Journal of Approximation Theory **102**, 217–242 (2000)
Article ID jath.1999.3399, available online at http://www.idealibrary.com on IDEAL®

Continuous Algorithms in n-Term Approximation and Non-Linear Widths

Dinh Dung

Institute of Information Technology, Nghia Do, Cau Giay, Hanoi, Vietnam
E-mail:ddung@ioit.ncst.ac.vn

Communicated by Charles A. Micchelli

Received August 27, 1998; accepted June 16, 1999

In the present paper we investigate optimal continuous algorithms in n-term approximation based on various non-linear n-widths, and n-term approximation by the dictionary **V** formed from the integer translates of the mixed dyadic scales of the tensor product multivariate de la Vallée Poussin kernel, for the unit ball of Sobolev and Besov spaces of functions with common mixed smoothness. The asymptotic orders of these quantities are given. For each space the asymptotic orders of non-linear n-widths and n-term approximation coincide. Moreover, these asymptotic orders are achieved by a continuous algorithm of n-term approximation by **V**, which is explicitly constructed. © 2000 Academic Press

1. INTRODUCTION

Let X be a quasi-normed linear space and $\Phi = \{\varphi_k\}_{k=1}^{\infty}$ a family of elements in X (a quasi-norm $\|\cdot\|$ is defined as a norm except that the triangle inequality is substituted by: $\|f+g\| \leqslant C(\|f\| + \|g\|)$ with C an absolute constant). Consider n-term approximation of elements $f \in X$ by linear combinations of the form

$$\varphi = \sum_{k \in Q} a_k \varphi_k,$$

where Q is a set of natural numbers with $|Q| = n$. Here and later $|Z|$ denotes the cardinality of the set Z. It is convenient to assume that some elements of Φ can coincide, in particular, Φ can be a finite set, i.e., the number of distinct elements of Φ is finite. Denote by $\mathbf{M}_n(\Phi)$ the set of all these linear combinations. Notice that the set $\mathbf{M}_n(\Phi)$ is not linear. If the family Φ is bounded, i.e., $\|\varphi_k\| \leqslant C$, $k = 1, 2, ...,$ and the span of Φ is dense in X, then Φ is called a *dictionary*.

0021-9045/00 $35.00

If $W \subset X$, we can put

$$\sigma_n(W, \Phi, X) := \sup_{f \in W} \sigma_n(f, \Phi, X) = \sup_{f \in W} \inf_{\varphi \in \mathbf{M}_n(\Phi)} \|f - \varphi\|. \quad (1)$$

The quantities $\sigma_n(f, \Phi, X)$ and $\sigma_n(W, \Phi, X)$ are called the n-term approximation by the family Φ of f and W, respectively.

There has recently been great interest in both the theoretical and practical aspects of n-term approximation. It is directly related to non-linear approximation by trigonometric polynomials, by splines with free knots and by wavelet decompositions. There are special applications of n-term approximation to image and signal processing, numerical methods of PDE and statistical estimation (see [1] for details). It is easy to check that if X is separable and Φ is dense in the unit ball of X, then $\sigma_n(f, \Phi, X) = 0$ for any $f \in X$. Thus, the definition (1) is not suitable for dense dictionaries in separable spaces. Such dictionaries are not practical and for many well-known dictionaries with good properties the n-term approximation $\sigma_n(W, \Phi, X)$ has reasonable lower bounds for functions sets with common smoothness. Such a dictionary will be considered in our paper. In general, to obtain lower bounds on $\sigma_n(W, \Phi, X)$ for well-known classes W of functions families, Φ should be restricted by some "minimality properties" which at least well-known dictionaries would satisfy. This approach was considered in [10], [7].

An other way to deal with n-term approximation by $\mathbf{M}_n(\Phi)$ is to impose continuity assumptions on the algorithms of n-term approximation. This assumption which has its origin in the classical Alexandroff n-width is quite natural: the closer objects are the closer their reconstructions should be. On the one hand, any continuity assumption decreases the possibilities of approximation. On the other hand, it tends to guarantee a lower bound for n-term approximation. Moreover, it does not weaken the rate of the corresponding n-term approximation for many well-known dictionaries and functions classes. Namely, it is known that the best n-term approximation and n-term approximation by continuous algorithm have the same asymptotic order. This is shown again in our paper for the dictionary formed from the integer translates of the mixed dyadic scales of the tensor product multivariate de la Vallée Poussin kernel, and the unit ball of Sobolev and Besov spaces of functions with a common mixed smoothness. Just as the continuity assumption on the algorithms of approximation by "complexes" leads to the Alexandroff n-width (see definition (4) below), the continuity assumption on the algorithms of n-term approximation leads to various continuous non-linear n-widths. Let us introduce some of them.

A (continuous) algorithm in n-term approximation from Φ, is represented as a (continuous) mapping S from W into $\mathbf{M}_n(\Phi)$. We can restrict the approximations by elements of $\mathbf{M}_n(\Phi)$ only to those using continuous algorithms and in addition only from families Φ from $\mathscr{F}(X)$, which we define as the set of all bounded Φ whose intersection $\Phi \cap L$ with any finite dimensional subspace L in X, is a finite set. The n-term approximation with these restrictions leads to the non-linear n-width $\tau_n(W, X)$ which is given by

$$\tau_n(W, X) := \inf_{\Phi, S} \sup_{f \in W} \|f - S(f)\|, \tag{2}$$

where the infimum is taken over all continuous mappings S from W into $\mathbf{M}_n(\Phi)$ and all families $\Phi \in \mathscr{F}(X)$. Similar to $\tau_n(W, X)$ is the non-linear n-width $\tau'_n(W, X)$ which is defined by formula (2), but where the infimum taken over all continuous mappings S from W into a finite subset of $\mathbf{M}_n(\Phi)$, or equivalently, over all continuous mappings S from W into $\mathbf{M}_n(\Phi)$ and all finite families Φ in X. Note that the restrictions on the families Φ in the definitions of τ_n and τ'_n are quite natural. All well-known approximation systems satisfy them.

Another non-linear n-width, introduced in [7], is based on restrictions to continuous algorithms of n-term approximation. Before recalling this notion let us motivate it. Let l_∞ be the normed linear space of all bounded sequences of numbers $x = \{x_k\}_{k=1}^\infty$, equipped with the norm

$$\|x\|_\infty := \sup_{1 \leqslant k < \infty} |x_k|,$$

and $\mathbf{M}_n$ the subset in l_∞ of all $x \in l_\infty$ for which $x_k = 0$, $k \notin Q$, for some set of natural numbers Q with $|Q| = n$. Consider the mapping R_Φ from the metric space $\mathbf{M}_n$ into X defined by

$$R_\Phi(x) := \sum_{k \in Q} x_k \varphi_k,$$

if $x = \{x_k\}_{k=1}^\infty$ and $x_k = 0$, $k \notin Q$, for some Q with $|Q| = n$. From the definitions we can easily see that if the family Φ is bounded, then R_Φ is a continuous mapping from $\mathbf{M}_n$ into X and $\mathbf{M}_n(\Phi) = R_\Phi(\mathbf{M}_n)$. Thus, in this sense, $\mathbf{M}_n(\Phi)$ is a non-linear set in X, parametrized continuously by $\mathbf{M}_n$. On the other hand, any algorithm of n-term approximation of the elements in W by Φ can be treated as a composition $S = R_\Phi \circ G$ for some mapping G from W into $\mathbf{M}_n$. Therefore, if G is required to be continuous, then the algorithm S will also be continuous. These preliminary remarks are a basis for the notion of the non-linear n-width $\alpha_n(W, X)$ which is given by

$$\alpha_n(W, X) := \inf_{\Phi, G} \sup_{f \in W} \|f - R_\Phi(G(f))\|, \tag{3}$$

where the infimum is taken over all continuous mappings G from W into $\mathbf{M}_n$ and all bounded families Φ in X. In what follows, the families Φ in the definitions (2)–(3) are conveniently represented in the form $\Phi = \{\varphi_k\}_{k \in Q}$ where Q is an at most countable set of indices.

There are other notions of non-linear n-widths. We would like to recall some of these which are based on continuous algorithms of non-linear approximations different from n-term approximation, and related to problems discussed in the present paper.

The well-known and very old Alexandroff non-linear n-width $a_n(W, X)$ is defined by

$$a_n(W, X) := \inf_{G, K} \sup_{f \in W} \|f - G(f)\|, \tag{4}$$

where the infimum is taken over all complexes $K \subset X$ of dimensions $\leqslant n$ and all continuous mappings G from W into K. See, e.g., [18], [3], [9] for details regarding a_n. The non-linear manifold n-width $\delta_n(W, X)$ [2, 11] is defined by

$$\delta_n(W, X) := \inf_{R, G} \sup_{f \in W} \|f - R(G(f))\|, \tag{5}$$

where the infimum is taken over all continuous mappings G from W into $\mathbf{R}^n$ and R from $\mathbf{R}^n$ into X. The interested reader is referred to [3], [9] for brief surveys on the non-linear n-widths a_n and δ_n of the classical Sobolev and Besov classes.

The non-linear n-width $\beta_n(W, X)$ is defined by

$$\beta_n(W, X) := \inf_{R, G} \sup_{f \in W} \|f - R(G(f))\|, \tag{6}$$

where the infimum is taken over all continuous mappings G from W into $\mathbf{M}_n$ and R from $\mathbf{M}_n$ into X. This non-linear n-width has been introduced in [7].

The non-linear n-widths introduced in (2)–(6) are different. However, they possess some common properties and are closely related. Let W be a compact subset of a quasi-normed linear space X. Then the following inequalities hold

$$a_n(W, X) \leqslant \beta_n(W, X) \leqslant \alpha_n(W, X), \tag{7}$$

$$\delta_{2n+1}(W, X) \leqslant a_n(W, X) \leqslant \beta_n(W, X) \leqslant \delta_n(W, X), \tag{8}$$

(see [9], [7]), and

$$\tau_{n+1}(W, X) \leqslant \tau'_{n+1}(W, X) \leqslant a_n(W, X) \leqslant \tau'_n(W, X),$$

and in addition

$$\alpha_n(W, X) = \tau_n(W, X) = \tau'_n(W, X)$$

for finite dimensional X (see Lemma 4).

Our attention is primarily focused on continuous algorithms in n-term approximation and the relevant non-linear n-widths α_n, τ_n, τ'_n for classes of functions with common mixed smoothness. Interesting ideas concerning non-linear n-widths, which are not based on continuous algorithms, have been recently introduced in [13] and [16]. For other notions of non-linear n-widths, see [18], [3]. Non-continuous algorithms of n-term approximation and the n-term approximation for classes of functions with bounded mixed derivatives or differences, have been considered in [10], [17]. The reader can also consult [1] for a detailed survey of various aspects of non-linear approximation and applications, especially of n-term approximation.

