Institut für Numerische Simulation
Rheinische Friedrich-Wilhelms-Universität Bonn
Wegelerstraße 6 • 53115 Bonn • Germany
phone +49 228 73-3427 • fax +49 228 73-7527
www.ins.uni-bonn.de

M. Griebel, A. Hullmann

On a Multilevel Preconditioner and its Condition Numbers for the Discretized Laplacian on Full and Sparse Grids in Higher Dimensions

INS Preprint No. 1301

January 2013

On a Multilevel Preconditioner and its Condition Numbers for the Discretized Laplacian on Full and Sparse Grids in Higher Dimensions

M. Griebel and A. Hullmann

Abstract We first discretize the d-dimensional Laplacian in $(0,1)^d$ for varying d on a full uniform grid and build a new preconditioner that is based on a multilevel generating system. We show that the resulting condition number is bounded by a constant that is independent of both, the level of discretization J and the dimension d. Then, we consider so-called sparse grid spaces, which offer nearly the same accuracy with far less degrees of freedom for function classes that involve bounded mixed derivatives. We introduce an analogous multilevel preconditioner and show that it possesses condition numbers which are at least as good as these of the full grid case. In fact, for sparse grids we even observe falling condition numbers with rising dimension in our numerical experiments. Furthermore, we discuss the cost of the algorithmic implementations. It is linear in the degrees of freedom of the respective multilevel generating system. For completeness, we also consider the case of a sparse grid discretization using prewavelets and compare its properties to those obtained with the generating system approach.

1 Introduction

In this paper, we deal with the preconditioning of finite element system matrices that stem from elliptic partial differential equations (PDEs) of second order. Here, we are especially interested in the higher-dimensional case. For example, high-dimensional Poisson problems and high-dimensional convection diffusion equations result from diffusion approximation techniques or the Fokker–Planck approach. Examples are the description of queueing networks [Mit97, SCDD02], reaction mechanisms in molecular biology [Sjö07, SLE09], or various models for the pricing of financial derivatives [Kwo08, Rei04]. Furthermore, homogenization with multiple scales [All92, CDG02, Mat02, HS05] as well as stochastic uncertainty quantifica-

Institute for Numerical Simulation, University of Bonn, Wegelerstr. 6, 53115 Bonn, Germany
e-mail: {griebel, hullmann}@ins.uni-bonn.de

tion [HSS08, BNT10, JR08, NTW08b, NTW08a, BNTT11, MK10, CDS11] result in high-dimensional PDEs. Next, we find quite high-dimensional problems in quantum mechanics and particle physics. There, the dimensionality of the Schrödinger equation [Mes65] grows with the number of considered electrons and nuclei. Then, problems in statistical mechanics lead to the Liouville equation or the Langevin equation and related phase space models where the dimension depends on the number of particles [Bal97]. Furthermore, reinforcement learning and stochastic optimal control in continuous time give rise to the Hamilton–Jacobi–Bellman equation in high dimensions [SB98, Mun00, BGGK12]. Finally data mining problems involve differential operators as smoothing or regularization terms (priors) whose dimension grows with the number of features of the data [GJP95, GGT01, SS01, Heg03, Gar04].

We want to derive multilevel preconditioners with condition numbers that are bounded independently of both, the discretization level J *and* the dimension d. Furthermore, they should possess linear cost complexity with respect to the degrees of freedom.

We will focus on the model problem of the d-dimensional Laplacian, which has been intensively analyzed in numerical analysis, albeit mostly for fixed dimension d. To this end, we first consider the simple case of a discretization based on a uniform grid using, e.g., piecewise d-linear finite elements. The solution of the resulting system of linear equations is computed iteratively. This involves the cost of a matrix-vector multiplication times the number of iterations needed to achieve a given accuracy. Here, a sparse system matrix can usually be applied with a number of floating point operations that is linear in the number of degrees of freedom. An optimal iteration count which is independent of the number of degrees of freedom is typically achieved by multiplicative multigrid methods [Yse93, BL11, Hac85, Gri94b], the additive BPX preconditioner [BPX90, Osw92, Osw94] or wavelet-based methods. But even if the overall additive or multiplicative preconditioned matrix-vector product is linear in the number of degrees of freedom and the number of iterations is independent of the mesh width, the involved order constants are in general still dependent on the dimension d, which can be an issue in the higher-dimensional case.

Furthermore, the number of degrees of freedom itself is subject to the curse of dimension [Bel61]. One remedy is the use of so called sparse-grid discretizations. To this end, regular sparse grids, energy sparse grids [Bun92a, BG99] and general sparse grids [GK00, Kna00, GK09, Ham09] have been employed with good success. Furthermore, space- and dimension-adaptive extensions exist [GG03, Feu10]. However, the condition number of the resulting system and the cost of a matrix-vector multiplication are now more difficult to reduce than in the regular full-grid case. For example, already for a straightforward regular sparse grid discretization, cf. [GO94], a simple diagonal scaling similar to the case of the BPX-preconditioner does *not* result in asymptotically bounded condition numbers in dimensions $d \geq 3$. Here, more complicated basis functions like prewavelets offer a solution [GO95]. Furthermore, the system matrix is not inherently sparse, and a dimension-recursive algorithm based on the so-called unidirectional principle [BZ96, Bun92b] is needed to perform the matrix-vector-multiplication in linear time.

In this paper, we present a new additive preconditioner that is based on the multilevel idea and relies on isotropic and anisotropic subspaces. We show for the full grid case that the resulting condition number is bounded independently of the level J of the discretization and that it is also independent of the dimension d. The cost complexity is linear in the number of degrees of freedom of the enlarged generating system with a constant that grows at most polynomially in the dimension. However, it needs to be mentioned that the enlarged generating system has a factor of about 2^d more degrees of freedom than there are on the finest mesh. Our preconditioner is applicable to sparse grid discretizations as well, and the resulting condition number is now also bounded independently of J for $d \geq 3$. Furthermore, it is bounded independently of d and we even observe a falling condition number with rising dimension d. The new preconditioner can also be applied to prewavelet discretizations and then produces exactly the same condition numbers.

In Sect. 2, we introduce a multilevel discretization, and we present a norm equivalence with dimension-independent constants. Then, in Sect. 3, we introduce the full grid preconditioner with dimension-independent condition numbers for our enlarged generating system and discuss its costs. The new approach is extended to sparse grids in Sect. 4. In Sect. 5 we show that the same results can be obtained for prewavelet discretizations as well. In Sect. 6 we give numerical results that support our theory. In fact, for sparse grids, we even observe falling condition numbers with rising dimension d. Final remarks in Sect. 7 conclude the paper.

2 Discretization

We denote the unit interval by $\Omega = (0,1)$ and its d-fold tensor product by Ω^d. The Poisson problem on Ω^d for a given right-hand side $f : \Omega^d \to \mathbb{R}$ with $\Gamma = \partial\Omega^d$ and homogeneous boundary conditions reads as

$$\begin{aligned} -\Delta u &= f \qquad \text{on} \quad \Omega^d\,, \\ u &= 0 \qquad \text{on} \quad \Gamma\,. \end{aligned} \tag{1}$$

2.1 Discretization by an Isotropic Full-Grid

Our aim is to discretize problem (1) by piecewise polynomials on a uniform grid and to precondition the resulting system of linear equations optimally not only with respect to the number of degrees of freedom, but also with respect to the dimension d. As usual, we define the bilinear form $a : H^1(\Omega^d) \times H^1(\Omega^d) \to \mathbb{R}$ as

$$a(u,v) = \int_{\Omega^d} \nabla u \cdot \nabla v \,\mathrm{d}\mathbf{x}$$

and the right-hand side $b \in H^1(\Omega^d)^*$ as

$$b(v) = \int_{\Omega^d} f\, v\, \mathrm{d}\mathbf{x}\, .$$

The weak formulation then reads: Find a solution $u \in H_0^1(\Omega^d)$ that satisfies

$$a(u,v) = b(v) \quad \text{for all } \ v \in H_0^1(\Omega^d)\, . \tag{2}$$

We discretize $H_0^1(\Omega^d)$ by the d-fold tensor product of one-dimensional function spaces. To this end, we first consider a one-dimensional multiresolution scale of subspaces, i.e

$$V_1 \subset V_2 \subset V_3 \subset \dots\, , \tag{3}$$

for which $\overline{V}^{\|\cdot\|_1} = H_0^1([0,1])$ holds with $V = \cup_{l=1}^{\infty} V_l$. Here, we assume

$$V_l = \mathrm{span}\{\phi_{l,i} : 1 \le i \le n_l\} \tag{4}$$

with $n_l = \mathcal{O}(2^l)$ locally supported basis functions $\phi_{l,i}, 1 \le i \le n_l$, on level l. We define the d-dimensional tensor product space

$$V^d = V \otimes \cdots \otimes V$$

and the spaces

$$V_l^d = V_l \otimes \cdots \otimes V_l\, , \tag{5}$$

which are spanned by the functions

$$\phi_{l,\mathbf{i}} = \phi_{l,i_1} \cdots \phi_{l,i_d} \tag{6}$$

for $\mathbf{i} = (i_1, \dots, i_d) \in \mathbb{N}^d$ with $1 \le i_p \le n_l, p = 1, \dots, d$.

On level J, the weak problem (2) for V_J^d then leads to the system

$$\mathbf{A}_{d,J}\mathbf{x}_{d,J} = \mathbf{b}_{d,J} \tag{7}$$

of $N_{d,J} := (n_J)^d$ linear equations with

$$\mathbf{A}_{d,J} \in \mathbb{R}^{N_{d,J} \times N_{d,J}}, (\mathbf{A}_{d,J})_{\mathbf{i},\mathbf{j}} = a(\phi_{J,\mathbf{i}}, \phi_{J,\mathbf{j}})$$

and

$$\mathbf{x}_{d,J}, \mathbf{b}_{d,J} \in \mathbb{R}^{N_{d,J}}\ , (\mathbf{b}_{d,J})_{\mathbf{i}} = (f, \phi_{J,\mathbf{i}})_{L^2(\Omega^d)}$$

again for $\mathbf{i},\mathbf{j} \in \mathbb{N}^d$ with $1 \le i_p, j_p \le n_J, p = 1, \dots, d$. Note that, with a lexicographic ordering of the degrees of freedom, the system matrix can be expressed as a sum of Kronecker product matrices, i.e.

$$\mathbf{A}_{d,J} = \mathbf{A}_{1,J} \otimes \mathbf{M}_{d-1,J} + \sum_{p=2}^{d-1} \mathbf{M}_{p-1,J} \otimes \mathbf{A}_{1,J} \otimes \mathbf{M}_{d-p,J} + \mathbf{M}_{d-1,J} \otimes \mathbf{A}_{1,J}\, , \tag{8}$$

where $\mathbf{A}_{1,J} \in \mathbb{R}^{n_J \times n_J}$ is the stiffness matrix of the one-dimensional problem

$$(\mathbf{A}_{1,J})_{ij} = \Big(\frac{\partial \phi_{J,i}}{\partial x}, \frac{\partial \phi_{J,j}}{\partial x}\Big)_{L^2(\Omega)} \quad \text{for} \quad 1 \le i, j \le n_J \ ,$$

and $\mathbf{M}_{p,J} \in \mathbb{R}^{(n_J)^p \times (n_J)^p}$ also has Kronecker product structure

$$\mathbf{M}_{p,J} = \bigotimes_{q=1}^{p} \mathbf{M}_{1,J}$$

with $\mathbf{M}_{1,J} \in \mathbb{R}^{n_J \times n_J}$ and

$$(\mathbf{M}_{1,J})_{ij} = (\phi_{J,i}, \phi_{J,j})_{L^2(\Omega)} \quad \text{for} \quad 1 \le i, j \le n_J \ . \tag{9}$$

2.2 The Multilevel Approach

The system matrix $\mathbf{A}_{d,J}$ in (7) for, e.g., linear splines, possesses a condition number that is of the order $\mathscr{O}(2^{2J})$. Thus, classical iterative solution methods for (7) like the Jacobi method, the steepest descent approach or the conjugate gradient technique converge successively slower for rising values of J. The same is true for the Gauss-Seidel and the SOR methods. This problem is remedied by a multigrid method or a multilevel preconditioner. Then, the number of iterations necessary to obtain a prescribed accuracy is bounded independently of J, cf. [Hac85, Xu92, BL11, Bra07]. To this end, besides the grid and the basis functions on the finest scale J, also the grids and basis functions on all coarser isotropic scales are included in the iterative process, i.e. the multiscale generating system

$$\bigcup_{l=1}^{J} \{\ \phi_{l,\mathbf{i}} : 1 \le i_p \le n_l, p = 1, \ldots, d\ \}$$

is employed. Note that there is work that relates classical multigrid theory to multiplicative iterative algorithms operating on such a generating system [Gri94b, Gri94a]. Furthermore, the BPX-preconditioner [BPX90] can be identified with one step of the additive Jacobi iteration. Both methods guarantee asymptotically optimal convergence rates that are independent of J. However, the corresponding rates still depend on the dimension d.

To overcome this issue, we follow a different approach which relies on *all* coarser isotropic and anisotropic scales. To this end, we define the spaces

$$V_{\mathbf{l}} = V_{l_1} \otimes \cdots \otimes V_{l_d} \tag{10}$$

for the multiindices $\mathbf{l} = (l_1, \ldots, l_d) \in \mathbb{N}^d$. Next, we define the index sets

$$\chi_{\mathbf{l}} = \{1,\ldots,n_{l_1}\} \times \cdots \times \{1,\ldots,n_{l_d}\}$$

and the associated basis functions

$$\phi_{\mathbf{l},\mathbf{i}} = \phi_{l_1,i_1} \cdot \cdots \cdot \phi_{l_d,i_d} \quad \text{for} \quad \mathbf{i} = (i_1,\ldots,i_d) \in \chi_{\mathbf{l}} \,. \tag{11}$$

Obviously, it holds $V_{\mathbf{l}} = \operatorname{span}\{\phi_{\mathbf{l},\mathbf{i}} : \mathbf{i} \in \chi_{\mathbf{l}}\}$. From now on, $n_{\mathbf{l}} := |\chi_{\mathbf{l}}|$ denotes the number of degrees of freedom of the subspace $V_{\mathbf{l}}$. The isotropic spaces (5) can be expressed in this setting by $V_l^d = V_{\mathbf{l}}$, where $\mathbf{l} = (l,\ldots,l)$, and the isotropic functions (6) are given as $\phi_{l,\mathbf{i}} = \phi_{\mathbf{l},\mathbf{i}}$ for $\mathbf{i} \in \chi_{\mathbf{l}}$.

