N-Widths and ε-dimensions for high-dimensional approximations

Dinh Dũng[a]*, Tino Ullrich[b]

[a] Vietnam National University, Hanoi, Information Technology Institute
144 Xuan Thuy, Hanoi, Vietnam
dinhzung@gmail.com

[b]Hausdorff-Center for Mathematics and Institute for Numerical Simulation
53115 Bonn, Germany

Dedicated to the memory of Professor S.M. Nikol'skij

March 4, 2013 -- Version R4

Abstract

In this paper, we study linear trigonometric hyperbolic cross approximations, Kolmogorov n-widths $d_n(W, H^\gamma)$, and ε-dimensions $n_\varepsilon(W, H^\gamma)$ of periodic d-variate function classes W with anisotropic smoothness, where d may be large. We are interested in finding the accurate dependence of $d_n(W, H^\gamma)$ and $n_\varepsilon(W, H^\gamma)$ as a function of two variables n, d and ε, d, respectively. Recall that n, the dimension of the approximating subspace, is the main parameter in the study of convergence rates with respect to n going to infinity. However, the parameter d may seriously affect this rate when d is large. We construct linear approximations of functions from W by trigonometric polynomials with frequencies from hyperbolic crosses and prove upper bounds for the error measured in isotropic Sobolev spaces H^γ. Furthermore, in order to show the optimality of the proposed approximation, we prove upper and lower bounds of the corresponding n-widths $d_n(W, H^\gamma)$ and ε-dimensions $n_\varepsilon(W, H^\gamma)$. Some of the received results imply that the curse of dimensionality can be broken in some relevant situations.

*Corresponding author. Email: dinhzung@gmail.com.

Keywords High-dimensional approximation · Trigonometric hyperbolic cross space · Kolmogorov n-widths · ε-dimensions · Sobolev space · Function classes with anisotropic smoothness.

Mathematics Subject Classifications (2010) 42A10; 41A25; 41A63

Communicated by Wolfgang Dahmen

1 Introduction

In recent decades, there has been increasing interest in solving problems that involve functions depending on a large number d of variables. These problems arise from many applications in mathematical finance, chemistry, physics, especially quantum mechanics, and meteorology. It is not surprising that these problems can almost never be solved analytically such that one is interested in a proper framework and efficient numerical methods for an approximate treatment. Classical methods suffer the "curse of dimensionality" coined by Bellmann [2]. In fact, the computation time typically grows exponentially in d, and the problems become intractable already for mild dimensions d without further assumptions. A classical model, widely studied in literature, is to impose certain smoothness conditions on the function to be approximated; in particular, it is assumed that mixed derivatives are bounded. This is the typical situation for which "hyperbolic crosses" are made for. Trigonometric polynomials with frequencies in hyperbolic crosses have been widely used for approximating functions with a bounded mixed derivative or difference. These classical trigonometric hyperbolic crosses date back to Babenko [1]. Let us also mention "sparse grids" in this context which can be seen as the counterpart of hyperbolic crosses on the spatial domain. Sparse grids are discrete point sets consisting of significantly fewer points than a full tensor product grid. First considered by Smolyak [31] they turned out to be suitable for sampling recovery of functions and numerical integration. For further sources on hyperbolic crosses and sparse grids in this classical context we refer to [13, 12, 14, 38, 34, 42] and the references therein. Later on, these terminologies were extended to approximations by wavelets [8, 35], to B-splines [15, 36], and even to algebraic polynomials where frequencies are replaced by dyadic scales or the degree of algebraic polynomials [6, 7]. Hyperbolic cross and sparse grid techniques have applications in quantum mechanics and PDEs [45, 46, 47, 20], finance [18], numerical solution of stochastic PDEs [6, 7, 32, 33], and data mining [17] to mention just a few (see also the surveys [4] and [19] and the references therein).

In this paper, we study linear trigonometric hyperbolic cross approximations, Kolmogorov n-widths $d_n(W, H^\gamma)$, and ε-dimensions $n_\varepsilon(W, H^\gamma)$ of d-variate function classes W with anisotropic smoothness properties where d may be large. The approximation error is measured in an isotropic Sobolev space H^γ, which includes the L_2-metric as a special case. We are interested in finding the accurate dependence of $d_n(W, H^\gamma)$ and $n_\varepsilon(W, H^\gamma)$ as a function of two variables n, d and ε, d, respectively. Recall that n, the dimension of the approximating subspace, is the main parameter in the study of convergence rates with respect to n going to infinity. However, the parameter d may seriously affect this rate when d is large.

Recall the notion of the Kolmogorov n-widths [23] and linear n-widths introduced by Tikhomirov [39]. If X is a normed space and W a subset in X then the Kolmogorov n-width $d_n(W,X)$ is given by

$$d_n(W,X) := \inf_{L_n} \sup_{f \in W} \inf_{g \in L_n} \|f-g\|_X,$$

where the outer inf is taken over all linear manifolds L_n in X of dimension at most n. A different worst-case setting is represented by the linear n-width $\lambda_n(W,X)$ given by

$$\lambda_n(W,X) := \inf_{\Lambda_n} \sup_{f \in W} \|f - \Lambda_n(f)\|_X$$

where the inf is taken over all linear operators Λ_n in X with rank at most n. It represents a characterization of the best linear approximation error. There is a vast amount of literature on optimal linear approximations and the related Kolmogorov and linear n-widths [40], [30], especially for d-variate function classes [38]. In this paper, we are interested in measuring the approximation error in H^γ, therefore we can assume X to be a Hilbert space H. In this case both concepts coincide, i.e.,

$$d_n(W,H) \ = \ \lambda_n(W,H)$$

holds true. Indeed, orthogonal projections onto a finite dimensional space in H give the best approximation by its elements. Hence, it is sufficient to investigate linear approximations in H^γ and the optimality of the approximation in terms of $d_n(W,H^\gamma)$.

Let us recall some classical results in this direction. For the unit balls U^β and $U^{\alpha\mathbf{1}}$ of the periodic d-variate isotropic Sobolev space H^β, $\beta > 0$, and the space $H^{\alpha\mathbf{1}}$ with mixed smoothness $\alpha > 0$, the following well-known estimates hold true. Note, that we have the coincidences $H^\beta = H^{0,\beta}$ and $H^{\alpha\mathbf{1}} = H^{\alpha,0}$ with respect to (2.6) below. It holds

$$A(\beta,d)n^{-\beta/d} \ \leq \ d_n(U^\beta, L_2) \ \leq \ A'(\beta,d)n^{-\beta/d}, \tag{1.1}$$

and

$$B(\alpha,d)n^{-\alpha}(\log n)^{\alpha(d-1)} \ \leq \ d_n(U^{\alpha\mathbf{1}}, L_2) \ \leq \ B'(\alpha,d)n^{-\alpha}(\log n)^{\alpha(d-1)}. \tag{1.2}$$

Here, $A(\beta,d)$, $A'(\beta,d)$, $B(\alpha,d)$, $B'(\alpha,d)$ denote certain constants which are usually not computed explicitly. The inequalities (1.1) are a direct generalization of the first result on n-widths proved by Kolmogorov [23] (see also [24, 186–189]) where the exact values of n-widths were obtained for the univariate case. The inequalities (1.2) were proved by Babenko [1] already in 1960, where a linear approximation on hyperbolic cross spaces of trigonometric polynomial is used. These estimates are quite satisfactory if d, the number of variables, is small.

In computational mathematics, the so-called ε-dimension $n_\varepsilon = n_\varepsilon(W,H)$ is used to quantify the problem's complexity. In our setting it is defined as the inverse of $d_n(W,H)$. In fact, the quantity $n_\varepsilon(W,H)$ is the minimal number n_ε such that the approximation of W by a suitably chosen n_ε-dimensional subspace L in H (measured in terms of Kolmogorov n-widths) yields the approximation error $\le \varepsilon$ (see [10], [11], [16]). We provide upper and lower bounds of this quantity together with the corresponding n-widths in this paper. The quantity n_ε represents a special case of the information complexity which is defined as the minimal number $n(\varepsilon,d)$ of information needed to solve the d-variate problem within error ε (see [26, 4.1.4]). It is the key to study tractability of various multivariate problems. We refer the reader to the monographs [26, 29] for surveys and further references in this direction. In fact, in high-dimensional settings, i.e., if d is large, it turns out that the smoothness of the isotropic Sobolev class U^β is not suitable. In (1.1) the curse of dimensionality occurs since here $n_\varepsilon \ge C(\beta,d)\varepsilon^{-d/\beta}$. However, the class $U^{\alpha\mathbf{1}}$ is more appropriate for high-dimensional problems [4] since we have $n_\varepsilon = O(\varepsilon^{-1/\alpha}|\log\varepsilon|^{d-1})$. In this paper, we extend and refine existing estimates. In particular, we give the lower and upper bounds for constants $B(\alpha,d)$, $B'(\alpha,d)$ in (1.2) with regards to α, d.

We are especially concerned with measuring the approximation error in the isotropic smoothness space H^γ. To motivate this issue let us consider a Galerkin method for approximating the solution of a general elliptic variational problem. Let $a : H^\gamma \times H^\gamma \to \mathbb{R}$ be a bilinear symmetric form and $f \in H^{-\gamma}$, where $H^\gamma = H^\gamma(\mathbb{T}^d)$ and $\mathbb{T}^d$ is the d-dimensional torus. Assume that

$$a(u,v) \le \lambda \|u\|_{H^\gamma}\|v\|_{H^\gamma} \text{ and } a(u,u) \ge \mu\|u\|_{H^\gamma}^2.$$

Then, $a(\cdot,\cdot)$ generates the so called energy norm equivalent to the norm of H^γ. Consider the problem of finding an element $u \in H^\gamma$ such that

$$a(u,v) \;=\; (f,v) \text{ for all } v \in H^\gamma. \tag{1.3}$$

In order to get an approximate numerical solution we can consider the same problem on a finite dimensional subspace V_h in H^γ

$$a(u_h,v) \;=\; (f,v) \text{ for all } v \in V_h. \tag{1.4}$$

By the Lax-Milgram theorem [25], the problems (1.3) and (1.4) have unique solutions u^* and u_h^*, respectively, which by Céa's lemma [5], satisfy the inequality

$$\|u^* - u_h^*\|_{H^\gamma} \;\le (\lambda/\mu) \inf_{v\in V_h} \|u^* - v\|_{H^\gamma}.$$

Here a naturally arising question is how to choose optimal n-dimensional subspaces V_h and linear finite element approximation algorithms for the problem (1.4). This certainly leads to the problems of optimal linear approximation in H^γ of functions from U and Kolmogorov n-widths $d_n(U,H^\gamma)$, where U is a class of functions u having in some sense more regularity than the class H^γ. The regularity of the class U (in high-dimensional settings) is usually measured by L_2-boundedness of mixed derivatives of higher order or other anisotropic derivatives (a mixed derivative is sometimes referred to as an anisotropic derivative). Finite element approximation spaces based on hyperbolic cross frequency domains are suitable for this framework. It is well-known that the cost of approximately solving Poisson's equation in d dimensions in the Sobolev space H^1 is exponentially growing in d. Standard finite element methods lead to a cost $n_\varepsilon = O(\varepsilon^{-d})$. If we know in advance that the solution belongs to a space of functions with dominating mixed first derivative, and if we use hyperbolic cross spaces for finite element methods, then this requires the cost of $n_\varepsilon \le C(d)\,\varepsilon^{-1}|\log\varepsilon|^{d-1}$. Here and below, $C(d,...)$ are various constants depending on d and other parameters. In [3] it was shown how to get rid of the additional logarithmic term by the use of a subspace of the hyperbolic cross spaces. This results in energy norm based hyperbolic cross spaces and H^1-norm approximation of functions with dominating mixed second derivative. Then the total cost for the solution of Poisson's equation is of the order $n_\varepsilon \le C(d)\,\varepsilon^{-1}$, see also [41] for a generalization. In [21], [22] Griebel and Knapek generalized the construction of [3] to the elliptic variational problem (1.3). By use of tensor-product biorthogonal wavelet bases, they constructed for finite element methods so-called optimized sparse grid subspaces of lower dimension than the standard full-grid spaces. These subspaces preserve the approximation order of the standard full-grid spaces, provided that the solution possesses $H^{\alpha,\beta}$-regularity. To this end, the authors measured the approximation error in the energy H^γ-norm and estimated it from above by terms involving the $H^{\alpha,\beta}$-norm of the solution. The smoothness of spaces $H^{\alpha,\beta}$ is a "hybrid" of isotropic smoothness β and mixed smoothness α [22, Def. 2.1]. It turns out that the necessary dimension n_ε of the optimized sparse grid space for the approximation with accuracy ε does not exceed $C(d,\alpha,\gamma,\beta)\,\varepsilon^{-(\alpha+\beta-\gamma)}$ if $\alpha > \gamma-\beta > 0$. Due to the construction, the optimized sparse grid spaces can be considered as an extension of hyperbolic cross spaces.

The curse of dimensionality is not sufficiently clarified unless "constants" such as $B(\alpha,d)$, $B'(\alpha,d)$ in (1.2) for d_n or $C(d)$ and $C(d,\alpha,\gamma,\beta)$ in the above inequalities for n_ε are not completely determined. We are interested, so far possible, in explicitly determining these constants. The aim of the present paper is to compute $d_n(U,H^\gamma)$ and $n_\varepsilon(U,H^\gamma)$ where U is the unit ball $U^{\alpha,\beta}$ in $H^{\alpha,\beta}$ or its subsets $U^{\alpha,\beta}_*$ and the below characterized class $U^{\alpha,\beta}_\nu$ for $\alpha > \gamma-\beta \ge 0$. The function class $U^{\alpha,\beta}_*$ is the set of all functions $f \in U^{\alpha,\beta}$ such that $\hat{f}(s) = 0$ whenever $\prod_{j=0}^d s_j = 0$. In [21, 22], the authors considered a counterpart of the class $U^{\alpha,\beta}_*$ defined via a biorthogonal wavelet decomposition, see Section 5 in the present paper. They investigated the approximation of functions from this class by optimized sparse grid spaces. We complement their investigations by establishing sharp lower and upper bounds in an explicit form of all relevant components depending on α,β,γ and d,n,ν. This includes the case (1.2) and its modifications when $\alpha > \gamma = \beta = 0$. In contrast to [21, 22] we also obtain lower bounds and prove therefore that trigonometric hyperbolic cross approximations are optimal in terms of Kolmogorov n-widths. For the case $\alpha > \gamma-\beta > 0$, we prove that the hyperbolic cross approximation spaces from [21, 22] are optimal for $d_n(U^{\alpha,\beta}_*,H^\gamma)$. Moreover, the modifications given in the present paper are optimal for $d_n(U^{\alpha,\beta},H^\gamma)$ and $d_n(U^{\alpha,\beta}_\nu,H^\gamma)$. In the case $\alpha > \gamma-\beta = 0$, we prove that classical hyperbolic cross spaces (see, e.g.,

[38]) and their modifications in this paper are optimal for $d_n(U_*^{\alpha,\beta}, H^\gamma)$, $d_n(U^{\alpha,\beta}, H^\gamma)$ and $d_n(U_\nu^{\alpha,\beta}, H^\gamma)$.