A central problem in studying non-linear n-widths and the n-term approximation $\sigma_n(W, \Phi, X)$ of classes of functions is to compute their asymptotic order if these classes are defined by a common smoothness. In the present paper we investigate optimal algorithms of n-term approximation based on the non-linear n-widths α_n, τ_n, τ'_n and the n-term approximation $\sigma_n(W, \mathbf{V}, X)$ by the dictionary $\mathbf{V}$ formed from the integer translates of the mixed dyadic scales of the tensor product multivariate de la Vallée Poussin kernel V_m, for the unit ball of Sobolev and Besov spaces of functions with a common mixed smoothness. Because of the close relationship between α_n, τ_n, τ'_n, β_n, δ_n and a_n, and because they are asymptotically equivalent it is quite useful and natural to study them together.

Let us give a brief description of the main results of this paper. Throughout this paper we will assume that A is a given fixed finite subset of $\mathbf{R}^d$. For $0 < p, \theta \leqslant \infty$, let $\mathbf{B}^A_{p,\theta}$ denote the Besov space of all functions on the n-dimensional torus $\mathbf{T}^d := [0, 2\pi]^d$, for which the quasi-norm

$$\|f\|_{\mathbf{B}^A_{p,\theta}} := \|f\|_p + \sum_{\alpha \in A} |f|_{B^\alpha_{p,\theta}} \tag{9}$$

is finite, where $\|\cdot\|_p$ is the usual p-integral norm in $L_p := L_p(\mathbf{T}^d)$ and $|\cdot|_{B^\alpha_{p,\theta}}$ the Besov semi-quasi–norm determining the mixed smoothness of order α. We will use the abbreviation for the special case $A = \{0\}$: $\mathbf{B}_{p,\theta} := \mathbf{B}^{\{0\}}_{p,\theta}$. The Sobolev space $\mathbf{W}^A_p$ is defined similarly by replacing $|f|_{B^\alpha_{p,\theta}}$ in (9) by $|f|_{W^\alpha_p} := \|f^{(\alpha)}\|_p$, where $f^{(\alpha)}$ is the mixed derivative in the sense of Weil of order α. (The definitions of $|\cdot|_{B^\alpha_{p,\theta}}$ and $f^{(\alpha)}$ are given in Section 3.) Note that the classical Besov and Sobolev spaces are special cases of $\mathbf{B}^A_{p,\theta}$ and $\mathbf{W}^A_p$. The main results which are proved in the present paper are the asymptotic orders of the non-linear n-widths (2)–(6) and the n-term approximation (1) by the dictionary $\mathbf{V}$, of the unit ball of the Sobolev space $\mathbf{W}_p^A$ and the Besov space $\mathbf{B}_{p,\theta}^A$ in the space L_q.

It turns out that these asymptotic orders are closely related to the linear problem

$$(\mathbf{1}, x) \to \sup, \; x \in A_+^o, \tag{10}$$

where $A_+^o := \{x \in \mathbf{R}_+^d : (\alpha, x) \leqslant 1, \; \alpha \in A\}$, $\mathbf{1} := (1, 1, ..., 1) \in \mathbf{R}^d$. Here and later, we use the notation: $(x, y) := x_1 y_1 + \cdots + x_d y_d$ and $\mathbf{R}_+^d := \{x \in \mathbf{R}^d : x_j \geqslant 0, \; j = 1, ..., d\}$, where x_j is the jth coordinate of $x \in \mathbf{R}^d$, i.e., $x := (x_1, ..., x_d)$.

Let $1/r$ be the optimal value of (10) and v the linear dimension of the set of solutions of (10), i.e.,

$$1/r := \sup \{(\mathbf{1}, x) : x \in A_+^o\}, \qquad v := \dim \{x \in A_+^o : (\mathbf{1}, x) = 1/r\}. \tag{11}$$

We use the notation $F \asymp F'$ if $F \ll F'$ and $F' \ll F$, and $F \ll F'$ if $F \leqslant CF'$ with C an absolute constant. Denote by γ_n any one of α_n, τ_n, τ'_n, β_n, a_n and δ_n. Let

$$\mathbf{SB}_{p,\theta}^A := \{f \in \mathbf{B}_{p,\theta}^A : \|f\|_{\mathbf{B}_{p,\theta}^A} \leqslant 1\}$$

and

$$\mathbf{SW}_p^A := \{f \in \mathbf{W}_p^A : \|f\|_{\mathbf{W}_p^r} \leqslant 1\}$$

be the unit balls in $\mathbf{B}_{p,\theta}^r$ and $\mathbf{W}_p^A$, respectively.

For $1 < p, q < \infty, 2 \leqslant \theta \leqslant \infty$ and A a finite subset in $\mathbf{R}^d$, with some restrictions on A and p, q we have

$$\gamma_n(\mathbf{SB}_{p,\theta}^A, L_q) \asymp (n/\log^v n)^{-r} (\log^v n)^{1/2 - 1/\theta}, \tag{12}$$

$$\gamma_n(\mathbf{SW}_p^A, L_q) \asymp (n/\log^v n)^{-r}. \tag{13}$$

The asymptotic order (13) has been proven in [6] for a_n and δ_n. The upper bound of (12)–(13) is given by a continuous non-linear algorithm of approximation by the dictionary $\mathbf{V}$ which is constructed as follows.

For $m \in \mathbf{Z}_+^d := \{k \in \mathbf{Z}^d : k_j \geqslant 0, \; j = 1, ..., d\}$, we let the tensor product de la Vallée Poussin kernel V_m of order m be defined by

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

where

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

is the univariate de la Vallée Poussin kernel of order m. We put

$$S_m(x) := \prod_{j=1}^{d} (2/3m_j)\, V_m(x), \tag{14}$$

and

$$Q_k := \{s \in \mathbf{Z}_+^d : s_j < 3 \times 2^{k_j+1},\ j = 1, ..., d\}; \qquad h^k := 3^{-1}\pi(2^{-k_1}, ..., 2^{-k_d}). \tag{15}$$

We define the family $\mathbf{V}$ by

$$\mathbf{V} := \{\varphi_s^k\}_{s \in Q_k,\, k \in \mathbf{Z}_+^d}, \qquad \varphi_s^k := S_{2^{k+1}}(\cdot - sh^k). \tag{16}$$

Here and later we use the notation: $2^x := (2^{x_1}, ..., 2^{x_d})$ and $xy := (x_1 y_1, ..., x_d y_d)$ for $x, y \in \mathbf{R}^d$.

From well-known properties of de la Valleé Poussin kernels it follows that $\mathbf{V} \in \mathscr{F}(L_q)$, $0 < q \leqslant \infty$, and $\mathbf{V}$ is a dictionary. To establish the upper bounds of (12)–(13) we explicitly construct a positive homogeneous continuous mapping G^*: $Y \to \mathbf{M}_n$ such that

$$\sup_{f \in SY} \|f - R_{\mathbf{V}}(G^*(f))\|_q \ll E(n). \tag{17}$$

where $E(n)$ is the right-hand side of either (12) or (13), Y is either $\mathbf{B}_{p,\theta}^A$ or $\mathbf{W}_p^A$, respectively, and SY is the unit ball in Y.

Clearly from (17) we also obtain a upper bound for the n-term approximation by $\mathbf{V}$ of SY. Moreover, we prove that under the same conditions as those for (12)–(13)

$$\sigma_n(\mathbf{SB}_{p,\theta}^A, \mathbf{V}, L_q) \asymp (n/\log^\nu n)^{-r} (\log^\nu n)^{1/2 - 1/\theta}, \tag{18}$$

$$\sigma_n(\mathbf{SW}_p^A, \mathbf{V}, L_q) \asymp (n/\log^\nu n)^{-r}. \tag{19}$$

This means that $\sigma_n(SY, \mathbf{V}, L_q)$ and $\gamma_n(SY, L_q)$ have the same asymptotic order which is achieved by the continuous algorithm $S^* = R_\Phi \circ G^*$ of n-term approximation by $\mathbf{V}$. The asymptotic orders of the n-term approximation $\sigma_n(W, U^d, X)$ by the dictionary U^d formed from the integer translates of the mixed dyadic scales of the tensor product multivariate Dirichlet kernel for the classes of functions with bounded symmetric mixed derivatives or differences, have been obtained in [17]. In addition, these orders are achieved by a greedy type algorithm, a non-continuous algorithm of n-term approximation.

The main results of the present paper were announced without proofs in [8].

Our paper is organized as follows. In Section 2 we prove some estimates for non-linear n-widths and n-term approximation by the canonical basis in spaces of sequences with mixed norms, and also some equalities and inequalities between non-linear n-widths. Other auxiliary facts concerning the spaces $\mathbf{B}^A_{p,\theta}$ and $\mathbf{W}^A_p$ are given in Section 3. Section 4 is devoted to the proofs of (12)–(13) and (17). In Section 5 we give the proofs of (18)–(19) and an example of the asymptotic orders (12)–(13) and (18)–(19) for the well-known Besov and Sobolev spaces with one bounded derivative and difference, respectively.

2. NON-LINEAR WIDTHS AND n-TERM APPROXIMATION IN SPACES OF SEQUENCES

For $0 < p \leqslant \infty$ denote by l^m_p the space of all sequences $x = \{x_k\}^m_{k=1}$ of (complex) numbers, 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 the max norm when $p = \infty$. Denote by B^m_p the unit ball in l^m_p.

LEMMA 1. *Let $1 \leqslant p, q \leqslant \infty$ and $m > n \geqslant 1$. Denote by γ_n either one of $\alpha_n, \beta_n, \tau_n, \tau'_n$ and a_n. Then we have*

$$\gamma_n(B^m_p, l^m_q) \asymp A_{p,q}(m, n),$$

where

$$A_{p,q}(m,n) = \begin{cases} n^{1/q - 1/p}, & \text{for} \quad p < q \\ 1, & \text{for} \quad p = q \\ (m-n)^{1/q - 1/p}, & \text{for} \quad p > q. \end{cases}$$

Moreover, we can explicitly construct a common positive homogeneous continuous mapping $G: l^m_p \to \mathbf{M}_n$ so that the asymptotic order of the n-width $\gamma_n(B^m_p, l^m_q)$ is achieved by the algorithm $S := R_{\mathscr{E}} \circ G$, i.e.,

$$\sup_{x \in B^m_p} \|x - S(x)\|_{l^m_q} \ll A_{p,q}(m, n),$$

where $\mathscr{E}$ is the canonical basis in l^m_q.

