DFG-Schwerpunktprogramm 1324

„Extraktion quantifizierbarer Information aus komplexen Systemen"

Best m-Term Approximation and Tensor Products of Sobolev and Besov Spaces – the Case of Non-compact Embeddings

M. Hansen, W. Sickel

Preprint 39

Edited by

AG Numerik/Optimierung
Fachbereich 12 - Mathematik und Informatik
Philipps-Universität Marburg
Hans-Meerwein-Str.
35032 Marburg

DFG-Schwerpunktprogramm 1324

„Extraktion quantifizierbarer Information aus komplexen Systemen"

Best m-Term Approximation and Tensor Products of Sobolev and Besov Spaces – the Case of Non-compact Embeddings

M. Hansen, W. Sickel

Preprint 39

The consecutive numbering of the publications is determined by their chronological order.

The aim of this preprint series is to make new research rapidly available for scientific discussion. Therefore, the responsibility for the contents is solely due to the authors. The publications will be distributed by the authors.

Best m-Term Approximation and Tensor Products of Sobolev and Besov Spaces – the Case of Non-compact Embeddings

Markus Hansen & Winfried Sickel
Friedrich-Schiller-University Jena

March 3, 2010

Abstract

We shall investigate the asymptotic behaviour of the widths of best m-term approximation with respect to non-compact embeddings of tensor products of Sobolev as well as Besov spaces into L_p spaces. In almost all cases our approach leads to final results.

1 Introduction

Let $\Phi := (\psi_j)_j$ denote a tensor product wavelet basis satisfying some additional smoothness, integrability, and moment conditions, see Subsection 3.2.2 for an exact definition. We consider best m-term approximation with respect to Φ, i.e., we investigate the quantity

$$\sigma_m(f,\Phi)_X := \inf \left\{ \| \, f - \sum_{j\in\Lambda} c_j \, \psi_j \, \|_X : \quad |\Lambda| \le m \, , \quad c_j \in \mathbb{C} \, , \, j \in \Lambda \right\} , \qquad m \in \mathbb{N}_0 \, .$$

Associated widths are defined as follows. Let X and Y be quasi-Banach spaces such that $Y \hookrightarrow X$. Then we define

$$\sigma_m(Y,X,\Phi) := \sup \left\{ \sigma_m(f,\Phi)_X : \quad \| \, f \, |Y\| \le 1 \right\} , \qquad m \in \mathbb{N}_0 \, . \tag{1}$$

Usually one concentrates on $X = L_p(\mathbb{R}^d)$. We shall divide our investigations into two different cases. In a first case we shall study the asymptotic behaviour of $\sigma_m(Y,X,\Phi)$ for pairs (X,Y), where $X = L_{p_1}([0,1]^d)$ and Y is either the tensor product of Besov spaces $B^{r_0}_{p_0,p_0}([0,1])$ or of Sobolev spaces $H^{r_0}_{p_0}([0,1])$ (Bessel potential spaces). Then we continue by investigating the same problem for $[0,1]$ replaced by $\mathbb{R}$. As indicated in the title, we concentrate on non-compact embeddings, i.e. we consider $r_0 := \frac{1}{p_0} - \frac{1}{p_1} > 0$ for spaces on the cube and $r_0 \geq \frac{1}{p_0} - \frac{1}{p_1} > 0$ if we consider spaces defined on $\mathbb{R}^d$. Whenever $1 < p_1 < \infty$ holds and Y is a tensor product Besov space, then our approach leads to a final characterization of the asymptotic behaviour of $\sigma_m(Y, X, \Phi)$. If Y is a tensor product Sobolev space, this is also true but with some additional restrictions. Of course, it is well-known that the property

$$\lim_{m\to\infty} \sigma_m(Y, X, \Phi) = 0$$

is not related to the compactness of the embedding $Y \hookrightarrow X$, see, e.g. [7] and [12] for examples given by isotropic Besov spaces.

For us it is convenient to use the coincidence (in the sense of equivalent norms) of the above tensor product spaces with special cases of the scales of function spaces of dominating mixed smoothness, see Section 3 for details. The present paper is a continuation of [12], where we have investigated best m-term approximation with respect to (isotropic) Lizorkin-Triebel spaces.

Concerning the wavelet system Φ some remarks are in order. First of all, we use different systems for spaces on $\mathbb{R}^d$ and for spaces on $[0, 1]^d$. Exact definitions are given below in (11) and (18), respectively. When we deal with the widths $\sigma_m(Y, X, \Phi)$ it is assumed that Y and X allow a characterization by means of the same system Φ, see Propositions 4, 5 for sufficient conditions. In case of spaces on domains we will suppose that the associated spaces on $\mathbb{R}^d$ allow a characterization by Φ.

The paper is organized as follows. In Section 2 we state and comment on our main results. The next section is devoted to tensor product spaces and their relations to the scales of Besov and Lizorkin-Triebel spaces of dominating mixed smoothness. In Section 4 we investigate approximation spaces with respect to sequence spaces and we determine the asymptotic behaviour of the width of best m-term approximation with respect to embeddings of those sequence spaces. This will be an essential tool for us in the following section. However, we believe, it is of some self-contained interest as well. All proofs will be collected in Sections 5 and 6. The first step in our proofs will always be the application of a wavelet isomorphism. This reduces the problem for distribution spaces to a problem for sequence spaces. The appropriate wavelet isomorphisms will be described in Subsection 3.2.2. From this it follows immediately that the main job has to be done on the level of sequence spaces, for which we refer to Section 5. In Section 6 we collect consequences for the widths of best m-term approximation with respect to pairs of spaces of dominating mixed smoothness. In particular, we determine the asymptotic behaviour of $\sigma_m\Big(S^{r_0}_{p_0,q_0}B(\mathbb{R}^d), S^0_{p_1,q_1}B(\mathbb{R}^d), \Phi\Big)$, $r_0 = 1/p_0 - 1/p_1$, in all reasonable situations, see Cor. 8.

Notation

As usual, $\mathbb{N}$ denotes the natural numbers, $\mathbb{Z}$ the integers and $\mathbb{R}$ the real numbers. Further we use $\mathbb{N}_0$ in place of $\mathbb{N} \cup \{0\}$. For a real number a we put $a_+ := \max(a, 0)$. By $[a]$ we denote the integer part of a. With $\mathbb{R}^d$, $\mathbb{Z}^d$ and $\mathbb{N}_0^d$ we denote the d-dimensional counterparts. If $j \in \mathbb{N}_0^d$, then

$$|j|_1 := |(j_1, \ldots, j_d)|_1 = j_1 + \ldots + j_d \,.$$

If X and Y are two quasi-Banach spaces, then the symbol $Y \hookrightarrow X$ indicates that the embedding is continuous. As usual, the symbol c denotes positive constants which depend only on the fixed parameters r, p, q and probably on auxiliary functions, unless otherwise stated; its value may vary from line to line. Sometimes we will use the symbols "$\lesssim$" and "$\gtrsim$" instead of "$\leq$" and "$\geq$", respectively. The meaning of $A \lesssim B$ is given by: there exists a constant $c > 0$ such that $A \leq c\, B$. Similarly $\gtrsim$ is defined. The symbol $A \asymp B$ will be used as an abbreviation of $A \lesssim B \lesssim A$. For a discrete set ∇ the symbol $|\nabla|$ denotes the cardinality of this set. We shall use the multiindex convention that for two vectors $\alpha, \beta \in \mathbb{N}_0^d$ the inequality $\alpha \leq \beta$ means $\alpha_i \leq \beta_i$, $i = 1, \ldots, d$.

Tensor products of Besov and Sobolev spaces are investigated in [28], [26] and [27]. General information about Besov and Lizorkin-Triebel spaces of dominating mixed smoothness can be found, e.g., in [1, 25, 23, 34, 32] ($S^r_{p,q}B(\mathbb{R}^d)$, $S^r_{p,q}F(\mathbb{R}^d)$). We will not give definitions here. However, the wavelet characterizations, recalled in Subsection 3.2.2, can be taken as definitions as well. The reader, who is interested in more elementary descriptions of these spaces, e.g., by means of differences, is refered to [1, 25, 32] and [33].

Agreement: We shall deal with function and sequence spaces depending on three parameters r, p, q. If there is given no additional restriction then it is assumed that $r \in \mathbb{R}$ and $0 < p, q \leq \infty$. Furthermore, Ω always denotes an open, nontrivial subset of $\mathbb{R}^d$. $\Omega = \mathbb{R}^d$ is admissible. Finally, if we consider $S^r_{p,q}F(\Omega)$, $S^r_{p,q}F(\mathbb{R}^d)$, $s^r_{p,q}f(\Omega)$, or $s^r_{p,q}f$, then it is always assumed that $p < \infty$.

2 The asymptotic behaviour of the widths of best m-term approximation

Our main interest lies in the asymptotic behaviour of the widths $\sigma_m(Y, X, \Phi)$ for different choices of the spaces X and Y. Tensor products of Sobolev spaces are denoted by $S^r_pH(I^d)$ and I is either the interval $[0, 1]$ or $I = \mathbb{R}$. Tensor products of Besov spaces are denoted by $S^r_pB(I^d)$ and I is as before. Definitions are given in Section 3. To begin with we study the situation on the cube $[0,1]^d$. Afterwards we investigate the same problem on $\mathbb{R}^d$.

2.1 Widths of best m-term approximation on the cube

Since $Y \hookrightarrow X$ is a necessary condition we recall the needed embedding assertions.

Lemma 1. *Let $1 < p_1 < \infty$ and $r_0 > 0$.*
(i) *Let $1 < p_0 < \infty$. Then $S^{r_0}_{p_0}H([0,1]^d) \hookrightarrow L_{p_1}([0,1]^d)$ holds if, and only if,*

$$r_0 \geq \frac{1}{p_0} - \frac{1}{p_1}\,. \tag{2}$$

(ii) *Let $0 < p_0 \leq \infty$. Then $S^{r_0}_{p_0}B([0,1]^d) \hookrightarrow L_{p_1}([0,1]^d)$ holds if, and only if, (2) holds.*
(iii) *The embeddings in (i) and (ii) are compact if, and only if, the inequality in (2) is strict.*

Remark 1. Concerning parts (i) and (ii) we refer to [25, 2.4.1], see also [24]. A proof of Lemma 1(iii) may be found in [34].

Surprisingly, whenever we have a continuous embedding s.t. $r_0 > 0$, the associated widths of best m-term approximation form a null sequence. Roughly speaking, the widths $\sigma_m(S^{r_0}_{p_0}Y([0,1]^d), L_{p_1}([0,1]^d, \Phi)$ of best m-term approximation behave like m^{-r_0} times a lower order term. Only the lower order term is influenced by $Y \in \{B, W\}$, p_0 and p_1.

Let Φ be as in (18). Furthermore, we assume that Prop. 5 can be applied with this system Φ for the spaces $S^{r_0}_{p_0,p_0}F(\mathbb{R}^d)$, $S^{r_0}_{p_0,2}F(\mathbb{R}^d)$, and $S^{0}_{p_1,2}F(\mathbb{R}^d)$.

Theorem 1. *We assume* $\max(1,p_0) < p_1 < \infty$, *and* $r_0 := \frac{1}{p_0} - \frac{1}{p_1}$.
(i) *In case of tensor product Besov spaces we have*

$$\sigma_m\Big(S^{r_0}_{p_0}B([0,1]^d), L_{p_1}([0,1]^d), \Phi\Big) \asymp m^{-r_0}\,(\log m)^{(d-1)(r_0 - \frac{1}{p_0} + \frac{1}{2})_+} \tag{3}$$

for all $m \geq 2$.
(ii) *Let $1 < p_0 \leq 2 \leq p_1 < \infty$. In case of tensor product Sobolev spaces we have*

$$\sigma_m\Big(S^{r_0}_{p_0}H([0,1]^d), L_{p_1}([0,1]^d), \Phi\Big) \asymp m^{-r_0}\,(\log m)^{(d-1)r_0} \tag{4}$$

for all $m \geq 2$.

Remark 2. (i) Observe that in case $0 < p_0 < 2$ and $p_1 = 2$ we have

$$\sigma_m\Big(S^{1/p_0-1/2}_{p_0}B([0,1]^d), L_2([0,1]^d), \Phi\Big) \asymp m^{-r_0}\ ,$$

for all $m \geq 2$. This has been known before, see Nitsche [19]. However, it seems that Nitsche was not aware of Proposition 6.

(ii) There are several references dealing with best m-term approximation of functions belonging to spaces of dominating mixed smoothness on the d-dimensional torus, we refer to Bazarkhanov [5], Oswald [20], Temlyakov [30], Dinh Dung [8, 9] and Romanyuk [22]. In all these articles only compact embeddings are investigated.

2.2 Widths of best m-term approximation on $\mathbb{R}^d$

Again we start by having a look on the possible embeddings. Of course, since the underlying domain has infinite measure, the conditions are more restrictive than in Lemma 1.

Lemma 2. *Let* $1 < p_1 < \infty$ *and* $r_0 > 0$.
(i) *Let* $1 < p_0 < \infty$. *Then* $S^{r_0}_{p_0}H(\mathbb{R}^d) \hookrightarrow L_{p_1}(\mathbb{R}^d)$ *holds if, and only if,* $p_0 \le p_1$ *and (2) is satisfied.*
(ii) *Let* $0 < p_0 \le \infty$. *Then* $S^{r_0}_{p_0}B(\mathbb{R}^d) \hookrightarrow L_{p_1}(\mathbb{R}^d)$ *holds if, and only if,* $p_0 \le p_1$ *and (2) is satisfied.*
(iii) *The embeddings in (i) and (ii) are never compact.*

Remark 3. Again we refer to [25, 2.4] for a proof of (i) and (ii), see also [24]. Part (iii) is obvious.

In this subsection we assume that Φ is as in (11). Furthermore, we assume that Prop. 5 can be applied with this system Φ for the spaces $S^{r_0}_{p_0,p_0}F(\mathbb{R}^d)$, $S^{r_0}_{p_0,2}F(\mathbb{R}^d)$, and $S^{0}_{p_1,2}F(\mathbb{R}^d)$. We have two consider two different cases. First, we study the so-called non-limiting case, given by $r_0 > 1/p_0 - 1/p_1$.

Theorem 2. *Let* $1 < p_1 < \infty$ *and* $r_0 > 0$. *We suppose*

$$r_0 > \frac{1}{p_0} - \frac{1}{p_1}. \tag{5}$$

Then we have

$$\sigma_m\Big(S^{r_0}_{p_0}Y(\mathbb{R}^d), L_{p_1}(\mathbb{R}^d), \Phi\Big) \asymp m^{-\frac{1}{p_0}+\frac{1}{p_1}}\,, \tag{6}$$

for all $m \ge 2$ *and* $Y \in \{W, B\}$.

Remark 4. Observe that the asymptotic behaviour of $\sigma_m\Big(S^{r_0}_{p_0}Y(\mathbb{R}^d), L_{p_1}(\mathbb{R}^d), \Phi\Big)$ does not depend on r_0.

Next we turn to the limiting case, i.e. $r_0 = 1/p_0 - 1/p_1$.

Theorem 3. *We suppose* $1 < p_1 < \infty$ *and*

$$r_0 := \frac{1}{p_0} - \frac{1}{p_1} > 0. \tag{7}$$

(i) *Then we have*

$$\sigma_m\Big(S^{r_0}_{p_0}B(\mathbb{R}^d), L_{p_1}(\mathbb{R}^d), \Phi\Big) \asymp m^{-r_0}\,(\log m)^{(d-1)(r_0-\frac{1}{p_0}+\frac{1}{2})_+}\,,$$

for all $m \geq 2$.

(ii) *If* $1 < p_0 \leq 2 \leq p_1 < \infty$*, then we have*

$$\sigma_m\Big(S^{r_0}_{p_0}H(\mathbb{R}^d), L_{p_1}(\mathbb{R}^d), \Phi\Big) \asymp m^{-r_0}\,(\log m)^{(d-1)r_0}\,,$$

for all $m \geq 2$.

3 Tensor products of Sobolev and Besov spaces and their relation to Besov and Lizorkin-Triebel spaces of dominating mixed smoothness

The most prominent classes within these spaces occuring in the title of this section are the tensor product Sobolev spaces.

3.1 Tensor products of Sobolev spaces

Before we are turning to these classes we introduce Sobolev spaces of dominating mixed smoothness.