It seems that smoothness is not enough for ridding the curse of dimensionality. However, by imposing some additional restrictions on functions in $U^{\alpha,\beta}$ this is possible. In fact, $U_\nu^{\alpha,\beta}$ is the set of all functions $f \in U^{\alpha,\beta}$ actually depending on at most ν (unknown) variables by formally being a d-variate function. For this function class, the curse of dimensionality is broken. For instance, in Theorem 4.7 in Section 4, for the case $\alpha > \gamma - \beta > 0$, we obtain the relations

$$\begin{aligned} \frac{1}{2^{\rho+3\delta}}\nu^\delta \left(1 + \frac{d}{\nu(2^{\rho/\delta}-1)}\right)^{\delta\nu} n^{-\delta} \ &\leq \ d_n(U_\nu^{\alpha,\beta}, H^\gamma) \\ &\leq \ \left(\frac{\alpha}{\delta}\right)^\delta 2^{2\rho+\delta}\nu^\delta \left(1 + \frac{d}{2^{\rho/\delta}-1}\right)^{\delta\nu} n^{-\delta}, \end{aligned}$$

if $n \geq \frac{\alpha}{\delta}\nu 2^{\nu(2\alpha/\delta+1)}(1 + d/(2^{\rho/\delta}-1))^\nu$, where $\delta := \alpha + \beta - \gamma$ and $\rho := \gamma - \beta$. A corresponding result for the ε-dimension n_ε (see Theorem 4.8 in Section 4) states that the number $n_\varepsilon(U_\nu^{\alpha,\beta}, H^\gamma)$ is bounded polynomially in d and ε^{-1} from above. As a consequence, according to [26, (2.3)], we obtain that the problem is polynomially tractable. In addition, the case $\gamma = \beta$, which contains the classical situation with $U_\nu^{\alpha\mathbf{1}}$ instead of $U^{\alpha\mathbf{1}}$ in (1.2), gives as well the polynomial tractability, see Theorems 4.10 and 4.11. Let us mention the relation to the results of Novak and Woźniakowski on weighted tensor product problems with finite order weights [26, 5.3]. Their approach also limits the number ν of active variables in a function via a finite order weight sequence (of order ν). However, since in this paper in most cases neither the spaces $H^{\alpha,\beta}$ of the functions to be approximated, nor the space H^γ, where the approximation error is measured, are tensor products of univariate spaces [35], our results are not included in [26, Theorem 5.8]. Apart from that, totally different approaches for the approximation of functions depending on just a few variables in high dimensions are given in [9], [44].

The paper is organized as follows. In Section 2, we describe a dyadic harmonic decomposition of periodic functions from $H^{\alpha,\beta}$ used for norming these classes suitably for high-dimensional approximations. In Section 3, we prove upper bounds for hyperbolic cross approximations of functions from $U = U^{\alpha,\beta}$, $U_*^{\alpha,\beta}$ and $U_\nu^{\alpha,\beta}$ by linear methods, and for the dimensions of the corresponding approximation spaces. By means of these results, we are able to estimate $d_n(U, H^\gamma)$ and $n_\varepsilon(U, H^\gamma)$ from above. In Section 4, we prove the optimality of these approximations by establishing lower bounds for $d_n(U, H^\gamma)$. In Section 5, we discuss the extension of our results to biorthogonal wavelets and more general decompositions.

2 Dyadic decompositions

Let $\mathbb{N}$ denote the natural numbers, $\mathbb{Z}$ the integers, $\mathbb{Z}_+ = \mathbb{N} \cup \{0\}$ the natural numbers including zero, $\mathbb{R}$ the real numbers, and $\mathbb{C}$ the complex numbers. The number d is always reserved for the number of variables of the functions under consideration. Indeed, we will consider functions on $\mathbb{R}^d$ which are 2π-periodic in each variable, as functions defined on the d-dimensional torus $\mathbb{T}^d := [-\pi, \pi]^d$. Denote by $L_2 := L_2(\mathbb{T}^d)$ the Hilbert space of functions on $\mathbb{T}^d$ equipped with the inner product

$$(f, g) := (2\pi)^{-d} \int_{\mathbb{T}^d} f(x)\overline{g(x)}\, dx.$$

As usual, the norm in L_2 is $\|f\| := (f, f)^{1/2}$. For $s \in \mathbb{Z}^d$, let $\hat{f}(s) := (f, e_s)$ be the sth Fourier coefficient of f, where $e_s(x) := e^{i(s,x)}$.

Let $\mathcal{S}(\mathbb{T}^d)$ be the space of functions on $\mathbb{T}^d$ whose Fourier coefficients form a rapidly decreasing sequence, and $\mathcal{S}'(\mathbb{T}^d)$ the space of distributions which are continuous linear functionals on $\mathcal{S}(\mathbb{T}^d)$. It is well-known that, if $f \in \mathcal{S}'(\mathbb{T}^d)$, then the Fourier coefficients $\hat{f}(s), s \in \mathbb{Z}^d$, of f form a tempered sequence (see, e.g., [37, 40]). A function in L_2 can be considered as an element of $\mathcal{S}'(\mathbb{T}^d)$. For $f \in \mathcal{S}'(\mathbb{T}^d)$, we use the identity

$$f = \sum_{s \in \mathbb{Z}^d} \hat{f}(s) e_s$$

holding in the topology of $\mathcal{S}'(\mathbb{T}^d)$. Denote by $[d]$ the set of natural numbers from 1 to d, and by $\sigma(x) := \{i \in [d] : x_i \neq 0\}$ the support of the vector $x \in \mathbb{R}^d$. For $r \in \mathbb{R}^d$, the rth derivative $f^{(r)}$ of a distribution f is defined as the distribution in $\mathcal{S}'(\mathbb{T}^d)$ given by the identification

$$f^{(r)} := \sum_{s \in \mathbb{Z}^d_0(r)} (is)^r \hat{f}(s) e_s, \tag{2.1}$$

where $(is)^r := \prod_{j=1}^d (is_j)^{r_j}$, $(ia)^b := |a|^b e^{(i\pi b \,\mathrm{sign}\, a)/2}$ for $a, b \in \mathbb{R}$, and $\mathbb{Z}^d_0(r) := \{s \in \mathbb{Z}^d : s_j \neq 0,\ j \in \sigma(r)\}$.

Let us recall the definition of some well known function spaces with isotropic and anisotropic smoothness. The isotropic Sobolev space H^γ, $\gamma \in \mathbb{R}$. For $\gamma \geq 0$, H^γ is the subspace of functions in L_2, equipped with the norm

$$\|f\|^2_{H^\gamma} := \|f\|^2 + \sum_{j=0}^d \|f^{(\gamma \epsilon^j)}\|^2,$$

where $\epsilon^j := (0, ..., 0, 1, 0, ..., 0)$ is the jth unit vector in $\mathbb{R}^d$. For $\gamma < 0$, we define H^γ as the L_2-dual space of $H^{-\gamma}$.

The space H^r of mixed smoothness $r \in \mathbb{R}^d$ is defined as the tensor product of the spaces H^{r_j}, $j \in [d]$:

$$H^r := \bigotimes_{j=1}^d H^{r_j}.$$

where H^{r_j} is the univariate Sobolev space in variable x_j.

For a finite set $A \subset \mathbb{R}^d$, denote by H^A the normed space of all distributions f for which the following norm is finite

$$\|f\|_{H^A}^2 := \sum_{r \in A} \|f\|_{H^r}^2.$$

For $\alpha, \beta \in \mathbb{R}$, let us define the space $H^{\alpha,\beta}$ as follows. If $\beta \geq 0$, we put $H^{\alpha,\beta} := H^A$, where

$$A \ = \ \{(\alpha \mathbf{1} + \beta \epsilon^j) : \ j \in [d]\} \tag{2.2}$$

and $\mathbf{1} := (1, 1, \cdots, 1) \in \mathbb{R}^d$. If $\beta < 0$, we define $H^{\alpha,\beta}$ as the L_2-dual space of $H^{-\alpha,-\beta}$. The space $H^{\alpha,\beta}$ has been introduced in [22]. Notice that $H^{\alpha,0} = H^{\alpha\mathbf{1}}$ and $H^{0,\beta} = H^\beta$.

We will need a dyadic harmonic decomposition of distributions. We define for $k \in \mathbb{Z}_+$,

$$P_k := \ \{s \in \mathbb{Z} : 2^{k-1} \leq |s| < 2^k\}, \ k > 0, \quad P_0 := \ \{0\},$$

and for $k \in \mathbb{Z}_+^d$,

$$P_k := \prod_{j=0}^d P_{k_j}.$$

For distributions f and $k \in \mathbb{Z}_+^d$, let us introduce the following operator:

$$\delta_k(f) := \sum_{s \in P_k} \hat{f}(s) e_s.$$

If $f \in L_2$, we have by Parseval's identity

$$\|f\|^2 \ = \ \sum_{k \in \mathbb{Z}_+^d} \|\delta_k(f)\|^2. \tag{2.3}$$

Moreover, the space L_2 can be decomposed into pairwise orthogonal subspaces $W_k, \ k \in \mathbb{Z}_+^d$, by

$$L_2 \ = \ \bigoplus_{k \in \mathbb{Z}_+^d} W_k,$$

with

$$\dim W_k \ = \ |P_k| \ = \ 2^{|k|_1},$$

where W_k is the space of trigonometric polynomials g of the form

$$g = \sum_{s \in P_k} c_s e_s.$$

and $|Q|$ denotes the cardinality of the set Q.

Put $|k|_1 := \sum_{j=0}^d k_j$ and $|k|_\infty := \max_{1 \leq j \leq d} k_j$ for $k \in \mathbb{Z}_+^d$.

Lemma 2.1 *For any $\alpha, \beta \in \mathbb{R}$, we have the following norm equivalence*

$$\|f\|^2_{H^{\alpha,\beta}} \asymp \sum_{k \in \mathbb{Z}^d_+} 2^{2(\alpha|k|_1 + \beta|k|_\infty)} \|\delta_k(f)\|^2.$$

Proof. We need the following preliminary norms equivalence for $r \in \mathbb{R}^d$,

$$\|f\|^2_{H^r} \asymp \sum_{k \in \mathbb{Z}^d_+} 2^{2(r,k)} \|\delta_k(f)\|^2. \tag{2.4}$$

Indeed, for the univariate case ($d = 1$), by the definition $\|f\|_{H^r}$ is the norm of the isotropic Sobolev space H^γ for $\gamma = r$. Consequently, by (2.3)

$$\|f\|^2_{H^r} \asymp \sum_{k \in \mathbb{Z}_+} \|\delta_k(f)\|^2_{H^\gamma}.$$

Observe that $\|\delta_k(f)\|^2_{H^\gamma} \asymp 2^{2\gamma|k|_1} \|\delta_k(f)\|^2$. This inequality is implied from the definition (2.1) for $\gamma \geq 0$, and from the L_2-duality of H^γ for $\gamma < 0$. Hence, we prove (2.4) for the univariate case. Since in the multivariate case, H^r is the tensor product of isotropic Sobolev spaces it is easy derive (2.4) from the univariate case.

Let us prove the lemma. We first consider the case $\beta \geq 0$. Taking A for the definition of $H^{\alpha,\beta}$ as in (2.2), by (2.4) we get

$$\begin{aligned} \|f\|^2_{H^A} &\asymp \max_{r \in A} \sum_{k \in \mathbb{Z}^d_+} \|f\|^2_{H^r} \\ &\asymp \max_{r \in A} \sum_{k \in \mathbb{Z}^d_+} 2^{2(r,k)} \|\delta_k(f)\|^2 \\ &\leq \sum_{k \in \mathbb{Z}^d_+} 2^{2\max_{r \in A}(r,k)} \|\delta_k(f)\|^2. \end{aligned} \tag{2.5}$$

Let us decompose $\mathbb{Z}^d_+$ into the subsets $\mathbb{Z}^d_+(r)$, $r \in A$, such that

$$\mathbb{Z}^d_+ = \bigcup_{r \in A} \mathbb{Z}^d_+(r), \quad \mathbb{Z}^d_+(r) \cap \mathbb{Z}^d_+(r') = \emptyset, \ r' \neq r,$$

and

$$\max_{r' \in A}(r', k) = (r, k), \ k \in \mathbb{Z}^d_+(r).$$

(Obviously, such a decomposition is easily constructed and some of $\mathbb{Z}^d_+(r)$ may be empty set). Then we have

$$\begin{aligned}\max_{r\in A}\sum_{k\in\mathbb{Z}^d_+}2^{2(r,k)}\|\delta_k(f)\|^2 &= \max_{r\in A}\sum_{r'\in A}\sum_{k\in\mathbb{Z}^d_+(r')}2^{2(r,k)}\|\delta_k(f)\|^2\\ &\geq \sum_{r'\in A}\sum_{k\in\mathbb{Z}^d_+(r')}2^{2(r',k)}\|\delta_k(f)\|^2\\ &= \sum_{k\in\mathbb{Z}^d_+}2^{2\max_{r\in A}(r,k)}\|\delta_k(f)\|^2.\end{aligned}$$

This and (2.5) show that

$$\|f\|^2_{H^A} \asymp \sum_{k\in\mathbb{Z}^d_+}2^{2\max_{r\in A}(r,k)}\|\delta_k(f)\|^2.$$

By a direct computation one can verify that $\max_{r\in A}(r,k)=\alpha|k|_1+\beta|k|_\infty$. This proves the lemma for the case $\beta\geq 0$.

If $\beta<0$, by the definition, the L_2-duality and (2.3)

$$\begin{aligned}\|f\|^2_{H^{\alpha,\beta}} &\asymp \sum_{k\in\mathbb{Z}^d_+}2^{-2(-\alpha|k|_1-\beta|k|_\infty)}\|\delta_k(f)\|^2\\ &= \sum_{k\in\mathbb{Z}^d_+}2^{2(\alpha|k|_1+\beta|k|_\infty)}\|\delta_k(f)\|^2.\end{aligned}$$

□

On the basis of Lemma 2.1, let us redefine the space $H^{\alpha,\beta}$, $\alpha,\beta\in\mathbb{R}$ as the space of distributions f on $\mathbb{T}^d$ for which the following norm is finite

$$\|f\|^2_{H^{\alpha,\beta}} := \sum_{k\in\mathbb{Z}^d_+}2^{2(\alpha|k|_1+\beta|k|_\infty)}\|\delta_k(f)\|^2. \tag{2.6}$$

With this definition we have $H^{0,0}=L_2$. We put $H^{0,\beta}=H^\beta$ and $H^{\alpha,0}=H^{\alpha\mathbf{1}}$ as in the traditional definitions. Denote by $U^{\alpha,\beta}$ the unit ball in $H^{\alpha,\beta}$.

Regarding (2.6) it is worth mentioning the following important thing. In traditional approximation problems where the parameter d is small and fixed, the convergence rates with respect to different equivalent norms only differ by moderate constants. The picture completely changes for high-dimensional approximation problems where we stress the importance of finding an accurate dependence of the convergence rate on the number d of variables and the dimension n of the approximation space. In fact, it essentially depends on the choice of the norm of a function class (i.e., its unit ball) and a norm measuring the approximation error. In some high-dimensional problems it is more convenient to define the function spaces based on a mixed dyadic decomposition in terms of (2.6). The problem itself changes if we use another characterization for $H^{\alpha,\beta}$ and H^γ instead of (2.6). For instance, one can define the following equivalent norm of $H^{\alpha,\beta}$ in terms of the Fourier coefficients by using an ANOVA-type decomposition

$$\|f\|^2_{\tilde{H}^{\alpha,\beta}} := \sum_{s\in\mathbb{Z}^d} |\hat{f}(s)|^2 + \sum_{e\subset[d],\, e\neq\varnothing} \sum_{s\in\mathbb{Z}^{d,e}} \sum_{j\in e} |s_j|^{2(\alpha+\beta)} \left(\prod_{\ell\in e,\ \ell\neq j} |s_\ell|^{2\alpha} \right) |\hat{f}(s)|^2,$$

where $\mathbb{Z}^{d,e} := \{s \in \mathbb{Z}^d : \sigma(s) = e\}$. Note, that $H^{\alpha,\beta}$ and $\tilde{H}^{\alpha,\beta}$ coincide as function spaces. However, if d is large the unit balls with respect to the norms of these spaces differ significantly.

