CONVERGENCE ANALYSIS OF A MULTIGRID SOLVER FOR A FINITE ELEMENT METHOD APPLIED TO CONVECTION-DIFFUSION PROBLEMS

MAXIM A. OLSHANSKII* AND ARNOLD REUSKEN †

Abstract. The paper presents a convergence analysis of a multigrid solver for a system of linear algebraic equations resulting from the disretization of a convection-diffusion problem using a finite element method. We consider piecewise linear finite elements in combination with a streamline diffusion stabilization . We analyze a multigrid method that is based on canonical inter-grid transfer operators, a "direct discretization" approach for the coarse-grid operators, a smoother of line-Jacobi type and one nonstandard component which is called a "local presolver". A robust (diffusion and h-independent) bound for the contraction number of the two-grid method and the multigrid W-cycle are proved for a special class of convection-diffusion problems, namely with Neumann conditions on the outflow boundary, Dirichlet conditions on the rest of the boundary and a flow direction that is constant and aligned with gridlines. Our convergence analysis is based on modified smoothing and approximation properties. The arithmetic complexity of one multigrid iteration is optimal up to a logarithmic term.

AMS Subject Classification: 65F10, 65N22, 65N30, 65N55.

Key Words: multigrid, streamline diffusion, convection-diffusion

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 is well developed (cf. [3, 7, 33, 34]). 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

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

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_k. Note that in this discrete problem we have three interesting parameters: h_k (mesh size), ε (convection-diffusion ratio) 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_k, 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, matrix-dependent prolongations and restrictions and semicoarsening techniques. For an explanation of these methods we refer to [7, 31, 2, 11, 12, 14, 15, 20, 35]. These

* Dept. Mechanics and Mathematics, Moscow State University, Moscow 119899, Russia. The work was partially supported by the Russian Foundation for Basic Research: grant No 02-01-0615 linked to project 02-01-00592.

†Institut für Geometrie und Praktische Mathematik, RWTH-Aachen, D-52056 Aachen, Germany

modifications are based on heuristic arguments and empirical studies and rigorous convergence analysis proving robustness is still missing for most of these modifications.

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 perturbation arguments [4, 7, 13, 30]. 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 [9, 22] 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 [19, 23] 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 [19] the problem (1.1) with periodic boundary conditions and $\phi = 0$ is studied. A bound for V-cycle contraction number is proved which is uniform in ε and h_k provided $\varepsilon \leq ch_k$ is satisfied with c a positive constant that does not depend on ε or h_k. In [23] a two-grid method for solving a first order upwinding 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 ε, h_k, ϕ. In [1] 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_k. In [24] the convergence of a multigrid method applied to a standard finite difference discretization of the problem (1.1) with $\phi = 0$ is analyzed. This method is based on semicoarsening, a matrix-dependent prolongation and restriction and a line smoother. It is proved that the multigrid W-cycle has a contraction number (in the euclidean norm) smaller than one independent of h_k and ε. The analysis is based on linear algebra arguments only.

In the present paper we consider the convection-diffusion problem

$$\begin{aligned} -\varepsilon\Delta u + u_x &= f \quad \text{in} \quad \Omega := (0,1)^2 \\ \frac{\partial u}{\partial x} &= 0 \quad \text{on} \quad \Gamma_E := \{\, (x,y) \in \overline{\Omega} \mid x = 1 \,\} \\ u &= 0 \quad \text{on} \quad \partial\Omega \setminus \Gamma_E \end{aligned} \tag{1.2}$$

In this problem we have Neumann boundary conditions on the outflow boundary and Dirichlet boundary conditions on the remaining part of the boundary. Hence, the solution may have parabolic layers but exponential boundary layers at the outflow boundary do not occur. For this case an a priori regularity estimate of the form $\|u\|_{H^2} \leq c\,\varepsilon^{-1}\|f\|_{L^2}$ holds, whereas for the case with an exponential boundary layer one only has $\|u\|_{H^2} \leq c\,\varepsilon^{-\frac{3}{2}}\|f\|_{L^2}$. The fact that for the case with Neumann outflow boundary conditions the regularity bound is significantly better is important for our analysis. We remark that Neumann outflow conditions are often used in practice (cf. [26]).

Due to the Dirichlet boundary conditions a Fourier analysis is not applicable.

For the discretization we use conforming linear finite elements. As far as we know there is no multigrid convergence analysis for convection-dominated problems known in the literature that can be applied in a finite element setting. In this paper we present an analysis which partly fills this gap. It is well-known that a standard

Galerkin finite element discretization is not suitable for the convection-dominated case. We use the streamline diffusion finite element method (SDFEM). We remark that SDFEM ensures a higher order of accuracy (cf. [25, 36]) than a first order upwind finite difference method. In SDFEM a mesh-dependent anisotropic diffusion, which acts only in streamline direction, is added to the discrete problem. Such anisotropy is important for the high order of convergence of this method and also plays a crucial role in our convergence analysis of the multigrid method. In this paper we only treat the case of a uniform triangulation which is taken such that the streamlines are aligned with gridlines. Whether our analysis can be generalized to the situation of an unstructured triangulation is an open question.

We briefly discuss the different components of the multigrid solver.
For the *prolongation and restriction* we use the canonical inter-grid transfer operators that are induced by the nesting of the finite element spaces, i.e., for the prolongation we use linear interpolation and for the restriction operator we take the adjoint of the prolongation.
The hierarchy of *coarse grid discretization operators* is constructed by applying the SDFEM on each grid level. Note that due to the level-dependent stabilization term we have level-dependent bilinear forms and the Galerkin property $A_{k-1} = r_k A_k p_k$ does not hold.
Related to the *smoother* we note the following. First we emphasize that due to a certain crosswind smearing effect in the finite element discretization the x-line Jacobi or Gauss-Seidel methods do *not* yield robust smoothers (i.e., they do not result in a direct solver in the limit case $\varepsilon = 0$, cf. [7]). This is explained in more detail in remark 1 in section 6. In the present paper we use a smoother of x-line-Jacobi type which is not robust.
The only nonstandard component in our multigrid algorithm is a *local presolver*, which is applied before the coarse-grid correction. This step consists of solving a local subproblem of relatively small dimension on a subdomain adjacent to the inflow boundary. The arithmetic complexity of this presolver is less than the arithmetic complexity of a matrix-vector multiplication. The reason for using this presolver will be explained below.
These components are combined in a standard manner resulting in a multigrid W-cycle algorithm (cf. Algorithm 1 on page 17). In this method we use ν postsmoothing iterations (of x-line-Jacobi type) but instead of a presmoother we apply a local presolver.

We now outline the convergence analysis of the multigrid method. Let A_k denote the stiffness matrix on level k and $S_k = I - W_k^{-1} A_k$ the iteration matrix of the postsmoother. Furthermore, for the iteration matrix of the local presolver we use the notation Q_k. The two-grid iteration matrix is given by $T_k = S_k^\nu (I - p_k A_{k-1}^{-1} r_k A_k) Q_k$. After the local presolve operation the defect is zero on a small subdomain adjacent to the inflow boundary. This yields the relation $A_k Q_k = J_k A_k Q_k$, where J_k is a diagonal projection matrix with $(J_k)_{ii} \in \{0,1\}$ and $(J_k)_{ii} = 0$ iff the grid point corresponding to index i lies in this subdomain. We use the euclidean norm denoted by $\|\cdot\|$ and the level-dependent A-norm $\|x\|_A = \|A_k x\|$. For the A-norm contraction number of the two-grid method we have the bound

$$\|T_k\|_A \leq \|A_k S_k^\nu W_k^{-1}\| \, \|W_k C_k J_k\| \, \|Q_k\|_A \ , \quad C_k := A_k^{-1} - p_k A_{k-1}^{-1} r_k \qquad (1.3)$$

For the analysis we restrict ourselves to the convection-dominated case $\varepsilon \leq c h_k$ ($h_k = 2^{-k}$: mesh size on level k). We derive bounds for the three terms on the righthand

side in (1.3).
For the first term in (1.3) we prove $\|A_k S_k^\nu W_k^{-1}\| \leq c\,\nu^{-\frac{1}{2}}$. Here and in the remainder a constant c is always independent of k and ε. This result is closely related to the smoothing-property as introduced by Hackbusch [7, 8].
For the second term in (1.3) we prove uniform boundedness, $\|W_k C_k J_k\| \leq c$. This can be interpreted as a modified approximation property. Note that the approximation property known from the literature (and used for diffusion-dominated problems) is of the form $\|C_k\| \leq c\|A_k\|^{-1}$, whereas we consider the term $\|W_k C_k J_k\|$. Our approach is different in two respects. Firstly, in our analysis the preconditioner W_k, which comes from the smoother, plays an essential role in the approximation property. If we would use the splitting $\|W_k C_k J_k\| \leq \|W_k\|\|C_k J_k\| \leq c\|A_k\|\|C_k J_k\|$, then sharp bounds for the terms $\|A_k\|$ and $\|C_k J_k\|$ do not imply uniform boundedness of $\|W_k C_k J_k\|$. Only the combined effect of the preconditioner W_k and the operator $C_k J_k$ results in uniform boundedness. This is related to the fact that we do not have a robust smoother. Secondly, we use the operator J_k which comes from the local presolver and which projects errors to zero on a local subdomain adjacent to the inflow boundary. The main reason for using this operator (and thus the presolver) is the following. As is usually done in the analysis of the approximation property we use finite element error bounds combined with regularity results. In the derivation of a L^2 bound for the finite element discretization error we use a duality argument. However, the formal dual problem has poor regularity properties, since the inflow boundary of the original problem ($x = 0$) is the outflow boundary of the dual problem. Thus Dirichlet outflow boundary conditions would appear and we obtain poor estimates due to the poor regularity. To avoid this, we consider a dual problem with Neumann outflow and Dirichlet inflow conditions. To be able to deal with the inconsistency caused by these "wrong" boundary conditions we assume the righthand side to be zero near the boundary $x = 0$. In order to satisfy this assumption we use the local presolver in the multigrid algorithm which then yields the projection term J_k.
For the third term in (1.3) we obtain, unfortunately, a bound which is k-dependent: $\|Q_k\|_A \leq c\,k$. To compensate this, the number of smoothings has to be taken level-dependent: $\nu = \nu_k \sim k^2$. This then results in a two-grid method with a contraction number $\|T_k\|_A \leq c < 1$ and a complexity $\mathcal{O}(N_k(\ln N_k)^2)$, with $N_k = h_k^{-2}$.
Using standard arguments we obtain a similar convergence result for the multigrid W-cycle.

The remainder of this paper is organized as follows. In section 2 we give the weak formulation of the problem (1.2) and describe the SDFEM. In section 3 some useful properties of the stiffnes matrix are derived. In section 4 we prove some a priori estimates for the continuous and the discrete solution. In section 5 we derive quantitative results concerning the upstream influence of a righthand side on the solution. These results are needed in the proof of the modified approximation property. *Section 6 contains the main results of this paper*. In this section we describe the multigrid algorithm and present the convergence analysis. This analysis is based on four important auxiliary results (smoothing property, modified approximation property, bound for $\|Q_k\|_A$ and a stability result for the prolongation operator). These auxiliary results are formulated in section 6 and using these we derive a bound for the multigrid W-cycle contraction number. The proofs of the auxiliary results are given in the sections 7–10.

2. The continuous problem and its discretization. For the weak formulation of the problem (1.2) we use the $\mathrm{L}^2(\Omega)$ scalar product which is denoted by

$(\cdot,\cdot)$. For the corresponding norm we use the notation $\|\cdot\|$. With the Sobolev space $\mathbf{V} := \{ v \in \mathrm{H}^1(\Omega) \mid v = 0 \ \text{ on } \ \partial\Omega \setminus \Gamma_E \}$ the weak formulation is as follows: find $u \in \mathbf{V}$ such that

$$a(u,v) := \varepsilon(u_x,v_x) + \varepsilon(u_y,v_y) + (u_x,v) = (f,v) \quad \text{for all} \quad v \in \mathbf{V} \tag{2.1}$$

From the Lax-Milgram lemma it follows that a unique solution of this problem exists. For the discretization we use linear finite elements on a uniform triangulation. For this we use a mesh size $h_k := 2^{-k}$ and grid points $x_{i,j} = (ih_k, jh_k)$, $0 \le i,j \le h_k^{-1}$. A uniform triangulation is obtained by inserting diagonals that are oriented from south-west to north-east. Let $\mathbb{V}_k \subset \mathbf{V}$ be the space of continuous functions that are piecewise linear on this triangulation and have zero values on $\partial\Omega \setminus \Gamma_E$. For the discretization of (2.1) we consider the streamline-diffusion finite element method: find $u_k \in \mathbb{V}_k$ satisfying

$$(\varepsilon+\delta_k h_k)((u_k)_x,v_x)+\varepsilon((u_k)_y,v_y)+((u_k)_x,v) = (f,v+\delta_k h_k v_x) \quad \text{for all} \quad v \in \mathbb{V}_k \tag{2.2}$$

with

$$\delta_k = \begin{cases} \bar{\delta} & \text{if } \frac{h_k}{2\,\varepsilon} \ge 1 \\ 0 & \text{otherwise} \end{cases} \tag{2.3}$$

The stabilization parameter $\bar{\delta}$ is a given constant of order 1. For an analysis of the streamline diffusion finite element method we refer to [25, 6].

