Adaptive algorithms in sampling recovery

Dinh Dũng

Information Technology Institute, Vietnam National University, Hanoi
144 Xuan Thuy, Cau Giay, Hanoi, Vietnam
dinhdung@vnu.edu.vn

February 26, 2011 – Version 0.2

Abstract

We study optimal algorithms in adaptive sampling recovery of smooth functions defined on the unit d-cube \( \mathbb{I}^d := [0,1]^d \). The recovery error is measured in the quasi-norm \( \| \cdot \|_q \) of \( L_q := L_q(\mathbb{I}^d) \). For \( B \) a subset in \( L_q \), we define a sampling recovery algorithm with the free choice of sample points and recovering functions from \( B \) as follows. For each \( f \) from the quasi-normed Besov space \( B_{p,\theta}^\alpha \), we choose \( n \) sample points. This choice defines \( n \) sampled values. Based on these sample points and sampled values, we choose a function from \( B \) for recovering \( f \). The choice of \( n \) sample points and a recovering function from \( B \) for each \( f \in B_{p,\theta}^\alpha \) defines a \( n \)-sampling algorithm \( S_n^B \) by functions in \( B \). We suggest a new approach to investigate the optimal adaptive sampling recovery by \( S_n^B \) in the sense of continuous non-linear \( n \)-widths which is related to \( n \)-term approximation. If \( \Phi = \{ \varphi_k \}_{k \in K} \) is a family of elements in \( L_q \), let \( \Sigma_n(\Phi) \) be the non-linear set of linear combinations of \( n \) free terms from \( \Phi \), that is \( \Sigma_n(\Phi) := \{ \varphi = \sum_{j=1}^n a_j \varphi_{k_j} : k_j \in K \} \). Denote by \( \mathcal{G} \) the set of all families \( \Phi \) in \( L_q \) such that the intersection of \( \Phi \) with any finite dimensional subspace in \( L_q \) is a finite set, and by \( C(B_{p,\theta}^\alpha, L_q) \) the set of all continuous mappings from \( B_{p,\theta}^\alpha \) into \( L_q \). We define the quantity

\[
\nu_n(B_{p,\theta}^\alpha, L_q) := \inf_{\Phi \in \mathcal{G}} \inf_{S_n^B \in C(X,L_q); B=\Sigma_n(\Phi)} \sup_{\| \cdot \|_{B_{p,\theta}^\alpha} \leq 1} \| f - S_n^B(f) \|_q.
\]

Let \( 0 < p, q, \theta \leq \infty \) and \( \alpha > d/p \). Then we prove the asymptotic order

\[
\nu_n(B_{p,\theta}^\alpha, L_q) \asymp n^{-\alpha/d}.
\]

We also obtained the asymptotic order of quantities of optimal recovery by \( S_n^B \) in terms of best \( n \)-term approximation as well of other non-linear \( n \)-widths.

Keywords Adaptive sampling recovery · \( n \)-sampling algorithm · B-spline quasi-interpolant representation · B-spline · Besov space

Mathematics Subject Classifications (2000) 41A46 · 41A05 · 41A25 · 42C40
1 Introduction

The purpose of the present paper is to investigate optimal algorithms in adaptive sampling recovery of functions defined on the unit $d$-cube $I^d := [0,1]^d$. Let $L_q := L_q(I^d)$, $0 < q \leq \infty$, denote the quasi-normed space of functions on $I^d$ with the usual $q$th integral quasi-norm $\| \cdot \|_q$ for $0 < q < \infty$, and the normed space $C(I^d)$ of continuous functions on $I^d$ with the max-norm $\| \cdot \|_{\infty}$ for $q = \infty$. For $0 < p, \theta, q \leq \infty$ and $\alpha > 0$, let $B^\alpha_{p,\theta}$ be the quasi-normed Besov space with smoothness $\alpha$, equipped with the quasi-norm $\| \cdot \|_{B^\alpha_{p,\theta}}$ (see Section 2 for the definition). We consider problems of adaptive sampling recovery of functions from $B^\alpha_{p,\theta}$. The recovery error will be measured in the quasi-norm $\| \cdot \|_q$.

We first recall some well-known non-adaptive sampling recovery algorithms. Let $X$ be a quasi-normed space of functions defined on $I^d$, such that the linear functionals $f \mapsto f(x)$ are continuous for any $x \in I^d$. We assume that $X \subset L_q$ and the embedding $\Id : X \to L_q$ is continuous, where $\Id(f) := f$. Suppose that $f$ is a function in $X$ and $\xi_n = \{x^n_k\}_{k=1}^n$ are $n$ points in $I^d$. We want to approximately recover $f$ from the sampled values $f(x^1), f(x^2), \ldots, f(x^n)$. A classical linear sampling algorithm of recovery is

$$L_n(f) = L_n(\Phi_n, \xi_n, f) := \sum_{k=1}^n f(x^k) \varphi_k,$$  (1.1)

where $\Phi_n = \{\varphi_k\}_{k=1}^n$ are given $n$ functions in $L_q$. A more general sampling algorithm of recovery can be defined as

$$R_n(f) = R_n(H_n, \xi_n, f) := H_n(f(x^1), \ldots, f(x^n)),$$  (1.2)

where $H_n$ is a given mapping from $\mathbb{R}^n$ to $L_q$. Such a sampling algorithm is, in general, non-linear. 

Denote by $SX$ the unit ball in the quasi-normed space $X$. To study optimal sampling algorithms of recovery for $f \in X$ from $n$ their values by algorithms of the form (1.2), one can use the quantity

$$g_n(X, L_q) := \inf_{H_n, \xi_n} \sup_{f \in SX} \| f - R_n(H_n, \xi_n, f) \|_q,$$

where the infimum is taken over all sequences $\xi_n = \{x^n_k\}_{k=1}^n$ and all mappings $H_n$ from $\mathbb{R}^n$ into $L_q$.

We use the notations: $x_+ := \max(0, x)$ for $x \in \mathbb{R}$; $A_n(f) \ll B_n(f)$ if $A_n(f) \leq C B_n(f)$ with $C$ an absolute constant not depending on $n$ and/or $f \in W$, and $A_n(f) \asymp B_n(f)$ if $A_n(f) \asymp B_n(f)$ and $A_n(f) \ll A_n(f)$. It is known the following result (see [13, 23, 26, 27, 31] and references there). If $0 < p, \theta, q \leq \infty$ and $\alpha > d/p$, then there is a linear sampling recovery method $L^*_n$ of the form (1.1) such that

$$g_n(B^\alpha_{p,\theta}, L_q) \asymp \sup_{f \in S B^\alpha_{p,\theta}} \| f - L^*_n(f) \|_q \asymp n^{-\alpha/d+(1/p-1/q)_+}.  \tag{1.3}$$

This result says that the linear sampling algorithm $L^*_n$ is asymptotically optimal in the sense that any sampling algorithm $R_n$ of the form (1.2) does not give the rate of convergence better than $L^*_n$.

Sampling algorithms of recovery of the form (1.2) which may be linear or non-linear are non-adaptive, i.e., the points $\xi_n = \{x^n_k\}_{k=1}^n$ at which the values $f(x^1), \ldots, f(x^n)$ are sampled, and the
method for construction of recovering functions are the same for all functions \( f \in X \). Let us introduce a setting of adaptive sampling recovery.

If \( B \) is a subset in \( L_q \), we define a sampling algorithms of recovery with the free choice of sample points and recovering functions from \( B \) as follows. For each \( f \in X \) we choose a set of \( n \) sample points. This choice defines a collection of \( n \) sampled values. Based on the information of these sampled values, we choose a function from \( B \) for recovering \( f \). The choice of \( n \) sample points and a recovering function from \( B \) for each \( f \in X \) defines a sampling algorithms of recovery \( S_n^B \) by functions in \( B \). More precisely, a formal definition of \( S_n^B \) is given as follows. Denote by \( T^n \) the set of subsets \( \xi \) in \( \mathbb{I}^d \) of cardinality at most \( n \), \( V^n \) the set of subsets \( \eta \) in \( \mathbb{R} \times \mathbb{I}^d \) of cardinality at most \( n \). Let \( T_n \) be a mapping from \( X \) into \( T^n \). Then \( T_n \) generates an \( n \)-sampling operator \( I_n \) from \( X \) into \( V^n \) which is defined as follows. If \( T_n(f) = \{x^1, ..., x^n\} \) then \( I_n(f) = \{(f(x^1), x^1), ..., (f(x^n), x^n)\} \). Let \( P_n^B \) a mapping from \( V^n \) into \( B \). Then the pair \((I_n, P_n^B)\) generates the mapping \( S_n^B \) from \( X \) into \( B \), by the formula

\[
S_n^B(f) := P_n^B(I_n(f)),
\]

which defines a \( n \)-sampling algorithm with the free choice of \( n \) sample points and approximant from \( B \). We call the mapping \( P_n^B \) a recovering operator.

Clearly, a linear sampling algorithm \( L_n(\Phi_n, \xi_n, \cdot) \) defined in (1.1) is a particular case of \( S_n^B \). We are interested in adaptive \( n \)-sampling algorithms \( S_n^B \) of special form which are an extension of \( L_n(\Phi_n, \xi_n, \cdot) \) to an \( n \)-sampling algorithm with the free choice of \( n \) sample points and \( n \) functions \( \Phi_n = \{\varphi_k\}_{k=1}^n \) for each \( f \in X \). To this end we let \( \Phi = \{\varphi_k\}_{k \in K} \) be a family of elements in \( L_q \), and consider the non-linear set \( \Sigma_n(\Phi) \) of linear combinations of \( n \) free terms from \( \Phi \), that is

\[
\Sigma_n(\Phi) := \{ \varphi = \sum_{j=1}^n a_j \varphi_{k_j} : \ k_j \in K \}. 
\]

Then for \( B = \Sigma_n(\Phi) \), an \( n \)-sampling algorithm \( S_n^B \) is of the following form

\[
S_n^B(f) = \sum_{k \in Q(\eta)} a_k(\eta) \varphi_k, \tag{1.5}
\]

where \( \varphi_k \in \Phi, \eta = I_n(f), Q(\eta) \subset K \) with \( |Q(\eta)| \leq n \) and \( a_k \) are functions on \( V^n \).

