Multigrid and Multiscale Decompositions

Wolfgang Dahmen Siegfried Müller* Thomas Schlinkmann†

1 Introduction

Multigrid methodology has caused a major breakthrough in the numerical analysis of partial differential equations. For standard elliptic problems (with well–behaved coefficients) it is by now well understood and known to have asymptotically optimal complexity.

Major challenges are still posed by problems that exhibit strong anisotropies in their coefficients or when ellipticity degenerates. A typical example of model problem type is the *convection-diffusion-reaction* equation

$$\begin{aligned} \mathcal{L}\,u := -\mathrm{div}\,(\boldsymbol{\varepsilon}\,\nabla\,u) + \boldsymbol{\beta}^T\,\nabla\,u + a\,u &= f \qquad \text{in } \Omega \subset \mathrm{I\!R}^d \\ u &= 0 \qquad \text{on } \partial\Omega, \end{aligned} \tag{1}$$

where Ω is a given sufficiently regular domain in $\mathrm{I\!R}^d$, $\boldsymbol{\beta}(\boldsymbol{x}) \in \mathrm{I\!R}^d$, $a(\boldsymbol{x}) \geq 0$ on Ω and $\boldsymbol{\varepsilon}(\boldsymbol{x}) \in \mathrm{I\!R}^{d\times d}$ satisfies

$$\boldsymbol{\xi}^T\,\boldsymbol{\varepsilon}(\boldsymbol{x})\,\boldsymbol{\xi} \geq \alpha\,|\boldsymbol{\xi}|^2, \qquad \boldsymbol{x} \in \overline{\Omega},\ \boldsymbol{\xi} \in \mathrm{I\!R}^d, \tag{2}$$

for some $\alpha > 0$. Equations of this type form core ingredients of more complex systems like Navier–Stokes equations or equations modelling porous media and multi-phase flow problems. But already in the simple form shown above standard multigrid versions would fail to retain their ideal efficiency or even to converge at all when $\|\boldsymbol{\beta}\| \gg \|\boldsymbol{\varepsilon}\|$. Therefore substantial research effort has been documented in the literature in search of remedies. We will have opportunity to comment briefly on some approaches below. Nevertheless, it is fair to say that the problem is still far from being completely solved, in particular, when dealing with more than two spatial variables.

The primary objective of this paper is to outline a general framework for tackling problems of the above type from a somewhat different angle motivated by recent developments in the context of multiscale decomposition techniques and wavelet analysis [CDP96]. A more detailed analysis of the case of dominated convection can be found in [DMS97]. Therefore we will spend some effort on relating these concepts to multigrid methodology. This will also allow us to relate the present setting to existing methods.

The discussion will be illustrated by some preliminary numerical experiments which mainly concern the case of convection-diffusion equations. In general the flow direction $\boldsymbol{\beta}$ as well as the right hand side may depend on u. However, subsequent numerical experiments will be confined at this stage to the simple linear case where, however, the convective term may vary and strongly dominate the diffusive principle part, i.e., $\|\boldsymbol{\beta}\| \gg \|\boldsymbol{\varepsilon}\|$.

*The author is supported by SFB 401 of the DFG (German Research Foundation).
†The author is supported by the DFG in the Graduiertenkolleg "Analyse und Konstruktion in der Mathematik".

2 Some Preliminary Remarks

Difficulties of different types arise depending on the mutual interrelation of the coefficients in (1). Cases of interest are strong fluctuation in $\boldsymbol{\varepsilon}(\boldsymbol{x})$ or strong variation of $a(\boldsymbol{x})/\|\boldsymbol{\varepsilon}(\boldsymbol{x})\|$ and finally, dominating convection $\|\boldsymbol{\beta}(\boldsymbol{x})\| \gg \|\boldsymbol{\varepsilon}(\boldsymbol{x})\|$ for suitable norms $\|\ \|$. Much of the subsequent discussion is motivated by the latter case.

In the case of dominating convection one encounters several different sources of trouble. While under suitable assumptions one still knows that the weak formulation

$$a(v,u) = (v,f) \quad \forall\, v \in H_0^1(\Omega). \tag{3}$$

with $a(v,u) := (\mathcal{L}\,u, v) := \int_\Omega (\mathcal{L}\,u)(\boldsymbol{x})v(\boldsymbol{x})\,d\boldsymbol{x}$, has a unique solution u in $H_0^1(\Omega)$, a corresponding *Galerkin* approximation u_h in some finite dimensional trial space S_h defined by

$$a(v,u_h) = (v,f) \quad \forall\, v \in S_h, \tag{4}$$

may exhibit strong (unphysical) oscillations when $\|\boldsymbol{\beta}\| \gg \|\boldsymbol{\varepsilon}\|$. This highly undesirable phenomenon depends on the *grid Peclet number* which in the simple model case $\varepsilon(x) = \varepsilon\,\boldsymbol{I}$, for some $\varepsilon > 0$, is given by

$$P_h := \|\boldsymbol{\beta}\|h/2\varepsilon. \tag{5}$$

When the mesh size h becomes small enough, so that P_h decreases below one, no such oscillations will appear while for large h and hence large P_h the solution of the Galerkin problem may lack any meaning (see Figs. 1 – 4).

A correspondingly high resolution of the discretization would prohibitively increase computational and storage cost although in most parts of the domain a much coarser mesh would do.

A possible remedy is offered by stabilization methods such as streamline diffusion or Petrov–Galerkin formulations [CGMZ76, GL79, HB79, Joh87]. However, this often causes a serious loss of accuracy. One may be solving a discrete problem in a stable way that does no longer model the original physical process properly.

The second difficulty that has attracted much attention concerns the *efficiency* of the solution process. For $\boldsymbol{\beta} = \boldsymbol{0}$ and α in (2) not too small classical multigrid methods exhibit asymptotically optimal complexity. Within an accuracy tolerance of the order of the discretization error the Galerkin solution can be computed at computational and storage expense that is proportional to the size of the problem, i.e., to the number of unknowns. However, with degenerating ellipticity the efficiency of standard versions of such schemes degrades until around $P_h \sim 0.5$ (in our experiments) convergence no longer holds. Retaining the efficiency of a solver for a possibly wide range of parameters $\boldsymbol{\varepsilon}$, $\boldsymbol{\beta}$, a is therefore an issue of central importance. In fact, when $\|\boldsymbol{\varepsilon}\|/\|\boldsymbol{\beta}\|$ varies strongly over the domain Ω, it may be a primary concern that the algorithm does not break down even if, according to the above remarks, the accuracy of the solution degrades locally. *Robustness* has been significantly improved by combining streamline numberings of the nodes with ILU smoothers [HP97, Wit89]. This works well in two space dimensions and in absense of circular flows.

Here we will focus on strategies that *do not* make use of special numberings. Moreover, our primary concern is a proper interrelation of robustness and accuracy, i.e., correctness

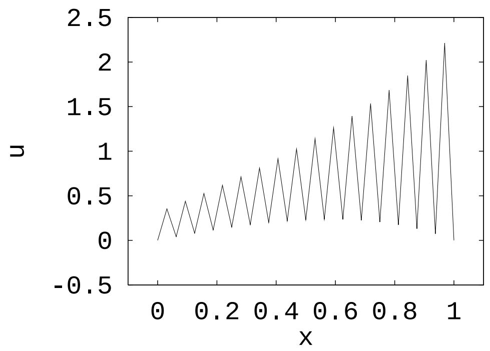

Figure 1: $\varepsilon = 2^{-11},\ \beta = f = 1,\ P_h = 32$

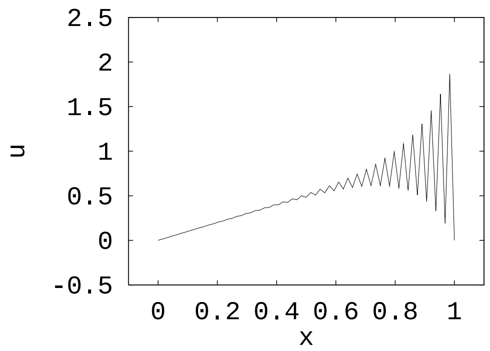

Figure 2: $\varepsilon = 2^{-11},\ \beta = f = 1,\ P_h = 16$

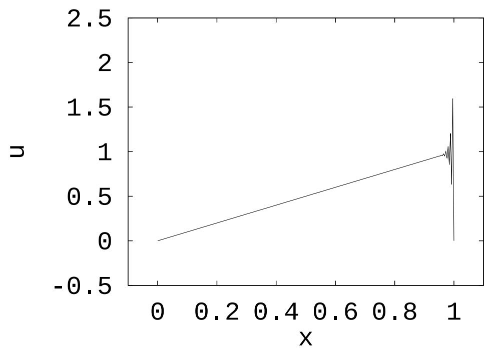

Figure 3: $\varepsilon = 2^{-11},\ \beta = f = 1,\ P_h = 4$

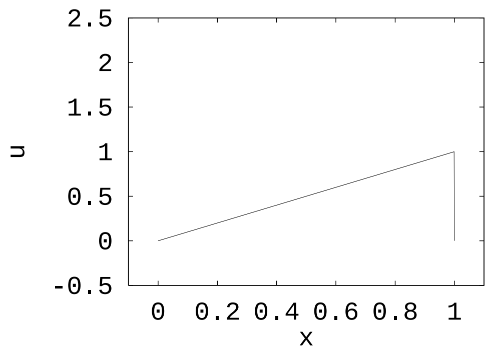

Figure 4: $\varepsilon = 2^{-11},\ \beta = f = 1,\ P_h = 1$

of the discrete model. To this end, recall that stabilization techniques in their standard form refer to some fixed mesh size. On the other hand, the above interrelation between P_h and oscillations stresses the well-known fact that, roughly speaking, the (discretization of the) operator $\boldsymbol{\mathcal{L}} = -\mathrm{div}\ (\boldsymbol{\varepsilon}\,\nabla) + \boldsymbol{\beta}^T\,\nabla$ behaves differently on different scales of resolution. On fine scales it is essentially elliptic, i.e., diffusion dominates, while on low scales only the transport part is resolved. This suggests keeping the multigrid ingredients like prolongation, restriction and smoothing *scale dependent*, see e.g. [Bra77, Bra82], which, of course, is to be accomplished by algebraic multigrid in an automatic way.

The second point concerns a possible stabilizing effect of adaptive refinements even if the (unbiased) Galerkin formulation is to be used as the basic discretization setting.

Let us now briefly describe the discretization setting for (1). To this end, it will be convenient to use the following shorthand notation. For any finite collection of functions Φ and any vector $\boldsymbol{u}$ of corresponding size we write

$$\Phi^T \boldsymbol{u} := \sum_{\phi\in\Phi} \boldsymbol{u}_\phi \phi.$$

Moreover, for any bilinear form $b(\cdot,\cdot)$ and any two collections of functions Φ, Θ let

$$b(\Phi,\Theta) := (b(\phi,\theta))_{\phi\in\Phi,\theta\in\Theta}$$

denote the corresponding generalized Gramian. Finally, we express the fact that a quantity a is uniformly bounded by a constant multiple of another quantity b by $a \lesssim b$ and write

$a \sim b$ to mean $a \lesssim b \ \wedge b \lesssim a$. Suppose that $S_h, \tilde{S}_h$ is a suitable pair of finite dimensional trial and test spaces spanned by bases $\Phi_h = \{\phi_{h,k} : k \in I_h\}$, $\tilde{\Phi}_h = \{\tilde{\phi}_{h,k} : k \in I_h\}$. The variational problem: find $u_h \in S_h$ such that $a(v, u_h) = (v, f)$ for $v \in \tilde{S}_h$ is equivalent to the linear system

$$a(\tilde{\Phi}_h, \Phi_h)\boldsymbol{u}_h = (\tilde{\Phi}_h, f) \tag{6}$$

where $u_h = \Phi_h^T \boldsymbol{u}_h$.

