Multilevel frames for sparse tensor product spaces

H. Harbrecht[2] , R. Schneider[3] and C. Schwab

Research Report No. 2007-06
September 2007

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

[2]Institut für Angewandte Mathematik, Universität Bonn, Wegelerstr. 6, 53115 Bonn, Germany

[3]Institut für Informatik, Christian–Albrechts–Universität zu Kiel, Olshausenstr. 40, 24098 Kiel, Germany

Multilevel frames for sparse tensor product spaces

H. Harbrecht[2] , R. Schneider[3] and C. Schwab

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

Research Report No. 2007-06 September2007

Abstract

For $Au = f$ with an elliptic differential operator $A : \mathcal{H} \to \mathcal{H}'$ and stochastic data f, the m-point correlation function $\mathcal{M}^m u$ of the random solution u satisfies a deterministic, hypoelliptic equation with the m-fold tensor product operator $A^{(m)}$ of A. Sparse tensor products of hierarchic FE-spaces in $\mathcal{H}$ are known to allow for approximations to $\mathcal{M}^m u$ which converge at essentially the rate as in the case $m = 1$, i.e. for the deterministic problem. They can be realized by wavelet-type FE bases [28]. If wavelet bases are not available, we show here how to achieve log-linear complexity computation of sparse approximations of $\mathcal{M}^m u$ for Galerkin discretizations of A by multilevel frames such as BPX or other multilevel preconditioners of any standard FEM approximation for A. Numerical examples illustrate feasibility and scope of the method.

[2]Institut für Angewandte Mathematik, Universität Bonn, Wegelerstr. 6, 53115 Bonn, Germany

[3]Institut für Informatik, Christian–Albrechts–Universität zu Kiel, Olshausenstr. 40, 24098 Kiel, Germany

1. Introduction

Up to now sparse tensor product spaces were constructed with the help of wavelet or multilevel bases, e.g. hierarchical bases, see [5, 6, 7, 17, 34]. However, there are some drawbacks at least from the practical point of view: on the one hand, wavelets are not easy to construct on complicated domains or manifolds. On the other hand, the computation of system matrices in wavelet coordinates is a nontrivial task since then also interactions of functions on different levels will appear.

In the present paper we are going to use a multilevel frame instead of wavelet or multilevel bases to approximate functions from the sparse tensor product space. Especially, we have in mind standard multigrid hierarchies and traditional finite elements. The frame construction is based on the BPX-preconditioner (see e.g. [4, 11, 18, 19, 27]) and related generating systems (see e.g. [16, 17]).

We intend to present a self-contained paper on multilevel frames for sparse tensor product spaces. The application of frames to partial differential equations has been established in very recent years, see [9, 10, 33]. Most, if not all results have analogues in the context of generating systems as considered in the pioneering publications of Griebel and Oswald [16, 17, 18, 19, 27]. Under the perspective of frames the earlier results on generating systems and stable splittings gain a new interpretation and actuality.

Since the representation of functions with respect to a frame is non-unique, stiffness matrices corresponding to bijective operators have a nontrivial, in general large kernel. However, the load vector with respect to the frame lies in the image of the stiffness matrix. Therefore, Krylov subspace methods, such as the conjugate gradient method or GMRES, will converge without further modifications (see, e.g., [10, 16, 17, 25]). This is due to the fact that (in exact arithmetic) the Krylov subspace, and thus the residuum, stays orthogonal to the kernel.

In particular, preconditioning becomes obsolete provided that the frame is normalized with respect to the energy space $\mathcal{H}$ of the elliptic operator A. In the present paper, we use a finite element frame for elliptic differential operators of positive order derived from the BPX-preconditioner.

Sparse tensor product spaces allow linear complexity computation of deterministic solution statistics for elliptic partial differential equations with stochastic data, see for example [20, 28, 31, 32].

As an example, the two-point and m-point correlation functions of second order elliptic problems with stochastic source terms are known to satisfy a deterministic, hypoelliptic partial differential equation of order $2m$ with m-fold tensor product of the elliptic operator on the m-fold cartesian product Ω^m of the bounded physical domain $\Omega \subset \mathbb{R}^n$, i.e. in a computational domain of dimension mn.

Specifically, let A denote a linear, second order elliptic partial differential operator that maps the Hilbert space $\mathcal{H}$ onto its dual $\mathcal{H}'$. Typically one might think of $\mathcal{H}$ being a Sobolev space with dual $\mathcal{H}'$.

For a given stochastic load vector $f \in \mathcal{H}'$ with known expectation and two-point correlation, we consider the stochastic operator equation

$$Au = f.$$

Then the random solution's expectation $\mathrm{E}(u)$ satisfies the mean field equation

$$A\,\mathrm{E}(u) = \mathrm{E}(f) \tag{1.1}$$

while its two-point correlation is given by

$$(A \otimes A)\,\mathrm{Cor}(u) = \mathrm{Cor}(f), \tag{1.2}$$

see [28, 31] for details; here $E(u)$ denotes the expectation (or "ensemble average") for the random field u taking values in $\mathcal{H}$, and $\mathrm{Cor}(u) = E(u \otimes u)$ where now $E(\cdot)$ denotes the expectation with respect to the product measure on the tensor product space $\mathcal{H} \otimes \mathcal{H}$ (see [28, 31]). Similarly, deterministic hypoelliptic equations can be obtained also for m-th order moments of the random field u (see [32]).

Regularity results for tensor product operators resp. boundary value problems, particularly non-hypoelliptic equations, can be found in [14, 22, 23, 24].

In the present paper, we present a new algorithm for the approximate solution of (1.2) with a complexity that stays essentially proportional to the number of unknowns N required for discretizing the domain or manifold of the mean field equation (1.1). Similar algorithmical techniques have been used in [1, 2, 29]. Unlike e.g. [28, 32], our algorithm does not require explicit hierarchical or wavelet bases for the "increment" spaces in the multilevel hierarchy used in Galerkin approximation of the mean-field equation (1.1), but rather works with standard finite element shape functions.

Here and throughout the paper, "essentially" means up to powers of $\log N$ resp., in the context of convergence rates, up to powers of $|\log h|$. Our algorithm involves only prolongations, restrictions, and finite element stiffness matrices associated with the elliptic operator A in a standard, nonhierarchical FE-Basis. These ingredients are provided by standard finite element tools.

We emphasize that the algorithm in the present form is only feasible for partial differential operators, that is, for *local* operators. Nonlocal operators, like integral operators, induce densely populated system matrices such that the approach is no longer efficient. Here, one has to apply modern boundary element methods like the multipole method [15], adaptive cross approximation [3] or the $\mathcal{H}$-matrix approach [21] to approximate the dense system matrices. In case of wavelet matrix compression for nonlocal operators (see [13] and the references therein) one can use the wavelet basis also for the sparse tensor product approximation of (1.2), see e.g. [20, 28].

The outline of the paper is as follows. In Section 2 we give the mathematical theory of the multilevel frame. We introduce the finite element spaces and define the frame in one coordinate. Then, we show that the tensor product of this frame is still a frame. We obtain the frame for the sparse tensor product space $\widehat{V}_J$ by restriction of the given set of basis functions to all functions that are contained in $\widehat{V}_J$. Section 3 is dedicated to the efficient solution of tensor product equations like (1.2). We develop an algorithm that computes the solution of (1.2) in essentially the complexity needed to discretize the corresponding mean field equation. Finally, in Section 4 we present two numerical examples to demonstrate our method.

Throughout the paper, in order to avoid the repeated use of generic but unspecified constants, by $C \lesssim D$ we mean that C can be bounded by a multiple of D, independently of parameters which C and D may depend on. Obviously, $C \gtrsim D$ is defined as $D \lesssim C$, and $C \sim D$ as $C \lesssim D$ and $C \gtrsim D$.

2. Multilevel frames

2.1. **Frames.** Let $\mathcal{H}$ be a separable Hilbert space with dual $\mathcal{H}'$. We denote the duality product on $(\mathcal{H}, \mathcal{H}')$ by $\langle\cdot,\cdot\rangle$. Typically we have here in mind a Sobolev space and the L^2-inner product. According to [8, 12, 33] a frame for $\mathcal{H}$ is defined as follows:

Definition 1. A countable collection $\Phi = \{\varphi_i : i \in \Delta\} \subset \mathcal{H}$ is called a frame for $\mathcal{H}$ if there holds

$$C\|f\|_{\mathcal{H}'}^2 \leq \sum_{i\in\Delta} |\langle f, \varphi_i\rangle|^2 \leq D\|f\|_{\mathcal{H}'}^2 \tag{2.1}$$

for all $f \in \mathcal{H}'$.

In what follows the collection Φ of functions will be viewed as a row vector, which means that for $\mathbf{f} = [f_i]_{i\in\Delta} \in \ell^2(\Delta)$, the function $f = \Phi\mathbf{f}$ is defined as $v = \sum_{i\in\Delta} f_i\varphi_i$.