In this paper we assume

$$\bar{\delta} \in [\frac{1}{3}, 1] \ . \tag{2.4}$$

The value $\frac{1}{3}$ for the lower bound is important for our analysis. The choice of 1 for the upper bound is made for technical reasons and this value is rather arbitrary.
The finite element formulation (2.2) gives rise to the (stabilized) bilinear form

$$a_k(u,v) := (\varepsilon + \delta_k h_k)(u_x,v_x) + \varepsilon(u_y,v_y) + (u_x,v), \quad u,v \in \mathbf{V} \tag{2.5}$$

Note the following relation for the bilinear form $a_k(\cdot,\cdot)$:

$$a_k(v,v) = \varepsilon\|v_y\|^2 + (\varepsilon + \delta_k h_k)\|v_x\|^2 + \frac{1}{2}\int_{\Gamma_E} v^2\,dy \quad \text{for} \quad v \in \mathbf{V}. \tag{2.6}$$

The main topic of this paper is a convergence analysis of a multigrid solver for the algebraic system of equations that corresponds to (2.2). In this convergence analysis the particular form of the righthand side in (2.2), which is essential for consistency in the streamline diffusion finite element method, does not play a role. Therefore we will consider the discrete problem, with a righthand side induced by an arbitrary $f \in \mathrm{L}^2(\Omega)$: find $u_k \in \mathbb{V}_k$ such that

$$a_k(u_k,v_k) = (f,v_k) \quad \text{for all} \quad v_k \in \mathbb{V}_k \tag{2.7}$$

we will also use the following auxiliary continuous problem: find $u \in \mathbf{V}$ such that

$$a_k(u,v) = (f,v) \quad \text{for all} \quad v \in \mathbf{V}. \tag{2.8}$$

Note that u and u_k depend on the stabilization term in the bilinear form and that these solutions differ from those in (2.1) and (2.2).

3. Representation of the stiffness matrix. We now derive a representation of the stiffness matrix corresponding to the bilinear form $a_k(\cdot,\cdot)$ that will be used in the analysis below. The standard nodal basis in $\mathbb{V}_k$ is denoted by $\{\phi_\ell\}_{1\le\ell\le N_k}$ with N_k the dimension of the finite element space, $N_k := h_k^{-1}(h_k^{-1}-1)$. Define the isomorphism:

$$P_k : X_k := \mathbb{R}^{N_k} \to \mathbb{V}_k, \quad P_k x = \sum_{i=1}^{N_k} x_i \phi_i .$$

On X_k we use a scaled Euclidean scalar product: $\langle x, y\rangle_k = h_k^2 \sum_{i=1}^{N_k} x_i y_i$ and corresponding norm denoted by $\|\cdot\|$ (note that this notation is also used to denote the $\mathrm{L}^2(\Omega)$ norm). The adjoint $P_k^* : \mathbb{V}_k \to X_k$ satisfies $(P_k x, v) = \langle x, P_k^* v\rangle_k$ for all $x \in X_k,\ v \in \mathbb{V}_k$. The following norm equivalence holds

$$C^{-1}\|x\| \le \|P_k x\| \le C\|x\| \quad \text{for all } x \in \mathrm{X}_k, \tag{3.1}$$

with a constant C independent of k. The stiffness matrix A_k on level k is defined by

$$\langle A_k x, y\rangle_k = a_k(P_k x, P_k y) \quad \text{for all } x, y \in \mathrm{X}_k. \tag{3.2}$$

In an interior grid point the discrete problem has the stencil

$$\frac{1}{h_k^2}\begin{bmatrix} 0 & -\varepsilon & 0 \\ -\varepsilon_k & 2(\varepsilon_k+\varepsilon) & -\varepsilon_k \\ 0 & -\varepsilon & 0 \end{bmatrix} + \frac{1}{h_k}\begin{bmatrix} 0 & -\frac{1}{6} & \frac{1}{6} \\ -\frac{1}{3} & 0 & \frac{1}{3} \\ -\frac{1}{6} & \frac{1}{6} & 0 \end{bmatrix}, \qquad \varepsilon_k := \varepsilon + \delta_k h_k . \tag{3.3}$$

For a matrix representation of the discrete operator we first introduce some notation. Let $n_k := h_k^{-1}$ and

$$D_x := \frac{1}{h_k}\mathrm{tridiag}(-1,1,0) = \frac{1}{h_k}\begin{pmatrix} 1 & & & \\ -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{pmatrix} \in \mathbb{R}^{n_k \times n_k} ,$$

$$A_x := D_x^T D_x = \frac{1}{h_k^2}\begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 1 \end{pmatrix} \in \mathbb{R}^{n_k \times n_k} ,$$

$$A_y := \frac{1}{h_k^2}\mathrm{tridiag}(-1,2,-1) \in \mathbb{R}^{(n_k-1)\times(n_k-1)} ,$$

$$S_x := \frac{1}{h_k}\begin{pmatrix} -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \\ & & & 0 \end{pmatrix} \in \mathbb{R}^{n_k \times n_k} ,$$

$$J := \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & \frac{1}{2} \end{pmatrix} \in \mathbb{R}^{n_k \times n_k},$$

$$T := \mathrm{tridiag}(0,0,1) \in \mathbb{R}^{n_k \times n_k} .$$

Furthermore, let I_k be the $k \times k$ identity matrix. We finally introduce the following $N_k \times N_k$ matrices ($N_k = (n_k - 1)n_k$)

$$\hat{D}_x := I_{n_k-1} \otimes D_x \ , \quad \hat{A}_x := I_{n_k-1} \otimes A_x = \hat{D}_x^T \hat{D}_x \ , \quad \hat{A}_y := A_y \otimes J$$

and the $N_k \times N_k$ blocktridiagonal matrix

$$\hat{B} := \text{blocktridiag}(I_{n_k}, 4I_{n_k}, T) \ .$$

Using all this notation we obtain the following representation for the stiffness matrix A_k in (3.2):

$$\begin{aligned} A_k &= (\varepsilon + \delta_k h_k)\hat{A}_x + \varepsilon \hat{A}_y + \frac{1}{6}\text{blocktridiag}(D_x, -2h_k A_x + 4D_x, S_x) \\ &= (\varepsilon + (\delta_k - \frac{1}{3})h_k)\hat{A}_x + \varepsilon \hat{A}_y + \frac{1}{6}\text{blocktridiag}(D_x, 4D_x, S_x) \\ &= (\varepsilon + (\delta_k - \frac{1}{3})h_k)\hat{A}_x + \varepsilon \hat{A}_y + \frac{1}{6}\hat{B}\hat{D}_x \end{aligned} \tag{3.4}$$

The latter decomposition can be written in stencil notation as

$$\frac{\bar{\varepsilon}_k}{h_k^2}\begin{bmatrix} 0 & 0 & 0 \\ -1 & 2 & -1 \\ 0 & 0 & 0 \end{bmatrix} + \frac{\varepsilon}{h_k^2}\begin{bmatrix} 0 & -1 & 0 \\ 0 & 2 & 0 \\ 0 & -1 & 0 \end{bmatrix} + \frac{1}{6h_k}\begin{bmatrix} 0 & -1 & 1 \\ -4 & 4 & 0 \\ -1 & 1 & 0 \end{bmatrix} \tag{3.5}$$

with $\bar{\varepsilon}_k = \varepsilon + (\delta_k - \frac{1}{3})h_k > 0$.

Some properties of the matrices used in the decomposition (3.4) are collected in the following lemma.

For $B, C \in \mathbb{R}^{n\times n}$ we write $B \geq C$ iff $x^T B x \geq x^T C x$ for all $x \in \mathbb{R}^n$.

LEMMA 3.1. *The following inequalities hold*

$$\hat{A}_x \hat{D}_x^{-1} \geq 0 \tag{3.6}$$

$$\hat{A}_y \hat{D}_x^{-1} \geq 0 \tag{3.7}$$

$$\hat{B} \geq 2I \tag{3.8}$$

$$A_k \hat{D}_x^{-1} \geq \frac{1}{3}I \tag{3.9}$$

$$\|\hat{D}_x A_k^{-1}\| \leq 3 \tag{3.10}$$

Proof. To check (3.6) observe

$$\hat{A}_x \hat{D}_x^{-1} = \hat{D}_x^T \hat{D}_x \hat{D}_x^{-1} = \hat{D}_x^T.$$

Now note that $\hat{D}_x^T + \hat{D}_x$ is symmetric positive definite.

To prove (3.7) it suffices to show that $\hat{D}_x^T \hat{A}_y \geq 0$ holds. We have

$$K := \hat{D}_x^T \hat{A}_y = (I_{n_k-1} \otimes D_x^T)(A_y \otimes J) = A_y \otimes \tilde{D}_x^T \ ,$$

with the matrix

$$\tilde{D}_x^T = \frac{1}{h_k}\begin{pmatrix} 1 & -1 & & \\ & \ddots & \ddots & \\ & & 1 & -\frac{1}{2} \\ & & & \frac{1}{2} \end{pmatrix}$$

Hence in the matrix $K + K^T = 2A_y \otimes (\tilde{D}_x^T + \tilde{D}_x)$ both factors A_y and $\tilde{D}_x^T + \tilde{D}_x$ are symmetric positive definite. From this the result follows.

To prove (3.8) we define $R := \hat{B} - 4I$ and note that $\|R\|^2 \le \|R\|_\infty \|R\|_1 \le 4$. Using this we get

$$\langle \hat{B}u, u \rangle_k = 4\|u\|^2 + \langle Ru, u \rangle_k \ge 4\|u\|^2 - \|R\| \|u\|^2 \ge 2\|u\|^2$$

which proves the desired result.

Inequality (3.9) follows immediately from the representation of A_k in (3.4) and inequalities (3.6)–(3.8).

From the result in (3.9) it follows that $\hat{D}_x^T A_k \ge \frac{1}{3} \hat{D}_x^T \hat{D}_x$. This implies $\|\hat{D}_x z\|^2 \le 3\langle A_k z, \hat{D}_x z \rangle_k \le 3\|A_k z\| \|\hat{D}_x z\|$ for all z and thus estimate (3.10) is also proved.
□

4. A priori estimates. In this paper we study the convergence of a multigrid method for solving the system of equations

$$A_k x_k = b \,, \tag{4.1}$$

with A_k the stiffness matrix from the previous section. As already noted in the introduction our analysis relies on smoothing and approximation properties. For establishing a suitable approximation property we will use regularity results and a priori estimates for solutions of the continuous and the discrete problems. Such results are collected in this section.

In the remainder of the paper we restrict ourselves to the convection-dominated case:

ASSUMPTION 1. *We only consider values of k and ε such that*

$$\varepsilon \le \frac{1}{2} h_k$$

If instead of the factor $\frac{1}{2}$ in this assumption we take another constant C, our analysis can still be applied but some technical modifications are needed (to distinguish between $\delta_k = \bar{\delta}$ and $\delta_k = 0$) which make the presentation less transparent.
We consider this convection-dominated case to be the most interesting one. Many results that will be presented also hold for the case of an arbitrary positive ε but the proofs for the diffusion-dominated case often differ from those for the convection-dominated case. In view of the presentation we decided to treat only the convection-dominated case. Note that then

$$\delta_k = \bar{\delta} \in [\frac{1}{3}, 1] \;\; \text{and} \;\; \frac{1}{3} h_k \le \varepsilon_k = \varepsilon + \bar{\delta} h_k \le \frac{3}{2} h_k \tag{4.2}$$

hold.