A common approach to solving such problems is to write the above linear system (upon a possible reordering of unknowns) as a 2×2 block matrix $\mathbf{A} = (\mathbf{A}_{i,l})_{i,l=0}^{1}$ by choosing some subset $I_H \subset I_h$ representing the *coarse grid*. This is also a conceptual starting point for *algebraic multigrid* when the choice of the coarse and fine grid is solely based on information extracted from the matrix, see e.g. [Axe86, RS86]. Following [DE88], a general strategy for generating then suitable multigrid ingredients like prolongations, restrictions and smoothers can be sketched as follows. Noting that

$$\mathbf{A} = \begin{pmatrix} \boldsymbol{I} & \mathbf{A}_{0,1} \\ \mathbf{0} & \mathbf{A}_{1,1} \end{pmatrix} \begin{pmatrix} (\mathbf{A}/\mathbf{A}_{1,1}) & \mathbf{0} \\ \mathbf{A}_{1,1}^{-1}\mathbf{A}_{1,0} & \boldsymbol{I} \end{pmatrix}, \tag{7}$$

where

$$(\mathbf{A}/\mathbf{A}_{1,1}) = \mathbf{A}_{0,0} - \mathbf{A}_{0,1}\mathbf{A}_{1,1}^{-1}\mathbf{A}_{1,0} \tag{8}$$

is the *Schur complement*, and defining the (matrix dependent) *prolongation* and *restriction operators*

$$\boldsymbol{P} := \begin{pmatrix} \boldsymbol{I} \\ -\mathbf{A}_{1,1}^{-1}\mathbf{A}_{1,0} \end{pmatrix}, \quad \boldsymbol{R} := (\boldsymbol{I} \ \mid \ -\mathbf{A}_{0,1}\mathbf{A}_{1,1}^{-1}), \tag{9}$$

one readily verifies that the corresponding *coarse grid operator* agrees with the Schur complement

$$\boldsymbol{R}\mathbf{A}\boldsymbol{P} = (\mathbf{A}/\mathbf{A}_{1,1}). \tag{10}$$

Moreover, the following *two-grid scheme* can easily be checked to be an exact solver based on the above factorization (7) of $\mathbf{A}$, see e.g. [DE88]:

- Let $\boldsymbol{v}$ be any approximate solution.
- Perform a *smoothing* step consisting of the application of $\begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{1,1}^{-1} \end{pmatrix}$ to the residual $\mathbf{r} := \boldsymbol{f} - \mathbf{A}\boldsymbol{v}$ with starting vector zero.
- Restrict the new residual by $\boldsymbol{R}$ to the coarse grid.
- Solve the coarse grid residual equation.
- Prolongate the solution by $\boldsymbol{P}$ and update the current solution.

Replacing the inverse $\mathbf{A}_{1,1}^{-1}$ in the above exact two-grid scheme by some approximate inverse $\mathbf{Q}$ gives rise to an approximation of $\mathbf{A}$. Moreover, if $\boldsymbol{Q}$ is efficiently executable, the corresponding modified prolongations, restrictions and smoother become practicable and the analogous two-grid cycle may be viewed as one step of an iteration scheme.

The subsequent development essentially fits into this framework. However, we are not aiming at a purely algebraic scheme but rather wish to exploit information about the underlying trial and test spaces. Consequently, one is *not* forced to derive the above basic blocking of the stiffness matrix only from an appropriate numbering. In order to motivate the generalizations we have in mind, note that the selection of a subset $I_H \subset I_h$ corresponds to a *decomposition* of the trial and test spaces into *direct summands*

$$S_h = S_H \oplus W, \quad \tilde{S}_h = \tilde{S}_H \oplus \tilde{W}, \tag{11}$$

where

$$S_H := \operatorname{span}\{\phi_{h,k} : k \in I_H\}, \quad W := \operatorname{span}\{\phi_{h,k} : k \in J := I_h \setminus I_H\}, \tag{12}$$

and likewise for $\tilde{S}_H, \tilde{W}$. This immediately suggests the following generalization which corresponds to HBMG (hierarchical basis multigrid), [BDY88]. In general the coarse spaces are not just spanned by subsets of the given bases but by appropriate *linear combinations* of basis functions

$$\Phi_H^T := \Phi_h^T \boldsymbol{P}_0, \quad \Psi^T := \Phi_h^T \boldsymbol{P}_1, \tag{13}$$

and likewise for $\tilde{S}_h$. The fact that $S_h = \operatorname{span}\Phi_H \oplus \operatorname{span}\Psi$ is then equivalent to saying that the composed matrix

$$\boldsymbol{P} := (\boldsymbol{P}_0, \boldsymbol{P}_1) \tag{14}$$

is invertible, i.e., $\boldsymbol{P}^{-1} =: \boldsymbol{R} = \binom{\boldsymbol{R}_0}{\boldsymbol{R}_1}$ exists, see Section 3 below. Since by (13), any $v_H \in S_H$ has the form

$$v_H = \Phi_H^T \boldsymbol{v}_H = \Phi_h^T \boldsymbol{P}_0 \boldsymbol{v}_H, \tag{15}$$

the matrix $\boldsymbol{P}_0$ represents a *prolongation* operator. Thus any such $\boldsymbol{P}$ (and likewise any $\tilde{\boldsymbol{P}}$ for the test spaces) transform the original system (6) into an equivalent one of the form

$$\tilde{\boldsymbol{P}}^T a(\tilde{\Phi}_h, \Phi_h) \boldsymbol{P} \boldsymbol{R} \boldsymbol{u}_h = \tilde{\boldsymbol{P}}^T (\tilde{\Phi}_h, f) \tag{16}$$

in the new unknowns $\binom{\boldsymbol{u}_H}{\boldsymbol{d}} := \boldsymbol{R}\boldsymbol{u}_h$. By (13), this becomes

$$\begin{pmatrix} \mathbf{A}_{0,0} & \mathbf{A}_{0,1} \\ \mathbf{A}_{1,0} & \mathbf{A}_{1,1} \end{pmatrix} \begin{pmatrix} \boldsymbol{u}_H \\ \boldsymbol{d} \end{pmatrix} = \begin{pmatrix} \mathbf{f}_0 \\ \mathbf{f}_1 \end{pmatrix}, \tag{17}$$

where now

$$\begin{aligned} \mathbf{A}_{0,0} &:= a(\tilde{\Phi}_H, \Phi_H), \qquad & \mathbf{A}_{0,1} &:= a(\tilde{\Phi}_H, \Psi), \\ \mathbf{A}_{1,0} &:= a(\tilde{\Psi}, \Phi_H), \qquad & \mathbf{A}_{1,1} &:= a(\tilde{\Psi}, \Psi), \end{aligned} \tag{18}$$

and

$$\mathbf{f}_0 := (\tilde{\Phi}_H, f), \quad \mathbf{f}_1 := (\tilde{\Psi}, f). \tag{19}$$

Now the system (17) could be used as a starting point for an algebraic strategy in the above spirit, where hopefully it is now easier to find a good approximation to $\mathbf{A}_{1,1}^{-1}$ and hence obtain appropriate adapted prolongations and smoothers. Of course, the central problem remains to find good splittings $\boldsymbol{P}_0, \boldsymbol{P}_1,\ \tilde{\boldsymbol{P}}_0, \tilde{\boldsymbol{P}}_1$.

Note that so far nothing particular, such as a *refinement history*, has been assumed yet about the trial and test spaces. In fact, when S_h is a finite element space with respect to some unstructured mesh no such refinement history may exist so that classical geometric multigrid is not directly applicable. An important issue is then to use information about the matrix $\mathbf{A}$ to construct suitable $\boldsymbol{P}_0$'s [BG97a, BG97b].

In the following we take here primarily the opposite point of view, namely that there exists already a "natural" ancestor S_H of S_h spanned by some (nodal) finite element basis Φ_H which naturally determines $\boldsymbol{P}_0$ in (13). For instance, in the case of standard bivariate piecewise linear finite elements (assuming that the basis functions are normalized to have L_2-norms of order one) column k of $\boldsymbol{P}_0$ corresponding to $\phi_{H,k}$ contains the values $\frac{1}{2}\phi_{H,k}(l)$ where l are the fine grid nodal points on the edges emanating from $k \in I_H$. What is not so clear is the choice of $\boldsymbol{P}_1$ that completes $\boldsymbol{P}_0$ to an invertible matrix. Nevertheless, in case of nodal bases there is a particularly simple choice derived from taking as in (12)

$$\Psi := \{\phi_{h,k} : k \in J := I_h \setminus I_H\}, \tag{20}$$

i.e., the nodal basis functions at new nodal points are taken to span a complement space. This corresponds to the *hierarchical basis* [Yse86]. Of course, this choice is independent of the problem at hand and may not do yet the intended job of creating an easily invertible *complement block* $\mathbf{A}_{1,1}$. However, there is another favorable feature about hierarchical bases namely that in this case the inverse $\boldsymbol{R}$ of the matrix $\boldsymbol{P}$ can be determined explicitly, see e.g. [Dah97]. This will be seen to open up the use of some concepts developed in the context of wavelet analysis. Pursuing this line differs, besides the above mentioned similarities, from the above mentioned other approaches. The point is that once a pair $\boldsymbol{P}, \boldsymbol{R}$ is known there is a systematic way of generating *all* other *completions* $\check{\boldsymbol{P}}_1$ of $\boldsymbol{P}_0$ that induce a proper space splitting.

A central objective of this paper is to outline a general setting where these mechanisms can be applied for generating problem adapted multigrid ingredients. Moreover, the explicit representations of the corresponding complement spaces will be used to construct suitable smoothers as well as to control *adaptive refinements*. Of course, there are interesting situations where a nested sequence of trial spaces is available while this is not the case for the test spaces. For instance, in connection with box schemes the test functions are piecewise constants with respect to a partition which is dual to the triangulation for the trial spaces. Subspaces are then formed by agglomoration.

The layout of the paper is as follows. In Section 3 we briefly review the background of *stable completions* [CDP96]. In Section 4 we reinterpret stabilization techniques for the hierarchical basis preconditioner in these terms. In Section 5 we focus on multiplicative variants with emphasis on convection dominated problems [DMS97]. After discussing some numerical experiments in Section 6 we conclude in Section 7 with indicating adaptive mechanisms based on stable completions again supported by some numerical results.

3 Stable Completions

Let us now take a more systematic look at the situation considered above. As indicated in the previous section, we are interested in a hierarchy of finite element spaces $S_j = S_{h_j}$

$$S_0 \subset S_1 \subset S_2 \subset \ldots \subset H_0^1(\Omega), \tag{21}$$

where each space is spanned by a *stable* nodal basis $\Phi_j = \{\phi_{j,k} : k \in I_j\}$. Here 'stable' means that

$$\|\boldsymbol{u}\|_{\ell_2} \sim \|\Phi_j^T \boldsymbol{u}\|_{L_2(\Omega)}, \tag{22}$$

where $\|\boldsymbol{u}\|_{\ell_2} := \left(\sum_{k\in I_j} |u_k|^2\right)^{1/2}$ and as above $\Phi_j^T\, \boldsymbol{u} := \sum_{k\in I_j} u_k\, \phi_{j,k}$ is a shorthand notation for expansion in terms of a given basis.

As has been already pointed out in Section 2, nestedness implies that each basis function on a coarse level j can be written as a linear combination of those on the next finer level $j+1$ which can be expressed as in (13) by

$$\Phi_j^T = \Phi_{j+1}^T\, \boldsymbol{P}_{j,0}, \tag{23}$$

where the k–th column of $\boldsymbol{P}_{j,0}$ contains the coefficients in the expansion of $\phi_{j,k}$ in terms of Φ_{j+1}. Thus when dealing with nodal bases $\boldsymbol{P}_{j,0}$ is usually very sparse as shown by the example of standard piecewise linear elements. As mentioned in the previous section, (23) says that $v_j = \Phi_j^T\, \boldsymbol{v}_j \in S_j$ can be written as a function in S_{j+1} as

$$v_j = \Phi_j^T \boldsymbol{v}_j = \Phi_{j+1}^T\, (\boldsymbol{P}_{j,0}\, \boldsymbol{v}_j) \tag{24}$$

with new coefficients $\boldsymbol{v}_{j+1} := \boldsymbol{P}_{j,0}\, \boldsymbol{v}_j$. Thus the refinement matrix $\boldsymbol{P}_{j,0}$ acts as a *prolongation* operator.