As a consequence of (2.1), the frame operator

$$F : \mathcal{H}' \to \ell^2(\Delta), \qquad f \mapsto [\langle f, \varphi_i\rangle]_{i\in\Delta}$$

and its dual

$$F' : \ell^2(\Delta) \to \mathcal{H}, \qquad \mathbf{f} \mapsto \Phi\mathbf{f}$$

are bounded with constant $\sqrt{D}$. The composition $F'F : \mathcal{H}' \to \mathcal{H}$ is boundedly invertible and we have

$$\ell^2(\Delta) = \operatorname{ran} F \overset{\perp}{\oplus} \ker F'.$$

In particular, the collection $\widetilde{\Phi} := \Phi(F'F)^{-1}$ is a frame for $\mathcal{H}'$, the so-called canonical dual frame, satisfying

$$\frac{1}{D}\|f\|_{\mathcal{H}}^2 \leq \sum_{i\in\Delta} |\langle f, \widetilde{\varphi}_i\rangle|^2 \leq \frac{1}{C}\|f\|_{\mathcal{H}}^2$$

for all $f \in \mathcal{H}$. The associated frame operators are given by

$$\widetilde{F} = F(F'F)^{-1}, \qquad \widetilde{F}' = (F'F)^{-1}F'.$$

Functions $u \in \mathcal{H}$ and $v \in \mathcal{H}'$ own the unique representations

$$u = \Phi\widetilde{F}u = \sum_{i\in\Delta}\langle u, \widetilde{\phi}_i\rangle\phi_i, \qquad v = \widetilde{\Phi}Fv = \sum_{i\in\Delta}\langle v, \phi_i\rangle\widetilde{\phi}_i.$$

The operator

$$\mathbf{Q} := F(F'F)^{-1}F' = \widetilde{F}(\widetilde{F}'\widetilde{F})^{-1}\widetilde{F}' : \ell^2(\Delta) \to \ell^2(\Delta)$$

is the orthogonal projector onto $\operatorname{ran} F = \operatorname{ran} \widetilde{F}$.

We shall consider a linear elliptic operator $A : \mathcal{H} \to \mathcal{H}'$ and a load vector $f \in \mathcal{H}'$. Then, the solution $u = F'\mathbf{u}$ of the partial differential equation $Au = f$ satisfies the infinite dimensional discrete equation

$$\mathbf{A}\mathbf{u} = \mathbf{f}, \qquad \mathbf{A} = FAF' = \langle A\Phi, \Phi\rangle, \qquad \mathbf{f} = Ff = \langle f, \Phi\rangle. \tag{2.2}$$

Herein, $\mathbf{A}|_{\operatorname{ran} F} : \operatorname{ran} F \to \operatorname{ran} F$ is boundedly invertible, whereas $\ker \mathbf{A} = \ker F$. More precisely, we have the lemma:

Lemma 2. *Let $A : \mathcal{H} \to \mathcal{H}'$ be a given elliptic partial differential operator satisfying*

$$\gamma\|u\|_{\mathcal{H}}^2 \leq |\langle Au, u\rangle|, \qquad |\langle Au, v\rangle| \leq \Gamma\|u\|_{\mathcal{H}}\|v\|_{\mathcal{H}},$$

for $0 < \gamma \leq \Gamma$. Then, there holds

$$\big\|\mathbf{A}|_{\operatorname{ran} F}^{-1}\big\|_{\ell^2(\Delta)\to\ell^2(\Delta)} \leq \frac{1}{\gamma C}, \qquad \|\mathbf{A}\|_{\ell^2(\Delta)\to\ell^2(\Delta)} \leq D\Gamma.$$

Proof. On the one hand we have for $\mathbf{v} \in \operatorname{ran} F$ that

$$\begin{aligned}\|\mathbf{A}\mathbf{v}\|_{\ell^2(\Delta)} = \|FAF'\mathbf{v}\|_{\ell^2(\Delta)} &= \sup_{\mathbf{u}\in\ell^2(\Delta)\setminus\{\mathbf{0}\}} \frac{\langle AF'\mathbf{v}, F'\mathbf{u}\rangle}{\|\mathbf{u}\|_{\ell^2(\Delta)}} \geq \gamma\frac{\|F'\mathbf{v}\|_{\mathcal{H}'}^2}{\|\mathbf{v}\|_{\ell^2(\Delta)}}\\ &\geq \gamma C\|\mathbf{v}\|_{\ell^2(\Delta)},\end{aligned}$$

which implies the first estimate. On the other hand there holds for all $\mathbf{v} \in \ell^2(\Delta)$

$$\begin{aligned}\|\mathbf{A}\mathbf{v}\|_{\ell^2(\Delta)} = \|FAF'\mathbf{v}\|_{\ell^2(\Delta)} &\leq \|F\|_{\ell^2(\Delta)\to\mathcal{H}'}\|A\|_{\mathcal{H}\to\mathcal{H}'}\|F'\|_{\mathcal{H}\to\ell^2(\Delta)}\|\mathbf{v}\|_{\ell^2(\Delta)} \\ &\leq D\Gamma\|\mathbf{v}\|_{\ell^2(\Delta)}.\end{aligned}$$

□

With the help of this lemma we conclude that (2.2) is well-posed in $\operatorname{ran} F \subset \ell^2(\Delta)$ since $\mathbf{f} \in \operatorname{ran} F$. But in general the frame Φ consists of an infinite number of functions. Therefore, we choose an appropriate finite index set $\nabla \subset \Delta$ and consider the collection $\Psi = \{\varphi_i : i \in \nabla\} \subset \Phi$ for numerical approximation:

Definition 3. We call the collection $\Psi = \{\varphi_i : i \in \nabla\} \subset \mathcal{H}$ a subframe of Φ if there holds

$$C\|f\|^2_{\mathcal{H}'} \leq \sum_{i\in\nabla} |\langle f, \varphi_i\rangle|^2 \leq D\|f\|^2_{\mathcal{H}'} \tag{2.3}$$

for all $f \in V := \operatorname{span} \Psi$.

With the frame operators

$$\begin{aligned} G &: \mathcal{H}' \to \ell^2(\nabla), \qquad f \mapsto [\langle f, \varphi_i\rangle]_{i\in\nabla} \\ G' &: \ell^2(\nabla) \to \mathcal{H}, \qquad \mathbf{f} \mapsto \Psi\mathbf{f}\end{aligned}$$

it follows by the same reasoning as above that the Galerkin matrix $\mathbf{B} := \langle A\Psi, \Psi\rangle$ satisfies

$$\|\mathbf{B}|^{-1}_{\operatorname{ran} G}\| \leq \frac{1}{\gamma C}, \qquad \|\mathbf{B}\| \leq D\Gamma.$$

This means that the Galerkin system $\mathbf{B}\mathbf{u} = \mathbf{g}$ is well-posed due to $\mathbf{g} := \langle f, \Psi\rangle \in \operatorname{ran} G$.

2.2. **Multiresolution analyses.** Let $\Omega \subset \mathbb{R}^n$ be a sufficiently smooth, bounded domain. We consider a dense, nested sequence of finite dimensional subspaces

$$V_0 \subset V_1 \subset \ldots \subset V_j \ldots \subset L^2(\Omega), \tag{2.4}$$

consisting of piecewise polynomial ansatz functions $V_j = \operatorname{span}\{\varphi_{j,k} : k \in \Delta_j\}$, such that $\dim V_j \sim 2^{jn}$ and

$$L^2(\Omega) = \overline{\bigcup_{j\in\mathbb{N}_0} V_j}, \qquad V_0 = \bigcap_{j\in\mathbb{N}_0} V_j. \tag{2.5}$$

Since we are going to use the spaces V_j as test and trial spaces for the approximate solution in a Galerkin method, we shall assume that the following Jackson and Bernstein type estimates hold for $s \leq t < \gamma$, $t \leq q \leq d$,

$$\inf_{v_j\in V_j} \|u - v_j\|_{H^t(\Omega)} \lesssim h_j^{q-t}\|u\|_{H^q(\Omega)}, \quad u \in H^q(\Omega), \tag{2.6}$$

and

$$\|v_j\|_{H^t(\Omega)} \lesssim h_j^{s-t}\|v_j\|_{H^s(\Omega)}, \quad v_j \in V_j, \tag{2.7}$$

uniformly in j, where we set $h_j := 2^{-j}$. Notice that the parameter $d > 0$ refers to the maximal degree of polynomials which are locally contained in V_j while

$$\gamma := \sup\left\{t \in \mathbb{R} : V_j \subset H^t(\Omega)\right\} > 0$$

indicates the regularity or smoothness of the functions contained in the spaces V_j. The quantity h_j corresponds to the mesh width of the mesh associated with the subspace V_j on Ω. Note that for ansatz functions based on cardinal B-splines there holds $\gamma = d - 1/2$ and for the standard, piecewise polynomial, continuous Finite Element shape functions $\gamma = 3/2$.

We observe that the L^2-orthogonal projection $P_j : L^2(\Omega) \to V_j$ is the Galerkin approximation of the equation $Ix = y$. Applying an Aubin-Nitsche duality argument to the corresponding Galerkin equation, the above assumptions imply for $-d \le t \le \gamma$, $t \le q$, $0 \le q \le d$

$$\|(I - P_j)v\|_{H^t(\Omega)} \lesssim h_j^{q-t}\|v\|_{H^q(\Omega)}, \quad v \in H^q(\Omega).$$

We shall also introduce projectors on the “detail” or “increment” spaces:

$$Q_j := P_j - P_{j-1}, \quad j \in \mathbb{N}, \qquad\qquad Q_0 := P_0.$$

There holds the following lemma.

Lemma 4. *The operators Q_j are $L^2(\Omega)$-orthogonal projectors which satisfy*

$$Q_j P_\ell = \begin{cases} 0, & j > \ell, \\ Q_j, & j \le \ell. \end{cases}$$

Proof. For $\ell \le j$ it follows from $V_\ell \subset V_j$ that $P_\ell P_j = P_\ell$ and likewise $P_j P_\ell = P_\ell$. Therefore, we find

$$\begin{aligned} Q_j^2 &= (P_j - P_{j-1})(P_j - P_{j-1}) \\ &= P_j^2 - P_j P_{j-1} - P_{j-1} P_j + P_{j-1}^2 \\ &= P_j - P_{j-1} \\ &= Q_j. \end{aligned}$$

Since the projectors P_j are orthogonal, it follows

$$Q_j^\star = P_j^\star - P_{j-1}^\star = P_j - P_{j-1} = Q_j,$$

i.e., Q_j is the $L^2(\Omega)$-orthogonal projector onto $W_j = V_j \overset{\perp}{\ominus} V_{j-1}$.

