2(1/\alpha-1/r')}\nu_2^{s'}\\
&+ \sum_{\xi<\nu_1+\nu_2/\alpha\le\xi^*}\left(2^{-\delta(\nu_1+\nu_2/\alpha-\xi)}2^\xi\xi^s(\nu_1+\nu_2/\alpha-\xi)^s 2^{-\nu_2(1/\alpha-1/r')}\nu_2^{s'} + \log\binom{m_\nu}{n_\nu}\right).
\end{aligned}$$

Stirling's formula gives

$$\begin{aligned}
\log\binom{m_\nu}{n_\nu} &\le n_\nu\log\frac{bm_\nu}{n_\nu}\\
&\le 2^{-\delta(\nu_1+\nu_2/\alpha-\xi)}2^\xi\xi^s(\nu_1+\nu_2/\alpha-\xi)^s 2^{-\nu_2(1/\alpha-1/r')}\nu_2^{s'}(b + (1\delta)(\nu_1+\nu_2/\alpha-\xi)),
\end{aligned}$$

where b is a constant. Hence,

$$\log|B| \le C'2^{\xi}\xi^{s}\sum_{t=0}^{\infty} 2^{-\delta t}t^{s},$$

where C' is an absolute constant. Setting $C'' := C'\sum_{s=0}^{\infty} 2^{-\delta t}t^{s}$, we obtain $\log|B| \le n$, and consequently $|B| \le 2^{n}$. Let $V^* = \cup_\nu V_\nu^*$, where $V_\nu^* = \{\varphi_{k,s}\}_{s\in Q_k,\ k\in D_\nu}$. By construction, it follows that V^* is a finite subset of V and B is a subset of $M_n(V^*)$.

Summing up, we have constructed a subset B in $M_n(V^*)$ having cardinality does not exceed 2^n and a sampling recovery method $S_n^B := S$ of the form (2) satisfying the inequality (27) and therefore, the upper bound of (16) and (17). ∎

Proof of Theorem 3.1. Notice that

$$\|.\|_{q_1} \ll \|.\|_{q_2},\ q_1 \le q_2. \tag{28}$$

From (28), it is sufficient to prove (15) for $q > 2$. By (12), we can verify that

$$e_n(U^a_{p,\theta},\ L_q) \ll e_n(U^a_{p,\theta},\ B_{q,\min\{q,2\}}).$$

Using this inequality and Theorem 3.2, we get the upper bound of $e_n(U^a_{p,\theta}, L_q)$. ∎

The lower bound of $\rho(U^a_{p,\theta}, L_q)$ in obtained from the following theorem.

Theorem 3.3. *Let $1 < p,\ q < \infty$, $0 < \theta \le \infty$ and $r > 1/p$. Then we have*

$$\rho(U^a_{p,\theta}, L_q) \gg (n/\log^s n)^{-r}(\log n)^{s(1/2-1/\theta)}.$$

Proof. Denote by $U^{a^*}_{p,\theta}(\mathbb{T}^{s+1})$ the unit ball in the space $B^{a^*}_{p,\theta}(\mathbb{T}^{s+1}) \subset L_q(\mathbb{T}^{s+1})$, where $a^* := (a_1, a_2, \ldots, a_{s+1}) = (r, r, \ldots, r) \in \mathbb{R}^{s+1}_+$. In [3] it was proven that

$$\rho_n(U^{a^*}_{p,\theta}(\mathbb{T}^{s+1}), B_{q,\tau}(\mathbb{T}^{s+1})) \gg n^{-r}(\log n)^{s(r+1/2-1/\theta)}.$$

Notice that for any function $f \in L_q(\mathbb{T}^{s+1})$, the function $g : \mathbb{T}^d \to \mathbb{R}$ which is defined by $g(x_1, x_2, \ldots, x_d) = f(x_1, \ldots, x_{s+1})$, belongs to $L_q(\mathbb{T}^d)$. Moreover, if $f \in U^{a^*}_{p,\theta}(\mathbb{T}^{s+1})$, then $g \in U^a_{p,\theta}(\mathbb{T}^d)$. Hence we deduce that

$$\rho_n(U^a_{p,\theta}(\mathbb{T}^d), B_{q,\tau}(\mathbb{T}^d)) \ge \rho_n(U^{a^*}_{p,\theta}(\mathbb{T}^{s+1}), B_{q,\tau}(\mathbb{T}^{s+1})).$$

Therefore,

$$\rho_n(U^a_{p,\theta}(\mathbb{T}^d), B_{q,\tau}(\mathbb{T}^d)) \gg (n/\log^s n)^{-r}(\log n)^{s(1/2-1/\theta)}.$$

The proof is complete. ∎

We now can state and prove the main results (4) and (5) as follows.