Our enlarged generating system includes all basis functions

$$\bigcup_{\mathbf{l} \in \mathscr{F}_J^d} \{\phi_{\mathbf{l},\mathbf{i}} : \mathbf{i} \in \chi_{\mathbf{l}}\} \,, \tag{12}$$

where the index set

$$\mathscr{F}_J^d = \{\mathbf{l} \in \mathbb{N}^d : |\mathbf{l}|_\infty \le J\} \tag{13}$$

contains the multiindices $\mathbf{l}$ of all coarser scales, i.e. $V_{\mathbf{l}} \subset V_J^d$ for $\mathbf{l} \in \mathscr{F}_J^d$. Next, the weak problem (2) for V_J^d leads with (12) to the enlarged system

$$\widehat{\mathbf{A}}_{d,J} \widehat{\mathbf{x}}_{d,J} = \widehat{\mathbf{b}}_{d,J} \tag{14}$$

of linear equations, with $\widehat{\mathbf{A}}_{d,J} \in \mathbb{R}^{\widehat{N}_{d,J} \times \widehat{N}_{d,J}}$ and $\widehat{\mathbf{x}}_{d,J}, \widehat{\mathbf{b}}_{d,J} \in \mathbb{R}^{\widehat{N}_{d,J}}$, where $\widehat{N}_{d,J} := \left(\sum_{l=1}^J n_l\right)^d$. The matrix $\widehat{\mathbf{A}}_{d,J}$ is block-structured with blocks $(\widehat{\mathbf{A}}_{d,J})_{\mathbf{l},\mathbf{k}} \in \mathbb{R}^{n_{\mathbf{l}} \times n_{\mathbf{k}}}$ for $\mathbf{l},\mathbf{k} \in \mathscr{F}_J^d$, where

$$((\widehat{\mathbf{A}}_{d,J})_{\mathbf{l},\mathbf{k}})_{\mathbf{i},\mathbf{j}} = a(\phi_{\mathbf{l},\mathbf{i}}, \phi_{\mathbf{k},\mathbf{j}}) \quad \text{for} \quad \mathbf{i} \in \chi_{\mathbf{l}}, \mathbf{j} \in \chi_{\mathbf{k}}$$

and the right-hand side vector $\widehat{\mathbf{b}}_{d,J}$ consists of blocks $(\widehat{\mathbf{b}}_{d,J})_{\mathbf{l}} \in \mathbb{R}^{n_{\mathbf{l}}}$, $\mathbf{l} \in \mathscr{F}_J^d$, with

$$((\widehat{\mathbf{b}}_{d,J})_{\mathbf{l}})_{\mathbf{i}} = (\phi_{\mathbf{l},\mathbf{i}}, f)_{L^2(\Omega^d)} \quad \text{for} \quad \mathbf{i} \in \chi_{\mathbf{l}} \,.$$

Note that the non-unique representation of functions in the enlarged generating system (12) results in a non-trivial kernel of $\widehat{\mathbf{A}}_{d,J}$. Thus $\widehat{\mathbf{A}}_{d,J}$ is not invertible. But the system (14) is nevertheless solvable since the right-hand side $\widehat{\mathbf{b}}_{d,J}$ lies in the range of the system matrix. A solution can be generated by any semi-convergent iterative method [BP94]. Many convergence results, e.g., for the steepest descent or conjugate gradient method, also apply to the semi-definite case, cf. [Gri94b]. There, the usual condition number κ is no longer defined, but the *generalized* condition number $\tilde{\kappa}$, i.e. the ratio of the largest and the smallest *non-zero* eigenvalue, is now decisive for the speed of convergence.

Just like in (8), we can express our enlarged system matrix as the sum of Kronecker product matrices, i.e.

$$\widehat{\mathbf{A}}_{d,J} = \widehat{\mathbf{A}}_{1,J} \otimes \widehat{\mathbf{M}}_{d-1,J} + \sum_{p=2}^{d-1} \widehat{\mathbf{M}}_{p-1,J} \otimes \widehat{\mathbf{A}}_{1,J} \otimes \widehat{\mathbf{M}}_{d-p,J} + \widehat{\mathbf{M}}_{d-1,J} \otimes \widehat{\mathbf{A}}_{1,J} \ ,$$

where $\widehat{\mathbf{A}}_{1,J} \in \mathbb{R}^{(\sum_{l=1}^{J} n_l)\times(\sum_{l=1}^{J} n_l)}$ is the multilevel stiffness matrix of the one-dimensional bilinear form and reads

$$(\widehat{\mathbf{A}}_{1,J})_{(l,i),(k,j)} = \left(\frac{\partial \phi_{l,i}}{\partial x}, \frac{\partial \phi_{k,j}}{\partial x} \right)_{L^2(\Omega)} \quad \text{for} \quad 1 \le i \le n_l, 1 \le j \le n_k, 1 \le l,k \le J \ .$$

Furthermore, $\widehat{\mathbf{M}}_{p,J} \in \mathbb{R}^{(\sum_{l=1}^{J} n_l)^p \times (\sum_{l=1}^{J} n_l)^p}$ also has Kronecker product structure

$$\widehat{\mathbf{M}}_{p,J} = \bigotimes_{q=1}^{p} \widehat{\mathbf{M}}_{1,J}$$

with $\widehat{\mathbf{M}}_{1,J} \in \mathbb{R}^{(\sum_{l=1}^{J} n_l)\times(\sum_{l=1}^{J} n_l)}$ and

$$(\widehat{\mathbf{M}}_{1,J})_{(l,i),(k,j)} = (\phi_{l,i}, \phi_{k,j})_{L^2(\Omega)} \quad \text{for} \quad 1 \le i \le n_l, 1 \le j \le n_k, 1 \le l,k \le J \ .$$

Of course, at some point, we need to be able to transform the non-unique solution $\widehat{\mathbf{x}}_{d,J}$ of (14) to the unique solution $\mathbf{x}_{d,J}$ of (7). To this end, we assume to have matrices $\mathbf{I}_l^k \in \mathbb{R}^{n_k \times n_l}$, which are one-dimensional restrictions from level l to level k for $l > k$, prolongations from level l to level k for $l < k$ and the identity matrix for $l = k$. Note here that the $\mathbf{I}_l^k, k \neq l \pm 1$ can be expressed as just a product of successive 2-level restrictions and prolongations, respectively, i.e. we have

$$\mathbf{I}_l^k = \mathbf{I}_{k+1}^k \cdots \mathbf{I}_l^{l-1} \quad \text{for} \quad l > k \quad \text{and} \quad \mathbf{I}_l^k = \mathbf{I}_{k-1}^k \cdots \mathbf{I}_l^{l+1} \quad \text{for} \quad l < k \ . \tag{15}$$

Naturally, the multi-dimensional case is obtained by the product construction

$$\mathbf{I}_{\mathbf{l}}^{\mathbf{k}} = \bigotimes_{p=1}^{d} \mathbf{I}_{l_p}^{k_p} \ . \tag{16}$$

Then, for $\mathbf{l}, \mathbf{k} \in \mathscr{F}_J^d$, we can express any block $(\widehat{\mathbf{A}}_{d,J})_{\mathbf{l},\mathbf{k}} \in \mathbb{R}^{n_{\mathbf{l}} \times n_{\mathbf{k}}}$ and any part $(\widehat{\mathbf{b}}_{d,J})_{\mathbf{l}}$ as

$$(\widehat{\mathbf{A}}_{d,J})_{\mathbf{l},\mathbf{k}} = \mathbf{I}_{\mathbf{J}}^{\mathbf{l}} \mathbf{A}_{d,J} \mathbf{I}_{\mathbf{k}}^{\mathbf{J}} \quad \text{and} \quad (\widehat{\mathbf{b}}_{d,J})_{\mathbf{l}} = \mathbf{I}_{\mathbf{J}}^{\mathbf{l}} \mathbf{b}_{d,J} \ , \tag{17}$$

respectively, where $\mathbf{J} = (J, \ldots, J)$ is the multiindex that describes the finest level on the isotropic scale. In the special case of identity matrices, we sometimes abbreviate $\mathbf{I}_l^l$ by $\mathbf{I}_l$ and $\mathbf{I}_{\mathbf{l}}^{\mathbf{l}}$ by $\mathbf{I}_{\mathbf{l}}$, i.e. we drop the superscript if it is equal to the subscript.

Furthermore, let us define the rectangular block-structured matrix $\widehat{\mathbf{S}}_{1,J} \in \mathbb{R}^{n_J \times (\sum_{l=1}^{J} n_l)}$ by

$$\widehat{\mathbf{S}}_{1,J} = (\, \mathbf{I}_1^J \mid \ldots \mid \mathbf{I}_J^J \,) \ .$$

Then, we can express the block-structured matrix $\widehat{\mathbf{S}}_{d,J}$ as

$$\widehat{\mathbf{S}}_{d,J} = \bigotimes_{p=1}^{d} \widehat{\mathbf{S}}_{1,J} \ ,$$

and with (17) we obtain

$$\widehat{\mathbf{A}}_{d,J} = \widehat{\mathbf{S}}_{d,J}^T \mathbf{A}_{d,J} \widehat{\mathbf{S}}_{d,J} \quad \text{and} \quad \widehat{\mathbf{b}}_{d,J} = \widehat{\mathbf{S}}_{d,J}^T \mathbf{b}_{d,J} \ .$$

As a result, we see that $\mathbf{x}_{d,J} = \widehat{\mathbf{S}}_{d,J} \widehat{\mathbf{x}}_{d,J}$ solves (7), if $\widehat{\mathbf{x}}_{d,J}$ is any solution to (14). Note that we will never set up the matrices $\widehat{\mathbf{S}}_{d,J}$ and $\widehat{\mathbf{S}}_{d,J}^T$ in our implementation, but compute their application to vectors by a straightforward algorithm in $\mathscr{O}(d \cdot \widehat{N}_{d,J})$ floating point operations using (15) and (16).

In Sect. 3, we will propose a matrix $\widehat{\mathbf{C}}_{d,J}$ that can be applied cheaply to a vector and acts as a preconditioner on the enlarged system (14) with

$$\tilde{\kappa}(\widehat{\mathbf{C}}_{d,J} \widehat{\mathbf{A}}_{d,J}) = \mathscr{O}(1) \tag{18}$$

independently of the level J and the dimension d. Since

$$\tilde{\kappa}(\widehat{\mathbf{C}}_{d,J} \widehat{\mathbf{A}}_{d,J}) = \tilde{\kappa}(\widehat{\mathbf{C}}_{d,J} \widehat{\mathbf{S}}_{d,J}^T \mathbf{A}_{d,J} \widehat{\mathbf{S}}_{d,J}) = \kappa(\widehat{\mathbf{S}}_{d,J} \widehat{\mathbf{C}}_{d,J} \widehat{\mathbf{S}}_{d,J}^T \mathbf{A}_{d,J}) \ ,$$

we can deduce that $\mathbf{C}_{d,J} := \widehat{\mathbf{S}}_{d,J} \widehat{\mathbf{C}}_{d,J} \widehat{\mathbf{S}}_{d,J}^T$ is thus a preconditioner for $\mathbf{A}_{d,J}$ with a resulting condition number that is bounded independently of J *and* d. Before we can present this preconditioner in Sect. 3, we need to discuss a specific norm equivalence in the next subsection.

2.3 A Norm Equivalence Based on Orthogonal Subspaces

The multiresolution scale of subspaces (3) induces a sequence of L^2-orthogonal complement spaces $(W_l)_{l=1}^{\infty}$ with

$$V_l = V_{l-1} \oplus_{L^2} W_l \quad \text{for} \quad l \geq 1, \quad \text{and} \quad V_0 := \{0\} \ . \tag{19}$$

A recursive application of (19) then yields $V_l = \oplus_{k=1}^{l} W_k$. Analogously to the anisotropic full-grid subspaces $V_{\mathbf{l}}$ in (10), we can now define anisotropic orthogonal complement spaces by the d-fold tensor products

$$W_{\mathbf{l}} = W_{l_1} \otimes \cdots \otimes W_{l_d} \ , \tag{20}$$

which satisfy $W_{\mathbf{l}} \subset V_{\mathbf{l}}$ and $W_{\mathbf{l}} \perp_{L^2} W_{\mathbf{k}}$ for $\mathbf{l} \neq \mathbf{k}$.

Now, we assume that, due to Jackson- and Bernstein-inequalities [Dah96, Osw94] for the spaces $(V_l)_{l=1}^{\infty}$, we have a one-dimensional equivalence

$$\lambda_{\min} \sum_{l \in \mathbb{N}} 2^{2l} \|w_l\|_{L^2(\Omega)}^2 \leq \left\| \frac{\partial u}{\partial x} \right\|_{L^2(\Omega)}^2 \leq \lambda_{\max} \sum_{l \in \mathbb{N}} 2^{2l} \|w_l\|_{L^2(\Omega)}^2 \tag{21}$$

for $u \in H_0^1(\Omega)$ with $u = \sum_{l\in\mathbb{N}} w_l$, where $w_l \in W_l, l \in \mathbb{N}$, and $0 < \lambda_{\min} \le \lambda_{\max} < \infty$. In the following, we will use the symbol $\simeq$ to indicate such an equivalence and call $\lambda_{\min}$ and $\lambda_{\max}$ *norm equivalence constants*. The next theorem shows that a similar equivalence exists in higher dimensions with dimension-independent constants.

Theorem 1. *For $u \in H_0^1(\Omega^d)$, it holds that*

$$a(u,u) \simeq \sum_{\mathbf{l}\in\mathbb{N}^d} \Big(\sum_{p=1}^{d} 2^{2l_p}\Big) \|w_\mathbf{l}\|^2_{L^2(\Omega^d)} \quad \textit{for} \quad u = \sum_{\mathbf{l}\in\mathbb{N}^d} w_\mathbf{l} \quad \textit{with} \quad w_\mathbf{l} \in W_\mathbf{l}, \mathbf{l} \in \mathbb{N}^d\,, \tag{22}$$

where the constants $\lambda_{min}^{(d)}$ and $\lambda_{max}^{(d)}$ associated with (22) *are the same as in* (21)*, i.e. $\lambda_{min}^{(d)} = \lambda_{min}$ and $\lambda_{max}^{(d)} = \lambda_{max}$.*

Proof. In (4), we have introduced $(\phi_{J,i})_{i=1}^{n_J}$ as a basis for the space V_J. Of course, there also exists a L^2-orthonormal basis $(\psi_{J,i})_{i=1}^{n_J}$ of V_J. Furthermore, we need the orthogonal decomposition $(\omega_{l,i})_{l=1}^{J}$ of $\psi_{J,i} \in V_J$ for all $i = 1,\dots,n_J$ with $\omega_{l,i} \in W_l, l = 1,\dots,J$ and

$$\psi_{J,i} = \sum_{l=1}^{J} \omega_{l,i}\,.$$

Next, analogously to (11), we define

$$\psi_{\mathbf{J},\mathbf{i}}(\mathbf{x}) = \psi_{J,i_1}(x_1)\dots\psi_{J,i_d}(x_d) \quad \text{and} \quad \omega_{\mathbf{l},\mathbf{i}}(\mathbf{x}) = \omega_{l_1,i_1}(x_1)\cdots\omega_{l_d,i_d}(x_d)$$

for all $\mathbf{i} \in \chi_\mathbf{J}$ and $\mathbf{l} \in \mathscr{F}_J^d$. This opens a direct way to find orthogonal decompositions of functions $u = \sum_{\mathbf{i}\in\chi_\mathbf{J}} \alpha_\mathbf{i}\psi_{\mathbf{J},\mathbf{i}} \in V_J^d$ by

$$u = \sum_{\mathbf{i}\in\chi_\mathbf{J}} \alpha_\mathbf{i} \sum_{\mathbf{l}\in\mathscr{F}_J^d} \omega_{\mathbf{l},\mathbf{i}} = \sum_{\mathbf{l}\in\mathscr{F}_J^d}\sum_{\mathbf{i}\in\chi_\mathbf{J}} \alpha_\mathbf{i}\omega_{\mathbf{l},\mathbf{i}} = \sum_{\mathbf{l}\in\mathscr{F}_J^d} w_\mathbf{l}$$

with

$$w_\mathbf{l} = \sum_{\mathbf{i}\in\chi_\mathbf{J}} \alpha_\mathbf{i}\omega_{\mathbf{l},\mathbf{i}} \in W_\mathbf{l} \tag{23}$$

for all $\mathbf{l} \in \mathscr{F}_J^d$.

