Preconditioners for Sparse Grid Discretizations

Frank Koster, Bonn

Abstract

In this paper we deal with preconditioners for sparse grid finite-difference– and Petrov-Galerkin–discretizations of the Poisson equation. We analyse the Jacobi-preconditioner for the simple setting of non-adaptive grids and periodic boundary conditions. The analysis shows that the resulting condition numbers mainly depend on the underlying tensor product Wavelets. For example, high order Lifting-Interpolets lead to l_2-condition numbers which are essentially independent of the finest mesh size. Based on this observation we introduce a so-called Lifting-preconditioner for discretizations which use Interpolets as trial-functions. Numerical examples show the efficiency of the preconditioners for cases which are not covered by our analysis, e.g., adaptive grids.

Key words. finite differences, Petrov-Galerkin, sparse grids, Wavelets, Interpolets, Lifting
AMS subject classification. 33F05, 41A25, 65B05, 65D05, 65M06, 65D25

1 Introduction

Hyperbolic tensor product Wavelets allow for a very efficient non-linear approximation of functions with bounded mixed derivatives [11, 30] or anisotropic smoothness [18, 22]. But, also functions from general Besov-spaces may be approximated efficiently [18]. Thus, it is quite appealing to use these Wavelets as trial-functions for the adaptive discretization of partial differential equations (PDEs). Such a method was introduced in [32], where tensor products of piecewise linear hat functions were used for a Galerkin-scheme. Later, this approach was extended to piecewise polynomials of a higher degree [2]. The finite difference method on adaptive sparse grids [10, 16, 26] is another discretization scheme based on tensor product Wavelets. The main advantages of this method over the Galerkin scheme are the much cheaper operator evaluation and the simplicity with which special discretizations, e.g., WENO schemes for convection problems, may be incorporated [20]. Originally, the finite difference (FD) method was built on tensor products of piecewise linear hat functions. It was then extended in [13, 19] to tensor products of high order generalizations of the hat function, so-called Interpolets [8, 9]. The high order functions are smoother than the hat functions. Thus, they also allow for Petrov-Galerkin–schemes, where the dual Wavelets are used as test-functions. In some cases these Petrov-Galerkin–schemes are nothing but FD schemes with special stencils [19].

In this paper we deal with the efficient preconditioning of linear systems which arise from the FD/Petrov-Galerkin–discretization of the Poisson-equation

$$-\Delta u = f \quad \text{in} \quad]0,1[^d \tag{1}$$

on adaptive sparse grids. The respective coefficient matrices have a quite different structure in comparison to the Galerkin-stiffness matrices. For example, they are much sparser and they are always non-symmetric. Consequently, the well-known statements on preconditioners for Wavelet-Galerkin–schemes [14] do not directly carry over to FD discretizations. Some steps towards efficient preconditioners were made in [26]. A multigrid like approach has been considered in [27]. However, these results left much space for improvements. We follow a different approach here. We investigate the Jacobi-preconditioner by means of some modifications of the analysis for the Galerkin case [14]. Like there, the estimates rest upon the Kronecker product structure of the coefficient matrix and norm-equivalences for the Sobolev-spaces $\mathcal{L}^2$ and $\mathcal{H}^1$. Since the Interpolets do not fulfill the norm-equivalences for $\mathcal{L}^2$ in the strict sense, we obtain condition numbers which depend on the maximal level of refinement. However, certain modifications of the Interpolets, the Lifting-Interpolets, fulfill the norm-equivalences for $\mathcal{L}^2$ and $\mathcal{H}^1$. Therefore, we get optimal condition numbers in this case. Based on this observation we introduce in Section 4 a new preconditioner which consists of a basis transform from/to Lifting-Interpolets and the Jacobi preconditioner. With this technique optimal condition numbers are possible even for the Interpolet-based discretizations.

This paper is organized as follows. In the next section we briefly introduce the FD- and the Petrov-Galerkin–schemes for general underlying univariate

Wavelets. The analysis of the Jacobi-preconditioner is given in Section 3 for the special case of (Lifting-)Interpolets, non-adaptive grids, and periodic boundary conditions. In Section 4 we describe the Lifting-preconditioner and some algorithmical details of the involved basis transforms. We finish this paper with Section 5, where numerical experiments are presented. Here, the preconditioners are used within a BiCGstab iterative solver [34]. They also seem to be very effective for situations which are not covered by our analysis, e.g., for adaptive grids.

2 Biorthogonal Petrov-Galerkin– and finite-difference–discretizations on sparse grids

We define some generic notation for general biorthogonal Wavelets first and then describe the Petrov-Galerkin– and FD-schemes. By $\mathcal{H}^t(\Omega), t \in \mathbb{R}$, we denote Sobolev-spaces over a domain Ω. $\mathcal{L}^2(\Omega)$ is the space of square-integrable functions. The respective norms are $\|.\|_{\mathcal{H}^t}$. The $\mathcal{H}^1$–semi-norm is $|.|_{\mathcal{H}^1}$. Vectors and multiindices are written boldface with small letters, for example, $\mathbf{l} = (l_1, ..., l_d)$. For such vectors or multiindices we set $|\mathbf{l}|_1 := \sum_i |l_i|$ and $|\mathbf{l}|_2 := (\sum_i l_i^2)^{1/2} = (\mathbf{l}^t\mathbf{l})^{1/2}$. Especially for multiscale functions it is convenient to index them with tuples like (l, s). $l \in \mathbb{N}$ is the level index and $s \in \mathbb{N}$ is a spatial index. We will use such generalized indices also for the entries of the stiffness matrix. For example, if $A = (A_{l,t,k,s})_{(l,t),(k,s)\in\mathcal{T}}$ is the stiffness matrix with respect to a finite-dimensional multiscale basis defined by the index set $\mathcal{T}$ and $\mathbf{u} = (u_{k,s})_{(k,s)\in\mathcal{T}}$ is a compatible vector, then $A\mathbf{u}$ is the vector with components $(A\mathbf{u})_{l,t} = \sum_{(k,s)\in\mathcal{T}} A_{l,t,k,s}u_{k,s}$. In a similar way, the l_2-norm of a vector $\mathbf{u}$ may be written as $|\mathbf{u}|_2 := (\sum_{(k,s)\in\mathcal{T}} u_{k,s}u_{k,s})^{1/2} = (\mathbf{u}^t\mathbf{u})^{1/2}$.
We write $X \lesssim Y$ if there exists a constant $c > 0$, independent of any parameters X or Y may depend on, such that $X \leq cY$. $X \sim Y$ means that $X \lesssim Y$ and $Y \lesssim X$.

2.1 Biorthogonal Wavelets

Let $V^{l_0} \subset V^{l_0+1} \subset ... \subset \mathcal{L}^2([0, 1])$ be a sequence of approximation spaces which are spanned by univariate scaling functions $\phi^{l,s}$, i.e. $V^l = \text{span}\{\phi^{l,s}\}_{s\in S(l)}$. Here, l is the level of refinement, l_0 is a certain coarsest level which results from the construction of scaling functions on the interval, compare e.g. [12], and $S(l)$ is the index set $\{0, ..., 2^l\}$. If the scaling functions shall satisfy homogeneous Dirichlet- or Neumann-boundary conditions in 0 and 1, then $S(l)$ is $\{1, ..., 2^l - 1\}$, see [12]. The Wavelets $\{\psi_{l,t}\}_{t\in T(l)}$, $T(l) := \{0, ..., 2^{l-1} - 1\}$, are bases of complementary spaces W_l of V^{l-1} in V^l for $l > l_0$, i.e. $V^l = V^{l-1} \oplus W_l$. For reasons of a simpler notation, we extend the definition of the Wavelets to the coarsest level and set $\psi_{l_0,t} := \phi^{l_0,t}$ for $t \in T(l_0) := S(l_0)$. Except for boundary modifications, the scaling functions and the Wavelets are supposed to be dilates and translates of single mother functions. Furthermore we assume that they are normalized to $\|\phi^{l,s}\|_{\mathcal{L}^2} = \|\psi_{l,t}\|_{\mathcal{L}^2} = 1$.

Associated to the primal scaling functions and Wavelets are dual scaling functions $\{\tilde{\phi}^{l,s}\}_{s\in S(l)}$ and dual Wavelets $\{\tilde{\psi}_{l,t}\}_{t\in T(l)}$ which satisfy the biorthogonality relations

$$\begin{aligned} \langle\tilde{\phi}^{l-1,s}, \phi^{l-1,t}\rangle &= \delta_{s,t} \quad , \quad \langle\tilde{\phi}^{l-1,s}, \psi_{l,t}\rangle = 0 \ , \\ \langle\tilde{\psi}_{l,s} \quad , \psi_{l,t} \quad \rangle &= \delta_{s,t} \quad , \quad \langle\tilde{\psi}_{l,s}, \phi^{l-1,t}\rangle = 0 \end{aligned} \tag{2}$$

for $l > l_0$. If both the primal and the dual functions are in $\mathcal{L}^2$ then $\langle .,. \rangle$ denotes the $\mathcal{L}^2$-inner product. However, the dual functions of Interpolets are linear combinations of the Dirac-functional. In this case $\langle .,. \rangle$ denotes the dual pairing.

For later explanations it is convenient to define the following row vectors of basis functions

$$\mathbf{\Phi}^l := (\phi^{l,s})_{s\in S(l)} \ , \quad \mathbf{\Psi}_l := (\psi_{l,t})_{t\in T(l)} \ \text{ and } \ \mathbf{\Psi}^l := (\mathbf{\Psi}_{l_0}, ..., \mathbf{\Psi}_l) \ .$$

The corresponding coefficient vectors (columns) are written as

$$\mathbf{u}^l := \Big((u^{l,s})_{s\in S(l)}\Big)^t \quad \text{and} \quad \mathbf{u}_l := \Big((u_{l,t})_{t\in T(l)}\Big)^t,$$

respectively. The vectors $\tilde{\mathbf{\Phi}}^l$, $\tilde{\mathbf{\Psi}}_l$ and $\tilde{\mathbf{\Psi}}^l$ are defined in the same way. Note that we use superscript indices for the scaling functions ($\sim$coefficients) and subscript indices for the Wavelets ($\sim$coefficients). Multivariate Wavelets for the domain $[0,1]^d$ are obtained by tensor products:

$$\psi_{\mathbf{l},\mathbf{t}}(\mathbf{x}) := \prod_{i=1}^{d} \psi_{l_i,t_i}(x_i) \ .$$

The respective dual Wavelets are $\tilde{\psi}_{\mathbf{l},\mathbf{t}}(\mathbf{x}) := \prod_{i=1}^{d} \tilde{\psi}_{l_i,t_i}(x_i)$.

