

# Optimized approximation spaces for operator equations

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 pseudodifferential operators of arbitrary order. Based on the framework of tensor-product biorthogonal wavelet bases and stable subspace splittings, we construct operator adapted finite element subspaces with less dimension than the standard full grid spaces that keep the approximation order of the standard full grid spaces provided that certain additional regularity assumptions on the solution are fulfilled. Specifically for operators of positive order, the dimension is  $O(2^J)$  independent of the dimension  $n$  of the problem compared to  $O(2^{Jn})$  for the full grid space. Also, for operators of negative order the overall complexity is significantly in favor of the new approximation spaces. We give complexity estimates for the case of continuous linear information. We show these results in a constructive manner by proposing a finite element method together with optimal preconditioning. The theory covers elliptic boundary value problems as well as boundary integral equations.

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

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

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Motivation and Assumptions</b>	<b>4</b>
<b>3</b>	<b>Norm equivalences</b>	<b>10</b>
<b>4</b>	<b>Optimized approximation spaces for Sobolev spaces</b>	<b>14</b>
4.1	Approximation spaces for problems with constraint on the solution . . . . .	14
4.1.1	Estimates on the order of approximation for the spaces $V_J^{-\infty}$ and $V_J^0$ . . .	14
4.1.2	Definition and order of approximation of the approximation spaces $V_J^T$ . .	16
4.2	Dimension of the approximation spaces $V_J^T$ . . . . .	20
4.3	Optimization procedures and subspace selection . . . . .	24
<b>5</b>	<b>Complexity estimates</b>	<b>27</b>
<b>6</b>	<b>Applications</b>	<b>29</b>
<b>7</b>	<b>Concluding remarks</b>	<b>33</b>

# 1 Introduction

This paper is concerned with the construction of finite element spaces for the approximate solution of elliptic problems in Sobolev spaces  $\mathcal{H}^s(\Omega)$ ,  $s \in \mathbb{R}$ . This introductory section is a summary of relevant background information on existing methods for the construction of optimized grids for discretization purposes.

It is well-known that the  $\epsilon$ -complexity [TWW88, We91] of solving Poissons equation in  $n$  dimensions and in the Sobolev space  $\mathcal{H}_0^1$  on a bounded domain is  $\epsilon^{-n}$ , i.e., it is exponentially dependent on  $n$ . Hence for higher-dimensional problems a direct numerical solution on a regular standard triangulation is prohibitive. To overcome the exponential dependence on  $n$  the idea of reducing the dimension of the solution space considering only functions of higher regularity has been presented in [Z91]. Using a hierarchical subspace decomposition based on tensor-products of one-dimensional hierarchical bases of piecewise linear splines and the space of functions with bounded mixed second derivative as solution space, subspaces with relatively large dimensions that contribute only “little” to the error reduction can be identified and omitted from the approximation space. This idea has been well known for some time in approximation and interpolation theory [Bab60, BDJ92, Del82, Go69, Sm63, Te93, Te89, WW95] and attracted much attention in last time [Ka96, DKT98, Si98, SS98]. Recently (in connection with linear splines as basis functions) in [BG98] an approximation space for the discretization of Poissons equation with dimension of the order  $O(2^J)$  has been constructed. Here,  $J$  is the number of levels involved in the multilevel discretization scheme. It has been shown that no loss in approximation power occurs compared to the full grid, required the error is measured in the energy norm, i.e., the  $\mathcal{H}^1$ -norm. If measured in the  $\mathcal{H}^0$ -norm, the order of approximation deteriorates from  $O(2^{-2J})$  for the full grid to  $O(2^{-2J} J^{n-1})$ .

Note that this construction of approximation spaces (i.e., the selection of subspaces) hinges on the additional freedom that is provided by the tensor-product ansatz and on additional regularity requirements. The resulting sequence of grid points and supports has also been used successfully in connection with different hierarchical basis functions, for example prewavelets and higher order splines [GO95, Bu96, BD97, Bu98]. Recently, [GOS98] used the sparse grid approach together with prewavelets for the solution of boundary integral equations of order between  $-1/2$  and  $1/2$ . Note that the application of the sparse grid approach to boundary integral equations is particularly tempting, as the matrix compression step, ensuring optimal work count, might be avoided. Specifically the authors dealt with the single layer potential equation. They noticed, that for operators of negative order, there appears a loss of approximation power due to the fact, that the embedding  $\mathcal{H}_{mix}^s \subset \mathcal{H}^s$  of the Sobolev space  $\mathcal{H}_{mix}^s$  with bounded mixed  $s$ -th derivative in the isotropic Sobolev space  $\mathcal{H}^s$  for positive  $s$  is reversed in the negative case.

In this paper we extend these ideas to pseudodifferential operators of arbitrary order and to approximation spaces spanned by biorthogonal wavelet systems. The sparse grids presented in the literature thus far are based on  $\mathcal{L}^p$ -norms or on the  $\mathcal{H}^1$ -norm, i.e., on the Laplace operator. We will base the construction of approximation spaces on the  $\mathcal{H}^s$ -norm,  $s \in \mathbb{R}$ , and will thus extend these ideas to the whole scale of Sobolev spaces. This enables us to define optimized approximation spaces for elliptic operators acting on arbitrary  $\mathcal{H}^s$  spaces. In the construction procedure of the approximation spaces there is a need for the decoupling of the subspaces arising from the tensor-product ansatz and semicoarsing in coordinate directions. In [Bu92, BG98] this has been done by simply applying the triangle inequality. In this paper we rely directly on norm estimates and norm equivalences that allow us to decouple the subspaces and moreover ensure the stability of the resulting subspace splittings. Hence norm equivalences are not only important for preconditioning, but can also be used as a source of information for discretization

and subspace selection. This is of particular interest for operators with large ellipticity constants, that lead to large constants in the estimates of the approximation error. Hence the constants in the asymptotic estimates may dominate the error approximation for practical problem sizes.

In [GOS98] 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 phenomenon has been studied theoretically in a simplified setting via  $N$ -term best approximation. It turned out that a selection of a  $N$ -term linear combination of tensor-product Haar basis leads to extremely good approximation rates. The essential idea is to split the global solution  $u$  into a regular part  $u^{reg}$  and a singular part  $u^{sing}$ , that is

$$u = u^{reg} + u^{sing},$$

and to approximate the regular part on a rather sparse grid. The singular part has to be treated via additional grid points where it is necessary. The idea is that a few wavelets of high level clustered around the singularity will suffice, while relatively sparse grids are enough to treat the smooth parts of the solution. The singular part has to be treated either by adaptive refinement techniques using a-posteriori local error estimators or indicators or (if possible) by adapting the approximation space to the solution by using a-priori available information about for example edge or corner singularities of the solution. Adaptive refinement techniques using a-posteriori local error estimators or indicators have been studied especially with respect to finite element discretizations of partial differential equations [BM87, BR78, BW85, V94] and integral equations [St96, Dahl96] and have proven to be very powerful especially for problems with low global Sobolev smoothness due to edge or corner singularities or singular perturbed operators or not sufficiently smooth right hand sides. We refer to the recent papers [Dahm97, DDD97, DeV98] where excellent results were obtained using  $N$ -term best approximation. However, these techniques pose huge technical problems connected with mesh refinement strategies especially in higher dimensions. Therefore a combination of adaptivity to capture the non-smooth parts of the solution and a relatively sparse grid for the smooth parts seems to be a very promising approach, compare [GOS98, O98]. In this paper we focus on finding approximation spaces for problems with regular solution (or for the smooth part of the solution) and on determining the complexity of obtaining such a solution.

The remainder of this paper is as follows. Section 2 introduces some notation, collects basic facts about biorthogonal wavelet bases and tensor-product spaces and describes the motivations for the construction of optimized grids. Section 3 contains some theory about norm equivalences in Sobolev spaces. In Section 4 we make use of the results of the two preceding sections. There the optimized spaces are defined and estimates on the dimension of the optimized spaces and their order of approximation are given. Section 5 contains remarks on the complexity of solving elliptic equations for the case that continuous linear information is permissible, that is, stiffness matrix as well as load vector can be computed exactly. 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. Section 6 discusses applications to two elliptic problems, the Poisson problem and the single layer potential equation. At the end of the paper we indicate further modifications of the construction of the optimized grids and summarize the main conclusions and open problems. Specifically we derive modifications of the optimized spaces by incorporating additional information from the operators considered. This leads to the definition of for example anisotropic sparse grids [R91]. Furthermore we discuss the potential possibilities of incorporating a-priori known information about singularities of the solution into the construction process of optimized grids.

## 2 Motivation and Assumptions

Let us denote by  $\mathcal{H}^t(T^n)$ ,  $t \in \mathbb{R}$ , a scale of Sobolev spaces on the  $n$ -dimensional torus, and  $\mathcal{L}^2(T^n)$  the space of  $\mathcal{L}^2$ -integrable functions on  $T^n$ , see [Ad75]. For ease of presentation and in order to avoid restrictions on the smoothness exponent  $t$  we restrict ourselves to the  $n$ -dimensional torus in the first parts of this paper. Applications to non-periodic problems will be discussed in section 6. We represent  $T^n$  by the  $n$ -dimensional cube

$$T := [0, 1], \quad T^n = T \times T \times \cdots \times T$$

where opposite faces are identified. When  $t < 0$ ,  $\mathcal{H}^t(T^n)$  is defined as the dual of  $\mathcal{H}^{-t}(T^n)$ , i.e.,

$$\mathcal{H}^t(T^n) := (\mathcal{H}^{-t}(T^n))'. \quad (1)$$

When 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.

Consider an elliptic variational problem: Given  $f \in \mathcal{H}^{-s}$ , find  $u \in \mathcal{H}^s$  such that

$$a(u, v) = (f, v) \quad \forall v \in \mathcal{H}^s, \quad (2)$$

where  $a$  is a symmetric positive definite form satisfying

$$a(v, v) \approx \|v\|_{\mathcal{H}^s}^2. \quad (3)$$

Here  $x \approx y$  means that there exist  $C_1, C_2$  independent of any parameters  $x$  or  $y$  may depend on, such that

$$C_1 \cdot y \leq x \leq C_2 \cdot y.$$

(Clearly, the lower estimate  $a(u, u) \geq \alpha \cdot \|u\|_{\mathcal{H}^s}^2$  in (3) is in general not fulfilled for problems on the torus without additional constraints ensuring uniqueness of the solution of (2). In the following we will assume that the solution of the variational problem (2) is unique. Note however that for the construction of optimized grids, we will only need the upper estimate in (3).)

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, but does not depend on the number of levels  $J$ . In the following multi-indices (vectors) are written boldface, for example  $\mathbf{j}$  for  $(j_1, \dots, j_n)$ . Inequalities like  $\mathbf{l} \leq \mathbf{t}$  or  $\mathbf{l} \leq 0$  are to be understood componentwise.

Model examples for (2) would be the variational form of

- the biharmonic equation ( $s = 2$ )

$$\Delta^2 u = f,$$

which has applications in plate bending and shell problems,

- the (anisotropic) Helmholtz equation ( $s = 1$ )

$$-\nabla \cdot K \nabla u + cu = f \quad \text{on } T^n, \quad (4)$$

where  $K(x) \approx I$  and  $\exists C > 0 : 0 \leq c(x) \leq C$ , modelling for example the single phase flow in a porous medium with permeability  $K$ , or a diffusion process in a (possibly) anisotropic medium characterized by the diffusion tensor  $K$ ,

- the hypersingular equation ( $s = \frac{1}{2}$ )

$$\frac{1}{c} \int_{T^n} \frac{\partial}{\partial n_x} \frac{\partial}{\partial n_y} \left( \frac{1}{|x - y|} \right) \cdot g(y) dy = f(x),$$

- Fredholm equations of the second kind ( $s = 0$ )

$$g(x) - \int_{T^n} k(x, y) g(y) dy = f(x),$$

with given kernel function  $k$  defined on  $T^n \times T^n$ , specifically the double layer potential equation

$$g(x) - \frac{1}{c} \int_{T^n} \frac{n_y \cdot (y - x)}{|x - y|^3} g(y) dy = f(x),$$

arising from a reformulation of Laplace's equation via the indirect method,

- the single layer potential equation ( $s = -\frac{1}{2}$ )

$$\frac{1}{c} \int_{T^n} \frac{g(y)}{|x - y|} dy = f(x). \quad (5)$$

The Galerkin method to solving problem (2) numerically is to select a finite dimensional subspace from  $\mathcal{H}^s \cap \mathcal{L}^2$  and to solve the variational problem in this subspace instead of  $\mathcal{H}^s$ . It is well-known that the most efficient way of solving such problems exploits the interaction of several scales of discretization. These multilevel schemes use a sequence of closed nested subspaces  $S_0 \subset S_1 \subset \dots \subset \mathcal{H}^s \cap \mathcal{L}^2$  of the basic Hilbert space  $\mathcal{H}^s$ , whose union is dense in  $\mathcal{H}^s$ . Fixing a basis of  $S_J$  finally leads to a linear system of equations

$$A_J x_J = b_J \quad (6)$$

of dimension  $\dim(S_J)$ . Here  $A_J$  is called stiffness matrix and  $b_J$  is the load vector. Storage requirements and computation time mostly exclude the use of direct solvers, since  $\dim(S_J)$  is usually very large. Specifically for full grid spaces with subdivision rate two it holds  $\dim(S_J) = O(2^{J \cdot n})$ . That is, the dimension of  $S_J$  grows exponentially with the dimension  $n$ .

In order to iteratively solve (2) or (6), respectively, the following problems and questions arise. Accuracy requirements necessitate a fine partitioning of  $T^n$ , i.e.,  $\dim(S_J)$  is large. Is it possible to select  $S_J$  as a subspace of the full grid space with  $\dim(S_J)$  only polynomially dependent on the dimension  $n$ , compared to an exponential dependence on  $n$  of the dimension of the full grid space? Such a choice of a finite element space would require that one can identify those basis functions that add most to an accurate representation of the solution of the variational problem. For differential operators, the resulting linear systems are sparse if the basis functions have local support. However, the discretization of integral operators results in most cases in discrete systems that are dense. I.e., on a regular full grid, with  $O(2^{nJ})$  unknowns the discrete operator has  $O(2^{2nJ})$  entries. This makes matrix-vector multiplications, as they are needed in iterative methods, prohibitively expensive for large  $n$  or large  $d$  and enforces the use of bases that result in nearly sparse matrices, e.g. biorthogonal wavelet bases with a sufficient number of vanishing moments. Then, most entries in these matrices are close to zero and can be replaced by zero without destroying the order of approximation (compression) [DJP92, DPS93].

Let us recall the definition of the tensor-product of two separable Hilbert-spaces  $H$  with associated bilinear form  $a(., .)$  and  $\hat{H}$  with bilinear form  $\hat{a}(., .)$ , see for example [Wei80]. Let  $\{e_j\}_{j=1}^m$ ,  $\{\hat{e}_i\}_{i=1}^{\hat{m}}$  be complete orthonormal systems in  $H$  and  $\hat{H}$ . Then  $\{e_j \otimes \hat{e}_i\}$  is a complete orthonormal system in

$$H \otimes \hat{H} := \left\{ \sum_{j,i} \gamma_{ij} e_j \otimes \hat{e}_i : \sum_{j,i} \gamma_{ij}^2 < \infty \right\} \quad (7)$$

with scalar product

$$a \otimes \hat{a} \left( \sum_{j,i} \gamma_{ij} e_j \otimes \hat{e}_i, \sum_{k,\ell} \gamma'_{k\ell} e_k \otimes \hat{e}_\ell \right) = \sum_{j,i} \gamma_{ji} \gamma'_{ji}.$$

We identify the tensor-product  $H \otimes \hat{H}$  with a function space over the corresponding product domain via the mapping

$$f \otimes \hat{f} \mapsto f(x) \hat{f}(\hat{x}).$$

E.g., a basis in  $H \otimes \hat{H}$  is given by  $\{\psi_j(x) = e_{j_1}(x_1) \hat{e}_{j_2}(x_2) : 1 \leq j \leq (m, \hat{m})\}$ . These definitions extend naturally to higher dimensions  $n > 2$ .

The finite element spaces considered here are tensor-products of univariate function spaces. Starting from a one-dimensional splitting  $\mathcal{L}^2 = \bigoplus_{j \geq 0} S_j$  we assume that the complement spaces

$$W_j = S_j \ominus S_{j-1} \tag{8}$$

of  $S_{j-1}$  in  $S_j$  are spanned by some  $\mathcal{L}^2$ -stable bases

$$W_j = \text{span}\{\psi_{jk}, k \in \tau_j\}, \tag{9}$$

where  $\tau_j$  is some finite dimensional index set defined from the subdivision rate of successive refinement levels. Here we stick to dyadic refinement. Furthermore we assume that

$$\left\| \sum_{k \in \tau_j} C_k \psi_{jk} \right\|_{\mathcal{L}^2} \approx \|\{C_k\}_k\|_{\ell^2(\tau_j)} \tag{10}$$

where as usual  $\|\sum_{k \in \tau_j} C_k \psi_{jk}\|_{\mathcal{L}^2}$  denotes the norm induced from the scalar product on  $\mathcal{L}^2$  and  $\|\{C_k\}_k\|_{\ell^2(\tau_j)}^2 = \sum_{k \in \tau_j} |C_k|^2$ .

Let there be given a biorthogonal system  $\cup\{\tilde{\psi}_{jk}, k \in \tau_j, j \in \mathbb{N}_0\}$ , i.e.,

$$\langle \psi_{jk}, \tilde{\psi}_{j'k'} \rangle = \delta_{jj'} \delta_{kk'}, \quad j, j' \in \mathbb{N}_0, k \in \tau_j, k' \in \tau_{j'}. \tag{11}$$

Assuming that  $\cup\{\psi_{jk}, k \in \tau_j, j \in \mathbb{N}_0, \}$  forms a Riesz-basis in  $\mathcal{L}^2$ , i.e.,

$$\left\| \sum_{j,k \in \tau_j} C_{jk} \psi_{jk} \right\|_{\mathcal{L}^2} \approx \|\{C_{jk}\}_{jk}\|_{\ell^2(j \in \mathbb{N}, k \in \tau_j)}, \tag{12}$$

every  $u \in H$  has a unique expansion

$$u = \sum_{j=0}^{\infty} \sum_{k \in \tau_j} \langle u, \tilde{\psi}_{jk} \rangle \psi_{jk} = \sum_{j=0}^{\infty} \sum_{k \in \tau_j} \langle u, \psi_{jk} \rangle \tilde{\psi}_{jk} \tag{13}$$

and the biorthogonal system also forms a Riesz-basis in  $\mathcal{L}^2$ .