Now, we show that the norm equivalence (22) holds for any $u \in V_J^d$ with the constants $\lambda_{\max}$ and $\lambda_{\min}$ from (21). We have

$$a(u,u) = \sum_{p=1}^{d} \Big(\frac{\partial}{\partial x_p}\sum_{\mathbf{i}\in\chi_\mathbf{J}} \alpha_\mathbf{i}\psi_{\mathbf{J},\mathbf{i}}, \frac{\partial}{\partial x_p}\sum_{\mathbf{j}\in\chi_\mathbf{J}} \alpha_\mathbf{j}\psi_{\mathbf{J},\mathbf{j}}\Big)_{L^2(\Omega^d)} \tag{24}$$

$$= \sum_{p=1}^{d}\sum_{\mathbf{i}\in\chi_\mathbf{J}}\sum_{\mathbf{j}\in\chi_\mathbf{J}} \Big(\frac{\partial}{\partial x_p}\alpha_\mathbf{i}\psi_{J,i_p}, \frac{\partial}{\partial x_p}\alpha_\mathbf{j}\psi_{J,j_p}\Big)_{L^2(\Omega)} \prod_{\substack{q=1\\q\neq p}}^{d} (\psi_{J,i_q},\psi_{J,j_q})_{L^2(\Omega)} \tag{25}$$

$$= \sum_{p=1}^{d}\sum_{\substack{\mathbf{i}'=\mathbf{i}\ominus\{i_p\}\\ \mathbf{i}\in\chi_\mathbf{J}}} \Big(\frac{\partial}{\partial x_p}\sum_{i_p=1}^{n_J}\alpha_{\mathbf{i}'\oplus\{i_p\}}\psi_{J,i_p}, \frac{\partial}{\partial x_p}\sum_{j_p=1}^{n_J}\alpha_{\mathbf{i}'\oplus\{j_p\}}\psi_{J,j_p}\Big)_{L^2(\Omega)}\,. \tag{26}$$

We obtain (25) by repeated application of the distributive law and by using the product structure of the L^2-scalar product. Then, the orthonormal basis property of the $(\psi_{J,i})_{i=1}^{n_J}$ cancels all terms for $i_q \neq j_q, q \neq p$, and we get (26). Note that

$$\begin{aligned}\mathbf{i}' := \mathbf{i} \ominus \{i_p\} = (i_1, \ldots, i_{p-1}, i_{p+1}, \ldots, i_d) \quad \text{and} \\ \mathbf{i}' \oplus \{i_p\} = (i_1, \ldots, i_{p-1}, i_p i_{p+1}, \ldots, i_d)\,.\end{aligned}$$

We can apply the one-dimensional norm equivalence (21) to (26) and get the upper bound

$$\cdots \leq \sum_{p=1}^{d} \lambda_{\max} \sum_{\substack{\mathbf{i}'=\mathbf{i}\ominus\{i_p\} \\ \mathbf{i}\in\chi_\mathbf{J}}} \sum_{l_p=1}^{J} 2^{2l_p} \Big(\sum_{i_p=1}^{n_J} \alpha_{\mathbf{i}'\oplus\{i_p\}} \omega_{l_p,i_p}, \sum_{j_p=1}^{n_J} \alpha_{\mathbf{i}'\oplus\{j_p\}} \omega_{l_p,j_p} \Big)_{L^2(\Omega)} \tag{27}$$

$$= \lambda_{\max} \sum_{p=1}^{d} \sum_{\mathbf{i}\in\chi_\mathbf{J}} \sum_{\mathbf{j}\in\chi_\mathbf{J}} \sum_{l_p=1}^{J} 2^{2l_p} (\alpha_\mathbf{i} \omega_{l_p,i_p}, \alpha_\mathbf{j} \omega_{l_p,j_p})_{L^2(\Omega)} \cdot \prod_{\substack{q=1 \\ q\neq p}}^{d} (\psi_{J,i_q}, \psi_{J,j_q})_{L^2(\Omega)} \tag{28}$$

$$= \lambda_{\max} \sum_{p=1}^{d} \sum_{\mathbf{i}\in\chi_\mathbf{J}} \sum_{\mathbf{j}\in\chi_\mathbf{J}} \sum_{\mathbf{l}\in\mathscr{F}_J^d} 2^{2l_p} (\alpha_\mathbf{i} \omega_{l_p,i_p}, \alpha_\mathbf{j} \omega_{l_p,j_p})_{L^2(\Omega)} \cdot \prod_{\substack{q=1 \\ q\neq p}}^{d} (\omega_{l_q,i_q}, \omega_{l_q,j_q})_{L^2(\Omega)} \tag{29}$$

$$= \lambda_{\max} \sum_{\mathbf{l}\in\mathscr{F}_J^d} \Big(\sum_{p=1}^{d} 2^{2l_p} \Big) \Big(\sum_{\mathbf{i}\in\chi_\mathbf{J}} \alpha_\mathbf{i} \omega_{\mathbf{l},\mathbf{i}}, \sum_{\mathbf{j}\in\chi_\mathbf{J}} \alpha_\mathbf{j} \omega_{\mathbf{l},\mathbf{j}} \Big)_{L^2(\Omega^d)}\,. \tag{30}$$

In (27) and (28), we used the distributive law again and reintroduced the terms we dropped previously. In (29), we replaced the ψ_{J,i_q} and ψ_{J,j_q} by the decompositions $\sum_{l_q=1}^{J} \omega_{l_q,i_q}$ and $\sum_{l_q=1}^{J} \omega_{l_q,j_q}$, respectively. Then, in (30), we recombined the product of d one-dimensional L^2-scalar products to one d-dimensional L^2-scalar product. Note that the lower bound with $\lambda_{\min}$ can be proven in the same way. Now, in combination with (23), we know that (22) is a norm equivalence with constants $\lambda_{\max}^{(d)} \leq \lambda_{\max}$ and $\lambda_{\min}^{(d)} \geq \lambda_{\min}$.

Next, our goal is to prove the sharpness of the estimates, i.e. we will show that indeed $\lambda_{\max}^{(d)} = \lambda_{\max}$ and $\lambda_{\min}^{(d)} = \lambda_{\min}$. Since (21) holds for $\lambda_{\max}$ and $\lambda_{\min}$ on V^d, it also holds on $V_J^d \subset V^d$ with optimal constants $\lambda_{\max}(J) \leq \lambda_{\max}$ and $\lambda_{\min}(J) \geq \lambda_{\min}$. We now choose $u_{\max,J} \in V_J$ associated with the constant $\lambda_{\max}(J)$ of (21), and plug the multivariate function

$$u(\mathbf{x}) = u_{\max,J}(x_1) \cdots u_{\max,J}(x_d)$$

into (24). This results in an equality instead of an upper bound in (27) with the constant $\lambda_{\max}(J)$ instead of $\lambda_{\max}$. Because of

$$\lambda_{\max}^{(d)} \geq \lambda_{\max}(J) \nearrow \lambda_{\max} \quad \text{for} \quad J \to \infty\,,$$

we can conclude that $\lambda_{\max}^{(d)} = \lambda_{\max}$. The $\lambda_{\min}^{(d)}$-case can be shown analogously. □

The norm equivalence (22) can be found in, e.g., [GK09] or in [GO95] for the $2d$-case, but so far no special attention was paid to the dimension-independence of the equivalence constants. A remark in that direction can also be found in [CS12].

3 A Dimension-Independent Full Grid Preconditioner

The norm equivalence (22) holds for orthogonal subspaces $(W_{\mathbf{l}})_{\mathbf{l}\in\mathscr{F}_J^d}$. In order to make this result available to our discretization, which is based on the subspaces $(V_{\mathbf{l}})_{\mathbf{l}\in\mathscr{F}_J^d}$, see Sect. 2, we need an orthogonalization operator, which will be defined in the next subsection. Then, in Subsect. 3.2, we can finally present our new preconditioner.

So far, we have used d and J as subscripts to indicate the dependence on the dimension and the discretization level. For the following operators and matrices this dependence is still present, but we will omit these subscripts for better readability.

3.1 Orthogonalization Operator

We now consider the whole multivariate sequence of subspaces $V_{\mathbf{l}}, \mathbf{l} \in \mathscr{F}_J^d$, which we denote as

$$\widehat{V}_J^d = (V_{\mathbf{l}})_{\mathbf{l}\in\mathscr{F}_J^d} .$$

For $\widehat{u}, \widehat{v} \in (V_{\mathbf{l}})_{\mathbf{l}\in\mathscr{F}_J^d}$, we define the scalar product

$$(\widehat{u}, \widehat{v})_{\widehat{V}_J^d} = \sum_{\mathbf{l}\in\mathscr{F}_J^d} (u_{\mathbf{l}}, v_{\mathbf{l}})_{L^2(\Omega^d)} \quad \text{for} \quad \widehat{u} = (u_{\mathbf{l}})_{\mathbf{l}\in\mathscr{F}_J^d} \quad \text{and} \quad \widehat{v} = (v_{\mathbf{l}})_{\mathbf{l}\in\mathscr{F}_J^d} .$$

Then, we define the operator $\widehat{P} : \widehat{V}_J^d \to \widehat{V}_J^d$ by

$$\widehat{P}\widehat{u} = (Q_{W_{\mathbf{l}}} u_{\mathbf{l}})_{\mathbf{l}\in\mathscr{F}_J^d} ,$$

where $Q_{W_{\mathbf{l}}} : V^d \to W_{\mathbf{l}}$ is the standard L^2-projection into $W_{\mathbf{l}}$, i.e. it holds

$$(Q_{W_{\mathbf{l}}} u, w_{\mathbf{l}})_{L^2(\Omega^d)} = (u, w_{\mathbf{l}})_{L^2(\Omega^d)} \quad \text{for all } \; w_{\mathbf{l}} \in W_{\mathbf{l}} \tag{31}$$

for $u \in V^d$. The following well-known Lemma 1 is the basis for an efficient computation of $Q_{W_{\mathbf{l}}} u, \mathbf{l} \in \mathscr{F}_J^d$ without an explicit discretization of the spaces $W_{\mathbf{l}}$.

Lemma 1. *There holds the identity*

$$Q_{W_{\mathbf{l}}} = (Q_{V_{l_1}} - Q_{V_{l_1-1}}) \otimes \cdots \otimes (Q_{V_{l_d}} - Q_{V_{l_d-1}}) ,$$

where $Q_{V_l} : V \to V_l$ denotes the one-dimensional standard L^2-projection into the space V_l.

Proof. We abbreviate $z_{\mathbf{l}} = (Q_{V_{l_1}} - Q_{V_{l_1-1}}) \otimes \cdots \otimes (Q_{V_{l_d}} - Q_{V_{l_d-1}})u$. First, we have to show that $z_{\mathbf{l}} \in W_{\mathbf{l}}$. It is obvious that $z_{\mathbf{l}} \in V_{\mathbf{l}}$, but we also have to establish the orthogonality to all $v_{\mathbf{k}} \in V_{\mathbf{k}}$ for $\mathbf{k} \leq \mathbf{l}, \mathbf{k} \neq \mathbf{l}$. To this end, let us pick an index $i \in \{1, \ldots, d\}$ with $k_i < l_i$. Then, we have

$$\begin{aligned}(z_{\mathbf{l}}, v_{\mathbf{k}})_{L^2(\Omega^d)} &= (\cdots \otimes (Q_{V_{l_i}} - Q_{V_{l_i-1}}) \otimes \ldots u, v_{\mathbf{k}})_{L^2(\Omega^d)} \\ &= (\cdots \otimes (Q_{V_{l_i}} - Q_{V_{l_i}}) \otimes \ldots u, v_{\mathbf{k}})_{L^2(\Omega^d)} = 0 \,. \end{aligned} \tag{32}$$

In (32), we used the d-dimensional generalization of the equality

$$(Q_{V_{l-1}} u, v_k)_{L^2(\Omega)} = (u, v_k)_{L^2(\Omega)} = (Q_{V_l} u, v_k)_{L^2(\Omega)} \quad \text{for all} \quad v_k \in V_k \text{ with } l-1 \geq k \,,$$

which holds since $V_k \subset V_{l-1} \subset V$. Now, we know that $z_{\mathbf{l}} \in W_{\mathbf{l}}$, but we still need to show (31). Due to the L^2-orthogonality of $w_{\mathbf{l}} \in W_{\mathbf{l}}$ to all functions in $V_{\mathbf{k}}$ with $\mathbf{k} \leq \mathbf{l}, \mathbf{k} \neq \mathbf{l}$, it holds that

$$\begin{aligned}(z_{\mathbf{l}}, w_{\mathbf{l}})_{L^2(\Omega^d)} &= ((Q_{V_{l_1}} - Q_{V_{l_1-1}}) \otimes \cdots \otimes (Q_{V_{l_d}} - Q_{V_{l_d-1}})u, w_{\mathbf{l}})_{L^2(\Omega^d)} \\ &= ((Q_{V_{l_1}} \otimes \cdots \otimes Q_{V_{l_d}})u, w_{\mathbf{l}})_{L^2(\Omega^d)} \\ &= (u, w_{\mathbf{l}})_{L^2(\Omega^d)} \,, \end{aligned}$$

and thus we have proven that $z_{\mathbf{l}} = Q_{W_{\mathbf{l}}} u$. □

The operator $\widehat{P}$ can be given in block-diagonal matrix form as $\widehat{\mathbf{P}} : \mathbb{R}^{\widehat{N}_{d,J} \times \widehat{N}_{d,J}}$ with blocks $(\widehat{\mathbf{P}})_{\mathbf{l},\mathbf{k}} \in \mathbb{R}^{n_{\mathbf{l}} \times n_{\mathbf{k}}}$ and

$$(\widehat{\mathbf{P}})_{\mathbf{l},\mathbf{k}} = \begin{cases} \mathbf{Q}_{W_{\mathbf{l}}} & \text{for} \quad \mathbf{l} = \mathbf{k} \,, \\ 0 & \text{else} \end{cases} \tag{33}$$

for all $\mathbf{l}, \mathbf{k} \in \mathscr{F}_J^d$, where $\mathbf{Q}_{W_{\mathbf{l}}} \in \mathbb{R}^{n_{\mathbf{l}} \times n_{\mathbf{l}}}$ is the matrix representation of the operator $Q_{W_{\mathbf{l}}}$ restricted to the subspace $V_{\mathbf{l}}$. According to Lemma 1, the matrices $\mathbf{Q}_{W_{\mathbf{l}}}$ can be expressed by

$$\mathbf{Q}_{W_{\mathbf{l}}} = (\mathbf{I}_{l_1} - \mathbf{I}_{l_1-1}^{l_1} (\mathbf{M}_{1,l_1-1})^{-1} \mathbf{I}_{l_1}^{l_1-1} \mathbf{M}_{1,l_1}) \otimes \cdots \otimes (\mathbf{I}_{l_d} - \mathbf{I}_{l_d-1}^{l_d} (\mathbf{M}_{1,l_d-1})^{-1} \mathbf{I}_{l_d}^{l_d-1} \mathbf{M}_{1,l_d}) \,, \tag{34}$$

where $\mathbf{M}_{1,l}$ are the non-hierarchical isotropic mass matrices from (9) with $J = l$. Note that, besides the simple 2-level restrictions and prolongations, d applications of one-dimensional mass matrices and d applications of the inverse of one-dimensional mass matrices are employed. Both operations can be cheaply executed since only band matrices are involved here, e.g., tridiagonal matrices for linear splines.

Note furthermore that $\widehat{\mathbf{P}}$ possesses the overall Kronecker product structure

$$\widehat{\mathbf{P}} = \bigotimes_{p=1}^{d} \widehat{\mathbf{P}}_1$$

with a block-diagonal $\widehat{\mathbf{P}}_1 \in \mathbb{R}^{N_{1,J} \times N_{1,J}}$, where

$$\widehat{\mathbf{P}}_1 = \mathrm{diag}(\mathbf{I}_1^1, \mathbf{I}_2^2 - \mathbf{I}_1^2(\mathbf{M}_{1,1})^{-1}\mathbf{I}_2^1\mathbf{M}_{1,2}, \ldots, \mathbf{I}_J^J - \mathbf{I}_{J-1}^J(\mathbf{M}_{1,J-1})^{-1}\mathbf{I}_J^{J-1}\mathbf{M}_{1,J}) \ .$$

The block-diagonal structure of $\widehat{\mathbf{P}}$ in combination with the Kronecker product structure (34) allows for an efficient application of $\widehat{\mathbf{P}}$ in our generating system, which involves $\mathscr{O}(d \cdot \widehat{N}_{d,J})$ floating point operations. A more detailed cost discussion will be given in Subsect. 3.3.

Note that even though the matrix $\widehat{\mathbf{P}}$ is block-diagonal, it is not symmetric since its blocks on the diagonal are not symmetric. This is even more remarkable as the corresponding operator $\widehat{P} : \widehat{V}_J^d \to \widehat{V}_J^d$ is self-adjoint. In fact, the non-symmetry is a property only of the matrix representation.

For our preconditioner, we also need to apply $\widehat{\mathbf{P}}^T$ efficiently. To obtain a favorable representation of $\widehat{\mathbf{P}}^T$, we first consider the mapping

$$\widehat{Z} : \mathbb{R}^{\widehat{N}_{d,J}} \to \widehat{V}_J^d$$

that maps a block-structured vector $\widehat{\mathbf{x}}_{d,J} = (x_{\mathbf{l},\mathbf{i}})_{\mathbf{i}\in\chi_\mathbf{l},\mathbf{l}\in\mathscr{F}_J^d}$ of the enlarged generating system to a collection of subspaces by

$$\widehat{Z} : \widehat{\mathbf{x}}_{d,J} \mapsto (\sum_{\mathbf{i}\in\chi_\mathbf{l}} x_{\mathbf{l},\mathbf{i}}\phi_{\mathbf{l},\mathbf{i}})_{\mathbf{l}\in\mathscr{F}_J^d} \ . \tag{35}$$

Note that $\widehat{P}$ and $\widehat{\mathbf{P}}$ are linked by $\widehat{\mathbf{P}} = \widehat{Z}^{-1}\widehat{P}\widehat{Z}$.