Proof. For the proofs of this lemma for α_n, β_n and a_n see [7] and [9]. Its proofs for τ_n and τ'_n are similar. ∎

Let $0 < p, \theta \leqslant \infty$, $\mathbf{N} = \{N_k\}_{k \in Q}$ be a sequence of natural numbers, and $\mathbf{\Lambda} = \{\lambda_k\}_{k \in Q}$ a sequence of positive numbers with Q an at most countable set of indices. Denote by $\mathbf{b}^{\mathbf{N}}_{p,\theta}(\mathbf{\Lambda})$ 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

$$\|x\|_{\mathbf{b}^{\mathbf{N}}_{p,\theta}(\mathbf{\Lambda})} := \left(\sum_{k \in Q} (\|x^k\|_{X^k}/\lambda_k)^\theta \right)^{1/\theta}, \ \theta < \infty,$$

is finite (the sum is changed to supremum for $\theta = \infty$), where $X^k := l_p^{N_k}$. Let $S^{\mathbf{N}}_{p,\theta}(\mathbf{\Lambda})$ be the unit ball in $\mathbf{b}^{\mathbf{N}}_{p,\theta}(\mathbf{\Lambda})$. If $\lambda_k = 1$ for all $k \in Q$, then we use the abbreviations: $\mathbf{b}^{\mathbf{N}}_{p,\theta} := \mathbf{b}^{\mathbf{N}}_{p,\theta}(\mathbf{\Lambda})$ and $\mathbf{S}^{\mathbf{N}}_{p,\theta} := \mathbf{S}^{\mathbf{N}}_{p,\theta}(\mathbf{\Lambda})$. We also denote by SX the unit ball in the linear normed space X.

LEMMA 2. *Let $0 < p, q, \theta, \tau \leqslant \infty$ and let $\mathbf{N} = \{N_k\}_{k \in Q}$ be a sequence of natural numbers, $\mathbf{\Lambda} = \{\lambda_k\}_{k \in Q}$ and $\mathbf{\Lambda}' = \{\lambda'_k\}_{k \in Q}$ sequences of positive numbers, and $\{n_k\}_{k \in Q}$ a sequence of non-negative integers such that $\sum_{k \in Q} n_k < \infty$. Denote by γ_n any one of α_n, β_n, τ_n, τ'_n and a_n. Assume that*

$$\gamma_{n_k}(B_p^{N_k}, l_q^{N_k}) \leqslant b_k, \quad k \in Q,$$

for the sequence of non-negative numbers $\{b_k\}_{k \in Q}$, and $Q = \{k_j\}_{j=1}^m$ is ordered so that

$$\mu_{k_1} \geqslant \mu_{k_2} \geqslant \cdots \mu_{k_j} \geqslant \cdots,$$

where $\mu_k := b_k \lambda_k / \lambda'_k$ and $m = |Q|$. For any natural number $s \leqslant m$, define

$$F_{\theta,\tau}(s) = \begin{cases} \mu_{k_s}, & \text{for} \quad \theta \leqslant \tau \\ (\sum_{j=s}^m \mu_{k_j}^\rho)^{1/\rho}, & \text{for} \quad \theta > \tau, \end{cases} \tag{20}$$

with $\rho := \theta\tau/(\theta - \tau)$. Then we have

$$\gamma_n(\mathbf{S}^{\mathbf{N}}_{p,\theta}(\mathbf{\Lambda}), \mathbf{b}^{\mathbf{N}}_{q,\tau}(\mathbf{\Lambda}')) \leqslant F_{\theta,\tau}(s),$$

where

$$n := \sum_{j=1}^{s-1} N_{k_j} + \sum_{j=s}^m n_{k_j}. \tag{21}$$

In addition, we can explicitly construct a positive homogeneous continuous mapping $G: \mathbf{b}^{\mathbf{N}}_{p,\theta}(\mathbf{\Lambda}) \to \mathbf{M}_n$ *such that*

$$\sigma_n(\mathbf{S}^{\mathbf{N}}_{p,\theta}(\mathbf{\Lambda}), \mathscr{E}, \mathbf{b}^{\mathbf{N}}_{q,\tau}(\mathbf{\Lambda}')) \leqslant \sup_{x \in \mathbf{S}^{\mathbf{N}}_{p,\theta}(\mathbf{\Lambda})} \|x - S(x)\|_{\mathbf{B}^{\mathbf{N}}_{q,\tau}(\mathbf{\Lambda}')} \leqslant F_{\theta,\tau}(s), \tag{22}$$

where $S := R_{\mathscr{E}} \circ G$ *and* $\mathscr{E}$ *is the canonical basis in* $\mathbf{b}^{\mathbf{N}}_{q,\tau}(\mathbf{\Lambda}')$.

Proof. Since $\mathbf{S}^{\mathbf{N}}_{p,\theta'}(\mathbf{\Lambda}) \subset \mathbf{S}^{\mathbf{N}}_{p,\theta}(\mathbf{\Lambda})$ if $\theta' < \theta$, it suffices to prove Lemma 2 for $\theta \geqslant \tau$. We prove the case $\tau \leqslant \theta < \infty$, $m = \infty$ and $\gamma_n = \alpha_n$ of the lemma. The other cases can be treated in a similar way with a slight modification. Obviously, without loss of generality we can also assume $\mathbf{b}^{\mathbf{N}}_{q,\tau}(\mathbf{\Lambda}') = \mathbf{b}^{\mathbf{N}}_{q,\tau}$. By Lemma 1 there are positive homogeneous continuous mappings $G_k: X^k \to \mathbf{M}_{n_k}$ such that the widths $\alpha_{n_k}(SX^k, Y^k)$ are achieved by the algorithms S_k, where $S_k := R_{\mathscr{E}_k} \circ G_k$, $X^k := l^{N_k}_p$, $Y^k := l^{N_k}_q$ and $\mathscr{E}_k$ is the canonical basis in Y^k (the case $n_k \geqslant N_k$ which is not included in Lemma 1 is trivial). This implies that for any $x^k \in X^k$

$$\|x^k - S_k(x^k)\|_{Y^k} \leqslant b_k \|x^k\|_{X^k}. \tag{23}$$

Let us also use the abbreviations: $X := \mathbf{b}^{\mathbf{N}}_{p,\theta}(\mathbf{\Lambda})$, $Y := \mathbf{b}^{\mathbf{N}}_{q,\tau}$ and $B := \mathbf{S}^{\mathbf{N}}_{p,\theta}(\mathbf{\Lambda})$. We represent X^k and Y^k as subspaces of X and Y, and the sequence $\{\mathbf{M}_{n_k}\}_{k \in Q}$ as a subset of $\mathbf{M}_n$, respectively. Thus we can identify the sequence $\{\mathscr{E}_{n_k}\}_{k \in Q}$ with $\mathscr{E}$.

We first define a positive homogeneous continuous mapping $y = S(x)$ from X into Y with $x = \{x^k\}_{k \in Q}$ $y = \{y^k\}_{k \in Q}$, by putting

$$y^{k_j} := \begin{cases} x^{k_j}, & \text{for } \ j = 1, ..., s-1 \\ S_{k_j}(x^{k_j}), & \text{for } \ j = s, s+1, ... \end{cases}$$

By (23) we have for $x \in X$

$$\|x - S(x)\|^{\tau}_{Y} = \sum_{j=s}^{\infty} \|x^{k_j} - S_k(x^{k_j})\|^{\tau}_{Y^{k_j}}$$

$$\leqslant \sum_{j=s}^{\infty} b^{\tau}_{k_j} \|x^{k_j}\|^{\tau}_{X^{k_j}} \leqslant \sum_{j=s}^{\infty} (\|x^{k_j}\|_{X^{k_j}}/\lambda_{k_j})^{\tau} \mu^{\tau}_{k_j}.$$

Hence, for $x \in B$

$$\|x - S(x)\|_{Y} \leqslant F_{\theta,\tau}(s) \tag{24}$$

in the case $\theta=\tau$, and by the Hölder inequality

$$\|x-S(x)\|_Y \leqslant \left(\sum_{j=s}^{\infty} \mu_{k_j}^{\rho}\right)^{1/\rho} \left(\sum_{j=s}^{\infty} (\|x^{k_j}\|_{X^{k_j}}/\lambda_{k_j})^{\theta}\right)^{1/\theta} \leqslant F_{\theta,\tau}(s) \qquad (25)$$

in the case $\theta>\tau$.

Obviously, S is a positive homogeneous continuous mapping from X into Y. Let us represent $\mathbf{b}^{\mathbf{N}}_{\infty,\infty}$ as l_∞ or a subspace in l_∞. Then S can be represented via the composition $S=R_{\mathscr{E}}\circ G$, where the mapping $G\colon X\to \mathbf{M}_n$ is given by $G(x):=S(x)$ for $x\in X$, and n is defined in (21). Clearly, G is positive homogeneous, continuous and satisfies (22). The lemma is proved. ∎

Lemma 3. *Under the assumptions and notation of Lemma 2 let $p<\infty$, $\theta<\tau$, Q be a finite set with $|Q|=m$, and $N_k\leqslant N^*$, $n_k\leqslant n^*$, $k\in Q$. Then we have for any natural number $s\leqslant m$*

$$\gamma_n(\mathbf{S}^{\mathbf{N}}_{p,\theta}(\mathbf{\Lambda}), \mathbf{b}^{\mathbf{N}}_{\infty,\tau}(\mathbf{\Lambda}')) \leqslant \left(\sum_{j=1}^{s} \mu_{k_j}^{\rho}\right)^{1/\rho},$$

where $n:=(s-1)\,N^*+(m-s+1)\,n^*$. In addition, we can explicitly construct a positive homogeneous continuous mapping $G:\mathbf{b}^{\mathbf{N}}_{p,\theta}(\mathbf{\Lambda})\to\mathbf{M}_n$ such that

$$\sigma_n(\mathbf{S}^{\mathbf{N}}_{p,\theta}(\mathbf{\Lambda}), \mathscr{E}, \mathbf{b}^{\mathbf{N}}_{\infty,\tau}(\mathbf{\Lambda}')) \leqslant \sup_{x\in \mathbf{S}^{\mathbf{N}}_{p,\theta}(\mathbf{\Lambda})} \|x-S(x)\|_{\mathbf{b}^{\mathbf{N}}_{\infty,\tau}(\mathbf{\Lambda}')} \leqslant \left(\sum_{j=1}^{s} \mu_{k_j}^{\rho}\right)^{1/\rho},$$

where $S:=R_{\mathscr{E}}\circ G$.

