OPTIMIZED TENSOR-PRODUCT APPROXIMATION SPACES

M. GRIEBEL AND S. KNAPEK
INSTITUT FÜR ANGEWANDTE MATHEMATIK
ABTEILUNG FÜR WISSENSCHAFTLICHES RECHNEN UND NUMERISCHE SIMULATION
UNIVERSITÄT BONN
D-53115 BONN
{GRIEBEL,KNAPEK}@IAM.UNI-BONN.DE

Abstract. This paper is concerned with the construction of optimized grids and approximation spaces for elliptic differential and integral equations. The main result is the analysis of the approximation of the embedding of the intersection of classes of functions with bounded mixed derivatives in standard Sobolev spaces. Based on the framework of tensor-product biorthogonal wavelet bases and stable subspace splittings, the problem is reduced to diagonal mappings between Hilbert sequence spaces. We construct operator adapted finite-element subspaces with a lower dimension than the standard full-grid spaces. These new approximation spaces preserve the approximation order of the standard full-grid spaces, provided that certain additional regularity assumptions are fulfilled. The form of the approximation spaces is governed by the ratios of the smoothness exponents of the considered classes of functions. We show in which cases the so called curse of dimensionality can be broken. The theory covers elliptic boundary value problems as well as boundary integral equations. [1]

AMS subject classification. 41A25, 41A30, 41A65, 45L10, 65N12, 65N30, 65Y20, 68Q25

Key words. biorthogonal wavelets, optimized approximation spaces, hyperbolic cross, sparse grids, norm equivalences, subspace splittings, partial differential equations, integral equations, elliptic problems

1. Introduction.

In this paper we deal with the construction of finite element spaces for the approximate solution of elliptic problems in Sobolev spaces $\mathcal{H}^s(\Omega), s \in \mathbb{R}$. It is well-known that the ϵ-complexity [TWW88, Wer91] of solving Poissons equation in n dimensions in the Sobolev space $\mathcal{H}^1_0 \cap \mathcal{H}^2$ on a bounded domain is $O(\epsilon^{-n})$, i.e., it is exponentially dependent on n. This dependence on n is called the curse of dimensionality. Hence, for higher-dimensional problems, a direct numerical solution on a regular standard triangulation is prohibitive [TWW88]. However, if the solution is in the space of functions with dominating mixed second derivative the complexity reduces to $O(\epsilon^{-1})$ [BG98]. Compare also [Wer95] for related results concerning the complexity of the Poisson equation with right-hand side in the space of functions with dominating mixed derivative.

Now, standard finite element algorithms that use regular full grids for the discretization lead to a cost of $O(\epsilon^{-n})$. Hence they are not suited for such problems. A corresponding algorithm that realizes the cost of $O(\epsilon^{-1})$ up to an additional logarithmic term has been presented in [Z91, Gr91]. It uses tensor-products of linear splines as basis functions, and the so called sparse grid space/hyperbolic cross space as approximation space in the finite element method. Together with an optimal multigrid solver for sparse grid discretizations [GO95, Gri94a, Pf98], the cost of this algorithm is $O(\epsilon^{-1} \cdot |\ln \epsilon|^{n-1})$. See also the subsequent papers [GSZ92, Bu92]. In [BG98] it was shown shown how to get rid of the additional logarithmic term by the use of a subspace of the sparse grid space. Then the overall cost for the solution of Poissons equation is indeed of the order $O(\epsilon^{-1})$.

---

[1] M. Griebel, S. Knapek, Optimized tensor-product approximation spaces, Constructive Approximation, to appear

The sparse grid scheme is well established in approximation and interpolation theory [Bab60, BDJ92, Del82, Go69, Sm63, Te93, Te89, WW95]. It continues to attract significant attention [Ka96, DKT98, SS98] and has also been used successfully in connection with other hierarchical basis functions, for example prewavelets [GO95], interpolets [GKo99] and higher order splines [BD97]. Generalizations of hyperbolic trigonometric approximation may be found in [D83, D88, G78].
Only recently, [GOS99] used the sparse grid approach together with prewavelets for the solution of boundary integral equations of order between $-1/2$ and $1/2$. It has been observed that adaptive sparse grid spaces lead to good approximation rates for the single layer potential equation on a square. In [O98] this has been studied theoretically in a simplified setting via N-term best approximation. Note that for operators of negative order algorithms that use the sparse grid approach have not yet been studied in depth, although the application of the sparse grid approach to boundary integral equations seems to be particularly tempting, as it helps to avoid the matrix compression step to some extent. For a study of the complexity of Fredholm integral equations of the second kind with smooth kernels we refer to [Pe88, Pe89, FHP96]. In contrast to our investigation they use sparse grid type spaces for the kernel approximation and not as ansatz- and test-spaces in the finite element discretization.
In this paper we generalize the construction of [BG98] to operators of arbitrary order. We construct operator adapted finite-element subspaces of lower dimension than the standard full-grid spaces which preserve the approximation order of the standard full-grid spaces, provided that certain additional regularity assumptions are fulfilled. We analyse the approximation of the embedding $\mathcal{H}^{t,l}_{mix} \to \mathcal{H}^s$ on the n-dimensional torus. Here $\mathcal{H}^{t,l}_{mix}$ is a certain intersection of classes of functions with bounded mixed derivatives, see (4) below, and $\mathcal{H}^s$ is the standard Sobolev space. The parameter l in $\mathcal{H}^{t,l}_{mix}$ governs the isotropic smoothness whereas t governs the dominating mixed smoothness. We use norm equivalences to facilitate the decoupling of the subspaces arising from the tensor-product ansatz and to ensure the stability of the resulting subspace splittings. Hence, the analysis is reduced to diagonal mappings between Hilbert sequence spaces. It turns out that the optimal approximation space in terms of the ratio of cost versus accuracy is only dependent on the quotient $\frac{s-l}{t}$ of isotropic smoothness $s-l$ to dominating mixed smoothness t. We identify those approximation spaces that lead to algorithms with minimal cost. Specifically we show that one can break the curse of dimension in the case $\frac{s-l}{t} > 0$ and get rid of all asymptotic dependencies on n. In the case $\frac{s-l}{t} < 0$ (for example for operators of negative order and spaces with dominating mixed derivative) there remains a certain moderate dependence on the dimension.
The remainder of this paper is as follows. Section 2 introduces the necessary notation, summarizes basic facts about biorthogonal wavelet bases and tensor-product spaces, introduces the Sobolev spaces under consideration and presents some results about norm equivalences in these Sobolev spaces. In Section 3 we define our new approximation spaces and give estimates on their dimension and their order of approximation. Section 4 contains remarks on the cost of solving elliptic equations in case that continuous linear information is permissible. We show these results in a constructive manner by proposing a finite element method working with the optimized approximation spaces together with

a multilevel preconditioned iterative solver. Some concluding remarks close the paper.

