Generalized hierarchical basis multigrid methods for convection-diffusion problems

Michael Griebel* and Frank Kiefer

Institut für Angewandte Mathematik, Abteilung für Wissenschaftliches Rechnen und Numerische Simulation, Universität Bonn, Wegelerstr. 6, D-53115 Bonn, {griebel, kiefer}@iam.uni-bonn.de

SUMMARY

We consider the efficient solution of discrete convection-diffusion problems with dominating convection. It is well known that the standard hierarchical basis multigrid method (HBMG) leads neither to optimal (w.r.t. mesh size) nor to robust solvers for discrete operators arising from singularly perturbed convection-diffusion problems. Its performance is strongly dependent on the coefficients in the differential equation (e.g. strength of convection) *and* the mesh size. (Pre-)Wavelet splittings of the underlying function spaces allow for efficient algorithms which can be viewed as generalized HBMG methods. They can be interpreted as ordinary multigrid methods which employ a special kind of multiscale smoother and show an optimal convergence behavior for the respective non-perturbed equations similar to classical multigrid. Here, we present a general Petrov–Galerkin multiscale approach which makes use of problem-dependent coarsening strategies known from robust multigrid techniques (matrix-dependent prolongations, algebraic coarsening) together with certain (pre-)wavelet-like and hierarchical multiscale decompositions of the trial and test spaces on the finest grid. The presented numerical results show that by this choice generalized HBMG methods can be constructed which result in robust yet efficient solvers.

KEY WORDS: convection-diffusion; Petrov–Galerkin multiscale methods; hierarchical basis; wavelets; robustness; algebraic multigrid

AMS SUBJECT CLASSIFICATIONS: 65F10, 65N30, 65N55

1. Introduction

We consider the efficient solution of large sparse non-symmetric linear systems that arise from discretizations of the following singularly perturbed stationary convection-diffusion problem with homogeneous Dirichlet boundary conditions on the unit-square:

$$Tu := -\Delta u + \vec{b} \cdot \nabla u = f \text{ in } \Omega :=]0,1[^2 \quad \text{and} \quad u = 0 \text{ on } \partial\Omega. \tag{1.1}$$

*Correspondence to: Michael Griebel, Institut für Angewandte Mathematik, Universität Bonn, Wegelerstr. 6, D-53115 Bonn, griebel@iam.uni-bonn.de

Contract/grant sponsor: DFG-Projekt *Mehrgitteralgorithmen für unsymmetrische und heterogene Probleme*, Sonderforschungsbereich 256 Nichtlineare Partielle Differentialgleichungen, Projekt C *Numerik und graphische Darstellung*

We allow for a convective vector field $\vec{b} = (b_1, b_2)$ with large $\mathcal{L}^\infty(\Omega)$-norm and for $\vec{b}$ to vary within a wide range. Problem (1.1) can variationally be formulated as

$$\text{Find } u \in \mathcal{H}_0^1(\Omega): \quad a(u,v) = (f,v)_{\mathcal{L}^2(\Omega)} \quad \forall v \in \mathcal{H}_0^1(\Omega). \tag{1.2}$$

Under suitable assumptions on $\vec{b}$, the bilinear-form $a(u,v) := \int_\Omega \nabla u \cdot \nabla v + b_1(\partial_x u)v + b_2(\partial_y u)v\,dxdy$ is elliptic and the Lax-Milgam lemma [11] thus implies the existence of a unique solution to (1.2) for $f \in \mathcal{H}^{-1}(\Omega)$. The continuous problem leads after some suitable discretization [28, 45], e.g. by a variational Petrov–Galerkin method with trial and test spaces $\mathcal{V}_0$ and $\mathcal{S}_0$, to a large sparse (usually) non-symmetric system of linear equations

$$T_0 u_0 = f_0. \tag{1.3}$$

Often these linear systems are solved by non-symmetric iterative methods [29, 45, 47]. But here, the following two common major difficulties arise: The convergence

- deteriorates with decreasing mesh size of the underlying mesh and
- is strongly dependent on the convective vector field $\vec{b}$.

The hierarchical basis multigrid method (HBMG) [4, 9], which is based on the hierarchical decomposition of the space of (bi-)linear splines, has been proven to be an efficient solver for nicely elliptic problems. However, for convection-diffusion equations it is well known that standard HBMG suffers from both of the above obstructions. It neither leads to an optimal (w.r.t. mesh size) nor to a robust method when applied to the discrete singularly perturbed problems – either directly or as a preconditioner for a non-symmetric iterative solver. Here, the performance is still strongly dependent on the coefficients of the differential equation (e.g. strength of convection) [5]. A main reason for the lack of robustness is that the coarse-grid operators tend to be increasingly unstable when the convective term starts to dominate on coarser scales just like in standard multigrid methods. The hereby computed subspace-"corrections" finally destroy the convergence of the iteration [5, 24, 58]. Furthermore, the criterion of "robust smoothers" from robust multigrid strategies cannot be met since the employed hierarchical smoothers will not degenerate in the limit case of pure convection to a direct solver for each level [28]. Hierarchical smoothers compute corrections only for parts of the unknowns.

Recently there have been several attempts to improve on the robustness of the corresponding algorithm. One approach introduces generalized hierarchical bases which are better suited to the problem at hand [7, 17, 41]. Here, one benefits from an intimate connection between hierarchical bases and an approximate Gaussian elimination [8, 24], which implicitly leads to problem-adapted and physically meaningful coarse-grid operators. However, these methods still suffer from the bad stability properties (now w.r.t. mesh size) of the underlying hierarchical splittings and the conceptual weakness of hierarchical smoothers. From our numerical examples we see that the convergence rates of the corresponding iterations even seem to deteriorate when the perturbation increases.

Another approach constructs stabilized algorithms (w.r.t. mesh size) that use recursively defined polynomial Schur-complement approximations to improve on the coarse-grid corrections of the classical HBMG method [2, 3]. They may be interpreted as HBMG algorithms with a more general cyclic structure than the standard V-cycle. Indeed, one obtains the same effect from a hierarchical W-cycle, which can be implemented with linear complexity and also leads to an optimal method [24, 34].

Finally, (pre-)wavelet splittings of the underlying function spaces allow for efficient algorithms which can be viewed as generalized HBMG methods [34, 39, 55]. They can also be interpreted as ordinary multigrid methods which use a special kind of *multiscale smoother* [34] and show an optimal convergence behavior for the respective non-perturbed equations similar to classical multigrid [53]. From a theoretical point of view this is due to the enhanced stability properties of the corresponding splittings [13, 38]. From the multigrid interpretation the reason is the stronger effect of the respective smoothing iteration compared to the hierarchical one. However, (pre-)wavelet based multiscale smoothers still compute corrections only for parts of the unknowns due to the direct subspace decompositions. Hence, *robust* multiscale smoothers are on principle out of reach.

A means left to construct robust solvers based on direct splittings of the underlying function spaces is the choice of the multiscale decompositions itself. In our approach we start from a rather general Petrov–Galerkin multiscale concept. We apply problem-dependent coarsening strategies known from robust multigrid techniques (matrix-dependent prolongations, algebraic coarsening) [1, 12, 19, 20, 24, 31, 46, 50, 58, 59] together with special (pre-)wavelet-like and hierarchical multiscale decompositions of the trial and test spaces $\mathcal{V}_0$ and $\mathcal{S}_0$ on the finest grid. The main idea to overcome the difficulties mentioned above is to choose one complement space hierarchically (to approximately eliminate the coarse-grid unknowns from the transformed equations for the fine-without-coarse-grid degrees of freedom and thus to obtain physically meaningful coarse-grid operators) and the other one wavelet-like (to stabilize the multiscale setting w.r.t. mesh size and to strengthen the smoother). Our numerical results show that by this choice generalized HBMG methods can be constructed which result in robust yet efficient solvers.

The remainder of the paper is organized as follows. In Section 2 we introduce our general Petrov–Galerkin multiscale setting. After the presentation of the discrete linear system to be solved we first study two-scale methods which result from two-scale decompositions. Then, we shortly explain two equivalent multiscale representations of the stiffness matrix, the *standard-* and *non-standard form* [10], and derive our basic multiplicative multiscale algorithm. In Sections 3 and 4 we discuss problem-dependent multiscale decompositions which are crucial besides a Petrov–Galerkin approach to obtain effective solvers. The subspace splittings from Section 3 are based on the standard geometric coarsening of the underlying fine grid and are therefore called *geometric multiscale decompositions*. After the discussion of the problem-dependent choices of the coarse-grid spaces on the trial and test side we explain their interplay with hierarchical and wavelet-like choices for the corresponding complement spaces. We first consider one-dimensional problems and then show how the one-dimensional techniques carry over to two-dimensional cases using a tensor product approach. In Section 4 we discuss the construction of *algebraic multiscale decompositions* which apply essentially non-stationary subdivisions based on principles from algebraic multigrid (AMG) [46, 50]. AMG is especially suitable for *non-separable* problems where the corresponding operators do not decompose into directional parts, e.g. in case of a circular convection [23]. Here, we propose a wavelet-like stabilization of the obvious AMG-based hierarchical bases which leads to so-called *AMGlets*. These new types of multiscale bases are build in a purely algebraic way reflecting again the philosophy of AMG. In Section 5 we present the numerical results we obtained from our methods for four classes of test problems of increasing difficulty. Finally, we conclude with some remarks on our strategy in Section 6.

2. Multiplicative multiscale methods

In this section we present our general Petrov–Galerkin multiscale approach. We first study a two-scale method based on an approximate block-factorization of the two-scale transformed linear system to be solved. Then we explain two equivalent multiscale representations of the stiffness matrix, the standard- and non-standard form [10], and derive our multiplicative subspace-correction method with a cyclic structure. It generalizes the classical HBMG method. Furthermore, we show a way to improve the standard algorithm by incorporating an explicit Schur-complement approximation into the computation of the projected residuals for the different scales.

2.1. Representation of the discrete problem on the finest level

In this subsection we deduce the linear system (1.3) that results from a general Petrov–Galerkin discretization of the continuous problem (1.1) with respect to Finite Element (FE) trial and test spaces $\mathcal{V}_0$ and $\mathcal{S}_0$ on the finest grid. To this end we define different discrete vector-like representations of functions from these finite-dimensional spaces by introducing bases. Furthermore, we define the corresponding stiffness matrix, which can also be interpreted as a discrete weak form of the continuous operator T.

Let $\Omega_0 := \{x_i^0 \in \Omega \,:\, i = 1, ..., N_0\}$ ($N_0 = \dim(\mathcal{V}_0)$) be the set of grid points on the finest grid (also denoted by Ω_0) where basis functions $\varphi_{i,\mathcal{V}}^0$ and $\varphi_{i,\mathcal{S}}^0$ spanning $\mathcal{V}_0$ and $\mathcal{S}_0$ are located. Hence, we assume that there is a one to one correspondence between grid points of the underlying mesh and basis functions on the trial and test side. We further assume that these basis functions are localized, i.e. each has compact support and is centered in its corresponding grid point. The *nodal representation* of the solution $u_0(x) = \sum_{i=1}^{N_0} u_0(x_i^0)\varphi_{i,\mathcal{V}}^0(x)$ to the discrete problem is then given by the vector

$$u_0^{NB_\mathcal{V}} := (u_0(x_i^0))_{i=1,...,N_0},$$

i.e. $u_0 \in \mathcal{V}_0$ will be identified with the vector $u_0^{NB_\mathcal{V}} \in \mathbb{R}^{N_0}$. The *momentum representation* of the right hand side is defined by the vector

$$f_0^{MOM_\mathcal{S}} := ((f, \varphi_{i,\mathcal{S}}^0)_{\mathcal{L}^2(\Omega)})_{i=1,...,N_0},$$

Finally, the *stiffness matrix* according to the bases $\{\varphi_{i,\mathcal{V}}^0 \,:\, i = 1, ..., N_0\}$ and $\{\varphi_{i,\mathcal{S}}^0 \,:\, i = 1, ..., N_0\}$ of $\mathcal{V}_0$ and $\mathcal{S}_0$ is given by

$$T_0^{NB_\mathcal{V} \to MOM_\mathcal{S}} := (a(\varphi_{j,\mathcal{V}}^0, \varphi_{i,\mathcal{S}}^0))_{i,j}.$$

Overall, we obtain the discrete linear system

$$T_0^{KB_\mathcal{V} \to MOM_\mathcal{S}} u_0^{NB_\mathcal{V}} = f_0^{MOM_\mathcal{S}} \tag{2.1}$$

as a Petrov–Galerkin discretization of (1.1). Here, the discrete operator $T_0^{KB_\mathcal{V} \to MOM_\mathcal{S}}$ maps nodal representations of functions from the trial space $\mathcal{V}_0$ to momentum representations of functions from the test space $\mathcal{S}_0$. The same distinctions arise when we transform the fine-grid linear system into a multiscale one, which will be presented in the subsequent subsections. These distinctions have to be considered during the design and the implementation of corresponding multiscale solvers. However, in the following we will suppress the cumbersome indications of representations and mapping properties for ease of notation.

2.2. Two-scale decompositions and multiplicative two-scale methods

Now, we consider two-scale decompositions of the initial trial and test spaces

$$\mathcal{V}_0 = \mathcal{W}_1 \oplus \mathcal{V}_1 \quad \text{and} \quad \mathcal{S}_0 = \mathcal{F}_1 \oplus \mathcal{S}_1 \tag{2.2}$$

into coarse-grid spaces $\mathcal{V}_1$ and $\mathcal{S}_1$ and the respective complement spaces $\mathcal{W}_1$ and $\mathcal{F}_1$. The complement spaces contain detail information which is neglected when passing from the finer to the next coarser scale. Here, we assume that we have localized bases for both the coarse-grid spaces and the complement spaces. Two examples are the classical hierarchical and prewavelet decompositions of the space of linear splines [38]. In both cases the fine- and coarse-grid spaces are spanned by scaled and dilated linear splines with compact support, the well known 'hat' functions, which are located in the fine- and coarse-grid points. For the hierarchical decomposition the corresponding complement spaces are spanned by fine-grid basis functions and in the case of the prewavelet decomposition by prewavelets. The basis functions of the complement spaces are located in the fine-without-coarse-grid points for both kinds of decomposition.