Let us recall the notion of vanishing moments. In one dimension  $\psi_{jk}$  and  $\tilde{\psi}_{jk}$  are said to have vanishing moments of order  $N, \tilde{N}$  respectively, if

$$\int_{\mathbb{R}} x^r \psi_{jk}(x) dx = 0, r = 0, \dots, N-1, \quad \int_{\mathbb{R}} x^r \tilde{\psi}_{jk}(x) dx = 0, r = 0, \dots, \tilde{N}-1. \tag{14}$$

Note that due to the biorthogonality of the basis functions (i.e. due to (13)) the number of vanishing moments  $N$  of the biorthogonal basis  $\{\tilde{\psi}_{jk}\}$  is exactly the order of polynomial reproduction of the wavelet basis  $\{\psi_{jk}\}$  and vice versa. It is well known [DPS93, Sc98] that

the number of vanishing moments governs the compression capacity of a wavelet and that the order of polynomial reproduction governs the approximation power. Estimates of the order of approximation are mainly based on the local  $\mathcal{L}^2$ -stability (3), and an inequality of Jackson type which in turn depends on estimates of the coefficients  $\langle u, \psi_{jk} \rangle$ , i.e. on a moment condition for the dual wavelet. 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  $\psi_{jk}$  [DPS94, DJP92].

One of the merits of biorthogonal wavelets is that the number of vanishing moments can be chosen independently of the order of polynomial exactness. We will see later on that it is the number of vanishing moments of the dual wavelets  $\tilde{\psi}_{jk}$  that governs the form of the resulting optimized grids, if we pose specific assumptions on the solution of the variational problem.

Let

$$S = \cup_{i=0}^{\infty} S_i \text{ and } \tilde{S} = \cup_{i=0}^{\infty} \tilde{S}_i$$

with  $\tilde{S}_i := \cup_{j=0}^i \{\tilde{\psi}_{jk}, k \in \tau_j\}$ . Moreover, we assume that the  $\psi_{jk}$  and  $\tilde{\psi}_{jk}$  are scaled and delated versions of single scale functions (mother wavelets)  $\psi_0$  and  $\tilde{\psi}_0$ , i.e.

$$\psi_{jk}(x) = 2^{\frac{j}{2}} \psi_0\left(\frac{x-k}{2^{-j}}\right) \text{ and } \tilde{\psi}_{jk}(x) = 2^{\frac{j}{2}} \tilde{\psi}_0\left(\frac{x-k}{2^{-j}}\right). \quad (15)$$

We assume the following conditions to hold:

- direct estimate (estimate of Jackson type, approximation order  $m$ )

$$\inf_{u_j \in S_j} \|u - u_j\|_{\mathcal{L}^2} \leq C 2^{-jm} |u|_{\mathcal{H}^m} \quad \forall u \in \mathcal{H}^m \quad (16)$$

for some  $m \in \mathbb{N}$  with  $1 \leq m$ ,

- inverse estimate (Bernstein inequality)

$$\|u_j\|_{\mathcal{H}^r} \leq C 2^{jr} \|u_j\|_{\mathcal{L}^2} \quad \forall u_j \in S_j \quad (17)$$

for some  $r \in \mathbb{R}$  with  $r \in (0, m]$ .

We also assume that similar relations hold for the dual system  $\tilde{S}$  with parameters  $\tilde{m}$  and  $\tilde{r}$ . Then the validity of the following norm equivalences can be inferred from (16) and (17), see [Dahm96, O92].

$$\|u\|_{\mathcal{H}^t}^2 \approx \sum_{j=0}^{\infty} \|w_j\|_{\mathcal{H}^t}^2 \approx \sum_{j=0}^{\infty} 2^{2tj} \|w_j\|_{\mathcal{L}^2}^2 \text{ for } t \in (-\tilde{r}, r) \quad (18)$$

where

$$u = \sum_{j=0}^{\infty} w_j, w_j \in W_j. \quad (19)$$

(For  $t < 0$  the  $\mathcal{L}^2$ -convergence in (19) has to be replaced by distributional convergence.) Note that (18) with  $t = 0$  together with the local stability (10) enforces the global stability

$$\|u\|_{\mathcal{L}^2} \approx \|\{\langle u, \tilde{\psi}_{jk} \rangle\}_{jk}\|_{\ell^2(j \in \mathbb{N}, k \in \tau_j)},$$

i.e., (12). The two-sided estimate (18) allows to characterize smoothness properties of a function from the properties of a multiscale decomposition. It is a consequence of approximation theory

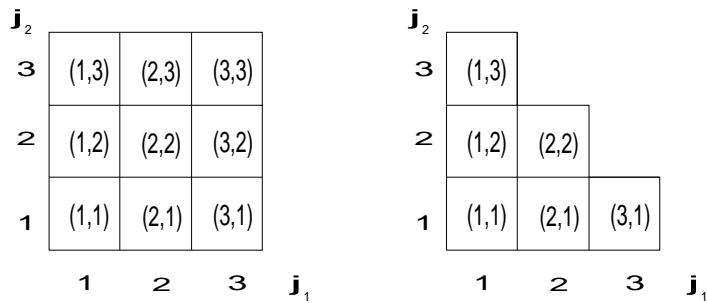


Figure 1: Index sets of the full grid space  $V_3^{-\infty}$  (left) and of the sparse grid space  $V_3^0$  (right)

in Sobolev spaces together with interpolation and duality arguments [O92, Dahm96]. Moreover, it states that bilinear forms  $a(\cdot, \cdot)$  satisfying the two-sided estimate (3) are spectrally equivalent to the sum of the bilinear forms  $2^{2sj}(\cdot, \cdot)_{\mathcal{L}^2}$  on  $W_{\mathbf{j}} \times W_{\mathbf{j}}$  induced from the right hand side of (18). A similar result holds for the analogous construction using the dual wavelets instead of the primal ones. This leads to the range  $t \in (-r, \tilde{r})$ . See [Dahm97] for an overview over multiscale methods dealing with biorthogonal wavelets.

For the higher-dimensional case  $n > 1$ , let  $\mathbf{j} \in \mathbb{Z}_+^n, \mathbf{j} \equiv (j_1, \dots, j_n)$ , 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.

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

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

Given an index set  $I_J \subset \mathbb{Z}_+^n, J \in \mathbb{N}$ , we consider the approximation spaces

$$V_J := \sum_{\mathbf{j} \in I_J} W_{\mathbf{j}}. \quad (20)$$

Here,  $J$  is the maximal level in  $V_J$ , i.e.  $j_i \leq J, i = 1, \dots, n \forall \mathbf{j} \in I_J$ . Associated with rectangular index sets  $I_J^{-\infty} := \{\|\mathbf{j}\|_{\infty} \leq J\}$  are the full grid spaces

$$V_J^{-\infty} := \bigoplus_{\|\mathbf{j}\|_{\infty} \leq J} W_{\mathbf{j}}, J > 0. \quad (21)$$

The so called sparse grid space

$$V_J^0 := \bigoplus_{\|\mathbf{j}\|_1 \leq J+n-1} W_{\mathbf{j}}, J > 0 \quad (22)$$

is associated with the index set  $I_J^0 := \{\|\mathbf{j}\|_1 \leq J+n-1\}$ . The approximation spaces  $V_J^{-\infty}$  and  $V_J^0$  will turn out to be special choices of a family of approximation spaces  $V_J^T$  that are adapted to Sobolev spaces. Specifically,  $V_J^0$  will turn out to be the appropriate choice for  $\mathcal{H}^0$ . See Figure 1 for the index sets of the full and the sparse grid spaces  $V_3^{-\infty}$  and  $V_3^0$  in the two-dimensional case.

The dimensions of  $W_{\mathbf{j}}, V_J^{-\infty}$  and  $V_J^0$  are (note that we count only interior grid points)

$$|W_{\mathbf{j}}| = 2^{|\mathbf{j}|_1 - n}, \quad (23)$$



$$|V_J^{-\infty}| = (2^J - 1)^n = O(2^{Jn}) \quad (24)$$

and

$$|V_J^0| = 2^J \left( \frac{J^{n-1}}{(n-1)!} + O(J^{n-2}) \right), \quad (25)$$

see [Bu92, BG98], e.g. The estimates of  $|W_J|$  and  $|V_J^{-\infty}|$  are clear. The estimate of  $|V_J^0|$  is straightforward and will follow as a byproduct of the estimate of the dimensions of the spaces from a more general class of spaces in section 4.2.

In this paper we introduce index sets that are optimized with respect to Sobolev norms and spaces with bounded mixed derivatives. Note that the dimension of the sparse grid space  $V_J^0$  compares favorable 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 in section 4 that for certain classes of smooth functions and for operators of positive order, the order of approximation of the sparse grid space is the same as for the space  $V_J^{-\infty}$ . It is even possible to use a subspace of the sparse grid space without losing the optimal order of approximation. Compare also [Bu92, BG98] for partial results in this direction for the case  $s = 1$  and finite element spaces of piecewise linear splines. For operators of negative order, we will show that the interplay between the dimension of the approximation space and the order of approximation is still in favor for the optimized approximation spaces.

Now, we define the spaces  $\mathcal{H}_{mix}^{t,l}$ . They fix the smoothness assumptions we consider. Note that we consider smoothness assumptions on the solution  $u$  or on the right hand side  $f$  (that in turn leads to smoothness assumptions on  $u$ ) of the variational problem. This leads us to the definition of more general spaces than the standard Sobolev spaces  $\mathcal{H}^t$ .

**Definition:** Let  $t \in \mathbb{R}_0^+$ ,  $l \in \mathbb{R}$ ,  $t + l \geq 0$ ,  $\mathbf{1} = (1, \dots, 1)$  and  $\mathbf{e}_i = (0, \dots, 0, 1, 0, \dots, 0)$  the  $i$ -th unit-vector in  $\mathbb{R}^n$ .

$$\mathcal{H}_{mix}^{t,l}(T^n) := \mathcal{H}_{mix}^{t+\mathbf{1}\mathbf{e}_1}(T^n) \cap \dots \cap \mathcal{H}_{mix}^{t+\mathbf{1}\mathbf{e}_n}(T^n), \quad (26)$$

where

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

Furthermore we write

$$\mathcal{H}_{mix}^t(T^n) := \mathcal{H}^t(T) \otimes \dots \otimes \mathcal{H}^t(T), t \geq 0. \quad (27)$$

These are spaces of dominating mixed derivative. For  $\mathbf{t} \in \mathbb{N}^n$  the space  $\mathcal{H}_{mix}^{\mathbf{t}}$  possesses the equivalent norm

$$\|u\|_{\mathcal{H}_{mix}^{\mathbf{t}}}^2 \approx \sum_{\mathbf{0} \leq \mathbf{k} \leq \mathbf{t}} \|u^{(\mathbf{k})}\|_{\mathcal{L}^2}^2. \quad (28)$$

Here,  $u^{(\mathbf{k})}$  is the generalized mixed derivative  $\frac{\partial^{|\mathbf{k}|}}{\partial^{k_1} \dots \partial^{k_n}} u$ . For example  $u^{(t, \dots, t)}$  is the  $nt$ -th order mixed derivative and describes the additional smoothness requirements for the space  $\mathcal{H}_{mix}^t$  compared to the larger isotropic Sobolev space  $\mathcal{H}^t$ .

Note that the relations

$$\mathcal{H}_{mix}^t \subset \mathcal{H}^t \subset \mathcal{H}_{mix}^{t/n} \text{ for } t \geq 0$$

and

$$\mathcal{H}_{mix}^{t/n} \subset \mathcal{H}^t \subset \mathcal{H}_{mix}^t \text{ for } t \leq 0$$

hold. See [ST87] for problems connected with these spaces and further references.

The spaces  $\mathcal{H}_{mix}^{t,l}$  are special cases of the spaces

$$\mathcal{H}_{mix,\cap}^{(\mathbf{t}^1, \dots, \mathbf{t}^n)}(T^n) := \mathcal{H}_{mix}^{\mathbf{t}^1}(T^n) \cap \dots \cap \mathcal{H}_{mix}^{\mathbf{t}^n}(T^n), \quad (29)$$

where  $\mathbf{t}^i \in \mathbb{R}^n, \mathbf{t}^i \geq 0, 1 \leq i \leq n$ . On the other hand the standard Sobolev spaces  $\mathcal{H}^t(T^n)$  as well as the spaces  $\mathcal{H}_{mix}^t(T^n)$  with dominating mixed derivative are special cases of the spaces  $\mathcal{H}_{mix}^{t,l}(T^n)$  defined in (26). We have

$$\mathcal{H}^t(T^n) = \mathcal{H}_{mix}^{0,t}(T^n) \quad (30)$$

and

$$\mathcal{H}_{mix}^t(T^n) = \mathcal{H}_{mix}^{t,0}(T^n). \quad (31)$$

Indeed, for  $t \in \mathbb{R}_0^+$  we have the representation

$$\begin{aligned} \mathcal{H}_{mix}^{0,t}(T^n) &= \mathcal{H}_{mix}^{(t,0,\dots,0)}(T^n) \cap \dots \cap \mathcal{H}_{mix}^{(0,\dots,0,t)}(T^n) \\ &= \mathcal{H}_{mix}^{t\mathbf{e}_1}(T^n) \cap \dots \cap \mathcal{H}_{mix}^{t\mathbf{e}_n}(T^n) \\ &= \mathcal{H}^t(T^n) \end{aligned} \quad (32)$$

where

$$\mathcal{H}_{mix}^{(0,\dots,0,1,0,\dots,0)}(T^n) := \mathcal{L}^2(T) \otimes \dots \otimes \mathcal{L}^2(T) \otimes \mathcal{H}^t(T) \otimes \mathcal{L}^2(T) \otimes \dots \otimes \mathcal{L}^2(T). \quad (33)$$

To prove the last equality in (32), choose an orthogonal basis of  $\mathcal{H}^t(T)$  and use the definition of the tensor-product via orthonormal systems (7). More precisely, using periodic continuation to  $\mathbb{R}$  and the fact that for example  $\{\sin(n(2x\pi - \pi))\}$  defines a complete orthonormal system in  $\mathcal{L}^2(T)$  and  $\mathcal{H}^t(T)$  it is clear that every  $u \in \mathcal{H}^t(T^n)$  can be represented as a Fourier sine series and (32) follows directly from the definition of the tensor product (7) and the definition of intersection of Hilbert spaces. Note that similar results hold for problems with Dirichlet or Neumann boundary conditions and certain cases of mixed boundary conditions. See [GO95] for more details and some examples. Equation (31) is clear from the definition of  $\mathcal{H}_{mix}^t(T^n)$  in (27). A norm on  $\mathcal{H}_{mix}^{t,l}(T^n)$  can be defined directly via

$$\|u\|_{\mathcal{H}_{mix}^{t,l}}^2 \approx \sum_{1 \leq i \leq n} \|u\|_{\mathcal{H}_{mix}^{t\mathbf{1} + l\mathbf{e}_i}}^2.$$

Hence, the spaces  $\mathcal{H}_{mix}^{t,l}$  from (26) give a unified framework for the study of the special cases  $\mathcal{H}^t = \mathcal{H}_{mix}^{0,t}$  and  $\mathcal{H}_{mix}^t = \mathcal{H}_{mix}^{t,0}$ .

### 3 Norm equivalences

To get norm equivalences analogous to (18) in  $n \geq 2$  dimensions, we use the above representation of  $\mathcal{H}^t$  and  $\mathcal{H}_{mix}^t$  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  $a(\cdot, \cdot)$ . Consider a collection of Hilbert spaces  $H_l, l = 1, \dots, n, n \in \mathbb{N}$  and a collection of closed subspaces  $V_l \subset H_l$  such that topologically  $H_l = \sum_i V_l$ . An additive subspace splitting  $\{H_l; a_l\} = \sum_i \{V_l; b_l\}$  is called stable if the norm equivalence  $a_l(u, u) \approx |||u|||^2 \equiv \inf_{u_i \in H_l: u = \sum_i u_i} (b_l(u_i, u_i))$  holds true, i.e. the characteristic numbers

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

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*

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

*are stable and possess the condition number  $\kappa_l$ , respectively, then the tensor-product splitting*

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

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

See [GO95] for a proof in the 2D case. The extension to the n-dimensional case is straightforward. Therefore we omit it here.

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

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

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

$$\{H_1 \cap \dots \cap H_n; \alpha_1 a_1 + \dots + \alpha_n a_n\} = \sum_i \{V_i; (\alpha_1 \alpha_{1i} + \dots + \alpha_n \alpha_{ni})b\} \quad (34)$$

*is stable with condition number  $\kappa \leq \frac{\max\{\lambda_{max,1}, \dots, \lambda_{max,n}\}}{\min\{\lambda_{min,1}, \dots, \lambda_{min,n}\}}$ .*

Proof: See [GO95] for a proof in the 2D case. The n-dimensional case is analogous. We give it here for the sake of completeness.

Since the splittings are into direct sums, every  $u \in H_1 \cap \dots \cap H_n$  has a unique representation  $u = \sum_i u_i$  with respect to  $\{V_i\}$  such that

$$\lambda_{max,l}^{-1} a_l(u, u) \leq \sum_i \alpha_{li} b(u_i, u_i) \leq \lambda_{min,l}^{-1} a_l(u, u), l = 1, \dots, n.$$

Multiplying with  $\alpha_l$  and adding these two-sided inequalities gives

$$\sum_l \alpha_l \lambda_{max,l}^{-1} a_l(u, u) \leq \sum_i (\alpha_1 \alpha_{1i} + \dots + \alpha_n \alpha_{ni}) b(u_i, u_i) \leq \sum_l \alpha_l \lambda_{min,l}^{-1} a_l(u, u).$$

Hence

$$\frac{\sum_l \alpha_l a_l(u, u)}{\max_l \{\lambda_{max,l}\}} \leq \sum_i (\alpha_1 \alpha_{1i} + \dots + \alpha_n \alpha_{ni}) b(u_i, u_i) \leq \frac{\sum_l \alpha_l a_l(u, u)}{\min_l \{\lambda_{min,l}\}}$$

which shows that the splitting (34) is stable with characteristic numbers

$$\lambda_{min} \geq 1 / \max_l \{\lambda_{max,l}\}, \lambda_{max} \leq 1 / \min_l \{\lambda_{min,l}\}, \kappa = \frac{\max_l \{\lambda_{max,l}\}}{\min_l \{\lambda_{min,l}\}}.$$

□

Combining the above representation (32) of  $\mathcal{H}^t(T^n), t \geq 0$ , with these Propositions and the stability result (18) in one dimension we come up with the following norm equivalence and stable splitting of  $\mathcal{H}^t(T^n)$ .