We want to choose an \( n \)-sampling algorithm \( S_n^B \) 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 \( f \). To investigate the optimality of (non-continuous) \( n \)-term recovery of functions \( f \) from the quasi-normed space \( X \) by \( n \)-sampling algorithms \( S_n^B \) of the form (1.5), we introduce the quantity \( s_n(X, \Phi, L_q) \) as follows:

\[
s_n(X, \Phi, L_q) := \inf_{s_B^B : B = \Sigma_n(\Phi)} \sup_{f \in SX} \|f - S_n^B(f)\|_q. \tag{1.6}
\]

The definition (1.6) corrects a definition of \( s_n(X, \Phi, L_q) \) which has been introduced and denoted by \( \nu_n(SX, \Phi)_q \) and \( s_n(SX, \Phi)_q \) in [17] and [18], respectively. The quantity \( s_n(X, \Phi, L_q) \) is directly related to non-linear \( n \)-term approximation. We refer the reader to [7], [32] for surveys on various aspects in the last direction.
The quantity $s_n(X, \Phi, L_q)$ depends on the family $\Phi$ and therefore, is not absolute in the sense of $n$-widths or optimal algorithms. We suggest an approach to investigate the optimal adaptive sampling recovery by $S^n_B$ in the sense of continuous non-linear $n$-widths which is related to $n$-term approximation too. Namely, we consider the optimality in the restriction with only $n$-sampling algorithms of recovery $S^n_B$ of the form (1.5) and with a continuity assumption on them. Continuity assumptions on approximation and recovery algorithms have their origin in the very old Alexandroff $n$-width which characterizes best continuous approximation algorithm by $n$-dimensional topological complexes $[1]$ (see also $[33]$ for details). Later on, (continuous) manifold $n$-width was introduced by in $[8]$, $[24]$, and investigated in $[12]$, $[9]$, $[20]$, $[14]$, $[15]$, $[16]$. Several continuous $n$-widths based on continuous algorithms of $n$-term approximation, were introduced and studied in $[14]$, $[15]$, $[16]$. The continuity assumption is quite natural: the closer objects are the closer their reconstructions should be. A first look seems that a continuity restriction may decrease the choice of approximants. However, in most cases it does not weaken the rate of the corresponding approximation. Continuous and non-continuous algorithms of nonlinear approximation give the same asymptotic order. This motivate us to consider continuous $n$-sampling algorithms of recovery $S^n_B$. Since we assume that functions to be recovered are living in the quasi-normed space $X$ and the recovery error is measured in the quasi-normed space $L_q$, the requirement that $S^n_B \in C(X, L_q)$ is quite reasonable. (Here and in what follows, $C(X, Y)$ denotes the set of all continuous mappings from $X$ into $Y$ for the quasi-metric spaces $X, Y$). This leads to the following definition.

Denote by $G$ the set of all families $\Phi$ in $L_q$ such that the intersection of $\Phi$ with any finite dimensional subspace in $L_q$ is a finite set. We define the quantity

$$\nu_n(X, L_q) := \inf_{\Phi \in G} \sup_{B = \Sigma_n(\Phi)} \| f - S^n_B(f) \|_q.$$ 

The restriction $\Phi \in G$ in the definition of $\nu_n(X, L_q)$ is minimal and natural for all well-known approximation systems.

Another way to study optimal adaptive (non-continuous) $n$-sampling algorithms of recovery $S^n_B$ in the sense of nonlinear $n$-widths has been proposed in $[17]$, $[18]$. In this approach, $B$ is required to have a finite capacity which is measured by their cardinality or pseudo-dimension. Given a family $B$ of subsets in $L_q$, we consider optimal sampling recoveries by $B$ from $B$ in terms of the quantity

$$R_n(W, B)_q := \inf_{B \in B} \sup_{f \in W} \| f - S^n_B(f) \|_q.$$ 

We assume a restriction on the sets $B \in B$, requiring that they should have, in some sense, a finite capacity. In the present paper, the capacity of $B$ is measured by its cardinality or pseudo-dimension. This reasonable restriction would provide nontrivial lower bounds of asymptotic order of $R_n(W, B)_q$ for well known function classes $W$. Denote $R_n(W, B)_q$ by $e_n(W)_q$ if $B$ in (1.7) is the family of all subsets $B$ in $L_q$ such that $|B| \leq 2^n$, where $|B|$ denotes the cardinality of $B$, and by $r_n(W)_q$ if $B$ in (1.7) is the family of all subsets $B$ in $L_q$ of pseudo-dimension at most $n$. The definition (1.7) corrects definitions of $e_n(W)_q$ and $r_n(W)_q$ introduced in $[18]$.

The quantity $e_n(W)_q$ is related to the entropy $n$-width (entropy number) $\varepsilon_n(W)_q$ which is the functional inverse of the classical $\varepsilon$-entropy introduced by Kolmogorov and Tikhomirov $[22]$. The
quantity \( r_n(W)_q \) is related to the non-linear \( n \)-width \( \rho_n(W)_q \) introduced recently by Ratsaby and Maiorov \[29\]. (See the definition of \( \varepsilon_n(W)_q \) and \( \rho_n(W)_q \) in Section 5).

The pseudo-dimension of a set \( B \) of real-valued functions on a set \( \Omega \), is defined as follows. For a real number \( t \), let \( \text{sgn}(t) \) be 1 for \( t > 0 \) and \(-1\) otherwise. For \( x \in \mathbb{R}^n \), let \( \text{sgn}(x) = (\text{sgn}(x_1), \text{sgn}(x_2), \ldots, \text{sgn}(x_n)) \). The pseudo-dimension of \( B \) is defined as the largest integer \( n \) such that there exist points \( a_1, a_2, \ldots, a_n \) in \( \Omega \) and \( b \in \mathbb{R}^n \) such that the cardinality of the set

\[
\{ \text{sgn}(y) : y = (f(a_1) + b_1, f(a_2) + b_2, \ldots, f(a_n) + b_n), \ f \in B \}
\]

is \( 2^n \). If \( n \) is arbitrarily large, then the the pseudo-dimension of \( B \) is infinite. Denote the pseudo-dimension of \( B \) by \( \text{dim}_p(B) \). The notion of pseudo-dimension was introduced by Pollard \[28\] and later Haussler \[21\] as an extension of the VC-dimension \[34\], suggested by Vapnik-Chervonenkis for sets of indicator functions. The pseudo-dimension and VC-dimension measure the capacity of a set of functions and are related to its \( \varepsilon \)-entropy (see also \[29\], \[30\]). If \( B \) is a \( n \)-dimensional linear manifold of real-valued functions on \( \Omega \), then \( \text{dim}_p(B) = n \) (see \[21\]).

We say that \( p, q, \theta, \alpha \) satisfy Condition (1.8) if

\[
0 < p, q, \theta, \alpha \leq \infty, \ 0 < \alpha < \infty, \ \text{and there holds one of the following restrictions:}
\]

(i) \( \alpha > d/p; \)

(ii) \( \alpha = d/p, \ \theta \leq \min(1, p), \ p, q < \infty. \) \hspace{1cm} (1.8)

Let \( M \) be the set of B-splines which are the tensor product of integer translated dilations of the centered cardinal spline of order \( 2r \). (see the definition in Section 2).

The main results of the present paper are read as follows.

**Theorem 1.1** Let \( p, q, \theta, \alpha \) satisfy the Condition (1.8) and \( \alpha < 2r \). Then for the \( d \)-variable Besov space \( B_{p,\theta}^{\alpha} \), there is the following asymptotic order

\[
s_n(B_{p,\theta}^{\alpha}, M, L_q) \asymp \nu_n(B_{p,\theta}^{\alpha}, L_q) \asymp r_n(SB_{p,\theta}^{\alpha})_q \asymp e_n(SB_{p,\theta}^{\alpha})_q \asymp n^{-\alpha/d}. \hspace{1cm} (1.9)
\]

Comparing this asymptotic order with (1.3), we can see that for \( 0 < p < q \leq \infty \), the asymptotic order of optimal adaptive sampling recovery in terms of the quantities \( s_n, \nu_n, e_n \) and \( r_n \), is better than the asymptotic order of any non-adaptive \( n \)-sampling algorithm of recovery of the form (1.2).

To prove the upper bound for (1.9), we use a B-spline quasi-interpolant representation of functions in the Besov space \( B_{p,\theta}^{\alpha} \) associated with some equivalent discrete quasi-norm \[17\], \[18\]. On the basis of this representation we construct corresponding asymptotically optimal \( n \)-sampling algorithms of recovery which give the upper bound for (1.9). The lower bound of (1.9) is established by the lower estimating of the smaller related \( n \)-widths and the quantity of \( n \)-term approximation.

The paper is organized as follows.

In Section 2, we give a definition of quasi-interpolant for functions on \( \mathbb{I}^d \), describe a B-spline quasi-interpolant representation for Besov spaces \( B_{p,\theta}^{\alpha} \) with a discrete quasi-norm in terms of the
coefficient functionals. The proof of the asymptotic order of \( \nu_n(B_{p, \theta}^\alpha, L_q) \) in Theorem 1.1 is given in Sections 3 and 4. More precisely, in Section 3 we construct asymptotically optimal adaptive \( n \)-sampling algorithms of recovery which give the upper bound for \( \nu_n(B_{p, \theta}^\alpha, L_q) \) (Theorem 3.1). In Section 4 we prove the lower bound for \( \nu_n(B_{p, \theta}^\alpha, L_q) \) (Theorem 4.1). In Section 5, we prove the asymptotic order of \( s_n(B_{p, \theta}^\alpha, M, L_q) \), \( r_n(B_{p, \theta}^\alpha, q) \) and \( e_n(B_{p, \theta}^\alpha, q) \) in Theorem 1.1.