A basic concept in wavelet analysis is to write a function as a sum of successive updates of some coarse scale version. These updates sit in suitable complements W_j of S_j in S_{j+1}

$$S_{j+1} = S_j \oplus W_j. \tag{25}$$

Since $W_j \subset S_{j+1}$, there must exist another matrix $\boldsymbol{P}_{j,1}$ such that

$$\Psi_j^T = \Phi_{j+1}^T\, \boldsymbol{P}_{j,1} \tag{26}$$

is a basis for W_j. In fact, a necessary and sufficient condition for $\Phi_j \cup \Psi_j$ being also a stable basis for S_{j+1} is that the composed matrices $\boldsymbol{P}_j = (\boldsymbol{P}_{j,0}, \boldsymbol{P}_{j,1})$ are invertible and the $\boldsymbol{P}_j$ and their inverse $\boldsymbol{R}_j = \begin{pmatrix} \boldsymbol{R}_{j,0} \\ \boldsymbol{R}_{j,1} \end{pmatrix}$ are uniformly spectrally bounded [CDP96]

$$\|\boldsymbol{P}_j\|,\ \|\boldsymbol{R}_j\| = \mathcal{O}(1). \tag{27}$$

$\boldsymbol{P}_{j,1}$ is called a *stable completion* of $\boldsymbol{P}_{j,0}$. It also represents a prolongation but into the complement space W_j.

Note that

$$\boldsymbol{P}_j\, \boldsymbol{R}_j = \boldsymbol{I} \tag{28}$$

is equivalent to saying that

$$\Phi_{j+1}^T = \Phi_j^T\, \boldsymbol{R}_{j,0} + \Psi_j^T\, \boldsymbol{R}_{j,1} \tag{29}$$

so that $v_{j+1} = \Phi_{j+1}^T\, \boldsymbol{v}_{j+1}$ can be written, in view of (29), as

$$\boldsymbol{v}_{j+1} = \Phi_j^T\, (\boldsymbol{R}_{j,0}\, \boldsymbol{v}_{j+1}) + \Psi_j^T\, (\boldsymbol{R}_{j,1}\, \boldsymbol{v}_{j+1}) = \Phi_j^T\, \boldsymbol{v}_j + \Psi_j^T\, \boldsymbol{d}_j \tag{30}$$

in terms of a coarse scale part $v_j = \Phi_j^T\, \boldsymbol{v}_j$ and a *detail* part $w_j = \Psi_j^T\, \boldsymbol{d}_j$. Thus the matrices $\boldsymbol{R}_{j,0}$, $\boldsymbol{R}_{j,1}$ act as *restriction operators* into the coarse trial space S_j and the complement W_j.

By construction, the collection $\Psi^J := \Phi_0 \cup \bigcup_{j=0}^{J-1} \Psi_j$ is also a basis for S_J. In particular, when the complement bases Ψ_j have the form (20), one obtains the hierarchical basis.

Given any element in S_J represented in *multiscale form* $v_J = \boldsymbol{d}_J^T \Psi^J$, one easily concludes from (23) and (26) that the transformation $\mathbf{T}_J$ which takes the multiscale coefficients $\boldsymbol{d}_J$ of v_J into its nodal coefficients $\boldsymbol{v}_J = \mathbf{T}_J \boldsymbol{d}_J$ is schematically given by

$$\begin{array}{ccccccccc}
 & \boldsymbol{P}_{0,0} & & \boldsymbol{P}_{1,0} & & & \boldsymbol{P}_{J-1,0} & & \\
\boldsymbol{v}_0 & \rightarrow & \boldsymbol{v}_1 & \rightarrow & \boldsymbol{v}_2 & \rightarrow \cdots & \rightarrow & \boldsymbol{v}_J & \\
 & \boldsymbol{P}_{0,1} & & \boldsymbol{P}_{1,1} & & & \boldsymbol{P}_{J-1,1} & & , \\
 & \nearrow & & \nearrow & & \nearrow \cdots & \nearrow & & \\
\boldsymbol{d}^0 & & \boldsymbol{d}^1 & & \boldsymbol{d}^2 & & \boldsymbol{d}^{J-1} & &
\end{array} \tag{31}$$

where $\boldsymbol{d}_J^T := (\boldsymbol{v}_0^T, (\boldsymbol{d}^0)^T, \ldots, (\boldsymbol{d}^{J-1})^T)$. To rephrase this in matrix form, let

$$\mathbf{T}_{J,j} := \begin{pmatrix} \boldsymbol{P}_j & \mathbf{0} \\ \mathbf{0} & \boldsymbol{I} \end{pmatrix},$$

where $\boldsymbol{I}$ is the identity block of size $\#\Phi_J - \#\Phi_{j+1}$. Then one has

$$\mathbf{T}_J = \mathbf{T}_{J,J-1} \cdots \mathbf{T}_{J,0}. \tag{32}$$

Prolongations of the type $\boldsymbol{P}_{j,0}$ are standard ingredients of multigrid concepts. One usually does not care, however, about those for the complement spaces, since the complements represent high *frequency components* that are to be eliminated by smoothing anyway. However, it will be seen that the $\boldsymbol{P}_{j,1}$ will play an important role for the construction of suitable smoothers as well as for controlling local refinements.

Stable completions are, of course, by no means unique. Those giving rise to orthogonal splittings $S_j \perp W_j$ would give rise to *wavelets* in the classical sense. Here it is important to note that by (28) varying the stable completion will change the restriction operators. Thus stable completions can be used as tuning knobs for constructing restrictions. Of course, it is by no means clear beforehand how to get stable completions with desireable properties. A key is provided by the very simple but useful observation from [CDP96] that, once one knows *one* (initial) stable completion *all* others can be parametrized by two families of matrices. In fact, for any $(\# I_k) \times (\# J_k)$ matrix $\check{\boldsymbol{L}}$ and invertible $(\# J_k) \times (\# J_k)$ matrix $\check{\boldsymbol{K}}$ we infer from (28) that

$$\boldsymbol{I} = \boldsymbol{P}_j \boldsymbol{R}_j = \boldsymbol{P}_j \begin{pmatrix} \boldsymbol{I} & \check{\boldsymbol{L}} \\ \mathbf{0} & \check{\boldsymbol{K}} \end{pmatrix} \begin{pmatrix} \boldsymbol{I} & -\check{\boldsymbol{L}}\check{\boldsymbol{K}}^{-1} \\ \mathbf{0} & \check{\boldsymbol{K}}^{-1} \end{pmatrix} \boldsymbol{R}_j =: \check{\boldsymbol{P}}_j \check{\boldsymbol{R}}_j, \tag{33}$$

where

$$\check{\boldsymbol{P}}_{j,0} = \boldsymbol{P}_{j,0}, \quad \check{\boldsymbol{P}}_{j,1} = \boldsymbol{P}_{j,0}\check{\boldsymbol{L}} + \boldsymbol{P}_{j,1}\check{\boldsymbol{K}} \tag{34}$$

and

$$\check{\boldsymbol{R}}_{j,0} = \boldsymbol{R}_{j,0} - \check{\boldsymbol{L}}\check{\boldsymbol{K}}^{-1}\boldsymbol{R}_{j,1}, \quad \check{\boldsymbol{R}}_{j,1} = \check{\boldsymbol{K}}^{-1}\boldsymbol{R}_{j,1}. \tag{35}$$

Thus, one obtains new restriction operators and a new prolongation to a correspondingly new complement space

$$\check{W}_j := \operatorname{span} \check{\Psi}_j, \quad \check{\Psi}_j^T := \Phi_{j+1}^T \check{\boldsymbol{P}}_{j,1}, \tag{36}$$

while retaining the prolongation $\boldsymbol{P}_{j,0}$ to S_{j+1}.

Likewise, there is a *dual* way of changing stable completions according to

$$\boldsymbol{I} = \boldsymbol{R}_j^T \boldsymbol{P}_j^T = \boldsymbol{R}_j^T \begin{pmatrix} \boldsymbol{I} & \hat{\boldsymbol{L}} \\ \mathbf{0} & \hat{\boldsymbol{K}} \end{pmatrix} \begin{pmatrix} \boldsymbol{I} & -\hat{\boldsymbol{L}}\hat{\boldsymbol{K}}^{-1} \\ \mathbf{0} & \hat{\boldsymbol{K}}^{-1} \end{pmatrix} \boldsymbol{P}_j^T =: \hat{\boldsymbol{R}}_j^T \hat{\boldsymbol{P}}_j^T \tag{37}$$

which yields further modifications

$$\hat{\boldsymbol{R}}_{j,0} = \boldsymbol{R}_{j,0}, \quad \hat{\boldsymbol{R}}_{j,1} = \hat{\boldsymbol{L}}^T \boldsymbol{R}_{j,0} + \hat{\boldsymbol{K}}^T \boldsymbol{R}_{j,1} \tag{38}$$

and

$$\hat{\boldsymbol{P}}_{j,0} = \boldsymbol{P}_{j,0} - \boldsymbol{P}_{j,1}\hat{\boldsymbol{K}}^{-T}\hat{\boldsymbol{L}}^T, \quad \hat{\boldsymbol{P}}_{j,1} = \boldsymbol{P}_{j,1}\hat{\boldsymbol{K}}^{-T}. \tag{39}$$

This time the restriction stays the same while the prolongations change.

Such modifications can be different for each j. As long as the corresponding matrices $\hat{\boldsymbol{L}}$, $\check{\boldsymbol{L}}$, $\hat{\boldsymbol{K}}$, $\hat{\boldsymbol{K}}^{-1}$, $\check{\boldsymbol{K}}$, $\check{\boldsymbol{K}}^{-1}$ have uniformly bounded spectral norms the corresponding modified bases stay stable [CDP96].

We will explain next how this mechanism can be used to adapt multigrid ingredients to a given problem.

4 Stabilization of Hierarchical Bases

Let us begin with reinterpreting a recent approach to multilevel preconditioning proposed in [VW97] in the above terms. It concerns the special case $\boldsymbol{\beta} = 0$ in (1) and standard Galerkin discretizations by piecewise linear (or multilinear) finite elements, i.e., $S_j = \tilde{S}_j$. Thus Φ_j are corresponding nodal bases and the indices $k \in I_j$ represent the nodal points so that $J_j := I_{j+1} \setminus I_j$. For Ψ_j defined in analogy to (20) Ψ^J denotes the hierarchical basis of S_J. The hierarchical basis preconditioner is based on the fact that a certain (levelwise) symmetric diagonal scaling of the stiffness matrix $\mathbf{A}_{\psi^J} := a(\Psi^J, \Psi^J)$ with respect to the hierarchical basis gives rise (for $\boldsymbol{\beta} = \mathbf{0}, a = 0$) to spectral condition numbers that in the bivariate case only grow like J^2 [Yse86]. Since the low level elements in Ψ^J have larger supports $\mathbf{A}_{\Psi^J}$ is less sparse than $a(\Phi_J, \Phi_J)$. However, repeating (16) shows that

$$\mathbf{A}_{\Psi^J} = \mathbf{T}_J^T a(\Phi_J, \Phi_J)\mathbf{T}_J.$$

In case of the hierarchical basis the transformation $\mathbf{T}_J$ is very simple and efficient and so is the application of a scaled version of $\mathbf{A}_{\Psi^J}$ [Yse86].

The drawbacks of this scheme are that condition numbers grow like $2^{j(d-2)}$ in the d–dimensional case, when $a = 0$. When $a > 0$ the efficiency further deteriorates with increasing $\|a\|/\|\boldsymbol{\varepsilon}\|$. It is known that choosing stable complement bases for *orthogonal decompositions* $S_{j+1} = S_j \oplus^\perp W_j$ would offer a remedy for both difficulties. The idea pursued in [VW97] is to turn the hierarchical basis into another multiscale basis which corresponds *approximately* to orthogonal splittings, while maintaining the computational advantages of hierarchical bases to a possibly large extent.

To describe this in present terms let us first see how to determine *exact* orthogonal complements between successive trial spaces. As mentioned before, for hierarchical bases the matrices $\boldsymbol{P}_j$, $\boldsymbol{R}_j$ are easily identified (see e.g. [Dah97]). They are both very sparse. Now consider

$$\check{\boldsymbol{P}}_{j,1} := \Big(\boldsymbol{I} - (\boldsymbol{P}_{j,0}(\Phi_j, \Phi_j)^{-1}(\Phi_j, \Phi_{j+1}))\Big)\, \boldsymbol{P}_{j,1}. \tag{40}$$

In fact, one readily checks that

$$\begin{aligned}(\Phi_j, \Phi_{j+1})\check{\boldsymbol{P}}_{j,1} &= (\Phi_j, \Phi_{j+1})\boldsymbol{P}_{j,1} - (\Phi_j, \Phi_{j+1})\boldsymbol{P}_{j,0}(\Phi_j, \Phi_j)^{-1}(\Phi_j, \Phi_{j+1})\boldsymbol{P}_{j,1} \\ &= (\Phi_j, \Phi_{j+1})\boldsymbol{P}_{j,1} - (\Phi_j, \Phi_{j+1})\boldsymbol{P}_{j,1} = \mathbf{0}\end{aligned}$$

where we have used that, by (23),

$$(\Phi_j, \Phi_{j+1})\boldsymbol{P}_{j,0} = (\Phi_j, \Phi_j).$$

Thus the functions in $\check{\Psi}_j = \check{\boldsymbol{P}}_{j,1}^T \Phi_{j+1}$ are indeed orthogonal to the elements of Φ_j and thus to S_j.