Using once more that $V_\ell \subset V_j$ implies $P_\ell P_j = P_\ell$ and $P_j P_\ell = P_\ell$, we conclude the second part of the assertion from

$$Q_j P_\ell = (P_j - P_{j-1})P_\ell = \begin{cases} P_\ell - P_\ell, & j > \ell, \\ P_j - P_{j-1}, & j \le \ell. \end{cases}$$

□

Property (2.5) implies for all $f \in L^2(\Omega)$ the identity

$$f = \sum_{j\in\mathbb{N}_0} Q_j f, \tag{2.8}$$

while (2.6) and (2.7) yield the multilevel norm equivalence

$$\|f\|^2_{H^s(\Omega)} \sim \sum_{j\in\mathbb{N}_0} 2^{2sj}\|Q_j f\|_{L^2(\Omega)}, \qquad -\gamma \le s \le \gamma. \tag{2.9}$$

In the sequel the space $H^q(\Omega)$, $0 < q \le \gamma$, will be considered as energy space V of an elliptic operator equation to be solved by the Galerkin method. Hence, we suppose that the basis $\Phi_j = \{\varphi_{j,k} : k \in \Delta_j\}$ is normalized with respect to this space:

$$\|\varphi_{j,k}\|_{H^q(\Omega)} \sim 1.$$

For the efficient discretization it is important that the basis functions are compactly supported in terms of

$$\operatorname{diam} \varphi_{j,k} \sim 2^{-j}.$$

Finally, the basis is assumed to be *stable*, i.e., there holds the following Riesz-property

$$\sum_{k\in\Delta_j} |\langle f, \varphi_{j,k}\rangle|^2 \sim 2^{-2qj}\|P_j f\|^2_{L^2(\Omega)}. \tag{2.10}$$

For example, one might think of the energy space $H^1(\Omega)$ and properly scaled continuous nodal basis functions defined on a multigrid hierarchy resulting from uniform refinement of a given coarse grid.

2.3. **Multilevel frames for $H^q(\Omega)$.** We show next that the collection

$$\Phi = \{\varphi_{j,k} : k \in \Delta_j,\ j \in \mathbb{N}_0\}$$

defines a frame for $\mathcal{H} = H^q(\Omega)$. Notice that this frame underlies the construction of the so-called BPX preconditioner, see e.g. [4, 11, 27].

Theorem 5. *The collection of functions* $\Phi = \{\varphi_{j,k} : k \in \Delta_j,\ j \in \mathbb{N}_0\}$ *defines a frame for the energy space* $H^q(\Omega)$.

Proof. We have to show that

$$\sum_{j\in\mathbb{N}_0}\sum_{k\in\Delta_j}|\langle f,\varphi_{j,k}\rangle|^2\sim\|f\|^2_{H^{-q}(\Omega)}\qquad\forall f\in H^{-q}(\Omega).$$

Using (2.10) we find

$$\sum_{j\in\mathbb{N}_0}\sum_{k\in\Delta_j}|\langle f,\varphi_{j,k}\rangle|^2\sim\sum_{j\in\mathbb{N}_0}2^{-2qj}\|P_jf\|^2_{L^2(\Omega)}.$$

By (2.8) and Lemma 4 we conclude

$$P_jf=\sum_{\ell\in\mathbb{N}_0}Q_\ell P_jf=\sum_{\ell=0}^{j}Q_\ell f.$$

This yields together with (2.9)

$$\sum_{j\in\mathbb{N}_0}\sum_{k\in\Delta_j}|\langle f,\varphi_{j,k}\rangle|^2\sim\sum_{j\in\mathbb{N}_0}2^{-2qj}\bigg\|\sum_{\ell=0}^{j}Q_\ell f\bigg\|^2_{L^2(\Omega)}\sim\sum_{j\in\mathbb{N}_0}2^{-2qj}\sum_{\ell=0}^{j}\|Q_\ell f\|^2_{L^2(\Omega)}.$$

Since there further holds

$$\{(j,\ell):0\le j<\infty,\,0\le\ell\le j\}=\{(j,\ell):0\le\ell<\infty,\,\ell\le j<\infty\},\tag{2.11}$$

we obtain finally

$$\begin{aligned}\sum_{j\in\mathbb{N}_0}\sum_{k\in\Delta_j}|\langle f,\varphi_{j,k}\rangle|^2&\sim\sum_{\ell\in\mathbb{N}_0}\sum_{j=\ell}^{\infty}2^{-2qj}\|Q_\ell f\|^2_{L^2(\Omega)}\\&\sim\sum_{\ell\in\mathbb{N}_0}2^{-2q\ell}\|Q_\ell f\|^2_{L^2(\Omega)}\\&\sim\|f\|^2_{H^{-q}(\Omega)},\end{aligned}$$

where the last step follows again by the norm equivalence (2.9). □

Remark 6. As one readily verifies the collection $\{\varphi_{j,k}:k\in\Delta_j,\,j\le J\}$ is a subframe of the frame Φ.

2.4. **Multilevel frames for tensor product spaces.** We are now going to consider the discretization of functions defined on the product domain $\Omega^m=\Omega\times\cdots\times\Omega$. To this end, we first introduce some notation. For multi-indices $\mathbf{s}=(s_1,s_2,\ldots,s_m)$, $\mathbf{t}=(t_1,t_2,\ldots,t_m)\in\mathbb{R}^m$ we will write

$$\mathbf{s}\le\mathbf{t}\quad:\Longleftrightarrow\quad s_1\le t_1,s_2\le t_2,\ldots,s_m\le t_m.$$

Obviously, for $\mathbf{s}\ge\mathbf{0}$ there holds

$$\|\mathbf{s}\|_{\ell^1}=\sum_{i=1}^{m}s_i,\qquad\|\mathbf{s}\|_{\ell^\infty}=\max_{i=1}^{m}s_i.$$

We define the tensor product Sobolev spaces

$$H^{\mathbf{s}}(\Omega^m) := H^s(\Omega) \otimes H^s(\Omega) \otimes \cdots \otimes H^s(\Omega) \subset L^2(\Omega^m), \quad s \geq 0.$$

Their duals with respect to the "pivot" space $L^2(\Omega)$ satisfy

$$H^{-\mathbf{s}}(\Omega^m) = \big(H^{\mathbf{s}}(\Omega^m)\big)' = H^{-s}(\Omega) \otimes \cdots \otimes H^{-s}(\Omega).$$

For a multiindex $\mathbf{j} \in \mathbb{N}_0^m$ we set

$$\Delta_{\mathbf{j}} := \Delta_{j_1} \times \Delta_{j_2} \times \cdots \times \Delta_{j_m}.$$

Then, tensor products of the basis functions

$$\varphi_{\mathbf{j},\mathbf{k}}(\mathbf{x}) := \varphi_{j_1,k_1}(x_1)\varphi_{j_2,k_2}(x_2)\cdots\varphi_{j_n,k_n}(x_n), \quad \mathbf{j} \in \mathbb{N}_0^m, \quad \mathbf{k} \in \Delta_{\mathbf{j}}, \quad \mathbf{x} \in \Omega^m,$$

generate the tensor product spaces

$$V_{\mathbf{j}} := V_{j_1} \otimes V_{j_2} \otimes \cdots \otimes V_{j_m} = \operatorname{span}\{\varphi_{\mathbf{j},\mathbf{k}} : \mathbf{k} \in \Delta_{\mathbf{j}}\}.$$

We denote the L^2-orthogonal projection onto $V_{\mathbf{j}}$ by $P_{\mathbf{j}} = P_{j_1} \otimes P_{j_2} \otimes \cdots \otimes P_{j_m}$ and we define $Q_{\mathbf{j}} := Q_{j_1} \otimes Q_{j_2} \otimes \cdots \otimes Q_{j_m}$. Obviously, in view of (2.10), the collection $\{\varphi_{\mathbf{j},\mathbf{k}}\}_{\mathbf{k}\in\Delta_{\mathbf{j}}}$ is a stable basis of $V_{\mathbf{j}}$ in the energy space $H^{\mathbf{q}}(\Omega^m)$:

$$\sum_{\mathbf{k}\in\Delta_{\mathbf{j}}} |\langle f, \varphi_{\mathbf{j},\mathbf{k}}\rangle|^2 \sim 2^{-2q\mathbf{j}} \|P_{\mathbf{j}} f\|^2_{L^2(\Omega^m)}. \tag{2.12}$$

Moreover, we get from (2.8) that

$$f = \sum_{\mathbf{j}\in\mathbb{N}_0^m} Q_{\mathbf{j}} f. \tag{2.13}$$

Equation (2.9) and standard tensor product arguments imply

$$\|f\|^2_{H^{\mathbf{s}}(\Omega^m)} \sim \sum_{\mathbf{j}\in\mathbb{N}_0^m} 2^{2s\mathbf{j}} \|Q_{\mathbf{j}} f\|^2_{L^2(\Omega^m)}, \qquad -\gamma \leq s \leq \gamma. \tag{2.14}$$