For the inflow boundary (i.e., the one on the west side) we use the notation $\Gamma_W := \{ (x,y) \in \overline{\Omega} \mid x = 0 \}$. For the continuous solution u the following a-priori estimates hold

THEOREM 4.1. *For $f \in L_2(\Omega)$ let u be the solution of (2.8). There is a constant c independent of k and ε such that:*

$$\|u\| + \|u_x\| \le c\|f\|, \tag{4.3}$$

$$\sqrt{\varepsilon}\|u_y\| \le c\|f\|, \tag{4.4}$$

$$h_k\|u_{xx}\| + \sqrt{\varepsilon h_k}\|u_{xy}\| + \varepsilon\|u_{yy}\| \le c\|f\|, \tag{4.5}$$

$$\int_{\Gamma_E} u^2 \, dy + h_k \int_{\Gamma_W} u_x^2 \, dy + \varepsilon \int_{\Gamma_E} u_y^2 \, dy \le c\|f\|^2. \tag{4.6}$$

Proof. Since $f \in L_2(\Omega)$, the regularity theory from [5] ensures that the solution u of (2.8) belongs to $\mathrm{H}^2(\Omega)$. Hence we can consider the strong formulation of (2.8)

$$-\varepsilon u_{yy} - \varepsilon_k u_{xx} + u_x = f \tag{4.7}$$

with boundary conditions as in (1.2). Now we multiply (4.7) with u_x and integrate by parts. Taking boundary conditions into account, we get the following terms:

$$\begin{aligned}
-\varepsilon(u_{yy}, u_x) &= \frac{\varepsilon}{2}((u_y^2)_x, 1) = \frac{\varepsilon}{2}\int_{\Gamma_E} u_y^2\, dy, \\
-\varepsilon_k(u_{xx}, u_x) &= \frac{\varepsilon_k}{2}((u_x^2)_x, 1) = \frac{\varepsilon_k}{2}\int_{\Gamma_W} u_x^2\, dy \geq c\, h_k \int_{\Gamma_W} u_x^2\, dy, \quad \text{(we use (4.2))} \\
(u_x, u_x) &= \|u_x\|^2 \geq \|u\|^2, \\
(f, u_x) &\leq \frac{1}{2}\|f\|^2 + \frac{1}{2}\|u_x\|^2.
\end{aligned}$$

From these relations the results (4.3) and (4.6), except the bound for $\int_{\Gamma_E} u^2\, dy$, easily follow. Next we multiply (4.7) with u and integrate by parts to obtain

$$\varepsilon\|u_y\|^2 + \varepsilon_k\|u_x\|^2 + \frac{1}{2}\int_{\Gamma_E} u^2\, dy = (f, u) \leq \|f\|\|u\| \leq c\,\|f\|^2. \quad \text{(we use (4.3))}$$

Estimate (4.4) and the remainder of (4.6) now follow.
To prove (4.5) we introduce $F = f - u_x$. Due to (4.3) we have $\|F\| \leq c\,\|f\|$. Moreover

$$-\varepsilon u_{yy} - \varepsilon_k u_{xx} = F.$$

If we square both sides of this equality and integrate over Ω we obtain

$$\varepsilon^2\|u_{yy}\|^2 + 2\varepsilon\varepsilon_k(u_{yy}, u_{xx}) + \varepsilon_k^2\|u_{xx}\|^2 = \|F\|^2 \leq c\,\|f\|^2. \tag{4.8}$$

Further note that for any sufficiently smooth function v, satisfying boundary conditions from (1.2), the relations

$$v_{xx}(x,0) = v_{xx}(x,1) = 0,\ x \in (0,1), \quad v_y(0,y) = v_{xy}(1,y) = 0,\ y \in (0,1),$$

hold, and thus

$$(v_{yy}, v_{xx}) = -(v_y, v_{xxy}) = (v_{xy}, v_{xy}).$$

Using a standard density argument we conclude that for $u \in \mathrm{H}^2(\Omega)$ the relation $(u_{yy}, u_{xx}) = (u_{xy}, u_{xy})$ holds. Now (4.8) gives

$$\varepsilon^2\|u_{yy}\|^2 + 2\varepsilon\varepsilon_k\|u_{xy}\|^2 + \varepsilon_k^2\|u_{xx}\|^2 \leq c\,\|f\|^2.$$

In combination with (4.2) this yields (4.5).
□

The next lemma states that the x-derivative of the *discrete* solution is also uniformly bounded if the righthand side is from $\mathbb{V}_k$.

LEMMA 4.2. *For $f_k \in \mathbb{V}_k$ let $u_k \in \mathbb{V}_k$ be a solution to (2.7), then*

$$\|(u_k)_x\| \leq c\,\|f_k\|. \tag{4.9}$$

Proof. The result in (4.9) follows from the estimate (3.10) in lemma 3.1. To show this we need some technical considerations.

First we show how the size of the x-derivative of a finite element function $v \in \mathbb{V}_k$ can be determined from its corresponding coefficent vector $P_k^{-1}v \in X_k$. Let $\mathcal{I}$ be the index set $\{(i,j) \mid 0 \le i \le n_k - 1, \;\; 1 \le j \le n_k - 1\}$. For $(i,j) \in \mathcal{I}$ let $T^l_{(i,j)}$ and $T^u_{(i,j)}$ be the two triangles in the triangulation which have the line between the grid points $x_{i,j}$ and $x_{i+1,j}$ as a common edge. Let $v \in \mathbb{V}_k$ be given. For $1 \le j \le n_k - 1$ we introduce the vector $\mathbf{v}_j = (v(x_{1,j}), \dots, v(x_{n_k,j}))^T$. We then obtain

$$\begin{aligned}
\|v_x\|^2 &= \sum_{(i,j)\in\mathcal{I}} \Big(\int_{T^l_{(i,j)}} v_x^2 \, dxdy + \int_{T^u_{(i,j)}} v_x^2 \, dxdy \Big) \\
&= \sum_{(i,j)\in\mathcal{I}} \Big(\frac{v(x_{i+1,j}) - v(x_{i,j})}{h_k} \Big)^2 h_k^2 \\
&= h_k^2 \sum_{1 \le j \le n_k - 1} (D_x \mathbf{v}_j)^T (D_x \mathbf{v}_j) \\
&= h_k^2 \big(\hat{D}_x P_k^{-1} v\big)^T \big(\hat{D}_x P_k^{-1} v\big) = \|\hat{D}_x P_k^{-1} v\|^2.
\end{aligned}$$

Therefore

$$\|v_x\| = \|\hat{D}_x P_k^{-1} v\| \quad \text{for any} \;\; v \in \mathbb{V}_k. \tag{4.10}$$

For the discrete solution of (2.7) with $f = f_k$ we have the representation $u_k = P_k A_k^{-1} P_k^* f_k$. Now from (3.10) and (4.10) it follows that

$$\|(u_k)_x\| = \|\hat{D}_x A_k^{-1} P_k^* f_k\| \le 3\,\|P_k^* f_k\| \le c\,\|f_k\|$$

with a constant c independent of k and ε. □

The next lemma gives some bounds on the difference between discrete and continuous solutions

LEMMA 4.3. *Define the error $e_k = u - u_k$, where u and u_k are solutions of the problems (2.8) and (2.7) with righthand side $f = f_k \in \mathbb{V}_k$. Then the following estimates hold*

$$\|(e_k)_x\| \le c\|f_k\| \tag{4.11}$$

$$\varepsilon\|(e_k)_y\|^2 + \frac{1}{2}\int_{\Gamma_E} e_k^2 \, dy \le c\frac{h_k^2}{\varepsilon}\|f_k\|^2. \tag{4.12}$$

Proof. Estimate (4.11) directly follows from (4.3) and (4.9) by a triangle inequality. The proof of (4.12) is based on standard arguments: the Galerkin orthogonality, approximation properties of $\mathbb{V}_k$ and a priori estimates from (4.5). Indeed

$$\begin{aligned}
&\varepsilon\|(e_k)_y\|^2 + (\varepsilon + \bar{\delta} h_k)\|(e_k)_x\|^2 + \frac{1}{2}\int_{\Gamma_E} e_k^2 \, dy = a_k(e_k, e_k) = \inf_{v_k \in \mathbb{V}_k} a_k(e_k, u - v_k) \\
&\le \varepsilon\|(e_k)_y\|\|(u - v_k)_y\| + (\varepsilon + \bar{\delta} h_k)\|(e_k)_x\|\|(u - v_k)_x\| + \|(e_k)_x\|\|u - v_k\| \\
&\le c\,\big(\varepsilon\, h_k\|(e_k)_y\|\|u\|_{H^2} + h_k^2\|(e_k)_x\|\|u\|_{H^2}\big) \\
&\le c\,\big(h_k\|(e_k)_y\|\|f_k\| + \frac{h_k^2}{\varepsilon}\|f_k\|^2\big) \le \frac{\varepsilon}{2}\|(e_k)_y\|^2 + c\frac{h_k^2}{\varepsilon}\|f_k\|^2.
\end{aligned}$$

The estimate (4.12) follows.

□

5. Upstream influence of the streamline diffusion method. Consider the continuous problem (2.8). The goal of this section is to estimate the upstream influence of the righthand side function f on the solution u. The same will be done for the corresponding discrete problem. In the literature results of such type are known for the problem with Dirichlet boundary conditions and typically formulated in the form of estimates on the (discrete) Greens function (see, e.g., [29, 16]). A typical result is that the value of the solution at a point x is essentially determined by the values of the righthand side in a "small" strip that contains x. This strip has a crosswind width of size $O(\varepsilon^*|\ln h|)$, where $\varepsilon^* = \max\{\varepsilon, h^{\frac{3}{2}}\}$, and in the streamline direction it ranges from the inflow boundary to a $\mathcal{O}(h|\ln h|)$ upstream distance from x. A proof of such estimates is rather technical and involves the formal adjoint problem. We need similar estimates regarding the upstream influence for the problem with Neumann outflow boundary conditions. Below we present elementary proofs of the results that are needed for the multigrid convergence analysis further on. Our analysis avoids the use of an adjoint problem and is based on the following lemma.

LEMMA 5.1. *For $\varepsilon_k = \varepsilon + \bar{\delta} h_k$ assume a function $\phi \in H^1_\infty(0,1)$, such that $0 \le -\varepsilon_k \phi_x \le \phi$. Denote by $\|\cdot\|_\phi$ a semi-norm induced by the scalar product $(\phi\cdot,\cdot)$. Then the solution u of (2.8) satisfies*

$$\|u_x\|_\phi \le 2\|f\|_\phi \tag{5.1}$$

$$\varepsilon_k\ \phi(0)\int_{\Gamma_W} u_x^2\,dy \le \|f\|_\phi^2 \tag{5.2}$$

$$\frac{1}{4}\|u\|^2_{-\phi_x} + \varepsilon\|u_y\|^2_\phi \le (\phi f, u). \tag{5.3}$$

Proof. We consider the strong formulation (4.7) and multiply it with ϕu_x and integrate by parts. We then get the following terms

$$\begin{aligned}
-\varepsilon(u_{yy},\phi u_x) &= \frac{\varepsilon}{2}\|u_y\|^2_{-\phi_x} + \frac{\varepsilon}{2}\phi(1)\int_{\Gamma_E} u_y^2\,dy \ge 0\ ,\\
-\varepsilon_k(u_{xx},\phi u_x) &= -\frac{\varepsilon_k}{2}\|u_x\|^2_{-\phi_x} + \frac{\varepsilon_k}{2}\phi(0)\int_{\Gamma_W} u_x^2\,dy\\
&\ge -\frac{1}{2}\|u_x\|^2_\phi + \frac{\varepsilon_k}{2}\phi(0)\int_{\Gamma_W} u_x^2\,dy\ ,\\
(u_x,\phi u_x) &= \|u_x\|^2_\phi\ ,\\
(f,\phi u_x) &\le \|f\|_\phi\|u_x\|_\phi \le \|f\|^2_\phi + \frac{1}{4}\|u_x\|^2_\phi.
\end{aligned}$$

Now (5.1) and (5.2) immediately follow.
To obtain the estimate (5.3) we multiply (4.7) with $\phi\, u$ and integrate by parts. We get the following terms:

$$\begin{aligned}
-\varepsilon(u_{yy},\phi\, u) &= \varepsilon\|u_y\|^2_\phi,\\
-\varepsilon_k(u_{xx},\phi\, u) &= \varepsilon_k\|u_x\|^2_\phi + \varepsilon_k(u_x,\phi_x\, u) \ge \varepsilon_k\|u_x\|^2_\phi - \varepsilon_k^2\|u_x\|^2_{-\phi_x} - \frac{1}{4}\|u\|^2_{-\phi_x}\\
&\ge -\frac{1}{4}\|u\|^2_{-\phi_x}.\\
(u_x,\phi\, u) &= \frac{1}{2}\|u\|^2_{-\phi_x} + \frac{\phi(1)}{2}\int_{\Gamma_E} u^2\,dy.
\end{aligned}$$

Thus (5.3) follows. □

For arbitrary $\xi \in [0,1]$ consider the function

$$\phi_\xi(x) = \begin{cases} 1 & \text{for } x \in [0,\xi], \\ \exp\left(-\frac{x-\xi}{\varepsilon_k}\right) & \text{for } x \in (\xi,1]. \end{cases}$$

For any ξ the function $\phi_\xi(x)$ satisfies the assumptions of lemma 5.1. Define the domains

$$\begin{aligned} \Omega_\xi &= \{(x,y) \in \Omega :\, x < \xi\}\,, \\ \Omega_\eta &= \{(x,y) \in \Omega :\, x > \eta\}\,, \quad \eta \in (0,1). \end{aligned} \tag{5.4}$$

Direct application of lemma 5.1 with $\phi = \phi_\xi$ gives the following corollary.

COROLLARY 5.2. *Consider $f \in L_2(\Omega)$ such that* $\mathrm{supp}(f) \in \Omega_\eta$ *and let u be the corresponding solution of problem (2.8). Assume $\eta - \xi \geq 2\,\varepsilon_k\, p\,|\ln h_k|$, $p > 0$. Then we have*

$$\|u_x\|_{L_2(\Omega_\xi)} \leq h_k^p \|f\|, \tag{5.5}$$

$$\varepsilon_k \int_{\Gamma_W} u_x^2\, dy \leq h_k^{2p}\|f\|^2, \tag{5.6}$$

$$\sqrt{\varepsilon}\|u_y\|_{L_2(\Omega_\xi)} \leq \sqrt{\varepsilon_k}\, h_k^p\|f\|. \tag{5.7}$$

Proof. The estimate

$$\|f\|_\phi^2 = (\phi f, f)_{\Omega_\eta} \leq \phi(\eta)\|f\|_{\Omega_\eta}^2 = h_k^{2p}\|f\|^2$$

and (5.1), (5.2) imply the results (5.5) and (5.6). We also have

$$\begin{aligned} (\phi f, u) = (\phi f, u)_{\Omega_\eta} &\leq \varepsilon_k \|f\|_\phi^2 + \frac{1}{4\varepsilon_k}(\phi u, u)_{\Omega_\eta} = \varepsilon_k\|f\|_\phi^2 + \frac{1}{4}(-\phi_x u, u)_{\Omega_\eta} \\ &\leq \varepsilon_k\|f\|_\phi^2 + \frac{1}{4}\|u\|_{-\phi_x}^2. \end{aligned}$$

Together with (5.3) this yields (5.7).
□

We need an analogue of estimate (5.1) for the finite element solution u_k of (2.7). To this end consider a vector $\phi = (\phi_0, \ldots, \phi_{n_k})$, such that $\phi_i > 0$ for all i and

$$c_0\phi_i \geq -\bar{\varepsilon}_k \frac{\phi_i - \phi_{i-1}}{h_k} \geq 0, \quad i = 1, \ldots, n_k \tag{5.8}$$

with a constant $c_0 \in (0, \frac{4}{9})$ and $\bar{\varepsilon}_k = \varepsilon + (\bar{\delta} - \frac{1}{3})h_k$.