Proof. We use the notation in the proof of Lemma 2. We first define a positive homogeneous mapping $y=S(x)$ from X into Y with $x=\{x^k\}_{k\in Q}$ and $y=\{y^k\}_{k\in Q}$, as follows. Let $u=\{u^k\}_{k\in Q}$ and $v=\{v^k\}_{k\in Q}$ be given by $u^k=S_k(x^k)$ and $v^k=x^k-u^k$. For $x=\{x^k\}_{k\in Q}\in X$, we define

$$D_k=D_k(x):=\mu_k^{-\rho/\tau}\,\|v^k\|_{Y^k},$$

and rearrange the set of indices Q so that $D_{r_1}\geqslant D_{r_2}\geqslant\cdots\geqslant D_{r_m}$. Then, the mapping $y=S(x)$ is defined by setting

$$y_i^{r_j}:=\begin{cases} x_i^{r_j}-\mu_{r_j}^{\rho/\tau}\,D_{r_s}\,\operatorname{sign} v_i^{r_j}, & \text{if } \ |v_i^{r_j}|\geqslant \mu_{r_j}^{\rho/\tau} D_{r_s} \\ u_i^{r_j}, & \text{if } \ |v_i^{r_j}|<\mu_{r_j}^{\rho/\tau} D_{r_s} \end{cases}, \quad j=1,2,...,s, \quad i=1,...,N_{r_j},$$

and

$$y^{r_j}:=u^{r_j}, \quad j=s+1,...,m.$$

It is easy to check that S is continuous. We estimate $\|x-S(x)\|_Y^\tau$ by using a technique of [14]. We have

$$\|x-S(x)\|_Y^\tau = D_{r_s}^\tau F^\rho + \sum_{j=s+1}^{m} \|v^{r_j}\|_{Y^{r_j}}^\tau, \tag{26}$$

where $F := (\sum_{j=1}^s \mu_{r_j}^\rho)^{1/\rho}$. Let $x = \{x^k\}_{k\in Q} \in B$ and put

$$\sum_{j=s+1}^{m} (\|x^{r_j}\|_{X^{r_j}}/\lambda_{r_j})^\theta =: \varepsilon. \tag{27}$$

Then, we have

$$\sum_{j=1}^{s} (\|x^{r_j}\|_{X^{r_j}}/\lambda_{r_j})^\theta \leqslant 1-\varepsilon. \tag{28}$$

From (23), (28) and the inequality $D_{r_j} \geqslant D_{r_s}$ for $j=1, ..., s$, we obtain

$$\begin{aligned} D_{r_s}^\theta F^\rho &\leqslant \sum_{j=1}^{s} \mu_{r_j}^\rho D_{r_j}^\theta = \sum_{j=1}^{s} \mu_{r_j}^\rho \mu_{r_j}^{-\rho\theta/\tau} \|v^{r_j}\|_{Y^{r_j}}^\theta \\ &\leqslant \sum_{j=1}^{s} \mu_{r_j}^{-\theta} b_{r_j}^\theta \|x^{r_j}\|_{X^{r_j}}^\theta = \sum_{j=1}^{s} (\|x^{r_j}\|_{X^{r_j}}/\lambda_{r_j})^\theta \leqslant 1-\varepsilon. \end{aligned}$$

Thus the following estimate has been proven

$$D_{r_s} \leqslant (1-\varepsilon)^{1/\theta} F^{-\rho/\theta}. \tag{29}$$

On the other hand, using (23), (27) and the inequality $D_{r_j} \leqslant D_{r_s}$ for $j=s+1, ..., m$, we have

$$\sum_{j=s+1}^{m} \|v^{r_j}\|_{Y^{r_j}}^\tau = \sum_{j=s+1}^{m} \|v^{r_j}\|_{Y^{r_j}}^\theta \|v^{r_j}\|_{Y^{r_j}}^{\tau-\theta}$$

$$\leqslant \sum_{j=s+1}^{m} (\|x^{r_j}\|_{X^{r_j}}^\theta \mu_{r_j}^{(\theta-\tau)\rho/\tau})^{1/\rho} D_{r_j}^{\tau-\theta} \leqslant D_{r_s}^{\theta-\tau} \sum_{j=s+1}^{m} (\|x^{r_j}\|_{X^{r_j}}/\lambda_{r_j})^\theta = \varepsilon D_{r_s}^{\theta-\tau}.$$

Hence, by (26) and (29)

$$\begin{aligned} \|x-S(x)\|_Y^\tau &\leqslant \varepsilon D_{r_s}^{\tau-\theta} + D_{r_s}^\tau F^\rho \\ &\leqslant \varepsilon\{(1-\varepsilon)^{1/\theta} F^{-\rho/\theta}\}^{\tau-\theta} + F\{(1-\varepsilon)^{1/\theta} F^{-\rho/\theta}\}^\tau \\ &= F^\tau\{\varepsilon(1-\varepsilon)^{\tau/\theta-1} + (1-\varepsilon)^{\tau/\theta}\} \leqslant F^\tau \leqslant \left(\sum_{j=1}^{s} \mu_{k_j}^\rho\right)^{\tau/\rho}. \end{aligned}$$

From the last estimate, similarly to the proof of Lemma 2, we complete the proof of Lemma 3. ∎

LEMMA 4. *Let the linear space L be equipped with two equivalent quasi-norms $\|\cdot\|_X$ and $\|\cdot\|_Y$, and W be a subset of L. Denote by γ_n either one of $\alpha_n, \tau_n, \tau'_n, \beta_n, \delta_n$ and a_n. Assume that W is compact in these norms and $\gamma_m(W, X) > 0$. Then we have*

$$\gamma_{n+m}(W, Y) \leqslant \gamma_n(SX, Y)\, \gamma_m(W, X).$$

Proof. For the proofs of Lemma 4 for $\alpha_n, \beta_n, \delta_n$ and a_n see [7] and [9]. Its proofs for τ_n and τ'_n are similar. For completeness of the present paper we give the proof for τ_n. We put $\tau_n := \tau_n(SX, Y)$ and $\tau_m := \tau_m(W, X)$. We can assume $\tau_n < \infty$, because the opposite case is trivial. Given arbitrary $\varepsilon > 0$, by definition there are families $\Phi = \{\varphi_k\}_{k \in Q}$ and $\Phi' = \{\varphi_k\}_{k \in Q'}$ from $\mathscr{F}(L)$, and continuous mappings S: $W \to \mathbf{M}_m(\Phi)$ and S': $SX \to \mathbf{M}_n(\Phi')$ such that

$$\|f - S(f)\|_X \leqslant \tau_m + \varepsilon, \qquad f \in W, \tag{30}$$

and

$$\|f - S'(f)\|_Y \leqslant \tau_n + \varepsilon, \qquad f \in SX. \tag{31}$$

We put $G(f) := f - S(f)$, $\lambda := \sup_{f \in W} \|G(f)\|_X$. By the assumptions and (30) we have

$$0 < \tau_m \leqslant \lambda \leqslant \tau_m + \varepsilon < \infty. \tag{32}$$

Note that $\lambda^{-1} G(f) \in SX$ for any $f \in W$. We define $\Phi^* := \{\varphi_k\}_{k \in Q \cup Q'}$ and the continuous mapping S^*: $W \to \mathbf{M}_{m+n}(\Phi^*)$ by $S^*(f) := S(f) + \lambda S'(\lambda^{-1} G(f))$. It is easily seen that $f - S^*(f) = \lambda(\lambda^{-1} G(f) - S'(\lambda^{-1} G(f)))$ for any $f \in W$. Hence, we obtain by (31)–(32)

$$\begin{aligned}\tau_{n+m}(W, Y) &\leqslant \sup_{f \in W} \|f - S^*(f)\|_Y \leqslant \lambda \sup_{f \in W} \|\lambda^{-1} G(f) - S'(\lambda^{-1} G(f))\|_Y \\ &\leqslant \lambda \sup_{f \in SX} \|f - S'(f)\|_Y \leqslant (\tau_m + \varepsilon)(\tau_n + \varepsilon)\end{aligned}$$

for arbitrary $\varepsilon > 0$. This proves the lemma. ∎

LEMMA 5. *Let X be a quasi-normed linear space and W a compact subset in X. Then we have*

$$\tau_{n+1}(W, X) \leqslant \tau'_{n+1}(W, X) \leqslant a_n(W, X) \leqslant \tau'_n(W, X).$$

In addition if X is finite dimensional, we have

$$\alpha_n(W, X) = \tau_n(W, X) = \tau'_n(W, X).$$

Proof. The inequalities $\tau_{n+1}(W, X) \leqslant \tau'_{n+1}(W, X)$ and $a_n(W, X) \leqslant \tau'_n(W, X)$ follow from the definitions. Let K be an n-dimensional complex in X and $\Psi = \{\psi_k\}_{k=1}^{m}$ $(m > n)$ the set of nodes of K. Then obviously, $K \subset \mathbf{M}_{n+1}(\Psi)$. Hence we obtain the inequality $\tau'_{n+1}(W, X) \leqslant a_n(W, X)$.

Let X be finite dimensional. Consider a bounded family $\Phi = \{\varphi_k\}_{k \in Q} \in \mathscr{F}(X)$ and a continuous mapping S: $W \to \mathbf{M}_n(\Phi)$. Since X is finite dimensional, without loss of generality we can assume Q is a finite subset in $\mathbf{N}$. Let the continuous mapping G: $W \to \mathbf{M}_n$ be defined by $G(f) := \{x_k\}_{k \in \mathbf{N}}$ with $x_k = a_k$ for $k \in Q$, and $x_k = 0$ for $k \notin Q$, if $S(f) = \sum_{k \in Q} a_k \varphi_k$. Then the mapping S can be represented by the composition $S = R_\Phi \circ G$. This implies the inequality $\alpha_n(W, X) \leqslant \tau_n(W, X)$. On the other hand, let $\Phi = \{\varphi_k\}_{k \in Q}$ be a bounded family in X with $\|\varphi_k\| \leqslant C$, $k \in Q$, and G: $W \to \mathbf{M}_n$ a continuous mapping. Given arbitrary $\varepsilon > 0$, by the compactness of the ball $\{f \in X : \|f\| \leqslant C\}$, there exists a finite family $\Phi' = \{\varphi_k\}_{k \in Q'}$ in X with the following property. For any $\varphi \in \Phi$ there is a $\varphi' \in \Phi'$ such that $\|\varphi - \varphi'\| \leqslant \varepsilon/n$. Hence, it is easy to check that $\|f - R_{\Phi'}(G(f))\| \leqslant \|f - R_\Phi(G(f))\| + \varepsilon$ for any $f \in W$. This implies that $\tau'_n(W, X) \leqslant \alpha_n(W, X) + \varepsilon$ for arbitrary $\varepsilon > 0$. Therefore, we have $\alpha_n(W, X) = \tau_n(W, X) = \tau'_n(W, X)$. ▮

3. SOBOLEV AND BESOV SPACES AND OTHER AUXILIARIES

Let us first make the notions of the Sobolev and Besov spaces $\mathbf{B}^A_{p,\theta}$ and $\mathbf{W}^A_p$ precise by defining $|\cdot|_{B^\alpha_{p,\theta}}$ and $|\cdot|_{W^\alpha_p}$ for $\alpha \in \mathbf{R}^d$ and $0 < p, \theta \leqslant \infty$.

