CONVERGENCE ANALYSIS OF A MULTIGRID METHOD FOR CONVECTION–DIFFUSION EQUATIONS

ARNOLD REUSKEN*

Abstract. This paper is concerned with the convergence analysis of robust multigrid methods for convection-diffusion problems. We consider a finite difference discretization of a 2D model convection-diffusion problem with constant coefficients and Dirichlet boundary conditions. For the approximate solution of this discrete problem a multigrid method based on semicoarsening, matrix-dependent prolongation and restriction and line smoothers is applied. For a multigrid W-cycle we prove an upper bound for the contraction number in the euclidean norm which is smaller than one and independent of the mesh size and the diffusion/convection ratio. For the contraction number of a multigrid V-cycle a bound is proved which is uniform for a class of convection-dominated problems. The analysis is based on linear algebra arguments only.

Key words. multigrid, convection–diffusion, convergence analysis

AMS subject classifications. 65F10, 65F50, 65N22, 65N55

1. Introduction. Concerning the theoretical analysis of multigrid methods different fields of application have to be distinguished. For linear selfadjoint elliptic boundary value problems the convergence theory has reached a mature, if not its final state (cf. [5, 9, 28, 29]). In other areas the state of the art is (far) less advanced. For example, for convection-diffusion problems the development of a multigrid convergence analysis is still in its infancy. In this paper we present a convergence analysis of a multilevel method for a special class of 2D convection-diffusion problems.

An interesting class of problems for the analysis of multigrid convergence is given by

$$\begin{cases} -\varepsilon\Delta u + b\cdot\nabla u &= f \quad \text{in} \quad \Omega=(0,1)^2 \\ u &= g \quad \text{on} \quad \partial\Omega, \end{cases} \tag{1.1}$$

with $\varepsilon > 0$ and $b = (\cos\phi, \sin\phi)$, $\phi \in [0, 2\pi)$. The application of a discretization method (e.g., a finite difference method with upwinding or a streamline diffusion finite element method) results in a large sparse linear system which depends on a mesh size parameter h. Note that in this discrete problem we have three interesting parameters: h (mesh size), ε (convection-diffusion ration) and ϕ (flow direction). For the approximate solution of this type of problems *robust* multigrid methods have been developed which are efficient solvers for a large range of relevant values for the parameters h, ε, ϕ. To obtain good robustness properties the components in the multigrid method have to be chosen in a special way because in general the "standard" multigrid approach used for a diffusion problem does not yield satisfactory results when applied to a convection-dominated problem. To improve robustness several modifications have been proposed in the literature, such as "robust" smoothers (smoothers which try to follow the flow direction), matrix-dependent prolongations and restrictions and semicoarsening techniques. For an explanation of these methods we refer to [9, 27, 3, 7, 13, 14, 16, 17, 21, 30]. These modifications are based on heuristic arguments and empirical studies.

Related to the theoretical analysis of multigrid applied to convection-diffusion problems we note the following. In the literature one finds convergence analyses of multigrid methods for nonsymmetric elliptic boundary value problems which are based on

*Institut für Geometrie und Praktische Mathematik, RWTH Aachen, Templergraben 55, D-52056 Aachen, Germany.

perturbation arguments [6, 9, 15, 26]. If these analyses are applied to the problem in (1.1) the constants in the estimates depend on ε and the results are not satisfactory for the case $\varepsilon \ll 1$, i.e., for convection-dominated problems. In [1, 11, 23] multigrid convergence for a 1D convection-diffusion problem is analyzed. The results show robustness of two- and multigrid methods. These analyses, however, are restricted to the 1D case. In [18, 24] convection-diffusion equations as in (1.1) with *periodic* boundary conditions are considered. A Fourier analysis is applied to analyze the convergence of two- or multigrid methods. In [18] the problem (1.1) with periodic boundary conditions and $\phi = 0$ is studied. A V-cycle contraction result is proved which is uniform in ε and h provided $\frac{\varepsilon}{h} \leq c$ is satisfied with c a positive constant that does not depend on ε or h. In [24] a two-grid method for solving a finite difference discretization of the problem (1.1) with periodic boundary conditions is analyzed and it is proved that the two-grid contraction number is bounded by a constant smaller than one which does not depend on any of the parameters $\varepsilon,\ h,\ \phi$. In [2] the application of the hierarchical basis multigrid method to a finite element discretization of problems as in (1.1) is studied. The analysis there shows how the convergence rate depends on ε and on the flow direction, but the estimates are not uniform with respect to the mesh size parameter h. Recently, in [19] an analysis of the robustness of a multigrid method applied to a class of 2D convection-diffusion problems with Dirichlet boundary conditions has been presented. In our opinion, this analysis contains some loose ends and is not convincing.

As one of the simplest model cases for a 2D discrete convection-diffusion problem one can take the problem as in (1.1) with $\phi = 0$ and discretized by a stable finite difference method on a tensor product grid. In none of the analyses known from the literature the robustness of multigrid, with respect to variation in ε and h, applied to this problem has been proved. In this paper such robustness results will be presented. Note that for this discrete problem we have, for ε small, a "flow" in the x-direction from left to right aligned with the grid lines. We study a multigrid method in which the following components for smoothing and coarse grid correction are used. The coarse grids are obtained by semicoarsening in the x-direction and for the construction of the coarse grid operators we use the Galerkin approach. The prolongation and restriction are matrix-dependent and result from a tensor product construction applied to 1D matrix-dependent intergrid transfer operators known from the literature. For the smoother we take a y-line Jacobi or a left-to-right y-line Gauss-Seidel method. Note that the y-lines are in crosswind direction and thus the y-line Jacobi method is not a robust smoother. We use y-line smoothers to make the analysis work. The convergence analysis that will be presented is based on *linear algebra arguments only*. The main idea is as follows. The discrete operator (block tridiagonal matrix) has the tensor product form $A = I \otimes A_x + A_y \otimes I$, where A_y is a symmetric positive definite tridiagonal matrix (corresponds to $-\varepsilon \frac{\partial^2}{\partial y^2}$) and A_x is a tridiagonal M-matrix (corresponds to $-\varepsilon \frac{\partial^2}{\partial x^2} + \frac{\partial}{\partial x}$). Using an orthogonal eigenvector basis of A_y we can transform A to block *diagonal* form where the diagonal blocks are *tridiagonal* matrices. The diagonal blocks depend on the eigenvalues of A_y. These eigenvalues are treated as unknown positive parameters. This block diagonalization can be applied to all components in our two-grid method. It then turns out that each diagonal block of the transformed two-grid iteration matrix can be interpreted as the iteration matrix of a two-grid method applied to a discrete *one*-dimensional convection-diffusion-reaction problem. In this one-dimensional setting with tridiagonal matrices we can prove a smoothing property and approximation property which are robust with respect to variation in the relevant parameters (ε, h and the eigenvalues of A_y). This block diagonalization technique is also used in [4] for the analysis of robustness of multigrid applied to a symmetric anisotropic diffusion problem. Using this technique we obtain a bound for the two-grid contraction number which depends only on ν, the number of smoothing iterations. Moreover, this bound is smaller than one for ν sufficiently large. Using a standard technique as in [9] we obtain a similar bound for the multigrid W-cycle. For the class of convection-dominated problems with $\frac{\varepsilon}{h} \leq \frac{1}{2}$ we will prove a robustness result for the multigrid V-cycle with left-to-right y-line Gauss-Seidel smoothing.

The remainder of this paper is organized as follows. In Sect. 2 we introduce the class of discrete problems that will be considered and describe the two-grid method. It is shown that for this two-grid method the contraction number corresponding to a two-dimensional convection diffusion problem can be bounded by the maximum of the two-grid contraction numbers for a certain class of one-dimensional convection-diffusion-reaction problems. In Sect. 3 we consider this class of one-dimensional problems and prove a smoothing- and approximation property in which the constants are independent of all relevant parameters. As a consequence of this we obtain a robustness result for the two-grid method. In Sect. 4 the W-cycle convergence is analyzed along the lines as in [9]. Finally, in Sect. 5 a robustness result for the V-cycle multigrid method applied to convection-dominated problems is proved.

2. Two-grid method and preliminary results. We start with a description of the class of problems that will be considered. Let $\Omega_{k,m}$ be a two-dimensional tensor product grid, with an arbitrary but fixed mesh size $h_y = \frac{1}{m}$ in the y-direction and a variable mesh size $h_k = 2^{-k},\ k = 1, 2, \ldots$, in the x-direction. We use the notation

$$\begin{aligned} & n_k := 2^k - 1, \\ & I_k \ :\ n_k \times n_k \text{ identity matrix },\\ & I_y \ :\ (m-1)\times(m-1) \text{ identity matrix }. \end{aligned}$$

We assume that A_y is a given *symmetric positive definite* matrix of order $m-1$. Let T_γ be the tridiagonal matrix

$$T_\gamma = \begin{bmatrix} 1 & -\gamma & & \\ -(1-\gamma) & \ddots & \ddots & \\ & \ddots & \ddots & -\gamma \\ & & -(1-\gamma) & 1 \end{bmatrix}, \quad \gamma \in (0, \frac{1}{2}) \ . \tag{2.1}$$

The dimension of T_γ will be clear from the context. Now a class of tensor product matrices is introduced:

$$\hat{L}_k(\gamma) = \{\ \alpha I_y \otimes T_\gamma + \beta A_y \otimes I_k \mid \alpha > 0,\ \beta \geq 0\ \}\ . \tag{2.2}$$

In (2.2) we use the tensor product notation: For matrices $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{m\times m}$, $A \otimes B \in \mathbb{R}^{nm\times nm}$ is a block-matrix with (i,j)-block given by $a_{ij}B$, $1 \leq i,j \leq n$.

REMARK 2.1. We show that finite difference methods applied to certain 2D convection–diffusion equations yield matrices that are in $\hat{L}_k(\gamma)$. Consider the boundary value problem

$$\begin{cases} -\varepsilon \frac{\partial^2 u}{\partial x^2} - \frac{\partial}{\partial y}(b(y)\frac{\partial u}{\partial y}) + \frac{\partial u}{\partial x} &= f \quad \text{in } \ \Omega = (0,1)^2 \\ \qquad\qquad\qquad\qquad u &= 0 \quad \text{on } \ \partial\Omega, \end{cases} \tag{2.3}$$

with $\varepsilon > 0$, $b \in C^1[0,1]$ and $b(y) > 0$ for all $y \in [0,1]$. An interesting special case of (2.3) is $b(y) = \varepsilon$ for all y. For the second order derivatives central differences on the grid $\Omega_{k,m}$ are used. The differential operator in y-direction then results in the symmetric positive definite matrix

$$A_y = \frac{1}{h_y^2}\begin{bmatrix} b_1+b_2 & -b_2 & & & \\ -b_2 & b_2+b_3 & -b_3 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & -b_{m-1} \\ & & & -b_{m-1} & b_{m-1}+b_m \end{bmatrix}, \quad b_j := b((j-\frac{1}{2})h_y). \tag{2.4}$$

If for the differential operator in x-direction, $u \to -\varepsilon\frac{\partial^2 u}{\partial x^2} + \frac{\partial u}{\partial x}$, we use *full upwind* differences for the term $\frac{\partial u}{\partial x}$ we obtain the matrix

$$\alpha T_\gamma \quad \text{with } \alpha = \frac{2\varepsilon}{h_k^2} + \frac{1}{h_k}, \quad \gamma = \frac{\frac{\varepsilon}{h_k}}{2\frac{\varepsilon}{h_k}+1} \in (0, \frac{1}{2}) \ . \tag{2.5}$$