Note that (40) correspondes to the choice $\check{\boldsymbol{K}} = \boldsymbol{I}$ and

$$\check{\boldsymbol{L}} = -(\Phi_j, \Phi_j)^{-1}(\Phi_j, \Phi_{j+1})\boldsymbol{P}_{j,1} \tag{41}$$

in (34). Hence the *new* complement bases $\check{\Psi}_j = \check{\boldsymbol{P}}_{j,1}^T \Phi_{j+1}$ are of the form

$$\check{\Psi}_j^T = \Phi_{j+1}^T(\boldsymbol{P}_{j,1} + \boldsymbol{P}_{j,0}\check{\boldsymbol{L}}) = \Psi_j^T + \Phi_j^T\check{\boldsymbol{L}}, \tag{42}$$

that is, they are obtained by a coarse grid correction of the original complement bases. Accordingly, the corresponding new multiscale transformation can be efficiently realized in two stages namely by the original one followed by coarse scale corrections provided that the corrections are local. Algorithmic details are discussed in [Swe97].

Of course, due to the appearance of the inverse $(\Phi_j, \Phi_j)^{-1}$ of the Gramian of the nodal basis the coarse grid correction given by (40) is by no means local. By the above comments a core ingredient of the preconditioning process is the application of $\check{\boldsymbol{P}}_{j,1}$ to a vector $\boldsymbol{v}$ which amounts to applying the initial, explicitly given $\boldsymbol{P}_{j,1}$ followed by the update $\boldsymbol{P}_{j,0}\check{\boldsymbol{L}}\boldsymbol{v} = \boldsymbol{z}$. Thus the crucial part is the evaluation of $\check{\boldsymbol{L}}\boldsymbol{v} = \boldsymbol{y}$, which according to (41) requires solving

$$(\Phi_j, \Phi_j)\boldsymbol{y} = -(\Phi_j, \Phi_{j+1})\boldsymbol{P}_{j,1}\boldsymbol{v}. \tag{43}$$

The uniform stability of the Φ_j is equivalent to the fact that the mass matrices (Φ_j, Φ_j) have uniformly bounded condition numbers. Thus a few Gauss–Seidel iterations suffice to produce an approximation to $\check{\boldsymbol{L}}\boldsymbol{v}$ corresponding to an almost orthogonal decomposition [VW97].

We conclude this section by another stabilization scheme [DK92]. This is based on constructing first a dual basis Ξ_J to Φ_J, i.e., $(\Phi_J, \Xi_J) = \mathbf{I}$. The $\xi_{J,k}$ in Ξ_J consist of discontinuous piecewise linear functions having the same support as the corresponding nodal basis functions. This gives rise to L_2-bounded projectors of the form $Q_J v := (v, \Xi_J)\Phi_J$. Instead of orthogonal complements one considers now the complement spaces $W_j := (Q_{j+1} - Q_j)S_{j+1}$. One can show that these spaces are spanned by the bases $\check{\Psi}_j^T := \Phi_{j+1}^T \check{\boldsymbol{P}}_{j,1}$ [CDP96], where in this case

$$\check{\boldsymbol{P}}_{j,1} = \left(\boldsymbol{I} - \boldsymbol{P}_{j,0}(\Xi_j, \Phi_{j+1})\right)\boldsymbol{P}_{j,1},$$

so that the new complement basis is again of the form (42) and the above comments on the algorithmic realization apply in the same way as above. For a regular bivariate triangulation the coarse grid correction involves four terms. Moreover, in this case the corresponding multiscale basis can be shown to be a *Riesz basis* which means that any $v \in L_2$ has a unique expansion $v = \check{\Psi}^T\boldsymbol{d}$ and the Euclidean norm of the coefficient array $\boldsymbol{d}$ behaves as the L_2-norm of the function

$$\|v\|_{L_2} \sim \|\boldsymbol{d}\|_{\ell_2},$$

see [LO97]. As a consequence this stabilization produces a robust preconditioner for the Galerkin discretizations of (1) with $\boldsymbol{\beta} = \mathbf{0}$ but essentially arbitrary a [Dah97].

5 Multiplicative Schemes

In this section we turn to multiplicative multigrid versions in the sense of Section 2. In fact, we confine the discussion first mainly to a two–grid scheme. Thus we will adhere to the notation used in Section 2 and consider two successive trial spaces $S_H \subset S_h$. We ease the notation by largely supressing level–indices. If necessary the roles of $j,\ j+1$ will be played by $H,\ h$ respectively.

We will assume that some initial decomposition

$$S_h = S_H \oplus W \tag{44}$$

is available where as before $\Phi_h,\ \Phi_H,\ \Psi$ denote the (stable) bases of $S_h,\ S_H,\ W$, respectively. Depending on the underlying finite element spaces Ψ could be a small support wavelet basis from [Ste96] or the hierarchical complement basis described in the previous section, so that the prolongation $\boldsymbol{P}$ and restriction $\boldsymbol{R}$ are sparse and explicitly known. For simplicity, we will confine the discussion to the Galerkin case, i.e., $S_h = \tilde{S}_h$.

As mentioned before, the first step is common to a number of different approaches [Yse86, BDY88, BG97a, BG97b, VW97] namely to write (4) as a linear system of equations relative to the basis $\Phi_H \cup \Psi$ of S_h. Relative to the fine scale basis Φ_h relation (4) is then equivalent to the linear system

$$a(\Phi_h, \Phi_h)\boldsymbol{u}_h = \boldsymbol{f}_h := (\Phi_h, f). \tag{45}$$

As in (16) we will transform this system into the equivalent one

$$\hat{\boldsymbol{P}}^T a(\Phi_h, \Phi_h)\check{\boldsymbol{P}}\check{\boldsymbol{R}}\boldsymbol{u}_h = \hat{\boldsymbol{P}}^T \boldsymbol{f}_h \tag{46}$$

where, however, now $\hat{\boldsymbol{P}}, \check{\boldsymbol{P}}, \hat{\boldsymbol{R}}, \check{\boldsymbol{R}}$ are defined by (34), (35), (38) and (39) as *modified* stable completions relative to the initially given $\boldsymbol{P},\ \boldsymbol{R}$. By the same relations this can be written as

$$\begin{pmatrix} \boldsymbol{I} & -\hat{\boldsymbol{L}}\hat{\boldsymbol{K}}^{-1} \\ \boldsymbol{0} & \hat{\boldsymbol{K}}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{A}_{0,0} & \mathbf{A}_{0,1} \\ \mathbf{A}_{1,0} & \mathbf{A}_{1,1} \end{pmatrix} \begin{pmatrix} \boldsymbol{I} & \check{\boldsymbol{L}} \\ \boldsymbol{0} & \check{\boldsymbol{K}} \end{pmatrix} \begin{pmatrix} \check{\boldsymbol{u}}_H \\ \check{\boldsymbol{d}} \end{pmatrix} = \begin{pmatrix} \hat{\boldsymbol{f}}_0 \\ \hat{\boldsymbol{f}}_1 \end{pmatrix}, \tag{47}$$

where

$$\check{\boldsymbol{u}}_H = \check{\boldsymbol{R}}_0\boldsymbol{u}_h, \qquad \check{\boldsymbol{d}} = \check{\boldsymbol{R}}_1\boldsymbol{u}_h,$$

and for $\Psi^0 := \Phi_H,\ \Psi^1 := \Psi$ we have set

$$\mathbf{A}_{i,l} := a(\Psi^i, \Psi^l), \quad i, l = 0, 1.$$

The two by two block matrix $\mathbf{A} = (\mathbf{A}_{i,l})_{i,l=0}^1$ is the stiffness matrix relative to the initial two–level basis $\Phi_H \cup \Psi$, see (17). Specifically, when Ψ is the hierarchical complement basis, this is the starting point in several closely related studies [Yse86, BDY88, AV89, BG97b]. Straightforward calculations show that the matrix on the left hand side of (47) is given by

$$\hat{\check{\boldsymbol{A}}} = \begin{pmatrix} \mathbf{A}_{0,0} - \hat{\boldsymbol{L}}\hat{\boldsymbol{K}}^{-1}\mathbf{A}_{1,0} & \check{\boldsymbol{A}}_{0,1} - \hat{\boldsymbol{L}}\hat{\boldsymbol{K}}^{-1}\check{\boldsymbol{A}}_{1,1} \\ \hat{\boldsymbol{K}}^{-1}\mathbf{A}_{1,0} & \hat{\boldsymbol{K}}^{-1}\check{\boldsymbol{A}}_{1,1} \end{pmatrix}. \tag{48}$$

Here we have used the notation

$$\check{\boldsymbol{A}} = \left(\check{\boldsymbol{A}}_{i,l}\right)_{i,l=0,1} = \mathbf{A}\begin{pmatrix} \boldsymbol{I} & \check{\boldsymbol{L}} \\ \boldsymbol{0} & \check{\boldsymbol{K}} \end{pmatrix}.$$

The central objective is now to choose $\hat{\boldsymbol{L}},\ \hat{\boldsymbol{K}}, \check{\boldsymbol{L}},\ \check{\boldsymbol{K}}$ depending on the coefficients in (1) in such a way that multigrid based on this splitting works well.

The strategy pursued in [DMS97] can be summarized as follows

(I) Choose $\check{\boldsymbol{L}},\ \check{\boldsymbol{K}}$ such that $\check{\boldsymbol{A}}_{1,1} := a(\Psi, \check{\Psi})$ can be inverted efficiently. Recall that the change from Ψ to $\check{\Psi}$ on the trial side correspondes to modified restriction operators.

(II) Choose $\hat{\boldsymbol{L}},\ \hat{\boldsymbol{K}}$ so that the coarse grid problem decouples, i.e.,

$$\hat{\check{\boldsymbol{A}}}_{0,1} = \mathbf{0}, \tag{49}$$

which means

$$\check{\boldsymbol{A}}_{0,1} = \hat{\boldsymbol{L}}\hat{\boldsymbol{K}}^{-1}\check{\boldsymbol{A}}_{1,1}. \tag{50}$$

Specially, choosing $\hat{\boldsymbol{K}} := \boldsymbol{I}$ this yields

$$\hat{\boldsymbol{L}} = \check{\boldsymbol{A}}_{0,1}\check{\boldsymbol{A}}_{1,1}^{-1}, \tag{51}$$

and (48) becomes

$$\hat{\check{\boldsymbol{A}}} = \begin{pmatrix} \mathbf{A}_{0,0} - \hat{\boldsymbol{L}}\mathbf{A}_{1,0} & \mathbf{0} \\ \mathbf{A}_{1,0} & \check{\boldsymbol{A}}_{1,1} \end{pmatrix}. \tag{52}$$

(III) Construct a smoother that effectively reduces *only* the error components in the complement space (high frequency part). To this end, we consider the ansatz

$$\boldsymbol{S} = \left(\boldsymbol{I} - \boldsymbol{C}\hat{\check{\boldsymbol{A}}}\right), \quad \boldsymbol{C} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \boldsymbol{M} & \boldsymbol{Q} \end{pmatrix}$$

which means

$$\boldsymbol{S} = \begin{pmatrix} \boldsymbol{I} & \mathbf{0} \\ -\left(\boldsymbol{M}\hat{\check{\boldsymbol{A}}}_{0,0} + \boldsymbol{Q}\mathbf{A}_{1,0}\right) & \boldsymbol{I} - \boldsymbol{Q}\check{\boldsymbol{A}}_{1,1} \end{pmatrix}. \tag{53}$$

In these terms the objective is to determine $\boldsymbol{M},\ \boldsymbol{Q}$ such that

$$\left\| -\left(\boldsymbol{M}\hat{\check{\boldsymbol{A}}}_{0,0} + \boldsymbol{Q}\mathbf{A}_{1,0}\right)\boldsymbol{e}_0 + \left(\boldsymbol{I} - \boldsymbol{Q}\check{\boldsymbol{A}}_{1,1}\right)\boldsymbol{e}_1\right\| < \eta\|\boldsymbol{e}_1\| \tag{54}$$

for some $\eta < 1$. Of course, an ideal choice of $\boldsymbol{Q}$ would be

$$\boldsymbol{Q} = \check{\boldsymbol{A}}_{1,1}^{-1}.$$

On account of (I) therefore application of $\boldsymbol{Q}$ to $\boldsymbol{v}$ means the execution of a few iterations on the system $\check{\boldsymbol{A}}_{1,1}\,\boldsymbol{x} = \boldsymbol{y}$. Likewise, the ideal choice of $\boldsymbol{M}$ would be $\boldsymbol{M} = -\boldsymbol{Q}\mathbf{A}_{1,0}\hat{\check{\boldsymbol{A}}}_{0,0}^{-1}$. Thus evaluating $\boldsymbol{M}\boldsymbol{v}$ means

(i) solve $\hat{\check{\boldsymbol{A}}}_{0,0}\boldsymbol{x} = \boldsymbol{v}$ approximately;

(ii) compute $\boldsymbol{y} := \mathbf{A}_{1,0}\boldsymbol{x}$;

(iii) apply $\boldsymbol{Q}\boldsymbol{y}$.