Lemma 2. *The adjoint $\widehat{Z}^* : \widehat{V}_J^d \to \mathbb{R}^{\widehat{N}_{d,J}}$ of* (35) *is given by*

$$\widehat{Z}^* : \widehat{u} \mapsto \widehat{\mathbf{x}}_{d,J} \quad \textit{with} \quad \widehat{\mathbf{x}}_{d,J} = ((u_\mathbf{l}, \phi_{\mathbf{l},\mathbf{i}})_{L^2(\Omega^d)})_{\mathbf{i}\in\chi_\mathbf{l},\mathbf{l}\in\mathscr{F}_J^d} \quad \textit{for} \quad \widehat{u} = (u_\mathbf{l})_{\mathbf{l}\in\mathscr{F}_J^d} \ .$$

Proof. For any $\widehat{v} = (v_\mathbf{l})_{\mathbf{l}\in\mathscr{F}_J^d} \in \widehat{V}_J^d$ and $\widehat{\mathbf{x}}_{d,J} \in \mathbb{R}^{\widehat{N}_{d,J}}$, we have

$$\begin{aligned}(\widehat{Z}\widehat{\mathbf{x}}_{d,J}, \widehat{v})_{\widehat{V}_J^d} = \sum_{\mathbf{l}\in\mathscr{F}_J^d} (\sum_{\mathbf{i}\in\chi_\mathbf{l}} x_{\mathbf{l},\mathbf{i}}\phi_{\mathbf{l},\mathbf{i}}, v_\mathbf{l})_{L^2(\Omega^d)} &= \sum_{\mathbf{l}\in\mathscr{F}_J^d}\sum_{\mathbf{i}\in\chi_\mathbf{l}} x_{\mathbf{l},\mathbf{i}}(v_\mathbf{l}, \phi_{\mathbf{l},\mathbf{i}})_{L^2(\Omega^d)} \\ &= (\widehat{\mathbf{x}}_{d,J}, \widehat{Z}^*\widehat{v})_{\ell^2} \ . \qquad \square\end{aligned}$$

Now, having $\widehat{Z}$ and $\widehat{Z}^*$, we are able to give a computationally efficient representation of $\widehat{\mathbf{P}}^T$.

Lemma 3. *It holds that*

$$\widehat{\mathbf{P}}^T = \widehat{\mathbf{G}}\widehat{\mathbf{P}}\widehat{\mathbf{G}}^{-1} \ ,$$

where $\widehat{\mathbf{G}} : \mathbb{R}^{\widehat{N}_{d,J}\times\widehat{N}_{d,J}}$ is a block-diagonal matrix with blocks $(\widehat{\mathbf{G}})_{\mathbf{l},\mathbf{k}} \in \mathbb{R}^{n_\mathbf{l}\times n_\mathbf{k}}$ and

$$(\widehat{\mathbf{G}})_{\mathbf{l},\mathbf{k}} = \begin{cases} \mathbf{M}_{\mathbf{l}} & \textit{for} \quad \mathbf{l} = \mathbf{k} \,, \\ 0 & \textit{else} \end{cases}$$

for all $\mathbf{l},\mathbf{k} \in \mathscr{F}_J^d$ *with the mass matrices* $\mathbf{M}_{\mathbf{l}} = \bigotimes_{p=1}^d \mathbf{M}_{1,l_p}$.

Proof. It holds that

$$(\widehat{Z}^* \widehat{Z} \widehat{\mathbf{x}}_{d,J}, \widehat{\mathbf{y}}_{d,J})_{\ell^2} = (\widehat{Z}\widehat{\mathbf{x}}_{d,J}, \widehat{Z}\widehat{\mathbf{y}}_{d,J})_{\widehat{V}_J^d} = \sum_{\mathbf{l} \in \mathscr{F}_J^d} \sum_{\mathbf{i},\mathbf{j} \in \chi_{\mathbf{l}}} x_{\mathbf{l},\mathbf{i}} (\phi_{\mathbf{l},\mathbf{i}}, \phi_{\mathbf{l},\mathbf{j}})_{L^2(\Omega^d)} y_{\mathbf{l},\mathbf{j}} = \widehat{\mathbf{x}}_{d,J}^T \widehat{\mathbf{G}} \widehat{\mathbf{y}}_{d,J} \,,$$

and thus $\widehat{Z}^*\widehat{Z} = \widehat{\mathbf{G}}$. Then, we can infer

$$(\widehat{\mathbf{P}})^T = (\widehat{Z}^{-1}\widehat{P}\widehat{Z})^* = \widehat{Z}^*\widehat{P}(\widehat{Z}^{-1})^* = \widehat{\mathbf{G}}\widehat{Z}^{-1}\widehat{P}\widehat{Z}\widehat{\mathbf{G}}^{-1} = \widehat{\mathbf{G}}\widehat{\mathbf{P}}\widehat{\mathbf{G}}^{-1} \,. \qquad \square$$

Note furthermore that the operator $\widehat{P}$ is a projection, i.e. $\widehat{P}\widehat{P} = \widehat{P}$. The same is true for $\widehat{\mathbf{P}}$ since

$$\widehat{\mathbf{P}}\widehat{\mathbf{P}} = \widehat{Z}^{-1}\widehat{P}\widehat{Z}\widehat{Z}^{-1}\widehat{P}\widehat{Z} = \widehat{Z}^{-1}\widehat{P}\widehat{P}\widehat{Z} = \widehat{Z}^{-1}\widehat{P}\widehat{Z} = \widehat{\mathbf{P}} \,.$$

Finally, we need the following Lemma.

Lemma 4. *For a block-diagonal scaling matrix* $\widehat{\mathbf{D}} \in \mathbb{R}^{\widehat{N}_{d,J} \times \widehat{N}_{d,J}}$ *with blocks* $(\widehat{\mathbf{D}})_{\mathbf{l},\mathbf{k}} \in \mathbb{R}^{n_{\mathbf{l}} \times n_{\mathbf{k}}}$ *for* $\mathbf{l},\mathbf{k} \in \mathscr{F}_J^d$ *and*

$$(\widehat{\mathbf{D}})_{\mathbf{l},\mathbf{k}} = \begin{cases} c_{\mathbf{l}} \mathbf{I}_{\mathbf{l}} & \textit{for} \quad \mathbf{l} = \mathbf{k} \,, \\ 0 & \textit{else} \,, \end{cases}$$

the matrix $\widehat{\mathbf{D}}$ *commutes with any other block-diagonal matrix* $\widehat{\mathbf{B}} \in \mathbb{R}^{\widehat{N}_{d,J} \times \widehat{N}_{d,J}}$*, i.e. a block-structured matrix with blocks* $(\widehat{\mathbf{B}})_{\mathbf{l},\mathbf{k}} \in \mathbb{R}^{n_{\mathbf{l}} \times n_{\mathbf{k}}}$ *for* $\mathbf{l},\mathbf{k} \in \mathscr{F}_J^d$*, where*

$$(\widehat{\mathbf{B}})_{\mathbf{l},\mathbf{k}} = \begin{cases} \mathbf{B}_{\mathbf{k}} & \textit{for} \quad \mathbf{l} = \mathbf{k} \,, \\ 0 & \textit{else} \,, \end{cases}$$

and $\mathbf{B}_{\mathbf{k}} \in \mathbb{R}^{n_{\mathbf{k}} \times n_{\mathbf{k}}}$ *are general matrices.*

Proof. For ease of notation, we use Kronecker's δ in this short proof. It holds that

$$\begin{aligned} (\widehat{\mathbf{D}}\widehat{\mathbf{B}})_{\mathbf{l},\mathbf{k}} &= \sum_{\mathbf{m} \in \mathscr{F}_J^d} (\widehat{\mathbf{D}})_{\mathbf{l},\mathbf{m}} (\widehat{\mathbf{B}})_{\mathbf{m},\mathbf{k}} = \sum_{\mathbf{m} \in \mathscr{F}_J^d} \delta_{\mathbf{l},\mathbf{m}} c_{\mathbf{l}} \delta_{\mathbf{m},\mathbf{k}} \mathbf{B}_{\mathbf{k}} = \delta_{\mathbf{l},\mathbf{k}} c_{\mathbf{l}} \mathbf{B}_{\mathbf{k}} \\ &= \delta_{\mathbf{l},\mathbf{k}} \mathbf{B}_{\mathbf{l}} c_{\mathbf{k}} = \sum_{\mathbf{m} \in \mathscr{F}_J^d} \delta_{\mathbf{l},\mathbf{m}} \mathbf{B}_{\mathbf{m}} \delta_{\mathbf{m},\mathbf{k}} c_{\mathbf{m}} = \sum_{\mathbf{m} \in \mathscr{F}_J^d} (\widehat{\mathbf{B}})_{\mathbf{l},\mathbf{m}} (\widehat{\mathbf{D}})_{\mathbf{m},\mathbf{k}} = (\widehat{\mathbf{B}}\widehat{\mathbf{D}})_{\mathbf{l},\mathbf{k}} \,, \end{aligned}$$

and thus $\widehat{\mathbf{D}}\widehat{\mathbf{B}} = \widehat{\mathbf{B}}\widehat{\mathbf{D}}$. $\square$

Obviously, Lemma 4 can be applied to, e.g., $\widehat{\mathbf{B}} = \widehat{\mathbf{P}}$ or $\widehat{\mathbf{B}} = \widehat{\mathbf{G}}$.

3.2 Preconditioner

Now we will present our new preconditioner for the operator matrix $\widehat{\mathbf{A}}_{d,J}$. To this end, the most important ingredient is the norm equivalence (22). From Theorem 1 we already know that its constants are independent of the dimension d and bounded independently of J.

Theorem 2. *Let $\widehat{\mathbf{D}} \in \mathbb{R}^{\widehat{N}_{d,J} \times \widehat{N}_{d,J}}$ be a diagonal block-structured scaling matrix with blocks $(\widehat{\mathbf{D}})_{\mathbf{l},\mathbf{k}} \in \mathbb{R}^{n_\mathbf{l} \times n_\mathbf{k}}$ and*

$$(\widehat{\mathbf{D}})_{\mathbf{l},\mathbf{k}} = \begin{cases} (\sum_{p=1}^{d} 2^{2l_p})\mathbf{I}_\mathbf{l} & \textit{for} \quad \mathbf{l} = \mathbf{k}\,, \\ 0 & \textit{else} \end{cases}$$

for all $\mathbf{l},\mathbf{k} \in \mathscr{F}_J^d$. Then, the generalized condition number of the symmetric matrix

$$\widehat{\mathbf{L}}^{-1}\widehat{\mathbf{P}}^T\widehat{\mathbf{D}}^{-1/2}\widehat{\mathbf{A}}_{d,J}\widehat{\mathbf{D}}^{-1/2}\widehat{\mathbf{P}}\widehat{\mathbf{L}}^{-T} \tag{36}$$

is bounded asymptotically with respect to J and is completely independent of the dimension d. Here, $\widehat{\mathbf{L}}$ is the Cholesky-factor of $\widehat{\mathbf{G}}$, i.e. $\widehat{\mathbf{G}} = \widehat{\mathbf{L}}\widehat{\mathbf{L}}^T$.

Proof. For any block-structured vector $\widehat{\mathbf{x}}_{d,J} \in \operatorname{im}\widehat{\mathbf{P}}$, we have

$$\begin{aligned}
\widehat{\mathbf{x}}_{d,J}^T\widehat{\mathbf{P}}^T\widehat{\mathbf{A}}_{d,J}\widehat{\mathbf{P}}\widehat{\mathbf{x}}_{d,J} &= \widehat{\mathbf{x}}_{d,J}^T\widehat{\mathbf{A}}_{d,J}\widehat{\mathbf{x}}_{d,J} && (37)\\
&= a\Big(\sum_{\mathbf{l}\in\mathscr{F}_J^d}\sum_{\mathbf{i}\in\chi_\mathbf{l}} x_{\mathbf{l},\mathbf{i}}\phi_{\mathbf{l},\mathbf{i}}, \sum_{\mathbf{l}\in\mathscr{F}_J^d}\sum_{\mathbf{i}\in\chi_\mathbf{l}} x_{\mathbf{l},\mathbf{i}}\phi_{\mathbf{l},\mathbf{i}}\Big) \\
&\simeq \sum_{\mathbf{l}\in\mathscr{F}_J^d}\Big(\sum_{p=1}^{d} 2^{2l_p}\Big)\Big\|\sum_{\mathbf{i}\in\chi_\mathbf{l}} x_{\mathbf{l},\mathbf{i}}\phi_{\mathbf{l},\mathbf{i}}\Big\|_{L^2(\Omega^d)}^2 && (38)\\
&= \sum_{\mathbf{l}\in\mathscr{F}_J^d}\Big(\sum_{p=1}^{d} 2^{2l_p}\Big)\mathbf{x}_\mathbf{l}^T\mathbf{M}_\mathbf{l}\mathbf{x}_\mathbf{l} \\
&= \widehat{\mathbf{x}}_{d,J}^T\widehat{\mathbf{D}}\widehat{\mathbf{G}}\widehat{\mathbf{x}}_{d,J}\,. && (39)
\end{aligned}$$

In (37), we have used $\widehat{\mathbf{P}}\widehat{\mathbf{x}}_{d,J} = \widehat{\mathbf{x}}_{d,J}$ and in (38), we have applied the norm equivalence (22). The levelwise summation of the mass matrix products was then expressed using the matrix $\widehat{\mathbf{G}}$ in (39). In the following, we need the block-diagonal factor $\widehat{\mathbf{L}}$ of the Cholesky-decomposition

$$\widehat{\mathbf{G}} = \widehat{\mathbf{L}}\widehat{\mathbf{L}}^T\,.$$

We set $\widehat{\mathbf{y}}_{d,J} = \widehat{\mathbf{L}}^T\widehat{\mathbf{D}}^{1/2}\widehat{\mathbf{x}}_{d,J}$ and obtain with $\widehat{\mathbf{D}}\widehat{\mathbf{G}} = \widehat{\mathbf{D}}^{1/2}\widehat{\mathbf{G}}\widehat{\mathbf{D}}^{1/2}$, see Lemma 4, the equation

$$\widehat{\mathbf{x}}_{d,J}^T\widehat{\mathbf{D}}\widehat{\mathbf{G}}\widehat{\mathbf{x}}_{d,J} = \widehat{\mathbf{x}}_{d,J}^T\widehat{\mathbf{D}}^{1/2}\widehat{\mathbf{L}}\widehat{\mathbf{L}}^T\mathbf{D}^{1/2}\widehat{\mathbf{x}}_{d,J} = \widehat{\mathbf{y}}_{d,J}^T\widehat{\mathbf{y}}_{d,J}\,.$$

Then, using the equivalence of (37) and (39), and $\widehat{\mathbf{x}}_{d,J} = \widehat{\mathbf{D}}^{-1/2}\widehat{\mathbf{L}}^{-T}\widehat{\mathbf{y}}_{d,J}$, we obtain the relation

$$\widehat{\mathbf{y}}_{d,J}^T\widehat{\mathbf{L}}^{-1}\widehat{\mathbf{D}}^{-1/2}\widehat{\mathbf{P}}^T\widehat{\mathbf{A}}_{d,J}\widehat{\mathbf{P}}\widehat{\mathbf{D}}^{-1/2}\widehat{\mathbf{L}}^{-T}\widehat{\mathbf{y}}_{d,J} \simeq \widehat{\mathbf{y}}_{d,J}^T\widehat{\mathbf{y}}_{d,J} \tag{40}$$

for all $\widehat{\mathbf{y}}_{d,J} \in \operatorname{im}\widehat{\mathbf{L}}^T\widehat{\mathbf{D}}^{1/2}\widehat{\mathbf{P}}$ with the same favorable constants as in (22). With the commuting of the matrices $\widehat{\mathbf{P}}$ and $\widehat{\mathbf{D}}^{-1/2}$, see Lemma 4, the left-hand side of (40) leads to (36).