If we use *central* differences for the term $\frac{\partial u}{\partial x}$ we obtain the matrix

$$\alpha T_\gamma \quad \text{with } \alpha = \frac{2\varepsilon}{h_k^2}, \quad \gamma = \frac{\frac{\varepsilon}{h_k} - \frac{1}{2}}{2\frac{\varepsilon}{h_k}} \ , \tag{2.6}$$

with $\gamma \in (0, \frac{1}{2})$ if the usual stability condition $h_k < 2\varepsilon$ is fulfilled. Using a left-to-right, bottom-to-top ordering of the grid points both (2.4), (2.5) and (2.4), (2.6) result in a discrete problem with a matrix of the form

$$\alpha I_y \otimes T_\gamma + A_y \otimes I_k \in \hat{L}_k(\gamma). \tag{2.7}$$

We note that many other known finite difference approximations of the differential operator $-\varepsilon\frac{\partial^2}{\partial x^2} + \frac{\partial}{\partial x}$, like e.g. the Il'in scheme, combined with (2.4) yield a matrix which is an element of $\hat{L}_k(\gamma)$ for a suitable $\gamma \in (0, \frac{1}{2})$. □

We now introduce a two–grid method for solving a system with matrix $\hat{A} \in \hat{L}_k(\gamma)$. This method uses semicoarsening in the x-direction, a matrix-dependent prolongation and restriction and a Galerkin approach. For the smoother a (damped) block-Jacobi or block-Gauss-Seidel iteration is used. For the description of these two-grid components we need the following matrices

$$r_\gamma : \ \mathbb{R}^{n_k} \to \mathbb{R}^{n_{k-1}}, \quad r_\gamma = \frac{1}{2} r_{\text{inj}}(2I_k - T_\gamma), \tag{2.8a}$$

$$p_\gamma : \ \mathbb{R}^{n_{k-1}} \to \mathbb{R}^{n_k}, \quad p_\gamma = (2I_k - T_\gamma) r_{\text{inj}}^T, \tag{2.8b}$$

where $r_{\text{inj}}(x_1, x_2, \ldots, x_{n_k})^T = (x_2, x_4, \ldots, x_{n_k-1})^T$ is the trivial injection. Note that r_γ and p_γ are standard matrix-dependent transfer operators corresponding to T_γ, with stencils $[r_\gamma] = \frac{1}{2}[1-\gamma \quad 1 \quad \gamma]$ and $[p_\gamma] = [\gamma \quad 1 \quad 1-\gamma]$ (cf. [9, 23]). Let

$$\hat{A} = \alpha I_y \otimes T_\gamma + \beta A_y \otimes I_k \in \hat{L}_k(\gamma) \tag{2.9}$$

be given. We define the matrix-dependent restriction and prolongation

$$\hat{r}_\gamma = I_y \otimes r_\gamma, \quad \hat{p}_\gamma = I_y \otimes p_\gamma \ , \tag{2.10}$$

and the coarse grid matrix $\hat{A}_H$

$$\hat{A}_H = \alpha \hat{r}_\gamma (I_y \otimes T_\gamma)\hat{p}_\gamma + \beta A_y \otimes I_{k-1} \ . \tag{2.11}$$

Using $\hat{r}_\gamma(I_y \otimes T_\gamma)\hat{p}_\gamma = I_y \otimes r_\gamma T_\gamma p_\gamma$ and $r_\gamma T_\gamma p_\gamma = \frac{1}{2}\mathrm{tridiag}(-(1-\gamma)^2, 1-2\gamma(1-\gamma), -\gamma^2)$ it follows that

$$\hat{A}_H \in \hat{L}_{k-1}(\tilde{\gamma}) \quad \text{with} \quad \tilde{\gamma} = \frac{\gamma^2}{1-2\gamma(1-\gamma)} \in (0,\gamma). \tag{2.12}$$

Note that in a coarse grid correction based on (2.10), (2.11) we use semicoarsening in the x-direction. Furthermore, the interpolation and restriction operators have 3-point stencils . The smoothers we use have an iteration matrix of the form

$$\hat{S}_\theta = I_y \otimes I_k - \theta \hat{W}^{-1}\hat{A}, \qquad \theta \in (0,1) \ .$$

We will consider a y-line-Jacobi method, i.e.

$$\hat{W} = \hat{W}_{\mathrm{J}} = \alpha I_y \otimes I_k + \beta A_y \otimes I_k \tag{2.13}$$

and a left-to-right y-line-Gauss-Seidel method, i.e.,

$$\hat{W} = \hat{W}_{\mathrm{GS}} = \alpha I_y \otimes L_\gamma + \beta A_y \otimes I_k \ , \tag{2.14}$$

where

$$L_\gamma = \begin{bmatrix} 1 & & & \\ -(1-\gamma) & \ddots & \emptyset & \\ & \ddots & \ddots & \\ & & -(1-\gamma) & 1 \end{bmatrix} \tag{2.15}$$

is the lower triangular part of T_γ. In this paper we consider the following two smoothers: A damped Jacobi method with iteration matrix

$$\hat{S}^{(\nu)} = \hat{S}_{\mathrm{J}}^{(\nu)} = (I - \frac{1}{2}\hat{W}_{\mathrm{J}}^{-1}\hat{A})^\nu \tag{2.16}$$

and a Gauss-Seidel method with iteration matrix

$$\hat{S}^{(\nu)} = \hat{S}_{\mathrm{GS}}^{(\nu)} = (I - \hat{W}_{\mathrm{GS}}^{-1}\hat{A})(I - \frac{1}{2}\hat{W}_{\mathrm{GS}}^{-1}\hat{A})^{\nu-1} \ . \tag{2.17}$$

Note that in (2.17) we use a combination of a Gauss-Seidel method with and without damping. The reason for this originates from the analysis of the smoothing property below (Theorem 3.8).

A two-grid method with the components described above has a corresponding iteration matrix

$$\hat{M}_k = \hat{M}_k(\gamma) = (I - \hat{p}_\gamma \hat{A}_H^{-1}\hat{r}_\gamma \hat{A})\hat{S}^{(\nu)} \ , \tag{2.18}$$

with $I = I_y \otimes I_k$. Note that $\hat{M}_k(\gamma)$ depends only on $\hat{A} \in \hat{L}_k(\gamma)$, the parameter ν and on the choice of the smoother (Jacobi (2.16) or Gauss-Seidel (2.17)). We call $\hat{M}_k(\gamma)$ the *two-grid iteration matrix induced by* $\hat{A}$. Because the dependence of $\hat{M}_k$ on γ plays an important role in the remainder we use the notation $\hat{M}_k(\gamma)$.

Let $Q_y \in \mathbb{R}^{(m-1)\times(m-1)}$ be an orthogonal matrix such that

$$Q^T A_y Q = \mathrm{diag}(\mu_i)_{1\le i\le m-1} \ .$$

Note that $\mu_i > 0$ for all i. We define the orthogonal matrix

$$\hat{Q}_k := Q \otimes I_k, \quad k = 1, 2, \dots \ . \tag{2.19}$$

Using $\hat{Q}_k$ we can transform all components in the iteration matrix $\hat{M}_k$ to block-diagonal form:

LEMMA 2.2. *For* $\hat{A}$, $\hat{p}_\gamma$, $\hat{r}_\gamma$, $\hat{W}_{\mathrm{J}}$, $\hat{W}_{\mathrm{GS}}$ *as in* (2.9), (2.10), (2.13), (2.14) *the following holds:*

$$\hat{Q}_k^T \hat{A} \hat{Q}_k = \mathrm{blockdiag}(\alpha T_\gamma + \beta\mu_i I_k)_{1\le i\le m-1}, \tag{2.20}$$

$$\hat{Q}_{k-1}^T \hat{r}_\gamma \hat{Q}_k = \mathrm{blockdiag}(r_\gamma)_{1\le i\le m-1} \quad \textit{with } r_\gamma \quad \textit{as in } (2.8\mathrm{a}), \tag{2.21}$$

$$\hat{Q}_k^T \hat{p}_\gamma \hat{Q}_{k-1} = \mathrm{blockdiag}(p_\gamma)_{1\le i\le m-1} \quad \textit{with } p_\gamma \quad \textit{as in } (2.8\mathrm{b}), \tag{2.22}$$

$$\hat{Q}_{k-1}^T \hat{A}_H \hat{Q}_{k-1} = \mathrm{blockdiag}(\alpha r_\gamma T_\gamma p_\gamma + \beta\mu_i I_{k-1})_{1\le i\le m-1}, \tag{2.23}$$

$$\hat{Q}_k^T \hat{W}_{\mathrm{J}} \hat{Q}_k = \mathrm{blockdiag}((\alpha + \beta\mu_i) I_k)_{1\le i\le m-1}, \tag{2.24}$$

$$\hat{Q}_k^T \hat{W}_{\mathrm{GS}} \hat{Q}_k = \mathrm{blockdiag}(\alpha L_\gamma + \beta\mu_i I_k)_{1\le i\le m-1}. \tag{2.25}$$

Proof. We show how the results in (2.20) and (2.21) can be proved. The proofs for the other results are similar. Note that

$$\begin{aligned}
\hat{Q}_k^T \hat{A} \hat{Q}_k &= (Q^T \otimes I_k)(\alpha I_y \otimes T_\gamma + \beta A_y \otimes I_k)(Q \otimes I_k) \\
&= \alpha I_y \otimes T_\gamma + \beta Q^T A_y Q \otimes I_k \\
&= \alpha I_y \otimes T_\gamma + \beta \mathrm{diag}(\mu_i)_{1\le i\le m-1} \otimes I_k \\
&= \alpha \mathrm{blockdiag}(T_\gamma)_{1\le i\le m-1} + \beta \mathrm{blockdiag}(\mu_i I_k)_{1\le i\le m-1} \\
&= \mathrm{blockdiag}(\alpha T_\gamma + \beta\mu_i I_k)_{1\le i\le m-1}
\end{aligned}$$

and thus (2.20) holds. The result in (2.21) follows from

$$\begin{aligned}
\hat{Q}_{k-1}^T \hat{r}_\gamma \hat{Q}_k &= (Q^T \otimes I_{k-1})(I_y \otimes r_\gamma)(Q \otimes I_k) \\
&= I_y \otimes r_\gamma = \mathrm{blockdiag}(r_\gamma)_{1\le i\le m-1}. \qquad \square
\end{aligned}$$

Note that the diagonal blocks in (2.20), (2.24), (2.25) are (tri)diagonal matrices of dimension $n_k \times n_k$. Motivated by the results in Lemma 2.2 we introduce the following class of $n_k \times n_k$ tridiagonal matrices:

$$L_k(\gamma) = \{\ \alpha T_\gamma + \delta I_k \mid \alpha > 0,\ \delta \ge 0\ \}, \qquad \gamma \in (0, \frac{1}{2}).$$

Let

$$A = \alpha T_\gamma + \delta I_k \in L_k(\gamma)$$

be given. We define the coarse grid matrix

$$A_H = \alpha r_\gamma T_\gamma p_\gamma + \delta I_{k-1} \in L_{k-1}(\tilde{\gamma}) \quad \text{with} \quad \tilde{\gamma} = \frac{\gamma^2}{1-2\gamma(1-\gamma)} \tag{2.26}$$

and a smoother with iteration matrix

$$S_\theta = I_k - \theta W^{-1} A, \qquad \theta \in (0,1)\ .$$

We consider a Jacobi smoother with iteration matrix

$$S^{(\nu)} = S_{\mathrm{J}}^{(\nu)} = (I - \frac{1}{2} W_{\mathrm{J}}^{-1} A)^\nu, \quad W_{\mathrm{J}} = \mathrm{diag}(A) = (\alpha+\delta) I_k \tag{2.27}$$

and a Gauss-Seidel smoother with iteration matrix

$$S^{(\nu)} = S_{\mathrm{GS}}^{(\nu)} = (I - W_{\mathrm{GS}}^{-1} A)(I - \frac{1}{2} W_{\mathrm{GS}}^{-1} A)^{\nu-1}, \quad W_{\mathrm{GS}} = \alpha L_\gamma + \delta I_k\ . \tag{2.28}$$

These components yield a corresponding two-grid iteration matrix

$$M_k = M_k(\gamma) = (I_k - p_\gamma A_H^{-1} r_\gamma A) S^{(\nu)}\ , \tag{2.29}$$

which is called the *two-grid iteration matrix induced by* A.