Note that (i) requires solving the coarse grid problem.

The above choice $\hat{\boldsymbol{K}} = \boldsymbol{I}$ is only one among possible alternatives. In fact, according to (48) $\hat{\boldsymbol{K}}$ can also be used to precondition $\hat{\check{\boldsymbol{A}}}_{1,1}$. We refer to [DMS97] for a more detailed discussion.

Recall that the transformations in (47) involves modifying the restriction operator in the trial space. In addition the *prolongation* in the test space is changed whose transpose becomes a *restriction* acting on right hand sides (46).

There is an interesting variant which is based on employing the dual way of changing stable completions (37) *also* for the trial space. Relation (47) then takes the form

$$\begin{pmatrix} \boldsymbol{I} & -\hat{\boldsymbol{L}}\hat{\boldsymbol{K}}^{-1} \\ \boldsymbol{0} & \hat{\boldsymbol{K}}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{A}_{0,0} & \mathbf{A}_{0,1} \\ \mathbf{A}_{1,0} & \mathbf{A}_{1,1} \end{pmatrix} \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ -\check{\boldsymbol{K}}^{-T}\check{\boldsymbol{L}}^{T} & \check{\boldsymbol{K}}^{-T} \end{pmatrix} \begin{pmatrix} \check{\boldsymbol{u}}_H \\ \check{\boldsymbol{d}} \end{pmatrix} = \begin{pmatrix} \hat{\boldsymbol{f}}_0 \\ \hat{\boldsymbol{f}}_1 \end{pmatrix}, \tag{55}$$

so that

$$\hat{\check{\boldsymbol{A}}} = \begin{pmatrix} \check{\boldsymbol{A}}_{0,0} - \hat{\boldsymbol{L}}\hat{\boldsymbol{K}}^{-1}\check{\boldsymbol{A}}_{1,0} & \check{\boldsymbol{A}}_{0,1} - \hat{\boldsymbol{L}}\hat{\boldsymbol{K}}^{-1}\check{\boldsymbol{A}}_{1,1} \\ \hat{\boldsymbol{K}}^{-1}\check{\boldsymbol{A}}_{1,0} & \hat{\boldsymbol{K}}^{-1}\check{\boldsymbol{A}}_{1,1} \end{pmatrix}, \tag{56}$$

where

$$\check{\boldsymbol{A}} = \begin{pmatrix} \mathbf{A}_{0,0} - \mathbf{A}_{0,1}\check{\boldsymbol{K}}^{-T}\check{\boldsymbol{L}}^{T} & \mathbf{A}_{0,1}\check{\boldsymbol{K}}^{-T} \\ \mathbf{A}_{1,0} - \mathbf{A}_{1,1}\check{\boldsymbol{K}}^{-T}\check{\boldsymbol{L}}^{T} & \mathbf{A}_{1,1}\check{\boldsymbol{K}}^{-T} \end{pmatrix}. \tag{57}$$

The choice

$$\hat{\boldsymbol{K}} = \boldsymbol{I},\ \check{\boldsymbol{L}}^{T} = \check{\boldsymbol{A}}_{1,1}^{-1}\mathbf{A}_{1,0},\ \hat{\boldsymbol{L}} = \mathbf{A}_{0,1}\check{\boldsymbol{A}}_{1,1}^{-1} \tag{58}$$

would eliminate both blocks $\hat{\check{\boldsymbol{A}}}_{0,1}$, $\hat{\check{\boldsymbol{A}}}_{1,0}$ where again $\check{\boldsymbol{K}}$ has to be chosen so that $\check{\boldsymbol{A}}_{1,1} = \mathbf{A}_{1,1}\check{\boldsymbol{K}}^{-T}$ can be efficiently inverted. In this case the coarse grid and complement problem are completely decoupled. The corresponding smoother has then an even simpler form (compare (53))

$$\boldsymbol{S} = \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I} - \boldsymbol{Q}\check{\boldsymbol{A}}_{1,1} \end{pmatrix} \tag{59}$$

and acts *only* on the complement block.

Let us briefly relate this to the general setting described in Section 2. Straightforward calculations reveal that the coarse grid operator

$$\begin{aligned} \hat{\check{\boldsymbol{A}}}_{0,0} &= \mathbf{A}_{0,0} - \check{\boldsymbol{A}}_{0,1}\check{\boldsymbol{A}}_{1,1}^{-1}\mathbf{A}_{1,0} = (\mathbf{A}/\mathbf{A}_{1,1}) \\ &= \check{\boldsymbol{A}}_{0,0} - \check{\boldsymbol{A}}_{0,1}\check{\boldsymbol{A}}_{1,1}^{-1}\check{\boldsymbol{A}}_{1,0} = (\check{\boldsymbol{A}}/\check{\boldsymbol{A}}_{1,1}) \end{aligned} \tag{60}$$

agrees with the Schur complement in both blockings $\mathbf{A}$ and $\check{\boldsymbol{A}}$, see (8) and (10).

Similarly, the second strategy, provides after some straightforward calculations

$$\begin{aligned} \hat{\check{\boldsymbol{A}}}_{0,0} &= \check{\boldsymbol{A}}_{0,0} - \mathbf{A}_{0,1}\check{\boldsymbol{A}}_{1,1}^{-1}\check{\boldsymbol{A}}_{1,0} = (\mathbf{A}/\mathbf{A}_{1,1}) \\ &= \check{\boldsymbol{A}}_{0,0} - \check{\boldsymbol{A}}_{0,1}\check{\boldsymbol{A}}_{1,1}^{-1}\check{\boldsymbol{A}}_{1,0} = (\check{\boldsymbol{A}}/\check{\boldsymbol{A}}_{1,1}). \end{aligned} \tag{61}$$

Recall also that the smoother described in Section 2 in connection with the algebraic two-grid scheme agrees exactly with (59). In this sense the above mechanism based on space decompositions simply offers some tuning parameters for generating suitable blockings and then approximate the Schur complements in an appropriate way.

Of course, as in the general two-grid scheme of Section 2, (51) and (58) reflect the ideal situation. In practice $\check{\boldsymbol{L}}$ and $\hat{\boldsymbol{L}}$ are determined only approximately by performing, quite in the spirit of Section 4, a few iterations for solving linear systems with the matrix $\check{\boldsymbol{A}}_{1,1}$ which in turn is supposed to have been preconditioned before. Note that the transformation in (47) offers two degrees of freedom in terms of the matrices $\check{\boldsymbol{L}}$, $\check{\boldsymbol{K}}$ to affect the matrix $\check{\boldsymbol{A}}_{1,1}$ while in (55) one can only exploit $\check{\boldsymbol{K}}$.

We are now ready to describe the multigrid scheme based on the first strategy. The second version is similar.

Multigrid–Cycle:

- **current solution – ℓth iterate in S_h:**

$$\begin{pmatrix} \boldsymbol{u}_H \\ \boldsymbol{d} \end{pmatrix} = \check{\boldsymbol{R}}\boldsymbol{u}_h^\ell$$

- **Pre–smoothing**:

$$\begin{pmatrix} \boldsymbol{u}_H \\ \tilde{\boldsymbol{d}} \end{pmatrix} = \boldsymbol{S}\begin{pmatrix} \boldsymbol{u}_H \\ \boldsymbol{d} \end{pmatrix} + \begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{M}\hat{\boldsymbol{f}}_0 + \boldsymbol{Q}\hat{\boldsymbol{f}}_1 \end{pmatrix} \tag{62}$$

  or equivalently

$$\tilde{\boldsymbol{d}} = \left(\boldsymbol{I} - \boldsymbol{Q}\check{\boldsymbol{A}}_{1,1}\right)\boldsymbol{d} + \boldsymbol{M}\left(\hat{\boldsymbol{f}}_0 - \hat{\check{\boldsymbol{A}}}_{0,0}\boldsymbol{u}_H\right) + \boldsymbol{Q}\left(\hat{\boldsymbol{f}}_1 - \mathbf{A}_{1,0}\boldsymbol{u}_H\right) \tag{63}$$

- **coarse grid correction:**
  - **calculate the residual**

$$\begin{pmatrix} \hat{\boldsymbol{f}}_0 \\ \hat{\boldsymbol{f}}_1 \end{pmatrix} - \hat{\check{\boldsymbol{A}}}\begin{pmatrix} \boldsymbol{u}_H \\ \tilde{\boldsymbol{d}} \end{pmatrix} =: \begin{pmatrix} \boldsymbol{r}_0 \\ \boldsymbol{r}_1 \end{pmatrix}$$

  - **solve the defect equation by μ cycles of the multigrid scheme applied to**

$$\hat{\check{\boldsymbol{A}}}_{0,0}\boldsymbol{e}_H = \boldsymbol{r}_0$$

  - **add the defect to the coarse grid approximation**

$$\tilde{\boldsymbol{u}}_H = \boldsymbol{u}_H + \boldsymbol{e}_H$$

- **Post–smoothing**:

$$\begin{pmatrix} \tilde{\boldsymbol{u}}_H \\ \tilde{\tilde{\boldsymbol{d}}} \end{pmatrix} = \boldsymbol{S}\begin{pmatrix} \tilde{\boldsymbol{u}}_H \\ \tilde{\boldsymbol{d}} \end{pmatrix} + \begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{M}\hat{\boldsymbol{f}}_0 + \boldsymbol{Q}\hat{\boldsymbol{f}}_1 \end{pmatrix} \tag{64}$$

  or equivalently

$$\tilde{\tilde{\boldsymbol{d}}} = \left(\boldsymbol{I} - \boldsymbol{Q}\check{\boldsymbol{A}}_{1,1}\right)\tilde{\boldsymbol{d}} + \boldsymbol{M}\left(\hat{\boldsymbol{f}}_0 - \hat{\check{\boldsymbol{A}}}_{0,0}\tilde{\boldsymbol{u}}_H\right) + \boldsymbol{Q}\left(\hat{\boldsymbol{f}}_1 - \mathbf{A}_{1,0}\tilde{\boldsymbol{u}}_H\right) \tag{65}$$

- **new solution** – $(\ell+1)$**st iterate**:

$$\boldsymbol{u}_h^{\ell+1} = \check{\boldsymbol{P}} \begin{pmatrix} \tilde{\boldsymbol{u}}_H \\ \tilde{\tilde{\boldsymbol{d}}} \end{pmatrix}$$

Note that the prolongation operator $\check{\boldsymbol{P}}$ is only applied in the final step of the cycle. Note also that we could have applied $\boldsymbol{S}$ ν times with $\nu > 1$ as has been done in several experiments below. But damping the error in the complement space is governed by the quality of $\boldsymbol{Q}$ and $\boldsymbol{M}$, i.e., by the effort spent in solving the corresponding subsystems, typically (although not necessarily in an adaptive context) again by an iterative process. This is an important point since this effort can be adapted to the form of the equations depending on the current scale or even *locally* in connection with adaptive refinements. Moreover, the choice of $\boldsymbol{Q}$ is not necessarily tied to the approximate calculation of $\hat{\boldsymbol{L}}$ in (51).