2.2 Petrov-Galerkin– and Finite-Difference–Schemes

For the biorthogonal Petrov-Galerkin–scheme a finite dimensional subspace $\mathrm{span}\{\psi_{\mathbf{l},\mathbf{t}}\}_{(\mathbf{l},\mathbf{t})\in\boldsymbol{\mathcal{T}}}$ is used as trial-space and the corresponding dual Wavelets $\{\tilde{\psi}_{\mathbf{l},\mathbf{t}}\}_{(\mathbf{l},\mathbf{t})\in\boldsymbol{\mathcal{T}}}$ are used as test-functions. We assume that trial- and test-functions are sufficiently smooth such that the method is well-defined. The index set $\boldsymbol{\mathcal{T}}$ which actually determines the trial- and test-spaces is supposed to be given. The problem of finding a proper index set, i.e. the adaptive refinement strategy, is not in the scope of this paper.
The Petrov-Galerkin approach leads to a linear system of equations

$$\mathbf{A}\mathbf{u} = \mathbf{f} \ , \tag{3}$$

where $\mathbf{A} := (\langle\tilde{\psi}_{\mathbf{l},\mathbf{t}}, -\Delta\psi_{\mathbf{k},\mathbf{s}}\rangle)_{(\mathbf{l},\mathbf{t}),(\mathbf{k},\mathbf{s})\in\boldsymbol{\mathcal{T}}}$ is the stiffness matrix and $\mathbf{u} = (u_{\mathbf{l},\mathbf{t}})_{(\mathbf{l},\mathbf{t})\in\boldsymbol{\mathcal{T}}}$ and $\mathbf{f} = (f_{\mathbf{l},\mathbf{t}})_{(\mathbf{l},\mathbf{t})\in\boldsymbol{\mathcal{T}}}$ are the Wavelet coefficients of the numerical solution and the right hand side, respectively. Matrix $\mathbf{A}$ is very sparse and the matrix-vector multiplication essentially boils down to a number of matrix-vector multiplies with univariate stiffness matrices like $\langle\tilde{\mathbf{\Psi}}^{lt}, \partial_{xx}\mathbf{\Psi}^{l}\rangle$. The reason is the biorthogonality (2) of trial- and test-functions: $\mathbf{A}$ is the sum of matrices $(\langle\tilde{\psi}_{\mathbf{l},\mathbf{t}}, \partial_{x_i x_i}\psi_{\mathbf{k},\mathbf{s}}\rangle)_{(\mathbf{l},\mathbf{t}),(\mathbf{k},\mathbf{s})\in\boldsymbol{\mathcal{T}}}$ for $i = 1...d$. However, in these matrices $(\mathbf{l},\mathbf{t})$ and $(\mathbf{k},\mathbf{s})$ are coupled only if $l_j = k_j$ and $s_j = t_j$, for $j \neq i$, since

$$\begin{aligned} \langle\tilde{\psi}_{\mathbf{l},\mathbf{t}}, \partial_{x_i x_i}\psi_{\mathbf{k},\mathbf{s}}\rangle &= \langle\tilde{\psi}_{l_i,t_i}, \partial_{x_i x_i}\psi_{k_i,s_i}\rangle \prod_{j\neq i}\langle\tilde{\psi}_{l_i,t_i}, \psi_{k_i,s_i}\rangle \\ &= \langle\tilde{\psi}_{l_i,t_i}, \partial_{x_i x_i}\psi_{k_i,s_i}\rangle \prod_{j\neq i}\delta_{l_i,k_i}\delta_{t_i,s_i} \ . \end{aligned}$$

This saves a lot of computational work and explains the interest in such Petrov-Galerkin–schemes. With some modifications, these simplifications carry over to more general operators [20]. If the index set $\boldsymbol{\mathcal{T}}$ has the following simple form

$$\boldsymbol{\mathcal{T}} = \{(\mathbf{l},\mathbf{t}) : \ \mathbf{t} \in T(l_1) \times ... \times T(l_d) \ \text{ for } \ |\mathbf{l}|_\infty \leq L\} \ , \tag{4}$$

i.e. $\boldsymbol{\mathcal{T}}$ corresponds to a uniform refinement up to level L, then $\mathbf{A}$ reads

$$\mathbf{A} = \sum_{i=1}^{d} \underbrace{I \otimes .. \otimes I}_{i-1 \text{ times}} \otimes A \otimes \underbrace{I \otimes .. \otimes I}_{d-i \text{ times}} \ . \tag{5}$$

Here, I is the $|S(L)| \times |S(L)|$ identity matrix, $A := \langle \tilde{\mathbf{\Psi}}^{Lt}, \partial_{xx} \mathbf{\Psi}^L \rangle$ is the stiffness matrix for the univariate case and $\otimes$ denotes the Kronecker-product of matrices [33]. There is a close relation of this Petrov-Galerkin scheme to FD schemes: Let T and $\tilde{T}$ be the matrix representations of the inverse Wavelet transform for the primal and dual Wavelets which are defined by $\mathbf{\Phi}^L T = \mathbf{\Psi}^L$ and $\tilde{\mathbf{\Phi}}^L \tilde{T} = \tilde{\mathbf{\Psi}}^L$. Because of the biorthogonality relations, there holds $I = \langle \tilde{\mathbf{\Psi}}^{Lt}, \mathbf{\Psi}^L \rangle = \tilde{T}^t \langle \tilde{\mathbf{\Phi}}^{Lt}, \mathbf{\Phi}^L \rangle T = \tilde{T}^t T$ and, thus,

$$\tilde{T}^t = T^{-1} . \tag{6}$$

Therefore, A takes the form

$$A = \tilde{T}^t B T = T^{-1} B T , \qquad \text{where} \qquad B = \langle \tilde{\mathbf{\Phi}}^{Lt}, \partial_{xx} \mathbf{\Phi}^L \rangle . \tag{7}$$

Now, B is in some sense the matrix representation of a special FD scheme with stencils in the nodes $2^{-L} S(L)$. That means the matrix-vector multiply by A is nothing else but an inverse Wavelet transform (T) which yields nodal values of the given function, then the application of a FD scheme (B) and finally a Wavelet transform (T^{-1}). The order of consistency of the FD scheme for B depends on the underlying scaling functions, see for example [19]. However, this scheme is not very attractive in general, since the stencil width is larger than that for the classical central FD schemes of the same order of consistency.

At this point the FD method on sparse grids comes into play. The main idea is to replace $\langle \tilde{\mathbf{\Phi}}^{Lt}, \partial_{xx} \mathbf{\Phi}^L \rangle$ in (7) by another matrix which corresponds to a more efficient FD scheme. That means, the FD stiffness matrix has exactly the structure of (5) and (7) whenever simple index sets like (4) are used. Therefore, we can give a unified analysis of the Jacobi preconditioner in the next section which covers the Petrov-Galerkin–scheme as well as the FD discretization.

For general index sets, i.e. the fully adaptive case, the FD scheme on sparse grids also consists of the basic steps: inverse Wavelet transform with respect to one coordinate direction only, application of the underlying FD scheme to the nodal values on univariate adaptive grids and a Wavelet transform. A more detailed description would be rather technical and is beyond the scope of the present paper. Here we refer the reader to e.g. [10, 13, 16, 26].

3 Analysis of the Jacobi preconditioner

In this section we estimate the l_2-condition number κ of $\mathbf{AD}^{-1}$, where $\mathbf{D}$ is the diagonal matrix with entries $D_{(\mathbf{l},\mathbf{t}),(\mathbf{l},\mathbf{t})} = \sum_{i=1}^{d} 4^{l_i}$. To this end, we give lower and upper bounds for factors $c(L)$ and $C(L)$ which fulfill

$$c(L)|\mathbf{Du}|_2 \;\le\; |\mathbf{Au}|_2 \;\le\; C(L)|\mathbf{Du}|_2 \quad \forall \mathbf{u}\ . \tag{8}$$

Obviously, then there holds $\kappa \le C(L)/c(L)$. Our analysis relies on the Kronecker-product structure (5) of $\mathbf{A}$. Hence, we consider the simple index sets of (4) only. Furthermore, we assume periodic boundary conditions for the model problem (1). In this case we may assume that B is a symmetric Toeplitz matrix. Another prerequisite we need is the stability of the underlying discretization in the form

$$\mathbf{u}^t B\mathbf{u} \;\sim\; \mathbf{u}^t\langle \partial_x \mathbf{\Phi}^{Lt}, \partial_x \mathbf{\Phi}^{L}\rangle \mathbf{u} \;=\; \int_0^1 |\partial_x \mathbf{\Phi}^L \mathbf{u}|^2 dx \quad \forall \mathbf{u}\ . \tag{9}$$

This condition is easy to verify for a given discretization. For periodic boundary conditions all scaling functions are simple dilates and translates of a single mother function. Hence, both $M := \langle \partial_x \mathbf{\Phi}^{Lt}, \partial_x \mathbf{\Phi}^{L}\rangle$ and B are circulant Toeplitz matrices. Therefore, B and M have the same orthogonal eigenvectors. The respective eigenvalues $\{\lambda_k\}_{0\le k<2^L}$ and $\{\mu_k\}_{0\le k<2^L}$ can be evaluated independent on the level L using the matrix symbols. The stability condition holds if $\lambda_k \sim \mu_k \sim k^2$ for $0 \le k < 2^{L-1}$. As an example, we plot in Figure 1 the quotients λ_k/k^2 and μ_k/k^2, $0 < k < 2^{L-1}$, for Interpolets of order 6 and the stiffness matrices of the Petrov-Galerkin–scheme and a fourth order central FD scheme, respectively. As usual the wave number k is scaled by $2\pi/2^L$ to get ordinates which are independent of L. For the biorthogonal Petrov-Galerkin scheme and most common FD schemes (9) is satisfied.

The bounds for $c(L)$ and $C(L)$ are shown by means of one-sided norm-estimates for the univariate primal and dual Wavelets, respectively. The range of Sobolev-spaces for which such norm estimates hold is determined by the Sobolev-regularity γ and by the polynomial exactness N of these basis functions:

$$\begin{aligned} \gamma &:= \sup\{s : \psi_{l,t} \in \mathcal{H}^s \ \text{ for all } \ (l,t)\} \\ N &:= \max\{k \in \mathbb{N} : \ \text{each polynomial of degree less than } k \\ &\qquad\qquad \text{is in } \operatorname{span}\{\phi^{l_0,t}\}_{t\in S(l_0)}\}\ . \end{aligned}$$

The parameters $\tilde\gamma$ and $\tilde N$ are defined analogously for the dual functions.

