INTEGRAL OPERATORS ON SPARSE GRIDS

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

Abstract. This paper is concerned with the construction and use of wavelet approximation spaces for the fast evaluation of integral expressions. Based on the framework of anisotropic tensor-product biorthogonal wavelet bases we introduce sparse grid (hyperbolic cross) type approximation spaces and discuss additional compression schemes on top of these discretizations. Numerical examples for the Laplace equation with Dirichlet boundary conditions and an additional integral term with smooth kernel validate our theoretical findings.

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

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

1. Introduction.

A naive Galerkin discretization of an integral operator

$$Au(\mathbf{x}) := \int K(\mathbf{x},\mathbf{y})u(\mathbf{y})d\mathbf{y} \tag{1.1}$$

with global kernel K leads to a dense stiffness matrix. Hence, on a uniform full grid with $O(2^{nJ})$ unknowns (n dimension, J maximal level in a multiscale discretization), 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 J. A standard strategy to reduce this cost is to exploit the decay properties of the kernel that, for singular kernels, are typically of the form

$$\exists\ C = C_{\boldsymbol{\alpha\beta}} < \infty : \left|\partial_x^{\boldsymbol{\alpha}}\partial_y^{\boldsymbol{\beta}}K(\mathbf{x},\mathbf{y})\right| \le C\cdot\left|\mathbf{x}-\mathbf{y}\right|^{-f(\boldsymbol{\alpha},\boldsymbol{\beta})},\ f(\boldsymbol{\alpha},\boldsymbol{\beta})>0,\ \mathbf{x}\neq\mathbf{y}. \tag{1.2}$$

Such decay properties do hold for pseudo-differential operators [24]. Together with (isotropic) biorthogonal wavelets with a sufficient number of vanishing moments, most entries in the corresponding stiffness matrices are then close to zero and can be replaced by zero without destroying the order of approximation [4, 9, 12].

Another approach to reduce the cost of the integral evaluation is to replace the full grid approximation space by the so-called sparse grid space [30]. The idea is to use an anisotropic tensor-product basis together with a specific subset of the full grid approximation space. It has been shown that under some assumptions on the approximation and smoothness properties of the underlying basis functions and provided that certain additional regularity assumptions are fulfilled, the resulting approximation spaces exhibit the same order of approximation as the full grid space, while having less dimension, see for example [32].

The sparse grid approach (that also appeared under the names hyperbolic cross approximation or boolean blending schemes) is well established in approximation and interpolation theory, see e.g. [1, 11, 13, 30, 33, 34]. First approaches to use sparse grids for integral operators can be found in [16, 18, 20, 27]. In [20] it has been observed that a discretization with adaptive sparse grid spaces leads to good approximation rates for the single layer potential equation on a square. In [18] theoretical results on the discretization of elliptic operators of arbitrary order with sparse grids and biorthogonal wavelets have been presented. In both papers sparse grids are used as test- and ansatz-spaces in a Galerkin discretization. Hence, without further compression, the resulting discrete systems are dense. This rises the question, whether additional compression schemes along the lines of [10, 29] are applicable for this type of tensor-product discretizations.

For integral operators with smooth kernel, sparse grids may be used directly for the approximation of the kernel. The theoretical background to this approach is that the smoothness of the kernel translates into decay properties of its multiscale coefficients. Then one can construct (modifications of) sparse grid spaces such that the evaluation of integral expressions of the type (1.1) can be performed with much lower cost but the same accuracy as with the full grid approximation space. This approach has already been used for complexity estimates of Fredholm integral equations of the second kind with smooth kernels for the special case with isotropic Sobolev smoothness or dominating mixed smoothness for K und u [15, 16, 27].

In this paper we develop sparse grid type approximation spaces for the approximate evaluation of (1.1). We start with smoothness classes including isotropic Sobolev spaces as well as spaces of dominating mixed derivative, and construct sparse grid type approximation spaces for the kernel that are adapted to the smoothness of K and u from (1.1), as well as the norm in which the error of the integral evaluation is measured. Moreover we develop compression schemes for anisotropic tensor-product discretizations for kernels that obey (1.2). Note that although these two types of compression are based on different properties of the kernel, it is instructive to present both approaches in a somewhat unified framework.

The remainder of the paper is as follows. Section 2 introduces the necessary notation and summarizes the basic facts about biorthogonal wavelet bases, tensor-product spaces, norm equivalences and the smoothness classes we consider. They are certain intersections of classes of functions with dominating mixed derivatives, see (2.1) below. In Section 3 we investigate strategies for the fast evaluation of (1.1). We introduce our new approximation spaces and give estimates on the order of approximation obtained with these spaces and discuss compression schemes for anisotropic tensor-product discretizations for kernels that obey (1.2). In Section 4 we present numerical examples for the integral evaluation and for the numerical solution of the Laplace equation with Dirichlet boundary conditions and an additional integral term with smooth kernel.

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

2.1. Sobolev spaces. Let us denote by $\mathcal{H}^t(I^n), t \in \mathbb{R}$, a scale of standard Sobolev spaces on $I^n = [0,1]^n$ and by $\mathcal{L}^2(I^n)$ the space of $\mathcal{L}^2$-integrable functions on I^n.

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

DEFINITION 1. *Let* $t \in \mathbb{R}, l \in \mathbb{R}_0^+$, $\mathbf{1} := (1, \dots, 1)$ *and* $\mathbf{e_i} := (0, \dots, 0, 1, 0, \dots, 0)$ *the*

i-th unit-vector in $\mathbb{R}^n$. Then we define

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

where $\mathcal{H}^{\mathbf{k}}_{mix}(I^n) := \mathcal{H}^{k_1}(I) \otimes \ldots \otimes \mathcal{H}^{k_n}(I)$ for $\mathbf{k} \in \mathbb{R}^n$. For $l < 0$ we define $\mathcal{H}^{t,l}_{mix}(I^n)$ as dual of $\mathcal{H}^{-t,-l}_{mix}(I^n)$.

This definition includes the class of functions of dominating mixed derivative $\mathcal{H}^{t,0}_{mix}(I^n)$ as well as the standard isotropic Sobolev spaces $\mathcal{H}^t(I^n) = \mathcal{H}^{0,t}_{mix}(I^n)$. For additional information on these spaces, specifically on the definition via Fourier transform, the treatment of boundary conditions and the connections between spaces of bounded mixed derivative and these anisotropic tensor-product spaces see [18, 22] and [21]. When the meaning is clear from the context, we will sometimes drop I^n in $\mathcal{H}^{t,l}_{mix}(I^n)$ and write $\mathcal{H}^{t,l}_{mix}$.

2.2. Biorthogonal wavelet bases. The approximation spaces considered here stem from anisotropic tensor-products of univariate function spaces. We start from a one-dimensional multi-resolution analysis $\mathcal{L}^2 = \bigcup_{j\geq 0} V_j$ and we assume that the complement spaces $W_j = V_j \ominus V_{j-1}$ of V_{j-1} in V_j are spanned by some multiscale basis functions such that we have $W_j = \operatorname{span}\{\psi_{jk}, k \in \tau_j\}$, where τ_j is some index set defined from the subdivision rate of successive refinement levels. We will consider dyadic refinement throughout the paper. Moreover, we assume that $\{\psi_{jk}, k \in \tau_j\}$ forms a Riesz-basis of W_j and that there exists a dual system $\{\tilde{\psi}_{jk}, k \in \tau_j, j \in \mathbb{N}_0\}$, such that

$$\langle \psi_{jk}, \tilde{\psi}_{j'k'} \rangle_{\mathcal{L}^2} = \delta_{jj'}\delta_{kk'}, \;\; j, j' \in \mathbb{N}_0, k \in \tau_j, k' \in \tau_{j'}$$

holds. We assume $\|\psi_{jk}\|_{\mathcal{L}^2} = 1$ and $diam(supp(\psi_{jk})) \lesssim 2^{-j}$. In the following let ψ_{jk} and $\tilde{\psi}_{jk}$ have N and $\tilde{N}$ vanishing moments, respectively. Moreover we write $\tilde{W}_j = \operatorname{span}\{\tilde{\psi}_{jk}, k \in \tau_j\}$ for the complement spaces spanned by the dual wavelets. Under these conditions every $u \in \mathcal{L}^2$ has unique expansions

$$u = \sum_{j=0}^{\infty} \sum_{k\in\tau_j} \langle u, \tilde{\psi}_{jk} \rangle_{\mathcal{L}^2} \psi_{jk} = \sum_{j=0}^{\infty} \sum_{k\in\tau_j} \langle u, \psi_{jk} \rangle_{\mathcal{L}^2} \tilde{\psi}_{jk}. \tag{2.2}$$