We define the subsets $U^{\alpha,\beta}_*$ and $U^{\alpha,\beta}_\nu$, $1 \leq \nu \leq d-1$, in $U^{\alpha,\beta}$ as follows. $U^{\alpha,\beta}_*$ is the subset in $U^{\alpha,\beta}$ of all f such that

$$\delta_k(f) = 0 \text{ if } \prod_{j=0}^d k_j = 0.$$

The subset $U^{\alpha,\beta}_\nu$ is the set of all $f \in U^{\alpha,\beta}$ such that

$$\delta_k(f) = 0 \text{ if } |\sigma(k)| > \nu.$$

By the definitions we have

$$1 \geq \|f\|^2_{H^{\alpha,\beta}} = \sum_{k\in\mathbb{N}^d} 2^{2(\alpha|k|_1+\beta|k|_\infty)} \|\delta_k(f)\|^2, \quad f \in U^{\alpha,\beta}_*,$$

and

$$1 \geq \|f\|^2_{H^{\alpha,\beta}} = \sum_{k\in\mathbb{Z}^{d,\nu}_+} 2^{2(\alpha|k|_1+\beta|k|_\infty)} \|\delta_k(f)\|^2, \quad f \in U^{\alpha,\beta}_\nu,$$

where $\mathbb{Z}^{d,\nu}_+ := \{k \in \mathbb{Z}^d_+ :\ |\sigma(k)| \leq \nu\}$.

The function class $U^{\alpha,\beta}_*$ can also be seen as the subset in $U^{\alpha,\beta}$ of all f such that $\hat{f}(s) = 0$ whenever $\prod_{j=0}^d s_j = 0$. In case that $H^{\alpha,\beta}$ is a subspace of $L_2(\mathbb{T}^d)$ (recall that it is formally defined as a space of distributions), then every $f \in U^{\alpha,\beta}_*$ has zero mean value in each variable, i.e., we have almost everywhere (in $\mathbb{T}^{d-1}$) the identities

$$\int_{\mathbb{T}} f(x)d\,x_j = 0, \quad j \in [d].$$

The function class $U^{\alpha,\beta}_\nu$ can also be seen as the set of all $f \in U^{\alpha,\beta}$ such that $\hat{f}(s) = 0$ if $|\sigma(s)| > \nu$. It can be interpreted as the set of of all $f \in U^{\alpha,\beta}$ such that f are functions of at most ν variables:

$$f(x) \ = \sum_{e\subset[d]:\,|e|=\nu} f_e(x^e), \quad x^e = (x_j)_{j\in e}.$$

In some high-dimensional problems, objects (functions) only depend on a few variables ν (or represent sums of such objects), where ν is fixed and much smaller than d, the total number of variables. The class $U^{\alpha,\beta}_\nu$ represents a model of such functions.

3 Upper bounds for d_n and n_ε

3.1 Linear trigonometric hyperbolic cross approximations

Let $\alpha, \beta, \gamma \in \mathbb{R}$ be given. For $\xi \geq 0$, we define the subspace in L_2

$$V^d(\xi) := \bigoplus_{k\in J^d(\xi)} W_k,$$

where

$$J^d(\xi) := \ \{k \in \mathbb{Z}^d_+ : \alpha|k|_1 - (\gamma-\beta)|k|_\infty \leq \xi\}.$$

Notice that $\dim V^d(\xi) < \infty$ for all $\xi \geq 0$ if and only if $\alpha - (\gamma - \beta) > 0$. If the last condition is fulfilled, $V^d(\xi)$ is the space of trigonometric polynomials g of the form

$$g = \sum_{k\in J^d(\xi)} \delta_k(g).$$

We define also the subspaces $V^d_*(\xi)$ and $V^d_\nu(\xi)$ in $V^d(\xi)$ by

$$V^d_*(\xi) := \bigoplus_{k\in J^d_*(\xi)} W_k, \quad V^d_\nu(\xi) := \bigoplus_{k\in J^d_\nu(\xi)} W_k,$$

where

$$J^d_*(\xi) := \ \{k \in \ \mathbb{N}^d : \alpha|k|_1 - (\gamma-\beta)|k|_\infty \leq \xi\},$$
$$J^d_\nu(\xi) := \ \{k \in \mathbb{Z}^{d,\nu}_+ : \alpha|k|_1 - (\gamma-\beta)|k|_\infty \leq \xi\}.$$

For a distribution f, we define the linear operator S_ξ as

$$S_\xi(f) := \sum_{k\in J^d(\xi)} \delta_k(f).$$

Obviously, the restriction of S_ξ on L_2 is the orthogonal projection onto $V^d(\xi)$. Put

$$H^d(\xi) := \bigcup_{k \in J^d(\xi)} P_k, \quad H^d_*(\xi) := \bigcup_{k \in J^d_*(\xi)} P_k, \quad H^d_\nu(\xi) := \bigcup_{k \in J^d_\nu(\xi)} P_k.$$

We call the sets $H^d(\xi)$, $H^d_*(\xi)$, $H^d_\nu(\xi)$ (step) hyperbolic cross due to their geometric form. If $\alpha - (\gamma - \beta) > 0$, then $V^d(\xi)$, $V^d_*(\xi)$, $V^d_\nu(\xi)$ are space of trigonometric polynomials with frequencies from $H^d(\xi)$, $H^d_*(\xi)$, $H^d_\nu(\xi)$, respectively. We call them "trigonometric hyperbolic cross spaces", whereas an approximation with respect to these spaces is called "trigonometric hyperbolic cross approximation". In fact, by definition we have

$$S_\xi(f) := \sum_{s \in H^d(\xi)} \hat{f}(s) e_s \,,$$

which represents a trigonometric hyperbolic cross approximation to f.

Before presenting precise approximation results, let us mention an important connection to singular numbers of operators between Hilbert spaces. Let the linear operator $A : H^\gamma \to H^\gamma$ be defined by

$$A(\phi_s) := \ 2^{-(\alpha|k|_1 + (\beta-\gamma)|k|_\infty)} \phi_s, \ s \in P_k, \ k \in \mathbb{Z}^d_+,$$

where the functions $\phi_s := 2^{-\gamma|k|_\infty} e_s$, $s \in P_k$, $k \in \mathbb{Z}^d_+$, are an orthonormal basis in H^γ. Then the Kolmogorov n-widths $d_n(H^{\alpha,\beta}, H^\gamma)$ and linear n-widths $\lambda_n(H^{\alpha,\beta}, H^\gamma)$ coincide with the Kolmogorov numbers of A. Hence, if $\sigma_1(A) \ge \sigma_2(A)... \ge \sigma_j(A) \ge ...$ denote the singular numbers of the operator A, then $d_n(H^{\alpha,\beta}, H^\gamma) = \sigma_{n+1}(A)$ (see, e.g., [26, Theorem 4.11, Corollary 4.12] for details). This reduces the evaluation of $d_n(H^{\alpha,\beta}, H^\gamma)$ to the problem of evaluating the cardinality of the sets $H^d(\xi)$. Similar reductions also hold for $d_n(H^{\alpha,\beta}_*, H^\gamma)$ and $d_n(H^{\alpha,\beta}_\nu, H^\gamma)$. However, in the sequel we want to directly give upper bounds and show that the trigonometric hyperbolic cross spaces $V^d(\xi)$, $V^d_*(\xi)$, $V^d_\nu(\xi)$ are optimal for $d_n(H^{\alpha,\beta}, H^\gamma)$, $d_n(H^{\alpha,\beta}_*, H^\gamma)$, and $d_n(H^{\alpha,\beta}_\nu, H^\gamma)$, respectively.

The following lemma and corollary give upper bounds with regard to ξ for the error of these approximations.

Lemma 3.1 *Let $\alpha, \beta, \gamma \in \mathbb{R}$ be given. Then for arbitrary $\xi \ge 0$,*

$$\|f - S_\xi(f)\|_{H^\gamma} \ \le \ 2^{-\xi} \|f\|_{H^{\alpha,\beta}} \,, \quad f \in H^{\alpha,\beta}.$$

Proof. Indeed, we have for every $f \in H^{\alpha,\beta}$,

$$\begin{aligned} \|f - S_\xi(f)\|^2_{H^\gamma} \ &= \ \sum_{k \notin J^d(\xi)} 2^{\gamma|k|_\infty} \|\delta_k(f)\|^2 \\ &\le \ \sup_{k \notin J^d(\xi)} 2^{-2(\alpha|k|_1 - (\gamma-\beta)|k|_\infty)} \sum_{k \notin J^d(\xi)} 2^{2(\alpha|k|_1 + \beta|k|_\infty)} \|\delta_k(f)\|^2 \\ &\le \ 2^{-2\xi} \|f\|^2_{H^{\alpha,\beta}}. \end{aligned}$$

∎

Corollary 3.2 *Let $\alpha, \beta, \gamma \in \mathbb{R}$ satisfy the condition $\alpha > \gamma - \beta \geq 0$. Then for arbitrary $\xi \geq 0$,*

$$\sup_{f \in U^{\alpha,\beta}} \inf_{g \in V^d(\xi)} \|f - g\|_{H^\gamma} \ \leq \ \sup_{f \in U^{\alpha,\beta}} \|f - S_\xi(f)\|_{H^\gamma} \ \leq \ 2^{-\xi};$$

$$\sup_{f \in U_*^{\alpha,\beta}} \inf_{g \in V_*^d(\xi)} \|f - g\|_{H^\gamma} \ \leq \ \sup_{f \in U_*^{\alpha,\beta}} \|f - S_\xi(f)\|_{H^\gamma} \ \leq \ 2^{-\xi};$$

$$\sup_{f \in U_\nu^{\alpha,\beta}} \inf_{g \in V_\nu^d(\xi)} \|f - g\|_{H^\gamma} \ \leq \ \sup_{f \in U_\nu^{\alpha,\beta}} \|f - S_\xi(f)\|_{H^\gamma} \ \leq \ 2^{-\xi}.$$

In the next two subsections, we establish upper bounds for Kolmogorov n-widths $d_n(U^{\alpha,\beta}, H^\gamma)$, $d_n(U_*^{\alpha,\beta}, H^\gamma)$ and $d_n(U_\nu^{\alpha,\beta}, H^\gamma)$ as well their inverses $n_\varepsilon(U^{\alpha,\beta}, H^\gamma)$, $n_\varepsilon(U_*^{\alpha,\beta}, H^\gamma)$ and $n_\varepsilon(U_\nu^{\alpha,\beta}, H^\gamma)$ on the basis of Lemma 3.1 and Corollary 3.2, and upper bounds of the dimension of the spaces $V^d(\xi)$, $V_*^d(\xi)$ and $V_\nu^d(\xi)$.

3.2 The case $\alpha > \gamma - \beta > 0$

For a given $\theta > 1$, we put $C_\theta := 1$ if $\theta > 2$, and $C_\theta := 1 + \frac{1}{\theta - 1}$ if $1 < \theta \leq 2$. For $\eta \geq 0$, we define

$$I_\eta^d := \ \{k \in \mathbb{N}^d : \theta |k|_1 - |k|_\infty \leq (\theta - 1)\eta + \theta(d-1)\}.$$

For $a \geq 0$, denote by $\lfloor a \rfloor$ the largest integer which is equal or smaller then a, and by $\lceil a \rceil$ the smallest integer which is equal or larger than a. To give an upper estimate of the dimension of the spaces $V^d(\xi)$, $V_*^d(\xi)$ and$V_\nu^d(\xi)$ we need the following lemma.

Lemma 3.3 *Let $\theta > 1$ be a fixed number. Then for any $\eta \geq 0$ the following inequality holds true*

$$\sum_{k \in I_\eta^d} 2^{|k|_1} \ \leq \ C_\theta 2^{1/(\theta-1)} d 2^{d-1} (1 - 2^{-1/(\theta-1)})^{-d} 2^\eta.$$

Proof. Notice that it is enough to prove the lemma for nonnegative integer $\eta = n$. Otherwise, we can treat it for $n = \lfloor \eta \rfloor$. Consider the subsets $\bar{I}_n^d(j), j \in [d]$, in I_n^d defined by

$$\bar{I}_n^d(j) := \{k \in I_n^d : |k|_\infty = k_j\}.$$

Obviously,

$$\sum_{k \in I_n^d} 2^{|k|_1} \ \leq \ \sum_{j=1}^d \ \sum_{k \in \bar{I}_n^d(j)} 2^{|k|_1}.$$

Due to the symmetry, all the sums $\sum_{k \in \bar{I}_n^d(j)} 2^{|k|_1}$, $j \in [d]$, are equal. Thus, in order to prove the lemma it is enough to show for instance, that

$$\sum_{k \in \bar{I}_n^d(d)} 2^{|k|_1} \ \leq \ C_\theta 2^{1/(\theta-1)} 2^{d-1} (1 - 2^{-1/(\theta-1)})^{-d} 2^n. \tag{3.1}$$

Observe that for $k \in \bar{I}_n^d(d)$, $|k|_1$ can take the values $d, ..., n+d-1$. Put $|k|_1 = n+d-1-m$ for $m = 0, 1, ..., n-1$. Fix a nonnegative integer m with $0 \le m \le n-1$. Assume that $|k|_1 = n+d-1-m$. Then clearly, $k \in \bar{I}_n^d(d)$ if and only if $k_d \ge n - \theta m$. It is easy to see that the number of all such $k \in \bar{I}_n^d(d)$, is not larger than

$$\binom{(n+d-1-m) - \lceil n - \theta m \rceil}{d-1} = \binom{d-1+\lceil (\theta-1)m \rceil}{d-1}.$$

Indeed, for the combinatorial identities behind this statement we refer to the proofs of the Lemmas 3.8 and 3.10 below. We obtain

$$\begin{aligned} \sum_{k \in \bar{I}_n^d(d)} 2^{|k|_1} &\le \sum_{m=0}^{n-1} 2^{n+d-1-m} \binom{d-1+\lceil (\theta-1)m \rceil}{d-1} \\ &= 2^{n+d-1} \sum_{m=0}^{n-1} 2^{-m} \binom{d-1+\lceil (\theta-1)m \rceil}{d-1} =: 2^{n+d-1} D(n). \end{aligned} \tag{3.2}$$

Put $\varepsilon := 1/(\theta-1)$ and $N := \lceil (\theta-1)(n-1)) \rceil$. Replacing m by $\tau := m/\varepsilon$ in $D(n)$, we obtain

$$D(n) = \sum_{\tau \in \varepsilon^{-1}\{0,1,...,n-1\}} 2^{-\varepsilon\tau} \binom{d-1+\lceil \tau \rceil}{d-1} \le \sum_{\tau \in \varepsilon^{-1}\{0,1,...,n-1\}} 2^{-\varepsilon(\lceil \tau \rceil - 1)} \binom{d-1+\lceil \tau \rceil}{d-1}.$$

We first consider the case $\theta \ge 2$. For this case, $\varepsilon \le 1$. Since the step length of τ is $1/\varepsilon \ge 1$, we have

$$D(n) \le 2^{\varepsilon} \sum_{s=0}^{N} 2^{-\varepsilon s} \binom{d-1+s}{d-1}. \tag{3.3}$$

Now we consider the case $1 < \theta < 2$. For this case, the step length of τ is $1/\varepsilon < 1$. Notice that then the number of all integers s such that $s = \lceil \tau \rceil$, is not larger than $1 + \varepsilon$. Hence,

$$D(n) \le (1+\varepsilon)\, 2^{\varepsilon} \sum_{s=0}^{N} 2^{-\varepsilon s} \binom{d-1+s}{d-1}.$$

It was proved by Griebel and Knapek [22, p.2242–2243] that

$$\begin{aligned} \sum_{s=0}^{N} 2^{-\varepsilon s} \binom{d-1+s}{d-1} &= (1-t)^{-d} \left[1 - t^{d+N+1} \sum_{s=0}^{d-1} \binom{d+N}{s} \left(\frac{1-t}{t} \right)^s \right] \Bigg|_{t=2^{-\varepsilon}} \\ &\le (1-2^{-\varepsilon})^{-d}. \end{aligned} \tag{3.4}$$

