Computing 56, 303–322 (1996)

Computing

On a Robust Multigrid Solver

A. Reusken, Eindhoven

Received February 28, 1995; revised July 4, 1995

Abstract — Zusammenfassung

On a Robust Multigrid Solver. We consider a two-grid method based on approximation of the Schur complement. We study the dependence of the two-grid convergence rate on certain problem parameters. As test problems we take the rotated anisotropic diffusion equation and the convection-diffusion equation. Using Fourier analysis we show that for both test problems the two-grid method is robust w.r.t. variation in the relevant problem parameters. For the multigrid method we use a standard W-cycle on coarse grids. This multigrid method then has the same algorithmic structure as a standard multigrid method and is fairly efficient. Moreover, when applied to the two test problems then, as in the two-grid method, we have a strong robustness w.r.t. variation of the problem parameters.

AMS Subject Classification: 65N20, 65N30, 65N55

Key words: Multigrid, elliptic problems, anisotropy, convection-dominated.

Über einen robusten Mehrgittersolver. Wir betrachten eine Zweigittermethode, die auf der Approximation des Schurkomplements basiert. Wir untersuchen die Abhängigkeit der Zweigitter-Konvergenzrate von bestimmten Parametern des Problems. Als Testprobleme wählen wir die rotierte anisotrope Diffusionsgleichung und die Konvektions-Diffusions-Gleichung. Mittels Fourieranalyse zeigen wir, daß die Zweigittermethode robust bezüglich Variationen in den relevanten Parametern des Problems ist. Als Mehrgittermethode verwenden wir einen üblichen W-Zyklus auf groben Gittern. Diese Mehrgittermethode hat dann dieselbe algorithmische Struktur wie eine Standard-Mehrgittermethode und ist effizient. Weiters erhalten wir bei Anwendung auf die zwei Testprobleme ebenso wie für die Zweigittermethode starke Robustheit bezüglich Variation der Problemparameter.

1. Introduction

Today multigrid methods are used in nearly every field where partial differential equations are solved by numerical methods. Concerning the theoretical analysis of multigrid methods different fields of application have to be distinguished. For selfadjoint and coercive linear elliptic boundary value problems the convergence theory has reached a mature, if not its final state, cf. [14], [15]. In other areas the state of the art is (far) less advanced. For example, for convection-dominated problems the development of a satisfactory theoretic analysis is still in its infancy.

In recent years extensive research has been devoted to the development and analysis of multigrid methods for elliptic boundary value problems which con-

verge with an optimal rate, independent of the regularity of the solution (cf. [14], [15], and the references therein). Clearly this is a form of robustness. Another interesting topic with respect to robustness is the dependence of the multigrid convergence rate on certain problem parameters. An example of such a problem parameter is the ratio of anisotropy in an elliptic boundary value problem. There are some recent papers (e.g. [6], [7], [9], [11]) in which multigrid methods are treated that are robust with respect to variation in this anisotropy parameter.

In this paper we treat multigrid robustness with respect to variation in problem parameters. Two important test problems in this field are the discrete rotated anisotropic diffusion equation and the discrete convection-diffusion equation (cf. [13]). In the discrete rotated anisotropic diffusion problem we have parameters h (mesh size), β (rotation angle) and ϵ (ratio of anisotropy). In the discrete convection-diffusion problem we have parameters h (mesh size), β (convection angle) and ϵ/h (Peclet number). An important difference between the two test problems is that the rotated anisotropic diffusion equation yields a symmetric operator, whereas the operator corresponding to the convection-diffusion problem is (strongly) nonsymmetric.

We study a two-grid method which is strongly related to the two-level methods underlying the hierarchical basis multigrid method (HBMG, cf. [4]) and the algebraic multilevel iteration (AMLI, cf. [2,3]). We assume two nested grids ("coarse" and "fine") and on the fine mesh the new mesh points are ordered first and then the coarse grid points. This yields the usual block two by two partitioning of the fine grid stiffness matrix:

$$A_h = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}. \tag{1.1}$$

Block LU-decomposition results in the following identity:

$$J_h \left(I - \begin{bmatrix} \emptyset \\ I \end{bmatrix} \mathcal{S}_H^{-1} [\emptyset \; I] A_h \right) J_h = 0, \tag{1.2a}$$

$$J_h := I - \begin{bmatrix} A_{11}^{-1} & \emptyset \\ \emptyset & \emptyset \end{bmatrix} A_h, \quad \mathcal{S}_H := A_{22} - A_{21} A_{11}^{-1} A_{12}. \tag{1.2b}$$

We consider a two-grid method that results from (1.2) if we solve the A_{11} system in J_h only approximately and if we replace $\mathcal{S}_H^{-1}$ by ωA_H^{-1}, with A_H the coarse grid stiffness matrix and ω a scaling parameter. The main topic of this paper is the robustness of this two-grid method with respect to variation in the parameters of the two test problems. In many cases a robust and efficient approximate solver for the A_{11} system is not hard to find. For example on a rectangular mesh one may use a line solver or a variant of ILU. For the analysis in this paper we assume that the A_{11} systems are solved exactly. Then the spectral radius of the two-grid iteration matrix is given by $\rho(I - \omega A_H^{-1} \mathcal{S}_H)$. We take periodic boundary

conditions in the test problems and apply Fourier analysis to estimate $\rho(I - \omega A_H^{-1}\mathcal{S}_H)$. The Fourier analysis yields that $\rho(I - 0.7A_H^{-1}\mathcal{S}_H) \leq 0.6$ holds for both test problems. This is an important result in this paper, because it shows that for both test problems A_H is a robust preconditioner for the Schur complement $\mathcal{S}_H$. We note that for the rotated anisotropic diffusion equation with a finite element discretization robustness results can be derived from [8], [9], [12], too.

To obtain a feasible method we solve the A_{11} systems approximately by a robust basic iterative solver, and we apply a standard W-cycle on coarse grids. With respect to the implementation we note that the algorithmic structure is the same as in standard multigrid. In the "smoothing" phase (i.e. J_h) only unknowns corresponding to "new" nodes in the fine grid are updated. This is similar to the approach in the HBMG method (cf. [4]). Note, however, that in HBMG a V-cycle is used, whereas we consider a W-cycle. Numerical experiments with the multigrid W-cycle applied to (a large class of) problems similar to the two test problems show convergence factors between 0.3 and 0.6. Apparently the robustness property of the two-grid method also yields a robust multigrid W-cycle. Unfortunately, we do not have a convergence analysis of this W-cycle method.

The remainder of this paper is organized as follows. In Section 2 we derive a two-grid method based on a Schur complement approximation. In Section 3 we present the results of a Fourier analysis for this two-grid method applied to the two test problems. In Section 4 we show the results of numerical experiments with the multigrid W-cycle.

2. Derivation of the Two- and Multigrid Method

In this section we discuss some algebraic aspects of a two- and multigrid method based on a Schur complement approximation. The setting presented here is rather general. Concrete examples are given in Sections 3 and 4.

We consider a second order elliptic linear boundary value problem on a plane polygonal domain Ω. Let Ω_H be a given "coarse" mesh on Ω consisting of triangles or quadrilaterals. By Ω_h we denote the corresponding "fine" mesh that results after a standard refinement of Ω_H. The space of grid functions on $\Omega_H(\Omega_h)$ is denoted by $U_H(U_h)$. In U_H and U_h we use the standard nodal basis. The ordering of the basis functions in U_h is chosen such that the basis functions corresponding to nodes in $\Omega_h \setminus \Omega_H$ are taken first. This induces a partitioning of $u \in U_h$ into two blocks. We assume a given finite element or finite difference discretization method on Ω_h, resulting in a linear system

$$A_h x_h = b_h, \tag{2.1}$$

with $A_h : U_h \to U_h$. The ordering of the nodes yields a block partitioning

$$A_h = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \tag{2.2}$$

in which $[A_{11}\ A_{12}]$ corresponds to the equations in the points of $\Omega_h \setminus \Omega_H$. It is assumed that A_{11} is invertible.

