frontiers | Frontiers in Applied Mathematics and Statistics

TYPE Original Research
PUBLISHED 24 August 2023
DOI 10.3389/fams.2023.1216331

OPEN ACCESS

EDITED BY
Frank Filbir,
Helmholtz Association of German Research Centres (HZ), Germany

REVIEWED BY
Feng Dai,
University of Alberta, Canada
Jan Vybiral,
Czech Technical University in Prague, Czechia

*CORRESPONDENCE
Serhii Stasyuk
✉ stasyuk@imath.kiev.ua

RECEIVED 03 May 2023
ACCEPTED 08 June 2023
PUBLISHED 24 August 2023

CITATION
Byrenheid G, Stasyuk S and Ullrich T (2023) L_p-Sampling recovery for non-compact subclasses of L_∞.
Front. Appl. Math. Stat. 9:1216331.
doi: 10.3389/fams.2023.1216331

L_p-Sampling recovery for non-compact subclasses of L_∞

Glenn Byrenheid[1], Serhii Stasyuk[2,3]* and Tino Ullrich[2]

[1]Institute of Mathematics, Friedrich-Schiller-University Jena, Jena, Germany, [2]Faculty of Mathematics, Chemnitz University of Technology, Chemnitz, Germany, [3]Department of Theory of Functions, Institute of Mathematics of NAS of Ukraine, Kyiv, Ukraine

In this article, we study the sampling recovery problem for certain relevant multivariate function classes on the cube $[0,1]^d$, which are not compactly embedded into $L_\infty([0,1]^d)$. Recent tools relating the sampling widths to the Kolmogorov or best m-term trigonometric widths in the uniform norm are therefore not applicable. In a sense, we continue the research on the small smoothness problem by considering limiting smoothness in the context of Besov and Triebel-Lizorkin spaces with dominating mixed regularity such that the sampling recovery problem is still relevant. There is not much information available on the recovery of such functions except for a previous result by Oswald in the univariate case and Dinh Dũng in the multivariate case. As a first step, we prove the uniform boundedness of the ℓ_p-norm of the Faber coefficients at a fixed level by Fourier analytic means. Using this, we can control the error made by a (Smolyak) truncated Faber series in $L_q([0,1]^d)$ with $q < \infty$. It turns out that the main rate of convergence is sharp. Thus, we obtain results also for $S^1_{1,\infty}F([0,1]^d)$, a space "close" to $S^1_1W([0,1]^d)$, which is important in numerical analysis, especially numerical integration, but has rather poor Fourier analytical properties.

KEYWORDS

sampling recovery, limiting smoothness, non-compact embedding, Faber basis, mixed smoothness

1. Introduction

In this article, we continue to study the approximation power of Smolyak sparse grid sampling recovery for multivariate function classes with small mixed smoothness $S^r_{p,\theta}B([0,1]^d)$ and $S^r_{p,\theta}F([0,1]^d)$ based on the multivariate Faber representation [1, 2] of f. The advantage of such a representation is the fact that the coefficient functionals only use discrete functional values of f, see (2.1) and (2.3) below, which allow for various applications. For instance, Kempka et al. [3] used the Faber system to analyze the path regularity of the Brownian motion, where a small smoothness setting is also required.

It provides a powerful tool for studying various situations of the sampling recovery problem where errors are measured in L_q. Surprisingly, the proposed method turned out to be sharp in several regimes. A systematic discretization of multivariate functions with mixed smoothness in terms of Faber coefficients is given [1, 2, 4, 5], see also [6] for further history. In this article, we study an endpoint situation for the sampling recovery problem on the cube $[0,1]^d$, where $r = 1/p$ in the source space and the fine parameter $\theta \le 1$ is sufficiently small. This still allows for an embedding into the continuous functions (therefore making function evaluations possible). However, the embedding into L_q is only compact if $q < \infty$.

Recent observations regarding the problem of optimal sampling recovery of function classes in L_2 bring classes with mixed smoothness to the focus again since several newly developed techniques only work for Hilbert-Schmidt operators [7–9] or, more generally, in situations where certain asymptotic characteristics (approximation numbers) are square summable [10, 11]. We need new techniques for situations where this is not the case. In Temlyakov and Ullrich [12, 13], the authors consider the range of small smoothness where one is far away from square summability of the corresponding widths. However, we have the compact embedding into L_∞ for those examples. This embedding seems to be of crucial importance when (non-linear) sampling widths $(\varrho_m)_m$, defined by

$$\varrho_m(\mathbf{F},X) := \inf_{x^1,\dots,x^m} \inf_{R:\mathbb{C}^m\to X} \sup_{f\in\mathbf{F}} \|f - R(f(x^1),\dots,f(x^m))|X\|,$$

with $X = L_2$ are related to certain asymptotic characteristics such as Kolmogorov $(d_m)_m$ or best m-term trigonometric widths $(\sigma_m)_m$ in L_∞. It has been shown by Temlyakov [14], Bartel et al. [9] and Jahn et al. [15] that the following inequalities hold for relevant function classes $\mathbf{F}$

$$\varrho^{lin}_{bm}(\mathbf{F},L_2) \leq \frac{C}{(b-1)^{3/2}} d_m(\mathbf{F},L_\infty), \qquad (1.1)$$
$$\varrho_{\lceil Dm(\log m)^3\rceil}(\mathbf{F},L_2) \lesssim \sigma_m(\mathbf{F},L_\infty)\,.$$

Here, we have some constants $C, D > 0$ and oversampling parameter $1 < b \leq 2$ in the first relation. We speak about *linear sampling widths* denoted by $(\varrho^{lin}_m)_m$ if the reconstruction mapping R is linear. The second relation might not need a compact embedding into L_∞ at first glance. However, the proof heavily relies on certain compactness properties of the embedding, see [15]. Clearly, a compact embedding into L_∞ allows us to use the decaying Kolmogorov widths for controlling the linear sampling widths. As discussed by Temlyakov and Ullrich [12], sparse grid techniques [16] perform asymptotically worse by a log factor for classes with mixed smoothness compactly embedded into L_∞.

In this article, we continue our research in this direction. Note that there are several relevant (multivariate) function classes $\mathbf{F}$ which are continuously but not compactly embedded into L_∞. Thus, at least the first relation in (1.1) is useless. In fact, only few results have been published on reconstructing functions from samples, which only satisfy a Besov regularity with smoothness $r = 1/p$ or Sobolev type regularity with $r = 1$ and $p = 1$. Our aim is to investigate this problem systematically in the Fourier analytic context. As a first step, we prove the following relation for the Faber coefficients $d^2_{j,k}(f)$ of f, namely

$$\sup_{j\in\mathbb{N}^d_{-1}} \Big(\sum_{k\in\mathbb{Z}^d} |d^2_{j,k}(f)|^p\Big)^{1/p} \lesssim \|f|S^{1/p}_{p,\min\{p,1\}}B(\mathbb{R}^d)\|$$

in case $1/2 \leq p \leq \infty$. Our contribution is a proof that works for Fourier analytic defined spaces and allows for incorporating also the extreme cases $S^2_{1/2,1/2}B(\mathbb{R}^d)$ and $S^0_{\infty,1}B(\mathbb{R}^d)$ [in contrast to Dũng [17], Theorem 3.2]. It represents an extreme case of the considered limit situation, in which the Faber approximation can benefit from the highest regularity of smoothness equals 2. The above relation directly implies that the truncated Faber representation still works well when we consider errors in L_q with $q < \infty$. We make progress toward the solution of an open problem mentioned [[18], Section 3.2]. The univariate class $B^{1/p}_{p,1}([0,1])$ and its approximation by equidistant samples in $[0,1]$ have been considered by Oswald [19] in the beginning of the 80s. In 2011, Dũng [[17], Theorem 3.2] obtained results for the multivariate situation in the framework of Besov spaces with a bounded mixed difference. In contrast to the spaces considered by Dũng [17], we highlight that the Besov and Triebel-Lizorkin spaces considered here are defined using Fourier analytic building blocks. Note that in the considered limiting situation and $p < 1$, it is not yet known whether these spaces coincide with the ones considered by Dũng [see [20], Remark 2.3.4/2].

Notation. In general, $\mathbb{N}$ denotes the natural numbers, $\mathbb{N}_0 = \mathbb{N}\cup\{0\}$, $\mathbb{N}_{-1} = \mathbb{N}_0\cup\{-1\}$, $\mathbb{Z}$ denotes the integers, $\mathbb{R}$ denotes the real numbers, and $\mathbb{C}$ denotes the complex numbers. The letter d is always reserved for the underlying dimension in $\mathbb{R}^d, \mathbb{Z}^d$, etc. For $a \in \mathbb{R}$, we denote $a_+ := \max\{a,0\}$. For $0 < p \leq \infty$ and $x \in \mathbb{R}^d$, we denote $|x|_p = (\sum_{i=1}^d |x_i|^p)^{1/p}$ with the usual modification in the case $p = \infty$. We further denote $x_+ := ((x_1)_+,\dots,(x_d)_+)$ and $|x|_+ := |x_+|_1$. By $(x_1,\dots,x_d) > 0$, we mean that each coordinate is positive. If X and Y are two (quasi-)normed spaces, the (quasi-)norm of an element x in X will be denoted by $\|x|X\|$. The symbol $X\hookrightarrow Y$ indicates that the identity operator is continuous. For two sequences a_n and b_n, we will write $a_n \lesssim b_n$ if there exists a constant $c > 0$ such that $a_n \leq c\, b_n$ for all n. We will write $a_n \asymp b_n$ if $a_n \lesssim b_n$ and $b_n \lesssim a_n$.

2. The tensor Faber basis

As a main tool, we will use decompositions of functions in terms of a Faber series expansion.