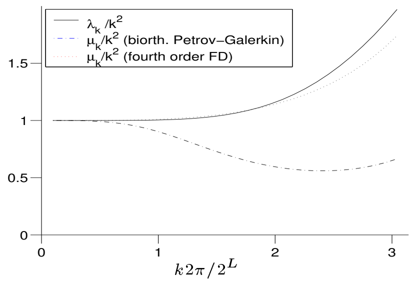


Figure 1: *Quotients λ_k/k^2 and μ_k/k^2 for Interpolets $N = 6$ and various discretizations.*

Theorem 3.1 (One-sided norm estimates [25]). *For $-\tilde{N} < s < \gamma$ there are constants c_s, such that for all $L \geq l_0$ and all $u = \sum_{l=l_0}^{L} \mathbf{\Psi}_l \mathbf{u}_l$ there holds*

$$\|u\|_{\mathcal{H}^s}^2 \;\leq\; c_s \sum_{l=l_0}^{L} 4^{ls} |\mathbf{u}_l|_2^2 \;.$$

For $s = -\tilde{N}$ there holds

$$\|u\|_{\mathcal{H}^{-\tilde{N}}}^2 \;\leq\; c_{-\tilde{N}} L \sum_{l=l_0}^{L} 4^{-l\tilde{N}} |\mathbf{u}_l|_2^2 \;.$$

Similar relations hold for the dual Wavelets and $-N \leq s < \tilde{\gamma}$.

By means of duality arguments [25] one derives the following norm equivalences.

Theorem 3.2 (Norm equivalences). *For $-\tilde{\gamma} < s < \gamma$ there holds*

$$\|u\|_{\mathcal{H}^s}^2 \;\sim\; \sum_{l=l_0}^{L} 4^{ls} |\mathbf{u}_l|_2^2 \;.$$

Two particular univariate Wavelet families are especially interesting for Petrov-Galerkin or FD schemes on sparse grids. The first family are Interpolets with an even order N of polynomial exactness [8]. The most simple Interpolet basis ($N = 2$) is the well-known piecewise linear hierarchical basis. For tensor products of Interpolets it is very simple to compute nodal values of e.g. $u = \sum_{\mathbf{l}} \mathbf{\Psi}_{\mathbf{l}} \mathbf{u}_{\mathbf{l}}$ on certain adaptive grids from the Wavelet coefficients. It is also simple to recover the Wavelet coefficients from these nodal values, i.e. to find the unique Wavelet interpolant. This allows for a fast approximative evaluation of non-linear terms.

However, the dual Wavelets to Interpolets are linear combinations of the Dirac-functional [13]. Thus, $\tilde{N} = 0$ and $\tilde{\gamma} = -0.5$. Hence, the norm estimates of Theorems 3.1 and 3.2 are valid for a relatively small range of Sobolev spaces only.

The second Wavelet family considered in this paper does not have these problems. The so-called Lifting-Interpolets are essentially Interpolets, but modified by coarse level scaling functions [3, 28]. For a given Interpolet basis $\{\psi_{l,t}\}$ they are defined by

$$\overline{\psi}_{l,t} := \psi_{l,t} - \frac{\sqrt{2}}{4}(\phi^{l-1,t} + \phi^{l-1,t+1}) \quad \text{for } l > l_0 \tag{10}$$

and $\overline{\psi}_{l_0,t} := \phi^{l_0,t}$. The idea of this definition is to enforce vanishing first and second moments of the Wavelets, i.e. $\int \overline{\psi}_{l,t}(x)dx = \int \overline{\psi}_{l,t}(x)xdx = 0$ for $l > l_0$. It is well-known that changing the primal Wavelets also leads to new dual Wavelets [5]. For the Lifting-Interpolets, these are no longer Diracs, but more regular functions. Their order of polynomial exactness is 2 and they are in $\mathcal{H}^s$ for some $s > 0$. In Table 1 numerical estimates for the Sobolev regularity of the (Lifting-)Interpolets and the associated dual Wavelets are shown. The values were computed with the techniques of [7]. Because of the smoother duals, the Lifting-Interpolets allow for a much larger range of norm equivalences than Interpolets.

primal (γ)				dual ($\tilde{\gamma}$)			
N	2	4	6	N	2	4	6
Interpolets	1.5	2.171	2.755	Interpolets	-0.5	-0.5	-0.5
Lifting-I.	"	"	"	Lifting-I.	0.171	0.290	0.358

Table 1: *Sobolev-regularity of (Lifting-)Interpolets and their dual Wavelets.*

Next, we give bounds for the condition numbers of some building blocks of $\mathbf{A}^t\mathbf{A}$. These bounds are consequences of Theorems 3.1 and 3.2 and the stability (9) of B.

Theorem 3.3. *Let $\mathbf{\Psi}^L$ be an Interpolet- or Lifting-Interpolet basis and D be the diagonal matrix with entries $D_{(l,t),(l,t)} = 4^l$. Then* *(i) there are factors* $\underline{a}(L) \leq \overline{a}(L)$ *such that*

$$\underline{a}(L)\mathbf{u}^t\mathbf{u} \ \leq \ \mathbf{u}^t T^t T\mathbf{u} \ \leq \ \overline{a}(L)\mathbf{u}^t\mathbf{u} \ .$$

For Interpolets the bounds $2^{-L} \lesssim \underline{a}(L)$ *and* $\overline{a}(L) \lesssim L$ *hold, while for Lifting-Interpolets we have* $\underline{a}(L) \sim \overline{a}(L) \sim 1$.

(ii) there holds $\mathbf{u}^t D\mathbf{u} \sim \mathbf{u}^t T^t BT\mathbf{u}$ *for coefficient vectors* $\mathbf{u} = (\mathbf{u}_{l_0}^t, ..., \mathbf{u}_L^t)^t$ *with*

$$\int_0^1 \sum_{l=l_0}^{L} (\mathbf{\Psi}_l \mathbf{u}_l)(x)dx = 0 \ . \tag{11}$$

(iii) there are factors $\underline{b}(L) \leq \overline{b}(L)$ *such that*

$$\underline{b}(L)\mathbf{u}^t D^2\mathbf{u} \ \leq \ \mathbf{u}^t(BT)^t(BT)\mathbf{u} \ \leq \ \overline{b}(L)\mathbf{u}^t D^2\mathbf{u} \tag{12}$$

for all coefficients $\mathbf{u}$ *with (11). If* $N \geq 4$ *then* $\underline{b}(L) \sim \overline{b}(L) \sim 1$. *For* $N = 2$ *we have*

$$L^{-1} \lesssim \underline{b}(L) \quad \text{and} \quad \overline{b}(L) \lesssim 4^{(2-\gamma+\epsilon)L} \quad \text{for all } \epsilon > 0 \ .$$

(iv) there are factors $\underline{c}(L) \leq \overline{c}(L)$ *such that*

$$\underline{c}(L)\mathbf{u}^t D^2\mathbf{u} \ \leq \ \mathbf{u}^t(T^{-1}BT)^t(T^{-1}BT)\mathbf{u} \ \leq \ \overline{c}(L)\mathbf{u}^t D^2\mathbf{u} \ ,$$

for all coefficients $\mathbf{u}$ *with (11). The factors are bounded by* $\underline{b}(L)/\overline{a}(L) \leq \underline{c}(L)$ *and* $\overline{c}(L) \leq \overline{b}(L)/\underline{a}(L)$.

Condition (11) says that the coefficient vectors are orthogonal to the null space of the discrete operators. This constraint is necessary, since D is positive definite. The proof for Theorem 3.3 is not really needed for an understanding of the further explanations. Therefore, we postpone it to the end of this section. In Table 2 numerical estimates for $\underline{a}(L), ..., \overline{c}(L)$ for piecewise linear Interpolets with $N = 2$ and the respective Lifting-Interpolets are shown. The numerical values have been computed by means of the eigenvalues of the operator matrices using MatLab. In order to check the agreement with the estimates of Theorem 3.3 we show also the accordingly re-scaled values. Obviously the agreement is quite good.

Interpolets $N = 2$

L	$\underline{a}2^L$	$\overline{a}/L$	$\underline{b}L$	$\overline{b}/2^L$	$\underline{c}L^2$	$\overline{c}/4^L$
4	2.71	0.55	8.66	4.13	1.88	1.17
6	2.55	0.54	8.49	4.22	1.80	1.30
8	2.52	0.53	8.34	4.25	1.74	1.33
10	2.51	0.52	8.24	4.25	1.70	1.34

Lifting-Interpolets $N = 2$

L	$\underline{a}$	$\overline{a}$	$\underline{b}L$	$\overline{b}/2^L$	$\underline{c}L$	$\overline{c}/2^L$
4	0.34	1.00	7.56	4.21	9.8	11.4
6	0.18	1.00	7.60	4.33	10.4	18.2
8	0.13	1.00	7.64	4.36	10.5	22.6
10	0.11	1.00	7.66	4.36	10.5	24.8

Table 2: *Equivalence factors for the 1D operators, re-scaled according to Theorem 3.3.*

For the proof of (8) we use the preliminaries of Theorem 3.3 and the Kronecker-product structure of $\mathbf{A}$, see (5), and $\mathbf{D}$ which reads

$$\mathbf{D} = \sum_{i=1}^{d} \underbrace{I \otimes .. \otimes I}_{i-1 \text{ times}} \otimes D \otimes \underbrace{I \otimes .. \otimes I}_{d-i \text{ times}} \ .$$

The link between the two aspects are estimates for the field of values of Kronecker-products of symmetric matrices by the field of values of the factors, see [33]. Here we need the following modification:

Lemma 3.4. *Let M_1 and M_2 be symmetric, positive semi-definite $n \times n$ matrices whose null space is spanned by $\mathbf{m}_1$ and $\mathbf{m}_2$, i.e. $M_i\mathbf{m}_i = 0$ for $i = 1, 2$. The smallest and the largest non-vanishing eigenvalues of M_i are $\underline{\alpha}_i$ and $\overline{\alpha}_i$, respectively. Then, for all $\mathbf{u} \in \mathbb{R}^{n\times n}$ with*

$$\mathbf{u}^t(\mathbf{m}_1 \otimes \mathbf{q}) = \mathbf{u}^t(\mathbf{q} \otimes \mathbf{m}_2) = 0 \quad \forall \ \ \mathbf{q} \in \mathbb{R}^n$$

there holds

$$\underline{\alpha}_1\underline{\alpha}_2\mathbf{u}^t\mathbf{u} \ \leq \ \mathbf{u}^t(M_1 \otimes M_2)\mathbf{u} \ \leq \ \overline{\alpha}_1\overline{\alpha}_2\mathbf{u}^t\mathbf{u} \ .$$

The proof follows the same lines as the proof of Theorem 6.14 in [33]. With these preliminaries we now may estimate $|\mathbf{Au}|_2$ by $|\mathbf{Du}|_2$.

Theorem 3.5. *For coefficients* $\mathbf{u} = (u_{\mathbf{l},\mathbf{t}})_{(\mathbf{l},\mathbf{t})\in\boldsymbol{\mathcal{T}}}$ *with*

$$\int_{[0,1]^d} \sum_{(\mathbf{l},\mathbf{t})\in\boldsymbol{\mathcal{T}}} u_{\mathbf{l},\mathbf{t}}\psi_{\mathbf{l},\mathbf{t}}(\mathbf{x})d\mathbf{x} = 0 \tag{13}$$

there holds

$$|\mathbf{Au}|_2^2 \;\lesssim\; \overline{c}(L)|\mathbf{Du}|_2^2 \;. \tag{14}$$

Note, the hidden constants in (14) may depend on the dimension d. However, the only factor which may depend on L, i.e. $\overline{c}(L)$, enters this estimate *independent* of d.