LEMMA 5.3. *For the solution u_k of the problem (2.7) with $f = f_k \in \mathbb{V}_k$ the estimate*

$$\sum_{i=1}^{n_k}\sum_{j=1}^{n_k-1} h_k^2 \phi_i \left(\frac{u_{i,j} - u_{i-1,j}}{h_k}\right)^2 \leq C \sum_{i=1}^{n_k}\sum_{j=1}^{n_k-1} h_k^2\phi_i (M_k\hat{f})_{i,j}^2 \tag{5.9}$$

holds. Here u_{ij} is the nodal value of u_k at the grid point $x_{i,j}$, $\hat{f}$ is the vector of nodal values of f_k and M_k is the mass matrix.

Proof. Let $\hat{u}_k = P_k^{-1} u_k \in X_k$ be the vector of nodal values of u_k, then

$$A_k \hat{u}_k = M_k \hat{f} =: b_k. \tag{5.10}$$

We introduce the diagonal matrices $\Phi = \mathrm{diag}(\phi_i)_{1\le i\le n_k}$ and $\hat{\Phi} := I_{n_k-1} \otimes \Phi$. The statement of the lemma is equivalent to

$$\langle \hat{\Phi}\hat{D}_x \hat{u}_k, \hat{D}_x \hat{u}_k\rangle_k \;\le\; c\,\langle \hat{\Phi} b_k, b_k\rangle_k$$

with a constant c that is independent of b_k. This is the same as

$$\|\hat{D}_x A_k^{-1}\|_\phi \le c \tag{5.11}$$

where the norm $\|\cdot\|_\phi$ is induced by the scalar product $\langle \hat{\Phi}\cdot,\cdot\rangle_k =: \langle\cdot,\cdot\rangle_\phi$ on X_k. Note that (5.11) is a generalization of the result in (3.10). Here we use similar arguments as in the proof of (3.10). We use the representation (3.4) of the stiffness matrix:

$$A_k = \bar{\varepsilon}_k \hat{A}_x + \varepsilon \hat{A}_y + \frac{1}{6}\hat{B}\hat{D}_x \;.$$

Note that

$$\hat{D}_x^T \hat{\Phi} \hat{A}_y = (I_{n_k-1}\otimes D_x^T)(I_{n_k-1}\otimes \Phi)(A_y \otimes J) = A_y \otimes D_x^T \Phi J.$$

The matix A_y is symmetric positive definite. Using $\phi_i \le \phi_{i-1}$ and a Gershgorin theorem it follows that $D_x^T\Phi J + J\Phi D_x$ is symmetric positive definite, too. Hence, $\hat{D}_x^T \hat{\Phi}\hat{A}_y \ge 0$ holds, i.e.,

$$\langle \hat{A}_y z, \hat{D}_x z\rangle_\phi \ge 0 \quad \text{for all} \;\; z \in X_k. \tag{5.12}$$

From the assumption on ϕ it follows that $\phi_{i-1} \le (1 + \frac{c_0 h_k}{\bar{\varepsilon}_k})\phi_i$ for all i. Using this and the relation

$$\frac{1}{2}(\Phi^{\frac{1}{2}} D_x^T \Phi^{-\frac{1}{2}} + \Phi^{-\frac{1}{2}} D_x \Phi^{\frac{1}{2}}) = \frac{1}{2h_k}\mathrm{tridiag}\Big(\sqrt{\frac{\phi_{i-1}}{\phi_i}}\,,\,2\,,\,\sqrt{\frac{\phi_i}{\phi_{i+1}}}\Big)$$

it follows that

$$\hat{\Phi}^{\frac{1}{2}} \hat{D}_x^T \hat{\Phi}^{-\frac{1}{2}} \ge \frac{1}{2h_k}\Big(2 - 2\sqrt{1 + \frac{c_0 h_k}{\bar{\varepsilon}_k}}\Big)\, I \;\ge\; -\frac{c_0}{2\bar{\varepsilon}_k} I$$

holds. And thus

$$\bar{\varepsilon}_k \langle \hat{A}_x z, \hat{D}_x z\rangle_\phi = \bar{\varepsilon}_k \langle \hat{\Phi}\hat{D}_x^T \hat{D}_x z, \hat{D}_x z\rangle \ge -\frac{1}{2} c_0 \langle \hat{D}_x z, \hat{D}_x z\rangle_\phi \quad \text{for all} \;\; z \in X_k. \tag{5.13}$$

We decompose $\hat{B}$ as $\hat{B} = 4I - R$. A simple computation yields

$$\|\hat{\Phi}^{\frac{1}{2}} R \hat{\Phi}^{-\frac{1}{2}}\|_1 \le 1 + \sqrt{1 + \frac{c_0 h_k}{\bar{\varepsilon}_k}} \le 1 + \sqrt{1 + 3c_0} \le 2 + \frac{3}{2}c_0.$$

Similarly we get $\|\hat{\Phi}^{\frac{1}{2}} R \hat{\Phi}^{-\frac{1}{2}}\|_\infty \le 2 + \frac{3}{2}c_0$ and thus $\|\hat{\Phi}^{\frac{1}{2}} R \hat{\Phi}^{-\frac{1}{2}}\| \le 2 + \frac{3}{2}c_0$. From this we obtain

$$\hat{\Phi}^{\frac{1}{2}} \hat{B} \hat{\Phi}^{-\frac{1}{2}} \ge \big(4 - (2 + \frac{3}{2}c_0)\big) I = (2 - \frac{3}{2}c_0) I$$

and thus

$$\frac{1}{6}\langle \hat{B}\hat{D}_x z, \hat{D}_x z\rangle_\phi \geq (\frac{1}{3}-\frac{1}{4}c_0)\langle \hat{D}_x z, \hat{D}_x z\rangle_\phi \quad \text{for all} \quad z \in X_k \tag{5.14}$$

Combination of the results in (5.12), (5.13) and (5.14) yields

$$\langle A_k z, \hat{D}_x z\rangle_\phi \geq (\frac{1}{3}-\frac{3}{4}c_0)\langle \hat{D}_x z, \hat{D}_x z\rangle_\phi \geq c\langle \hat{D}_x z, \hat{D}_x z\rangle_\phi \quad \text{for all} \quad z \in X_k$$

with a constant $c > 0$ (use that $c_0 \in (0, \frac{4}{9})$). From this we have

$$\|\hat{D}_x z\|_\phi^2 < \frac{1}{c}\langle A_k z, \hat{D}_x z\rangle_\phi \leq \frac{1}{c}\|A_k z\|_\phi \|\hat{D}_x z\|_\phi \quad \text{for all} \quad z \in X_k,$$

and thus $\|\hat{D}_x z\|_\phi \leq \tilde{c}\|A_k z\|_\phi$ for all z. Hence we have proved the result in (5.11).
□

In the discrete case we consider the vector

$$\phi_i^\xi = \begin{cases} 1 & \text{for } ih_k \in [0,\xi], \\ \exp\left(-\frac{ih_k-\xi}{4h_k}\right) & \text{for } ih_k \in (\xi,1]. \end{cases} \quad \text{for } i = 0,\ldots,n_k$$

It is straightforward to check

$$-(\phi_i^\xi - \phi_{i-1}^\xi) = (\exp(\frac{1}{4}) - 1)\,\phi_i^\xi.$$

Therefore, using $\bar{\varepsilon}_k \leq \frac{3}{2}h_k$,

$$0 \leq -\bar{\varepsilon}_k \frac{\phi_i^\xi - \phi_{i-1}^\xi}{h_k} \leq \frac{3}{2}(\exp(\frac{1}{4})-1)\phi_i^\xi, \quad i = 1,\ldots,n_k.$$

For any ξ the vector ϕ_i^ξ satisfies the assumptions of lemma 5.3 with $c_0 = \frac{3}{2}(\exp(\frac{1}{4})-1)$. This constant is less than $\frac{4}{9}$. We get the following corollary from lemma 5.3.

COROLLARY 5.4. *Consider* $f_k \in \mathbb{V}_k$ *such that* $\operatorname{supp}(f_k) \in \Omega_\eta$ *and let* u_k *be a the corresponding solution of problem (2.7). Assume* $\eta - \xi \geq 8\,h_k\,p\,|\ln h_k|,\ p > 0$, *then*

$$\|(u_k)_x\|_{L_2(\Omega_\xi)} \leq c\,h_k^p \|f_k\|, \tag{5.15}$$

$$\|(u_k)_y\|_{L_2(\Omega_\xi)} \leq c\,\xi\,h_k^{p-1}\|f_k\|. \tag{5.16}$$

Proof. Estimate (5.15) is a consequence of (5.9). Indeed, observe the following inequalities:

$$\begin{aligned}\|(u_k)_x\|_{L_2(\Omega_\xi)} &\leq c \sum_{i:\,ih\leq\xi}\sum_{j=1}^{N-1} h_k^2 \left(\frac{u_{i,j}-u_{i-1,j}}{h_k}\right)^2 \\ &= c \sum_{i:\,ih\leq\xi}\sum_{j=1}^{N-1} h_k^2\phi_i \left(\frac{u_{i,j}-u_{i-1,j}}{h_k}\right)^2 \leq c\sum_{i=1}^{N}\sum_{j=1}^{N-1} h_k^2\phi_i (M_k\hat{f})_{i,j}^2 \\ &\leq c\,(\max_{ih\geq\eta}\phi_i)\sum_{i=1}^{N}\sum_{j=1}^{N-1} h_k^2 (M_h\hat{f})_{i,j}^2 \leq c\,(\max_{ih\geq\eta}\phi_i)\|f_k\|^2 \leq c\,h_k^{2p}\|f_k\|^2.\end{aligned}$$

Estimate (5.16) follows from an inverse inequality, the Friedrichs inequality and (5.15):

$$\|(u_k)_y\|_{L_2(\Omega_\xi)} \leq c\, h_k^{-1}\|u_k\|_{L_2(\Omega_\xi)} \leq c\,\xi\, h_k^{-1}\|(u_k)_x\|_{L_2(\Omega_\xi)} \leq c\,\xi h_k^{p-1}\|f\|.$$

□

COROLLARY 5.5. *Consider* $f_k \in \mathbb{V}_k$ *such that* $\mathrm{supp}(f_k) \in \Omega_\eta$. *Let* u *and* u_k *be the solutions (2.8) and (2.7), respectively. Assume* $\eta - \xi \geq 8\, h_k\, p\,|\ln h_k|,\ p > 0$. *Then for* $e_k = u - u_k$ *we have*

$$\|(e_k)_x\|_{L_2(\Omega_\xi)} \leq c\, h_k^p \|f_k\|,$$
$$\|(e_k)_y\|_{L_2(\Omega_\xi)} \leq c\, \max\{\sqrt{\frac{\varepsilon_k}{\varepsilon}};\ \frac{\xi}{h_k}\}\, h_k^p\|f_k\|.$$

Proof. Direct superposition of estimates in the corollaries 5.2 and 5.4. □

The result in corollary 5.5 shows that the (H^1-norm of) errors close to the inflow boundary can be made arbitrarily small if the righthand side is zero on a sufficiently large subdomain ($\Omega \setminus \Omega_\eta$) that is adjacent to this inflow boundary. In the proof of the approximation property in section 10 we will need these estimates for the case $\xi = h_k$ and $p = \frac{1}{2}$. Hence we take $\eta = 4h_k|\ln h_k| + h_k$. Note that for the results in the previous corollaries to be applicable we need righthand side functions f_k which are zero in $\Omega \setminus \Omega_\eta$. For technical reasons we take Ω_η such that the right boundary of the domain $\Omega \setminus \Omega_\eta$ coincides with a grid line. We use $|\ln h_k| = k \ln 2$ and thus $4h_k|\ln h_k| + h_k \leq (3k+1)h_k$ and introduce the following auxiliary domains for each grid level

$$\overline{\Omega}_k^{in} := \{\, (x,y) \in \overline{\Omega} \mid x \leq (3k+1)h_k \,\}. \tag{5.17}$$

As a direct consequence of the previous corollary we then obtain

COROLLARY 5.6. *Consider* $f_k \in \mathbb{V}_k$ *such that* f_k *is zero on the subdomain* Ω_k^{in}. *Let* u *and* u_k *be the solutions of (2.8) and (2.7), respectively. Then for* $e_k = u - u_k$ *we have*

$$\|(e_k)_x\|_{L_2(\Omega_{h_k})} \leq c\, h_k^{\frac{1}{2}}\|f_k\|, \tag{5.18}$$
$$\|(e_k)_y\|_{L_2(\Omega_{h_k})} \leq c\, \frac{h_k}{\sqrt{\varepsilon}}\|f_k\|. \tag{5.19}$$

6. Multigrid method and convergence analysis. In this section we describe the multigrid method for solving a problem of the form $A_k x_k = b$ with the stiffness matrix A_k from section 2 and present a convergence analysis.

