COMPRESSION OF ANISOTROPIC TENSOR-PRODUCT DISCRETIZATIONS

S. KNAPEK
D-80333 MÜNCHEN
{STKNAPEK}@YAHOO.DE

Abstract. This paper is concerned with the discretization and compression of integral operators. It continues the work of [19]. Based on the framework of tensor-product biorthogonal wavelet bases we construct compression schemes for operator adapted sparse grid type discretizations. These discretization and compression schemes preserve the approximation order of the standard full-grid spaces. We give detailed information on the cases in which the curse of dimension can be broken. We show that our compression schemes are optimal up to logarithmic factors.

AMS subject classification. 41A17, 41A25, 41A30, 41A65, 45L05, 45L10, 65D99, 65N12, 65N30, 65N38, 65N55

Key words. biorthogonal wavelets, compression, optimized approximation spaces, sparse grids, hyperbolic cross approximation, integral equations

1. Introduction. Consider a variational problem: Given $f \in \mathcal{H}^s$, $s \in \mathbb{R}$, find $u \in \mathcal{H}^{-s}$ such that

$$a(u,v) = (f,v) \ \forall v \in \mathcal{H}^{-s}. \tag{1.1}$$

The Galerkin method for solving problem (1.1) numerically is to select a finite-dimensional subspace V_J from $\mathcal{H}^{-s} \cap \mathcal{L}^2$ and to solve the variational problem therein. Fixing a basis of V_J finally leads to a linear system of equations

$$A_J x_J = b_J \tag{1.2}$$

of dimension $dim(V_J)$. The work count to solve (1.2) to a prescribed accuracy can often be made proportional to the number of non-zero entries in A_J, see e.g. [5, 21]. For conventional spaces that belong to uniform full grids with subdivision rate two, this number is at least of the order $dim(V_J) = O(2^{nJ})$, i.e. it grows exponentially with the dimension n. This is particularly problematic in the case of integral operators with global kernel K,

$$Au(\mathbf{x}) := \int K(\mathbf{x},\mathbf{y})u(\mathbf{y})d\mathbf{y}, \tag{1.3}$$

where a naive Galerkin discretization on a uniform full grid with $O(2^{nJ})$ unknowns leads to a discrete operator with $O(2^{2nJ})$ entries. This makes matrix vector multiplications, as they are needed in iterative methods, prohibitively expensive for large n or large J.

Approaches to reduce the work count for approximating the solution of (1.1) are either based on the smoothness properties of the solution u (and eventually the kernel K), i.e. on decay properties in Fourier-space, or on the decay properties of the kernel K in physical space.

The original idea of the first approach applied to (1.1) was to exploit smoothness properties of the solution u. This leads to the replacement of the full grid approximation space by a so-called sparse grid space (that also appeared under the names hyperbolic cross approximation or boolean blending schemes) [1, 8, 10, 13, 14, 15, 24, 25, 26, 27]. It has been shown that under some assumptions on the approximation and smoothness properties of the underlying anisotropic tensor-product basis functions and provided

that certain additional regularity assumptions on the solution u of (1.1) are fulfilled, the resulting approximation spaces exhibit the same order of approximation as the full grid space, while having significantly less dimension. This approach has been applied successfully to differential [4, 15, 27] as well as integral operators [15, 16].
The full potential of the sparse grid approach for integral operators is revealed when it is applied to the kernel K, exploiting K's smoothness properties. This leads to a significant further reduction of the work count, see [12, 18, 19, 22], and hence, reduces the need for additional compression schemes on top of these discretizations, compare section 3.
Another well known strategy to reduce the cost of matrix vector multiplications is to exploit eventual decay properties of the kernel in physical space. Such decay properties do hold for pseudo-differential operators and hence for example for the single and double layer potential operator as well as for classical differential operators [20]. For isotropic biorthogonal wavelets it has been shown that, with a sufficient number of vanishing moments, most entries in the corresponding stiffness matrices can be replaced by zero without destroying the order of approximation [3, 6, 7, 9, 23]. Then for isotropic tensor-product wavelets on a full grid the number of entries in the stiffness matrix after compression is only $O(J^k 2^{nJ})$ with some $k \in \mathbb{R}^+$ or even $O(2^{nJ})$ [23]. Corresponding investigations concerning anisotropic tensor-product discretizations seem still to be missing, especially together with sparse grid type discretizations.
In this paper we continue the work of [19] and develop compression schemes for anisotropic tensor-product type discretisations of (1.3). We show that in many cases a suitable discretization scheme renders additional compression unnecessary. We identify the cases, where additional compression is worthwhile. For the relevant cases we we develop a suitable compression scheme along the lines of [23] and show how optimal compression (up to logarithmic factors) can be obtained.
The remainder of the paper is as follows. Section 2 introduces the necessary notation and summarizes the basic facts about biorthogonal wavelet bases, tensor-product spaces, norm equivalences and the smoothness classes we consider (they are certain intersections of classes of functions with dominating mixed derivatives, see (2.1) below). In Section 3 we discuss anisotropic tensor-product discretizations for integral operators including full and sparse grid type discretizations. In Section 4 we introduce and discuss our compression scheme.

2. Preliminaries. Multi-indices (vectors) are written boldface, for example $\mathbf{j} = (j_1, \dots, j_n)$. Inequalities like $\mathbf{l} \leq \mathbf{t}$ or $\mathbf{l} \leq \mathbf{0}$ are to be understood componentwise. We write $x \simeq y$ if there exist C_1, C_2 independent of any parameters x or y may depend on such that $C_1 \cdot y \leq x \leq C_2 \cdot y$. In the rest of the paper C denotes a generic constant which may depend on the smoothness assumptions and on the dimension n of the problem. Moreover, for $\mathbf{j}, \mathbf{k} \in \mathbb{N}^n$ we write $|\mathbf{j}|_\infty := \max_{1 \leq i \leq n}(|j_i|)$, $|\mathbf{j}|_1 := \sum_{i=1}^n |j_i|$, $|(\mathbf{j}, \mathbf{k})|_\infty := \max(|\mathbf{j}|_\infty, |\mathbf{k}|_\infty)$ and $|(\mathbf{j}, \mathbf{k})|_1 := |\mathbf{j}|_1 + |\mathbf{k}|_1$. With $dist(\mathbf{x}, \mathbf{y})$ we denote the Euclidian distance between $\mathbf{x}$ and $\mathbf{y}$.