Proof. We consider the two-dimensional case only, as the proof directly generalizes to higher dimensions. We use the representation (5) of $\mathbf{A}$ and set $E = T^{-1}BT$. Then,

$$\begin{aligned} |\mathbf{Au}|_2^2 &= \mathbf{u}^t(E\otimes I + I\otimes E)^t(E\otimes I + I\otimes E)\mathbf{u} \\ &= \mathbf{u}^t(E^tE\otimes I + I\otimes E^tE)\mathbf{u} + \mathbf{u}^t(E^t\otimes E + E\otimes E^t)^t\mathbf{u} \\ &\leq 2\mathbf{u}^t(E^tE\otimes I + I\otimes E^tE)\mathbf{u}\;. \end{aligned}$$

Bounds for $\mathbf{u}^t(E^tE\otimes I)\mathbf{u}$ and $\mathbf{u}^t(I\otimes E^tE)\mathbf{u}$, respectively, follow from Theorem 3.2 and Lemma 3.4. However, because of the constraint (11) required for the univariate estimates, we need a special decomposition of $\mathbf{u}$. Let $\mathbf{n} = T^{-1}\mathbf{1}$ be the vector of Wavelet coefficients of the constant function $x \mapsto 1$. The polynomial exactness of the Wavelets implies that the components of $\mathbf{n}$ are non-zero for the coarsest level only:

$$n_{l,t} = \begin{cases} 2^{-l_0/2} & \text{for } l = l_0 \text{ and } t\in T(l_0) \\ 0 & \text{for } t \in T(l) \end{cases} \;.$$

Therefore, we have $D\mathbf{n} = 2^{3l_0/2}\mathbf{n}$. In addition, $\mathbf{n}$ is the null vector of E, since for consistent discretizations there holds $B\mathbf{1} = 0$. Now, let $\mathbf{w}$ be the vector with components $w_{l,t} = \sum_{s\in T(l_0)} u_{(l_0,l),(s,t)}$. Then, $\mathbf{u}$ may be written as $\mathbf{u} = \mathbf{v} + \mathbf{n}\otimes\mathbf{w}$, where $\mathbf{v}$ is just $\mathbf{u} - \mathbf{n}\otimes\mathbf{w}$ and the tensor product $\mathbf{n}\otimes\mathbf{w} =: \mathbf{p}$ is a vector with components $p_{(l_1,l_2),(t_1,t_2)} = n_{l_1,t_1}w_{l_2,t_2}$. This decomposition is defined in such a way that for all $\mathbf{q} = (q_{l,t})$ there holds $(\mathbf{n}\otimes\mathbf{q})^t\mathbf{v} = 0$. Putting all this together we end up with $\mathbf{u}^t(E^tE\otimes I)\mathbf{u} =$

$$= \mathbf{v}^t(E^tE\otimes I)\cdot\mathbf{v} + \underbrace{\mathbf{v}^t(E^tE\cdot\mathbf{n}\otimes\mathbf{w}) + (E^tE\cdot\mathbf{n}\otimes\mathbf{w})^t\mathbf{v} + (\mathbf{n}^tE^tE\mathbf{n})\mathbf{w}^t\mathbf{w}}_{=0}$$

$$
\begin{aligned}
&\leq \; \overline{c}(L)\mathbf{v}^t(D^2\otimes I)\mathbf{v}\\
&\leq \; \overline{c}(L)\mathbf{v}^t(D^2\otimes I)\mathbf{v}+\underbrace{\mathbf{v}^t(D^2\mathbf{n}\otimes\mathbf{w})+(D^2\mathbf{n}\otimes\mathbf{w})^t\mathbf{v}}_{=0}+\underbrace{(\mathbf{n}^tD^2\mathbf{n})\mathbf{w}^t\mathbf{w}}_{\geq 0}\\
&= \; \overline{c}(L)(\mathbf{v}+\mathbf{n}\otimes\mathbf{w})^t(D^2\otimes I)(\mathbf{v}+\mathbf{n}\otimes\mathbf{w})\\
&= \; \overline{c}(L)\mathbf{u}^t(D^2\otimes I)\mathbf{u}\;.
\end{aligned}
$$

Likewise one shows $\mathbf{u}^t(I\otimes E^tE)\mathbf{u}\leq\overline{c}(L)\mathbf{u}^t(I\otimes D^2)\mathbf{u}$. Then, $|\mathbf{Au}|_2^2\leq 2\overline{c}(L)\mathbf{u}^t(D^2\otimes I+I\otimes D^2+2D\otimes D)\mathbf{u}=2\overline{c}(L)\mathbf{u}^t\mathbf{D}^2\mathbf{u}$. □

Unfortunately, for reasons which will be explained later, we can not use the idea for the proof of (14) also for lower bounds of $|\mathbf{Au}|_2$. Therefore, some other technique will be used here which leads to bounds which may strongly depend on the dimension.

Theorem 3.6. *With the prerequisites of Theorem 3.5, there holds*

$$
\frac{\min(1,\underline{b}(L)\underline{a}(L)^{d-1})}{\overline{a}(L)^d}|\mathbf{Du}|_2^2 \;\lesssim\; |\mathbf{Au}|_2^2\;.
$$

Proof. Again, we consider the two-dimensional case only. The translation of the proof to higher dimensions is straight forward. Using the decomposition (7) of A in $\mathbf{A}$, compare (5), we see that $\mathbf{A}$ may be written in the form

$$
\mathbf{A}=(T^{-1}\otimes T^{-1})(B\otimes I+I\otimes B)(T\otimes T)\;.
$$

Thus, $|\mathbf{Au}|_2^2=$

$$
\begin{aligned}
&= \; \mathbf{u}^t(T\otimes T)^t(B\otimes I+I\otimes B)^t(TT^t)^{-1}\otimes(TT^t)^{-1}(B\otimes I+I\otimes B)(T\otimes T)\cdot\mathbf{u}\\
&\geq \; \overline{a}(L)^{-d}\mathbf{u}^t(T\otimes T)^t(B\otimes I+I\otimes B)^t(B\otimes I+I\otimes B)(T\otimes T)\cdot\mathbf{u}\\
&= \; \overline{a}(L)^{-d}\mathbf{u}^t(T^tBBT\otimes T^tT+2T^tBT\otimes T^tBT+T^tT\otimes T^tBBT)\mathbf{u}
\end{aligned}
$$

Now, we can proceed as in the proof of Theorem 3.5. With Lemma 3.4 and Theorem 3.2 we obtain

$$
\begin{aligned}
\mathbf{u}^tT^tBT\otimes T^tBT\mathbf{u} \;&\gtrsim\; \mathbf{u}^tD\otimes D\mathbf{u} \qquad \text{and}\\
\mathbf{u}^tT^tBBT\otimes T^tT\mathbf{u} \;&\geq\; \underline{b}(L)\underline{a}(L)^{d-1}\mathbf{u}^tD^2\otimes I\mathbf{u}\;.
\end{aligned}
$$

This yields

$$
\begin{aligned}
|\mathbf{Au}|_2^2 \;&\gtrsim\; \frac{\min(1,\underline{b}(L)\underline{a}(L)^{d-1})}{\overline{a}(L)^d}\mathbf{u}^t(D^2\otimes I+2D\otimes D+I\otimes D^2)\mathbf{u}\\
&= \; \frac{\min(1,\underline{b}(L)\underline{a}(L)^{d-1})}{\overline{a}(L)^d}\mathbf{u}^t\mathbf{D}^2\mathbf{u}\;.
\end{aligned}
$$

$\square$

In the case of Lifting-Interpolets, the factors $\underline{a}(L)$ and $\overline{a}(L)$ are ~ 1. Therefore, Theorem 3.6 yields bounds which are essentially independent of the dimension. But, for Interpolets the estimates $\underline{a}(L) \sim 2^{-L}$ and $\overline{a}(L) \sim L$ result in strongly dimension- and level-dependent estimates for the l_2-condition number of $\mathbf{AD}^{-1}$. This is much to pessimistic. For example, in the numerical experiments of Section 5 we observe that the Jacobi-preconditioner works for three-dimensional problems as good as for two-dimensional problems. The critical point for dimension-independent lower bounds of $|\mathbf{Au}|_2^2$ are estimates for the mixed terms like $\mathbf{u}^t(E^t \otimes E)\mathbf{u}$. What we would need here is an estimate from below, e.g.,

$$\mathbf{u}^t(E^t \otimes E + E \otimes E^t)\mathbf{u} \ \geq \ -c\mathbf{u}^t(E^tE \otimes I + I \otimes E^tE)\mathbf{u} \ , \tag{15}$$

for some $c < 1$. Then, we could continue with

$$\begin{aligned} |(E \otimes I + I \otimes E)\mathbf{u}|_2^2 \ &\geq \ (1-c)\mathbf{u}^t(E^tE \otimes I \ + \ I \otimes E^tE)\mathbf{u} \\ &\geq \ (1-c)\underline{c}(L)\mathbf{u}^t(D^2 \otimes I \ + \ I \otimes D^2)\mathbf{u} \\ &\gtrsim \ \underline{c}(L)\mathbf{u}^t(D \otimes I \ + \ I \otimes D)^2\mathbf{u} \ . \end{aligned}$$

That means, the level-dependent term $\underline{c}(L)$ enters the lower bound *independent* of the dimension – similar to the upper bound. Therefore, we conjecture that there holds

$$\underline{c}(L)|\mathbf{Du}|_2^2 \ \lesssim \ |\mathbf{Au}|_2^2 \ . \tag{16}$$

It is convenient to collect the estimates of Theorem 3.3, 3.5 and 3.6 to get some clear statements on the l_2-condition number of $\mathbf{AD}^{-1}$. For Interpolets we do not use the pessimistic lower bounds of Theorem 3.6 for $c(L)$ from (8) but those of conjecture (16). The results are shown in Table 3. Obviously, high order (Lifting-)Interpolets lead to better condition numbers than low order basis functions and the preconditioner is almost optimal for Lifting-Interpolets.

We finish the section with the proof of Theorem 3.2.