So far we have only outlined a general setting and the corresponding degrees of freedom for problem adaptation. A more detailed discussion of the concrete choice of the various ingredients is given in [DMS97]. Some of the relevant observations will be discussed below in connection with several preliminary numerical experiments.

We conclude this section with some brief comments on several related approaches in the literature. Obviously, the above scheme can be viewed as a Schur-complement-technique and is therefore closely related, for instance, to AMLI [AV89, AV90], or MW–HBMG [VW97].

The use of hierarchical basis decompositions in combination with symmetric block Gauss–Seidel relaxation leads to the HBMG method [BDY88] which has proven to work well for classical elliptic problems. Moreover, recent modifications developed for convection dominated problems combined with an ILU smoother correspond to transformations of the form (55) [BG97a, BG97b]. It correspondes to (55) with $\check{\boldsymbol{K}} = \hat{\boldsymbol{K}} = \boldsymbol{I}$, i.e., only the parameters $\check{\boldsymbol{L}}$ and $\hat{\boldsymbol{L}}$ are incorporated. These parameters are determined by the complete decoupling conditions $\hat{\check{\boldsymbol{A}}}_{0,1} = \boldsymbol{0}$ *and* $\hat{\check{\boldsymbol{A}}}_{1,0} = \boldsymbol{0}$. In the 1D test case the resulting modified restriction and prolongation operators as well as the coarse grid matrix $\hat{\check{\boldsymbol{A}}}_{0,0}$ coincide with those derived by Hackbusch ([Hac85], pp. 213). As already mentioned in the context of [DMS97] the basis functions $\hat{\Phi}$ in the test space, generated by the first strategy have an upwind shark fin shape (see Fig. 5). It is interesting to note that the modified basis functions $\check{\Phi}$ in the trial space have a down–biased shark fin shape when the second strategy is employed. With regard to higher dimensions, the ansatz (48) has turned out to be preferable in our test calculations since convergence rates appear to be better in all examples while in the second strategy (56) efficiency degrades due to a higher fill in for the coarse grid operator.

Likewise the approach in [Reu95, Reu96] aims explicitly at reducing the error components in complement spaces by applying block Jacobi smoother of the form (59).

6 Numerical Experiments for Uniform Refinements

The 1D case is certainly exceptional and one has to be extremely careful about extrapolating results. Nevertheless, since it provides some useful information on the basic strategy, in particular with regard to the stabilizing effect of our adaptive refinement strategy, and also exhibits interesting connections with earlier results we include a brief discussion and some examples. Of course, it is very special in that the complement block $\mathbf{A}_{1,1}$ is diagonal so that the ideal choice (51) for $\hat{\boldsymbol{L}}$ can be calculated explicitly and the inversion of $\check{\boldsymbol{A}}_{1,1} = \mathbf{A}_{1,1}$ becomes trivial. Moreover, starting with piecewise linear continuous basis functions, it can be easily verified that the support of the modified basis functions $\hat{\Phi}$ for the test spaces coincides with those of the initial basis functions Φ. Moreover, they become shark fin shaped with decreasing grid level (see Fig. 5). This results in an increasingly upwind discretization on the coarser levels corresponding to the coarse grid stiffness matrix $\hat{\check{\boldsymbol{A}}}_{0,0}$ (cf. [DMS97]). We emphasize that in the case (47), (48) the trial space is not modified since $\check{\boldsymbol{K}} = \boldsymbol{I}$, $\check{\boldsymbol{L}} = \mathbf{0}$, i.e., the basis functions remain symmetric. As expected the exact invertibility of the complement block gives rise to extremely good error reduction independently of ε and hence to unconditional robustness, see Tab. 2.

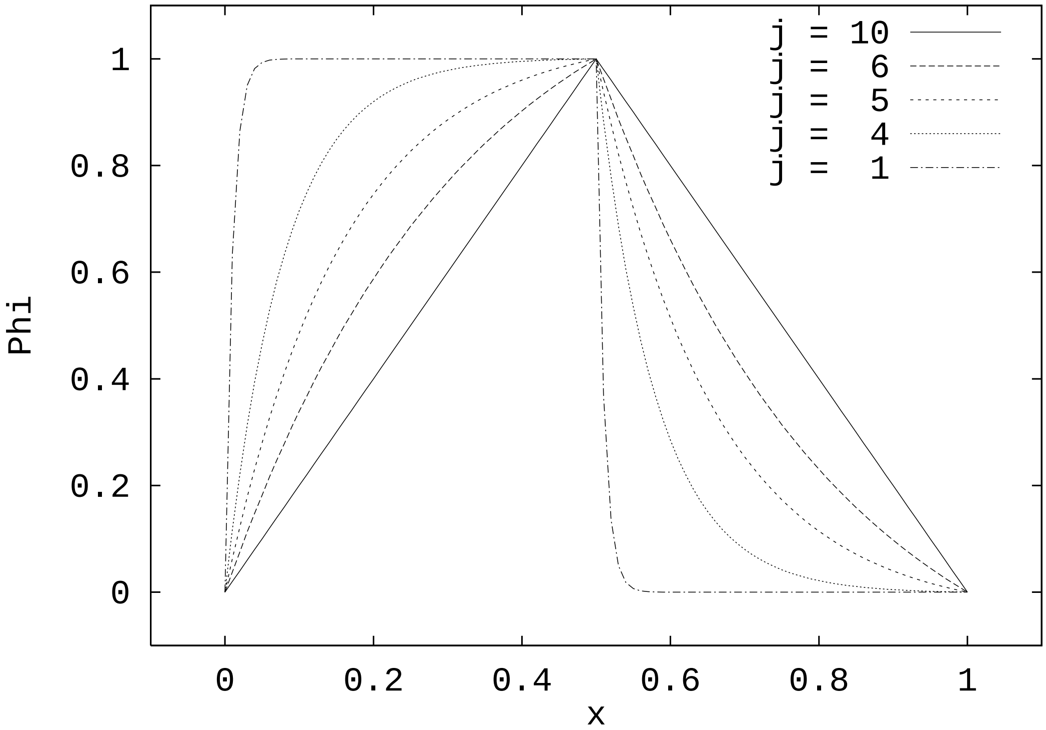

Figure 5: Modified test functions (normalized and shifted), finest level $J = 10$, coarsest level $j = 1$, $\varepsilon = 1$, $\beta = 100$, $d = 1$

It is interesting to note that in the context of interface problems, matrix–dependent prolongation operators are proposed in [Hac85], pp. 213, replacing linear interpolation by certain unsymmetric interpolation schemes. Again the restriction operator is assumed to be the adjoint of the prolongation operator. Hackbusch recommends these operators also for convection dominated problems, since they lead to a stabilization of the multigrid scheme, see [Hac85], pp. 221, as is verified in various other studies, see e.g. [HKWd83, Hd85, Den82, Den83]. It was somewhat surprising to realize that this restriction operator coincides with our modified operator. Moreover, the resulting stiffness matrix of the coarse grid problem are the same, although we apply a Petrov–Galerkin approximation using the initial hierarchical basis for the trial space, i.e., the prolongation operator results from linear interpolation.

Below we have recorded some results of our numerical experiments for the following test examples

(T1) $d = 1$, $\beta = 1$, $f = 1$;

(T2) $d = 1$, $\beta = 1 - 2x$, $f = 1$;

(T3) Rotating flow

$$\beta_1(x,y) = \begin{cases} \sin(\pi(y - \frac{1}{3}))\ \cos(\pi(x - \frac{1}{3})) & , \quad (x,y) \in \Omega_R := B_{(\frac{1}{3},\frac{1}{3})}\left(\frac{1}{4}\right) \\ 0 & , \quad (x,y) \in \Omega_R^C := [0,1]^2 \backslash \Omega_R \end{cases}$$

$$\beta_2(x,y) = \begin{cases} -\cos(\pi(y - \frac{1}{3}))\ \sin(\pi(x - \frac{1}{3})) & , \quad (x,y) \in \Omega_R \\ 0 & , \quad (x,y) \in \Omega_R^C \end{cases}$$

$$f = 0;$$

(T4) $\beta_1(x,y) = xy$, $\beta_2(x,y) = -0.4((1-6x^2)y^2(1-y^2) + (1-6y^2)x^2(1-x^2))$, $f = 1$.

The multigrid version that have been tested are specified in Tab. 1.

Example T1 gives rise to the shark fin test bases displayed in Fig. 5. Example T2 treats still the one dimensional case but with variable $\beta(x)$. It is mainly to support the principal validity of the concept. As expected the exact invertibility of the complement block gives rise to extremely good error reduction independently of ε and hence to unconditional robustness, see Tab. 2.

A frequently employed test example is the rotating flow in T3. Since the exact solution is zero, robustness and efficiency of the scheme is reflected by the speed at which the iterations based on an arbitrary initial guess approach zero. Our first experiments do not exploit the full flexibility of the setting yet. In fact, we have set $\check{\boldsymbol{K}} = \boldsymbol{I}$, $\check{\boldsymbol{L}} = \mathbf{0}$ which leaves the restriction operators for the trial spaces unchanged. Thus there is still room for further adaptation. In spite of this crude choice the results exhibit significant differences. While the modification of test spaces alone increases robustness significantly in comparison with standard variants, employing in addition problem–dependent smoother of the type (53), extends the range of admissible grid Peclet numbers again significantly, as shown in Tab. 3.

T4 presents another bivariate example (see Figs. 6, 7)) with variable $\boldsymbol{\beta}(\boldsymbol{x})$ that will be used later for comparisons with adaptive schemes.

7 An Adaptive Full Multigrid Scheme

The superb efficiency of the above scheme for the 1D case hinges, of course, on the artefact that the complement block $\check{\mathbf{A}}_{1,1}$ is diagonal which will certainly not prevail for more spatial dimensions. Thus a more judicious choice of the modifications of stable completions has to be taken. Again we refer to a more detailed discussion of this issue to [DMS97]. Nevertheless, the current relatively crude choices already provide multigrid efficiency for grid Peclet numbers well above one, see Tab. 3. Since our central objective is to combine robustness with sufficient accuracy this will actually turn out to be fully sufficient in the following context. We wish to explore the stabilizing effect of local refinements in the sense that sufficient resolution corresponding to *local* grid Peclet numbers *only* in layer regions suffice to supress oscillations even when in other regions the mesh is much coarser entailing local grid Peclet numbers that are larger by several orders of magnitude. We wish to achieve this goal by employing full multigrid where initially on coarser levels a problem with added artificial viscosity is treated that will then be successively lowered when moving to higher scales through adaptive refinements. Roughly this scheme can be described as follows:

- Suppose that $u_j = u_j(\varepsilon_j)$ is an approximate solution on j-th grid for a problem with artificial viscosity ε_j such that $P_j \leq 1$.
- Set $\varepsilon_{j+1} = \max\{\varepsilon_j/2, \varepsilon\}$ which defines problem $(j+1)$. Decompose $u_j(\varepsilon_j)$ into coarse part $\check{u}_{j-1}$ and wavelet part $\check{w}_{j-1}$ and apply the problem–dependent smoother for problem $(j+1)$ on the j-th grid to obtain a smoothed approximation u_{j-1}, w_{j-1}.
- At this point the explicit representation of the complement space is exploited. The wavelet coefficients of the high frequency part are direct indicators where the spaces have to be (possibly only) locally refined. Thus one determines the grid refinement from coefficients of the smoothed wavelet part w_{j-1}. This leads to the $(j+1)$-st grid. Since we simply add fine scale nodal basis functions the corresponding meshes contain *hanging nodes* which greatly simplifies the refinement and coarsening process and switching efficiently between local nodal bases and wavelet representations.
- Use the prolongation of $u_{j-1} + w_{j-1}$ to the $(j+1)$-st grid as starting solution for problem $(j+1)$ on the $(j+1)$-st grid and perform a μ-cycle with respect to the current grid hierarchy.
- Repeat until error tolerance is met and the added viscosity has been completely removed.