To simplify things we assume in the following the validity of the norm equivalences

$$\|u\|^2_{\mathcal{H}^t} \simeq \sum_{j=0}^{\infty} \|w_j\|^2_{\mathcal{H}^t} \simeq \sum_{j=0}^{\infty} 2^{2tj} \|w_j\|^2_{\mathcal{L}^2} \text{ for } t \in (-\tilde{r}, r), \tag{2.3}$$

where $u = \sum_{j=0}^{\infty} w_j, w_j \in W_j$, with $r, \tilde{r} \in \mathbb{R}^+_0$ and an analogous relation for the dual wavelet (with $t \in (-r, \tilde{r})$). Such two-sided estimates can be inferred from the validity of direct estimates (estimates of Jackson type) and inverse estimates (Bernstein inequalities) for the primal and the dual wavelets as a consequence of approximation theory in Sobolev spaces together with interpolation and duality arguments, see e.g. [8, 25].

For the higher-dimensional case $n > 1$, let $\mathbf{j} \in \mathbb{N}^n$ be given, and consider the tensor-product partition with uniform step size 2^{-j_i} into the i-th coordinate direction. By $W_{\mathbf{j}}$ and $\tilde{W}_{\mathbf{j}}$ we denote the corresponding tensor-product function spaces $W_{\mathbf{j}} := W_{j_1} \otimes$

$\ldots \otimes W_{j_n}$ and $\tilde{W}_{\mathbf{j}} := \tilde{W}_{j_1} \otimes \ldots \otimes \tilde{W}_{j_n}$. A basis of $W_{\mathbf{j}}$ is then given by $\cup_{\mathbf{k}\in\tau_{\mathbf{j}}}\{\psi_{\mathbf{jk}}(\mathbf{x}) = \psi_{j_1 k_1}(x_1)\cdot \ldots \cdot \psi_{j_n k_n}(x_n)\}$, with $\tau_{\mathbf{j}} := \tau_{j_1} \times \ldots \times \tau_{j_n}$.
Under the assumption of the validity of (2.3) and an analogous relation for the dual wavelet, it is then straightforward to prove the following norm equivalences

$$\|u\|^2_{\mathcal{H}^{t,l}_{mix}} \simeq \sum_{\mathbf{j}} 2^{2t|\mathbf{j}|_1 + 2l|\mathbf{j}|_\infty} \|w_{\mathbf{j}}\|^2_{\mathcal{L}^2}, \text{ for } u = \sum_{\mathbf{j}} w_{\mathbf{j}} \in \mathcal{H}^{t,l}_{mix}, w_{\mathbf{j}} \in W_{\mathbf{j}}, \tag{2.4}$$

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

$$\|u\|^2_{\mathcal{H}^{t,l}_{mix}} \simeq \sum_{\mathbf{j}} 2^{2t|\mathbf{j}|_1 + 2l|\mathbf{j}|_\infty} \|\tilde{w}_{\mathbf{j}}\|^2_{\mathcal{L}^2}, \text{ for } u = \sum_{\mathbf{j}} \tilde{w}_{\mathbf{j}} \in \mathcal{H}^{t,l}_{mix}, \tilde{w}_{\mathbf{j}} \in \tilde{W}_{\mathbf{j}}, \tag{2.5}$$

$t, t+l \in (-r, \tilde{r})$, see [18]. For $t = 0$ or $l = 0$ we regain the (standard) norm equivalences for the isotropic Sobolev space $\mathcal{H}^t$ and the Sobolev space with dominating mixed derivative $\mathcal{H}^{t,0}_{mix}$. The different factors $2^{2t|\mathbf{j}|_1}$ and $2^{2l|\mathbf{j}|_\infty}$ in these equivalences reflect the different smoothness requirements.
One of the merits of the validity of a norm equivalence for $\mathcal{H}^s$ is the fact that it leads directly to optimal preconditioning for $\mathcal{H}^s$-elliptic problems and hence to fast iterative methods with convergence rates independent of the number of unknowns, see [8].
Moreover, (2.4) and (2.5) may be used to analyse the approximation of the embedding $\mathcal{H}^{t,l}_{mix} \to \mathcal{H}^s$ and to construct optimized sparse grid type Finite Element approximation spaces for elliptic variational problems including integral operators [18]. However, for integral operators the resulting stiffness matrices are dense. In the next Section we develop schemes to avoid the additional cost due to the denseness of the matrices.

3. Tensor-product approximation and compression of integral operators. The aim of this Section is to efficiently approximate the coefficients $h_{\mathbf{jk}}$ of

$$h(\mathbf{x}) := Au(\mathbf{x}) = \sum_{\mathbf{j}\in\mathbb{N}^n} \sum_{\mathbf{k}\in\tau_{\mathbf{j}}} h_{\mathbf{jk}} \tilde{\psi}_{\mathbf{jk}}(\mathbf{x}). \tag{3.1}$$

To this end, we represent K and u from (1.1) in the dual and primal basis, respectively,

$$K(\mathbf{x},\mathbf{y}) = \sum_{(\mathbf{j},\mathbf{k})\in\mathbb{N}^{2n}} \sum_{(\mathbf{l},\mathbf{m})\in\tau_{\mathbf{jk}}} a_{\mathbf{jklm}} \tilde{\psi}_{\mathbf{jl}}(\mathbf{x}) \tilde{\psi}_{\mathbf{km}}(\mathbf{y}), \qquad u(\mathbf{x}) = \sum_{\mathbf{j}\in\mathbb{N}^n} \sum_{\mathbf{l}\in\tau_{\mathbf{k}}} u_{\mathbf{jl}} \psi_{\mathbf{jl}}(\mathbf{x}),$$

where $a_{\mathbf{jklm}} = \langle K, \psi_{\mathbf{jl}}(\mathbf{x})\psi_{\mathbf{km}}(\mathbf{y})\rangle_{\mathcal{L}^2} = \langle A\psi_{\mathbf{jl}}, \psi_{\mathbf{km}}\rangle_{\mathcal{L}^2}$, $u_{\mathbf{jl}} = \langle u, \tilde{\psi}_{\mathbf{jl}}\rangle_{\mathcal{L}^2}$, see (2.2), and $\tau_{\mathbf{jk}} := \tau_{\mathbf{j}} \times \tau_{\mathbf{k}}$. Here we assume that eventual boundary conditions are implemented into the definition of the primal and dual wavelets, compare Section 4.
Given an index set $\mathcal{I} \subset \mathbb{N}^{2n}$ we consider the approximation space

$$\tilde{V}_{\mathcal{I}} := \bigoplus_{(\mathbf{j},\mathbf{k})\in\mathcal{I}} \tilde{W}_{\mathbf{j}} \otimes \tilde{W}_{\mathbf{k}}. \tag{3.2}$$

Typical examples are the full grid spaces

$$\tilde{V}_{\{|(\mathbf{j},\mathbf{k}|_\infty \le J\}} = \bigoplus_{|(\mathbf{j},\mathbf{k})|_\infty \le J} \tilde{W}_{\mathbf{j}} \otimes \tilde{W}_{\mathbf{k}}, J > 0, \tag{3.3}$$

and the sparse grid (or hyperbolic cross) spaces

$$\tilde{V}_{\{|(\mathbf{j},\mathbf{k})|_1 \le J\}} = \bigoplus_{|(\mathbf{j},\mathbf{k})|_1 \le J} \tilde{W}_{\mathbf{j}} \otimes \tilde{W}_{\mathbf{k}}, J > 0, \tag{3.4}$$