We show now how to transform the linear system (2.1) according to the direct decompositions (2.2). Due to our assumptions on the bases and the inclusions $\mathcal{V}_1 \subset \mathcal{V}_0$ and $\mathcal{W}_1 \subset \mathcal{V}_0$ there exist rectangular sparse matrices $P^0_{1,\mathcal{V}}$ and $Q^0_{1,\mathcal{V}}$ whose columns describe how basis functions $\varphi^1_{i,\mathcal{V}}$ of $\mathcal{V}_1$ and $\psi^1_{i,\mathcal{V}}$ of $\mathcal{W}_1$ may be represented by those of $\mathcal{V}_0$ according to scaling and refinement equations [10, 13]. The composed square matrix $W^0_{1,\mathcal{V}} := [Q^0_{1,\mathcal{V}}, P^0_{1,\mathcal{V}}]$ (for convenience we use notation from the Matlab programming language) describes the change of basis for a function from its two-scale nodal representation with respect to the direct decomposition $\mathcal{W}_1 \oplus \mathcal{V}_1$ with basis $\{\theta^1_{i,\mathcal{V}} : i = 1, ..., N_0\} := \{\psi^1_{i,\mathcal{V}} : x^0_i \in \Omega_0 \backslash \Omega_1\} \cup \{\varphi^1_{i,\mathcal{V}} : x^0_i \in \Omega_1\}$ to its nodal representation with fine-scale basis $\{\varphi^0_{i,\mathcal{V}} : x^0_i \in \Omega_0\}$ of $\mathcal{V}_0$. The matrix $P^0_{1,\mathcal{V}}$ is also known as *prolongation* in the context of multigrid methods [28]. It describes by means of interpolation coarse-grid basis functions of the level 1 with respect to the next finer scale of level 0. The matrix $Q^0_{1,\mathcal{V}}$ does not appear explicitly in multigrid methods because the multigrid smoothing procedure affects all degrees of freedom of the actual fine grid. However here, with its help basis functions of the complement space may be represented on the next finer grid. The inverse change of basis from the fine-grid nodal representation of a function from $\mathcal{V}_0$ into the nodal representation with respect to the direct decomposition $\mathcal{W}_1 \oplus \mathcal{V}_1$ is given by rectangular matrices $P^1_{0,\mathcal{V}}$ and $Q^1_{0,\mathcal{V}}$ of transposed dimensions. We obtain $W^1_{0,\mathcal{V}} := (W^0_{1,\mathcal{V}})^{-1} := [Q^1_{0,\mathcal{V}}; P^1_{0,\mathcal{V}}]$. In contrast to $P^0_{1,\mathcal{V}}$ and $Q^0_{1,\mathcal{V}}$ the matrices $P^1_{0,\mathcal{V}}$ and $Q^1_{0,\mathcal{V}}$ are now generally not sparse anymore. However, similar to classical HBMG methods, the matrices $P^1_{0,\mathcal{V}}$ and $Q^1_{0,\mathcal{V}}$ are never needed explicitly for the formulation and implementation of our algorithms.

One obtains analogous transformations and matrices for the decomposition on the test side and, together with these, the original system (2.1) can be transformed equivalently to

$$\begin{aligned} (W^0_{1,\mathcal{S}})^t T_0 W^0_{1,\mathcal{V}} \tilde{u}_0 &= (W^0_{1,\mathcal{S}})^t f_0, \\ \Longleftrightarrow: \qquad \tilde{T}_0 \tilde{u}_0 &= \tilde{f}_0. \end{aligned} \tag{2.3}$$

Here $\tilde{u}_0 := (W^0_{1,\mathcal{V}})^{-1} u_0$ denotes the two-scale nodal representation of u_0 and $\tilde{f}_0$ the corresponding momentum representation according to the decompositions (2.2). With the definitions

$$A_1 := (Q^0_{1,\mathcal{S}})^t T_0 Q^0_{1,\mathcal{V}}, \qquad B_1 := (Q^0_{1,\mathcal{S}})^t T_0 P^0_{1,\mathcal{V}},$$
$$C_1 := (P^0_{1,\mathcal{S}})^t T_0 Q^0_{1,\mathcal{V}}, \qquad T_1 := (P^0_{1,\mathcal{S}})^t T_0 P^0_{1,\mathcal{V}},$$

$[d_1^u; s_1^u] := (W^0_{1,\mathcal{V}})^{-1} u_0$ and $[d_1^f; s_1^f] := (W^0_{1,\mathcal{S}})^t f_0$ the transformed system (2.3) can be rewritten in block-form

$$\begin{pmatrix} A_1 & B_1 \\ C_1 & T_1 \end{pmatrix} \begin{pmatrix} d_1^u \\ s_1^u \end{pmatrix} = \begin{pmatrix} d_1^f \\ s_1^f \end{pmatrix}. \tag{2.4}$$

For a better understanding of the multiplicative multiscale method presented in the next subsection we briefly examine its two-scale analogue first. We consider the iteration

$$\tilde{u}_0^{i+1} = \tilde{u}_0^i + \tilde{M}_0^{-1}(\tilde{f}_0 - \tilde{T}_0 \tilde{u}_0^i), \quad (i = 0, 1, 2, ...), \tag{2.5}$$

for the transformed system where $\tilde{M}_0$ is defined by the following approximate block-factorization of $\tilde{T}_0$:

$$\tilde{M}_0 \quad := \quad \begin{pmatrix} A_1 & \mathbf{0} \\ C_1 & T_1 \end{pmatrix} \begin{pmatrix} \mathbf{1} & A_1^{-1} B_1 \\ \mathbf{0} & \mathbf{1} \end{pmatrix} =: \tilde{L}_0 \tilde{U}_0.$$

Compared with the exact block-factorization of $\tilde{T}_0$ the first lower block-triangular factor $\tilde{L}_0$ of $\tilde{M}_0$ approximates the arising Schur-complement operator

$$S_1 := \tilde{T}_0 / T_1 := T_1 - C_1 A_1^{-1} B_1$$

simply by its first term, the coarse-grid operator T_1. The application of the inverse $\tilde{M}_0^{-1} = \tilde{U}_0^{-1} \tilde{L}_0^{-1}$ to the residual $\tilde{f}_0 - \tilde{T}_0 \tilde{u}_0^i$ can be realized by a two-scale forward- and backward-substitution. This involves once the inversion of the block T_1 and twice the inversion of A_1. Iteration (2.5) is equivalent to a symmetric block-Gauß–Seidel iteration for the two-scale transformed linear system [6] since

$$\tilde{M}_0 = \left[\begin{pmatrix} \mathbf{0} & \mathbf{0} \\ C_1 & \mathbf{0} \end{pmatrix} + \begin{pmatrix} A_1 & \mathbf{0} \\ \mathbf{0} & T_1 \end{pmatrix} \right] \begin{pmatrix} A_1 & \mathbf{0} \\ \mathbf{0} & T_1 \end{pmatrix}^{-1} \left[\begin{pmatrix} \mathbf{0} & B_1 \\ \mathbf{0} & \mathbf{0} \end{pmatrix} + \begin{pmatrix} A_1 & \mathbf{0} \\ \mathbf{0} & T_1 \end{pmatrix} \right].$$

It is easy to see that the iteration (2.5) is equivalent to a successive subspace-correction method [56] in the case of an exact inversion of the block A_1. Starting from the system (2.1) the iteration (2.5) can also be performed without use of the generally dense matrices $P^1_{0,\mathcal{V}/\mathcal{S}}$ and $Q^1_{0,\mathcal{V}/\mathcal{S}}$ and may then be formulated as a two-scale V(1,1)-cycle. This will be presented for the case of multiscale decompositions in the next subsection. The block A_1 is well conditioned for many decompositions since it contains only couplings between the subspaces $\mathcal{W}_1$ and $\mathcal{F}_1$ of detail information [22]. Hence, it is often sufficient to invert A_1 only approximately, e.g. by an iterative method or an incomplete factorization (ILU). Note however that the iterates from the symmetric block-Gauß–Seidel iteration for the two-scale transformed linear system and from the successive subspace-correction method will be different if we approximate A_1^{-1}. This is due to a slight difference when computing corrections of the residual projections for the post-correction steps. But this does not affect the convergence behavior significantly [34].

2.3. Multiscale decompositions and multiplicative multiscale methods

By the recursive decomposition of the coarse-grid spaces $\mathcal{V}_1$ and $\mathcal{S}_1$ to a coarsest level l we obtain the following multiscale decompositions of the initial trial and test spaces $\mathcal{V}_0$ and $\mathcal{S}_0$:

$$\mathcal{V}_0 = \bigoplus_{k=1}^{l} \mathcal{W}_k \oplus \mathcal{V}_l \quad \text{and} \quad \mathcal{S}_0 = \bigoplus_{k=1}^{l} \mathcal{F}_k \oplus \mathcal{S}_l. \tag{2.6}$$

Now, a canonical representation of the finest-grid stiffness matrix T_0 relative to the multiscale decompositions (2.6) is induced by the corresponding change of basis. Let $W^0_{l,\mathcal{V}}$ be the transformation from the multiscale nodal representation to the finest-grid nodal representation on the trial side. Then the transposed matrix $(W^0_{l,\mathcal{S}})^t$ will be the corresponding change of basis on the test side with respect to moments. From these two transformation matrices we obtain the *standard form* $\tilde{T}_{0,s} := (W^0_{l,\mathcal{S}})^t T_0 W^0_{l,\mathcal{V}}$ and the associated linear system

$$(W^0_{l,\mathcal{S}})^t T_0 W^0_{l,\mathcal{V}} \tilde{u}_{0,s} = (W^0_{l,\mathcal{S}})^t f_0, \tag{2.7}$$

where $\tilde{u}_{0,s} := (W^0_{l,\mathcal{V}})^{-1} u_0$ [10]. From an algorithmic point of view it is not advantageous to work directly with the standard form. If T_0 is sparse with $\mathcal{O}(N_0)$ entries, $\tilde{T}_{0,s}$ will have in general at least $\mathcal{O}(N_0(l+1))$ entries. If $l = \log_2(h_0^{-1})$, which is the case for the standard geometric coarsening of grids generated by dyadical refinement, one will loose linear complexity. The *non-standard form* introduced in [10] in the context of sparse representations of integral operators is an alternative multiscale representation of the stiffness matrix T_0. The non-standard form is generated by the recursive transformation of the coarse-grid operators T_k for $k = 1, ..., l-1$ only. Then, we obtain

$$\tilde{T}_{0,ns} := \{\{A_k, B_k, C_k\}_{k=1,\ldots,l}, T_l\}. \tag{2.8}$$

The non-standard form gives no ordinary matrix anymore, however, its blocks may be imbedded into a generating system representation of T_0 where also the blocks of the standard form appear [34]. If T_0 is sparse with $\mathcal{O}(N_0)$ entries, then its non-standard form $\tilde{T}_{0,ns}$ will have at most $\mathcal{O}(N_0)$ entries due to a geometric series argument. Therefore the non-standard form allows for the efficient implementation of multiplicative multiscale methods, e.g. consider the iteration

$$u_0^{i+1} = \text{ms\_v\_cycle}(l, 1, u_0^i, f_0), \tag{2.9}$$

where the function "ms_v_cycle" is recursively defined using the blocks of the non-standard form (2.8) by

Algorithm (Basic multiscale V(1,1)-cycle)

`function` $x = \text{ms\_v\_cycle}(l, k, x, f)$;

1. $[d^{res}; s^{res}] = [W^{(k-1)}_{k,\mathcal{S}}]^t (f - T_{k-1}x)$; (decomposition)
2. $d_v = A_k^{-1} d^{res}$; (pre-correction)
3. $s^{res}_m = s^{res} - C_k d_v$; (modification of the coarse-scale part of the residual)
4. `if` $k == l$
   $s_v = T_l^{-1} s^{res}_m$;
   `else`

$s_v = \mathrm{ms\_v\_cycle}(l, k+1, \mathbf{0}, s_m^{res});$
`end`

5. $d_n = A_k^{-1}(d^{res} - A_k d_v - B_k s_v);$ (post-correction)
6. $x = x + W_{k,\mathcal{V}}^{(k-1)}[d_v + d_n; s_v];$ (back-transformation and correction)

We assume here that the vectors x, f and $\mathbf{0}$ have appropriate dimensions. The above basic multiscale algorithm has a structure very similar to the standard multigrid V(1,1)-cycle [28]. Note that also the classical HBMG method can be formulated within our general framework [34].

We give several remarks on the algorithm:

i) It can equivalently be realized by multiscale forward- and backward-substitutions that rely on lower and upper triangular factors $\tilde{L}_{0,ns}$ and $\tilde{U}_{0,ns}$ of an approximate LU-factorization of $\tilde{T}_{0,ns}$. If the blocks A_k are inverted in steps 2 and 5 only approximately by $\breve{A}_k^{-1}$, e.g. by an iterative method or by ILU, we will obtain slightly different iterates as already remarked for the two-scale method [34].

ii) Using an incomplete factorization of the blocks A_k the multiscale forward- and backward-substitutions from $\tilde{L}_{0,ns}$ and $\tilde{U}_{0,ns}$ are equivalent to forward- and backward-substitutions with respect to lower and upper triangular factors $\tilde{L}_{0,s}$ and $\tilde{U}_{0,s}$ of an incomplete factorization of the standard form $\tilde{T}_{0,s}$. Hence, our multiscale V(1,1)-cycle may also be interpreted as an efficient implementation of an iteration for the standard-transformed system (2.7) based on an incomplete factorization of $\tilde{T}_{0,s}$ [22, 34].

iii) It is equivalent to an ordinary multigrid V(1,1)-cycle, using the special *multiscale smoothing iteration*

$$x = x + Q_{k,\mathcal{V}}^{k-1} \breve{A}_k^{-1} (Q_{k,\mathcal{S}}^{k-1})^t (f - T_{k-1}x). \tag{2.10}$$

Here, $\breve{A}_k^{-1}$ again denotes an approximate inverse of A_k as in i). The method therefore allows for more general cycles, e.g. a multiscale W(1,1)-cycle, which can be implemented with linear complexity in the case of geometric coarsening of the fine grid.

iv) The implementation is possible without the complete computation of $\tilde{T}_{0,ns}$ because for $k = 1, ..., l$ we have the representations

$$B_k = (Q_{k,\mathcal{S}}^{k-1})^t T_{k-1} P_{k,\mathcal{V}}^{k-1} \quad \text{and} \quad C_k = (P_{k,\mathcal{S}}^{k-1})^t T_{k-1} Q_{k,\mathcal{V}}^{k-1} \tag{2.11}$$

for the off-diagonal blocks. The resulting iterations are the so-called *standard* multiscale methods. If the blocks A_k are (approximately) inverted by an iterative method, where only its applications to vectors are needed, one can also use the representations

$$A_k = (Q_{k,\mathcal{S}}^{k-1})^t T_{k-1} Q_{k,\mathcal{V}}^{k-1}.$$

v) Only the application of the off-diagonal blocks is needed for the correction of the projected residuals. We therefore propose for $k = 2, ..., l$ the following improved representations (compare (2.11))

$$B_k^S = (Q_{k,\mathcal{S}}^{k-1})^t \breve{S}_{k-1} P_{k,\mathcal{V}}^{k-1} \quad \text{and} \quad C_k^S = (P_{k,\mathcal{S}}^{k-1})^t \breve{S}_{k-1} Q_{k,\mathcal{V}}^{k-1}. \tag{2.12}$$

Here, the matrices $\breve{S}_{k-1} := T_{k-1} - B_{k-1}\breve{A}_{k-1}^{-1}C_{k-1}$ are feasible approximations of the exact Schur-complement operators S_{k-1}, e.g. using incomplete factorizations of the blocks A_{k-1}. Besides, the Schur-complement approximations $\breve{S}_{k-1}$ may be used within the decomposition step 1 of the above algorithm. Thus, fine-scale features may be transported more accurately to coarser scales. We call the corresponding iterations *improved* multiscale methods. Depending on the approximations $\breve{S}_{k-1}$ the computational cost for our improved methods can be quite high. This will be investigated in Section 5. For an interesting approach, which uses Schur-complement approximations also for the multiscale decomposition steps, see [37]. However, in [37] the multiscale decomposition is obtained only by a multilevel ordering of the unknowns.

3. Problem-dependent geometric multiscale decompositions

Starting from a rather general Petrov–Galerkin multiscale concept we developed in the previous section multiplicative subspace-correction methods with a cyclic structure that generalize the classical HBMG method. It is clear that the quality and robustness of our algorithms essentially depend on the transformation matrices $P^k_{k+1,\mathcal{V}/\mathcal{S}}$ and $Q^k_{k+1,\mathcal{V}/\mathcal{S}}$ $(k = 0, ..., l-1)$, which define the multiscale decompositions and hereby implicitly the underlying coarse-grid spaces and complement spaces. They determine the blocks of the non-standard form $\tilde{T}_{0,ns}$ as well as the projections of the residual vectors into subspaces and the final back-transformations of the computed corrections.