As usual, $\hat{f}(k)$ denotes the k-th Fourier coefficient, in the distributional sense, of $f \in L_p$. The αth mixed derivative $f^{(\alpha)}$, in the sense of Weil, of f is defined by

$$f^{(\alpha)} := \sum_{k \in \mathbf{Z}^d_o} \hat{f}(k)(ik)^\alpha \, e^{i(k, \cdot)},$$

where $\mathbf{Z}^d_o := \{k \in \mathbf{Z}^d : k_j \neq 0,\ j = 1, ..., d\}$; $(ik)^\alpha := (ik_1)^{\alpha_1} \cdots (ik_d)^{\alpha_d}$; $(ix)^y := |x|^y \, e^{(i\pi y \operatorname{sign} x)/2}$. If $d = 1$ and α is an integer, then $f^{(\alpha)}$ coincides with the usual αth derivative of f for $\alpha > 0$, $f - (2\pi)^{-1} \hat{f}(0)$ for $\alpha = 0$, and the usual αth primitive of f with the zero mean value for $\alpha < 0$. Recall that $|\cdot|_{W^\alpha_p}$ is defined by

$$|f|_{W^\alpha_p} := \|f^{(\alpha)}\|_p.$$

While $|\cdot|_{B^{\alpha}_{p,\theta}}$ is defined as

$$|f|_{B^{\alpha}_{p,\theta}} := \left(\int_{\mathbf{T}^d} \prod_{j=1}^{d} h_j^{-1-\theta\beta_j} \|\Delta_h^l f^{(s)}\|_p^\theta \, dh \right)^{1/\theta}, \quad \theta < \infty,$$

(the integral changed to the supremum for $\theta = \infty$) for some triple $l \in \mathbf{N}^d$ and $\beta, s \in \mathbf{Z}^d$, satisfying the condition $\beta + s = \alpha$; $l_j > \beta_j > 0$, $j = 1, ..., d$, where Δ_h^l denotes the operator of lth mixed difference with step $h \in \mathbf{T}^d$. It is well-known that different triples l, p, s satisfying the last condition determine the same quasi-norm $\|\cdot\|_{B^{\alpha}_{p,\theta}}$.

We formulate the well-known Littlewood-Paley Theorem which plays a basic role in multivariate trigonometric polynomial approximation. We define the operator δ_k, $k \in \mathbf{Z}^d_+$, by

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

where $\mathbf{Z}^d_+ := \{k \in \mathbf{Z}^d : k_j \geqslant 0, \ \ j = 1, ..., d\}$ and $P_k := \{s \in \mathbf{Z}^d : [2^{k_j - 1}] \leqslant |s_j| < 2^{k_j}, \ j = 1, ..., d\}$ ($[a]$ denotes the integer part of a). The Littlewood–Paley Theorem (see, e.g., [12]) states that for $1 < p < \infty$, there holds the following norms equivalence

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

This theorem can be generalized for the norm $\|\cdot\|_{\mathbf{W}^A_p}$ as follows. Let $S(B, x) := \sup_{\alpha \in B} (\alpha, x)$ be the support function of a subset B of $\mathbf{R}^d$, and

$$\mu(B) := \inf \{ t > 0 : te^j \in \operatorname{conv}(B \cup \{0\}), \ j = 1, ..., d\},$$

where $\operatorname{conv} G$ denotes the convex hull of G and $\{e^j\}_{j=1}^d$ is the canonical basis in $\mathbf{R}^d$.

Lemma 6. *Let $1 < p < \infty$ and A be a finite subset of $\mathbf{R}^d$. Then we have*

$$\|f\|_{\mathbf{W}^A_p} \asymp \left\| \sum_{k \in \mathbf{Z}^d_+} (|2^{S(A, k)} \delta_k f|^2)^{1/2} \right\|_p .$$

Proof. For the proof of this lemma see [6]. ∎

We now give descriptions of quasi-norm equivalences for $\|\cdot\|_{\mathbf{B}^A_{p,\theta}}$. For univariate functions $f \in L_p(\mathbf{T})$, the convolution $V_m f := f * V_m$ defines the de la Vallée Poussin sum of f. Next, we put

$$v_0 f := V_1 f; \ v_k f := V_{2^k} f - V_{2^{k-1}} f, \ k = 1, 2, ...$$

For multivariate functions $f \in L_p(\mathbf{T}^d)$, the mixed operator $v_k, k \in \mathbf{Z}^d_+$, is defined by

$$v_k f := v_{k_d} \cdots v_{k_1} f,$$

where the univariate operator v_{k_j} is applied to the variable x_j. Note that $v_k f$ is a trigonometric polynomial of order $< 2^{k_j+1}$ in the variable x_j, $j = 1, ..., d$.

LEMMA 7. *Let $1 \leqslant p \leqslant \infty$, $0 < \theta \leqslant \infty$, and A be a finite subset of $\mathbf{R}^d$. Then we have for $1 < p < \infty$*

$$\|f\|_{\mathbf{B}^A_{p,\theta}} \asymp \left(\sum_{k \in \mathbf{Z}^d_+} (2^{S(A,k)} \|\delta_k f\|_p)^\theta \right)^{1/\theta}, \; \theta < \infty,$$

and for $1 \leqslant p \leqslant \infty$

$$\|f\|_{\mathbf{B}^A_{p,\theta}} \asymp \left(\sum_{k \in \mathbf{Z}^d_+} (2^{S(A,k)} \|v_k f\|_p)^\theta \right)^{1/\theta}, \; \theta < \infty,$$

with the change to supremum for $\theta = \infty$.

Proof. The proof of this lemma is similar to those of Theorem 2.1 in [4] and of Theorem 1.1 in Chapter II of [15]. ▮

If $\mu(A) > 0$, then the set

$$\Gamma(\xi) := \{k \in \mathbf{Z}^d_+ : S(A,k) \leqslant \xi\}$$

is a finite subset of $\mathbf{Z}^d_+$ for any $\xi \geqslant 0$ (see [4]). Put $\bar{\Gamma}(\xi) := \mathbf{Z}^d_+ \setminus \Gamma(\xi)$ and $A_t := A - t\mathbf{1}$ for a real number t. The following estimates of sums of exponents, taken over the elements from $\Gamma(\xi)$ and $\bar{\Gamma}(\xi)$, were also proved in [4].

LEMMA 8. *Let $\delta, \varepsilon > 0$, and A be a finite subset of $\mathbf{R}^d$ with $\mu(A) > \varepsilon d$. Then we have*

$$\sum_{k \in \Gamma(\xi)} 2^{(\mathbf{1},k)} \asymp 2^{\xi/r} \xi^\nu, \qquad \sum_{k \in \bar{\Gamma}(\xi)} 2^{-\delta S(A_\varepsilon, k)} \asymp 2^{-\delta(1-\varepsilon/r)\xi} \xi^\nu.$$

Moreover, $r > \varepsilon$ if and only if $\mu(A) > \varepsilon d$.

In what follows we use the notation $a_+ := \max\{a, 0\}$. Let the linear projection $P(\xi)$ be defined by

$$P(\xi, f) := \sum_{k \in \Gamma(\xi)} v_k f.$$

LEMMA 9. *Let $1 \leqslant p, q \leqslant \infty$, $0 < \theta, \tau \leqslant \infty$ and A be a finite subset of $\mathbf{R}^d$ with $\mu(A) > (d/p - d/q)_+$. Then we have for every $f \in \mathbf{B}^A_{p,\theta}$*

$$\lim_{\xi \to \infty} \|f - P(\xi, f)\|_{\mathbf{B}_{q,\tau}} = 0.$$

Proof. Clearly, by virtue of the standard inequality $\|\cdot\|_{\mathbf{B}^A_{p,\theta}} \ll \|\cdot\|_{\mathbf{B}^A_{q,\tau}}$, for $p \geqslant q$ and $\theta \leqslant \tau$, it suffices to prove the lemma for the case $p < q$ and $\theta > \tau$. By Lemma 7 we have

$$\|f - P(\xi, f)\|^\tau_{B_{q,\tau}} \asymp \sum_{k \in \bar{\Gamma}(\xi)} \|v_k f\|^\tau_q.$$

Since $v_k f \in \mathscr{T}_{2^k}$, the well-known Nikol'skii inequality (see [12]) gives $\|v_k f\|_q \ll 2^{\varepsilon(\mathbf{1}, k)} \|v_k f\|_p$, where $\varepsilon := d/p - d/q$. Therefore, by the Hölder inequality

$$\begin{aligned}\|f - P(\xi, f)\|^\tau_{B_{q,\tau}} &\ll \sum_{k \in \bar{\Gamma}(\xi)} (2^{\varepsilon(\mathbf{1}, k)} \|v_k f\|_p)^\tau \\ &\leqslant \left(\sum_{k \in \bar{\Gamma}(\xi)} (2^{S(A_\varepsilon, k)} \|v_k f\|_p)^\theta \right)^{\tau/\theta} \left(\sum_{k \in \bar{\Gamma}(\xi)} 2^{-\rho S(A_\varepsilon, k)} \right)^{\tau/\rho},\end{aligned}$$

where $\rho := \theta\tau/(\theta - \tau) > 0$. Hence, by Lemmas 7 and 8

$$\|f - P(\xi, f)\|_{\mathbf{B}_{q,\tau}} \ll 2^{-(1 - \varepsilon/r)\xi} \xi^{\nu/\rho} \|f\|_{\mathbf{B}^A_{p,\theta}}.$$

Since $\varepsilon < r$ by Lemma 8, the right-hand side in the last inequality tends to zero as $\xi \to \infty$ for $f \in \mathbf{B}^A_{p,\theta}$. The lemma is proved. ∎

For $m \in \mathbf{Z}^d$, denote by $\mathscr{T}_m$ the space of all trigonometric polynomials of order $\leqslant m_j$ in the variable x_j, $j = 1, ..., d$. It is easy to check that for every $f \in \mathscr{T}_m$

$$f = \sum_{k \in U(m)} f(hk)\, S_m(\cdot - hk), \tag{33}$$

and in addition there hold the Marcinkiewicz type inequalities

$$\|f\|_p \asymp \prod_{j=1}^{d} m_j^{-1/p} \|\{f(hk)\}\|_{l_p^{s(m)}}, \; 1 \leqslant p \leqslant \infty, \tag{34}$$

(see [5]) where the kernel S_m is defined in (14), $h := (2\pi/3)(m_1^{-1}, ..., m_d^{-1})$, $U(m) := \{k \in \mathbf{Z}_+^d : k_j < 3m_j, j = 1, ..., d\}$ and $s(m) := \prod_{j=1}^{d} (3m_j)$.