**Theorem 1** Let  $u \in \mathcal{H}^t(T^n)$ ,  $u = \sum_{\mathbf{j}} w_{\mathbf{j}}$ ,  $w_{\mathbf{j}} \in W_{\mathbf{j}}$  (for  $t < 0$  with distributional convergence) and let the above assumptions (16) and (17) on the validity of a Jackson and a Bernstein inequality for the primal as well as the dual system hold. Then

$$\|u\|_{\mathcal{H}^t}^2 \approx \sum_{\mathbf{j}} 2^{2t|\mathbf{j}|_{\infty}} \|w_{\mathbf{j}}\|_{\mathcal{L}^2}^2 \text{ for } t \in (-\tilde{r}, r), \quad (35)$$

where  $|\mathbf{j}|_{\infty} = \max_{1 \leq i \leq n} j_i$ .

**Proof:** In the one-dimensional case we have from (18)

$$\|u\|_{\mathcal{H}^t(T)}^2 \approx \sum_{j=0}^{\infty} 2^{2tj} \|w_j\|_{\mathcal{L}^2(T)}^2, \quad 0 \leq t < r, \quad u = \sum_{j=0}^{\infty} w_j, \quad w_j \in W_j, \quad u \in \mathcal{H}^t(T)$$

and from (12)

$$\|u\|_{\mathcal{L}^2} \approx \|\{u, \tilde{\psi}_{jt}\}_{jt}\|_{\ell^2(j \in \mathbb{N}, t \in I_j)}, \quad u = \sum_{j=0}^{\infty} w_j, \quad w_j \in W_j, \quad u \in \mathcal{L}^2(T).$$

This shows the stability of the 1D splittings

$$\{\mathcal{H}^t(T); \|\cdot\|_{\mathcal{H}^t(T)}^2\} = \sum_j \{W_j; 2^{2tj} \|\cdot\|_{\mathcal{L}^2(T)}^2\}$$

and

$$\{\mathcal{L}^2(T); \|\cdot\|_{\mathcal{L}^2(T)}^2\} = \sum_j \{W_j; \|\cdot\|_{\mathcal{L}^2(T)}^2\}.$$

From Proposition 1 we obtain the stability of the splittings

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

Now we represent  $\mathcal{H}^t(T^n)$  as in (32) and we apply Proposition 2. Then, we obtain the stability of the splitting

$$\{\mathcal{H}^t(T^n); \|\cdot\|_{\mathcal{H}^t}^2\} = \sum_{\mathbf{j}} \{W_{\mathbf{j}}; (\sum_{i=1}^n 2^{2tj_i}) \|\cdot\|_{\mathcal{L}^2}^2\} \quad (36)$$

for nonnegative  $t < r$ . Because of  $2^{2t|\mathbf{j}|_{\infty}} \leq \sum_{i=1}^n 2^{2tj_i} \leq n 2^{2t|\mathbf{j}|_{\infty}}$  for  $t \geq 0$  we then have (35) for positive  $t$ . To obtain the validity of (35) 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 a similar result for the spaces spanned by the dual wavelets for  $0 \leq t < \tilde{r}$ . By duality  $(\mathcal{H}^t)' = \mathcal{H}^{-t}$  the assertion follows then for the range  $-\tilde{r} < t < 0$  and hence for the whole range  $t \in (-\tilde{r}, r)$ .  $\square$

For the space  $H_{mix}^t$  the following norm equivalence holds.

**Theorem 2** Let  $u \in \mathcal{H}_{mix}^t$ ,  $u = \sum_{\mathbf{j}} w_{\mathbf{j}}$ ,  $w_{\mathbf{j}} \in W_{\mathbf{j}}$  and let the assumptions (16) and (17) on the validity of a Jackson and a Bernstein inequality for the primal as well as the dual system hold. Then

$$\|u\|_{\mathcal{H}_{mix}^t}^2 \approx \sum_{\mathbf{j}} 2^{2t|\mathbf{j}|_1} \|w_{\mathbf{j}}\|_{\mathcal{L}^2}^2 \text{ for } t \in (-\tilde{r}, r). \quad (37)$$

**Proof:** The two-sided estimate (37) is a direct consequence of Proposition 1 and the definition of the space  $\mathcal{H}_{mix}^t$  as tensor-product of one-dimensional Hilbert spaces. Again we use the stable 1D splittings

$$\{\mathcal{H}^t(T); \|\cdot\|_{\mathcal{H}^t(T)}^2\} = \sum_j \{W_j; 2^{2tj} \|\cdot\|_{\mathcal{L}^2}^2\}$$

and

$$\{\mathcal{L}^2(T); \|\cdot\|_{\mathcal{L}^2}^2\} = \sum_j \{W_j; \|\cdot\|_{\mathcal{L}^2}^2\}$$

(which we get from (18) and (12)) and Proposition 1 to obtain the stability of the splitting

$$\begin{aligned} & \{\mathcal{H}_{mix}^t; a(\cdot, \cdot) \otimes \dots \otimes a(\cdot, \cdot)\} \\ &= \{\mathcal{H}^t(T) \otimes \dots \otimes \mathcal{H}^t(T); a(\cdot, \cdot) \otimes \dots \otimes a(\cdot, \cdot)\} \\ &= \sum_{\mathbf{j}} \{W_{j_1} \otimes \dots \otimes W_{j_n}; 2^{2tj_1}(\cdot, \cdot)_{\mathcal{L}^2} \otimes \dots \otimes 2^{2tj_n}(\cdot, \cdot)_{\mathcal{L}^2}\} \\ &= \sum_{\mathbf{j}} \{W_{j_1} \otimes \dots \otimes W_{j_n}; 2^{2t|\mathbf{j}|_1}(\cdot, \cdot)_{\mathcal{L}^2} \otimes \dots \otimes (\cdot, \cdot)_{\mathcal{L}^2}\}. \end{aligned}$$

This shows (37).

□

Note that under the assumptions of the Theorems 1 and 2 there hold similar relations for the subspace splittings induced by the dual wavelets.

**Remark 1** The norm equivalences in Theorems 1 and 2 are special cases of norm equivalences for the spaces  $\mathcal{H}_{mix, \Omega}^{(\mathbf{t}^1, \dots, \mathbf{t}^n)}$  from (29). Again using Propositions 1 and 2 it is straightforward to show that

$$\|u\|_{\mathcal{H}_{mix, \Omega}^{(\mathbf{t}^1, \dots, \mathbf{t}^n)}}^2 \approx \sum_{\mathbf{j}} \left( \sum_{i=1}^n 2^{2\langle \mathbf{t}^i, \mathbf{j} \rangle} \right) \|w_{\mathbf{j}}\|_{\mathcal{L}^2}^2 \text{ for } \mathbf{t}^i \geq 0, -\tilde{r} < \mathbf{t}^i < r. \quad (38)$$

Specifically for the spaces  $\mathcal{H}_{mix}^{t,l}, t+l \geq 0, -\tilde{r} < t+l < r, 0 \leq t < r$ , the norm equivalence (38) reads

$$\|u\|_{\mathcal{H}_{mix}^{t,l}}^2 \approx \sum_{\mathbf{j}} \left( \sum_{i=1}^n 2^{2t|\mathbf{j}|_1 + 2lj_i} \right) \|w_{\mathbf{j}}\|_{\mathcal{L}^2}^2 \approx \sum_{\mathbf{j}} 2^{2t|\mathbf{j}|_1 + 2l|\mathbf{j}|_\infty} \|w_{\mathbf{j}}\|_{\mathcal{L}^2}^2. \quad (39)$$

Compared to (35), (37), the additional factors  $2^{2t|\mathbf{j}|_1}$  or  $2^{2l|\mathbf{j}|_\infty}$  in (39) reflect the different smoothness requirements. Note that for  $t = 0$  or  $l = 0$  we regain (35) from Theorem 1 and (37) from Theorem 2, respectively. Analogous relations hold for the dual spaces.

**Remark 2** For the construction of optimized approximation spaces, we will use the upper estimate from (35) and the lower estimate in (37) and (39).

**Remark 3** One of the merits of the norm equivalences (35), (37) or the more general one (39) is the fact that they lead directly to optimal preconditioning. For example, if one chooses the scaled system  $\{2^{-s|\mathbf{l}|_\infty} \psi_{\mathbf{l}\mathbf{k}} : |\mathbf{l}|_\infty \leq J, \mathbf{k} \in \tau_{\mathbf{l}}\}$  as the basis in the finite element approximation space  $V_J^{-\infty}$ , then the spectral condition numbers  $\kappa(A_J)$  of the discretization matrices  $A_J = \{2^{-s|\mathbf{l}|_\infty + l'|\mathbf{l}'|_\infty} a(\psi_{\mathbf{l}\mathbf{k}}, \psi_{\mathbf{l}'\mathbf{k}'})\}_{\mathbf{l}, \mathbf{l}', \mathbf{k}, \mathbf{k}'}$  are bounded uniformly in  $J$ , i.e.

$$\kappa(A_J) = O(1), \quad (40)$$

see [Be93, DK92, J92]. This leads to fast iterative methods with convergence rates independent of the number of unknowns of the approximation space. Note that this result can be trivially extended to the case of discretization matrices built from arbitrary collections of scaled basis functions.

## 4 Optimized approximation spaces for Sobolev spaces

Suppose a symmetric elliptic variational problem (2) and its variational formulation

$$a(u_{FE}, v) = (f, v) \quad \forall v \in V_{FE} \quad (41)$$

on a finite element approximation space  $V_{FE} \subset \mathcal{H}^s$  are given. Then, we have due to the  $\mathcal{H}^s$ -ellipticity condition (3) and Cea's Lemma

$$\sqrt{a(u - u_{FE}, u - u_{FE})} \approx \|u - u_{FE}\|_{\mathcal{H}^s} \approx \inf_{v \in V_{FE}} \|u - v\|_{\mathcal{H}^s}$$

for the error  $\sqrt{a(u - u_{FE}, u - u_{FE})}$  between the solution  $u$  of the continuous problem (2) and the solution  $u_{FE}$  of the approximate problem (41) measured in the energy norm. In this section we give 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

$$u \in \mathcal{H}_{mix}^{t,l}, \text{ where } -\tilde{r} < s < t + l < r, 0 \leq t < r \text{ and } t + l \geq 0.$$

Specifically we are interested in the cases

$$u \in \mathcal{H}^t \text{ and } u \in \mathcal{H}_{mix}^t, \text{ where } -\tilde{r} < s < t < r, t \geq 0.$$

We define grids and associated approximation spaces that are adapted to the parameter  $s$  and to the constraint on the smoothness of the solution and give estimates on their dimension and the order of approximation. The definition of the grids is motivated by the results of section 3, specifically on the norm equivalence (39) and the special cases (35) and (37). We are particularly interested in constructing approximation spaces that break the curse of dimensionality, that is whose dimension is at most polynomially dependent on  $n$ .

### 4.1 Approximation spaces for problems with constraint on the solution

We first deal with the cases  $u \in \mathcal{H}^t$  and  $u \in \mathcal{H}_{mix}^t$ . More general cases will be discussed at the end of section 4.1.2, see Theorem 3. In this section let  $u = \sum_{\mathbf{j}} w_{\mathbf{j}}$ , where  $w_{\mathbf{j}} \in W_{\mathbf{j}}$ . Furthermore let  $-\tilde{r} < s < t < r$ . Then  $\mathcal{H}^t \subset \mathcal{H}^s$ . For notational convenience we restrict ourselves to the case  $t > 0$ . Note that the case  $t < 0$  could be covered with analogous reasoning.

#### 4.1.1 Estimates on the order of approximation for the spaces $V_J^{-\infty}$ and $V_J^0$

First of all we consider the order of approximation for the full grid case. Let  $u \in \mathcal{H}^s$ . Applying the norm equivalence (35) gives us

$$\begin{aligned} \inf_{v \in V_J^{-\infty}} \|u - v\|_{\mathcal{H}^s}^2 &\leq \|u - \sum_{|\mathbf{j}|_{\infty} \leq J} w_{\mathbf{j}}\|_{\mathcal{H}^s}^2 \stackrel{(35)}{\approx} \sum_{|\mathbf{j}|_{\infty} > J} 2^{2s|\mathbf{j}|_{\infty}} \|w_{\mathbf{j}}\|_{\mathcal{L}^2}^2 \\ &= \sum_{|\mathbf{j}|_{\infty} > J} 2^{2(s-t)|\mathbf{j}|_{\infty}} 2^{2t|\mathbf{j}|_{\infty}} \|w_{\mathbf{j}}\|_{\mathcal{L}^2}^2 \\ &\leq \max_{|\mathbf{j}|_{\infty} > J} 2^{2(s-t)|\mathbf{j}|_{\infty}} \sum_{|\mathbf{j}|_{\infty} > J} 2^{2t|\mathbf{j}|_{\infty}} \|w_{\mathbf{j}}\|_{\mathcal{L}^2}^2. \end{aligned} \quad (42)$$

To continue, we assume additional smoothness of the solution, i.e.  $u \in \mathcal{H}^t$ . Then we can apply (35) once more, now with  $u \in \mathcal{H}^t$ . This yields

$$\max_{|\mathbf{j}|_\infty > J} 2^{2(s-t)|\mathbf{j}|_\infty} \sum_{|\mathbf{j}|_\infty > J} 2^{2t|\mathbf{j}|_\infty} \|w_{\mathbf{j}}\|_{\mathcal{L}^2}^2 \stackrel{(35)}{\leq} C \cdot \max_{|\mathbf{j}|_\infty > J} 2^{2(s-t)|\mathbf{j}|_\infty} \|u\|_{\mathcal{H}^t}^2 \leq C \cdot 2^{2(s-t)(J+1)} \|u\|_{\mathcal{H}^t}^2.$$

Altogether we have the standard error estimate

$$\inf_{v \in V_J^-} \|u - v\|_{\mathcal{H}^s}^2 \leq C \cdot 2^{2(s-t)(J+1)} \|u\|_{\mathcal{H}^t}^2 \text{ for } u \in \mathcal{H}^t \text{ and } -\tilde{r} < s < t < r. \quad (43)$$

From the exponent on the right hand side we get  $O((s-t)J)$  as order of approximation. It is easy to see that the order of approximation does not change when  $u \in \mathcal{H}_{mix}^t \subset \mathcal{H}^t$ , i.e.

$$\inf_{v \in V_J^-} \|u - v\|_{\mathcal{H}^s}^2 \leq C \cdot 2^{2(s-t)(J+1)} \|u\|_{\mathcal{H}_{mix}^t}^2 \text{ for } -\tilde{r} < s < t < r.$$

(For  $t < 0$  we would have to assume  $u \in \mathcal{H}_{mix}^t \cap \mathcal{H}^s$  here.) Note that we are implicitly using several times the vanishing moment condition of the dual wavelets, which is implicitly contained in the Jackson inequality (16).

Changing from the full grid space  $V_J^-$  to the approximation space  $V_J^0$  changes the situation significantly. Applying again the norm equivalence (35) gives for  $u \in \mathcal{H}^s$

$$\begin{aligned} \inf_{v \in V_J^0} \|u - v\|_{\mathcal{H}^s}^2 &\leq \|u - \sum_{|\mathbf{j}|_1 \leq J+n-1} w_{\mathbf{j}}\|_{\mathcal{H}^s}^2 \stackrel{(35)}{\approx} \sum_{|\mathbf{j}|_1 > J+n-1} 2^{2s|\mathbf{j}|_\infty} \|w_{\mathbf{j}}\|_{\mathcal{L}^2}^2 \\ &= \sum_{|\mathbf{j}|_1 > J+n-1} 2^{2(s-t)|\mathbf{j}|_\infty} 2^{2t|\mathbf{j}|_\infty} \|w_{\mathbf{j}}\|_{\mathcal{L}^2}^2 \leq \max_{|\mathbf{j}|_1 > J+n-1} 2^{2(s-t)|\mathbf{j}|_\infty} \sum_{|\mathbf{j}|_1 > n+J-1} 2^{2t|\mathbf{j}|_\infty} \|w_{\mathbf{j}}\|_{\mathcal{L}^2}^2. \end{aligned}$$

Now we again require  $u$  to be of higher regularity, i.e.  $u \in \mathcal{H}^t$ . This yields

$$\begin{aligned} \max_{|\mathbf{j}|_1 > J+n-1} 2^{2(s-t)|\mathbf{j}|_\infty} \sum_{|\mathbf{j}|_1 > n+J-1} 2^{2t|\mathbf{j}|_\infty} \|w_{\mathbf{j}}\|_{\mathcal{L}^2}^2 \\ \stackrel{(35)}{\leq} C \cdot \max_{|\mathbf{j}|_1 > J+n-1} 2^{2(s-t)|\mathbf{j}|_\infty} \|u\|_{\mathcal{H}^t}^2 \leq C \cdot 2^{2(s-t)(1-\frac{1}{n})} 2^{2(s-t)\frac{J}{n}} \|u\|_{\mathcal{H}^t}^2 \end{aligned}$$

where we used in the last but one step that the maximum is obtained for  $|\mathbf{j}|_\infty = \lceil \frac{J+n-1}{n} \rceil$ . Altogether we have

$$\inf_{v \in V_J^0} \|u - v\|_{\mathcal{H}^s}^2 \leq C \cdot 2^{2(s-t)(1-\frac{1}{n})} 2^{2(s-t)\frac{J}{n}} \|u\|_{\mathcal{H}^t}^2 \text{ for } u \in \mathcal{H}^t \text{ and } -\tilde{r} < s < t < r. \quad (44)$$

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

However, 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 is replaced by the space  $V_J^0$ . This is due to the fact that we can apply norm equivalence (37) instead of (35) (remember the different exponents of the forefactors in (35) and (37)). We apply (35) for functions from  $\mathcal{H}^s$  and (37) for  $u \in \mathcal{H}_{mix}^t$  and get

$$\inf_{v \in V_J^0} \|u - v\|_{\mathcal{H}^s}^2 \leq \|u - \sum_{|\mathbf{j}|_1 \leq J+n-1} w_{\mathbf{j}}\|_{\mathcal{H}^s}^2 \stackrel{(35)}{\approx} \sum_{|\mathbf{j}|_1 > J+n-1} 2^{2s|\mathbf{j}|_\infty} \|w_{\mathbf{j}}\|_{\mathcal{L}^2}^2$$

$$\begin{aligned}
&= \sum_{|\mathbf{j}|_1 > J+n-1} 2^{2s|\mathbf{j}|_\infty - 2t|\mathbf{j}|_1} 2^{2t|\mathbf{j}|_1} \|w_{\mathbf{j}}\|_{\mathcal{L}^2}^2 \\
&\leq \max_{|\mathbf{j}|_1 > J+n-1} 2^{2s|\mathbf{j}|_\infty - 2t|\mathbf{j}|_1} \sum_{|\mathbf{j}|_1 > n+J-1} 2^{2t|\mathbf{j}|_1} \|w_{\mathbf{j}}\|_{\mathcal{L}^2}^2 \\
&\stackrel{(37)}{\leq} C \cdot \max_{|\mathbf{j}|_1 > J+n-1} 2^{2s|\mathbf{j}|_\infty - 2t|\mathbf{j}|_1} \|u\|_{\mathcal{H}_{mix}^t}^2 \tag{45} \\
&\leq C \cdot 2^{-2tn} 2^{2s} 2^{2(s-t)J} \|u\|_{\mathcal{H}_{mix}^t}^2 \text{ for } u \in \mathcal{H}_{mix}^t, \tag{46}
\end{aligned}$$

where we used in the last step that  $2^{2s|\mathbf{j}|_\infty - 2t|\mathbf{j}|_1}$  takes its maximum in  $(J+1, 1, \dots, 1)$ . Altogether we have

$$\inf_{v \in V_J^0} \|u - v\|_{\mathcal{H}^s}^2 \leq C \cdot 2^{-2tn} 2^{2s} 2^{2(s-t)J} \|u\|_{\mathcal{H}_{mix}^t}^2 \text{ for } u \in \mathcal{H}_{mix}^t \text{ and } -\tilde{r} < s < t < r, t > 0.$$

That is there appears no loss in the order of approximation compared to the result for the full grid approximation space.