In this section we discuss problem-dependent choices for the matrices $P^k_{k+1,\mathcal{V}/\mathcal{S}}$, which are crucial to obtain physically meaningful coarse-grid operators. We will see that problem-dependent transformations $P^k_{k+1,\mathcal{V}/\mathcal{S}}$ can be derived from an operator induced interpolation by solving *continuous* homogeneous (transposed) problems locally. This may also be interpreted by means of operator-dependent FE basis functions (L-splines), which fulfill a generalized scaling equation. With help of these local Green's functions equivalent matrix-dependent prolongations can be derived from the resulting stiffness matrix by solving the underlying *discrete* homogeneous (transposed) problems locally. This is equivalent to a local Gaussian elimination. After the discussion of the problem-dependent choices of the prolongation matrices $P^k_{k+1,\mathcal{V}/\mathcal{S}}$ we investigate their interplay with hierarchical and wavelet-like choices for the matrices $Q^k_{k+1,\mathcal{V}/\mathcal{S}}$ and the hereby induced complement bases. We first discuss one-dimensional problems and then show how the one-dimensional techniques carry over to two-dimensional cases by means of a tensor product ansatz. The corresponding problem-dependent multiscale decompositions for 2D problems rely essentially on the geometric coarsening of the underlying grid. They are therefore called *geometric multiscale decompositions.*

3.1. Problem-dependent prolongations $P^k_{k+1,\mathcal{V}/\mathcal{S}}$ for 1D problems

To derive a problem-dependent prolongation $P^k_{k+1,\mathcal{V}}$ on the trial side we consider the one-dimensional homogeneous convection-diffusion problem with constant scalar convection $b \neq 0$ and inhomogeneous Dirichlet boundary conditions

$$-u'' + bu' = 0 \quad \text{on} \quad]x_i^{k+1}; x_{i+1}^{k+1}[, \quad u(x_i^{k+1}) = \alpha \quad \text{and} \quad u(x_{i+1}^{k+1}) = \beta. \tag{3.1}$$

Here, the interval $]x_i^{k+1}; x_{i+1}^{k+1}[\subset \Omega :=]0,1[$ is supposed to be a finite element of length $h_{k+1} = 2h_k$ with respect to a uniform mesh with grid points $\Omega_{k+1} := \{x_i^{k+1} \in \Omega : x_i^{k+1} := ih_{k+1}$ for $0 < i < 1/h_{k+1}\}$. Furthermore, let us assume that the mesh Ω_{k+1} is imbedded into a nested sequence of dyadically refined meshes $\Omega_l \subset \cdots \subset \Omega_{k+1} \subset \Omega_k \subset \cdots \subset \Omega_0$. Now, the solution to the one-dimensional boundary value problem (3.1) can be easily computed as

$$u(x) = c_1 + c_2 \exp(bx) \tag{3.2}$$

with constants

$$c_1 = \frac{\alpha \exp(bx_{i+1}^{k+1}) - \beta \exp(bx_i^{k+1})}{\exp(bx_{i+1}^{k+1}) - \exp(bx_i^{k+1})} \quad \text{and} \quad c_2 = \frac{\beta - \alpha}{\exp(bx_{i+1}^{k+1}) - \exp(bx_i^{k+1})}.$$

From the solution u to (3.1) we are now able to define a problem-adapted interpolation at the midpoint of the interval $x_{i+1/2}^{k+1} := x_{2i+1}^k = (x_i^{k+1} + x_{i+1}^{k+1})/2 \in \Omega_k$. To this end we assume that the interpolation of (3.2) from the boundary values α and β with some weights p_W and p_E is exact at $x_{i+1/2}^{k+1}$:

$$u(x_{i+1/2}^{k+1}) \stackrel{!}{=} p_E \alpha + p_W \beta.$$

After some computations we obtain for these weights $p_W = \frac{1}{2}\left(1 - \tanh\left(\frac{bh_k}{2}\right)\right)$ and $p_E = \frac{1}{2}\left(1 + \tanh\left(\frac{bh_k}{2}\right)\right)$. This leads to the one-dimensional prolongation stencil

$$(P_{k+1,\mathcal{V}}^k)_* = [p_W \quad p_C \quad p_E] = \frac{1}{2}\left\{\begin{bmatrix}1 & 2 & 1\end{bmatrix} + \tanh\left(\frac{bh_k}{2}\right)\begin{bmatrix}-1 & 0 & 1\end{bmatrix}\right\}, \tag{3.3}$$

which defines a problem-dependent interpolation from the solution u to (3.1). Let us now consider a FE interpretation of the molecule (3.3). To this end we assume that $\mathcal{V}_{k+1}$ is a FE space spanned by local FE basis functions $\varphi_{i,\mathcal{V}}^{k+1}$ centered in the grid points of Ω_{k+1}. An approximation u_{k+1} of the continuous solution u may then be written as a linear combination

$$u_{k+1}(x) = \alpha \varphi_{i,\mathcal{V}}^{k+1}(x) + \beta \varphi_{i+1,\mathcal{V}}^{k+1}(x) \tag{3.4}$$

of the two basis functions located at the boundary points of the interval $]x_i^{k+1}; x_{i+1}^{k+1}[$. Now, if we piecewise define operator-dependent basis functions by the exact solutions to the homogeneous problems with boundary values 0 and 1

$$\varphi_{i,\mathcal{V}}^{k+1}(x) := \begin{cases} \dfrac{\exp(bx) - \exp(bx_{i-1}^{k+1})}{\exp(bx_i^{k+1}) - \exp(bx_{i-1}^{k+1})}, & x \in [x_{i-1}^{k+1}, x_i^{k+1}[, \\ \dfrac{\exp(bx_{i+1}^{k+1}) - \exp(bx)}{\exp(bx_{i+1}^{k+1}) - \exp(bx_i^{k+1})}, & x \in [x_i^{k+1}, x_{i+1}^{k+1}[, \\ 0, & \text{else}, \end{cases} \tag{3.5}$$

we will find that the approximation (3.4) is exact, i.e. $u_{k+1}(x) = u(x)$ on $]x_i^{k+1}; x_{i+1}^{k+1}[$. The so defined local Green's functions are the well known *L-splines* [45] and the weights of the problem-dependent interpolation can equally be derived from the values of the L-spline basis functions at $x_{i-1/2}^{k+1}$, x_i^{k+1} and $x_{i+1/2}^{k+1}$. Furthermore, these L-splines fulfill the generalized scaling equation

$$\varphi_{i,\mathcal{V}}^{k+1}(x) = p_W \varphi_{2i-1,\mathcal{V}}^k(x) + p_C \varphi_{2i,\mathcal{V}}^k(x) + p_E \varphi_{2i+1,\mathcal{V}}^k(x). \tag{3.6}$$

The prolongation molecule (3.3) may therefore be interpreted as the *filter* for a generalized multiresolution analysis given by the nested sequence of L-spline FE spaces $\mathcal{V}_l \subset \cdots \subset \mathcal{V}_{k+1} \subset \mathcal{V}_k \subset \cdots \subset \mathcal{V}_0 (\subset \cdots \subset \mathcal{L}^2(\Omega))$ [34]. Let us assume that there are localized bases $\{\varphi_{i,\mathcal{S}}^k : i =$

$1,...,N_k\}$ for the test spaces $\mathcal{S}_k$ $(k=0,...,l)$ such that $\operatorname{supp}(\varphi^k_{i,\mathcal{S}}) \subset \operatorname{supp}(\varphi^k_{i,\mathcal{V}})$. Then the computation of the stiffness matrix with respect to the bilinear-form $a(u,v)=\int_\Omega u'v'+bu'v\,dx$ leads after appropriate scaling to the well known finite differences scheme due to Ilin [33]

$$(T_k)_* := [t_w \quad t_c \quad t_e] := \frac{\gamma_k}{h_k}\Big[-1 \quad 2 \quad -1\Big] + \frac{b}{2}\Big[-1 \quad 0 \quad 1\Big], \tag{3.7}$$

where $\gamma_k := \frac{bh_k}{2}\coth\left(\frac{bh_k}{2}\right)$ is a scale-dependent viscosity factor. Note that the entries of the stiffness matrix (3.7) are independent of the basis of the test space. The result of the prolongation of a coarse-grid function u_{k+1} in the fine-without-coarse-grid points $\Omega^f_k := \Omega_k\backslash\Omega_{k+1}$ shall be given by the solution of the homogeneous local problems. This can now be written in discrete form as

$$(T_k P^k_{k+1,\mathcal{V}} u_{k+1})(x^k_i) = 0 \quad \text{for} \quad x^k_i \in \Omega^f_k. \tag{3.8}$$

Let $u_k := P^k_{k+1,\mathcal{V}} u_{k+1}$ be the interpolation of a coarse-grid function to the next finer grid. Then one obtains from (3.8) together with the boundary conditions $u_k(x^{k+1}_{i-1})=0$ and $u_k(x^{k+1}_{i+1})=0$

$$t_w u_k(x^{k+1}_{i-1}) + t_c u_k(x^{k+1}_{i-1/2}) + t_e u_k(x^{k+1}_i) = 0, \quad \text{i.e. } u_k(x^{k+1}_{i-1/2}) = -\frac{t_e}{t_c}u_k(x^{k+1}_i),$$
$$t_w u_k(x^{k+1}_i) + t_c u_k(x^{k+1}_{i+1/2}) + t_e u_k(x^{k+1}_{i+1}) = 0, \quad \text{i.e. } u_k(x^{k+1}_{i+1/2}) = -\frac{t_w}{t_c}u_k(x^{k+1}_i).$$

The comparison with (3.3) shows that the weights $p^m_W := -(t_e/t_c)$ and $p^m_E := -(t_w/t_c)$ of the so defined matrix-dependent prolongation equal those of the operator-dependent prolongation (3.3).

The above statements will also hold for a non-constant convection $b=b(x)$, if the solutions to the homogeneous continuous equations are computed exactly [34]. However, then the stencils $(P^k_{k+1,\mathcal{V}})_*$ and $(T_k)_*$ will not only be scale-dependent but also local. We finally remark that the prolongations $P^k_{k+1,\mathcal{S}}$ on the test side might be chosen from a linear interpolation or from an interpolation dependent on the transposed problem. The latter choice will be motivated in the next subsection.

3.2. Choice of the complement bases for 1D problems

A simple but effective choice for a basis of the complement space $\mathcal{F}_{k+1}$ on the test side is given by the fine-grid basis functions $\varphi^k_{i,\mathcal{S}}$ which are centered in the fine-without-coarse-grid points from Ω^f_k. It can be implemented by the trivial molecule $(Q^k_{k+1,\mathcal{S}})_* = [\quad 1 \quad]$ located in these points. We call it the *hierarchical choice.* For this choice we obtain that the entries within the block $B_{k+1} = (Q^k_{k+1,\mathcal{S}})^t T_k P^k_{k+1,\mathcal{V}}$ of the two-scale transformed stiffness matrix $\tilde{T}_k = (W^k_{k+1,\mathcal{S}})^t T_k W^k_{k+1,\mathcal{V}}$ are zero. We have

$$\begin{aligned}(B_k)_{i+1,i} = a(\varphi^{k+1}_{i,\mathcal{V}}, \varphi^k_{2i+1,\mathcal{S}}) &= a(p_W\varphi^k_{2i-1,\mathcal{V}} + 1\cdot\varphi^k_{2i,\mathcal{V}} + p_E\varphi^k_{2i+1,\mathcal{V}}, \varphi^k_{2i+1,\mathcal{S}}) \\ &= 1\cdot a(\varphi^k_{2i,\mathcal{V}}, \varphi^k_{2i+1,\mathcal{S}}) + p_E a(\varphi^k_{2i+1,\mathcal{V}}, \varphi^k_{2i+1,\mathcal{S}}) \\ &= t_w(x^k_{2i+1}) - \frac{t_w(x^k_{2i+1})}{t_c(x^k_{2i+1})}t_c(x^k_{2i+1}) = 0,\end{aligned}$$

$$\begin{aligned}(B_k)_{i-1,i} = a(\varphi_{i,\mathcal{V}}^{k+1}, \varphi_{2i-1,\mathcal{S}}^k) &= a(p_W \varphi_{2i-1,\mathcal{V}}^k + 1 \cdot \varphi_{2i,\mathcal{V}}^k + p_E \varphi_{2i+1,\mathcal{V}}^k, \varphi_{2i-1,\mathcal{S}}^k) \\ &= p_W a(\varphi_{2i-1,\mathcal{V}}^k, \varphi_{2i-1,\mathcal{S}}^k) + 1 \cdot a(\varphi_{2i,\mathcal{V}}^k, \varphi_{2i-1,\mathcal{S}}^k) \\ &= -\frac{t_e(x_{2i-1}^k)}{t_c(x_{2i-1}^k)} t_c(x_{2i-1}^k) + t_e(x_{2i-1}^k) = 0,\end{aligned}$$

due to the scaling relation (3.6). Hence, by the hierarchical choice for the complement space on the test side the coarse-grid unknowns have been eliminated from the transformed equations for the fine-without-coarse-grid degrees of freedom. Turning the analysis around we see that the decoupling relation $B_{k+1} = \mathbf{0}$ is in fact equivalent to the use of L-splines as trial functions together with a hierarchical choice on the test side.

We will now see from a simple transposition argument that also the lower left off-diagonal block $C_{k+1} = (P_{k+1,\mathcal{S}}^k)^t T_k Q_{k+1,\mathcal{V}}^k$ of $\tilde{T}_k$ may vanish, which means a complete decoupling of the blocks of the transformed system. To this end we additionally choose $P_{k+1,\mathcal{S}}^k$ dependent on the transposed stiffness matrix T_k^t and make a hierarchical choice for $Q_{k+1,\mathcal{V}}^k$. If we interchange the trial and test side, denoted by a lower index $_{\mathcal{V}\circlearrowleft\mathcal{S}}$, and compute a corresponding two-scale transformation of the transposed stiffness matrix $T_{k,\mathcal{V}\circlearrowleft\mathcal{S}} = T_k^t$, we will find that

$$\begin{aligned}\tilde{T}_{k,\mathcal{V}\circlearrowleft\mathcal{S}} &:= (W_{k+1,\mathcal{V}}^k)^t T_k^t W_{k+1,\mathcal{S}}^k \\ &= \begin{pmatrix} (Q_{k+1,\mathcal{V}}^k)^t T_k^t Q_{k+1,\mathcal{S}}^k & (Q_{k+1,\mathcal{V}}^k)^t T_k^t P_{k+1,\mathcal{S}}^k \\ (P_{k+1,\mathcal{V}}^k)^t T_k^t Q_{k+1,\mathcal{S}}^k & (P_{k+1,\mathcal{V}}^k)^t T_k^t P_{k+1,\mathcal{S}}^k \end{pmatrix},\end{aligned}$$

where $B_{k+1} = (Q_{k+1,\mathcal{V}}^k)^t T_k^t P_{k+1,\mathcal{S}}^k = \mathbf{0}$. Hence, we obtain the desired block-diagonal form

$$\begin{aligned}\tilde{T}_k &= ((W_{k+1,\mathcal{V}}^k)^t T_k^t W_{k+1,\mathcal{S}}^k)^t \\ &= \begin{pmatrix} (Q_{k+1,\mathcal{S}}^k)^t T_k Q_{k+1,\mathcal{V}}^k & \mathbf{0} \\ \mathbf{0} & (P_{k+1,\mathcal{S}}^k)^t T_k P_{k+1,\mathcal{V}}^k \end{pmatrix}\end{aligned}$$

by transposition of $\tilde{T}_{k,\mathcal{V}\circlearrowleft\mathcal{S}}$ together with the former $(Q_{k+1,\mathcal{S}}^k)^t T_k P_{k+1,\mathcal{V}}^k = \mathbf{0}$. We see that $A_{k+1} = (Q_{k+1,\mathcal{S}}^k)^t T_k Q_{k+1,\mathcal{V}}^k$ is diagonal due to the hierarchical choices for the bases of $\mathcal{W}_{k+1}$ and $\mathcal{F}_{k+1}$. Furthermore, the coarsening process preserves the form of a three-point stencil for the coarse-grid operator $T_{k+1} = (P_{k+1,\mathcal{S}}^k)^t T_k P_{k+1,\mathcal{V}}^k$ and the matrix-dependent computation of the prolongations can be continued recursively. With the proposed problem-dependent multiscale decompositions our multiscale V(1,1)-cycle leads to a direct solver because the non-standard form $\tilde{T}_{0,ns}$ is diagonal.