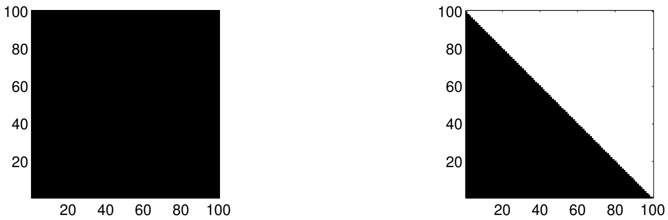

FIGURE 3.1. *Index sets corresponding to the full (left) and the sparse grid space (right) with maximal level $J = 100$, in one dimension ($n = 1$).*

from [30]. These are associated with rectangular index sets and with the index sets $\{|(\mathbf{j},\mathbf{k})|_1 \leq J\}$, respectively, compare Figure 3.1. The dimensions of the full grid and the sparse grid approximation space are $O(2^{2nJ})$ and $O(2^J J^{2n-1})$, respectively, see [14, 30]. Note that the dimension of the sparse grid space compares favourably with the dimension of the full grid space.
Given an additional index set $\tau_{\mathbf{jk}}^{comp} \subset \tau_{\mathbf{jk}}$ we define the following two approximations of the kernel K,

$$K_{\mathcal{I}}(\mathbf{x},\mathbf{y}) := \sum_{(\mathbf{j},\mathbf{k})\in\mathcal{I}} \sum_{(\mathbf{l},\mathbf{m})\in\tau_{\mathbf{jk}}} a_{\mathbf{jklm}} \tilde{\psi}_{\mathbf{jl}}(\mathbf{x}) \tilde{\psi}_{\mathbf{km}}(\mathbf{y}) \tag{3.5}$$

and

$$K_{\mathcal{I}}^{comp}(\mathbf{x},\mathbf{y}) := \sum_{(\mathbf{j},\mathbf{k})\in\mathcal{I}} \sum_{(\mathbf{l},\mathbf{m})\in\tau_{\mathbf{jk}}^{comp}} a_{\mathbf{jklm}} \tilde{\psi}_{\mathbf{jl}}(\mathbf{x}) \tilde{\psi}_{\mathbf{km}}(\mathbf{y}). \tag{3.6}$$

In (3.5) the kernel K is approximated by tensor-product approximation and by setting all coefficients $a_{\mathbf{jklm}}$ with $(\mathbf{j},\mathbf{k}) \notin \mathcal{I}$ to zero. The second approximation step (3.6) consists of (additionally) setting some of the coefficients $a_{\mathbf{jklm}}$ to zero.
The corresponding integral operators are the uncompressed operator A from (1.1) and the two operators

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

Now, to obtain schemes for an efficient evaluation of the integral expression Au, we have to choose the index sets $\mathcal{I}$ and $\tau_{\mathbf{jk}}^{comp}$ in such a way, that on the one hand a good approximation of the integral is obtained, and on the other hand, the cost of the integral evaluation remains as small as possible. The rest of this Section is devoted to this task.

3.1. Tensor product approximation. The following Theorem estimates the approximation error $\|(A - A_{\mathcal{I}})u\|_{\mathcal{H}^s}$ for a given index set $\mathcal{I}$.

THEOREM 1. *Let $(A - A_{\mathcal{I}})u \in \mathcal{H}^s, u \in \mathcal{H}_{mix}^{p,q}, K \in \mathcal{H}_{mix}^{t,l}$. We assume that the norm equivalences (2.4) and (2.5) hold. Moreover let the parameters $\tilde{r}, r$ from (2.4) and (2.5) be such that $-\tilde{r} < s < r$, $p \in [0,r), p+q \in [0,r)$ and $t \in [0,\tilde{r}), t+l \in [0,\tilde{r})$ holds. Then,*

$$\begin{aligned} &\|(A - A_{\mathcal{I}})u\|_{\mathcal{H}^s} \\ &\quad \leq C \cdot \max_{(\mathbf{j},\mathbf{k})\notin\mathcal{I}} \left(2^{s|\mathbf{j}|_\infty - l|\mathbf{j},\mathbf{k}|_\infty - q|\mathbf{k}|_\infty - t|\mathbf{j}|_1 - (t+p)|\mathbf{k}|_1}\right) \cdot \|K\|_{\mathcal{H}_{mix}^{t,l}} \cdot \|u\|_{\mathcal{H}_{mix}^{p,q}}. \end{aligned} \tag{3.7}$$

Proof: We use the biorthogonality between the dual and the primal wavelets and the validity of (2.4) and (2.5). It holds

$$
\begin{aligned}
\|(A-A_{\mathcal{I}})u\|^2_{\mathcal{H}^s} &= \| \int \left(\sum_{(\mathbf{j},\mathbf{k})\notin\mathcal{I}} \sum_{(\mathbf{l},\mathbf{m})\in\tau_{\mathbf{jk}}} a_{\mathbf{jklm}} \tilde{\psi}_{\mathbf{jl}}(\mathbf{x}) \tilde{\psi}_{\mathbf{km}}(\mathbf{y}) \right) \cdot \\
&\qquad \cdot \left(\sum_{\mathbf{o}} \sum_{\mathbf{p}\in\tau_{\mathbf{o}}} u_{\mathbf{op}} \psi_{\mathbf{op}}(\mathbf{y}) \right) d\mathbf{y} \|^2_{\mathcal{H}^s} \\
&= \| \sum_{(\mathbf{j},\mathbf{k})\notin\mathcal{I}} \sum_{(\mathbf{l},\mathbf{m})\in\tau_{\mathbf{jk}}} a_{\mathbf{jklm}} u_{\mathbf{km}} \tilde{\psi}_{\mathbf{jl}}(\mathbf{x}) \|^2_{\mathcal{H}^s} \\
&= \| \sum_{\mathbf{j}} \sum_{\mathbf{l}\in\tau_{\mathbf{j}}} \left(\sum_{\mathbf{k}:(\mathbf{j},\mathbf{k})\notin\mathcal{I}} \sum_{\mathbf{m}\in\tau_{\mathbf{k}}} a_{\mathbf{jklm}} u_{\mathbf{km}} \right) \tilde{\psi}_{\mathbf{jl}}(\mathbf{x}) \|^2_{\mathcal{H}^s} \\
&\overset{(2.5)}{\simeq} \sum_{\mathbf{j}} \sum_{\mathbf{l}\in\tau_{\mathbf{j}}} 2^{2s|\mathbf{j}|_\infty} \left(\sum_{\mathbf{k}:(\mathbf{j},\mathbf{k})\notin\mathcal{I}} \sum_{\mathbf{m}\in\tau_{\mathbf{k}}} a_{\mathbf{jklm}} u_{\mathbf{km}} \right)^2 \\
&\le \max_{(\mathbf{j},\mathbf{k})\notin\mathcal{I}} \left(2^{s|\mathbf{j}|_\infty - l|\mathbf{j},\mathbf{k}|_\infty - q|\mathbf{k}|_\infty - t|\mathbf{j}|_1 - (t+p)|\mathbf{k}|_1} \right)^2 \\
&\qquad \cdot \sum_{\mathbf{j}} \sum_{\mathbf{l}\in\tau_{\mathbf{j}}} \left(\sum_{\mathbf{k}:(\mathbf{j},\mathbf{k})\notin\mathcal{I}} \sum_{\mathbf{m}\in\tau_{\mathbf{k}}} 2^{l|\mathbf{j},\mathbf{k}|_\infty + t|\mathbf{j},\mathbf{k}|_1} a_{\mathbf{jklm}} \cdot 2^{q|\mathbf{k}|_\infty + p|\mathbf{k}|_1} u_{\mathbf{km}} \right)^2 \\
&\le \max_{(\mathbf{j},\mathbf{k})\notin\mathcal{I}} \left(2^{s|\mathbf{j}|_\infty - l|\mathbf{j},\mathbf{k}|_\infty - q|\mathbf{k}|_\infty - t|\mathbf{j}|_1 - (t+p)|\mathbf{k}|_1} \right)^2 \\
&\qquad \cdot \sum_{\mathbf{j}} \sum_{\mathbf{l}\in\tau_{\mathbf{j}}} \left(\sum_{\mathbf{k}} \sum_{\mathbf{m}\in\tau_{\mathbf{k}}} 2^{2l|\mathbf{j},\mathbf{k}|_\infty + 2t|\mathbf{j},\mathbf{k}|_1} a^2_{\mathbf{jklm}} \right) \cdot \left(\sum_{\mathbf{k}} \sum_{\mathbf{m}\in\tau_{\mathbf{k}}} 2^{2q|\mathbf{k}|_\infty + 2p|\mathbf{k}|_1} u^2_{\mathbf{km}} \right) \\
&\overset{(2.4)}{\simeq} \max_{(\mathbf{j},\mathbf{k})\notin\mathcal{I}} \left(2^{s|\mathbf{j}|_\infty - l|\mathbf{j},\mathbf{k}|_\infty - q|\mathbf{k}|_\infty - t|\mathbf{j}|_1 - (t+p)|\mathbf{k}|_1} \right)^2 \cdot \|K\|^2_{\mathcal{H}^{t,l}_{mix}} \cdot \|u\|^2_{\mathcal{H}^{p,q}_{mix}} .
\end{aligned}
$$