For operators of negative order, i.e.  $s < 0$ , the situation is different. Here, compared to the estimate (43) 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. Applying (35) for functions from  $\mathcal{H}^s$  and (37) we have for  $u \in \mathcal{H}_{mix}^t$

$$\begin{aligned}
\inf_{v \in V_J^0} \|u - v\|_{\mathcal{H}^s}^2 &\leq \|u - \sum_{|\mathbf{j}|_\infty > J+n-1} w_{\mathbf{j}}\|_{\mathcal{H}^s}^2 \stackrel{(35)}{\approx} \sum_{|\mathbf{j}|_1 > J+n-1} 2^{2s|\mathbf{j}|_\infty} \|w_{\mathbf{j}}\|_{\mathcal{L}^2}^2 \\
&= \sum_{|\mathbf{j}|_1 > J+n-1} 2^{2s|\mathbf{j}|_\infty - 2t|\mathbf{j}|_1} 2^{2t|\mathbf{j}|_1} \|w_{\mathbf{j}}\|_{\mathcal{L}^2}^2 \leq \max_{|\mathbf{j}|_1 > J+n-1} 2^{2s|\mathbf{j}|_\infty - 2t|\mathbf{j}|_1} \sum_{|\mathbf{j}|_1 > n+J-1} 2^{2t|\mathbf{j}|_1} \|w_{\mathbf{j}}\|_{\mathcal{L}^2}^2 \\
&\stackrel{(37)}{\leq} C \cdot \max_{|\mathbf{j}|_1 > J+n-1} 2^{2s|\mathbf{j}|_\infty - 2t|\mathbf{j}|_1} \|u\|_{\mathcal{H}_{mix}^t}^2 \leq C \cdot 2^{2s(1-\frac{1}{n})} 2^{-2tn} 2^{2(s/n-t)J} \|u\|_{\mathcal{H}_{mix}^t}^2, \tag{47}
\end{aligned}$$

where we used in the last step that  $2^{2s|\mathbf{j}|_\infty - 2t|\mathbf{j}|_1}$  takes its maximum for  $|\mathbf{j}|_\infty = \lceil \frac{J+n-1}{n} \rceil$  and  $|\mathbf{j}|_1 = J+n$  for  $s < 0$ . That is, although the order of approximation is improved when changing from  $\mathcal{H}^t$  to  $\mathcal{H}_{mix}^t$  there still appears a loss in the order of approximation of  $s(1-1/n)$  compared to the full grid. This fact has been described already in [GOS98] 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  $W_{\mathbf{j}}$ ), where this behavior is explained in more detail. In summary we have that for operators with  $s \geq 0$  the order of approximation is kept for  $u \in \mathcal{H}_{mix}^t$ ,  $s < t$ , when changing from the approximation space  $V_J^{-\infty}$  to the sparse grid space  $V_J^0$ . For operators of negative order a deterioration of the order of approximation appears.

#### 4.1.2 Definition and order of approximation of the approximation spaces $V_J^T$

In the following we construct approximation spaces for functions from  $\mathcal{H}_{mix}^{t,l}$ ,  $-\tilde{r} < s < t+l < r$ ,  $t+l \geq 0$ , and operators of positive or negative order by carefully selecting subspaces of the full grid space. They are chosen in such a way that the order of approximation of the full grid space is kept. The sparse grid space  $V_J^0$  and the full grid space  $V_J^{-\infty}$  are special cases. We start with the space  $\mathcal{H}_{mix}^t = \mathcal{H}_{mix}^{t,0}$ . Inequality

$$\max_{\mathbf{j} \notin I_J^0} 2^{2s|\mathbf{j}|_\infty - 2t|\mathbf{j}|_1} \|u\|_{\mathcal{H}_{mix}^t}^2 \leq C \cdot 2^{2(s-t)J} \|u\|_{\mathcal{H}_{mix}^t}^2 \text{ for } u \in \mathcal{H}_{mix}^t, 0 \leq s < t,$$



from (45) reveals that for  $s \geq 0$  one could discard indices from the index set  $I_J^0$  without destroying the optimal order of approximation. Consider an index set  $I_J \subset I_J^0$  such that

$$\max_{\mathbf{j} \notin I_J} 2^{2s|\mathbf{j}|_\infty - 2t|\mathbf{j}|_1} \leq C \cdot 2^{2(s-t)J} \quad (48)$$

where  $C \neq C(s, t, J)$ . Then the order of approximation is kept for the approximation space defined from the index set  $I_J$ . Taking logarithms on both sides of (48) and dividing by  $2t$  (remember that we have  $t > 0$ ) shows that (48) is equivalent to

$$\mathbf{j} \in I_J \Leftrightarrow -|\mathbf{j}|_1 + \frac{s}{t}|\mathbf{j}|_\infty \geq -J + \frac{s}{t}J + c, \quad (49)$$

where  $c \neq c(\mathbf{j}, J)$  is essentially the logarithm of the constant  $C$  on the right hand side of the asymptotic estimate (48). For operators of negative order we deduce from (47) that we have to add indices to the index set  $I_J^0$  to keep the optimal order of approximation. Again, the order is kept if  $I_J$  is such that (48) and hence (49) holds.

Therefore we define the optimized grid as the minimal index set for which (49) holds. Fixing  $(J, 1, \dots, 1)$  to be the index with maximal  $|\cdot|_\infty$ -norm to be included into the index sets leads to  $c = n - 1$  and the index sets

$$I_J^{\frac{s}{t}} := \{\mathbf{j} : -|\mathbf{j}|_1 + \frac{s}{t}|\mathbf{j}|_\infty \geq -(n + J - 1) + \frac{s}{t}J\}.$$

They are dependent on the parameter  $J$  and on the quotient  $s/t$ .

In order to give the results more flexibility we parametrize the index sets with a new parameter  $T$  and get finally

$$I_J^T := \{\mathbf{j} : -|\mathbf{j}|_1 + T|\mathbf{j}|_\infty \geq -(n + J - 1) + TJ\} \quad (50)$$

with the related approximation spaces

$$\begin{aligned} V_J^T &:= \bigoplus_{\mathbf{j} \in I_J^T} W_{\mathbf{j}} \\ &= \bigoplus_{-|\mathbf{j}|_1 + T|\mathbf{j}|_\infty \geq -(n+J-1)+TJ} W_{\mathbf{j}}. \end{aligned} \quad (51)$$

The new parameter  $T$  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. In the following we will consider terms like

$$\inf_{v \in V_J^T} \|u - v\|_{\mathcal{H}^s}^2,$$

now for varying  $T$ , where we assume again that  $u \in \mathcal{H}^t$  or  $u \in \mathcal{H}_{mix}^t$ . Definition (50) ensures that the optimal order of approximation is kept for  $T \leq \frac{s}{t}$  and functions from  $\mathcal{H}_{mix}^t$  (compare (48) and (49)). For  $T > \frac{s}{t}$  the order of approximation deteriorates. We discuss this point in more detail below.

Note that for  $T = 0$  we have  $V_J^T = V_J^0$  and for  $T \rightarrow -\infty$  we have  $V_J^T \rightarrow V_J^{-\infty}$ , i.e. the full grid space. Furthermore we have the natural restriction to  $T \leq 1$ . Obviously the inclusions

$$V_J^1 \subset V_J^{T_1} \subset V_J^{T_2} \subset V_J^0 \subset V_J^{T_3} \subset V_J^{T_4} \subset V_J^{-\infty} \text{ for } T_4 \leq T_3 \leq 0 \leq T_2 \leq T_1 \leq 1 \quad (52)$$

hold. Schematically the behavior of the index sets  $I_J^T$  is depicted in Figure 2 with varying  $T$  for the two-dimensional case. Figures 3-6 show some examples for the two-dimensional case.



Figure 2: Index sets  $I_j^T$  for  $T > 0$ ,  $T = 0$  and  $T < 0$

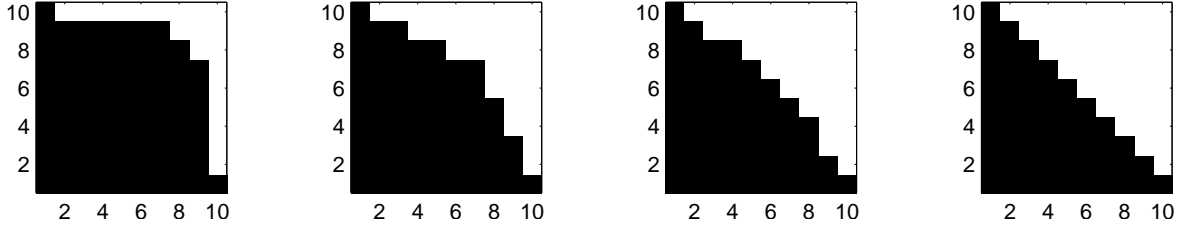


Fig. 3: Index sets  $I_{10}^{-5}$ ,  $I_{10}^{-1}$ ,  $I_{10}^{-1/2}$  and  $I_{10}^{-1/8}$ , from left to right.

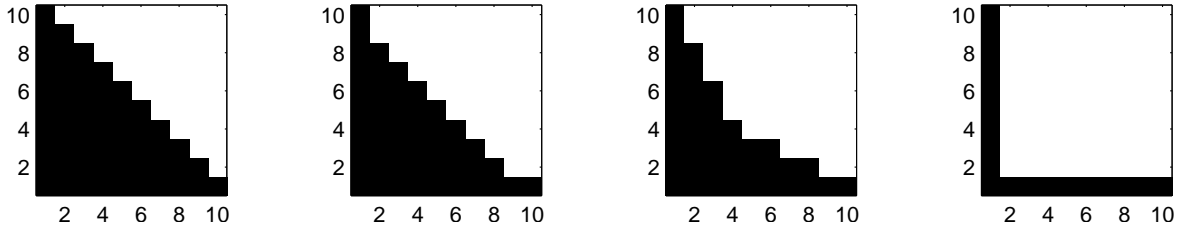


Fig. 4: Index sets  $I_{10}^0$ ,  $I_{10}^{1/8}$ ,  $I_{10}^{1/2}$  and  $I_{10}^1$ , from left to right.

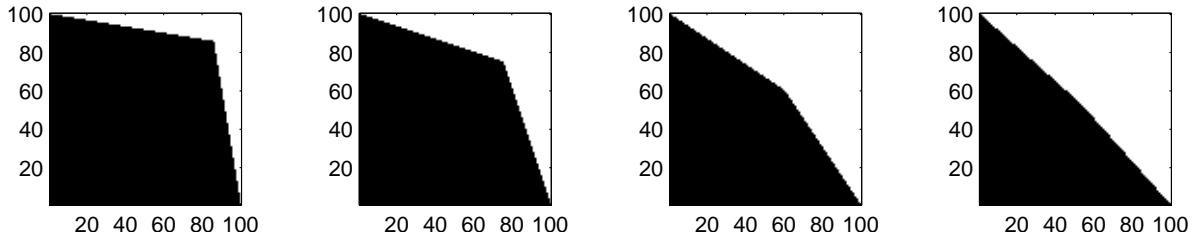


Fig. 5: Index sets  $I_{100}^{-5}$ ,  $I_{100}^{-1}$ ,  $I_{100}^{-1/2}$  and  $I_{100}^{-1/8}$ , from left to right.

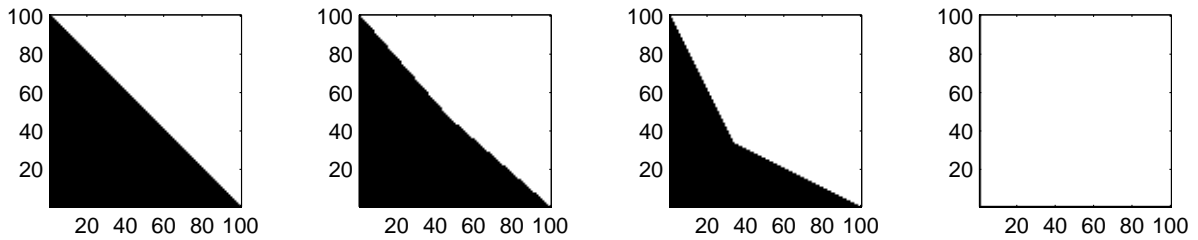


Fig. 6: Index sets  $I_{100}^0$ ,  $I_{100}^{1/8}$ ,  $I_{100}^{1/2}$  and  $I_{100}^1$ , from left to right.

We now discuss the dependence of the order of approximation of the approximation space  $V_J^T$  on the parameter  $T$  in more detail. Let us first consider the case  $u \in \mathcal{H}^t$ . Remember that  $\mathcal{H}^t \subset \mathcal{H}^s$ . Then we have (similar to (44))

$$\begin{aligned}
\inf_{v \in V_J^T} \|u - v\|_{\mathcal{H}^s}^2 &\leq \|u - \sum_{\mathbf{j} \in I_J^T} w_{\mathbf{j}}\|_{\mathcal{H}^s}^2 \stackrel{(35)}{\leq} C \cdot \max_{\mathbf{j} \notin I_J^T} 2^{2(s-t)|\mathbf{j}|_\infty} \|u\|_{\mathcal{H}^t}^2 \\
&\stackrel{(50)}{=} C \cdot \max_{T|\mathbf{j}|_\infty - |\mathbf{j}|_1 < TJ - (n+J-1)} 2^{2(s-t)|\mathbf{j}|_\infty} \|u\|_{\mathcal{H}^t}^2 \\
&= C \cdot 2^{2(s-t)\frac{(1-T)J-n+1}{n-T}} \|u\|_{\mathcal{H}^t}^2 = C \cdot 2^{2(s-t)\frac{J}{n} \frac{1-T}{1-T/n}} \|u\|_{\mathcal{H}^t}^2. \quad (53)
\end{aligned}$$

In the last but one step we used that  $\max_{T|\mathbf{j}|_\infty - |\mathbf{j}|_1 < TJ - (n+J-1)} 2^{2(s-t)|\mathbf{j}|_\infty}$  takes its maximum at  $|\mathbf{j}|_\infty = \lceil \frac{(1-T)J-n+1}{n-T} \rceil$ . Compared to the result (44) for the space  $V_J^0$ , the order of approximation deteriorates in case  $T > 0$ , when changing from the space  $V_J^0$  to the space  $V_J^T$ , now by the factor  $\frac{1-T}{1-T/n}$ . For  $T < 0$  the order of approximation is improved by the factor  $\frac{1-T}{1-T/n}$ . Compared to the full grid  $V_J^\infty$  equation (53) indicates a loss in the order of approximation by the factor  $\frac{1-T}{n-T}$ . Note that for  $T = 0$  we regain estimate (44).

For  $u \in \mathcal{H}_{mix}^t$  we have (compare (48) and remember that  $\mathcal{H}_{mix}^t \subset \mathcal{H}^s$ )

$$\begin{aligned}
\inf_{v \in V_J^T} \|u - v\|_{\mathcal{H}^s}^2 &\leq \|u - \sum_{\mathbf{j} \in I_J^T} w_{\mathbf{j}}\|_{\mathcal{H}^s}^2 \stackrel{(35),(37)}{\leq} C \cdot \max_{\mathbf{j} \notin I_J^T} 2^{2s|\mathbf{j}|_\infty - 2t|\mathbf{j}|_1} \|u\|_{\mathcal{H}_{mix}^t}^2 \\
&= C \cdot \max_{T|\mathbf{j}|_\infty - |\mathbf{j}|_1 < TJ - (n+J-1)} 2^{2s|\mathbf{j}|_\infty - 2t|\mathbf{j}|_1} \|u\|_{\mathcal{H}_{mix}^t}^2. \quad (54)
\end{aligned}$$

It is straightforward to show that for  $T \geq \frac{s}{t}$  the maximum is obtained for  $|\mathbf{j}|_\infty = \lceil \frac{(1-T)J-n+1}{n-T} \rceil$ , and for  $T \leq \frac{s}{t}$  the maximum is obtained in  $\mathbf{j} = (J+1, 1, \dots, 1)$ . We continue (54) and have

$$\begin{aligned}
&\inf_{v \in V_J^T} \|u - v\|_{\mathcal{H}^s}^2 \\
&\leq \begin{cases} C \cdot 2^{2(s-nt)\frac{(1-T)J-n+1}{n-T}} \|u\|_{\mathcal{H}_{mix}^t}^2 = C \cdot 2^{2(s-t+(Tt-s)\frac{n-1}{n-T})J} \|u\|_{\mathcal{H}_{mix}^t}^2 & \text{for } T \geq \frac{s}{t} \\ C \cdot 2^{-2t(n-1)} 2^{2(s-t)J} \|u\|_{\mathcal{H}_{mix}^t}^2 = C \cdot 2^{2(s-t)J} \|u\|_{\mathcal{H}_{mix}^t}^2 & \text{for } T \leq \frac{s}{t}. \end{cases} \quad (55)
\end{aligned}$$

Note that for  $T = \frac{s}{t}$  both estimates give the same result.

The estimates from (55) show once more that for  $u \in \mathcal{H}_{mix}^t$  there appears no loss of asymptotic approximation power if the full grid is replaced by an optimized grid induced by the index set  $I_J^T$  with  $T \leq \frac{s}{t}$ . Note that  $I_J^T$  is of lower dimension than the index set  $I_J^\infty$  of the full grid. However, a further reduction of the number of grid points by using an index set  $I_J^T$  with  $T > \frac{s}{t}$  results in a deterioration of the order of approximation. In this case the order of approximation is reduced by  $(Tt - s)\frac{n-1}{n-T}$ .