Unfortunately the above hierarchical choice for the complement spaces is not as efficient for higher dimensional problems since

- the elimination of the off-diagonal blocks is only approximate [24, 34] and
- hierarchical multiscale algorithms suffer from bad stability properties of the hierarchical decompositions in higher dimensional spaces [38].

However, using here (pre-)wavelet-like functions to span the complement spaces, it can be proven for symmetric positive definite problems that the multiscale transformed stiffness matrix, i.e. the standard form $\tilde{T}_{0,s}$, is spectrally equivalent to a diagonal matrix [38]. Now, our multiscale V(1,1)-cycle will also show an optimal convergence behavior similar to classical multigrid. This is due to the enhanced stability properties of the corresponding decompositions. From the multigrid interpretation of the multiscale V(1,1)-cycle the reason is the stronger effect of the corresponding multiscale smoother (2.10) compared to the hierarchical one.

A simple way to construct (pre-)wavelet-like decompositions for higher dimensions is by means of a tensor product approach which uses one-dimensional splittings. We therefore want to close this discussion with a brief comment on the construction of (pre-)wavelet-like complement bases in 1D. Let $\mathcal{V}_k$ and $\mathcal{V}_{k+1}$ be the spaces spanned by one-dimensional linear splines with respect to the grids Ω_k and Ω_{k+1}. *Linear prewavelets* are defined and constructed to be localized basis functions for $\mathcal{W}_{k+1}$ that are $\mathcal{L}^2(\Omega)$-orthogonal to $\mathcal{V}_{k+1}$ [27]. They can be characterized in fine-without-coarse-grid points which are not directly adjacent to the boundaries 0 and 1 by the molecule $(Q^k_{k+1,\mathcal{V}})_* = [1 \quad -6 \quad \underline{10} \quad -6 \quad 1]/10$ centered there. (The weight for the respective degree of freedom from Ω^f_k will be underlined.) For the western and eastern points from Ω^f_k right next to the boundaries one obtains the modified molecules $(Q^k_{k+1,\mathcal{V},w})_* = [\underline{9} \quad -6 \quad 1]/10$ and $(Q^k_{k+1,\mathcal{V},e})_* = [1 \quad -6 \quad \underline{9}]/10$. If $\mathcal{V}_k$ and $\mathcal{V}_{k+1}$ are spanned by one-dimensional L-splines, which result from a problem-dependent prolongation process as described in the previous subsection, corresponding *L-spline prewavelets* can be constructed (even locally) in the same way. One finds that the one-dimensional stencils for the multiscale transformations tend for increasing convection b continuously to those characterizing the Haar system [18]. This is also evident from the convergence $[1 \quad 2 \quad 1]/2 \longrightarrow [0 \quad 1 \quad 1]$ for $(P^k_{k+1,\mathcal{V}})_*$ when $b \to +\infty$.

3.3. Problem-dependent prolongations $P^k_{k+1,\mathcal{V}/\mathcal{S}}$ for 2D problems

In this subsection we generalize the derivation of the problem-dependent prolongations from the one-dimensional to the two-dimensional case. We first assume that the vector field of convection $\vec{b} = (b_1, b_2)$ is constant. Then it is possible to split the continuous convection-diffusion operator $T = -\Delta + b_1\partial_x + b_2\partial_y$ additively into two directional parts $T_x := -\partial_x^2 + b_1\partial_x$ and $T_y := -\partial_y^2 + b_2\partial_y$ with

$$T = T_x + T_y. \tag{3.9}$$

Let us consider a grid point $(x_i^{k+1}, y_j^{k+1}) \in \Omega :=]0,1[^2$ of a square coarse grid $\Omega_{k+1} := \{(x_i^{k+1}, y_j^{k+1}) \in \Omega \ : \ x_i^{k+1} := ih_{k+1} \text{ and } y_j^{k+1} := jh_{k+1} \text{ for } 0 < i,j < 1/h_{k+1}\}$ with $h_{k+1} = 2h_k$. Again, we assume that Ω_{k+1} is imbedded into a nested sequence of dyadically refined uniform square meshes $\Omega_l \subset \cdots \subset \Omega_{k+1} \subset \Omega_k \subset \cdots \subset \Omega_0$. Using a classical four-color subdivision the next finer grid Ω_k can be subdivided into the four disjoint sets

$$\begin{aligned}
\Omega_k^{(1,1)} &:= \{(x_i^k, y_j^k) \in \Omega \, : x_i^k = ih_k, y_j^k = jh_k, \ 0 < i,j < 1/h_k, i \text{ odd}, \, j \text{ odd}\},\\
\Omega_k^{(0,1)} &:= \{(x_i^k, y_j^k) \in \Omega \, : x_i^k = ih_k, y_j^k = jh_k, \ 0 < i,j < 1/h_k, i \text{ even}, \, j \text{ odd}\},\\
\Omega_k^{(1,0)} &:= \{(x_i^k, y_j^k) \in \Omega \, : x_i^k = ih_k, y_j^k = jh_k, \ 0 < i,j < 1/h_k, i \text{ odd}, \, j \text{ even}\},\\
\Omega_k^{(0,0)} &:= \{(x_i^k, y_j^k) \in \Omega \, : x_i^k = ih_k, y_j^k = jh_k, \ 0 < i,j < 1/h_k, i \text{ even}, \, j \text{ even}\}\\
&= \Omega_{k+1}.
\end{aligned}$$

If we apply the one-dimensional technique separately to T_x and T_y, we will obtain for the two directions the one-dimensional prolongation stencils

$$\begin{aligned}(P^{k,x}_{k+1,\mathcal{V}})_* &= [p_W \quad p_{C,x} \quad p_E] &= \frac{1}{2}\Big\{\Big[1 \quad 2 \quad 1\Big] + \tanh\Big(\frac{b_1 h_k}{2}\Big)\Big[-1 \quad 0 \quad 1\Big]\Big\},\\ (P^{k,y}_{k+1,\mathcal{V}})_* &= [p_N \quad p_{C,y} \quad p_S] &= \frac{1}{2}\Big\{\Big[1 \quad 2 \quad 1\Big] + \tanh\Big(\frac{b_2 h_k}{2}\Big)\Big[-1 \quad 0 \quad 1\Big]\Big\}.\end{aligned}$$

They define the weights of a problem-dependent interpolation already for grid points from $\Omega^{(1,0)}_k$ and $\Omega^{(0,1)}_k$. Again, this may be interpreted by means of L-splines $\varphi^{k+1,x}_{(i,j),\mathcal{V}}$ and $\varphi^{k+1,y}_{(i,j),\mathcal{V}}$, which span one-dimensional operator-dependent FE spaces $\mathcal{V}^{x,j}_{k+1}$ and $\mathcal{V}^{y,i}_{k+1}$ along the grid lines $y = y^{k+1}_j = const$ and $x = x^{k+1}_i = const$ of the fine grid Ω_k. (In the following one-dimensional basis functions from the spaces $\mathcal{V}^x_{k+1}$ and $\mathcal{V}^y_{k+1}$ and its complements will be indexed by the corresponding grid-points from the fine grid Ω_k.) Now, we form abstract one-dimensional spaces $\mathcal{V}^x_{k+1}$ and $\mathcal{V}^y_{k+1}$ from these grid-aligned spaces by ordering the different grid lines with respect to the x- and y-direction. By tensor product construction one is lead to the two-dimensional operator-dependent FE space

$$\mathcal{V}^{(0,0)}_{k+1} = \Big\langle \varphi^{k+1}_{(i,j),\mathcal{V}} : \varphi^{k+1}_{(i,j),\mathcal{V}}(x,y) = \varphi^{k+1,x}_{(i,j),\mathcal{V}}(x)\varphi^{k+1,y}_{(i,j),\mathcal{V}}(y),\ 0 < i,j < 1/h_k, i \text{ even}, j \text{ even}\Big\rangle.$$

Here, the two-dimensional basis functions from the space $\mathcal{V}_{k+1}$ carry the same index as the one-dimensional splines from which they are build. The weights of a two-dimensional operator-dependent prolongation molecule $(P^k_{k+1,\mathcal{V}})_*$ for the adjacent fine-grid points of $(x^k_i, y^k_j) = (x^{k+1}_{i/2}, y^{k+1}_{j/2}) \in \Omega_k$ can now be determined from the values of $\varphi^{k+1}_{(i,j),\mathcal{V}}$ in these points. We obtain the nine-point stencil

$$(P^k_{k+1,\mathcal{V}})_* = [p_N \quad p_{C,y} \quad p_S]^t \cdot [p_W \quad p_{C,x} \quad p_E]. \tag{3.10}$$

Similar to (3.6) its entries may be interpreted as the coefficients of a scaling equation for a two-dimensional generalized multiresolution analysis which is given by a nested sequence of tensor product L-spline FE spaces. To construct a matrix-dependent prolongation we assume that there exist test spaces $\mathcal{S}_k$ which have also been constructed by the tensor product of one-dimensional spaces $\mathcal{S}^x_k$ and $\mathcal{S}^y_k$ that meet the requirements from the one-dimensional case. The Petrov–Galerkin discretizations of the directional parts T_x and T_y and the one-dimensional identity operations with respect to the one-dimensional spaces $\mathcal{V}^{x,j}_k$ and $\mathcal{S}^{x,j}_k$ as well as $\mathcal{V}^{y,i}_k$ and $\mathcal{S}^{y,i}_k$ lead to the following one-dimensional stencils:

$$\begin{aligned}(T^x_k)_* &:= [t_w \quad t_{c,x} \quad t_e], &\qquad (T^y_k)_* &:= [t_n \quad t_{c,y} \quad t_s],\\ (M^x_k)_* &:= [m_w \quad m_{c,x} \quad m_e], &\qquad (M^y_k)_* &:= [m_n \quad m_{c,y} \quad m_s].\end{aligned}$$

The stencil for the two-dimensional stiffness matrix which corresponds to the weak form $a(u,v) = \int_\Omega \nabla u \cdot \nabla v + b_1(\partial_x u)v + b_2(\partial_y u)v\, dxdy$ can now be computed as the sum of dyadic products of the above one-dimensional stencils:

$$(T_k)_* = (M^y_k)^t_* \cdot (T^x_k)_* + (T^y_k)^t_* \cdot (M^x_k)_*. \tag{3.11}$$

This is due to the tensor product ansatz and the separability (3.9) of T. An equivalent matrix-dependent prolongation can be derived from the computed stiffness matrix and the following well known 'lumping' procedure. Let us first define one-dimensional three-point stencils $(\breve{T}_k^x)_*$ and $(\breve{T}_k^y)_*$ from the two items of (3.11) by a diagonal lumping of the stencils $(M_k^x)_*$ and $(M_k^y)_*$:

$$\begin{aligned}(\breve{T}_k^x)_* &:= m_y \cdot (T_k^x)_*, \quad \text{where} \quad m_y := m_n + m_{c,y} + m_s,\\ (\breve{T}_k^y)_* &:= m_x \cdot (T_k^y)_*, \quad \text{where} \quad m_x := m_w + m_{c,x} + m_e.\end{aligned}$$

This may also be interpreted as a column- and row-wise summation of the entries of the nine-point stencil $(T_k)_*$ because the sum of the entries of the stencils $(T_k^x)_*$ and $(T_k^y)_*$ is zero. The condensed stencils $(\breve{T}_k^x)_*$ and $(\breve{T}_k^y)_*$ are just scaled versions of $(T_k^x)_*$ and $(T_k^y)_*$. The resulting axis-oriented weights $p_W^m := -\breve{t}_e/\breve{t}_{c,x}$, $p_E^m := -\breve{t}_w/\breve{t}_{c,x}$, $p_N^m := -\breve{t}_s/\breve{t}_{c,y}$ and $p_S^m := -\breve{t}_n/\breve{t}_{c,y}$ equal the above defined operator-dependent weights since the scaling factors cancel out. To define the remaining weights p_{NW}^m, p_{NE}^m, p_{SW}^m and p_{SE}^m let us consider the five-point stencil

$$(\breve{T}_k)_* := (\breve{T}_k^x)_* + (\breve{T}_k^y)_*^t. \tag{3.12}$$

For points from $\Omega_k^{(1,1)}$ we locally solve the discrete homogeneous problems with $\breve{T}_k$ where together with zeros the already defined axis-oriented weights serve as discrete inhomogeneous boundary values. We arrive for example for p_{SW}^m at the equation

$$\breve{t}_c p_{SW}^m + \breve{t}_n p_W^m + \breve{t}_e p_S^m = 0.$$

If we insert the known weights $p_W^m = -\breve{t}_e/\breve{t}_{c,x}$ and $p_S^m = -\breve{t}_n/\breve{t}_{c,y}$ we will finally obtain

$$p_{SW}^m = \frac{1}{\breve{t}_c}\left[\breve{t}_n \frac{\breve{t}_e}{\breve{t}_{c,x}} + \breve{t}_e \frac{\breve{t}_n}{\breve{t}_{c,y}}\right] = \frac{\breve{t}_e}{\breve{t}_{c,x}} \frac{\breve{t}_n}{\breve{t}_{c,y}} = p_S^m p_W^m.$$

One finds analogously $p_{SE}^m = p_S^m p_E^m$, $p_{NW}^m = p_N^m p_W^m$ and $p_{NE}^m = p_N^m p_E^m$ and the hereby determined matrix-dependent prolongation is identical to (3.10). The test space $\mathcal{S}_k$ influences only the entries of the one-dimensional stencils $(M_k^x)_*$ and $(M_k^y)_*$ and for that reason only the scaling of the condensed stencils $(\breve{T}_k^x)_*$ and $(\breve{T}_k^y)_*$. Hence, the matrix-dependent prolongation and, equivalently, also the operator-dependent prolongation are independent of the choice of $\mathcal{S}_k$.

The above statements will also hold for the case of a non-constant separable convective vector field, i.e. $\vec{b}(x,y) = (b_1(x), b_2(y))$, if the solutions to the homogeneous continuous equations are computed exactly. From the separability of the two-dimensional operator (3.9) the exact solution to these two-dimensional local problems can be computed as the product of the continuous solutions to the one-dimensional problems with respect to the directional parts T_x and T_y [34]. Again, the two-dimensional stencils $(P_{k+1,\mathcal{V}}^k)_*$ and $(T_k)_*$ will not only be scale- but also space-dependent. Eventhough in non-separable cases the above statements do not hold exactly anymore this technique can still be applied to obtain a problem-dependent prolongation. Here, we use the original local operator stencils $(T_k)_*$ instead of the lumped versions (3.12) to compute a correction of the NW-, NE-, SW- and SE-entries of the 9-point molecule $(P_{k+1,\mathcal{V}}^k)_*$. This leads to a non-conforming FE ansatz for the coarse-grid spaces [24, 34]. Besides, there exist refined techniques by de Zeeuw, which use the decomposition of the local operator stencils into its symmetric and skew-symmetric parts, see [58, 59] for details. We finally remark that the prolongations $P_{k+1,\mathcal{S}}^k$ on the test side might be chosen similarly to the one-dimensional case from a bilinear interpolation or an interpolation that depends on the transposed problem. However, the numerical examples in [30] show that interpolations which are dependent on the transposed problem may give rise to unphysical oscillations and corresponding instabilities already for the fine-grid discretizations.