Finally, we have to show that no $\widehat{\mathbf{v}}_{d,J} \in \mathbb{R}^{\widehat{N}_{d,J}}$ with $\widehat{\mathbf{v}}_{d,J} \perp \widehat{\mathbf{y}}_{d,J}$ affects the spectrum. From the Fundamental Theorem of Linear Algebra, from $\widehat{\mathbf{P}}^T = \widehat{\mathbf{G}}\widehat{\mathbf{P}}\widehat{\mathbf{G}}^{-1}$, see Lemma 3, and from $\widehat{\mathbf{P}}^T\widehat{\mathbf{D}}^{1/2} = \widehat{\mathbf{D}}^{1/2}\widehat{\mathbf{P}}^T$ we know that

$$\begin{aligned}\widehat{\mathbf{v}}_{d,J} \in\ & \ker\widehat{\mathbf{P}}^T\widehat{\mathbf{D}}^{1/2}\widehat{\mathbf{L}} = \ker\widehat{\mathbf{D}}^{1/2}\widehat{\mathbf{G}}\widehat{\mathbf{P}}\widehat{\mathbf{G}}^{-1}\widehat{\mathbf{L}} = \ker\widehat{\mathbf{D}}^{1/2}\widehat{\mathbf{G}}\widehat{\mathbf{P}}\widehat{\mathbf{L}}^{-T}\widehat{\mathbf{L}}^{-1}\widehat{\mathbf{L}} \\ &= \ker\widehat{\mathbf{P}}\widehat{\mathbf{L}}^{-T} \ .\end{aligned} \tag{41}$$

We dropped the matrix $\widehat{\mathbf{D}}^{1/2}\widehat{\mathbf{G}}$ from the kernel in the last equality (41), as it is a full-rank matrix and thus has no effect on the kernel. Obviously, if $\widehat{\mathbf{v}}_{d,J} \in \ker\widehat{\mathbf{P}}\widehat{\mathbf{L}}^{-T}$, then $\widehat{\mathbf{v}}_{d,J}$ belongs to the kernel of the preconditioned system (36). This finally proves the theorem. □

As a result of Theorem 2, we can express our preconditioner for $\widehat{\mathbf{A}}_{d,J}$ as

$$\widehat{\mathbf{C}}_{d,J} := \widehat{\mathbf{P}}\widehat{\mathbf{D}}^{-1}\widehat{\mathbf{G}}^{-1}\widehat{\mathbf{P}}^T \ .$$

Moreover, this approach also gives us with $\widehat{\mathbf{A}}_{d,J} = \widehat{\mathbf{S}}_{d,J}^T\mathbf{A}_{d,J}\widehat{\mathbf{S}}_{d,J}$ the preconditioner

$$\mathbf{C}_{d,J} := \widehat{\mathbf{S}}_{d,J}\widehat{\mathbf{P}}\widehat{\mathbf{D}}^{-1}\widehat{\mathbf{G}}^{-1}\widehat{\mathbf{P}}^T\widehat{\mathbf{S}}_{d,J}^T \tag{42}$$

for the fine grid system matrix $\mathbf{A}_{d,J}$. The preconditioned system possesses the same condition number, i.e. it is also independent of d and bounded independently of J.

3.3 Cost Discussion for the New Preconditioner

So far, we obtained a preconditioner with condition numbers independent of d and bounded independently of J. Of course, the question is now how high its computational costs are. Remember that a perfect preconditioner would be $\mathbf{A}_{d,J}^{-1}$ anyway, but it involves way too many computations. With (42) we now have a preconditioner $\mathbf{C}_{d,J}$ which comes, up to a d- and J-independent constant, close to $\mathbf{A}_{d,J}^{-1}$, but involves only a number of floating point operations that is linear in the number of degrees of freedom $\widehat{N}_{d,J}$ of the enlarged system.

We will now give a short discussion of the required matrix-vector multiplications and their costs, also with respect to the dimension d. As stated earlier, the application of the matrices $\widehat{\mathbf{S}}_{d,J}$ and $\widehat{\mathbf{S}}_{d,J}^T$ onto a vector can be implemented by a simple algorithm that exploits (15) in $\mathscr{O}(d \cdot \widehat{N}_{d,J})$ floating point operations. The application of $\widehat{\mathbf{D}}^{-1}$ is obviously possible with the same cost complexity. The matrix $\widehat{\mathbf{G}}^{-1}$ needs however a more elaborate discussion. As it is block-diagonal, its action can be implemented with an algorithm that works subspace by subspace. On every $V_{\mathbf{l}}, \mathbf{l} \in \mathscr{F}_J^d$, the mass matrix $\mathbf{M}_{\mathbf{l}} = \bigotimes_{p=1}^d \mathbf{M}_{1,l_p}$ must be inverted. As these matrices have Kronecker product structure, the inversion can be realized by the application of $\mathbf{M}_{1,l_p}^{-1}$ to the dimension p for $p = 1, \ldots, d$. We assume the functions $\{\phi_{\mathbf{l},\mathbf{i}}\}_{\mathbf{i} \in \chi_{\mathbf{l}}}$ to be of finite element type (h-version with fixed polynomial degree) having local support. Consequently, the associated one-dimensional matrices $\mathbf{M}_{1,l_p}$ have band matrix structure with constant band size and are thus invertible with linear costs[1]. As a result, we have a cost of $\mathscr{O}(d \cdot n_{\mathbf{l}})$ on each subspace and obtain a cost complexity of $\mathscr{O}(d \cdot \widehat{N}_{d,J})$ in total. The same argumentation holds for $\widehat{\mathbf{P}}$, which has a somewhat more complicated form, see (33) and (34), but also works subspace by subspace, where we can again exploit a Kronecker product structure. In total, we arrive at costs of $\mathscr{O}(d \cdot \widehat{N}_{d,J})$ for our preconditioner. The application of $\mathbf{A}_{d,J}$ is directly possible[2] in $\mathscr{O}(d^2 \cdot N_{d,J})$ due to the representation of the system matrix as a sum of Kronecker product matrices (8). In comparison, our preconditioner (42) is slightly more expensive since its costs depend on the *enlarged* system with $\widehat{N}_{d,J}$ degrees of freedom. However, a geometric series argument shows that

$$\widehat{N}_{d,J} = \mathscr{O}(2^d N_{d,J}) = \mathscr{O}(2^d 2^{Jd}) = \mathscr{O}(2^{(J+1)d}) = \mathscr{O}(N_{d,J+1}) ,$$

and thus the costs for our preconditioner on level J compare simply to the costs for a regular fine grid system on level $J+1$.

4 Sparse Grids

So far, we have dealt with the preconditioning of an isotropic full grid with $\mathscr{O}(N_{d,J})$ degrees of freedom. They scale exponentially with the dimension d and are thus impossible to deal with for $d > 4$ anyway. Under some additional smoothness requirements, sparse grids [BG04] remove this curse of dimension to some extent. Then, the multivariate multilevel structure is a fundamental necessity for both, a good preconditioner and the discretization itself. The implementation of a sparse grid multilevel discretization was already dealt with in [BG04, Feu10]. In the following, we discuss our new preconditioner for the sparse grid case in detail.

[1] Non-local basis functions (p-version) are likely to result in a Toeplitz-type matrix, which can be inverted in log-linear time.

[2] Note that it is even possible to execute this matrix-vector product in $\mathscr{O}(d \cdot N_{d,J})$ operations by the successive multiplications of $\mathbf{M}_{d,J}$ and of $\mathbf{A}_{d,J}\mathbf{M}_{d,J}^{-1} = \sum_{p=1}^d (\bigotimes_{q=1}^{p-1} \mathbf{I}_J) \otimes \mathbf{A}_{1,J}\mathbf{M}_{1,J}^{-1} \otimes (\bigotimes_{q=p+1}^d \mathbf{I}_J)$.

4.1 Definition

We can use an index set $\mathscr{I} \subset \mathbb{N}^d, |\mathscr{I}| < \infty$, which defines the subspaces included in some discretization by

$$V_{\mathscr{I}} = \sum_{\mathbf{l} \in \mathscr{I}} V_{\mathbf{l}} \,.$$

A proper choice of $\mathscr{I}$ now depends – besides the error we want to achieve – on the smoothness of the function class[3] for which we want to approximate.

For example, the full grid space V_J^d from (5) can be described by the index set $\mathscr{F}_J^d$ from (13), i.e. $V_J^d = V_{\mathscr{F}_J^d}$, and has the approximation property[4]

$$\inf_{v \in V_{\mathscr{F}_J^d}} \|u - v\|^2_{H^s(\Omega^d)} \leq c(d) 2^{-2(t-s)J} \|u\|^2_{H^t(\Omega^d)}$$

with rate $t-s$ and $u \in H_0^t(\Omega^d)$. Its number of degrees of freedom is of the order $\mathscr{O}(2^{Jd})$. Thus, the accuracy as function of the degrees of freedom deteriorates exponentially with rising d, which resembles the well-known 'curse of dimensionality', cf. [Bel61, BG04].

The sparse grid index set

$$\mathscr{S}_J^d = \{\mathbf{l} \in \mathbb{N}^d : |\mathbf{l}|_1 \leq J + d - 1\} \tag{43}$$

circumvents this problem to some extent provided that additional mixed smoothness $u \in H^t_{0,\text{mix}}(\Omega^d)$ is present. For details, see [BG04]. An example for the function system associated to $\mathscr{S}_J^d$ is given in Fig. 1 (right) for the two-dimensional case. The associated rate of best approximation

$$\inf_{v \in V_{\mathscr{S}_J^d}} \|u - v\|^2_{H^s(\Omega^d)} \leq c(d) 2^{-2(t-s)J} \|u\|^2_{H^t_{\text{mix}}(\Omega^d)}$$

is the same[5] as for the full grid space, i.e. $t-s$, but the number of degrees of freedom now is only of the order $\mathscr{O}(2^J J^{d-1})$ in J. This is a substantial improvement of the asymptotics in J in comparison to the full grid case. A-priori H^s-optimized sparse grids need, depending on the available smoothness class, even less degrees of freedom. For further details, cf. [BG04, GK09].

It is furthermore possible to adapt the index set $\mathscr{I}$ a-posteriori to a given function by means of a proper error estimation and a successive refinement procedure. This approach results in adaptively refined sparse grids, see e.g. [Feu10, GG03]. Note that for both, practical and theoretical reasons, our index set $\mathscr{I}$ needs to satisfy the

---

[3] In this paper, we restrict ourselves to homogeneous boundary conditions and do not introduce functions at the boundary. However, by $u = u_{\Omega^d} + u_\Gamma$ with $u_{\Omega^d}|_\Gamma = 0$ and $-\Delta u_{\Omega^d} = f + \Delta u_\Gamma$, we cover any case with Dirichlet boundary functions $u_\Gamma = g$ on Γ.

[4] This holds for a range of parameters $0 \leq s < t \leq r$ with r being the order of the spline of the space construction. In our case of linear splines $r = 2$ holds.

[5] Note that an additional logarithmic term appears in the error estimate for $s = 0$, cf. [BG04].

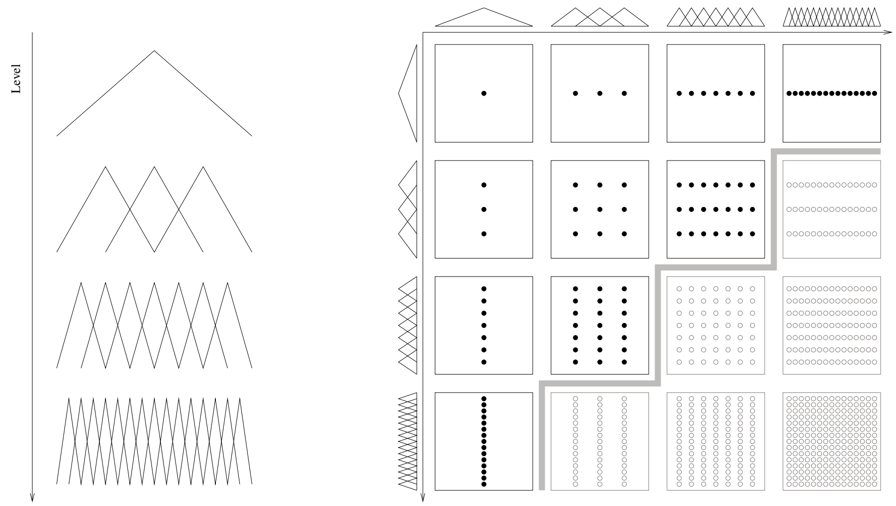

Fig. 1 The first four levels of a one-dimensional multilevel generating system based on linear splines (left). Two-dimensional tensorization and the sparse subspace (right)

admissibility condition

$$\mathbf{l} \in \mathscr{I}, \mathbf{k} \in \mathbb{N}^d, \mathbf{k} \leq \mathbf{l} \Rightarrow \mathbf{k} \in \mathscr{I} \ . \tag{44}$$

The number of degrees of freedom in the enlarged system for the regular sparse grid space $V_{\mathscr{S}_J^d}$ is $\widehat{N}_{d,J}^{\mathrm{SG}} = \sum_{\mathbf{l} \in \mathscr{S}_J^d} n_{\mathbf{l}}$. Note that here again some redundancy is involved, but the asymptotics in J of the number of degrees of freedom remains the same as for the sparse grid approach based on, e.g., the hierarchical basis [BG04], that is $\widehat{N}_{d,J}^{\mathrm{SG}} = \mathscr{O}(2^J J^{d-1})$ in J.

The weak problem (2) on $V_{\mathscr{S}_J^d}$ with the generating system

$$\bigcup_{\mathbf{l} \in \mathscr{S}_J^d} \{\phi_{\mathbf{l},\mathbf{i}} : \mathbf{i} \in \chi_{\mathbf{l}}\} \tag{45}$$

now leads to the equation

$$\widehat{\mathbf{A}}_{d,J}^{\mathrm{SG}} \widehat{\mathbf{x}}_{d,J}^{\mathrm{SG}} = \widehat{\mathbf{b}}_{d,J}^{\mathrm{SG}} \ . \tag{46}$$

Here, the matrix $\widehat{\mathbf{A}}_{d,J}^{\mathrm{SG}}$ is block-structured with blocks $(\widehat{\mathbf{A}}_{d,J}^{\mathrm{SG}})_{\mathbf{l},\mathbf{k}} \in \mathbb{R}^{n_{\mathbf{l}} \times n_{\mathbf{k}}}$ for $\mathbf{l}, \mathbf{k} \in \mathscr{S}_J^d$, where

$$((\widehat{\mathbf{A}}_{d,J}^{\mathrm{SG}})_{\mathbf{l},\mathbf{k}})_{\mathbf{i},\mathbf{j}} = a(\phi_{\mathbf{l},\mathbf{i}}, \phi_{\mathbf{k},\mathbf{j}}) \quad \text{for} \quad \mathbf{i} \in \chi_{\mathbf{l}}, \mathbf{j} \in \chi_{\mathbf{k}}$$

and the right-hand side vector $\widehat{\mathbf{b}}_{d,J}^{\mathrm{SG}}$ consists of blocks $(\widehat{\mathbf{b}}_{d,J}^{\mathrm{SG}})_{\mathbf{l}} \in \mathbb{R}^{n_{\mathbf{l}}}$, $\mathbf{l} \in \mathscr{S}_J^d$, with

$$((\widehat{\mathbf{b}}_{d,J}^{\mathrm{SG}})_{\mathbf{l}})_{\mathbf{i}} = (\phi_{\mathbf{l},\mathbf{i}}, f)_{L^2(\Omega^d)} \quad \text{for} \quad \mathbf{i} \in \chi_{\mathbf{l}} \ .$$

Similar to the full grid case (14), the non-unique representation in an enlarged sparse grid generating system (45) results in a non-trivial kernel of $\widehat{\mathbf{A}}^{\mathrm{SG}}_{d,J}$. Thus, $\widehat{\mathbf{A}}^{\mathrm{SG}}_{d,J}$ is not invertible. But, again, the system (46) is solvable since the right-hand side $\widehat{\mathbf{b}}^{\mathrm{SG}}_{d,J}$ lies in the range of the system matrix and a solution can be generated by any semi-convergent iterative method.