Proof. (Theorem 3.2).
(ii): The stability constraint (9) and the norm estimates of Theorem 3.1

κ	Lifting-Interpolets	Interpolets*
$N=2$ ($\gamma=1.5$)	$\sim \left(L2^{(1+\epsilon)L}\right)^{1/2}$	$\sim L2^{(1+\epsilon)L}$
$N\geq 4$ ($\gamma>2$)	~ 1	$\sim \left(L2^{L}\right)^{1/2}$

Table 3: *Estimates for the l_2-condition numbers of $\mathbf{AD}^{-1}$ for the Jacobi-preconditioner. * We used conjecture (16) here.*

together with the Poincare-Friedrichs inequality [1] yield

$$\mathbf{u}^t T^t B T \mathbf{u} \sim \mathbf{u}^t T^t \langle \partial_x \mathbf{\Phi}^{Lt}, \partial_x \mathbf{\Phi}^L \rangle T \mathbf{u} = |\sum_{l=l_0}^{L} \mathbf{\Psi}_l \mathbf{u}_l|^2_{\mathcal{H}^1} \sim \|\sum_{l=l_0}^{L} \mathbf{\Psi}_l \mathbf{u}_l\|^2_{\mathcal{H}^1} \ \sim\ \mathbf{u}^t D \mathbf{u}\,,$$

i.e. assertion.
(iii): If $N \geq 4$, i.e. $\gamma > 2$, we can proceed as above using the norm equivalences of Theorem 3.2 for $\mathcal{H}^2$. However, for the first step we need the stability condition

$$\mathbf{v}^t B B \mathbf{v} \sim \mathbf{v}^t \langle \partial_{xx} \mathbf{\Phi}^{Lt}, \partial_{xx} \mathbf{\Phi}^L \rangle \mathbf{v} \quad \text{for all } \ \mathbf{v}\ . \tag{17}$$

For our simple setting, where B is a symmetric Toeplitz matrix, this condition is a consequence of the stability (9). We know that the eigenvalues λ_k of B are equivalent to k^2, i.e. $\lambda_k \sim k^2$. Therefore, the eigenvalues of BB scale like k^4. But, so do the eigenvalues of $\langle \partial_{xx} \mathbf{\Phi}^{Lt}, \partial_{xx} \mathbf{\Phi}^L \rangle$. Thus (17) holds.
The case of piecewise linear Interpolets, i.e. $N = 2$ and $\gamma < 2$, is a little bit more complicated. We start with bounds for $\overline{b}(L)$. Let $\mathbf{u}_l$ be the partial vector of $\mathbf{u}$ for level l, $\mathbf{u}_l := (u_{l,t})_{t\in T(l)}$, and R_l be the restriction operator defined by $R_l \mathbf{u} := \mathbf{u}_l$. Then, conversely $\mathbf{u} = \sum_{l=l_0}^{L} R_l^t \mathbf{u}_l$. Furthermore, let $\mathbf{v}^l := T R_l^t \mathbf{u}_l$ and $\mathbf{\Phi}_4^L$ be the vector of scaling functions for 4th order Interpolets. These scaling functions are in $\mathcal{H}^2$, see Table 1. Therefore, we have the stability

$$\mathbf{w}^t B B \mathbf{w} \ \sim\ \mathbf{w}^t \langle \partial_{xx} \mathbf{\Phi}_4^{Lt}, \partial_{xx} \mathbf{\Phi}_4^L \rangle \mathbf{w} \quad \text{for all } \mathbf{w}\ .$$

Together with Bernstein's inequality this implies

$$\mathbf{v}^{lt} B B \mathbf{v}^l \ \lesssim\ \|\mathbf{\Phi}_4^L \mathbf{v}^l\|^2_{\mathcal{H}^2} \ \lesssim\ 4^{(2-\gamma+\epsilon)L} \|\mathbf{\Phi}_4^L \mathbf{v}^l\|^2_{\mathcal{H}^{\gamma-\epsilon}}\ .$$

The next step is a bound for $\|\mathbf{\Phi}_4^L \mathbf{v}^l\|_{\mathcal{H}^{\gamma-\epsilon}}$ by $\|\mathbf{\Phi}^L \mathbf{v}^l\|_{\mathcal{H}^{\gamma-\epsilon}}$. $\mathbf{\Phi}_4^L \mathbf{v}^l$ is just the interpolant of the (less smooth) function $\mathbf{\Phi}^L \mathbf{v}^l$ in the nodes $\{0, 2^{-L}, ..., 1-$

$2^{-L}\}$, since $\mathbf{\Phi}_4^L$ as well as $\mathbf{\Phi}^L$ are interpolating scaling functions. A further consequence of Bernstein- and Jackson-inequalities for these functions are uniform bounds for the $\mathcal{H}^{\gamma-\epsilon} \to \mathcal{H}^{\gamma-\epsilon}$-norm of the interpolation projector onto span$\mathbf{\Phi}_4^L$ [25]. This leads to $\|\mathbf{\Phi}_4^L\mathbf{v}^l\|_{\mathcal{H}^{\gamma-\epsilon}} \lesssim \|\mathbf{\Phi}^L\mathbf{v}^l\|_{\mathcal{H}^{\gamma-\epsilon}}$. Therefore,

$$\begin{aligned}\mathbf{v}^{lt}BB\mathbf{v}^l &\lesssim 4^{(2-\gamma+\epsilon)L}\|\mathbf{\Phi}^L\mathbf{v}^l\|^2_{\mathcal{H}^{\gamma-\epsilon}} = 4^{(2-\gamma+\epsilon)L}\|\mathbf{\Psi}_l\mathbf{u}_l\|^2_{\mathcal{H}^{\gamma-\epsilon}}\\ &\lesssim 4^{(2-\gamma+\epsilon)L}4^{(\gamma-\epsilon)l}|\mathbf{u}_l|_2^2 \ .\end{aligned}$$

Thus, we have

$$\begin{aligned}\mathbf{u}^tBB\mathbf{u} = \sum_{l,k=l_0}^{L}\mathbf{v}^{kt}BB\mathbf{v}^l &\leq \sum_{l,k}|B\mathbf{v}^k|_2|B\mathbf{v}^l|_2\\ &\lesssim \sum_{l,k}2^{(2-\gamma+\epsilon)2L}2^{(\gamma-\epsilon)(l+k)}|\mathbf{u}_k|_2|\mathbf{u}_l|_2\\ &\leq \sum_{l,k}2^{(2-\gamma+\epsilon)(2L-(l+k))}|4^k\mathbf{u}_k|_2|4^l\mathbf{u}_l|_2\\ &\lesssim 2^{(2-\gamma+\epsilon)2L}\sum_{l}|4^l\mathbf{u}_l|_2^2 \ = \ 2^{(2-\gamma+\epsilon)2L}\mathbf{u}^tD^2\mathbf{u} \ .\end{aligned}$$

This proves the bound for $\overline{b}(L)$.
In order to prove the lower bound for $\underline{b}(L)$ we rewrite the left inequality in (12) in an equivalent form which allows to use one-sided norm estimates for the dual Wavelets. Obviously, the assertion is equivalent to

$$\mathbf{v}^tT^{-t}D^2T^{-1}\mathbf{v} \ \lesssim \ c\mathbf{v}^tBB\mathbf{v} \quad \text{for all } \ \mathbf{v} \ \text{ with } \ \int_0^1(\mathbf{\Phi}^L\mathbf{v})(\mathbf{x})dx = 0 \ , \qquad (18)$$

where c is L if $N=2$ and $c=1$ otherwise. The next transformation uses the pseudo inverse B^+ of B which may be defined via the Schur-decomposition

$$B^+ := U^t\begin{pmatrix}0 & & & \\ & \lambda_1^{-1} & & \\ & & \ddots & \\ & & & \lambda_{2^L-1}^{-1}\end{pmatrix}U \quad \text{where} \quad B = U^t\begin{pmatrix}0 & & & \\ & \lambda_1 & & \\ & & \ddots & \\ & & & \lambda_{2^L-1}\end{pmatrix}U \ .$$

Because of the constraint on $\mathbf{v}$, (18) is equivalent to

$$\mathbf{w}^t(B^+T^{-t}D)(DT^{-1}B^+)\mathbf{w} \ \lesssim \ c\mathbf{w}^t\mathbf{w} \quad \text{for all } \ \mathbf{w} \ .$$

That means, we have to find bounds for the largest eigenvalue of Q^tQ for $Q = DT^{-1}B^+$. The largest eigenvalue of Q^tQ is also the largest eigenvalue of QQ^t and vice versa [17]. Therefore, it is sufficient to show

$$\mathbf{q}^t T^{-1}(BB)^+ T^{-t}\mathbf{q} \;\lesssim\; c\mathbf{q}^t D^{-2}\mathbf{q} \quad \text{for all } \;\mathbf{q}\;. \tag{19}$$

From (6) we know that T^{-t} is exactly the transformation matrix $\tilde{T}$ for the dual Wavelets. Now, if the left hand side of (19) would be $\lesssim \|\tilde{\mathbf{\Psi}}^L\mathbf{q}\|^2_{\mathcal{H}^{-2}}$, then (19) would follow from the one-sided norm estimates in Theorem 3.1. For an arbitrary vector $\mathbf{p}$ we have

$$\begin{aligned}
\|\tilde{\mathbf{\Phi}}^L\mathbf{p}\|^2_{\mathcal{H}^{-2}} &= \sup_{r\in\mathcal{H}^2} \frac{\langle\tilde{\mathbf{\Phi}}^L\mathbf{p}, r\rangle^2}{\|r\|^2_{\mathcal{H}^2}} \geq \sup_{\mathbf{r}:\ \mathbf{1}^t\mathbf{r}=0} \frac{\langle\tilde{\mathbf{\Phi}}^L\mathbf{p}, \mathbf{\Phi}^L_4\mathbf{r}\rangle^2}{\|\mathbf{\Phi}^L_4\mathbf{r}\|^2_{\mathcal{H}^2}} \\
&\gtrsim \sup_{\mathbf{r}:\ \mathbf{1}^t\mathbf{r}=0} \frac{(\mathbf{p}^tK\mathbf{r})^2}{\mathbf{r}^tBB\mathbf{r}} \quad, \qquad \text{where } K := \langle\tilde{\mathbf{\Phi}}^{Lt}, \mathbf{\Phi}^L_4\rangle \\
&= \sup_{\mathbf{s}} \frac{(\mathbf{p}^tKB^+\mathbf{s})^2}{\mathbf{s}^t\mathbf{s}} = \mathbf{p}^tKB^+B^+K^t\mathbf{p}\;.
\end{aligned}$$