Theorem 3.4. *Let* $1 < p,\ q < \infty,\ 0 < \theta \leq \infty$ *and* $r > 1/p$. *Then*

$$\epsilon_n(U^a_{p,\theta}, L_q) \asymp \rho_n(U^a_{p,\theta}, L_q) \asymp n^{-r}(\log n)^{s(r+1/2-1/\theta)}.$$

Moreover, we have also the asymptotic order of optimal methods of adaptive sampling recovery following

$$e_n(U^a_{p,\theta}, L_q) \asymp r_n(U^a_{p,\theta}, L_q) \asymp n^{-r}(\log n)^{s(r+1/2-1/\theta)}.$$

Proof. By Theorem 3.1, Theorem 3.3 and (14), we have

$$\epsilon_n(U^a_{p,\theta}, L_q) \geq \rho_n(U^a_{p,\theta}, L_q) \gg n^{-r}(\log n)^{s(r+1/2-1/\theta)}$$

and

$$\rho_n(U^a_{p,\theta}, L_q) \leq \epsilon_n(U^a_{p,\theta}, L_q) \ll n^{-r}(\log n)^{s(r+1/2-1/\theta)}.$$

Hence

$$\epsilon_n(U^a_{p,\theta}, L_q) \asymp \rho_n(U^a_{p,\theta}, L_q) \asymp n^{-r}(\log n)^{s(r+1/2-1/\theta)}.$$

Using Theorem 3.1 and (14), we get

$$r_n(U^a_{p,\theta}, L_q) \leq e_n(U^a_{p,\theta}, L_q) \ll n^{-r}(\log n)^{s(r+1/2-1/\theta)}.$$

Since Theorem 3.3 and (13), we obtain

$$r_n(U^a_{p,\theta}, L_q) \geq \rho_n(U^a_{p,\theta}, L_q) \gg n^{-r}(\log n)^{s(r+1/2-1/\theta)}.$$

By the last two inequalities, we get

$$e_n(U^a_{p,\theta}, L_q) \asymp r_n(U^a_{p,\theta}, L_q) \asymp n^{-r}(\log n)^{s(r+1/2-1/\theta)}.$$

■

4. CONCLUSION

In this paper, we extend the results in [3] to multivariate Besov-type classes $U^a_{p,\theta}$ of functions having nonuniform mixed smoothness $a \in \mathbb{R}^d_+$ and the problems of entropy numbers $\epsilon_n(U^a_{p,\theta}, L_q)$ and non-linear widths $\rho_n(U^a_{p,\theta}, L_q)$. We obtain the asymptotic order of entropy numbers $\epsilon_n(U^a_{p,\theta}, L_q)$ and non-linear widths $\rho_n(U^a_{p,\theta}, L_q)$. Moreover, we construct corresponding asymptotically optimal methods of nonlinear approximations. In result we obtain the asymptotic order of optimal methods of adaptive sampling recovery of functions in $U^a_{p,\theta}$ by sets of a finite capacity which is measured by their cardinality or pseudo-dimension. In the future we shall consider the above problems in the space $B^A_{p,\theta}$, which is the intersection of spaces $B^a_{p,\theta}$, where A is a finite subset in $\mathbb{R}^d_+$.

ACKNOWLEDGMENT

This work is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under Grant No. 102.01-2017.05.

REFERENCES

[1] Nguyen Manh Cuong and Mai Xuan Thao, "Adaptive sampling recovery of functions with bounded modulus of smoothness," *Acta Mathematica Vietnamica*, vol. 42, pp. 113–127, 2017.

[2] Dinh Dũng, "Continous algorithms in n term approximation and non-linear n-widths," *J. Approx. Theory,* vol. 102, pp. 217–242, 2000.

[3] Dinh Dũng, "Non-linear approximations using sets of finite cardinality or finite pseudo-dimension," *J Complex.*, vol. 17, pp. 467-492, 2001.

[4] Dinh Dũng, "Non-linear sampling recovery based on quasi-interpolant wavelet representations," *Adv. Comput. Math.*, vol. 30, pp. 375-401, 2009.

[5] Dinh Dũng, "Optimal adaptive sampling recovery," *Adv. Comput. Math.,* vol. 31, pp. 1–41, 2011.

[6] Dinh Dũng, "Erratum to: Optimal adaptive sampling recovery," *Adv. Comput. Math.,* vol. 36, pp. 605–606, 2012.

[7] Dinh Dũng, V. N. Temlyakov and T. Ullrich, "Hyperbolic cross approximation," *Advanced Courses in Mathematics* - CRM Barcelona, Birkhauser, 2018.

[8] D. Haussler, "Decision theoretic generalization of the PAC model for neural net and other learning applications," *Inform. Comput.,* vol. 100, pp.78–150, 1982.

[9] A. N. Kolmogorov and V. M. Tikhomirov, "ε-entropy and ε-capacity of sets in function space," *Uspekhi Mat. Nauk,* vol. 14, pp. 3-86, 1959; English transl. in *Amer. Math. Soc. Transl.*, vol. 17, no. 2, 1961.

[10] S.Nikol'skii, *Approximation of Functions of Several Variables and Embedding Theorems*, Springer-Verlag, Berlin, 1975.

[11] D. Pollard, "NSF-CBMS Regional Conference Series in Probability and Statistics", *Empirical Processes: Theory and Applications,* vol. 2, 1990.

[12] J. Ratsaby, V. Maiorov, "The degree of approximation of sets in Euclidean space using sets with bounded Vapnik–Chervonekis dimension," *Discrete Applied Math.*, vol. 86, pp. 81–93, 1998.

[13] J. Ratsaby, V. Maiorov, "On the degree of approximation by manifolds of finite pseudo-dimension," *Constr. Approx.,* vol. 15, pp. 291–300, 1999.

[14] V. N. Vapnik and A. Ya. Chervonekis, "Necessary and sufficient conditions for the uniform convergence of means to their expectations," *Theory Prob. Appl.,* vol. 26, pp. 264-280, 1981.

Received on January 21, 2019
Revised on April 07, 2019