For the prolongation and restriction in the multigrid algorithm we use the canonical choice:

$$\begin{aligned} p_k &: \mathrm{X}_{k-1} \to \mathrm{X}_k, \quad p_k = P_k^{-1}P_{k-1} \\ r_k &: \mathrm{X}_k \to \mathrm{X}_{k-1}, \quad r_k = P_{k-1}^*(P_k^*)^{-1} = \tfrac{1}{4}p_k^T. \end{aligned} \tag{6.1}$$

Let $W_k : \mathrm{X}_k \to \mathrm{X}_k$ be a nonsingular matrix. We consider a smoother of the form

$$x^{\mathrm{new}} = \mathcal{S}_k(x^{\mathrm{old}}, b) = x^{\mathrm{old}} - \omega W_k^{-1}(A_k x^{\mathrm{old}} - b), \quad \text{for } x^{\mathrm{old}}, b \in \mathrm{X}_k, \tag{6.2}$$

with corresponding iteration matrix denoted by

$$S_k = I - \omega W_k^{-1} A_k. \tag{6.3}$$

The preconditioner W_k we use is of line-Jacobi type:

$$W_k = \alpha \frac{\varepsilon}{h_k^2} I + \hat{D}_x, \quad \alpha > 0 \ . \tag{6.4}$$

Note that W_k is a blockdiagonal matrix with diagonal blocks that are $n_k \times n_k$ bidiagonal matrices. The parameters ω and α are independent of k and ε. A suitable choice for these parameters follows from the analysis below.

REMARK 1. In the literature one is often recommended to apply a so-called *robust smoother* for solving singularly perturbed elliptic problem using multigrid. Such a smoother should have the property that it becomes a direct solver if the singular perturbation parameter tends to zero (cf. [7], chapter 10). In the formulation (6.2) one then must have a splitting such that $A_k - W_k = \mathcal{O}(\varepsilon)$ (the constant in $\mathcal{O}$ may depend on k). Such robust smoothers are well-known for some anisotropic problems. For anisotropic problems in which the anisotropy is *aligned with the gridlines* one can use a line (Jacobi or Gauss-Seidel) method or an ILU factorization as a robust smoother. Theoretical analyses of these methods can be found in [27, 28, 32].
If the convection-diffusion problem (1.2) is discretized using standard finite differences it is easy to see that an appropriate line solver yields a robust smoother. However, in the finite element setting such line methods do *not yield a robust smoother*. This is clear from the stencil in (3.3). For $\varepsilon \to 0$ the diffusion part yields an x-line difference operator which can be represented exactly by an x-line smoother, but in the convection stencil the $[0 \ \ -\frac{1}{6} \ \ \frac{1}{6}]$ and $[-\frac{1}{6} \ \ \frac{1}{6} \ \ 0]$ parts of the difference operator are not captured by such a smoother. It is not clear to us how for the finite element discretization, with a stencil as in (3.3), a robust smoother can be constructed.
In multigrid analyses for reaction-diffusion or anisotropic diffusion problems one usually observes a ε^{-1} dependence in the standard approximation property that is then compensated by an ε factor from the smoothing property (cf. [17, 18, 27, 28, 32]). However, we can not apply a similar technique, due to the fact that for our problem class a robust smoother is not available. Instead, we use another splitting of the iteration matrix of the two-grid method, leading to modified (ε-independent) smoothing and approximation properties. $\square$

We now introduce a nonstandard component in the multigrid algorithm. The convergence analysis will be based on the framework of the smoothing and approximation property. For the analysis of the approximation property we derive finite element error estimates in the L^2-norm (section 10). There we use an Aubin-Nitsche duality argument. For the convection-diffusion equation that we consider the corresponding dual problem can have a boundary layer at the boundary Γ_W, which then gives rise to large discretization errors in the dual problem. This effect results in a strong increase (for $\varepsilon \downarrow 0$) of the theoretical bound in the approximation property and yields unsatisfactory convergence results for the two- and multi-grid method. To overcome this problem we introduce a *"local presolver"* in which a block solver on the local inflow domain Ω_k^{in} defined in (5.17) is applied.

We now describe this local solver operation more precisely. For technical reasons we need a block solver on a slightly larger domain, namely Ω_k^{in} with one additional vertical strip of triangles:

$$\overline{\Omega}_k^{in,+} := \{ \, (x,y) \in \overline{\Omega} \mid x \leq (3k+2)h_k \, \}$$

For the local solver on this subdomain we introduce the local bilinear form

$$a_{in}(u,v) := \int_{\Omega_k^{in,+}} (\varepsilon + \bar{\delta} h_k) u_x v_x + \varepsilon u_y v_y + u_x v \, dxdy, \quad u,v \in \mathbb{V}_k$$

and the corresponding $N_k \times N_k$ stiffness matrix

$$\langle \tilde{A}_k x, y \rangle_k = a_{in}(P_k x, P_k y) \quad \text{for all } x,y \in \mathrm{X}_k. \tag{6.5}$$

Note that this corresponds to the discretization of the given convection-diffusion problem on the local domain $\Omega_k^{in,+}$ with a Neumann boundary condition at the outflow boundary of this local domain.
We introduce the local grid, which is that part of the global grid that lies in the local subdomain $\Omega_k^{in,+}$:

$$G_k^{in,+} := \{x_{i,j} \mid 1 \le i \le h_k^{-1}, 1 \le j < h_k^{-1}\} \cap \Omega_k^{in,+}$$

We assume an ordering of the grid points in this local grid, using a numbering $1, 2, \ldots, m_k$, with $m_k = |G_k^{in,+}|$. Let

$$\bar{P}_k : \mathbb{R}^{m_k} \to \mathbb{R}^{N_k}$$

be the trivial injection operator, i.e., $(\bar{P}_k)_{ij} = 1$ iff grid point j in the local domain has the number i in the global domain. Furthermore let $\bar{R}_k := \bar{P}_k^T$. Using this notation the local solve on the subdomain $G_k^{in,+}$ is given by

$$x^{\text{new}} = \mathcal{R}_k(x^{\text{old}}, b) := x^{\text{old}} - \bar{P}_k(\bar{R}_k \tilde{A}_k \bar{P}_k)^{-1} \bar{R}_k (A_k x^{\text{old}} - b), \quad \text{for } x^{\text{old}}, b \in \mathrm{X}_k. \tag{6.6}$$

For the corresponding iteration matrix we use the notation

$$Q_k := I - \bar{P}_k(\bar{R}_k \tilde{A}_k \bar{P}_k)^{-1} \bar{R}_k A_k.$$

With the components introduced above a standard multigrid algorithm is defined in which we use ν_k post-smoothing iterations with the method in (6.2) and one pre-smoothing iteration (which is actually not a smoother) with the local presolver (6.6). In quasi-algorithmic formulation this multigrid method is as follows:

ALGORITHM 1 (Multigrid method).

```
function MGM_k(x_k, b_k)
{
    if  k = k_0  then
        x_k := A_k^{-1} b_k;                  // solve coarse grid problem
    else
    {
        x_k := R_k(x_k, b_k);                 // local block solver
        d_{k-1} := r_k (b_k - A_k x_k);       // restriction of defect
        e^0_{k-1} := 0;
        for  i = 1  to  τ  do                 // recursion
            e^i_{k-1} := MGM_{k-1}(e^{i-1}_{k-1}, d_{k-1});
        x_k := x_k + p_k e^τ_{k-1};           // add coarse grid correction
        x_k := S_k^{ν_k}(x_k, b_k);           // postsmoothing
    }
    return x_k;
}
```

The two-grid method then has an iteration matrix

$$T_k = S_k^{\nu_k}(I - p_k A_{k-1}^{-1} r_k A_k)Q_k,$$

for the corresponding multigrid W-cycle (i.e., $\tau = 2$) the iteration matrix (cf. [8]) is

$$M_{k_0}^{\mathrm{mgm}} := 0, \quad M_k^{\mathrm{mgm}} = T_k + S_k^{\nu_k} p_k (M_{k-1}^{\mathrm{mgm}})^2 A_{k-1}^{-1} r_k A_k Q_k \ , \tag{6.7}$$

$k = k_0 + 1, k_0 + 2, \ldots$. Note that in the presolver we solve a problem on the domain $\Omega_k^{in,+}$. If $(3k+2)h_k \geq 1$ then $\Omega_k^{in,+} = \bar{\Omega}$ and the two-grid method on level k is a direct solver due to $Q_k = 0$. Hence in the multigrid method we take k_0 sufficiently large (but independent of ε). In the remainder we only consider $k \geq k_0$ such that $\Omega_k^{in,+} \subsetneq \bar{\Omega}$.

After the block solve on the subdomain $\Omega_k^{in,+}$ the *defect* will be equal to zero on the slightly smaller domain Ω_k^{in} (this is the reason why we introduced the enlarged domain $\Omega_k^{in,+}$). This will play an important role in our convergence analysis for the multigrid method. To formulate this "zero defect property" in a more precise way we introduce the diagonal matrix $J_k \in \mathbb{R}^{N_k \times N_k}$ with $(J_k)_{ii} = 0$ iff the grid point with index i in the (global) grid lies in Ω_k^{in}, and $(J_k)_{ii} = 1$ otherwise. Using this projection we can formulate the following

LEMMA 6.1. *The following holds*

$$A_k Q_k = J_k A_k Q_k$$

Proof. One easily verifies that $(A_k)_{ij} = (\tilde{A}_k)_{ij}$ if the grid point with index i lies in Ω_k^{in}. Hence

$$(I - J_k)(\tilde{A}_k - A_k) = 0 \tag{6.8}$$

holds. From the definitions of $\bar{P}_k$ and J_k it follows that

$$(I - J_k)(I - \bar{P}_k \bar{R}_k) = 0 \tag{6.9}$$

Using the results in (6.8) and (6.9) we get

$$\begin{aligned}(I - J_k)A_k Q_k A_k^{-1} &= (I - J_k)(I - A_k \bar{P}_k (\bar{R}_k \tilde{A}_k \bar{P}_k)^{-1} \bar{R}_k) \\ &= (I - J_k)(I - \tilde{A}_k \bar{P}_k (\bar{R}_k \tilde{A}_k \bar{P}_k)^{-1} \bar{R}_k) \\ &= (I - J_k) - (I - J_k)(I - \bar{P}_k \bar{R}_k)\tilde{A}_k \bar{P}_k (\bar{R}_k \tilde{A}_k \bar{P}_k)^{-1} \bar{R}_k \\ &\quad - (I - J_k)\bar{P}_k \bar{R}_k \tilde{A}_k \bar{P}_k (\bar{R}_k \tilde{A}_k \bar{P}_k)^{-1} \bar{R}_k \\ &= (I - J_k) - 0 - (I - J_k)\bar{P}_k \bar{R}_k = (I - J_k)(I - \bar{P}_k \bar{R}_k) = 0 \ .\end{aligned}$$

From this we obtain $(I - J_k)A_k Q_k = 0$. □

We now formulate three main results on which the convergence analysis will be based. The proofs of these results will be given further on.

THEOREM 6.2. *If in (6.4) we take $\alpha > 0$ sufficiently large, then there exists a constant $c_1 = c_1(\alpha)$ independent of k and ε such that*

$$W_k A_k^{-1} \geq c_1 I \quad \textit{for} \quad k = 1, 2, \ldots \tag{6.10}$$

Proof. Given in section 7. It yields (6.10) for any $\alpha \geq 1$ and a corresponding $c_1 \geq 0.1$. □

THEOREM 6.3. *For any $\alpha > 0$ there exists a constant $c_2 = c_2(\alpha)$ independent of k and ε such that*

$$\|W_k(A_k^{-1} - p_k A_{k-1}^{-1} r_k) J_k\| \leq c_2 \quad \textit{for} \quad k = 2, 3, \ldots \tag{6.11}$$

Proof. Given in section 10. □

THEOREM 6.4. *There exists a constant c_3 independent of k and ε such that*

$$\|A_k Q_k A_k^{-1}\| \leq c_3 k \quad \textit{for} \quad k = 1, 2, \ldots \tag{6.12}$$

Proof. Given in section 8. □

Using these results we can prove a two-grid convergence result in the A-norm: $\|B\|_A := \|A_k B A_k^{-1}\|$ for $B \in \mathbb{R}^{N_k \times N_k}$. First we prove a smoothing property in the A-norm.

LEMMA 6.5. *From (6.10) it follows that*

$$\|I - 2c_1 W_k^{-1} A_k\|_A \leq 1$$

holds.