2.1. The univariate Faber basis

Let us briefly recall the basic facts about the Faber basis taken from [[4], 3.2.1 and 3.2.2]. For $j \in \mathbb{N}_0$ and $k \in \mathbb{D}_j := \{0,1,\dots,2^j-1\}$, we denote the dyadic interval by $I_{j,k}$ given by

$$I_{j,k} = [2^{-j}k, 2^{-j}(k+1)]\,.$$

Definition 2.1. [The univariate Faber system] Let

$$h(t) = \begin{cases} 1 : t\in[0,1/2)\,, \\ -1 : t\in(1/2,1]\,, \textit{and} \\ 0 : \text{otherwise}\,, \end{cases}$$

the Haar function and $v(x)$ be the integrated Haar function, i.e.,

$$v(x) := 2\int_0^x h(t)\,dt, \quad x\in\mathbb{R}\,,$$

and for $j \in \mathbb{N}_0$ and $k \in \mathbb{D}_j$, then

$$v_{j,k}(\cdot) := v(2^j\cdot - k)\,.$$

For notational reasons, we let $v_{-1,0} = x$ and $v_{-1,1} := 1 - x$ for $j = -1$ and obtain the univariate Faber basis

$$\{v_{j,k} : j\in\mathbb{N}_{-1}, k\in\mathbb{D}_j\}\,,$$

where $\mathbb{D}_{-1} := \mathbb{D}_1 = \{0,1\}$.

Faber [21] observed that every continuous (non-periodic) function f on $[0,1]$ can be represented as

$$f(x) = f(0)\cdot(1-x) + f(1)\cdot x - \frac{1}{2}\sum_{j=0}^{\infty}\sum_{k=0}^{2^j-1}\Delta^2_{2^{-j-1}}(f, 2^{-j}k)v_{j,k}(x), \tag{2.1}$$

with uniform convergence [see e.g., [4], Theorem 2.1, Step 4]. The analysis of Besov and Triebel-Lizorkin spaces on $\mathbb{R}$ as defined in Section 4 requires a version of the Faber representation acting on $\mathbb{R}$. For this purpose, we extend the number of translations to the whole integers and obtain

$$f(x) = \sum_{k\in\mathbb{Z}} f(k)v_{-1,k}(x) - \frac{1}{2}\sum_{j=0}^{\infty}\sum_{k\in\mathbb{Z}}\Delta^2_{2^{-j-1}}(f, 2^{-j}k)v_{j,k}(x), \tag{2.2}$$

where $v_{-1,k}(\cdot) := v_{0,0}((\cdot + 1 + k)/2)$.

2.2. The multivariate Faber basis

Let $f(x_1,\ldots,x_d)$ be a d-variate function, $f \in C(\mathbb{R}^d)$. By fixing all variables except x_i, we obtain by $g(\cdot) = f(x_1,\ldots,x_{i-1},\cdot,x_{i+1},\ldots,x_d)$, a univariate continuous function. By applying (2.2) in every such component, we obtain the representation

$$f(x) = \sum_{j\in\mathbb{N}^d_{-1}}\sum_{k\in\mathbb{Z}^d} d^2_{j,k}(f)v_{j,k}(x) \tag{2.3}$$

in $C(K)$, $K\subset\mathbb{R}^d$ compact, where

$$d^2_{j,k}(f) := (-2)^{-|e(j)|}\Delta^{2,e(j)}_{2^{-(j+1)}}(f, x_{j,k}), \quad j\in\mathbb{N}^d_{-1},\ k\in\mathbb{Z}^d,$$

$$v_{j,k}(x_1,\ldots,x_d) := v_{j_1,k_1}(x_1)\cdots v_{j_d,k_d}(x_d), \quad j\in\mathbb{N}^d_{-1},\ k\in\mathbb{Z}^d,$$

and $|e(j)|$ denotes the cardinality of $e(j)$. Here we put $e(j) = \{i : j_i \neq -1\}$ and $x_{j,k} = (2^{-(j_1)_+}k_1,\ldots,2^{-(j_d)_+}k_d)$.

In Section 6, we apply the Faber series expansion for functions on the d-variate unit cube $[0,1]^d$. For this purpose, we simply truncate the series expansion to all translations whose support has a non-empty intersection with $[0,1]^d$. That is

$$\mathbb{D}_j := \mathbb{D}_{j_1}\times\cdots\times\mathbb{D}_{j_d}, \quad j=(j_1,\ldots,j_d)\in\mathbb{N}^d_{-1}.$$

With similar arguments as above, we obtain for $f\in C([0,1]^d)$ the representation

$$f(x) = \sum_{j\in\mathbb{N}^d_{-1}}\sum_{k\in\mathbb{D}_j} d^2_{j,k}(f)v_{j,k}(x). \tag{2.4}$$

3. Faber coefficients and bandlimited functions

In the sequel we deal with two tensor domains. On the one hand the d-variate unit cube $\mathbb{I}^d = [0,1]^d$ and on the other hand the d-variate Euclidean space $\mathbb{R}^d$. We use the notation

$$\|f\|_p := \|f|L_p(\mathbb{I}^d)\| := \Big(\int_{\mathbb{I}^d}|f(x)|^p\,dx\Big)^{1/p} < \infty,$$

with the usual modification in case $p=\infty$. The space $C(\mathbb{I}^d)$ is often used as a replacement for $L_\infty(\mathbb{I}^d)$. It denotes the collection of all continuous and bounded d-variate functions equipped with the uniform norm. The computation of the Fourier transform (and its inverse) of an L_1-integrable d-variate function is performed by the integrals ($\xi\in\mathbb{R}^d$)

$$\mathcal{F}f(\xi) = \frac{1}{(2\pi)^{d/2}}\int_{\mathbb{R}^d} f(x)e^{-i\xi\cdot x}\,dx,$$
$$\mathcal{F}^{-1}f(\xi) = \frac{1}{(2\pi)^{d/2}}\int_{\mathbb{R}^d} f(x)e^{i\xi\cdot x}\,dx,$$

where $\xi\cdot x := \xi_1x_1+\cdots+\xi_dx_d$. To begin with, we recall the concept of a dyadic decomposition of the unity. The space $C_0^\infty(\mathbb{R})$ consists of all infinitely many times differentiable compactly supported functions.

Definition 3.1. Let $\Phi(\mathbb{R})$ be the collection of all systems $\varphi = \{\varphi_n(x)\}_{n=0}^\infty \subset C_0^\infty(\mathbb{R})$ satisfying

(i) $\operatorname{supp}\varphi_0 \subset \{x : |x|\le 2\}$,

(ii) $\operatorname{supp}\varphi_n \subset \{x : 2^{n-1}\le|x|\le 2^{n+1}\}$, $\quad n=1,2,\ldots,$

(iii) for all $\ell\in\mathbb{N}_0$, it holds $\sup_{x,n} 2^{n\ell}\,|D^\ell\varphi_n(x)| \le c_\ell < \infty$, and

(iv) $\sum_{n=0}^{\infty}\varphi_n(x) = 1$ for all $x\in\mathbb{R}$.

Now we fix a system $\varphi = \{\varphi_n\}_{n\in\mathbb{N}_0} \in \Phi(\mathbb{R})$. for $\ell = (\ell_1,\ldots,\ell_d)\in\mathbb{N}_0^d$, let the building blocks f_ℓ be given by

$$f_\ell(x) = \mathcal{F}^{-1}[\varphi_{\ell_1}(\xi_1)\cdots\varphi_{\ell_d}(\xi_d)\mathcal{F}f(\xi)](x), \quad x\in\mathbb{R}^d. \tag{3.1}$$

Because of the Paley-Wiener theorem, the functions f_ℓ are entire analytic functions and therefore continuous. The goal of this section is to derive bounds for fixed "levels" j of the Faber expansion (2.3) of such a bandlimited function f_ℓ. To be more precise, we aim at bounds for $\|\sum_{k\in\mathbb{Z}^d} d^2_{j,k}(f_\ell)v_{j,k}(\cdot)\|_p$. Clearly, due to the compact support of $v_{j,k}$, we may replace $v_{j,k}$ by the characteristic function $\chi_{j,k}$ of the parallelepiped $[2^{-j_1}k_1, 2^{-j_1}(k_1+1)]\times\cdots\times[2^{-j_d}k_d, 2^{-j_d}(k_d+1)]$. Note that for any continuous function $f\in C(\mathbb{R}^d)$

$$\Big\|\sum_{k\in\mathbb{Z}^d} d^2_{j,k}(f)v_{j,k}(\cdot)\Big\|_p \asymp \Big\|\sum_{k\in\mathbb{Z}^d} d^2_{j,k}(f)\chi_{j,k}(\cdot)\Big\|_p$$
$$\asymp \Big(2^{-|j|_1}\sum_{k\in\mathbb{Z}^d}|d^2_{j,k}(f)|^p\Big)^{1/p},$$

where $0<p\le\infty$. To perform this, we need some tools from harmonic analysis. We state a mixed version of the Peetre maximal inequality, proved in [[20], 1.6.4].

Lemma 3.2. [Peetre maximal inequality] Let $0<p\le\infty$ and $a>1/p$ ($a>0$ in case $p=\infty$). Furthermore, let $f\in L_1(\mathbb{R}^d)$ such that $\operatorname{supp}\mathcal{F}f \subset [-b_1,b_1]\times\cdots\times[-b_d,b_d]$, where $b=(b_1,\ldots,b_d)\in\mathbb{R}^d_+$. Then there is a constant $c>0$, only depending on a and p but not on f and b, such that

$$\Big\|\sup_{y\in\mathbb{R}^d}\frac{|f(x+y)|}{(1+b_1|y_1|)^a\ldots(1+b_d|y_d|)^a}\Big\|_p \le c\,\|f\|_p.$$

The following univariate pointwise estimate connecting differences between bandlimited functions and Peetre maximal operator is taken from [[22], Lemma 3.3.1].