2.1. Sobolev spaces. Let us denote by $\mathcal{H}^t(I^n), t \in \mathbb{R}$, a scale of standard Sobolev spaces on $I^n := [0,1]^n$ and by $\mathcal{L}^2(I^n)$ the space of $\mathcal{L}^2$-integrable functions on I^n. When $t < 0, \mathcal{H}^t(I^n)$ is defined as the dual of $\mathcal{H}^{-t}(I^n)$, i.e., $\mathcal{H}^t := (\mathcal{H}^{-t})'$ (incorporating eventual boundary conditions).
Now we fix the smoothness assumptions we consider. Let $t \in \mathbb{R}, l \in \mathbb{R}_0^+$, $\mathbf{1} := (1, \dots, 1)$ and $\mathbf{e_i} := (0, \dots, 0, 1, 0, \dots, 0)$ the i-th unit-vector in $\mathbb{R}^n$. Then we define

$$\mathcal{H}^{t,l}_{mix}(I^n) := \mathcal{H}^{t\mathbf{1}+l\mathbf{e_1}}_{mix}(I^n) \cap \dots \cap \mathcal{H}^{t\mathbf{1}+l\mathbf{e_n}}_{mix}(I^n), \tag{2.1}$$

where $\mathcal{H}^{\mathbf{k}}_{mix}(I^n) := \mathcal{H}^{k_1}(I) \otimes \ldots \otimes \mathcal{H}^{k_n}(I)$ for $\mathbf{k} \in \mathbb{R}^n$.

This definition includes the class of functions with dominating mixed derivative $\mathcal{H}^{t,0}_{mix}(I^n)$ as well as the standard isotropic Sobolev spaces $\mathcal{H}^t(I^n) = \mathcal{H}^{0,t}_{mix}(I^n)$. Note that $\mathcal{H}^{t,l}_{mix}$ is a mixture of classes of functions with dominating mixed derivative $\mathcal{H}^t_{mix}(T^n) = \mathcal{H}^{t,0}_{mix}(T^n)$ and standard Sobolev spaces $\mathcal{H}^t(T^n) = \mathcal{H}^{0,t}_{mix}(T^n)$. For additional information on these spaces, specifically on the definition via Fourier transform, the treatment of boundary conditions and the connections between spaces of bounded mixed derivative and these anisotropic tensor-product spaces see [15, 17, 18]. In the rest of the paper we will drop I^n in $\mathcal{H}^{t,l}_{mix}(I^n)$ and write $\mathcal{H}^{t,l}_{mix}$.

2.2. Biorthogonal wavelet bases. The approximation spaces considered here are anisotropic tensor-products of univariate function spaces. We start from a one-dimensional multi-resolution analysis $\mathcal{L}^2 = \bigcup_{j\geq 0} V_j$ and we assume that the complement spaces $W_j = V_j \ominus V_{j-1}$ of V_{j-1} in V_j are spanned by some multiscale basis functions such that we have $W_j = \operatorname{span}\{\psi_{jk}, k \in \tau_j\}$, where τ_j is some index set defined from the subdivision rate of successive refinement levels. We will consider dyadic refinement throughout the paper. Moreover, we assume that $\{\psi_{jk}, k \in \tau_j\}$ forms a Riesz-basis of W_j and that there exists a dual system $\{\tilde{\psi}_{jk}, k \in \tau_j, j \in \mathbb{N}_0\}$, such that $\langle \psi_{jk}, \tilde{\psi}_{j'k'}\rangle_{\mathcal{L}^2} = \delta_{jj'}\delta_{kk'},\ j, j' \in \mathbb{N}_0, k \in \tau_j, k' \in \tau_{j'}$ holds. We assume $\|\psi_{jk}\|_{\mathcal{L}^2} = 1$ and $diam(supp(\psi_{jk})) \lesssim 2^{-j}$. In the following let ψ_{jk} and $\tilde{\psi}_{jk}$ have N and $\tilde{N}$ vanishing moments, respectively. Moreover we write $\tilde{W}_j = \operatorname{span}\{\tilde{\psi}_{jk}, k \in \tau_j\}$ for the complement spaces spanned by the dual wavelets.

For the higher-dimensional case $n > 1$, let $\mathbf{j} \in \mathbb{N}_0^n$ be given. Consider the tensor-product partition with uniform step size 2^{-j_i} into the i-th coordinate direction. By $W_{\mathbf{j}}$ and $\tilde{W}_{\mathbf{j}}$ we denote the corresponding tensor-product function spaces $W_{\mathbf{j}} := W_{j_1} \otimes \ldots \otimes W_{j_n}$ and $\tilde{W}_{\mathbf{j}} := \tilde{W}_{j_1} \otimes \ldots \otimes \tilde{W}_{j_n}$. A basis of $W_{\mathbf{j}}$ is then given by $\cup_{\mathbf{k}\in\tau_{\mathbf{j}}}\{\psi_{\mathbf{jk}}(\mathbf{x}) = \psi_{j_1k_1}(x_1) \cdot \ldots \cdot \psi_{j_nk_n}(x_n)\}$, with $\tau_{\mathbf{j}} = \tau_{j_1} \times \ldots \times \tau_{j_n}$.
To simplify things we assume in the following the validity of the norm equivalences

$$\|u\|^2_{\mathcal{H}^{t,l}_{mix}} \simeq \sum_{\mathbf{j}} 2^{2t|\mathbf{j}|_1+2l|\mathbf{j}|_\infty}\|w_{\mathbf{j}}\|^2_{\mathcal{L}^2},\ \text{ for } u = \sum_{\mathbf{j}} w_{\mathbf{j}} \in \mathcal{H}^{t,l}_{mix}, w_{\mathbf{j}} \in W_{\mathbf{j}}, \tag{2.2}$$

$t, t+l \in (-\tilde{r}, r)$ and

$$\|u\|^2_{\mathcal{H}^{t,l}_{mix}} \simeq \sum_{\mathbf{j}} 2^{2t|\mathbf{j}|_1+2l|\mathbf{j}|_\infty}\|\tilde{w}_{\mathbf{j}}\|^2_{\mathcal{L}^2},\ \text{ for } u = \sum_{\mathbf{j}} \tilde{w}_{\mathbf{j}} \in \mathcal{H}^{t,l}_{mix}, \tilde{w}_{\mathbf{j}} \in \tilde{W}_{\mathbf{j}}, \tag{2.3}$$

$t, t+l \in (-r, \tilde{r})$, see [15]. Such two-sided estimates can be inferred from the validity of norm equivalences in one dimension which in turn can be inferred from direct estimates (estimates of Jackson type) and inverse estimates (Bernstein inequalities) for the primal and the dual wavelets as a consequence of approximation theory in Sobolev spaces together with interpolation and duality arguments, see [5, 15, 21].
Note that for $t = 0$ or $l = 0$ we regain the (standard) norm equivalences for the isotropic Sobolev space $\mathcal{H}^l$ and the Sobolev space with dominating mixed derivative $\mathcal{H}^{t,0}_{mix}$. The different factors $2^{2t|\mathbf{j}|_1}$ and $2^{2l|\mathbf{j}|_\infty}$ in these equivalences reflect the different smoothness requirements.