2 Preliminary background

For a given natural number \( r \), let \( M \) be the centered B-spline of even order \( 2r \) with support \([-r, r]\) and knots at the integer points \(-r, ..., 0, ..., r\) and define the B-spline

\[
M_{k,s}(x) := M(2^k x - s),
\]

for a non-negative integer \( k \) and \( s \in \mathbb{Z} \). To get the \( d \)-variable B-spline \( M_{k,s} \) for a non-negative integer \( k \) and \( s \in \mathbb{Z}^d \), we let

\[
M(x) := \prod_{i=1}^d M(x_i), \quad x = (x_1, x_2, ..., x_d),
\]

and

\[
M_{k,s}(x) := M(2^k x - s).
\]

Denote by \( \mathbf{M} \) the set of all \( M_{k,s} \) which do not vanish identically on \( \mathbb{R}^d \).

Let \( \Lambda = \{ \lambda(j) \}_{j \in P^d(\mu)} \) be a finite even sequence, i.e., \( \lambda(-j) = \lambda(j) \), where \( P^d(\mu) := \{ j \in \mathbb{Z}^d : |j_i| \leq \mu, \ i = 1, 2, ..., d \} \). We define the linear operator \( Q \) for functions \( f \) on \( \mathbb{R}^d \) by

\[
Q(f, x) := \sum_{s \in \mathbb{Z}^d} \Lambda(f, s) M(x - s),
\]

where

\[
\Lambda(f, s) := \sum_{j \in P^d(\mu)} \lambda(j) f(s - j).
\]

The operator \( Q \) is bounded in \( C(\mathbb{R}^d) \) and

\[
\|Q(f)\|_{C(\mathbb{R}^d)} \leq \|\Lambda\| \|f\|_{C(\mathbb{R}^d)}
\]

for each \( f \in C(\mathbb{R}^d) \), where

\[
\|\Lambda\| = \sum_{j \in P^d(\mu)} |\lambda(j)|.
\]

Moreover, \( Q \) is local in the following sense. There is a positive number \( \delta > 0 \) such that for any \( f \in C(\mathbb{R}^d) \) and \( x \in \mathbb{R}^d \), \( Q(f, x) \) depends only on the value \( f(y) \) at a finite number of points \( y \) with \( |y_i - x_i| \leq \delta, \ i = 1, 2, ... d \). We will require \( Q \) to reproduce the space \( P^d_{2r-1} \) of polynomials of order at most \( 2r - 1 \) in each variable \( x_i \), that is,

\[
Q(p) = p, \ p \in P^d_{2r-1}.
\]
An operator \( Q \) of the form (2.1)–(2.2) reproducing \( P_{2r-1} \), is called a \textit{quasi-interpolant in} \( C(\mathbb{R}^d) \).

There are many ways to construct quasi-interpolants. A method of construction via Neumann series was suggested by Chui and Diamond [4] (see also [3] p. 100–109). De Bore and Fix [5] introduced another quasi-interpolant based on the values of derivatives. The reader can see also the books [3], [6] for surveys on quasi-interpolants. The most important cases of \( d \)-variate quasi-interpolants \( Q \) are those where the functional \( \Lambda \) is the tensor product of such \( d \) univariate functionals. Let us give some examples of univariate quasi-interpolants. The simplest example is a piecewise linear quasi-interpolant is defined for \( r = 1 \) by

\[
Q(f, x) := \sum_{s \in \mathbb{Z}} f(s)M(x - s),
\]

where \( M \) is the symmetric piecewise linear B-spline with support \([-1, 1]\) and knots at the integer points \(-1, 0, 1\). This quasi-interpolant is also called nodal and directly related to the classical Faber-Schauder basis [19]. Another example is the cubic quasi-interpolant defined for \( r = 2 \) by

\[
Q(f, x) := \sum_{s \in \mathbb{Z}} \frac{1}{6} \{-f(s - 1) + 8f(s) - f(s + 1)\}M(x - s),
\]

where \( M \) is the symmetric cubic B-spline with support \([-2, 2]\) and knots at the integer points \(-2, -1, 0, 1, 2\).

Let \( \Omega = [a, b]^d \) be a \( d \)-cube in \( \mathbb{R}^d \). Denote by \( L_p(\Omega) \) the quasi-normed space of functions on \( \Omega \) with the usual \( p \)th integral quasi-norm \( \| \cdot \|_{p, \Omega} \) for \( 0 < p < \infty \), and the normed space \( C(\Omega) \) of continuous functions on \( \Omega \) with the max-norm \( \| \cdot \|_{\infty, \Omega} \) for \( p = \infty \). If \( \tau \) be a number such that \( 0 < \tau \leq \min(p, 1) \), then for any sequence of functions \( \{f_k\} \) there is the inequality

\[
\left\| \sum f_k \right\|_{p, \Omega}^\tau \leq \sum \|f_k\|_{p, \Omega}^\tau. \tag{2.3}
\]

We introduce Besov spaces of smooth functions and give necessary knowledge of them. The reader can read this and more details about Besov spaces in the books [2], [25], [10]. Let

\[
\omega_l(f, t)_p := \sup_{|h| < t} \| \Delta^l_h f \|_{p, \mathbb{I}^d(lh)}
\]

be the \( l \)th modulus of smoothness of \( f \) where \( \mathbb{I}^d(lh) := \{x : x + lh \in \mathbb{I}^d\} \), and the \( l \)th difference \( \Delta^l_h f \) is defined by

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

For \( 0 < p, \theta \leq \infty \) and \( 0 < \alpha < l \), the Besov space \( B^\alpha_{p, \theta} \) is the set of functions \( f \in L_p \) for which the Besov quasi-semi-norm \( |f|_{B^\alpha_{p, \theta}} \) is finite. The Besov quasi-semi-norm \( |f|_{B^\alpha_{p, \theta}} \) is given by

\[
|f|_{B^\alpha_{p, \theta}} := \begin{cases} 
(f_0 \{t^{-\alpha} \omega_l(f, t)_p \theta dt/t \}^{1/\theta}), & \theta < \infty, \\
\sup_{t > 0} t^{-\alpha} \omega_l(f, t)_p, & \theta = \infty.
\end{cases}
\]

7
The Besov quasi-norm is defined by

\[ B(f) = \|f\|_{B^\alpha_{p,\theta}} := \|f\|_p + |f|_{B^\alpha_{p,\theta}}. \]

If \( \{f_k\}_{k=0}^\infty \) is a sequence whose component functions \( f_k \) are in \( L_p \), for \( 0 < p, \theta \leq \infty \) and \( \beta \geq 0 \) we use the \( b^\beta_{\theta}(L_p) \) “quasi-norms”

\[ \|\{f_k\}\|_{b^\beta_{\theta}(L_p)} := \left( \sum_{k=0}^\infty (2^{\beta k} \|f_k\|_p)^\theta \right)^{1/\theta} \]

with the usual change to a supremum when \( \theta = \infty \). When \( \{f_k\}_{k=0}^\infty \) is a positive sequence, we replace \( \|f_k\|_p \) by \( |f_k| \) and denote the corresponding quasi-norm by \( \|\{f_k\}\|_{b^\beta_{\theta}} \).

For the Besov space \( B^\alpha_{p,\theta} \), there is the following quasi-norm equivalence

\[ B(f) \asymp B_1(f) := \|\{\omega(f, 2^{-k})\}\|_{b^\alpha_{\theta}} + \|f\|_p. \]

In the present paper, we study the sampling recovery of functions from the Besov space \( B^\alpha_{p,\theta} \) with some restriction on the smoothness \( \alpha \). Namely, we assume that \( \alpha > d/p \). This inequality provides the compact embedding of \( B^\alpha_{p,\theta} \) into \( C(\mathbb{I}^d) \). In addition, we also consider the restriction \( \alpha = d/p \) and \( \theta \leq \min(1, p) \) which is a sufficient condition for the continuous embedding of \( B^\alpha_{p,\theta} \) into \( C(\mathbb{I}^d) \). In both these cases, \( B^\alpha_{p,\theta} \) can be considered as a subset in \( C(\mathbb{I}^d) \).

If \( Q \) of is a quasi-interpolant of the form (2.1)–(2.2), for \( h > 0 \) and a function \( f \) on \( \mathbb{R}^d \), we define the operator \( Q(\cdot; h) \) by

\[ Q(f; h) = \sigma_h \circ Q \circ \sigma_{1/h}(f), \]

where \( \sigma_h(f, x) = f(x/h) \). By definition it is easy to see that

\[ Q(f, x; h) = \sum_k \Lambda(f, k; h) M(h^{-1}x - k), \]

where

\[ \Lambda(f, k; h) := \sum_{j \in P^d(\mu)} \lambda(j) f(h(k - j)). \]

The operator \( Q(\cdot; h) \) has the same properties as \( Q \): it is a local bounded linear operator in \( \mathbb{R}^d \) and reproduces the polynomials from \( P^d_{2r-1} \). Moreover, it gives a good approximation of smooth functions [6, p. 63–65]. We will also call it a quasi-interpolant for \( C(\mathbb{R}^d) \).

The quasi-interpolant \( Q(\cdot; h) \) is not defined for a function \( f \) on \( \mathbb{I}^d \), and therefore, not appropriate for an approximate sampling recovery of \( f \) from its sampled values at points in \( \mathbb{I}^d \). An approach to construct a quasi-interpolant for a function on \( \mathbb{I}^d \) is to extend it by interpolation Lagrange polynomials. This approach has been proposed in [17] for the univariate case. Let us recall it.