3.4. Choice of the complement bases for 2D problems

For two-dimensional problems and a hierarchical choice for $Q^k_{k+1,\mathcal{V}/\mathcal{S}}$, matrix-dependent prolongations $P^k_{k+1,\mathcal{V}/\mathcal{S}}$ lead only to an approximate block Gaussian elimination of $\tilde{T}_k$. Besides, the hierarchical construction of the complement spaces suffers from bad stability properties (w.r.t. mesh size) of the corresponding decompositions in higher dimensions. This goes hand in hand with the conceptual weakness of the hereby constructed hierarchical smoothers. The main idea to overcome both difficulties within our general Petrov–Galerkin multiscale approach is to choose one complement basis hierarchically but the other one wavelet-like. By the hierarchical choice on one side we will still approximately decouple fine-without-coarse-grid from coarse-grid unknowns if we choose the prolongation on the respective other side in a matrix-dependent way. Furthermore, the matrix-dependent prolongation ensures physically meaningful coarse-grid operators. From the wavelet-like decomposition on the other side we obtain a stabilization of the multiscale setting with respect to mesh size and corresponding stronger multiscale smoothers. As a result $\tilde{T}_{0,ns}$ will be approximately lower or upper triangular and may – we hope robustly – be inverted by a multiplicative method. In this subsection we present the general construction for $Q^k_{k+1,\mathcal{V}}$ on the trial side. The construction of $Q^k_{k+1,\mathcal{S}}$ is analogous.

We start from the direct splittings $\mathcal{V}^{x,j}_k = \mathcal{W}^{x,j}_{k+1} \oplus \mathcal{V}^{x,j}_{k+1}$ and $\mathcal{V}^{y,i}_k = \mathcal{W}^{y,i}_{k+1} \oplus \mathcal{V}^{y,i}_{k+1}$ of the one-dimensional FE spaces along the grid lines $y = y^k_j = const$ and $x = x^k_i = const$ of the fine grid Ω_k, where

$$\begin{aligned}
\mathcal{W}^{x,j}_{k+1} &:= \left\langle \psi^{k+1,x}_{(i,j),\mathcal{V}} : 0 < i < 1/h_k, i \text{ odd} \right\rangle, \quad \mathcal{V}^{x,j}_{k+1} := \left\langle \varphi^{k+1,x}_{(i,j),\mathcal{V}} : 0 < i < 1/h_k, i \text{ even} \right\rangle, \\
\mathcal{W}^{y,i}_{k+1} &:= \left\langle \psi^{k+1,y}_{(i,j),\mathcal{V}} : 0 < j < 1/h_k, j \text{ odd} \right\rangle, \quad \mathcal{V}^{y,i}_{k+1} := \left\langle \varphi^{k+1,y}_{(i,j),\mathcal{V}} : 0 < j < 1/h_k, j \text{ even} \right\rangle.
\end{aligned}$$

This leads to the abstract one-dimensional decompositions $\mathcal{V}^x_k = \mathcal{W}^x_{k+1} \oplus \mathcal{V}^x_{k+1}$ and $\mathcal{V}^y_k = \mathcal{W}^y_{k+1} \oplus \mathcal{V}^y_{k+1}$. Inserting these into the tensor product representation $\mathcal{V}_k = \mathcal{V}^y_k \otimes \mathcal{V}^x_k$ we obtain the following decomposition of $\mathcal{V}_k$

$$\begin{aligned}
\mathcal{V}_k &= (\mathcal{W}^y_{k+1} \oplus \mathcal{V}^y_{k+1}) \otimes (\mathcal{W}^x_{k+1} \oplus \mathcal{V}^x_{k+1}) \\
&= (\mathcal{W}^y_{k+1} \otimes \mathcal{W}^x_{k+1}) \oplus (\mathcal{W}^y_{k+1} \otimes \mathcal{V}^x_{k+1}) \oplus (\mathcal{V}^y_{k+1} \otimes \mathcal{W}^x_{k+1}) \oplus (\mathcal{V}^y_{k+1} \otimes \mathcal{V}^x_{k+1}) \\
&=: \mathcal{W}^{(1,1)}_{k+1} \oplus \mathcal{W}^{(0,1)}_{k+1} \oplus \mathcal{W}^{(1,0)}_{k+1} \oplus \mathcal{V}_{k+1}.
\end{aligned}$$

We have the following natural basis representations for the different parts of the complement space $\mathcal{W}_{k+1} = \mathcal{W}^{(1,1)}_{k+1} \oplus \mathcal{W}^{(0,1)}_{k+1} \oplus \mathcal{W}^{(1,0)}_{k+1}$

$$\begin{aligned}
\mathcal{W}^{(1,1)}_{k+1} &= \left\langle \psi^{k+1,(1,1)}_{(i,j),\mathcal{V}} : \psi^{k+1,(1,1)}_{(i,j),\mathcal{V}}(x,y) = \psi^{k+1,x}_{(i,j),\mathcal{V}}(x)\psi^{k+1,y}_{(i,j),\mathcal{V}}(y),\ 0 < i,j < 1/h_k, i \text{ odd}, j \text{ odd} \right\rangle, \\
\mathcal{W}^{(0,1)}_{k+1} &= \left\langle \psi^{k+1,(0,1)}_{(i,j),\mathcal{V}} : \psi^{k+1,(0,1)}_{(i,j),\mathcal{V}}(x,y) = \psi^{k+1,x}_{(i,j),\mathcal{V}}(x)\varphi^{k+1,y}_{(i,j),\mathcal{V}}(y),\ 0 < i,j < 1/h_k, i \text{ even}, j \text{ odd} \right\rangle, \\
\mathcal{W}^{(1,0)}_{k+1} &= \left\langle \psi^{k+1,(1,0)}_{(i,j),\mathcal{V}} : \psi^{k+1,(1,0)}_{(i,j),\mathcal{V}}(x,y) = \varphi^{k+1,x}_{(i,j),\mathcal{V}}(x)\psi^{k+1,y}_{(i,j),\mathcal{V}}(y),\ 0 < i,j < 1/h_k, i \text{ odd}, j \text{ even} \right\rangle.
\end{aligned}$$

The basis functions of the above complement spaces can also be characterized by local two-dimensional stencils. These result from the dyadic products of local one-dimensional prolongation molecules and wavelet-like detail filters for the two directions.

We obtain for grid points

$$\begin{aligned}
(x_i^k, y_j^k) \in \Omega_k^{(1,1)} \quad &: \quad (Q^{k,y}_{k+1,(i,j),\mathcal{V}})_*^t \cdot (Q^{k,x}_{k+1,(i,j),\mathcal{V}})_*, \\
(x_i^k, y_j^k) \in \Omega_k^{(0,1)} \quad &: \quad (Q^{k,y}_{k+1,(i,j),\mathcal{V}})_*^t \cdot (P^{k,x}_{k+1,(i,j),\mathcal{V}})_*, \\
(x_i^k, y_j^k) \in \Omega_k^{(1,0)} \quad &: \quad (P^{k,y}_{k+1,(i,j),\mathcal{V}})_*^t \cdot (Q^{k,x}_{k+1,(i,j),\mathcal{V}})_*.
\end{aligned}$$

A detail filter with respect to the x-direction is chosen for grid points from $\Omega_k^{(0,1)}$ because they are next to coarse-grid points which are located on x-grid lines from Ω_{k+1}. In the same way a detail filter with respect to the y-direction is applied for grid points from $\Omega_k^{(1,0)}$ since they are next to coarse-grid points which are located on y-grid lines from Ω_{k+1}. The complete transformation matrix which characterizes the complement space $\mathcal{W}_{k+1}$ can now be composed from the resulting transformation matrices $Q^{k,(1,1)}_{k+1,\mathcal{V}}$, $Q^{k,(0,1)}_{k+1,\mathcal{V}}$ and $Q^{k,(1,0)}_{k+1,\mathcal{V}}$ to $Q^k_{k+1,\mathcal{V}} = [Q^{k,(1,1)}_{k+1,\mathcal{V}}, Q^{k,(0,1)}_{k+1,\mathcal{V}}, Q^{k,(1,0)}_{k+1,\mathcal{V}}]$.

Unfortunately, tensor product L-spline prewavelets do not lead to robust algorithms even for separable problems. This is probably due to the bad stability and regularity properties of the Haar system in the limit case of pure convection. However, the results of our numerical experiments show that constant one-dimensional standard detail filters, e.g. $(Q^k_{k+1,\mathcal{V}})_* = [-1 \quad \underline{2} \quad -1]/2$ (*pre-prewavelet filter*) or $(Q^k_{k+1,\mathcal{V}})_* = [1 \quad -6 \quad \underline{10} \quad -6 \quad 1]/10$ (*linear prewavelet filter*) and their boundary modifications, lead to more robust methods when used with one-dimensional matrix-dependent prolongation molecules.

4. Problem-dependent algebraic multiscale decompositions

The above tensor product multiscale decompositions are not feasible anymore for problems in complicated domains. For simple non-perturbed problems in complex geometries geometric multiscale decompositions have been proposed in [16, 21, 35, 48, 49]. However, the corresponding constructions are quite involved and the question arises whether they can be avoided. A related question is, whether reasonable multiscale decompositions can still be defined when there is no more underlying physical grid that belongs to the linear system under consideration. A multigrid variant which is especially suited for problems with complex geometries is the class of *algebraic multigrid methods* (AMG) [12, 46, 50]. In contrast to geometric multigrid AMG first selects suitable coarse-grid points based on the concept of algebraically smooth errors with the help of two parameters α and β. After the selection of a coarse grid corresponding prolongations and restrictions are constructed. Note that by the coarse-grid selection the corresponding fine-without-coarse-grid points will be automatically determined. For both, the coarse-grid selection and the computation of prolongations and restrictions, only information from the matrix of the linear system is used.

AMG also provides robust solvers for problems with complicated coefficients of the differential operator, e.g. convection-diffusion equations with a non-separable circular convective vector field [23]. From the multigrid interpretation (2.10) of our multiscale cycle it seems to be quite hopeless to attack such difficult problems on the basis of the above tensor product techniques. Now, because AMG is especially designed for simple smoothers, the natural question arises: Is it possible to construct robust multiscale solvers in our hierarchical sense using AMG-based *algebraic multiscale decompositions*?

An obvious answer to this question is given by an AMG-based hierarchical decomposition. Here, on each level, after selecting the coarse-grid points and computing the basis functions for the coarse-grid spaces with the help of the AMG-prolongation, the complement spaces are determined by fine-grid basis functions located in the fine-without-coarse-grid points. However, for the non-perturbed Laplacian AMG approximately leads back to geometric coarsening and bilinear interpolations. Hence, we do not expect an AMG-based *hierarchical* multiscale solver to be optimal. For the approximate cyclic reduction method [41], which is similar in spirit to AMG-based HBMG, this drawback is also conceded in [43]. Of course, a corresponding hierarchical W-cycle would stabilize the method similarly to an AMLI approach [2, 3], but for complicated problems it can be much too expensive. This is due to the fact that the number of coarse-grid points does not necessarily decay in a geometric fashion since the coarse grids are selected algebraically.

Therefore, we finally come to the question of wavelet-like stabilizations of AMG-based hierarchical bases. The explicit construction of AMG-based prewavelets by orthogonalization with respect to the next coarser scale is quite involved. Furthermore, the success of such AMG-based prewavelet decompositions is questionable from our experience with the explicitly constructed L-spline prewavelet decompositions. One might also try the lifting technique [51] but this construction suffers from similar difficulties in our case. In contrast to this, our proposal is very simple and can be derived from the following elementary observation:

Consider a trivial one-dimensional convection-diffusion problem like (3.1), now on $\Omega =]0,1[$ together with appropriate boundary conditions. For discretization we use a simple upwinding [45]. In the case of vanishing convection $b = 0$ the 3-point operator stencil equals the pre-prewavelet filter $[-1 \quad \underline{2} \quad -1]/2$ up to a scaling factor. The same is true for the opposite limit case $b = \infty$ of pure convection, where the explicitly constructed L-spline prewavelet filters $[0 \quad -1 \quad \underline{1} \quad 0 \quad 0]$ are identical to the pure upwind stencils.

Therefore, we suggest to construct AMG-based wavelet-like filters simply by operator stencils that belong to fine-without-coarse-grid points. We set for $(x_i^k, y_j^k) \in \Omega_k^f$

$$(Q^k_{k+1,(i,j),\mathcal{V}})_* := (T_{k,(i,j)})_*.$$

Compare also [32] for a similar approach which concerns the choice of $P^k_{k+1,\mathcal{V}}$. It is well known that the coarse-grid operator stencils computed by the Galerkin-product $(P^k_{k+1,\mathcal{V}})^t T_k P^k_{k+1,\mathcal{V}}$ enlarge during the algebraic coarsening process [50]. The supports of the now constructed AMG-based wavelet-like basis functions, which we call *AMGlets*, increase correspondingly. To avoid this growth, one can use a threshold parameter γ_P to cut off rows of the AMG-prolongations. After that a rescaling has to be performed [50]. Besides, we add the smallest entries of the resulting operator stencils to the central entry with help of a further parameter γ_Q (lumping). The parameter γ_Q limits the relative size of entries which will be discarded, and we finally obtain the matrices $Q^k_{k+1,\mathcal{V}}$ from

$$(Q^k_{k+1,(i,j),\mathcal{V}})_* = [(T_{k,(i,j)})_*]_{\gamma_Q} \quad \text{for} \quad (x_i^k, y_j^k) \in \Omega_k^f, \tag{4.14}$$

where $[\quad]_{\gamma_Q}$ indicates the lumping procedure.

In summary, we propose the following procedure for determining the two-scale transformations $P^k_{k+1,\mathcal{V}}$ and $Q^k_{k+1,\mathcal{V}}$ for a wavelet-like decomposition of the trial side from the respective fine-grid operator T_k and AMG(α,β):

1. coarse-grid selection by AMG(α,β) using T_k;
2. computation of the AMG prolongation $P^k_{k+1,\mathcal{V}}$;
3. γ_P-thresholding and rescaling of $P^k_{k+1,\mathcal{V}}$;
4. computation of $Q^k_{k+1,\mathcal{V}}$ from (4.14).

Our implementation of AMG is based on [23]. However, we use the refined *standard prolongation* proposed in [50].

5. Numerical examples

In this section we present the results from numerical experiments with our multiplicative multiscale methods motivated and discussed in the previous sections for examples with increasing difficulty. We consider the following four convection-diffusion operators together with homogeneous Dirichlet boundary conditions on the unit-square:

$$\begin{aligned} T^\alpha &:= -\Delta + b\,\partial_x, & T^\beta &:= -\Delta + \frac{b}{\sqrt{2}}\,(\partial_x + \partial_y)\,, \\ T^\gamma &:= -\Delta + b\cos\omega\,\partial_x + b\sin\omega\,\partial_y, & T^\delta &:= -\Delta - 4b\,x(x-1)(1-2y)\,\partial_x \\ & & &\quad + 4b\,y\,(y-1)(1-2x)\,\partial_y, \end{aligned}$$

where b is a constant ($b = 0, 10^1, ..., 10^6, 10^8$). The angle $\omega \in [0, \frac{\pi}{2}]$ describes for T^γ the direction of the convective vector field with respect to the x-axis.