Note that smoothness assumptions on the right hand side  $f$  in the variational problem (2) imply smoothness properties of the solution. Consider for example the case of a differential operator. Then for example  $f \in \mathcal{H}_{mix}^t$  implies  $u \in \mathcal{H}_{mix}^{t,s}$ . Therefore we now deal also with the more general case  $u \in \mathcal{H}_{mix}^{t,l}$ ,  $t+l \geq 0$ ,  $-\tilde{r} < s < t+l < r$ . More general choices of smoothness conditions can be handled in the same spirit required that norm equivalences for these spaces and associated bilinear forms are known. We summarize the discussion in a Theorem.

**Theorem 3** Let  $-\bar{r} < s < t + l < r, t + l \geq 0$  and  $0 \leq t < r$ . Then for  $u \in \mathcal{H}_{mix}^{t,l}$  it holds

$$\inf_{v \in V_J^T} \|u - v\|_{\mathcal{H}^s}^2 \leq \begin{cases} C \cdot 2^{2(s-l-t+(Tt-s+l)\frac{n-1}{n-T})J} \|u\|_{\mathcal{H}_{mix}^{t,l}}^2 & \text{for } T \geq \frac{s-l}{t} \\ C \cdot 2^{2(s-l-t)J} \|u\|_{\mathcal{H}_{mix}^{t,l}}^2 & \text{for } T \leq \frac{s-l}{t}. \end{cases} \quad (56)$$

Specifically for  $u \in \mathcal{H}^t = \mathcal{H}_{mix}^{0,t}$  it holds

$$\inf_{v \in V_J^T} \|u - v\|_{\mathcal{H}^s}^2 \leq C \cdot 2^{2(s-t)\frac{1-T}{n-T}J} \|u\|_{\mathcal{H}^t}^2 \quad (57)$$

and for  $u \in \mathcal{H}_{mix}^t = \mathcal{H}_{mix}^{t,0}$  it holds

$$\inf_{v \in V_J^T} \|u - v\|_{\mathcal{H}^s}^2 \leq \begin{cases} C \cdot 2^{2(s-t+(Tt-s)\frac{n-1}{n-T})J} \|u\|_{\mathcal{H}_{mix}^t}^2 & \text{for } T \geq \frac{s}{t} \\ C \cdot 2^{2(s-t)J} \|u\|_{\mathcal{H}_{mix}^t}^2 & \text{for } T \leq \frac{s}{t}. \end{cases} \quad (58)$$

**Proof:** Let  $u \in \mathcal{H}_{mix}^{t,l}$ . To proof (56) we use the upper estimate from the norm equivalence (35) and the lower estimate from (39). Then

$$\begin{aligned} \inf_{v \in V_J^T} \|u - v\|_{\mathcal{H}^s}^2 &\leq \|u - \sum_{\mathbf{j} \in I_J^T} w_{\mathbf{j}}\|_{\mathcal{H}^s}^2 \stackrel{(35)}{\approx} \sum_{\mathbf{j} \notin I_J^T} 2^{2s|\mathbf{j}|_{\infty}} \|w_{\mathbf{j}}\|_{\mathcal{L}^2}^2 \\ &\leq \max_{\mathbf{j} \notin I_J^T} 2^{2(s-l)|\mathbf{j}|_{\infty} - 2t|\mathbf{j}|_1} \cdot \sum_{\mathbf{j} \notin I_J^T} 2^{2l|\mathbf{j}|_{\infty} + 2t|\mathbf{j}|_1} \|w_{\mathbf{j}}\|_{\mathcal{L}^2}^2 \\ &\stackrel{(39)}{\leq} C \cdot \max_{\mathbf{j} \notin I_J^T} 2^{2(s-l)|\mathbf{j}|_{\infty} - 2t|\mathbf{j}|_1} \|u\|_{\mathcal{H}_{mix}^{t,l}}^2. \end{aligned}$$

Evaluating the maximum with respect to  $I_J^T$  shows (56). The inequalities (57) and (58) are special cases of the inequality (56), with  $t = 0$  and  $l = 0$ , respectively. See also inequalities (53) and (55) together with the above discussion.

□

Theorem 3 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  $T \leq (s-l)/t$ . The use of approximation spaces  $V_J^T$  with  $T > (s-l)/t$  leads to a deterioration of the optimal order of convergence. Hence, for purposes of discretization of large scale problems with solution in the space  $\mathcal{H}_{mix}^{t,l}$ , the spaces  $V_J^{(s-l)/t}$  with  $T \leq (s-l)/t$  are well suited. From the nestedness of the spaces  $V_J^T$  we conclude that the choice  $T = (s-l)/t$  will lead to most economical algorithms. This holds true especially in higher dimensions where the benefits of the spaces  $V_J^T$  with  $T$  large become most obvious, as we will see in section 4.2. There we deal with the dimension of the approximation spaces  $V_J^T$  and compare it with the dimension of the full grid space  $V_J^{-\infty}$ .

## 4.2 Dimension of the approximation spaces $V_J^T$

The following Lemma discusses the dimension of the spaces  $V_J^T$ . We split the basis functions into two sets. One with those functions that correspond to the interior of the unit cube and the other with those functions that correspond to the boundary. For ease of exposition we restrict ourselves to homogeneous Dirichlet boundary conditions in this section, that is, we count only

those basis functions/indices that correspond to the interior of the unit cube. Hence the index  $\mathbf{j}$  with minimal  $|\cdot|_\infty$  and  $|\cdot|_1$ -norm in an index set  $I_f^T$  is  $\mathbf{j} = (1, \dots, 1)$ . Note that other boundary conditions could be dealt with analogous reasoning. But we would have to count also indices  $\mathbf{j}$  with  $j_i = 0$  for some  $1 \leq i \leq n$ .

**Lemma 1** *It holds*

$$\dim(V_f^T) \leq \begin{cases} n \cdot 2^J & \text{for } T = 1, \\ \frac{n}{2} \left( \frac{1}{1-2^{-1/T-1}} \right)^n \cdot 2^J = O(2^J) & \text{for } 0 < T < 1, \\ O(2^{\frac{T-1}{T/n-1}J}) & \text{for } T < 0. \end{cases} \quad (59)$$

The case  $T = 0$  is covered by the estimate

$$\dim(V_f^T) \leq \left( \frac{J^{n-1}}{(n-1)!} + O(J^{n-2}) \right) \cdot 2^J = O(2^J J^{n-1}) \text{ for } 0 \leq T \leq 1/J. \quad (60)$$

**Proof:**

**The case  $T \geq 0$**

Let  $|\mathbf{j}|_1 = n + J - 1 - i$  and  $0 < T \leq 1$ . Then  $W_{\mathbf{j}} \subset V_f^T \Leftrightarrow -|\mathbf{j}|_1 + T|\mathbf{j}|_\infty \geq -(n + J - 1) + TJ \Leftrightarrow |\mathbf{j}|_\infty \geq J - \frac{1}{T}i$ . Since  $\sum_{|\mathbf{j}|_1 = n+J-1-i} 1 = \binom{|\mathbf{j}|_1-1}{n-1}$  and

$$\sum_{\substack{|\mathbf{j}|_1 = n+J-1-i, \\ |\mathbf{j}|_\infty = \ell_1 \geq J-i/T}} 1 \leq \binom{\lceil |\mathbf{j}|_1 - (J - i/T) \rceil}{n-1} = \binom{\lceil n-1 + (1/T-1)i \rceil}{n-1} \quad (61)$$

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

$$\begin{aligned} |V_f^T| &= \sum_{i=0}^{J-1} \sum_{\substack{|\mathbf{j}|_1 = n+J-1-i, \\ |\mathbf{j}|_\infty = \ell_1 \geq J-i/T}} |W_{\mathbf{j}}| \\ &\leq \sum_{i=0}^{J-1} 2^{J-1-i} n \binom{\lceil n-1 + (1/T-1)i \rceil}{n-1} \\ &= 2^{J-1} n \sum_{i=0}^{J-1} 2^{-i} \binom{\lceil n-1 + (1/T-1)i \rceil}{n-1}. \end{aligned} \quad (62)$$

In the case  $T - 1 = 0$  we get  $|V_f^T| \leq 2^{J-1} n \sum_{i=0}^{J-1} 2^{-i} \leq 2^J n$ , hence (59). For  $T < 1$  the substitution  $i \rightarrow \frac{i}{1/T-1}$  leads to

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

Since  $(x^{n-1+i})^{(n-1)} = \frac{(n-1+i)!}{i!} x^i \quad \forall x \in \mathbf{R}$ , we get

$$|V_f^T| \leq 2^{J-1} n \frac{1}{(n-1)!} \sum_{i=0}^{\lceil (1/T-1)(J-1) \rceil} (x^{n-1+i})^{(n-1)} \Big|_{x=2^{-\frac{1}{1/T-1}}}$$

$$\begin{aligned}
&= 2^{J-1} n \frac{1}{(n-1)!} \left( x^{n-1} \sum_{i=0}^{\lceil (1/T-1)(J-1) \rceil} x^i \right)^{(n-1)} \Big|_{x=2^{-\frac{1}{1/T-1}}} \\
&= 2^{J-1} n \frac{1}{(n-1)!} \left( x^{n-1} \frac{1-x^{\lceil (1/T-1)(J-1) \rceil + 1}}{1-x} \right)^{(n-1)} \Big|_{x=2^{-\frac{1}{1/T-1}}} \\
&= 2^{J-1} n \frac{1}{(n-1)!} \left[ \left( \frac{x^{n-1}}{1-x} \right)^{(n-1)} - \left( \frac{x^{\lceil (1/T-1)(J-1) \rceil + n}}{1-x} \right)^{(n-1)} \right] \Big|_{x=2^{-\frac{1}{1/T-1}}}. \tag{63}
\end{aligned}$$

Since

$$\begin{aligned}
\frac{1}{(n-1)!} \left( x^k \frac{1}{1-x} \right)^{(n-1)} &= \frac{1}{(n-1)!} \sum_{i=0}^{n-1} \binom{n-1}{i} (x^k)^{(i)} \left( \frac{1}{1-x} \right)^{(n-i-1)} \\
&= \frac{1}{(n-1)!} \sum_{i=0}^{n-1} \binom{n-1}{i} \frac{k!}{(k-i)!} x^{k-i} (n-1-i)! \left( \frac{1}{1-x} \right)^{n-i} \\
&= x^{k-n} \sum_{i=0}^{n-1} \binom{k}{i} \left( \frac{x}{1-x} \right)^{n-i} \tag{64}
\end{aligned}$$

we get

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

Hence we obtain (59).

To proof (60) let again  $|\mathbf{j}|_1 = n + J - 1 - i$  and  $T \leq 1/J$ . Then  $W_{\mathbf{j}} \subset V_J^T \Leftrightarrow -|\mathbf{j}|_1 + T|\mathbf{j}|_{\infty} \geq -(n + J - 1) + TJ \Leftrightarrow |\mathbf{j}|_{\infty} \geq 0$ . I.e., every  $W_{\mathbf{j}}$  with  $|\mathbf{j}|_1 \leq n + J - 1 - i$ , is in  $V_J^T$ . Hence

$$\begin{aligned}
|V_J^T| &= \sum_{|\mathbf{j}|_1 \leq n+J-1} |W_{\mathbf{j}}| = \sum_{i=0}^{J-1} 2^{J-1-i} \sum_{|\mathbf{j}|_1 = n+J-1-i} 1 \\
&= \sum_{i=0}^{J-1} 2^{J-1-i} \binom{n-1+J-1-i}{n-1} = \sum_{i=0}^{J-1} 2^i \binom{n-1+i}{n-1}.
\end{aligned}$$

This results in (see [BG98], proof of Lemma 7 for details)

$$\begin{aligned}
|V_J^T| &= (-1)^n + 2^J \sum_{i=0}^{n-1} \binom{n+J-1}{i} (-2)^{n-1-i} \\
&= \left( \frac{J^{n-1}}{(n-1)!} + O(J^{n-2}) \right) \cdot 2^J.
\end{aligned}$$

This completes the proof for the case  $T \geq 0$ .

**The case  $T < 0$**

Now we deal with the approximation spaces  $V_J^T$ ,  $T < 0$ . We introduce an auxiliary index set  $I_J^T$  with  $I_J^T \subset I_J^0$  given by

$$I_J^T = \{\mathbf{j} : -|\mathbf{j}|_1 + T|\mathbf{j}|_\infty \geq -(n + J - 1) + TJ/n\} \quad (65)$$

and the related approximation spaces

$$V_J^T := \bigoplus_{\mathbf{j} \in I_J^T} W_{\mathbf{j}}. \quad (66)$$

Note that  $I_J^T$  is just a shifted version of  $I_J^0$ . See Figure 7 for a schematical comparison of the index sets  $I_J^T$  and  $I_J^0$  in the 2D case.

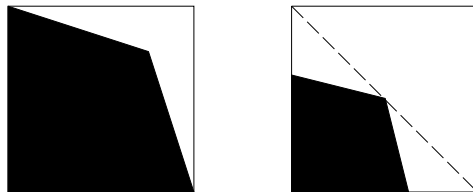


Figure 7: Schematical representation of  $I_J^T$  (left) and  $I_J^0$  (right) for  $T < 0$ , 2D case

Obviously  $\dim(I_J^T) = O(2^J)$ . Equation (47) shows that the order of approximation of the approximation space  $V_J^T$ ,  $T \leq \frac{s}{t}$ , for functions from  $\mathcal{H}_{mix}^t$  is the same as for the space  $V_J^0$ , i.e.  $O(2^{(s/n-t)J})$ . On the other hand we have from inequality (58) that the order of approximation of the space  $V_{\hat{J}}^{s/t}$  is  $O(2^{(s-t)\hat{J}})$ . This shows that  $O(2^{(s/n-t)J}) = O(2^{(s-t)\hat{J}})$  must hold. Hence we have that  $J = \frac{s-t}{s/n-t}\hat{J} + C$  and therefore  $\dim(V_J^{s/t}) = O(2^{\frac{s-t}{s/n-t}\hat{J}})$  and  $\dim(V_J^{s/t}) = O(2^{\frac{s-t}{s/n-t}J})$ . This completes the proof.

□

Note that the coefficient in the asymptotic estimate of the first inequality in (59) is unbounded for  $T \rightarrow 0$  whereas the coefficient in the estimate (60) remains bounded. Asymptotically, for  $T > 0$ , the estimate (59) is sharper than (61). However, for computationally relevant sizes of  $J$ , (61) might be sharper than (59) for  $T$  near 0. Similar results have been obtained recently in [BG98] for  $s = 1, t = 2$  and approximation spaces spanned by piecewise linear functions.

The estimates (59),(60) should be compared to the results for the full grid spaces  $V_J^{-\infty}$  with dimension  $\dim(V_J^{-\infty}) = (2^J - 1)^n$ . The first two estimates in (59) show that for  $T > 0$  the dependence of the dimension of the approximation space on the dimensionality  $n$  of the problem has been reduced from  $2^{nJ}$  to  $nC^n \cdot 2^J$ , with some constant  $C$  independent of  $n$  and  $J$ . Note that  $C$  is explicitly given by Lemma 1 for this case. For the case  $T < 0$  we have again that using the above spaces as finite element spaces in the Galerkin method leads to a significant reduction of the numbers of unknowns, and hence the number of entries in the stiffness matrices. Note that  $\dim(V_J^T) \ll \dim(V_J^{-\infty})$  for large  $n$  or large  $T$ . Here we did not compute the prefactors explicitly, as the asymptotic estimate depends itself on the dimension  $n$ .

Using the above spaces as finite element spaces in the Galerkin method leads to a significant reduction of the numbers of unknowns and hence the number of entries in the stiffness matrices.

Table 1: Number of interior grid points for various approximation spaces  $V_J^T$  in  $n = 2, 3$  and 4 dimensions

	$n = 2$		$n = 3$		$n = 4$	
	$J = 10$	$J = 20$	$J = 10$	$J = 20$	$J = 10$	$J = 20$
$V_J^{-\infty}$	$1.0465 \cdot 10^6$	$1.0995 \cdot 10^{12}$	$1.0706 \cdot 10^9$	$1.1529 \cdot 10^{18}$	$1.095 \cdot 10^{12}$	$1.2095 \cdot 10^{24}$
$V_J^{-1}$	$2.8673 \cdot 10^4$	$3.3135 \cdot 10^8$	$3.676 \cdot 10^5$	$1.9813 \cdot 10^{10}$	$2.9983 \cdot 10^6$	$5.4044 \cdot 10^{11}$
$V_J^{-1/2}$	$1.4337 \cdot 10^4$	$7.7595 \cdot 10^7$	$1.2698 \cdot 10^5$	$2.1349 \cdot 10^9$	$7.20897 \cdot 10^5$	$3.1113 \cdot 10^{10}$
$V_J^0$	$9.217 \cdot 10^3$	$1.9923 \cdot 10^7$	$4.71 \cdot 10^4$	$2.0028 \cdot 10^8$	$1.7818 \cdot 10^5$	$1.4145 \cdot 10^9$
$V_J^{1/4}$	$4.865 \cdot 10^3$	$7.0779 \cdot 10^6$	$1.6639 \cdot 10^4$	$3.8011 \cdot 10^7$	$4.5825 \cdot 10^4$	$1.6640 \cdot 10^8$
$V_J^{1/2}$	$3.585 \cdot 10^3$	$4.145 \cdot 10^6$	$8.895 \cdot 10^3$	$1.2095 \cdot 10^7$	$1.8561 \cdot 10^4$	$3.0724 \cdot 10^7$

Table 1 gives the dimension of the approximation spaces  $V_J^T$  for various  $T$ , levels  $J$  and dimensions  $n$ . Clearly, for higher values of  $n$  or  $J$  the full grid approximation space  $V_J^{-\infty}$  becomes impractical and the benefit of approximation spaces  $V_J^T$  with  $T$  around 0 can be seen. Figure 8 shows a graph of  $\dim(V_J^T)/\dim(V_J^1)$  for  $T \in \{-\infty, -2, -0.5, 0, 0.25\}$  against the number of levels  $J$ . Note that the dimension of  $V_J^T$  is scaled with the dimension of the space  $V_J^1$ . For higher values of  $J$  the merits of using the approximation spaces  $V_J^T$  with large  $T$  become obvious. Figure 9 shows a graph of the dimension of the approximation spaces  $V_J^T$ ,  $T \in \{-\infty, -0.5, 0, 0.25\}$  against the dimension  $n$ . Again, for higher values of  $n$ , the merits of the approximation spaces  $V_J^T$  with large  $T$  become obvious.

In summary, Theorem 3 and Lemma 1 show that for approximation problems with  $u \in \mathcal{H}_{mix}^{t,l}$ ,  $t+l \geq 0$ ,  $-\tilde{r} < s < t+l < r$ , the use of the approximation spaces  $V_J^T$  with  $T \leq (s-l)/t$  leads to a significant reduction of the number of degrees of freedom compared to the full grid, while the order of approximation remains the same as for the full grid. This will become even more clear in section 5 where we consider the overall complexity of solving the operator equations up to a prescribed tolerance.