Lemma 3.3. Let $a, b > 0$ and $f \in L_1(\mathbb{R})$ with $\operatorname{supp} \mathcal{F}f \subset [-b, b]$. Then there exists a constant $C > 0$ such that

$$|\Delta_h^m f(x)| \le C \min\{1, |bh|^m\} \max\{1, |bh|^a\} \sup_{y\in\mathbb{R}} \frac{|f(x+y)|}{(1+b|y|)^a}.$$

The bound may be slightly improved when replacing pointwise estimates with estimates involving L_p-norms. We have the following.

Lemma 3.4. Let $j, \ell \in \mathbb{N}_0$, $a > 0$, $0 < p \le \infty$, and $f \in C(\mathbb{R})$. Then we have

$$\begin{aligned}\|\sup_{|h|\le 2^{-j}} |f(x+h)|\|_p &\lesssim 2^{\ell/p} \|\sup_{|h|\le 2^{-(j+\ell)}} |f(x+h)|\|_p \\ &\lesssim 2^{\ell/p} \Big\| \sup_{y\in\mathbb{R}} \frac{f(x+y)}{(1+|2^{j+\ell}y|)^a} \Big\|_p\end{aligned}$$

independent of $x \in \mathbb{R}$, j, ℓ, and f.

Proof. We start with a pointwise estimate

$$\begin{aligned}\sup_{|h|\le 2^{-j}} |f(x+h)| &\le \sup_{|k|\le 2^{\ell}-1} \sup_{|h|\le 2^{-(j+\ell)}} |f(x+k2^{-(j+\ell)}+h)| \\ &\le \Big(\sum_{|k|\le 2^{\ell}-1} \sup_{|h|\le 2^{-(j+\ell)}} |f(x+k2^{-(j+\ell)}+h)|^p \Big)^{1/p}.\end{aligned}$$

Taking L_p-norms on both sides gives

$$\begin{aligned}\|\sup_{|h|\le 2^{-j}} |f(x+h)|\|_p &\le \Big(\sum_{|k|\le 2^{\ell}-1} \int_{\mathbb{R}} \sup_{|h|\le 2^{-(j+\ell)}} |f_{j+\ell}(x+k2^{-(j+\ell)}+h)|^p\, dx \Big)^{1/p} \\ &\lesssim 2^{\ell/p} \|\sup_{|h|\le 2^{-(j+\ell)}} |f(x+h)|\|_p.\end{aligned}$$

Finally, we trivially observe

$$\begin{aligned}\Big\| \sup_{|h|\le 2^{-(j+\ell)}} |f(x+h)| \Big\|_p &= \Big\| \sup_{|h|\le 2^{-(j+\ell)}} \frac{2^a |f(x+h)|}{(1+1)^a} \Big\|_p \\ &\lesssim \Big\| \sup_{|h|\le 2^{-(j+\ell)}} \frac{|f(x+h)|}{(1+2^{j+\ell}|h|)^a} \Big\|_p \qquad (3.2) \\ &\le \Big\| \sup_{y\in\mathbb{R}} \frac{|f(x+y)|}{(1+2^{j+\ell}|y|)^a} \Big\|_p.\end{aligned}$$

In the next lemma, we combine both univariate bounds and derive a multivariate estimate via iteration with respect to coordinate directions. Let $f \in L_1(\mathbb{R}^d)$ and $f_{j+\ell}$ denotes the bandlimited function from (3.1).

Lemma 3.5. Let $0 < p \le \infty$, $j \in \mathbb{N}_{-1}^d$, and $\ell \in \mathbb{Z}^d$. Then we have

$$\Big\| \sum_{k\in\mathbb{Z}^d} d_{j,k}^2(f_{j+\ell})\chi_{j,k}(\cdot) \Big\|_p \lesssim \|f_{j+\ell}\|_p \prod_{i=1}^d \min\{2^{2\ell_i}, 1\} \max\{2^{\ell_i/p}, 1\}.$$

Proof. Step 1. To provide a technically transparent proof of this lemma, we start with the univariate case ($d = 1$). In the second part of this proof, we deal with the multivariate case, which requires more involved notation. For $x \in \mathbb{R}$, we define

$$F_{j,\ell}(x) := \sum_{k\in\mathbb{Z}} d_{j,k}^2(f_{j+\ell})\chi_{j,k}(x).$$

Let $x \in [2^{-j}k, 2^{-j}(k+1)]$. For this x, we have

$$\begin{aligned}|d_{j,k}^2(f_{j+\ell})\chi_{j,k}(x)| &\lesssim |f_{j+\ell}(k\cdot 2^{-j} + 2\cdot 2^{-(j+1)})| \\ &\quad + 2|f_{j+\ell}(k\cdot 2^{-j} + 2^{-(j+1)})| + |f_{j+\ell}(k\cdot 2^{-j})| \\ &\lesssim \sup_{|h|\lesssim 2^{-j}} |f_{j+\ell}(x+h)|.\end{aligned}$$

Since $\chi_{j,k}(x)$ do not overlap, we receive

$$|F_{j,\ell}(x)| = \Big| \sum_{k\in\mathbb{Z}} d_{j,k}^2(f_{j+\ell})\chi_{j,k}(x) \Big| \lesssim \sup_{|h|\lesssim 2^{-j}} |f_{j+\ell}(x+h)|.$$

By Lemma 3.4, we find

$$\|F_{j,\ell}\|_p \lesssim \max\{2^{\ell/p}, 1\} \Big\| \sup_{y\in\mathbb{R}} \frac{f_{j+\ell}(x+y)}{(1+|2^{j+\ell}y|)^a} \Big\|_p \qquad (3.3)$$

for some $a > 0$, which is at our disposal.

In case $\ell \le 0$, we may continue arguing pointwise. First of all, we have

$$|F_{j,\ell}(x)| \le \sup_{|y|\lesssim 2^{-j}} |\Delta_{2^{-j-1}}^2 f_{j+\ell}(x+y)|.$$

Using Lemma 3.3 and the fact that $\ell \le 0$, we obtain

$$|F_{j,\ell}(x)| \le 2^{2\ell} \sup_{y\in\mathbb{R}} \frac{|f(x+y)|}{(1+|2^{j+\ell}y|)^a}. \qquad (3.4)$$

Combining (3.3) and (3.4) gives

$$\|F_{j,\ell}(x)\|_p \lesssim \Big\| \sup_{y\in\mathbb{R}} \frac{|f(x+y)|}{(1+|2^{j+\ell}y|)^a} \Big\|_p \max\{1, 2^{\ell/p}\} \min\{1, 2^{2\ell}\}.$$

Choosing $a > 1/p$ and applying the Peetre maximal inequality in Lemma 3.2 gives the result for $d = 1$.

Step 2. We deal with the multivariate case and start with a pointwise estimate of

$$F_{j,\ell}(x) := \sum_{k\in\mathbb{Z}^d} d_{j,k}^2(f_{j+\ell})\chi_{j,k}(x),$$

where we apply the above procedure in every direction. In order not to drown in notation, we introduce the following direction-wise maximal operator

$$\mathfrak{M}_j^i f(x) := \sup_{|h|\lesssim 2^{-j}} |f(x+e_i h)|,$$

where $e_i = (\underbrace{0,\ldots,0}_{i-1}, 1, \underbrace{0,\ldots,0}_{d-i})$. Clearly, for $x \in \mathbb{R}^d$, we have

$$|F_{j,\ell}(x)| \lesssim \Big(\prod_{i:\ell_i>0} \mathfrak{M}_{j_i}^i \Big) \circ \Big(\prod_{i:\ell_i\le 0} \mathfrak{M}_{j_i}^i \Big) \circ \Big(\prod_{i:\ell_i\le 0} \Delta_{2^{-j_i-1}}^{2,i} \Big) f_{j+\ell}(x).$$

Here we use the fact that we have in every direction

$$|\Delta^{2,i}_{2^{-j-1}} f_{j+\ell}| \le \mathfrak{M}^i_{j_i} f_{j+\ell}(x),$$

including the case $j = -1$, where the difference is replaced by the function value at the respective point. This case is included in $\left(\prod_{i:\ell_i>0} \mathfrak{M}^i_j\right)$ since for $\ell_i < 0$ there is nothing to prove in this case. We use the triangle inequality in order to estimate the difference by point evaluations. Taking sup over the step length of this absolutely valued point evaluations leads to the direction-wise maximal operator. In case $\ell_i \le 0$, we keep the direction-wise difference. Nevertheless, in order to get rid of the characteristic function, we have to additionally apply a direction-wise sup also in this case. Since it holds that

$$\mathfrak{M}^i_{j_i} f(x) \le \mathfrak{M}^i_{j_i} g(x)$$

with $|f| \le |g|$, we first estimate

$$\Big|\Big(\prod_{i:\ell_i\le 0} \Delta^{2,i}_{2^{-j-1}}\Big) f_{j+\ell}(x)\Big|$$

pointwise from above using Lemma 3.3 iteratively. This gives

$$\Big(\prod_{i:\ell_i\le 0} \Delta^{2,i}_{2^{-j-1}}\Big) f_{j+\ell}(x) \lesssim \sup_{\substack{y_i\in\mathbb{R}\\ i:\ell_i\le 0}} \frac{|f_{j+\ell}(x+\sum_{i:\ell_i\le 0} y_i e_i)|}{\prod_{i:\ell_i\le 0}(1+2^{\ell_i+j_i}|y_i|)^a} \prod_{i:\ell_i\le 0} 2^{2\ell_i},$$

for any $a > 0$. The maximal operators $(\prod_{i:\ell_i\le 0} \mathfrak{M}^i_{j_i})$ go on the same coordinates. Since $\ell_i \le 0$ we clearly also have

$$\Big(\prod_{i:\ell_i\le 0} \mathfrak{M}^i_{j_i}\Big) \circ \Big(\prod_{i:\ell_i\le 0} \Delta^{2,i}_{2^{-j-1}}\Big) f_{j+\ell}(x) \lesssim \sup_{\substack{y_i\in\mathbb{R}\\ i:\ell_i\le 0}} \frac{|f_{j+\ell}(x+\sum_{i:\ell_i\le 0} y_i e_i)|}{\prod_{i:\ell_i\le 0}(1+2^{\ell_i+j_i}|y_i|)^a} \prod_{i:\ell_i\le 0} 2^{2\ell_i}.$$