Proof. We use the notation $B := A_k W_k^{-1}$. Note that

$$\|I - 2c_1 W_k^{-1} A_k\|_A^2 = \|I - 2c_1 B\|^2 = \rho((I - 2c_1 B^T)(I - 2c_1 B)) \ .$$

From (6.10) it follows that $B^T B \leq \frac{1}{2c_1}(B^T + B)$ holds and thus

$$0 \leq (I - 2c_1 B^T)(I - 2c_1 B) = I - 2c_1(B^T + B) + 4c_1^2 B^T B \leq I \ .$$

From this the result follows. □

ASSUMPTION 2. *We take a fixed $\alpha > 0$ sufficiently large such that (6.10) holds. In the smoother (6.2) we take*

$$\omega := c_1 \ .$$

We then obtain the following

COROLLARY 1. *The following smoothing property holds:*

$$\|S_k^{\nu_k} W_k^{-1} A_k\|_A \leq \frac{4}{c_1\sqrt{2\pi\nu_k}} \ . \tag{6.13}$$

Proof. Follows from lemma 6.5 and theorem 10.6.8 in [8] (or results in [10, 21]). □

We now obtain a two-grid convergence result:

THEOREM 6.6. *Assume that (6.10), (6.11) and (6.12) hold. For the two-grid method we then have*

$$\|T_k\|_A \leq c_{\text{tg}} \frac{k}{\sqrt{\nu_k}}, \quad \textit{with} \quad c_{\text{tg}} := \frac{4c_2\, c_3}{c_1\sqrt{2\pi}}.$$

Proof. Note that

$$\begin{aligned}
\|T_k\|_A &= \|S_k^{\nu_k}(I - p_k A_{k-1}^{-1} r_k A_k) Q_k\|_A \\
\text{(we use lemma 6.1)} \quad &= \|\big(A_k S_k^{\nu_k} W_k^{-1}\big)\big(W_k(A_k^{-1} - p_k A_{k-1}^{-1} r_k) J_k\big)\big(A_k Q_k A_k^{-1}\big)\| \\
&\le \|S_k^{\nu_k} W_k^{-1} A_k\|_A \|W_k(A_k^{-1} - p_k A_{k-1}^{-1} r_k) J_k\| \|A_k Q_k A_k^{-1}\| \\
&\le \frac{4}{c_1\sqrt{2\pi\nu_k}}\, c_2\, c_3\, k\,.
\end{aligned}$$

Thus the result is proved.
□
Note that from the result in this theorem one obtains a uniform bound $\|T_k\|_A \le c < 1$ for the two-grid contraction number, provided the number of smoothing iterations is sufficiently large: $\nu_k \sim c\,k^2$.
We now derive a convergence result for the multigrid W-cycle. We use that for the canonical restriction operator the inequality

$$\|r_k\| \le c_r$$

holds with a constant c_r independent of k. We need a stability result that will be proved further on:

THEOREM 6.7. *There exists a constant c_4 independent of k and ε such that*

$$\|A_k p_k A_{k-1}^{-1}\| \le c_4 \quad \textit{for} \quad k = 2, 3, \ldots \tag{6.14}$$

Proof. Given in section 9 □

¿From the two-grid result of theorem 6.6 we derive a multigrid convergence result using standard arguments:

THEOREM 6.8. *In addition to the assumptions of theorem 6.6 we assume that (6.14) holds and the number of smoothing steps on every grid level is sufficiently large:*

$$\nu_k \ge c_\nu\, k^4$$

with a suitable constant c_ν (which follows from the proof). Then for the contraction number of the multigrid W-cycle the inequality

$$\|M_k^{\rm mgm}\|_A \le \frac{\xi^*}{k}\ , \quad k > k_0\ , \tag{6.15}$$

holds, with a constant $\xi^* < 1$.

Proof. From the recursion relation (6.7) for $M_k^{\rm mgm}$ immediately follows

$$\|M_k^{\rm mgm}\|_A \le \|T_k\|_A + \|S_k\|_A^{\nu_k} \|A_k p_k A_{k-1}^{-1}\| \|M_{k-1}^{\rm mgm}\|_A^2 \|r_k\| \|A_k Q_k A_k^{-1}\|.$$

The results in theorem 6.4, lemma 6.5, theorem 6.6 and theorem 6.7 yield

$$\|M_k^{\rm mgm}\|_A \le c_{\rm tg} \frac{k}{\sqrt{\nu_k}} + c_r\, c_3\, c_4\, k\, \|M_{k-1}^{\rm mgm}\|_A^2, \quad \text{for} \quad k > k_0\ . \tag{6.16}$$

Define $c_0 := \big(\frac{k_0}{k_0-1}\big)^2$ and consider the recursion

$$\beta_0 = 0, \quad \beta_k = \xi + c_r\, c_3\, c_4\, c_0\, \beta_{k-1}^2, \quad k = 1, 2, \ldots$$

and assume that ξ is sufficiently small such that this recursion has a fixed point $\xi^* < 1$. Then $\beta_k \leq \xi^* \; \forall k$ holds. Now take the number of smoothing iterations on level k sufficiently large such that

$$\nu_k \geq c_{\rm tg}^2 \, \xi^{-2} k^4$$

is satisfied. For $\mu_k := k\|M_k^{\rm mgm}\|_A$ the bound $\mu_k \leq \xi^*$ holds, i.e., we obtain the convergence result (6.15). □

REMARK 2. We briefly discuss the arithmetic work needed in one W-cycle iteration. First note that the arithmetic work for a matrix vector multiplication on level k is of order $\mathcal{O}(N_k) = \mathcal{O}(n_k^2)$. For the block solver iteration $x^{\rm new} = \mathcal{R}_k(x^{\rm old}, b)$ we need a local defect calculation (on the subdomain $\Omega_k^{in,+}$) and a direct solver for the matrix $\bar{R}_k \tilde{A}_k \bar{P}_k$. This is a $m_k \times m_k$ matrix ($m_k = \mathcal{O}(n_k \ln n_k)$) with bandwidth $\mathcal{O}(\ln n_k)$. Hence for solving a system with such a matrix the arithmetic work needed by a direct method is of order $n_k(\ln n_k)^3$. We conclude that (asymptotically) the arithmetic work in the presolver procedure is negligible compared to the matrix vector multiplication $A_k x$. The work needed in one smooting iteration $x^{\rm new} = \mathcal{S}_k(x^{\rm old}, b)$ is of order $\mathcal{O}(N_k)$. The number of smoothings behaves like $\nu_k \sim k^4$. Hence the postsmoothing procedure requires $\mathcal{O}\big(N_k(\ln N_k)^4\big)$ operations. Using a standard recursive argument (cf. [8] chapter 10) it follows that for a multigrid W-cycle iteration the arithmetic complexity is of the order $N_k(\ln N_k)^4$. We conclude that this multigrid method has suboptimal complexity.
The fact that the method does not have the optimal complexity $\mathcal{O}(N_k)$ is caused by the k-dependence of the bound in theorem 6.4. If we would have a local presolver (which should yield a zero defect on the local subdomain) with $\mathcal{O}(N_k)$ complexity per iteration and such that the A-norm of the corresponding iteration matrix is uniformly bounded, this would result in a multigrid method with optimal complexity.

7. Proof of theorem 6.2. We recall the representation of the stiffness matrix in (3.4)

$$A_k = (\varepsilon + (\bar{\delta} - \frac{1}{3})h_k)\hat{A}_x + \varepsilon \hat{A}_y + \frac{1}{6}\hat{B}\hat{D}_x.$$

We will need the following lemma

LEMMA 7.1. *The following holds*

$$\hat{B}\hat{D}_x \geq 0$$

Proof. The matrix $\frac{1}{6}\hat{B}\hat{D}_x - \frac{1}{3}h_k\hat{A}_x$ is the stiffness matrix corresponding to the bilinear form

$$(u,v) \to \int_\Omega u_x v \; dxdy$$

For any $z \in X_k$ we get

$$\frac{1}{6}\langle \hat{B}\hat{D}_x z, z\rangle_k - \frac{1}{3}\langle h_k \hat{A}_x z, z\rangle_k = \int_\Omega (P_k z)_x (P_k z) \; dxdy = \frac{1}{2}\int_{\Gamma_E} (P_k z)^2 \; dxdy \geq 0$$

Since the matrix $\hat{A}_x$ is symmetric positive definite the result now follows.
□

We now consider the preconditioner $W_k = \alpha\frac{\varepsilon}{h_k^2}I + \hat{D}_x$, as in (6.4).

THEOREM 7.2. *For $\alpha \geq 1$ the following holds:*

$$W_k A_k^{-1} \geq \frac{1}{10} I$$

Proof. First note that

$$h_k D_x D_x^T = D_x + D_x^T - \frac{1}{h_k}(1,0,\ldots,0)^T(1,0,\ldots,0) \leq D_x + D_x^T$$

and thus $h_k D_x^T D_x D_x^T D_x \leq D_x^T(D_x + D_x^T)D_x$ holds. Using $A_x = D_x^T D_x$ this results in $h_k A_x^2 \leq 2D_x^T A_x$ and thus:

$$\frac{1}{2}h_k \hat{A}_x^2 \leq \hat{D}_x^T \hat{A}_x \tag{7.1}$$

Note that the following inequality holds for any $a, b, c \in \mathbb{R}$ and $\sigma_1, \sigma_2, \sigma_3 > 0$:

$$(a+b+c)^2 \leq (1+\sigma_2+\sigma_3^{-1})a^2 + (1+\sigma_3+\sigma_1^{-1})b^2 + (1+\sigma_1+\sigma_2^{-1})c^2.$$

We apply this inequality with $\sigma_3 = 2, \sigma_1 = \sigma_2 = 1$. Also using $\|\hat{A}_y\| \leq 4h_k^{-2}$ and $\|\hat{B}\| \leq 6$ we get

$$\begin{aligned} \|A_k z\|^2 &\leq \frac{5}{2}\varepsilon^2\|\hat{A}_y z\|^2 + 4\bar{\varepsilon}_k^2\|\hat{A}_x z\|^2 + 3\|\frac{1}{6}\hat{B}\hat{D}_x z\|^2 \\ &\leq 10\big(\frac{\varepsilon}{h_k}\big)^2\langle \hat{A}_y z, z\rangle_k + 4\bar{\varepsilon}_k^2\|\hat{A}_x z\|^2 + 3\|\hat{D}_x z\|^2 \end{aligned} \tag{7.2}$$

We recall that $\bar{\varepsilon}_k = \varepsilon_k - \bar{\delta}h_k \leq \frac{7}{6}h_k$. Now apply the result (7.1) and the estimates in lemma 3.1, lemma 7.1 to obtain

$$\begin{aligned} \langle W_k z, A_k z\rangle_k &= \langle \alpha\frac{\varepsilon}{h_k^2}z + \hat{D}_x z\,,\, \varepsilon\hat{A}_y z + \bar{\varepsilon}_k\hat{A}_x z + \frac{1}{6}\hat{B}\hat{D}_x z\rangle_k \\ &\geq \alpha\big(\frac{\varepsilon}{h_k}\big)^2\langle \hat{A}_y z, z\rangle_k + \bar{\varepsilon}_k\langle \hat{D}_x z, \hat{A}_x z\rangle_k + \langle \hat{D}_x z, \frac{1}{6}\hat{B}\hat{D}_x z\rangle_k \\ &\geq \alpha\big(\frac{\varepsilon}{h_k}\big)^2\langle \hat{A}_y z, z\rangle_k + \frac{3}{7}\bar{\varepsilon}_k^2\|\hat{A}_x z\|^2 + \frac{1}{3}\|\hat{D}_x z\|^2 \\ &\geq \frac{1}{10}\Big(10\alpha\big(\frac{\varepsilon}{h_k}\big)^2\langle \hat{A}_y z, z\rangle_k + \frac{30}{7}\bar{\varepsilon}_k^2\|\hat{A}_x z\|^2 + \frac{10}{3}\|\hat{D}_x z\|^2\Big) \\ &\geq \frac{1}{10}\Big(10\big(\frac{\varepsilon}{h_k}\big)^2\langle \hat{A}_y z, z\rangle_k + 4\bar{\varepsilon}_k^2\|\hat{A}_x z\|^2 + 3\|\hat{D}_x z\|^2\Big) \end{aligned}$$

Combination of this with the inequality in (7.2) proves the theorem.
□
The previous theorem shows that theorem 6.2 holds.