2. Tensor-product multiscale bases, Sobolev spaces and norm equivalences. Let us denote by $\mathcal{H}^t(T^n), t \in \mathbb{R}$, a scale of standard Sobolev spaces on the n-dimensional torus $T^n = T \times T \times \cdots \times T$, $T := [0,1]$ (opposite faces identified), and by $\mathcal{L}^2(T^n)$ the space of $\mathcal{L}^2$-integrable functions on T^n, [Ad75]. When $t < 0, \mathcal{H}^t(T^n)$ is defined as the dual of $\mathcal{H}^{-t}(T^n)$, i.e., $\mathcal{H}^t(T^n) := (\mathcal{H}^{-t}(T^n))'$. If the meaning is clear from the context, we will write $\mathcal{H}^t$ instead of $\mathcal{H}^t(T^n)$ and we proceed analogously for other function spaces. In the following multi-indices (vectors) are written in boldface, for example $\mathbf{j}$ for $(j_1, \ldots, j_n)$. Inequalities like $\mathbf{l} \leq \mathbf{t}$ or $\mathbf{l} \leq 0$ are to be understood componentwise. We define $|\mathbf{j}|_\infty = \max\limits_{1 \leq i \leq n} j_i$ and $|\mathbf{j}|_1 = \sum_{i=1}^n |j_i|$. C denotes a generic constant, which may depend on the smoothness assumptions and on the dimension n of the problem. $x \approx y$ means that there exist constants C_1, C_2 independent of any parameters x or y may depend on, such that $C_1 \cdot y \leq x \leq C_2 \cdot y$.
Consider an elliptic variational problem: Given $f \in \mathcal{H}^{-s}$, $s \in \mathbb{R}$, find $u \in \mathcal{H}^s$ such that

$$a(u, v) = (f, v) \ \ \forall v \in \mathcal{H}^s, \tag{1}$$

where a is a symmetric positive definite form satisfying

$$a(v, v) \approx \|v\|^2_{\mathcal{H}^s}. \tag{2}$$

The Galerkin method for solving problem (1) numerically is to select a finite-dimensional subspace S_J from $\mathcal{H}^s \cap \mathcal{L}^2$ and to solve the variational problem therein. Fixing a basis of S_J finally leads to a linear system of equations

$$A_J x_J = b_J \tag{3}$$

of dimension $dim(S_J)$. Storage requirements and computation time exclude the use of direct solvers, since $dim(S_J)$ is usually very large. It is well-known that the most efficient methods to solve such problems exploit the interaction of several scales of discretization. These multilevel schemes use a nested dense sequence of closed subspaces $S_0 \subset S_1 \subset \ldots \subset \mathcal{H}^s \cap \mathcal{L}^2$ of the basic Hilbert space $\mathcal{H}^s$. The work count to solve (3) to a prescribed accuracy is then only proportional to the number of non-zero entries in A_J provided that $\{dim(S_j)\}_j$ forms a geometric series. But still, specifically for conventional spaces that belong to uniform full grids with subdivision rate two, this number is at least of the order $dim(S_J) = O(2^{nJ})$, i.e. it grows exponentially with n. This makes matrix-vector multiplications, as they are needed in iterative methods, prohibitively expensive for large n or large J.
This is particularly problematic in the case of integral operators, where the discretization results, in most cases, in discrete systems that are dense. That is, on a regular full grid with $O(2^{nJ})$ unknowns, the discrete operator has $O(2^{2nJ})$ entries.

2.1. Sobolev spaces. Now, we define the spaces $\mathcal{H}^{t,l}_{mix}$. They fix the smoothness assumptions we consider.

DEFINITION 1. *Let $t \in \mathbb{R}_0^+, l \in \mathbb{R}, t + l \geq 0,$ $\mathbf{1} = (1, \ldots, 1)$ and $\mathbf{e_i} = (0, \ldots, 0, 1, 0, \ldots, 0)$ the i-th unit-vector in $\mathbb{R}^n$.*

$$\mathcal{H}^{t,l}_{mix}(T^n) := \mathcal{H}^{t\mathbf{1}+l\mathbf{e_1}}_{mix}(T^n) \cap \cdots \cap \mathcal{H}^{t\mathbf{1}+l\mathbf{e_n}}_{mix}(T^n), \tag{4}$$

where

$$\mathcal{H}^{\mathbf{k}}_{mix}(T^n) := \mathcal{H}^{k_1}(T) \otimes \ldots \otimes \mathcal{H}^{k_n}(T).$$

$\mathcal{H}^{t,l}_{mix}$ is a mixture of classes of functions of dominating mixed derivative

$$\mathcal{H}^{t}_{mix}(T^n) = \mathcal{H}^{t,0}_{mix}(T^n). \tag{5}$$

and standard Sobolev spaces

$$\mathcal{H}^{t}(T^n) = \mathcal{H}^{0,t}_{mix}(T^n). \tag{6}$$

To prove (6) choose an orthogonal basis of $\mathcal{H}^t(T)$ and $\mathcal{L}^2(T)$, use periodic continuation to $\mathbb{R}$ and use the definition of the tensor-product via orthonormal systems, [Wei80]. Note that similar results hold for problems on I^n with Dirichlet or Neumann boundary conditions and certain cases of mixed boundary conditions.
Note that the relations $\mathcal{H}^t_{mix} \subset \mathcal{H}^t \subset \mathcal{H}^{t/n}_{mix}$ for $t \geq 0$ and $\mathcal{H}^{t/n}_{mix} \subset \mathcal{H}^t \subset \mathcal{H}^t_{mix}$ for $t \leq 0$ hold. See [ST87] and [Ho99] for more information on the spaces $\mathcal{H}^t_{mix}$.

2.2. Tensor-product bases. The approximation spaces considered here are tensor-products [Wei80] of univariate function spaces. We start from a one-dimensional splitting $\mathcal{L}^2 = \bigoplus_{j \geq 0} S_j$ and we assume that the complement spaces $W_j = S_j \ominus S_{j-1}$ of S_{j-1} in S_j are spanned by some biorthogonal wavelets ψ_{jk}, i.e. $W_j = span\{\psi_{jk}, k \in \tau_j\}$, where τ_j is some index set defined from the subdivision rate of successive refinement levels. Moreover, we assume that the $\{\psi_{jk}, k \in \tau_j\}$ form a Riesz-basis of W_j, i.e.

$$\| \sum_{k \in \tau_j} C_k \psi_{jk} \|_{\mathcal{L}^2} \approx \| \{C_k\}_k \|_{\ell^2(\tau_j)} \tag{7}$$

where, as usual, $\|.\|_{\mathcal{L}^2}$ denotes the norm induced from the scalar product on $\mathcal{L}^2$ and $\|\{C_k\}_k\|^2_{\ell^2(\tau_j)} = \sum_{k \in \tau_j} |C_k|^2$. Here we stick to dyadic refinement.
The dual system is denoted by a tilde. Then every $u \in \mathcal{L}^2$ has a unique expansion

$$u = \sum_{j=0}^{\infty} \sum_{k \in \tau_j} \langle u, \tilde{\psi}_{jk} \rangle \psi_{jk} = \sum_{j=0}^{\infty} \sum_{k \in \tau_j} \langle u, \psi_{jk} \rangle \tilde{\psi}_{jk}$$

and the dual system also forms a Riesz-basis in $\mathcal{L}^2$.
Estimates on the approximation order are mainly based on the local $\mathcal{L}^2$-stability (7), and an inequality of Jackson type, which in turn depends on estimates of the coefficients $\langle u, \tilde{\psi}_{jk} \rangle$, i.e. on a moment condition for the dual wavelets. For purposes of compression, one usually assumes specific decay properties of the Schwarz-kernel of the pseudodifferential operator under consideration. Then estimates of the size of the entries $a(\psi_{jk}, \psi_{lm})$

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 ψ_{jk}.
We assume the validity of the following norm equivalences

$$\|u\|^2_{\mathcal{H}^t} \approx \sum_{j=0}^{\infty} \|w_j\|^2_{\mathcal{H}^t} \approx \sum_{j=0}^{\infty} 2^{2tj} \|w_j\|^2_{\mathcal{L}^2} \text{ for } t \in (-\tilde{r}, r), \text{ where } u = \sum_{j=0}^{\infty} w_j, w_j \in W_j. \quad (8)$$