The scheme works well for variable flow directions $\boldsymbol{\beta}$ including rotating flows. Already simple univariate examples with variable β exhibit the interesting phenomenon that the scale dependent variation of the problem creates an instationary effect in that layers start to move towards their final position. Accordingly, it is not sufficient to refine but also to *coarsen* the mesh during the process. Again wavelet coefficients of the current solution provide the relevant information. In this context several options suggest themselves which are currently investigated. One could use the particular multiscale decomposition created through the multigrid scheme or one could employ some other multiscale basis which is known to have particularly good stability properties with regard to the norm by which the error is measured. In either case critical ingredients of the scheme are fast local basis transforms between various representations. It is however important to note that all these transforms have complexity which stays proportional to the current number of degrees of freedom. More details can be found in [DMS97].

The following numerical examples include again a 1D case. The tremendeous reduction of complexity is essentially independent of the viscosity. The bivariate examples confirm that the scheme works as well for variable $\boldsymbol{\beta}$ showing, in principle, the same effect as in the 1D case. Figures 8 and 9 display the adapted mesh generated by the above scheme for example T1. The convergence behaviour is described by Tab. 4. The columns headed by $\#\Phi_j^{adapt}$ and $\#\Phi_j^{HB}$ show the respective degrees of freedom. ε_j denotes the artificial viscosity on level j, $P(h_j, \varepsilon_j)$ is the local grid Peclet number at locations of current highest resolution while $P(h_4, \varepsilon_j)$ denotes the grid Peclet number on the coarsest level.

We have compared uniform versus adaptive refinements for example T4. Since the complement blocks $\check{\boldsymbol{A}}_{1,1}$ are no longer trivial more cycles than in the univariate case are needed. We emphasize though that our tuning has so far been very crude and leaves much room for further improvement. Some alternative strategies are currently under investigation. In principle, this is also true for the thresholding criteria controling the refinement process. Nevertheless the significant gain in efficiency and accuracy of the adaptive variant is evident, see Tabs. 5, 6. Figures 10, 11 represent the solution for T4 obtained by the adaptive scheme which is to be compared with Figs. 6, 7.

Figure 12 shows the (L_∞–) error of the solution from Fig. 6 (with respect to a numerical solution on a higher level). Likewise Fig. 13 displays the error produced by the adaptive scheme. Although the number of degrees of freedom is lower by a factor 8 the error is smaller by a factor 4. There still seems to be room for equilibrating the error in a better way by further improving the adaptation. The difference in complexity becomes more enhanced when proceeding to higher levels as shown by Figs. 14 and 15, see also Tab. 6.

Decreasing ε to $\varepsilon = 1/192$ the uniform refinements up to level 7 would still create oscillations in the layer region shown by Fig. 16. The multigrid scheme is still in the efficient range. In contrast the solution displayed in Fig. 17 is adequate. It is obtained by the adaptive scheme where the highest level is also 7. It should be noted that no oscillations appear.

scheme	parameters
MG	damped Jacobi smoother ($\omega = 2/3$) standard restriction and prolongation Smoothing Steps: $\nu_1 = 2$ (pre), $\nu_2 = 1$ (post)
UMG	Jacobi smoother upwind restriction and prolongation with $\check{\boldsymbol{L}} = \mathbf{0}$, $\hat{\boldsymbol{K}} = \check{\boldsymbol{K}} = \boldsymbol{I}$ approximation of the inverse of $\mathbf{A}_{1,1} = \hat{\check{\boldsymbol{A}}}_{1,1}$ by one step with the damped Jacobi smoother with $\omega = 2/3$ Smoothing Steps: $\nu_1 = 2$ (pre), $\nu_2 = 1$ (post)
USMG	problem–dependent smoother upwind restriction and prolongation with $\check{\boldsymbol{L}} = \mathbf{0}$, $\hat{\boldsymbol{K}} = \check{\boldsymbol{K}} = \boldsymbol{I}$ approximation of the inverse of $\mathbf{A}_{1,1} = \hat{\check{\boldsymbol{A}}}_{1,1}$ by two steps with the damped Jacobi smoother with $P_j \leq 1 : \omega = 1$, $P_j > 1 : \omega = 2/3$ Smoothing Steps: $\nu_1 = 2$ (pre), $\nu_2 = 1$ (post)
FUSMG	adaptive grid refinements nested iterations with V–cycle USMG

Table 1: Different types of multigrid schemes

j	spectral radius	P_j
4	3.56×10^{-16}	53.33
5	5.11×10^{-17}	29.33
6	1.82×10^{-15}	15.33
7	1.05×10^{-15}	7.83
8	1.07×10^{-15}	3.96
9	1.84×10^{-15}	1.99
10	6.71×10^{-16}	1.00
11	6.58×10^{-17}	0.50
12	6.11×10^{-16}	0.25

Table 2: T2, $\varepsilon = 2^{-11}$, V–cycle USMG, coarsest level = 3

		V–cycle			W–cycle		
P_J	ε	MG	UMG	USMG	MG	UMG	USMG
0.1	0.05000	0.153	0.095	0.075	0.110	0.098	0.075
0.2	0.02500	0.153	0.095	0.075	0.110	0.098	0.075
0.3	0.01667	0.153	0.095	0.075	0.110	0.098	0.075
0.4	0.01250	0.168	0.095	0.075	0.110	0.098	0.075
0.5	0.01000	0.619	0.095	0.075	0.130	0.098	0.075
0.6	0.00833	1.280	0.095	0.075	1.590	0.098	0.075
0.7	0.07143		0.096	0.075		0.098	0.075
0.8	0.00625		0.106	0.075		0.098	0.075
0.9	0.00556		0.130	0.075		0.100	0.075
1.0	0.00500		0.156	0.080		0.129	0.075
1.1	0.00455			0.180			0.180
$\vdots$	$\vdots$			$\vdots$			$\vdots$
2.0	0.00250			0.180			0.180

Table 3: Error reduction number $(\|u_{20}\|_2/\|u_0\|_2)^{1/20}$, initial guess $u_0 = 1$, T3, finest level $J = 6$, coarsest level $j = 2$

level j	cycles	$\#\Phi_j^{adapt}$	$\#\Phi_j^{HB}$	ε_j	$P(h_j, \varepsilon_j)$	$P(h_4, \varepsilon_j)$
4	1	15	15	2^{-4}	0.5	0.5
5	1	28	31	2^{-5}	0.5	1
6	1	41	63	2^{-6}	0.5	2
7	1	54	127	2^{-7}	0.5	4
8	1	67	255	2^{-8}	0.5	8
9	1	80	511	2^{-9}	0.5	16
10	1	93	1023	2^{-10}	0.5	32
11	1	106	2047	2^{-11}	0.5	64
12	1	119	4095	2^{-11}	0.25	64

Table 4: T1, $\varepsilon = 2^{-11}$, FUSMG, coarsest level = 3

level j	cycles	$\#\Phi_j^{adapt}$	$\#\Phi_j^{HB}$	ε_j	$P(h_j, \varepsilon_j)$	$P(h_3, \varepsilon_j)$
3	14	49	49	1/12	0.74	0.74
4	15	225	225	1/24	0.74	1.48
5	15	730	961	1/48	0.74	2.96
6	14	1224	3969	1/96	0.74	5.92
7	13	2127	16129	1/96	0.37	5.92

Table 5: T4, $\varepsilon = 1/96$, FUSMG, coarsest level = 2

level j	cycles	$\#\Phi_j^{adapt}$	$\#\Phi_j^{HB}$	ε_j	$P(h_j, \varepsilon_j)$	$P(h_3, \varepsilon_j)$
3	14	49	49	1/12	0.74	0.74
4	16	225	225	1/24	0.74	1.48
5	14	730	961	1/48	0.74	2.96
6	14	1224	3969	1/96	0.74	5.92
7	13	2127	16129	1/192	0.37	11.84
8	13	3815	65025	1/192	0.37	11.84

Table 6: T4, $\varepsilon = 1/192$, FUSMG, coarsest level = 2

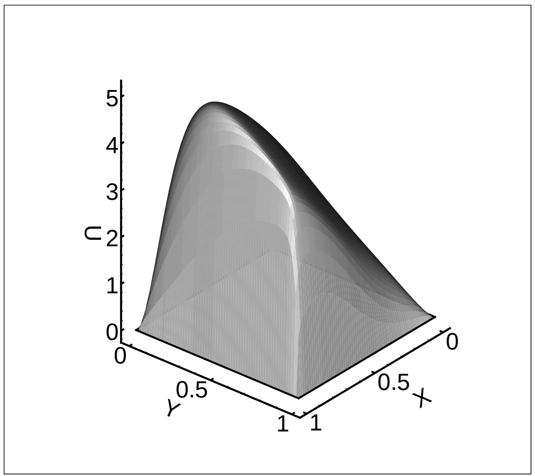

Figure 6: T4, $\varepsilon = 1/96$, V–cycle USMG, $J = 6$ (fine)

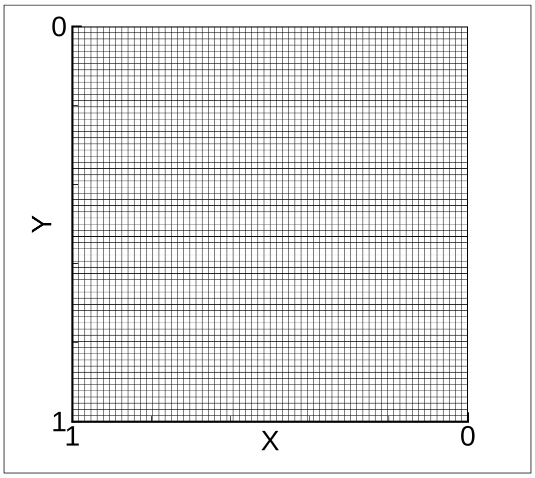

Figure 7: T4, $\varepsilon = 1/96$, V–cycle USMG, $J = 6$ (fine)

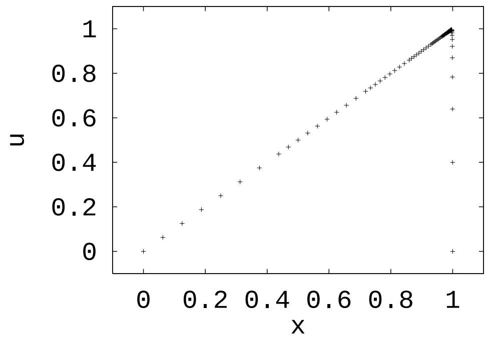

Figure 8: T1, $\varepsilon = 2^{-11}$, FUSMG, $J = 12$

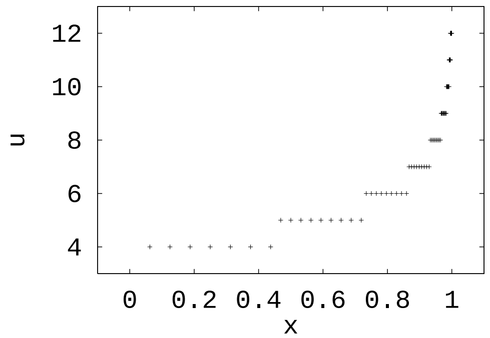

Figure 9: T1, $\varepsilon = 2^{-11}$, FUSMG, $J = 12$

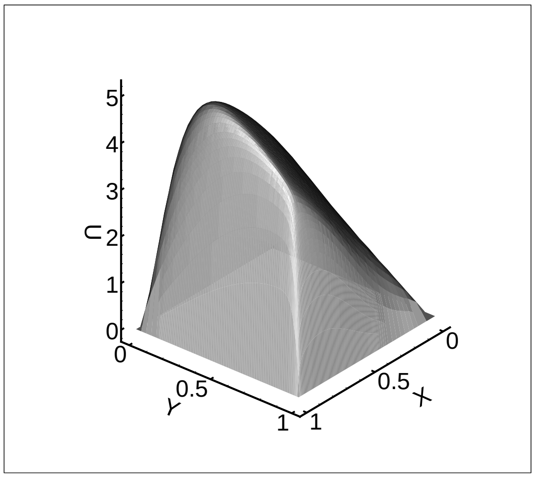

Figure 10: T4, $\varepsilon = 1/96$, FUSMG, $J = 7$ (fine)