□

Now, for a given bound ϵ of the error $\|(A-A_{\mathcal{I}})u\|_{\mathcal{H}^s}$ we determine minimal index sets $\mathcal{I}$. To simplify things, we assume in the rest of this Section $l,t,p,q \geq 0$ and $l+t \geq s \geq -(p+q)$. Then, if we use in (3.7) for $\mathcal{I}$ the index set $\{(\mathbf{j},\mathbf{k}) \in \mathbb{N}^{2n} : |\mathbf{j},\mathbf{k}|_\infty \leq J\}$ which corresponds to the standard full grid approximation space (3.3) for the kernel K, we obtain

$$\|(A-A_{\mathcal{I}})u\|_{\mathcal{H}^s} \leq C \cdot 2^{(s-l-t)J} \cdot \|K\|_{\mathcal{H}^{t,l}_{mix}} \cdot \|u\|_{\mathcal{H}^{p,q}_{mix}} . \tag{3.8}$$

Note that the order of approximation in (3.8) does not explicitely depend on the smoothness of u (p and q only restrict the range of l and t). For example for a kernel with product structure $K(\mathbf{x},\mathbf{y}) = k_1(\mathbf{x}) \cdot k_2(\mathbf{y})$ the smoothness of Au is only dependent on the smoothness of k_1.

A closer look on (3.7) and (3.8) reveals that one could discard indices from the full grid index set without destroying the order of approximation from (3.8). This motivates the following definition of index sets.

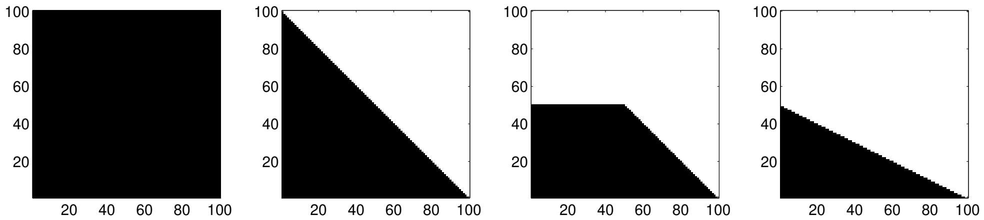

FIGURE 3.2. *Index sets* $\mathcal{I}_{100}(0,0,1,0,0), \mathcal{I}_{100}(0,1,0,0,0), \mathcal{I}_{100}(0,0,1,0,1)$ *and* $\mathcal{I}_{100}(0,1,0,0,1)$ *for* $n = 1$ *from left to right.*

DEFINITION 2. *We introduce the parametrized index sets*

$$\begin{aligned} \mathcal{I}_J(s,t,l,p,q) := \{(\mathbf{j},\mathbf{k}) \in \mathbb{N}^{2n} : \\ -s|\mathbf{j}|_\infty + l|\mathbf{j},\mathbf{k}|_\infty + t|\mathbf{j},\mathbf{k}|_1 + q|\mathbf{k}|_\infty + p|\mathbf{k}|_1 \le (l+t-s)J\}. \end{aligned} \tag{3.9}$$

According to (3.2) the related approximation spaces are

$$\tilde{V}_{\mathcal{I}_J(s,t,l,p,q)} = \sum_{(\mathbf{j},\mathbf{k}) \in \mathcal{I}_J(s,t,l,p,q)} \tilde{W}_\mathbf{j} \otimes \tilde{W}_\mathbf{k}.$$

When the meaning is clear from the context, we sometimes drop the parameters (s,t,l,p,q) and write $\mathcal{I}_J$ instead of $\mathcal{I}_J(s,t,l,p,q)$. For $s = l = p = q = 0, t > 0$ and $s = t = p = q = 0, l > 0$ we regain the index sets corresponding to the sparse grid spaces (3.4) and the standard full grid spaces (3.3), respectively, compare Figure 3.2 (first and second) and Figure 3.1. Additional smoothness in u, that is $p > 0$ or $q > 0$, further reduces the number of entries in the index sets $\mathcal{I}_J$. Figure 3.2 (third and fourth) gives two examples for the one-dimensional case.
Due to the biorthogonality of the primal and dual wavelets, not all coefficients of the wavelet representation of u are required for the evaluation of $A_\mathcal{I} u$ or $A_\mathcal{I}^{comp} u$. For example, the coefficients of $h^{\mathcal{I}_J} := A_{\mathcal{I}_J} u$ read

$$h_{\mathbf{jl}}^{\mathcal{I}_J} = \sum_{\mathbf{k}:(\mathbf{j},\mathbf{k}) \in \mathcal{I}_J} \sum_{\mathbf{m} \in \tau_\mathbf{k}} a_{\mathbf{jklm}} u_{\mathbf{km}}. \tag{3.10}$$

Therefore, we need the coefficients $u_{\mathbf{km}}$ only if $\exists \mathbf{k} \in \mathbb{N}^n : (\mathbf{0},\mathbf{k}) \in \mathcal{I}_J$. In this way, $\mathcal{I}_J$ also induces approximation spaces for u,

$$V_{\mathcal{U}_{\mathcal{I}_J}} := \bigoplus_{\mathbf{k} \in \mathcal{U}_{\mathcal{I}_J}} W_\mathbf{k} \ ,$$

where

$$\begin{aligned} \mathcal{U}_{\mathcal{I}_J(s,t,l,p,q)} &:= \{\mathbf{k} \in \mathbb{N}^n : (\mathbf{0},\mathbf{k}) \in \mathcal{I}_J(s,t,l,p,q)\} \\ &= \{\mathbf{k} \in \mathbb{N}^n : (l+q)|\mathbf{k}|_\infty + (t+p)|\mathbf{k}|_1 \le (l+t-s)J\}. \end{aligned} \tag{3.11}$$

Similarly, for Au the definition of $\mathcal{I}_J$ implies the index set

$$\{\mathbf{j} \in \mathbb{N}^n : (\mathbf{j},\mathbf{0}) \in \mathcal{I}_J(s,t,l,p,q)\} = \{\mathbf{j} \in \mathbb{N}^n : (l-s)|\mathbf{j}|_\infty + t|\mathbf{j}|_1 \le (l+t-s)J\}.$$

Note that this index set is optimal for the approximation of functions from $\mathcal{H}_{mix}^{t,l}$ if the error is measured in the $\mathcal{H}^s$-norm, compare [18]. Moreover, it depends only on $\frac{s-l}{t}$.