If the underlying Wavelets $\mathbf{\Psi}^L$ are the hat functions, then the dual scaling functions $\tilde{\mathbf{\Phi}}^L$ are translates of the Dirac functional which are also biorthogonal to $\mathbf{\Phi}^L_4$. Then, K is the identity matrix and, thus, $\|\tilde{\mathbf{\Phi}}^L\mathbf{p}\|^2_{\mathcal{H}^{-2}} \gtrsim \mathbf{p}^t(BB)^+\mathbf{p}$. Otherwise, K is a symmetric Toeplitz matrix with generator $\{a_s := \int_0^1 \tilde{\phi}(x)\cdot \phi_4(x-s)dx\}_{s\in\mathbb{Z}}$. Here, $\tilde{\phi}$ and ϕ_4 are the mother scaling functions for the dual Wavelets and the 4th order Interpolets, respectively. A possible way to compute these generators is the eigenvector technique of [6]. The refinement masks of $\tilde{\phi}$ and ϕ_4 needed there are explicitly known. They have been used e.g. for estimates of the Sobolev-regularity of $\tilde{\phi}$ and ϕ_4. By means of the generators of K we computed the matrix symbol and found that all eigenvalues of K are between 1.0 and 1.15 independent on the refinement level L. This implies $\mathbf{p}^tKB^+B^+K^t\mathbf{p} \gtrsim \mathbf{p}^t(BB)^+\mathbf{p}$. Therefore, we end up with

$$\mathbf{q}^tT^{-1}(BB)^+T^{-t}\mathbf{q} \lesssim \|\tilde{\mathbf{\Phi}}^LT^{-t}\mathbf{q}\|^2_{\mathcal{H}^{-2}} \lesssim c\mathbf{q}^tD^{-2}\mathbf{q} \quad \text{for all } \;\mathbf{q}$$

which proves the assertion.
(i): Assertion *(i)* is a consequence of the stability of the single scale mass matrix:

$$\mathbf{v}^t\langle\mathbf{\Phi}^{Lt}, \mathbf{\Phi}^L\rangle\mathbf{v} \;\sim \mathbf{v}^t\mathbf{v} \quad \text{for all } \;\mathbf{q}\;. \tag{20}$$

For the hat functions the stability is well-known. For other scaling functions one may use the above mentioned procedure via the matrix symbols to verify the stability. Now, in case of Lifting-Interpolets, *(ii)* follows from (20) and the norm equivalences for $\mathcal{H}^0$. For the simple Interpolets, we get from Theorem 3.1 the upper bound

$$\mathbf{u}^t T^t T \mathbf{u} \;\sim\; \mathbf{u}^t T^t \langle \mathbf{\Phi}^{Lt}, \mathbf{\Phi}^L \rangle T \mathbf{u} = \|\mathbf{\Psi}^L \mathbf{u}\|^2_{\mathcal{L}^2} \;\lesssim\; L \mathbf{u}^t \mathbf{u} \ .$$

For the lower bound $\underline{a}(L)$ the proof follows the lines of the proof for *(iii)*. Assertion *(iv)* is a consequence of *(i)* and *(iii)*:

$$\begin{aligned} \mathbf{u}^t (T^{-1}BT)^t (T^{-1}BT)\mathbf{u} = \mathbf{u}^t (BT)^t (TT^t)^{-1} (BT)\mathbf{u} \;&\lesssim\; \frac{1}{\underline{a}(L)} \mathbf{u}^t (BT)^t (BT)\mathbf{u} \\ &\lesssim\; \frac{\overline{b}(L)}{\underline{a}(L)} \mathbf{u}^t D^2 \mathbf{u} \ . \end{aligned}$$

□

4 The Lifting-preconditioner

In the previous section it was shown that the simple Jacobi-preconditioner works better for discretizations based on Lifting-Interpolets than for discretizations based on Interpolets. On the other hand, Interpolets allow for a quite simple evaluation of non-linear terms. So one might wonder, whether the advantages of both bases may be combined somehow. In fact this is possible. Because of their close relation (10) one can find for given $u = \sum_{(\mathbf{l},\mathbf{t}) \in \boldsymbol{\mathcal{T}}} u_{\mathbf{l},\mathbf{t}} \psi_{\mathbf{l},\mathbf{t}}$ the alternative representation $u = \sum_{(\mathbf{l},\mathbf{t}) \in \boldsymbol{\mathcal{T}}} \overline{u}_{\mathbf{l},\mathbf{t}} \overline{\psi}_{\mathbf{l},\mathbf{t}}$ in terms of the Lifting-Interpolets. Note that throughout this section $\psi_{\mathbf{l},\mathbf{t}}$ denote Interpolets and $\overline{\psi}_{\mathbf{l},\mathbf{t}}$ Lifting-Interpolets. Therefore, whenever needed, we can switch to that representation which fits best to the current task. An application of this idea is the Lifting-preconditioner. Here, the Interpolet-based discretization is kept and just for the preconditioning we switch the basis. For a more formal description, let $\mathbf{C}$ be the matrix for the 'Interpolet $\to$ Lifting-Interpolet' transform, i.e.

$$\sum_{(\mathbf{l},\mathbf{t}) \in \boldsymbol{\mathcal{T}}} u_{\mathbf{l},\mathbf{t}} \psi_{\mathbf{l},\mathbf{t}} = \sum_{(\mathbf{l},\mathbf{t}) \in \boldsymbol{\mathcal{T}}} \overline{u}_{\mathbf{l},\mathbf{t}} \overline{\psi}_{\mathbf{l},\mathbf{t}} \quad \Leftrightarrow \quad \overline{\mathbf{u}} = \mathbf{C}\mathbf{u} \ . \tag{21}$$

Then, the Lifting-preconditioner for (3) consists of a multiplication by $\mathbf{C}$ from the left and a multiplication by $\mathbf{C}^{-1}\mathbf{D}^{-1}$ from the right, i.e. (3) turns into

$$\mathbf{C}\mathbf{A}\mathbf{C}^{-1}\mathbf{D}^{-1}\overline{\mathbf{u}} = \mathbf{C}\mathbf{f} \ .$$

For a motivation of this approach, let us consider the case of simple index sets (4) first. In this case, the transformation matrix $\mathbf{C}$ takes the form

$$\mathbf{C} = (\overline{T}^{-1}T) \otimes ... \otimes (\overline{T}^{-1}T) \ ,$$

where T is the matrix for the Interpolet transform and $\overline{T}$ is the respective matrix for the Lifting-Interpolet transform, defined by $\mathbf{\Phi}^L\overline{T} = \overline{\mathbf{\Psi}}^L$. Hence, with (5) and (7) we obtain

$$\mathbf{CAC}^{-1} = \sum_{i=1}^{d} \underbrace{I \otimes .. \otimes I}_{i-1 \text{ times}} \otimes (\overline{T}^{-1}B\overline{T}) \otimes \underbrace{I \otimes .. \otimes I}_{d-i \text{ times}} .$$

This is the stiffness matrix for the Lifting-Interpolet based discretization. Therefore, the Lifting-preconditioner yields $O(1)$ condition numbers for high order Interpolets and $O(\sqrt{L2^{(1+\epsilon)L}})$ for piecewise linear Interpolets. The numerical experiments of the next section show that the Lifting-preconditioner works as well also for more general index sets.

4.1 Algorithms for the Basis Transforms

In this subsection we show how the multivariate transform boils down to univariate transforms. Then, these are considered in detail. Besides algorithms we also discuss a necessary condition on the index sets $\boldsymbol{\mathcal{T}}$ required for the transforms. We begin with 'Interpolets $\rightarrow$ Lifting-Interpolets'. As a first step, we pick a particular univariate partial function

$$w(x) := \sum_{(l_1,t_1):\ ((l_1,\bar{\mathbf{l}}),(t_1,\bar{\mathbf{t}}))\in\boldsymbol{\mathcal{T}}} u_{(l_1,\bar{\mathbf{l}}),(t_1,\bar{\mathbf{t}})}\psi_{l_1,t_1}(x)$$

for some $(d-1)$-dimensional multiindices $\bar{\mathbf{l}} = (l_2, ..., l_d)$ and $\bar{\mathbf{t}} = (t_2, ..., t_d)$, respectively. Now, assume we can find for all such partial functions the alternative representation

$$w(x) = \sum_{(l_1,t_1):\ ((l_1,\bar{\mathbf{l}}),(t_1,\bar{\mathbf{t}}))\in\boldsymbol{\mathcal{T}}} \overline{u}_{(l_1,\bar{\mathbf{l}}),(t_1,\bar{\mathbf{t}})}\overline{\psi}_{l_1,t_1}(x)$$

in terms of the Lifting-Wavelets. Consequently, we can find a representation of u in terms of the *mixed* basis functions $\overline{\psi}_{l_1,t_1}(x_1) \cdot \psi_{l_2,t_2}(x_2) \cdot ... \cdot$

$\psi_{l_d,t_d}(x_d)$, $(\mathbf{l},\mathbf{t}) \in \boldsymbol{\mathcal{T}}$. In this way univariate transforms, successively applied to all d coordinate directions, form the desired change of bases. This also explains why we need to consider univariate transforms

$$w = \sum_{(l,t)\in\mathcal{T}} w_{l,t}\psi_{l,t} \quad \rightarrow \quad w = \sum_{(l,t)\in\mathcal{T}} \overline{w}_{l,t}\overline{\psi}_{l,t} \ , \tag{22}$$

in the following, only. Note, the univariate index sets are denoted by $\mathcal{T}$ instead of $\boldsymbol{\mathcal{T}}$. The transform (22) essentially consists of two steps. First, nodal values of w on a certain (non-uniform) grid are computed by means of the inverse Interpolet-transform from $\{u_{l,t}\}$. Then, we obtain the coefficients $\overline{w}_{l,t}$ by a modified Interpolet-transform. An easy to read description of this second step is given e.g. in [29] for the non-adaptive case. However, a few things need to be taken into account for the adaptive case. For example, $\mathcal{T}$ has to fulfill a so-called cone-condition [10]:

$$(l,t) \in \mathcal{T} \ \text{ and } \ \text{supp}\overline{\psi}_{l,t} \cap \text{supp}\overline{\psi}_{l-1,s} \neq \emptyset \ \text{ imply that } \ (l-1,s) \in \mathcal{T} \ . \tag{23}$$

If $\mathcal{T}$ does not meet this condition, it may happen that for given $\{w_{l,t}\}$ there are no coefficients $\{\overline{w}_{l,t}\}$ such that (22) holds. In this case the basis transform is not possible. Because of the recursive structure of the cone-condition it is always possible to enrich an arbitrary finite index set to fulfill (23) in a top-bottom manner. The question is, whether the artificial fill-in significantly increases the size of the index set. In fact, this may happen in some pathological limit cases, but in general the number of indices increases by a certain fraction at most [4]. The cone-condition must be generalized to higher dimensions in the following way:

> For $1 \leq i \leq d$ and all $(l_1, .., l_{i-1}, l_{i+1}, .., l_d)$ and $(t_1, .., t_{i-1}, t_{i+1}, .., t_d)$
> the index set $\mathcal{T} := \{(l_i, t_i) : \ (\mathbf{l},\mathbf{t}) \in \boldsymbol{\mathcal{T}}\}$ must meet (23) .