Note that for the (large) tensor product matrices we use a notation with "$\hat{}$" (e.g. $\hat{A},\ \hat{W}$), whereas for the (small) tridiagonal matrices we use a notation without "$\hat{}$" (e.g. $A,\ W$). Using these definitions and Lemma 2.2 we obtain the following

THEOREM 2.3. *Take $\hat{A} \in \hat{L}_k(\gamma)$ as in* (2.9) *and let $\hat{M}_k(\gamma)$ be the two-grid iteration matrix induced by $\hat{A}$ (cf.* (2.18)*). Then*

$$\hat{Q}_k^T \hat{M}_k(\gamma) \hat{Q}_k = \mathrm{blockdiag}(M_k^{(i)}(\gamma))_{1 \le i \le m-1}\ ,$$

where $M_k^{(i)}(\gamma)$ is the two-grid iteration matrix as in (2.29) *induced by the matrix $A = \alpha T_\gamma + \beta \mu_i I_k \in L_k(\gamma)$.*

Proof. Note that

$$\hat{Q}_k^T \hat{M}_k(\gamma) \hat{Q}_k = (I - (\hat{Q}_k^T \hat{p}_\gamma \hat{Q}_{k-1})(\hat{Q}_{k-1}^T \hat{A}_H \hat{Q}_{k-1})^{-1} (\hat{Q}_{k-1}^T \hat{r}_\gamma \hat{Q}_k)(\hat{Q}_k^T \hat{A} \hat{Q}_k))(\hat{Q}_k^T \hat{S}^{(\nu)} \hat{Q}_k),$$
$$\text{and} \quad \hat{Q}_k^T \hat{S}_\theta \hat{Q}_k = I - \theta (\hat{Q}_k^T \hat{W} \hat{Q}_k)^{-1} (\hat{Q}_k^T \hat{A} \hat{Q}_k).$$

Now use the results of Lemma 2.2. □

In this theorem and in Theorem 2.4 below we assume that the smoother $S^{(\nu)}$ used in $M_k(\gamma)$ corresponds to the smoother $\hat{S}^{(\nu)}$ used in $\hat{M}_k(\gamma)$, i.e., $\hat{S}^{(\nu)} = \hat{S}_{\mathrm{J}}^{(\nu)}$ corresponds to $S^{(\nu)} = S_{\mathrm{J}}^{(\nu)}$ and $\hat{S}^{(\nu)} = \hat{S}_{\mathrm{GS}}^{(\nu)}$ corresponds to $S^{(\nu)} = S_{\mathrm{GS}}^{(\nu)}$.

The result in the following theorem yields a possibility to analyze the convergence of the two-grid method induced by a given matrix $\hat{A} \in \hat{L}_k(\gamma)$ by considering two-grid convergence for a class of tridiagonal matrices.

THEOREM 2.4. *Let $\hat{A} \in \hat{L}_k(\gamma)$ be given and let $\hat{M}_k(\gamma)$ be the two-grid iteration matrix induced by $\hat{A}$. Then the following holds:*

$$\|\hat{M}_k(\gamma)\|_2 \le \max_{A \in L_k(\gamma)} \|M_k(\gamma)\|_2\ , \tag{2.30}$$

where $M_k(\gamma)$ is the two-grid iteration matrix (2.29) *induced by A.*

Proof. From Theorem 2.3 we obtain

$$\|\hat{M}_k(\gamma)\|_2 = \max_{1 \le i \le m-1} \|M_k^{(i)}(\gamma)\|_2 \ ,$$

with $M_k^{(i)}(\gamma)$ the two-grid iteration matrix induced by $A = \alpha T_\gamma + \beta \mu_i I_k \in L_k(\gamma)$. Hence the inequality in (2.30) holds because the maximum over $A \in L_k(\gamma)$ is taken. ∎

COROLLARY 2.5. The iteration matrix $M_k(\gamma)$ does not depend on the scaling of the matrix $A \in L_k(\gamma)$. Hence we obtain

$$\|\hat{M}_k(\gamma)\|_2 \le \max_{A \in L_k(\gamma)} \|M_k(\gamma)\|_2 = \max_{A \in L_k^*(\gamma)} \|M_k(\gamma)\|_2 \ ,$$

with

$$L_k^*(\gamma) = \{ \ T_\gamma + \delta I_k \mid \delta \ge 0 \ \} \ . \qquad \square$$

3. Convergence analysis of the two-grid method. In this section we will derive bounds for

$$\max_{A \in L_k^*(\gamma)} \|M_k(\gamma)\|_2 \ . \tag{3.1}$$

Note that $A \in L_k^*(\gamma)$ is a tridiagonal matrix, which can be interpreted as a finite difference discretization of a 1D convection-diffusion-reaction equation (cf. Remark 2.1). The analysis below will be based on an approximation property (given in Theorem 3.4) and a smoothing property (proved in Theorems 3.7, 3.8). These two properties combined yield a two-grid convergence result (Theorem 3.9). We emphasize that, because we are in a 1D setting (A is tridiagonal), the *approximation property can be proved using only linear algebra arguments*. This is crucial for the convergence analysis presented in this paper. Note that the convergence analyses of multigrid methods for second order elliptic boundary value problems, which are known in the literature and based on an approximation- and smoothing property, make use of a $H^{1+\delta}$-regularity assumption for the underlying boundary value problem, with $\delta > 0$. As is well-known, even for very simple (1D) convection-diffusion problems the constants in such regularity estimates tend to infinity if the ratio between diffusion and convection tends to zero (cf. [25]).

For the analysis below we need the permutation

$$P_k : \ \mathbb{R}^{n_k} \to \mathbb{R}^{n_k}, \quad P_k(x_1, x_2, \ldots, x_{n_k})^T = (x_1, x_3, \ldots, x_{n_k}, x_2, x_4, \ldots, x_{n_k-1})^T$$

which transforms the lexicographical ordering to a red-black ordering. Using this renumbering we obtain:

$$P_k T_\gamma P_k^T =: \begin{bmatrix} I & -T_{rb} \\ -T_{br} & I_{k-1} \end{bmatrix} \ , \tag{3.2}$$

$$r_\gamma P_k^T =: \frac{1}{2}\begin{bmatrix} T_{br} & I_{k-1} \end{bmatrix} \ , \tag{3.3}$$

$$P_k p_\gamma =: \begin{bmatrix} T_{rb} \\ I_{k-1} \end{bmatrix} \ , \tag{3.4}$$

$$r_\gamma T_\gamma p_\gamma = \frac{1}{2}(I_{k-1} - T_{br}T_{rb}) = \frac{1}{2}\text{tridiag}(-(1-\gamma)^2, 1-2\gamma(1-\gamma), -\gamma^2) \ . \tag{3.5}$$

We start with a lemma in which a few results are collected that will be useful in the remainder of this section.

LEMMA 3.1. *For the matrices T_γ, p_γ and r_γ the following holds:*

$$\|T_\gamma\|_2 \le 2 \ , \tag{3.6}$$

$$\|p_\gamma\|_2 \le \sqrt{2} \ , \quad \|r_\gamma\|_2 \le \frac{1}{2}\sqrt{2} \ , \tag{3.7}$$

$$\|r_\gamma T_\gamma p_\gamma\|_2 \le 1 \ , \tag{3.8}$$

$$\|T_\gamma(T_\gamma + \delta I_k)^{-1}\|_2 \le \min\{1, \frac{2}{\delta}\} \quad \textit{for} \ \ \delta > 0 \ , \tag{3.9}$$

$$\|(r_\gamma T_\gamma p_\gamma + \delta I_{k-1})^{-1}\|_2 \le \frac{1}{\delta} \quad \textit{for} \quad \delta > 0 \ , \tag{3.10}$$

$$r_\gamma p_\gamma = I_{k-1} - r_\gamma T_\gamma p_\gamma \ . \tag{3.11}$$

Proof. The result in (3.6) follows from $\|T_\gamma\|_2^2 \le \|T_\gamma\|_\infty \|T_\gamma\|_1 \le 4$. Note that

$$\|p_\gamma\|_2^2 = \|P_k p_\gamma\|_2^2 = \| \begin{bmatrix} T_{rb} \\ I_{k-1} \end{bmatrix} \|_2^2 \le \| \begin{bmatrix} T_{rb} \\ I_{k-1} \end{bmatrix} \|_\infty \| \begin{bmatrix} T_{rb} \\ I_{k-1} \end{bmatrix} \|_1 \le 2 \ ,$$

hence $\|p_\gamma\|_2 \le \sqrt{2}$. With a similar argument one can prove $\|r_\gamma\|_2 \le \frac{1}{2}\sqrt{2}$. Using (3.5) we get

$$\|r_\gamma T_\gamma p_\gamma\|_2^2 \le \|r_\gamma T_\gamma p_\gamma\|_\infty \|r_\gamma T_\gamma p_\gamma\|_1 \le 1 \ ,$$

i.e., the result (3.8) holds. For the result in (3.9) we note that $T_\gamma + T_\gamma^T$ is positive definite and that

$$\begin{aligned} \max_{x \ne 0} \frac{\|T_\gamma(T_\gamma + \delta I_{k-1})^{-1}\|_2^2}{\|x\|_2^2} &= \max_{y \ne 0} \frac{\|T_\gamma y\|_2^2}{\|(T_\gamma + \delta I_{k-1})y\|_2^2} \\ &= \max_{y \ne 0} \frac{\|T_\gamma y\|_2^2}{\|T_\gamma y\|_2^2 + 2\delta(T_\gamma y, y) + \delta^2\|y\|_2^2} \\ &\le \max_{y \ne 0} \frac{\|T_\gamma y\|_2^2}{\|T_\gamma y\|_2^2 + \delta^2\|y\|_2^2} \le \min\{1, \left(\frac{\|T_\gamma\|_2}{\delta}\right)^2\} \ . \end{aligned}$$

Using the result in (3.6) we obtain (3.9). Note that $r_\gamma T_\gamma p_\gamma$ is positive definite (cf. (3.5)) and thus (3.10) follows from

$$\|(r_\gamma T_\gamma p_\gamma + \delta I_{k-1})^{-1}\|_2^2 = \max_{x \ne 0} \frac{\|x\|^2}{\|r_\gamma T_\gamma p_\gamma x\|_2^2 + 2\delta(r_\gamma T_\gamma p_\gamma x, x) + \delta^2\|x\|_2^2} \le \frac{1}{\delta^2} \ .$$

From (3.3), (3.4) and (3.5) it follows that

$$r_\gamma p_\gamma = \frac{1}{2}(T_{br}T_{rb} + I_{k-1}) = I_{k-1} - \frac{1}{2}(I_{k-1} - T_{br}T_{rb}) = I_{k-1} - r_\gamma T_\gamma p_\gamma \ ,$$

and thus the result (3.11) holds. $\square$

In the next lemma we prove a uniform bound for the iteration matrix of the coarse grid correction $I_k - p_\gamma A_H^{-1} r_\gamma A$ for $A \in L_k^*(\gamma)$.