(For $t < 0$ the $\mathcal{L}^2$-convergence has to be replaced by distributional convergence.) Two-sided estimates like (8) can be inferred from the validity of a direct estimate (estimate of Jackson type, approximation order m) and an inverse estimate (Bernstein inequality) as a consequence of approximation theory in Sobolev spaces together with interpolation and duality arguments, see [D96, O92]. This allows the characterization of the smoothness properties of a function from the decay properties of the coefficients in a multiscale decomposition. A similar result holds for the analogous construction using the dual wavelets instead of the primal ones. This leads to the range $t \in (-r, \tilde{r})$ in (8). See [D97] for an overview on multiscale methods dealing with biorthogonal wavelets.
For the higher-dimensional case $n > 1$, let $\mathbf{j} \in \mathbb{Z}^n_+, \mathbf{j} \equiv (j_1, \ldots, j_n)$, be given, and consider the tensor-product partition with uniform step size 2^{-j_i} into the i-th coordinate direction. By $W_{\mathbf{j}}$ we denote the corresponding function space of tensor-products of one dimensional function spaces, i.e.

$$W_{\mathbf{j}} := W_{j_1} \otimes \ldots \otimes W_{j_n}.$$

A basis of $W_{\mathbf{j}}$ is then given by $\cup_{\mathbf{k} \in \tau_{\mathbf{j}}} \{\psi_{\mathbf{jk}}(\mathbf{x}) = \psi_{j_1 k_1}(x_1) \cdot \ldots \cdot \psi_{j_n k_n}(x_n)\}$.

2.3. Norm equivalences. To obtain norm equivalences analogous to (8) in $n \geq 2$ dimensions, we use the above representation of $\mathcal{H}^{t,l}_{mix}$ as tensor-products of 1D spaces and intersections.
We use the notation $\{V; a\}$ to denote a Hilbert space V equipped with the scalar-product $a(.,.)$. Consider a collection of Hilbert spaces $H_l, l = 1, \ldots, n, n \in \mathbb{N}$ and a collection of closed subspaces $V_{li} \subset H_l$ such that topologically $H_l = \sum_i V_{li}$. An additive subspace splitting $\{H_l; a_l\} = \sum_i \{V_{li}; b_{li}\}$ is called stable if the norm equivalence $a_l(u, u) \approx |||u|||^2 \equiv \inf_{u_i \in V_{li}: u = \sum_i u_i} \sum_i (b_{li}(u_i, u_i))$ holds true, i.e. the characteristic numbers

$$\lambda_{min,l} = \min_{0 \neq u \in H_l} \frac{a_l(u, u)}{|||u|||^2}, \ \lambda_{max,l} = \max_{0 \neq u \in H_l} \frac{a_l(u, u)}{|||u|||^2}, \ \kappa_l = \frac{\lambda_{max,l}}{\lambda_{min,l}}$$

are finite and positive. We cite two Propositions from [GO95]. (Note that we extend them from the two-dimensional to the n-dimensional case.)
PROPOSITION 1. *If the splittings*

$$\{H_l; a_l\} = \sum_i \{V_{li}; b_{li}\}, l \in \{1, \ldots, n\}, n \in \mathbb{N}$$

are stable and possess the condition number κ_l, respectively, then the tensor-product splitting

$$\{H_1 \otimes H_2 \otimes \ldots \otimes H_n; a_1 \otimes \ldots \otimes a_n\} = \sum_{i_1} \ldots \sum_{i_n} \{V_{1i_1} \otimes \ldots \otimes V_{ni_n}; b_{1i_1} \otimes \ldots \otimes b_{ni_n}\}$$

is also stable and possesses the condition number $\prod_{l=1}^{n} \kappa_l$.

PROPOSITION 2. *Let there be given sequences* $\{\alpha_{li}\}_i, l = 1,\ldots,n, n \in \mathbb{N}$. *Suppose that the splittings*

$$\{H_l; a_l\} = \sum_i \{V_i; \alpha_{li} b\}, l = 1,\ldots,n,$$

are stable and that the sums are direct. Then, for all $\alpha_l > 0, l = 1,\ldots,n$, *the splitting*

$$\{H_1 \cap \ldots \cap H_n; \alpha_1 a_1 + \ldots + \alpha_n a_n\} = \sum_i \{V_i; (\alpha_1\alpha_{1i} + \ldots + \alpha_n \alpha_{ni})b\}$$

is stable with condition number $\kappa \le \frac{\max\{\lambda_{max,1},\ldots,\lambda_{max,n}\}}{\min\{\lambda_{min,1},\ldots,\lambda_{min,n}\}}$.

Combining the above representation (6) of $\mathcal{H}^t(T^n)$ for $t \ge 0$, with these Propositions and the stability in one dimension we obtain the following norm equivalence and stable splitting of $\mathcal{H}^t(T^n)$.

THEOREM 1. *Let* $u \in \mathcal{H}^t(T^n), u = \sum_{\mathbf{j}} w_{\mathbf{j}}, w_{\mathbf{j}} \in W_{\mathbf{j}}$, *and let (8) hold for the primal as well as the dual system. Then*

$$\|u\|^2_{\mathcal{H}^t} \quad \approx \quad \sum_{\mathbf{j}} 2^{2t|\mathbf{j}|_\infty} \|w_{\mathbf{j}}\|^2_{\mathcal{L}^2} \textit{ for } t \in (-\tilde{r}, r). \tag{9}$$

Proof: From (8) we have the stability of the splitting $\{\mathcal{H}^t(T); \|.\|^2_{\mathcal{H}^t(T)}\} = \sum_j \{W_j; 2^{2tj}\|.\|^2_{\mathcal{L}^2}\}$. Then, from Proposition 1, we obtain the stability of the splittings

$$\begin{aligned}\{\mathcal{H}_{mix}^{(0,\ldots,0,t,0,\ldots,0)}; \quad & (.,.)_{\mathcal{L}^2} \otimes \ldots \otimes (.,.)_{\mathcal{L}^2} \otimes a(.,.) \otimes (.,.)_{\mathcal{L}^2} \otimes \ldots \otimes (.,.)_{\mathcal{L}^2}\} \\ & = \textstyle\sum_{\mathbf{j}} \{W_{j_1} \otimes \ldots \otimes W_{j_n}; 2^{2tj_i}(.,.)_{\mathcal{L}^2} \otimes \ldots \otimes (.,.)_{\mathcal{L}^2}\}.\end{aligned}$$

Representing $\mathcal{H}^t(T^n)$ as in (6) and applying Proposition 2, we obtain the stability of the splitting $\{\mathcal{H}^t(T^n); \|.\|^2_{\mathcal{H}^t}\} = \sum_{\mathbf{j}} \{W_j; (\sum_{i=1}^n 2^{2tj_i})\|.\|^2_{\mathcal{L}^2}\}$ for nonnegative $t < r$. To obtain the validity of (9) for $-\tilde{r} < t < 0$ note that the same reasoning as above applied to the representation of u in the dual wavelet system shows that we have an analogous result for the spaces spanned by the dual wavelets for $0 \le t < \tilde{r}$. The assertion follows by duality $(\mathcal{H}^t)' = \mathcal{H}^{-t}$.