5.1. Test environment

As discretization we use a five-point finite difference stencil with simple upwinding to obtain stable discretizations already with respect to the finest grid Ω_0 [28]. This can be interpreted as a discretization which results from a Petrov–Galerkin FE method with operator-dependent trial and test spaces $\mathcal{V}_0$ and $\mathcal{S}_0$. For reasons of numerical stability we scale the above operators in our computations by a factor $\varepsilon = 1/b$ for $b \neq 0$.

If not stated differently we apply standard multiscale V- or W-cycles or improved multiscale V-cycles (c.f. remark v) from Section 2.3) together with a complete decomposition of the spaces $\mathcal{V}_0$ and $\mathcal{S}_0$. Hence, the coarsest grid consists of one interior grid point only. The approximation $\breve{A}_k^{-1}$ of A_k^{-1} used in our experiments is computed by a forward- and backward-substitution from an incomplete factorization (ILU(0)) of the blocks A_k ($k = 1, ...l$).

We start our iterations with a normalized random vector u_0^0 and consider without loss of generality the discrete solution $u_0 = \mathbf{0}$ to a zero right hand side $f_0 = \mathbf{0}$ on the finest grid Ω_0 [29].

Our measure for the speed of convergence is

$$\varrho_{it,1} := \left(\frac{||e_0^{it}||_0}{||e_0^1||_0}\right)^{\frac{1}{it-1}}, \tag{5.1}$$

i.e. the average error reduction within it steps with respect to the discrete l^2-norm [29]. We start to measure after one initial iteration to be free of effects from the starting vector. The stopping criterion is 10^{-10} for the norm of the absolute error. If this has not been reached with 30 iterations, the iteration is also stopped.

5.2. Computational costs

Let $\mathcal{C}_S$, $\mathcal{C}_R$ and $\mathcal{C}_C$ be the normalized work counts for the multiscale smoothing, residual restriction and correction steps within the equivalent multigrid implementation of our standard multiscale cycles. If we replace T_{k-1} by $\breve{S}_{k-1}$ for $k > 1$ we will obtain the corresponding work counts $\mathcal{C}_{S,imp}$, $\mathcal{C}_{R,imp}$ for the improved multiscale cycles using (2.12). Then, in the case of geometric multiscale decompositions ($N_k \leq N_{k-1}/4$) and coarsest scale l, we have the following standard cost estimates for multigrid-like implementations of our proposed cycles [28]:

- Standard multiscale V(1,1)-cycle: $\frac{4}{3}\mathcal{C}_{std} + \mathcal{O}((1/4)^l)$,
- Improved multiscale V(1,1)-cycle: $\mathcal{C}_{std} + \frac{1}{3}\mathcal{C}_{imp} + \mathcal{O}((1/4)^l)$,
- Standard multiscale W(1,1)-cycle: $2\mathcal{C}_{std} + \mathcal{O}((1/2)^l)$,

where $\mathcal{C}_{std} := 2\mathcal{C}_S + \mathcal{C}_R + \mathcal{C}_C$ and $\mathcal{C}_{imp} := 2\mathcal{C}_{S,imp} + \mathcal{C}_{R,imp} + \mathcal{C}_C$. From the representation

$$\begin{aligned} \breve{S}_{k-1} &= T_{k-1} - C_{k-1}\breve{A}_{k-1}^{-1}B_{k-1} \\ &= T_{k-1} - (P_{k-1,\mathcal{S}}^{k-2})^t T_{k-2} Q_{k-1,\mathcal{V}}^{k-2} \breve{A}_{k-1}^{-1} (Q_{k-1,\mathcal{S}}^{k-2})^t T_{k-2} P_{k-1,\mathcal{V}}^{k-2} \end{aligned}$$

one sees that the application of $\breve{S}_{k-1}$ within the improved cycles is at least by a factor of nine more expensive than the application of T_{k-1} alone. Here, we use 9-point stencils and consider only the costs for the application of T_{k-1} and T_{k-2}. Hence, the improved versions are more expensive than a standard multiscale W-cycle.

Regarding the costs for our AMG-based multiscale cycles there are unfortunately no predictive cost analyses available. Therefore, we define as *grid* and *non-standard operator complexities* (adapted to an AMG-based non-standard form) the two numbers

$$C^g := \frac{\sum_{k=0}^{l} \dim(\mathcal{V}_k)}{\dim(\mathcal{V}_0)} \quad \text{and} \quad C_{ns}^o := \frac{\mathrm{nnz}(T_0) + \sum_{k=1}^{l} \left[\mathrm{nnz}(A_k) + \mathrm{nnz}(T_k)\right]}{\mathrm{nnz}(T_0)},$$

where $\mathrm{nnz}(M)$ denotes the number of nonzero entries of a matrix M. Grid complexity C^g gives an accurate measure of the total storage required for the residual and correction vectors with respect to the different scales. Non-standard operator complexity C_{ns}^o indicates the total storage space necessary for the blocks of the non-standard form as well as for the fine-grid operator that are used within the multiscale cycles. It generalizes the well known AMG operator complexity [12]. Furthermore, C_{ns}^o gives a measure of the work involved in the solve phase in the case of multiscale V(1,1)-cycles. Here, we assume that the costs for the detail correction (essentially ILU(0)(A_k)) and for the computation of the residuals are proportional to the number of nonzero entries of the corresponding blocks of the non-standard form.

5.3. Test example T^α: Axis oriented convection

If we choose for both decompositions on the trial and test side the classical hierarchical basis transformation [57], the Galerkin-like multiscale V(1,1)-cycle will soon break down when b increases. A reason for this is that the resulting coarse-grid operators become more and more unstable (w.r.t. convection) and physically meaningless. The hereby computed "corrections" finally destroy the convergence of the method [24, 59]. Therefore, we apply our matrix-dependent coarsening strategy from Section 3. Table 1 shows the average error reduction rates for a standard multiscale V(1,1)-cycle. Here, we use a matrix-dependent hierarchical

TABLE 1: *Example T^α: Standard multiscale V(1,1)-cycle, average error reduction $\varrho_{it,1}$.*

$h_0^{-1} \backslash b$	0.0	10^1	10^2	10^3	10^4	10^5	10^6	10^8
8	0.29	0.38	0.58	0.56	0.58	0.63	0.64	0.64
16	0.36	0.51	0.71	0.78	0.79	0.80	0.78	0.79
32	0.52	0.57	0.81	0.87	0.87	0.87	0.88	0.88
64	0.60	0.67	0.85	0.91	0.93	0.93	0.93	0.94
128	0.64	0.69	0.89	0.93	0.93	0.95	0.96	0.96
256	0.69	0.69	0.90	0.93	0.94	0.96	0.97	0.97
512	0.72	0.75	0.86	0.93	0.94	0.96	0.98	0.98
1024	0.75	0.76	0.88	0.92	0.96	0.95	0.98	0.99

$P^k_{k+1,\mathcal{V}}$: matrix-dependent $Q^k_{k+1,\mathcal{V}}$: hierarchical filter [1]
$P^k_{k+1,\mathcal{S}}$: bilinear $Q^k_{k+1,\mathcal{S}}$: hierarchical filter [1]

TABLE 2: *Example T^α: Standard multiscale V(1,1)-cycle, average error reduction $\varrho_{it,1}$.*

$h_0^{-1} \backslash b$	0.0	10^1	10^2	10^3	10^4	10^5	10^6	10^8
8	0.17	0.13	0.08	0.07	0.07	0.07	0.07	0.07
16	0.20	0.16	0.11	0.08	0.08	0.08	0.08	0.08
32	0.24	0.19	0.14	0.11	0.08	0.07	0.07	0.08
64	0.25	0.21	0.17	0.13	0.09	0.08	0.07	0.07
128	0.26	0.23	0.19	0.17	0.11	0.08	0.08	0.08
256	0.27	0.24	0.20	0.19	0.16	0.10	0.08	0.08
512	0.27	0.25	0.21	0.21	0.21	0.13	0.09	0.08
1024	0.28	0.26	0.21	0.22	0.24	0.19	0.11	0.08

$P^k_{k+1,\mathcal{V}}$: matrix-dependent $Q^k_{k+1,\mathcal{V}}$: linear prewavelet filter
$P^k_{k+1,\mathcal{S}}$: bilinear $Q^k_{k+1,\mathcal{S}}$: hierarchical filter [1]

decomposition of the trial space together with the classical hierarchical decomposition of the test space. The rates increase slightly with increasing problem size, which is already clear from the behavior of the classical HBMG method with respect to the non-perturbed equation (c.f. the first column, which is the Poisson problem with $b = 0$). However, $\varrho_{it,1}$ grows also with increasing size of b. This is somewhat surprising since the coarse-grid operators by our method are identical to those of robust multigrid iterations. The same behavior can even be observed if the blocks A_k are inverted exactly. We realized this by applying an iterative solver to invert these blocks. We omit the corresponding results. Hence, we see from the multigrid interpretation (2.10) of the hierarchical multiscale method that the loss of robustness must be due to a principal defect of the hierarchical smoother.

However, we will obtain a fairly robust method, if we use a wavelet-like decomposition on the trial side instead of the hierarchical one. The wavelet-like decomposition is based on the classical linear prewavelet filter $[1 \; -6 \;\; \underline{10} \; -6 \;\; 1]/10$ and its modifications $(Q^k_{k+1,\mathcal{V},w})_* = [\underline{9} \;\; -6 \;\; 1]/10$ and $(Q^k_{k+1,\mathcal{V},e})_* = [1 \;\; -6 \;\; \underline{9}]/10$ at the western and eastern boundaries. The robustness of the method is clear from the numbers displayed in Table 2. One observes that the rates even decrease with increasing size of b. Note that we obtained similar results in further experiments when using tensor product L-spline prewavelets.

TABLE 3: *Example T^β: Standard multiscale V(1,1)-cycle, average error reduction $\varrho_{it,1}$.*

$h_0^{-1} \backslash b$	0.0	10^1	10^2	10^3	10^4	10^5	10^6	10^8
8	0.13	0.07	0.09	0.14	0.15	0.15	0.16	0.15
16	0.15	0.09	0.13	0.22	0.24	0.25	0.24	0.24
32	0.16	0.10	0.15	0.27	0.29	0.29	0.29	0.29
64	0.17	0.12	0.17	0.29	0.32	0.32	0.32	0.32
128	0.17	0.13	0.18	0.32	0.36	0.36	0.37	0.37
256	0.17	0.14	0.19	0.35	0.42	0.42	0.42	0.43
512	0.17	0.14	0.20	0.38	0.50	0.53	0.53	0.53
1024	0.17	0.14	0.20	0.40	0.62	0.67	0.68	0.68

$P^k_{k+1,\mathcal{V}}$: matrix-dependent $Q^k_{k+1,\mathcal{V}}$: pre-prewavelet filter
$P^k_{k+1,\mathcal{S}}$: bilinear $Q^k_{k+1,\mathcal{S}}$: hierarchical filter [1]

TABLE 4: *Example T^β: Improved multiscale V(1,1)-cycle, average error reduction $\varrho_{it,1}$.*

$h_0^{-1} \backslash b$	0.0	10^1	10^2	10^3	10^4	10^5	10^6	10^8
8	0.13	0.06	0.08	0.10	0.10	0.10	0.10	0.10
16	0.14	0.07	0.10	0.13	0.13	0.13	0.13	0.13
32	0.15	0.09	0.12	0.15	0.16	0.16	0.16	0.16
64	0.15	0.10	0.12	0.17	0.18	0.18	0.18	0.18
128	0.15	0.11	0.13	0.19	0.20	0.20	0.20	0.20
256	0.15	0.11	0.13	0.20	0.24	0.25	0.25	0.25
512	0.15	0.12	0.14	0.21	0.32	0.33	0.33	0.33
1024	0.15	0.12	0.14	0.24	0.41	0.48	0.49	0.49

$P^k_{k+1,\mathcal{V}}$: matrix-dependent $Q^k_{k+1,\mathcal{V}}$: pre-prewavelet filter
$P^k_{k+1,\mathcal{S}}$: bilinear $Q^k_{k+1,\mathcal{S}}$: hierarchical filter [1]

TABLE 5: *Example T^β: Standard multiscale W(1,1)-cycle, average error reduction $\varrho_{it,1}$.*

$h_0^{-1} \backslash b$	0.0	10^1	10^2	10^3	10^4	10^5	10^6	10^8
8	0.07	0.06	0.08	0.13	0.15	0.15	0.15	0.15
16	0.08	0.07	0.11	0.23	0.24	0.25	0.24	0.24
32	0.08	0.08	0.11	0.25	0.29	0.29	0.30	0.28
64	0.08	0.08	0.09	0.24	0.30	0.31	0.30	0.31
128	0.08	0.08	0.08	0.20	0.30	0.31	0.31	0.31
256	0.08	0.08	0.08	0.14	0.28	0.31	0.31	0.31
512	0.08	0.08	0.08	0.10	0.26	0.30	0.30	0.30
1024	0.08	0.08	0.08	0.09	0.20	0.28	0.29	0.29

$P^k_{k+1,\mathcal{V}}$: matrix-dependent $Q^k_{k+1,\mathcal{V}}$: pre-prewavelet filter
$P^k_{k+1,\mathcal{S}}$: bilinear $Q^k_{k+1,\mathcal{S}}$: hierarchical filter [1]

5.4. Test example T^β: Diagonally oriented convection

For the case of the diagonally oriented convective vector field we directly present our results for the classical hierarchical basis decomposition on the test and the matrix-dependent prolongation together with a wavelet-like decomposition on the trial side. Here, the decomposition of the trial space is based on the smaller pre-prewavelet filter $[-1 \quad \underline{2} \quad -1]/2$ and its boundary modifications $(Q^k_{k+1,\mathcal{V},w})_* = [\underline{2} \quad -1 \quad 0]/2$ and $(Q^k_{k+1,\mathcal{V},e})_* = [0 \quad -1 \quad \underline{2}]/2$. Linear prewavelet filters produce similar results, but tensor product L-spline prewavelet filters

TABLE 6: *Example T^γ: Standard multiscale W(1,1)-cycle, average error reduction $\varrho_{it,it-4}$.*

$\omega\backslash\, b$	0.0	10^1	10^2	10^3	10^4	10^5	10^6	10^8
				$h_0 = 1/256$				
0	0.09	0.09	0.09	0.09	0.09	0.04	0.01	0.01
$\pi/8$	0.09	0.09	0.09	0.13	0.24	0.26	0.26	0.26
$\pi/4$	0.09	0.09	0.09	0.15	0.30	0.33	0.33	0.33
$3\pi/8$	0.09	0.09	0.09	0.13	0.23	0.25	0.25	0.25
$\pi/2$	0.09	0.09	0.09	0.08	0.08	0.03	0.01	0.01
				$h_0 = 1/512$				
0	0.09	0.09	0.09	0.09	0.09	0.07	0.02	0.01
$\pi/8$	0.09	0.09	0.09	0.11	0.19	0.22	0.23	0.23
$\pi/4$	0.09	0.09	0.09	0.11	0.28	0.33	0.33	0.33
$3\pi/8$	0.09	0.09	0.09	0.10	0.21	0.24	0.25	0.25
$\pi/2$	0.09	0.09	0.09	0.08	0.09	0.06	0.02	0.01
				$h_0 = 1/1024$				
0	0.09	0.09	0.09	0.09	0.09	0.11	0.05	0.01
$\pi/8$	0.09	0.09	0.09	0.09	0.18	0.19	0.22	0.21
$\pi/4$	0.09	0.09	0.09	0.10	0.23	0.31	0.32	0.32
$3\pi/8$	0.09	0.09	0.09	0.09	0.18	0.23	0.24	0.24
$\pi/2$	0.09	0.09	0.09	0.08	0.09	0.11	0.05	0.01

$P^k_{k+1,\mathcal{V}}$: matrix-dependent $\quad Q^k_{k+1,\mathcal{V}}$: pre-prewavelet filter
$P^k_{k+1,\mathcal{S}}$: bilinear $\quad Q^k_{k+1,\mathcal{S}}$: hierarchical filter [1]

only work for the lower convective range. We omit the corresponding results. From Table 3 we see clearly the stabilizing effect (w.r.t. mesh size) of the wavelet-like decomposition and the standard multiscale V(1,1)-cycle when b is small ($b \leq 10^2$). However, the method is not robust. Now, Table 4 shows that one obtains much better results for an improved multiscale V(1,1)-cycle. Here, the residuals and the corrections of the projected residuals are computed with the help of Schur-complement approximations that use incomplete factorizations of the blocks A_k ($k = 1, ..., l-1$). Fine-scale features are transported more accurately to coarser scales. However, we have to pay for the improved convergence behavior with substantially larger computational costs. Furthermore, the method still does not seem to be overall robust, which can be seen from the comparison of the results for $h_0 = 1/512$ and $h_0 = 1/1024$ for large b. Therefore, we also applied a standard multiscale W(1,1)-cycle using the same transformations. Then, one obtains a truly robust method which is suggested by Table 5. The repeated recursive calls of coarse-grid corrections produce some extra smoothing which finally implies robustness.