Lemma 3.2. *Take $A = T_\gamma + \delta I_k \in L_k^*(\gamma)$ and define $A_H := r_\gamma T_\gamma p_\gamma + \delta I_{k-1}$ as in* (2.26). *Then the following holds:*

$$\|I_k - p_\gamma A_H^{-1} r_\gamma A\|_2 \leq \sqrt{2} + 1 \ .$$

Proof. We use the identities in (3.2)-(3.5) and the notation $T_H := r_\gamma T_\gamma p_\gamma = A_H - \delta I_{k-1}$, $C_T := I - P_k p_\gamma T_H^{-1} r_\gamma T_\gamma P_k^T$. Now note

$$\begin{aligned} P_k(I - p_\gamma A_H^{-1} r_\gamma A)P_k^T &= C_T + P_k p_\gamma A_H^{-1}(A_H - T_H)T_H^{-1} r_\gamma T_\gamma P_k^T - \delta P_k p_\gamma A_H^{-1} r_\gamma P_k^T \\ &= C_T + \delta P_k p_\gamma A_H^{-1}(T_H^{-1} r_\gamma T_\gamma P_k^T - r_\gamma P_k^T) \ , \end{aligned} \tag{3.12}$$

and

$$T_H^{-1} r_\gamma T_\gamma P_k^T = T_H^{-1} r_\gamma P_k^T P_k T_\gamma P_k^T = T_H^{-1} \begin{bmatrix} \emptyset & T_H \end{bmatrix} = \begin{bmatrix} \emptyset & I_{k-1} \end{bmatrix} \ . \tag{3.13}$$

Using the latter in (3.12) yields

$$\|I - p_\gamma A_H^{-1} r_\gamma A\|_2 \leq \|C_T\|_2 + \frac{1}{2}\delta \|p_\gamma\|_2 \|A_H^{-1}\|_2 \| \begin{bmatrix} -T_{br} & I_{k-1} \end{bmatrix} \|_2 \ . \tag{3.14}$$

For the first term on the right handside in (3.14) we get, using (3.13),

$$\begin{aligned} \|C_T\|_2 = \|P_k C_T P_k^T\|_2 &= \|I - \begin{bmatrix} T_{rb} \\ I_{k-1} \end{bmatrix} \begin{bmatrix} \emptyset & I_{k-1} \end{bmatrix} \|_2 = \| \begin{bmatrix} I & -T_{rb} \\ \emptyset & \emptyset \end{bmatrix} \|_2 \\ &\leq \| \begin{bmatrix} I & -T_{rb} \end{bmatrix} \|_1^{\frac{1}{2}} \| \begin{bmatrix} I & -T_{rb} \end{bmatrix} \|_\infty^{\frac{1}{2}} \leq \sqrt{2} \ . \end{aligned} \tag{3.15}$$

Furthermore,

$$\| \begin{bmatrix} -T_{br} & I_{k-1} \end{bmatrix} \|_2 \leq \| \begin{bmatrix} -T_{br} & I_{k-1} \end{bmatrix} \|_1^{\frac{1}{2}} \| \begin{bmatrix} -T_{br} & I_{k-1} \end{bmatrix} \|_\infty^{\frac{1}{2}} \leq \sqrt{2} \ . \tag{3.16}$$

The inequality $\|A_H^{-1}\| \leq \frac{1}{\delta}$ is given in (3.10). Using this and the results (3.15), (3.16) in (3.14) yields

$$\|I - p_\gamma A_H^{-1} r_\gamma A\|_2 \leq \sqrt{2} + \frac{1}{2}\delta\sqrt{2}\frac{1}{\delta}\sqrt{2} = \sqrt{2} + 1. \qquad \square$$

The main ingredient for the approximation property (Theorem 3.4 below) is given in the following lemma:

Lemma 3.3. *The following inequality holds:*

$$\|(I - p_\gamma r_\gamma)T_\gamma^{-1}\|_2 \leq \frac{1}{2}\sqrt{3 + \sqrt{5}} \ .$$

Proof. First note that for $\tilde{T}_\gamma := P_k T_\gamma P_k^T$ we have

$$\tilde{T}_\gamma = \begin{bmatrix} I & -T_{rb} \\ -T_{br} & I_{k-1} \end{bmatrix} = \begin{bmatrix} I & -T_{rb} \\ \emptyset & I_{k-1} \end{bmatrix} \begin{bmatrix} I - T_{rb}T_{br} & \emptyset \\ -T_{br} & I_{k-1} \end{bmatrix} ,$$

and thus

$$\tilde{T}_\gamma^{-1} = \begin{bmatrix} (I - T_{rb}T_{br})^{-1} & \emptyset \\ * & * \end{bmatrix} \begin{bmatrix} I & T_{rb} \\ \emptyset & I_{k-1} \end{bmatrix} ,$$

with appropriate blocks $*$. Using this we get

$$\begin{aligned} P_k(I - p_\gamma r_\gamma)T_\gamma^{-1}P_k^T &= (I - P_k p_\gamma r_\gamma P_k^T)\tilde{T}_\gamma^{-1} \\ &= (I - \frac{1}{2}\begin{bmatrix} T_{rb}T_{br} & T_{rb} \\ T_{br} & I_{k-1} \end{bmatrix})\tilde{T}_\gamma^{-1} \\ &= \frac{1}{2}(\tilde{T}_\gamma + \begin{bmatrix} I - T_{rb}T_{br} & \emptyset \\ \emptyset & \emptyset \end{bmatrix})\tilde{T}_\gamma^{-1} \\ &= \frac{1}{2}(I + \begin{bmatrix} I & T_{rb} \\ \emptyset & \emptyset \end{bmatrix}) = \frac{1}{2}\begin{bmatrix} 2I & T_{rb} \\ \emptyset & I_{k-1} \end{bmatrix} . \end{aligned}$$

From this we obtain

$$\|(I - p_\gamma r_\gamma)T_\gamma^{-1}\|_2^2 \le \frac{1}{4}\| \begin{bmatrix} 2I & T_{rb} \\ \emptyset & I_{k-1} \end{bmatrix} \|_1 \| \begin{bmatrix} 2I & T_{rb} \\ \emptyset & I_{k-1} \end{bmatrix} \|_\infty \le 1\frac{1}{2} .$$

A slightly sharper bound can be obtained as follows. Note that

$$\|(I - p_\gamma r_\gamma)T_\gamma^{-1}\|_2^2 = \frac{1}{4}\rho(B) \quad \text{with} \quad B = \begin{bmatrix} 4I + T_{rb}T_{rb}^T & T_{rb} \\ T_{rb}^T & I_{k-1} \end{bmatrix} . \tag{3.17}$$

A straightforward computation yields that $\lambda \in \sigma(B)$ implies

$$\frac{\lambda^2 - 5\lambda + 4}{\lambda} \in \sigma(T_{rb}T_{rb}^T) .$$

Using $\|T_{rb}T_{rb}^T\|_\infty = \|T_{rb}T_{rb}^T(1, 1, \ldots, 1)^T\|_\infty \le 1$ we obtain that $\sigma(T_{rb}T_{rb}^T) \subset [0, 1]$ holds. Hence $0 < \lambda^2 - 5\lambda + 4 \le \lambda$ must hold for all eigenvalues $\lambda \in \sigma(B)$. From this one obtains $\lambda \le 3 + \sqrt{5}$ for all $\lambda \in \sigma(B)$. Using this in (3.17) proves the result of this lemma. ▯

Note that for the result in Lemma 3.3 to hold, it is essential that p_γ and r_γ are appropriate matrix-dependent prolongations and restrictions.

THEOREM 3.4 (Approximation Property). *Take $A = T_\gamma + \delta I_k \in L_k^*(\gamma)$ and let $A_H := r_\gamma T_\gamma p_\gamma + \delta I_{k-1}$. Then the following holds:*

$$\|A^{-1} - p_\gamma A_H^{-1} r_\gamma\|_2 \le 5\frac{1}{2}\min\{1, \frac{2}{\delta}\} .$$

Proof. We start with

$$\begin{aligned} &\|A^{-1} - p_\gamma A_H^{-1} r_\gamma\|_2 = \|(I - p_\gamma A_H^{-1} r_\gamma A)(p_\gamma r_\gamma + (I - p_\gamma r_\gamma))T_\gamma^{-1}T_\gamma A^{-1}\|_2 \\ &\le \big(\|(I - p_\gamma A_H^{-1} r_\gamma A)p_\gamma r_\gamma T_\gamma^{-1}\|_2 + \|I - p_\gamma A_H^{-1} r_\gamma A\|_2 \|(I - p_\gamma r_\gamma)T_\gamma^{-1}\|_2\big) \|T_\gamma A^{-1}\|_2 . \end{aligned} \tag{3.18}$$

For the term $\|T_\gamma A^{-1}\|_2$ the inequality

$$\|T_\gamma A^{-1}\|_2 \leq \min\{1, \frac{2}{\delta}\} \tag{3.19}$$

is given in (3.9). For the second term in the upper bound in (3.18) we obtain, using Lemma 3.2 and Lemma 3.3,

$$\|I - p_\gamma A_H^{-1} r_\gamma A\|_2 \|(I - p_\gamma r_\gamma)T_\gamma^{-1}\|_2 \leq (\sqrt{2}+1)\frac{1}{2}\sqrt{3+\sqrt{5}}\ . \tag{3.20}$$

We finally consider the first term in the upper bound in (3.18). For this we introduce $\tilde{A} := (1+\delta)T_\gamma + \delta I_k = A + \delta T_\gamma$. From the result (3.11) it follows that

$$A_H = r_\gamma T_\gamma p_\gamma + \delta I_{k-1} = r_\gamma T_\gamma p_\gamma + \delta(r_\gamma p_\gamma + r_\gamma T_\gamma p_\gamma) = r_\gamma \tilde{A} p_\gamma\ .$$

Using this Galerkin relation between $\tilde{A}$ and A_H we obtain

$$\begin{aligned}(I - p_\gamma A_H^{-1} r_\gamma A)p_\gamma r_\gamma T_\gamma^{-1} &= (I - p_\gamma A_H^{-1} r_\gamma(\tilde{A} - \delta T_\gamma))p_\gamma r_\gamma T_\gamma^{-1} \\ &= (I - p_\gamma A_H^{-1} r_\gamma \tilde{A})p_\gamma r_\gamma T_\gamma^{-1} + \delta p_\gamma A_H^{-1} r_\gamma T_\gamma p_\gamma r_\gamma T_\gamma^{-1} \\ &= \delta p_\gamma A_H^{-1} r_\gamma T_\gamma p_\gamma r_\gamma T_\gamma^{-1}\ .\end{aligned}$$

Using this result and those in (3.10), (3.6), (3.7) and Lemma 3.3 it follows that

$$\begin{aligned}\|(I - p_\gamma A_H^{-1} r_\gamma A)p_\gamma r_\gamma T_\gamma^{-1}\|_2 &= \delta\|p_\gamma A_H^{-1} r_\gamma T_\gamma p_\gamma r_\gamma T_\gamma^{-1}\|_2 \\ &\leq \delta\|p_\gamma\|_2\|A_H^{-1}\|_2\|r_\gamma T_\gamma(I - p_\gamma r_\gamma)T_\gamma^{-1} - r_\gamma\|_2 \\ &\leq \delta\sqrt{2}\frac{1}{\delta}(\|r_\gamma\|_2\|T_\gamma\|_2\|(I - p_\gamma r_\gamma)T_\gamma^{-1}\|_2 + \|r_\gamma\|_2) \\ &\leq \frac{1}{2}\sqrt{2}\sqrt{3+\sqrt{5}} + 1\ .\end{aligned} \tag{3.21}$$

Combination of (3.18)-(3.21) results in

$$\begin{aligned}\|A^{-1} - p_\gamma A_H^{-1} r_\gamma\|_2 &\leq \left\{\frac{1}{2}\sqrt{2}\sqrt{3+\sqrt{5}} + 1 + (\sqrt{2}+1)\frac{1}{2}\sqrt{3+\sqrt{5}}\right\}\min\{1, \frac{2}{\delta}\} \\ &= \left\{1 + \frac{1}{2}(2\sqrt{2}+1)\sqrt{3+\sqrt{5}}\right\}\min\{1, \frac{2}{\delta}\} \leq 5\frac{1}{2}\min\{1, \frac{2}{\delta}\}\ ,\end{aligned}$$

which completes the proof of the theorem. ∎

REMARK 3.5. Note that in the proof of the approximation property we use *only linear algebra arguments*. A relation between the tridiagonal matrix A and an underlying two-point boundary value problem is not used. □

We now turn to the analysis of the smoothing property. We will use a result from [20, 22] (cf. also [12] or Theorem 10.6.8 in [10]). For completeness we cite a special case of this result, which will be used below.