□

Similarly, by combining the representation (4) of $\mathcal{H}^{t,l}_{mix}(T^n)$ with the above Propositions and the stability in one dimension we obtain

$$\|u\|^2_{\mathcal{H}^{t,l}_{mix}} \approx \sum_{\mathbf{j}} \left(\sum_{i=1}^n 2^{2t|\mathbf{j}|_1 + 2lj_i} \right) \|w_{\mathbf{j}}\|^2_{\mathcal{L}^2} \approx \sum_{\mathbf{j}} 2^{2t|\mathbf{j}|_1 + 2l|\mathbf{j}|_\infty} \|w_{\mathbf{j}}\|^2_{\mathcal{L}^2}, \tag{10}$$
$$t \ge 0, 0 \le t + l < r.$$

Note that for $t = 0$ or $l = 0$ we regain the (standard) norm equivalence (9) for the Sobolev space $\mathcal{H}^t$. For the space $\mathcal{H}^t_{mix}$ we obtain the equivalence $\|u\|^2_{\mathcal{H}^t_{mix}} \approx \sum_{\mathbf{j}} 2^{2t|\mathbf{j}|_1} \|w_{\mathbf{j}}\|^2_{\mathcal{L}^2}$. The different factors $2^{2t|\mathbf{j}|_1}$ and $2^{2l|\mathbf{j}|_\infty}$ in these equivalences reflect the different smoothness requirements.

One of the merits of the norm equivalence (9) is the fact that it leads directly to optimal preconditioning. If one chooses the scaled system $\{2^{-s|\mathbf{l}|_\infty}\psi_{\mathbf{lk}} : |\mathbf{l}|_\infty \le J, \mathbf{k} \in \tau_{\mathbf{l}}\}$ as the basis in the finite element approximation space $V_J^{-\infty}$, then the spectral condition numbers $\kappa(A_J)$ of the discretization matrices $A_J = \{2^{-s|\mathbf{l}+\mathbf{l}'|_\infty} a(\psi_{\mathbf{lk}}, \psi_{\mathbf{l}'\mathbf{k}'})\}_{\mathbf{l},\mathbf{l}',\mathbf{k},\mathbf{k}'}$ are bounded uniformly in J, i.e. $\kappa(A_J) = O(1)$, see [DK92, J92, Be93]. This leads to fast iterative methods with convergence rates independent of the number of unknowns of the approximation space.

3. Optimized approximation spaces for Sobolev spaces. Suppose now that a symmetric elliptic variational problem (1) and its variational formulation $a(u_{FE}, v) = (f, v)\ \forall v \in V_{FE}$ on a finite element approximation space $V_{FE} \subset \mathcal{H}^s$ are given. In this section we derive bounds on the term $\inf_{v\in V_{FE}} \|u - v\|_{\mathcal{H}^s}$ for various choices of the approximation space V_{FE}, under the constraint $u \in \mathcal{H}^{t,l}_{mix}$. We define grids and associated approximation spaces which are adapted to the parameters s, t and l and give estimates on their dimension and the order of approximation.
Given index sets $I_J \subset \mathbb{Z}^n_+, J \in \mathbb{N}$, we consider the approximation spaces $V_J := \sum_{\mathbf{j}\in I_J} W_{\mathbf{j}}$. Here, J is the maximal level in V_J, i.e. $j_i \le J, i = 1, \ldots, n\ \forall \mathbf{j} \in I_J$. The full grid space

$$V_J^{-\infty} := \bigoplus_{|\mathbf{j}|_\infty \le J} W_{\mathbf{j}}, J > 0.$$

is associated with the rectangular index set $I_J^{-\infty} := \{|\mathbf{j}|_\infty \le J\}$. The so called sparse grid space (hyperbolic cross space)

$$V_J^0 := \bigoplus_{|\mathbf{j}|_1 \le J+n-1} W_{\mathbf{j}}, J > 0,$$

is associated with the index set $I_J^0 := \{|\mathbf{j}|_1 \le J + n - 1\}$. The dimensions of $W_{\mathbf{j}}$, $V_J^{-\infty}$ and V_J^0 are (note that we count only interior grid points) $|W_{\mathbf{j}}| = 2^{|\mathbf{j}|_1 - n}$, $|V_J^{-\infty}| = (2^J - 1)^n = O(2^{Jn})$, and $|V_J^0| = 2^J\left(\frac{J^{n-1}}{(n-1)!} + O(J^{n-2})\right)$, see [Bu92, BG98].
The dimension of the sparse grid space V_J^0 compares favorably with the dimension of the full grid space, especially for higher dimensions. Statements about the usefulness of the sparse grid space for discretization purposes require estimates of the approximation power. It will be shown that, for certain classes of smooth functions and for $s > 0$, the order of approximation of the sparse grid space is the same as for the space $V_J^{-\infty}$. In some cases it is even possible to use a further subspace of the sparse grid space without loosing the optimal order of approximation.
To see this let $u = \sum_{\mathbf{j}} w_{\mathbf{j}}$, where $w_{\mathbf{j}} \in W_{\mathbf{j}}$ and let $-\tilde{r} < s < t < r$. For notational convenience we restrict ourselves to the case $t > 0$. The case $t < 0$ can be covered with analogous reasoning.
Let $u \in \mathcal{H}^t$. We apply the norm equivalence (9) and obtain with the approximation space $V_J^{-\infty}$ the standard error estimate

$$\inf_{v\in V_J^{-\infty}} \|u - v\|^2_{\mathcal{H}^s} \le C \cdot 2^{2(s-t)J} \|u\|^2_{\mathcal{H}^t} \text{ for } u \in \mathcal{H}^t \text{ and } -\tilde{r} < s < t < r. \tag{11}$$

From the exponent on the right hand side we obtain $t - s$ as order of approximation.

If we instead use the approximation space V_J^0 the situation is changed significantly. For $u \in \mathcal{H}^t$ we obtain

$$\inf_{v \in V_J^0} \|u - v\|_{\mathcal{H}^s}^2 \leq \|u - \sum_{|\mathbf{j}|_1 \leq J+n-1} w_{\mathbf{j}}\|_{\mathcal{H}^s}^2 \overset{(9)}{\approx} \sum_{|\mathbf{j}|_1 > J+n-1} 2^{2s|\mathbf{j}|_\infty} \|w_{\mathbf{j}}\|_{\mathcal{L}^2}^2$$
$$= \sum_{|\mathbf{j}|_1 > J+n-1} 2^{2(s-t)|\mathbf{j}|_\infty} 2^{2t|\mathbf{j}|_\infty} \|w_{\mathbf{j}}\|_{\mathcal{L}^2}^2 \leq \max_{|\mathbf{j}|_1 > J+n-1} 2^{2(s-t)|\mathbf{j}|_\infty} \sum_{|\mathbf{j}|_1 > n+J-1} 2^{2t|\mathbf{j}|_\infty} \|w_{\mathbf{j}}\|_{\mathcal{L}^2}^2$$
$$\overset{(9)}{\leq} C \cdot \max_{|\mathbf{j}|_1 > J+n-1} 2^{2(s-t)|\mathbf{j}|_\infty} \|u\|_{\mathcal{H}^t}^2 \leq C \cdot 2^{2(s-t)(1-\frac{1}{n})} 2^{2(s-t)\frac{J}{n}} \|u\|_{\mathcal{H}^t}^2 \text{ for } -\tilde{r} < s < t < r.$$

Compared to the result for the full grid approximation space, the order of approximation deteriorates from $t - s$ to $t - s/n$.