3. Anisotropic tensor-product discretization. Given an index set $\mathcal{I} \subset \mathbb{N}_0^{2n}$ we consider the approximation space

$$V_{\mathcal{I}} := \bigoplus_{(\mathbf{j},\mathbf{k})\in\mathcal{I}} \tilde{W}_{\mathbf{j}} \otimes \tilde{W}_{\mathbf{k}}$$

and the corresponding Galerkin stiffness matrix

$$(\mathcal{A}_{\mathcal{I}})_{\mathbf{jklm}} := \begin{cases} a_{\mathbf{jklm}} & \text{for } (\mathbf{j},\mathbf{k}) \in \mathcal{I} \text{ and } (\mathbf{l},\mathbf{m}) \in \tau_{\mathbf{jk}}, \\ 0 & \text{else,} \end{cases}$$

with $a_{\mathbf{jklm}} := \langle A\psi_{\mathbf{jl}}, \psi_{\mathbf{km}}\rangle_{\mathcal{L}^2}$ and $\tau_{\mathbf{jk}} := \tau_{\mathbf{j}} \times \tau_{\mathbf{k}}$. Here we assume that eventual boundary conditions are implemented into the definition of the primal and dual wavelets. The corresponding integral operator is given by

$$A_{\mathcal{I}}u(\mathbf{x}) := \int K_{\mathcal{I}}(\mathbf{x},\mathbf{y})u(\mathbf{y})d\mathbf{y}$$

with

$$K_{\mathcal{I}}(\mathbf{x},\mathbf{y}) := \sum_{(\mathbf{j},\mathbf{k})\in\mathcal{I}} \sum_{(\mathbf{l},\mathbf{m})\in\tau_{\mathbf{jk}}} a_{\mathbf{jklm}}\tilde{\psi}_{\mathbf{jl}}(\mathbf{x})\tilde{\psi}_{\mathbf{km}}(\mathbf{y}).$$

Typical approximation spaces are defined via the index sets

$$\begin{aligned} \mathcal{I}_J &= \{(\mathbf{j},\mathbf{k}) \in \mathbb{N}_0^{2n} : |\mathbf{j},\mathbf{k}|_\infty \le J\}, \\ \mathcal{I}_J &= \{(\mathbf{j},\mathbf{k}) \in \mathbb{N}_0^{2n} : |\mathbf{j}|_1 \le J \text{ and } |\mathbf{k}|_1 \le J\} \text{ or} \\ \mathcal{I}_J &= \{(\mathbf{j},\mathbf{k}) \in \mathbb{N}_0^{2n} : |\mathbf{j},\mathbf{k}|_1 \le J\}, \end{aligned} \tag{3.1}$$

$J \in \mathbb{N}$, leading to the full grid space with dimension $O(2^{2nJ})$, the tensor product of two sparse grid spaces in n dimensions with dimension $O(2^{2J}J^{2n-2})$ and the sparse grid space in $2n$ dimensions with dimension $O(2^J J^{2n-1})$, respectively. Note that the dimension of the sparse grid spaces compares favorably with the dimension of the full grid space.

In the rest of the paper we assume $l, t, p, q \ge 0$ and $l + t \ge s \ge -(p+q)$.

Then the approximation error when using the full grid approximation space induced by the index set (3.1) is

$$\|(A - A_{\mathcal{I}_J})u\|_{\mathcal{H}^s} \le C \cdot 2^{(s-l-t)J} \cdot \|K\|_{H^{t,l}_{mix}} \|u\|_{H^{p,q}_{mix}} \text{ for } u \in H^{p,q}_{mix}, K \in H^{t,l}_{mix}, \tag{3.2}$$

compare [19]. From the results in [19] we also obtain the following Theorem.

Theorem 1. *We assume that the norm equivalences (2.2) and (2.3) hold. Let the parameters $\tilde{r}, r$ from (2.2) and (2.3) fulfill $-\tilde{r} < s < r$ and $p \in [0, r), p + q \in [0, r)$, $t \in [0, \tilde{r}), t + l \in [0, \tilde{r})$ Let the index set $\mathcal{I}_J(s, t, l, p, q)$ be defined by*

$$\begin{aligned} \mathcal{I}_{J(s,t,l,p,q)} := \{&(\mathbf{j},\mathbf{k}) \in \mathbb{N}_0^{2n} : \\ &-s|\mathbf{j}|_\infty + l|\mathbf{j},\mathbf{k}|_\infty + t|\mathbf{j},\mathbf{k}|_1 + q|\mathbf{k}|_\infty + p|\mathbf{k}|_1 \le (l + t - s)J\}. \end{aligned} \tag{3.3}$$

We denote the corresponding integral operator by $A_{\mathcal{I}_J(s,t,l,p,q)}$. Moreover, let $(A - A_{\mathcal{I}_J(s,t,l,p,q)})u \in \mathcal{H}^s$, $u \in \mathcal{H}^{p,q}_{mix}$, $K \in \mathcal{H}^{t,l}_{mix}$. Then

$$\|(A - A_{\mathcal{I}_J(s,t,l,p,q)})u\|_{\mathcal{H}^s} \le C \cdot 2^{(s-l-t)J} \cdot \|K\|_{H^{t,l}_{mix}} \|u\|_{H^{p,q}_{mix}}. \tag{3.4}$$

That is, the order of approximation of the full grid approximation space is preserved when using the approximation space $V_J(s, t, l, p, q)$ induced by the index set (3.3).

The number of elements in the approximation $V_{\mathcal{I}_J(s,t,l,p,q)}$ space is given by

$$
\dim(V_{\mathcal{I}_J(s,t,l,p,q)}) = \begin{cases} O(2^J) & \text{for } l-s<0, \\ O(2^{nJ\frac{s-l-t}{s-l-nt}}) & \text{for } 0<l-s<q+np, \\ O(2^{2nJ\frac{s-l-t}{s-l-2nt-q-np}}) & \text{for } 0\le q+np<l-s. \end{cases} \tag{3.5}
$$

In the extremal cases $l-s=0$ and $0<l-s=q+np$ there appear additional logarithmic terms. These estimates follow directly from general estimates in [11].
Combining (3.4) with (3.5) we see that the number of non-zero elements $\mathcal{N}$ in the stiffness matrix $\mathcal{A}_{\mathcal{I}_J(s,t,l,p,q)}$ needed to obtain $\|(A-A_{\mathcal{I}_J(s,t,l,p,q)})u\|_{\mathcal{H}^s}\le\epsilon, \epsilon\in\mathbb{R}^+$, is bounded by

$$
\mathcal{N} \le \begin{cases} O(\epsilon^{\frac{1}{s-l-t}}) & \text{for } l-s<0, \\ O(\epsilon^{\frac{1}{\frac{s-l}{n}-t}}) & \text{for } 0<l-s<q+np, \\ O(\epsilon^{\frac{2}{\frac{s-l-q}{n}-p-2t}}) & \text{for } 0\le q+np<l-s. \end{cases} \tag{3.6}
$$