By combining (3.2) – (3.4) we obtain (3.1). $\square$

Remark 3.4 *Lemma 3.3 corrects the last inequality on the bottom of Page 2242 in [22, Lemma 4.2] from which we adapted some proof techniques.*

From now on, for given $\alpha, \beta, \gamma \in \mathbb{R}$, we will frequently use the notations

$$\delta := \alpha - (\gamma - \beta) \quad \text{and} \quad \rho := \gamma - \beta\,. \tag{3.5}$$

Lemma 3.5 *Let $\alpha, \beta, \gamma \in \mathbb{R}$ satisfy the conditions $\alpha > \gamma - \beta > 0$. Then we have*

(i) for any $\xi \geq \alpha(d-1)$,

$$\dim V^d(\xi) \;\leq\; C_{\alpha/\rho} 2^{2\rho/\delta} d(1 + 1/(2^{\rho/\delta} - 1))^d 2^{\xi/\delta},$$

(ii) for any $\xi \geq \alpha(d-1)$,

$$\dim V^d_*(\xi) \;\leq\; C_{\alpha/\rho} 2^{2\rho/\delta} d(2^{\rho/\delta} - 1)^{-d} 2^{\xi/\delta},$$

(iii) for any $\xi \geq \alpha(\nu-1)$,

$$\dim V^d_\nu(\xi) \;\leq\; C_{\alpha/\rho} 2^{2\rho/\delta} \nu(1 + d/(2^{\rho/\delta} - 1))^\nu 2^{\xi/\delta}.$$

Proof. Put $\theta := \alpha/\rho$ and $\eta := (\xi - \alpha(d-1))/\delta$. We have $J^d_*(\xi) = I^d_\eta$. Hence, by Lemma 3.3

$$\begin{aligned}
\dim V^d_*(\xi) \;=\; \sum_{k \in J^d_*(\xi)} 2^{|k|_1} \;&\leq\; \sum_{k \in I^d_\eta} 2^{|k|_1} \\
&\leq\; C_\theta 2^{1/(\theta-1)} d 2^{d-1} (1 - 2^{-1/(\theta-1)})^{-d} 2^\eta \\
&\leq\; C_{\alpha/\rho} 2^{2\rho/\delta} d(2^{\rho/\delta} - 1)^{-d} 2^{\xi/\delta}.
\end{aligned}$$

Inequality (ii) has been proved.

Let us prove the remaining inequalities of the lemma. For a subset $e \in [d]$, put $J^{d,e}(\xi) := \{k \in J^d(\xi) : k_j \neq 0, j \in e,\ k_j = 0, j \notin e\}$. Clearly, $J^{d,e}(\xi) \cap J^{d,e'}(\xi) = \emptyset,\ e \neq e'$, and

$$J^d(\xi) \;=\; \bigcup_{e \subset [d]} J^{d,e}(\xi), \quad J^d_\nu(\xi) \;=\; \bigcup_{|e| \leq \nu} J^{d,e}(\xi),$$

Hence,

$$V^d(\xi) \;=\; \bigoplus_{e \subset [d]} V^{d,e}(\xi), \quad V^d_\nu(\xi) \;=\; \bigoplus_{|e| \leq \nu} V^{d,e}(\xi).$$

where

$$V^{d,e}(\xi) := \left\{ g = \sum_{k \in J^{d,e}(\xi)} \delta_k(g) \right\}.$$

From the last equation and Inequality (ii) of the lemma it follows that

$$\begin{aligned}
\dim V^d(\xi) &= \sum_{e\subset[d]} \dim V^{d,e}(\xi) \\
&= \sum_{k=0}^{d} \sum_{|e|=k} \dim V^{d,e}(\xi) \\
&= \sum_{k=0}^{d} \binom{d}{k} \dim V_*^k(\xi) \\
&\le \sum_{k=0}^{d} \binom{d}{k} C_{\alpha/\rho} 2^{2\rho/\delta} k (2^{\rho/\delta}-1)^{-k} 2^{\xi/\delta} \\
&\le C_{\alpha/\rho} 2^{2\rho/\delta} d 2^{\xi/\delta} \sum_{k=0}^{d} \binom{d}{k} (2^{\rho/\delta}-1)^{-k} \\
&= C_{\alpha/\rho} 2^{2\rho/\delta} d (1+1/(2^{\rho/\delta}-1))^d 2^{\xi/\delta}.
\end{aligned} \tag{3.6}$$

Inequality (iii) can be proved in a similar way. Indeed, it can be shown that

$$\dim V_\nu^d(\xi) = \sum_{k=0}^{\nu} \binom{d}{k} \dim V_*^k(\xi),$$

and hence, applying Inequality (ii) gives

$$\begin{aligned}
\dim V_\nu^d(\xi) &\le \sum_{k=0}^{\nu} \binom{d}{k} C_{\alpha/\rho} 2^{2\rho/\delta} k (2^{\rho/\delta}-1)^{-k} 2^{\xi/\delta} \\
&\le C_{\alpha/\rho} 2^{2\rho/\delta} \nu 2^{\xi/\delta} \sum_{k=0}^{\nu} \binom{\nu}{k} \frac{d!(\nu-k)!}{\nu!(d-k)!} (2^{\rho/\delta}-1)^{-k} \\
&\le C_{\alpha/\rho} 2^{2\rho/\delta} \nu 2^{\xi/\delta} \sum_{k=0}^{\nu} \binom{\nu}{k} d^k (2^{\rho/\delta}-1)^{-k} \\
&= C_{\alpha/\rho} 2^{2\rho/\delta} \nu (1+d/(2^{\rho/\delta}-1))^\nu 2^{\xi/\delta}.
\end{aligned}$$

□

Theorem 3.6 *Let $\alpha, \beta, \gamma \in \mathbb{R}$ satisfy the conditions $0 < \rho = \gamma - \beta < \alpha$. Then, we have*

(i) for any integer $n \ge C_{\alpha/\rho} 2^{\rho/\delta} d 2^{\alpha d/\delta} (1+1/(2^{\rho/\delta}-1))^d$,

$$d_n(U^{\alpha,\beta}, H^\gamma) \le C_{\alpha/\rho}^\delta 2^{2\rho+\delta} d^\delta \left(1+1/(2^{\rho/\delta}-1)\right)^{\delta d} n^{-\delta},$$

(ii) for any integer $n \geq C_{\alpha/\rho} 2^{\rho/\delta} d 2^{\alpha d/\delta}(2^{\rho/\delta}-1)^{-d}$,

$$d_n(U_*^{\alpha,\beta}, H^\gamma) \ \leq \ C_{\alpha/\rho}^\delta 2^{2\rho+\delta} d^\delta \left(2^{\rho/\delta}-1\right)^{-\delta d} n^{-\delta},$$

(iii) for any integer $n \geq C_{\alpha/\rho} 2^{\rho/\delta} \nu 2^{\alpha\nu/\delta}(1+d/(2^{\rho/\delta}-1))^{\nu}$,

$$d_n(U_\nu^{\alpha,\beta}, H^\gamma) \ \leq \ C_{\alpha/\rho}^\delta 2^{2\rho+\delta} \nu^\delta \left(1+d/(2^{\rho/\delta}-1)\right)^{\delta\nu} n^{-\delta}.$$

Proof. We prove the upper bound in Inequality (i) for $d_n(U^{\alpha,\beta}, H^\gamma)$. The other upper bounds can be proved in a similar way.

Put $\varphi(\xi) := \dim V^d(\xi)$. Then φ is a step function in the variable ξ. Moreover, there are sequences $\{\xi_m\}_{m=1}^\infty$ and $\{\eta_m\}_{m=1}^\infty$ such that

$$\varphi(\xi) \ = \ \eta_m, \quad \xi_m \leq \xi < \xi_{m+1}. \tag{3.7}$$

Notice that

$$\xi_{m+1} - \xi_m \ \leq \ \delta. \tag{3.8}$$

Indeed, let

$$\xi_m = \alpha|k|_1 - \rho|k|_\infty$$

for some $k \in J^d(\xi)$. Without loss of generality we can assume that $|k|_\infty = k_d$. Define $k' \in \mathbb{Z}_+^d$ by $k'_d = k_d + 1$ and $k'_j = k_j, \ j \neq d$. Then we have

$$\begin{aligned} \xi_{m+1} - \xi_m \ &\leq \ \alpha|k'|_1 - \rho|k'|_\infty - (\alpha|k|_1 - \rho|k|_\infty) \\ &= \ \alpha(|k'|_1 - |k|_1) - \rho(|k'|_\infty - |k|_\infty) \\ &= \ \alpha - \rho \ = \ \delta. \end{aligned}$$

For a given n satisfying the condition for Inequality (i) of the theorem, let m be the number such that,

$$\dim V^d(\xi_m) \ \leq \ n \ < \ \dim V^d(\xi_{m+1}). \tag{3.9}$$

Hence, by the corresponding restriction on n in the theorem it follows that $\xi_{m+1} \geq \alpha(d-1)$. Putting $\xi := \xi_m$ we obtain by Lemma 3.5 and (3.8)

$$\begin{aligned} n \ &\leq \ C_{\alpha/\rho} 2^{2\rho/\delta} d(1+1/(2^{\rho/\delta}-1))^d 2^{\xi_{m+1}/\delta} \\ &\leq \ C_{\alpha/\rho} 2^{2\rho/\delta+1} d(1+1/(2^{\rho/\delta}-1))^d 2^{\xi/\delta}, \end{aligned}$$

or, equivalently,

$$2^{-\xi} \ \leq \ C_{\alpha/\rho}^\delta 2^{2\rho+\delta} d^\delta \left(1+1/(2^{\rho/\delta}-1)\right)^{\delta d} n^{-\delta}. \tag{3.10}$$

On the other hand, by the definitions, (3.9) and Corollary 3.2,

$$d_n(U_\nu^{\alpha,\beta}, H^\gamma) \ \leq \ \sup_{f \in U^{\alpha,\beta}} \|f - S_\xi(f)\|_{H^\gamma} \ \leq \ 2^{-\xi}.$$

The last relations combined with (3.10) prove the desired inequality. $\square$

Theorem 3.7 *Let $\alpha, \beta, \gamma \in \mathbb{R}$ satisfy the conditions $0 < \gamma - \beta < \alpha$. Then we have*

(i) for any $0 < \varepsilon \leq 1$,

$$n_\varepsilon(U^{\alpha,\beta}, H^\gamma) \ \leq \ C_{\alpha/\rho} 2^{2\rho/\delta} d(1 + 1/(2^{\rho/\delta} - 1))^d \varepsilon^{-1/\delta},$$

(ii) for any $0 < \varepsilon \leq 2^{-\alpha(d-1)}$,

$$n_\varepsilon(U_*^{\alpha,\beta}, H^\gamma) \ \leq \ C_{\alpha/\rho} 2^{2\rho/\delta} d(2^{\rho/\delta} - 1)^{-d} \varepsilon^{-1/\delta},$$

(iii) for any $0 < \varepsilon \leq 2^{-\alpha(\nu-1)}$,

$$n_\varepsilon(U_\nu^{\alpha,\beta}, H^\gamma) \ \leq \ C_{\alpha/\rho} 2^{2\rho/\delta} \nu(1 + d/(2^{\rho/\delta} - 1))^\nu \varepsilon^{-1/\delta}.$$

Proof. The inequalities (i)–(iii) in the theorem can be proved in the same way. Let us prove for instance (i). For a given $0 < \varepsilon \leq 2^{-\alpha d}$, putting $\xi := |\log \varepsilon|$, we get by the definitions and Corollary 3.2,

$$\sup_{f \in U^{\alpha,\beta}} \inf_{g \in V^d(\xi)} \|f - g\|_{H^\gamma} \ \leq \ \sup_{f \in U^{\alpha,\beta}} \|f - S_\xi(f)\|_{H^\gamma} \ \leq \ 2^{-\xi} \ \leq \ \varepsilon.$$

Consequently, Lemma 3.5(i) yields

$$\begin{aligned} n_\varepsilon(U^{\alpha,\beta}, H^\gamma) \ &\leq \ \dim V^d(\xi) \\ &\leq \ C_{\alpha/\rho} 2^{2\rho/\delta} d(1 + 1/(2^{\rho/\delta} - 1))^d 2^{\xi/\delta} \\ &\leq \ C_{\alpha/\rho} 2^{2\rho/\delta} d(1 + 1/(2^{\rho/\delta} - 1))^d \varepsilon^{-1/\delta}. \end{aligned}$$

□

3.3 The case $\alpha > \gamma - \beta = 0$

For $m \in \mathbb{N}$, we define

$$K_*^d(m) := \ \{k \in \mathbb{N}^d : |k|_1 \leq m\}.$$

The following estimates have already been used in [43, Lemma 7]. For convenience of the reader we will give a prove.

Lemma 3.8 *For any $m \geq d$, there hold true the inequalities*

$$2^m \binom{m-1}{d-1} \ < \ \dim V_*^d(\alpha m) \ = \ \sum_{k \in K_*^d(m)} 2^{|k|_1} \ \leq \ 2^{m+1} \binom{m-1}{d-1}.$$

Proof. Observe that for $k \in K_*^d(m)$, $|k|_1$ can take the values $d, ..., m$. It is easy to check that the number of all such $k \in K_*^d(m)$ that $|k|_1 = j$, is

$$\binom{j-1}{d-1}.$$

Hence,

$$\begin{aligned}\sum_{k \in K_*^d(m)} 2^{|k|_1} &= \sum_{j=d}^{m} \binom{j-1}{d-1} 2^j \\ &\le \binom{m-1}{d-1} \sum_{j=0}^{m} 2^j \le 2^{m+1} \binom{m-1}{d-1},\end{aligned}$$

and,

$$\sum_{k \in K_*^d(m)} 2^{|k|_1} = \sum_{j=d}^{m} \binom{j-1}{d-1} 2^j > 2^m \binom{m-1}{d-1}.$$

$\square$

We will use several times the following well-known inequalities for any nonnegative integers n, m with $n \le m$

$$\left(\frac{m}{n}\right)^n \le \binom{m}{n} \le \left(\frac{em}{n}\right)^n. \tag{3.11}$$

Remark 3.9 *From Lemma 3.8 together with the relations* (3.11) *we have*

$$2^m \left(\frac{m-1}{d-1}\right)^{d-1} < \sum_{k \in K_*^d(m)} 2^{|k|_1} \le 2^{m+1} \left(\frac{e(m-1)}{d-1}\right)^{d-1}.$$

This, in particular, sharpens and improves Lemma 3.6 in [4].

For $m \in \mathbb{Z}_+$, we define

$$K^d(m) := \{k \in \mathbb{Z}_+^d : |k|_1 \le m\}.$$

Lemma 3.10 *For any $d \in \mathbb{N}$ and $m \in \mathbb{Z}_+$, there holds true the inequality*

$$2^m \binom{m+d-1}{d-1} < \dim V^d(\alpha m) = \sum_{k \in K^d(m)} 2^{|k|_1} \le 2^{m+1} \binom{m+d-1}{d-1}.$$

Proof. Let $G(d, j)$, $j \in \mathbb{Z}_+$, be the number of all $k \in K^d(n)$ such that $|k|_1 = j$. Observe that $G(d, j)$ coincides with the number of all $k \in \mathbb{N}^d$ such that $|k|_1 = j + d$, and consequently,

$$G(d, j) = \binom{j+d-1}{d-1}.$$

Hence,

$$\begin{aligned}\sum_{k\in K^d(m)} 2^{|k|_1} &= \sum_{j=0}^{m}\binom{j+d-1}{d-1}2^j \\ &\le \binom{m+d-1}{d-1}\sum_{j=0}^{m}2^j \;\le\; 2^{m+1}\binom{m+d-1}{d-1},\end{aligned}$$

and,

$$\sum_{k\in K^d(m)} 2^{|k|_1} = \sum_{j=0}^{m}\binom{j+d-1}{d-1}2^j > 2^m\binom{m+d-1}{d-1}.$$

∎

Let us define the index set $K_\nu^d(m)$ given by

$$K_\nu^d(m) := \{k\in\mathbb{Z}_+^{d,\nu} : |k|_1 \le m\}$$

for some $1\le\nu\le d$ and $m\in\mathbb{Z}_+$.