Let $1 < p < \infty$ and $r \in \mathbb{N}$. Then $S^r_pW(\mathbb{R}^d)$ is the collection of all functions in $L_p(\mathbb{R}^d)$ s.t.

$$\|\, f\,|S^r_pW(\mathbb{R}^d)\| := \sum_{\alpha\leq(r,\dots,r)} \|\, D^\alpha f|L_p(\mathbb{R}^d)\| < \infty\,.$$

Here $D^\alpha f$ denotes the distributional derivative of order α of f. The derivative $D^\alpha f$ of the highest order is the mixed one, given by $\alpha = (r,\dots,r)$. This explains the name *Sobolev space of dominating mixed smoothness.*

The connection to tensor product spaces is as follows, see [26]. For the basics of tensor products of Banach spaces we refer to [18].

Proposition 1. *Let* $d \geq 1$, $1 < p < \infty$ *and* $r \in \mathbb{N}$. *Let* α_p *denote the* p*-nuclear tensor norm, see e.g. [18]. Then*

$$S^r_pW(\mathbb{R}^2) = W^r_p(\mathbb{R}) \otimes_{\alpha_p} W^r_p(\mathbb{R}) \tag{8}$$

and

$$S^r_pW(\mathbb{R}^{d+1}) = S^r_pW(\mathbb{R}^d) \otimes_{\alpha_p} W^r_p(\mathbb{R}) = W^r_p(\mathbb{R}) \otimes_{\alpha_p} S^r_pW(\mathbb{R}^d)\,. \tag{9}$$

Remark 5. For $p = 2$ Proposition 1 has been folklore. Let us mention that many times $S^r_2W(\mathbb{R}^d)$ is also denoted as $H^r_{mix}(\mathbb{R}^d)$.

Proposition 1 has a fractional order counterpart. Sobolev spaces $H^r_p(\mathbb{R})$ of fractional order $r > 0$ as well as the Sobolev spaces $W^r_p(\mathbb{R})$, $r \in \mathbb{N}$ itself can be interpreted as special cases of the Lizorkin-Triebel scale, see, e.g., [31, 2.5.6]. If $1 < p < \infty$ and $r > 0$, then it holds

$$H^r_p(\mathbb{R}) = F^r_{p,2}(\mathbb{R}) \qquad \text{(in the sense of equivalent norms).}$$

There is another variant of Sobolev spaces of fractional order, usually called Slobodeckij spaces and denoted by $W^r_p(\mathbb{R}^d)$. These spaces coincide with $B^r_{p,p}(\mathbb{R}^d)$ if $r \notin \mathbb{N}$. Furthermore, there is a well-developed theory of Lizorkin-Triebel spaces of dominating mixed smoothness $S^r_{p,q}F(\mathbb{R}^d)$, see Prop. 5 below. Defining for $r > 0$

$$S^r_pH(\mathbb{R}^d) := S^r_{p,2}F(\mathbb{R}^d)$$

then the following supplement to Proposition 1 has been proved in [26].

Proposition 2. *Let $d \geq 1$, $r > 0$ and $1 < p < \infty$. Then*

$$S^r_pH(\mathbb{R}^2) = H^r_p(\mathbb{R}) \otimes_{\alpha_p} H^r_p(\mathbb{R})$$

and

$$S^r_pH(\mathbb{R}^{d+1}) = S^r_pH(\mathbb{R}^d) \otimes_{\alpha_p} H^r_p(\mathbb{R}) = H^r_p(\mathbb{R}) \otimes_{\alpha_p} S^r_pH(\mathbb{R}^d)\,.$$

Finally, we wish to mention that all these tensor product formulas carry over from $\mathbb{R}$ to intervals. For simplicity let $S^r_pH([0,1]^d)$ be defined by restrictions, see Subsection 3.2.3. Then the following is proved in [27].

Proposition 3. *Let $d \geq 1$, $1 < p < \infty$ and $r > 0$. Then*

$$S^r_pH([0,1]^2) = H^r_p([0,1]) \otimes_{\alpha_p} H^r_p([0,1])$$

and

$$S^r_pH([0,1]^{d+1}) = S^r_pH([0,1]^d) \otimes_{\alpha_p} H^r_p([0,1]) = H^r_p([0,1]) \otimes_{\alpha_p} S^r_pH([0,1]^d)\,.$$

Remark 6. Let I be either $[0,1]$ or $\mathbb{R}$. With the interpretation as an iterated tensor product we may write

$$S^r_pH(I^d) \quad = \quad \underbrace{H^r_p(I) \otimes_{\alpha_p} \ldots \otimes_{\alpha_p} H^r_p(I)}_{d-\text{times}}\,.$$

3.2 Isomorphisms associated to tensor product wavelet systems and Besov-Lizorkin-Triebel spaces of dominating mixed smoothness

As mentioned in the Introduction we will not give definitions of the above classes. However, the characterizations by means of tensor product wavelet systems may be taken as a definition.

3.2.1 Sequence spaces

Now we introduce sequence spaces related to the characterization of Besov and Triebel-Lizorkin spaces of dominating mixed smoothness in terms of wavelet coefficients.
Let $\mathcal{X}$ denote the characteristic function of the interval $[0,1]$. Then we put

$$\mathcal{X}_{j,k}(x) := \mathcal{X}(2^{j_1}x_1 - k_1)\cdots\mathcal{X}(2^{j_d}x_d - k_d)\,, \qquad x = (x_1,\ldots,x_d) \in \mathbb{R}^d\,, \tag{10}$$

where $j = (j_1,\ldots,j_d) \in \mathbb{N}_0^d$ and $k = (k_1,\ldots,k_d) \in \mathbb{Z}^d$. In other words, $\mathcal{X}_{j,k}$ is the characteristic function of the dyadic rectangle

$$Q_{j,k} = \big(2^{-j_1}([0,1]+k_1)\big) \times \cdots \times \big(2^{-j_d}([0,1]+k_d)\big)\,.$$

Definition 1. *Let $d > 1$, $0 < q \le \infty$ and $r \in \mathbb{R}$. Let $\nabla = (\nabla_j)_{j\in\mathbb{N}_0^d}$ be a sequence of nontrivial subsets of $\mathbb{Z}^d$.*

(i) *Let $0 < p \le \infty$. Then $s^r_{p,q}b(\nabla)$ consists of all sequences $a = (a_{j,k})_{j,k}$ such that*

$$\big\| \, a \, |s^r_{p,q}b(\nabla)\big\| := \Bigg(\sum_{j\in\mathbb{N}_0^d} 2^{|j|_1\,(r+\frac{1}{2}-\frac{1}{p})q}\,\Big(\sum_{k\in\nabla_j} |a_{j,k}|^p\Big)^{\frac{q}{p}}\Bigg)^{\frac{1}{q}} < \infty\,.$$

(ii) *Let $0 < p < \infty$. Then $s^r_{p,q}f(\nabla)$ consists of all sequences $a = (a_{j,k})_{j,k}$ such that*

$$\big\| \, a \, |s^r_{p,q}f(\nabla)\big\| := \Bigg\|\Big(\sum_{j\in\mathbb{N}_0^d}\sum_{k\in\nabla_j} 2^{|j|_1\,(r+\frac{1}{2})q}\,|a_{j,k}|^q\,\mathcal{X}_{j,k}(\,\cdot\,)\Big)^{\frac{1}{q}}\,\Big|L_p(\mathbb{R}^d)\Bigg\| < \infty\,.$$

Remark 7.

(i) Many times we shall use the notation $s^r_{p,q}x(\nabla)$, $x \in \{b,f\}$. If, then we always assume $p < \infty$ if $x = f$.

(ii) Two special cases of sequences ∇ are of particular importance. The first one is simply $\nabla_j = \mathbb{Z}^d$ for all j, and we will denote the corresponding spaces by $s^r_{p,q}b$ and $s^r_{p,q}f$, respectively. In the second one we will choose ∇_j in dependence of a bounded open (nontrivial) subset of $\mathbb{R}^d$. The exact definition will be given below, see Def. 2.

(iii) Obviously we have $s^r_{p,p}b(\nabla) = s^r_{p,p}f(\nabla)$.

3.2.2 Spaces of dominating mixed smoothness

For a definition of spaces of dominating mixed smoothness in Fourier-analytic terms we refer to the monographs [1] and [25]. Characterizations in terms of atoms and wavelets have been given in papers by Bazarkhanov [2, 3, 4] and Vybiral [34]. Here we are going to recall a few results from [34].
Let ϕ denote an univariate scaling function associated with the wavelet ψ such that $\phi, \psi \in C^s(\mathbb{R})$ are compactly supported, the wavelet satisfies the moment condition

$$\int_{-\infty}^{\infty} t^\ell\,\psi(t)\,dt = 0\,, \quad 0 \le \ell < s\,,$$

and the collection of functions, consisting of

$$
\begin{aligned}
\psi_{0,k}(t) &:= \phi(t-k)\,, \qquad k \in \mathbb{Z}\,,\\
\psi_{j,k}(t) &:= 2^{(j-1)/2}\,\psi(2^{j-1}t-k)\,, \qquad k \in \mathbb{Z}, \quad j \in \mathbb{N}\,,
\end{aligned}
$$

forms an orthonormal basis for the space $L_2(\mathbb{R})$. Here $s \in \mathbb{N}$ will be chosen later on. Next we need the corresponding tensor product systems. We put

$$
\psi_{j,k}(x) := \psi_{j_1,k_1}(x_1)\cdot\ldots\cdot\psi_{j_d,k_d}(x_d)\,, \qquad x = (x_1,\ldots,x_d)\,.
$$

Then the tensor product wavelet system Φ is the collection

$$
\psi_{j,k}\,, \qquad j \in \mathbb{N}_0^d, \quad k \in \mathbb{Z}^d\,. \tag{11}
$$

For the following propositions we refer to [34]. To begin with we deal with Lizorkin-Triebel spaces.

Proposition 4. *Let $r \in \mathbb{R}$, $0 < p < \infty$ and $0 < q \le \infty$. If φ, ψ are satisfying the above conditions for sufficiently large $s = s(r,p,q)$ then we have the following: For every $f \in S^r_{p,q}F(\mathbb{R}^d)$, we have*

$$
f = \sum_{j\in\mathbb{N}_0^d}\sum_{k\in\mathbb{Z}^d}\langle f,\,\psi_{j,k}\rangle\,\psi_{j,k}\,, \tag{12}
$$

convergence in $\mathcal{S}'(\mathbb{R}^d)$ (and in $S^r_{p,q}F(\mathbb{R}^d)$ if $q<\infty$), and

$$
\|\, f\,|S^r_{p,q}F(\mathbb{R}^d)\| \asymp \|\,(\langle f,\,\psi_{j,k}\rangle)_{j,k}\,|s^r_{p,q}f\|\,. \tag{13}
$$

The counterpart in case of Besov spaces reads as follows.

Proposition 5. *Let $r \in \mathbb{R}$, $0 < p \le \infty$ and $0 < q \le \infty$. If φ, ψ are satisfying the above conditions for sufficiently large $s = s(r,p,q)$ then we have the following: For every $f \in S^r_{p,q}B(\mathbb{R}^d)$, the identity (12) holds with convergence in $\mathcal{S}'(\mathbb{R}^d)$ (and in $S^r_{p,q}B(\mathbb{R}^d)$ if $\max(p,q)<\infty$), and*

$$
\|\, f\,|S^r_{p,q}B(\mathbb{R}^d)\| \asymp \|\,(\langle f,\,\psi_{j,k}\rangle)_{j,k}\,|s^r_{p,q}b\|\,. \tag{14}
$$

Remark 8. Here we are not interested in optimal conditions with respect to Φ. However, we wish to mention the following. If we have two fixed triples (r_0,p_0,q_0) and (r_1,p_1,q_1), then there always exists a system Φ s.t. Prop. 4 (Prop. 5) can be applied simultaneously with respect to both spaces. Propositions 4 and 5 will allow us to transfer the major part of our calculations from function spaces to sequence spaces.

3.2.3 Wavelets and domains

Let $\Omega \subset \mathbb{R}^d$ be a bounded and open set. Then we define the spaces $S^r_{p,q}F(\Omega)$ and $S^r_{p,q}B(\Omega)$ by restrictions. More exactly, we put

$$\begin{aligned} S^r_{p,q}X(\Omega) &:= \Big\{ f \in D'(\Omega): \quad f = g_{|\Omega} \quad \text{for some } g \in S^r_{p,q}X(\mathbb{R}^d) \Big\} \\ \| \, f \, | S^r_{p,q}X(\Omega) \| &:= \inf \| \, g \, | S^r_{p,q}X(\mathbb{R}^d) \| \, , \end{aligned}$$

where the infimum is taken over all $g \in S^r_{p,q}X(\mathbb{R}^d)$ such that $f = g_{|\Omega}$. Here $X \in \{F, B\}$. For our purpose it is enough to observe the following. The univariate scaling function ϕ and the associated wavelet ψ are compactly supported, say

$$\Big(\operatorname{supp} \phi \, \cup \, \operatorname{supp} \psi \Big) \subset [-N, N]$$

for some $N > 0$. For given $f \in S^r_{p,q}X(\Omega)$ let $\mathcal{E}f$ denote an extension of f such that

$$\| \, \mathcal{E}f \, | S^r_{p,q}X(\mathbb{R}^d) \| \le 2 \, \| \, f \, | S^r_{p,q}X(\Omega) \| \le 2 \, \| \, \mathcal{E}f \, | S^r_{p,q}X(\mathbb{R}^d) \| \, .$$

Then

$$\mathcal{E}f = \sum_{j \in \mathbb{N}_0^d} \sum_{k \in \mathbb{Z}^d} \langle \mathcal{E}f, \, \psi_{j,k} \rangle \, \psi_{j,k} \, .$$

Hence, also

$$\mathcal{E}^* f := \sum_{j \in \mathbb{N}_0^d} \sum_{\operatorname{supp} \psi_{j,k} \cap \Omega \neq \emptyset} \langle \mathcal{E}f, \, \psi_{j,k} \rangle \, \psi_{j,k} \tag{15}$$

is an extension of f such that

$$\| \, \mathcal{E}^* f \, | S^r_{p,q}X(\mathbb{R}^d) \| \asymp \| \, f \, | S^r_{p,q}X(\Omega) \| \, . \tag{16}$$

Moreover, we have

$$\operatorname{supp} \mathcal{E}^* f \subset \Gamma := \{ x \in \mathbb{R}^d : \ \operatorname{dist}(x, \Omega) < 2N \} \, . \tag{17}$$

For spaces on bounded domains we define Φ to be the collection of all functions $\psi_{j,k}$ such that

$$\Omega \cap \operatorname{supp} \psi_{j,k} \neq \emptyset \, , \qquad j \in \mathbb{N}_0^d, \ k \in \mathbb{Z}^d \, . \tag{18}$$

With (15) and (16) we do not get an intrinsic characterization of $S^r_{p,q}X(\Omega)$.
Now we are ready to define related sequence spaces.

Definition 2. *Let $\Omega \subset \mathbb{R}^d$ be a bounded open (nontrivial) set. Let r, p, q as in Def. 1. Let $(\psi_{j,k})_{j,k}$ satisfy the restrictions in Prop. 4 (f-case) or in Prop. 5 (b-case). Then we define $\nabla = \nabla(\Omega)$ according to (18) by*

$$\nabla_j = \Big\{ k \in \mathbb{Z}^d : \quad \operatorname{supp} \, \psi_{j,k} \cap \Omega \neq \emptyset \Big\} \, . \tag{19}$$

The corresponding sequence spaces will be denoted by $s^r_{p,q}b(\Omega) := s^r_{p,q}b(\nabla(\Omega))$ and $s^r_{p,q}f(\Omega) := s^r_{p,q}f(\nabla(\Omega))$ respectively.

Remark 9. We want to mention the following fact, which we will make constant use of without always explicitely mentioning it. There exist positive constants $C_1 = C_1(\Omega)$ and $C_2 = C_2(\Omega)$ and an appropriate integer $J = J(\Omega)$, such that

$$C_1 \le 2^{-|j|_1} \, |\nabla_j(\Omega)| \le C_2, \qquad |j|_1 \ge J. \tag{20}$$

For simplicity we will always assume $J = 0$.