Again, in the extremal cases $l-s=0$ and $0<l-s=q+np$ there appear additional logarithmic terms. These estimates need to be compared with

$$
\mathcal{N} = O(\epsilon^{\frac{2n}{s-l-t}}) \tag{3.7}
$$

for the full grid approximation space.
Hence, for a kernel with dominating mixed smoothness, i.e. $t>0$, we have the following results: In the case $l-s<0$ the number of non-zero elements in the resulting stiffness matrix does not depend on the dimension n. In the case $0\le l-s<q+np$ there appears a slight n-dependence, which is further increasing in the case $0\le q+np<l-s$.
In the case $t=0$, i.e. for a kernel with isotropic smoothness, the above approximation scheme in most cases still significantly reduces the number of non-zero entries in the stiffness matrix compared to the full grid scheme. Consider for example the case $t=p=0$, $0<l-s<q$, i.e., the case of fully isotropic smoothness. Then the dependence on the dimension n of the number of non-zero elements is reduced from $2n$ in the full grid case to n. Therefore, in many cases the discretization scheme induced by the index set (3.3) renders additional compression schemes unnecessary. Only in the case $0\le q+np<l-s$ there can be a relatively strong dependence on the dimension n reflected by the exponent $2n$ in the third estimates of (3.5) and (3.6). Here additional compression may be worthwhile. The rest of the paper is concerned with the question whether the dependence on n in this case can be reduced significantly via additional compression.

4. Compression. An additional reduction of the cost of evaluating (1.3) approximately may be obtained by exploiting the decay properties of the kernel K in physical space. The idea is again to drop small entries in the stiffness matrix without destroying the order of approximation.
Here one usually assumes specific decay properties of the Schwarz-kernel of the pseudodifferential operator under consideration. For singular kernels these are typically of the form

$$
\exists\, C=C_{\boldsymbol{\alpha\beta}}<\infty : \left|\partial_x^{\boldsymbol{\alpha}}\partial_y^{\boldsymbol{\beta}}K(\mathbf{x},\mathbf{y})\right| \le C\cdot\left|\mathbf{x}-\mathbf{y}\right|^{-f(\boldsymbol{\alpha},\boldsymbol{\beta})},\ f(\boldsymbol{\alpha},\boldsymbol{\beta})>0,\ \mathbf{x}\ne\mathbf{y}. \tag{4.1}
$$

Here f is typically of the form $f(\boldsymbol{\alpha},\boldsymbol{\beta}) = |\boldsymbol{\alpha}+\boldsymbol{\beta}+m\mathbf{1}|_{\mathbf{1}}$ with some $m \in \mathbb{R}^+$. Then estimates of the size of the entries $a_{\mathbf{jklm}}$ of the Galerkin stiffness matrix are obtained by expansions of the Schwarz-kernel in a polynomial basis together with the cancellation properties of the primal wavelets. A Taylor expansion of the kernel together with the cancellation properties of the primal wavelets together with (4.1) shows

$$\langle A\psi_{\mathbf{jl}}, \psi_{\mathbf{km}}\rangle_{\mathcal{L}^2} \leq C \cdot \frac{2^{-|\mathbf{l},\mathbf{k}|_1 (N+\frac{3}{2})}}{dist(supp(\psi_{\mathbf{jl}}), supp(\psi_{\mathbf{km}}))^{f(\mathbf{N+1},\mathbf{N+1})}}, \tag{4.2}$$

for $dist(supp(\psi_{\mathbf{jl}}), supp(\psi_{\mathbf{km}})) > 0$.
Given an additional index set $\tau_{\mathbf{jk}}^{comp} \subset \tau_{\mathbf{jk}}$ we can replace the stiffness matrix $\mathcal{A}_{\mathcal{I}}$ by $\mathcal{A}_{\mathcal{I}}^{comp}$ with

$$(\mathcal{A}_{\mathcal{I}}^{comp})_{\mathbf{jklm}} = \begin{cases} (\mathcal{A}_{\mathcal{I}})_{\mathbf{jklm}} & \text{for } (\mathbf{j},\mathbf{k}) \in \mathcal{I} \text{ and } (\mathbf{l},\mathbf{m}) \in \tau_{\mathbf{jk}}^{comp}, \\ 0 & \text{for } (\mathbf{j},\mathbf{k}) \in \mathcal{I} \text{ and } (\mathbf{l},\mathbf{m}) \notin \tau_{\mathbf{jk}}^{comp}. \end{cases}$$

The corresponding integral operator is given by

$$A_{\mathcal{I}}^{comp} u(\mathbf{x}) := \int K_{\mathcal{I}}^{comp}(\mathbf{x},\mathbf{y}) u(\mathbf{y}) d\mathbf{y}$$

with

$$K_{\mathcal{I}}^{comp}(\mathbf{x},\mathbf{y}) := \sum_{(\mathbf{j},\mathbf{k})\in\mathcal{I}} \sum_{(\mathbf{l},\mathbf{m})\in\tau_{\mathbf{jk}}^{comp}} a_{\mathbf{jklm}} \tilde{\psi}_{\mathbf{jl}}(\mathbf{x}) \tilde{\psi}_{\mathbf{km}}(\mathbf{y}).$$

The remaining task is to define the $\tau_{\mathbf{jk}}^{comp}$ in such a way that the order of approximation does not deteriorate because of this compression.
The following result tells us how to define the index sets $\tau_{\mathbf{jk}}^{comp}$ in order to balance the overall complexity and the error after compression.

THEOREM 2. *Let the assumptions of Theorem 1 hold. Moreover, let the decay property (4.2) hold. Let* $(A_{\mathcal{I}_J(s,t,l,p,q)} - A_{\mathcal{I}_J(s,t,l,p,q)}^{comp})u \in \mathcal{H}^s$. *Here* $\mathcal{I}_J(s,t,l,p,q)$ *is again given by (3.3). For* $(\mathbf{j},\mathbf{k}) \in \mathcal{I}_J(s,t,l,p,q)$ *let*

$$B_{\mathbf{jk}} := J^{\frac{n}{f(N+1,N+1)}} \cdot 2^{\frac{(t+l-s)J + s|\mathbf{j}|_\infty - q|\mathbf{k}|_\infty - p|\mathbf{k}|_1 - (N+1)|\mathbf{j},\mathbf{k}|_1}{f(N+1,N+1)}}.$$

We define

$$\tau_{\mathbf{jk}}^{comp} := \{(\mathbf{l},\mathbf{m}) \in \tau_{\mathbf{jk}} : dist(supp(\psi_{\mathbf{jl}}), supp(\psi_{\mathbf{km}})) \leq B_{\mathbf{jk}}\}. \tag{4.3}$$

Then

$$\|(A - A_{\mathcal{I}_J(s,t,l,p,q)}^{comp})u\|_{\mathcal{H}^s} \leq C \cdot 2^{(s-l-t)J} \cdot \|u\|_{\mathcal{H}_{mix}^{p,q}}.$$