LEMMA 3.6. *Take $A, W \in \mathbb{R}^{n\times n}$ with W invertible and such that $\|I - W^{-1}A\|_2 \leq 1$ is satisfied. Define $S := I - \frac{1}{2}W^{-1}A$. Then the following inequality holds:*

$$\|W^{-1}AS^\nu\|_2 \leq 2\sqrt{\frac{2}{\pi\nu}} \quad for \quad \nu = 1, 2, \ldots\ . \tag{3.22}$$

We first consider the case with damped Jacobi smoothing (2.27): for $A \in L_k^*(\gamma)$, $S^{(\nu)} = S_{\mathrm{J}}^{(\nu)} := (I - \frac{1}{2}\mathrm{diag}(A)^{-1}A)^{\nu}$.

THEOREM 3.7 (Smoothing Property). *For $A = T_\gamma + \delta I_k \in L_k^*(\gamma)$ the following holds:*

$$\|AS_{\mathrm{J}}^{(\nu)}\|_2 \leq (1+\delta)\eta_{\mathrm{J}}(\nu) \quad for \quad \nu = 1, 2, \dots \ ,$$
$$with \quad \eta_{\mathrm{J}}(\nu) = 2\sqrt{\frac{2}{\pi\nu}} \ .$$

Proof. We use the notation $W_{\mathrm{J}} = \mathrm{diag}(A) = (1+\delta)I_k$ and $R = W_{\mathrm{J}} - A = \mathrm{tridiag}(1-\gamma, 0, \gamma)$. It follows that

$$\|I_k - W_{\mathrm{J}}^{-1}A\|_2 = \frac{1}{1+\delta}\|R\|_2 \leq \frac{1}{1+\delta}(\|R\|_\infty \|R\|_1)^{\frac{1}{2}} \leq \frac{1}{1+\delta} \leq 1$$

holds. Application of Lemma 3.6 results in

$$\frac{1}{1+\delta}\|AS_{\mathrm{J}}^{(\nu)}\|_2 \leq 2\sqrt{\frac{2}{\pi\nu}} \ . \qquad \square$$

We now analyze the Gauss-Seidel smoother (2.28): for $A = T_\gamma + \delta I_k \in L_k^*(\gamma)$, $S^{(\nu)} = S_{\mathrm{GS}}^{(\nu)} = (I_k - W_{\mathrm{GS}}^{-1}A)(I_k - \frac{1}{2}W_{\mathrm{GS}}^{-1}A)^{\nu-1}$, with $W_{\mathrm{GS}} = L_\gamma + \delta I_k$. In Lemma 3.6 a damping factor $\theta = \frac{1}{2}$ is used (we note that in the more general analysis in [8] it suffices to have arbitrary $\theta < 1$). On the other hand, the Gauss-Seidel method without damping, i.e. $\theta = 1$, is a so-called robust smoother in the sense that it becomes a direct solver in the limit case $\gamma = 0$. Here we want to benefit both from the smoothing effect as formulated in Lemma 3.6 and from the fact that for $\gamma \downarrow 0$ the Gauss-Seidel method with $\theta = 1$ is a fast solver. This motivates the use of a Gauss-Seidel smoother where we first apply $\nu - 1$ damped Gauss-Seidel iterations with $\theta = \frac{1}{2}$ and then one Gauss-Seidel iteration with $\theta = 1$. For this method one can prove a smoothing property as given in the following theorem.

THEOREM 3.8 (Smoothing Property). *For $A = T_\gamma + \delta I_k \in L_k^*(\gamma)$ the following holds:*

$$\|AS_{\mathrm{GS}}^{(\nu)}\|_2 \leq \gamma\eta_{\mathrm{GS}}(\nu) \quad for \quad \nu = 1, 2, \dots \ , \tag{3.23}$$
$$with \quad \eta_{\mathrm{GS}}(1) = 2, \quad \eta_{\mathrm{GS}}(\nu) = 2\sqrt{\frac{2}{\pi(\nu-1)}} \quad for \quad \nu = 2, 3, \dots \ .$$

Proof. We use the notation $R = W_{\mathrm{GS}} - A = \mathrm{tridiag}(0, 0, \gamma)$. Note that

$$A(I - W_{\mathrm{GS}}^{-1}A) = (W_{\mathrm{GS}} - R)W_{\mathrm{GS}}^{-1}R = RW_{\mathrm{GS}}^{-1}A \tag{3.24}$$

holds. And thus

$$\|AS_{\mathrm{GS}}^{(\nu)}\|_2 \leq \|R\|_2 \|W_{\mathrm{GS}}^{-1}A(I_k - \frac{1}{2}W_{\mathrm{GS}}^{-1}A)^{\nu-1}\|_2 \leq \gamma \|W_{\mathrm{GS}}^{-1}A(I_k - \frac{1}{2}W_{\mathrm{GS}}^{-1}A)^{\nu-1}\|_2 \ . \tag{3.25}$$

To be able to apply Lemma 3.6 we now prove that $\|I_k - W_{\mathrm{GS}}^{-1}A\|_2 \leq 1$ holds. A standard result for the Gauss-Seidel method applied to irreducibly diagonally dominant matrices yields $\|I_k - W_{\rm GS}^{-1}A\|_\infty < 1$. To bound $\|I_k - W_{\rm GS}^{-1}A\|_1$ we use an explicit representation of $W_{\rm GS}^{-1}$. With $J_k := \text{tridiag}(1,0,0) \in \mathbb{R}^{n_k\times n_k}$ we have $W_{\rm GS} = (1+\delta)I_k - (1-\gamma)J_k = (1+\delta)(I_k - \frac{1-\gamma}{1+\delta}J_k)$ and thus

$$W_{\rm GS}^{-1} = \frac{1}{1+\delta}\sum_{m=0}^{n_k-1}\left(\frac{1-\gamma}{1+\delta}\right)^m J_k^m \ .$$

From this we obtain, with $\mathbf{1} = (1,1,\dots,1)^T$,

$$\begin{aligned}\|W_{\rm GS}^{-1}R\|_1 &= \|R^T W_{\rm GS}^{-T}\|_\infty = \|R^T W_{\rm GS}^{-T}\mathbf{1}\|_\infty \\ &\le \gamma\|W_{\rm GS}^{-T}\mathbf{1}\|_\infty = \gamma\frac{1}{1+\delta}\sum_{m=0}^{n_k-1}\left(\frac{1-\gamma}{1+\delta}\right)^m \\ &\le \gamma\frac{1}{1+\delta}\frac{1}{1-\frac{1-\gamma}{1+\delta}} = \frac{\gamma}{\delta+\gamma} \le 1 \ .\end{aligned}$$

It now follows that $\|W_{\rm GS}^{-1}R\|_2^2 \le \|W_{\rm GS}^{-1}R\|_\infty\|W_{\rm GS}^{-1}R\|_1 < 1$. Application of Lemma 3.6 yields for $\nu > 1$:

$$\|W_{\rm GS}^{-1}A(I_k - \frac{1}{2}W_{\rm GS}^{-1}A)^{\nu-1}\|_2 \le 2\sqrt{\frac{2}{\pi(\nu-1)}} \ . \tag{3.26}$$

For $\nu = 1$ we have (cf. (3.24)):

$$\|AS_{\rm GS}^{(1)}\|_2 \le \|R\|_2\|W_{\rm GS}^{-1}A\|_2 \le \|R\|_2(\|I\|_2 + \|I - W_{\rm GS}^{-1}A\|_2) \le 2\gamma \ . \tag{3.27}$$

Combining (3.25), (3.26) and (3.24) proves the theorem. ∎

Note that the upper bound in (3.23) tends to zero for $\gamma \downarrow 0$. As a direct consequence of the approximation property and smoothing property proved above we obtain a two-grid convergence result:

THEOREM 3.9 (Two-grid convergence). *For $A \in L_k^*(\gamma)$, let $M_k(\gamma) = (I_k - p_\gamma A_H^{-1} r_\gamma A)S^{(\nu)}$ be the two-grid iteration matrix* (2.29) *induced by A. For the two-grid method with damped Jacobi smoothing, i.e., $S^{(\nu)} = S_{\rm J}^{(\nu)}$ the inequality*

$$\max_{A\in L_k^*(\gamma)}\|M_k(\gamma)\|_2 \le 16\frac{1}{2}\eta_{\rm J}(\nu) \quad \textit{for} \quad \nu = 1,2,\dots \tag{3.28}$$

holds. For the two-grid method with (damped) Gauss-Seidel smoothing, i.e., $S^{(\nu)} = S_{\rm GS}^{(\nu)}$ the inequality

$$\max_{A\in L_k^*(\gamma)}\|M_k(\gamma)\|_2 \le 5\frac{1}{2}\gamma\eta_{\rm GS}(\nu) \quad \textit{for} \quad \nu = 1,2,\dots \ . \tag{3.29}$$

holds.

Proof. The result in (3.28) follows directly from Theorem 3.4 and Theorem 3.7 and the observation $\max_{\delta\ge 0}\{(1+\delta)\min\{1,\frac{2}{\delta}\}\} = 3$. The result in (3.29) is a direct consequence of Theorem 3.4 and Theorem 3.8. ∎

COROLLARY 3.10. Using Theorem 3.9 in Corollary 2.5 yields a convergence result for the two-grid method applied to a system with a tensor product matrix $\hat{A} \in \hat{L}_k(\gamma)$. For $\hat{A} \in \hat{L}_k(\gamma)$ let $\hat{M}_k(\gamma)$ be the two-grid iteration matrix (2.18) induced by $\hat{A}$. If damped y-line Jacobi smoothing is used (cf. (2.16)), i.e., $\hat{S}^{(\nu)} = (I - \frac{1}{2}\hat{W}_{\mathrm{J}}^{-1}\hat{A})^{\nu}$, then

$$\|\hat{M}_k(\gamma)\|_2 \leq 16\frac{1}{2}\eta_{\mathrm{J}}(\nu) \quad \text{for} \quad \nu = 1, 2, \ldots \tag{3.30}$$

holds. If (damped) y-line Gauss-Seidel smoothing is used (cf. (2.17)), i.e., $\hat{S}^{(\nu)} = (I - \hat{W}_{\mathrm{GS}}^{-1}\hat{A})(I - \frac{1}{2}\hat{W}_{\mathrm{GS}}^{-1}\hat{A})^{\nu-1}$, then

$$\|\hat{M}_k(\gamma)\|_2 \leq 5\frac{1}{2}\gamma\eta_{\mathrm{GS}}(\nu) \quad \text{for} \quad \nu = 1, 2, \ldots \tag{3.31}$$

holds. □

4. W-cycle convergence. In this section we analyze the convergence of the multigrid W-cycle along the lines as in [10] Section 10.6.5. We consider a given problem

$$\hat{A}x = b \quad \text{with} \quad \hat{A} \in \hat{L}_{k_{\max}}(\gamma), \quad \gamma = \gamma_{k_{\max}} \in (0, \frac{1}{2})\ . \tag{4.1}$$