3.3 Tensor products of Besov spaces

By γ_p we denote the projective tensor p-norm, see [26]. We define

$$\delta_p := \begin{cases} \alpha_p & \text{if} \quad 1 < p < \infty\,, \\ \gamma_p & \text{if} \quad 0 < p \le 1\,. \end{cases}$$

The counterpart of Propositions 1, 2 and 3 in case of Besov spaces reads as follows, see [26] and [27].

Proposition 6. *Let $d \ge 1$, $0 < p < \infty$ and $r \in \mathbb{R}$. Then*

$$S^r_{p,p}B(I^2) = B^r_{p,p}(I) \otimes_{\delta_p} B^r_{p,p}(I)$$

and

$$S^r_{p,p}B(I^{d+1}) = S^r_{p,p}B(I^d) \otimes_{\delta_p} B^r_{p,p}(I) = B^r_{p,p}(I) \otimes_{\delta_p} S^r_{p,p}B(I^d)\,,$$

where I is either $\mathbb{R}$ or $I = [0,1]$.

Remark 10. (i) Let I be either $\mathbb{R}$ or $I = [0,1]$. Instead of $S^r_{p,p}B(I^d)$ we shall use many times the shorter form $S^r_pB(I^d)$. By using this abbreviation we obtain the following comparison between tensor product Sobolev and tensor product Besov spaces

$$S^r_pB(I^d) \hookrightarrow S^r_pH(I^d) \qquad \text{if} \quad 1 < p \le 2$$

and

$$S^r_pH(I^d) \hookrightarrow S^r_pB(I^d) \qquad \text{if} \quad 2 \le p < \infty\,.$$

(ii) As in Remark 6 we may write

$$S^r_pB(I^d) \quad = \quad \underbrace{B^r_{p,p}(I) \otimes_{\delta_p} \ldots \otimes_{\delta_p} B^r_{p,p}(I)}_{d-\text{times}}\,,$$

with the interpretation as an iterated tensor product.

4 Sequence spaces and approximation spaces

In this section, we deal with abstract approximation spaces as well as the behaviour of best m-term approximation with respect to embeddings of vector-valued ℓ_p spaces.

4.1 Approximation spaces

Let $\mathcal{D}$ be a subset of the quasi-Banach space X. Then we define

$$\sigma_m\big(a, X, \mathcal{D}\big) := \inf\left\{ \Big\| a - \sum_{n\in\Lambda} c_n\, \psi_n \Big\|_X \ : \ |\Lambda| \le m\,, \ c_n \in \mathbb{C}\,, \ \psi_n \in \mathcal{D}\,, \ n \in \Lambda \right\} ,$$

Obviously $\sigma_0(a, X, \mathcal{D}) = \|\, a\,\|_X$. We are interested in approximation spaces relative to σ_m. Let $\mathcal{A}^s_q(X, \mathcal{D})$ be the collection of all elements $a \in X$, such that

$$\begin{aligned} &\|\, a \ \,|\mathcal{A}^s_q(X,\mathcal{D})\| \\ &\quad := \begin{cases} \Big(\sum\limits_{m=0}^{\infty} \Big[(m+1)^s\, \sigma_m\big(a, X, \mathcal{D}\big)\Big]^q \dfrac{1}{m+1}\Big)^{1/q} & \text{if} \quad 0 < q < \infty\,, \\ \sup\limits_{m=0,1,\ldots} \ (m+1)^s\, \sigma_m\big(a, X, \mathcal{D}\big) & \text{if} \quad q = \infty\,, \end{cases} \end{aligned}$$

where $s > 0$. We want to mention two further facts, which are almost trivial but nonetheless important for our later considerations. Let $s > 0$ and $0 < u_0 \le u_1 \le \infty$. Then we have the embedding

$$\mathcal{A}^s_{u_0}(X, \mathcal{D}) \hookrightarrow \mathcal{A}^s_{u_1}(X, \mathcal{D})$$

(use the monotonicity of σ_m with respect to m and switch to dyadic subsequences). Moreover, if X and Y are two quasi-Banach spaces with $\mathcal{D} \subset X$ and $X \hookrightarrow Y$, then it holds

$$\mathcal{A}^s_u(X, \mathcal{D}) \hookrightarrow \mathcal{A}^s_u(Y, \mathcal{D})$$

for all $s > 0$ and $0 < u \le \infty$.

We need some general assertions on best m-term approximation on the level of sequence spaces. In this connection we concentrate on best m-term approximation with respect to the canonical orthonormal basis of $\ell_2(I)$, where I is a fixed infinite, but countable, index set. We put

$$\mathcal{B} := \{e^j \ : \ j \in I\}\,, \quad e^j := (e^j_k)_k\,, \quad e^j_k := \delta_{j,k}\,, \ j,k \in I.$$

By $\ell_{p,u}(I)$ we denote the Lorentz sequence spaces. Here $\ell_{p,u}(I)$ is the collection of all sequences $a = (a_j)_{j\in I}$, such that

$$\|\, a\, |\ell_{p,u}(I)\| := \Big\| \big(n^{\frac{1}{p}-\frac{1}{u}} a^*_n\big)_{n\in\mathbb{N}} \Big| \ell_u(\mathbb{N}) \Big\| < \infty, \qquad 0 < p, u \le \infty\,,$$

where $a^* = (a^*_n)_n$ denotes the non-increasing rearrangement of a.

We shall investigate best m-term approximation with respect to the following spaces of vector-valued sequences.

Definition 3. *Let I be an arbitrary nonempty, countable index set, and let $J = (J_i)_{i\in I}$ be a family of nonempty, countable index sets. Let $0 < p, q \le \infty$. Then we define the space $\ell_q\big(I, \ell_p(J)\big)$ as the collection of all sequences*

$$a = \big\{a_{i,j} : a_{i,j} \in \mathbb{C}, i \in I, j \in J_i\big\}\,,$$

such that

$$\big\|a\big|\ell_q\big(I, \ell_p(J)\big)\big\| := \Bigg(\sum_{i\in I}\Bigg(\sum_{j\in J_i}|a_{i,j}|^p\Bigg)^{\frac{q}{p}}\Bigg)^{\frac{1}{q}} < \infty\,.$$

Our point of departure is the following nice result of Pietsch [21, Ex. 1].

Proposition 7. *Let $0 < p_1, u \le \infty$. Let I be a fixed index set. Then $a \in \ell_{p_1}(I)$ belongs to the approximation space $\mathcal{A}^s_u(\ell_{p_1}(I), \mathcal{B})$, if and only if $a \in \ell_{p_0,u}(I)$, where $1/p_0 := s + 1/p_1$. Furthermore,*

$$\big\|\, a \,\big|\mathcal{A}^s_u\big(\ell_{p_1}(I), \mathcal{B}\big)\big\| \asymp \big\|\, a \,\big|\ell_{p_0,u}(I)\big\|\,, \tag{21}$$

where the constants of equivalence do not depend on I.

The observation $s^r_{p,p}b(\nabla) = s^r_{p,p}f(\nabla) = \ell_p(\mathbb{N}^d_0 \times \nabla)$ if $r + \frac{1}{2} - \frac{1}{p} = 0$ allows us to apply Prop. 7.

Corollary 1. *Let ∇ be as in Definition 1. Let $0 < p_0 < p_1$. Then*

$$\mathcal{A}^{\frac{1}{p_0}-\frac{1}{p_1}}_{p_0}\big(\ell_{p_1}(\mathbb{N}^d_0 \times \nabla), \mathcal{B}\big) = s^{\frac{1}{p_0}-\frac{1}{2}}_{p_0,p_0}b(\nabla)\,,$$

in the sense of equivalent quasi-norms.

We formulate a second (equivalent) variant.

Corollary 2. *Let ∇ be as in Definition 1. Let $0 < p_0 < p_1$ and $r \in \mathbb{R}$. Then we have*

$$\mathcal{A}^{\frac{1}{p_0}-\frac{1}{p_1}}_{p_0}\big(s^{r+\frac{1}{p_1}-\frac{1}{2}}_{p_1,p_1}b(\nabla), \mathcal{B}\big) = s^{r+\frac{1}{p_0}-\frac{1}{2}}_{p_0,p_0}b(\nabla)\,,$$

in the sense of equivalent quasi-norms.

Proof. We consider the mapping $a \mapsto b$ defined by $b_{j,k} = 2^{|j|_1 r}\, a_{j,k}$. Then

$$\sigma_m\big(a, s^{r+\frac{1}{p_1}-\frac{1}{2}}_{p_1,p_1}b(\nabla), \mathcal{B}\big) = \sigma_m\big(b, s^{\frac{1}{p_1}-\frac{1}{2}}_{p_1,p_1}b(\nabla), \mathcal{B}\big) = \sigma_m\big(b, \ell_{p_1}(\mathbb{N}^d_0 \times \nabla), \mathcal{B}\big)$$

and

$$\big\|\, a \,\big|s^{r+\frac{1}{p_0}-\frac{1}{2}}_{p_0,p_0}b(\nabla)\big\| = \big\|\, b \,\big|s^{\frac{1}{p_0}-\frac{1}{2}}_{p_0,p_0}b(\nabla)\big\| = \big\|\, b \,\big|\ell_{p_0}(\mathbb{N}^d_0 \times \nabla)\big\|\,.$$

This proves the claim. ■

Remark 11. We are mainly interested in the case $r + \frac{1}{p_1} - \frac{1}{2} = 0$. Then we find

$$\mathcal{A}^{\frac{1}{p_0}-\frac{1}{p_1}}_{p_0}\big(s^0_{p_1,p_1}b(\nabla), \mathcal{B}\big) = s^{\frac{1}{p_0}-\frac{1}{p_1}}_{p_0,p_0}b(\nabla)\,,$$

again in the sense of equivalent quasi-norms.

4.2 Widths of best m-term approximation for embeddings of ℓ_p-spaces

The results we are proving in this subsection have preparatory character. However, they are of interest on their own.

Theorem 4. *Let* $0 < p, q \le \infty$ *and* $r > 0$. *Then there is a positive constant* c, *such that*

$$\sigma_m\Big(\ell_{q_r}(I, \ell_{p_r}(J)), \ell_q(I, \ell_p(J)), \mathcal{B}\Big) \le c\, m^{-r}\,, \qquad m \in \mathbb{N}\,, \tag{22}$$

where

$$\frac{1}{p_r} := r + \frac{1}{p} \qquad \textit{and} \qquad \frac{1}{q_r} := r + \frac{1}{q}\,. \tag{23}$$

Moreover, the constant c *does neither depend on* I *nor on* J.

Proof. Assume $0 < p, q < \infty$, and let $a \in \ell_{q_r}(I, \ell_{p_r}(J))$. Furthermore, we shall use the notation $a^i = (a_{i,j})_{j \in J_i}$, $i \in I$. Let

$$K_l(\varepsilon) := \Big\{ i \in I : \quad 2^l \varepsilon < \big\| a^i \big| \ell_{p_r}(J_i) \big\| \le 2^{l+1}\varepsilon \Big\}, \quad l \in \mathbb{Z}.$$

Obviously, the $(K_l)_l$ form a pairwise disjoint covering of I. By Prop. 7 we already know $\ell_{p_r}(J_i) = \mathcal{A}^r_{p_r}\big(\ell_p(J_i)\big) \hookrightarrow \mathcal{A}^r_\infty\big(\ell_p(J_i)\big)$ and hence

$$\sigma_m(a^i, \ell_p(J_i), \mathcal{B}) \le c\, m^{-r} \big\| a^i \big| \ell_{p_r}(J_i) \big\|\,,$$

but we shall present another more explicit version of this result.

Step 1. We fix some $i \in I$. We shall prove

$$\sigma_m(a^i, \ell_p(J_i), \mathcal{B}) = \left(\sum_{n=m+1}^{\infty} \Big((a^i)^*_n \Big)^p \right)^{1/p} \le (m+1)^{-r} \big\| a^i \big| \ell_{p_r}(J_i) \big\|\,, \tag{24}$$

where $(a^i)^*$ denotes the non-increasing rearrangement of a^i. An easy calculation shows

$$\sigma_{n-1}(a^i, \ell_\infty(J_i), \mathcal{B}) = (a^i)^*_n \le n^{-1/p_r} \big\| a^i \big| \ell_{p_r}(J_i) \big\|\,, \qquad n \in \mathbb{N}\,, \tag{25}$$

and hence

$$\begin{aligned}
\left(\sum_{n=m+1}^{\infty} \Big((a^i)^*_n \Big)^p \right)^{1/p} &= \left(\sum_{n=m+1}^{\infty} \Big((a^i)^*_n \Big)^{p-p_r} \Big((a^i)^*_n \Big)^{p_r} \right)^{1/p} \\
&\le \left(\sum_{n=m+1}^{\infty} \Big(n^{-1/p_r} \big\| a^i \big| \ell_{p_r}(J_i) \big\| \Big)^{p-p_r} \Big((a^i)^*_n \Big)^{p_r} \right)^{1/p} \\
&\le (m+1)^{-\frac{1}{p_r}+\frac{1}{p}} \big\| a^i \big| \ell_{p_r}(J_i) \big\|^{1-\frac{p_r}{p}} \left(\sum_{n=m+1}^{\infty} \Big((a^i)^*_n \Big)^{p_r} \right)^{1/p} \\
&\le (m+1)^{-r} \big\| a^i \big| \ell_{p_r}(J_i) \big\|^{1-\frac{p_r}{p}} \big\| a^i \big| \ell_{p_r}(J_i) \big\|^{\frac{p_r}{p}}\,.
\end{aligned}$$

This proves (24). Further, we wish to mention that for every $m \in \mathbb{N}$ there is a set $\Lambda_i^m \subset J_i$ of cardinality at most m, such that

$$\sigma_m(a^i, \ell_p(J_i), \mathcal{B}) = \left(\sum_{n=m+1}^{\infty} \left((a^i)_n^* \right)^p \right)^{1/p} = \left(\sum_{j \not\in \Lambda_i^m} |a_{i,j}|^p \right)^{1/p} . \qquad (26)$$

Step 2. We shall use these sets Λ_i^m to define a suitable approximation of a. We define

$$T_\varepsilon a^i := \sum_{j \in \Lambda_i^{[2^{lq_r}]}} a_{i,j} e^{i,j} , \qquad i \in I ,$$

and

$$S_\varepsilon a := \sum_{l=0}^{\infty} \sum_{i \in K_l(\varepsilon)} T_\varepsilon a^i .$$

Then $S_\varepsilon a$ is a m_ε-term approximation with

$$m_\varepsilon \le \sum_{l=0}^{\infty} |K_l(\varepsilon)| \, 2^{lq_r} .$$

Furthermore, by definition of the sets $K_l(\varepsilon)$, it holds

$$\begin{aligned} \left\| a \middle| \ell_{q_r}(I, \ell_{p_r}(J)) \right\|^{q_r} &= \sum_{l \in \mathbb{Z}} \sum_{i \in K_l(\varepsilon)} \left\| a^i \middle| \ell_{p_r}(J_i) \right\|^{q_r} \qquad (27) \\ &\ge \sum_{l \in \mathbb{Z}} \sum_{i \in K_l(\varepsilon)} \left(2^l \varepsilon \right)^{q_r} \ge \varepsilon^{q_r} \sum_{l=0}^{\infty} |K_l(\varepsilon)| \, 2^{lq_r} , \end{aligned}$$

and hence

$$m_\varepsilon \le \varepsilon^{-q_r} \left\| a \middle| \ell_{q_r}(I, \ell_{p_r}(J)) \right\|^{q_r} .$$

Step 3. Using the definition of $S_\varepsilon a$ we find

$$\left\| a - S_\varepsilon a \middle| \ell_q(I, \ell_p(J)) \right\|^q = \sum_{l=-\infty}^{-1} \sum_{i \in K_l(\varepsilon)} \left\| a^i \middle| \ell_p(J_i) \right\|^q + \sum_{l=0}^{\infty} \sum_{i \in K_l(\varepsilon)} \left\| a^i - T_\varepsilon a^i \middle| \ell_p(J_i) \right\|^q . \qquad (28)$$