Proof: For shortness we use the abbreviations $\mathcal{I}$ for $\mathcal{I}_J(s,t,l,p,q)$, $A_{\mathcal{I}}$ for $A_{\mathcal{I}_J(s,t,l,p,q)}$ and analogously $A_{\mathcal{I}}^{comp}$ for $A_{\mathcal{I}_J(s,t,l,p,q)}^{comp}$. Let $u_{\mathbf{jk}} := \langle u, \tilde{\psi}_{jk}\rangle_{\mathcal{L}^2}$ be the unique wavelet coefficients of $u \in \mathcal{L}^2$, i.e., $u = \sum_{\mathbf{j}=0}^{\infty} \sum_{\mathbf{k}\in\tau_{\mathbf{j}}} u_{\mathbf{jk}}\psi_{jk}$. We use the biorthogonality between the dual and the primal wavelets and the stability of the dual wavelets in $\mathcal{H}^s$ and the primal wavelets in $\mathcal{H}^{p,q}_{mix}$, i.e. the validity of (2.2) and (2.3). It holds

$$\|(A_{\mathcal{I}} - A_{\mathcal{I}}^{comp})u\|^2_{\mathcal{H}^s} = \| \sum_{(\mathbf{j},\mathbf{k})\in\mathcal{I}} \sum_{(\mathbf{l},\mathbf{m})\in\tau_{\mathbf{jk}}\setminus\tau_{\mathbf{jk}}^{comp}} a_{\mathbf{jklm}} u_{\mathbf{km}} \tilde{\psi}_{\mathbf{jl}}(\mathbf{x})\|^2_{\mathcal{H}^s}$$

$$= \| \sum_{\mathbf{j}} \sum_{\mathbf{l}\in\tau_{\mathbf{j}}} \left(\sum_{\mathbf{k}:(\mathbf{j},\mathbf{k})\in\mathcal{I}} \sum_{\mathbf{m}:(\mathbf{l},\mathbf{m})\in\tau_{\mathbf{jk}}\setminus\tau_{\mathbf{jk}}^{comp}} a_{\mathbf{jklm}} u_{\mathbf{km}} \right) \tilde{\psi}_{\mathbf{jl}}(\mathbf{x})\|^2_{\mathcal{H}^s}$$

$$\overset{(2.3)}{\simeq} \sum_{\mathbf{j}} \sum_{\mathbf{l}\in\tau_{\mathbf{j}}} 2^{2s|\mathbf{j}|_\infty} \left(\sum_{\mathbf{k}:(\mathbf{j},\mathbf{k})\in\mathcal{I}} \sum_{\mathbf{k}:(\mathbf{l},\mathbf{m})\in\tau_{\mathbf{jk}}\setminus\tau_{\mathbf{jk}}^{comp}} a_{\mathbf{jklm}} u_{\mathbf{km}} \right)^2$$

$$= \sum_{\mathbf{j}} \sum_{\mathbf{l}\in\tau_{\mathbf{j}}} 2^{2s|\mathbf{j}|_\infty} \left(\sum_{\mathbf{k}:(\mathbf{j},\mathbf{k})\in\mathcal{I}} \sum_{\mathbf{m}:(\mathbf{l},\mathbf{m})\in\tau_{\mathbf{jk}}\setminus\tau_{\mathbf{jk}}^{comp}} 2^{-q|\mathbf{k}|_\infty - p|\mathbf{k}|_1} a_{\mathbf{jklm}} 2^{q|\mathbf{k}|_\infty + p|\mathbf{k}|_1} u_{\mathbf{km}} \right)^2$$

$$\leq \sum_{\mathbf{j}} \sum_{\mathbf{l}\in\tau_{\mathbf{j}}} 2^{2s|\mathbf{j}|_\infty} \left(\sum_{\mathbf{k}:(\mathbf{j},\mathbf{k})\in\mathcal{I}} \sum_{\mathbf{m}:(\mathbf{l},\mathbf{m})\in\tau_{\mathbf{jk}}\setminus\tau_{\mathbf{jk}}^{comp}} 2^{-2q|\mathbf{k}|_\infty - 2p|\mathbf{k}|_1} a^2_{\mathbf{jklm}} \right)$$

$$\cdot \left(\sum_{\mathbf{k}\in\mathbb{N}^n} \sum_{\mathbf{m}\in\tau_{\mathbf{k}}} 2^{2q|\mathbf{k}|_\infty + 2p|\mathbf{k}|_1} u^2_{\mathbf{km}} \right)$$

$$\overset{(2.2)}{\leq} C \cdot \left(\sum_{(\mathbf{j},\mathbf{k})\in\mathcal{I}} \sum_{(\mathbf{l},\mathbf{m})\in\tau_{\mathbf{jk}}\setminus\tau_{\mathbf{jk}}^{comp}} 2^{2s|\mathbf{j}|_\infty - 2q|\mathbf{k}|_\infty - 2p|\mathbf{k}|_1} a^2_{\mathbf{jklm}} \right) \cdot \|u\|^2_{\mathcal{H}^{p,q}_{mix}}.$$

With the definition of $\tau_{\mathbf{jk}}^{comp}$ together with (4.2) we then obtain (for shortness we use the abbreviation $f = f(\mathbf{N}+\mathbf{1}, \mathbf{N}+\mathbf{1})$)

$$\|(A_{\mathcal{I}} - A_{\mathcal{I}}^{comp})u\|^2_{\mathcal{H}^s} \leq C \cdot \Bigg(\sum_{(\mathbf{j},\mathbf{k})\in\mathcal{I}} 2^{2s|\mathbf{j}|_\infty - 2q|\mathbf{k}|_\infty - 2p|\mathbf{k}|_1} 2^{-2|\mathbf{j},\mathbf{k}|_1 (N+\frac{3}{2})}$$

$$\cdot \sum_{\substack{(\mathbf{l},\mathbf{m})\in\tau_{\mathbf{jk}},\\ dist(supp(\psi_{\mathbf{jl}}),supp(\psi_{\mathbf{km}}))>B_{\mathbf{jk}}}} dist(supp(\psi_{\mathbf{jl}}), supp(\psi_{\mathbf{km}}))^{-2f} \Bigg) \cdot \|u\|^2_{\mathcal{H}^{p,q}_{mix}}$$

$$\leq C \cdot \Bigg(\sum_{(\mathbf{j},\mathbf{k})\in\mathcal{I}} 2^{2s|\mathbf{j}|_\infty - 2q|\mathbf{k}|_\infty - 2p|\mathbf{k}|_1} 2^{-2|\mathbf{j},\mathbf{k}|_1 (N+\frac{3}{2})}$$

$$\cdot B_{\mathbf{jk}}^{-2f} \sum_{\substack{(\mathbf{l},\mathbf{m})\in\tau_{\mathbf{jk}},\\ dist(supp(\psi_{\mathbf{jl}}),supp(\psi_{\mathbf{km}}))>B_{\mathbf{jk}}}} 1 \Bigg) \cdot \|u\|^2_{\mathcal{H}^{p,q}_{mix}}.$$