It remains to apply $(\prod_{i:\ell_i>0} \mathfrak{M}^i_{j_i})$ and take the L_p-norm, $p < \infty$. Here we use Lemma 3.4 iteratively:

$$\begin{aligned}
&\int_{\mathbb{R}^d} |F_{j,\ell}(x)|^p\,dx \\
&\lesssim \int_{\{(x_i)_{i:\ell_i\le 0}\}}\cdots\int \int_{\{(x_i)_{i:\ell_i>0}\}}\cdots\int \Big|\Big(\prod_{i:\ell_i>0}\mathfrak{M}^i_{j_i}\Big)\circ\Big(\prod_{i:\ell_i\le 0}\mathfrak{M}^i_{j_i}\Big) \\
&\circ \Big(\prod_{i:\ell_i\le 0}\Delta^{2,i}_{2^{-j-1}}\Big) f_{j+\ell}(x)\Big|^p \prod_{i:\ell_i>0} dx_i \prod_{i:\ell_i\le 0} dx_i \\
&\lesssim \prod_{i=1}^d \min\{2^{2p\ell_i},1\}\max\{2^{\ell_i},1\}\times \\
&\times \int_{\{(x_i)_{i:\ell_i\le 0}\}}\cdots\int\int_{\{(x_i)_{i:\ell_i>0}\}}\cdots\int \Big(\Big(\prod_{i:\ell_i>0}\mathfrak{M}^i_{j_i+\ell_i}\Big)\sup_{\substack{y_i\in\mathbb{R}\\ i:\ell_i\le 0}} \frac{|f_{j+\ell}(x+\sum_{i:\ell_i\le 0} y_i e_i)|}{\prod_{i:\ell_i\le 0}(1+2^{\ell_i+j_i}|y_i|)^a}\Big)^p \\
&\prod_{i:\ell_i>0} dx_i \prod_{i:\ell_i\le 0} dx_i\,.
\end{aligned}$$

Finally, we use the trivial estimate [already known from (3.2)] to replace the operators $\mathfrak{M}^i_{j_i+\ell_i}$ by the Peetre maximal function. This gives

$$\begin{aligned}
\int_{\mathbb{R}^d} |F_{j,\ell}(x)|^p\,dx &\lesssim \prod_{i=1}^d \min\{2^{2p\ell_i},1\}\max\{2^{\ell_i},1\} \\
&\int_{\mathbb{R}^d}\Big(\sup_{y\in\mathbb{R}^d}\frac{|f_{j+\ell}(x+y)|}{\prod_{i=1}^d(1+2^{\ell_i+j_i}|y_i|)^a}\Big)^p dx \\
&\lesssim \prod_{i=1}^d \min\{2^{2p\ell_i},1\}\max\{2^{\ell_i},1\}\|f_{j+\ell}\|_p^p\,,
\end{aligned}$$

if we choose $a > 1/p$. For the sake of completeness, let us additionally consider the case $p = \infty$. We have

$$\begin{aligned}
\sup_{x\in\mathbb{R}^d}|F_{j,\ell}(x)| &\lesssim \sup_{x\in\mathbb{R}^d}\Big|\Big(\prod_{i:\ell_i>0}\mathfrak{M}^i_{j_i}\Big)\circ\Big(\prod_{i:\ell_i\le 0}\mathfrak{M}^i_{j_i}\Big) \\
&\circ\Big(\prod_{i:\ell_i\le 0}\Delta^{2,i}_{2^{-j-1}}\Big) f_{j+\ell}(x)\Big| \\
&\lesssim \prod_{i=1}^d \min\{2^{2\ell_i},1\}\sup_{x\in\mathbb{R}^d}\sup_{y\in\mathbb{R}^d}\frac{|f_{j+\ell}(x+y)|}{\prod_{i=1}^d(1+2^{\ell_i+j_i}|y_i|)^a} \\
&\lesssim \prod_{i=1}^d \min\{2^{2\ell_i},1\}\|f_{j+\ell}\|_\infty\,,
\end{aligned}$$

if we choose $a > 0$.

4. Besov and Triebel-Lizorkin spaces with mixed smoothness

For the definition of the corresponding function spaces on $\mathbb{R}^d$, we refer to [1, 20, 23]. The corresponding function spaces on $[0,1]^d$ are defined via restrictions of functions on $\mathbb{R}^d$ [see [1], Section 3.4]. In this section, we mainly focus on the definition of Besov and Triebel-Lizorkin spaces with dominating mixed (in the sequel only called mixed) smoothness on $\mathbb{R}^d$ since they are crucial for our subsequent analysis. We closely follow [[20], Chapter 2] and use the building blocks $f_j(\cdot)$ defined in (3.1).

Definition 4.1. [Mixed Besov and Triebel-Lizorkin spaces] (i) Let $0 < p, \theta \le \infty$, and $r > (1/p - 1)_+$. If $\theta \le \min\{p, 1\}$, we admit $r = (1/p - 1)_+$. Then $S^r_{p,\theta}B(\mathbb{R}^d)$ is defined as the collection of all $f \in L_{\max\{p,1\}}(\mathbb{R}^d)$ such that

$$\|f|S^r_{p,\theta}B(\mathbb{R}^d)\|^\varphi := \Big(\sum_{j\in\mathbb{N}_0^d} 2^{|j|_1 r\theta}\|f_j\|_p^\theta\Big)^{1/\theta}$$

is finite (with the usual modification if $\theta = \infty$).
(ii) Let $0 < p < \infty$, $0 < \theta \le \infty$, and $r > (1/p - 1)_+$. Then $S^r_{p,\theta}F(\mathbb{R}^d)$ is defined as the collection of all $f \in L_{\max\{p,1\}}(\mathbb{R}^d)$ such that

$$\|f|S^r_{p,\theta}F(\mathbb{R}^d)\|^\varphi := \Big\|\Big(\sum_{j\in\mathbb{N}_0^d} 2^{|j|_1 r\theta}|f_j(x)|^\theta\Big)^{1/\theta}\Big\|_p$$

is finite (with the usual modification if $\theta = \infty$).

It is noted that this definition is independent of the chosen system φ in the context of equivalent (quasi-)norms. Moreover, in the case $\min\{p,\theta\} \geq 1$, the defined spaces are Banach spaces, whereas they are quasi-Banach spaces in the case $\min\{p,\theta\} < 1$. For details, we refer to [[20], Section 2.2.4]. In the next lemma, there appears the condition $r > (1/p - 1)_+$, which is caused by the parameter range in our Definition 4.1. All the subsequent embeddings, of course, also hold true for the general situation, where $r \in \mathbb{R}$. We have the following elementary embeddings [see [20], Section 2.2.3].

Lemma 4.2. Let $0 < p < \infty$ (including $p = \infty$ in the B-case), $0 < \theta \leq \infty$, and $r > (1/p - 1)_+$. If $\theta \leq \min\{p, 1\}$, we admit $r = (1/p - 1)_+$ in the B-case. Furthermore, let $A \in \{B, F\}$.

(i) If $\varepsilon > 0$ and $0 < v \leq \infty$, then

$$S^{r+\varepsilon}_{p,\theta} A(\mathbb{R}^d) \hookrightarrow S^r_{p,v} A(\mathbb{R}^d) .$$

(iia) If $p < u < \infty$, $0 < \theta, \theta_1, \theta_2 \leq \infty$, and $r - 1/p = t - 1/u$, then

$$S^r_{p,\theta} B(\mathbb{R}^d) \hookrightarrow S^t_{u,\theta} B(\mathbb{R}^d),$$
$$S^r_{p,\theta_1} F(\mathbb{R}^d) \hookrightarrow S^t_{u,\theta_2} F(\mathbb{R}^d) .$$

(iib) (Jawerth-Franke embedding I) If $0 < p < u < \infty$, $0 < w \leq \infty$, and $r - 1/p = t - 1/u$, then

$$S^r_{p,u} B(\mathbb{R}^d) \hookrightarrow S^t_{u,w} F(\mathbb{R}^d) .$$

(iic) (Jawerth-Franke embedding II) If $0 < p < u \leq \infty$, $0 < \theta \leq \infty$, and $r - 1/p = t - 1/u$, then

$$S^r_{p,\theta} F(\mathbb{R}^d) \hookrightarrow S^t_{u,p} B(\mathbb{R}^d) .$$

(iiia) If $r > 1/p$ (including $p = \infty$, $r > 0$), then

$$S^r_{p,\theta} B(\mathbb{R}^d) \hookrightarrow C(\mathbb{R}^d) .$$

(iiib) If $r = 1/p$ and $\theta \leq 1$, then we still have

$$S^{1/p}_{p,\theta} B(\mathbb{R}^d) \hookrightarrow C(\mathbb{R}^d) ,$$

and especially the limiting case

$$S^0_{\infty,\theta} B(\mathbb{R}^d) \hookrightarrow C(\mathbb{R}^d) .$$

(iiic) It holds

$$S^1_{1,\infty} F(\mathbb{R}^d) \hookrightarrow C(\mathbb{R}^d) .$$

(iv) If $1 < p < \infty$ and $r > 0$, then

$$S^r_{p,\min\{p,\theta\}} B(\mathbb{R}^d) \hookrightarrow S^r_{p,\theta} F(\mathbb{R}^d) \hookrightarrow S^r_{p,\max\{p,\theta\}} B(\mathbb{R}^d) .$$

Proof. The embeddings (i), (iia), (iiia), (iiib), and (iv) are standard and can be found in [[20], Chapter 2] especially we refer to [[20], Remark 2, p. 132], which includes the limiting case $p = \infty$ in (iiib). As for the Jawerth-Franke type embeddings, we refer to [[24], Theorem 1.2 and 1.4] and the summary [[6], Lemma 3.4.2 and 3.4.3]. See also [[6], Rem. 3.4.4] for further references, especially for the mixed smoothness case. Finally, the embedding in (iiic) is a consequence of (iic) and (iiib).