### 4.3 Optimization procedures and subspace selection

In this section we present another way of obtaining the approximation spaces  $V_J^T$ . The idea is to explicitly use an optimization procedure to select subspaces. We describe this briefly in the following. See [BG98] for a longer discussion in the case of  $s = 1$  and basis functions of piecewise linear splines.

As we already noticed several times we have from the norm equivalence (35) and the ellipticity condition (3) together with the local stability (10) the two-sided estimates

$$a(u, u) \stackrel{(3)}{\approx} \|u\|_{\mathcal{H}^s}^2 \stackrel{(35)}{\approx} \sum_{\mathbf{j}} 2^{2s|\mathbf{j}|_\infty} \|w_{\mathbf{j}}\|_{\mathcal{L}^2}^2 \stackrel{(10)}{\approx} \sum_{\mathbf{j}} 2^{2s|\mathbf{j}|_\infty} \left( \sum_{m \in \tau_{\mathbf{j}}} \langle u, \tilde{\psi}_{\mathbf{j}\mathbf{m}} \rangle^2 \right) \text{ for } s \in (-\tilde{r}, r). \quad (67)$$

Following (67), the contribution of the subspace  $W_{\mathbf{j}}$  to  $a(u, u)$  is bounded by  $Work_{\mathbf{j}} \cdot C$ , where

$$Work_{\mathbf{j}} := 2^{2s|\mathbf{j}|_\infty} \|w_{\mathbf{j}}\|_{\mathcal{L}^2}^2 \approx 2^{2s|\mathbf{j}|_\infty} \sum_{m \in \tau_{\mathbf{j}}} \langle u, \tilde{\psi}_{\mathbf{j}\mathbf{m}} \rangle^2. \quad (68)$$



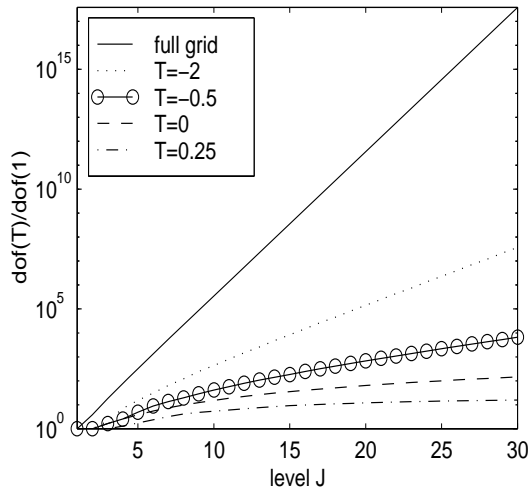


Figure 8: Semi-logarithmic plot of the dimension of the approximation space  $V_J^T$  scaled by the dimension of  $V_J^1$  for  $T \in \{-\infty, -2, -0.5, 0, 0.25\}$  against the number of levels  $J$  in three dimensions

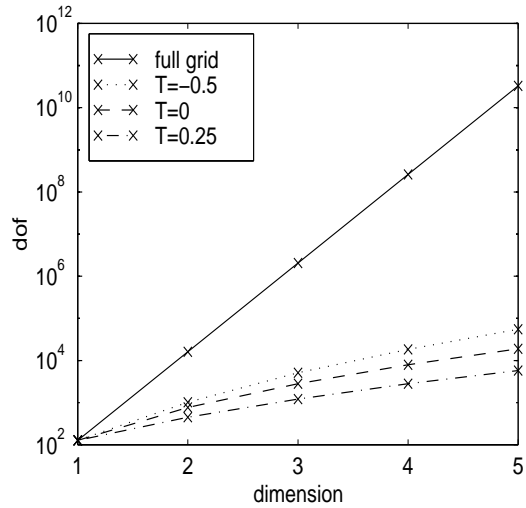


Figure 9: Semi-logarithmic plot of the dimension of the approximation spaces  $V_7^T$ ,  $T \in \{-\infty, -0.5, 0, 0.25\}$  against the dimension  $n$

Together with an upper estimate of  $\|w_{\mathbf{j}}\|_{\mathcal{L}^2}^2$  or of the coefficients  $\langle u, \tilde{\psi}_{\mathbf{j}\mathbf{m}} \rangle$ , the resulting upper estimate of  $Work_{\mathbf{j}}$  can be considered a measure of the profit gained in approximation power when  $W_{\mathbf{j}}$  is included into the approximation space. Note that such an estimate of  $\|w_{\mathbf{j}}\|_{\mathcal{L}^2}^2$  or an upper estimate of  $\langle u, \tilde{\psi}_{\mathbf{j}\mathbf{m}} \rangle$  by  $\|u\|_{\mathcal{H}_{mix}^{t,l}}^2$  can be obtained easily for elements of the considered smoothness classes by exploiting the vanishing moment condition on the dual wavelets  $\tilde{\psi}_{\mathbf{j}\mathbf{m}}$  (compare the Jackson inequality). Implicitly we used this several times in the last sections. On the other hand the inclusion of  $W_{\mathbf{j}}$  into the approximation space causes some cost in the discretization and hence in the solution procedure. The easiest measure for this cost is the dimension of the subspace  $|W_{\mathbf{j}}|$ . The task is now to find a grid (i.e. to select subspaces  $W_{\mathbf{j}}$ ) such that a given error bound gets minimal for some fixed cost, that is, the dimension of the approximation space is bounded by some given value  $b$ . This problem of deciding which subspaces should be included into the approximation space, given some prescribed overall cost can be reformulated as a classical binary knapsack problem. Restricting the range of possible subspaces  $W_{\mathbf{j}}$  to  $|\mathbf{j}|_{\infty} \leq J$  for some integer  $J$ , and arranging the possible indices  $\mathbf{j}$  in some linear order, the optimization problem reads as follows:

Find a binary vector  $y \in \{0, 1\}^{n \times J}$  such that

$$\sum_{|\mathbf{j}|_{\infty} \leq J} Work_{\mathbf{j}} \cdot y_{\mathbf{j}} \text{ constrained to } \sum_{|\mathbf{j}|_{\infty} \leq J} |W_{\mathbf{j}}| \cdot y_{\mathbf{j}} \leq b \quad (69)$$

is maximal.

Here the binary array  $y$  indicates which subspaces are to be included into the approximation space. See Figure 10 for an example of the binary array  $y$  in the case  $J = 3$  and two dimensions. Unfortunately such a binary knapsack problem is NP-hard. However, the situation changes,

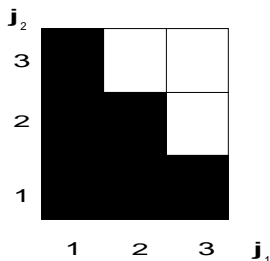


Figure 10: Example for the binary vector  $y$  in the case  $J = 3$ : Let the linear order be lexicographical from lower left to upper right. Then  $y = (1, 1, 1, 1, 1, 0, 1, 0, 0)$ .

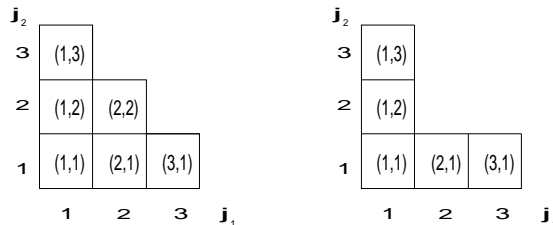


Figure 11: Example for  $J = 3$ : Depending on the profit/cost quotients different spaces  $W_{\mathbf{j}}$  are selected from the possible spaces allowed by the constraint  $\|\mathbf{j}\|_{\infty} \leq 3$ . On the left hand side the spaces with index  $\mathbf{j} \in \{(1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (3, 1)\}$  and on the right hand side the spaces with index  $\mathbf{j} \in \{(1, 1), (1, 2), (1, 3), (2, 1), (3, 1)\}$  are selected.

when we allow the array  $y$  to be a rational array in  $([0, 1] \cap \mathbb{Q})^{n \times J}$ . Then we know that the solution can be obtained by the following algorithm [MT90]:

1. Arrange the possible indices in some linear order such that  $\{Work_{\mathbf{j}}/|W_{\mathbf{j}}|\}_{\mathbf{j}}$  is decreasing in size, that is

$$\frac{Work_{i_1}}{|W_{i_1}|} \geq \frac{Work_{i_2}}{|W_{i_2}|} \geq \dots$$

2. Let  $M := \max\{i : \sum_{k=1}^i |W_{j_k}| \leq b\}$ .

3. The solution of the rational knapsack problem is given by

$$\begin{aligned} y_1 &= \dots = y_M = 1; \\ y_{M+1} &= \frac{b - \sum_{i=1}^M |W_{j_i}|}{|W_{j_{M+1}}|}; \\ y_{M+2} &= y_{M+3} = \dots = 0; \end{aligned}$$

Hence,  $y_{M+1}$  may be rational in  $[0, 1]$ . Therefore the solution of the rational knapsack problem is in general no solution of the binary knapsack problem. However, we still have the freedom of slightly changing the size of the cost  $b$ . We can do this in such a way that  $y_{M+1}$  is in  $\{0, 1\}$ . Then,  $y$  is also a solution of the corresponding binary knapsack problem. We refer to [BG98, Bu98] for more details and for a continuous analog of this discrete optimization procedure. The optimization process can thus be reduced to the discussion of the profit/cost quotients of the subspaces (or upper bounds of these)

$$\gamma_{\mathbf{j}} := \frac{Work_{\mathbf{j}}}{|W_{\mathbf{j}}|}. \quad (70)$$

That is, for an optimal grid in this sense one has to include those  $W_{\mathbf{j}}$  into the approximation space that have  $\gamma_{\mathbf{j}}$  bigger than some threshold. See Figure 11 for an example. Note that the optimization has to be performed with the use of upper bounds for  $Work_{\mathbf{j}}$  and not with the exact (but unknown) values. Hence, the optimization procedure is optimal only in this sense. Combining (68), (70) and using the moment condition on the dual wavelets together with the smoothness assumptions on the solution, we end up with the same grids  $V_f^T$  as in section 4.1.2.

## 5 Complexity estimates

In this section we deal with the complexity of solving the elliptic variational problem (2) up to some prescribed error when using the approximation spaces  $V_J^T$  and preconditioners arising from the norm equivalences from section 3, compare Remark 3. We consider the worst case setting, that is the error of an approximation  $u_{FE}$  from a finite element approximation space  $V_{FE}$  compared to the exact solution  $u$  is measured in the  $\mathcal{H}^s$ -norm. The cost of computing an approximation to the solution of the variational problem (2) can be divided into two parts, namely the cost to obtain the discrete system (6) and the cost to compute an approximate solution to this discrete system. The price for these two parts is often called informational and combinatory cost, respectively.

Note that due to the larger supports of the wavelets from coarser scales, the resulting stiffness matrices  $\mathcal{A}_J$  are rather densely populated. Here we have to distinguish two cases, namely integral and differential operators. In the case of integral operators  $\mathcal{A}_J$  is dense and thus has  $O(\dim(V_{FE})^2)$  entries. In the case of differential operators  $\mathcal{A}_J$  has  $O(\dim(V_{FE}) \ln(\dim(V_{FE}))^n)$  entries and is therefore much sparser than in the case of integral operators.

Lets take a closer look at the case of integral operators first. There are techniques to estimate the size of the entries in the stiffness matrix a-priori and to avoid the computation of entries below a prescribed threshold [DPS93, Sc98, GKS98]. Specifically, the authors showed in [GKS98] that it is worth-while to use compression also for optimized spaces, that is on top of dimension reduction. See [GOS98] for first numerical experiments regarding compression with respect to the single layer potential equation and approximation spaces built with the index sets  $I_J^0$ . Here we refrain from incorporating the effect of additional compression on the overall complexity as not to mix the effects of the use of the approximation spaces  $V_J^T$  and of compression. Note moreover that additional compression provides us with purely asymptotic estimates only, whereas the choice of optimized approximation spaces pays already for computationally relevant problem sizes especially in higher dimensions.

For differential operators it is important to note that one need not assemble the stiffness matrix, because all that is required in an iterative scheme is the application of the preconditioned stiffness matrix to a vector. Exploiting the pyramid structure of the multiscale transformations and the tensor-product structure of our wavelet basis functions the matrix-vector product can then be performed with  $O(\dim(V_{FE}))$  operations for example for differential operators with constant or separable coefficients, especially for the Laplace operator. The same holds true in the case of general coefficient functions on uniform grids, i.e. for the approximation space  $V_J^{-\infty}$ . Note however that the implementation of the matrix-vector product with linear complexity is a very involved and delicate task. In the case of adaptively refined grids or differential equations with more general coefficient functions and approximation spaces  $V_J^T \neq V_J^{-\infty}$  it is unclear whether the matrix-vector product can be carried out with a number of operations that scales linearly with the number of unknowns. In the following we assume that the matrix-vector product can be performed with  $O(\dim(V_{FE}))$  operations. For all other cases we get additional logarithmic terms in the overall complexity.

In order to be able to make precise statements about the complexity of the algorithms we propose, we have to describe our framework more carefully. We have to specify

- the information on the right hand side  $f$  and the bilinear form  $a(., .)$  that is allowed, and
- the model of computation.

Concerning the model of computation we assume for simplicity that comparison of real numbers as well as standard arithmetic operations, such as addition and scalar multiplication and division,

are performed with unit cost.

Concerning the permissible information there are two approaches used in the theory of information based complexity, namely linear and standard information [TWW88, We91]. Assuming continuous linear information means that any continuous linear functional is permissible information. In contrast, standard information of cardinality  $R$  means that the only information available consists of the result of  $R$  distinct samples. As an example the information about  $f$  would be  $f(x_1), f(x_2), \dots, f(x_R)$  at  $R$  different points  $x_i \in [0, 1]^n$ . Clearly standard information makes sense only for (at least) continuous functions  $f$ .

In the following we assume that arbitrary continuous linear information is permissible, i.e. we assume that the stiffness matrix as well as the load vector have been computed exactly (or at least with sufficient accuracy).

Once the stiffness matrix and the load vector have been computed, we are left with the issue of proposing an algorithm for the approximate solution of the discrete problem. We discuss an algorithm whose complexity is  $O(\dim(V_{FE}))$  for differential and  $O(\dim(V_{FE})^2)$  for integral operators.

Concerning the computational cost we mentioned already in Remark 3 that a simple diagonal scaling of the stiffness matrix is enough to obtain optimal preconditioning, if the related norm equivalences hold. This allows to construct solvers with a complexity of the number of entries in the stiffness matrix. To be a bit more precise, let us estimate the cost to solve (41) up to discretization error  $\epsilon$ , which is of order  $O(2^{-cJ})$ , with some  $c > 0$  depending on the order of approximation of the wavelet basis. From (40) in section 3 we have that the preconditioned (diagonally scaled) Galerkin stiffness matrix  $\{2^{-s|l+l'|} a(\psi_{\mathbf{l}\mathbf{k}}, \psi_{\mathbf{l}'\mathbf{k}'})\}_{\mathbf{l}, \mathbf{l}', \mathbf{k}, \mathbf{k}'}$  has a condition number which is bounded independent of the number of levels involved. Hence, the convergence rate  $\xi$  of gradient methods is independent of the dimension of the finite element approximation space  $V_{FE}$  if the stiffness matrix is symmetric. Applied to the preconditioned system, the initial error is reduced at least by the factor  $\xi$  in every iteration step and the number of iterations necessary to obtain an approximation within the prescribed accuracy is then  $|\log_{\xi}(\epsilon)| = cJ$ . Hence, the overall  $\epsilon$ -complexity of computing a solution of the variational problem (2) within discretization accuracy  $\epsilon$  is  $O(J \cdot \dim(V_{FE})^2)$  if the stiffness matrix is dense and  $O(J \cdot \dim(V_{FE}))$  if the matrix-vector product can be performed with  $O(\dim(V_{FE}))$  operations. Note that it is possible to get rid of the  $J$ -term in the complexity estimate by embedding the solver in a nested iteration scheme [KD71]. The idea is to compute a suitable start value by first applying some iteration steps to the problem on a coarser level and to apply this procedure recursively starting from the coarsest level. This makes the optimized spaces defined in section 4 good candidates for the approximation space  $V_{FE}$  provided that the required regularity assumptions on the solution of the variational problem hold.

To obtain an approximation of the exact solution that has an error of  $O(\epsilon)$  in the energy-norm the number of levels  $J$  has to be chosen such that the approximation error is smaller than  $O(\epsilon)$ . Combining the results about the approximation error from Theorem 3 with the estimate of the dimension of the space  $V_J^{(s-l)/t}$  in section 4.2 gives us the  $\epsilon$ -complexities. Tables 2 and 3 summarize the above discussion. There the complexity of solving the problem (2) in the space  $\mathcal{H}_{mix}^{t,l}$  up to an error of the order of  $\epsilon$  is given for positive and for negative smoothness parameters  $s$ . Tables 2 and 3 show that for problems with solution  $u \in \mathcal{H}_{mix}^{t,l}$  and optimized approximation spaces, the asymptotic complexity is independent of the dimension  $n$  if  $s - l > 0$ . For fixed dimension  $n$  and  $1 - (s - l)/t < n$ , the complexity is in favor of the approximation space  $V_J^{(s-l)/t}$  also for the case  $s - l < 0$ .

Note that the complexities for integral operators in Table 3 are not yet optimal, as we made no use of the potential of further compression of the stiffness matrix [GKS98].

Table 2: Complexity of solving an  $\mathcal{H}^s$ -elliptic variational problem with a differential operator up to an error of  $O(\epsilon)$  under the constraint that the solution is in  $\mathcal{H}_{mix}^{t,l}$  for the approximation spaces  $V_J^{(s-l)/t}$  and  $V_J^{-\infty}$ ; we assume that the matrix-vector product can always be carried out with  $O(\dim(V_J^{\frac{s-l}{t}}))$  or  $O(\dim(V_J^{-\infty}))$  operations, respectively; the error is measured in the  $\mathcal{H}^s$ -norm;

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

Table 3: Complexity of solving an  $\mathcal{H}^s$ -elliptic variational problem with an integral operator up to an error of  $O(\epsilon)$  under the constraint that the solution is in  $\mathcal{H}_{mix}^{t,l}$  for the approximation spaces  $V_J^{(s-l)/t}$  and  $V_J^{-\infty}$ ; the error is measured in the  $\mathcal{H}^s$ -norm;

	$V_J^{(s-l)/t}$	$V_J^{-\infty}$
$s > l$	$O(\epsilon^{\frac{2}{s-l-t}})$	$O(\epsilon^{\frac{2n}{s-l-t}})$
$s = l$	$O\left(\epsilon^{-\frac{2}{t}} \left(\ln(\epsilon^{-\frac{1}{t}})\right)^{2(n-1)}\right)$	$O(\epsilon^{-\frac{2n}{t}})$
$s < l$	$O(\epsilon^{\frac{2}{(s-l)/n-t}})$	$O(\epsilon^{\frac{2n}{s-l-t}})$

Concerning the case of standard information, note that standard information is a proper subset of continuous linear information [We95]. Hence the problem complexity will be larger when standard information is permissible.