Lemma 3.11 *Let $\nu, d\in\mathbb{N}$, $m\in\mathbb{Z}$ and $m, d\ge\nu$. Then*

$$2^m\binom{d}{\nu}\binom{m-1}{\nu-1} < \dim V_\nu^d(\alpha m) = \sum_{k\in K_\nu^d(m)} 2^{|k|_1} \le 2^{m+1}\sum_{j=1}^{\nu}\binom{d}{j}\binom{m-1}{j-1}. \tag{3.12}$$

Moreover, if $a>0$ is a fixed number and $b := \frac{a}{a+\sqrt{a+1}}$, then for $\nu, d, m\in\mathbb{N}$, such that $\nu\le b\min(d,m)$, we have

$$\dim V_\nu^d(\alpha m) = \sum_{k\in K_\nu^d(m)} 2^{|k|_1} \le (1+a)2^{m+1}\binom{d}{\nu}\binom{m-1}{\nu-1}. \tag{3.13}$$

Proof. Put

$$K^{d,e}(m) := \{k\in K^d(m) : k_j\ne 0, j\in e,\ k_j = 0, j\notin e\}$$

for a subset $e\subset[d]$. Clearly, we have that

$$\begin{aligned}\sum_{k\in K_\nu^d(m)} 2^{|k|_1} &= \sum_{i=0}^{m}2^i\sum_{j=1}^{\nu}\sum_{\substack{e\subset[d]\\|e|=j}}|K^{d,e}(i)| \\ &= \sum_{j=1}^{\nu}\binom{d}{j}\sum_{i=j}^{m}2^i\binom{i-1}{j-1} \\ &\le \sum_{j=1}^{\nu}\binom{d}{j}\binom{m-1}{j-1}\sum_{i=j}^{m}2^i \\ &\le 2^{m+1}\sum_{j=1}^{\nu}\binom{d}{j}\binom{m-1}{j-1}.\end{aligned} \tag{3.14}$$

The lower bound in (3.12) follows from the second line in (3.14). Next, let us prove the inequality (3.13) by induction on ν. It is trivial for $\nu = 1$. Suppose that it is true for $\nu - 1 \geq 0$. Put

$$S(\nu) := \sum_{j=1}^{\nu} \binom{d}{j} \binom{m-1}{j-1}.$$

We have by the induction assumption

$$\begin{aligned} S(\nu) &= S(\nu-1) + \binom{d}{\nu}\binom{m-1}{\nu-1} \\ &\leq (1+a)\binom{d}{\nu-1}\binom{m-1}{\nu-2} + \binom{d}{\nu}\binom{m-1}{\nu-1} \\ &\leq \frac{(1+a)\nu(\nu-1)}{(d+1-\nu)(m+1-\nu)}\binom{d}{\nu}\binom{m-1}{\nu-1} + \binom{d}{\nu}\binom{m-1}{\nu-1}. \end{aligned} \tag{3.15}$$

By the inequality $\nu \leq b\min(d,m)$, one can immediately verify that $\frac{\nu\sqrt{a+1}}{d+1-\nu} \leq \sqrt{a}$ and $\frac{(\nu-1)\sqrt{a+1}}{m+1-\nu} \leq \sqrt{a}$. Hence, by (3.15) we prove (3.13). □

Remark 3.12 *For a practical application, if we take $\nu \leq \min(d/2, m/2)$, then from Lemma 3.11 we have*

$$2^m\binom{d}{\nu}\binom{m-1}{\nu-1} \leq \dim V_\nu^d(\alpha m) \;\leq\; (\sqrt{5}+3)2^m\binom{d}{\nu}\binom{m-1}{\nu-1}. \tag{3.16}$$

Theorem 3.13 *Let $\alpha, \beta, \gamma \in \mathbb{R}$ satisfy the conditions $\alpha > \gamma - \beta = 0$. Then the following relations hold true.*

(i) For any $n \in \mathbb{N}$,

$$d_n(U^{\alpha,\beta}, H^\gamma) \;\leq\; 4^\alpha \left(\frac{d-1}{e}\right)^{-\alpha(d-1)} n^{-\alpha}(d+\log n)^{\alpha(d-1)},$$

and for any $n \geq 2^d$,

$$d_n(U^{\alpha,\beta}, H^\gamma) \;\leq\; 4^\alpha \left(\frac{d-1}{2e}\right)^{-\alpha(d-1)} n^{-\alpha}(\log n)^{\alpha(d-1)}.$$

(ii) For any integer $n \geq 2^d$,

$$d_n(U_*^{\alpha,\beta}, H^\gamma) \;\leq\; 4^\alpha \left(\frac{d-1}{e}\right)^{-\alpha(d-1)} n^{-\alpha}(\log n)^{\alpha(d-1)}.$$

(iii) If in addition $\nu \le d/2$, then for any $n \ge \frac{\sqrt{5}+3}{2}\binom{d}{\nu}\binom{2\nu-1}{\nu-1}2^{2\nu+1}$,

$$d_n(U_\nu^{\alpha,\beta}, H^\gamma) \;\le\; [2(\sqrt{5}+3)]^\alpha \Big(\frac{\nu-1}{e}\Big)^{-\alpha(\nu-1)} \Big(\frac{\nu}{e}\Big)^{-\alpha\nu} d^{\alpha\nu} n^{-\alpha} (\log n)^{\alpha(\nu-1)}.$$

Proof. We prove the inequality for $d_n(U_*^{\alpha,\beta}, H^\gamma)$ in Relation (ii). The other inequalities in Relations (i) and (iii) can be proved in a similar way. For a given $n \ge 2^d$, by Lemma 3.8 there is a unique $m \ge d$ such that,

$$\dim V_*^d(\alpha m) \;\le\; n \;<\; \dim V_*^d(\alpha(m+1)). \tag{3.17}$$

Again, from Lemma 3.8 we get

$$2^m \binom{m-1}{d-1} \;\le\; \sum_{k \in K_*^d(m)} 2^{|k|_1} \;=\; \dim V_*^d(\alpha m) \;\le\; n$$

and

$$n \;<\; \dim V_*^d(\alpha(m+1)) \;=\; \sum_{k \in K_*^d(m+1)} 2^{|k|_1} \;\le\; 2^{m+2}\binom{m}{d-1}.$$

Hence, by (3.11) we obtain

$$2^m \;=\; 2^{m+2}\Big(\frac{em}{d-1}\Big)^{d-1} \frac{1}{4}\Big(\frac{em}{d-1}\Big)^{-(d-1)} \;\ge\; \frac{1}{4} n \Big(\frac{em}{d-1}\Big)^{-(d-1)}.$$

From the last inequalities we derive

$$\begin{aligned} 2^{-\alpha m} \;&\le\; 4^\alpha \left([(d-1)/e]^{-\alpha(d-1)}\right) n^{-\alpha} m^{\alpha(d-1)} \\ &\le\; 4^\alpha [(d-1)/e]^{-\alpha(d-1)} n^{-\alpha} (\log n)^{\alpha(d-1)}. \end{aligned} \tag{3.18}$$

On the other hand, by the definitions, (3.17) and Corollary 3.2,

$$d_n(U_*^{\alpha,\beta}, H^\gamma) \;\le\; \sup_{f \in U_*^{\alpha,\beta}} \|f - S_{\alpha m}(f)\|_{H^\gamma} \;\le\; 2^{-\alpha m}.$$

This combined with (3.18) proves the desired inequality. □

Theorem 3.14 *Let $\alpha, \beta, \gamma \in \mathbb{R}$ satisfy the conditions $\alpha > \gamma - \beta = 0$. Then the following relations hold true.*

(i) For any $0 < \varepsilon \le 1$,

$$n_\varepsilon(U^{\alpha,\beta}, H^\gamma) \;\le\; 4\Big(\frac{d-1}{e}\Big)^{-(d-1)} [\alpha^{-1}|\log \varepsilon| + d]^{d-1} \varepsilon^{-1/\alpha},$$

and for $0 < \varepsilon \le 2^{-\alpha d}$,

$$n_\varepsilon(U^{\alpha,\beta}, H^\gamma) \;\le\; 4\Big(\frac{\alpha(d-1)}{2e}\Big)^{-(d-1)} \varepsilon^{-1/\alpha} |\log \varepsilon|^{d-1}.$$

(ii) For any $0 < \varepsilon \le 2^{-\alpha d}$

$$n_\varepsilon(U_*^{\alpha,\beta}, H^\gamma) \ \le \ 4\Big(\frac{\alpha(d-1)}{e}\Big)^{-(d-1)} \varepsilon^{-1/\alpha} |\log \varepsilon|^{d-1}.$$

(iii) If in addition $\nu \le d/2$ *then for any* $0 < \varepsilon \le 2^{-2\alpha\nu}$

$$n_\varepsilon(U_\nu^{\alpha,\beta}, H^\gamma) \ \le \ 2(\sqrt{5}+3)\Big(\frac{\alpha(\nu-1)}{e}\Big)^{-(\nu-1)} (\nu/e)^{-\nu} d^\nu \varepsilon^{-1/\alpha} |\log \varepsilon|^{\nu-1}.$$

Proof. Let us prove (ii). The other assertions can be proved in a similar way. For a given $\varepsilon \le 2^{-\alpha d}$ we take $m > d$ such that

$$2^{-\alpha m} \le \ \varepsilon \ < \ 2^{-\alpha(m-1)}.$$

The right inequality gives

$$2^m \le \ 2\varepsilon^{-1/\alpha} \quad \text{and} \quad m \ \le \ \alpha^{-1}|\log \varepsilon| + 1.$$

On the other hand, by the definitions, (3.17) and Corollary 3.2,

$$\sup_{f \in U_*^{\alpha,\beta}} \ \inf_{g \in V_*^d(\alpha m)} \|f - g\|_{H^\gamma} \ \le \ \sup_{f \in U_*^{\alpha,\beta}} \|f - S_{\alpha m}(f)\|_{H^\gamma} \ \le \ 2^{-\alpha m} \ \le \ \varepsilon.$$

Consequently, by Lemma 3.8

$$\begin{aligned} n_\varepsilon(U_*^{\alpha,\beta}, H^\gamma) \ &\le \ 2^{m+1}\binom{m-1}{d-1} \\ &\le \ 4\varepsilon^{-1/\alpha}\binom{\lfloor \alpha^{-1}|\log \varepsilon| \rfloor}{d-1} \\ &\le \ 4\Big(\frac{\alpha(d-1)}{e}\Big)^{-(d-1)} \varepsilon^{-1/\alpha} |\log \varepsilon|^{d-1}. \end{aligned}$$

∎

4 Optimality and lower bounds for for d_n and n_ε

In this section, we give lower bounds for Kolmogorov n-widths $d_n(U^{\alpha,\beta}, H^\gamma)$, $d_n(U_*^{\alpha,\beta}, H^\gamma)$ and $d_n(U_\nu^{\alpha,\beta}, H^\gamma)$ as well their inverses $n_\varepsilon(U^{\alpha,\beta}, H^\gamma)$, $n_\varepsilon(U_*^{\alpha,\beta}, H^\gamma)$ and $n_\varepsilon(U_\nu^{\alpha,\beta}, H^\gamma)$ by applying an abstract result on Kolmogorov n-widths of the unit ball, Bernstein type inequalities, and lower bounds of the dimension of the spaces $V^d(\xi)$, $V_*^d(\xi)$ and $V_\nu^d(\xi)$. We show that the trigonometric hyperbolic cross spaces $V^d(\xi)$, $V_*^d(\xi)$, $V_\nu^d(\xi)$ optimal for optimal $d_n(H^{\alpha,\beta}, H^\gamma)$, $d_n(H_*^{\alpha,\beta}, H^\gamma)$, $d_n(H_\nu^{\alpha,\beta}, H^\gamma)$, respectively. We place the upper bounds of these quantities next to their lower bounds to show the optimality of the linear trigonometric hyperbolic cross approximations with respect to $V^d(\xi)$, $V_*^d(\xi)$ and $V_\nu^d(\xi)$ in the high-dimensional setting.

4.1 Some preparation

The following lemma on Kolmogorov n-widths of the unit ball has been proved in [39, Theorem 1].

Lemma 4.1 *Let L_{n+1} be an $n+1$-dimensional subspace in a Banach space X, and $B_{n+1}(r) := \{f \in L_{n+1} : \ \|f\|_X \le r\}$. Then*

$$d_n(B_{n+1}(r), X) \ = \ r.$$

Next, we prove a Bernstein type inequality.

Lemma 4.2 *Let $\alpha, \beta, \gamma \in \mathbb{R}$ be given. Then for arbitrary $\xi \ge 0$,*

$$\|f\|_{H^{\alpha,\beta}} \ \le \ 2^{\xi} \|f\|_{H^{\gamma}}, \quad f \in V^d(\xi).$$

Proof. Indeed, we have for every $f \in V^d(\xi)$,

$$\begin{aligned}
\|f\|^2_{H^{\alpha,\beta}} \ &= \ \sum_{k \in J^d(\xi)} 2^{2(\alpha|k|_1 + \beta|k|_\infty)} \|\delta_k(f)\|^2 \\
&\le \ \sup_{k \in J^d(\xi)} 2^{2(\alpha|k|_1 - (\gamma-\beta)|k|_\infty)} \sum_{k \in J^d(\xi)} 2^{2\gamma|k|_\infty} \|\delta_k(f)\|^2 \\
&\le \ 2^{2\xi} \|f\|^2_{H^\gamma}.
\end{aligned}$$

□

4.2 The case $\alpha > \gamma - \beta > 0$

Lemma 4.3 *Let $0 < t \le 1/2$ and k, n be integers such that $0 \le k \le n/2$. Then*

$$t^n \sum_{s=0}^{k} \binom{n}{s} \left(\frac{1-t}{t}\right)^s \ \le \ \frac{1}{2}.$$

Proof. Since $(1-t)/t \ge 1$, $\binom{n}{s} = \binom{n}{n-s}$ and $0 \le k \le n/2$, we have

$$\binom{n}{s} \left(\frac{1-t}{t}\right)^s \ \le \ \binom{n}{n-s} \left(\frac{1-t}{t}\right)^s, \ s = 0, ..., k.$$

Hence,

$$\sum_{s=0}^{k} \binom{n}{s} \left(\frac{1-t}{t}\right)^s \ \le \ \sum_{s=0}^{k} \binom{n}{n-s} \left(\frac{1-t}{t}\right)^s,$$

and consequently,

$$t^n \sum_{s=0}^{k} \binom{n}{s} \left(\frac{1-t}{t}\right)^s \ \le \ \frac{1}{2} t^n \sum_{s=0}^{n} \binom{n}{s} \left(\frac{1-t}{t}\right)^s \ = \ \frac{1}{2}.$$

$\square$

Lemma 4.4 *Let $1 < \theta \le 2$. Then for any natural numbers d and n satisfying the condition*

$$d \ \le \ \frac{\theta-1}{2\theta-1} n + \frac{2}{\theta}, \tag{4.1}$$

there holds true the inequality

$$\sum_{k \in I_n^d} 2^{|k|_1} \ \ge \ 2^{-1/(\theta-1)} d 2^{d-2} (1 - 2^{-1/(\theta-1)})^{-d} 2^n.$$