We will now describe the enlarged sparse grid system (46) as a submatrix and a subvector of the enlarged full grid system (14). Note that this is done here for theoretical purposes only. In our implementation we of course avoid the full grid system with $\widehat{N}_{d,J}$ degrees of freedom. In fact, our computational costs stay proportional to $\widehat{N}^{\mathrm{SG}}_{d,J}$, which is substantially smaller, cf. Subsect. 4.3.

Like in (17), we can express the blocks of $\widehat{\mathbf{A}}^{\mathrm{SG}}_{d,J}$ and $\widehat{\mathbf{b}}^{\mathrm{SG}}_{d,J}$ with respect to (7) by

$$(\widehat{\mathbf{A}}^{\mathrm{SG}}_{d,J})_{\mathbf{l},\mathbf{k}} = \mathbf{I}^{\mathbf{l}}_{\mathbf{J}}\mathbf{A}_{d,J}\mathbf{I}^{\mathbf{J}}_{\mathbf{k}} \quad \text{and} \quad (\widehat{\mathbf{b}}^{\mathrm{SG}}_{d,J})_{\mathbf{l}} = \mathbf{I}^{\mathbf{l}}_{\mathbf{J}}\mathbf{b}_{d,J}$$

for $\mathbf{l},\mathbf{k} \in \mathscr{S}^d_J$. Now, we can express our sparse grid operator matrix by

$$\widehat{\mathbf{A}}^{\mathrm{SG}}_{d,J} = \widehat{\mathbf{R}}_{d,J}\widehat{\mathbf{A}}_{d,J}\widehat{\mathbf{R}}^T_{d,J}\ , \tag{47}$$

and our right-hand side by

$$\widehat{\mathbf{b}}^{\mathrm{SG}}_{d,J} = \widehat{\mathbf{R}}_{d,J}\widehat{\mathbf{b}}_{d,J}\ ,$$

where $\widehat{\mathbf{R}}_{d,J} \in \mathbb{R}^{\widehat{N}^{\mathrm{SG}}_{d,J} \times \widehat{N}_{d,J}}$ is a rectangular block-structured matrix with

$$(\widehat{\mathbf{R}}_{d,J})_{\mathbf{l},\mathbf{k}} = \begin{cases} \mathbf{I}_{\mathbf{l}} & \text{for} \quad \mathbf{k} = \mathbf{l}\,, \\ 0 & \text{else}\,, \end{cases}$$

for $\mathbf{l} \in \mathscr{S}^d_J, \mathbf{k} \in \mathscr{F}^d_J$. Note that $\widehat{\mathbf{R}}^T_{d,J}\widehat{\mathbf{R}}_{d,J} \in \mathbb{R}^{\widehat{N}_{d,J} \times \widehat{N}_{d,J}}$ is a block-diagonal scaling matrix in the enlarged full grid system which simply sets all vector blocks to zero that belong to $\mathbf{l} \in \mathscr{F}^d_J \setminus \mathscr{S}^d_J$, and $\widehat{\mathbf{R}}_{d,J}\widehat{\mathbf{R}}^T_{d,J} \in \mathbb{R}^{\widehat{N}^{\mathrm{SG}}_{d,J} \times \widehat{N}^{\mathrm{SG}}_{d,J}}$ is simply the identity matrix on $\mathbb{R}^{\widehat{N}^{\mathrm{SG}}_{d,J}}$.

4.2 Sparse Grid Submatrix and Preconditioner

The proof of Theorem 2 showed that the condition number of $\mathbf{C}_{d,J}\mathbf{A}_{d,J}$ is independent of the dimension d and bounded independently of J. We will now extend this result to the sparse grid case by a submatrix argument. For reasons of simplicity, we stick here to the case of the regular sparse grid space $V_{\mathscr{S}^d_J}$ and the associated matrix $\widehat{\mathbf{A}}^{\mathrm{SG}}_{d,J}$, i.e. to the index set $\mathscr{S}^d_J$ of (43). But note that the following proof works with any index set $\mathscr{I} \subset \mathbb{N}^d$ with $\mathscr{I} \subset \mathscr{F}_{d,J}$ for which condition (44) is fulfilled.

Theorem 3. *The generalized condition number of the symmetric matrix*

$$\widehat{\mathbf{R}}_{d,J}\widehat{\mathbf{L}}^{-1}\widehat{\mathbf{P}}^T\widehat{\mathbf{D}}^{-1/2}\widehat{\mathbf{R}}_{d,J}^T\widehat{\mathbf{A}}_{d,J}^{SG}\widehat{\mathbf{R}}_{d,J}\widehat{\mathbf{D}}^{-1/2}\widehat{\mathbf{P}}\widehat{\mathbf{L}}^{-T}\widehat{\mathbf{R}}_{d,J}^T \in \mathbb{R}^{\widehat{N}_{d,J}^{SG}\times\widehat{N}_{d,J}^{SG}} \tag{48}$$

is less than or equal to the condition number of the preconditioned system $\mathbf{C}_{d,J}\mathbf{A}_{d,J}$ *of the full grid with same dimension* d *and level* J*. Thus, the generalized condition number of* $\widehat{\mathbf{C}}_{d,J}^{SG}\widehat{\mathbf{A}}_{d,J}^{SG}$ *with*

$$\widehat{\mathbf{C}}_{d,J}^{SG} := \widehat{\mathbf{R}}_{d,J}\widehat{\mathbf{P}}\widehat{\mathbf{D}}^{-1}\widehat{\mathbf{G}}^{-1}\widehat{\mathbf{P}}^T\widehat{\mathbf{R}}_{d,J}^T \tag{49}$$

is bounded asymptotically with respect to J *and* d*.*

Proof. We recall (40) from the proof of Theorem 2, i.e.

$$\widehat{\mathbf{y}}_{d,J}^T\widehat{\mathbf{L}}^{-1}\widehat{\mathbf{D}}^{-1/2}\widehat{\mathbf{P}}^T\widehat{\mathbf{A}}_{d,J}\widehat{\mathbf{P}}\widehat{\mathbf{D}}^{-1/2}\widehat{\mathbf{L}}^{-T}\widehat{\mathbf{y}}_{d,J} \simeq \widehat{\mathbf{y}}_{d,J}^T\widehat{\mathbf{y}}_{d,J} \tag{50}$$

for $\widehat{\mathbf{y}}_{d,J} \in \operatorname{im}\widehat{\mathbf{L}}^T\widehat{\mathbf{D}}^{1/2}\widehat{\mathbf{P}}$. We obtain equivalence constants, which are at least as good as those in (40), by the stronger condition

$$\widehat{\mathbf{y}}_{d,J} \in \operatorname{im}\widehat{\mathbf{R}}_{d,J}^T\widehat{\mathbf{R}}_{d,J}\widehat{\mathbf{L}}^T\widehat{\mathbf{D}}^{1/2}\widehat{\mathbf{P}} \subset \operatorname{im}\widehat{\mathbf{L}}^T\widehat{\mathbf{D}}^{1/2}\widehat{\mathbf{P}} \subset \mathbb{R}^{\widehat{N}_{d,J}} .$$

The image of $\widehat{\mathbf{R}}_{d,J}^T$ is not enlarged by block-diagonal matrices, and we can safely replace $\widehat{\mathbf{A}}_{d,J}$ by $\widehat{\mathbf{R}}_{d,J}^T\widehat{\mathbf{A}}_{d,J}^{\mathrm{SG}}\widehat{\mathbf{R}}_{d,J}$ on $\operatorname{im}\widehat{\mathbf{R}}_{d,J}^T$. This gives us

$$\widehat{\mathbf{y}}_{d,J}^T\widehat{\mathbf{L}}^{-1}\widehat{\mathbf{D}}^{-1/2}\widehat{\mathbf{P}}^T\widehat{\mathbf{R}}_{d,J}^T\widehat{\mathbf{A}}_{d,J}^{\mathrm{SG}}\widehat{\mathbf{R}}_{d,J}\widehat{\mathbf{P}}\widehat{\mathbf{D}}^{-1/2}\widehat{\mathbf{L}}^{-T}\widehat{\mathbf{y}}_{d,J} \simeq \widehat{\mathbf{y}}_{d,J}^T\widehat{\mathbf{y}}_{d,J}$$

with the same constants as in (50). Setting $\widehat{\mathbf{z}}_{d,J}^{\mathrm{SG}} = \widehat{\mathbf{R}}_{d,J}\widehat{\mathbf{y}}_{d,J}$ results in

$$\widehat{\mathbf{y}}_{d,J} = \widehat{\mathbf{R}}_{d,J}^T\widehat{\mathbf{R}}_{d,J}\widehat{\mathbf{y}}_{d,J} = \widehat{\mathbf{R}}_{d,J}^T\widehat{\mathbf{z}}_{d,J}^{\mathrm{SG}} ,$$

and we obtain

$$(\widehat{\mathbf{z}}_{d,J}^{\mathrm{SG}})^T\widehat{\mathbf{R}}_{d,J}\widehat{\mathbf{L}}^{-1}\widehat{\mathbf{D}}^{-1/2}\widehat{\mathbf{P}}^T\widehat{\mathbf{R}}_{d,J}^T\widehat{\mathbf{A}}_{d,J}^{\mathrm{SG}}\widehat{\mathbf{R}}_{d,J}\widehat{\mathbf{P}}\widehat{\mathbf{D}}^{-1/2}\widehat{\mathbf{L}}^{-T}\widehat{\mathbf{R}}_{d,J}^T\widehat{\mathbf{z}}_{d,J}^{\mathrm{SG}} \simeq (\widehat{\mathbf{z}}_{d,J}^{\mathrm{SG}})^T\widehat{\mathbf{z}}_{d,J}^{\mathrm{SG}}$$

on

$$\widehat{\mathbf{z}}_{d,J}^{\mathrm{SG}} \in \operatorname{im}\widehat{\mathbf{R}}_{d,J}\widehat{\mathbf{R}}_{d,J}^T\widehat{\mathbf{R}}_{d,J}\widehat{\mathbf{L}}^T\widehat{\mathbf{D}}^{1/2}\widehat{\mathbf{P}} = \operatorname{im}\widehat{\mathbf{R}}_{d,J}\widehat{\mathbf{L}}^T\widehat{\mathbf{D}}^{1/2}\widehat{\mathbf{P}} \subset \mathbb{R}^{\widehat{N}_{d,J}^{\mathrm{SG}}} .$$

It is left to show that vectors $\widehat{\mathbf{v}}_{d,J}^{\mathrm{SG}}$ with $\widehat{\mathbf{v}}_{d,J}^{\mathrm{SG}} \perp \widehat{\mathbf{z}}_{d,J}^{\mathrm{SG}}$ are indeed in the kernel of (48). We obtain this by

$$\begin{aligned}\widehat{\mathbf{v}}_{d,J}^{\mathrm{SG}} \in \ker(\widehat{\mathbf{R}}_{d,J}\widehat{\mathbf{L}}^T\widehat{\mathbf{D}}^{1/2}\widehat{\mathbf{P}})^T &= \ker\widehat{\mathbf{P}}^T\widehat{\mathbf{D}}^{1/2}\widehat{\mathbf{L}}\widehat{\mathbf{R}}_{d,J}^T = \ker\widehat{\mathbf{D}}^{1/2}\widehat{\mathbf{G}}\widehat{\mathbf{P}}\widehat{\mathbf{G}}^{-1}\widehat{\mathbf{L}}\widehat{\mathbf{R}}_{d,J}^T \\ &= \ker\widehat{\mathbf{D}}^{1/2}\widehat{\mathbf{G}}\widehat{\mathbf{P}}\widehat{\mathbf{L}}^{-T}\widehat{\mathbf{R}}_{d,J}^T = \ker\widehat{\mathbf{P}}\widehat{\mathbf{L}}^{-T}\widehat{\mathbf{R}}_{d,J}^T ,\end{aligned}$$

where we have used similar arguments as in the proof of Theorem 2. Altogether, this proves that the matrix (48) has a generalized condition number that is at least as good as that for the full grid case, i.e. that of $\mathbf{C}_{d,J}\mathbf{A}_{d,J}$. Finally, we can rewrite the preconditioner in the form (49) since

$$
\begin{aligned}
&\tilde{\kappa}(\widehat{\mathbf{R}}_{d,J}\widehat{\mathbf{L}}^{-1}\widehat{\mathbf{P}}^T\widehat{\mathbf{D}}^{-1/2}\widehat{\mathbf{R}}_{d,J}^T\widehat{\mathbf{A}}_{d,J}^{\mathrm{SG}}\widehat{\mathbf{R}}_{d,J}\widehat{\mathbf{D}}^{-1/2}\widehat{\mathbf{P}}\widehat{\mathbf{L}}^{-T}\widehat{\mathbf{R}}_{d,J}^T)\\
&= \tilde{\kappa}(\widehat{\mathbf{R}}_{d,J}\widehat{\mathbf{D}}^{-1/2}\widehat{\mathbf{P}}\widehat{\mathbf{L}}^{-T}\widehat{\mathbf{R}}_{d,J}^T\widehat{\mathbf{R}}_{d,J}\widehat{\mathbf{L}}^{-1}\widehat{\mathbf{P}}^T\widehat{\mathbf{D}}^{-1/2}\widehat{\mathbf{R}}_{d,J}^T\widehat{\mathbf{A}}_{d,J}^{\mathrm{SG}}) \qquad (51)\\
&= \tilde{\kappa}(\widehat{\mathbf{R}}_{d,J}\widehat{\mathbf{P}}\widehat{\mathbf{D}}^{-1}\widehat{\mathbf{G}}^{-1}\widehat{\mathbf{P}}^T\widehat{\mathbf{R}}_{d,J}^T\widehat{\mathbf{A}}_{d,J}^{\mathrm{SG}}) = \tilde{\kappa}(\widehat{\mathbf{C}}_{d,J}^{\mathrm{SG}}\widehat{\mathbf{A}}_{d,J}^{\mathrm{SG}})\,. \qquad (52)
\end{aligned}
$$

In (51), we used that $\tilde{\kappa}(\mathbf{EF}) = \tilde{\kappa}(\mathbf{FE})$ for arbitrary square matrices $\mathbf{E}$ and $\mathbf{F}$, and in (52) we used that $\widehat{\mathbf{R}}_{d,J}^T\widehat{\mathbf{R}}_{d,J}$ is the identity on $\operatorname{im}\widehat{\mathbf{R}}_{d,J}^T$, that $\widehat{\mathbf{L}}^{-T}\widehat{\mathbf{L}}^{-1} = \widehat{\mathbf{G}}^{-1}$ and, finally, that block-diagonal scaling matrices commute with block-diagonal matrices, see Lemma 4. This proves the theorem. □

4.3 Cost Discussion

It is of course important not to implement the matrix $\widehat{\mathbf{A}}_{d,J}^{\mathrm{SG}}$ nor the preconditioner $\widehat{\mathbf{C}}_{d,J}^{\mathrm{SG}}$ from (49) naively, if we want to keep their computational costs proportional to the number of degrees of freedom $\widehat{N}_{d,J}^{\mathrm{SG}}$ of the sparse grid. First, let us consider $\widehat{\mathbf{C}}_{d,J}^{\mathrm{SG}}$ in more detail. In fact, all the matrices $\widehat{\mathbf{P}}$, $\widehat{\mathbf{D}}^{-1}$, $\widehat{\mathbf{G}}^{-1}$ and $\widehat{\mathbf{P}}^T$ are block-diagonal, which means that they can be implemented as subspace-wise operations. As for $\widehat{\mathbf{x}}_{d,J} \in \operatorname{im}\widehat{\mathbf{R}}_{d,J}^T$ all vector blocks $(\widehat{\mathbf{x}}_{d,J})_{\mathbf{l}}$ with $\mathbf{l} \in \mathscr{F}_J^d \setminus \mathscr{S}_J^d$ are zero and get removed by the final application of $\widehat{\mathbf{R}}_{d,J}$ anyway, they do not need to be considered in the implementation. By the same arguments as in Subsect. 3.3, we obtain that our preconditioner can indeed be applied in $\mathscr{O}(d \cdot \widehat{N}_{d,J}^{\mathrm{SG}})$.