Due to the embedding $\ell_{p_r}(J_i) \hookrightarrow \ell_p(J_i)$, it follows for the first sum in (28)

$$\begin{aligned} \sum_{l=-\infty}^{-1} \sum_{i \in K_l(\varepsilon)} \left\| a^i \middle| \ell_p(J_i) \right\|^q &\le \sum_{l=-\infty}^{-1} \sum_{i \in K_l(\varepsilon)} \left\| a^i \middle| \ell_{p_r}(J_i) \right\|^q \le \sum_{l=-\infty}^{-1} |K_l(\varepsilon)| \left(2^{l+1} \varepsilon \right)^q \\ &\le \varepsilon^{q-q_r} \, 2^{q_r} \sum_{l=-\infty}^{-1} |K_l(\varepsilon)| \left(2^l \varepsilon \right)^{q_r} \le \varepsilon^{q-q_r} \, 2^{q_r} \left\| a \middle| \ell_{q_r}(I, \ell_{p_r}(J)) \right\|^{q_r} , \end{aligned}$$

see (27). For the second sum in (28), we use (26) and find by (24), (27) and $1-rq_r = \frac{q_r}{q}$

$$\begin{aligned}
\sum_{l=0}^{\infty} \sum_{i\in K_l(\varepsilon)} \big\|a^i - T_\varepsilon a^i \big| \ell_p(J_i)\big\|^q &= \sum_{l=0}^{\infty} \sum_{i\in K_l(\varepsilon)} (\sigma_{[2^{lq_r}]}(a^i, \ell_p(J_i), \mathcal{B}))^q \\
&\le \sum_{l=0}^{\infty} \sum_{i\in K_l(\varepsilon)} \big([2^{lq_r}]+1\big)^{-rq} \big\|a^i\big|\ell_{p_r}(J_i)\big\|^q \\
&\le \sum_{l=0}^{\infty} |K_l(\varepsilon)|\, 2^{-lq_r rq} \left(2^{l+1}\varepsilon\right)^q \\
&\le 2^q\, \varepsilon^{q-q_r} \sum_{l=0}^{\infty} |K_l(\varepsilon)|\, 2^{lq_r} \varepsilon^{q_r} \\
&\le 2^q\, \varepsilon^{q-q_r} \big\|a\big|\ell_{q_r}(I, \ell_{p_r}(J))\big\|^{q_r} .
\end{aligned}$$

Altogether, we have proved

$$\begin{aligned}
\sigma_{m_\varepsilon}(a, \ell_q(I, \ell_p(J)), \mathcal{B})^q &\le \big\|a - S_\varepsilon a \big| \ell_q(I, \ell_p(J))\big\|^q \\
&\le \Big(2^{q_r} + 2^q\Big)\, \varepsilon^{q-q_r} \big\|a\big|\ell_{q_r}(I, \ell_{p_r}(J))\big\|^{q_r} .
\end{aligned}$$

Choosing $\varepsilon = m^{-1/q_r}\big\|a\big|\ell_{q_r}(I, \ell_{p_r}(J))\big\|$, we obtain $m_\varepsilon \le m$ and

$$\sigma_m\big(a, \ell_q\big(I, \ell_p(J)\big), \mathcal{B}\big) \le c(q,r)\, m^{-r} \big\|a\big|\ell_{q_r}(I, \ell_{p_r}(J))\big\| .$$

Step 4. It remains to study the case $\max(p,q) = \infty$. The restriction $p < \infty$ is used only in Step 1. If $p = \infty$, then we simply apply (25). The restriction $q < \infty$ has been used in Step 3. Here standard modifications can be used to prove (22). ■

Remark 12. We wish to mention that the basic ideas of the proof are picked up from the proof of a corresponding result for homogeneous isotropic Besov spaces in Kyriazis [17].

We continue with two consequences of Thm. 4. The first one is a generalization.

Corollary 3. *Let $0 < p, q \le \infty$ and $r > 0$. Then there exists a positive constant c, such that*

$$\sigma_m\Big(\ell_{q_r}(I, \ell_{p_r,\infty}(J)), \ell_q(I, \ell_p(J)), \mathcal{B}\Big) \le c\, m^{-r}, \qquad m \in \mathbb{N},$$

where p_r and q_r are as in (23). Moreover, the constant c does neither depend on I nor on J.

Proof. In comparison with the above given proof only very few modifications are necessary. Of course this time we use

$$K_l(\varepsilon) := \Big\{i \in I : \quad 2^l\varepsilon < \big\|a^i\big|\ell_{p_r,\infty}(J_i)\big\| \le 2^{l+1}\varepsilon\Big\}, \quad l \in \mathbb{Z}.$$

The second modification concerns (24). This time we get

$$\begin{aligned}\Big(\sum_{n=m+1}^{\infty}\Big((a^i)_n^*\Big)^p\Big)^{1/p} &\le \Big(\sum_{n=m+1}^{\infty}\Big(n^{-1/p_r}\big\|a^i\big|\ell_{p_r,\infty}(J_i)\big\|\Big)^p\Big)^{1/p}\\ &\le \Big(\int_m^\infty x^{-p/p_r}\,dx\Big)^{1/p}\,\big\|a^i\big|\ell_{p_r,\infty}(J_i)\big\|\\ &\le (pr)^{-1/p}\,m^{-r}\,\big\|a^i\big|\ell_{p_r,\infty}(J_i)\big\|\,.\end{aligned}$$

All the formulas in Step 2 remain true by replacing $\ell_{p_r}(J_i)$ by $\ell_{p_r,\infty}(J_i)$. In Step 3 we use the continuous embedding $\ell_{p_r,\infty}(J_i) \hookrightarrow \ell_p(J_i)$. Here we have to notice that the norm of the mapping $\mathrm{Id}:\ \ell_{p_r}(J_i) \longrightarrow \ell_p(J_i)$ is not 1, however, uniformly bounded in i. No further arguments are needed. ■

For the next conclusion we need a slight restriction on the index sets I and J.

Corollary 4. *Let $0 < p_0 \le p_1 \le \infty$ and $0 < q_0 \le q_1 \le \infty$. Let I be an infinite index set, and let the family $J = (J_i)_{i\in I}$ of index sets fulfil*

$$\sup_{i\in I} |J_i| = \infty\,.$$

Furthermore, let

$$r = \min\Big(\frac{1}{p_0} - \frac{1}{p_1}\,,\,\frac{1}{q_0} - \frac{1}{q_1}\Big)\,.$$

Then

$$\sigma_m\Big(\ell_{q_0}\big(I,\ell_{p_0}(J)\big),\ell_{q_1}\big(I,\ell_{p_1}(J)\big),\mathcal{B}\Big) \asymp m^{-r}\,,$$

where the constants of equivalence do not depend on I or J.

Proof. *Step 1.* Estimates from above. We shall distinguish into two case, namely $r = 0$ and $r > 0$.

Substep 1.1. Let $r = 0$. Under the given restrictions we have $\ell_{q_0}\big(I,\ell_{p_0}(J)\big) \hookrightarrow \ell_{q_1}\big(I,\ell_{p_1}(J)\big)$, where the embedding operator has norm 1. Thus, also

$$\sigma_m\Big(\ell_{q_0}\big(I,\ell_{p_0}(J)\big),\ell_{q_1}\big(I,\ell_{p_1}(J)\big),\mathcal{B}\Big) \le 1$$

is true.

Substep 1.2. Let $r > 0$. We split our considerations into three cases. If $r = \frac{1}{p_0} - \frac{1}{p_1} = \frac{1}{q_0} - \frac{1}{q_1}$, then the result follows immediately from the previous theorem. If $r = \frac{1}{p_0} - \frac{1}{p_1} < \frac{1}{q_0} - \frac{1}{q_1}$, then we define q_* by

$$\frac{1}{q_*} := \frac{1}{p_0} - \frac{1}{p_1} + \frac{1}{q_1}\,.$$

Hence $q_0 < q_*$, and by the monotonicity of the ℓ_p-norms we have $\ell_{q_0}\big(I,\ell_{p_0}(J)\big) \hookrightarrow \ell_{q_*}\big(I,\ell_{p_0}(J)\big)$. But as $r = \frac{1}{p_0} - \frac{1}{p_1} = \frac{1}{q_*} - \frac{1}{q_1}$, the desired estimate now follows from the first case.

If finally $r = \frac{1}{q_0} - \frac{1}{q_1} < \frac{1}{p_0} - \frac{1}{p_1}$, then we define p_* by $\frac{1}{p_*} = \frac{1}{q_0} - \frac{1}{q_1} + \frac{1}{p_1}$. Hence $p_0 < p_*$, $\ell_{q_0}\big(I, \ell_{p_0}(J)\big) \hookrightarrow \ell_{q_0}\big(I, \ell_{p_*}(J)\big)$ and we may argue as before.

Step 2. For the estimates from below we shall discuss two types of sequences. The first one is defined by

$$a^m := \sum_{j \in \Lambda_m} e^{i,j}\,, \qquad \text{i.e.} \qquad a^m_{i,j} := \begin{cases} 1, & i = i_m,\ j \in \Lambda_m, \\ 0, & \text{else}, \end{cases} \tag{29}$$

where i_m is chosen such that $|J_{i_m}| \geq 2m$ and $\Lambda_m \subset J_{i_m}$ is arbitrary satisfying $|\Lambda_m| = 2m$. Clearly,

$$\big\| a^m | \ell_{q_0}\big(I, \ell_{p_0}(J)\big)\big\| = (2m)^{1/p_0}\,.$$

Due to the special structure of the sequences the best m-term approximation is easy to determine. Any partial sum x^m with exactly m terms is optimal. Hence, we find

$$\big\| a^m - x^m | \ell_{q_1}\big(I, \ell_{p_1}(J)\big)\big\| = m^{1/p_1}\,.$$

This finally gives the estimate

$$\sigma_m\Big(\ell_{q_0}\big(I, \ell_{p_0}(J)\big), \ell_{q_1}\big(I, \ell_{p_1}(J)\big), \mathcal{B}\Big) \geq 2^{-1/p_0}\, m^{-1/p_0 + 1/p_1}\,.$$

Now we turn to the second example. We choose $i_l \in I$ pairwise distinct and $j_l \in J_{i_l}$, $l \in \mathbb{N}$. Define sequences b^m by

$$b^m_{i,j} := \begin{cases} 1, & i = i_l,\ j = j_l, \quad 1 \leq l \leq m, \\ 0, & \text{else}\,. \end{cases} \tag{30}$$

Then we find

$$\big\| b^m | \ell_{q_0}\big(I, \ell_{p_0}(J)\big)\big\| = m^{1/q_0}$$

as well as

$$\big\| b^{2m} - b^m | \ell_{q_1}\big(I, \ell_{p_1}(J)\big)\big\| = m^{1/q_1}\,.$$

Of course, b^m is a best m-term approximation for b^{2m}. Hence, we obtain

$$\sigma_m\Big(\ell_{q_0}\big(I, \ell_{p_0}(J)\big), \ell_{q_1}\big(I, \ell_{p_1}(J)\big), \mathcal{B}\Big) \geq 2^{-1/q_0}\, m^{-1/q_0 + 1/q_1}\,.$$

This proves the claim. ∎

Remark 13. One can prove even more than stated in Thm. 4 and Cor. 3. Let $0 < p, q \leq \infty$, $r > 0$ and p_r, q_r as before. Then

$$\mathcal{A}^r_{q_r}\big(\ell_q(I, \ell_p(J)), \mathcal{B}\big) = \ell_{q_r}\big(I, \ell_{p_r, q_r}(J)\big)\,.$$

We omit details and refer to [11]. A closely connected formula has been proved in [16].

5 Widths of best m-term approximation for embeddings of $s^t_{p,q}x(\nabla)$-spaces

Within this section we deal with the behaviour of the widths of best m-term approximation with respect to pairs $(s^t_{p_0,q_0}x(\nabla), s^0_{p_1,q_1}y(\nabla))$, $x, y \in \{b, f\}$.

5.1 Estimates from below

Let Ω be a bounded open and nontrivial subset of $\mathbb{R}^d$ and let $\nabla = \nabla(\Omega)$ be the associated subset of $\mathbb{N}_0^d \times \mathbb{Z}^d$, see (19). We shall use the following abbreviations. For $\mu \in \mathbb{N}$ we define

$$\begin{aligned} M\big(\mu, d\big) &:= \big\{ j \in \mathbb{N}_0^d :\ |j|_1 = \mu \big\}\,, \\ \nabla_\mu &:= \Big\{ (j,k) \in \mathbb{N}_0^d \times \mathbb{Z}^d :\ j \in M\big(\mu, d\big),\ k \in \nabla_j \Big\}\,, \\ S\big(\mu, d\big) &:= |M\big(\mu, d\big)| \qquad \text{and} \qquad D_\mu := |\nabla_\mu|\,. \end{aligned}$$

Obviously, $S\big(\mu, d\big) \asymp \mu^{d-1}$ and from (20) we conclude $D_\mu \asymp S\big(\mu, d\big) 2^\mu \asymp \mu^{d-1} 2^\mu$. Furthermore,

$$\mathcal{B} := \Big\{ e^{j,k} :\quad j \in \mathbb{N}_0^d,\ k \in \nabla_j \Big\}$$

denotes the canonical basis with respect to ∇.

Lemma 3. *Let $t \in \mathbb{R}$. With $x, y \in \{b, f\}$ we have: for each m there exists a finite sequence a such that $\|\, a\,|s^t_{p_0,q_0}x(\Omega)\| = 1$ and*

$$\sigma_m\Big(a, s^0_{p_1,q_1}y(\Omega), \mathcal{B}\Big) \gtrsim m^{-t}\big(\log m\big)^{(d-1)(t-\frac{1}{q_0}+\frac{1}{q_1})_+}\,, \qquad m \in \mathbb{N}.$$

Here the constant behind $\gtrsim$ does not depend on m.

Proof. *Step 1.* Due to the monotonicity properties of σ_m it is sufficient to consider $m = 2^M$ for some $M \in \mathbb{N}_0$. Let $j \in \mathbb{N}_0^d$ be some arbitrary vector, such that $|j|_1 = M$. Let C_1 and C_2 be the constants in (20). Additionally, let K be the smallest natural number such that $C_1 2^K \geq 2$. This means that $|\nabla_{j+Ke^1}| \geq 2m$. For brevity we put $\square := \nabla_{j+Ke^1}$. Now consider the finite sequence

$$a := \sum_{k \in \square} |\square|^{-1/p_0}\, 2^{-(M+K)(t+\frac{1}{2}-\frac{1}{p_0})}\, e^{j+Ke^1,k}\,.$$

Consequently we find

$$\big\|\, a\,|s^t_{p_0,q_0}b(\Omega)\big\| = \big\|\, a\,|s^t_{p_0,q_0}f(\Omega)\big\| = 1 \qquad \text{for any} \quad 0 < q_0 \leq \infty\,.$$

Now, let $\Gamma \subset \square$ s.t. $|\Gamma| = m$. Then, for all q_1, we obtain

$$
\begin{aligned}
\sigma_m\big(a, s^0_{p_1,q_1}b(\Omega), \mathcal{B}\big) &= \Big\| \sum_{k\in\square\setminus\Gamma} |\square|^{-1/p_0}\, 2^{-(M+K)(t+\frac{1}{2}-\frac{1}{p_0})}\, e^{j+Ke^1,k} \Big| s^0_{p_1,q_1}b(\Omega)\Big\| \\
&= |\square|^{-1/p_0}\, 2^{-(M+K)(t-\frac{1}{p_0}+\frac{1}{p_1})}\, |\square\setminus\Gamma|^{1/p_1} \\
&\geq \big(C_2 2^{M+K}\big)^{-1/p_0}\, 2^{-(M+K)(t-\frac{1}{p_0}+\frac{1}{p_1})}\, \big(2^M(C_1 2^K - 1)\big)^{1/p_1} \\
&\geq C_2^{-1/p_0}\, 2^{-K(t+1/p_1)}\, 2^{-Mt} \\
&\gtrsim m^{-t}\,.
\end{aligned}
$$

In addition we observe $\sigma_m\big(a, s^0_{p_1,q_1}f(\Omega), \mathcal{B}\big) = \sigma_m\big(a, s^0_{p_1,q_1}b(\Omega), \mathcal{B}\big)$. This proves the claim if $t - \frac{1}{q_0} + \frac{1}{q_1} \leq 0$.