The following Lemma summarizes the approximation properties of our new approximation spaces. It is a direct consequence of (3.7) and the definition of $\mathcal{I}_J$ in (3.9).

LEMMA 1. *Let $(A - A_{\mathcal{I}_J})u \in \mathcal{H}^s$. Then, under the assumptions of Theorem 1 it holds*

$$\|(A - A_{\mathcal{I}_J(s,t,l,p,q)})u\|_{\mathcal{H}^s} \leq C \cdot 2^{(s-l-t)J} \cdot \|K\|_{\mathcal{H}^{t,l}_{mix}} \cdot \|u\|_{\mathcal{H}^{p,q}_{mix}}. \tag{3.12}$$

Inequality (3.12) shows that the use of the index set $\mathcal{I}_J(s,t,l,p,q)$ leads to the same rate of convergence as the use of the full grid index set, compare (3.8), although in most cases the number of elements in $\mathcal{I}_J(s,t,l,p,q)$ is much smaller.

Note that the validity of (3.7) and (3.12) may be extended to the maximal range $t, t+l \in [0, \tilde{N}]$ and $p, p+q \in [0, N]$ because of the wider range of the $\geq$-estimate in (2.3), see [29] (with eventual additional logarithmic terms in the extremal case and special cases of the index sets).

If the optimal approximation order is bounded by the approximation order of the wavelets and not by the smoothness of K or u (i.e. for example for the case $t > \tilde{N}$), then the optimal index set $\mathcal{I}_J$ is obtained by setting $t = \tilde{N}$, $p = N$ and $l = q = 0$. Then, for example for the $\mathcal{L}^2$-norm, i.e. $s = 0$, the optimal index set is $\mathcal{I}_J(0, \tilde{N}, 0, N, 0)$.

Together with an estimate of the number of elements in $\mathcal{I}_J(s,t,l,p,q)$ (which may be obtained from general results in [14]), it is straightforward to obtain estimates for the cost of approximating Au. For example in the case $t > 0$, i.e. for a kernel with dominating mixed smoothness, the number of elements $\mathcal{N}$ needed to obtain an approximation with $\|(A - A_{\mathcal{I}_J(s,t,l,p,q)})u\|_{\mathcal{H}^s} \leq \epsilon, \epsilon \in \mathbb{R}^+$, is bounded by

$$\mathcal{N} \leq \begin{cases} O(\epsilon^{\frac{1}{s-l-t}}) & \text{for } s-l > 0, \\ O(\epsilon^{\frac{1}{s-l-t}} \ln(\epsilon^{-1})^{n-1}) & \text{for } s-l = 0, \\ O\left(\epsilon^{\frac{1}{\frac{s-l}{n}-t}}\right) & \text{for } s-l < 0. \end{cases} \tag{3.13}$$

In the case of isotropic smoothness, i.e. $t = p = 0$, the number of elements $\mathcal{N}$ needed to obtain an approximation $A_{\mathcal{I}_J}u$ with $\|(A - A_{\mathcal{I}_J})u\|_{\mathcal{H}^s} \leq \epsilon$ is bounded by

$$\mathcal{N} \leq \begin{cases} O(\epsilon^{\frac{n}{s-l}}) & \text{for } s-l+q > 0, \\ O(\epsilon^{\frac{n}{s-l}} \ln(\epsilon^{-1})) & \text{for } s-l+q = 0, \\ O\left(\epsilon^{\frac{2n}{s-l-q}}\right) & \text{for } s-l+q < 0. \end{cases} \tag{3.14}$$

The above estimates have to be compared with the corresponding number of elements

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

when using the full grid approximation spaces.

Hence, for a kernel with dominating mixed smoothness and $s > l$, an ϵ-approximation may be obtained with a cost that is asymptotically independent of the dimension n, compare (3.13). For $s \leq l$, there appears a slight n-dependence. This is similar to the approximation of functions from $\mathcal{H}^{t,l}_{mix}$ in $\mathcal{H}^s$. Indeed, in this case, the integral evaluation is asymptotically not more expensive than the approximation of a function from $\mathcal{H}^{t,l}_{mix}$, compare [18].

For K and u with isotropic smoothness, the cost is dependent on n, compare (3.14). But in any case, the integral evaluation with the above scheme is less costly than the evaluation with the use of full grid approximation spaces. See also [26, 27], where special cases of this phenomenon have been described previously.

Note however, that without further compression the above sparse grid approach still leads to dense discrete operators.

3.2. Additional compression. Additional compression may be obtained via a special choice of the index sets $\tau_{\mathbf{jk}}^{comp}$ from (3.6). The idea is to drop small entries in the stiffness matrix without destroying the order of approximation. Here, simple thresholding, i.e. dropping those entries in the stiffness matrix that are under a certain threshold, is in general not feasible, as this approach requires the computation of all entries. Moreover, this approach does not take into consideration the smoothness of u and the norm in which the error is measured.
Special compression schemes for isotropic discretizations that take these considerations into account can be found in [10, 29]. Then for isotropic tensor-product wavelets on a full grid the number of entries in the stiffness matrix after compression is only of order $O(J^k 2^{Jn})$ with some $k \in \mathbb{R}^+$ or even $O(2^{Jn})$. Corresponding investigations concerning anisotropic tensor-product discretizations seem still to be missing, especially together with sparse grid type discretizations. Along the lines of [10] we obtain the following results.
Decay properties of the kernel such as (1.2) introduce decay properties of the stiffness matrix entries. Specifically, if the decay property (1.2) holds for all $\boldsymbol{\alpha}, \boldsymbol{\beta} \le N+1$, a Taylor expansion of the kernel together with the cancellation properties of the primal wavelets shows

$$\langle A\psi_{\mathbf{lk}}, \psi_{\mathbf{l'k'}}\rangle_{\mathcal{L}^2} \le C \cdot \frac{2^{-|\mathbf{l},\mathbf{l'}|_1 (N+\frac{3}{2})}}{dist(supp(\psi_{\mathbf{lk}}), supp(\psi_{\mathbf{l'k'}}))^{f(N+1,N+1)}}, \tag{3.15}$$

for $dist(supp(\psi_{\mathbf{lk}}), supp(\psi_{\mathbf{l'k'}})) > 0$, see [23]. That is, the size of the entries depends on the distance of the supports of the basis functions $\psi_{\mathbf{lk}}$, $\psi_{\mathbf{l'k'}}$ and on their scales. The rate of the decay of the entries is governed by the number of vanishing moments. The following Lemma estimates the error of setting entries in the stiffness matrix to zero.

Lemma 2. *Let $(A_{\mathcal{I}} - A_{\mathcal{I}}^{comp})u \in \mathcal{H}^s$ und $u \in \mathcal{H}^{p,q}_{mix}$. Moreover let the norm equivalences (2.4) and (2.5) hold. The parameters $\tilde r, r$ from (2.4) and (2.5) shall fulfil $-\tilde r < s < r$ and $0 \le p < r, 0 \le p+q < r$. Then,*

$$\begin{aligned} \|(A_{\mathcal{I}} - A_{\mathcal{I}}^{comp})u\|^2_{\mathcal{H}^s} \tag{3.16} \\ \le C \cdot \left(\sum_{(\mathbf{j},\mathbf{k})\in\mathcal{I}} \sum_{(\mathbf{l},\mathbf{m})\in\tau_{\mathbf{jk}}\setminus\tau_{\mathbf{jk}}^{comp}} 2^{2s|\mathbf{j}|_\infty - 2q|\mathbf{k}|_\infty - 2p|\mathbf{k}|_1} a^2_{\mathbf{jklm}} \right) \cdot \|u\|^2_{\mathcal{H}^{p,q}_{mix}}. \end{aligned}$$