An efficient matrix-vector multiplication with the operator matrix $\widehat{\mathbf{A}}_{d,J}^{\mathrm{SG}}$ is however far more complicated than in the full grid case. One reason is that the index $\mathscr{S}_J^d$ has, unlike $\mathscr{F}_J^d$, no representation as a Cartesian product, which means that $\widehat{\mathbf{A}}_{d,J}^{\mathrm{SG}}$ has no Kronecker product structure like $\widehat{\mathbf{A}}_{d,J}$ does. Of course, we must not use (47), which was given only for theoretical reasons to allow for the submatrix argument of the last subsection. Instead, we resort to a quite sophisticated dimension-recursive algorithm based on the so-called unidirectional principle [BZ96, Bun92b] to perform the matrix-vector-multiplication linearly in the number of degrees of freedom $\widehat{N}_{d,J}^{\mathrm{SG}}$. Typically the associated dimension-dependent constant in the costs is proportional to 2^d. This factor can however be reduced to d^2 in the case of the Laplacian by exploiting the L^2-orthogonality between subspaces, see [Feu05]. Then, the total algorithmic cost of one application of $\widehat{\mathbf{C}}_{d,J}^{\mathrm{SG}}\widehat{\mathbf{A}}_{d,J}^{\mathrm{SG}}$ is of the order $\mathscr{O}(d^2 \cdot \widehat{N}_{d,J}^{\mathrm{SG}})$.

So far, we have expressed the computational effort with respect to the enlarged sparse grid system with $\widehat{N}_{d,J}^{\mathrm{SG}}$ degrees of freedom, which has by a factor of about 2^d more degrees of freedom than the hierarchical basis [BG04]. We consider this acceptable, because the number of degrees of freedom of regular sparse grids is of the order $\mathscr{O}(2^J J^{d-1})$, which is exponential in d anyway. Moreover, the number of degrees of freedom of energy sparse grids is of the order $\mathscr{O}(2^J)$ in J, but which involves a constant that is also exponential in d, cf. [Gri06].

Note that it is possible to remove the redundancy of our multilevel discretization via the generating system (45) by using a prewavelet discretization. This seems to eliminate the 2^d-factor by construction. It however still appears hidden in the setup of the discrete right-hand side for general functions f. The prewavelet approach and a new improved preconditioner will be discussed in the next section.

5 Prewavelets

The enlarged generating system introduced some additional difficulties like a non-trivial kernel of the operator matrix and the need for an orthogonalization operator $\widehat{\mathbf{P}}$. This can be avoided in the first place if a direct discretization of the orthogonal subspaces $W_{\mathbf{l}}$ is available, which is just the case for so-called prewavelets and for wavelets.

Let us first consider the full grid case. To this end, let us assume that we have basis functions $(\psi_{\mathbf{l},\mathbf{i}})_{\mathbf{i}\in\xi_{\mathbf{l}},\mathbf{l}\in\mathscr{F}_J^d}$ with

$$W_{\mathbf{l}} = \operatorname{span}\{\psi_{\mathbf{l},\mathbf{i}} : \mathbf{i} \in \xi_{\mathbf{l}}\} \quad \text{for} \quad \mathbf{l} \in \mathscr{F}_J^d , \tag{53}$$

and set $\bar{n}_{\mathbf{l}} := |\xi_{\mathbf{l}}|$. Note here that we have L^2-orthogonality between different levels by definition, but we have not necessarily L^2-orthogonality within one level. The multilevel matrix $\bar{\mathbf{A}}_{d,J} \in \mathbb{R}^{N_{d,J}\times N_{d,J}}$ with

$$(\bar{\mathbf{A}}_{d,J})_{(\mathbf{l},\mathbf{i}),(\mathbf{k},\mathbf{j})} = a(\psi_{\mathbf{l},\mathbf{i}}, \psi_{\mathbf{k},\mathbf{j}}) \quad \text{for} \quad \mathbf{i} \in \xi_{\mathbf{l}}, \mathbf{j} \in \xi_{\mathbf{k}}, \mathbf{l}, \mathbf{k} \in \mathscr{F}_J^d \tag{54}$$

results just from the system matrix $\mathbf{A}_{d,J}$ of (7) by a change of the basis, since

$$\bigoplus_{\mathbf{l}\in\mathscr{F}_J^d} W_{\mathbf{l}} = V_J^d .$$

Thus,

$$\bar{\mathbf{A}}_{d,J} = \mathbf{T}^T \mathbf{A}_{d,J} \mathbf{T} ,$$

where $\mathbf{T}$ maps from $\{\psi_{\mathbf{l},\mathbf{i}}\}_{\mathbf{i}\in\xi_{\mathbf{l}},\mathbf{l}\in\mathscr{F}_J^d}$ to $\{\phi_{\mathbf{J},\mathbf{i}}\}_{\mathbf{i}\in\chi_{\mathbf{J}}}$. The analogue holds for the right-hand side $\bar{\mathbf{b}}_{d,J} \in \mathbb{R}^{N_{d,J}}$ with

$$(\bar{\mathbf{b}}_{d,J})_{\mathbf{l},\mathbf{i}} = (f, \psi_{\mathbf{l},\mathbf{i}})_{L^2(\Omega^d)}, \mathbf{i} \in \xi_{\mathbf{l}}, \mathbf{l} \in \mathscr{F}_J^d , \quad \text{i.e.} \quad \bar{\mathbf{b}}_{d,J} = \mathbf{T}^T \mathbf{b}_{d,J} .$$

Note here that, in contrast to $\widehat{\mathbf{A}}_{d,J}$, the system matrix $\bar{\mathbf{A}}_{d,J}$ is now invertible, since the functions in (53) form a basis of $V_{\mathscr{F}_J^d}$.

5.1 Preconditioner

Now, we will present our preconditioner for prewavelets and discuss its resulting condition number. To this end, we need a diagonal scaling matrix $\bar{\mathbf{D}} \in \mathbb{R}^{N_{d,J} \times N_{d,J}}$ with blocks $(\bar{\mathbf{D}})_{\mathbf{l},\mathbf{k}} \in \mathbb{R}^{\bar{n}_\mathbf{l} \times \bar{n}_\mathbf{k}}$ and

$$(\bar{\mathbf{D}})_{\mathbf{l},\mathbf{k}} = \begin{cases} (\sum_{p=1}^{d} 2^{2l_p})\bar{\mathbf{I}}_\mathbf{l} & \text{for} \quad \mathbf{l} = \mathbf{k}\,, \\ 0 & \text{else}\,, \end{cases} \tag{55}$$

where the $\bar{\mathbf{I}}_\mathbf{l} \in \mathbb{R}^{\bar{n}_\mathbf{l} \times \bar{n}_\mathbf{l}}$ denote identity matrices on the subspaces. Furthermore, we need the subspace-wise mass matrix $\bar{\mathbf{G}} \in \mathbb{R}^{N_{d,J} \times N_{d,J}}$ with blocks $(\bar{\mathbf{G}})_{\mathbf{l},\mathbf{k}} \in \mathbb{R}^{\bar{n}_\mathbf{l} \times \bar{n}_\mathbf{k}}$, where

$$(\bar{\mathbf{G}})_{\mathbf{l},\mathbf{k}} = \begin{cases} \bar{\mathbf{M}}_\mathbf{l} & \text{for} \quad \mathbf{l} = \mathbf{k}\,, \\ 0 & \text{else}\,. \end{cases}$$

Here, $\bar{\mathbf{M}}_\mathbf{l} \in \mathbb{R}^{\bar{n}_\mathbf{l} \times \bar{n}_\mathbf{l}}$ denotes the mass matrix

$$(\bar{\mathbf{M}}_\mathbf{l})_{\mathbf{i},\mathbf{j}} = (\psi_{\mathbf{l},\mathbf{i}}, \psi_{\mathbf{l},\mathbf{j}}) \quad \text{for} \quad \mathbf{i},\mathbf{j} \in \xi_\mathbf{l}\,.$$

Then, we have the following theorem:

Theorem 4. *The condition number of*

$$\bar{\mathbf{D}}^{-1}\bar{\mathbf{G}}^{-1}\bar{\mathbf{A}}_{d,J} \tag{56}$$

is bounded asymptotically with respect to J and is completely independent of the dimension d.

Proof. We translate the norm equivalence (22) into the matrix-vector setting for $\bar{\mathbf{x}}_{d,J} \in \mathbb{R}^{N_{d,J} \times N_{d,J}}$ and obtain

$$\begin{aligned} \bar{\mathbf{x}}_{d,J}^T \bar{\mathbf{A}}_{d,J} \bar{\mathbf{x}}_{d,J} &= a\big(\sum_{\mathbf{l}\in\mathscr{F}_J^d}\sum_{\mathbf{i}\in\xi_\mathbf{l}} \bar{x}_{\mathbf{l},\mathbf{i}}\psi_{\mathbf{l},\mathbf{i}}, \sum_{\mathbf{l}\in\mathscr{F}_J^d}\sum_{\mathbf{i}\in\xi_\mathbf{l}} \bar{x}_{\mathbf{l},\mathbf{i}}\psi_{\mathbf{l},\mathbf{i}}\big) \\ &\simeq \sum_{\mathbf{l}\in\mathscr{F}_J^d} \Big(\sum_{p=1}^{d} 2^{2l_p}\Big) \Big\| \sum_{\mathbf{i}\in\xi_\mathbf{l}} \bar{x}_{\mathbf{l},\mathbf{i}}\psi_{\mathbf{l},\mathbf{i}} \Big\|_{L^2(\Omega^d)}^2 \\ &= \sum_{\mathbf{l}\in\mathscr{F}_J^d} \Big(\sum_{p=1}^{d} 2^{2l_p}\Big) \bar{\mathbf{x}}_\mathbf{l}^T \bar{\mathbf{M}}_\mathbf{l} \bar{\mathbf{x}}_\mathbf{l} \\ &= \bar{\mathbf{x}}_{d,J}^T \bar{\mathbf{D}}\bar{\mathbf{G}}\bar{\mathbf{x}}_{d,J}\,. \end{aligned} \tag{57}$$

From Theorem 1 we know that the constants in (57) are independent of the dimension d and bounded independently of J. This concludes the proof. □

In the case of a regular sparse grid with $\mathscr{I} = \mathscr{S}_J^d$, the equality

$$\bigoplus_{\mathbf{l}\in\mathscr{S}_J^d} W_\mathbf{l} = \sum_{\mathbf{l}\in\mathscr{S}_J^d} V_\mathbf{l} \tag{58}$$

holds. The analogue is valid for a general sparse grid with any arbitrary index set $\mathscr{I}$ for which the condition (44) is satisfied. Thus, the regular sparse grid space $\sum_{\mathbf{l}\in\mathscr{S}_J^d} V_\mathbf{l}$ or the general sparse grid space $\sum_{\mathbf{l}\in\mathscr{I}} V_\mathbf{l}$ can both be expressed by the left-hand side of (58), i.e. with the help of $W_\mathbf{l}$-subspaces. The resulting linear system matrix $\bar{\mathbf{A}}_{d,J}^{\mathrm{SG}}$ is just a submatrix of the full matrix[6] $\bar{\mathbf{A}}_{d,J}$. Consequently, the condition number for the sparse grid system is at least as good as the one of (56). Analogously the resulting right-hand side $\bar{\mathbf{b}}_{d,J}^{\mathrm{SG}}$ is just a subvector of $\bar{\mathbf{b}}_{d,J}$.

Note here that prewavelets have been frequently used in the past as the basis functions of sparse grid discretizations [GO95, Feu10] but mostly no special attention was paid to the dependence of the condition number on the dimension. A simple Jacobi-diagonal scaling of $\bar{\mathbf{A}}_{d,J}$ is equivalent to replacing the subspace-wise inversion of the mass matrices $\bar{\mathbf{G}}^{-1}$ in (56) by the identity and the $\bar{\mathbf{D}}$ from (55) by $\mathrm{diag}(\bar{\mathbf{A}}_{d,J})$. This does not affect the asymptotics in J for L^2-stable basis functions, but the condition number of the operator matrix grows now exponentially with the dimension [Feu05]. Sometimes $(\bar{\mathbf{D}})_{\mathbf{l},\mathbf{l}} = 2^{2|\mathbf{l}|_\infty}\,\mathrm{diag}(\bar{\mathbf{M}}_\mathbf{l})$ is chosen, see [GO94, GK09], which also results in condition numbers that grow with the dimension d.

5.2 Cost Discussion

We now have the preconditioner $\bar{\mathbf{C}}_{d,J} := \bar{\mathbf{D}}^{-1}\bar{\mathbf{G}}^{-1}$ for $\bar{\mathbf{A}}_{d,J}$ in prewavelet discretization. At first sight, this approach looks simpler and more efficient than the more complicated discretizations $\widehat{\mathbf{A}}_{d,J}$ and $\widehat{\mathbf{A}}_{d,J}^{\mathrm{SG}}$ using the enlarged generating system and the associated preconditioners $\widehat{\mathbf{C}}_{d,J}$ and $\widehat{\mathbf{C}}_{d,J}^{\mathrm{SG}}$, respectively. This is due to the fact that the prewavelet system $\{\psi_{\mathbf{l},\mathbf{i}} : \mathbf{i}\in\xi_\mathbf{l}\}_{\mathbf{l}\in\mathscr{I}}$ forms a basis and therefore exhibits no redundancies. Thus, by a factor of about 2^d less degrees of freedom are involved than for the corresponding generating system.

However, there are additional difficulties to be faced in the prewavelet approach, which should not be underestimated and may give the generating system method a practical advantage. First, prewavelets are less local than, e.g. the corresponding multilevel spline basis. Thus, the mass matrix inversions in $\bar{\mathbf{G}}^{-1}$ become more involved. From a programming perspective, the more complicated basis functions and different types of prewavelet functions near the boundary make the application of the matrix $\bar{\mathbf{A}}_{d,J}$ to a vector more difficult. The efficient application of the action of the sparse grid system matrix $\bar{\mathbf{A}}_{d,J}^{\mathrm{SG}}$ onto a vector is even more involved, since the unidirectional principle strongly relies on the nestedness of the subspaces. If this is no longer the case, the one-dimensional operators have to be tailored to the

[6] Here, the level J of the full grid is to be equal to the level J of the regular sparse grid or equal to the finest level involved in $\mathscr{I}$ for the general sparse grid.

specific discretization [Feu10] or the algorithm must switch to a generating system anyway [Zei11].

Finally, as already mentioned at the end of Subsect. 4.3, the cost complexity of the setup of the right-hand side $\bar{\mathbf{b}}^{\mathrm{SG}}_{d,J}$ also has at least a 2^d-factor if the corresponding integrations are realized by the interpolation of the function f from (1) in our prewavelet sparse grid space and a subsequent multiplication by the mass matrix to account for the necessary numerical quadrature. As stated in [Feu05], for general functions f, this approach requires the inclusion of boundary functions in the interpolation step (even if the solution u of our Poisson problem has homogeneous boundary conditions). Since the d-dimensional hypercube Ω^d has 2^d faces, an additional factor of the order 2^d enters the cost complexity for the setup of the right-hand side. The dependence of the cost complexity on the dimension d of other techniques for the assembly of the right-hand side for wavelets and prewavelets with sufficient accuracy, e.g. by the solution of an eigenvector-moment problem associated with the coefficients of the refinement equation [DM93], is unknown to us. We however believe that also these methods involve a factor of at least 2^d in the d-dimensional case due to the tensor product construction.

Altogether, the generating system approach from (36) and (48) can be seen as a simple form of implementation of the prewavelet approach and, indeed, both methods give exactly the same condition numbers for the piecewise linear case.

6 Numerical Experiments

Now, we give the results of numerical experiments for our new full and sparse grid preconditioners (42) and (49), respectively. We consider the d-dimensional Laplace operator on the domain $\Omega^d = (0,1)^d$ with vanishing Dirichlet boundary conditions. As locally supported basis functions in (4), we choose on level l the $n_l = 2^l - 1$ hat functions

$$\phi_{l,i}(x) = \max(1 - 2^l \left|x - x_{l,i}\right|, 0) \,,$$

which are centered at the points of an equidistant mesh

$$x_{l,i} = 2^{-l} i$$

for $i = 1, \ldots, n_l$. The resulting space $\cup_{l=1}^{\infty} V_l$ is indeed equal to the underlying Sobolev space $H^1_0(\Omega)$ up to completion with the H^1-norm, see [BG04].