Now we apply $\sum_{(\mathbf{l},\mathbf{m})\in\tau_{\mathbf{jk}}} 1 = |\tau_{\mathbf{jk}}| \lesssim 2^{|\mathbf{j},\mathbf{k}|_1}$ and plug in the definition of $B_{\mathbf{jk}}$. We obtain

$$\|(A_{\mathcal{I}} - A_{\mathcal{I}}^{comp})u\|^2_{\mathcal{H}^s} \leq C \cdot \left(\sum_{(\mathbf{j},\mathbf{k})\in\mathcal{I}} 2^{2s|\mathbf{j}|_\infty - 2q|\mathbf{k}|_\infty - 2p|\mathbf{k}|_1} 2^{-2|\mathbf{j},\mathbf{k}|_1 (N+1)} B_{\mathbf{jk}}^{-2f} \right) \cdot \|u\|^2_{\mathcal{H}^{p,q}_{mix}}$$

$$\leq C \cdot 2^{2(s-l-t)J} J^{-2n} \left(\sum_{(\mathbf{j},\mathbf{k})\in\mathcal{I}} 1 \right) \cdot \|u\|^2_{\mathcal{H}^{p,q}_{mix}}$$

$$(4.4) \qquad \le C \cdot 2^{2(s-l-t)J} \cdot \|u\|^2_{\mathcal{H}^{p,q}_{mix}}.$$

Now we combine (3.4) from Theorem 1 with (4.4) and obtain

$$\begin{aligned}\|(A - A^{comp}_{\mathcal{I}})u\|_{\mathcal{H}^s} &\le \|(A - A_{\mathcal{I}})u\|_{\mathcal{H}^s} + \|(A_{\mathcal{I}} - A^{comp}_{\mathcal{I}})u\|_{\mathcal{H}^s} \\ &\le C \cdot 2^{(s-l-t)J} \cdot \|K\|_{H^{t,l}_{mix}} \|u\|_{H^{p,q}_{mix}} + C \cdot 2^{(s-l-t)J} \cdot \|u\|_{\mathcal{H}^{p,q}_{mix}} \\ &\le C \cdot 2^{(s-l-t)J} \cdot \|u\|_{\mathcal{H}^{p,q}_{mix}}.\end{aligned}$$

□

Theorem 2 shows that the order of approximation of the full grid approximation space is preserved. Note that the resulting stiffness matrix $\mathcal{A}_{\mathcal{I}_J(s,t,l,p,q)}$ will in general be non-symmetric. A symmetric operator may be derived by taking the maximum of $B_{\mathbf{jk}}$ and $B_{\mathbf{kj}}$ in the definition of $\tau^{comp}_{\mathbf{jk}}$.

The following Lemma gives an upper estimate for the number of remaining non-zero elements in the compressed stiffness matrix.

Lemma 1. *Let $0 \le q + np < l - s$. Then the number of non-zero elements $\mathcal{N}$ in the stiffness matrix $\mathcal{A}^{comp}_{\mathcal{I}_J(s,t,l,p,q)}$ is bounded by*

$$\begin{aligned}\mathcal{N} &\le CJ^{n(2+\frac{n}{f})}\ dim\left(V_{\mathcal{I}_J(s,t,l,p,q)}\right)^{1-\frac{n}{f}(N+1-t-l)} \\ &\le CJ^{n(2+\frac{n}{f})} 2^{nJ\frac{s-l-t}{s-l-2nt-q-np}2\left(1-\frac{n}{f}(N+1-t-l)\right)}.\end{aligned}$$

Specifically for $f(\mathbf{N+1}, \mathbf{N+1}) = 2n(N+1+m), m \in \mathbb{R}^+$, it holds

$$\mathcal{N} \le O(J^n) \cdot 2^{nJ\frac{s-l-t}{s-l-2nt-q-np}\left(1+\frac{t+l+m}{N+1+m}\right)}.$$

Proof: For shortness we use again the abbreviations $f = f(\mathbf{N+1}, \mathbf{N+1})$ as well as $\mathcal{I}_J := \mathcal{I}_J(s,t,l,p,q)$ and $V_{\mathcal{I}_J} := V_{\mathcal{I}_J(s,t,l,p,q)}$ for the index set (3.3) and the corresponding approximation space. The number of non-zero elements in the stiffness matrix $\mathcal{A}^{comp}_{\mathcal{I}_J}$ is given by

$$(4.5)\ \mathcal{N} = \sum_{(\mathbf{j},\mathbf{k})\in\mathcal{I}_\mathbf{J}} \sum_{(\mathbf{l},\mathbf{m})\in\tau^{comp}_{\mathbf{jk}}} 1 \overset{(4.3)}{=} \sum_{(\mathbf{j},\mathbf{k})\in\mathcal{I}_\mathbf{J}} \sum_{\{(\mathbf{l},\mathbf{m})\in\tau_{\mathbf{jk}}: dist(supp(\psi_{\mathbf{jl}}), supp(\psi_{\mathbf{km}}))\le B_{\mathbf{jk}}\}} 1.$$

Now, for every fixed index pair $(\mathbf{j}, \mathbf{l})$ the number of pairs $(\mathbf{k}, \mathbf{m})$ with

$$dist(supp(\psi_{\mathbf{jl}}), supp(\psi_{\mathbf{km}})) \le B_{\mathbf{jk}}$$

is bounded by

$$\begin{aligned}C \cdot \prod_{i=1}^n \max\left(1, \frac{2^{-j_i} + B_{\mathbf{jk}}}{2^{-k_i}}\right) &= C \cdot \prod_{i=1}^n \max\left(1, \frac{2^{-j_i}}{2^{-k_i}} + \frac{B_{\mathbf{jk}}}{2^{-k_i}}\right) \\ &\le C \cdot \sum_{\substack{\bar\mu,\bar\nu:\bar\mu\cup\bar\nu=\{1,\dots,n\},\\ \bar\mu\cap\bar\nu=\emptyset}} \left\{\prod_{i\in\bar\mu} \max\left(1, \frac{2^{-j_i}}{2^{-k_i}}\right) \cdot \prod_{i\in\bar\nu} \max\left(1, \frac{B_{\mathbf{jk}}}{2^{-k_i}}\right)\right\}\end{aligned}$$

$$(4.6) \qquad \leq C \cdot \max_{\substack{\bar\mu,\bar\nu:\bar\mu\cup\bar\nu=\{1,\ldots,n\},\\ \bar\mu\cap\bar\nu=\emptyset}} \left\{ \prod_{i\in\bar\mu} \max\left(1, \frac{2^{-j_i}}{2^{-k_i}}\right) \cdot \prod_{i\in\bar\nu} \max\left(1, \frac{B_{\mathbf{jk}}}{2^{-k_i}}\right) \right\}.$$