Theorem 7. *The collection of functions $\Phi^m := \{\varphi_{\mathbf{j},\mathbf{k}} : \mathbf{k} \in \Delta_{\mathbf{j}}, \mathbf{j} \in \mathbb{N}_0^m\}$ defines a frame for the energy space $H^{\mathbf{q}}(\Omega^m)$, that is, there holds*

$$\sum_{\mathbf{j}\in\mathbb{N}_0^m} \sum_{\mathbf{k}\in\Delta_{\mathbf{j}}} |\langle f, \varphi_{\mathbf{j},\mathbf{k}}\rangle|^2 \sim \|f\|^2_{H^{-\mathbf{q}}(\Omega^m)} \qquad \forall f \in H^{-\mathbf{q}}(\Omega^m).$$

Proof. We just have to extend the proof of Theorem 5: Due to (2.12) there holds

$$\sum_{\mathbf{j}\in\mathbb{N}_0^m} \sum_{\mathbf{k}\in\Delta_{\mathbf{j}}} |\langle f, \varphi_{\mathbf{j},\mathbf{k}}\rangle|^2 \sim \sum_{\mathbf{j}\in\mathbb{N}_0^m} 2^{-2q\mathbf{j}} \|P_{\mathbf{j}} f\|^2_{L^2(\Omega^m)}.$$

Since $V_{\mathbf{j}} \supset V_{\boldsymbol{\ell}}$ for $\boldsymbol{\ell} \leq \mathbf{j}$ we have $P_{\boldsymbol{\ell}} P_{\mathbf{j}} = P_{\boldsymbol{\ell}}$ and likewise $P_{\mathbf{j}} P_{\boldsymbol{\ell}} = P_{\boldsymbol{\ell}}$. Hence, we find

$$P_{\mathbf{j}} f = \sum_{\boldsymbol{\ell}\in\mathbb{N}_0^m} Q_{\boldsymbol{\ell}} P_{\mathbf{j}} f = \sum_{\mathbf{0}\leq\boldsymbol{\ell}\leq\mathbf{j}} Q_{\boldsymbol{\ell}} f.$$

This identity leads because of (2.14) to

$$\sum_{\mathbf{j}\in\mathbb{N}_0^m}\sum_{\mathbf{k}\in\Delta_{\mathbf{j}}}|\langle f,\varphi_{\mathbf{j},\mathbf{k}}\rangle|^2 \sim \sum_{\mathbf{j}\in\mathbb{N}_0^m}\sum_{\mathbf{0}\le\boldsymbol{\ell}\le\mathbf{j}} 2^{-2q\mathbf{j}}\|Q_{\boldsymbol{\ell}}f\|_{L^2(\Omega^m)}^2.$$

We reorder the index sets analogously to (2.11) and conclude by (2.14)

$$\begin{aligned}\sum_{\mathbf{j}\in\mathbb{N}_0^m}\sum_{\mathbf{k}\in\Delta_{\mathbf{j}}}|\langle f,\varphi_{\mathbf{j},\mathbf{k}}\rangle|^2 &\sim \sum_{\boldsymbol{\ell}\in\mathbb{N}_0^m}\sum_{\mathbf{j}\ge\boldsymbol{\ell}} 2^{-2q\mathbf{j}}\|Q_{\boldsymbol{\ell}}f\|_{L^2(\Omega^m)}^2\\ &\sim \sum_{\boldsymbol{\ell}\in\mathbb{N}_0^m} 2^{-2q\boldsymbol{\ell}}\|Q_{\boldsymbol{\ell}}f\|_{L^2(\Omega^m)}^2\\ &\sim \|f\|_{H^{-\mathbf{q}}(\Omega^m)}^2.\end{aligned}$$

□

2.5. **Sparse tensor product spaces.** Instead of the full tensor product space

$$V_{J,J,\ldots,J} = V_J\otimes V_J\otimes\cdots\otimes V_J = \sum_{\|\mathbf{j}\|_{\ell^\infty}=J} V_{\mathbf{j}} \subset H^{\mathbf{q}}(\Omega^m)$$

we shall consider the *sparse* tensor product space

$$\widehat{V}_{J,J,\ldots,J} = \sum_{\|\mathbf{j}\|_{\ell^1}=J} V_{\mathbf{j}} \subset H^{\mathbf{q}}(\Omega^m).$$

Abbreviating $N_J := \dim V_J$ there holds $\widehat{N}_{\mathbf{J}} = \dim \widehat{V}_{J,J,\ldots,J} \sim N_J \log^{m-1} N_J$, cf. [7], which is substantially smaller than the dimension N_J^m of the full tensor product space $V_{J,J,\ldots,J}$.

The following lemma, proven in [28, 32], tells us that the approximation power in the sparse tensor product spaces is nearly as good as in the full tensor product spaces, provided that the given function has bounded mixed derivatives of order d.

Lemma 8. *We denote the L^2-orthogonal projection onto the sparse tensor product space $\widehat{V}_{J,J,\ldots,J}$ by*

$$\widehat{P}_{J,J,\ldots,J} = \sum_{\|\mathbf{j}\|_{\ell^1}\le J} Q_{\mathbf{j}}.$$

Then, for $0\le s<\gamma$, $s\le t\le d$ there holds

$$\big\|u-\widehat{P}_{J,J,\ldots,J}u\big\|_{H^{\mathbf{s}}(\Omega^m)} \lesssim \begin{cases} 2^{J(s-d)}J^{(m-1)/2}\|u\|_{H^{\mathbf{d}}(\Omega^m)}, & \text{if } t=d,\\ 2^{J(s-t)}\|u\|_{H^{\mathbf{t}}(\Omega^m)}, & \text{otherwise.}\end{cases}$$

In the sparse tensor product spaces we like to represent functions by the collection $\widehat{\Phi}^m := \{\varphi_{\mathbf{j},\mathbf{k}} : \mathbf{k}\in\Delta_{\mathbf{j}},\ \|\mathbf{j}\|_{\ell^1}\le J\}$.

Theorem 9. *The collection $\widehat{\Phi}^m = \{\varphi_{\mathbf{j},\mathbf{k}} : \mathbf{k}\in\Delta_{\mathbf{j}},\ \|\mathbf{j}\|_{\ell^1}\le J\}$ defines a subframe of Φ^m.*

Proof. Analogously to the proof of Theorem 7 we get

$$\sum_{\|\mathbf{j}\|_{\ell^1}\leq J}\sum_{\mathbf{k}\in\Delta_{\mathbf{j}}}|\langle f,\varphi_{\mathbf{j},\mathbf{k}}\rangle|^2\sim\sum_{\|\mathbf{j}\|_{\ell^1}\leq J}\sum_{\mathbf{0}\leq\boldsymbol{\ell}\leq\mathbf{j}}2^{-2q\mathbf{j}}\|Q_{\boldsymbol{\ell}}f\|^2_{L^2(\Omega^m)}$$

for all $f\in\widehat{V}_{J,J,\ldots,J}$. Next, from the identity

$$\{\|\mathbf{j}\|_{\ell^1}\leq J,\,\mathbf{0}\leq\boldsymbol{\ell}\leq\mathbf{j}\}=\{\|\mathbf{j}\|_{\ell^1},\|\boldsymbol{\ell}\|_{\ell^1}\leq J,\,\boldsymbol{\ell}\leq\mathbf{j}\}$$

we conclude by (2.14)

$$\begin{aligned}\sum_{\|\mathbf{j}\|_{\ell^1}\leq J}\sum_{\mathbf{k}\in\Delta_{\mathbf{j}}}|\langle f,\varphi_{\mathbf{j},\mathbf{k}}\rangle|^2&\sim\sum_{\|\boldsymbol{\ell}\|_{\ell^1}\leq J}\sum_{\substack{\|\mathbf{j}\|_{\ell^1}\leq J\\ \boldsymbol{\ell}\leq\mathbf{j}}}2^{-2q\mathbf{j}}\|Q_{\boldsymbol{\ell}}f\|^2_{L^2(\Omega^m)}\\&\sim\sum_{\|\boldsymbol{\ell}\|_{\ell^1}\leq J}2^{-2q\boldsymbol{\ell}}\|Q_{\boldsymbol{\ell}}f\|^2_{L^2(\Omega^m)}\\&\sim\|f\|^2_{H^{-\mathbf{q}}(\Omega^m)}.\end{aligned}$$

□

Concerning the cardinality of the subframe $\widehat{\Phi}^m$ we have the following statement.

Theorem 10. *With $N_J:=\dim V_J$ there holds*

$$\operatorname{card}\big(\{\varphi_{\mathbf{j},\mathbf{k}}:\mathbf{k}\in\Delta_{\mathbf{j}},\,\|\mathbf{j}\|_{\ell^1}\leq J\}\big)\lesssim N_J\log^{m-1}N_J.$$

Proof. Since $\operatorname{card}(\Delta_{\mathbf{j}})\sim 2^{n\|\mathbf{j}\|_{\ell^1}}$ we conclude

$$\begin{aligned}\operatorname{card}\big(\{\varphi_{\mathbf{j},\mathbf{k}}:\mathbf{k}\in\Delta_{\mathbf{j}},\,\|\mathbf{j}\|_{\ell^1}\leq J\}\big)&\sim\sum_{\ell=0}^{J}\sum_{\|\mathbf{j}\|_{\ell^1}=\ell}2^{n\|\mathbf{j}\|_{\ell^1}}\sim\sum_{\ell=0}^{J}2^{n\ell}\log^{m-1}(2^{n\ell})\\&\lesssim\log^{m-1}(2^{nJ})\sum_{\ell=0}^{J}2^{n\ell}\sim 2^{nJ}\log^{m-1}(2^{nJ}).\end{aligned}$$

This finishes the proof due to $N_J\sim 2^{nJ}$. □

By using Lemma 8 and applying an Aubin-Nitsche duality argument one deduces the following approximation result for the Galerkin discretization of operator equations in $H^{\mathbf{q}}(\Omega^m)$.