Step 2. For a real number x we denote by $[x]$ its integer part. Again we concentrate on a subsequence with respect to m. Let $m = \big[\frac{D_\mu}{2}\big]$ for some $\mu \in \mathbb{N}$. This time our test sequence is given by

$$
\beta := S\big(\mu, d\big)^{-1/q_0} \sum_{(j,k)\in\nabla_\mu} 2^{-|j|_1(t+\frac{1}{2})}\, e^{j,k}\,. \tag{31}
$$

We immediately find $\big\|\beta\big|s^t_{p_0,q_0}b(\Omega)\big\| = 1$. Let $\gamma_1 := \min(p_1, q_1)$. Then, for any set $\Gamma \subset \nabla_\mu$, satisfying $|\Gamma| = m$, we obtain

$$
\begin{aligned}
&\Big\| S\big(\mu, d\big)^{-1/q_0} \sum_{(j,k)\in\nabla_\mu\setminus\Gamma} 2^{-|j|_1(t+\frac{1}{2})}\, e^{j,k} \Big| s^0_{p_1,q_1}b(\Omega)\Big\| \\
&\qquad = \; 2^{-\mu(t+\frac{1}{p_1})}\, S\big(\mu, d\big)^{-1/q_0} \Big(\sum_{j\in M(\mu,d)} |\{k\,:\,(j,k)\in\nabla_\mu\setminus\Gamma\}|^{q_1/p_1}\Big)^{1/q_1}.
\end{aligned}
$$

Next we employ Hölder's inequality and find

$$
\begin{aligned}
&\Big(\sum_{j\in M(\mu,d)} |\{k\,:\,(j,k)\in\nabla_\mu\setminus\Gamma\}|\Big)^{1/\gamma_1} \\
&\qquad\lesssim\; 2^{\mu(\frac{1}{\gamma_1}-\frac{1}{p_1})} \Big(\sum_{j\in M(\mu,d)} |\{k\,:\,(j,k)\in\nabla_\mu\setminus\Gamma\}|^{\gamma_1/p_1}\Big)^{1/\gamma_1} \\
&\qquad\lesssim\; 2^{\mu(\frac{1}{\gamma_1}-\frac{1}{p_1})}\, S\big(\mu, d\big)^{\frac{1}{\gamma_1}-\frac{1}{q_1}} \Big(\sum_{j\in M(\mu,d)} |\{k\,:\,(j,k)\in\nabla_\mu\setminus\Gamma\}|^{q_1/p_1}\Big)^{1/q_1}.
\end{aligned}
$$

Inserting this into the previous identity we get

$$
\begin{aligned}
\sigma_m\big(\beta, s^0_{p_1,q_1}b(\Omega), \mathcal{B}\big) \;&\gtrsim\; 2^{-\mu(t+1/p_1)} S\big(\mu, d\big)^{-1/q_0}\, 2^{-\mu(1/\gamma_1-1/p_1)} S\big(\mu, d\big)^{-(1/\gamma_1-1/q_1)} \\
&\qquad\times\; \Big(\sum_{j\in M(\mu,d)} |\{k\,:\,(j,k)\in\nabla_\mu\setminus\Gamma\}|\Big)^{1/\gamma_1} \\
&=\; 2^{-\mu(t+1/\gamma_1)} S\big(\mu, d\big)^{-1/q_0-1/\gamma_1+1/q_1} |\nabla_\mu\setminus\Gamma|^{1/\gamma_1} \\
&\geq\; 2^{-\mu(t+1/\gamma_1)} S\big(\mu, d\big)^{-1/q_0-1/\gamma_1+1/q_1} \big(\tfrac{1}{2}D_\mu\big)^{1/\gamma_1} \\
&\gtrsim\; 2^{-\mu(t+1/\gamma_1)} S\big(\mu, d\big)^{-1/q_0-1/\gamma_1+1/q_1} S\big(\mu, d\big)^{1/\gamma_1} 2^{\mu/\gamma_1} \\
&=\; 2^{-\mu t} S\big(\mu, d\big)^{-1/q_0+1/q_1}.
\end{aligned}
$$

By taking into account $S(\mu,d) \asymp \mu^{d-1}$ and $m \asymp \mu^{d-1}\, 2^\mu$ this yields

$$\sigma_m\big(\beta, s^0_{p_1,q_1}b(\Omega), \mathcal{B}\big) \gtrsim m^{-t}\,\big(\log m\big)^{(d-1)(t-\frac{1}{q_0}+\frac{1}{q_1})}\,.$$

Together with Step 1 this proves the claim for $x=y=b$.

Step 3. We deal with $x=y=f$. Again we work with the finite sequence β defined in (31).

Substep 3.1. We claim $\big\|\,\beta\,|s^t_{p_0,q_0}f(\Omega)\big\| \asymp 1$ (with constant independent of m). First we consider the model case

$$\nabla_j := \{k \in \mathbb{Z}^d : \quad 0 \le k_\ell < 2^{j_\ell}\,,\ 1 \le \ell \le d\}\,, \qquad j \in \mathbb{N}_0^d\,.$$

Then

$$\sum_{k\in\nabla_j} \mathcal{X}_{j,k}(x) = \begin{cases} 1 & \text{if} \quad x \in [0,1]^d\,,\\ 0 & \text{otherwise}\,, \end{cases} \qquad \text{a.e.}$$

Hence

$$\begin{aligned} \|\,\beta\,|s^t_{p_0,q_0}f(\Omega)\| &= S\big(\mu,d\big)^{-1/q_0}\Big(\int\Big(\sum_{(j,k)\in\nabla_\mu}\mathcal{X}_{j,k}(x)\Big)^{p_0/q_0}dx\Big)^{1/p_0}\\ &= S\big(\mu,d\big)^{-1/q_0}\Big(\int_{[0,1]^d}|S\big(\mu,d\big)|^{p_0/q_0}\,dx\Big)^{1/p_0} = 1\,. \end{aligned}$$

Now we turn to the case of a general bounded open set Ω. We may assume that the origin is contained in Ω. Then there are two cubes Q_0 and Q_1 s.t. $Q_0 \subset \Omega \subset Q_1$, $|Q_i| = 2^{K_i d}$, $i=0,1$, for some $K_i \in \mathbb{Z}$ and

$$\begin{aligned} \{k \in \mathbb{Z}^d : \quad 0 \le k_\ell < 2^{j_\ell+K_0}\,,\ 1\le \ell\le d\} \ &\subset\ \nabla_j(\Omega)\\ &\subset\ \{k \in \mathbb{Z}^d : \quad 0 \le k_\ell < 2^{j_\ell+K_1}\,,\ 1\le \ell\le d\}\,. \end{aligned}$$

Then, by arguing as in the model case, we obtain

$$2^{K_0 d/q_0} \le \|\,\beta\,|s^t_{p_0,q_0}f(\Omega)\| \le 2^{K_1 d/q_0}\,.$$

Substep 3.2. Let $\gamma_1 := \min(p_1,q_1)$ as in Step 2. Also the set Γ is chosen as in Step 2. We put

$$\Gamma_j := \{k\ :\ (j,k) \in \nabla_\mu \setminus \Gamma\}\,, \qquad |j|_1 = \mu\,,$$

and

$$f_j(x) := \sum_{k\in\Gamma_j}\mathcal{X}_{j,k}(x)\,, \qquad j \in M(\mu,d)\,.$$

Obviously, $\sum_{j\in M(\mu,d)} f_j(x) \le S(\mu,d)$ a.e.. This, together with Hölder's inequality, implies

$$\begin{aligned} \Big(\int\sum_{|j|_1=\mu} f_j(x)\,dx\Big)^{1/\gamma_1} &\lesssim S\big(\mu,d\big)^{1/\gamma_1-1/q_1}\Big(\int\Big(\sum_{|j|_1=\mu} f_j^{q_1}(x)\Big)^{\gamma_1/q_1}dx\Big)^{1/\gamma_1}\\ &\lesssim S\big(\mu,d\big)^{1/\gamma_1-1/q_1}\,|\Omega|^{1/\gamma_1-1/p_1}\Big(\int\Big(\sum_{|j|_1=\mu} f_j(x)\Big)^{p_1/q_1}dx\Big)^{1/p_1}\,. \end{aligned}$$

Since $\left(\int \sum_{|j|_1=\mu} \sum_{k\in\Gamma_j} \mathcal{X}_{j,k}(x)\,dx\right)^{1/\gamma_1} \asymp \mu^{(d-1)/\gamma_1}$ we conclude

$$\begin{aligned} \sigma_m\big(\beta, s^0_{p_1,q_1}f(\Omega), \mathcal{B}\big) &\gtrsim 2^{-\mu t}\, S\big(\mu, d\big)^{-1/q_0+1/q_1-1/\gamma_1}\, \mu^{(d-1)/\gamma_1} \\ &= 2^{-\mu t}\, S\big(\mu, d\big)^{-1/q_0+1/q_1}\,. \end{aligned}$$

This proves the claim if $x = y = f$.

Step 4. For the mixed situations we simply have to combine Step 2 and Step 3. The proof is complete. ■

By obvious modifications in case $\Omega = \mathbb{R}^d$ Lemma 8 yields the following.

Proposition 8. *Let Ω be either a bounded open and nontrivial subset of $\mathbb{R}^d$ or $\Omega = \mathbb{R}^d$. We suppose $s^t_{p_0,q_0}x(\Omega) \hookrightarrow s^0_{p_1,q_1}y(\Omega)$. With $x, y \in \{b, f\}$ it follows*

$$\sigma_m\Big(s^t_{p_0,q_0}x(\Omega), s^0_{p_1,q_1}y(\Omega), \mathcal{B}\Big) \gtrsim m^{-t}\big(\log m\big)^{(d-1)(t-\frac{1}{q_0}+\frac{1}{q_1})_+}\,, \qquad m \in \mathbb{N}.$$

5.2 The widths of best m-term approximation for non-compact embeddings

Our approach is based on a generalization of an inequality due to Wojtaszczyk [35], which itself has generalized a well-known inequality of Temlyakov [29].

5.2.1 Some basic inequalities

We need some further notations. The set of all dyadic cubes Q with $|Q| \le 1$ will be denoted by $\mathcal{D}^*$, i.e.

$$\mathcal{D}^* := \Big\{Q = 2^{-j}\big([0,1]^d + k\big)\ :\ j \in \mathbb{N}_0\,, k \in \mathbb{Z}^d\Big\}\,.$$

By $\chi^{(p)}_Q$ we denote the p-normalized characteristic function of Q, i.e., $\chi^{(p)}_Q = |Q|^{-1/p}\,\chi_Q$. Then the following result is well-known, see Lemmas 2.1, 2.2 in [29], Lemma 1 in [6] or Theorem 11.2 in [15].

Lemma 4. *Let $0 < p < \infty$ and let $\Lambda \subset \mathcal{D}^*$ be a set with $|\Lambda| = m$. Then it holds*

$$\Bigg\| \sum_{Q\in\Lambda} \chi^{(p)}_Q \Bigg| L_p(\mathbb{R}^d)\Bigg\| \asymp m^{1/p}\,.$$

To prove the generalization we have in mind we need to consider a different set of rectangles. Let

$$\mathcal{D} := \Big\{Q_{j,k}\ :\quad j \in \mathbb{N}^d_0\,, k \in \mathbb{Z}^d\Big\}\,,$$

see (10). If necessary, we shall indicate the dimension by writing $\mathcal{D}(d)$. Further we shall use the following abbreviation

$$\|\, a \,\| := \left(\int_{\mathbb{R}^d} \Big(\sum_{Q\in\mathcal{D}} \big| a_Q \, \chi_Q^{(p)}(s) \big|^q \Big)^{p/q} ds \right)^{1/p} = \|\, a \,| s_{p,q}^{\frac{1}{p}-\frac{1}{2}} f \| \, .$$

The first step toward the desired generalization of Lemma 4 is the following estimate for finite sequences, i.e. sequences with only finitely many nonvanishing components a_Q.

Lemma 5. *Let a be a finite sequence, $a = \sum_{Q\in\Lambda} a_Q e^Q$ with $|\Lambda| = m \geq 2$. Then it holds for $0 < p \leq q < \infty$*

$$\big(\log m\big)^{d(\frac{1}{q}-\frac{1}{p})} \Big(\sum_{Q\in\Lambda} |a_Q|^p \Big)^{1/p} \lesssim \|\, a \,\| \leq \Big(\sum_{Q\in\Lambda} |a_Q|^p \Big)^{1/p} , \tag{32}$$

and for $0 < q \leq p < \infty$ we obtain

$$\Big(\sum_{Q\in\Lambda} |a_Q|^p \Big)^{1/p} \leq \|\, a \,\| \lesssim \ \big(\log m\big)^{d(\frac{1}{q}-\frac{1}{p})} \Big(\sum_{Q\in\Lambda} |a_Q|^p \Big)^{1/p} . \tag{33}$$

All occurring constants depend on p, q and d only.

Remark 14. As mentioned above, the proofs of this lemma and the successive proposition follow closely the arguments given in [35].

Proof. *Step 1.* First we consider the case $q = 1$.
Substep 1.1. Let $0 < p \leq 1$. The prove of the right-hand side estimate in (32) follows immediately from the monotonicity of ℓ_p-quasinorms. We obtain

$$\begin{aligned} \|\, a \,\| &= \left(\int_{\mathbb{R}^d} \Big(\sum_{Q\in\Lambda} |a_Q \chi_Q^{(p)}(s)| \Big)^p ds \right)^{1/p} \\ &\leq \left(\int_{\mathbb{R}^d} \sum_{Q\in\Lambda} \Big(|a_Q \chi_Q^{(p)}(s)| \Big)^p ds \right)^{1/p} = \Big(\sum_{Q\in\Lambda} |a_Q|^p \Big)^{1/p} . \end{aligned}$$

Substep 1.2. For the prove of the right hand side inequality in (33) for $1 \leq p < \infty$ we consider the case $d = 1$ first. For the convenience of the reader we repeat the arguments from [35]. Let $\pi : \{1, \ldots, m\} \longrightarrow \Lambda$ be a bijection, such that $|a_{\pi(j)}|$ is a non-increasing sequence. Furthermore, let M be the uniquely determined integer such that $2^{M-1} \leq m < 2^M$, and define

$$g_k := \sum_{j=2^{k-1}}^{2^k-1} |a_{\pi(j)} \chi_{\pi(j)}^{(p)}| , \qquad k = 1, \ldots, M \, .$$

Then the triangle inequality yields

$$\begin{aligned}\|\, a \,\| &= \left\| \sum_{k=1}^{M} g_k \Big| L_p(\mathbb{R}) \right\| \le \sum_{k=1}^{M} \|\, g_k \,|L_p(\mathbb{R})\| = \sum_{k=1}^{M} \left\| \sum_{j=2^{k-1}}^{2^k-1} a_{\pi(j)} e^{\pi(j)} \right\| \\ &\le \sum_{k=1}^{M} \left\| \sum_{j=2^{k-1}}^{2^k-1} a_{\pi(2^{k-1})} e^{\pi(j)} \right\| \lesssim \sum_{k=1}^{M} 2^{(k-1)/p} \big| a_{\pi(2^{k-1})} \big| \,.\end{aligned}$$

The last two estimates follow from the lattice structure of $\|\cdot\|$ and from Lemma 4. On the other hand, we obtain from Hölder's inequality with respect to $1 = \frac{1}{p} + (1 - \frac{1}{p})$

$$\begin{aligned}\sum_{Q\in\Lambda} |a_Q|^p &= \sum_{j=1}^{m} |a_{\pi(j)}|^p \ge \sum_{k=1}^{M} 2^{k-1} |a_{\pi(2^k-1)}|^p \ge M^{1-p} \left(\sum_{k=1}^{M} 2^{(k-1)/p} |a_{\pi(2^k-1)}| \right)^p \\ &\ge 2^{-p} M^{1-p} \left(\sum_{k=1}^{M-1} 2^{(k-1)/p} \big| a_{\pi(2^k)} \big| + \big| a_{\pi(1)} \big| \right)^p \\ &\ge 2^{1-p} M^{1-p} \left(\sum_{k=1}^{M} 2^{(k-1)/p} \big| a_{\pi(2^{k-1})} \big| \right)^p .\end{aligned}$$

Combining both estimates yields

$$\|\, a \,\| \lesssim M^{1-1/p} \left(\sum_{Q\in\Lambda} |a_Q|^p \right)^{1/p} \lesssim \big(\log m\big)^{1-1/p} \left(\sum_{Q\in\Lambda} |a_Q|^p \right)^{1/p} .$$

Substep 1.3. The case $d \ge 2$ will be proven by induction over d. Given a finite set of rectangles $\Lambda \subset \mathcal{D}(d)$, we can rewrite every $Q \in \Lambda$ as $Q = Q' \times Q''$ with $Q' \in \mathcal{D}(1)$ and $Q'' \in \mathcal{D}(d-1)$, and accordingly $\chi_Q^{(p)} = \chi_{Q'}^{(p)} \otimes \chi_{Q''}^{(p)}$. Note that for $|\Lambda| = m \ge 2$ there are at most m different intervals Q' and at most m rectangles Q'' occurring in this way. Then we find

$$\begin{aligned}\|\, a \,\|^p &= \int_{\mathbb{R}} \int_{\mathbb{R}^{d-1}} \left(\sum_{Q=Q'\times Q''\in\Lambda} |a_Q \chi_{Q'}^{(p)}(t)| \cdot |\chi_{Q''}^{(p)}(s)| \right)^p ds\, dt \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}^{d-1}} \left(\sum_{Q''} \left(\sum_{Q' : Q'\times Q''\in\Lambda} |a_Q \chi_{Q'}^{(p)}(t)| \right) |\chi_{Q''}^{(p)}(s)| \right)^p ds\, dt \,.\end{aligned}$$

Now we can apply the induction hypothesis to the $(d-1)$-dimensional integral (the inner sums serve as coefficients for fixed $t \in \mathbb{R}$). In this way we obtain

$$\|\, a \,\|^p \lesssim \big(\log m\big)^{(d-1)(p-1)} \int_{\mathbb{R}} \sum_{Q''} \left(\sum_{Q' : Q'\times Q''\in\Lambda} |a_Q \chi_{Q'}^{(p)}(t)| \right)^p dt \,. \tag{34}$$

At this point we further apply the result for the case $d = 1$. We end up with

$$\begin{aligned}\|\, a \,\|^p &\lesssim \big(\log m\big)^{(d-1)(p-1)} \big(\log m\big)^{(p-1)} \sum_{Q''\in\mathcal{D}(d-1)} \sum_{Q' : Q'\times Q''\in\Lambda} \big| a_Q \big|^p \\ &= \big(\log m\big)^{d(p-1)} \sum_{Q\in\Lambda} \big| a_Q \big|^p .\end{aligned}$$

This proves the right hand side of (33).