Now let us consider the case $u \in \mathcal{H}_{mix}^t$. For the smaller space $\mathcal{H}_{mix}^t \subset \mathcal{H}^t$ and operators of positive order, i.e., $s \geq 0$, no loss in the order of approximation occurs, if the full grid space $V_J^{-\infty}$ is replaced by the space V_J^0. This is due to the fact that we can apply norm equivalence (10) for $\mathcal{H}_{mix}^t$ instead of $\mathcal{H}^t$ (remember the different exponents of the forefactors in (10)). We get for $u \in \mathcal{H}_{mix}^t$ and $-\tilde{r} < s < t < r, t > 0$,

$$\inf_{v \in V_J^0} \|u - v\|_{\mathcal{H}^s}^2 \overset{(10)}{\leq} C \cdot \max_{|\mathbf{j}|_1 > J+n-1} 2^{2s|\mathbf{j}|_\infty - 2t|\mathbf{j}|_1} \|u\|_{\mathcal{H}_{mix}^t}^2 \leq C \cdot 2^{-2tn} 2^{2s} 2^{2(s-t)J} \|u\|_{\mathcal{H}_{mix}^t}^2. \quad (12)$$

Here the factor $2^{-2tn}2^{2s}$ is only a constant with respect to the asymptotics in J. Hence, there appears no loss in the order of approximation compared to the result (11) for the full grid approximation space.

For operators of negative order, i.e., $s < 0$, the situation is different. Here, compared to the estimate (11) for $u \in \mathcal{H}^t$, the order of approximation improves when changing to the space $\mathcal{H}_{mix}^t$, but in contrast to the case $s \geq 0$, the optimal order of convergence cannot be retained. If we apply again (10) we obtain for $u \in \mathcal{H}_{mix}^t$

$$\inf_{v \in V_J^0} \|u - v\|_{\mathcal{H}^s}^2 \overset{(10)}{\leq} C \cdot \max_{|\mathbf{j}|_1 > J+n-1} 2^{2s|\mathbf{j}|_\infty - 2t|\mathbf{j}|_1} \|u\|_{\mathcal{H}_{mix}^t}^2 \leq C \cdot 2^{2s(1-\frac{1}{n})} 2^{-2tn} 2^{2(s/n-t)J} \|u\|_{\mathcal{H}_{mix}^t}^2.$$

Here the order of approximation is improved when changing from $\mathcal{H}^t$ to $\mathcal{H}_{mix}^t$ but there is still a loss in the order of approximation of $s(1-1/n)$ compared to the full grid. This fact is described in [GOS99] for the case $-1 \leq s < 0$ and prewavelets (i.e., wavelets that are $\mathcal{L}^2$-orthogonal between different subspaces $W_{\mathbf{j}}$ and build a Riesz basis in the subspaces $W_{\mathbf{j}}$). In summary we have that for operators with $s \geq 0$ the order of approximation is preserved for $u \in \mathcal{H}_{mix}^t$, $s < t$, when changing from the approximation space $V_J^{-\infty}$ to the sparse grid space V_J^0. In contrast, for operators of negative order a slight deterioration of the order of approximation appears.

3.1. Optimized approximation spaces. A closer look on the last inequality in (12) reveals that for $s \geq 0$ one could even discard further indices from the index set I_J^0 without destroying the optimal order of approximation. This motivates the following definition of index sets and approximation spaces.

Definition 2. *Let $T \in [-\infty, 1], J \in \mathbb{N}$. We define the parametric index sets*

$$I_J^T := \{\mathbf{j} \in \mathbb{N}_0^n : -|\mathbf{j}|_1 + T|\mathbf{j}|_\infty \geq -(n + J - 1) + TJ\} \quad (13)$$

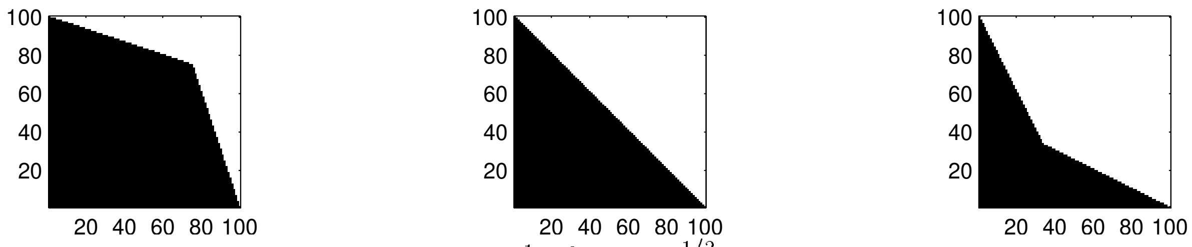

Fig. 1: Index sets I_{100}^{-1}, I_{100}^{0} and $I_{100}^{1/2}$, from left to right.

and the related approximation spaces

$$V_J^T \quad := \quad \bigoplus_{\mathbf{j}\in I_J^T} W_{\mathbf{j}}. \tag{14}$$

The new parameter T in (13), (14) allows us to decouple the definition of the index sets and the resulting grids from the smoothness parameters s and t and to investigate more closely into the relation between smoothness assumptions, the choice of approximation space and the order of approximation.
For $T = 0$ and $T = -\infty$ we regain the sparse and full grid spaces V_J^0 and $V_J^{-\infty}$. Obviously the inclusions $V_J^{T_1} \subset V_J^{T_2}$ for $T_2 \leq T_1$ hold. Figure 1 shows some examples of spaces V_J^T for the two-dimensional case.
For the approximation properties of the spaces V_J^T we have the following Theorem.
THEOREM 2. *Let $-\tilde{r} < s < t + l < r, t + l \geq 0$ and $0 \leq t < r$. Then for $u \in \mathcal{H}_{mix}^{t,l}$ it holds*

$$\inf_{v\in V_J^T} \|u - v\|_{\mathcal{H}^s}^2 \quad \leq \quad \begin{cases} C \cdot 2^{2\left(s-l-t+(Tt-s+l)\frac{n-1}{n-T}\right)J}\|u\|_{\mathcal{H}_{mix}^{t,l}}^2 & \textit{for } T \geq \frac{s-l}{t} \\ C \cdot 2^{2(s-l-t)J}\|u\|_{\mathcal{H}_{mix}^{t,l}}^2 & \textit{for } T \leq \frac{s-l}{t}. \end{cases} \tag{15}$$

Proof: To show (15) we use the norm equivalence (10). Then

$$\begin{aligned} \inf_{v\in V_J^T} \|u - v\|_{\mathcal{H}^s}^2 \quad &\leq \quad \|u - \sum_{\mathbf{j}\in I_J^T} w_{\mathbf{j}}\|_{\mathcal{H}^s}^2 \overset{(9)}{\approx} \sum_{\mathbf{j}\notin I_J^T} 2^{2s|\mathbf{j}|_\infty}\|w_{\mathbf{j}}\|_{\mathcal{L}^2}^2 \\ &\leq \quad \max_{\mathbf{j}\notin I_J^T} 2^{2(s-l)|\mathbf{j}|_\infty - 2t|\mathbf{j}|_1} \cdot \sum_{\mathbf{j}\notin I_J^T} 2^{2l|\mathbf{j}|_\infty + 2t|\mathbf{j}|_1}\|w_{\mathbf{j}}\|_{\mathcal{L}^2}^2 \\ &\overset{(10)}{\leq} \quad C \cdot \max_{\mathbf{j}\notin I_J^T} 2^{2(s-l)|\mathbf{j}|_\infty - 2t|\mathbf{j}|_1}\|u\|_{\mathcal{H}_{mix}^{t,l}}^2. \end{aligned}$$

Evaluating the maximum with respect to I_J^T gives (15).
□

Theorem 2 shows that the optimal order of approximation of a function in $\mathcal{H}_{mix}^{t,l}$ is kept when changing from the full grid approximation space $V_J^{-\infty}$ to an approximation space V_J^T with $T \leq (s-l)/t$. However, the use of approximation spaces V_J^T with $T > (s-l)/t$ leads to a deterioration of the optimal order of convergence. More general choices of smoothness conditions can be handled in the same spirit if norm equivalences for these spaces and associated bilinear forms are known.