Because of

$$(4.7) \qquad \prod_{i\in\bar\mu} \max\left(1, \frac{2^{-j_i}}{2^{-k_i}}\right) \leq C \cdot \max_{\{\mu:\mu\subseteq\bar\mu\}} 2^{\sum_{i\in\mu}(k_i-j_i)}$$

and similarly

$$(4.8) \qquad \prod_{l\in\bar\nu} \max\left(1, \frac{B_{\mathbf{jk}}}{2^{-k_l}}\right) \leq C \cdot \max_{\{\nu:\nu\subseteq\bar\nu\}} 2^{\sum_{l\in\nu} k_l} B_{\mathbf{jk}}^{|\nu|}$$

we obtain

$$\begin{aligned}
\mathcal{N} &= \sum_{(\mathbf{j},\mathbf{k})\in\mathcal{I}_\mathbf{J}} \sum_{\{(\mathbf{l},\mathbf{m})\in\tau_{\mathbf{jk}}: dist(supp(\psi_{\mathbf{jl}}),supp(\psi_{\mathbf{km}}))\leq B_{\mathbf{jk}}\}} 1 \\
&\overset{(4.6)}{\leq} C \sum_{(\mathbf{j},\mathbf{k})\in\mathcal{I}_\mathbf{J}} 2^{|\mathbf{j}|_\mathbf{1}} \max_{\substack{\bar\mu,\bar\nu:\bar\mu\cup\bar\nu=\{1,\ldots,n\},\\ \bar\mu\cap\bar\nu=\emptyset}} \left\{ \prod_{i\in\bar\mu} \max\left(1, \frac{2^{-j_i}}{2^{-k_i}}\right) \cdot \prod_{l\in\bar\nu} \max\left(1, \frac{B_{\mathbf{jk}}}{2^{-k_l}}\right) \right\} \\
&\overset{(4.7),(4.8)}{\leq} C \sum_{(\mathbf{j},\mathbf{k})\in\mathcal{I}_\mathbf{J}} 2^{|\mathbf{j}|_\mathbf{1}} \max_{\substack{\mu,\nu:\mu\cup\nu\subseteq\{1,\ldots,n\},\\ \mu\cap\nu=\emptyset}} \left\{ 2^{\sum_{i\in\mu}(k_i-j_i)}) 2^{\sum_{i\in\nu} k_i} B_{\mathbf{jk}}^{|\nu|} \right\} \\
&\leq C \sum_{(\mathbf{j},\mathbf{k})\in\mathcal{I}_\mathbf{J}} \max_{\substack{\mu,\nu:\mu\cup\nu\subseteq\{1,\ldots,n\},\\ \mu\cap\nu=\emptyset}} \left\{ 2^{(\sum_{i\in\mu\cup\nu} k_i)+(\sum_{i\notin\mu} j_i)} B_{\mathbf{jk}}^{|\nu|} \right\} \\
&\leq CJ^{2n} \max_{(\mathbf{j},\mathbf{k})\in\mathcal{I}_\mathbf{J}} \max_{\substack{\mu,\nu:\mu\cup\nu\subseteq\{1,\ldots,n\},\\ \mu\cap\nu=\emptyset}} \left\{ 2^{(\sum_{i\in\mu\cup\nu} k_i)+(\sum_{i\notin\mu} j_i)} B_{\mathbf{jk}}^{|\nu|} \right\}.
\end{aligned}$$

Now we plug in the definition of $B_{\mathbf{jk}}$ and obtain

$$\mathcal{N} \leq CJ^{n(2+\frac{n}{f})} 2^K$$

with

$$\begin{aligned}
K := \max_{(\mathbf{j},\mathbf{k})\in\mathcal{I}_\mathbf{J}} \max_{\substack{\mu,\nu:\mu\cup\nu\subseteq\{1,\ldots,n\},\\ \mu\cap\nu=\emptyset}} &\left\{ \sum_{i\in\mu\cup\nu} k_i + \sum_{i\notin\mu} j_i \right. \\
&\left. + \frac{|\nu|}{f} \left[(t+l-s)J + s|\mathbf{j}|_\infty - q|\mathbf{k}|_\infty - p|\mathbf{k}|_1 - (N+1)|\mathbf{j},\mathbf{k}|_\mathbf{1} \right] \right\}.
\end{aligned}$$

The maximum in μ, ν is obtained for $\mu = \emptyset, \nu = \{1, \ldots, n\}$. Hence

$$K \leq \max_{(\mathbf{j},\mathbf{k})\in\mathcal{I}_\mathbf{J}} \left\{ |\mathbf{j},\mathbf{k}|_1 + \frac{n}{f} \left[(t+l-s)J + s|\mathbf{j}|_\infty - q|\mathbf{k}|_\infty - p|\mathbf{k}|_1 - (N+1)|\mathbf{j},\mathbf{k}|_\mathbf{1} \right] \right\}$$

$$
\begin{aligned}
&\le \max_{(\mathbf{j},\mathbf{k})\in\mathcal{I}_J} \Bigg\{ \left(1-\frac{n}{f}(N+1-t-l)\right)|\mathbf{j},\mathbf{k}|_1 \\
&\qquad\qquad + \frac{n}{f}\left[(-t-l)|\mathbf{j},\mathbf{k}|_1 + s|\mathbf{j}|_\infty - q|\mathbf{k}|_\infty - p|\mathbf{k}|_1 + (t+l-s)J\right]\Bigg\} \\
&\le \max_{(\mathbf{j},\mathbf{k})\in\mathcal{I}_J} \Bigg\{ \left(1-\frac{n}{f}(N+1-t-l)\right)|\mathbf{j},\mathbf{k}|_1 \\
&\qquad\qquad + \frac{n}{f}\left[s|\mathbf{j}|_\infty - q|\mathbf{k}|_\infty - p|\mathbf{k}|_1 - t|\mathbf{j},\mathbf{k}|_1 - l|\mathbf{j},\mathbf{k}|_\infty + (t+l-s)J\right]\Bigg\}.
\end{aligned}
$$

Using the definition of $\mathcal{I}_J$ in (3.3) we obtain

$$
K \le \max_{(\mathbf{j},\mathbf{k})\in\mathcal{I}_\mathbf{J}} \left(1-\frac{n}{f}(N+1-t-l)\right)|\mathbf{j},\mathbf{k}|_\mathbf{1}
$$

and hence together with the third inequality in (3.5)

$$
\begin{aligned}
\mathcal{N} &\le CJ^{n(2+\frac{n}{f})} \dim(V_{\mathcal{I}_J})^{1-\frac{n}{f}(N+1-t-l)} \\
&\le CJ^{n(2+\frac{n}{f})} 2^{nJ\frac{s-l-t}{s-l-2nt-q-np}2\left(1-\frac{n}{f}(N+1-t-l)\right)}.
\end{aligned}
$$

□

Lemma 1 shows that for $N+1 \approx t+l$ there will be no compression effect. However, for $N+1 > t+l$ there will be a significant reduction of the number of non-zero entries in the stiffness matrix. Combining Theorem 2 with the estimate for the number of non-zero elements in the stiffness matrix in Lemma 1 we obtain the following result.