Step 2. Now we consider general q. We obtain for $0 < p \le q$ from Step 1, applied to $0 < p/q \le 1$,

$$\begin{aligned}\bigg(\int_{\mathbb{R}^d}\Big(\sum_{Q\in\Lambda}\big|a_Q\chi_Q^{(p)}(s)\big|^q\Big)^{\frac{p}{q}}ds\bigg)^{\frac{1}{p}} &= \bigg(\int_{\mathbb{R}^d}\Big(\sum_{Q\in\Lambda}|a_Q|^q\chi_Q^{(p/q)}(s)\Big)^{\frac{p}{q}}ds\bigg)^{\frac{q}{p}\cdot\frac{1}{q}}\\ &\le \bigg(\sum_{Q\in\Lambda}\big(|a_Q|^q\big)^{p/q}\bigg)^{\frac{q}{p}\cdot\frac{1}{q}} = \bigg(\sum_{Q\in\Lambda}|a_Q|^p\bigg)^{1/p}.\end{aligned}$$

Similarly we find for $0 < q \le p$ from Step 1, applied to $1 \le p/q < \infty$,

$$\begin{aligned}\bigg(\int_{\mathbb{R}^d}\Big(\sum_{Q\in\Lambda}|a_Q\chi_Q^{(p)}(s)|^q\Big)^{\frac{p}{q}}ds\bigg)^{\frac{1}{p}} &\lesssim \bigg(\big(\log m\big)^{d(1-q/p)}\Big(\sum_{Q\in\Lambda}\big(|a_Q|^q\big)^{p/q}\Big)^{\frac{q}{p}}\bigg)^{\frac{1}{q}}\\ &= \big(\log m\big)^{d(1/q-1/p)}\Big(\sum_{Q\in\Lambda}|a_Q|^p\Big)^{1/p}.\end{aligned}$$

Step 3. We prove the estimates from below.

Step 3.1. Preparation. Here we are going to use the following duality statement, see, e.g., [31, Prop. 2.11.1] or [10, 8.20.5]. Let $1 < p, q < \infty$. It holds $g \in \big(L_p(\ell_q)\big)'$ if, and only if, it can be represented uniquely as

$$g(f) = \sum_{k\in\mathbb{N}_0^d}\int_{\mathbb{R}^d} g_k(x)\, f_k(x)dx\,, \qquad f = \big\{f_k\big\}_{k\in\mathbb{N}_0^d} \in L_p(\ell_q)\,,$$

where $\big\{g_k\big\}_{k\in\mathbb{N}_0^d} \in L_{p'}(\ell_{q'})$. Moreover, it holds

$$\big\|g\big\| = \big\|g_k\big|L_{p'}(\ell_{q'})\big\|$$

for the usual operator norm on $\big(L_p(\ell_q)\big)'$. Of course, p', q' are the usual conjugated indices.

Step 3.2. Let $1 < p, q < \infty$. For every finite sequence $a = \sum_{Q\in\Lambda} a_Q\, e^Q$, $|\Lambda| = m$, with $a_Q \neq 0$, $Q \in \Lambda$, define another finite sequence b by $b_Q = \frac{|a_Q|^p}{a_Q}$ for $Q \in \Lambda$ and zero otherwise. Then we have $f_a = \big(a_Q\chi_Q^{(p)}\big)_{Q\in\mathcal{D}} \in L_p(\ell_q)$ and $f_b = \big(b_Q\chi_Q^{(p')}\big)_{Q\in\mathcal{D}} \in L_{p'}(\ell_{q'})$. We begin with the case $1 < q \le p < \infty$ and hence $1 < p' \le q' < \infty$. In view of the above duality relation we find that f_b generates a functional on $L_p(\ell_q)$. Applying the characterization of these functionals we obtain

$$\begin{aligned}\sum_{Q\in\Lambda}|a_Q|^p &= \Big|\sum_{Q\in\Lambda} a_Q b_Q\Big| = \Big|\int_{\mathbb{R}^d}\sum_{Q\in\Lambda}|Q|^{-1}a_Q b_Q\chi_Q(s)ds\Big|\\ &= \Big|\int_{\mathbb{R}^d}\sum_{Q\in\mathcal{D}}\Big(a_Q\chi_Q^{(p)}(s)\Big)\Big(b_Q\chi_Q^{(p')}(s)\Big)ds\Big|\\ &\le \big\|\, f_a\,\big|L_p(\ell_q)\big\|\cdot\big\|\, f_b\,\big|L_{p'}(\ell_{q'})\big\|\,.\end{aligned}$$

Moreover, Step 2 yields

$$\| f_b \,|L_{p'}(\ell_{q'})\|^{p'} = \int_{\mathbb{R}^d} \Big(\sum_{Q\in\Lambda} \Big| b_Q \chi_Q^{(p')}(s) \Big|^{q'} \Big)^{p'/q'} ds \le \sum_{Q\in\Lambda} |b_Q|^{p'} = \sum_{Q\in\Lambda} |a_Q|^p ,$$

where we used $(p-1)p' = p$. Combining both estimates we now conclude

$$\begin{aligned} \| f_a \,|L_p(\ell_q)\| &= \Big(\int_{\mathbb{R}^d} \Big(\sum_{Q\in\Lambda} \Big| a_Q \chi_Q^{(p)}(s) \Big|^{q} \Big)^{p/q} ds \Big)^{1/p} \\ &\ge \| f_b \,|L_{p'}(\ell_{q'})\|^{-1} \sum_{Q\in\Lambda} |a_Q|^p \ge \Big(\sum_{Q\in\Lambda} |a_Q|^p \Big)^{1/p} . \end{aligned}$$

This proves the lower estimate in (33) if $1 < q \le p < \infty$. Now we turn to the condition $1 < p \le q < \infty$. This implies $1 < q' \le p' < \infty$, and due to Step 2 we find

$$\| f_b \,|L_{p'}(\ell_{q'})\|^{p'} \lesssim \big(\log m\big)^{d(1/q'-1/p')p'} \sum_{Q\in\Lambda} |b_Q|^{p'} = \big(\log m\big)^{d(1/p-1/q)p'} \sum_{Q\in\Lambda} |a_Q|^p .$$

Consequently

$$\Big(\int_{\mathbb{R}^d} \Big(\sum_{Q\in\Lambda} \Big| a_Q \chi_Q^{(p)}(s) \Big|^{q} \Big)^{p/q} ds \Big)^{1/p} \ge \big(\log m\big)^{d(1/q-1/p)} \Big(\sum_{Q\in\Lambda} |a_Q|^p \Big)^{1/p} .$$

Substep 3.3. Using a similar argumentation as in Step 2 we remove the restrictions $1 < q \le p < \infty$ and $1 < p \le q < \infty$ for the lower estimates. ■

It is a bit surprising that we can partly improve (32) and (33) by restricting to the sequence $a_Q = 1$, $Q \in \Lambda$, and $a_Q = 0$, $Q \notin \Lambda$.

Proposition 9. *Let Λ be a finite and nonempty subset of $\mathbb{N}_0^d \times \mathbb{Z}^d$. Then it holds for $0 < p \le q < \infty$*

$$\big(1+\log m\big)^{(d-1)(1/q-1/p)} m^{1/p} \lesssim \Big(\int_{\mathbb{R}^d} \Big(\sum_{(j,k)\in\Lambda} (2^{|j|_1/p}\, \chi_{j,k}(s))^q \Big)^{p/q} ds \Big)^{1/p} \lesssim m^{1/p} , \quad (35)$$

and for $0 < q \le p < \infty$ we find

$$m^{1/p} \lesssim \Big(\int_{\mathbb{R}^d} \Big(\sum_{(j,k)\in\Lambda} (2^{|j|_1/p} \chi_{j,k}(s))^q \Big)^{p/q} ds \Big)^{1/p} \lesssim \big(1+\log m\big)^{(d-1)(1/q-1/p)} m^{1/p} . \quad (36)$$

Proof. We proceed as in proof of Lemma 5. First, we have to modify Substep 1.2. Our starting point for the induction is simply Lemma 4. This is enough to prove the estimates from above. The estimates from below can be derived now by the same duality argument as in proof Lemma 5. ■

Remark 15. The above inequalities have been proved in [35] if $q = 2$. Wojtaszczyk considers linear combinations of elements of the tensor product Haar system. But using $|h_{Q_{j,k}}|^2 = \mathcal{X}_{j,k}\, 2^{|j|_1}$ it is easily seen that the set of inequalities, proved in [35, Sect. 4], coincides with (35), (36).

5.2.2 Best m-term widths for embeddings of f-spaces

Theorem 5. *Let $p_0 < p_1$, $q_1 < p_1$ and $p_0 \le q_0 < \infty$. Let $t := \frac{1}{p_0} - \frac{1}{p_1}$. Then we have*

$$\sigma_m\Big(s^t_{p_0,q_0}f(\nabla), s^0_{p_1,q_1}f(\nabla), \mathcal{B}\Big) \asymp m^{-\frac{1}{p_0}+\frac{1}{p_1}} \,(\log m)^{(d-1)(\frac{1}{p_0}-\frac{1}{p_1}-\frac{1}{q_0}+\frac{1}{q_1})}\,,$$

if $m \ge 2$. Moreover, if $q_0 \le p_0 < p_1 \le q_1$ then we find for $m \in \mathbb{N}$

$$\sigma_m\Big(s^t_{p_0,q_0}f(\nabla), s^0_{p_1,q_1}f(\nabla), \mathcal{B}\Big) \sim m^{-\frac{1}{p_0}+\frac{1}{p_1}}\,.$$

Proof. *Step 1.* First we consider the case $q_1 < p_1$ and $p_0 \le q_0$. Let $a \in s^t_{p_0,q_0}f(\nabla)$ s.t. $\big\| \, a \,|s^t_{p_0,q_0}f(\nabla)\big\| = 1$. For $i \in \mathbb{Z}$ and $\mu \in \mathbb{N}_0$ we define

$$\Lambda_{\mu,i} := \big\{ (j,k) \in \nabla_\mu : \quad 2^{-i} < 2^{|j|_1(\frac{1}{2}-\frac{1}{p_1})}\, |a_{j,k}| \le 2^{-i+1} \big\}\,.$$

Obviously,

$$\begin{aligned} 2^{|j|_1(\frac{1}{2}-\frac{1}{p_1})p_0}\, |a_{j,k}|^{p_0} &= 2^{|j|_1(t+\frac{1}{2}-\frac{1}{p_0})p_0}\, |a_{j,k}|^{p_0} \\ &= \int_{\mathbb{R}^d} 2^{|j|_1(t+\frac{1}{2})p_0}\, |a_{j,k}|^{p_0} \chi_{j,k}(x)\, dx \le \|\, a\, |s^t_{p_0,q_0}f(\nabla)\|^{p_0}\,. \end{aligned}$$

Hence $\Lambda_{\mu,i} = \emptyset$ if $i \le 0$. For $M \in \mathbb{N}$ we put

$$\Lambda^M := \bigcup_{i=1}^{M} \Lambda_i\,, \qquad \Lambda_i := \bigcup_{\mu=0}^{M} \Lambda_{\mu,i}\,,$$

and an associated approximation is defined by

$$T_M a := \sum_{(j,k)\in\Lambda^M} a_{j,k}\, e^{j,k}\,.$$

Substep 1.1. Estimates of the cardinality of Λ^M. For $\mu \in \mathbb{N}$ we define the restriction operator R_μ by

$$\big(R_\mu a\big)_{j,k} := \begin{cases} a_{j,k}, & \text{if} \quad j \in M(\mu,d)\,, k \in \nabla_j\,, \\ 0, & \text{else}\,. \end{cases}$$

Then we remark that

$$\big\| \, R_\mu a \,\big| s^r_{p,w}f(\Omega)\big\| \le S(\mu,d)^{\frac{1}{w}-\frac{1}{q}} \big\| \, R_\mu a \,\big| s^r_{p,q}f(\Omega)\big\|\,, \qquad w \le q\,. \tag{37}$$

The latter inequality follows from the observation that for a fixed x the sum $\sum_{k\in\nabla_j} a_{j,k}\,\chi_{j,k}(x)$ consists of exactly one summand. Hence, the cardinality of the summands in $\sum_{|j|_1=\mu}\sum_{k\in\nabla_j} a_{j,k}\,\chi_{j,k}(x)$ is uniformly bounded by $S(\mu,d)$. Hölder's inequality now yields (37). Using this inequality and $p_0 \le q_0$ we find

$$\begin{aligned} |\Lambda_{\mu,i}| &= \sum_{(j,k)\in\Lambda_{\mu,i}} 1 \le \sum_{(j,k)\in\Lambda_{\mu,i}} \bigg(\frac{2^{|j|_1(\frac{1}{2}-\frac{1}{p_1})}\, |a_{j,k}|}{2^{-i}}\bigg)^{p_0} \\ &\le 2^{ip_0} \int_{\mathbb{R}^d} \sum_{(j,k)\in\nabla_\mu} 2^{\mu(\frac{1}{2}+\frac{1}{p_0}-\frac{1}{p_1})p_0}\, |a_{j,k}|^{p_0}\, \chi_{j,k}(x)\, dx \\ &\le 2^{ip_0}\, S\big(\mu,d\big)^{(1/p_0-1/q_0)p_0} \,\big\| \, R_\mu a \,\big| s^{\frac{1}{p_0}-\frac{1}{p_1}}_{p_0,q_0} f(\nabla)\big\|^{p_0} < \infty\,. \end{aligned}$$