The proof of Lemma 2 is similar to the proof of Theorem 1. It uses again the biorthogonality of the primal and the dual wavelet and the stability of the dual wavelets in $\mathcal{H}^s$, see (2.5), and of the primal wavelets in $\mathcal{H}^{p,q}_{mix}$, see (2.4).
Now, we need to define $\tau_{\mathbf{jk}}^{comp}$ in such a way that the order of approximation does not deteriorate because of this compression. Assume for a moment that we have $\|(A - A_{\mathcal{I}})u\|_{\mathcal{H}^s} \simeq O(2^{-RJ})$ with some $R \in \mathbb{R}_0^+$. Then we require that $\|(A_{\mathcal{I}} - A_{\mathcal{I}}^{comp})u\|_{\mathcal{H}^s}$ is of the same order, i.e. $\|(A_{\mathcal{I}} - A_{\mathcal{I}}^{comp})u\|_{\mathcal{H}^s} \simeq O(2^{-RJ})$. According to Lemma 2 it is sufficient to require

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

The following Theorem tells us how to define the index sets $\tau_{\mathbf{jk}}^{comp}$. Note that the choice of the $\tau_{\mathbf{jk}}^{comp}$ has to balance the overall complexity and the error after compression in an optimal way.

THEOREM 2. *For* $(\mathbf{j},\mathbf{k}) \in \mathcal{I}$ *let*

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

where $J := \max\{|\mathbf{j},\mathbf{k}|_\infty : (\mathbf{j},\mathbf{k}) \in \mathcal{I}\}$. *We define*

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

Under the above assumptions ((3.15) and Lemma 2) we then have

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

Proof: Combination of (3.15), Lemma 2 and the Definition of $\tau_{\mathbf{jk}}^{comp}$ shows (for shortness we use the abbreviation $f = f(N+1, N+1)$)

$$\begin{aligned}
\|(A_{\mathcal{I}} - A_{\mathcal{I}}^{comp})u\|_{\mathcal{H}^s}^2 &\leq C \cdot \Bigg(\sum_{(\mathbf{j},\mathbf{k})\in\mathcal{I}} 2^{2s|\mathbf{j}|_\infty - 2q|\mathbf{k}|_\infty - 2p|\mathbf{k}|_1} 2^{-2|\mathbf{j},\mathbf{k}|_1(N+\frac{3}{2})} \\
&\quad \cdot \sum_{\substack{(\mathbf{l},\mathbf{m})\in\tau_{\mathbf{jk}},\\ dist(supp(\psi_{\mathbf{jl}}), supp(\psi_{\mathbf{km}}))>B_{\mathbf{jk}}}} dist(supp(\psi_{\mathbf{jl}}), supp(\psi_{\mathbf{km}}))^{-2f}\Bigg) \cdot \|u\|_{\mathcal{H}^{p,q}_{mix}}^2 \\
&\leq C \cdot \Bigg(\sum_{(\mathbf{j},\mathbf{k})\in\mathcal{I}} 2^{2s|\mathbf{j}|_\infty - 2q|\mathbf{k}|_\infty - 2p|\mathbf{k}|_1} 2^{-2|\mathbf{j},\mathbf{k}|_1(N+\frac{3}{2})} \cdot B_{\mathbf{jk}}^{-2f} \sum_{(\mathbf{l},\mathbf{m})\in\tau_{\mathbf{jk}}} 1\Bigg) \cdot \|u\|_{\mathcal{H}^{p,q}_{mix}}^2.
\end{aligned}$$

Now we apply $\sum_{(\mathbf{l},\mathbf{m})\in\tau_{\mathbf{jk}}} 1 = |\tau_{\mathbf{jk}}| \simeq 2^{|\mathbf{j},\mathbf{k}|_1}$ and obtain

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

□

The resulting stiffness matrix will in general be non-symmetric. A symmetric operator may be derived by taking the maximum of $B_{\mathbf{jk}}$ and $B_{\mathbf{kj}}$ in the definition of $\tau_{\mathbf{jk}}^{comp}$. At this point we need an estimate of the number of remaining non-zero elements in the compressed stiffness matrix. A full analysis of this problem is presently missing. Note however, it must be possible to obtain a compression with anisotropic wavelets which is optimal up to a logarithmic factor: Each compactly supported univariate scaling function on level j has a representation by $O(j)$ compactly supported univariate wavelets. Hence, each isotropic tensor product wavelet on level j has a representation by at most $O(j^{n-1})$ anisotropic tensor product wavelets. Therefore, the compression schemes of [10, 29] translate into compression schemes for anisotropic wavelets which result in stiffness matrices with $O(J^{2n-1}J^k 2^{Jn})$ entries.

4. Numerical examples. In this Section we present two numerical examples in 2D ($n = 2$) which show the efficiency of our scheme for the approximate integral evaluation and for the solution of integro-differential equations. Here we use orthogonal Daubechies wavelets on the interval with $N = \tilde{N} = 3$, see e.g. [7, 19]. Now, there is a certain coarsest level $j_0 = 4$ such that the approximation space V_{j_0} cannot be further decomposed into V_{j_0-1} and W_{j_0}. In the multivariate setting of our numerical experiments, this leads to some modifications in the definition of the index sets $\mathcal{I}_J$, $\mathcal{U}_{\mathcal{I}_J}$ for the approximation of K and u. Here, we use

$$(4.1)\quad \mathcal{I}_J := \{(\mathbf{j},\mathbf{k}) \in \mathbb{N}^4 : \ |\mathbf{j}|_1 + 2|\mathbf{k}|_1 \le 6j_0 + J\}, \quad \mathcal{U}_J := \{\mathbf{k} \in \mathbb{N}^2 : \ |\mathbf{k}|_1 \le 2j_0 + J\}$$

for the optimized scheme and

$$(4.2)\quad \mathcal{I}_J^f := \{(\mathbf{j},\mathbf{k}) \in \mathbb{N}^4 : \ |\mathbf{j},\mathbf{k}|_\infty \le j_0 + J\}, \quad \mathcal{U}_J^f := \{\mathbf{k} \in \mathbb{N}^2 : \ |\mathbf{k}|_\infty \le j_0 + J\}$$

for the classical full grid scheme. Except for the correction terms $6j_0$ and $2j_0$, the index sets $\mathcal{I}_J$ and $\mathcal{U}_J$ correspond to the index sets $\mathcal{I}_J(0,N,0,N,0)$, $\mathcal{U}_{\mathcal{I}_{2J}(0,N,0,N,0)}$ which are optimized for smooth K and u (where the order of approximation is limited only by N) and for $\|.\|_{\mathcal{L}^2}$-norm estimates of the error. In our experiments K and u are arbitrarily smooth, see subsections 4.1 and 4.2. Table 4.1 shows the dimension of the approximation spaces $\tilde{V}_{\mathcal{I}}$ for the different index sets. It holds $dim(\tilde{V}_{\mathcal{I}_J^f}) \simeq 16^J$ and $dim(\tilde{V}_{\mathcal{I}_J}) \simeq (J + C)^3 2^J$.

TABLE 4.1
Dimension of approximation spaces for $n = 2$ and the index sets from (4.1) and (4.2).