We introduce the following notation, in which we use a block partitioning as in (2.2):

$$J_h := I - \begin{bmatrix} A_{11}^{-1} & \emptyset \\ \emptyset & \emptyset \end{bmatrix} A_h \quad (\text{"block Jacobi"}) \tag{2.3a}$$

$$\mathcal{S}_H := A_{22} - A_{21} A_{11}^{-1} A_{12} \quad (\text{"Schur complement"}) \tag{2.3b}$$

$$p := \begin{bmatrix} * \\ I \end{bmatrix} \quad (\text{"prolongation"}) \tag{2.3c}$$

$$r := [*\ \ I] \quad (\text{"restriction"}). \tag{2.3d}$$

In (2.3c, d) the "$*$" blocks are chosen arbitrarily. A block LU-factorization of A_h yields the identity

$$J_h \big(I - p \mathcal{S}_H^{-1} r A_h \big) J_h = 0. \tag{2.4}$$

The structure in (2.4) is as in a classical two-grid method, but clearly (2.4) does not yield a feasible method. We may try to use an approximation $\tilde{J}_h$ of J_h (e.g. by solving the A_{11} system approximately) and an approximation $\tilde{\mathcal{S}}_H$ of $\mathcal{S}_H$. In general it is rather easy to obtain an efficient approximation of J_h. This will be discussed further in Section 4. For the analysis in this paper we assume $\tilde{J}_h = J_h$, i.e. the systems with A_{11} are solved exactly. A suitable approximation $\tilde{\mathcal{S}}_H$ of $\mathcal{S}_H$ is much harder to obtain. In this paper we mainly consider the obvious choice

$$\tilde{\mathcal{S}}_H = A_H / \omega, \tag{2.5}$$

with A_H a discretization of the same boundary value problem but now on the coarse mesh Ω_H, and ω a scaling factor. Another option is the Galerkin approach

$$\tilde{\mathcal{S}}_H = \hat{r} A_h \hat{p} / \omega, \tag{2.6}$$

with a given prolongation $\hat{p}$ and restriction $\hat{r}$. In Section 3 we will comment on the approach in (2.6).

Because $r A_h J_h$ and $J_h p$ are independent of the choice for the "$*$" blocks in (2.3c, d) we may take

$$p = p_{inj} = \begin{bmatrix} \emptyset \\ I \end{bmatrix}, \qquad r = r_{inj} = [\emptyset \quad I].$$

If we take a Schur complement approximation as in (2.5), this results in a two-grid iteration matrix

$$W_h = J_h \big(I - \omega p_{inj} A_H^{-1} r_{inj} A_h \big) J_h. \tag{2.7}$$

The spectrum of a matrix A is denoted by $\sigma(A)$.

Proposition 2.1. *For W_h as in* (2.7) *the following holds*

$$W_h^k = \begin{bmatrix} -A_{11}^{-1}A_{12} \\ I \end{bmatrix} \left(I - \omega A_H^{-1}\mathcal{S}_H\right)^k r_{inj} \quad (k \in \mathbb{N}), \tag{2.8a}$$

$$\sigma(W_h) = \sigma\left(I - \omega A_H^{-1}\mathcal{S}_H\right) \cup \{0\}. \tag{2.8b}$$

Proof: Note that J_h can be written as

$$J_h = \begin{bmatrix} -A_{11}^{-1} A_{12} \\ I \end{bmatrix} r_{inj}.$$

Thus we get $J_h^2 = J_h$, $r_{inj}A_h J_h = \mathcal{S}_H r_{inj}$. Using this it is easy to show that the results in (2.8a, b) hold. ■

As expected, the two-grid convergence is determined by the quality of A_H as a preconditioner for $\mathcal{S}_H$. In Section 3 we will analyze $\sigma(A_H^{-1}\ \mathcal{S}_H)$ for a specified class of problems.

For the multigrid method we use the standard W-cycle approach (cf. [5]). So if $\bar{W}_H$ is the iteration matrix on the coarse grid, then for the iteration matrix $\bar{W}_h$ of the W-cycle we have

$$\bar{W}_h = J_h\left(I - \omega p_{inj}\left(I - \bar{W}_H^2\right)A_H^{-1} r_{inj} A_h\right) J_h. \tag{2.9}$$

In the remarks below we compare the approach from above with similar other methods.

Remark 2.2. Comparing (2.9) with a standard multigrid W-cycle (cf. [5]) we note that the algorithmic structure of a method based on (2.9) is the same as for a standard multigrid W-cycle. However, instead of the usual smoothing iterations we now apply a (approximate) block Jacobi iteration in the points of $\Omega_h \setminus \Omega_H$. Another difference with standard multigrid is that in (2.9) the correction from the coarse grid is scaled with a factor ω. Finally we note that in standard multigrid one should not use p_{inj}/r_{inj} for the prolongation/restriction.

Remark 2.3. The approach above is similar to the algebraic multilevel preconditioning methods (AMLI) of Axelsson and Vassilevski (cf. [2, 3]). For comparison with AMLI we formulate the multigrid method based on (2.9) in terms of a preconditioner of A_h. If we define the preconditioner M_H of A_H by $\bar{W}_H = I - M_H^{-1}A_H$ ($\bar{W}_H$ iteration matrix on the coarse grid as in (2.9)) and M_h by $\bar{W}_h = I - M_h^{-1}A_h$ ($\bar{W}_h$ as in (2.9)), then some rewriting yields

$$M_h = \begin{bmatrix} A_{11} & \emptyset \\ A_{21} & \tilde{A}_H \end{bmatrix} \begin{bmatrix} I & A_{11}^{-1} A_{12} \\ \emptyset & I \end{bmatrix},$$

with $\tilde{A}_H^{-1} = \omega(I - p_2(M_H^{-1} A_H))A_H^{-1}$, $p_2(t) = (1-t)^2$. Only if $\omega = 1$, this is of the form as in version (ii) in [2]. Below in Section 3 we will see that the parameter ω plays an important role in our method and that we will use $\omega \neq 1$. Note that w.r.t. implementation our approach is different from AMLI. The

method based on (2.9) is implemented as a standard multigrid W-cycle (cf. §4), whereas the AMLI method is based on a recursively defined polynomial preconditioner, using a shifted Chebychev polynomial of degree 2 or 3.

Remark 2.4. The method corresponding to (2.7), (2.9) can also be derived using a suitable basis transformation (cf. hierarchical basis two-grid method, e.g. [4]). If we introduce the matrix dependent basis transformations

$$S_L := \begin{bmatrix} I & \emptyset \\ -A_{21}A_{11}^{-1} & I \end{bmatrix}, \quad S_R := \begin{bmatrix} I & -A_{11}^{-1}A_{12} \\ \emptyset & I \end{bmatrix},$$

then on the transformed bases we have a block diagonal matrix:

$$B_h := S_L A_h S_R = \begin{bmatrix} A_{11} & \emptyset \\ \emptyset & \mathscr{S}_H \end{bmatrix}.$$

To the transformed problem with matrix B_h we apply the iterative method with iteration matrix

$$I - \begin{bmatrix} A_{11}^{-1} & \emptyset \\ \emptyset & \omega A_H^{-1} \end{bmatrix} B_h.$$

Reformulation of this method in terms of the original nodal basis yields the method with iteration matrix as in (2.7). With respect to implementation we note that in the method based on (2.9) in the "smoothing" phase (i.e. J_h) only unknowns corresponding to the "new" nodes in the fine grid are updated. This is similar to the approach in the hierarchical basis multigrid method (HBMG, [4]). Note, however, that HBMG is a multigrid V-cycle, whereas we consider a multigrid W-cycle.

3. Two-Grid Fourier Analysis

Numerical experiments (e.g. as in Section 4, cf. also [9], [10]) have shown that the method based on (2.9) has interesting robustness properties both with respect to the amount of anisotropy and the amount of convection in the problem. For a first analysis of this phenomenon we take two relevant classes of test problems as in [13] and we apply a two-grid Fourier analysis. We note that by means of an analysis of $\sigma(A_H^{-1}\mathscr{S}_H)$ and using Proposition 2.1 we can do a Fourier analysis of the complete two-grid iteration, whereas in [13] only the smoothing part of a two-grid method is analyzed using Fourier analysis. On $\Omega = [-1, 1] \times [-1, 1]$ we consider the *rotated anisotropic diffusion equation*

$$-(\epsilon\hat{c}^2 + \hat{s}^2)u_{xx} - 2(\epsilon - 1)\hat{c}\hat{s}u_{xy} - (\epsilon\hat{s}^2 + \hat{c}^2)u_{yy} = 0, \tag{3.1}$$

and the *convection-diffusion equation*

$$-\epsilon(u_{xx} + u_{yy}) + \hat{c}u_x + \hat{s}u_y = 0, \tag{3.2}$$

with $\hat{c} = \cos(\beta)$, $\hat{s} = \sin(\beta)$, $\beta \in [0, \frac{1}{2}\pi]$, $\epsilon \in (0, 1]$.