4.1. Spaces on domains

We aim for approximating functions defined on the unit cube $[0,1]^d$ with the above regularity assumptions. This requires the definition of function spaces on domains. The domain $\Omega \subset \mathbb{R}^d$ represents an open connected set. Later, when dealing with continuous bounded functions, we may use as well compact sets like $[0,1]^d$.

Definition 4.3. Let Ω be a domain in $\mathbb{R}^d$.

1. Furthermore, let $0 < p, \theta \leq \infty$, and $r > (1/p - 1)_+$. If in case $\theta \leq \min\{p, 1\}$, we admit $r = (1/p - 1)_+$. Then we define

$$S^r_{p,\theta} B(\Omega) := \{f \in L_{\max\{p,1\}}(\Omega) : \exists g \in S^r_{p,\theta} B(\mathbb{R}^d) \text{ with } g|_\Omega = f\} ,$$

where

$$\|f|S^r_{p,\theta} B(\Omega)\| := \inf\{\|g|S^r_{p,\theta} B(\mathbb{R}^d)\| : g \in S^r_{p,\theta} B(\mathbb{R}^d),\ g|_\Omega = f\}.$$

2. In case $0 < p < \infty$, $0 < \theta \leq \infty$, and $r > (1/p - 1)_+$, we define

$$S^r_{p,\theta} F(\Omega) := \{f \in L_{\max\{p,1\}}(\Omega) : \exists g \in S^r_{p,\theta} F(\mathbb{R}^d) \text{ with } g|_\Omega = f\} ,$$

where

$$\|f|S^r_{p,\theta} F(\Omega)\| := \inf\{\|g|S^r_{p,\theta} F(\mathbb{R}^d)\| : g \in S^r_{p,\theta} F(\mathbb{R}^d),\ g|_\Omega = f\}.$$

On bounded domains Ω, all the embeddings in Lemma 4.2 keep valid. In addition, we have the following embeddings. If $0 < p_2 < p_1 \leq \infty$ (F-case: $p_i < \infty$), $0 < \theta \leq \infty$, and $|\Omega| < \infty$, then

$$S^r_{p_1,\theta} F(\Omega) \hookrightarrow S^r_{p_2,\theta} F(\Omega),$$

and

$$S^r_{p_1,\theta} B(\Omega) \hookrightarrow S^r_{p_2,\theta} B(\Omega).$$

Clearly, this is a trivial consequence of the embedding

$$L_{p_1}(\Omega) \hookrightarrow L_{p_2}(\Omega).$$

It is well-known that spaces $S^r_{p,\theta} B([0,1]^d)$ with sufficiently large smoothness, namely $r > 1/p$, are compactly embedded into $L_\infty([0,1]^d)$. This is a direct consequence of results on entropy numbers of the classes $S^r_{\infty,\theta} B([0,1]^d)$ in $L_\infty([0,1]^d)$ [see [25], Cor. 23, (iii) or [12], Theorem 6.2], and the embeddings stated in Lemma 4.2 above.

However, in case $r = 1/p$, we do not have a compact embedding. For the convenience of the reader, we give a direct proof.

Lemma 4.4. [Non-compactness of limiting embeddings] (i) Let $0 < p \leq \infty$ and $\theta \leq \min\{p, 1\}$. Then the embedding

$$S^{1/p}_{p,\theta}B([0,1]^d) \hookrightarrow L_\infty([0,1]^d) \qquad (4.1)$$

is not compact.
(ii) If $p \leq 1$ and $0 < \theta \leq \infty$, then the embedding

$$S^{1/p}_{p,\theta}F([0,1]^d) \hookrightarrow L_\infty([0,1]^d) \qquad (4.2)$$

is not compact.

Proof. We show the non-compactness of the embedding (4.1) first in the case $d = 1$. Clearly, by standard (tensorization) arguments, this would also imply the non-compactness in higher dimensions. Note further that the non-compactness of (4.2) is implied by the Jawerth-Franke embedding [see Lemma 4.2, (iib)]

$$S^{1/p_1}_{p_1,p_2}B([0,1]^d) \hookrightarrow S^{1/p_2}_{p_2,\theta}F([0,1]^d)$$

together with the non-compactness of (4.1). So, it remains to be proved the non-compactness of $B^{1/p}_{p,\theta}([0,1]) \hookrightarrow L_\infty([0,1])$ for any $\theta \leq \min\{p, 1\}$ and $0 < p \leq \infty$. This fact is certainly known and can be found in the literature. However, we would like to present a direct argument here since it fits the scope of the article. The idea is to find a sequence $(g_j)_j$ of functions with $\|g_j|B^{1/p}_{p,\theta}([0,1])\| \leq 1$ for all j, which is not convergent in $L_\infty([0,1])$. A straightforward choice for the g_j is the Faber basis functions (L_∞-normalized hat functions) for different levels j. To be more precise, we let $j = 0, 1, 2, \ldots$ and consider the sequence $(g_j)_j := (v_{j,0})_{j\in\mathbb{N}_0}$. Clearly, we always have

$$\sup_{|h|\leq 2^{-k}} \|\Delta^m_h v_{j,0}\|_p \lesssim 2^{-j/p}\,.$$

In case $k > j$, we even obtain (due to cancellation)

$$\sup_{|h|\leq 2^{-k}} \|\Delta^m_h v_{j,0}\|_p \lesssim 2^{-k/p}2^{j-k},$$

see, for instance [[2], (3.19) and (3.20)]. According to [[26], p. 110] we can equivalently describe the norm of $B^s_{p,\theta}(\mathbb{R})$ in terms of differences. Note that here we have to use differences of sufficiently high order $m > s$ since $1/p$ may get large. The estimate from above implies

$$\begin{aligned}\|v_{j,0}|B^s_{p,\theta}([0,1])\| &\asymp \|v_{j,0}\|_p + \Big(\sum_{k=0}^{\infty}[2^{ks}\sup_{|h|\leq 2^{-k}}\|\Delta^m_h v_{j,0}\|_p]^\theta\Big)^{1/\theta}\\ &\asymp 2^{-j/p} + \Big(\sum_{k\leq j}[2^{ks}2^{-j/p}]^\theta + \sum_{k>j}[2^{ks}2^{-k/p}2^{j-k}]^\theta\Big)^{1/\theta}\\ &\asymp 1\end{aligned}$$

in case $s \leq 1/p$. Hence, the elements $v_{j,0}$ have uniformly (in j) bounded quasi-norm in $B^s_{p,\theta}([0,1])$. However, it holds

$$\|g_j - g_\ell\|_\infty \geq 1,$$

if $j \neq \ell$, see, for instance, Figure 1. This directly disproves the compactness of the unit ball of $B^{1/p}_{p,\theta}([0,1])$ in $L_\infty([0,1])$.

Remark 4.5. We need to formulate the definition of the Kolmogorov widths to demonstrate some relations with approximative characteristics considered here. For a compact set $\mathbf{F} \hookrightarrow X$ of a Banach space X, we define the Kolmogorov widths as follows:

$$d_m(\mathbf{F}, X) := \inf_{\{u_i\}_{i=1}^m\subset X}\sup_{f\in\mathbf{F}}\inf_{c_i}\Big\|f - \sum_{i=1}^m c_i\,u_i\Big|X\Big\|, \quad m = 1, 2, \ldots,$$

and

$$d_0(\mathbf{F}, X) := \sup_{f\in\mathbf{F}}\|f|X\|.$$

Clearly, considering $\mathbf{F}$ as the unit ball in $S^{1/p}_{p,\theta}B([0,1]^d)$ with $\theta \leq \min\{p, 1\}$ and $X = L_\infty([0,1]^d)$ then

$$d_m(S^{1/p}_{p,\theta}B([0,1]^d), L_\infty([0,1]^d)) \nrightarrow 0$$

by Lemma 4.4.

5. The decay of the Faber coefficients

Now we are ready for proving our main tool: an assertion about the decay of the Faber coefficients of functions from $S^{1/p}_{p,\min\{p,1\}}B(\mathbb{R}^d)$. In order to do so we need to define the following space of doubly indexed sequences.

Definition 5.1. Let $0 < p, \theta \leq \infty$ and $r \geq 1/p$.
(i) The sequence space $s^r_{p,\theta}b$ is the collection of all doubly indexed sequences $\{\lambda_{j,k}\}_{j\in\mathbb{N}^d_{-1},k\in\mathbb{Z}^d}$ such that

$$\big\|\lambda_{j,k}|s^r_{p,\theta}b\big\| := \Big[\sum_{j\in\mathbb{N}^d_{-1}} 2^{|j|_1(r-1/p)\theta}\Big(\sum_{k\in\mathbb{Z}^d}|\lambda_{j,k}|^p\Big)^{\theta/p}\Big]^{1/\theta}$$

is finite (with the usual modification if $\max\{p,\theta\} = \infty$).
(ii) Let $\Omega \subset \mathbb{R}^d$ be a compact domain. We define the index set $\mathbb{D}_j(\Omega)$ to be the set of all $k \in \mathbb{Z}^d$ such that $x_{j,k} \in \Omega$. The space $s^r_{p,\theta}b(\Omega)$ is defined as the space of all doubly indexed sequences $\{\lambda_{j,k}\}_{j\in\mathbb{N}^d_{-1},k\in\mathbb{D}_j(\Omega)}$ such that

$$\big\|\lambda_{j,k}|s^r_{p,\theta}b(\Omega)\big\| := \Big[\sum_{j\in\mathbb{N}^d_{-1}} 2^{|j|_1(r-1/p)\theta}\Big(\sum_{k\in\mathbb{D}_j(\Omega)}|\lambda_{j,k}|^p\Big)^{\theta/p}\Big]^{1/\theta}$$

is finite (with the usual modification if $\max\{p,\theta\} = \infty$).