Proof. Consider the subsets $I_n^d(j), j \in [d]$, in I_n^d defined by

$$I_n^d(j) := \left\{ k \in I_n^d : |k|_\infty = k_j, \ |k|_1 \ge \frac{2(\theta-1)}{2\theta-1} n + d - 1 + 2/\theta \right\}.$$

We prove that $I_n^d(j) \cap I_n^d(j') = \emptyset$ for $j \ne j'$. Fix $j \in [d]$ and let k be an arbitrary element in $I_n^d(j)$. Then by the definitions we have

$$\begin{aligned} k_j \ &\ge \ \theta |k|_1 - (\theta-1)n - \theta(d-1) \\ &\ge \ \frac{2\theta(\theta-1)}{2\theta-1} n + \theta(d-1) + 2 - (\theta-1)n - \theta(d-1) \\ &= \ \frac{\theta-1}{2\theta-1} n + 2. \end{aligned} \tag{4.2}$$

On the other hand,

$$\theta(|k|_1 - k_j) + (\theta-1)k_j \ = \ \theta|k|_1 - |k|_\infty \ \le \ (\theta-1)n + \theta(d-1).$$

Hence,

$$\begin{aligned} |k|_1 - k_j \ &\le \ \theta^{-1}[(\theta-1)n + \theta(d-1)] - \theta^{-1}(\theta-1)\left[\frac{\theta-1}{2\theta-1} n + 2\right] \\ &= \ \frac{\theta-1}{2\theta-1} n + d - 3 + 2/\theta. \end{aligned}$$

Take an arbitrary $j' \in [d]$ such that $j \ne j'$. Then, since $k_i \ge 1$ for any $i \in [d]$, and $\theta > 1$, from the last inequality and (4.2) we get

$$\begin{aligned} k_{j'} \ &\le \ |k|_1 - k_j - (d-2) \\ &\le \ \frac{\theta-1}{2\theta-1} n + d - 3 + 2/\theta - (d-2) \\ &= \ \frac{\theta-1}{2\theta-1} n + 2/\theta - 1 \\ &< \ k_j. \end{aligned}$$

This proves that $I^d_n(j) \cap I^d_n(j') = \emptyset$ for $j \neq j'$. Therefore, there holds true the inequality

$$\sum_{k \in I^d_n} 2^{|k|_1} \;\geq\; \sum_{j=1}^{d} \sum_{k \in I^d_n(j)} 2^{|k|_1}.$$

Due to the symmetry, all the sums $\sum_{k \in I^d_n(j)} 2^{|k|_1}$, $j \in [d]$ are equal. Thus, in order to prove the lemma it is enough to show for instance, that

$$\sum_{k \in I^d_n(d)} 2^{|k|_1} \;\geq\; 2^{-1/(\theta-1)} 2^{d-2} (1 - 2^{-1/(\theta-1)})^{-d} 2^n. \tag{4.3}$$

Observe that for $k \in I^d_n(d)$, $|k|_1$ can take the values $\lceil (2(\theta-1)/(2\theta-1))n + d - 1 + 2/\theta \rceil, ..., n+d-1$. Put $|k|_1 = n+d-1-m$ for $m = 0, 1, ..., M$, where $M := n+d-1-\lceil (2(\theta-1)/(2\theta-1))n + d - 1 + 2/\theta \rceil$. Fix a nonnegative integer m with $0 \leq m \leq M$. Assume that $|k|_1 = n+d-1-m$. Then clearly, $k \in I^d_n(d)$ if and only if $k_d \geq n - \theta m$. It is easy to see that the number of all such $k \in I^d_n(d)$ is not smaller than

$$\binom{(n+d-1-m) - \lfloor n - \theta m \rfloor}{d-1} \;=\; \binom{d-1+\lfloor (\theta-1)m \rfloor}{d-1}$$

We have

$$\begin{aligned} \sum_{k \in I^d_n(d)} 2^{|k|_1} \;&\geq\; \sum_{m=0}^{M} 2^{n+d-1-m} \binom{d-1+\lfloor (\theta-1)m \rfloor}{d-1} \\ &=\; 2^{n+d-1} \sum_{m=0}^{M} 2^{-m} \binom{d-1+\lfloor (\theta-1)m \rfloor}{d-1} \;=:\; 2^{n+d-1} A(n). \end{aligned} \tag{4.4}$$

Put $\varepsilon := (\theta-1)^{-1}$ and $N := \lfloor (\theta-1)M) \rfloor$. Replacing m by $\tau := m/\varepsilon$ in $A(n)$, we obtain

$$A(n) \;=\; \sum_{\tau \in \varepsilon^{-1}\{0,1,...,M\}} 2^{-\varepsilon\tau} \binom{d-1+\lfloor \tau \rfloor}{d-1} \;\geq\; \sum_{\tau \in \varepsilon^{-1}\{0,1,...,M\}} 2^{-\varepsilon(\lfloor \tau \rfloor + 1)} \binom{d-1+\lfloor \tau \rfloor}{d-1}.$$

Since $1 < \theta \leq 2$, the step length of τ is $1/\varepsilon \leq 1$. Therefore, we have

$$A(n) \;\geq\; 2^{-\varepsilon} \sum_{s=0}^{N} 2^{-\varepsilon s} \binom{d-1+s}{d-1}. \tag{4.5}$$

By (3.4) we have

$$B(n) := \sum_{s=0}^{N} 2^{-\varepsilon s} \binom{d-1+s}{d-1} \;=\; (1-t)^{-d} \left[1 - t^{d+N+1} \sum_{s=0}^{d-1} \binom{N+d}{s} \left(\frac{1-t}{t} \right)^s \right] \Bigg|_{t=2^{-\varepsilon}}.$$

By the assumptions of the lemma $0 < 2^{-\varepsilon} \le 1/2$ and $d-1 \le (d+N+1)/2$. Applying Lemma 4.3 gives

$$B(n) \;\ge\; \frac{1}{2}(1-t)^{-d}|_{t=2^{-\varepsilon}} \;=\; \frac{1}{2}(1-2^{-1/(\theta-1)})^{-d}. \tag{4.6}$$

Combining (4.4) – (4.6) proves (4.3). $\square$

Remark 4.5 *By using weaker assumptions we can prove the following slightly worse lower bound compared to Lemma 4.4. If $1 < \theta \le 2$, then for any natural numbers d and n, there holds true the inequality*

$$\sum_{k \in I_n^d} 2^{|k|_1} \;\ge\; 2^{-1/(\theta-1)} 2^{d-2} (1-2^{-1/(\theta-1)})^{-d} 2^n.$$

Lemma 4.6 *Let $\alpha, \beta, \gamma \in \mathbb{R}$ satisfy the conditions $2(\gamma-\beta) \ge \alpha > \gamma - \beta > 0$ and $1 \le \nu \le d-1$. Then we have*

(i) for any $\xi \ge (2\alpha+\delta)(d-1)$,

$$\dim V^d(\xi) \;\ge\; \frac{1}{4} 2^{-\rho/\delta} d \Big[1 + \frac{1}{2^{\rho/\delta}-1}\Big]^d 2^{\xi/\delta},$$

(ii) for any $\xi \ge (2\alpha+\delta)(d-1)$,

$$\dim V_*^d(\xi) \;\ge\; \frac{1}{4} d (2^{\rho/\delta}-1)^{-d} 2^{\xi/\delta},$$

(iii) for any $\xi \ge (2\alpha+\delta)(\nu-1)$,

$$\dim V_\nu^d(\xi) \;\ge\; \frac{1}{4} 2^{-\rho/\delta} \nu \Big[1 + \frac{d}{\nu(2^{\rho/\delta}-1)}\Big]^\nu 2^{\xi/\delta}.$$

Proof. We first prove the second inequality in the lemma. Put $\theta := \alpha/\rho$ and $n := \lfloor(\xi - \alpha(d-1))/\delta\rfloor$. We have $J_*^d(\xi) \supset I_n^d$. Since $\xi \ge (2\alpha+\delta)(d-1)$, Condition (4.1) is satisfied. Hence, by Lemma 4.4,

$$\begin{aligned}
\dim V_*^d(\xi) \;=\; \sum_{k \in J_*^d(\xi)} 2^{|k|_1} \;&\ge\; \sum_{k \in I_n^d} 2^{|k|_1} \\
&\ge\; 2^{-1/(\theta-1)} d 2^{d-2} (1-2^{-1/(\theta-1)})^{-d} 2^n \\
&\ge\; 2^{-1/(\theta-1)} d 2^{d-2} (1-2^{-1/(\theta-1)})^{-d} 2^{(\xi-\alpha(d-1))/\delta - 1} \\
&\ge\; \frac{1}{4} d (2^{\rho/\delta}-1)^{-d} 2^{\xi/\delta}.
\end{aligned}$$

For proving (i) we start similar as in the proof of Lemma 3.5 (see the first three equations in (3.6)) and conclude by using the previous relation

$$\begin{aligned}
\dim V^d(\xi) &\geq \sum_{k=1}^{d} \binom{d}{k} \frac{1}{4} k (2^{\rho/\delta} - 1)^{-k} 2^{\xi/\delta} \\
&= \frac{1}{4} 2^{\xi/\delta} \sum_{k=1}^{d} \frac{d!}{(d-k)!(k-1)!} (2^{\rho/\delta} - 1)^{-k} \\
&= \frac{1}{4} 2^{\xi/\delta} (2^{\rho/\delta} - 1)^{-1} d \sum_{k=0}^{d-1} \binom{d-1}{k} (2^{\rho/\delta} - 1)^{-k} \\
&= \frac{1}{4} (2^{\rho/\delta} - 1)^{-1} d [1 + 1/(2^{\rho/\delta} - 1)]^{d-1} 2^{\xi/\delta} \\
&= \frac{1}{4} 2^{-\rho/\delta} d [1 + 1/(2^{\rho/\delta} - 1)]^{d} 2^{\xi/\delta} \, .
\end{aligned} \tag{4.7}$$

Finally, we prove (iii) with a similar computation as done in (4.7). Indeed, we obtain

$$\begin{aligned}
\dim V^d_\nu(\xi) &\geq \sum_{k=1}^{\nu} \binom{d}{k} \frac{1}{4} k (2^{\rho/\delta} - 1)^{-k} 2^{\xi/\delta} \\
&= \frac{1}{4} 2^{\xi/\delta} \sum_{k=1}^{\nu} \binom{\nu}{k} k \frac{d!(\nu-k)!}{(d-k)!\nu!} (2^{\rho/\delta} - 1)^{-k} \\
&\geq \frac{1}{4} 2^{\xi/\delta} \sum_{k=1}^{\nu} \binom{\nu}{k} k \left(\frac{d}{\nu}\right)^k (2^{\rho/\delta} - 1)^{-k} \\
&= \frac{1}{4} \frac{d}{\nu} \nu (2^{\rho/\delta} - 1)^{-1} 2^{\xi/\delta} \sum_{k=0}^{\nu-1} \binom{\nu-1}{k} \left(\frac{d}{\nu}\right)^k (2^{\rho/\delta} - 1)^{-k} \\
&= \frac{1}{4} (2^{\rho/\delta} - 1)^{-1} d \Big[1 + \frac{d}{\nu(2^{\rho/\delta} - 1)}\Big]^{\nu-1} 2^{\xi/\delta} \\
&= \frac{1}{4} (2^{\rho/\delta} + d/\nu - 1)^{-1} d \Big[1 + \frac{d}{\nu(2^{\rho/\delta} - 1)}\Big]^{\nu} 2^{\xi/\delta} \\
&\geq \frac{1}{4} 2^{-\rho/\delta} \nu \Big[1 + \frac{d}{\nu(2^{\rho/\delta} - 1)}\Big]^{\nu} 2^{\xi/\delta} \, .
\end{aligned}$$

$\square$

Theorem 4.7 *Let $\alpha, \beta, \gamma \in \mathbb{R}$ satisfy the conditions $2(\gamma - \beta) \geq \alpha > \gamma - \beta > 0$. Then, we have* *(i) for any integer* $n \geq \frac{\alpha}{\delta} d 2^{d(2\alpha/\delta+1)}(1+1/(2^{\rho/\delta}-1))^d$,

$$\begin{aligned} \frac{1}{2^{\rho+3\delta}} d^\delta \left(1+\frac{1}{2^{\rho/\delta}-1}\right)^{\delta d} n^{-\delta} \ &\leq \ d_n(U^{\alpha,\beta}, H^\gamma) \\ &\leq \ \left(\frac{\alpha}{\delta}\right)^\delta 2^{2\rho+\delta} d^\delta \left(1+\frac{1}{2^{\rho/\delta}-1}\right)^{\delta d} n^{-\delta}, \end{aligned}$$

(ii) for any integer $n \geq \frac{\alpha}{\delta} d 2^{d(2\alpha/\delta+1)}(2^{\rho/\delta}-1)^{-d}$,

$$\left(\frac{1}{8}\right)^\delta d^\delta \left(2^{\rho/\delta}-1\right)^{-\delta d} n^{-\delta} \ \leq \ d_n(U_*^{\alpha,\beta}, H^\gamma) \ \leq \ \left(\frac{\alpha}{\delta}\right)^\delta 2^{2\rho+\delta} d^\delta \left(2^{\rho/\delta}-1\right)^{-\delta d} n^{-\delta},$$

(iii) for any integer $n \geq \frac{\alpha}{\delta} \nu 2^{\nu(2\alpha/\delta+1)}(1+d/(2^{\rho/\delta}-1))^\nu$,

$$\begin{aligned} \frac{1}{2^{\rho+3\delta}} \nu^\delta \left(1+\frac{d}{\nu(2^{\rho/\delta}-1)}\right)^{\delta\nu} n^{-\delta} \ &\leq \ d_n(U_\nu^{\alpha,\beta}, H^\gamma) \\ &\leq \ \left(\frac{\alpha}{\delta}\right)^\delta 2^{2\rho+\delta} \nu^\delta \left(1+\frac{d}{2^{\rho/\delta}-1}\right)^{\delta\nu} n^{-\delta}. \end{aligned}$$

Proof. Due to Theorem 3.6, we have to prove the lower bounds in this theorem. Let us prove the lower bound for $d_n(U^{\alpha,\beta}, H^\gamma)$. The other lower bounds can be proved in a similar way. It has been shown in the proof of Theorem 3.6 that the function $\varphi(\xi) := \dim V^d(\xi)$ in variable ξ satisfies (3.7) and (3.8). For a given n satisfying the condition in (i) of the theorem, let ξ_m be the number such that

$$\dim V^d(\xi_m) \ \geq \ n+1 \ > \ \dim V^d(\xi_{m-1}). \tag{4.8}$$

Hence, by the corresponding restriction on n in the theorem it follows that $\xi \geq (2\alpha+\delta)(d-1)$. By Lemma 4.6 and (3.8) we obtain

$$\begin{aligned} n \ &\geq \ \frac{1}{4} 2^{-\rho/\delta} d(1+1/(2^{\rho/\delta}-1))^d 2^{\xi_{m-1}/\delta} \\ &\geq \ \frac{1}{8} 2^{-\rho/\delta} d(1+1/(2^{\rho/\delta}-1))^d 2^{\xi_m/\delta}, \end{aligned}$$

or, equivalently,

$$2^{-\xi_m} \ \geq \ (1/2^{\rho/\delta+3})^\delta d^\delta \left(1+1/(2^{\rho/\delta}-1)\right)^{\delta d} n^{-\delta}. \tag{4.9}$$

Consider the set $B(m) := \{f \in V^d(\xi) : \ \|f\|_{H^\gamma} \leq 2^{-\xi_m}\}$ in H^γ. By Lemma 4.2 $B(m) \subset U^{\alpha,\beta}$ and consequently, by Lemma 4.1 and (4.8)

$$d_n(U^{\alpha,\beta}, H^\gamma) \ \geq \ d_n(B(m), H^\gamma) \ \geq \ 2^{-\xi_m}.$$

The last inequalities combining with (4.9) prove the desired inequality. $\square$