For a non-negative integer \( m \), we put \( x_j = j 2^{-m}, j \in \mathbb{Z} \). If \( f \) is a function on \( \mathbb{I} \), let \( U_m(f) \) and \( V_m \) be the \((2r-1)\)th Lagrange polynomials interpolating \( f \) at the \( 2r \) left end points \( x_0, x_1, \ldots, x_{2r-1} \),
and 2r right end points $x_{2m−2r+1}, x_{2m−2r+3}, ..., x_{2m}$, of the interval $I$, respectively. The function $f_m$ is defined as an extension of $f$ on $\mathbb{R}$ by the formula

$$f_m(x) := \begin{cases} 
U_m(f, x), & x < 0, \\
f(x), & 0 \leq x \leq 1, \\
V_m(f, x), & x > 1.
\end{cases}$$

Obviously, if $f$ is continuous on $I$, then $f_m$ is a continuous function on $\mathbb{R}$. Let $Q$ be a quasi-interpolant of the form (2.1)-(2.2) in $C(\mathbb{R})$. We introduce the operator $Q_m$ by putting

$$Q_m(f, x) := Q(f_m(x; 2^{−m}), x \in I,$$

for a function $f$ on $I$. By definition we have

$$Q_m(f, x) = \sum_{s \in J(m)} a_{m,s}(f) M_{m,s}(x), \forall x \in I,$$

where $J(m) := \{ s \in \mathbb{Z} : -r < s < 2^m + r \}$ is the set of $s$ for which $M_{m,s}$ do not vanish identically on $I$, and

$$a_{m,s}(f) := \Lambda(f, s; 2^{−m}) = \sum_{|j| \leq \mu} \lambda(j) f_m(2^{−m}(s − j)).$$

The multivariate operator $Q_m$ is defined for functions $f$ on $I^d$ by

$$Q_m(f, x) := \sum_{s \in J(m)} a_{m,s}(f) M_{m,s}(x), \forall x \in I^d,$$

where $J^d(m) := \{ s \in \mathbb{Z}^d : -r < s_i < 2^m + r, i = 0, 1, ..., d \}$ is the set of $s$ for which $M_{m,s}$ do not vanish identically on $I^d$, and

$$a_{m,s}(f) = a_{m,s_1}(a_{m,s_2}(...a_{m,s_d}(f))), \quad (2.4)$$

where the univariate functional $a_{m,s}$ is applied to the univariate function $f$ by considering $f$ as a function of variable $x_i$ with the other variables held fixed. Moreover, the number of the terms in $Q_m(f)$ is of the size $\approx 2^{dm}$.

The operator $Q_m$ is a local bounded linear mapping in $C(I^d)$ and reproducing $\mathcal{P}_{2r−1}^d$. In particular,

$$\|Q_m(f)\|_{C(I^d)} \leq C\|\Lambda\|\|f\|_{C(I^d)} \quad (2.5)$$

for each $f \in C(I^d)$, with a constant $C$ not depending on $m$, and,

$$Q_m(p^*) = p, \quad p \in \mathcal{P}_{2r−1}^d, \quad (2.6)$$

where $p^*$ is the restriction of $p$ on $I^d$. The multivariate operator $Q_m$ is called a quasi-interpolant in $C(I^d)$.

From (2.5) and (2.6) we can see that

$$\|f - Q_m(f)\|_{C(I^d)} \to 0, \quad m \to \infty. \quad (2.7)$$
Put $M(m) := \{ M_{m,s} \in \mathbf{M} : s \in J^d(m) \}$ and $V(m) := \text{span}M(m)$. If $0 < p \leq \infty$, for all non-negative integers $m$ and all functions

$$g = \sum_{s \in J^d(m)} a_s M_{m,s}  \tag{2.8}$$

from $V(m)$, there is the norm equivalence

$$\| g \|_p \asymp 2^{-dm/p} \| \{ a_s \} \|_{p,m}, \tag{2.9}$$

where

$$\| \{ a_s \} \|_{p,m} := \left( \sum_{s \in J^d(m)} |a_s|^p \right)^{1/p}$$

with the corresponding change when $p = \infty$.

For non-negative integer $k$, let the operator $q_k$ be defined by

$$q_k(f) := Q_k(f) - Q_{k-1}(f) \quad \text{with} \quad Q_1(f) := 0. \tag{2.10}$$

From (2.6) and (2.7) it is easy to see that a continuous function $f$ has the decomposition

$$f = \sum_{k=0}^{\infty} q_k(f)$$

with the convergence in the norm of $C(\mathbb{I}^d)$. By using the B-spline refinement equation, one can represent the component functions $q_k(f)$ as

$$q_k(f) = \sum_{s \in J^d(k)} c_{k,s}(f) M_{k,s}, \tag{2.11}$$

where $c_{k,s}$ are certain coefficient functionals of $f$, which are defined as follows. For the univariate case, we put

$$c_{k,s}(f) := a_{k,s}(f) - a_{k,s}^0(f), \quad k > 0, \tag{2.12}$$

and

$$a_{k,s}^0(f) := 2^{-2r+1} \sum_{(m,j) \in C(k,s)} \binom{2r}{j} a_{k-1,m}(f), \quad k > 0, \quad a_{0,s}^0(f) := 0.$$

and

$$C(k,s) := \{(m,j) : 2m + j - r = s, \ m \in J(k-1), \ 0 \leq j \leq 2r, \ k > 0, \ C(0,s) := \{0\}. \tag{2.13}$$

For the multivariate case, we define $c_{k,s}$ in the manner of the definition (2.11) by

$$c_{k,s}(f) := c_{k,s_1}(c_{k,s_2}(\ldots c_{k,s_d}(f))). \tag{2.14}$$

For functions $f$ on $\mathbb{I}^d$, we introduce the quasi-norms:

$$B_2(f) := \| \{ q_k(f) \} \|_{b^0_\infty (L_p)};$$

$$B_3(f) := \left( \sum_{k=0}^{\infty} (2^{(a-d)p})^k \| \{ c_{k,s}(f) \} \|_{p,k}^\theta \right)^{1/\theta}. \tag{2.15}$$

The following theorem has been proven in [18].
Theorem 2.1 Let $0 < p, \theta \leq \infty$ and $d/p < \alpha < 2r$. Then the hold the following assertions.

(i) A function $f \in B^\alpha_{p,\theta}$ can be represented by the mixed B-spline series

$$f = \sum_{k=0}^{\infty} q_k(f) = \sum_{k=0}^{\infty} \sum_{s \in J^d(k)} c_{k,s}(f) M_{k,s},$$

satisfying the convergence condition

$$B_2(f) \asymp B_3(f) \ll B(f),$$

where the coefficient functionals $c_{k,s}(f)$ are explicitly constructed by formula (2.11)–(2.12) as linear combinations of at most $N$ function values of $f$ for some $N \in \mathbb{N}$ which is independent of $k, s$ and $f$.

(ii) If in addition, $\alpha < \min(2r, 2r - 1 + 1/p)$, then a continuous function $f$ on $I^d$ belongs to the Besov space $B^\alpha_{p,\theta}$ if and only if $f$ can be represented by the series (2.13). Moreover, the Besov quasi-norm $B(f)$ is equivalent to one of the quasi-norms $B_2(f)$ and $B_3(f)$.

3 Adaptive continuous sampling recovery

In this section, we construct asymptotically optimal algorithms which give the upper bound of $\nu_n(B^\alpha_{p,\theta}, L_q)$ in Theorem 1.1. We need some auxiliary lemmas.

Lemma 3.1 Let $p, q, \theta, \alpha$ satisfy Condition (1.8) and $\alpha < 2r$. Then $Q_m \in C(B^\alpha_{p,\theta}, L_q)$ and for any $f \in B^\alpha_{p,\theta}$, we have

$$\|Q_m(f)\|_q \leq C f_{B^\alpha_{p,\theta}},$$

$$\|f - Q_m(f)\|_q \leq C' 2^{-(\alpha - d(1/p - 1/q)_+)} m \|f\|_{B^\alpha_{p,\theta}},$$

with some constants $C, C'$ depending at most on $d, r, p, q, \theta$ and $\|\Lambda\|$.

Proof. We first prove (3.1). The case when the Condition (ii) holds has been proven in [18]. Let us prove the case when the Condition (i) takes place. We put $\alpha' := \alpha - d(1/p - 1/q)_+ > 0$. For an arbitrary $f \in B^\alpha_{p,\theta}$, by the representation (2.13) and (2.3) we have

$$\|f - Q_m(f)\|_q^\tau \ll \sum_{k > m} \|q_k(f)\|_q^\tau$$

with any $\tau \leq \min(q, 1)$. From (2.10) and (2.8)–(2.9) we derive that

$$\|q_k(f)\|_q \ll 2^{(1/p - 1/q)_+ k} \|q_k(f)\|_p$$

(3.3)
Therefore, if \( \theta \leq \min(q,1) \), then we get
\[
\|f - Q_m(f)\|_q \ll \left( \sum_{k>m} \|q_k(f)\|_{q}^\theta \right)^{1/\theta} 
\leq \left( \sum_{k>m} \{2^{(1/p-1/q)}|k\|q_k(f)\|_{p}\}^\theta \right)^{1/\theta} 
\leq 2^{-\alpha'm} \left( \sum_{k>m} \{2^{\alpha k}\|q_k(f)\|_{p}\}^\theta \right)^{1/\theta} 
\ll 2^{-\alpha'm} B_2(f) \ll 2^{-\alpha'm} \|f\|_{B_p^\alpha}. 
\]
Further, if \( \theta > \min(q,1) \), then from (3.2) and (3.3) it follows that
\[
\|f - Q_m(f)\|_{q^*} \ll \sum_{k>m} \|q_k(f)\|_{q^*} 
\ll \sum_{k>m} \{2^{\alpha k}\|q_k(f)\|_{q}\}^{q^*} \{2^{-\alpha'k}\}^{q^*}, 
\]
where \( q^* = \min(q,1) \). Putting \( \nu := \theta/q^* \) and \( \nu' := \nu/(\nu - 1) \), by Hölder’s inequality obtain
\[
\|f - Q_m(f)\|_{q^*} \ll \left( \sum_{k>m} \{2^{\alpha k}\|q_k(f)\|_{q}\}^{q^*\nu} \right)^{1/\nu} \left( \sum_{k>m} \{2^{-\alpha'k}\}^{q^*\nu'} \right)^{1/\nu'} 
\ll \{B_2(f)\}^{q^*} \{2^{-\alpha'm}\}^{q^*} \ll \{2^{-\alpha'm}\}^{q^*} \|f\|_{B_p^\alpha}^{q^*}. 
\]
Thus, the inequality (3.1) is completely proven.