Remark 5.2. (i) In case $\Omega = [0,1]^d$ we have $\mathbb{D}_j(\Omega) = \mathbb{D}_j$, which was defined right before (2.4). In a certain sense the elements of $s^r_{p,\theta}b(\Omega)$ are restrictions of elements in $s^r_{p,\theta}b$ to indices related to Ω.
(ii) These sequence spaces already appeared in [[23], Def. 2.1, 3.2]. Let $0 < p, \theta \leq \infty$ and $r \in \mathbb{R}$. The spaces $s^r_{p,\theta}b$ and $s^r_{p,\theta}b(\Omega)$ are Banach spaces if $\min\{p,\theta\} \geq 1$. In case $\min\{p,\theta\} < 1$ the space $s^r_{p,\theta}b$ is a quasi-Banach space. Moreover, if $u := \min\{p, \theta, 1\}$ it is a u-Banach space, i.e.,

$$\|\lambda + \mu|s^r_{p,\theta}b\|^u \leq \|\lambda|s^r_{p,\theta}b\|^u + \|\mu|s^r_{p,\theta}b\|^u, \quad \lambda, \mu \in s^r_{p,\theta}b\,.$$

Here is the first main result.

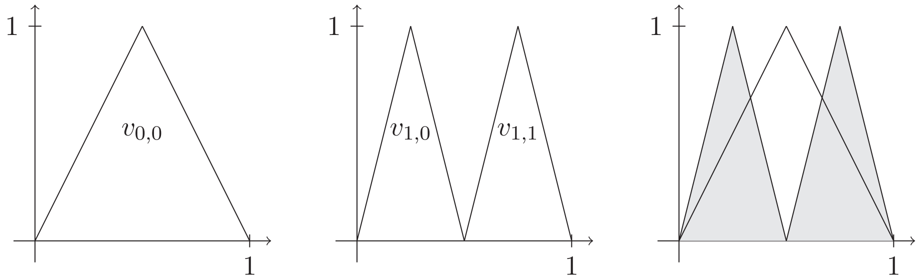

FIGURE 1
Univariate hierarchical Faber basis on $[0, 1]$ for levels $j \in \{0, 1\}$ and their union.

Proposition 5.3. Let $1/2 \le p \le \infty$. Then there exists a constant $c > 0$ (independent of f) such that

$$\|d^2_{j,k}(f)|s^{1/p}_{p,\infty}b\| \le c\,\|f|S^{1/p}_{p,\min\{p,1\}}B(\mathbb{R}^d)\| \tag{5.1}$$

for all $f \in S^{1/p}_{p,\min\{p,1\}}B(\mathbb{R}^d)$.

Proof. Let us put $u = \min\{p, 1\}$. We make use of the decomposition (2.3) in a slightly modified way. For fixed $j \in \mathbb{N}^d_{-1}$ we write $f = \sum_{\ell\in\mathbb{Z}^d} f_{j+\ell}$. Putting this into (5.1) and using the u-triangle inequality yields

$$\begin{aligned}\|d^2_{j,k}(f)|s^{1/p}_{p,\infty}b\|^u &\asymp \sup_{j\in\mathbb{N}^d_{-1}} 2^{u|j|_1/p}\Big\|\sum_{k\in\mathbb{Z}^d} d^2_{j,k}(f)\chi_{j,k}(x)\Big\|^u_p\\ &\le \sup_{j\in\mathbb{N}^d_{-1}} 2^{u|j|_1/p}\sum_{\ell\in\mathbb{Z}^d}\Big\|\sum_{k\in\mathbb{Z}^d} d^2_{j,k}(f_{j+\ell})\chi_{j,k}(x)\Big\|^u_p.\end{aligned}$$

Applying Lemma 3.5 we obtain

$$\begin{aligned}\|d^2_{j,k}(f)|s^{1/p}_{p,\infty}b\|^u &\lesssim \sup_{j\in\mathbb{N}^d_{-1}}\sum_{\ell\in\mathbb{Z}^d} 2^{u|j+\ell|_1/p}\|f_{j+\ell}\|^u_p\\ &\quad \prod_{i=1}^d 2^{-u\ell_i/p}\min\{2^{2u\ell_i},1\}\max\{2^{u\ell_i/p},1\}\\ &\lesssim \Big(\sum_{\ell\in\mathbb{Z}^d} 2^{u|j+\ell|_1/p}\|f_{j+\ell}\|^u_p\Big)\\ &\quad \sup_{\ell\in\mathbb{Z}^d}\prod_{i=1}^d 2^{-u\ell_i/p}\min\{2^{2u\ell_i},1\}\max\{2^{u\ell_i/p},1\}\\ &\lesssim \|f|S^{1/p}_{p,u}B(\mathbb{R}^d)\|^u,\end{aligned}$$

since the sup stays finite since $1/2 \le p \le \infty$.

As a direct consequence we have the following restricted version.

Corollary 5.4. Let $1/2 \le p \le \infty$. Then there exists a constant $c > 0$ such that

$$\|d^2_{j,k}(f)|s^{1/p}_{p,\infty}b(\Omega)\| \le c\,\|f|S^{1/p}_{p,\min\{p,1\}}B(\Omega)\|$$

for all $f \in S^{1/p}_{p,\min\{p,1\}}B(\Omega)$.

Proof. Assume $f \in S^{1/p}_{p,\min\{p,1\}}B(\Omega)$. Then, by Definition 4.3 there is a $g \in S^{1/p}_{p,\min\{p,1\}}B(\mathbb{R}^d)$ with $g|_\Omega = f$. By Proposition 5.3 we obtain

$$\|d^2_{j,k}(g)|s^{1/p}_{p,\infty}b\| \le c\,\|g|S^{1/p}_{p,\min\{p,1\}}B(\mathbb{R}^d)\|.$$

Since $d^2_{j,k}(g) = d^2_{j,k}(f)$ for $j \in \mathbb{N}^d_{-1}$ and $k \in \mathbb{D}_j(\Omega)$ we have

$$\|d^2_{j,k}(f)|s^{1/p}_{p,\infty}b(\Omega)\| \le \|d^2_{j,k}(g)|s^{1/p}_{p,\infty}b\|$$

we obtain

$$\|d^2_{j,k}(f)|s^{1/p}_{p,\infty}b(\Omega)\| \le c\|g|S^{1/p}_{p,\min\{p,1\}}B(\mathbb{R}^d)\|.$$

The last inequality holds for every extension g of f. Taking the infimum yields the result.

The last corollary can be interpreted as a generalization of [[2], Proposition 3.4] to the limiting smoothness case $r = 1/p$. Related results concerning non-limiting smoothness were obtained in [1, 5, 27].

6. Application for sampling recovery in L_q with $q < \infty$

In this section, we would like to apply the Faber embedding in Corollary 5.4 for sampling recovery on the unit cube $\Omega = [0, 1]^d$. As we have shown in Lemma 4.4, we cannot expect an error decay of a sampling recovery operator in the worst case when we measure the error in $L_\infty([0, 1]^d)$. Hence, we focus on the recovery in $L_q([0, 1]^d)$ with $q < \infty$. Based on the Faber representation, we will use a sparse grid truncation in order to obtain a recovery operator. Let

$$I_n f = \sum_{|j|_1\le n}\sum_{k\in\mathbb{D}_j} d^2_{j,k}(f)v_{j,k} \tag{6.1}$$

with the notation from Section 2.

Lemma 6.1. The following estimates hold true for $\alpha > 0$

(i)

$$\sum_{|j|_1>n} 2^{-\alpha|j|_1} \asymp 2^{-\alpha n}n^{d-1} \quad \textit{and}$$

(ii)

$$\sum_{|j|_1 \le n} 2^{|j|_1} \asymp 2^n n^{d-1}.$$

Proof. We refer to [[28], p. 10, Lemma D].

We first consider the case where $p = q$.

Theorem 6.1. Let $1/2 \le p < \infty$. Then there is a constant $C > 0$ (independent of n and f) such that

$$\|f - I_n f\|_p \le C\, 2^{-n/p} n^{(d-1)/\min\{p,1\}} \|f|S^{1/p}_{p,\min\{p,1\}}B([0,1]^d)\|$$

holds for all $n \in \mathbb{N}$.

Proof. The representation in (2.4) allows us to express and estimate the error by $u = \min\{p,1\}$

$$\begin{aligned}\|f - I_n f\|_p^u &= \Big\| \sum_{|j|_1 > n} \sum_{k \in \mathbb{D}_j} d^2_{j,k}(f) v_{j,k} \Big\|_p^u \\ &\le \sum_{|j|_1 > n} \Big\| \sum_{k \in \mathbb{D}_j} d^2_{j,k}(f) v_{j,k} \Big\|_p^u 2^{u|j|_1/p}\, 2^{-u|j|_1/p} \\ &\le \Big(\sup_j 2^{|j|_1/p} \Big\| \sum_{k \in \mathbb{D}_j} d^2_{j,k}(f) v_{j,k} \Big\|_p \Big)^u \sum_{|j|_1 > n} 2^{-u|j|_1/p}.\end{aligned}$$

Finally applying Proposition 5.3 or Corollary 5.4 yields

$$\begin{aligned}\|f - I_n f\|_p^u &\lesssim \|d^2_{j,k}(f)|s^{1/p}_{p,\infty}b([0,1]^d)\|^u \sum_{|j|_1 > n} 2^{-u|j|_1/p} \\ &\lesssim 2^{-un/p} n^{d-1} \|f|S^{1/p}_{p,\min\{p,1\}}B([0,1]^d)\|^u,\end{aligned}$$

where the sum is estimated by Lemma 6.1.