For the Fourier analysis we assume periodic boundary conditions. We use a uniform square grid with mesh size h:

$$\Omega_h := \{(x,y) \in \Omega | x = \nu h,\ y = \mu h,\ -N+1 \le \nu, \mu \le N\}, \tag{3.3}$$

with $N := 1/h$. The differential operators are replaced by standard difference stars as in [13]:

$$-\frac{\partial^2}{\partial x^2} \to h^{-2}[-1 \quad 2 \quad -1], \quad -\frac{\partial^2}{\partial y^2} \to h^{-2}\begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix}$$

$$-\frac{\partial^2}{\partial x \partial y} \to \tfrac{1}{4}h^{-2}\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix},$$

$$\frac{\partial}{\partial x} \to h^{-1}[-1 \quad 1 \quad 0], \quad \frac{\partial}{\partial y} \to h^{-1}\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}.$$

Thus for the rotated anisotropic diffusion equation we obtain, after scaling with h^2, a discrete system of equations with stencil

$$[A_h] = (\epsilon\hat{c}^2 + \hat{s}^2)[-1 \quad 2 \quad -1] + (\epsilon\hat{s}^2 + \hat{c}^2)\begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix}$$

$$+ \tfrac{1}{2}(\epsilon - 1)\hat{c}\hat{s}\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}. \tag{3.4}$$

For the Fourier analysis we use the standard approach (cf. [5]). In $\ell^2(\Omega_h)$, with $Nh = 1$, we introduce $4N^2$ basis vectors $e_h^{\nu\mu}$ with

$$e_h^{\nu\mu}(x,y) = \tfrac{1}{2}e^{\pi i(\nu x + \mu y)},\ (x,y) \in \Omega_h,\ -N+1 \le \nu,\ \mu \le N. \tag{3.5}$$

For the coarse grid space we define $H := 2h$, $N_H := N/2$ and Ω_H as in (3.3) with h and N replaced by H and N_H respectively. In $\ell^2(\Omega_H)$ we use the Fourier basis

$$e_H^{\nu\mu}(x,y) = \tfrac{1}{2}e^{\pi i(\nu x + \mu y)}, \quad (x,y) \in \Omega_H, \quad -N_H + 1 \le \nu, \quad \mu \le N_H. \tag{3.6}$$

The vectors in (3.5) form an orthonormal basis w.r.t. a scaled Euclidean inner product, and thus the Fourier transform

$$Q_h : (\alpha_{\nu\mu})_{-N+1 \le \nu,\mu \le N} \to \sum_{\nu,\mu=-N+1}^{N} \alpha_{\nu,\mu} e_h^{\nu\mu}$$

is unitary. Every "low" frequency (ν, μ) with $-N_H + 1 \le \nu, \mu \le N_H$ is associated with the "high" frequencies (ν', μ), (ν, μ'), (ν', μ') where ν', μ' are defined by

$$\nu' = \begin{cases} \nu + N & \text{if } \nu \le 0 \\ \nu - N & \text{if } \nu > 0 \end{cases}, \quad \mu' = \begin{cases} \mu + N & \text{if } \mu \le 0 \\ \mu - N & \text{if } \mu > 0. \end{cases}$$

Clearly $\ell^2(\Omega_h)$ is a direct sum of the N^2 subspaces $U_h^{\nu\mu} := \operatorname{span}\{e_h^{\nu\mu}, e_h^{\nu'\mu}, e_h^{\nu\mu'}, e_h^{\nu'\mu'}\}$, $-N_H + 1 \le \nu, \mu \le N_H$. By $Q_h^{\nu\mu}$ we denote the $4N^2 \times 4$ matrix with columns these basis vectors of $U_h^{\nu\mu}$:

$$Q_h^{\nu\mu} := \left[e_h^{\nu\mu} \; e_h^{\nu'\mu} \; e_h^{\nu\mu'} \; e_h^{\nu'\mu'} \right].$$

Now note that we have $(Q_h^{\nu\mu})^* A_h Q_h^{\nu\mu} = \operatorname{diag}(d_1^{\nu\mu}, d_2^{\nu\mu}, d_3^{\nu\mu}, d_4^{\nu\mu})$ (we use the adjoint w.r.t. the scaled Euclidean inner product). For the operator A_h as in (3.4) we obtain the following formulas for the eigenvalues $d_j^{\nu\mu}$ $(-N_H + 1 \le \nu, \mu \le N_H)$:

$$d_1^{\nu\mu} = 4\left(K_{xx} s_\nu^2 + K_{yy} s_\mu^2 + K_{xy} s_\nu c_\nu s_\mu c_\mu \right) \tag{3.7a}$$

$$d_2^{\nu\mu} = 4\left(K_{xx} c_\nu^2 + K_{yy} s_\mu^2 - K_{xy} s_\nu c_\nu s_\mu c_\mu \right) \tag{3.7b}$$

$$d_3^{\nu\mu} = 4\left(K_{xx} s_\nu^2 + K_{yy} c_\mu^2 - K_{xy} s_\nu c_\nu s_\mu c_\mu \right) \tag{3.7c}$$

$$d_4^{\nu\mu} = 4\left(K_{xx} c_\nu^2 + K_{yy} c_\mu^2 + K_{xy} s_\nu c_\nu s_\mu c_\mu \right) \tag{3.7d}$$

with

$$K_{xx} = \epsilon \hat{c}^2 + \hat{s}^2, \quad K_{yy} = \hat{c}^2 + \epsilon \hat{s}^2, \quad K_{xy} = 2(\epsilon - 1)\hat{c}\hat{s},$$

$$s_k = \sin\left(\tfrac{1}{2} k\pi h\right), \quad c_k = \cos\left(\tfrac{1}{2} k\pi h\right). \tag{3.7e}$$

Note that for $-N_H + 1 \le k \le N_H$ we have $s_k \in [-\frac{1}{2}\sqrt{2}, \frac{1}{2}\sqrt{2}]$, $c_k \in [\frac{1}{2}\sqrt{2}, 1]$, $|s_k| \le c_k$.

We now introduce the *harmonic mean* of the eigenvalues $d_j^{\nu\mu}$, $j = 1, 2, 3, 4$. For $(\nu, \mu) \neq (0, 0)$ we define

$$\mathscr{H}_{\nu\mu} = 4\left(\sum_{j=1}^{4} 1/d_j^{\nu\mu} \right)^{-1} \quad (-N_H + 1 \le \nu, \mu \le N_H). \tag{3.8}$$

If $A_h : \ell^2(\Omega_h) \to \ell^2(\Omega_h)$ would be nonsingular, then the identity

$$\mathscr{S}_H = \left([\emptyset \; I] A_h^{-1} \begin{bmatrix} \emptyset \\ I \end{bmatrix} \right)^{-1}$$

holds. Fourier transformation of p_{inj}, A_h, r_{inj} then results in block-diagonal matrices with diagonal blocks

$$\tfrac{1}{4} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} d_1^{\nu\mu} & & & \emptyset \\ & d_2^{\nu\mu} & & \\ & & d_3^{\nu\mu} & \\ \emptyset & & & d_4^{\nu\mu} \end{bmatrix} \quad \text{and} \quad [1\,1\,1\,1],$$

respectively. This then yields $\mathscr{S}_H e_H^{\nu\mu} = \mathscr{K}_{\nu\mu} e_H^{\nu\mu}$. However, in the case with periodic boundary conditions A_h is singular and a special treatment of the constant grid function e_H^{00} is needed. This then leads to the following result (as in [10]).

Lemma 3.1. *The Fourier mode* $e_H^{\nu\mu}$ $(-N_H + 1 \le \nu,\ \mu \le N_H,\ (\nu,\mu) \neq (0,0))$ *is an eigenvector of* $\mathscr{S}_H$ *with corresponding eigenvalue the harmonic mean of* $d_1^{\nu\mu}$, $d_2^{\nu\mu}$, $d_3^{\nu\mu}$, $d_4^{\nu\mu}$, *i.e.*:

$$\mathscr{S}_H e_H^{\nu\mu} = \mathscr{K}_{\nu\mu} e_H^{\nu\mu}. \tag{3.9}$$

Proof: The grid function in $\ell^2(\Omega_h)$ ($\ell^2(\Omega_H)$) with value 1 in all grid points of Ω_h (Ω_H) is denoted by 1_h (1_H). We define

$$r := \left[-A_{21} A_{11}^{-1}\ I\right], \quad p = \begin{bmatrix} -A_{11}^{-1} A_{12} \\ I \end{bmatrix}.$$

Due to $\operatorname{rank}(A_h) = \operatorname{rank}(A_{11}) + \operatorname{rank}(\mathscr{S}_H)$ we have that $\dim(\operatorname{Ker}(\mathscr{S}_H)) = 1$. Note that $\mathscr{S}_H = r A_h p$ and $A_h 1_h = A_h^T 1_h = 0$, $p 1_H = 1_h$, $r^T 1_h = 1_h$. Using this we obtain $\operatorname{Ker}(\mathscr{S}_H) = \langle 1_H \rangle$, $\mathscr{R}(\mathscr{S}_H) = \operatorname{Ker}(\mathscr{S}_H^T)^{\perp} = 1_H^{\perp}$ and thus $\mathscr{S}_H : 1_H^{\perp} \to 1_H^{\perp}$ is regular; the inverse is denoted by $\mathscr{S}_H^{-1}$. Now define $W : 1_h^{\perp} \to \ell^2(\Omega_h)$ by

$$W = p \mathscr{S}_H^{-1} r + \begin{bmatrix} I \\ \emptyset \end{bmatrix} A_{11}^{-1} [I\ \emptyset].$$