Next, by use of the inequality
\[
\|Q_m(f)\|_{q^*} \ll \sum_{k\leq m} \|q_k(f)\|_{q^*} 
\]
with any \( \tau \leq \min(q,1) \), in a similar way we can prove that \( \|Q_m(f)\|_q \ll \|f\|_{B_p^\alpha}^\tau \) and therefore, the inclusion \( Q_m \in C(B_p^\alpha, L_q) \).

Lemma 3.2 For functions \( f \) on \( \mathbb{I}^d \), \( Q_k \) defines a linear \( n \)-sampling algorithm of the form (1.1). More precisely,
\[
Q_k(f) = L_n(f) = \sum_{s \in I^d(k)} f(2^{-k}j)\psi_{k,j},
\]
where \( n := (2^k + 1)^d \), \( \psi_{k,j} \) are explicitly constructed as linear combinations of at most \( (2\mu + 2)^d \) B-splines \( M_{k,s} \), and \( I^d(k) := \{s \in \mathbb{Z}_+^d : 0 \leq s_i \leq 2^k, \ i = 1, ..., d\} \).
Proof. For univariate functions the coefficient functionals $a_{k,s}(f)$ can be rewritten as

$$a_{k,s}(f) = \sum_{|s-j| \leq \mu} \lambda(s-j)f(2^{-k}j) = \sum_{j \in P(k,s)} \lambda_{k,s}(j)f(2^{-k}j),$$

where $\lambda_{k,s}(j) := \lambda(s-j)$ and $P(k,s) = P_s(\mu) := \{ j \in \{0,2^k\} : s-j \in P(\mu) \}$ for $\mu \leq s \leq 2^k - \mu$; $\lambda_{k,s}(j)$ is a linear combination of no more than $\max(2^r, 2\mu + 1) \leq 2\mu + 2$ coefficients $\lambda(j), j \in P(\mu)$, for $s < \mu$ or $s > 2^k - \mu$ and

$$P(k,s) \subset \begin{cases} P_s(\mu) \cup \{0,2r-1\}, & s < \mu, \\ P_s(\mu) \cup \{2^k-2r+1,2^k\}, & s > 2^k - \mu. \end{cases}$$

If $j \in P(k,s)$, we have $|j-s| \leq \max(2r, \mu) \leq 2\mu + 2$. Therefore, $P(k,s) \subset P_s(\mu)$, and we can rewrite the coefficient functionals $a_{k,s}(f)$ in the form

$$a_{k,s}(f) = \sum_{j-s \in P(2\mu+2)} \lambda_{k,s}(j)f(2^{-k}j)$$

with zero coefficients $\lambda_{k,s}(j)$ for $j \notin P(k,s)$. Therefore, for any $k \in \mathbb{Z}_+$, we have

$$Q_k(f) = \sum_{s \in J(k)} a_{k,s}(f)M^r_{k,s} = \sum_{s \in J(k)} \sum_{j \in P(2\mu+2)} \lambda_{k,s}(j)f(2^{-k}j)M^r_{k,s}$$

$$= \sum_{j \in I(k)} f(2^{-k}j) \sum_{s-j \in P(2\mu+2)} \gamma_{k,j}(s)M^r_{k,s},$$

for certain coefficients $\gamma_{k,j}(s)$. Thus, the univariate $q_k(f)$ is of the form

$$Q_k(f) = \sum_{j \in I(k)} f(2^{-k}j)\psi_{k,j},$$

where

$$\psi_{k,j} := \sum_{s-j \in P(2\mu+2)} \gamma_{k,j}(s)M_{k,s},$$

are a linear combination of no more than the absolute number $2\mu + 2$ of B-splines $M_{k,s}$, and the size $|I(k)|$ is $2^k$. Hence, the multivariate $q_k(f)$ is of the form

$$Q_k(f) = \sum_{j \in I^d(k)} f(2^{-k}j)\psi_{k,j},$$

where

$$\psi_{k,j} := \prod_{i=1}^d \psi_{k,j_i}$$

are a linear combination of no more than the absolute number $(2\mu + 2)^d$ of B-splines $M_{k,s}$, and the size $|I^d(k)|$ is $2^{dk}$. \qed
For $0 < p \leq \infty$, denote by $\ell_p^m$ the space of all sequences $x = \{x_k\}_{k=1}^m$ of numbers, equipped with the quasi-norm
\[ \|x\|_{\ell_p^m} := \left( \sum_{k=1}^m |x_k|^p \right)^{1/p} \]
with the change to the max norm when $p = \infty$. Denote by $B_p^m$ the unit ball in $\ell_p^m$. Let $E = \{e_k\}_{k=1}^m$ be the canonical basis in $\ell_q^m$, i.e., $x = \sum_{k=1}^m x_k e_k$.

For $x = \{x_k\}_{k=1}^m \in \ell_q^m$, we let the set $\{k_j\}_{j=1}^m$ be ordered so that
\[ |x_{j_1}| \geq |x_{j_2}| \geq \cdots \geq |x_{j_s}| \geq \cdots \geq |x_{j_m}|. \]

We define the algorithm $P_n$ for the $n$-term approximation with regard to the basis $E$ in the space $\ell_q^m$ ($n \leq m$) as follows. For $x = \{x_k\}_{k=1}^m$ we let the set $\{k_j\}_{j=1}^m$ be ordered so that
\[ |x_{j_1}| \geq |x_{j_2}| \geq \cdots \geq |x_{j_s}| \geq \cdots \geq |x_{j_m}|. \]

Then, for $n < m$ we define
\[ P_n(x) := \sum_{j=1}^n (x_{k_j} - |x_{n+1}| \text{sign } x_{k_j}) e_{k_j}. \]

For a proof of the following lemma see [16].

**Lemma 3.3** The operator $P_n \in C(\ell_p^m, \ell_q^m)$ for $0 < p, q \leq \infty$. If $0 < p < q \leq \infty$, then we have for any positive integer $n < m$
\[ \sup_{x \in B_p^m} \|x - P_n(x)\|_{\ell_q^m} \leq n^{1/q - 1/p}. \]

The following theorem gives the upper bound of (1.9) in Theorem 1.1.

**Theorem 3.1** Let $p, q, \theta, \alpha$ satisfy Condition (1.8). Then for the $d$-variable Besov space $B_{p,\theta}^\alpha$, there is the following upper bound
\[ \nu_n(B_{p,\theta}^\alpha, L_q) \ll n^{-\alpha/d}. \] (3.4)

If in addition, $\alpha < 2r$, we can find an positive integer $k^*$ and a continuous $n$-sampling recovery algorithm $S_n^B \in C(B_{p,\theta}^\alpha, L_q)$ of the form (1.4) with $A = \sum_n (M(k^*))$, such that
\[ \sup_{f \in S_n^B} \|f - S_n^B(f)\|_q \ll n^{-\alpha/d}. \] (3.5)

**Proof.** We will prove (3.5) and therefore, (3.4). We first consider the case $p \geq q$. For any integer $k^*$, by Lemmas 3.2 and 3.1 we have
\[ \sup_{f \in S_n^B} \|f - Q_{k^*}(f)\|_q \ll 2^{-\alpha k^*}. \] (3.6)
The number of sampled values of $f$ in $Q_{k^*}(f)$ is $n^* := (2k^* + 1)^d$. For a given integer $n$ (not smaller than $2^d$), define $k^*$ by the condition

$$Cn \leq n^* = (2k^* + 1)^d \leq n,$$

(3.7)

with $C$ an absolute constant. By Lemma 3.1 $Q_{k^*} \in C(B^\alpha_{p,\theta}, L_q)$. By the choice of $k^*$, $Q_{k^*}(f) = S^B_n(f)$ is a linear $n$-sampling algorithm $S^B_n(f)$ is of the form (1.1) with $A = \Sigma_n(M(k^*))$ and $M(k^*) \in \mathcal{G}$ as a finite family. Therefore, by (3.6) and (3.7) we receive (5.5) for the case $p \geq q$.

We next treat the case $p < q$. For arbitrary positive integer $m$, a function $f \in SB^\alpha_{p,\theta}$ can be represented by a series

$$f = \sum_{k=0}^{m} \sum_{s \in J^d(k)} a_{k,s}(f)M_{k,s} + \sum_{k=m+1}^{\infty} \sum_{s \in J^d(k)} c_{k,s}(f)M_{k,s}$$

(3.8)

converging in the norm of $B^\alpha_{p,\theta}$ or, equivalently,

$$f = Q_m(f) + \sum_{k=m+1}^{\infty} q_k(f)$$

(3.9)

with the component functions

$$q_k(f) = \sum_{s \in J^d(k)} c_{k,s}(f)M_{k,s}$$

(3.10)

from the subspace $V(k)$. Moreover, $q_k(f)$ satisfy the condition

$$\|q_k(f)\|_p \asymp 2^{-dk/p}\|\{c_{k,s}(f)\}\|_{p,k} \ll 2^{-\alpha k}, \quad k = m + 1, m + 2, ...$$

(3.11)

The representation (3.8)–(3.11) follows from Theorem 2.1 for the case (i) in Condition (1.8), and from Lemma 3.1 for the case (ii) in Condition (1.8).