Proposition 11. *Let Ω be sufficiently smooth and consider the operator equation*

$$A^{(m)}u=f,\tag{2.15}$$

where

$$A^{(m)}:=A\otimes\ldots\otimes A:\quad H^{\mathbf{q}}(\Omega^m)\to H^{-\mathbf{q}}(\Omega^m)$$

is a linear elliptic operator and $f \in H^{\mathbf{d}-2\mathbf{q}}(\Omega^m)$*. Then, the discrete solution*

$$\widehat{u}_J = \sum_{\|\mathbf{j}\|_{\ell^1} \leq J} \sum_{\mathbf{k} \in \Delta_\mathbf{j}} [\widehat{\mathbf{u}}_J]_{\mathbf{j},\mathbf{k}} \varphi_{\mathbf{j},\mathbf{k}}$$

of the Galerkin approximation $\widehat{\mathbf{A}}_J \widehat{\mathbf{u}}_J = \widehat{\mathbf{f}}_J$*, where*

$$\widehat{\mathbf{A}}_J = [\langle A\varphi_{\mathbf{j},\mathbf{k}}, \varphi_{\mathbf{j}',\mathbf{k}'}\rangle]_{\|\mathbf{j}\|_{\ell^1}, \|\mathbf{j}'\|_{\ell^1} \leq j, \mathbf{k} \in \Delta_\mathbf{j}, \mathbf{k}' \in \Delta'_\mathbf{j}}, \qquad \widehat{\mathbf{f}}_J = [\langle f, \varphi_{\mathbf{j},\mathbf{k}}\rangle]_{\|\mathbf{j}\|_{\ell^1} \leq j, \mathbf{k} \in \Delta_\mathbf{j}}$$

satisfies the error estimate

$$\|u - \widehat{u}_J\|_{H^{\mathbf{s}}(\Omega^m)} \lesssim \begin{cases} 2^{J(s-t)} \|u\|_{H^{\mathbf{t}}(\Omega^m)}, & 2q - d < s \leq q \leq t < d, \\ 2^{J(s-t)} J^{(m-1)/2} \|u\|_{H^{\mathbf{t}}(\Omega^m)}, & 2q - d < s \leq q, \ t = d, \\ 2^{J(s-t)} J^{(m-1)/2} \|u\|_{H^{\mathbf{t}}(\Omega^m)}, & 2q - d = s, \ q \leq t < d, \\ 2^{J(s-t)} J^{m-1} \|u\|_{H^{\mathbf{t}}(\Omega^m)}, & 2q - d = s, \ t = d. \end{cases}$$

Proof. Due to Galerkin orthogonality we find

$$\|u - \widehat{u}_J\|^2_{H^{\mathbf{q}}(\Omega^m)} \lesssim |\langle A(u - \widehat{u}_J), u - \widehat{v}_J\rangle| \lesssim \|u - \widehat{u}_J\|_{H^{\mathbf{q}}(\Omega^m)} \|u - \widehat{v}_J\|_{H^{\mathbf{q}}(\Omega^m)}$$

for all $\widehat{v}_J \in \widehat{V}_{J,J,\ldots,J}$, i.e,

$$\|u - \widehat{u}_J\|_{H^{\mathbf{q}}(\Omega^m)} \lesssim \inf_{\widehat{v}_J \in \widehat{V}_{J,J,\ldots,J}} \|u - \widehat{v}_J\|_{H^{\mathbf{q}}(\Omega^m)}.$$

Consequently, by Lemma 8 we obtain

$$\|u - \widehat{u}_J\|_{H^{\mathbf{q}}(\Omega^m)} \lesssim \begin{cases} 2^{J(q-t)} J^{(m-1)/2} \|u\|_{H^{\mathbf{d}}(\Omega^m)}, & t = d, \\ 2^{J(q-t)} \|u\|_{H^{\mathbf{t}}(\Omega^m)}, & t < d. \end{cases} \tag{2.16}$$

Next, denoting the solution of the adjoint equation by $\psi^g = (A^{(m)})^{-\star} g \in H^{\mathbf{d}}(\Omega^m)$, an Aubin-Nitsche duality argument gives

$$\begin{aligned} \|u - \widehat{u}_J\|_{H^{\mathbf{s}}(\Omega^m)} &= \sup_{\|g\|_{H^{-\mathbf{s}}(\Omega^m)} = 1} \langle u - \widehat{u}_J, g\rangle \\ &= \sup_{\|g\|_{H^{-\mathbf{s}}(\Omega^m)} = 1} \langle A(u - \widehat{u}_J), \psi^g\rangle \\ &= \sup_{\|g\|_{H^{-\mathbf{s}}(\Omega^m)} = 1} \langle A(u - \widehat{u}_J), \psi^g - \widehat{v}_J\rangle \\ &\lesssim \sup_{\|v\|_{H^{-\mathbf{s}}(\Omega^m)} = 1} \|u - \widehat{u}_J\|_{H^{\mathbf{q}}(\Omega^m)} \|\psi^g - \widehat{v}_J\|_{H^{\mathbf{q}}(\Omega^m)}. \end{aligned}$$

By virtue of Lemma 8 we conclude

$$\|\psi^g - \widehat{v}_J\|_{H^{\mathbf{q}}(\Omega^m)} \lesssim \begin{cases} 2^{J(s-q)} J^{(m-1)/2} \|\psi^g\|_{H^{2\mathbf{q}-\mathbf{s}}(\Omega^m)}, & s = 2q - d, \\ 2^{J(s-q)} \|\psi^g\|_{H^{2\mathbf{q}-\mathbf{s}}(\Omega^m)}, & s > 2q - d. \end{cases}$$

Due to $\|g\|_{H^{-\mathbf{s}}(\Omega^m)} \sim \|\psi^g\|_{H^{2\mathbf{q}-\mathbf{s}}(\Omega^m)}$, the combination of this estimate with (2.16) yields the assertion. $\square$

3. Application of tensor product operators

3.1. **Prolongations and restrictions.** Let $A : H^q(\Omega) \to H^{-q}(\Omega)$ be a given linear elliptic partial differential operator of order $2q$, where $0 < q < \gamma$, and $\{\varphi_{j,k} : k \in \Delta_j,\ 0 \le j \le J\}$ a frame for $H^q(\Omega)$ in accordance with the previous section.

The system matrix with respect to the frame $\{\varphi_{j,k} : k \in \Delta_j,\ 0 \le j \le J\}$ has a block structure corresponding to the hierarchy (2.4) of subspaces. If we denote the blocks corresponding to trial and test spaces at scales j and j', respectively, by $\mathbf{A}_{j,j'}$, we have $\mathbf{A} = [\mathbf{A}_{j,j'}]_{0\le j,j'\le J}$, where

$$\mathbf{A}_{j,j'} := \langle A\Phi_{j'}, \Phi_j\rangle = [\langle A\varphi_{j',k'}, \varphi_{j,k}\rangle]_{k\in\Delta_j, k'\in\Delta_{j'}}.$$

Standard finite element tools provide only the system matrices $\mathbf{A}_{j,j'}$ for $j = j'$. However, this information is sufficient when using restrictions and prolongations. We denote the restriction of the function

$$f_j = \sum_{k\in\Delta_j} f_{j,k}\varphi_{j,k} = \Phi_j \mathbf{f}_j \in V_j$$

to the space V_ℓ, $\ell < j$, by I_j^ℓ. The corresponding discrete operator will be denoted by $\mathbf{I}_j^\ell$, that is

$$I_j^\ell f_j = \Phi_\ell \mathbf{I}_j^\ell \mathbf{f}_j \in V_\ell.$$

Conversely, I_ℓ^j resp. $\mathbf{I}_\ell^j$ denotes the prolongation of $f_\ell = \Phi_\ell \mathbf{f}_\ell \in V_\ell$ onto V_j. Both, the application of the restriction $\mathbf{I}_\ell^j$ and the prolongation $\mathbf{I}_j^\ell$, to a vector is of complexity $\mathcal{O}(2^{nj}) = \mathcal{O}(N_j)$.

Invoking restriction and prolongation we obviously have

$$\mathbf{A}_{j,j'} = \begin{cases} \mathbf{I}_{j'}^j \mathbf{A}_{j',j'}, & j \le j', \\ \mathbf{A}_{j,j}\mathbf{I}_{j'}^j, & j > j'. \end{cases} \tag{3.1}$$

Since the operator A is a local operator the system matrices $\mathbf{A}_{j,j}$ have only $\mathcal{O}(1)$ nonzero coefficients per column and row, independently of the level j. Thus, employing (3.1), the matrix-vector multiplication $\mathbf{A}_{j,j'}\mathbf{x}$ can be performed in $\mathcal{O}(2^{n\max\{j,j'\}})$ operations, which is order-optimal.

3.2. **Fast two-factor matrix-vector multiplication.** We first focus on second moments, i.e. the case $m = 2$ in (2.15) and address the case $m > 2$ below.