8. Proof of theorem 6.4. We consider the presolver with iteration matrix

$$Q_k := I - \bar{P}_k(\bar{R}_k\tilde{A}_k\bar{P}_k)^{-1}\bar{R}_k A_k$$

and introduce

$$\tilde{D}_x := \bar{R}_k\hat{D}_x\bar{P}_k.$$

Note that $\tilde{D}_x$ represents the difference operator $\frac{1}{h}[-1 \quad 1 \quad 0]_x$ in all grid points x in the local domain $\Omega^{in,+}$. The matrix $\bar{R}_k\tilde{A}_k\bar{P}_k$ results from the streamline diffusion finite element discretization on this local domain with homogeneous Neumann outflow boundary conditions. Now we can use the same arguments as in section 4 (with A_k replaced by $\bar{R}_k\tilde{A}_k\bar{P}_k$ and $\hat{D}_x$ replaced by $\tilde{D}_x$ to prove (cf. lemma 3.1)

$$\|\tilde{D}_x(\bar{R}_k\tilde{A}_k\bar{P}_k)^{-1}\| \leq 3$$

The matrix $\tilde{D}_x$ has (after some permutation) block diagonal form with diagonal blocks $\frac{1}{h_k}\text{tridiag}(-1,1,0)$ of dimension $(3k+1)\times(3k+1)$. Hence,

$$\|\tilde{D}_x^{-1}\| \leq (3k+1)h_k$$

holds. Using this and $\|A_k\| \leq ch_k^{-1}$ we get

$$\|A_kQ_kA_k^{-1}\| \leq 1 + \|A_k\|\|\bar{P}_k\|\|\tilde{D}_x^{-1}\|\|\tilde{D}_x(\bar{R}_k\tilde{A}_k\bar{P}_k)^{-1}\|\|\bar{R}_k\| \leq c\,k\ .$$

This proves the result in theorem 6.4

9. Proof of theorem 6.7. Let $g_{k-1} \in X_{k-1}$ be given and define $\tilde{g}_{k-1} := (P_{k-1}^*)^{-1}g_{k-1} \in \mathbb{V}_{k-1}$. Let $u_{k-1} \in \mathbb{V}_{k-1}$ be such that

$$a_{k-1}(u_{k-1}, v_{k-1}) = (\tilde{g}_{k-1}, v_{k-1}) \quad \text{for all} \quad v_{k-1} \in \mathbb{V}_{k-1}.$$

Then $A_{k-1}^{-1}g_{k-1} = P_{k-1}^{-1}u_{k-1}$ holds. The corresponding continuous solution $u \in \mathbf{V}$ is given by

$$a_{k-1}(u, v) = (\tilde{g}_{k-1}, v) \quad \text{for all} \quad v \in \mathbf{V}.$$

Now note that

$$\begin{aligned}\|A_kp_kA_{k-1}^{-1}g_{k-1}\| &= \max_{y\in X_k}\frac{\langle A_kp_kP_{k-1}^{-1}u_{k-1}, y\rangle_k}{\|y\|} \leq c\max_{v_k\in\mathbb{V}_k}\frac{a_k(u_{k-1},v_k)}{\|v_k\|} \\ &\leq c\max_{v_k\in\mathbb{V}_k}\frac{a_{k-1}(u_{k-1},v_k)}{\|v_k\|} + c\max_{v_k\in\mathbb{V}_k}\frac{a_k(u_{k-1},v_k)-a_{k-1}(u_{k-1},v_k)}{\|v_k\|}\end{aligned} \tag{9.1}$$

Define $e_{k-1} := u - u_{k-1}$. For the first term in (9.1) we get, using the results of lemma 4.3 :

$$\begin{aligned}a_{k-1}(u_{k-1}, v_k) &\leq |a_{k-1}(e_{k-1}, v_k)| + |a_{k-1}(u, v_k)| \\ &\leq ch_k\|(e_{k-1})_x\|\|(v_k)_x\| + \varepsilon\|(e_{k-1})_y\|\|(v_k)_y\| + \|(e_{k-1})_x\|\|v_k\| + |(\tilde{g}_{k-1}, v_k)| \\ &\leq c\big(\|(e_{k-1})_x\| + \frac{\varepsilon}{h_k}\|(e_{k-1})_y\|\big)\|v_k\| + \|\tilde{g}_{k-1}\|\|v_k\| \\ &\leq c\|\tilde{g}_{k-1}\|\|v_k\| \leq c\|g_{k-1}\|\|v_k\|\end{aligned} \tag{9.2}$$

For the second term in (9.1) we have, using lemma 4.2 :

$$\begin{aligned}|a_k(u_{k-1}, v_k) - a_{k-1}(u_{k-1}, v_k)| &= \bar{\delta}h_k|((u_{k-1})_x, (v_k)_x)| \\ &\leq c\|(u_{k-1})_x\|\|v_k\| \\ &\leq c\|\tilde{g}_{k-1}\|\|v_k\| \leq c\|g_{k-1}\|\|v_k\|\end{aligned} \tag{9.3}$$

Combination of the results in (9.1), (9.2) and (9.3) yields

$$\|A_kp_kA_{k-1}^{-1}g_{k-1}\| \leq c\|g_{k-1}\|$$

and thus the result in theorem 6.7 holds. □

10. Proof of theorem 6.3. We briefly comment on the idea of the proof. As usual to prove an estimate for the error in the L^2-norm we use a duality argument. However, the formal dual problem has poor regularity properties, since in this dual problem Γ_E is the "inflow" boundary and Γ_W is the "outflow" boundary. Thus Dirichlet outflow boundary conditions would appear and we obtain poor estimates due to the poor regularity. To avoid this, we consider a dual problem with Neumann outflow and Dirichlet inflow conditions. To be able to deal with the inconsistency caused by these "wrong" boundary conditions we *assume* the righthand side is zero near the boundary Γ_W. In order to satisfy this assumption we use the local presolver in the multigrid algorithm.
A further problem we have to deal with is the fact that due to the level dependent stabilization term we have to treat k-dependent bilinear forms.

We introduce the space

$$\mathbb{V}_k^0 := \{ v_k \in \mathbb{V}_k \mid v_k(x) = 0 \quad \text{for all} \;\; x \in \Omega_k^{in} \}$$

Let $b_k \in X_k$ be given. In view of theorem 6.3 we must prove an estimate $\|W_k(A_k^{-1} - p_k A_{k-1}^{-1} r_k) J_k b_k\| \le c\|b_k\|$ with a constant c that is independent of k, ε and b_k. Note that

$$(P_k^*)^{-1} J_k b_k =: f_k \in \mathbb{V}_k^0$$

holds. For this $f_k \in \mathbb{V}_k^0$ we define corresponding discrete solutions and continuous solutions as follows:

$$\begin{array}{rll} u_k \in V_k: & a_k(u_k, v_k) = (f_k, v_k) & \text{for all} \;\; v_k \in V_k \\ u \in \mathbf{V}: & a_k(u, v) = (f_k, v) & \text{for all} \;\; v \in \mathbf{V} \\ u_{k-1} \in V_{k-1}: & a_{k-1}(u_{k-1}, v_{k-1}) = (f_k, v_{k-1}) & \text{for all} \;\; v_{k-1} \in V_{k-1} \\ \tilde{u} \in \mathbf{V}: & a_{k-1}(\tilde{u}, v) = (f_k, v) & \text{for all} \;\; v \in \mathbf{V} \end{array} \tag{10.1}$$

In the proof of lemma 4.2 we showed that $\|v_x\| = \|\hat{D}_x P_k^{-1} v\|$ holds for all $v \in \mathbb{V}_k$. We use that $W_k = \alpha \frac{\varepsilon}{h_k^2} I + \hat{D}_x$ and obtain

$$\begin{aligned} &\|W_k(A_k^{-1} - p_k A_{k-1}^{-1} r_k) J_k b_k\| \\ \le\; & c\frac{\varepsilon}{h_k^2} \|(A_k^{-1} - p_k A_{k-1}^{-1} r_k) J_k b_k\| + \|\hat{D}_x A_k^{-1} J_k b_k\| + \|\hat{D}_x p_k A_{k-1}^{-1} r_k J_k b_k\| \\ &\le c\big(\frac{\varepsilon}{h_k^2} \|u_k - u_{k-1}\| + \|(u_k)_x\| + \|(u_{k-1})_x\| \big) \\ \le\; & c\Big(\frac{\varepsilon}{h_k^2} \big(\|u - u_k\| + \|\tilde{u} - u_{k-1}\| + \|u - \tilde{u}\| \big) + \|(u_k)_x\| + \|(u_{k-1})_x\| \Big) \end{aligned} \tag{10.2}$$

From lemma 4.2 we get

$$\|(u_k)_x\| + \|(u_{k-1})_x\| \le c\|f_k\| \tag{10.3}$$

From the result in theorem 10.1 below it follows that

$$\|u_k - u\| + \|u_{k-1} - \tilde{u}\| \le c\, \frac{h_k^2}{\varepsilon} \, \|f_k\| \tag{10.4}$$

Finally, from theorem 10.4 we have

$$\|u - \tilde{u}\| \le c\, h_k \|f_k\| \tag{10.5}$$

If we insert the results (10.3),(10.4) and (10.5) in (10.2) we get

$$\|W_k(A_k^{-1} - p_k A_{k-1}^{-1} r_k) J_k b_k\| \leq c\|f_k\| \leq c\|b_k\|$$

and thus the result of theorem 6.3 is proved.
It remains to prove the results in the theorems 10.1 and 10.4.

THEOREM 10.1. *For $f_k \in \mathbb{V}_k^0$ let u and u_k be as defined in (10.1). Then*

$$\|u - u_k\| \leq c\,\frac{h_k^2}{\varepsilon}\,\|f_k\| \tag{10.6}$$

holds.

Proof. Define $e_k := u - u_k$. Let $w \in H^2(\Omega)$ be such that

$$-\varepsilon w_{yy} - \varepsilon_k w_{xx} - w_x = e_k \tag{10.7}$$

with

$$w_x = 0 \text{ on } \Gamma_W\ , \quad w = 0 \text{ on } \Gamma \setminus \Gamma_W. \tag{10.8}$$

Note that for this problem Γ_E is the "inflow" boundary and Γ_W is the "outflow" boundary. We multiply (10.7) with e_k and integrate by parts to get

$$\begin{aligned}\|e_k\|^2 &= \varepsilon((e_k)_y, w_y) + \varepsilon_k((e_k)_x, w_x) - \varepsilon_k \int_{\Gamma_E} w_x e_k\, dy + ((e_k)_x, w)\\ &= a_k(e_k, w) - \varepsilon_k \int_{\Gamma_E} w_x e_k\, dy\end{aligned}$$

We use (4.6) with w and e_k instead of u and f, respectively, and (4.12) to estimate

$$\left|\varepsilon_k \int_{\Gamma_E} w_x e_k\, dy\right| \leq \varepsilon_k^{\frac{1}{2}} \left(\varepsilon_k \int_{\Gamma_E} w_x^2\, dy\right)^{\frac{1}{2}} \left(\int_{\Gamma_E} e_k^2\, dy\right)^{\frac{1}{2}} \leq c\, h_k^{\frac{1}{2}} \|e_k\| \frac{h_k}{\sqrt{\varepsilon}} \|f_k\|. \tag{10.9}$$

From this estimate and the Galerkin orthogonality for the error it follows that for any $v_k \in \mathbb{V}_h$

$$\|e_k\|^2 \leq \varepsilon\,((e_k)_y, (w - v_k)_y) + \varepsilon_k\,((e_k)_x, (w - v_k)_x) + ((e_k)_x, w - v_k) + c\,\|e_k\| \frac{h_k^{\frac{3}{2}}}{\sqrt{\varepsilon}} \|f_k\|. \tag{10.10}$$

Let $\Omega_h := \Omega_{h_k}$ be as defined in (5.4), i.e., Ω_h is the set of triangles with at least one vertex on Γ_W. In the remainder of the domain, $\omega = \Omega \backslash \Omega_h$, we take v_k as a nodal interpolant to w and we put $v_k = 0$ on Γ_W to ensure $v_k \in \mathbb{V}_h$. Note that v_k is a proper interpolant of w everywhere in Ω except in Ω_h. Therefore we will estimate scalar products in (10.10) over ω and Ω_h, separately. We continue (10.10) with:

$$\begin{aligned}\|e_k\|^2 &\leq c\,\varepsilon\, h_k\, \|(e_k)_y\|_\omega\, \|w\|_{H^2(\omega)} + c\,\varepsilon_k\, h_k\, \|(e_k)_x\|_\omega\, \|w\|_{H^2(\omega)}\\ &\quad + c\, h_k^2\, \|(e_k)_x\|_\omega\, \|w\|_{H^2(\omega)} + c\, \|e_k\| \frac{h_k^{\frac{3}{2}}}{\sqrt{\varepsilon}} \|f_k\| + \mathrm{I}_{\Omega_h}\\ &\leq c\, h_k^2 \|f_k\| \frac{1}{\varepsilon} \|e_k\| + \mathrm{I}_{\Omega_h}.\end{aligned} \tag{10.11}$$

The term I_{Ω_h} collects integrals over Ω_h:

$$\mathrm{I}_{\Omega_h} = \varepsilon\left((e_k)_y, (w - v_k)_y\right)_{\Omega_h} + \varepsilon_k\left((e_k)_x, (w - v_k)_x\right)_{\Omega_h} + \left((e_k)_x, w - v_k\right)_{\Omega_h}$$

To estimate I_{Ω_h} we use corollary 5.6 and the following auxiliary estimate, with $\omega_h = \{(x, y) \in \Omega : x \in (h_k, 2h_k)\}$:

$$\begin{aligned}
\|v_k\|_{\Omega_h} &\leq c\|v_k\|_{\omega_h} \leq c(\|w\|_{\omega_h} + \|v_k - w\|_\omega) \\
&= c\left(\left(\int_0^1 \int_{h_k}^{2h_k} \left[w(0, y) + \int_0^x w_\eta(\eta, y)\, d\eta \right]^2 dx\, dy \right)^{\frac{1}{2}} + \|v_k - w\|_\omega \right) \\
&\leq c(h_k^{\frac{1}{2}}(\int_{\Gamma_W} w^2\, dy)^{\frac{1}{2}} + h_k\|w_x\| + h_k^2\|w\|_{H^2(\omega)}) \leq c(h_k^{\frac{1}{2}} + \frac{h_k^2}{\varepsilon})\|e_k\|.
\end{aligned}$$