Theorem 4.8 *Let $\alpha, \beta, \gamma \in \mathbb{R}$ satisfy the conditions $2(\gamma - \beta) \geq \alpha > \gamma - \beta > 0$. Then we have*

(i) for any $\varepsilon \leq 2^{-(2\alpha+\delta)(d-1)}$,

$$\frac{1}{2^{\rho/\delta+2}} d \left(1 + \frac{1}{2^{\rho/\delta} - 1}\right)^d \varepsilon^{-1/\delta} \ \leq \ n_\varepsilon(U^{\alpha,\beta}, H^\gamma) \ \leq \ \frac{\alpha}{\delta} 2^{2\rho/\delta} d \left(1 + \frac{1}{2^{\rho/\delta} - 1}\right)^d \varepsilon^{-1/\delta},$$

(ii) for any $\varepsilon \leq 2^{-(2\alpha+\delta)(d-1)}$,

$$\frac{1}{4} d (2^{\rho/\delta} - 1)^{-d} \varepsilon^{-1/\delta} \ \leq \ n_\varepsilon(U_*^{\alpha,\beta}, H^\gamma) \ \leq \ \frac{\alpha}{\delta} 2^{2\rho/\delta} d (2^{\rho/\delta} - 1)^{-d} \varepsilon^{-1/\delta},$$

(iii) for any $\varepsilon \leq 2^{-(2\alpha+\delta)(\nu-1)}$,

$$\frac{1}{2^{\rho/\delta+2}} \nu \left[1 + \frac{d}{\nu(2^{\rho/\delta} - 1)}\right]^\nu \varepsilon^{-1/\delta} \ \leq \ n_\varepsilon(U_\nu^{\alpha,\beta}, H^\gamma) \ \leq \ \frac{\alpha}{\delta} 2^{2\rho/\delta} \nu \left[1 + \frac{d}{2^{\rho/\delta} - 1}\right]^\nu \varepsilon^{-1/\delta}.$$

Proof. Due to Theorem 3.7, we have to prove the lower bounds in this theorem. Let us prove the lower bound for $n_\varepsilon(U^{\alpha,\beta}, H^\gamma)$. The other lower bounds can be proved in a similar way. For a given $\varepsilon \leq 2^{-(2\alpha+\delta)(d-1)}$, put $\xi = |\log \varepsilon|$. Consider the set $B_*(\xi) := \{f \in V_*^d(\xi) : \|f\|_{H^\gamma} \leq 2^{-\xi}\}$ in the subspace $V_*^d(\xi)$ of H^γ. By Lemma 4.2 $B_*(\xi) \subset U_*^{\alpha,\beta}$. Hence, by (4.12) and Lemma 4.1 we have

$$d_n(U_*^{\alpha,\beta}, H^\gamma) \ \geq \ d_n(B_*(\xi), H^\gamma) \ \geq \ 2^{-\xi} \ = \ \varepsilon,$$

where $n := \dim V_*^d(\xi) - 1$. Therefore, by the definition and Lemma 4.6(ii),

$$\begin{aligned} n_\varepsilon(U_*^{\alpha,\beta}, H^\gamma) \ &\geq \ \dim V_*^d(\xi) - 1 \\ &\geq \ \frac{1}{4} d (2^{\rho/\delta} - 1)^{-d} 2^{\xi/\delta} \\ &\geq \ \frac{1}{4} d (2^{\rho/\delta} - 1)^{-d} \varepsilon^{-1/\delta}. \end{aligned}$$

The last inequalities combined with (4.11) concludes the proof. ∎

Remark 4.9 (a) *Note, that in Theorems 4.7 and 4.8 the upper and lower bounds (almost) match. Therefore, we have sharp bounds for every d, large enough n and small enough ε (depending on d).*

(b) *Note, that we have exponentially growing d-dependence in the constants in Theorems 4.7, 4.8/(i),(ii). However, this does not imply the curse of dimensionality here. In fact, it is easy to see that*

$$\frac{\log(n_{\varepsilon(d)})}{\varepsilon(d)^{-1} + d} \xrightarrow[d \to \infty]{} 0\,,$$

where $\varepsilon(d) := 2^{-(2\alpha+\delta)(d-1)}$ *in (i). According to the notions in [26, 29] this, surprisingly, indicates weak tractability here. So far this is not a proof since Theorem 4.8 does not give the behavior of* n_ε *for "larger"* ε *and therefore the quantity*

$$\lim_{\varepsilon^{-1}+d\to\infty} \frac{\log(n_\varepsilon)}{\varepsilon^{-1}+d}$$

is not yet proven to be zero. This issue will be clompletely clarified in a forthcoming paper of Ullrich and coauthors, see also [28].

(c) *Note, that in Theorem 3.6(ii), depending on* ρ *and* δ*, the constant* $\left(2^{\rho/\delta}-1\right)^{-\delta d}$ *might decay exponentially in* d*. However, the statement is given for* $n > C_{\alpha/\rho}2^{\rho/\delta}d2^{\alpha d/\delta}(2^{\rho/\delta}-1)^{-d}$ *where* $\alpha > \rho$*. Hence, if the constant decays exponentially one might have to wait exponentially long (with respect to* d*). Therefore, the above result so far does not imply a break of the curse of in Theorems 4.7 and 4.8dimensionality. In fact, this refers to the "footnote" in [22] on page 2224 where the opposite is stated.*

(d) *In contrast to* d *we assume that* ν *is a fixed parameter. Due to the upper bound in Theorem 3.6(iii) we can break the curse of dimensionality here.*

(e) *Based on Theorems 3.7 and 4.8, we have similar statements on the curse of dimensionality in terms of* n_ε*. We mention here the related paper [27] discussing the intractability of* L_∞*-approximation of infinitely differentiable functions on* $\mathbb{I}^d$*.*

4.3 The case $\alpha > \gamma - \beta = 0$

Theorem 4.10 *Let* $\alpha, \beta, \gamma \in \mathbb{R}$ *satisfy the conditions* $\alpha > \gamma - \beta = 0$*. Then the following relations hold true.*

(i) For any $d \geq 2$ *and* $n \geq 2^d$

$$\begin{aligned} 4^{-\alpha}[(1+\log e)(d-1)]^{-\alpha(d-1)} n^{-\alpha}(\log n)^{\alpha(d-1)} &\leq d_n(U^{\alpha,\beta}, H^\gamma) \\ &\leq 4^\alpha \left(\frac{d-1}{2e}\right)^{-\alpha(d-1)} n^{-\alpha}(\log n)^{\alpha(d-1)}. \end{aligned}$$

(ii) For any integer $n \geq 2^{d+1}$ *and* $d \geq 4$*,*

$$\begin{aligned} 4^{-\alpha}[(1+\log e)(d-1)]^{-\alpha(d-1)} n^{-\alpha}(\log n)^{\alpha(d-1)} &\leq d_n(U_*^{\alpha,\beta}, H^\gamma) \\ &\leq 4^\alpha \left(\frac{d-1}{e}\right)^{-\alpha(d-1)} n^{-\alpha}(\log n)^{\alpha(d-1)}. \end{aligned}$$

(iii) If in addition $4 \le \nu \le d/2$*, then for any* $n \ge \frac{\sqrt{5}+3}{2}\binom{d}{\nu}\binom{2\nu-1}{\nu-1}2^{2\nu+1}$,

$$\begin{aligned} & 4^{-\alpha}[(1+\log e)(\nu-1)]^{-\alpha(\nu-1)}\nu^{-\alpha\nu}d^{\alpha\nu}n^{-\alpha}(\log n)^{\alpha(\nu-1)} \\ \le\ & d_n(U^{\alpha,\beta}_\nu, H^\gamma) \\ \le\ & [2(\sqrt{5}+3)]^\alpha\Big(\frac{\nu-1}{e}\Big)^{-\alpha(\nu-1)}\Big(\frac{\nu}{e}\Big)^{-\alpha\nu}d^{\alpha\nu}n^{-\alpha}(\log n)^{\alpha(\nu-1)}\,. \end{aligned}$$

Proof. Due to Theorem 3.13, we have to prove the lower bounds in this theorem. Let us prove the lower bound for $d_n(U^{\alpha,\beta}, H^\gamma)$. The other lower bounds can be proved in a similar way.

For a given $n \ge 2^{d+1}$, there is an unique $m \ge d+1$ such that,

$$(d-1)^{-(d-1)}2^m(m-1)^{d-1}\ \ge\ n+1\ >\ (d-1)^{-(d-1)}2^{m-1}(m-2)^{d-1}. \tag{4.10}$$

Hence, by the inequality $a\log e > \log(1+a),\ a \ge 0$, we obtain

$$m + (d-1)\left(\frac{m-1}{d-1} - 1\right)\log e\ \ge\ m + \log\left(\frac{m-1}{d-1}\right)^{d-1}\ \ge\ \log(n+1),$$

and consequently,

$$m\ \ge\ \frac{\log(n+1) + d\log e}{1+\log e},$$

From the last inequality and (4.10) we derive

$$\begin{aligned} 2^{-\alpha m}\ &\ge\ 2^{-\alpha(m-1)}(d-1)^{\alpha(d-1)}(m-2)^{-\alpha(d-1)}2^{-\alpha}(d-1)^{-\alpha(d-1)}(m-2)^{\alpha(d-1)} \\ &\ge\ 2^{-\alpha}(d-1)^{-\alpha(d-1)}\left[\frac{\log(n+1)+d\log e}{1+\log e} - 2\right]^{\alpha(d-1)}(n+1)^{-\alpha}. \end{aligned}$$

We have $(n+1)^{-\alpha} \ge 2^{-\alpha}n^{-\alpha}$. Moreover, by the inequality $d \ge 2 + 2/\log e$ one can verify that

$$\frac{\log(n+1)+d\log e}{1+\log e} - 2\ \ge\ \frac{\log n}{1+\log e},$$

and consequently,

$$2^{-\alpha m}\ \ge\ 4^{-\alpha}[(1+\log e)(d-1)]^{-\alpha(d-1)}n^{-\alpha}(\log n)^{\alpha(d-1)}\,. \tag{4.11}$$

From Lemma 3.8 and (3.11), we obtain

$$\dim V^d_*(\alpha m)\ =\ \sum_{k\in K^d_*(m)} 2^{|k|_1}\ \ge\ (d-1)^{-(d-1)}2^m(m-1)^{d-1}\ \ge\ n+1. \tag{4.12}$$

Consider the set $B_*(m) := \{f \in V_*^d(\xi) : \ \|f\|_{H^\gamma} \le 2^{-\alpha m}\}$ in the subspace $V_*^d(\alpha m)$ of H^γ. By Lemma 4.2 we have $B_*(m) \subset U_*^{\alpha,\beta}$. Hence, by (4.12) and Lemma 4.1 we obtain

$$d_n(U_*^{\alpha,\beta}, H^\gamma) \ \ge \ d_n(B_*(m), H^\gamma) \ \ge \ 2^{-\alpha m}.$$

The last inequality combined with (4.11) finishes the proof of the desired lower bound. □

Theorem 4.11 *Let $\alpha, \beta, \gamma \in \mathbb{R}$ satisfy the conditions $\alpha > \gamma - \beta = 0$. Then the following relations hold true.*

(i) For any $0 < \varepsilon \le 2^{-\alpha d}$,

$$\frac{1}{2}[\alpha(d-1)]^{-(d-1)}\varepsilon^{-1/\alpha}|\log\varepsilon|^{d-1} \ \le \ n_\varepsilon(U^{\alpha,\beta}, H^\gamma) \ \le \ 4\Big(\frac{\alpha(d-1)}{2e}\Big)^{-(d-1)}\varepsilon^{-1/\alpha}|\log\varepsilon|^{d-1}.$$

(ii) For any $0 < \varepsilon \le 2^{-\alpha d}$ and $d \ge 4$,

$$\frac{1}{2}[2\alpha(d-1)]^{-(d-1)}\varepsilon^{-1/\alpha}|\log\varepsilon|^{d-1} \ \le \ n_\varepsilon(U_*^{\alpha,\beta}, H^\gamma) \ \le \ 4\Big(\frac{\alpha(d-1)}{e}\Big)^{-(d-1)}\varepsilon^{-1/\alpha}|\log\varepsilon|^{d-1}.$$

(iii) If in addition $2 \le \nu \le d/2$ then for any $0 < \varepsilon \le 2^{-2\alpha\nu}$,

$$\begin{aligned} &\frac{1}{2}[2\alpha(\nu-1)]^{-(\nu-1)}\nu^{-\nu}d^\nu\varepsilon^{-1/\alpha}|\log\varepsilon|^{\nu-1} \\ &\le n_\varepsilon(U_\nu^{\alpha,\beta}, H^\gamma) \ \le \ 2(\sqrt{5}+3)\Big(\frac{\alpha(\nu-1)}{e}\Big)^{-(\nu-1)}(\nu/e)^{-\nu}d^\nu\varepsilon^{-1/\alpha}|\log\varepsilon|^{\nu-1}. \end{aligned}$$

Proof. Due to Theorem 3.14, we have to prove the lower bounds in this theorem. Let us prove the lower bound for $n_\varepsilon(U_*^{\alpha,\beta}, H^\gamma)$. The other lower bounds can be proved in a similar way. For a given $\varepsilon \le 2^{-\alpha d}$ we take $m \ge d \ge 4$ such that

$$2^{-\alpha m} \ \ge \ \varepsilon \ > \ 2^{-\alpha(m+1)}.$$

The right-hand inequality gives

$$2^m \ge \ \frac{1}{2}\varepsilon^{-1/\alpha} \quad \text{and} \quad m \ \ge \ \alpha^{-1}|\log\varepsilon| - 1.$$

Consider the set $B_*(m) := \{f \in V_*^d(\xi) : \ \|f\|_{H^\gamma} \le 2^{-\alpha m}\}$ in the subspace $V_*^d(\alpha m)$ of H^γ. By Lemma 4.2 it holds $B_*(m) \subset U_*^{\alpha,\beta}$. Hence, by (4.12) and Lemma 4.1 we have

$$d_n(U_*^{\alpha,\beta}, H^\gamma) \ \ge \ d_n(B_*(m), H^\gamma) \ \ge \ 2^{-\alpha m} \ \ge \ \varepsilon,$$

where $n := \dim V_*^d(\alpha m)-1$. Therefore, by Lemma 3.8, (3.11) and the inequality $|\log\varepsilon| \geq 4\alpha$ we get

$$\begin{aligned} n_\varepsilon(U_*^{\alpha,\beta}, H^\gamma) &\geq \dim V_*^d(\alpha m) - 1 \\ &\geq \frac{1}{2} 2^m \binom{m-1}{d-1} \\ &\geq \frac{1}{2}(d-1)^{-(d-1)}(m-1)^{d-1}\varepsilon^{-1/\alpha} \\ &\geq \frac{1}{2}(d-1)^{-(d-1)}(\alpha^{-1}|\log\varepsilon| - 2)^{d-1}\varepsilon^{-1/\alpha} \\ &\geq \frac{1}{2}[2\alpha(d-1)]^{-(d-1)}\varepsilon^{-1/\alpha}|\log\varepsilon|^{d-1}. \end{aligned}$$

$\square$

Remark 4.12 *Note, that in [22] the authors did not prove any lower bounds for the dimensions of the optimized sparse grid spaces and the approximation error for the linear approximation in H^γ of functions from the class $U_*^{\alpha,\beta}$ (which is defined via a biorthogonal wavelet decomposition, see next section).*

5 Biorthogonal wavelet decompositions

In this section, we discuss how to extend the results for the dyadic harmonic decomposition in this paper to periodic biorthogonal wavelet decompositions and other periodic and non-periodic decompositions.