Note that for problems with large ellipticity constants it may be advisable to spare the detour via the $\mathcal{H}^s$-norm and to make use of norm estimates applied directly to $a(.,.)$. Then a further adaptation of the approximation space to the operator at hand can be obtained. This is of importance for preconditioning purposes also. Specifically for anisotropic elliptic problems

$$-\sum_{i=1}^{n} d_i \; \frac{\partial^2}{\partial^2 x_i} \; u = f, d_i > 0,$$

tensor-product approximation spaces are well suited since they easily allow for anisotropic refinement. Let $a(.,.)$ denote the corresponding $\mathcal{H}^1$-elliptic variational form. In this case, the condition number of the corresponding Galerkin stiffness matrix for an isotropic full grid depends linearly on $\max_{1\le i\le n}(d_i)/\min_{1\le i\le n}(d_i)$. The same is true for the coefficient in the asymptotic estimate of the approximation error. Hence, for fixed refinement level J, the corresponding convergence rate of iterative methods as well as the approximation error depend on d_i.
A consideration analogously to that in the proof of Theorem 1 shows

$$a(u,u) \approx \sum_{\mathbf{j}} \left(\sum_{i=1}^{n} d_i \, 2^{2j_i} \right) \|w_j\|^2_{\mathcal{L}^2} \approx \sum_{\mathbf{j}} \max_{1\le i\le n} (d_i 2^{2j_i}) \|w_{\mathbf{j}}\|^2_{\mathcal{L}^2}$$

for $u \in \mathcal{H}^1, u = \sum_{\mathbf{j}} w_{\mathbf{j}}$, see also [GO95]. Compared to the norm equivalence (10) (set $s = 1$) the weight $2^{2|\mathbf{j}|_\infty}$ is substituted by $\max_{1\le i\le n}(d_i 2^{2j_i})$. Then, analogously to the above definition, approximation spaces can be constructed using this norm equivalence instead of (10), see [GKn98]. The resulting approximation spaces keep the order of approximation of the full grid approximation space and the resulting preconditioners are robust with respect to the coefficients d_i. In the case of extreme anisotropy, the resulting grid consists of extremely stretched grids in the direction of the anisotropy.

3.2. Dimension of the approximation spaces V_J^T. The following lemma discusses the dimension of the spaces V_J^T.
Lemma 1. *It holds*

$$dim\left(V_J^T\right) \;\; \le \;\; \begin{cases} n \cdot 2^J & \textit{for } T = 1, \\ \frac{n}{2}\left(\frac{1}{1-2^{-\frac{1}{1/T-1}}}\right)^n \cdot 2^J = O(2^J) & \textit{for } T \in (1/J, 1), \\ O(2^J J^{n-1}) & \textit{for } T \in [0, 1/J], \\ O(2^{\frac{T-1}{T/n-1}J}) & \textit{for } T < 0. \end{cases} \tag{16}$$

Proof: We need to estimate $\sum_{\mathbf{j}\in I_J^T} |W_{\mathbf{j}}| \approx \sum_{\mathbf{j}\in I_J^T} 2^{|\mathbf{j}|_1}$. Let $K(J,T) = \sup\{|\mathbf{j}|_1 : \mathbf{j} \in I_J^T\}$. Then we know from [D83], Theorem 6, that

$$\sum_{\mathbf{j}\in I_J^T} 2^{|\mathbf{j}|_1} \approx 2^K J^r$$

with $r(0) = n - 1$ and $r(T) = 0$ for $T \neq 0$. For fixed T, the evaluation of $K(J,T)$ gives the asymptotic estimates of the cases $T = 1$ and $T, 0$ in (16).

For fixed J however, T acts more like a discrete parameter and the result $dim(V_J^0) = O(2^J J^{n-1})$ for $T = 0$ extends to $T \in [0, 1/J]$, compare [GKn98]. Thus, the analysis of the coefficient in the case $1/J < T < 1$ remains.
To this end we let $|\mathbf{j}|_1 = n + J - 1 - i$ and $1/J < T \le 1$, such that $W_{\mathbf{j}} \subset V_J^T \Leftrightarrow -|\mathbf{j}|_1 + T|\mathbf{j}|_\infty \ge -(n+J-1) + TJ \Leftrightarrow |\mathbf{j}|_\infty \ge J - \frac{1}{T}i$. Since $\sum_{|\mathbf{j}|_1 = n+J-1-i} 1 = \binom{|\mathbf{j}|_1 - 1}{n-1}$ and

$$\sum_{\substack{|\mathbf{j}|_1 = n+J-1-i, \\ |\mathbf{j}|_\infty = \ell_1 \ge J - i/T}} 1 \le \binom{\lceil |\mathbf{j}|_1 - (J - i/T) \rceil}{n-1} = \binom{\lceil n - 1 + (1/T - 1)i \rceil}{n-1}, \tag{17}$$

the number of subspaces $W_{\mathbf{j}} \subset V_J^T$ with $|\mathbf{j}|_1 = n + J - 1 - i$ is bounded by $n\binom{\lceil n-1+(\frac{1}{T}-1)i \rceil}{n-1}$. Hence, with the definition of V_J^T,

$$|V_J^T| \;=\; \sum_{i=0}^{J-1} \sum_{\substack{|\mathbf{j}|_1 = n+J-1-i, \\ |\mathbf{j}|_\infty = \ell_1 \ge J - i/T}} |W_j| \le 2^{J-1} n \sum_{i=0}^{J-1} 2^{-i} \binom{\lceil n - 1 + (1/T - 1)i \rceil}{n-1}. \tag{18}$$

In the case $T - 1 = 0$ we get $|V_J^T| \le 2^{J-1} n \sum_{i=0}^{J-1} 2^{-i} \le 2^J n$. For $T < 1$ the substitution $i \to \frac{i}{1/T-1}$ leads to

$$|V_J^T| \le 2^{J-1} n \sum_{i=0}^{\lceil (1/T-1)(J-1) \rceil} 2^{-\frac{i}{1/T-1}} \binom{n-1+i}{i}.$$

Since $(x^{n-1+i})^{(n-1)} = \frac{(n-1+i)!}{i!} x^i \;\; \forall x \in \mathbb{R}$, we obtain

$$\begin{aligned}
|V_J^T| &\le 2^{J-1} n \, \frac{1}{(n-1)!} \sum_{i=0}^{\lceil (1/T-1)(J-1) \rceil} (x^{n-1+i})^{(n-1)} \Big|_{x = 2^{-\frac{1}{1/T-1}}} \\
&= 2^{J-1} n \frac{1}{(n-1)!} \left[\left(\frac{x^{n-1}}{1-x} \right)^{(n-1)} - \left(\frac{x^{\lceil (1/T-1)(J-1) \rceil + n}}{1-x} \right)^{(n-1)} \right] \Big|_{x = 2^{-\frac{1}{1/T-1}}} \\
&\le 2^{J-1} n \left(\frac{1}{1-x} \right)^n \Big|_{x = 2^{-\frac{1}{1/T-1}}} = 2^{J-1} n \left(\frac{1}{1 - 2^{-\frac{1}{1/T-1}}} \right)^n.
\end{aligned}$$

This completes the proof.
□

Compare the estimate in (16) to the results for the full grid spaces $V_J^{-\infty}$. For $T > 1/J$ the dependence of the dimension of the approximation space on the dimensionality n of the problem has been reduced from $O(2^{nJ})$ to $O(2^J)$. Note however, that the constants do still dependent on n, compare (16).
In summary, Theorem 2 and Lemma 1 show that for approximation problems with $u \in \mathcal{H}_{mix}^{t,l}$, the use of the approximation spaces V_J^T with $T \le (s-l)/t$ leads to a significant reduction of the number of degrees of freedom compared to the full grid, while the order of approximation is preserved. Specifically in the case $T \in (1/J, (s-l)/t)$, the dependence of the (asymptotic) dimension of V_J^T on n is completely removed, while the order of approximation is the same as in the full grid case.