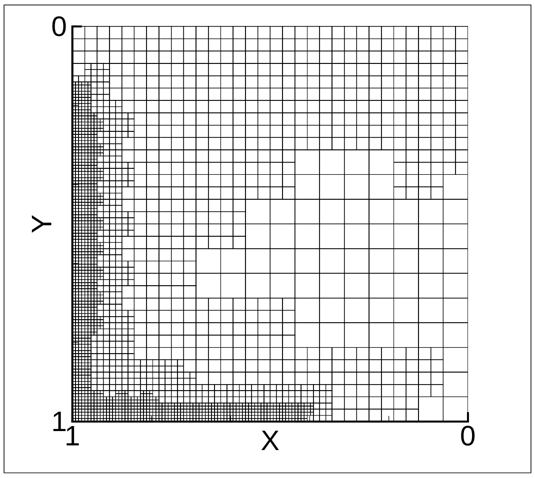

Figure 11: T4, $\varepsilon = 1/96$, FUSMG, $J = 7$ (fine)

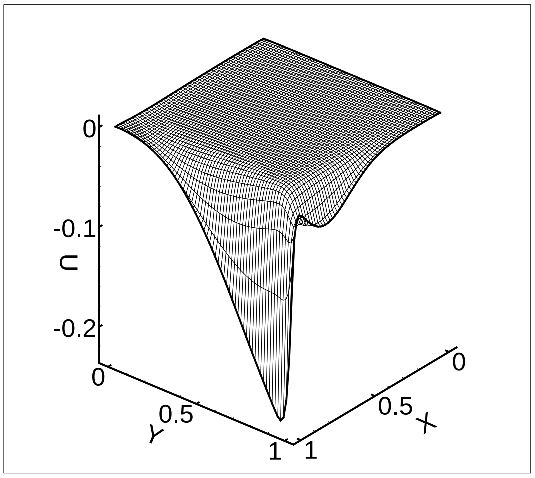

Figure 12: T4, $\varepsilon = 1/96$, V–cycle USMG, $J = 7$ (fine)

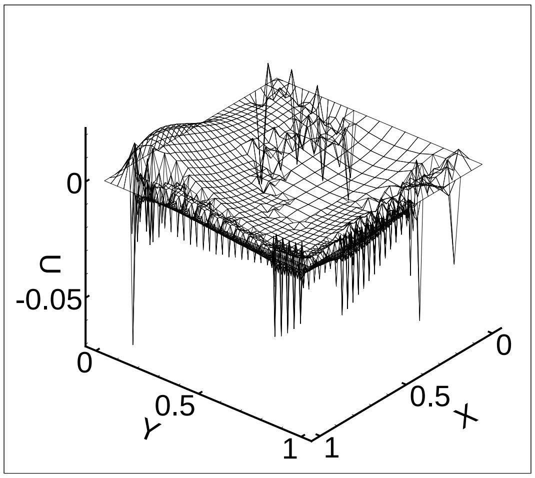

Figure 13: T4, $\varepsilon = 1/96$, FUSMG, $J = 7$ (fine)

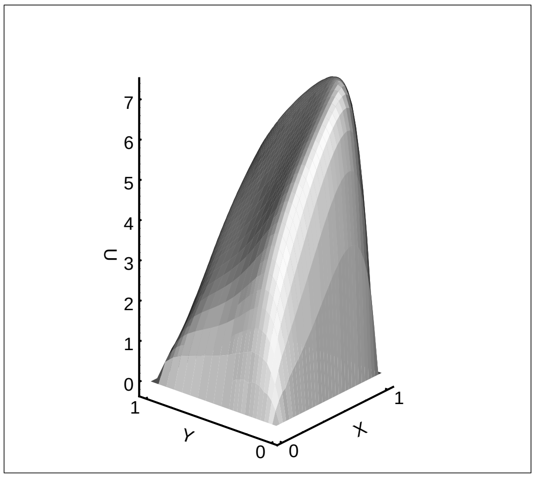

Figure 14: T4, $\varepsilon = 1/192$, FUSMG, $J = 8$ (fine)

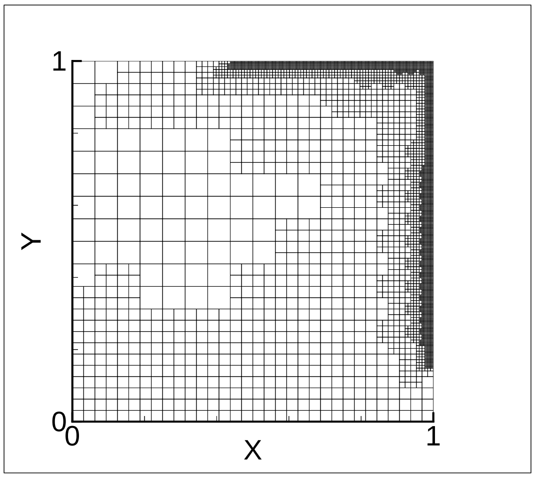

Figure 15: T4, $\varepsilon = 1/192$, FUSMG, $J = 8$ (fine)

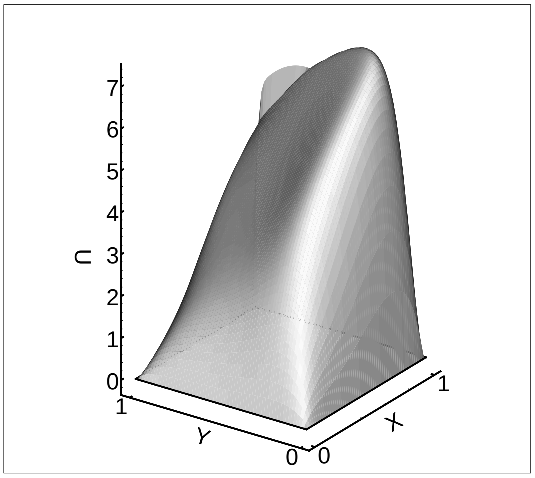

Figure 16: T4, $\varepsilon = 1/192$, FUSMG, $J = 7$ (fine)

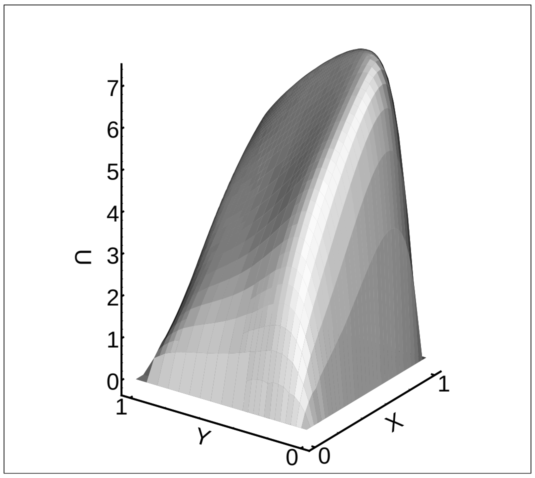

Figure 17: T4, $\varepsilon = 1/192$, FUSMG, $J = 7$ (fine)

References

[AV89] O. Axelsson and P. S. Vassilevski. Algebraic multilevel preconditioning methods I. *Numer. Math.*, 56:157–177, 1989.

[AV90] O. Axelsson and P. S. Vassilevski. Algebraic multilevel preconditioning methods II. *SIAM J. Numer. Anal.*, 27:1569–1590, 1990.

[Axe86] O. Axelsson. A general incomplete block-matrix factorization method. *Linear Algebra and its Applications*, 74:179–190, 1986.

[BDY88] R. E. Bank, T. F. Dupont, and H. Yserentant. The hierarchical basis multigrid method. *Numer. Math.*, 52:427–458, 1988.

[BG97a] R. E. Bank and S. Gutsch. The generalized hierarchical basis two–level method for the convection–diffusion equation on a regular grid. Internal report, Mathematical Seminar II, Christian–Albrechts University Kiel, Germany, 1997.

[BG97b] R. E. Bank and S. Gutsch. Hierarchical basis for the convection–diffusion equation on unstructered meshes. In P. Bjørstad, M. Espedal, and D. Keyes, editors, *Ninth International Symposium on Domain Decomposition Methods for Partial Differential Equations*, New York, 1997. J. Wiley and Sons. To appear.

[Bra77] A. Brandt. Multi–level adaptive solutions to boundary value problems. *Math. Comp.*, 31:333–350, 1977.

[Bra82] A. Brandt. Guide to multigrid development. In W. Hackbusch and U. Trottenberg, editors, *Multigrid Methods*, volume 960 of *Lecture Notes in Mathematics*. Springer, 1982.

[CDP96] J. M. Carnicer, W. Dahmen, and J. M. Peña. Local Decomposition of refinable Spaces and Wavelets. *Appl. Comput. Harmon. Anal.*, 3:127–153, 1996.

[CGMZ76] I. Christie, D. Griffiths, A. Mitchell, and O. Zienkiewicz. Finite element method for second order differential equation with significant first derivatives. *Int. J. Num. Math. Eng.*, 10:1389–1396, 1976.

[Dah97] W. Dahmen. Wavelet and multiscale methods for operator equations. *Acta Numerica*, 6:55–228, 1997.

[DE88] W. Dahmen and L. Elsner. Algebraic multigrid methods and the Schur Complement. In W. Hackbusch, editor, *Robust Multi–Grid Methods*, volume 23 of *Notes on Numerical Fluid Mechanics*, pages 58–68, Braunschweig, 1988. Vieweg.

[Den82] J. E. Dendy Jr. Black box multigrid. *J. Comp. Phys.*, 48:366–386, 1982.

[Den83] J. E. Dendy Jr. Black box multigrid for nonsymmetric problems. *Appl. Math. Comput.*, 13:261–283, 1983.

[DK92] W. Dahmen and A. Kunoth. Multilevel Preconditioning. *Numer. Math.*, 63(2):315–344, 1992.

[DMS97] W. Dahmen, S. Müller, and Th. Schlinkmann. Convection dominated equations, multigrid and stable completions. in preperation, 1997.

[GL79] P. M. Gresho and R. L. Lee. Don't suppress the wiggles — they are telling you something. In Hughes, editor, *Finite Element Methods for Convection Dominated Flows*, volume 34 of *AMD*. ASME, New York, 1979.

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

[HB79] T. J. R. Hughes and A. N. Brooks. A multidimensional upwind scheme with no crosswind diffusion. In Hughes, editor, *Finite Element Methods for Convection Dominated Flows*, volume 34 of *AMD*. ASME, New York, 1979.

[Hd85] P. W. Hemker and P. M. deZeeuw. Some implementations of multigrid linear system solvers. In H. Holstein and D. J. Paddon, editors, *Short Course on multigrid methods. Proceedings, Bristol, Sept. 1983*. Oxford Press, 1985.

[HKWd83] P. W. Hemker, R. Kettler, P. Wesseling, and P. M. deZeeuw. Multigrid methods: Development of fast solvers. *Appl. Math. Comput.*, 13:311–326, 1983.

[HP97] W. Hackbusch and Th. Probst. Downwind Gauss–Seidel smoothing for convection dominated problems. *Num. Lin. Alg. with Applications*, 4(1):85–102, 1997.

[Joh87] C. Johnson. *Numerical solution of partial differential equations by the finite element method*. Cambridge University Press, Cambridge, 1987.

[LO97] R. Lorentz and P. Oswald. Multilevel finite element Riesz bases in Sobolev spaces. In P. Bjørstad, M. Espedal, and D. Keyes, editors, *DD9 Proceedings*, 1997. To appear.

[Reu95] A. Reusken. Fourier analysis of a robust multigrid method for convection–diffusion equations. *Numer. Math.*, 71:365–397, 1995.

[Reu96] A. Reusken. On a robust multigrid solver. *Computing*, 56:303–322, 1996.

[RS86] J. Ruge and K. Stüben. Algebraic multigrid. In S. F. McCormick, editor, *Frontiers in Applied Mathematics*, volume 5. SIAM, Philadelphia, 1986.

[Ste96] R. P. Stevenson. The frequency decomposition multilevel method: A robust additive hierarchical basis preconditioner. *Math. Comp.*, 65:983–997, 1996.

[Swe97] W. Sweldens. The lifting scheme: A construction of second generation wavelets. To appear in SIAM J. Numer. Anal., 1997.

[VW97] P. S. Vassilevski and J. Wang. Stabilizing the hierarchical basis by approximate wavelets I: Theory. *Num. Lin. Alg. with Applications*, 4(1):103–126, 1997.

[Wit89] G. Wittum. On the robustness of ILU smoothing. *SIAM J. SCI. Statist. Comput.*, 10(4):699–717, 1989.

[Yse86] H. Yserentant. On the multilevel splitting of finite element spaces. *Numer. Math.*, 49:379–412, 1986.