THEOREM 3. *Under the assumptions of Theorem 2 and* $0 \le q+np < l-s$ *the number of non-zero elements in the stiffness matrix* $\mathcal{A}^{comp}_{\mathcal{I}_J(s,t,l,p,q)}$ *needed to obtain* $\|(A - A^{comp}_{\mathcal{I}_J(s,t,l,p,q)})u\|_{\mathcal{H}^s} \le \epsilon$, $\epsilon \in \mathbb{R}^+$, *is bounded by*

$$
\mathcal{N} \le O\left(\epsilon^{\frac{2}{\frac{s-l-q}{n}-p-2t}(1-\frac{n}{f}(N+1-t-l))} \ln(\epsilon^{-1})^{n^2}\right). \tag{4.9}
$$

Specifically for $f(\mathbf{N+1},\mathbf{N+1}) = 2n(N+1+m)$ *with* $m \in \mathbb{R}^+$ *it holds*

$$
\mathcal{N} \le O\left(\epsilon^{\frac{1}{\frac{s-l-q}{n}-p-2t}(1+\frac{t+l+m}{N+1+m})} \ln(\epsilon^{-1})^{n}\right). \tag{4.10}
$$

A comparison of (4.9) with the last estimate in (3.6) shows that compression reduces the quadratic exponential factor in the right hand side of (3.6) to $2-2\frac{n}{f}(N+1-t-l))$. Additionally, (4.10) shows that for $t+l+m/N+1+m \to 0$ we obtain optimal compression up to a logarithmic factor.

Note however, that generalizations to more general geometries are not as easy as in the isotropic case.

REFERENCES

[1] K. I. Babenko, *Approximation by trigonometric polynomials in a certain class of periodic functions of several variables*, Dokl. Akad. Nauk SSSR 132 (1960), 672-675.

[2] G. Baszenski, *N-th order polynomial spline blending*, in Multivariate approximation theory III, W. Schemp, K. Zeller (eds.), Birkhäuser, Basel, 1985, 35-46.

[3] G. Beylkin, R. Coifman, V. Rokhlin, *Fast wavelet transforms and numerical algorithms I*, Comm. Pure and Appl. Math. 44 (1991), 141-183.

[4] H.-J. Bungartz, M. Griebel, D. Röschke, C. Zenger, *Pointwise convergence of the combination technique for Laplace's equation*, East-West J. of Numerical Mathematics 1(2) (1994), 21-45.

[5] W. Dahmen, *Stability of Multiscale Transformations*, Journal of Fourier Analysis and Applications 2 (1996), 341-361.

[6] W. Dahmen, *Wavelet and Multiscale Methods for Operator equations*, Acta Numerica (1997), 55-228.

[7] W. Dahmen, S. Prößdorf, R. Schneider, *Wavelet approximation methods for pseudodifferential equations II: Matrix compression and fast solution*, Advances in computational Mathematics 1 (1993), 259-335.

[8] F.-J. Delvos, *d-Variate Boolean interpolation*, J. Approx. Theory 34 (1982), 99-114.

[9] R. A. DeVore, B. Jawerth, V. Popov, *Compression of Wavelet decompositions*, Amer. J. Math. 114 (1992), 737-785.

[10] Din' Zung, *The approximation of classes of periodic functions of many variables*, Russion Math. Surveys 38 (1983), 117-118.

[11] Din' Zung, *Number of integral points in a certain set and the approximation of functions of several variables*, Math. Notes of the Academy of Sciences of the USSR 36(4) (1984), 736-744.

[12] K. Frank, S. Heinrich, S. Pereverzev, *Information complexity of multivariate Fredholm equations in sobolev classes*, J. Complexity 12 (1996), 17-34.

[13] E. M. Galeev, *Approximation by fourier sums of functions with several bounded derivatives*, Math. Notes 23 (1978), 109-117.

[14] W. J. Gordon, *Blending-function methods for bivariate and multivariate interpolation and approximation*, SIAM J. Numerical Analysis 8 (1971), 158-177.

[15] M. Griebel, S. Knapek, *Optimized tensor-product approximation spaces*, Constructive Approximation 16(4) (2000), 525-540.

[16] M. Griebel, P. Oswald, T. Schiekofer, *Sparse grids for boundary integral equations*, Numer. Math. 83(2) (1999), 279-312.

[17] R. Hochmuth, S. Knapek, G. Zumbusch, *Tensor-products of Sobolev spaces and Applications*, SFB-256 Bericht Nr. 685, Universität Bonn, 2000.

[18] S. Knapek, *Hyperbolic cross approximation of integral operators with smooth kernel*, SFB-256 Bericht Nr. 665, Universität Bonn, 2000.

[19] S. Knapek, F. Koster *Integral operators on sparse grids*, SIAM J. Numer. Anal. 39(5) (2002), 1794-1809.

[20] Y. Meyer, R. Coifman *Wavelets: Calderon-Zymund and multilinear operators*, Cambridge studies in advanced mathematics 48, Cambridge University Press (1997).

[21] P. Oswald, *On discrete norm estimates related to multilevel preconditioners in the finite element method*, in: Constructive Theory of Functions, K.G. Ivanov, P. Petrushev, B. Sendov, (eds.), Proc. Int. Conf. Varna, 1991, Bulg. Acad. Sci., Sofia, 1992, 203-214.

[22] S. V. Pereverzev, *On the complexity of finding solutions of Fredholm equations of the second kind with differentiable kernels, I*, Ukrain. Mat. Zh. 40 (1988), 84-91.

[23] R. Schneider, *Multiskalen- und Wavelet-Kompression: Analysisbasierte Methoden zur effizienten Lösung großer vollbesetzter Gleichungssysteme*, Advances in Numerical Analysis, Teubner, Stuttgart, 1998.

[24] S. A. Smolyak, *Quadrature and interpolation formulas for tensor products of certain classes of functions*, Dokl. Akad. Nauk SSSR 4 (1963), 240-243.

[25] V. N. Temlyakov, *Approximation of periodic functions*, Nova Science, New York, 1993.

[26] J. F. Traub, G. W. Wasilkowski, H. Wozniakowski, *Information-Based Complexity*, Academic Press, New York, 1988.

[27] C. Zenger, *Sparse grids*, Parallel Algorithms for Partial Differential Equations, Notes on Num. Fluid Mech. 31, W. Hackbusch (ed.), Vieweg, Braunschweig, 1991.