4. Cost of resulting finite element algorithms. Now we consider the cost of solving the elliptic variational problem (1) up to some prescribed error in the worst case setting (i.e. the error of an approximation is measured in the $\mathcal{H}^s$-norm) [Wer91]. Due to the larger supports of the wavelets in the approximation spaces V_J^T from coarser scales, the resulting stiffness matrices $\mathcal{A}_J$ are rather densely populated.

However, usually we need not to assemble the stiffness matrix. Instead, all that is required in an iterative scheme is the application of the preconditioned stiffness matrix to a vector. The pyramid structure of the multiscale transformations and the tensor-product structure of our wavelet basis functions then often allow the matrix-vector product to be performed with $O(dim(V_{FE}))$ operations. This can directly be seen for the Laplace operator and for differential operators with constant or separable coefficients. However, in the case of equations with more general coefficient functions and general V_J^T-spaces or adaptively refined grids, the implementation of the matrix-vector product with linear complexity is a very involved and delicate task.

We assume that comparison of real numbers as well as standard arithmetic operations are performed with unit cost and that the stiffness matrix as well as the load vector have been computed with sufficient accuracy.

If the norm equivalence (9) is valid, simple diagonal scaling of the stiffness matrix leads to optimal preconditioning. Hence, in each iteration step, the error is reduced at least by a factor $\rho < 1$ which is independent of the problem size $dim(V_J^T)$. Let e^{it} denote the error after the i-th iteration step. Then we get $\|e^{it}\| \leq \rho^{it}\|e^0\|$. Therefore the number of iteration steps that are necessary to compute an approximation with $\|e^{it}\| \leq O(\epsilon)$ is of the order $O(|\ln \epsilon|)$. Now, because it does not make sense to iterate beyond the discretization error, we let $\epsilon = O(2^{-\alpha J})$ for some $\alpha > 0$. Hence, the cost of computing a solution of the variational problem (1) within accuracy ϵ is $O(J \cdot dim(V_{FE})^2)$ if the stiffness matrix is dense and it is $O(J \cdot dim(V_{FE}))$ if the matrix-vector product can be performed with $O(dim(V_{FE}))$ operations. We note that it is usually possible to remove the J-term in the complexity estimate by embedding the solver in a nested iteration scheme [KD71]. Then, if the solution is in $\mathcal{H}_{mix}^{t,l}$ and the matrix-vector product can be performed with $O(dim(V_{FE}))$ operations, the cost of solving an $\mathcal{H}^s$-elliptic variational problem with the approximation space $V_J^{(s-l)/t}$ is

$s > l$	$s = l$	$s < l$
$O(\epsilon^{\frac{1}{s-l-t}})$	$O\left(\epsilon^{-\frac{1}{t}}\left(\ln(\epsilon^{-\frac{1}{t}})\right)^{n-1}\right)$	$O(\epsilon^{\frac{1}{(s-l)/n-t}})$

These results should be compared to the $O(\epsilon^{\frac{n}{s-l-t}})$ cost for the full grid case. If the stiffness matrix is dense, i.e. it has $O(dim(V_{FE})^2)$ entries (as is the case for example for integral operators), these results have to be squared. In this case, additional compression, along the lines of [DPS93, DJP92] could lead to a further significant reduction of the cost.

Note however, that the constants in these asymptotic estimates in general still depend exponentially on the dimension n. This may be seen from the the constants in the estimate of the dimension of the spaces V_J^T, compare Lemma 1. Also Proposition 1 indicates that the condition number of the space decompositions and therefore of the stiffness matrices is exponentially dependent on n.

5. Concluding remarks. The constructions of the approximation spaces presented in this paper are not restricted to biorthogonal wavelets as basis functions but can be carried over to other multiscale basis functions. Specifically, the above construction of optimized grids is not limited to stable multilevel splittings. Only certain decay properties of the coefficients in the multiscale representation are needed. Then, applying the triangle inequality together with a Bernstein and a direct estimate yields corresponding results. This may also be used as the starting point for enlarging the range of the validity of the estimates presented in this paper. Especially the upper range of the parameters t and l can be enlarged. Apart from eventual logarithmic terms in the cases $\frac{s-l}{t} \leq 0$, the results remain the same.

To be able to repeat the above reasoning for function spaces fulfilling other boundary conditions (which are consistent with the tensor-product approach) appropriate multiscale basis functions have to be constructed. Then the argumentation can be repeated with obvious modifications.

The approximation spaces V_J^T can also be obtained via the use of an optimization procedure. From

$$a(u,u) \overset{(2)}{\approx} \|u\|_{\mathcal{H}^s}^2 \overset{(9)}{\approx} \sum_{\mathbf{j}} 2^{2s|\mathbf{j}|_\infty} \|w_{\mathbf{j}}\|_{\mathcal{L}^2}^2 \overset{(7)}{\approx} \sum_{\mathbf{j}} 2^{2s|\mathbf{j}|_\infty} \left(\sum_{m \in \tau_{\mathbf{j}}} \langle u, \tilde{\psi}_{\mathbf{jm}} \rangle^2 \right) \text{ for } s \in (-\tilde{r}, r), \qquad (19)$$

we conclude that $\gamma_{\mathbf{j}} := 2^{2s|\mathbf{j}|_\infty} \|w_{\mathbf{j}}\|_{\mathcal{L}^2}^2 / |W_{\mathbf{j}}|$ is a measure of the benefit of including $W_{\mathbf{j}}$ into the approximation space. Here $|W_{\mathbf{j}}|$ is a measure for the cost and the numerator is a measure for the profit of including $W_{\mathbf{j}}$. Therefore, given some prescribed overall cost, the problem to determine which subspaces should be included into the approximation space can be reformulated as a classical knapsack problem [MT90]. Moreover, following [BG98], the optimization process can be reduced to the discussion of these profit/cost quotients. For an optimal grid in this sense one has to include those $W_{\mathbf{j}}$ into the approximation space which have $\gamma_{\mathbf{j}}$ greater than a specified threshold. A short consideration shows that the moment condition on the dual wavelets together with the smoothness assumptions leads to the same optimized grids [GKn98].

Note finally that generalizations to more general geometries are not easy to handle. Research in this direction is mainly based either on domain transformation techniques [BD97] or on some kind of domain decomposition approach [Sch99]. However, we note that $[0,1]^n$ is the natural computational domain for many higher-dimensional physical applications anyway.

REFERENCES

[Ad75] R.A. Adams, *Sobolev spaces*, Academic Press, New York, 1975.

[Bab60] K. I. Babenko, *Approximation by trigonometric polynomials in a certain class of periodic functions of several variables*, Soviet Math. Dokl. 1 (1960), 672–675.

[BDJ92] G. Baszenski, F.-J. Delvos, S. Jester, *Blending approximation with sine functions*, in Numerical Methods of Approximation 9, D. Braess and L. L. Schumaker, eds., Int. Ser. Num. Math., 105, 1992, 1–19.

[Be93] G. Beylkin, *Wavelets and fast numerical algorithms*, in: Proceedings of Symposia in Applied Mathematics, 47, 1993.