Note that W is well-defined due to $r(1_h^{\perp}) \subset 1_H^{\perp}$. A simple calculation shows that $A_h W = I_{|1_h^{\perp}}$. Using $p_{inj}(1_H^{\perp}) \subset 1_h^{\perp}$ we see that $W p_{inj} : 1_H^{\perp} \to \ell^2(\Omega_h)$ is well-defined. From the definition of W we now conclude that $r_{inj} W p_{inj|1_H^{\perp}} = \mathscr{S}_H^{-1}$. Note that

$$r_{inj} e_h^{\nu\mu} = r_{inj} e_h^{\nu'\mu} = r_{inj} e_h^{\nu\mu'} = r_{inj} e_h^{\nu'\mu'} = e_H^{\nu\mu} \quad (-N_H + 1 \le \nu, \mu \le N_H). \tag{3.10}$$

We take a Fourier mode $e_H^{\nu\mu} \in 1_H^{\perp}$, i.e. $(\nu,\mu) \neq (0,0)$. Using $p_{inj} = r_{inj}^T = \frac{1}{4} r_{inj}^*$ and (3.10) we obtain

$$p_{inj} e_H^{\nu\mu} = \tfrac{1}{4}\left(e_h^{\nu\mu} + e_h^{\nu'\mu} + e_h^{\nu\mu'} + e_h^{\nu'\mu'}\right). \tag{3.11}$$

From $A_h W = I_{|1_h^{\perp}}$ we get $A_h W e_h^{\nu\mu} = e_h^{\nu\mu}$ and thus $W e_h^{\nu\mu} = 1/d_1^{\nu\mu}\, e_h^{\nu\mu} + \alpha_{\nu\mu} 1_h$ for a certain $\alpha_{\nu\mu} \in \mathbb{C}$. Similar relations hold for $e_h^{\nu'\mu}$, $e_h^{\nu\mu'}$ and $e_h^{\nu'\mu'}$. Combining this with (3.11) yields

$$W p_{inj} e_H^{\nu\mu} = \tfrac{1}{4}\left(1/d_1^{\nu\mu}\, e_h^{\nu\mu} + 1/d_2^{\nu\mu}\, e_h^{\nu'\mu} + 1/d_3^{\nu\mu}\, e_h^{\nu\mu'} + 1/d_4^{\nu\mu}\, e_h^{\nu'\mu'}\right) + \beta_{\nu\mu} 1_h$$

for a certain $\beta_{\nu\mu} \in \mathbb{C}$. Using (3.10) and $r_{inj} 1_h = 1_H$ we obtain

$$r_{inj} W p_{inj} e_H^{\nu\mu} = \left(\frac{1}{4} \sum_{j=1}^{4} 1/d_j^{\nu\mu}\right) e_H^{\nu\mu} + \beta_{\nu\mu} 1_H.$$

Finally, because $r_{inj} W p_{inj|1_H^{\perp}} = \mathscr{S}_H^{-1} : 1_H^{\perp} \to 1_H^{\perp}$ we conclude $\beta_{\nu\mu} = 0$ and

$$\mathscr{S}_H^{-1} e_H^{\nu\mu} = \left(\frac{1}{4} \sum_{j=1}^{4} 1/d_j^{\nu\mu}\right) e_H^{\nu\mu}.$$

■

Lemma 3.1 makes it possible to apply a Fourier analysis to $A_H^{-1}\mathscr{S}_H:(e_H^{00})^\perp \to (e_H^{00})^\perp$. For the rotated anisotropic diffusion equation the coarse grid operator A_H, corresponding to A_h in (3.4), has eigenvalues (cf. (3.7)):

$$d_H^{\nu\mu} := 16\Big\{K_{xx}s_\nu^2c_\nu^2 + K_{yy}s_\mu^2c_\mu^2 + K_{xy}s_\nu c_\nu\big(1-2s_\nu^2\big)s_\mu c_\mu\big(1-2s_\mu^2\big)\Big\}. \quad (3.12)$$

Now, as an example, consider the case with $\beta = 0$, i.e.

$$[A_h] = \epsilon[-1\ 2\ -1] + \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix}.$$

We only consider $(\nu, \mu) \neq (0,0)$ with $-N_H + 1 \leq \nu, \mu \leq N_H$. It is easy to see that $1 + d_1^{\nu\mu}(1/d_2^{\nu\mu} + 1/d_3^{\nu\mu} + 1/d_4^{\nu\mu}) \in (1,4)$ and thus $\mathscr{K}_{\nu\mu}/d_1^{\nu\mu} \in (1,4)$, i.e. the (ν,μ)-th eigenvalue of the Schur complement $(\mathscr{K}_{\nu\mu})$ is of the same order of magnitude as the (ν,μ)-th eigenvalue of A_h $(d_1^{\nu\mu} = 4(\epsilon s_\nu^2 + s_\mu^2))$, with constants independent of the parameters ϵ, h. Because $d_1^{\nu\mu}$ is the eigenvalue corresponding to a low frequency, it can be expected that $d_1^{\nu\mu}$ is of the same order of magnitude as $d_H^{\nu\mu}$. Indeed, we have

$$\frac{d_1^{\nu\mu}}{d_H^{\nu\mu}} = \frac{4\big(\epsilon s_\nu^2 + s_\mu^2\big)}{16\big(\epsilon s_\nu^2 c_\nu^2 + s_\mu^2 c_\mu^2\big)} \in \Big[\tfrac{1}{4}, \tfrac{1}{2}\Big].$$

Thus we have that for $(\nu,\mu) \neq (0,0)$ $A_H^{-1}\mathscr{S}_H e_H^{\nu\mu} = \xi^{\nu\mu} e_H^{\nu\mu}$ with $\xi^{\nu\mu} \in [\frac{1}{4}, 2]$, i.e. A_H is a preconditioner of $\mathscr{S}_H$ that is robust w.r.t. variation of ϵ and h. Analogous results are derived for the rotated anisotropic diffusion equation with $\beta \neq 0$ in Lemma 3.2, 3.3 and Theorem 3.4. In essence similar results hold for the convection-diffusion equation, too, but then we have to take into account that the eigenvalues are complex. This will be discussed below.

Lemma 3.2. *For all* $(\nu,\mu) \neq (0,0)$ *with* $-N_H + 1 \leq \nu, \mu \leq N_H$ *the following holds, with* $d_1^{\nu\mu}$ *and* $\mathscr{K}_{\nu\mu}$ *as in* (3.7), (3.8):

$$\mathscr{K}_{\nu\mu}/d_1^{\nu\mu} \in \Big[1 - \tfrac{1}{2}\sqrt{2}, 4\Big].$$

Proof: First note that $\mathscr{K}_{\nu\mu}/d_1^{\nu\mu} = 4(1 + d_1^{\nu\mu}/d_2^{\nu\mu} + d_1^{\nu\mu}/d_3^{\nu\mu} + d_1^{\nu\mu}/d_4^{\nu\mu})^{-1}$. Using $d_j^{\nu\mu} > 0$ for $j = 1, 2, 3, 4$ we obtain $\mathscr{K}_{\nu\mu}/d_1^{\nu\mu} \leq 4$. Furthermore:

$$\begin{aligned} d_1^{\nu\mu} &= 4\Big\{K_{xx}s_\nu^2 + K_{yy}s_\mu^2 + K_{xy}s_\nu c_\nu s_\nu c_\mu\Big\} \\ &\leq 4\Big\{K_{xx}c_\nu^2 + K_{yy}c_\mu^2 + K_{xy}s_\nu c_\nu s_\mu c_\mu\Big\} = d_4^{\nu\mu}. \end{aligned} \quad (3.13)$$

We now derive an upper bound for $d_1^{\nu\mu}/d_2^{\nu\mu}$. Due to $\epsilon \in (0,1]$ we have $|K_{xy}| \leq 2|\hat{c}|\,|\hat{s}|$. Using this and $|s_\nu| \leq \frac{1}{2}\sqrt{2}$ we get

$$\begin{aligned} |K_{xy}s_\nu c_\nu s_\mu c_\mu| &\leq \tfrac{1}{2}\sqrt{2}\ 2|\hat{s}c_\nu|\,|\hat{c}s_\mu| \\ &\leq \tfrac{1}{2}\sqrt{2}\left(\hat{s}^2c_\nu^2 + \hat{c}^2s_\nu^2\right) \leq \tfrac{1}{2}\sqrt{2}\left(K_{xx}c_\nu^2 + K_{yy}s_\mu^2\right). \end{aligned} \quad (3.14)$$