We proceed estimating terms from I_{Ω_h}:

$$\begin{aligned}
\varepsilon\left((e_k)_y, (w - v_k)_y\right)_{\Omega_h} &\leq \varepsilon\, \|(e_k)_y\|_{\Omega_h} \left(\|w_y\| + \|(v_k)_y\|_{\Omega_h} \right) \\
&\leq c\,\varepsilon^{\frac{1}{2}} h_k^{\frac{1}{2}} \varepsilon_k^{\frac{1}{2}} \|f_k\|\ (\varepsilon^{-\frac{1}{2}}\|e_k\| + h_k^{-1}\|v_k\|_{\Omega_h}) \\
&\leq c\,\varepsilon^{\frac{1}{2}} h_k \|f_k\| (\varepsilon^{-\frac{1}{2}} + h_k^{-\frac{1}{2}} + \frac{h_k}{\varepsilon})\|e_k\| \leq c\,(h_k + \frac{h_k^2}{\sqrt{\varepsilon}})\|f_k\|\|e_k\|, \\
\varepsilon_k\left((e_k)_x, (w - v_k)_x\right)_{\Omega_h} &\leq \varepsilon_k\, \|(e_k)_x\|_{\Omega_h} (\|w_x\| + \|(v_k)_x\|_{\Omega_h}) \\
&\leq c\, h_k^{\frac{1}{2}} \varepsilon_k \|f_k\|\ (\|e_k\| + h_k^{-1}\|v_k\|_{\Omega_h}) \leq c\,(h_k + \frac{h_k^{\frac{5}{2}}}{\varepsilon})\|f_k\|\|e_k\|, \\
\left((e_k)_x, w - v_k\right)_{\Omega_h} &\leq \|(e_k)_x\|_{\Omega_h} (\|w\|_{\Omega_h} + \|v_k\|_{\Omega_h}) \\
&\leq c\, h_k^{\frac{1}{2}} \|f_k\|\ (h_k^{\frac{1}{2}} \left((\int_{\Gamma_W} w^2\, dy)^{\frac{1}{2}} + h_k\|w_x\|_{\Omega_h} + \|v_k\|_{\Omega_h} \right) \\
&\leq c\,(h_k + \frac{h_k^{\frac{5}{2}}}{\varepsilon})\|f_k\|\|e_k\|.
\end{aligned}$$

Inserting this estimates into (10.11) and using $\varepsilon \leq \frac{1}{2}h_k$ we obtain

$$\|e_k\|^2 \leq c\,\frac{h_k^2}{\varepsilon}\|f_k\|\|e_k\| + c\,(h_k + \frac{h_k^2}{\sqrt{\varepsilon}} + \frac{h_k^{\frac{5}{2}}}{\varepsilon})\|f_k\|\|e_k\| \leq c\,\frac{h_k^2}{\varepsilon}\|f_k\|\|e_k\|.$$

and thus the theorem is proved.
□

For the proof of theorem 10.4 we first formulate two lemmas.

LEMMA 10.2. *Consider a function $g \in \mathrm{H}^1(\Omega)$. The solution of*

$$-\varepsilon u_{yy} - \varepsilon_k u_{xx} + u_x = g_x \tag{10.12}$$

with boundary conditions as in (1.2) satisfies

$$\int_{\Gamma_E} u^2\, dy \leq c\left(h_k^{-1}\, \|g\|^2 + \int_{\Gamma_E} g^2\, dy + h_k\, \|g_x\|^2 \right). \tag{10.13}$$

Proof. We multiply (10.12) with u and integrate by parts to get

$$\varepsilon\, \|u_y\|^2 + \varepsilon_k\, \|u_x\|^2 + \frac{1}{2}\int_{\Gamma_E} u^2\, dy = -(g, u_x) + \int_{\Gamma_E} g\, u\, dy. \tag{10.14}$$

For the righthand side in (10.14) we have

$$|(g, u_x)| \leq \|g\| \, \|u_x\| \leq c \, \|g\| \, \|g_x\| \leq c \, \left(h_k^{-1} \, \|g\|^2 + h_k \, \|g_x\|^2 \right)$$

and

$$\int_{\Gamma_E} g \, u \, dy \leq \int_{\Gamma_E} g^2 \, dy + \frac{1}{4} \int_{\Gamma_E} u^2 \, dy.$$

Combining these estimates and (10.14) we prove the lemma.
□

LEMMA 10.3. *For $g \in \mathbb{V}_k^0$ let u be the corresponding solution of (10.12). Then the following holds:*

$$\|u\| \leq c \, \|g\|. \tag{10.15}$$

(Note that the standard a priori estimates would give only $\|u\| \leq c \, \|g_x\|$.)

Proof. Consider the auxiliary function $v(x, y) := \int_0^x u(\xi, y) \, d\xi$. It satisfies

$$-\varepsilon v_{yy} - \varepsilon_k v_{xx} + v_x = g + \varepsilon_k \, u_{in}, \tag{10.16}$$

with $u_{in}(x, y) = \frac{\partial u}{\partial x}(0, y)$. The corresponding boundary conditions are

$$\frac{\partial v}{\partial x} = u(1, y) \text{ on } \Gamma_E \ , \qquad v = 0 \text{ on } \partial\Omega \setminus \Gamma_E. \tag{10.17}$$

Then the estimate (10.15) is equivalent to

$$\|v_x\| \leq c \, \|g\|. \tag{10.18}$$

The estimate (10.18) is proved by the following arguments. We multiply (10.16) with v_x and integrate by parts to obtain

$$\begin{aligned} \|v_x\|^2 &+ \frac{\varepsilon}{2} \int_{\Gamma_E} (v_y)^2 \, dy + \frac{\varepsilon_k}{2} \int_{\Gamma_W} (v_x)^2 \; dy \\ &= (g, v_x) + \varepsilon_k (u_{in}, v_x) + \frac{\varepsilon_k}{2} \int_{\Gamma_E} (v_x)^2 \; dy. \end{aligned} \tag{10.19}$$

Note that for a function g from the finite element space $\mathbb{V}_h$ estimate (10.13) yields due to finite element inverse inequalities

$$\int_{\Gamma_E} (v_x)^2 \; dy = \int_{\Gamma_E} u^2 \, dy \leq c \, \left(h_k^{-1} \, \|g\|^2 + \int_{\Gamma_E} g^2 \, dy + h_k \; \|g_x\|^2 \right) \leq c_1 \, h_k^{-1} \, \|g\|^2. \tag{10.20}$$

Now we use the fact that $g = 0$ in Ω_k^{in}. Due to the choice of Ω_k^{in} (cf. (5.17)) we can apply corollary 5.2 with $\xi = h_k$, $\eta = 2\varepsilon_k |\ln h_k| + h_k$ and $p = 1$. Using (5.6) we get

$$\begin{aligned} \varepsilon_k (u_{in}, v_x) &\leq \varepsilon_k \|u_{in}\|^2 + \frac{1}{4} \, \|v_x\|^2 = \varepsilon_k \int_{\Gamma_W} (u_x)^2 \; dy + \frac{1}{4} \, \|v_x\|^2 \\ &\leq c \, h_k^2 \; \|g_x\|^2 + \frac{1}{4} \, \|v_x\|^2 \leq c \, \|g\|^2 + \frac{1}{4} \, \|v_x\|^2 \, . \end{aligned} \tag{10.21}$$

Now (10.15) follows from (10.19) by applying the Cauchy inequality, estimate (10.20) and (10.21).
□

Using these lemmas we can prove the final result we need.

THEOREM 10.4. *Let u and $\tilde{u}$ be the continuous solutions defined in (10.1). Then the following holds*

$$\|u - \tilde{u}\| \leq c\, h_k \|f_k\|. \tag{10.22}$$

Proof. The difference $e := u - \tilde{u}$ solves the equation

$$-\varepsilon e_{yy} - \varepsilon_k e_{xx} + e_x = g_x \ , \tag{10.23}$$

with $g = -\bar{\delta} h_k \tilde{u}_x$ and boundary conditions as in (1.2). Now the result of lemma 10.3 apply. We obtain

$$\|e\| \leq c\,\|g\| = c\, h_k\, \|\tilde{u}_x\| \leq c_1\, h_k \|f_k\|.$$

□

REFERENCES

[1] Bank, R.E., Benbourenane, M. (1992): The hierarchical basis multigrid method for convection-diffusion equations. Numer. Math. **61**, 7–37

[2] Bey, J., Wittum, G. (1997): Downwind numbering: robust multigrid for convection-diffusion problems. Appl. Numer. Math. **23**, 177–192

[3] Bramble, J.H. (1993): Multigrid methods. Longman, Harlow

[4] Bramble, J.H., Pasciak, J.E., Xu, J. (1988): The analysis of multigrid algorithms for nonsymmetric and indefinite problems. Math. Comp. **51**, 389–414

[5] Grisvard, P. (1985): Elliptic problems in nonsmooth domains. Pitman, Boston

[6] Johnson, C., Schatz, A.H., Wahlbin, L.B. (1987): Crosswind smear and pointwise errors in streanline diffusion finite element methods. Math. Comp. **49**, 25–38

[7] Hackbusch, W. (1985): Multi-Grid Methods and Applications. Springer, Berlin

[8] Hackbusch, W. (1994): Iterative Solution of Large Sparse Systems of Equations. Springer, New-York

[9] Hackbusch, W. (1984): Multigrid convergence for a singular perturbation problem. Linear Algebra Appl. **58**, 125–145.

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

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

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

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

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

[15] 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

[16] Niijima, K. (1990): Poinwise error estimates for a streamline diffusion finite element scheme. Numer. Math. **56**, 707–719

[17] Olshanskii, M.A. Reusken, A. (2000): On the convergence of a multigrid method for linear reaction-diffusion problem, Computing, **65**, 193–202

[18] Olshanskii, M.A. Reusken, A. (2002): Navier-Stokes equations in rotation form: A robust multigrid solver for the velocity problem, SIAM J. Sci. Comput., **23**, 1683–1706

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

[20] 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

[21] Reusken, A. (1992): On maximum norm convergence of multigrid methods for two-point boundary value problems. SIAM J. Numer. Anal. **29**, 1569–1578
[22] Reusken, A. (1994): Multigrid with matrix-dependent transfer operators for a singular perturbation problem. Computing **50**, 199–211
[23] Reusken, A. (1995): Fourier analysis of a robust multigrid method for convection-diffusion equations. Numer. Math. **71**, 365–397
[24] Reusken, A. (2002): Convergence analysis of a multigrid method for convection-diffusion equations. Numer. Math. **91**, 323–349
[25] Roos, H.-G., M. Stynes, and L. Tobiska (1996): Numerical Methods for Singularly Perturbed Differential Equations. Springer, Berlin
[26] Sani, R.L., Gresho P.M.(1994): Resume and remarks on open boundary conditions minisimposium. Int.J.Numer.Meth.Fluids **18**, 983–1008
[27] Stevenson, R.P. (1993): New estimates of the contraction number of V-cycle multi-grid with applications to anisotropic equations. In: Incomplete decompositions (W. Hackbusch and G. Wittum, eds.), Proceedings of the eight GAMM Seminar. Notes on Numerical Fluid Mechanics **41**, 159–167
[28] Stevenson, R.P. (1993): Robustness of multi-grid applied to anisotropic equations on convex domains and on domains with re-entrant corners. Numer. Math. **66**, 373–398
[29] Wahlbin, L.B. (1991): Local behavior in finite element methods. In: Handbook of Numerical Analysis, Vol.II Finite element methods (eds. P.G. Ciarlet, J.L. Lions), North-Holland, Amsterdam, 353–522
[30] Wang, J. (1993): Convergence analysis of multigrid algorithms for nonselfadjoint and indefinite elliptic problems. SIAM J. Numer. Anal. **30**, 275–285
[31] Wesseling, P. (1992): An Introduction to Multigrid Methods. Wiley, Chichester
[32] Wittum, G. (1989): On the robustness of ILU smoothing, SIAM J. Sci. Stat. Comput. **10**, 699–717
[33] Xu, J. (1992): Iterative methods by space decomposition and subspace correction. SIAM Review **34**, 581–613
[34] Yserentant, H. (1993): Old and new convergence proofs for multigrid methods. Acta Numerica 1993, 285–326
[35] Zeeuw, P.M. de (1990): Matrix-dependent prolongations and restrictions in a blackbox multigrid solver. J. Comput. Appl. Math. **33**, 1–27
[36] Zhou, G. (1997): How accurate is the streamline diffusion finite element method. Math. Comput. **66**, 31–44

Maxim A. Olshanskii
Department of Mechanics and Mathematics
Moscow M.V.Lomonosov University
Moscow 119899, Russia
e-mail: Maxim.Olshanskii@mtu-net.ru

Arnold Reusken
Institut für Geometrie und Praktische Mathematik
RWTH Aachen
Templergraben 55, D-52056 Aachen, Germany
e-mail: reusken@igpm.rwth-aachen.de