## 6 Applications

In this section we give examples for applications of the above ideas. We deal with the Laplace problem with homogeneous Dirichlet boundary conditions and with the screen problem. These are two prominent examples of the class of elliptic problems that show the conceptual ideas and may therefore serve as a guideline for dealing with other elliptic variational problems. First of all, we are looking for candidates of univariate wavelet bases that fulfill our requirements. Note that because of our tensor-product ansatz we can reduce the questions to the one-dimensional case. Specifically, those basis functions whose support intersects with the boundary have to fulfill special boundary conditions. We do not aim at a general framework for this matter. Therefore we do not take up this issue here, but refer the interested reader to the literature and state that these problems can be settled. See [AHJP94, DKU96, CDV93, CW92] for appropriate constructions of localized functions and their boundary adaptation.

Sobolev spaces of interest for the study of integral and differential equations on the  $n$ -dimensional

unit square  $I^n = [0, 1]^n$  are defined by

$$\mathcal{H}^s(I^n) = \{f \in \mathcal{D}'(I^n) : \exists g \in \mathcal{H}^s(\mathbb{R}^n) : g|_{I^n} = f \text{ and } \|f\|_{\mathcal{H}^s(I^n)} = \inf_{f=g|_{I^n}} \|g\|_{\mathcal{H}^s(\mathbb{R}^n)}\} \quad (71)$$

and

$$\tilde{\mathcal{H}}^s(I^n) = \{f = g|_{I^n} : g \in \mathcal{H}^s(\mathbb{R}^n) \text{ and } \text{supp } g \subset I^n\} \quad (72)$$

equipped with the norm

$$\|f\|_{\tilde{\mathcal{H}}^s(I^n)} = \|g\|_{\mathcal{H}^s(\mathbb{R}^n)},$$

see [Tr78, LM72], e.g. The interpolation spaces  $\mathcal{H}^s(I^n)$  and  $\tilde{\mathcal{H}}^s(I^n)$  are dual to each other, i.e.

$$(\mathcal{H}^s(I^n))' = \tilde{\mathcal{H}}^{-s}(I^n), \quad (\tilde{\mathcal{H}}^s(I^n))' = \mathcal{H}^{-s}(I^n), \quad -\infty < s < \infty. \quad (73)$$

Furthermore

$$\tilde{\mathcal{H}}^s(I^n) = \mathcal{H}_0^s(I^n) \equiv \text{clos}_{\mathcal{H}^s(I^n)} \mathcal{C}_0^\infty(I^n) \text{ for } s > \frac{1}{2}, s \neq k + \frac{1}{2}, k \in \mathbb{N},$$

i.e.  $\tilde{\mathcal{H}}^s(I^n)$  is the appropriate space for problems with homogeneous essential boundary conditions and  $\tilde{\mathcal{H}}^s(I^n) = \mathcal{H}^s$  for  $-1/2 < s < 1/2$ . Which of these spaces is appropriate depends on the application. For example  $\tilde{\mathcal{H}}^1(I^n) = \mathcal{H}_0^1(I^n)$  is the appropriate space for the Laplace problem with homogeneous essential boundary conditions. For the screen problem the space  $\tilde{\mathcal{H}}^{1/2}$  is appropriate.

Sobolev spaces of functions in other spaces of interest, like those with dominating mixed derivative on  $I^n$  are analogously defined. For example we have

$$\mathcal{H}_{mix}^s(I^n) := \mathcal{H}^s(I) \otimes \dots \otimes \mathcal{H}^s(I)$$

and

$$\tilde{\mathcal{H}}_{mix}^s(I^n) := \tilde{\mathcal{H}}^s(I) \otimes \dots \otimes \tilde{\mathcal{H}}^s(I).$$

To be able to repeat the above reasoning function spaces fulfilling the required boundary conditions and a Jackson and a Bernstein inequality have to be constructed. Then the argumentation of section 4 can be repeated with obvious modifications.

Here, we concentrate on semi-orthogonal linear spline wavelets (prewavelets) on uniform dyadic grids as introduced in [CW92]. Figure 12 shows a prewavelet in the interior of the domain. Concerning our cases of interest, suitable boundary constructions have been given for example in [Au92, CDJV93, GO95] and [GOS98], respectively.

**Example 1:** The Poisson equation

We consider the Poisson problem

$$\Delta u = f \quad (74)$$

with Dirichlet boundary conditions in its variational form on  $\mathcal{H}_0^1(I^n)$ . In this case we have  $s = 1$ . Estimates of the  $\epsilon$ -complexities of solving (74) for  $u \in \mathcal{H}_{mix}^2$  and continuous linear information have been given in [BG98]. The authors constructed a finite element method using tensor-products of piecewise linear splines and index sets that are asymptotically equal to  $I_J^{2/5}$ . They proposed the use of a multilevel method to solve the resulting discrete problems. The resulting overall complexity is then

$$O(\epsilon^{-1}) \quad (75)$$

because of the optimality of the proposed multilevel method and  $\dim(I_J^{2/5}) = O(2^J)$ . Estimates of the  $\epsilon$ -complexities of solving (74)  $f \in \mathcal{H}_{mix}^t$  and standard information can be found in [We95].

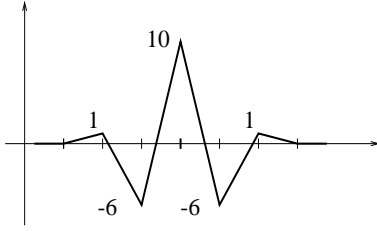


Figure 12: Semi-orthogonal linear spline prewavelet

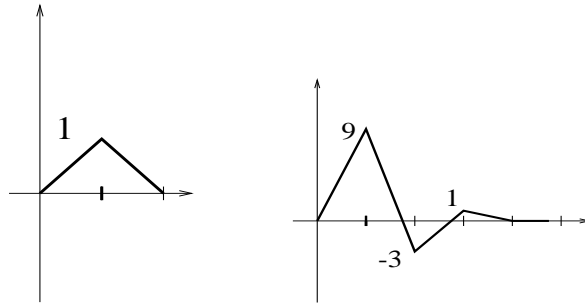


Figure 13: Nodal basis function  $\psi_{1,0}$  corresponding to the coarsest level  $W_1$  (left) and boundary wavelet for the left boundary for homogeneous Dirichlet boundary conditions (right)

The author proposed a finite element method on the index set  $I_j^0$  and showed that the solution can be obtained with a complexity of

$$O\left(\epsilon^{-\frac{1}{t}} \left(\ln(\epsilon^{-1})\right)^{(1+1/t)(n+1)}\right). \quad (76)$$

Note that  $f \in \mathcal{H}_{mix}^t$  is a stronger assumption than  $u \in \mathcal{H}_{mix}^t$ . Therefore the complexity (75) for the case  $u \in \mathcal{H}_{mix}^2$  and continuous linear information and the complexity (76) cannot be compared directly.

Let us discuss our method. A suitable choice for the boundary prewavelets is shown in Figure 17. The basis function assigned to the coarsest level is the usual nodal basis function, see Figure 13. The orthogonal complement spaces  $W_j, j \geq 2$ , are spanned by scaled and delayed versions of the functions shown in Figure 12 for the interior grid points and Figure 13 for the left boundary and an analogous construction for the right boundary. The resulting multilevel system incorporating homogeneous Dirichlet boundary conditions is a semi-orthogonal Riesz basis in  $\mathcal{H}_0^s(I)$  for  $0 \leq s < 3/2$ .

The resulting approximation spaces can be visualized by associating a grid-point with the midpoint of the support of every basis function. See Figure 14 for a visualization of different basis functions and Figure 15 for the representation of the full grid corresponding to the index set  $I_3^{-\infty}$  and the grid corresponding to the index set  $I_3^0$ .

We assume that the solution of the variational problem is in the space  $\mathcal{H}_{mix}^{t,l}$  for some parameters  $t, l$  with  $t + l \geq 1$ . From Table 2 we take the following  $\epsilon$ -complexities.

$$comp(\epsilon) \leq \begin{cases} O(\epsilon^{\frac{1}{1-t-t}}) & \text{for } l < 1, \\ O\left(\epsilon^{-\frac{1}{t}} \left(\ln(\epsilon^{-\frac{1}{t}})\right)^{n-1}\right) & \text{for } l = 1, \\ O(\epsilon^{\frac{1}{(1-l)/n-t}}) & \text{for } l > 1. \end{cases}$$

Specifically for the cases  $u \in \mathcal{H}^2 = \mathcal{H}_{mix}^{0,2}, \mathcal{H}_{mix}^2 = \mathcal{H}_{mix}^{2,0}$  and  $\mathcal{H}_{mix}^{1,1}$  we obtain the complexities

$$comp(\epsilon) \leq \begin{cases} O(\epsilon^{-1}) & \text{for } u \in \mathcal{H}_{mix}^2, \\ O\left(\epsilon^{-1} \left(\ln(\epsilon^{-1})\right)^{n-1}\right) & \text{for } u \in \mathcal{H}_{mix}^{1,1}, \\ O(\epsilon^{-n}) & \text{for } u \in \mathcal{H}^2. \end{cases}$$

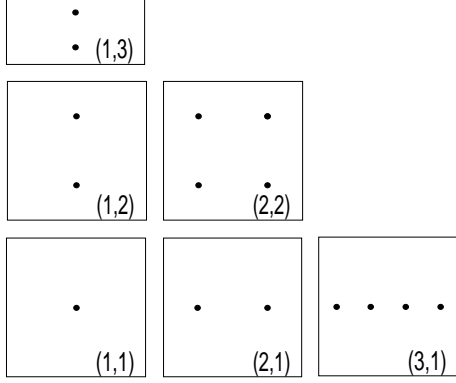


Figure 14: Grid point sets with corresponding indices  $\mathbf{j} = (j_1, j_2)$  from the index set  $I_j^0$

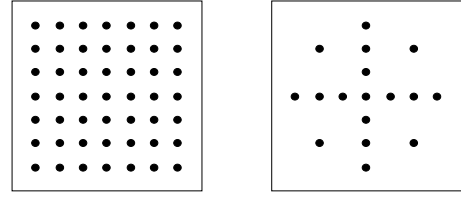


Figure 15: Grid point sets of grid  $I_3^{-\infty}$  (left) and  $I_3^0$  (right)

Hence we regain the result (75) of [BG98] as a special case. It is interesting to note that for  $u \in \mathcal{H}_{mix}^{t,1}$  the resulting optimized approximation space is  $V_j^0$ . Hence the complexity is

$$O\left(\epsilon^{-\frac{1}{t}}(\ln(\epsilon^{-\frac{1}{t}}))^{n-1}\right). \quad (77)$$

This result should be compared to the result (76) from [We95] for the case of standard information. A short computation shows that for  $t \geq 1$  the asymptotic complexity of the case with standard information (76) is larger than in the case of continuous linear information (77).

**Example 2:** Single layer potential equation

The second example we consider is the single layer potential equation

$$\frac{1}{c} \int_{I^n} \frac{g(y)}{|x-y|} dy = f(x)$$

in its variational form on  $\tilde{\mathcal{H}}^{-1/2}(I^n)$ . Here we have  $s = -\frac{1}{2}$ . The corresponding bilinear form

$$a(u, v) = \left(\frac{1}{c} \int_{I^n} \frac{g(y)}{|x-y|} dy, v\right)_{\mathcal{H}^{1/2} \times \tilde{\mathcal{H}}^{-1/2}}, \quad u, v \in \tilde{\mathcal{H}}^{-1/2}(I^n),$$

is symmetric and  $\tilde{\mathcal{H}}^{-1/2}$ -elliptic.

For problems in  $\tilde{\mathcal{H}}^{-\frac{1}{2}}$  the basis does not have to fulfill special boundary conditions. The bases for  $W_0$  and  $W_1$  are shown in Figure 16. The orthogonal complement spaces  $W_j, j \geq 2$  are spanned by scaled and delated versions of the functions shown in Figure 12 for the interior grid points and Figure 17 for the left boundary and an analogous construction for the right boundary. The resulting multilevel system is a semi-orthogonal Riesz basis in  $\mathcal{H}^s(I)$  for  $0 \leq s < 3/2$ , and for  $\tilde{\mathcal{H}}^s(I)$  for  $-3/2 < s < 0$ . Hence this example is fully covered by the theory of sections 3 and 4. Especially the preconditioning and approximation results and the complexity estimates of section 5 can be applied. Regularity theory for the screen problem shows that if the right hand side vector  $f$  is smooth enough, then the solution  $u$  can be decomposed into a regular part  $u^{reg}$  and a singular part due to corner and edge singularities, see [HMS94], compare also [GOS98]. Here we restrict ourselves to an approximation of the regular part of the solution. For a treatment of the singular parts see [GOS98, O98].



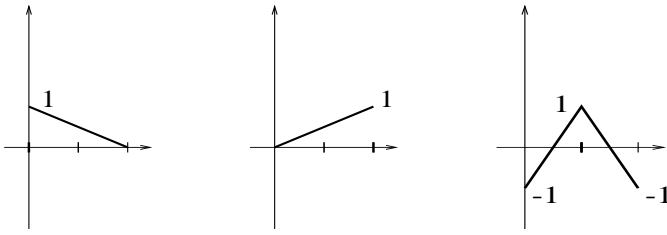


Figure 16: Basis functions  $\psi_{0,0}$ ,  $\psi_{0,1}$  and  $\psi_{1,0}$  corresponding to the spaces  $W_0$  and  $W_1$

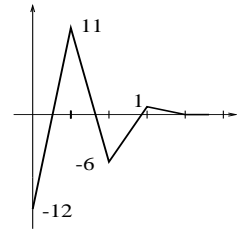


Figure 17: Boundary wavelet for the left boundary

Hence we assume that the solution of the variational problem is in the space  $\tilde{\mathcal{H}}_{mix}^{t,l}$  for some parameters  $t, l$  with  $t + l \geq 1$ . From Table 3 we take the following  $\epsilon$ -complexities.

$$comp(\epsilon) \leq \begin{cases} O(\epsilon^{-\frac{2}{1/2+t+l}}) & \text{for } l < -1/2, \\ O\left(\epsilon^{-\frac{2}{l}} \left(\ln(\epsilon^{-\frac{1}{l}})\right)^{2(n-1)}\right) & \text{for } l = -1/2, \\ O(\epsilon^{-\frac{2}{(1/2+l)/n+l}}) & \text{for } l > -1/2. \end{cases}$$

Note that a further reduction of the complexity can be achieved by compression strategies as described in [GKS98].

## 7 Concluding remarks

In this paper we constructed Finite Element approximation spaces for elliptic variational problems with solution in  $\mathcal{H}_{mix}^{t,l}$ ,  $r > t+l > s > -\tilde{r}$ ,  $t+l \geq 0$ ,  $0 \leq t < r$ . We gave complexity estimates for the case of continuous linear information. We showed these results in a constructive manner by proposing a finite element method together with optimal preconditioning. Specifically, we identified smoothness assumptions that make it possible to choose the Finite Element approximation space in such a way that the number of grid points is  $O(2^J)$  compared to  $O(2^{nJ})$  for the full grid space, while keeping the optimal order of approximation.

We like to give some hints on extensions to problems with large ellipticity constants and non-stable splittings.

- For problems with large ellipticity constants, the constants in the estimates of the approximation error do become large and poison the behavior of the approximation in actual implementations as the constants may dominate the error approximation for practical problem sizes. In these cases, the asymptotic estimates do not provide full insight into the behavior of the approximants. It is 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. As a simple example consider the anisotropic elliptic problem

$$-\sum_{i=1}^n d_i \frac{\partial^2}{\partial^2 x_i} u = f, d_i > 0, \quad (78)$$

in its variational form on  $\mathcal{H}^1(T^n)$ . Tensor-product approximation spaces are well suited for such problems as they allow easily for anisotropic refinement. Let  $a(.,.)$  denote the corresponding

$\mathcal{H}^1$ -elliptic variational form. The problem with the numerical solution of (78) is that the condition number of the Galerkin stiffness matrix on a isotropic full grid is linearly dependent on  $\max_{1 \leq i \leq n}(d_i) / \min_{1 \leq i \leq n}(d_i)$ . The same is true for the coefficient in the asymptotic estimate of the approximation error. Hence, for fixed refinement level  $J$  and varying coefficients  $d_i$ , the convergence rate of iterative methods as well as the error of approximation depend on  $d_i$ . For problems with large anisotropies this leads to a slow down of convergence and a deterioration of approximation. It is well known that some kind of semi-coarsening in the subspace splittings or in the construction of the approximation spaces can remedy these problems. These ideas can also be used for the approximation spaces defined here. It amounts to the use of a norm estimate on  $a(\cdot, \cdot)^{\frac{1}{2}}$  instead of  $\|\cdot\|_{\mathcal{H}^1}$ .