[Bu92] H.-J. Bungartz, *Dünne Gitter und deren Anwendung bei der adaptiven Lösung der dreidimensionalen Poisson-Gleichung*, Dissertation, TU München, Institut für Informatik, 1992.

[BG98] H.-J. Bungartz, M. Griebel, *A note on the complexity of solving Poisson's equation for spaces of bounded mixed derivatives*, J. Complexity 15 (1999), 167-199.

[BD97] H.-J. Bungartz, T. Dornseifer, *Sparse grids: Recent developments for elliptic partial differential equations*, Report TUM-I9702, Institut für Informatik, TU München, 1997.

[D96] W. Dahmen, *Stability of Multiscale Transformations*, Journal of Fourier Analysis and Applications 2 (1996), 341-361.

[D97] W. Dahmen, *Wavelet and Multiscale Methods for Operator equations*, Acta Numerica (1997), 55-228.

[DK92] W. Dahmen, A. Kunoth, *Multilevel preconditioning*, Numer. Math., 63 (1992), 315–344.

[DPS93] 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.

[Del82] F.-J. Delvos, *d-Variate Boolean interpolation*, J. Approx. Theory, 34 (1982), 99–114.

[DJP92] R.A. DeVore, B. Jawerth, V. Popov, *Compression of Wavelet decompositions*, Amer. J. Math. 114 (1992), 737-785.

[DKT98] R.A. DeVore, S.V. Konyagin, V.N. Temlyakov, *Hyperbolic Wavelet Approximation*, Constructive Approximation 14 (1998), 1-26.

[D83] Din' Zung, *The approximation of classes of periodic functions of many variables.* Russian Math. Surveys 38 (1983), 117-118.

[D84] Din' Zung, *Number of integral points in a certain set and the approximation of functions of several variables*, Math. Notes (1984), 736-744.

[D88] Din' Zung, *Approximation by trigonometric polynomials of functions of several variables on the torus*, Math. USSR Sbornik 59 (1988), 247-267.

[FHP96] K. Frank, S. Heinrich, S. Pereverzev, *Information complexity of multivariate Fredholm equations in sobolev classes*, J. Complexity 12 (1996), 17-34.

[G78] E.M. Galeev, *Approximation by fourier sums of functions with several bounded derivatives*, Math. Notes 23 (1978), 109-117.

[Go69] W. J. Gordon, *Distributive lattices and the approximation of multivariate functions*, in Approximation with Special Emphasis on Spline Functions, I. J. Schoenberg, ed., Academic Press, New York, 1969, 223–277.

[Gr91] M. Griebel, *A parallelizable and vectorizable multi-level algorithm on sparse grids*, Parallel Algorithms for Partial Differential Equations, Notes on Num. Fluid Mech. 31, W. Hackbusch (ed.), Vieweg, Braunschweig, 1991.

[Gri94a] M. Griebel, *Multilevelmethoden als Iterationsverfahren über Erzeugendensystemen*, Teubner Skripten zur Numerik, Teubner, Stuttgart, 1994.

[GKn98] M. Griebel, S. Knapek, *Optimized approximation spaces for operator equations*, SFB 256, report Nr. 568, Universität Bonn, 1998.

[GKo99] M. Griebel, F. Koster, *The adaptive solution of PDEs with interpolating wavelets*, SFB 256, report, Universität Bonn, 1999.

[GO95] M. Griebel, P. Oswald, *Tensor product type subspace splittings and multilevel iterative methods for anisotropic problems*, Advances in Computational Mathematics 4 (1995), 171-206.

[GOS99] M. Griebel, P. Oswald, T. Schiekofer, *Sparse grids for boundary integral equations*, Numerische Math. 83(2) (1999), 279-312.

[GSZ92] M. Griebel, M. Schneider, C. Zenger, *A combination technique for the solution of sparse grid problems*, Iterative Methods in Linear Algebra, P. de Groen and R. Beauwens (eds.), Elsevier, Amsterdam, 1992, 263-281.

[Ho99] R. Hochmuth, *Wavelet Bases in Numerical Analysis and Restricted Nonlinear Approximation*, Habilitation, Freie Universität Berlin, 1999.

[J92] S. Jaffard, *Wavelet methods for fast resolution of elliptic equations*, SIAM J. Numer. Anal. 29 (1992), 965-986.

[Ka96] A. Kamont, *On Hyperbolic Summation and Hyperbolic Moduli of Smoothness*, Constructive Approximation 12 (1996), 111-125.

[KD71] L. Kronsjö, G. Dahlquist, *On the design of nested iteration for elliptic differential equations*, BIT 11 (1971), 63-71.

[MT90] S. Martello, P. Toth, *Knapsack Problems: Algorithms and Computer Implementations*, John Wiley & Sons, Chichester, 1990.

[O92] 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.

[O98] P. Oswald, *Best N-term approximation of singularity functions in two Haar bases*, Bell Labs, Lucent Technologies, 1998, preprint.

[Pe88] 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.

[Pe89] S.V. Pereverzev, *On the complexity of finding solutions of Fredholm equations of the second kind with differentiable kernels, II*, Ukrain. Mat. Zh. 41 (1989), 169-173.

[Pf98] C. Pflaum, *A multilevel algorithm for the solution of second order differential equations on sparse grids*, Numer. Math. 79 (1998), 141-155.

[ST87] H.-J. Schmeisser, H. Triebel, *Fourier Analysis and Functions spaces*, John Wiley, Chichester, 1987.

[Sc98] R. Schneider, *Multiskalen- und Wavelet-Kompression: Analysisbasierte Methoden zur effizienten Lösung großer vollbesetzter Gleichungssysteme*, Advances in Numerical Analysis, Teubner, Stuttgart, 1998.

[Sch99] T. Schiekofer, *Die Methode der Finiten Differenzen auf dünnen Gittern zur Lösung elliptischer und parabolischer partieller Differentialgleichungen*, Dissertation, Universität Bonn, 1999.

[SS98] W. Sickel, F. Sprengel, *Interpolation on sparse grids and tensor products of Nikol'skij-Besov spaces*, J. Computational Analysis and Applications 1 (1999), 261-286.

[Sm63] S.A. Smolyak, *Quadrature and interpolation formulas for tensor products of certain classes of functions*, Dokl. Akad. Nauk SSSR 4 (1963), 240–243.

[Te93] V. N. Temlyakov, *On approximate recovery of functions with bounded mixed derivative*, J. Complexity, 9 (1993), 41–59.

[Te89] V. N. Temlyakov, *Approximation of functions with bounded mixed derivative*, Proceedings of the Steklov Institute of Mathematics, 1 (1989).

[TWW88] J. F. Traub, G. W. Wasilkowski, H. Wozniakowski, *Information-Based Complexity*, Academic Press, New York, 1988.

[WW95] G. W. Wasilkowski, H. Wozniakowski, *Explicit cost bounds of algorithms for multivariate tensor products*, J. Complexity, 11 (1995), 1–56.

[Wei80] J. Weidmann, *Linear operators in Hilbert spaces*, Springer, New York, 1980.

[Wer91] A. G. Werschulz, *The computational complexity of differential and integral equations, An information based approach*, Oxford University Press, New York, 1991.

[Wer95] A. G. Werschulz, *The computational complexity of the Poisson problem for spaces of bounded mixed derivative*, Technical Report CUCS-016-95, 1995.

[Z91] C. Zenger, *Sparse grids*, Parallel Algorithms for Partial Differential Equations, Notes on Num. Fluid Mech. 31, W. Hackbusch (ed.), Vieweg, Braunschweig, 1991.