This yields

$$d_2^{\nu\mu} = 4\left\{K_{xx}c_\nu^2 + K_{yy}s_\mu^2 - K_{xy}s_\nu c_\nu s_\mu c_\mu\right\}$$
$$\geq 4\left(1 - \tfrac{1}{2}\sqrt{2}\right)\left(K_{xx}c_\nu^2 + K_{yy}s_\mu^2\right). \tag{3.15}$$

Using (3.14), (3.15) results in the following

$$\frac{d_1^{\nu\mu}}{d_2^{\nu\mu}} \leq \frac{1}{1 - \frac{1}{2}\sqrt{2}} \frac{K_{xx}s_\nu^2 + K_{yy}s_\mu^2 + K_{xy}s_\nu c_\nu s_\mu c_\mu}{K_{xx}c_\nu^2 + K_{yy}s_\mu^2}$$
$$\leq \frac{1}{1 - \frac{1}{2}\sqrt{2}}\left(1 + \tfrac{1}{2}\sqrt{2}\right) = 3 + 2\sqrt{2}\,.$$

A symmetry argument yields $d_1^{\nu\mu}/d_3^{\nu\mu} \leq 3 + 2\sqrt{2}$. Thus we finally have

$$\mathscr{K}_{\nu\mu}/d_1^{\nu\mu} \geq 4\left(2 + 2\left(3 + 2\sqrt{2}\right)\right)^{-1} = 1 - \tfrac{1}{2}\sqrt{2}\,.$$

□

Lemma 3.3. For all $(\nu, \mu) \neq (0,0)$ with $-N_H + 1 \leq \nu, \mu \leq N_H$ the following holds, with $d_1^{\nu\mu}$ and $d_H^{\nu\mu}$ as in (3.7), (3.12):

$$d_1^{\nu\mu}/d_H^{\nu\mu} \in \left[\tfrac{1}{16}, 1\right]. \tag{3.16}$$

Proof: Take $(\nu, \mu) \neq (0,0)$, $-N_H + 1 \leq \nu, \mu \leq N_H$. We use the notation $z := K_{xy}s_\nu c_\nu s_\mu c_\mu$. Then we have:

$$d_1^{\nu\mu} = 4\left\{K_{xx}s_\nu^2 + K_{yy}s_\mu^2 + z\right\},$$
$$d_H^{\nu\mu} = 16\left\{K_{xx}s_\nu^2 c_\nu^2 + K_{yy}s_\mu^2 c_\mu^2 + z\left(1 - 2s_\nu^2\right)\left(1 - 2s_\mu^2\right)\right\}.$$

Note that

$$z \leq 2|\hat{s}s_\nu|\,|\hat{c}s_\mu| \leq \hat{s}^2 s_\nu^2 + \hat{c}^2 s_\mu^2 \leq K_{xx}s_\nu^2 + K_{yy}s_\mu^2 \tag{3.17}$$

holds. First we consider the case $z \geq 0$. Then we have, using (3.17) and $c_\nu^2, c_\mu^2 \geq \frac{1}{2}$:

$$d_1^{\nu\mu} \leq 8\left(K_{xx}s_\nu^2 + K_{yy}s_\mu^2\right)$$
$$\leq 16\left(K_{xx}s_\nu^2 c_\nu^2 + K_{yy}s_\mu^2 c_\mu^2 + z\left(1 - 2s_\nu^2\right)\left(1 - 2s_\mu^2\right)\right)$$
$$= d_H^{\nu\mu} \leq 16\left(K_{xx}s_\nu^2 + K_{yy}s_\mu^2 + z\right) = 4d_1^{\nu\mu}.$$

So then we obtain

$$d_1^{\nu\mu}/d_H^{\nu\mu} \in \left[\tfrac{1}{4}, 1\right]. \tag{3.18}$$

We now assume $z < 0$. Then we get

$$d_1^{\nu\mu} \leq 4\left\{K_{xx}s_\nu^2 + K_{yy}s_\mu^2 + z\left(1 - 2s_\nu^2\right)\left(1 - 2s_\mu^2\right)\right\}$$
$$= \tfrac{1}{4}d_H^{\nu\mu} + 4\left(K_{xx}s_\nu^4 + K_{yy}s_\mu^4\right). \tag{3.19}$$

Also we note that

$$\begin{aligned} d_H^{\nu\mu} &= 16\Big\{K_{xx}s_\nu^2c_\nu^2 + K_{yy}s_\mu^2c_\mu^2 + 2(\epsilon-1)\big[\hat{s}s_\nu c_\nu(1-2s_\nu^2)\big]\big[\hat{c}s_\mu c_\mu(1-2s_\mu^2)\big]\Big\} \\ &\geq 16\Big\{K_{xx}s_\nu^2c_\nu^2 + K_{yy}s_\mu^2c_\mu^2 - \Big(\hat{s}^2s_\nu^2c_\nu^2(1-2s_\nu^2)^2 + \hat{c}^2s_\mu^2c_\mu^2\big(1-2s_\mu^2\big)^2\Big)\Big\} \\ &\geq 16\Big\{K_{xx}s_\nu^2c_\nu^2\big(1-(1-2s_\nu^2)^2\big) + K_{yy}s_\mu^2c_\mu^2\Big(1-\big(1-2s_\mu^2\big)^2\Big)\Big\} \\ &= 16\big\{4K_{xx}s_\nu^4c_\nu^4 + 4K_{yy}s_\mu^4c_\mu^4\big\} \\ &\geq 16\big(K_{xx}s_\nu^4 + K_{yy}s_\mu^4\big). \end{aligned} \tag{3.20}$$

Combination of (3.19), (3.20) yields that for $z<0$ we have

$$d_1^{\nu\mu}/d_H^{\nu\mu} \leq \tfrac{1}{2}. \tag{3.21}$$

On the other hand we have, using $-z \leq K_{xx}s_\nu^2c_\mu^2 + K_{yy}s_\mu^2c_\nu^2$:

$$\begin{aligned} d_H^{\nu\mu} &= 16\Big\{K_{xx}s_\nu^2 + K_{yy}s_\mu^2 + z\big(1-2\big(s_\nu^2c_\mu^2 + s_\mu^2c_\nu^2\big)\big) - \big(K_{xx}s_\nu^4 + K_{yy}s_\mu^4\big)\Big\} \\ &= 4d_1^{\nu\mu} - 32z\big(s_\nu^2c_\mu^2 + s_\mu^2c_\nu^2\big) - 16\big(K_{xx}s_\nu^4 + K_{yy}s_\mu^4\big) \\ &\leq 4d_1^{\nu\mu} + 32\big(K_{xx}s_\nu^2c_\mu^2 + K_{yy}s_\mu^2c_\nu^2\big)\big(s_\nu^2c_\mu^2 + s_\mu^2c_\nu^2\big) - 16\big(K_{xx}s_\nu^4 + K_{yy}s_\mu^4\big) \\ &= 4d_1^{\nu\mu} + 16K_{xx}s_\nu^4\big(2c_\mu^4 - 1\big) + 16K_{yy}s_\mu^4\big(2c_\nu^4 - 1\big) + 32\big(K_{xx}+K_{yy}\big)s_\nu^2c_\nu^2s_\mu^2c_\mu^2 \\ &\leq 4\big\{d_1^{\nu\mu} + 4K_{xx}s_\nu^4 + 4K_{yy}s_\mu^4 + 8\big(K_{xx}+K_{yy}\big)s_\nu^2s_\mu^2\big\}. \end{aligned} \tag{3.22}$$

Now note that

$$|z| \leq 2|\hat{s}s_\nu c_\nu|\,|\hat{c}s_\mu c_\mu| \leq K_{xx}s_\nu^2c_\nu^2 + K_{yy}s_\mu^2c_\mu^2,$$

and also

$$|z| \leq 2|\hat{s}s_\nu c_\mu|\,|\hat{c}s_\mu c_\nu| \leq K_{xx}s_\nu^2c_\mu^2 + K_{yy}s_\mu^2c_\nu^2.$$

This yields

$$d_1^{\nu\mu} \geq 4\Big\{K_{xx}s_\nu^2\big(1-c_\nu^2\big) + K_{yy}s_\mu^2\big(1-c_\mu^2\big)\Big\} = 4\big(K_{xx}s_\nu^4 + K_{yy}s_\mu^4\big), \tag{3.23}$$

and also

$$d_1^{\nu\mu} \geq 4\Big\{K_{xx}s_\nu^2\big(1-c_\mu^2\big) + K_{yy}s_\mu^2\big(1-c_\nu^2\big)\Big\} = 4\big(K_{xx}+K_{yy}\big)s_\nu^2s_\mu^2. \tag{3.24}$$

Using (3.23) and (3.24) in (3.22) yields

$$d_H^{\nu\mu} \leq 4\big\{d_1^{\nu\mu} + d_1^{\nu\mu} + 2d_1^{\nu\mu}\big\} = 16d_1^{\nu\mu}.$$