Let $\Phi := \{\varphi_{k,s}\}_{k\in\mathbb{Z}_+, s\in Q_k}$ and $\tilde\Phi := \{\tilde\varphi_{k,s}\}_{k\in\mathbb{Z}_+, s\in Q_k}$ be biorthogonal systems in L_2, where $Q_k := \{s \in \mathbb{Z} : 0 \leq s < 2^k\}$. We will assume that $\{\varphi_{k,s}\}_{k\in\mathbb{Z}_+, s\in Q_k}$ forms a Riesz basis for L_2, that is

$$\left\| \sum_{k\in\mathbb{Z}_+} c_{k,s}\varphi_{k,s} \right\|^2 \asymp \sum_{k\in\mathbb{Z}_+} \sum_{s\in Q_k} |c_{k,s}|^2.$$

Therefore, every $f \in L_2$ has a unique representation

$$f = \sum_{k\in\mathbb{Z}_+} \sum_{s\in Q_k} (f, \tilde\varphi_{k,s})\varphi_{k,s},$$

and there hold true the dyadic biorthogonal wavelet decomposition

$$f = \sum_{k\in\mathbb{Z}_+} q_k(f),$$

with the norm equivalence

$$\|f\|^2 \asymp \sum_{k \in \mathbb{Z}_+} \|q_k(f)\|^2,$$

where

$$q_k(f) := \sum_{s \in Q_k} (f, \tilde{\varphi}_{k,s}) \varphi_{k,s}$$

One of the most important cases of biorthogonal systems in L_2 which has wide applications are wavelet biorthogonal systems. Univariate periodic wavelet biorthogonal systems $\Phi := \{\varphi_{k,s}\}_{k \in \mathbb{Z}_+, s \in Q_k}$ and $\tilde{\Phi} := \{\tilde{\varphi}_{k,s}\}_{k \in \mathbb{Z}_+, s \in Q_k}$ are of the form

$$\varphi_{k,s}(x) = \varphi^k(x - 2\pi 2^{-k} s), \quad \tilde{\varphi}_{k,s}(x) = \tilde{\varphi}^k(x - 2\pi 2^{-k} s),$$

where $\{\varphi^k\}_{k \in \mathbb{Z}_+}$ and $\{\tilde{\varphi}^k\}_{k \in \mathbb{Z}_+}$ are the sequences of mother wavelets which in particular, can be received from the mother wavelets ψ and $\tilde{\psi}$ of univariate nonperiodic wavelet biorthogonal systems by the periodization formula

$$\varphi^k(x) = \sum_{s \in \mathbb{Z}} \psi(2^k(x + 2\pi s)), \quad \tilde{\varphi}^k(x) = \sum_{s \in \mathbb{Z}} \tilde{\psi}(2^k(x + 2\pi s)).$$

We assume the following conditions on Φ. There hold the Jackson type inequality

$$\inf_{g \in \Sigma_k} \|f - g\|_{L_2} \leq C 2^{-mk} \|f\|_{H^r},$$

for some $m \in \mathbb{N}$, and the Bernstein type inequality

$$\|f\|_{H^l} \leq C 2^{lk} \|f\|_{L_2}, \quad f \in V_k,$$

for some $l \leq r$ with $0 < r \leq m$, where $V_k := \operatorname{span}\{\varphi_{k,s} : s \in Q_k\}$. We also assume that similar inequalities hold for the dual system $\tilde{\Phi}$ with parameters $\tilde{m}$ and $\tilde{r}$.

For distributions f and $k \in \mathbb{Z}^d_+$, let us introduce the following operator:

$$q_k(f) := \prod_{j=0}^{d} q_{k_j}(f)$$

where the univariate operator q_{k_j} is applied to f as a univariate function in variable x_j while the other variables are held fixed.

If $f \in L_2$, we have

$$\|f\|^2 \asymp \sum_{k \in \mathbb{Z}^d_+} \|q_k(f)\|^2.$$

Let us use the notation $H_*^{\alpha,\beta}$ to denote the subspace in $H^{\alpha,\beta}$ of all functions f having the biorthogonal wavelet decomposition

$$\|f\|^2 \asymp \sum_{k\in\mathbb{N}^d} \|q_k(f)\|^2.$$

The following lemma has been proved in [22].

Lemma 5.1 *Let $\alpha,\beta \in \mathbb{R}$ satisfy the restrictions $0 \le \alpha < r$ and $0 \le \alpha+\beta < r$, where r is the parameter in the Jackson and Bernstein type inequalities. Then there holds true the following norm equivalence*

$$\|f\|^2_{H_*^{\alpha,\beta}} \ \asymp \ \sum_{k\in\mathbb{N}^d} 2^{2(\alpha|k|_1+\beta|k|_\infty)}\|q_k(f)\|^2_{L_2}.$$

Lemma 2.1 and Lemma 5.1 show that functions $f \in H_*^{\alpha,\beta}$ have similar dyadic harmonic and biorthogonal wavelet decompositions with the same equivalent norms. There are other analogous decompositions of $H_*^{\alpha,\beta}$ not only for periodic functions but for non-periodic functions defined on a d-dimensional cube. Indeed, we can treat spaces $H_*^{\alpha,\beta}$ as well $H^{\alpha,\beta}$, $H_\nu^{\alpha,\beta}$ and classes $U^{\alpha,\beta}$, $U_*^{\alpha,\beta}$, $U_\nu^{\alpha,\beta}$ in a more general form which are suitable for different applications.

Let H be a separable Hilbert space and H have the following dyadic decomposition. Namely, H is decomposed into pairwise orthogonal subspaces $W_k, \ k \in \mathbb{Z}^d_+$,

$$H \ = \ \bigoplus_{k\in\mathbb{Z}^d_+} W_k,$$

with

$$\dim W_k \ = \ 2^{|k|_1}.$$

Then every $f \in H$ can be decomposed into a series

$$f \ = \ \sum_{k\in\mathbb{Z}^d_+} p_k(f), \quad p_k(f) \in W_k,$$

with

$$\|f\|^2 \ = \ \sum_{k\in\mathbb{Z}^d_+} \ \|p_k(f)\|^2.$$

We define $H^{\alpha,\beta}$ as the Hilbert space of formal series

$$f \ = \ \sum_{k\in\mathbb{Z}^d_+} p_k(f), \quad p_k(f) \in W_k,$$

for which the following norm is finite

$$\|f\|^2_{H^{\alpha,\beta}} = \sum_{k\in\mathbb{Z}^d_+} 2^{2(\alpha|k|_1+\beta|k|_\infty)}\|p_k(f)\|^2.$$

With this definition we have $H^{0,0} = H$. For $\alpha = 0$, we put $H^{0,\beta} = H^\gamma$ for $\beta = \gamma$.

We define the subspaces $H^{\alpha,\beta}_*$ and $H^{\alpha,\beta}_\nu$, $1 \le \nu \le d-1$, in $H^{\alpha,\beta}$ as follows. The subspace $H^{\alpha,\beta}_*$ is the set of all $f \in H^{\alpha,\beta}$ such that such that

$$p_k(f) = 0 \text{ if } \prod_{j=0}^d k_j = 0.$$

The subspace $H^{\alpha,\beta}_\nu$ is the set of all $f \in H^{\alpha,\beta}$ such that

$$p_k(f) = 0 \text{ if } |\sigma(s)| > \nu.$$

Denote by $U^{\alpha,\beta}$, $U^{\alpha,\beta}_*$ and $U^{\alpha,\beta}_\nu$ the unit ball in $H^{\alpha,\beta}$, $H^{\alpha,\beta}_*$ and $H^{\alpha,\beta}_\nu$, respectively. (For convenience, here we use the same notations $H^{\alpha,\beta}$, $U^{\alpha,\beta}$, $U^{\alpha,\beta}_*$, $U^{\alpha,\beta}_\nu$ as in Section 2 for the harmonic dyadic decomposition.) For the above defined function spaces and function sets all the results in Sections 3 and 4 remain true.

Acknowledgments The work of the first named author was supported by Grant 102.01-2012.15 of the National Foundation for Development of Science and Technology (Vietnam). The both authors would like to thank the Hausdorff Research Institute for Mathematics (HIM) and the organizers of the HIM Trimester Program “Analysis and Numerics for High Dimensional Problems”, where this paper was initiated, for providing a fruitful research environment and additional financial support. Last but not least, the authors would like to thank the referees for a critical reading of the manuscript and for several valuable suggestions which helped to improve its presentation.

References

[1] K.I. Babenko, On the approximation of periodic functions of several variables by trigonometric polynomials, Dokl. Akad. Nauk USSR **132**(1960), 247–250; English transl. in Soviet Math. Dokl. 1 (1960).

[2] R. Bellmann, Dynamic Programming, Princeton University Press, Princeton, 1957.

[3] H.-J. Bungartz and M. Griebel, A note on the complexity of solving Poisson’s equation for spaces of bounded mixed derivatives, J. Complexity, **15**(1999), 167–199.

[4] H.-J. Bungartz and M. Griebel, Sparse grids, Acta Numer., **13**(2004),147–269.

[5] J. Céa, Approximation variationnelle des problémes aux limites, Annales de l'institut Fourier, **14**(1964), 345–444.

[6] A. Chernov, Sparse polynomial approximation in positive order Sobolev spaces with bounded mixed derivatives and applications to elliptic problems with random loading, Appl. Numer. Math., **62**(2012), 360-377.

[7] A. Chernov and C. Schwab, Sparse p-version BEM for first kind boundary integral equations with random loading, Appl. Numer. Math., **59**(2009), 2698–2712.

[8] R. A. DeVore, S. V. Konyagin, and V. N. Temlyakov, Hyperbolic wavelet approximation, Constr. Approx., **14**(1998), 1–26.

[9] R. DeVore, G. Petrova and P. Wojtaszczyk, Approximating functions of few variables in high dimensions, Constructive Approximation, **33**(2011), 125–143.

[10] D. Dũng, Some approximative characteristics of the classes of smooth functions of several variables in the metric of L_2, Uspekhi Mat. Nauk, **34**(1979), 189–190.

[11] D. Dũng, Mean ε-dimension of the functional class $B_{G,p}$, Mat. Zametki, **28**(1980), 727–736.

[12] D. Dũng, Approximation of functions of several variables on a torus by trigonometric polynomials, Mat. Sb. (N.S.), **131**(173)(2)(1986), 251–271.

[13] D. Dũng, On optimal recovery of multivariate periodic functions, In: Harmonic Analysis (S. Igary ed.), Springer-Verlag 1991, Tokyo-Berlin, pp. 96-105.

[14] D. Dũng, Optimal recovery of functions of a certain mixed smoothness, Vietnam J. Math., **20**(2)(1992), 18–32.

[15] D. Dũng, B-spline quasi-interpolant representations and sampling recovery of functions with mixed smoothness, Journal of Complexity, **27**(2011), 541-567.

[16] D. Dũng and D. Dryanov, The minimal number of sample values for recovery of non-bandlimited functions, Atti Semin. Mat. Fis. Univ. Modena, **39**(1991), 423–431.

[17] J. Garcke, M. Griebel, and M. Thess, Data mining with sparse grids, Computing, **67**(2001), 225–253.

[18] T. Gerstner, Sparse Grid Quadrature Methods for Computational Finance, Habilitation, Institute for Numerical Simulation, University of Bonn, 2007.

[19] T. Gerstner and M. Griebel, Sparse grids, In R. Cont, editor, Encyclopedia of Quantitative Finance, John Wiley and Sons, Feb. 2010.

[20] M. Griebel and J. Hamaekers, Tensor product multiscale many-particle spaces with finite-order weights for the electronic Schrödinger equation, Zeitschrift für Physikalische Chemie, **224**(2010), 527–543.

[21] M. Griebel and S. Knapek, Optimized tensor-product approximation spaces, Constructive Approximation, **16**(2000), 525–540.

[22] M. Griebel and S. Knapek, Optimized general sparse grid approximation spaces for operator equations, Math. Comp., 78(268):2223–2257, 2009.

[23] A.N. Kolmogorov, Über die beste Annäherung von Funktionen einer Funktionklasse, Ann. Math., **37**(1936), 107–111.

[24] A.N. Kolmogorov, Selected papers, Mathematics and Mechanics, Volume I, Nauka, Moscow 1985 (in Russian).

[25] P.D. Lax and A.N. Milgram, "Parabolic equations". Contributions to the theory of partial differential equations, Ann. of Math. Studies, **33**(1954), 167–190.

[26] E. Novak and H. Woźniakowski, Tractability of Multivariate Problems, Volume I: Linear Information, EMS Tracts in Mathematics, Vol. 6. Eur. Math. Soc. Publ. House, Zürich 2008.

[27] E. Novak and H. Woźniakowski, Approximation of infinitely differentiable multivariate functions is intractable, Journal of Complexity **25**(2009), 398–404.

[28] E. Novak and H. Woźniakowski, Optimal order of convergence and (in)tractability of multivariate approximation of smooth functions, Constr. Appr. **30**(2009), 457-473.

[29] E. Novak and H. Woźniakowski, Tractability of Multivariate Problems, Volume II: Standard Information for Functionals EMS Tracts in Mathematics, Vol. 12, Eur. Math. Soc. Publ. House, Zürich 2010.

[30] A. Pinkus, N-Widths in Approximation Theory, Springer-Verlag, New York, 1985.

[31] S.A. Smolyak, Quadrature and interpolation formulas for tensor products of certain classes of functions, Dokl. Akad. Nauk, **148**(1963), 1042–1045.

[32] C. Schwab and R.-A. Todor, Sparse finite elements for elliptic problems with stochastic loading, Numer. Math. **95**(2003), 707–734.

[33] C. Schwab and R.-A. Todor, Sparse finite elements for stochastic elliptic problems higher order moments, Computing **71**(2003), 43–63.

[34] W. Sickel and T. Ullrich, The Smolyak algorithm, sampling on sparse grids and function spaces of dominating mixed smoothness, East J. Approx. **13**(2007), 387–425.

[35] W. Sickel and T. Ullrich, Tensor products of Sobolev-Besov spaces and applications to approximation from the hyperbolic cross, J. Approx. Theory **161**(2009), 748–786.

[36] W. Sickel and T. Ullrich, Spline interpolation on sparse grids, Applicable Analysis **90**(2011), 337-383.

[37] L. Schwartz, Théorie des Distributions, Hermann & Cie, Paris, 1966.

[38] V. N. Temlyakov, Approximation of Periodic Functions, Computational Mathematics and Analysis Series, Nova Science Publishers Inc., Commack, NY, 1993.

[39] V. M. Tikhomirov, Widths of sets in function spaces and the theory of best approximations, Uspekhi Mat. Nauk **15** No.3 (93)(1960), 81–120. English translation in Russian Math. Survey, 15, 1960.

[40] V. M. Tikhomirov, Some Problems in Approximation Theory (in Russian), Moscow State University, 1985.

[41] R.-A. Todor, A new approach to energy-based sparse finite-element spaces, IMA J Numer Anal **29**(1)(2009),72-85.

[42] T. Ullrich, Smolyak's algorithm, sampling on sparse grids and Sobolev spaces of dominating mixed smoothness, East J. Approx., **14**(2008), 1–38.

[43] G. Wasilkowski and H. Woźniakowski, Explicit cost bounds of algorithms for multivariate tensor product problems, J. of Complexity, **11**(1995), 1–56.

[44] P. Wojtaszczyk, Complexity of approximation of functions of few variables in high dimensions, J. of Complexity, **27**(2011), 141–150.

[45] H. Yserentant, On the regularity of the electronic Schrödinger equation in Hilbert spaces of mixed derivatives, Numer. Math., **98**(2004), 731–759.

[46] H. Yserentant, Sparse grid spaces for the numerical solution of the electronic Schrödinger equation, Numer. Math., **101**(2005), 381–389.

[47] H. Yserentant, Regularity and Approximability of Electronic Wave Functions, volume 2000 of Lecture Notes in Mathematics. Springer-Verlag, Berlin, 2010.