As noted in Remark 2.1, such a problem is obtained if a stable finite difference scheme is applied to a convection-diffusion problem as in (2.3) on a grid $\Omega_{k_{\max},m}$ with grid sizes $h_y = \frac{1}{m}$ and $h_{k_{\max}} = 2^{-k_{\max}}$ in the y- and x-direction, respectively. To the problem (4.1) we apply a W-cycle based on semicoarsening in the x-direction. More precisely, we use a Galerkin approach as described in (2.11). If

$$\hat{A}_k = \alpha_k I_y \otimes T_{\gamma_k} + \beta A_y \otimes I_k \in \hat{L}_k(\gamma_k) \quad \text{with} \quad \gamma_k \in (0, \frac{1}{2}) \tag{4.2}$$

then for $\hat{A}_{k-1} := \alpha_k \hat{r}_{\gamma_k}(I_y \otimes T_{\gamma_k})\hat{p}_{\gamma_k} + \beta A_y \otimes I_{k-1}$ (as in (2.11)) we have:

$$\hat{A}_{k-1} = \alpha_{k-1} I_y \otimes T_{\gamma_{k-1}} + \beta A_y \otimes I_{k-1} \quad \text{with} \tag{4.3}$$

$$\alpha_{k-1} = \frac{1}{2}\alpha_k(1 - 2\gamma_k(1-\gamma_k)) > 0, \quad \gamma_{k-1} = \frac{\gamma_k^2}{1 - 2\gamma_k(1-\gamma_k)}\ , \tag{4.4}$$

i.e., $\hat{A}_{k-1} \in \hat{L}_{k-1}(\gamma_{k-1})$ with $0 < \gamma_{k-1} < \gamma_k$. Hence this Galerkin technique *can be applied recursively* and results in $\hat{A}_k \in \hat{L}_k(\gamma_k)$ for $k = k_{\max} - 1, k = k_{\max} - 2, \ldots, 1$. Note that the matrices $\hat{A}_1 \in \hat{L}_1(\gamma_1)$ are of the form $\alpha I_y + \beta A_y \in \mathbb{R}^{(m-1)\times(m-1)}$. Hence if A_y is a bandmatrix (e.g. tridiagonal) the computational costs for solving a system with matrix $\hat{A}_1$ are low. In the multigrid method, for the intergrid transfer operators between level k and level $k-1$ we use the matrix-dependent prolongation and restriction $\hat{p}_\gamma = \hat{p}_{\gamma_k}$, $\hat{r}_\gamma = \hat{r}_{\gamma_k}$ as in (2.10). As in the two-grid method in Section 2 we consider damped y-line Jacobi smoothing with iteration matrix $\hat{S}_{\mathrm{J}}^{(\nu)}$ as in (2.16) and (damped) y-line Gauss-Seidel method with iteration matrix $\hat{S}_{\mathrm{GS}}^{(\nu)}$ as in (2.17). We use the notation $\hat{S}^{(\nu)}$ to denote $\hat{S}_{\mathrm{J}}^{(\nu)}$ or $\hat{S}_{\mathrm{GS}}^{(\nu)}$. Note that a linear iterative method for solving a system $\hat{A}_k x = b_k$ with iteration matrix $\hat{S}^{(\nu)}$ can be represented as $x^{\mathrm{new}} = \hat{S}^{(\nu)}x^{\mathrm{old}} + (I - \hat{S}^{(\nu)})\hat{A}_k^{-1}b_k$. We use the notation $x^{\mathrm{new}} = \mathcal{S}^{(\nu)}(x^{\mathrm{old}}, b_k)$.

Using the components described above we introduce a multigrid method for solving a problem $\hat{A}_k x = b_k$ with matrix $\hat{A}_k \in \hat{L}_k(\gamma_k)$, $1 \le k \le k_{\max}$:

$$
\begin{array}{l}
\textit{procedure}\ \ MGM_k^{(\nu)}(x_k, b_k) \\
\textit{if}\ \ k = 1\ \ \textit{then}\ \ x_1 := \hat{A}_1^{-1} b_1\ \ \textit{else} \\
\textit{begin}\quad x_k := \mathcal{S}^{(\nu)}(x_k, b_k); \\
\quad d_{k-1} := \hat{r}_{\gamma_k}(\hat{A}_k x_k - b_k); \\
\quad e_{k-1}^{(0)} := 0; \\
\quad \textit{for}\ \ j := 1\ \ \textit{to}\ \ \tau\ \ \textit{do}\ \ e_{k-1}^{(j)} := MGM_{k-1}^{(\nu)}(e_{k-1}^{(j-1)}, d_{k-1}); \\
\quad x_k := x_k - \hat{p}_{\gamma_k} e_{k-1}^{(\tau)}; \\
\quad MGM_k^{(\nu)} := x_k; \\
\textit{end};
\end{array}
\tag{4.5}
$$

In this section we consider the W-cycle, i.e., $\tau = 2$ in the algorithm above. For the iteration matrix $\hat{W}_k = \hat{W}_k(\gamma_k)$ of the W-cycle we have the recursion:

$$
\begin{cases}
\hat{W}_1 = 0 \\
\hat{W}_k = (I_k - \hat{p}_{\gamma_k}(I_{k-1} - \hat{W}_{k-1}^2)\hat{A}_{k-1}^{-1}\hat{r}_{\gamma_k}\hat{A}_k)\hat{S}^{(\nu)}, \quad 2 \le k \le k_{\max}\ .
\end{cases}
\tag{4.6}
$$

REMARK 4.1. Consider $\hat{A}_k$ as in (4.2) with A_y a $(m-1)\times(m-1)$ bandmatrix with at most M nonzero entries per row. The constant M is assumed to be independent of m and k. Due to the semicoarsening the arithmetic costs (number of arithmetic operations) needed in one W-cycle iteration are not proportional to the number of unknowns. Let $N_k := n_k \times (m-1)$ be the number of interior grid points (= the number of unknowns) on level k. Note that $N_{k-1} = \frac{1}{2}N_k$. The arithmetic costs of a matrix-vector multiplication $\hat{A}_k x$ and, for fixed ν, of the application of the smoother $\mathcal{S}^{(\nu)}$ on level k are both proportional to N_k. The arithmetic costs for one application of the intergrid transfer operators $\hat{r}_{\gamma_k}$ and $\hat{p}_{\gamma_k}$ are also proportional to N_k. Using a standard argument (cf. [10], Section 10.4.4) it follows that the total arithmetic costs for one W-cycle iteration on level k are bounded by $CN_k \log N_k$, with C independent of k and m. □

LEMMA 4.2. *Let $\hat{W}_k = \hat{W}_k(\gamma_k)$ as in (4.6) be the iteration matrix of the W-cycle and let $\hat{M}(\gamma_k)$ as in (2.18) be the iteration matrix of the two-grid method induced by $\hat{A}_k \in \hat{L}_k(\gamma_k)$. Then for $2 \le k \le k_{\max}$, we have:*

$$\|\hat{W}_k(\gamma_k)\|_2 \le \|\hat{M}(\gamma_k)\|_2 + (1+\sqrt{2})\|\hat{W}_{k-1}(\gamma_{k-1})\|_2^2\ .$$

Proof. From

$$\hat{W}_k(\gamma_k) = \hat{M}(\gamma_k) + \hat{p}_{\gamma_k}\hat{W}_{k-1}(\gamma_{k-1})^2\hat{A}_{k-1}^{-1}\hat{r}_{\gamma_k}\hat{A}_k\hat{S}^{(\nu)}$$

we obtain

$$\|\hat{W}_k(\gamma_k)\|_2 \le \|\hat{M}(\gamma_k)\|_2 + \|\hat{p}_{\gamma_k}\|_2\|\hat{W}_{k-1}(\gamma_{k-1})\|_2^2\|\hat{A}_{k-1}^{-1}\hat{r}_{\gamma_k}\hat{A}_k\|_2\|\hat{S}^{(\nu)}\|_2\ . \tag{4.7}$$

Using the orthogonal transformations as in Lemma 2.2 it follows that

$$\|\hat{p}_{\gamma_k}\|_2 = \|p_{\gamma_k}\|_2 \leq \sqrt{2} \tag{4.8}$$

and, using the notation $T_H := r_{\gamma_k} T_{\gamma_k} p_{\gamma_k}$,

$$\begin{aligned}\|\hat{A}_{k-1}^{-1}\hat{r}_{\gamma_k}\hat{A}_k\|_2 &\leq \max_{\delta\geq 0} \|(T_H+\delta I_{k-1})^{-1} r_{\gamma_k}(T_{\gamma_k}+\delta I_k)\|_2 \\ &\leq \max_{\delta\geq 0} \left\{ \|(T_H+\delta I_{k-1})^{-1}T_H\|_2 \|T_H^{-1} r_{\gamma_k} T_{\gamma_k}\|_2 + \delta \|(T_H+\delta I_{k-1})^{-1}\|_2 \|r_{\gamma_k}\|_2 \right\} \ .\end{aligned}$$

Using the results in (3.13), (3.10), (3.7) and $\|(T_H+\delta I_{k-1})^{-1}T_H\|_2 \leq 1$ we get

$$\|\hat{A}_{k-1}^{-1}\hat{r}_{\gamma_k}\hat{A}_k\|_2 \leq 1 + \frac{1}{2}\sqrt{2} \ . \tag{4.9}$$

We now consider $\hat{S}^{(\nu)} = \hat{S}_{\mathrm{J}}^{(\nu)} = (I - \frac{1}{2}\hat{W}_{\mathrm{J}}^{-1}\hat{A}_k)^\nu$. In the proof of Theorem 3.7 it is shown that $\|I - W_{\mathrm{J}}^{-1}A\|_2 \leq 1$ holds for all $A \in L_k^*(\gamma_k)$. Using this and the results in (2.20), (2.24) it follows that for $\hat{A}_k \in \hat{L}_k(\gamma_k)$

$$\begin{aligned}\|I - \frac{1}{2}\hat{W}_{\mathrm{J}}^{-1}\hat{A}_k\|_2 &\leq \max_{A\in L_k^*(\gamma_k)} \|I - \frac{1}{2}W_{\mathrm{J}}^{-1}A\|_2 \\ &\leq \frac{1}{2} + \frac{1}{2}\max_{A\in L_k^*(\gamma_k)} \|I - W_{\mathrm{J}}^{-1}A\|_2 \leq 1 \ .\end{aligned}$$

Hence $\|\hat{S}_{\mathrm{J}}^{(\nu)}\|_2 \leq 1$ holds. Finally, we consider $\hat{S}^{(\nu)} = \hat{S}_{\mathrm{GS}}^{(\nu)}$. In the proof of Theorem 3.8 it is shown that $\|I - W_{\mathrm{GS}}^{-1}A\|_2 \leq 1$ holds for all $A \in L_k^*(\gamma_k)$. Using a similar argument as for the Jacobi method one can prove that $\|\hat{S}_{\mathrm{GS}}^{(\nu)}\|_2 \leq 1$ holds. Hence, both for the Jacobi and Gauss-Seidel method the inequality

$$\|\hat{S}^{(\nu)}\|_2 \leq 1 \tag{4.10}$$

holds. The results in (4.7)-(4.10) prove the statement of this lemma. ☐

For the convergence of the W-cycle method we obtain the following result:

THEOREM 4.3. *Let $\hat{W}_{k_{\max}}(\gamma_{k_{\max}})$ be the iteration matrix of the W-cycle method, (4.5) with $\tau = 2$, applied to the problem (4.1). Assume that ν is sufficiently large such that $66(1+\sqrt{2})\eta_{\mathrm{J}}(\nu) < 1$ is satisfied. Then for the W-cycle method with damped Jacobi smoothing ($\hat{S}^{(\nu)} = \hat{S}_{\mathrm{J}}^{(\nu)}$) the inequalities*

$$\|\hat{W}_{k_{\max}}(\gamma_{k_{\max}})\|_2 \leq \frac{33\eta_{\mathrm{J}}(\nu)}{1+\sqrt{1-66(1+\sqrt{2})\eta_{\mathrm{J}}(\nu)}} < 33\eta_{\mathrm{J}}(\nu) < 1 \tag{4.11}$$

hold. Assume that ν is sufficiently large such that $22\gamma_{k_{\max}}(1+\sqrt{2})\eta_{\mathrm{GS}}(\nu) < 1$ is satisfied. Then for the W-cycle method with Gauss-Seidel smoothing ($\hat{S}^{(\nu)} = \hat{S}_{\mathrm{GS}}^{(\nu)}$) the inequalities

$$\|\hat{W}_{k_{\max}}(\gamma_{k_{\max}})\|_2 \leq \frac{11\gamma_{k_{\max}}\eta_{\mathrm{GS}}(\nu)}{1+\sqrt{1-22\gamma_{k_{\max}}(1+\sqrt{2})\eta_{\mathrm{GS}}(\nu)}} < 11\gamma_{k_{\max}}\eta_{\mathrm{GS}}(\nu) < 1 \tag{4.12}$$

hold.