Table 1 shows the generalized condition numbers of the preconditioned matrices (36) and (48) of the enlarged generating systems in the full and sparse grid case for different dimensions d and levels J. We clearly observe that the full grid condition numbers are bounded from above by a constant independently of the level J. Moreover, they are perfectly independent of the dimension as our theory suggests. The sparse grid condition numbers are even smaller than the corresponding full grid ones for $d > 1$, which is in accordance with our submatrix argument from Theorem 3. In Table 2 (top) we give the number of degrees of freedom $\widehat{N}^{\mathrm{SG}}_{d,J}$ for various values of J and d. Finally, Table 2 (bottom) reveals that the condition numbers of the sparse grid even *decrease* with rising dimension d for a fixed level J. For a sparse grid discretization of the Poisson problem, these numbers clearly show that we are altogether able to efficiently solve the associated linear systems of equations for both, quite large values of J and larger dimensions d.

Table 1 Degrees of freedom $\widehat{N}_{d,J}$ and $\widehat{N}_{d,J}^{\mathrm{SG}}$ and condition numbers $\tilde{\kappa}$ of the preconditioned matrices (36) and (48) for the Laplacian on the unit hypercube with a full- and sparse-grid discretization for the generating system approach based on linear splines

		DOFs $\widehat{N}_{d,J}$ and $\widehat{N}_{d,J}^{\mathrm{SG}}$		condition number $\tilde{\kappa}$	
	level J	full grid	sparse grid	full grid	sparse grid
dim = 1	2	4	4	3.40	3.40
	3	11	11	4.67	4.67
	4	26	26	5.17	5.17
	5	57	57	5.84	5.84
	6	120	120	6.37	6.37
	7	247	247	6.80	6.80
	8	502	502	7.16	7.16
	9	1013	1013	7.47	7.47
	10	2036	2036	7.74	7.74
	11	4083	4083	7.96	7.96
	12	8178	8178	8.16	8.16
	13	16369	16369	8.33	8.33
dim = 2	2	16	7	3.40	2.99
	3	121	30	4.67	4.46
	4	676	102	5.17	5.06
	5	3249	303	5.84	5.65
	6	14400	825	6.37	6.20
dim = 3	2	64	10	3.40	2.71
	3	1331	58	4.67	4.28
	4	17576	256	5.17	5.00
dim = 4	2	256	13	3.40	2.51
	3	14641	95	4.67	4.12
dim = 5	2	1024	16	3.40	2.36

Finally note that the prewavelet approach from Sect. 5 results in exactly the same condition numbers.

7 Concluding Remarks

We presented preconditioners $\mathbf{C}_{d,J}$ (42) and $\widehat{\mathbf{C}}_{d,J}^{\mathrm{SG}}$ (49) for an isotropic full grid and an enlarged sparse grid system, respectively. Both result in condition numbers that are bounded independently of the level J and are constant or, in the sparse grid case, even decreasing for rising dimension d.

The computational costs of the preconditioner remained linear in the number of degrees of freedom. The size of the enlarged systems grows by a factor of about 2^d

Table 2 Degrees of freedom and generalized condition numbers of the preconditioned sparse grid system (48) for different dimensions d and levels J

	degrees of freedom with respect to the level J											
dim	2	3	4	5	6	7	8	9	10	11	12	13
1	4	11	26	57	120	247	502	1013	2036	4083	8178	16369
2	7	30	102	303	825	2116	5200	12381				
3	10	58	256	955	3178	9740						
4	13	95	515	2310	9078							
5	16	141	906	4746								
6	19	196	1456	8722								
7	22	260	2192	14778								
8	25	333	3141									
9	28	415	4330									
10	31	506	5786									

	condition number with respect to the level J											
dim	2	3	4	5	6	7	8	9	10	11	12	13
1	3.40	4.67	5.17	5.84	6.37	6.80	7.16	7.47	7.74	7.96	8.16	8.33
2	2.99	4.46	5.06	5.65	6.20	6.65	7.04	7.36				
3	2.71	4.28	5.00	5.49	6.06	6.53						
4	2.51	4.12	4.94	5.35	5.95							
5	2.36	3.97	4.88	5.23								
6	2.24	3.83	4.82	5.17								
7	2.15	3.71	4.77	5.15								
8	2.07	3.60	4.71									
9	2.00	3.50	4.66									
10	1.94	3.41	4.61									

compared to the corresponding basis. This seems a fair price to pay. Better constants in the respective norm equivalence and associated condition numbers reduce the number of iterations of Krylov methods and also make corresponding residual-based error estimates more reliable.

Our new preconditioners can be applied to differential operators other than the Laplacian. The approach works straightforwardly for constant coefficients or variable coefficients which are separable, i.e. can be written as a product of one-dimensional diffusion functions, or can be well approximated by a low-rank representation. But then the equivalence constants of

$$a(u,u) \simeq \|u\|^2_{H^1(\Omega^d)} ,$$

i.e. the ellipticity constants, cf. [CS12], at least partly enter the condition estimate of the resulting system. If they grow exponentially with the dimension d we may run into problems, though.

In a similar way, equivalences to H^s-norms can be dealt with by using diagonal scaling matrices

$$(\widehat{\mathbf{D}})_{\mathbf{l},\mathbf{k}} = \begin{cases} (\sum_{p=1}^{d} 2^{2sl_p})\mathbf{I}_{\mathbf{l}} & \text{for} \quad \mathbf{l} = \mathbf{k} \,, \\ 0 & \text{else} \,, \end{cases}$$

and the associated $\widehat{\mathbf{G}}^{-1}$ if the regularity of the employed basis functions is sufficient.

References

[All92] G. Allaire. Homogenization and two-scale convergence. *SIAM Journal on Mathematical Analysis*, 23(6):1482–1518, 1992.

[Bal97] R. Balescu. *Statistical Dynamics: Matter Out of Equilibrium*. Imperial College Press, 1997.

[Bel61] R. Bellman. *Adaptive Control Processes: A Guided Tour*. Princeton University Press, 1961.

[BG99] H. Bungartz and M. Griebel. A note on the complexity of solving Poisson's equation for spaces of bounded mixed derivatives. *J. Complexity*, 15:167–199, 1999.

[BG04] H. Bungartz and M. Griebel. Sparse grids. *Acta Numerica*, 13:1–123, 2004.

[BGGK12] O. Bokanowski, J. Garcke, M. Griebel, and I. Klompmaker. An adaptive sparse grid semi-Lagrangian scheme for first order Hamilton-Jacobi Bellman equations. *J. of Scientific Computing*, 2012.

[BL11] A. Brandt and O. Livne. *Multigrid Techniques: 1984 Guide with Applications to Fluid Dynamics*. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics, 2011.

[BNT10] I. Babuška, F. Nobile, and R. Tempone. A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM Review*, 52(2):317–355, 2010.

[BNTT11] J. Bäck, F. Nobile, L. Tamellini, and R. Tempone. Stochastic spectral Galerkin and collocation methods for PDEs with random coefficients: A numerical comparison. In J. Hesthaven and E. Rønquist, editors, *Spectral and High Order Methods for Partial Differential Equations*, volume 76 of *Lecture Notes in Computational Science and Engineering*, pages 43–62. Springer Berlin Heidelberg, 2011.

[BP94] A. Berman and R. Plemmons. *Nonnegative Matrices in the Mathematical Sciences*. Society for Industrial and Applied Mathematics, 1994.

[BPX90] J. Bramble, J. Pasciak, and J. Xu. Parallel multilevel preconditioners. *Math. Comp.*, 55(191):1–22, 1990.

[Bra07] D. Braess. *Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics*. Cambridge University Press, 2007.

[Bun92a] H. Bungartz. An adaptive Poisson solver using hierarchical bases and sparse grids in iterative methods in linear algebra. In P. de Groen and R. Beauwens, editors, *Proceedings of the IMACS International Symposium, Brussels, 2.-4. 4. 1991*, pages 293–310, Amsterdam, 1992. Elsevier.

[Bun92b] H. Bungartz. *Dünne Gitter und deren Anwendung bei der adaptiven Lösung der dreidimensionalen Poisson-Gleichung*. Dissertation, Fakultät für Informatik, Technische Universität München, 1992.

[BZ96] R. Balder and C. Zenger. The solution of multidimensional real Helmholtz equations on sparse grids. *SIAM J. Sci. Comput.*, 17:631–646, 1996.

[CDG02] D. Cioranescu, A. Damlamian, and G. Griso. Periodic unfolding and homogenization. *Comptes Rendus Mathematique*, 335(1):99 – 104, 2002.

[CDS11] A. Cohen, R. DeVore, and C. Schwab. Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDEs. *Anal. Appl.*, 9:11–47, 2011.

[CS12] N. Chegini and R. Stevenson. The adaptive tensor product wavelet scheme: sparse matrices and the application to singularly perturbed problems. *IMA Journal of Numerical Analysis*, 32(1):75–104, 2012.

[Dah96] W. Dahmen. Stability of multiscale transformations. *J. Fourier Anal. Appl*, 2:341–361, 1996.

[DM93] W. Dahmen and C. Micchelli. Using the refinement equation for evaluating integrals of wavelets. *SIAM Journal on Numerical Analysis*, 30(2):507–537, 1993.

[Feu05] C. Feuersänger. Dünngitterverfahren für hochdimensionale elliptische partielle Differentialgleichungen. Diplomarbeit, Institut für Numerische Simulation, Universität Bonn, 2005.

[Feu10] C. Feuersänger. *Sparse Grid Methods for Higher Dimensional Approximation*. Dissertation, Institut für Numerische Simulation, Universität Bonn, 2010.

[Gar04] J. Garcke. *Maschinelles Lernen durch Funktionsrekonstruktion mit verallgemeinerten dünnen Gittern*. Doktorarbeit, Institut für Numerische Simulation, Universität Bonn, 2004.

[GG03] T. Gerstner and M. Griebel. Dimension-adaptive tensor-product quadrature. *Computing*, 71(1):65–87, 2003.

[GGT01] J. Garcke, M. Griebel, and M. Thess. Data mining with sparse grids. *Computing*, 67(3):225–253, 2001.

[GJP95] F. Girosi, M. Jones, and T. Poggio. Regularization theory and neural networks architectures. *Neural Computation*, 7:219–269, 1995.

[GK00] M. Griebel and S. Knapek. Optimized tensor-product approximation spaces. *Constructive Approximation*, 16(4):525–540, 2000.

[GK09] M. Griebel and S. Knapek. Optimized general sparse grid approximation spaces for operator equations. *Mathematics of Computations*, 78(268):2223–2257, 2009.

[GO94] M. Griebel and P. Oswald. On additive Schwarz preconditioners for sparse grid discretizations. *Numer. Math.*, 66:449–464, 1994.

[GO95] M. Griebel and P. Oswald. Tensor product type subspace splitting and multilevel iterative methods for anisotropic problems. *Adv. Comput. Math.*, 4:171–206, 1995.

[Gri94a] M. Griebel. Multilevel algorithms considered as iterative methods on semidefinite systems. *SIAM Int. J. Sci. Stat. Comput.*, 15(3):547–565, 1994.

[Gri94b] M. Griebel. *Multilevelmethoden als Iterationsverfahren über Erzeugendensystemen*. Teubner Skripten zur Numerik. Teubner, Stuttgart, 1994.

[Gri06] M. Griebel. Sparse grids and related approximation schemes for higher dimensional problems. In L. Pardo, A. Pinkus, E. Suli, and M.J. Todd, editors, *Foundations of Computational Mathematics (FoCM05), Santander*, pages 106–161. Cambridge University Press, 2006.

[Hac85] W. Hackbusch. *Multi-Grid Methods and Applications*. Springer Series in Computational Mathematics. Springer, 1985.

[Ham09] Jan Hamaekers. *Tensor Product Multiscale Many–Particle Spaces with Finite–Order Weights for the Electronic Schödinger Equation*. Dissertation, Institut für Numerische Simulation, Universität Bonn, 2009.

[Heg03] M. Hegland. Adaptive sparse grids. In K. Burrage and Roger B. Sidje, editors, *Proc. of 10th Computational Techniques and Applications Conference CTAC-2001*, volume 44, pages C335–C353, 2003.

[HS05] V. Hoang and C. Schwab. High-dimensional finite elements for elliptic problems with multiple scales. *Multiscale Modeling & Simulation*, 3(1):168–194, 2005.

[HSS08] H. Harbrecht, R. Schneider, and C. Schwab. Sparse second moment analysis for elliptic problems in stochastic domains. *Numer. Math.*, 109:385–414, 2008.

[JR08] J. Jakeman and S. Roberts. Stochastic Galerkin and collocation methods for quantifying uncertainty in differential equations: a review. *ANZIAM J.*, 50((C)):C815–C830, 2008.

[Kna00] S. Knapek. *Approximation und Kompression mit Tensorprodukt-Multiskalenräumen*. Doktorarbeit, Universität Bonn, 2000.

[Kwo08] Y. Kwok. *Mathematical Models of Financial Derivatives*. Springer Finance. Springer London, 2008.

[Mat02] A. Matache. Sparse two-scale FEM for homogenization problems. *Journal of Scientific Computing*, 17:659–669, 2002.

[Mes65] A. Messiah. *Quantum mechanics*. North-Holland, 1965.

[Mit97] U. Mitzlaff. *Diffusionsapproximation von Warteschlangensystemen*. Doktorarbeit, TU Clausthal, 1997.

[MK10] O. Maître and O. Knio. *Spectral Methods for Uncertainty Quantification: With Applications to Computational Fluid Dynamics*. Scientific Computation. Springer, 2010.

[Mun00] R. Munos. A study of reinforcement learning in the continuous case by the means of viscosity solutions. *Mach. Learn.*, 40(3):265–299, 2000.

[NTW08a] F. Nobile, R. Tempone, and C. Webster. An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM J. Numer. Anal.*, 46(5):2411–2442, 2008.

[NTW08b] F. Nobile, R. Tempone, and C. Webster. A sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM J. Numer. Anal.*, 46(5):2309–2345, 2008.

[Osw92] P. Oswald. On discrete norm estimates related to multilevel preconditioners in the finite element method. In *Constructive Theory of Functions, Proc. Int. Conf. Varna 1991*, pages 203–214. Bulg. Acad. Sci., Sofia, 1992.

[Osw94] P. Oswald. *Multilevel Finite Element Approximation: Theory and Applications*. Teubner Skripten zur Numerik. Teubner, 1994.

[Rei04] C. Reisinger. *Numerische Methoden für hochdimensionale parabolische Gleichungen am Beispiel von Optionspreisaufgaben*. Dissertation, Universität Heidelberg, 2004.

[SB98] R. Sutton and A. Barto. *Reinforcement Learning: An Introduction (Adaptive Computation and Machine Learning)*. The MIT Press, 1998.

[SCDD02] X. Shen, H. Chen, J. Dai, and W. Dai. The finite element method for computing the stationary distribution of an SRBM in a hypercube with applications to finite buffer queueing networks. *Queueing Syst. Theory Appl.*, 42(1):33–62, 2002.

[Sjö07] P. Sjöberg. Partial approximation of the master equation by the Fokker–Planck equation. In B. Kågström, E. Elmroth, J. Dongarra, and J. Waśniewski, editors, *Applied Parallel Computing. State of the Art in Scientific Computing*, volume 4699 of *Lecture Notes in Computer Science*, pages 637–646. Springer Berlin Heidelberg, 2007.

[SLE09] P. Sjöberg, P. Lötstedt, and J. Elf. Fokker–Planck approximation of the master equation in molecular biology. *Computing and Visualization in Science*, 12:37–50, 2009.

[SS01] B. Schölkopf and A. Smola. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. MIT Press, Cambridge, MA, USA, 2001.

[Xu92] J. Xu. Iterative methods by space decomposition and subspace correction. *SIAM Review*, 34(4):pp. 581–613, 1992.

[Yse93] H. Yserentant. Old and new convergence proofs for multigrid methods. *Acta Numerica*, 2:285–326, 1993.

[Zei11] A. Zeiser. Fast matrix-vector multiplication in the sparse-grid Galerkin method. *J. Sci. Comput.*, 47(3):328–346, 2011.