5.5. *Test example T^γ: Dependence on the angle of convection*

To check the dependence on the angle of convection of the convergence behavior of the above multiscale W(1,1)-cycle we apply exactly the same multiscale setting to T^γ as for the last example. We use the classical hierarchical basis transformation for the decomposition of the test spaces and matrix-dependent prolongations together with pre-prewavelet-based decompositions on the trial sides. Here, averages $\varrho_{it,it-4}$ have been derived from the last five iterations to get a better impression of the asymptotic behavior of the iterations. From Table 6 we conclude that one obtains robust solvers also for angles different from $\omega = 0$, $\omega = \frac{\pi}{4}$ and $\omega = \frac{\pi}{2}$ with the multiscale W(1,1)-cycles. The rates for the corresponding axis-oriented cases are of course significantly better compared to other angles.

TABLE 7: *Example T^δ: Standard multiscale V(1,1)-cycle, average error reduction $\varrho_{it,1}$.*

$h_0^{-1} \setminus b$	0.0	10^1	10^2	10^3	10^4	10^5	10^6	10^8
8	0.02	0.03	0.14	0.06	0.02	0.01	0.01	0.01
16	0.08	0.09	0.18	0.23	0.16	0.14	0.17	0.16
32	0.17	0.17	0.26	0.30	0.28	0.22	0.24	0.25
64	0.21	0.30	0.46	0.52	0.37	0.33	0.29	0.28
128	0.35	0.33	0.43	0.46	0.46	0.32	0.40	0.37
256	0.37	0.42	0.51	0.46	0.52	0.43	0.43	0.49
512	0.45	0.41	0.53	0.51	0.59	0.51	0.46	0.60
1024	0.48	0.52	0.57	0.61	0.59	0.63	0.61	0.66

$P^k_{k+1,\mathcal{V}}$: AMG$(0.1, 0.35, 0.0)$ $Q^k_{k+1,\mathcal{V}}$: hierarchical filter [1]
$P^k_{k+1,\mathcal{S}}$: $P^k_{k+1,\mathcal{V}}$ $Q^k_{k+1,\mathcal{S}}$: hierarchical filter [1]

TABLE 8: *Example T^δ: Standard multiscale V(1,1)-cycle, average error reduction $\varrho_{it,1}$.*

$h_0^{-1} \setminus b$	0.0	10^1	10^2	10^3	10^4	10^5	10^6	10^8
8	0.03	0.04	0.06	0.03	0.01	0.01	0.01	0.01
16	0.07	0.07	0.08	0.07	0.10	0.05	0.05	0.05
32	0.10	0.09	0.10	0.12	0.09	0.10	0.10	0.10
64	0.11	0.09	0.10	0.18	0.11	0.10	0.10	0.11
128	0.10	0.10	0.12	0.16	0.14	0.11	0.13	0.12
256	0.13	0.11	0.13	0.15	0.16	0.14	0.14	0.12
512	0.12	0.13	0.12	0.14	0.18	0.18	0.14	0.14
1024	0.13	0.12	0.11	0.25	0.32	0.25	0.17	0.22

$P^k_{k+1,\mathcal{V}}$: AMG$(0.1, 0.35, 0.0)$ $Q^k_{k+1,\mathcal{V}}$: $[(T_{k-1,(i,j)})_*]_{0.125}$
$P^k_{k+1,\mathcal{S}}$: $P^k_{k+1,\mathcal{V}}$ $Q^k_{k+1,\mathcal{S}}$: hierarchical filter [1]

TABLE 9: *Example T^δ: Multigrid V(1,1)-cycle by AMG$(0.1, 0.35, 0.0)$, average error reduction $\varrho_{it,1}$, standard prolongation, pointwise Gauß–Seidel smoother based on the fine to coarse ordering of the unknowns, stopping criterion: Reduction of the l_2-norm of the corresponding residuals below 10^{-10}.*

$h_0^{-1} \setminus b$	0.0	10^1	10^2	10^3	10^4	10^5	10^6	10^8
8	0.04	0.04	0.02	0.02	0.00	0.00	0.00	0.00
16	0.05	0.05	0.04	0.05	0.03	0.03	0.03	0.03
32	0.08	0.07	0.06	0.09	0.05	0.04	0.04	0.04
64	0.09	0.09	0.09	0.12	0.06	0.11	0.05	0.05
128	0.12	0.11	0.11	0.25	0.09	0.08	0.08	0.07
256	0.14	0.14	0.07	0.16	0.12	0.09	0.10	0.10
512	0.15	0.15	0.09	0.22	0.21	0.10	0.10	0.11
1024	0.16	0.17	0.11	0.18	0.38	0.14	0.13	0.12

5.6. Test example T^δ: Circular convection

For the challenging problem T^δ of a circular convection none of our multiscale solvers works satisfactory on the basis of the geometric multiscale decompositions from Section 3. Therefore, we switch to solvers based on the algebraic multiscale decompositions that were presented in Section 4. Due to the algebraic selection of the coarse-grid points we only apply multiscale

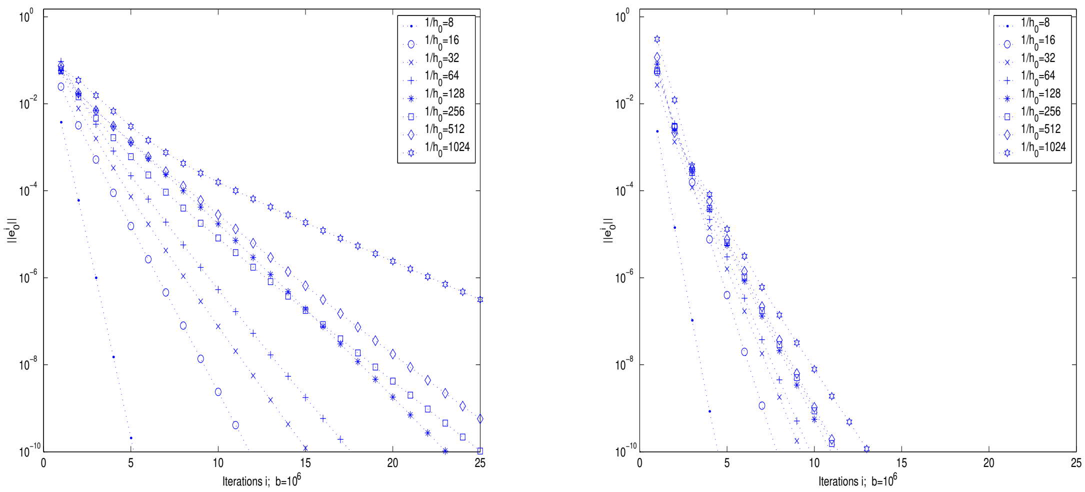


FIGURE 1: *Example T^δ: Error norms w.r.t. standard multiscale V(1,1)-cycles, different fine grid sizes h_0 and $b = 10^6$, AMG-based hierarchical decomposition (left) and stabilized decomposition (right).*

TABLE 10: *Example T^δ: Non-standard operator complexity $\mathcal{C}^o_{ns}$.*

$h_0^{-1} \backslash b$	0.0	10^1	10^2	10^3	10^4	10^5	10^6	10^8
8	2.93	2.94	3.44	3.18	3.18	3.18	3.18	3.18
16	3.37	3.36	5.38	4.37	4.18	4.17	4.17	4.17
32	3.77	3.78	7.12	5.51	5.12	5.29	5.23	5.23
64	4.11	4.03	5.20	8.88	5.52	5.64	5.62	5.60
128	4.17	4.17	5.05	10.87	6.33	6.20	6.20	6.22
256	4.34	4.32	4.57	11.81	7.59	6.68	6.66	6.66
512	4.35	4.44	4.44	9.78	10.54	7.19	7.13	7.15
1024	4.38	4.41	4.33	6.58	12.15	7.47	7.60	7.83

$P^k_{k+1,\mathcal{V}}$: AMG(0.1, 0.35, 0.0) $\quad Q^k_{k+1,\mathcal{V}}$: hierarchical filter [1]
$P^k_{k+1,\mathcal{S}}$: $P^k_{k+1,\mathcal{V}}$ $\quad Q^k_{k+1,\mathcal{S}}$: hierarchical filter [1]

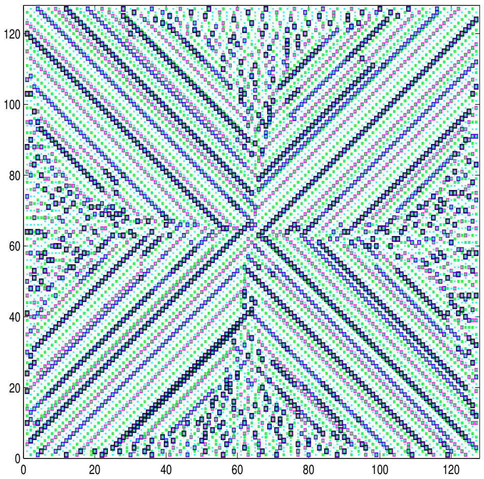
TABLE 11: *Example T^δ: Non-standard operator complexity $\mathcal{C}^o_{ns}$.*

$h_0^{-1} \backslash b$	0.0	10^1	10^2	10^3	10^4	10^5	10^6	10^8
8	3.94	3.95	4.11	3.72	3.72	3.72	3.72	3.72
16	4.94	4.92	6.51	5.09	4.93	4.91	4.91	4.91
32	5.66	5.67	8.66	6.53	5.98	6.16	6.11	6.11
64	6.22	6.15	7.09	10.24	6.56	6.62	6.62	6.61
128	6.48	6.46	7.14	12.62	7.43	7.24	7.23	7.24
256	6.73	6.70	6.80	13.80	8.84	7.76	7.76	7.74
512	6.79	6.86	6.80	11.91	12.11	8.31	8.24	8.25
1024	6.84	6.87	6.76	8.77	14.04	8.65	8.73	8.94

$P^k_{k+1,\mathcal{V}}$: AMG(0.1, 0.35, 0.0) $\quad Q^k_{k+1,\mathcal{V}}$: $[(T_{k-1,(i,j)})_*]_{0.125}$
$P^k_{k+1,\mathcal{S}}$: $P^k_{k+1,\mathcal{V}}$ $\quad Q^k_{k+1,\mathcal{S}}$: hierarchical filter [1]

TABLE 12: *Example T^δ: Grid complexity $\mathcal{C}^g$ and number of scales l produced by AMG* $(0.1, 0.35, 0.0)$.

$h_0^{-1} \backslash b$	0.0	10^1	10^2	10^3	10^4	10^5	10^6	10^8
8	1.88 (5)	1.94 (6)	2.20 (6)	2.22 (6)	2.22 (6)	2.22 (6)	2.22 (6)	2.22 (6)
16	1.73 (6)	1.74 (6)	2.13 (9)	2.25 (8)	2.26 (8)	2.26 (8)	2.26 (8)	2.26 (8)
32	1.70 (7)	1.71 (8)	2.02 (9)	2.22 (11)	2.39 (11)	2.36 (10)	2.36 (10)	2.36 (10)
64	1.69 (8)	1.69 (9)	1.81 (13)	2.22 (14)	2.25 (12)	2.31 (13)	2.32 (12)	2.32 (12)
128	1.68 (9)	1.68 (11)	1.74 (15)	2.17 (16)	2.30 (16)	2.34 (15)	2.37 (16)	2.37 (15)
256	1.67 (10)	1.67 (11)	1.69 (15)	2.07 (19)	2.26 (17)	2.32 (17)	2.32 (16)	2.34 (18)
512	1.67 (11)	1.67 (12)	1.67 (16)	1.90 (22)	2.20 (19)	2.33 (21)	2.33 (21)	2.32 (22)
1024	1.67 (13)	1.67 (13)	1.67 (17)	1.75 (23)	2.15 (23)	2.29 (23)	2.32 (23)	2.34 (24)



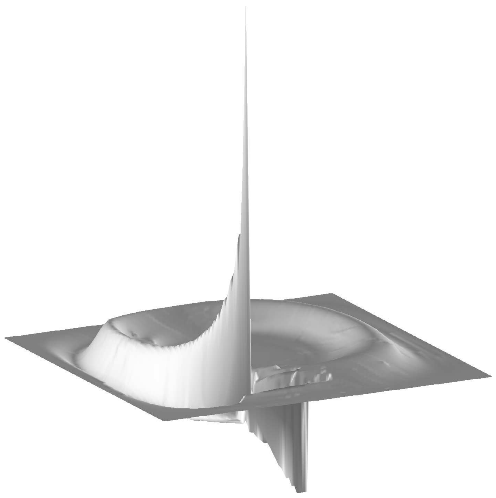


FIGURE 2: *Example T^δ: Left picture: The first five coarse-grids selected by AMG$(0.1, 0.35, 0.0)$ (coarser grids correspond to thicker points), $h_0 = 128$ and $b = 10^6$; right picture: AMGlet generated by the stabilization of the AMG-based hierarchical basis on the trial side, same parameters.*

V(1,1)-cycles to obtain iteration schemes that can still be implemented with linear complexity. In all experiments we use the same prolongations for the transformation matrices $P^k_{k+1,\mathcal{V}}$ and $P^k_{k+1,\mathcal{S}}$ which result from AMG with parameters $(\alpha, \beta, \gamma_P)$. The decompositions on the trial and test side are chosen hierarchically or wavelet-like and the detail corrections are always computed via incomplete factorizations of the resulting blocks A_k $(k = 1, ..., l)$.

Table 7 shows the average error reduction rates from standard multiscale V(1,1)-cycles based on a hierarchical decomposition of both the trial and test side resulting from AMG$(0.1, 0.35, 0.0)$. From the first column we see that the method is not independent of the fine-grid mesh size. This is typical for HBMG-like methods. The overall convergence behavior of the method is neither optimal nor robust. Of course, using the cycle only as a preconditioner for a Krylov subspace method may soften the dependence on h_0 [43].