Similar to the univariate case the multivariate cone-condition may be enforced for arbitrary finite index sets $\boldsymbol{\mathcal{T}}^0$ by an appropriate fill up. Such a scheme is shown below:

```
for i = 1 to d do
    set 𝓣^i := ∅
    for all (l_1,..,l_{i-1},l_{i+1},..,l_d) and (t_1,..,t_{i-1},t_{i+1},..,t_d)
    such that 𝒯 := {(l_i,t_i) : (l,t) ∈ 𝓣^{i-1}} ≠ ∅ do
        if necessary, enrich 𝒯 such that(23) is fulfilled
        𝓣^i = 𝓣^i ∪ {(l,t) : (l_i,t_i) ∈ 𝒯}
    end
end
𝓣 := 𝓣^d
```

It seems that the problem, whether the artificial fill-in destroys the advantages of adaptivity or not, is even worse here than in the univariate case. E.g. fill-in for $i = 1$ may itself induce further fill-in for $i > 1$. However, in all our practical applications we found that the final index set $\boldsymbol{\mathcal{T}}$ is larger only by 10% - 30% than $\boldsymbol{\mathcal{T}}^0$ which is acceptable.

For a description of the algorithm for (22), it is convenient to introduce some further notation. We separate the indices $(l, t) \in \mathcal{T}$ according to their level and set $\mathcal{T}(l) := \{t : (l, t) \in \mathcal{T}\}$. L is defined as the maximal level of $\mathcal{T}$, i.e. $\mathcal{T}(L) \neq \emptyset$ and $\mathcal{T}(L + 1) = \emptyset$. The irregular grid Ω^L on which w is evaluated by the inverse Interpolet-transform is recursively defined by

$$\Omega^l := \begin{cases} 2^{-l_0}\mathcal{T}(l_0) & \text{for } l = l_0 \\ \Omega^{l-1} \cup \{2^{-l}(2t + 1) : t \in \mathcal{T}(l)\} & \text{else} \end{cases} .$$

Now, consider the first step of the Lifting-Interpolet transform. Equation (22) just means that

$$\sum_l \sum_{t\in\mathcal{T}(l)} w_{l,t}\psi_{l,t} \stackrel{!}{=} \sum_{t\in\mathcal{T}(l_0)} \overline{w}_{l_0,t}\psi_{l_0,t} + \sum_{l>l_0} \sum_{t\in\mathcal{T}(l)} \overline{w}_{l,t}\Big(\psi_{l,t} - \frac{\sqrt{2}}{4}(\phi^{l-1,t} + \phi^{l-1,t+1})\Big) .$$

If we test this equation by dual Wavelets $\tilde{\psi}_{L,t}$ of the maximal level, we obtain

$$w_{L,t} = \overline{w}_{L,t} \quad \text{for} \quad t \in \mathcal{T}(L) .$$

That means, we get the coefficients $\{\overline{w}_{L,t}\}$ via the Interpolet high pass filter from the nodal values $w(\Omega^L)$. To proceed with this scheme on level $L-1$ we need the nodal values of

$$w' := w - \sum_{t\in\mathcal{T}(L)} \overline{w}_{L,t}\overline{\psi}_{L,t}$$

on the grid Ω^{L-1}. Here, some properties of the Interpolets help: the scaling functions $\phi^{l,s}$ are interpolating on the grid $2^{-l}S(l)$, i.e. $\phi^{l,s}(2^{-l}t) = \delta_{s,t}$. Furthermore, the Interpolets are just the odd numbered scaling functions, $\psi_{l,t} = \phi^{l,2t+1}$ for $l > l_0$. This implies that the Wavelets $\psi_{L,t}$ vanish on Ω^{L-1}. Consequently, for $x := 2^{-(L-1)}s \in \Omega^{L-1}$ there holds

$$w'(x) = \begin{cases} w(x) & \text{if } t \notin \mathcal{T}(L) \text{ and } t-1 \notin \mathcal{T}(L) \\ w(x) - \frac{\sqrt{2}}{4}\overline{w}_{L,t} & \text{if } t \in \mathcal{T}(L) \text{ and } t-1 \notin \mathcal{T}(L) \\ w(x) - \frac{\sqrt{2}}{4}\overline{w}_{L,t-1} & \text{if } t \notin \mathcal{T}(L) \text{ and } t-1 \in \mathcal{T}(L) \\ w(x) - \frac{\sqrt{2}}{4}(\overline{w}_{L,t} + \overline{w}_{L,t-1}) & \text{if } t \in \mathcal{T}(L) \text{ and } t-1 \in \mathcal{T}(L) \end{cases} . \tag{24}$$

This scheme of high pass filters and subsequent corrections of nodal values on the coarser grid now repeats for the lower levels. In this way we obtain all desired coefficients $\{\overline{w}_{l,t}\}$.

The algorithm for the inverse transform 'Lifting-Interpolet $\rightarrow$ Interpolet' just inverts the above scheme. In the first step the nodal values of w are computed from the coefficients $\{\overline{w}_{l,t}\}$. This is done by a modified inverse Interpolet transform. The modification consists of undoing the Lifting-correction (24) on each level before the synthesis filters are used to compute the nodal values on the next finer grid. Then, we obtain the coefficients $\{w_{l,t}\}$ by means of the Interpolet transform.

5 Numerical Experiments

The analysis of Section 3 was restricted to simple index sets and periodic boundary conditions only. In this section we present some numerical experiments which go beyond these limitations. Of main interest are the following questions: 1. Are there differences in the performance of the preconditioners for non-adaptive and adaptive grids? 2. Does the dimension of the domain influence the preconditioning effect? 3. The estimates of Section 3 show that smoother Wavelets lead to better condition numbers than less smooth Wavelets like the hat functions. Is this the case?

In our experiments we consider the model problem

$$-\Delta u = d9\pi^2 \prod_{i=1}^{d} \sin(3\pi x_i) \quad \text{in} \quad]0,1[^d \quad \text{and} \quad u = 0 \;\text{ on }\; \partial[0,1]^d$$

for $d = 2$ and $d = 3$. For the discretization we use 2nd and 4th order Interpolets, $N = 2$, $\gamma = 1.5$ and $N = 4$, $\gamma = 2.171$, respectively, and the standard second order FD scheme. The index sets $\boldsymbol{\mathcal{T}}$ were either full grid index sets (4) or the following sparse grid index sets

$$\boldsymbol{\mathcal{T}} = \{(\mathbf{l}, \mathbf{t}) :\ \mathbf{t} \in T(l_1) \times ... \times T(l_d) \ \text{ for } \ \mathbf{l} \ \text{ such that } \ l_1 + ... + l_d \leq L\}$$

for maximal levels L from l_0 up to 10. Because of memory limitations, we could run the experiments with the three-dimensional full grids up to level $L = 7$, only. Note, although the sparse grid index sets have a much smaller dimension than the full grid index sets with the same maximal level, $O(L^{d-1}2^L)$ in comparison to $O(2^{dL})$, the resulting discretization error is almost the same [19].

Because of the non-symmetric coefficient matrices which are not explicitly available we use a BiCGStab iterative solver with the simple Jacobi-preconditioner and the more sophisticated Lifting-preconditioner for the solution of the linear systems. For the calculation of the convergence histories, first the linear systems were solved up to a very high accuracy to obtain the coefficients $\mathbf{u}$ of the true numerical solution. Then, the linear systems were solved for a second time with initial guess $\mathbf{0}$. Within this iteration some accurate approximations of the $\mathcal{L}^2$-error

$$e^{(it)} := \Big(\int_{[0,1]^d} |\Psi_{\boldsymbol{\mathcal{T}}}(\mathbf{u} - \mathbf{u}^{(it)})|^2 d\mathbf{x} \Big)^{1/2}$$

of the current iterates $\mathbf{u}^{(it)}$ were computed.

In Figures 2 and 3 the convergence histories of $e^{(it)}$ are shown for the various test cases. Obviously there is no major difference in the convergence rates between full and sparse grids or between the two- and the three-dimensional case. This is particularly interesting for the Jacobi-preconditioner, as the estimates of Theorem 3.5 indicate that the convergence rate could deteriorate with increasing dimension. But as we see here, this is not the case. This supports the conjecture made in Section 3. Another interesting point is that the convergence rates for the Jacobi-preconditioner and high order Interpolets are really much better than for the 2nd order Interpolets. This behaviour is also in agreement with the conjecture in Section 3: $\kappa \lesssim L2^L$ for $N = 2$ and $\kappa \lesssim \sqrt{L2^L}$ for high order Interpolets.

6 Conclusions.

In this paper preconditioners for general FD and biorthogonal Petrov-Galerkin discretizations on adaptive sparse grids were considered. For the simple case of full grid index sets and periodic boundary conditions estimates for the resulting l_2-condition numbers have been given. For the Jacobi-preconditioner in connection with Lifting-Interpolets or the Lifting-preconditioner we found the bounds $\kappa \lesssim \sqrt{L2^L}$ if $N = 2$ and $\kappa \sim 1$ if $N \geq 4$. These bounds hold independent of the dimension d. This result is supported by our numerical experiments.

The estimates for just Interpolets and the Jacobi-preconditioner seem not to be optimal. Here, we conjecture that the condition numbers are also independent of the dimension and that they scale moderately with the maximal level. This behaviour has been observed in numerical experiments.

Acknowledgements: The author has been supported by the Deutsche Forschungsgemeinschaft, GR 1144/7-2.

References

[1] Braess D., *Finite Elements: Theory, Fast Solvers and Applications in Solid Mechanics*; Cambridge University Press, Cambridge, 1997

[2] Bungartz H., *Finite elements of higher order on sparse grids*; Habil. thesis, TU München (1998)

[3] Carnicer J.M., Dahmen W., Pena J.M., *Local Decomposition of Refinable Spaces*; Appl. Comp. Harm. Anal. 3, (1996), pp. 127–153

[4] Cohen A., Dahmen W., Daubechies I., DeVore R., *Tree Approximation and Optimal Encoding*; IGPM Report, RWTH Aachen, September 1999.

[5] Dahmen W., *Wavelet and Multiscale Methods for Operator Equations*; Acta Numerica, Vol 6., (1997), pp. 55–228

[6] Dahmen W., Michelli C.A., *Using the refinement equation for evaluating integrals of wavelets*; SIAM J. Num. Analysis 30, (1993), pp. 507-537

[7] Daubechies I., *Ten lectures on Wavelets*; CBMS-NSF Regional Conference Series in Appl. Math. 61 (1992)

[8] Deslaurier G., Dubuc S., *Symmetric iterative interpolation processes*; Constr. Approx. 5 (1989), pp. 49-68

[9] Donoho D., *Interpolating Wavelet transforms*; report, Stanford University, (1992)

[10] Griebel M., *Adaptive sparse grid multilevel methods for elliptic PDEs based on finite differences*; Computing 61 No. 2 (1998), pp. 151-180

[11] Griebel M., Knapek S., *Optimized approximation spaces for operator equations*; Preprint Universität Bonn (1998), accepted for publication in Constr. Approx.