J	$\tilde{V}_{\mathcal{I}_J}$	$V_{\mathcal{U}_J}$	$\tilde{V}_{\mathcal{I}_J^f}$	$V_{\mathcal{U}_J^f}$
0	83521	289	83521	289
1	240737	833	1185921	1089
2	786369	2177	17850625	4225
3	2007105	5377	-	-
4	5272193	12801	-	-
5	12627073	29697	-	-

In the previous Sections the exact projections $K_{\mathcal{I}}$ and

$$u_{\mathcal{I}} := \sum_{\mathbf{j}\in\mathcal{U}_{\mathcal{I}}} \sum_{\mathbf{l}\in\tau_{\mathbf{j}}} u_{\mathbf{jl}}\psi_{\mathbf{jl}}$$

were used for the analysis. However, in numerical applications $K_{\mathcal{I}}$ and $u_{\mathcal{I}}$ are usually not explicitly known and have to be approximated by, say, $K_{\mathcal{I}}^{'}$ and $u_{\mathcal{I}}^{'}$. This is an essential part in applications and needs to be considered also for accuracy and complexity estimates. For full grid approximation spaces one usually employs tensor products of one-dimensional quadrature schemes, see [31], to approximate the scalar products of K and u with the scaling functions on the finest level of refinement. In all our experiments we used a quadrature scheme of order $M = 2N$. Then, by the wavelet transform one obtains the coefficients of $K_{\mathcal{I}}^{'}$ and $u_{\mathcal{I}}^{'}$. For the sparse grid spaces $\tilde{V}_{\mathcal{I}_J}$ we use Smolyak's blending scheme [2, 11, 17, 30]. Then, for example, for $u \in \mathcal{H}_{mix}^{M+1,0}$ and $\mathcal{U}_{\mathcal{I}_{2J}(0,N,0,N,0)} = \{\mathbf{k} : \ |\mathbf{k}|_1 \le J\}$, one can show

$$\|u_{\mathcal{I}_{2J}}^{'} - u_{\mathcal{I}_{2J}}\|_{\mathcal{L}^2} \le C J^{n-1} 2^{-(M+1/2)J} \|u\|_{\mathcal{H}_{mix}^{M+1,0}} \ .$$

For $K \in \mathcal{H}_{mix}^{M+1,0}$ and the index set $\mathcal{I}_J(0,N,0,N,0) = \{(\mathbf{j},\mathbf{k}) : |\mathbf{j}|_1 + 2|\mathbf{k}|_1 \le J\}$ we obtain

$$\|K'_{\mathcal{I}_J} - K_{\mathcal{I}_J}\|_{\mathcal{L}^2} \le C J^{2n-1} 2^{-(M+1/2)J/2} \|K\|_{\mathcal{H}_{mix}^{M+1,0}} \ .$$

Note, that this error is of order $O(J^{2n-1}2^{-(M+1/2)J/2})$. However, since $M = 2N$, the overall error is dominated by $\|(A - A_{\mathcal{I}_J})u\|_{\mathcal{H}^s}$ and not by the approximation of $K_{\mathcal{I}_J}$ and $u_{\mathcal{I}_{2J}}$. Similar estimates (without the logarithmic terms) hold for the full grid case.

For the calculation of errors the nodal values of the numerical solution $h^{\mathcal{I}}$ and the exact solution h are computed on a grid Ω with mesh size 2^{-12} for the x- and y-direction. Then, the exact (semi-) norms $\|h - h^{\mathcal{I}}\|_{\mathcal{L}^2}$ and $|h - h^{\mathcal{I}}|_{\mathcal{H}^1}$ are approximated using the discrete (semi-) norms

$$\|u\|_0^2 := 2^{-24} \sum_{\mathbf{x}\in\Omega} |u(\mathbf{x})|^2 \ , \quad |u|_1^2 := \frac{1}{4} \sum_{i=1}^{n} \sum_{\mathbf{x}\in\Omega} |u(\mathbf{x} + 2^{-12}\mathbf{e}_i) - u(\mathbf{x} - 2^{-12}\mathbf{e}_i)|^2 \ .$$

These discrete norms yield accurate estimates of the true norms of the error.

4.1. Integral evaluation. We consider the integral evaluation (1.1) with

$$K(\mathbf{x},\mathbf{y}) := \sin\Big(\prod_{i=1}^{2}(1-x_i)(1-y_i)\Big)\Big(1 + 20\big((y_1 - \frac{1}{3})^2 + (y_2 - \frac{1}{8})^2\big)\Big) \tag{4.3}$$

and

$$u(\mathbf{y}) := \frac{y_1(1-y_1)y_2(1-y_2)}{\Big(1 + 20\big((y_1 - \frac{1}{3})^2 + (y_2 - \frac{1}{8})^2\big)\Big)} \ , \quad \mathbf{x},\mathbf{y} \in [0,1]^2 \ . \tag{4.4}$$

The exact result is

$$h(\mathbf{x}) = \frac{X Si(X) + \cos(X) - 2 + 2Ci(X) - 2\gamma - 2\ln(X)}{X^3}, \tag{4.5}$$

where $X = (1-x_1)(1-x_2)$ and $\gamma = 0.577...$ is Euler's constant. Furthermore, Si and Ci denote the *Sinus* and *Cosinus Integralis*. See Figure 4.5 for a plot of u and h.

Figure 4.1 shows the (discrete) $\mathcal{L}^2$- and $\mathcal{H}^1$-norms of the error. The new scheme with index sets (4.1) yields approximately the same accuracy as the classical full grid scheme – with a much smaller work count, of course. Furthermore, the rate of error reduction agrees very well with the rate predicted by Theorem 1. Plots of error vs. number of degrees of freedom are given in Figure 4.2. The $\mathcal{L}^2$-optimized scheme leads to superior convergence rates in both norms. In all experiments the calculation of $a'_{\mathbf{jklm}}$ was by far the most expensive part of the algorithm. Table 4.2 lists cpu-times for the numerical quadrature, the wavelet transforms involved in the blending scheme and the matrix-vector product (3.10) for the $\mathcal{L}^2$-optimized scheme. The measurements were carried out on a SGI O2000 with a 195 MHz R10k processor. Due to the blending scheme the numerical work count and, therefore, the cpu-times are only asymptotically proportional to $dim(\tilde{V}_{\mathcal{I}_J})$. This explains the growth of the normalized cpu-time shown in the fifth column of Table 4.2.

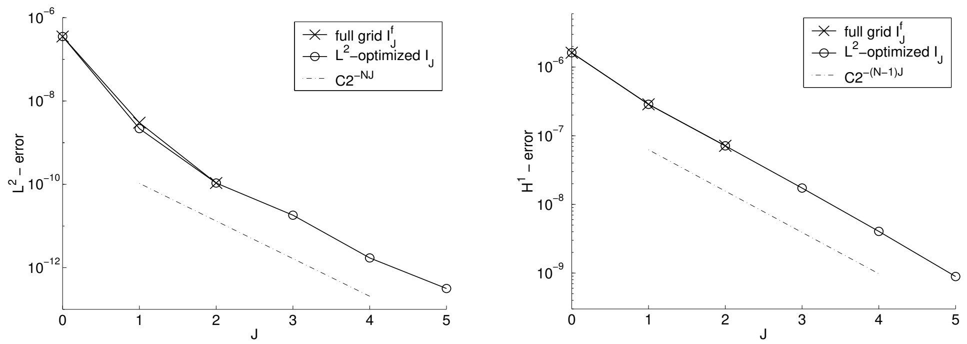

FIGURE 4.1. *Discrete $\mathcal{L}^2$- and $\mathcal{H}^1$-errors against the maximal level J for the integral evaluation.*

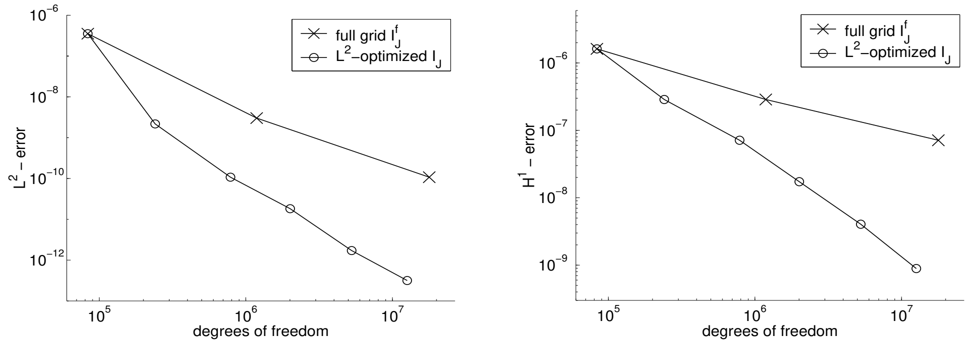

FIGURE 4.2. *Discrete $\mathcal{L}^2$- and $\mathcal{H}^1$-errors against $dim(\tilde{V}_{\mathcal{I}})$ for the integral evaluation.*

TABLE 4.2

$\mathcal{L}^2$-optimized scheme: cpu-times in [s] for the calculation of the coefficients $a'_{\mathbf{jklm}}$ (quadratures (Q), wavelet transforms (WT)), the multiplications (M) according to (3.10) and cpu-times for $Q+WT$ and M normalized with the dimensions of $\tilde{V}_{\mathcal{I}_J}$.