Remark 1. A one-dimensional version of the above result has been proven by Oswald [19] about four decades ago. Dũng proved in [[17], Theorem 3.2] that a related result for slightly different Besov-type function spaces, especially not including the case $p = 1/2$.

In the situation $p < q$, we loose in the main rate. This phenomenon has been observed earlier in the literature [see [6]]. We will use Jawerth-Franke type embeddings to improve the order of the logarithmic term.

Theorem 6.2. Let $1/2 \le p < q < \infty$. Then

$$\|f - I_n f\|_q \le C\, 2^{-n/q} n^{(d-1)/q} \|f|S^{1/p}_{p,\min\{p,1\}}B([0,1]^d)\|.$$

Proof. Using Jawerth-Franke I (see Lemma 4.2), we obtain

$$\begin{aligned}\|f - I_n f\|_q &\lesssim \|f - I_n f|S^{1/p-1/q}_{p,q}B([0,1]^d)\| \\ &\lesssim \Big(\sum_{|j|_1 > n} 2^{q|j|_1[1/p-1/q]} \Big\| \sum_{k \in \mathbb{D}_j} d^2_{j,k}(f) \chi_{j,k} \Big\|_p^q \Big)^{1/q} \\ &\lesssim \Big[\sup_{j \in \mathbb{N}^d_{-1}} 2^{|j|_1/p} \Big\| \sum_{k \in \mathbb{D}_j} d^2_{j,k}(f) \chi_{j,k} \Big\|_p \Big] \Big(\sum_{|j|_1 > n} 2^{-|j|_1} \Big)^{1/q},\end{aligned}$$

where we used the inverse Faber characterization for spaces with positive smoothness [cf. [1], Theorem 4.18]. This reference deals with the $\mathbb{R}^d$ case but can be easily extended by standard arguments as shown, for instance, in the proof of Corollary 5.4 or following the arguments in [[1], Theorem 4.25] to the unit cube setting. Finally applying Proposition 5.3 yields

$$\|f - I_n f\|_q \lesssim 2^{-n/q} n^{(d-1)/q} \|f|S^{1/p}_{p,\min\{p,1\}}B([0,1]^d)\|.$$

Let us now deal with the space $S^1_{1,\infty}F([0,1]^d)$, which is embedded into $C([0,1]^d)$, as shown in Lemma 4.2, (iiic). By Jawerth-Franke embedding II [(Lemma 4.2, (iic)], we even know that for every $1 < p < \infty$, we have $S^1_{1,\infty}F([0,1]^d) \hookrightarrow S^{1/p}_{p,1}B([0,1]^d)$. As a direct corollary of Theorems 6.1 and 6.2, we make the new observations.

Theorem 6.3. (i) It holds for any small $\varepsilon > 0$

$$\|f - I_n f\|_1 \lesssim_\varepsilon 2^{-n(1-\varepsilon)} \|f|S^1_{1,\infty}F([0,1]^d)\|.$$

(ii) For any $1 < q < \infty$ we have

$$\|f - I_n f\|_q \lesssim 2^{-n/q} n^{(d-1)/q} \|f|S^1_{1,\infty}F([0,1]^d)\|.$$

Proof. By the embedding $S^1_{1,\infty}F([0,1]^d) \hookrightarrow S^{1/p}_{p,1}B([0,1]^d)$ for $p > 1$, we may apply Theorem 6.2 to obtain the result. For (i), we simply choose small $q > 1$ and use the fact that the L_q-norm dominates the L_1-norm. Clearly, the logterm can be dropped in this regime.

6.1. Sampling widths

Let us focus our considerations toward the problem of optimal sampling recovery. We compare the number m of samples to the resulting error in an algorithm and call the quantity

$$\begin{aligned}&\varrho_m(S^r_{p,\theta}A([0,1]^d), L_q([0,1]^d)) \\ &:= \inf_{\substack{X_m \subset [0,1]^d, |X_m| = m \\ \varphi : \mathbb{C}^m \to L_q([0,1]^d)}} \sup_{\|f|S^r_{p,\theta}A([0,1]^d)\| \le 1} \big\| f - \varphi(f(X_m)) \big| L_q([0,1]^d) \big\|,\end{aligned}$$

the m-th sampling width. If, in addition, the mapping $\varphi : \mathbb{C}^n \to L_q([0,1]^d)$ is linear, then we obtain the linear sampling widths

$$\begin{aligned}&\varrho^{lin}_m(S^r_{p,\theta}A([0,1]^d), L_q([0,1]^d)) \\ &:= \inf_{\substack{X_m \subset [0,1]^d, |X_m| = m \\ \varphi : \mathbb{C}^m \to L_q([0,1]^d) \\ linear}} \sup_{\|f|S^r_{p,\theta}A([0,1]^d)\| \le 1} \big\| f - \varphi(f(X_m)) \big| L_q([0,1]^d) \big\|,\end{aligned}$$

where $A \in \{B, F\}$. Therefore, according to the definitions mentioned above, we have

$$\varrho_m(S^r_{p,\theta}A([0,1]^d), L_q([0,1]^d)) \le \varrho^{lin}_m(S^r_{p,\theta}A([0,1]^d), L_q([0,1]^d)).$$

In the next theorems, we apply the linear algorithm $I_n f$ [cf. (6.1)] to obtain upper bounds for ϱ^{lin}_m and ϱ_m.

Theorem 6.4. Let $1/2 \le p < \infty$. Then

$$\varrho^{lin}_m(S^{1/p}_{p,\min\{p,1\}}B([0,1]^d), L_p([0,1]^d)) \lesssim m^{-1/p} (\log^{d-1} m)^{1/p + 1/\min\{p,1\}}$$

for all $m \in \mathbb{N}$.

Proof. The upper bound is due to Theorem 6.1 recognizing that the algorithm I_n in (6.1) samples f in $m \asymp 2^n n^{d-1}$ nodes. This can be trivially checked by applying Lemma 6.1, (ii). For the sake of completeness, we refer to [[1], Section 5.1] where further properties of this operator were studied.

Theorem 6.5. Let $1/2 \le p < q < \infty$. Then

(i)

$$\varrho_m^{lin}(S^{1/p}_{p,\min\{p,1\}}B([0,1]^d), L_q([0,1]^d)) \lesssim m^{-1/q}(\log^{d-1} m)^{2/q}$$

and

(ii)

$$\varrho_m^{lin}(S^1_{1,\infty}F([0,1]^d), L_q([0,1]^d)) \lesssim m^{-1/q}(\log^{d-1} m)^{2/q}$$

for all $m \in \mathbb{N}$.

Proof. The upper bound in (i) is due to Theorem 6.2 taking the number of sampling nodes into account, see Theorem 6.4. The upper bound in (ii) is due to Theorem 6.3.

6.2. Lower bounds

The linear width of class $\mathbf{F}$ in a normed space X has been introduced by Tikhomirov [29] more than 60 years ago. It is defined by

$$\lambda_m(\mathbf{F}, X) := \inf_{\substack{A:\,X\to X \text{ linear}\\ \operatorname{rank} A \le m}} \sup_{f\in\mathbf{F}} \left\| f - A(f) | X \right\| .$$

Romanyuk [30, 31] proved for $\mathbf{F}$, the unit ball in $S^r_{p,1}B([0,1]^d)$, that in case $1 \le p \le q \le 2$ and $q > 1$

$$\begin{aligned}\lambda_m(S^{1/p}_{p,\theta}B([0,1]^d), L_q([0,1]^d)) &\ge d_m(S^{1/p}_{p,\theta}B([0,1]^d), L_q([0,1]^d))\\ &\gtrsim (m^{-1}\log^{d-1} m)^{1/q} .\end{aligned} \tag{6.2}$$

We obtain the following lower bounds for the (linear) sampling widths.

Theorem 6.6. (i) Let $1 \le p \le q < \infty$. Then we have

$$\varrho_m(S^{1/p}_{p,1}B([0,1]^d), L_q([0,1]^d)) \gtrsim m^{-1/q} .$$

(ii) If, additionally, $1 \le p \le q \le 2$ and $q > 1$, then

$$\varrho_m^{lin}(S^{1/p}_{p,1}B([0,1]^d), L_q([0,1]^d)) \gtrsim (m^{-1}\log^{d-1} m)^{1/q} .$$

Proof. The result in (ii) follows immediately from (6.2). For the lower bound in (i), we use a fooling function that is constructed as the simple tensor function, containing the univariate fooling function obtained from [32] in the first direction smooth compactly supported (bump) functions in all remaining directions. We obtain our result by considering that all corresponding norms have product properties related to simple tensor functions.

7. Outlook and discussion

In this article, we have shown that for the sampling recovery problem the compact embedding into $L_\infty([0,1]^d)$ is not necessary. There are several relevant multivariate function classes that fall under this scope, such as limiting mixed Besov and Triebel-Lizorkin spaces with smoothness $r = 1/p$ and further parameter conditions to ensure the embedding into the class of continuous functions. We were able to give upper bounds for the sampling widths in $L_q([0,1]^d)$ with $q < \infty$, which are sharp in the polynomial main rate. As for the right order of the logarithm, the situation is completely open.