A consideration analogously to that in the Proof of Theorem 1 shows (use again Propositions 1 and 2)

$$a(u, u) \approx \sum_{\mathbf{j}} \left( \sum_{i=1}^n d_i 2^{2j_i} \right) \|w_{\mathbf{j}}\|_{\mathcal{L}^2}^2 \approx \sum_{\mathbf{j}} \max_{1 \leq i \leq n} (d_i 2^{2j_i}) \|w_{\mathbf{j}}\|_{\mathcal{L}^2}^2 \quad (79)$$

for  $u \in \mathcal{H}^1, u = \sum_{\mathbf{j}} w_{\mathbf{j}}$ . (See [GO95] for a proof in the case of prewavelets. There this norm estimate was used to obtain robust preconditioners for anisotropic problems.) Compared to the norm equivalence (35) (set  $s = 1$ ) the weight  $2^{2|\mathbf{j}|_{\infty}}$  is substituted by the weight  $\max_{1 \leq i \leq n} (d_i 2^{2j_i})$  including information from the anisotropy. Let  $u \in \mathcal{H}_{mix}^{t,l}, r > t + l \geq 1, 0 \leq t < r, I_{\mathbf{j}}$  a subset of  $I_J^{-\infty}$  and  $V_{FE}$  the corresponding approximation space. Then (79) together with (39) shows

$$\inf_{v \in V_{FE}} a(u - v, u - v) \leq C \cdot \max_{\mathbf{j} \notin I_{\mathbf{j}}} \left( \max_{1 \leq i \leq n} (d_i 2^{2j_i}) \cdot 2^{-2t|\mathbf{j}|_{\infty} - 2t|\mathbf{j}|_1} \right) \|u\|_{\mathcal{H}_{mix}^{t,l}}^2. \quad (80)$$

Without loss of generality we may assume that

$$d_1 = \operatorname{argmax}_{1 \leq i \leq n} \{d_i\} \text{ and } d_n = \operatorname{argmin}_{1 \leq i \leq n} \{d_i\}.$$

Fixing  $(J, 1, \dots, 1)$  to be the index with maximal  $|\cdot|_{\infty}$ -norm to be included into the index sets leads to  $c = \frac{1}{2t} \ln_2(d_1 2^{2J}) - n + 1 - \frac{1}{t}J$  and the index sets

$$I_J^{l,t,d} := \left\{ \mathbf{j} : -|\mathbf{j}|_1 - \frac{l}{t}|\mathbf{j}|_{\infty} + \frac{1}{2t} \ln_2 \left( \max_{1 \leq i \leq n} \left( \frac{d_i}{d_1} 2^{2(j_i - 2J)} \right) \right) \geq -(n + J - 1) - \frac{l}{t}J \right\}$$

where the index  $d$  indicates the dependence on the parameters  $d_i, 1 \leq i \leq n$ . Fixing  $(1, \dots, 1, J)$  to be the index with maximal  $n$ -th component to be included into the index sets leads to  $c = \frac{1}{2t} \ln_2(d_n 2^{2J}) - n + 1 - \frac{1}{t}J$  and the index sets

$$\widehat{I}_J^{l,t,d} := \left\{ \mathbf{j} : -|\mathbf{j}|_1 - \frac{l}{t}|\mathbf{j}|_{\infty} + \frac{1}{2t} \ln_2 \left( \max_{1 \leq i \leq n} \left( \frac{d_i}{d_n} 2^{2(j_i - 2J)} \right) \right) \geq -(n + J - 1) - \frac{l}{t}J \right\}$$

Then the corresponding approximation spaces keep the order of approximation of the full grid approximation space. Estimates on the dimension and the order of approximation can be derived in the spirit of the preceding sections. We obtain the same orders of approximation as for the approximation spaces  $V_J^{(1-l)/t}$  from the preceding section but with different coefficients. Note that in the case of the index set  $I_J^{l,t,d}$  the coefficient is dependent on  $d_1 = \operatorname{argmax}_{1 \leq i \leq n} \{d_i\}$  and the number of unknowns is reduced further, as  $I_J^{l,t,d} \subset I_J^{(1-l)/t}$ . For the case  $\widehat{I}_J^{l,t,d}$  the number of unknowns is increased compared to  $I_J^{(1-l)/t}$ , but the coefficient is only depending on  $d_n = \operatorname{argmin}_{1 \leq i \leq n} \{d_i\}$ . The norm equivalence (79) leads also to robust preconditioners,

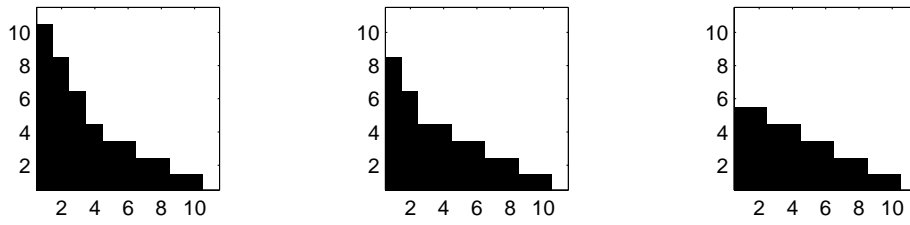


Figure 18: Index sets  $I_{10}^{2,0,d}$  for  $d_1/d_2 = 1$ ,  $d_1/d_2 = 10$  and  $d_1/d_2 = 1000$  from left to right, two dimensional case

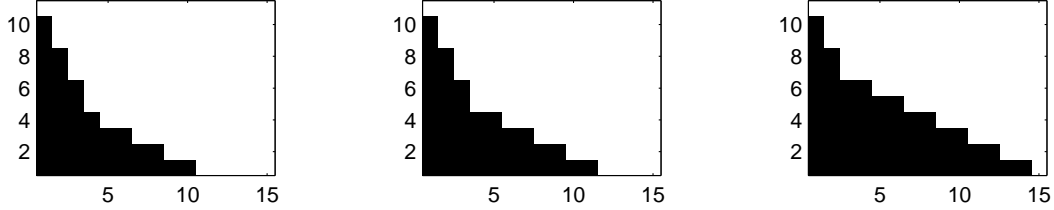


Figure 19: Index sets  $\hat{I}_{10}^{2,0,d}$  for  $d_1/d_2 = 1$ ,  $d_1/d_2 = 10$  and  $d_1/d_2 = 1000$  from left to right, two dimensional case

compare [GO95]. In the case of extreme anisotropy, the resulting grid consists of extremely stretched grids in the direction of the anisotropy, corresponding to semicoarsening. Figures 18 and 19 show some examples in two dimensions.

- To be able to apply the theory presented here, the resulting algorithms need to be complemented with some special treatment of singular parts of the solution of the variational problem. The idea is that a few wavelets of high level clustered around the singularity will suffice, while the optimized grids of the above sections are enough to treat the smooth parts of the solution. Solutions of elliptic boundary value problems in non-smooth domains mostly exhibit a singular behavior at the boundary. This also holds in the case of integral equations. Moreover, singular right hand sides or singularly perturbed operators may also be a source of singularities in the solution. It is well known that the rate of convergence of the Galerkin method is governed by the regularity of the solution. Singularities reduce the regularity and hence impair convergence. This holds particularly true for the discretizations constructed here, where additional regularity is required. First experiments in [GOS98] indicate that these tensor-product constructions are well suited for the approximation of edge singularities, when refinement towards the edges is used.

Here, it is helpful to refine the selection criteria to the atomic level, i.e., to allow for single basis functions/grid points to be selected. From (70) together with (68) we obtain the profit/cost quotient of a single basis function

$$\gamma_{\mathbf{j}\mathbf{l}} := 2^{2s|\mathbf{j}\mathbf{l}|_\infty} \langle u, \tilde{\psi}_{\mathbf{j}\mathbf{l}} \rangle^2. \quad (81)$$

Suppose for example the leading singularity component  $\chi$  of the solution  $u$  is known. Decomposing  $\chi$  with respect to the given basis, we can use the weights  $|\langle \chi, \tilde{\psi}_{\mathbf{j}\mathbf{l}} \rangle|$  in (81) instead of the weights  $|\langle u, \tilde{\psi}_{\mathbf{j}\mathbf{l}} \rangle|$ . This leads to the definition of grids adapted to  $\chi$  by choosing those indices that have

$$\gamma_{\mathbf{j}\mathbf{l}} := 2^{2s|\mathbf{j}\mathbf{l}|_\infty} \langle \chi, \tilde{\psi}_{\mathbf{j}\mathbf{l}} \rangle^2 \quad (82)$$

bigger than some threshold. This a-priori adaptivity should lead to a relatively high degree of adaptivity without complicated mesh refinement strategies especially for problems in higher dimensions. Nevertheless, for singularly perturbed problems with large ellipticity constants and problems that exhibit boundary singularities, a-posteriori adaptivity is still necessary. We refer to [Dahm97, DDD97] and [DeV98] for very promising results on nonlinear approximation and adaptivity.

- 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 as well. Specifically, the above construction of optimized grids is not limited to stable multilevel splittings, that is to multilevel finite element spaces that possess norm equivalences like those described in section 4.

For the construction of optimized grids the validity of norm equivalences like (35) and (37) is therefore not really essential. Keeping this in mind, it is possible, to obtain optimized grids also for non stable splittings. Consider for example the important case of an  $\mathcal{H}^1$ -elliptic operator and multiscale basis functions of tensor-products of piecewise linear splines  $\phi_{\mathbf{j}\mathbf{k}}$ . Let  $W_{\mathbf{j}} = \text{span}\{\phi_{\mathbf{j}\mathbf{k}}, \mathbf{k} \in \tau_{\mathbf{j}}\}$  denote the hierarchical difference space between two successive spaces spanned by piecewise  $n$ -linear functions. It is easy to see that in this case there holds a Bernstein inequality

$$\|w_{\mathbf{j}}\|_{\mathcal{H}^1} \leq C2^{|\mathbf{j}|_{\infty}} \|w_{\mathbf{j}}\|_{\mathcal{L}^2} \quad \forall w_{\mathbf{j}} \in W_{\mathbf{j}}, \mathbf{j} \in \mathbb{Z}^n, \quad (83)$$

and there holds an estimate

$$\|w_{\mathbf{j}}\|_{\mathcal{L}^2} \leq C2^{-2|\mathbf{j}|_1} |u|_{\mathcal{H}_{mix}^2} \quad \forall u = \sum_{\mathbf{j}} w_{\mathbf{j}} \in \mathcal{H}_{mix}^2, \quad (84)$$

see [BG98]. Inequality (84) can again be inferred from decay properties of the coefficients in the representation of  $u$  in the bases of piecewise linear splines. Then, applying the triangle inequality together with (83) and (84) yields for example for  $u \in \mathcal{H}_{mix}^2$

$$\inf_{v \in V_J^T} \|u - v\|_{\mathcal{H}^1} \leq \left\| \sum_{\mathbf{j} \notin I_J^T} w_{\mathbf{j}} \right\|_{\mathcal{H}^1} \leq \sum_{\mathbf{j} \notin I_J^T} \|w_{\mathbf{j}}\|_{\mathcal{H}^1} \stackrel{(83),(84)}{\leq} C \sum_{\mathbf{j} \notin I_J^T} 2^{|\mathbf{j}|_{\infty} - 2|\mathbf{j}|_1} \|u\|_{\mathcal{H}_{mix}^2}.$$

Summing up gives after a longer calculation for  $T < 1/2$  a generalized Jackson inequality

$$\inf_{v \in V_J^T} \|u - v\|_{\mathcal{H}^1} \leq C2^{-J} \|u\|_{\mathcal{H}_{mix}^2}, \quad (85)$$

where  $C = C(T)$ . Hence the optimal order of approximation is kept as long as  $T < 1/2$ . That is, we obtain a similar result for a multilevel approximation space without the direct use of norm equivalences. This can 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  $l$  which were restricted from above by  $t + l < r$  and  $t < r$  could be enlarged to the whole range  $t + l \leq m$  and  $t \leq m$ , see (17). Apart from eventual logarithmic terms in the extremal cases, the results remain the same.

A disadvantage of the approaches described in this paper is 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 [DS98] where the computational domain is decomposed locally and transformed to unit cubes. On these local domains the wavelet techniques can be applied. We refer the reader to these papers. Note however, that  $[0, 1]^n$  is the natural computational domain for many higher-dimensional

physical applications. Consider for example the Schrödinger equation with  $p > 1$  particles and consequently  $n = 3p$  dimensions. We see the advantages of our method especially for the range  $n \in 4, 5, \dots, 10$ .

Numerical examples on the theory presented here and on theoretical results about compression of tensor-product discretizations on full and optimized grids can be found in a companion paper [GKS98].

## References

- [Ad75] R.A. Adams, *Sobolev spaces*, Academic Press, New York, 1975.
- [AHJP94] L. Andersson, N. Hall, B. Jawerth, G. Peters, *Wavelets on closed subsets of the real line*, in: Topics in the Theory and Applications of Wavelets, L.L. Schumaker and G. Webb (eds.), Academic Press, Boston, 1994, 1-61.
- [Au92] P. Auscher, *Wavelets with boundary conditions on the interval*, in C.K. Chui (ed.), *Wavelets: A tutorial in theory and applications*, Acad. Press, Boston, 1992, 217-236.
- [Bab60] K. I. Babenko, *Approximation by trigonometric polynomials in a certain class of periodic functions of several variables*, Dokl. Akad. Nauk SSSR, 132 (1960), 982-985 (Russian); 672-675 (English translation).
- [BM87] I. Babuska, A. Miller, *A feedback finite element method with a-posteriori error estimation: Part I. The finite element method and some basic properties of the a-posteriori error estimator*, Comput. Methods Appl. Mech. Engrg. 61 (1987), 1-40.
- [BR78] I. Babuska, W.C. Rheinboldt, *Error estimates for adaptive finite element computations*, SIAM J. Numer. Anal. 15 (1978), 736-754.
- [BW85] R.E. Bank, A. Weiser, *Some a-posteriori error estimates for elliptic partial differential equations*, Math. Comp. 44 (1985), 283-301.
- [BDJ92] G. Baszenski, F.-J. Delves, 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.
- [BCR91] G. Beylkin, R.R. Coifman, V. Rokhlin, *Fast wavelet transforms and numerical algorithms I*, Comm. Pure and Appl. Math. 44 (1991), 141-183.
- [Be91] G. Beylkin, *On the representation of operators in bases of compactly supported wavelets*, SIAM J. Numer. Anal. 29 (1992), 1716-1740.
- [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.
- [Bu96] H.-J. Bungartz, *Concepts for Higher Order Finite Elements on Sparse Grids*, In: A.V. Ilin and L.R. Scott (eds.) Houston Journal of Mathematics: *Proceedings of the 3rd Int. Conf. on Spectral and High Order Methods*, Houston (1996), 159-170.

- [Bu98] H.-J. Bungartz, *Finite Elements of Higher order on Sparse Grids*, Habilitation, TU München, Institut für Informatik, 1998.
- [BG98] H.-J. Bungartz, M. Griebel, *A note on the complexity of solving Poisson's equation for spaces of bounded mixed derivatives*, J. Complexity (submitted), 1998.
- [BD97] H.-J. Bungartz, T. Dornseifer, *Sparse grids: Recent developments for elliptic partial differential equations*, SFB-Bericht 342/02/97, TU München, Institut für Informatik, 1997.
- [CW92] C.K. Chui, J.Z. Wang, *On compactly supported spline wavelets*, Trans. Amer. Math. Soc. 330 (1992), 903-915.
- [CDJV93] A. Cohen, I. Daubechies, B. Jawerth, P. Vial, *Multiresolution analysis, wavelets and fast algorithms on the interval*, C.R. Acad. Sci. Paris 316, (1993), 417-421.
- [CDV93] A. Cohen, I. Daubechies, P. Vial, *Wavelets on the interval and fast wavelet transforms*, Appl. Comput. Harm. Anal. 1 (1993), 54-81.
- [Dahl96] S. Dahlke, *Construction Principles and Applications to the Numerical Treatment of Operator Equations*, Habilitation, Aachen, 1996.
- [DDD97] S. Dahlke, W. Dahmen, R.A. DeVore, *Nonlinear approximation and adaptive techniques for solving elliptic operator equations*, in: *Multiscale Wavelet Methods for PDEs*, W. Dahmen, A. Kurdila, P. Oswald, (eds.), Academic Press, 1997, 237-284.
- [Dahm96] W. Dahmen, *Stability of Multiscale Transformations*, Journal of Fourier Analysis and Applications 2 (1996), 341-361.
- [Dahm97] 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.
- [DKU96] W. Dahmen, A. Kunoth, K. Urban, *Biorthogonal spline-wavelets on the interval - Stability and moment conditions*, IGPM-Report 129, RWTH Aachen, 1996.
- [DPS94] W. Dahmen, S. Pröβdorf, R. Schneider, *Wavelet approximation methods for pseudo-differential equations I: Stability and convergence*, Math. Z. 215 (1994), 583-620.
- [DPS93] W. Dahmen, S. Pröβdorf, R. Schneider, *Wavelet approximation methods for pseudo-differential equations II: Matrix compression and fast solution*, Advances in computational Mathematics 1 (1993), 259-335.
- [DS98] W. Dahmen, R. Schneider, *Wavelets on manifolds I: Construction and Domain Decomposition*, IGPM-Preprint 149, January 1998.
- [Del82] F.-J. Delves, *d-Variate Boolean interpolation*, J. Approx. Theory, 34 (1982), 99-114.
- [DeV98] R.A. DeVore, *Nonlinear Approximation*, Acta Numerica (1998), 51-150.
- [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.
- [F09] G. Faber: *Über stetige Funktionen*, Mathematische Annalen, 66 (1909), 81–94.
- [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.
- [GKS98] M. Griebel, S. Knapek, T. Schiekofer, *Galerkin discretizations of pseudodifferential operators using wavelets of tensor-product type: Dimension reduction and compression*, 1998, in preparation.
- [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.
- [GOS98] M. Griebel, P. Oswald, T. Schiekofer, *Sparse grids for boundary integral equations*, Numerische Math. (submitted), 1998.
- [GSZ92] M. Griebel, M. Schneider, C. Zenger, *A combination technique for the solution of sparse grid problems*, in Iterative Methods in Linear Algebra, P. de Groen and R. Beauwens, eds., Elsevier, Amsterdam, 1992, 263–281.
- [HMS94] N. Heuer, M. Maischak, E.P. Stephan, *The hp-version of the boundary element method for screen problems*, Preprint, University Hannover, 1994.
- [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.
- [LM72] J.-L. Lions, E. Magenes, *Non-Homogeneous Boundary Value Problems and Applications I*, Springer, Berlin, 1972.
- [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.
- [R91] D. Röschke, *Über eine Kombinationstechnik zur Lösung partieller Differentialgleichungen*, Diplomarbeit, Institut für Informatik, TU München, 1991.
- [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; also: Habilitation, Technische Hochschule Darmstadt, 1995.
- [Si98] W. Sickel, *Tensor products of periodic Besov spaces*, J. of Computational Analysis and Applications, to appear.
- [SS98] W. Sickel, F. Sprengel, *Interpolation on sparse grids and Nikol'skij-Besov spaces of dominating mixed derivative*, J. of Computational Analysis and Applications, to appear.
- [Sm63] S.A. Smolyak, *Quadrature and interpolation formulas for tensor products of certain classes of functions*, Dokl. Akad. Nauk SSSR 4 (1963), 240–243.
- [St96] E.P. Stephan, *The h-p boundary element method for solving 2- and 3-dimensional problems*, Preprint, University Hannover, 1996.
- [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).
- [Tr78] H. Triebel, *Interpolation Theory, Functions Spaces, Differential Operators*, Dt. Verlag. Wiss., Berlin 1978; North Holland, Amsterdam, 1978.
- [TWW88] J. F. Traub, G. W. Wasilkowski, H. Wozniakowski, *Information-Based Complexity*, Academic Press, New York, 1988.
- [V94] R. Verfürth, *A posteriori error estimates and adaptive mesh refinement techniques*, J. Comput. and Appl. Math. 50 (1994), 67–83.
- [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.
- [We91] A. G. Werschulz, *The computational complexity of differential and integral equations, An information based approach*, Oxford University Press, New York, 1991.
- [We95] A. G. Werschulz, *The complexity of the Poisson problem for spaces of bounded mixed derivatives*, Technical Report CUCS-016-95, 1995.
- [Z91] C. Zenger, *Sparse grids*, In Hackbusch, W. (ed.): Parallel Algorithms for Partial Differential Equations, *Notes on Num. Fluid Mech. 31*, Vieweg, Braunschweig, 1991.