In combination with (3.21) we obtain that for $z<0$ we have $d_1^{\nu\mu}/d_H^{\nu\mu} \in [\frac{1}{16}, \frac{1}{2}]$. Together with the result in (3.18) this yields a proof of the lemma. □

Theorem 3.4. *For all* $(\nu,\mu) \neq (0,0)$ *with* $-N_H+1 \leq \nu, \mu \leq N_H$ *the following holds, with* $\mathscr{K}_{\nu\mu}$, $d_H^{\nu\mu}$ *as in* (3.8), (3.12):

$$\mathscr{K}_{\nu\mu}/d_H^{\nu\mu} \in \big[\tfrac{1}{16}\big(1-\tfrac{1}{2}\sqrt{2}\big), 4\big]. \tag{3.25}$$

Proof: Direct consequence of the results in Lemma 3.2, 3.3. □

The result of Theorem 3.4 shows that for the rotated anisotropic diffusion equation A_H is a preconditioner of $\mathscr{S}_H$ which is robust with respect to variation in the problem parameters h, ϵ, β. If the bounds in (3.25) would be sharp then the condition number can be quite large and we would need a significant acceleration to obtain a reasonable convergence rate. A simple W-cycle on coarse grids would not be sufficient then.

To obtain an indication of the dependence of $\text{cond}(A_H^{-1}\mathscr{S}_H)$ on the parameters ϵ and β we computed the following quantities.

We take $h = 1/64$ and different values for β and ϵ: $\beta = \beta_k = k\pi/64$, $k = 0, 1, 2, \ldots, 64$ and $\epsilon = \epsilon_l = 10^{(1-l)/4}$, $l = 1, 2, \ldots, 25$. We introduce:

$$M(k,l) = \max\{\mathscr{K}_{\nu\mu}/d_H^{\nu\mu} | h = 1/64,\ \beta = \beta_k,\ \epsilon = \epsilon_l,$$
$$-31 \le \nu, \mu \le 32,\ (\nu,\mu) \ne (0,0)\},$$
$$m(k,l) = \min\{\mathscr{K}_{\nu\mu}/d_H^{\nu\mu} | h = 1/64,\ \beta_k,\ \epsilon = \epsilon_l,$$
$$-31 \le \nu, \mu \le 32,\ (\nu,\mu) \ne (0,0)\},$$
$$\text{Cond}(k,l) = \text{cond}\left(A_H^{-1}\mathscr{S}_H\right) = M(k,l)/m(k,l).$$

The computations resulted in the following. For all k, l we have $0.99 \le M(k,l) \le 1$, i.e. $\rho(A_H^{-1}\mathscr{S}_H) \approx 1$ independent of k, l. Furthermore $2 \le \text{Cond}(k,l) \le 4$ for all k, l. The graph of $\text{Cond}(k,l)$ is shown in Fig. 1. These results confirm the robustness proved in Theorem 3.4. Also these results indicate that $\rho(I - \omega A_H^{-1}\mathscr{S}_H) \le 0.6$ with $\omega = 8/5$ for all values of β and ϵ, i.e. a reasonable convergence rate with a fixed ω.

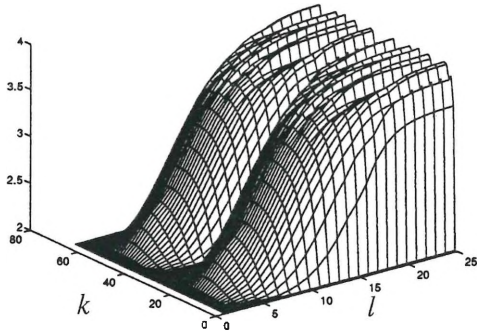

Figure 1. $\text{Cond}(k,l)$

Note that if the discrete problems are rescaled such that both the discrete operator on the fine grid and the discrete operator on the coarse grid are consistent with the differential operator, then $M(k,l)$ $(= \rho(A_H^{-1}\mathcal{S}_H))$ will be approximately 4 (instead of ≈ 1) and we should rescale the parameter ω, too.

It is easy to verify that for the problem (3.1) with Dirichlet boundary conditions and $\beta = 0$ we have $\mathrm{cond}(A_H^{-1}\mathcal{S}_H) \downarrow 1$ for $\epsilon \downarrow 0$. This is not observed in Fig. 1. It is noticed in [13] that for Dirichlet boundary conditions often a better agreement with Fourier analysis results is obtained if we only consider frequencies (ν, μ) with $\nu \neq 0$ and $\mu \neq 0$. So we also computed

$$\hat{M}(k,l) = \max\{\mathcal{K}_{\nu\mu}/d_H^{\nu\mu} | h = 1/64,\ \beta = \beta_k,\ \epsilon = \epsilon_l,$$
$$-31 \leq \nu, \mu \leq 32,\ \nu \neq 0 \text{ and } \mu \neq 0\},$$
$$\hat{m}(k,l) = \min\{\mathcal{K}_{\nu\mu}/d_H^{\nu\mu} | h = 1/64,\ \beta = \beta_k,\ \epsilon = \epsilon_l,$$
$$-31 \leq \nu, \mu \leq 32,\ \nu \neq 0 \text{ and } \mu \neq 0\},$$
$$\hat{\mathrm{Cond}}(k,l) = \hat{M}(k,l)/\hat{m}(k,l).$$

The results for $\hat{\mathrm{Cond}}(k,l)$ are shown in Fig. 2.

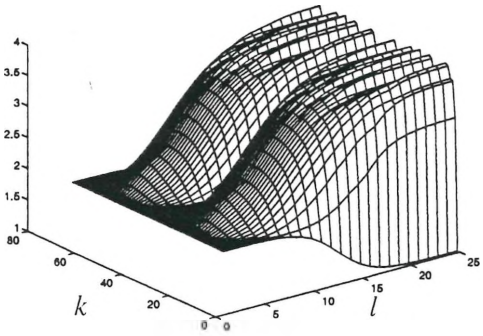

Figure 2. $\hat{\mathrm{Cond}}(k,l)$

Note that the results in Fig. 1 and Fig. 2 are very similar. In Figure 2 we indeed observe that $\hat{\mathrm{Cond}}(k,l)$ is close to 1 for $\beta = 0$, and $\epsilon \ll 1$.

Remark 3.5. The discretization for the rotated anisotropic diffusion equation in (3.4) is *not* a finite element discretization. In a finite element setting one can use the strenghtened CBS inequality to derive estimates for $\mathrm{cond}(A_H^{-1}\mathcal{S}_H)$. For symmetric positive definite problems there is a direct relation between the CBS constant γ^2 and $\mathrm{cond}(A_H^{-1}\mathcal{S}_H)$: $\mathrm{cond}(A_H^{-1}\mathcal{S}_H) = (1-\gamma^2)^{-1}$(cf. [1]). The theory as in [8] shows $\gamma^2 \leq \frac{3}{4}$ independent of the anisotropy, resulting in $\mathrm{cond}(A_H^{-1}\mathcal{S}_H) \leq 4$ (cf. [9] for a robust AMLI method based on this result). Note that this bound is the same as the bound observed in Figs. 1 and 2.

As discussed in Remark 3.5, for the rotated anisotropic diffusion equation, discretized using finite elements, one can prove a robustness result for $\mathrm{cond}(A_H^{-1}\mathcal{S}_H)$ using the theory of the strengthened CBS inequality. Clearly, for the convection-diffusion test problem (3.2) such an analysis is not available. However, a Fourier analysis is still possible. In the Fourier analysis some technical complications arise due to the fact that we now have complex eigenvalues.