Put $m(k) := |J^d(k)| = (2^k + 2r - 1)^d$. Let $\bar{k}, k^*$ be non-negative integers with $\bar{k} < k^*$, and \{n(k)\}_{k=\bar{k}+1}^{k^*} a sequence of non-negative integers with $n(k) \leq m(k)$. We will construct a recovering function of the form

$$G(f) := \sum_{s \in J^d(k)} a_{k,s}(f)M_{k,s} + \sum_{k=\bar{k}+1}^{k^*} \sum_{s \in J^d(k)} c_{k,s_j}(f)M_{k,s_j},$$

(3.12)

with $s_{k,j} \in J^d(k)$, or equivalently,

$$G(f) = Q_{\bar{k}}(f) + \sum_{k=\bar{k}+1}^{k^*} G_k(f).$$

(3.13)

The algorithms $G_k$ are constructed as follows. For a $f \in SB^\alpha_{p,\theta}$, we take the sequence of coefficients $\{c_{k,s}(f)\}_{s \in J^d(k)}$ and reorder the indexes $s \in J^d(k)$ as $s_j^{m(k)}$ so that

$$|c_{k,s_1}(f)| \geq |c_{k,s_2}(f)| \geq \cdots \geq |c_{k,s_n}(f)| \geq \cdots \geq |c_{k,m(k)}(f)|,$$

15
and then define
\[ G_k(f) := \sum_{j=1}^{n(k)} \left( c_{k,j}(f) - |c_{k,j}(f)| \operatorname{sign} c_{k,j}(f) \right) M_{k,s_j}. \]

We prove that \( G \in \mathcal{C}(B^\alpha_{p,\theta}, L_q) \). For \( 0 < \tau \leq \infty \), denote by \( \mathbf{V}(k)_\tau \) the quasi-normed space of all functions \( f \in \mathbf{V}(k) \), equipped with the quasi-norm \( L_\tau \). Then by Lemma 3.1, \( q_k \in \mathcal{C}(B^\alpha_{p,\theta}, \mathbf{V}(k)_p) \).

Consider the sequence \( \{c_{k,s}(f)\}_{s \in J^d(k)} \) as an element in \( \ell_p^{m(k)} \) and let the operator \( D_k : \mathbf{V}(k)_p \to \ell_p^{m(k)} \) be defined by \( g \mapsto \{a_s\}_{s \in J^d(k)} \) if \( g \in \mathbf{V}(k)_q \) and \( g = \sum_{s \in J^d(k)} a_s M_{k,s} \). Obviously, by (2.8)–(2.9), \( D_k \in \mathcal{C}(\Sigma(k)_p, \ell_p^{m(k)}) \). For \( x = \{x_{k,s}\}_{s \in J^d(k)} \in \ell_p^{m(k)} \), we let the set \( \{k_j\}_{j=1}^{m(k)} \) be ordered so that
\[ |x_{j_1}| \geq |x_{j_2}| \geq \cdots \geq |x_{j_{m(k)}}| \]
and define
\[ P_{n(k)}(x) := \sum_{j=1}^{n(k)} (x_{k,j} - |x_{n(k)+1}| \operatorname{sign} x_{k,j}) e_{k,j}. \]

Temporarily denote by \( H \) the quasi-metric space of all \( x = \{x_{k,s}\}_{s \in J^d(k)} \in \ell_q^{m(k)} \) for which \( x_k = 0, k \notin Q \), for some subset \( Q \subset J^d(k) \) with \( |Q| = n(k) \). The quasi-metric of \( H \) is generated by the quasi-norm of \( \ell_q^{m(k)} \). By Lemma (3.3) we have \( P_{n(k)} \in \mathcal{C}(\ell_p^{m(k)}, H) \). Consider the mapping \( R_{M(k)} \) from \( H \) into \( \Sigma_{n(k)}(M(k)) \) defined by
\[ R_{M(k)}(x) := \sum_{s \in Q} x_{k,s} M_{k,s}, \]
if \( x = \{x_{k,s}\}_{s \in J^d(k)} \in H \) and \( x_k = 0, k \notin Q \), for some \( Q \) with \( |Q| = n(k) \). Since the family \( M(k) \) is bounded in \( L_q \), it is easy to verify that \( R_{M(k)} \in \mathcal{C}(H, L_q) \). We have
\[ G_k = R_{M(k)} \circ P_{n(k)} \circ D_k \circ q_k. \]
Hence, \( G_k \in \mathcal{C}(B^\alpha_{p,\theta}, L_q) \) as the supercomposition of continuous operators. Since by Lemma 3.1, \( Q_{\tilde{k}}(f) \in \mathcal{C}(B^\alpha_{p,\theta}, L_q) \), from 3.13 it follows \( G \in \mathcal{C}(B^\alpha_{p,\theta}, L_q) \).

Notice that in the operator \( G \), the quasi-interpolant \( Q_{\tilde{k}}(f) \) is the main non-adaptive linear part. Its adaptive non-linear part is a sum of continuous algorithms \( G_k \) for a continuous adaptive approximation of each component function \( q_k(f) \) in the \( k \)th scale subspaces \( \mathbf{V}(k) \), \( \tilde{k} < k \leq k^* \).

Let \( m \) be the number of the terms in the sum (3.12). Then, \( G(f) \in \Sigma_m(M(k^*)) \) and
\[ m = (2^{\tilde{k}} + r - 1)^d + \sum_{k=\tilde{k}+1}^{k^*} n(k). \]

Moreover, the number of sampled values defining \( G(f) \) does not exceed
\[ m' := (2^{\tilde{k}} + 1)^d + (2\mu + 2r)^d \sum_{k=\tilde{k}+1}^{k^*} n(k). \]
Let us select \( k, k^* \) and a sequence \( \{n(k)\}_{k=k+1}^{k^*} \). We define an integer \( \bar{k} \) from the condition

\[
C_1 2^{\bar{k}k} \leq n < C_2 2^{\bar{k}k},
\]

where \( C_1, C_2 \) are absolute constants which will be chosen below.

Notice that under the hypotheses of Theorem 1.1 we have \( 0 < \delta < \alpha \). Further, we fix a number \( \varepsilon \) satisfying the inequalities

\[
0 < \varepsilon < (\alpha - \delta)/\delta,
\]

where \( \delta := d(1/p - 1/q) \). An appropriate selection of \( k^* \) and \( \{n(k)\}_{k=k+1}^{k^*} \) is

\[
k^* := \lceil \varepsilon^{-1} \log(\lambda n) \rceil + \bar{k} + 1.
\]

and

\[
n(k) = \lceil \lambda n 2^{-\varepsilon(k-\bar{k})} \rceil, \quad k = \bar{k}, \bar{k} + 1, \ldots, k^*,
\]

with a positive constant \( \lambda \). Here \([a]\) denotes the integer part of the number \( a \). It is easy to find constants \( C_1, C_2 \) in (3.14) and \( \lambda \) in (3.17) so that \( n(k) \leq m(k), k = \bar{k} + 1, \ldots, k^* \), \( m \leq n \) and \( m' \leq n \). Therefore, \( G \) is an \( n \)-sampling algorithm \( S_n^B \) of the form (1.4) with \( B = \Sigma_{\alpha}(M(k^*)) \) and \( M(k^*) \in G \) as a finite family. Let us give an upper bound for \( \| f - S_n^B(f) \|_q \). For a fixed number \( 0 < \tau \leq \min(p,1) \), we have by (2.3),

\[
\| f - S_n^B(f) \|_q \leq \sum_{k=\bar{k}+1}^{k^*} \| q_k(f) - G_k(q_k(f)) \|_q^\tau + \sum_{k > k^*} \| q_k(f) \|_q^\tau.
\]

By (2.8)-(2.9) and (3.11) we have for all \( f \in SB_p^{\alpha,\theta} \)

\[
\| q_k(f) \|_q \ll 2^{-(\alpha - \delta)k}, \quad k = k^* + 1, +2, \ldots
\]

Further, we will estimate \( \| q_k(f) - G_k(q_k(f)) \|_q \) for all \( f \in SB_p^{\alpha,\theta} \) and \( k = \bar{k} + 1, \ldots, k^* \). From Lemma 3.3 we get

\[
\left( \sum_{j=n(k)+1}^{m(k)} |c_{k,s_j}(f)|^q \right)^{1/q} \leq \{n(k)\}^{-\delta} \| \{c_{k,s}(f)\} \|_{p,k}
\]

By (2.8)-(2.9), (3.19) and (3.20) we obtain for all \( f \in SB_p^{\alpha,\theta} \) and \( k = \bar{k} + 1, \ldots, k^* \)

\[
\| q_k(f) - G_k(q_k(f)) \|_q \ll 2^{-k/q} \left( \sum_{j=n(k)+1}^{m(k)} |c_{k,s_j}(f)|^q \right)^{1/q} \ll 2^{-k/q} \{n(k)\}^{-\delta} \| \{c_{k,s}(f)\} \|_{p,k} \ll 2^{-\alpha k} 2^{\delta k} \{n(k)\}^{-\delta}.
\]
From (3.18) by using (3.21), (3.19), (3.14)–(3.17) and the inequality \( \alpha > \delta \), we derive that for all functions \( f \in SB_{\alpha}^{p,\theta} \)

\[
\| f - S_B^n(f) \|_q \ll \sum_{k=k+1}^{k^*} 2^{-\tau \alpha k^* 2^{\tau \delta k}} \{ n(k) \}^{-\tau \delta} + \sum_{k=k^*+1}^{\infty} 2^{-\tau \alpha k^* 2^{\tau \delta k}}
\]

\[
\ll n^{-\tau \delta} 2^{-\tau (\alpha - \delta) \bar{k}} \sum_{k=k+1}^{k^*} 2^{-\tau (\alpha - \delta) (k-k^*)} + 2^{-\tau (\alpha - \delta) k^*} \sum_{k=k^*+1}^{\infty} 2^{-\tau (\alpha - \delta) (k-k^*)}
\]

\[
\ll n^{-\tau \delta} 2^{-\tau (\alpha - \delta) \bar{k}} + 2^{-\tau (\alpha - \delta) k^*} \ll n^{-\tau \alpha/d}.
\]

Summing up, we have proven that the constructed \( n \)-sampling algorithm \( G = S_B^n(f) \in C(B_{\alpha}^p, \ell_q) \)

and is of the form (1.4) with \( A = \sum_{M(k^*)} \), and \( M(k^*) \in G \) as a finite family for which the inequality \( (5.5) \) holds true for the case \( p < q \).