Thus, also the sets Λ_i have finite cardinality. Because of $p_0 \le q_0 < \infty$ Prop. 9 in combination with the definition of Λ_i implies

$$
\begin{aligned}
|\Lambda_i| &= |\Lambda_i| \big(1+\log(1+|\Lambda_i|)\big)^{(d-1)(1/q_0-1/p_0)p_0} \big(1+\log(1+|\Lambda_i|)\big)^{(d-1)(1-p_0/q_0)} \\
&\lesssim \big(1+\log(1+|\Lambda_i|)\big)^{(d-1)(1-p_0/q_0)} \int_{\mathbb{R}^d} \Big(\sum_{(j,k)\in\Lambda_i} \big(2^{|j|_1/p_0}\, \chi_{j,k}(x)\big)^{q_0} \Big)^{p_0/q_0} dx \\
&\le \big(1+\log(1+|\Lambda_i|)\big)^{(d-1)(1-\frac{p_0}{q_0})}\, 2^{ip_0} \int_{\mathbb{R}^d} \Bigg(\sum_{(j,k)\in\Lambda_i} \Big(2^{|j|_1(\frac{1}{p_0}-\frac{1}{p_1}+\frac{1}{2})}\, |a_{j,k}|\, \chi_{j,k}(x)\Big)^{q_0} \Bigg)^{\frac{p_0}{q_0}} dx \\
&\le \big(1+\log(1+|\Lambda_i|)\big)^{(d-1)(1-\frac{p_0}{q_0})}\, 2^{ip_0}\, \big\|\, a \,\big|\, s^{\frac{1}{p_0}-\frac{1}{p_1}}_{p_0,q_0} f(\nabla) \big\|^{p_0} .
\end{aligned}
$$

This means, if $|\Lambda_i| \ge 1$, we have

$$
|\Lambda_i| \big(1+\log(1+|\Lambda_i|)\big)^{(d-1)(\frac{p_0}{q_0}-1)} \lesssim 2^{ip_0} ,
$$

which trivially remains true for $\Lambda_i = \emptyset$. For $i \ge 1$ this can be reformulated as

$$
\log(1+|\Lambda_i|) \lesssim i \qquad \text{and} \qquad |\Lambda_i| \lesssim 2^{ip_0}\, i^{(d-1)(1-p_0/q_0)} . \tag{38}
$$

Hence we conclude that $T_M a$ is an m-term approximation of a with

$$
m := \big[c_0 2^{Mp_0} M^{(d-1)(1-p_0/q_0)}\big] . \tag{39}
$$

Substep 1.2. Taking into account $\Lambda_{\mu,i} = \emptyset$ if $i \le 0$, we find

$$
\begin{aligned}
\big\|\, a - T_M a \,\big|\, s^0_{p_1,q_1} f(\nabla) \big\|^{p_1} &= \int \Big(\sum_{i=M+1}^{\infty} \sum_{(j,k)\in\Lambda_i} \big(2^{|j|_1/2}\, |a_{j,k}|\, \chi_{j,k}\big)^{q_1} \Big)^{p_1/q_1} dx \\
&\le 2^{p_1} \int \Big(\sum_{i=M+1}^{\infty} \sum_{(j,k)\in\Lambda_i} \big(2^{|j|_1/p_1}\, 2^{-i}\, \chi_{j,k}\big)^{q_1} \Big)^{p_1/q_1} dx .
\end{aligned}
$$

Because of $p_1 > q_1$ there exists some $\delta > 0$, such that $p_1(q_1-\delta)/q_1 > p_0$. Applying Hölder's inequality with respect to $1 = \frac{p_1-q_1}{p_1} + \frac{q_1}{p_1}$ to the integrand yields

$$
\begin{aligned}
\sum_{i=M+1}^{\infty} \sum_{(j,k)\in\Lambda_i} \big(2^{|j|_1/p_1}\, 2^{-i}\, \chi_{j,k}\big)^{q_1} &= \sum_{i=M+1}^{\infty} 2^{-i\delta} \sum_{(j,k)\in\Lambda_i} 2^{|j|_1 q_1/p_1}\, 2^{-i(q_1-\delta)} \chi_{j,k} \\
&\le \Big(\sum_{i=M+1}^{\infty} 2^{-i\delta\frac{p_1}{p_1-q_1}} \Big)^{\frac{p_1-q_1}{p_1}} \Bigg(\sum_{i=M+1}^{\infty} \Big(\sum_{(j,k)\in\Lambda_i} 2^{|j|_1 q_1/p_1}\, 2^{-i(q_1-\delta)}\, \chi_{j,k} \Big)^{\frac{p_1}{q_1}} \Bigg)^{\frac{q_1}{p_1}} \\
&\lesssim 2^{-M\delta} \Bigg(\sum_{i=M+1}^{\infty} \Big(\sum_{(j,k)\in\Lambda_i} 2^{|j|_1 q_1/p_1}\, 2^{-i(q_1-\delta)}\, \chi_{j,k} \Big)^{\frac{p_1}{q_1}} \Bigg)^{\frac{q_1}{p_1}}
\end{aligned}
$$

Hence, with Prop. 9, (38) and by the choice of δ we finally obtain

$$\begin{aligned}
&\big\| a - T_M a \,|s^0_{p_1,q_1} f(\nabla)\big\|^{p_1} \\
&\lesssim 2^{-M\delta p_1/q_1} \int \sum_{i=M+1}^{\infty} \Bigg(\sum_{(j,k)\in\Lambda_i} 2^{|j|_1 q_1/p_1}\, 2^{-i(q_1-\delta)} \chi_{j,k} \Bigg)^{p_1/q_1} dx \\
&= 2^{-M\delta p_1/q_1} \sum_{i=M+1}^{\infty} 2^{-ip_1(q_1-\delta)/q_1} \int \Bigg(\sum_{(j,k)\in\Lambda_i} 2^{|j|_1 q_1/p_1}\, \chi_{j,k} \Bigg)^{p_1/q_1} dx \\
&\lesssim 2^{-M\delta p_1/q_1} \sum_{i=M+1}^{\infty} 2^{-ip_1(q_1-\delta)/q_1} \Big(|\Lambda_i|^{1/p_1} \big(1+\log|\Lambda_i|\big)^{(d-1)(1/q_1-1/p_1)} \Big)^{p_1} \\
&\lesssim 2^{-M\delta p_1/q_1} \sum_{i=M+1}^{\infty} 2^{-ip_1(q_1-\delta)/q_1} 2^{ip_0}\, i^{(d-1)(1-p_0/q_0)}\, i^{(d-1)(p_1/q_1-1)} \\
&\lesssim 2^{-M\delta \frac{p_1}{q_1}}\, 2^{-M(p_1\frac{q_1-\delta}{q_1}-p_0)} M^{(d-1)(\frac{p_1}{q_1}-\frac{p_0}{q_0})} = 2^{-M(p_1-p_0)}\, M^{(d-1)(\frac{p_1}{q_1}-\frac{p_0}{q_0})}\,.
\end{aligned}$$

A simple calculation and the definition of m, see (39), shows, that the right-hand side, taken to the power $1/p_1$, is equivalent to $m^{-\frac{1}{p_0}+\frac{1}{p_1}} \big(\log m\big)^{(d-1)(\frac{1}{p_0}-\frac{1}{p_1}-\frac{1}{q_0}+\frac{1}{q_1})}$. For the remaining natural numbers m, not covered by (39), we use a monotonicity argument.
Step 2. The case $q_0 \le p_0 < p_1 \le q_1$ follows simply by monotonicity from Cor. 2, see also Rem. 11. Indeed, we have

$$s^t_{p_0,q_0} f(\nabla) \hookrightarrow s^t_{p_0,p_0} f(\nabla) = \mathcal{A}^{\frac{1}{p_0}-\frac{1}{p_1}}_{p_0} \big(s^0_{p_1,p_1} f(\nabla)\,, \mathcal{B}\big) \hookrightarrow \mathcal{A}^{\frac{1}{p_0}-\frac{1}{p_1}}_{\infty} \big(s^0_{p_1,q_1} f(\nabla)\,, \mathcal{B}\big)\,.$$

Step 3. In both cases, the estimates from below are consequences of Prop. 8. ■

Remark 16. The method applied in Step 1 of the proof can be applied also for large q_1. However, it seems that it does not lead to optimal results in general. Without going into detail we mention

$$\sigma_m\Big(s^t_{p_0,q_0} f(\nabla), s^0_{p_1,q_1} f(\nabla), \mathcal{B}\Big) \lesssim m^{-\frac{1}{p_0}+\frac{1}{p_1}} \big(\log m\big)^{(d-1)(\frac{1}{p_0}-\frac{1}{q_0})}\,, \qquad m \ge 2\,, \tag{40}$$

if $p_0 < p_1$, $p_0 \le q_0 < \infty$ and $p_1 \le q_1 < \infty$. As above $t := 1/p_0 - 1/p_1$.

5.2.3 Best m-term widths for embeddings of b-spaces

Because we prepared ourselves quite well in Section 4 it is now very simple to characterize the asymptotic behaviour of $\sigma_m\Big(s^{\frac{1}{p_0}-\frac{1}{p_1}}_{p_0,q_0} b(\nabla), s^0_{p_1,q_1} b(\nabla), \mathcal{B}\Big)$.

Theorem 6. *Let $p_0 < p_1$ and $q_0 \le q_1$. Let*

$$t := \frac{1}{p_0} - \frac{1}{p_1} \qquad \textit{and} \qquad r := \min\Big(\frac{1}{p_0} - \frac{1}{p_1}\,, \frac{1}{q_0} - \frac{1}{q_1}\Big)\,.$$

Then

$$\sigma_m\Big(s^t_{p_0,q_0} b(\nabla), s^0_{p_1,q_1} b(\nabla), \mathcal{B}\Big) \asymp m^{-r}\,, \qquad m \in \mathbb{N}\,.$$

Proof. The identity

$$\ell_q\big(I, \ell_p(J)\big) = s^{\frac{1}{p}-\frac{1}{2}}_{p,q} b(\nabla)\,,$$

with $I = \mathbb{N}_0^d$ and $J_i = \nabla_i$ is obvious. Proposition 4 yields

$$\sigma_m\Big(s^{\frac{1}{p_0}-\frac{1}{2}}_{p_0,q_0} b(\nabla), s^{\frac{1}{p_1}-\frac{1}{2}}_{p_1,q_1} b(\nabla), \mathcal{B}\Big) \asymp m^{-r}\,.$$

The proof is complete. ∎

Remark 17. (i) The restriction $q_0 \le q_1$ is necessary in this context, since we have

$$s^{\frac{1}{p_0}-\frac{1}{p_1}}_{p_0,q_0} b(\Omega) \hookrightarrow s^0_{p_1,q_1} b(\Omega) \qquad \Longleftrightarrow \qquad q_0 \le q_1\,.$$

This follows immediately from the necessity of $q_0 \le q_1$ for the embedding $S^{\frac{1}{p_0}-\frac{1}{p_1}}_{p_0,q_0} B(\mathbb{R}^d) \hookrightarrow S^0_{p_1,q_1} B(\mathbb{R}^d)$, see [24].
(ii) The m-term approximation for this case shows exactly the same decay behaviour as in the isotropic setting. Also the proof is remarkably similar, see [12].

5.2.4 Best m-term widths in case of non-limiting embeddings

It remains to deal with embeddings such that $t > 1/p_0 - 1/p_1$. The answer is simple and a bit surprising.

Theorem 7. *We suppose $p_0 \le p_1$ and*

$$t > \frac{1}{p_0} - \frac{1}{p_1}\,.$$

Then, for $x, y \in \{b, f\}$, we have

$$\sigma_m\Big(s^t_{p_0,q_0} x, s^0_{p_1,q_1} y, \mathcal{B}\Big) \asymp m^{-\frac{1}{p_0}+\frac{1}{p_1}}\,, \qquad m \in \mathbb{N}\,.$$

Proof. *Step 1.* First we consider the estimate from above.
Substep 1.1. We suppose $p_0 < p_1$. Let

$$\varepsilon := \frac{1}{2}\Big(t - d\Big(\frac{1}{p_0} - \frac{1}{p_1}\Big)\Big)\,.$$

Then, with $x, y \in \{b, f\}$, we find

$$s^t_{p_0,q_0} x \hookrightarrow s^{t-\varepsilon}_{p_0,p_0} b = \mathcal{A}^{\frac{1}{p_0}-\frac{1}{p_1}}_{p_0}\Big(s^{t-\varepsilon-d(\frac{1}{p_0}-\frac{1}{p_1})}_{p_1,p_1} b, \mathcal{B}\Big) \hookrightarrow \mathcal{A}^{\frac{1}{p_0}-\frac{1}{p_1}}_{\infty}\big(s^0_{p_1,q_1} y, \mathcal{B}\big)\,,$$

where we have used Corollary 2 and elementary monotonicity properties of the approximation spaces.
Substep 1.2. Let $p_0 = p_1$. Then $\sigma_m\Big(s^t_{p_0,q_0} x, s^0_{p_0,q_1} y, \mathcal{B}\Big) \lesssim 1$, $m \in \mathbb{N}$, follows from the continuous embedding $s^t_{p_0,q_0} x \hookrightarrow s^0_{p_0,q_1} y$.

Step 2. For the estimate from below, we use sequences a^m as in (29), where $i_m = 0$ for all $m \in \mathbb{N}$. One easily obtains

$$\big\|a^m\big|s^t_{p_0,q_0}x\big\| = (2m)^{1/p_0} \qquad \text{and} \qquad \sigma_m\big(a^m, s^0_{p_1,q_1}y, \mathcal{B}\big) = m^{1/p_1}$$

with the same arguments as in the proof of Prop. 4. ■

Remark 18. As in the isotropic case the behaviour of $\sigma_m\Big(s^t_{p_0,q_0}x, s^0_{p_1,q_1}y, \mathcal{B}\Big)$ does not depend on t. Only integrability properties count.

6 The widths of best m-term approximation for spaces of dominating mixed smoothness

Using Prop. 4 and Prop. 5 we can transfer Thm. 5 (and (40)) and Thm. 6 from sequence spaces to function spaces of dominating mixed smoothness. In addition we shall use the Littlewood-Paley assertions

$$S^0_{p,2}F(\mathbb{R}^d) = L_p(\mathbb{R}^d)\,, \qquad 1 < p < \infty\,,$$

and

$$S^r_{p,2}F(\mathbb{R}^d) = S^r_pH(\mathbb{R}^d)\,, \qquad 1 < p < \infty\,,$$

we refer to [25, Prop. 2.3.1, Thm. 2.3.1] and the references given there. Let Φ be as in Subsection 2.2. Then this procedure leads to the following.

Corollary 5. *Let $p_0 < p_1$, $q_1 \le p_1$ and $p_0 \le q_0 < \infty$. Let $t := \frac{1}{p_0} - \frac{1}{p_1}$. Then we have*

$$\sigma_m\Big(S^t_{p_0,q_0}F(\mathbb{R}^d), S^0_{p_1,q_1}F(\mathbb{R}^d), \Phi\Big) \asymp m^{-\frac{1}{p_0}+\frac{1}{p_1}}\,\big(\log m\big)^{(d-1)(\frac{1}{p_0}-\frac{1}{p_1}-\frac{1}{q_0}+\frac{1}{q_1})}\,,$$

if $m \ge 2$. Moreover, if $q_0 \le p_0 < p_1 \le q_1$ then we find for $m \in \mathbb{N}$

$$\sigma_m\Big(S^t_{p_0,q_0}F(\mathbb{R}^d), S^0_{p_1,q_1}F(\mathbb{R}^d), \Phi\Big) \sim m^{-\frac{1}{p_0}+\frac{1}{p_1}}\,.$$

Remark 19. (i) Observe, if $p_0 = q_0$ and $1 < p_1 < \infty$, $q_1 = 2$, we obtain

$$\begin{aligned}
\sigma_m\Big(S^t_{p_0,p_0}F(\mathbb{R}^d), S^0_{p_1,2}F(\mathbb{R}^d), \Phi\Big) &\asymp \sigma_m\Big(S^t_{p_0,p_0}B(\mathbb{R}^d), L_{p_1}(\mathbb{R}^d), \Phi\Big) \\
&\asymp m^{-\frac{1}{p_0}+\frac{1}{p_1}}\,\big(\log m\big)^{(d-1)(-\frac{1}{p_1}+\frac{1}{2})_+}\,,
\end{aligned}$$

if $m \ge 2$. This proves Thm. 3(i).

(ii) This time we choose $q_0 = q_1 = 2$. Then the second assertion in Cor. 5 can not be applied. Hence, we only obtain

$$\begin{aligned}
\sigma_m\Big(S^t_{p_0,2}F(\mathbb{R}^d), S^0_{p_1,2}F(\mathbb{R}^d), \Phi\Big) &\asymp \sigma_m\Big(S^t_{p_0,2}F(\mathbb{R}^d), L_{p_1}(\mathbb{R}^d), \Phi\Big) \\
&\asymp m^{-\frac{1}{p_0}+\frac{1}{p_1}}\,\big(\log m\big)^{(d-1)(\frac{1}{p_0}-\frac{1}{p_1})}\,,
\end{aligned}$$

if $p_0 < p_1$, $2 \le p_1$ and $p_0 \le 2$. This proves Thm. 3(ii).