For the test problem (3.2) we consider the following difference star:

$$[A_h] = \alpha_h \begin{bmatrix} & -1 & \\ -1 & 4 & -1 \\ & -1 & \end{bmatrix} + \hat{c}[-1 \quad 1 \quad 0] + \hat{s} \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \quad \alpha_h = \epsilon/h. \tag{3.26}$$

The corresponding eigenvalues are (s_k, c_k as in (3.7e)):

$$d_1^{\nu\mu} = 4\alpha_h\left(s_\nu^2 + s_\mu^2\right) + 2\hat{c}\left(s_\nu^2 + is_\nu c_\nu\right) + 2\hat{s}\left(s_\mu^2 + is_\mu c_\mu\right) \tag{3.27a}$$

$$d_2^{\nu\mu} = 4\alpha_h\left(c_\nu^2 + s_\mu^2\right) + 2\hat{c}\left(c_\nu^2 - is_\nu c_\nu\right) + 2\hat{s}\left(s_\mu^2 + is_\mu c_\mu\right) \tag{3.27b}$$

$$d_3^{\nu\mu} = 4\alpha_h\left(s_\nu^2 + c_\mu^2\right) + 2\hat{c}\left(s_\nu^2 + is_\nu c_\nu\right) + 2\hat{s}\left(c_\mu^2 - is_\mu c_\mu\right) \tag{3.27c}$$

$$d_4^{\nu\mu} = 4\alpha_h\left(c_\nu^2 + c_\mu^2\right) + 2\hat{c}\left(c_\nu^2 - is_\nu c_\nu\right) + 2\hat{s}\left(c_\mu^2 - is_\mu c_\mu\right). \tag{3.27d}$$

The coarse grid discrete analogon A_H of A_h in (3.26) has eigenvalues

$$\begin{aligned} d_H^{\nu\mu} = {} & 8\alpha_h\left(s_\nu^2 c_\nu^2 + s_\mu^2 c_\mu^2\right) + 4\hat{c}\left(2s_\nu^2 c_\nu^2 + is_\nu c_\nu\left(c_\nu^2 - s_\nu^2\right)\right) \\ & + 4\hat{s}\left(2s_\mu^2 c_\mu^2 + is_\mu c_\mu\left(c_\mu^2 - s_\mu^2\right)\right). \end{aligned} \tag{3.28}$$

In a recent paper [10] we applied a two-grid Fourier analysis to the convection-diffusion problem in (3.26), but then with a Schur complement approximation resulting from a Galerkin approach as in (2.6). In [10] it is shown that with a suitable fixed ω one can obtain $\rho(I - \omega(\hat{r}A_h\hat{p})^{-1}\mathcal{S}_H) \le C \le 1$ with C independent of h, ϵ, β. Using tools as in [10] one can prove a similar robustness result if we use the coarse grid discretization (cf. (3.28)) as an approximation of the Schur complement. However, a proof would require many technical manipulations (as in [10]) and therefore we will not present this here. We do present numerical results comparable to the results for the first test problem shown in Fig. 1.

We take $h = 1/64$, $\beta = \beta_k = k\pi/64$, $k = 0, 1, \ldots, 32$ and $\alpha_h = \alpha_{h,l} = 10^{(5-l)/4}$, $l = 1, 2, \ldots, 25$ (i.e. $\beta \in [0, \frac{1}{2}\pi]$, $\alpha_h \in [10^{-5}, 10]$). We introduce

$$\begin{aligned} \sigma_{k,l}\left(A_H^{-1}\mathcal{S}_H\right) = \{ & \mathcal{H}_{\nu\mu}/d_H^{\nu\mu} \mid h = 1/64,\ \beta = \beta_k,\ \alpha_h = \alpha_{h,l}, \\ & -31 \le \nu, \mu \le 32,\ (\nu, \mu) \ne (0,0)\}, \end{aligned}$$

with $\mathcal{H}_{\nu\mu}$ the harmonic average of the eigenvalues in (3.27) and $d_H^{\nu\mu}$ as in (3.28). This results in a spectrum that is contained in an ellipsoid in the complex right

half plane (cf. [10]). To measure the quality of A_H as a preconditioner for $\mathcal{S}_H$ we define

$$\text{Radius}(k,l) := \max\{|1-\omega_{k,l}\lambda| \mid \lambda \in \sigma_{k,l}(A_H^{-1}\mathcal{S}_H)\}, \tag{3.29}$$

where we use a shift parameter $\omega_{k,l}$ given by

$$\omega_{k,l} = 2\Big(\min_{\lambda} \text{Re}(\lambda) + \max_{\lambda} \text{Re}(\lambda)\Big)^{-1}, \tag{3.30}$$

in which the minimum and maximum is taken over $\lambda \in \sigma_{k,l}(A_H^{-1}\mathcal{S}_H)$. The graph of Radius$(k,l)$ is shown in Fig. 3. For the shift parameter $\omega_{k,l}$ we have $0.66 \le \omega_{k,l} \le 0.76$ for all k,l.

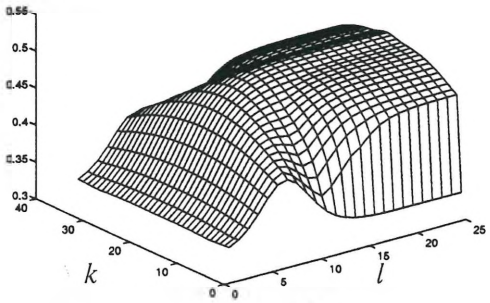

Figure 3. Radius(k,l): $\tilde{\mathcal{S}}_H = A_H/\omega$

Note that here we have a scaling parameter $\omega \approx 0.7$, which is about a factor 2 smaller than the scaling parameter for the rotated anisotropic diffusion equation. This is related to the different scaling of the problems in (3.4) and in (3.26).

We now compare the Schur complement preconditioner $\tilde{\mathcal{S}}_H = A_H/\omega$ from above with the Schur complement preconditioner $\hat{\mathcal{S}}_H = \hat{r}A_h\hat{p}/\omega$ as proposed in [10]. Based on the identity

$$\mathcal{S}_H = [0\ I]A_h\begin{bmatrix} -A_{11}^{-1}A_{12} \\ I \end{bmatrix}$$

we use $\hat{r} = [0\ \text{I}]$ and for $\hat{p}$ a matrix dependent prolongation explained in [10]. We take $h = 1/64$ and use parameter values $\beta = \beta_k$, $\alpha_h = \alpha_{h,l}$ as above. For given β_k and $\alpha_{h,l}$ the spectrum of $(\hat{r}A_h\hat{p})^{-1}\mathcal{S}_H$ is denoted by $\sigma_{k,l}((\hat{r}A_h\hat{p})^{-1}\mathcal{S}_H)$. We computed Radius$(k,l)$ as in (3.29) but now with $\sigma_{k,l}((\hat{r}A_h\hat{p})^{-1}\mathcal{S}_H)$ instead of $\sigma_{k,l}(A_H^{-1}\mathcal{S}_H)$. The results are shown in Fig. 4. In this case for the shift parameter $\omega_{k,l}$ (cf. (3.30)) we have $0.7 \le \omega_{k,l} \le 0.9$ for all k,l. We conclude that both in Fig. 3 and Fig. 4 we observe robustness with respect to variation in α_h and β. Finally, note that with $\omega \approx 0.7$ we have $\rho(I - \omega A_H^{-1}\mathcal{S}_H) \le 0.5$ and that the preconditioner based on the Galerkin property seems to be better than the standard coarse grid discretization preconditioner A_H.

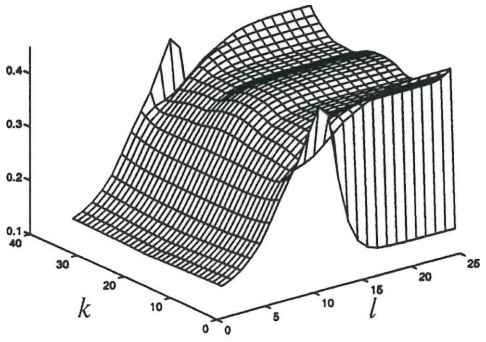

Figure 4. Radius(k,l): $\tilde{\mathcal{S}}_H = \hat{r}A_H\hat{p}/\omega$

4. Numerical Experiments

In this section we apply a multigrid W-cycle based on (2.9) to the following test problems (cf. (3.1), (3.2)):

$$\begin{cases} -(\epsilon\hat{c}^2+\hat{s}^2)u_{xx} - 2(\epsilon-1)\hat{c}\hat{s}u_{xy} - (\epsilon\hat{s}^2+\hat{c}^2)u_{yy} = f \text{ in } \Omega=(0,1)\times(0,1) \\ \qquad\qquad u = g \text{ on } \partial\Omega \end{cases} \tag{4.1}$$

$$\begin{cases} -\epsilon\Delta u + a(x,y)u_x + b(x,y)u_y = f \text{ in } \Omega=(0,1)\times(0,1) \\ \qquad\qquad u = g \text{ on } \partial\Omega, \end{cases} \tag{4.2}$$

with $\hat{c} = \cos(\beta)$, $\hat{s} = \sin(\beta)$, $\beta \in [0, \frac{1}{2}\pi]$, $\epsilon \in (0,1]$.

We use standard finite difference discretization on a square mesh Ω_h as in Section 3, resulting in a discrete problem $A_h x_h = b_h$. The scaling of the discrete equations is as in (3.4), (3.26). The finest mesh always has mesh size $h = 1/128$ and the coarsest mesh size is $h = 1/4$. For the approximation of the Schur complement we use the approach of (2.6) i.e. $\tilde{\mathcal{S}}_H = A_H/\omega$.

We use the notation $\Omega_h^c = \Omega_h \setminus \Omega_H$ (i.e. "new" nodes).