Consider the numerical solution of the operator equation

$$(A \otimes A)u = f \tag{3.2}$$

with the two-factor tensor product operator $A \otimes A : H^{\mathbf{q}}(\Omega) \to H^{-\mathbf{q}}(\Omega)$. Equation (3.2) can efficiently be discretized in the sparse tensor product spaces $\widehat{V}_J$, see e.g. [20, 28, 32]. But at first glance it is not clear how to solve the corresponding system

$$\widehat{\mathbf{A}}_J \widehat{\mathbf{x}}_J = \widehat{\mathbf{f}}_J, \quad \widehat{\mathbf{A}}_J = [\mathbf{A_j} \otimes \mathbf{A_{j'}}]_{0 \le \|\mathbf{j}\|_{\ell^1}, \|\mathbf{j'}\|_{\ell^1} \le J}, \quad \widehat{\mathbf{f}}_J = [\langle f, \varphi_{\mathbf{j},\mathbf{k}} \rangle]_{\|\mathbf{j}\|_{\ell^1} \le J, \mathbf{k} \in \Delta_{\mathbf{j}}} \tag{3.3}$$

where $\mathbf{j} = (j_1, j_2)$ and $\mathbf{j}' = (j_1', j_2')$ in an efficient way since the system matrix is not sparse.

Even though the system matrix $\widehat{\mathbf{A}}_J$ has a possibly large kernel it is well conditioned in terms of

$$\min\big\{\lambda \in \sigma\big(\widehat{\mathbf{A}}_J^\star \widehat{\mathbf{A}}_J\big) : \lambda > 0\big\} \sim \max\{\sigma\big(\widehat{\mathbf{A}}_J^\star \widehat{\mathbf{A}}_J\big)\big\} \sim 1$$

according to Lemma 2. Thus, iterative solvers like Krylov subspace methods, for example the conjugate gradient method or GMRES, will converge with a convergence rate that is independent of the discretization level J (e.g. [10, 16, 17, 25]).

We fix some notation. For a matrix

$$\mathbf{A} = \Big[\ \mathbf{a}_1 \ \Big|\ \mathbf{a}_2 \ \Big|\ \cdots \ \Big|\ \mathbf{a}_n \ \Big] \in \mathbb{R}^{m \times n}, \qquad \mathbf{a}_i \in \mathbb{R}^m$$

we define $\mathrm{vec}(\mathbf{A})$ as the column vector

$$\mathrm{vec}(\mathbf{A}) = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_n \end{bmatrix} \in \mathbb{R}^{m \cdot n}.$$

Then, for given matrices $\mathbf{B} \in \mathbb{R}^{k \times \ell}$, $\mathbf{X} \in \mathbb{R}^{\ell \times m}$, $\mathbf{A} \in \mathbb{R}^{n \times m}$, and $\mathbf{Y} \in \mathbb{R}^{k \times n}$, there holds the identity

$$\mathrm{vec}(\mathbf{Y}) = (\mathbf{A} \otimes \mathbf{B})\, \mathrm{vec}(\mathbf{X}). \quad \Longleftrightarrow \mathbf{B}\mathbf{X}\mathbf{A}^\top = \mathbf{Y} \tag{3.4}$$

We will use the equivalence (3.4) to develop a fast matrix-vector multiplication. To this end, we assume that the vector $\widehat{\mathbf{x}}_J = [\widehat{\mathbf{x}}_{j_1,j_2}]_{0 \le j_1 + j_2 \le J}$ that consists of the coefficients associated with the sparse tensor product basis $\{\varphi_{\mathbf{j},\mathbf{k}} : \mathbf{k} \in \Delta_{\mathbf{j}},\ \|\mathbf{j}\|_{\ell^1} \le J\}$ is blockwise stored in matrix form, i.e. $\widehat{\mathbf{x}}_{j_1,j_2} \in \mathbb{R}^{j_1 \times j_2}$. Then, for the matrix-vector multiplication (3.4) we compute products of the form

$$\mathrm{vec}(\mathbf{z}) = (\mathbf{A}_{j_1,j_1'} \otimes \mathbf{A}_{j_2,j_2'})\, \mathrm{vec}(\widehat{\mathbf{x}}_{j_1',j_2'}).$$

Using (3.4), this means that

$$\mathbf{z} = \mathbf{A}_{j_2,j_2'} \widehat{\mathbf{x}}_{j_1',j_2'}^\top \mathbf{A}_{j_1,j_1'}^\top \qquad j_1 + j_2 \le J, \quad j_1' + j_2' \le J.$$

To minimize the $\log N$-powers in the complexity bound, we have to be careful with the order in which we perform the multiplications. We perform the matrix-vector multiplication according to

$$\mathbf{z} = \begin{cases} \mathbf{A}_{j_2,j_2'}(\widehat{\mathbf{x}}_{j_1',j_2'}^\top \mathbf{A}_{j_1,j_1'}^\top), & j_1 + j_2' \le j_1' + j_2, \\ (\mathbf{A}_{j_2,j_2'}\widehat{\mathbf{x}}_{j_1',j_2'}^\top)\mathbf{A}_{j_1,j_1'}^\top, & j_1 + j_2' > j_1' + j_2. \end{cases}$$

Observing (3.1), in case of $j_1 + j_2' \le j_1' + j_2$ we compute

$$\mathbf{y} := \mathbf{A}_{j_1,j_1'}\widehat{\mathbf{x}}_{j_1',j_2'} = \begin{cases} \mathbf{I}_{j_1'}^{j_1}\mathbf{A}_{j_1',j_1'}\widehat{\mathbf{x}}_{j_1',j_2'}, & j_1 \le j_1', \\ \mathbf{A}_{j_1,j_1}\mathbf{I}_{j_1'}^{j_1}\widehat{\mathbf{x}}_{j_1',j_2'}, & j_1 > j_1', \end{cases} \tag{3.5}$$

$$\mathbf{z} := \mathbf{A}_{j_2,j_2'}\mathbf{y}^\top = \begin{cases} \mathbf{I}_{j_2'}^{j_2}\mathbf{A}_{j_2',j_2'}\mathbf{y}^\top, & j_2 \le j_2', \\ \mathbf{A}_{j_2,j_2}\mathbf{I}_{j_2'}^{j_2}\mathbf{y}^\top, & j_2 > j_2'. \end{cases} \tag{3.6}$$

Therein, one needs $\mathcal{O}(2^{\max\{j_1,j_1'\}n+j_2'n})$ operations to get $\mathbf{y}$ by (3.5) and additional $\mathcal{O}(2^{\max\{j_2,j_2'\}n+j_1n})$ operations to derive $\mathbf{z}$ by (3.6). We abbreviate $j := j_1 + j_2$ and and $j' = j_1' + j_2'$. Then, we see that the complexity is bounded by $\mathcal{O}(2^{\max\{j,j',j_1+j_2'\}n})$. Since $j_1 + j_2' \le j_1' + j_2$ implies

$$j_1 + j_2' \le j' - j_2' + j - j_1$$

we conclude $j_1 + j_2' \le \max\{j, j'\}$. Consequently, the computation of $\mathbf{z}$ is of complexity $\mathcal{O}(2^{\max\{j_1+j_2,j_1'+j_2'\}n})$ if $j_1 + j_2' \le j_1' + j_2$.

In case of $j_1 + j_2' > j_1' + j_2$ we compute

$$\mathbf{y} := \mathbf{A}_{j_2,j_2'}\widehat{\mathbf{x}}_{j_1',j_2'}^\top = \begin{cases} \mathbf{I}_{j_2'}^{j_2}\mathbf{A}_{j_2',j_2'}\widehat{\mathbf{x}}_{j_1',j_2'}^\top, & j_2 \le j_2', \\ \mathbf{A}_{j_2,j_2}\mathbf{I}_{j_2'}^{j_2}\widehat{\mathbf{x}}_{j_1',j_2'}^\top, & j_2 > j_2', \end{cases} \tag{3.7}$$

$$\mathbf{z}^\top := \mathbf{A}_{j_1,j_1'}\mathbf{y}^\top = \begin{cases} \mathbf{I}_{j_1'}^{j_1}\mathbf{A}_{j_1',j_1'}\mathbf{y}^\top, & j_1 \le j_1', \\ \mathbf{A}_{j_1,j_1}\mathbf{I}_{j_1'}^{j_1}\mathbf{y}^\top, & j_1 > j_1', \end{cases}. \tag{3.8}$$

Using the same argument as above we find that $\mathbf{z}$ is also computed in complexity $\mathcal{O}(2^{\max\{j_1+j_2,j_1'+j_2'\}n})$ if $j_1 + j_2' > j_1' + j_2$.