\section{Lower bounds of \( \nu_n(B_{p,\theta}^\alpha, \ell_q) \)}

To prove the lower bound Theorem \( \square \) we compare \( \nu_n(B_{p,\theta}^\alpha, \ell_q) \) with a related non-linear \( n \)-width which is defined on the basis of continuous algorithms in \( n \)-term approximation.

Let \( X, Y \) be quasi-normed spaces and \( X \) is a linear subspace of \( Y \). Let \( W \) be a subset in \( X \) and \( \Phi = \{ \varphi_k \}_{k \in K} \) a family of elements in \( Y \). Denote by \( G(Y) \) the set of all bounded families \( \Phi \subset Y \)

whose intersection \( \Phi \cap L \) with any finite dimensional subspace \( L \) in \( Y \) is a finite set. We define the non-linear \( n \)-width \( \tau_X^n(W, Y) \) by

\[
\tau_X^n(W, Y) := \inf_{\Phi \in \mathcal{G}(Y)} \inf_{S \in \mathcal{C}(X, Y)} \sup_{f \in W} \| f - S(f) \|_Y.
\]

Since all quasi-norms in a finite dimensional linear space are equivalent, we will drop ”\( X \)” in the notation \( \tau_X^n(W, Y) \) for the case where \( Y \) is finite dimensional.

Denote by \( SX \) the unit ball in the quasi-normed space \( X \). By definition we have

\[
\nu_n(B_{p,\theta}^\alpha, \ell_q) \geq \tau_B^n(SB_{p,\theta}^\alpha, \ell_q),
\]

where we use the abbreviation: \( B := B_{p,\theta}^\alpha \).

\textbf{Lemma 4.1} Let the linear space \( L \) be equipped with two equivalent quasi-norms \( \| \cdot \|_X \) and \( \| \cdot \|_Y \), \( W \) a subset of \( L \). If \( \tau_X^n(W, Y) > 0 \), we have

\[
\tau_X^{n+m}(W, X) \leq \tau_X^n(W, Y) \tau_m(SY, X).
\]

\textbf{Proof.} This lemma can be proven is a way similar to the proof of Lemma 4 in \( \square \).

\textbf{Lemma 4.2} Let \( 0 < q \leq \infty \). Then we have for any positive integer \( n < m \)

\[
\tau_n(B_{\infty}^m, \ell_q^m) \geq \frac{1}{2} (m - n - 1)^{1/q}.
\]

18
Proof. We need the following inequality. If $W$ is a compact subset in the finite dimensional normed space $Y$, then we have

$$2\tau_n(W,Y) \geq b_n(W,Y), \tag{4.2}$$

where the Bernstein $n$-width $b_n(W,Y)$ is defined by

$$b_n(W,Y) := \sup_{L_{n+1}} \sup \{ t > 0 : tSY \cap L_{n+1} \subset W \}$$

with the outer supremum taken over all $(n+1)$-dimensional linear manifolds $L_{n+1}$ in $Y$.

By definition we have $b_{m-1}(B^m_{\infty},\ell^m_{\infty}) = 1$. Hence, by (4.2), Lemmas 3.3 and 4.1 we derive that

$$1 = b_{m-1}(B^m_{\infty},\ell^m_{\infty}) \leq 2\tau_{m-1}(B^m_{\infty},\ell^m_{\infty}) \leq 2\tau_n(B^m_{\infty},\ell^m_{q}) \tau_{m-n-1}(B^m_{q},\ell^m_{\infty}) \leq 2(m-n-1)^{-1/q}\tau_n(B^m_{\infty},\ell^m_{q}).$$

This proves the lemma. \qed

Theorem 4.1 Let $0 < p,q,\theta \leq \infty$ and $\alpha > 0$. Then we have

$$\nu_n(B^\alpha_{p,\theta},L_q) \gg n^{-\alpha/d}. \tag{4.3}$$

Proof. By (4.1) the theorem follows from the inequality

$$\tau_n^B(SB^\alpha_{p,\theta},L_q) \gg n^{-\alpha/d}. \tag{4.3}$$

To prove (4.3) we will need an additional inequality. Let $Z$ be a subspace of the quasi-normed space $Y$ and $W$ a subset of the quasi-normed space $X$. If $P : Y \rightarrow Z$ is a linear projection such that $\|P(f)\|_Y \leq \lambda\|f\|_Y (\lambda > 0)$ for every $f \in Y$, then it is easy to verify that

$$\tau_n^X(W,Y) \geq \lambda^{-1}\tau_n^X(W,Z). \tag{4.4}$$

Because of the inclusion $U := SB^\alpha_{\infty,\theta} \subset SB^\alpha_{p,\theta}$, we have

$$\tau_n^B(SB^\alpha_{p,\theta},L_q) \geq \tau_n^B(U,L_q). \tag{4.5}$$

Fix an integer $r$ with the condition $\alpha < \min(2r,2r-1+1/p,2r)$. Let $U(k) := \{ f \in V(k) : \|f\|_\infty \leq 1 \}$. For each $f \in V(k)$, there holds the Bernstein inequality $\|f\|_{B^\alpha_{\infty,\theta}} \leq C2^\alpha_k\|f\|_\infty$, where $C > 0$ does not depend on $f$ and $k$. Hence, $C^{-1}2^{-\alpha k}U(k)$ is a subset in $U$. This implies the inequality

$$\tau_n^B(U,L_q) \gg 2^{-\alpha k}\tau_n^B(U(k),L_q). \tag{4.6}$$
Denote by \( V(k)_q \) the quasi-normed space of all functions \( f \in V(k) \), equipped with the quasi-norm \( L_q \). Let \( T_k \) be the bounded linear projector from \( L_q \) onto \( V(k)_q \) constructed in [11] such that \( \|T_k(f)\|_q \leq \lambda \|f\|_q \) for every \( f \in L_q \), where \( \lambda \) is an absolute constant. Therefore, by (4.4)

\[
\tau_n^B(U(k), L_q) \gg \tau_n^B(U(k), V(k)_q) = \tau_n(U(k), V(k)_q).
\] (4.7)

Observe that \( m := |J^d(k)| = \dim V(k)_q = (2^k + 2r - 1)^d \gg 2^{dk} \). For a non-negative integer \( n \), define \( m = m(n) \) from the condition

\[
n \gg 2^{dk} \gg m > 2n.
\] (4.8)

Consider the quasi-normed space \( \ell^m_q \) of all sequences \( \{a_s\}_{s \in J^d(k)} \). Let the natural continuous linear one-to-one mapping \( \Pi \) from \( V(k)_q \) onto \( \ell^m_q \) be defined by

\[
\Pi(f) := \{a_s\}_{s \in J^d(k)}
\]

if \( f \in V(k)_q \) and \( f = \sum_{s \in J^d(k)} a_s M_{k,s} \). We have by (2.8)–(2.9) \( \|f\|_\infty \gg \|\Pi(f)\|_{\ell^m_q} \) and \( \|f\|_q \gg 2^{-dk/q} \|\Pi(f)\|_{\ell^m_q} \). Hence, we obtain by Lemma 4.2

\[
\tau_n(U(k), V(k)_q) \gg 2^{-dk/q} \tau_n(B_m^\infty, \ell^m_q) \gg 2^{-dk/q} (m - n - 1)^{1/q} \gg 1.
\]

Combining the last estimates and (4.5)–(4.8) completes the proof of (4.3).

## 5 Adaptive non-continuous sampling recovery

In this section, we prove the asymptotic order of \( s_n(B^\alpha_{p,\theta}, M, L_q) \), \( r_n(SB^\alpha_{p,\theta})_q \) and \( e_n(SB^\alpha_{p,\theta})_q \) in Theorem 1.1.

Let \( W \) and \( B \) be subsets in \( L_q \). For approximation of elements from \( W \) by \( B \), the quantity

\[
E(W, B)_q := \sup_{f \in W} \inf_{\varphi \in B} \|f - \varphi\|_q
\]
gives the worst case error of approximation.

Let \( \Phi = \{\varphi_k\}_{k \in K} \) be a family of elements in \( L_q \). The quantity of \( n \)-term approximation \( \sigma_n(W, \Phi)_q \) with regard to \( \Phi \), is defined by

\[
\sigma_n(W, \Phi)_q := E(W, \Sigma_n(\Phi))_q.
\]

Given a family \( B \) of subsets in \( L_q \), we can consider the best approximation by \( B \) from \( B \) in terms of the quantity

\[
d(W, B)_q := \inf_{B \in B} E(W, B)_q.
\] (5.1)
Notice the following useful identities

\[ \sigma_n(W, \Phi)_q = \inf_{S: W \to \Sigma_n(\Phi)} \sup_{f \in W} \| f - S(f) \|_q. \]

and

\[ d(W, B)_q = \inf_{B \in B} \inf_{S^B: W \to \Sigma_n(\Phi)} \sup_{f \in W} \| f - S^B(f) \|_q. \] (5.2)

The quantity \( d(W, B)_q \) is called the entropy \( n \)-width (entropy number) \( \varepsilon_n(W)_q \) if \( B \) in (5.1) is the family of all subsets \( B \) of \( L_q \) such that \( |B| \leq 2^n \). The non-linear \( n \)-width \( \rho_n(W)_q \) is defined only when \( L_q \) is a space of real-valued functions on a set \( \Omega \), if \( B \) in (5.1) is the family of all subsets in \( L_q \) of pseudo-dimension at most \( n \).