We now discuss the approximation used in the block Jacobi method with iteration matrix

$$J_h = I - \begin{bmatrix} A_{11}^{-1} & \emptyset \\ \emptyset & \emptyset \end{bmatrix} A_h.$$

In general the matrix A_{11} has a condition number $O(1)$ and then, in principle, any basic iterative method for solving $A_{11}y = z$ can be used. However, if we have strong alignment then cond(A_{11}) deteriorates. Our main interest in this paper is on robustness, so we should use a robust solver for the A_{11} systems. Probably the ILU method will yield a good compromise between robustness and

efficiency. In the method we implemented some efficiency has been sacrificed and we used a simple line Jacobi method. One iteration of this method consists of a sweep over the "odd" horizontal lines followed by a sweep over the "odd" vertical lines (these lines together form the pattern of $\Omega_h \setminus \Omega_H$). The result of μ iterations of such a line Jacobi method with starting vector 0 applied to $A_{11} y = z$ is denoted by $\mathcal{J}^{\mu}(A_{11}; 0; z)$.

The two-grid method based on (2.7) is as follows:

1. a) $d_{|\Omega_h^c} := (A_h x_h - b_h)_{|\Omega_h^c}$: compute defect on Ω_h^c.
   b) $\hat{x}_h := \mathcal{J}^{\mu}(A_{11}; 0; d_{|\Omega_h^c})$: line Jacobi for solving A_{11} system.
   c) $x_{h|\Omega_{h^c}} := (x_{h|\Omega_{h^c}} - \hat{x}_h)$: add correction on Ω_h^c.
2. a) $d_{|\Omega_H} := (A_h x_h - b_h)_{|\Omega_H}$: compute defect on Ω_H.
   b) $A_H v_H = d_{|\Omega_H}$: solve coarse grid problem.
   c) $x_{h|\Omega_H} := x_{h|\Omega_H} - \omega v_H$: add correction on Ω_H.
3. Repeat 1a, b, c.

This algorithm has the same structure as a standard two-grid algorithm. As in the standard approach, we use two recursive calls in 2b) to obtain a multigrid W-cycle, with iteration matrix as in (2.9).

In the experiments below we always take the data such that the exact discrete solution is equal to zero and we take an arbitrary starting vector. As a measure for the error reduction we computed $r := (\|e_{20}\|_2 / \|e_0\|_2)^{1/20}$, with e_k the error in the k-th iteration. For μ, i.e. the number of line Jacobi iterations, we take $\mu = 3$. Experiments have shown that this yields sufficiently accurate approximations when solving the A_{11} systems; often even $\mu = 2$ is sufficient. Based on the Fourier analysis we take $\omega = 1.4$ in Experiment 1 (rotated anisotropic diffusion equation) and $\omega = 0.7$ in Experiments 2 and 3 (convection-diffusion equation). As noted already in Section 3, the value of ω is related to the scaling of the discrete problems. In the two discrete test problems the scaling differs by a factor h (cf. (3.4), (3.26)). Using the same scaling in the discretization of the two test problems would result in *one* fixed ω for all three experiments below.

Experiment 1 (rotated anisotropic diffusion equation). We apply the multigrid W-cycle to the discrete version of (4.1). In Table 1 the resulting r are given for several values of β and ϵ.

Table 1

ϵ \ β	0	$\pi/10$	$2\pi/10$	$3\pi/10$	$4\pi/10$	$5\pi/10$
10^0	0.30					
10^{-1}	0.31	0.29	0.30	0.30	0.30	0.31
10^{-2}	0.31	0.35	0.37	0.37	0.35	0.31
10^{-3}	0.31	0.44	0.45	0.45	0.44	0.31
10^{-4}	0.35	0.46	0.46	0.46	0.46	0.35

Experiment 2 (convection-diffusion). We apply the multigrid W-cycle to the discrete version of (4.2) with $a(x, y) = \cos(\beta)$, $b(x, y) = \sin(\beta)$. In Table 2 the resulting r are given for several values of β and ϵ.

Table 2

ϵ \ β	0	$\pi/10$	$2\pi/10$	$3\pi/10$	$4\pi/10$	$5\pi/10$
10^{-1}	0.29	0.29	0.29	0.29	0.29	0.29
10^{-2}	0.30	0.30	0.30	0.30	0.30	0.30
10^{-3}	0.30	0.39	0.43	0.43	0.39	0.30
10^{-4}	0.33	0.38	0.46	0.46	0.38	0.33
10^{-5}	0.37	0.38	0.46	0.46	0.38	0.37

Table 3

ϵ	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}
r	0.29	0.31	0.44	0.53	0.55

Experiment 3 (rotating flow). We define $\Omega_R := \{(x, y) | ((x - \frac{1}{3})^2 + (y - \frac{1}{3})^2) \leq \frac{1}{16}\}$, and

$$\begin{cases} a(x, y) = \sin\left(\pi\left(y - \frac{1}{3}\right)\right) \cos\left(\pi\left(x - \frac{1}{3}\right)\right) \text{ if } (x, y) \in \Omega_R, \text{ and zero otherwise;} \\ b(x, y) = -\cos\left(\pi\left(y - \frac{1}{3}\right)\right) \sin\left(\pi\left(x - \frac{1}{3}\right)\right) \text{ if } (x, y) \in \Omega_R, \text{ and zero otherwise.} \end{cases}$$

We apply the multigrid W-cycle to the discrete version of (4.2) with these functions a, b. The results are given in Table 3.

In these experiments we see a convergence behaviour of the multigrid W-cycle that is similar to the convergence behaviour of the two-grid method in Section 3. As in the two-grid method we have robustness w.r.t. variation in the problem parameters and contraction numbers between 0.3 and 0.6. Finally we note that in this paper the main topic is robustness and not efficiency. The method we used above can be improved w.r.t. efficiency, for example by using an ILU solver instead of the line Jacobi solver for the A_{11} systems or by using a Galerkin approach as in (2.6) for approximation of the Schur complement.

References

[1] Axelsson, O.: On multigrid methods of the two-level type. In: Multigrid methods (Hackbusch, W., Trottenberg, U., eds.), pp. 352–367. Berlin Heidelberg New York: Springer 1982.

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

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

[4] Bank, R. E., Dupont, T. F., Yserentant, H.: The hierarchical basis multigrid method. Numer. Math. *52*, 427–458 (1988).
[5] Hackbusch, W.: Multigrid methods and applications. Berlin Heidelberg New York: Springer 1985.
[6] Hackbusch, W.: The frequency decomposition multi-grid method. I. Application to anisotropic equations. Numer. Math. *56*, 229–245 (1989).
[7] Hackbusch, W.: The frequency decomposition multi-grid method. II. Convergence analysis based on the additive Schwarz method. Numer. Math. *63*, 433–453 (1992).
[8] Maitre, J. F., Musy, F.: The contraction number of a class of two-level methods; an exact evaluation for some finite element subspaces and model problems. In: Multigrid methods (Hackbusch, W., Trottenberg, U., eds.), pp. 201–216. Berlin Heidelberg New York: Springer 1982.
[9] Margenov, S. D., Vassilevski, P. S.: Algebraic multilevel preconditioning of anisotropic elliptic problems. SIAM J. Sci. Comput. *15*, 1026–1037 (1994).
[10] Reusken, A.: Fourier analysis of a robust multigrid method for convection-diffusion equations. Numer. Math. *71*, 365–397 (1995).
[11] Stevenson, R.: Robustness of the additive and multiplicative frequency decomposition multi-level method. Computing *54*, 331–346 (1995).
[12] Vassilevski, P. S., Etova, M. H.: Computation of constants in the strengthened Cauchy inequality. SIAM J. Sci. Statist. Comp. *13*, 655–665 (1992).
[13] Wesseling, P.: An introduction to multigrid methods. Chichester: Wiley 1992.
[14] Xu, J.: Iterative methods by space decomposition and subspace correction. SIAM Rev. *34*, 581–613 (1992).
[15] Yserentant, H.: Old and new convergence proofs for multigrid methods. Acta Numerica 285–326 (1993).

Arnold Reusken
Department of Mathematics and Computing Science
Eindhoven University of Technology, PO Box 513
5600 MB Eindhoven,
The Netherlands
e-mail: wsanar@win.tue.nl

Verleger: Springer-Verlag KG, Sachsenplatz 4–6, A-1201 Wien. – Herausgeber: Prof. Dr. Hans J. Stetter, Institut für Angewandte und Numerische Mathematik der Technischen Universität Wien, Wiedner Hauptstrasse 6–10, A-1040 Wien. – Redaktion: Wiedner Hauptstrasse 6–10, A-1040 Wien. – Satz und Umbruch: Compuscript Ltd., Shannon; Offsetdruck: Novographic, Ing. W. Schmid, Maurer-Lange-Gasse 64, A-1238 Wien. – Verlagsort: Wien. – Herstellungsort: Wien. – Printed in Austria.