The Cor. 5 and Remark 19 carry over to bounded domains Ω, see Subsection 3.2.3, as long as we define the spaces $S^r_{p,q}X(\Omega)$, $X \in \{F, B\}$, by restrictions. Now Φ has to be defined according to (18).

Corollary 6. *Let $p_0 < p_1$, $q_1 \le p_1$ and $p_0 \le q_0 < \infty$. Let $t := \frac{1}{p_0} - \frac{1}{p_1}$. Then we have*

$$\sigma_m\Big(S^t_{p_0,q_0}F(\Omega), S^0_{p_1,q_1}F(\Omega), \Phi\Big) \asymp m^{-\frac{1}{p_0}+\frac{1}{p_1}} \, \big(\log m\big)^{(d-1)(\frac{1}{p_0}-\frac{1}{p_1}-\frac{1}{q_0}+\frac{1}{q_1})} \, ,$$

if $m \ge 2$. Moreover, if $q_0 \le p_0 < p_1 \le q_1$ then we find for $m \in \mathbb{N}$

$$\sigma_m\Big(S^t_{p_0,q_0}F(\Omega), s^0_{p_1,q_1}F(\Omega), \Phi\Big) \sim m^{-\frac{1}{p_0}+\frac{1}{p_1}} \, .$$

Remark 20. The Cor. 6 has to be supplemented by Prop. 3, 6. This yields Thm. 1.

Next we translate Theorem 7. This time Φ has to be as in Subsection 2.2, see also (11).

Corollary 7. *We suppose $p_0 \le p_1$ and*

$$t > \frac{1}{p_0} - \frac{1}{p_1} \, .$$

Then, for $X, Y \in \{B, F\}$, we have

$$\sigma_m\Big(S^t_{p_0,q_0}Y(\mathbb{R}^d), S^0_{p_1,q_1}X(\mathbb{R}^d), \Phi\Big) \asymp m^{-\frac{1}{p_0}+\frac{1}{p_1}} \, , \qquad m \in \mathbb{N} \, .$$

Remark 21. By specializing $X = F$, $q_1 = 2$ and using the above Littlewood-Paley assertion we obtain Thm. 2.

Finally, we formulate the counterpart of Thm. 6. Let Φ be as in (18).

Corollary 8. *Let $p_0 < p_1$ and $q_0 \le q_1$. Let*

$$t := \frac{1}{p_0} - \frac{1}{p_1} \qquad \textit{and} \qquad r := \min\left(\frac{1}{p_0} - \frac{1}{p_1}, \frac{1}{q_0} - \frac{1}{q_1}\right) .$$

Then

$$\sigma_m\Big(S^t_{p_0,q_0}B(\Omega), S^0_{p_1,q_1}B(\Omega), \Phi\Big) \asymp m^{-r} \, , \qquad m \in \mathbb{N} \, .$$

References

[1] T.I. Amanov (1976): Spaces of differentiable functions with dominating mixed derivatives. Nauka Kaz. SSR, Alma-Ata.

[2] D.B. Bazarkhanov (2003): Characterizations of Nikol'skij-Besov and Lizorkin-Triebel function spaces of mixed smoothness. *Proc. Stekloc Inst.*, **243**, 46-58.

[3] D.B. Bazarkhanov (2005): Equivalent (quasi)normings of some function spaces of generalized mixed smoothness. *Proc. Stekloc Inst.*, **248**, 21-34.

[4] D.B. Bazarkhanov (2005): Wavelet representations and equivalent normings of some function spaces of generalized mixed smoothness. *Math. Zh.*, **5**, 12-16.

[5] D.B. Bazarkhanov (2008): Order sharp estimates of some approximation characteristics of certain function spaces with generalized mixed smoothness. Lecture, given at the int. conf, Freyburg.

[6] R.A. DeVore (1998): Nonlinear Approximation. *Acta numerica* **7**, 51-150.

[7] R.A. DeVore, B. Jawerth, and V. Popov (1992): Compression of wavelet decompositions. *Amer. J. Math.* **114**, 737–785.

[8] Dinh Dung (2001): Non-linear approximations using sets of finite cardinality or finite pseudo-dimension. *J. Compl.* **17**, 467-492.

[9] Dinh Dung (2001): Asymptotic orders of optimal non-linear approximations. *East J. Approx.* **7**, 55-76.

[10] R.E. Edwards (1965): Functional Analysis. Theory and Applications. New York, Holt, Rinehardt and Winston.

[11] M. Hansen (2010): Nonlinear approximation and function spaces of dominating mixed smoothness. Thesis, Friedrich-Schiller-Univ. Jena, Jena.

[12] M. Hansen, W. Sickel (2009): Best m-term approximation and Lizorkin-Triebel spaces. Preprint **22**, DFG-SPP 1324, Marburg.

[13] M. Hansen, W. Sickel (2010): Best m-term approximation and Sobolev-Besov spaces of dominating mixed smoothness – the case of compact embeddings. Preprint (in preparation), Jena.

[14] M. Hansen and J. Vybiral (2009): The Jawerth-Franke embedding for spaces with dominating mixed smoothness. Georgian Math. J. **16**, 667-682.

[15] C.-c. Hsiao, B. Jawerth, B.J. Lucier, and X.M. Yu (1994): Near optimal compression of orthonormal wavelet expansions. In: *Wavelets: mathematics and applications*, 425–446, Stud. Adv. Math., CRC, Boca Raton.

[16] B. Jawerth and M. Milman (1992): Wavelets and best approximation in Besov spaces. *Interpolation spaces and related topics (Haifa, 1990)*, Israel Math. Conf. Proc. **5**, 107–112.

[17] G. KYRIAZIS (2001): Non-linear approximation and interpolation spaces. *JAT* **113**, 110-126.

[18] W.A. LIGHT AND E.W. CHENEY (1985): *Approximation theory in tensor product spaces*, Lecture Notes in Math. **1169**, Springer, Berlin.

[19] P.-A. NITSCHE (2006): Best N-term approximation spaces for tensor product wavelet bases. *Constr. Approx.* **24**, 49-70.

[20] P. OSWALD (1999): On N-term approximation by Haar functions in H^s-norms. In: Metric function Theory and related topics in analysis (S.M. Nikol'skij, B.S. Kashin, A.D. Izaak, eds.), AFC, Moscow, pp. 137-163 (translated into russian).

[21] A. PIETSCH (1980): Approximation spaces. JAT **32**, 115-134.

[22] A.S. ROMANYUK (2003): Best M-term trigonometric approximations of Besov classes of periodic functions of several variables. *Izv. Math.*, **67** (2003), 265-302.

[23] H.-J. SCHMEISSER (2007): Recent developments in the theory of function spaces with dominating mixed smoothness. Proc. Conf. NAFSA-8, Prague 2006, (ed. J. Rakosnik), Inst. of Math. Acad. Sci., Czech Republic, pp. 145-204, Prague.

[24] H.-J. SCHMEISSER AND W. SICKEL (2004): Spaces of functions of mixed smoothness and their relations to approximation from hyperbolic crosses. JAT **128**, 115-150.

[25] H.-J. SCHMEISSER AND H. TRIEBEL (1987): Topics in Fourier analysis and function spaces. Wiley, Chichester.

[26] W. SICKEL UND T. ULLRICH (2009): Tensor products of Sobolev-Besov spaces and applications to approximation from the hyperbolic cross. *JAT* **161**, 748-786.

[27] W. SICKEL UND T. ULLRICH (2009): Spline interpolation on sparse grids. *Preprint*, Jena, Bonn.

[28] F. SPRENGEL (1999): A tool for approximation in bivariate periodic Sobolev spaces. In *Approximation Theory IX*, Vol. **2**, Vanderbilt Univ. Press, Nashville, 319-326.

[29] V.N. Temlyakov (1998): The best m-term approximations and greedy algorithms. *Adv. Comp. Math.* **8**, 249–265.

[30] V.N. Temlyakov (2000): *Greedy algorithms with regard to multivariate systems with special structure*. Constr. Approx. **16**, 399-425.

[31] H. Triebel (1983): *Theory of function spaces*. Birkhäuser, Basel.

[32] H. Triebel (2010): *Bases in function spaces, sampling, discrepancy, numerical integration*. EMS Publ. House, Zürich, 2010.

[33] T. Ullrich(2006): Function spaces with dominating mixed smoothness. Characterizations by differences. *Jenaer Schriften zur Mathematik und Informatik* Math/Inf/05/06, Jena.

[34] J. Vybiral (2006): Function spaces with dominating mixed smoothness. *Dissertationes Math.* **436**, 73 pp.

[35] P. Wojtaszczyk (2000): Greedy algorithm for general biorthoganal systems. *JAT* **37**, 293-314.

Preprint Series DFG-SPP 1324

http://www.dfg-spp1324.de

Reports

[1] R. Ramlau, G. Teschke, and M. Zhariy. A Compressive Landweber Iteration for Solving Ill-Posed Inverse Problems. Preprint 1, DFG-SPP 1324, September 2008.

[2] G. Plonka. The Easy Path Wavelet Transform: A New Adaptive Wavelet Transform for Sparse Representation of Two-dimensional Data. Preprint 2, DFG-SPP 1324, September 2008.

[3] E. Novak and H. Woźniakowski. Optimal Order of Convergence and (In-) Tractability of Multivariate Approximation of Smooth Functions. Preprint 3, DFG-SPP 1324, October 2008.

[4] M. Espig, L. Grasedyck, and W. Hackbusch. Black Box Low Tensor Rank Approximation Using Fibre-Crosses. Preprint 4, DFG-SPP 1324, October 2008.

[5] T. Bonesky, S. Dahlke, P. Maass, and T. Raasch. Adaptive Wavelet Methods and Sparsity Reconstruction for Inverse Heat Conduction Problems. Preprint 5, DFG-SPP 1324, January 2009.

[6] E. Novak and H. Woźniakowski. Approximation of Infinitely Differentiable Multivariate Functions Is Intractable. Preprint 6, DFG-SPP 1324, January 2009.

[7] J. Ma and G. Plonka. A Review of Curvelets and Recent Applications. Preprint 7, DFG-SPP 1324, February 2009.

[8] L. Denis, D. A. Lorenz, and D. Trede. Greedy Solution of Ill-Posed Problems: Error Bounds and Exact Inversion. Preprint 8, DFG-SPP 1324, April 2009.

[9] U. Friedrich. A Two Parameter Generalization of Lions' Nonoverlapping Domain Decomposition Method for Linear Elliptic PDEs. Preprint 9, DFG-SPP 1324, April 2009.

[10] K. Bredies and D. A. Lorenz. Minimization of Non-smooth, Non-convex Functionals by Iterative Thresholding. Preprint 10, DFG-SPP 1324, April 2009.

[11] K. Bredies and D. A. Lorenz. Regularization with Non-convex Separable Constraints. Preprint 11, DFG-SPP 1324, April 2009.

[12] M. Döhler, S. Kunis, and D. Potts. Nonequispaced Hyperbolic Cross Fast Fourier Transform. Preprint 12, DFG-SPP 1324, April 2009.

[13] C. Bender. Dual Pricing of Multi-Exercise Options under Volume Constraints. Preprint 13, DFG-SPP 1324, April 2009.

[14] T. Müller-Gronbach and K. Ritter. Variable Subspace Sampling and Multi-level Algorithms. Preprint 14, DFG-SPP 1324, May 2009.

[15] G. Plonka, S. Tenorth, and A. Iske. Optimally Sparse Image Representation by the Easy Path Wavelet Transform. Preprint 15, DFG-SPP 1324, May 2009.

[16] S. Dahlke, E. Novak, and W. Sickel. Optimal Approximation of Elliptic Problems by Linear and Nonlinear Mappings IV: Errors in L_2 and Other Norms. Preprint 16, DFG-SPP 1324, June 2009.

[17] B. Jin, T. Khan, P. Maass, and M. Pidcock. Function Spaces and Optimal Currents in Impedance Tomography. Preprint 17, DFG-SPP 1324, June 2009.

[18] G. Plonka and J. Ma. Curvelet-Wavelet Regularized Split Bregman Iteration for Compressed Sensing. Preprint 18, DFG-SPP 1324, June 2009.

[19] G. Teschke and C. Borries. Accelerated Projected Steepest Descent Method for Nonlinear Inverse Problems with Sparsity Constraints. Preprint 19, DFG-SPP 1324, July 2009.

[20] L. Grasedyck. Hierarchical Singular Value Decomposition of Tensors. Preprint 20, DFG-SPP 1324, July 2009.

[21] D. Rudolf. Error Bounds for Computing the Expectation by Markov Chain Monte Carlo. Preprint 21, DFG-SPP 1324, July 2009.

[22] M. Hansen and W. Sickel. Best m-term Approximation and Lizorkin-Triebel Spaces. Preprint 22, DFG-SPP 1324, August 2009.

[23] F.J. Hickernell, T. Müller-Gronbach, B. Niu, and K. Ritter. Multi-level Monte Carlo Algorithms for Infinite-dimensional Integration on $\mathbb{R}^{\mathbb{N}}$. Preprint 23, DFG-SPP 1324, August 2009.

[24] S. Dereich and F. Heidenreich. A Multilevel Monte Carlo Algorithm for Lévy Driven Stochastic Differential Equations. Preprint 24, DFG-SPP 1324, August 2009.

[25] S. Dahlke, M. Fornasier, and T. Raasch. Multilevel Preconditioning for Adaptive Sparse Optimization. Preprint 25, DFG-SPP 1324, August 2009.

[26] S. Dereich. Multilevel Monte Carlo Algorithms for Lévy-driven SDEs with Gaussian Correction. Preprint 26, DFG-SPP 1324, August 2009.

[27] G. Plonka, S. Tenorth, and D. Roşca. A New Hybrid Method for Image Approximation using the Easy Path Wavelet Transform. Preprint 27, DFG-SPP 1324, October 2009.

[28] O. Koch and C. Lubich. Dynamical Low-rank Approximation of Tensors. Preprint 28, DFG-SPP 1324, November 2009.

[29] E. Faou, V. Gradinaru, and C. Lubich. Computing Semi-classical Quantum Dynamics with Hagedorn Wavepackets. Preprint 29, DFG-SPP 1324, November 2009.

[30] D. Conte and C. Lubich. An Error Analysis of the Multi-configuration Time-dependent Hartree Method of Quantum Dynamics. Preprint 30, DFG-SPP 1324, November 2009.

[31] C. E. Powell and E. Ullmann. Preconditioning Stochastic Galerkin Saddle Point Problems. Preprint 31, DFG-SPP 1324, November 2009.

[32] O. G. Ernst and E. Ullmann. Stochastic Galerkin Matrices. Preprint 32, DFG-SPP 1324, November 2009.

[33] F. Lindner and R. L. Schilling. Weak Order for the Discretization of the Stochastic Heat Equation Driven by Impulsive Noise. Preprint 33, DFG-SPP 1324, November 2009.

[34] L. Kämmerer and S. Kunis. On the Stability of the Hyperbolic Cross Discrete Fourier Transform. Preprint 34, DFG-SPP 1324, December 2009.

[35] P. Cerejeiras, M. Ferreira, U. Kähler, and G. Teschke. Inversion of the noisy Radon transform on $SO(3)$ by Gabor frames and sparse recovery principles. Preprint 35, DFG-SPP 1324, January 2010.

[36] T. Jahnke and T. Udrescu. Solving Chemical Master Equations by Adaptive Wavelet Compression. Preprint 36, DFG-SPP 1324, January 2010.

[37] P. Kittipoom, G. Kutyniok, and W.-Q Lim. Irregular Shearlet Frames: Geometry and Approximation Properties. Preprint 37, DFG-SPP 1324, February 2010.

[38] G. Kutyniok and W.-Q Lim. Compactly Supported Shearlets are Optimally Sparse. Preprint 38, DFG-SPP 1324, February 2010.

[39] M. Hansen and W. Sickel. Best m-Term Approximation and Tensor Products of Sobolev and Besov Spaces – the Case of Non-compact Embeddings. Preprint 39, DFG-SPP 1324, March 2010.