[12] Koster F. Griebel M., Orthogonal Wavelets on the interval; University Bonn, (1998) preprint No. 576

[13] Griebel M., Koster F., *Adaptive Wavelet Solvers for the Unsteady Incompressible Navier-Stokes Equations*; to appear in Malek J., Rokyta M. (eds.), Advanced Mathematical Theories in Fluid Mechanics, Springer (2000)

[14] Griebel M., Oswald P., *Tensor product type subspace splitting and multilevel iterative methods for anisotropic problems*; Adv. Comp. Math. 4 (1995), pp. 171-206

[15] Hemker P., Pflaum C., *Approximation on partially ordered sets of regular grids*; Applied Numerical Mathematics: Trans. of IMACS, (1997)

[16] Hemker P., Sprengel F., *On the representation of functions and finite difference operators on adaptive sparse grids*; Report CWI (1999)

[17] Hackbusch W., *Iterative Solution of Large Sparse Systems of Equations*; Springer Verlag, New York (1993)

[18] Hochmuth R., *Wavelet Bases in Numerical Analysis and Restricted Nonlinear Approximation*; Habilitationsschrift, Freie Universität Berlin, (1999)

[19] Koster F., *A Proof of the Consistency of the Finite Difference Technique on Sparse Grids*; Computing 65 (2000), pp. 247-261

[20] Koster F., *Numerische Lösung partieller Differentialgleichungen*; PhD thesis, Universtät Bonn, (2001)

[21] Noordmans J., Hemker P., *Application of an adaptive sparse grid technique to a model singular perturbation problem*; Report CWI (1999)

[22] Oswald P., *Best-N-Term capacitance approximation on sparse grids*; in Proc. 11th Intern. Conf. on Domain Decomposition Methods in Science and Engineering, T. Chan et al. (eds.), (2000)

[23] Pflaum, C., *Diskretisierung elliptischer Differentialgleichungen mit dünnen Gittern*; PhD thesis, TU München (1996)

[24] Pflaum, C., *Convergence of the combination technique for the solution of second-order elliptic differential equations*; SIAM J. Numer. Anal. 34 (1997), pp. 2431-2455

[25] Schneider R., *Multiskalen- und Wavelet-Matrixkompression.*; Advances in Numerical Mathematics, Teubner Verlag (1998)

[26] Schiekofer T., *Die Methode der finiten Differenzen auf dünnen Gittern zur Lösung elliptischer und parabolischer PDEs*; PhD thesis, Universität Bonn (1998)

[27] Sprengel F., *Multilevel Algorithms for Finite-Difference Discretizations on Sparse Grids*; Numer. Algorithms 26, (2001), pp. 111-121

[28] Sweldens W., *The Lifting Scheme: a Construction of Second Generation Wavelets*; SIAM J. Math. Anal., Vol. 29,2 (1997), pp. 511–546

[29] Sweldens W., *Wavelets and the lifting scheme: A 5 minute tour*; ZAMM 76, (1996), pp. 41-44

[30] Temlyakov V., *Approximation of periodic functions*; Nova Science Publ. New York (1994)

[31] Triebel H., *Theory of function spaces II*; Birkhäuser Verlag (1992)

[32] Zenger C., *Sparse Grids*; in Hackbusch W. (ed.), Notes on Numerical Fluid Mechanics Vol. 31, Vieweg Verlag (1991)

[33] Zhang F., *Matrix theory*; UTX Springer-Verlag, (1999)

[34] Barret R., Berry M., Chan T. et al., *Templates for the solution of linear systems: Building Blocks for Iterative Methods*; SIAM, Philadelphia, (1994)

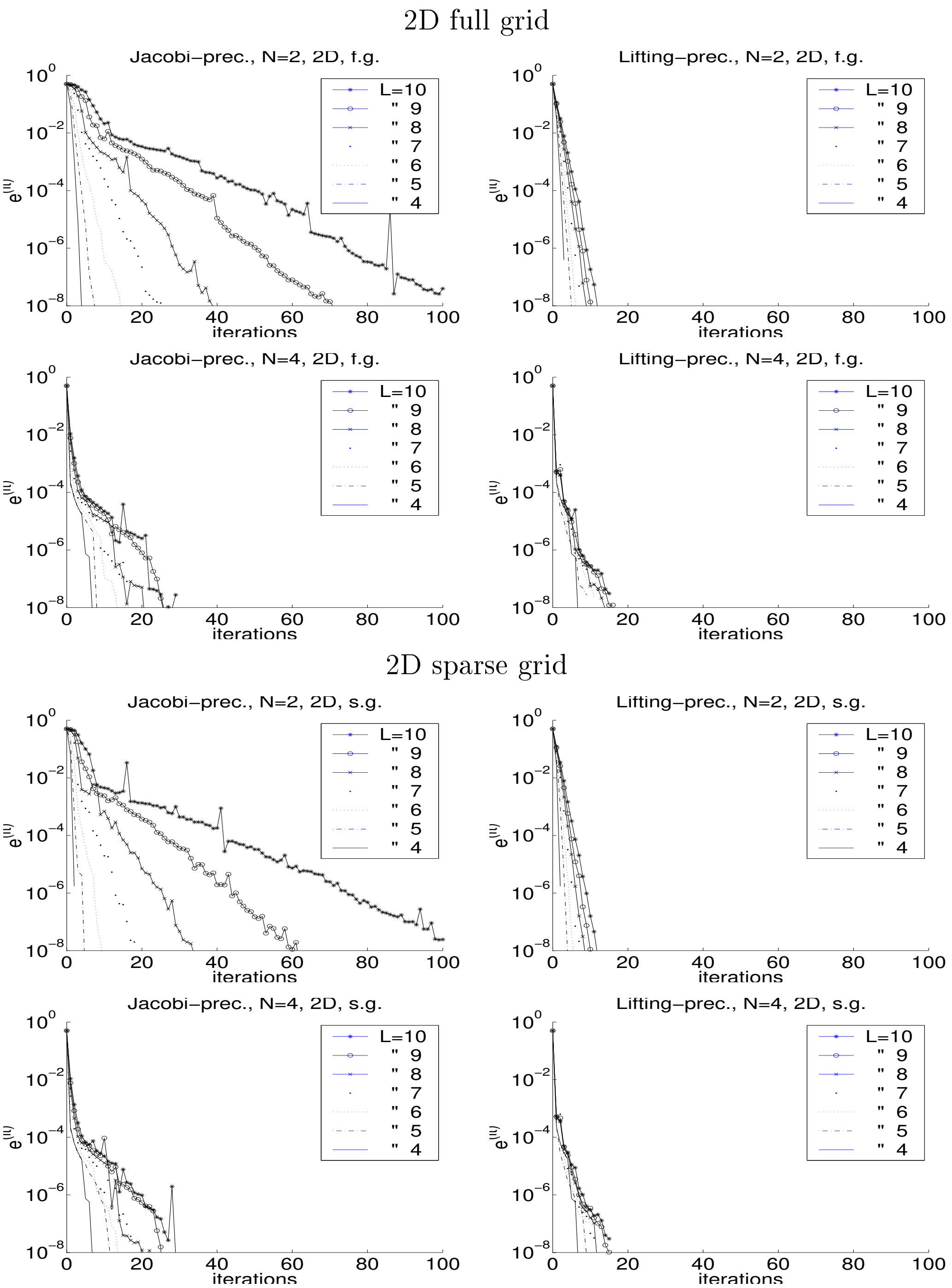
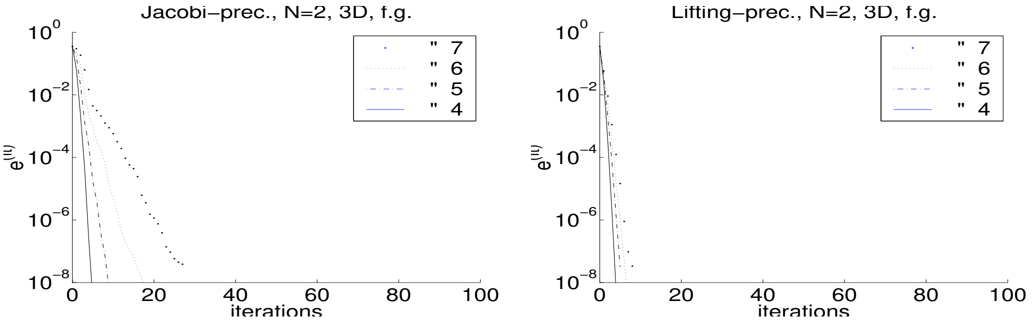


Figure 2: *Left: diagonal-, right: Lifting-preconditioner. First and 3rd row: $N = 2$. second and 4th row $N = 4$.*

3D full grid



3D sparse grid

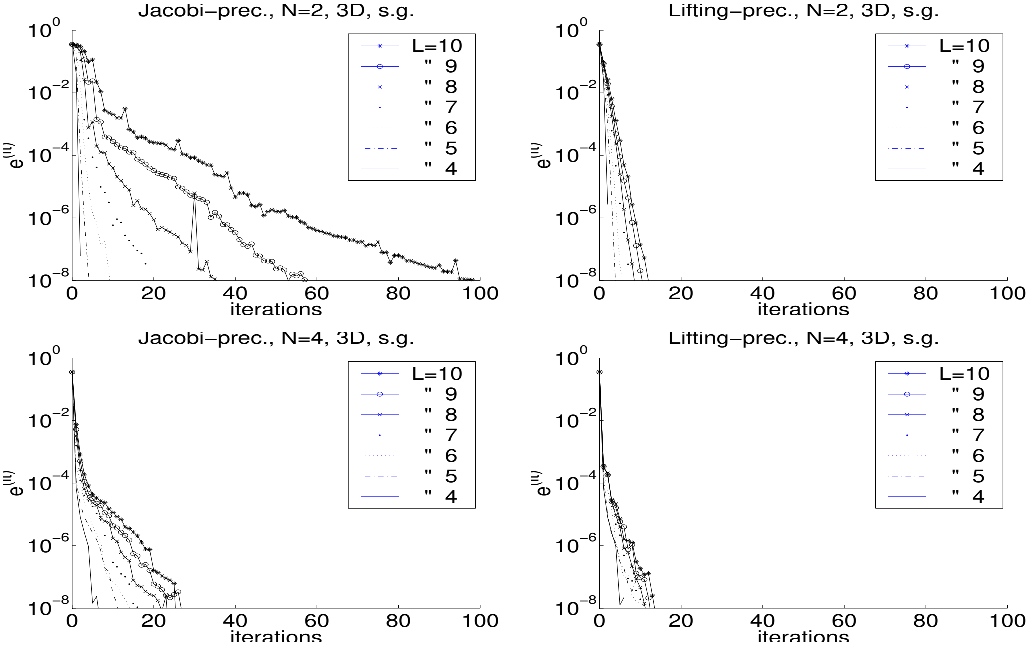


Figure 3: *Left: diagonal-, right: Lifting-preconditioner. First and 3rd row:* $N = 2$. *second and 4th row* $N = 4$.