Proof. Define $T := \max_{1\le k\le k_{\max}} \|\hat{M}(\gamma_k)\|_2$. Assume that

$$4(1+\sqrt{2})T < 1 \tag{4.13}$$

holds. Then it follows from Lemma 4.2 that

$$\|\hat{W}_{k_{\max}}(\gamma_{k_{\max}})\|_2 \le \xi^* := \frac{2T}{1+\sqrt{1-4(1+\sqrt{2})T}} < 2T\ , \tag{4.14}$$

with ξ^* the fixed point of the iteration

$$x_1 = 0, \quad x_i = T + (1+\sqrt{2})x_{i-1}^2, \quad i \ge 2\ .$$

For the Jacobi method it follows from (3.30) that $T = T_{\mathrm{J}} \le 16\frac{1}{2}\eta_{\mathrm{J}}(\nu)$. Hence (4.13) holds if $66(1+\sqrt{2})\eta_{\mathrm{J}}(\nu) < 1$ is satisfied. The result in (4.11) then follows from (4.14). For the Gauss-Seidel method it follows from (3.31) and from $\gamma_k < \gamma_{k_{\max}}$ for $k < k_{\max}$ that $T = T_{\mathrm{GS}} \le 5\frac{1}{2}\gamma_{k_{\max}}\eta_{\mathrm{GS}}(\nu)$. Hence (4.13) holds if $\gamma_{k_{\max}}22(1+\sqrt{2}) < 1$ is satisfied. The result in (4.12) then follows from (4.14). $\square$

Note that the bound $33\eta_{\mathrm{J}}(\nu)$ in (4.11) *depends only on* ν. Hence, for ν sufficiently large we have a bound for the contraction number of the W-cycle with Jacobi smoothing which is smaller than one and *independent of* m, $k_{\max}$ *and* $\gamma_{k_{\max}}$. If the problem (4.1) corresponds to a discrete convection-diffusion problem as in (2.3) this means that the W-cycle method with Jacobi smoothing is a *robust solver* in the sense that for ν sufficiently large its contraction number is smaller than one and independent of both the grid size and the parameter ε.
Due to $\gamma_{k_{\max}} < \frac{1}{2}$ the result in (4.12) yields the bound $5\frac{1}{5}\eta_{\mathrm{GS}}(\nu)$ for the contraction number of the W-cycle with Gauss-Seidel smoothing. Hence this Gauss-Seidel method results in a robust solver, too. Note, however, that the bound in (4.12) improves if $\gamma_{k_{\max}}$ becomes smaller. We comment on this in the following remark.

REMARK 4.4. It is clear that if $\gamma \downarrow 0$ the y-line Gauss-Seidel method with iteration matrix $\hat{S}_{\mathrm{GS}}^{(1)} = I - \hat{W}_{\mathrm{GS}}^{-1}\hat{A}$ becomes an exact solver. Hence, for γ "sufficiently small" we do not need a coarse grid correction because the smoother is already a fast solver. Here we want to be more precise about what "sufficiently small" means and show that even for "small" γ the use of a coarse grid correction may result in a significantly faster convergence. We consider the boundary value problem as in (2.3) with $b(y) = \varepsilon$ for all y. A uniform square grid with mesh size $h = h_k = h_y$ in both directions is used. We use the notation $n = n_k = \frac{1}{h} - 1$. The term $\frac{\partial u}{\partial x}$ is discretized using full upwind differences. This results in a discrete problem with matrix

$$\hat{A} = \hat{A}_k = \alpha I_k \otimes T_\gamma + A_y \otimes I_k$$

as in (2.7) with $\gamma = \frac{\varepsilon}{h}(2\frac{\varepsilon}{h}+1)^{-1}$, $\alpha = \frac{1}{h}(2\frac{\varepsilon}{h}+1) = \frac{\varepsilon}{h^2}\frac{1}{\gamma}$. Note that $\sigma(A_y) = \{\mu_i\}_{1\le i\le n}$ with $\mu_1 := \min_{1\le i\le n}\mu_i = \varepsilon\pi^2(1+\mathcal{O}(h))$. We define $\delta_i = \frac{\mu_i}{\alpha}$. Note that $\delta_1 = \gamma\pi^2h^2(1+\mathcal{O}(h)) \le \frac{1}{2}\pi^2h^2(1+\mathcal{O}(h))$. Here and in the following the constants in the $\mathcal{O}(h)$ terms are independent of all parameters. Using the orthogonal transformation as in Lemma 2.2 yields, with $R = \mathrm{tridiag}(0,0,1)$,

$$\begin{aligned}\|\hat{S}_{\mathrm{GS}}^{(1)}\|_2 &= \max_{1\le i\le n}\|I - (L_\gamma + \delta_i I_k)^{-1}(T_\gamma + \delta_i I_k)\|_2 \\ &= \gamma \max_{1\le i\le n}\|(L_\gamma + \delta_i I_k)^{-1}R\|_2 = \gamma\|(L_\gamma + \delta_1 I_k)^{-1}R\|_2\ .\end{aligned}$$

We now analyze the dependence of the contraction number $\|\hat{S}^{(1)}_{\rm GS}\|_2$ on the parameters h and γ. Using $RR^T = I_k - e_n e_n^T$, with $e_n = (0,\ldots,0,1)^T$, and $(L_\gamma + \delta_1 I_k)^{-1} e_n = (1+\delta_1)^{-1} e_n$ it follows that

$$
\begin{aligned}
\|\hat{S}^{(1)}_{\rm GS}\|_2^2 &= \gamma^2 \|L_\gamma + \delta_1 I_k)^{-1} RR^T (L_\gamma + \delta_1 I_k)^{-T}\|_2 \\
&= \gamma^2 \|((L_\gamma^T + \delta_1 I_k)(L_\gamma + \delta_1 I_k))^{-1} - (1+\delta_1)^{-2} e_n e_n^T\|_2
\end{aligned}
$$

and thus

$$
\begin{aligned}
&\gamma^2(\|((L_\gamma^T + \delta_1 I_k)(L_\gamma + \delta_1 I_k))^{-1}\|_2 - 1) \le \|\hat{S}^{(1)}_{\rm GS}\|_2^2 \le \\
&\gamma^2(\|((L_\gamma^T + \delta_1 I_k)(L_\gamma + \delta_1 I_k))^{-1}\|_2 + 1) \ .
\end{aligned}
\tag{4.15}
$$

A longer straightforward computation results in

$$
(L_\gamma^T + \delta_1 I_k)(L_\gamma + \delta_1 I_k) = (1-\gamma)(1+\delta_1)B - (1-\gamma)^2 e_n e_n^T + (\gamma+\delta_1)^2 I_k \ ,
$$

with $B = \text{tridiag}(-1,2,-1)$. From this and $\lambda_{\min}(B) = ch^2(1+\mathcal{O}(h))$, $\lambda_{\min}(B - e_n e_n^T) = \tilde{c}h^2(1+\mathcal{O}(h))$ with positive constants c and $\tilde{c}$ it follows that there are constants $c_1, c_2 > 0$ independent of h and γ such that

$$
\lambda_{\min}((L_\gamma^T + \delta_1 I_k)(L_\gamma + \delta_1 I_k)) \in [c_1 h^2 + \gamma^2,\ c_2 h^2 + \gamma^2]
$$

holds. Using this in (4.15) one finally obtains

$$
\frac{\gamma^2}{c_2 h^2 + \gamma^2}(1 - c_2 h^2 - \gamma^2) \le \|\hat{S}^{(1)}_{\rm GS}\|_2^2 \le \frac{\gamma^2}{c_1 h^2 + \gamma^2}(1 + c_1 h^2 + \gamma^2) \ . \tag{4.16}
$$

This result shows the dependence of the contraction number $\|\hat{S}^{(1)}_{\rm GS}\|_2$ on h and γ. We compare this with the result of Theorem 4.3 for the W-cycle with *one* Gauss-Seidel iteration ($\nu = 1$) in the smoother. If γ is sufficiently small such that $22\gamma(1+\sqrt{2})\eta_{\rm GS}(1) < 1$ is satisfied then for the contraction number of the W-cycle the estimate

$$
\|\hat{W}_k(\gamma)\|_2 < c\gamma\ < 1 \tag{4.17}
$$

holds with c independent of h and γ. There is an interesting difference between the bounds in (4.16) and in (4.17). As a measure for the convection-diffusion ratio we take the parameter $\frac{\varepsilon}{h}$. Note that for $\frac{\varepsilon}{h} \le \frac{1}{4}$ we have $\frac{2}{3}\frac{\varepsilon}{h} \le \gamma \le \frac{\varepsilon}{h}$ and an upper bound in (4.17) of the form $c\frac{\varepsilon}{h}$. Hence for $\frac{\varepsilon}{h}$ sufficiently small the error reduction in a W-cycle with $\nu = 1$ must be large. On the other hand, the result in (4.16) shows that for the Gauss-Seidel method the contraction number tends to one if $\frac{\varepsilon}{h} \downarrow 0$ and $\frac{\varepsilon}{h^2} \to \infty$. Roughly speaking, the Gauss-Seidel method has a small contraction number iff $\frac{\varepsilon}{h^2}$ is sufficiently small, whereas the W-cycle has a small contraction number if $\frac{\varepsilon}{h}$ is sufficiently small. Hence even for the convection-dominated case with $\frac{\varepsilon}{h} \ll 1$ a multigrid Gauss-Seidel method (W-cycle) can have a significantly higer rate of convergence than a one-grid Gauss-Seidel method. $\Box$

5. V-cycle convergence. In this section we study the convergence of the V-cycle algorithm (4.5) with $\tau = 1$ applied to the problem (4.1). We consider only Gauss-Seidel smoothing with iteration matrix $\hat{S}^{(\nu)} = \hat{S}^{(\nu)}_{\rm GS}$. It will be proved that for the *convection-dominated* case ($\gamma_{k_{\max}} \le \frac{1}{4}$) a bound for the contraction number similar to the one for the W-cycle in (4.12) holds. For the analysis it is crucial that for the two-grid method with Gauss-Seidel smoother the upper bound in (3.31) for the contraction number on level k contains a factor γ_k and that in the Galerkin method we have $\gamma_{k-1} = \frac{\gamma_k^2}{1-2\gamma_k(1-\gamma_k)}$ for $k = k_{\max}, k_{\max}-1, \ldots, 2$. Due to this, if $\gamma_{k_{\max}} \le \hat{\gamma} < \frac{1}{2}$ there is a significant decrease in the bound for the two-grid contraction number if the level number k decreases. The analysis does not yield satisfactory results for the diffusion-dominated case $\gamma_k \approx \frac{1}{2}$ because then $\gamma_{k-1} \approx \gamma_k$ and the bound for the two-grid contraction number is approximately constant as a function of the level number k.