With the above preparations at hand we can formulate the following algorithm which performs the matrix-vector multiplication $\widehat{\mathbf{y}}_J = \widehat{\mathbf{A}}_J\widehat{\mathbf{x}}_J$:

Algorithm 12 (Sparse tensor product matrix-vector multiplication)**.**

input: *finite element matrices* $\mathbf{A}_{j,j}$ $(0 \leq j \leq J)$

vector $\widehat{\mathbf{x}}_J = [\widehat{\mathbf{x}}_{j,j'}]_{0\leq j+j'\leq J}$

output: *blockwise stored vector* $\widehat{\mathbf{y}} = [\widehat{\mathbf{y}}_{j,j'}]_{0\leq j+j'\leq J}$

for all $0 \leq j_1 + j_2 \leq J$ **do begin**

  initialize $\widehat{\mathbf{y}}_{j_1,j_2} := \mathbf{0}$

  for all $0 \leq j_1' + j_2' \leq J$ **do begin**

    if $(j_1 + j_2' \leq j_1' + j_2)$ **then**

      update $\widehat{\mathbf{y}}_{j_1,j_2} := \widehat{\mathbf{y}}_{j_1,j_2} + \mathbf{z}$ **with z from** (3.5), (3.6)

    else

      update $\widehat{\mathbf{y}}_{j_1,j_2} := \widehat{\mathbf{y}}_{j_1,j_2} + \mathbf{z}$ **with z from** (3.7), (3.8)

    end

  end

end

Theorem 13. *Algorithm 12 computes the matrix-vector product* $\widehat{\mathbf{y}}_J = \widehat{\mathbf{A}}_J\widehat{\mathbf{x}}_J$ *for the second moments in* $\mathcal{O}(N_J \log^3 N_J)$ *operations.*

Proof. As we have seen both block matrix-vector products (3.5), (3.6) and (3.7),(3.8) have complexity $\mathcal{O}(2^{\max\{j_1+j_2,j_1'+j_2'\}n})$. We abbreviate $j = j_1 + j_2$, $j' = j_1' + j_2'$ and estimate the work $\mathcal{W}(N_J)$ $(N_J \sim 2^{Jn})$ required by Algorithm 12:

$$\begin{aligned}
\mathcal{W}(N_J) &\lesssim \sum_{j_1+j_2\leq J}\ \sum_{j_1'+j_2'\leq J} 2^{\max\{j_1+j_2,j_1'+j_2'\}n} \\
&= \sum_{j=0}^{J}\sum_{j'=0}^{J}(j+1)(j'+1)2^{\max\{j,j'\}n} \\
&= \sum_{j=0}^{J}(j+1)\left\{\sum_{j'=0}^{j-1}(j'+1)2^{jn} + \sum_{j'=j}^{J}(j'+1)2^{j'n}\right\} \\
&\lesssim \sum_{j=0}^{J}(j+1)\{2^{jn}j^2 + J2^{Jn}\} \\
&\lesssim J^3 2^{Jn}.
\end{aligned}$$

This is the desired assertion due to $J \sim \log N_J$. □

Together with the well-posedness of the Galerkin system $\widehat{\mathbf{A}}_J\widehat{\mathbf{u}}_J = \widehat{\mathbf{f}}_J$ we can realize an algorithm which solves the given operator equations in essentially the complexity required for discretizing the domain Ω.

3.3. **Moments $\mathcal{M}^m u$, $m \geq 3$.** If moments of order $m \geq 3$ are to be computed, we can proceed in a similar way. Instead of the tensor pde (3.2), we now have to solve the m-factor tensor pde

$$(A \otimes \cdots \otimes A)\mathcal{M}^m u = \mathcal{M}^m f. \tag{3.9}$$

As in the case of second moments, we obtain the matrix equation (3.3), where, however, now $\mathbf{j} = (j_1, \dots, j_m)$ and likewise for $\mathbf{j}'$.

We shall obtain a matrix-vector multiplication of log-linear complexity by recursive reduction of the case $m > 2$ to the previous one. Specifically, writing

$$A \otimes \cdots \otimes A = A \otimes (\underbrace{A \otimes \cdots \otimes A}_{(m-1)\text{-times}}) = A \otimes B,$$

we may use Algorithm 12 with

$$B = \underbrace{A \otimes \cdots \otimes A}_{(m-1)\text{-times}},$$

and obtain from Theorem 13 a matrix-vector multiplication in log-linear complexity, *provided that a matrix-vector multiplication of the matrix B with a vector is available.* Recursively applying Algorithm 12 $(m-2)$-times to the matrix-vector product $A \otimes B$, we obtain a corresponding matrix-vector multiplication for m-fold tensor operators.

Backward induction of the work bound in Theorem 13 over $m, m-1, \dots, 3, 2$ yields the log-linear complexity bound for the work of this recursion.

4. Numerical Examples

4.1. **First example.** In our first example we consider $\Omega = [0,1]$ and $A = -\Delta$. In other words, we are looking for the solution $u \in H_0^1([0,1]) \otimes H_0^1([0,1])$ of the equation

$$\frac{\partial^4 u(x,y)}{\partial x^2 \partial y^2} = f(x,y), \qquad (x,y) \in [0,1]^2$$

where $f \in (H_0^1([0,1]) \otimes H_0^1([0,1]))'$ is a given load. We choose

$$f(x,y) = 9\pi^4 \sin(\pi x)\sin(3\pi y)$$

which yields the solution

$$u(x,y) = \sin(\pi x)\sin(3\pi y).$$

We compute the solution numerically with respect to the sparse tensor product frame based on tensor products of piecewise linear hat functions on the interval,

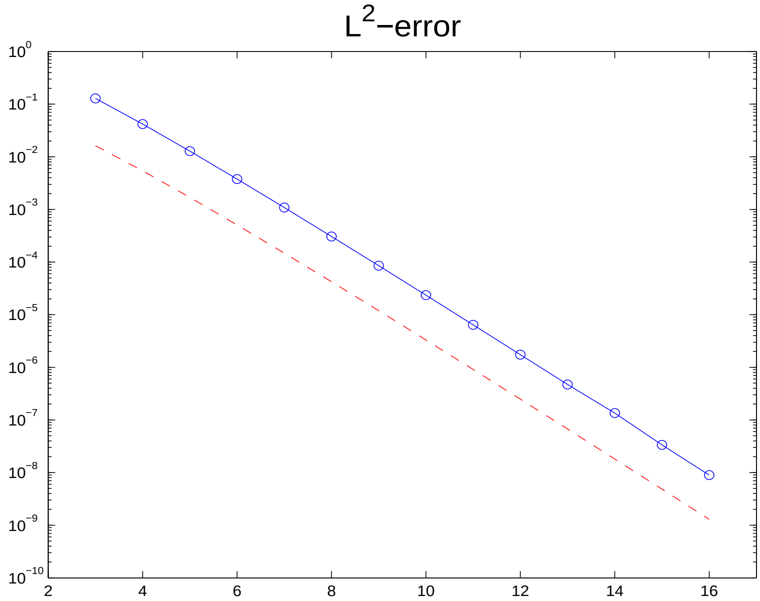

FIGURE 1. Discretization error the sparse tensor product approximation.

Level	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Iterations	8	13	16	20	23	26	28	30	32	33	35	36	37	37

TABLE 1. Number of iterations of the conjugate gradient method.

that is

$$\varphi_{j,k}(x) := \begin{cases} 2^{-j/2}(2^j x - k + 1), & 2^{-j}(k-1) \le x < 2^{-j}k, \\ 2^{-j/2}(k + 1 - 2^j x), & 2^{-j}k \le x < 2^{-j}(k+1), \\ 0, & \text{otherwise.} \end{cases}$$

Notice that $\|\varphi_{j,k}\|_{H_0^1([0,1])} = 8/3$ independently of j. We monitored the L^2-error of the approximate sparse tensor product solution in Figure 1. The dashed line corresponds to the expected rate of convergence $\mathcal{O}(h_j|\log h_j|)$ $(h_j = 2^{-j})$, see Proposition 11.

Figure 2 shows the over-all computing time, i.e., the time consumed to assemble the load vector and the finite element system matrices, and to solve the linear system of equations iteratively by the conjugate gradient method up to 10^{-6} accuracy, using Algorithm 12. The dashed line indicates the asymptotics $\mathcal{O}(N_j \log^3 N_j)$ $(N_j \sim 2^j)$, see Theorem 13. The number of iteration steps required by the conjugate gradient method is listed in Table 1. In fact, the required number of iterations is bounded by some constant.

4.2. **Second example.** In our second example we do consider the polygonal domain $\Omega \in \mathbb{R}^2$ shown in Figure 3 and again the Laplace operator $A = -\Delta$ with

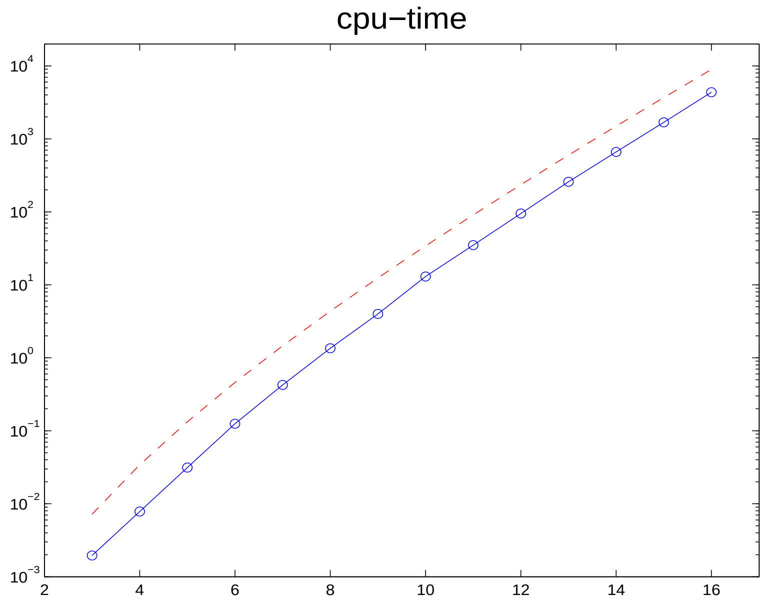

FIGURE 2. Over-all computing times.

homogeneous boundary conditions:

$$\begin{aligned}(\Delta_{\mathbf{x}} \otimes \Delta_{\mathbf{y}})u(\mathbf{x},\mathbf{y}) &= f(\mathbf{x},\mathbf{y}), \quad && (\mathbf{x},\mathbf{y}) \in \Omega \times \Omega,\\ u(\mathbf{x},\mathbf{y}) &= 0, && (\mathbf{x},\mathbf{y}) \in \partial\Omega \times \partial\Omega.\end{aligned}$$

The multigrid hierarchy is defined via standard uniform subdivision of a triangle into four sons. The coarse triangulation (level $j = 0$) and the triangulation after three subdivision steps (level $j = 3$) are depicted in Figure 3. We use canonical piecewise linear finite elements (they are $H^1(\Omega)$-normalized), defined on the given sequence of meshes.