J	Q	WT	M	(Q + WT)/$dim(\tilde{V}_{\mathcal{I}_J})$	M/$dim(\tilde{V}_{\mathcal{I}_J})$
0	0.4	0.3	0.1	8.4e-6	1.43e-6
1	1.7	1.4	0.3	12.9e-6	1.44e-6
2	7.5	6.8	1.0	18.2e-6	1.33e-6
3	25.5	24.0	2.6	24.6e-6	1.32e-6
4	77.8	75.0	6.8	29.0e-6	1.28e-6
5	217.1	213.8	16.0	34.0e-6	1.27e-6

4.2. Solution of integro-differential equations. We consider the numerical solution of the integro-differential equation

$$-\Delta u(\mathbf{x}) + 1000 \int_{\mathbf{y}\in[0,1]^2} K(\mathbf{x},\mathbf{y})u(\mathbf{y})d\mathbf{y} = f(\mathbf{x}) \quad \text{in } (0,1)^2 , \qquad u = 0 \ \text{ on } \partial(0,1)^2 . \tag{4.6}$$

Here, K is given by (4.3) and the right hand side f is chosen such that the exact solution is (4.4), see Figure 4.5. The trial and test functions for the Galerkin discretization are Daubechies wavelets with homogeneous Dirichlet boundary conditions

and $N = 3$ vanishing moments. A BiCGStab(2) iterative solver was used for the solution of the resulting linear system. The number of iterations required to reduce the residual to machine precision was between 40 and 50 (essentially independent of the maximal level J).
As shown in Figure 4.3 the optimized scheme yields approximately the same accuracy as the classical scheme and the theoretical convergence rates are matched well. Again the new scheme leads to dramatically reduced errors when compared to the full grid scheme, see Figure 4.4.

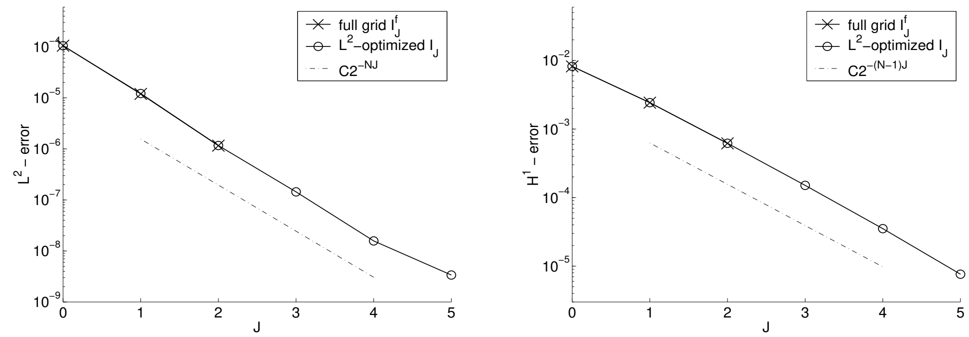

FIGURE 4.3. *Discrete $\mathcal{L}^2$- and $\mathcal{H}^1$-errors against the maximal level J for the solution of (4.6).*

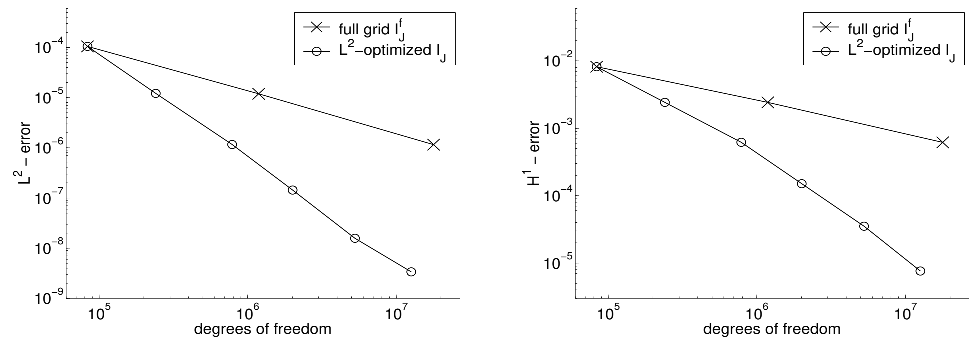

FIGURE 4.4. *Discrete $\mathcal{L}^2$- and $\mathcal{H}^1$-errors against $dim(\tilde{V}_{\mathcal{I}})$ for the solution of (4.6).*

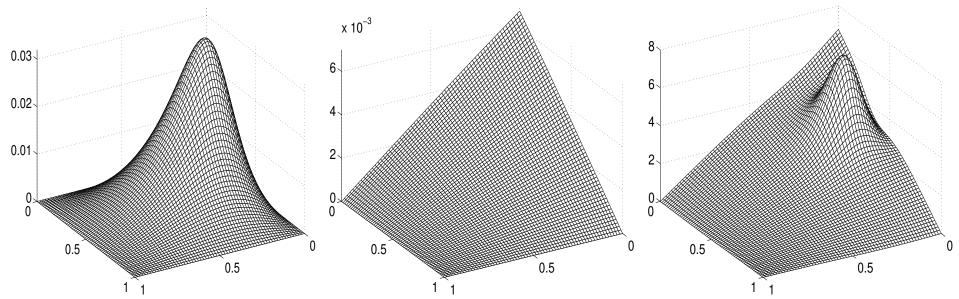

FIGURE 4.5. *Left: u, (4.4). Middle: h, (4.5). Right: f, (4.6)*

Acknowledgements: The authors would like to thank the referees for their useful comments. The second author has been supported by the Deutsche Forschungsgemeinschaft, GR 1144/7-2.

REFERENCES

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

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

[3] S. Bertoluzza, C. Canuto, K. Urban, *On the adaptive computation of integrals of wavelets*, Appl. Numer. Math. 34 (2000), 13-38.

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

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

[6] J. C. Cavendish, W. J. Gordon, C. H. Hall, *Ritz-Galerkin Approximations in Blending Function Spaces*, Numer. Math. 26 (1976), 155-178.

[7] A. Cohen, I. Daubechies, P. Vial, *Wavelets on the interval and fast wavelet transforms*, Appl. Comput. Harm. Anal. 1 (1993), 54-81.

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

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

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

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

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

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

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

[15] K. Frank, *Complexity of local solution of multivariate integral equations*, J. Complexity 11 (1995), 416-434.

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

[17] T. Gerstner, M. Griebel, *Numerical Integration using Sparse Grids*, Numerical Algorithms 18 (1998), 209-232.

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

[19] M. Griebel, F. Koster, *Orthogonal Wavelets on the Interval*, SFB-256 Bericht Nr. 576, Universität Bonn, 1998.

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

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

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

[23] S. Knapek, *Approximation und Kompression mit Tensorprodukt-Multiskalen-Räumen*, Dissertation, Inst. für angew. Math., Universität Bonn, 2000.

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

[25] 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.

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

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

[28] C. Pflaum, A. Zhou, *Error analysis of the combination technique*, Numer. Math. 84 (1999), 327-350.

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

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

[31] W. Sweldens, R. Piessens, *Quadrature formulae and asymptotic error expansions for wavelet approximations of smooth functions*, SIAM J. Numer. Anal. 32 (1994), 1240-1264.

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

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

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