Let us comment on the particular case of L_1-smoothness spaces. Smoothness spaces built upon $L_1([0,1]^d)$ with smoothness $r = 1$ play an important role in numerical integration. This includes, for instance, the space $S^1_1W([0,1]^d)$ defined using weak derivatives. This space cannot be described using Fourier analytical means and is therefore difficult to handle. However, these spaces are considered in the scope of this article since we also have $r = 1/p$ and a non-compact embedding into $L_\infty([0,1]^d)$. A Faber characterization is shown, similar to above, including

$$\sup_{j\in\mathbb{N}^d_{-1}} \sum_{k\in\mathbb{Z}^d} |d^2_{j,k}(f)| \lesssim \left\| f | S^1_1W(\mathbb{R}^d) \right\| . \tag{7.1}$$

In particular, this would imply the following so-called sampling inequality

$$\sum_{k\in\mathbb{Z}^d} |f(k)| \lesssim \left\| f | S^1_1W(\mathbb{R}^d) \right\| .$$

This extends the result in [[33], Prop. 2] in several directions. On the one hand, we consider the multivariate case, and on the other hand, the space $S^1_1W(\mathbb{R}^d)$ is larger than $S^1_{1,1}B(\mathbb{R}^d)$.

Having (7.1) at hand, the following slightly sharper version of Theorem 6.3 is immediate.

Theorem 7.1. For any $1 \le q < \infty$, we have

$$\|f - I_nf\|_q \lesssim 2^{-n/q} n^{(d-1)/q} \left\| f | S^1_1W([0,1]^d) \right\| .$$

A proper Fourier analytical replacement of the spaces $S^1_1W([0,1]^d)$ are the spaces $S^1_{1,q}F([0,1]^d)$. However, these spaces are not really comparable, especially when $q = \infty$. Using the methods in [[34], Theorem 1.9], there is strong evidence for proving a version of (7.1) also for the spaces $S^1_{1,\infty}F(\mathbb{R}^d)$. This would imply a version of Theorem 7.1. By well-known arguments, a cubature formula with performance

$$\left| \int_{[0,1]^d} f(x)\,dx - \sum_{i=1}^N \lambda_i f(x^i) \right| \lesssim N^{-1}(\log N)^{2(d-1)} \left\| f | S^1_{1,\infty}F([0,1]^d) \right\|$$

can be constructed by integrating the approximand. Note that there were efforts made in the literature to treat such limiting cases, see, for instance [[35], Cor. 6.5]. Suboptimal bounds were proven there. Note that results for such limiting cases are related to the Kokhsma-Hlawka inequality, showing that QMC-cubature in $S_1^1W([0,1]^d)$ is related to the star discrepancy of the cubature nodes.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

SS was supported by the German Academic Exchange Service (DAAD, Grant 57588362) and by the Philipp Schwartz Initiative of the Alexander von Humboldt Foundation.

Acknowledgments

The authors would like to thank the referees for their useful remarks for improving the quality of this manuscript. Additionally, the authors thank Winfried Sickel and Dinh Dũng for discussion and their comments on earlier versions of this manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. Byrenheid G. *Sparse representation of multivariate functions based on discrete point evaluations* [Dissertation]. Institut für Numerische Simulation; Universität Bonn, Bonn, Germany (2018).

2. Hinrichs A, Markhasin L, Oettershagen J, Ullrich T. Optimal quasi-Monte Carlo rules on order 2 digital nets for the numerical integration of multivariate periodic functions. *Numer Math.* (2016) 134:163–96. doi: 10.1007/s00211-015-0765-y

3. Kempka H, Schneider C, Vybiral J. Path regularity of Brownian motion and Brownian sheet. *Constr Approx.* (2022). doi: 10.1007/s00365-023-09647-z

4. Triebel H. *Bases in Function Spaces, Sampling, Discrepancy, Numerical Integration.* Vol. 11 of EMS Tracts in Mathematics. Zürich: European Mathematical Society (EMS) (2010). doi: 10.4171/085

5. Triebel H. *Faber Systems and Their Use in Sampling, Discrepancy, Numerical Integration.* EMS Series of Lectures in Mathematics. Zürich: European Mathematical Society (EMS) (2012). doi: 10.4171/107

6. Dũng D, Temlyakov VN, Ullrich T. Hyperbolic Cross Approximation. In: Tikhonov S, editor. *Advanced Courses in Mathematics. CRM Barcelona.* Cham: Birkhäuser; Springer (2018). doi: 10.1007/978-3-319-92240-9

7. Krieg D, Ullrich M. Function values are enough for L_2-approximation. *Found Comp Math.* (2021) 21:1141–51. doi: 10.1007/s10208-020-09481-w

8. Nagel N, Schäfer M, Ullrich T. A new upper bound for sampling numbers. *Found Comp Math.* (2022) 22:445–68. doi: 10.1007/s10208-021-09504-0

9. Bartel F, Schäfer M, Ullrich T. Constructive subsampling of finite frames with applications in optimal function recovery. *Appl Comput Harmon Anal.* (2023) 65:209–248. doi: 10.1016/j.acha.2023.02.004

10. Krieg D, Ullrich M. Function values are enough for L_2-approximation. II. *J Complex.* (2021) 66:14. doi: 10.1016/j.jco.2021.101569

11. Dolbeault M, Krieg D, Ullrich M. A sharp upper bound for sampling numbers in L_2. *Appl Comput Harmon Anal.* (2023) 63:113–34. doi: 10.1016/j.acha.2022.12.001

12. Temlyakov VN, Ullrich T. Approximation of functions with small mixed smoothness in the uniform norm. *J Approx Theory.* (2022) 277:105718. doi: 10.1016/j.jat.2022.105718

13. Temlyakov V, Ullrich T. Bounds on Kolmogorov widths and sampling recovery for classes with small mixed smoothness. *J Complex.* (2021) 67:101575. doi: 10.1016/j.jco.2021.101575

14. Temlyakov V. On optimal recovery in L_2. *J Complex.* (2021) 65:101545. doi: 10.1016/j.jco.2020.101545

15. Jahn T, Ullrich T, Voigtlaender F. Sampling numbers of smoothness classes via ℓ^1-minimization. *J Complex.* (2022). doi: 10.48550/arXiv.2212.00445

16. Bungartz HJ, Griebel M. Sparse grids. *Acta Numer.* (2004) 13:147–269. doi: 10.1017/S0962492904000182

17. Dũng D. B-spline quasi-interpolant representations and sampling recovery of functions with mixed smoothness. *J Complex.* (2011) 27:541–67. doi: 10.1016/j.jco.2011.02.004

18. Ullrich T. *Smolyak's algorithm, sparse grid approximation and periodic function spaces with dominating mixed smoothness* (PhD thesis). Friedrich-Schiller-Universität Jena, Jena, Germany (2007).

19. Oswald P. L^p-Approximation durch Reihen nach dem Haar-Orthogonalsystem und dem Faber-Schauder-System. *J Approx Theory.* (1981) 33:1–27. doi: 10.1016/0021-9045(81)90086-1

20. Schmeisser HJ, Triebel H. *Topics in Fourier Analysis and Function Spaces.* A Wiley-Interscience Publication. Chichester: John Wiley & Sons, Ltd. (1987).

21. Faber G. Über stetige Funktionen. *Math Ann.* (1909) 66:81–94. doi: 10.1007/BF01450912

22. Ullrich T. Function spaces with dominating mixed smoothness, characterization by differences. *Tech Rep Jenaer Schriften zur Math Inform.* (2006).

23. Vybiral J. Function spaces with dominating mixed smoothness. *Dissert Math.* (2006) 436:73. doi: 10.4064/dm436-0-1

24. Hansen M, Vybíral J. The Jawerth–Franke embedding of spaces with dominating mixed smoothness. *Georg Math J.* (2009) 16:667–82. doi: 10.1515/GMJ.2009.667

25. Mayer S, Ullrich T. Entropy numbers of finite dimensional mixed-norm balls and function space embeddings with small mixed smoothness. *Constr Approx.* (2021) 53:249–79. doi: 10.1007/s00365-020-09510-5

26. Triebel H. *Theory of Function Spaces. Modern Birkhäuser Classics.* Basel: Birkhäuser/Springer Basel AG (2010).

27. Ullrich T. Optimal cubature in Besov spaces with dominating mixed smoothness on the unit square. *J Complex.* (2014) 30:72–94. doi: 10.1016/j.jco.2013.09.001

28. Temlyakov V. Approximation of functions with bounded mixed derivative. *Proc Steklov Inst Math.* (1986) 20:173

29. Tikhomirov VM. Diameters of sets in functional spaces and the theory of best approximations. *Russ Math Surv.* (1960) 15:75–111. doi: 10.1070/RM1960v015n03ABEH004093

30. Romanyuk AS. Approximation of the Besov classes of periodic functions of several variables in the space L_q. *Ukr Math J*. (1991) 43:1297–306. doi: 10.1007/BF01061817

31. Romanyuk AS. Kolmogorov and trigonometric widths of the Besov classes $B^r_{p,\theta}$ of multivariate periodic functions. *Sb Math*. (2006) 197:69–93. doi: 10.1070/SM2006v197n01ABEH003747

32. Novak E, Triebel H. Function spaces in Lipschitz domains and optimal rates of convergence for sampling. *Constr Approx*. (2006) 23:325–50. doi: 10.1007/s00365-005-0612-y

33. Schmeisser HJ, Sickel W. Sampling theory and function spaces. In: *Applied Mathematics Reviews*. Vol. 1. River Edge, NJ: World Science Publication (2000). p. 205–84. doi: 10.1142/9789812792686_0008

34. Garrigós G, Seeger A, Ullrich T. Haar frame characterizations of Besov-Sobolev spaces and optimal embeddings into their dyadic counterparts. *J Fourier Anal Appl*. (2023) 29:39. doi: 10.1007/s00041-023-10013-7

35. Ullrich M, Ullrich T. The role of Frolov's cubature formula for functions with bounded mixed derivative. *SIAM J Numer Anal*. (2016) 54:969–93. doi: 10.1137/15M1014814