For the V-cycle algorithm, (4.5) with $\tau = 1$, the iteration matrix $\hat{V}_k = \hat{V}_k(\gamma_k)$ satisfies the recursion

$$\begin{cases} \hat{V}_1 = 0 \\ \hat{V}_k = (I_k - \hat{p}_{\gamma_k}(I_{k-1} - \hat{V}_{k-1})\hat{A}_{k-1}^{-1}\hat{r}_{\gamma_k}\hat{A}_k)\hat{S}^{(\nu)}, \quad 2 \le k \le k_{\max} \ . \end{cases}$$

THEOREM 5.1. *Let $\hat{V}_{k_{\max}}(\gamma_{k_{\max}})$ be the iteration matrix of the V-cycle with $\hat{S}^{(\nu)} = \hat{S}_{\mathrm{GS}}^{(\nu)}$ applied to problem* (4.1). *Then for $\gamma_{k_{\max}} \le \frac{1}{4}$ the following holds:*

$$\|\hat{V}_{k_{\max}}(\gamma_{k_{\max}})\|_2 \le C\gamma_{k_{\max}}\eta_{\mathrm{GS}}(\nu) \ , \tag{5.1}$$

with a constant C which does not depend on any of the parameters.

Proof. Note that for $2 \le k \le k_{\max}$

$$\hat{V}_k(\gamma_k) = \hat{M}(\gamma_k) + \hat{p}_{\gamma_k}\hat{V}_{k-1}(\gamma_{k-1})\hat{A}_{k-1}^{-1}\hat{r}_{\gamma_k}\hat{A}_k\hat{S}_{\mathrm{GS}}^{(\nu)} \ .$$

Hence, with $m_k := \|\hat{M}_k(\gamma_k)\|_2$:

$$\|\hat{V}_k(\gamma_k)\|_2 \le m_k + \|\hat{p}_{\gamma_k}\|_2\|\hat{V}_{k-1}(\gamma_{k-1})\|_2\|\hat{A}_{k-1}^{-1}\hat{r}_{\gamma_k}\hat{A}_k\|_2\|\hat{S}_{\mathrm{GS}}^{(\nu)}\|_2 \ .$$

In the proof of Lemma 4.2 it is shown that $\|\hat{p}_{\gamma_k}\|_2\|\hat{A}_{k-1}^{-1}\hat{r}_{\gamma_k}\hat{A}_k\|_2\|\hat{S}_{\mathrm{GS}}^{(\nu)}\|_2 \le 1 + \sqrt{2}$ holds. Thus we obtain

$$\|\hat{V}_k(\gamma_k)\|_2 \le m_k + (1+\sqrt{2})\|\hat{V}_{k-1}(\gamma_{k-1})\|_2$$

and using the bound for $\|\hat{M}_k(\gamma_k)\|_2$ in (3.31)

$$\|\hat{V}_{k_{\max}}(\gamma_{k_{\max}})\|_2 \le \sum_{k=0}^{k_{\max}-2}(1+\sqrt{2})^k m_{k_{\max}-k} \le 5\frac{1}{2}\eta_{\mathrm{GS}}(\nu)\sum_{k=0}^{k_{\max}-2}(1+\sqrt{2})^k\gamma_{k_{\max}-k} \ . \tag{5.2}$$

From $\gamma_{k-1} = \frac{\gamma_k^2}{1-2\gamma_k(1-\gamma_k)}$ it follows that $\gamma_{k_{\max}-k}$ is a decreasing function of k. Using the notation $a_k = (1+\sqrt{2})^k\gamma_{k_{\max}-k}$ one gets

$$\begin{aligned} \frac{a_{k+1}}{a_k} &= (1+\sqrt{2})\frac{\gamma_{k_{\max}-k-1}}{\gamma_{k_{\max}-k}} = (1+\sqrt{2})\frac{\gamma_{k_{\max}-k}}{1-2\gamma_{k_{\max}-k}(1-\gamma_{k_{\max}-k})} \\ &\le (1+\sqrt{2})\frac{\gamma_{k_{\max}}}{1-2\gamma_{k_{\max}}(1-\gamma_{k_{\max}})} =: \beta < 1 \quad \text{for} \ \ k = 0, 1, \ldots, k_{\max}-2 \ . \end{aligned}$$

The inequality $\beta < 1$ holds due to the assumption $\gamma_{k_{\max}} \le \frac{1}{4}$. From $a_{k+1} \le \beta a_k$ we get $a_k \le \beta^k a_0$ and

$$\sum_{k=0}^{k_{\max}-2} a_k < \frac{a_0}{1-\beta} = \gamma_{k_{\max}}\hat{c}_\gamma \tag{5.3}$$

with $\hat{c}_\gamma = (1-\beta)^{-1} = 1 + \gamma_{k_{\max}}(1+\sqrt{2})(1-(3+\sqrt{2})\gamma_{k_{\max}} + \gamma_{k_{\max}}^2)^{-1}$, which is bounded for all $\gamma_{k_{\max}} \in [0, \frac{1}{4}]$. Using (5.3) in (5.2) proves the theorem. ∎

REMARK 5.2. A standard argument as in Remark 4.1 shows that the arithmetic costs for one V-cycle iteration are proportional to the number of unknowns. □

COROLLARY 5.3. As a direct consequence of Theorem 5.1 we obtain the following result. Consider the matrix class $\hat{L}_k^c := \cup_{\gamma \leq \frac{1}{4}} \hat{L}_k(\gamma)$. Then for the V-cycle with Gauss-Seidel smoothing applied to $\hat{A}_k \in \hat{L}_k^c$ the following bound for the contraction number holds:

$$\|\hat{V}_k\|_2 \leq K\eta_{\mathrm{GS}}(\nu) \ ,$$

with a constant K that does not depend on any of the parameters. Thus for ν sufficiently large we have a bound for the contraction number which is smaller than one uniformly in k and in $\gamma \leq \frac{1}{4}$. □

In Theorem 5.1 the assumption $\gamma_{k_{\max}} \leq \frac{1}{4}$ is used. With a few more technicalities a V-cycle bound similar to the one in (5.1) can be proved under the assumption $\gamma_{k_{\max}} \leq \hat{\gamma} < \frac{1}{2}$. The constant C in (5.1) then depends on $\hat{\gamma}$. The analysis as in the proof of Theorem 5.1, however, does not yield a satisfactory bound for the V-cycle contraction number if we only assume $\gamma_{k_{\max}} < \frac{1}{2}$. Hence, a strong robustness result as has been shown to hold for the W-cycle in Section 4, has not been proven for the V-cycle, yet.
Note that for the discrete convection-equation as in Remark 2.1 with γ as in (2.5) the condition $\gamma \leq \frac{1}{4}$ is equivalent to the condition $\frac{\varepsilon}{h} \leq \frac{1}{2}$. Hence $\cup_k \hat{L}_k^c$ can be considered as the class of discretization matrices corresponding to *all convection-dominated* problems. In this setting the result in Corollary 5.3 states that, for the V-cycle contraction number with ν sufficiently large, there is a uniform bound smaller than one for all convection-dominated problems.

REFERENCES

[1] Bank, R.E., Benbourenane, M. (1991): A Fourier analysis of the two-level hierarchical basis multigrid method for convection-diffusion equations. In: Proceedings of the Fourth International Symposium on Domain Decomposition Methods for Partial Differential Equations, 1–8, SIAM, Philadelphia
[2] Bank, R.E., Benbourenane, M. (1992): The hierarchical basis multigrid method for convection-diffusion equations. Numer. Math. **61**, 7–37
[3] Bey, J., Wittum, G. (1997): Downwind numbering: robust multigrid for convection-diffusion problems. Appl. Numer. Math. **23**, 177–192
[4] Börm, S., Hiptmair, R. (1999): Analysis of tensor product multigrid. Preprint, Mathematisches Seminar, University of Kiel
[5] Bramble, J.H. (1993): Multigrid methods. Longman, Harlow
[6] Bramble, J.H., Pasciak, J.E., Xu, J. (1988): The analysis of multigrid algorithms for nonsymmetric and indefinite problems. Math. Comp. **51**, 389–414
[7] Dendy, J.E. Jr.(1983): Black box multigrid for nonsymmetric problems. Appl. Math. Comput. **13**, 261–283
[8] Ecker, A., Zulehner, W. (1996): On the smoothing property of multigrid methods in the non-symmetric case. Numer. Linear Algebra Appl. **3**, 161–172
[9] Hackbusch, W. (1985): Multi-Grid Methods and Applications. Springer, Berlin
[10] Hackbusch, W. (1994): Iterative Solution of Large Sparse Systems of Equations. Springer, New-York
[11] Hackbusch, W. (1984): Multigrid convergence for a singular perturbation problem. Linear Algebra Appl. **58**, 125–145.

[12] Hackbusch, W. (1995): A note on Reusken's Lemma. Computing **55**, 181–189

[13] Hackbusch, W., and T. Probst (1997): Downwind Gauss-Seidel smoothing for convection dominated problems. Numer. Linear Algebra Appl. **4**, 85–102

[14] Johannsen, K. (1999): Robust smoothers for convection-diffusion problems. Preprint IWR, University of Heidelberg

[15] Mandel, J. (1986): Multigrid convergence for nonsymmetric, indefinite variational problems and one smoothing step. Appl. Math. Comput. **19** , 201–216

[16] Mulder, W. (1989): A new multigrid approach to convection problems. J. Comput. Phys. **83**, 303–323

[17] Naik, N.H., van Rosedale, J. (1993): The improved robustness of multigrid elliptic solvers based on multiple semicoarsened grids. SIAM Numer. Anal. **30**, 215–229

[18] Persson, I., Samuelsson, K. , Szepessy, A. (1999): On the convergence of multigrid methods for flow problems. ETNA **8**, 46–87

[19] Pflaum, C. (2000): Robust convergence of multilevel algorithms for convection-diffusion equations. SIAM J. Numer. Anal. **37**, 443–469

[20] Reusken, A. (1992): The smoothing property for regular splittings. In: Incomplete Decompositions (ILU)–Algorithms, Theory, and Applications (W. Hackbusch and G. Wittum editors), Vieweg, Braunschweig, 130–138

[21] Reusken, A. (1994): Multigrid with matrix-dependent transfer operators for convection-diffusion problems. In: Multigrid Methods 4, Proceedings of the fourth multigrid conference (eds. P.W. Hemker, P. Wesseling), International series of Numerical Mathematics **116**, 269–280

[22] Reusken, A. (1992): On maximum norm convergence of multigrid methods for two-point boundary value problems. SIAM J. Numer. Anal. **29**, 1569–1578

[23] Reusken, A. (1994): Multigrid with matrix-dependent transfer operators for a singular perturbation problem. Computing **50**, 199–211

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

[25] Roos, H.-G., M. Stynes, and L. Tobiska (1996): Numerical Methods for Singularly Perturbed Differential Equations. Springer, Berlin

[26] Wang, J. (1993): Convergence analysis of multigrid algorithms for nonselfadjoint and indefinite elliptic problems. SIAM J. Numer. Anal. **30**, 275–285

[27] Wesseling, P. (1992): An Introduction to Multigrid Methods. Wiley, Chichester

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

[29] Yserentant, H. (1993): Old and new convergence proofs for multigrid methods. Acta Numerica 1993, 285–326

[30] Zeeuw, P.M. de (1990): Matrix-dependent prolongations and restrictions in a blackbox multigrid solver. J. Comput. Appl. Math. **33**, 1–27