Table 8 shows the benefits of the wavelet-like stabilization as described in Section 4 in a striking way. Here, only the transformations $Q^k_{k+1,\mathcal{V}}$ have been changed according to formula (4.14) where we used $\gamma_Q = 0.125$ as cut-off parameter. By this choice we obtain an AMGlet decomposition of the trial space that leads to an overall robust convergence behavior. For $h_0 \leq 1/512$ the average error reduction rates stay below 0.2 uniformly w.r.t. b. For $h_0 = 1/1024$ we observe a slight growth and a maximal rate of 0.32 when $b = 10^4$. Note however that also the ordinary AMG algorithm seems to "feel" the difficulty. This can be seen from Table 9, which shows the average error reduction rates for an ordinary multigrid V(1,1)-cycle based on AMG$(0.1, 0.35, 0.0)$ and a pointwise Gauß-Seidel smoother. Here, we obtain $\varrho_{it,1} = 0.38$ for the corresponding problem. It is clear that our method can not do essentially better than ordinary AMG. The stabilizing effect of the AMGlet approach is also evident from Figure 1 which shows the error norms from iterations by the purely hierarchical (left) and the stabilized method (right) for different fine-grid sizes h_0 with $b = 10^6$. The choice of the complement bases has no influence on the AMG coarsening process. Hence, AMG produces for both kinds of decompositions the same number of scales and one obtains the same grid complexities. They behave quite reasonable which can be observed from Table 12. The non-standard operator complexities are in the lower convective range about 56% higher for the AMGlet decomposition of the trial space compared to those which result from the hierarchical decompositions on both sides. This can be seen from the comparison of the last three rows of Tables 10 and 11. For $b > 10^3$ the difference is much smaller, i.e. only about 16%. If we compare this with the respective error reduction rates, the AMGlet decomposition – although expensive – will be preferable. However, it is not completely clear from Table 10 that the solver is indeed robust.

On the left of Figure 2 the hierarchy of grids is displayed which is produced by AMG$(0.1, 0.35, 0.0)$ after five coarsening steps for the case $b = 10^6$ and mesh size $h_0 = 1/128$ of the fine grid. Finally, it is interesting to take a look at the complementary basis functions which are implicitly generated by our approach. For the same parameters as above an AMGlet is shown on the right of Figure 2. Apparently the AMGlet follows the convective flow and takes positive *and* negative values as a wavelet should.

6. Concluding remarks

The present work can be seen as a practical answer to the following question which extends the central topic of [52]:

> *How far could one go in developing* robust *multiscale techniques for convection-diffusion problems exploiting direct decompositions of the discretization spaces?*

According to the three major problems which have been raised in the introduction with respect to the above question (dependence of h_0, dependence of b and stability of the coarse-grid problems involved) our answer consists of three main ingredients.

First, the problem-dependent choice of the prolongation operators $P^k_{k+1,\mathcal{V}/\mathcal{S}}$ within our Petrov-Galerkin multiscale approach is indispensable to obtain physically reasonable coarse-grid problems, which are used for the multiscale solution process. Considering the challenging test example T^δ of a circular convection we even had to resort to techniques from AMG. Second, starting from hierarchical decompositions of the trial and test side a change to a wavelet-like complement basis on one side is necessary to enhance the stability of the decomposition with respect to mesh size and to repair (at least to some extend) the defects of the hierarchical smoothers. This is even true for the essentially one-dimensional problem T^α. However note that the resulting multiscale smoothers still do not fulfill the criterion of robust smoothers from robust multigrid. The third reason why the proposed method works, is the combination of the hierarchical decomposition on the test side along with the problem-dependent coarsening on the trial side. As we have seen this is equivalent to a decoupling of the transformed linear system for one-dimensional problems. In the two-dimensional case the decoupling relation from this choice is only approximate. Further experiments showed that it is nevertheless necessary for non-symmetric problems.

By a Petrov–Galerkin multiscale concept it is possible to bring together all of the above ideas which are of course closely intertwined. Therefore, our approach extends by far other existing methods which also rely on direct subspace decompositions, e.g. those in [2, 3, 22, 41, 42]. The wavelet based direct solver in [22] does not work for our singularly perturbed problems since it does not apply problem adapted multiscale decompositions. Many other traditional wavelet approaches for solving partial differential equations face the same difficulty [14, 15, 22, 40, 44, 54]. This is also true for the AMLI method which starts from classical hierarchical decompositions and employs a stabilization by inner iterations to guarantee optimal preconditioning. The generalized hierarchical basis multigrid methods in [7, 17] are based on matrix dependent prolongations and hierarchical decompositions within a Petrov–Galerkin multiscale setting. However, they suffer from the principal weakness of the corresponding hierarchical smoothers which can also be seen from the numerical results in [36]. Although the method of approximate cyclic reduction [41, 42] applies algebraically defined hierarchical subspace decompositions, the quality of the resulting preconditioner still seems to be slightly dependent on the mesh size of the finest grid. Our new AMGlet approach, which is also based on purely algebraical principles, provides here a simple remedy. Furthermore, with this approach we are able to leave the context of tensor product constructions which only work well for non-separable problems. This paths the way from model problems to problems which are relevant in practice.

We think that due to all the interrelations and especially due to the conceptual weakness of the corresponding multiscale smoothers it is much harder to design robust multiscale methods starting from direct decompositions than from a true generating system approach as for example in multigrid [25, 26]. This is true, even if one has a clear separation of the correction steps by direct decompositions, which might be advantageous for developing a satisfying convergence theory in the future.

ACKNOWLEDGEMENT

We thank M. A. Schweitzer for stimulating discussions and for providing us his AMG-code.

REFERENCES

1. R. E. Alcouffe, A. Brandt, J. E. Dendy and J. W. Painter. The multi-grid method for the diffusion equation with strongly discontinuous coefficients. *SIAM J. Sci. Comput.*, 2, 430–454, 1981.
2. O. Axelsson and P. Vassilevski. Algebraic multilevel preconditioning methods. I. *Numer. Math.*, 56, 157–177, 1989.
3. O. Axelsson and P. Vassilevski. Algebraic multilevel preconditioning methods. II. *SIAM J. Numer. Anal.*, 27, 1569–1590, 1990.
4. R. E. Bank. Hierarchical bases and the finite element method. *Acta Numerica*, 5, 1–43, 1996.
5. R. E. Bank and M. Benbourenane. The hierarchical basis multigrid method for convection-diffusion equations. *Numer. Math.*, 61, 7–37, 1992.

6. R. E. Bank and T. F. Dupont. Analysis of a two-level scheme for solving finite element equations. Technical Report CNA-159, Center for Numerical Analysis, The University of Texas at Austin, 1980.
7. R. E. Bank and S. Gutsch. Hierarchical basis for the convection-diffusion equation on unstructured meshes. In P. Bjørstad, M. Espedal and D. Keyes, editors, *Ninth International Conference on Domain Decomposition Methods*, 1998.
8. R. E. Bank and J. Xu. The hierarchical basis multigrid method and incomplete lu decomposition. In D. Keyes and J. Xu, editors, *Seventh International Conference on Domain Decomposition Methods*, 1994.
9. R. E. Bank, T. F. Dupont and H. Yserentant. The hierarchical basis multigrid method. *Numer. Math.*, 52, 427–458, 1988.
10. G. Beylkin, R. Coifman and V. Rokhlin. Fast wavelet transforms and numerical algorithms. I. *Comm. Pure and Appl. Math.*, 44, 141–183, 1991.
11. D. Braess. *Finite Elemente.* Springer, Berlin, Heidelberg, New York, 1992.
12. W. L. Briggs, V. E. Henson and S. F. McCormick. *A Multigrid Tutorial.* SIAM, Phiadelphia, 2000.
13. J. M. Carnicer, W. Dahmen and J. M. Peña. Local decomposition of refinable spaces. *Appl. Comput. Harm. Anal.*, 3, 127–153, 1996.
14. T. F. Chan, W. P. Tang and W. L. Wan. Wavelet sparse approximate inverse preconditioners. *BIT*, 37, 644–660, 1997.
15. A. Cohen and R. Masson. Wavelet methods for second-order elliptic problems, preconditioning and adaptivity. *SIAM J. Sci. Comput.*, 21, 1006–1026, 1999.
16. W. Dahmen and R. Schneider. Wavelets on manifolds I: Construction and domain decomposition. Technical Report 149, RWTH Aachen, 1999.
17. W. Dahmen, S. Müller and T. Schlinkmann. On a robust adaptive multigrid solver for convection-dominated problems. Technical Report 171, RWTH Aachen, 1999.
18. I. Daubechies. *Ten Lectures on Wavelets.* SIAM, Phiadelphia, 1992.
19. J. E. Dendy. Black box multigrid. *J. Comput. Phys.*, 48, 366–386, 1982.
20. J. E. Dendy. Black box multigrid for nonsymmetric problems. *Appl. Math. Comput.*, 13, 261–283, 1983.
21. M. Floater and E. Quak. Piecewise linear prewavelets on arbitrary triangulations. *Numer. Math.*, 82, 221–252, 1999.
22. D. L. Gines, G. Beylkin and J. Dunn. LU factorization of non-standard-forms and direct multiresolution solvers. Technical Report 278, Program in Applied Mathematics, University of Colorado at Boulder, 1996.
23. T. Grauschopf, M. Griebel and H. Regler. Additive multilevel-preconditioners based on bilinear interpolation, matrix dependent geometric coarsening and algebraic multigrid coarsening for second order elliptic PDEs. *Appl. Numer. Math.*, 23, 63–96, 1997.
24. M. Griebel. Zur Lösung von Finite-Differenzen- und Finite-Element-Gleichungen mittels der Hierarchischen Transformations-Mehrgitter-Methode. Technical Report 342/4/90 A, TU München, 1990.
25. M. Griebel. Multilevel algorithms considered as iterative methods on semidefinite systems. *SIAM J. Sci. Comput.*, 15, 547–565, 1994.
26. M. Griebel. *Multilevelmethoden als Iterationsverfahren über Erzeugendensystemen.* Teubner, Stuttgart, 1994.
27. M. Griebel and P. Oswald. Tensor product type subspace splitting and multilevel iterative methods for anisotropic problems. *Adv. Comput. Math.*, 4, 171–206, 1995.
28. W. Hackbusch. *Multigrid-Methods and Applications.* Springer, Berlin, Heidelberg, New York, 1985.
29. W. Hackbusch. *Iterative Lösung großer schwachbesetzter Gleichungssysteme.* Teubner, Stuttgart, 1993.
30. A. F. Hegarty, E. O'Riordan and M. Stynes. A comparison of uniformly convergent difference schemes for two-dimensional convection-diffusion problems. *J. Comp. Phys.*, 105, 24–32, 1993.
31. P. W. Hemker, R. Kettler, P. Wesseling and P. M. de Zeeuw. Multigrid methods: Development of fast solvers. *Appl. Math. Comput.*, 13, 311–326, 1983.
32. T. Huckle. Matrix multilevel methods and preconditioning. Technical Report 342/11/98A, Institut für Informatik, TU München, 1998.
33. A. M. Ilin. Differencing scheme for a differential equation with a small parameter affecting the highest derivative. *Math. Notes Acad. Sc. USSR*, 6, 596–602, 1969.
34. F. Kiefer. *Multiskalen-Verfahren für Konvektions-Diffusions Probleme.* PhD thesis, University of Bonn, 2001.
35. U. Kotyczka and P. Oswald. Piecewise linear prewavelets of small support. In C. Chui and L. Schumaker, editors, *Approximation Theory VIII.* World Scientific, Singapore, 1995.
36. S. Le Borne. *Multigrid Methods for Convection-Dominated Problems.* PhD thesis, University of Kiel, 1999.
37. Y. Notay. A robust algebraic preconditioner for finite difference approximations of convection-diffusion equations. Technical Report GANMN 99–01, University of Brussels, 1999.
38. P. Oswald. *Multilevel Finite Element Approximation.* B. G. Teubner, Stuttgart, 1994.
39. C. Pflaum. Fast and Robust Multilevel Algorithms. Habilitation, University of Würzburg, 1998.

40. C. Pflaum. Robust convergence of multilevel algorithms for convection-diffusion equations. *SIAM J. Numer. Anal.*, 37, 443–469, 2000.
41. A. Reusken. On the approximate cyclic reduction preconditioner. Technical Report 144, Institut für Geometrie und Praktische Mathematik, RWTH Aachen, 1997.
42. A. Reusken. Approximate cyclic reduction preconditioning. In W. Hackbusch and G. Wittum, editors, *Multigrid Methods V*, Lecture Notes in Computational Science and Engineering Vol. 3. Proceedings of the Fifth European Multigrid Conference, Springer, Berlin, Heidelberg, New York, 1998.
43. A. Reusken. An algebraic multilevel preconditioner for symmetric positive definite and indefinite problems. In A. Frommer, T. Lippert, B. Meldeke and K. Schilling, editors, *Numerical Challenges in Lattice Quantum Chromodynamics*, Lecture Notes in Computational Science and Engineering Vol. 15. Springer, Berlin, Heidelberg, New York, 2000.
44. A. Rieder, R. O. Wells and X. Zhou. A wavelet approach to robust multilevel solvers for anisotropic elliptic problems. *Appl. Comput. Harm. Anal.*, 1, 355–367, 1994.
45. H.-G. Roos, M. Stynes and L. Tobiska. *Numerical Methods for Singularly Perturbed Equations.* Springer, Berlin, Heidelberg, New York, 1996.
46. J. W. Ruge and K. Stüben. Algebraic multigrid. In S. F. McCormick, editor, *Multigrid Methods.* SIAM, Phiadelphia, 1987.
47. Y. Saad. *Iterative Methods for Sparse Linear Systems.* PWS Publishing, Boston, 1996.
48. R. Stevenson. Piecewise linear (pre-)wavelets on non-uniform meshes. In W. Hackbusch and G. Wittum, editors, *Multigrid Methods V*, Lecture Notes in Computational Science and Engineering Vol. 3. Proceedings of the Fifth European Multigrid Conference, Springer, Berlin, Heidelberg, New York, 1998.
49. R. Stevenson. Stable three-point wavelet bases on general meshes. *Numer. Math.*, 80, 131–158, 1998.
50. K. Stüben. Algebraic multigrid (AMG): An introduction with applications. Technical Report 53, GMD-Forschungszentrum Informationstechnik GmbH, St. Augustin, 1999.
51. W. Sweldens. The lifting scheme: A construction of second generation wavelets. *SIAM J. Math. Anal.*, 29, 511–546, 1997.
52. P. S. Vassilevski. On two ways of stabilizing the hierarchical basis multilevel methods. *SIAM Rev.*, 39, 18–53, 1997.
53. P. S. Vassilevski and J. Wang. Stabilizing the hierachical basis by approximate wavelets. I: Theory. *Numer. Lin. Algebra Appl.*, 4, 103–126, 1997.
54. P. S. Vassilevski and J. Wang. Wavelet-like methods in the design of efficient multilevel preconditioners for elliptic PDEs. In W. Dahmen, A. Kurdila and P. Oswald, editors, *Multiscale Wavelet Methods*, Wavelet Analysis and its Applications Vol. 15. Academic Press, San Diego, 1997.
55. P. S. Vassilevski and J. Wang. Stabilizing the hierachical basis by approximate wavelets, II: Implementation and numerical results. *SIAM J. Sci. Comput.*, 20, 490–514, 1998.
56. J. Xu. Iterative methods by space decomposition and subspace correction. *SIAM Rev.*, 34, 581–613, 1992.
57. H. Yserentant. On the multi-level splitting of finite element spaces. *Numer. Math.*, 49, 379–412, 1986.
58. P. M. de Zeeuw. Matrix-dependent prolongations and restrictions in a black-box multigrid solver. *J. Comput. Appl. Math.*, 33, 1–27, 1990.
59. P. M. de Zeeuw. *Acceleration of Iterative Methods by Coarse Grid Corrections.* PhD thesis, University of Amsterdam, 1997.