4. NON-LINEAR WIDTHS OF CLASSES OF FUNCTIONS

Again, denote by γ_n any one of α_n, τ_n, τ'_n, β_n, a_n and δ_n. In this section we find the asymptotic orders of γ_n of $\mathbf{SW}_p^A$ and $\mathbf{SB}_{p,\theta}^A$ in the space L_q.

THEOREM 1. *Let $1 < p, q < \infty$, $2 \leqslant \theta \leqslant \infty$ and A be a finite subset of $\mathbf{R}^d$ with $\mu(A) > \max\{0, d/p - d/q, d/p - d/2\}$. Then we have*

$$\gamma_n(\mathbf{SB}_{p,\theta}^A, L_q) \asymp (n/\log^\nu n)^{-r} (\log^\nu n)^{1/2 - 1/\theta}, \tag{35}$$

$$\gamma_n(\mathbf{SW}_p^A, L_q) \asymp (n/\log^\nu n)^{-r}. \tag{36}$$

In addition, we can explicitly construct a positive homogeneous continuous mapping G^*: $Y \to \mathbf{M}_n$ such that the asymptotic order $E(n)$ is achieved by the continuous algorithm S of n-term approximation by $\mathbf{V}$, i.e.,

$$\sup_{f \in SY} \|f - S(f)\|_q \ll E(n),$$

where $S := R_{\mathbf{V}} \circ G^*$. $E(n)$ is the right-hand side of either (35) or (36) and Y is either $\mathbf{B}_{p,\theta}^A$ or $\mathbf{W}_p^A$, respectively.

Theorem 1 will be proved from the following

THEOREM 2. *Let $1 \leqslant p, q \leqslant \infty$, $1 \leqslant \tau \leqslant \theta \leqslant \infty$ and A be a finite subset of $\mathbf{R}^d$ with $\mu(A) > (d/p - d/q)_+$. Then we have*

$$\gamma_n(\mathbf{SB}_{p,\theta}^A, \mathbf{B}_{q,\tau}) \asymp E_{\theta,\tau}(n), \tag{37}$$

where

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

In addition we can explicitly construct a positive homogeneous continuous mapping $G^*\colon \mathbf{B}^A_{p,\theta} \to \mathbf{M}_n$ *such that the asymptotic order* $E_{\theta,\tau}(n)$ *is achieved by the continuous algorithm* S *of* n*-term approximation by* $\mathbf{V}$, *i.e.*,

$$\sup_{f \in \mathbf{SB}^A_{p,\theta}} \|f - S(f)\|_q \ll E_{\theta,\tau}(n), \tag{38}$$

where $S := R_{\mathbf{V}} \circ G^*$.

Proof of Theorem 2. We first construct a positive homogeneous continuous mapping $G^*\colon \mathbf{B}^A_{p,\theta} \to \mathbf{M}_n$ satisfying (38), and, therefore by the definition of α_n, τ_n, Lemma 5 and the inequalities (7)–(8) prove the upper bound of (37). We carry this out for the case of α_n and $p<q$, $\theta>\tau$. The remaining cases can be similarly proved with slight modifications. Without lost of generality we can assume that $n=2m$, an even number. From (33)–(34) and Lemmas 7 and 9 one can prove that a function f belongs to $\mathbf{B}^A_{p,\theta}$ if and only if f can be represented by a series

$$f = \sum_{k \in \mathbf{Z}^d_+} \sum_{s \in Q_k} f_{k,s}\, \varphi^k_s,$$

converging in the norm of $\mathbf{B}_{q,\tau}$, and, in addition

$$\|f\|_{\mathbf{B}^A_{p,\theta}} \asymp \|D(f)\|_{\mathbf{b}^{\mathbf{N}}_{p,\theta}(\boldsymbol{\Lambda})}, \quad \|f\|_{\mathbf{B}_{q,\tau}} \asymp \|D(f)\|_{\mathbf{b}^{\mathbf{N}}_{q,\tau}(\boldsymbol{\Lambda}')}, \tag{39}$$

where Q_k and h^k are in (15), $\mathbf{N} := \{N_k\}_{k \in \mathbf{Z}^d_+}$, $N_k = |Q_k|$,

$$\boldsymbol{\Lambda} := \{\lambda_k\}_{k \in \mathbf{Z}^d_+},\ \lambda_k = 2^{-S(A,k)+(\mathbf{1},k)/p}, \quad \boldsymbol{\Lambda}' := \{\lambda'_k\}_{k \in \mathbf{Z}^d_+},\ \lambda'_k = 2^{(\mathbf{1},k)/q}\},$$

and D is the positive homogeneous continuous mapping from $\mathbf{B}^A_{p,\theta}$ into $\mathbf{B}^{\mathbf{N}}_{q,\tau}(\boldsymbol{\Lambda}')$, given by

$$D(f) := \{x^k\}_{k \in \mathbf{Z}^d_+}, \quad x^k := \{f_{k,s}\}_{s \in Q_k}.$$

Take a function $\xi = \xi(n)$ satisfying the condition

$$3(4^d + 2^d)\, J(\xi) \leqslant m < C\, 2^{\xi/r} \xi^{\nu}, \tag{40}$$

where C is an absolute constant whose value will be chosen below, and $J(\xi) := \sum_{k \in \Gamma(\xi)} 2^{(\mathbf{1},k)}$. Let the sequence $\{n_k\}_{k \in \mathbf{Z}^d_+}$ be given by

$$n_k := \begin{cases} |Q_k|, & \text{for} \quad k \in \Gamma(\xi) \\ [3C^{-1} m 2^{(1-\varepsilon/r)\xi} \xi^{-\nu} 2^{-S(A_\varepsilon,k)}], & \text{for} \quad k \in \bar{\Gamma}(\xi), \end{cases}$$

and the sequence $\{b_k\}_{k \in \mathbf{Z}^d_+}$ be given by

$$b_k := \begin{cases} 0, & \text{for} \quad k \in \Gamma(\xi) \\ (3C^{-1} m 2^{(1-\varepsilon/r)\xi} \xi^{-\nu} 2^{-S(A_\varepsilon,k)})^{-\varepsilon}, & \text{for} \quad k \in \bar{\Gamma}(\xi), \end{cases}$$

where $\varepsilon := 1/p - 1/q > 0$. By Lemma 1 $\alpha_n(B_p^{|Q_k|}, l_q^{|Q_k|}) \leqslant b_k$, $k \in \mathbf{Z}_+^n$. By using Lemma 8 we can easily verify that $\sum_{k \in \mathbf{Z}_+^d} n_k < \infty$. Note that $0 < \varepsilon < r$ by Lemma 8, and $\alpha_{n_k}(B_p^{|Q_k|}, l_q^{|Q_k|}) = 0$ for $k \in \Gamma(\xi)$. Hence, $\alpha_{n_k}(B_p^{|Q_k|}, l_q^{|Q_k|}) \leqslant b_k$ for $k \in \mathbf{Z}_+^d$. Denote by $H(\xi)$ the subset of $\bar{\Gamma}(\xi)$ which consists of all $k \in \bar{\Gamma}(\xi)$ such that the set $\{x \in \mathbf{R}_+^d : k_j - 1 \leqslant x_j < k_j,\ j = 1, ..., d\}$ has a non-empty intersection with the set $\{x/\xi \in A_+^o : (\mathbf{1}, x) = 1/r\}$. We put $s = s(n) := |H(\xi)|$, $\xi = \xi(n)$. Suppose that $\mathbf{Z}_+^d$ is rearranged as $\mathbf{Z}_+^d = \{k_j\}_{j=1}^{\infty}$ so that

$$\mu_{k_1} \geqslant \mu_{k_2} \geqslant \cdots \mu_{k_j} \geqslant \cdots, \tag{41}$$

where $\mu_k := b_k \lambda_k / \lambda'_k$. By Lemma 2 we construct a positive homogeneous continuous mapping $G\colon \mathbf{b}_{p,\theta}^{\mathbf{N}}(\mathbf{\Lambda}) \to \mathbf{M}_{n'}$ so that

$$\|x - R_{\mathscr{E}}(G(x))\|_{\mathbf{B}_{q,\tau}^{\mathbf{N}}(\mathbf{\Lambda}')} \leqslant F_{\theta,\tau}(s) \|x\|_{\mathbf{B}_{p,\theta}^{\mathbf{N}}(\mathbf{\Lambda})}, \tag{42}$$

where $\mathscr{E}$ is the canonical basis in $\mathbf{b}_{q,\tau}^{\mathbf{N}}(\mathbf{\Lambda}')$ and $F_{\theta,\tau}(s)$ is defined by the formula (20) for our case, and

$$n' = \sum_{j=1}^{s-1} N_{k_j} + \sum_{j=s}^{\infty} n_{k_j}.$$

We define the positive homogeneous continuous mapping $G^*\colon \mathbf{B}_{p,\theta}^{A} \to \mathbf{M}_{n'}$ by $G^* := G \circ D$. Note that $D(f - R_{\mathbf{V}}(G^*(f))) = D(f) - R_{\mathscr{E}}(G(D(f)))$. Hence, by (39) and (40) it is easy to check that

$$\|f - R_{\mathbf{V}}(G^*(f))\|_{\mathbf{B}_{q,\tau}} \ll F_{\theta,\tau}(s)\ \|f\|_{\mathbf{B}_{p,\theta}^{A}}. \tag{43}$$

We will check the inequalities $n' \leqslant n$ for C large enough and

$$F_{\theta,\tau}(s) \leqslant E_{\theta,\tau}(n). \tag{44}$$

This implies that for the positive homogeneous continuous mapping G^* from $\mathbf{B}_{p,\theta}^{A}$ into $\mathbf{M}_n$, there holds the estimate (38).

By the definition of the quantity v in (11) we see at once that $|H(\xi)| \asymp \xi^v$ and, moreover, $\mu_k = \sup_{k' \in \bar{\Gamma}(\xi)} \mu_{k'}$ for any $k \in H(\xi)$. Hence according to (41) we have by Lemma 2

$$n' \leqslant \sum_{k \in H(\xi)} |Q_k| + \sum_{k \in \Gamma(\xi)} |Q_k| + \sum_{k \in \bar{\Gamma}(\xi)} 3C^{-1} m 2^{(1-\varepsilon/r)\xi} \xi^{-v} 2^{-S(A_\varepsilon, k)}.$$

From the inclusion $k-\mathbf{1}\in\Gamma(\xi)$, $k\in H(\xi)$, the inequality $|H(\xi)|\leqslant|\Gamma(\xi)|$, Lemma 8 and (40), we can continue this estimation as follows

$$n'\leqslant 2^d\sum_{k\in H(\xi)}3.2^{\langle\mathbf{1},k\rangle}+\sum_{k\in\Gamma(\xi)}3.2^{\langle\mathbf{1},k+\mathbf{1}\rangle}+3C^{-1}m2^{(1-\varepsilon/r)\,\xi}\xi^{-\nu}\sum_{k\in\bar{\Gamma}(\xi)}2^{-S(A_\varepsilon,k)}$$
$$\leqslant 3(4^d+2^d)\,J(\xi)+3aC^{-1}m\leqslant n,$$

for $C=C(d,a)$ large enough where a is an absolute constant (see Lemma 8). Thus the inequality $n'\leqslant n$ has been proved.