From (5.2) we have

\[ \varepsilon_n(W)_q = \inf_{|B| \leq 2^n} \inf_{S^B: W \to \Sigma_n(\Phi)} \sup_{f \in W} \| f - S^B(f) \|_q, \]

and

\[ \rho_n(W)_q = \dim_{p, B \leq n} \inf_{S^B: W \to \Sigma_n(\Phi)} \sup_{f \in W} \| f - S^B(f) \|_q. \]

Therefore, we can take the last identities as alternative definitions of \( \varepsilon_n(W)_q \) and \( \rho_n(W)_q \).

**Theorem 5.1** Let \( p, q, \theta, \alpha \) satisfy Condition (1.8) and \( \alpha < 2r \). Then for the \( d \)-variable Besov class \( SB^\alpha_{p, \theta} \), we can explicitly construct an \( n \)-sampling algorithm \( S^B_n \) with \( B = \Sigma_n(M) \) so that

\[ \sup_{f \in SB^\alpha_{p, \theta}} \| f - S^B_n(f) \|_q \asymp s_n(B^\alpha_{p, \theta}, M, L_q) \asymp n^{-\alpha/d}. \] (5.3)

**Proof.** In [18, Corollary 2.3, Theorem 3.2] an \( n \)-sampling algorithm \( S^B_n \) with \( B = \Sigma_n(M) \) was explicitly constructed such that

\[ \sup_{f \in SB^\alpha_{p, \theta}} \| f - S^B_n(f) \|_q \ll n^{-\alpha/d}. \]

This proves the upper bound of (5.3).

The lower bound follows from the inequality \( s_n(B^\alpha_{p, \theta}, M, L_q) \geq \sigma_n(SB^\alpha_{p, \theta}, M)_q \) and the inequality

\[ \sigma_n(SB^\alpha_{p, \theta}, M)_q \gg n^{-\alpha/d}. \]

which was proven in [18, Theorem 5.1].

**Theorem 5.2** Let \( p, q, \theta, \alpha \) satisfy Condition (1.8). Then for the \( d \)-variable Besov class \( SB^\alpha_{p, \theta} \), there is the following asymptotic order

\[ r_n(SB^\alpha_{p, \theta})_q \asymp n^{-\alpha/d}. \] (5.4)
If in addition, $\alpha < 2r$, we can explicitly construct a subset $B$ in $\Sigma_n(M)$ having $\dim_p(B) \leq n$, and a sampling recovery method $S_n^B$ of the form (1.4), such that
\[
\sup_{f \in SB_{p,\theta}^\alpha} \|f - S_n^B(f)\|_q \ll n^{-\alpha/d}.
\] (5.5)

**Proof.** The inequality (5.5) and therefore, the upper bound of (5.4) was proven in [18, Theorem 3.1].

The lower bound follows from the inequality $r_n(SB_{p,\theta}^\alpha)q \geq \rho_n(SB_{p,\theta}^\alpha)q$ and the inequality
\[
\rho_n(SB_{p,\theta}^\alpha)q \gg n^{-\alpha/d}.
\]
which was proven in [18, Theorem 5.3] □

**Theorem 5.3** Let $p, q, \theta, \alpha$ satisfy Condition (1.8). Then for the $d$-variable Besov class $SB_{p,\theta}^\alpha$, there is the following asymptotic order
\[
e_n(SB_{p,\theta}^\alpha)q \asymp n^{-\alpha/d}.
\] (5.6)

If in addition, $\alpha < 2r$, we can explicitly construct a subset $B$ in $\Sigma_n(M)$ having $|B| \leq 2^n$, and a sampling recovery method $S_n^B$ of the form (1.4), such that
\[
\sup_{f \in SB_{p,\theta}^\alpha} \|f - S_n^B(f)\|_q \ll n^{-\alpha/d}.
\] (5.7)

**Proof.** The inequality (5.7) and therefore, the upper bound of (5.6) was proven in [18, Theorem 4.1].

The lower bound follows from the inequality $e_n(SB_{p,\theta}^\alpha)q \geq \varepsilon_n(SB_{p,\theta}^\alpha)q$ and the inequality
\[
\varepsilon_n(SB_{p,\theta}^\alpha)q \gg n^{-\alpha/d}.
\]
which was proven in [18, Theorem 5.5] □

**Acknowledgments.** This work was supported by Vietnam National Foundation for Science and Technology Development (NAFOSTED).

**References**

[1] P.S. Alexandrov, Über die Urysohnschen Konstanten, Fund. Math. 20(1933), 140–150.

[2] O.V. Besov, V.P. Il’in, S.M. Nikol’ skii, Integral representations of functions and embedding theorems, Winston & Sons, Washington D.C.; Halsted Press [John Wiley & Sons], New York-Toronto, Ont.-London, (1978) (Vol. I), (1979) (Vol. II).
3. C.K. Chui, An Introduction to Wavelets, Academic Press, New York, 1992.
4. C.K. Chui, H. Diamond A natural formulation of quasi-interpolation by multivariate splines, Proc. Amer. Math. Soc. 99(1987), 643–646.
5. C. de Boor, G.J. Fix, Spline approximation by quasiinterpolants, J. Approx. Theory 8(1973), 19–45.
6. C. de Bore, K. Höllig, S. Riemenschneider, Box Spline, Springer-Verlag, Berlin, 1993.
7. R.A. DeVore, Nonlinear approximation, Acta Numerica 7(1998), 51–150.
8. R. DeVore, R. Howard and C. Micchelli, Optimal non-linear approximation, Manuscripta Math. 63(1993), 194–214.
9. R. DeVore, G. Kyriazis, D. Leviatan and V. Tikhomirov, Wavelet compression and non-linear n-widths, Adv. Comp. Math. 1993, 469–478.
10. R.A. DeVore, G.G. Lorentz, Constructive approximation, Springer-Verlag, New York, 1993.
11. R.A. DeVore, V.A. Popov, Interpolation of Besov spaces, Transaction Amer. Math. Soc. 305(1988), 397–413.
12. R.A. DeVore, X.M. Yu, Nonlinear n-widths in Besov space, In: Approximation Theory VI: v.1, Academic Press, 1989, 203–206.
13. Dinh Dung, On interpolation recovery for periodic functions, In: Functional Analysis and Related Topics (Ed. S. Koshi), World Scientific, Singapore 1991, pp. 224–233.
14. Dinh Dung, On nonlinear n-widths and n-term approximation, Vietnam J. Math. 26(1998), 165–176.
15. Dinh Dung, Continuous algorithms in n-term approximation and non-linear n-widths, J. Approx. Theory 102(2000), 217–242.
16. Dinh Dung, Asymptotic orders of optimal non-linear approximations, East J. Approx. 7(2001), 55–76.
17. Dinh Dung, Non-linear sampling recovery based on quasi-interpolant wavelet representations, Adv. in Comp. Math. 30(2009), 375–401.
18. Dinh Dung, Optimal adaptive sampling recovery, Adv. in Comput. Math. 34(2011), 1–41.
19. Dinh Dung, B-spline quasi-interpolant representations and sampling recovery of functions with mixed smoothness accepted for publication in J. Complexity.
20. Dinh Dung and Vu Quoc Thanh, On nonlinear n-widths, Proc. of Amer. Math. Soc. 124(1996), 3357–3365.
21. D. Haussler, Decision theoretic generalization of the PAC model for neural net and other learning applications, Inform. Comput. 100(1)(1982), 78–150.
[22] A. N. Kolmogorov, V.M. Tikhomirov, $\varepsilon$-entropy and $\varepsilon$-capacity of sets in function space, Uspekhi Mat. Nauk 14(1959), 3–86; English transl. in Amer. Math. Soc. Transl. (2) 17(1961).

[23] S.N. Kydryatsev, The best accuracy of reconstruction of finitely smooth functions from their values at a given number of points, Izv. Math. 62(1998), 19–53.

[24] P. Mathé, s-Number in information-based complexity, J. Complexity 6(1990), 41–66.

[25] S. Nikol’skii, Approximation of functions of several variables and embedding theorems, Springer-Verlag, Berlin, 1975.

[26] E. Novak, Deterministic and Stochastic Error Bounds in Numerical Analysis, Lecture Notes in Mathematics 1349, Springer, Berlin, 1988.

[27] E. Novak, H. Triebel, Function spaces in Lipschitz domains and optimal rates of convergence for sampling, Constr. Approx. 23(2006), 325-350.

[28] D. Pollard, Empirical processes: theory and applications, NSF-CBMS Regional Conference Series in Probability and Statistics, vol. 2, Inst. Math., Stat. and Ann. Stat. Assoc., Providence, RI, 1989.

[29] J. Ratsaby, V. Maiorov, The degree of approximation of sets in Euclidean space using sets with bounded Vapnik–Chervonenkis dimension, Discrete Applied Math. 86(1998), 81–93.

[30] J. Ratsaby, V. Maiorov, On the degree of approximation by manifolds of finite pseudo-dimension, Constr. Approx. 15(1999), 291–300.

[31] V. Temlyakov, Approximation of periodic functions, Nova Science Publishers, Inc., New York, 1993.

[32] V. Temlyakov, Nonlinear methods of approximation, Fund. Comput. Math. 3(2003), 33–107.

[33] V. Tikhomirov, Some topics in approximation theory, Moscow State Univ., Moscow, 1976.

[34] V. N. Vapnik, A. Ya. Chervonenkis, Necessary and sufficient conditions for the uniform convergence of means to their expectations, Theory Probab. Appl. 26(1981), 264–280.