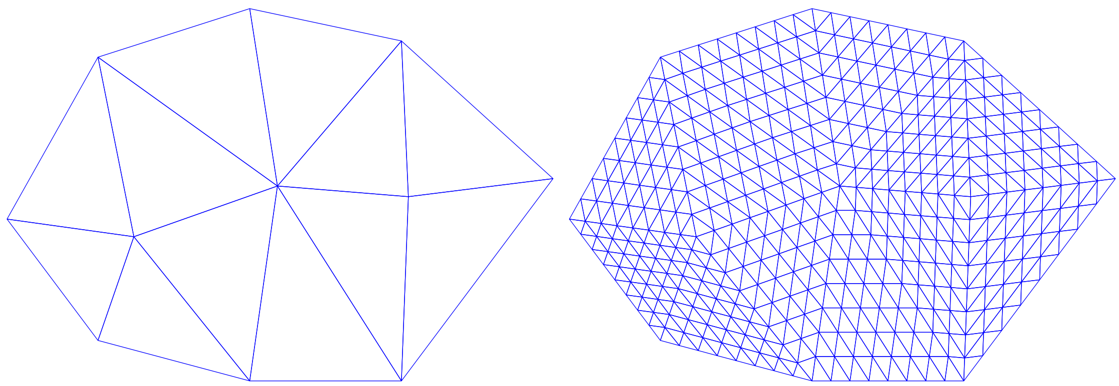

FIGURE 3. The domain Ω with coarse grid triangulation (left) and the mesh on level 3 (right).

We choose $f = 1$ and compute the approximate solution by the sparse tensor product ansatz. Since we do not know the problem's solution analytically, we compare the sparse tensor product approximation with the full tensor product approximation which is computable since $u(\mathbf{x},\mathbf{y}) = v(\mathbf{x}) \cdot v(\mathbf{y})$ with $v = A^{-1}1$. Notice that the ℓ^2-difference of the discrete coefficients corresponds to the norm in the energy space $H^{1,1}(\Omega \times \Omega)$ since the frame is for this space.

level	unknowns	H^1-error		iterations	cpu-time
1	847	$1.2 \cdot 10^{-1}$	(—)	10	0 s
2	4422	$9.4 \cdot 10^{-2}$	(1.3)	14	0 s
3	21054	$6.4 \cdot 10^{-2}$	(1.5)	18	1 s
4	96567	$4.2 \cdot 10^{-2}$	(1.5)	21	13 s
5	434481	$2.7 \cdot 10^{-2}$	(1.6)	24	110 s
6	1930540	$1.6 \cdot 10^{-2}$	(1.6)	26	896 s
7	8.5 Mio	$9.4 \cdot 10^{-3}$	(1.7)	29	2 h
8	37 Mio	$5.4 \cdot 10^{-3}$	(1.7)	31	15 h

TABLE 2. Number of iterations of the conjugate gradient method.

In Table 2 we tabulate the numerical results. Notice that on level 8 we have 400 000 finite elements per domain Ω. This corresponds to $16 \cdot 10^{10}$ unknowns in the full tensor product space. We still observe the logarithmical factors in the growth of the number of unkowns (and thus in the cpu-times) as well as in the rate of convergence. Nevertheless, the present approach is obviously suitable to treat such large problems.

References

[1] R. Balder. Adaptive Verfahren für elliptische und parabolische Differentialgleichungen auf dünnen Gittern. PhD thesis, Institut für Informatik, TU München, 1994.

[2] R. Balder and C. Zenger. The solution of multidimensional real Helmholtz equations on sparse grids. *SIAM J. Sci. Comput.* **17** (1996) 631–646.

[3] M. Bebendorf and S. Rjasanow. Adaptive low-rank approximation of collocation matrices. *Computing* **70** (2003) 1–24.

[4] J. Bramble, J. Pasciak, and J. Xu. Parallel multilevel preconditioners. *Mathematics of Computation* **55** (1990) 1–22.

[5] H.J. Bungartz. A multigrid algorithm for higher order finite elements on sparse grids. *ETNA, Electron. Trans. Numer. Anal.* **6** (1997) 63–77.

[6] H.J. Bungartz and T. Dornseifer. Sparse grids: Recent developments for elliptic partial differential equations. Multigrid methods V. Proceedings of the 5th European multigrid conference, held in Stuttgart, Germany, October 1–4, 1996 (W. Hackbusch et al., eds.). Berlin: Springer. Lect. Notes Comput. Sci. Eng. 3, 45–70 (1998).

[7] H.J. Bungartz asnd M. Griebel. Sparse Grids. *Acta Numerica* **13** (2004) 147–269.

[8] O. Christiansen. *An Introduction to Frames and Riesz Bases.* Applied and Numerical Harmonic Analysis. Birkhäuser Boston, Inc., Boston, 2002.

[9] S. Dahlke, M. Fornasier, and T. Raasch. Adaptive Frame Methods for Elliptic Operator Equations *Adv. Comput. Math.* **27** (2007) 27–63.

[10] S. Dahlke, M. Fornasier, T. Raasch, R. Stevenson, and M. Werner. Adaptive Frame Methods for Elliptic Operator Equations: The Steepest Descent Approach. *IMA J. Numer. Math.*, to appear.

[11] W. Dahmen. Wavelet and multiscale methods for operator equations. *Acta Numerica* **6** (1997) 55–228.

[12] I. Daubechies. *Ten lectures on wavelets.* CBMS-NSF Regional Conference Series in Applied Mathematics, 61. SIAM, Philadephia, 1992.

[13] W. Dahmen, H. Harbrecht, and R. Schneider. Compression techniques for boundary integral equations — optimal complexity estimates. *SIAM J. Numer. Anal.* **43** (2006) 2251–2271.

[14] K. Doppel and R. Hochmuth. On the regularity of solutions of a homogeneous Dirichlet problem for a nonhypoelliptic linear partial differential operator. *Ann. Acad. Sci. Fenn. A.I. Math.* **16** (1991) 183–198.

[15] L. Greengard and V. Rokhlin. A fast algorithm for particle simulation. *J. Comput. Phys.* **73** (1987) 325–348.

[16] M. Griebel. Multilevel algorithms considered as iterative methods on semidefinite systems. *SIAM J. Sci. Comput.* **15** (1994), 547–565.

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

[18] M. Griebel and P. Oswald. On additive Schwarz preconditioners for sparse grid discretizations. *Numer. Math.* **66** (1994) 449–463.

[19] M. Griebel and P. Oswald. Tensor product type subspace splittings and multilevel iterative methods for anisotropic problems. *Adv. Comput. Math.* **4** (1995) 171–206.

[20] H. Harbrecht, R. Schneider, and C. Schwab. Sparse Second Moment Analysis for Elliptic Problems in Stochastic Domains. *Preprint No. 304*, Berichtsreihe des SFB 611, Universität Bonn, 2006. submitted to Numer. Math.

[21] W. Hackbusch. A sparse matrix arithmetic based on $\mathcal{H}$-matrices. Part I: Introduction to $\mathcal{H}$-matrices. Computing **64** (1999) 89–108.

[22] R. Hochmuth. An inhomogeneous Dirichlet problem for a non-hypoelliptic linear partial differential operator. *Ann. Acad. Sci. Fenn. Math.* **21** (1996) 179–187.

[23] R. Hochmuth. Wavelets in Numerical Analysis and Restricted Nonlinear Approximation. Habilitation-Thesis (1999).

[24] R. Hochmuth. Nonlinear anisotropic boundary value problems - regularity results and multiscale discretizations. *Nonlinear Analysis: Series A, Theory and Methods* **46** (2001) 1–18.

[25] E.F. Kaasschieter. Preconditioned conjugate gradients for solving singular systems. J. of Comput. Appl. Math. **24** (1988) 265–275.

[26] P. Oswald. *Multilevel finite element approximation. Theory and applications.* Teubner Skripten zur Numerik. B.G. Teubner, Stuttgart, 1994.

[27] P. Oswald. On function spaces related to Finite Element Approximation Theory. Z. Anal. Anwendungen **9** (1990) 43–64.

[28] T. von Petersdorff and C. Schwab. Sparse wavelet methods for operator equations with stochastic data. Applications of Mathematics **51**(2) (2006) 145–180.

[29] A. Pfaffinger. Funktionale Beschreibung und Parallelisierung von Algorithmen auf dünnen Gittern. PhD thesis, Institut für Informatik, TU München, 1997.

[30] C. Schwab and R. Stevenson. Adaptive Wavelet Algorithms for Elliptic PDEs on Product Domains. Report 2006-10, Seminar for Applied Mathematics, ETH Zürich http://www.sam.math.ethz.ch/reports/2006/10 (to appear in Mathematics of Computation, 2007)

[31] C. Schwab and R. Todor. Sparse finite elements for elliptic problems with stochastic loading. *Numer. Math.* **95** (2003) 707–734.

[32] C. Schwab and R. Todor. Sparse finite elements for stochastic elliptic problems-higher order moments Computing **71** (2003) 43–63.

[33] R. Stevenson. Adaptive solution of operator quations using wavelet frames. *SIAM J. Numer. Anal.* **41** (2003) 1074–1100.

[34] C. Zenger. Sparse grids. Parallel algorithms for partial differential equations, Proc. 6th GAMM- Semin., Kiel/Ger. 1990, Notes Numer. Fluid Mech. 31, 241–251 (1991).