Put $\rho:=\theta\tau/(\theta-\tau)>0$. From the equality $S(A,\cdot)-\varepsilon(\mathbf{1},\cdot)=S(A_\varepsilon,\cdot)$, Lemmas 2 and 8 we have

$$F^{\rho}_{\theta,\tau}(s)\leqslant\sum_{k\in\bar{\Gamma}(\xi)}\mu_k^{\rho}\leqslant\sum_{k\in\bar{\Gamma}(\xi)}(b_k2^{-S(A_\varepsilon,k)})^{\rho}$$
$$\leqslant(3C^{-1}m2^{(1-\varepsilon/r)\,\xi}\xi^{-\nu})^{-\rho\varepsilon}\sum_{k\in\bar{\Gamma}(\xi)}2^{-\rho(1-\varepsilon)\,S(A_\varepsilon,k)}$$
$$\asymp(m2^{(1-\varepsilon/r)\,\xi}\xi^{-\nu})^{-\rho\varepsilon}\,2^{-\rho(1-\varepsilon)(1-\varepsilon/r)\,\xi}\xi^{\nu}.$$

Hence using (40), by a simple computation we obtain (44).

Let us prove the lower bound of (37). By the inequalities (7)–(8) and Lemma 4 it is sufficient to prove this lower bound for τ_n. We will need some additional inequalities. If W is a compact subset in the finite dimensional normed space X, from the inequality $2a_n(W,X)\geqslant b_n(W,X)$ [18, p. 220] and Lemma 5 it follows that

$$2\tau_n(W,X)\geqslant b_n(W,X). \tag{45}$$

Here the Bernstein n-width $b_n(W,X)$ is defined by

$$b_n(W,X):=\sup_M\sup\{t>0:tSX\cap M\subset W\},$$

where the outer supremum is taken over all $(n+1)$-dimensional linear subspaces M of X. The proof of the following assertion is similar to that of Lemma 2.3 in [9]. If Y is a subspace of the normed linear space X, W is a subset of X and $P\colon X\to Y$ is a linear projection with $\|P(f)\|\leqslant\lambda\,\|f\|(\lambda>0)$ for every $f\in X$, then

$$\tau_n(W,X)\geqslant\lambda^{-1}\tau_n(W,Y). \tag{46}$$

We now proceed with a proof of the lower bound of (37) for τ_n. Because of the inclusion $\mathbf{SB}^A_{\infty,\theta}\subset\mathbf{SB}^A_{p,\theta}$, it is sufficient to treat the case $p=\infty$. Put

$w(\xi) := \max_{k \in \Gamma(\xi)} (\mathbf{1}, k)$, and $Q(\xi) := \{k \in \Gamma(\xi) : (\mathbf{1}, k) \geqslant w(\xi) - d\}$. Denote by $B(\xi)$ the space of all trigonometric polynomials f of the form

$$f = \sum_{k \in Q(\xi)} \sum_{s \in Q_k} f_{k,s} \varphi_s^k,$$

and for $0 < \zeta, \eta \leqslant \infty$ denote by $B(\xi)_{\zeta,\eta}$ the subspace in $\mathbf{B}_{\zeta,\eta}$ which consists of all $f \in B(\xi)$. Lemma 7 and the inequality $S(A, k) \leqslant \xi$, $k \in Q(\xi)$, give $\|f\|_{\mathbf{B}^A_{\infty,\theta}} \asymp 2^\xi \|f\|_{\mathbf{B}_{\infty,\theta}}$, $f \in B(\xi)_{\infty,\theta}$. This implies $a' 2^{-\xi} SB(\xi)_{\infty,\theta} \subset \mathbf{SB}^A_{\infty,\theta}$, with some absolute constant $a' > 0$. Therefore,

$$\tau_n(\mathbf{SB}^A_{\infty,\theta}, \mathbf{B}_{q,\tau}) \gg 2^{-\xi} \tau_n(SB(\xi)_{\infty,\theta}, \mathbf{B}_{q,\tau}).$$

Using (39) and applying (46) to the linear projection

$$P(\xi, f) = \sum_{k \in Q(\xi)} \sum_{s \in Q_k} f_{k,s} \varphi_s^k$$

in the space $\mathbf{B}_{q,\tau}$, we obtain

$$\tau_n(SB(\xi)_{\infty,\theta}, \mathbf{B}_{q,\tau}) \gg \tau_n(SB(\xi)_{\infty,\theta}, B(\xi)_{q,\tau}),$$

and consequently,

$$\tau_n(\mathbf{SB}^A_{\infty,\theta}, \mathbf{B}_{q,\tau}) \gg 2^{-\xi} \tau_n(SB(\xi)_{\infty,\theta}, B(\xi)_{q,\tau}). \tag{47}$$

Let us now give a lower bound for $\tau_n(SB(\xi)_{\infty,\theta}, B(\xi)_{q,\tau})$ by using Lemma 4 with $X = B(\xi)_{q,\tau}$, $Y = B(\xi)_{\infty,\theta}$ and $W = SB(\xi)_{\infty,\theta}$. Note that in our case all these spaces are finite dimensional. Therefore their norms are equivalent and $SB(\xi)_{\infty,\theta}$ is compact in these norms. Thus we have

$$\tau_{n+m}(SB(\xi)_{\infty,\theta}, B(\xi)_{\infty,\theta}) \leqslant \tau_n(SB(\xi)_{\infty,\theta}, B(\xi)_{q,\tau})\, \tau_m(SB(\xi)_{q,\tau}, B(\xi)_{\infty,\theta}), \tag{48}$$

for any m and ξ satisfying the condition

$$\dim B(\xi) > n + m, \tag{49}$$

which obviously implies the inequality $\tau_n(SB(\xi)_{\infty,\theta}, B(\xi)_{q,\tau}) > 0$. Below we will define such an $m = m(n)$. By the definition $b_{n+m}(SB(\xi)_{\infty,\theta}, B(\xi)_{\infty,\theta}) \geqslant 1$, and consequently by (45) $2\tau_{n+m}(SB(\xi)_{\infty,\theta}, B(\xi)_{\infty,\theta}) \geqslant 1$, for any m with the condition (49). This and (48) give

$$2\tau_n(SB(\xi)_{\infty,\theta}, B(\xi)_{q,\tau}) \geqslant 1/\tau_m(SB(\xi)_{q,\tau}, B(\xi)_{\infty,\theta}). \tag{50}$$

This inequality implies that a lower estimate for $\tau_n(SB(\xi)_{\infty,\theta}, B(\xi)_{q,\tau})$ can be obtained from an upper estimate for $\tau_m(SB(\xi)_{q,\tau}, B(\xi)_{\infty,\theta})$. We get this upper estimate by applying the method use to establish the upper bound for (37). We outline a brief proof of this bound. By (39) and the inequalities $\xi/r - d \leqslant w(\xi) \leqslant \xi/r$ we have

$$\|f\|_{B(\xi)_{\infty,\theta}} \asymp \|D(f)\|_{\mathbf{b}^{\mathbf{N}}_{\infty,\theta}}, \quad \|f\|_{B(\xi)_{q,\tau}} \asymp 2^{-\xi/rq} \|D(f)\|_{\mathbf{b}^{\mathbf{N}}_{q,\tau}}, \tag{51}$$

where $\mathbf{N} := \{N_k\}_{k \in Q(\xi)}$, $N_k = |Q_k|$ (for simplicity we use the same notation as in the upper bound for different sequences and quantities). Given n, we define $\xi = \xi(n)$ as any function of the variable n, satisfying the inequalities

$$n < 4^{-d-1} |Q(\xi)| \, 2^{w(\xi)} \ll n. \tag{52}$$

Let the sequence $\{n_k\}_{k \in Q(\xi)}$ be given by $n_k := n^* = [2^{w(\xi)-2d-2}]$ and the sequence $\{b_k\}_{k \in Q(\xi)}$ by $b_k := (2^{w(\xi)-d-2})^{-1/q}$, where $\xi = \xi(n)$. Note that $N_k \leqslant N^* = 6^d . 2^{w(\xi)}$, $k \in Q(\xi)$, and by Lemma 1 $\tau_{n_k}(B^{N_k}_q, l^{N_k}_\infty) \leqslant b_k$, $k \in Q(\xi)$. Let the set $Q(\xi) = \{k_1, ..., k_l\}$, $l = |Q(\xi)|$, be ordered so that $\mu_{k_1} \geqslant \mu_{k_2} \geqslant \cdots \geqslant \mu_{k_l}$, where $\mu_k = b_k$ (note that $l < \infty$). Similarly to (43), by using Lemma 2 for the case $\theta = \tau$ and Lemma 3 for the case $\theta > \tau$ from (51) we establish the following estimate:

$$\tau_m(SB(\xi)_{q,\tau}, B(\xi)_{\infty,\theta}) \ll 2^{\xi/rq} F'(s),$$

where $s = s(n) := [2^{-5d-2} |Q(\xi)|]$, $\quad m = (s-1) N^* + (l - s + 1) n^*$ and

$$F'(s) = \begin{cases} (\sum_{j=1}^{s} \mu_{k_j}^{\rho'})^{1/\rho'}, & \text{for} \quad \theta > \tau \\ \mu_{k_s}, & \text{for} \quad \theta = \tau, \end{cases}$$

with $\rho' := \tau\theta(\tau - \theta)$. Hence,

$$\tau_m(SB(\xi)_{q,\tau}, B(\xi)_{\infty,\theta}) \ll \xi^{-\nu(1/\tau - 1/\theta)}. \tag{53}$$

Note that $B(\xi)$ contains the space of dimension $|\Delta(\xi)| \, 2^{w(\xi)-d}$ of all trigonometric polynomials f of the form $f = \sum_{k \in \Delta(\xi)} \delta_k f$, where $\Delta(\xi) := \{k \in \Gamma(\xi) : (\mathbf{1}, k) = w(\xi)\}$. Hence, $\dim B(\xi) > |\Delta(\xi)| \, 2^{w(\xi)-d} \geqslant 4^{-d} |Q(\xi)| \, 2^{w(\xi)}$. On the other hand, it is easy to verify that $m \leqslant 2^{-2d-1} |Q(\xi)| \, 2^{w(\xi)}$. Hence, by (52) we can see that m satisfies (49). Combining (47), (50), (53) gives

$$\tau_n(\mathbf{SB}^{\Delta}_{\infty,\theta}, \mathbf{B}_{q,\tau}) \gg 2^{-\xi} \xi^{\nu(1/\tau - 1/\theta)} \asymp E_{\theta,\tau}(n).$$

The lower bound of (37) is proved. ∎

Proof of Theorem 1. By using the Littlewood-Paley Theorem and Lemmas 6–7 it is easy to verify the inequalities

$$\|f\|_{\mathbf{B}_{\eta(q),2}} \ll \|f\|_q \ll \|f\|_{\mathbf{B}_{\zeta(q),2}} \quad \|f\|_{\mathbf{B}^A_{\eta(p),2}} \ll \|f\|_{\mathbf{W}^A_p} \ll \|f\|_{\mathbf{B}^A_{\zeta(p),2}},$$

for $1 < p, q < \infty$ where $\eta(t) := \min\{t, 2\}$, $\zeta(t) := \max\{t, 2\}$. Hence it follows that

$$\gamma_n(\mathbf{SB}^A_{p,\theta}, \mathbf{B}_{\eta(q),2}) \ll \gamma_n(\mathbf{SB}^A_{p,\theta}, L_q) \ll \gamma_n(\mathbf{SB}^A_{p,\theta}, \mathbf{B}_{\zeta(q),2}),$$
$$\gamma_n(\mathbf{SB}^A_{\zeta(p),2}, \mathbf{B}_{\eta(q),2}) \ll \gamma_n(\mathbf{SW}^A_p, L_q) \ll \gamma_n(\mathbf{SB}^A_{\eta(p),2}, \mathbf{B}_{\zeta(q),2}).$$

This and Theorem 2 imply Theorem 1. ∎

5. BEST n-TERM APPROXIMATION OF CLASSES OF FUNCTIONS

In this section we obtain the asymptotic order of the best n-term approximation by $\mathbf{V}$ of $\mathbf{SW}^A_p$ and $\mathbf{SB}^A_{p,\theta}$ in the space L_q.

THEOREM 3. *Under the conditions of Theorem* 1 *we have*

$$\sigma_n(\mathbf{SB}^A_{p,\theta}, \mathbf{V}, L_q) \asymp (n/\log^\nu n)^{-r} (\log^\nu n)^{1/2-1/\theta},$$
$$\sigma_n(\mathbf{SW}^A_p, \mathbf{V}, L_q) \asymp (n/\log^\nu n)^{-r}.$$

As in the proof of Theorem 1, this theorem is obtained from the following

THEOREM 4. *With the conditions and notation of Theorem* 2 *we have*

$$\sigma_n(\mathbf{SB}^A_{p,\theta}, \mathbf{V}, \mathbf{B}_{q,\tau}) \asymp E_{\theta,\tau}(n).$$

To prove Theorem 4 we need some auxiliary lemmas.

LEMMA 10. *Let the linear space* L *be* (*quasi-*) *normed by two* (*quasi-*) *norms* $\|\cdot\|_X$ *and* $\|\cdot\|_Y$, *and* W *be a subset of* L. *Assume that* Φ *is a family of elements in* X *such that* $\sigma_m(W, \Phi, X) > 0$. *Then we have*

$$\sigma_{n+m}(W, \Phi, Y) \leqslant \sigma_n(SX, \Phi, Y)\, \sigma_m(W, \Phi, X).$$

Proof of Lemma 10. The proof of this lemma is similar to that of Lemma 4. ∎

LEMMA 11. *Let* $1 \leqslant p, \theta \leqslant \infty$, $\mathbf{N} = \{N_k\}_{k=1}^s$ *be a finite sequence of natural numbers with* $\beta N \leqslant N_k \leqslant N$, $k = 1, 2, ..., s$, *for some* $\beta > 0$. *Let*

$\Phi=\{\varphi_k\}_{k=1}^{M}$, $M=\sum_{k=1}^{s} N_k$, *be any family of elements of* $l^{\mathbf{N}}_{p,\theta}$, *and* $K_{p,\theta}$ *be an arbitrary m-dimensional cross-section of the unit ball* $\mathbf{S}^{\mathbf{N}}_{p,\theta}$. *Assume that* $\lambda M \leqslant m \leqslant M$ *for some* $\lambda>0$, *and* $n \leqslant m/2$. *Then we have*

$$\sigma_n(K_{p,\theta}, \Phi, l^{\mathbf{N}}_{p,\theta}) \geqslant C,$$

where $C=C(\beta,\lambda)>0$.

Proof of Lemma 11. Lemma 2 in [10] gives

$$\sigma_n(K_{\infty,\infty}, \Phi, \mathbf{b}^{\mathbf{N}}_{1,1}) \geqslant C(\lambda)\, M. \tag{54}$$

Hence we obtain Lemma 11. Indeed, by the inequality $\|\cdot\|_{\mathbf{b}^{\mathbf{N}}_{p,\theta}} \leqslant s^{1/\theta} N^{1/p}\, \|\cdot\|_{\mathbf{b}^{\mathbf{N}}_{\infty,\infty}}$, we have $K_{p,\theta} \supset s^{-1/\theta} N^{-1/p} K_{\infty,\infty}$. On the other hand, it is easy to check that $\|\cdot\|_{\mathbf{b}^{\mathbf{N}}_{p,\theta}} \geqslant s^{1/\theta-1} N^{1/p-1}\, \|\cdot\|_{\mathbf{b}^{\mathbf{N}}_{1,1}}$. Therefore, by (54)

$$\sigma_n(K_{p,\theta}, \Phi, \mathbf{b}^{\mathbf{N}}_{p,\theta}) \geqslant (s^{-1/\theta} N^{-1/p})(s^{1/\theta-1} N^{1/p-1})\, \sigma_n(K_{\infty,\infty}, \Phi, \mathbf{b}^{\mathbf{N}}_{1,1}) \geqslant \beta C(\lambda). \quad \blacksquare$$

Proof of Theorem 4. The upper bound for this theorem follows from (38) in Theorem 2. The lower bound can be proved in a manner similar to the proof of the lower bound for (37) in Theorem 2, using Lemmas 10 and 11. ▮

Example. Let $\alpha \in \mathbf{R}^d$ with positive coordinates. Denote by $\mathring{W}^{\alpha}_p$ and $\mathring{B}^{\alpha}_{p,\theta}$ the subspaces of $\mathbf{W}^{\{\alpha\}}_p$ and $\mathbf{B}^{\{\alpha\}}_{p,\theta}$, respectively, which consists of all functions f such that

$$\int_{-\pi}^{\pi} f(x)\, dx_j = 0, \qquad j=1, ..., d.$$

Without loss of generality we can assume that

$$0<r=\alpha_1=\cdots=\alpha_v=\alpha_{v+1}<\alpha_{v+2}\leqslant \cdots \leqslant \alpha_d \quad (0 \leqslant v \leqslant d-1).$$

If $1<p, q<\infty$, $2 \leqslant \theta \leqslant \infty$ and $r>\max\{0, d/p-d/2, d/p-d/q\}$, then from Theorems 1 and 3 one can easily deduce that

$$\gamma_n(S\mathring{B}^{\alpha}_{p,\theta}, L_q) \asymp \sigma_n(S\mathring{B}^{\alpha}_{p,\theta}, \mathbf{V}, L_q) \asymp (n/\log^{v} n)^{-r}\, (\log^{v} n)^{1/2-1/\theta},$$

$$\gamma_n(S\mathring{W}^{\alpha}_{p}, L_q) \asymp \sigma_n(S\mathring{W}^{\alpha}_{p}, \mathbf{V}, L_q) \asymp (n/\log^{v} n)^{-r}.$$

REFERENCES

1. R. DeVore, Nonlinear approximation, *Acta Numerica* **7** (1998), 51–150.
2. R. DeVore, R. Howard, and C. Micchelli, Optimal non-linear approximation, *Manuscripta Math.* **63** (1989), 469–478.
3. R. DeVore, G. Kyriazis, D. Leviatan, and V. Tikhomirov, Wavelet compression and nonlinear n-widths, *Adv. Comp. Math.* **1** (1993), 194–214.
4. Dinh Dung (Din' Zung), Approximation by trigonometric polynomials of functions of several variables on the torus, *Mat. Sb.* **131**(173) (1986), 251–267. [English transl. in *Math. USSR Sb.* **59** (1988), 247–267]
5. Dinh Dung, On optimal recovery of multivariate periodic functions, *in* "Harmonic Analysis" (S. Igari, Ed.), Springer-Verlag, Tokyo/Berlin/Heidelberg, 1991, pp. 96–105.
6. Dinh Dung, Optimal non-linear approximation of functions with a mixed smoothness, *East J. Approx.* **4** (1998), 101–112.
7. Dinh Dung, On nonlinear n-widths and n-term approximation, *Vietnam J. Math.* **26** (1998), 165–176.
8. Dinh Dung, On best continuous methods in n-term approximation, *Vietnam J. Math.* **26** (1998), 367–371.
9. Dinh Dung and Vu Quoc Thanh, On non-linear n-widths, *Proc. Amer. Math. Soc.* **124** (1996), 2757–2765.
10. B. Kashin and V. Temlyakov, On best m-term approximation and the entropy of sets in the space L^1, *Mat. Zametki* **56** (1994), 57–86. [English translation in *Math. Notes* **56** (1994), 1137–1157]
11. P. Mathé, s-Number in information-based complexity, *J. Complexity* **6** (1990), 41–66.
12. S. Nikol'skii, "Approximation of Functions of Several Variables and Embedding Theorems," Springer-Verlag, Berlin, 1975.
13. J. Ratsaby and V. Maiorov, On the value of partial information for learning from examples, *J. of Complexity* **13** (1997), 509–544.
14. M. Stesin, Alexandroff widths of finite dimensional sets and classes of smooth functions, *Dokl. Akad. Nauk SSSR* **220** (1974), 38–41.
15. V. Temlyakov, Approximation of functions with bounded mixed derivatives, *Trudy Mat. Inst. Steklov* **178** (1986), 1–112. [English transl. in *Proc. Steklov Inst. Math.* **178** (1989), 1–121]
16. V. Temlyakov, Nonlinear Kolmogorov's widths, *Mat. Zametki* **63** (1998), 891–902.
17. V. Temlyakov, Greedy algorithms with regard to the multivariate systems with a special structure, *Constr. Approx.*, to appear.
18. V. Tikhomirov, "Some Topics in Approximation Theory," Moscow State Univ., Moscow, 1976.